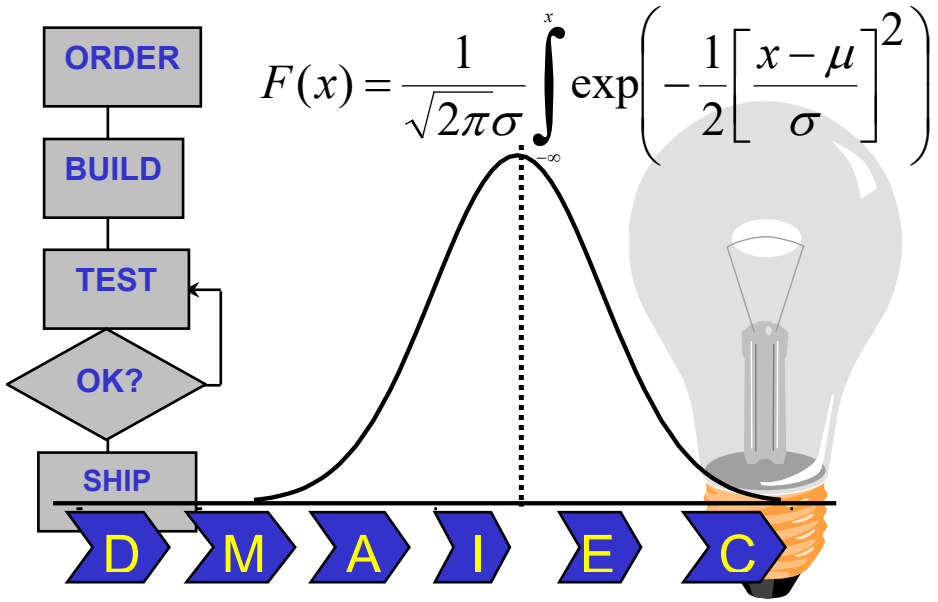
The Black Belt Six Sigma Toolkit



Section	Topic
1.0	**Introduction and Purpose**
2.0	**Improvement Systems**
2.1	Quality Management Systems
2.2	Improving Existing Products and Services
2.3	Managing Processes
2.4	Designing New Products and Services
2.5	Business Strategy Planning Process
2.6	Six Sigma "Belts"
3.0	**Team Facilitation and Management**
3.1	Working With Teams
3.2	Idea Generation & Decision Making
3.3	Exercises
4.0	**Obtaining the Voice of the Customer**
4.1	Core Customer Research Methods
4.2	Exercises
5.0	**Process Management & Analysis**
5.1	Process Thinking
5.2	Pictures of the Process
5.3	Process Management Methods
5.4	Process Analysis Methods
5.5	Lean Manufacturing
5.6	Exercises
6.0	**Measuring Performance & Variability**
6.1	Developing Performance Indicators
6.2	Data Collection
6.3	Core Data Displays
6.4	Introduction to Control Charts
6.5	Measurement Control Charts
6.6	Attribute Control Charts
6.7	Measurement System Analysis
6.8	Process Capability Analysis
6.9	Additional Control Chart Topics
6.10	Exercises
7.0	**Stratification & Prioritization**
7.1	Pie, Bar & Radar Charts
7.2	Pareto Analysis
7.3	Exercises
8.0	**Cause & Effect**
8.1	Cause and Effect Analysis
8.2	Exercises
9.0	**Detecting Differences**
9.1	Foundations of Probability and Statistics
9.2	Hypothesis Testing
9.3	Sampling Theory
9.4	Exercises
10.0	**Relationships Between Variables**
10.1	Scatter Diagrams and Correlation Analysis
10.2	Regression Analysis
10.3	Analysis of Variance (ANOVA)
10.4	Exercises
11.0	**Experimentation**
11.1	Designing and Running Experiments
11.2	Exercises

Section	Topic
12.0	**Changing the Process**
12.1	Selecting & Implementing Countermeasures
12.2	Financial Analysis of Changes
12.3	Exercises
13.0	**Changing**
13.1	Change Management
13.2	Exercises
14.0	**Design Management**
14.1	Defining Product/Service Requirements
14.2	Conceptual Design
14.3	Benchmarking
14.4	Taguchi Design Approach
14.5	Multi-Generational Product Planning
14.6	Exercises
15.0	**Reliability Management**
15.1	Reliability Concepts and Management
15.2	Failure/Error Modes & Effects Analysis
15.3	Fault Tree Analysis
15.4	Quantifying Reliability
15.5	Root Cause Analysis
15.6	Exercises
16.0	**Planning & Review Tools**
16.1	Seven Planning Tools
16.2	Operating Reviews
16.3	Exercises

Section	Topic
Appendices	**A. Probability Distributions**
	B. Sigma Conversion Table
	C. Forms and Templates
	D. Answers to Selected Exercises
	Glossary of Statistical Terms
	Bibliography
	Index

1.0 Introduction & Purpose

Introduction & Purpose

If one term is synonymous with Six Sigma, it is the *Black Belt.* The Black Belt is the core engine of improvement – a *lean, mean quality machine* whose efforts drive a company toward the ambitious goals of Six Sigma performance. Since you are reading this introduction, we'll assume that you are a candidate Black Belt or perhaps have already been through "Belt" training.

What will you need, then, to do your job (see Section 2 for a more detailed job description of the Black Belt)? Well, there are four basic principles that you will be applying in your efforts. We'll use these to briefly describe the tools and methods contained in this manual:

Customer Satisfaction – Although it appears elementary, your company survives and grows by satisfying customers with its products and services. To satisfy our customers, we first must understand their needs, wants and requirements. In some cases, you will have to obtain feedback on existing products and services, in others, you will need to gather information to support the design of a new product or service. In both cases, you will need methods that allow you to "talk to" and "listen to" the customer. Section 4 describes *Voice of Customer* methods. Section 15 will describe how to collect and analyze data on one of the *key characteristics* of your products – *reliability.*

Manage With Facts – It is natural, when faced with a problem or opportunity, to want to develop a solution. In many cases, we can solve problems based on a combination of our past experiences and our logic. In other cases, though, we should stop before "jumping" to solution. Whenever we are unclear about the causes of a problem, we should insert an *analysis* step into our problem-solving efforts. Many of the tools in this kit support you in this area. One of the difficult issues we face is that there exists *variation* in all our processes. Learning how to understand process variation and act appropriately is a key element of practicing this principle. Sections 5 through 11 will provide you with the necessary tools and methods.

Plan-Do-Check-Act (PDCA) – The PDCA cycle is a simple one, but hard for organizations to practice. In essence, PDCA asks us to plan our work, do the work, check the results of the work and then act to revise the plan if there are gaps between the actual and desired outcomes. Organizations often have disconnects between these steps – deliberate processes must be put in place to practice PDCA. Section 2 presents the "systems" intended to improve your company's implementation of this principle. Section 14 focuses on how to manage design of products and services.

Respect for People – Finally, we recognize the people "dimension" of quality. There are several aspects to address here. First, at the core, modern quality management adopts the assumption that people are "good." They don't come to work intending to produce failures or defects. Second, to improve, we will ask our staff to change, which is not always comfortable. Third, you will be involved with teams of people who will help you solve problems and improve your business processes. You will need skills to effectively lead and facilitate improvement efforts. Sections 3, 12, 13 & 16 provide you with methods to support this principle.

Good luck with your Black Belt training, qualification and projects. We are sure you will find this time of your career exciting, challenging and rewarding – both for you personally and for your company.

John O'Neill

Sigma Quality Management

Recognition – There are too many people to thank for their contribution and input to this manual. A few, though, that we cannot fail to mention include*:*

Our Counselors at Florida Power & Light: Dr. Teiichi Ando, Prof. Hideo Iwasaki, Dr. Kazuyuki Suzuki, Dr. Hajime Makabe, Dr. Noriaki Kano, Dr. Yoshio Kondo, Professor Asaka,

FPL Thought and Application Leaders: Bob Young, Bill Hensler, Bob Fritz, Cathy Lindbergh, Bruce Sharp, Marie Daverio, Tom Gilmore, Bob Wernly, Brendan Collins, Rick Dobbins, Don Paxson, Kent Sterett

A Few Special Friends: "Dr. Bob" Abernethy, David Wilkerson, Ed Rhew, Ken Maynard, Barbara Reusser, Linda Mills, Bill Lindenfelder, Eric Mattenson, John Willis, Brian Farison and John Gilbert.

2.0 Management Systems

Unit	Description	Page
2.1	Quality Management Systems	2.1 - 1
2.2	Improving Existing Products and Services	2.2 – 1
2.3	Managing Processes	2.3 – 1
2.4	Designing New Products and Services	2.4 – 1
2.5	Strategic Planning Process	2.5 – 1
2.6	Black Belts and Master Black Belts	2.6 - 1

2.1 Quality Management Systems

Learning Objectives

- To understand the purpose, principles and practice of quality management

Unit Contents

- Quality Management Defined
- Quality Management Approaches
- A Baldrige-Award Based Approach

2.1.1 Quality Management Defined

Unfortunately, the word *quality* has several different meanings in the English language. Likewise, the term *quality management (or assurance)* has the same problem. To some, it means simply preventing defects in products or services from reaching the customer. This involves the *negative* aspect of quality management and invokes images of inspection and activities designed to ensure the production process conforms to specifications. However, there is a *positive* aspect to quality management – assuring that the company's products and services offered to the consumer satisfy their *needs and expectations.* Now the image broadens to include corporate planning, market research studies and the design process. This manual will take a *broad view* of quality management:

> *Quality Management includes those systematic activities designed to ensure that consumers can purchase products and services that economically and reliably meet their needs. Quality Management also ensures that the company receives sufficient profit to pay employees, invest in the future and provide an attractive return on shareholders' investment.*[1] *In short, Quality Management is the way we ensure the success of our business.*

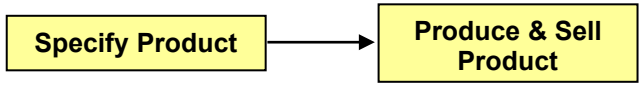
Quality Management operates on just a few basic principles:

Customer and Quality First – The role of any company is to produce products and provide services for their customers. This principle requires that the company focus their efforts on the long-term goal of producing *quality* products and services, not the short-term goal of profits. This kind of company would not ship products or provide services that they knew were defective to meet the short-term goal of monthly production or sales targets. By adopting a long-term focus, the company will assure itself of long-term competitiveness and profits.

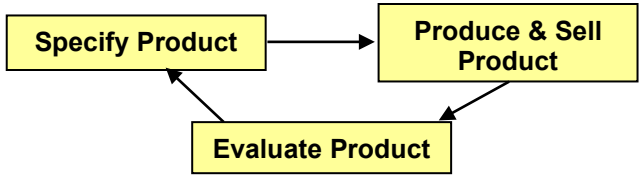
Management by Fact – Decisions in a quality-managed company are based on facts. Experience is still a valuable knowledge commodity, but a scientific and inclusive approach to decision making is necessary. If the customers say they want "vanilla," but management decides to provide "chocolate," then this principle is not being applied. If the problem's solution is the one proposed by the loudest voice in the meeting, then this principle is not being applied. As Dr. Deming would say, "In God We Trust, All Others Must Bring Data."

[1] As one company states their mission: *Completely Satisfying Customers, Profitably.*

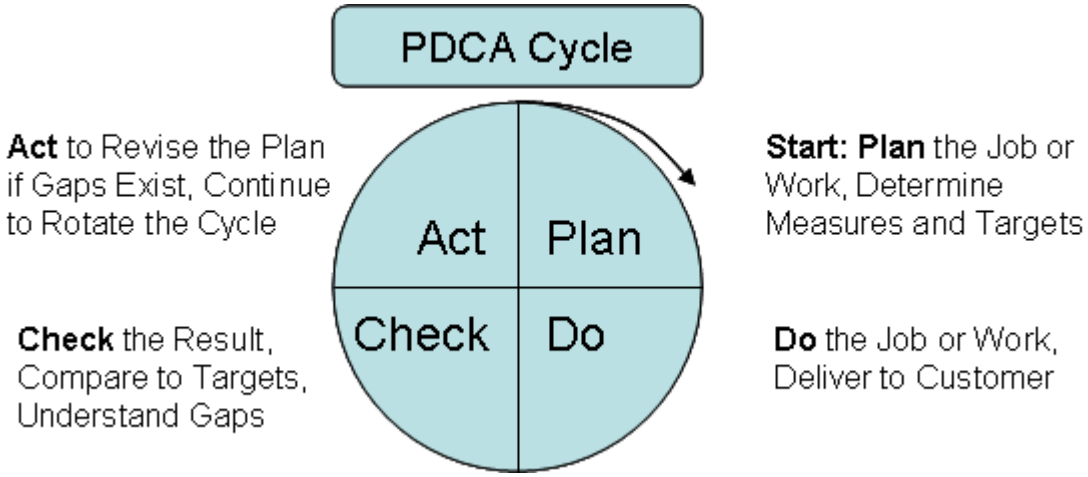
Plan-Do-Check-Act (PDCA) – This is the simplest of the principles, but the most difficult to practice. Companies often employ a linear, "market-out" process to product/service production:



In the 1920's, Walter Shewhart turned this linear process into a feedback loop, adding a "Check" step:



When Dr. Deming introduced the Shewhart cycle to the Japanese in the 1950's, they translated the cycle into the PDCA loop and renamed it the Deming Cycle. They also generalized the cycle to incorporate any kind of work, not just product production:



Respect for People – This last principle has several elements. The first is that all employees must be engaged in quality management. It is not just the job of the "quality" department. Second, a "people are good" assumption pervades all quality management practices. As Deming pointed out, over 80% of problems in the workplace are the fault of the process, not the workers. He clearly lays the prime responsibility for quality at the feet of senior management. Third, the

company must balance the rewards of its endeavors among management, staff and shareholders. Finally, the company owes its employees a safe working environment, as free as possible from injury potential.

One Company that has managed to blend all these into an operating philosophy is Johnson & Johnson. Their Credo is known (and applied!) by all employees. The Corporation has drawn heavily on the strength of the Credo for guidance through the years, and at no time was this more evident than during the TYLENOL® crises of 1982 and 1986, when the company's product was adulterated with cyanide and used as a murder weapon. With Johnson & Johnson's good name and reputation at stake, company managers and employees made countless decisions that were inspired by the philosophy embodied in the Credo. The company's reputation was preserved, and the TYLENOL® acetaminophen business was regained.

Today, company employees participate in a periodic survey and evaluation of just how well the company performs its Credo responsibilities. These assessments are then fed back to the senior management, and where there are shortcomings, corrective action is promptly taken. It is interesting that General Robert Wood Johnson first penned the Credo in 1943 (see next page).

JOHNSON & JOHNSON COMPANY CREDO

We believe our first responsibility is to the doctors, nurses and patients, to mothers and fathers and all others who use our products and services. In meeting their needs everything we do must be of high quality. We must constantly strive to reduce our costs in order to maintain reasonable prices. Customers' orders must be serviced promptly and accurately. Our suppliers and distributors must have an opportunity to make a fair profit.

We are responsible to our employees, the men and women who work with us throughout the world. Everyone must be considered as an individual. We must respect their dignity and recognize their merit. They must have a sense of security in their jobs. Compensation must be fail and adequate, and working conditions clean, orderly and safe. We must be mindful of ways to help our employees fulfill their family responsibilities. Employees must feel free to make suggestions and complaints. There must be equal opportunity for employment, development and advancement for those qualified. We must provide competent management, and their actions must be just and ethical.

We are responsible to the communities in which we live and work and to the world community as well. We must be good citizens – support good works and charities and bear our fair share of taxes. We must encourage civic improvements and better health and education. We must maintain in good order the property we are privileged to use, protecting the environment and natural resources.

Our final responsibility is to our stockholders. Business must make a sound profit. We must experiment with new ideas. Research must be carried on, innovative programs developed and mistakes paid for. New equipment must be purchased, new facilities provided and new products launched. Reserves must be created to provide for adverse times. When we operate according to these principles, the stockholders should realize a fair return.

2.1.2 Quality Management Approaches

The problem of quality management has been with us as long as humans have engaged in economic activity. Here we will provide a brief overview of different approaches. For those interested in the history of quality management, Dr. Joseph Juran has written *The History of Quality* – a fascinating documentary of this topic.

"Old-Fashioned" Approach

Prior to the development of mass-production methods, a large fraction of human economy occurred on a one-to-one basis. A customer would meet with a craftsman and describe what they wanted (e.g. a silversmith or goldsmith for jewelry or a blacksmith for a plough or other tool). The craftsman embodied all corporate "functions" in one person – sales, planning, design, production, and service. This approach helped ensure that the customer's needs were incorporated into the product or service. Since the products were generally produced one at a time, variation between parts was not a problem. The craftsman also acted as the "quality control" function, inspecting the product for flaws or defects.

Inspection Based

With the advent of mass-production and the modern, functionally divided organization, the close connection between the producer and consumer became fragmented. A worker assembling engines in an automobile factory would never see the ultimate customer of the car. To communicate requirements, specifications were developed. To account for inevitable variation in parts, tolerance limits were incorporated into the specifications. Inspection and sorting of parts based on a "go, no-go" conformance to specifications was employed to prevent defects in the product.

Standard Based

Along with the development of inspection-based quality control, the idea of standards for products became widespread. Two major drivers of standards included consumer safety (e.g. explosions in steam boilers on riverboats prompted the development of what's now known as the ASME Boiler and Pressure Vessel Code) and mass-production (e.g. interchangeability of parts such as light bulbs and sockets, train track gage, electric sockets, etc.).

Statistical Approach

In the 1920's, Walter Shewhart, of Bell Laboratories, developed the *statistical* approach to control of quality. His approach incorporates the idea that variation exists in all production processes and that a state of *control* can be achieved through systematic elimination of assignable causes of variation – that due to materials, methods, people, or machines. The incorporation of statistical quality control into the US' wartime "Z" standards is credited as one of the major factors leading to the allied victory (interestingly, Japanese quality texts almost always cite this effect). Through statistical quality control, a rifle produced in one factory could fire bullets produced in another plant.

Deming's Approach

Although W. Edwards Deming's roots are found in the application of statistical quality control on the shop floor, he recognized that quality was the main responsibility of senior management. Without their commitment to continuous improvement, efforts at lower levels in the organization would be fragmented and ineffective. Rather than focus on the "mechanics" of quality management, Deming evolved a set of principles that he stated could be applied by any organization, regardless of what they "produced:"

Deming's 14 Points

1. Create constancy of purpose toward improvement of product and service, with the aim to become competitive and to stay in business, and to provide jobs.
2. Adopt the new philosophy. Western management must awaken to the challenge, must learn their responsibilities, and take on leadership for a change.
3. Cease dependence on inspection to achieve quality. Eliminate the need for inspection on a mass basis by building quality into the product in the first place.
4. End the practice of awarding business on the basis of price tag. Instead, minimize total cost. Move toward a single supplier for any one of item, on a long-term relationship of loyalty and trust.
5. Improve constantly and forever the system of production and service, to improve quality and productivity and thus constantly decrease costs.
6. Institute training on the job.
7. Institute leadership. The aim of supervision should be to help people and machines and gadgets to do a better job. Supervision of management is in need of overhaul, as well as supervision of production workers.
8. Drive out fear, so that everyone may work effectively for the company.

9. Break down barriers between departments. People in research, design, sales and production must work as a team, to foresee problems of production and in use that may be encountered with the product or service.
10. Eliminate slogans, exhortations and targets for the work force asking for zero defects and new levels of productivity. Such exhortations only create adversarial relationships, as the bulk of the causes of low quality and low productivity belong to the system and thus lie beyond the power of the work force.
11. Eliminate work standards (quotas) on the factory floor. Eliminate management by objective. Eliminate management by numbers, numerical goals. Substitute leadership.
12. Remove barriers that rob the hourly worker of his right to pride of workmanship. The responsibility of supervisors must be changed from sheer numbers to quality. Remove barriers that rob people in management and in engineering of their right to pride of workmanship. This means, *inter alia,* abolishment of the annual or merit rating and of management by objective.
13. Institute a vigorous program of education and self-improvement.
14. Put everybody to work in the company to work to accomplish the transformation. The transformation is everybody's job.

Deming's principles also include the following "deadly diseases" and obstacles to improvement:

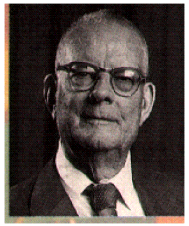
Deadly Diseases

1. Lack of constancy of purpose to plan product and service that will have a market and keep the company in business, and provide jobs.
2. Emphasis on short-term profits: short-term thinking (just the opposite from constancy of purpose to stay in business), fed by fear of unfriendly takeover and by push from bankers and owners for dividends.
3. Evaluation of performance, merit rating, or annual review.
4. Mobility of management; job-hopping.
5. Management by use only of visible figures, with little or no consideration of figures that are unknown or unknowable.
6. Excessive medical costs (unique to the US).
7. Excessive costs of liability, swelled by lawyers that work on contingency fees.

Obstacles to Improvement

1. Hope for instant pudding.
2. The supposition that solving problems, automation, gadgets and new machinery will transform industry.
3. Search for examples without guiding principles.

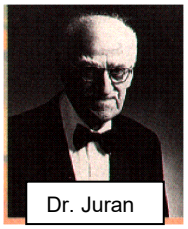
4. "Our problems are different."
5. Obsolescence in schools.
6. Poor teaching of statistical methods in industry.
7. Use of tables and methods for acceptance sampling of incoming or outgoing product.
8. "Our quality control department takes care of all our problems of quality."
9. "Our troubles lie entirely in the work force."
10. False starts.
11. "We installed quality control."
12. The unmanned computer.
13. The supposition that it is only necessary to meet specifications.
14. The fallacy of zero defects.
15. Inadequate testing of prototypes.
16. "Anyone who comes to try to help us must understand all about our business."



Dr. Deming

One of Dr. Deming's last books, *Out of the Crisis,* should be read by all people interested in managing for quality.

Feigenbaum, Juran, Quality Systems and the Japanese Approach



Dr. Juran

Armand Feigenbaum and Joe Juran also recognized that quality management required the cooperation and engagement of the entire organization. Contrasting to Deming's development of fundamental principles, Feigenbaum and Juran took a more application-oriented approach. Feigenbaum coined the term "Total Quality Management" to describe a holistic approach to achieving quality and financial performance. In his book of the same name, Feigenbaum outlines the responsibilities, quality systems, tasks and activities of quality management. Dr. Juran has long been noted for the *Juran Quality Handbook,* the "Bible" of quality management. The quality practitioner can find just about every quality tool in existence described there.

Deming, Juran and Feigenbaum were all influential in steering the post-war Japanese quality efforts. The Japanese credit Deming for providing them quality theory but they credit Juran & Feigenbaum for providing them the practical methods. While American industry ignored the quality message in the 50's and 60's, the Japanese applied their knowledge to key industries, such as automotive and electronics. The impact this has had on the US balance of trade is well known. The Japanese are also well known for their own development of a great deal of quality "technology," such as Quality Function Deployment, Taguchi's methods, Hoshin Planning, Kanban production and others. It is also worthy to note that the Japanese definition (and application!) of Total Quality Management (TQM) is consistent with that of Feigenbaum – a

holistic, organization-wide approach to quality. During the 1980's and 1990's, although the term TQM was used in the US, for the most part, organizations only focused on "local" problem-solving/process improvement, only one component of TQM.

Six Sigma

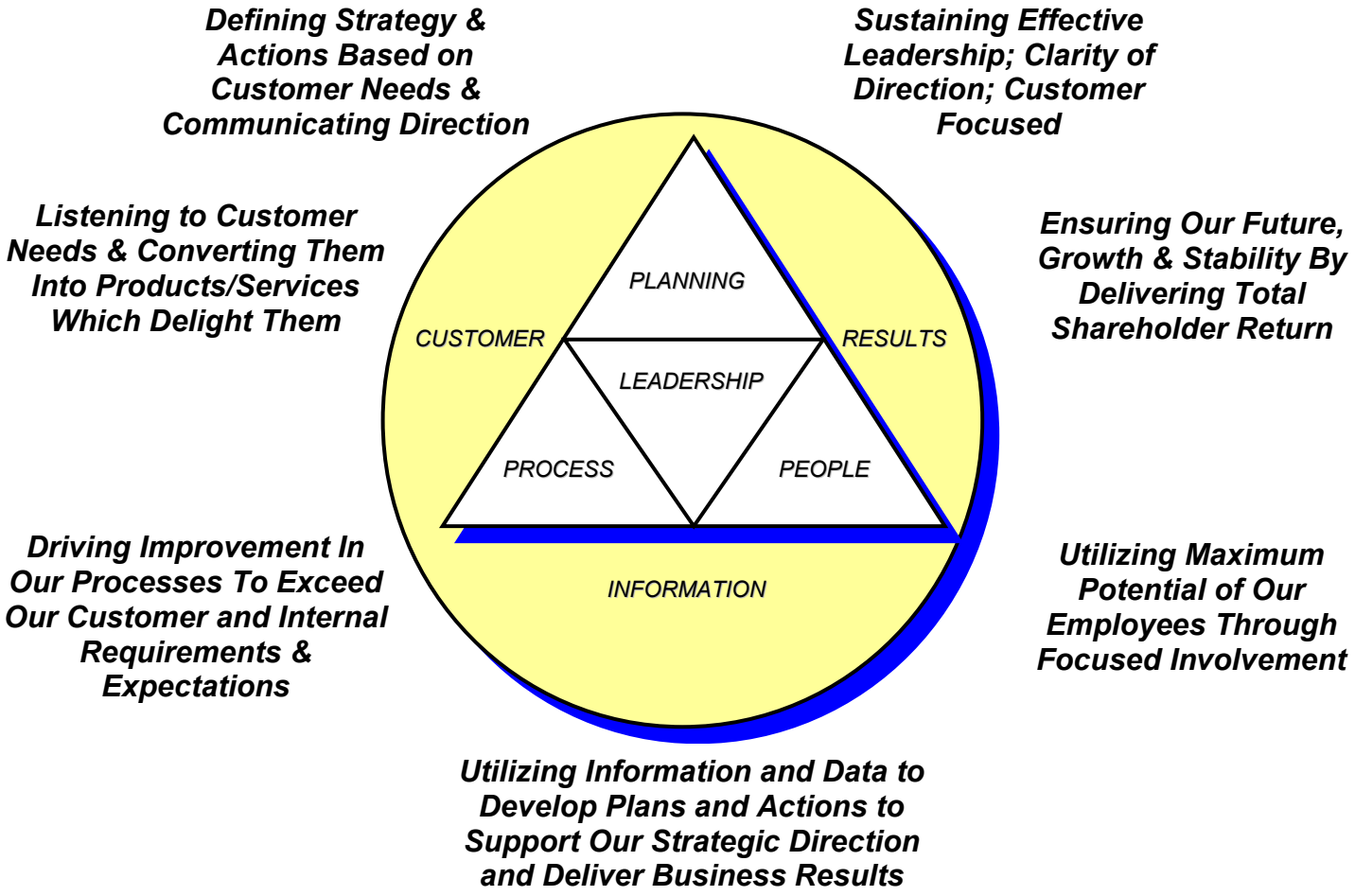
In the 1980's, a new angle on quality management appeared under the banner of Six Sigma. Although Six Sigma's core includes the traditional statistical and quality techniques, the original approach developed by Motorola added a few wrinkles worthy of note. The term Six Sigma simply refers to a process that operates at a short-term process capability of 2.0 – that is the process' standard deviation is $1/6^{th}$ the distance from the target to the specification limit. Over the long-term, such a process can be expected to produce less than 3.4 defects per million opportunities. There is nothing "magic" about this defect rate – it is simply a benchmark that Motorola observed was being achieved by best-in-class companies (typically Japanese!). The Six Sigma term does seem to have a certain appeal to executives as a worthy goal. The fact that Jack Welch of GE has successfully promoted the pursuit of Six Sigma has also enhanced its marketability and (currently) Wall Street looks favorably on companies that announce their pursuit of Six Sigma.

Another element of the Six Sigma approach is the use of a dedicated resource applied to significant corporate problems and opportunities – the Black Belt (see Section 2.6 for more details). The Black Belt model supports two aims. First, the dedicated resource embodied in the Black Belt helps achieve an improvement "velocity" in the organization that is does not occur via other models (e.g. training lots of staff and then engaging them in part time projects). Second, the rotation of the Black Belt back into the line organization after a "tour of duty" can help embed the quality culture in the organization.

The "original" Six Sigma program implemented at Motorola focused on improvement of existing product or service. GE and others have expanded the Six Sigma umbrella to include product and service design and process management. GE's process management method even includes some elements of Hoshin planning. Thus, although there will be different "flavors" of Six Sigma, at least in some companies it is evolving toward a holistic approach, like Total Quality Management.

2.1.3 One Approach – A Baldrige-Based Quality System

The following depicts a structure for a quality system based on the Malcolm Baldrige Quality Award. The Baldrige criteria provide a company with a holistic approach to assessing and improving their quality system.



A Brief Organizational Assessment:

The Malcolm Baldrige National Quality Award applies the following criteria in judging applicant companies. How does your company's current quality system "stack-up" against these criteria?

Category	***Items***
Leadership	Leadership System – *describe the company's leadership system and how senior leaders guide the company in setting directions and in developing and sustaining effective leadership throughout the organization.* Company Responsibility & Citizenship – *describe how the company addresses its responsibilities to the public and how the company practices good citizenship.*
Strategic Planning	Strategy Development Process – *describe how the company sets strategic directions to strengthen its business performance and competitive position.* Company Strategy – *summarize the company's strategy and action plans, how they are deployed and how performance is tracked.*
Customer & Market Focus	Customer and Market Knowledge – *describe how the company determines longer-term requirements, expectations and preferences of target and/or potential customers and markets. Describe this information is used to understand and anticipate needs and to develop business opportunities.* Customer Satisfaction & Relationship Enhancement – *describe how the company determines and enhances the satisfaction of its customers to build relationships, to improve current offerings, and to support customer- and market-related planning.*
Information & Analysis	Selection & Use of Information & Data – *describe how the company determines and enhances the satisfaction of its customers to build relationships, to improve current offerings and to support customer- and market-related planning.* Selection & Use of Comparative Information & Data – *describe the company's selection, management and use of information and data needed to support key company processes and action plans and to*

Category	*Items*
	improve company performance. Analysis & Review of Company Performance – *describe how the company analyzes and reviews overall performance to assess progress relative to plans and goals and to identify key areas for improvement.*
Human Resource Focus	Work Systems – *describe how all employees contribute to achieving the company's performance and learning objectives, through work design, compensation and recognition approaches.* Employee Education, Training & Development – *describe how the company's education and training support the accomplishment of key company action plans and address company needs, including building knowledge, skills and capabilities, and contribute to improved employee performance and development.* Employee Well-Being & Satisfaction – *describe how the company maintains a work environment and climate that support the well-being, satisfaction and motivation of employees.*
Process Management	Management of Product & Service Processes – *describe how products and services are designed, implemented and improved. Describe how production/delivery processes are designed, implemented, managed and improved.* Management of Support Processes – *describe how the company's key support processes are designed, implemented, managed and improved.* Management of Supplier and Partnering Processes – *describe how the company's supplier and partnering processes and relationships are designed, implemented, managed and improved. Describe how supplier and partner performance is managed and improved.*
Business Results	Customer Satisfaction Results – *summarize the company's customer satisfaction and dissatisfaction results.* Financial & Market Results – *summarize the company's key financial and marketplace performance results.*

Category	*Items*
	Human Resource Results – *summarize the company's human resource results, including employee well being, satisfaction, development and work system performance.* Supplier & Partner Results – *summarize the company's supplier and partner performance results.* Company-Specific Results – *summarize company operational performance results that contribute to the achievement of key company performance goals – customer satisfaction, product and service quality, operational effectiveness and financial/ marketplace performance.*

2.2 Improving Existing Products & Services

Learning Objectives

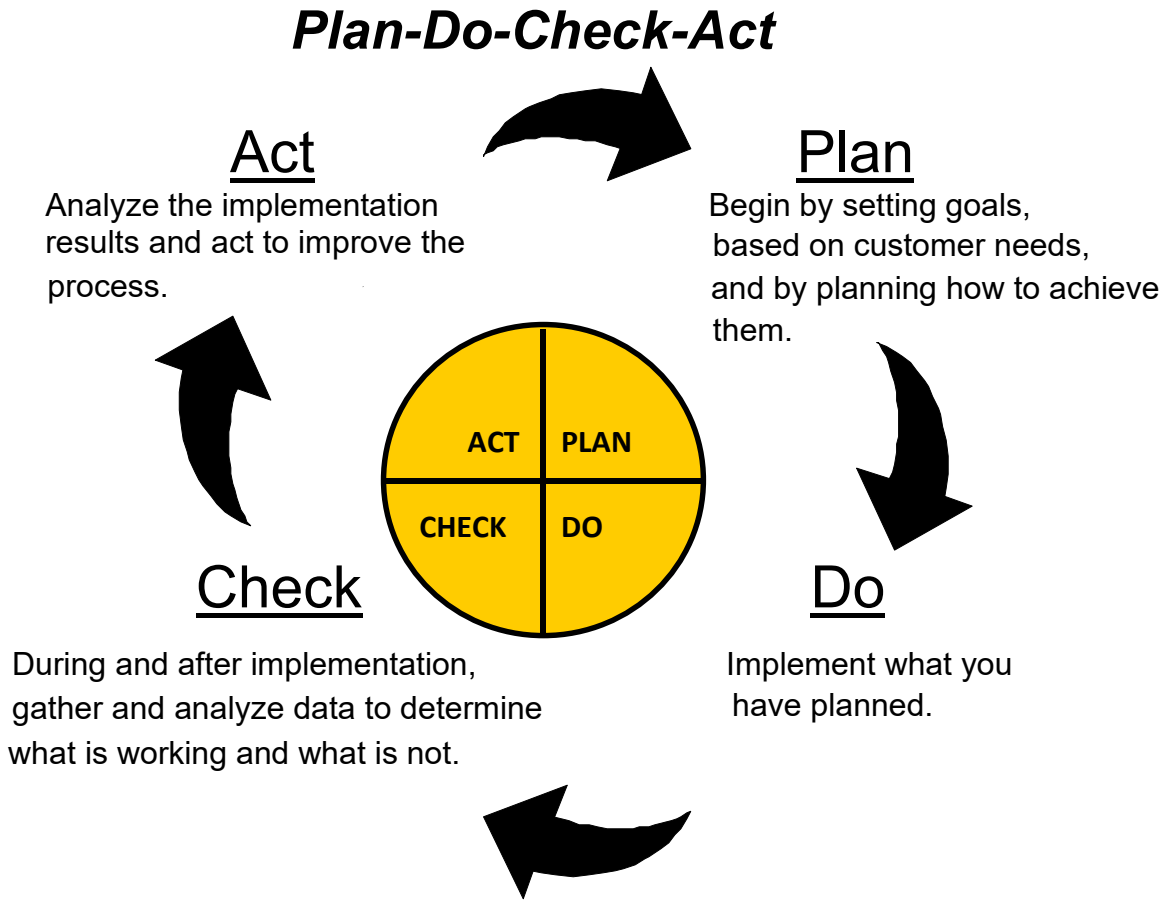
- Understand the DMAIEC Improvement Process

Unit Contents

- Plan-Do-Check-Act
- DMAIEC Improvement Process

Continuous improvement is the goal. As Jack Welch, CEO of General Electric notes, *"If the rate of change on the outside is greater than the rate of change on the inside, then the end is near."* The word *continuous* means ongoing, endless, unbroken, and is figuratively associated with the circle, which embodies these characteristics.

We use the Plan-Do-Check-Act (PDCA) cycle as the core "theory" of our improvement method. In our years working with companies, we've noted that PDCA is easy to understand, but hard to practice.



The DMAIEC Improvement Process

We've translated the PDCA cycle into a practical 6-step approach for teams and individuals to employ during problem solving/process improvement efforts. The method is described on the next few pages. The PDCA wheel is "rotated" more than once in this six-step approach.



Define

- Launch The Project
- Define Outcomes
- Identify Stakeholders
- Select Team
- Determine Project Approach
- Create Project Plan

Measure

- Define The Current Process
- Address "Low-Hanging Fruit"
- Obtain Customer CTQ's
- Gather Initial Metrics, Assess Process Stability
- Determine Current "Sigma"
- Stratify Data
- Determine Initial Value Proposition

Analyze

- Develop Cause & Effect Hypotheses
- Gather Causal Data
- Determine & Validate Root Causes (X's)

Identify

- Identify Breakthroughs
- Select Practical Approaches
- Design Future State
- Predict New "Sigma"
- Perform C/B & Risk Analysis

Execute

- Develop Control Methods
- Develop Dashboards and Scorecards
- Train
- Execute
- Measure Results
- Manage Change

Control

- Report Dashboard and Scorecard Data
- Create Feedback Loop & Adjust Process
- Identify Replication Opportunities
- Develop Future Plans

Define the Reason for Improvement

PURPOSE:	There are many problems could be addressed. You must build a case for why this problem is important to address **now**. Does the problem relate to the most important product/service of your department or is it strategically important to your organization? What is the "gap," what's the "pain?" If the problem is not seen as important, there won't be much enthusiasm to work on it. Also, in this step, the project is planned, team members identified, resources approved.
METHODS:	• **Launch The Project** - Brainstorm a list of problems. Select the most important one to address. Collect customer feedback to identify problems with your products and services. Select the most important one to address. Collect performance data on your products and services (quality, cost, delivery, safety). Pick a product/service with the largest performance "gap." Obtain a copy of your organization's strategic plan. Where and how can your department contribute to the plan? Which of your products/services must be improved? As the output of this step, develop a "theme" or "mission" statement for the effort. • **Define Outcomes** – How will the success of the project be measured? What aspect of the product or service needs improvement - quality, cost, delivery or safety? • **Identify Stakeholders** – Who are the key people who will be impacted by (or who can influence) the project's direction and success? Where do they stand relative to changes that may occur as a result of this project? • **Select Team** – Who should be represented? Full-time? Part-time? • **Determine Project Approach** – DMAIIC provides a general project approach; what specific modifications or additions are needed? • **Create Project Plan** – Develop a work breakdown structure, PERT and/or Gantt chart.
TOOLS:	• Customer Feedback/Complaint Data • Line Graphs, Run Charts, Control Charts • Organization Strategic Plan • Project Planning Worksheet
NOTES:	The first few efforts (practice time!) at improvement may address problems that are not the most important. As improvement skill increases, the problems can become more challenging.

Measure & Observe the Current Situation

<table>
<tr><td>PURPOSE:</td><td colspan="2">This is the “clue-gathering” step. How does the process work that “produces” the product or service to be improved? Here, we should understand the 5W1H (Who, What, Where, When, Why and How) about the process. The problem should be broken down into different aspects or categories - these may be ranked by priority and perhaps only one selected for solution in this improvement cycle.</td></tr>
<tr><td>METHODS:</td><td colspan="2">• Define The Current Process - Clarify how the process that “produces” the product or service works. Develop flowcharts or other “pictures” of the process.
• Address “Low-Hanging Fruit” - Opportunities to improve the process may be identified at this point. Clean up the obvious problems, but don’t make changes unless the root causes are obvious.
• Obtain Customer Needs, Develop CTQ’s – If its not clear what the customer wants from this process, do the necessary research. Interview or survey the customer. Translate the customer’s needs into measurable characteristics of your product or service. If these CTQs differ greatly from your initial theme, discuss changing the project’s direction with your champion or sponsor.
• Gather Initial Metrics – Measure current performance relative to the CTQs. Is the process in control?
• Determine Current “Sigma” - Determine the capability of the process. Express this as a “Sigma.”
• Stratify Data - Examine the problem from different “angles.” Study the variation in the problem. Does the problem occur more often on one shift, with one machine or operator? Look for differences by time, place, type & symptom of the problem. Pareto Analysis can be extremely helpful in isolating one aspect of the problem to address.
• Determine Initial Value Proposition - Clarify the Problem Statement. Often the initial theme or mission statement is very broad. After this step, you should have a more specific problem on which you will continue to work. Write this down before the Analysis step. Try to estimate dollar savings or revenue enhancements based on what you know about the problem now. Determine or refine improvement targets.</td></tr>
<tr><td>TOOLS:</td><td>• Process Flowcharting, Layout Diagramming, Process Watch
• Pareto</td><td>• Variation Studies (Control Charts, Histograms, Capability Analyses)</td></tr>
<tr><td>NOTES:</td><td colspan="2">Don’t try to answer the “why” question here. We sometimes refer to this step as the process “immersion.” When we work with clients on improvement projects, we spend time in this step just watching the work and asking questions along the 5W1H train of thought.</td></tr>
</table>

Analyze the Process

PURPOSE:	This is the "why" or diagnostic step. Where the *Current Situation* step helped us understand the 5W1H of the process, here we will develop hypotheses regarding the variables that are causing the problem or "gap." These hypotheses must then be confirmed or refuted and the "true causes" of the problem identified.
METHODS:	• **Develop Cause & Effect Hypotheses** - Develop hypotheses about why the problem occurs. These may include *Material, Machine/Equipment, Method, People, Measurement* and *Environment* factors. Cause and Effect analysis is the most basic (but powerful!) approach to developing these hypotheses. • **Gather Causal Data** – Plan how you will gather evidence to support your hypotheses. Gather evidence to establish the "guilt" or "innocence" of the different factors. This may be done through analysis of product or service outputs and "production" process factors, or through experiments performed that deliberately change the value of factors in the "production" process. • **Determine & Validate Root Causes (X's)** – Study the results of your cause-and-effect analysis. Which of the potential root causes contribute most to the problem you are attempting to solve. If you eliminate the root cause, how much will the problem be reduced?
TOOLS:	• Cause and Effect Analysis • Pareto Analysis of Causes • Histograms • Scatter Diagram • Process Analysis • Value Analysis/Value Engineering • Twenty Questions • Error Modes and Effects Analysis • Design of Experiments
NOTES:	Understanding *Cause and Effect* is fundamental to the PDCA cycle. In some "advanced" organizations, we've heard the company's everyday language change. When a problem occurred, people used to ask, "*Well, do you know **who** did it?* Now they ask "*Do you understand cause and effect?"*

Identify & Evaluate the Countermeasures

PURPOSE:	Here, changes will be identified that impact the important variables discovered during *Analysis* and that we think will improve performance. The changes should be evaluated for their benefits, costs and possible side effects. The changes must be "sold," planned and then implemented.
METHODS:	• **Identify Breakthroughs** - Identify possible countermeasures to address the process variables affecting performance. Select one or more that have the highest likelihood (and lowest cost) of impacting the variables. Benchmark "best" practices and select the aspects of these that address your situation. Once the countermeasures have been selected, they must be "sold" to the stakeholders (customers, staff, management, etc.). Then, detailed planning and implementation follow. A pilot or demonstration effort may occur prior to "full-scale" implementation. • **Select Practical Approaches** – Translate the countermeasure to a set of changes that can be implemented. *Experiments* may be performed to determine the best "level" for the key causal factors. • **Design Future State** - Design a new product/service or associated production process. In some cases, either the existing product or service is inadequate, or the "production" process cannot produce at the required quality and cost levels. A "clean sheet" design effort may be necessary (see Section 2.4 – *Designing New Products & Services*). • **Predict New "Sigma"** – Given what you know about the countermeasures, what improvement do you expect to see? Will the problem be reduced by 40%, 90%? What will the new "Sigma" of the process be? • **Perform C/B & Risk Analysis** – Are the changes you are suggesting justified by the economics? What risks (business, technical, legal, etc.) are created by the changes? How will the important risks be prevented or mitigated?
TOOLS:	• Root Cause/Countermeasure Matrix • Benchmarking • Cost/Benefit, Risk Analysis • Design Process • Project Planning Worksheet
NOTES:	Making the changes is often the hardest part of the project. Develop a plan to address the expected change resistance. Revisit your stakeholder analysis performed in the *Define* step.

Execute Countermeasures & Check the Results

PURPOSE:	After the changes are made, what effect have they had on performance - has the "gap" closed, or has the problem been eliminated? Do we understand that the changes we made caused the change in performance?
METHODS:	• **Develop Control Methods** – Create or revise the necessary procedures, protocols, drawings, instructions, specifications or other methods employed to control the process. • **Develop Dashboards and Scorecards** – Determine how you will measure the results. The CTQs you have focused on should be measured. Process variables and supplier metrics may also be required. • **Train** – Train workers on the changes to the process. • **Execute** – Implement the changes. You may first make the changes on a pilot scale, prior to full-scale implementation. • **Measure Results** - Collect and Analyze Performance Data to determine if the change has had a measurable impact. Collect data on both the *output* - that aspect of the product or service that you were trying to improve (quality, cost, etc.) and on the *variables* that you changed through the *countermeasures.* Conduct Customer Interviews/Collect Customer Feedback to determine if the problem addressed has "gone away" or has been reduced in frequency. Determine if the *results* (observed changes in performance) are because of the changes you made to the process (sometimes other variables may be acting on the process that are outside your control). Three outcomes are possible here: 1. The results are due to our changes and performance is as expected. Here, move to the *Control* step. 2. The results are much less than expected. Here, go back to *Analyze* and understand why. 3. The results are much better than expected. Here, too, go back to *Analyze* and understand why. • **Manage Change** – Make sure that the necessary changes are being implemented. Address sources of resistance; try to ensure a "win-win" for process stakeholders.
TOOLS:	• Line Graphs, Run Charts, Control Charts • Pareto Analysis • Histograms, Capability Analyses • Procedures, Instructions
NOTES:	One of the most common problems with this step is that organizations do not establish a "baseline" performance - what was performance before the changes were made?

Control the Process & Plan Next Steps

PURPOSE:	The changes may have been made on a pilot basis, or under temporary procedures. If the changes improved the process, then we must ensure that they are repeated each time the product or service is "produced." They must be built into the PLAN, training & education performed and responsibilities clarified. Monitoring tools should be put in place.
METHODS:	• **Report Dashboard and Scorecard Data** – Continue to measure and report on process performance. On-going measurements may occur less frequently and with fewer measurement points than during the pilot phase of the improvement. Monitor *performance* to ensure that the changes aren't *Teflon-coated,* i.e. that they don't "stick." • **Create Feedback Loop & Adjust Process** - Ensure that the performance metrics are acted upon if they go awry. Help staff understand the difference between actions to address process *instability* (e.g. special causes) and process *incapability* (e.g. process not centered or excessive variation relative to specifications). • **Identify Replication Opportunities** – Given that the process improvement has worked well, are there other products/services/processes that could benefit from the changes? • **Develop Future Plans** – What portion of the original problem remains? Is there benefit to be gained by tackling the next "bar on the Pareto?" At the end of the project, don't forget to reward the team and celebrate!
TOOLS:	• Procedures, Protocols, Standards • Training • Line Graph, Run Chart, Control Chart • Quality Improvement Story Review Form • Project Planning Worksheet
NOTES:	"Old habits die hard." Enough said!

2.3 Managing Processes

Learning Objectives

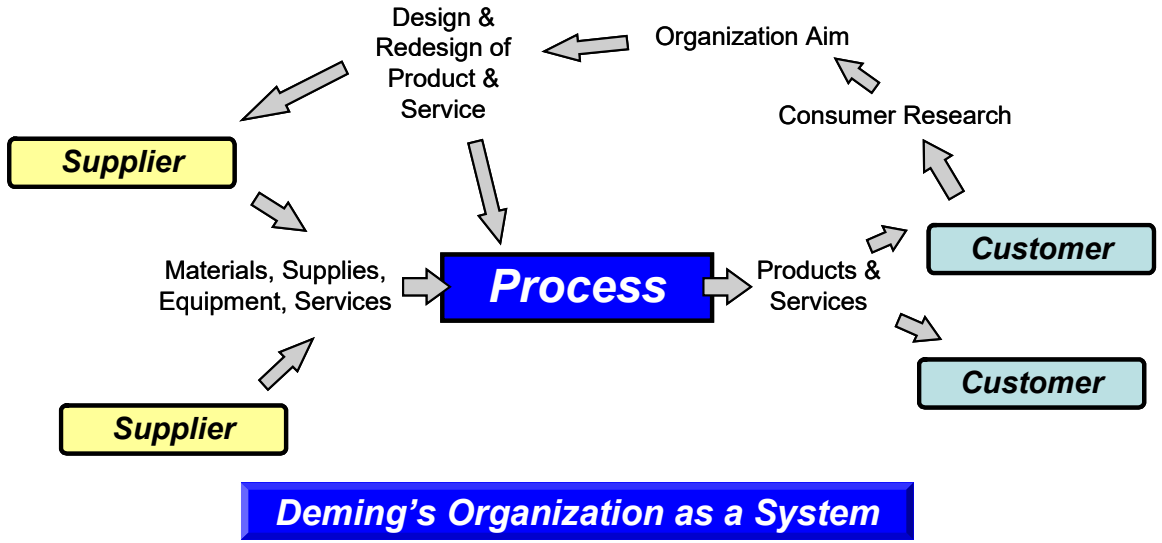
- Understand the Purpose of a Process Management System
- Be able to “Build,” Implement and “Run” a Process Management System
- Link Process Management to Process Improvement and Design

Unit Contents

- Process Management Purpose
- Building, Implementing and Running a Process Management System
- Process Management System Example

2.3.1 Process Management Purpose

In most businesses, managers and supervisors are responsible for organizational *functions,* such as sales, engineering and manufacturing. In many organizations, these functions operate as "silos of excellence" – each trying to be the best it can be. While this behavior is laudable (and often well rewarded!), individual functions must operate together for the organization to be successful. As Dr. Deming pointed out, though, it's the *process* that is responsible for producing products and services that meet the needs of the customer. His *Organization as a System* provides a model for how products and services reach the customer:



Deming's Organization as a System

In many organizations, *basic* work processes are not well defined; the "how-to" do a process is passed from worker to worker, *how well* a particular process is performing is not understood and, when things go wrong, reactions often focus on individual employees rather than the factors in the process responsible. Often, process improvement is difficult because each worker performs the job differently.

Process Management, then, has several quality-related purposes:

Locally – to ensure that work processes are planned and conducted to meet the objectives of the process – at the basic level, to satisfy the requirements of the customers of the process, and

Organizationally – to ensure that the ultimate, external customer of the company is receiving the value they seek from the products and services offered by the company.

Process Management helps the organization's *constitution*, by clarifying responsibilities and accountabilities for company activities. For example, who in your company is responsible for *reliability* of the products? Is it Engineering, Manufacturing, Service (or *all of the above*)?

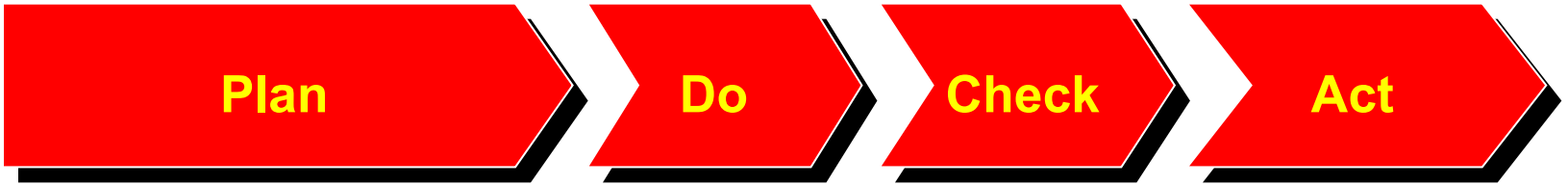
Finally, as Dr. Ishikawa points out, "Without control there can be no improvement, without improvement there can be no control." Any organization that seeks to improve must address the control or management of processes, unless they want to "reinvent the improvement wheel" time and again.



Dr. Kaoru Ishikawa

2.3.2 Building, Implementing and Running a Process Management System

Section Five describes the methods of Process Management. The Plan-Do-Check-Act cycle shows how the concept of continual improvement is embedded in Process Management.



Plan

- Identify/Prioritize Business Processes
- Assign Process Owners
- Define Customer Value & CTQs
- Define Process- Specific Functions & Goals
- Define Process Metrics
- "As-Build" Business Processes
- Prepare Process Control Methods:
  - Procedures
  - Dashboards
  - Response Plan
- Prepare Implementation Plan

Do

- Implement Process Control
- Train & Educate Staff
- Perform the Work
- Monitor Performance
- Take Immediate Remedies for Defects

Check

- Identify Performance Gaps
- Perform Root Cause Analysis (DMAIIC)
- Redesign Processes (DMEDVI)
- Conduct Process Management Reviews

Act

- Conduct System Performance Reviews
- Review For Linkage To Strategic Objective(s)

The basic elements of a Process Management System include:

Process Owners – Managers, Supervisors or Individuals responsible for the outcome of a company process. In some cases, especially for broad corporate processes (e.g. Order to Receipt), a team of managers will be assigned as Process Owners.

Process Purpose, Definition, Measurement and Action Plan – The company should define (and record) why the process exists – who are the customers of the process, what are their needs & requirements, what *key characteristics* of the process must be assured to meet the customers' needs? Process Definition usually includes some graphic picture such as a flowchart defining how the process operates (and, often, who is accountable for the various process steps). Measurement of both output-type variables (e.g. quantifying the quality, cost, delivery and safety *key characteristics*) and important input-type variables (key factors influencing the output variables) is put in place. *Performance dashboards* are often used to summarize the overall performance of the process. Action plans are developed to describe *immediate remedies* when the process variables do not exhibit a state of statistical control or when they produce output beyond the process' specification limits as well as plans to *prevent the recurrence* of chronic process performance gaps. Here we see the link to process improvement. Problem-solving teams can be assigned to analyze the root causes of these performance gaps and develop/implement actions to address the root causes.

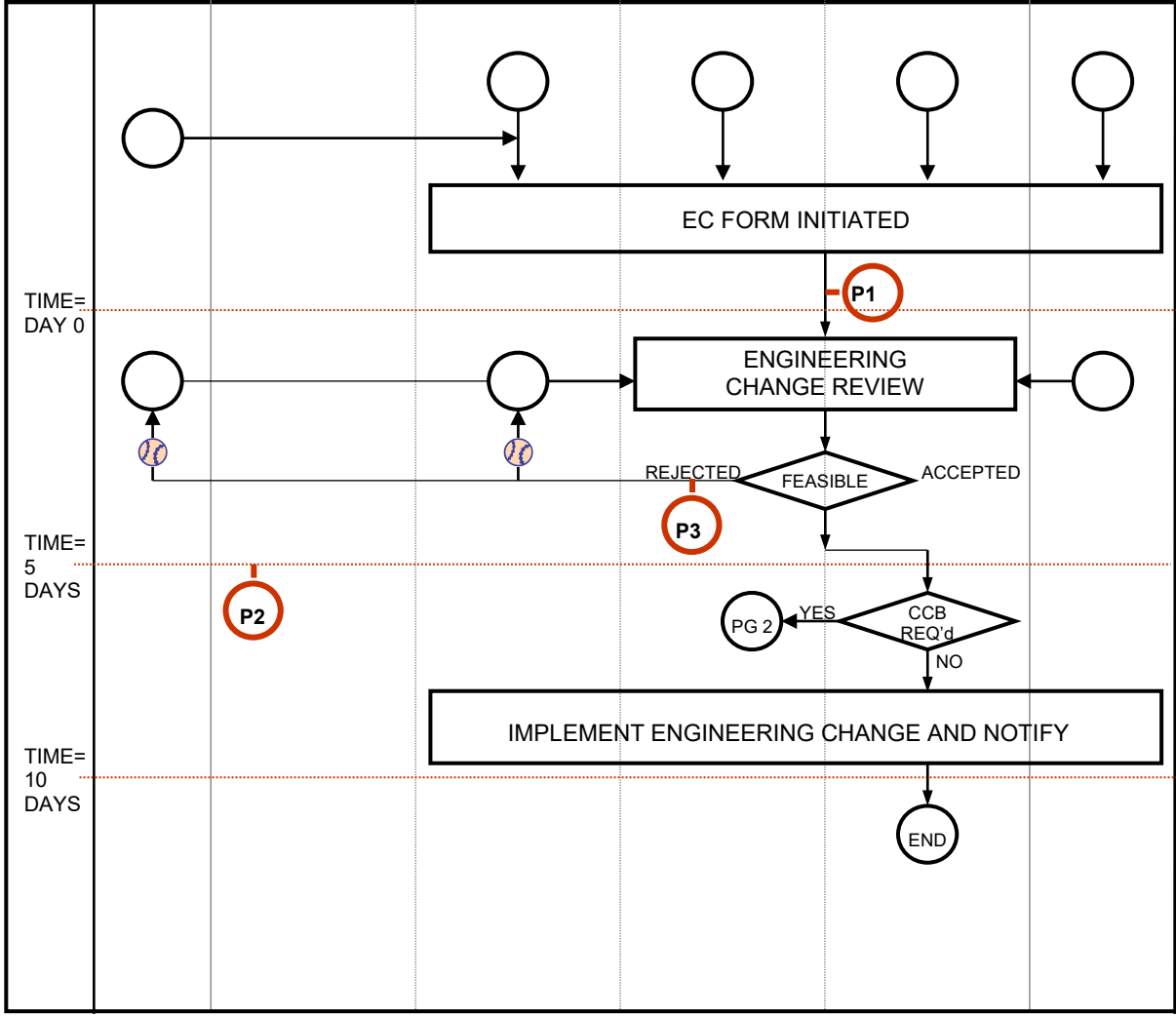
Training, Education, Application and Review – Completing a flowchart and developing a few process performance indicators is not Process Management. Often, the hardest part of Process Management is to educate management and staff in the *why, how, where* and *when* to employ the Process Management system. Dr. Ishikawa is clear in his expectation that management train workers in the production methods. Dr. Deming relates how many control charts wind up as "wallpaper" on company bulletin boards. The discipline of process management is hard, but worth the investment. Periodically, the company should examine how well process management activities are proceeding, look for and analyze gaps and then take corrective action to improve the process management system.

2.3.3 Process Management System Example

The following pages present an example of a Process Management System being developed for an Engineering Change Request Process. The Process Management Charts describe the process, accountabilities for process activities, how performance of the process is to be measured, and what actions occur when the process' performance experiences gaps from target:

TITLE:	PROCESS CUSTOMER:	CUSTOMER VALID REQUIREMENTS:	PROCESS OUTCOME MEASURE[S]:
ENGINEERING CHANGE MANAGEMENT SYSTEM	PROD. CUSTOMERS	RAPID IDENTIFICATION AND INCORPORATION OF DESIGN CHANGES	

PROCESS FLOW CHART



PROCESS INDICATORS		CHECKING PLAN			NOTES
REF. #	TARGET RANGE	WHAT TO CHECK	WHEN TO CHECK	WHO TO CHECKS	
P1	?	# OF EC REQ	1 PER MO.	ECR STEER COMM	
P2	?	TIME IN FROM REQ.	1 PER EC REQ.	ECR STEER COMM	
P3	?	# OF EC REQ	1 PER EC REQ.	ECR STEER COMM	

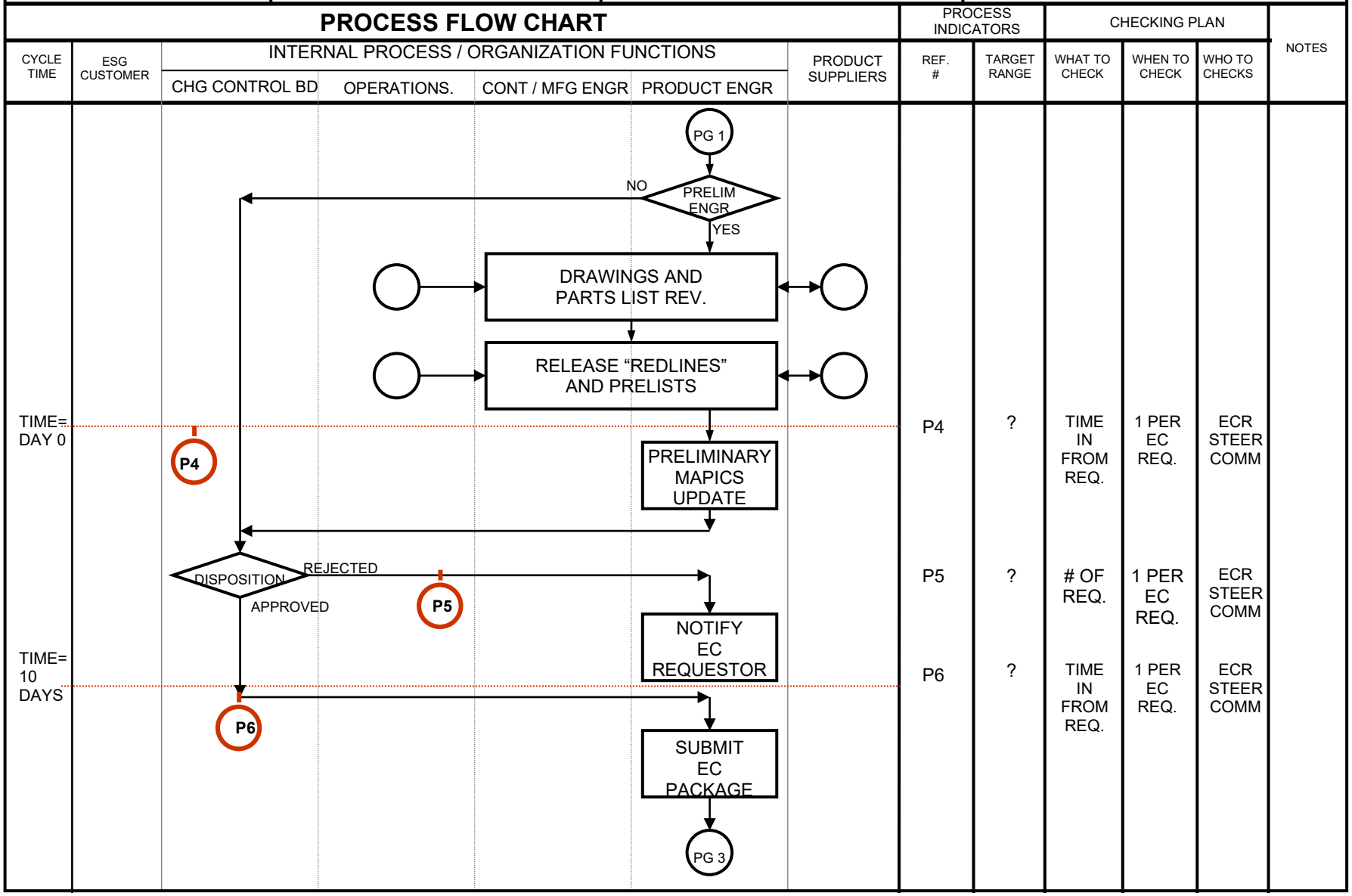
PROCESS CONTROL FORM

PROCESS OWNER: Vice President Operations

REVISED: 5/8/XX PAGE 1 OF 4

TITLE:	PROCESS CUSTOMER:	CUSTOMER VALID REQUIREMENTS:	PROCESS OUTCOME MEASURE[S]:
ENGINEERING CHANGE MANAGEMENT SYSTEM	PRODUCT CUSTOMERS	RAPID IDENTIFICATION AND INCORPORATION OF DESIGN CHANGES	

PROCESS FLOW CHART



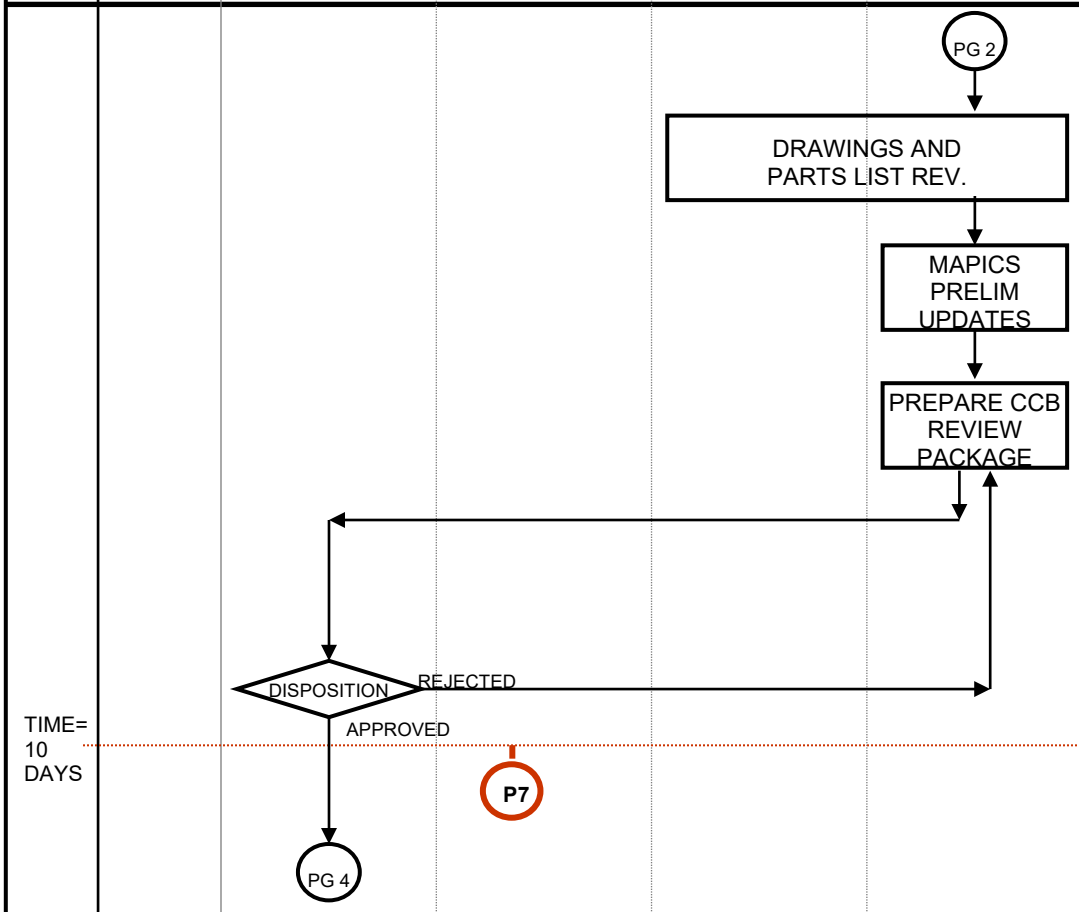
PROCESS CONTROL FORM

PROCESS OWNER: Vice President Operations

REVISED 5/8/XX PAGE 2 OF 4

TITLE: ENGINEERING CHANGE MANAGEMENT SYSTEM	PROCESS CUSTOMER: PRODUCT CUSTOMERS	CUSTOMER VALID REQUIREMENTS: RAPID IDENTIFICATION AND INCORPORATION OF DESIGN CHANGES	PROCESS OUTCOME MEASURE[S]:

PROCESS FLOW CHART

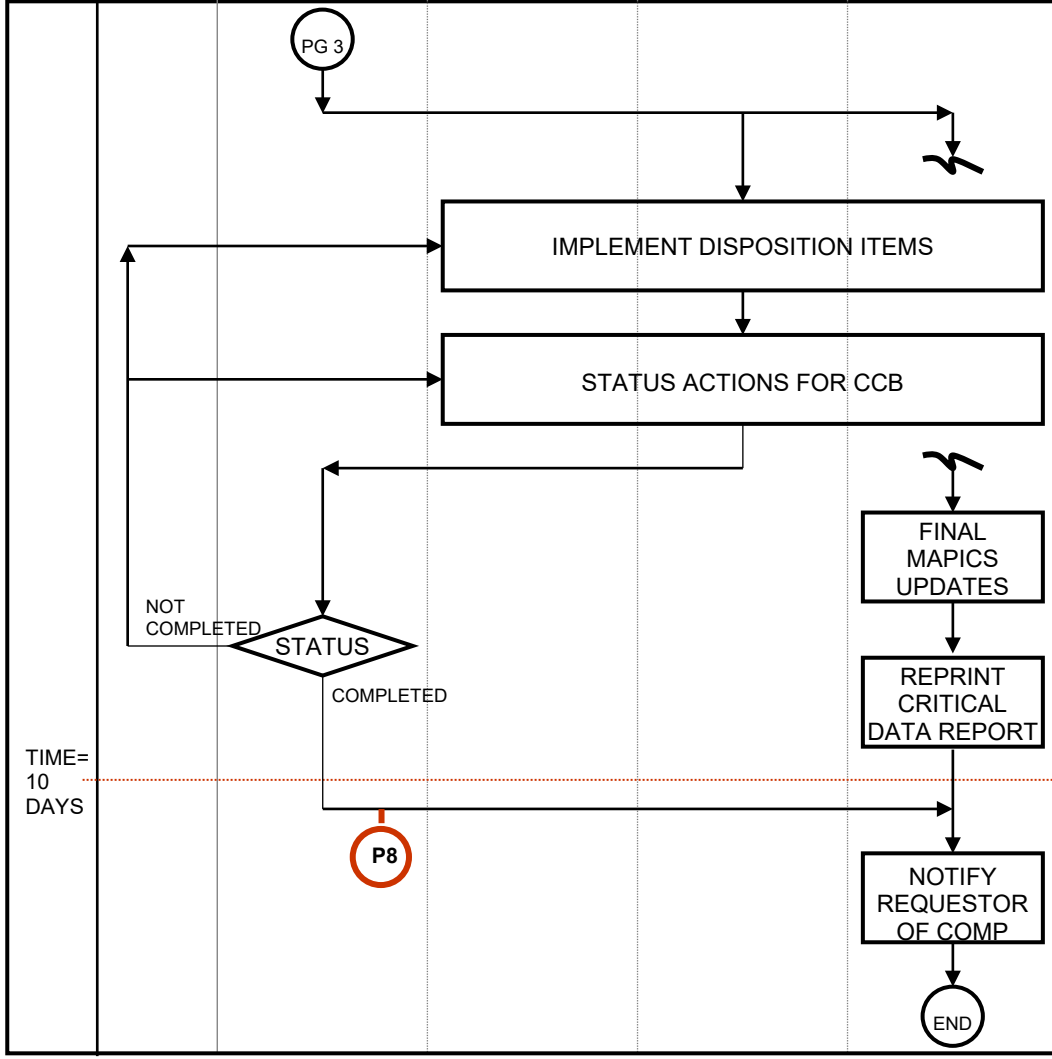


PROCESS INDICATORS		CHECKING PLAN			NOTES
REF. #	TARGET RANGE	WHAT TO CHECK	WHEN TO CHECK	WHO TO CHECKS	
P7	?	TIME IN FROM REQ.	1 PER EC REQ.	ECR STEER COMM	

PROCESS CONTROL FORM | PROCESS OWNER: Vice President Operations | REVISED 5/8/XX PAGE 3 OF 4

TITLE: ENGINEERING CHANGE MANAGEMENT SYSTEM	PROCESS CUSTOMER: PRODUCT CUSTOMERS	CUSTOMER VALID REQUIREMENTS: RAPID IDENTIFICATION AND INCORPORATION OF DESIGN CHANGES	PROCESS OUTCOME MEASURE[S]:

PROCESS FLOW CHART



PROCESS INDICATORS		CHECKING PLAN			NOTES
REF. #	TARGET RANGE	WHAT TO CHECK	WHEN TO CHECK	WHO TO CHECKS	
P8	?	TIME IN FROM REQ.	1 PER EC REQ.	ECR STEER COMM	

PROCESS CONTROL FORM — PROCESS OWNER: Vice President Operations — REVISED 5/8/XX PAGE 4 OF 4

2.4 Designing New Products & Services

Learning Objectives

- Be able to apply the design process to develop new products and services.

Unit Contents

- The DMEDVI Design Process

2.4.1 Introduction

This unit presents a process for designing and redesigning products and services. The design/redesign process is fundamentally different than the improvement process (e.g. DMAIEC). Instead of the “narrowing” approach employed in problem solving to identify root causes of product or process performance problems, the design process is a “broadening” one. To develop a new product or service, we must identify **all** the customers (prioritizing may be necessary in each of these steps), then identify **all** needs and expectations that we are trying to meet, identify the characteristics of the product or service that will enable the product/service to achieve the desired customer needs, design the product or service itself and then design the "production" process that will enable the product or service to be produced.

Literally hundreds or thousands of quality characteristics must be considered and there may be hundreds of thousands of variables that need to be planned, designed, and controlled/managed. Often, needs and expectations of different customer groups will be opposed to each other. For instance, the customer of a home air conditioning system wants the system to both cool his/her home quickly, but also do so without consuming a great deal of electricity. The "broad" approach must include provisions to *balance* these opposing needs.

Some products and services turn out to be wholly inadequate in meeting customer requirements. This could occur because of a poor or degraded product or service or because of changing customer requirements. In some cases, an organization will recognize the need to develop an entirely new product or service, based on their understanding of current or future customer requirements. In this case, we must adopt an approach that is somewhere between the “narrowing” improvement process and the “broadening” design process.

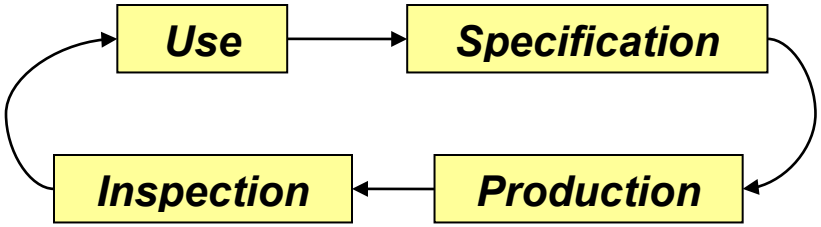
Although many of the methods and tools used to improve quality using the "narrowing" approach are common to those used to address a process "broadly" (i.e. flowcharting, cause and effect, Pareto), the approach taken is quite different.

This unit presents a "generic" path for designing or redesigning products and services (as well as the processes producing these products and services). The concept behind this path is to implement the "market-in" philosophy of developing products and services that meet the needs and expectations of their consumers and customers. The "Plan-Do-Check-Act" cycle is also firmly embedded in this path. In 1924, Walter Shewhart turned the linear, "product-out" process of developing products & services into a loop where information from the market is gathered, analyzed and fed back into the design's specifications:

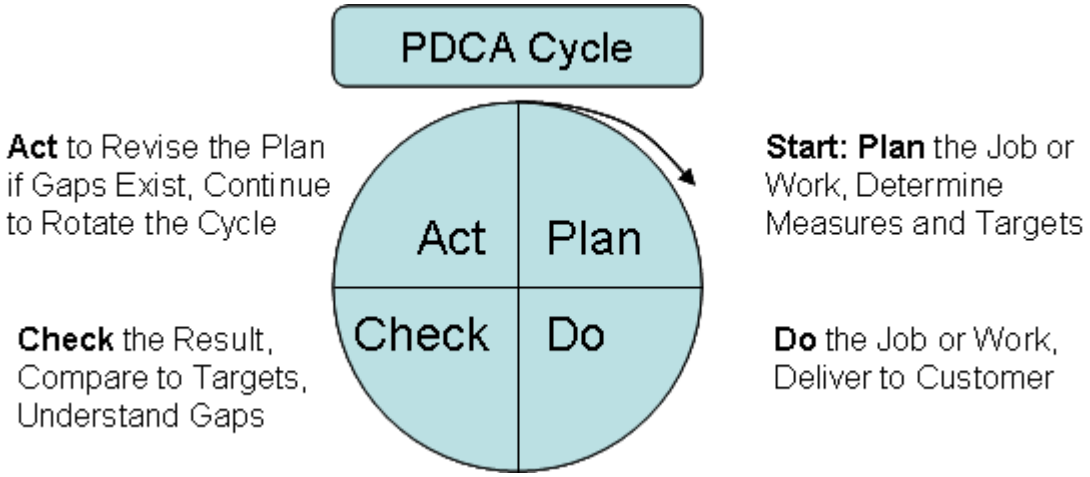
From:



To:



This feedback loop is best known today as the "PLAN-DO-CHECK-ACT" or PDCA cycle.



2.4.2 The Design Process

The figure on the next page is a "generic" design process, to be applied to either designing or redesigning products or services. Some general notes applicable to this process are described below. Starting on page 2.4-6, each of the Design Process steps is discussed as a "mini-process," with purpose, inputs, actions, outputs and suggested methods.

Tailoring - Each design project should develop a plan to implement this design process that is *tailored* to the specific project. For example, the plan associated with designing a *nuclear power plant* would be much different than the plan to develop a *mall,* a *toothbrush*, or a new *screw compressor*. The descriptions of the individual steps that follow the process flowchart are intended to be *very* general. In practice, the *tailored plan* would contain a subset of the activities described herein.

Timing - The design process is shown as a linear flow. In practice, the major steps both overlap and are often iterative in nature. Recent efforts have been directed at shortening the design cycle (time to develop/implement the design). *Concurrent engineering* attempts to perform as many steps in parallel as possible as well as integrate the product design with production process design.

Team - This design process invokes the concept of a "design team" drawn from the organization's product planning, market research, design, research and development, production engineering, and other departments (perhaps supplemented with internal or external consultants, vendors, or customers). This team is formed specifically for the design project and disbands when the product or service has been turned over to the production forces. The composition of the team may, of course, vary as the design progresses and as the project moves from design to implementation.

Terminology - The language used here is consistent with other quality "lingo" that has been explained previously. One distinction we would like to make is the difference between the **product/service** and the **production process** responsible for producing the product/service. It's easy to see the difference between a product like an automobile and its production process. Our design/redesign effort might focus on the product (i.e. designing the new Mustang) or it might focus on the production process (i.e. implementing a *Just in Time* inventory system). For services, the distinction between "service" and "production process" blurs. Dr. Juran, though, strongly recommends that in this "designing process," we try to distinguish between *what* we want to achieve (the *service*) and *how* we provide that service (the *production process*). For instance, the *service* we would like to provide to you may be described as *transferring knowledge and skills associated with designing products and services.* That's the *what.* In part, we have decided to write this section. That's the *how.* We have found this *what* vs. *how* distinction useful and employ it here.



Define

- Launch The Project
- Define Outcomes
- Scope Project
- Identify Stakeholders
- Select Team
- Determine Project Approach
- Create Project Plan
- Define Project Controls

Measure

- Identify Customers
- Define State of Current Customer Knowledge
- Develop & Implement Customer Research Plan
- Translate Customer Needs to Product/Service CTQ's
- Specify Targets, Tolerance Limits & Sigma Targets

Explore

- Develop Product/ Service Necessary Functions
- Develop Conceptual Product/ Service Designs
- Develop High-Level Production Processes
- Predict Capability & Evaluate Gaps

Design

- Develop Detailed Product & Service Designs
- Develop Detailed Production Processes
- Refine Capability & Gap Evaluation, Perform Tradeoffs
- Develop Process Control & Validation Plans

Validate

- Build Pilot Processes
- Validate Pilot Readiness
- Perform Pilot Testing
- Analyze Gaps, Determine Root Causes
- Evaluate Scale-up Potential
- Develop Implementation & Transition Plans

Implement

- Build Full-Scale Processes, Train Staff
- Perform Start-up Testing
- Analyze Gaps, Determine Root Causes
- Transition to Process Owners
- Evaluate & Close Design Project

◆ - Design Review

Define the Product/Service to be Designed

PURPOSE:	The *Define* Step of DMEDVI is like that of DMAIEC. A clear link to the company's product development priorities (perhaps articulated in the Business Plan) is established. By the end of this phase, the product or service to be designed is clarified, the overall scope of the project defined, the project team is in place and necessary plans and design controls developed.	
INPUTS:	• Business Plans • Competitive Information	• Customer Needs (High-Level) • Market Research
STEPS:	• **Launch The Project** – Decide that this product or service should be designed/redesigned (based on market research, company strategy, customer input). Assign overall responsibility for the project. • **Define Outcomes** – Determine how the success of the project will be measured (typically from a business standpoint). Will the design/redesign reduce costs, increase revenue or market share? • **Scope Project** – Determine the boundaries of the project. Determine the project deliverables, what is in and out of scope for the project. Product/Service designs may be divided into "generations." • **Identify Stakeholders** – Who will be impacted by the new design, who can impact the success of the design project? • **Select Team** – Determine full and part-time members of the team. Which disciplines or departments should be involved? • **Determine Project Approach** – DMEDVI provides a generic framework; determine how DMEDVI will be tailored to the specific project. • **Create Project Plan** – Develop a work breakdown structure, PERT and/or Gantt chart. • **Define Project Controls** – Develop communication plans, change control (for the design), change management (for stakeholders, staff), review plans (design and tollgate), risk management processes.	
OUTPUTS:	• Project Charter • Project Plans	• Design Process Controls
TOOLS:	• Process Capability Studies • Competitive Analyses • Benchmark Studies • Market Research Studies	• Analysis of Governmental/Regulatory Requirements • Multi-Generation Product/Service Planning • New Product/Service Introduction Process
NOTES:	Don't jump too quickly into the design project before ensuring there is a sound business case for the product/service. Also, don't forget to establish and implement the design control processes (including communication and change management).	

Measure the Customer Requirements

PURPOSE:	In the *Measure* step, you will obtain the "voices" of the various customers of the product or service. These will include those customers external to the business, perhaps internal customers and the stakeholders who will be impacted or may impact the success of the project (e.g. management, regulatory bodies, others). The goal of this step, though, is to develop a set of requirements (some of which will be *critical-to-quality, i.e. CTQs*) that the design team can use as inputs to their design processes. A clear linkage between the "voices" and the requirements must be established in this step.	
INPUTS:	• Project Charter • Preliminary Voice of Customer • Existing Customer Information	• Market Research Studies • Multi-Generational Plan
STEPS:	• **Identify Customers** – Determine external, internal customers; review stakeholder list generated in *Define Step.* • **Define State of Current Customer Knowledge** – Review existing customer information, including complaints, complements, and market research studies. • **Develop & Implement Customer Research Plan** – Determine what information must be collected, determine appropriate Voice of Customer methods (interviews, focus groups, surveys). • **Translate Customer Needs to Product/Service CTQ's** – The Voice of the Customer is generally obtained in their "language." A filtering and *translation* process takes the customers' voices as input and develops a set of requirements stated in the technical language of the product/service. • **Specify Targets, Tolerance Limits & Sigma Targets** – Numerical goals are set for the product/service requirements. Allowable variation and defect rates (i.e. sigma targets) are established to help the design team objectively judge their design.	
OUTPUTS:	• Product/Service Requirements (a subset of which are the *Critical-to-Quality Requirements (CTQs)*). • Design Scorecard – CTQ level	
TOOLS:	• Voice of Customer Tools • Quality Function Deployment • Kano Analysis • Conjoint Analysis	• Market Research • Affinity Sort • Structure Tree

Explore Design Concepts

PURPOSE:	Here, the design team will identify and evaluate possible design concepts to meet the requirements defined in *Measure.* The decisions made in this step will determine a large percentage of the ultimate quality and cost of the product or service. Moving too quickly through this important step can limit the potential market for and ultimate success of the product or service. Once the "best" concept has been selected, the team will begin to develop the production version of the product and design the necessary production processes. Before "too much" of the design energy is spent, the team will attempt to verify if their design will meet its requirements, through capability assessments.
INPUTS:	• Project Charter • Multi-Generation Product/Service Plan • Product/Service Requirements • Design Scorecard – CTQ level
STEPS:	• **Develop Product/ Service Necessary Functions** – Functional analysis takes a complex product or service and breaks down the "what's" that must occur for the requirements (e.g. CTQs) to be met. This analysis sets the stage for identification of product/service concepts. • **Develop Conceptual Product/ Service Designs** – Benchmarking, structured invention (e.g. TRIZ) and other creative methods are employed to identify concepts for the product/service functions. The various functional concepts are "assembled" into an overall product/service concept. Alternative concepts are evaluated and a "best-fit" selected. • **Develop High-Level Production Processes** – To ensure that the product/service can be "built," the production process elements of *process, information system, human, facility, equipment,* and *supplies* are developed at a high-level. • **Predict Capability & Evaluate Gaps** – Depending on the product/service requirements, analyses, predictions and prototype tests are made to assess the ability of the concept to meet requirements.
OUTPUTS:	• Functional Analyses, correlated to Goals/Needs, • Conceptual Designs, • Design Drawings (Layouts, Flowcharts, Schematics), • Bills of Material, • Supporting Analyses/Test Results: • Test Plans/Results, • Calculations, • Experimental Results • Predictive Analyses (FMECA, EMEA, FTA, Stress Analyses, Manufacturability, Assembly), • Regulatory Impact/Environmental Impact Analyses, • Product/Service Specifications, • R&D Results, • Models/Prototypes of the Product/Service, • Value Analysis, Value Engineering Studies, • Cost Estimates to produce product/service.

TOOLS:	• Quality Function Deployment • Customer Needs/Functions Matrix • Functional Analysis Breakdown (Tree Diagram) • Benchmark Analysis (functional) • Calculations • Modeling/Prototypes/Simulations • Design of Experiments	• Testing • Design Control/Configuration Control • Value Analysis/Value Engineering • Design Review (see next step) • Reliability Methods • Economic Decision Analyses • Design Reviews
NOTES:	During this phase, the decisions that are made will determine about 80% of the “ultimate” quality achieved by the product/service.	

Design Detailed Product, Service and/or Processes

PURPOSE:	Here, the "rest" of the design is developed. The production version of the product is finalized, as well as that of the production processes. Verification activities are refined and completed; the product has been determined to meet its requirements; tradeoffs are made where necessary. In preparation for validation of the design, process controls and validation plans are developed.	
INPUTS:	Detailed: • Functional Analyses, correlated to Goals/Needs, • Conceptual Designs, • Design Drawings (Layouts, Flowcharts, Schematics), • Bills of Material, • Supporting Analyses/Test Results: • Test Plans/Results, • Calculations, • Experimental Results	• Predictive Analyses (FMECA, EMEA, FTA, Stress Analyses, Manufacturability, Assembly), • Regulatory Impact/Environmental Impact Analyses, • Product/Service Specifications, • R&D Results, • Models/Prototypes of the Product/Service, • Value Analysis, Value Engineering Studies, • Cost Estimates to produce product/service.
STEPS:	• **Develop Detailed Product & Service Designs** – The work done in the Explore step is continued at the detailed level. By this step's completion, the design will be developed to the point where it can be produced using production equipment and processes. • **Develop Detailed Production Processes** – Likewise, the production process design is complete. • **Refine Capability & Gap Evaluation, Perform Tradeoffs** – Final testing and product verification activities are completed. • **Develop Process Control & Validation Plans** – In preparation for pilot testing and validation efforts, the necessary process controls – procedures, protocols, bills of material, device master record, etc. are developed.	
OUTPUTS:	Detailed: • Functional Analyses, correlated to Goals/Needs, • Conceptual Designs, • Design Drawings (Layouts, Flowcharts, Schematics), • Bills of Material, • Supporting Analyses/Test Results:	• Predictive Analyses (FMECA, EMEA, FTA, Stress Analyses, Manufacturability, Assembly), • Regulatory Impact/Environmental Impact Analyses, • Product/Service Specifications, • Procedures, Protocols

	• Test Plans/Results, • Calculations, • Experimental Results	• Value Analysis, Value Engineering Studies, • Cost Estimates to produce product/service.
TOOLS:	• Quality Function Deployment • Customer Needs/Functions Matrix • Functional Analysis Breakdown (Tree Diagram) • Benchmark Analysis (functional) • Calculations • Modeling/Prototypes/Simulations • Design of Experiments	• Testing • Design Control/Configuration Control • Value Analysis/Value Engineering • Design Review (see next step) • Reliability Methods • Economic Decision Analyses • Design Reviews

Validate Product, Service & Process

PURPOSE:	Whereas *verification* confirms the product meets its requirements, *validation* confirms the product (and processes) meet the needs of the customers. Pilot testing is a key part of the product/service's validation. Based on the results of these activities, the decision to scale up to full production is made; implementation and transition plans to support scale-up are developed.
INPUTS:	• Design Outputs from Previous Steps • Process Control Plans • Product/Process Validation Plans
STEPS:	• **Build Pilot Processes** – Production facilities, equipment, information systems, etc. are procured and constructed in preparation for pilot tests. • **Validate Pilot Readiness** – Startup testing of the production processes is completed. The processes are tested to determine if they can produce the product/service. • **Perform Pilot Testing** – Production version product (or service) is produced. The product or service is offered to customers; validation that the product/service meets the needs of the users is performed. • **Analyze Gaps, Determine Root Causes** – Problems experienced by the customer are identified, root causes determined and the product/service/process revised to eliminate the gaps. • **Evaluate Scale-up Potential** – A business decision is made to scale-up the product/service to "full-scale." • **Develop Implementation & Transition Plans** – Plans to fully implement the product/service are developed.
OUTPUTS:	• Validated production processes • Validated product/service • Implementation/transition plans
TOOLS:	• Pilot Testing • Design Reviews • Root Cause Analysis Tools • Project Management Tools
NOTES:	*Validation* tests the ability of the product/service to meet customer needs; previous *verification* activities have tested the product/service against requirements derived from customer needs.

Implement the New Product or Service

PURPOSE:	Here the product or service is launched. The design is *transitioned* to the operating forces, e.g. for products, the *Device Master Record* is completed and transferred to production. Although further commercialization of the product or service may occur, and the design of the next product/service generation has begun, the close of this design project is at hand. Lessons learned are documented, as well as the history of the design. As appropriate, the design team is rewarded & recognized.	
INPUTS:	• Validated production processes • Validated product/service	• Implementation/transition plans
STEPS:	• **Build Full-Scale Processes, Train Staff** – For many products/service, existing facilities are adapted to support the new processes. In some cases, though, new production facilities/processes will be required. • **Perform Start-up Testing** – Necessary testing of the new production processes is performed. Production is ramped up to full-scale. • **Analyze Gaps, Determine Root Causes** – Problems noted with early production units/processes are identified, root causes determined and appropriate countermeasures implemented. • **Transition to Process Owners** – As the new product/service enters production, the design team performs a turnover to operating forces. Bills of material, device master records, process procedures, and control plans are completed. Design history files are updated. **Evaluate & Close Design Project** – Before the design team disbands and begins to work on the next products/services, lessons learned are generated, good practices recognized, improvement opportunities identified. Both of these should be fed back to the "owners" of the design process to improve the overall design processes.	
OUTPUTS:	• Commercialized Product/Service • Design History Files	• Lessons Learned • Updated Multi-Generation Plans
TOOLS:	• Project Management • Process Control Methods	• Root Cause Analysis Methods • Configuration/Change Management

Design Reviews

Purpose – Design reviews are an effective means of reviewing the developing design at key stages to determine:

- Which conceptual design to pursue,
- Conformance of the design to goals/needs & expectations,
- Ability to fabricate/assemble or manufacture the design,
- Results of product/service "prototype" testing,
- Decision to proceed to implement the design.

Design review is used to communicate the progress of the design team to management and operating forces to solicit comments and suggestions and obtain "buy-in" from the people who will produce and use the new product or service.

Types of Design Review include:

Conceptual Design Review - the main purpose of this review is to decide which conceptual design to pursue into detailed design.

High-Level Design Review – here, the product/service design is reviewed against the high-level production processes – will the design be “produceable” – fabrication, assembly, process capability are addressed.

Detailed Design Review - the main purpose of this review is to assure that the design meets the goals of the project and the product/service requirements.

Interface Design Review - many products/services are "shoehorned" into an existing system. The objective of this review is to communicate how the new product/service will interface with other products/services, upstream, downstream and supporting the new design.

Final Design Review - this review's purpose is to gain management acceptance (and budgetary approval) of the new design to proceed to the next phase, *Designing Processes to Produce Product/Service.*

Outputs - As a result of these reviews:

- *The new Product/Service is "accepted" by both management and the operating forces,*
- *Interfaces between the new product/service and existing products/services are identified.*

Produce Product/Service - Check Performance

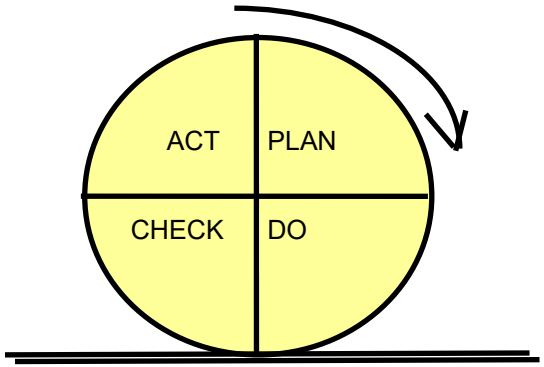
By the end of the DMEDVI methodology, the newly designed/redesigned process has been turned over to the operating forces. "Production" process management (i.e. quality control) controls the quality of the new product or service.

The "Check-Act" cycle must be implemented to obtain knowledge of customer reactions and their judgment of the new product/service's quality. Continual improvement (i.e. through the “usual” improvement methods) now addresses continued process improvement of the new design.

Although not something we wish to contemplate, there is the possibility that the new or redesigned product or service does not achieve its quality goals (remember the Edsel, "New" Coca-Cola, the movie *Waterworld*?) or dramatically exceeds its goals.

In these cases, an analysis of *why* the goals were not met (or were exceeded) should be performed. This analysis should focus on why the design process did not perform as expected. The investigation should result in improvements to this design process.

If the goals are not met, a decision, of course, will be made on whether to *redesign* the new product or service.



2.5 Business Planning Process

Learning Objectives

- Understand the purpose and process associated with setting corporate priorities

Unit Contents

- The Need for a Business Planning Process
- Business Planning Process Elements & Steps
- Business Planning Calendar

2.5.1 The Need for a Business Planning Process

Dr. Deming's often focused on the problem of companies and organizations *sub-optimizing* their efforts. He noted that improvements made to "optimize" one department's performance often hurt the performance of supplier or customer departments. Shigeo Shingo, one of the key developers of the Toyota Production System, states that the overall *system* must be analyzed and improved first, then improvements at local operations can occur. If I decrease the cycle time required to produce a part or assembly at my station, but the result is inventory stacking up in front of your machine, then the *system* has not improved.

The business planning process is thus intended to focus the organization's energies on the vital improvements that will benefit the customer and therefore translate to bottom-line results for the company. Read the following quote from a noted CEO to see the importance he places on this key process:

> *The business planning process is the best tool we have for communicating our strategies and objectives throughout the organization, aligning all activities with these strategies and objectives, and making sure we are on track to achieve them. . . Business planning is our primary vehicle for continuous improvement of those processes that will make us one of the world's premier companies. Our success in this year and beyond depends on our ability to move as one toward our vision. The business planning process gives us that ability.*

There are many different methods that can be employed to develop and implement a strategic plan. One such method is described on the following pages. There are a couple of simple tests that can be applied to see if the business planning process is effective. The first – simply walk around a company and ask a sample of managers and staff if they know the most important problems affecting their company. Then ask how they are contributing to solving those problems.

2.5.2 Business Planning Process Elements & Steps

Six Sigma companies make no distinction between their budget plans, business plans, quality plans, or management plans. Their key business drivers, long- and short-term strategies, key focus areas, action plans, and key performance indicators exist in one plan, a business plan that:

- Sets the course for the company and its business units/geographical units
- Focuses all units and employees on those areas that are most critical to the company
- Aligns all activities with the company's key business drivers
- Promotes continuous improvement of these critical areas

As a Black Belt, you can expect to be leading or facilitating projects that support achievement of the Business Plan. You may be asked to review the logic and data behind improvement efforts conducted in support of the Plan. Management may look to you to perform analyses needed to develop the plan. Finally, you may be assigned a project to improve the company's Business Planning process.

Business planning is a closed-loop process, without beginning or end. For example, considering the business planning steps shown on the following page, creation of the Annual Operating Plan (AOP) may seem like the "start" of the process. The AOP development is simply the start of an annual cycle. The AOP represents the newest version of an ongoing plan that continues to look at least three years ahead. Reviews of the company's performance on last year's plan are one input into this year's AOP; likewise, a periodic market "pulse-taking" feeds the strategic plan and then the AOP. In this way, the current AOP is neither at the beginning or end of the planning process but rather reflects the progress the company has made and the goals they need to reach.

The business planning process is the vehicle the company's Leadership Team, and all business units and geographic units use for setting priorities, allocating resources, reviewing performance, and driving continuous improvement. The four-step process shown on the next page captures the key steps a company must take to move as one toward their vision of being one of the world's premier companies.

The business planning process has four steps:

1. **Strategic Review.** At the corporate, business unit, and geographic unit levels, the company gathers relevant data and information about customers and markets, competitors, strengths/weaknesses/opportunities/threats (SWOT), people,

internal capabilities, supplier and partner capabilities, and performance to the current plan. This information is analyzed to set strategic directions that strengthen our performance and competitive position.

2. **Annual Operating Plan.** The Leadership Team develops an Annual Operating Plan that includes key business drivers, strategies, focus areas, and key performance indicators.
3. **Implementation.** Each business unit cascades the Annual Operating Plan throughout its organization. This includes identifying Business Unit objectives, strategies, action plans, key performance indicators, and benchmarks. It also includes communicating the plan throughout the organization, aligning all activities with the plan, and implementing the actions.
4. **Plan Reviews.** Corporate, business unit, and geographic unit leaders review performance to plan on a regular basis. The reviews focus on factual data and information and include support and assistance where needed.



For Corporate & Business Units Assess:

- Customers & Markets
- Competitive Environment
- SWOT Analysis
- People
- Business Unit Capabilities
- Supplier/Partner Capabilities
- Performance on Current Plan

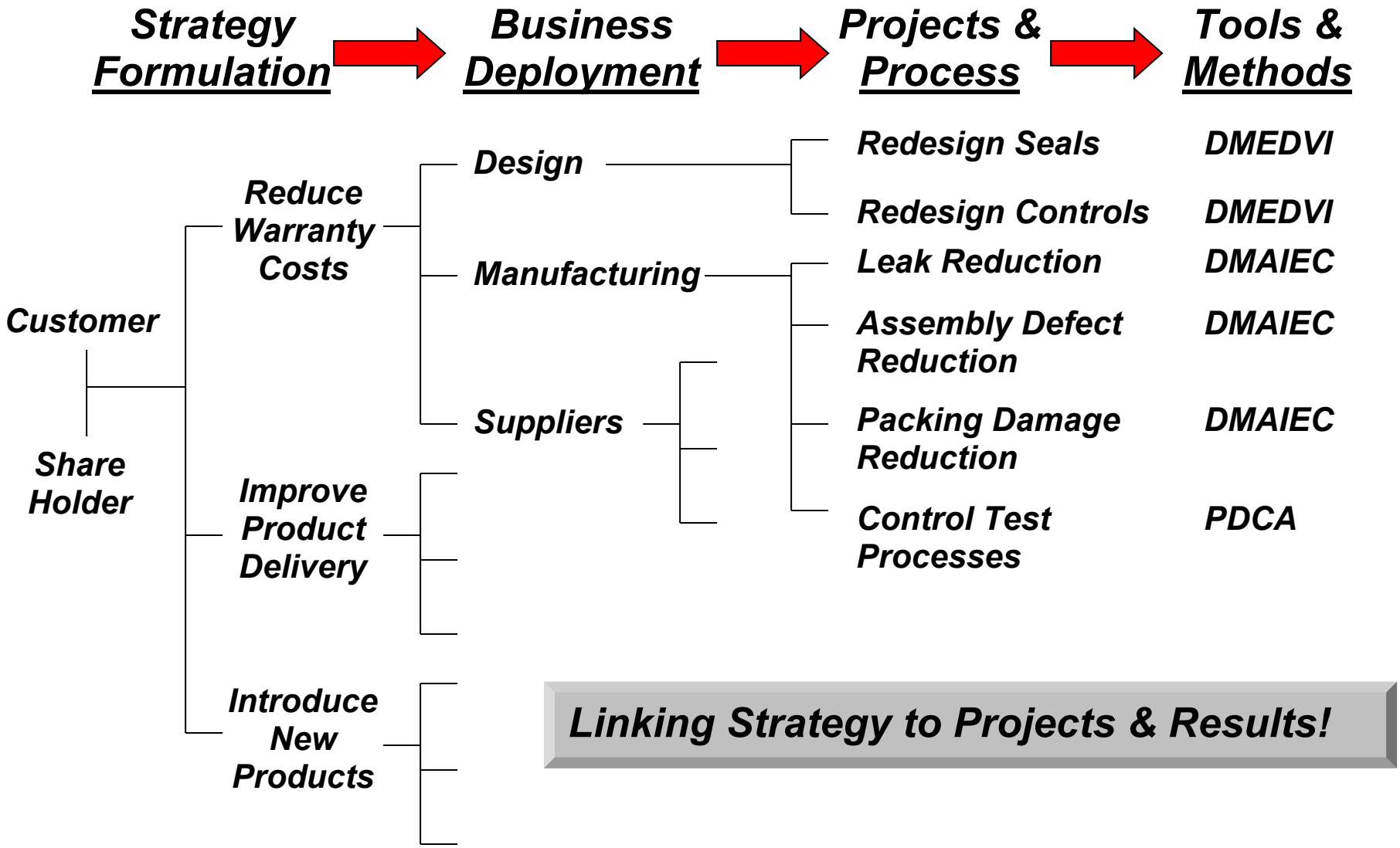
- Analyze Current Performance
- Explain Key Inputs to Plan
- State Key Business Drivers
- Identify Objectives, Strategies, Indicators & Benchmarks
- Develop Budget/Resource Allocation
- Outline Plans for Cascading & Reviewing Plan

Each Business Unit:

- Identify Business Unit Objectives, Strategies, Action Plans, Key Performance Indicators & Benchmarks
- Develop Process for Cascading BU Plan
- Review Plans with Leadership Team
- Communicate Plans Internally
- Implement Plans
- Verify Alignment of Divisions, Departments, Work Teams with BU Plan

- Review Performance on Key Indicators
- Review Performance to Plan
- Update/Revise Forecast and Action Plans
- Communicate Performance to Leadership Team
- Analyze and Improve Review Process

By the time the current year's AOP is developed, the company should have aligned and linked their strategy to specific projects. Black Belts should find that they are leading or facilitating improvement (DMAIEC), design (DMEDVI) or process control (PDCA) projects. The charters for these projects should clearly link the project to the strategy and quantify the contribution of the project to accomplishing some "piece" of the strategy.



2.5.3 Sample Business Planning Calendar

Planning Process	Leadership Team	Finance Department	Quality Department	Business Unit
Step 1: Strategic Reviews	Assess: Customers/markets Competitive environment SWOT People Corporate capabilities Supplier/partner capabilities Performance on current plan	Provide input and support	Provide input and support	Assess: Customers/markets Competitive environment SWOT People BU/GU capabilities Supplier/partner capabilities Performance on current plan
	Letter of Direction from CEO		Distribute Business Planning Handbook	
Step 2: Annual Operating Plan	Announce preliminary strategies and 1999 focus areas	Publish AOP assumptions		Begin analysis of improvement objectives, key performance indicators, targets, benchmarks
		Provide economic forecast		
	Develop 1999 AOP			Complete preliminary business plan (without budget)

Planning Process	Leadership Team	Finance Department	Quality Department	Business Unit
	Refine business plans; analyze cost/benefit of capital projects; develop budget; resolve cross-functional issues			
	Review complete business plans and make adjustments			Complete final business plan (with budget)
	Prepare final AOP and submit to Board of Directors for approval			
	Communicate AOP to Business Units			
Step 3: Implementation	Cascade AOP at corporate level		Black Belt support for projects	Cascade plan throughout unit
Step 4: Plan Reviews	Review performance; update/revise forecast and action plans			Review performance; update/revise forecast and action plans Communicate to Leadership Team

2.6 Six Sigma "Belts"

Learning Objectives

- To understand the role of the Green Belt, Black Belt and Master Black Belt
- To understand the skills required of a Green, Black and Master Black Belt
- To understand how Green, Black and Master Black Belts are certified
- To provide a report-out on your specific Black Belt project

Unit Contents

- Green Belt/Black Belt/Master Black Belt Roles and Responsibilities
- Supporting Roles
- Qualification/Certification Path
- Black Belt Project Description

2.6.1 Introduction

The focus of this section is simple – to describe the roles and responsibilities of Green Belts, Black Belts and Master Black Belts, the skills should they have and how a company could plan to develop these individuals.

2.6.2 Black Belt Roles and Responsibilities

Let's start with the "workhorse" of Six Sigma - the Black Belt. Basically, a Black Belt can be defined as:

> *A full time individual skilled in quality management systems, tools and methods deployed to work on important business problems or opportunities alone or with teams in pursuit of Six Sigma Performance.*

Expanding on this definition:

A full time individual . . . Dr. Juran notes that a company has to "budget" for improvement. Many organizations have attempted to improve by training "hordes" of staff in quality methods and then expecting them to apply these in their daily work. While this model can work, it assumes that the staff will be able to squeeze time from their daily work to participate on improvement efforts. This has proved difficult, especially where organizations have not aligned their reward and recognition systems to achieving improvements. The Black Belt model creates individuals whose entire effort is devoted to improvement. Reward and recognition are thereby aligned with "normal" activities.

. . skilled in quality management systems. . The Black Belt will engage in product and process improvement and design efforts, will "build" and implement process control systems, and will assist management in achieving corporate objectives through Policy Management.

. . tools and methods . . . The Black Belt will be skilled in applying core quality methods. These include methods to understand customer requirements (interviewing, surveying, etc.), methods to understand and analyze processes (flowcharting, statistical methods such as Pareto, Histogram, Control Charts, Sampling and Design of Experiments), and product assurance methods such as Quality Function Deployment, Failure Modes & Effects Analysis, statistical tolerance deployment and reliability testing.

. . deployed to work on important . . problems or opportunities . . Companies typically invest a great deal in the Black Belts and should expect a significant return on investment. One measure of the Black Belt "ROI" is the savings (reduction in expense) or revenue (increased sales) generated because of Black Belt projects. Companies such as General Electric expect Black Belts to complete about 10 projects a year, with an average project benefit of $50,000 for an annual benefit of about $500,000. A defense industry organization calculated that the average Black Belt project would cost about $75,000 (salaries and project expenses – not including investments identified by the project). Based on a benchmark

return on investment of 2 – 1 (their research indicated 2-1 to 4-1 ROIs – they set the "bar" low for their first year of Six Sigma implementation), they set an average project benefit goal of $150,000 per project. Although the initial project assigned to a Black Belt during training may not be a "home run," subsequent projects must be *high impact*.

. . alone or with teams . . The Black Belt will be skilled in leading teams through improvement and design efforts. The BB will manage project schedules and budgets, will facilitate teams in the application of quality tools and methods, and will be effective at implementing changes in the organization. The Black Belt will coach management and individuals to learn and apply quality systems, tools and methods, management systems, tools & methods and is able to apply these individually and within a team structure to effect improvement within the company's culture and organization.

A typical Black Belt job description follows:

Job Title: Black Belt

- *Location*: Assigned Company location
- *Reporting to*: Functional VP/Director in current Division/Staff Function; dotted-line reporting to VP Six Sigma
- *Grade/Job Level:* Determined by Division; with concurrence from VP Six Sigma
- *Job Duration*: Minimum of 18 months - 2 years, full time

Job Description

Job Profile:

- Lead multiple Six Sigma projects per year, each delivering an significant bottom-line improvement
- Lead, train and mentor Green Belts and network with peers in the use of Six Sigma tools and techniques
- Facilitate in the selection of Green Belt projects
- Support Six Sigma training activities, as required
- Carry out other duties and tasks, as requested, by the Functional VP/Director or VP Six Sigma.

Personal Characteristics:

- Complete Black Belt training
- Achieve Black Belt certification as determined by Six Sigma project office – Black Belt certification requires the successful completion of two projects
- Self-starter who can work on own initiative with minimum supervision
- Effective communicator, at all levels
- Able to influence and lead teams; effectively able to work at multiple levels within the organization
- Able to use the full range of Six Sigma tools – e.g., simple brainstorming, detailed statistical analysis of data, use of statistical software (e.g., Minitab)
- Computer-literate and competent in mathematics and elementary statistics
- Ability to lead, train, mentor, and work in a team
- Energy, enthusiasm, with a passion for excellence
- Potential to develop within Company

2.6.3 Green Belt Roles and Responsibilities

Green Belts primarily differ from Black Belts in their being "part-time" dedication to Six Sigma projects, in the amount of training they receive and, in the value of the project they may be assigned.

Some typical figures follow:

Dedication – Green Belts will typically spend about 20% of their time (1 day a week equivalent) working on Six Sigma projects. One company scopes the Green Belt projects so that they can be accomplished with minimal help from others. Another company encourages Green Belts to lead project teams.

Training – A typical Black Belt curriculum occupies about 4 weeks. Green Belt curricula are usually about 2 weeks. Johnson & Johnson has adopted the "bridge" concept. Their Green Belt curriculum is two weeks – DMAIEC and basic improvement tools are covered. After successful completion of a project, the Green Belt may choose to "bridge" to Black Belt. Additional training covering Black Belt topics is then received.

Project Value – As noted above, Green Belt projects have smaller scopes than Black Belt. We had a conversation with a Green Belt around the scope of their project once. The Green Belt was convinced they could accomplish a much larger scope of effort (the GB's manager was also of the opinion that the scope was too large). After some discussion, we finally asked the manager when they wanted the project complete – the answer was about 4 – 6 months. The next question – how much of "Andrea's" time could be spend on the project – about 1 day per week. Some simple multiplication yielded the range of 16 – 24 person-days to complete the effort. The Green Belt finally saw that they were trying to bite off much too much. A typical savings goal for Green Belt projects - $50,000. The Green Belt can then work about 2 – 3 projects per year for a total average benefit of $100 – 150k per year.

The value of a Green Belt program goes beyond the dollar savings associated with the projects on which they work. The Green Belt plays a key role in accelerating the adoption of Six Sigma into the company's "DNA." Instead of a job definition, a typical Green Belt *role* description follows:

Role: Green Belt

- *Location*: Assigned Company location
- *Reporting to*: Functional Manager/Director
- *Grade/Job Level:* Determined by Division
- *Role Duration*: N/A – expected to support and/or lead improvement projects as part of normal job duties

Role Description

Role Profile:

- Lead or participate in one or more Six Sigma projects per year, each delivering an significant bottom-line improvement
- Apply Green Belt skills as part of normal job duties
- Promote the application of Six Sigma within their functional department

Personal Characteristics:

- Complete Green Belt training
- Achieve Green Belt certification as determined by Six Sigma project office – Green Belt certification requires the successful completion of one project
- Self-starter who can work on own initiative with minimum supervision
- Effective communicator
- Able to influence and lead teams; effectively able to work at multiple levels within the organization
- Able to use a defined range of Six Sigma tools – e.g., project management, change management, basic process definition & analysis methods, basic statistical analysis tools
- Computer-literate and competent in mathematics and elementary statistics
- Ability to lead, and work in a team
- Energy, enthusiasm, with a passion for excellence
- Potential to develop within Company

2.6.4 Master Black Belt Roles and Responsibilities

The Master Black Belt plays a key role in the Six Sigma initiative. The Master Black Belt is an advanced Black Belt – expected to do everything that a Black Belt can do. The Master Black Belt will be skilled in more advanced design, statistical and reliability tools and methods, to be applied to "harder" improvement challenges. The Master Black Belt also plays a role teaching and coaching Black Belts and management/staff in quality systems and methods. "Full grown" MBBs also play the quality research and development role, identifying and developing methods to support the company's ongoing improvement efforts.

Six Sigma Implementation – Early in the Six Sigma initiative, a company may hire or contract experienced Master Black Belts to support the planning, launch and implementation of the effort (believe us, there's a *lot* of work to get a Six Sigma effort off the ground!). Developing training materials, management/executive briefings, working to develop the Six Sigma infrastructure (chartering, project tracking, financial benefit assessment, etc.), integrating the Six Sigma projects into the business strategy, coaching initial Black and Green Belt projects are typical early MBB duties.

Black Belt Coaching & Development – The Master Black Belt is typically assigned a group of Black Belts. Often an MBB will be assigned to a specific business unit – here, all the BBs will fall within their responsibility. The MBB will work with the BB and their sponsor to select appropriate Six Sigma project and to ensure that the DMAIEC methodology and associated tools are correctly applied by the Black Belt and their team. If the company has adopted a Black Belt certification program, the MBB will work with the trained BB to develop and implement a plan to accomplish their certification. MBBs will identify and coach Black Belts who are interested and capable of bridging to Master Black Belt.

High Strategic Value Projects – While many Six Sigma projects are run by Black and Green Belts, occasionally projects arise that require more skill/experience. Master Black Belts may be assigned as project managers. For example, a GE Capital business embarked on an SAP software implementation and business process redesign. Four Master Black Belts were assigned to support this very large scope project.

Six Sigma Growth – The Master Black Belts (collectively) monitor the growth and health of the Six Sigma initiative. They will identify weak areas in the organization, analyze their causes and develop/support implementation of process improvements. They will often coach and advise executives on ways and means to improve the Six Sigma initiative.

An example job description for a Master Black Belt follows:

Job Title: Master Black Belt

- Location: Per assigned Division
- Reporting to: VP Six Sigma and/or VP/GM or Staff VP
- Grade/Job Level: Determined by Division
- Job Duration: Minimum of 2 years, full time

Job Profile:

- Support improvement activities at locations, and at suppliers & customers, as required; overseas for short periods
- Provide mentoring and support, as required, to Black Belts, Green Belts – coach local teams to use the improvement tools appropriate to the problem
- Master Six Sigma theory and application; able to train/coach company staff and with customers/suppliers, as required
- Liaison, as required, with external agencies in the delivery of Six Sigma training
- Promote and support improvement activities in all business areas – manufacturing, engineering, services, finance, HR
- Network with other Master Black Belts
- Execute other duties and tasks, as defined by the VP Six Sigma

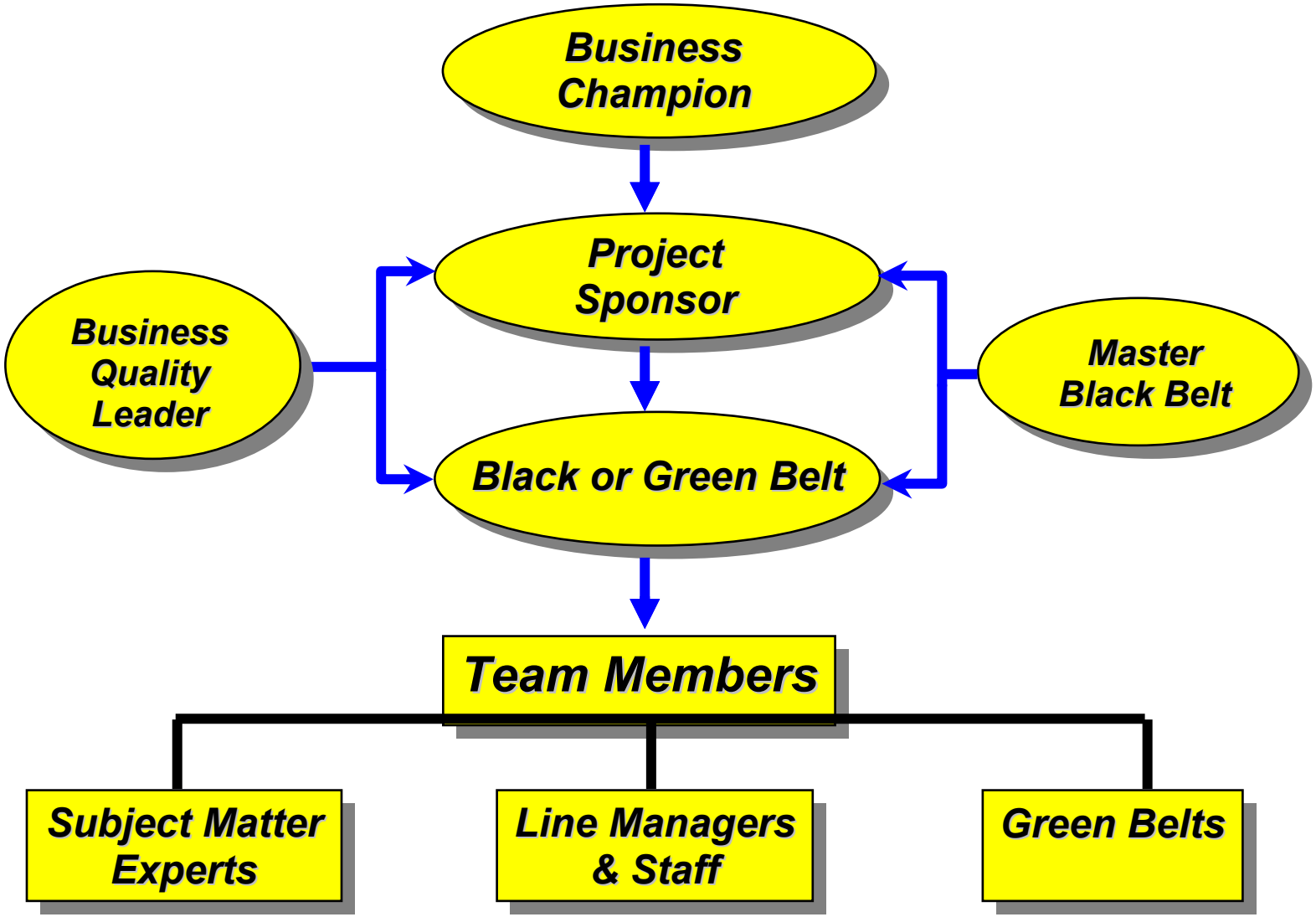
Personal characteristics:

- Certified Six Sigma Black Belt, and Complete Master Black Belt training, or demonstrate completion of a similarly structured program
- Educated to degree level or equivalent
- 2+ years experience or thorough and proven working knowledge of Six Sigma
- Technically strong in mathematics, statistics and use of statistical software (e.g., Minitab)
- Willingness to embrace change and new ideas
- Tough, resilient, and able to persuade others
- Able to work at multiple levels within the organization; politically savvy
- Energy, enthusiasm, with a passion for excellence
- Proactive leadership style; able to communicate at all levels
- Ability to promote the key messages of pace, results and sustainability in all activities
- Able to quickly grasp the bigger picture of Company business drivers and infrastructure
- Ability to build consensus, and work collaboratively as part of the world-wide Six Sigma team
- Ability to travel as required depending on business needs

2.6.5 Supporting Roles

Companies adopting the "Belt" model also assign responsibilities to *Champions, Sponsors* and *Quality Leaders.* Briefly, the *Quality Leaders* are responsible for organizing and "running" the quality organization. The *Quality Leader* is typically a director or vice president. *Project Sponsors* are line managers responsible for specific improvement (or design) efforts. *Business Champions* are senior leaders who charter "strategic" level projects, review progress toward company goals and generally promote the implementation of Six Sigma throughout the business. They may also be given responsibility as *Process Owners* for the performance and improvement of major corporate processes.

Roles and Responsibilities



A Brief Reflection:

Why were you picked to be a Black Belt?

What business unit do you represent?

What is your experience and education?

What is your present position?

What are your personal reasons for being a Black Belt?

Your Current Understanding:

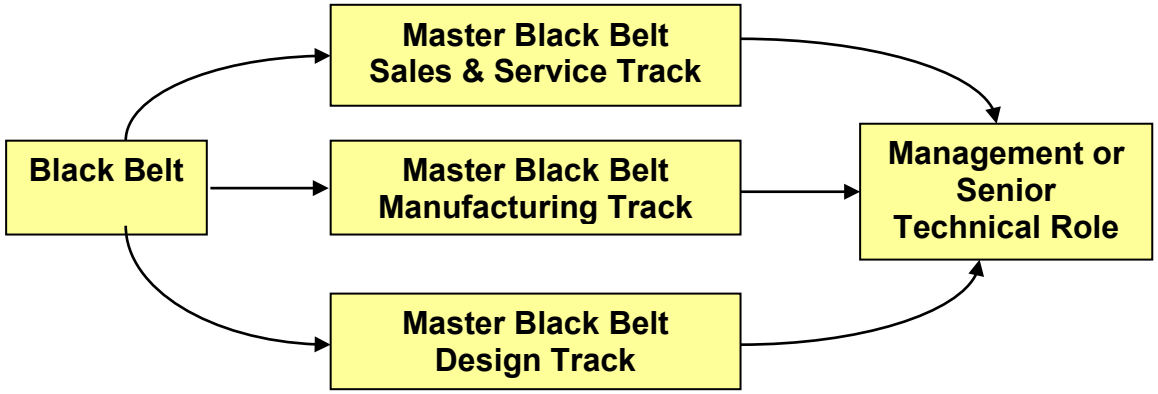
2.6.6 Qualification/Certification Path

The following table depicts the relationship between organizational functions and associated quality management responsibilities. Note that the first three responsibilities are core to all functions. Black Belts will therefore all be qualified in these quality areas. Three Master Black Belts "tracks" are also identified, to provide specialized and advanced skills for individual business functions. Most companies select their Black Belts with an eye toward future promotion. Companies such as General Electric rotate their Black Belts and Master Black Belts back into the business both as part of their professional development and to provide a means of instilling the quality culture into the business.

	Quality Responsibility						
Function	***Product/Process Improvement & Core Tools***	***Process Management***	***Strategic Planning/ Management***	***Customer Research***	***Product Planning***	***Reliability***	***Advanced Statistics***
Marketing	◉	◉	◉	◉	○		○
Design	◉	◉	◉	○	◉	◉	◉
Manufacturing	◉	◉	◉		◉	◉	◉
Sales & Marketing	◉	◉	◉	◉	○		
Service	◉	◉	◉		○	○	○
Supply Chain/ Suppliers	◉	◉	◉	○	◉	◉	◉

A Common Set of Core Competencies with Specialized Knowledge to Support Business Areas

Some companies have also developed specialized "tracks" for Black Belts and Master Black Belts. The model below shows specialization by three different company functions.



Training/Qualification - General

A typical training and qualification path for a Black Belt appears below.

Training Sessions
- 4 - Five Day Sessions
- Training "progresses" through improvement method
- Training will include project reviews and coaching
- Benchmarking visits to other companies will occur, when practical
- Tools practice using "hand" calculations first, then computer application
- Training will include Evening Homework
- Continuing Education will be a part of the process

Black Belts
- Black Belt – Waves 1 – 4 concurrent with project
- Candidate Black Belts arrive at training with an improvement project
- Company will certify candidates as Black Belts at completion of project (see later in this unit for specific criteria)

Master Black Belts
- Master Black Belt – Black Belt plus Wave 5 concurrent with project
- Master Black Belts will have three flavors – Sales & Service, Design, Manufacturing
- Company will certify candidates as Master Black Belts at completion of project (TBD)

The Wave training curriculum appears on the following pages. Some of the topics will be introduced in one Wave and then revisited during succeeding Waves. The matrix shows which Wave will treat which topic.

Wave 1 – Intro to Problem Solving & Identify the Problem

	Wave				
Topic	**1**	**2**	**3**	**4**	**5**
Black Belt Roles and Responsibilities	X				
Qualification/Certification Plan	X				
Improving Existing Products and Services	X	X			
Voice of Customer Feedback	X				
Developing Indicators	X				
Basic Data Collection	X				
Measurement System Analysis	X		X		
Line Graphs	X				
Run Charts	X				
Sampling	X				
Histograms	X				
Process Capability (including Six Sigma)	X		X		
Process Flow Charts	X				
Process Analysis Methods	X				
Bar Charts	X				
Pareto Analysis	X				
Pie Charts	X				
Radar Charts	X				
Cause and Effect Analysis	X	X	X	X	
Project Chartering	X				
Project Reporting and Reviews	X				

Focus on Getting Your Project Started; Performance Measurement, Stratification

Wave 2 – Analyzing & Improving the Process

	Wave				
Topic	**1**	**2**	**3**	**4**	**5**
Minitab Statistical Software		X			
Improving Existing Products and Services	X	X			
Cause & Effect Analysis	X	X	X	X	
Contingency Analysis		X			
Scatter Diagrams		X			
Correlation Analysis		X			
Regression Analysis – Simple, Linear		X			
Probability Distributions		X			
Hypothesis Testing		X			
Parameter Estimation & Confidence Intervals		X			
Sampling		X			
Single Factor Experiments		X			
Reliability Terms and Definitions		X			
Reliability Management		X			
Failure Modes & Effects Analysis		X			
Fault Tree Analysis		X			
Weibull Analysis		X			

Focus on Understanding Cause & Effect, Verifying Root Causes

Wave 3 – Managing & Controlling the Process

	Wave				
Topic	**1**	**2**	**3**	**4**	**5**
Selecting and Implementing Process Changes			X		
Cost-Benefit Analysis			X		
Evaluating the Effects of Changes/ Standardization & Replication			X		
Controlling Processes			X		
Process Management Charts			X		
Control Charts			X		
Process Capability (including Six Sigma)	X		X		
Measurement System Analysis	X		X		

Focus on Process Control to Ensure Consistent Quality Outcomes

Wave 4 – Designing & Delivering Products/Services

	Wave				
Topic	**1**	**2**	**3**	**4**	**5**
Designing New Products and Services				X	
Obtaining Voice of the Customer				X	
Developing Product/Service Requirements – QFD				X	
Creativity Methods				X	
Performance & Process Benchmarking				X	
Pugh Concept Design Selection				X	
Tolerance Development & Analysis				X	
Analysis of Variation (ANOVA)				X	
Design of Experiments				X	
Taguchi Approach to Design				X	
Reliability Testing/Accelerated Testing				X	

Focus on Design of New Products & Services; Advanced Improvement Methods

Wave 5 – Business Planning & Advanced Topics

	Wave				
Topic	**1**	**2**	**3**	**4**	**5**
Business Planning Process (SRP-AOP-LRP)					X
Company-Wide Process Management					X
Operating Reviews					X
Flag Systems					X
Indicator Families					X
Seven Planning Tools					X
Team Types					X
Team & Meeting Management					X
Team Roles & Responsibilities					X
Idea Generation Methods					X
Decision Making Methods					X
Conflict Management & Interventions					X
Change Management Approach					X
Change Management Tools and Methods					X
Facilitating Improvement					X

Focus on Business Planning; Link to Quality Projects and Processes, Working With Teams

Black Belt Certification

Objective: To certify Black Belts by assuring their capability to use the methods and tools taught in the Black Belt curriculum. Certification as a Black Belt at requires the following:

- Attendance at all 5 waves of the Black Belt training.
- Completion of all in class work and homework assigned during the classes.
- The completion of an improvement project, assigned by a sponsor that results in significant, **proven**, improvement. The results must be reported in financial terms.
- The completion of a process control system that is implemented and used to manage a process.
- Application of the methods and tools in actual projects and using actual data gathered from the projects. (See methods and tools certification requirements checklist)
- Sign-off by the mentor for each of the requirements
- Presentation of the improvement project and process control system to the sponsor and a panel of mentors.

*The certification of the Black Belt may extend past the completion of the training classes.

Certification Criteria

To be certified, Black Belts must demonstrate the proper application of the methods and tools taught in the training. Below is a list of the methods and tools and the demonstration requirement that must be signed off by the candidate's mentor. In cases where application of the tool or method is not appropriate the mentor may use example data or problems as part of the certification.

Method or Tool	Certification Requirement
Improvement Process (DMAIIC)	Candidates are required to utilize the improvement process to complete a project. The project must result in significant improvement, reported in financial terms and meet the improvement story checkpoints.
Process control system	Candidates are required to develop a process control system that is implemented and used to control and monitor a process.
Design Method (Optional)	Candidates may complete a design project to obtain a special certification for design.
Customer Survey or interview	Candidates are required to determine customer needs through the use of an interview or survey.
Pareto Chart	Candidates are required to use Pareto charts to focus improvement efforts on significant problems.
Histogram	Candidates are required to plot a histogram of actual data they obtained in their job function.
Process capability calculation	Candidates are required to calculate process capability for a process in their job function.
Cause & Effect Diagram	Candidates are required to develop a cause-and-effect diagram to determine potential root causes in their project.
Control charts	Candidates are required to use an appropriate control chart to evaluate process performance and interpret stability.
Hypothesis test • Test of means • Test of proportions • Test of variability	Candidates are required to use hypothesis tests to determine significant differences between data sets within their project for means, proportions or variability.
Gantt chart	Candidates are required to develop a Gantt chart for their project.
Cost Benefit analysis	Candidates are required to calculate the cost benefit of countermeasures in their project and the cost impact of their improvements.
Sampling and data	Candidates must devise a data collection and sampling plan for their project. The plan must

Method or Tool	Certification Requirement
collection.	include the confidence level of their sample and the sampling technique used to limit bias.
Quality Function Deployment	Candidates may use QFD to determine specific requirements for a special design certification.
Deployment Flowchart	Candidates are required to develop a deployment flow chart for an actual process.
Scatter Plots	Candidates are required to graph a scatter plot to identify possible correlation between two factors.
Correlation coefficient	Candidates are required to calculate the correlation coefficient for two factors that appear correlated.
Regression analysis	Candidates are required to calculate the regression equation for correlated variables.
Failure modes and effects analysis. (FMEA)	Candidates are required to develop an FMEA for a product or service.
Design of Experiments	Candidates may perform a DOE on data from their job function for a special certification.
Weibull Analysis	Candidates may perform a Weibull analysis on data from their job function for a special certification.
Functional Analysis	Candidates may perform a functional analysis to obtain a special design certification.
Fault Tree Analysis	Candidates are required to perform a fault tree analysis on failure that has occurred in their job function.
Affinity Diagram	Candidates are required to construct an affinity diagram to group unstructured inputs.
Relations diagram	Candidates are required to draw a relations diagram to determine leverage points of a group of actions on an objective.
Matrices	Candidates are required to use matrices to show relationship of projects to priorities.
ANOVA	Candidates are required to perform an ANOVA to detect factor differences.

Additional Self-Study:

The following texts form a "starter" library to supplement this text.

Leadership/Organization:

- The Prince - Machiavelli
- The Art of War – Sun-Tzu
- War as I Knew It – Patton
- Success is a Choice – Pitino
- Hope is Not a Method – Sullivan
- Jack Welch Speaks - Lowe
- The Spirit of St. Louis – Charles Lindbergh
- Rocket Boys - Hickam
- Managing Transitions - Bridges

Quality Systems:

- Introduction to Quality Control - Ishikawa
- Out-of-the-Crisis – Deming
- Total Quality Control – Feigenbaum
- Kaizen – Imai

Additional Technical:

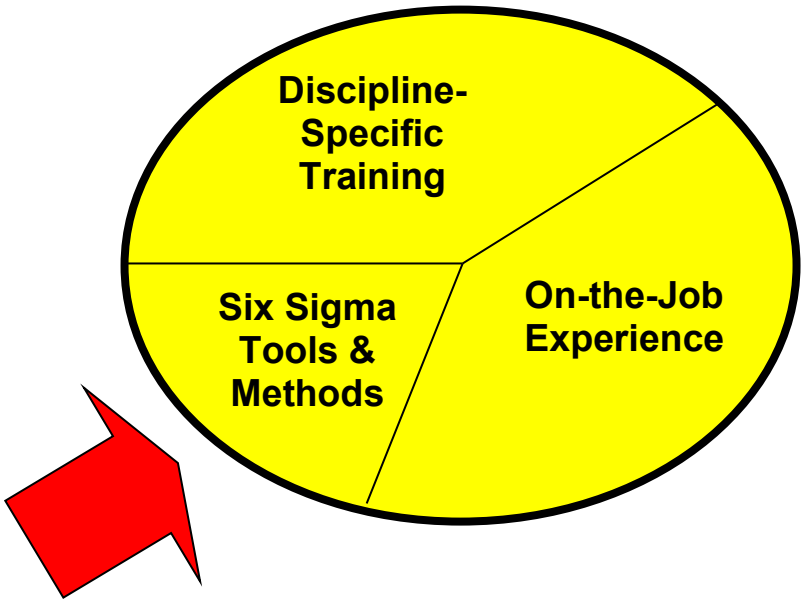
- Economic Control of Quality of Manufactured Product – Shewhart
- Tolerance Design - Crevelling
- Quality Function Deployment - Cohen
- Software Quality - Jones
- Total Productive Maintenance - Nakajima
- Applied Life Data Analysis – Nelson
- Corporate Financial Analysis - Harrington
- Creating Innovative Products Using Total Design – Pugh
- Design for Manufacturing and Assembly – Boothroyd, Dewhurst, Knight
- Design for Six Sigma - Crevelling
- Design and Management of Service Processes – Ramaswamy

The Black Belt Transformation:

"Regular" Skill Sets:



BB/MBB Skill Sets:



A High Bar for You!

Cecila Shallenberger is a Black Belt at TRW Systems, a defense and government contractor. We received this email from her:

"I trust all is well with you. All is well here.

I realize Six Sigma is all about teamwork, but just this once I thought I'd brag just a little. I certainly do realize that the following could not have been accomplished without the help of many, many people.

I closed 10 charters last year, total benefit $11.84M (although the books show a conservative $10.6M). 62 GBs were trained, and 44 of those were certified. So far this year, I've got 5 charters in process, and 5 being drafted as we speak. 3 of those include the customer as GBs on the teams.

I have about 32 people left on the TCS contract to get onto charters and through training. GBs are kicking down my door wanting to lead charters.

This year, my focus is on creating a lead GB community with 3 BBs being mentored - they will lead the charters, I will lead them. As you can see, we're doing well. I'm now going after my BB certification through ASQ to compliment the TRW cert.

My thanks go to you for your inspiration. You've created a monster!!!

PS We won 100% of our award fee, . . . "

2.6.7 Black Belt Project Description

Describe Your Project to the Class:

- What is the Product or Process?

- What is the Problem? Why is it a Problem? How Long Has It Been a Problem?

- What Data Do You Have to Support the Problem?
    - Quality, Cost, Warranty, Delays, etc.

- Who "Owns" the Product/Process?

- Who Will Sponsor Your Project?

- Who Will Help You With Your Project?

- When Should The Project Be Complete?

3.0 Team Facilitation & Management

Unit	Description	Page
3.1	Working With Teams	3.1 - 1
3.2	Idea Generation & Decision Making	3.2 - 1
3.3	Exercises	3.3 - 1

3.1 Working With Teams

Learning Objectives

- Understand the types of Teams that may be used in improvement projects
- Be able to organize a team
- Be able to develop a project charter
- Be able to develop a project plan
- Be able to plan and conduct team meetings
- Be able to plan and conduct team reviews
- Be able to recognize and address team conflict
- Be able to close a team

Unit Contents

- Teams
- Team Processes

3.1.1 Teams

The columnist George Will, once commented, "*Football combines the two worst features of American life. It is violence punctuated by committee meetings."*

You've probably been on at least one committee that just doesn't seem to be much more than a forum for gossip, whining, psychoanalyzing each other, or reading the minutes from the last committee meeting.

Worse are those committees "chartered" to make some decision when it turns out that the decision has already been made by the "higher-ups."

Even though we've seen the "worst" of teams, we've also been on some pretty darn good teams. We've been fortunate to have been part of teams that have been given a *mission,* the *responsibility, authority* and *resources* to get it done, and have "crossed the goal line," successfully completing the mission.

When the *right organizational conditions* are in place, teams can be a very good way of making improvements *happen.* There's nothing better than the feeling of having accomplished something with a group of strangers who have turned into your friends.

This section will describe our "philosophy" of teams and provide you with some methods that can help your team experiences be positive *and* productive.

Team Philosophy

Let's establish one thing right away. A quality improvement team is a *method* of accomplishing something. Teams don't exist just for fun.

In fact, there really is *no such thing* as a team. There is, however, a group of people that have come together for a purpose, hopefully a *common* one. We always try to remember that it's the *people* who are going to accomplish the goal. How can we balance their needs with the needs of the improvement effort?

Teams also *consume resources.* Meeting time, data collection and analysis work all take time. Therefore, without sounding too much like a bean counter, we should expect some *return on this investment.*

You should consider the *efficiency* of your teams, just as you would any other *production process.* If your teams are taking a year or more to solve simple problems, then something is rotten in Denmark! What's going on? Teams and the improvement process can and should be improved, just like any other.

Having made the preceding nasty, business-like statements, we'll back off a bit. If you are just starting to use teams for quality improvement in your organization, we'll "allow" for a growth curve. It takes people a while to get used to working together, to practicing the steps of quality improvement, to using statistical tools and improvement methods.

Florida Power and Light started the teams' program in the early 1980's. It took us about six or seven years to get to where our teams were "efficiently" producing quality improvements. Along the way, much improvement occurred, and when we compared the early QI Stories to those of the late 1980's, it was like Stone Age to Space Age.

Miliken, the textile company, shared a similar experience. One of their managers told us that early on, they had a very formal "chartering" process for teams, requiring approvals of management and help from designated "facilitators." After a few years, though, if a problem arose that needed a team, it became a "natural" action to gather a group of staff that could address the problem and "*just do it.*"

Before Starting a Team

Carefully consider the need for a team before you start one. From a management perspective, here are some criteria we've applied for this decision:

Problem known, but cause of problem is unknown - This *used to* be our toughest reason to start a team. As a manager, it was an admission that we didn't know the answer to a problem. These days, though, we enjoy it when we can say, "I don't know." It means that there's an opportunity to learn something new.

Time restraints and resource requirements - This is an obvious one. We need to accomplish something, and one person can't do it, or there are a variety of skills necessary for the project to succeed.

Need to leverage problems - One of our heroes is a nurse who runs an "AM Admit" unit at a hospital. She is always working on making the unit a better place for her patients and staff. But she's only got so much time in the day. She could accelerate the unit's improvement journey if she could get her staff involved in identifying and making improvements (see the lab director's strategy, for example, in *Types of Teams*).

Need to solve cross-departmental quality problems - Teams are practically the only effective way of addressing quality issues that cross department boundaries.

Philosophy of delegating authority - Many organizations have found success when authority is delegated to the lowest level possible. Modern quality management incorporates this delegation "philosophy." Teams are simply one means of practicing this philosophy. The idea of delegation, though, can sometimes be a difficult balance for management. For example, although the manager has delegated authority to a team, he or she still retains *responsibility* for the outcome of the team's work. The manager cannot approve of a solution that he/she knows will have a negative impact on the organization. On the other hand, managers should be flexible and learn to accept solutions that can work but are not necessarily the same as they might have chosen. Speaking from experience, although we were sometimes skeptical of our teams' solutions, we were "forced" to learn that their solutions often worked better than our ideas.

Want to help staff develop technically and personally - One oft-neglected responsibility of management is to develop and mentor their people. Teams can be one way of accomplishing this. The team leader, for instance, will develop management and leadership skills. Critical thinking skills develop. Project management skills develop. The staff comes

to work with their heads engaged, instead of just "their hands and feet." One of our greatest pleasures is to see someone stretch and develop beyond their current capabilities.

On the other hand, here are some reasons *to not form a team:*

Have a solution that you want proved - On our very first team, the manager handed us a solution and then told us to prove that this was *the* solution. We didn't even know what the problem was!

Want a team to "rubber-stamp" a decision - In organizations where teams are popular, some managers may think that having a team reach their conclusion will lend credence to the decision or action.

When you don't intend to act on the problem - We worked at a nuclear plant that had developed a wonderful, but short-term solution to plant safety issues. Whenever the regulator (the NRC) came around, management could always point to a team that was working on a particular problem. After a while, though, the regulator began to wonder where the *products* of all this team activity were. Management had used the teams as a delaying tactic, but it only worked for a while.

Like to have consensus on all decisions, think a team will achieve this - We worked for a manager once who seemed to think that *consensus* was necessary on all things. Teams weren't formed to *analyze* a problem, but as the manager's way of *forcing consensus* on his decision. He'd just keep talking about the issue until we gave up and "agreed with" his conclusion.

Can't make a decision yourself - If you have a hard time making critical management decisions, don't think that a team will help.

Like to get a team started, then confused, then "save them." - Some managers like to play the role of the "cavalry." They will start a team, give them poor direction or a fuzzy objective to begin with, watch the team flounder around[1] for a while, and then come in and "save" the team. We suppose their ego gets a boost, but . . .

Organization says you must have a certain number of teams - Early in many organizations' quality journeys, they will measure the number of teams doing something in the organization. Of course, this puts pressure on managers to have teams, regardless of whether they need them or not.

[1]One of our friends has a strange expression for this: *"The team was flopping around like a dead mackerel!"*

Beginning Team Activities

Do you remember the TV program *Mission: Impossible?* The team always got their mission through a tape-recorded message: "*Your mission, should you choose to accept it, is. . . .* " Of course, they always accepted the mission. There wouldn't be a show if they threw the tape recorder away and said, "*Nah, let's go to the beach today!"*

Beginnings are a delicate time for teams. Let's examine some of the issues you'll have to address. We're going to suggest some general pointers, but you decide for your organization and specific team what you think will work best. There's no *right* answer, only suggestions for what we've seen work and not work in the past:

Before the First Meeting

Who will lead the Team? - It's easy to say*, "Get your* best *person to lead the team*," but what does that mean? Our best definition is "*somebody who can herd cats."* Leadership is an art. We've seen a variety of leadership styles, we've seen people of whom little was expected achieve great things, we've seen "hot runners"[2] miss the target. Some general characteristics we look for: *is respected by, and respects others, can focus on the mission, flexible, can handle the "rough and tumble" of the team, sense of humor.*

Sometimes, it makes sense to think first about who will be on the team, and then ask the question, "*Who might be capable of leading this team?"* One of our poorer choices was to put a young secretary in charge of an administrative team. Although she was capable, an older, dominating member made life very difficult for the leader. We really didn't "engineer success" into her leadership opportunity.

We'll give you the other side of this coin. Pick the leader, explain the mission and then trust them to decide who needs to participate.

Who will participate? - Probably the most important criterion is to identify people who know something about the problem being addressed. We've been on teams where it's been difficult for us to participate simply because of our lack of technical knowledge. This is hard for a leader, but don't just pick your "buddies" to be on the team. One NASA manager learned to like having at least one "left fielder" on his teams. These were the people who challenged everything, who kept the team away from "groupthink" and who sometimes came up with the off-the-wall ideas that proved fruitful.

[2]Submariners' term for a live torpedo. Often applied to hot shot young officers.

Try to keep the team as small as possible. Although we've seen teams of 20 people who were successful, the old saying, "*More than three people can't figure out where to go to lunch!"* is often true.

If you can, it's always nice to ask someone if they want to participate in the improvement effort, rather than sending a memo or e-mail message telling him or her to be in such-and-such a meeting on a certain date.

"Chartering" – When a company begins to promote the use of teams, a "formal" chartering process is usually needed for almost all teams. There may be a need to identify which teams are out there, what they're doing, and what progress they're making. As the teams' process matures, we recommend that you consider segmenting which teams require a "formal" charter. For day-to-day improvements requiring more than one person, we recommend that you move as soon as possible to the Miliken model described in *Team Philosophy.* For large scope or multi-team projects, charters are essential. See later in this section for more on chartering and a charter template.

How much time is required and who's going to do the "real work" - Dr. Juran talks about the need to "budget for improvement." For many organizations, there is no "slack time" provided for improvement. Companies where work occurs in shifts or on assembly lines, or where customer needs cannot be interrupted for any significant time (healthcare falls into this category) are challenged to make time for improvement work.

Senior management needs to consider how this issue will be addressed in their organization. Some companies have decided that the work of improvement is important enough to pay overtime to shift employees; physicians have been compensated for time spent on improvement efforts for the hospital, etc. The solutions are generally not complicated; they just need to be identified and implemented.

The First Few Meetings

How do you break the ice? - While we're not trying to mix alcohol and quality improvement, one of the best icebreaking meetings we attended was at the lounge of the *Holiday Inn* in Homestead, Florida. Sure, the team had plenty of subsequent battles, but it was a good way to start. On the other hand, in one of the worst icebreakers we experienced, the team leader asked us all how we *felt* about the issue and then had us play some cute, "getting to know you" exercise. We'll also never forget facilitating a kick-off meeting for a project, getting an hour into the meeting and suddenly realizing that most people in the room *didn't know why they were there!*

Figure out some way to break the ice that's appropriate for your organization and its current *culture.*

Gaining commitment to the mission - Make sure this is clear before you go charging off after the windmills. Oftentimes, management will hand your team a broad charter. When you realize this, go back and get some clarity about what they are *really* looking for.

One team was charged by management to *reduce supply expenses.* Well, it turns out, they didn't have much control over a major piece of supply expense, since most of the supply contracts were negotiated through the corporate office and they were part of one division. They did have control over utilization of supplies in their local area and could make improvements in this aspect of supply expense.

Jim Walden, former VP of Power Resources at Florida Power & Light, had a favorite phrase: "*The GOTTAWANNA."* Management and the team leader need to consider how to motivate the people on the team to tackle their problems. How did Moses get the Israelites to wander in the desert for 40 years?

What's in it for me? - This issue is a very careful balancing act and relates closely to the *commitment* issue above. Recent research on motivation has shown that the American worker on a team wants the project to succeed but also wants to *shine* individually as a result. This is contrary to management's typical expectation that the project succeeds to achieve corporate success (this motivation style is often applied in Japanese companies with success in *their culture*).

Here's a contrast: One company designed a quality tools handbook where each team member's signature appeared on the inside cover. Another company produced a similar handbook where several professional quality "gurus" were thanked for their wonderful influence, but the forward was "signed" by the "Staff of XYZ, Inc." Which approach do you prefer?

How will the team work together? - Early on, the team should decide how they want to work together. Issues such as meetings (should we?, how often, when, where, how long, etc., also, see *Team Meeting Process*), work assignments, confidentiality, etc. can be captured in a set of ground rules adopted by the team.

"Penalties" for breaking the ground rules are encouraged. Many teams have a consequence for showing up late at the meeting - a dollar or more in the team "shot-pot," or having to bring the "munchies" for the next meeting are common. Our favorite ground rule is the one on celebration of accomplishments or milestones.

It's the little things that get you. Often, teams are crammed into a long, thin conference room (we worked at power plant sites were the offices were single-wide trailers). It's hard to have a sense of cohesiveness when people are stretched out

on both sides of a long table.[3] Try to get a room where a round table or at least a square arrangement of tables can be achieved. Everybody's facing each other and side conversations tend to be limited under these conditions.

How much time is required? **-** Make sure everybody recognizes and accepts their commitment. Even more important, make sure their management is aware of the commitment. If somebody can't support the time commitment, they may still be able to contribute as a "guest" (see *Team Organization*).

What support will the team need? - Many organizations provide some sort of support structure for their improvement teams. The *facilitator* (see *Team Organization*) is a popular method of providing the team with guidance through their first improvement project. The facilitator may offer advice on tools, methods, or team "dynamics" issues.

Additional support may include information from the data processing department, laboratory support, vendor support and others. Despite what Phil Crosby preaches, quality is *not* free!

Planning the project - One VP of Quality could see the difference in productivity between teams who did not plan their projects and those who did (see *Project Planning*). His insightful comment: "*The teams that did a project plan got results, those that did not floundered."*

"Outside" Issues

Those "left out" - What about those people in the department who are not on the team? Dr. Kaoru Ishikawa used to comment, "*The whole department is on the team; there are just some who go to the meetings."* There are several strategies to address this.

The Storyboard posted in a public area can serve as a *two-way* communication vehicle. From the team to the rest of the department, it communicates progress made on the problem. From the department to the team, "sticky notes" can be left on the storyboard with suggestions or ideas.

During periodic department meetings, the team can spend a few minutes (again, the Storyboard is helpful, here) presenting the progress they're making, and "alligators" they're wrestling with.

[3]Have you ever noticed that most boardroom meeting tables are long and narrow?

A pharmacy director from a North Florida hospital once made a presentation that brought a few tears of "quality joy" to our eyes. This woman did not seem like a strong advocate of quality improvement, but her story was simple and inspiring. Her hospital had started a quality improvement effort, and she was "volunteered" to be on the first team. She described how their team "fussed through" an improvement in filling Crash Carts,[4] leading to a breakthrough in performance for this process.

Then she "got to figurin'" how she could apply this in her department. She asked her staff what they thought were the key problems in the department and they picked one. She put a team together to address this issue. Well, the rest of the department got to see how much fun this team was having. One day, a pharmacist came up to her and asked her if she could start a team on another problem. Within about six months, the director had become the hospital's improvement champion!

She concluded her presentation with the comment that "it was sure easier to get to sleep at night, since she wasn't the only one worrying about these problems anymore!"

Communication – The team should develop a communication plan. The scope of this plan will depend on factors such as the scope of the project, who the "stakeholders" are, how much of the organization will be affected by this project, etc.

We were fortunate to lead an improvement effort where one of the team members taught us about "salesmanship." Since our project was not well understood by most of the organization, we scheduled as many meetings as we could to explain what we were doing and what progress we were making (the potential to impact the organization was large!). By the time our recommendations came out, everybody was comfortable with our work, and the recommendations were accepted without any objections. *Salesmanship!*

Reviews - This message is to both management and the team. Make sure that the project is reviewed, even informally, every so often (if more than a month goes by and nobody asks you about the project, that's a warning sign!). Use the *QI Story Review Form*, in Appendix C - Forms & Templates.

Reviews should be a simple presentation of the project (that ol' Storyboard helps here, too) followed by a question-and-answer session to clarify issues. Action items should be recorded at the review and followed up as soon as possible.

[4]"Crash Carts" are filled with equipment and supplies needed for emergency patient care (i.e. when a patient "codes" on a nursing unit).

Team Types

Four basic kinds of *improvement* teams are observed "in nature." The matrix summarizes these:

Project	*Departmental*	*Cross-Departmental*
Picked by Team	**A**	**C**
Assigned to Team	**B**	**D**

A - *Departmental Team, Picks own Project* - One of our good friends was a laboratory director at a hospital in Georgia. He maintains a "Wish Board" in the lab's break area. On this board, people pin small slips of paper that start with "*I wish we could (fill in the blank)."* Most of these are minor process problems that get in the way of the staff doing their best work. He encourages his laboratory staff to organize small (two or three people) teams, pick any one of the *"I wish"* statements and work them through to completion. There are always four or five small teams at work on the "*I wish"* projects.

What our friend has created is a very simple, yet effective *quality circle* program in his lab. The people on these teams are all from one *department* or *function,* and they *pick their own projects.* Now George has several purposes for this program. First, his people learn how to work together. Second, they learn how to *practice* process improvement. Third, they solve local departmental problems.

The order of these purposes is important. For quality circle teams, the first two are most important. *Education* and *practice* are the key words here. In some organizations, these are "standing" teams, continuing from project to project. For each project, a different leader may be picked, again, to develop leadership skills in the group. People in the department may rotate on and off the team, depending on the problem being tackled.

B - *Departmental Team, Assigned a Project* - Because the improvement is important to the *department* or to the *organization* (but involves just the one department), management has assigned a group of people to work on the effort. This is a common application of quality improvement teams.

Much of the remaining discussion in this section will pertain to these and type "D" teams.

C - *Cross Departmental, Picks Own Project* - Typically, this kind of team will be a group of managers working on *cross-functional* issues that affect the organization. These teams can be a very important means of improving the organization's overall *quality assurance system,* as well as addressing important quality issues.

For example, an Engineering department formed the *Nuclear Cross-Functional Team*, to identify and address issues that were common to all the engineering disciplines. This was a group of supervisors, who self-selected their projects and reported progress periodically to the engineering management. They accomplished some major improvements in the department, such as standardizing the format of engineering "packages" developed by the department for nuclear plant modifications.

In one hospital, two groups of managers formed *service line teams.* These cross-functional teams addressed quality, cost and service issues for the *cardiac* and *perinatal* service lines. Often, this kind of team will "spin-off," or charter improvement projects that support the major themes they are addressing.

D - *Cross Departmental, Assigned a Project* - Improvements worked on by these teams will cross department boundaries, and are often the most important affecting the organization. These teams are often chartered as part of the *strategic improvement process*, which focuses the organization's resources on the highest priority improvement needs.

Team Organization

Some basic team roles are defined below:

Team Leader - The Leader of the team coordinates and directs the work of the team as it moves through the quality improvement effort.

Team Member - The team members share responsibility for the work of the team (both inside and outside team meetings) during the project. Some specific team member duties *during* team meetings are described below:

Guests - People may be invited from time to time to participate in one meeting, or to work with team members for a short time during the project. Physicians, for instance, who may not be able to participate in the entire project, can be involved in clinical quality improvement efforts as guests. When the project is completed, make sure these people receive the recognition they deserve.

Facilitator - The facilitator is a team advisor or consultant who has expertise in the improvement process, tools and methods. The facilitator supports the team leader in planning next steps and in providing support and feedback to the team's effort.

Recorder - The recorder is a rotated meeting role assigned to help keep the record of the team's work. The recorder logs significant meeting content on a flip chart in front of the team.

Timekeeper - The timekeeper is a rotated meeting role assigned to help the team manage time. The timekeeper informs the team of when agenda items have run their "budgeted" time and may also call out the time remaining for agenda items.



Charters - Communicating the Team's Mission

The core elements of a charter include:

- **Objective**- A statement that describes what the team is being asked to accomplish.
- **Indicators** and targets for improvement
- **Impact**- benefits to the company of the improvement.
- **Process and Boundaries**- The beginning and ending points of the process to be improved.
- **Limitations**- Specified constraints of the project. Deadlines, budget, regulations, paybacks, etc.
- **Key Assumptions** - Assumptions of the sponsors as to the outcome of the project; includes the deliverables.
- **Resources** - The people, equipment, space, support, data, etc. that projects expects to use.
- **Milestones** – key points in the project with time goals. The DMAIEC steps can be used as milestones.

Other items may be added to the charter as the situation warrants. Some teams have included the business reasons for doing the project and the customers whose needs are to be met.

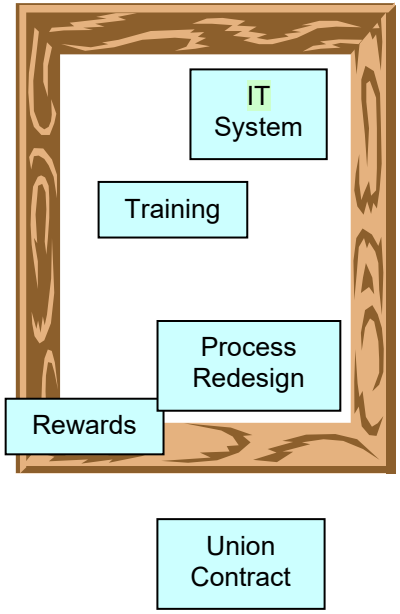
Project Scope

Defining Scope

Have you ever been on a "boil the ocean" or "fix world hunger" project? These can be some of the most frustrating experiences for a team. Defining the scope of the project up-front is critical to success. "Having the conversation" with your champion regarding the scope is one good way to start the definition process.

The Framing Tool

One group technique that can be used to define scope is the *Framing Tool.* The process is described below:



1. On a flipchart or whiteboard, draw a picture frame (or a box if you are not artistically inclined!).

2. Brainstorm possible deliverables or elements of the project. Record these on sticky notes.

3. One-by-one, decide whether the deliverable/element is in or out of scope. Place the in-scope sticky notes inside the frame and the out-of-scope notes outside the frame. If you can't decide, place the note *on-the frame* for now.

4. Review the results. You should have clarified what is clearly in and out of scope (this often occurs quickly). You can now discuss the "on-the-frame" ideas; this may be a good time to bring your champion/sponsor in to review the results and help you move the "on-the-frame" ideas in or out of scope.

Multi-Phased Efforts

Another technique for managing scope is to divide the project into phases. Phased implementation allows the organization to begin to see project benefits quicker and to make decisions regarding the need to proceed to later phases. See also Multi-Generational Planning for design projects (Unit 10.1).

Business Case

Dr. Juran notes that while workers usually speak in the "technical" language of the company, senior management speaks in financial terms. Middle management and team leaders (i.e. Black Belts) then need to be "bi-lingual." As Black Belts, we "know" that reducing defects and failures will save money for the company; the business case challenges us to put a number to these savings (projected early in the project, verified by project close).

One company developed the following statement regarding the business case for Six Sigma projects:

- All projects, direct and indirect, strategic, hard and soft, DMAIEC and DFSS, engineering process, accounting transactions, etc., must create financial benefits now or in the future
- Target benefit for each project of $150K or $1M in cash flow improvement/year.
  - This achieves industry standard of 2-4:1 savings
  - Recovers training and program office costs
- Black Belts, Champions, and Money Belts have key roles in estimating and validating project benefits.

As noted above, the main objective for quantifying the project's potential benefit is to demonstrate to project sponsors and other stakeholders that the project is worth the attention of a Black Belt or Green Belt's effort. Typical "opportunity" numbers are $100-200K financial benefit per Black Belt project and $50-75K benefit per Green Belt project.

Developing a Business Case/Opportunity

1. From your charter, identify the "**pain**" associated with the current performance of the process (e.g. manufacturing or software defects, inability to bill a customer, lengthy cycle time/excessive resources required for the process, errors that must be corrected, vendor/subcontractor "excessive" costs, etc.).

2. What is the **cost** of the "unit pain?" For example, what does it cost to fix a defect (e.g. it may only be a few dollars if the defect is found in the plant; several thousand if found in the field)?

3. How many times does the "pain" occur? If we did nothing, how often will the pain occur (over the next year, or if program-related, how many times will the process be "executed" next year)?

4. What is the **cost** of the overall, predicted "pain" if nothing were done to improve the process? How much do you think we can reduce the "pain?" Note that this latter question will generally require some judgment – e.g. I think we should be able to cut the defect rate in half.

5. Again, although it's generally too early to accurately tell, what will it take to reduce the pain (ballpark **cost** of possible process changes)?

6. Develop the **Potential Benefit** for the project.

7. As new information is developed, refine the potential benefit. For example, following a Pareto analysis of the defects, your team decides to focus on the defect category contributing to 60% of the problem. Refine the potential benefit based on a) your decision to focus on the one defect category and b) your estimate of how much of the 60% associated with that category can be eliminated.

Assumptions & Constraints

Often, you will have to make some assumptions to support the opportunity quantification. Your company will also place constraints on the benefits you can claim. Typical values appear below; these should be validated with your financial staff/office:

- The cost of money is 8% per annum
- A burdened labor hour is $150.
- The opportunity should be time-bound to a one/two-year maximum

Financial Spreadsheet

The opportunity may be spread or divided among several categories. A typical spreadsheet used to allocate the opportunity appears below with associated category definitions.

	Annualized Impact									
			Gross Benefit		Costs Required		Net Benefit			
Line Item Detail	Potential Net Benefit	Sales Increase	Hard	Soft	Hard	Soft	Hard	Soft	PBT	Headcount Reduction
Direct										
Direct Labor/RPE										
Direct ODC										
Direct Material										
Direct Subk										
Indirect										
Indirect Labor/RPE										
Space Costs										
All Other Costs										
G&A										
Indirect Labor/RPE										
Space Costs										
All Other Costs										
Total Expense	0	0	0	0	0	0	0	0	0	0
Capital										
Working Capital/Receivables										
Other Assets										
Total Assets	0		0	0	0	0	0	0	0	

Definitions

Category	Type	Description
Annualized Impact		Estimated annual impact of Six Sigma projects on each of the Benefit/Cost items
Potential Net Benefit		Estimated net benefit of Six Sigma project in the Opportunity phase as required for initial business case
Sales Increase		Revenue changes due to implementation of Six Sigma projects
Savings	**Hard**	Identifiable changes in the cost of operations due to Six Sigma project implementation. Savings can result from headcount reduction, changes to planned expense for an ongoing activity
	Soft	Productivity improvements not immediately identifiable to hard cost reductions, expense reduction/cost avoidance
Cost	**Hard**	One time expenses associated with implementation of Six Sigma project recommendations, annualized depreciation expense, S/W purchases, vendor or Subcontractor cost, etc
	Soft	Cost associated with the implementation of productivity improvements and other cost avoidance actions
Net Benefit		Net impact of Six Sigma project (Benefit - Cost incurred). Measure applies to Hard and Soft Benefit
Profit Before Tax		Profit Before Tax (PBT) impact as a result of Six Sigma project implementation
Capital		Changes to capital expenditures as a result of Six Sigma project implementation
Working Capital/ Receivables		Changes in receivables, short-term assets and payables as a result of Six Sigma actions
Other		Changes in other assets as a result of Six Sigma actions

Business Case Examples

Here are a few examples of business cases developed during the chartering process.

Proposal process improvement opportunity – experience shows that for fixed price bids; the proposal estimation process under-predicts development costs by 2% on average (**the pain**). The average cost of contracts for the last year is $50M and we expect to win about 10 bids in this division next year. The opportunity is then 0.02 x $50M x 10 bids = $10M – assuming the improvement can eliminate the under-prediction – a clear Black Belt project!

Unbillable line items – invoice errors result in about 3000 line-items a month that can't be billed to the customer (**the pain**). The average cost of a line item is $20. The opportunity associated with this project is then 3000 items/month x $20/item x 12 months = $720,000 – again, a clear Black Belt project.

Alliance/Partnership lost opportunities – the current clearance process discourages/prevents use of lower cost alliance labor (**the pain**). Assuming equal productivity; the differential labor rate is $7/hour. The difficulty here is in estimating the number of potential hours that could be used from the alliance resource pool. However, if we could predict that the improvement would allow us to use about 11 more person-years of alliance labor (in the next year), then the opportunity would be $7/hour x 2000 hrs/year x 11 years >= $150,000 – in the neighborhood of a Six Sigma project for a Black Belt.

Improving New Hire Orientation – newly hired employees have difficulty/take longer "navigating" in the business environment (i.e. finding the right people, resources, knowing what tools are available) (**the pain**). A few assumptions need to be "strung together" here. How often does the employee seek information, how long does it take them (vs. an experienced employee), what's the labor cost/hour, how many new employees are/will be hired in a year? Suppose that the employee "needs" information once a day and that it takes them 30 minutes to find it (vs. 10 minutes for an experienced employee). At a labor rate of $100/hour, if we could improve their productivity to the same as an experienced employee, a savings of $33/employee/day, or about $6600/year. So, if the division hires more than 23 employees a year (23 x $6600 ~ $152K), then this project meets the Black Belt criteria. Note that although the opportunity fits, this doesn't include costs of improving the process. Although it's too early in the project to predict the actual figures, the conservative Black Belt will make sure the opportunity is sufficiently large to cover investment costs (e.g. an IT system change).

Software Defect Containment – the current software development process "allows" defects to escape to later phases (**the pain**). Industry statistics show that if a defect can be "caught" in the design phase, the cost of repair is about $100

(about one labor hour) vs. $5000 if detected in the field (assumption: 50/1 ratio for costs). The challenge here is to estimate how many defects "escape" to the field (and perhaps normalize this number for the size of the development program). Suppose for large software development efforts, 20 defects are discovered in the field. If all of these could be detected in the design phase, a savings of $4900 x 20 = $98,000 per program could be achieved. If the business is scheduled to produce four software releases in the next year, the opportunity is about $400,000 – a Black Belt project.

Delayed Contract Closeouts – The current process produces delays in closing a contract; subsequently delaying the issuance of a final bill (**the pain**). Improving this process will impact the business cash flow. Again, a few assumptions backed up with some initial data collection can help establish a business case. Suppose the contract bill could be issued and paid in month "X" (ideal case). On average, the current process results in the bill being issued and paid in month "X+3." The opportunity, then relates to the time value of money –the present value of the "X+3" month payment. Assuming an annual cost of money of 10%, $1000 paid in month "X+3" is worth $975 in month "X." In other words, our cash flow is impacted by $25 for each $1000 of the delayed payment. If the improved process was applied to 10 contracts in the next year, each worth $1,000,000, then the opportunity is roughly ($1,000,000 - $975,000) x 10 = $250,000.

Sponsor's Role:

The team sponsor (or champion) plays a critical role

- Usually a member of the lead team. Typically, they have authority over the process being addressed.
- Coordinates with the lead team.
- Assists team leader but does not attend all the meetings. (Helps define the scope).
- Supports the team in obtaining resources, information, and removing barriers.
- Helps resolve cross-functional issues.
- Guides team leader in all matters related to the project.

What does the sponsor want from the team?

3.1.2 Team Processes

Team Meeting Process

One of our best teams *hated* meetings. They loved doing the work associated with the project, communicated frequently in their offices, the hallway, and after work at the local “watering hole.” As the project leader, we tried to oblige them by *minimizing* the number of “whole team” meetings.

If you feel you simply *must* meet, though, here’s a process that can help make your meetings short, sweet and productive. An example meeting agenda appears on the next page.

Seven Steps of a Meeting:

1. Clarify Objectives: Make sure that everybody present has a clear understanding of what is to be accomplished in this meeting.

2. Review Meeting Roles: Assign the roles of recorder and timekeeper. Decide how time feedback will be given.

3. Review of the Agenda: Review the items listed in step 4. Make sure that everybody agrees with the agenda and the items are consistent with the objective.

4. Work through the Agenda: Try to stick to the agenda. Manage the time spent on each agenda item and "bank" or "borrow" time consciously.

5. Review of the Meeting Record: Review the flip chart or other records generated during the meeting. Decide which ones represent the "record" of the meeting.

6. Plan Next Steps & Next Meeting Agenda: Decide what actions are needed before the next meeting. Determine the objective and agenda items for the next meeting.

7. Evaluate the Meeting: How well did the meeting go? What improvements could the team make to the meeting process? Build these ideas into the next meeting.

Meeting Agenda Form – Example

Team **The Leakers** Date **8/17/XX**
Place **B Conference Room** Time **2:00 – 3:00 PM**

Item Time	Content
5 Min.	**1. Clarify Meeting's Objective** Begin building Cause and Effect Understanding of Air Compressor Head leaks
	2. Review Roles A. Einstein Team Leader B. Franklin Recorder A. Lincoln Timekeeper T. A. Edison Facilitator
	3. Review Agenda Items
	4. Agenda Items:
15 Min.	A. Review Lab Analysis
5 Min.	B. Clarify Problem Statement
25 Min.	C. Develop Air Comp. Head Leak Cause and Effect Diagram
10 Min.	**5. Review Meeting Record**
	6. Plan Next Steps & Meeting Agenda
	7. Evaluate Meeting

Planning & Managing the Project

Project management is an art and science to be applied to a team effort. The "usual" focus of project management is on the cost and schedule of the project. This project plan should align with the charter milestones.

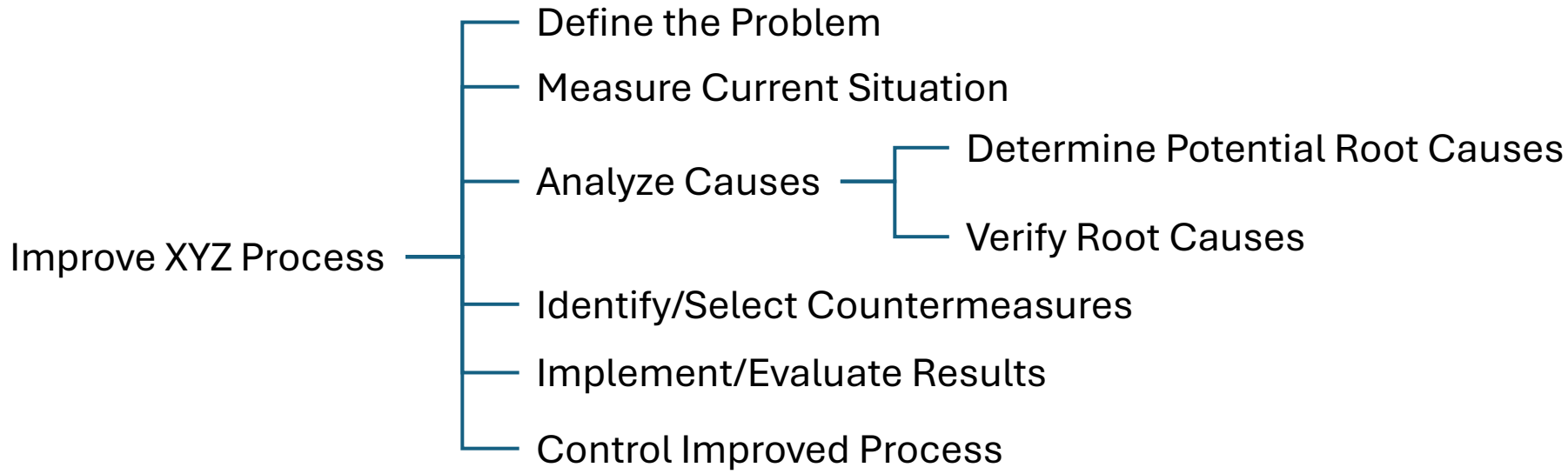
Identifying the tasks to complete a project - The project tasks are typically identified one level at a time using brainstorming and an organizing tool called a *work breakdown structure*.

Steps to develop a work breakdown structure.

1. Clearly define the objective of the project.

2. Identify the major categories of tasks needed to complete the project.

3. Divide the major categories into more detailed tasks, one level at a time, until the individual tasks can be performed in less than <u>one week</u> by the resource it is assigned to.



Sequencing and linking the tasks to complete a project

- The work breakdown structure helped identify the tasks needed to complete a project. A network diagram can be used to put the tasks in order.

- The network diagram can be used at high and low levels of detail.

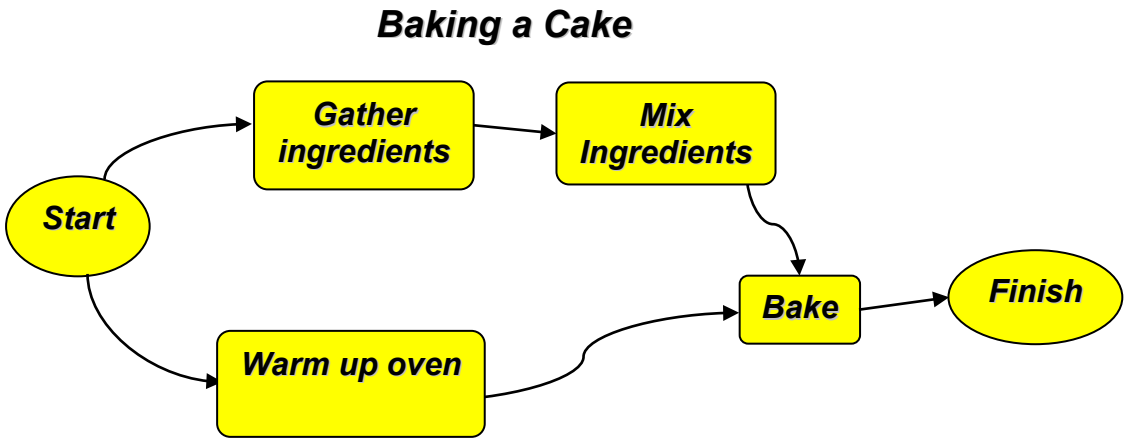
Developing a Network Diagram (see also Arrow (PERT) diagram, Section 16)

1. Define the objective for the plan.

2. List each of the tasks needed on a sticky note.

3. Starting on the left put the tasks that have no preceding tasks for them. Next put the tasks that follow the initial ones and so on.

4. Draw arrows from a node labeled start to the initial tasks and then to the next tasks until you have completed all the tasks. Draw arrows from the last tasks to a node labeled finish.



Gantt Charts

Once the tasks and their relationships have been identified, the project schedule can be developed. Gantt charts are one of the most popular schedule depictions. A simple Gantt chart example is shown below. The plan should be updated as progress is made (gray bars represent completed tasks). Project management software such as Microsoft Project can help you manage complex projects. However, don't forget that "managing the plan" takes time and effort. Often, a simple chart drawn on a flipchart or whiteboard is most useful.

PROJECT PLAN			
Project:	**Locomotive Low Voltage Wire Insulation Failure Analysis**		
TEAM MEMBERS:	B. J. TRIM	C. E. RASHER	
	R. L. YOUNG	J. B. HARSTAD (F)	
T - TEAM LEADER	R. L. HAVRANEK	C. M. LAIN	
F - FACILITATOR	J. F. MASTERSON (T)		

PROJECT TASKS	**WEEKS BEGINNING**										
	7/5	7/12	7/19	7/26	8/ 2	8/ 9	8/ 16	8/ 23	OCT	NOV	DEC
Gather Field Failure Info											
Obtain Failed Parts											
Physical Exam of Parts											
Lab Analysis (if necessary)											
Cause & Effect Analysis											
Identify Solutions											
Cost Benefit Analysis											
Present to Steering Committee											
Implement Solutions											
Track Results/ Standardize											

Managing a Project

Managing a project includes getting the tasks completed but also includes modifying the plan when necessary. In the role of project manager, you should monitor the completion of tasks and be available to adjust the plan.

What are the options for adjusting the plan?

Reports will be requested on projects. It is best to set up the communication plan as part of the plan.

How could you report progress on a project?

Many companies adopt a *storyboard* approach to communicating improvement project status. See Appendix C for a sample DMAIEC storyboard template.

Team Reviews (see also Section 16)

- Local management to monitor the progress of team process improvement efforts conducts improvement Team Reviews.
- These reviews provide an opportunity for coaching, feedback, and recognition.
- Team Reviews may be conducted at any time, but should at least be held monthly at a regularly scheduled time, e.g., the last Thursday of the month at 1:00 p.m.
- Ongoing reviews establish a "predictable event" that aids both the presenting team and local leadership with preparation planning.

Generally, a Team Review lasts 30-45 minutes, and consists of:

Sample Team Review Agenda		
Who	**What**	**Time**
Team	Present progress to date in problem solving and applying tools and techniques	10-15 minutes
Team	Addresses any open action items	5 minutes
Reviewers	Ask questions of the presenting team	5-10 minutes
Team	Respond to sponsor questions	5 minutes
Team	Discusses next steps	5 minutes
Reviewers	Summarize review and feedback	5 minutes

Strategies for assuring effective reviews:

Strategy One – Help Develop a Positive Self Image

- Acceptance – Accept the person and his/her work as worthwhile.
- Appreciation – Demonstrate sincere recognition for effort.
- Approval – Offer pleasant, honest feedback - intended to build, not tear down.

Strategy Two – Help Develop Self confidence

- Time – Taking time for reviews emphasizes the importance of the work being done. Is more time required?
- Talents – Recognize the talents that people have.
- Trust – Build a trusting relationship.

Strategy Three – Help develop a Foundation of Trust

- Meaningful – As reviews become regular events, work to ensure their effectiveness.
- Measured – We are on a journey; consider the time and dosage of new information to be shared.
- Model – We learn from each other; walk the talk.

Team Review Guidelines:

Team Uses DMAEIC Process Improvement Method	• During and between meetings, team members document their improvement activities using the DMAEIC steps and tools
Team prepares for Presentation	• Teams prepare review material in the Improvement Story format • Prior to review, presenting teams provide their sponsor with supporting documentation
Reviewers prepare for presentation	• Sponsors provide copies to other reviewers • Reviewers prepare questions for the team
Team makes presentation	• Team addresses Improvement story checkpoints
Reviewers evaluate team using checkpoints	• Reviewers use the Improvement Story checkpoints to evaluate the team's presentation and guide their feedback

Getting Ready for a Review

Planning and preparation will help ensure effective reviews.

Plan
- Administrative details
- Room setup
- Equipment/materials
- Who will present/involve team

Prepare
- Presentation materials
- Anticipate questions/prepare responses
- Agenda
- Examples
- Expert resources

Practice
- Review content
- Confirm timing
- Check details

Perform
- Work your plan
- Stay on schedule
- Respond to questions with data

Conflict and Other Team Wonders

Conflict

Quality improvement means *CHANGE.* Change is difficult; and it just may bring out some conflict from time to time. Now here, we're talking about *unhealthy conflict.* The healthy interchange of ideas in a team, where people are free to express opinions (the statistician's credo, though is *In God we trust, all others must bring data!*) is not conflict. In fact, if your team members don't have disagreements from time to time, then something else is wrong! Now we will admit to *not* being experts on conflict resolution. But we have learned (painfully!) to ***not*** avoid conflict in team situations. One way or the other, we'll try to surface it, understand it, and address it through:

Solution - Make a decision that addresses the problem's symptoms, but not its causes (some corporate personal relationships have had years to fester, we are not going to turn enemies into friends in one or two months).

Resolution - Try to reach a compromise that satisfies both sides (everybody gives in, but nobody is quite happy), or

Dissolution - Remove the conditions that caused the conflict (as close to *WIN-WIN* as practical).

There is one last option that we've exercised only once - we had to *dismember* an individual from the team (actually, he dismembered himself).

Loss of Mission Accident

Machiavelli pointed out that a walled fortress needs only stock provisions for one year. The invading force would usually lose interest if the siege lasted longer than that. He has been proven correct repeatedly. Most quality improvement efforts should take much less than one year. But be careful of the passage of time, personnel changes, etc. that can cause the team to wake up one day and ask: "*Why in the heck are we working on this project"*

One of our friends is in the painful position of trying to implement the last pieces of an organizational restructuring effort. She's about the only one left who remembers why they decided to do the restructuring and why many of the change decisions were made. She's also about the only champion of the change in the organization. It's an uphill battle for her, every step of the way.

Closing Out the Team

Once the team has achieved "victory," the only remaining task is to close out the team. This is usually not difficult, but a few points from experience:

Recognize their work and that the project is over - Humans need closure on things. A recognition ceremony can accomplish both. It's a chance to look back on the project, relive the "glory days," and recognize that it's time to face the future.

The ceremony is also a good way to promote the idea that teams can be a good way of accomplishing improvements.

Let them rest on their laurels, *at least a few days* - After a two-year, intense project, a team that we led was faced with another arduous project. For about two weeks, though, we let the team members "putter around," cleaning up project files, their offices, and generally getting things in order. Not a lot of "productive" work was accomplished.

But the team needed to take a break and gather their energy for the next project. There was no way we could have jumped into the next project the day after the first was complete.

Take time to evaluate the project - Every time we go through an improvement project, there's something we recognize that could have gone better. We also think about the things that went particularly well, so we can repeat them in the future when appropriate.

3.2 Idea Generation & Decision Making

Learning Objectives

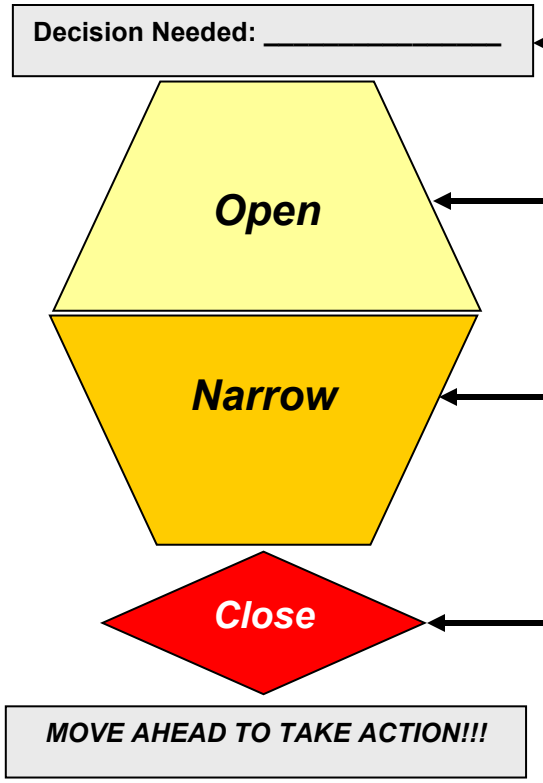
- Be able to lead a group to generate ideas
- Be able to lead a group to a decision

Unit Contents

- Idea Generation & Creativity Methods
- Decision Making

3.2.1 The Team Decision Process

Teams must make decisions. Unfortunately, wags have observed, "*more than three people can't figure out where to go to lunch.*" We've seen too many teams that can't reach decisions or take excessive amounts of time to reach a decision. Recognizing that there is a *process* associated with decision-making can help the team here. One simple model is presented below:



Decision Needed - What is the decision that must be reached? Clarify specifically what must be decided. Write it on a flipchart so everybody knows the goal!

Open for Ideas – Generate as many ideas as possible (or are needed) to provide decision options. Employ techniques such as brainstorming (see 3.2.2) to generate the ideas quickly and efficiently.

Narrow the List of Ideas – Don't go too quickly for the final decision. Identify the most likely ideas from the initial list. Use methods such as Multivoting (see 3.2.3) to narrow the list.

Close to Reach the Decision – Using either group techniques or data analysis (e.g. for a root cause verification, data should be employed), come to a final decision. Group techniques include consensus (see 3.2.3), but sometimes voting is necessary to move ahead. Try to avoid "Teflon consensus" – where the decision doesn't *stick!"*

In 3.2.2, we'll present several techniques to facilitate generating ideas. Then, in 3.2.3, we'll provide you with means of reducing the number of ideas and reaching agreement on the "best" ideas.

3.2.2 Idea Generation & Creativity Methods

Brainstorming

Often a team will need to generate ideas as part of the quality improvement effort. Developing the Cause-and-Effect Diagram is one example where many ideas are needed. Brainstorming allows a group to quickly develop a large list of ideas without spending time "beating each idea to death."

Brainstorming Steps:

1. Clearly state the purpose of the Brainstorming session.

2. Select Recorder(s) to capture ideas on flip charts.

3. Call out ideas in a "round robin" style (each person takes a turn, going around the group - it's OK to "Pass").

   - *Don't discuss or criticize ideas (sometimes, the ideas "from left field" turn out to be the most useful),*
   - *Build on ideas of others. Listen to others' ideas; you may be inspired!*
   - *Note: A variation of Brainstorming asks each member to write ideas down before the session begins.*

4. When the "round robin" has slowed down, open the brainstorming session up to any additional ideas.

5. When the brainstorm has ended, review the list. Clarify the remaining ideas (add additional words) making sure that everybody understands each idea. Delete any duplicate ideas.

Attribute Listing

Attribute listing is a technique for ensuring all possible aspects of a problem have been examined. Attribute listing breaks the problem down into smaller bits and then discovers what happens. Let's say your business is making flashlights. You are under pressure from your competition and need to improve the quality of your product. By breaking the flashlight down into its component parts - casing, switch, battery, bulb and the weight - the attributes of each one - you can develop a list of ideas to improve each one.

Attribute Listing - Improving Our Flashlight

Part/Feature	Attribute	Ideas
Casing	Plastic	Metal
Switch	On/Off	On/Off low beam
Battery	Power	Rechargeable
Bulb	Brass	Plastic
Weight	Heavy	Light

Attribute listing is a very useful technique for quality improvement of complicated products, procedures for services. It is a good technique to use in conjunction with some other creative techniques, especially idea-generating ones like brainstorming. This allows you to focus on one specific part of a product or process before generating a whole lot of ideas.

Imitation

How many ideas are really original? It is quite valid to imitate other ideas as a preparatory step to original thinking. Try what all the "great" creators have done: *imitate, imitate, imitate*. After you have imitated enough, you will find your preferences shape what you are doing into a distinct style. Originality is a natural result of sincere creative pursuit. Isaac Newton said:

"If I have seen farther, it is by standing on the shoulder of giants".

Just as the Beatles started out playing cover tunes, J.S. Bach went blind in his old age copying scores of other musicians (for personal study), Beethoven played on the themes of his time, and Jazz musicians insert popular melodies into the middle of bizarre atonal solos (for an interesting book linking this theme and creativity, see *Jamming,* by John Kao). Ideas are constantly on the move, much to the annoyance of patent & copyright lawyers! Certainly, ideas may be exploited by the materially minded, just like anything else. But if you truly comprehend an idea, it is yours.

"What is originality? Undetected plagiarism."
Dean William R. Inge

“The immature poet imitates; the mature poet plagiarizes.”
T. S. Eliot

Assumption Smashing

A useful idea generating technique is to list the assumptions of the problem and then explore what happens as you drop each of these assumptions individually or in combination. For example, say you work in the Customer Service division of a software company. When customers purchase software, they are encouraged to purchase support agreements for a cost of 15% of the software value. The revenue from this maintenance funds the support personnel who answer telephones. The assumptions of this situation are:

- Customers purchase maintenance agreements
- Customers pay 15% of the software's worth for support
- Support is a product and should therefore be sold
- The software vendor provides helpful, timely support

Now think about the situations as each attribute is dropped.

What happens if support is free? - Maybe the software price should be increased and the support given away, creating the impression of free support.

Don't support the product - Don't offer support. The vendor doesn't have to support it, so doesn't have to employ support staff. If anyone rings for help, tell them to buzz off! This could lead to customers forming their own support groups (user groups) or turning to other areas such as the Internet, bulletin boards, newsletters, independent support specialists and so on.

Even more assumptions could be dropped.

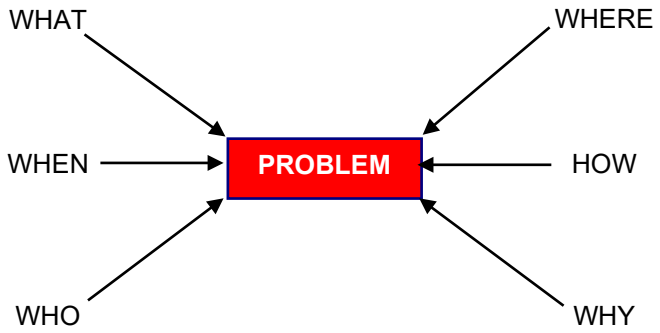
What if the vendor **gave away** the software? You probably have a copy of Google Chrome or Adobe Acrobat. Did you buy that software? How do you think Google makes money if most people don't pay for the browser?

The Six Universal Questions

Idea Generators should be aware of a simple universal truth. There are only six questions that one human can ask another:

What?
Where?
When?
How?
Why?
Who?

You may want to draw a mind map of the problem with these six words as nodes on the map.



Checklists

Alex Osborn in his pioneering book **Applied Imagination** talks about "Questions as spurs to ideation", and outlines about 75 idea-spurring questions. The simplest set of questions comes from the six basic questions described above:

- **Why** is it necessary?
- **Where** should it be done?
- **When** should it be done?
- **Who** should do it?
- **What** should be done?
- **How** should it be done?

What other uses? is a good question. By adding uses we can often add value. By piling up alternatives by way of other uses, a still **better** use is likely to come to light. Osborn went on with the following questions:

- Adapt?
- Modify?
- Substitute?
- Magnify/Maximize?
- Minimize/Eliminate?
- Rearrange?
- Reversal?
- Combine?

DO IT

This technique is fully described in the book **The Art of Creative Thinking** by Robert W. Olson. The name is based on the following abbreviation:

Define Open Identify Transform

The pattern of the DO IT process emphasizes the need to **Define** problems, **Open** yourself to many possible solutions, **Identify** the best solution and then **Transform** it into action effectively. The ten DO IT catalysts, designed to help us creatively define, open, identify and transform, are...

- Define
  - Mind Focus
  - Mind Grip
  - Mind Stretch
- Open
  - Mind Prompt
  - Mind Surprise
  - Mind Free
  - Mind Synthesize
- Identify
  - Mind Integrate
  - Mind Strengthen
  - Mind Synergize
- Transform

The DO IT Process and Catalysts

The DO IT catalysts may be used effectively separately for quick problem solving, or together as a process when very important or difficult problems need to be solved. They are designed to accelerate and strengthen your natural creative problem-solving ability and to stimulate many good, diverse ideas for solutions to your problems.

Write down a statement of the problem!

Define the problem carefully to make sure you are solving the real problem and to help engage your unconscious and conscious minds to the problem.

Mind Focus	Ask why the problem exists. This may lead to a broader statement of the problem. Try to subdivide the problem into smaller problems. This may lead to a narrower restatement of the problem.
Mind Grip	Write down at least three two-word statements of the problem objective. Select the combination of words that best represents the precise problem you want to solve. Use this to write a new, more optimal and effective restatement of the problem.
Mind Stretch	List the goals, objectives and/or criteria that the solution of the problem is to satisfy. (Think of the obstacles that must be overcome.) Then stretch each goal, objective or criterion and write down any ideas that are stimulated.

Write down the most optimal statement of the problem

Open yourself to consider many diverse solution ideas. Delay judgment on ideas generated until the Identify step. First, list any ideas that are on your mind. Then....

Mind Prompt	Ask other people with diverse backgrounds, knowledge and intelligence for solutions to your problem. Use their solutions as prompters for your own ideas.
Mind Surprise	List ridiculous, laughable ideas. Use them to trigger more reasonably, possible usable solutions to your problem.
Mind Free	Stimulate fresh ideas by forcing similarities between your problem and things that aren't logically related to your problem. 1 - Write down the name of a physical object, picture, plant or animal. 2 - List its characteristics in detail 3 - Use the listed characteristics to stimulate insights into and ideas for the solution to your problem.
Mind Synthesize	Bring the ideas together. Draw them visually and draw connections between the ideas.

Circle the best of ideas generated so far during the Define and Open steps

Identify the best solution to your problem and modify it until you are ready to transform your idea into action.

Mind Integrate	Review your goals, objectives and/or criteria then trust your own gut-level feeling to select the best idea from the already circled ideas.
Mind Strengthen	List the negative aspects of your idea. Be vicious! Try to positive the negatives. Then **modify** the **solution** to reduce the negative aspects.
Mind Energize	Exaggerate the worst and best potential consequence that might result from the implementation of your solution. Modify your solution to minimize bad consequences and maximize good consequences. Proceed to the transformation step if you are sufficiently energized.

Carefully write down a statement of your final solution idea

Transform your solution idea into action. Use the DO IT process and catalysts again to help creatively solve the problem that you now have of "How to transform your solution idea into action."

Note: When time allows, take advantage of incubation (unconscious thinking) and research processes (find out what ideas have already been tried). Most of our everyday personal and professional problems are solved in a few minutes or instantly. Therefore, you will probably find it advantageous to use only one or a few of the catalysts at a time.

Forced Analogy

Forced analogy is a very useful and fun-filled method of generating ideas. The idea is to compare the problem with something else that has little or nothing in common and gaining new insights as a result. You can force a relationship between almost anything, and get new insights - companies and whales, management systems and telephone networks, or your relationship and a pencil.

Forcing relationships is one of the most powerful ways to develop new insights and new solutions. A useful way of developing relationships is to have a selection of objects or cards with pictures to help you generate ideas. Choose an object or card at random and see what relationships you can force. Use mind mapping or a matrix to record the attributes and then explore aspects of the problem at hand. An example follows:

Marriage as a pencil - Betty Edwards in her book Drawing on the Artist Within shows the example of a pencil used to examine aspects of a marriage.

Pencil	**Marriage**
Gold Ring	Remember promises
Blue Ring	Clean the tub. I share depression too often with family
Yellow	Too timid. Harold needs to know my true feelings
Flat side	Dull daily routine. Change activities
Six sides	6 things to do: Budget, Take a class, Improve discipline, be more assertive, improve communications and start now!
Eraser	Rub him out! Forgive and forget past mistakes
Money	Spend too much. Need a budget. Take a job
Superior	I feel inferior to my husband
Wood shaft	Feel closed in. Need other interests. Am I getting shafted?
Lead	Get the lead out! Do It! If I press any harder I will break.
Write	Send a note telling Harold that I love him.

Problem Reversal

The world is full of opposites. Of course, any attribute, concept or idea is meaningless without its opposite. Lao-tzu wrote *Tao-te Ching* that stresses the need for the successful leader to see opposites all around:

The wise leader knows how to be creative. To lead, the leader learns to follow. To prosper, the leader learns to live simply. In both cases, it is the interaction that is creative. All behavior consists of opposites...Learn to see things backwards, inside out, and upside down.

The method

1. State your problem in reverse. Change a positive statement into a negative one. For example, if you are dealing with Customer Service issues, list all the ways you could make customer service *bad*. You will be pleasantly surprised at some of the ideas you will come up with.

2. Try to define what something is *not*. Figure out what everybody else is *not* doing. For example, Apple Computer did what IBM didn't, Japan made small, fuel-efficient cars.

3. Use the "What If" Compass - Just ask yourself, "What if I ....." and plug in each one of the opposites. A small sample:

- Stretch it/Shrink It
- Freeze it/Melt it
- Personalize it/De-personalize it

4. Change the direction or location of your perspective. This can include a physical change of perspective, managing by walking around, or doing something different.

5. Flip-flop results - If you want to increase sales, think about decreasing them. What would you have to do?

6. Turn defeat into victory or victory into defeat - If something turns out bad, think about the positive aspects of the situation. If I lost all the files off this computer, what good would come out of it? Maybe I would spend more time with my family! Who knows!

The Six Hats (DeBono)

DeBono has developed a model based on six metaphorical hats. The thinker can put on or take off one of these hats to indicate the type of thinking being used. This putting on and taking off is essential. The hats must never be used to categorize individuals, even though their behavior may seem to invite this. When done in a group, everybody wears the same hat at the same time.

"Thinking Hat"	Description
White	This covers facts, figures, information needs and gaps. "I think we need some white hat thinking at this point..." means Let's drop the arguments and proposals and look at the data base."
Red	This covers intuition, feelings and emotions. The red hat allows the thinker to put forward an intuition without any need to justify it. "Putting on my red hat, I think this is a terrible proposal." Usually, feelings and intuition can only be introduced into a discussion if they are supported by logic. Usually, the feeling is genuine but the logic is spurious. The red hat gives full permission to a thinker to put forward his or her feelings on the subject at the moment.
Black	This is the hat of judgment and caution. It is a most valuable hat. It is not in any sense an inferior or negative hat. The black hat is used to point out why a suggestion does not fit the facts, the available experience, the system in use, or the policy that is being followed. The black hat must always be logical.
Yellow	This is the logical positive. Why something will work and why it will offer benefits. It can be used in looking forward to the results of some proposed action but can also be used to find something of value in what has already happened.
Green	This is the hat of creativity, alternatives, proposals, what is interesting, provocations and changes.
Blue	This is the overview or process control hat. It looks not at the subject itself but at the 'thinking' about the subject. "Putting on my blue hat, I feel we should do some more green hat thinking at this point." In technical terms, the blue hat is concerned with meta-cognition.



3.2.3 Decision Making

Once you've generated a set of ideas (the OPEN part of the decision-making process), it is time to begin to narrow down the list. You enter the NARROW and CLOSE parts of the process.

Reducing the Number of Ideas (Multivoting)

Let's say that your brainstorming session has identified twenty different countermeasures that could be applied to the root causes of a problem. Rather than discuss each one, in turn, the Multivoting method can be used to reduce the list to a more manageable size. If the Multivoting process leaves you with about 10 ideas, then you can use Rank Ordering to further reduce the list (see next page).

Multivoting Steps

1. Clarify the purpose of the Multivoting activity.

2. Decide the criteria to be applied to the voting (most cost-beneficial, most probable root causes).

3. Decide how many votes each member gets (usually 20 - 25% of the total number of ideas, for example, if you brainstormed a list of 25 ideas, each member would get 5 or 6 votes.).

4. Each member votes for the ideas that best fit the criteria.

5. Votes are recorded - the ideas that get most votes are circled and pursued further.

- *Voting may occur again, if a large list remains, or Rank Ordering can be used.*

- *Never Multivote down to one item - Use Consensus to decide on the one item from the multivoted list.*

Reducing the Number of Ideas (Rank Ordering)

Rank Ordering is often used to reduce a list of 10 or *fewer* ideas. Rank Ordering can be used by itself or following a Multivoting session.

Rank Ordering can be used to reduce the list to three to five ideas that will be pursued further in detail through a consensus process.

Rank Ordering Steps

1. Clarify the purpose of the Rank Ordering activity.

2. Decide the criteria to be applied to the ranking (most cost-beneficial, most probable root causes).

3. Label each idea with a *letter.*

4. Each member ranks the ideas from "best fit" (rank of 1) to "least fit" (rank of "*n*," where *n* is the total number of ideas). All ideas on the list are ranked.

5. The rankings are recorded from each team member and summed by idea - the ideas that get *fewest* votes are circled and pursued further.

- *Never Rank Order down to one item - Use Consensus to decide on the one item from the ranked list.*

Reaching Agreement on One Idea (Consensus)

Deciding on the best Countermeasure from several alternatives can be difficult for a group. Consensus is a method used to obtain support for an idea from the members of the group and their agreement to help carry it out, if it requires that action be taken.

Consensus is sometimes hard to achieve and takes time (more than that needed by one person to make a decision or for the group to vote on a decision), but it is worthwhile, since the agreement is generally considered to be a WIN-WIN for everybody in the group.

Consensus Process

1. Clarify what is to be decided and why consensus is important for the decision.

2. Members prepare their own positions, using the facts and data available (this is usually done prior to a consensus meeting).

3. Members share their positions (and the supporting facts and data), with the group actively listening and note taking.

4. General discussion then follows, until agreement is reached.

   - *From time-to-time, stop, seek out, and record the issues that the group agrees upon, also record the points of difference. Further discussion, then, can focus on the differences.*

   - *Avoid "giving-in," just to save time. Remember, the group will have to live with the decision.*

   - *A facilitator, detached from the emotions of the decision, can help support the group trying to reach consensus.*

3.3 Exercises

Team Situations

In which of these following situations do you think a team should be formed to improve quality? If a team is needed, what *type* should it be? If you don't think a team is needed, how could or should the situation be addressed?

- A railroad has been experiencing water leaks on its locomotives' diesel engines. There are about 1000 locomotives in the railroad's fleet. The engineer's failure report includes where the leak is observed, but not why it occurred.

- An architectural firm has been receiving complaints from customers that "they are not responsive" to the customers' needs. The firm has four design groups, each acting as a design team for projects.

- A small manufacturing company wishes to improve its employee safety record. The company president wants to form a team, but the Safety Officer tells him that he can solve the problem with a new training program for proper lifting techniques.

- An unacceptably high defect rate of integrated circuits has plagued a small electronics firm for the last few weeks. The reliability engineer is working on a test plan, the design engineers are preparing changes to the IC design and manufacturing is changing their "clean room" procedures.

- A hospital's case managers have identified one physician as being "high" on both patient Length of Stay and Cost per Case for a certain diagnosis.

- Nurse managers have been complaining to the chief nurse executive about delays in receiving laboratory "stat" specimen reports. The lab director says the orders are only being sent to the lab twice a shift.

- A physician called plant maintenance about dust blowing into one of her examining rooms from an air conditioning vent. The problem has existed for three days now.

- Two employees on the evening shift at a plastics plant are chronically late. The other shift members are angry at having to carry their "load" when they are late.

- A manufacturer of ceramics for hobbyists found that their product sales were declining. Projections indicated that the manufacturer would suffer a $10 million loss if the current trend continues.

Team Appropriateness

Comment on the following "team" situations described below. Was the use of a team appropriate? What issues do you see that may lead (or did lead) to the success or failure of these efforts?

- A manager of an engineering division told a group of engineers to investigate computerizing a certain reference document. He told them to make sure and "prove" that the computerization was necessary so the necessary budget approvals could be obtained.

- A new chief engineer of a nuclear engineering department identified a "laundry list" of engineering practice problems. The chief assigned a group of engineering managers to form a team, prioritize the problems and start working on fixing them.

- The senior managers of a bank have just been through quality improvement training and are excited to begin improvement efforts. They assigned 10 projects to branch office and "back office" staff. The branch and "back" office managers were not consulted before these assignments were made.

- Factory management assigned a group of maintenance workers, purchasing and receiving personnel to work on reducing the time to obtain "non-stocked" spare parts for plant equipment. Three weeks after the team began, they realized a new parts inventory database was being installed in the next month.

- A manager of a nursing unit assigned a group of nurses and nurse assistants to improve morale and communication in the unit. She thought that would help reduce turnover in the unit, which was running at 40% annually.

- A president of a small consulting firm had decided to expand the company office space. He assembled a team of clerical support staff to determine the best strategy to "handle the increased need for product inventory space." The team came back to him with the recommendation to let the consultants "telecommute" from their homes and use their office space for the product inventory. The president disagreed and proceeded to lease additional space.

- A hospital initiated a number of teams whose purpose was to improve the clinical quality of care for patients. Physicians were invited to participate on these teams. Although some were initially interested, the meetings were held during the day and, gradually, the doctors stopped coming to meetings.

- Nurses and quality assurance/utilization review staff were assigned to develop “Clinical Pathways,” standardized methods of patient care. After reviewing the first five Pathways developed, the physicians told the chief nurse executive that the “Pathways were worthless, they weren’t going to practice ‘cookbook’ medicine.”

- The corporate quality improvement department told power plant managers that they needed to have a certain number of teams “running” by years end. Over 80% of plant personnel work shifts that only allow for short breaks and lunch. By the end of the year, the only “functioning” teams were composed of administrative clerks.

- A new car dealer assigned members of his sales force to develop “best practices” for selling cars to customers. After three months of meeting, the team had not made any progress. The sales personnel are paid based on commission.

- A hospital’s administration has decided to decentralize the respiratory therapy function to the patient care units. The leader of the team is the current department director. The patient care unit managers don’t want the additional responsibility of respiratory therapy, and the department director is reluctant to give up his “power.”

Exercise - Project Planning

1. Review your project charter.
2. Divide the project into phases if appropriate. (You may decide that there is only one phase)
3. Using brainstorming and the work breakdown structure identify the tasks for completing part of the project.
4. Use the network diagram to sequence and link the tasks.
5. Prepare a flipchart and be prepared to show your exercise to the group.

Scoping a Project

List a set of criteria you'd employ to determine if a project's scope is too large:

Given the following scope, how could you multi-phase it?

Decrease the cycle time from order entry to order fulfillment in all plants.

Party Brainstorm

With a small group, *Brainstorm* a list of parties that you might like to throw (Valentine's Day, Clambake, Birthday, Toga, etc.). If the list has more than 10 ideas, then employ *Multivoting* to narrow the list. If necessary, then *Rank Order* the remaining list and reach *Consensus* on the one you will plan.

Develop a *Project Plan* to prepare for the party. Decide what needs to be done, who will do it, when it will be done, etc.

Idea Generation and Decision Making

Here are some simple "decisions" that can be reached with the help of the Idea Generating and Group Decision-Making methods. Use these to practice before getting into a "real" situation. Don't forget to set criteria before you employ the decision-making tools. (*Note: these can be conducted as meetings, using the 7 step meeting process*):

- Where to have the department picnic.
- What to do for pediatric cancer patients for Christmas.
- What to do for the nursing home patients for Christmas.
- What two-mile stretch of road your organization will "adopt."
- The "best" (you define "best") sitcom of all time (corollary - which episode is "best.").
- Why women love shopping (top 3 reasons) - except for tools and hardware.
- Why men hate shopping (top 3 reasons) - except for tools and hardware.
- Which are the top ten movies of all time.
- Who are the top ten baseball players of all time.
- Who has the "worse" (you define "worse") TV talk show.
- Who would make the "best" (you define "best") next President.
- What to do for your secretary on Secretary's Day.
- What to do for your "boss" on Bosses' Day.
- What's the most important problem facing your department (Oops!, sorry, that's a real one!)

Team Dynamics

Comment on the "dynamics" occurring in these team situations. What would you do to address the situation?

- Six members of a team are working on an improvement project. One of the members, although a solid contributor, is generally negative about the project, and is convinced that management will not adopt the team's recommendations. The other members of the team have taken to avoiding him, except one who vigorously "counterattacks" him every possible chance. You are the team leader.

- A team is working to reduce supply expenses in the surgical department. The team leader often comes to the meetings late and frequently must develop the agenda in the first few minutes of the meeting. Team members have begun to avoid coming to meetings. You are a team member.

- A team assigned to design a new engineering scheduling process has several very strong personalities. One member frequently verbally attacks other members, another "acts out" and often leaves the meeting for a few minutes or wanders around the meeting room. One member heads the current scheduling department and is not happy about having others "meddle in her process." Although the team has designed a new process, planning to implement the changes is proceeding slowly. You are a team member.

- A team of assistants has been formed and asked to select their own improvement project. The team leader is a young, shy male. One of the team members is an older, domineering female who thinks this "quality stuff" doesn't apply to her and that she doesn't need to be on the team. You are the young assistant's Black Belt and he has come to you for advice.

- A development specialist has been assigned to lead a team that will design a quality improvement video on control charts. The company president has assigned members to the team, two of whom are consultants who are always joking around. Their language is sometimes "borderline crude" and they often make disparaging comments about the project. However, they have been the main two contributors to the ideas and work of the project. You are the team leader.

- You have been assigned to a team working on resolving a safety issue at a nuclear power plant. The team has frequent meetings with management. As the work continues, you start to suspect that management is not willing to address the issue and that your team's real role is to justify the existing situation.

- A fossil power plant manager is assigned to lead a team of the company's "best" quality improvement people. Their assignment is to help the company's poorly performing nuclear plant "turn-around." The team members are all strong personalities, frequently disagreeing with the team leader. The team leader is good at developing a "vision" of where he wants to go, but not so good at the details. A small subgroup takes on the role of "translators," developing coherent, practical plans to implement the leader's vision.

- You have been brought in as the new Chief Operating Officer of a large, metropolitan hospital. You assess that each of the Vice Presidents has been running their own operations, with little communication between the departments. In fact, department directors have been told *not* to talk with those directors "belonging" to other VPs. All questions or problems have to come up through the "chain of command."

4.0 Obtaining the Voice of the Customer

Unit	Description	Page
4.1	Core Customer Research Methods	4.1 -1
4.2	Exercises	4.2 - 1

4.1 Core Customer Research Methods

Learning Objectives

- Understand the Need to Listen to the Customer
- Understand and Apply a Voice of the Customer Process
- Plan, Conduct and Analyze:
  - Interviews
  - Focus Groups
  - Surveys

Unit Contents

- Why Listen to Customers?
- A Voice of the Customer Listening Process
  - Developing a Listening Strategy
  - Listening to Customers
  - Organize and Analyze Data
  - Communicate the Learning
  - Drive Business Activities

4.1.1 Why Listen to Customers?

Research shows that ***Six Sigma Companies***:

- Define key customer segments and key customers that are critical to achieve their strategy.
- Have the market's pulse and are flexible to meet changing demands.
- Strategically define contribution of key customers in terms of:
  – New opportunities
  – Profitability
  – Market share opportunities
  – Competitive advantage
- Gather customer data and turn into meaningful, actionable information:
  - Design of new products/services,
  - Enhancement of Existing Products/Services
  - Improvement of process capability.
- Establish the voice of customer (VOC) process as a critical enabling process:
  - Identify ownership for the VOC process
  - Communicate and share VOC information throughout the organization,
  - Apply the VOC information to the management of core processes, and
  - Incent & reward the organization to listen to the customer

VOC Challenge and Opportunity:

- Customers “Talk” Through Their
  - – Behaviors
  - – Reactions to specific products and services
  - – Silence

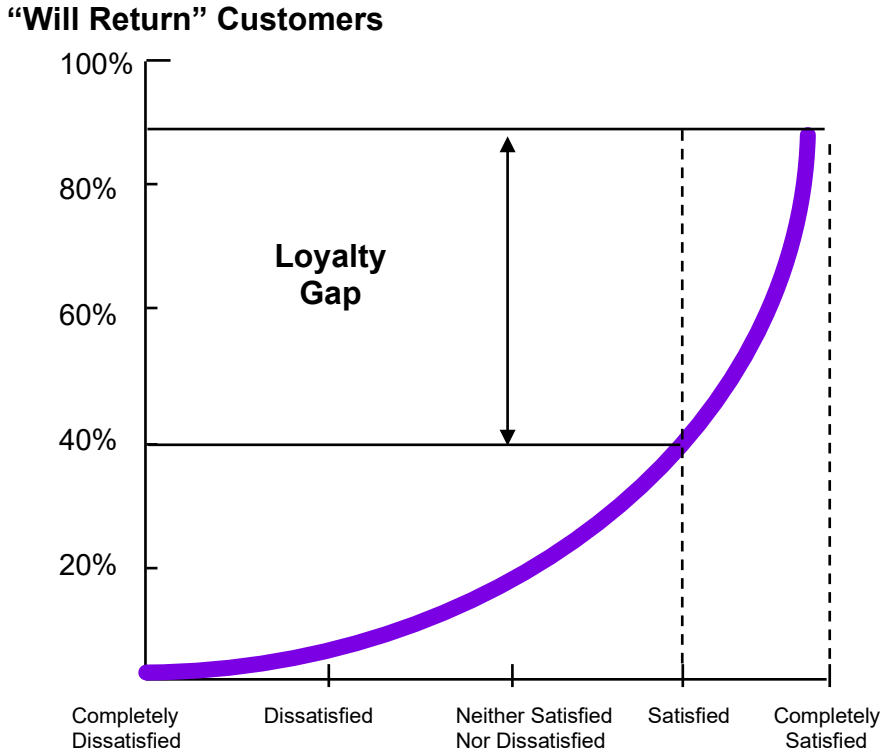
- Customer “Talk” Is “Noise”

- The Challenge For Businesses Is To Convert Customer “Noise” Into Meaning

- The Opportunity For Businesses Is To Make Listening Strategic
  - – Customer loyalty (retention, satisfaction)
  - – New markets/opportunities
  - – Increased market share

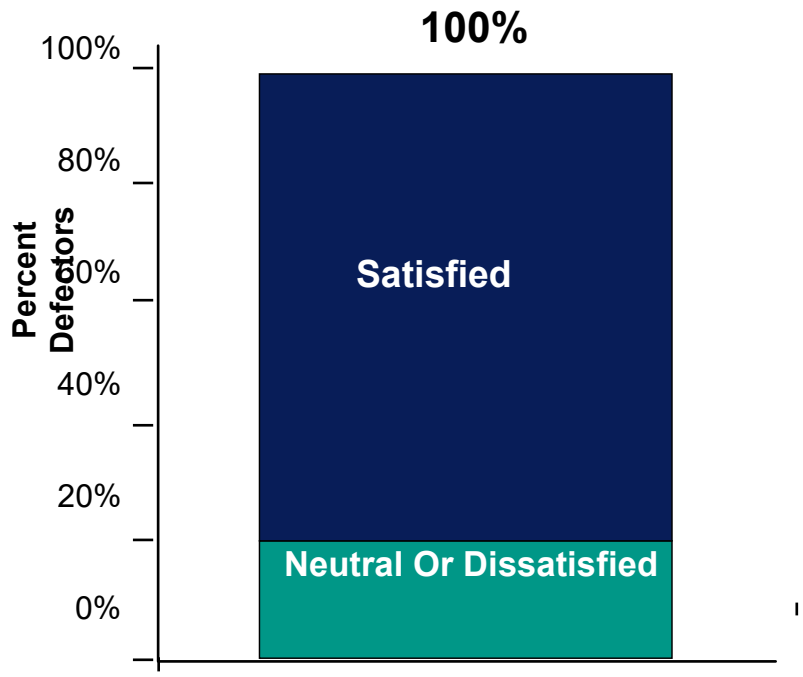
Relationship - Customer Satisfaction And Loyalty In Highly Competitive Industries

- As satisfaction goes up, so does loyalty – but the relation is not simple.
- Any drop from *total satisfaction* results in major drop in loyalty.
- In competitive markets, there is a tremendous difference between the loyalty of *satisfied* and *completely satisfied* customers. This difference is hard to achieve and is a moving target.
- It is more effective to move customers from *Satisfied* to *Completely Satisfied* than to focus on customers who are below *Dissatisfied*.



Most Defecting Customers Were "Satisfied"

- Customers Want To Be Completely Satisfied. When They Aren't Completely Satisfied, They Have Reasons…
- Most Managers Should Be Concerned If The Majority Of Their Customers Fall Into The Satisfied Category
- The Key To Keeping, Finding, And Winning Customers Is To Understand What Customers Are Saying. . .



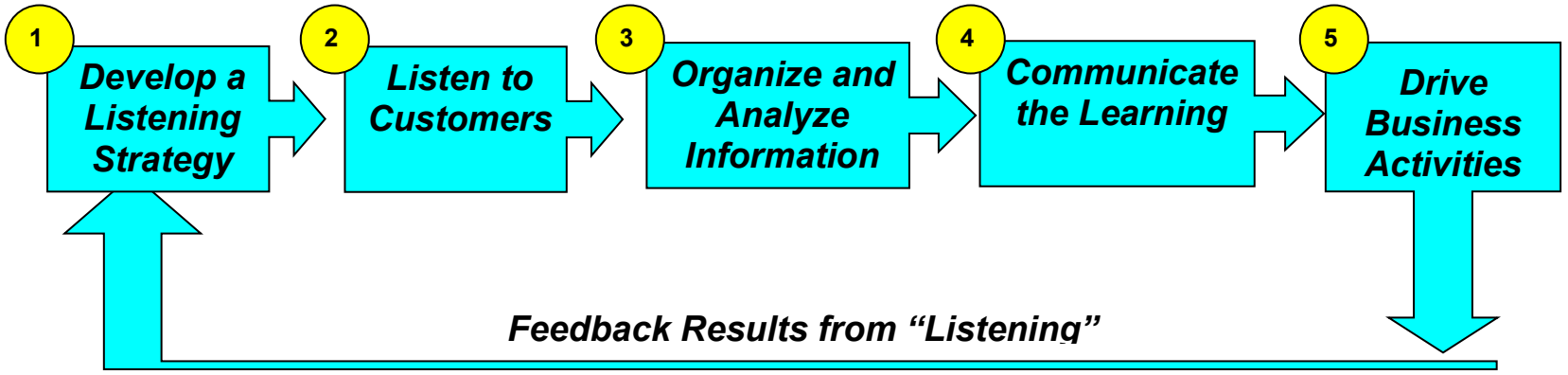
4.1.2 A Voice-Of-The-Customer Listening Process

The VOC Process Is a Continuous, Strategy-Driven Set of Activities Focused on Establishing a Learning Relationship Between Providers and Customers That Drives Business Results.

Characteristics of a Successful VOC Process include:

- We Know Our Key Customers.
- We Know How Satisfied Current Customers Are With Our Business Products And Services.
- Our Listening Process Provides The Data To Support Our Business Strategy.
- We Know Why Customers Like Products And Services From Our Competitors.
- We Know What Information We Already Have About Customers.
- We Know How To Gather And Translate Customer "Noise" Into Meaningful Data.
- We Have A Strategy For Sharing Learning About Customers With Our Customer – Both External And Internal.
- We Have A VOC Listening Process In Place That Links Listening To Business Improvement And Innovation.

Key VOC Process Steps Include:

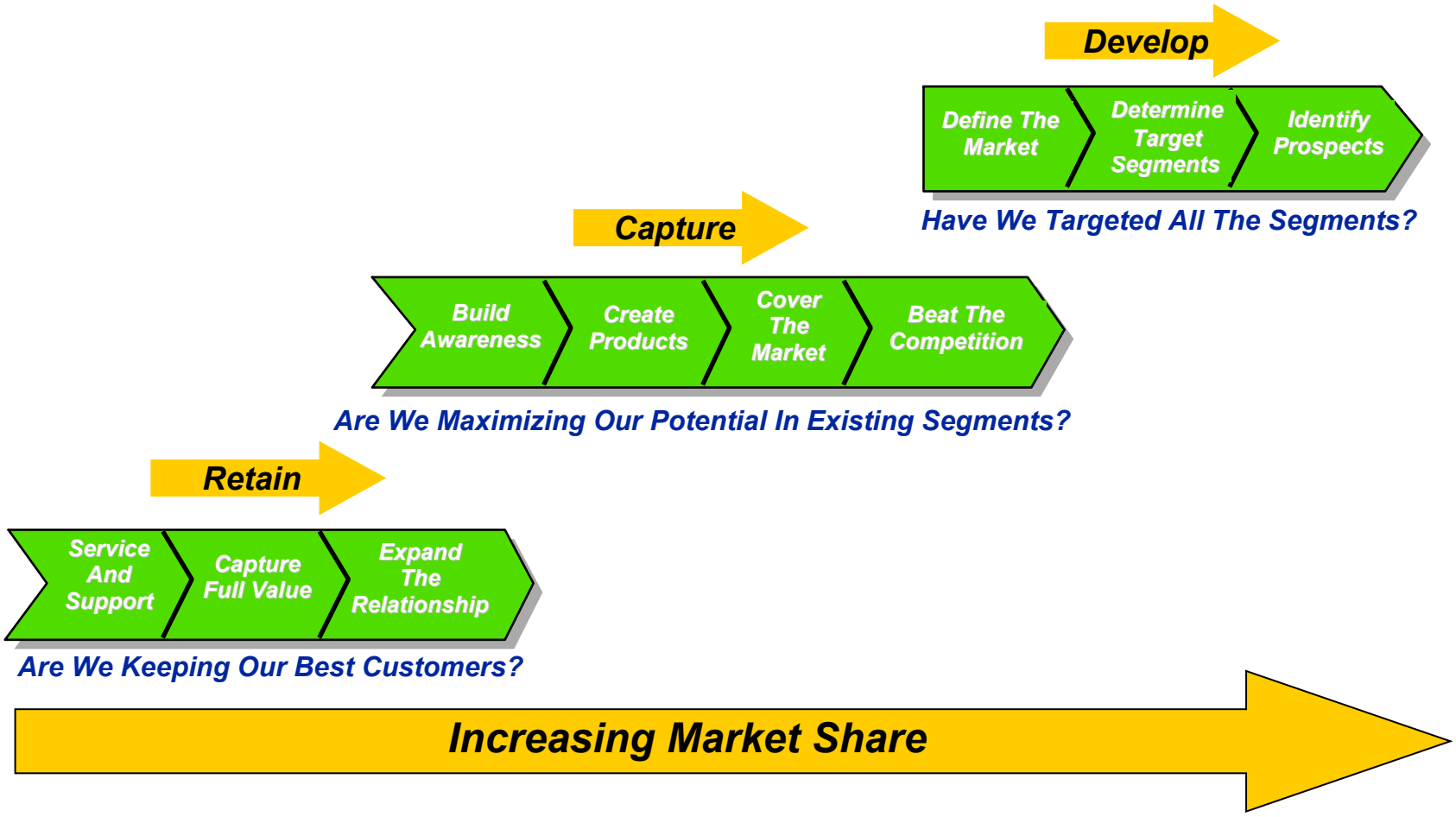


Step 1. Developing a Listening Strategy

<table>
<tr><td colspan="2">Purpose:
• To define scope, purpose, and goals of customer listening activities.
• To identify what kind of information needs to be collected from which customer segment.</td></tr>
<tr><td colspan="2">Key Activities and Sub-Activities</td></tr>
<tr><td>Understand the business need.</td><td>• Understand the business strategy.
• Identify business needs/ opportunities.</td></tr>
<tr><td>Target the customers to address business needs.</td><td>• Determine logical customer segments.
• Describe why information from these segments is critical to achieving business strategy.</td></tr>
<tr><td>Determine the information that must be gathered from customer segments.</td><td>• Identify/review existing information available on targeted customers.
• Determine what information needs to be gathered from important customer segments.</td></tr>
<tr><td colspan="2">Supporting Tasks</td></tr>
<tr><td colspan="2">• Identify current customers and what’s important to them.
• Identify ”lost” customers and why they no longer buy from you.
• Assess recent sales and marketing results.
• Talk to sales and operations staff to identify issues and concerns.
• Gather existing internal intelligence on customers.
• Review market trends.
• Review findings from past VOC listening activities.
• Review activities implemented because of previous findings.
• Conduct benchmarking study related to products and services.</td></tr>
<tr><td colspan="2">Possible Tools and Resources</td></tr>
<tr><td>Tools:
• Customer Identification Worksheet.
• Strategy Development Worksheet.
• Knowledge Needs Worksheets.</td><td>Resources:
• The Customer Driven Company - Richard Whiteley
• Customer Centered Growth - Richard Whiteley, Diane Hessan
• Keeping Customers - Harvard Business Review Book</td></tr>
</table>

A Strategy for Attracting and Retaining Customers

Six Sigma companies clearly define their customer "attraction" strategy. Retaining good customers is obviously a key first step in improving market share. Cost ratios of 5:1 – 10:1 (and some higher!) have been quantified for the resources required to attract a new customer vs. keeping an existing customer. Capturing new customers (primarily through improvement against the competition) is next. This often has limits and the develop strategy is then employed to find new markets for products and services. Customer research activities should reflect the company's strategy. A retention strategy should drive conversations with existing (and lost) customers. A capture strategy should drive discussions/ research with the competition's customers. A develop strategy may drive research in several areas – i.e. existing customers for whom we are developing a new product/service.



Who are the Customers of the Product/Service/Process?

Customers Defined:

- External Customers – those outside the company who purchase products and services (also, "bystanders" who may be affected by the product/service – noise, pollution)
- Internal Customers – "The next process is your customer." Who receives the output of your process?
- Stakeholders – Others who may be affected by the process or have a "stake" in the outcome of the process (management, employees, regulators, shareholders).

Segmentation:

Are There Different Customer **Segments**?

- A Fundamental Way To Segment Customers Is According To Customers' Similar Need For Products And Services. Other Typical Segmentation Strategies Include:
  - Revenue/Customer
  - Deal Size
  - Geographic

- Businesses Should Focus Products And Services On The Customer Segment They Have Chosen To Achieve Their Business Strategy

- A Business Should Choose Their Customer Segments Based Upon:
  - Their Capability To Serve Customer Needs Profitably Today
  - Businesses' Desire To Develop The Capability To Serve Different Customer Needs Profitably Tomorrow

Example: In a Sandwich Shop in Palmetto, Florida, A Conversation with an Air Conditioning Repairman -

"What Do You Think About Company Y's Air Conditioners?"
"They're too complicated, they put a fancy circuit board in where Brand X just puts a starting relay. Also, their coils are aluminum. They pit real quick *here in Florida* and I have a hard time fixin' them."

Types of Current Customer Voices

- Complaints
- Hang-ups
- Product Returns
- Contract Cancellation
- Loss Of Market Share
- Customer Defections
- Loss Of Potential Repeat Business
- High Rejects On Sales Calls
- Behavior

Potential "Locations" For Gathering VOC

Customer Research – Formal collection of information from customers about their "wants."
Examples: Market studies, behavioral observations

Transactions – All interactions between the business and paying customers that are the direct result of what the customer has paid the business to provide.
Examples: Installation of equipment, service calls

Casual Contacts – Unplanned contacts between customers and the business that yield information relevant to the continued satisfaction of wants, needs and unperceived delighters.
Examples: Meeting accidentally or socially with a customer

Inbound Communications – Customers contacting the business for information or requests for assistance that fall outside what the customer has paid the business to provide.
Examples: A customer calls in about a product, or visits the company's web site

Outbound Communications – The business contacting customers for information, assistance or to offer a product/service that falls out of products/services already sold/contracted out.
Examples: Official sales call, customer satisfaction survey

Internal Intelligence – Knowledge of customers needs, competitors and/or the market environment that exists within individual members of the business that can be systematically captured.
Examples: "Mining" sales staff information, competitive benchmarking, industry research, and the reservoir of data collected over time because of business contacts

Step 2. Listening to Customers

<table>
<tr><td colspan="2">Purpose:
• To identify approach for gathering information from targeted customers.
• To develop an action plan for implementing listening activities.</td></tr>
<tr><td colspan="2">Key Activities and Sub-Activities</td></tr>
<tr><td>Evaluate and select data collection methods.</td><td>• Review alternatives for gathering information.
• Select appropriate tool combinations.
• Identify needed resources.</td></tr>
<tr><td>Design tools and schedule data collection activities.</td><td>• Turn knowledge needs into measurable device by crafting questions.
• Develop action plan for gathering and tracking data.</td></tr>
<tr><td>Collect data.</td><td>• Implement action plan.
• Monitor progress.
• Adjust plan as needed.</td></tr>
<tr><td colspan="2">Supporting Tasks</td></tr>
<tr><td colspan="2">• Review listening tools and devices used previously (including best practices).
• Review the findings of these tools and devices.
• Identify population and sample size of customers and others as required.
• Select external and /or internal resources for crafting, overseeing, and conduction listening activities.
• Assess and/or set up data information tracking system.
• Determine responsibilities and accountabilities for developing and implementing listening plan.</td></tr>
<tr><td colspan="2">Possible Tools and Resources</td></tr>
<tr><td>Tools:
• Interviews
• Focus Groups
• Surveys
• Conjoint Analysis
• Kano Analysis
• VOC Data Collection Plan.</td><td>Resources:
• Elemental Survey Sampling, Scheaffer, Mendenhall, Ott
• Marketing Research, Lehmann, Gupta, Steckel
• Hearing the Voice of the Market, Vincent Barrabba, Gerald Zaltman
• Marketing Research, An Applied Approach, Thomas Kinnear and James R. Taylor
• Memory Jogger</td></tr>
</table>

Effective Customer Listening Skills/Process:

- Develop the Research Questions
- Select The Appropriate Listening Tools
- Evaluate The Application Of These Tools
- Select Sample Size
- Build A Specific Data Collection Plan

Listening Tools:

Active Methods ("Asking" the Customer):

- Interviewing
- Focus Groups
- Surveys
- Lost Customer Analysis

Passive Methods ("Listening" to the Customer):

- Complaints
- Listening Posts (Sales, Service)
- Warranty Claims
- Customer Observation
- Be a Customer

Interviewing

Purpose: To identify problems, needs & requirements from process customers, useful when "little" is known about customer needs. For internal customer/supplier relationships, the interview can also help break down functional barriers.

Process:

1. Determine The Knowledge Needed And Question Areas
2. Determine The Objective Of The Interview
3. Determine The Number Of Customers to Interview
4. Determine The Number Of Interviews
5. Draft Interview Guide
6. Determine The Time Limit For Interviewers
7. Test Interview Guide
8. Train Interviewees
9. Finalize The Interview Guide And Gain Approval
10. Schedule Appointments
11. Conduct Interviews
12. Analyze Interview Results

Analysis:

"Two-Way" Analysis (Within a Given Interview, Across Interviews), Extracting Verbatim Comments

Interview Types

Individual

- Individual Unique Perspectives
- Senior Level
- Large Volume Customer

Group

- Group Similar Products And Services
- Mid-To Lower-Level
- One Organization

Telephone

- Telephone Customers Are Widely Dispersed
- Basic Or Simple Issues

Example: "Wing-to-Wing" Customer Interview Guide

Introduction

- Objectives of session:
- Better understand your entire business process so we can explore ways we can bring more value to you (e.g., save time/costs, additional profits or productivity improvements)
- Provide definition and example of "Wing to Wing" concept and how it fits in with our Quality initiative (getting our focus from customers)
- Provide (or ask for) wing-to-wing example for your business/industry

1. What is most important to your customers? What key areas do you focus on/measure for your customers? (Provide examples, if necessary. Push them "out of the box" to think about things analogous to the to/from wing cycle time, not just the wing repair cycle time.

2. What part of your process affects the outcome (end product or service), yet you have limited direct control over? For example, where do you rely on sub-contractors or third parties? (May need to have them think about one division or product line)

3. When you decide to acquire equipment or services, when/how does your process start? (It may be helpful to create a simple process flow)

4. What do you consider to be the beginning of our relationship with you? What do you consider to be the end point? Should it be different that what you described?

5. *When selecting an equipment-financing source, what is most important to you? (If not mentioned, ask:)*

6. *How important is it for the source to meet promised dates for credit decision? How about meeting promised dates for funding? When does the clock start and stop?*

7. Where can we improve to provide better focus on <u>your</u> processes and operating practices? Think about the existing parts of the process where we interact now AND areas where we currently do not interact but you feel we may be able to help.

8. What company do you consider "best in class" for meeting your needs? Why? Where do they start in your capital management/equipment acquisition process?

9. Would an E-commerce application (e.g., apply online for financing, equipment takedowns, account maintenance) benefit your business and your relationship with a financial services provider? [Probe for details]

10. In what ways might we provide additional value beyond what we typically provide today? How can we help you make more money or better service your customers? Please think broadly.

11. What other companies do you think might be able to help with parts of this broader effort/process to serve your customers?

Closing - Thank you very much for your time and input.

Example: Loyalty/Disloyalty Interview Guide

INTRODUCTION
- My name, department, role.
- This is not a sales call; it is an opportunity to learn how to meet your needs and expectations.
- Your comments are confidential; you are part of a representative group of customers.

How long have you been working with E-LEASE.COM? How did you first encounter them?

In the spirit of improvement, what could E-LEASE.COM do better or what are you dissatisfied with?

Have you used other financing companies in 1996?

If no…What does E-LEASE.COM do well? (label LOYAL).

If yes…Which companies?

If multiple…How was your financing divided among these companies?

Were these leases or loans? What was the term?

What actions did the company take that led to your purchase decision? Why were they attractive?

Why did you not choose or consider E-LEASE.COM in 1996? Which factors led you to not using or decreasing your share with E-LEASE.COM?

Rank each factor - Give a percentage weighting based on the impact it had on your decision not to use E-LEASE.COM.

What actions could E-LEASE.COM have taken that would have resulted in being selected (or gotten more of your business)?

If E-LEASE.COM had taken the associated actions, would they have kept their account share?

I'd like to verify that your total year's equipment financing was $xxx.

CLOSE
- Thank you for your time providing this valuable information. We are committed to providing the best products and services in the industry.
- May I have the option to call back briefly if there is need for clarification?

Interviewing Techniques

Customers "talk" in many ways. To encourage customers to talk at a deeper level, there are basic and advanced techniques that an interviewer can use. Some of these techniques are listed below.

Techniques:	**Why Used:**
1. Start with General Questions.	To provide warm-up. To start the conversation where the customer is. To allow the customer to learn from their own words.
2. Listening actively by asking "the five whats" in a row: (e.g., "What does that mean?" "What does that mean?"…)	To push at customers' responses when they say things like, "I expect great service." To get a deeper response.
3. Listen for tensions by setting up questions with two points of view (e.g., "some customers we've talked with like type A, others seem to like type B better. What do you make of this?")	To see how people resolve tensions. To get beyond pat answers.
4. Avoid putting customers on the stops with questions like "Why do you think that?"	To prevent customers from giving you easy answers or saying what they might like they should say.
5. Always listen for opportunities to get more thinking from the customer. At these times, ask questions like, "Tell me more about that," or in some cases, "What do you make of that."	To provide the customer an opportunity to weave together additional thoughts and to give you a deeper sense of what he/she really feels and wants.

Interviews – Advantages & Disadvantages

Advantages

- Flexibility: Able To Obtain More Detailed Explanations
- Greater Complexity: Able To Administer Highly – Complex Questionnaires/ Surveys
- Able To Reach All Population Types; Able To Interview Populations That Are Difficult Or Impossible To Reach By Other Methods
- High Response Rate: Degree To Which Survey Is Fully Completed Is Higher
- Assurance That Instructions Are Followed

Disadvantages

- High Cost: Process Of Administering Is Costly
- Interviewer Bias: The Least Reliable Form Of Data Collection – The Interviewer Will Most Likely Influence The Responses To The Questionnaire
- Less Anonymity
- Personal Safety
- Limit To 15-20 Minutes (Business-To-Business 45-50 Minutes)
- Difficult To Analyze
- Positive Response Bias (People Give Higher Ratings In Personal Interviews)

Customer Focus Groups

Definition & Purpose: A carefully planned discussion designed to obtain perceptions on a defined area of interest in a non-threatening environment. Focus groups gather detailed information about relatively few topics and are useful when some preliminary needs information is available, and to test concepts and get feedback. The Focus Group is an exploratory research method used to help companies gain a deeper understanding of their customers' and prospects' perceptions, feelings, motivations and desires.

The most frequent applications of Focus Groups are:

- Exploring opinions, attitudes and preferences about products/services and the purchase and use of those products/services.
- Understanding consumer emotions regarding purchase decisions
- Searching for questions, vocabulary, and perceptions of buyers, and users of a product category.
- Analyzing target consumer reaction to copy and advertising methods.
- Exploring customer/prospect reactions to new product/service concepts.
- Formulating hypotheses that can be tested with quantitative surveys.

Benefits of Focus Groups

- Allow respondents to express detailed feelings, opinions and attitudes.
- It is possible to ask "What if..." type questions.
- Discover hidden feelings and motives.
- Focus Groups are economical
- It is possible to use visual or audio props in a Focus Group study
- Participants give immediate reactions.
- Learn what the consumers "out there" really think.
- Offers the client immediate contact with current and prospective customers.
- Focus Groups can help guide marketers in providing better goods and services.

Use Focus Groups When:

- Insights Are Needed In Exploratory Or Preliminary Studies
- A Communication Gap Exists Between Groups Of People
- Insight Is Needed Into Complicated Topics Where Opinions And Attitudes Are Conditional

- Synergy Among Individuals Is Needed To Create Ideas
- Information Is Needed To Prepare For A Large Scale Study
- A **High Value** Is Placed On Capturing Open-Ended Comments From The Customer Segments

Do Not Use Focus Groups When:

- The Environment Is Emotionally Charged And More Information Of Any Type Is Likely To Intensify The Conflict
- Quantitative (Measure) Statistical Projections Are Needed
- The Confidentiality Of Sensitive Information Can Not Be Ensured
- You Are Trying to Sell Products

Process:

1. Determine focus group(s) Purpose
2. Identify Topics, Develop Question Guide
3. Determine Who Will Moderate, Secure Required Facilities
4. Determine Group Characteristics, Select & Invite Representatives
5. Conduct focus group(s), Gather Data
6. Analyze Results

Focus Groups bring eight to twelve people together for a round table discussion lasting from one to two hours. Qualified participants are typically recruited by telephone - offering an incentive to attend each group. Participants can be qualified by specific segments or drawn by random selection to match targeted demographics.

Prior to holding a group, discussion topics are developed using an outline that is prepared in consultation with the client. Often participants of Focus Groups are asked to fill out a questionnaire relating to the main topic. Focus Group sessions will last from 2 to 4 hours and are usually recorded by video and audiotape for further analysis (often, the customers of the focus group information will observe participants through a one-way mirror and may interact with the participants through the moderator).

Typically, a minimum of two sessions with different groups will be conducted. This not only ensures confidence and eliminates bias but also provides more valuable information than a single session by allowing comparisons between groups. Typically, customers are compared with prospects, although more specific sampling is possible.

Focus Groups have advantages over other research methods. One advantage is that focus groups allow respondents to express detailed opinions - usually telephone or mail surveys limit respondents to responses that can be expressed only in a few words.

With Focus Groups it is possible to ask "What if..." type questions. In the give and take of a lively discussion, people can raise questions that the researcher had not thought of and might never have raised during a few individual interviews. It is possible to use visual and/or audio props in a focus group study - participants in the focus group can be exposed to examples of advertising and give immediate reactions.

Finally, Focus Groups give the client a chance to learn what their customers and prospects really think. No other research method offers the client this immediate contact with current and prospective customers.

Analysis:

1. Extracting Verbatims
2. Needs Prioritization (not statistical)

Typical Focus Group Costs:

Recruitment - (the process of getting participants to come to a session)
Ranges from $75-$150 per recruit based on difficulty of recruit.
i.e.: Professionals=$150, Students=$75.

Incentives for participants- *(paying participants for their time)*
Similar to recruitment - average cost= $75/participant
Again, this will vary according to difficulty of recruit.
i.e. Professionals =$100-$150, Consumers =$60-$75

Facility & refreshments
Average Focus Group facility rental is $500/group
Participant meal average cost is $100. Client meal average cost is $25 per person

Moderation / Consultation / Management
$1,150/group - includes top-line summary report

Videotaping - $150/group

Verbatim Transcript - $225/group

Reporting - $150/hour – 4-hour minimum.

Average cost range per Focus Group session $4,500 - $6,500

Tips for making the most of a Focus Group project

- Arrange for key management to attend and actively use the information.
- Involve key management in deciding objectives and topics for group discussions.
- Segment and qualify group participants.
- Offer a proper incentive for participants to attend and verify recruiting at least three times.
- Use pre-discussion questionnaires to gather basic information such as demographics and other non-discussion questions
- Use props and audio-visual aids when possible.
- Use written exercises within the groups to break up the pace and capture unbiased preferential information.
- Check audio and video clarity during course of discussion.
- Make sure there are plenty of refreshments for both the participants and the clients. M&M's should always be plentiful in focus groups.

Customer Surveys

Purpose: To quantify or prioritize customer feedback or needs, to measure change or causality, to develop statistically valid information, to collect efficiently a considerable amount of information from a large population.

Process:

1. Identify the information needed and develop the questions to provide that information
2. Look at all previous surveys, focus group findings, etc.
3. Identify the total number of customers or prospects (the population).
4. Identify the subgroups or stratifiers needed.
5. Determine sample size and approach (e.g., random sample or census *see Section 9.3 – Sampling Theory*).
6. Determine if samples are identified or unidentified.
7. Determine if responses will be attributed (confidentiality).
8. Draft the questions (will require several iterations) and get input from team and other key stakeholders. Don't ask for information you already have (e.g., annual revenue).
9. Pilot the questionnaire internally to test the flow and the timing. Revise based on feedback.
10. Set up a survey "red alert" process to capture leads and/or service issues.
11. Train the interviewers so they are comfortable with the flow of questions and the customer's terminology.
12. Consider offering a token gift of appreciation (of nominal value) for the respondent's time.
13. Send a thank you letter that could include major findings and high-level actions.

Analysis:

Means, Standard Deviations, Proportions, Confidence Intervals
Cross-Tabulation, Regression, Prioritization, Utility Measures (See Section 9.3 for Sampling Methods and Analysis)

Survey Development Questions

1. ***What are your objectives?***
   - Problem solution/action planning
   - Cultural change
   - Communication event
   - Management information
   - Customer feedback
   - Individual assessment
   - Group differences
   - Team improvement
   - Program assessment
   - Union avoidance
   - Loss assessment

2. ***Is a survey the best method?***
   - Existing data bases
   - Behavioral indicators
   - Focus groups
   - Individual interviews
   - Outside sources

3. ***What types of people do you want to survey?***
   - Job types
   - Demographic groups
   - Union/non-union
   - Functions
   - Family members
   - Customers

4. ***How do you want to group the results?***
   - Organizations
   - Geography
   - Demographics
   - Customer Segments

5. ***How accurate do you want the results to be?***
   - Accuracy of smallest group drives sample size
   - Census: best accuracy; big feedback commitment
   - Sample: OK for large groups; “left out” syndrome
   - Stratified Sample: can target accuracy; complicated summaries

6. ***What do you want to ask?***
   - Decide on topics (management, experts, focus groups)
   - Beg, borrow, invent items (single-concept, 5 or 7 point balanced scales)
   - Pretest items with small group of people
   - Pilot test items (item analysis)
   - Write-ins: are you going to read them all?

7. ***How are you going to analyze the information?***
   - Inside resources ($2 per person)
   - Outside resources ($5-$25 per person)
   - Telling a good result from a bad result
     -absolute 60/40 rule; 30+ corollary
     -group comparisons (statistical/rule of thumb)
     -external comparisons
     -history
     -model testing
   - Multivariate methods demand huge samples
   - Don’t over analyze

Continued on Next Page . . .

8. How do you want to collect the information?
- Paper (mail? /group? /voting day?)
- "Rapid response"
- Telephone
- Touch-tone entry
- On-line server or main frame

9. How do you get a good return rate?
- Anonymity
- Minimize demographics
- "Objective" processor
- Credible sponsor
- Reminders
- Group administration
- Incentives

10. Who is going to be the sponsor?
- Senior business executive
- Outside researcher

11. Who will get the information? When?
- No surprises!
- No surprises!
- Most immediate owner gets it first; bubble up
- Don't delay in the name of micro-analysis

12. What are you going to feed back?
- Narrative only
- Representative items
- Total disclosure
- Thematic analysis of comment

13. How are you going to feed back the results?
- Managers (training?)
- HR role
- All employee meetings
- Written summary

14. How will participants be involved?
- Development
- Feedback
- Action-planning
- Follow-up
- Communication

15. Do you plan to do this again?
- Don't ask again until improvements have occurred!

16. What are the risks?
- Expectations
- Legal
- Report card
- Surprises

17. What are the mega-issues?
- Link between employee and customer satisfaction
- Role of Report Card
- Use of 360º's
- Surveys as evidence
- Too many surveys

Guidelines for Writing Survey Questions

Types of Questions

There are several common question types that are used on questionnaires. The first three types, *Yes/No, Multiple Choice* and *Ranking* are closed-ended questions. The response is limited to pre-determined choices. The *Fill-in-the-Blank* is an open-ended question - any response offered by the participant is captured.

Yes/No - These are the "black and white" questions of the survey and are useful where the expected response is "binomial" - only one of two choices. Analysis of responses is usually limited to calculating a fraction or percentage.

If the survey is repeated over time, changes in the fraction or percentage may be noted on a run or control (*p* or *np*) chart. If two populations are surveyed with the same question, hypothesis testing may be used to detect significant differences in the populations' response.

Multiple Choice - These are used where the response to the question can be predicted to fall within one of several pre-defined answers. Even so, many multiple-choice questions will include a "Don't Know" or "Not Applicable" answer as a possible choice. Again, analysis of responses will consist of the fraction or percentage response for each category. A Pie Chart of the results may be a good display of the results.

Multiple-Choice may be used to elicit some sensitive information. Personal income is often determined by offering respondents a multiple-choice question that places their response into a range ($0 - $10,000, $10,001 - $20,000, etc.).

Ranking - These questions are often used to obtain judgments from respondents. The responses may be "word-based" (Excellent, Good, Average, Poor) or numerical (Likert Scale 1 - 5, or 1 - 7).

When using a numerical scale, the issue of resolution must be addressed. If the range of possible responses is too small (i.e. 1 to 3), the results may not be useful, especially if several alternatives are being compared through ranking questions (What's your opinion of Choice A, B, C, etc.). On the other hand, if the range of responses is too large (i.e. 1 to 19), respondents may have difficulty distinguishing the "shades of gray" offered by the question. The most effective ranges for ranking questions seem to be either 1 to 5 or 1 to 7.

Be careful of building a pattern of *expected responses* into your questionnaire and then reversing the pattern. For example, you include a series of ranking questions where 1 = Poor and 5 = Excellent. The questionnaire has thus built up

a pattern of expected response from the participants. If you suddenly reverse the ranking (1 = Excellent and 5 = Poor), the respondent may not notice the change and response errors can occur.

Bar Charts, Histograms or Frequency Charts of the responses can be constructed. Means and Standard Deviations can be calculated (for "word-based" questions, the responses must be equated to numbers - e.g. Excellent = 100, Good = 66, Average = 33, Poor = 0). Hypothesis tests can be performed to identify significant differences between alternatives or populations' responses.

Fill-in-the-Blank Questions - These are questions designed to obtain verbal or written responses from participants and are most useful where the responses cannot be predicted or pre-determined. Fill-in-the-Blank questions should be used sparingly, especially in a long questionnaire, as they lengthen the time required to complete the survey and contribute to respondent fatigue. Responses from these questions may be recorded on index cards and affinitized to look for patterns.

One difficulty with these questions is non-response, especially to "generic" fill-in-the-blank questions such as "Additional Comments?" Even if the respondent has comments, they may not take the time to articulate and record them on the questionnaire. More specific questions may help elicit a response.

Writing and Sequencing the Questions

Writing the questions is a critical step because the results of the survey depend on the answers given to each question. The question wording must be clear and comprehensible to most respondents to minimize biasing of the survey results.

In addition to writing the questions, the designer must sequence them in a natural order that will flow smoothly from one topic to another. For example, a survey administered to hospital patients may be sequenced in the chronological order of the patient's stay - admission, initial nursing care, post-surgical care (including nutrition, patient education, etc.) and finally discharge/follow-up.

The flow may be improved by using *screening questions* and *skip patterns.* For example, screening questions may be used to determine if respondents have children before leading them through a series of child-related questions. These are used when the survey is administered verbally, i.e. a telephone survey. Skip patterns are used in the same way, except for written surveys. The response to a "skip question" determines if the respondent completes the next set of questions or "skips" to the next topic.

Sources Of Survey Bias Or Error:

- Leading Questions – suggests a preferred response
- Unclear Concepts, Acronyms, Definitions And Vocabulary
- Ambiguous Wording – terms which may have multiple meanings
- Double-Barreled Questions – two questions combined in one
- Loaded or Sensitive Questions – responses may be biased due to emotionalism or fear of embarrassment (these may be placed later in the survey after some comfort/trust has been established)
- Over Specificity/Over Generalization – too detailed (or broad) information requested; often beyond ability of respondent to remember
- Questionnaire Sequence And Flow/Previous Question Context – response to question influenced by previous question
- Non-Random Sample – see Unit 9.3 for random sampling methods
- Non-Response Error – fraction of surveys returned by respondents; what do you know (or don't know) about those that didn't respond
- Expected Response Pattern – sequence of questions where "high" value is expected; with one "low" response embedded in the sequence.

Examples of poorly worded questions appear in *ITALICS* and good questions appear in **BOLD.**

1. Questions should not be biased or leading.

   Would you agree that the company has an excellent product range?
   Would you say the product range is:

Excellent	**1**
Very good	**2**
Good	**3**
Fair	**4**
Poor	**5**

2. Avoid jargon, acronyms and uncommon words; use language of your customer.

   GNP, ROI, salient, etc.

3. Avoid ambiguous words.

Usually, frequently, etc.
Do you frequently use the Internet?
How frequently do you use the Internet?

4. Questions should be as short as possible. Remove any unnecessary words.

5. Do not build 2 questions into one (called “double-barreled questions”).

   Do you regularly use AI or does your manager?
   Do you regularly use AI? Does your manager use AI?

6. Questions should be very specific.

   Did you use AI last year?
   Did you use AI in the last 12 months?

7. Keep the number of meaningful words to a minimum.

   What motivates and inspires you in the selection or specification of a new supplier?
   What, above else, influences your choice of a new supplier?

8. Do not use words that could be misheard or spell them out.

   15, I mean one five
   50, five oh

9. Desensitize questions by using response bands.

   I will read a number of revenue size bands. Would you tell me which your company fits into?

-	**Less than $1MM**	**1**
-	**Between $1MM and $10MM**	**2**
-	**Between $11 and $20MM**	**3**
-	**Over $20MM**	**4**

10. Allow for “other” responses in fixed response questions.

11. Consider “softening” knowledge-based questions.

Do you know the number of calls you have received last February?
Do you happen to know …

12. Be careful of interactions between questions, due to sequencing within the survey.

Poor Sequence (2nd question's response is affected by 1st):
What do you think about the streetlights in the area?
How safe do you think the neighborhood is?

13. Keep questions within abilities of people's memory, knowledge or skill.

What did you have for dinner on the first Tuesday of last month?

Survey Delivery Methods

	Survey Delivery Method					
Survey Characteristic	**Mail**	**Phone Interview**	**Phone Automated Call Back**	**Interviewer Administrated**	**Group Sessions – Written**	**Internet (e.g., SurveyMonkey)**
Data Collection Costs	Low	Medium	Medium	High	Medium	Low
Time Required to Collect	High	Low	Medium	High	Medium	Low
Response	Low	High	Medium	High	High	High
Interviewer Bias	None	Medium	None	High	Low	None
Acceptable Length of Survey	Long	Short (Max. 15 minutes)	Medium	Medium to Long	Long	Short (5 – 10 min.)
Ability to Obtain Open Ended Questions	Low	Low	Low	High	High	Medium
Perceived Anonymity	High	Low	High	Low	Medium	None

Example Survey for Financial Services Customers

A major credit card services company employs the following survey. Its purpose is to determine customers' satisfaction with the service processes (e.g. when a customer calls to request some service such as questioning a charge, or to increase a credit limit).

CONSUMER ACCOUNT HOLDERS

Screener/Questionnaire
May I please speak with Mr./Ms. _____?

IF ASKED WHO'S CALLING, SAY:
Hello, my name is _____ calling on behalf of Credit Cards 'R' Us.

ONCE PERSON IS ON THE PHONE, SAY:
Hello, my name is _____. I am conducting a customer satisfaction survey on behalf of Credit Cards 'R' Us. I am calling regarding your recent call about your Credit Cards 'R' Us Credit Card. We would like to ask you a few questions about your experience with the Credit Service Center so that we can improve the quality of our customer service to accountholders.

ONLY IF ASKED, SAY: These questions will take about ten minutes of your time.
S1. Do you or does anyone in your household work for Credit Cards 'R' Us or a market research firm?
1 YES — THANK AND TERMINATE
2 NO — CONTINUE
97 DON'T KNOW — THANK AND TERMINATE
98 REFUSED — THANK AND TERMINATE

S2. Our records show that you called the Credit Cards 'R' Us Credit Service Center about your Credit Cards 'R' Us credit card on _____ [STATE DATE OF CALL]. Is that correct?
1 YES — CONTINUE TO QS2d
2 NO — CONTINUE TO QS2a
97 DON'T KNOW — CONTINUE TO QS2a
98 REFUSED — THANK AND TERMINATE

S2a. Is there anyone else in your household who has a Credit Cards 'R' Us credit card who might have called the Credit Cards 'R' Us Credit Service Center?
1 YES — CONTINUE TO QS2b
2 NO — THANK AND TERMINATE
7 DON'T KNOW — THANK AND TERMINATE
8 REFUSED — THANK AND TERMINATE

S2b. May I have his/her name and may I speak to him/her?
1 YES — RECORD NAME AND GO BACK TO INTRO
2 NOT NOW — CONTINUE TO QS2c
3 NO — THANK AND TERMINATE
7 DON'T KNOW — THANK AND TERMINATE
8 REFUSED — THANK AND TERMINATE

S2c. When would he/she be available?

RECORD TIME & DAY, SCHEDULE CALLBACK FOR TODAY. IF NOT TODAY, THANK AND TERMINATE

S2d. AFTER CONFIRMING THAT WE HAVE THE CORRECT PERSON ON THE PHONE, PLEASE ENTER NAME HERE _____

ASK S3 FOR EVERYONE

S3. Did you speak with a customer service representative?
1 YES — CONTINUE
2 NO — TERMINATE
97 DON'T KNOW — TERMINATE
98 REFUSED — TERMINATE

Main Questionnaire

Now I would like to ask you some questions concerning the service you received when you called the Credit Cards 'R' Us Credit Service Center. Please think about your experience with the Credit Cards 'R' Us credit card and not the in-store service.

Q1. When you called the customer service center about your Credit Cards 'R' Us credit card in the past week, what was the purpose of the call? [PROBE: Are there any other reasons?][DO NOT READ LIST]
1 STATEMENT/ BILLING QUESTIONS
2 QUESTION/PROBLEM ABOUT A PURCHASE/ RETURN
3 PAYMENT DELAYS/ PAYMENT ISSUES/ TO CONFIRM PAYMENT
4 BALANCE INQUIRY/ CREDIT LINE INQUIRY/ PAYMENT DATE
5 QUESTION/ PROBLEM ABOUT INTEREST CHARGES/ LATE CHARGES
6 GENERAL INFORMATION REQUEST/ ACCOUNT CHANGE/ CANCEL CARD
7 TO GET CREDIT LINE INCREASE/ CORRECTIONS
8 CHANGE OF ADDRESS/ NAME
9 QUESTIONS/ PROBLEM ABOUT MAIL ORDER/ DELIVERY/ TO PLACE AN ORDER
10 CREDIT/ CASH ADVANCE/ PURCHASE DENIED/ CARD WOULDN'T WORK
11 TO GET A NEW CARD/ OPEN A NEW ACCOUNT

12 DIDN'T RECEIVE CARD/ RECEIVED WRONG CARD
13 LOST/ STOLEN CARD
14 NOT CREDITING ACCOUNT CORRECTLY
15 QUESTION/ PROBLEM ABOUT PROMOTION/ OFFER/ COUPON
16 CHANGE/ ADD AUTHORIZATION OF ACCOUNTS
17 TAX EXEMPTION/ INCORRECT TAXING
96 OTHER [PLEASE SPECIFY:_____]
97 DON'T KNOW
98 REFUSED

Q2. Thinking of your recent telephone call to the Credit Cards 'R' Us Credit Service Center, on a scale from 1 to 10 where a 10 means you are extremely satisfied and a 1 means you are not at all satisfied, how would you rate your overall satisfaction with the customer service you received?

Not at all satisfied								Extremely satisfied		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

Q3. Why did you rate your call experience that way? [PROBE FOR THE MAIN REASON]

Q4. Next, I would like you to comment specifically on the credit representative that handled your call by answering yes, no, or does not apply to the following. Did the representative . . . ? [READ LIST]

1 YES
2 NO
3 DOESN'T APPLY
7 DON'T KNOW
8 REFUSED

[ROTATE ITEMS]

	BCE	Guidelines
a. Introduce herself or himself to you	2	Greeting
b. Apologize for mistakes if there were any and take ownership of them by saying something like, "I'm sorry that we made this mistake"		Reflect
c. Present you with options to solve your concerns or questions	7	Explore Options
d. Recap the decision made and check for your understanding	4	Get Agreement
e. Set expectations during your call; for example, did he/she tell you when you would receive an answer		Realistic Expectations
f. Offer additional assistance, for example by asking, "Is there anything else we can do for you today?"		Check for Satisfaction
g. Use your name multiple times during the discussion	2	Tone & Technique
h. Thank you for calling		Close

Q5. How would you rate the quality of your Credit Cards 'R' Us credit representative in the following areas? Please use a 1 to 10 scale where a 10 means the representative is excellent in that area and a 1 means the representative is unacceptable in that area. [READ ITEM.]

ROTATE ITEMS

		BCE	Guidelines
a.	Used a friendly, warm, and caring tone	1	Tone & Technique
b.	Listened carefully to you about your questions or concerns	3	Listen
c.	Asked questions to better understand your concern		Ask
d.	Gave explanations that were easy to understand	4	Tone & Technique
e.	Had the knowledge to help resolve your questions or concerns	5	Explore Options
f.	Took responsibility for your questions or concerns	6	Explore Options
g.	Gave you the feeling that you are valued as a customer	8	Close
h.	Responded with empathy and a genuinely caring attitude	3	Reflect

Unacceptable								Excellent		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

Q6a. Were you put on hold by a customer service representative at any time during your call?
- 1 YES — CONTINUE TO Q6b
- 2 NO — SKIP TO Q7a
- 7 DON'T KNOW — SKIP TO Q7a
- 8 REFUSED — SKIP TO Q7a

IF Q6a=1, ASK Q6b, OTHERWISE SKIP TO Q7a

Q6b. Did the customer service representative ask your permission to put you on hold and wait for your approval?
- 1 YES
- 2 NO
- 7 DON'T KNOW
- 8 REFUSED

Q7a. Were you transferred by a customer service representative at any time during your call?
- 1 YES — CONTINUE TO Q7b AND Q7c
- 2 NO — SKIP TO Q8
- 7 DON'T KNOW — SKIP TO Q8
- 8 REFUSED — SKIPTO Q8

IF Q7a=1, ASK Q7b AND Q7c, OTHERWISE SKIP TO Q8

Q7b. Did the customer service representative ask your permission to transfer your call and wait for your approval?
1 YES
2 NO
7 DON'T KNOW
8 REFUSED

Q7c. Did the customer service representative hand off your call so that you did not have to repeat the reason for your call?
1 YES
2 NO
7 DON'T KNOW
8 REFUSED

Q8a. Was your concern or question resolved?
1 YES — CONTINUE TO Q8b
2 NO — SKIP TO Q8c
3 DOESN'T APPLY — SKIP TO Q9
7 DON'T KNOW — SKIP TO Q9
8 REFUSED — SKIP TO Q9

Q8b. On a scale from 1 to 10 where a 10 means you were extremely satisfied and a 1 means you were not at all satisfied, how satisfied were you with the resolution?

Not at all satisfied								Extremely satisfied		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

IF NO TO Q8a, ASK Q8c AND Q8e; IF 1 THROUGH 7 TO Q8b, ASK Q8d AND Q8e;
OTHERWISE SKIP TO Q9
Q8c. In what way was your concern or question not resolved? (SKIP TO Q8e)

Q8d. Why would you rate it that way? (CONTINUE TO Q8e)

Q8e. Would you like to be contacted by someone at Credit Cards 'R' Us Credit Services to discuss this?

ASK EVERYONE
Q9. As a result of your recent call to the Credit Cards 'R' Us Credit Service Center, would you say that your overall satisfaction with Credit Cards 'R' Us has . . . ? [READ CATEGORIES]
1 Increased
2 Decreased
3 Remained the same

97 DON'T KNOW
98 REFUSED

Q10. In terms of your recent experience with the Credit Cards 'R' Us Credit Service Center, please tell me how much you agree or disagree with the following statement. Please use a scale from 1 to 10 where a 10 means you strongly agree and a 1 means you strongly disagree.

"Credit Cards 'R' Us is delivering a caring customer service experience that exceeds my expectations." [READ CATEGORIES]

INTERVIEWER: READ ONLY IF NECESSARY: Please focus on your credit experience and not your experience with in-store service.

Strongly disagree								Strongly agree		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

Now I would like you to think about your relationship with Credit Cards 'R' Us' credit card.

Q11. Thinking of your Credit Cards 'R' Us credit card, using the same 1 to 10 scale where 10 means extremely satisfied and 1 means not at all satisfied, how would you rate your overall satisfaction?

INTERVIEWER: READ IF NECESSARY: Please focus on your credit experience and not your experience with in-store service.

Not at all satisfied								Extremely satisfied		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

Q11a. Overall, how would you rate Credit Cards 'R' Us' service in opening your Credit Cards 'R' Us credit card? Please use a scale from 1 to 10 where 10 means it is excellent and 1 means it is unacceptable.

Unacceptable								Excellent		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

PROGRAMMER: IF RATING IS 1 THRU 7, ASK Q11b, OTHERWISE SKIP TO Q12

Q11b. What would Credit Cards 'R' Us have to do to receive a better rating? [PROBE FOR THE MAIN REASON]

[READ IF NECESSARY: Please think about your experience with the Credit Cards 'R' Us credit card and not the in-store service.]

Q12. How would you rate your overall experience using your Credit Cards 'R' Us credit card to make purchases [IF COMMERCIAL: for your company]?

IF NECESSARY, SAY: Please use a scale from 1 to 10 where 10 means it is excellent and 1 means it is unacceptable.

[READ IF NECESSARY: Please think about your experience with the Credit Cards 'R' Us credit card and not the in-store service.]

Unacceptable								Excellent		DON'T KNOW	REF-USED
1	2	3	4	5	6	7	8	9	10	97	98

[PROGRAMMER: IF RATING IS 1 THRU 7, ASK Q12a, OTHERWISE SKIP TO Q13

Q12a. What would Credit Cards 'R' Us have to do to receive a better rating? [PROBE FOR THE MAIN REASON]

[READ ONLY IF NECESSARY: Please think about your experience with the Credit Cards 'R' Us credit card and not the in-store service.]

Q13. Now thinking only about how Credit Cards 'R' Us handles your payments on your account, how would you rate the overall quality of service Credit Cards 'R' Us provides in handling your credit card payments? [READ CATEGORIES]

IF NECESSARY, SAY: Please use a scale from 1 to 10 where 10 means it is excellent and 1 means it is unacceptable.

Unacceptable								Excellent		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

PROGRAMMER: IF RATING IS 1 THRU 7, ASK Q13a, OTHERWISE SKIP TO Q14

Q13a. What would Credit Cards 'R' Us have to do to receive a better rating? [PROBE FOR THE MAIN REASON]

[READ ONLY IF NECESSARY: Please think about your experience with the Credit Cards 'R' Us credit card and not the in-store service.]

Q14. On a scale of 1 to 10 where 10 means you are extremely likely and 1 means you are not at all likely, how likely are you to recommend a Credit Cards 'R' Us credit card to your friends and family?

Unacceptable								Excellent		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

Q14a. Why do you give that rating?

Q15. How likely are you to continue to use your Credit Cards 'R' Us credit card?

IF NECESSARY, SAY: Please use the same 1 to 10 scale where a 10 means you are extremely likely and a 1 means you are not at all likely.

Unacceptable								Excellent		DON'T KNOW	REFUSED
1	2	3	4	5	6	7	8	9	10	97	98

Q16. How can Credit Cards 'R' Us improve the service on your credit card account? [PROBE: Anything else?]

READ ONLY IF NECESSARY: Please think about your experience with the Credit Cards 'R' Us credit card and not the in-store service.

NOTE: Q17a AND 17b, DELETED

Q18. Is there anything you would change about Credit Cards 'R' Us' credit program? [PROBE: Anything else?]
1 ENTER COMMENT
2 NO — CONTINUE TO Q19

Q19. Of the types of things that you typically can buy at Credit Cards 'R' Us, what percent of your office supply shopping do you do at Credit Cards 'R' Us?

INTERVIEWER, IF NECESSARY, SAY: For example, if you did half of your office supply shopping at Credit Cards 'R' Us and the other half at other locations, you would say that 50% of your purchasing is done at the Credit Cards 'R' Us.

Q20a. How often do you use your Credit Cards 'R' Us credit card?
1 Always – GO TO CLOSING
2 Sometimes – GO TO Q21
3 Rarely – GO TO Q21
7 DON'T KNOW – GO TO Q21
8 REFUSED – GO TO Q21

IF Q20a=2 OR 3, ASK Q20b

Q20b. Why do you (INSERT ANSWER FROM Q20a) use your Credit Cards 'R' Us credit card?

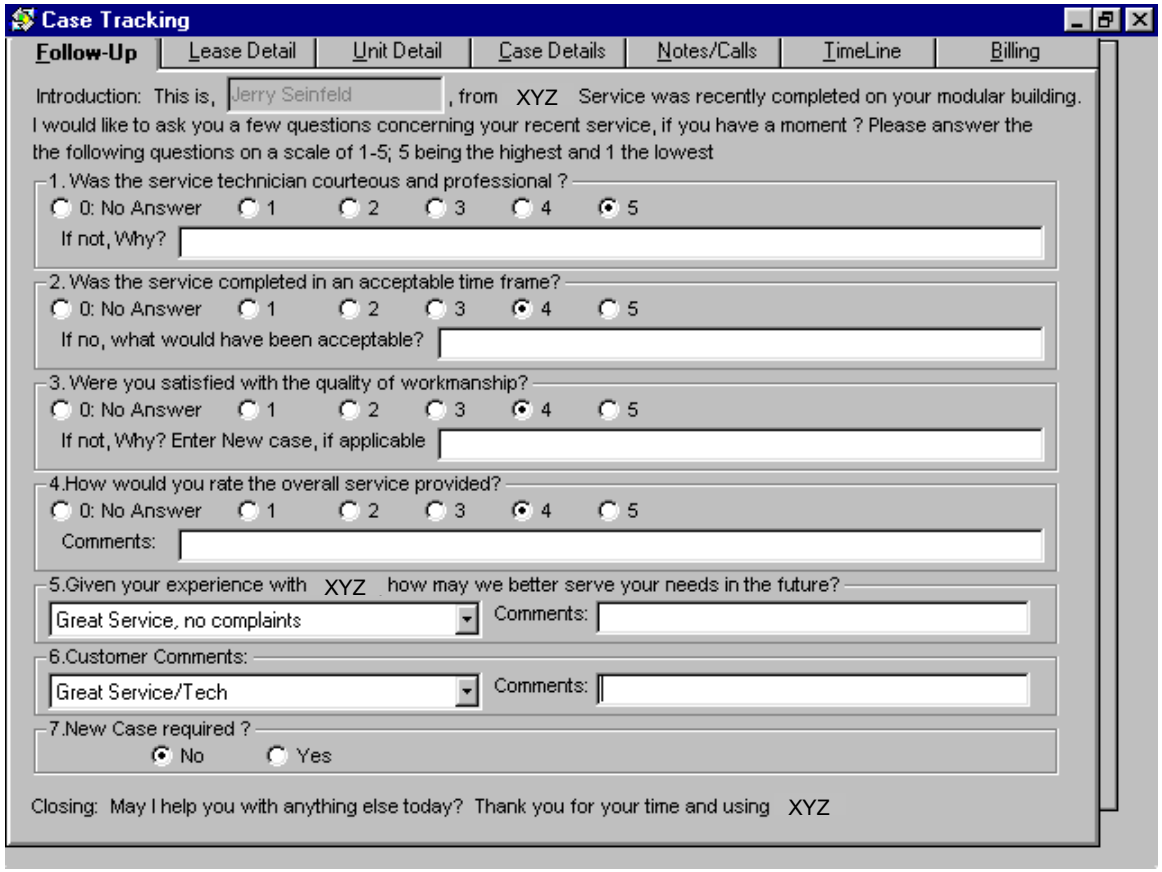
Q21.And for the purchases when the Credit Cards 'R' Us credit card is not used at the store, how do you usually pay for those purchases?
1 Cash
2 Check
3 Other credit card
5 Other (SPECIFY)
7 DON'T KNOW
8 REFUSED

CLOSING
Thank you for your time, Mr./Ms. _____, and have a pleasant afternoon/evening. Thanks again for participating!

Example: Post Service Survey - Modular Building Leasing

This Post-Service Survey is designed to compile customer feedback ratings collected by Customer Service Representatives utilizing a follow-up process for 20% (randomly selected) of all service cases once the service case has been closed. The customers are asked to answer 4 questions with a ranking 1-5 (Very Dissatisfied – Very Satisfied) regarding their service experience with the leasing company. The last two questions are open-ended and capture customer verbatim. The customer feedback data is input directly into the company's case management application (CIS – Customer Information System). The information is gathered to help them better understand the individual needs of their customers. A Voice of the Customer (VOC) module was designed to utilize the data collected. Any employee with access to the case management system (CIS) can query the feedback data to display graphical summaries or detailed reports of the customer feedback data specific to a date range and location (company, region and branch). In addition, the stored data is convertible into Excel for greater data analysis.

Follow up tab in case tracking screen:



Step 3. Organize and Analyze Data

Purpose:

- To identify Critical-to-Quality Characteristics/requirements (CTQs).
- To establish an ongoing system for measuring customer satisfaction related to business strategy.

Key Activities and Sub-Activities	
Organize data into logical requirements.	• Separate quantitative from qualitative data. • Arrange data into logical hierarchy.
Prioritize key findings.	• Separate out “satisfied” and “completely satisfied” requirements.
Determine Critical-to-Quality Characteristics & Indicators that can be measured on an ongoing basis.	• Review and validate key findings. • Correlate key findings to achieving business strategy (Find More, Win More, Keep More). • Identify CTQs and Indicators.

Supporting Tasks

- Review previously identified CTQs.
- Test findings for consistency and validity.
- Discuss potential CTQs with sales and operations.
- Begin assessing/testing potential opportunities for process improvement and product/service innovations.

Possible Tools and Resources	
Tools: • Affinity Diagram. • Structure Tree. • Critical-to-Quality Matrix. • QFD/House of Quality.	**Resources** See list of resources from step three.

Organizing and Prioritizing Customer Data

Customer Data Classification

Qualitative Data (Verbatims, Comments):

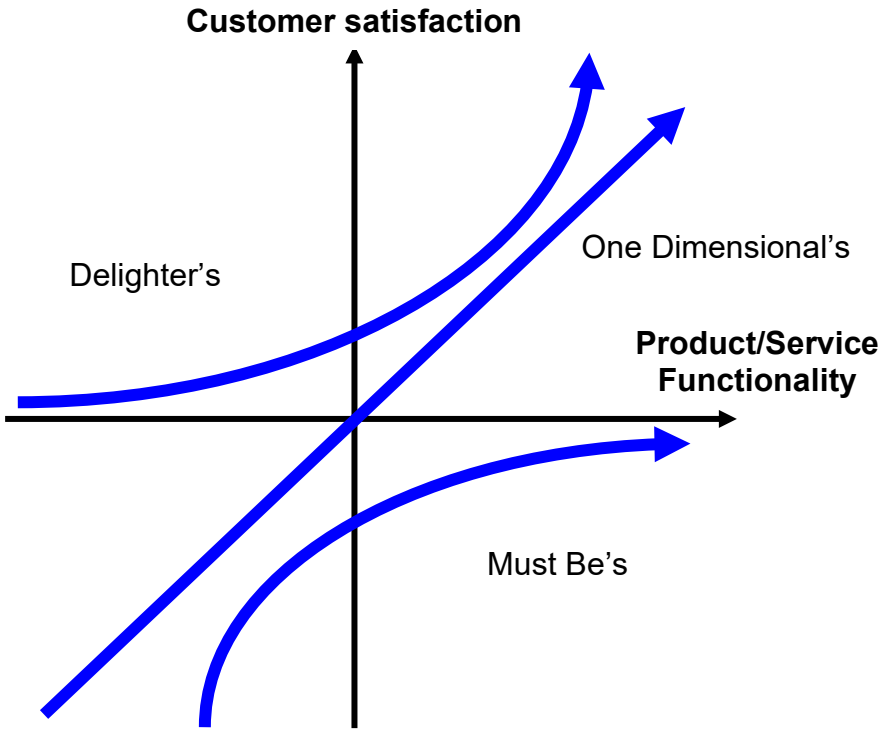
- Affinitize Data Into “Natural” Groupings
- Develop Structure or Hierarchy of Needs
  See Section 16.1 for Tool Use and Analysis

Quantitative Analysis:

- Prioritization (e.g. Surveys or Conjoint Analysis (product features))
- Means, Standard Deviations, Proportions, Confidence Intervals
- Correlation/Regression (*See Units 10.1, 10.2)*

Dr. Kano's Hierarchy of Needs

Dr. Noriaki Kano has developed both a structure of customer needs (also applied to product/service characteristics) and a means of determining into which category the needs fall. The graph below shows the reaction of a customer to meeting or not meeting their needs through the product or service.



Must Be Needs – Needs that have to be met. Needs associated with the basic functions of the product or service. An air conditioner must provide cooling. Often "safety" needs fall into this category. For example, "landing safely" is a must be need associated with airplane travel. If the need is met, the customer feels (at best) a neutral reaction – they expected the need to be met. However, if the need is not met, dissatisfaction occurs.

One Dimensional Needs – Needs that can be met in varying degrees, leading to more or less satisfaction with the product or service. Examples: Leg room in an airplane, efficiency of an air conditioner. Here, the customer can feel either positively satisfied if the need is met, or dissatisfied if the need is not met.

Delighted Needs – Needs that, if met, will delight the customer. Often the customer does not know what will delight them – this is the responsibility of the producer or provider to discover. Examples – Cup holders in automobiles, warm cookies on an airplane. If this need is not met, the customer will feel neutral, but if met, these needs lead to high satisfaction with the product or service.

Also Consider:

Indifferent – Issues that the customer doesn't care about. Examples – Colored cases on laptop computers, 90% of computer software functions.

Reverse – Issues to which the customer responds negatively. Examples – Waiters that introduce themselves and tell you their life stories to be "friendly." The paperclip "Help" feature that was once included in Microsoft Office.

How to Complete a Kano Analysis

This questionnaire includes pairs of multiple-choice questions about potential product capabilities. Half of each pair of questions asks how you would feel if the product included a particular capability; the other half of each pair asks how you would feel if the capability was not provided.

	I would be delighted to find it that way (1)	I expect it to be that way (2)	I am neutral (3)	I would not like it that way, but I can live with it that way (4)	It must not be that way (5)
1a. If you receive feedback that is more detectable from the device than you do today, how would you feel?					
1b. If you received less detectable feedback, how would you feel?					X
2a. If you have a greater field of vision of the surgical site than you have today, how would you feel?					
2b. If you have less, how would you feel				X	

Processing the Kano Survey

Customer Requirement Dysfunctional

1. I'd be delighted to find it that way.
2. I expect it to be that way.
3. I am neutral.
4. I would not like it that way, but I could live with it that way.
5. It must not be that way.

Customer Requirement Functional

1. I'd be delighted to find it that way.
2. I expect it to be that way.
3. I am neutral.
4. I would not like it that way, but I could live with it that way.
5. It must not be that way.

Two-Dimension Evaluation Table

Functional \ Dysfunctional	1. delighted	2. should-be	3. neutral	4. live with	5. must not be
1. delighted	Q	A	A	A	O
2. should-be	R	I	I	M	M
3. neutral	R	I	I	I	M
4. live with	R	I	I	I	M
5. must not be	R	R	R	R	Q

Key:

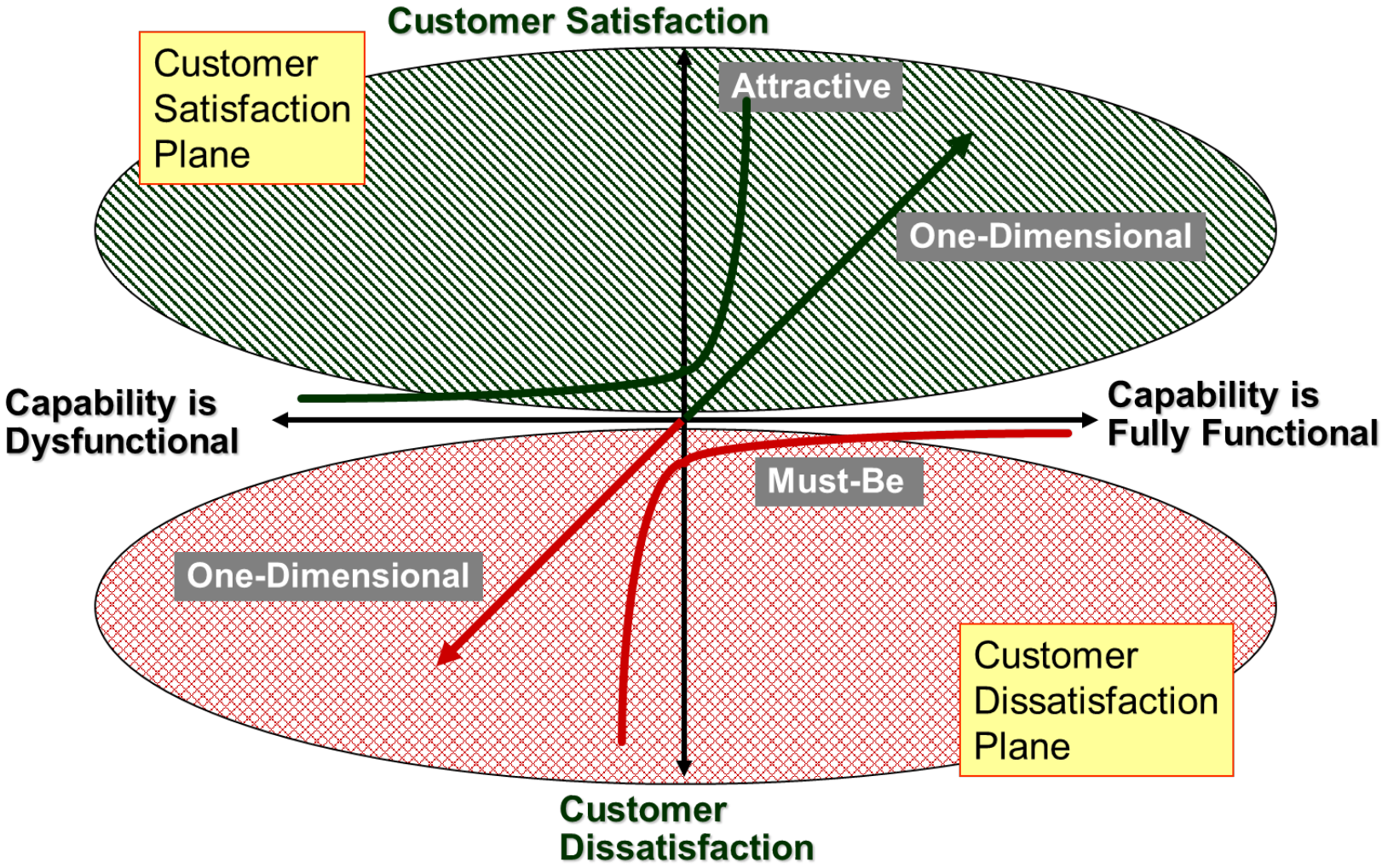
- **A: Attractive**
- **I: Indifferent**
- **O: One-dimensional**
- **M: Must-Be**
- **Q: Questionable results**
- **R: Reversed**

Reference: Shiba, Shoji, David Walden and Alan Graham, A New American TQM: Four Practical Revolutions in Management, *1993, ch. 8*

Converting Kano Results into Weighting Factors

- Weighting according to customer's perception of the requirement's capability to provide satisfaction
- Since relationship between functionality and satisfaction is highly non-linear, one weighting factor per requirement is inadequate. Define two factors for each requirement .
  - First weighting factor: applied when functionality is provided at a level beyond that to which the customer is currently accustomed.
    - Referred to as the "IF BETTER THAN" factor or IBT
  - Second weighting factor: applied when functionality is provided at a level below that to which the customer is currently accustomed.
    - Referred to as the "IF WORSE THAN" factor or IWT

Kano Chart with Satisfaction & Dissatisfaction



Deriving Results for Kano Mapping

Scores:
A: Attractive = 20
I: Indifferent = 9
O: One-dimensional = 57
M: Must-Be = 3

FORMULAE

$$If\ Better\ Than\ (IBT) = \frac{(A+O)\ x\ 100}{n}$$

$$If\ Worse\ Than\ (IWT) = \frac{(M+O)\ x\ 100}{n}$$

where n = total number of responses, A + O + M + I

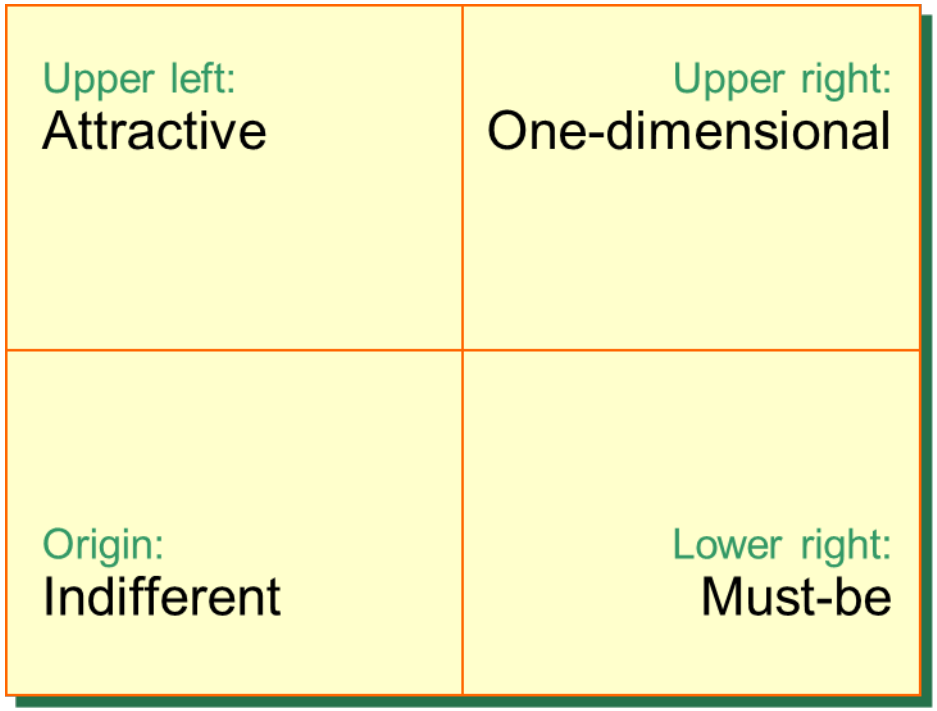
EXAMPLE

$$If\ Better\ Than\ (IBT) = \frac{(20+57)\ x\ 100}{20+57+3+9} = \frac{7700}{89} = 87$$

$$If\ Worse\ Than\ (IWT) = \frac{(3+57)\ x\ 100}{20+57+3+9} = \frac{6000}{89} = 67$$

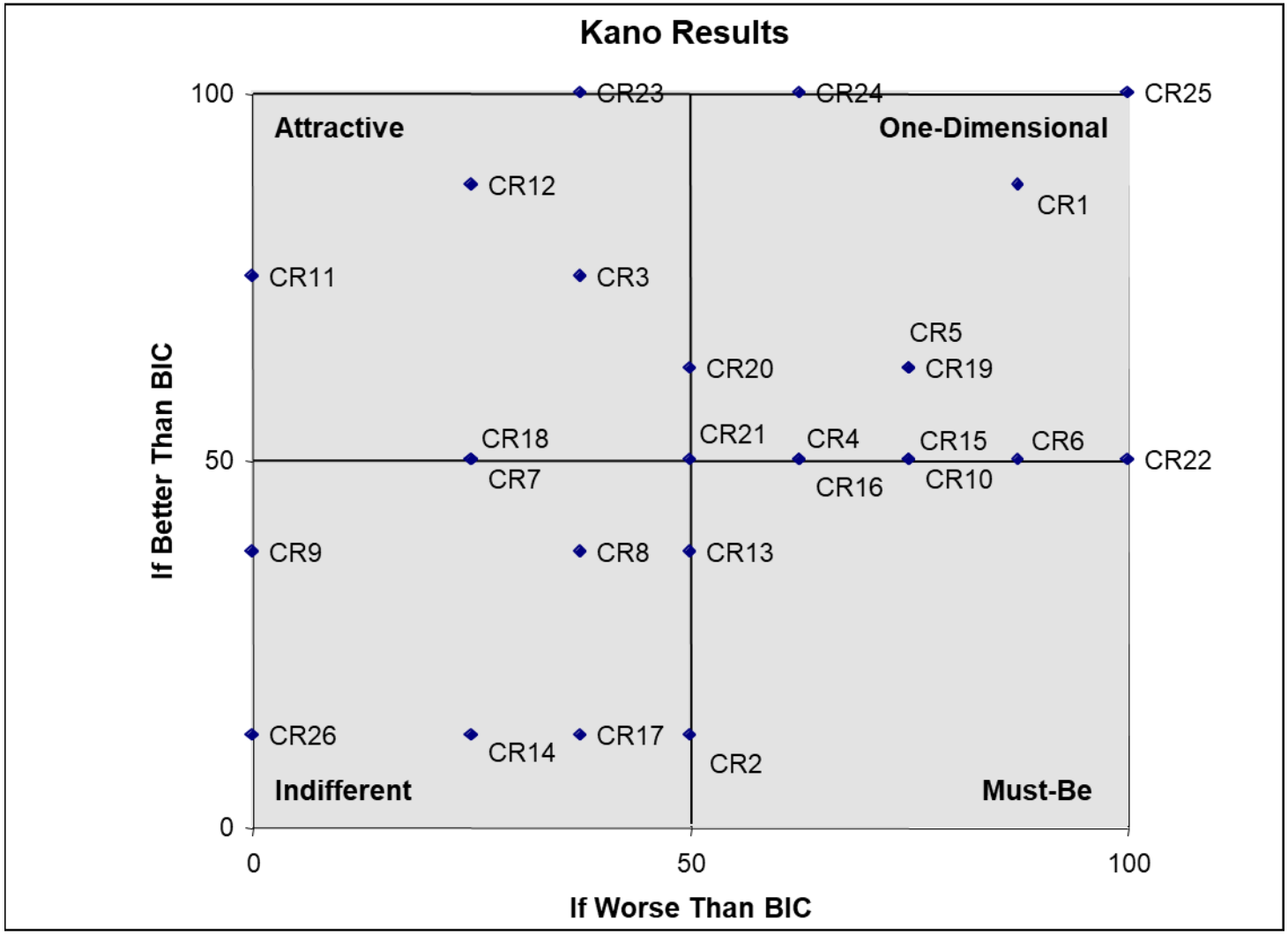
Mapping Kano Satisfaction Factors

- Plot Is Worse Than (IWT) versus Is Better Than (IBT) for each customer need
- Each of the four quadrants has a distinctive meaning:



- Discover new market segments by generating a different graph for each customer type

IBT Versus IWT Kano Plot Example



Step 4. Communicate the Learning

<table>
<tr><td colspan="2">Purpose:
• To identify key messages from listening to the customer.
• To develop and implement a plan for communicating these messages that helps to drive business strategy.</td></tr>
<tr><td colspan="2">Key Activities and Sub-Activities</td></tr>
<tr><td>Define the strategy for communicating with internal and external customers.</td><td>• Identify the findings that are important to business and customer perceptions.
• Identify the key audiences.
• Confirm “the why” (the purpose and objectives) for communicating findings.
• Form the key messages that need to be communicated.</td></tr>
<tr><td>Develop the communication plan for internal and external customers.</td><td>• Define key messages for target audience.
• Identify the appropriate vehicles for communicating (meetings, newsletters, press releases, on-line systems, training, work groups, displays, etc.).
• Determine the frequency and sequence of vehicles.
• Identify the accountability for each vehicle.
• Identify the means for measuring performance.</td></tr>
<tr><td>Implement the communication plan.</td><td>• Monitor performance.
• Adjust as necessary</td></tr>
<tr><td colspan="2">Supporting Tasks</td></tr>
<tr><td colspan="2">• Review the results from previous customer communications.
• Assess how the listening activities’ findings relate to customers’ perception of value.
• Use internal and external customer groups to formulate and test potential messages.
• Seek employee input as to strengthen understanding, buy in, and improved performance around CTQs.
• Identify ways to link training, measurement, recognition and reward systems to key messages.
• Support key messages through your personal actions and communications.
• Remember: internal and external customers’ perception of value affects individual behaviors and business performance.</td></tr>
<tr><td colspan="2">Possible Tools and Resources</td></tr>
<tr><td>Tools:
• Communication Strategy Worksheet.
• Communication Plan Worksheet.</td><td>Resources:</td></tr>
</table>

Step 5. Drive Business Activities

<table>
<tr><td colspan="2">Purpose:
• Ensure that process improvement, design and control efforts are driven by VOC activities.</td></tr>
<tr><td colspan="2">Key Activities and Sub-Activities</td></tr>
<tr><td>Identify and manage processes to achieve CTQs.</td><td>• Identify what existing business processes directly and/or indirectly affect customers' satisfaction and value CTQ.
• Utilize PM for existing processes to ensure CTQs.
• Identify obstacles to achieving CTQs.</td></tr>
<tr><td>Improve existing process and develop new processes to achieve CTQs.</td><td>• Use DMAIIC to work to improve existing processes.
• Use Design Process to develop new processes.</td></tr>
<tr><td>Assess the effectiveness of CTQs in achieving business strategy.</td><td>• Correlate CTQ performance to the achievement of business strategy.
• Revise CTQs, as needed, as well as the business and support processes to achieve business strategy.</td></tr>
<tr><td>Reinforce and reward importance of CTQs.</td><td>• Identify and communicate specific performance required to achieve CTQ.
• Train, develop, measure, and provide feedback related to this performance.</td></tr>
<tr><td colspan="2">Supporting Tasks</td></tr>
<tr><td colspan="2">• Establish and empower teams to drive process improvement around CTQs.
• Identify ways to achieve buy-in to changes related to achieving CTQs.
• Identify the gap between current performance and that required to achieve CTQs.
• Assess the training and development required to address this gap.
• Identify leadership behaviors/actions required to drive performance related to achieving CTQs and business strategy.
• Encourage individuals to assess their own performance related to achieving CTQs and business strategy.
• Assess policies, reward and recognition systems.</td></tr>
<tr><td colspan="2">Possible Tools and Resources</td></tr>
<tr><td>Tools:
Linking to Business Strategy Worksheet.</td><td>Resources:
The Customer Driven Company - Richard Whiteley
Customer Centered Growth - Richard Whitely, Diane Hessan</td></tr>
</table>

4.2 Exercises

Objective: To assess your company's current Voice of the Customer Process

Instructions:

1. Review the characteristics of a successful VOC process (Section 4.1, page 6).
2. Turn these into questions and assess your company's current VOC activities.
3. Develop a list of strengths weaknesses, opportunities & threats (SWOT) analysis regarding VOC at your company.

Time: 30 minutes

Objective: To consider the nature of “customer” for a product or service.

Instructions:

1. Pick a product or service for which you are responsible.
2. Brainstorm a list of customers for this product or service.
3. Classify them as External, Internal Customers, and/or Stakeholders.
4. Review the list for completeness.
5. Take one or two of the external customer groups. What segmentation strategies are possible for these groups?

Time: 20 minutes

Objective: To identify current sources of customer “voices.”

Instructions:

1. Pick one of the customer groups from the previous exercise.
2. How do you currently “hear” their voice? (Review the list of potential voices, Section 4.1, page 9)
3. How are the data currently analyzed? Acted upon?

Time: 20 minutes

Objective: To develop an interview guide.

Instructions:

1. Pick a product or service for which you are responsible.
2. Develop a list of questions to determine how one customer group “feels” about the product/service (satisfaction/dissatisfaction with product features, performance, associated services).
3. Organize these questions into an interview guide.
4. Determine how you will analyze the results of the interviews.

Time: 30 minutes

Objective: To practice interviewing techniques.

Instructions:

1. Review the interviewing techniques (Section 4.1, page 16).
2. Using the interview guide developed in the last exercise, practice executing the guide with a partner (and vice versa).
3. Give each other feedback on the interviewing techniques observed (pluses and deltas).

Time: 30 minutes

Objective: To practice planning and conducting a focus group.

Instructions:

1. Pick a product or service for which you are responsible.
2. Pick two or three specific topics for which you would like customer reactions.
3. Develop specific questions to address these topics.
4. Plan who you would include in the Focus Group(s).

Time: 30 minutes

Objective: To critique an existing customer survey.

Instructions:

1. Obtain an existing survey conducted by your company (or use the sample survey starting on page 31 of Section 4.1).
2. Review the survey by answering the questions on page 23-24 of Section 4.1.
3. Develop a list of strengths and weaknesses for the survey (note: if the sample survey is employed, try to determine how you would answer questions that are not obvious from the survey – e.g. how to analyze the data, reporting, etc.).

Time: 30 minutes

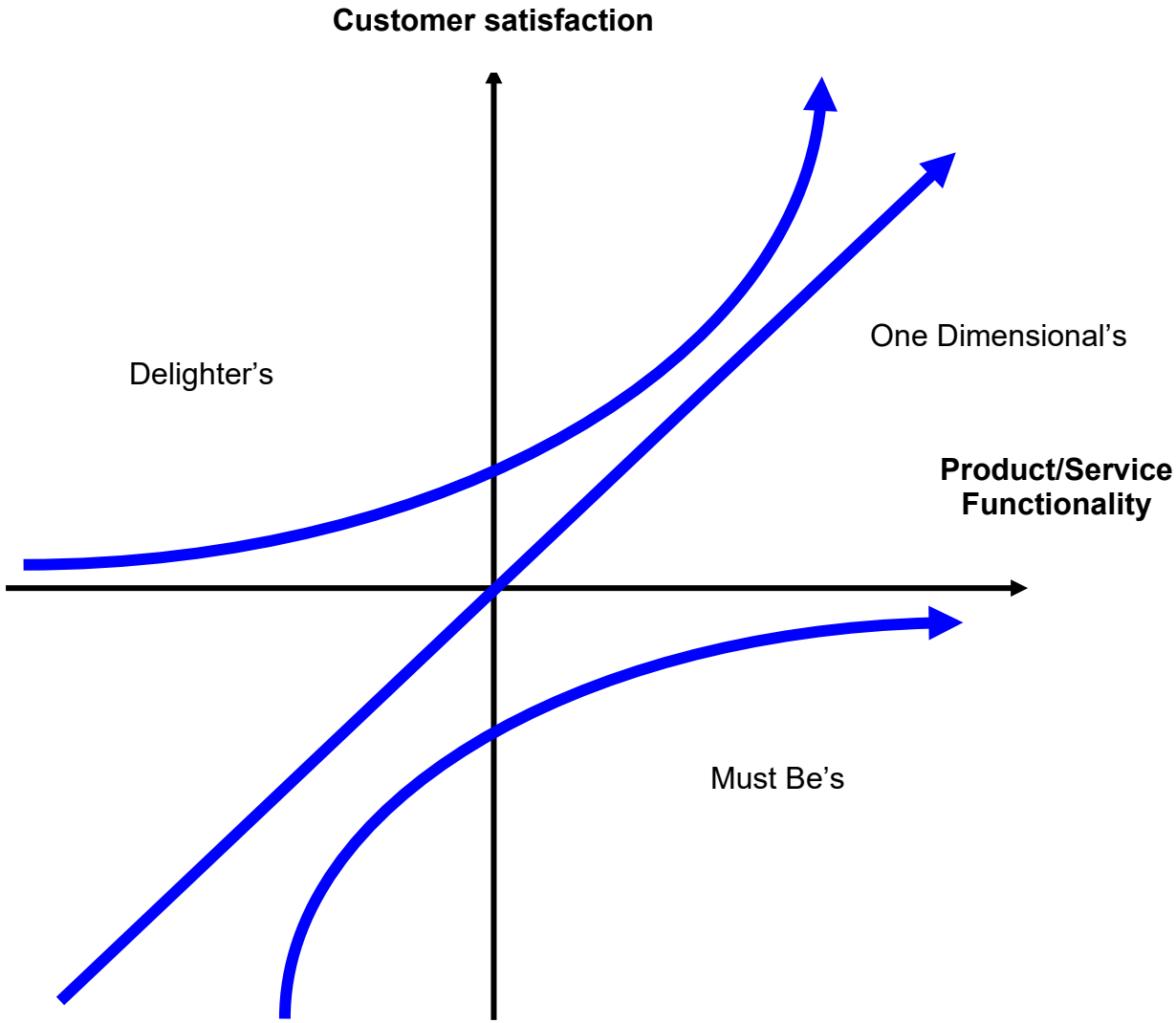
Objective: Testing Dr. Kano's Model of Needs.

Instructions:

1. Recall some products you've used or service you've recently experienced.
2. Brainstorm a list of the characteristics/features of the product or service.
3. Classify these according to the five categories (must-be, one-dimensional, delighters, indifferent, reverse).
4. How well did the company/provider of the product or service meet your needs. Plot your reaction on the graph below (each point represents one need; the X-axis is the "how well" and the Y-axis is your satisfaction response.

Time: 30 minutes

Kano Model of Customer Needs



5.0 Process Management & Analysis

Unit	Description	Page
5.1	Process Thinking	5.1 - 1
5.2	Pictures of the Process	5.2 - 1
5.3	Process Management Methods	5.3 - 1
5.4	Process Analysis Methods	5.4 - 1
5.5	Cause and Effect Analysis	5.5 - 1
5.6	Exercises	5.6 - 1

In this Section, we will present methods to understand and analyze your business processes. One of the key principles of quality management process is that of *management by fact.* To improve an existing process, we need to understand the *process variables* that affect the *key characteristics* of our product or service. In some cases, our experience will help us identify these variables, in others, we must search for the causes by analyzing the process for the “driving” variables, or root causes of poor performance.

5.1 Process Thinking

Learning Objectives

- Understand the concept and elements of a process
- Be able to develop a profile of a process
- Be able to identify process customers and their requirements

Unit Contents

- Process Basics
- Process Customers and Requirements

Why Process?

Let's start off by asking *you* a personal question. Which of the following statements best describes your approach to management (come on, be honest!):

1. Just get me the results; I don't care how.
2. We'll get results, and we'll do it my way.
3. We'll get results, but we'll get them through the process.
4. The process is most important.
5. All we have to do is follow procedures and inspect the results.

Now for the critique: If you answered "1," you probably play golf with Attila the Hun. Look carefully at the faces of your staff the next time you issue an order. Do they seem scared and not sure what to do? Is that really the way you want them to do their work? If you answered "2," you're best buddies with Frederick Taylor. His philosophy was basically to have engineers design the production system, insert the workers and push the button - not very complementary to the workers and their abilities!

If you answered "3," you get a gold star. This statement reflects the basic premise of our discussion:

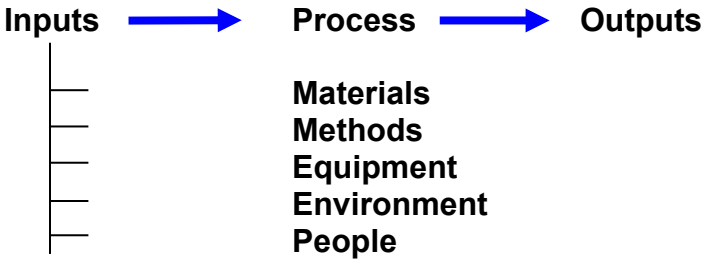
> *The products and services that we "make" in our daily work are the result of* ***processes****.*
> Their quality and cost depend on the quality and cost of the "production" processes.



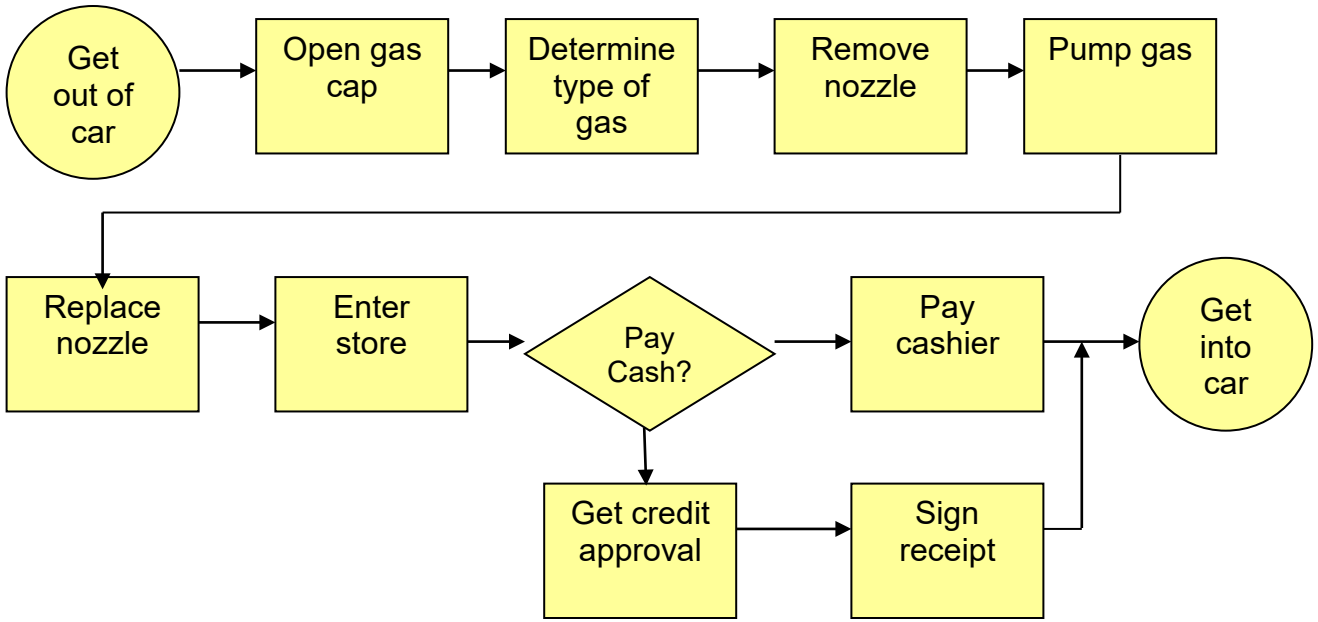
How about answer "4?" Well, there's a lot of *soft* quality management out there. We've heard some of the "Deming Lemmings" (people who think they're quality experts because they attended one of his seminars!) advocate getting rid of all targets and goals and telling people "Just focus on the process and everything will be all right." Not here. Finally, how about answer "5?" We'll have to admit to following this philosophy for some years, having worked in the US nuclear power industry. We've come to realize, though, that this is an outdated quality approach. Inspection has its place, but relying on inspection is too costly. Procedures are important, but they're not enough.

What is a Process?

A process is a group of logically related activities that utilize resources such as people, equipment, information systems, tools, procedures, and material to meet customer requirements and achieve business results. Processes are the way people work, taking inputs, carrying out a set of interrelated activities, and producing specific outputs for one or more customers.



A Simple Process -This process flowchart lays out the *how* of the process. What materials, equipment, etc. are also involved?



What are the consequences of not focusing *on the process?*

To manage current performance and to achieve better results from our process, we must understand which process elements contribute to the quality of the product or service (i.e. to understand *cause & effect* relationships). Let's review some cases where this kind of understanding *did not* exist:

Perception	*A Medical Records director assembles her record coders (workers who enter billing and other info from the manual record to a computer) for a meeting. She tells them in this meeting about all the errors and incomplete codings that have been noticed lately and that they better improve.*
Reality	A later analysis revealed that new coding requirements had been put in place and that only the first shift coders had been trained

Perception	*A maintenance supervisor for a large railroad assembles the overhaul mechanics for a meeting. The maintenance facility has been getting complaints from the locomotive engineers about leaks in their engine cooling systems. The supervisor tells the mechanics they need to be more careful in their engine reassembly work.*
Reality	An analysis of the most commonly occurring leaks revealed that bolts of improper length were being furnished to the workers assembling a critical flange on the engine.

Perception	*At a nuclear power plant, an important valve was found in the wrong position (closed instead of opened). Work orders were reviewed and the person who last performed a procedure that involved the valve was identified. The worker was counseled about following proper work procedures and given three days off without pay to "think about" their error.*
Reality	Further investigation revealed that the procedure, while calling for a step to close the valve, was missing the "open valve" step. Workers were being held to "verbatim compliance" with procedures at this plant.

Perception	*A hospital was faced with increasing supply expenses. The hospital administration changed the purchasing process to require a VP signature for any purchase over $500. Supply expenses decreased for the first two months and then continued their upward climb.*
Reality	An investigation revealed that doctors were increasing the number of "stents"[1] placed in patients during cardiac procedures. Each stent costs about $700.

[1]A "stent" is a tube-like device designed to hold open a patient's weakened blood vessel.

Perception	*A consulting firm (specializing in process improvement!) sent out 10,000 brochures one month describing newly developed services. In the next month, about 300 responses were obtained. The president was all excited about the prospect of new business and revenue.*
Reality	Unfortunately, the one person assigned to follow-up the calls could not keep up with all the responses. By the time she could call back, the prospect had either found some other firm or was upset at the firm's "lackadaisical" response. The campaign cost $60,000; estimated sales from the campaign were $50,000.

What do you think about these scenarios? Do they sound like something you'd do, or have you experienced management actions like those above? What can we learn from these cases? A few observations are listed below. You probably have additional thoughts:

- *People vs. Process* - In the first three cases, "management" placed the blame for operational problems on the workers. In each case, though, the system created by management was the actual problem. The late Dr. W. Edwards Deming estimated that over 90% of workplace problems are due to the *management system*, with less than 10% the "fault" of the workers.

- *Reliance on Procedures* - Procedures and methods **are** important elements of a process. But just because an organization can point to a shelf full of policies and procedures does not mean everything is going all right. This is a balancing act we'll have to explore. One of the best philosophies we've heard describing the use of procedures goes like this - *The "standard" (work method) is there so everybody knows how to do the job (along with training and education). But if the procedure is not constantly followed and improved, then it is useless. If your procedures' revision dates are more than six months old, then they are probably not being used and most assuredly not being improved.*

- *Process Understanding* - In the supply expense case, the administration was being pressured by corporate offices to reduce costs. They felt forced into a "Ready, Fire, Aim" reaction. Understanding of the underlying process that "produces" supply expense was not developed. The results confirmed this.

- *Customer/Supplier Relationships* - In the consulting firm case, although the firm proclaimed their "process-orientation," they failed to consider the "downstream" implications of the sales brochures being shipped. They did not predict that there was a "bottleneck" in their system. The "customer" of the brochures (the follow-up process) could not handle all the inputs from the brochure "supplier" (mailing process).

What's unfortunate is that these are all real cases, although somewhat disguised to protect the "guilty." Many, many more occur daily. It should be clear from these cases, too, that we must understand **both** the variables that make up our processes and their *dynamics*, that is, how do the variables interact with each other. Searching for bottlenecks, understanding critical pathways, root cause analysis, study of variation, workflow analysis, etc. are tools that will help us in this endeavor. In many cases, simple models of a process such as a flowchart or cause/effect diagram, when combined with some equally simple statistical tools will yield all the understanding we need to learn what drives the performance of our processes.

Defining the Process

Here are some basic concepts and tools used to define business processes:

- Process Levels
- Process Boundaries & Links
- Process Ownership
- Process Mapping/Flowcharts
- Process Profile sheet
- Process inventory sheet
- Customer prioritization table
- Customer requirements table

Process Levels

We can think of a business process on several different levels. From a "50,000 foot" level, most organizations look about the same. They all have Marketing, Sales, Product/Service Development, "Production" and Delivery *systems*. As we drill down to the "10,000 foot" level, more definition appears. As you begin to define your business processes, a useful strategy is to start first with the "50,000 foot" view – what are the five – seven major pieces of your process. This provides you with a skeleton on which the "meat" can be drawn. It's surprising how difficult it is to get people to agree on what the skeleton looks like, much less the details.

Level	***Definition***	***Example***
System	A group of logically related processes.	Information System Development
Process	A group of logically related activities or sub-processes that utilize resources (people, equipment, methods, and material) to meet customer requirements.	Program Development
Sub-process	A group of logically related activities within a process.	Requirements Definition
Activity	A series of tasks within a process or sub-process, i.e., what you do.	User Interview
Task	The smallest unit of action within a process or sub-process that is practical or reasonable to study, i.e., how you do it.	Needs Recording

Process Boundaries

Process boundaries define:

- Where a process **begins**
- What activities a process **includes**
- Where a process **ends**

Knowing the boundaries of a process:

- Clarifies responsibilities.
- Defines where inputs enter and outputs exit a process.
- Helps establish what should be measured / tracked.
- Specifies opportunities for cross-functional coordination.

Process Links

Process links supply something to or receive something from the process under consideration.

Examples:

- **Supplier Links** - Inputs (information, products/services) into a process.
- **Support Links** – Activities, usually outside the boundaries of the process, that review, check, have procedural or legislative impact on the flow, e.g., legal dept.
- **Customer Links** - Processes that receive outputs.

Process Ownership

A process owner has responsibility for all aspects of a process within its boundaries, wherever they may exist horizontally and vertically within the organization.

Responsibilities of a process owner:

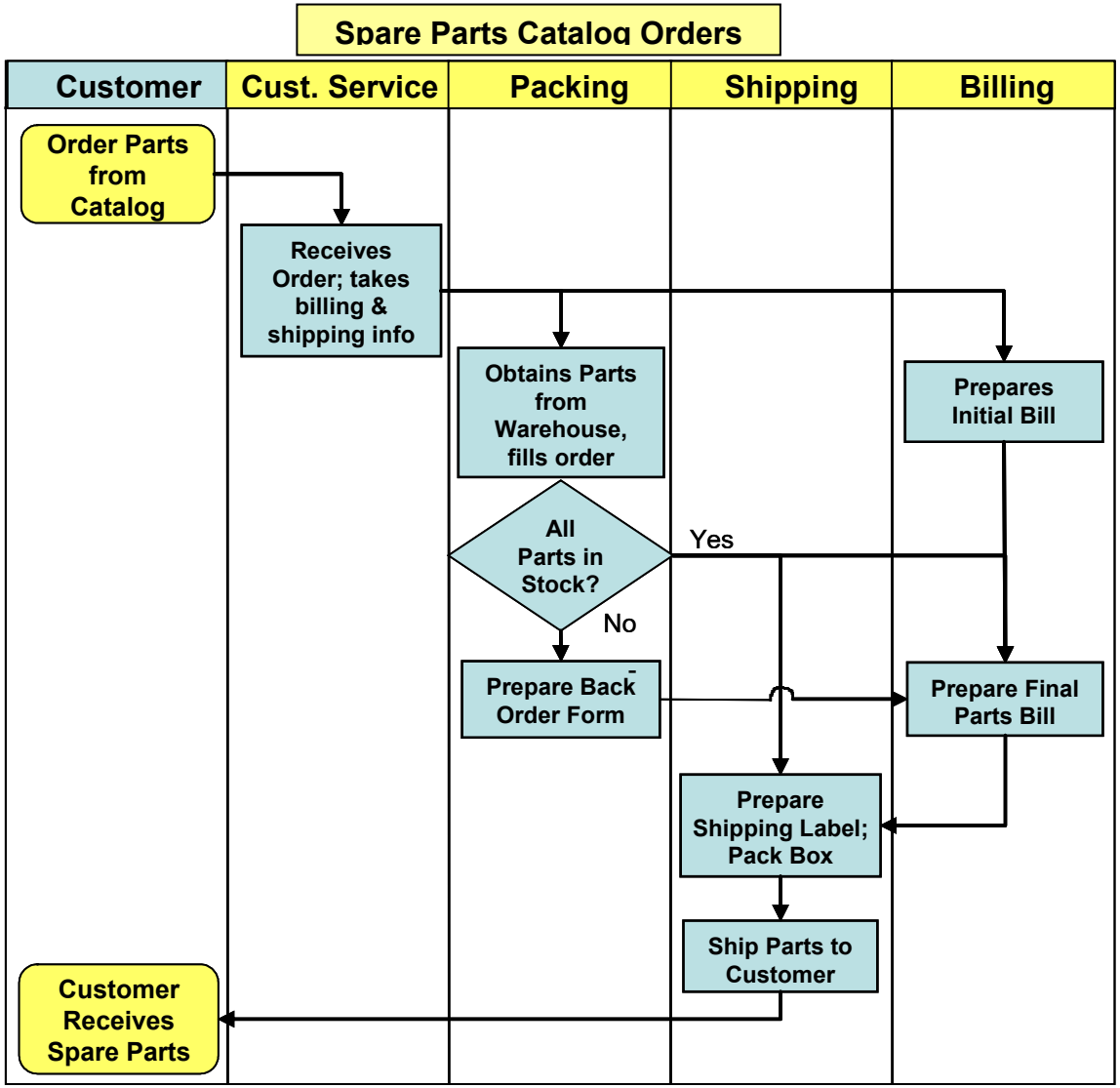
- Identify and document customer requirements and other information required for the process to be effective
- Define sub-processes and assign ownership
- Ensure people who work within the process understand what customers expect.
- Define key linkages necessary to meet the needs of the organization/work unit over time.
- Establish indicators and set targets to monitor process effectiveness
- Ensure the integrity of information throughout the process
- Resolve or bubble up cross functional issues
- Ensure the process satisfies agreed upon customer requirements by involving others who can impact/improve results

Process Mapping - Cross-functional Processes

Unit 5.2 will present various "pictures" you can develop of your process. Page 3 of this unit depicts one of these: the "humble" flowchart." The cross-functional (or *responsibilities)* flowchart shown below allows us to add the *ownership* dimension to the flowchart:

The cross-functional flowchart is often drawn at a high-level, with each activity the possible subject of a more detailed flowchart. The main point is to show the activities and their relationships (inputs and outputs) across departments, divisions or individuals.

By convention, the *Customer* is listed in the left-most column; any *Suppliers* external to your organization would be assigned the right-most column(s).



Process Profile Worksheet

This worksheet can be used to summarize the concepts discussed above for your process:

PROCESS PROFILE WORKSHEET
EXAMPLE

Process:	*Finding Unit Leaks*
Process Owner:	*Quality Control Supervisor*
Sub-processes:	*Various tests performed at different times during production*
Boundaries:	***Starts when:*** *Test request is received* ***Ends when:*** *Test results are logged.*
Links:	*Welding, Fabrication, Assembly, and Painting*

Process Inventory Sheet

This worksheet provides a means of listing the inputs, outputs, materials, machines, people, information and environment of the process.

Process Inventory Sheet

Process:	Water box welding for marine chillers
Outputs:	Welded water boxes
Inputs:	Steel components from suppliers and shop
Materials:	Weld wire and shielding gases
Machines:	Robot welder, fit up jig, clamps, squares, tape measure.
Information:	Job Order Number, Drawing X-123, Robot Setup Procedure RM-231, Hot Work Permit, QC Weld Inspection Procedure QW-3345
Environment:	Inside, open to outside temperature and humidity.
Skills:	Fit up person with measurement and tack welding skills. Robot welding person trained in its operation. All with several years of experience

Process Customers & Requirements[2]

Types of Customers

A customer is anyone who is impacted by the product or process. Customers can be internal or external.

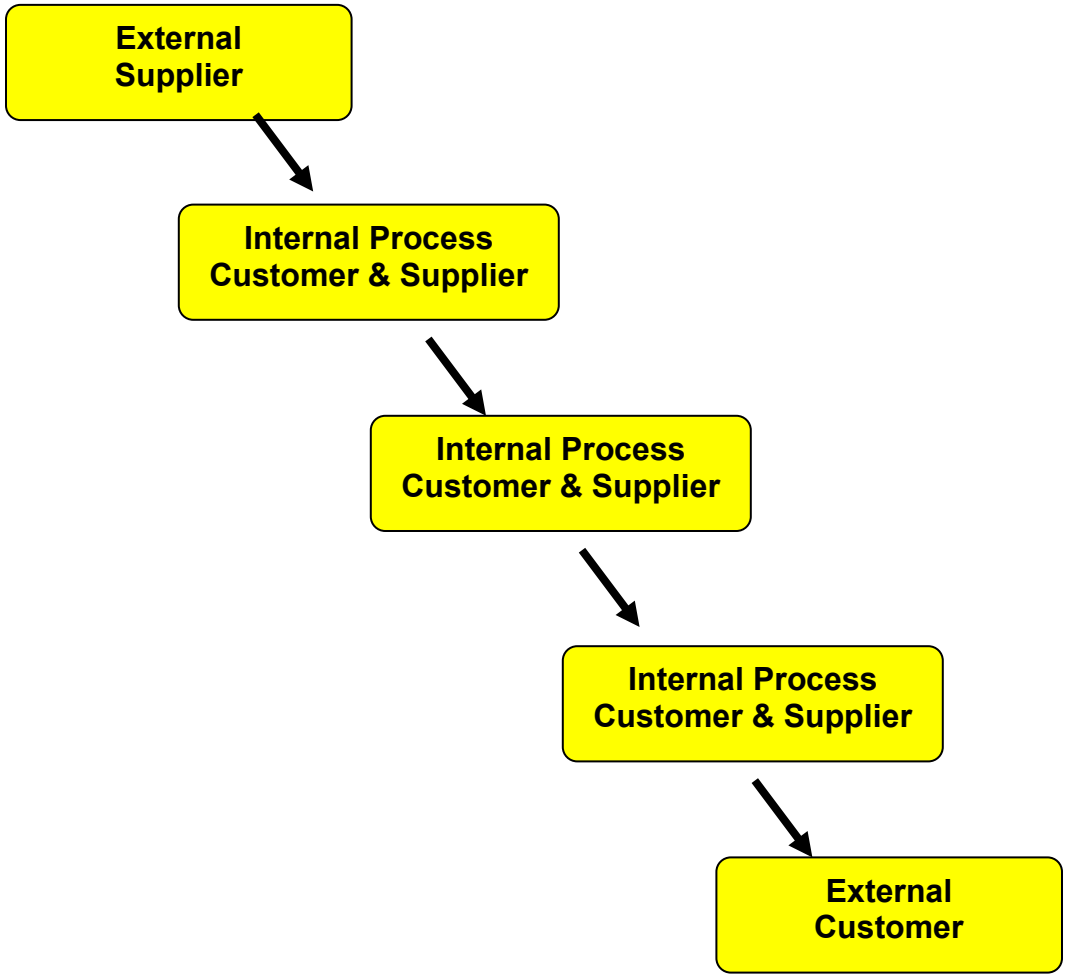
- **External** - people who receive and pay for the company's products/services
- **Internal** - people in departments within our company who use the output of our work (the next process)

Supplier(s)	Product	Customer(s)
East Elbonia Plant Airside Products Group	Air Conditioning Units	General Contractor (external) Owner (external)
Coil Group East Elbonia Plant Airside Products Group	Refrigeration Coils	East Elbonia Plants: Salisbury, Crisfield, Nashville, Wilmington (internal)
Accounts Payable – General Machines Co.	Payment for Material & Services Payment Reports	Company Vendors (external) Management (internal)

[2] Note: See Section 3.1 for a more detailed treatment of obtaining Voice of the Customer information.

Customer Supplier Chain

Most work processes include a chain of customer/supplier interfaces. The chain begins when a supplier provides something to a customer and continues as that customer adds value to what was received, and becomes a supplier for the next customer until the chain finally reaches the external customer.



Points to Remember

- The customer/supplier chain extends from our **external customers** through our company to our **external suppliers.**
- At each step in the **customer/ supplier chain** there are work processes.
- Internally, your customer is **the next process** that receives your output.
- **Suppliers** have responsibility for understanding and satisfying their customer's requirements.
- Instead of re-working or scrapping process inputs, a customer should make certain the supplier understands his/her **requirements**.

Identifying Customers & Suppliers

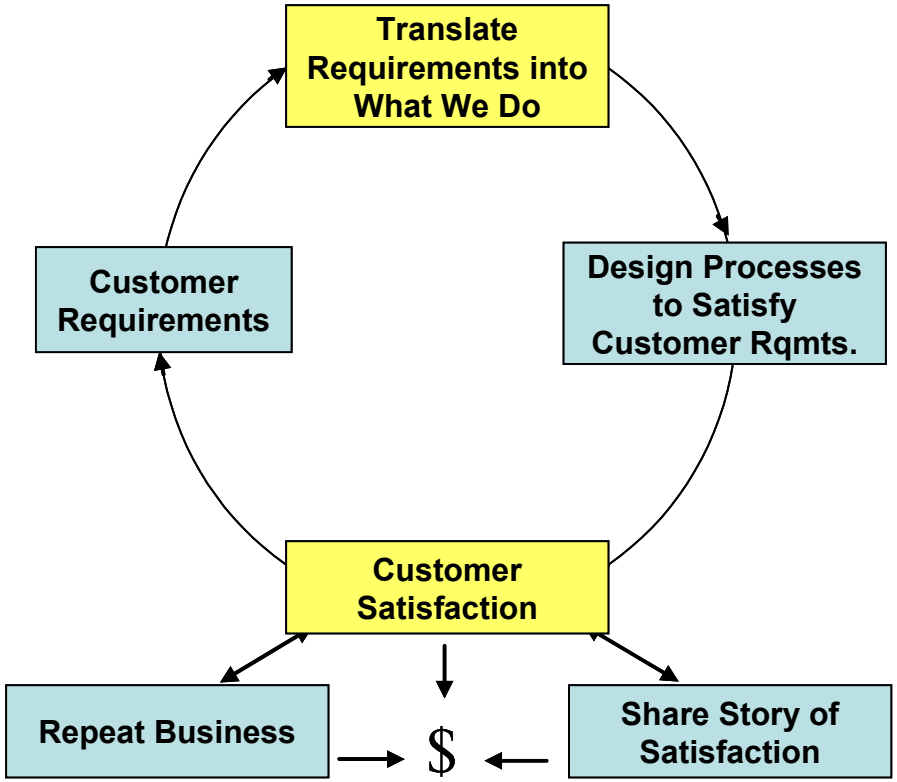
Since a process may have several customers, ensure your resources are used to the best advantage. Apply the Pareto principle (80/20 rule) to classify the list of customers as being part of the vital few or the “useful” many.

Criteria to consider when classifying customers:

- Revenue
- Volume
- Criticality to Other Processes and Strategies
- Number of People (Employees/Customers) Impacted

Responding to Requirements

- We must be responsive to the requirements of our internal and external customers.
- For external customers, our response determines customer satisfaction levels, which translates into whether the customer will buy from us again.
- For internal customers, our response determines our competitiveness in terms of productivity, quality, cost, and delivery.
- Both responses translate into revenue for the company.



Typical Requirements

We are all customers. We feel satisfied as customers and suppliers get our business, when our needs have been met. We call these needs ***requirements***.

Example: 4 doors, air conditioning, automatic transmission, white with blue interior, average MPG of 23, AM/FM Stereo Radio with SiriusXM, 36 month warranty, competitive price.

Most of us want:

- Products and services that are perceived as the best value
- Professional, friendly service from someone who will listen
- To be asked for feedback regarding products and service
- Quick delivery; not kept waiting
- Easy to contact
- User friendly products and instructions
- Problems resolved quickly without a lot of hassle
- Clear and accurate correspondence

Valid Requirements

Valid requirements are process output standards agreed upon by both customers and suppliers. Valid requirements are simple statements that describe how a supplier will provide a product or service to a customer.

Example: Job Status form, legibly filled out, and submitted within 3 working days of job start.

Typical Requirements (continued)

Most requirements fall into one of the following categories:

Quality - Free of errors, defects, and mistakes

Cost - Value exceeds or equals price

Delivery - Output received when needed

Safety - Safe to use; safe for employees

Environment - Not hazardous to environment

Examples of Requirements

Customer	Requirements
Distributor	Sales Support, Competitive Products, Reliable Service, Availability
Consulting Engineer	Product Information, Pre-sale Support Design Tools
Design-Build Contractor	Full Line of Products, Project Support
Contractor	Reliable Service, Easy Installation, Availability, Delivery
Owner/Operator	Value, Comfort, Reliability

Tools to Identify Requirements

The best way to identify requirements is to communicate. This can be accomplished through interviews when the number of customers or suppliers is small, or through surveys when the number is large, making face-to-face communication impractical. The sample interview guide below can be used to discuss requirements with customers and prioritize their importance and your current performance. See Section 3 for more information about interviews and other Voice of Customer methods.

Remember to stay in touch once you have identified a set of valid requirements. Needs can and do change!

Interview Guide

Name/Location/Phone # _____
Process (Product Service) Name _____
Interviewer Name/Date of Interview _____

Importance	Requirements / Metric	Performance
Circle One 1 2 3 4 5	1	Circle One 1 2 3 4 5
1 2 3 4 5	2	1 2 3 4 5
1 2 3 4 5	3	1 2 3 4 5
1 2 3 4 5	4	1 2 3 4 5
1 2 3 4 5	5	1 2 3 4 5
1 Not Important 2 Somewhat Important 3 Important 4 Very Important 5 Extremely Important	Notes:	1 Poor 2 Fair 3 Good 4 Very Good 5 Extremely Good

Customer/Supplier Interview Guide

Using the Interview Guide will help:

- Clarify customer/supplier requirements
- Identify the importance of requirements
- Provide feedback on performance

Customer Interview Process:

1. Before the interview, list the customer requirements and metrics, as you understand them (list your requirements/metrics for suppliers).

2. At the beginning of the interview explain, “I want to ensure a shared understanding of your (for customers) my (for suppliers) needs.” Then, discuss your requirement/metric list, editing as appropriate.

3. Agree on metrics to be used for measuring process performance.

4. Ask the customer (or supplier) to rate and rank each requirement.

5. For importance ratings of 3 and higher and performance ratings of 3 and lower ask, “Do you have suggestions for improvement?”

6. For importance ratings of 3 and higher and performance ratings of 5 ask, “Please provide reasons for this rating.”

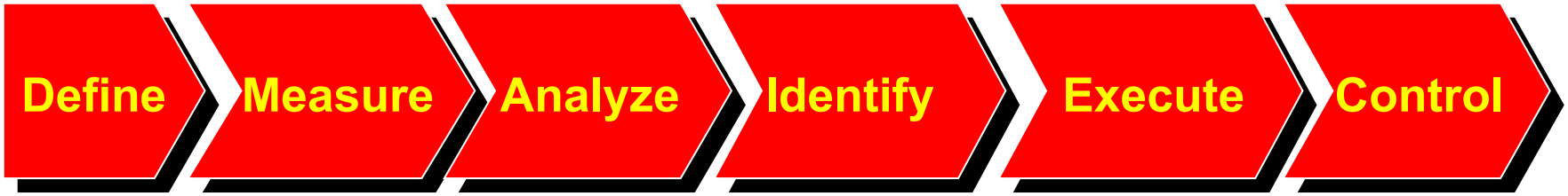
7. Ask for any additional comments. Make notes as required. Then, thank the customer (or supplier) for their time, and suggest that this open dialogue be maintained in the future.

Interviewing Tips

- Remain objective, not defensive
- Listen carefully
- Make note of what is actually said
- Be prepared

Improving Processes

As you interview your customers, you will likely find “gaps” between your current performance and where they would like that performance to be. At this point, you may decide to embark on a process improvement journey. See Unit 2.2 for more information on the DMAIEC method:



5.2 Pictures of the Process

Learning Objectives

- Be able to flowchart a process
- Be able to create a responsibilities flowchart of a process
- Be able to develop a layout diagram of a process
- Be able to conduct a process watch

Unit Contents

- SIPOC
- The Humble Flowchart
- The Responsibility Flowchart
- Layout Diagrams
- Combos & Other Process Pictures

5.2.1 SIPOC (Suppliers-Inputs-Process-Outputs-Customer)

Purpose

The SIPOC flowchart provides a high-level view of the overall process producing the defects. To the "basic" flowchart, SIPOC includes the steps of the process and adds customer and supplier information as well as what inputs and outputs are associated with the process.

Application

SIPOC flowcharts are often used early in DMAIIC to help:

- Scope the process that is producing the defects
- Identify customers – for planning Voice of Customer activities
- Identify process variables that may contribute to the defects

Construction

1. Name the process.

2. Clarify the start and the stop (boundaries) of the process.

3. List key outputs and customers.

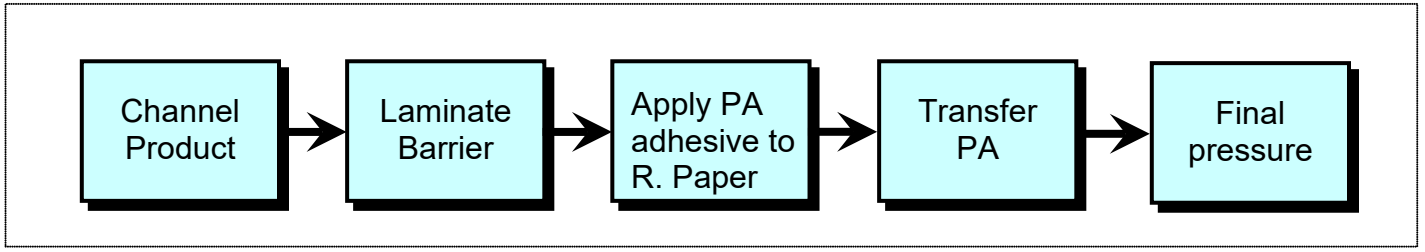
4. List key inputs and suppliers.

5. Identify, name, and order the major process steps (guideline: 5 – 7 maximum).

This SIPOC describes the high level process for applying adhesive to a Sanitary Napkin

Supplier	Input	Process	Outputs	Customers
Raw Material	Web	Apply Position Adhesive (PA) to a Sanitary Napkin web	Sanitary Napkin	Consumer
W+D	Machinery			R&D
	Setting			QA
HVAC	Environment			
Union	Personnel			

Process



5.2.2 The Humble Flowchart

Purpose

The flowchart is one of the simplest, yet most powerful tools for understanding how a product or service is *produced.* The flowchart is a picture that shows the steps or tasks required to accomplish something, whether that "something" is building an automobile engine, obtaining a travel reservation, or changing the dilithium crystals on the starship *Enterprise* (every 5 years or 50,000 light-years!).

Application

Improvement (Measure/Analyze Steps) - The Flowchart is used to understand how the production process is currently performed. We've seen flowcharts on basically any work process that exists:

- *Admitting patients to a hospital,*
- *Manufacturing an axle shaft,*
- *Cleaning a hotel room,*
- *Transferring funds in a bank's "back-office,"*
- *Purchasing "non-stocked" spare parts,*
- *Telephone repairs in a large corporate headquarters building,*
- *Setting temporary security posts at a nuclear power plant,*
- *"Harvesting" organs and tissues from a deceased donor.*[1]

Improvement (Identify/Implement Steps) - The Flowchart is also helpful when redesigning an existing production process or developing a new one. Here, the steps of the process can be laid out and the design "optimized" for whatever quality characteristics are important (time, accuracy, volume, etc.).

Standardization - Many companies have transformed their old policy and procedures manuals into simple flowcharts describing the work processes. They have found flowchart-based procedures easier to develop, easier to understand, and easier for training new employees. At *Sigma Quality Management*, we have all our *core* and *support* work processes laid out on flowcharts. It works for us.

[1]No, we won't be showing that one here!

Construction

"Basic" construction of a flowchart is relatively simple; however, *make sure you know the purpose of your flowcharting exercise*. For example, flowcharting a process to identify *wastes* may require a much greater level of detail than developing one to use as an instruction guide. There are some problems that people typically encounter when flowcharting the first time; we'll point these out after the construction steps.

1. Identify the process to be flowcharted. Most of the time, you're dealing with the *production process* that makes some *product* or *service*.

2. Identify the *ending* point of the process. For a *production process,* this will generally be the point where the service is provided to the customer, or the product is made or received by the customer.

3. Identify the *starting* point of the process. For service-type processes, this point is generally where the customer requests a particular service. For manufacturing-type processes, it may be where the raw material enters the "factory,"[2] or it may be where parts are gathered for assembly.

*Steps 2 and 3 clarify the **boundaries** of the process being flowcharted.*

4. Identify the major steps of the process. If a very large, complex process is being flowcharted, you'll want to start at the "50,000 foot" level. Pretend you're viewing the process from a high altitude. What do you "see" from this height? Generally, the high-level *functions* are identified first.

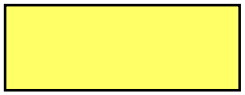
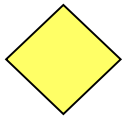
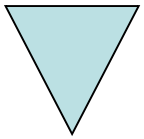
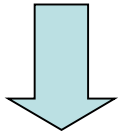
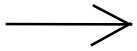
Example: For a manufacturing process, Casting, Machining, Heat Treatment, Assembly might be high-level functions. For a service process, Order Receipt/Entry, Order Filling, Packing, Shipping, and Billing are high-level functions.

Once these are clarified, the "onion-peeling" approach can be taken. Depending on the purpose of your flowchart, the major steps required to perform the functions can then be identified, and then the tasks required for each step identified.

[2]"Factory" must be in quotes, here. We once described a Medical Records department as a "factory," since they took the raw material of patient charts, "processed" and "assembled" the charts, and "manufactured" a completed chart as well as a patient bill. Thinking about the process in this way gave us great insight on how the process worked and what we could do to improve it.

Think of these detail levels as a book. The high-level functions are the book's chapters, the steps are the paragraphs, and the tasks are the sentences. Some organizations will assign "owners" to these high-level processes, generally at a vice-president, or director level. Managers and supervisors then assume ownership for the next level processes.

5. Organize the activities into a flowchart. The basic symbols of a flowchart are presented below:

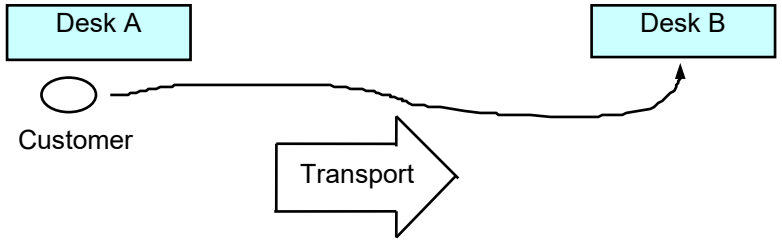
BASIC FLOWCHARTING SYMBOLS	
	CIRCLE - Used to indicate starting and endpoints of the flowchart. The circle should always enclose the first step and the last step.
	RECTANGLE - Used to indicate activities performed as part of the process. The statement inside the rectangle should begin with a verb, an action being taken.
	DIAMOND - Used to indicate decision or inspection points of the process. There will always be two directional arrows - one for "Yes," one for "No" decisions.
	TRIANGLE - Used to indicate where some "Thing" is stored during the process.
	ARROW - Used to indicate transport or movement of some "Thing" during the process.
	LINE - Used to connect flowchart symbols.

There are several approaches to developing the flowchart. Two of the most used are discussed below:

The Walkthrough - This is often used when an existing process is being flowcharted. Draw a circle and describe the starting point inside the circle. Then, ask what happens next. If this is an activity, draw a rectangle and describe the activity. Continue to "walkthrough" the process, from start to end point (the ending activity is also drawn in a circle).

Where there are decisions to be made (options, or inspection points), draw a diamond with the decision question written inside. The flowchart will now branch in two. Remember to complete all branches (i.e. resolve both options).

The table shows symbols for "Storage" and "Transport." Consider the meaning of these symbols broadly. A customer waiting to be served is, at a minimum, being "stored." Similarly, if a customer must pick up a form at Desk ***A***, then fill it in and deliver it to Desk ***B***, a "transport" step is involved.

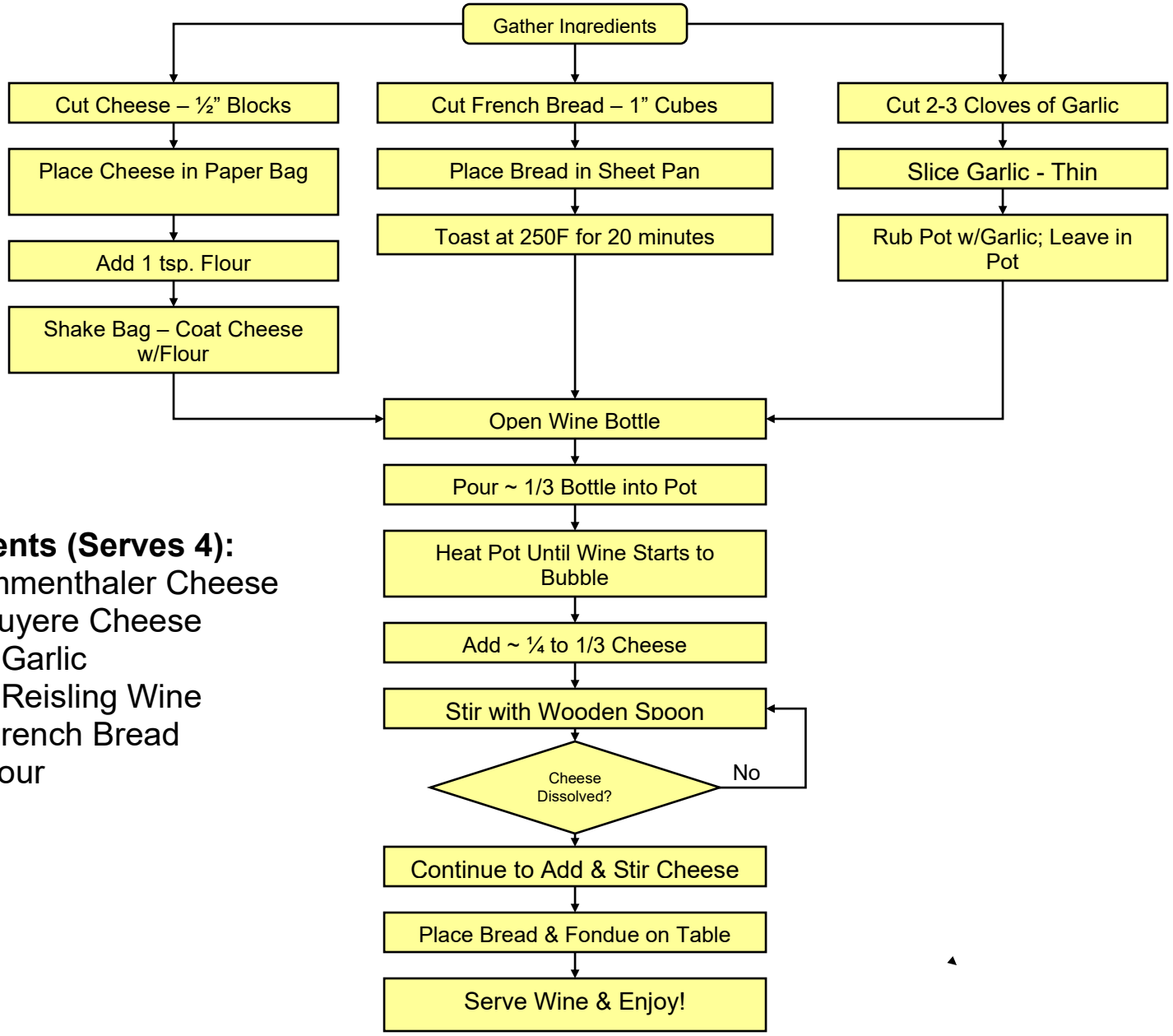


A variation on this process combines the "onion-peel" with the "walkthrough." After the high-level functions are identified, each function is flowcharted individually (often, sub-teams can be assigned to flowchart each function). They are then joined to develop the complete flowchart.

The Brainstorm - This approach is often used when designing a new process via flowchart. Brainstorm all the activities/steps that are necessary to "produce" the product or service. Then arrange the steps logically, considering which steps must occur before others.

6. Label and date the completed flowchart. Undoubtedly, you will be changing the flowchart as you make improvements to the production process. It's helpful to know which is the most current method of producing the product or service.

An example follows that we hope you'll find both illustrative of flowcharting *and* delicious:



Ingredients (Serves 4):

½ lb. Emmenthaler Cheese
½ lb. Gruyere Cheese
1 Clove Garlic
1 Bottle Reisling Wine
1 Loaf French Bread
2 tsp. Flour

Some Notes on the Example:

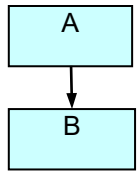
1. This is a *manufacturing-type* process. The aim of this process is to take raw materials and turn them into a finished "product."

2. The purpose of this flowchart is to show *how* the ingredients are processed and assembled. Hence, the *starting point* for the process was chosen to be "Gather Ingredients." We did not include the activities associated with purchasing or shipping the "raw materials."

3. The processing steps for the *cheese, bread,* and *garlic* are shown as parallel paths. This indicates that they can be done simultaneously (if there is only one "chef," though, the available "resources" may cause these activities to be done in series, i.e. one after another).

This is an important point for flowcharting that, unfortunately, is often ignored. When we connect one activity to another in *series,* the following implicit statement is made:

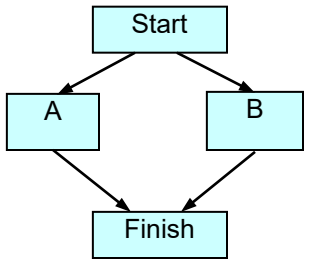
The preceding activity (A) must be completed before the succeeding activity (B) can be started.[3]



For example, the cheese and flour must be in the paper bag *before* the bag is shaken.

On the other hand, activities shown in parallel can be performed at the same time:

Activities 'A' and 'B' can occur at the same time:



[3]In certain circumstances, we'll relax this requirement - the succeeding activity can't be *finished* until the preceding activity is *finished*.

4. There are only two decision points in this example, and both are essentially "DO" loops (Do this activity until some result is achieved). Often, though, decision points will result in different activity paths being followed.

Flowcharting Tips

1. *Boundaries* - In the fondue example, above, we made an easy decision about where to start the flowchart, based on the *purpose* of the flowchart. This guideline is applicable to most flowcharts. Clarify the purpose of the chart!

If the chart is being prepared as part of an improvement effort, review the charter or *reason for improvement* before starting the flowcharting effort. What are you trying to improve?

Sometimes, data can help pinpoint the portion of the production process that needs to be flowcharted. For example, a team working on improving an outpatient admissions process initially set their starting boundary where the patients signed in at the admitting desk.

When they asked the patients about problems with the process, though, they found that many patients were having trouble just getting to the admitting desk. The team then revised the starting point to where the *patient was driving up to the outpatient facility.*

2. *Level of Detail* - This is one of the most difficult areas of flowcharting. Often, teams will get bogged down in the "flowchart session from hell," as they detail every step of the process. We mentioned the "onion-peeling" philosophy, above. This is one of the best strategies to follow.

When we work with teams, we try to get them to describe the process in about 10 - 15 steps, initially. This is the outline or "skeleton" of the process. At this point, we'll try to decide if the entire process needs to be fleshed out, or should we gather some data to determine what part of the process needs further study.

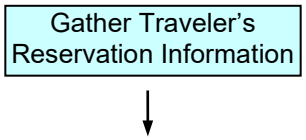
For example, if errors are a concern, gather data on where in the process most of the errors seem to be occurring (see Pareto Analysis, Section 7). Usually, two or three segments of the process will be responsible for up to 80 percent of the errors. Chart these steps in detail.

What ***NOT*** to do: One of our friends decided that his team needed to flowchart *in detail* the procurement process for nuclear power plant spare parts. The flowchart took one month to prepare and stretched forty feet in length! When we asked him what he was going to do with it, his only response was "Hang it on a long wall and look at it."

3. *Multiple Choices* - Service-type processes often have multiple paths or branches, depending on choices made by the customer, or situations that arise in the service process. For example, making a reservation for a hotel room involves the following decisions:

- *How many people?*
- *Adults? Children?*
- *Number of Rooms?*
- *Handicapped-Equipped Room Needed?*
- *Arrival Date? Time?*
- *Departure Date?*
- *Type of Bed(s) desired?*
- *Smoking/Non-Smoking?*
- *Guarantee with Credit Card?*

Now it would be possible to flowchart all these decisions and the resulting choices, but the *purpose* of our flowchart may not require this detail. For example, if we are trying to understand how the reservation process works to improve the *timeliness* of this service, then all of the above questions may be "collapsed" into one activity on the flowchart:



Even if you want to flowchart various alternative paths for your production process, consider the *Pareto principle* and flowchart only those paths that occur 80 - 90% of the time. Those paths that occur "once in a blue moon," while interesting, can distract you from the work of improving the main process.

4. *Sticky Notes* - One of the quickest ways to develop a flowchart (either with a team or individually) is to use sticky notes such as *Post-it Notes™*. Write the process steps on the notes and then arrange them in proper order on a flipchart or other flat surface. Square sticky notes can be turned 45 degrees to make decision diamonds.

5.2.3 The Cross-Functional or Responsibility Flowchart

Purpose

Most organizations divide *responsibility* for various functions by department or section (Billing, Shipping, Purchasing, etc.). While this division of labor may bring certain efficiencies, it can also be the source of delays and errors. In fact, every time a "hand-off" occurs in a production process, there is an opportunity for a delay, for "inventory" buildup and, in many cases, for error.

An important variation of the "humble" flowchart, the *responsibility flowchart* can provide a good picture of not only ***what*** is done, but also ***who*** does it. The responsibility flowchart adds a second dimension to the flowchart, the department or person responsible.

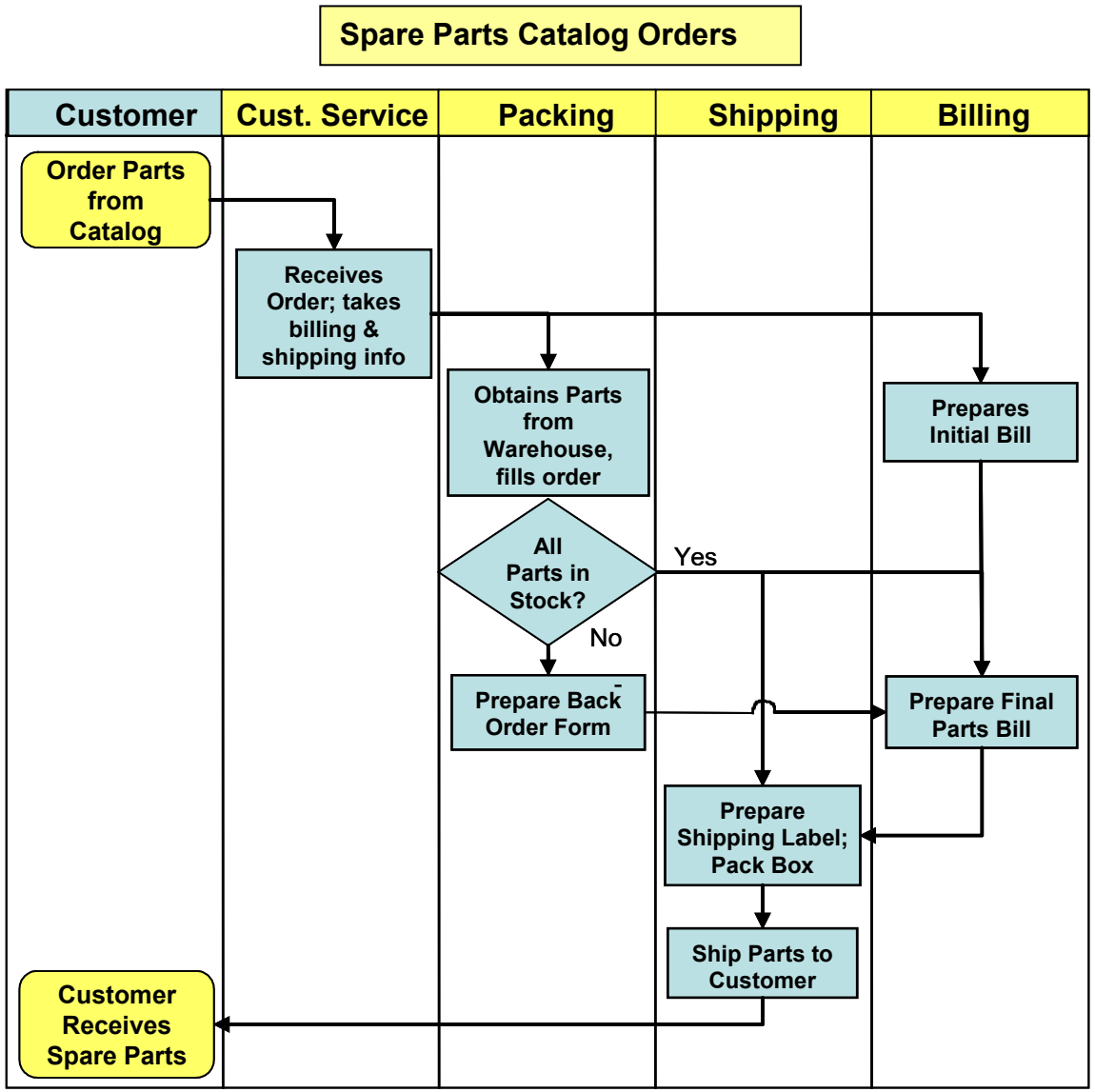
Application

Whenever more than one individual is responsible for a production process, the *responsibility flowchart* is useful. Examples of *responsibility flow charts* include:

- *Manufacturing,*
- *Patient Care,*
- *Purchasing,*
- *Budget Preparation,*
- *Mortgage Loan Processing,*
- *Quality Assurance Systems,*
- *Product Design,*
- *Legislative/Lawmaking*
- *Stock Trading*

Construction

The flowcharting steps for the responsibility flowchart are no different than the process described previously. The difference is in the "second dimension" added by creating columns for the responsible departments or individuals:



The responsibilities flowchart is often drawn at a high-level, with each activity the possible subject of a more detailed flowchart. The main point is to show the activities and their relationships (inputs and outputs) across departments, divisions or individuals.

By convention, the *Customer* is given the left-most column; any *Suppliers* external to your organization would be assigned the right-most column(s).

5.2.4 Layout Diagrams

Purpose

The flow chart shows the steps required to do the work. Of equal interest, though, is *where* the work is performed and how the workflow occurs. The *layout diagram* is a graphical method of picturing the physical flow of work through a process. The location of workstations, storage facilities and transportation requirements can be clearly shown on the layout diagram. The layout diagram also appears as various *drawings* of a product to picture equipment or component arrangements. Like the flowchart, once the physical layout is understood, opportunities for improvement often become apparent.

Application

Virtually any process can be pictured on a layout diagram (it is a good idea to do both a flowchart and a layout diagram, they are complementary pictures). The layout diagram should be developed as part of the *Current Situation/Analysis* steps of improvement to understand the current process. If changes to the production process include physically rearranging equipment or other aspects, these may be "tested" using a layout diagram.

Some layout diagrams we've seen include:

- *Locomotive Overhaul Layout*
- *Plastic Component Part Fabrication, Assembly and Packing Plant*
- *Hospital Emergency Room Layout*
- *Same Day Surgery Nurse and Patient Flow*
- *Laboratory Specimen Processing Flow*
- *Naval Recruiting Station - Recruit Testing Area*

The most common application for a layout diagram is manufacturing plants. Often, plant equipment is located by *function*, all the milling machines are in one area, the lathes in another, the heat treatment equipment and assembly areas in another.

One of our quality friends worked in an aerospace plant, where he measured the distance a rocket engine turbine rotor had to travel from raw material to finished product. The result - *5 miles!* The company was able to rearrange its equipment and drop this distance to *1200 feet*, with a corresponding reduction in storage and transport required.

Construction

The most common layout diagrams of work processes are those that picture the process from a "birds-eye" view, looking "down" on the process. Don't forget, though, that the work area is a *three-dimensional* volume. Other "slices" of the work area may provide a better picture for you.

1. Sketch the general area of interest, including walls, equipment, fixtures, desks, doors, and other items as necessary.

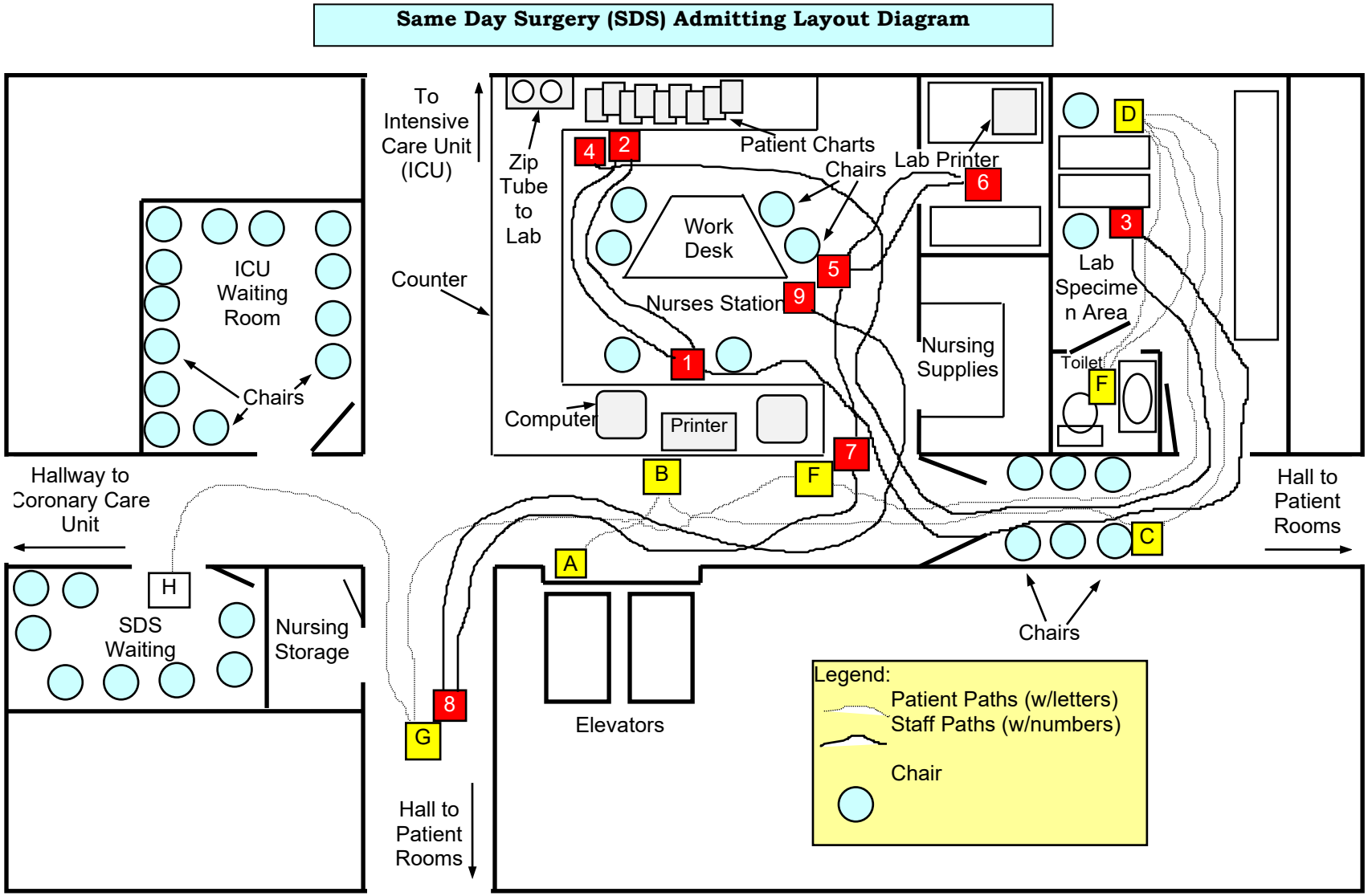
2. Observe the flow of work through the area. Draw arrows to indicate how the product (for manufacturing-type processes) or customer (for service-type processes) flows through the area.

Use the flowchart symbols (activity = rectangle, transport = arrow, storage = triangle, inspection = diamond) to distinguish the "things" that happen to the product or customer.

3. (Optional) Measure various characteristics of the workflow. The two most common include *time to complete steps* and *distance traveled* through the process. Label the diagram with these measurements (average and range values may be calculated from several observations).

4. Place a title and date on the layout diagram. Changes to the workflow should be reflected on the layout diagram.

Example Layout Diagram



Notes on the Layout Diagram Example:

1. We chose the Same Day Surgery (SDS) Admitting Process as an example of a service-type process, since it makes many manufacturing processes seem simple in comparison. Let's go through the play-by-play in this process:

The Patient & Family's Path:		
From	**To**	**Description**
A	B	The patient & family exit the elevator and try to get to the nurses' station counter. There's usually a large group of people in front of the counter.
B	C	After the patient is registered, they are sent to get lab specimens taken. The family waits in the hallway (clogging up the hall).
C	D	The patient enters the lab specimen area. A blood sample is taken.
D	E	The patient goes into the bathroom to obtain a urine specimen. They then return to the lab specimen area.
D	F	The patient (picking up their family along the way) returns to the nurses' station.
F	G	A nurse takes them to their room, where they are prepped for surgery.
G	H	When the patient has been prepped and transported to surgery, the family goes to the SDS waiting room, until the patient returns from surgery.

The Staffs' Path:		
From	**To**	**Description**
1 2	2 1	After the patient arrives and is confirmed on the schedule, the nurse retrieves the patient's medical chart. Additional paperwork is completed and necessary signatures are obtained.
1	3	The chart is walked to the lab specimen area so the lab techs can determine what specimens are required.
3	4	After the specimens are collected, the lab tech walks the specimens to the "zip tube" (for transport to the lab) and returns the chart to the nurses station.
5	6	When the lab work has been analyzed, the lab sends the results to the lab printer. The nurse picks up the results and places them in the chart at the work desk.
5	7	When the patient returns from the lab specimen area, a nurse comes to escort them to their room.
7	8	The nurse takes the patient and family to the room where they will be prepped for surgery.
8	9	The nurse returns to the station and prepares to repeat the process on the next patient.

2. A *Process Watch* was performed at the request of the SDS Supervisor. The layout chart was developed as a result (see *Process Watch* later in this Unit). She knew that there were "production" problems in her process and needed an

outside set of eyes to view the work. When we reviewed the layout chart with she and her staff, it was like playing the children's game, *How Many Things Are Wrong in this Picture?* They came up with over 20 suggestions to improve their work process. How many can you identify?

5.2.5 Combos & Other Process Pictures

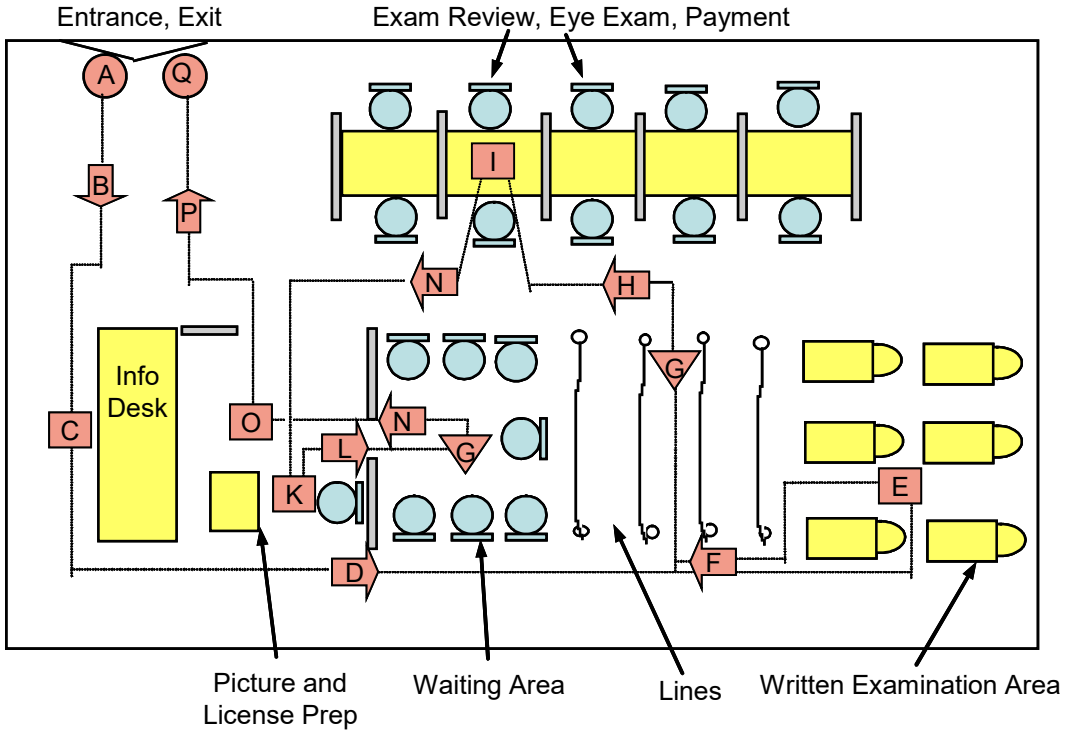
The Flowchart and Layout Diagram are the most used and most useful pictures of processes. Combinations or variations of these pictures have been used, though. Here are a few suggestions that may be useful to you or that may give you ideas about how best to picture your process.

- *Flowchart/Layout Diagram*
- *The Process Watch*
- *Physical Models*
- *Computer Simulation*

Flowchart/Layout Diagram

The flowchart and layout diagrams can be combined to provide a model of *what happens* and *where it happens.* Here's a simple example:

Driver's License Office Layout



This production process can be summarized in terms of its *value-added* and *non-value added* activities. There are five value-added activities from the driver's perspective:

C - Get directions and pick up examination,
E - Complete written examination,
I - Review exam, take eye test and pay for license,
K - Have picture taken, and
O - Receive completed license

The remaining *storage (G & M)*, and *transport steps (B, D, F, H, J, L, N, & P)* are all *non-value-added steps*, that is, they do not contribute to accomplishing the process of getting a driver's license.

This "combo" picture is the start of an analysis of the process, where you would begin to look for ways of minimizing or eliminating *non-value-added activities* from your process.

The Process Watch

The *Process Watch* is not so much a picture of the production process, as it is a way of obtaining a detailed understanding of the process. The purpose of a process watch is simple:

> *Immerse yourself in the production process. Understand everything there is to know about the process. Include the 5W1H (Why, What, Who, When, Where, & How).*

George Butts, former vice president of Chrysler manufacturing, tells this story to illustrate the concept of a process watch:

"One day, I received a request from two young Japanese engineers to review the Chrysler car body painting process. Since we had a technical exchange agreement with their company, I approved the request. I expected them to spend about a half day touring the paint shop.

Well, two weeks later, they visited me to say 'thanks' and 'good-bye!' When I asked them what they'd learned, I was astounded. They had walked through the entire painting process, asking questions and taking notes all the way. For example, they asked the man who wipes the body panels prior to priming what kind of cloth he used, where he obtained it, why did he use a side-to-side instead of up-and-down motion, how much pressure he used, how long he used the cloth before getting a new one, etc., etc., etc. Every other detail of our process was noted as well.

Although they expressed themselves with typical Japanese humility, I could sense their pride when they said, 'Chrysler has the best body paint of all US manufacturers. But you do not understand why. We now understood your process and will take it back to Japan and make it better.' I believed that they would do just as they said."[4]

Application

Improvement (Measure/Analyze Steps) - The process watch is used to help understand how the current production process operates. The Chrysler Painting process watch is a good example of this application.

Improvement (Improve/Implement Steps) - Before "reinventing the wheel," an improvement effort may decide to investigate other, similar production processes and gather ideas that may help redesign their process. Some authors refer to this application of the process watch as *process benchmarking.*

[4]From notes taken at a Palm Beach, Florida ASQC meeting in 1988.

Planning and Executing a Process Watch

The five key questions that must be part of any process watch are simply: *Why, What, Who, When, Where, & How.* Some planning and execution suggestions follow:

Develop a high-level understanding of the process, first. This may be as simple as laying out the key functions of the process in a "macro-flowchart." This helps provide some organization to the process watch. Prepare a notebook, whose tabs are labeled with each of the functions.

If you are going to develop a layout diagram, sketch the physical area on paper before the actual watch. In the Same Day Surgery Admitting process example, we prepared the area sketch the night before. The morning of the process watch, we could then begin to trace the paths of staff, patients and families easily.

Get permission to do the process watch. Explain to the workers you are watching or interviewing *why* the watch is being done. Sometimes, misunderstanding or fear will lead them to act "by the book" instead of how they normally do the work.

Pretend you are a customer or the "product." For service-type processes, this is a very useful strategy. A management engineer followed fourteen customers through a process one day, what she learned that day was more useful than all the data the team had collected up to that point.

Process the results of your watch as soon as possible. Despite all the notes you will take, there will be a great deal of information that will be caught not on paper, but in your mind. Draw out the flowcharts, layout diagrams, and prepare your write-ups immediately. Note any questions that you still have and go back to the process to answer these questions. Don't assume anything or make up information that you thought you saw.

Review the results of the process watch with the workers. Make sure you've captured exactly what goes on. Change any details that were captured incorrectly.

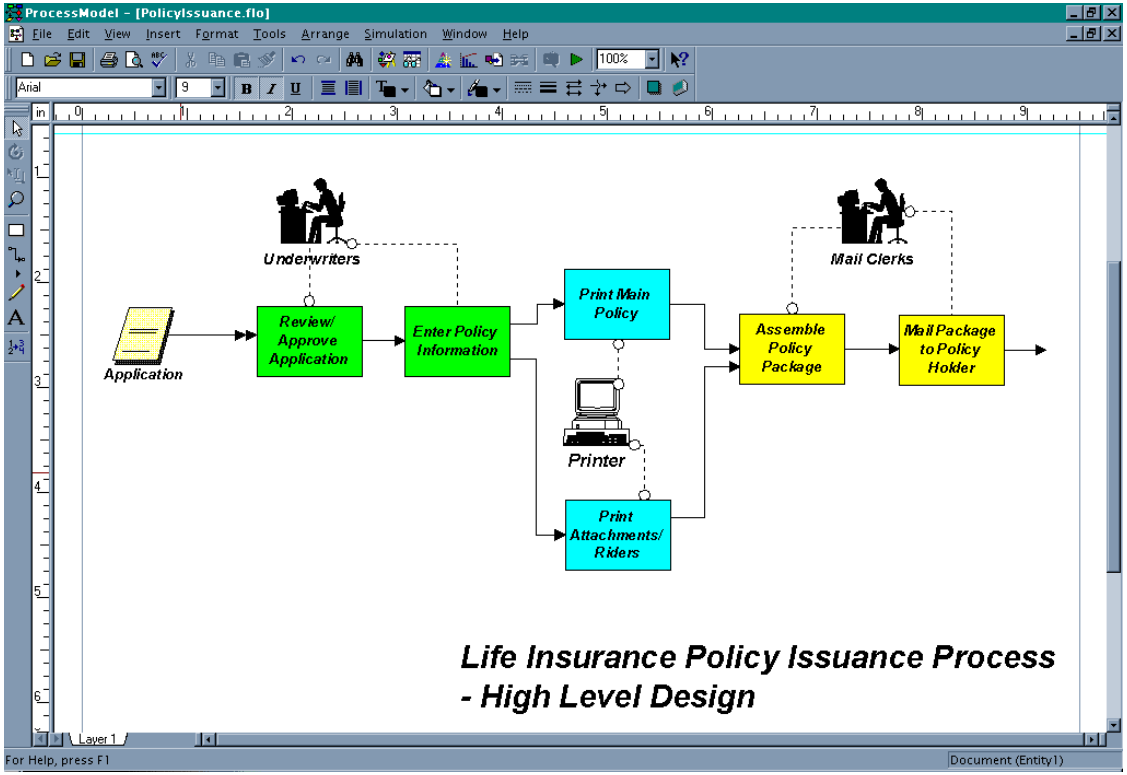
Physical Models

In the "old days," scale or full-size models of complex engineered systems would be built from wood, plastic or other materials. These models provide a three-dimensional view of the system that can be checked for features, such as constructability (interferences), maintainability (how are we going to repair the system), and operability (can we operate this system). Admiral Hyman Rickover had full-scale wood and string mockups built of nuclear reactor propulsion systems to determine if the operators could run and maintain the plant inside the cramped hull of a submarine. Physical modeling is still done today, but the trend is toward "virtual" modeling.

Computer Simulation

Boeing's 777 aircraft was the first to be designed completely on computer. Engineers could view any section of the aircraft, from any angle as the design progressed. Many previous problems with design, such as interferences (two objects "designed" into the same physical volume) were eliminated by this innovative approach.

PC-based programs are available to simulate simpler systems. Architectural software is available that allows a home or office to be created on computer, the prospective owner or occupant can then take a "virtual walkthrough," examining the structure from any angle. For processes, Process Model's *ProcessModel* software can be used to simulate the performance of manufacturing and service work processes.



ProcessModel Software Screen Shot

5.3 Process Management Methods

Learning Objectives

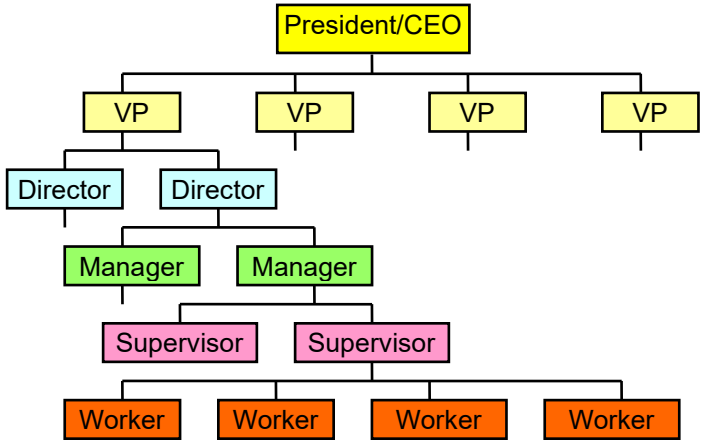
- Understand how the Organization is a System of Processes
- Understand and Apply the PDCA Concept
- “Build” and Use a Process Management System

Unit Contents

- The Organization as a System of Processes
- PDCA – The “Engine” of Process Management
- Process Management Systems

5.3.1 The Organization as a System of Processes

How do you "picture" an organization? Is the organization chart the first image that comes to mind?



What does the "org chart" show us? Well, it mainly shows the organization's reporting relationships and how the organization has divided its functions. You know where your "box" is and how you fit in the chain of command. You may also get some idea from the chart about your department or division's responsibilities.

Now we're not trying to beat up the "traditional" organization. There are advantages to this structure - it is the most efficient from a communication perspective - when the job is to quickly (and orally![1]) pass orders *down* the chain of command. That's why the Roman Army adopted it a few dozen centuries ago.

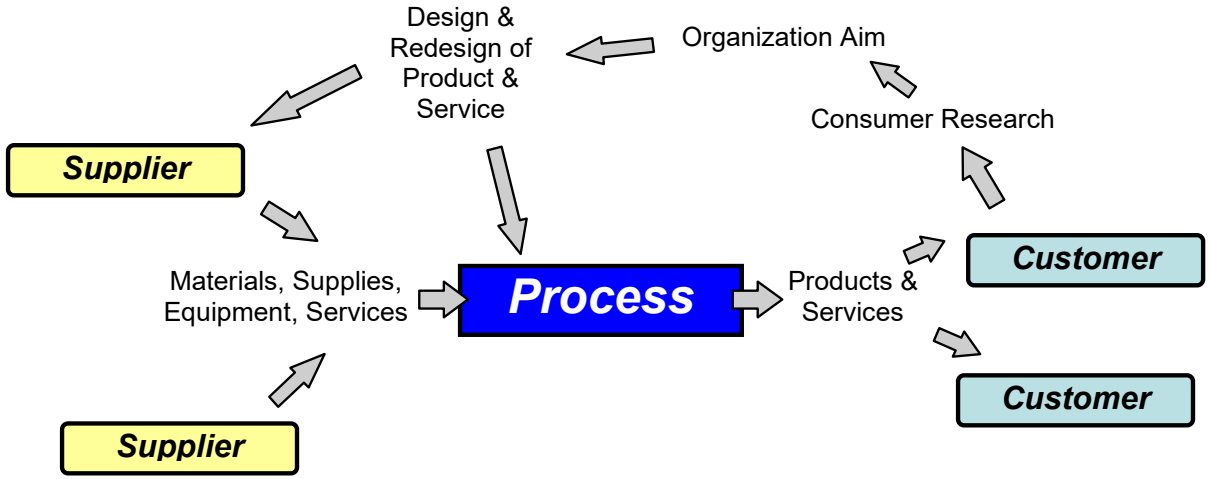
When we employ this (or any other) organizational structure, though, we must recognize that it comes with some "baggage." Departments have an unfortunate tendency to become "fiefdoms" and forget that they serve a purpose **within** a system. We've worked in the "Engineering Dukedom," and have witnessed the effects of the "Power Plant Kingdom," the "Nursing Tribe," the "Nation of Physicians" and others.

There are several other problems with this as a "picture" of our organization. How does the work flow through this picture? Can you see the customers of the organization (are the managers the customers?); can you see the suppliers of the organization? Why are they "outside" this picture? Don't they help you assure the quality of your products and services through their inputs of raw material, supplies, information, etc.?

In 1950, Dr. W. Edwards Deming was talking to a group of Japanese industrialists. He was aware of this departmental tendency and how it could prevent an organization from achieving its aims.

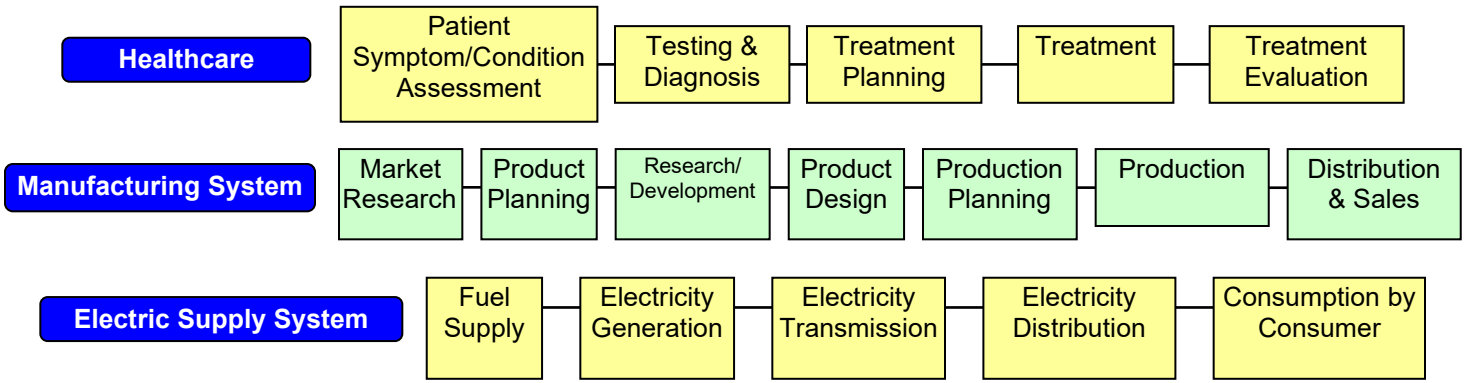
[1]We're talking BE days here - "Before E-Mail."

He drew a simple picture on the blackboard to show them how they must view their organization *as a system with a purpose or aim of providing products and services to satisfy customers:*



Deming's Organization as a System

Now this may seem like a *manufacturing-oriented* diagram. But it works for virtually any organization. Consider these "Production Processes" for different industries:



What does the Production Process for your organization look like? Sketch it out:

Your System of Production

Organization as a System – Notes:

Feedback Loop - The Organization as a System is not a one-way street. It is a cycle. Depending on where you start (e.g. Consumer Research), if you follow the arrows (Design and Redesign of Product, Service, Production Process on through to the actual production processes and the customer) you will return to your starting point. The Organization as a System is fundamentally a feedback loop (more about this in Section 6 - Measurement).

Deming used this cyclic picture to illustrate the need for **continual improvement** of quality, where the focus of improvement may be on the product or service or its production system (including suppliers and vendors).

Workflow Across Departments - Notice, too, that this picture shows the flow of work through the organization. It emphasizes the connections between the functions that we often divide into departments (Engineering, Manufacturing, Shipping, Nursing, Pharmacy, Purchasing, Planning, etc.). Dr. Kaoru Ishikawa coined a phrase to describe these connections:

"The next process is your customer."

Many of us who work in organizations today never "see" the **ultimate customer -** the person who buys the product or service of our organization. In the "old days," when work was much simpler, the person who made the product and the customer often talked face to face. The oft-quoted example of the blacksmith discussing with the noble knight his needs for a new suit of armor illustrates this. Today, the "blacksmith" toils at one station of an assembly line or sits in front of a computer screen all day and rarely sees the customer.

However, we can "see" the person who directly receives our output - the **immediate customer** - and can identify their needs and requirements. Often, though, they will be in the "next" department. Are there "forces" or motivators in your organization to cause the individual departments to identify their internal suppliers and customers and work with them to improve the system of production? We give this activity the term - **cross-functional management** and it is a key element of process-focused management.

Alignment of "Vectors" - We'll never forget one Saturday in February 1988. Florida Power and Light's (FPL's) Nuclear Energy Department management had been called to Japan for a "command performance" in front of Dr. Tetsuchi Asaka, FPL's head quality counselor. A manager from the Power Plant Engineering department (not part of Nuclear Energy Dept.) had just finished describing an improvement he was working on to support the "nucs." Dr. Asaka asked him one simple question:

How was the need for this project communicated to you from Nuclear Energy management?

The manager started to answer that his department always looked for ways to support Nuclear Energy, but then Dr. Asaka cut him off. He turned to the Nuclear Energy management and asked how they "deployed" this improvement to the Engineering manager.

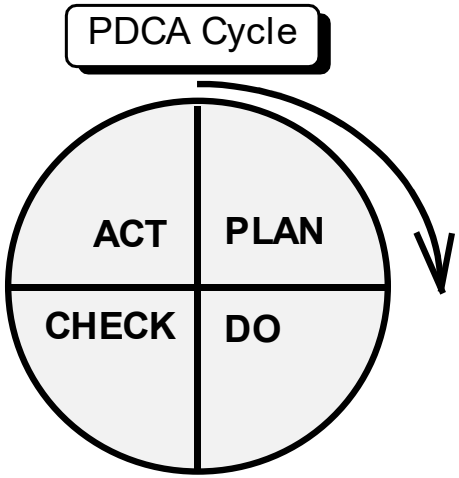
They couldn't answer. In fact, there was no **alignment** of the organization toward a common **aim.** Dr. Asaka's harsh judgment - Nuclear Energy's management system "was nonsense."

Deming made this point repeatedly. Everybody must be involved and it is the job of management to align the organization toward the common *aim* - one based on customer knowledge and the mission of the organization.

5.3.2 PDCA – The "Engine" of Process Management

Deming's Organization as a System gives us the "big picture" of process-focused management. But you may be faced with a more "local" problem - *how do I manage the day-to-day processes for which I'm responsible?*

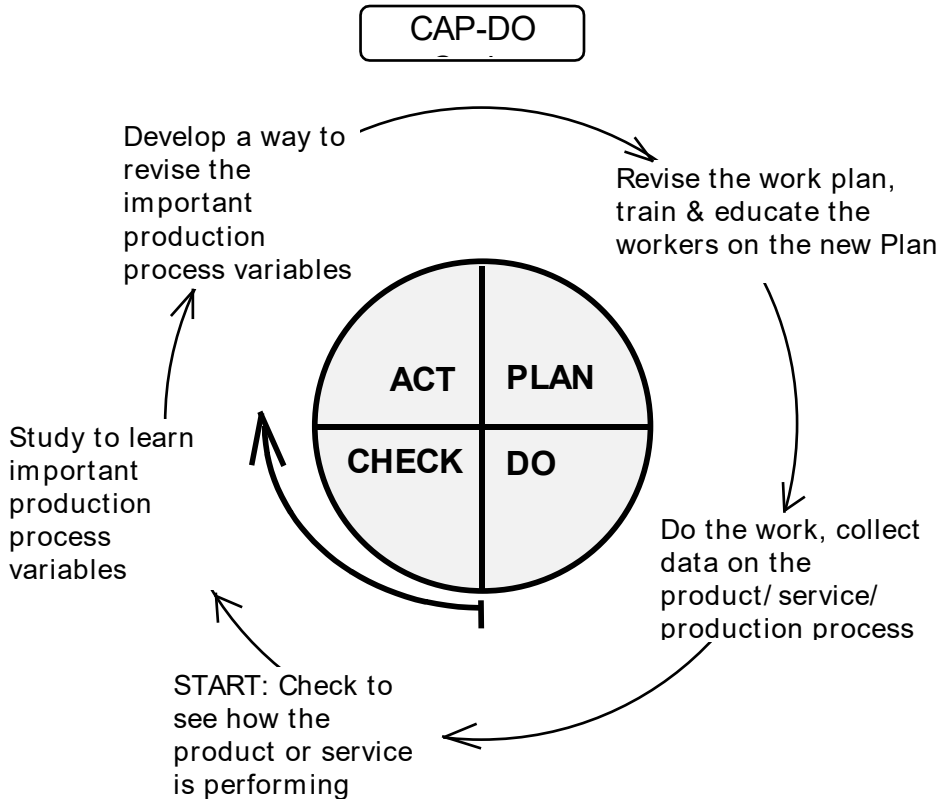
There is another version of Deming's idea - the **Plan-Do-Check-Act** cycle - that will help you.



The PDCA cycle is the "basic engine" that drives process-focused management. How does it work? There are two ways we can "rotate the PDCA cycle:"

- CAP-DO
- PDCA

The CAP-DO Cycle



In your current work, you already produce many products and services. You are already "DO-ing" the work. The CAP-DO cycle first drives you to "CHECK" the performance of the product or service and, based on "gaps" in its Quality, Cost, Delivery or Safety characteristics, you will study the process and its variables.

You will then "ACT" to change the important variables, thus revising the "PLAN." "DO-ing" the work brings you back to the "CHECK" phase, where you will see how well your process changes worked. Sound simple? It really is, but it will take discipline to practice.

For instance, we see many different performance measures in place in organizations. Although there may be an understanding that a "gap" exists, the organization may be reluctant to analyze why the "gap" is there (you probably know them - those chronic problems your organization has suffered with for years). Or there may be reluctance to act even if the important process variables are understood (sometimes, "politics" stands in the way). Even if action is taken and changes are made, the "CHECK" step is often skipped (a change does not necessarily equal improvement!).

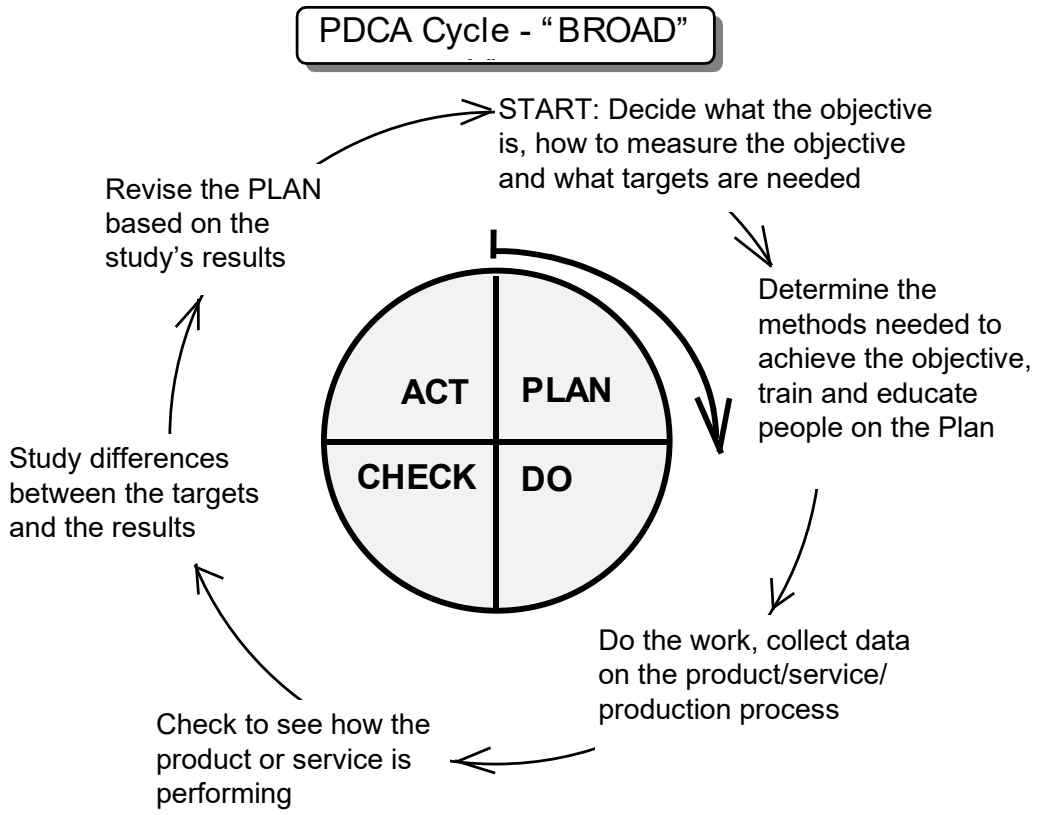
We know several organizations that like to practice the "PCA" cycle: *They PLAN, PLAN, PLAN and then they CHECK the Plan and ACT to revise the PLAN. They never seem to get to the DO phase!*

The PDCA Cycle (or S-DCA Cycle)

There is a "broader" view of PDCA. The CAP-DO cycle helps us *re*design existing products, services or their production processes. For a *new* product or service, though, we will start in the PLAN phase and rotate the PDCA wheel from here.

This approach can also be applied to an *existing* process. Sometimes a product or service is being "produced" without a planned production process.

We often see this exist in the service-functions of organizations. The phrase "*Paving a Goat Path"* can often be applied to their production processes - they've just grown up willy-nilly. Here, it may be worthwhile to start in the PLAN phase to "STANDARDIZE" (i.e. develop a method of achieving the objective) the production processes (hence the SDCA cycle).



Practicing the PDCA (or CAP-DO) cycle in your daily work is the key to successful process-focused management. We'll expand on this in the next few pages.

5.3.3 Process Management Systems

We can combine the elements of *customer, process, variation* and the *Plan-Do-Check-Act* cycle to ensure that we consistently meet the customer-required quality. They come together in a tool called a **Process Control** or **Process Management System.**[2]

There are several aims of a "control" system:

- To establish and communicate the objectives of a production process and to "standardize" the production methods, materials, machines, information and environment,
- To provide a feedback mechanism so that the performance of the product and service can be understood as well as that of variables critical to the performance of the production process,
- To provide a basis for continual improvement of the product, service and associated production processes, and
- To "hold the gains" achieved through the hard work and effort of improvement.

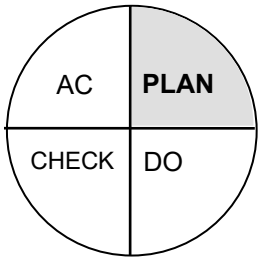
Here, we'll present the elements of a process control system, show how to "construct" a control system and manage using this system. See Section 6 for one of the key elements of process-focused management - *measurement.*

[2]Both "Control" and "Management" terms come with some negative baggage. The object of a Process *Control/Management* System is to assure the quality of our products and services. It is **not** designed to *control* people, nor is it the sole the province of management - *everybody's involved!*

Process Management System Elements

Here, we'll expand the steps of PDCA into the elements of a process management system.

The PLAN



1. The first step in PLAN is to determine the objectives of the process through our customer research. What products and services should be produced? What are the customers' needs and expectations (both "stated" and "hidden" -those that they don't know about)?

2. Next, how can we meet these with our product or service? As we begin to develop the product or service, which characteristics should be emphasized (consider all aspects: quality, cost, delivery, sales & service, safety, corporate responsibility)? At what quality level should the products or services be produced?

3. Based on knowledge of the process' objectives, targets or goals must be set. Quality control is impossible without knowing the "level" at which the process must perform. The targets should not be arbitrarily set but must be a "negotiated settlement" considering what the customer's needs and expectations are and what is currently "technologically" feasible. Since the customer has both *needs* (must-be requirements) and *expectations* (requested or delighted requirements), some of the process targets must be met, others are desirable to meet.

Target or goal setting cannot be considered in isolation by one department. Their processes are part of a system, and the targets should aim to optimize the system. For instance, if *Marketing* sets sales targets that the *Production* people cannot meet, shipments will be delayed or backlogged, quality may suffer in the rush to fill orders.

4. The process objectives/targets should be put in writing and communicated widely. Decisions should be made on how to measure the performance of the process (i.e. what indicators are to be set, how these are to be measured. A Quality Table can help summarize this phase of the PLANNING cycle. Information from this Quality Table can be combined with the actual process flow and work standards to construct a Quality Process Chart.

QUALITY TABLE

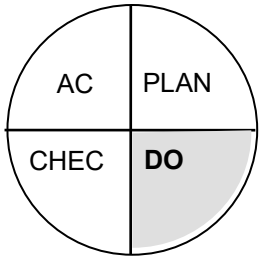
PRODUCT/SERVICE: *CONTROL CHART TRAINING COURSE*					
Customer	***Needs/ Expectations***	***Quality Characteristics***	***Indicator/ Target***	***Process***	***Responsible Department***
Black Belts	Control Chart Theory	• Variation • Measures of Central Tendency, Variation • Sub-grouping Concept • Special Cause Rules	Included in Course Material, Easy to Understand (Pre, Post Test)	Course Development Process	Corporate Training & Quality Services Dept's
	Control Chart Construction	- -	- -	- -	- -
	Control Chart Application to Manufacturing	- -	- -	- -	- -

5. The next step in PLAN is to determine what methods (or "standards") will be used to achieve the objectives. For an existing process, a *flowchart* of the process will at least document **how** the process works and maybe **who** performs the work (a Tischer or responsibility chart shows this explicitly). A *Cause-and-Effect Diagram* can help organize knowledge about the important factors that affect the quality of the process. Since the process may have multiple quality characteristics, multiple cause and effect analyses may be needed.

There are many, many factors that can affect the output of the process. The Pareto principle tells us that there are generally only a few *key* process variables that must be managed to ensure the quality of the product or service. The quality control system must clearly identify these key process variables and the methods established to manage them (i.e. make sure they are at the correct level or "setting," or to minimize their variation). These key process variables may also be measured as part of the control system.

As with the process' objectives/targets, the process methods must also be documented. The process flowchart and Cause and Effect diagram(s) are basic documentation. Procedures or instructions manuals are also helpful.

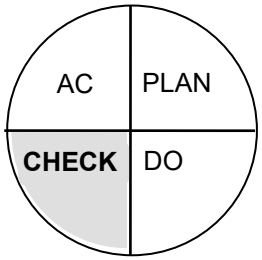
The DO



1. The first step in DO is to educate and train workers on the methods developed as part of the PLAN. This is a major job of management and, if possible, should not be delegated to the "training department." There are two reasons for this: a) by giving management a "teacher" role, they can learn what the methods are requiring the workers to do. This can often help accelerate improvement by itself. b) If management assumes the "teacher" role, the workers will understand the importance of following the established methods. This too helps clearly point out where the established methods are deficient.

2. The second step of DO is simply to do the work according to the methods/ standards. Data is collected to determine how the process is performing and how the key process variables are "behaving." This helps promote "upstream" control of the process.

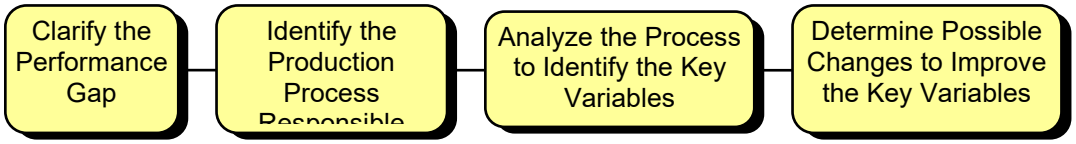
The CHECK



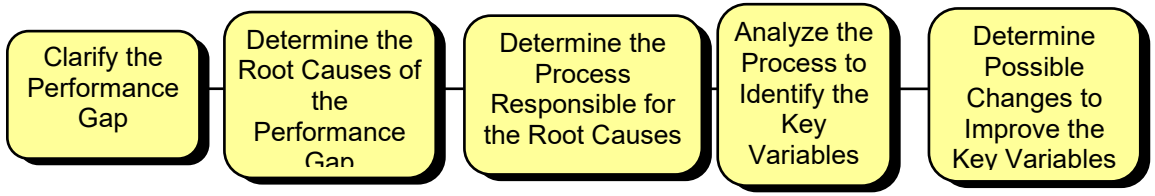
1. Based on the data collected from the process (quality characteristics and process variables), compare the process's performance to the objectives/targets. Is there a "gap?" Are there special causes of variation present?

If there is a "gap," or if special causes of variation are detected, then the process should be studied to understand why the gap exists and what to do about it. From our experience, there are two general approaches to this type of study:

"Gap" Exists, Process Known - For one class of process problems, the "gap" is known, and the process that causes the gap is also known. For example, patients have complained about the registration process taking too long. Or the Billing process is producing too many errors. Here, our study will focus on which variable(s) in these processes are not performing well. The process study will likely follow these steps:

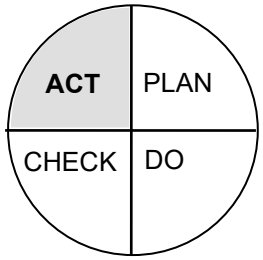


"Gap" Exists, Process Unknown - For another class of process problems, the "gap" is known, but the process that produces the gap is not "immediately obvious to the casual observer." Product failures are typical examples of this class of problem. We are first challenged to determine the root cause of the failures, and then determine which process is at fault. For example, is the failure a Design, Manufacturing, Installation, Operation or Maintenance issue? For this class of problems, the analysis will likely proceed through the following steps:



We'll have to choose which process study method best fits our situation The DMAIEC method (see Unit 2.2) is "biased" toward the *process known* case – you will have to tailor your approach if your problem starts with *process unknown* (Deming said – "All models are wrong, some are useful!).

The ACT



1. Once an understanding is reached of how to close the gap, action must be taken. Depending on the process variables at work, this could include improved training, revising the methods, machines, materials, information, etc. This is where the "narrow" PDCA cycle kicks in. The change must be planned, implemented and checked to see if it is an improvement.

Constructing the Process Management System

Here are some questions that will help you "build" a process management system.

DETERMINING THE OBJECTIVES

1. What is the purpose of this process? What products/services are "produced?"

2. Who are the customers of the product/service? What are their needs and expectations?

3. Translate these needs and expectations into quality characteristics. Which are "must-be's," which are "requested" or "delighted?" Which are most important (Dr. Kano's structure of needs and characteristics)?

4. How can these quality characteristics be measured? At what level should these quality characteristics be set? What are the target values?

DETERMINING THE METHODS

1. How is the current process performed? What are the important factors that affect the important quality characteristics?

2. If necessary, revise the current method to accommodate "new" quality characteristics identified above. Are there characteristics of the existing process that are not necessary? Delete these.

3. How will the important or "key" quality characteristics and "key" process variables be measured or checked? What will be the allowable defect rate or *sigma level* for the characteristic?

EDUCATING AND TRAINING

1. How are the people currently working in the process educated and trained in their jobs? Does the training match the method used to do the work? What new or additional training is needed?

CHECKING THE RESULTS

1. How does the process perform relative to the objectives/targets? How are the process variables performing? Are there special causes of variation present? Why?

2. What is the best method to study the gap? Team, individual, or management effort?

3. What process variables are key to affecting the process output?

ACTING

1. What alternatives exist to modify the process variables? Costs? Benefits?

2. How will the chosen alternative be measured for its effectiveness?

3. What effects were seen from the process change?

The Process Management Book

You may find it helpful to write down and organize the answers to the questions posed above. While the goal is to minimize the "bureaucracy," many managers and supervisors (and teams) have found it useful to keep this information in a binder, known as a "Process Management Book." The book should be available for anyone's' review or learning and will contain the following:

1. Customer Survey Information (Needs and Expectations, current process performance feedback).

2. List of Prioritized Quality Characteristics.

3. Process Flowchart.

4. Process Cause and Effect Diagram(s).

5. Current Measures of Performance (Graphs, Charts of Key Quality Characteristics, Key Process Variables).

6. Applicable Procedures/Instructions.

7. Training Records.

8. Current Process Action Plan - What are the current problem priorities, the status of improvements being analyzed, and the results from changes implemented.

This book is useful for several reasons. It helps the manager keep track of what's "most important." It is invaluable for management reviews and helps put a perspective on things when the latest "flap" emerges. It serves as the "corporate history" of the process, documenting the problems addressed and improvements made to the process.

Process Management Charts

When a power plant manager at FPL was describing to Dr. Noriaki Kano his plans for a major upgrade to the power plant (a conversion from oil-fired to both oil and gas-fired capability), Dr. Kano asked him where his Quality Process Chart was. Our plant manager didn't know what he was talking about, but all Dr. Kano was looking for was the PLAN (the "5W1H") for the project. Many Process Management Chart examples that we've seen look confusing, since they're depicting the details of specific industrial processes (and often just a portion of a complex process). Let's take a very simple "production" process to illustrate the concept and application of the Process Management Chart - we'll make instant coffee!

Process Management Chart Elements

Since we're the customers, we'll ask ourselves how we like our coffee (the customer needs):

Customer Needs: Hot, Sweet (but not too sweet), Smooth, not Bitter, served in a Mug, Slight "Chocolaty" taste, and enough to last through Breakfast. Also, the price per cup should be less than 50 cents/cup.

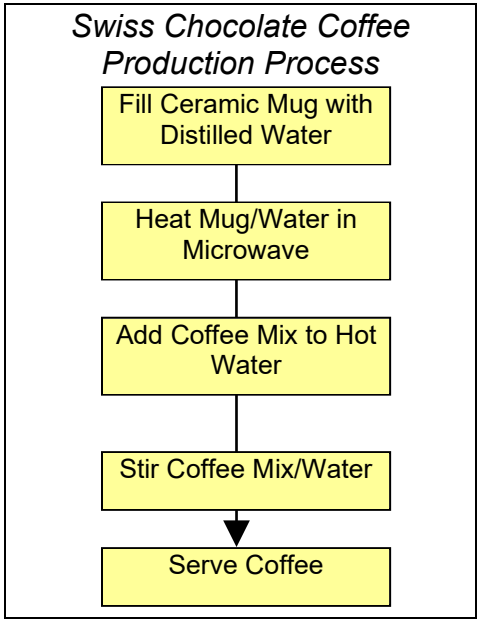
Product Quality Characteristics: We've developed a product called *Swiss Chocolate Coffee* that we think meets "our" customer's needs. The quality characteristics of the product include:

Customer Need	Quality Characteristic	Measure/Target/Specification Limits	*Sigma Target*
Sweet (but not too)	Sweet	3 +/- 0.5 (on 1 - 5 Sweet Scale)	6
Smooth	Smoothness	5 + 0, -0.5 (on 1 - 5 Smooth Scale)	5
Not Bitter	Bitterness	1 + 0.5, - 0 (on 1 - 5 Bitter Scale)	5
Chocolaty Taste	Chocolateness	2 +/- 0.5 (on 1 - 5 Chocolate Scale)	4.5
Hot	Temperature	160 F, $\pm$ 10 F	3
Enough	Volume	7 oz, $\pm$ 0.5 oz.	6
Served in a Mug	Served in a Mug	8 oz Mug + 1, - 0.2 oz.	5
Price	Sales Price	25 cents/cup	N/A

Now that we know what we're trying to produce, we move to the production process. We've identified a supplier (Major Foods) who makes an instant coffee powder mix that meets the first four quality characteristics.

The city water supply has too high a mineral content, we've done experiments to show that we can't meet the *Bitterness* quality characteristic with city water, so we have decided to purchase distilled water in bottles to meet this need (OOPS, the cost just went up!). We've also purchased a nice 8 oz. ceramic mug that has a picture of our favorite commercial air conditioning system on the side.

Having obtained the raw materials, the production process can now be laid out:



This picture assumes the availability of the coffee mix and distilled water. We could add the steps of obtaining these supplies from our vendors, but let's keep the example simple.

The next steps in constructing our Process Chart are iterative. Since we've established the *How*, we could address the *Who.* For this example, we'll assemble the following workforce: *Filler/Heater* and *Mixer/Server.*[3]

With this "organization," we have assigned authority for the production operations and can then think about how to control production so that the quality characteristics (and their targets) are achieved.

Let's take each step and determine what needs to be controlled there and how control will be accomplished.

Fill Ceramic Mug with Distilled Water - The quality characteristics that this step affects are *Served in a Mug* and *Volume.* Our *Filler/Heater* must then obtain the correct mug and fill the mug with 7 oz. of distilled water. These become his *control points*. One of his *checkpoints* may include "Mug with Commercial Air Conditioner Picture" to assure the *correct mug* is obtained. Production engineers have decided to mount the distilled water bottle on a stand with a spigot. Since the same mug is used every time, they've also scribed a *fill line* inside the mug - this is the *Filler/Heater's* check point to assure that 7 oz. of water are added.

[3]We know that you "downsizers" want to combine these jobs, but come on, it's just an example!

Heat Mug/Water in Microwave - Here, the *Filler/Heater* uses the microwave to heat the mug and water. Experiments have determined that the *Filler/Heater* must heat the water to 175F so that the actual *served* temperature is 160F (for temperature losses due to adding room temperature coffee mix and the ambient losses during the mixing/serving process steps). The 175F becomes his control point for this operation; his checkpoint is to set the microwave at "2 minutes, High Power setting."

Add Coffee Mix to Hot Water - Given that the vendor has provided us with the right coffee mix, our *Mixer/Server* is responsible for adding the correct *quantity* of mix to the heated water. We've determined that three teaspoons of coffee mixed in 7 oz. of hot water will satisfy the taste-related characteristics. The amount then becomes the *Mixer/Server's* checkpoint.

Stir Coffee Mix/Water - This step's main purpose is to assure that the coffee mix is dissolved in the hot water. Since the coffee mix has been found to float on the water surface, the mixing must continue until no "lumps" of coffee mix are on the surface. This is the control point for this process. The *Mixer/Server* achieves this by first pushing the mix into the water and then stirring the mix.

Serve Coffee - This is a transportation step. Control items include *not spilling* the coffee, *not dropping* the mug and *delivery to the right customer*.

Note that for these last three steps, the *processing time* is a control item. If this is too long, the water will cool too much, if too short, it may be too hot for the customer. Again, experiments have determined that processing time should be 1 minute (plus/minus 10 seconds) to meet the quality characteristic of coffee temperature. Note that the *"factory environment"* plays a role here - the ambient temperature of the production area is a major factor influencing heat loss during this time - does it need to be "controlled" or could we make the production process "robust"[4] by insulating the mug during production?). Measurements of the current process' capability indicate that this can be achieved. If not, the previous process' target level for temperature could have been increased.

All this information can now be summarized on our Process Management Chart for *Swiss Chocolate Coffee Production:*

[4]Sorry, we're slipping into the quality engineering philosophy of Genichi Taguchi with this issue!

Process Management Chart Example

Process: Swiss Chocolate Coffee Production							**Date:** 1 Oct --	**Rev:** 2
Process Step	**Control Items**		**Measurements**				**Control Methods**	
	Description	**Q,C,D (1)**	**Sampling**	**How**	**Target**	**Chart/ Report**	**Who**	**How**
Fill Ceramic Mug with Distilled Water	Served in a Mug	Q	Each Serving	Checklist	0 Incorr. Mugs	Failure Rpt.	Filler/Heater	Process Standard SCC-001
	Volume	D	Every 10th Serving	Calibrated Measuring Cup	7 oz. ± 0.5 oz.	X,mR Control Cht.		
Heat Mug/Water in Microwave	Water Temperature	Q	Every 10th Serving	Calibrated Digital Thermometer	175F ± 10F	X, mR Control Chart	Filler/Heater	Process Standard SCC-002
Add Coffee Mix to Hot Water	Mix Amount	Q	Each Serving	Standard Teaspoon	3 Level Tsp.	X, mR Control Chart	Mixer/Server	Process Standard SCC-003
	Process Time	D	Every 10th Serving (2)	Kitchen Digital Clock	60 ± 10 Sec.	X,mR Control Chart		
Stir Coffee Mix/Water	Surface Lumps	Q	Each Serving	Visual	None Visible	C Control Chart	Mixer/Server	Process Standard SCC-003
	Process Time	D	Every 10th Serving (2)	See Above				
Serve Coffee	Spills	D	Each Serving	Visual	None	C Control Chart	Mixer/Server	Process Standard SCC-004
	Dropped Mug	QCD	Each Serving	Visual	None	Failure Rpt.		
	Process Time	D	Every 10th Serving (2)	See Above				

Notes: 1) QCD - Quality, Cost, Delivery
2) - Process Time to be measured for total of these three process steps

Now, although this chart looks complicated, you can track how the word descriptions of each process step are summarized on the Process Chart. The chart is a form of "shorthand" for the Who, What, When, Where, Why and How (5W1H) of the process. With a little practice, these charts can become an effective way of designing a production process or standardizing a current process. A few notes on the Process Management Chart:

1. **Do Not** get trapped into one format (or even try to copy the one shown above). The basic questions you want to answer on the chart are the 5W1H but think about how each application should be designed. Some process management charts also include a "response" column. How will special causes of variation be addressed? How will process capability improvements be prioritized and worked?

2. For manufacturing-type processes, the Process Chart may be developed in two stages - during design, the production process is "roughed out" on a Process Chart; as the production planning moves closer to actual production, the details are fleshed out in the actual production Process Chart.

3. In Unit 6.1, we will discuss the difference between **control** and **checkpoints.** Briefly, control points are *effects*; checkpoints are *causes*. These need to be clearly differentiated when preparing a process chart. In the example above, the microwave time and power settings are checkpoints - they are the causes that result in the effect (control point) of water temperature. This is important.

4. In the process chart shown above, the actual "how-to" do a certain process step was listed as a Work Standard (SCC-001, SCC-002, etc.). The development of both the process chart and the work standards needs to be coordinated.

5. The Process Chart helps us design a new process, but it is also something to be followed during actual "production." As such, it should identify not only what should happen when the process is going along smoothly, but also who and what actions are to be taken when something "bad" happens. A while back there was a Saturn™ TV commercial where the production worker talks about the first time he "pulled the cord" and shut the production process down. This is an example of authority delegated to the best place on-the-line. Response plans for assignable causes of variation may be different from those associated with common causes.

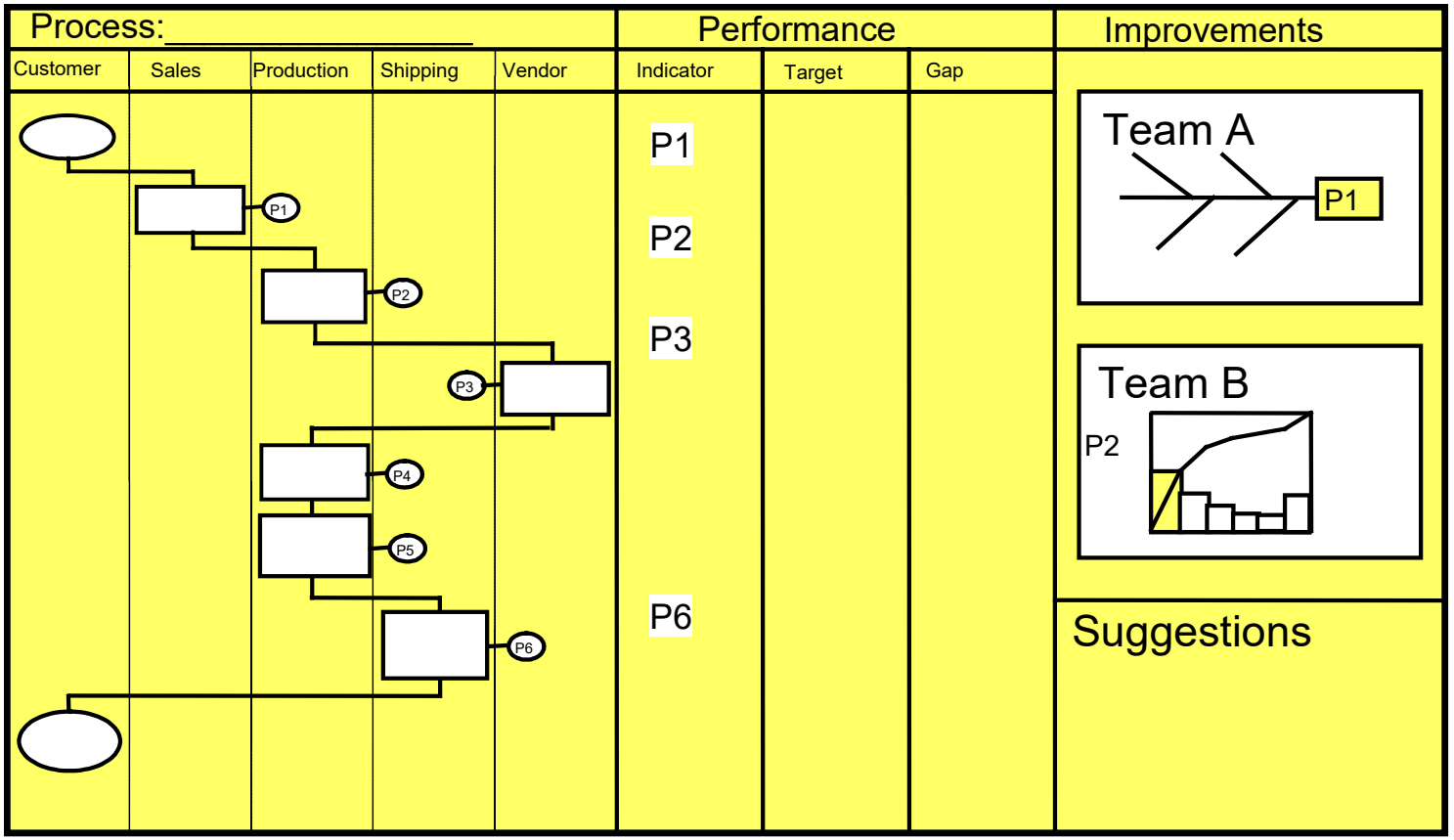
6. Notice that, although there are many factors that could affect the quality of our *Swiss Chocolate Coffee*, only the most important factors have been identified on the process chart. Here's the Pareto Principle at work.

7. We've included the QCD (Quality, Cost, Delivery) column in our process chart. We put it there to help remind us that QUALITY is multi-dimensional and that we need to think about how we control all the important dimensions.

The Process Storyboard

Bruce Sharp, of Duncan Enterprises, gave us this idea for a simple form of a Process Management Chart. The main purposes of this chart are to publicly display the current process (how), the indicators being used to measure performance (and any "gaps"), as well as the status of improvements being worked.

This is a simple approach to the Process Chart that may be useful when you're just getting started in process management. Note how it "links" the two pillars of quality *improvement* and quality *control.* Some organizations also employ quality storyboards that visually document the progress of improvement efforts. If a large area is available, these two tools can be merged.



5.4 Process Analysis Methods

Learning Objectives

- Be able to analyze a process for non-value added activity
- Be able to determine if the process *can* be followed reliably.
- Be able to conduct a comparative process analysis.
- Be able to develop the critical path through a multi-branch process

Unit Contents

- Analyzing the Process' Pictures
- Cycle Time Analysis

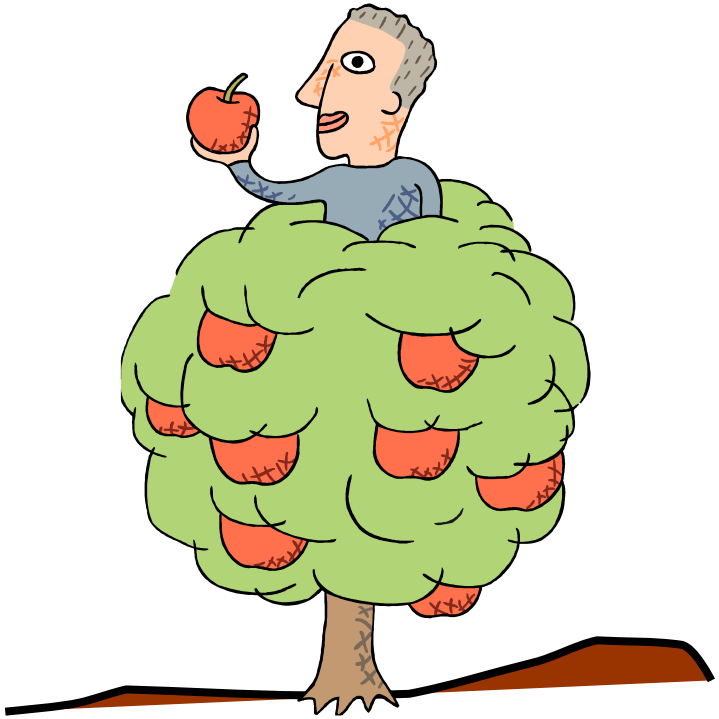
5.4.1 Analyzing the Process' Pictures

Perhaps you've heard of the "Quality Tree?" On this tree grow fruit that represent our opportunities for improvement. Now some of the fruit is lying on the ground (be careful, don't step on these, but *do* pick them up!), some fruit is on the lower branches, easy to reach, some grows on the higher branches, and we'll have to climb a ladder to get to this fruit.

When you *first* develop a picture of your process, we recommend that you examine the picture(s) critically. Experience has shown that you will often identify the "low-hanging fruit" of improvement, here.

Several types of process analysis are presented here; they are characterized by the fact that "only" a picture of the process is needed to support the analysis.

- *Low Hanging Fruit*
- *Was the Process Followed?*
- *Comparative Process Analysis*
- *Twenty Questions*



Low Hanging Fruit

There are some general categories of "inefficiencies" that you should look for in your process. Don't just look for the "big stuff." One company held a campaign where any suggestion that *reduced wasted motion by at least 0.6 seconds* was considered.

Inefficiency	Description/Example	Action
Duplication	When we make reservations with most hotel chains, we provide them with our name, address, etc. When we arrive at the actual hotel, the probability that we will have to give that same information is usually very high.	Eliminate unnecessary duplication
Misplaced Activity	At a famous car manufacturer, the glove box was installed at point "A" on the line. At point "C," down the line, the glove box was removed to install some electrical wiring.	Reorder process steps
Storage	Stored material, supplies, etc., takes up space and incurs cost. Examples range from manufacturing *in-process inventory* to the "vast" amount of pens, pencils, and stationery "stored" in office buildings.	Minimize storage, implement Just-in-Time delivery
Transport	Transportation does not add value to the material or product being transported (consider also the "transport" of customers through your processes)	Minimize wasted transports
Motion	Often people do things a certain way because of habit, not because it's the best way to perform a task.	Minimize wasted motion
Inactivity	There's an old saying, "*A person watching a machine is not working."* One of our quality friends was curious to know why an automated production line was shut down when the workers went on break.[1]	Minimize or eliminate.

[1]It turned out there *was* a "good" reason - the automatic line produced so many defective items, it needed to be constantly watched!

The Five "S's"

One approach to obtaining quick improvements and to prepare the process for further analysis is to perform the Five "S's." See Unit 5.5 for the application of these techniques in Lean Manufacturing.

Sort	Go through the workplace and *sort* all materials, equipment, etc. that is not needed to perform the work. Get rid of this *stuff!*
Shine	Clean the workplace. Remove grease and dirt from equipment, scraps and offal lying around. Hang necessary tools in appropriate places – make it easy to obtain tools for the job.
Set-in Place	Establish responsibilities for maintaining the clean, clutter-free workplace (e.g. Joe and Jill clean up the lathe area once a shift, Jack and Jean clean up the tool bin).
Standardize	Develop a common method for performing the process.
Sustain	Ensure that the common method is employed each time the process is performed and that improvements made to the process are built into the common method.

Was the Process Followed?

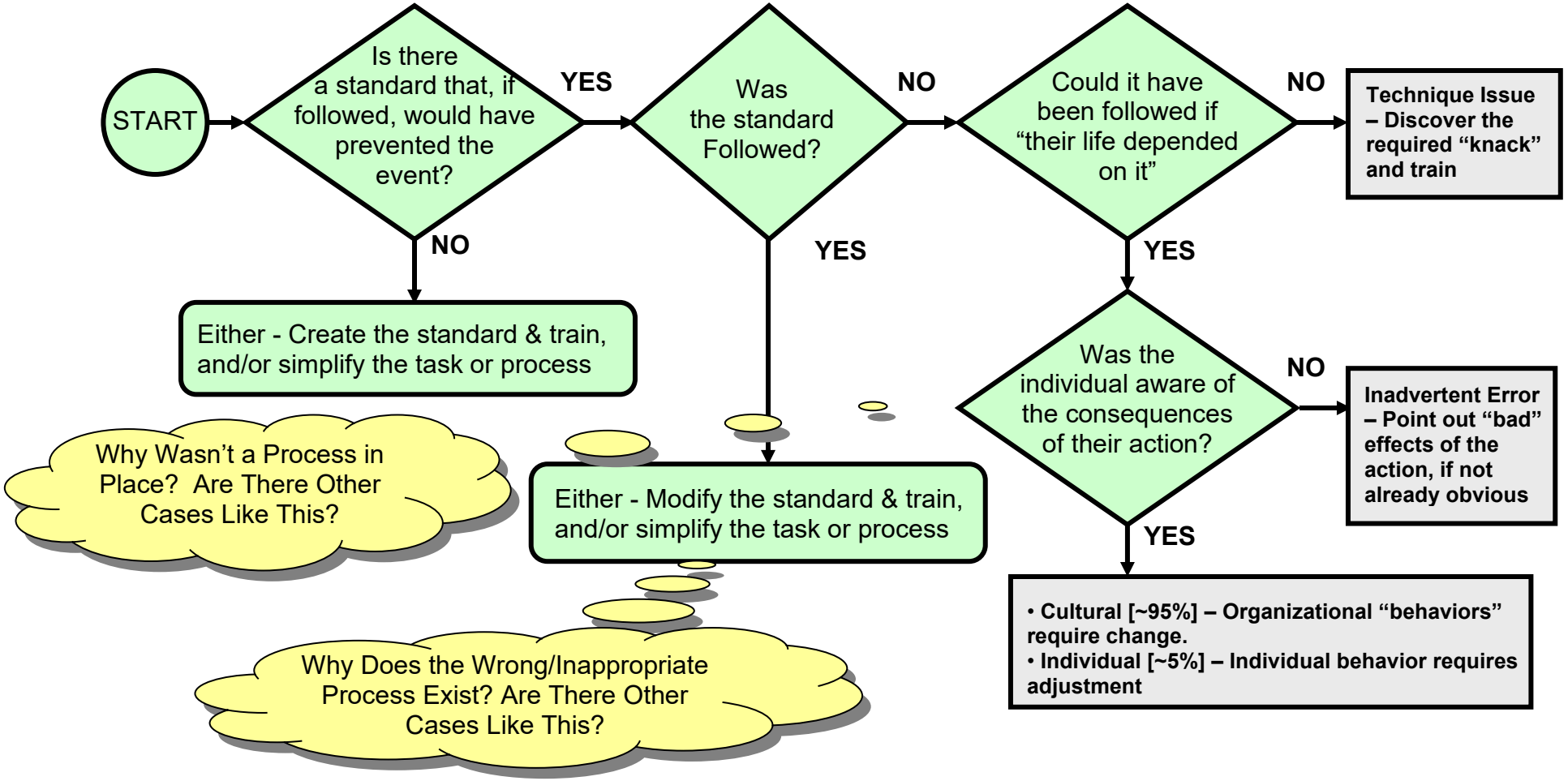
Many improvement opportunities arise because either:

a. a process does not exist to ensure the required quality of the product or service, or

b. a process does exist but is either not followed or is too difficult to follow.

The flowchart on the next page presents a series of questions when “problems” occur. Notice that there is both a *specific* response to the problem being addressed and a *broader* response that questions whether other situations like this exist in the organization. These broader questions often arise when an organization is just beginning to implement the philosophy of quality control & improvement.

Was the Process Followed?



STANDARD means policies, procedures, training, checklists and other items that are supposed to prescribe the way specific tasks are to be done.

The items shown in the gray boxes are "human factors" issues. Appropriate countermeasures would be generated directly from this analysis without the need for additional root cause analysis work.

Comparative Process Analysis

Comparative Process Analysis is often used when a specific problem occurs as a result of some work process. The analysis has three steps:

1. Identify how the process operated when the problem occurred,

2. Identify how the process operates when *no* problem occurs, and

3. Seek out the differences between the two. Adopt the best method of performing the process.

This type of analysis is especially useful in accident or injury investigations, although, it has broader application.

Examples

Sand Barge Overturning - During a "fill" operation of a barge with river sand, the barge capsized and sank. Only a few minutes before, a barge crewman was on the deck of the barge. Although the crane operator was experienced, a team of operators identified several key differences between other fill processes and the one that resulted in the capsizing. As a result of the investigation, the operators adopted several changes. Some of the changes improved the fill operation efficiency; some addressed the safety of the process.

Medium Voltage Electric Lines - Lineman Injury - During a routine repair of insulators on a 13.6 kV distribution line, a lineman fell from the pole. His harness caught him, but he still sustained a back injury. The process he used during the repair operation was documented and compared to that of other linemen. Several differences were noted; the injured lineman adopted changes in his work process.

Fuel Injector Failures During Application to Locomotive Diesel Engines - As part of the locomotive overhaul process, rebuilt fuel injectors are "applied" to the diesel engine. During a one-week period, it was noticed that a large fraction of the injectors appeared to be frozen when the engine was tested. All injectors came from one railroad vendor. The vendor's injector rebuild process was documented and compared to other vendors. It was found that they had adopted a new shipping container (required by the railroad) that packed the injectors in an upright position (previously, they had been shipped "loose."). Fuel oil introduced to the injector during the vendor's test was not draining from the injector when stored upright as it did when they were shipped "loose."

Twenty Questions

"Twenty Questions" is a method of breaking down a process and asking the *What, When, Where, Who* and *How* questions for each step. This describes the current reasoning behind the step. Each step is then subject to the *Why* question, i.e. *why* do the step, *why* do it then, *why* do it there, *why* do these people do it, and *why* do it this way?

The next round of questioning challenges the improvement team to either eliminate the step, do it another way, at another time, by another individual, etc. The results of the process step questions are then integrated into the new process. *Twenty Questions* is used to address both the quality of a service or production process, as well as the delivery (timeliness) of the process.

"Twenty" Questions
(Applied to Each Task of the Process)

	Present Method	**Reason**	**Other Choices**	**Method Chosen**	**Improvement**
Purpose	*What Happens?*	*Why do it?*	*Can something else be done?*	*What should be done?*	*Changes to the current process listed here.*
Place	*Where is it done?*	*Why there?*	*Can it be done some-where else?*	*Where should it be done?*	
Time	*When is it done?*	*Why is it done at that time?*	*Can it be done another time?*	*When should it be done?*	
Person	*Who does it?*	*Why that person?*	*Can another do it?*	*Who should do it?*	
Procedure	*How is it done?*	*Why do it this way?*	*Is there another way?*	*What should be done?*	

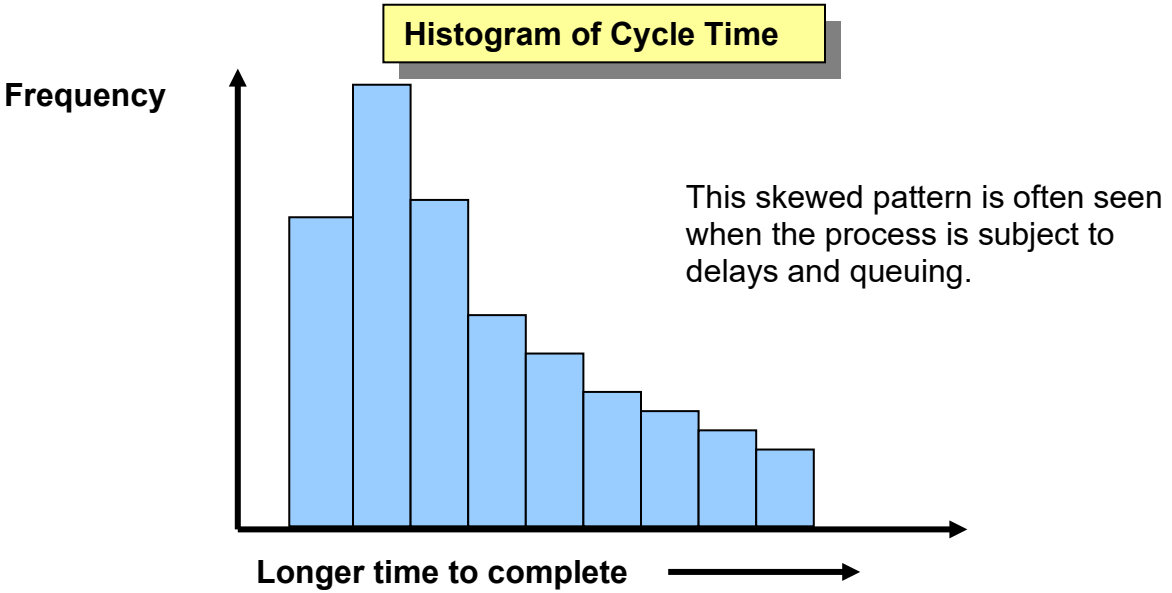
5.4.2 Cycle Time Analysis

Many projects focus on cycle time reduction and so we will include several methods for analyzing and improving this key characteristic.

General Considerations

The overall cycle time for a product or service must be operationally defined. Although there are many possible start and stop points, make sure that you consider the customer's point of view. For example, a factory may consider the "start" of a process to be when they receive an order. The customer's "clock" though, started when they placed the order with the salesperson.

Consider the variation in performing the overall process and its steps. The data should be plotted on a graph showing performance over time, i.e. a line graph, run or control chart. A histogram can also be helpful in showing the spread and may be a clue to causes of delay.



Value Added/Non-Value-Added Analysis

Often, a process contains a mixture of value-added activities – those steps that contribute to transforming the process inputs into a product or service of value to the customer and the *non-value-added* steps – those that do not contribute to increasing the value of the inputs. Typical non-value-added steps include *transportation, storage, waits, set-ups, rework.*

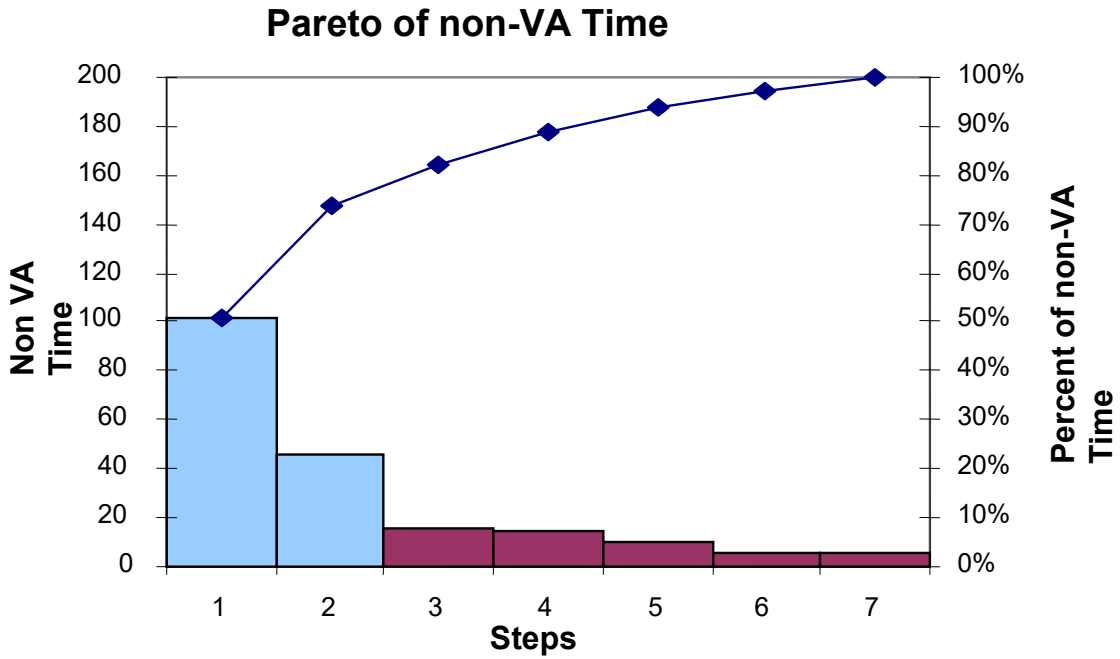
The following table can be used to categorize your process steps. Make sure that your flowchart identifies *all process steps;* sometimes it's easy to ignore these non-value-added steps. You can do this analysis on a step-by-step basis, or you may group several steps together into a segment. This analysis is often conducted as part of "leaning" a process (See Unit 5.5).

Value Analysis Table

Process Step/Segment	**1** Gather Receipts	**2** Open Exp. Account	**3** Record Event	**4** Enter Expenses	**5** Enter Charges	**6** Print Invoice	**7** Address Envelope
Time Analysis:							
Overall Time	30	10	20	60	30	30	30
VA Time	15	4	20	40	15	5	20
Non VA time	15	6	0	20	15	25	10
Step Category:							
Internal failure time (rework)							
External failure time (rework)							
Wait/delay time							
Prep time							
Transport time							
Value Adding							

Pareto Analysis

Pareto analysis (see Unit 7.2) is a valuable tool for prioritizing your cycle time investigation. Develop the Pareto of the *non-value-added time.* If you simply break down the time required to perform each step, you may focus your efforts on a step that takes a long time but is completely value-added time. You should try to reduce the non-value-added time before trying to speed up a value-added activity.



The Critical Pathway

When we discussed the flowchart, the notion of *series* and *parallel* activities was presented. This concept can be taken one step further to develop a picture of a process that is specifically designed to focus on the *time (and/or resources) required to complete a process.*

The Critical Pathway Method was originally developed to help manage large development projects. Critical Pathways are tools that allow project managers to "model" the development effort, arrange the development tasks in the most "optimum" manner, and understand how well the project is proceeding as the different tasks are completed. The *critical path* is defined as the sequence of specific activities that must occur for the project to finish and that takes the *longest* to complete. Let's illustrate this concept with a simple example. Every morning, we complete the following "project" - *Getting Breakfast Ready.* The following tasks are involved, with their times included:

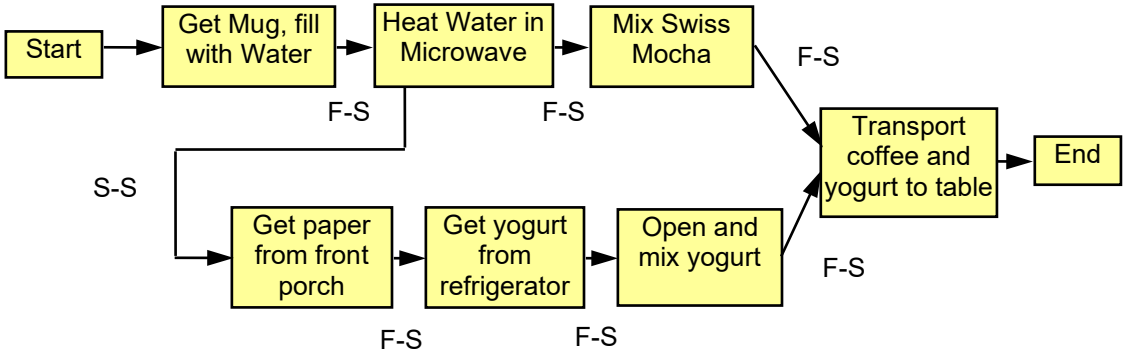
Task	Time
Get Newspaper from front porch	20 seconds
Get Yogurt from refrigerator	10 seconds
Open and stir yogurt	20 seconds
Get coffee mug, fill with water	15 seconds
Heat coffee water in microwave	120 seconds
Mix Swiss Mocha in coffee	20 seconds
Transport coffee & yogurt to table	10 seconds

Now let's say all these tasks are done in *series*, that is, the second task would not start until we'd finished the first (this is called a *finish-to-start* relationship), the third doesn't start until the second is done, etc. The total time to finish all the tasks would then be the sum of the individual task times: 215 seconds, or a little over three and a half minutes. This *series* sequence of tasks would be the current *critical path*.

But must all the tasks be done in *series*? Perhaps, while the water is heating in the microwave, some of the other things could be done, like getting the paper, the yogurt, etc. We've now put some of the tasks in *parallel* and have cut the total time. There are still some *finish-to-start* relationships, i.e. the Swiss Mocha can't be mixed until the water is heated, but we've taken some of the tasks *off the critical path.* There is a graphical picture called a PERT[2] diagram (some call it an

[2]PERT - Program Evaluation and Review Technique.

Arrow diagram) that you can construct for your processes, analyze and improve and manage by as you proceed through your projects. Here's what our "improved" Breakfast preparation process would look like on a PERT chart:



How do you interpret the PERT chart? Well, the boxes are all tasks/activities that require some *duration.* The arrows, combined with the *F-S, S-S, etc.* notes describe the *relationship* between tasks. For example, we would not START heating the water until we've FINISHED filling the mug. *F-S* stands for *Finish-to-Start, S-S* stands for *Start-to-Start, F-F* for *Finish-to-Finish*.[3] Note how we've described the relationship between heating the water and getting the paper as a *Start-to-Start.* As soon as we put the water in the microwave, we can get the paper. We don't have to wait for the heating task to *finish* before we *start* after the paper. We can predict the effect of our improvement from this picture. There are essentially two *paths* that must be completed for our breakfast to be ready. By calculating how long each of these paths take and finding the *longest time (i.e. the new Critical Path),* we will know how long the entire process takes, start to finish.

Although you might want to check our work (by calculating the time through the other path), the new *Critical Path* is:

Critical Path Task	Time
Get coffee mug, fill with water	15 seconds
Heat coffee water in microwave	120 seconds
Mix Swiss Mocha in coffee	20 seconds
Transport coffee & yogurt to table	10 seconds

[3]No, we won't discuss any *Finish-to-Swedish* relationships!

The time to perform these tasks is 165 seconds, or about two and three quarters minutes. We've cut the process time by about 50 seconds.

Analyzing Your Process Using *Critical Path Thinking*

You can directly apply this kind of thinking to your production processes, especially relating to the *time* quality characteristic:

Evaluating Improvements - If time is an important quality characteristic for your process, then the PERT diagram (combined with Critical Path thinking) can be used to evaluate ideas for improvement.

Suppose our "significant other" offers to get the paper while we work on the yogurt. Does this cut the overall process time? Here, the answer is no, because we **have not changed any of the critical path tasks!** The only changes that will improve the overall time are those that affect the critical path tasks or their times!

Slack Times - While it takes 140 seconds to heat the water and mix the coffee, it only takes 50 seconds for us to get the paper and the yogurt. There are 90 seconds of *slack time* in this path.

Often, a PERT chart will be analyzed for these slack times when there are constraints on the *resources* available to complete the work. People or machines can be shifted from project paths with slack time to those on the critical path (we need to be careful that such shifting doesn't create a *new critical path!)*.

Wasted Time - Suppose we were reading the newspaper when the microwave bell rings? If we continue to read and not begin the mixing task, then we're adding to the critical path time. Delays, "down-time," etc. are all factors that increase the critical path time, without adding any value to the process. These must be eliminated or minimized in both duration and frequency.

In the "real world," coordination between two departments often contributes to delays and wasted time.

Resources - The PERT diagram can be "resource-loaded," that is, specific departments, personnel, machines, etc. can be included in this process model.

Notice in the PERT chart that if we're mixing the yogurt when the hot water is ready, we can't really start mixing the coffee until the yogurt is finished. This is an example of how a resource can constrain the completion of a project, even though there is no relationship between the actual tasks.

Unnecessary Tasks - This one seems so obvious, but we'll include it anyway. Let's say we have a habit of pulling a dirty coffee mug out of the dishwasher. This would add an unnecessary task of cleaning the cup before we filled it with water. This task should be done the night before, not during our "breakfast project's" critical path.

Differences between Similar "Projects" - Of course, if there are several of us out there getting breakfast ready, and there are differences in our processes' performance, we can compare the critical paths to see what's happening within the process.

5.5 Lean Manufacturing

Learning Objectives

- Understand the differences between traditional mass production and lean production
- Understand the principles and practice of lean production
- Know how to diagnose the maturity of a lean plant in transition

Unit Contents

- Lean Manufacturing/Lean Production Overview
- Lean Principles and Practice
- Manufacturing Plant Maturity During Lean Implementation
- Summary of Lean “Rules”
- Software Support for Lean Manufacturing
- Lean Manufacturing Resources

Lean Manufacturing/Lean Production Overview

Lean Manufacturing (AKA the Toyota Production System) is, in its most basic form, the systematic elimination of waste (*muda* in Japanese), and implementing the concepts of *flow* and *pull* into production systems. The benefits of lean production systems can be as high as 50% lower production costs, 50% less personnel, 50% less time to field new products, higher quality, higher profitability, higher system flexibility, and others. The basic elements of Lean Manufacturing are waste elimination, continuous one-piece workflow, and customer pull. When these elements are focused in the areas of cost, quality, delivery, safety, and morale, the journey towards lean production has begun.

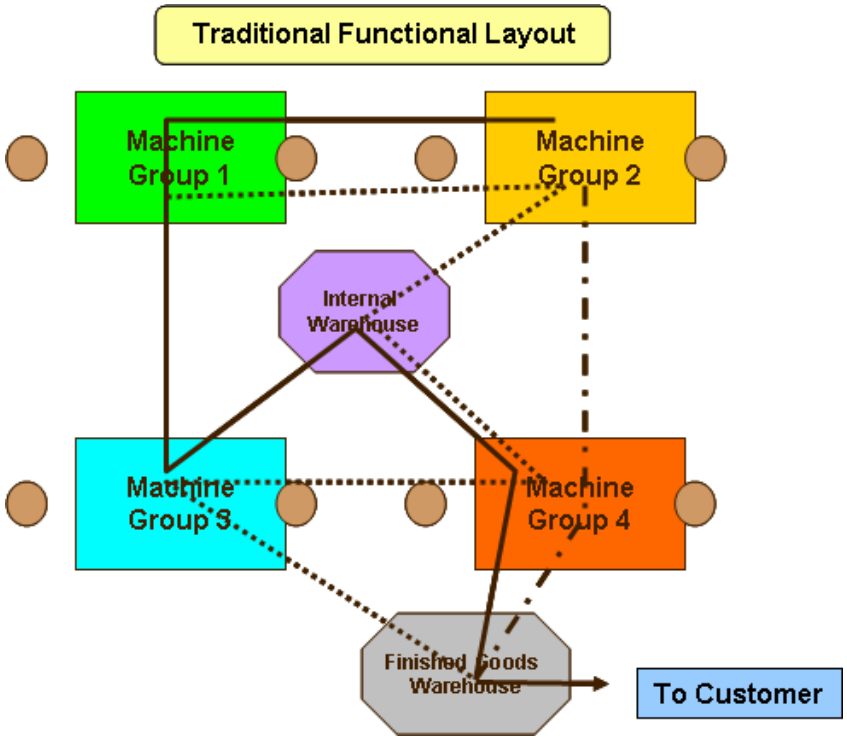
The "Lean Enterprise," a broader view of Lean Manufacturing, encompasses the entire production system, beginning with the customer, and includes the product sales outlet, the final assembler, product design, and all tiers of the supply chain (to include raw material mining and processing). Any truly 'lean' system is highly dependent on the demands of its customers and the reliability of its suppliers. No implementation of lean manufacturing can reach its full potential without including the entire 'enterprise' in its planning.

The Traditional Manufacturing Situation

Many manufacturing plants are organized around the following principles: functional layout – co-located processing machines such as lathes and presses to provide flexibility in making a wide variety of products, product routing through the plant, large batch manufacturing to achieve economies of scale and production schedules developed to meet projected demand.

The "Lean" Situation

By comparison, a lean manufacturing plant has the following characteristics: a process layout where processes are replicated and distributed throughout the plant and organized so that products can flow sequentially through the processes necessary to make them with little or no waiting time, single piece rather than batch flow. Instead of scheduling, the lean plant allows the customer to pull products from the system, sending a chain of work authorizations through the system in



the reverse direction of product flow, thereby directly linking production activity to demand. Quick changeover/set-ups are essential to support small lot production.

Lean production also implies a decentralized approach to process control and automation following a sequence of steps, where none is bypassed. Each step is taken when it is economically justified, and the automation level of the different machines is kept consistent. The improvements made to the shop floor impact a number of support activities:

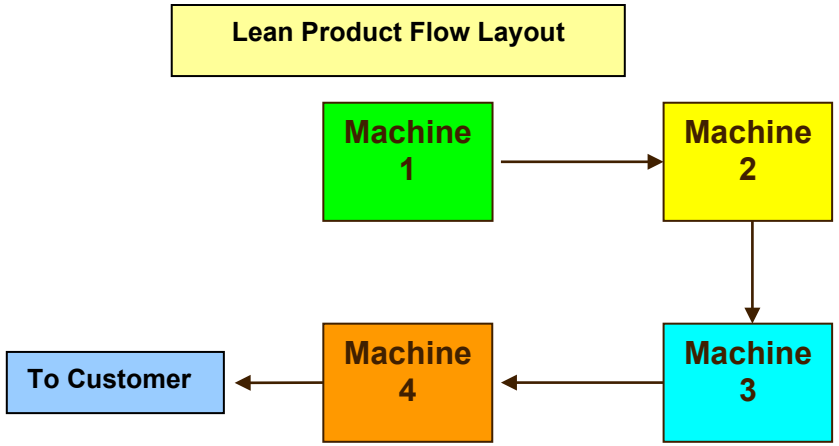
- For production planning and scheduling, this involves working back from customer order to leveled final assembly schedules, and Kanbans or timetables to assure parts supply.

- For quality assurance, the approach combines 100% go/no-go parts checking integrated into the process, with mistake-proofing (poke-yoke), an organization to collect and use customer claim information as well as help suppliers solve their quality problems.

- For maintenance, housekeeping and routine equipment checks are delegated to operators, and technicians are organized as "general practitioners" in charge of specific areas, while refurbishment or overhaul work is outsourced.

- The wage system is reorganized to reward ability as well as performance, to support the move to multifunction operators. Productivity improvement bonuses are awarded on a group basis to promote teamwork.

The conversion of a plant can be a multi-year effort but starts with a few pilot projects on the shop floor that pay back in a few months. The pilot projects teams are then used to seed new teams for a larger number of projects. Later, other aspects of the conversion require other implementation methods, such as kaizen teams, kaizen events, or a plant-wide coordinated effort involving all employees.



Lean Principles and Practice

1. Specify Value

The first step involves specifying what the customer *values*, so that the *wastes* (or *muda)* currently in the process can be eliminated. A company's customers judge whether value has been created. Therefore, one category of *muda* is having the "right" process for a product or service that the customer doesn't want. Lean companies therefore work to *define value in terms of specific products with specific capabilities offered at specific prices through a dialogue with specific customers (e.g. Voice of Customer processes).* In other words, they work to understand and deliver what the customer wants to buy. Lean companies often restructure based on product line, organizing managers and employees into product teams.

2. Identify the Value Stream

Once *value* has been defined it is time to begin the process of identifying and removing the waste that is involved in providing the products to the customer. The value stream can be defined as "*The set of all the specific actions required to bring a specific product through the three critical management tasks of any business: … problem solving, … information management, … physical transformation."* As you gain an understanding of what the value stream for a product is, you will discover three categories of activities:

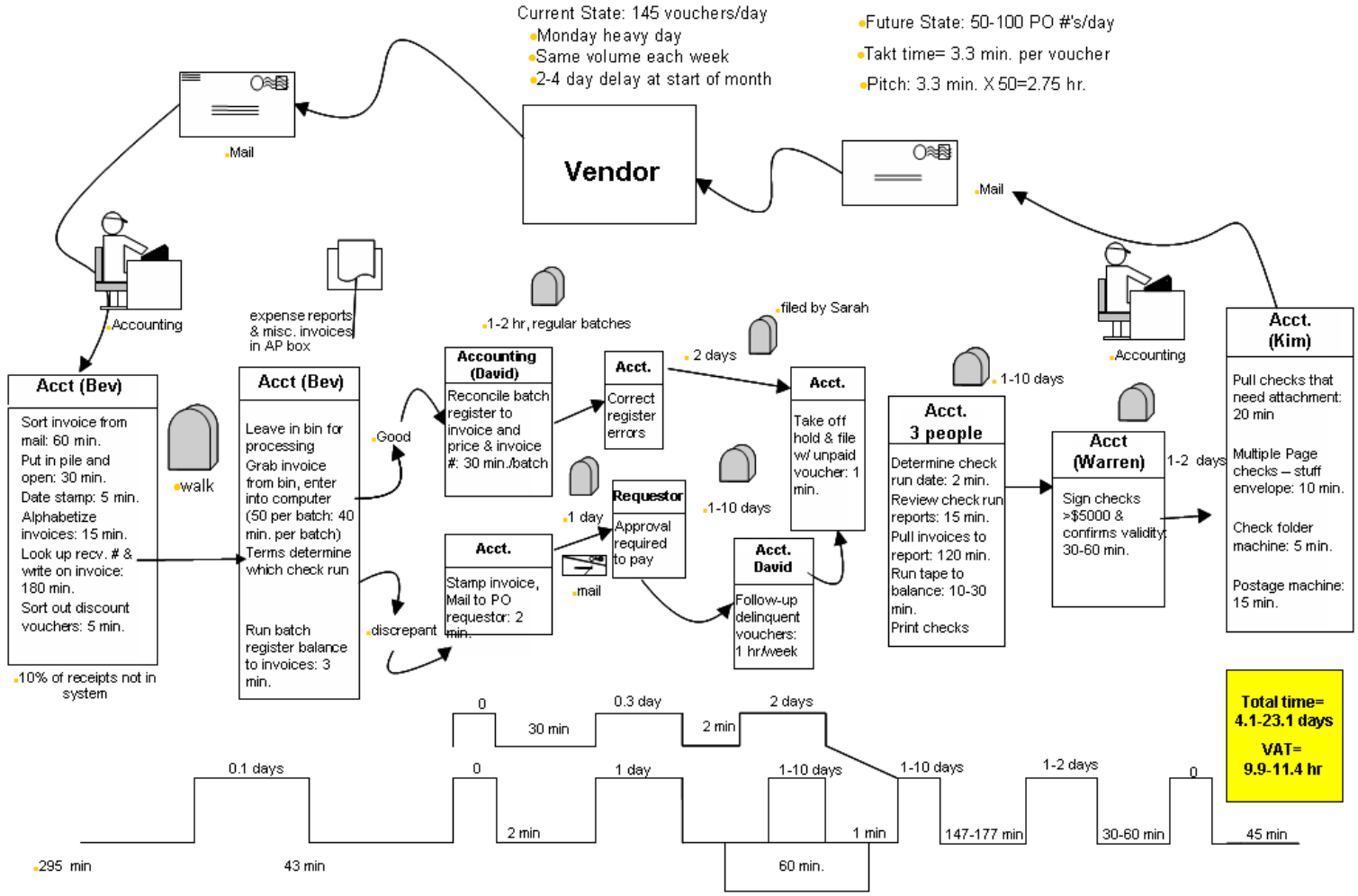
- *Steps that create value*. In the manufacturing process, these are the steps that are transforming the fit, form, or function of the raw material, and bring it a step closer to the finished product. Shigeo Shingo notes that only the actual transforming action is value-added. For example, in a bending operation, only the actual bending motion is value added. All other operations such as movement and setup of the piece are non-value added.

- *Steps that create no value but are necessary, due to current state of the system.* These might include inspection, waiting, and some transportation steps.

- *Steps that create no value and can be immediately eliminated*. If the activity clearly does not fall into one of the above categories, it should be stopped.

During this step in the process of becoming lean, detailed process flow diagrams are created for each product, highlighting all the steps that are considered to be *muda*. This is usually done in the context of *Kaikaku*—lean's term for radical improvement. Contrasted with kaizen, or continuous improvement, *kaikaku*, also known as *breakthrough kaizen*,

is an intense questioning and reexamining of every aspect of a process. Any steps that can be eliminated immediately are stopped. Any activities that are identified as "non-value but currently necessary" become targets for improvement. The "20 Questions" approach to process analysis (see Unit 5.4) can be employed to analyze the production process here, or the team can create a *value stream map*- one which shows the current process steps, who performs them and which are value-added or non-value added (the example below is a value-stream map created for an Accounts Payable Process).



This is also the point at which "target costing" is implemented. Target costing is a methodology in which the cost of a product is established based on its "*muda*-free" process. *What if we didn't have scrap? What if we didn't have to conduct receipt inspections?* This is now the cost that the company strives to achieve through the elimination of *muda*. As it gets closer to the target cost, the lean philosophy suggests that the company will then be able to enjoy increased profits, or to reduce its selling prices to its customers, thereby increasing value in the customers' eyes.

3. Flow

To document the process, Lean teams will physically walk the process, noting the distance the product must travel in order to go through its entire process. Some very small operations report that their process is over a hundred miles long, and it is estimated that the process of producing aircraft is tens of thousands of miles long! Even typical hospital patients are transported an average of 3-4 miles during a typical length of stay (not counting the distances their medications, lab samples, food, etc. must travel!).

With the process-specific *muda* identified and on its way to elimination, the purpose of this step is to encourage organizations to focus on *rapid product flow,* unencumbered by the walls and the physical distance that exist between typical functional departments. "Lean enterprises" are created for each product. The physical layout of the people and equipment involved in the process is changed. Factory floors are laid out in cells rather than in functional groupings, thereby reducing the distance parts must travel. Where before there were departments for engineering, scheduling, and customer service, lean enterprises have teams of people from each of those disciplines comprising the team responsible for the business of specific products.

Here, the **5S** principles are implemented to reduce the slack hidden in plants. **5S** is comprised of the activities listed below, which collectively translate to a cleanup activity at the workplace. The intent of **5S** is to remove the *muda* associated with clutter and disorganization.

- **Sort (Seiri)** – separate the necessary things from the unnecessary and discard the unnecessary,
- **Set in Order (Seiton)** – Neatly arrange and identify things for ease of use (a place for everything, and everything in its place),
- **Shine (Seiso)** – to always clean up; to maintain tidiness and cleanliness—to clear your workplace thoroughly.
- **Standardize (Shitsuke)** – to have workers make a habit of always conforming to rules.
- **Sustain (Seiketsu)** – To constantly maintain the 3S mentioned above, Seri, Seiton, and Seiso. Keeping a clean workplace without rubbish or oil leakage is an example of Seiketsu.

4. Pull

In the lean enterprise, inventory is waste. Therefore, producing anything that is not sold is waste as well, for if it's produced but not sold, it remains as finished goods inventory. Thus, it is important that real customer demand *pull* product through the system. This is in contrast with the traditional push approach to manufacturing where the system encourages each resource to produce as much as possible, *pushing* products through its system. Once the first three steps are implemented, this concept is especially important. Because the process is shortened when wasteful steps, wasteful activity within steps, and distance parts must travel is removed, lean organizations usually find themselves with the capability to produce more than before. In a push environment, such capability would translate into increased inventory—not exactly lean. In a pull environment, this tendency to overproduce is controlled. Activities may then be directed toward either removing excess capacity or increasing the rate of pull.

Today's information technology makes it possible for more systems to transition from the push mentality embodied in the traditional approach of manufacturing and distributing products based on forecasts (see Lean Software below). In an era of dynamic markets, where last year's demand in no way reflects what will happen this year, the traditional push approach places undue weight on historically based forecasts. In today's world, to the extent that the manufacturing and distribution system is responsive, it is far more effective to manufacture based on actual customer demand. Point of sale terminals provide the capability to capture in detail exactly what was sold and pass that information back through the supply chain to the distributors, manufacturers, and even to vendors. The practice of pull is made operational in lean enterprises with two methods, *takt time* and *kanban*.

Takt Time - *Takt time* is used to set the pace of production by matching production rate with the rate of customer demand. The *takt* time is calculated by dividing the available production time by the rate of customer demand. For example, for a plant that operates on a single 8-hour shift (480 minutes) with a demand of 240 units/day, the *takt* time is two minutes. Knowing this time is significant, in that it provides a sense of the desired pace of a plant's output. The point is always to define *takt* time precisely at a given point in time in relation to demand and to run the production sequence to *takt* time. In a lean enterprise, the goal of every resource at every step along the process is to produce at the rate demanded by *takt* time. Often the *takt* time and each resource's progress relative to this target is posted and displayed. Meanwhile, many manufacturing environments currently lack the flexibility to respond to frequent changes in *takt* time. The variation is *muda*, and becomes a candidate for improvement teams to eliminate.

Single Piece Flow - The following quote describes material flow after a lean implementation in a bicycle plant. “In the continuous-flow layout, the production steps are arranged in a sequence, usually within a single cell, and the product moves from one step to the next, one bike at a time with *no* buffer of work-in-process in between, using a range of techniques generically labeled “single-piece flow.” The lean philosophy considers any idle inventory to be muda. With the combination of *takt* time and single piece flow, the lean enterprise strives to achieve no idle inventory. Often, companies implementing lean begin with kanban systems. Kanban places small queues of inventory that are of a predetermined size at every resource. The job of each resource is to work to fill the queues to their predetermined size. When the queue is full, the preceding resource stops. In single piece flow, the queue is zero. The preceding operation works when the resource it feeds is ready to pull the piece directly. Single piece flow and kanban enable “pull” by effectively “stopping the push.” Workers know when to start working, and they know when to stop working. Idle time on the part of workers is *muda* in a lean environment. When a worker is prevented from working due to a full queue in a kanban system, or a downstream resource not ready to pull a part in a single piece flow system, idle time occurs. Elimination of this idle time is yet another candidate for “*muda*-attack” in a lean environment.

5. Perfection

The initial successes that are achieved because of implementing the first four steps highlight new opportunities for improvement in reducing effort, time, space, cost, and mistakes while offering products and services that more closely reflect what the customer really wants. This step serves to remind us that continuous improvement is possible and is the desired state of any change in any environment. To keep the pump primed for perfection, mature lean organizations practice open book management and work hard to instill the spirit of *kaizen* or continuous improvement.

Lean Personnel Model

As lean is implemented in a manufacturing environment, processes will eventually require fewer workers. In effect, there won't be enough work to keep everyone on the shop floor gainfully employed producing parts, unless significant growth occurs. In this situation, layoffs must be avoided at all costs. How, then, are these 'excess' personnel best utilized?

- One common approach is to form an office or department to coordinate all continuous improvement activities. Some organizations call this the 'Lean Office', or Continuous Improvement Office, or something similar. This department should be made up of a director/manager, and several Improvement Teams.
- Determine what the current minimum acceptable personnel level is. This is defined as the minimum number of personnel needed to perform production operations under current conditions. Remove the excess personnel

immediately (taking the best personnel out of production) and attach them to the Lean Office (they will form the initial core of the Improvement Teams).

- The Lean Office should begin a rotation involving 1/3 of their time in 'in-house' lean training, 1/3 of their time in continuous improvement activities, and 1/3 of their time participating in outside activities (ie, touring other organizations, etc...)

The Director oversees several Improvement Teams, coordinates the training and outside activities of the teams, and sets improvement goals (which support organizational goals). The teams will vary in number, according to the number of manufacturing cells, product lines, or functional departments in your organization. They are made up of a Senior Engineer and a Junior Engineer (these are the only permanent staff of the Lean Office), and 4 to 7 shop personnel. The Continuous Improvement (CI) Engineers are subject matter experts in various 'lean' subjects, and the totality of the CI Engineers make up the training staff for the teams. The staff trains the team in all aspects of lean manufacturing (as a group) and the CI Engineer teams coordinate the CI activity in the manufacturing cells, product lines, or functional areas. In a Six Sigma organization, Black Belts may be assigned to these CI engineer roles.

The members of the Improvement Teams (those from the shop floor) should be assigned to the teams for a predetermined period (a minimum of 3 months, a max of 6 months). During this rotation, the teams should participate in Lean/CI activities in all or most of the manufacturing cells, product lines, and functional areas. At the end of this period, these personnel should be rotated back into the workplace, and the next best individuals are rotated out into the CI teams/Lean Office.

By this time, the company has created a group of 'lean supermen', who will form the basis for a long-lasting, continuous improvement effort on the shop floor (even without the direction of the Lean Office). The constant rotation of personnel in and out of the Lean Office will ensure a workforce that fully understands the goals and practices of lean manufacturing, and will support the company’s efforts.

Some notes -

1. The acceptable or target personnel levels must constantly be watched and kept up to date (preferably by the Director of CI or the Lean Manufacturing Manager)

2. Eventually, the Lean Office will have very few shop personnel, due to the attrition factor, but will still conduct improvement activities within the cells.

3. It is recommended that the Senior CI Engineer of each team focus on waste reduction as his primary focus (which is 80% of what lean manufacturing is) and the Junior CI Engineer focus on other areas, such as implementing flow and pull systems.

4. It is important to assign the BEST personnel to your Lean Office initially. These are typically those who have high motivation, critical thinking skills, understand the product and processes, etc... This will help guarantee the success of the effort and ensure the momentum of the program.

5. The 'rotation' period (within the Lean Office) ensures that ideas, exposure, and best practices from all areas of the company are adequately circulated (this is truly use of 'intellectual capital').

A Tactical Approach to Evaluating Lean Progress

The lean transition is one that may take several years to fully implement. The plant will progress through a number of changes and improvements during this time. The "maturity" of the plant relative to a truly lean organization can be tracked and reported. The Production System Design Laboratory of MIT (Massachusetts Institute of Technology) has developed a diagnostic model that supports this goal.

The diagnostic (double-click on the embedded file below) shows the level of successful implementation for the lean principles listed below. The upper-level functional requirements are listed for each principle. Then, the functional requirement that is to be evaluated is stated in the Evaluation Criteria row of each example. Below that, there are 6 descriptions that correspond to the 6 levels of achievement described below (see *Manufacturing Plant Maturity During Lean Achievement)*. The state of the plant is matched with the closest description. "Partial" implementation in one part of the plant may also be assessed.

Lean Principles

Maximize Sales Revenue - Produce to Maximize Customer Satisfaction

- Deliver no Defects - Defect Free Production

- Stabilize Processes - Eliminate Assignable Causes of Variation
    - Eliminate machine assignable causes
    - Eliminate operator assignable causes
    - Eliminate Method Assignable Causes
    - Eliminate Material Assignable Causes

- Deliver Products on Time/Throughput Time Variation (σ) Reduction
    - Respond Rapidly to Production Disruptions
    - Minimize Production Disruptions

- Deliver Products on Time/Throughput Time Mean (X-bar) Reduction
    - 1.1.3.1 Reduce Run Size Delay
    - 1.1.3.2 Reduce Process Delay

  - Reduce Lot Delay
  - Reduce Transportation Delay
  - Reduce Systematic Operational Delays

Minimize Direct Labor Cost - Eliminate non-value adding sources of cost

- Reduce Waste in Direct Labor Cost - Eliminate non-value adding manual tasks
  - Eliminate Operators Waiting on Machines
  - Eliminate Wasted Motion by Operators

- Reduce Waste in Indirect Labor Costs - Reduce Indirect Labor Tasks
  - Eliminate Managerial Tasks
  - Eliminate Information Disruptions

Minimize Production Investment

The Pie Chart Scoring Method assesses the level of "leanness" that a Production System Design has achieved. Under each column, there are 6 descriptions that correspond to the 6 levels of achievement. What is seen in the plant is matched with the closest description. The pies represent the percentage of the plant at the level indicated.

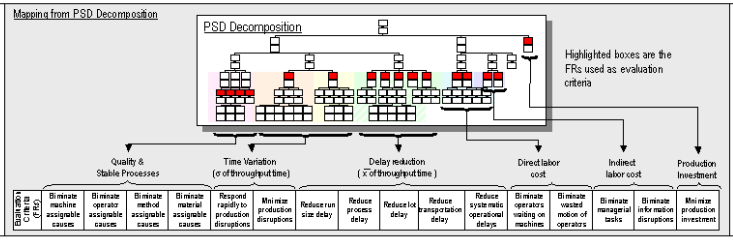
Upper level PSD Decomposition (FRs/DPs)	FR: Maximize long term return on investment / DP: Manufacturing system design																
	FR: Maximize sales revenue / DP: Production to maximize customer satisfaction												FR: Minimize direct labor cost / DP: Elimination of non-value adding sources of cost				Minimize production investment
	FR: Deliver no defects / DP: Defect free production				FR: Deliver products on time / DP: Throughput time variation reduction		FR: Meet customer expected lead time / DP: Mean throughput time reduction						FR: Reduce waste in indirect labor / DP: Elimination of non-value adding manual tasks		FR: Reduce waste in indirect labor / DP: Reduction of indirect labor tasks		
	FR: Stabilize Process / DP: Elimination of assignable causes of variation																
Evaluation Criteria (FRs)	Eliminate machine assignable causes	Eliminate operator assignable causes	Eliminate method assignable causes	Eliminate material assignable causes	Respond rapidly to production disruptions	Minimize production disruptions	Reduce run size delay	Reduce process delay	Reduce lot delay	Reduce transportation delay	Reduce systematic operational delays	Eliminate operators waiting on machines	Eliminate wasted motion of operators	Eliminate managerial tasks	Eliminate information disruptions	
Level of Achievement (DPs) 1	Poor quality output from machines due to unknown causes of variation (unable to hold mean). No maintenance of machine to ensure quality.	Workers learn tasks by watching others. Tasks may be done differently each time.	Methods are unstable and ill defined. Variation in methods is arbitrary and is not visible.	High variation in incoming parts which cause quality problems. Materials are damaged in storage and transport	Production disruptions occur frequently. Operators work around these disruptions so they are hidden.	Unpredictable resources. Disruptions are frequent and impact delivery.	System is designed to operate based on forecast demand, not actual demand. Production in large run size batches to avoid the long setup times.	Machine capacity and rate independent of demand (maximize output). Large and unpredictable levels of WIP between processes to manage system and avoid starvation.	Large transportation lot sizes between machines or processes to reduce transportation costs.	Mfg. Process focused layout. Machines arranged by function in isolated departments (job shop type). Complex material flow.	Processes must be interrupted frequently for routine tasks such as material handling, machine maintenance, chip removal etc.	One person, one machine design. Operator watches machine run.	Excessive walking required to search for tools and materials.	Vertical organization with many levels of management. Changes are slow to implement as they require review and authorization from many people.	No systematic information system in place. Difficult to understand production status.	Machines dedicated to part type are designed to run as fast as possible. Flexibility for future design and volume changes are not considered.
2	Some assignable causes of variation are identified. Maintenance is occasional and is not scheduled.	Limited amounts of formal skills training. Workers learn what tasks to do from instructions.	Methods are known but not documented (shop floor "tribal knowledge")	Parts arrive with questionable quality and must be inspected before use. Entire lots are sent back if there are bad parts	End of line inspection is used so quality problems are found late. Slow response to problems.	Disruptions are frequent but do not impact delivery often due to large buffer sizes. Frequency and type of disruptions is unknown.	System is designed for production plan based on forecast demand. Run size is based on <1 month's forecast demand.	Machines/processes in functional departments are arranged for product flow. High levels of inventory required between departments (carrying production rates)	Single piece flow in only some areas (like assembly). Upstream processes still deliver materials in large lots.	Process focused layout with departments grouped to reflect product family sequence of operations. Parallel processing occurs. Routing is unclear. (focused factory)	Routine tasks are designed so that they may be done infrequently - loading lots of material, large reservoirs for chips, infrequent maintenance.	One person, one machine design. Operator waits on machine when running, or does "fill-in" work when available.	Excessive walking required to obtain tools and materials that are not located at point of use. Few improvement activities.	Management and support groups organized by function. Lengthy time is required for tooling changes, facility improvements etc.	Information system (MRP) in place but not all employees understand or trust it. Constantly erroneous or out of date.	High speed batch production machines dedicated to part families enables development of lower cost tools and fixtures
3	Most causes of variation are identified but are still not eliminated. Maintenance is only in response to quality problems.	Formal skills training program in place. Workers learn tasks from senior workers. Work standards exist but methods still vary.	Methods have been defined and standards exist but are not always followed/ updated.	Supplier responsible for meeting specifications. Little inspection of incoming parts required. Some parts are still damaged within the plant.	Quick response to production disruptions (when they are found) to continue production. Root cause is not eliminated so problems may reoccur	Machine disruptions (MTTF, MTTR) are recorded and used to determine lead time required for predictable delivery.	System is designed for production plan based on forecast demand. Run size is based on <1 week's forecast demand.	Assembly or transfer line designs running at high speeds feeding multiple customers. Large amount of inventory before and after lines to manage product flow	Single piece flow in only some areas. Upstream processes deliver materials in lots based on standard inventory.	Product or customer oriented material flow. Machines/stations in cellular design with some batch processes/monument machines intermixed (partial cells)	Many routine tasks are scheduled so that they may be done after hours (like maintenance). Production still must stop regularly for other activities.	Operator runs more than one machine of the same type of manufacturing process. Machines do not have self-stopping capability.	Workers isolated to stations to avoid walking. Workers must prepare materials before adding value. Poor ergonomics causes worker to constantly reposition.	Support groups organized by product (IPT) and are co-located. Team members still evaluated by discipline.	Information system in place (MRP) to plan and schedule production. System requires team of expeditors and rescheduling to control production.	Equipment is reconfigurable to different part types. Machines still not designed to run at takt time in product oriented flow layouts.
4	Most causes of variation eliminated, some causes are still unable to be removed.	Manual tasks are defined so that they are done the same way each time. Standard work instructions indicate how tasks are performed.	Methods are well defined and repeatable. Standardized and followed.	Supplier responsible for meeting specifications. Little inspection of incoming parts required.	Production disruptions are addressed quickly and the root cause is eventually addressed. In-process inspection.	Disruptions from machines/ equipment, people, parts and information availability are known. Systems being developed to make resources more predictable.	System is not based on actual demand. Run size is based on a schedule that repeats at one or greater than one day. External setup tasks are reduced.	Customers grouped to achieve effective takt times. Machines and people are capable of working to takt time. Some parallel processing/stations exist.	Single piece flow within cells/sub-systems. Large lots transferred between sub-systems.	Material flow/cellular design with machines/stations close together. No monument machines that complicate material flow (well defined cells)	Machine/processes designed so they do not have to be interrupted for routine material replenishment, chip removal and work preparation.	Operator runs multiple machines of different types with self-stopping capability. No operator-process work routine defined.	5S program implemented so that parts, tools and equipment are where they are required. Some improvements in ergonomics are made.	Members of support group teams are dedicated to products. Teams formed on the shop floor and understand their basic performance measures.	Visual displays are installed. Not all employees understand how the display represents the factory status.	Ergonomics improved so that several machines may be operated by one person. Machines allow cells but are not right sized and so cause excessive walking.
5	Causes of variation reduced so that machine output is stabilized and mean shifts rarely occur. Maintenance is scheduled and performed on time.	Formal training is extended beyond skills to OJT by certified instructors. Standards are followed and upgraded by workers.	When methods are improved or updated, they are documented and implemented.	Collaboration with suppliers to ensure quality. Material handling and storage designed to maintain quality of products.	System designed so that production disruptions are visible. In-process checks by operators so quality issues are found quickly). Good root cause analysis.	Systems designed so that disruptions from all resources are reduced. Includes perfect attendance, TPM, std. material supply and reliable information systems.	System supports actual demand and expected peaks. Produce exactly what is consumed by the customer on a daily or shift basis. Internal setup tasks are reduced.	Cells/sub-systems running at takt time with standard inventory of one between stations. Machines and people are capable of working to minimum takt time.	All cells/sub-systems using single piece flow. Transfer lots sizes between sub-systems being reduced based on demand interval.	Material flow oriented layout design applies throughout value stream. Minimum handling with process close to receiving and shipping docks. (linked cells)	Workers continually make improvements in eliminating interference between people, material handling, maintenance etc.	Multi-skilled operator runs several machines/ processes. The operator-process work routine time graph defined and used.	Workers continually make improvements to decrease wasted motion and improve ergonomics. Machines are small and close together to allow reduced walking distances (in cells).	Work teams track their own performance and accept responsibility for maintaining and improving them.	A standard system for visual management is in place. All employees are trained in use of the system.	Machines support layout and operation (cells, automation). Equipment has product flexibility and is reusable.
6	Machines able to maintain mean, within tolerances. All assignable causes of variation eliminated or controlled through a regular maintenance program throughout system.	In addition to level 5, Any mistakes are not translated to defects through mistake proofing (poka-yokes)	Methods are continually being improved and implemented throughout organization. All employees are knowledgeable about the most current methods.	Collaboration with suppliers to improve quality and involvement in developing specifications. Parts transferred and stored to prevent damage.	Co-location of cause and effect (simplified material flow) and systematic method for communicating, and solving problems. Line stop methods in use (Andon)	Production disruptions rarely occur. Throughput time variation is very low and predictable.	Production of the desired mix and quantity during each demand interval using Heijunka. Almost no setup required between part types.	Production balanced to takt time throughout value stream. Some flexibility to produce to different takt times. Minimum WIP between processes & sub-systems/cells.	Single piece flow of parts throughout the factory.	Reduced transportation throughout supply chain. Production near customer Heijunka box. Material flow oriented layout throughout plant.	Processes rarely ever stop for routine activities.	The operator- process work routine defined and used. Number of operators may be varied to achieve a range of takt times. Machines run autonomously	Level 5 applied throughout company in addition to 5S understood by all employees and applied beyond the shop floor.	Self directed work teams are fully accountable for their performance. Applied across shop floor and support groups.	Most information about the system may be derived visually and understood by all the employees. Abnormalities get quick reaction by the correct people.	Machine designed to run at a range of takt times. In-house capability to modify standard machines quickly when required.
Comments	Refers to quality reliability of the machines. Assignable causes are those that cause the process to go out of control and may be: tool wear/breakage, bearing failures, etc. Maintenance in this branch refers to that which maintains quality instead of those that prevent breakdowns.	Refers to attaining predictable quality output from the workers. This is done through training, defining and following standard work and preventing common human errors from translating to defects/quality issues.	Method's are how process are done and include assembly tasks and process plans for machining etc. To prevent variation, these methods must be defined and followed. (standardized)	Material problems may be from suppliers or from handling within the plant.	Production disruptions refer to machine breakdowns, quality issues, information, worker availability. This branch deals with how these problems are addressed and eliminated.	Predictable resources are required for minimal production disruptions. Total preventative maintenance (TPM) programs, perfect attendance, reliable part supply and and information are required.	Delay due to inventory when producing different part types. Setup time is the time to change between part types within a station. Knowledge of the part types required is necessary or large inventories of different parts are required. With accurate and timely info, only the parts required are produced.	Process delay occurs when arrival rate > service rate (unbalanced rates). Large and unpredictable levels are indicative of process delay. It is reduced by balancing to a constant and predictable takt time. Flexibility to produce to different takt times keeps system balanced when demand changes.	Lot delay refers to parts waiting on other parts in the lot before they are transported together. This is avoided with single piece transport. Reducing transportation distance is important to achieve this.	To reduce transportation delay, the amount of transportation should be minimized. In system design, this is a consideration in factory layout. This is also applicable from a geographical view as well as within the plant	Operational delays are routine disruptions designed into the system (processes stopping for maintenance, material replenishment and other processes). It may be eliminated through machine and station design (Chip removal, control access from rear of station), operator work routine design.	Applies in fabrication or assembly when parts require multiple processes using different machines. Autonomous machines stop automatically when cycle ends or when a problem is detected.	To eliminate wasted motion, walking distances are decreased, 5S applied so tools, materials are readily available and ergonomics are improved.	This branch deals with the development work-teams, reducing the amount of management required. In the support groups, teams refer to the integration of engineers, production control, management, quality, scheduling. Their ability to make changes in order to support production is important.	To reduce the amount of tasks required to process information, visual management is used to simplify the system. This only works well when everyone understands how it works so that they can keep it accurate and respond to abnormalities when required.	Investment decisions are largely dependent on how the system is designed. Equipment should support this design and have the flexibility for expected volume changes, design changes and layout reconfiguration changes (cycle time/product flexible & small movable machines)
Metrics	# of non-conformances caused by machines	# of non-conformances caused by operators	# of non-conformances caused by method used	# of non-conformances caused by incoming parts	Average time required to eliminate root cause	# and time impact of production disruptions	Average run size	Cycle time / takt time of 1 for each station / system	Transfer lot size	Distance parts must travel through system	% of time down for routine tasks	% of time waiting on machine	% of time required for movement	% of employees in self directed work teams	% of system managed visually	% of value-added capital investment % of machines operating at or less than takt time
	Process capability Number of defects				% on-time delivery of parts / orders and variation of throughput time		Customer lead time and mean throughput time					% of value adding operator time		% of indirect employees		
	Sales Revenue											Production Cost				
	Return on investment															

Motivation

This Evaluation Tool is based on the "Production System Design Decomposition v5.0" which reflects the design of a Lean system as demonstrated by the Toyota Production System. This decomposition however, has identified how each element of TPS integrates into a system to satisfy business goals that are applicable throughout repetitive discrete production systems. The production system is evaluated in terms of how elements of the system satisfy the requirements defined for a "Lean Production System" The objectives are to:

- Show how well a system is designed and operated (current status)
- Identify areas for improvement when current design does not satisfy the objectives of a "lean" system (current problems).
- Indicate how system requirements may be satisfied using "lean" methods (future objectives)
- Provide a method to track the progress of improving the system design

Mapping from PSD Decomposition



Instructions

Steps:

1. The system to be evaluated must be first defined; it may be a cell, product line, or entire plant.
   Definition of system:_____
2. Evaluation is done column by column. If a description matches a percentage of the system, the percentage is indicated by filling in the appropriate portion of the pie chart in box. A typical evaluation is shown to the right as an example.
3. To make quantitative assessments, the metrics for each FR may be used.
4. As changes are implemented, the impact on the metrics and design evaluation may be tracked.

Example of Evaluating a Column

	Reduce process delay
1	Within the plant is still organized as a job shop with machines running at independent rates.
2	
3	Within the plant the A) is running as a transfer line feeding cells B, C and D. Large amounts of inventory after the line are required so that the cells are not starved.
4	One half the plant has been organized into cells running at takt time but have not achieved standard inventory of one between stations.
5	
6	

PSD Production System Design Laboratory
Laboratory for Manufacturing and Productivity
Massachusetts Institute of Technology
Prof. David S. Cochran
77 MASSACHUSETTS AVENUE, BUILDING 35-135
CAMBRIDGE, MA 02139
PHONE: (617) 253-1811 FAX: (617) 258-7488

Production System Design Evaluation Tool v5.0

D. S. Cochran, A. Wang

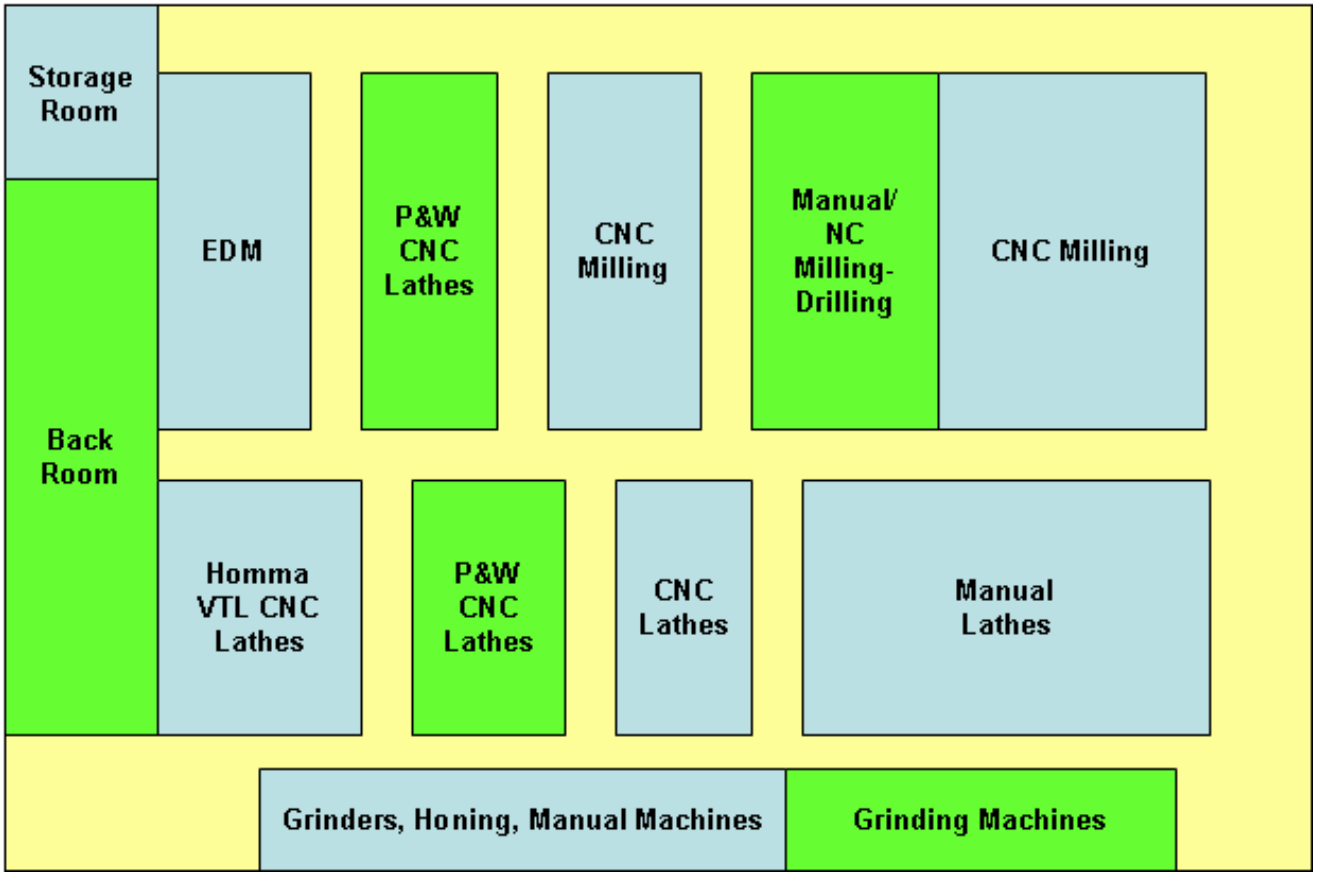
©MIT 1999

DATE/TIME	FILENAME	PAGE	OF PAGES
5/17/99 11:46 AM	PSD Evaluation Tool v5.0.ppt	1	1

Manufacturing Plant Maturity - Lean Implementation

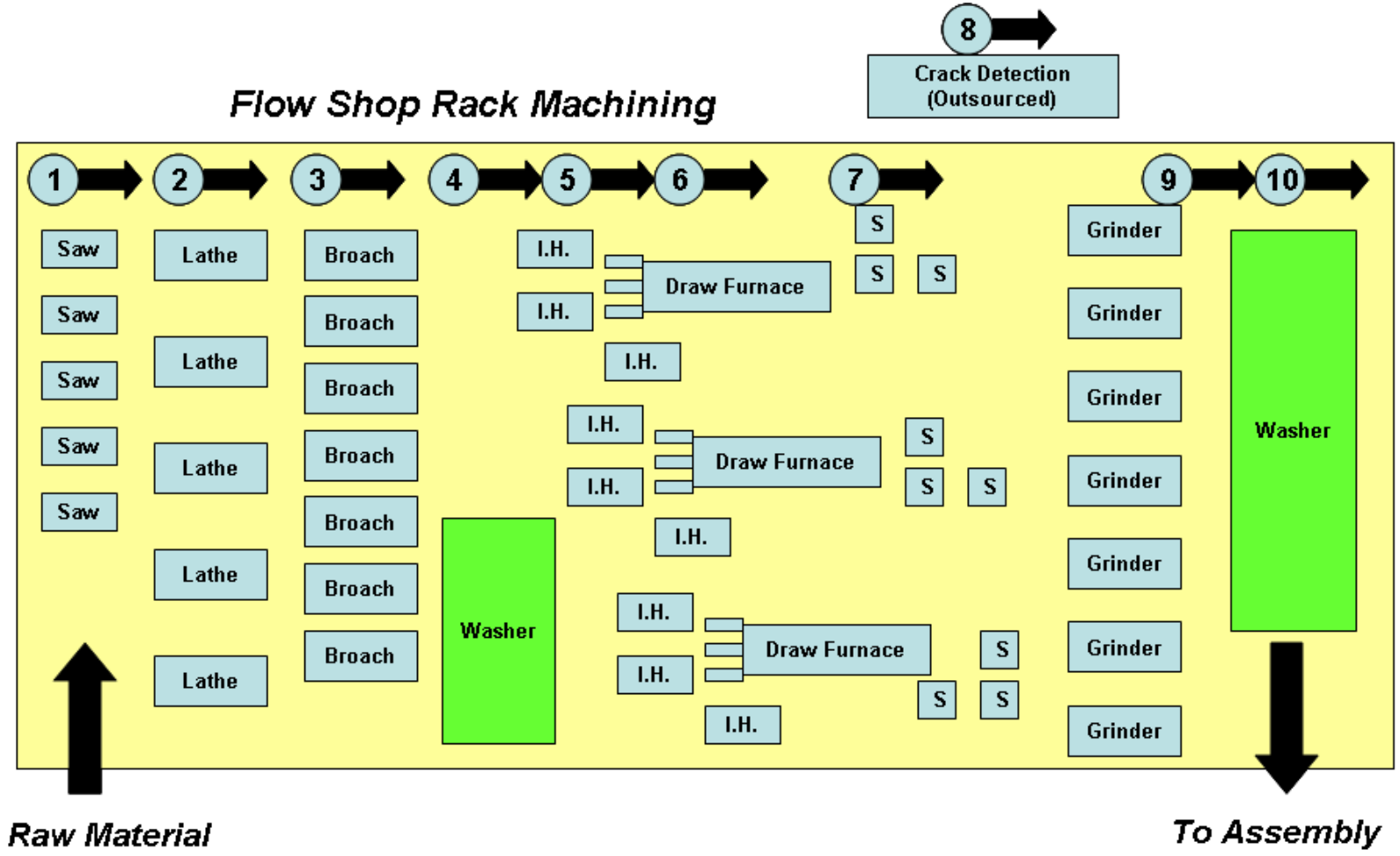
Level 1: Job Shop or Departmental Layout

This is the “traditional” layout – machines have been organized by their type, not by how the workflows through the shop.



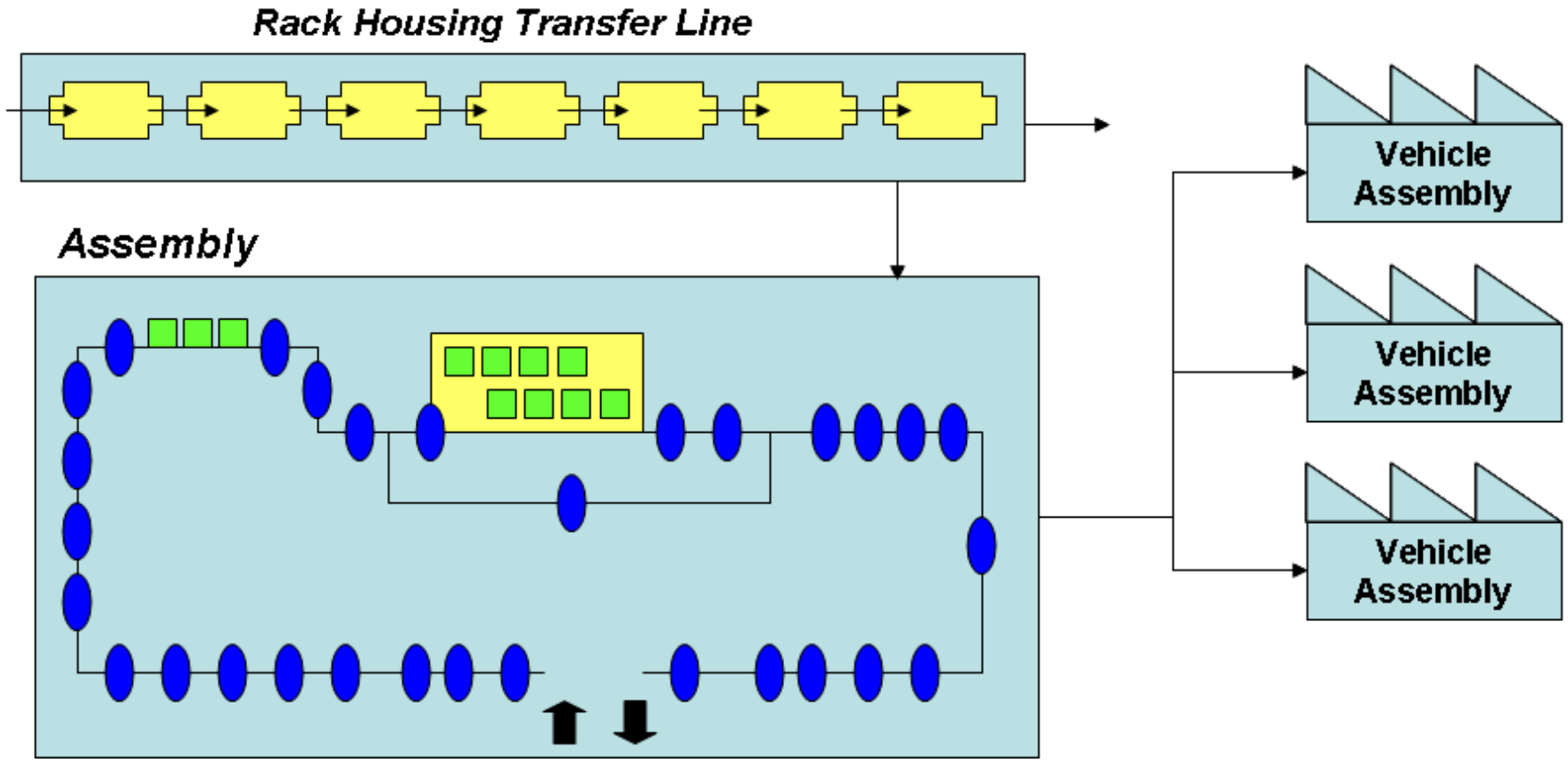
Level 2: Departments Arranged by Product Flow

Although the machines are still grouped together in this layout, they are now located by the order in which the product flows through the plant. Reductions in travel distance are achieved, but in-process queues may still be large.



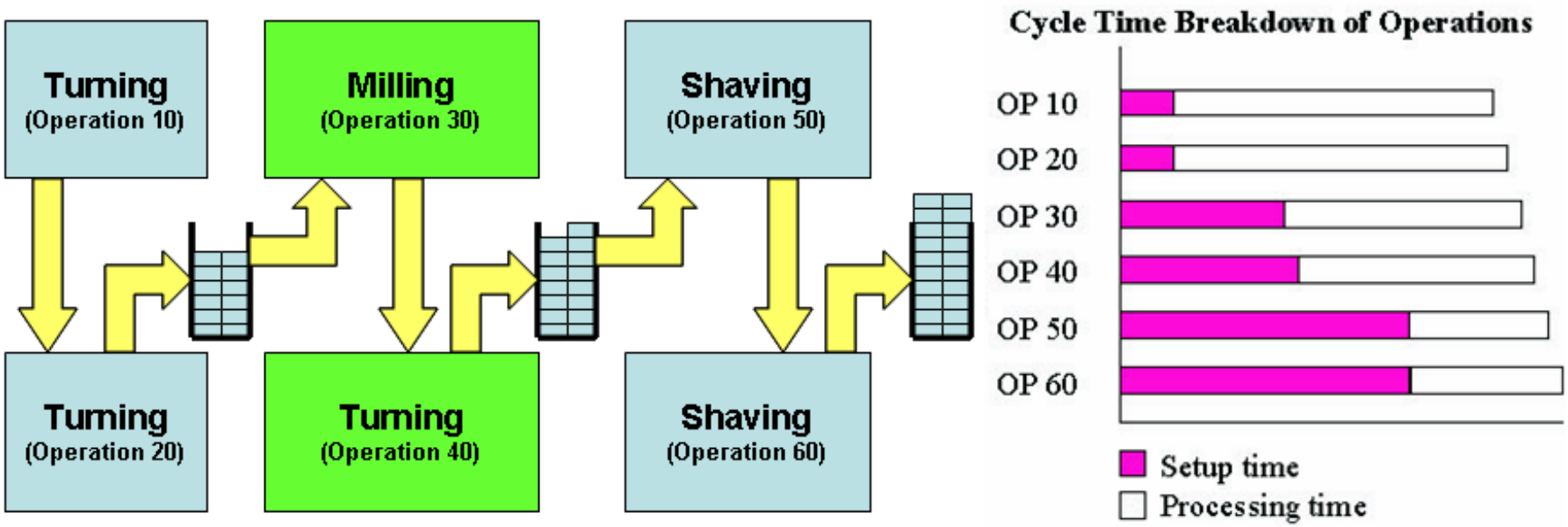
Level 3: Assembly Line or Transfer Line

Single piece flow may be occurring in some areas; however, between-process inventories are large and demand is still based on a forecast.



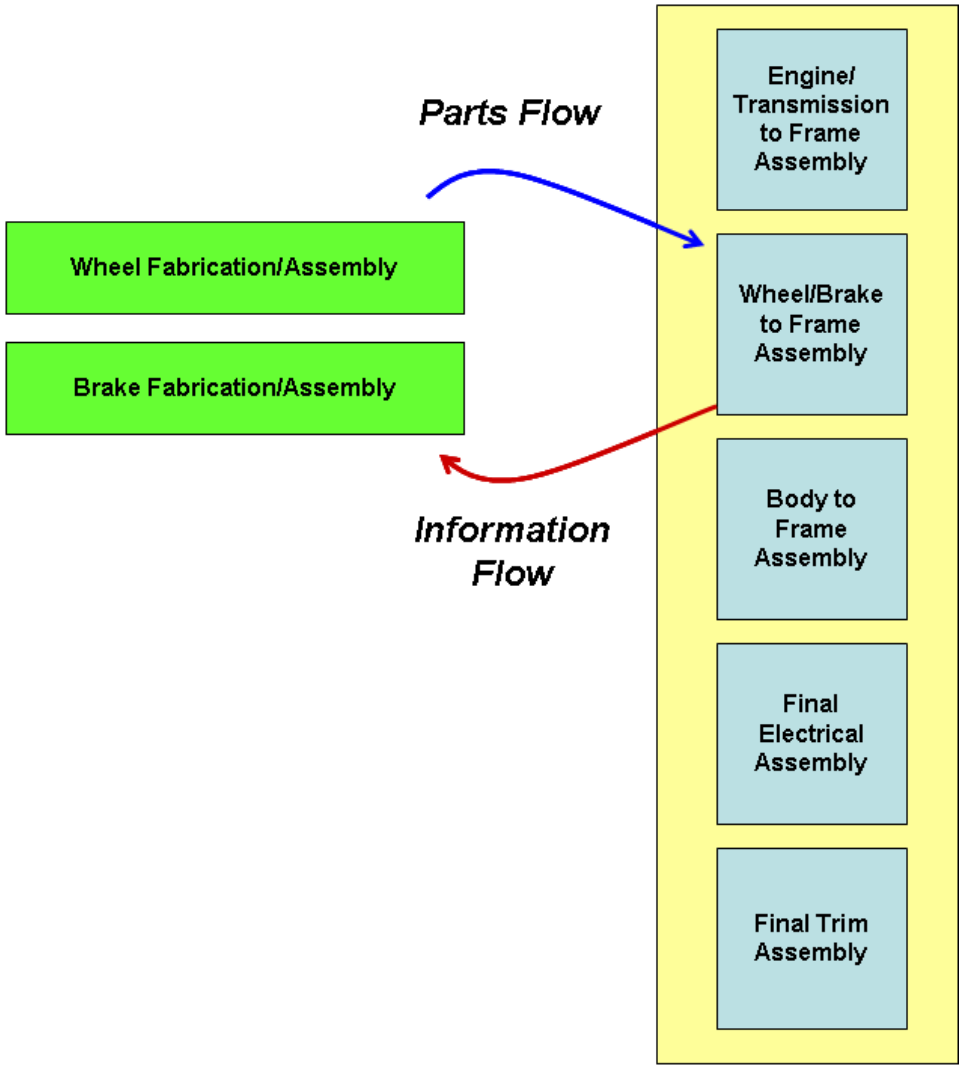
Level 4: Pseudo Cell

The operations are arranged in order of product flow, in-process inventory is controlled by output “buckets” – the process step produces enough to fill the bucket and then stops until the bucket is emptied.



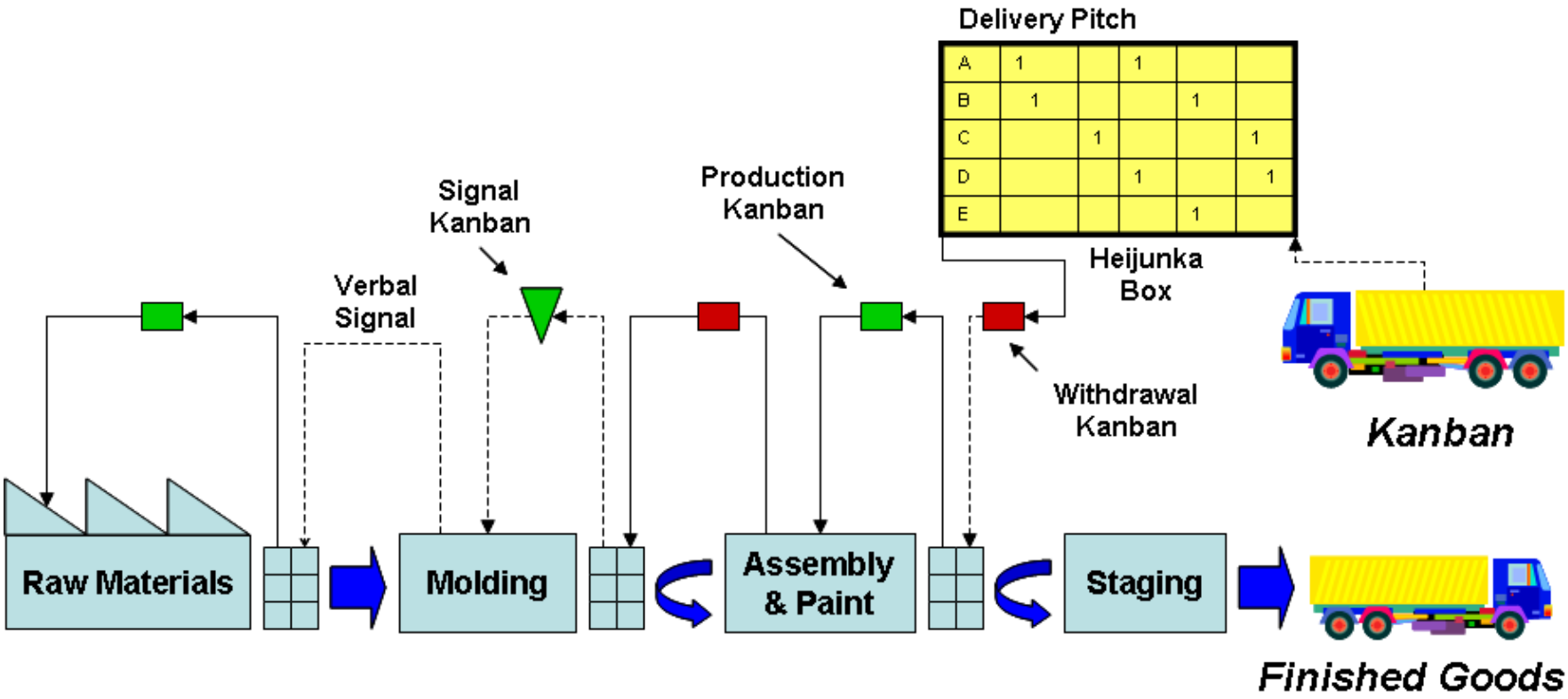
Level 5: Assembly or Machining Cells

Parts machining have been separated from the assembly flow; parts are produced in response to demands from the assembly cell.



Level 6: Linked-Cell Manufacturing System

The entire production process is organized by product flow; signals (kanban) are employed to *pull* the product through the plant based on market demand.



Summary of Lean "Rules"

Kanban Rules

- A kanban is always attached to the product
- No production or withdrawal without a kanban
- No production to keep workers busy
- Only standard size containers having the standard number of parts each
- Follow the kanban delivery sequence
- No defective parts sent forward
- Strive to reduce the number of kanban

Implementation Issues

- Human Costs of Just-in-Time
- Require an atmosphere of close cooperation and mutual trust between the work force and management
- Require daily production schedules that are virtually identical for extended periods
- Require daily production to closely approximate the daily schedule
- Cannot respond rapidly to changes in product design, product mix, or large demand volumes
- Require many production setups and frequent shipments of purchased items from suppliers
- Require parts to be produced and moved in the smallest containers possible
- Not well suited for irregularly used parts
- May require layout changes
- May require changes in reward systems
- Require revision of purchase agreements as the number of suppliers shrinks

Supplier Concerns

- Desire for diversification – supplier is concerned about all business stemming from a single customer
- Poor customer scheduling – supplier is concerned that customer will not be able to develop a consistent schedule
- Engineering changes – supplier is concerned that customer will promulgate frequent engineering changes with little lead time
- Quality Assurance – supplier may consider zero defects unrealistic
- Small lot sizes – supplier hasn't worked with small lot sizes
- Proximity – delivery of small lots over long distances is not economical

5.6 Exercises

Exercise - Process Inventory Sheet:

Complete a process inventory sheet for your project or a process with which you are familiar:

Process Inventory Sheet

Process:	
Outputs:	
Inputs:	
Materials:	
Machines:	
Information:	
Skills:	
Environment:	

Exercise – Classifying Customers & Suppliers

For the process identified in the previous exercise, identify the customers and their "value" to you. If your customers are primarily internal to your organization, try to estimate the "revenue" you receive or your "sales" to these customers.

Customer	Revenue	Volume	Criticality	# People

Criticality: Customer has High Impact = 3, Some Impact = 2, Low Impact = 1

Exercise – Customer and Supplier Requirements

This activity will help your team identify customer and supplier requirements for the process that your team is to improve.

Instructions:

1. Identify and classify the customers of your process. List the vital few customers and their requirements on the next page.
2. Identify and classify the suppliers of your process. List the vital few suppliers and your requirements on the next page.
3. Rate the requirements in terms of importance (1 – low, 5 – high).
4. Rate the requirements in terms of performance (1 – low, 5 – high).
5. Ask, yourselves, “Are these requirements valid?”
6. Plan future actions to confirm this initial assessment (e.g. interviews, surveys).
7. Be prepared to discuss your findings with others.

Process Name:				
Customer	**Their Requirements**	**Valid?**	**Rank**	**Performance**
Supplier	**Your Requirements**	**Valid?**	**Rank**	**Performance**

Exercise – Developing a Flowchart

Identify some products or services that your department produces. Develop an appropriate "picture" of the production process. This could be a flowchart, layout diagram, arrow diagram or "combined" chart.

Next, examine this "picture." Thinking about what is important to the customer (i.e. timeliness, accuracy, quality, cost, etc.), are there any ideas that you get about how to improve this production process?

Exercise – Flowcharting Process Descriptions

Flowchart one or more of the following processes. After flowcharting the process, try to analyze it to determine if there are ways the process might be improved. The most important quality characteristics are noted in parentheses. If "accuracy" is most important, try to determine where and what errors might be made (and what error-proofing measures might be taken). If timeliness is important, create a critical path diagram (estimate times to complete some of the steps) and look for ways to improve the critical path:

Travel Reservations (Accuracy, Time to Prepare) *Two weeks before traveling, a quality consultant checks the airline schedules on his on-line service. He copies down some likely flights and calls the airline reservations number. After making the airline reservations, he decides whether he'll need a hotel room and for how many nights. If a hotel is needed, he consults his list of favorite hotels for cities previously visited. If he has no "favorites," he calls the client and asks for some recommendations.*

Then he calls the hotel reservation number and makes his reservation. If they are booked, he tries the alternate hotel. Then, he calls the car rental agency to reserve a car. If they don't have anything available, he calls an alternate agency. After all the reservations are made, he prepares a travel folder for that trip, including typing and printing two copies of his itinerary. When he receives the airline tickets in the mail, these go in the folder.

Take-Out Pizza Ordering and Pickup (Accuracy, Timeliness - try this one as a responsibilities flowchart!) *A customer calls the pizza shop with a take-out order. The cashier takes the call and writes down the order on the order form. She then walks back to the kitchen area and tapes the order to the front of the shelf over the pizza "fixins" area.*

The pizza chef looks at the order and begins to make the pizza. The dough is first prepared, then sauce and cheese are applied. He glances at the order before putting the toppings on the pizza. Secret spices are the last to go on the pizza. If an oven is available, he'll pop the pizza in, if not, it waits until a previous order is removed. Occasionally, he'll check the pizza and turn it around in the oven. When the pizza is done, he removes it from the oven, checks it one last time and boxes it. He tapes the order to the pizza box and puts it on top of the oven to stay warm.

When the customer arrives, they tell the cashier their name and the cashier checks to see if the pizza is done. If it's not ready, she'll ask the chef how long it takes and then tell the customer. If it's ready, she will bring it up to the cash register. She'll open the box for the customer to inspect the pizza. If everything's OK, the customer pays for the pizza and leaves.

If not, she'll find out what and try to "disposition" the pizza (customer's willing to take it, but at a discount, customer not willing to take it, wants another, etc.).

Pam's Pack and Ship (Accuracy, Timeliness) *Pam takes orders for quality improvement videos. The customer will call to order. Pam takes the ordering information and inputs it into the computer. If the customer charges the order on a credit card, she takes the card number and calls to verify the card's validity. Then, she prints an invoice and shipping label.*

She takes these to the packing area and gathers the customer's videos from the inventory shelves. She packs these in a shipping box, inserts the invoice and tapes the box closed. Then, she tapes the shipping label to the box, weighs the box and puts it on the shipping shelf by the back door. Finally, she fills in the shipping log for the parcel service.

Operating Room Turnover (Cleanliness, Timeliness, Accuracy) *After the surgeon or the assistant closes the incision and the patient is wheeled to recovery, the turnover process begins. First, the crew disposes of liquids and sharps (syringes, suture needles, etc.). If there are any specimens to be analyzed, these are prepared. Everything else is placed in a red disposal bag and on the operating table. The table is then moved to "Central" and the crew returns to the room.*

The room is then washed down (tables, lights, kick panels) and the floor is mopped. A new suction canister is obtained for the next procedure. If one is not available, they call Materials Management to send one up "pronto." Linens are obtained and a new red disposal bag brought in. Anesthesia equipment is then obtained and readied. The bed is made up. The next surgical kit is obtained, opened and inspected. If there are items missing, these are obtained. The operating room tech then scrubs and performs the sterile setup of the surgical kit. The "counts" are done and meds are placed on the back table. Turnover concludes when the next patient is called.

Fast Valve-Open Power Plant Startup (timeliness, fuel costs, protection of turbine from overstress) *Power plants that are used in daily "start/stop" operation must be capable of quickly starting and shutting down to meet the needs of the utility's customers. At the same time, startup procedures must be mindful of equipment limits, such as occur when metal is heated too quickly and experiences thermal stresses. This procedure was developed to quickly start a turbine and minimize stresses.*

Initially, turbine metal temperature is at ambient temperature (90° - 100°F). The turbine control valves are opened. The operator applies steam to the turbine steam seals. Circulator and condensate pumps are started, and a vacuum is pulled in the turbine condenser. Steam formed in the boiler passes through the superheater, increasing temperature, but

decreasing pressure as it moves toward the condenser - it becomes superheated. Steam temperature is about 200°F when it reaches the control valves, with about 50-60°F of superheat (the condenser's saturation temperature at a vacuum of 22 inches of Mercury is about 150°F). The turbine experiences superheated steam before it rolls. The turbine is accelerated at about 25-50 rpm/minute, while the boiler is fired at a 2°F/min. rate. Turbine oil temperature must be monitored during this process. Spec limits for oil temperature are as follows:

Startup Condition	**Oil Temp.**
Turning Gear Operation	70°F
Turbine Roll	80°F
Turbine Speed @ 1000 rpm	90°F
Turbine Speed @ 3000 rpm	100°F

If oil temperature doesn't track with the startup condition, the turbine control valves must be throttled to control turbine speed and the "open valve" procedure aborted for a "traditional" startup. When turbine speed increases above 3000 rpm, a "hold" occurs to allow the Intermediate Pressure turbine blade temperature to increase above 250°F. This also allows the turbine rotor bore temperature to increase above 175°F.

Between 3000-3300 rpm, control is transferred from manual to turbine computer. The computer then controls turbine speed and pressure during the final ramp up to synchronization and generator breaker closing. The plant is now "on-line."

Locomotive Dynamic Brake Testing (Accuracy) *This is a portion of a test designed to determine if a locomotive's dynamic brakes are functional. The dynamic brakes operate to slow the locomotive by turning the drive motors into electric generators. The electricity generated is dissipated through grids mounted on top of the locomotive with fans blowing air across the grids.*

At the start of the procedure, the locomotive's diesel engine is running. The tech connects a motor generator set to two leads on a terminal block. A voltmeter is then connected across the motor generator leads to read voltage (0 - 500 VDC). Another voltmeter (0 - 150 VDC) is connected across terminals of the rate control panel.

The motor generator voltage is increased until a voltage is indicated on the rate control panel voltmeter (when contact 8-2 opens). MG set voltage should be between 305 - 310 volts. If not, then, reduce MG set voltage to zero and adjust

rheostat RH10. Repeat the test until the proper MG voltage conditions on contact 8-2 opening are achieved. Finally, reduce MG voltage to zero. Disconnect the rate control panel voltmeter. Proceed to the next test.

Diesel Engine Cylinder, Piston and Rod Inspection (accuracy) *This procedure inspects the condition of a locomotive diesel engine's piston and rod assembly. The inspection is performed periodically, or when trouble is suspected in the engine. Prior to the inspection, the engine is shut down and air box and oil pan inspection covers removed. All cylinder test valves are opened to allow rotation of the crankshaft.*

For the cylinder being inspected, the crankshaft is rotated (using the turning jack) until the piston is at bottom center. The technician inspects the top of the piston and the cylinder wall. If the piston crown is wet, the fuel injector may be leaking. Water leaks and cylinder scoring are noted. If these are present the entire power assembly may require replacement.

The crankshaft is then rotated toward "top dead center" until the piston compression rings are visible through the liner ports. The tech checks the following ring conditions: Side clearance of the Number 1 compression ring, condition of chrome grooves in ring (if little chrome is left, rings should be replaced), broken rings and ring blow-by (vertical brown streaks on the face of the ring - replace if condition is severe).

The piston skirt is inspected for scoring or buffing and the air box is inspected for foreign material and signs of water or oil leakage. The tech records the inspection results on the M-form and returns them to the scheduler's office.

Organ and Tissue Donation (timeliness, accuracy, family consent) *The process starts with a potential donor. If the donor's heart is still beating (only this path will be followed here), the Organ Procurement Team (OPT) is called to the hospital. The nurse and team then ask the family for their donation approval.*

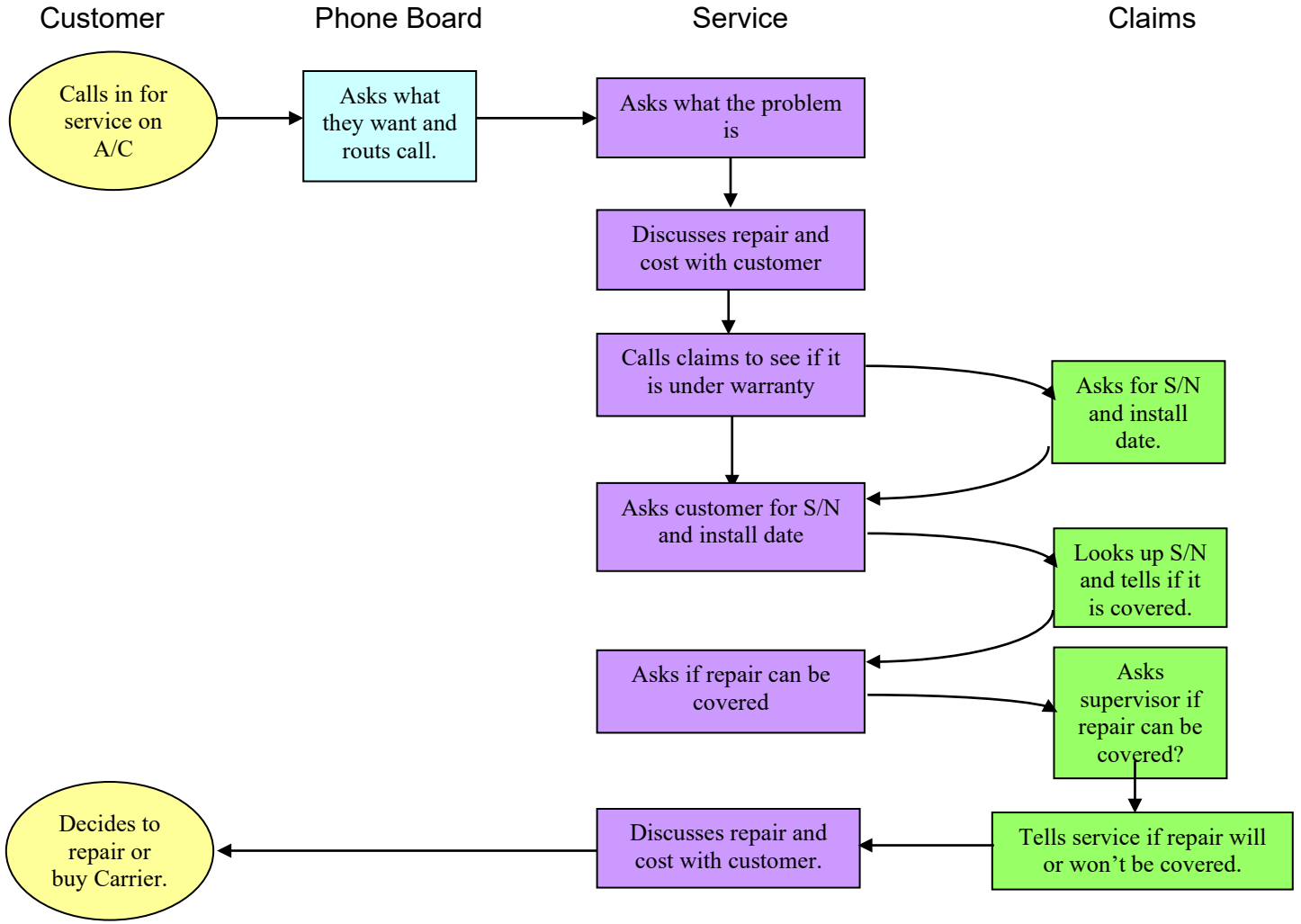
If the family agrees, a "Release of Body" and forms NG212 and NG213 are completed. The nurse works with the OPT to save the organs and the unit nurse calls the Operating Room circulating nurse.

The "Release of Body" form is pinned to the body by the unit nurse and the OPT and body go to the Operating Room. Organs are retrieved, packed and shipped. The body is taken to the morgue. The transporter fills in the "Release of Body Log" in the morgue. Additional retrieval of tissues may occur in the morgue. The body is then sent to the funeral home.

The morgue signs the "Release of Body Log" to release the body to the funeral home and the funeral home representative signs the "Release of Body Form." If the donation was not approved by the family, the nurse completes form NG212 and, after death, the body is transported to the morgue.

Exercise – Quick and Easy Improvement:

Examine the following flowchart for obvious improvements that would not cost much money or resources to implement. Do they also meet the criteria for easily reversed and their benefit quickly seen?



Project Assignment – Quick and Easy Improvements:

1. Review your flowchart for Quick & Easy improvements.
2. If you find opportunities, plan how you would make the process changes.
3. Be prepared to discuss and show your work in the morning.

Exercise – Critical Path Analysis:

Perform critical path analysis for the cycle time of your process.

Introduction to the Card Drop Shop (Table Exercise)

Objective: To establish performance of an “unstandardized” process.

Instructions: 1. You will be provided with the following production equipment and materials:

- Sheet of paper with “X”
- Deck of Cards
- Measuring Tape

2. Setup the production process as follows:

- Pick an operator, a materials handler an inspector and a recorder.
- Place the “X” paper on the floor
- Position the operator so that with their arm outstretched, their hand is over the “X.”
- Provide the materials handler with the deck of cards.
- Provide the inspector with the measuring tape.
- Position the recorder at a flipchart with a marker.

3. Produce 25 “units” of product as follows:

- Materials handler provides one card at a time to the operator.
- Operator drops the card over the target
- Inspector measures the distance from the card to the “X”
- Recorder writes the distance on the flipchart.

4. Quality Analysis – Develop a histogram of the process output, calculate the mean distance and also the standard deviation.

Time: 20 minutes

Standardizing a Current Process (Table Exercise)

Objective: To create a standard process and document same via a process management chart.

Instructions: 1. Review Units 2.6, 2.8 and the Card Drop Shop production process.

2. Standardize the current process and document this on a process management chart (Note: the instructor is the "customer" of the process. They are available to answer any questions you have about their needs).

3. Run the standardized production process (25 units). Did standardizing the process result in any improvement?

Time: 30 minutes

Standardizing Your Process

Objective: To create a standard process and document same via a process management chart.

Instructions: 1. Pick a process that you own or that is the subject of your improvement project

2. Standardize the current process and document this on a process management chart.

3. Run the standardized production process. Did standardizing the process result in any improvement?

Objective: To understand the concept of *muda* in a production process.

Instructions:
1. Pick a production process (or small “piece”).
2. Develop a detailed flowchart for the process. Include all current steps that are taken to perform the process.
3. Classify the steps as value or non-value added. Make sure that only *operations – those steps that transform the material, product or service* are classified as value added.
4. Identify at least five ideas that could be used to eliminate the *muda* from that process.
5. (Extra credit!) Implement these ideas.

Time: 60 minutes

Objective: Understanding the current production process from a lean perspective.

Instructions:

1. Identify one of your current production process.
2. Develop an understanding of how the product or service is produced (high-level flow chart is a possible deliverable).
3. Develop an understanding of how the product or service is *moved* through the process – is it currently a *push* or *pull* system.
4. How are current production equipment being utilized – all the time, in response to demand, other?
5. Look for areas in the current process where inventory accumulates. Where are these? What current methods are employed to manage inventory?
6. What improvements does this review suggest?

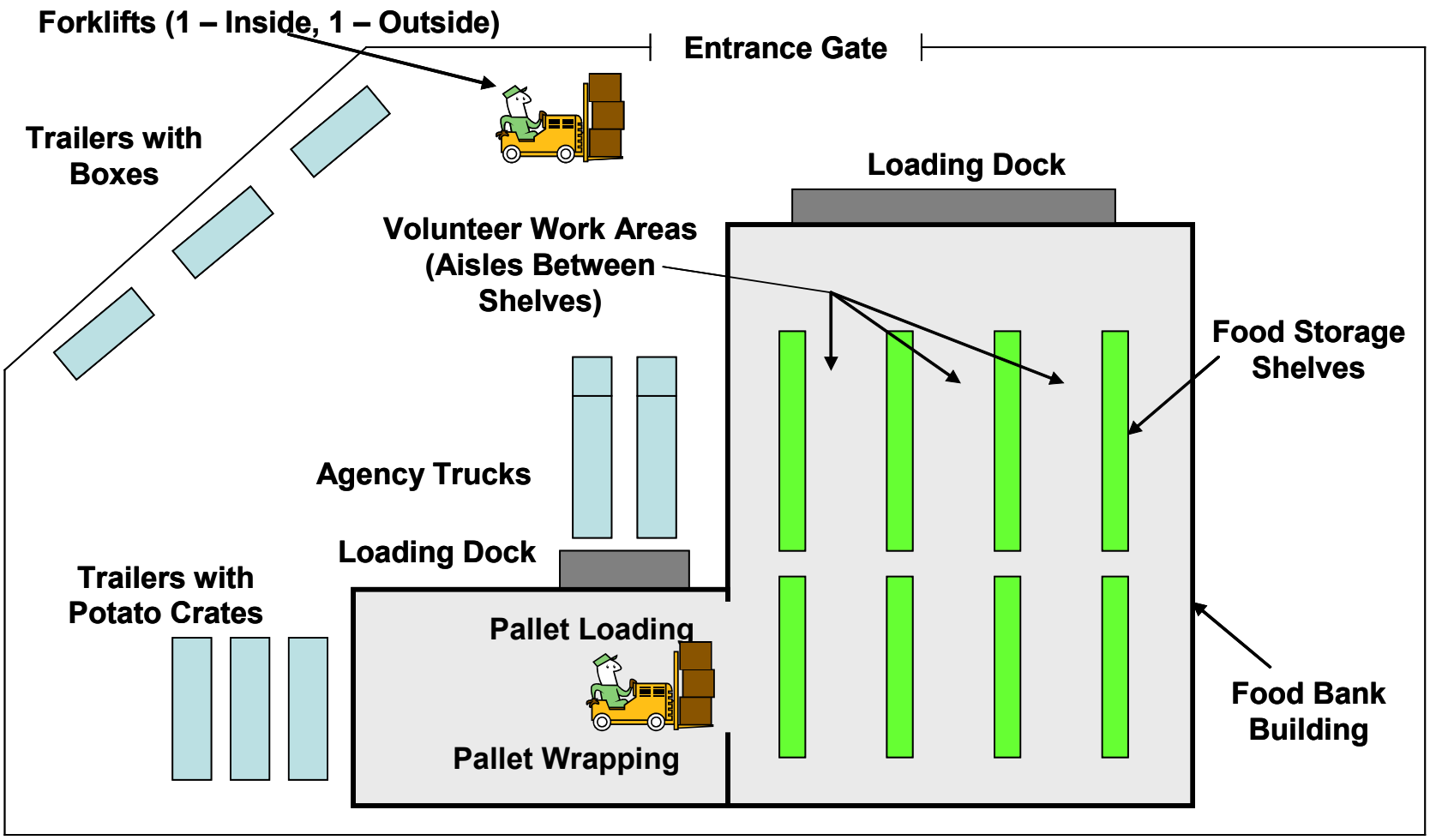
Objective: "Lean out" a production process.

Scenario: Each Saturday, volunteers arrive at a Food Bank to help sort groceries into bundles to be delivered to specific relief agencies. One of the key "production processes" involves packing sweet potatoes into boxes. The process starts with a forklift delivering a large (6' x 6' x 4') crate of potatoes from a truck trailer to a group of about 5-6 volunteers (one Saturday, there were 50 volunteers at the Center). Another forklift delivers a pallet of boxes to the volunteers (also from a truck trailer). The volunteers then take the potatoes from the large crate and fill the boxes. If a potato is "bad" (rotten, full of holes, etc.) it is thrown into a blue plastic crate. When this crate is filled, it is dumped into the "defect" crate. When a box of potatoes is filled, it is placed on a pallet. The pallet is stacked in a 3 x 2 arrangement – 3 boxes with their long sides adjacent on one side of the pallet, 2 boxes with their short sides adjacent on the other side of the pallet. When 4 layers of boxes are filled, the forklift takes the pallet away to a wrapping location where it is wrapped in plastic for shipping to a distribution agency. The agency's truck then transports to pallet of potatoes for distribution to the needy.

Instructions:

1. Review the sweet potato production process (above) and the physical layout of the Food Bank (next page).
2. Using the principles described in Unit 5.5, develop a lean production process for the potatoes.

Food Bank Physical Layout:



6.0 Measuring Performance and Variability

Unit	*Description*	*Page*
6.1	Developing Performance Indicators	6.1 - 1
6.2	Data Collection	6.2 – 1
6.3	Core Data Displays	6.3 – 1
6.4	Introduction to Control Charts	6.4 – 1
6.5	Measurement Control Charts	6.5 – 1
6.6	Attribute Control Charts	6.6 – 1
6.7	Measurement System Analysis	6.7 – 1
6.8	Process Capability Analysis	6.8 – 1
6.9	Additional Control Chart Topics	6.9 – 1
6.10	Exercises	6.10 - 1

The focus of this Section is on methods used to measure the performance of a product or process. We start with the issue of identifying performance indicators that correspond to meeting customer needs, will then discuss planning data collection activities and finally present a number of tools to display and analyze the indicator. One of the major themes in this section is that of *variability.* We live in a world of variability – all our processes exhibit variability. However, variability in product or service is generally *undesirable*, for the most part, we will treat variability as an ***enemy!***

6.1 Developing Performance Indicators

Learning Objectives

- Understand Why Measurement and Improvement are Linked
- Identify Indicator(s) for Improvement Projects
- Operationally Define Measurements

Unit Contents

- Why Should We Measure?
- Process Measurement “Philosophy”
- What to Measure – The Customer Perspective
- Process Management
- Critical to Quality Characteristics Defined
- Control Points and Checkpoints
- Selecting Project Indicators – General Process

6.1.1 Why Should We Measure?

To many people, data really is a "four-letter" word. Some people are uncomfortable dealing with numbers, especially when those numbers are a reflection of their or their department's performance.

We've all been "beat up" with numbers at some point in our lives, all the way from grades in school to last month's financial report. However, the collection and use of data is essential to process management. To paraphrase one of our heroes, Lord Kelvin,

If you don't measure it, you don't understand it.

In this unit, we'll explore some of the "philosophy" of process measurement and develop several *very* useful tools that will help us understand the performance of our processes. Although these tools (run and control charts) were initially developed for manufacturing, they have been extended to virtually every kind of process management setting with great effectiveness.

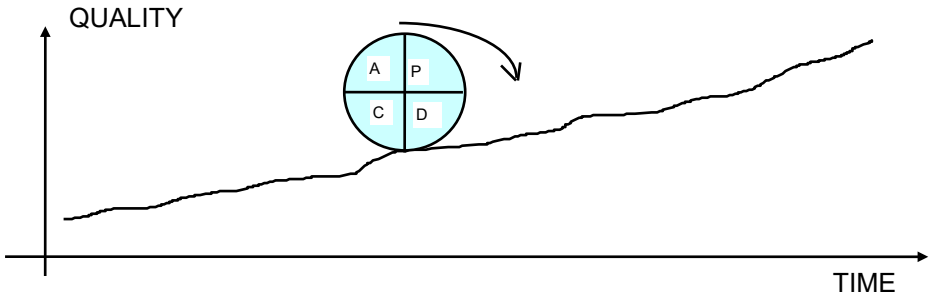
One key aspect of measurement should be introduced here. The customer of our production process cares, of course, about the individual product or service they receive regardless of how well we are doing "on-average." By caring about the *statistical behavior* of our process' output, we will manage the individual products and services received by the customer. Another of our heroes, Dr. Teiichi Ando, introduces his statistical training seminars with a gentle reminder that "we must move from the world of averages to the world of *dispersion.*"

Any efforts made to control the quality of product or service may be named *process control.* When we add measurement and understanding of variability to our efforts, we practice *Statistical Process Control.*

6.1.2 Process Measurement "Philosophy"

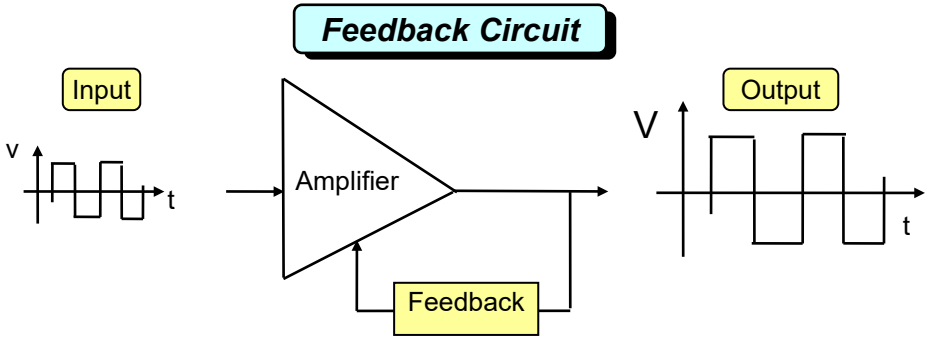
PDCA And Measurement

The PDCA cycle is essentially a feedback loop. If we keep at it, this loop doesn't simply keep the process aimed at one level. Over time, the continual application of PDCA will "nudge" the process towards increasing levels of quality.



Now "feedback" is a common term in today's language, but it's not that old. "Feedback" has its origin in radio design. Engineers found that they could control the output of a circuit by taking a small "piece" of the output signal and feeding it back to the circuit.

To apply this concept to our processes, we, too, will have to take a "small piece" of the process' output. We will use the feedback in the "Check-Act" phases of PDCA to make decisions about the need to adjust our process, to understand when certain "special" factors are influencing our process and to determine if our process is "capable" of making products and services at the desired level and variation. This is a key aspect of measurement – its use to take action on the process. This text will not advocate simply taking data to judge the state of the process without an appropriate action being taken.

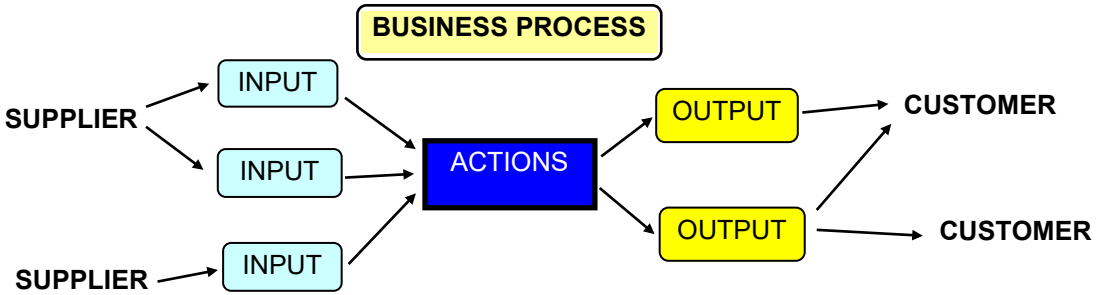


6.1.3 What to Measure – The Customer Perspective

Many college "stats" books start off with a discussion of measures of central tendency, dispersion, skewness, etc., but they never seem to talk about ***what*** we need to measure. We see this problem in the real world when people attend a quality course and then go back to their departments and say, *"Now what?"*

Because of this *observed* behavior, we'll start with the "basics." There's no single answer to the question "*What to measure?"* but we should be able to provide enough guidance to get you going.

Recall our picture of a business process:



There are three major components to our process - **outputs, actions** and **inputs.** These are all candidates for *measurement.* However, before we jump into identifying a measure, we should make sure we understand what the *customer* wants from this product or service. Then, we can identify which aspect of the product or service is most in need of improvement.

Output Type Measures

Let's start with the **outputs** (i.e. products or services). Here, we will identify the customers, their needs and expectations associated with our product or service (i.e. to listen to the **Voice of the Customer,** see Section 3). A natural *measurement* question is

"*How well are we meeting the customer' needs and expectations?"*

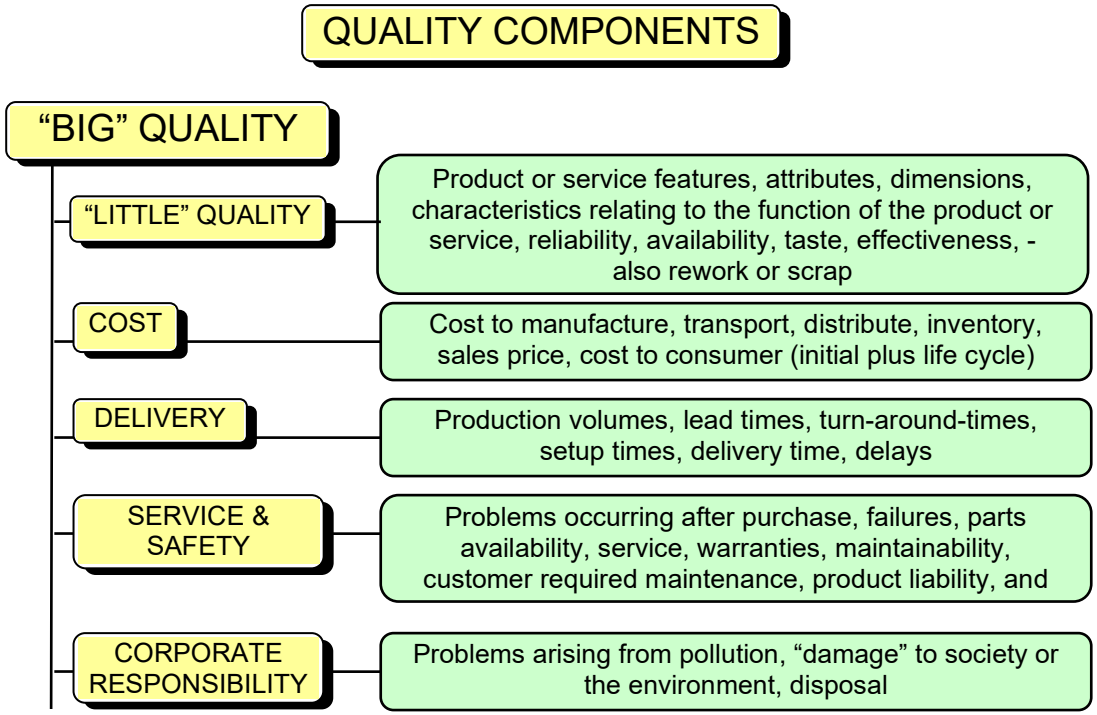
A simple, "popularity-type" of output feedback is then *sales*. This might be stated in terms of either *volume* or *dollars.* Think of *sales* in a broad sense. If you are a department whose customers are all *internal* to your organization, you still have "sales." For example, one quality-resources group at a large healthcare system developed a poor reputation for their "services." They were too theoretical in their consulting practice, often quoting Deming, but not providing any practical advice. Requests for help from the individual hospitals that made up the system just dried up - in effect, this department's "sales" were ZERO!

Sales volume is not the only customer feedback we might obtain. Customer surveys regarding their satisfaction with our products and services may be taken, however, in many cases, simply talking to our customers can reveal valuable feedback.[1] Observing how the customer uses our product or service can be an extremely valuable *measurement.*

Once we understand their needs, these can be "translated" into the **quality characteristics** of our product or service. These quality characteristics are the *causes* that result in the *effect* of customer satisfaction (or dissatisfaction) with our product or service. Although it's sometimes a challenge, quality characteristics can *always* be measured (the ghost of Lord Kelvin backs us up here!). Unit 14.1 describes a rigorous method, Quality Function Deployment (or QFD) for performing this "translation" function.

When we say ***Quality* characteristics,** we're using QUALITY in a "Broad" sense. We can break this "Big Q" down into five components:

[1]It can be hard to sit across the table from your customer and get frank feedback - especially if they are another department that has been considered the "enemy" for some time!



Often, we can focus on and measure several "key" or critical quality characteristics to get an overall picture of the process' performance (here, we will call these *Critical to Quality Characteristics,* or *CTQs*).

High-Level Examples:

- A supplier's "Big Quality" can be summarized as "delivering the *right parts* to the *right customer* at the *right time*, for the *least cost* with no *defective parts.*" Embedded in this statement are five quality characteristics that can be quantified.

- A maintenance department's "Big-Q" is to maintain *reliable* equipment in a *short* time at a *minimum* cost. Three key quality characteristics that can be measured are embedded here.

- An air conditioning manufacturer's "Big Q" is to provide *reliable*, low *cost*, air conditioning and refrigeration equipment *on time*. Three key quality characteristics are found here.

Notice how these "high-level" quality characteristics are obtained from statements that describe the *mission* or *function* of the production process.

Low-Level Examples

- A fuse must *reliably* open at a certain *over-current* or *over-voltage* condition in an electric circuit. It must fit into a socket of specified *diameter* and *length.* The fuse must *not open* below the specified voltage or current.

- A home air conditioner must remove a certain amount of *heat/hour* from a house (BTU/hour) while operating at a certain *noise level* (decibels) and consuming less than a certain *power* (Watts).

- A hard drive provides *reliable*, *error-free* digital data storage. It must fit into the PC's frame (*length, width* and *height* dimensions).

Again, although these are "low-level" products, knowledge of the required functions provides us a path to understanding the critical to quality characteristics of the product.

One last note: make sure you define the characteristic with the *customer* in mind. Often companies fail to consider the "***wing-to-wing***[2]" nature of the customer's need. One financial services company measured the time it took from when a lease was approved to when it was shipped to the customer. Since the leases were approved at various times during the day and all leases were sent out at 4 pm each day, there was a lot of "variability" in the process. They spent a great deal of energy trying to improve the process; however, when asked, the customer was completely satisfied with the company's performance.

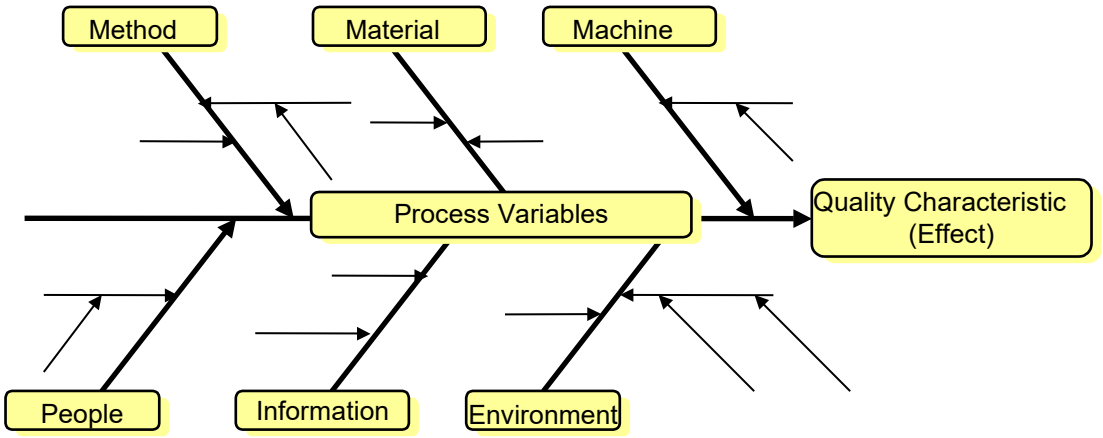
From another perspective, wing-to-wing means to "reach" into your customers' and suppliers' processes to help them improve. One GE Capital business worked with their customers to significantly reduce the time it took for the customer to be paid for their products. This *improvement partnership* led to GE Capital receiving a long-term contract worth millions of dollars.

[2] The "Wing-to-wing" term comes from an aircraft jet maintenance example. One company focused on improving the time it took to repair an engine *in their shop.* The customer (the airline), though, was concerned with how long it took from when the engine was removed from the wing to when it was returned to the wing of the craft.

Action Type Measures

Products or services are the **effects** (outputs) of our production processes. The process then can be considered a system of **causes** that produce the desired effect. The Cause & Effect (or Ishikawa) shows this relationship:

Cause & Effect (Ishikawa) Diagram



All processes are composed of literally hundreds of causes. In most cases, though, there are only a few that have a major impact on the desired effect.[3] If we can identify these key causes, we will achieve two goals: 1) we are well on our way to effective process management and 2) we will probably have identified some key action-type measurement points.

Manufacturing-type processes - In-process measurements are quite common in fabrication and assembly processes. Important dimensions, weights, and physical properties are all candidates for measurement.

Service-type processes - Here, in-process quantities, timeliness and error rates are candidates for measurements. For example, if an accurate sales order is not entered, a customer shipment may be incorrect. If a service technician is not dispatched within a given time, restoration of service is delayed.

[3]This is known as the PARETO Principle or 80/20 Rule.

Input-Type Measures

If the "success" of your product or service also depends on SUPPLIERS, then you will consider **input-type measurements.** This is a lot more fun than identifying the output-type measurements, because now YOU are the customer! In an ideal world, your suppliers will come to you to ask you about your needs and expectations, and they will identify appropriate measures of their performance.

In many situations, as the purchaser of products and services, you will identify targets and specifications for your suppliers to meet. These can generally be measured. Just like the action-type measurements, the most important characteristics of "purchased" products and services should be identified as candidates for measurement.

Prioritizing Measures

Early in an improvement project (e.g. in the *Measure* Step of DMAIEC), you may struggle to identify an *appropriate* (i.e. not too many, not too few) set of measurements. A qualitative cause and effect analysis can help funnel your list of potential indicators. This funneling analysis makes use of your team's process knowledge to explore the relationships between potential action and input-type measures and the output measures you are attempting to improve.

Funneling Steps

1. Identify the output measures (using Voice of the Customer, existing process specifications, etc.). If there are more than a few output measures, prioritize the measures (1 – low importance, 5 – high importance).

2. Brainstorm potential action and input measures.

3. Identify (qualitatively) the relationship between the action/input measures and the output measures using a correlation matrix (see example below). If there is a **strong** relationship between the action/input measure and the output measure, use a bullet (◉), and score the relationship a 9. For **medium** relationships use a circle (◯) and a score of 3; for weak relationships, use a triangle (△) and a score of 1. If there is no relationship, leave the matrix cell blank.

4. Prioritize the action/input variables. If the output measures have not been prioritized, simply add up the relationship scores for each action/input variable (sum the column). If the output measures have been prioritized, multiply the

relationship by the output measure's priority and sum up these products. The example below illustrates the latter situation.

Prioritizing Action/Input Measures – Invoicing Example

Measure Funneling Matrix

Output Measures	***Rating***	***Action/Input Measures***													
		Computer Downtime	Invoice Accuracy	Time – Check Generation to Check Receipt	Time – Invoice Approval to Check Generation	Time – Invoice Generation to Invoice Receipt	Time – Invoice Receipt to Invoice Approval	Time – Shipment to Invoice Generation	Time of Month Invoice Sent	Time to Correct Invoice Errors	Contract Setup at Customer	Bill of Material Accuracy	Contract Information Accuracy	Customer Information Availability	Receipt Availability
Time to Receive Payment	5	△	◉	○	◉	○	◉	○	◉	◉	○	○	○	○	○
Correct Payment Received	5		◉								○	◉		○	◉
$ Value Non-Receivables	3		○									◉			◉
# Invoice Disputes	2		◉								○	◉	◉	○	◉
Importance		5	**117**	15	**45**	15	**45**	15	**45**	**45**	36	**105**	33	36	**105**

With help from the prioritization matrix, the team can decide to collect data on the most important action/input measures.

6.1.4 Critical to Quality Characteristics Defined

A Critical to Quality Characteristic (CTQ) is a measurable characteristic of a product or service with associated targets and specification or tolerance limits that correlates to meeting an important customer need. Often, a *sigma* (or allowable defect rate) target will also be set. For example, Six Sigma corresponds to less than 3.4 defects per million opportunities.

CTQ Characteristic Table:

	CTQ Characteristic				
Customer Need	**Characteristic**	**Measure (operational definition)**	**Target**	**Specification**	**Allowable Defect Rate**
Correct order taken	Accuracy	Number of incorrect orders/ numbers of orders x10^6 (DPMO)	N/A	N/A	3.4 DPMO
Order confirmed promptly	Timeliness	Process time in hours- Start; Order information completed by sales Stop; Client receives delivery confirmation	24 Hours	Lower spec: N/A Upper spec: 48 hours	3.4 DPMO

Operational Definitions

Consider the following statements:

- Service should be performed *on time.*
- Weld gas should be without contaminants.
- These air conditioners are *highly reliable.*

These are typical quality characteristics of products or services. But what do they mean? What is "on-time," how do you know if something is "contaminant-free," or "reliable?" These terms have meaning only if we can agree on their **operational definitions**. In essence, we must sit down and define what "on-time" really means.[4] These definitions don't have to be fancy; as Shewhart put it, an operational definition is one "that reasonable men can agree on." Deming suggests that there are three components to an operational definition:

- *A Sampling Method*
- *A Test*
- *A Judgment Criterion*

[4]Trying to define *when* surgery starts can be an interesting exercise. Is it when the patient enters the operating room, anesthesia is administered, the surgeon enters the room or the "first cut?"

In essence, a measurement is an output of a *process*. We must be careful how we manage the quality of the measurement process. For example, how could we agree that an event started "on time?" Every evening, the news "predicts" the next day's sunrise, say at 6:32 am. How could we determine if the sun rose "on time?"

Sampling - In this case, we are interested in a specific event, tomorrow's sunrise. For a production process, we may take a periodic sample of items from the process or inspect a random sample of items from a lot of finished goods.

Test - What test shall we apply to the event? How shall we define a "sunrise?" Is it when the sun's orb first crosses the horizon, or should the bottom of its orb cross the horizon? Which horizon shall we use? That viewed from our back yard? Or that viewed from the deck of a ship? Does it matter if you are 5 feet off the ground or 50 feet? Perhaps the simplest answer would be to call the TV station and ask them how they define a sunrise.

Criterion: In this case, the criterion is a simple decision: Did the sun rise "on time" (as defined by our test? Yes or No?)

One notion that we will have to abandon is that of a "true value." When an event occurs, we observe the event through some instrument (even if it is our own eyes). The results of any measurement depend on the instrument and the observer. We may agree that this instrument and that observer are preferred over that instrument and this observer, but that's the best we can do. Deming and Shewhart both cite the example of the speed of light. There is no "true value" for this physical quantity. The results obtained depend on the process used to measure the speed.


Event


Measuring Device


Observer

As a test of this true value business, try to establish what is the "true value" of your weight or height?

6.1.5 Process Management

Focus on the Outputs

How should we attempt to control the process? We'll combine the notion of measurement and feedback to answer this question. Let's start with a simple process where quality is controlled by measuring only the **output** of the process.

You are driving down the road in your car. You do not want to get a speeding ticket, so you select "speed" as the key quality characteristic. How is the process' quality *controlled* in this case? Here, measurement is simple. Every so often, you glance at the speedometer. If your speed is above the limit, you lighten your foot pressure on the gas pedal (the key process variable). You "check-act" the results of this change by checking the process' output, *speed*. If it is still above the limit (and you don't see flashing lights in your rear-view mirror), you further lighten the foot pressure. If your speed is too far below the limit, you may increase your foot pressure.

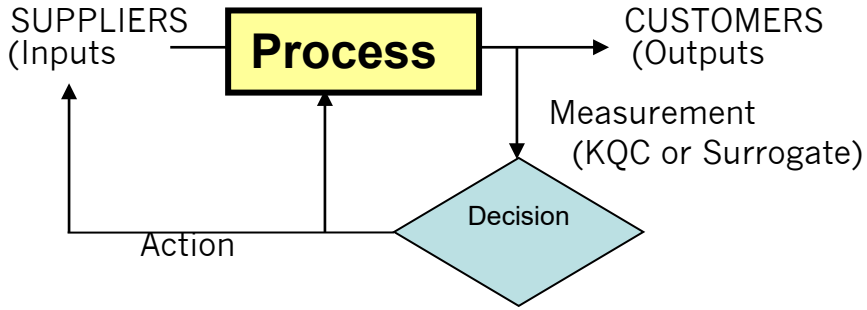
One concept that's important to grasp is that of "Checking *through* the Output." In our example, when the output was not where we wanted it, we went back *into the process* to adjust the **causes** that affect the output.

Many *inspection-oriented* organizations do not employ this concept. They check the output and accept or reject the output by comparing it to specifications or standards. If the output "fails," effort is not expended to go back to the process and uncover why it failed. *This is key to effective use of measures!!* The closer the measurement is to the process which produced the defect (given that a feedback loop occurs, the more effective the process management system will be in controlling quality.

What if the key quality characteristic is difficult to measure directly? Sometimes, a "surrogate" indicator can be identified, one that has a strong relationship to the key quality characteristic. Some time ago, we were riding in a car whose speedometer was out of order. The car was equipped with a tachometer, though, and the driver had established a relationship between the rpm indicated on the tach and the car's speed. The *control* of quality proceeded in the same manner; the tach simply replaced the speedometer as the measuring instrument.

Is there some way to improve this process if you find yourself consistently too far above or below the speed limit? What methods, machines, people, material, information or environmental process elements could be changed to improve the process?

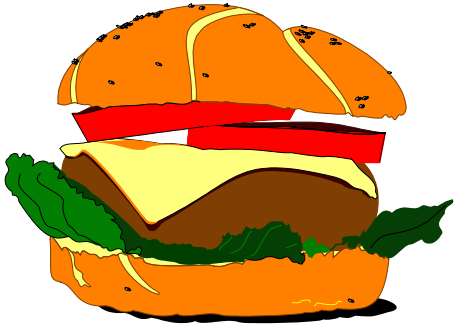
Graphically, the process we have been describing is shown below:



NOTE: "decision" means to study variation in the process and act "appropriately"

Focus on Both Outputs and Inputs

For many processes, the control of quality is more effective when measurement addresses both the output *and* inputs to the process (or process variables). Let's switch from cars to the kitchen to illustrate this point. One of our favorite Saturday dinners is "Cheeseburgers in Paradise."[5] What are the quality characteristics here? Well, they include *enough* burgers (quantity), *taste, temperature,* etc.



How do we manage quality for this "product?" We will still measure the results and "Check through the Output" (Boy, these are great burgers!). Let's also look at the "PLAN" for this product.

The type of ingredients, their amount and the sequence of "assembling" the ingredients have been "standardized." The "workers" have been trained to perform the process, and a recipe is used every time as the process' procedure (a flowchart would work equally well!).

What additional role does measurement play here? Each time the process occurs, the ingredients are measured, this is part of the PLAN. But every time we make the burgers, they do not come out the same. To manage the repeated

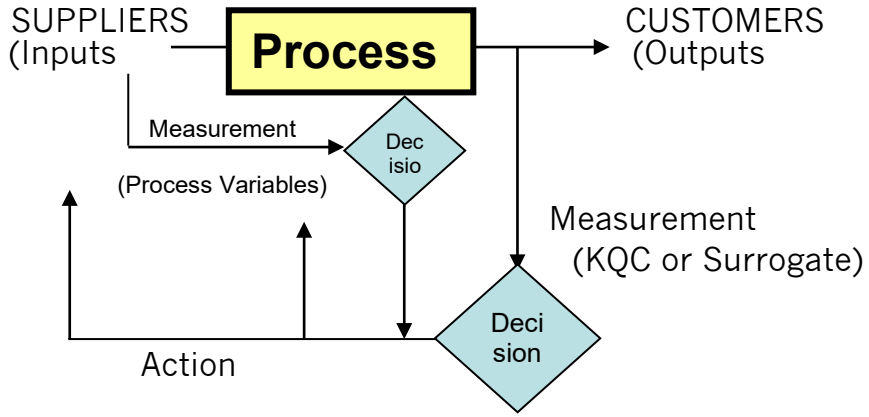
[5]Our apologies to Jimmy Buffett!

performance of the process, *measurement* would focus on monitoring the key process outputs *and* inputs (variables) over time. This variation would be studied and the PDCA wheel rotated to reduce variation.

For instance, the burger “customers” have indicated that they like their buns "lightly toasted." The time the buns are on the grill has been found to be the key process variable affecting this quality characteristic. Experiments have helped identify the optimal time to toast the buns and the process recipe includes this time. Here, we would measure the variation in both the “toastedness” of the buns and the time on the grill each time the burgers were made.

Of course, if there are any suppliers to this process, we may identify important quality characteristics of their products and services and apply measurements to these.

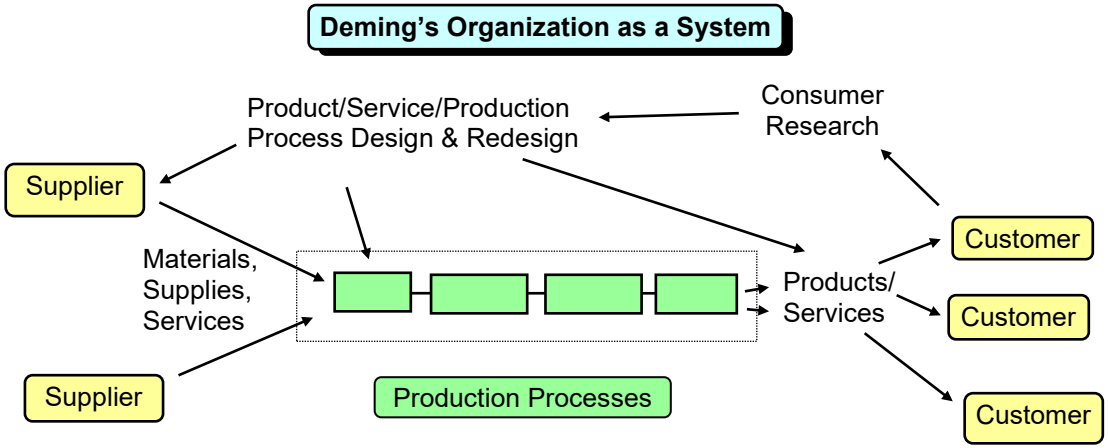
Graphically, the measurement process described here looks like this:



NOTE: "decision" means to study variation in the process and act “appropriately”

6.1.6 Control Points and Checkpoints

We've described how measurement might be applied to the outputs, actions and inputs that form the cause-and-effect system for a product or service. There's another way we could apply these principles, one that focuses on "combining" the Organization Chart *and* the Organization as a System.



The organization itself is a *cause-and-effect system.* The President/CEO will have certain goals or *effects* that he/she will wish to occur. If we *measure* these (i.e. set up indicators) for these desired effects, then they are the President/CEO's **control points.**

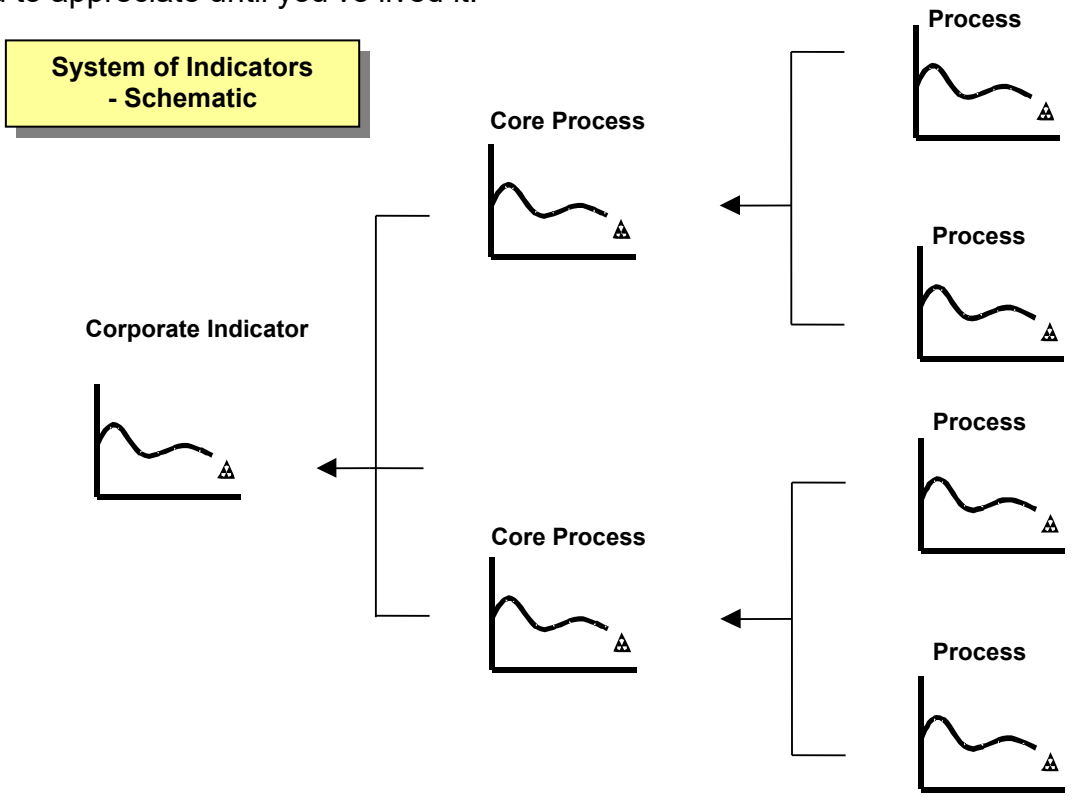
The President/CEO cannot directly affect these, but relies on the VP's, Directors, Managers, Supervisors and workers to accomplish the desired effects. Let's say that, to accomplish the President's goals, he/she asks the VPs to do certain things (i.e. assigns responsibility for Marketing, Planning, R&D, Design, Manufacturing, Sales, etc. For a *process-focused organization,* instead of the key functions, *process-owners* are identified). When they develop methods (causes) and *measures* for these responsibilities, they have identified *their* control points.

But the VP's control points are the President/CEO's **checkpoints.** If one or more of the President's control points is not performing well, then the President can look to see which of his/her checkpoints is responsible.

This measurement system can be extended throughout the organization. The VPs control points are deployed to the Directors and Managers; theirs are deployed to the Supervisors and Workers. The philosophy of controlling *through* the process still applies. The measurement system should be based on the "BIG-Q" concept - too often only the financial indicators are "deployed" throughout the organization. A System of Indicators can be developed, linking all corporate, business unit and departmental indicators together (see next page).

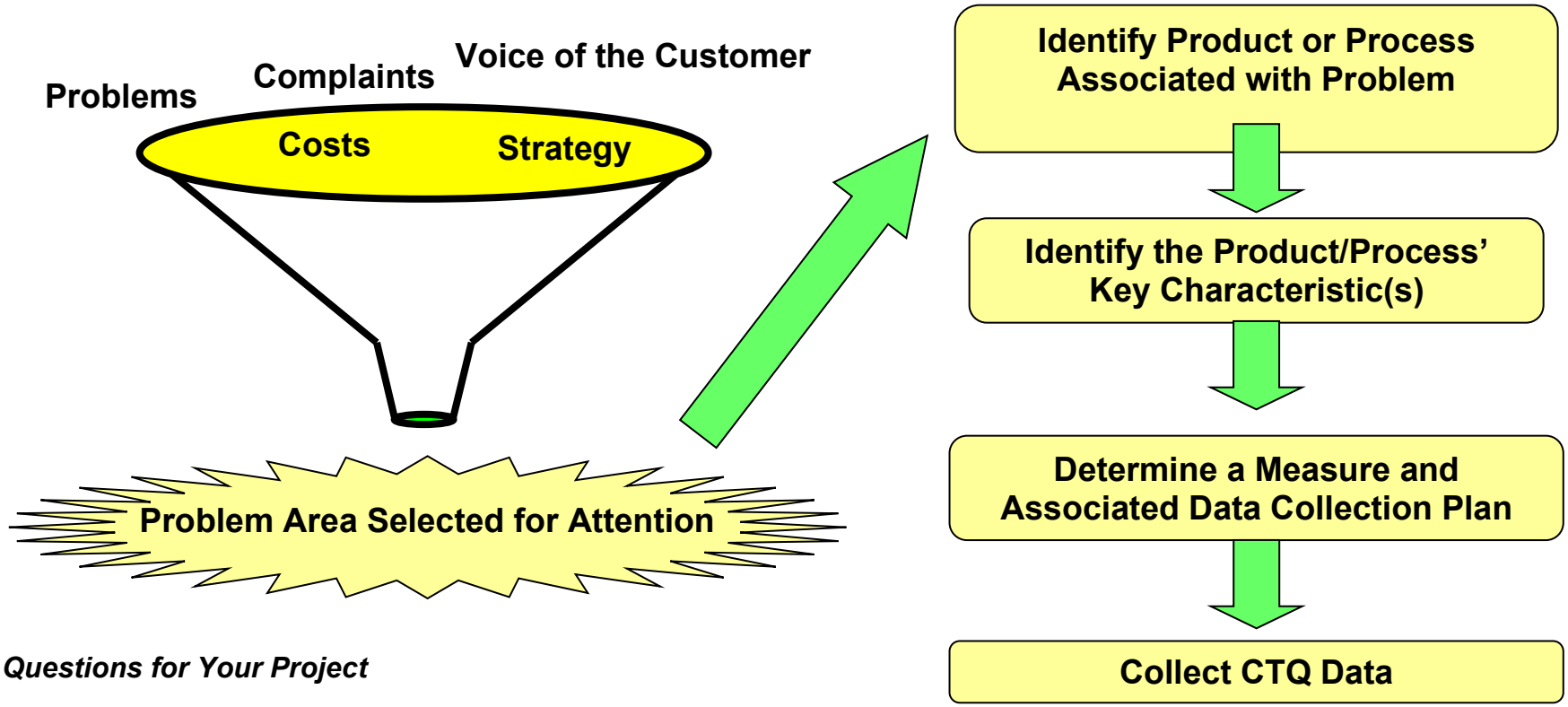
By 1989, Florida Power & Light had developed this measurement system, with the best example occurring in the Power Resources (fossil-fueled power plants) Department. A maintenance worker knew the impact of "turning a wrench" the right way (checkpoint) on preventing leaks from steam valves (control point for the worker).

The maintenance supervisor knew that preventing steam leaks (her checkpoints) would prevent maintenance-related plant shutdowns (her control point). In turn, the plant manager's control points included forced plant shutdowns, which was a checkpoint for the VP of Power Resources (responsible for all FPL fossil plants' performance). The power of this system is hard to appreciate until you've lived it!



6.1.7 Selecting Project Indicators – General Process

For your project, you will select one or more project indicators that reflect the problem being addressed. Here is a simple schematic to help you in this process:



Questions for Your Project

- Why are you working on this project?
- Are you trying to reduce costs, improve quality, reduce defects, or improve safety?
- Who are the customers of the product or service you're trying to improve?
- What needs improvement from their perspective?
- How will you know if the project is a success?

6.2 Data Collection

Learning Objectives

- Plan a Data Collection Effort
- Draw Random and Interval Samples from a Process
- Design Checksheets to Support Data Collection

Unit Contents

- Data Collection – General Process
- Data Collection Principles
- Sampling
- Checksheets

6.2.1 Data Collection – General Process

Here is the general process you should follow in collecting data. Make sure that you clearly address the first – your goals. In our experience, this is the area where most data collection efforts "go south."

Clarify Data Collection Goals

- What questions do you want to answer?
- Link customer requirements to measures.



Develop operational definitions and procedures

- Develop operational definitions
- New vs. existing data
- Types of data
- Plan how to collect and record data
- Develop sampling procedure



Plan for data consistency and stability

- Validate the measurement system
- Consider gage R&R
- Train data collectors
- Test data collection process



Begin data collection

- Collect and analyze data
- Monitor data collection activities
- Provide feedback to data collectors



Continue improving measurement consistency

- Improve on-going data collection activities
- Assign responsibility for measurement system
- Consider sampling after improvement is seen

6.2.2 Data Collection Principles

We always get nervous when preparing to discuss this topic in front of a class or team. Inevitably, the questions boil down to "How much data do I need?" This is the least favorite question of statisticians, since everybody wants a simple thumbrule that they can apply to every situation. We'd like to oblige, folks, but unfortunately, it's not that simple.

Some general principles for collecting process data:

- **More is Better** - This is the "First Rule of Statistics." The more "good" data you have, the better off you are.

- **Cost is an Issue** - This is the corollary to the First Rule. Reality kicks in and tells you that you can't have "all the data."

- **How Fast Does Your Process "Produce" Data?** - If your process produces 20,000 items a year, you will likely need to collect more data than if it produces 20 items a year. A corollary to this is that, for some types of sampling, the amount of data required does not increase "linearly" with the amount produced.[1]

- **"Tool" Guidance** – The improvement tools often have minimum "feeding" requirements. For example, to put a "good" histogram together, you'll need a *minimum* of about 25 – 30 data points.

- **Sampling** - Sampling is almost always an option, but, in our experience, people seem reluctant to use this handy, labor-saving data collection device.

[1]Case in Point: Most political or opinion surveys require only about 1200 - 1500 people to get a pretty good picture of what the 250 million US population will say or do.

6.2.3 Sampling

Sampling (taking a portion of the data from a process) is sometimes employed when the process "produces" a lot of data, and it's too expensive or time-consuming to look at all the data. Sampling can improve the accuracy of your estimates of process characteristics or variables if collecting the data is boring or tedious.

Some sampling situations may include:

- A hospital "produces" over 30,000 Medical Records a year. The Medical Records department is interested in the accuracy of their coding process.
- A manufacturing plant produces 50,000 feet of copper tubing a week. The Quality Control department is interested in the defect rate of the tubing.
- The maintenance department "produces" about 1000 work orders a month. The maintenance supervisor is interested in the number of work orders that were held up waiting for spare parts.
- An engineering department is developing the design of a new screw compressor. They are trying to optimize performance of several key characteristics of the compressor.
- An engineer wants to determine if a new type of bearing will have a longer life on large motors.
- A pharmacologist wants to determine if a new procedure will reduce the "trough" level of a certain antibiotic in sick newborn babies.

One of your first decisions in the sampling arena is the type of study or question(s) you have. In the first three examples, the question being raised was "How Many?" How many records have errors, how many pieces are defective, how many work orders are held up for parts? For these situations, you should employ some type of Random Sampling method. We'll present two commonly used techniques in this unit and introduce several more in Unit 9.3.

For the last three situations, the question is of a "Why" or "How" variety. Here, experiments will be designed to collect data to confirm or refute some theory. Although the experiments may be performed in a randomized order, we will not be taking any random samples from the process. We are looking for the differences: with and without.

Simple Random Sampling

Purpose

Simple Random Sampling is a way to collect a portion of the data produced by a process objectively. We need to make a strong point here. There is no way to collect a random sample from an ongoing process. The conditions we'll outline below will make this obvious.

The situation arises, though, where we have a "bunch" of things already produced by the process. Using one of the previous examples, we can go to the Medical Records department and review last year's records. This "bunch" is usually called a ***lot*** in statistical work. We can take a random sample from this "bunch" or lot of records.

Simple Random Sampling will help ensure that each item in the "bunch" had an equal chance of being selected into the sample. This minimizes the chance that only an isolated portion of the process' output is contained in the sample.

Application

Simple Random Samples could be taken of the following "bunches:"

- Employee records (to see what percentage were up to date),
- A box of electronic components (what percentage meet specifications),
- A drawer full of maintenance records (to determine what types of equipment are most often failing),
- A group of patients who have been seen at an outpatient facility in the last month (to determine satisfaction levels).

Procedure for Simple Random Sampling

1. Create a numbering system for the items to be sampled. Each item must be given a unique number.

2. Select an appropriate sample size. The tool or analysis you are trying to conduct will often guide this decision. For example, to construct a "good" histogram, at least 30 points are needed. (*See Section 9 for a more detailed treatment of the "how many" issue*).

3. Select random numbers that can range from 1 to the highest number in your numbering system. This can be done from a random number table, or a random number generator, found on many calculators. For example, if the highest number in your system is 980, then you'll want to select three-digit random numbers. Select as many random numbers as you need to meet your sample size of Step 2. If duplicate random numbers appear, or numbers higher than the highest number in your system (i.e. 995), just pick another.

4. Associate the random numbers with the items' numbers. Pick these items and measure the characteristics of interest to you.

Random Number Table - Example

16408818991415353381794 01
18629819530552091962047 39
73115351014749887637990 16
57491167032316749323450 21
30405839462379214422150 59
16631350068590098275323 88
91227211993193527022840 67
50001381406632119924721 63
65390052247295828609814 06
27504961318394441575105 73

To use this table, close your eyes and put your pencil down anywhere on the table. Say you need random digits of size two. Pick the two digits next to your pencil and pick additional digits by going down, left, right, up, diagonally, any way you want. The numbers you pick are random.

Interval (Systematic) Sampling

Purpose

Interval Sampling is a process by which items are selected for the sample at some regular interval. The first item in the sample is usually selected at random.

Interval Sampling is a kind of "hybrid" sampling technique. Like Simple Random Sampling, it can be used when a "bunch" of items is being sampled. But Interval Sampling can also be used to collect data from an ongoing process. For example, every tenth coil of tubing that is received by the plant can be included in the Interval Sample.

Application

The same examples presented under Simple Random Sampling are candidates for an Interval Sample.

In addition, the Interval Sample can be applied to these type situations:

- Every third passenger entering an airplane is asked to fill out a survey,
- Every hour, one item is pulled off the assembly line and inspected,
- Every shift, readings are taken from a set of instrumentation installed on a process,
- Every fifth customer phone call is monitored for "quality assurance purposes."

Procedure for Interval Sampling

1. Identify the number of items from which a sample will be taken (N).

2. Determine the size of the sample desired (n). (*See Unit 9.3 for a more detailed treatment of the "how many" question.*).

3. Determine the sampling interval (k) by dividing the number of items by the sample size (k = N/n) and rounding up.

Note: This procedure applies when the “bunch” already exists. It can be modified slightly for collecting process data by estimating the number of items (N) to be “produced” that day, week or whatever time period is of interest.

4. Randomly select the first item in the sample between 1 and k. Call this item “j.”

5. Pick items j, (j + k), (j + 2k), (j + 3k), etc. until you’ve obtained your sample size. You may have to “cycle back” to the beginning of the item numbers to get the last sample item.

Note: Interval Sampling can lead to a distorted picture if there is any “periodicity” in your data. If the interval equals the period of the data, then the data will not be random. Silly example: Say you collected temperature data at an interval of 24 hours. This would not well represent the “average” daily temperature.

6.2.4 Checksheets

Purpose

Today, much information is collected directly into electronic form. For example, manufacturing inspection and process monitoring data can be measured on a gauge whose output is sent directly into a computer database. In healthcare, "electronic charts" have making their way into hospitals. Bar coding has greatly facilitated automatic data collection. Even data collected through review of manual charts or records can be entered directly into a spreadsheet, without the need for an intermediate manual form.

There still exists, though, a need for temporary data collection forms for improvement projects or permanent forms where computers are not yet available. The Checksheet is still an important tool of quality improvement.

Types

There are various types of checksheets used to collect data. As mentioned above, with the widespread use of computers, the Summary Checksheet described below has essentially been replaced by the modern spreadsheet:

Individual Events - These checksheets are designed to collect data on individual events. They may be as simple as the short customer survey found by the cash register in a chain restaurant, or as complicated as medical records forms or power plant maintenance record forms.

Summary Checksheets - Summary checksheets are used to summarize data collected from many individual events checksheets, or to record one or more variables from multiple products or services. For example, inspectors will sample several "widgets" from a box, measure one or two quality characteristics and record the data on a summary checksheet.

Concentration Checksheets - These checksheets not only serve to collect the data, but also to analyze the data. The concentration checksheet usually is a "picture" of the area where events of interest may occur. Several examples of concentration checksheets include:

- Power plant boiler tube failures are noted on maps of the boiler tube arrangements. Areas of high concentration can provide clues as to why the failures are occurring.

- Integrated circuit lead failures are noted on a map of the chip, in a search for processing problems.

- Chiller leaks are plotted on a drawing to show the exact location of a leak and help determine the cause and countermeasure.

- Locomotive low voltage wiring failures are mapped to determine if high temperature in the engine compartment or other variable may be causing the failures.

Designing a Checksheet

Each checksheet is different, depending on the data that is to be collected. Some general principles apply to the design and use of a checksheet:

- Keep it simple, statistician (KISS principle) - people who work in the "production" process do most of the manual data collection. They do not have time to fill out complicated forms.

- Use "checkoffs" where possible - If you are collecting performance data and also category information (e.g. certain pre-defined defect or non-conformity categories), provide check boxes for the category data rather than having the collector write in the category each time. Leave another space for the collector to write in a category that was not identified in advance.

- Follow the "flow" of the work - We designed a data input form for an improvement team that was collecting schedule information. When we tried the form, we realized that we had placed the data elements out-of-sequence; the collectors had to jump around on the form to fill in the required information.

- Leave room for comments - Leave a block for the data collector to record unusual or explanatory remarks.

- Test the form - Try the form out before you go "live" with the data collection effort.

- Train the data collectors - Explain to the collectors why they will be collecting this data, how to use the form, how the data will be used. Ask them if they have any suggestions on how to make the form better, easier, etc.

6.3 Core Data Displays

Learning Objectives

- Understand the difference between Common and Assignable Variation
- Act appropriately to address Common vs. Assignable Variation
- Calculate Measures of Central Tendency and Variation
- Prepare and Interpret Line and Run Graphs
- Prepare and Interpret Frequency Diagrams and Histograms
- Combine Part and Process Steps Average and Variation

Unit Contents

- Understanding Variation
- Data and Statistics Concepts
- Line Graphs and Run Charts
- Frequency Charts and Histograms

6.3.1 Understanding Variation

Turning Data into Information

There are many "DRIP" organizations out there today. With the advent of computers, data is taken and recorded on many different events that occur and stored for "posterity." But much of this data just sits in files or computer disks - the organization is ***D**ata **R**ich, **I**nformation **P**oor*.

One hospital captures data on every surgical procedure that occurs in their facility: surgeon, anesthesiologist, surgical staff, patient, procedure, and many different times (scheduled start time, start and end of anesthesia, start and end of procedure, etc., etc.).

Nuclear plants generate tremendous amounts of data. Every operation is documented, every maintenance action recorded, every test generates a test report. There's an old joke that every nuclear plant should have a paper-burning power plant built next to it. Before electronic records, "Nukes" even devoted a significant space in their administrative buildings to this "dead" data - the QA Vault.

Part of the problem we face is the overload of information. Applying the "philosophy" described in the previous unit (identifying **key** quality characteristics and **key** process variables) should help you focus on the most important data to help you manage your operations.

Another problem that we face is with the *presentation* of the data. Again, the computer has made the spreadsheet a very popular tool for management. Now the spreadsheet itself is not bad, in fact, we keep most of our data on spreadsheets. But how do we "extract" information from the spreadsheet? Too often, it is in a form like this:

Data	This Month	Budget	Variance	Last Month	Variance	This Month Last Year	Variance	Notes
Salary Expense	$3685	$3200	($485)	$4100	$415	$3510	($175)	Over Budget
Overtime	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
Supply Expense	$12500	$11000	($1500)	$13400	$900	$10700	($1800)	Over Budget
Volumes	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
Etc.	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
Etc.	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .

What can we learn from such a display? Well, we get this month's performance and how well we did against "budget." So that makes us happy or sad. What about the rest of the columns and comparisons? Does it help you manage to know that *this month last year* your supply costs were less than this month? Who cares?

This format can lead to the practice of "*Red Circle Management*." When your manager gets a report on your department's performance, have they ever circled one or more numbers in red, and written a big red "WHY???" next to the red circle? Have you ever been a "Red Circle Manager?" Come on, be honest!

What happens when you're subjected to Red Circle Management? From our experience, people usually reach into their "excuse file" and pull out one that hasn't been used in a while. When you ask doctors why their costs per patient are higher than the "average," they will *always* respond with "Well, my patients are sicker than others." Everybody's got similar "excuses" to answer the Red Circle WHY???

So, what's going on here? Well, as humans, we like to react to differences, whether the difference is "real" or not. The "manager" in us seems particularly interested in picking out differences. Admiral Hyman Rickover[1] used to run an experiment to "test" whether a person had an "engineer" or "manager" orientation:

> *Take two glasses (same type) and fill them each "about" halfway with water. Ask the person being "tested" to describe the glasses. The "engineer" will try to describe the* ***similarities*** *he or she sees in the glasses; the "manager" will describe the* ***differences.***

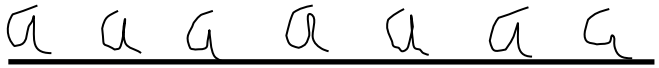
Is one or the other "right?" No, that's not the point. When we consider our processes' outputs (even filling the two glasses is a process!), they result from a large system of causes. Each output *will* be different, because the inputs will differ each time we "produce" a product or service. We say that there is **variation** in our processes. Our challenge is to understand when these differences are important and when they are not.

Fortunately, a fellow named Walter Shewhart was faced with this problem in the early days of mass-production of telephones and telephone system components. He gave us tools to objectively determine if there are "real" differences. To employ Shewhart's **run** and **control charts**, we must understand the nature of this thing called **variation**. The next few pages will provide the background that you need.

[1] "Father" of the US Navy's Nuclear Power Program and noted educator.

Variation in a Product or Service

Variation exists everywhere you look. No two products are exactly alike. Each time a service is performed, there will be differences. Even if we write seven ”identical” letters on a line, no two will be exactly alike:



A “production” process formed these letters. In this case we used a mouse, printer and a computer screen. There are many factors at work to make these letters. Each of them will vary and the sum of the factors’ variation results in overall variation.

There is profit, though, in studying the variation in our products and services. In fact, this is the key idea behind statistical quality improvement. The data you collect from your products and services is the “Voice of the Process.”

If you become skilled at listening to this voice, you’ll be able to discover the factors that are responsible for variation in products and services. You’re on the road to improvement!

This unit will provide you with some basic tools that help you understand the variation that exists in your products and services. After you become acquainted with these tools, we hope that you’ll never again be satisfied with just an average value from a set of data. One of our highly respected quality teachers, Dr. Teiichi Ando, began his lectures with this statement:

“We must leave the world of Averages
and learn to enter the world of Dispersion!”

There’s one caution we need to insert here. The tools of this unit, frequency chart and histogram are “static” pictures of variation. They should be used in combination with run or control charts to be most effective.

On the Subject of Variation

Take a coin out of your pocket. Flip it. What's the result? Did you get Heads or Tails? Why did you get Heads (Tails)? You'll probably answer, "Toss of the Coin." Well, you've just executed a process that exhibits **variation**.

Flip the coin nine more times. Record the results here:

Flip	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Result (H or T)										

Number of Heads: _____ Fraction: _____

Number of Tails: _____ Fraction: _____

You've just observed the **statistical** behavior of a process. What did you expect to get? Did you think that exactly five heads and five tails would occur? No, you probably realize that in only ten flips, you're likely to get 4 heads and 6 tails or maybe even 3 heads and 7 tails without getting suspicious about the coin. How lopsided does the result have to be before you do get "suspicious?"

Flip the coin ten more times. Did you get the same result as the first "experiment?" Can you answer WHY (remember our Red Circle Management!)?

We could progress through the evolutionary ladder of these kinds of experiments (tossing die, pulling black and white balls out of an urn, pulling cards out of a deck, Deming's Red Bead Experiment), but let's jump right to your processes.

Whenever your process "operates," the output will be a function of all the process' causes. As these causes differ from "operation" to "operation," so will the output. But we can't force these causes to be "exactly" the same each time. It's impossible. We must accept that we live in a world of **variation**.

In the early 1900's, physicists such as Albert Einstein, Werner Heisenberg, Schrodinger, Dirac and others discovered that **variation** is part of the fundamental nature of our universe at the microscopic level. This was hard for many to accept. Einstein was one of the most resistant; his famous quote: *"God does not play dice with the universe!"* spoke to his

difficulty with this issue. In the last few years, the topic of *chaos theory* has revealed the variation inherent in *macro* systems, such as weather and organizations.

But how does this affect us in our day-to-day work? We will have to accept that the outputs of our process vary and that the *best we can do* is to "control" the process within some *limits of variability*.

Action on a Process: By circling one or more numbers, our Red Circle Manager questioned the differences or variability in the process' output. Presumably, he or she first wanted an explanation for the differences and then to take some action on the process to correct the next output (i.e. next month's salary expense).

But if all our processes exhibit this thing called **variation**, then is it "wrong" to question these differences?

In some cases, the answer is "YES." If the process is exhibiting "natural" variation (that due to normal fluctuations in the input variables - we call this **common cause variation**), then we should *not* question the **point-to-point** differences in the process' output. We will only chase our tails and frustrate ourselves with this line of thinking.

Now, let's clarify one thing. We emphasized the term "point-to-point" above and that's important. It may be, that when we consider the overall performance of our process, we find there is too much variation, or that it is not operating at the right "level." We are absolutely justified in asking "WHY?" about this overall performance, but you will *not* try to explain the point-to-point differences if you conclude that your process is subject to only common cause variation.

In other cases, though, there will be a strong enough "signal" coming from the process that indicates something unusual has happened to the process. A variable within the process may have changed "significantly," a new variable may have entered the process, or something "outside" the process may be acting on the process. We call these signals **assignable** (or **special**) **cause variation.** Here it is "acceptable" to question "WHY?"

This is a fundamental concept in process management. Let's approach it from another angle. When we look at the output of a process, there are two possible CORRECT conclusions we could reach and two possible ERRORS we could make:

Correct Decision 1 - Only Common Cause Variation Present - If we conclude that there is only common cause variation present and this is the case, then we've made a correct judgment. Our *ACTIONS* in this case would include:

1. *Not* ask why there are point-to-point differences,

2. Ask whether the process was operating where it should be (based on customer needs and expectations),

3. Investigate the process variables to discover why the process was not performing where it should be.

Correct Decision 2 - Common and Assignable Cause Variation Present - If we conclude that there is an assignable cause of variation present (and there is), we've made a correct decision. Our *ACTIONS* would include:

1. Investigate the events when the assignable cause signal was present,

2. Determine what the assignable cause variable was, and

3. Determine ways to "eliminate" the assignable cause from the process.[2]

Error 1 - Just Common Cause Variation Present - NOT! - Here, we are not reacting to the presence of an assignable cause in our process. We think there is only common cause present. The error here is one of omission. These assignable causes are increasing the variation in our process and we are just sitting around doing nothing about it! Some refer to this as a "Type B" (or **B**lockhead) error.

Error 2 - Assignable Cause Variation Present - NOT! - Here, we are reacting to common cause variation as if it was due to assignable causes (the typical Red Circle Manager!). If our reaction includes changing the process variables, we can increase the variation in the process. This error is often termed a "Type A" (or H**a**sty) error.

The classic example of this is a thermostat in a room with one "hot" and one "cold" person. The thermostat controls room temperature within "limits." But as the room warms up, the "hot" person will turn down the thermostat. The room cools down even more than the thermostat would have allowed. But as the room cools, the "cold" person will turn the thermostat up, allowing the room to heat up more that the normal limits of the thermostat.

[2]If the assignable cause worsens the process' performance, then "elimination" equals preventing it from reoccurring or adding it to the list of variables we are controlling. If the assignable cause makes the process perform better, then "elimination" means to try to build it into the process as a common cause.

Deming's Funnel Experiment illustrates the different reactions of a process to "over-control" - reacting to common cause variation as if it was due to assignable causes. All these reactions *increase* the variability in the process, not reduce it!

To help us make "correct" decisions in a world of variability, we will need some tools to help us decide when we have only common cause variation present or when there is an assignable cause in our process. In fact, it would be even more useful if these tools could help us in an active search for assignable causes, rather than just waiting for them to appear as signals. The run and control charts developed by Shewhart meet this need.

Summary of Actions:

		"Our" Interpretation	
		Only Common Causes	**Common *Plus* Assignable Causes**
"True" Situation	**Only Common Causes**	Correct Decision – If Process is Not *Capable,* Act to Understand Process Variables and Improve Process	Wrong Decision – You are Overreacting to Point-to-Point Variation!
	Common *Plus* Assignable Causes	Wrong Decision – You Are Ignoring Possible Opportunities to Eliminate Assignable Causes from the Process	Correct Decision – Understand and Eliminate Assignable Causes from Process

6.3.2 Data and Statistics Concepts

We've mentioned some introductory data and statistical concepts (such as measurement and count data, mean, median, etc.) in the preceding units. Here's a refresher on these if you've seen them before or an introduction if they're new to you.

The Nature of Data

The first question we ask when presented with data is "What kind of data is it?" Here, we want to classify the data into one of two distinct types: Measurement or Count.

Our improvement tools will generally treat these types of data differently. Some tools work well with one kind of data. For instance, we found it's generally easier to create a Line Graph with measurement data than with count data.

Measurement data (synonyms for Measurement include Variables, Continuous, Analog) can be subdivided infinitely and often requires some device or instrument to measure its value. Time, speed, costs, length, pressure are examples of measurement data.

Count data (synonyms for Count include Attribute, Discrete, Digital) comes to us in the form of individual events that we count. There are some basic units that cannot be divided further. The number of errors, volumes (number of sales or procedures), defects, defectives, and number of employees are examples of count data. Count data is often "disguised" as ratios, proportions or rates.

Don't get confused by the precision with which you measure your data in this classification exercise. For example, if you measure sick days to the nearest day, you might be tempted to think of this as count data (0, 1, 2, 3, 4, . . . etc. days). Time, though, is always measurement data.

Single Point Measures of Process Performance

Line Graphs, Run Charts and others are "pictures" of the process' performance. In addition to these pictures, we often characterize a set of data by measures that help us understand where the **center** of a set of data lies (Central Tendency) and how much the data **varies** (Variability). Three popular measures for each of these are presented on the following page.

Measures of Central Tendency

Measure	Description & Use	How to Calculate
Mean	The Average of a set of numbers. The most used measure of the data's center. Remember, when you calculate an average, about half of the raw data will be above and half will be below the average - this does not translate into one half good and one half bad!!	$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ where: Σ is the symbol for "sum of" n is the number of data, and x_i are the data values
Median	The midpoint of a set of numbers placed in rank order. The median is a preferred measure of the data's center when there are very large or small values, i.e. when the data is skewed.	for an odd number of data: $x_{(n+1)/2}$ for an even number of data: $\frac{x_{n/2+1} + x_{n/2}}{2}$
Mode	The most frequently appearing number(s) in a set of data. Useful when data displays wide variation, perhaps due to mixed processes.	For the data set: 1,2,2,3,3,3,3,4,4,5,5,6,7 three is the mode

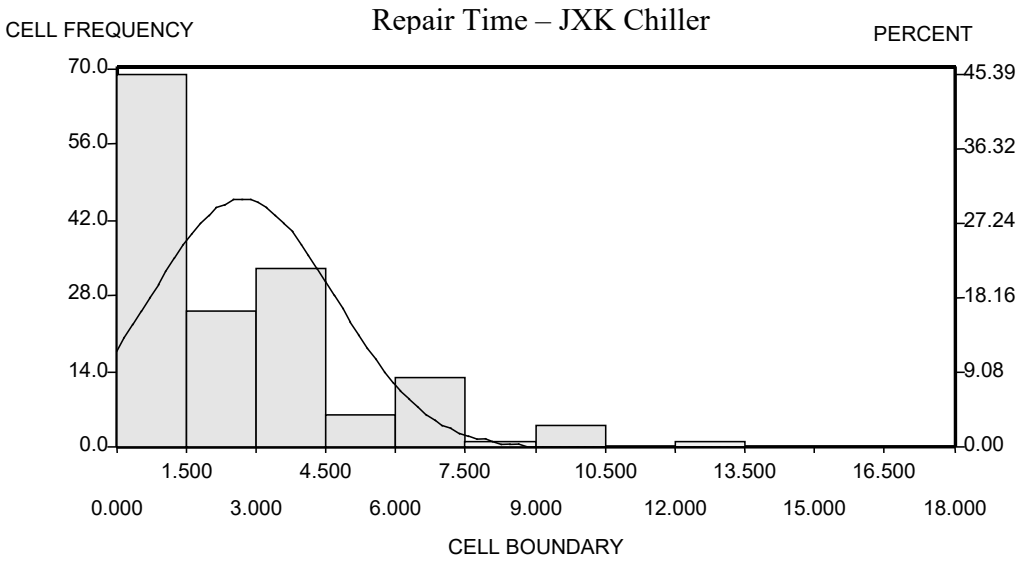
Measures of Variability

Measure	Description & Use	How to Calculate
Range	The difference between the largest and smallest values in a data set.	$R = x_{max} - x_{min}$
Variance	The sum of the squared differences of the data from the mean, divided by the number of data less one.[3] Forms the basis for the standard deviation.	$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$
Standard Deviation	The square root of the variance. This is the "best" measure of variability, since it considers all the data from the sample. The Standard Deviation can be thought of as a "distance" measure - showing how far the data are away from the mean value.	$s = \sqrt{s^2}$

[3]This is the *sample* standard deviation. If the entire *population* is known, there is no need to subtract one from *n,* the number of data.

Skewness and Kurtosis

Here, we discuss two additional statistics used to describe a set of data, *skewness* and *kurtosis.* These are oft reported, but little understood statistics. For example, the following histogram was obtained from a sample of repair times for a certain model chiller (measured in hours). A popular statistical software package was used to develop this "picture." Notice the table of statistics appearing below the histogram. What value are the skewness and kurtosis in statistical quality control?



Fitted curve is a Normal. K-S test: 0.000. Lack of fit is significant.

Total number analyzed	=	152
Group Range:		1-152
Average	=	2.625
Process sigma	=	1.984
Population sigma	=	2.314
Sample sigma	=	2.321
Standard error of mean	=	0.188
Skewness	**=**	**1.5887**
Kurtosis	**=**	**2.5109**

Moments of a Set of Data

The histogram provides a very useful graphical representation of a set of data, but there are strong advantages to being able to characterize a set of data *numerically.* Comparisons are much easier when a few numbers can be looked at and differences examined. By now, the reader is already familiar with the calculation and use of the *mean*, *range, standard deviation (and variance)* and *proportions* or *rates*. We have never, though, defined the origin of these statistics.

Some of these are easily motivated. The *range*, for example, provides us with a simple measure of the data's spread. Whenever a situation arises where *m* of *n* items has some characteristic, or *m* events have occurred in a given time, it is natural to consider a *proportion* or a *rate.* But how about the *mean* and *standard deviation?* Is there any rational for how these quantities are calculated? The answer lies in the concept of a *moment*. People trained in the physical sciences are familiar with various *moments*, such as the *center of mass* and *moment of inertia*. These quantities provide information that helps the analyst quantify the characteristics of a body or mass of certain shape, size and density. For any set of data (x1, x2, x3, . . . xn), the "kth" moment of the data around the origin (0) is defined to be:

$$m'_k = \frac{1}{n}\sum_{i=1}^{n} x_i^k$$

$$where:$$

$$m'_k \text{ - "kth" moment around the origin}$$

$$n \text{ - number of data in sample}$$

Since these moments are based on samples of data, their "formal" name is the *sample moment.*[4] Now the *first* moment of a data set is simply the *sample mean*, which may be considered analogous to the "center of mass" of the data:

$$m'_1 = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

[4]Moments for populations are calculated slightly differently, these are not addressed here since the analyst will rarely be dealing with population data. See Hoel, Introduction to Mathematical Statistics, John Wiley & Sons for a detailed discussion of moments and their generating function.

Now the second and higher moments can be developed, but it is more useful to define *moments about the mean of the sample:*

$$m_k = \frac{1}{n}\sum_{I=1}^{N}(x_i - \bar{x})^K$$

The second moment about the mean begins to look familiar:

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

This moment, adjusted for bias, is the sample variance:

$$s^2 = \frac{n}{n-1}m_2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \; or, \; s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$where:$$

$$s \text{ - is the sample standard deviation}$$

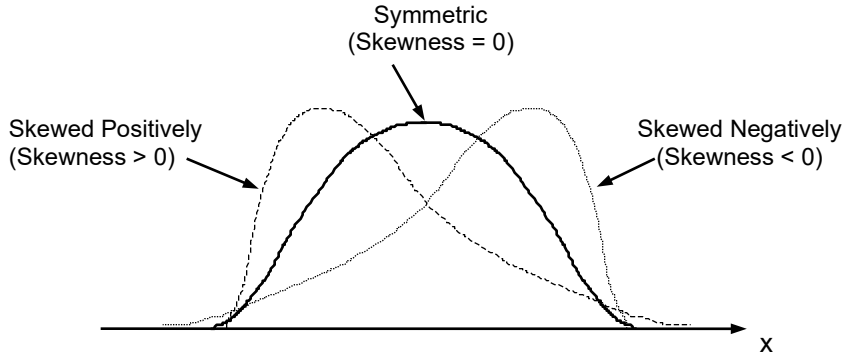
The third and fourth moments can be developed into the *skewness* and *kurtosis* statistics, as shown below. Higher moments (> 4) can be calculated, but these are of little practical interest.

Skewness

The third moment about the mean is known as the *absolute skewness:*

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$$

For a symmetric distribution, this *skewness* will be zero. If the data are not symmetric, perhaps with a long tail to the right, the skewness will be a positive value, and the data is *positively skewed.* On the contrary, if the data's tail is to the left, the skewness will be negative, and the data will be *negatively skewed.*



The absolute skewness is used rarely; a *relative skewness* is most discussed. Shewhart (ref. I2) presents this formula for relative skewness:

$$k = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}} = \frac{\sqrt{n}}{(n-1)^{3/2}s^3}\sum_{i=1}^{n}(x_i - \bar{x})^3$$

The transformation of the absolute skewness by dividing it with the second moment (taken to the 3/2 power) changes the values of the statistic, but not the interpretation (symmetric = 0, positively skewed - +, negatively skewed - -).

Kurtosis (Flatness)

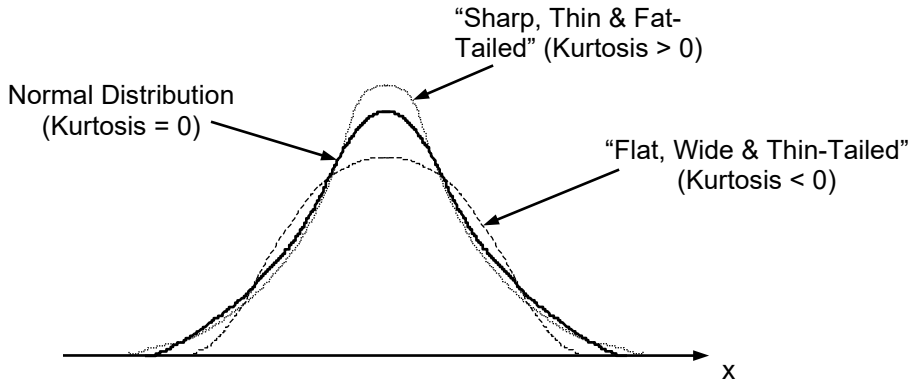
Where the third moment measures the skewness of the data, the fourth moment measures the *flatness* or *kurtosis* of the data.[5] Like skewness, *kurtosis* can be calculated in either absolute or relative form. The *relative* kurtosis is shown below:

$$\beta_2 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^2} = \frac{n}{(n-1)^2 s^4}\sum_{i=1}^{n}(x_i - \bar{x})^4$$

The analyst must be careful when interpreting the *kurtosis.* Normally distributed data will have a kurtosis equal to 3. Some texts (and presumably, software programs) measure kurtosis *relative* to the normal distribution, by subtracting 3 from the kurtosis. The following interpretation is based on kurtosis as measured by this correction.

$$\beta_2' = \beta_2 - 3$$

An "adjusted, relative kurtosis" of 0 indicates the data is shaped like a normal distribution. Kurtosis greater than 0 indicates the data will have a sharper peak, thinner "shoulders" and fatter tails than the normal. Kurtosis less than 0 means the data will have a flatter peak, wider shoulders and thinner tails than the normal.



[5]In fact, Shewhart termed *kurtosis,* "flatness."

Application to Quality Improvement

Normality Tests

One application of skewness and kurtosis measures is found in certain *normality tests.* These hypothesis test procedures start with a set of data that, on a histogram, may appear to be normally distributed. Skewness and kurtosis statistics are calculated and compared to reference values (dependent on sample size and α-risk). A decision to reject or not reject the hypothesis of normality is then reached.

If the normality hypothesis *is not rejected*, then further statistical analysis of the data *based on the assumption of normality* will likely occur. Note what is being stated here. If enough evidence *cannot* be found to "convict" the process of non-normality, then the assumption is made that it is normal. This test, then, is the key that opens the door to all statistical tests that presume the data to be normally distributed. This logic is a bit like the case where, since a person was *not* convicted of a crime, they are then elected to chief of police because of their "trustworthiness."

Role in Describing Data Distributions

Shewhart went to great pains to show that the information contained in the first two moments (the mean and standard deviation) was essentially all you needed to know to characterize a set of data. He discusses the problem of recreating a distribution of data from knowledge of just the basic statistics and an assumption of the frequency distribution. He shows, that, while the skewness adds a "little," *for practical purposes,* the mean and standard deviation alone will recreate the distribution.

He further concludes that moments higher than the second are of "little value unless there is also given some function involving these statistics. . . " Shewhart was referring here, to an assumption of some probability or frequency distribution function.

6.3.3 Line Graphs and Run Charts

Line Graphs

Purpose

Line graphs are basically graphs of your performance measure taken over time. They help you see where the "center" of the data tends to be, the variability in performance, trends, cycles and other patterns. Line graphs are very simple to construct. One of the most important factors to keep in mind for line graphs is that the data must be plotted in the order in which it occurs. Losing this order will prevent you from seeing patterns that are time dependent.

Application

Virtually any data can be placed on a line graph (as long as you've kept it in order of occurrence). Some typical line graph applications include:

- Quality Indicators - Turn-around Times, Errors, Defect Rates, Defective Proportions, Physical parameters - Condenser Vacuum, Machine Start Times, Setup Times, Pressure, temperature readings taken periodically, chemical or drug concentrations (peak and trough levels),
- Personal data - Weight, heart rate,
- Financial Data - Salary Expense, Supply Costs, Sales, Volumes.

Construction of Line Graphs

1. Draw a vertical and a horizontal axis on a piece of graph paper.

2. Label the vertical axis with the variable being plotted.

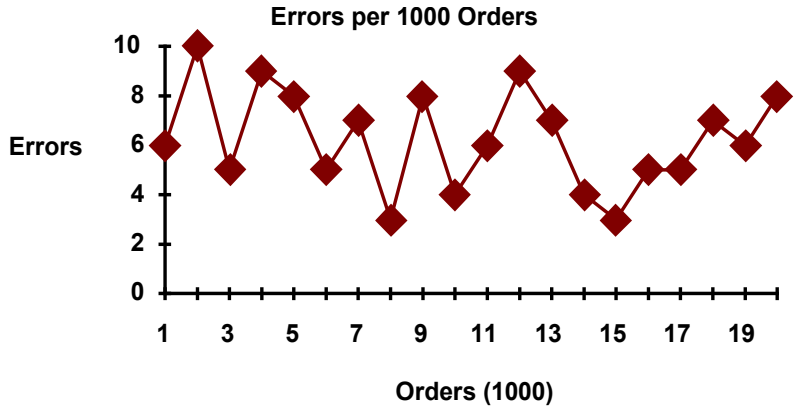
3. Label the horizontal axis with the unit of time or order in which the numbers were collected (i.e. Day 1, 2, 3, . . ., Customer 1, 2, 3, . . . etc.).

4. Determine the scale of the vertical axis. The top of this axis should be about 20 percent larger than the largest data value. The bottom of this axis should be about 20 percent lower than the smallest data value. This let's you see the best picture of the process' variability. Label the axis in convenient intervals between these numbers.

5. Plot the data values on the graph number by number, preserving the order in which they occurred.

6. Connect the points on the graph.

7. (Optional) Calculate the mean of the data and draw this as a solid line through the data. This turns the line graph into a run chart - trends and patterns are often easier to see with a run chart.



Construction Notes

Try to get about twenty-five (25) data points to get a line graph running. If you have less, go ahead and plot them anyway. It's good to start trending performance no matter how many points you currently have (i.e. if you process only "produces" one data per month - salary expense, supply costs, etc. - don't wait two years to start your line graph!).

Now, here's how you get these 25 points for a line graph. If you are dealing with **measurement data** (time, cost, etc.), then each event you measure represents a data point to be plotted. Each patient's temperature measurement could be plotted on a line graph: 98.5, 98.7, 98.6, 99.0, 98.4, etc.

If you are dealing with **count data**, though (or even worse, percentages made up of count data), then there are a few guidelines that may cause some data "heartburn." For typical count data, the guideline is that the **mean** of the data you plot should at least equal to 5, and no less than 1. Let's say you are counting the number of errors that occur daily. You get these numbers for a week's worth of errors: 7, 10, 6, 5, 8, 7, and 6. The mean number of errors (daily) is 7. This number is greater than 5, so you can plot the daily values as individual point.

We apply this rule for two reasons. First, to "see" variation in the process, we need to keep the data away from the horizontal (0 value) axis. The second reason lies in why you are taking the data in the first place: to take action. If you want to detect whether your change has had an effect, you'll want to see its impact on the line graph.

Now let's look at a different set of values. In counting orders for a particular specialty magazine (again, daily), a publications distributor finds that their first week gives this data: 0, 1, 0, 2, 1, 0, 1. Here, the daily mean value is less than 1. This mean doesn't meet the guidelines and plotting these data won't produce a very useful line graph.

The distributor could group the data by combining enough days to make the mean equal or better than 5. In this case, there are 5 orders occurring per week. So, instead of plotting the daily occurrence of orders, they plot the weekly orders. To get a line graph going here, note that they are going to have to observe at least 125 events (25 points x 5 - mean).

This is difficult since it's now going to take them about 25 weeks to get a complete line graph instead of only 25 days. This kind of thing happens often when we start to stratify processes that are low volume to begin with for the company down to an individual department. The process just doesn't give us enough data for a line graph.

One way of getting around this problem is to plot the time between events. For example, one company was studying employee injuries. They measured the time between injuries. Since this is measurement data, it "only" took 26 injuries to get a good line graph going.

Percentage (or Proportion) Data - The guideline for plotting one point on a line graph (where the percentage is count data divided by count data, i.e. errors per 1000 orders) is that the numerator's mean should be greater than or equal to 3 and the denominator's mean should be greater than or equal to 50. You can see the implications of this on the amount of data and time needed to get a line graph going.

Run Charts

Run charts are graphs of data over time or sequence. They are used to display variation and determine if special cause and/or common causes of variation are present.

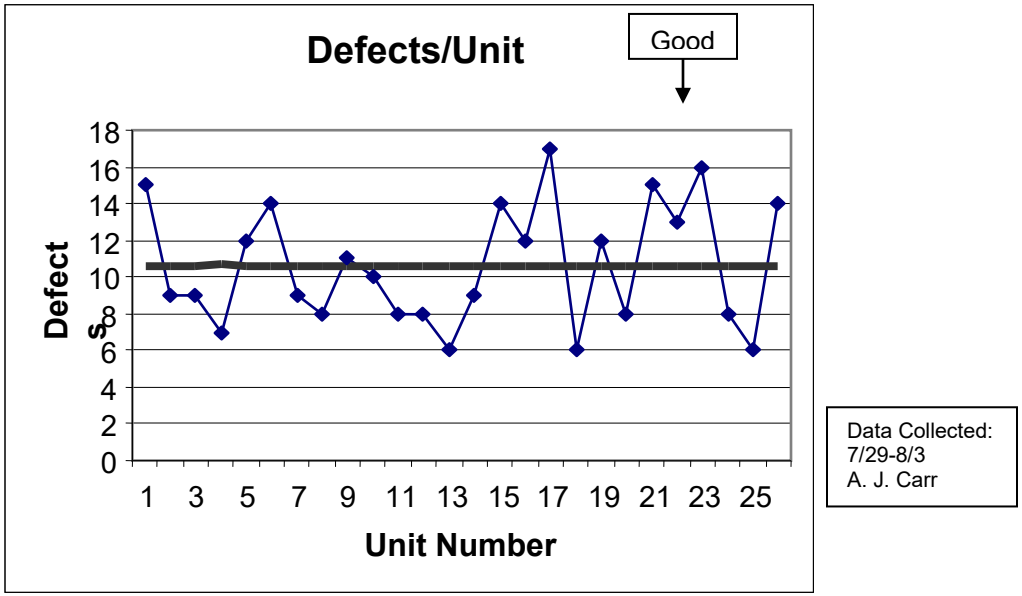
Construction of Run Charts

1. Draw a set of axis and label them with the time or sequence of the data on the X-axis and the measure on the Y-axis.

2. Scale the Y-axis so it shows values 20% above and zero or 20% below the values to be plotted.

3. Plot the data in the sequence they occurred and connect the data points with lines. These lines denote that the data is sequential. To evaluate process variation, at least 25 data points are needed. More is better.

4. Calculate the mean and plot it on the graph as a reference. Be sure to label the graph and show the source of the data. Note the mean value should be at least 5 or greater to be able to interpret the run chart for special causes. It's always good to show a target and the direction of improvement on graphs.

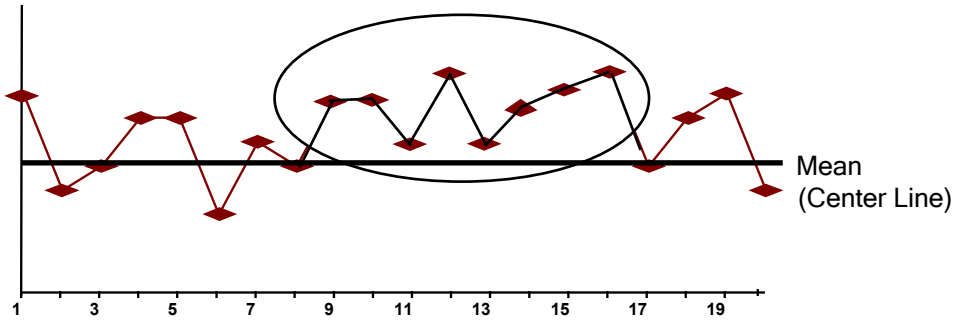


Run Chart Interpretation

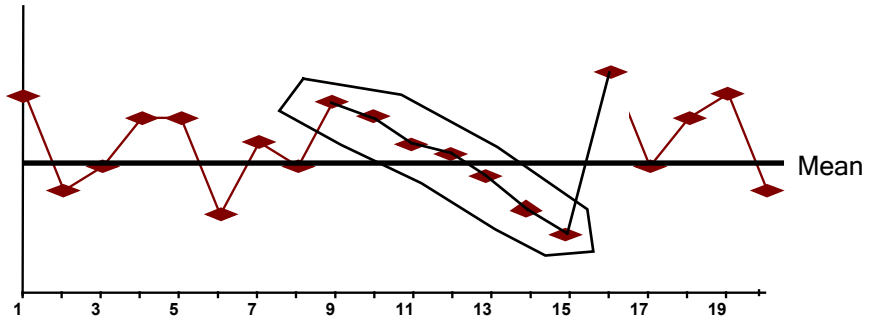
Random patterns of data on run charts depict common cause variation. Common cause variation is always present. Non-random patterns note special (assignable) causes or that something has changed in the process.

Patterns to look for are:

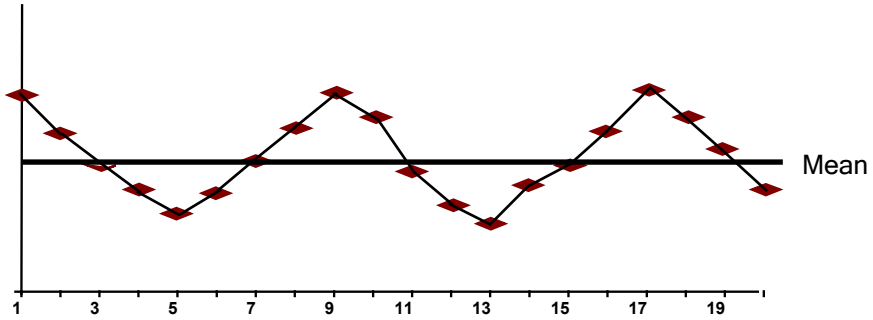
- Shifts, 8 or more consecutive data points either above or below the centerline. Points on the centerline are ignored and do not make or break a shift.



- Trends, 7 or more consecutive data points going up or down. Equal consecutive values are ignored and do not make or break a trend.



- Repeating patterns, any non-random pattern may be a special cause signal. Generally, if the pattern occurs 8 or more times it should be investigated.



- Extreme Values, isolated values that are extremely high or low with respect to the rest of the values may be a special cause. Single values can be difficult to interpret with run charts. Control charts are better for identifying single points.

Dealing with special causes

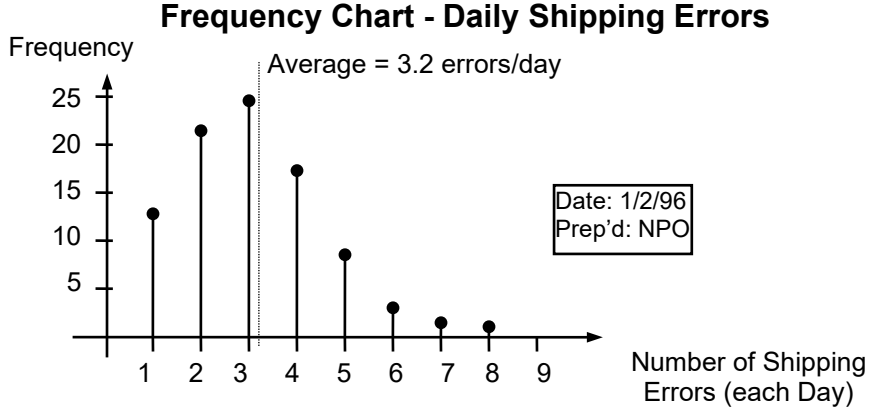
When a special cause is noted in the data, you should investigate what caused it. A change in the process is a prime suspect. Look for different materials, equipment, people or procedures. Special causes may not be “bad,” they could be something that you want to repeat.

6.3.4 Frequency Charts and Histograms

The Frequency Chart

Purpose

The frequency chart helps you display the variation in a set of count data. The count data could be the number of defects or defective items identified in samples (or lots) of your product. In general, when you record the number of events that occur in a given time, or in a given sample, and you do this repeatedly, the frequency chart will be useful. You might display the variation in volumes (i.e. number of surgeries performed each day, number of customers, number of telephone calls, etc.) on a frequency chart.



Here, the Shipping Department has kept a daily tally of the number of errors made. The horizontal axis of the chart shows the range of errors, each day had at least one error, but no day had more than eight errors. The vertical lines show how many times a day that number of errors occurred. For example, it appears that in thirteen days one error was made, on 22 days two errors were made, etc. Stepping back and looking at all the lines, you get a picture of the shape or distribution of the data. Note that if your count data can assume many different values, the histogram might be a better tool. Generally, if there are more than about 20 - 25 different data values, the histogram should be used.

Application

Frequency charts are applied in various steps of quality improvement:

Identify the Problem - The frequency chart can help show that there is a need to improve a product or service. If the average number of errors is too high, or if the pattern shows some unusual shape (such as outliers), there may be a need to improve the process.

Analyze Causes - Frequency charts may be prepared for various strata of the process. Different machines, methods, personnel, plants or departments may be examined. Here, you are beginning to break down the variation that you see in the process' output.

Implement/Evaluate Results - Before and after frequency charts will show the effects of changes made to the process.

Construction

1. Collect the count data to be displayed on the frequency chart. At least 25 - 30 data should be available, preferably closer to 50 data. Be careful that the events you're recording come from approximately the same area of opportunity. In the example above, if the number of shipments varied widely from day to day, the shipping errors (events) would not be coming from the same area of opportunity. The Shipping Department might want to display the number of errors per 100 shipments to make the area of opportunity the same.

2. Determine the range of the events - the smallest number and the highest number. Develop a tally sheet to record the number of times each value appears:

SHIPPING ERRORS - TALLY SHEET

# Errors	Tally	Frequency
1	𝍸 𝍸 ///	13
2	𝍸 𝍸 𝍸 𝍸 //	22
3	𝍸 𝍸 𝍸 𝍸 𝍸	25
4	𝍸 𝍸 𝍸 ///	18
5	𝍸 ///	8
6	////	4
7	//	2
8	//	2

3. Draw horizontal and vertical axes. Label the horizontal axis with the values (i.e. number of errors) and the vertical axis with a convenient scale to display the frequencies.

4. For each value, draw a vertical line from the horizontal axis to the appropriate frequency value. Draw a small circle at the top of the line.

5. Title and label the chart. Include the date and who prepared the chart.

6. Optional: Calculate the average number of events. Draw this as a dotted line on the frequency chart. If you are calculating the average by hand, the tally sheet can help simplify the calcs:

# Errors	Frequency	# Errors x Frequency
1	13	1 x 13 = 13
2	22	2 x 22 = 44
3	25	3 x 25 = 75
4	18	4 x 18 = 72
5	8	5 x 8 = 40
6	4	6 x 4 = 24
7	2	7 x 2 = 14
8	2	8 x 2 = 16
Totals	94	298
Average = 298/94 = 3.2 errors/day		

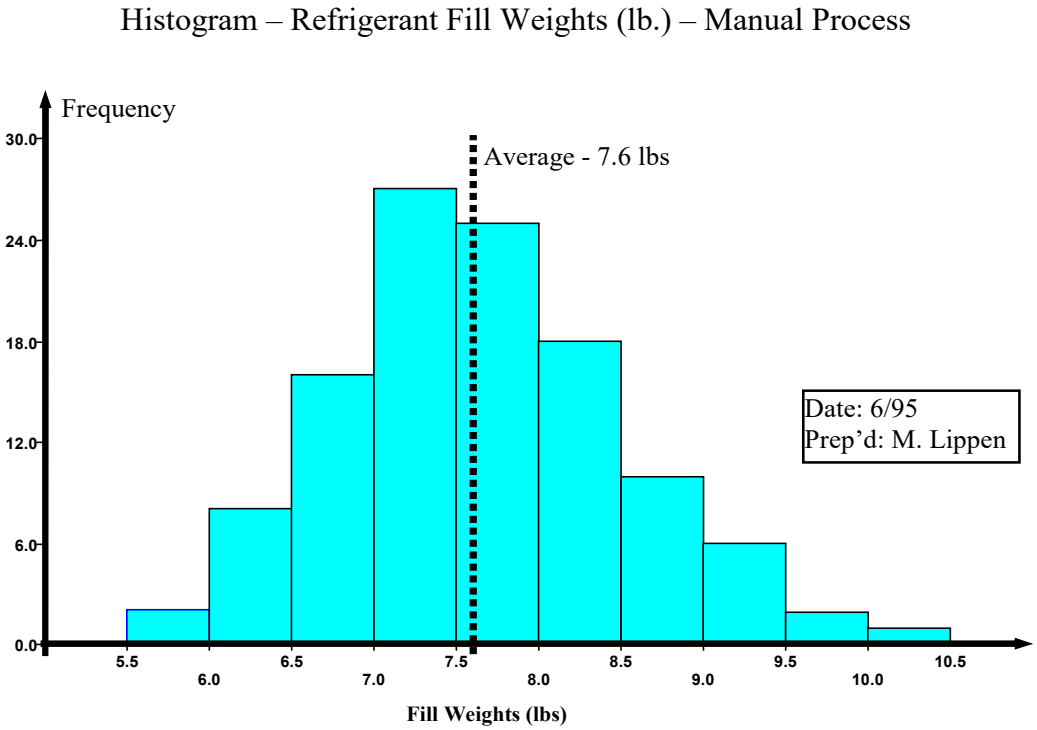
7. Interpret the frequency chart (see **Interpretation**, later in this unit).

The Histogram

Purpose

The histogram helps you display the variation in a set of measurement data (or if your count data can assume many different values). The histogram provides a picture of the distribution of the data.

In this example, a one-month sample of refrigerant fill weights was collected. The horizontal axis shows the range of fill weights, from 5.5 to 10.5 lb. The height of each "cell" represents the number of cooling units whose fill weight fell in the range of that cell. For example, 18 units were filled with refrigerant weighing between 8.0 and 8.5 lb.



In this example, the shape of the data is symmetric around the average value of 7.6 lb. and tails off rapidly as we move away from the average on either side. This kind of data may be described (or modeled) by the Normal Distribution. If this data had a different shape, another distribution such as the Lognormal, Weibull or Exponential could be used to model the process.

These models are mathematical descriptions of the data and are the next step beyond the histogram. If the process that produced this data is influenced only by the random variation of its factors, then we may be able to predict the future of the process.

We could make statements about the likelihood of a unit being filled with a certain refrigerant weight range, or what fraction of units will be filled above or below a certain weight. These are important applications in the study of variation.

Application

Histograms are applied in various steps of quality improvement:

Identify the Problem - The histogram can help show that there is a need to improve a product or service as part of a Process Capability Study (i.e. is the process capable of meeting customer requirements?).

If the average value is too high (compared to the customer's requirement or specification), or if the variation (spread) of the data is too high, or if the pattern shows some unusual shape, there may be a need to improve the process.

Analyze the Cause - Histograms may be prepared for various strata of the process. Different machines, methods, personnel, plants or departments may be examined to break down the variation seen in the process' output.

Implement/Evaluate Results - Before and after histograms will show the effects of changes made to the process.

Construction

Note: The histogram's construction is bit complicated, due to two important issues that relate to obtaining the best picture of your data: 1) applying the "right" number of cells to the data, and 2) ensuring that the data falls into the cells "appropriately."

1. Collect the measurement data to be displayed on the histogram. At least 25 - 30 data should be available, preferably closer to 50 data. Count the number of data (we'll call this n). Also, note the measurement unit of your data - did you measure to the nearest whole unit (i.e. 1 pound), or to the nearest tenth or hundredth (i.e. 0.1 lb., or 0.01 lb.)?

2. Calculate the range of the data:

$$Range = x_{\max} - x_{\min}$$

where:

$x_{\max}$ - Largest Data Value

$x_{\min}$ - Smallest Data Value

3. Calculate the approximate number of cells:

$$\#\ \text{Cells} = \sqrt{n}$$

(approximate - don't round off!)

where:

n - number of data in sample

This rule works very well. As the number of data increases, the number of cells will increase, but at a slower rate:

# Data	# Cells (Approx.)
25	5
50	7
100	10
150	12
200	14

4. Steps 2 and 3 are now used to determine the width of each cell. This is a two-step process. First, we'll calculate an approximate cell width:

$$\text{Cell Width (approx.)} = \frac{Range}{\#Cells\,(approx.)}$$

Next, we'll round this off to the nearest multiple of your data's measurement unit. Here are a few examples:

Cell Width (approx.)	Measurement Unit	Cell Width (Corrected)
0.54 lb.	0.1 lb.	0.5 lb.
14.3 minutes	1 minute	14 min.
0.263 inches	0.05 inch	0.25

5. Now that we know how wide the cells are, we'll determine where to start the first cell. This is called the Lower Bound of the First Cell (LBFC):

$$LBFC = x_{\min} - \frac{Data\ Precision}{2}$$

This correction factor prevents any of your data from falling on a cell boundary.

6. Now, prepare the following tally sheet to identify the range of each cell and record the number of times the data falls in each cell:

Histogram Tally Sheet

Cell #	Cell Boundaries	Tally	Frequency
1	LBFC - C2	𝍸 /	6
2	C2 - C3	𝍸 𝍸 ///	13
3	C3 - C4	𝍸 𝍸 𝍸 𝍸 𝍸	25
4	C4 - C5	𝍸 𝍸 𝍸	15
5	C5 - C6	𝍸 ///	8
6	C6 - C7	////	4
7	C7 - C8	//	2
8	C8 - C9	/	1

The first cell's lower boundary is the LBFC. Its upper boundary (C2) is the LBFC plus the cell width. Each of the remaining cell boundaries (C3, C4, etc.) is obtained by adding the cell width to the upper boundary of the previous cell.

Continue creating cells until the largest data value is "contained" by a cell. Tally the number of data that fall into each cell.

7. Draw a horizontal and vertical axis. Label the horizontal axis with the variable being measured and divide and scale the axis into the number of cells. Label the vertical axis with a convenient scale to display the cell frequencies.

8. For each cell, draw a bar from the horizontal axis to the appropriate frequency value.

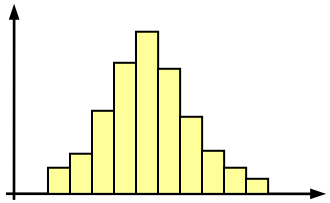
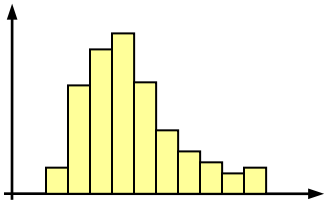
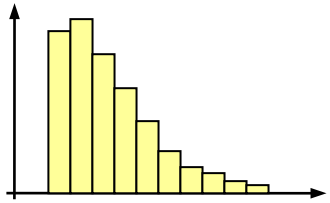
9. Title and label the chart. Include the date the chart was prepared and who prepared the chart.

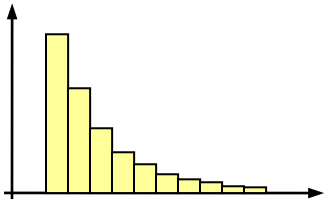
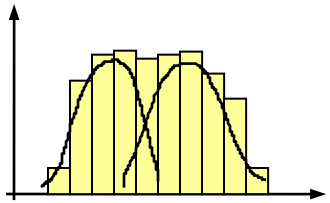
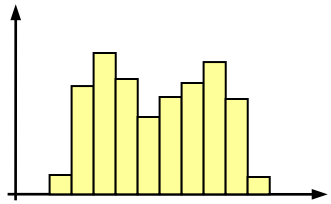
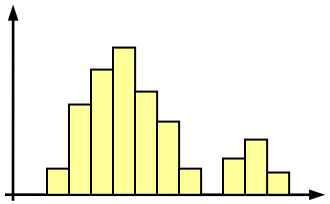
10. Optional: Calculate the average number of events. Draw this as a dotted line on the histogram. Calculate the standard deviation of the data. Record these three quantities on the histogram.

11. Interpret the histogram (see **Interpretation**, later in this unit).

Frequency Chart and Histogram Interpretation

The shapes of these charts can give us clues about what might be happening in our process. Here are some common shapes and their interpretation. The first four are shapes that "appear in nature," depending on the type of process that is at work. The last four are indications of something odd in either the data or the process.

Shape	Interpretation
Symmetrical 	Many processes' outputs take this shape, especially those where an attempt is being made to produce the product or service at some target or nominal value. If a data sample is periodically obtained from a random process and the average of the sample is calculated, the histogram of averages will always assume this shape.
Skewed 	The time to complete a process will often appear skewed. Most of the events will fall in a clump to the right or left, with only a few data in a "tail." Other data that often appear skewed are times or cycles to failure.
Extreme Skewness 	Here the data appears to be pushed up against some boundary. This is often the case when there is a lower or upper limit that the data can assume. For example, some time data can appear extremely skewed when it is possible to complete the process in very short times (close to zero). If a product is inspected and rejected if it does not meet a specification, the "good" products will take on this shape after inspection.

Shape	Interpretation
Exponential 	This shape can appear if there is a "birth-to-death" process being measured and the failure time is measured. This is also the shape that a radioactive decay process will produce and either the quantity of material remaining or the disintegration rate is measured periodically.
Plateau 	This is a shape where we suspect that two processes are being mixed together, whose mean values are not very far apart. Another time this shape can appear is when an automatic compensating control is fitted to a machine or process. When the process output reaches a certain value, the control adjusts the machine to a lower (or higher) value.
Twin-Peaked 	This is an example of two processes, whose outputs are mixed together. For example, two production lines' outputs are mixed and the histogram data is collected at a final inspection point. This is an invitation to stratify the data. Usually, one of the processes will perform better than the other. If you can understand why, then the other process can be changed to perform as well.
Outliers 	Sometimes, special circumstances will cause the production process to produce outliers. For example, during a colonoscopy procedure, some patients become rather resistant. The time required to "produce" these colonoscopies will show up as an outliers when combined with data from normal procedures.

Combining Factors' Variation

The Problem

The frequency chart and histogram provide us with pictures of the data; they tell you three key features of your process:

- The Center of the Process
- The Variability of the Process
- The Shape of the Data

With these three pieces of information, you've captured the essence of your process. Now, most of the time, improvement will drive us to try and break down the variation in the process; what causes the variation that we see? Occasionally though, you may want to combine the variation that exists in two or more factors:

- A manufacturing-type example: Two parts are to be joined into one component. If the individual parts are produced with the following dimensions, what will be the average and standard deviation of the combined parts? This will determine if the component can meet customer specifications.

Part	A	B
Average Value	1.500"	2.000"
Standard Deviation	0.015"	0.020"

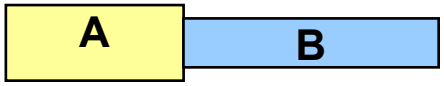
- A service-type example: A new fabrication procedure has three basic time segments, Setup, Fabricate, Inspect. If the individual segments are "produced" in the following times, can we predict the overall average time and standard deviation? This will be helpful in determining how to schedule for this procedure.

Segment	Setup	Fabricate	Inspect
Average Value	15 min.	35 min.	20 min.
Standard Deviation	5 min.	10 min.	7 min.

Combining Average Values

Average values can simply be added together (or subtracted, depending on how the variables are combined). For our two examples:

Combining Parts:



The average length of the component (A + B) is:

$$L_{A+B} = 1.500” + 2.000” = 3.500”$$

Combining Procedure Segments:



The average procedure length is:

$$T_{Setup + Fabricate + Inspect} = 15 \text{ min.} + 35 \text{ min.} + 20 \text{ min.} = 70 \text{ minutes}$$

Combining Variation

The process is not as easy if we want to combine variation. Adding the standard deviation values, as we did with the averages, may seem to be a practical method, but it turns out that this gives us too large an estimate of the combined variation.

The proper method of combining individual variations is to add the variances of the components. This holds regardless of whether the average values are being added or subtracted:

$$\sigma^2_{A+B} = \sigma^2_A + \sigma^2_B$$

where:

σ^2 - Variance of the Data

The standard deviation is then the square root of the combined individual variations. For our two examples:

Combining Parts' Variation:

$$\sigma^2_{Component} = (0.015)^2 + (0.020)^2$$
$$= 0.000625 inches^2$$
and
$$\sigma_{Component} = \sqrt{0.000625}$$
$$= 0.025"$$

Combining Fabrication Segments' Variation:

$$\sigma^2_{Procedure} = (5)^2 + (10)^2 + (7)^2$$
$$= 174 \min^2$$
and
$$\sigma_{Procedure} = \sqrt{174}$$
$$= 13.2 \min$$

From this information, we could make some predictions about the possible range of values for the combined parts or entire procedure. For example, we would expect very few components to be produced outside a band that extended +/- 3 standard deviations from the average size. The component variability could then be stated as:

3.500" +/- 3 x 0.025" or 3.500" +/- 0.075"

This could be compared to the customer's specifications to determine if the component can be produced inside the specs.

Likewise, very few fabrication procedures would be completed outside the following time band:

70 min. +/- 3 x 13.2 min. or 70 min. +/- 39.6 min.

This could be input to the scheduling system to determine how long to schedule a fabrication procedure.

Caution – *The method shown above for combining variation only applies when the relationship between the Y and the X's is linear. For non-linear situations, Monte Carlo simulation or other advanced methods are required.*

6.4 Introduction to Control Charts

Learning Objectives

- Understand the Purpose of the Control Chart
- Understand the Theory of the Control Chart

Unit Contents

- Introduction
- Control Charts
- “Simple” Explanation of Control Charts

6.4.1 Introduction

In 1931, Walter Shewhart published a book that should be required reading (at least twice!) for anyone who claims to be "expert" in statistical process control. The *Economic Control of Quality of Manufactured Product* did just what its title states; it laid the foundation and provides basic methods of *economically* controlling product (*and* service) quality.

Here, Dr. Shewhart tackles the problem of variation in the "production process." He first recognizes that there will always be variation in products or services produced by the process. In the 1920s, this was a radical departure from accepted philosophy. It was assumed then that only lack of knowledge of *all* the causes of variation was preventing the production of completely uniform, "identical" products/services.

With this inherent variation in product or service, Shewhart then proceeds to develop methods (the control chart) of distinguishing *assignable causes* of variation from the variation produced by a *constant (or common) cause system.* His control charts are *economical,* that is, they minimize the likelihood of our looking for "troubles" (his favorite term for *assignable causes*) when they are not present, and of ignoring "troubles" when they are present.

Dr. W. Edwards Deming encouraged Shewhart to turn a series of lectures delivered to the Graduate School of the Department of Agriculture into a subsequent book published in 1939, *Statistical Method from the Viewpoint of Quality Control*. This companion volume provides additional insight into Shewhart's 1931 work, particularly in *statistical control as a process, the role of measurement as a means of gaining knowledge,* including *the presentation of data,* and the role (or lack of!) of tests of significance and "true values" in science, engineering and the everyday work of production. Again, this volume is required reading, and, again, it probably needs to be read at least twice to understand its message.

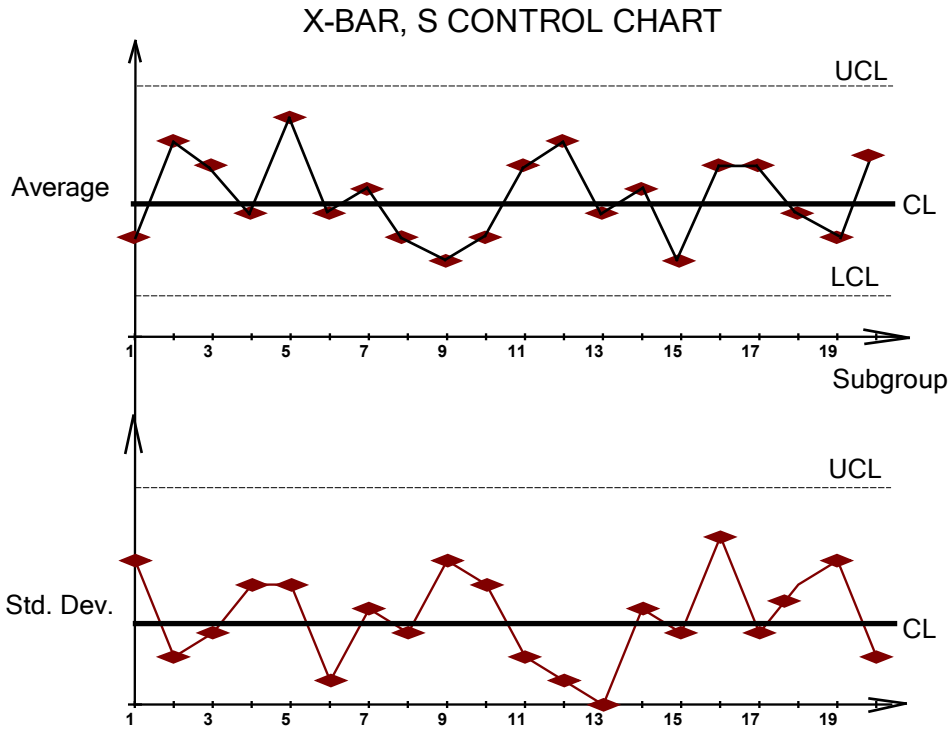
This Unit attempts to provide the background and theory of control charts and their role in the *operation of control.* We present a "simple" explanation of the control charts here. If you are interested in more detail, obtain a copy of Walter Shewhart's *Economic Control of the Quality of Manufactured Product* (for the original explanation) or one of Don Wheeler's texts on SPC.

6.4.2 Control Charts

To create a control chart, we'll start with a *run* chart and add a bit more complexity, and a *lot* more power. A Control Chart is (again) a graph of some quality characteristic or process variable. For Control Charts, though, we will add lines to the graph called *Control Limits (Upper and Lower)* that are calculated based on the data from our process.

Although they are most often used to track a process over time, we'll relax this requirement. When we use control charts to help us analyze a process, a more general concept called **subgrouping** will be invoked. In fact, Shewhart indicated that just putting the "over-time" data on a control chart was the least preferable strategy.

Here's what a typical control chart will look like:



Choosing the Best Control Chart

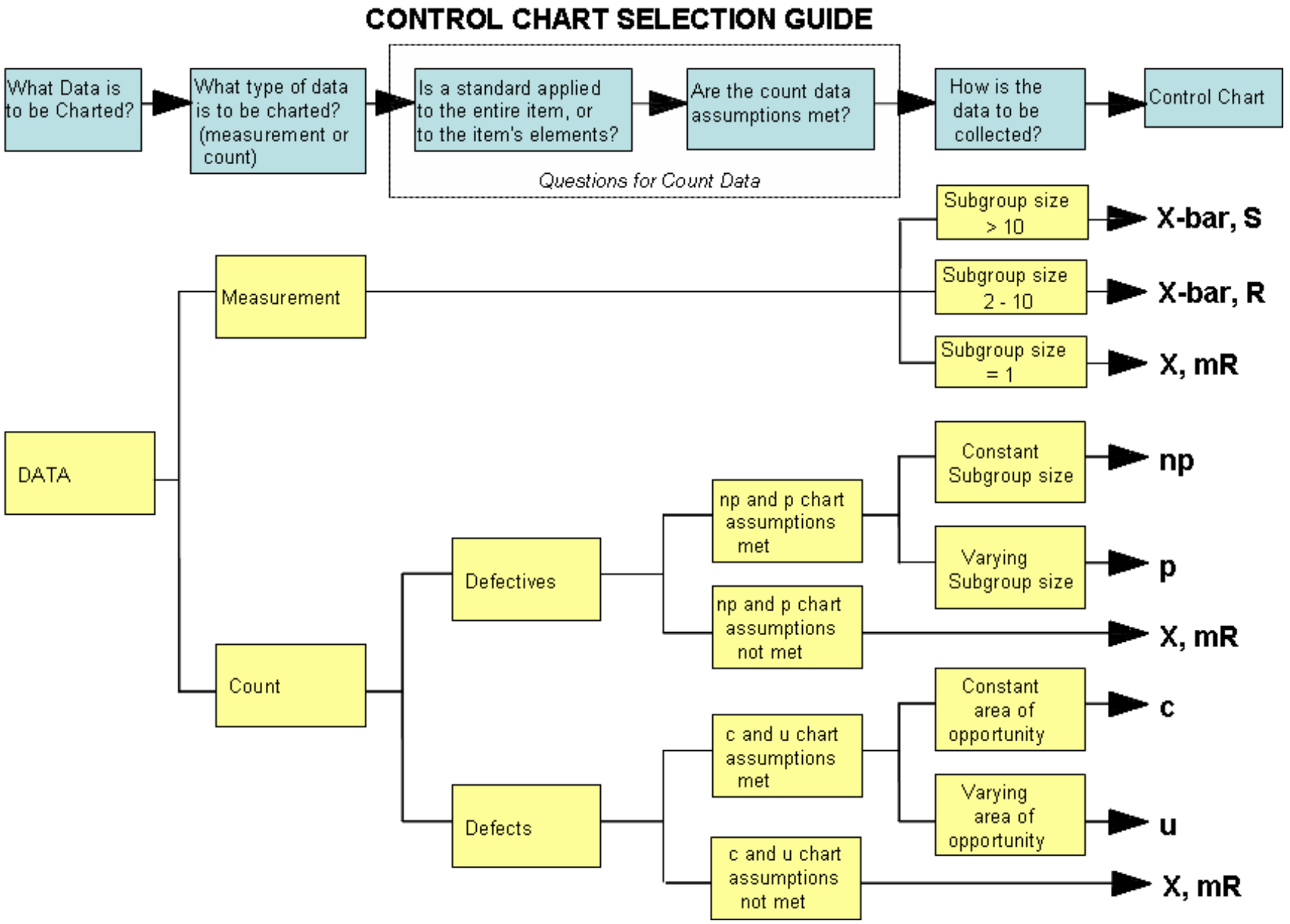
We will present seven different kinds of control charts here that are widely used. To choose the correct control chart for your application, you will have to answer several questions.

Type of Data - Is your data *Measurement* or *Count*? This is the major "divide" for the control charts. Three charts can be applied to measurement data situations: *X-bar, S*, *X-Bar, R* and *X, mR*. Four charts are applied to count data: *np, p, c,* and *u.*

Measurement Data Criteria - If you are dealing with measurement data, we'll ask you whether it makes sense to gather data in a "large" sample (size > 10 - leads to the *X-Bar, S* control chart), take a "small" sample (generally of size 4 - 5 - leads to the *X-Bar, R* control chart), or does it make more sense to collect the data one point at a time (the *X, mR* control chart).

Count Data Criteria - If you are working with count data, you need to think about whether you are dealing with **defectives** (leading to the *np* or *p* control charts) or with **defects** (leading the *c* or *u* control charts). We'll define the difference between defectives and defects later. For count data, the last question will involve the size of the sample you will take to get your data and, most importantly, does the sample size *change* from time to time. The latter criteria will lead us to choose between the members of the count data control chart families.

On the following page is a "cheat sheet" for selecting the best chart. Use it the first few times you need to select a control chart for your application and you'll quickly get the hang of the selection process.


CONTROL CHART SELECTION GUIDE
What Data is to be Charted?
What type of data is to be charted? (measurement or count)
Is a standard applied to the entire item, or to the item's elements?
Are the count data assumptions met?
How is the data to be collected?
Control Chart
Questions for Count Data
DATA
Measurement
Count
Subgroup size > 10
X-bar, S
Subgroup size 2 - 10
X-bar, R
Subgroup size = 1
X, mR
Defectives
Defects
np and p chart assumptions met
np and p chart assumptions not met
Constant Subgroup size
np
Varying Subgroup size
p
X, mR
c and u chart assumptions met
c and u chart assumptions not met
Constant area of opportunity
c
Varying area of opportunity
u
X, mR

Subgroup Strategies

This is such an important control chart concept, and we'll address it right up front. Shewhart coined the term **Rational Subgroup** to describe how to organize the data for a control chart.

The concept is quite easy, but it is a departure for those who are used only to plotting data over time, or in sequence. Let's illustrate Rational Subgrouping with a few examples:

Hypothesis: There is a difference in defects for units that are built on first shift vs. the second shift. Here, our "rationale" is that the first shift employees either are or are not part of a common cause system that "produce" defects. A sample of units produced by the first shift in our study would then be a rational subgroup. Comparisons would then be feasible **between** shifts. Does the proportion or rate of defects *really* differ between shifts?

Hypothesis: There is a difference in Metal Quality (i.e. Hardness, Toughness, or Ductility) among Suppliers to our Company. Here, too, our "rationing" is that the suppliers may not be a common cause system in their sheet metal production. We would obtain samples from each supplier, test them for their metal qualities and then the data would be plotted on the appropriate control chart. Questions regarding the central tendency or variability of the metal qualities could be answered. A side note: we'll also present ANOVA as another way to answer this question later.

Hypothesis: We want to see if Response Times to customer requests differ by type of request. The "rationale" here is that the type of request causes a difference in the response time. We would obtain samples of response times for each request type and plot these on the control chart.

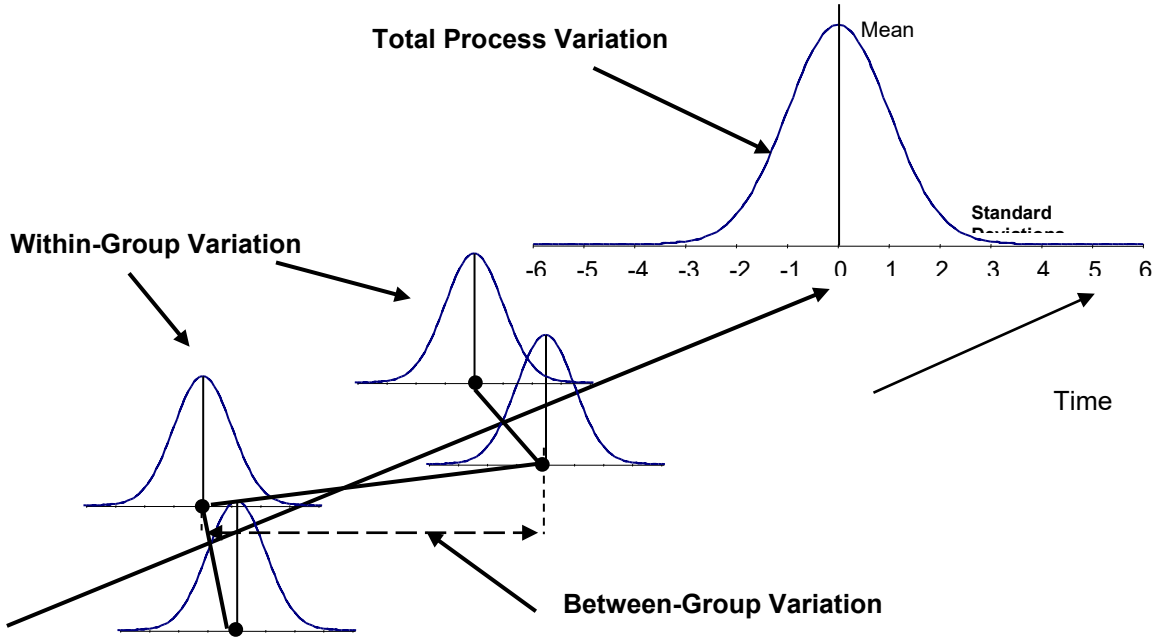
You can see that the idea of Rational Subgrouping is to identify possible factors that may be **assignable causes of variation**. As we organize the data, we'll keep data from these suspected causes in one subgroup and not mix data together from different subgroups.

A Subgroup's Impact on Control Chart Limits - Here's what goes on "inside" the control chart and how subgrouping impacts the "sensitivity" of our control chart.

Let's consider 5 numbers and pretend they are a sample from one of our processes taken, say, early in the day: 10, 24, 9, 18, and 13. The average is 14.8. We'll use the range as our measure of variability; in this case the range is 15 (max. - min. or 24 – 9 = 15). Let's focus on the range for a minute. This value of 15 is an estimate of the variation *within* these

numbers. We give this a specific term: the *within group variation*. There are certain process variables that influence this type of variation. If we took another sample from the process, say, later in the same day, we would get a different set of numbers and could calculate a new average and range. The range gives us a new estimate of the within group variation. But now (you guessed it), we can look at the two averages and think about what they are telling us in terms of the *between group variation.* There are process variables that can affect this kind of variation.

In choosing how we sample from the process (our *subgrouping strategy*), we want to try and get samples that are as homogeneous as possible. Our objective is to try and let the special causes show up as between group variation, not within group variation. In fact, we calculate our control limits with this philosophy. The upper and lower control limits are based on an estimate of the *within-subgroup variation*. The more homogeneous our subgroups are, the smaller this estimate of variation will be. Hence, the tighter our control limits will be. This increases the *sensitivity* of the control chart to detecting special causes!



Practical Example: Let's say we know the output of our process depends on shift. We would *not* want our subgroups (for one data point) to include data from two different shifts. We would take a sample from one shift, plot that point and then take another sample from the second shift and plot that point. This way, most of the variation will show up *between points* as *between group variation.*

This issue of subgrouping is one of the most important for successful control charters. Think about this issue when you are setting up your control chart. It may take some trial and error to identify a good subgrouping scheme for your process.

Several years ago, we decided to plot our gas mileage on a control chart. At that time, we were driving a Mustang GT. We decided to take samples of size 5, which represented about a month's worth of driving (5 fill-ups/month). This scheme worked. We were able to see a special cause in our process - *increased gas mileage traceable to a speeding ticket we had received (since we'd lightened up the ol' lead foot)* - show up as between group variation.

When we bought a more economical car, a Topaz, this scheme did not work as well. The Topaz gas mileage depended on a process variable - city driving vs. highway driving - to which the Mustang was not very sensitive. Since a month's worth of driving included city and highway trips, the sample was not very homogeneous. We wound up keeping two control charts, one for city driving, and one for highway trips.

Here's a summary of sub-grouping principles:

Principle	Comments
Never knowingly subgroup unlike things together	Each subgroup must be logically homogeneous Trim amount – right side, left side (should be separate) Signal – running record that hugs the center line
Minimize the variation within each subgroup	Defines routine background noise Makes chart more sensitive to signals
Maximize the opportunity for variation between the subgroups	If there is a possibility that two things differ, make sure they are in different subgroups Variation will show up in Average Chart (between Subgroups)
Average across noise, not signals	Concentrate the signal and make it easier to detect
Treat the chart in accordance with the data	Subgroup frequency should reflect monitoring and decision frequency appropriate for the process.

Principle	Comments
	Single points vs. multiple points
Establish operational definitions for the sampling procedure.	Eliminate subjectivity in measurement process (operator-to-operator) Values obtained under essentially the same conditions

6.4.3 "Simple" Explanation of Control Charts

The Problem of Variation

Variation exists in all things. We, of course, are concerned about variation in the quality of our products and services. We may want to make our products and services as uniform as possible (*if* this leads to higher quality or lower cost!), but we cannot eliminate all variation. We must accept that quality can only be controlled *within limits.* We also know that there are many *causes* in the "system of production." Methods, materials, equipment, people, measurement, etc. are all factors or variables that influence quality. We also know that there may be "significant" differences between these factors and that it may be profitable to identify and *eliminate* factors which move the average quality away from our target or increase the variability in quality.

The problem we are faced with is *how much variation can be left to chance (or constant) causes (i.e. the "system" of production)*. In a world of variation, how can we distinguish a system of constant causes from one in which assignable causes are also present? The answer, of course, lies in the application of the control chart to our *operation of controlling a process.* So how does the control chart perform its function?

Detecting Lack of Control

Shewhart spent over six years in research before publishing his 1931 book. He wrestled with different statistical measures (means, medians, variances, etc.), with different displays of process data (running records, histograms, control "ellipses"), and with different means of characterizing the "universe," or production process (probability-based, distribution-based, etc.). His work led him to develop five criteria for detecting the presence of assignable causes. For example, Criterion I is the basis for control charts as we know them:

Criterion I - General

Given a set of *n* data to determine whether they arise from a constant cause system, do the following:

1. Divide the *n* data into *m* rational subgroups (of constant or variable size).

2. Pick the *statistics* you will use to judge the data. The mean, standard deviation and proportion defective have been shown to be the most useful statistics for this purpose.

3. For each *statistic*, calculate (using the data) estimates of the *average* and *standard deviation* of the statistic, where these estimates satisfy as nearly as possible the following conditions:

a) If the quality characteristic from which the sample is drawn is *controlled* with average **X-Bar** and standard deviation σ, the estimates used should approach these values as the number of data *n* becomes very large (i.e. in the *statistical* limit),

b) If the quality characteristic is *not* controlled, the estimates used should be those that will be most likely to indicate the presence of trouble (i.e. assignable causes).

4. For each statistic, construct control charts with limits based on the *statistic's* estimated average plus/minus three times the statistic's estimated standard deviation.

5. If a point falls outside the limits of the control chart, take this as evidence of the presence of assignable causes, or lack of control.

Some Comments on The Criteria

Statistics vs. *Parameters* - Shewhart control charts make use of *statistics* calculated from the data. Examples of statistics include the *mean*, *standard deviation* and *proportion defective*. Statistics can be calculated from any set of data. Now you may also be familiar with the mean and standard deviation as *parameters* of a probability distribution such as the normal or Gaussian. If we say we are dealing with *parameters,* we are always assuming some sort of *probability distribution* as the underlying *model* of the data. Shewhart rejected the *parametric/probability distribution* approach to establishing the state of control for three major reasons:

a) The first argument is simply the difficulty of picking the *one unique* distribution that fits our production process data. Attempts by statisticians to claim their favorite distribution "fits" all or even most process situations simply do not work, whether that distribution function is the nice, symmetric *normal distribution,* or the "chameleon-like" general distribution of *Weibull.*

b) The second major argument lies in the fact that production processes are dynamic (they have produced products/services, they are producing, and they will go on producing) but are also finite. Distribution functions are

descriptive of an *infinite* sequence of numbers, not a finite sequence as obtained from a process. Even if a unique function did exist (contrary to a) above), establishing whether a state of control existed would require that a *large* number of data (the observed distribution) be compared to the theoretical distribution (e.g. through a *chi-square* test). This requirement for a *large* number of data would jeopardize the *practical* application of the statistical control process.

c) Third, the distribution approach ignores the *sequence* or *order* of the data. It is easy to show that different sequences of data can result in the same histogram (and hence, same distribution *model*). The functional form of a probability distribution *does not* depend on the order of the data. The distribution approach, then, would likely *not* be capable of detecting non-random patterns that fall *inside* whatever limits were established. In abandoning the *probability distribution* approach, Shewhart tells us that the practical problem faced in production is to establish whether a "universe" exists (i.e. system of constant causes), rather than the actual functional form of some distribution that *might* be used to describe the "universe."

Rational Subgroups - Shewhart often discusses the importance of dividing the data into rational subgroups. He calls for the process analyst to use their imagination, experience, intuition and all other intellectual powers in this subgrouping endeavor. Rational subgroups may be found in the methods, materials, equipment, measurement, people and environmental factors which can influence the process' outcome. Shewhart's goal was improvement of the system of production, and he viewed his control charts as a major tool to achieve this end. Rational subgrouping becomes our way of actively exploring the production process, looking for signals of the presence of assignable causes. Today, unfortunately, many control charts are relegated to the *passive*, "monitoring" role. It's as though Sherlock Holmes had abandoned his *active* detective work, instead deciding to sit out in front of his townhouse, waiting patiently for suspicious people (i.e. "signals") to wander by.

The Choice of "Three Sigma" - This is probably one of the simplest issues surrounding the control chart, but it is also one that has stirred up a lot of confusion and misinformation. The "usual" explanation invokes the normal distribution (and sometimes the *Central Limit Theorem*) and its property that 99.73% of normally distributed data fall within plus/minus three standard deviations of the mean. This explanation is hogwash! Shewhart picked the multiplier of ***3*** simply because it *works*. Over seventy years of experience has shown that when the control limits are set at the statistic's estimated average plus/minus three times the estimate of the standard deviation of the statistic, then the chances of looking for "troubles" when there are none and the chances of overlooking the presence of "troubles" when they are present are minimized. No probability distribution theory needs to be invoked to "justify" the multiplier of 3 (the opening scene of MacBeth comes to mind, replace the witches with statisticians, though!).

Shewhart did make use of probability theory to determine the *distribution* of various statistics to *estimate* the probabilities associated with a symmetric range (i.e. some value *t* times the standard deviation) around the statistic's average. But he always tested his choices against both a *known universe* (i.e. the *normal*) and against the conditions he expected to find in practice: the *unknown universe.* Tchebycheff's Inequality and the "better" Camp-Meidell Inequality provided him with assurance that his choices would stand up to the tests of the "real world." Research and experience have shown that his choices are very robust, that is, the three-*sigma* control limits are not sensitive to the shape of the data (normal or not!).

On Detecting Assignable Causes - Shewhart's rule was simple: if any subgroup statistic fell outside the control limits, then this was evidence enough to search for an assignable cause; the data did not arise from a constant cause system. Since no probability theory was invoked in setting the limits, then no attempt need be made to attach a probability to finding a point outside the limits. Those who invoke the normal distribution as a means of describing the basis of control charts often state that "*there is a 1 - 0.9973 = 0.0027, or about 3 in 1000 chance of a point being outside the limits*" - more hogwash! The "3" is based on *economy*, not *probability!!*

Shewhart went on in his 1939 book to describe the importance of data *order* and its role in detecting assignable causes that appear both outside and within control limits. He does not employ probability arguments to "justify" certain data sequences as having arisen from a non-random (i.e. non-constant system of causes). Rather, Shewhart relies on the "scientist's" ability to detect sequences that "*if observed in the course of actual experimental work . . . would not likewise be called random under normal circumstances."* Further, these observed sequences could only be checked for their non-randomness *"by making further experiments."* That is, by finding and eliminating the assignable causes from the process and observing the results (on the control chart!).

Summary of "Simple" Explanation

Shewhart's control charts (and the *operation of control*) are robust tools to be employed in the improvement and control of production processes. Although it is tempting to justify control charts with assumptions of normal probability distributions, the Central Limit Theorem and probability theory (perhaps because of the perception that invoking these "high-brow" statistics will lend credibility to the charts!), these justifications are not needed and, in general, may do harm to the practice of *economic control of quality*. "Three-sigma" works and has worked for the last 70 years.

6.5 Measurement Control Charts

Learning Objectives

- Construct an *X, Bar, R, X-Bar, S* and *X, mR* Control Charts
- Determine when to apply *X-Bar, R, X-Bar, S* and *X, mR* Control Charts
- Interpret a Control Chart to detect assignable vs. common cause variation

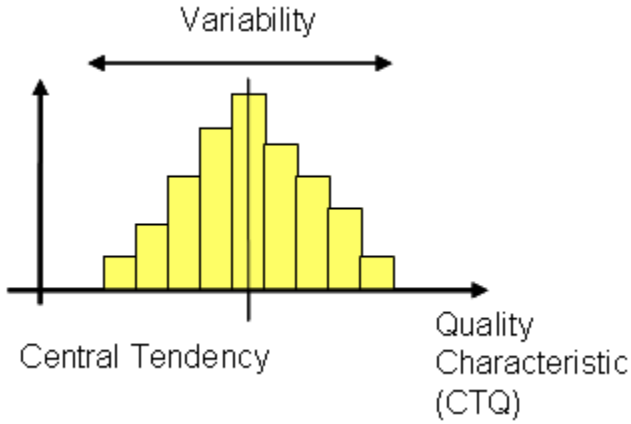
Unit Contents

- X-Bar, R Control Chart
- Interpreting the Control Chart
- X-Bar, S Control Chart
- X, mR Control Chart

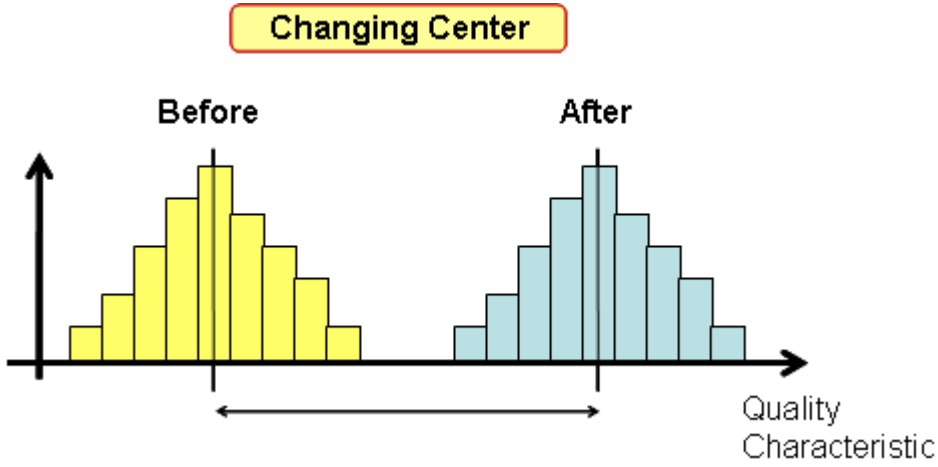
6.5.1 X-Bar, R Control Chart for Small Subgroup Sizes

Purpose

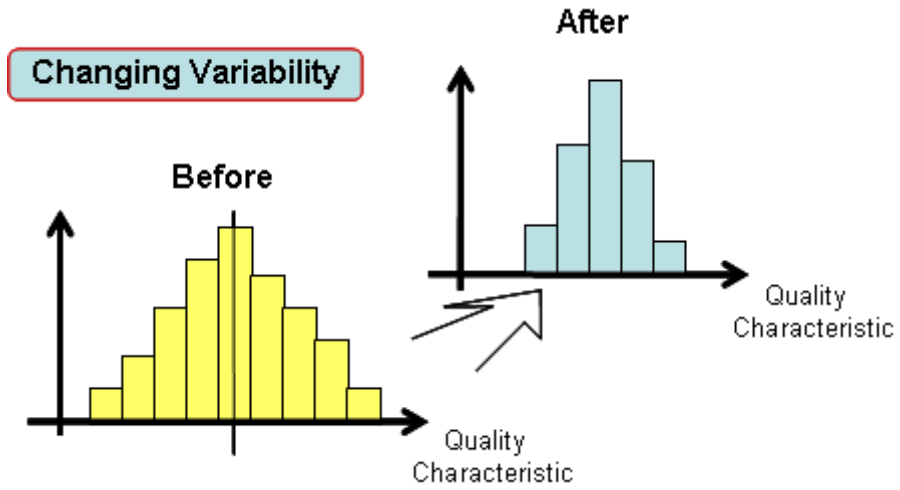
This chart is useful for measurement data, when we are taking small subgroups (generally, < 10) from the process. This chart can track changes in both the central tendency and variability of the process. The X-Bar, R chart is the “classic” manufacturing chart - every hour or so, an operator can measure some quality characteristic (i.e. a critical dimension) of four or five components coming “down-the-line.” These component measurements form a subgroup for the chart. The X-Bar, R chart and its “cousin” the X-Bar, S chart are the most powerful control charts for several reasons. The first reason is that the X-Bar, R (and S) chart consists of two graphs, the "X-Bar" graph tracks how the average (central tendency) of the data is doing, and the "R" (or Range) graph tracks how the variability is behaving.



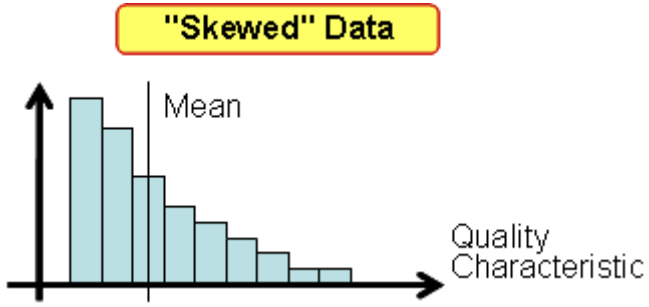
Changes in the process can result in either the center of the data moving or result in the dispersion of the data changing. The X-Bar, R chart can pick up either change. The X-Bar chart will detect changes in the center of the data:



The Range chart will detect changes in the process variability:

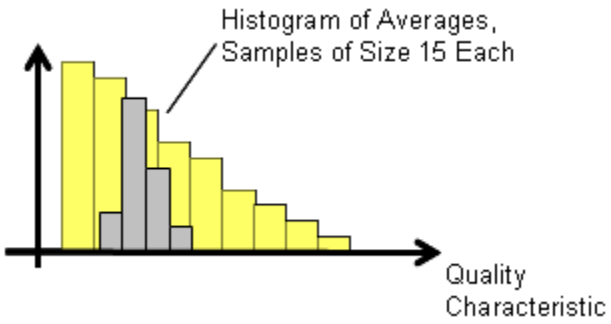


The second reason the X-Bar, R and S charts are powerful is that they track how the *average* of the process output is doing, not just the individual output values. Some process quality characteristics (especially those where we are interested in the time to complete an activity) may have some funny shapes when we plot the individual data values on a histogram:



Note that there are some events that take on high values compared to the mean. If we plotted some individual data from that process, we might be tempted to believe a special cause was present if one of the very high values occurred. Instead, our strategy will be to take several samples of data from this process and calculate the average for each sample.

We can expect that the histogram of these *averages* would have a rather smoother and tighter shape than that presented above:



By plotting the sample averages on a control chart, the chances of us thinking that a special cause is present when there really isn't one are reduced (this is a good thing!). X-Bar, R and X-Bar, S control charts make use of this property. Statisticians refer to the Central Limit Theorem to explain this – a sum of "enough" random variables will assume the shape of a Normal Distribution, regardless of the shape of the individual distributions.

Applications

The X-Bar, R control chart can be used in the following situations:

- *Manufacturing Production Line* - As described above, any high-volume production line is a candidate for monitoring via an X-Bar, R control chart.
- *High Volume Service "Production" Lines* - Similarly, services that are "produced" in high volumes are candidates. Examples include Turn-around Times for Laboratory results and Time to respond to customer requests.
- *Inspections of Received Material* - Each shipment or box of material received from a vendor may be considered a subgroup. Four or five items from each lot may be considered a subgroup.
- *High Volume Instrument Readings* - For continuous processing applications (power plant, chemical or refinery plants), a sample of four or five process variable readings (pressures, temperatures, concentrations, etc.) can be considered a subgroup. One caution here, though. These processes are sometimes auto-correlated - that is, the readings are not independent. There are ways of treating this kind of data that "extract" the auto-correlation factor. Fortunately, control charts are not very sensitive to autocorrelation.

Construction of X-Bar, R Chart

1. Collect the Data - Decide on the subgrouping strategy you will use (shift, day of week, etc.). Decide on the subgroup size you will use. Four or five data per sample is a good general rule, although X-bar, R charts have been developed with anywhere from 2 to 10 data per sample.

> *It's a good idea to use a constant sample size for your first few X-Bar, R control charts. You can develop a control chart with varying subgroup sample sizes, but the control limits must be calculated for each subgroup.*

Data Needed: Try to get at least enough data for 25 points. If your sample size per point is 4, then you will need to measure at least 100 events. Don't get too concerned if you only have 40 or 50 data. Go ahead and develop the X-Bar, R chart, but you should recalculate the control limits as you accumulate more data.

Arrange the data into a table that looks something like that pictured below. In addition to the table, get a piece of graph paper ready so you can plot the data.

Subgroup	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**.**	**.**	**.**	**.**	**.**	**.**	**25**
Data 1																
Data 2																
Data 3																
Data 4																
Subgroup Average																
Subgroup Range																

R Chart

We'll work on the R chart first, since several values calculated here are needed for the X-Bar chart.

1. Calculate the ranges for each subgroup. Record these ranges on the last row of the table.

$$R_j = X_{j-\max} - X_{j-\min}$$

where:

$X_{j-\max}$ - the largest value of the "jth" subgroup

$X_{j-\min}$ - the smallest value of the "jth" subgroup

R_j - "jth" subgroup Range

2. Calculate the average Range. Add up all the subgroup ranges and divide by the total number of subgroups you have.

$$\overline{R} = \frac{1}{k}\sum_{j=1}^{k} R_j$$

where:

R_j - "jth" Subgroup Range

k - number of Subgroups

$\overline{R}$ - Average Range

3. Now calculate the Upper and Lower Control Limits for the Range chart:

$$UCL_R = \overline{R} \times D_4$$

$$LCL_R = \overline{R} \times D_3$$

where:

D_4, D_3 - Coefficients

UCL_R - Upper Control Limit for Range Chart

LCL_R - Lower Control Limit for Range Chart

The values of the coefficients D_4 and D_3 depend on the size of the sample you are using for each subgroup. For a sample of size 4, D_3 is not applicable (this means that there is no Lower Control Limit) and D_4 is 2.282. Since there are several more coefficients, we have provided a summary table at the end of this procedure.

4. Draw the average Range as a **solid** line on your graph. Draw the Upper and Lower Control Limits as *dashed* lines on the graph. Plot the subgroup Ranges.

5. Check the Range chart for special causes (see below for control chart interpretation). If you see special causes here, it means that some process variable is causing the *dispersion* of your process to change over time.

> *NOTE: It's worthwhile to stop here and investigate this type of process variation. In the "old days" (pre-computers), statisticians would recommend that you not even bother proceeding with the X-Bar chart until your process' variation was under control.*

X-Bar Chart

6. Now, calculate the average of each subgroup and record it on the second to last row.

$$\bar{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{ij}$$

where:

n_j - "jth" subgroup size

x_{ij} - "ith" element of the "jth" subgroup

$\bar{x}_j$ - "jth" subgroup average

7. Now calculate the *grand average*. This is just the average of all the subgroup averages.

$$\bar{\bar{x}} = \frac{1}{k}\sum_{j=1}^{k} \bar{x}_j$$

where:

$\bar{\bar{x}}$ - Grand Average of Subgroups

8. Calculate the Upper and Lower Control Limits for X-Bar:

$$UCL_{X-Bar} = \overline{\overline{X}} + (\overline{R} \times A_2)$$

$$LCL_{X-Bar} = \overline{\overline{X}} - (\overline{R} \times A_2)$$

where:

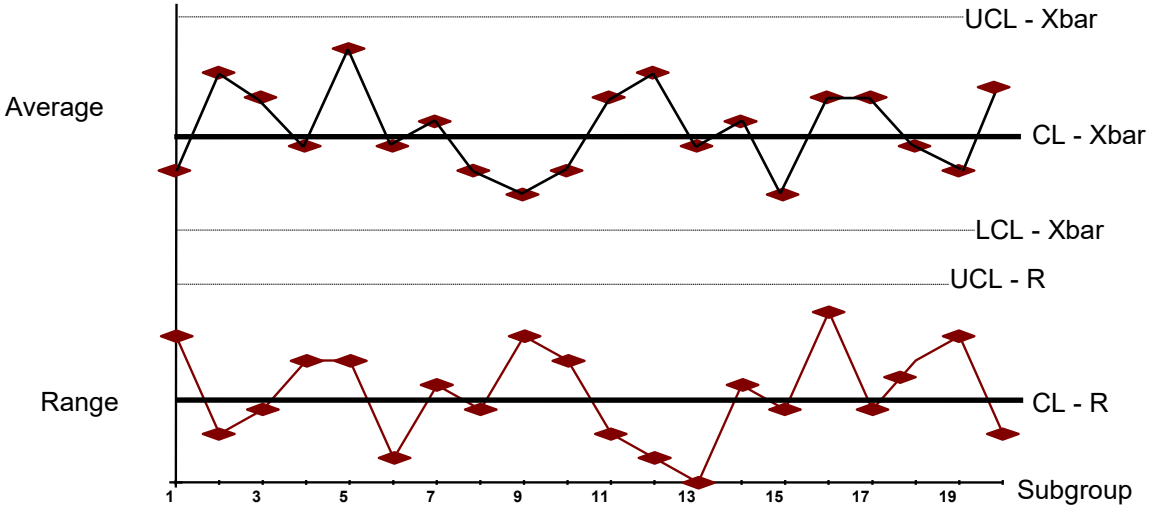
A_2 - Coefficient

UCL_{X-Bar} - Upper Control Limit for X - Bar

LCL_{X-Bar} - Lower Control Limit for X - Bar

Again, the coefficient A_2 will vary depending on your sample size. For a subgroup size of four, A_2 is 0.729.

9. Plot this grand average as a *solid* line on the X-Bar part of the chart. Draw the control limits as *dashed* lines on your graph. Plot the subgroup averages on your graph. Your graph should look like this:



10. Interpret the charts for assignable causes of variation. Use the rules presented at the end of this unit.

COEFFICIENTS FOR X-Bar, R CONTROL LIMITS

Sample Size (1)	A_2	D_3 (2)	D_4	d_2
2	1.880	-	3.268	1.128
3	1.023	-	2.574	1.693
4	0.729	-	2.282	2.059
5	0.577	-	2.114	2.326
6	0.483	-	2.004	2.534
7	0.419	0.076	1.924	2.704
8	0.373	0.136	1.864	2.847
9	0.337	0.184	1.816	2.970
10	0.308	0.223	1.777	3.078

Notes:

(1) This is the number of data points that are combined into one subgroup.

(2) For sample sizes 2 through 6, the Lower Control Limit of the Range is not applicable.

Once you have decided on your sample size, just use the coefficients from the corresponding row of this table. You might be wondering what purpose that last column serves. Well, remember that the Range is the value we are using to measure the dispersion of the process. There are other measures of dispersion, including variance and standard deviation. The d_2 coefficient can be used to give us a good estimate of the process' standard deviation:

$$\text{Standard Deviation} = \text{Average Range} / d_2$$

This is a useful conversion to have when we start to consider the capability of a process to meet customer specifications. We'll mention it here and return to this concept later.

X-Bar, R Control Chart Example

Scenario - A manufacturer of pressure-treated lumber tests for the chemical concentration (in percent) of the treatment. A sample of 4 pieces is taken from each batch after the treatment process and the concentration obtained. Data from the last few batches appears below:

	Batch					
Data	**A**	**B**	**C**	**D**	**E**	**F**
1	3.42	3.34	3.41	3.25	3.40	3.25
2	3.61	3.30	3.33	3.30	3.35	3.35
3	3.22	3.26	3.28	3.28	3.37	3.34
4	3.38	3.32	3.35	3.27	3.30	3.28

The basic calculations are straightforward. The subgroup Ranges are easily found:

Subgroup A - R = 3.61 - 3.22 = 0.39
Subgroup B - R = 3.34 - 3.26 = 0.08
|
Subgroup F - R = 3.35 - 3.25 = 0.10

The Average Range is then equal to 0.14. With this information, we can calculate the Upper Control Limit (for a subgroup of size 4, there is no Lower Control Limit):

$UCL_R = 0.14 \times 2.282 = 0.32$

We see that Subgroup A's Range is out of control and should investigate this instability in the *variability* of the treatment process. The Subgroup Averages are then calculated; we'll use Subgroup A as an example:

$$\bar{x}_A = \frac{3.42 + 3.61 + 3.22 + 3.38}{4} = \frac{13.63}{4} = 3.41$$

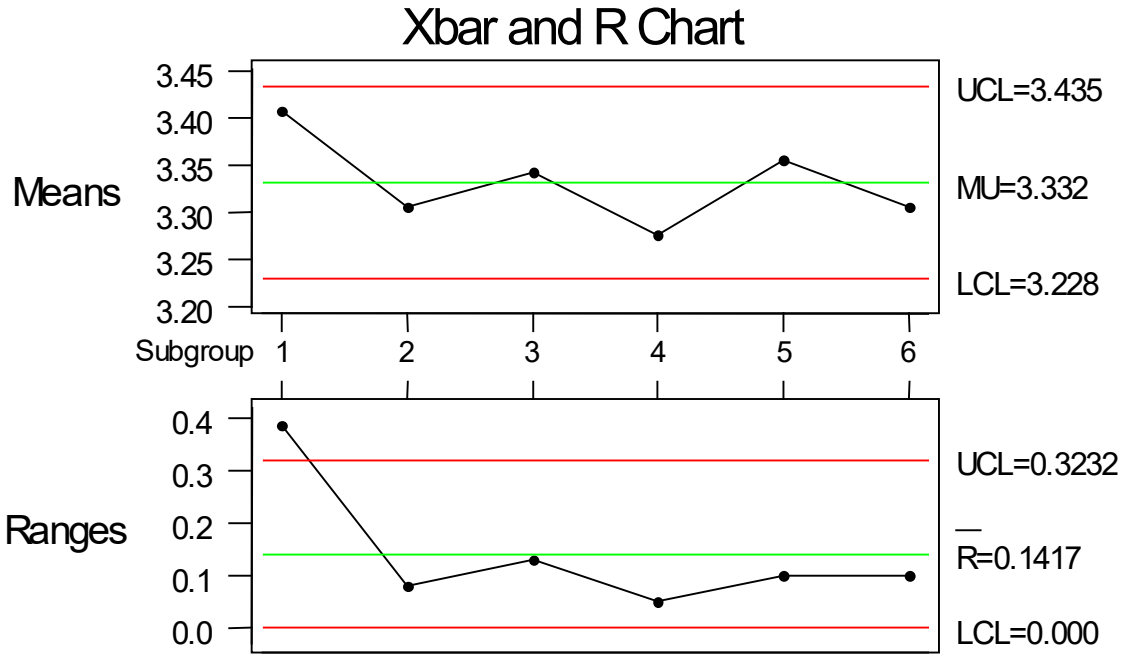
The remaining calculations are provided below:

Subgroup	A	B	C	D	E	F
Average	3.41	3.31	3.34	3.28	3.36	3.31
Range	0.39	0.08	0.13	0.05	0.10	0.10

Grand Average:	3.34
Upper Control Limit - Average:	3.44
Lower Control Limit - Average:	3.24

None of the Averages are out of control, but with the Range out of control, we can't be sure what's happening with the averages. You should investigate the Range Chart first!

Note: The Graph below was developed using Minitab statistical software. Some of the calculations may differ slightly from those in the text due to rounding.



6.5.2 Interpreting the Control Chart

To determine the state of our process, we will look for special patterns that appear in the data, just as we did when we used run charts. When using the measurement data charts (X-Bar, S, X-Bar, R or X, mR Chart), the rule is to first look at the chart that displays the process' **variability** (Standard Deviation or Range) for special causes of variation and work on eliminating these. Then examine the X-Bar or X charts for special causes and work on eliminating them (or stratifying them, see two paragraphs down).

If the variability of the process is being influenced by special causes of variation, then the variation we see in the averages of our samples is affected by both the *within-subgroup* and *between-subgroup* variation. Trying to determine what is going on in our process with both these kinds of variation present can be very difficult.

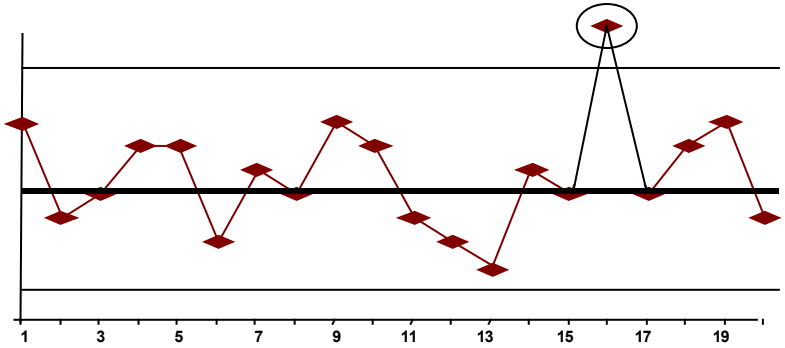
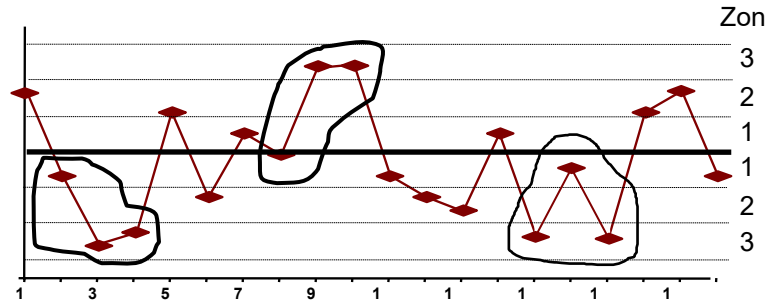
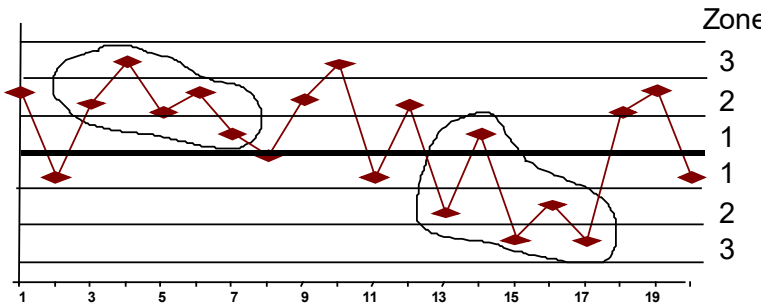
We said that we want our special causes to show up as *between group* variation. Well, with this philosophy, we want to try and get rid of one kind of variation first; the within-group type is the best place to start. Additionally, since our control limits are based on the within group variation, we want to make this as small as possible to improve our chances of detecting actual "out-of-control" conditions affecting the subgroup averages.

Note that sometimes "special cause" indications may arise when we combine the output of two processes on one graph or, equivalently, the output from one process that has multiple paths. The processes should be stratified in these cases and control charts prepared for the individual processes.

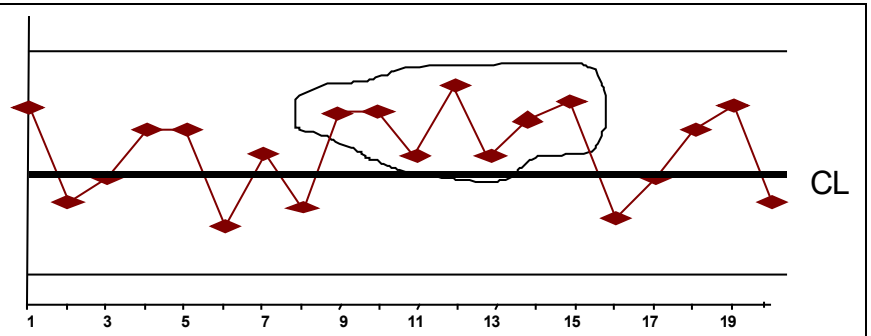
As an example, if we measure the time it takes to repair equipment, the type of equipment (pumps, valves, heat exchangers, circuit breakers, etc.) may be a significant process variable. Preparing a control chart that mixed these types of equipment together may be a mistake.

How do we detect special causes of variation using our new control charts? Here are eight rules that indicate the presence of special causes (three of which can be used with run charts).

The first four rules indicate assignable causes that range from Dominant to Small-but-Sustained process changes. Some of the rules require that we divide the distance between the chart's centerline and the control limits into Zones, each zone being one third of the distance from the centerline to the control limit (each zone, then, is one SIGMA!).

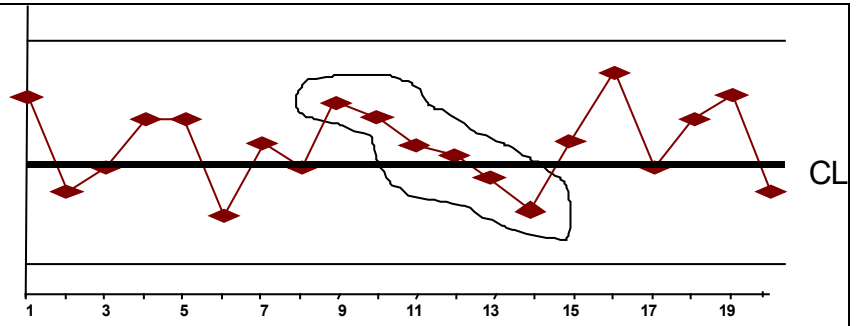
Rule 1 - Any point that is outside the control limits on either side. This is Shewhart's "classic" rule and may be applied to any control chart, whether it measures the central tendency or variability of the process. Here, a **DOMINANT** Assignable Cause has occurred.	
Rule 2 - Whenever two out of three successive values are both on the same side of the central line and are both more than two sigma units away from the central line, look for the cause of an **INTERMEDIATE** sized shift in the underlying process.	
Rule 3 - Whenever four out of five successive values are on the same side of the center line and are also more than one sigma unit away from the center line, look for the cause of a **MODERATE** shift in the underlying process.	

Rule 4 - Whenever seven successive values fall on the same side of the central line, you should look for the cause of a **SMALL BUT SUSTAINED** shift in the underlying process. This indicates that the process average has shifted or, (if the signal appears on the R or s-chart), that the process variability has shifted. Other runs that are interpreted as special cause signals include at least 10 of 11, 12 of 14 and 16 of 20 consecutive points on one side of the centerline.

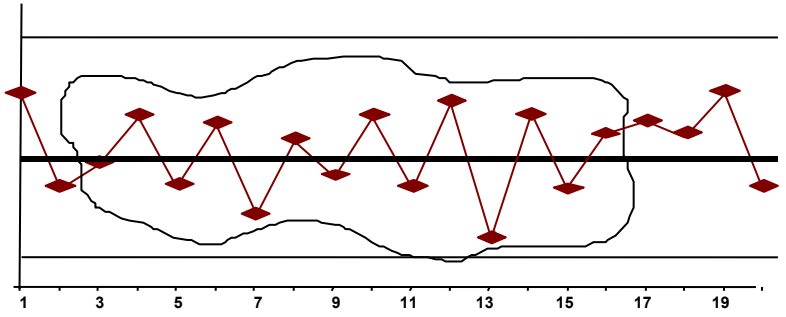


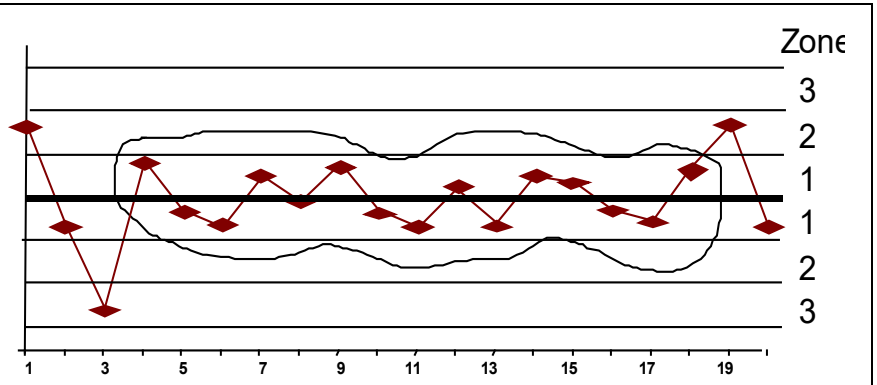
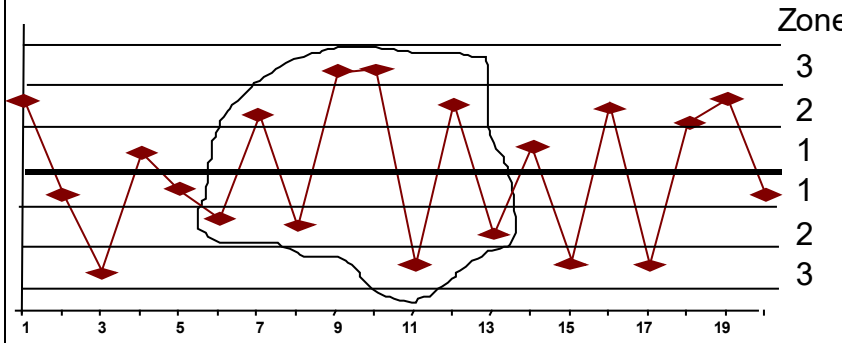
These remaining rules can also be applied but be careful of using too many rules. You'll find yourself chasing too many signals, many of which will turn out to be false alarms. Your Statistical Process Control effort will lose credibility!

Rule 5 - A trend of at least 6 consecutive points either increasing or decreasing. A drastic trend downward (even though not all points are consecutively decreasing) is also evidence of a special cause.

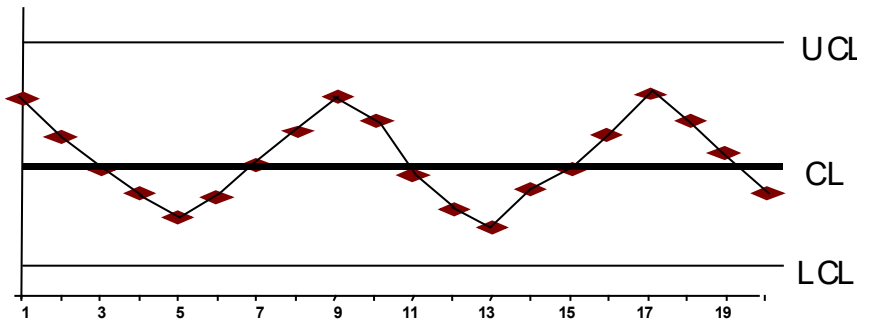


Rule 6 - Fourteen points in a row, alternating up and down. This often indicates two processes where the output from the first process is alternated with the second, back to the first, etc. This signal is often called the "hour-glass" effect - measuring the time for the sand to flow alternately from one glass to the other will produce this effect.



Rule 7 - Fifteen points in a row in Zone 1, above and below the centerline. This is known, affectionately in control chart circles as "hugging" (some refer to it as "stratification). It may seem at first that this is a good situation. This signal most often tells us that we have mixed two processes together, and that each process is "equally represented" in each subgroup (for simplicity, suppose a subgroup of size 2, where each data comes from one of the different processes). Since the control limits are based on the average range, the limits will be wide, and the data will appear to "hug" the centerline. Bottom Line: look for a mixture of two processes contributing to the data for this chart.	
Rule 8 - Eight points in a row on both sides of the centerline, with none in Zone 1. This signal is the same as rule 7; we've mixed processes together. This difference in the signal, though, comes from the fact that each subgroup consists of data entirely from one process or the other - no mixing of data within the subgroups has occurred. Again, look for two process streams being captured on this chart.	

This last rule is somewhat of a "catchall." There may be non-random patterns that exist in your data that do not generate the signals described above. Be careful, though, of trying to detect a signal that is not there. If you look at a data set on a chart long enough, you **will** detect something, even if the data is random.

Rule 9 - A repeating pattern over time is known as Cycling. Look for seasonal influences on the process. This often appears when weekday and weekend data is mixed (a five and two pattern), or when shift-to-shift data is mixed (a three pattern).	

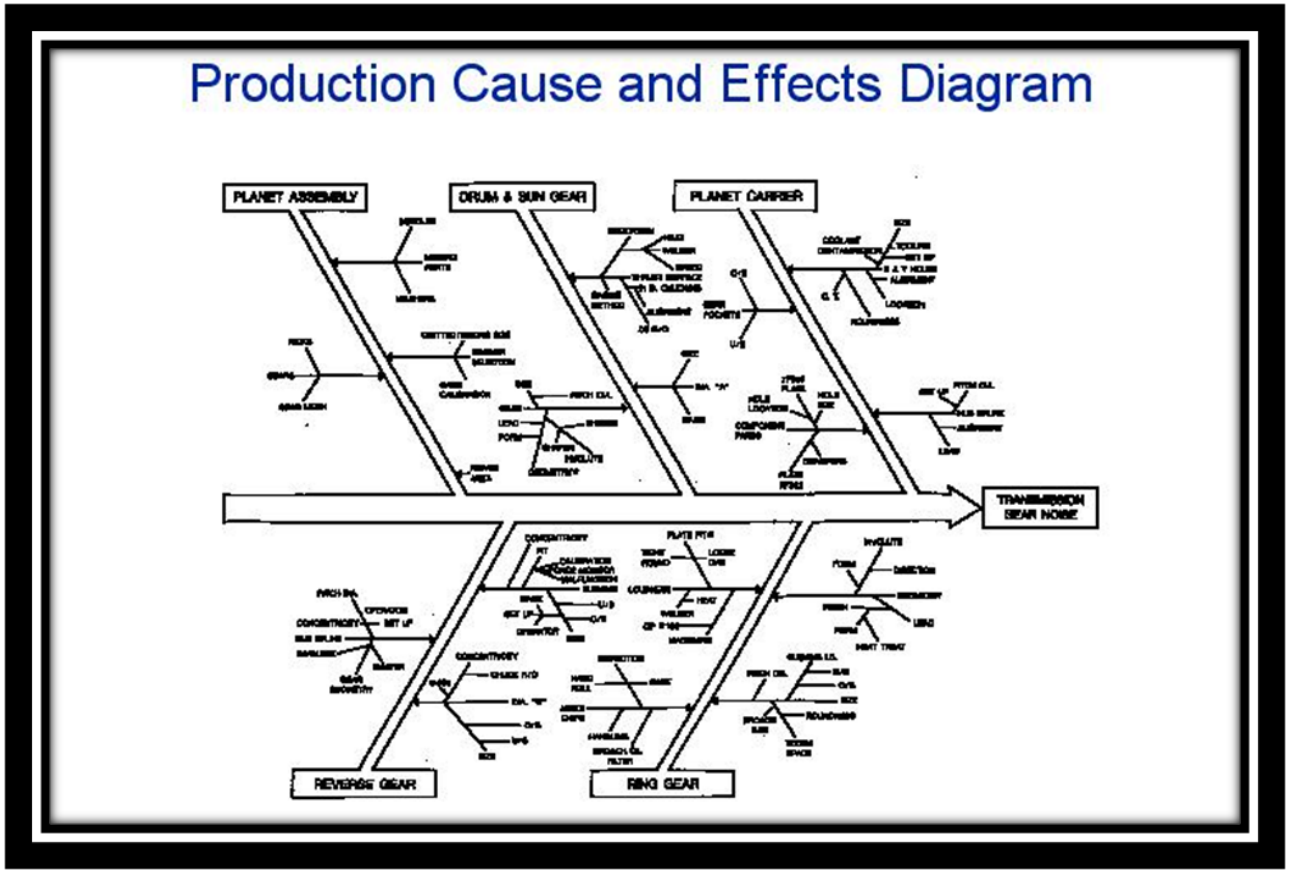
False Alarm Rate

You should expect to see some false alarms using control charts. Shewhart's argument is to *economically* control quality. We'll occasionally miss some assignable causes and raise the flag when there isn't an assignable cause.

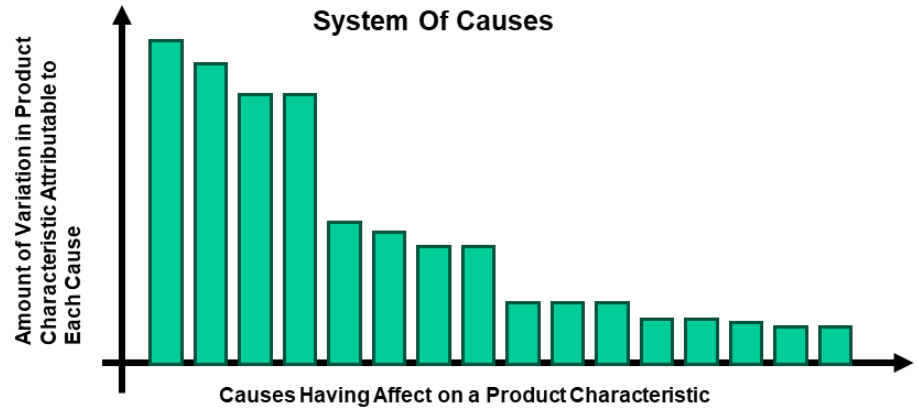
- **Rule One**: will yield about one false alarm every 370 subgroups
- **Rules One and Four**: will yield about one false alarm in every 153 subgroups
- **Rules One, Two, Three and Four**: will yield about one false alarm every 92 subgroups

The Nature of Assignable Causes

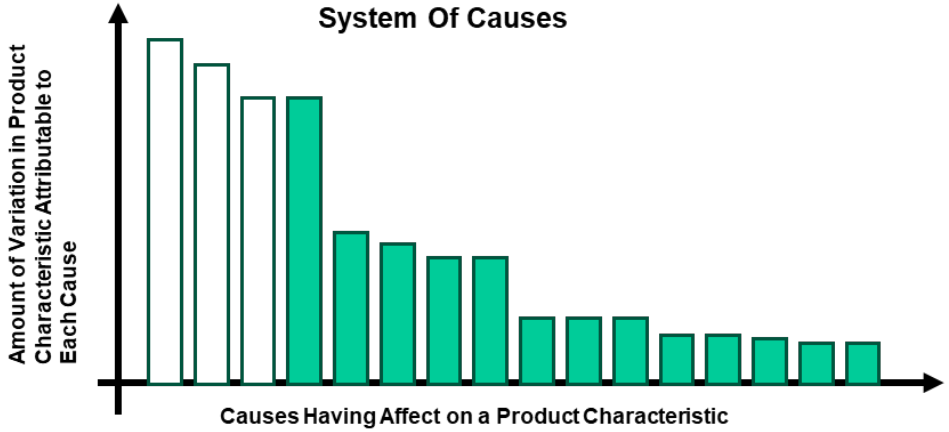
When we chart data from a process, there may be hundreds of causes that contribute to the observed variation. But the Pareto Principle (later!) tells us that most of the variation comes from a small subset of causes. The question is – *which ones have the greatest effect?*



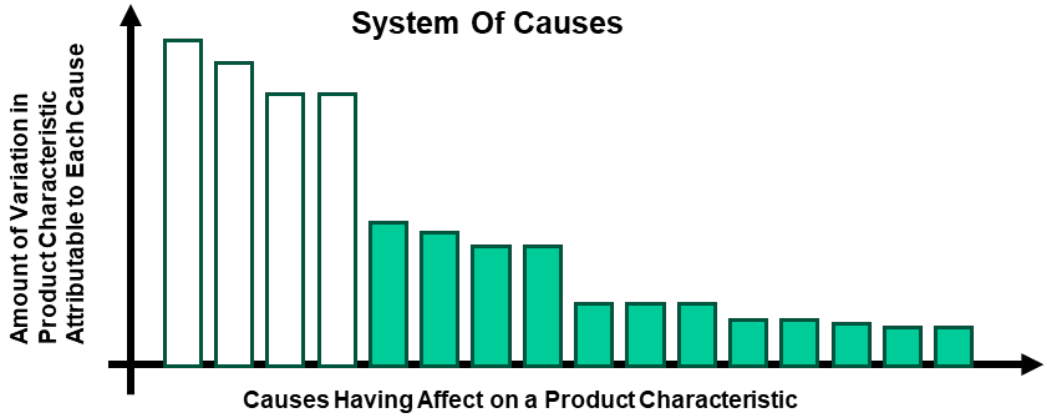
Earlier, we discussed establishing *process control systems*. We've made decisions about which factors/causes to control. Let's assume our initial system of causes looks like this:



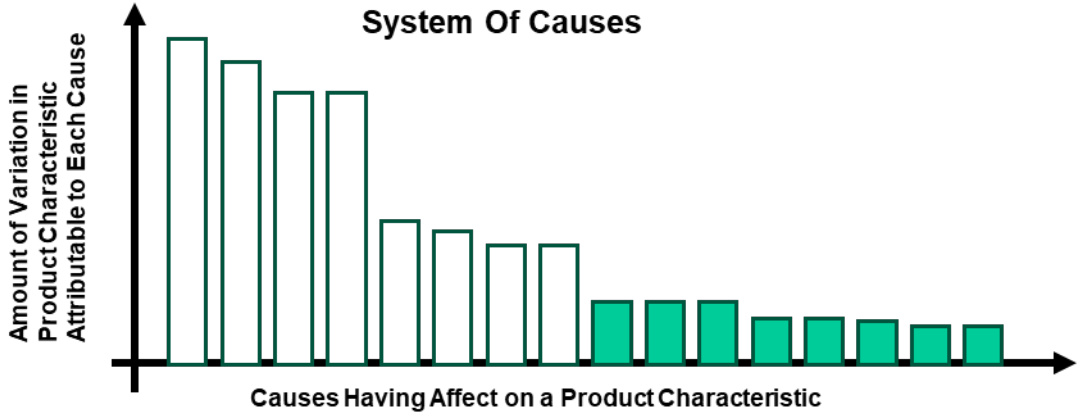
We've decided to control three causes that we believe have a large effect on the quality characteristic. Since they are controlled, their influence on process variation has been reduced:



But there is one dominant cause that is not controlled. This will eventually appear as an assignable cause (most likely a violation of Rule 1 – a point outside the limits). If we identify that cause and bring it under control, our system of causes now looks like this:



There are still some uncontrolled causes (INTERMEDIATE, MODERATE), and they will show up as Rules 2, 3 or 4 signals. If we identify these and bring them into the control system, eventually our cause system will look like this:



Finally, we are left with uncontrolled causes that are either too small to result in signals on our control chart or may show up as SMALL BUT SUSTAINED assignable causes. If we address these, our process will then be “as good as it gets.” Further improvement will require process re-design.

6.5.3 X-Bar, S Control Chart – For Large, Variable or Constant Subgroup Sizes

Purpose of the X-Bar, S Control Chart

The X-Bar, S chart is like the X-Bar, R control chart. This chart handles situations where either (or both) the subgroup size is large (> 10) or the subgroup size varies from group to group. This chart is applied to *measurement* data. Where the X-Bar, R chart is useful for *controlling* an ongoing production process, the X-Bar, S chart has added value when we are trying to *analyze* a production process for presence of assignable causes.

Applications

Typical applications of the X-bar, S Control Chart include:

- *Cycle Time* - Here, the subgroup size could vary if the subgroup was time-based (i.e. data plotted monthly) or factor-based (i.e. data stratified by shift, unit type).
- *Customer Survey Data* - Here, the subgroup size is generally both large (>10 responses) and varies (# responses per month or by unit).
- *Lot Inspection* - Where a "large" sample from a lot provided by a vendor (or an upstream process) is gathered and a measurement is made on each sample item.
- *High, Variable Volume Processes* - Any process that produces a relatively high volume of output which varies (i.e. by time period) such as Deficiency Reports, Order processing, Educational Course Feedback, Repair Procedures, etc.

Construction Steps

1. Collect the data in rational subgroups. If the chart is being used to *control* a process, the subgroups will likely be some unit of time (day, month, etc.), but if the chart is being used to *analyze* a process, process variables may help define the subgroups (i.e. workers, units, shift, machines, etc.).

Data Needed: If the chart is being used to **control** a process, try to get at least enough data for 25 points. If your subgroup size per point is 20, then you will need to measure at least 500 events. If you don't have enough, go ahead and develop the X-Bar, S chart, but you should recalculate the control limits as you accumulate more data. If the chart is being used in **analysis** to detect assignable causes, the number of subgroups may often be less than 25 (i.e. 7 or 8

machines who fabricate a certain part, 4 or 5 suppliers of a product or service). Here, try to accumulate at least 10 data per subgroup.

Arrange the data into a table that looks something like that pictured below.

Subgroup	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**.**	**.**	**.**	**.**	**.**	**.**	**.**	**25**
Data 1																	
Data 2																	
Data 3																	
Data 4																	
. . . . .																	
Data “n”																	
Subgroup Average																	
Subgroup Std. Dev.																	

SIGMA (S) Chart

2. We’ll start with the SIGMA or Standard Deviation part of the control chart. The data we calculate here will be needed to develop the X-Bar part of the chart. First, calculate the Average of each Subgroup:

$$\bar{x}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} x_{ij}$$

where:

n_j - "jth" subgroup size

x_{ij} - "ith" element of the "jth" subgroup

$\bar{x}_j$ - "jth" subgroup average

3. Calculate the Variance and Standard Deviation of each Subgroup:

$$s_j^2 = \frac{1}{n_j - 1}\sum_{i=1}^{n_j}(x_{ij} - \bar{x}_j)^2 \text{ and } s_j = \sqrt{s_j^2}$$

where:

n_j - *size of " jth" subgroup*

x_{ij} - *"ith" element of " jth" subgroup*

$\bar{x}_j$ - *" jth" subgroup average*

s_j^2 - *" jth" subgroup Variance*

s_j - *" jth" subgroup Standard Deviation*

4. Now, calculate the Average Standard Deviation for all subgroups:

- *For large (n > 25) and/or variable subgroup sizes:*

$$\bar{s} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + (n_3 - 1)s_3^2 + \ldots + (n_k - 1)s_k^2}{(n_1 + n_2 + n_3 \ldots + n_k) - k}}$$

where :

$\bar{s}$ - Average Standard Deviation

k - number of subgroups

- *For small (n < 25) and constant subgroup sizes:*

$$\bar{s} = \frac{1}{k}\sum_{j=1}^{k} s_j$$

where :

k - number of subgroups

5. Calculate the Average Subgroup Size (only if the subgroup sizes vary):

$$\overline{n} = \frac{1}{k}\sum_{j=1}^{k} n_j$$

where:

n_j - Size of " jth" Subgroup

k - Number of Subgroups

$\overline{n}$ - Average Subgroup Size

6. Calculate the Upper and Lower Control Limits (UCL, LCL) for the Standard Deviation:

$$UCL_s = \overline{s} + 3\overline{s} / \sqrt{2\overline{n}}$$
$$LCL_s = \overline{s} - 3\overline{s} / \sqrt{2\overline{n}}$$

Note: If the subgroup sizes vary significantly (more than +/-25% of the average subgroup size), use the individual subgroup sizes (n_j) in place of the average subgroup size and calculate individual control limits for each subgroup.

7. Plot the subgroup standard deviations, the average standard deviation (as a solid line), and the upper and lower control limits (as dashed lines). If the calculated Lower Control Limit (LCL) is negative, then there is no LCL for the standard deviation.

X-BAR Chart

8. Calculate the Grand Average of the Subgroups:

$$\overline{\overline{x}} = \frac{1}{N}\sum_{all\,i,j} x_{ij}$$

where:

x_{ij} - Individual Data from all Subgroups

N - Total number of data from all Subgroups

$\overline{\overline{x}}$ - Grand Average of Subgroups

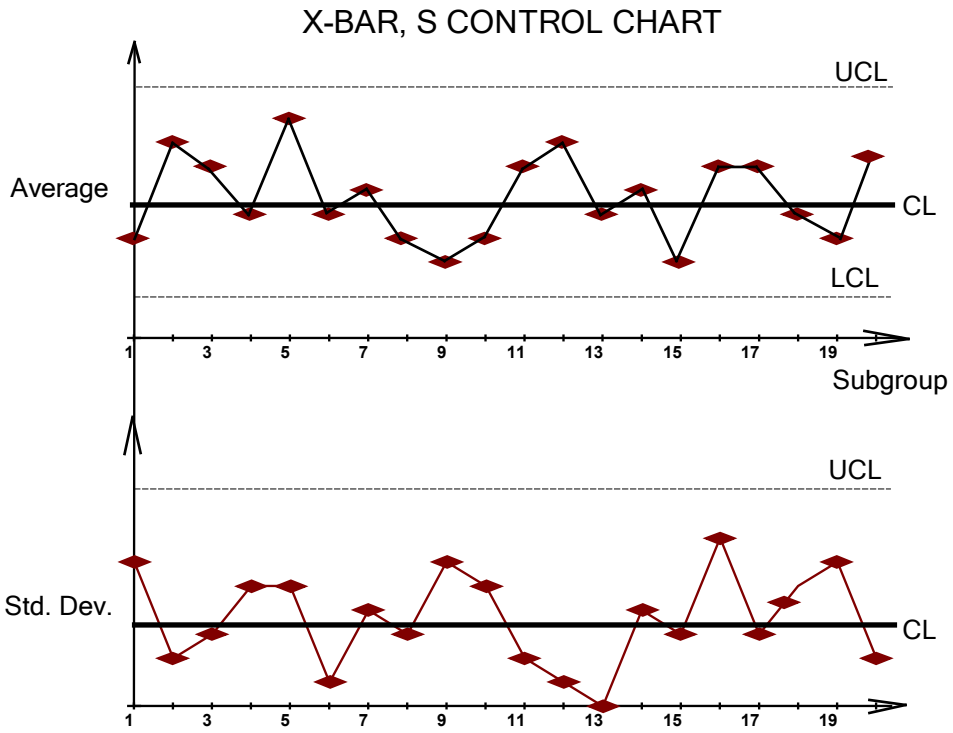
*Note: The Grand Average may be found by taking the average of the subgroup averages, **only** if the subgroup sizes are equal (or very nearly so).*

9. Calculate the Upper and Lower Control Limits (UCL, LCL) for the Averages:

$$UCL_{\bar{x}} = \bar{\bar{x}} + 3\bar{s} / \sqrt{\bar{n}}$$
$$LCL_{\bar{x}} = \bar{\bar{x}} - 3\bar{s} / \sqrt{\bar{n}}$$

Note: If the subgroup sizes vary significantly (more than +/-25% of the average subgroup size), use the individual subgroup sizes (n_j) in place of the average subgroup size in the equation. This requires calculating individual control limits for each subgroup.

10. Plot the subgroup averages, grand average (as a solid line), and the upper and lower control limits (as dashed lines). Your chart should look like this:



11. Interpret the chart using the special cause rules discussed above. Take action to eliminate assignable causes of variability or improve the performance of the process (reduce variability or change the central tendency).

Example X-Bar, S Control Chart

Scenario - For a maintenance supervisor, minimizing the time production waits before equipment is restored to service is very important. The supervisor thinks that the day of week makes a difference in how long they must wait, so he collects the following restoration times (in minutes):

MON	TUE	WED	THU	FRI
10	35	5	25	10
25	45	10	35	40
15	40	0[1]	15	20
45	30	10	0	30
20	50	15	20	25
5	30	20	35	15
60		10	15	20
40		15	15	35
15			10	20
20				0
				10

Note: 1. A “zero” time means that it took less than a minute to restore the equipment.

Let’s make sure we know how to calculate the building blocks of the X-Bar, S Chart:

Subgroup Averages - For Monday, the Subgroup Average is:

$$\bar{x}_{mon} = \frac{10+25+15+45+20+5+60+40+15+20}{10} = 255/10 = 25.5\,\text{min}.$$

Subgroup Standard Deviations - Again, using Monday’s restoration times, the subgroup standard deviation calculation would be:

$$s^2_{Mon} = \frac{(10-25.5)^2 + (25-25.5)^2 + (15-25.5)^2 + \ldots + (20-25.5)^2}{10-1}$$

$$= \frac{240.25 + 0.25 + 110.25 + \ldots + 30.25}{9} = 302.5$$

and

$$s_{Mon} = \sqrt{302.5} = 17.39 \min.$$

The only other calculation that is a bit tricky for this chart is the average Standard Deviation. Proceeding through the rest of the subgroups, we would develop a table (or spreadsheet) like the one below:

Subgroup	MON	TUE	WED	THU	FRI
Avg.	25.5	38.33	10.63	18.89	20.45
Std. Dev.	17.39	8.16	6.23	11.40	11.72
Std. Dev. Squared	302.41	66.59	38.81	129.96	137.36
n	10	6	8	9	11

Since the subgroup sizes are nearly constant (thumbrule: if the largest subgroup is less than twice the size of the smallest, the subgroups can be considered "constant."), we can employ the simple average Standard Deviation:

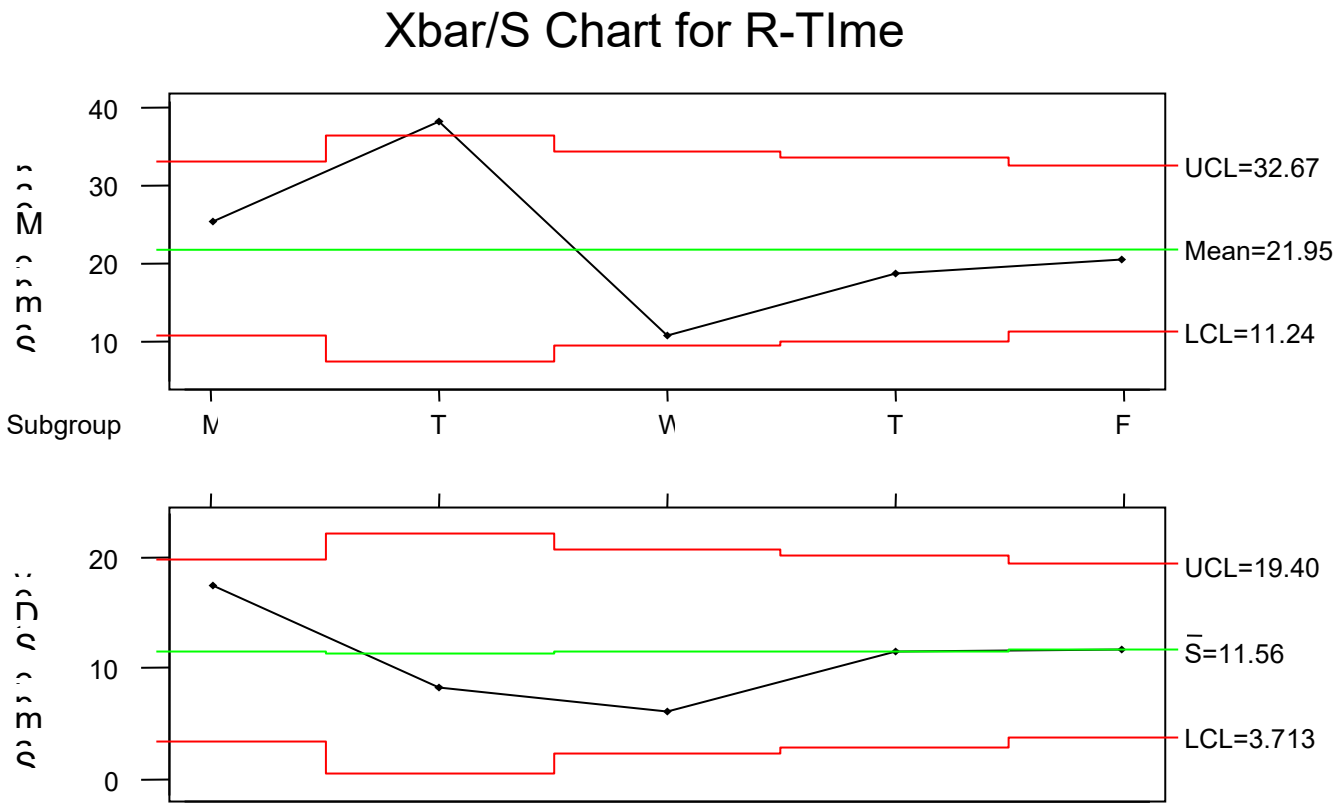
$$\bar{s} = \frac{17.39 + 8.16 + 6.23 + 11.40 + 11.72}{5} = \frac{54.90}{5} = 10.98 \min.$$

In case you'd like to check your math, here are the rest of the calcs for this example:

Average Subgroup Size:	8.80
Upper Control Limit - Std. Dev.:	18.83
Lower Control Limit - Std. Dev.:	3.13
Grand Average:	21.93
Upper Control Limit - X-Bar:	33.03
Lower Control Limit - X-Bar:	10.83

Notice that Tuesdays' Average Restoration Time (38.33 minutes) is Out-of-Control. When we use the +/- 25% thumb rule, Wednesday's Average (10.63 minutes) is also outside the limits - our supervisor should investigate why these days are different from the other days of the week.

However, the largest and smallest subgroups (6 and 11) differ by 32% from the average subgroup size, slightly larger than our 25% suggestion. When the data is entered into Minitab, it automatically calculates the varying control limits. Here, Tuesday is still out-of-control, but Wednesday "sneaks in" and would not be considered an assignable cause:



6.5.4 X, mR Control Chart for Individual Data

Purpose

The X-Bar, S and X-Bar, R charts are the most challenging to calculate (by hand, you will likely be using software) and there were a lot of issues to get on the table. From here on out, we'll talk mainly about the mechanics of constructing the charts. The issues of subgrouping and control chart interpretation are pretty much the same for all the remaining charts.

The X, mR (mR stands for *moving Range*, some books call this the sequential range) chart is like the other two measurement charts, except our subgroup size is going to be one (1).

The X, mR chart is useful when our process does not produce a large volume of data, maybe only one point a week or a month. The X, mR chart is composed of two graphs, the X and the mR charts. Sometimes, people refer to this chart as the *Individuals* chart.

The immediate question we must address is how to get a measure of within-group variation when our sample size is only one. We wind up "creating" subgroups by treating sequential "X" values as a subgroup. The X, mR chart is not quite as good at differentiating within group vs. between group variation as the X-Bar, R chart, but it is still very useful. We'll find that it can also be used when count data don't meet their control chart assumptions.

Application

- *Financial Report Data* - Virtually all financial data that periodically comes to you in spreadsheet form today can be converted to X, mR Control Charts. Salary & Supply Expense, Productivity measures, Sales Figures (volumes and dollars), etc. are all candidates.
- *Periodic Equipment Measurements* - If your operating procedures or preventive maintenance program includes periodic measures of equipment or system performance (e.g. once a shift, or weekly/monthly testing), the critical parameters can be plotted on an X, mR chart (in our experience, here the "X" part of the chart is often the only one used).
- *Process Data "Samples"* - Very often, our first attempt a "playing" with a set of data from a process will be to prepare an X, mR control chart of the data. This will directly identify special causes in the data as well as provide clues that may lead to further, "better" analysis. For example, if the X, mR control chart shows several runs above and/or below

the mean, there may be a subgrouping strategy lurking behind these runs. The data could then be examined for stratifications or displayed on an X-Bar, S chart.

Construction of X, mR Chart

1. Collect the data. Again, about 25 individual data values are recommended to get this chart going, but you can start with less. Recalculate your control limits as you collect more data. Subgrouping rears its ugly head here as a stratification question. Try not to mix data from different processes or from the individual paths of a process.

2. Organize the data on a table that looks like the one below:

Subgroup	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**.**	**.**	**.**	**.**	**.**	**.**	**.**	**25**
Data																	
Range	x																

3. Calculate the ranges by taking the absolute value of the difference between sequential "X" values. Note that the "first" range is associated with the second subgroup.

$$R_2 = |x_2 - x_1|$$
$$R_3 = |x_3 - x_2|$$
$$R_4 = |x_4 - x_3|$$

etc.

where:

$|x_2 - x_1|$ - Absolute Value of $x_2 - x_1$

R_i - "ith" Subgroup Range

4. Calculate the average Range. Note that you have one fewer ranges than you do "X" values, so divide the total of the ranges by "k - 1" to get the average range.

$$\overline{R} = \frac{1}{k-1}\sum_{i=2}^{k} R_i$$

where:

$\overline{R}$ - Average Range

k - Number of Subgroups

5. Calculate the Upper Control Limit for the Ranges (the Lower Control Limit is not applicable) as follows:

$$UCL = 3.268 \times \overline{R}$$

(3.268 is the "D_4" coefficient for the X,mR Chart)

6. Prepare the graph paper as you did for the X-Bar, R Chart. Plot the average range on the graph as a *solid* line. Plot the Upper Control Limit on the graph as a *dashed* line. Plot the ranges, with the first range appearing under the second "X" value.

7. Check the mR chart for special causes using the same "rules" you used for the X-Bar, R chart.

X Chart

8. Calculate the average of the individual values ("X's"):

$$\overline{X} = \frac{1}{k}\sum_{i=1}^{k} X_i$$

where:

$\overline{X}$ - Average of X_i's

k - Number of Subgroups

9. Calculate the Upper and Lower Control Limits for the X's:

$$UCL_x = \overline{X} + 2.66 \times \overline{R}$$

$$LCL_x = \overline{X} - 2.66 \times \overline{R}$$

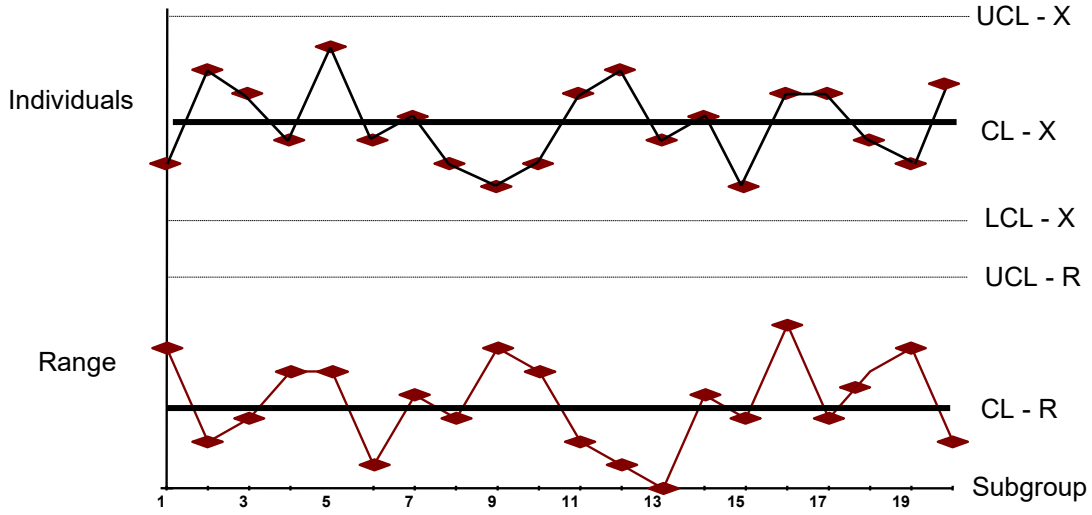
where:

UCL_X - Upper Control Limit for X

LCL_X - Lower Control Limit for X

(Again the "2.660" is the coefficient, except here it's known as E_2, not A_2.)

7. Now plot the average as a *solid* line on the X part of the chart. Plot the control limits as *dashed* lines on the X graph. Plot the X values and interpret the graph to determine the presence of special causes of variation in the process. Take action as appropriate.



A note on the X, mR Control Chart: If the data from your process is seriously skewed, then the chart may look a little funny. For example, if the data is skewed left, a broad gap may appear between the smallest values and the lower control limit. Take your data and plot it on a histogram to see if it is skewed. For more information, see Unit 6.9 – Additional Control Chart Topics – *Non-Normal Data and X, mR Charts.*

Example X, mR Control Chart

Scenario - A manufacturing plant monitors *offal* (waste or by-products) produced by a particular production line (data below divided by 10,000 lbs). One data point is collected every shift. The last six days' data is presented below:

Day	Mon		Tue		Wed		Thurs		Fri		Sat	
Shift	1	2	1	2	1	2	1	2	1	2	1	2
Offal (/10,000 lb)	0.10	0.08	0.10	0.12	0.11	0.13	0.13	0.14	0.17	0.18	0.17	0.18

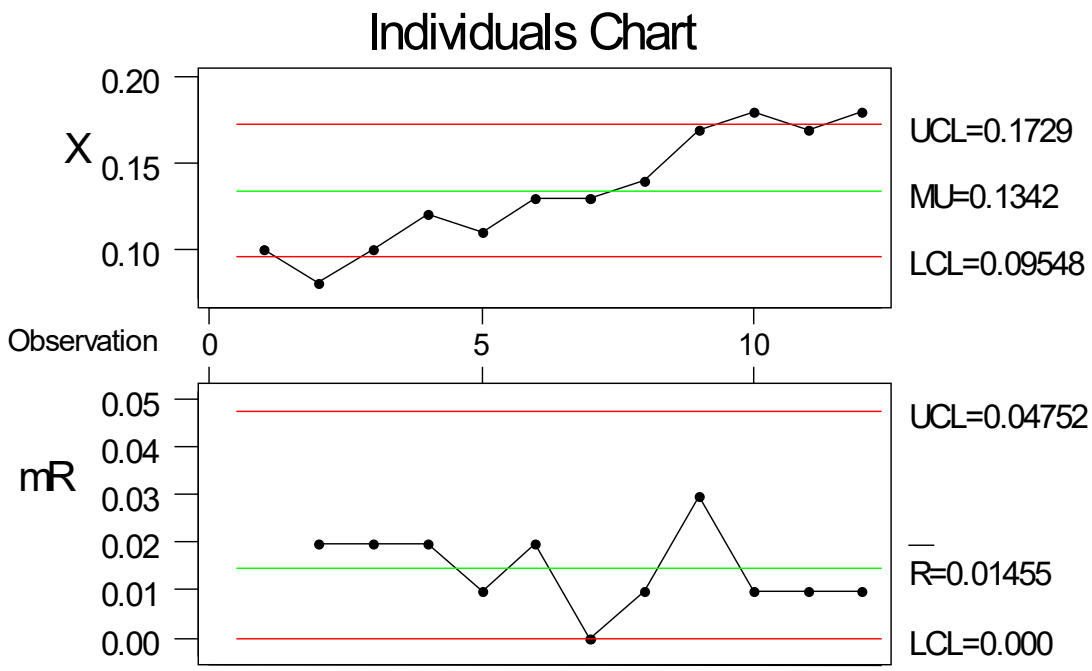
Here, the Range calculations are based on adjacent data points. The first Range is found by taking the absolute value of the difference between the first two points:

$$Range_1 = |0.10 - 0.08| = 0.02$$

The remaining calculations are shown below:

Average Range:	0.015
Upper Control Limit - Range:	0.048
Average of X's:	0.134
Upper Control Limit - X's:	0.173
Lower Control Limit - X's	0.095

Several points are out of control on the X-Chart. It's obvious from the data, though, that a long-term increasing trend is occurring. Investigate! Investigate!


Individuals Chart
X
0.20
0.15
0.10
UCL=0.1729
MU=0.1342
LCL=0.09548
Observation
0
5
10
mR
0.05
0.04
0.03
0.02
0.01
0.00
UCL=0.04752
R=0.01455
LCL=0.000

6.6 Attribute Control Charts

Learning Objectives

- Determine the appropriate distribution, Binomial, Poisson and others
- Construct and interpret *np* and *p* Control Charts for binomial distributions
- Construct and interpret *c and u* Control Charts for Poisson distributions

Unit Contents

- Control Charts for Attribute Data
- The *np* Control Chart
- The *p* Control Chart
- The *c* Control Chart
- The *u* Control Chart

6.6.1 Control Charts for Attribute Data

Now let's introduce the charts to use if you are dealing with count (synonyms: *count* and *discrete*) data. Two “families” of count data charts exist: the *np & p,* and the *c* & *u* charts. Before we start, there is one issue and one concept we need to address.

The issue has to do with how the control limits are calculated for these charts. They assume that the data being plotted can be accurately modeled by one of two types of probability distribution: the Binomial or the Poisson. In some cases, it will be difficult to assure that all the assumptions surrounding these distributions are met.

You have a "bail-out" option. The **X, mR** Chart can be used as an effective substitute for any of these charts (Review the control chart selection guide - you’ll see these paths). We’ll list the specific assumptions before each pair of count data charts.

Now let's examine the additional concept - the difference between *defects* and *defectives*.

Say that we have a standard (or specification) for a part dimension. If the part is manufactured within the specification limit, we treat it as OK, if not, then it's not OK and is reworked or scrapped. We could sample 100 parts, measure their dimensions and then count the number that did not meet the standard. Through this experiment we would have identified the number of *defectives*. The same concept applies to customer orders compared to a delivery time standard. Those that do not meet the standard (e.g. those that are not delivered on time) are *defective.*

In general, if we can look at an event or a thing and judge it to be OK or not OK, then we are dealing with defectives. Parts outside spec limits, errors, incorrect bills, late shipments, and “bad order” parts can all be considered as defectives. We will use the *np* or *p* control charts to deal with defectives.

Let's now consider a different situation. We’ll inspect a manufactured air conditioner. Are there any leaks? Are all the electrical wires terminated correctly? Are all the features the customer ordered present? Are there any coating scratches or undercoating areas? Here, we can count the number of problems with the unit. We will consider each of these problems to be a *defect*. The unit could have 0, 1, 2, 3 or more defects.

We can apply the same thinking to a Bill of Materials prepared by an engineer. Are all parts & materials listed? Are they all the correct parts/materials? Are the standard part numbers correct? Are the amounts correct? Here, too, we can count the number of *defects* on the Bill of Materials and the BOM could have 0, 1, 2, 3 or more defects.

To generalize, when we examine a "thing" and can count the number of things "wrong" with the thing, we are dealing with defects. The "thing" we are examining is given a special term: the *area of opportunity*. In our first example, the air conditioner was the area of opportunity, in the second, the Bill of Materials.

In many cases, especially in manufacturing situations, we can look at the surface area and count the number of defects. Take a painted surface of 1 square foot and count the number of scratches or pinholes; take a polished or finished surface and count the number of blemishes.

For situations where we are dealing with defects, the *c* and *u* charts will be our control charts of choice.

The terms *defective* and *defect* are from the manufacturing world. Let's consider these somewhat different examples from the *service world*:

1. *A customer service office was interested in the number of customers to use their "Hot Line" each day.*

2. *A warehouse was interested in the number of parts ordered by different technicians for a particular repair.*

3. *A manufacturer was interested in the number of customers who chose a certain type of option.*

Now these indicators are neither "defectives" nor "defects," but we would treat them using the concepts described above. The first two examples are count data that would be considered under the "defect" category, the third is an example of a "defective" type of data (they either did or did not choose the option).

Binomial Assumptions for np and p Control Charts

To test if your data meets the assumptions for a Binomial model, and hence is a candidate for the *np* or *p* control charts, ask these questions:

1. Can you identify a sample (subgroup) of items? (YES is Good)

2. Will each item in the sample (subgroup) be classified as either having, or not having some characteristic? (YES is Good)

3. Does the fact that one item possesses the characteristic affect the probability of the other items having the characteristic? (NO is good)

The first two tests are usually easy to pass. You've identified items to inspect, and you will either pass or fail each item in the group.

The third test is a little tricky. The concept here is one of **independence**. In many cases, the items *are* independent of each other. The fact that one item possesses the characteristic doesn't affect any others in the sample.

Here's an example where this Binomial test would not pass, though. A plant has a shutdown due to a tornado that causes several customer orders to be delayed. Here, the probability of one shipment being delivered late is not independent of the others.

Remember, if the data doesn't meet these assumptions, then the **X, mR** control chart may be a useful alternative.

Assignable Cause Tests for Attribute Control Charts

In Unit 6.5, we presented several tests for assignable causes. Of those, the following are applicable to count control charts:

- **Rule 1** - Points outside the control limits
- **Rule 4** – Seven points in a row on the same side of the center line
- **Rule 5**- Six points in a row, increasing or decreasing
- **Rule 6** - Fourteen points in a row, alternating up and down.

6.6.2 The np Control Chart for Yes/No Attributes – Constant Subgroup Size

The *np* chart is used where we are interested in the number of *defectives* resulting from a process. In addition, the number of items that we look at to count the number of defectives must be relatively constant. Here's two ways this can occur: 1) The volume of the process (say, weekly) stays about the same (i.e. every week, about 50 shipments are sent out) or, 2) we deliberately take samples of constant size from the process.

Why is this called an *np* chart? The *n* stands for the size of the subgroup that we take from the process. The *p* stands for the fraction of the process' output that is defective (0.2, 0.3, 0.002, etc.). The *np* taken together is simply the number of defectives.

Applications

The np chart is very popular in situations where we inspect the output of the process (or a sample of the output) and make a go/no go decision. In fact, this gives us a clue that we are using the correct control chart. When we look at the output from the process, at a *minimum*, all items could potentially be OK; at a *maximum*, all items could be not OK. Inspecting material received from a vendor, on-time order filling, incorrect bills, number of accounts receivable greater than 60 days old are candidates for an np chart.

A coil assembly process improvement team was concerned with the need to rework assemblies. To measure the performance of this process, the team took a subgroup of 100 coils assembled each week and counted the number of coils with leaks. They tracked this number using an *np* control chart.

Construction of np Chart

1. Collect the data. Here, it is important to make the number of items we will look at in each sample as constant as possible.

> *If you can set up a sampling procedure, then you can decide to look at a constant sample size each day or week or month. If you are looking at the entire process' output, then the volume of the process should be roughly constant from week to week or month to month. Here's a thumb rule: The individual sample sizes should not vary by more than about +/- 25% from the average sample size.*

If the average volume of a process is 100 "widgets" per week, and the individual weekly volumes are never more than 125 per week nor less than 75 per week, then you're OK. If this is a problem, use the p chart instead[1].

2. Count the number of defective items in each sample and record them in a table such as appears below. Notice that we still call each sample of data a "subgroup."

Subgroup	1	2	3	4	5	6	7	8	9	.	.	.	.	.	.	.	25
Number Defective																	
Size of Subgroup																	

3. Calculate the average number of defective items per subgroup. Add up all the defectives and divide by the number of subgroups to get this average.

$$\overline{np} = \sum_{i=1}^{k} np_i \Big/ k$$

where:

np_i - Number of Defective Items, "ith" subgroup

k - Number of Subgroups

$\overline{np}$ - Average Number of Defectives per subgroup

4. Calculate the Upper and Lower Control Limits for the np Chart. If your subgroup size has varied, calculate the average subgroup size first:

$$\overline{n} = \frac{1}{k}\sum_{i=1}^{k} n_i$$

where:

n_i - "ith" Subgroup Size

k - Number of Subgroups

and

[1]Statistical software generally doesn't cut you any slack on this. If your subgroup sizes are not the same, they'll force you to use the *p* chart.

$$UCL_{np} = \overline{np} + 3\sqrt{\overline{np}(1 - \overline{np}/\overline{n})}$$

$$LCL_{np} = \overline{np} - 3\sqrt{\overline{np}(1 - \overline{np}/\overline{n})}$$

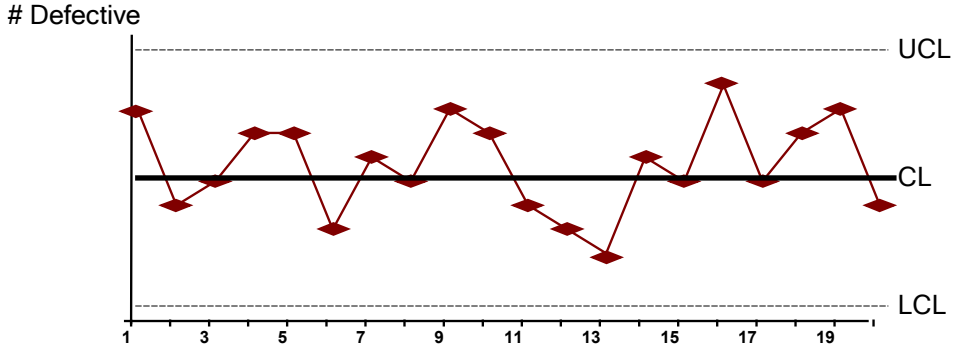
where:

UCL_{np} - Upper Control Limit

LCL_{np} - Lower Control Limit

$\overline{n}$ - Constant (or Average) Subgroup Size

4. Plot the average number of defectives as a *solid* line on your graph. Plot the Control Limits as *dashed lines* on the *np* chart. Plot the number of defectives on the graph for each subgroup. Note that sometimes the Lower Control Limit is calculated to be less than zero. In these cases, the Lower Control Limit is not applicable. Your control chart should look something like this:



6. Interpret the control chart. Look for special causes using the rules presented in Unit 6.5 (and applicable to count control charts). Take action as appropriate.

Example np Control Chart

Scenario – An electronic controller manufacturer screens components purchased from vendors. Each month, a batch of 10,000 Integrated Circuit chips is received from a particular vendor. Automatic machinery tests each chip, accepting or rejecting the chip. The number of rejected chips for the last few months is provided below:

Month	J	F	M	A	M	J	J	A	S	O	N	D
# Rejects	10	8	14	6	23	15	11	8	12	13	17	14

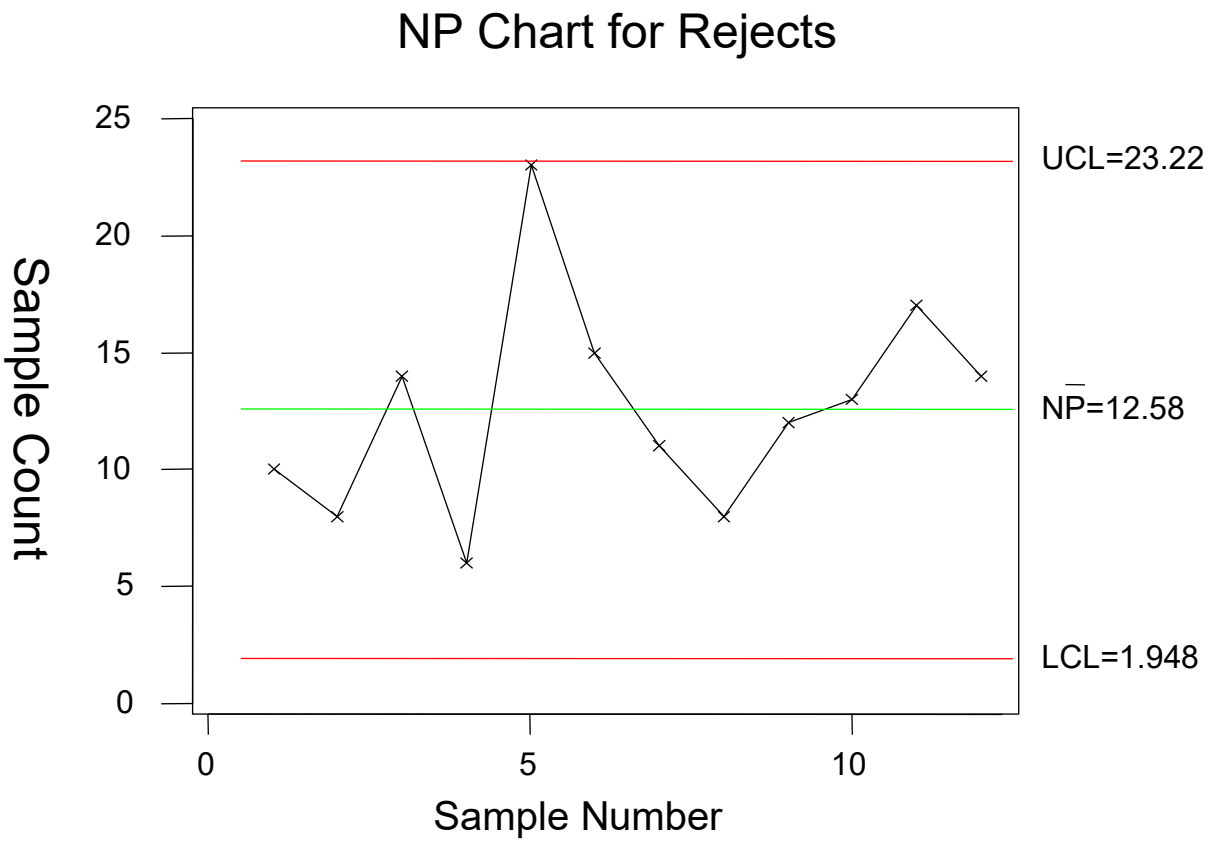
We first find the average number of defective chips:

$$\overline{np} = \frac{10+8+14+6+23+15+11+8+12+13+17+14}{12}$$
$$= \frac{151}{12}$$
$$\overline{np} = 12.58 \text{ defectives / batch}$$

The Upper and Lower Control Limit calcs are the only "interesting" part of this chart:

$$UCL = 12.58 + 3 \times \sqrt{12.58 \times (1 - 12.58/10000)}$$
$$= 12.58 + 3 \times \sqrt{12.58 \times (1 - 0.00126)}$$
$$= 12.58 + 3\sqrt{12.56}$$
$$= 12.58 + 3 \times 3.54$$
$$= 12.58 + 10.63$$
$$UCL = 23.21 \text{ defectives / batch}$$
$$\textit{and}$$
$$LCL = 12.58 - 3 \times \sqrt{12.58 \times (1 - 12.58/10000)}$$
$$= 12.58 - 10.63$$
$$LCL = 1.95 \text{ defectives / batch}$$

Even though one of the points (May - 23 defectives) is close to the UCL, we would call this process in-control. We can expect the in-coming defective percentage from this vendor to be about 0.13%, based on our screening test. The Minitab output for this data is shown below:



6.6.3 The p Chart for Yes/No Attributes – Variable Subgroup Size

Purpose

The *p* chart is a variation of the *np* chart. By tracking the *fraction defective*, it handles the situation where you either can't or don't want to keep the size of your sample constant. There's a price to pay for a varying sample size, though.

The Control Limits are going to vary from subgroup to subgroup. This makes calculations a bit more tedious and makes interpretation of the chart a bit more difficult. PC programs take the drudgery out of this task. If you find yourself doing lots of p-charts, you'll want to invest in a good SPC software package.

Applications

Any of the *np* applications can also be charted using a *p* chart, if the subgroup size varies. Some like to use the *p* chart even if the subgroup size is held constant, since the percent or fraction defective has more meaning to them.

- *Number of incorrect shipments per week* - The number of shipments (and, hence, the sample size) varies widely from week to week. The fraction defective - number of incorrect shipments divided by the number of shipments per week would be an appropriate application of the *p* chart.
- *Fraction of Lost Orders - Parts* - The number of orders received by a service location will vary from day to day. The *p* chart may be used to track this fraction.
- *Fraction of incorrect repairs performed by technicians* - When a repair is performed, a test is run to determine its effectiveness. On occasion the repair performed is incorrect. A comparison here may be done across technicians - each technician's repairs form a subgroup.

Construction of p Chart

1. Collect the data. Record both the number of defective items and subgroup size on a table such as appears below:

Subgroup	1	2	3	4	5	6	7	8	9	.	.	.	.	.	.	.	25
Number Defective																	
Subgroup Size																	
Fraction Defective																	
Upper Control Limit																	
Lower Control Limit																	

2. Calculate and record the fraction defective for each subgroup. Divide the number defective by the subgroup size to obtain these values. If you want, you can change these to *percent defective*. The calculations are slightly different:

Fraction Defective (or Percent Defective):

$$p_i = \frac{np_i}{n_i} \qquad p_i = \frac{np_i}{n_i} \times 100\%$$

where:

np_i - Number defective - "ith" subgroup

n_i - Subgroup size - "ith" subgroup

p_i - Fraction defective - "ith" subgroup

3. Now calculate the average fraction defective. Sum the number of defectives from each subgroup. Sum the subgroup sizes. Divide the first by the second[2]:

$$\overline{p} = \sum_{i=1}^{k} np_i \Big/ \sum_{i=1}^{k} n_i$$

where:

$\overline{p}$ - Average fraction defective

or, percent defective $= \overline{p} \times 100\%$

[2]Remember, with different subgroup sizes, you can't average the individual subgroup fractions or percentages!

4. Finally, calculate and record the Upper and Lower Control Limits for **each** subgroup:

$$UCL_p = \overline{p} + 3 \times \sqrt{\overline{p}(1-\overline{p})/n_i} \text{ and } LCL_p = \overline{p} - 3 \times \sqrt{\overline{p}(1-\overline{p})/n_i}$$

or, for percent defective:

$$UCL_p = \overline{p} + 3 \times \sqrt{\overline{p}(100-\overline{p})/n_i} \text{ and } LCL_p = \overline{p} - 3 \times \sqrt{\overline{p}(100-\overline{p})/n_i}$$

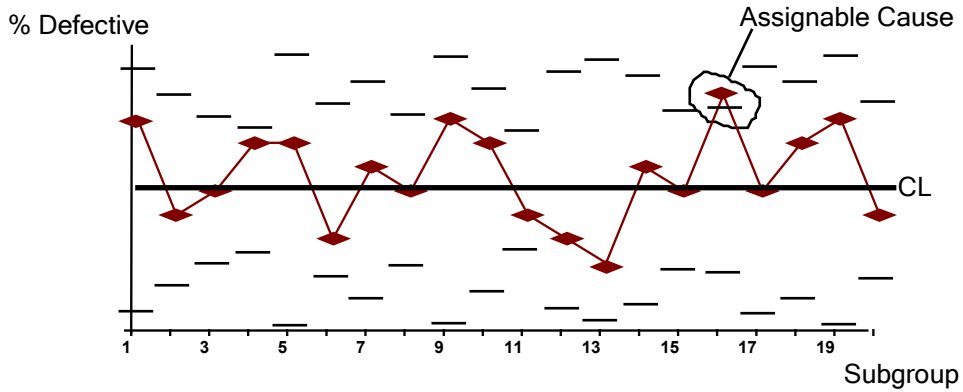
where:

UCL_p, LCL_p - Upper & Lower Control Limits

n_i - "ith" subgroup size

Notice the equation for the Upper and Lower Control Limits. Since the subgroup size is in the denominator, the larger the subgroup size, the tighter the control limits. The larger the subgroup size, the less uncertainty we have about the "true" value of the fraction defective coming from our process. If a special cause of variation does change the process's fraction defective, the more data we collect, the easier it will be for us to see the effect of the change (i.e. as a point outside of the control limits).

5. Plot the average fraction defective as a *solid* line on your graph. Plot the fraction defectives. Draw the individual control limits as *dashed lines* above and below the points. Again, if the Lower Control Limit is calculated to be less than zero, then it is not applicable. Your control chart should look something like this:



6. Interpret the chart. Use the rules for detecting special causes of variation discussed in Unit 6.5 (and that are applicable to count control charts). Take action as appropriate.

Example p Control Chart

Scenario - A QC group was concerned about their efficiency. One of their indicators is the *Percentage of Units Delayed.* For the last few weeks, they've collected data on the number of inspections performed in the plant and the number that were delayed (operationally defined to have taken more than 15 minutes to inspect). Here's their data:

Week	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
# Delayed	8	10	11	5	14	8	6	6	8	12	11	9
# Inspections	44	60	54	35	60	48	63	72	42	38	49	58

The calcs are a bit more complicated here, since we must keep straight when we're dealing with the individual subgroups and when we must combine all the data. Here's the procedure:

First, calculate the subgroup fractions (or percentages). Here, we'll use the data from the individual subgroups:

$$\text{Week 1: } p_1 = \frac{8}{44} = 0.182 \text{ or } 18.2\%$$

$$\text{Week 2: } p_2 = \frac{10}{60} = 0.167 \text{ or } 16.7\%$$

. . . . . .

$$\text{Week 12: } p_{12} = \frac{9}{58} = 0.155 \text{ or } 15.5\%$$

Now, to calculate the average fraction delayed, we go back to the raw data:

$$\overline{p} = \frac{8+10+11+5+14+8+6+6+8+12+11+9}{44+60+54+35+60+48+63+72+42+38+49+58} = \frac{108}{623}$$

$$\overline{p} = 0.173 \text{ or } 17.3\%$$

The Upper and Lower Control Limits get calculated for each subgroup:

$$\text{Week 1:}$$

$$UCL_1 = 0.173 + 3 \times \sqrt{\frac{0.173(1-0.173)}{44}} = 0.173 + 3 \times \sqrt{\frac{0.143}{44}}$$

$$= 0.173 + 3 \times \sqrt{0.00325} = 0.173 + 3 \times 0.0570$$

$$= 0.173 + 0.171$$

$$UCL_1 = 0.370 \text{ or } 37.0\%$$

$$LCL_1 = 0.173 - 3 \times \sqrt{\frac{0.173(1-0.173)}{44}}$$

$$= 0.173 - 0.171$$

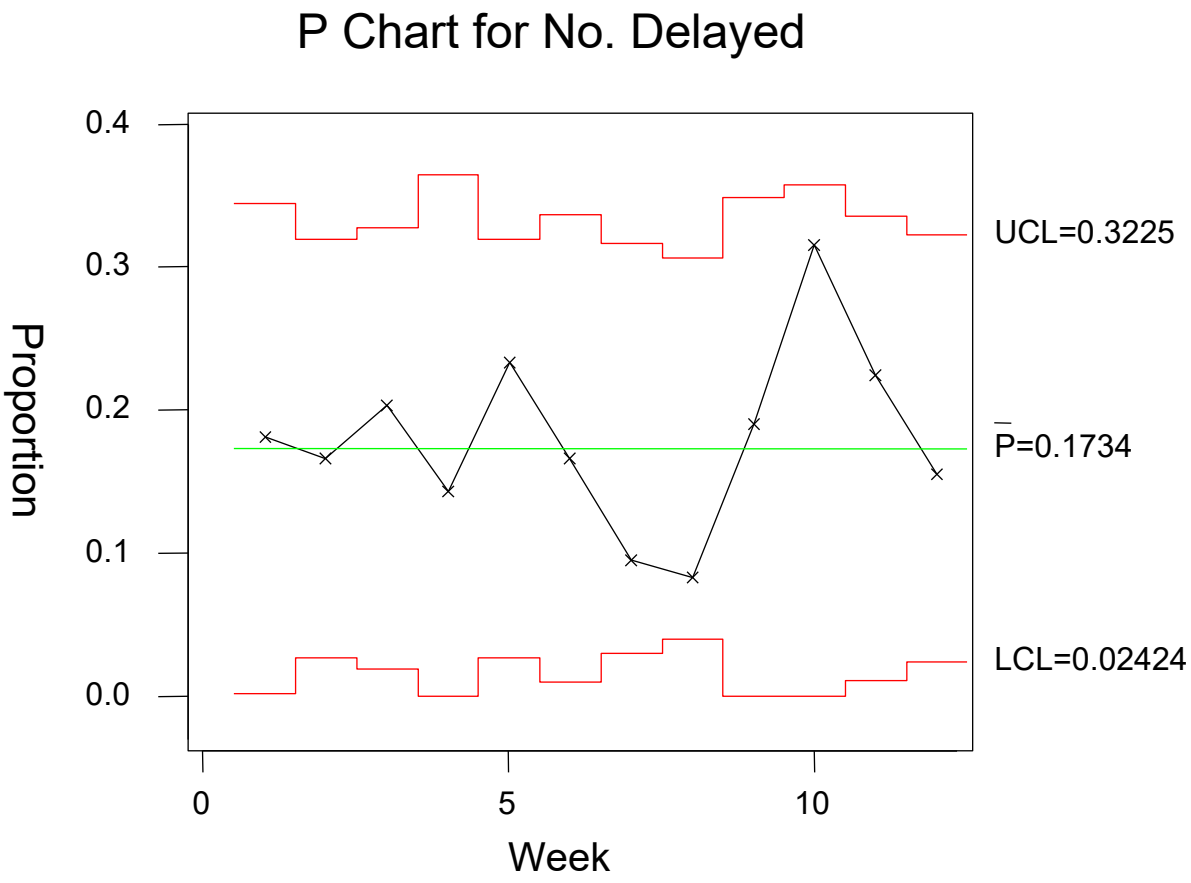
$$LCL_1 = 0.002 \text{ or } 0.2\%$$

Here's a summary of the remaining calculations for this control chart:

Week	1	2	3	4	5	6	7	8	9	10	11	12
# Delayed	8	10	11	5	14	8	6	6	8	12	11	9
# Cases	44	60	54	35	60	48	63	72	42	38	49	58
Fraction Defective	0.182	0.167	0.204	0.143	0.233	0.167	0.095	0.083	0.19	0.316	0.224	0.155
UCL	0.344	0.319	0.327	0.365	0.319	0.337	0.316	0.307	0.348	0.357	0.335	0.322
LCL	0.002	0.027	0.019	NA	0.027	0.009	0.03	0.039	NA	NA	0.011	0.024

None of the data fall outside the varying control limits here. Based on this evidence, we would declare this system to be stable at "producing" delayed inspections and look for the common causes of variation present in the system.

The Minitab output for this data appears below:



Poisson Assumptions for c and u Control Charts

To test if your data meets the assumptions of a Poisson model, and hence is a candidate for the *c* or *u* control charts, ask these questions:

1. Are you counting discrete events? (YES is Good)

2. Do these events occur within some well-defined area of opportunity (characterized by at least a *spatial* dimension and possibly by a *time* dimension)? (YES is Good)

3. Does the occurrence of one event affect the likelihood of another? (NO is Good)

4. Are the events rare? (YES is Good)

The first assumption is usually easy to address. The second one (area of opportunity) was discussed earlier - defining it well may take some thought about the factors influencing the occurrence of the events (i.e. if we are not brazing today then it will be difficult to observe brazing defects).

The third assumption is like the **independence** issue raised for the Binomial model.

The fourth assumption takes a little thought. One way to test this assumption is to consider the "theoretical" number of events that could occur within "one" area of opportunity. For instance, if we considered a square foot of painted surface, there *could* be millions of scratches or nicks. In a one-month period, "*billions and billions"* of errors in records or warranty claims could occur. If the actual occurrence rate, though, is less than 10% of the "theoretical," then the events may be considered "rare."

If these assumptions are not met, the **X, mR** control chart may be used as an alternative.

6.6.4 The c Chart for "How Many" Attributes – Constant Area of Opportunity

Purpose

The last two charts help us track the number (or rate) of *defects* produced by our processes. Remember the difference between defects and defectives. When we look at process output and decide whether it's OK or not OK, we are working with defectives. When we look at process output and count the number of problems with the output, we are dealing with *defects*. The *c* Chart is used when we are counting the number of defects, *and* the *area of opportunity* is constant from subgroup to subgroup. What does this mean?

Applications

Employee injuries can be candidates for a c chart. *Billing errors* are also candidates. What is the area of opportunity for these quality characteristics? How can we make the area of opportunity "constant?"

The *area of opportunity* usually includes both *time* and *place*. For the injuries, the place could be the warehouse, and the time could be one week or one month. Our indicator would be the *number of injuries per week in the warehouse.*

For billing errors, the "place" is the bill and "time" could be a sample of 100 bills. An indicator could be the *number of errors per 100 bills.* These could be considered as constant areas of opportunity.

You might be thinking that there are some other factors besides time and place that may affect the of employee injuries, such as the number of employees. We might better define the injury indicator to include some measure of this number (i.e. injuries per 100,000 hours worked). This addition might improve the "constancy" of our area of opportunity.

Defining the area of opportunity is the most important aspect of both the c and u charts.

Construction of the c Chart

1. Collect the data. As discussed above, define the area of opportunity as carefully as possible. Sometimes sampling from the process output can help make the area of opportunity "constant." For instance, if we are tracking the number of errors per customer record, we might choose to sample 100 records per week and count the number of errors in these records. Record the data on a form like the one below:

Subgroup	1	2	3	4	5	6	7	8	9	.	.	.	.	.	25
Number of Defects															

2. Calculate the average number of defects. Add up all the defects and divide by the number of subgroups.

$$\bar{c} = \frac{1}{k}\sum_{i=1}^{k} c_i$$

where:

c_i - Number of defects, "ith" subgroup

k - Number of subgroups

$\bar{c}$ - Average number of defects

3. Calculate the Control Limits. They are easy for this chart:

$$UCL_c = \bar{c} + 3 \times \sqrt{\bar{c}}$$

$$LCL_c = \bar{c} - 3 \times \sqrt{\bar{c}}$$

where:

UCL_c - Upper Control Limit

LCL_c - Lower Control Limit

4. Draw the average number of defects as the *solid* Center Line on the graph. Plot the data. Plot the control limits as *dashed* lines on your graph. This chart should look about the same as the *np* chart.

5. Interpret the control chart. Use the rules discussed in Unit 6.5 (and that are applicable to count control charts). Take action as appropriate.

Example c Control Chart

Scenario - A manufacturer of industrial X-Ray film has been receiving complaints of defects found on the film by customers. As they begin to analyze the problem, they sample film rolls from the production process and inspect the film for defects. Here's the data they collected from 24 rolls of X-Ray film:

Film	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
# Def.	8	16	14	19	11	15	8	11	21	12	23	16
Film	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
# Def.	9	25	15	9	9	14	11	9	10	22	7	28

The c control chart is the simplest of all. We first find the average number of defects/roll:

$$\bar{c} = \frac{8+16+14+19+11+15+8+11+21+12+23+16+\ldots+10+22+7+28}{24}$$

$$= \frac{342}{24}$$

$$\bar{c} = 14.25 \text{ defects / roll}$$

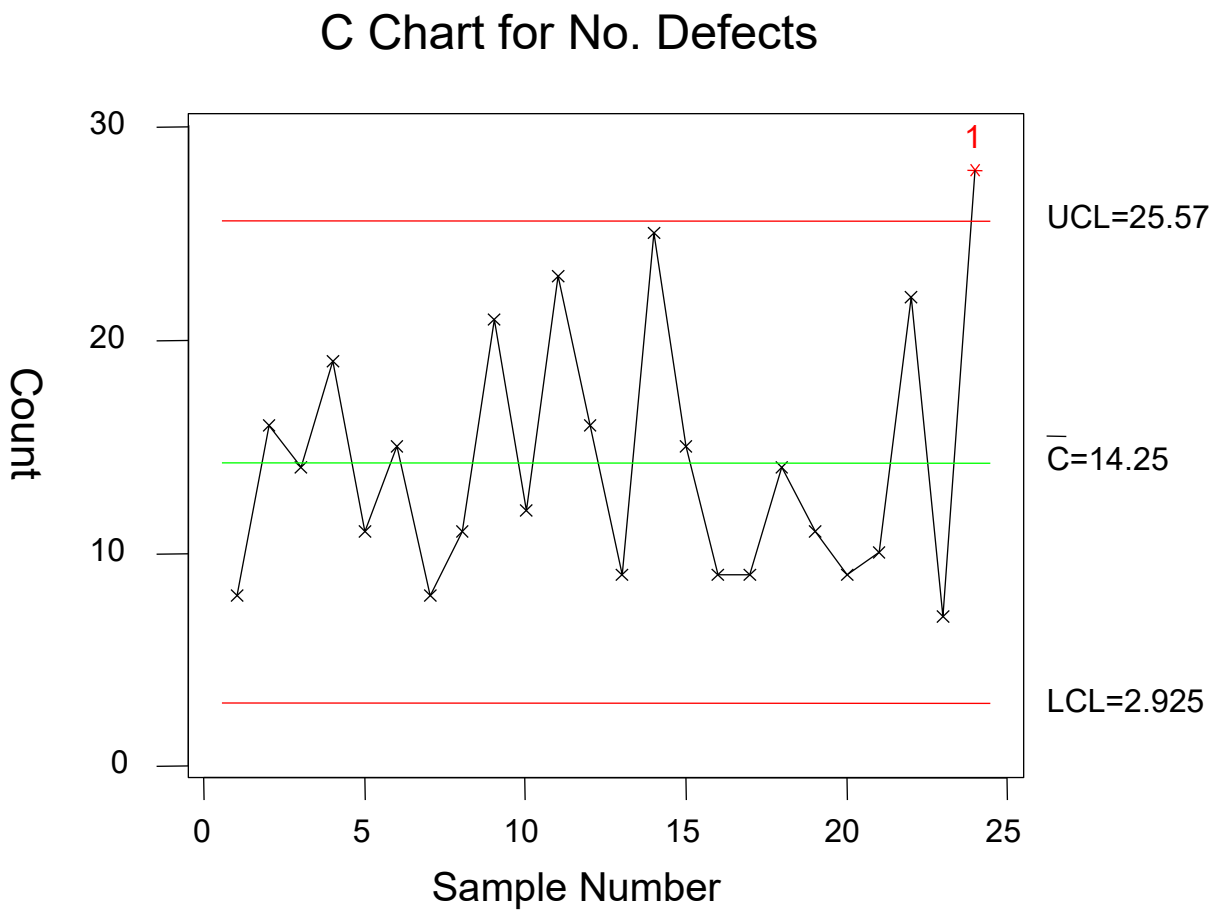
The upper and lower control limits are then very easy to calculate:

$$LCL = 14.25 + 3 \times \sqrt{14.25}$$
$$= 14.25 - 11.31$$
$$LCL = 2.94$$

$$UCL = 14.25 + 3 \times \sqrt{14.25}$$
$$= 14.25 + 3 \times 3.77$$
$$= 14.25 + 11.31$$
$$UCL = 25.56$$

The last X-Ray roll's number of defects (28) is above the Upper Control Limit and should be investigated as an assignable cause. The rest of the points fall inside the limits (and there are no patterns) indicating that film defects are ordinarily produced by the "production" system.

The Minitab output for this data appears below:



6.6.5 The u Chart for "How Many" Attributes – Variable Area of Opportunity

Purpose

The *u* chart tracks the *rate* of defects occurring in our processes. The *u* chart and *c* chart are related the same way as the *p* and *np* charts. The *p* chart was used when we could not or did not want to keep the subgroup size the same. The *u* chart is used instead *of* the *c* chart when we cannot or do not want to keep the area of opportunity the same from subgroup to subgroup.

The *u* chart suffers from the same difficulty as the *p* chart in that the control limits will have to be calculated for each subgroup, and the interpretation of the chart is a bit more difficult. Again, an SPC software package will make this job easier.

Applications

Any application that fits the *c* chart can also be charted on a *u* chart.

Let's take an example that we used for the *c* chart and turn it into a *u* chart. If we were interested in developing a control chart for billing errors that were occurring and wanted to plot a point each day, we would be faced with the issue of a varying number of bills per day.

To address this issue, we could record both the number of billing errors per day, **and** the number of bills prepared per day. From these data, we could calculate the *rate* of billing errors (total billing errors per day). Even though one aspect of our area of opportunity varies, the u chart can handle the problem.

Construction of the u Chart

1. Collect the data. Here, we must collect both the number of defects and some measure of the changing area of opportunity. For example, we could collect the number of patient falls as the defects and the daily patient census as the area of opportunity. Record the data on a table that looks like the one below:

Subgroup	1	2	3	4	5	6	7	8	9	.	.	.	.	.	.	.	25
Number of Defects																	
Subgroup Size																	
Defect Rate																	
Upper Control Limit																	
Lower Control Limit																	

2. Calculate and record the defect rate for each subgroup. Divide the number of defects for each subgroup by the size of that subgroup:

$$u_i = \frac{c_i}{n_i}$$

where:

c_i - Number of defects, "ith" subgroup

n_i - Area of opportunity, "ith"subgroup

u_i - Defect rate, "ith" subgroup

3. Calculate the average number of defects. Sum the number of defects. Sum the subgroup sizes. Divide the first by the second:

$$\overline{u} = \sum_{i=1}^{k} c_i \Big/ \sum_{i=1}^{k} n_i$$

where:

$\overline{u}$ - Average number of defects

k - Number of subgroups

4. Calculate the Upper and Lower Control Limits for each point:

$$UCL_u = \overline{u} + 3 \times \sqrt{\overline{u} / n_i}$$

$$LCL_u = \overline{u} - 3 \times \sqrt{\overline{u} / n_i}$$

where:

UCL_u - Upper Control Limit

LCL_u - Lower Control Limit

5. Draw the average defect rate as a *solid* Center Line on the graph. Plot the individual defect rates and the individual upper and lower control limits as *dashed* lines above and below the points. This chart should look like the *p* chart.

If any of the Lower Control Limits are negative, disregard these values.

6. Interpret the Control Chart. Use the rules for detecting special causes of variation discussed in Unit 6.5 (and that are applicable to count control charts). Take action as appropriate.

Example u Control Chart

Scenario - An insurance company has gathered employee injury data from 8 plants operated by a company. The insurance company has ranked the injury rates from highest to lowest and wants the company to come up with a plan to reduce the rates at the top three plants. As the risk manager for the company, how would you respond? Here is the data:

Plant	North Hills	Jonson	Foggy Bottom	Fairview	Crab Apple	Ithaca	Ricker's Corners	Davis Hill
Injuries	23	20	19	17	16	16	14	10
Employee Hours	200,000	210,000	185,000	230,000	170,000	190,000	350,000	314,000

Before we jump to conclusions about the "top three," let's see if there's evidence to single these plants out. We'll construct a *u* control chart of this data:

As with the *p* control chart, we must be careful about how we calculate the *u* control chart's components. We'll first calculate the injury rates:

$$North\ Hills\ Plant:$$

$$u_{North\,Hills} = \frac{23}{200{,}000}$$

$$u_{North\,Hills} = 0.00012\ injuries/hour$$

Then, we'll calculate the average injury rate:

$$\bar{u} = \frac{23+20+19+17+16+16+14+10}{200{,}000+210{,}000+185{,}000+230{,}000+170{,}000+190{,}000+350{,}000+314{,}000}$$

$$= 135/1{,}849{,}000$$

$$= 0.000073\ injuries/hour$$

Now we will calculate the Upper and Lower Control Limits for each subgroup:

North Hills Plant :

$$UCL_{North\,Hills} = 0.000073 + 3 \times \sqrt{\frac{0.000073}{200{,}000}} = 0.000073 + 3 \times \sqrt{3.65E-10}$$

$$UCL_{North\,Hills} = 0.000073 + 3 \times 0.000019 = 0.000073 + 0.000057$$

$$UCL_{North\,Hills} = 0.000130$$

$$LCL_{North\,Hills} = 0.000073 - 3 \times \sqrt{\frac{0.000073}{200{,}000}} = 0.000073 - 0.000057$$

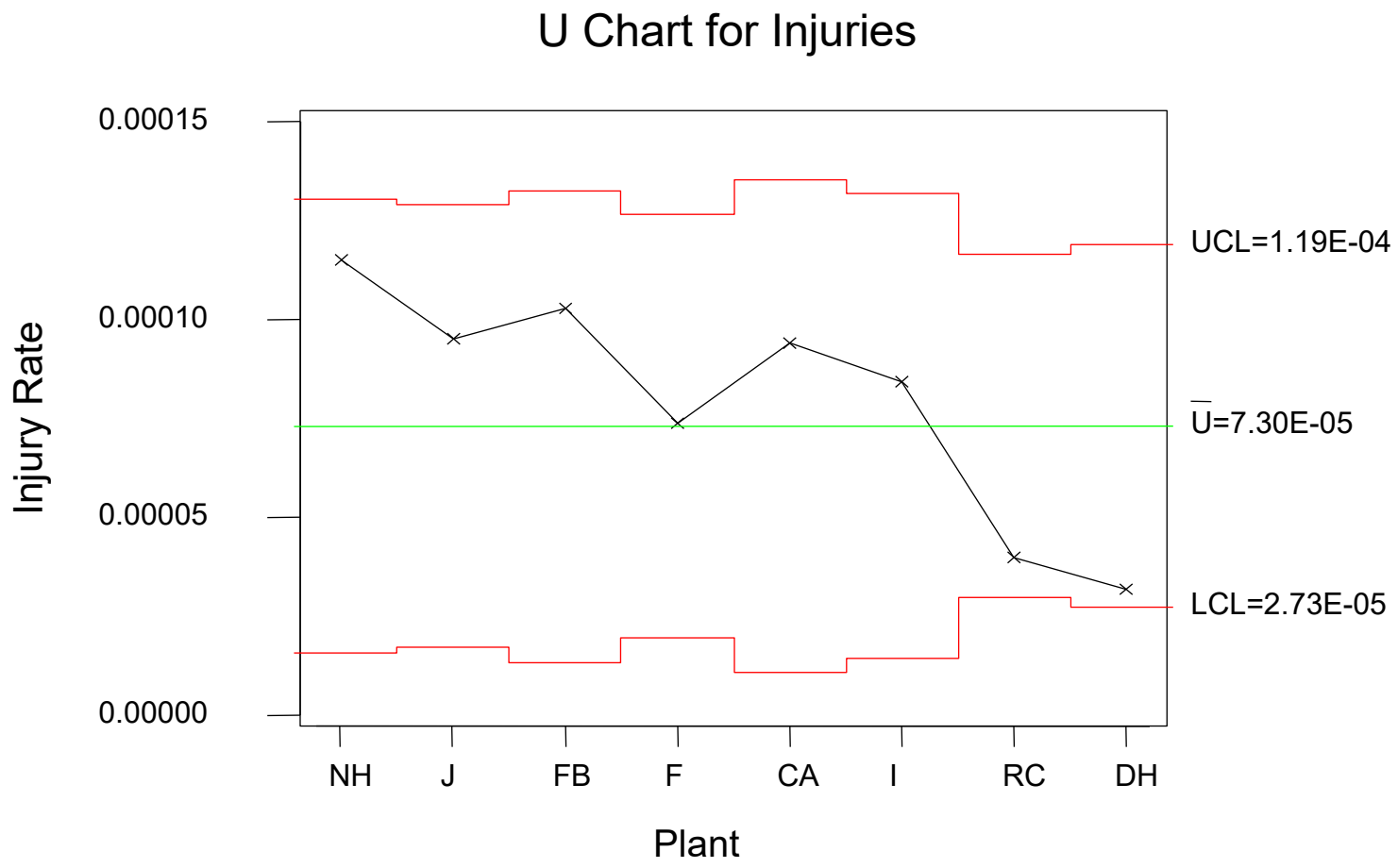
$$LCL_{North\,Hills} = 0.000016$$

You can see that the calculations get messy when we're dealing with very small rates such these injury rates. We could have applied a multiplier (i.e. injuries per 1000 hours) to make the calcs easier. Here are the rest of the calculations:

Plant	North Hills	Jonson	Foggy Bottom	Fairview	Crab Apple	Ithaca	Ricker's Corners	Davis Hill
Injuries	23	20	19	17	16	16	14	10
Employee Hours	200,000	210,000	185,000	230,000	170,000	190,000	350,000	314,000
u	0.000115	0.000095	0.000103	0.000074	0.000094	0.000084	0.00004	0.000033
UCL	0.00013	0.000129	0.000133	0.000126	0.000135	0.000132	0.000116	0.000119
LCL	0.000016	0.000017	0.000013	0.00002	0.000011	0.000014	0.00003	0.000027

None of the plants fall outside the control limits. Our response to the insurance company would be that it is inappropriate to rank the plants and demand that the "top three" improve their rates. They are making a "Hasty" type of error; treating common cause variation as if it was special cause. We might decide, though, to begin a program of improvement across **all** the plants to reduce employee injuries.

The Minitab output for this data appears below:



6.7 Measurement System Analysis

Learning Objectives

- To Understand the Variation/Errors associated with a Measurement System
- To Conduct a Gauge R&R Study
- To Analyze and Correct Problems with the Measurement System

Unit Contents

- The Measurement System as a Source of Variation
- Measurement System Properties
- Measurement System Studies

6.7.1 The Measurement System as a Source of Variation

We have established that all process outputs vary, regardless of whether the process' purpose is to manufacture a chiller or air handler, or to obtain a sales order or bill a customer. Some of the variation is due to the process itself, the methods, materials, operators, equipment, etc. that make up the process. But the *system* used to *measure* the process is also a source of variation or error. Observed variation is then a function of both process and measurement system variation:



$$X = P + E$$

For a collection of measures:

$$Avg.Measurement = Avg.of\ Product\ Values + Avg.Measurement\ Error$$

$$Mean(X) = Mean(P) + Mean(E)$$

$$We\ Want\ Mean(X) = Mean(P), so\ Mean(E)\ needs\ to\ be\ zero\ (0) - \boldsymbol{unbiased}$$

We use measurement systems in four different ways:

Way	Description
Description	To describe the measured item. Answers the "how much" or "how many" question. Need to understand the uncertainty attached to the value. Primary source = measurement system.
Characterization	To characterize the measured item relative to specifications. Measurement is used to take action on the item measured. Accept or reject the item. Primary source = measurement system.
Representation	To represent the product *not* measured (sampling, subgrouping). Object is to characterize the *product stream* with regard to specifications. Need to extrapolate from product measured to

	product not measured. Sources = measurement system, product variation, selection procedure.
Prediction	To characterize future process outcomes. Past behavior is applied to future performance. Concerned with measurement error, predictability of the process and combined variation from these sources (continuation over time)

A measurement system is more than just a measuring device, such as a micrometer or pressure gage, it includes the entire measurement process:

- The instrument or measuring device
- The human operator of the device
- The product (or process event) itself
- The measurement process (including environment)



Event



Instrument



Observer

We will be concerned with the following sources of error inherent in the measurement system:

Bias – If the length of a standard, reference block of metal is 1.000" and repeated measurements using a caliper produce an average of 1.015," then bias exists. Bias may be defined as the difference between the observed average of a series of measurements and the reference or master value (e.g. a value established by a "higher" level of measuring equipment – as applied at a national laboratory.). Although bias is sometimes referred to as accuracy, we will not use the accuracy

term, due to its numerous common meanings. Bias often exists when instruments are not subject to periodic calibration (*note, though, that a calibration program is just one element of a measurement system program designed to reduce both bias and variation).*

Variation – Repeated measurements of the reference block of metal by a particular caliper and operator will not be the same. The spread of these repeated measurements (as displayed on a histogram and summarized by a standard deviation) represents the variation in the measurement system. Variation is caused by meter or gauge friction, instrument wear, deterioration and/or environmental conditions. Four components of variation exist:

Repeatability – If one operator, using the same instrument measures a given characteristic of a part, the variation in these measurements is known as the repeatability.

Reproducibility – If *different* operators, all using the same instrument measure a given characteristic of a part, the variation in their *average* measurements is known as reproducibility.

From experience, *repeatability* and *reproducibility* tend to be the largest contributors to measurement system variation, especially one where the instrument has already been determined suitable to measure the process. Hence, “Gauge R&R” (repeatability and reproducibility) studies are often conducted to assess the measurement system.

Stability - Over time, an instrument may drift. The difference in average measurements taken of the same part over time by a given measurement system is known as stability. Note that there is a difference between the meaning of stability here and the notion of *statistical stability* described elsewhere in this manual. Two measurement systems may be *statistically stable,* yet one system may display a higher drift over time and therefore be less stable as a measurement system. In either case, though, we will require that the measurement system displays statistical stability as one of its *key characteristics.*

Linearity – Finally, a particular instrument is designed to measure parts of varying sizes (or other characteristics). As the instrument is exercised over its operating range (again, measuring reference or master values), the difference in the *bias* values is known as linearity.

One important issue associated with a measurement system is its ability to *discriminate* differences in part or process variation. Discrimination here is defined as the ability to detect and faithfully indicate (or *resolve*) “small” changes in the characteristic being measured. Inadequate discrimination of a measurement system may invalidate its use for a given

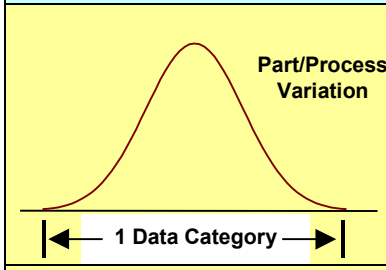
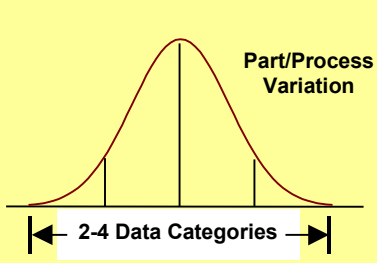
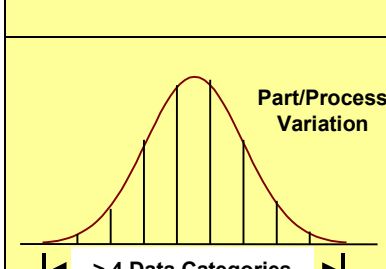
part or process. Most standards recommend that the resolution of the measurement system be one-tenth of the total process variation (expressed as six times the process standard deviation); a less conservative recommendation is one-tenth the total tolerance spread (Upper Specification Limit – Lower Specification Limit).

The above discussion focused on *measurements* of dimensions, pressures or time. Some measurement systems are of the "GO/NO-GO" type – these systems measure part or process *attributes.* Although the analysis methods will differ, the same concepts apply to *attribute* measurement system analysis.

6.7.2 Measurement System Properties

Given that there is no "ideal" measurement system (e.g. one that always measures the "exact" value of a part or process' characteristic, we define the desirable properties of a measurement system in terms of its statistical behavior:

Discrimination/Resolution – The measurement system must be able to adequately resolve changes in the part or process characteristics. Since resolution is defined *relative* to the part/process variation (and/or tolerance width) and since a measurement system may be used to *control* or *analyze* a process, the following table indicates how a measurement system may be employed based on its ability to resolve part/process changes:

Resolution	Use to Control a Process	Use to Analyze a Process
	• Use only if part/process variation is small compared to tolerance width or if main source(s) of process variation result in a shift in the process mean	• Only good to determine if part or process is defective or not • Don't use to estimate process characteristics' mean/variation
	• Marginal if data displayed on measurement control charts (X-Bar, R, X-Bar, S, X, mR) – Range chart will not resolve the process' variation (only a few data values will be possible); X-Bar, X control limits will be artificially reduced by lower R-Bar value, resulting in false out-of-control signals.	• Will only provide coarse estimates of process characteristics' mean/ variation
	• Will support use of measurement control charts	• Will provide good estimates of process characteristics' mean/variation – can be used to analyze effects of changes, Designed Experiments

Statistical Stability – The measurement system must be in statistical control. No special causes of variation should be present (including those due to operator, the measurement process, the instrument itself or the process' environment). Note that the *contra* of this is to **not** overreact to changes in the measurement system. A good practice is to check the measurement system against a standard value – if the measurements display statistical stability (i.e. determined by plotting on a control chart), then do not recalibrate the instrument.

Variability – The variability of the system must be small compared to both the process' variability and the tolerance width. For a system that measures items over a range of values, the worst-case variability of the system must be small when compared to either the process' variability or the tolerance width. Acceptable measurement systems will "consume" less than 10-15% of the process variation and/or tolerance. A measurement system which consumes more than 30% of the process variation and/or tolerance is generally not acceptable (note – this must be evaluated on a case-by-case basis. For example, some chemical measurement systems employed in the pharmaceutical industry may consume over 50% of the process variation and still be considered "acceptable.").

Comparing Measurement System Variation – Process or Tolerances – The above discussion compared the measurement system to the process variation. You may also be interested in how the measurement system compares to the overall process specifications (as expressed in a tolerance width).

6.7.3 Measurement System Studies

Objectives

The key objective of a measurement system study is:

To Determine the Amount and Type of Measurement Error

Specific study objectives will be associated with the different elements of measurement system error:

- ***Resolution/Discrimination***
- ***Stability***
- ***Repeatability***
- ***Reproducibility***
- ***Bias***
- ***Linearity***

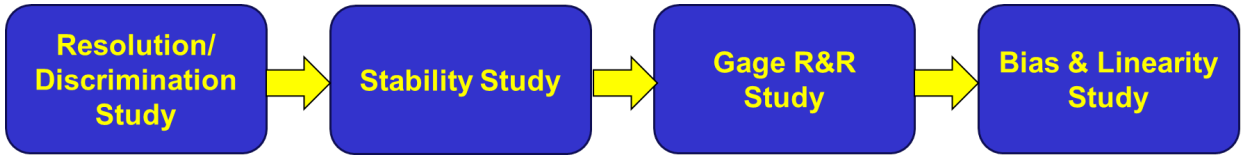
When a measurement system is being considered for use in a process, the early measurement system studies will focus on the suitability of the system to perform its function – e.g. does the measurement system possess the required *discrimination/resolution* and *linearity* properties? Once the system is judged to be capable of performing its function, additional, periodic studies will be performed to ensure that the system remains capable – e.g. gauge calibration to assess *bias*, Gage R&R studies to assess *repeatability* and *reproducibility.*

General questions to address in planning the study include:

- What approach should be used?
- Is a standard or reference value required (to assess bias)?
- How many sample parts are to be tested?
- How many operators are to be involved?
- How many repeat readings will be needed?

Procedures

The AIAG *Measurement Systems Analysis Reference Manual* is a good reference for general metrology procedures. Donald Wheeler's EMP III: Evaluating the Measurement Process & Using Imperfect Data is also an excellent presentation on how to use control charts to assess the measurement system. The general process involves four procedures, each with its specific purpose:



Step	Study	Purpose
1	Resolution / Discrimination	Can the measurement system see small enough differences to be useful for this tolerance or process variation? Are there enough decimal places vs tolerance?
2	Stability	If nothing about the part changes, does the gage slowly drift, jump, or stay stable? Is the gage consistent over time?
3	Gage R&R	"Is this measurement system precise enough to be useful at all?"
4	Bias & Linearity	Is it accurate across the range? Bias addresses vertical offset, linearity addresses constant slope.

"Typical" Gage Repeatability & Reproducibility Study (Gage R&R)

One study often performed to judge the quality of a measurement system is Gage Repeatability & Reproducibility, or simply the Gage R&R study. The results of this study will tell you how much of the variation in the process is due to the measurement system (overall) and to its components (i.e. the repeatability of the operator and the reproducibility across operators). If you include the specification limits in the analysis, then you will also learn what fraction of the tolerance is "consumed" by the measurement system. You can then determine whether the existing measurement system is adequate for your purposes (i.e. go/no go inspection, statistical process control) or whether the measurement system

requires improvement. Understanding which of the components (repeatability or reproducibility) is dominant, you will have diagnosed which part of the measurement system needs improvement.

The following provides a "typical" procedure for performing a Gage R&R study:

1. Identify the gage (instrument) to be assessed.
2. Identify the operators who will measure the parts (Joe, Sally, Harry)
3. Obtain parts from the production process (try to get parts whose characteristic (i.e. dimension) spans the width of the process variation).
4. *Check* – To make sure you have enough data for the analysis, the product of #operators (O) times the #parts (P) is greater than 15 (O x P > 15).
5. Obtain the specification limits for the process characteristic.
6. Have the operators measure the parts. Arrange the measurement order so that they don't know which part is being measured (i.e. randomize the order of the measurements). Ensure that each operator measures each part at least twice. If it is not practical to meet the O x P criteria above, you will need to increase the number of times the operators measure the parts.
7. Conduct ANOVA (Unit 10.3) to understand the total variation, and the components (parts, repeatability, reproducibility, error). Most quality statistical packages (Minitab) have special routines to provide you with both graphical analysis of the study and analytic (ANOVA, variance, standard deviation components).
8. Analyze the results and determine your course of action (measurement system OK, needs improvement).

Gage R&R Setup and Measurements

Here, three operators measured 10 parts twice. Note that OxP = 3 x 10 = 30.

Part	Operator	Response	Part	Operator	Response	Part	Operator	Response
1	Henry	0.65	1	Beth	0.55	1	Marilyn	0.50
1	Henry	0.60	1	Beth	0.55	1	Marilyn	0.55
2	Henry	1.00	2	Beth	1.05	2	Marilyn	1.05
2	Henry	1.00	2	Beth	0.95	2	Marilyn	1.00
3	Henry	0.85	3	Beth	0.80	3	Marilyn	0.80
3	Henry	0.80	3	Beth	0.75	3	Marilyn	0.80
4	Henry	0.85	4	Beth	0.80	4	Marilyn	0.80
4	Henry	0.95	4	Beth	0.75	4	Marilyn	0.80
5	Henry	0.55	5	Beth	0.40	5	Marilyn	0.45
5	Henry	0.45	5	Beth	0.40	5	Marilyn	0.50
6	Henry	1.00	6	Beth	1.00	6	Marilyn	1.00
6	Henry	1.00	6	Beth	1.05	6	Marilyn	1.05
7	Henry	0.95	7	Beth	0.95	7	Marilyn	0.95
7	Henry	0.95	7	Beth	0.90	7	Marilyn	0.95
8	Henry	0.85	8	Beth	0.75	8	Marilyn	0.80
8	Henry	0.80	8	Beth	0.70	8	Marilyn	0.80
9	Henry	1.00	9	Beth	1.00	9	Marilyn	1.05
9	Henry	1.00	9	Beth	0.95	9	Marilyn	1.05
10	Henry	0.60	10	Beth	0.55	10	Marilyn	0.85
10	Henry	0.70	10	Beth	0.50	10	Marilyn	0.80

Gage R&R Results – Example – Minitab Software Analysis

```
Gage name:      ExactoMeas - Dimensions
Date of study:  2/14/03
Reported by:    HDUNN
Tolerance:
Misc:
```

Description of Gage, Study "Demographics"

Two-Way ANOVA Table With Interaction

Source	DF	SS	MS	F	P
Part	9	2.05871	0.228745	39.7178	0.00000
Operator	2	0.04800	0.024000	4.1672	0.03256
Operator*Part	18	0.10367	0.005759	4.4588	0.00016
Repeatability	30	0.03875	0.001292		
Total	59	2.24912			

Analysis of Variance (ANOVA)
This analysis shows that the parts are a significant source of variation (good), there is a difference by operator and there is an interaction between the operator factor and the parts factor

Gage R&R

Source	VarComp	%Contribution (of VarComp)
Total Gage R&R	0.004437	10.67
Repeatability	0.001292	3.10
Reproducibility	0.003146	7.56
Operator	0.000912	2.19
Operator*Part	0.002234	5.37
Part-To-Part	0.037164	89.33
Total Variation	0.041602	100.00

Contribution to Variance
*Shows breakdown of **variance** into components – here, the measurement system represents about 11% of the total variance. Of this, 7% is due to reproducibility issues, 3% to repeatability (within the operator).*

Source	StdDev (SD)	Study Var (5.15*SD)	%Study Var (%SV)	%Tolerance (SV/Toler)
Total Gage R&R	0.066615	0.34306	32.66	34.31
Repeatability	0.035940	0.18509	17.62	18.51
Reproducibility	0.056088	0.28885	27.50	28.89
Operator	0.030200	0.15553	14.81	15.55
Operator*Part	0.047263	0.24340	23.17	24.34
Part-To-Part	0.192781	0.99282	94.52	99.28
Total Variation	0.203965	1.05042	100.00	105.04

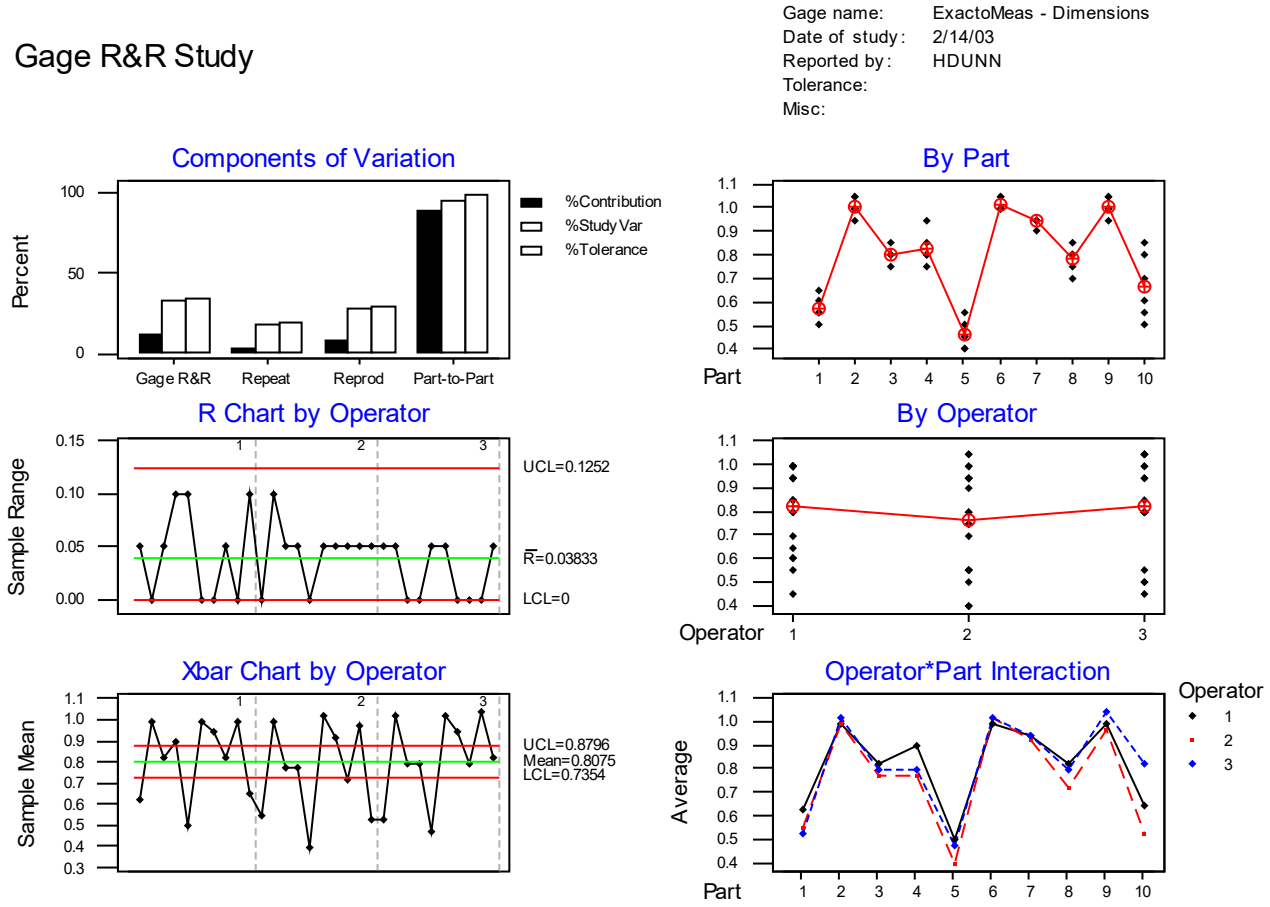
Contribution to Standard Deviation, Tolerance
*Shows breakdown of **standard deviation** and percentages relative to the total variation and to the tolerance width.*

```
Number of Distinct Categories = 4
```

Assesses the ***resolution*** of the system – 4 or higher is "adequate"

Graphical Output (Minitab)

The equivalent graphical output appears below. *Components of Variation* shows the contributions of the parts, gage and the "R&R" component. The *R (Range) Chart by Operator* should be in control, indicating the measurement variation is due to common causes. The *X-Bar Chart by Operator* should have most of the points *outside* the limits – indicating that the variation is due to the parts (the UCL/LCL are based on the average range of measurements by operator). The *By Part* graph shows how the operators measured each part. The *By Operator* chart shows consistency across operators and the *Operator*Part Interaction* chart shows if the operator's measurements depend on the part being measured.



Interpreting Measurement System Analyses

Diagnostics

High Bias – If there are large differences between the true measurement (standard or reference value) and the observed averages, look for these possible causes:

- Error in the master
- Worn Gauge
- Gauge made to wrong dimension
- Gauge is measuring wrong characteristic
- Gauge is not calibrated properly
- Gauge is being used improperly by the operator

Inadequate Repeatability – Generally a control chart will be prepared from the data. If the R chart is out of control, this indicates a problem with the *consistency* of the measurement process. Look for causes here.

Inadequate Reproducibility – Here, several operators will be compared. Again, the data will be presented on a control chart and the R chart is of interest. If one operator's R chart displays assignable causes, then investigate the measurement method used by that operator. If all operators' R charts display assignable causes, then the measuring instrument is sensitive to the operators' techniques and should be analyzed.

Lack of Stability – One strategy to assess a measurement system's stability is to perform measurements against some known standard or reference value at given intervals. These measurements can then be plotted on a control chart and analyzed for assignable cause signals. If none are present, then the system may be said to exhibit stability, as usual, those assignable causes should be investigated and corrected. Possible actions to take include calibration of the instrument. Other causes may include temperature variations that affect the instrument, a dirty or damaged standard/reference or the "usual" measurement process variables. One note regarding stability is important. If the control chart indicates stability, then re-calibration of the instrument can *introduce* variability – e.g. we are tampering with the system.

Poor Linearity – The results of linearity studies will be displayed on a scatter diagram and perhaps a regression line fitted to the data. The ideal state occurs when there is a one-to-one increase in the measured value as the reference value is

increased. Problems exist when the measurement system diverges from this, and when the scatter diagram appears to behave non-linearly. Causes to investigate include:

- Instrument not calibrated properly at lower and upper ends of the operating range,
- Error in the minimum or maximum master/reference value
- Worn instrument
- Instrument design problems

Part-to-Part Variation – Often several different parts are obtained over a period of time to capture the expected process variation. These parts' measurements are displayed on a control chart (e.g. X-Bar, R). Here, each data point represents repeated measurements of one part by one operator. The X-Bar chart is of interest in determining the measurement system's adequacy. In this case, we *expect* to see a large fraction of the data points fall *outside* the X-Bar chart's control limits, since the subgrouping strategy has focused on the *repeatability* and *reproducibility* of measurements. If this does not occur, then the measurement system's variation is hiding the process' variation and the system is likely to be inadequate. Note also, that when several operators are employed in the study, they should agree on which parts are outside of the control limits.

Acceptance Criteria

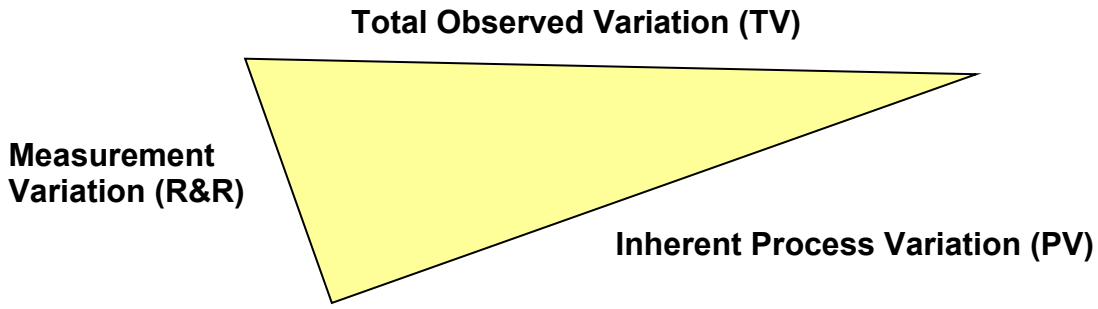
The basic issue to address is what fraction of the overall process variation or tolerance is being "consumed" by the measurement system? The AIAG manual provides calculation methods that allow you to separate the measurement system and process variation from the overall, observed variation.

The general criteria for acceptance of gauge repeatability and reproducibility are:

- *Under 15%* - Acceptable
- *16% to 30%* - May be acceptable based on importance of the application, gauge cost, cost of repairs, etc.
- *Over 30%* - Generally not acceptable. Take action to identify and correct the cause.

Process Capability and Measurement System Variation

When a process' capability is marginal, the measurement system variation should be "backed" out of the analysis. The additivity of variances principle can be employed to accomplish this:



$$TV^2 = PV^2 + R\&R^2 \quad (Additivity\ of\ Variances)$$

$$PV = \sqrt{TV^2 - R\&R^2}$$

6.8 Process Capability Analysis

Learning Objectives

- Explain Process Capability
- Calculate Process Capability Indices (Count Data, Cp, Cpk, Quality Loss Function, Sigma)
- Determine Stability and Capability
- Determine if improvement has occurred

Unit Contents

- Process Capability
- A Picture of Process Capability
- Measures of Process Capability
- Process Capability Studies
- Capability and Six Sigma
- Some Notes on Capability Calculations
- The Difference Between *Capa-* and *Sta-*bility

6.8.1 Process Capability

Run and *Control Charts* help you determine if your process is being influenced by common causes alone or by a combination of common and special causes - the “Voice of the Process.”

Up until now, though, we haven't talked much about the customer and what they would like to see from the products and/or services made by this process - the “Voice of the Customer.”

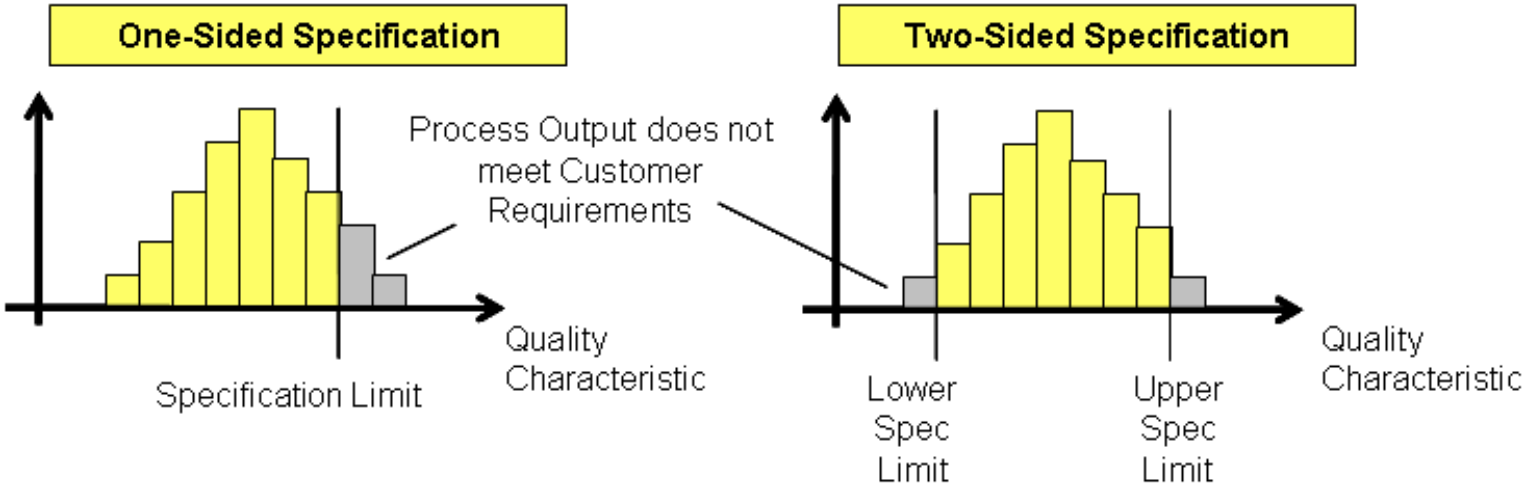
Process Capability brings us back to this central issue of quality management. The concept of Process Capability is simple:

“How well does our process' output (product or service) meet the valid requirements of the customer?”

Process Capability is not only a concept; it is also something we can measure. We’ll explore one way of graphically depicting Capability and three *measures* of Capability. Process Capability is one of the crucial links between **process control** and **process improvement.**

6.8.2 A Picture of Process Capability

The control chart is a good tool to determine the *stability* of our processes. On the other hand, the Histogram is the best tool to examine the *capability* of our processes, especially when we are dealing with measurement data. Here's the picture:



Through our knowledge of the customer, we have set a target or specification limit(s) for the individual outputs of our process. We see from this picture that some fraction of our process' output does not meet our customer's expectations. Sometimes, our customers have requirements that set both an upper and lower specification on our process (Papa Bear's porridge was too hot, Mama Bear's was too cold, but Baby Bear's porridge was just right!). In this situation, our process can fail in two ways to produce an output that meets customer requirements. The distance between these two specifications is known as the customer's *tolerance*. Let's turn these pictures into a measure that is called the Process Capability Index.

6.8.3 Measures of Process Capability

There are several measures of process capability. We will provide you with three that you can use to calculate a process capability index.

Inherent Process Capability Index – C_P

The first measure is called the *Inherent* Process Capability Index and is given the symbol "C_p." This index is a ratio of the customer's *tolerance* (i.e. maximum variation that the customer is currently willing to accept) to the process' *dispersion* (i.e. variation that our process is currently producing). If our customer cares about both an upper and lower specification limit for the outputs of our process, then the *tolerance* is simply the difference between these two values:

Customer's Tolerance = Upper Specification Limit - Lower Specification Limit

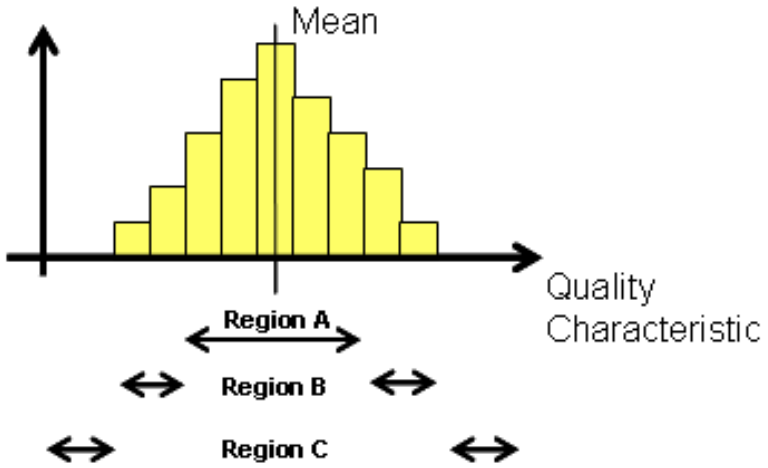
Previously, we mentioned three measures of *variability*: Range, Variance and the Standard Deviation. For this process capability index, we are going to use a special "flavor" of the Standard Deviation and will use a value of 6 times the standard deviation as the measure of our process' dispersion. Where does the value of "6" come from?

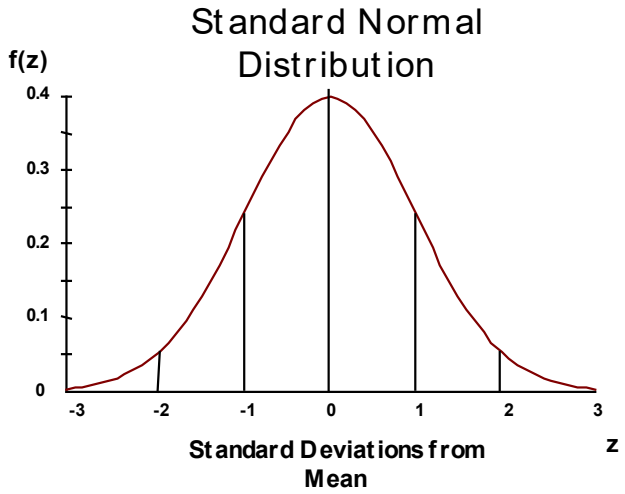
The Standard Deviation gives us a measure of the dispersion of the process. If one process has a standard deviation of 6 "arglebargles" and another a standard deviation of 3 "arglebargles," then we understand that the second process has less dispersion than the first.

The Standard Deviation can also help us in another important way. As we've seen, the individual outputs of our processes will tend to gather around some central point that we measure using the mean or the median. As we move away from this central point (on either side), the chances of finding an output from the process get smaller.

The process shown to the right is likely to produce an output in **Region A**, less likely to produce outputs in **Region B**, and highly unlikely to produce outputs in **Region C**. It turns out that we can estimate the chances of having an output of our process occur as a function of the *number of standard deviations we are away from the process mean or average.*

Look at the nice smooth, symmetrical (i.e. a normal or Gaussian or "Bell Curve") distribution that you see below. For this situation, consider a band that goes from one standard deviation to the left of the mean to one standard deviation to the right of the mean. Observing a process with this type of distribution, we would find that about 68% of the process outputs would fall into this band.





If we stretched this band to two standard deviations on either side of the mean, about 95% of our process output would be observed to fall here. Finally, stretching to three standard deviations on either side of the mean, about 99.7% of our process output would be found in this band. For this last case, only 0.3% (about 3 in 1000) of our process outputs would be found outside a band that includes plus or minus three standard deviations from the mean.

Now you can see how we chose 6 times the standard deviation as the measure of our process' dispersion. If the difference between the upper and lower specification limits of our process is the same as plus or minus three standard deviations around the mean, then only 3 in 1000 process outputs would **not** meet customer requirements. We would say that our process is “fairly” capable of meeting customer requirements.

Where to "Find" the Process Standard Deviation

If you have a sample of at least 25 data from a process, then you can simply calculate the standard deviation using the formula presented in Unit 6.3. If you have been using a control chart (one of the measurement charts: X-Bar S, X-Bar R, or XmR) then you can use the Standard Deviation or Range from the chart, with a minor conversion factor:

Standard Deviation Estimate from X-Bar, S Control Chart:

$$s = \bar{s} / c_4$$

where:

s - Process Standard Deviation Estimate

$\bar{s}$ - Average Standard Deviation (from s - Chart)

c_4 - Constant depending on Average Subgroup Size

A table of "c_4" values is given here:

n	2	3	4	5	6	7	8	9	10	11
c_4	.798	.886	.921	.940	.952	.959	.965	.969	.973	.975
n	12	13	14	15	16	17	18	19	20	>20
c_4	.978	.979	.981	.982	.983	.984	.985	.986	.987	≈1

n - Average Subgroup Size

You can see the correction factor, "c_4," is essentially equal to 1 for larger subgroup sizes - a "shop floor" process capability estimate can be obtained by taking the average standard deviation right from the X-Bar, S chart.

The process is similar for converting the Range to an estimate of process Standard Deviation. We alluded to this back in the X-Bar, R control chart discussion:

Standard Deviation Estimate from X-Bar, R or X, mR Control Charts:

$$s = \overline{R} / d_2$$

where:

s - Process Standard Deviation Estimate

$\overline{R}$ - Average Range (from R - Chart)

d_2 - Constant depending on Subgroup Size

A table of "d_2" values is given here:

n	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
d_2	1.13	1.69	2.06	2.33	2.53	2.70	2.85	2.97	3.08	3.17

n - Subgroup Size

When the X, mR chart is used, the subgroup size is obtained from the moving range and is equal to 2. Since the "d_2" value for a subgroup size of 2 is close to 1, a "shop floor" estimate of process capability may be based on setting the Process Standard Deviation equal to the Average Range.

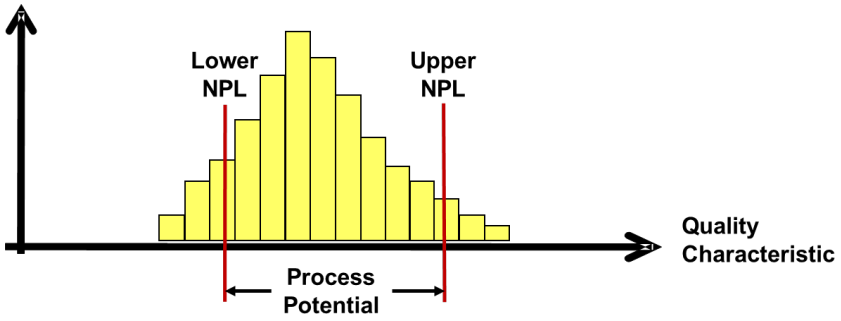
Note that the standard deviations calculated from the control chart values are different from the usual calculation in an important way – these provide us with a measure of the *within-subgroup variation*, whereas the usual calculation mixes *within* and *between-subgroup variation.* The *within-subgroup variation* represents the process's performance when it's living up to its full potential.

Natural Process Limits

If we take the Grand Average from the control chart and add/subtract 3 times the *within-subgroup* standard deviation, this provides us with Natural Process Limits – you would expect the process to deliver product within these limits a large fraction of the time if there is only within-subgroup variation present (no assignable causes!).

$$Natural\ Process\ Limits = \bar{\bar{X}} \pm 3 * \frac{\bar{R}}{d_2}$$

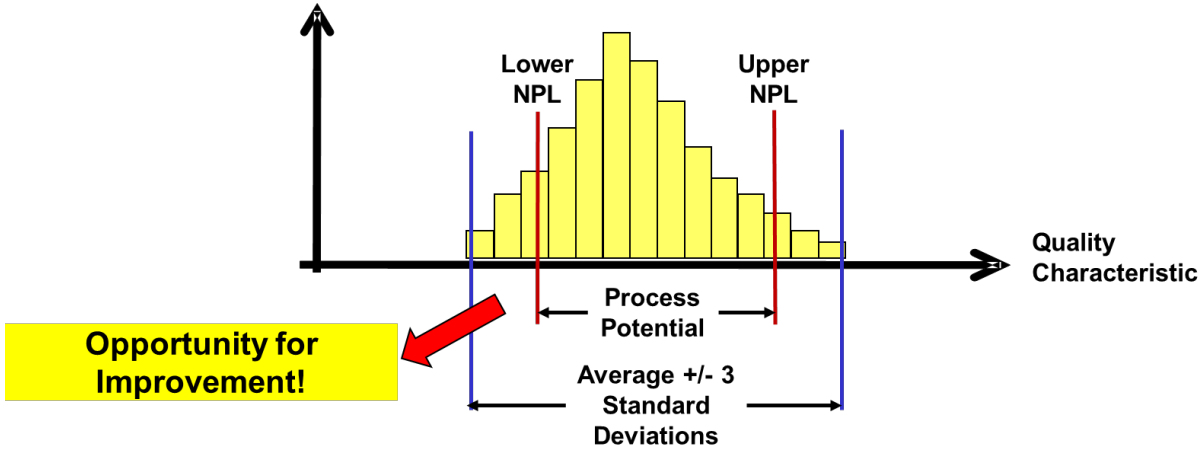
$$Natural\ Process\ Limits = \bar{\bar{X}} \pm 3 * Sigma(X)$$



Notice the difference between the Natural Process Limits and those found by taking the Grand Average plus/minus three times the "usual" standard deviation (sometimes called the "Global Standard Deviation"):

$$Natural\ Process\ Limits = \bar{\bar{X}} \pm 3 * \frac{\bar{R}}{d_2}$$

$$3\ std.dev.\ Limits = \bar{\bar{X}} \pm 3 * s$$



Comparing the two illustrates the opportunity for improvement – if we can eliminate assignable causes, this is the process living up to its full potential.

With the preceding introduction, the Inherent Process Capability Index is now easily defined:

$$\textit{Inherent Process Capability} = \frac{\textit{Upper Specification - Lower Specification}}{\textit{6 x Process Standard Deviation}}$$

If our process has only one specification (an upper or lower), we simply calculate the inherent process capability like this:

$$\textit{Inherent Process Capability (One Spec Only)} = \frac{|\textit{Specification - Process Mean}|}{\textit{3 x Process Standard Deviation}}$$

Note that the absolute value of the specification-process mean difference is taken for the one-sided spec case. You can see that the index is essentially a ratio of "distances" - the numerator is the "distance" (tolerance) of the customer, the denominator is the "distance" (variation) of our process.

Watershed Specifications

Specifications are a bit tricky. The customer may tell us they want the process centered at 100 with specifications at 95 and 105 (+/- 5). To operationalize these requirements, we must consider the measurement system's resolution – e.g., the value to which the product/parts are measured. Suppose our measurement system's resolution is "1." The possible measurements inside the 95-105 range are 96, 97, 98, 99, 100, 101, 102, 103, and 104 – only nine successful measurements when there really should be eleven (we need to allow 95 and 105).

To address this, we will subtract half of 1 from the lower specification to determine the *lower watershed specification (95 – ½ = 94.5)* and add half of 1 to the upper specification to determine the *upper watershed specification (105 + ½ = 105.5)*. Parts whose values are between 95 and 110 can be considered in-spec and there are now 11 possible successful measurements.

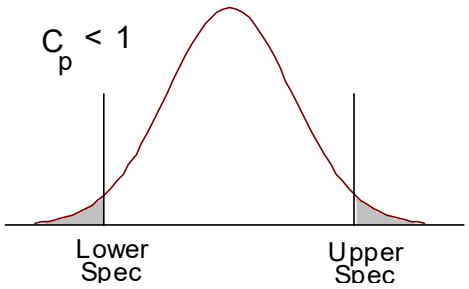
The watershed values account for the reality (here) of 11 acceptable values.

Values of Process Capability

When we first heard about process capability (mid-1980's), a process capability index of 1.33 was very good. Now, we read reports of process capabilities more than 2 or 3. For example, the *Six Sigma* approach developed at Motorola aims for a process capability of 2. If you reach this state for one of your processes, then you're probably among the world-class performers for this process and should probably consider picking another process to improve.

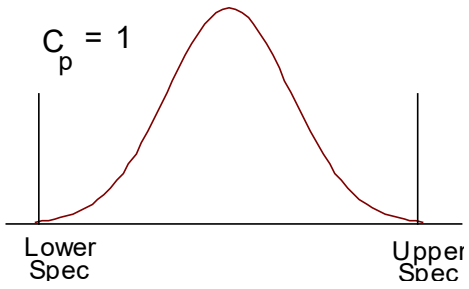
Ratio < 1

If this ratio for our process is less than 1, then that means that a significant fraction of our process output does not meet requirements, i.e. our process is not very capable.



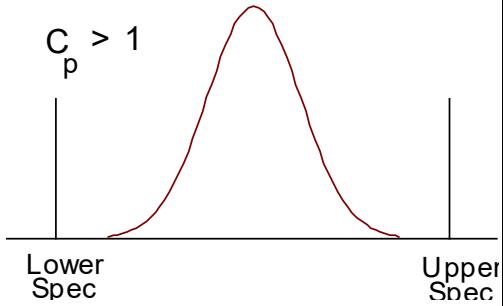
Ratio = 1

If the Process Capability Index is equal to 1, then, as we said above, only about 3 in 1000 process outputs do not meet customer requirements.



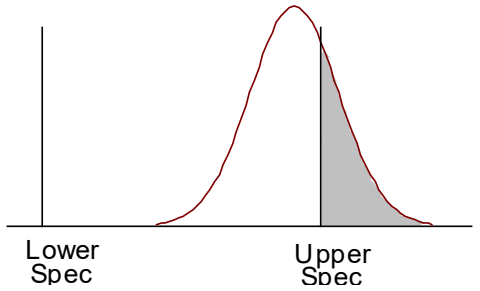
Ratio > 1

Finally, if the Process Capability Index is greater than 1, then there is a very small chance of our process producing an output that does not meet our customers' requirements.



Process Not Centered

Here, the inherent capability is good, but the picture shows that the process is producing defects due to centering problems.



The Operational Process Capability - C_{pk}

In the *Process Not Centered* situation, the *inherent* process capability index was greater than one, but there is a large fraction of the process being produced outside the upper specification. The process is said to have an *inherent* capability; if we could shift the process' mean and center it between the specification limits, then the process would truly be capable. This issue reinforces the Second Law of Statistics - "Draw the Picture." If a tabular report of *inherent* process capabilities came to your desk, you would not know which processes were centered and which were not.

The "centering" difficulty with the *inherent* process capability index leads us to the Operational Process Capability Index (C_{pk}). This measure is easy to calculate, and it handles the situation where the process is not centered between the spec limits. We will use a technique called the "Z-min Method" to calculate the Operational Process Capability (C_{pk}).

For a one-sided tolerance (only upper or lower specification limit), first calculate Z-min:

$$Z_{\min} = \left| \frac{SL - \overline{\overline{X}}}{s} \right|$$

where:

$Z_{\min}$ - number of standard deviations the process mean is from the specification limit

SL - Upper or Lower Specification Limit

$\overline{\overline{X}}$ - Process Average

(an X - Bar chart is assumed to be available here)

s - Process Standard Deviation

(estimated from $\overline{s} / c_4$ or $\overline{R} / d_2$)

For a two-sided tolerance interval (both upper and lower specifications), calculate Z-min as follows:

$$Z_{USL} = \frac{USL - \overline{\overline{X}}}{s} \qquad Z_{LSL} = \frac{\overline{\overline{X}} - LSL}{s}$$

and

$$Z_{\min} = Minimum(Z_{USL}, Z_{LSL})$$

where:

$Z_{\min}$ - minimum number of standard deviations the process mean is from a specification limit

USL - Upper Specification Limit

LSL - Lower Specification Limit

$\overline{\overline{X}}$ - Process Average

(an X - Bar chart is assumed to be available here)

s - Process Standard Deviation

(estimated from $\overline{s} / c_4$ or $\overline{R} / d_2$)

The Operational Process Capability Index is then calculated as follows:

$$C_{pk} = Z_{\min} / 3$$

where:

C_{pk} - Operational Process Capability

Note how this index compares to the Inherent Process Capability Index. If the Z-min value is 3.0, then the distance from the mean to the closest specification limit is 3 process standard deviations and the Operational Process Capability index is 1.0. This is the same result we would get if we calculated an Inherent Process Capability index for a one-sided specification limit. Although we can't say that less than 3 in 1000 items will be produced within the spec limits (because we've "ignored" the distance from the mean to the farther spec limit), the Operational Process Capability is similar in concept to the Inherent Process Capability index. Some organizations report a process capability as compared to the Upper Specification limit and the Lower Specification Limit. The Z-min formulae shown above (and divided by three) can be employed to calculate these capability indices.

6.8.4 Process Capability Studies

Process Capability Studies are a long-term evaluation of a total process. Data is collected periodically in subgroups of consecutive pieces. In many cases, it is convenient to use the data gathered for statistical process control purposes. The capability study will include the effect of all sources of variation:

- Raw material sources,
- Operators,
- Gauge users,
- Production rates, and
- Environmental conditions.

Also included in the total process variation will be the effect of the process control method. For example, we could be controlling the process with a control chart using subgroups of three. This will permit the process average to wander more without a high probability of detection than if we were using a subgroup of five or six. The process capability assessment estimates the variability of dimensions or characteristics and compares this variation to the specification in terms of Process Performance Index (P_{pk}). If the Process Performance Index is greater than or equal to 1.33, the specification will be satisfied. The larger the ratio, the less we are using of the available tolerance. For Process Performance Index, we use the computation:

$$P_{pk} = \frac{\left(USL - \bar{\bar{x}}\right)}{3\hat{\sigma}} \quad \text{or} \quad \frac{\left(\bar{\bar{x}} - LSL\right)}{3\hat{\sigma}_x}$$

The Process Performance Index considers both the special and common cause sources of variation over time. The variability of the processes and how well the process is centered are both considered. We use the minimum value to determine our index.

To compute P_{pk}, we should use the estimate of standard deviation based on individuals rather than average range. P_{pk} is based on standard deviation of individual measurements and estimates the total variation of the total process. This is the best estimate of what is coming from the process and being shipped to the assembly line or customer.

P_{pk} is computed with the standard deviation based on average range predicts the potential of the process based on existing common cause variation.

Steps for Assessing Process Capability (PC)

1. Accumulate data in subgroups of consecutive parts taken periodically from production runs. If the same process produces a variety of part numbers with different target dimensions, each different part should be treated as a separate process unless we have evidence that the variation about the targeted dimension is not affected by the normal value.

   For example, if we are grinding shafts and from previous machine capability studies we know that the inherent variation of the machine is a function of the nominal dimension, such that a 1.000" shaft has a standard deviation of 0.0002" but a 1.250" nominal shaft has a standard deviation of 0.0004", we must treat these two shafts as two distinct processes each with their own PC. However, if the machine standard deviation is unchanged regardless of which nominal diameter is ground, the overall PC for the entire family of shafts can be assessed. The easy way to do this is to record the data as a deviation from nominal dimensions. In other words, if we are processing shafts with a nominal print dimension of 1.000", a shaft measuring 1.002" would be recorded as a "plus .002". If another shaft had a nominal or targeted dimension of 1.200" and it measured 1.202", it too would be a "plus .002".

2. Data should be accumulated over a long enough period that all sources of variation have an opportunity to be exhibited (25 subgroups of 4 taken over a "long" period is a guide). Check for process stability.

3. Test the data for shape of distribution.

4. If the distribution is normal, compute the standard deviation based on individuals. If the data is not normal, either transform the data or perform a capability analysis using the Weibull distribution (see Minitab, Help topics, *process capability: non-normal data* for more information).

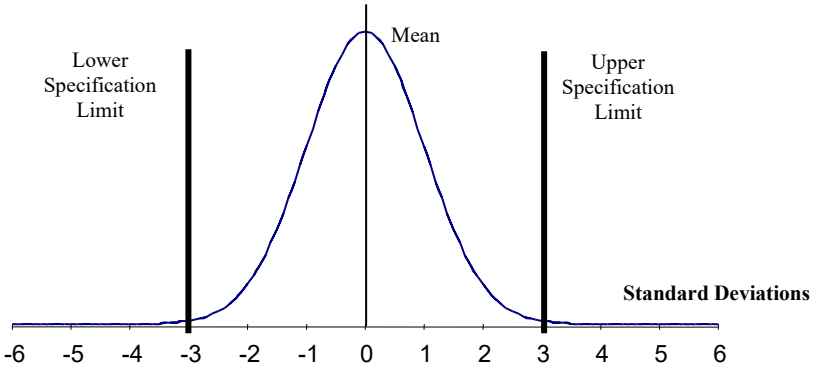
5. Calculate PC based on six standard deviations divided into the available tolerance.

6. Calculate PC using standard deviation based on the average range. Compare this with the value obtained in step 5 to see what is the potential for process improvement.

6.8.5 Capability and Six Sigma

Sigma is a measure of process capability that builds on the preceding discussions. Remember that the process' *distribution* is characterized by three main elements:

1) the *center* of the data (as measured by a *mean, median*, or *mode*),
2) the *spread* of the data (as measured by a *range, standard deviation,* or *variance*), and
3) the *shape* of the data (characterized by some mathematical function that best fits the data, such as the *exponential, normal, log-normal, Weibull,* or *Raleigh*).



For a given distribution, the *distance* from the mean to the specification limits expressed as *standard deviation* multiples will correspond to a given fraction defective. For the normal distribution, about 34% of a process' output will fall outside *one* standard deviation (above and below) the mean and only 3.4 parts per million fall outside 4.5 standard deviations. Since *sigma* (σ) is statistical shorthand for the standard deviation, we can express the capability of a process in terms of the *number of "sigmas."* The higher the *sigma* of a process, the lower the fraction defective, the more *capable* the process is of meeting its specifications.

Shifts and Drifts

We stated that only 3.4 parts per million (ppm) falls outside 4.5 standard deviations "units" from the mean. "Six Sigma," though, is associated with the fraction defective of 3.4 ppm. What's the difference? The answer lies in how data is collected from a process and the variation inherent in the process.

Often, data will be collected over a "short" time span. During this time, some, but not all, sources of variation will be present in the process. In a manufacturing process, a variable such as *tool wear* will generally not be seen in a short time. For a service process, there may be seasonal variation. Studies of process variation indicate that the process may *shift and drift* by up to 1.8 standard deviations over the long term. 1.5 standard deviations is a typical, widely used value of shift, although each process owner is encouraged to understand their process' "shifts and drifts".

If the *short-term* data shows that we are 6 sigma multiples from the specification limits, we will *subtract* a 1.5 sigma shift and state that the *long term* capability of the process is 6 – 1.5 = 4.5 sigma. This corresponds to a long-term fraction defective of 3.4 ppm. See the ***Additional Topics*** section below.

Measuring Capability With Sigmas

The following procedures show how to calculate the capability of a process in *sigma* units for a single characteristic of the process.

Basic Method for Continuous Variables

If the data is continuous, perform these calculations to estimate the characteristic's *sigma.*

Inputs – Upper/ Lower Specification Limits, a minimum of 25 –30 data points from the process.

Process

1. Create a histogram of the data. Does the data "look" normal (or roughly so)? If yes, proceed to step 2. If not, see **Additional Topics** below.

2. Calculate the mean and standard deviation of the data:

$$\text{Mean}: \bar{x} = \sum_{i=1}^{n} x_i / n$$

$$\text{Standard Deviation}: s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

where :

$x_i - data\ values$

$n - number\ of\ data$

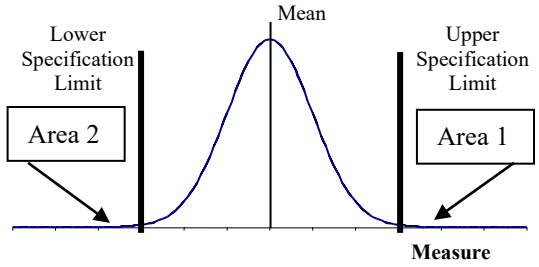
Note: in Excel, the AVERAGE and STDEV functions perform the above calculations.

3. Find the area under the normal distribution curve to the right and left of the Upper (USL) and Lower (LSL) Specification Limits (respectively):

$$Area - 1 = 1 - CumNorm\left(\frac{USL - \bar{x}}{s}\right)$$

$$Area - 2 = CumNorm\left(\frac{LSL - \bar{x}}{s}\right)$$

where : $CumNorm -$ Cumulative Normal Distribution



Notes:
a) Some processes may have only one specification limit. Here, calculate only the appropriate area outside the spec limit.
b) In Excel, the function NORMDIST(Value, 0, 1, TRUE) provides the cumulative normal distribution.

4. Add the two areas together (Total Area = Area-1 + Area-2) and calculate the Process Yield:

$$Yield = (1 - Total\ Area) \times 100\%$$

5. Look up Process Sigma in Table One (see Appendix B).

Basic Method for Discrete Variables

If the data is *discrete,* perform these calculations to estimate the characteristic's *sigma:*

Inputs – Number of Units (*n*), Number of Defects (*d*), Number of Defect Opportunities (*o*).

Process
1. Calculate the Defects per Million Opportunities (DPMO):

$$DPMO = \frac{d}{n \times o} \times 10^6$$

2. Look up Process Sigma in Table One (see Appendix B).

Notes:
a) Both the Number of Defects and *Non*-Defects (*n* – *d*) should be greater than 5. This ensures that the DPMO/Sigma conversion of Table One, which is based on the normal distribution, is valid.
b) The Number of Defects produced should include both those that were detected prior to receipt by the customer and those defects reaching the customer.
c) A Defect Opportunity is any event which can be measured and where there is a chance of not meeting a customer requirement. Many processes have multiple opportunities to create defects.

Calculating the Yield of a Process

Many products and services are "produced" through a series of processes, where each process has *opportunities* to create defects. Here, the producing organization has generally assigned accountability for the different processes to different levels of management (*Note: the organization may still want to adopt the process owner concept – an individual or committee that "owns" an end-to-end process, such as derivatives trading*). One more factor comes into play here – for what *purpose* are the capability measurements being taken? Does the purpose involve reporting to external agencies (i.e. stockholders), or is the purpose to prioritize improvement opportunities? Three capability measures are used to address these issues: *1) First Pass Yield, 2) Normalized Yield, and 3) Rolled Throughput Yield.*

First Pass Yield (Y_{FP})

First Pass Yield (Y_{FP}) is the fraction of units produced by a sub-process without a defect, considering all the defect opportunities and including all defects, whether or not they are detected prior to reaching the customer:

$$Y_{FP} = 1 - \frac{d}{n \times o}$$

For a given sub-process, then, the defect opportunities (*o*) are defined (based on customer needs & requirements), a given number of units (*n*) are chosen, and the total number of defects (*d*) counted. For discrete attributes, the number of opportunities where a "Pass/Fail" judgement was made is counted, for continuous attributes, the number of opportunities where the specification limit(s) were exceeded is counted. These are summed up to obtain the total number of defects. First Pass Yield will be used to calculate process sigmas, as shown below. This is consistent with the general Six Sigma philosophy of focusing on total quality costs and that the payback to the company is generally bigger to keep defects from occurring in the first place. For some reporting, though, a *Final Pass Yield* may be useful. This is simply the Yield of the process *after* detected defects have been reworked or otherwise corrected.

$$Y_{FINALPASS} = 1 - \frac{d - d'}{n \times o}$$

$$\text{where : } d' \text{ - number of defects detected and eliminated prior to reaching the customer}$$

Normalized Yield (Y_{NORM})

Normalized Yield (Y_{NORM}) is a "rolled up" weighted average of the sub-process First Pass Yields for an end-to-end process. This measure permits comparisons across different business processes and across processes of varying complexity. It is calculated as follows:

$$Y_{NORM} = 1 - \frac{\sum_i d_i}{n \times \sum_i o_i}$$

$$\text{where : } i \text{ - number of subprocesses}$$

A "rolled up" process sigma can be calculated from Normalized Yield (convert the yield value to a percentage and use Table One to obtain the *short-term sigma*). Y_{NORM} and associated sigma are used to communicate process capability *externally* and for *benchmarking* purposes.

Rolled-Throughput Yield (Y_{RTP})

Rolled-Throughput Yield (Y_{RTP}) is the probability of a "unit" going through all the processes without a defect. Y_{RTP} is the *product* of the First Pass Yields of each sub-process:

$$Y_{RTP} = \prod_i Y_{FPY-i}$$

Rolled-Throughput Yield is generally used for *internal monitoring of business processes* and for *prioritizing improvement projects.* It is considered the best measure of the effectiveness and efficiency of business processes. This measure is not converted to a sigma value generally because it results in a negative sigma.

Additional Topics

Process *Stability* vs. *Capability*

A process is *stable* if only *common-cause* sources of variation are present and a process that is ***not*** stable does not have a reliable *capability.* A *theoretical* capability may be calculated by removing the special cause data from the sample. *Run* or *control charts* are the best methods of determining if the process is stable.

Long-Term, Short Term

Continuous Attributes – Practically, the capability of continuous attributes such as time, cost, length, etc. can be calculated with "small" amounts of data (25 – 30 data). For a high-volume process, this data is then often representative of its *short-term* capability.

Discrete Attributes – When defects are "plentiful," the data required to assess capability may be considered as short-term (recall that at least 5 defects should be detected). When defects are "rare," the number of units produced may be large enough to consider the data long-term.

Reporting - The short-term capability includes the effects of the common or random sources of variation and is considered to reflect the *inherent* capability of the process, whereas the long-term capability includes the additional impact of assignable causes of variation - factors which influence the process from time-to-time. Short Term capability is generally reported.

Non-Normal Data

The sigma and yield relationships shown on Table One are based on the normal distribution. If the data is *non-normal* and the process characteristic is continuous, then an alternate method of calculating sigma must be used. In some cases, the data can be *transformed into* a normal distribution (e.g. time to complete a process is often skewed, this data may be transformed via logarithms to a normal distribution). *Note: Some misinterpret this issue to state that only normally distributed data can be used to calculate a sigma – this is* ***not*** *the case.*

Continuous vs. Discrete Data Measurement

Often, the customer desires that some target value be consistently achieved, although they are willing to *tolerate* some variation from the target. The simplest example is that of a plane schedule, where the need is to depart and arrive *on time.* The customer may have some *tolerance* for variability in this process; for example, arrival within 10 minutes of the schedule may be tolerable. 10 minutes, then, is the upper specification limit for the CTQ of arrival time.

Given this situation, it is easy for a company to shift its focus from *on time performance* to just *meeting specifications* - i.e. if the plane arrives before the 10 minute spec limit, then things are OK. Although process owners may become defensive if we focus on their performance *inside* the spec limits, this represents a "goalpost" mentality. Taguchi's *Quality Loss Function* challenges this by postulating that any departure from the target causes some *loss* to the customer. If the plane arrives 30 minutes late, each passenger will have suffered some loss (e.g. the value of 30 minutes' work at the office). In his model, the specification limit is merely the point at which deviation from the target becomes *unacceptable* to the customer.

The Six Sigma approach does incorporate Taguchi's thinking through its focus on variation reduction, but the "goalpost" mentality can creep in through the way process data is collected. In the arrival example, we could either record actual arrival times (*continuous data*) and compare them to the target and spec limits or simply record the number of times we failed to meet the specification limits (*discrete data*). Consider the following data:

Airline A

Flight	12/1	12/2	**12/3**	12/4	12/5	12/6	12/7	12/8	12/9	12/10
Δ	+9	+4	**+14**	+7	+2	+6	+9	+8	+3	+7

Airline B

Flight	12/1	12/2	12/3	12/4	12/5	12/6	12/7	12/8	**12/9**	12/10
Δ	0	0	0	+1	0	0	+2	0	**+12**	+1

Δ = Actual Departure Time - Scheduled Departure Time

Both airlines have a defect rate of 10% - one flight of ten left beyond the 10-minute departure spec limit. Inspection of the "Δ's," though, reveals that airline B typically departs close to schedule and that the variability of departure times is much less for B than A. Also, since Airline A's performance is often close to the spec limit, we would expect to see defects produced from their process. The 12/9 Δ of 12 minutes for Airline B appears to be a special instance, one not to be ordinarily expected from their process. To summarize, the *continuous data* gives us a better picture of the performance of the process than the *discrete data*. For reporting purposes, we may choose to use the defect rate since it provides a quick picture of performance; for analysis, the continuous data displayed on a histogram with associated targets and spec limits is preferable.

Process Sigma Calculation Example

Here, we'll examine a "simple" transaction process (such as selling a client an air conditioner), develop sigma/yield estimates for individual characteristics and the overall process. The three steps of the trading process are:



For simplicity, we will consider transactions involving existing products for existing clients. The *deal* is the unit produced; a few of the opportunities to create defects are listed below by process:

Process Step	Defect Opportunities	
Market Deal	• Misunderstand client requirements • Recording errors	• Inaccurate Price • Deal not compliant with Regulations
Execute Transaction	• Error-caused amendments • Recording Errors	• Sales/Order Transaction Mismatch
Complete Transaction	• Order Errors • Fulfillment Errors	• Confirmation Timeliness • Client/Company Confirmation Mismatch
Overall	• Lost Sale	

The following page shows the calculations performed to estimate process sigma. Note that this data is assumed to represent the *long-term* performance of the process, but short-term sigma values are *reported*, per the discussions above. The sigma values were calculated using the Excel NORMINV(Yield, 0, 1) function, to which 1.5 is added to obtain the short-term sigma value.

Transaction Process Sigma Spreadsheet

Process Step	Defect Opportunities	Data Type	No. of Units	No. of Opp's.	Defects	Avg.	Std. Dev.	Upper Spec	DPMO	Yield	Sigma (ST)	Step Yield	Step Sigma
Market Product	Misunderstand client requirements	D	1000	1	120				120000	88.00%	2.67	93.83%	3.04
	Recording errors	D	1000	1	90				90000	91.00%	2.84		
	Inaccurate Price	D	1000	1	25				25000	97.50%	3.46		
	Deal not compliant with Regulations	D	1000	1	12				12000	98.80%	3.76		
Execute Transaction	Error-caused amendments	D	1000	1	100				100000	90.00%	2.78	93.23%	2.99
	Recording Errors	D	1000	1	58				58000	94.20%	3.07		
	Sales/Order Transaction Mismatch	D	1000	1	45				45000	95.50%	3.20		
Complete Transaction	Order Errors	D	1000	1	8				8000	99.20%	3.91	97.58%	3.47
	Fulfillment Errors	D	1000	1	22				22000	97.80%	3.51		
	Confirmation Timeliness	C	1000	1		2.5	0.7	3	237525	76.25%	2.21		
	Client/Company Confirmation Mismatch	D	1000	1	67				67000	93.30%	3.00		
Overall	Lost Sale	D	1000	1	89				89000	91.10%	2.85		
												Overall (Normalized) Process Yield:	95.55%
												Rolled Throughput Yield:	77.76%
												Overall Process Sigma:	3.20

Calculated Using Basic Sigma Method for Discrete Data

Calculated Using First Pass Yield Formula and Short Term Sigma

Calculated Using Basic Sigma Method for Continuous Data

The Product of Sub-Process Yields (and Failed Trade Yield)

Calculated Using Normalized Yield Formula and Short Term Sigma

6.8.6 Some Notes on Capability Calculations

Although you can do the index calculations, if the process is not *stable*, then we really don't recommend examining the process' capability. This surfaces an “old” philosophy of process improvement that still has validity:

First understand the performance of the process (run or control chart). Then work to eliminate assignable causes of variability. When the process is stable, then assess the process’ capability and if it is not capable, work to identify the common causes of variation and improve the process from this perspective.

Yet another note: We have been dealing with measurement data for these calculations. For count data, the measure of process capability is easier. If we are dealing with *defective* items, then either the average number of defectives (per sample) or the average fraction (or percent) defective gives us a measure of the process' capability.

If we are dealing with *defects*, then the average number of defects (per area of opportunity) or the average defect rate gives us the capability of the process. Note that if you have a control chart for your process, the Center Line of the chart tells you your process' capability.

6.8.7 The Difference Between CAPA- and STA-bility

Now that we understand the concepts of stability and capability, let's spend a few minutes looking at the ways our process could be performing. Our processes can either be stable or unstable, capable or incapable. This gives us four possible states for the process:

State 1 - Not Stable, Not Capable (*Chaos*)

This is the worst of both worlds. There are special causes present in our process, it is not predictable **and** it does not reliably meeting customer requirements. This process cries out for improvement.

State 2 - Not Stable, Capable (*Brink of Chaos*)

Here, we are lucky. Special causes are present in our process: it is not predictable. For now, though, we can meet our customers' requirements (perhaps the customer has adjusted *their* tolerance band because of our process' past performance). Due to the special causes in our process, we may wind up not meeting their requirements in the future. Again, improvement is desired here.

State 3 - Stable, Not Capable (*Threshold*)

At least our process is predictable here, although not meeting customer requirements. We know that only common causes of variation are present. We can analyze these sources of variation and make process changes to minimize their impact on the process' output.

State 4 - Stable, Capable (*Ideal*)

This is the desired state. The process is stable, i.e. predictable and it meets customer requirements. At this point, though, we may go back to our customer and discuss the specification limits. It may be that our customers could improve *their* product if we can reliably produce *our* product within a tighter tolerance band.

Or, if the tolerance band is "rational," we may be able to use a cheaper raw material or ease up on our process' variation.

What, you might say, do you want me to increase variation? Yes, remember, the "game" is to produce quality products and services at the *least* cost to the consumer. Dr. Ishikawa noted that the job of technology is to create quality products from low quality (i.e. low cost) raw materials.

One caution, though. The "spec limit" approach to Process Capability is essentially a "goalpost" mentality. If the product or service goes through the spec limits, we assume everything is OK. This may not be the case.

Dr. Genichi Taguchi goes beyond the "goal post" mentality in his discussion and use of the *loss function* concept. His concept is that even though the product or service is being produced within the specification limits, deviation from the center point or target value results in some loss to the customer of our product or service. Continued improvement through rotating the PDCA cycle may still be desired and economically justified. This will be covered in Unit 14.4.

Even if you've reached the ideal state, the battle isn't over. *Entropy* tends to increase "disorder" in your process and can easily move the process back down to chaos.

Entropy Effects



Threshold State
- Predictable Outcomes
- Non-Conforming Product
- Must Either:
  - Change Process
  - Change Specifications
- Sorting is only a Temporary Fix
- Process Behavior Charts:
  - Maintain Predictability
  - Evaluate Improvement Efforts

Ideal State
- Predictable Outcomes
- 100% Conforming Product
- Process Behavior Charts:
  - Maintain Predictability
  - Give Timely Warning of any Troubles
  - Evaluate Improvement Efforts

Chaos Managers

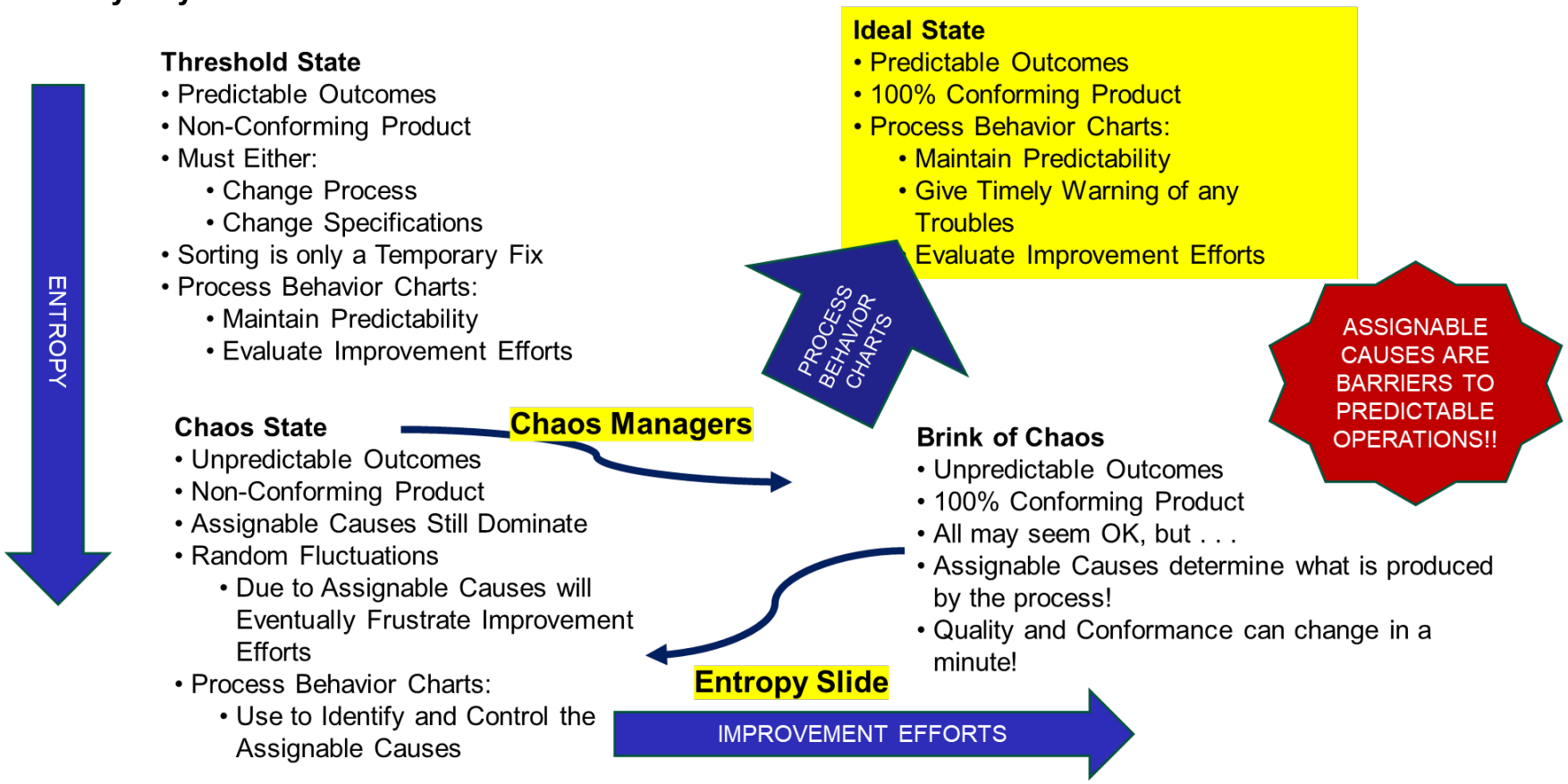
Brink of Chaos
- Unpredictable Outcomes
- 100% Conforming Product
- All may seem OK, but . . .
- Assignable Causes determine what is produced by the process!
- Quality and Conformance can change in a minute!

Chaos State
- Unpredictable Outcomes
- Non-Conforming Product
- Assignable Causes Still Dominate
- Random Fluctuations
  - Due to Assignable Causes will Eventually Frustrate Improvement Efforts
- Process Behavior Charts:
  - Use to Identify and Control the Assignable Causes

Entropy Slide

IMPROVEMENT EFFORTS

The Only Way Out



6.9 Additional Control Chart Topics

Learning Objectives

- Manage Sporadic Events with Control Charts
- Apply X, mR Charts to Non-Normal Data
- Perform an Analysis of Means (ANOM) with Control Charts
- Detect Small Average Shifts with the CUSUM Control Chart
- Use Control Charts to manage Short Runs of data
- Plot Auto-Correlated Data on Control Charts
- Calculate Variable Limits for X-Bar, R Control Charts

Unit Contents

- Managing with Control Charts
- Control Charts for Sporadic Events
- Non-Normal Data and X, mR Charts
- Control Charts and ANOM
- Detecting Small Average Shifts - the CUSUM Control Chart
- Short Run Control Charts
- Auto-Correlated Data on Control Charts
- Variable Limits for X-Bar, R Control Charts

6.9.1 Managing With Control Charts

We hate to make an obvious point, but there is too much of this going on to ignore. If you or your people are going to invest the time and effort to learn how to construct control charts and collect the data needed to feed the charts, please use them!

We guarantee that right after a company takes a course in control charting, there will be charts covering any available wall space in the building. After several months, though, cobwebs will start to appear, the last data point will have been taken a few weeks ago, special causes will be ignored, etc., etc., etc.

Even worse is the case where the charts' maintenance has been assigned to an individual and, as if a ritual, the data just keeps appearing on the charts (we know one fellow who managed to spend 40 hours a week just updating indicator charts - what a smart guy!).

Several topics of interest follow.

Analysis Versus Control

Subgrouping

Control Charts can be used to help analyze a process or to help control (or manage) a process. From our experience, this is one of the most *underutilized* applications of control charts. When Shewhart describes subgrouping, he de-emphasizes display of data over time (the usual focus of a typical SPC course). Shewhart uses subgrouping as a search for important process variables (generally, methods, machines, people, suppliers, materials, etc.). In the examples and exercises presented earlier in this manual, we've presented a few of these subgrouping examples. For example, the *u* chart exercise subgroups the data by X-Ray technician, searching for differences *among* the technicians. This is the preferred application of the control chart.

Data Collection

When we analyze the process (i.e. for improvement) we may perform some concentrated data collection and analysis via control charts, Pareto, histograms, etc. We will try to identify special causes and eliminate them from the process. We will introduce special causes of our own (process changes) to improve the performance of our process. The control chart is a very useful tool to help us measure what is happening in our process.

After a while (and a few turns of the old PDCA wheel), the process will be operating smoothly, and we will turn our attention to monitoring the process. Our control charts can help us here, too, but there may be differences in how we collect the data to feed the chart.

Instead of a concentrated data collection process occurring over several weeks or months, we will now sample from the process, taking maybe a few data points once a day or once a week and plotting these on the chart. We will be interested in assuring ourselves that the process improvements we worked so hard to obtain do not evaporate with time or that new factors are not influencing the performance of our process.

Investigate All Special Causes!

Some people are selective in terms of which special causes they investigate. If the process behaves better than you would expect based on past data, asking why that has happened is just as important as asking why it behaves worse than expected.

Another behavior that we've observed - Some people think that only points outside of the control limits are special causes. Although they may observe trends and shifts occurring (inside the limits), they do not react to them as special causes. INVESTIGATE, INVESTIGATE, AND INVESTIGATE!

One of the strangest behaviors we observed in managing with control charts came from a small Georgia hospital. The nursing staff showed us their control charts on productive hours per pay period. Several "Special Causes" were circled around the holidays. When we looked at the chart, though, we did not observe any signals, just common cause variation. Upon asking, the nurse managers stated that around the holidays, that indicator always went out of control! Although you may think that some factor will result in a special cause signal, listen to the control chart - it's giving you the "Voice of the Process!"

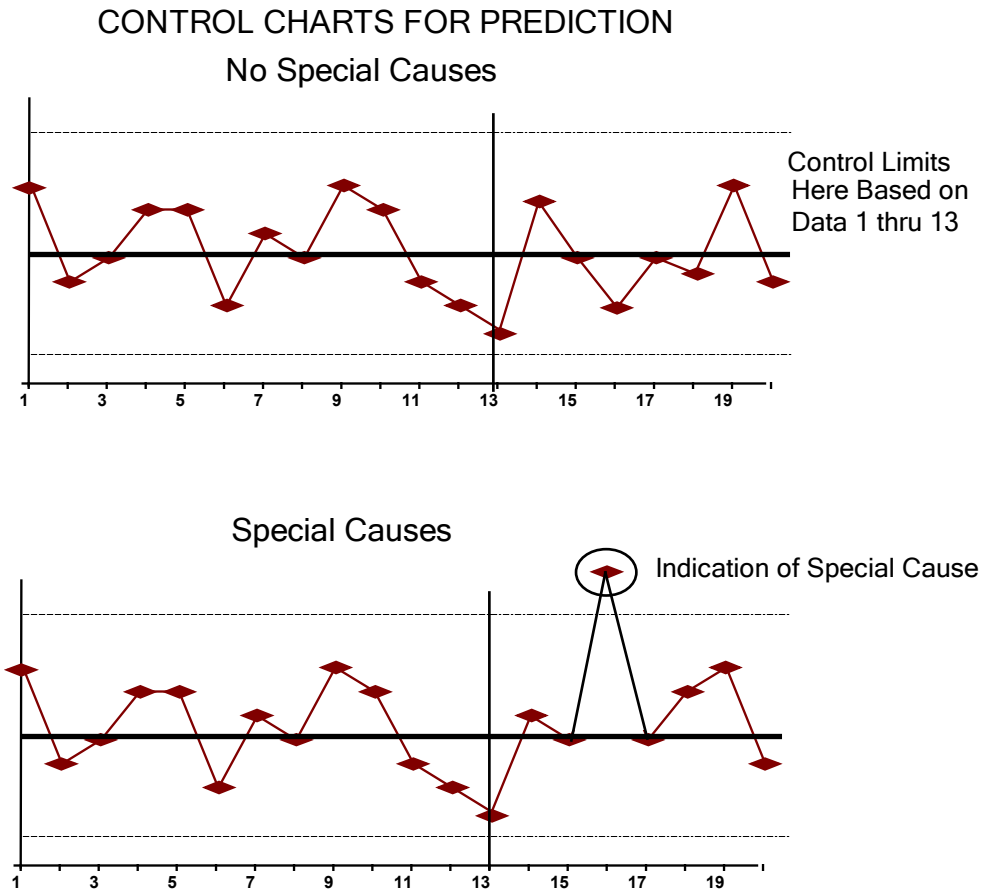
Predicting the Future

When we have gone through the effort of constructing a control chart and stabilizing a process (no special causes present), we gain a major benefit in that we can now predict the future. We now *expect* the process' output to be produced within the established Control Limits. We expect that the output will vary randomly about the Center Line and that no special causes patterns will appear.

Important Note: We cannot predict the value of the next output of the process. This will vary, depending on the interaction of all the process variables that affect this output.

Once the control chart indicates a stable process, we can test our prediction by extending the centerline and control limit lines that are based on the data we took from the past. As we gather future data from the process, we simply plot that data on the chart.

If the data falls within the control limits (and no patterns appear), then the same common causes that determined the process' performance in the past are still present. If a point appears outside the limits or if a pattern appears, then we have evidence that the process has changed. Here's how control charts look for these two cases:



Note that we use a "dot-dash-dot" line for control limits that are being applied currently and that were calculated based on past data.

When to Recalculate Limits

We'll mention a few times you may want to consider recalculating your control limits.

Special Cause Elimination - If you take some data on a process, calculate control limits and find special causes, your limits are based on a combination of both special and common causes at work. If you take action to eliminate the special causes and can assure yourself that they will not occur again, (at least in the foreseeable future), then recalculate your center line and control limits excluding the special cause data.

Insufficient Data - Let's say you only have ten points with which to start a control chart. Do you wait until you have the recommended number? No, plot those, calculate the limits and start evaluating your process. When you get another five points, recalculate your limits. Keep doing this. The fewer data you have to base the control limits, the more cautious you should be in your control chart interpretation. You may miss some special causes; you may chase after some phantom special causes.

Hugging - If you observe hugging on the control chart, then you should first think about your subgrouping strategy. Hugging usually means that we've mixed the output of two or more processes together. Look for whether the process is being performed the same way by the people working in the process (this is another argument for trying to establish the "best current method"), whether two or more machines are involved in the process, and, in general, ways of stratifying the data. Of course, we may just have made a math error in calculating the limits. No, that would never happen.

As Time Goes By - From time to time (maybe every month, maybe every quarter), recalculate the centerline and control limits based on recent process data.

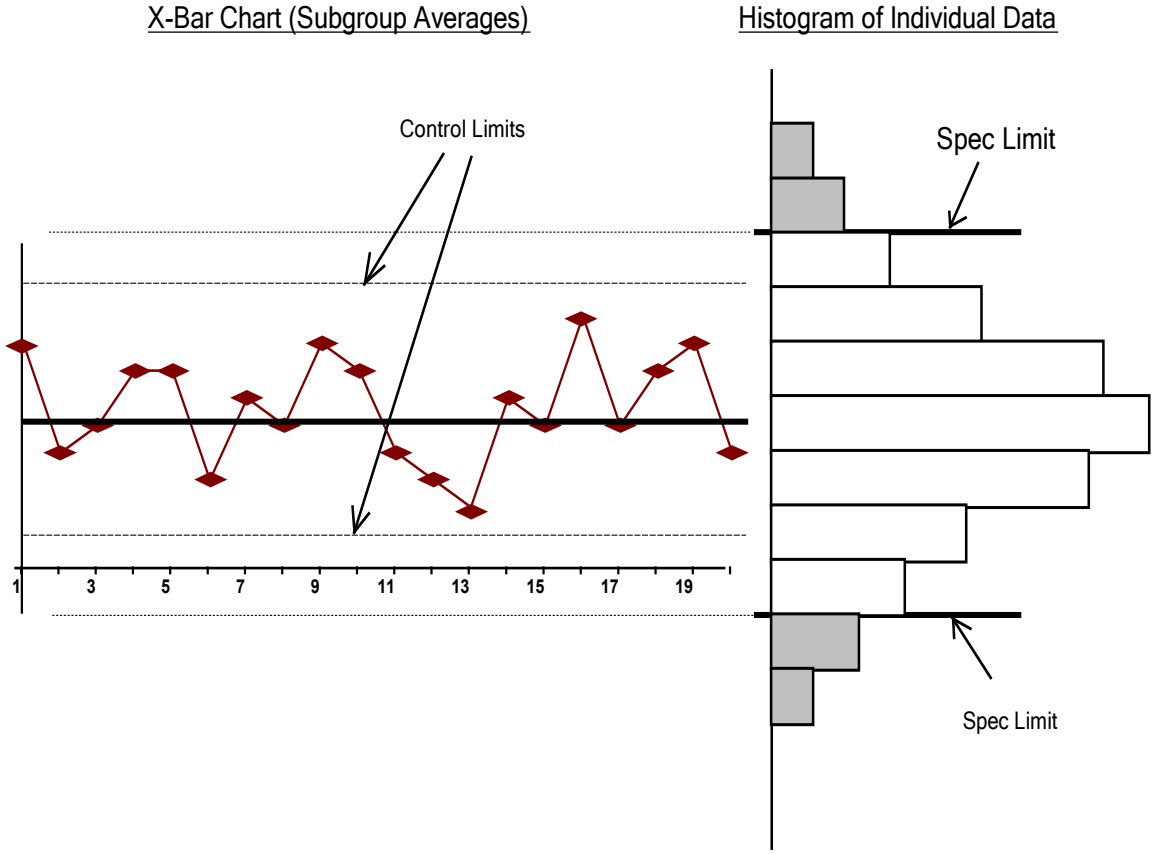
Subgrouping - We've presented the mechanics of control charting to you. Like most endeavors, though, there is a bit of art that you pick up only with experience. Try changing your subgrouping strategy from time to time. Slice the orange a different way and see what you learn. Recalculate the control limits. The strategy is to try and shift as much of the total variation in the process into the *between subgroup* component and away from the *within subgroup* component. Dr. Ishikawa says, "Stamp out R!"

Targets, Specifications and Control Limits

One last comment before we leave the exciting world of control charts. NEVER, EVER put a specification or target line on a control chart. This is important. The control limits are based on your process's data; the process is talking to you and your job is to listen to it. Targets and specifications are values that we impose on the process. If the output of our process falls outside of the target value, that has to do with the *capability* of the process (that's the next section). We do

not want to confuse taking action on a process because it is not stable, with taking action on a process because it is not capable of meeting customer requirements. They are two totally different issues.

There is another subtler issue. When we chart our data on an X-Bar, R control chart, the control limits on the X-Bar chart tell us how we would expect the **process average** to behave. How the individual data points that make up that average are behaving is quite another story. Putting specification limits on an X-Bar chart can cause us to fool ourselves. Notice how all the data on the X-Bar chart falls within the "spec limits," even though a good fraction of the individual data does not?



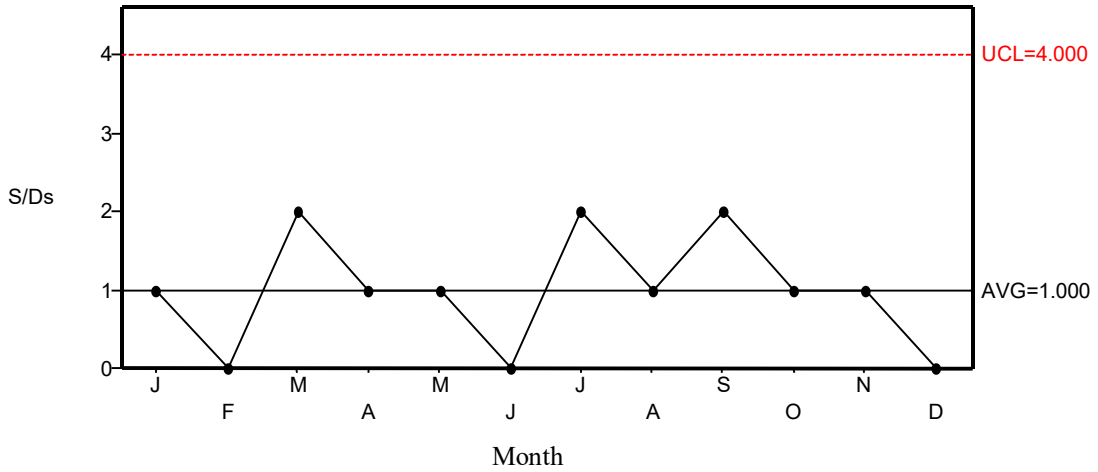
6.9.2 Control Charts for Sporadic Events

Control Charts were initially "designed" to control production processes, where, by their nature, many units are produced in each shift or day. In healthcare and other service industries, though, there are processes that "produce" events only sporadically. If the data obtained from the process is of the *variables* type, the X, mR Chart is used to track performance. A biweekly report on *salary expense* is such an example.

However, if the data obtained is *attribute-type,* a different solution must be pursued. Consider the following record of unplanned equipment shutdowns. The number of shutdowns is recorded for each month and the date of their occurrence noted:

MONTH	NO. OF SHUTDOWNS	OCCURRENCE DATE
JAN	1	1/13
FEB	0	-
MAR	2	3/8, 3/12
APR	1	4/28
MAY	1	5/22
JUN	0	-
JUL	2	7/5, 7/15
AUG	1	8/3
SEP	2	9/1, 9/12
OCT	1	10/31
NOV	1	11/19
DEC	0	-

Following the control chart selection guide, the data is *discrete,* it is of the *Poisson* type, and, assuming the *area of opportunity* remained constant (relatively constant operating demand), the choice of chart would be *c*. The centerline of the data equals *1.0* and the upper and lower control limits are *4* and *N/A*, respectively. The control chart appears below:



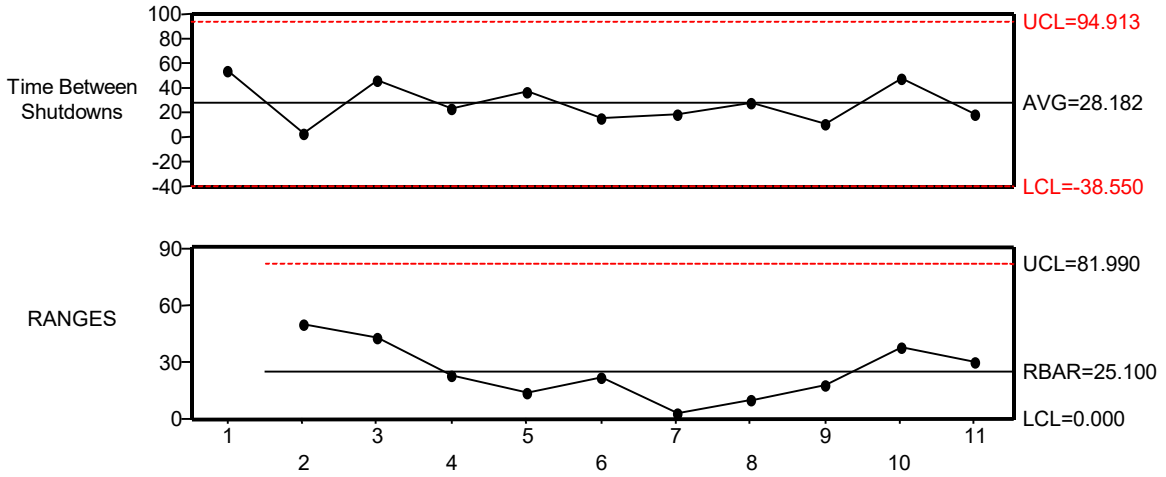
There is a problem with this chart, though. The center line (average number of shutdowns per month) is too close to zero, although a large increase in shutdowns will appear as shifts or other out-of-control signals, it will be difficult to detect improvements in the process. For example, eight or more months will have to go by without a shutdown before a shift signal is detected.

There is a simple alternative to this problem. Instead of tracking the *number* of events, the indicator should be *inverted*, tracking the time or events *between* the events of interest. Using this approach, every event (i.e. each unplanned shutdown) is plotted on the control chart, instead of the one point associated with the area of opportunity (in this case, the month). The X, mR control chart is then used since the data has been transformed to *variables* type.

An additional data collection burden is imposed. Instead of simply counting the number of events, the time between events must be captured (i.e. *days between shutdown*).

The revised unplanned shutdowns data would appear as follows on an X, mR control chart:

Date	1/13	3/8	3/12	4/28	5/22	7/5	7/15	8/3	9/1	9/12	10/31	11/19
Day of Year	013	067	071	118	142	180	196	215	244	255	304	323
Days Between	-	54	4	47	24	38	16	19	29	11	49	19



Unplanned Shutdown

The upper control limit for this data is 95 days. If a significant process improvement occurred, only three months would be required for an out-of-control signal to be detected using the X, mR chart, rather than the eight months required for the *c* control chart.

One caution should be noted regarding the *time/events between* approach. The area of opportunity should be chosen so that it does not change significantly over the period captured on the control chart. In the example above, *calendar time* was used as the area of opportunity. If, for instance, the equipment was shut down for some period due to low demand, then the *area of opportunity* for unplanned shutdowns would change. A strategy to deal with this issue is to eliminate the *down time* from the measure and track the *operating hours between shutdowns.*

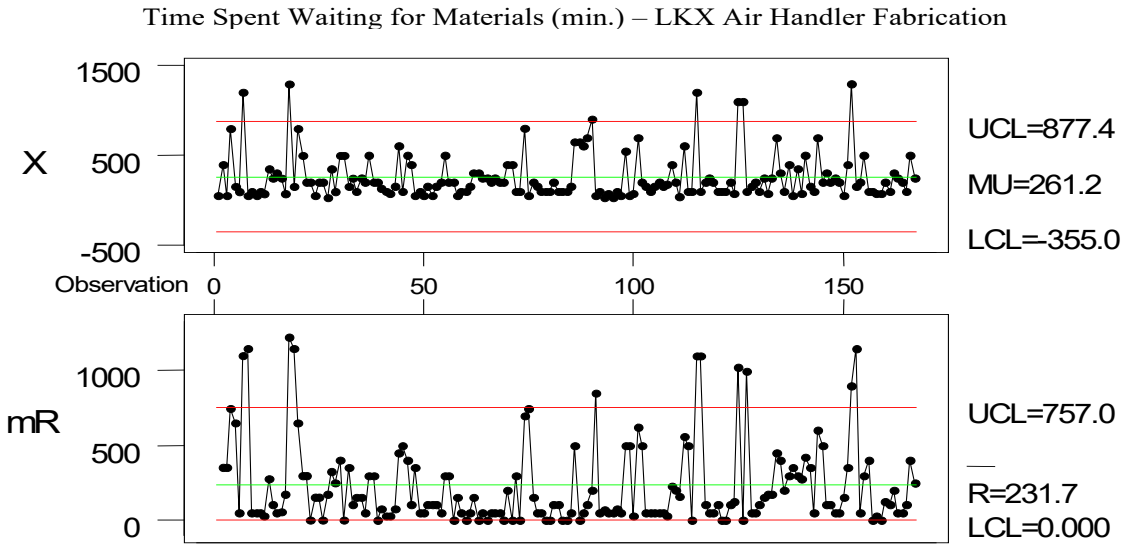
6.9.3 Non-Normal Data and X, mR Charts

The Problem

The Individuals, or XmR Control Chart is a popular one, since many processes "produce" one data point per day, week or month. However, many *time to complete "something"* quality characteristics are skewed left or right.[1] This produces an "interesting" picture of the data, where there appears to be gaps between the data and the control limits.

The following data represents the amount of time spent waiting for materials associated with fabricating a particular air handler. The data has been charted on an X, mR control chart. This process produces a *positively skewed* data distribution. Many of the events have very little in the way of delays, but there are cases where a significant delays occurred. Delays less than zero are not possible.

There are several out-of-control signals on this chart, but the feature of interest here is the "distance" between the smallest values of the data and the individuals' lower control limit of –355 cc's. This gap appears unusual. Is there some way to "close up the gap?" On the other hand, since the data is skewed, is it possible that points outside the upper control limit are not really due to assignable causes? Could they have been produced by the "system?" This topic will describe approaches to dealing with this problem.



[1]Process data are often skewed to the *right* when there is a lower boundary on the minimum time needed to complete a process, but no boundary on the maximum time. Data are sometimes skewed *left* when there is an upper target or specification on the process (e.g. April 15th for tax returns).

How to Detect Non-Normal Data Distributions

Before deciding to "treat" the problem of non-normal distributions, there must be evidence that the situation exists, i.e. is the data "naturally" non-symmetric. Several different approaches for variables data are available.

Frequency Chart/Histogram

The first and simplest approach is to take the suspected data set and create either a frequency chart (discrete data) or a histogram (variables data). This provides the analyst with the shape of the data. If this shape departs significantly from a symmetric, "normal" curve, then the data may be non-normal. Often, measurements of *time to complete* a task or process will show a *skewed shape.*

Probability Paper

This approach is like the frequency chart/histogram approach, except the data is plotted on special *normal probability paper.* If the plotted data departs significantly from a straight line on this paper, the conclusion of non-normality may be drawn.

Hypothesis Tests

There are several hypothesis-test procedures available. These include the χ^2 *goodness of fit test*, the *Kolmogorov-Smirnov test,* and the *Shapiro-Wilk test for normality.* These tests set up hypotheses stating that the data can or cannot be modeled as a normal distribution. "Passing" the test (i.e. not rejecting the null hypothesis of normality) means that there is not enough evidence to show that it does not.

Stratification

Last, but not least, an apparent non-symmetrical distribution can often be an invitation to stratify the data. For example, bill payment times were observed to be highly non-symmetric, i.e. skewed to the left. Stratification by type of customer (*"bill OK"* and *"bill disputed"*) accounted for much of the non-symmetry.

How to Approach Non-Normal Data on Control Charts

Given that evidence of non-normality exists, at least four suggestions on how to approach the problem are available:

Western Electric Handbook - "Modified limits are sometimes used on charts for individual measurements provided there is adequate knowledge about the shape of the distribution of individuals." No further guidance on *how* to make use of this "adequate knowledge" is provided.

Duncan - "It should be noted that control charts for individuals must be very carefully interpreted if the process shows evidence of marked departure from normality. In such cases the multiples of sigma used to set control limits might better be derived from a Pearson Type III distribution or some other distribution for which percentage points have been computed. . . or to chart some transformation of the data instead of the original data."

"Typical" transformations of data include taking the logarithm of the data, squaring the data, or taking the square root of the data. The transformed data is then plotted or charted.

Duncan does not provide any more detail on this procedure; however, it appears to be one where the control limits are set through a probability approach.

Software Approach – Some software allows the practitioner to determine how the control limits are calculated. For example, *Memory Jogger* statistical software provides five "analysis options" that allow one to select *how* the control limits will be calculated: *Best Fit (Johnson), Weibull, Normal, Folded Normal,* and *True-Position*). The "normal" option makes use of the standard coefficients; the *Johnson* and *Weibull* options fit a distribution to the data and calculate control limits based on a probability approach.

Wheeler – Don Wheeler recommends that the practitioner use the existing three-sigma limits. They are robust enough to handle significant non-normal data distributions. Wheeler bases his argument on both an *Empirical Rule,* describing how much of the data will be contained within the mean plus/minus multiples of the standard deviation, and work done by Irving W. Burr, showing the relative insensitivity of the d_2 and d_3 coefficients to non-normal distributions.[2]

[2]These values, used to obtain the relationship between the *Range* and the *Standard Deviation*, and the *Range* and it's *standard deviation,* are based on an assumption of normally distributed measurements.

Wheeler's Empirical Rule seems to be based on both experience and the Camp-Meidell inequality, a modification of the famous general Tchebycheff inequality. Camp-Meidell applies to distribution functions that are uni-modal, with the mode being "close to" or equal to the mean (i.e. not seriously skewed) and monotonic on either side of the mode. The Camp-Meidell inequality expresses the probability that an observation from such a distribution will fall within a given number of standard deviations from the mean:

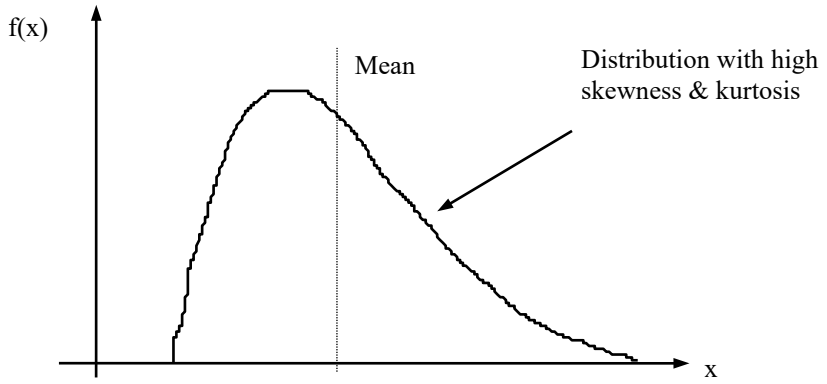
$$\Pr(\left|X - \overline{X}\right| \geq ks) \leq \frac{1}{2.25k^2}$$
$$for\ k > 1$$

This inequality is compared to Wheeler's empirical rule in the table below:

Number of Standard Deviations from Mean (t)	Wheeler's Empirical Rule	Camp-Meidell Inequality	Tchebycheff Inequality
1	60-75%	>56%	-
1.5	-	>80%	>56%
2	90-98%	>89%	>75%
3	99-100%	>95%	>89%

General agreement appears. Note that although Tchebycheff is more general (i.e. *any set of data, bimodal, extremely skewed, etc. will "obey" this inequality*), he also more conservative.

The work of Irving W. Burr involved taking 27 different distributions of various skewness and kurtosis, and calculating the theoretical values for the d_2 and d_3 coefficients for each distribution, these being the basis for calculating the upper and lower control limits for the individuals' charts. When a high degree of skewness and kurtosis exists (thus, pushing as much of the data into a "tail" as possible), the theoretical value of d_2 is about 10% smaller than that calculated for normally distributed data and the corresponding value of d_3 is about 10% larger.



Of what significance is this to the individuals' chart control limits? The control limits for the individuals data would be about 10% wider than those calculated using the "usual" d_2 coefficient. The upper control limit for the range would be about 20% wider (since this is influenced by both d_2 and d_3 coefficients).

Critique of Alternates

The basic difficulty with the first three alternatives (*Western Electric Handbook, Duncan,* and *Memory Jogger "analysis options"*) is that they resort to "invoking" a distribution model to develop the control limits. Given that the purpose of a control chart is always to detect uncontrolled variation in a process, the same difficulties arise with these approaches as were wrestled with by Shewhart in his work. Distribution theory cannot be applied to a process that is not in a state of control.

Wheeler demonstrates that the conservative nature of the three sigma limits, which make use of the estimated standard deviation of the individuals (and based on the average range, "converted" into the standard deviation), are not very sensitive to distributions where the data is forced into tails (Burr's work).

Even for extremely skewed data, the individual chart control limits would increase by only about 10% when the revised theoretical coefficient is used and the range limits would increase by about 20%. We conclude then that three sigma limits can be successfully applied to non-normal data distributions. No adjustments or distribution fitting need occur.

6.9.4 Control Charts and Analysis of Means (ANOM)

The Analysis of Means (ANOM) can be used to examine differences in process output when the cause-and-effect system is deliberately altered, as it is during experimentation. ANOM is a graphical approach like the control chart, but with changes that make it a more appropriate technique to use when conducting experiments, rather than examining the output of a production process. ANOM can also be used as a hypothesis test where we compare the rates of occurrence of some event (e.g. injuries, defects occurring in an area of opportunity)

Some writers refer to the process of comparing process subgroups *that do not have time as their basis* (i.e. comparing different processes, practitioners, equipment, etc.) as ANOM. We view these applications as traditional and, actually, *per* Shewhart, the *preferred* use of the control chart. The ANOM described here *does not make use of three sigma limits,* but rather *decision limits* based on the number of data subgroups and a sensitivity value set by the experimenter. Prior to discussing the actual procedure steps, let's spend a few minutes setting up the motivation for ANOM. The "impatient" reader may jump ahead to the ANOM procedure.

Review of the Control Chart's Purpose

We described above (6.9.1) the two common purposes of a control chart: *control* and *analysis*. These applications generally presume that data is being gathered from an existing process. For these purposes, the control chart makes use of three-sigma limits to detect out-of-control conditions, or *assignable causes* in a cause-and-effect system (i.e. *production system*). These three-sigma limits are based on an economic balance - *minimizing* the chances of both declaring out-of-control conditions to be present when they are not and ignoring the presence of out-of-control conditions when they are. The control chart is designed to be a *conservative* statistical procedure; it is one that is intended to produce very few false alarms. There is yet a *third* control chart purpose that must be addressed - *experimentation.* Here, the process is being deliberately altered (i.e. changes made to one or more process variables) in hopes of improving performance through either a change in the central tendency or a change in the variability of the process.

Control versus Analysis versus Experimentation

There are some differences between a process that is being *analyzed* or one that is being *controlled* to produce a product or service that is *on target with minimum variation* and an experiment that is being conducted to *detect differences* between categories, or between the levels of a certain factor.

The *controlled* process is assumed to be stable, unless the control chart shows evidence to the contrary. The *act of control* involves standardizing the important methods, materials, equipment, etc. so that as uniform a product or service may be produced as possible, given the current "technology." A typical process will "produce" as much data as there are products or services provided.

The *analyzed* process is not assumed to be stable; here we employ the subgrouping concept to *detect* the presence of assignable causes present in the process. Like the *controlled process* the *analyzed* process will "produce" as much data as there are products or services provided.

Experimentation, on the other hand, is conducted in hopes of obtaining a signal that something is different. The typical experiment will involve far fewer data than is available from the "production" process. Two different procedures may be compared, different materials may be compared, or different "levels" of one factor compared in relatively few experiments.[3]

For example, suppose that five manufacturers' epoxy glues are being evaluated for use. These represent five *levels* of one *factor*. Possible effects that might be of interest include their *adhesion* properties, or the *joint life* of the different glues.

In a healthcare setting, different therapy/medication routines may be evaluated for their effect on cardiac patients. Here, therapy and medication are both factors, the type of medication and therapy types are the factors' levels. Possible effects that might be of interest include quality outcome indicators (time to return to work, or regaining other functions), cost of treatment, safety (complications) or compliance with the treatments.

In these two cases, the experimenter is more concerned with detecting a difference. The control chart *can* be used to interpret the results of experiments. There is nothing "wrong" with its application, here. The only problem with the "experimental" control chart is that since it is *conservative,* it may miss some signals that are present. A less conservative, *exploratory* analytic approach can be justified when experimentation is being performed.

Now there exist many different statistical procedures to address this issue. Hypothesis tests, contingency tables, analysis of variance (ANOVA), and discriminant analysis are just a few. Many of these procedures, though, are complicated computationally, and require assumptions that may be difficult to justify in the workplace setting.

[3]Modern design of experiments seeks to maximize the amount of information obtained while *minimizing* the number of experiments required.

The *Analysis of Means* (ANOM) is a procedure, though, that is simple to apply by the "non-statistical" experimenter and one that has the same "look and feel" as the control chart. For organizations that are used to interpretation of the control chart, the ANOM will fit in easily.

Before the ANOM procedure is introduced, note that a single experiment does not *confirm* the presence of an important cause. *Replication* of the results is generally necessary to satisfy the conditions of the scientific method.

Experimental Process

ANOM is used to analyze and interpret data from an experiment. Any experimental process will identify certain factors that are suspected of (or *hypothesized to*) influencing some output variable. The basic experimental process will include the following steps:

1. State the problem to be solved - Determine what problem is being addressed. Is there only a *single* outcome or response of interest, or are there *multiple* outcomes?

2. Determine the objective of the experiment - Identify the performance characteristics (output variables) that are to be measured and the desired level of performance (i.e. target values) when the experiment is complete. Determine how the characteristics will be measured (operationally defined).

3. Determine the factors that are suspected of influencing the performance characteristic(s). A Cause & Effect diagram prepared by several people familiar with the process may be helpful here. Determine the number of levels and the values of these levels for each factor.

4. Determine the *risk*[4] of declaring that a signal is present when, in fact, there is only noise (the α for the experiment).

5. Design an experiment that considers these different factors, their levels and the performance characteristics. This may be as simple as a *one factor, two level* experiment, or as complex as a "nested" factor orthogonal array.

6. Conduct the experiments required by the design.

7. Analyze the data and interpret the results. Here is where the *statistical* technique of ANOM will be applied.

[4]The ANOM procedure will make use of a *sensitivity value* that represents this risk.

8. Run confirmatory experiments for the important factors (and levels) determined by the experiments. If the confirmatory experiments show similar results as the initial experiments, proceed to incorporate these changes in the production process. If not, return to the experimental process.

ANOM Procedure

As shown above, the ANOM method is employed to analyze and interpret the results of experiments performed. For the ANOM procedures described below, the assumptions regarding the *inputs* to the ANOM calculations will be explained.

ANOM for Single Factor Experiments

Introduction

Many experiments consider only the effect of one factor on some outcome variable or characteristic. The one factor may be set at two or more *levels* (or conditions) as part of the experiment. The ANOM for single factor experiments will analyze and interpret the data from this situation. For example, five different types of *glue* were tested for their effect on the characteristic *adhesion.* Here the factor is the *glue*, the different *types* are the levels, the *adhesion* is the outcome variable.

Like the control chart, *subgroups* of data will be collected, with each subgroup representing one level of the factor. The outcome variable, of course, will be measured for each experiment run. The ANOM will differ from the control chart, though, in several aspects:

1. Decision Limits (versus Control Limits) will be calculated from the data. These decision limits will be based on five factors:

- The Grand Average of the data,
- An estimate of the Standard Deviation of the subgroup,
- The number of subgroup averages being compared (i.e. the number of levels of the factor - five different glue brands are five *levels* of one factor: the glue),

- The number of *degrees of freedom* of the data (dependent on the number of subgroups and the subgroup size, obtained from a table lookup), and

- The *sensitivity* of the experiment - the more "exploratory" the experiment, the more sensitive it will be to detecting signals (this is termed the α value, and is set by the experimenter, but should not be confused with the "traditional" α and β *risk levels* of hypothesis testing).

2. The ANOM is *not* intended for use with on-going series of production data. The decision limits' dependency on the *degrees of freedom* of the data set precluded this. As more subgroups are added to the data set, the decision limits will change.

3. The only interpretation rule to be applied to an ANOM chart is a point outside of the decision limits (Wheeler indicates that points "close to" the decision limits may also be interpreted as signals). Runs, trends, stratifications, etc. do not apply to the ANOM.

Procedure - Variables Data, Subgroups of Size 10 or Less

The following ANOM procedure is applicable to variables data, where the subgroups are of size 10 or less. This should handle the "typical" experimental situation where a few replications (the subgroup size) of each level of a factor are obtained.

1. Perform the experiments, and obtain the *k* subgroups of size *n.*

2. Calculate the subgroup ranges and the average range. Prepare a Range chart using the X-Bar, R control chart procedure.

3. Calculate the subgroup averages and the grand average, using the X-Bar, R procedure.

4. Estimate the standard deviation of the subgroup averages as follows:

$$\hat{\sigma}_{\bar{X}} = \frac{\bar{R}}{d_2\sqrt{n}}$$

where:

$\bar{R}$ - average range

d_2 - coefficient dependent on subgroup size

n - subgroup size

$\hat{\sigma}_{\bar{X}}$ - estimate of averages' standard deviation

5. Determine the degrees of freedom (ν) for the given subgroup size and number of subgroups. Round this off to the nearest integer. Although tables of the degrees of freedom are published, an adequate approximation to be used *in lieu of* the tables is:

$$\nu = 0.90k(n-1)$$

6. Using the number of subgroups (*k*), the degrees of freedom (ν), and the sensitivity (α), obtain from the below tables the value of *H,* the multiplier for the decision limits.

7. Create the ANOM chart by plotting the subgroup averages, the grand average and the decision limits:

$$UDL = \bar{\bar{X}} + H\hat{\sigma}_{\bar{X}}$$

$$LDL = \bar{\bar{X}} - H\hat{\sigma}_{\bar{X}}$$

where:

UDL, LDL - Upper and Lower Decision Limits

$\bar{\bar{X}}$ - Grand Average of Subgroups

H - decision limit multiplier

$\hat{\sigma}_{\bar{X}}$ - subgroup averages' standard deviation

8. Interpret the chart for signals:

Range Chart - Points outside the range chart limit(s) are signals. Usually, the Range chart will show signals outside the upper control limit, indicating a subgroup whose variability is higher than the rest of the subgroups. This subgroup point should be eliminated from the calculations for two reasons:

a. A factor level whose variability is higher than the rest is usually not one that will be of interest from a quality standpoint (remember, the idea is to *reduce* variation).

b. Since this subgroup is not of practical interest, it should be eliminated. The higher range increases the estimate of the standard deviation of the subgroup averages, which makes the ANOM chart less sensitive.

X-Bar chart - Points outside (or close to) the decision limits are evidence of one or more factor levels being different from the rest. The factor level that produces results closest to the performance goal may be considered for further, confirmatory experiments.

ANOM for Multiple Factor Experiments

The preceding discussion addressed the *single factor, multiple levels* experiment. It is possible to "stretch" the ANOM procedure to analyze the results of *multiple factor, multiple level* experiments. When the experimentation reaches this level of sophistication, though, there are more appropriate techniques, such as Analysis of Variance (ANOVA) or Taguchi Signal-to-Noise Ratios that should be used. In these situations, too, the analyst must be careful to *design* the experiment so that they maximize the information obtained.

Decision Limit (H) Values

A. Sensitivity (α) = 0.10

Degrees of Freedom	Number of Means Being Compared (# Subgroups)						
	2	3	4	5	6	8	10
5	1.42	2.15	2.49	2.73	2.91	3.18	3.37
6	1.37	2.06	2.37	2.60	2.77	3.01	3.19
8	1.32	1.95	2.24	2.44	2.60	2.82	2.98
10	1.28	1.89	2.17	2.36	2.50	2.71	2.87
15	1.24	1.81	2.07	2.25	2.38	2.57	2.70
20	1.22	1.78	2.03	2.19	2.32	2.51	2.64
30	1.20	1.74	1.98	2.15	2.26	2.44	2.56
40	1.19	1.72	1.97	2.12	2.24	2.40	2.52
60	1.18	1.71	1.94	2.09	2.21	2.38	2.50
120	1.17	1.69	1.92	2.08	2.18	2.35	2.46
∞	1.16	1.67	1.90	2.05	2.15	2.31	2.42

B. Sensitivity (α) = 0.05

Degrees of Freedom	Number of Means Being Compared (# Subgroups)						
	2	3	4	5	6	8	10
5	1.82	2.65	3.06	3.33	3.54	3.84	4.07
6	1.73	2.59	2.94	3.19	3.37	3.58	3.78
8	1.63	2.39	2.71	2.92	3.09	3.33	3.45
10	1.58	2.29	2.58	2.78	2.93	3.15	3.31
15	1.51	2.16	2.42	2.60	2.74	2.93	3.07
20	1.48	2.10	2.35	2.52	2.64	2.83	2.96
30	1.44	2.04	2.28	2.44	2.56	2.73	2.86
40	1.43	2.01	2.25	2.40	2.52	2.69	2.80
60	1.41	1.98	2.21	2.36	2.48	2.64	2.76
120	1.40	1.95	2.18	2.33	2.44	2.60	2.71
∞	1.39	1.93	2.15	2.29	2.40	2.55	2.65

C. Sensitivity (α) = 0.01

Degrees of Freedom	Number of Means Being Compared (# Subgroups)						
	2	**3**	**4**	**5**	**6**	**8**	**10**
5	2.85	4.03	4.58	4.96	5.25	5.68	5.98
6	2.62	3.74	4.21	4.53	4.78	5.07	5.33
8	2.37	3.31	3.70	3.97	4.17	4.47	4.63
10	2.24	3.08	3.43	3.67	3.86	4.11	4.29
15	2.08	2.81	3.12	3.32	3.47	3.69	3.84
20	2.01	2.70	2.98	3.17	3.30	3.50	3.63
30	1.94	2.58	2.85	3.02	3.15	3.33	3.45
40	1.91	2.53	2.79	2.95	3.07	3.24	3.36
60	1.88	2.48	2.73	2.88	3.00	3.16	3.27
120	1.85	2.43	2.67	2.82	2.93	3.09	3.20
∞	1.82	2.39	2.61	2.76	2.87	3.02	3.12

ANOM Example

The managers of a truck fleet wished to determine if gasoline type made a difference in their mileage (miles/gallon). They identified three different gasoline brands and conducted experiments where each gasoline was used for six tanks of gas (the experiments were randomized over three drivers and three ambulances). Their results are shown below:

	Gasoline		
	PetrolUS	**KP-Extra**	**HCA-Lite**
	8	10	10
	9	12	7
	8	11	9
	10	10	6
	9	12	9
	8	13	9
R	2	3	4
Avg.	8.67	11.33	8.33

Calculations:

$$\overline{R} = 3.0$$

$$UCL_R = 2.004 \times 3.0 = 6.0$$

$$\hat{\sigma} = \frac{3.0}{2.534\sqrt{6}} = 0.483$$

$$\overline{\overline{X}} = 9.44$$

$$\nu = 0.9 \times 3 \times (6-1) = 13.5$$

$$Desired Sensitivity(\alpha) = 0.1$$

$$From Table 12B-1:$$

$$H = 1.834$$

$$Interpolated between \nu = 10 and 15$$

$$UDL = 9.443 + 1.834(0.483) = 10.32$$

$$LDL = 9.443 - 1.834(0.483) = 8.55$$

Interpretation: *The Ranges are in control. The subgroup averages, though, fall outside the control limits, except for PetrolUS. The KP-Extra has the highest average mileage (mile/gallon) and is significantly different than the other gasoline. If cost and other quality factors support the use of KP-Extra, the truck fleet should run confirmatory experiments and consider switching to this gasoline.*

6.9.5 Detecting Small Average Shifts – The CUSUM Control Chart

Purpose of the CUSUM Chart

The Cumulative Sum (CUSUM) Chart is often proposed as a more sensitive tool (than the control chart) to detect small shifts in a process' average. It is essentially a sequential test of hypothesis, where the detection of out-of-control conditions is based on all the data. E. S. Page, a British statistician developed the CUSUM chart.

Various forms of CUSUM charts will be presented in this topic and contrasted to the standard control chart approach. The reader will find that when the purpose of the control chart is to maintain the *status quo,* the CUSUM chart has some advantages. However, when the purpose is *continual improvement*, the CUSUM has little to offer over the "traditional" control chart.

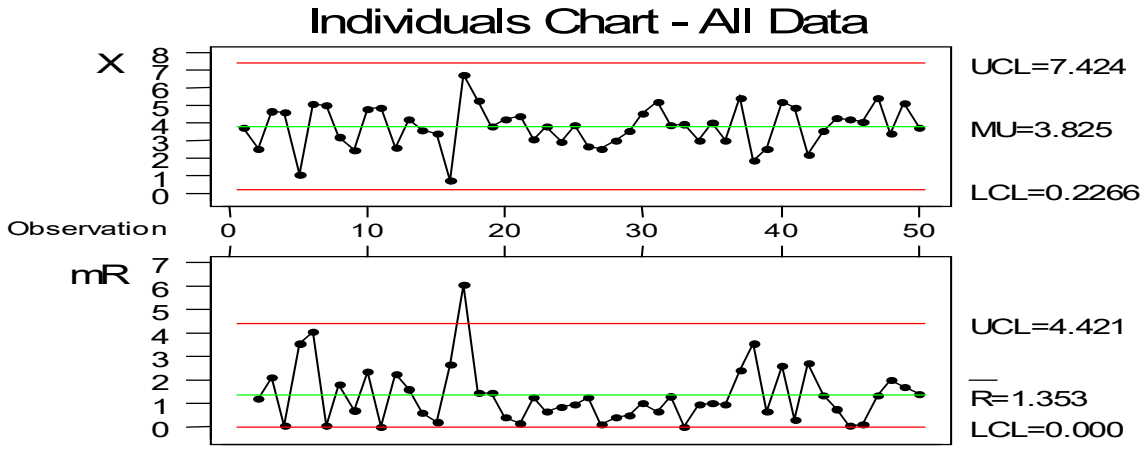
Motivation for the CUSUM Chart

We'll start by showing an example of the CUSUM chart's usefulness in detecting small shifts in the mean of a process that may not be detectable using a control chart. The following data was obtained from a random number generating BASIC program:

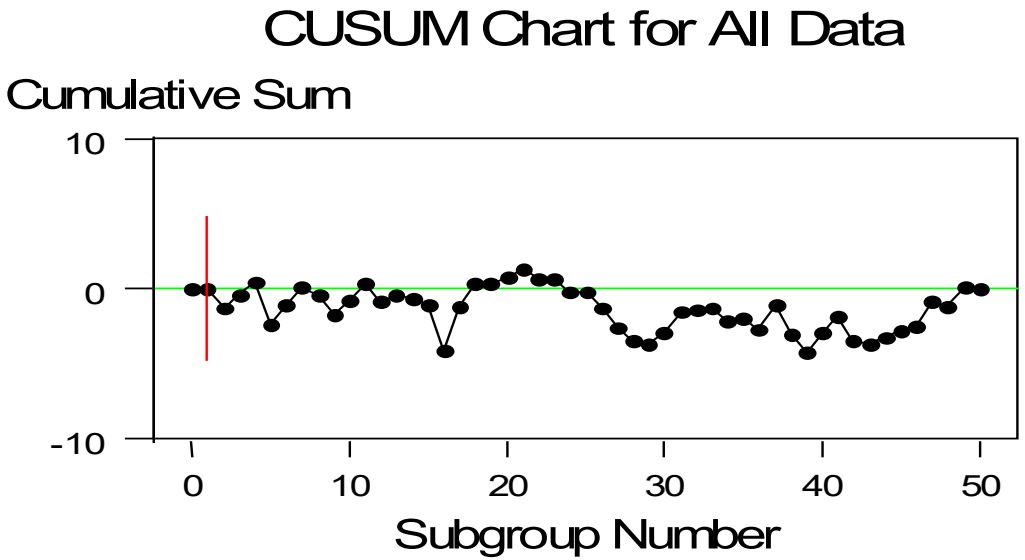
3.79	2.50	0.74	3.80	4.57	5.47	4.31
2.58	4.85	6.80	2.95	5.24	1.89	4.25
4.71	4.89	5.31	3.91	3.92	2.57	4.11
4.62	2.64	3.85	2.66 - X	3.95	5.21	5.47
1.04	4.24	4.25	2.56	2.99	4.91	3.44
5.11	3.64	4.40	3.00	4.03	2.21	5.17
5.04	3.42	3.11	3.54	3.05	3.56	3.75
3.24						

The first 25 points (read from top to bottom and left to right along the columns) are from a Weibull distribution, with shape parameter 3.4 and scale parameter 4 (roughly simulating a normal distribution, with mean of 4). At the "X," the program was modified to change the scale parameter to 4.7, with the remaining 25 points obtained from this run. An X, mR control chart was prepared using *all* of this data:

The range chart shows one range outside the upper control limit (point 16) and a run from points 19 - 35. On the X chart, point 16's out of control range appears as the difference between points 16 and 17 (these points are out-of-control using the 2/3 points outside zone 2 rule). No signals are present, which indicate the small shift in the process average initiated at point 25.



The following chart is a simple CUSUM chart of this same data. The points here are the cumulative sum of the *difference* between the data and the average value obtained from the X, mR control chart. The cumulative sum is calculated starting from the first difference:



On this CUSUM chart, the last twenty-five points clearly show up at a different level than the first 25 points. Simply "eyeballing" the chart leads the viewer to this conclusion, without any specific rules.

Two-Sided CUSUM Chart - V-Masks

The calculations required to create a "formal" CUSUM Chart can be performed with a spreadsheet program. The graphical interpretation requires the creation of a template called a "V-Mask." Scissors and a ruler are needed; a piece of heavy stock paper or light cardboard are the raw materials for the V-Mask.

1. Pick a Target Value for the data. The CUSUM chart is a sequential test of hypotheses. The null hypotheses (H_o) can be stated as "the process mean is equal to the target value," the alternate hypotheses (H_a) can be stated as "the process mean is not equal to the target value." Choice of the target value (*k*) is of some importance; it will influence the shape of the CUSUM chart greatly. Two possible choices for *k:*

a. If a "traditional" control chart of the data is available, use the average value of the data as *k.*[5]

b. If the process is being controlled to some target value (nominal or aim), use this as *k.* (the CUSUM chart is best for maintaining the *status quo*).

2. Calculate an Estimated Standard Deviation of the Data. If the CUSUM Chart is being developed for individual data, calculate the average moving range, just as is done for the X, mR control chart. If the CUSUM chart is being developed for subgroups (of size *n*), calculate the average range of the subgroups, just as is done for the X-Bar, R control chart. Then convert these values into estimates of the standard deviation:

For Individuals Data:

$$\hat{\sigma} = \overline{R} / d_2 = \overline{R} / 1.128$$

For Subgrouped Data:

$$\hat{\sigma} = \overline{R} / d_2\sqrt{n}$$

3. Calculate the Cumulative Sums for the Data:

[5]In the "old" days, with hand calculations being required to prepare the charts, resources may have led to having to choose one chart over another. With PC's and spreadsheets today, many different analyses can be performed in a short time.

$$S_i = \sum_{j=1}^{i} (X_j - k) = (X_i - k) + S_{i-1}$$

where:

X_j - "ith" individual data or subgroup average

k - target value

S_i - "ith" cumulative sum

The spreadsheet associated with these calculations would appear as follows:

Row	Data	Data - Target	Cumulative Sum (to be Plotted)
1	X_1	$X_1 - k$	$S_1 = X_1 - k$
2	X_2	$X_2 - k$	$S_2 = (X_2 - k) + S_1$
3	X_3	$X_3 - k$	$S_3 = (X_3 - k) + S_2$
.	.	.	.
.	.	.	.
i	X_i	$X_i - k$	$S_i = (X_i - k) + S_{i-1}$

4. Scale the axes of the CUSUM Chart. For the CUSUM chart, there is one important criterion: the *vertical and horizontal axis tick marks must be spaced the same distance.* If the *physical* distance between "tick marks" on the horizontal axis is "X," the *physical* distance between "tick marks" on the vertical axis must also be "X." A piece of quadrille paper, with squares of 0.25 inches would meet this requirement.

On the horizontal axis, of course, one data point will be plotted for each "tick mark." On the vertical axis, the scaling of the "tick marks" should be in intervals, "I," where "I" is a convenient value between one and two times the estimated standard deviation calculated in step 2:

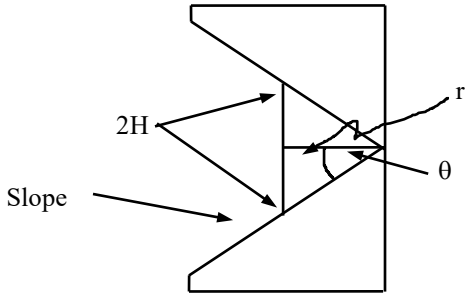
$$\hat{\sigma} \le I \le 2 \times \hat{\sigma}$$

For example, if the estimated standard deviation were 2.3, then the vertical axis interval would be between:

$$2.3 \leq I \leq 4.6$$

"I" could be 3, or 3.5 or 4, or whatever interval is convenient.

5. Prepare the V-Mask. Unlike control charts, where the "usual" out-of-control conditions are already established, the CUSUM chart user must make decisions that establish how *sensitive* the particular chart will be and what "risks" the user is willing to assume of making a wrong choice (assignable cause present when none exists and *vice versa*). The decisions are incorporated into the design of the V-Mask, whose format appears below:



a. *How large a shift in process mean to be detected by CUSUM chart?* - Typically, the CUSUM chart will be designed to detect a shift in the mean of from one to two times the estimated standard deviation ($\hat{\sigma}$). To set this detection ability, a value "D" is chosen to be one-half of the desired shift. For an estimated standard deviation of 2.3, and a desired "detectability" of one standard deviation (one sigma), "D" would be 1/2 x 2.3 or 1.15.

"D" is the slope of the V-mask, and, with the vertical axis interval, "I," determines the angle (θ) of the V-Mask's line:

$$\theta = \tan^{-1} (D/I)$$

b. The length of the reference line, *r,* is found as follows:

$$r \approx \frac{5 \times \hat{\sigma}^2}{2 \times D^2}$$

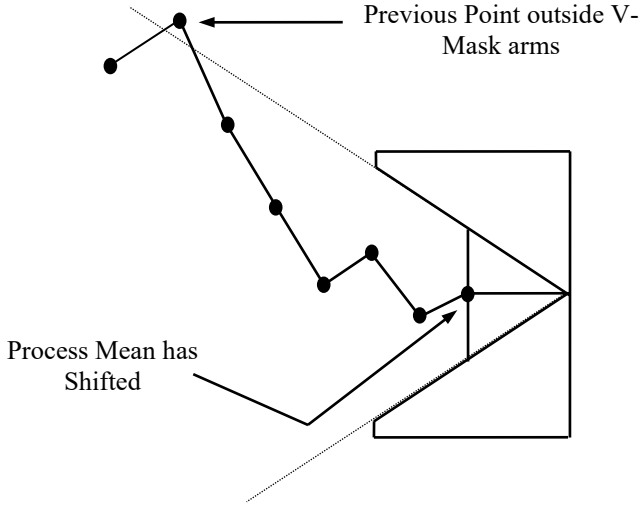
This approximation is obtained from a more complex formula involving the α and β risk levels, and results in a false alarm Average Run Length (ARL) of about 400. This value should be adequate for practical work.

If the analyst wishes to explore different values of *r*, it can be obtained as a function of risks (α - concluding a signal exists when it does not, and β - concluding no signal exists when there is one):

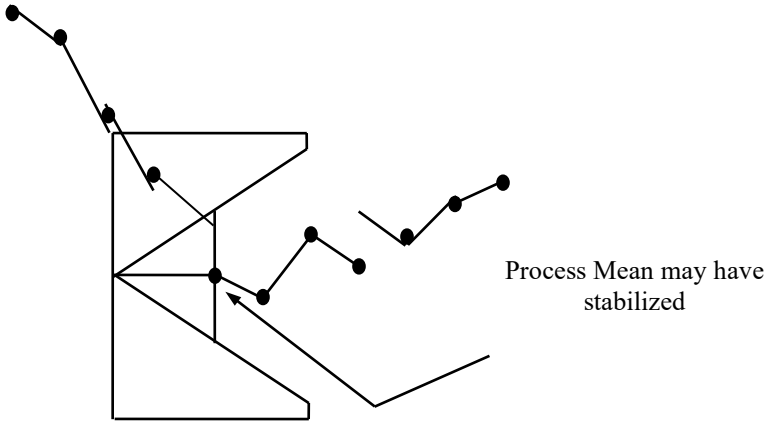
$$r = \frac{\hat{\sigma}^2}{2D^2} \ln\left(\frac{1-\beta}{\alpha}\right)$$

c. From *r* and θ, the V-Mask can be constructed. As a check, the critical distance, "H" is found by multiplying *r* and "D."

6. Plot the points and apply the V-Mask. The cumulative sums are plotted one by one. For each point, the V-Mask is applied with the reference line *parallel* to the horizontal axis of the chart. If any of the previous points fall outside the V-Mask's arms, the current point may be interpreted as a shift in the process mean.



The V-Mask may be reversed and applied to determine the beginning of the run, as follows:



Some Notes on the CUSUM Chart

1. The procedures above describe how to prepare a CUSUM chart for individual values or subgroup averages. The estimated standard deviations were obtained from the average moving range or average range (for subgroups). If the subgroups are large (> 10), then the standard deviation will likely be used as the measure of subgroup dispersion. The average subgroup standard deviation can then be used as the basis for the estimate of the standard deviation of the subgroup averages.

Likewise, if the data is *discrete*, the estimate of the standard deviation will be obtained from the Binomial or Poisson-based standard deviation. Remember, though, to test the discrete data against the *np, p, c* and *u* control chart assumptions before applying these probability models to the process' data. The discrete data may always be treated as *individuals data*.

2. The CUSUM chart shown above is termed a "two-sided" chart. That is, it will detect process mean shifts in either direction. If the purpose of the chart is only to detect an *upward shift* in the process mean (e.g. when tracking errors or defects), then a one-sided CUSUM chart can be prepared.

3. The CUSUM chart should still be used in conjunction with a range chart. Check this chart first before drawing conclusions about the process mean.

4. The CUSUM chart may not detect *gradual* changes in the process mean, or changes that may enter and leave the process in only a few subgroups (less than about 5). The X, or X-Bar chart can be used to detect these.

CUSUM Charts versus Shewhart Control Charts

The Average Run Length (ARL) is a probability-based comparison technique that can be used to compare the CUSUM Chart and the control chart. The ARL is a measure of how quickly a given technique will detect a process change. Wheeler demonstrates that when a control chart using only Rule One (a point outside of the control limits) is compared to a CUSUM Chart, the ARL for shifts of about one to two standard errors is much lower than that of the control chart. Thus, the CUSUM chart will detect process mean shifts quicker than the control chart.

The picture changes, though, when additional Rules are applied to the control chart. When Rules 2 (2/3 points outside zone 2), 3 (4/5 points outside zone 1) and 4 (8 successive points on one side of the centerline) are added to the control chart, the ARL curve for it and the CUSUM chart are virtually identical. The differences are not important for practical work.

The CUSUM chart has some practical difficulties (such as the need for large, vertical scales to track the changes in process mean) that can be overcome, but with additional effort.

After attempting to apply CUSUM charts, some practitioners conclude that there is no advantage to the CUSUM chart over the traditional control chart, others are extremely enthusiastic and recommend exclusively using the CUSUM chart in place of the control chart.

Wheeler points out that although the CUSUM chart is an effective technique, it assumes that the “notion of a distribution is well-defined, rather than examining the data for the existence of that stability which is a prerequisite for the use of distributional assumptions.” Thus, CUSUM moves a step away from the power of the control chart, which does not depend on any probability or distribution theory to justify its basis and application. Juran reinforces this difficulty in listing the required assumption for CUSUM charts that the population of individual measurements be normally distributed. This condition is often not found in practice.

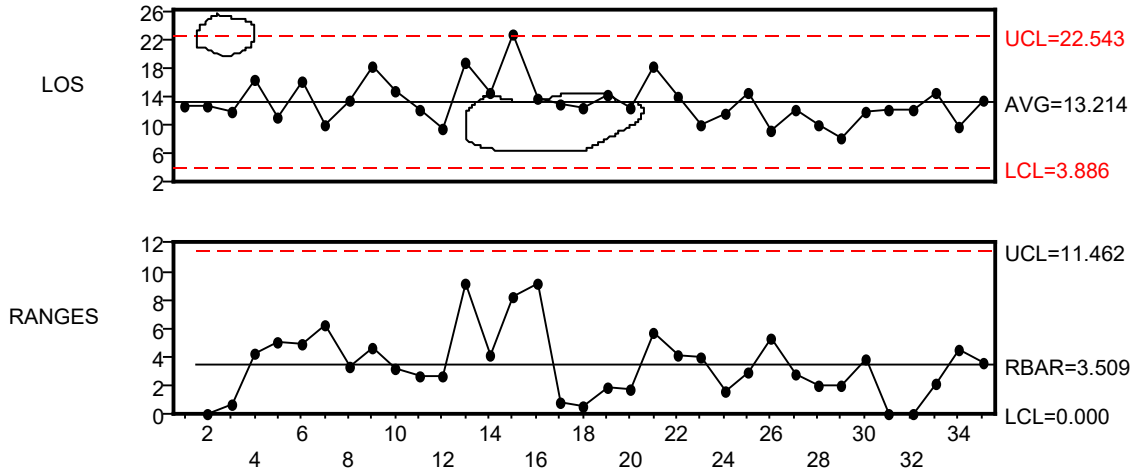
CUSUM Chart Example

The following Average Repair Time data was collected for home air conditioning systems. Efforts at improving the quality of maintenance have been ongoing for some time now. Have these improvement activities influenced Average Repair Time (ART)?

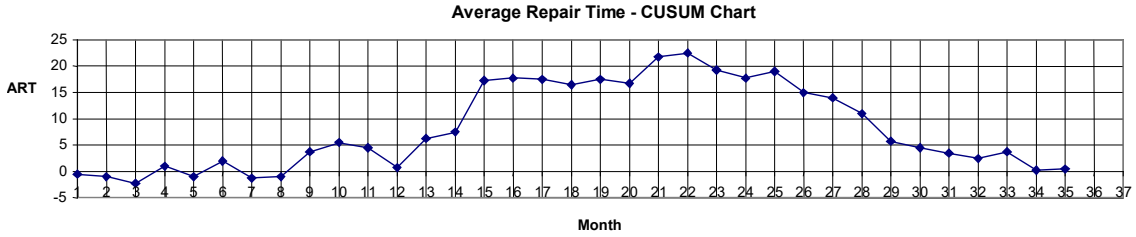
Average Repair Time (Hours)/Month

Month	ART	Month	ART
1	12.7	19	14.2
2	12.7	20	12.4
3	12.0	21	18.2
4	16.3	22	14.0
5	11.2	23	10.0
6	16.2	24	11.6
7	10.0	25	14.6
8	13.4	26	9.2
9	18.1	27	12.1
10	14.9	28	10.1
11	12.2	29	8.1
12	9.5	30	12.0
13	18.7	31	12.1
14	14.5	32	12.2
15	22.8	33	14.4
16	13.7	34	9.8
17	12.9	35	13.4
18	12.3		

The data was first plotted on an X, mR control chart. Assignable cause signals were noted on the chart, as shown below. It appears the last few months have seen a decline in the Average Repair Time.



Next, a CUSUM Chart was prepared using this same data.



Note that since the estimated standard deviation of the process is $\overline{R}/d_2$, or 3.51/1.128 = 3.11 days, the vertical axis is scaled at an interval (I) of 5 hours. This is between one and two times the estimated standard deviation.

The "V-Mask" is then constructed. To obtain a detectable shift of one sigma (3.11 hours), the "D" value is calculated to be 0.5 x 3.11 = 1.55. The V-Mask angle (θ) is calculated from the inverse tangent of D/I (or 1.55/5) and is 17.2 degrees. The length of the reference line, *r*, is $(5 \times 3.11^2)/(2 \times 1.55^2) = 10.$ The value of 10 represents the length of the reference line in time periods (as displayed on the graph above). From this information, the V-Mask is constructed and the CUSUM Chart is interpreted (this is left as an exercise for the student). In this case, does the CUSUM Chart provide more information than the X, mR control chart?

6.9.6 Short Run Control Charts

Purpose

In many manufacturing processes today, attempts are being made to emulate the "Just-in-Time" approach to manufacturing developed by Taiichi Ohno of Toyota (*note:* ref. Unit 5.5, Lean Manufacturing). There are two major aims of Just-in-Time:

- To reduce the cost associated with maintaining inventory ("carrying" charges, storage space, etc.),

- To allow the production process to quickly respond to market needs when the market demands a "little" of product "A" and a "little" of product "B," instead of mass quantities of product "C."

Any process that relies on a flow of supplies, equipment, linen, pharmaceuticals, etc. to "produce" a product or service should be familiar with the principles of Just-in-Time and be exploring its potential application.

In many of these "Just-in-Time" applications, the production runs of product "A" are too short to establish meaningful control limits, before product "B" replaces "A" on the line. Although different products are made, in many cases, essentially the same or similar equipment, methods, materials, etc. are employed to produce the products, i.e. essentially the same causal system is at work.

Short Run control charts allow a record of the process to be maintained despite these changes in product (or service) produced. They can be useful when, due to the low volume of any one product or service, meaningful control charts on the individual products or services cannot be maintained. Techniques for developing short run control charts, with application to "production" processes will be presented in this topic.

The Difference Chart

Purpose

The *Difference Chart* allows a stream of "mixed" products or services to be plotted on the same control chart. Through this chart, the process owner can obtain an understanding of how the underlying causal system is performing, despite the differences in the targeted or "nominal" values of the products and services.

The *Difference Chart* is used when there are differences in the products/services' *centers.* If there are also differences in the products/services' *variability*, then the ZED chart described below should be used. Range charts will be employed to make this distinction.

Application

Applications for the *Difference Chart* include:

- Machining similar parts of different lengths/dimensions
- Processing Time for orders of different equipment or materials
- Repair times for different failure modes on an air handler,
- QC Inspection procedure time (mixture of types of inspections),
- Laboratory turnaround time (small volume, different procedures),

Construction and Interpretation

The basic assumption here is that the production process does not produce "enough" of one product or service to establish meaningful control limits. Therefore, the X, mR control chart will be employed to track the "sequentially-mixed" output from the process. The *Difference Chart* takes the individual measurements and subtracts the appropriate "nominal" or target value. These differences are then used to create an X, mR control chart.

Subtracting the "nominal" or target value of the measurements simply removes the *known* differences from between or among products and services. Note that each short production run will generally have its own target or nominal value and that these values should be subtracted from the individual measurements as appropriate. Measurements A_1, A_2, A_3, A_4, A_5, etc. will have target "A" subtracted, measurements B_1, B_2, B_3, B_4, B_5, etc. will have target "B" subtracted.

The control limits are calculated in the usual way, using the moving ranges to develop an average range and the average range used to develop the upper and lower control limits of the individual *differences.*

$$X' = X - \text{Nominal (Target)}$$

(Specific to each Product or Service)

$$R_i = X'_i - X'_{i-1}$$

$$\overline{R} = \frac{1}{k-1}\sum_{i=2}^{k} R_i$$

$$UCL_R = 3.268 \times \overline{R}$$

$$\overline{X}' = \frac{1}{k}\sum_{i=1}^{k} X_i'$$

$$UCL, LCL - \overline{X}' \pm 2.66 \times \overline{R}$$

The out-of-control interpretation rules applied to the X, mR control chart are also applied here.

Some notes on the Difference Chart:

1. *Nominal* or *Target* or *What?* - The nominal value may be the average value of the (product or service *specific*) data used to develop the control chart or may be obtained from historic data. If the process has a target value, this also may be used. Target values should be used with caution, though, especially if the current *center* of the product/ service is relatively far away from the *target.*

2. If the current data is used to generate the *nominal* value, then the X chart will be centered on zero, and the calculation of X-Bar is not necessary.

3. The usual rule of maintaining the *order* of production still applies, even though there is a mixture of product/service A, A, A, A, B, B, B, B, B, A, B, C, B, A, etc.

4. The *Difference Chart* depends on the between product/service variability being about the same. Separate mR charts of the individual products/services can be prepared to visually compare the Ranges. A test to determine if the *between* product/service variability is the same appears below:

For two products / services (A and B):

Calculate $\overline{R}_A$ and $\overline{R}_B$

Calculate $\overline{R}$:

$$\overline{R} = \frac{(k_A \overline{R}_A + k_B \overline{R}_B)}{(k_A + k_B)}$$

where k_A, k_B are the number of subgroups for A & B.

Now, if $\overline{R}_A > \overline{R}_B$, calculate:

$\overline{R}_A / \overline{R}$ and $\overline{R} / \overline{R}_B$

or, if $\overline{R}_B > \overline{R}_A$, calculate:

$\overline{R}_B / \overline{R}$ and $\overline{R} / \overline{R}_A$

If the ratios are both greater than 1.2, then a difference may be said to exist in the product / services' variability.

If there is a difference in variabilities, then the *ZED Chart* should be used.

The ZED Chart

Purpose

The *Difference Chart* allows two or more products/services to be tracked on one chart, where the "production" process (or causal system) differs only in its *center* or *target* value. The *ZED Chart* handles situations where both the *center* and *variability* of the products or services vary. The *ZED Chart* operates by *normalizing* the data for both the differences in the center and variability of the production process.

Application

The same examples described above in the *Difference Chart* can be tracked on the *ZED Chart.*

Construction and Interpretation

Since both the product/services' *center* and *variability* are different, the process measurements are *normalized*. Generally, normalizing a variable consists of subtracting some estimate of the center (i.e. a mean) from the variable and dividing by an estimate of the standard deviation:

$$Z = \frac{X - \hat{\mu}}{\hat{\sigma}}$$

The *nominal* or *target* value of the product/service can be used as the estimate of the center and obtained as described for the *Difference Chart.* The estimate of the standard deviation must be obtained using a *within-subgroup* standard deviation estimator. As with the *Difference Chart,* the product or service data can be the basis of this standard deviation estimator, or historic process data can be used. For example, if the current product/service data is used, then a moving Range can be calculated and transformed into the estimate of the standard deviation:

$$\hat{\sigma} = \overline{R} / d_2 = \overline{R} / 1.128$$

The *Z-values* are calculated for each individual measurement and an X, mR control chart prepared using these *Z's.* An alternative to the *ZED Chart* is the *Z* Chart.* Here, the average Range (R-Bar) is used as the estimate of *within-subgroup* standard deviation. The *ZED Chart* (or *Z* Chart*) is interpreted using the X, mR chart out-of-control rules.

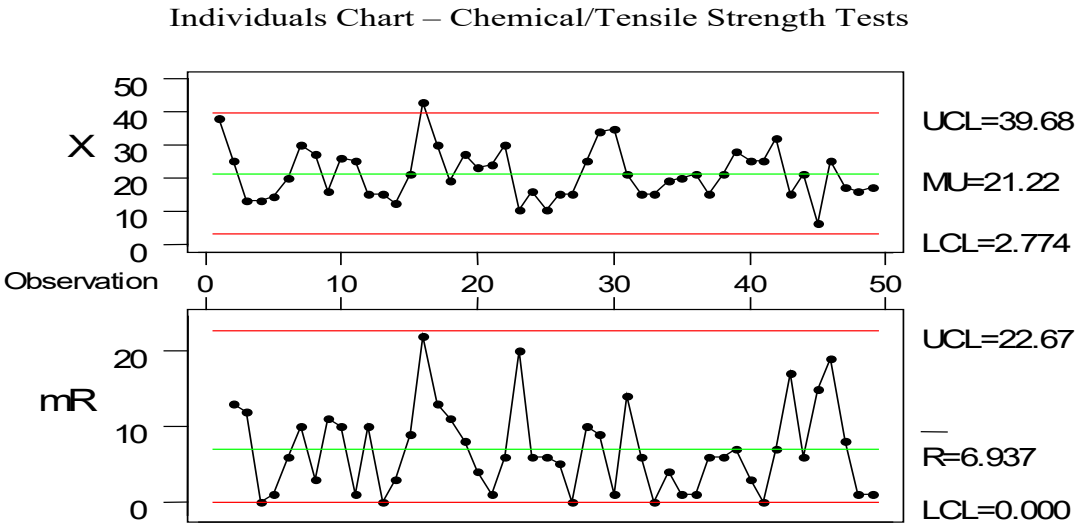
Difference Chart Example

A Testing Lab has been working to improve their schedule performance. As part of the improvement effort, the sample prep process is being analyzed. Each day, a mixture of chemical analyses and tensile strength tests are performed. The following data represents the time to prep the sample for the procedure.

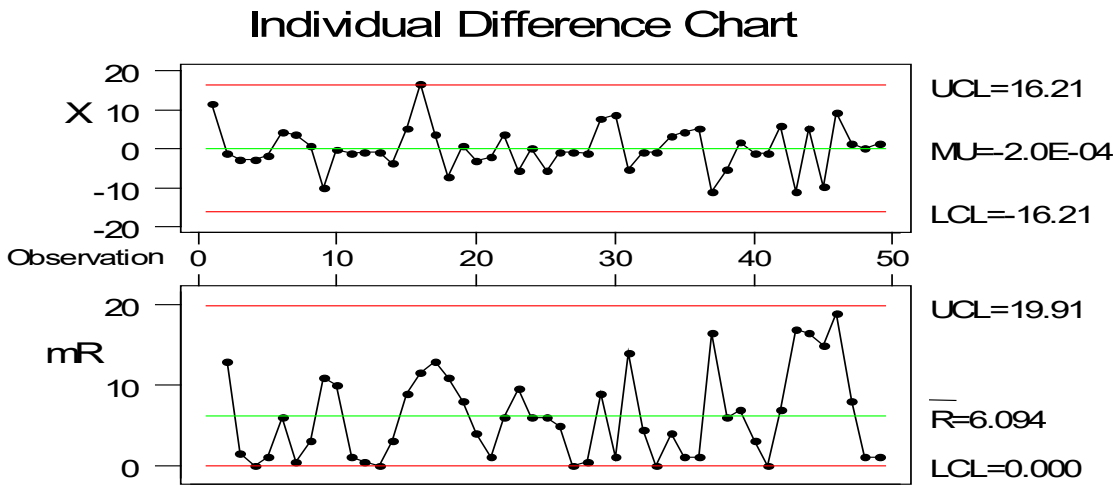
Chemical Analysis/Tensile Strength Procedure Prep Times (Minutes)

Proc. No.	Time	Proc. Type	Proc. No.	Time	Proc. Type	Proc. No.	Time	Proc. Type
1	38	Chemical	18	19	Chemical	34	19	Tensile
2	25	Chemical	19	27	Chemical	35	20	Tensile
3	13	Tensile	20	23	Chemical	36	21	Tensile
4	13	Tensile	21	24	Chemical	37	15	Chemical
5	14	Tensile	22	30	Chemical	38	21	Chemical
6	20	Tensile	23	10	Tensile	39	28	Chemical
7	30	Chemical	24	16	Tensile	40	25	Chemical
8	27	Chemical	25	10	Tensile	41	25	Chemical
9	16	Chemical	26	15	Tensile	42	32	Chemical
10	26	Chemical	27	15	Tensile	43	15	Chemical
11	25	Chemical	28	25	Chemical	44	21	Tensile
12	15	Tensile	29	34	Chemical	45	6	Tensile
13	15	Tensile	30	35	Chemical	46	25	Tensile
14	12	Tensile	31	21	Chemical	47	17	Tensile
15	21	Tensile	32	15	Tensile	48	16	Tensile
16	43	Chemical	33	15	Tensile	49	17	Tensile
17	30	Chemical						

The Testing Lab team first plotted the combined data on an X, mR control chart. They observed assignable causes, but these are traceable to the differences in the procedures.



They then prepared a difference chart by first calculating the average prep time for a chemical composition test (26.4 minutes) and the average prep time for a tensile strength test (16.1 minutes). The difference chart appears below:



The difference chart allows the team to see the variability in the process without the dominating effect of the differences between types of tests.

6.9.7 Auto-Correlated Data on Control Charts

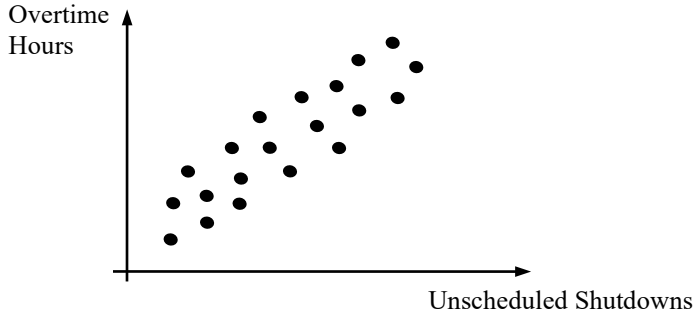
What is Auto-Correlated Data?

Variables are *correlated* when a change in one variable is accompanied by a change in the other variable. This correlation may or may not be the result of a cause-and-effect relationship between the variables.

Examples: Daily data was collected on the number of orders filled by a warehouse and units being built on the shop floor. The data was shown to be positively correlated, that is, as the units increased, the number of orders filled increased. Units being built, though, does not ***cause*** *the number of orders to increase, technicians ordering parts for the units is the causative factor.*

The number of overtime hours worked by plant maintenance staff was positively correlated to the number of unscheduled equipment shutdowns. Here, the additional workload imposed on the staff by the unscheduled shutdowns could be considered a causative factor.

The Scatter Diagram is the basic tool used to graphically determine if two variables are, or are not correlated:



In addition, the correlation between two variables may be measured through a *correlation coefficient:*

$$r = \frac{SS(xy)}{\sqrt{SS(x) \times SS(y)}}$$

where:

r - Simple Correlation Coefficient

$SS(xy)$ - Sum of Squares of xy

$SS(x)$ - Sum of Squares of x

$SS(y)$ - Sum of Squares of y

The Sums of Squares are calculated as follows:

$$SS(xy) = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}$$

$$SS(x) = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x\right)^2}{n}$$

$$SS(y) = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y\right)^2}{n}$$

where:

$\bar{x}$ - average of x's

$\bar{y}$ - average of y's

n - number of data

The correlation coefficient has a range from -1 to +1. *r* values close to + 1 indicate a high degree of positive correlation, *r* values close to -1 indicate negative correlation (as variable "x" increases, variable "y" decreases) and *r* values close to 0 indicate weak or no correlation between the variables. "Practical" values of the correlation coefficient will vary according to the industry or application. For industrial purposes, under controlled experimental conditions, *r*-values above 0.8 are interpreted as evidence of high correlation. For social science work, correlations as low as 0.3 may be indicative of important relationships.

How does this apply to process performance? For some processes, the individual events being measured are not independent of one another. For example, the waiting time for a part is dependent on the queue of orders in front of yours. In this case, the *time series* of data is *auto*-correlated. The *auto*-correlation observed here is between successive "products" or "services" produced by the process. For example, although your order's waiting time may be correlated to the number of other orders in line, or the waiting time for the order in front of yours, there would be little reason to believe that the waiting time today is influenced by waiting time, for parts yesterday.

This topic addresses the issue of auto-correlation, how the control chart is affected, and strategies to handle auto-correlated data.

Measuring Auto-Correlation

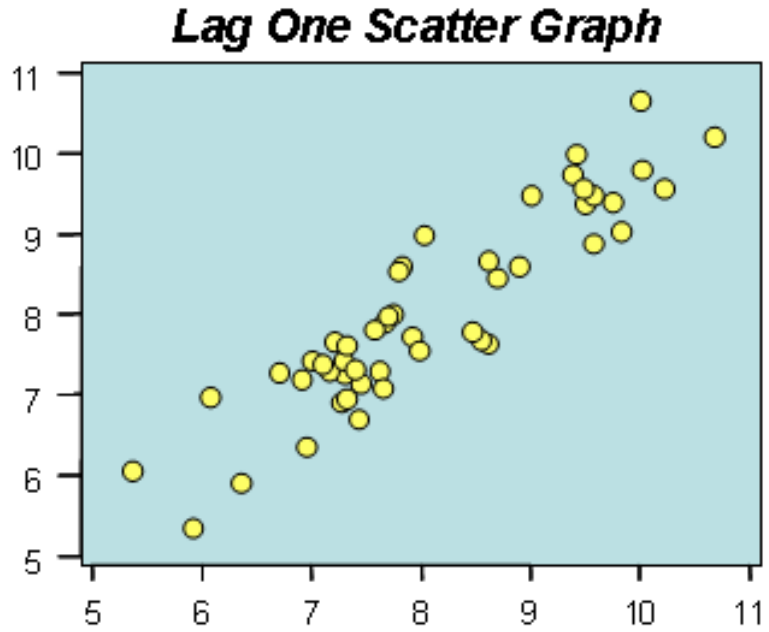
The correlation coefficient presented above will also be used to measure the auto-correlation of a time-series of data. The most common means of developing the "x, y" pairs of data is to consider a *lag* between data points "i" and "i + 1" or "i + 2," etc. That is, "nearby" data points are those that are suspected of being correlated, "farther away" data are not suspected of being correlated.

The following data was created through a simple BASIC program. The program creates a "random walk" pattern. Each data can vary at most by +/- 1 from the preceding data, with a random number (from -1 to 1) determining the actual amount of the random walk:

Random Walk Data

10.00	7.69	7.31	7.44	9.00
10.67	7.98	6.95	7.15	9.48
10.22	7.57	6.35	7.30	9.57
9.57	7.82	5.91	7.26	9.49
8.90	8.61	5.35	6.91	9.38
8.61	7.64	6.06	7.20	9.74
8.68	7.09	6.99	7.66	9.41
8.46	7.39	7.43	7.91	10.01
7.79	7.32	6.70	7.74	9.82
8.55	7.62	7.28	8.02	9.04

This data, then, should have at least a high “lag 1” correlation coefficient. The Scatter Diagram of the data pairs formed from the lag 1 pairs appears below:

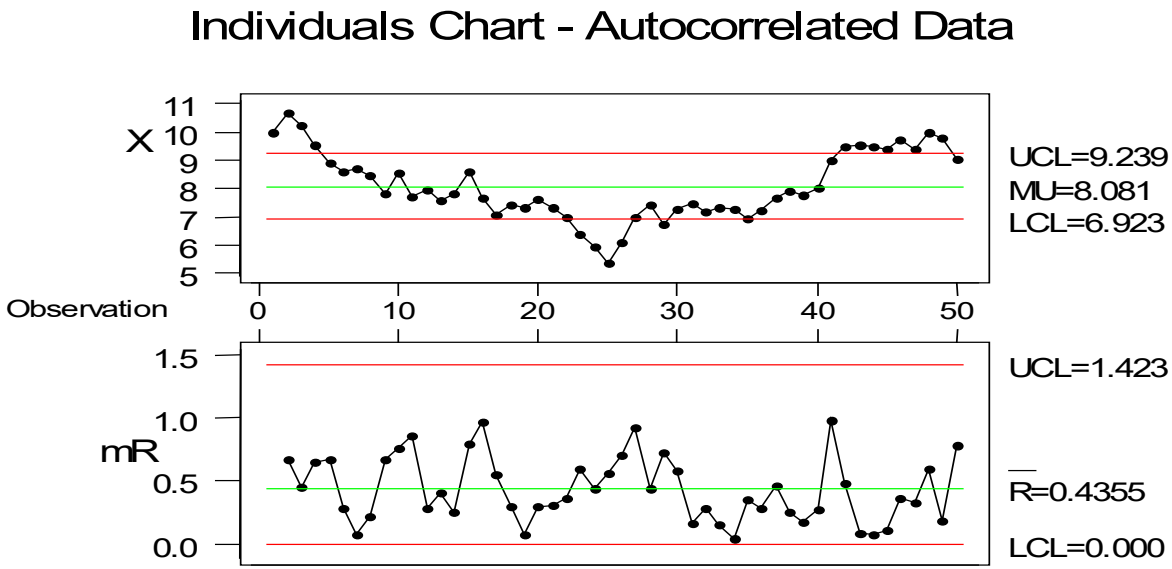


The picture of the data indicates the presence of correlation; the correlation coefficient is calculated to be 0.91, confirming the lag 1 auto-correlation. Examining the “lag 2” and “lag 3” auto-correlations are left as exercises for the interested reader.

The Effect of Auto-Correlated Data on a Control Chart

The data examined above is shown below on an X, mR control chart. In this case, the running record of the data by itself indicates that the process is not in a state of control. “Real world” processes will often give this same indication; the control limits applied to the individual values are not necessary to detect an out-of-control process. Notice the range chart. Since the original data was “constrained” to vary within a +/- 1 band from point to point, the range chart indicates a state of control. This is also behavior that can be expected of “real world” auto-correlated process data. Often, the *rate of*

change of the process will be constant from one time to another; this will produce a range chart that displays a state of control.



The lessons above can be summarized. If the process data is auto-correlated, then the individuals chart will often indicate this by displaying trends or "random walks" where the adjacent data do not vary significantly. This type of behavior should be a cause for considering the hypothesis of auto-correlation. Creating Scatter Diagrams (and calculating the associated correlation coefficient) for various "lags" can help confirm the presence of auto-correlation in the process. The analyst's experience and process knowledge should then be applied to determine the physical cause of the auto-correlation. This is one more signal that the process is not in a state of control. Examples of processes that may produce auto-correlated data include: queues or waiting lines, chemical processes (e.g. chemical concentration in a system following introduction, boiler corrosion chemical concentration), wear, fatigue, or deposition processes (e.g. tool sharpness, crack length in piping, tubing wall thickness.)

Methods of Handling Auto-Correlated Time Series Data

As mentioned above, the control limits for auto-correlated individuals data may not be necessary to detect the out-of-control behavior of the process. However, if they are to be used, an adjustment to the control limits' calculation can help correct for the auto-correlation.

If a “lag 1” auto-correlation exists, the moving ranges calculated from the individuals data tend to be smaller than those obtained from a process with little or no auto-correlation. This causes the average range to be smaller, which, in turn, produces calculated control limits that are also smaller (this should not be interpreted as a more “sensitive” control chart).

If the data is “lag 1” auto-correlated, the corrected, estimated standard deviation for the individual values can be approximated by the following:

$$\hat{\sigma}_X = \frac{\overline{R}}{d_2\sqrt{1-r^2}} = \frac{\overline{R}}{1.128\sqrt{1-r^2}}$$

where:

r - correlation coefficient, lag 1 data

This correction factor only becomes significant when the correlation coefficient is greater than about 0.7. A table of correction factors appears below for several values of *r:*

r	0.4	0.5	0.6	0.7	0.8	0.9	0.95
$1/\sqrt{1-r^2}$	1.1	1.15	1.25	1.4	1.6	2.3	3.2

This corrected, estimated standard deviation is used to calculate the upper and lower control limits for the individual values.

For cases where “lag 2,” or “lag 3” auto-correlation exists, the moving ranges likely will not be affected by the sequential dependency. No correction to the control limits is necessary.

A “lag 2” case may appear where a single waiting line is served by two clerks, who both take about the same time to process each individual. The time it takes an individual to wait is not so dependent on the person directly in front of them as it is on the *second* person in front of them.

Summary of Auto-correlated Process Data

If the process data is suspected of being auto-correlated, then the following analytic steps can be taken to “handle” this issue:

1. Develop a Scatter Diagram (and correlation coefficient) to determine if auto-correlation exists. Try a “lag 1” scatter first, then “lag 2,” and perhaps “lag 3.”

2. If a “lag 1” auto-correlation exists, and the correlation coefficient is higher than about 0.7, adjust the individual’s data control limits using the correction factor noted above.

3. Determine the source of the auto-correlation. Treat this as an out-of-control condition and take action to eliminate this factor, if possible.

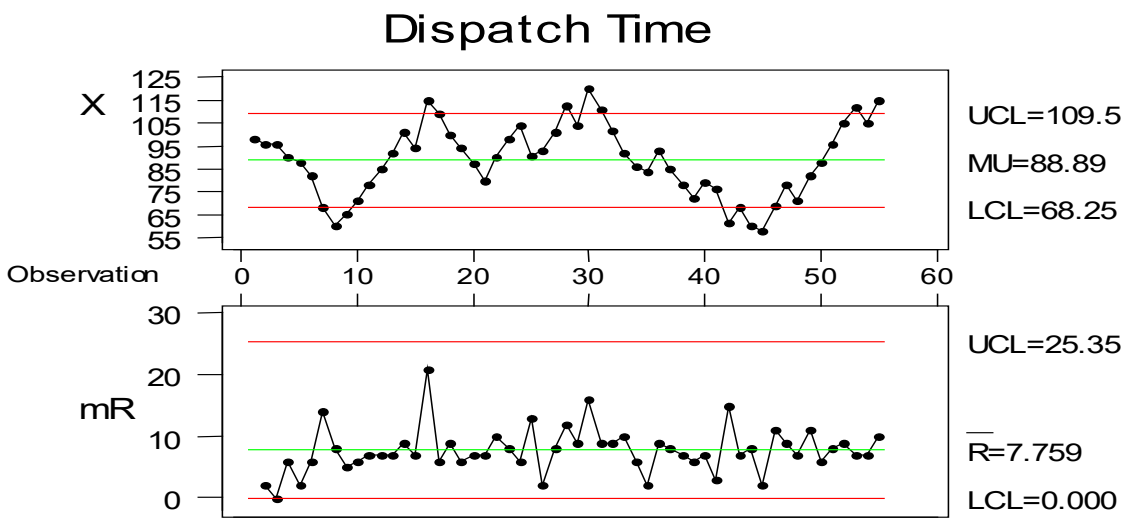
Auto-Correlation Example

A maintenance shop began tracking the time it took to dispatch a worker to the site of failed air conditioning equipment. The following data were collected over a two-day period:

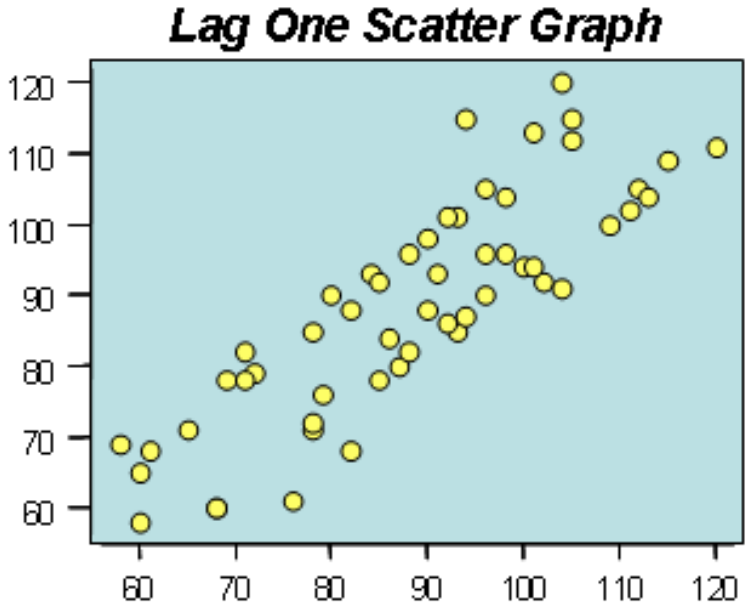
Total Dispatch Time (Minutes)

Failure	Dispatch Time	Failure	Dispatch Time	Failure	Dispatch Time
1	98	20	87	38	78
2	96	21	80	39	72
3	96	22	90	40	79
4	90	23	98	41	76
5	88	24	104	42	61
6	82	25	91	43	68
7	68	26	93	44	60
8	60	27	101	45	58
9	65	28	113	46	69
10	71	29	104	47	78
11	78	30	120	48	71
12	85	31	111	49	82
13	92	32	102	50	88
14	101	33	92	51	96
15	94	34	86	52	105
16	115	35	84	53	112
17	109	36	93	54	105
18	100	37	85	55	115
19	94				

The dispatcher took the data and prepared an X, mR control chart to examine the process for stability:



The data exhibited significant out-of-control conditions and appeared to be auto-correlated. The clerk tested this hypothesis by preparing a "lag 1" scatter diagram of the dispatch times:



The strong, positive correlation (r = 0.85) confirmed her suspicion of auto-correlation. Applying the correction factor (for r = 0.85, correction factor = 1.6) to the estimated standard deviation widens the control limits to UCL = 121.9 minutes and LCL = 55.9 minutes. No data are outside these adjusted limits.

6.9.8 Variable Control Limits for X-Bar, R Charts

Purpose

Although the X-Bar, S control chart can handle variable subgroup sizes, the calculations are somewhat daunting, especially when done by hand or with a non-programmable calculator. This topic shows how the X-Bar, R Control Chart can be modified to accommodate small, but variable subgroup sizes, usually of size 10 or less. For samples of size greater than 10, the X-Bar, S control chart is recommended[6].

Procedure

This procedure assumes that *k* subgroups of variables data have been collected from a process and that the *k* subgroups are of size 10 or less, but variable in size.

Subgroup	1	2	3	4	5	6	7	8	. . .	k
Data	x11 x12 x13 : : x1n	x21 x22 x23 : : : x2n	x31 x32 x33 : x3n	x41 x42 x43 : : : x4n	x51 x52 x53 : : x5n					xk1 xk2 xk3 : : xkn
Range	R1	R2	R3	R4	R5					Rk
Average	X1	X2	X3	X4	X5					Xk
Sub. Size	n1	n2	n3	n4	n5					nk
d2										
D3										
D4										

[6]Subgroup size 10 is not an arbitrary cutoff. Over 10, the range begins to lose its efficiency as a measure of variability (see Shewhart, Economic Control . . . , pg. 287, 288)

Range Chart

1. Calculate the Subgroup Ranges:

$$R_i = x_{i-\max} - x_{i-\min}$$

2. Calculate the Average Range of the subgroups:

$$\overline{R} = \sum_{i=1}^{k} n_i R_i \Big/ \sum_{i=1}^{k} n_i$$

3. Calculate the Upper and Lower Control Limits for each Subgroup Range. Use D_3 and D_4 coefficients specific to each subgroup (based on the subgroup size, n_i):

$$UCL_i = \overline{R} \times D_{4i}$$
$$LCL_i = \overline{R} \times D_{3i}$$

4. Plot the data, limits and centerline as usual. Interpret the chart.

X-Bar Chart

1. Calculate the Subgroup Averages:

$$\overline{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ji}$$

2. Calculate the *Pooled Standard Deviation* from the Ranges:

$$\sigma_p = \sqrt{\sum_{i=1}^{k} \frac{n_i R_i^2}{d^2{}_{2i}} \Big/ \sum_{i=1}^{k} n_i}$$

where:

d_{2i} - coefficient for converting Range (R)
to estimate of subgroup standard deviation
(varies as subgroup size varies)

3. Calculate the *Grand Average* of the Subgroups:

$$\overline{\overline{X}} = \frac{\sum_{i=1}^{k} n_i \overline{X}_i}{\sum_{i=1}^{k} n_i}$$

4. Calculate the Upper and Lower Control Limits for each subgroup:

$$\overline{\overline{X}} \pm 3\frac{\sigma_p}{\sqrt{n_i}}$$

5. Plot the data, limits and centerline as usual. Interpret the chart.

Variable Control Limits Example

A supervisor was comparing the cost per unit for suppliers (units are charged on a time & materials basis). She requested a report from Information Systems for the last two months and received the following information. Since the number of units built by each supplier was 10 or less, she decided to compare these costs on a variable subgroup size X-Bar, R control chart.

Supplier	Fitch	Hosmer	Gavankar	Stanton	Chaplin	Gibbs
Cost per	3077	1764	1024	1179	1238	7803
Unit ($):	5546	1255	2044	2459	1067	5287
	4134	916	2146	4616	1208	4995
	869	922	1472	1031	1342	3887
	901	5103	5105	6080	1063	4629
	2635	1016	989	1215	2505	6002
	2023	881	826	1384	2193	
	2274		2801	1173		
	2831		1203	1054		
			2096			
n:	9	7	10	9	7	6
Subgroup Average:	2698.9	1693.9	1970.6	2243.4	1516.6	5433.8
Subgroup Range:	4677	4222	4279	5049	1442	3916

The calculations' results that determine the Center Line and Control Limits for the Range and Subgroup Averages appear below:

Average Range: 4030.6

UCL - Range:	7319.6	7754.9	7162.4	7319.6	7754.9	8077.3
LCL - Range:	741.6	306.3	898.8	741.6	306.3	N/A

Pooled Standard Deviation: 1450.9

Grand Average: 2484.6

UCL - X-Bar	3935.5	4129.8	3861.0	3935.5	4129.8	4261.6
LCL - X-Bar	1033.7	839.4	1108.2	1033.7	839.4	707.6

The preparation of the control chart is left as an exercise for the student. By inspection, you can see that supplier Gibbs' is an assignable cause based on his average cost/unit being outside the upper control limit calculated for the subgroup averages.

All other suppliers should be considered part of the common cause system. The supervisor may want to investigate Gibbs' fabrication patterns to determine why the assignable cause exists.

6.10 Exercises

Note: The instructions for a number of these exercises reference the following Excel spreadsheet:



"Exercise Data.xls"

Objective: To develop definitions of Critical to Quality Characteristics (CTQs).

Instructions: 1. Determine one or more CTQ's for the products and services below:

Time: 20 minutes

Product/ Service	Customer need	Characteristic	Measure	Target	Specification(s)	Allowable Defect Rate
Fast Food	Quick Service					
Airline Travel	Luggage delivered to destination					
Air Conditioner	Reliability					
Hospital care	Correct medication					
Tuna Sandwich	Taste					

Objective: To develop operational definitions of some “common” quantities.

Instructions:

1. Review the definition of an Operational Definition.

2. Develop Operational Definitions for the 3 of the following:
   - Sunrise
   - 50% Cotton Shirt
   - On-Time Arrival of Airplane
   - Condenser Time to Failure
   - Blue
   - Speed of Light
   - Heat Exchanger Tube Failure (stress corrosion cracking)

Time: 20 minutes

Objective: To practice calculating basic descriptive statistics

Instructions: 1. Calculate the basic descriptive statistics from data collected in your class.

Time: 20 minutes

Data:

Measure	Calculation	Result
Mean		
Median		
Mode		
Range		
Variance		
Standard Deviation		

Objective: To practice calculating the Skewness and Kurtosis of a set of data.

Instructions:

1. For the set of data below, calculate the Skewness and Kurtosis values. Compare these values to those of a normal distribution.

2. Using Minitab or Excel; create a histogram of the data. Does the visual display agree with your interpretation?

Time: 20 minutes

18.92	18.63	16.08	16.15	20.64
18.76	16.43	19.63	16.33	26.78
18.24	14.38	17.58	22.34	17.71
18.89	22.41	15.02	16.14	17.59
24.13	22.83	17.05	17.39	39.17
19.38	10.41	23.96	12.50	19.59
19.02	11.59	23.20	26.61	19.07
20.21	21.94	28.95	13.77	22.77
10.85	11.26	17.81	16.07	15.55
10.62	18.52	16.91	16.13	21.22

Random Sample Exercise

From the data collected earlier about the class, select a random sample of 5 and determine the statistics.

Sample data

Measure	Calculation	Result
Mean		
Median		
Mode		
Range		
Variance		
Standard Deviation		

Interval Sample Exercise

From the data collected earlier about the class, select an interval sample of 5 and determine the statistics.

Sample data

Measure	Calculation	Result
Mean		
Median		
Mode		
Range		
Variance		
Standard Deviation		

Exercise - Line Graph:

A business unit track the number of jobs processed each week and the number of jobs that cannot pass inspection on the first attempt. Plot the number of jobs that cannot pass inspection on a line graph. What does this graph tell you?

Plot the fraction of jobs that cannot pass inspection. Does this graph tell a different story? Why?

First Pass Yield Job Log

Week	# Jobs/# Failing 1st time		Week	# Jobs/# Failing 1st time		Week	# Jobs/# Failing 1st time	
1	52	14	7	11	0	13	73	12
2	192	25	8	23	5	14	5	1
3	171	10	9	39	7	15	67	10
4	137	21	10	7	0	16	81	4
5	80	21	11	12	1	17	18	4
6	195	32	12	34	3	18	91	15

Exercise – Bar Graphs:

The following data were obtained from a consumer survey of products and services. Consumers were asked to categorize the products and services according to the "value" they thought they received. Plot the data on a bar chart. What conclusions do you reach?

Product or Service	Percent Saying "Good Value"
Doctor's Fees	34.7
Health Insurance	29.0
Hospital Charges	21.0
Lawyer's Fees	34.7
Poultry	66.4
Videotape Rentals	65.7
Women's Apparel	50.8

Exercise – Frequency Chart:

A quality improvement team is investigating the number of errors on warranty claim forms. They have collected the following data on the number of errors on each claim form. Plot this data on a frequency chart:

# Errors/Form	Frequency
0	50
1	40
2	72
3	116
4	52
5	23
6	12
7	3

Exercise – Histogram:

Piston Rings for Reciprocating Compressors are measured for width (in millimeters, outside diameter - inside diameter). Four measurements are taken, at 90-degree angles around the piston ring. Create a histogram of the entire data set. What does this tell you? Create histograms for each of the measurement positions. Are there any differences?

	Position (degrees)			
Ring	**0**	**90**	**180**	**270**
1	6.447	6.432	6.442	6.435
2	6.419	6.437	6.429	6.425
3	6.419	6.411	6.414	6.411
4	6.429	6.429	6.441	6.459
5	6.428	6.412	6.443	6.436
6	6.440	6.435	6.409	6.438
7	6.415	6.430	6.410	6.433
8	6.435	6.444	6.430	6.411
9	6.427	6.437	6.424	6.420
10	6.423	6.445	6.424	6.437
11	6.428	6.444	6.438	6.431
12	6.431	6.425	6.422	6.432
13	6.422	6.437	6.417	6.447
14	6.437	6.432	6.410	6.438
15	6.425	6.440	6.422	6.450
16	6.407	6.431	6.421	6.418
17	6.438	6.400	6.439	6.440
18	6.435	6.412	6.427	6.448
19	6.431	6.420	6.433	6.424
20	6.412	6.427	6.436	6.440
21	6.452	6.442	6.450	6.424
22	6.420	6.431	6.413	6.403
23	6.429	6.447	6.439	6.432
24	6.428	6.427	6.420	6.432
25	6.442	6.434	6.413	6.429

Exercise – Combining Variability:

Three components of a valve stem/gate assembly are produced. What is the expected length and standard deviation of the assembly? The three components are welded together in series:

Component	Mean	Std. Dev.
Valve Stem	18.00”	0.03”
Valve Disk	8.00”	0.02”
Valve Guide	4.00”	0.02”

If the specification calls for the assembly to be no longer than 30.10 inches, what is the current manufacturing process capability (C_p) of meeting the spec? (Note: typically, if the average plus/minus ***3*** times the standard deviation is within the spec limits, the process is considered OK).

Objective: To develop a Process Control Chart

Instructions:
1. Using John's pancake process, develop a process control system
2. Include the following elements: process steps, monitoring points (output, input, process), response (immediate remedy and recurrence prevention)

Time: 40 minutes

Process:	**CTQs:**
1. Gather Ingredients 2. Mix Dry Ingredients 3. Melt Butter (30 Sec. in Microwave) 4. Mix Butter, Eggs, Milk 5. Mix Liquid and Dry Ingredients 6. Heat Griddle to 375 F 7. Pour Batter in 3" Circles 8. Flip after ~ 1 minute 9. Serve when brown on both sides	• Size (USL = 6") • Thickness (USL = 0.5") • Served at Temperature (LSL = 130F) • Taste (Consumer Rating > 4.6/5) **Ingredients:** • 1 ½ Cup Flour • 3 Tbsp. Sugar • 1 ¾ Tsp. Baking Powder • 1 Tsp. Salt • 3 Tbsp. Butter • 1 – 2 Eggs • 1 – 1 ¼ Cup Milk

Objective: To develop and interpret an X-Bar, R control chart

Instructions:

3. Run your Card Drop Shop (using the standardized process) for 100 units.

4. Develop an X-Bar, R control chart for this data with a subgroup size of 4. Perform the calculations by hand; plot the points and limits on the control chart form. Interpret the control chart – are there assignable causes present?

5. Open Mini-Tab on your PC. Create the X-Bar, R control chart using the Card Drop Shop data. Compare results from Mini-Tab to your hand-drawn chart.

Time: 40 minutes

Objective: To develop and interpret an X-Bar, R control chart

Instructions:

1. Develop an X-Bar, R control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form. Interpret the control chart – are there assignable causes present?

2. Open Mini-Tab on your PC. Create the X-Bar, R control chart using the STUDS data file. Compare results from Mini-Tab to your hand-drawn chart.

Time: 40 minutes

Compressor Stud Lengths – A supplier fabricates studs for critical compressor applications. One key quality characteristic of the studs is their length. Their customer specifications call for a nominal value of 5.3750" with a tolerance of +/- 0.0005". The supplier pulls a subgroup of four studs each hour from the fabrication process and measures their length with a calibrated micrometer. In the table below, each row is a subgroup.

Subgroup	Stud Length (in.)			
1	5.37526	5.37525	5.37454	5.37464
2	5.37478	5.37525	5.37495	5.37411
3	5.37446	5.37476	5.37482	5.37492
4	5.37525	5.37485	5.37527	5.37506
5	5.37463	5.37430	5.37502	5.37523
6	5.37511	5.37473	5.37486	5.37478
7	5.37473	5.37475	5.37510	5.37480
8	5.37484	5.37497	5.37480	5.37498
9	5.37520	5.37457	5.37432	5.37484
10	5.37534	5.37487	5.37511	5.37517
11	5.37472	5.37433	5.37526	5.37486
12	5.37502	5.37501	5.37532	5.37502
13	5.37475	5.37542	5.37462	5.37473

Subgroup	Stud Length (in.)			
14	5.37482	5.37529	5.37539	5.37475
15	5.37499	5.37504	5.37515	5.37515
16	5.37464	5.37509	5.37458	5.37476
17	5.37465	5.37487	5.37456	5.37472
18	5.37515	5.37492	5.37504	5.37519
19	5.37440	5.37531	5.37504	5.37525
20	5.37436	5.37475	5.37516	5.37474
21	5.37493	5.37514	5.37471	5.37481
22	5.37463	5.37467	5.37511	5.37510
23	5.37511	5.37510	5.37530	5.37477
24	5.37436	5.37401	5.37525	5.37493
25	5.37483	5.37493	5.37448	5.37518

Objective: To practice developing and interpreting an X-Bar, S Control Chart.

Instructions:

1. Develop an X-Bar, S control chart for the data on the following pages. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the X-Bar, S control chart using the DELIVERY data files. Compare results from Mini-Tab to your hand-drawn charts.

Time: 40 minutes

Delivery Schedule Data by Month - To provide better on time delivery and to increase throughput a plant is monitoring DAYS LATE FOR DELIVERY. Take the following delivery data for Unit 1075B and prepare an X-Bar, S Control Chart. Each data point is for a late unit. Interpret the data when the month is used as a subgroup strategy:

APR	MAY	JUN	JUL	AUG
1	1	1	1	2
4	1	3	1	3
4	1	1	1	2
10	1	2	6	2
3	2	12	2	7
1	4	1	5	3
2	2	4	2	3
1	4	1	2	1
1	2	1	1	1
2	1	1	1	4
4	3	5	4	
1	1	4	1	
1	1	2	1	
	4	1	4	
	2	3	4	
	6	2	1	

REPAIR COSTS DATA BY TECHNICIAN

The following data represents the costs of repair for service technicians. Take the following data and create an X-Bar, S Control Chart (this is sheet REPAIR in your Excel exercise file). Interpret the results using subgrouping by technician strategy.

Cost ($)	Technician	Cost ($)	Technician
1520	JONES, ROBERT L.	2044	CONAN, MICHAEL B.
3227	JONES, ROBERT L.	2146	CONAN, MICHAEL B.
4618	RHEW, ROBERT I.	1472	CONAN, MICHAEL B.
8480	RHEW, ROBERT I.	5105	CONAN, MICHAEL B.
4239	RHEW, ROBERT I.	989	CONAN, MICHAEL B.
3077	BIRD, LARRY E.	826	CONAN, MICHAEL B.
5546	BIRD, LARRY E.	2801	CONAN, MICHAEL B.
4134	BIRD, LARRY E.	1203	CONAN, MICHAEL B.
869	BIRD, LARRY E.	2096	CONAN, MICHAEL B.
901	BIRD, LARRY E.	1121	CONAN, MICHAEL B.
2635	BIRD, LARRY E.	1609	CONAN, MICHAEL B.
2023	BIRD, LARRY E.	3175	MERTZ, MILDRED E.
2274	BIRD, LARRY E.	1794	MERTZ, MILDRED E.
2831	BIRD, LARRY E.	1300	MERTZ, MILDRED E.
1840	BIRD, LARRY E.	1179	EASTWOOD, DAVID H.
966	BIRD, LARRY E.	2459	EASTWOOD, DAVID H.
2367	BIRD, LARRY E.	4616	EASTWOOD, DAVID H.
981	BIRD, LARRY E.	1031	EASTWOOD, DAVID H.
2520	BIRD, LARRY E.	6080	EASTWOOD, DAVID H.
1764	FORD, DALE	1215	EASTWOOD, DAVID H.
1255	FORD, DALE	1384	EASTWOOD, DAVID H.
916	FORD, DALE	1173	EASTWOOD, DAVID H.
922	FORD, DALE	1054	EASTWOOD, DAVID H.
5103	FORD, DALE	2062	EASTWOOD, DAVID H.
1016	FORD, DALE	1668	EASTWOOD, DAVID H.
881	FORD, DALE	2103	EASTWOOD, DAVID H.
2612	CHRYSLER, ROBERT M.	4016	EASTWOOD, DAVID H.
2437	CHRYSLER, ROBERT M.	1563	EASTWOOD, DAVID H.

Cost ($)	Technician	Cost ($)	Technician
1024	CONAN, MICHAEL B.	1454	SCHRODINGER, STEVEN E.
2285	SCHRODINGER, STEVEN E.	1207	HOSOYAMADA, MICHAEL E.
720	SCHRODINGER, STEVEN E.	888	HOSOYAMADA, MICHAEL E.
1618	SCHRODINGER, STEVEN E.	2246	HOSOYAMADA, MICHAEL E.
1902	SCHRODINGER, STEVEN E.	1019	HOSOYAMADA, MICHAEL E.
857	SCHRODINGER, STEVEN E.	1508	HOSOYAMADA, MICHAEL E.
1620	SCHRODINGER, STEVEN E.	1665	HOSOYAMADA, MICHAEL E.
1093	SCHRODINGER, STEVEN E.	1311	HOSOYAMADA, MICHAEL E.
1273	SCHRODINGER, STEVEN E.	1718	WASHINGTON, BARRY
865	SCHRODINGER, STEVEN E.	933	WASHINGTON, BARRY
2609	SCHRODINGER, STEVEN E.	2266	HINCKLEY, ANDERSON M.
1981	SCHRODINGER, STEVEN E.	1979	HINCKLEY, ANDERSON M.
1171	SCHRODINGER, STEVEN E.	1132	HINCKLEY, ANDERSON M.
840	SCHRODINGER, STEVEN E.	876	HINCKLEY, ANDERSON M.
971	SCHRODINGER, STEVEN E.	1446	HINCKLEY, ANDERSON M.
1361	ARAFAT, YOUSSEF	2083	HINCKLEY, ANDERSON M.
1829	ARAFAT, YOUSSEF	1138	HINCKLEY, ANDERSON M.
4368	ARAFAT, YOUSSEF	995	HINCKLEY, ANDERSON M.
2795	ARAFAT, YOUSSEF	1212	HINCKLEY, ANDERSON M.
1632	ARAFAT, YOUSSEF	980	HINCKLEY, ANDERSON M.
1046	ARAFAT, YOUSSEF	1652	DIRAC JR., THOMAS O.
1335	ARAFAT, YOUSSEF	1298	DIRAC JR., THOMAS O.
2111	ARAFAT, YOUSSEF	1910	DIRAC JR., THOMAS O.
2141	ARAFAT, YOUSSEF	820	DIRAC JR., THOMAS O.
1814	NEWTON, JOHN D.	1270	DIRAC JR., THOMAS O.
3832	NEWTON, JOHN D.	1069	DIRAC JR., THOMAS O.
1238	NEWTON, JOHN D.	1112	DIRAC JR., THOMAS O.
1067	NEWTON, JOHN D.	3595	DIRAC JR., THOMAS O.
1208	NEWTON, JOHN D.	1139	DIRAC JR., THOMAS O.
1342	NEWTON, JOHN D.	1076	DIRAC JR., THOMAS O.
1063	NEWTON, JOHN D.	2867	DIRAC JR., THOMAS O.
2505	NEWTON, JOHN D.	4417	DIRAC JR., THOMAS O.
2193	SHERMAN, BRYAN		
6124	SHERMAN, BRYAN		

Objective: To practice developing and interpreting an X, mR Control Chart.

Instructions:

1. Develop an X, mR control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the X, mR control chart using the VALVE data file. Compare results from Mini-Tab to your hand-drawn charts.

Time: 25 minutes

Butterfly Control Valve - An air-operated butterfly control valve is used to control cooling water flow to heat exchangers in an air conditioning unit. The valve must close within ten seconds of receipt of the signal from the unit's protective circuitry. The valve is tested monthly and maintenance technical personnel record its closing time (in seconds):

2.87	4.2
1.96	4.82
2.22	4.58
1.51	5.81
5.04	4.27
3.67	2.22
2.62	2.65
4.61	4.52
4.46	3.62
3.95	2.86
4.36	3.81
4.16	3.91
4.08	

Objective: To practice developing and interpreting an X, mR Control Chart.

Instructions:

1. The following data are the results of opinion polls taken in the three months prior to the 2000 US election. Each data is the percent favoring Governor Bush minus the percent favoring Vice President Gore. Are there assignable causes present (i.e. due to convention "lift", kissing a wife, etc.)?

Time: 20 minutes

Bush minus Gore – Polling Data

2	3
17	-3
-1 (August)	0
1	0
-3	-3
2	-13
-7	7 (October)
-6	8
-7.5	4
-7	0
-5.5 (September)	-1
-9	0
-11	4
4	5
3	3
0	

Objective: To practice developing and interpreting an X, mR Control Chart.

Instructions:

1. The following data were taken from the Raleigh News and Observer newspaper. The article implied that a high number of births were due to the Hurricane Floyd which occurred in October of the previous year (the middle column shows the nine month "lag" between the hurricane and the births). Is there evidence that Hurricane Floyd is an assignable cause of variation?

Time: 20 minutes

Births by Month - Wake County		
Month	***9M Lag***	***Births***
Jan	Apr	44427
Feb	May	40821
Mar	Jun	44336
Apr	Jul	42101
May	Aug	43746
Jun	Sep	44076
Jul	Oct - Floyd	47085
Aug	Nov	47995
Sep	Dec	47050
Oct	Jan	45738
Nov	Feb	43087
Dec	Mar	45471

Objective: To practice developing and interpreting an X, mR Control Chart.

Instructions:

1. The following data were obtained at a company cafeteria in Osan, Republic of Korea. The cafeteria staff measured the quality of their food by how much was thrown away each day (average of food on the plates in grams).

2. Develop and interpret an X, mR control chart for this data.

Time: 20 minutes

Day	Waste (gms)	Day	Waste (gms)
1	28	13	28
2	26	14	26
3	16	15	40
4	18	16	25
5	20	17	27
6	21	18	19
7	25	19	22
8	34	20	25
9	18	21	24
10	26	22	26
11	36	23	25
12	28	24	24

Objective: To develop and interpret an *np* control chart

Instructions:

1. Run the Card Drop Shop for 100 units in subgroups of size 10. Record the number of cards that fall outside of 20 inches from the target as *defectives.*

2. Develop an *np* control chart for this data. Perform the calculations by hand; plot the points and limits on the control chart form.

3. Open Mini-Tab on your PC. Create the *np* control chart. Compare results from Mini-Tab to your hand-drawn charts.

Time: 40 minutes

Objective: To develop and interpret an *np* control chart

Instructions:

1. Develop an *np* control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *np* control chart using the MOTOR data file. Compare results from Mini-Tab to your hand-drawn charts.

Time: 30 minutes

Motor Rejects - A company that produces air conditioning units orders batches of the motors from a supplier. Due to past quality problems, the company inspects 20 motors from each batch. Each motor is accepted or rejected. Based on the number of motors rejected, a decision is made to either inspect the remaining motors or return the batch to the supplier for rework.

Number of Rejected Motors:

5	4
2	3
3	1
3	2
3	0
1	0
4	6
3	4
5	4
2	6
1	3
2	

Objective: To develop and interpret an *p* control chart

Instructions:

1. Develop a *p* control chart for the data on the following pages. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *p* control chart using the WELDS data file. Compare results from Mini-Tab to your hand-drawn charts.

Time: 30 minutes

Defective Full-Penetration Welds

A welding supervisor receives inspection reports by the Quality Control Department. The QC supervisor has recently called his attention to a seemingly high number of rejected full-penetration welds on critical high pressure piping systems. The welding supervisor begins his analysis of the situation by preparing a p-chart of rejected welds for the past six months.

Week	# Welds	# Defective		Week	# Welds	# Defective
1	476	41		14	352	36
2	379	40		15	415	39
3	412	42		16	557	60
4	424	48		17	581	51
5	483	44		18	466	57
6	415	48		19	584	54
7	541	55		20	573	66
8	544	50		21	471	51
9	466	39		22	305	49
10	439	37		23	383	44
11	428	40		24	379	47
12	363	31		25	526	59
13	463	57		26	543	66

Objective: To develop and interpret a *c* control chart

Instructions:

1. Develop a *c* control chart for the data on the following pages. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *np* control chart using the PINHOLES data file. Compare results from Mini-Tab to your hand-drawn charts.

Time: 25 minutes

Ceramic Paint Pinholes - A paint manufacturing company, which produces special paints used by hobbyists on ceramics, tests samples of their paint daily. They apply the paint to unfired ceramic plates, fire the plates in a kiln and then inspect the finished plates. Among other defect categories, they count the number of pinholes in each sample. The test manager has recently begun to track the number of pinholes obtained from each sample on a c chart.

Number of Pinholes per Sample:

18	14
8	15
17	17
16	13
20	17
10	17
19	16
19	13
13	6
10	16
21	19
12	22
13	14

Objective: To develop and interpret a *u* control chart

Instructions:

1. Develop a *u* control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *u* control chart using the same data using the DCRs data file. Compare results from Mini-Tab to your hand-drawn charts.

Time: 25 minutes

Design Change Requests (DCRs) - Engineers are responsible for developing custom designs of air conditioning systems. As they are built, manufacturing discovers problems with the designs and requests changes from Engineering (Design Change Request). The Engineering Manager was curious to see if there were significant differences between the engineers.

# Units	# DCRs	Engineer
40	97	Maynard
90	69	Kinney
90	153	Gibbs
70	125	Nichols
30	45	Fritz
50	66	Stone
40	62	Fielding
10	25	Adams
70	82	Pelham

Objective: To understand the types of tampering and their impact on process variation

Instructions:

1. Run the Card Drop Shop for 100 repetitions, employing the four adjustment rules described below (400 total runs):

a) Rule 1: Aim the card over the target for all runs. Measure the distance (z) and angle (Φ) from the target.
b) Rule 2: Measure the distance (z) and angle (Φ) from the first dropped card to the target. Move your aim from its initial point to one opposite this distance and angle. Apply this rule to successive runs.
c) Rule 3: Measure the distance (z) and angle (Φ) from the first dropped card to the target. Move your aim to a point opposite this distance and angle as measured from the target. Apply this rule to successive runs.
d) Rule 4: After each drop, move your aim to where the card just landed. Measure the distance (z) and angle (Φ) from the target for each card.

2. Analyze the data three ways:

a) Plot the data on a flipchart. Mark where each card lands relative to the target considering both distance and angle.
b) Using the distances, open Mini-Tab and create a histogram of the data. Calculate the mean distance and standard deviation
c) Compare the results of each rule. Which “produces” the minimum dispersion from the target?

3. Consider these rules. How are they applied in your business? Provide examples as part of the exercise debrief.

Time: 40 minutes

Pick-a-Chart (Control Chart Selection)

Often, one of the difficulties people face with control charts is the question of “Which is the right one?” This exercise is intended to give you some practice going through the logic of the Control Chart Selection Guide (Unit 6.4).

As you develop your answer, note your assumptions. There is more than one way many of these scenarios could be charted. We’ll start out with some “warm-up” exercises, move on to more complex situations:

Scenario	Control Chart
1. Each day, the number of units shipped is counted at 12:00 AM.	
2. The Sales department keeps track of the number of units sold each day, by type of unit.	
3. A laboratory gets a report, once a day, which provides the number of samples analyzed, the average processing time, and the standard deviation of the processing time.	
4. Each day, a technician measures the time she takes tubing a condenser on her shift.	
5. At a ballpark, a hot dog vendor counts the number of wieners sold each game day.	
6. A factory worker measures the diameter of valve stems after machining. She takes four stems at random from each hour's production.	
7. An engineer measures the cycle time for engineering change orders weekly.	
8. An administrative assistant tracks the number of days it takes customers to pay their bills. She keeps a chart for each of the company's top 6 customers.	
9. A quality consultant tracks the number of days she is on the road each month.	
10. A Sales supervisor has developed control charts for her clerks - they track the number of line items entered each day.	
11.The power of a motor is measured and is subject to a purchase specification.	
12. Coatings are purchased in tank car lots. The material is sampled and the chemical composition determined.	
13. The procedures group has noticed an increase in the number of comments made on their draft procedures being circulated for review. They are wondering if something unusual is going on in the procedure drafting/ review process.	
14. The LAN (Local Area Network) administrator has been trying to improve the reliability of the system. She is interested in seeing if the number of LAN "crashes" has decreased.	
15. This same LAN administrator has also been working on trying to reduce the time required to restore the LAN after it crashes.	
16. The Production Manager is interested in employee absenteeism, measured in days/employee. The corporate staff supplies her with a monthly report, which breaks down this measure into weekly increments.	
17. The Financial Officer is concerned about the utilization of company cars; he suspects that there are too many cars. He begins tracking the number of hours the cars are utilized for business purposes each week.	

Scenario	Control Chart
18. A production facility wishes to improve the set-up time required when products being produced are changed. They usually make the same product for about two days and then switch over to another product.	
19. A bolt manufacturer must ensure that the tensile strength of stainless steel bolts meets the customers' specifications. About 5000 bolts are produced daily.	
20. You have been troubled by the number of times your production facility has been stopped due to power interruptions by the local utility. You have records of all production stoppages for the last two years.	
21. A certain vendor provides you with bolts for your product. Before the bolts are used in your production process, you sample 50 from each box of 1000 and inspect them for defects.	
22. A large consulting firm prepares about 30 proposals per week for prospective clients. The Sales Department manager is interested in the number of proposals that are not accepted by clients.	
23. An engineering department prepares design changes to improve the performance of a chemical processing plant. They are interested in the number of *field change requests*, those changes that are requested by construction engineering because the design change cannot be implemented in the field.	
24. A Sales Manager tracks weekly sales volumes by number of items sold, dollar amount of sales and items sold per salesperson.	
25. An Automotive Manager is concerned about the quality of a particular brand of tire used on company cars. His primary concern is the possibility of a tire blowout. If the size of the company car fleet stays constant, how should he track this process?	
26. A Records department director is concerned about the errors made by her staff. She asks for help in determining the best chart to use. She tells you that the number of records varies significantly from week to week.	
27. Each month you receive a departmental budget variance report that, among other things, provides the dollar amount you are over or under salary budget, supply expense budget and overtime hours.	
28. A physician thinks that the complications associated with a particular surgical procedure varies from surgeon to surgeon. Each surgeon does a different number of these procedures each year.	
29. A company is interested in using a new vendor to supply control circuits that emit a specified signal for a specified time. They wish to determine if the process used to produce the circuits is in control. They are particularly interested in the signal’s duration.	

Control Chart Setup Exercises

For the following scenarios, determine how you would setup and run a control chart. Consider not only what type of control chart to use, but also how to collect and analyze the process data.

1. A chemical laboratory processes hundreds of samples daily and is concerned about the Turn-Around-Time (TAT) of their process. Although there are many different types of analyses, the "Pareto" principle applies – there are three most frequent analyses performed: a) metallurgical, b) tensile strength, and c) contaminants. The laboratory runs three shifts daily. Each sample arriving at the lab is bar-coded. When the sample arrives, the lab's computer collects the start time. When the sample is finished, the time is entered and the computer calculates the turn-around-time. A standard report provides the average turn-around time for each shift. Weekend volumes tend to be about one half the weekdays.

2. A sales distributor is concerned about errors made in processing orders for units for their clients. Sales personnel meet with the client and take the order. The order is then processed by the Engineering Department, who "translate" the customer order into specifications for the manufacturer. The Audit Department checks each order for accuracy and the manufacturer will also call the Engineering Department if there is a problem with the order. The distributor operates nationwide through local districts; they process about 1000 orders a year (note that the number of units associated with each order varies from 1 to 15).

3. A truck-leasing company is concerned about errors made in processing orders for new vehicles for their clients. Sales personnel meet with the client and take the truck order. The order is then processed by the Specifications Department, who "translate" the customer order into specifications for the truck manufacturer. An Audit Department checks each order for accuracy, and the manufacturer will also call the Specifications Department if there is a problem with the order. The leasing company operates nationwide through local districts; they process about 1000 truck orders a year (note that the number of trucks associated with each order varies from 10 to 150).

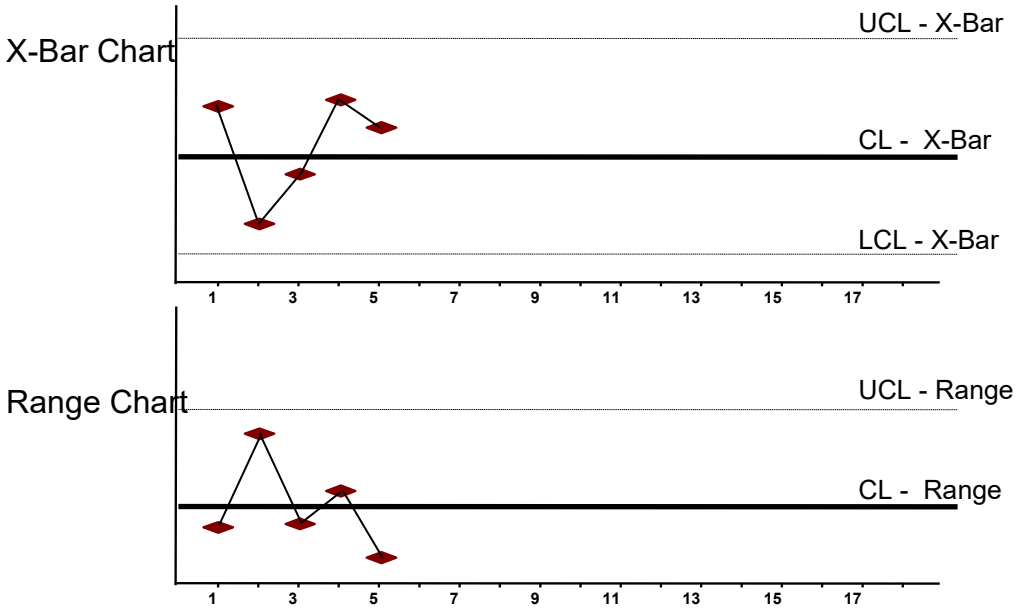
4. An air-conditioning manufacturer is concerned about failures occurring in brazing blades to compressor impellers. Twelve blades are brazed to each impeller wheel (2 wheels per impeller with the blades in the middle – 24 braze surfaces total). After the brazing process, the impeller is inspected using an ultrasonic process. The inspection provides the fraction of blade/impeller surface that has been brazed (if at least 75% of the surface is brazed, the braze passes). If there is insufficient brazing material *on at least one braze surface,* the impeller is rejected.

5. Sheet steel is delivered to a plant in railroad cars. The company's contract with the supplier specifies the thickness of the steel as well as the tensile strength. Each load consists of about 150 rolls of steel. Deliveries occur every three days.

6. Coal is delivered to a power plant in railroad cars. The utility's contract with the supplier specifies the average size of the coal as well as the maximum sulfur content. Each train load consists of about 150 cars. Deliveries occur every three days.

Control Chart Application Example

Consider the following scenario. A manufacturing plant runs a two-shift operation. Ten parts are produced each shift. The process control plan calls for maintaining the current process settings until the control chart displays assignable cause signals. All parts are measured and plotted *real-time* on an X-Bar, R control chart, with subgroup size = 2. The following control chart shows the data from the last shift. The control limits are based on the previous 5 days of production, not including the last shift. For each of the three scenarios described below, discuss and predict what the data would look like on the chart.



Scenario 1 – The instrument is checked at the beginning of the second shift. Due to a bias noted against the plant's standard, the gauge is adjusted noticeably higher (e.g. for a part previously measured to be 0.900", the new reading would be 1.000"), prior to production. Sketch the next two shifts.

Scenario 2 – With the second shift, a new gauge is introduced. Compared to the old gauge, the bias is the same, but the gauge variation is much less. Sketch the next two shifts.

Scenario 3 – A new operator starts work on the second shift. He tends to read the gauge lower than the other operators, although no one knows this since a measurement system study has not been performed. Sketch the next three shifts.

Objective: To apply control charts to detection of process improvement/change.

Instructions: 1. The following data shows a process' performance Before/After Improvement. Show three pictures of the data (this data is in the PROCEDURE worksheet of your Excel file):

a) Control Limits calculated for **all** the data,

b) Control Limits calculated based on just the "Before" data (but showing all the points), and

c) Control Limits calculated for the "Before" data as well as control limits for the "After" data (on the same graph).

Time: 20 minutes

Procedure Time (Minutes):

Before				After		
29	22	25		11	18	16
32	27	22		22	27	25
43	33	28		32	19	24
29	26	36		13	35	16
26	34	24		18	19	20
25	30	32		9	15	27
28	25	26		24	24	36
28	29	30		15	21	25
28	23	25		25	14	21
27	28	24		22	26	14
48	27	45		30	27	5
31	29	37		21	18	14
28	34			16	24	13
				27	23	19
				13	13	34

Objective: To apply control charts to detection of process improvement/change.

Instructions: 1. The following data represent the number of railroad crashes (involving one or more trains) that have occurred since the British Railway system was privatized. Can management claim to have improved safety on their lines?

Time: 20 minutes

Rail Crashes	Year
510	89
480	90
320	91
350	92
250	93
275	94
215	95
225	96
210	97
220	98
185	99

Objective: To have fun with control charts.

Instructions:

1. The following data were collected from a torque operation. Six spindles of a multi-driver tighten bolts on an assembly. The assembly is then dehydrated and the torque again measured. What questions could you generate about this process? What answers might come from looking at the data on a control chart? Try to answer your questions.

2. On the next page, similar data is listed. This data was collected from a machine being evaluated to replace the current multi-driver. Repeat question one on this data.

Time: 20 minutes

H29BU Thick Valve Plate Gasket, .034"

Torque Target: 25 ft-lb.

Machine A Line Multi-Driver																															
Assembly	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***13***	***14***	***15***	***16***	***17***	***18***	***19***	***20***	***21***	***22***	***23***	***24***	***25***	***26***	***27***	***28***	***29***	***30***	***31***	***32***
Before Dehydration																																
Spindle-1	21.0	21.0	21.0	22.0	21.0	21.0	20.0	20.0	20.0	19.0	20.0	22.0	20.0	20.0	21.0	22.0	18.0	21.0	21.0	21.0	20.0	20.0	21.0	18.0	20.0	21.0	20.0	21.0	20.0	25.0	23.0	20.0
Spindle-2	23.0	22.0	24.0	23.0	22.0	22.0	20.0	21.0	20.0	24.0	23.0	23.0	24.0	20.0	23.0	24.0	21.0	18.0	26.0	26.0	20.0	22.0	19.0	22.0	25.0	22.0	26.0	24.0	20.0	25.0	22.0	21.0
Spindle-3	24.0	21.0	23.0	21.0	23.0	20.0	20.0	21.0	20.0	20.0	21.0	22.0	23.0	20.0	23.0	22.0	20.0	20.0	22.0	22.0	19.0	22.0	22.0	22.0	22.0	23.0	21.0	23.0	20.0	25.0	22.0	22.0
Spindle-4	21.0	22.0	22.0	20.0	21.0	20.0	20.0	23.0	21.0	19.0	21.0	21.0	20.0	19.0	22.0	21.0	19.0	19.0	21.0	21.0	20.0	21.0	20.0	21.0	20.0	22.0	22.0	20.0	20.0	25.0	20.0	23.0
Spindle-5	21.0	23.0	23.0	21.0	23.0	20.0	23.0	19.0	22.0	20.0	21.0	21.0	20.0	21.0	22.0	19.0	21.0	19.0	21.0	20.0	19.0	18.0	20.0	20.0	20.0	19.0	20.0	15.0	22.0	25.0	21.0	20.0
Spindle-6	21.0	22.0	23.0	19.0	21.0	20.0	19.0	22.0	20.0	20.0	21.0	20.0	20.0	10.0	21.0	18.0	19.0	19.0	18.0	28.0	19.0	24.0	20.0	19.0	20.0	23.0	25.0	24.0	21.0	25.0	18.0	25.0
After Dehydration																																
Spindle-1	10.0	10.0	10.0	10.0	10.0	10.0	8.0	11.0	8.0	10.0	9.0	10.0	9.0	10.0	10.0	12.0	9.0	10.0	10.0	12.0	10.0	8.0	10.0	8.0	9.0	10.0	10.0	10.0	10.0	13.0	12.0	10.0
Spindle-2	13.0	12.0	13.0	12.0	13.0	11.0	7.0	12.0	11.0	14.0	16.0	13.0	16.0	9.0	12.0	15.0	10.0	6.0	19.0	13.0	13.0	10.0	8.0	12.0	18.0	12.0	15.0	14.0	8.0	14.0	13.0	10.0
Spindle-3	15.0	8.0	11.0	7.0	16.0	8.0	6.0	14.0	11.0	9.0	6.0	10.0	8.0	8.0	12.0	10.0	9.0	10.0	10.0	11.0	9.0	10.0	13.0	10.0	9.0	10.0	9.0	10.0	10.0	12.0	12.0	11.0
Spindle-4	11.0	15.0	13.0	10.0	10.0	11.0	10.0	15.0	11.0	7.0	10.0	11.0	10.0	10.0	14.0	14.0	11.0	12.0	8.0	11.0	13.0	12.0	12.0	14.0	10.0	14.0	15.0	10.0	11.0	17.0	10.0	15.0
Spindle-5	14.0	17.0	15.0	11.0	16.0	11.0	18.0	9.0	16.0	10.0	13.0	12.0	11.0	12.0	14.0	10.0	14.0	12.0	14.0	12.0	9.0	7.0	11.0	10.0	11.0	10.0	12.0	6.0	15.0	14.0	15.0	12.0
Spindle-6	10.0	12.0	15.0	8.0	10.0	8.0	8.0	12.0	8.0	8.0	11.0	8.0	9.0	7.0	12.0	8.0	10.0	8.0	7.0	15.0	8.0	16.0	11.0	8.0	11.0	15.0	15.0	14.0	14.0	16.0	10.0	20.0

Machine AAG #800864-05																																
Assembly	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*
Before Dehydration																																
Spindle-1	25.0	25.0	25.0	24.0	25.0	25.0	26.0	25.5	26.0	24.5	25.0	25.5	25.0	26.5	25.0	25.0	25.5	26.0	25.0	25.0	25.5	26.0	25.5	25.0	26.0	26.0	26.0	25.5	26.0	23.5	25.5	25.5
Spindle-2	26.5	26.5	26.0	26.0	26.0	25.5	26.0	26.5	26.0	26.0	27.0	27.0	24.5	27.0	28.0	26.5	27.5	25.5	26.5	27.5	28.0	27.0	27.0	26.0	27.0	28.0	27.5	26.0	28.0	27.0	28.0	26.5
Spindle-3	24.0	27.0	26.0	24.5	25.0	25.5	25.5	25.5	25.5	24.0	25.5	25.5	25.5	25.5	25.0	26.0	26.0	25.0	25.0	26.5	26.0	25.0	26.0	25.5	26.5	26.0	26.5	26.5	26.5	25.0	26.0	25.0
Spindle-4	24.5	25.0	25.0	25.0	24.5	24.5	25.0	25.5	25.5	25.0	25.5	26.0	25.5	25.0	25.0	25.5	26.0	24.5	26.0	26.0	26.0	25.5	26.0	26.0	25.0	26.0	27.0	25.5	26.0	25.5	26.0	26.5
Spindle-5	24.5	24.0	25.0	24.0	25.0	24.5	25.0	25.5	24.0	24.5	24.5	24.0	24.0	24.0	24.0	26.0	25.0	25.0	25.0	25.0	25.0	25.5	25.0	25.0	25.0	25.5	24.0	25.5	26.0	25.0	25.0	24.5
Spindle-6	24.0	25.0	25.5	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.5	24.5	25.0	25.0	25.0	25.0	25.5	26.0	26.0	25.0	25.5	25.0	25.5	25.5	26.0	26.0	25.0	26.0	24.5	26.0	26.0	25.5
After Dehydration																																
Spindle-1	18.0	20.0	18.0	23.0	17.0	19.0	20.0	15.0	22.0	23.0	15.0	20.0	21.0	20.0	22.0	19.0	18.0	19.0	17.0	21.0	20.0	19.0	23.0	22.0	18.0	21.0	17.0	18.0	19.0	24.0	16.0	18.0
Spindle-2	20.0	21.0	25.0	21.0	22.0	23.0	23.0	18.0	18.0	22.0	18.0	22.0	22.0	22.0	24.0	21.0	21.0	24.0	18.0	22.0	23.0	24.0	24.0	22.0	19.0	20.0	18.0	23.0	23.0	23.0	20.0	23.0
Spindle-3	22.0	20.0	20.0	23.0	21.0	18.0	15.0	16.0	18.0	21.0	18.0	20.0	19.0	18.0	16.0	18.0	19.0	20.0	17.0	21.0	19.0	16.0	18.0	20.0	16.0	18.0	17.0	19.0	19.0	22.0	18.0	17.0
Spindle-4	20.0	20.0	20.0	19.0	20.0	19.0	20.0	18.0	21.0	22.0	21.0	22.0	21.0	20.0	19.0	23.0	20.0	18.0	21.0	22.0	21.0	21.0	21.0	21.0	19.0	20.0	21.0	19.0	23.0	21.0	20.0	22.0
Spindle-5	20.0	20.0	21.0	19.0	22.0	20.0	18.0	19.0	16.0	18.0	20.0	21.0	20.0	20.0	19.0	20.0	20.0	19.0	18.0	20.0	20.0	22.0	28.0	19.0	20.0	18.0	20.0	20.0	22.0	21.0	20.0	20.0
Spindle-6	23.0	20.0	21.0	23.0	23.0	21.0	20.0	19.0	22.0	21.0	25.0	22.0	21.0	22.0	22.0	20.0	23.0	20.0	23.0	22.0	22.0	21.0	22.0	20.0	21.0	22.0	21.0	20.0	23.0	23.0	23.0	22.0

Objective: To have fun with control charts.

Instructions: 1. The following data were collected from a brazing operation. Two models are built on the line, the BU's and the B's. The plant runs four shifts (A – D). What questions could you generate about this process? What answers might come from looking at the data on a control chart? Try to answer your questions.

Time: 20 minutes

Models Built				*Leaks per Shift*					*Total*		*Leaks per Model*			
BU's	*B's*	*Total*		*A*	*B*	*C*	*D*	*Total*	*%*		*BU's*	*%*	*B's*	*%*
4158	11180	15338		39	9	0	0	48	0.31%		21	0.51%	27	0.24%
13850	33918	47768		13	5	25	21	64	0.13%		32	0.23%	32	0.09%
2062	40588	42650		1	5	22	19	47	0.11%		4	0.19%	43	0.11%
2076	36281	38357		13	17	8	4	42	0.11%		11	0.53%	31	0.09%
9563	32883	42446		41	25	26	14	106	0.25%		49	0.51%	57	0.17%
21054	25721	46775		33	2	35	27	97	0.21%		73	0.35%	24	0.09%
33026	8675	41701		26	3	73	46	148	0.35%		143	0.43%	5	0.06%
32388	9334	41722		6	3	80	26	115	0.28%		99	0.31%	16	0.17%
14040	25209	39249		42	15	7	3	67	0.17%		30	0.21%	37	0.15%
8807	34242	43049		44	9	12	3	68	0.16%		33	0.37%	35	0.10%
11092	35921	47013		24	0	15	19	58	0.12%		22	0.20%	36	0.10%
15614	28132	43746		14	19	27	29	89	0.20%		50	0.32%	39	0.14%
18267	20769	39036		72	15	5	13	105	0.27%		83	0.45%	22	0.11%
19659	18938	38597		56	36	1	6	99	0.26%		96	0.49%	3	0.02%
19566	23045	42611		58	25	6	14	103	0.24%		75	0.38%	28	0.12%
21278	27357	48635		8	1	78	25	112	0.23%		84	0.39%	28	0.10%
10408	30585	40993		24	3	9	13	49	0.12%		27	0.26%	22	0.07%
6200	33422	39622		42	16	16	11	85	0.21%		64	1.03%	21	0.06%
8865	33403	42268		25	12	9	8	54	0.13%		31	0.35%	23	0.07%
8506	36660	45166		1	3	55	19	78	0.17%		40	0.47%	38	0.10%

Models Built				Leaks per Shift					Total		Leaks per Model			
BU's	B's	Total		A	B	C	D	Total	%		BU's	%	B's	%
14180	31391	45571		11	5	36	27	79	0.17%		45	0.32%	34	0.11%
16351	27049	43400		25	6	47	36	114	0.26%		54	0.33%	60	0.22%
4888	26321	31209		13	11	0	19	43	0.14%		8	0.16%	35	0.13%
9401	31168	40569		20	22	20	37	99	0.24%		25	0.27%	74	0.24%
13647	30139	43786		15	2	97	22	136	0.31%		68	0.50%	68	0.23%
12761	35060	47821		10	0	54	43	107	0.22%		52	0.41%	55	0.16%
12187	29000	41187		40	7	19	5	71	0.17%		30	0.25%	41	0.14%
12807	26991	39798		84	34	24	15	157	0.39%		93	0.73%	54	0.20%
10984	18146	29130		51	16	12	0	79	0.27%		41	0.37%	38	0.21%
3078	9197	12275		0	0	16	0	16	0.13%		10	0.32%	6	0.07%
237	12973	13210		0	0	21	5	26	0.20%		0	0.00%	26	0.20%
0	10355	10355		10	8	0	0	18	0.17%		0	#DIV/0!	18	0.17%
9	10066	10075		0	11	0	0	11	0.11%		0	0.00%	11	0.11%
5437	7429	12866		0	0	7	21	28	0.22%		25	0.46%	3	0.04%
5779	5950	11729		0	0	11	9	20	0.17%		15	0.26%	5	0.08%

Objective: To perform basic capability calculations.

Instructions: 1. For the control chart exercises above, develop the picture of process capability, calculate Cp, Sigma and, if the picture indicates the need, Cpk.

Time: 40 minutes

Description	Specification Limits
Compressor Stud Length (page 13)	5.3750” +/- 0.0005” Nominal, Spec Limits
Days Late for Delivery (page 14)	Less than 2 Days Late
Cost per Unit Repair – Air conditioners (page 15)	Less than $2500.00
Butterfly Valve Closure Time (page 17)	Less than 10 Seconds
Motor Rejects (page 19)	None Defective
Defective Full-Penetration Welds (page 20)	None Defective
Ceramic Paint Pinholes (page 21)	No Defects
Design Change Requests (page 22)	1 per Design

Objective: To assess stability and capability of a production process.

Instructions: 1. The following data was gathered from production of HSG compressors. Assess the *stability* and *capability* of this production process for these critical parameters.

Time: 40 minutes

			TOLERANCE	TOLERANCE	Tolerance	Tolerance	Tolerance	Tolerance	Tolerance	Tolerance	Tolerance
			+/- 0.001	+/- 0.001	+/- 0.0005	+/- 0.0018	+/- 0.0008	+/- 0.0008	+/- 0.0008	+/- 0.0008	+/- 0.0008
HSG NO	DATE	MACH.NO	M RTR BORE	FEM RTR BORE	SLD VAL	Parallel A	M dowel pos A	M dowel pos opp	F dowel pos A	F dowel pos opp	PosSld val
1393	11/23/1999	G&L3	-0.0001	0.0003	0.0012	0.0005	0.0007	0.0005	0.0004	0.0005	0.001
1394	11/24/1999	G&L3	0.0002	0.0002	-0.0005	0.0009	0.0005	0.0004	0.0002	0.0005	0.0012
1397	11/26/1999	G&L3	0.0001	0.0004	-0.0002	0.0012	0.0005	0.0003	0.0006	0.0009	0.0008
1398	11/27/1999	G&L3	0	0.0002	-0.001	0.0006	0.001	0.0005	0.0001	0.0007	0.0014
1399	11/27/1999	G&L3	0.0003	0.0003	0.0008	0.0015	0.0006	0.0003	0.0001	0.001	0.0008
1440	3/23/2000	Orion 1	0.0002	-0.0007	0.0006	0.0016	0.0011	0.0014	0.0011	0.0008	0.0008
1443	4/4/2000	Orion 1	-0.0003	-0.0005	0.0006	0.0011	0.001	0.001	0.0001	0.0005	0.0006
1444	4/4/2000	Orion 1	0	-0.0004	0.0006	0.0013	0.0008	0.0008	0.0001	0	0.0008
1488	8/8/2000	G&L3	0.001	0.0007	0.0001	0.0009	0.0001	0.0002	0.0001	0.0001	0.0001
1492	8/10/2000	G&L3	0.0013	0.0011	0	0.001	0.0003	0.0003	0.0007	0.0001	0.0011
1493	8/10/2000	G&L3	0.0008	0.0004	0.0001	0.0009	0.0006	0.0005	0.0006	0.0004	0.0012
1504	9/12/2000	Orion 1	-0.0002	-0.001	-0.0014	0.0011	0.0005	0.001	0.0002	0.0009	0.0003
1506	9/12/2000	Orion 1	0.0002	0	0.0002	0.0013	0.0005	0.0005	0.0002	0.0012	0.0002
1507	9/6/2000	Orion 1	-0.0011	-0.0018	-0.0003	0.0019	0.0003	0.0007	0.0001	0.0008	0.0011
1508	9/6/2000	Orion 1	0.0003	-0.0002	-0.0003	0.0032	0.0004	0.0005	0.0002	0.0009	0.0004
1509	9/11/2000	Orion 1	0.0004	-0.0001	-0.0003	0.0013	0.0004	0.0009	0.0001	0.001	0.0004
1510	9/12/2000	Orion 1	0.0005	-0.0002	-0.0011	0.0013	0.0004	0.0007	0.0003	0.0009	0.0004
1512	9/12/2000	Orion 1	0	-0.0001	0	0.0004	0.0002	0.0006	0.0007	0.0002	0.0012
1514	9/13/2000	G&L3	0.0014	0.0013	-0.0018	0.0002	0.0002	0.0002	0.0003	0.0001	0.0014
1516	9/14/2000	G&L3	0	-0.0003	-0.0013	0.0003	0.0002	0.0002	0.0003	0.0001	0.0007
1517	9/15/2000	G&L3	-0.0003	-0.0006	-0.0006	0.0005	0.0004	0.0001	0.0001	0.0006	0.0008

Objective: To understand the concept and calculation of process yield.

Instructions: 1. For the following process, calculate the First Pass Yields of the process steps, the Normalized Yield of the overall process and Rolled Through-put Yield of the process. Based on the number of defects detected through inspection, calculate the Final Pass Yield of the process. See the AXLE worksheet on your Excel file for the data.

Time: 20 minutes

Axle Production - The following steps are performed to manufacture this axle:



Process Step	# of Units Produced	# of Defect Opportunities	# of Defects Produced	# Detected Prior to Customer
1. End Face Milling	10,000	1	82	82
2. Rough Machining (Lathe)	10,000	1	25	25
3. Finish Turning (Lathe)	10,000	1	650	500
Diameter Inspection				
4. Axle End Gear Cutting	10,000	6[1]	235	120
5. Cleaning	10,000	1	3	3
6. Heat Treat/Quenching	10,000	1	100	10
7. Axle Grinding	10,000	1	140	20
Diameter & Surface Finish Inspection				
8. Flange Machining (Automatic Lathe)	10,000	1	5	3
9. Axle Flange Drilling (6 Holes)	10,000	6[1]	256	30
Final Inspection				

Notes: 1. These operations are applied six times to the axle.

Objective: To understand the concept and calculation of process yield.

Instructions:

1. For the following process, calculate the First Pass Yields of the process steps, the Normalized Yield of the overall process and Rolled Through-put Yield of the process. Note that if the item fails initial inspection, it is scrapped or reworked until good. Note that after Step 3 - final assembly, the customer judges the quality of the product.

Time: 20 minutes

Step 1 – Roughing		**Step 2 – Finishing**		**Step 3 – Assembly**	
Measure	**Pass/Fail**	**Measure**	**Pass/Fail**	**Measure**	**Pass/Fail**
12	Pass	5	Pass	7	Pass
5	Pass	8	Pass	25	Fail
23	Pass	5	Pass	11	Pass
11	Pass	25	Fail	15	Fail
0.75	Pass	19	Fail	20	Fail
7.75	Pass	4	Pass	0	Pass
7	Pass	11	Pass	6	Pass
11	Pass	24	Fail	9.5	Pass
2	Pass	9	Pass	14	Fail
16	Pass	11	Pass	21	Fail
10	Pass	1	Pass	29.5	Fail
7.75	Pass	9	Pass	1	Pass
1.5	Pass	20	Fail	13	Fail
18	Pass	9	Pass	20.5	Fail
4	Pass	10	Pass	20	Fail
11	Pass	27	Fail	13	Fail
11	Pass	7	Pass	12	Pass
14.5	Pass	10	Pass	13	Fail

Step 1 – Roughing		Step 2 – Finishing		Step 3 – Assembly	
Measure	**Pass/Fail**	**Measure**	**Pass/Fail**	**Measure**	**Pass/Fail**
16.75	Pass	3	Pass	28	Fail
2.25	Pass	11	Pass	24.5	Fail
11	Pass	3	Pass	14.5	Fail
8.25	Pass	4	Pass	15	Fail
13	Pass	0	Pass	15	Fail
9	Pass	3	Pass	12.5	Fail
11	Pass	1	Pass	2.5	Pass
		3	Pass	8	Pass
		3	Pass	4	Pass
		6	Pass	4	Pass
		8	Pass	7	Pass
		5	Pass	4.5	Pass
				10.5	Pass
				4	Pass
				23	Fail
				10.5	Pass
				8	Pass
				25	Fail
				Customer	
				Defects	11
				Trials	25

Objective: To practice improving the information quality of data arising from sporadic events.

Instructions:

1. Chart the information below on a *c* chart (see the INJURY worksheet on your Excel file for the data).

2. Apply the lessons of the Sporadic Events special topic to improve the information "quality."

Time: 30 minutes

Employee Injuries - A manufacturing plant has been working on reducing employee injuries through root cause analysis and corrective actions on the processes "producing" injuries. At the beginning of the year (1998), the plant put several countermeasures in place to address back injuries. Have these helped reduce this class of injury?

1998		
Month	**# of Injuries**	**Date(s) of Injury**
Jan	3	5-Jan, 20-Jan, 28-Jan
Feb	3	9-Feb, 18-Feb, 27-Feb
Mar	4	9-Mar, 16-Mar, 24-Mar, 30-Mar
Apr	4	9-Apr, 14-Apr, 20-Apr, 26-Apr
May	2	4-May, 19-May
Jun	3	1-Jun, 11-Jun, 30-Jun
Jul	3	7-Jul, 17-Jul, 24-Jul
Aug	3	2-Aug, 8-Aug, 24-Aug
Sep	2	14-Sep, 25-Sep
Oct	5	2-Oct, 9-Oct, 16-Oct, 23-Oct, 30-Oct
Nov	3	4-Nov, 16-Nov, 24-Nov
Dec	4	2-Dec, 9-Dec, 17-Dec, 23-Dec

1999		
Month	**# of Injuries**	**Date(s) of Injury**
Jan	3	9-Jan, 18-Jan, 31-Jan
Feb	1	18-Feb
Mar	3	8-Mar,19-Mar, 28-Mar
Apr	2	14-Apr, 27-Apr
May	1	28-May
Jun	1	9-Jun
Jul	2	2-Jul, 19-Jul
Aug	3	3-Aug, 18-Aug, 30-Aug
Sep	2	13-Sep, 28-Sep
Oct	1	11-Oct
Nov	2	8-Nov, 25-Nov
Dec	2	10-Dec, 31-Dec

Objective: To apply the ANOM method of detecting differences in experimental data.

Instructions:

1. Chart the following experimental data on an X-Bar, R Control Chart (see the COATING worksheet on your Excel file). Do any of the factor levels produce a "signal?"

2. Take the same data and perform an ANOM. What difference does this produce?

Time: 30 minutes

Coating Process Improvement

A Six Sigma team has run experiments to increase the coating thickness for air conditioner housings in an attempt to provide a more durable surface finish. The coatings are sprayed on the air conditioner housing and then the housings are baked. Four different spray nozzles (A – D) were used in the experiment with the following results:

Coating Thickness by Nozzle (mils)			
A	**B**	**C**	**D**
2.9364	2.9882	3.2488	2.8960
3.0135	2.6863	3.2824	2.7380
3.1551	2.7986	3.3301	2.8240
2.9543	2.8324	3.2620	2.7837
2.9839	2.7991	3.3198	2.8050
3.0006	2.8375	3.2669	2.6654
3.1108	2.7202	3.2788	2.9812
3.0059	2.7531	3.2703	2.8110
2.9054	2.8139	3.3224	2.8543
2.9897	2.7728	3.3029	2.7546

Objective: To determine the relative advantage of the CUSUM chart over the X, mR control chart.

Instructions:

1. Chart the following data on an X, mR Control Chart (see the TOOL WEAR worksheet on your Excel file). Do any of the data produce a "signal?"

2. Take the same data and develop a CUSUM chart. What difference does this produce?

Time: 20 minutes

Suspected Tool Wear

Manufacturing engineers are trying to determine if tool wear is affecting a particular cutting machine. They have collected the following data from the process (order of data proceeds down column 1 then to column 2 etc.):

Length	Length	Length	Length	Length
1.50095	1.50071	1.49602	1.50050	1.49195
1.50386	1.49940	1.49257	1.49744	1.49588
1.50019	1.49810	1.49799	1.49277	1.49715
1.49223	1.50023	1.49716	1.50050	1.49673
1.50306	1.49668	1.49855	1.49689	1.48889
1.50209	1.49867	1.49470	1.49146	1.49444
1.49407	1.50278	1.50078	1.50424	1.49743
1.50309	1.49790	1.49685	1.49524	1.49518
1.50139	1.50367	1.49884	1.49874	1.49548
1.49664	1.49687	1.49618	1.49406	1.49505

Objective: To practice developing the Difference short run control chart.

Instructions:
1. Chart the following data on an X, mR Control Chart (see the SHAFT worksheet on your Excel file).
2. Take the same data and develop a Difference control chart. How does this help you produce better information from the data?

Time: 20 minutes

Short Run Motor Shafts - Compressor motor shafts are machined to support a Just-in-Time production operation. Manufacturing engineers believe that the machining process' variation doesn't change from shaft to shaft, however the shaft diameters differ (order of production proceeds down column 1 and then to column 2).

Part	Diameter	Part	Diameter
XB-4	1.2659	KJ-11	2.2544
XB-4	1.2604	KJ-11	2.2197
XB-4	1.2718	KJ-11	2.2586
XB-4	1.2431	KJ-11	2.2524
XB-4	1.2493	KJ-11	2.2536
XB-4	1.2543	KJ-11	2.2607
XB-4	1.2379	KJ-11	2.2485
XB-4	1.2621	KJ-11	2.2537
XB-4	1.2364	KJ-11	2.2508
XB-4	1.2418	XB-4	1.2477
XB-4	1.2622	XB-4	1.2458
XB-4	1.2573	XB-4	1.2561
XB-4	1.2464	XB-4	1.2595
XB-4	1.2525	XB-4	1.2334
KJ-11	2.2618	XB-4	1.2341
KJ-11	2.2359	XB-4	1.2600
KJ-11	2.2440	XB-4	1.2566

Objective: To practice developing the ZED short run control chart.

Instructions:
1. Chart the following data on an X, mR Control Chart (see the TUBE SHEET worksheet on your Excel file).
2. Take the same data and develop a ZED control chart. How does this help you produce better information from the data?

Time: 20 minutes

Tube Sheet Hole Drilling -The following data were collected from three commercial air conditioning tube sheets all drilled by the same machine. The tube diameters are different (Sheets 1 and 3 are drilled for 0.5 inch tubes, Sheet 2 is drilled for 0.75 inch tubes). Manufacturing engineering suspects that the variation in the tube diameters is different.

Tube Sheet 1	Diameter	Tube Sheet 2	Diameter	Tube Sheet 3	Diameter
AC-5	0.4905	AD-7	0.7507	AC-5	0.4905
AC-5	0.4898	AD-7	0.7504	AC-5	0.4898
AC-5	0.4898	AD-7	0.7508	AC-5	0.4895
AC-5	0.4897	AD-7	0.7498	AC-5	0.4896
AC-5	0.4895	AD-7	0.7493	AC-5	0.4907
AC-5	0.4901	AD-7	0.7501	AC-5	0.4902
AC-5	0.4896	AD-7	0.7506	AC-5	0.4900
AC-5	0.4907	AD-7	0.7503	AC-5	0.4902
AC-5	0.4898	AD-7	0.7506	AC-5	0.4903
AC-5	0.4902	AD-7	0.7500	AC-5	0.4899
AC-5	0.4904	AD-7	0.7497	AC-5	0.4893
AC-5	0.4899	AD-7	0.7502	AC-5	0.4908
AC-5	0.4906	AD-7	0.7502	AC-5	0.4901
AC-5	0.4900	AD-7	0.7498	AC-5	0.4901
AC-5	0.4895	AD-7	0.7505	AC-5	0.4900
AC-5	0.4903	AD-7	0.7502	AC-5	0.4898
AC-5	0.4899	AD-7	0.7500	AC-5	0.4900
AC-5	0.4897	AD-7	0.7501	AC-5	0.4902
AC-5	0.4902	AD-7	0.7512	AC-5	0.4901

Objective: To practice developing X-Bar, R control charts with varying subgroup sizes.

Instructions: 1. Chart the following data on an X-Bar, R Control Chart (see the LATHE worksheet on your Excel file).

Time: 20 minutes

Lathe Out-of-Service Times - Take the following Lathe Out-of-Service times (hours) data and create an X-Bar, R Control Chart. To accommodate the varying subgroup sizes, you will have to calculate limits for each subgroup, using the "A" and "D" coefficients. Also, don't forget to use the raw data to calculate the grand average; you can't average the subgroup averages (without weighting them!):

JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN
1	3	2	1	1	1	1	2	4	3	6	4	1
1	1	1	4	1	3	1	3	1	1	1	3	8
2	1	3	4	1	1	1	2	1	4	2	1	2
4	4	2	10	1	2	6	2	4	3	7	1	3
1	2	1	3	2	12	2	7	4	1	2	2	2
1	6	3	1	4	1	5	3	1	1	1	3	1
		1	2	2	4	2	3	7		1	2	2
		2		4	1	2	1	1		2		1
		1		2	1	1	1	2		1		
		1		1	1	1	4					
		1			5		3					
					4		6					

Objective: To differentiate between the idea of an Arithmetic Limit and a Statistical Limit.

Instructions:

1. Flip a coin 100 times and record the sequence of heads and tails.
2. Graph the fraction of tosses resulting in heads as a function of the number of coin tosses.
3. Observe the behavior of the fraction. Does it continuously approach the statistical limit of 0.5? How does it behave? What do you think would happen if you flipped the coin 100 times more (try it!)?

(Hint: Set this up as an Excel spreadsheet)

Time: 20 minutes

Objective: To perform a “Capstone” process capability exercise.

Instructions:

1. Review the information on the Hair Pin tubes and the new Tube Bender. (5 min.)
2. Develop a Plan to assess the capability of the new Tube Bender. (30 min.)
3. Review the “Actual” Plan, and the capability study notes. (10 min.)
4. Based on the “Actual” Plan, analyze the capability of the new Tube Bender. What conclusions/recommendations would you make? (30 min.)

See the CAPSTONE worksheets on your Excel data file.

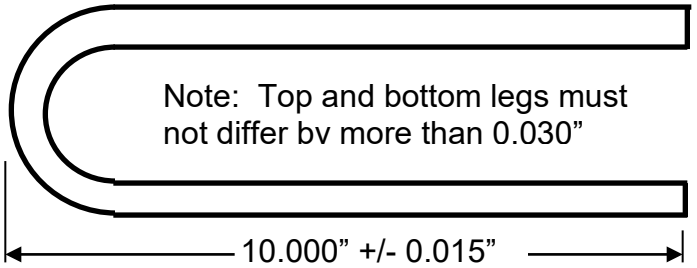
Time: 75 minutes

Capability Study Exercise

In air conditioning heat transfer assemblies and in radiator cores, there is a series of "U" shaped components fabricated from aluminum or copper tubing. They are sometimes called "hairpins" because of their resemblance to the hairpiece item by the same name. A number of these hairpins are inserted into heat transfer fins, and then expanded to obtain a mechanical contact with the fin. They are then connected to each other in series by brazing or soldering a return bend to the leg of one hairpin to the leg of an adjacent hairpin to route the fluid from one to the other. To properly fit into the assembly, the length of the hairpin is important. Also of importance is that the two legs of the hairpin not differ in length excessively or the return bend will not mate properly when connected to adjacent hairpins.

The sketch below is a simplified representation of the part, showing the length and tolerance on the legs and a supplemental dimensional constraint specifying that they must not differ in length by more than a given amount. This difference in leg length is sometimes referred to as "peg leg" for obvious reasons.

Hair Pin Tube



Specifying peg leg tolerance is redundant. If there is a tolerance on each of the leg lengths, then peg leg can theoretically be as much as the maximum in-tolerance variation of the two legs. If the peg leg tolerance is less than this amount, then the total tolerance on the legs is not available to the process. Nevertheless, peg leg is a dimension that is sometimes monitored.

A machine called a tube bender creates hairpins. Usually, several hairpins are made in one cycle by having multiple bending stations. A coil of tubing serves each of the stations. The fabrication cycle is as follows:

- With the stations empty, each of the coils feeds straightened tubing into its station until it reaches a positive stop.
- The tubes are clamped into position.

- Each tube is cut to a given length.

- Half of the table supporting the tubes then swings upward through an arc of 180 degrees and bends the tubing over a forming tool located midway along the length of the tube.

- The table returns to its initial position.

- The formed hairpins are then ejected or removed from the process.

Several factors may affect the length of each of the legs. Some of them might include:

- The cut length of the tubing.
- Diameter of the tubing.
- Metallurgy of tubing.
- Wall thickness of the tubing.
- Stability of the positive stop and clamping mechanism.
- Location of the forming tool at the midpoint of the tube length.
- Dimension of the forming tool.
- Bearings and linkages in the tube bender.
- Cycle time.

In this study, we will attempt to evaluate a new tube bender to see if it can adequately satisfy the print dimensions and tolerances. The machine being evaluated has five stations, producing five hairpins in each cycle. Assume that the production cycle-time will be 15 seconds. A full coil of tubing can produce 500 hairpins. An internal company specification for tubing exists - #EMR8890.

The length of each leg of the hairpin is measured by placing it in a fixture with the outside of the bend against a stop, and a dial indicator at the end of each leg mastered to a reading of zero for the nominal specified length. When the dial indicator has been zeroed, it is mechanically attached to the fixture. The gauging is owned by the purchasing company and taken to the equipment manufacturer where the capability study will be performed.

Planning the Capability Study

Develop a Plan to Study the Capability of the new Tube Bender. Consider the following questions during your planning. Complete the Capability Requirements Documentation forms on the next two pages.

- What level of capability should be required?
- What should we know about the measurement system?
- What characteristics, dimensions or features of the raw material should be documented?
- What will be specified for cycle-time?
- How many pieces should be evaluated?
- How many pieces should be run?
- Who will operate the process?
- What subgroup size will be specified?
- What will constitute a rational subgroup?
- Will a setup or warm up period be allowed?
- Will we evaluate leg length, peg leg or both?
- If peg leg were to be measured, do we record the absolute difference in length or use one leg as a reference? If we do the latter, how do we handle the sign of the readings?
- Should there be a constraint on how close the process mean should be to the nominal value in the specification?
- When the hairpins are removed from the process, how do we keep track of which was the top and bottom leg?
- Should the subgroups be consecutive or spaced over the entire run if the run quantity exceeds the sample quantity?
- Should we maintain the production sequence for the parts we will be measuring? How would we do this? Why would we do this?
- Who will do the measurement?
- If the C_{PK} does not meet our requirements, but the C_P does, will we have to re-run the study? Why?
- What analyses of the data will be made?
- In production, this machine will be required to make hairpin lengths ranging from 10’ to 30". Which one(s) should be evaluated?

Capability Requirements Documentation

Supplier	Ajax Machine Tool Company
Address	Cleveland, Ohio
Process / Machine	Tube Bender - Hairpins
Part Numbers / Operation	HC7789 Rev3, HC7666 Rev1, HC7344 Rev3

	Characteristic	Specification	Adjustable or Fixed Mean	Required Cpk
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

Additional Requirements

Source of material	
Material Specification	EMR8890
Quantity / Cycles to be run	
Quantity / Cycles to be studied	
Data to be evaluated by:	
Process to be operated by:	
Customer Attendees	
Supplier attendees	
Source of gauging	

Measurement Methods:

Measurement Requirements:

Analytical methods to be used:

Process Operation and Requirements:

Comments:

Company	Name	Dept.	Date	Company	Name	Dept.	Date

Don't Turn the Page Until You've Developed <u>Your</u> Plan!

Capability Requirements Documentation

Completed Plan

Supplier	Ajax Machine Tool Company
Address	Cleveland, Ohio
Process / Machine	Tube Bender - Hairpins
Part Numbers / Operation	HC7789 Rev3, HC7666 Rev1, HC7344 Rev3

	Characteristic	Specification	Adjustable or Fixed Mean	Required Cpk
1	Leg length	10.000" +/-0.015	Adjustable	1.43
2	Leg length	20.000" +/-0.015	Adjustable	1.43
3	Leg length	30.000" +/-0.015	Adjustable	1.43
4	Peg Leg	0.030" Max	Adjustable	1.43
5	Bend radius	2.000" +/-0.010	Fixed	1.43
6				
7				
8				
9				
10				

Additional Requirements
Process is to be operated at the rated cycle time of 15 sec.
Process mean must be in the middle third of the tolerance.
Drift caused by the equipment is not anticipated. However, over a run of 500 pieces, an Xbar-R chart
should not show evidence of drift.
Each of the five stations will be evaluated independently.
The customer will provide a sufficient number of coils of tubing.
The material will be verified as conforming to spec #EMR8890 by the customer.
Each tube used in the analysis will be identified by marking as to which station and which cycle of the run
of it represents.

Source of material	Material to be provided by the customer.
Material Specification	EMR8890
Quantity / Cycles to be run	500 cycles for each of the five stations.
Quantity / Cycles to be studied	25 subgroups of 5 for each station.
Data to be evaluated by:	Customer Manufacturing Engineering
Process to be operated by:	Ajax Machine Tool personnel
Customer Attendees	E. Smith, QC; B. Jones, ME; R. Dean, PE
Supplier attendees	To be determined.
Source of gauging	Customer

Measurement Methods: Customer will provide the fixture and gauging for data.
Gauging will consist of dial indicators.

Measurement Requirements:
Dial indicators reading to 0.001 will be used. Customer is responsible for documenting the calibration
and R&R of the indicators.

Analytical methods to be used:
X-Bar and R control chart analysis.
Equally spaced subgroups of 5 over the run off period will be used.
Determine the shape of the underlying distribution (if process exhibits stability).
Calculate Machine Capability Indices Cp, Cpk.
If distributions are not normal, other statistical methods for estimating capability will be used.

Process Operation and Requirements:
Ajax technicians will be allowed to center the process to their satisfaction prior to the formal run off.

Comments: No physical damage to the tubing attributable to the machinery will be accepted.

Running the Capability Tests

Based on the above capability plan, the tube bender was tested. Some notes from the study appear below:

- Once the process was adjusted for nominal, each of the five stations was monitored separately. A subgroup consisted of five consecutive pieces from each of the stations. The process was then allowed to produce an additional 15 pieces from each station, and then the next subgroups of five were taken.

- Before the tubes were removed, they were coded with spray paint so that each of the five stations was identified. Each subgroup of five was bundled and numbered in sequence.

- No process adjustments were allowed during the run.

- Only one coil of tubing was used for each station.

- After the measurements were taken, the identified bundles of tubes were set aside so dimensions could be reconfirmed should the data show any abnormalities.

- The dial indicators were calibrated prior to taking the measurements and the calibration was reconfirmed after the data was taken.

- A gauge R&R study was conducted on the dial indicators prior to the study. It was determined that the six standard deviation range for the indicator was 0.002"

Analyzing the Study Results

The following tables show the measurements taken during the capability study. Not all of this data is required to assess the tube bender's capability. Use whatever you decide is necessary. All measurements have been coded to show the number of 0.001" from the nominal for the leg lengths and for the difference in length between the two legs. Double-click on the tables to launch their Excel version:

Top Tube Length

Top Leg Length. Data Coded to Show Measurement in 0.001" From Nominal

Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg
1	-7	26	-3	51	0	76	1	101	-3
2	-1	27	0	52	0	77	3	102	0
3	2	28	3	53	-3	78	-1	103	5
4	-2	29	-4	54	2	79	-4	104	-2
5	-5	30	3	55	-1	80	7	105	-1
6	0	31	-1	56	-4	81	1	106	-4
7	3	32	-3	57	-2	82	-2	107	-1
8	5	33	5	58	-3	83	-3	108	4
9	-2	34	3	59	2	84	2	109	-4
10	6	35	-3	60	1	85	-4	110	0
11	2	36	0	61	0	86	2	111	1
12	7	37	-5	62	-2	87	2	112	2
13	-2	38	0	63	-2	88	-1	113	-4
14	-7	39	0	64	-1	89	2	114	4
15	-6	40	3	65	3	90	0	115	5
16	4	41	-1	66	-3	91	0	116	2
17	5	42	-1	67	1	92	2	117	-5
18	1	43	-1	68	0	93	-6	118	-3
19	2	44	4	69	-2	94	1	119	3
20	5	45	-5	70	-2	95	2	120	1
21	3	46	0	71	1	96	2	121	0
22	1	47	-1	72	-5	97	2	122	-5
23	5	48	3	73	1	98	3	123	-1
24	-1	49	5	74	2	99	0	124	0
25	1	50	-4	75	0	100	3	125	2

Bottom Tube Length

Bottom Leg Length. Data Coded to Show Measurement in 0.001" From Nominal

Meas Number	Length Bot Leg	Meas Number	Length Bot Leg	Meas Number	Length Bot Leg	Meas Number	Length Bot Leg	Meas Number	Length Bot Leg
1	-2	26	7	51	1	76	-3	101	2
2	2	27	12	52	-3	77	-1	102	-3
3	0	28	-2	53	7	78	-1	103	6
4	4	29	3	54	3	79	5	104	6
5	2	30	3	55	-4	80	3	105	-7
6	1	31	1	56	9	81	3	106	4
7	4	32	2	57	2	82	7	107	7
8	6	33	3	58	2	83	4	108	12
9	2	34	-5	59	-3	84	0	109	-1
10	2	35	2	60	6	85	5	110	9
11	-3	36	8	61	7	86	3	111	4
12	-3	37	6	62	3	87	-3	112	-1
13	-1	38	8	63	2	88	6	113	-2
14	4	39	2	64	4	89	5	114	-6
15	0	40	2	65	-6	90	2	115	4
16	3	41	8	66	1	91	-4	116	0
17	0	42	9	67	2	92	-3	117	10
18	9	43	7	68	-1	93	0	118	1
19	8	44	7	69	1	94	2	119	6
20	-4	45	5	70	7	95	4	120	5
21	1	46	-6	71	5	96	4	121	6
22	2	47	4	72	0	97	7	122	3
23	-1	48	7	73	-4	98	13	123	-1
24	2	49	6	74	7	99	-2	124	3
25	5	50	4	75	0	100	1	125	3

Longest Leg - Shortest Leg

Peg Leg.Absolute (Longest - Shortest) Data Coded to Show Measurement in 0.001"

Meas Number	Length Bot Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg
1	5	26	10	51	1	76	4	101	5
2	3	27	12	52	3	77	4	102	3
3	2	28	5	53	10	78	0	103	1
4	6	29	7	54	1	79	9	104	8
5	7	30	0	55	3	80	4	105	6
6	1	31	2	56	13	81	2	106	8
7	1	32	5	57	4	82	9	107	8
8	1	33	2	58	5	83	7	108	8
9	4	34	8	59	5	84	2	109	3
10	4	35	5	60	5	85	9	110	9
11	5	36	8	61	7	86	1	111	3
12	10	37	11	62	5	87	5	112	3
13	1	38	8	63	4	88	7	113	2
14	11	39	2	64	5	89	3	114	10
15	6	40	1	65	9	90	2	115	1
16	1	41	9	66	4	91	4	116	2
17	5	42	10	67	1	92	5	117	15
18	8	43	8	68	1	93	6	118	4
19	6	44	3	69	3	94	1	119	3
20	9	45	10	70	9	95	2	120	4
21	2	46	6	71	4	96	2	121	6
22	1	47	5	72	5	97	5	122	8
23	6	48	4	73	5	98	10	123	0
24	3	49	1	74	5	99	2	124	3
25	4	50	8	75	0	100	2	125	1

Top Tube Length - Bottom Tube Length

Peg Leg.Top vs. Bottom. Data Coded to Show Measurement in 0.001"

Meas Number	Length Bot Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg	Meas Number	Length Top Leg
1	-5	26	-10	51	-1	76	4	101	-5
2	-3	27	-12	52	3	77	4	102	3
3	2	28	5	53	-10	78	0	103	-1
4	-6	29	-7	54	-1	79	-9	104	-8
5	-7	30	0	55	3	80	4	105	6
6	-1	31	-2	56	-13	81	-2	106	-8
7	-1	32	-5	57	-4	82	-9	107	-8
8	-1	33	2	58	-5	83	-7	108	-8
9	-4	34	8	59	5	84	2	109	-3
10	4	35	-5	60	-5	85	-9	110	-9
11	5	36	-8	61	-7	86	-1	111	-3
12	10	37	-11	62	-5	87	5	112	3
13	-1	38	-8	63	-4	88	-7	113	-2
14	-11	39	-2	64	-5	89	-3	114	10
15	-6	40	1	65	9	90	-2	115	1
16	1	41	-9	66	-4	91	4	116	2
17	5	42	-10	67	-1	92	5	117	-15
18	-8	43	-8	68	1	93	-6	118	-4
19	-6	44	-3	69	-3	94	-1	119	-3
20	9	45	-10	70	-9	95	-2	120	-4
21	2	46	6	71	-4	96	-2	121	-6
22	-1	47	-5	72	-5	97	-5	122	-8
23	6	48	-4	73	5	98	-10	123	0
24	-3	49	-1	74	-5	99	2	124	-3
25	-4	50	-8	75	0	100	2	125	-1

7.0 Stratification

Unit	Description	Page
7.1	Pie, Bar & Radar Charts	7.1-1
7.2	Pareto Analysis	7.2-1
7.3	Exercises	7.3-1

One of the key principles of problem solving is that of stratification. Geologists are always studying strata of rock structures found around the Earth. A *stratum* is a layer; *strata* are layers. When we perform a stratification analysis, we are trying to stratify the problem or process, looking for the Vital Few factors that contribute the most to the problem.

When we find these Vital Few factors, we'll concentrate on finding their causes and eliminating these as sources of problems. We'll leave the rest of the factors, or Useful Many for later. This is the essence of stratification analysis - stratify and prioritize!

7.1 Bar, Pie & Radar Charts

Learning Objectives

- Be able to construct and interpret Bar, Pie and Radar Charts
- Be able to determine which chart is best for a specific situation

Unit Contents

- Bar Charts
- Pie Charts
- Radar Charts
- Comparison of Line, Bar, Pie and Radar Charts

7.1.1 Bar Charts

Purpose

The purpose of a bar graph is mainly to show differences between categories. Some special bar graphs can be used to show trends over time, but please don't use bar graphs where a line graph is more appropriate.

Application

Some typical bar graph applications are listed below. Today, newspapers and magazines are great sources of bar graph examples. We always look forward to seeing how USA Today™ will show the results of surveys or other data!

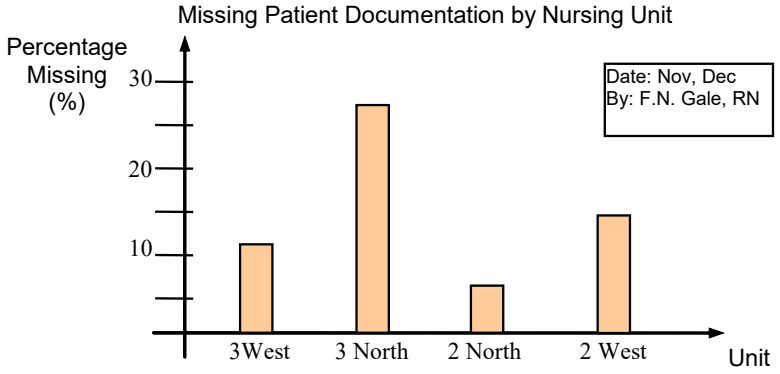
- Defects categorized by type, location, assembly stage, etc.
- Any variable stratified by categories: Hospital Length of Stay by physician, Sales by Region or Store, power generation by fuel (nuclear, coal, oil, gas), alcohol consumption by age group, etc., etc.

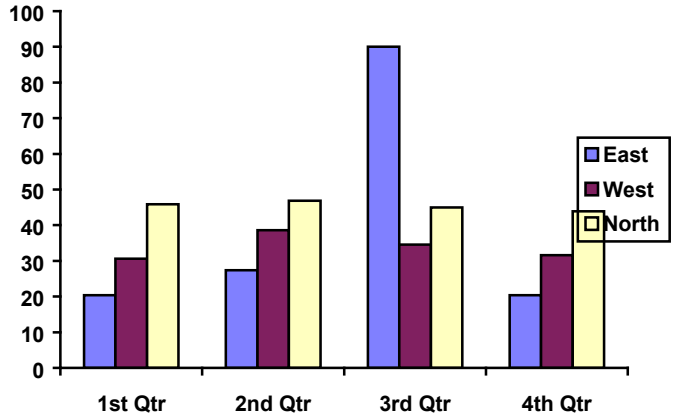
Construction & Examples

"Basic" construction steps are very simple:

1. Gather the data and sort it into categories of interest.
2. Draw a vertical and horizontal line on a piece of graph paper.
3. For a vertical bar graph, label the vertical axis with the performance measure. Label the horizontal axis with the categories.
4. Scale the vertical axis from zero to a value 10 - 15 % higher than the largest category.
5. Draw the bars, title and label the graph.

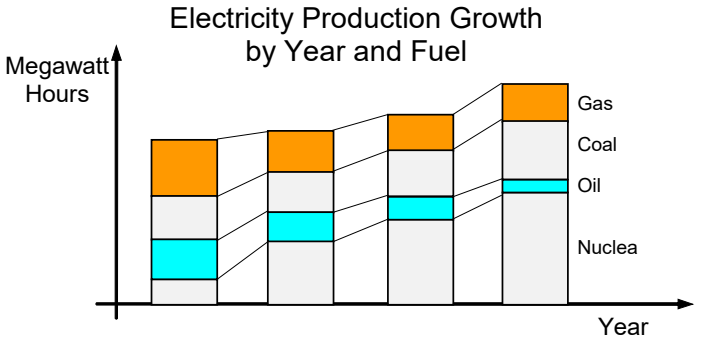
Here's a simple bar chart displaying differences between nursing units:





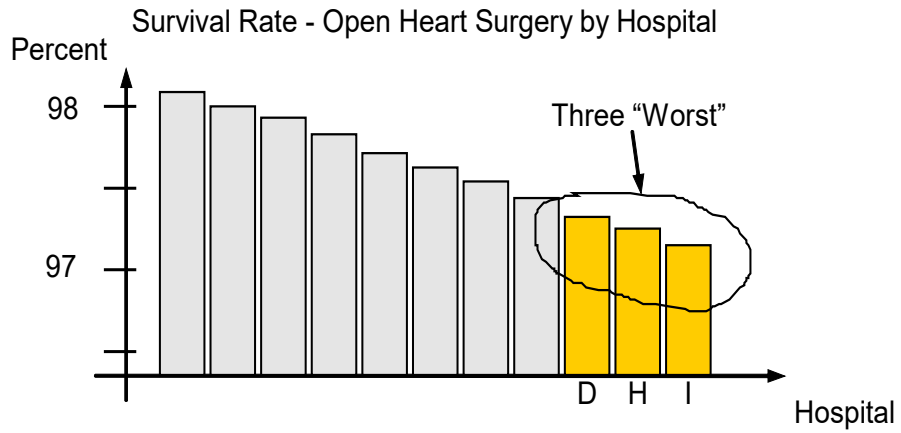
Combining two or more categories on a bar chart can communicate a great deal of information in a single picture as seen to the left.

Stacked Bar Charts are also compact, powerful communicators of information:



There's one caution we'd like to point out in using bar charts. It's very easy to enter some performance data and arrange it on a bar chart in increasing or decreasing order. The eye is then drawn to those categories that are the highest or lowest and we may think that the "top three" or "bottom three" categories need to be improved:

Would you be reluctant to have open-heart surgery at hospitals D, H, or I? Well, first, looking at the vertical scale, your survival chance at these hospitals is about 97.25%, versus about 98% at the three "best" hospitals. Three quarters of a percent isn't a big difference.



Second, there will be variation in every process. Are the survival rate differences significant, or are they just the result of random and expected variation?

We can order any set of data and calculate the average value. But remember this important and astonishing fact:

In any set of data, about half the data will be above average and half will be below the average!!!

We see advertisements like this all the time. Of course they're trying to sell us on something, but we've got to look beyond the surface.

7.1.2 Pie Charts

Purpose

The Pie Chart takes a set of data divided into categories and displays the relative proportions as “slices.” Categories appearing as larger slices make up a larger fraction of the whole. The Pie Chart is often used during improvement projects to help prioritize which piece of the problem will receive further attention and study.

Application

Any variable that can be broken down into categories is a candidate for analysis through a Pie Chart - needlesticks broken down by type, product assembly defects broken down by type, shift, or assembly point, injuries broken down by type, job, or department, total sales by region, department or store are just a few examples.

Construction

1. Collect the data and organize it by category. Total the events by category or sum the performance measure for each category (e.g. add the sales for each store within one region, then the next, etc.).

   Note: If there are some categories that contribute very little to the total, you can group their contributions into an “Other” category.

2. Calculate the category fractions by dividing each category’s total by the grand total (i.e. divide the Northern Region’s sales by the Total Sales).

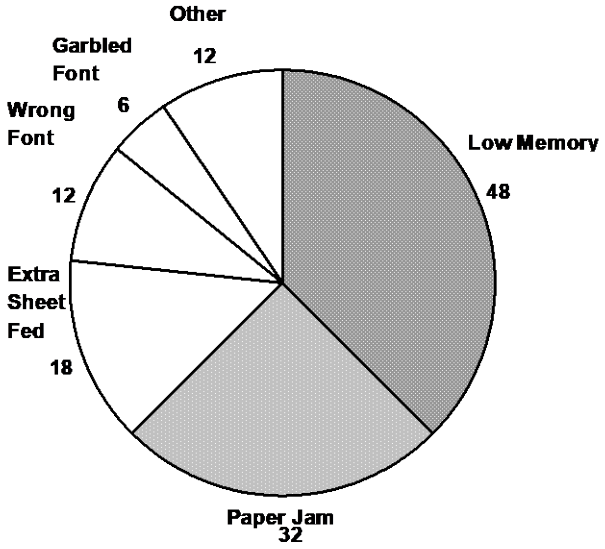
3. If you can divide a circle into one hundred parts, then multiply each fraction by one hundred. Starting at the 12:00 position, mark off the slices, from largest category to smallest (or the “Other” category).

   Note: If your circle is divided into 360 degrees, then multiply each fraction by 360. This will give you the number of degrees associated with each category’s slice.

4. Title and label the Pie Chart.

Here's a Pie Chart of printer errors that occurred during a development test of a new laser printer. From this test data, you can see that the engineers need to work on improving the printer's memory and paper handling qualities before sending Model PH-6 to production:

Pie Chart of PC Printer Errors - Model PH-6



7.1.3 Radar Charts

Purpose

Sometimes, we may want to display several different variables on one chart, perhaps to compare their relative performance. The Radar Chart helps us do this.

Application

Customer and employee surveys often measure several different variables or quality characteristics. The Radar Chart is used to display the performance of these individual characteristics and to look for relative strong or weak areas in the product or service's performance.

Construction

1. Collect the data for the different variables or quality characteristics.

2. Draw a circle and put a dot at its center. Draw a radius for each variable to be shown on the Radar Chart. Space these equally:

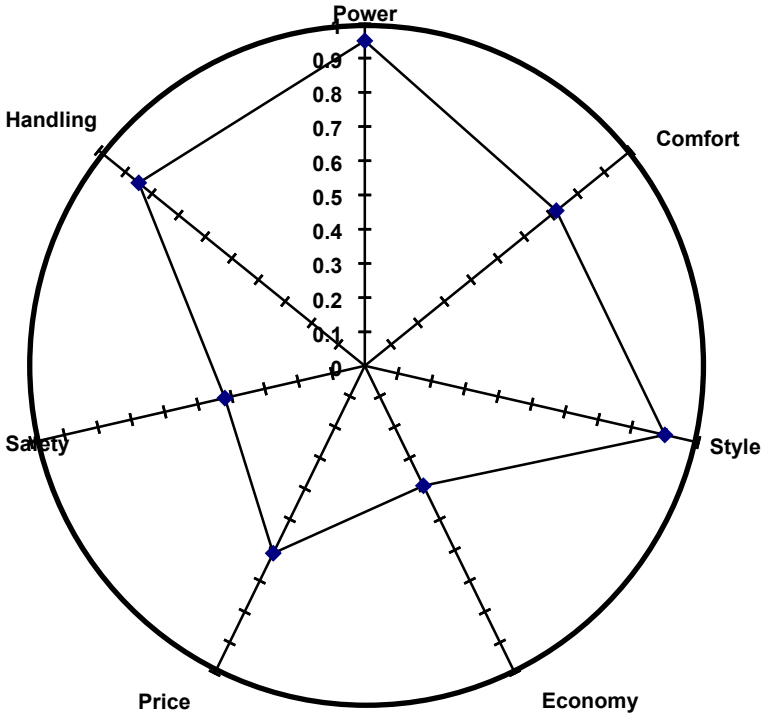
Number of Variables	Degrees between Radii
4	90
5	72
6	60
7	51.4
8	45
9	40
10	36
11	32.7

3. Scale each radius. For survey data, the responses are often obtained using a Likert scale (e.g. 1 to 5). For this data, scaling should start with the center of the circle labeled 1 and the circumference labeled 5. If the data is percentage type, a similar procedure is followed with 0% at the center and 100% at the circumference.

4. For each variable, draw its value as a point on its radius (the results of multiple surveys may be averaged and the averages plotted).

5. Join each point with a line, title and label the Radar Chart.

Rallye Motor Company
"Stallion" Sports Coupe Customer Survey



For most products and services, the Radar Chart should be "unbalanced." That is, customers perceive certain quality characteristics as stronger than others. These are the "selling points" of the product or service. The market research, planning and design processes should identify these selling points and build them in to the design. Of course, characteristics that are unusually or unexpectedly weak or poor performers are candidates for improvement.

Two or more products or services may be compared on one Radar Chart. This can be an effective way of comparing your performance to your competitors or an old product or service to a new one.

7.1.4 Comparison of Line, Bar, Pie and Radar Charts

Here's a quick comparison of the charts and graphs we've introduced in this unit:

Chart	Advantages	Disadvantages
Line Graph	Makes trends and data variation over time easy to track. Good for highlighting changes in some variable. Can be used to track more than one variable at a time.	Can lead to "overreaction" to changes that aren't there. Control charts are the best tools for studying variation.
Bar Chart	Good for comparing one category to another. Many different styles can be constructed (stacked bar, "3-dimensional" bar chart, etc.) Easy to construct.	Sometimes it is inappropriately used for tracking data over time. Similar danger to overreaction as line graph.
Pie Chart	Useful for showing relative proportion of each category to the whole. Several layers of stratification can be shown on one graph.	Should not be used for tracking data over time.
Radar Chart	Useful for showing performance of many variables or characteristics on one chart. Useful for comparing two or more products/services across many characteristics.	Not useful for tracking data over time, although "before & after" comparisons can be made.

7.2 Pareto Analysis

Learning Objectives

- Understand and Apply the Pareto Principle to Problem Solving
- Perform Contingency Table Analysis for Attribute Data

Unit Contents

- The Pareto Principle
- The Pareto Analysis Process
- The Pareto Chart
- Contingency Table Analysis

7.2.1 The Pareto Principle

Back in the late 1800's, an Italian economist, Vilfredo Frederico Pareto, came up with an interesting finding: About 80% of the wealth of his country was held by fewer than 20% of the people. Let's fast forward to the 1930's. In studying manufacturing defects, Dr. Joseph Juran observed that often over 80% of the problems were caused by only a few factors or variables in the production process. His familiarity with the work of V. F. Pareto led him to name this empirical observation The Pareto Principle.[1] This principle has broad application in quality improvement. A few examples of the Pareto Principle at work:

- A large fraction of power plant failures are due to boiler tube problems,
- Seventy percent of assembly defects on irrigation sprinklers are due to problems inserting two components,
- Over 75% of a printing company's sales are to just three customers.
- Wasted days in a hospital are due mainly to problems transferring patients to Skilled Nursing Facilities,
- Delays receiving spare parts are most often associated with one vendor,
    - Over 90% of Mutual Fund transaction errors fall into four categories,
- Sixty-five percent of employee injuries are back strains and sprains,

Although most of the examples apply to problems encountered in production processes, Pareto can also apply to sales, volumes, costs and other quality characteristics of a product or service.

Now we won't guarantee that the Pareto Principle will appear in every situation, but it does often. In fact, it appears often enough that we've found it worthwhile to include an attempt at Pareto Analysis in almost every process improvement effort in which we're involved. There's even a sort of Pareto Principle that's applied to quality tools: With just Pareto and Cause & Effect over 80% of quality problems can be addressed.

[1]Dr. Juran humorously laments that he didn't name it after himself. He says that he was there, he had the opportunity, but he blew it!

7.2.2 The Pareto Analysis Process

The Pareto Principle is simple and widely applicable, yet we've seen many quality improvement efforts that could have but did not employ Pareto Analysis. One of our hypotheses is that while the Pareto Chart is taught in most quality courses, Pareto Thinking is not.

Let's explore the basic thought process behind Pareto Analysis. The process is "exploratory" in nature; we may start out on a certain train of thought, and the data may or may not confirm our thinking. Although it's sometimes frustrating, if an initial Pareto Analysis doesn't pan out, you should view this positively - take another shot at it and you might learn something new!

What's the Effect or Problem? - Pareto Analysis starts with some effect or problem that you're investigating. The typical Pareto Analysis is done on an effect that is stated as a problem:

- Manufacturing Defects
- Medical Record Errors
- Power Plant Failures
- Ordering Mistakes
- Shipping Delays
- Employee Injuries
- Contracts not Awarded
- Construction Errors
- Customer Complaints
- Wrong Diagnoses (car repair, patients, TV/VCR's, etc.)

We'll generalize this below (see Types of Problems), but these examples are a good place to start.

How should the Effect be Measured? - The next question addresses how we'll measure the problem. Often, frequency is used, simply, how many times has the problem occurred? Cost, though, is a better measure of the problem.

For example, some manufacturing defects may be possible to rework, the cost of these defects is the rework cost. Others may require that the product be scrapped at a higher cost. Likewise, an ordering mistake that sends the customer a higher quantity than ordered may not be as costly as a mistake that sends the order to the wrong customer.

How can the Effect be Stratified? - If you are just beginning to understand your problem, then you'll be using the "4W" categories:

- What types of problem are there? Are there different types of problems that occur more frequently than others?
- When do the problems occur? Are there differences by day of week, shift, time of day?
- Where do the problems occur? Are there differences by location, area, plant, or production line?
- Who performs the work that produces the problems? Are there differences by operator, physician, teller, technician, or manager?

If you are trying to use Pareto Analysis to understand the causes of the problem, then you'll be looking for why or how categories. Your Cause-and-Effect diagram will have identified these why or how categories. This is an important distinction. Pareto is often used early in an improvement to stratify by phenomena or symptoms. Later, Pareto is used to stratify by cause (method, machine, material, personnel, etc.).

For many industrial problems, categories will already exist that you can use to start your Pareto Analysis. For example, employee injuries are categorized by type (or what) such as strains, sprains, lacerations, breaks, shocks, etc. Often, there will be some initial, logical categories you can use.

These categories are good to start with, but we've seen their use result in many "mechanical" Pareto analyses. Even though these categories don't lead to a good 80/20 split, the improvement effort will still simply pick the highest frequency category and attempt to work on preventing this category of problem.

You should try, though, to stratify the data in different ways. You may learn something. For instance, one company stratified employee injuries by where they occurred. They found that over 90% of the injuries were occurring in the office buildings, not in the "field" as they expected.

What does the Data tell you? - Collect or organize your "problem" data according to the categories you've chosen. Construct the Pareto Chart. Do you "see" an 80/20 split? Have you identified the Vital Few categories that contribute the most to your problem?

If so, then move on to find out why these vital few are occurring. Leave the Useful Many for later (unless some of these are very easy to address).

If not, then go back and think of your problem from a different angle. How else could you stratify the problem? As one of our friends says, ”Lather, rinse, repeat.”

Types of “Problems”

Pareto Analysis works best with a zero-type problem. This is a problem whose desired “level” is zero. Errors, defects, injuries, accidents are all zero-type problems. Ideally, we would like to have none of these problems occur (regardless of the current practicality of zero defects). If you have a zero-type problem, then Pareto Analysis can be used directly. Collect the data by category or strata and construct the Pareto Chart.

There are two other kinds of “problems” that you’ll deal with, though. The first is a decrease-type problem. For example, you may want to minimize the time required to perform some process. The “ideal” process time here is not zero, but you would like to eliminate any wasted time or non-value-added time.

You can still do a Pareto Analysis, but you’ll want to transform your decrease-type problem into a zero-type problem.

> Example: A hospital wanted to decrease Lengths of Stay for certain types of patients. They began their analysis by collecting “Lost Days” data - days where the patient was in the hospital, but didn’t need to be. Ideally, the number of Lost Days is zero, so they transformed a decrease-type problem into a zero-type problem.

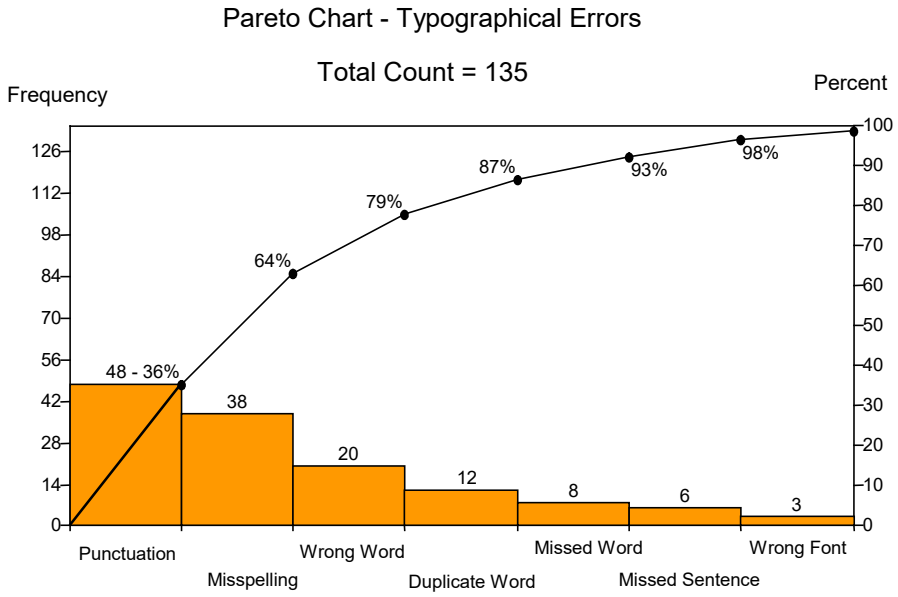
The other type of problem is an increase-type problem. Sales or volumes are examples of increase-type problems. Here, too, you’ll want to transform the increase-type problem into a zero-type problem as part of your Pareto Analysis.

> Example: A utility was promoting the sale of water heater heat pumps. Based on their market research, they had predicted the number of sales by geographic area, customer type and income. When the expected sales did not materialize, they performed Pareto Analysis by measuring the “gap” between actual vs. predicted sales, using the different categories. Measurement of the “gaps” turned the increase-type problem into a zero-type problem.

7.2.3 The Pareto Chart

Purpose

The Pareto Chart is a special kind of bar chart that displays the results of a Pareto Analysis. The Chart shows, immediately, the Vital Few categories that contribute the most to the problem. A simple Pareto Chart of typographical errors is shown below:



There are two parts to the Pareto Chart. The bar chart shows the contribution of the individual categories (in order) to the overall problem. The line graph shows the cumulative impact of the categories, from largest to smallest. Three types of typographical errors, Punctuation, Misspelling, and Wrong Word make up almost 80% of all errors. These are the Vital Few.

Some statistical tools are “Ah-hah!” tools. For instance, when you take some data and construct a histogram, there is an “Ah-hah!” the moment the histogram appears. The Pareto Chart is not like that. The “Ah-hah!” comes when you collect the data and organize it by category. The Pareto Chart’s purpose is to communicate the results of your analysis to others.

Application

The Pareto Chart is applied whenever a Pareto Analysis is performed. Generally, the Pareto Analysis will be performed during these steps of a quality improvement effort:

Identify the Problem - For the problem being addressed, which are the most important categories? Which are the Vital Few versus the Useful Many?

Analyze Causes - For the problem being addressed, which causes of the problem appear most often?

Implement and evaluate results - After changes have been made, has the problem been reduced in frequency or cost? Was the category of problem reduced? Has the cause of the problem been eliminated? You can see that Pareto is widely applicable.

Construction

Note: These construction steps assume the Pareto Chart is being prepared as part of a Pareto Analysis.

1. Collect data on the frequency or cost of the problem, stratified by the categories you think are important.

2. Order the categories from largest to smallest contributor to the problem. Note that if several categories are very small contributors to the problem, you can group them into an “Other” category. Just make sure that this “Other” category doesn’t make up more than about 20 - 25% of the total. Even if the “Other” category is larger than some of the individual categories, always put it last.

3. Add up the individual categories, from largest to smallest to obtain the cumulative values. Note that you can also calculate the cumulative percentages if you want to label the “cum line” with these.

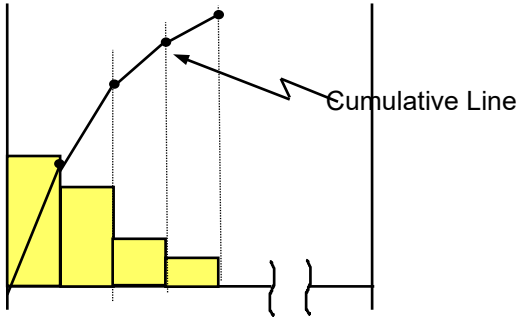
The following table summarizes these calculations:

Category	Contribution	Cumulative	Cumulative Percent
Punctuation	48	48	36%
Misspelling	38	86	64%
Wrong Word	20	106	79%
Duplicate Word	12	118	87%
Missed Word	8	126	93%
Missed Sentence	6	132	98%
Wrong Font	3	135	100%
Total	135	-	-

4. Draw the left vertical axis and scale it from zero to the total of all categories. Draw the horizontal axis, and divide it equally into the number of categories you have. Draw the right vertical axis and scale it from zero to 100 percent (make the 100% point even with the total on the left vertical axis).

5. Draw the individual categories as bars on a piece of graph paper. If you have grouped several categories into the "Other" category, draw this as the right-most bar.

6. Draw the cumulative line as a series of line segments, starting from the "0" point on the left vertical axis and finishing at the 100% point on the right vertical axis. The segments end at the right side of each bar:



7. Title, label and date the Pareto Chart. Note the dates the data was collected and who prepared the chart.

Interpretation and Action

Interpretation of the Pareto Chart is simple: Does the Pareto Principle appear with the stratification strategy you've employed? If it does, then you can take the next steps in your improvement effort, if not, try to identify a different stratification scheme.

Even if the Pareto Principle appears on your first try, you may want to examine the data from other angles. It never hurts.

After you've gone to the trouble of collecting the data and performing the Pareto Analysis, now what? Let's return to the basic purpose of Pareto Analysis - stratify and **prioritize!** If you're trying to pick the most important problem to address through an improvement effort, the "big bars" on the Pareto are the ones on which to focus. If you've gathered data on

the causes of your problem, these same “big bars” are the variables you should change to reduce the frequency or cost of your problem.

Pareto causes us to make two important and related choices:

We **will** work on the Vital Few factors, and we **will not** work on the Useful Many.

Too often, organizations expect their people to work on everything, with little prioritization. Pareto forces us to make choices.

Don’t try to improve more things than you have fingers on one hand (A good “thumb rule[2]”).

One of our CEO friends implemented this philosophy beautifully and simply - He asked each of his VP’s to identify the three most important issues in their departments each month and tell him what their plans were to address them.

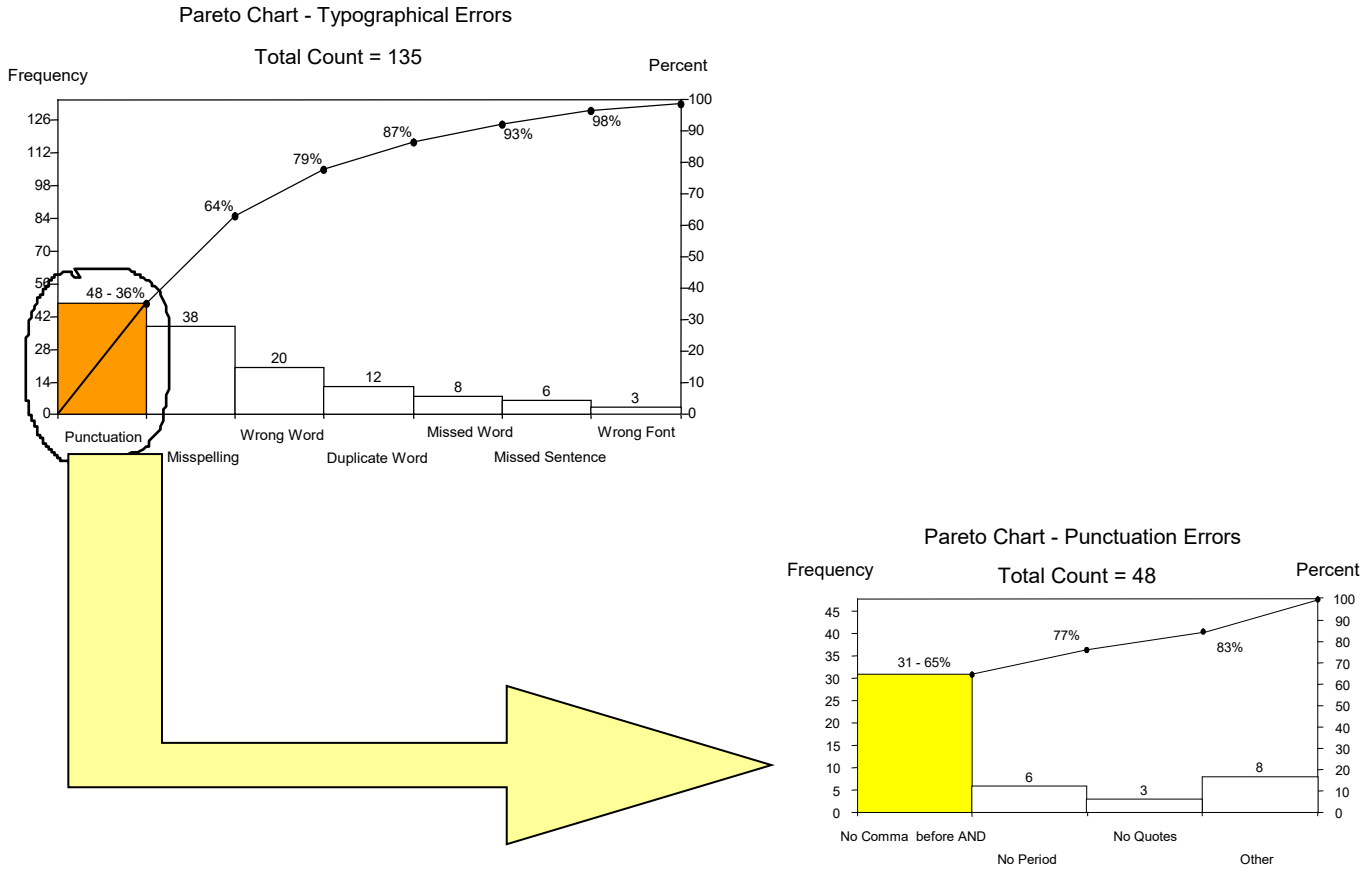
[2]Pun intended!

Pareto Pointers

Here are a few points on the practice of Pareto Analysis that we've found helpful in quality improvement work:

Multiple Stratifications

You may find that, after your first "level" of stratification, that the data can be stratified further. This can be a great strategy for really focusing on the problem. Be careful, though. If you stratify too many levels, you're in danger of entering the *Pareto Death Spiral*. You don't want to wind up focusing on a very tiny portion of the overall problem.



7.2.4 Contingency Table Analysis

Purpose of a Contingency Table

The Pareto Diagram allows us to perform a *one-way* stratification of data. We take some *effect*, identify categories and then see how much of the effect is due to each of the categories. A more general approach to attacking the stratification issue, especially if we are dealing with a *discrete* (or *count*) effect is the *Contingency Table*. This approach allows us to employ a *two-way* stratification of a group of items.

Contingency Tables and Their Use

The easiest way to introduce Contingency Tables is to show you a few examples and how they can be used in process improvement. In this first example, we are trying to determine if administering two different types of flu vaccine made a difference in the proportion of people contracting flu. The contingency table is used here to explore a *cause-and-effect* relationship:

2 x 2 Contingency Table - 2 Rows, 2 Columns

Flu Vaccine Type

	Shanghai	Malaysian	Total
Contracted Flu	673	816	1489
Did Not Contract Flu	2880	2194	5074
Total	3553	3010	6563

In this next example, the contingency table examines the differences in infection rates across hospital units. This may be done to help in understanding the *current situation:*

2 x 5 Contingency Table - 2 Rows, 5 Columns

Hospital Units

	2E	2W	3N	3W	4E	Total
Infections	5	3	6	4	7	25
No Infection	124	212	186	134	303	959
Total	129	215	192	138	310	984

In this last example, we are examining a *cause-and-effect* relationship, but the contingency table shows its power by allowing us to determine four different levels of the factor, *Number of Quality Improvement (QI) Courses*, against three different levels of the effect, *post-test performance:*

4 x 3 Contingency Table - 4 Rows, 3 Columns
Post-Test Performance

	Bad	Average	Good	Total
0 QI Courses	20	20	10	50
1 QI Courses	10	30	15	55
2 QI Courses	5	30	15	50
3 QI Courses	5	20	20	45
Total	40	100	60	200

Contingency Table Notation

Let's generalize the Contingency Tables examples shown above. The notation for the elements of the Contingency Table is shown below:

n COLUMNS

m ROWS

	A_1	A_2	...	A_n	Total
B_1	X_{11}	X_{12}	...	X_{1n}	$X_{1.}$
B_2	X_{21}	X_{22}	...	X_{2n}	$X_{2.}$
B_3	X_{31}	X_{32}	...	X_{3n}	$X_{3.}$
.	.	.	...	.	.
B_m	X_{m1}	X_{m2}	...	X_{mn}	$X_{m.}$
Total	$X_{.1}$	$X_{.2}$	...	$X_{.n}$	$X_{..}$

The *X's* are the values of the variable being measuring. Each element of the matrix is the value of the variable for the combination of *attributes* (*A's* and *B's*) we are exploring. In our first example, X_{11} = 673, this represents the number of people who contracted the flu *and* that received the Shanghai vaccine.

We'll use these *X's* in the Contingency Table Analysis to perform a hypothesis test, like those described in Section 9. Note the symbols used for the Row and Column totals. We define these as follows:

$$X_{\cdot j} = \sum_{i=1}^{m} X_{ij} \qquad X_{i\cdot} = \sum_{j=1}^{n} X_{ij} \qquad X_{\cdot\cdot} = \sum_{i=1}^{m}\sum_{j=1}^{n} X_{ij}$$

Although not strictly required, a good convention is to assign the *rows* to be the *effects*, or *dependent variable;* the *columns* then become the *factor* or *independent variable.*

Contingency Table Analysis

The contingency table analysis process is performed as follows:

1. Establish the Hypothesis:

   Null Hypothesis (H_O) **-** There is no relationship (i.e. independence exists) between Attributes A and B.

   Alternative Hypothesis (H_a) **-** There is a relationship (i.e. dependence exists) between Attributes A and B.

   In the flu vaccines example, the hypotheses can be stated as follows:

   Null Hypothesis (H_O) - There is no relationship between the type of flu vaccine administered and the occurrence of flu.

   Alternative Hypothesis (H_a) - There is a relationship between the type of flu vaccine administered and the occurrence of flu.

2. Choose the Significance Level of the Test (α).

3. Plan the Test:

   a) The Test Statistic is:

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

where:

$k =$ *number of cells in the table* $(m \times n)$

$O_i =$ *observed count for cell i*

$E_i =$ *expected count for cell i*

(assuming H_o *is true)*

To help perform this calculation (without a computer program), it is helpful to set up an *Expected Counts Table*, right below the Contingency Table:

Observed Counts
Flu Vaccine Type

	Shanghai	Malaysian	Total
Contracted Flu	673	816	1489
Did Not Contract Flu	2880	2194	5074
Total	3553	3010	6563

Expected Counts
Flu Vaccine Type

	Shanghai	Malaysian	Total
Contracted Flu	806.1	682.9	1489
Did Not Contract Flu	2746.9	2327.1	5074
Total	3553	3010	6563

The cell values in the *Expected Counts* table are those that would be expected to arise if there *were no difference in the treatments* (i.e. the null hypothesis).

The easiest way to obtain the *expected* cell values is to calculate the proportions from the *totals* column of the observed counts table and "back calculate" the *expected* counts' cell values:

Expected Proportion Contracting Flu = 1489/6563 = 0.227

leads to:

Expected Count for Shanghai Vaccine = 3553 x 0.227 = 806.1

Expected Count for Malaysian Vaccine = 3010 x 0.227 = 682.9

and:

Expected Proportion Not Contracting Flu = 5074/6563 = 0.773

leads to:

Expected Count for Shanghai Vaccine = 3553 x 0.773 = 2746.9

Expected Count for Malaysian Vaccine = 3010 x 0.773 = 2327.1

One condition we impose on this analysis is that the *expected cell counts* should be greater than or equal to 5. As we saw above, the *relative proportions* and the total number of events influences the expected cell counts. Practically, if the relative proportions are small (e.g. 0.001), then to meet this condition, a large sample size will be required (e.g. 0.001 x 5000 = 5).

b) Determine the Rejection Region:

Appendix A provides a table of the χ^2 distribution. Find the table value for *(m - 1)(n - 1)* degrees of freedom at the α level of significance. For example, for a *4 row, 3 column* contingency table, *m = 4, n = 3* and the χ^2 value for 6 {(4 - 1) x (3 - 1) = 3 x 2 = 6} degrees of freedom at the 0.05 level of significance would be obtained from the look up table (this value is 12.59). The flu example is a 2 x 2 table, which therefore has {(2 - 1) x (2 - 1)} = 1 degree of freedom. From the table, then, the critical value (at the 0.05 level of significance) is 3.84.

4. Collect the Data and calculate the Test Statistic. Data should then be collected and sorted into the cells and the expected counts table prepared. Now the test statistic can be calculated:

$$\chi_0^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi_0^2 = \frac{(673 - 806.1)^2}{806.1} + \frac{(816 - 682.9)^2}{682.9} + \frac{(2880 - 2746.9)^2}{2746.9} + \frac{(2194 - 2327.1)^2}{2327.1}$$

$$\chi_0^2 = 21.98 + 25.94 + 6.45 + 7.61$$

$$\chi_0^2 = 61.98$$

5. Draw the Conclusion. The last step is to compare the *calculated value* of the test statistic to the *table value* obtained from the chi-squared table in Appendix A. If the calculated value is greater than the table value, then it falls into the *rejection region.* The null hypothesis would then be rejected in favor of the alternative hypothesis.

In this example, 61.98 (calculated value) is greater than (>) 3.84 (table value). We would then *reject* the null hypothesis and conclude that there is a relationship between the type of flu vaccine and the occurrence of the flu.

On the other hand, if the calculated value were less than the table value, then we would conclude that we could not reject the null hypothesis. Note that we *do not* conclude the null hypothesis is *true,* merely that we have insufficient evidence to reject the hypothesis.

7.3 Exercises

Controller Failures

In the last six months, HVAC controllers from three manufacturers have failed while in service. As part of their improvement effort, a team identified how many controllers were installed (by manufacturer). They also counted the number of failures experienced:

Controller Failures

Manufacturer	# Installed	# Failed
Jonson	24	7
Airaid	32	9
BlowPulse	9	2

How would you display this data graphically? Do so. Do you think there is a difference in reliability by manufacturer? Perform a contingency table analysis of the data. What does this test tell you?

Consumer Satisfaction

The following data were obtained from a consumer survey of products and services. Consumers were asked to categorize the products and services according to the "value" they thought they received. Plot the data on a bar chart. What conclusions do you reach?

Product or Service	Percent Saying "Good Value"
Doctor's Fees	34.7
Health Insurance	29.0
Hospital Charges	21.0
Lawyer's Fees	34.7
Poultry	66.4
Videotape Rentals	65.7
Women's Apparel	50.8

Machine Setup

An improvement team collected data on the time segments that contribute to *machine setup* after the preceding operation and before the next. Plot these time segments on a pie chart. If you were attempting to reduce setup time, does this chart give you any clues about where to focus? Why or why not?

Machine Setup Sub-Processes
(Average of 25 setups)

Time Segment	Time (min.)
Clean Fixtures	3.96
Down Time	6.12
Setup	19.6
Fit Assembly Pieces	5.64

Employee Perception

A large engineering firm conducted a survey of employees one year and two years after introduction of their Total Quality Management system. The questions were designed to determine how employees perceived progress made by management in transforming their style and practice of management. Prepare a radar chart and plot both of these survey results on the same chart. What changed from year one to year two? Where is management strongest, weakest in TQM?

	Average Score	
Survey Question	**Year 1**	**Year 2**
1. Company culture supports quality.	6.2	7.5
2. Company uses data in decision-making.	4.0	4.5
3. Quality led by senior management.	6.0	6.5
4. All company employees involved.	3.3	7.5
5. Practices quality principles.	5.2	5.4
6. Teams used to achieve important goals.	5.8	7.8
7. Engages suppliers in improvement.	3.0	3.2
8. Customer input used to support decisions.	4.6	6.5
9. PDCA practiced in daily management.	5.7	5.7
10. Supports quality in community.	4.3	4.4
11. Proactive with regulatory agencies.	8.0	8.2
12. Promotes quality education.	4.5	7.8
13. Quality objectives clearly defined in strategic plan.	5.0	4.2

Customer Complaints

This is a common example of *available* data, often *not used* for improvement. A customer service manager keeps a Complaint Log, where every complaint by a customer is dutifully noted. The *immediate remedy* taken to resolve the complaint is also noted.

Over a six-month period, though, here are the *recurring* complaint types and their frequencies. Prepare a Pareto Chart of these complaints. If you were the manager "in search of" opportunities to improve service, which category would you address first?

Customer Complaint Log
(Jan-Jun '99)

Complaint Category	Frequency
Question Not Answered	72
Wrong Question Answered	56
Timeliness of Answer	102
Personnel Response	21
Wait Time on Telephone	25
Error in Information	42

Injury Reduction Efforts

a) The senior management of a factory set as a strategic objective the reduction of injuries to factory workers. They began by collecting data on the frequency and cost of injuries. Prepare Pareto Charts for this data by both of these measures. Which category(s) should management work on first?

Air Handler 'R' Us, Inc.
Employee Injuries – 1999

Injury Type	Frequency	Cost ($)
Cut	2	2742
Fiber in Eye	9	469
Lift	15	13,597
Pull	14	109,115
Puncture	1	1368
Slip/Fall	21	354,739
Struck Against	13	149,049
Struck By	12	2725

b) Slips and falls were then examined to determine if there was a particular type that resulted in employee injuries. The following data was collected:

Air Handler 'R' Us, Inc.
Employee Injuries - 1999
Slips and Falls Category

Type of Slip/Fall	Frequency
Wet Floor/Object on Floor	14
Steps	4
Platform	2
Trailer	1

Develop a Pareto Chart of this data. If you were in this company's situation, what would be your next steps?

Handwritten Checks

A payroll supervisor was working on reducing the number of "handwritten checks." These are employee payroll checks issued by hand, due to some error in the computer-generated check. She told us that each handwritten check was estimated to cost the company about $60.00. Develop a Pareto Chart of this data. Which category(s) would you work on first?

Handwritten Checks by Category
(Jul - Dec '95)

Check Category	Frequency
Hours not Keyed	8
Hours not Paid/Left Off	46
Cash Plus not Paid	8
Cash Plus Sellback	48
Shift Bonus	0
Paid Wrong Rate	12
System Error	22
Badge Error	4
Refunds	8
Check Issued/Void or Lost	8

Incorrect Billing

The Billing department criticized the shipping department in a plant for incorrect invoices. The shippers all had their own opinion why the bills were incorrect. They collected data over a one-week period, with the following results. Develop a Pareto Chart of this data. What would your next steps be?

Shipping Department
Incorrect Billing Categories

Category	Frequency
Bill of Materials and Order Ticket differ	22
Inconsistent charging for special features	45
Incorrect computer keying	3
Other	3

Treatment Costs

A hospital that tracked the Length of Stay data for the diagnosis, Coronary Bypass with Cardiac Catheterization, began an improvement effort to reduce the *unnecessary costs* of this diagnosis. They collected data on the charges associated with 13 patients who fell into this diagnosis. Prepare a Pareto Chart of this data.

Coronary Bypass with Cardiac Catheterization
DRG-106 (13 Patients)

Category	Charges ($)
Anesthesia	498
Cardiac Cath Lab	3170
Cardiac Diagnosis	546
ICU/CCU	3336
Lab & Blood	3183
Operating Room	6356
Other	347
Pharmacy	5182
Radiology	475
Regular Room	1602
Respiratory Therapy	2193
Step Down	438
Supplies	4863

Does this Pareto provide you with clues as to where to begin to reduce *unnecessary costs?* What's the problem with this Pareto? (Hint: what kind of "problem" is this?).

Safety System Failures

A nuclear industry “watchdog” group collected data on reports of failures occurring in nuclear plant safety systems. Prepare a Pareto Chart of these failures. What systems seem to need reliability improvement the most?

Nuclear Plant Safety System Failures

Safety System	No. Of Failures
Reactor Coolant System	24
Control Rod Drive Mechanism	5
Steam Generators	2
Reactor Water Cleanup	5
Feed Water	12
Main Steam	19
Normal AC Power	13
Emergency AC Power	62
Other	11

Light Rail Fatalities

In a recent USA Today newspaper article, the following table of fatalities associated with light-rail system fatalities was presented. What possible Pareto Analyses could you perform from this data? Do so. Which gives the "best" picture of the situation? Why?

City	Track Miles	Daily Ridership	Fatalities
LA	22	70	61
San Diego	47	75	22
Portland	38	81	14
Sacramento	21	31	14
San Jose	31	30	9
San Francisco	73	164	8
Philadelphia	69	84	7
Boston	51	231	6
Denver	17	35	6
Salt Lake City	18	28	5
Baltimore	29	284	4
Dallas	44	39	3
New Orleans	16	14	2
St. Louis	34	42	2
Pittsburgh	18	25	2
Buffalo	6	23	1
Cleveland	15	15	0
Newark	9	8	0

Production Scheduling

A plant scheduling team tracked the frequency of delays for large air handlers. Plot their data on a Pareto Chart. What suggestions do you have for their next steps? Is there a different way they could have performed the Pareto Analysis that might reveal a different picture?

Delay Category	Frequency
Equipment	31
Operators Missing	6
Welding	12
Fit up Time	97
Engineering	67
Can't find material	1
Supplier	20
Machine Setup Time	121
Repositioning	113
Quality Control	4
Procedure Added	6
Other	35

Labor and Delivery

A Labor & Delivery team is investigating the relationship between the mother's dilation when an epidural is administered and the C-Section rate. Four dilation ranges were identified and C-Section rates measured for two months. Perform a Contingency Table analysis of the data. Use $\alpha = 0.05$. Is there a difference?

	Dilation (cm)			
Delivery	**0 - 2.5 cm**	**2.5 - 5.0**	**5.0 - 7.5**	**7.5 to 10**
C-Section	48	51	28	12
Vaginal	142	219	272	228
Total	190	270	300	240

Electrical Wiring Errors

The following data represents the number of wiring errors stratified by operator. Is this a valid comparison, i.e. should we focus our improvement efforts on the operator with the highest number of errors?

Operator	**A**	**B**	**C**	**D**	**E**	**F**
# Errors	46	22	119	82	61	30

Welding Errors

A Black Belt is studying the occurrence of welding errors during pressure vessel fabrication. She questions whether the welder performing the procedure makes a difference. Perform a Contingency Table analysis on the following data. Use an alpha (α) of 5%.

Welder	**A**	**B**	**C**	**D**	**E**	**F**
Number of Welding Errors	14	25	10	21	21	32
Total Number of Welds	52	192	171	137	80	195

Compressor Failures

A Black Belt is studying compressor failures that have been occurring in air conditioning units. She wonders if the manufacturer of the compressor is a factor. Perform a contingency table analysis of the following data. Test for an α = 0.05.

	Presshard	Blows-a-lot	HighHead	Totals
Failed	7	9	5	21
Not Failed	24	32	18	74
Total	31	41	23	95

8.0 Cause & Effect

Unit	Description	Page
8.1	Cause and Effect Analysis	8.1 - 1
8.2	Exercises	8.2 - 1

Process Analysis gives us a start in determining cause and effect relationships. Here, we explore ways to develop hypotheses about potential process variables that can impact performance and how to determine (with facts) that these are the “true” or root causes of the performance defects.

8.1 Cause and Effect Analysis

Learning Objectives

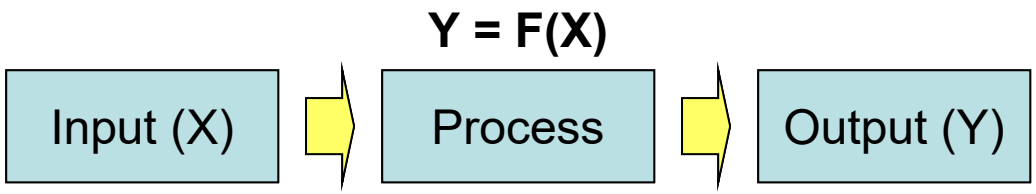
- Understand the Nature of Cause and Effect
- Develop Cause and Effect Diagrams for Problems
- Verify Cause and Effect Relationships

Unit Contents

- The Nature of Cause and Effect
- The Cause and Effect Diagram
- Root Causes Verification

8.5.1 The Nature of Cause and Effect

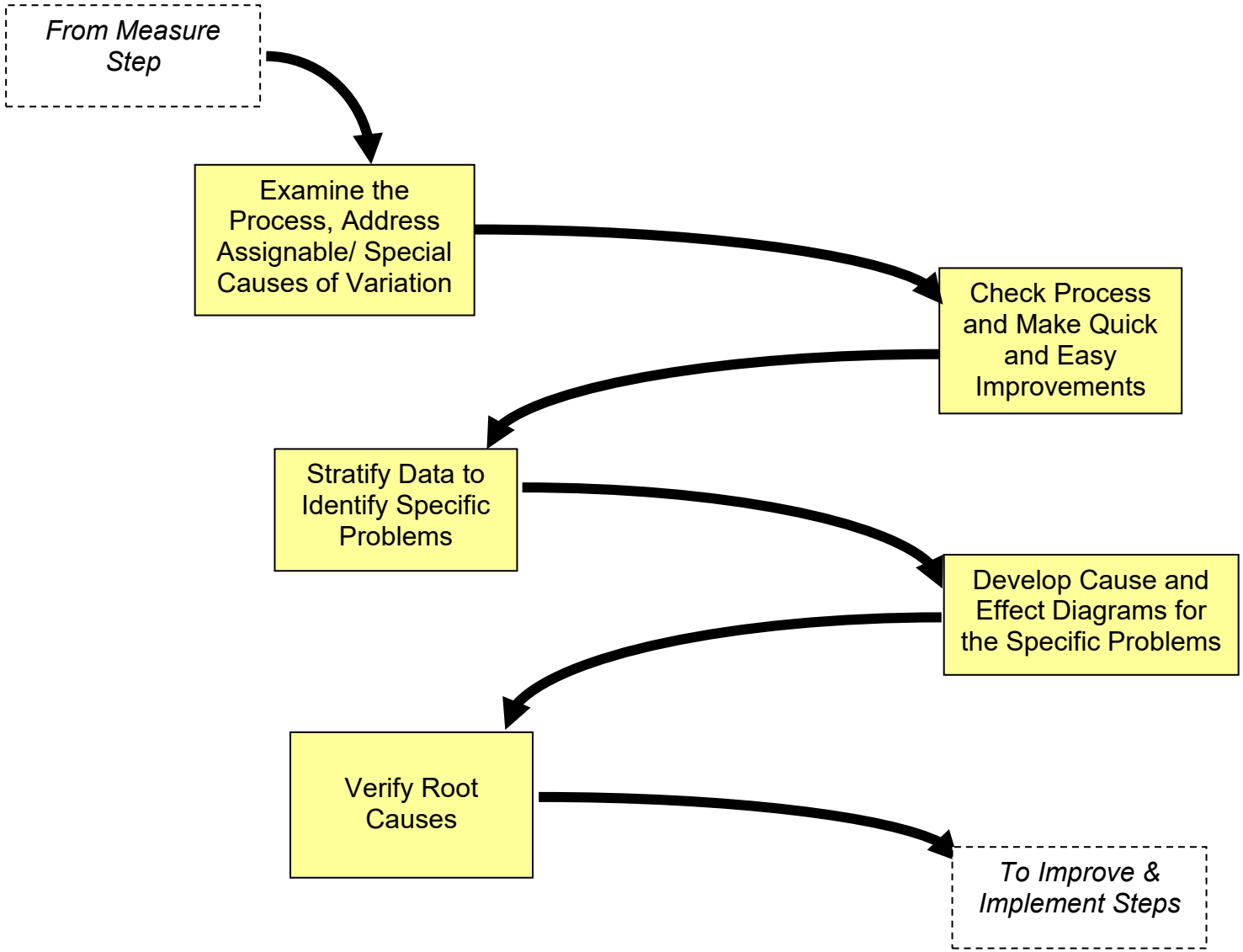
Cause and Effect is a critical element of quality improvement. To improve quality, we must understand the factors that contribute to variation in quality. Through this path we can identify countermeasures that have a high probability of improving the quality of our products and services. It's easy to "jump to solutions," harder to understand the real factors that affect performance. The process (methods, machines, materials, personnel, & environment) and inputs produces your products and services.



"X's" Include:

- **Methods**
- **Materials**
- **Machines**
- **Personnel**
- **Measurement**
- **"Mother Nature"**

Road Map for Analyzing Causes



Examining the process

"Simple" methods of analyzing the cause-and-effect relationships should be tried before the more sophisticated ones are employed (*See Unit 5.4 for "simple" process analysis methods.*):

- Many problems can be solved with a thorough examination of the process and just common sense. Before trying more sophisticated methods, observe the process in detail and talk to the operators. Get a clear picture of the process and its bottlenecks. A flow chart is probably the best tool for this along with questions. Ask who, what, why, when and how?

- Gather the process documents as well for comparison. Get copies of the specifications, drawings and procedures. If they are hard to obtain or do not exist, that may be a clue.

- Compare the actual process to the documents. See if there are disparities and ask why if there are. Determine if the process documents were followed, if the problem would go away. Then ask why the documents aren't being followed?

- See what training is given to operators and if they are qualified in their process.

- Gather data on the process' performance (Key Characteristics – quality, cost, delivery, safety). Examine the data for assignable/special cause signals. Address the causes of these signals (variation across machines or operators, non-standard processes in place, variation in supplier inputs).

Quick and Easy Improvements:

During this initial observation of the process, certain improvements may appear to be obvious. If they are agreed to by everyone, they do not cost much money or resources to implement, can be reversed easily and their improvement can be quickly seen, go for it. Quick and easy improvements may include items that:

- Eliminate redundancies
- Clarify unclear steps
- Re-order steps
- Eliminate unnecessary steps
- Decrease the amount of transportation or delays

Stratifying the Data

As part of *Identify the Problem*, you were encouraged to stratify the overall problem, looking for "leverage" points through Pareto Analysis. Even though you may have refined your problem in that step, there are always opportunities to employ stratification.

Involving the Right People

A cause-and-effect analysis starts by employing your current knowledge and experience and proceeds to further investigation of causal relationships using data. If you don't have experience in your problem area (e.g. you may be leading a project in which you have little technical knowledge) then the initial cause and effect "hypotheses" may not lead you in the right direction, resulting in wasted effort and project delays. As you begin this phase of problem solving, make sure that you involve the right people. **Who are these "right" people?**

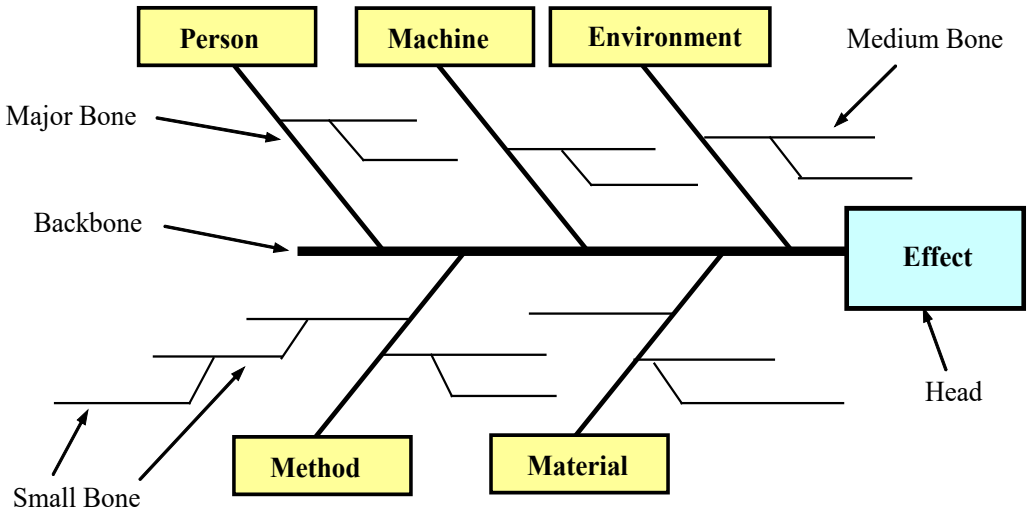
8.5.2 Cause and Effect Diagrams

The Cause-and-Effect Diagram (sometimes called the Fishbone or Ishikawa Diagram) is used as the starting point of a Cause-and-Effect Analysis. Here the Diagram is used to develop hypotheses about the causes of variability or poor performance of the product or service. The Cause-and-Effect Diagram is also used to record data and to note "discoveries" made during the verification step of Cause-and-Effect Analysis.

The advantage of the Cause-and-Effect Diagram is that it provides you with a picture of all the possible causes. Ideas from many different people can be captured on the diagram and the search for important causes then planned systematically. The Cause-and-Effect Diagram helps you avoid the tendency to think of only one possible cause at a time and go searching for that one cause.

Cause-and-Effect Diagram Form

The Cause-and-Effect Diagram is also called the Fishbone because of its appearance. The Effect is shown on the right side of the diagram. The "Major Bones" are general categories of causal factors, with the medium and small bones identifying more specific causes or factors:



The Cause-and-Effect Diagram can be adapted to many different situations. Several different types of Cause-and-Effect Diagrams are shown later in this section.

Application

The Cause-and-Effect Diagram is used most often in Analyze Causes, to discover the important factors that relate to performance of the product or service. Cause and Effect diagrams are also used to educate new employees in the process. They quickly summarize the key factors that are important to ensure the quality of the product or service.

Constructing the Cause-and-Effect Diagram:

1. State the problem as the Effect. Some preliminary data collection or process observation may help you focus this statement. A Pareto Analysis of the problem is often a good prelude to the Cause and Effect. A good "Effect" statement:

- states what is wrong, not why it is wrong,
- focuses on the gap between what is and what should be, and
- is measurable and specific.

2. Identify possible causes of the Effect. Review the problems' occurrences; try to understand how they occurred and what the process situation was when they occurred. Brainstorm a list of causes based on knowledge of the production process.

The phrase "Ask Why Five Times" is often applied to this step of the cause-and-effect analysis. The object is to identify causal factors that can be corrected by changing one or more process factors.

In developing the Cause-and-Effect Diagram, try to avoid:

- Solutions - Solutions are not causes. Solutions will be addressed after the root causes are understood.
- "Lack of" Statements - These are like solutions. The "Lack of X" statement implies that if "X" is present, then the problem will not occur.
- "Fuzzy" causes - Causes such as attitude and morale can be important issues, but the cause-and-effect analysis should try to focus on the process factors that contribute to the effect.

3. Define the major categories to be used on the Cause-and-Effect diagram. The "default" categories of Person, Machine, Material, Method, Information, and Environment are often used for small problems or when just starting to use the Diagram. Major categories such as the process steps associated with producing the product or service should be used for larger problems.

4. Draw the diagram. Write the effect clearly and objectively in a box. Build the diagram by organizing the brainstormed causes under appropriate categories. Lines should flow toward the effect. Refine the causes where necessary and continue asking:

- What causes this?
- Why does this condition exist?

5. When the diagram appears complete, walk through the logic in both directions: a) proceed from the effect to the causes, making sure that the effect can result from the causes and b) proceed from the causes to the effect making sure that the causes can result in the effect. Often, an illogical statement will not surface until the second direction is tried.

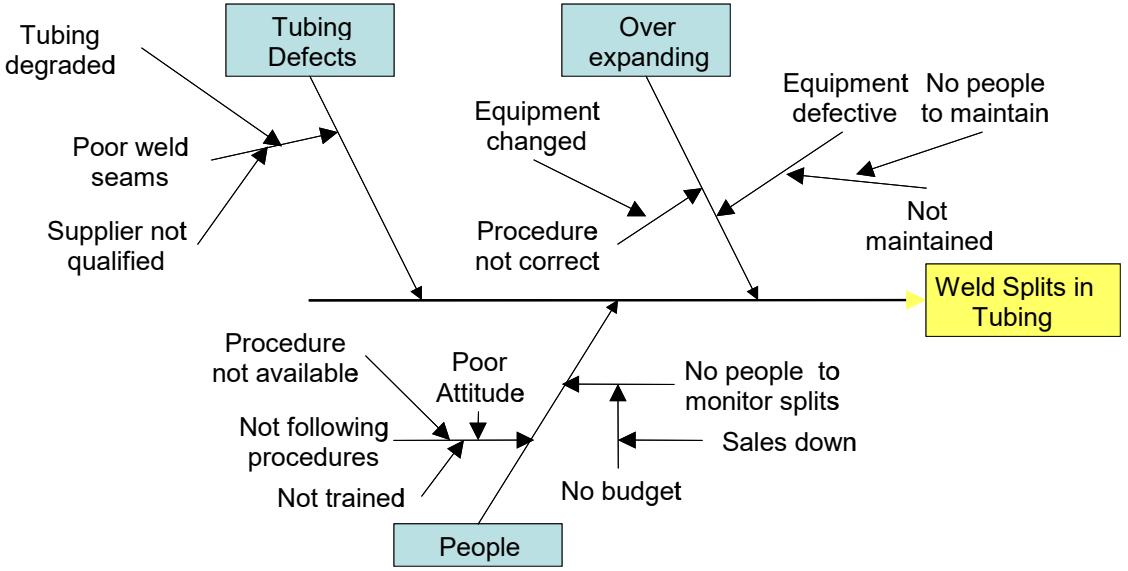
Types of Cause-and-Effect Diagram

Cause and Effect Diagrams are not limited to a specific form. Several different types are presented below:

Variation or Problem Analysis Type

This is one of the most common applications of Cause and Effect. The quality improvement project has identified a specific problem where the *cause(s) are not understood.* The purpose of this type is to determine the factors that could cause the problem or variability in the process' output. Its strength is that it quickly helps organize and relate the different potential causes of the problem.

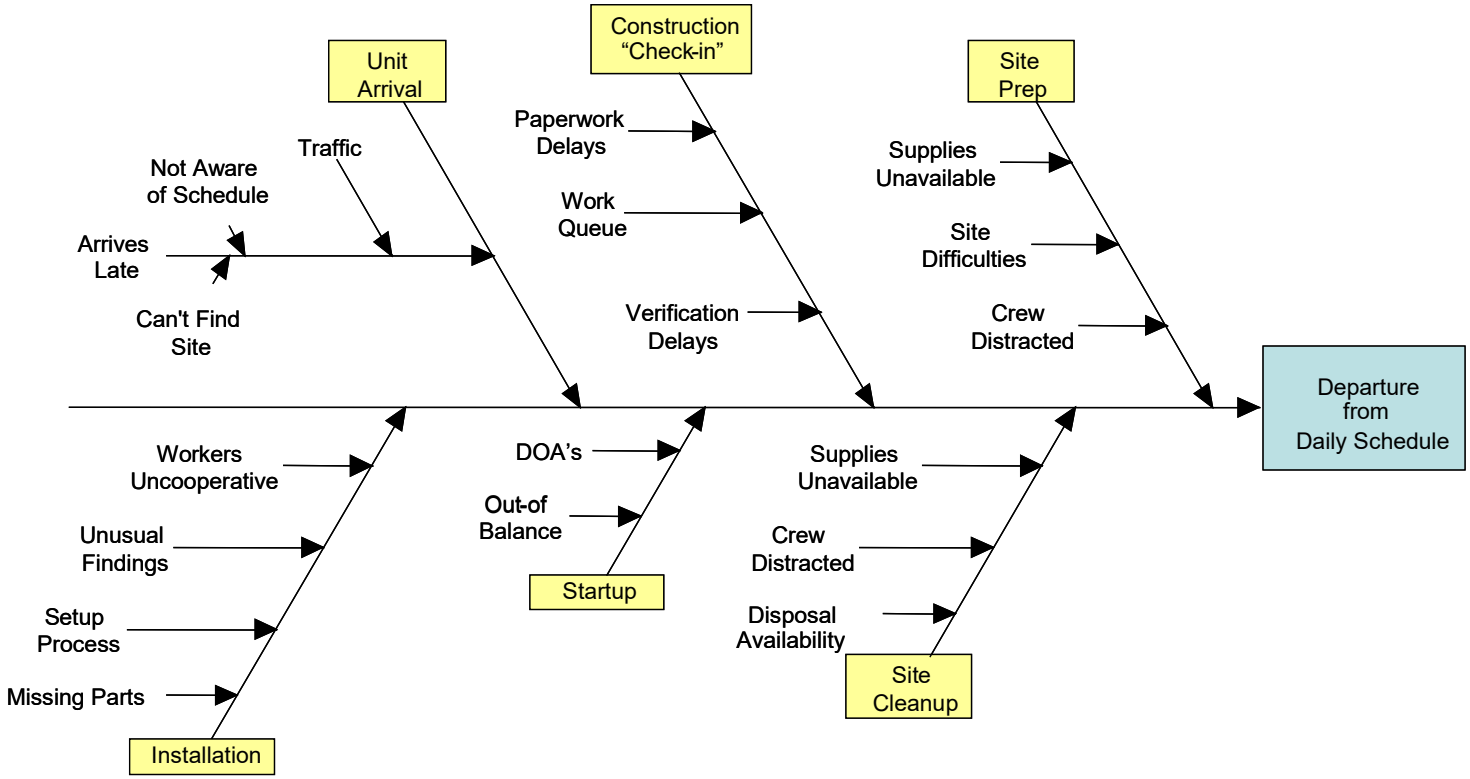
Here, the Cause & Effect Diagram is used to identify possible reasons why weld splits are occurring in tubing. Rather than focus on a "one-at-a-time" approach to diagnosing causes of the splits, the diagram provides the team with the "big-picture" of all possible causes.



Production Process Type

This type of Cause-and-Effect Diagram focuses on the steps of the process and tries to identify the factors in each step that can contribute to the problem or variability. Its strength is that it is easy to create and understand, since each Major Bone considers only one step of the process. It does have a weakness in that similar causes can appear repeatedly, and it's difficult to illustrate situations when the problem is due to more than one factor.

Here, an installation scheduling staff was exploring reasons why their schedule ran late each day. They identified the six segments of the process: *Unit Arrival, Construction "Check-in", Site Preparation, Installation, Startup & Site Cleanup.* The Cause-and-Effect Diagram was then organized with each process segment as a Major Bone.

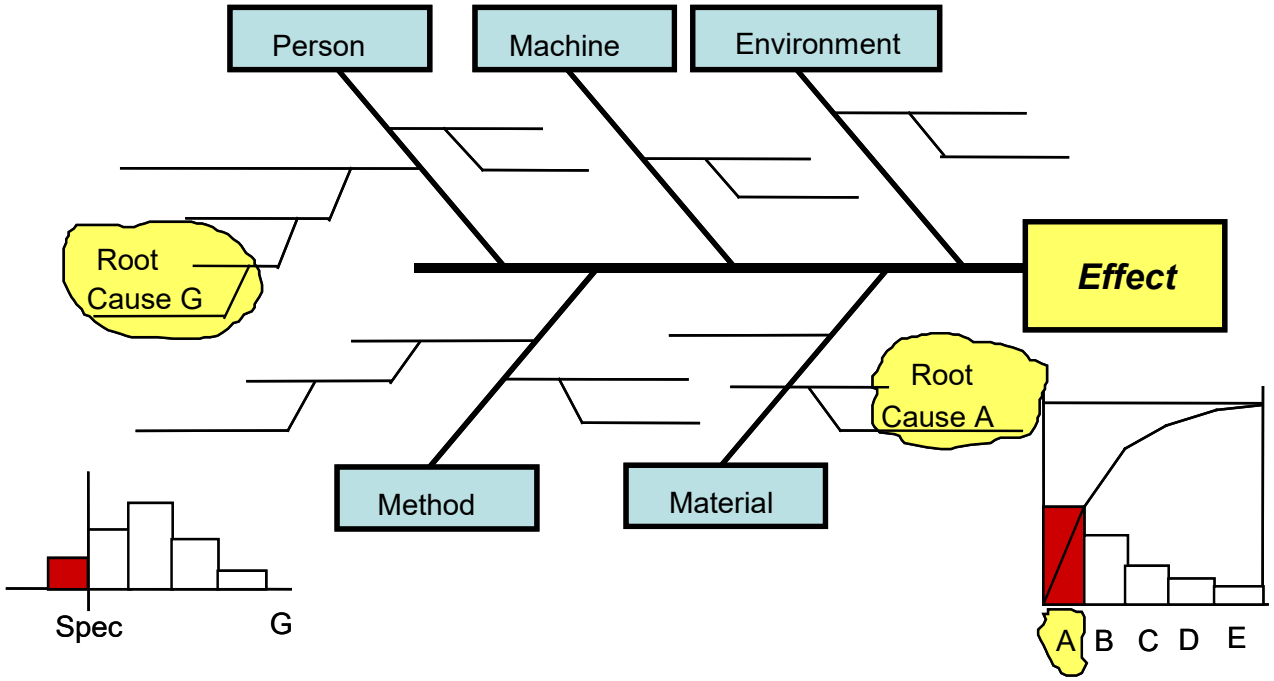


Comparative Cause & Effect Type

This is a strategy to help identify the factors that cause the problems to occur. Two cause and effect diagrams are created. The first lists the factors that are present when the problem occurs; the second lists the factors that are present when the problem *does not* occur. The differences can then be investigated to determine the important factors. Note that this is like *Comparative Process Analysis*, described in Unit 5.4.

8.5.3 Root Cause Verification

Verifying Root Causes is the second step of the Cause-and-Effect Analysis, here you are looking for evidence that one or more factors are contributing to the quality problem or variability. These key factors are often called the Root Causes of process performance. The Pareto Principle applies here: look for the Vital Few causes of variation or problems in your process.



Facts and data must now be gathered to help "convict" one or more potential causes as the actual (or most probable) causes. The verification step generally proceeds by selecting the most likely causes based on the evidence, our experience and "gutfeel."

Then, some method of proving or disproving the potential cause must be determined, data collected and analyzed and the decision made: "Have we found the cause(s)?" If not, then the next set of likely causes is identified and the proving/disproving process repeated.

Methods for Verifying Causes:

One of the first choices you'll make is whether to "play detective" or "play scientist." If you "play detective," then you'll generally be gathering data from the ongoing production process, looking for clues as to which factor is present when the problems occur. If you "play scientist," you will design and conduct experiments (see Unit 11.1), trying to determine which of your process' variables are important and to determine the best "level" at which to set the important variables.

In addition, these general strategies are used to determine the tool that may help you discover the important factors in your process:

Absence/Presence - If a potential cause is present when the problems (the effect) have occurred, and it is not present when the problems do not occur, this may be evidence of a cause/effect relationship. Examples include:

- A pump coupling failed due to the type of material used in the coupling. Mild Steel was used in a saltwater environment instead of Stainless Steel.

- Respiratory infections in an office decreased significantly when the air conditioning filters were cleaned periodically.

- Construction crews were not aware of the availability of an underground cable locator service, contributing to telephone cable cuts during trenching operations. Publishing this information periodically dramatically reduced the number of cable cuts.

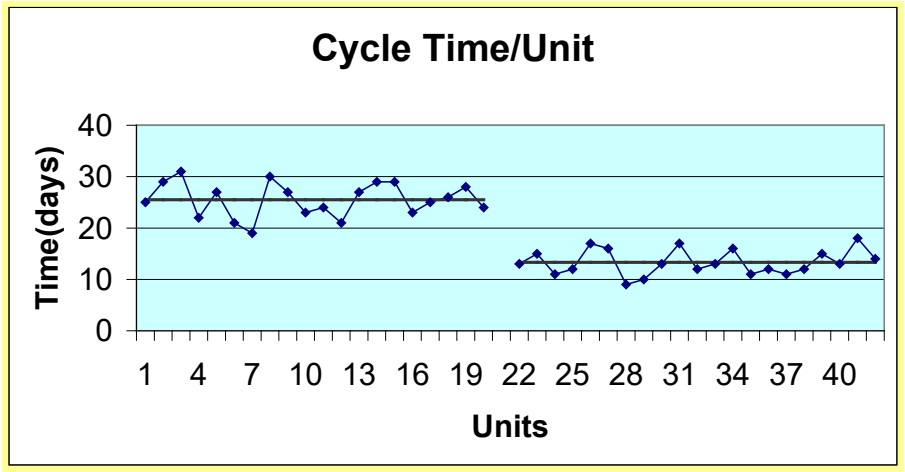
- Seventy-nine percent of electric meter-reader dog bites were due to the "dangerous dog present" field not being entered on the meter-reader's electronic recorder.

- Improper lifting technique was found to be the cause of back injuries in an installation department.

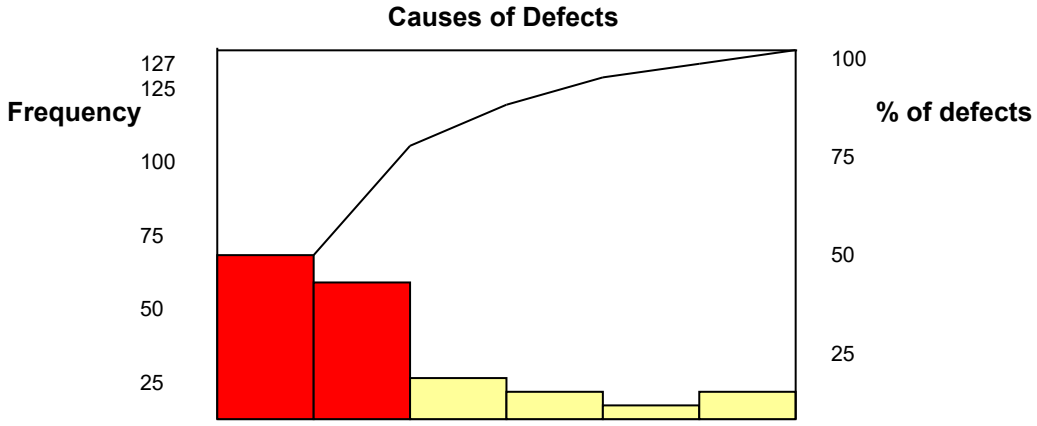
- Pipe leaks were caused by the application of longer bolts than specified. These bolts "bottomed out" before proper compression of the "Y" pipe-engine flange/gasket could occur. Locomotives with short bolts did not experience these "Y" pipe leaks.

Tools Supporting the Absence/Presence Method

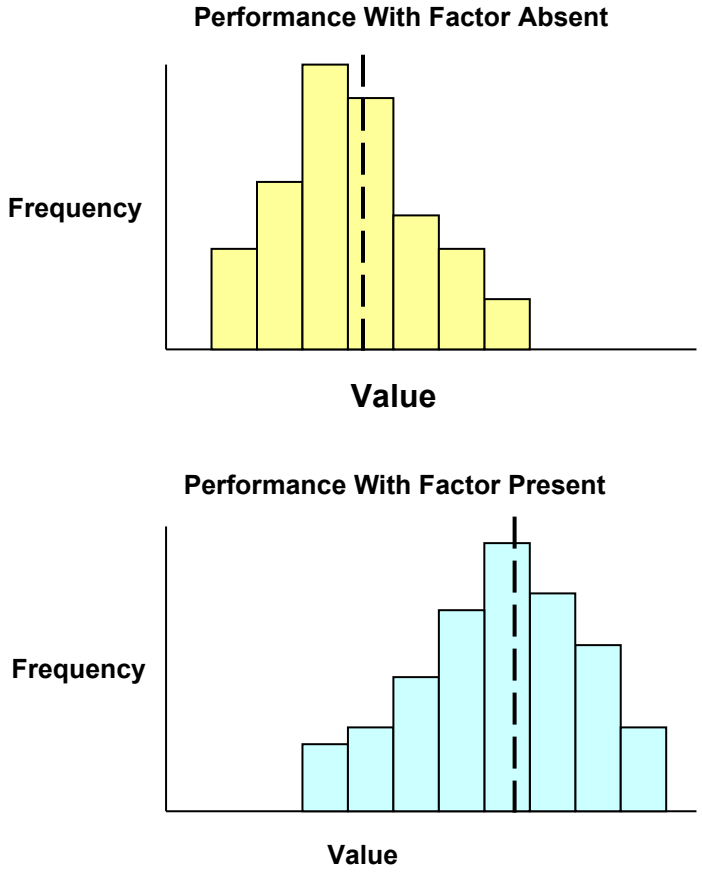
Line graphs or run charts and histograms can be used to show cause and effect relations in the absence/presence method by comparing performance with or without the causal factor present.



Pareto charts can be used to prioritize the known causal factors.



Comparative histograms can show the distribution of data with or without a cause. They are typically shown up and down with the same scale for ease of visual comparison. A frequency chart can be used in the same way for count data.



Hypothesis tests (see Unit 9.2) may help you distinguish between "significant" differences, versus those that arise due to sampling error. **Analysis of Variance** (ANOVA - see Unit 10.3) will help distinguish differences when you are dealing with multiple levels associated with one factor, or if you are dealing with multiple factors. A **Designed Experiment** (see Unit 11.1) will help you plan a rational series of tests to determine which factors are important and to detect if interactions exist between the factors.

Variable Level - The value (i.e. physical measurement or dimension) of a particular factor influences the occurrence of the problems (or effect). Many variables are correlated, that is, if one changes, the other also changes. Here, though, we are attempting to determine either the necessity (the first variable must change for the second to change) or sufficiency (everything else held "constant," changes in the first result in changes in the second). These are not always easy conditions to satisfy. Examples include:

- Cable insulation was found to degrade faster in high temperature locations within the heating unit (i.e. near burners and hot gas exhaust), requiring more frequent replacement to prevent grounds and shorts.

- An increased workload was found to lead to overheating of X-Ray tubes in a welding inspection machine, causing the tubes to fail prematurely.

- Reduced spacing of lightning arrestors on electricity distribution lines was found to reduce the frequency of lightning-related outages.

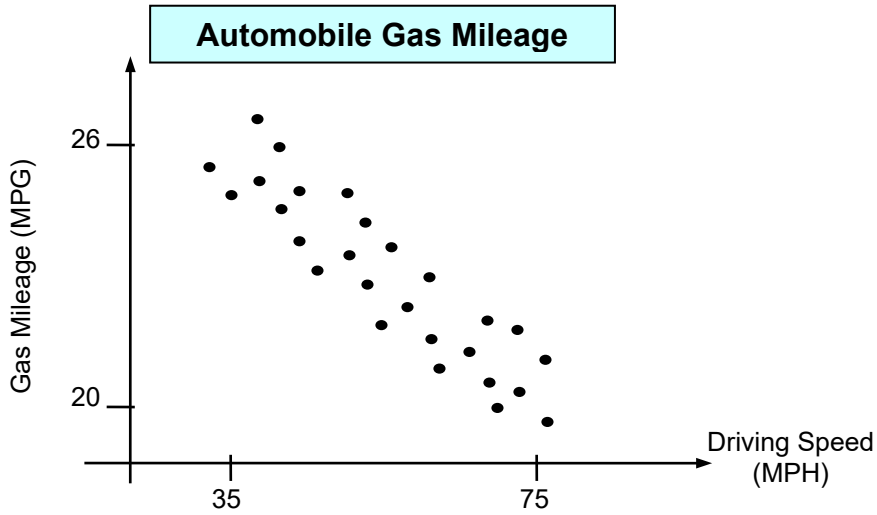
- Oil analysis revealed the presence of high levels of silicon (sand) in gas compressor bearings. This, combined with physical inspection of the inner race sleeves (pitting evidence) and metallurgical analysis of sleeve cracks led to the accumulation of sand into the bearings as the root cause of inner race sleeve failures.

- The boiler superheat temperature was found to be a factor that could be used to control the turbine metal temperature during power plant startup. This control was necessary to avoid high thermal stresses (leading to cracking) in the turbine.

- The number of stress cycles was found to be a factor contributing to leaks in a high-pressure control valve.

Tools Supporting the Variable Level Method

Scatter diagrams (see Unit 10.1) are simple graphical pictures showing relationships between variables. Here, the speed at which an automobile is driven is the causative (or independent) variable. The gas mileage (number of miles per gallon of gasoline) is the effect (or dependent variable). Each point on the scatter diagram represents an observation - for a given driving speed, what gas mileage was observed? From the diagram, you can see that there is a negative relationship[1] (or correlation) between these variables - as driving speed increases, the gas mileage decreases.



Notice that the points do not fall on a straight line. There are other sources of variability at work in this process. Very rarely will a "real world" process display perfect correlation between the variables. The "Punch line" of the Scatter Diagram is important. If there exists a correlation between two variables, then you should be able to change the performance of the effect (perhaps this is some important quality characteristic), by changing the independent variable. The example of gas mileage shows us that we could increase our gas mileage by decreasing our driving speed. Driving speed is something we can control. The Scatter Diagram also shows you how much benefit you'll get from changing the independent variable. In the gas mileage example, it appears that we could gain an additional 6 miles per gallon if we could control our speed at 35 MPH instead of 75 MPH.

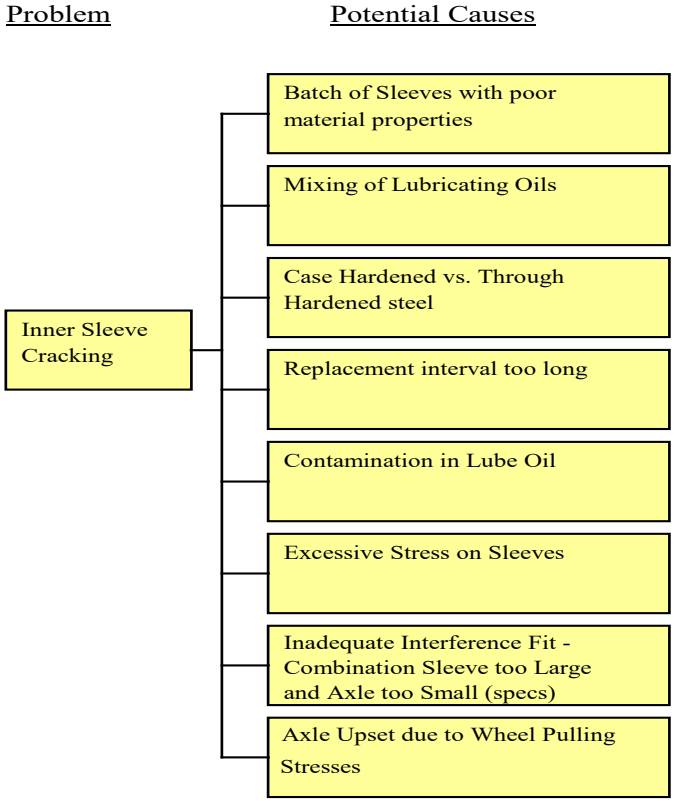
Correlation and Regression Analysis (see Units 10.1 and 10.2) help you explore the strength of the "X and Y" relationship and to develop a mathematical relationship between the two variables. A **Designed Experiment** (see Unit 11.1) can help you plan a rational set of tests to determine relationships when several factors are at work in your process.

[1]The term *negative* refers to the kind of relationship, as the independent variable *increases*, the dependent variable *decreases*. "Negative" doesn't mean it's a *bad* or undesirable relationship.

Cause and Effect Verification Matrix

The Verification Matrix can help plan and track root cause verification efforts. The potential root causes to be investigated are listed and appropriate tests and/or verification activities are then determined ("X's" mean the test in that column will *not* verify that root cause). Test results are recorded in the box that represents the intersection of the potential root cause and its associated verification test.

This example documents some of the verification tests conducted to uncover the cause of locomotive axle bearing “inner sleeve” cracking and failure occurring at one US railroad:



Tests/Verifications

Mat'l Test	Oil Sample	Weibull An'l.	Fracture An'l.	Dimensioning
Mat'ls Tested OK	X	X	X	X
X	Not a Problem for Journal Box Lubr.	X	X	X
Other Roads use Case, ours uses Through	X	X	X	X
X	X	B-10 Life is 38 Mo., we repl. on 60 Mo.	X	X
X	High Sand levels found in J.B. Oil	X	X	X
X	X	X	Fracture Rpt. indicates High Stress on Axle	X
X	X	X	X	1% Chance of too small inter. fit (30 data)
X	X	X	X	Axles meas'd. no taper or out of round

8.2 Exercises

Exercise – Cause and Effect Diagram Development:

1. Review the data from the case study and determine what cause & effect diagram(s) would be appropriate for the case.
2. Brainstorm possible causes using post it notes. Place these on a flip chart as they are generated.
3. Use the affinity process to group them and create categories for a fishbone diagram.
4. Layout the fishbone diagram using the post it notes and dig deeper into one of the branches until the root cause level is reached.
5. Check the logic of the root cause level.
6. Be prepared to display your fishbone to the group.

Exercise – Cause and Effect Diagram Development

Here are some "simple" effects that you can use to practice developing cause and effect diagrams. Remember to focus on the *process* that produces these effects:

- Waiting time in a doctor's office averages 45 minutes.
- Teenagers don't take garbage out 48% of required times.
- "Maintenance" backlog for house projects averages 10 items/month.
- Stove clock blinks "12:00" 80% of the time.
- Spouse snores on 75% of nights.
- Lights and TV left on in vacant "rec" room 63% of time.
- Employees are late for work 22% of workdays.
- Operating Room overtime averages 18% of payroll each pay period.
- Delay time of 15 minutes (average) between end of meal and receipt of check at a restaurant.
- Required Surgical gowns not available for 10% of days.
- 15% of apartment rent checks were received more than 10 days after due date.
- In the last four trips, only one bass was caught at Moore Lake.
- Hiring process takes 4 months from job posting to job filling.

Exercise – Cause & Effect in the Sports World

An article in the American Society for Quality's *Quality Progress* journal focused on quality tools' application to football and, specifically to the *placekicking process.* For example, some of the factors that affect the accuracy of this process include *# steps to kick, approach speed, approach angle, foot placement, leg swing, leg velocity, ball variation, ball placement, distance from target, angle from target, snap location, held ball angle, fan noise, weather (wind, etc.), field conditions.*

Take one of your favorite sports, identify a critical-to-quality characteristic and then develop a cause-and-effect diagram that identifies the factors affecting this CTQ. How would you go about identifying which factors are most important?

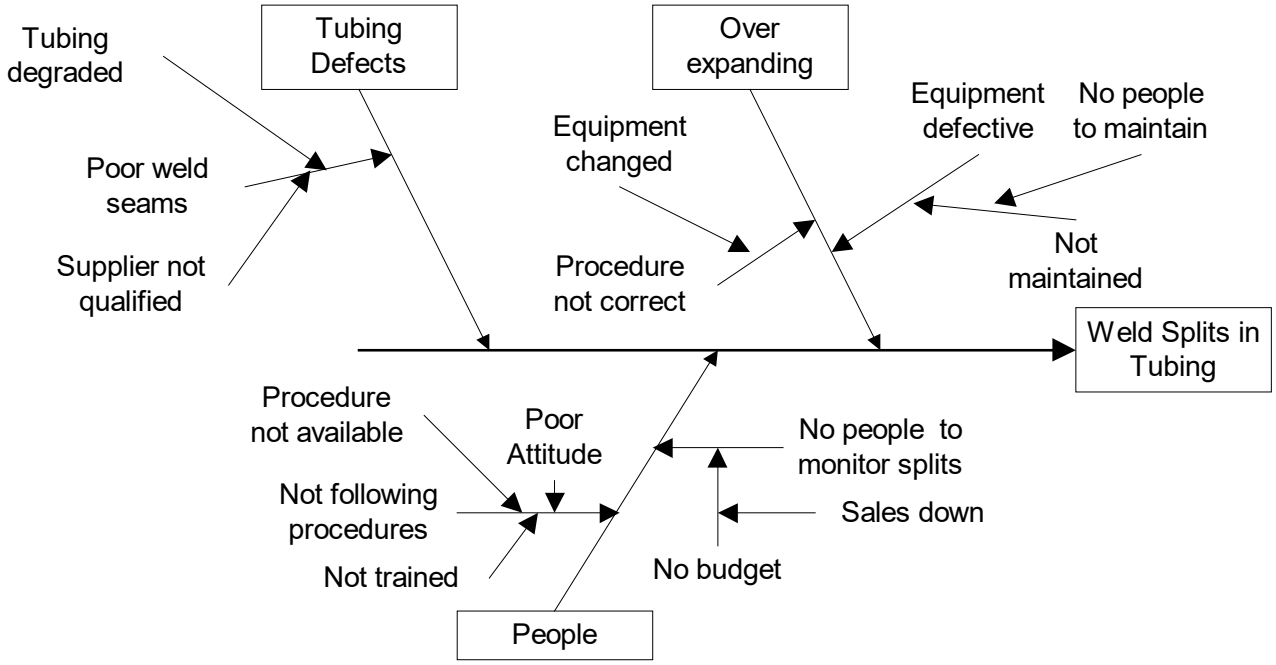
Exercise – With and Without Problem Cause and Effect Analysis

For these examples, develop cause and effect diagrams “with” and “without” the problem. What factors are different?

- Pleasant car buying experience/unpleasant car buying experience.
- Children late for school/children not late for school.
- Pleasant discussion of family finances/arguments during family finance discussion.
- Pleasant family vacation/”Clark Grizwald” family vacation (rent one of the “Vacation” tapes if you don’t know what we mean!).
- Successful installation of *TurboTax™* (or other program)/ unsuccessful or difficult installation of *TurboTax™*.
- Beneficial and Exciting Skills Training session/Boring, “Put you to Sleep” Skills Training session.

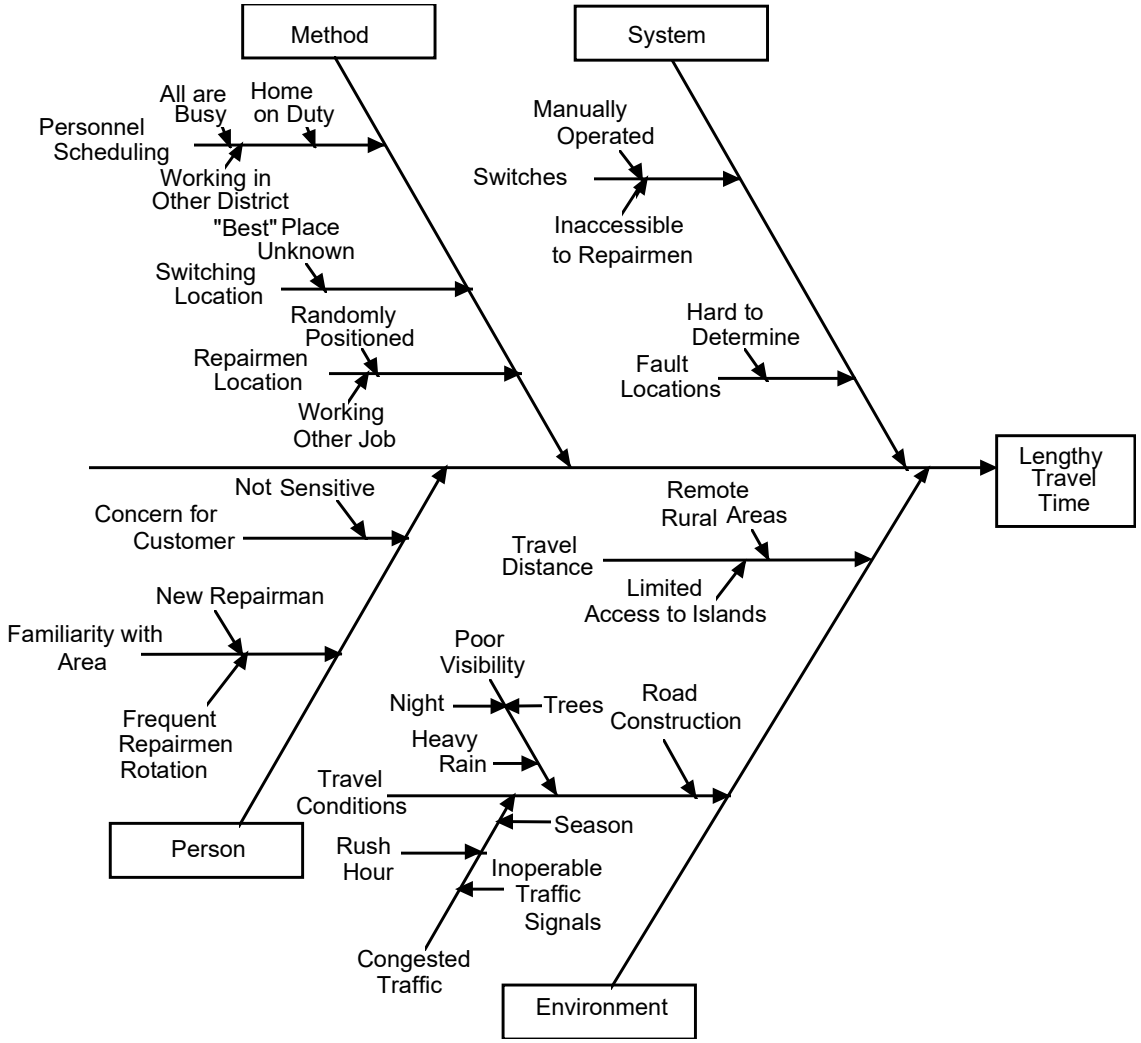
Exorcise – Cause and Effect Diagram Critique

Review the cause & effect diagram below. Comment on the effect and the potential causes.



Exercise – Travel Time Verification of Root Causes

A team of utility repairmen developed the following cause and effect diagram by focusing on lengthy travel time to restore customers' electricity. Consider how you would go about *verifying* which of these potential causes were most important:



Exercise – Verification by Pareto

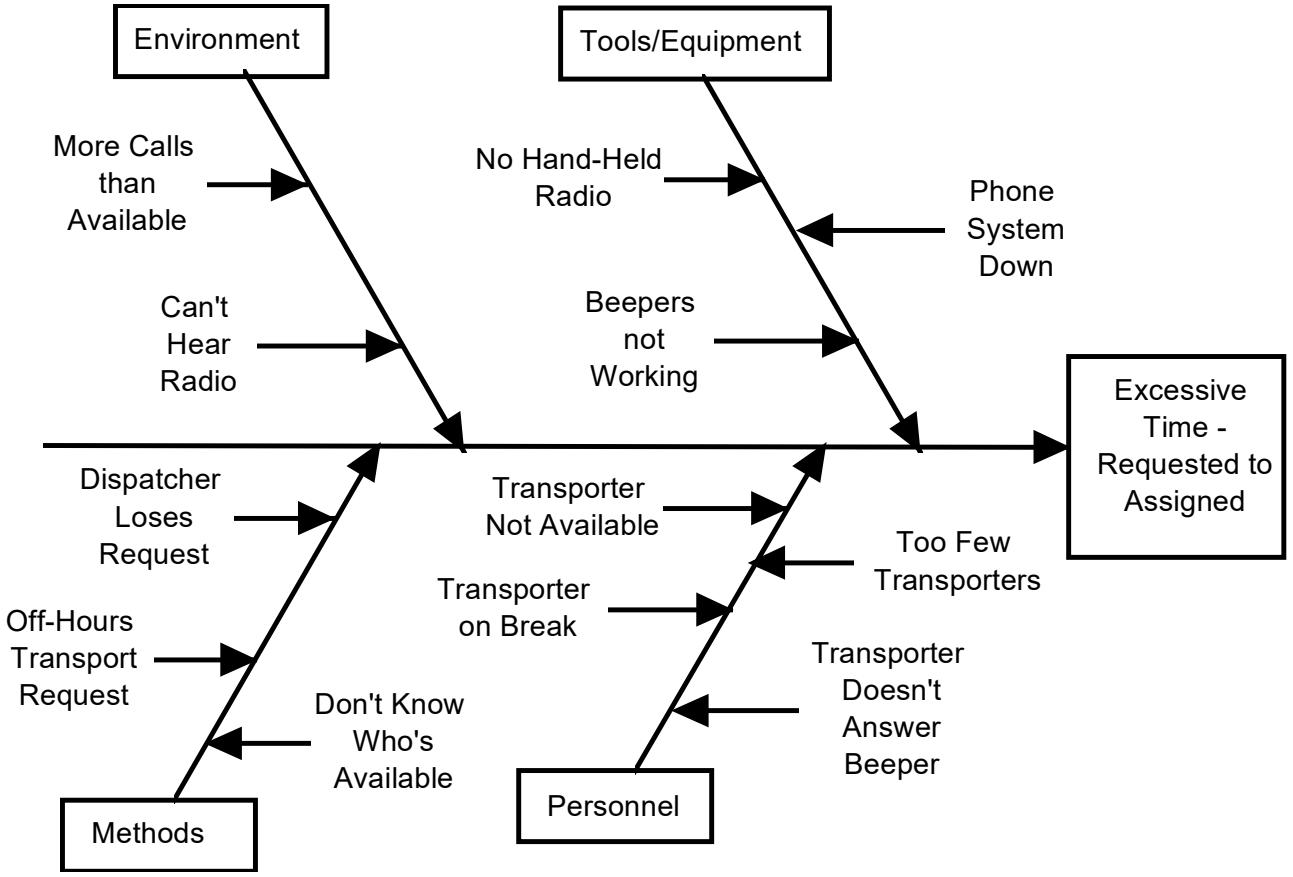
A team working on electrical splice failures analyzed 30 “pin and socket” splices that had failed in service. They developed the following Pareto Table and concluded that not meeting the clearance specification was the cause of the failures. They could only identify failed splices, as the remaining splices were buried underground. What could be wrong with their conclusion?

Pin and Socket Splice Failures
Pareto Table

Cause	Frequency	Cum. %
1/8” clearance specification not met	22	73.3
Unknown	3	83.3
Other	5	100.0

Exercise – Transport Time Verification of Root Causes

A hospital transport team was working on improving the *timeliness* of their response to requests to transport patients within the facility. Calls for transport came into a central dispatcher, who then assigned them to available transporters. The team had identified a time segment, *transport requested* to *transporter assigned,* as being the source of most of the delays. Consider how you would *verify* which of the potential causes are most important:



Project Assignment – Cause and Effect Development:

1. If you were able to stratify your data use it to determine which cause & effect diagram(s) to develop.
2. If you have sufficient knowledge of the causes, begin developing the cause & effect diagram.
3. Plan how you would deepen your understanding of the causes.

Project Assignment – Root Cause Verification:

1. Select several of the more probable root causes from your cause & effect diagram.
2. Decide whether the absence/presence or variable method of verification would be used.
3. Describe your verification method and the tools you would use to display/analyze the data.
4. Include the activities in your project plan.
5. Be ready to discuss and display your verification plan.

9.0 Detecting Differences

Unit	Description	Page
9.1	Foundations of Probability and Statistics	9.1-1
9.2	Hypothesis Testing	9.2-1
9.3	Sampling Theory	9.3-1
9.4	Exercises	9.4-1

In this section, we present a few methods that will help you detect *differences*. You may be interested in determining if there is a difference between the *means* of two processes (or a *before* and *after* situation), or you may wonder if a change will decrease the *variation* in your process. Since we live in a world of variation, we want to be as sure as possible that we are detecting actual differences, and not just variation inherent in the process.

9.1 Foundations of Probability & Statistics

Learning Objectives

- Understand Common Probability/Statistics Terms and Concepts
- Calculate Probabilities of Events
- Calculate Measures of Central Tendency and Variation
- Know and Interpret Common Probability Distributions
- Develop Point and Interval Estimates for Common Parameters

Unit Contents

- Probability Concepts and Methods
- Statistics

9.1.1 Introduction

Through our experience with “real-world” products and services, we’ve learned that we often must deal with the problem of *variation.* We’ve learned that variation in the *output* is a function of the variation in the *causal* factors of the *production system.* One of our respected teachers, Dr. Teiichi Ando, told us many years ago that we must *“move from the world of averages to the world of dispersion.”*

Much of process improvement involves building **models** that describe the behavior of the process’ variables – either output variables, input variables or relations between output and input variables. Variation drives us to use ***probabilistic* models** (rather than *deterministic*) as the best way of describing process behavior. This, in turn, drives us into the study of probability and statistics. Unit 9.2 presents tools and methods used to solve “advanced” variation-related problems. We present the background concepts and methods that support these tools here.

This section assumes the reader is starting with a minimal knowledge of "Prob & Stat." The material is divided into two major blocks, 9.1.2 - *Probability Concepts and Methods* and 9.1.3 - *Statistical Methods*. Even if the reader has had prior experience in these methods, it will be worthwhile reviewing this material. Many of the underlying concepts are just not those that are encountered in one's everyday experiences. It's easy to forget.

9.1.2 Probability Concepts and Methods

Terms and Concepts

We'll start by introducing some basic terms so we all can speak the same "probability language:"

Sets & Elements - A **set** is simply a collection of objects, events or numbers. The individual objects, events or numbers are the **elements** of the set. Examples include:

- The whole numbers from 1 to 10,
- The pills in a bottle of aspirin, and
- The outpatient visits during the month of March.

Experiments - When the term **experiment** is mentioned, you probably think of something a scientist does in a laboratory setting. We will broaden this concept to include the *output (or outcome) of any process*. *Measurement* of some characteristic(s) of the process' output/outcome is a part of the experiment. A certain type of experiment, a **random experiment**, is of particular interest in probability theory. The random experiment has the following characteristics:

- It can be repeated as many times as we care to, with the conditions of the experiment essentially unchanged,
- The particular output/outcome of any one experiment (especially the measured characteristic) cannot be predicted, although we can generally describe the set of the *possible outcomes* of the experiment,
- When the experiment is repeated many times, although the individual outcomes may appear to be haphazard, a pattern starts to emerge from our "looking at" the experiment's repetitions.

Examples of Random Experiments include:

1. The *time to failure* of a gas compressor,
2. The *power consumption* of an air conditioner,
3. The *number* that appears on the top face of a rolled die,
4. The number of *errors* on an expense account.

Sample Space - A set whose elements represent all possible outcomes of an experiment (or trial) in which the outcome depends on chance is called a **sample space.** For example, consider example # 3 - throwing the die. The six sides of the

die are embossed with one, two, three, four, five or six dots. Six outcomes are possible, corresponding to the die coming to rest with each of the six sides up. These six outcomes comprise the *sample space*. For the random experiments described above, the *sample spaces* are:

1. Time (months) > 0
2. Power (Kilowatts) > 0
3. Face Number = 1, 2, 3, 4, 5, or 6
4. Errors $\geq$ 0

Random Variable - When the experiment is run, some characteristic is measured, i.e. some *value* is obtained. A **random variable** is defined as a *quantity* that can be equal to any of the values in a sample space and is given the symbol x_i (in general, **X**). For the die example, the random variable x_i could be 1, 2, 3, 4, 5, or 6. Here, x_1 would represent the first outcome of throwing a die; x_2 would represent the second outcome etc. A random variable may be *discrete* (count) or *continuous* (measurement). The number on a die and the number of expense account errors are cases of *discrete* random variables. The gas compressor's life or the air conditioner's efficiency are cases of *continuous* random variables.

Events - In some cases, we may be more interested in the occurrence of an **event** rather than a specific element of a set. For example, we might be interested in a throw of a die resulting in a value < 3, or how many gas compressors last longer than 12 months.

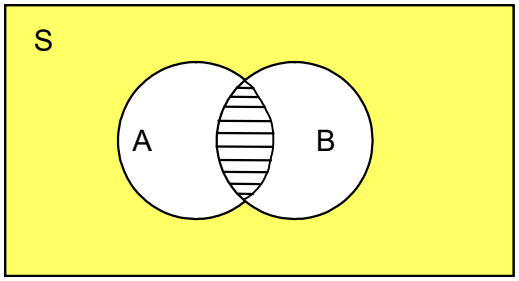
Each event is a subset of a sample space and would be assigned a collection of *elements*. Here, the random variable **Y** is used to represent the event. **Y** is then a function of several random variables $\mathbf{X_1, X_2, X_3, X_4}$, etc. that represent the collection of elements that compose the event.

This relationship helps us understand the quality characteristics of a *system*, which are expressed as a *function* of the system's factors or variables. We express this functional relationship as follows:

$$\mathbf{Y = f(X_1, X_2, X_3, X_4, \ldots X_n)}$$

Elementary Probability Operations

A Venn diagram can depict the relationship between events and their sample space. The sample space, *S,* is drawn as a *rectangle* and events are drawn by *circles* drawn inside the rectangle.



Venn Diagram

Using the die example, sample space **S** would be all the possible outcomes, 1 through 6. Let's consider the circle **"A"** to be the event of rolling a value less than **4** and circle **"B"** to be the event of rolling an **odd number**.

The **union** of events **A** and **B** is defined to be the event containing all the elements that belong to either or both events, that is, all values less than **4** and all **odd** numbers. This is shown graphically by the entire area inside the two circles and by the (*engineering)* notation:

$$A + B = \{1,2,3,5\}$$

The **intersection** of the events **A** and **B** is the event containing all elements that are common to *both* **A** and **B**. This would include all **odd** numbers less than **4**. The intersection is represented graphically by the shaded area and by the notation:

$$A \times B = \{1,3\}$$

Two events that cannot occur simultaneously are said to be **mutually exclusive**. These events have *no* intersection. Two or more events whose union is equal to the area of the sample space are said to be **collectively exhaustive**. For the sample space of possible outcomes when rolling a die, the two events 1) an even numbered outcome and 2) an odd numbered outcome are *mutually exclusive* and *collectively exhaustive*.

Two additional concepts you should understand are **independence** and **dependence**.

For events to be *independent*, the occurrence of one event cannot be affected by the fact that the other is either occurring or not occurring. If the event is flipping a coin, then successive flips of the same coin are *independent*; and the second flip's outcome is not affected by the first.

In many cases, outputs of the process can be considered as *independent*. For example, whether your printer works is most likely independent of whether the PC's monitor is functioning or not. Whether Mrs. Jones delivered twins is most likely independent of whether Mrs. Williams delivered twins or not. If the occurrence of one event causes another event to be more or less likely to occur, then they are said to be *dependent*. Two Black Belts left for the airport, one following the other. The probability that the second arrives on time is *dependent* on that of the first.

We've laid the groundwork for the definition of **probability**:

Probability is the likelihood of the occurrence of an event.

We will adopt the notation P(A) to represent the probability of event "A." Probabilities are "unit-less" and range from zero to one, or,

$$0 \leq P(A) \leq 1$$

The probability of the *sample space* event, **S**, is equal to one:

$$P(S) = 1$$

To calculate probabilities, we will adopt the notion of relative **frequency**. That is, if we do an experiment "**n**" times, and observe that event **"A"** occurs **"n_A"** times, the relative frequency is:

$$f_A = n_A/n$$

For now, we will claim that if we run this experiment "enough" times, f_A will converge in a *statistical* limit to the probability, P(A). There is an intuitive appeal to this definition.

For a simple experiment, such as tossing a coin, we expect that the relative frequency of obtaining the event *"heads"* will converge to the probability of obtaining a heads, or P(Heads). Here are the results from thirteen tosses of a quarter:

Coin Toss	1	2	3	4	5	6	7	8	9	10	11	12	13
Outcome	T	H	T	H	H	H	T	T	H	T	T	T	H
Relative Frequency	0.0	0.50	0.33	0.50	0.60	0.67	0.57	0.50	0.56	0.50	0.45	0.42	0.46

Notice that the relative frequency "wanders" around the "true value" of 0.5 (if the quarter is *fair*). For those of you familiar with the *arithmetic limit* of a function, this "wandering" doesn't occur. This is the difference between arithmetic and *statistical* limits. Try tossing a coin 100 times and plotting the relative frequency vs. coin tosses. We'll bet that the relative frequency continues to "wander" around the value of 0.5.

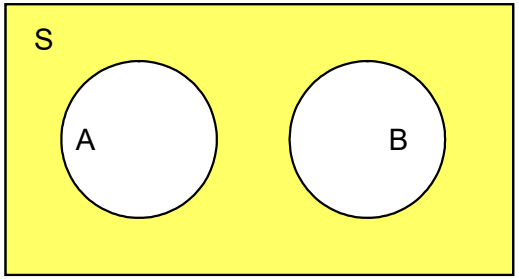
There is another school of thought that is based on a different concept of probability. The *Bayesian* approach considers probability as the *degree of belief* we have in the occurrence of some event. Bayesian proponents will "allow" a discussion of the probability that a certain candidate will be elected, since they are expressing their degree of belief that this event will occur. "Classical" probability proponents will not admit that a probability can be assigned to this event, since it is not an experiment that can be repeated. Bayes' Theorem will be discussed below.

Joint & Conditional Probability

Having introduced the concept of an individual *event's* probability, we now turn to the challenge of how to *combine* the probabilities of multiple events. Again, the motivation for this is mainly in our desire to predict a *system's* quality characteristics as a function of the *variables/factors'* characteristics. The *exclusivity* and *dependency* relationships were introduced above because they affect how we *combine* the individual probabilities. If two events are *mutually exclusive*, the probability of their *union* is:

$$P(A+B) = P(A) + P(B)$$

A simple example of mutually exclusive events is the "on or off" status of a piece of equipment (i.e. it can't be *both* on *and* off). When we judge the outcome of any process by a standard or specification, we make use of the mutually exclusive principle. An axle either meets its diameter specifications, or it doesn't. The correct medication is ordered, or it is not. The Venn diagram for mutually exclusive events is:

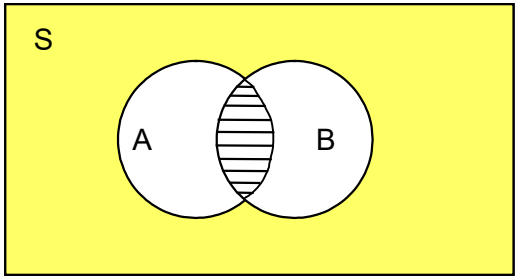


Venn Diagram – Mutually Exclusive Events

If the two events can occur at the same time (i.e. they are *not* mutually exclusive), then we determine the probability of their union by:

$$P(A + B) = P(A) + P(B) - P(A\ x\ B)$$

The last term is subtracted because it is common to both events and we don't want to double count this "piece" of probability. The Venn diagram below shows two *non*-mutually exclusive or *intersecting* events:



Venn Diagram – Intersecting Events

If one event, **A**, is dependent upon a second event, **B**, then we define the **conditional probability** of event **A** given event **B** as **P(A|B)**. The probability of event **A** occurring given event **B** is:

$$P(A/B) = \frac{P(A \times B)}{P(B)}$$

If **A** and **B** are *independent* events (i.e. *A* does **not** depend on *B*) then the probability of two independent events both occurring is the product of the probabilities of each of the events:

$$P(A\ x\ B) = P(A)\ x\ P(B)$$

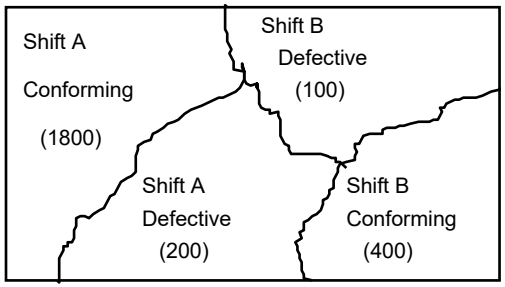
and:

$$P(A/B) = \frac{P(A \times B)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A)$$

Example: A team is investigating brazing impellor blades performed by two shifts. Suppose that within a six-month period Shift A has fabricated 2000 impellors of which 200 were defective while Shift B has fabricated 500 impellors with 100 of these being defective. We can summarize this information in the following table:

Shift	# Defective	# Conforming	Totals
A	200	1800	2000
B	100	400	500
Totals	300	2200	2500

The Venn diagram shows us that all the elements of this example are mutually exclusive:



Venn Diagram - Compressor Brazing

The overall probability of an impellor being defective is:

$$P(\text{defective}) = \frac{300}{2500} = 0.12$$

The individual *Shift* defective probabilities can also be calculated:

$$P(defective|Shift\ A) = 0.1$$

$$P(defective|Shift\ B) = 0.2$$

Let's pose this question: If an impellor is picked from the combined lot at random and found to be defective, what is the probability that it was fabricated by Shift B? Here, the conditional probability equation helps us answer this question:

$$P(B / defective) = \frac{P(B \times defective)}{P(defective)} = \frac{100/2500}{300/2500} = 0.33$$

Similarly, the probability that the impellor was fabricated by Shift A *given that it is defective* is 0.67. Let's ask a little harder question: If two impellors are picked from the combined lot, what is the probability that *both* are defective? First, let's rearrange the conditional probability equation:

$$P(A\ x\ B) = P(A|B)\ x\ P(B)$$

If we let **B** be the *first impellor* and **A** be the *second defective impellor*, then:

$$P(B) = \frac{300}{2500} = 0.1200\ and\ P(A|B) = \frac{299}{2499} = 0.1196$$

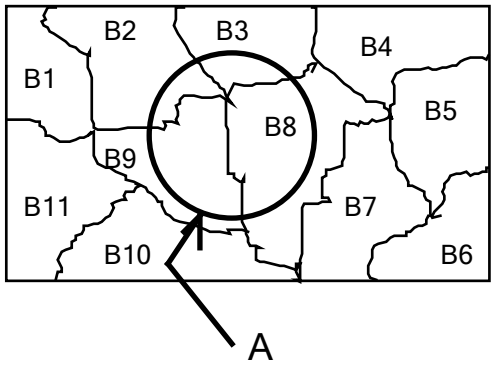
$$Therefore, P(AxB) = 0.1196\ x\ 0.1200 = 0.0144$$

Notice how the *P(A|B)* is calculated. Since the event being described is that the *second impellor is defective* ***given*** *the first is also defective*, we need to subtract one defective from the numerator and one fabrication (i.e. the defective) from the denominator. This reflects our "knowledge" of event **B**, the first impellor being defective.

In this impellor example, all of the elements of the sample space were mutually exclusive. The impellor could only be in one of four possible states. We can generalize this situation to the case where many different elements can exist and our problem is to find the probability of some event that is a subset of these elements. This leads us to the impressive sounding **Law of Total Probability:**

$$P(A) = P(A|B_1)\ x\ P(B_1) + \ P(A|B_2)\ x\ P(B_2) + \ldots P(A|B_n)\ x\ P(B_n)$$

The event **"A"** is the one we are interested in, the **"B_i"** are mutually exclusive elements of some sample space, **"S."** The Venn diagram shows the picture of this situation:



This relationship is very useful. There are many quality engineering situations where we are interested in the probability of **"A"** occurring (i.e. if **"A"** is a system failure), but have no way to compute this probability directly, or through combining the events through the "normal" union and intersection relationships. The law of total probability helps us through the additional knowledge that some event **"B_i"** has occurred.

Bayes' Theorem

The conditional probability law has the property of being *symmetric* with respect to the events **"A"** and **"B."** That is:

$$P(A/B) = \frac{P(A \times B)}{P(B)} \Rightarrow P(A \times B) = P(A/B)P(B) = P(B/A)P(A)$$

Let's take the last two equations and rearrange them slightly:

$$P(B_i/A) = \frac{P(A/B_i)P(B)}{P(A)}$$

If we substitute the *law of total probability* in the denominator of this expression, we obtain the discrete form of Bayes' Theorem:

$$P(B_i/A) = \frac{P(A/B_i)P(B_i)}{\sum_{i=1}^{n} P(A/B_i)P(B_i)}$$

This equation (and the concept behind it) has been put to many uses in the quality field, but it has also been the subject of a great deal of controversy. Although the algebra was straightforward to obtain this equation, the interpretation of the equation is not easy.

Here is one way of looking at this equation. Think of event **"A"** as the effect (i.e., the head of a fishbone diagram) and the **"B_i"** as a set of causes. If we run a series of experiments, **"A"** may or may not occur. When **"A"** does occur, further assume that we cannot observe which of the **"B_i"** resulted in the effect.

Given that **"A"** does occur, Bayes' Theorem allows us to compute the probability that each of the set of possible causes operated to cause **"A."** In effect, Bayes' Theorem allows us to reason retrospectively from effects to causes. This is one issue that has resulted in the controversy surrounding Bayes' Theorem.

Bayesian "partisans" call the $P(B_i)$ the **prior** probabilities, since they usually represent estimates of the probabilities before any objective data has been gathered. The $P(B_i/A)$ are called **posterior** probabilities, since they represent estimates after data has been observed.

Why bother with a complicated and controversial Theorem, like Bayes'? Let's discuss a typical application of Bayes' Theorem in quality management.

A system is being tested to determine its reliability characteristics. The system is "new," but engineers know that it is mostly composed of components with proven track records. There are some new components in their system, and the configuration of the components is new. Should our engineers assume that they have no knowledge about the reliability of this new system, or can they make use of the performance data they have for the proven components? Bayesians will claim that the engineers should take the latter course.

The Bayesian approach is to consider the prior knowledge and formulate **prior** distributions of the reliability parameters, such as percent defective or failure rate. Then, the system is tested, operating times and numbers of failures and successes are recorded.

With this data, the analyst then updates the prior distributions using Bayes' Theorem, to obtain a **posterior** distribution that incorporates both the prior (subjective) knowledge and the objective test data. This process may be repeated, continually modifying the reliability estimates until sufficient confidence is obtained in the value of the percent defective or failure rate. This becomes the basis for a decision to terminate testing and to accept or reject the "new" system's reliability performance.

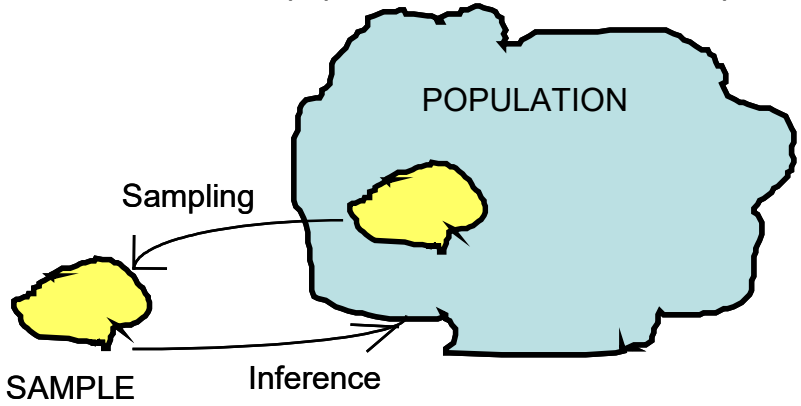
Because the engineers do not start with a state of (assumed) complete ignorance, the Bayesian approach tends to converge to a decision point more rapidly than would a traditional or "classical" approach to the problem. If the cost of testing is high, or if there is only a limited test time allowed before a decision is to be reached, Bayesian analysis can be a more cost-effective approach to the decision-making process.

Difficulties with Bayesian methods include such questions as: *How is subjective (prior) knowledge to be distinguished from bias and ignorance? What is the best way to quantify the prior knowledge and how much weight should it be given relative to the test data?*

9.1.3 Statistics

While *Probability* attempts to predict the future performance of systems, *Statistics* looks back and tries to understand what has already happened. Frequently, the two disciplines are blended, since statistics supply data that allow us to estimate the probability of future events. The **Basic Statistical Method** involves:

- taking a *sample* of data from a population,
- conducting analyses (i.e. calculating statistics) with the sample data and then,
- making inferences about the population from which the sample data was taken.



Populations may be *infinite* (or practically so) such as the population of the United States, or the number of molecules in a liter of saline solution. Populations may be finite, such as the number of J-Compressors produced in one year, or the number of staff in a customer service center. Practically speaking, when we try to *fit* a probabilistic model to data from the field or experiments, we perform the following process:

1. Understand where the central tendency of the data lies,
2. Determine how the data varies or is dispersed,
3. Develop a "picture" of the data (i.e. control chart, histogram, probability chart),
4. Fit the best probability model or distribution to the data, and
5. Develop uncertainty estimates for the population parameters.

Although there are other analyses we could do, these steps are among the most useful in turning data into *information.* Our discussion of statistics in this manual will focus on how we can develop and interpret *probabilistic models* for our product/service's characteristics. Let's explore the five steps of the modeling process:

Measures of Central Tendency

Understanding the **central tendency** of the data is the first step toward a probabilistic model. As its name implies, this is some point around which the data lies or can be grouped. By now, you are familiar with three measures of central tendency: the Mean, the Median, and the Mode.

Mean	The sum of the data set values, divided by the number of values in the data set. The mean is denoted by $\overline{x}$ for a sample. $$\overline{x} = \frac{x_1 + x_2 + x_3 + ... + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$
Median	The value that divides the data set in half. The median is calculated by: $$\tilde{x} = x_{(n+1)/2} \qquad \textit{if } n \textit{ is odd}$$ $$\tilde{x} = \frac{x_{n/2} + x_{(n/2)+1}}{2} \qquad \textit{if } n \textit{ is even}$$
Mode	The most frequently occurring value in the data set. If we are looking at a histogram of data or a continuous function, the highest bar or the peak of the distribution is the mode.

Any or all these measures of central tendency can be used to indicate the center of a data set. However, one may be more appropriate than another depending upon the data set. The mean is the most frequently used measure of central tendency and is often shown when constructing a histogram or frequency chart. However, the median may be a better indication of the center of the data set when there are extreme values present. You may recall seeing incomes or house prices reported in the newspaper as median values. The mode is best used, of course, with apple pie.

Measures of Dispersion

The next useful information is the **dispersion** or variability of the data. For example, the average of both sets of numbers {40, 60} and {49, 51} is 50, but there is a big difference in the dispersion of the two sets. There are four measures of dispersion we'll employ - the Range, Sum of Squares, Variance and Standard Deviation.

Range	The difference between the highest and lowest (greatest and smallest) values: $R = x_{max} - x_{min}$
Sum of Squares	The sum of the squared differences between the individual values and the mean of the data set. SS denotes the Sum of Squares. $SS = \sum_{i=1}^{n} (x_i - \overline{x})^2$
Sample Variance	The sum of squares divided by one less than the number of values in the data set. The variance is denoted by s^2. $s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n-1}$
Sample Standard Deviation	The square root of the variance. The standard deviation is denoted by s. $s = \sqrt{s^2}$

The smaller the standard deviation or variance, the "tighter" the values are located about the mean. To see why measures of dispersion are important consider the following example:

Two hospitals state the average length of stay (LOS) for a particular diagnosis is seven days. However, upon gathering data representing a sample of 20 patients from each hospital you find that the average LOS of hospital A is 7.1 days with a standard deviation of 1.2 days while the LOS at hospital B is 6.9 days with a standard deviation of 3.0.

Although the LOS at hospital B is slightly less than hospital A's, the dispersion is smaller at "A." It's worthwhile investigating *why* this is so.

Chebychev's and Camp-Meidell's Theorems

Once we know the mean and standard deviation of some population, we can begin to estimate probabilities. Two important theorems come into play here, *Chebychev's* and *Camp-Meidell's*. Both theorems place *upper bounds* on the probability that a random variable will assume a value *outside* a given number of standard deviations away from the mean. These bounds provide us with a "*quick 'n dirty*" way of estimating probability and they are also useful when the *functional* form of the probability distribution is not known (as described below).

Chebychev's Theorem - For any data distribution, the upper bound on the probability of a random variable assuming a value outside of the mean plus or minus *k* standard deviations (with *k > 1*) is:

$$P(|X| > \mu + k\sigma) < \frac{1}{k^2}$$

For example, if *k = 2*, then the probability of finding an $\mathbf{X_i}$ greater than two standard deviations away from the mean is less than 1/4 or 25%. For *k = 3*, the upper bound on the probability is 1/9 or about 11%.

Camp-Meidell's Theorem - This relationship is like Chebychev's, except that we have some knowledge about the data's distribution. If we have evidence that the distribution is *uni-modal* (that is, there is only one mode) and that the distribution decreases *monotonically* (i.e. the slope of the curve is always negative) on either side of the mode, then we can improve Chebychev's relationship. Here, the upper bound on the probability of a random variable assuming a value outside of the mean plus or minus *k* standard deviations (with *k > 1*) is:

$$P(|X| > \mu + k\sigma) < \frac{4}{9k^2} = \frac{1}{2.25k^2}$$

For example, if *k = 2*, then the probability of finding an $\mathbf{X_i}$ greater than two standard deviations away from the mean is less than 1/(2.25 x 4) or about 11%. For *k = 3*, the upper bound on the probability is 1/(2.25 x 9) or about 5%.

Probability Distributions - General

The third element of our model is the probability distribution. These distributions are mathematical functions that help us describe the behavior of our quality characteristics or process variables. There are two general classes of probability distributions. The first are used to model *discrete* data, the second are used for *continuous* data.

Probability Mass/Density Functions

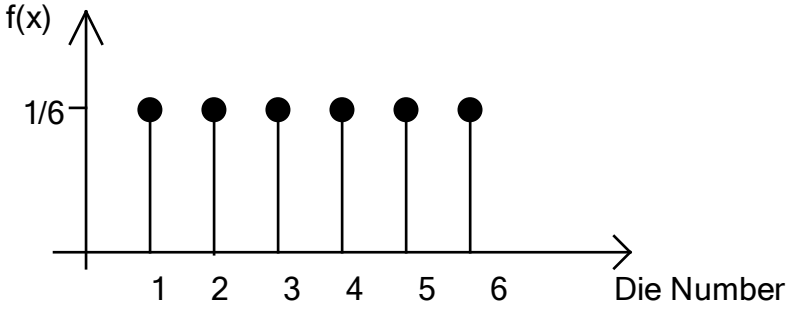
Frequently, our studies will lead us to question the probability that a specific number of events will occur. For example, how many medication errors do we predict will occur this year? To help us answer this question, we will introduce the notion of a **probability mass/density function**. In the Probability section, we defined a *random variable* as a quantity that can be equal to any of the values in the sample space. There are functions that associate probabilities with the values of a particular random variable. When we are dealing with discrete data, these functions are called *probability mass functions (pmf)*. In other words, for every possible value x_i of the random variable, the *pmf* specifies the probability of that value occurring. The function f(x) is a probability mass function for a discrete random variable if:

$$f(x) \geq 0 \quad \text{for all } x_i$$

$$\sum_{i=1}^{\infty} f(x_i) = 1$$

$$P(X = x_i) = f(x_i)$$

$P(X = x_i)$ is read "the probability that the random variable X assumes the value x_i." For our die example, the probability mass function would look like this:

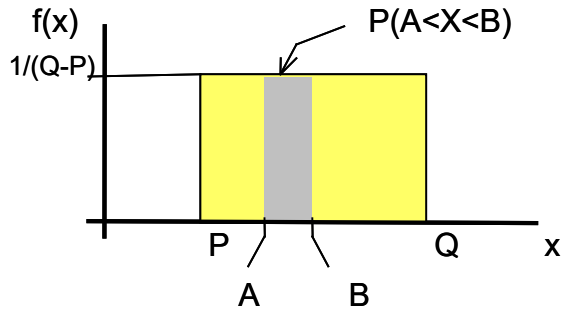


Similarly, for continuous data the *probability density function (pdf)* of x is a function *f(x)* where for any two numbers a and b:

$$for\ a \le b,\ \ P(a \le X \le b) = \int_a^b f(x)dx$$

That is, the probability the random variable X takes on a value between *A* and *B* is the area under the graph of the density function *f(x)* between *A* and *B*. For continuous data, we must specify an area under the curve because the probability of a particular value of X occurring is zero. This is illustrated using the *uniform* distribution function shown below:

Uniform Distribution -
Probability Density Function



Also note that *f(x)* is always greater than or equal to zero and that the total area under the curve is equal to one or

$$\int_{-\infty}^{\infty} f(x)dx = 1 \quad for\ -\infty \le x \le \infty$$

Don't get nervous about these summation and integral equations. Practically, to work with the distributions we will present in the next few sections, we'll either use look-up tables to obtain the necessary values (see Appendix A) or rely on a statistical software package.

Cumulative Distribution Functions

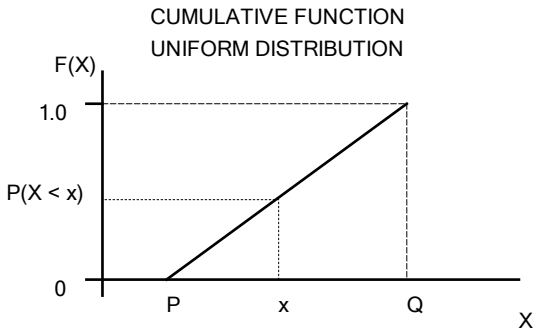
From the *pmf* or *pdf*, the **cumulative distribution function (cdf)** can be determined for both discrete and continuous data. The cumulative distribution function is simply the probability that a random variable X is less than or equal to a particular value x. For discrete data, this is given by:

$$F(X) = P(X \le x_i) = \sum_{t \le x} f(t)$$

and for continuous data by:

$$F(X) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$

Notice that this integral is evaluated from "-∞" to "x." This assumes that the random variable (i.e. the “x's") can range across the entire number line (i.e. the sample space is infinite). For many practical problems, the characteristic will “only” range from “0” to “+∞“ (i.e. times, lengths, volumes, costs) and the integral will be evaluated from “0” to “x.” The main point is that we must consider the sample space of the characteristic to determine the appropriate integral limits. The cumulative distribution function for the uniform distribution appears below:



Expected Values

We may have obtained some historical data and fit that data into a particular probability mass function or probability density function. For prediction purposes, we are often concerned about the **expected value** of the random variable. That is, we

may want to know how many events to "expect" during a given time period or we may want to know the "expected" time a given procedure may take. The expected value, E(X), of the random variable X is determined by the following:

$$E(X) = \sum_{all\ x} xf(x) \quad if\ X\ is\ discrete$$

$$E(X) = \int_{-\infty}^{\infty} xf(x)dx \quad if\ X\ is\ continuous$$

For example, let's consider our die once again. What is the expected value of our die throw?

$$E(X) = 1\ x\ 1/6 + 2\ x\ 1/6 + 3\ x\ 1/6 + 4\ x\ 1/6 + 5\ x\ 1/6 = 21/6 = 3.5$$

The expected value of the thrown die is the *average* or *mean* value. For some probability distributions, the expected value will be one of the parameters that describes the shape and location of the probability distribution. The normal distribution's expected value *is* the mean (μ), for instance. For other distributions, the expected value will be a function of the distribution's parameters.

Since the mean (and variance) are calculated from our sample data, we will show you the relationship between these two values and the distribution parameters for the specific distributions discussed below.

Discrete Probability Distributions

There are many different *discrete* (AKA *count or attribute*) probability distributions. We will present two that have been found most useful in quality management, the *binomial* and the *Poisson*.

Binomial Distribution

For many analysis, we may be conducting experiments with only two possible outcomes (e.g. Go/No-Go, success/failure). Experiments of this type are called **binomial** experiments, and they possess the following properties:

- The experiment consists of "n" repeated trials.
- Each trial results in an outcome that may be classified as a success or failure (or a yes/no).
- The probability of success, denoted by "p," remains constant from trial to trial.
- The repeated trials are independent.

The distribution used to describe the outcomes of such an experiment is called the **Binomial Distribution**. Binomial processes include:

- Tossing a coin (success = "heads"),
- Testing incoming parts from a vendor against a standard (each part classified as *pass* or *fail*),
- Component failure during the warranty period.

The probability mass function for the binomial distribution is:

$$f(x:n,p)=\binom{n}{x}p^x q^{n-x}$$

where:

$f(x:n,p)$ - the function f(x) given values n and p

n - number of repeated trials

p - probability of success

q = 1 - p - probability of failure

$$\binom{n}{x}=\frac{n!}{x!(n-x)!}$$ - number of combinations of n objects taken x at a time

n! = n(n - 1)(n - 2)…(3)(2)(1) - n factorial

Combination Example: For lottery fans - How many combinations of 49 items can be made taking 6 at a time?

$$\binom{49}{6} = \frac{49!}{6!(49-6)!} = \frac{49!}{6!(43!)} = \frac{49\times48\times47\times46\times45\times44\times43!}{6\times5\times4\times3\times2\times1(43!)}$$

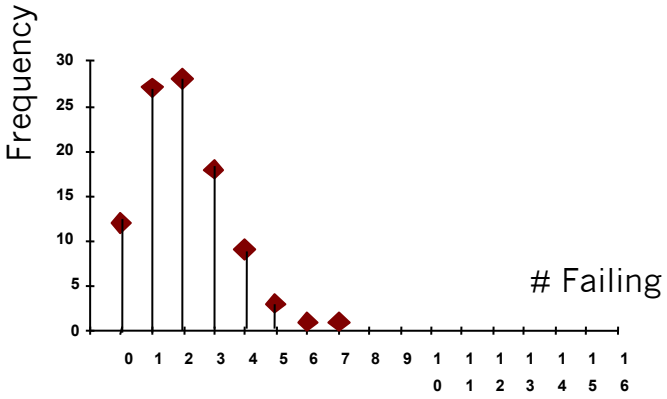
$$\binom{49}{6} = \frac{10{,}068{,}347{,}520}{720} = 13{,}983{,}816$$

This is where the odds statement comes from: If you hold a lottery ticket (and the lottery is a "fair game"), there is about a one in fourteen million chance that your ticket will match the winning combination.

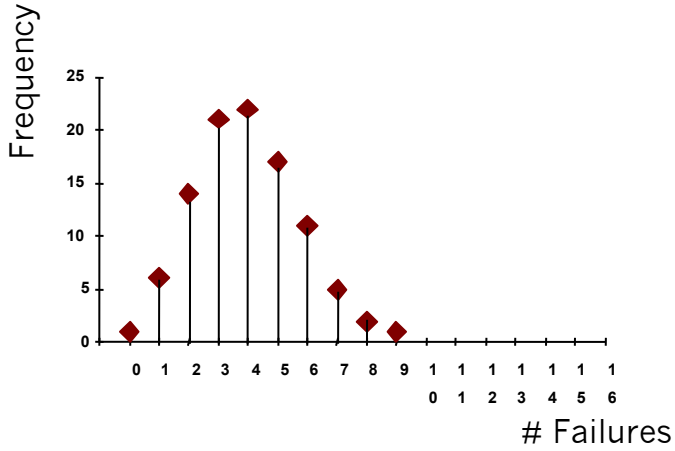
The mean of the binomial distribution is *np* and the variance is *np(1 - p)* or *npq*. The binomial cumulative distribution function is:

$$F(X;n,p) = \sum_{x=0}^{x=X}\binom{n}{x}p^x(1-p)^{n-x}$$

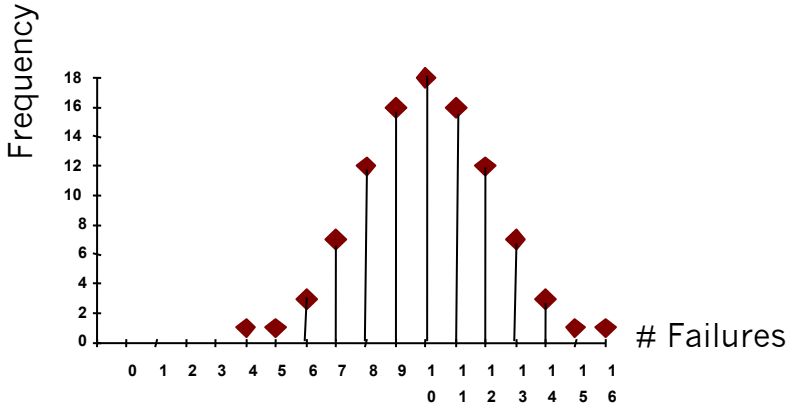
The binomial parameters, *n* and *p* affect the shape of the distribution. Let's say that 20 medical records (*n*) are inspected each week, and we know from experience that 10% (*p = 0.1*) fail to meet specifications. Although we might reasonably expect 2 out of each 20 to fail the inspection (20 x 0.10) on average, we also know that sometimes there might be 1 or 3 or 0 or 4 failures out of a particular sample of 20. If we made a frequency chart of how often we found 0, 1, 2, 3, 4, etc. in 100 samples of 20 it might look like this:



If we hold the sample size *n* constant and change the value of *p*, the shape of the frequency distribution will change, as the mean and standard deviation change. For example, if we change *p* to 0.20, then the chart might look like this:



And if *p* = 0.50 the chart will most likely look like this:



If *p* is changed to 0.90, the frequency chart looks like the "mirror image" of the chart when p=0.10.

Poisson Distribution

The **Poisson Distribution** is named for a 19th century French mathematician who derived it by studying records of the number of soldiers in the Prussian army who were kicked by horses each year. The Poisson is used in situations where the opportunities for an event of interest to occur are unlimited, but the probability of the event in any brief interval of time (or physical location) is very low. These conditions apply in situations where the events of interest are infrequent failures of devices or systems that are used continuously and are therefore subject to the risk of failure at any time. Examples of Poisson processes include:

- The number of paint scratches on a chiller panel.
- The number of employee injuries per month.
- The number of leaks in a tube header.
- The number of needlesticks in a hospital per month

The probability mass function of a Poisson variable for a given unit of time (or location) is:

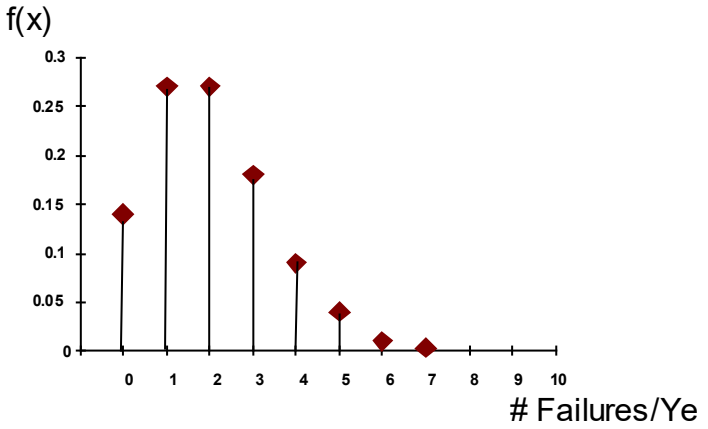
$$f(x{:}\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

for x = 0,1,2,3,.... and λ >0

The distribution depends on the single parameter λ (lambda). λ is the expected number of events and is both the mean and variance of the Poisson distribution. The cumulative Poisson distribution is

$$F(X{:}\lambda) = \sum_{x=0}^{x=X} \frac{\lambda^x e^{-\lambda}}{x!}$$

The shape of the Poisson distribution, like the binomial distribution, changes as its parameter changes. The shape of the Poisson distribution is nearly symmetric when λ is greater than 5.0. The frequency chart shown on the right is a Poisson distribution with λ =2.



The Poisson and binomial distributions are "related." As the sample size increases (*n* gets larger) and the individual event probability decreases (*p* gets smaller), the binomial distribution begins to look like the Poisson. The product *np* "approaches" the Poisson parameter λ. Consider the number of needlesticks in a hospital each month. This is really a binomial process. Each time a needle is handled, a "stick" either occurs or it does not. But the probability of a stick is very small with each handling (*p* is small) and the number of needles handled each month (*n*) is large. We can approximate this binomial process with the Poisson distribution. Of course, if the number of needles used each month varies widely from month to month (perhaps due to wide census fluctuations), this approximation may not be valid.

Continuous Distributions

There are many different distributions for *continuous* (AKA *measurement, variables)* random variables. Five "popular" continuous distributions are presented here: the Uniform, Normal, Lognormal, Exponential, and Weibull.

Uniform Distribution

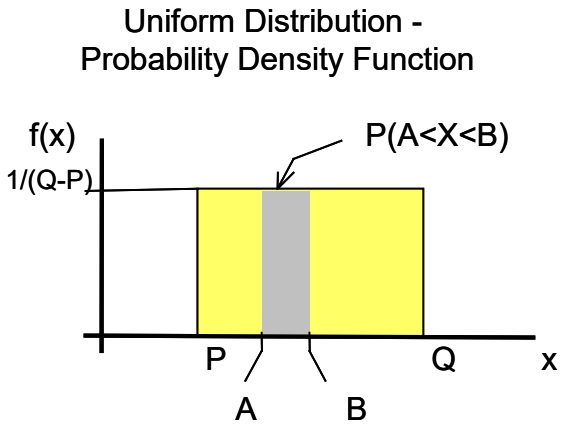
The **Uniform Distribution** is perhaps the simplest of all continuous distributions. It fits the situations where there is an equal probability of any value in a range from *p* to *q* occurring. The probability density function is:

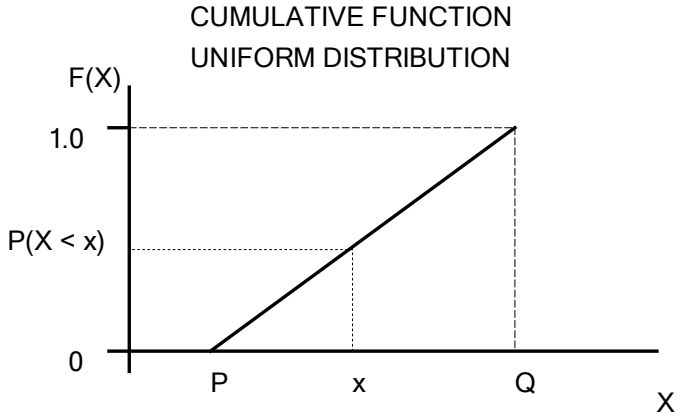
$$f(x) = 1/(q-p)$$
$$\text{where}: p \le x \le q$$

The mean and standard deviation for the uniform are:

$$\mu = (q-p)/2$$
$$\sigma = (q-p)/\sqrt{12}$$

The probability density and cumulative distributions are:





Normal Distribution

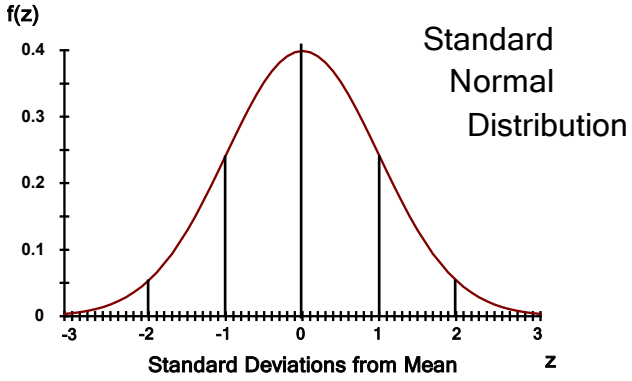
The **Normal Distribution** is used to model many processes which occur in nature. It arises from the actions of many small errors whose impacts are *additive* producing a total *error*. For example, many manufacturing processes fit this model. If five parts are assembled into a product, the total variation in the product's length is a function of the parts' variation. If these individual variations are *additive*, then the overall product's variation could be modeled by the normal distribution.

The normal can also be used to model the distribution of averages calculated from repeated samples from a constant process. This holds for very small samples sizes (as small as 4 or 5) regardless of the distribution from which the samples are drawn. This latter property is very important, since we often wish to bound the average value of a quality characteristic. The probability density function of the normal distribution is:

$$f(x:\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(1/2)\frac{(x-\mu)^2}{\sigma^2}}$$

where $-\infty \le x \le \infty$

Here μ and σ are the population mean and standard deviation, respectively. The probability density function of the normal distribution has the familiar "bell-shaped curve." If we let $z = (x - \mu)/\sigma$, the standardized normal distribution is obtained. The mean of this distribution is 0, and the variance is 1.0. The standard normal distribution appears below:



There are several distinctive traits of the normal distribution:

- It is symmetrical about its mean.
- The curve goes from concave to convex (i.e. the second derivative of the function is zero) at the mean plus and minus one standard deviation.
- About 68% of all points within a normal distribution fall within plus and minus 1 standard deviation of the mean.
- About 95% of all points are within plus or minus 2 standard deviations of the mean.
- About 99.7% of all points are within plus or minus 3 standard deviations of the mean.

The cumulative distribution function of the normal distribution is:

$$F(X:\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{X} e^{-(1/2)\frac{(x-\mu)^2}{\sigma^2}}\,dx$$

This integral doesn't have a closed form solution, Appendix A, Table A.1 tabulates the standard normal cumulative distribution function.

Example: A sample of bolts was tested for tensile strength. The test results demonstrated an average tensile strength of 62,000 lb. with standard deviation of 5,000 lb. If the specification requires a minimum strength of 50,000 lb., what percent defective could we expect from this manufacturer?

$$F(50{,}000: 62{,}000, 5{,}000) = F\left(\frac{50{,}000 - 62{,}000}{5000} : 0, 1\right) = F) - 2.4, 0, 1) = 0.0082$$

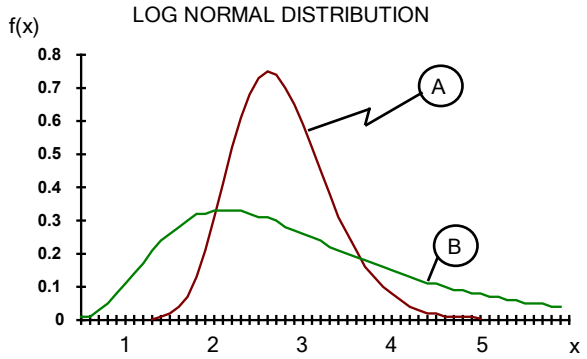
Therefore, based on the sample results, 0.82% of these bolts could be expected to fail the specification.

Log Normal Distribution

The **log normal distribution** is that of a random variable whose *natural logarithm* is normally distributed with parameters μ and σ. This distribution arises from the effect of many small errors, the effects of which are *multiplicative.* Some failure mechanisms, such as metal fatigue have been modeled successfully with a log normal distribution. Times to repair equipment have also been modeled using the log-normal. The probability density function for this distribution is:

$$f(x{:}\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2}(\frac{\ln x-\mu}{\sigma})^2}$$

Because only positive numbers have real-valued logarithms, the log normal distribution occupies only the positive portion of the real number line. It is not symmetrical and has a long right tail:



Curve A - Parameters: μ = 1.0, σ = 0.2
Curve B - Parameters: μ = 1.0, σ = 0.5

The parameters of this distribution are unit-less, unlike the normal distribution. The log normal Mean and Variance are:

$$E(x) = e^{\mu+(\sigma^2/2)}$$

and

$$Var(x) = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1)$$

A sample of data is usually "fitted" to a log normal distribution by plotting the data on special lognormal probability plotting paper. The parameters μ and σ are then determined graphically. Alternatively, many statistical software packages will fit a set of data to the log normal distribution. The cumulative log normal distribution is given by:

$$F(X{:}\mu,\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\int_0^X e^{-\frac{1}{2}(\frac{\ln x-\mu}{\sigma})^2}\frac{dx}{x}$$

Again, this integral is difficult to evaluate. Separate tables for the log normal distribution are not usually tabulated, since we can transform the log normally distributed random variable **X** into one that is normally distributed. The normal distribution's cumulative tables can then be consulted for the appropriate values. If we designate the cumulative distribution function for a lognormal random variable as $F_L(X, \mu, \sigma)$ and the same function for a normal random variable as $F_N(X, \mu, \sigma)$, then the following holds:

$$F_L(X, \mu, \sigma) = F_N(lnX, \mu, \sigma) = F_N\left(\frac{lnX - \mu}{\sigma}, 0, 1\right)$$

where the rightmost expression is the random variable *X* transformed into a standard normal deviate, *Z*. Table lookups are available for *Z (see Appendix A)*.

Example: What is the probability that a piece of equipment will fail in 10.0 months of operation if it's failure distribution is log normal with μ = 2.6 and σ = 0.13?

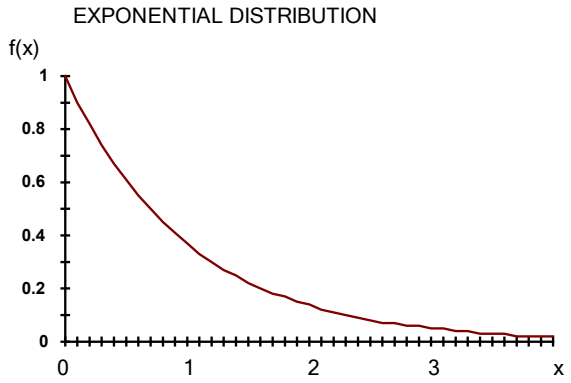
$$F_L(10.0, 2.6, 0.13) = F_N(\ln(10.0), 2.6, 0.13) = F_N\left(\frac{2.30 - 2.60}{0.13}, 0, 1\right)$$
$$F_N(-2.31, 0, 1) = 0.01 \; or, about \; 1 \; percent$$

Exponential Distribution

The **Exponential Distribution** is an important distribution in many areas of quality. Among other phenomena it describes the duration of telephone calls and the decay of radioactive materials. The exponential distribution is used extensively (and sometimes, inappropriately) in modeling the time to failure of components in reliability analysis. It is the distribution of times between successive Poisson events. Later, we will see that it is also a special case of the Weibull distribution. The density function of the exponential distribution is

$$f(t) = \lambda e^{-\lambda t} \qquad t \geq 0, \lambda > 0$$
$$= 0 \qquad \text{elsewhere}$$

The reciprocal of its single parameter, $1/\lambda$, is both its mean and standard deviation. The graph of the exponential distribution for λ = 1 would appear as follows:



An interesting property of the exponential distribution relates to the fraction surviving. After one mean, or one *1/λ*, only thirty-seven percent of the original population survives. After two means, or *2/λ*, only thirty-seven percent of the fraction that survived one mean is left (0.37 x 0.37 = 0.137). The exponential distribution models a process with a *constant failure or decay rate.*

A sample of radioactive material follows this behavior. Likewise, systems that are composed of many different failure modes can be modeled using the exponential distribution.

Many applications of the exponential distribution are based on its "memoryless property". The property applies to situations in which the history of past events does not influence the probability of occurrence of future events. For example, a system whose times to failure are exponentially distributed has no “memory” of its history. After operating successfully for 1000 hours, it is as reliable for the next hour of operation as it was when first placed in service.

The cumulative distribution function of the exponential distribution is:

$$F(t{:}\lambda) = 1 - e^{-\lambda t}$$

Example: LED screen failures were determined to fit an exponential model, with an average time to failure of 27,000 hours. What is the probability a screen will last one year without failure?

First, λ = 1/ 27,000 hours = 3.7 E-5 hr.$^{-1}$

$$F(8760\ hours{:}\ 3.7\ x\ E-5\ hr.^{-1} = 1-\exp(-8760\ hr.x\ 3.7\ x\ E-5\ hr.^{-1}) = 1-0.72 = 0.28$$

But this is the probability of failure. 0.72 (1 - 0.28) is the probability the screen will operate for one year without failure.

Weibull Distribution

The **Weibull Distribution** is one of the most widely used distributions to model survival behavior. The exponential and beta (not presented here) distributions are special cases of the Weibull distribution. The Weibull can even approximate the normal distribution. About the only continuous distribution that is useful in quality work and is not approximated by Weibull is the lognormal. The probability density function of the Weibull Distribution is

$$f(t{:}\beta,\eta,t_o) = \frac{\beta}{\eta}\left(\frac{t-t_o}{\eta}\right)^{\beta-1}\exp-\left(\frac{t-t_o}{\eta}\right)^{\beta}$$

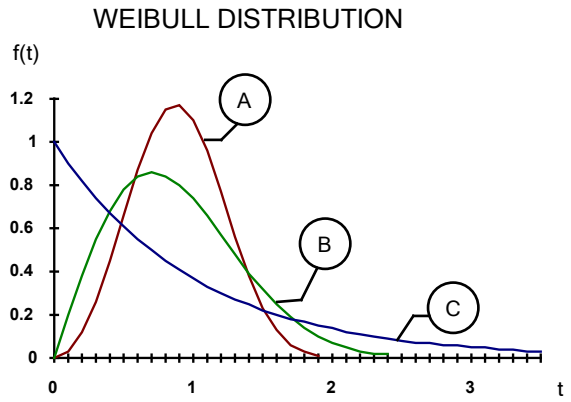
The Weibull has three parameters (β, η, and t_0). Each parameter has a physical interpretation (described in Section 15). The shape of the Weibull distribution changes as the parameters change. The mean and variance of the Weibull distribution are shown below as a function of **η, β,** and t_0:

$$E(t) = t_0 + \eta\Gamma\left(1+1/\beta\right)$$

$$Var(t) = \eta^2\left[\Gamma\left(1+2/\ \ \beta\right) - \Gamma^2\left(1+1/\beta\right)\right]$$

where $\Gamma(x)$ is the familiar "gamma" function.

Several curves are presented below showing varying **β's**. For these curves, η is constant and equal to 1.0 and t_0 is equal to 0.



A - β = 3, B - β = 2, C - β = 1

The cumulative density function of the Weibull Distribution is:

$$F(t:\eta,\beta,t_o) = 1 - e^{-[(t-t_0)/\eta]^{\beta}}$$

If you set t - t_0 = η in the cumulative density function, the cumulative probability of failure is now:

$$F(t:\eta,\beta,t_o) = 1 - e^{-[(\eta)/\eta]^{\beta}} = 1 - e^{-1^{\beta}} = 1 - e^{-1} = 0.632$$

So, regardless of the value of beta, eta (by definition) is the point at which 63.2% of the items have failed.

Practically, the Weibull distribution parameters are obtained through plotting the sample data on Weibull probability paper (see Section 15) or through statistical software packages.

Sampling Distributions & Parameter Interval Estimation

The fourth and last issue we will discuss is the concept of *uncertainty.* When we take a sample of data from a population, we use this information to develop a probabilistic model that best describes the population. Practically, this means that we must select one of the distributions (Weibull, for example), plug the data into this model and calculate or otherwise obtain estimates of the distribution's parameters.

Example: We have obtained the following life data for home air conditioners (time to unit disposal):

20, 21, 30, 15, 18, 24, 19, 20, 17, 17 years

We suspect that this data is normally distributed and we calculate the mean and standard deviation of this data.

$$Sample\ Mean = \bar{x} = \frac{20+21+30+15+18+24+19+20+17+17}{10}$$

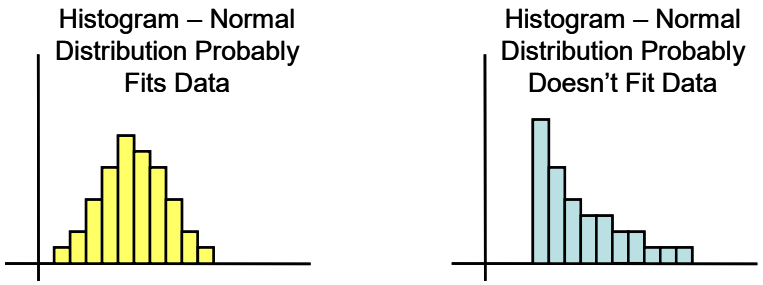
$$\bar{x} = 201/10 = 20.1\ years$$

$$Sample\ Standard\ Deviation = s = \sqrt{\frac{(20-20.1)^2+(21-20.1)^2+\ldots+(17-20.1)^2}{10-1}}$$

$$s = 4.28\ years$$

Notice that we are using the symbols for *parameter* ***estimates***. Based on these estimates, we could ask questions such as "How many (or what fraction) of our air conditioners will be replaced before 15 years?, How many will be replaced between 20 and 23 years?, etc." Before we answer these questions, though, there are two things we need to think about. First, how well does this data *fit* the normal distribution model we have chosen?

This is a two-part question. First, we must establish that the data arose from *some* constant (or common cause system). Control charts are the best tools to answer this question. Second, we need to determine whether the normal is a good model of the constant cause system. There are various statistical tests (such as the Andersen-Darling and Kolmolgorov-Smirnov procedures) that allow us to examine the "goodness of fit" of our chosen distribution. If enough data is available, a histogram may allow us to obtain a graphic estimate of goodness of fit.



If we use *probability paper*[1] to fit a distribution to the data, the graph gives us some confidence in our model. If the data do *not* fit the distribution line, then the fit is questionable. Additionally, a correlation coefficient can be calculated to measure the fit of the data to the distribution line.

The second issue revolves around the estimates of the distribution's parameters. What would happen if we took another sample of air conditioners? Our calculations would likely reveal mean and standard deviation estimates that are different from the first set - there is *variation* in our estimates of the population mean, as obtained from *samples*.

By adopting a *probabilistic model* for our quality characteristics, we have admitted that there is variation in the values of these characteristics. We are now going to take this probabilistic notion one step further. The *sample statistics* themselves are *random variables* and can be modeled with probability distributions.

Here are two questions to illustrate the issue - the first is associated with variation in a *population,* the second with variation in a *population parameter:*

A Population Question: What is the probability that any given air conditioner will survive 15 years, if the population of air conditioners can be modeled with a normal distribution with mean 23 years and standard deviation 6 years? To answer this question, we simply calculate the standard normal deviate, K_α, and look up the answer in the normal table.

A Population *Parameter* Question: What is the probability that the *population mean* is less than 20 years, if a sample of 10 air conditioners has a sample mean of 23 years, with a sample standard deviation of 3 years? To answer this question, we'll need to develop a model of the population mean's distribution. This distribution is not the same as the population's distribution.

[1] Probability paper is available for a wide variety of distributions. See Section 15 for an example where *Weibull Probability Paper* is applied.

We'll now present three distributions that are used to model the variation in our population *parameters*:

Normal Distribution for Sampling Means

Let's say we can obtain samples from a population where the *mean* of the population is not known, but the *variance* of the population is known. For example, suppose we want to change an existing process that has a predictable mean and variance. The change is expected to affect only the *mean* of the process, not the *variance*. We are then interested in what the new mean of the changed process will be.

Under these circumstances, the distribution of the sample means is **normal**, with the mean of this distribution equal to the new population mean and variance equal to the population variance divided by the size of our sample:

$$\mu_{sample\ means} = \mu_{population}$$

$$\sigma^2_{sample\ means} = \frac{\sigma^2_{population}}{n}, where\ n = sample\ size$$

If we take a sample from the new process, we can now use the cumulative normal distribution function to ask questions about the probability of the population parameter being greater or less than some particular value (i.e. the *old* population mean).

Example: Let's try to answer the question we raised about the air conditioner population's average life using this approach. We wanted to know the probability that the *average* life of the population was less than 20 years, given that our sample of 10 patients gave us an average life of 23 years with standard deviation of 3 years. Using the cumulative normal distribution and transforming our statistics into the standard normal form:

$$P(\overline{X} \le 20\ years{:}23{,}3/\sqrt{10}) = P(\frac{20-23}{3/\sqrt{10}}{:}0{,}1) = P(-3.16{:}0{,}1)$$

From Standard Normal Deviate Tables:

$$P(-3.16{:}0{,}1) = 0.00079$$

This answer tells us there is a very small probability that the average life is less than 20 years. If this sample of air conditioners included reliability improvements, the sample provides evidence of their effectiveness.

The "t" Distribution

We can use the normal distribution when the sample variance is known. If this is not the case, then we must turn to the "t" (or Student-t) distribution. The probability density function of the "t" distribution is:

$$f(t:\nu) = \frac{\Gamma\left[(\nu+1)/2\right]}{\sqrt{\pi\nu}\,\Gamma(\nu/2)}\left(1+t^2/\nu\right)^{-(\nu+1)/2}$$

The mean of the "t" distribution is *0* and its variance is $\nu/(\nu - 2)$. The "t" distribution's shape is like the normal distribution, except that the "t" function tails off less rapidly than the normal. ν is referred to as the number of "degrees of freedom." For a random variable of size "n" from a normal population with mean μ and variance σ^2, we can create the statistic:

$$F(\frac{\overline{x}-\mu}{s/\sqrt{n}}:n-1)$$

This is the cumulative "t" distribution with *n-1* degrees of freedom. Table 1.2 provides values of the cumulative "t" distribution at the end of this section. Let's see how the "t" distribution works:

Example: A medical monitoring device has been found to have an average life of 10,000 hours. An improvement was made and the manufacturer tested 16 "new" devices to failure. The average time to failure was reported to be 10,800 hours with a standard deviation of 615 hours. Can the manufacturer claim to have improved the average device life?

$$F(\frac{\overline{x}-\mu}{s/\sqrt{n}},n-1) = F(\frac{10{,}800-10{,}000}{615/\sqrt{16}},16-1) = F(5.2,15)$$

Referring to Appendix A, Table A.2, the Student's "t" distribution with 15 degrees of freedom, we find that the probability of getting a value even as large as 2.602 is only 0.01, or about one in a hundred.

The probability of getting a value as large as 5.2 is very small, so we can conclude that the manufacturer did improve the reliability of the device. (An operational question: Although the improvement is "statistically significant," how much more would you be willing to pay for this improvement in device lifetime?)

The Chi-square Distribution

The next sampling distribution we will present is the χ^2 (chi-square) distribution. The cumulative distribution for the chi-square is:

$$F(X,\nu) = \frac{1}{\Gamma(\nu/2)} \int_0^X (x/2)^{\nu/2-1} e^{-(x/2)} dx$$

The mean of the chi-square is ν, the variance is 2ν. The cumulative distribution is provided in Appendix A. The parameter ν is referred to as the number of degrees of freedom.

The chi-square distribution helps us understand the distribution of the sample standard distribution. The following statistic has a chi-square distribution with n-1 degrees of freedom:

$$(n-1)\,{s^2}/{\sigma^2}$$

We will use this shorthand for the cumulative distribution function of the chi-square:

$$F((n-1)\,{s^2}/{\sigma^2}, n-1)$$

Example: A critical O-ring must have a uniform thickness. The standard is set based on a maximum thickness standard deviation of 0.0075 inches. A random sample of 10 O-rings is taken from a lot and the standard deviation calculated to be 0.0036." Based on the sample, can we reject the possibility that these O-rings come from a population with a standard deviation of 0.0075 inches? We want to be 99% sure our decision is correct.

Calculating the chi-square statistic, $(n-1)\,{s^2}/{\sigma^2}$, we obtain a value of:

$$(10-1)\,{(0.0036)^2}/{(0.0075)^2} = 2.074$$

Comparing this value to Appendix A, Table A.3, the χ^2 critical-value for 9 degrees of freedom and α (our "sureness," or confidence) of 0.99 is 2.09. Since the chi-square statistic is less than the critical value, our O-rings do not come from a 0.0075-inch population.

Another application of the chi-square distribution is to model the parameter λ, which describes the exponential distribution. We may obtain a sample of failure data and wish to develop a conservative estimate of a component or system failure rate. The chi-square distribution handles this problem (See Section 15).

The F-Distribution

A distribution related to the chi-square distribution is the F-distribution (named for Sir R. A. Fisher, famous English statistician). The F-distribution is formed from the ratio of two chi-square distributions. Suppose that **A** and **B** are random variables whose distributions are chi-square, with ν_A and ν_B degrees of freedom. The random variable **F** has the F-distribution, with ν_A and ν_B degrees of freedom:

$$F = \frac{A / \nu_A}{B / \nu_B}$$

The F-distribution is used in comparing variances and is also used in Analysis of Variance (ANOVA) procedures. Appendix A, Table A.4 tabulates the cumulative F-distribution.

Point Estimates

Before we jump into *interval* estimates, let's review what we mean by a *point* estimate. The point estimate is simply our best shot at the value of the population parameter of interest, based on data we have taken in our sample. For instance, the sample mean, *x-bar*, is our point estimate of the population mean, μ. Similarly, the sample standard variance, s^2, is our point estimate of the population variance, σ^2.

There are four properties that statisticians are interested in when it comes to point estimates, *unbiasedness, consistency, efficiency,* and *sufficiency.* The probability distribution population estimators presented above are the ones that give the "best" estimates based on meeting these criteria.

The main point is this: If you use the formulae presented above to estimate some population parameter from a sample of data, you are on safe ground. If you use some other method of estimating the mean, for instance, by taking the average of the data's range, then you are on shaky ground statistically.

Interval Estimates

Although there are several different types of interval estimates, we will focus on one, the **Confidence Interval.** A confidence interval is an uncertainty interval on the value of a population *parameter.* There are three parts to a confidence interval: the point estimate, the interval and a statement of confidence.

Example: The average life of a water pump is 5000 hours, plus or minus 100 hours at a 95% level of confidence.

How do we translate this statement? The 5000 hours is our *point* estimate of the *average* pump life and the plus or minus 100 hours is the *range* in which we believe the population mean could lie (remember we are working from a sample of data and making an inference about the population of pumps). 4900 hours to 5100 hours is then the interval in which we believe the average life to lie. This interval is also called the *precision* or *error* of the estimate.

Notice that many product advertisements and engineering statements will only include one or at most two of these elements of a confidence interval. These are insufficient without the third element, the statement of the *confidence level.*

The classical statistician's way of interpreting the confidence level statement is as follows: If we drew many samples from the population of bearings and calculated the interval, *about 95 percent of the intervals would actually contain the population mean life*. On the contrary, about 5 percent of the intervals would not contain the population mean life. The higher the confidence level we set, the larger the interval will be and the more certain we will be that the interval contains the mean or whatever population parameter we are trying to estimate.

Bayesian statisticians interpret the confidence interval considering their definition of probability: There *is about a 95% chance that the calculated interval will include the population parameter of interest.*

How can we calculate a confidence interval for a population parameter? There are four inputs needed: the *point estimate* of the population parameter, the *variance* (or *standard deviation*) *estimate* of the population, the *number* of data in our sample, and the *confidence level* we desire (notice that we set the confidence level).

The basic procedure makes use of the sampling distribution associated with a particular population parameter. Recall that the distribution of the sampling mean, *x-bar* (with known standard deviation, σ), is normal, with mean, μ, and standard deviation, σ/n. We can create the interval:

$$-K_{\alpha/2} \le \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \le K_{\alpha/2}$$

or, equivalently,

$$\bar{x} - K_{\alpha/2}\sigma / \sqrt{n} \le \mu \le \bar{x} + K_{\alpha/2}\sigma / \sqrt{n}$$

Here, the quantity $K_{\alpha/2}$ is the standard normal deviate evaluated from minus infinity to $\alpha/2$. α is equal to one minus the confidence level we wish to set for this decision. Here are two more confidence intervals. The first is the interval constructed for the mean when we must estimate the population standard deviation with the sample standard deviation, *s*. Recalling that the t-distribution is the sampling distribution in this case:

$$-t_{\alpha/2,n-1} \le \frac{\bar{x} - \mu}{s / \sqrt{n}} \le t_{\alpha/2,n-1}$$

or, equivalently,

$$\bar{x} - t_{\alpha/2,n-1}s / \sqrt{n} \le \mu \le \bar{x} + t_{\alpha/2,n-1}s / \sqrt{n}$$

Below is a confidence interval that we can construct for the standard deviation, making use of the chi-square distribution:

$$\chi^2_{1-\alpha/2,n-1} \le \frac{(n-1)s^2}{\sigma^2} \le \chi^2_{\alpha/2,n-1}$$

or, equivalently,

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$$

We can construct a confidence interval for the *proportion* (i.e. fraction) of items possessing some characteristic. This interval assumes that the binomial distribution can be approximated by the normal distribution. To be “safe” in making this assumption, the products *np* and *n(1 - p)* should both be greater than 5:

$$-K_{\alpha/2} \le \frac{p - \mathbf{P}}{\sqrt{p(1-p)/n}} \le K_{\alpha/2}$$

or, equivalently,

$$p - K_{\alpha/2}\sqrt{p(1-p)/n} \le \mathbf{P} \le p + K_{\alpha/2}\sqrt{p(1-p)/n}$$

Finally, we can construct a confidence interval for a *rate* (i.e. a Poisson process). As above, this interval assumes the Poisson distribution can be approximated by the normal distribution, and therefore, we'll try to keep $\lambda > 5$:

$$-K_{\alpha/2} \le \frac{\hat{\lambda} - \lambda}{\sqrt{\hat{\lambda}/n}} \le K_{\alpha/2}$$

or, equivalently,

$$\hat{\lambda} - K_{\alpha/2}\sqrt{\hat{\lambda}/n} \le \lambda \le \hat{\lambda} + K_{\alpha/2}\sqrt{\hat{\lambda}/n}$$

Special Note: The distribution tables in Appendix A tabulate 1 minus the cumulative distribution, since these correspond to the α's used in hypothesis testing and most other statistical procedures. ***All*** *formulae that reference these tables have been set up so that you can directly look up the appropriate value to substitute in the equation.*

The "Rosetta Stone" for Probability and Statistics

You've probably noticed that there are a few equations and formulas used in probability and statistical work. We'll use the following popular notation to express the mathematics in this manual:

Quantity Type	General Notation	Examples
Population Parameter	Greek letter, or Bold, Capitalized English letter	μ - Population Mean σ - Population Standard Deviation **P** - Population Proportion λ- Population Rate of Occurrence
Statistic (Estimator of Population Parameter)	Small, English Letter, or "Hatted" Greek letter	$\bar{x}$ - Sample Mean s - Sample Standard Deviation p - Sample Proportion $\hat{\lambda}$ - Sample Rate of Occurrence

Summary of "Prob and Stat" Discussions

This review of probability and statistics basics is intended to support the remaining Sections of this manual. The *statistical* (or "*sadistical*") methods presented here are used to take our sample data and develop *probabilistic* models of characteristics of interest. These models will help us understand the dispersion of the product or service. If we decide to change the product/service's design, we will have a baseline to see the effects of our change.

We have discussed the various models that describe different reliability characteristics, the discrete distributions such as binomial, Poisson and the continuous distributions such as normal, exponential, lognormal and Weibull.

We have also explored the issue of uncertainty (at least the statistical kind). This helps us understand that the estimate from the sample data is just that, an estimate. We cannot know the *true* value of the quality characteristic, but we can bound the estimate and make a statement about where we think the true value lies.

Modern quality management incorporates **variation** as a central pillar of its theory. Ignoring the variation inherent in any quality characteristic can lead us to make wrong and costly decisions.

9.2 Hypothesis Testing

Learning Objectives

- To understand and apply Hypothesis Testing to detect differences

Unit Contents

- Hypothesis Testing Concept
- Hypothesis Testing Process
- Differences in the Center
- Differences in the Variation
- Differences in Proportions and Rates
- Power of the Test and Other Mysteries
- Non-Parametric Tests

9.2.1 Introduction

We'll introduce this section via two examples.

1. *Engineering Change Notice Process* - An Engineering Department (ED) sends change notices to the air handler production group. The current delivery process is manual, after the change notice is created; the ED clerk walks the documents over to the shift supervisor. The average delivery time is known to be 8 minutes, with a standard deviation of 2 minutes.

The department then decides to send the change notices via email. After this change is implemented, the process performance was again measured. The new average is 1 minute, with a standard deviation of 0.5 minutes.

Has the change resulted in an improvement? Here, we don't need any "powerful" statistical methods to help us decide. The large difference in results makes the improvement "obvious to the casual observer!" To show these results, either a control chart of the data or comparative histograms are sufficient.

2. *Oil Sample Processing* – A diagnostic laboratory has been trying to improve the time it takes to process the oil samples received from clients. The current performance averages 60 minutes, with a standard deviation of 3 minutes. A change was implemented and 50 oil samples were measured for turnaround times. The sample's average is 55 minutes, with a standard deviation of 2.8 minutes. Is there enough evidence to suggest that an improvement has been made?

Here, the answer is not as clear. If we just consider the average times, we might be tempted to conclude that five minutes have been shaved from the process. However, the laboratory director is concerned that the results could just be due to "random" variation in the process or in the sample of processing times collected. How can we address this concern?

This section's main topic is that of *hypothesis testing.* This statistical method provides us with the means of answering these types of questions and stating the "confidence" we have in the results.

9.2.2 Hypothesis Testing Concept

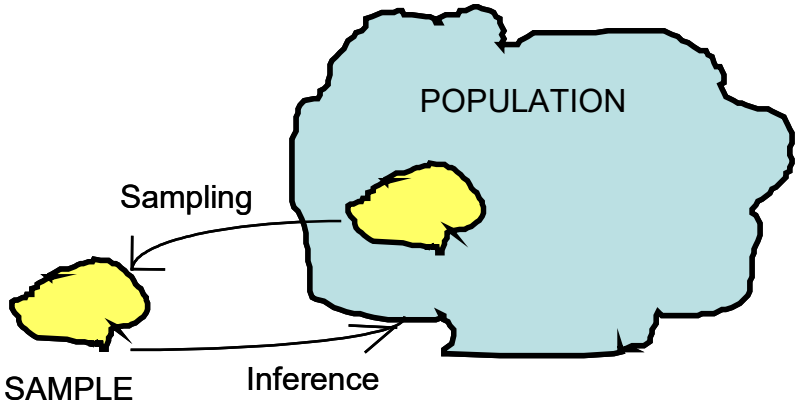
The basic question answered by hypothesis testing is actually very simple. We use control charts to help us answer a very similar question for production processes:[1]

> *Given that there is variation in the population, how can we distinguish between random variation and variation due to "special" or significant factors/differences?*

In Unit 9.1, we described two concepts that are key to understanding hypothesis testing:

- **The Basic Statistical Method**
- **Sampling Distributions**

The Basic Statistical Method provides us with a way of understanding the characteristics of a *population.* We *sample* from the population, measure the characteristic(s) of interest, and calculate *statistics* (means, standard deviations, and proportions) to summarize the data.



We then make *inferences* about the population(s) based on the statistics. For example, based on a *sample* of copper tubing, we make inferences regarding the *population* of tubing. Suppose we take 20 different tube samples (e.g. repeated

[1]As we present the concept and methods of hypothesis testing, we'll compare and contrast the control chart approach to this method of detecting differences.

samples of 50 tubes) from a given shipment. For each sample, we record the number of tubes that meet minimum wall thickness requirements and calculate the associated proportion. We may get results such as the following:

Sample Number	Proportion
1	86%
2	88%
3	90%
.	.
20	84%

Even though we are drawing the samples from the same population of tubes, the samples do not produce the same proportion, i.e. there is sampling variation (just as there is variation from subgroup to subgroup on the control chart of a common cause system). The *sampling distribution* provides us with a way of describing the "expected" variation in the statistic, given that the population parameter is some constant value.

Now let's say we are looking for a difference. For example, we measure the proportion of tubes meeting minimum wall thickness requirements this week (through a sample). Over the next month, a team from the tube manufacturer analyzes the production process and makes several changes to the process. Next month, we take a similar sample and attempt to detect a difference in the proportion of tubes that now meet the minimum wall thickness requirements (we set up a *hypothesis* that there is a difference).

If the new sample produces results that fall within the "expected" variation of the statistic, then we will conclude that the changes had no effect on the tubes' wall thickness. If the new sample produces results that fall outside the "expected" variation, then we may conclude that the changes *did* affect the thickness.

To transition from concept to practice, we see that hypothesis testing follows a standard *process*. To employ this process, we will need to understand the "expected" variation in the statistic. The *sampling distribution* will help us here. Each statistic has its own sampling distribution (i.e. the mean is *normally distributed* when the standard deviation is "known.").

9.2.3 Hypothesis Testing Process

A few cautions are in order before the hypothesis testing process is described. First, if the population we're investigating is not stable, any inferences we make from our samples are questionable. For example, if we define the population to be the output of a process for some time, we should examine the stability of the process first with a control chart. If the process exhibits special causes of variability, then hypothesis tests will not be valid.

When we perform a hypothesis test, the process will "force" us to make assumptions about the distribution of either the underlying data or the sampling distribution of the statistic. In many cases, these assumptions are easily justified (i.e. the sampling distribution of the mean will be *normal*, even if the sample size is only 4 or 5). In some cases, though, these assumptions may require additional study or analysis. For example, if two or more binomial populations are mixed, the mixture is *not* binomial. The point is that you should always check your assumptions!

There are six steps in the hypothesis test process:

1. **Establish the Hypothesis**
   a) Null Hypothesis (H_o)
   b) Alternative Hypothesis (H_a)

2. **Choose a Significance Level (α - "alpha")**

3. **Plan the Test:**
   a) Choose the Test Statistic (formula)
   b) Determine the Rejection Region.

4. **Collect Data and Calculate the Test Statistic**

5. **Draw a Conclusion**

6. **Estimate the Parameter of Interest**

Let's explore these steps on the following pages.

1. Establish the Hypothesis
a) Null Hypothesis (H_o)
b) Alternative Hypothesis (H_a)

We first establish the hypothesis, which consists of two components, the *null* and *alternative* hypotheses. Here are a few examples:

Laboratory Analysis

- *Null Hypothesis (H_o)* - The average time to complete an oil sample analysis is 60 minutes.
- *Alternative Hypothesis (H_a)* - The average time to complete an oil sample analysis is *less than* 60 minutes.

Compressor "Infant Mortality"

- *Null Hypothesis (H_o)* - The Dead-on-Arrival (DOA) rate for "Z-Compressors" is 12%.
- *Alternative Hypothesis (H_a)* - The compressor failure rate is *less than* 12%.

Braze Coverage

- *Null Hypothesis (H_o)* - There are no differences in braze coverage for flux A or B.
- *Alternative Hypothesis (H_a)* - Flux A's braze coverage is greater than flux B.

Machining Variability

- *Null Hypothesis (H_o)* – The current machining process can hold a tolerance of +/- 0.010."
- *Alternative Hypothesis (H_a)* - The new machining process can hold a tolerance of less than +/- 0.010."

Notes about the Null and Alternative Hypotheses:

1. Both hypotheses are statements about a *population parameter* (e.g. population mean, standard deviation, proportion, etc.). We never make hypotheses about the sample *statistic* (recall from the Basic Statistical Method that we are trying to make some inference about the *population,* and we use the sample statistics as our estimates of the population parameters).

2. The null hypothesis can arise from experience, a theory or model, design specifications or standards, or our goals & objectives.

3. Our desire is to *reject* the null hypothesis in favor of the alternative hypothesis. From the examples, you can see that the null hypothesis is always stated, "The XXX (population parameter) *is* some value." This represents our current experience or understanding. The alternative hypothesis is stated, "The XXX is (greater than, less than, not equal to) some value."

Rejecting a null hypothesis is a *strong* conclusion. It means we have established enough evidence that something is different. Failing to reject a null hypothesis is a *weak* conclusion. Here, all we can say is that there is not enough evidence to reject the null hypothesis. We have not "proved" the null hypothesis.

This hypothesis "philosophy" is very much like the American judicial system. We start with an "innocent until proven guilty" mind-set (the null hypothesis).

If we establish enough evidence to "convict" the person, we can claim that the person is "guilty," i.e. that we reject the innocent hypothesis in favor of the guilty hypothesis.

If we cannot establish enough evidence, we don't declare that the person is "innocent," we declare them to be "not guilty." The "not guilty" verdict is weaker than the "guilty" statement.

2. Choose a Significance Level (α - "alpha")

When we perform a hypothesis test, we reach a conclusion about the state of the "real world." Our conclusion may be either correct or incorrect. The following table shows the four possible outcomes:

	"Real World" Situation	
Test Conclusion:	**H_o is true**	**H_o is False**
Reject (H_o)	**Type I Error (α)**	**OK – Good Decision**
Do Not Reject (H_o)	**OK – Good Decision**	**Type II Error (β)**

A Type I error occurs when we *reject* the null hypothesis and in fact it is true. "Alpha" (α) is the probability of making a Type I error and is known as the *significance level* of the test.

A Type II error, on the other hand, occurs when we *accept* the null hypothesis (i.e. do not reject H_o) when it is false. "Beta" (β) is the probability of making a Type II error and 1 - β is known as the *power of the test.*

The size of both errors can be controlled. In general, they depend on the sample size and the variance of the underlying population(s). In practice, however, the sample size is often fixed (usually by cost or other resource constraint) and then the analyst sets the value of α. β is then determined by the other variables (see *The Power of the Test,* later in this unit for further discussion of β).

Alpha is determined by the *consequences* or *risks* associated with the decision we are trying to make. If the consequences associated with rejecting the null hypothesis in error are *low*, then α may be set at a large value (0.1 or 0.2). If the consequences associated with rejecting the null hypothesis in error are *high*, then α may be set at a low value (0.05 or 0.01 or lower). The larger the value of α, the easier it will be to reject the null hypothesis in favor of the alternative. On the contrary, the lower the α, the harder it will be to reject the null hypothesis.

Since taking risks is a management function, α should always be set with input from the responsible manager or decision maker. Many hypothesis tests "default" to using an $\alpha = 0.05$ (i.e. the test is significant at the 5% level). This can be bad *management* practice.[2]

Note also that you cannot set α to be zero and still *sample* from the population!! If you wish to have a zero risk of making an error, your only choice is to examine every item of the population! This assumes, of course, that your measurement system does not introduce errors! Very few measurement systems meet this criterion.

Here are a few examples of setting the Type I (α) error level:

Laboratory Analysis - The laboratory director wants to see if the team's changes have decreased the time to complete an oil sample analysis. Since she will have to "defend" this improvement to the management, she sets $\alpha = 0.05$. That is, she is willing to live with a 5% chance of being wrong in declaring the change to be an improvement.

Compressor Infant Mortality - Here, the Black Belt wishes to detect *any* decrease in failure rate. He is willing to live with a high Type I error level and sets $\alpha = 0.2$.

Braze Coverage - The supervisor knows that flux A costs about twice that of flux B, so she wants a high level of assurance that there really is a difference in coverage. She sets $\alpha = 0.01$.

Machining Variability - The shift supervisor has been getting "grief" from quality department about the high number of out of spec components. The supervisor wants to be able to provide reasonable assurance to the quality department that the process changes have been effective; she sets $\alpha = 0.05$.

Medical Device Experiment – A team is early in the concept phase of developing a new surgical device. They are trying to determine which design factors are important. They set up a Fractional Factorial design. Here, because they are looking for signals, they will set the risk level higher, e.g., $\alpha = 0.10$.

[2]As opposed to bad *statistical* practice!

3. Plan the Test:
a) Choose the Test Statistic (formula)
b) Determine the Rejection Region.

Choosing the Test Statistic

When we collect the sample data, we'll "crunch" it through a formula to obtain a value that will help us decide whether to reject the null hypothesis. The formula depends on the answers to several sequential questions:

1. Are there one, two or more populations?
2. Is the data discrete or continuous?
3. If the data is continuous, are we interested in the mean or the variance of the population?
4. If we are interested in the mean, do we know the population's variance?

Sections 9.2.4 through 9.2.6 provide the details associated with each path through these questions.

Select the Critical Region

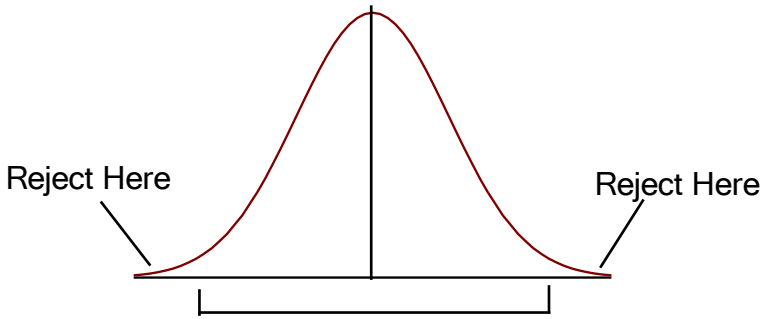
Rejecting or not rejecting the null hypothesis is a "GO/NO-GO" decision. We must establish two "regions" (actually, intervals on the number line). If the value of the test statistic falls in one of these "regions," then we will not reject the null hypothesis. If the value of the test statistic falls outside this first "region," then we *will* reject the null hypothesis.[3]

The test we conduct is built around the null hypothesis. For example, our null hypothesis may be that the average turnaround time for oil sample lab analyses is 60 minutes. If the null hypothesis is *true,* then we know the distribution into which the test statistic must fall.

If the actual value of the test statistic falls near the "middle" of the distribution, then we conclude that the sample data was obtained from the "null hypothesis population" (and do not reject the null hypothesis). If, on the other hand, the test statistic falls near the "tails" of the distribution, then we conclude that the sample data was *not* obtained from the "null hypothesis population" (and *reject* the null hypothesis).

Here's the picture:

[3]The control chart analogy can be drawn here. If a subgroup value falls outside the control limits, we "reject" the fact that it came from a system of common causes.



But how do we determine where the "middle" ends and the "tails" begin? There are three key factors that make this determination:

The Distribution of the Test Statistic - The distribution will determine the shape of the curve, its center and width.

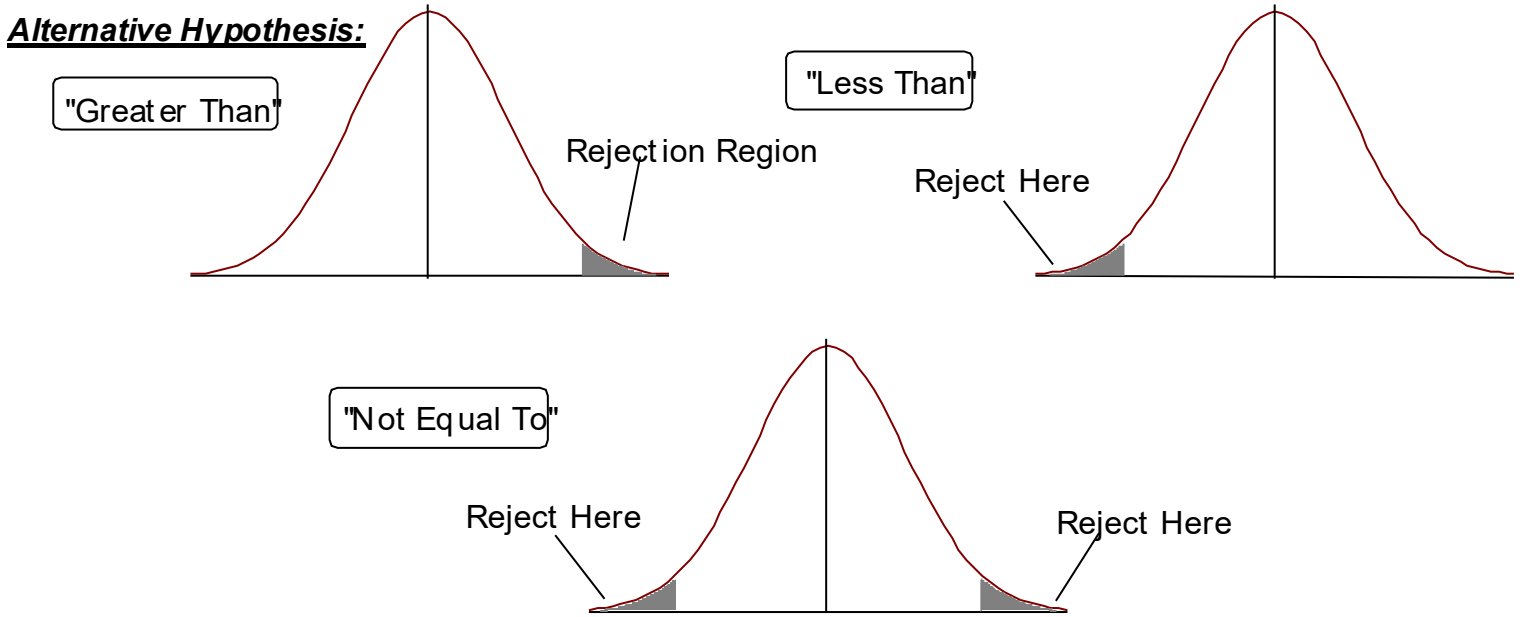
The Value of α - The smaller the value of α, the farther away from the middle we will need to be to reject the null hypothesis. Recall the definition of α:

α is the probability of rejecting the null hypothesis (H_o) when it is true.

To make this probability "small," we will want to make the "non-rejection" region as large as possible, i.e. include more of the area under the distribution. In fact, α is the area under the distribution curve *associated with the rejection region*.

The Alternative Hypothesis - There are three types of alternative hypothesis test we can generate: the *greater than,* the *less than*, and the *not equal to* statements.

For the first two statements, we will set up a "one-sided" test; for the third, we will set up a "two-sided" test:



The cumulative probability distribution tables (Appendix A) are constructed so that you can obtain these values easily. For the symmetric *normalized* distributions (normal and Student's - t), the tables are set up for the "Greater Than" Alternative Hypothesis. You enter the table with α (and, for the student's - t, the *degrees of freedom*). You then read from the table the value of K_α, the dividing line between the rejection and *non*-rejection regions.

If your test is a "Less Than" Alternative Hypothesis, the K_α is simply the negative of the "Greater Than" test value. If your test is *two-sided*, divide α by 2 and find the $K_{\alpha/2}$ value. The rejection *regions* are those to the right and left of the +/- values of $K_{\alpha/2}$.

For the chi-square (non-symmetric) distribution, the table provides you K_α values for both the right and left side of the distribution. For the F-distribution, only the "Greater Than" rejection region is provided, since this is generally all that is needed.

4. Collect Data and Calculate the Test Statistic

Although this seems like the easy part, there are a few comments to make about this step:

First, we've deliberately set up the procedure so that you make all the important decisions ahead of this step. There is always the temptation to try and make the data fit our beliefs. If we're honest statisticians and scientists, we'll go the other way - make our *beliefs* fit the *data*.

Even if you have the data in hand before you begin the hypothesis test process, try not to let your knowledge of the data affect the decisions you make in the first three steps.

Second, the *amount* of data you collect can also influence the hypothesis test. The Sampling Unit (9.3) will describe how to consider this issue.

The actual test statistic calculation will be based on the criteria described in Step 3 - *Plan the Test*. The specific test statistics are presented in Sections 9.2.4 through 9.2.6.

5. Draw a Conclusion

Here's where we bring it all together. The value of the test statistic is compared to the rejection region set up in Step 3. The conclusion is based on the following:

Test Statistic Value	Conclusion
Not in Rejection Region	Do Not Reject Null Hypothesis
In Rejection Region	Reject Null Hypothesis in Favor of Alternative Hypothesis

This may sound picky, but it's good practice to include all the elements of the decision in your conclusion statement:

> *"The test conclusion is to reject the null hypothesis that the oil sample analyses take an average of 60 minutes in favor of the alternative hypothesis that the average time is less than 60 minutes at a 0.05 level of significance."*

This statement provides your audience with the complete statement of the hypothesis test, including the *risk* you assigned to making a Type I error.

6. Estimate the Parameter of Interest

Step 6 is an "optional" part of the hypothesis test. By Step 5, we have concluded to reject or not reject the null hypothesis. If the null hypothesis is *not rejected,* we may decide to skip step 6 since there was no "new" knowledge gained.

However, if the null hypothesis was rejected, we're saying that our "old" knowledge needs to be replaced with "new" knowledge. Since we are trying to learn about the population (and its parameters), our next question will naturally be "What *is* the value of the population parameter?"

The statistic calculated from the sample data provides us with a *point estimate* of the parameter (i.e. mean, variance, proportion). But we must recognize that there is some uncertainty in this estimate.

We will construct an *interval estimate* that attempts to quantify this uncertainty. The interval will be a "plus/minus" band around the point estimate, with the width of the interval dependent on the test statistic's variance and a confidence level (1 - α) for the interval. This interval is known as either the "precision of the estimate" or the "error of the estimate."

For example, an interval statement might read:

> *"The oil sample analysis average turnaround time is 42 minutes, plus or minus 5 minutes at a 95% confidence level."*

How do we interpret such a statement? Well, if you subscribe to the classical school of statistics, your explanation will be: *"If we were to repeatedly sample from the population and construct such intervals, 95% of the intervals would 'contain' the population parameter."*

If you are of the Bayesian School, you will say, *"There is a 95% probability the population parameter is contained by the interval."*

Here, too, the sample size will influence the width of the interval. In general, the larger the sample size, the tighter the interval. One of the challenges here is to balance the desired precision with the cost of the study. Again, we will address these issues in Unit 9.3, Sampling.

9.2.4 Differences in the Center

Here, we will present tests associated with the *center* of the population. There are many different tests developed to investigate "center" issues, we will include six of the more common situations. In all these situations, the data we collect is *continuous*. The major groupings of these hypothesis tests are based on whether we are dealing with one or more populations:

Tests for One Population -

Typical questions here include the following:

- Has the population mean increased or decreased since the last time we measured it?
- Does this population meet our standard target?
- Does this process change affect the mean value?

There are two tests of interest here, depending on whether we "know" the population's variance (or standard deviation):

Population Variance (σ^2) known?	Test To Use	Section
Yes	Z - Test	9.2.4.1
No	t - Test	9.2.4.2

Tests for Two Populations

Typical questions here will include the following:

- Is there a difference in the means of these two populations (i.e. from two vendors or departments)?
- If we are doing a longitudinal study, is there a difference in the before and after populations?
  (this leads to the paired sample test)

The decision tree is a bit more complicated here. We first consider whether the samples are paired, i.e. are we taking repeated measurements on the same sample drawn from our population (e.g. eddy current measurements of tube thickness were measured in March and again in December after nine months of condenser operation. These two

measurements would constitute a paired sample.). Then, we ask questions about the population variances, are they known, and are they equal?

Paired Samples?	Population Variances known?	Population Variances Equal?	Test to Use	Section
Yes	N/A	N/A	Paired Sample t -Test	9.2.4.3
No	Yes	N/A	Two Population Z-Test	9.2.4.4
	No	Yes	Two - Population, Pooled Variance t-Test	9.2.4.5
		No	Two Population t-Test	9.2.4.6

Tests for More than Two Populations

We will cover these in the *Analysis of Variance* discussion (Unit 10.3).

9.2.4.1 Z - Test for Means, with Population Variance Known (1 Population)

Purpose and Description

This test is used to determine if the mean of a population (or lot) differs from some value. The "some value" could be a standard, or specification value, or it could arise from experience. For this test, we "know" the population variance. This assumption could arise, for example, from our knowledge that the process change would only affect the center of the distribution, not its width. If we have doubts about this assumption, then the *t - test* (Section 9.2.4.2) is more appropriate.

Using the Hypothesis Testing Process, we will develop our null and alternate hypotheses, decide on the risk of making an error and determine the rejection region. Then, we'll collect a sample of data from the population, calculate its mean and develop a *Z* test statistic. We will then compare this test statistic to the critical region and reject or not reject the null hypothesis.

Assumptions

1. This test is "best" if the distribution of the population is normal, or approximately so. Create a histogram of the sample data to check this assumption. Even if the population is skewed, if the sample size is large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

2. We "know" the population variance. As mentioned above, this assumption may be based on our knowledge of the effect of the change on the population or lot.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population mean is μ_0.
   b) Alternative Hypothesis (H_a) - The population mean is greater than/less than/not equal to μ_0.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

a) The Test Statistic is:

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

where:

$\bar{x}$ - *Sample Mean*

μ_0 - *Standard Mean*

σ - *Population StandardDeviation*

n - *Sample Size*

b) Determine the Rejection Region (K_α) - Using the value of α, consult the table of Standard Normal values to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** from the Population.*
b) Calculate the Mean of the Sample Data.
c) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Population Mean:

$$\bar{x} \pm K_{\alpha/2}\sigma / \sqrt{n}$$

Note that this provides a (1-α)% confidence interval for the mean.

Example - Difference of Means, Standard Deviation Known

A Supervisor has just received her productivity report and notices that it took 530 and 480 hours to build two air handlers last week. She wonders if this is natural variation, or if a significant change in the average time occurred. For the last 100 air handlers, the work hours averaged 450 hours with a standard deviation of 35 hours (based on this, we'll assume the standard deviation is "known"). At a 5% level of significance, can she say that the productivity has indeed decreased?

Hypothesis Test:

a) H_o: *Population Mean is 450 hours* ($\mu = 450 hrs.$)

H_a: *Population Mean is Greater Than 450 hours* ($\mu > 450 hrs.$)

b) $\alpha = 0.05$

c) *Test Statistic:* $Z = \dfrac{\overline{X} - \mu}{\sigma / \sqrt{n}}$ *Rejection Region:* $K_{0.05} = 1.645$ *(Normal)*

d) Calculations: $\overline{X} = (530 + 480) / 2 = 505 hours$

$$Z = \frac{505 - 450}{35 / \sqrt{2}} = 2.22$$

e) *Conclusion:* $2.22 > 1.645$, $\therefore$ *Reject* H_0 *in favor of* H_a *at the 5% significance level.*

f) Parameter Estimation: $\overline{X} \pm K_{\alpha/2}\sigma / \sqrt{2} = 505 \pm 1.96 \times 35 / \sqrt{2} = 505 \pm 48.5 hours$

9.2.4.2 t - Test for Means, Population Variance Unknown (1 Population)

Purpose and Description

This test is also used to determine if the mean of a population (or lot) differs from some value. The "some value" could be a standard, or specification value, or it could arise from experience. For this test, we do not "know" the population variance. This situation is more common than the one leading to the *Z - test.* The population variance will be estimated by the sample variance.

Using the Hypothesis Testing Process, we will develop our null and alternate hypotheses, decide on the risk of making an error and determine the rejection region. Then, we'll collect a sample of data from the population, calculate its mean and develop a *t* test statistic. We will then compare this test statistic to the critical region and reject or not reject the null hypothesis.

Assumptions

1. This test is "best" if the distribution of the population is normal, or approximately so. Create a histogram of the sample data to check this assumption. Even if the population is skewed, if the sample size is large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population mean is μ_0

   b) Alternative Hypothesis (H_a) - The population mean is greater than/less than/not equal to μ_0.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

a) The Test Statistic is:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

$$where:$$

$$\bar{x} \text{ - } Sample\ Mean$$

$$\mu_0 \text{ - } Standard\ Mean$$

$$s \text{ - } Sample\ Standard\ Deviation$$

$$n \text{ - } Sample\ Size$$

b) Determine the Rejection Region (K_α) - Using f, the degrees of freedom (f = n - 1) and the value of α, consult the student's t - distribution table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** from the Population*
b) Calculate the Mean of the Sample Data.
c) Calculate the Standard Deviation of the Sample Data.
c) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Population Mean:

$$\bar{x} \pm K_{\alpha/2} s / \sqrt{n}$$

Note that this provides a (1-α)% confidence interval for the mean.

Example - Difference of Means, Standard Deviation Unknown

A supplier of vendor payment services guarantees an average processing time of 24 hours from receipt of the invoice to preparing the electronic check for the vendor. Over the last 10 weeks, they have averaged 26 hours, with a standard deviation of 1 hour. Is this a significant departure from their guarantee? (Test at $\alpha = 0.05$)

Hypothesis Test:

a) H_o: *Population Mean is 24 hours* ($\mu = 24 hrs.$)

H_a: *Population Mean is Greater Than 24 hours* ($\mu > 24 hrs.$)

b) $\alpha = 0.05$

c) *Test Statistic:* $t = \dfrac{\overline{X} - \mu}{s / \sqrt{n}}$ *Rejection Region:* $K_{0.05} = 1.83$ *(t - dist., f = 10 - 1 = 9)*

d) Calculations: $t = \dfrac{26 - 24}{1 / \sqrt{10}} = 6.32$

e) *Conclusion: 6.32 > 1.83,* $\therefore$ *Reject* H_0 *in favor of* H_a *at the 5% significance level.*

f) Parameter Estimation: $\overline{X} \pm K_{\alpha/2} s / \sqrt{10} = 26 \pm 2.262 \times 1 / \sqrt{10} = 26 \pm 0.72 hours$

9.2.4.3 Paired Sample t- Test for Means (1 or 2 Populations)

Purpose and Description

In some cases, it is to our advantage to sample from the population and take two measurements from each sample element. Very often, the second measurement will occur after some period, or after some factor has been allowed to act upon the population. The author's first experience with this situation involved assessing the degradation in heat exchanger tubes that had occurred over an 18-month period. Each tube's degradation measurements were taken at the beginning and end of the 18 months, and these two numbers were a *paired sample.*

Generalizing this idea, we could take simultaneous measures of samples (i.e. under the same general conditions of temperature, humidity, and other environmental factors) from *two* populations and consider these measurements a paired sample. This strategy may allow us to eliminate some sources of variation in which we are not interested.

Assumptions

1. This test is "best" if the distribution of the paired sample *differences* is normal, or approximately so. Create a histogram of the differences to check this assumption. Even if this is skewed, if the sample size is large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The mean difference is 0.
   b) Alternative Hypothesis (H_a) - The mean difference is greater than/less than/not equal to 0.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is:

$$t = \frac{\bar{d}}{s / \sqrt{n}}$$

where:

$\bar{d}$ - *Sample Differences' Mean*

s - *Sample Standard Deviation*

n - *Sample Size*

b) Determine the Rejection Region (K_α) - Using f, the degrees of freedom (f = n - 1) and the value of α, consult the student's t-distribution table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

   *a) Draw a Random Sample(s) of size **n** from the Population(s).*
   b) Calcúlate differences (i.e. $d_1 = x_1 - y_1$, $d_2 = x_2 - y_2$, etc.).
   c) Calculate the Mean of the Differences.
   d) Calculate the Standard Deviation of the Differences.
   e) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Difference:

$$\bar{d} \pm K_{\alpha/2} s / \sqrt{n}$$

Example - Paired Samples

A new, less expensive pressure gage is being evaluated. Twenty-five pressure measurements were made with both the existing and new gages. The *differences* in the pressures are listed below. Can the new gage claim to be as accurate (here, accurate is defined as *unbiased*) as the existing one? (Test at $\alpha = 0.05$)

Measurement #	Pressure Difference (mm Hg)	Measurement #	Pressure Difference (mm Hg)	Measurement #	Pressure Difference (mm Hg)
1	-10	10	-10	19	-4
2	-15	11	-5	20	-7
3	-5	12	-4	21	+2
4	+5	13	+6	22	0
5	-6	14	-12	23	+12
6	+4	15	-10	24	-6
7	-13	16	-9	25	-2
8	-8	17	+8		
9	+7	18	0		

Hypothesis Test:

a) H_o : Population Difference is 0 mmHg ($d = 0 mmHg.$)

H_a : Population Difference is Not Equal to 0mmHg ($d \neq 0 mmHg.$)

b) $\alpha = 0.05$

c) Test Statistic : $t = \dfrac{\bar{d}}{s/\sqrt{n}}$ Rejection Region : $K_{0.05/2} = \pm 2.064$ (t - Dist., f = 25 - 1 = 24)

d) Calculations : $\bar{d} = (-10 + -15 + -5 + \ldots + -6 + -2)/25 = -3.28 mmHg$

$$s = \frac{1}{25-1}\left((-10-(-3.28))^2 + (-15-(-3.28))^2 + \ldots (-2-(-3.28))^2\right) = 7.24 \text{ and } t = \frac{-3.28}{7.24/\sqrt{25}} = -2.27$$

e) Conclusion : - 2.27 < -2.064, ∴ Reject H_0 in favor of H_a at the 5% significance level.

f) Parameter Estimation : $\bar{d} \pm K_{\alpha/2} s/\sqrt{25} = -3.28 \pm 2.064 \times 7.24/\sqrt{25}$

$= -3.28 \pm 2.99 mmHg$

9.2.4.4 Z - Test for Means, Population Variances Known (2 Populations)

Purpose and Description

Here, we are dealing with two populations and are questioning if their population means are equal or not. Situations may arise where two machines, shifts or vendors produce the same product or service, and we want to know if their average performance differs. In this Hypothesis Test, we will take two samples, one from each population. The population variances are known; they do not have to be equal to perform this test. We will present two cases for this test. The first is for *equal* sample sizes, the second, for *unequal* sample sizes.

Assumptions

1. This test is "best" if the distributions of the populations are normal, or approximately so. Create histograms of the sample data to check this assumption. Even though the populations are skewed, if the sample sizes are large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population means are equal ($\mu_A = \mu_B$)

   b) Alternative Hypothesis (H_a) - One population's mean is greater than/less than/not equal to the other.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

a.1) For Equal Sample Sizes, the Test Statistic is:	*a.2) For unequal sample sizes, the test statistic is:*
$$Z = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{1}{n}(\sigma_A^2 + \sigma_B^2)}}$$ *where*: $\bar{x}_i$ - *Sample Mean* σ_i - *Population Standard Deviation* n - *Sample Size*	$$Z = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$$ *where*: $\bar{x}_i$ - *Sample Mean* σ_i - *Population Standard Deviation* n_i - *Sample Size*

b) Determine the Rejection Region (K_α) - Using the value of α, consult the table of Standard Normal values to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

   *a) Draw Random Samples of size **n** (or n_A and n_B) from the Populations.*
   b) Calculate the Means of the Sample Data.
   c) Calculate the appropriate Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Population Means:

$$\bar{x}_i \pm K_{\alpha/2}\sigma_i / \sqrt{n_i}$$

Example – Test of Means, Population Variances Known

A lathe team has been experimenting with different methods of setting up a particular job. In 25 tests of one method (A), the team averaged 22 minutes. In 25 tests of method B, the team averaged 15 minutes. Assume that the standard deviations are known for both methods and are equal to 6 minutes. At a 5% level of significance, can they conclude that method B requires less time than method A?

Hypothesis Test:

$$a)\ H_o : \text{Population Means are Equal } (\mu_A = \mu_B)$$

$$H_a : \text{"A" Population Mean is Greater Than "B" } (\mu_A > \mu_B)$$

$$b)\ \alpha = 0.05$$

$$c)\ \text{Test Statistic : } Z = \frac{\overline{X}_A - \overline{X}_B}{\sqrt{(\sigma_A^2 + \sigma_B^2)/n}}$$

$$\text{Rejection Region : } K_{0.05} = 1.645 \text{ (Z - dist.)}$$

$$d)\ \text{Calculations : } Z = \frac{22 - 15}{\sqrt{(6^2 + 6^2)/25}} = 4.12 \qquad \text{(Assume Equal Variances here)}$$

$$e)\ \text{Conclusion : } 4.12 > 1.645, \therefore \text{Reject } H_0 \text{ in favor of } H_a \text{ at the 5\% significance level.}$$

$$f)\ \text{Parameter Estimation : } \quad \overline{X} \pm K_{\alpha/2} s/\sqrt{2} = 15 \pm 1.96 \times 6/\sqrt{25} = 15 \pm 2.35 \min.$$

9.2.4.5 t - Test for Means, Population Variances Unknown, but Equal (2 Populations)

Purpose and Description

Here again, we are dealing with two populations and are questioning whether their population means are equal or not. In this Hypothesis Test, we will take two samples, one from each population. The population variances are *not* known; but they are equal.[4] We will estimate the population variances with the sample variances. Sample sizes may be equal or not.

Assumptions

1. This test is "best" if the distributions of the populations are normal, or approximately so. Create histograms of the sample data to check this assumption. Even though the populations are skewed, if the sample sizes are large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population means are equal.
   b) Alternative Hypothesis (H_a) - One population mean is greater than/less than/not equal to the other.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

[4]This may be established by using the Hypothesis Test for Variances, Section 9.2.5.

a.1) For equal sample sizes, the Test Statistic is:	*a.2) For unequal sample sizes, the Test Statistic is:*
$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s_A^2 + s_B^2}{n}}}$$ *where*: $\bar{x}_i$ - *Sample Mean* s_i - *Sample Standard Deviation* n - *Sample Size*	$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\left(\frac{1}{n_A} + \frac{1}{n_B}\right)\frac{SS_A + SS_B}{n_A + n_B - 2}}}$$ *where*: $\bar{x}_i$ - *Sample Mean* SS_i - *Sample Sum of Squares (as follows):* $$SS_i = \sum_{i=1}^{n_i} x_i^2 - \frac{1}{n_i}\left(\sum_{i=1}^{n_i} x_i\right)^2 \text{ or } SS_i = (n_i - 1)s_i^2$$ n_i - *Sample Size*

b) Determine the Rejection Region (K_α) - Using f, the degrees of freedom (f = 2(n - 1) for equal sample sizes or f = n_A + n_B - 2 for unequal sample sizes) and the value of α, consult the student's t-distribution table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

   *a) Draw Random Samples of size **n** (or n_A and n_B) from the Populations.*
   b) Calculate the Means of the Sample Data.
   c) Calculate the Sample Sums of Squares.
   d) Calculate the appropriate Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Population Means:

$$\bar{x}_i \pm K_{\alpha/2} s_i / \sqrt{n_i}$$

where :

$$s_i - \text{Sample Standard Deviation (as follows)} : \mathrm{s_i} = \sqrt{SS_i/(n_i - 1)}$$

Example - Difference of Means, Same Sample Size

A lathe team has been experimenting with different methods of setting up a job. In 25 tests of one method (A), the team averaged 22 minutes, with a standard deviation of 6 minutes. In 25 tests of method B, the team averaged 15 minutes, with a standard deviation of 5.8 minutes. At a 5% level of significance, can they conclude that method B requires less time than method A?

Hypothesis Test:

a) H_o: *Population Means are Equal* ($\mu_A = \mu_B$)

H_a: *"A" Population Mean is Greater Than "B"* ($\mu_A > \mu_B$)

b) $\alpha = 0.05$

c) Test Statistic: $t = \dfrac{\overline{X}_A - \overline{X}_B}{\sqrt{(s_A^2 + s_B^2)/n}}$

Rejection Region: $K_{0.05} = 1.678$ *(t - dist., f = 2(25 - 1) = 48)*

d) Calculations: $t = \dfrac{22 - 15}{\sqrt{(6^2 + 5.8^2)/25}} = 4.19$ *(Assume Equal Variances here)*

e) Conclusion: 4.19 > 1.678, $\therefore$ *Reject* H_0 *in favor of* H_a *at the 5% significance level.*

f) Parameter Estimation: $\overline{X} \pm K_{\alpha/2} s/\sqrt{2} = 15 \pm 2.011 \times 5.8/\sqrt{25} = 15 \pm 2.33 \text{min.}$

Note that this example is like the case where the Standard Deviation is known. Contrast the difference in the calculations and result.

Example - Difference of Means, Different Standard Deviations

A billing supervisor is interested in reducing the cost of collecting overdue accounts. She tries two different methods of collection and finds the following:

	Company Collectors	Outside Contractors
Average -	$137.00	$149.00
Standard Deviation -	$19.00	$14.00
Number of Bills -	16	24

What conclusion (if any) can she draw about the relative cost of the two methods, at a 5% level of significance?

Hypothesis Test:

$a)$ H_o : Population Means are Equal $(\mu_{Company} = \mu_{Cont})$

H_a : Contractor Population Mean is Greater Than Company Mean $(\mu_{Cont} > \mu_{Company})$

$b)$ $\alpha = 0.05$

$c)$ Test Statistic : $$t = \frac{\overline{X}_{Cont} - \overline{X}_{Company}}{\sqrt{(1/n_{Cont} + 1/n_{York})\dfrac{SS_{Cont} + SS_{Company}}{n_{Cont} + n_{Company} - 2}}}$$

Rejection Region : $K_{0.05} = 1.687$ (t - dist., f = 24 + 16 - 2 = 38)

d) Calculations : $$t = \frac{149 - 137}{\sqrt{(1/24 + 1/16)\dfrac{(24-1)14^2 + (16-1)19^2}{24 + 16 - 2}}} = 2.3$$

(Assume the Variances are equal)

$e)$ Conclusion : $2.3 > 1.687, \therefore$ Reject H_0 in favor of H_a at the 5% significance level.

f) Parameter Estimation : $\overline{X}_{Company} \pm K_{\alpha/2} s_Y / \sqrt{n_Y} = \$137 \pm 2.131 \times \$19 / \sqrt{16} = \$137 \pm \$10.10$

We've assumed equal variances here. What if someone challenged that assumption? Here's the test for equality of variances (See 9.2.5.2):

$a)\ \mathrm{H_o}$: Population Variances are Equal $(\sigma^2_{\text{Company}} = \sigma^2_{Cont})$

H_a : Company Collectors' Variance is Greater Than Contractors' $(\sigma^2_{\text{Company}} > \sigma^2_{Cont})$

$b)\ \alpha = 0.05$

$c)$ Test Statistic : $\mathrm{F} = \dfrac{\mathrm{s}^2_{\text{Company}}}{s^2_{Cont}}$

Rejection Region : $\mathrm{K}_{0.05} = 2.13\ (\mathrm{F}\text{ - Dist.}, \mathrm{f_n} = 16-1 = 15, f_d = 24-1 = 23)$

d) Calculations : $\mathrm{F} = 19^2 / 16^2 = 1.41$

$e)$ Conclusion : $1.41 < 2.13, \therefore$ Do Not Reject H_0 in favor of $\mathrm{H_a}$ at the 5% significance level.

f) Parameter Estimation : N/A

9.2.4.6 t - Test for Means, Population Variances Unknown, Not Equal (2 Populations)

Purpose and Description

This last test of means is provided for the case where we wish to determine if two populations' means are different, but we do not *know* the population variances and the sample variance estimates lead us to believe they are different. Sample sizes may be equal or not; there is no difference in the test statistic's calculation.

Assumptions

1. This test is "best" if the distributions of the populations are normal, or approximately so. Create histograms of the sample data to check this assumption. Even though the populations are skewed, if the sample sizes are large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population means are equal
   b) Alternative Hypothesis (H_a) - One population mean is greater than/less than/not equal to the other.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is:

$$t = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{s_A^2 / n_A + s_B^2 / n_B}}$$

$$where:$$

$$\bar{x}_i \text{ - Sample Mean}$$

$$s_i \text{ - Sample Standard Deviation}$$

$$n_i \text{ - Sample Size}$$

b) Determine the Rejection Region (K_α) –

First, calculate c as follows:

$$c = \frac{s_A^2}{n_A} \Bigg/ \left\{ \frac{s_A^2}{n_A} + \frac{s_B^2}{n_B} \right\}$$

Then, calculate the number of degrees of freedom, f:

$$f = \left(\frac{c^2}{n_A - 1} + \frac{(1-c)^2}{n_B - 1} \right)^{-1}$$

Using the degrees of freedom, f, and the value of α, consult the student's t-distribution table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** (or n_A and n_B) from the Populations.*
b) Calculate the Means of the Sample Data.
c) Calculate the Sample Standard Deviations.
d) Calculate the appropriate Test Statistic shown above in Step 3a.
e) Determine the number of degrees of freedom as shown above.

5. Draw a Conclusion.

6. Estimate the Population Means:

$$\bar{x}_i \pm K_{\alpha/2} s_i / \sqrt{n_i}$$

Example – Means Comparisons – Standard Deviations Unknown, Not Equal

A scheduling group is experimenting with two different ways of estimating the time required to complete projects. Using method 1 on 15 projects, they find the difference between the estimated and actual project lengths to be 30 days with a standard deviation of 15 days. Using method 2 on 21 projects, they find the difference between the estimated and actual project lengths to be 21 days with a standard deviation of 8 days. Can the group conclude that method 2 is a better predictor of the average project time? (Test at $\alpha = 0.05$)

Hypothesis Test:

$a)\ \mathrm{H_o}$: Population Means are Equal $(\mu_1 = \mu_2)$

H_a : Population "1" Mean is Greater Than "2" $(\mu_1 > \mu_2)$

$b)\ \alpha = 0.05$

$c)$ Test Statistic : $t = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$

Rejection Region : DoF : $f = \left(\dfrac{c^2}{n_1 - 1} + \dfrac{(1-c)^2}{n_2 - 1} \right)^{-1}$ where : $c = \dfrac{s_1^2}{n_1} \Big/ \left\{ \dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2} \right\}$

$$c = \frac{15^2}{15} \Big/ \left\{ \frac{15^2}{15} + \frac{8^2}{21} \right\} = 0.83 \text{ and } f = \left(\frac{0.83^2}{15-1} + \frac{(1-0.83)^2}{21-1} \right)^{-1} = 19.74$$

Rejection Region : $K_\alpha : \alpha = .05, f = 20 : K_\alpha = 1.725$

d) Calculations : $t = \dfrac{30 - 21}{\sqrt{15^2/15 + 8^2/21}} = 2.12$

$e)$ Conclusion : t = 2.12 > 1.725, Therefore reject the null hypothesis in favor of the alternate.

f) Parameter Estimation : $21\,days \pm 2.086 \times 8/\sqrt{21} = 21\,days \pm 3.6\,days$

9.2.5 Differences in the Variation

In this section, we will present tests associated with the *variation* of the population. There are only two different tests we will consider here. The first deals with the situation where we have one population and we are comparing its standard deviation (actually, we'll deal with the *variance*) to some standard or known value. The second test allows us to compare two populations' variances to determine if they are equal or different. This latter test is a prelude to several of the tests of means presented in the previous section.

In all these situations, the data we collect is *continuous*. Hypothesis tests for more than two populations' variances are beyond the scope of this manual (a simple test is to construct an R-Chart or s-Chart and check to see if it is in or out of control).

Typical questions here include the following:

- Has the population variance increased or decreased since the last time we measured it?
- Does this population's variance meet our standard target?
- Does this process change affect the variation value?
- Are these two populations' variances equal?

Comparison	Test To Use	Section
Population Variance to a Standard	χ^2 - Test	9.2.5.1
Two Population Variances	F - Test	9.2.5.2

9.2.5.1 χ^2 - Test for Population Variance Comparison to a Standard Value

Purpose and Description

This test is used to determine if the variance of a population (or lot) differs from some value. The "some value" could be a standard, or specification value, or it could arise from experience.

Assumptions

1. This test is "best" if the distribution of the population is normal, or approximately so. Create a histogram of the sample data to check this assumption. Even if the population is skewed, if the sample size is large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population variance is σ_0^2

   .

   b) Alternative Hypothesis (H_a) - The population variance is greater than/less than/not equal to σ_0^2

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

where:

s - *Sample Standard Deviation*

σ_0 - *"Standard" Standard Deviation*

n - *Sample Size*

b) Determine the Rejection Region (K_α) - Using f, the degrees of freedom (f = n - 1), and the value of α, consult the χ^2 distribution table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** from the Population*
b) Calculate the Standard Deviation of the Sample Data.
c) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Population Variance:

s^2 - *Point Estimate*

$SS/\chi^2(n-1, 1-\alpha/2)$ - *Upper Confidence Limit*

$SS/\chi^2(n-1, \alpha/2)$ - *Lower Confidence Limit*

where:

$$SS = s^2 \times (n-1)$$

Note that this provides a (1-α)% confidence interval for the variance.

Example - Standard Deviation – Comparison to a Standard

An information systems manager has been receiving complaints about the time required to respond to customer problems. The average time to respond to a particular type of problem is 36 hours, which is within his standard of 40 hours. But the standard deviation of times to respond is 14 hours.

He charters a team to work on reducing the variation in the process and they implement process changes. A sample of 20 jobs after the changes reveals that the average time is still 36 hours, but the standard deviation is now 8 hours. Can the team claim that their changes have made an improvement in the process' variation? (Test at α = 0.05)

Hypothesis Test:

$a)\ \mathrm{H_o}: \text{Population Variance is } 14^2 \text{ hours}^2\ (\sigma_o^2 = 196)$

$H_a: \text{Population Variance is Less Than } 14^2 \text{ hours}^2\ (\sigma_o^2 < 196)$

$b)\ \alpha = 0.05$

$c)\ \text{Test Statistic}: \chi^2 = \dfrac{(n-1)s^2}{\sigma_o^2}$ Rejection Region : $\mathrm{K}_{0.95} = 10.12\ (\chi^2 \text{ dist, f} = 20 - 1 = 19)$

(Note that we enter the Chi - Square table at $1 - \alpha$, or 0.95 here)

$\text{d) Calculations}: \chi^2 = \dfrac{(20-1)8^2}{14^2} = 6.2$

$e)\ \text{Conclusion}: 6.2 < 10.12, \therefore \text{Reject } \mathrm{H_0} \text{ in favor of } \mathrm{H_a} \text{ at the 5\% significance level.}$

$\text{f) Parameter Estimation}: \quad s = 8\,hours, UCL = \sqrt{1216/8.91} = 11.68,$

$LCL = \sqrt{1216/32.9} = 6.08$

9.2.5.2 F - Test Comparison of Two Population's Variances

Purpose and Description

This test is used to determine if population variances differ. This test may be performed prior to the test of means where the population variances are unknown, and you wish to establish if they are equal or not.

Two cases are presented. The first is to be used when the sample sizes are equal, the second when they are not.

Assumptions

1. This test is "best" if the population distributions are normal, or approximately so. Create histograms of the sample data to check this assumption. Even if the populations are skewed, if the sample size is large enough (for "slightly" skewed distributions, the minimum sample size is 6, for "seriously" skewed distribution, the minimum size is 60), then this test can be used.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population variances are equal.

   b) Alternative Hypothesis (H_a) - One population variance (σ_A^2) is greater than the other (σ_B^2).

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) For equal or unequal sample sizes, the Test Statistic is:

$$F = s_A^2 / s_B^2$$

where :

s_i - Sample Standard Deviation

b.1) (Equal Sample Sizes) Determine the Rejection Region (K_α) - Using f_n and f_d, the degrees of freedom for the F - Distribution (f_n and f_d = n - 1) and the value of α, consult the F - table to determine K_α.
b.2) (Unequal Sample Sizes) Determine the Rejection Region (K_α) - Using f_n and f_d, the degrees of freedom for the F - Distribution ($f_n = n_A - 1$ and $f_d = n_B - 1$) and the value of α, consult the F - table to determine K_α.

4. Collect Data and Calculate the Test Statistic:

*a) Draw Random Samples of size **n** (or n_A and n_B) from the Populations.*
b) Calculate the Standard Deviations of the Sample Data. Use the larger standard deviation as s_A and the smaller as s_B.
c) Calculate the Test Statistic shown above in Step 3a.
d) Determine the Degrees of Freedom for the Critical Region.

5. Draw a Conclusion.

6. Estimate the Population Variances:

$$s^2 \text{ - } \textit{Point Estimate}$$

$$SS / \chi^2(n-1, 1-\alpha/2) \text{ - } \textit{Upper Confidence Limit}$$

$$SS / \chi^2(n-1, \alpha/2) \text{ - } \textit{Lower Confidence Limit}$$

where:

$$SS = s^2 \times (n-1)$$

Note that this provides (1-α)% confidence intervals for the populations' variances.

Example – Comparison of Two Standard Deviations

Two sheet metal stamping machines are being evaluated for purchase. As part of the purchase decision, the machines' ability to hold tolerances was tested. Each machine stamps 60 metal pieces. The standard deviation of the measurements for machine A is 0.005" and that of machine B is 0.003." Can the manufacturer of machine B claim that they have a better machine? Test at an α = 0.05.

Hypothesis Test:

$a)\ \mathrm{H_o}$: Population Variances are equal $(\sigma_A^2 = \sigma_B^2)$

H_a : Population B Variance is Less ThanPopulation A $(\sigma_B^2 < \sigma_A^2)$

$b)\ \alpha = 0.05$

$c)$ Test Statistic : $F = s_A^2 / s_B^2$

Rejection Region : $\mathrm{K}_{0.05} = 1.53$ (F dist, n = d = 60, dof $= 60 - 1 = 59$ (but use 60, since it is the closest value on the F - Table)

d) Calculations : $\mathrm{F} = 0.005^2 / 0.003^2 = 2.77$

$e)$ Conclusion : $2.77 > 1.53$, $\therefore$ Reject H_0 in favor of H_a at the 5% significance level.

f) Parameter Estimation : $s = 0.003\,in., UCL = \sqrt{59(0.003)/40.5} = 0.0044\,in.,$

$$LCL = \sqrt{59(.003)/83.3} = 0.0021\,in.$$

9.2.6 Differences in Proportions & Rates

In this section, we will present tests used when the quality characteristic is of the *discrete* (or count) data type. Many processes are measured using "GO/NO-GO" criteria, such as leaks, missing parts, employee injuries, etc. In other cases, we impose a standard on the process and count the number of items that do not meet the standard.

For these cases, our hypothesis tests will involve *proportions* or *fractions.* There are only two different tests we will consider here. The first deals with the situation where we have one population and we are comparing its proportion to some standard or known value. The second test allows us to compare two populations' proportions to determine if they are equal or different.

Typical questions here include the following:

- Has the fraction defective or non-conforming increased or decreased since the last time we measured it?
- Does this population's proportion meet our standard target?
- Does this process change affect the proportion value?
- Are these two populations' proportions equal?

Comparison	Test To Use	Section
Population Proportion to a Standard	Z - Test	9.2.6.1
Two Population Proportions	Z - Test (2 Pop's)	9.2.6.2

For more than two populations, *contingency* analysis (Unit 7.2) can be employed to detect differences. If you are measuring the rate of occurrence of some event (e.g. defects such as paint scratches per door panel, injuries per plant per month, errors per application), and wish to detect differences in rates, see Unit 6.9.2, Analysis of Means (ANOM). The ANOM procedure can be used to detect differences in Poisson processes.

9.2.6.1 Z - Test for One Population Proportion

Purpose and Description

This test is used to determine if one population's (or lot's) *proportion* differs from some value. The "some value" could be a standard, or specification value, or it could arise from experience.

Assumptions

1. This test is "best" if the distribution of the population's *proportion* is normal, or approximately so. This is generally satisfied if the product of the sample size (*n*) and the proportion (*p*) is at least equal to 5 (also, the product of sample size and 1 - *p* should be at least 5). This test assumes the binomial distribution can be approximated by the normal distribution. When *np* is very small this assumption is not met.

2. Try not to "mix" different binomial populations together. For example, the population of large chillers experiencing leaks may be different that those small or mid-size chillers experiencing leaks. If you mix these together to test if the overall leak rate has changed, the binomial distribution may not be a good model for the overall proportion.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population proportion is ***P_0.***
   b) Alternative Hypothesis (H_a) - The population proportion is greater than/less than/not equal to ***P_0.***

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is:

$$Z = \frac{p - \mathbf{P}_0}{\sqrt{\mathbf{P}_0(1 - \mathbf{P}_0) / n}}$$

where:

p - *Sample Proportion*

n - *Sample Size*

b) Determine the Rejection Region (K_α) - Using the value of α, consult the Standard Normal table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** from the Population.*
b) Calculate the Proportion of the Sample Data that has the characteristic of interest.
c) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Population Proportion:

$$p \pm K_{\alpha/2}\sqrt{p(1 - p) / n}$$

Note that this provides a (1-α)% confidence interval for the population proportion.

Example – Proportion Compared to a Standard

Prior to beginning a drug-testing program in a large company, the management felt that 10% would test positive. The program administrator feels that this estimate is too high. She randomly samples 100 employees and 5 employees test positive. Should she feel comfortable in informing management that their 10% estimate is too high? (Test at $\alpha = 0.01$)

Hypothesis Test:

a) H_o: *Population Proportion is 0.1* ($P_o = 0.1$.)
H_a: *Population Proportion is Less Than 0.1* ($P_o < 0.1$.)

b) $\alpha = 0.01$

c) *Test Statistic:* $Z = \dfrac{p - P_o}{\sqrt{P_o(1 - P_o) / n}}$ *Rejection Region:* $K_{0.01} = -2.326$ *(Normal)*

d) Calculations: $p = 5 / 100 = 0.05$

$$Z = \frac{0.05 - 0.1}{\sqrt{0.1(1 - 0.1) / 100}} = -1.67$$

e) *Conclusion:* $-1.67 > -2.326$, $\therefore$ *Do not Reject* H_0 *in favor of* H_a *at the 1% significance level.*

f) Parameter Estimation: Although we cannot reject the null hypothesis, our best estimate of the population parameter is based on the administrator's sample, i.e. $p = 0.05 \pm 0.04$ *(@ 5%)*

9.2.6.2 Z - Test for Two Populations' Proportions

Purpose and Description

This test is used to determine if one population's (or lot's) *proportion* differs from another population.

Assumptions

1. This test is "best" if the distributions of the populations' *proportions* are normal, or approximately so. This is generally satisfied if the product of the sample size (*n*) and the proportion (*p*) is at least equal to 5 (also, the product of sample size and 1 - *p* should be at least 5).

This test assumes the binomial distribution can be approximated by the normal distribution. When *np* is very small this assumption is not met.

2. Try not to "mix" different binomial populations together. For example, the population of large chillers experiencing leaks may be different that those small or mid-size chillers experiencing leaks. If you mix these together to test if the overall leak rate has changed, the binomial distribution may not be a good model for the overall proportion.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population proportions are equal.
   b) Alternative Hypothesis (H_a) - One population proportion is greater than/less than/not equal to the other.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is:

$$Z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

where:

p_i - *Sample Proportion*

n_i - *Sample Size*

$$p = \frac{x_1 + x_2}{n_1 + n_2}$$

x_i - *Number of Sample Items with the Characteristic of Interest*

b) Determine the Rejection Region (K_α) - Using the value of α, consult the Standard Normal table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** (or n_1 and n_2) from the Populations.*
b) Calculate the Proportions from the Sample Data, which have the characteristic of interest.
c) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

6. Estimate the Populations' Proportions:

$$p \pm K_{\alpha/2}\sqrt{p(1-p)/n}$$

Note that this provides a (1-α)% confidence interval for the population proportion.

Example - Difference of Proportions

Twenty out of 100 sales orders from company-employed sales reps contain missing information. Thirty of 200 sales orders from outside sales reps contain missing information. Is there a difference in the proportion? Construct a 90% confidence interval for the order error proportion.

Hypothesis Test:

$a)$ H_o : Population Proportions are Equal ($P_O = P_I$.)

H_a : The Population Proportions are Not Equal ($P_O \neq P_I$)

$b)$ $\alpha = 0.1$

$c)$ Test Statistic : $Z = \dfrac{P_O - P_I}{\sqrt{p(1-p)(1/n_O + 1/n_I)}}$ Rejection Region : $K_{0.05} = \pm 1.645$ (Normal)

d) Calculations : $p_O = 30/200 = 0.15,\ p_I = 20/100 = 0.2,\ p = \dfrac{30+20}{200+100} = 0.167$

$$Z = \frac{0.15 - 0.2}{\sqrt{0.167(1-0.167)(1/200 + 1/100)}} = -1.09$$

$e)$ Conclusion : -1.09 is within ± 1.645, $\therefore$ Do Not Reject H_0 in favor of H_a at the 10% significance level.

f) Parameter Estimation :

$$p \pm 1.645\sqrt{p(1-p)/n} = 0.167 \pm 1.645\sqrt{0.167(0.833)/300} = 0.167 \pm 0.04$$

9.2.7 The Power of the Test & Other Mysteries

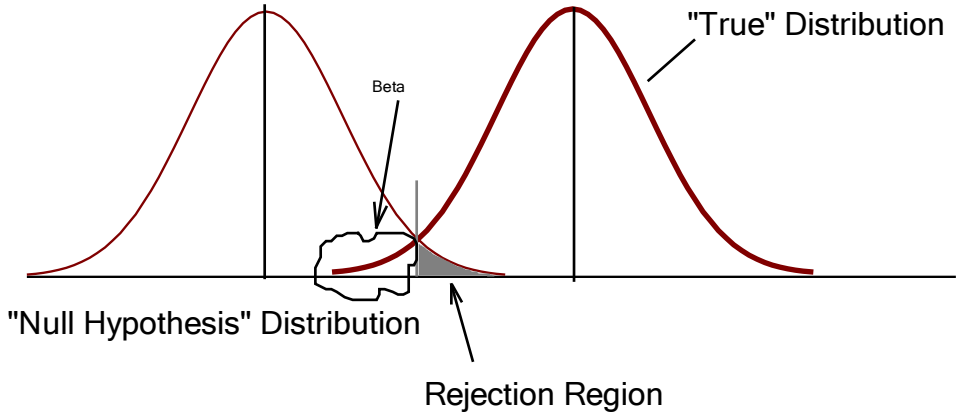
The Power of the Test

When we conduct a hypothesis test, there are two possible errors we can make. We have been primarily concerned with the Type I, or α error. But shouldn't we also be concerned with the Type II or β error? Let's recall the definition of the Type II error:

β is the probability of not rejecting the null hypothesis, when it is really false.

Here, we are saying that for one reason or another, we have not been able to "convict" a guilty person of their crime. Let's translate this concept into our picture of the hypothesis test:

"Accepting" a False Null Hypothesis:



You can see that, even though the "true" distribution is different from that postulated by our null hypothesis, there is still a chance that a sample drawn from the "true" distribution will not fall into the rejection region. The **power of the test** then, is the probability of correctly rejecting the null hypothesis, when the null is false; quantified, it is simply (1 - β).

The factors affecting the Type II error probability (i.e. β) are:

"Distance" between Parameter Values of Null Hypothesis and Actual Conditions - Obviously, the further the actual value of the mean, standard deviation or proportion is from the null hypothesis value, the less likely it will be to obtain a sample that does not fall into the rejection region. But we have to be careful about how we define this "further." The *dispersion* of the statistic is our "yardstick" here. For example, suppose that our null hypothesis is that the mean value is 10 cubits and the actual population's mean is 15 cubits.

If the standard deviation of the population is 1 cubit, and our sample size is 4, then the dispersion of the population mean will be $1 cubit / \sqrt{4} = 0.5 cubit$. The "distance" from the null hypothesis mean to the actual mean is then 10 standard deviations. The possibility of drawing a sample from the actual population that doesn't fall into the rejection region is *very small.*

Suppose, though, that the standard deviation of the population is 10 cubits, with a sample size of 4. Now, the dispersion of the population mean is *5 cubits.* The new "distance" from the null hypothesis mean to the actual mean is only 1 standard deviation. It's much more likely here that a sample drawn from the actual population will *not* fall into the rejection region.

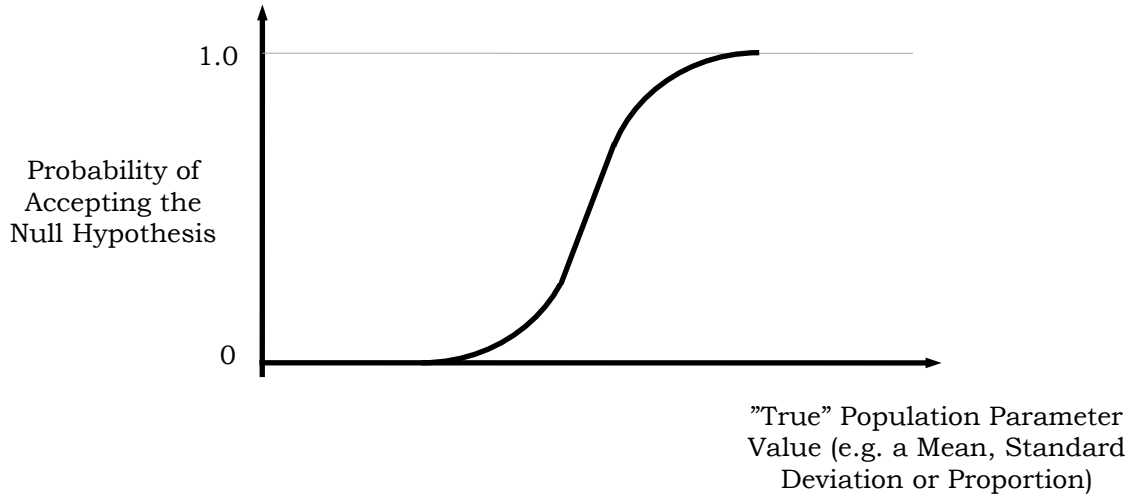
Sample Size - The larger the sample size, the better we will be able to detect actual differences between the hypothesis and the "true" population. Consider the second case presented above. With a sample size of only 4, the distance from the null hypothesis test to the actual mean was 1 standard deviation. What happens if we *quadruple* the sample size to 16?

The new dispersion of the mean is $10 cubits / \sqrt{16} = 2.5 cubits.$ Now, the "distance" from the null hypothesis mean to the actual mean is 5 cubits/2.5 cubits = 2 standard deviations. There's less of a chance that the sample mean will *not* fall into the rejection region. If we increase the sample size to 64, the "distance" increases to 4 standard deviations. And so on.

Type I Error - The two types of errors "balance" each other. If we set the Type I error probability (α) to be very small, then that will increase the Type II error probability. The smaller the Type I error probability, the harder it will be to reject the null hypothesis, thus, the easier it will be to *not reject* a false null hypothesis. Conversely, if you set the Type I error probability to be large, then it is harder to *not reject* the null hypothesis.

The Type II error concept is quantified and presented graphically as an *Operating Characteristic (OC) Curve.* OC curves show you the probability of *not rejecting* a Null Hypothesis as a function of the actual value of the population parameter. The shape and position of the OC curve depend on the factors discussed above.

In this example OC curve, the null hypothesis is that the population parameter equals some value, versus the alternative that it is less than that value. The smaller the "true" population value is, the smaller the probability of *not rejecting* the null hypothesis (i.e. smaller β).



Statistical Significance vs. Functional Significance

The hypothesis test is a “GO/NO-GO” test. We have either established enough evidence that something is different, or we have not. This, of course, is the *first step* in understanding “differences.” The second step is to examine the functional significance of the difference. A change to a process may produce a *statistically significant* result, but from a functional or operational standpoint, the change may be *insignificant.*

Always ask two questions when someone presents you with evidence supporting the effect of some change. First, is the change statistically significant? Have we managed to detect a signal amid the noise? Second, how important is this improvement from a functional standpoint. Is it cost-beneficial to make the change?

9.2.8 Non-Parametric Tests

The tests we've presented so far include assumptions about an underlying distribution of the data. For example, the Z-Test assumes that the population means can be modeled with a normal distribution. The t-test assumes that the population means can be modeled using the t-distribution. You may run across situations where meeting these assumptions is "iffy." Fortunately, there is a class of tests (called *non-parametric*) that may be employed. As their name implies, the tests do not invoke the use of a population parameter, such as the mean or variance. Although these tests may be employed in a variety of situations, their ***power*** (see last section) is generally not as high as the parametric tests.

Sample Comparison – Parametric vs. Non-Parametric Tests

Scenario	Appropriate Parametric Test	Non-Parametric Equivalent Test
2 Populations – Are their means different?	• 2-Population t-Test	• Mann-Whitney Test
1 Population – Are the Before & After means different?	• Paired t-Test	• Sign Test
Multiple Populations – Are the means different?	• Analysis of Variance (see Unit 10.3)	• Kruskal-Wallis Test

General Assumptions (or Lack Thereof!)

- The data values are not necessarily continuous. The data may include ***ordinal (frequency counts of scaled data – i.e. customer satisfaction scale of 1 – don't like, 3 – like, 5 – like very much)*** or ***nominal (frequency counts associated with categories – i.e. nuclear, gas, oil, hydro power plants)***.

- The inference being made doesn't involve a population parameter (previous hypothesis tests involved the ***mean, variance*** and ***proportion*** parameters).

- There are no assumptions about the underlying distribution of the sample – e.g. normality or equality of variance. Non-parametric tests may also be employed where the sample size is very small – in these cases, we don't have enough information to accurately determine the population.

9.2.8.1 One Sample Runs Test

Purpose and Description

The One Sample Runs test examines whether a series of items or events can be said to have occurred in a random sequence. We introduced this idea in Section 6 when we discussed control charts. Several of the assignable cause rules make use of the number of ***runs*** – *points above or below the center line of the control chart.*

Additional purposes of the One Sample Runs test include testing to see if a sample is truly *random*, and in regression analysis (Unit 10.2) to see if the regression residuals are random.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) – The sample or sequence of events is random.
   b) Alternative Hypothesis (H_a) - The sample or sequence of events is not random.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   *a) The Test Statistic is based on the number of data in the sequence and the number of runs present in the ordered data. If the event is binary (go no-go data), count the number of times the data changes from one value to another in the sequence (each time the data changes is called a **run**).*

   *For ordinal or continuous data, "convert" the data into two categories – data **above** the mean or median is in one category, **below** in the other category.*

   *b) Determine the Rejection Region (K_α) – For Small Samples (defined as the number of data in both categories less than or equal to 20), review the following table (here, α = 0.05) – count the number of data in each category, call the smaller number **r,** the larger number **s**. Look on the table for the **r** and **s** values (the columns are for **r**, the*

*rows are for **s**). The intersection of **r/s** is the limiting number of runs. If the total number of runs (sum of runs above/below) is smaller than the limiting value, then we can reject the null hypothesis.*

r/s	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
6	3														
7	4	4													
8	4	4	5												
9	4	5	5	6											
10	5	5	6	6	6										
11	5	5	6	6	7	7									
12	5	6	6	7	7	8	8								
13	5	6	6	7	8	8	9	9							
14	5	6	7	7	8	8	9	9	10						
15	6	6	7	8	8	9	9	10	10	11					
16	6	6	7	8	8	9	10	10	11	11	11				
17	6	7	7	8	9	9	10	10	11	11	12	12			
18	6	7	8	8	9	10	10	11	11	12	12	13	13		
19	6	7	8	8	9	10	10	11	12	12	13	13	14	14	
20	6	7	8	9	9	10	11	11	12	12	13	13	14	14	15

For Large Samples (those which don't meet the Small Sample criteria above), calculate the following statistic (courtesy of the Central Limit Theorem):

$$Z = \frac{r - \left(\frac{2n_1 n_2}{n_1 + n_2} + 1 \right)}{\sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}}}$$

where:

n_1 – number of data in first category

n_2 – number of data in second category

Critical values of the Z statistic can be found in Appendix A.

4. Collect Data and Calculate the Test Statistic as:

*a) Draw a Random Sample of size **n** from the Population or collect sequential data from the Process.*
b) Calculate the Number of Runs of the Sample Data.
c) Read or Calculate the Test Statistic shown above in Step 3.

5. Draw a Conclusion.

Example – 1 Sample Runs Test

A supplier has sent you a sample of data obtained from start-up tests of a machine they are developing for you. You wonder if the data is random.

Supplier Machine Data			
7.99	9.71	9.20	10.32
10.20	9.94	9.34	10.11
10.09	10.33	10.92	10.77
10.70	11.41	11.71	8.97
10.06	10.69	12.18	10.82

Null Hypothesis (H_o) – The sample or sequence of events is random.
Alternative Hypothesis (H_a) - The sample or sequence of events is not random.

The mean of the data is 10.274. Therefore, the "+" and "-" values associated with the data are as follows:

Supplier Machine Data			
-	-	-	+
-	-	-	-
-	+	+	+
+	+	+	-
-	+	+	+

First, there are 10 data above the mean and 10 data below the mean (hence r = s = 10). Counting the number of runs, we find that there are 5 runs below the mean and 6 runs above the mean (total runs = 11). Consulting the runs table, the limiting number of runs is 6, hence this is not unusual, and the null hypothesis cannot be rejected.

9.2.8.2 Sign Test

Purpose and Description

This test is a "crude' version of the paired-t test described earlier. We will deal with sets of paired observations (i.e. before and after customer impression of our company following a "treatment" such as an advertising campaign). All we will be able to conclude is that there is *more* of the "after" than there is of the "before."

Assumptions

1. Z_i is the difference between the "score" for one member of the paired observation (Y_i) and the other observation (X_i); that is $Z = Y_i - X_i$.

2. The model $Z_i = \Theta + e_i$ holds, where Θ is the unknown treatment effect and the e_i are mutually independent and come from a continuous distribution with median = 0 so that $P(e_i < 0) = P(e_i > 0) = ½$, for all i.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The treatment effect Θ is 0.
   b) Alternative Hypothesis (H_a) - The treatment effect Θ is greater than/less than/not equal to 0.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

4. *a) The Test Statistic for Small Sample Sizes makes use of the binomial distribution. Form the differences between the Y_i and X_i - $Z = Y_i - X_i$. If the difference is 0, drop the pair from the analysis (and reduce the sample size by 1). For the alternative hypothesis: H_a: $\Theta < 0$, count the number of negative differences (d-), for the alternative hypothesis H_a: $\Theta > 0$, count the number of positive differences (d+). For the not-equal alternative, use the number of the most frequently occurring difference (d- or d+).*

*b) Determine the Rejection Region (K_α) - Using the value of α and the binomial distribution (for p = 0.5), calculate the probability of finding d- or d+ events in a sample size of n (for the unequal alternative hypothesis, double the probabilities). If the probability is less than alpha (α), the null hypothesis may be rejected. A sample table which implements a slightly different approach follows. Here, for a given sample size **n** and a desired alpha (α), the null may be rejected if the test statistic is less than or equal to the table value. The table has been set up for 2-sided tests, for 1-sided tests, divide the column's alpha (α) in half:*

	α					α			
n	**1%**	**2%**	**5%**	**10%**	**n**	**1%**	**2%**	**5%**	**10%**
6	-	-	0	0	**21**	4	4	5	6
7	-	0	0	0	**22**	4	5	5	6
8	0	0	0	1	**23**	4	5	6	7
9	0	0	1	1	**24**	5	5	6	7
10	0	0	1	1	**25**	5	6	7	7
11	0	1	1	2	**26**	6	6	7	8
12	1	1	2	2	**27**	6	7	7	8
13	1	1	2	3	**28**	6	7	8	9
14	1	2	2	3	**29**	7	7	8	9
15	2	2	3	3	**30**	7	8	9	10
16	2	2	3	4	**31**	7	8	9	10
17	2	3	4	4	**32**	8	9	9	10
18	3	3	4	5	**33**	8	9	10	11
19	3	4	4	5	**34**	9	9	10	11
20	3	4	5	5	**35**	9	10	11	12

c) For large sample (n > 35), the normal approximation to the binomial can be used:

$$Z = \frac{K - n/2}{\sqrt{n/4}}$$

where k is the number of plus/minus signs

5. Collect Data and Calculate the Test Statistic:

   *a) Draw a Random Sample of size **n** from the Population.*
   b) Calculate the Signs of the Sample Data that has the characteristic of interest.
   c) Calculate the Test Statistic shown above in Step 3a or 3c.

6. Draw a Conclusion.

Example – Sign Test

A company is testing a new surgical instrument (to replace an existing model). 15 surgeons (familiar with the existing model) are asked to use the new instrument. They are asked to report whether the fatigue they experience using the instrument during surgery is less with the new model than the old. Twelve surgeons say they experience less fatigue, one the same amount of fatigue, and two said fatigue was higher with the new model. Is there evidence to suggest the new model reduces fatigue (alpha = .05)?

Let:

+ = fatigue higher with new instrument,
- = fatigue lower with new instrument,
0 = fatigue the same

Null Hypothesis: H_o: P(+) = P(-) = .5

Alternate Hypothesis: H_a: P(+) < P(-)

Test Significance Level – 0.05

Data:

Surgeon	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Sign	-	-	-	-	-	-	-	-	-	-	-	-	0	+	+

Rejection: Since this is a one-sided test, enter the table on the page 62 at the 10% column (associated with alpha = 0.05), and find the intersection with n = 14 (15 surgeons, one dropped for reporting “same” or sign = 0). Here, the critical value is 3 – that is if three or fewer surgeons reported a “+” with the rest reporting a “-,” then we can reject the null hypothesis. Here, there are two surgeons reporting a “+” – the null hypothesis can be rejected in favor of the alternative.

9.2.8.3 Wilcoxon Signed Ranks Test

Purpose and Description

This test is used on continuous variables, where the data can be treated as paired observations (similar to the paired t-test). You might use this test if the sample size is small ($n < 30$) and the data is drawn from a non-normal distribution. The Wilcoxon Signed Rank may also be used as a one-sample test, where each observation is compared to a hypothesized mean or median.

Assumptions

1. The model is found by obtaining $d_i = Y_i - X_i$ and then $d_i = \Theta + e_i$, where Θ is the unknown treatment effect and the e_i are mutually independent and come from a continuous distribution (not necessarily the same one) that is symmetric about 0.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population medians are equal (or $\Theta = 0$).
   b) Alternative Hypothesis (H_a) - The population medians are greater than/less than or not equal.
   In the one sample case, the hypothesis is modified to be the difference between the population median and some hypothesized μ.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is formed by first calculating the differences ($d_i = Y_i - X_i$ or $d_i = Y_i - \mu$) and taking their absolute values. The differences are then ranked from 1 to n.
   b) Add up the sum of the ranks for the positive differences (T+) and the sum of the ranks for the negative differences (T-).

c) Determine the Rejection Region (K_α) –

*1) Two-Sided Test – Reject the null hypothesis if either T+ or T- (whichever is smaller) is **smaller** than the critical value from the table on the next page for the desired alpha (table is set-up for $\alpha = 0.05$).*

*2) One-Sided Test – If the alternative hypothesis is that the median of the difference (Θ) is negative, reject the null if T+ is **less than** d from the table on the next page (for the given alpha). If the alternative hypothesis is that the median of the difference (Θ) is positive, reject the null if T- is **less than** d from the table.*

*3) Handling Ties – If one or more $d_i = Y_i - X_i = 0$, **exclude** these pairs from the test and reduce the sample size, n, accordingly.*

*4) Equal Differences – If two or more of the differences are equal (and therefore, their ranks), assign the **average of the ranks** which would have been assigned had the d_i differed slightly (see example below).*

5) Large Sample Approximation – For n > 25 calculate the critical d value as follows:

$$d = \frac{1}{2}\left[\frac{n(n+1)}{2} + 1 - Z\sqrt{\frac{n(n+1)(2n+1)}{6}}\right]$$

where Z is obtained from Appendix A - Normal Tables

Or, since T+ and T- can be approximately distributed by the normal distribution, calculate:

$$Z = \frac{T+ - \mu_{T+}}{\sigma_{T+}}, \mu_{T+} = \frac{n(n-1)}{4}, \sigma_{T+} = \frac{n(n+1)(2n+1)}{24}$$

Using the value of α, consult the Standard Normal table to determine K_α. Note that, for a given α, this value will differ based on the statement of the alternative hypothesis.

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of paired data of size **n** from the Population.*
b) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

Critical Values of Summed Ranks (α = 5%) *(n_1 – Column values, n_2 – Row Values)*

$n_2 \backslash n_1$	2	3	4	5	6	7	8	9	10	11	12	13	14	15
4			10											
5		6	11	17										
6		7	12	18	26									
7		7	13	20	27	36								
8	3	8	14	21	29	38	49							
9	3	8	15	22	31	40	51	63						
10	3	9	15	23	32	42	53	65	78					
11	4	9	16	24	34	44	55	68	81	96				
12	4	10	17	26	35	46	58	71	85	99	115			
13	4	10	18	27	37	48	60	73	88	103	119	137		
14	4	11	19	28	38	50	63	76	91	106	123	141	160	
15	4	11	20	29	40	52	65	79	94	110	127	145	164	185
16	4	12	21	31	42	54	67	82	97	114	131	150	169	
17	5	12	21	32	43	56	70	84	100	117	135	154		
18	5	13	22	33	45	58	72	87	103	121	139			
19	5	13	23	34	46	60	74	90	107	124				
20	5	14	24	35	48	62	77	93	110					
21	6	14	25	37	50	64	79	95						
22	6	15	26	38	51	66	82							
23	6	15	27	39	63	68								
24	6	16	28	40	55									
25	6	16	28	42										
26	7	17	29											
27	7	17												
28	7													

Example – Wilcoxon Signed Ranks Test

A doctor explored the effect of modifying patients' intake of fat. For seven weeks, patients ate a "standard" American diet – one rich in saturated fats and trans-fatty acids. Their cholesterol levels (LDL*) were measured during this time. The patients' diets were then changed to a "Mediterranean" diet – where the foods containing saturated fats were replaced by mono-saturated and polyunsaturated fats. Again, the cholesterol levels were monitored. The raw data appears below:

Patient	*Cholesterol Level*		*Difference*
	American Diet	*Mediterranean Diet*	
1	153	136	17
2	160	132	28
3	144	138	6
4	169	138	31
5	151	135	16
6	155	135	20
7	160	137	23
8	162	139	23
9	154	136	18
10	138	150	-12
11	169	175	-5

Notes:

- ** - LDL – Low-Density Lipoprotein – so-called "bad" cholesterol*
- *In real studies, the sample size for this type of experiment would be much higher!*

Null Hypothesis: The diet does not make a difference in cholesterol ($\Theta = 0$)
Alternative Hypothesis: The diet does make a difference in cholesterol ($\Theta \neq 0$)

Significance Level $\alpha = 0.05$

Ranked Differences:

Patient	*Difference*	*Absolute Value*	*Rank*
1	17	17	5
2	28	28	10
3	6	6	2
4	31	31	11
5	16	16	4
6	20	20	7
7	23	23	8.5*
8	23	23	8.5*
9	18	18	6
10	-12	12	3
11	-5	5	1

** - Since these are "ties," we assign an average rank – these would have been the 7th and 8th ranks; hence they are assigned the average rank of 8.5*

Sum of the ranks for the positive differences (T+) and the sum of the ranks for the negative differences (T-):

$$\textit{Positive differences (T+)} = 5 + 10 + 2 + 11 + 4 + 7 + 8.5 + 8.5 + 6 = 182$$

$$\textit{Negative differences (T-)} = 3+1 = 4$$

*Rejection Region: Find the intersection of 9 and 2 on the table (number of positive differences, number of negative differences) – this is **3**. Therefore, since the smaller of the positive/negative differences (4) is greater than the critical value of 3, the null hypothesis cannot be rejected.*

Conclusion: Despite the promising results, the physician cannot claim that the Mediterranean diet reduces cholesterol levels.

9.2.8.4 Mann-Whitney Test

Purpose and Description

This test is an alternative to the t-test for two independent groups when the data can't be assumed to be normally distributed. This is one of the most powerful of the non-parametric tests, even for moderately sized samples. The test analyzes the separation between two samples and gives the probability of obtaining the separation if both data sets are random samples from identical populations.

Assumptions

1. This test is "best" continuous or ordinal data.
2. Both samples are random and were drawn independently.
3. The samples are drawn from populations whose centers (mean or median) differ, but whose variances are equal.

Procedure

1. Establish the Hypothesis:
   a) Null Hypothesis (H_o) - The population medians are equal ($\Theta_1 = \Theta_2$).
   b) Alternative Hypothesis (H_a) - The median of population 1 (Θ_1) is greater than/less than/not equal to that of population 2 (Θ_2).

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is different for small or large sample sizes. "Large" is defined if the size of the sample drawn from either population is greater than 20:

      1) Small Sample Test Statistic – Combine the two samples and rank them from smallest to largest. Calculate S as the sum of the ranks of the observations from population 1 where population 1 is the group with the smallest sample size (n_1). If the groups are the same size, either may be designated as population 1. The test statistic is then:

$$T = S - \frac{n_1(n_1 - 1)}{2}$$

2) Large Sample Test Statistic – Use the following equation:

$$Z = \frac{T - \left(\frac{n_1 n_2}{2}\right)}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

b) Determine the Rejection Region:

1) Small Samples:

a) for H_a: $\Theta_1 \neq \Theta_2$, reject if T is sufficiently small or large – if T < W_α or > $W_{1-\alpha}$ where W_α is the critical T in Table A.5 – Two-Tailed Tests (Appendix A) and $W_{1-\alpha}$ = $n_1 n_2$ - W_α.

b) for H_a:, $\Theta_1 \leq \Theta_2$, reject for small values of T – if T < W_α where W_α is the critical T in Table A.5 – One-Tailed Tests (Appendix A).

c) for H_a: $\Theta_1 \geq \Theta_2$, reject for large values of T – if T > $W_{1-\alpha}$ where $W_{1-\alpha}$ = $n_1 n_2$ - W_α (using the One-Tailed Test Table).

2) Large Samples: Use the Normal Table in Appendix A to determine the appropriate critical value (i.e. for a H_a: $\Theta_1 \geq \Theta_2$, and alpha = 0.05, the critical value would be 1.645).

4. Collect Data and Calculate the Test Statistic:
   a) Draw Random Samples from the Populations.
   b) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

Example – Mann-Whitney Test

A manufacturer of luxury personal cleansing products wished to test a new soap concept. Group A washed with the new concept and Group B washed with the current product. They were then asked to rate their preference for the products (1 – 10 scale – 1 = don't like at all, 10 – like extremely well). The preference data appears below. We will designate Group B as Population 1 (G-B (1)), since it has the smaller sample size – 15:

Group A	8	8	8	8	8	10	9	8	8	9	9	7	8	7	8	7	10	9
Group B (G-B (1))	6	8	5	9	8	7	9	7	6	8	7	6	4	4	8			

Null Hypothesis (H_o) - The population medians are equal ($\Theta_B = \Theta_A$).
Alternative (H_a) - The median of Group B (existing product) (Θ_B) is less than that of Group A (new product) (Θ_A).

The significance level chosen for the test is 0.05.

The scores are then ranked and the sum of Group B (G-B (1)) ranks calculated. Note that we apply the average rank procedure of the Wilcoxon test here, since many of the scores are "ties:"

Score	Group	Rank	G-B (1) Ranks
4	G-B (1)	1	1.5
4	G-B (1)	2	1.5
5	G-B (1)	3	3
6	G-B (1)	4	5
6	G-B (1)	5	5
6	G-B (1)	6	5
7	G-A	7	
7	G-A	8	
7	G-A	9	
7	G-B (1)	10	9.5
7	G-B (1)	11	9.5
7	G-B (1)	12	9.5

Score	Group	Rank	G-B (1) Ranks
8	G-A	13	
8	G-A	14	
8	G-A	15	
8	G-A	16	
8	G-A	17	
8	G-A	18	
8	G-A	19	
8	G-A	20	
8	G-A	21	
8	G-B (1)	22	18.5
8	G-B (1)	23	18.5
8	G-B (1)	24	18.5

Score	Group	Rank	G-B (1) Ranks
8	G-B (1)	25	18.5
9	G-A	26	
9	G-A	27	
9	G-A	28	
9	G-A	29	
9	G-B (1)	30	28.5
9	G-B (1)	31	28.5
10	G-A	32	
10	G-A	33	
Total-G-B (1)			180.5

The test statistic is then: $T = S - \frac{n_1(n_1 - 1)}{2} = 180.5 - \frac{15(15-1)}{2} = 180.5 - 105 = 75.5$

Rejection Region: From *Table A.5 – One Tailed Tests, Appendix A*, find the intersection of $n_1 = 15$ and $n_2 = 18$, for alpha = 0.05. Here, 88 is the value of W_α. Since the alternate hypothesis is a "less than," the test statistic (75.5) is less than the critical value, and we can reject the null hypothesis.

Conclusion: The manufacturer may conclude that the new soap is preferred over the existing product. Next steps may include testing against competitive products (see Kruskal-Wallis test example below).

9.2.8.5 Kruskal-Wallis Test

Purpose and Description

So far, we have discussed tests that involve comparisons of one population parameter against a standard, or, at most, two population parameters. In Unit 10.3, we will introduce the topic of Analysis of Variance (ANOVA), a technique that will allow us to compare two or more population means. The Kruskal-Wallis (K-W) test is analogous to ANOVA. However, the K-W test relaxes ANOVA assumptions: that the data is drawn from normally distributed populations with equal variances. The K-W test may be applied when the data is ordinal or continuous.

Assumptions

1. The model for this test is that $X_{ij} = \mu + t_j + e_{ij}$, where μ is the overall mean, t_j the jth treatment effect, and e_{ij} the residual error. These latter are assumed to be mutually independent and to come from the same continuous distribution. The sum of the t_j is 0.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) - The population medians are equal: $t_1 = t_2 = \ldots t_j$.
   b) Alternative Hypothesis (H_a) - The population medians are not all equal.

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is found by ranking the $n = \sum_{j=1}^{k} n_j$ *observations from smallest to largest. Let* r_{ij} *be the rank of* X_{ij} *in this joint ranking. For j = 1, 2, . . . k, set* $R_j = \sum_{i=1}^{n_j} r_{ij}$, $\overline{R}_j = R_j / n_j$ *and* $\overline{R} = (n+1)/2$. *The Test Statistic is then:*

$$KW = \frac{12}{n(n+1)} \sum_{j=1}^{k} n_j (\overline{R}_j - \overline{R})^2$$

b) Determine the Rejection Region (K_α) – When k (the number of populations) ≥ 3, and the number of observations in each group is > 5, the KW sampling distribution is approximated by the chi-square distribution, with df = k – 1. Reject the null hypothesis if KW ≥ kw(α, n_1, n_2, n_3, . .).
c) Handling Ties: Use the average ranks as described in the Wilcoxon test and modify the KW statistic as follows:

$$KW' = KW \Big/ \left(1 - \left(\frac{\sum_{i=1}^{g} T_i}{(n^3 - n)} \right)\right)$$

where : g - number of tied groups, and $T_i = (t_i^3 - t_i)$ with t_i being the size of the tied group

d) Differences between population pairs. If the KW test statistic is significant, the conclusion is that at least one of the groups is significantly different from the others. To determine which specific differences are significant, calculate the absolute value of the differences between the $\overline{R}_i$ and compare that to the following critical value:

$$\left|\overline{R}_X - \overline{R}_Y\right| \geq Z_{\alpha / K(K-1)} \sqrt{\frac{N(N+1)}{12}\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}$$

Note that if the overall α is 0.05, this test statistic decreases the "look-up" α by K(K-1). For example, if there were K = 3 groups, 3(3-1) = 6 and the "look-up" α would be 0.05/6 ~ 0.01.

4. Collect Data and Calculate the Test Statistic:
   *a) Draw a Random Sample of size **n** from the Populations.*
   b) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

Example – *The manufacturer of luxury cleansing products has decided to see how the new product is preferred against competitive products. The design team selects two of the bestselling products on the market. Three groups are identified to wash with the soaps and provide their feedback (1 – 10 scale – 1 = don't like at all, 10 – like extremely well).*

The consumer preference data appear below:

New Product	6	8	6	6	8	9	6	10	7	9	8	10	8	10	6	8	9	7	9	7
Brand "X"	5	4	7	4	5	5	3	6	4	6	4	4	9	6	7	6	8	7	3	6
Brand "Z"	5	7	6	5	6	6	6	4	7	7	6	9	7	5	8	6	8	10	7	7

Null Hypothesis (H_o) - The population medians are equal ($\Theta_{NP} = \Theta_X = \Theta_Y$).
Alternative Hypothesis (H_a) - The medians are not equal.

The significance level chosen for the test is 0.05

The scores are then ranked and the sum of group ranks calculated. Note that we apply the average rank procedure of the Wilcoxon test here, since many of the scores are "ties:"

New Product		***Brand "X"***		***Brand "Y"***	
Score	***Rank***	***Score***	***Rank***	***Score***	***Rank***
6	22.5	3	1.5	4	5.5
6	22.5	3	1.5	5	11.5
6	22.5	4	5.5	5	11.5
6	22.5	4	5.5	5	11.5
6	22.5	4	5.5	6	22.5
7	36.5	4	5.5	6	22.5
7	36.5	4	5.5	6	22.5
7	36.5	5	11.5	6	22.5
8	46.5	5	11.5	6	22.5
8	46.5	5	11.5	6	22.5
8	46.5	6	22.5	7	36.5
8	46.5	6	22.5	7	36.5
8	46.5	6	22.5	7	36.5
9	53.5	6	22.5	7	36.5
9	53.5	6	22.5	7	36.5

	New Product		Brand "X"		Brand "Y"	
	Score	Rank	Score	Rank	Score	Rank
	9	53.5	7	36.5	7	36.5
	9	53.5	7	36.5	8	46.5
	10	58.5	7	36.5	8	46.5
	10	58.5	8	46.5	9	53.5
	10	58.5	9	53.5	10	58.5
R_j	844		387		599	
$\overline{R}_j$	42.2		19.35		29.95	
$\overline{R} = (n+1)/2 = (60+1)/2 = 30.5$						

The Test Statistic is then calculated: $KW = \frac{12}{n(n+1)} \sum_{j=1}^{k} n_j (\overline{R}_j - \overline{R})^2 = \frac{12}{20 \times 21} \times 20(136.89 + 124.32 + 0.30) = 149.4$

Since there are ties in the ranks, the KW statistic must be adjusted:

$$KW' = \frac{KW}{1 - \left(\frac{\sum_{i=1}^{g} T_i}{(n^3 - n)} \right)} = \frac{149.4}{1 - \left(\frac{6 + 210 + 210 + 4080 + 1716 + 504 + 210 + 60}{215940} \right)} = \frac{149.4}{0.968} = 154.4$$

Rejection Region: From the Chi-Square Table (Appendix A), for 3 – 1, or 2 degrees of freedom, the critical value is 5.99. The test statistic value (154.4) is greater than the critical value; reject the null in favor of the alternative hypothesis.

Conclusion: There is a difference between products. Due to the high score on the new product, the manufacturer may conclude that the new soap is preferred over the competitors' products (you can apply the differences between population pairs procedure to prove this!).

9.2.8.7 Spearman's rho Test

Purpose and Description

In Unit 10.1, *Scatter Diagrams & Correlation Analysis*, we will describe how to determine if there is a relationship between two variables – for example, between a process variable and a Critical-to-Quality characteristic. Spearman's rho is used where *rankings* of individual observations on two data are available or can be developed from the original data. For example, suppose data were collected from a consumer test of a freckle cream. The first consumer observed a 30% reduction in freckles and indicated a purchase intent (on a score of 1 – 10, 10 being "certain to buy") of 4. A second consumer's data is 50% and 6, the third is 20% and 3, etc. This data could then be ranked as follows:

Consumer	Freckles Removed (%)	Freckle Rank	Purchase Intent	Purchase Rank
1	30	2	4	2
2	50	3	6	3
3	20	1	3	1

Here the freckle and purchase ranks are in complete agreement.

Procedure

1. Establish the Hypothesis:

   a) Null Hypothesis (H_o) – There is no relationship between the ranks.
   b) Alternative Hypothesis (H_a) – There is a relationship between the ranks. The ranks may be positively correlated (as one increases the other increases), or negatively correlated (as one increases, the other decreases).

2. Choose a Significance Level (α - "alpha").

3. Plan the Test:

   a) The Test Statistic is found by first converting the original paired data (X's, Y's) into ranked pairs (R_X, R_Y). Calculate the differences $d = R_{Xi} - R_{Yi}$. Form the rho test statistic as follows:

$$r_S = 1 - \frac{6 \times \sum_{i=1}^{n} d_i^2}{n^3 - n}$$

b) Determine the Rejection Region (K_α) – When n > 30, calculate $z = r_S\sqrt{n-1}$ *and, using the value of* α*, consult the Standard Normal table to determine* K_α*. If z >* K_α*, then there is a relationship between the ranks. For n <= 30, compare the value of* r_s *to the critical value shown below. If the test statistic is greater than the critical value, you may reject the null hypothesis.*

n	α = .05	α = .01	n	α = .05	α = .01
8	.738	.881	20	.450	0.590
9	.683	.833	21	.434	0.577
10	.648	.794	22	.417	0.568
11	.623	.818	23	.399	0.562
12	.591	.780	24	.378	0.561
13	.566	.745	25	.355	0.564
14	.545	.716	26	.327	0.572
15	.525	.689	27	.296	0.586
16	.507	.666	28	.260	0.606
17	.490	.645	29	.219	0.632
18	.476	.625	30	.172	0.664
19	.462	.608			

4. Collect Data and Calculate the Test Statistic:

*a) Draw a Random Sample of size **n** from the Population.*
b) Calculate the Proportion of the Sample Data that has the characteristic of interest.
c) Calculate the Test Statistic shown above in Step 3a.

5. Draw a Conclusion.

Example – Spearman's rho Test

Let's continue the consumer test described in the introduction to this non-parametric test. Suppose the data obtained from a sample of consumers who tried the freckle cream and then provided the design team with their purchase intents. The data values have been ranked and the d_i^2 calculated for the ranks:

Consumer	Freckles Removed (%)	Freckle Rank	Purchase Intent	Purchase Rank	d_i^2
1	30	4	5	6.5	6.25
2	33	6	5	6.5	0.25
3	38	8	3	1	49
4	40	9	5	6.5	6.25
5	48	15	7	17.5	6.25
6	27	2	5	6.5	20.25
7	28	3	5	6.5	12.25
8	36	7	5	6.5	0.25
9	56	17	6	13.5	12.25
10	43	11	6	13.5	6.25
11	62	19	9	20	1
12	68	20	6	13.5	42.25
13	50	16	7	17.5	2.25
14	45	13	6	13.5	0.25
15	44	12	6	13.5	2.25
16	58	18	5	6.5	132.25
17	15	1	4	2.5	2.25
18	46	14	6	13.5	0.25
19	41	10	8	19	81
20	32	5	4	2.5	6.25

Null Hypothesis (H_o) – There is no relationship between the freckle and purchase intent ranks.
Alternative Hypothesis (H_a) – There is a relationship between the ranks.

Significance Level = 0.05

The Test Statistic is: $r_S = 1 - \frac{6 \times \sum_{i=1}^{n} d_i^2}{n^3 - n} = 1 - \frac{6 \times 389.25}{8000 - 20} = 1 - 0.293 = 0.707$

Rejection: From the table above, for n = 20 and alpha = 0.05, the critical value is 0.450. The test statistic is greater than the critical value; the null hypothesis may be rejected in favor of the alternative: ***There is a relationship between the ranks.***

9.2.9 Summary

To help you select the appropriate hypothesis test for a given situation, we will repeat all of the decision criteria presented earlier. Appendix C contains this same information in decision tree format.

Differences in the Means

Tests for One Population - Typical questions here include the following:

- Has the population mean increased or decreased since the last time we measured it?
- Does this population meet our standard target?
- Does this process change affect the mean value?

There are two tests of interest here, depending on whether we "know" the population's variance (or standard deviation):

Population Variance (σ^2) known?	Test To Use	Section
Yes	Z - Test	9.2.4.1
No	t - Test	9.2.4.2

Tests for Two Populations - Typical questions here will include the following:

- Is there a difference in the means of these two populations (i.e. from two vendors or departments)?
- If we are doing a longitudinal study, is there a difference in the before and after populations (this leads to the paired sample test)?

Paired Samples?	Population Variances known?	Population Variances Equal?	Test to Use	Section
Yes	N/A	N/A	Paired Sample t -Test	9.2.4.3
No	Yes	N/A	Two Population Z-Test	9.2.4.4
	No	Yes	Two - Population, Pooled Variance t-Test	9.2.4.5
		No	Two Population t-Test	9.2.4.6

Tests for More than Two Populations - We will cover these in the *Analysis of Variance* discussion.

Differences in the Dispersion

There are only two tests to consider here, depending on whether we are dealing with one or two populations:

Comparison	Test To Use	Section
Population Variance to a Standard	χ^2 - Test	9.2.5.1
Two Population Variances	F - Test	9.2.5.2

Differences in Proportions

Finally, here are the tests to employ if you are dealing with proportions. Again, the decision is based on whether one or two populations are being considered:

Comparison	Test To Use	Section
Population Proportion to a Standard	Z - Test	9.2.6.1
Two Population Proportions	Z - Test (2 Pop's)	9.2.6.2

Summary of Non-Parametric Tests

Hypothesis Type (Null Stated)	*Non-Parametric Test*	*"Equivalent" Parametric Test*	*Efficiency*
Data is in a random sequence	1 Sample Runs Test	None	-
No difference in mean difference	Sign Test	Paired t-test	.637
No difference in median difference	Wilcoxon Signed Rank Test	Paired t-test	.955
No difference in means of two populations	Mann-Whitney Test	2-Sample t-test	.955
No difference in medians among k independent populations	Kruskal-Wallis Test	1-Way ANOVA (See Unit 10.3)	.955
No relation between two rank-ordered variables	Spearman's rho	Pearson r (not covered here)	.912

9.3 Sampling Methods

Learning Objectives

- To be able to obtain and analyze the results of samples taken from populations.

Unit Contents

- Sampling – General
- Simple Random Sampling
- Interval (Systematic) Sampling
- Two-Stage Sampling
- Stratified Sampling
- Cluster Sampling

9.3.1 Introduction

The Statistical tools and methods of this manual take data from one or more populations and, after "processing" the data, allow us to make some inference or decision about the population. We've termed this process the *Basic Statistical Method.*

It's very rare that a study examines the *entire* population. Generally (and fortunately!), either cost or time or accuracy issues intervene. *Sampling* methods help us learn about the population, often through only a small fraction of the population's elements.

But the way we obtain these samples from the population is very important. It's easy to collect data that are biased in one way or another. Virtually all the *enumerative* statistical methods described in this manual assume that the data was obtained through a *random sampling process.* This section will describe several methods of obtaining random samples from populations.

The first sampling question asked of the statistician is "How Much Data Do I Need?" We will cover this issue here, providing you with methods to determine "How Much" for yourself.

9.3.2 Sampling - General

Before we jump into the sampling methods, let's step back and view the "sampling" issue broadly. Here are a few situations where a "sample" may help us answer a question:

- *A hospital "produces" over 30,000 Medical Records a year. The Medical Records department is interested in the accuracy of their coding process.*
- *A manufacturing plant produces 5000 plastic injection pieces a week. The Quality Control department is interested in the defect rate of the pieces.*
- *A utility maintenance department "produces" about 1000 work orders a month. The maintenance supervisor is interested in the number of work orders that were held up waiting for spare parts.*
- *A greeting card manufacturer has been experiencing delays in introducing new cards to the market. They develop about 300 new cards a year and want to improve the design cycle time.*
- *A railroad engineer wants to determine why a bearing has been failing on locomotive axles.*
- *A pharmacologist wants to determine if a new procedure will reduce the "trough" level of a certain antibiotic in sick newborn babies.*

One of your first decisions in the sampling arena is the *type* of study or question(s) you have. In the first three examples, the question raised is "*How many?"* How many records have errors, how many pieces are defective, how many work orders are held up for parts?

Dr. W. E. Deming described these studies as being of the *Enumerative* type. For these situations, you should employ some type of *Random Sampling* method (also known as *Probability Sampling* - since the probability of selecting an item is known prior to the sampling). We'll present several commonly used techniques here, including *Simple Random, Systematic, Two-Stage, Stratified,* and *Cluster Sampling*.

For the last three situations, the question is of a "*Why*" or "*How*" variety. Dr. Deming termed these *Analytic* studies. Here, data will be collected, or experiments will be designed to confirm or refute theories, perhaps obtained from a Cause and Effect diagram.

In these situations, *random* samples are not very useful. *Judgment* sampling, obtaining data from the population or process based on our theories is the way to proceed.

When we discuss *Designing Experiments* (Section 11), the concept of *randomization* will arise. This has to do with the *order* in which the experiments are conducted, not with obtaining a random sample.

Control charts and Cause/Effect analyses fall into the category of *analytic* tools. Although a subgroup taken for a control chart may appear to be a sample, it is not. The subgrouping strategy you use (i.e. *Who, what, when, where, how*) attempts to "force" the variation associated with some factor into the subgroup. If the factor's variation is strong enough, this will show up as an assignable cause signal. Here, we are using our judgment to determine what data will be collected.

Sampling Terms and Definitions

Here are several terms that will be used frequently in our sampling discussion:

Term	Definition
Target Population	A definition of the population to be studied through sampling; we generalize or *infer* our sample findings to the target population (i.e. all voters in Hanover County).
Lot	A group of items "collected" together for some reason. Today's processed invoices can be considered as a lot, likewise, a shipment of several boxes of syringes is a lot.
Element	The unit about which information is to be collected, i.e. in a marketing study, the element is a person or a family.
Frame	A list of *every* element of the population. Most frames are imperfect; they contain errors, omissions and outdated information. For example, a telephone directory would be an obvious frame for a survey of households in Fulton County. The directory, though, is not up to date, does not contain unlisted numbers and does not include households without phones and does include households with more than one telephone number.
Sampling Unit	In single stage sampling, a sampling unit is the same as an element. In multi-stage sampling, the sampling unit is composed of many elements.
Exhaustive Sampling	A sampling process where all elements of the population are selected for the sample.
Sampling Error	The difference between the *sample* (observed) value and the *true* value, resulting from chance. The maximum sampling error possible (at a given confidence level) is called the *precision* and may be expressed as a +/- percentage or in the same units as the measurement (i.e. grams, feet, etc.).
Sample	A set of elements drawn from a population. One element is not a sample; a sample is of size *n.*

Sampling – With and Without Replacement

Suppose we are examining yesterday's parts produced, stored in a bin. We want to sample the parts and measure the fraction of parts with defects. We could pull a part from the bin at random, examine it and then return it to the bin. Here, the part has been *replaced* and could be picked again. This is called **sampling with replacement** and is viewed as sampling from an *infinite* population. On the other hand, if we pick the part from the bin, examine it and then set it aside, we are conducting **sampling without replacement**. This is viewed as sampling from a *finite* population. Intuitively, as we continue to sample and "set aside" the sampled parts, there is less that we *don't* know about the population.

Finite Population Correction Factor

If we sample *without replacement*, and the sample size *n* is "large" compared to the population size, *N,* then we'll take advantage of this increased knowledge by making an adjustment to our estimate of the population *variance* (the estimate of the *mean* is not affected). The variance estimate is *reduced* by the following factor:

$$FPCF = \left(1 - \frac{n}{N}\right)$$

where:

FPCF - Finite Population Correction Factor

This factor is applied to both *discrete* and *continuous* estimates of the population variance (and, by extension, to the population mean's variance):

Continuous Characteristic Population Variance:

$$s_{FP}^2 = \left(1 - n/N\right)s^2$$

Binomial Proportion Population Variance:

$$s_{FP}^2 = \left(1 - n/N\right)\left(p(1-p)/n\right)$$

Sample Mean Variance:

$$s_{FP}^2 = \left(1 - n/N\right)s^2/n$$

A few notes on this correction factor:

- In this manual, parameter estimates are presented *without* the finite correction factor applied, unless otherwise noted.

- As *n* approaches *N*, the variance estimate approaches zero, i.e. the uncertainty in our estimates of the population parameters becomes smaller.

- As *N* approaches infinity, the correction factor approaches one. That is, the correction factor can be ignored for *infinite populations.* A rule of thumb: if *n/N* is less than 5%, the correction factor can be ignored.

- If we are dealing with a finite population and have some previous knowledge about the *population variance*, we can use the correction factor to determine how large a sample size is needed to guarantee a *sample mean variance* or standard deviation:

$$s_{\mu}^2 = \left(1 - n/N\right)s^2$$

Rearranging terms:

$$n = \frac{s^2}{s_{\mu}^2 + s^2/N}$$

This is useful if we wish to establish a certain level of precision around our estimate. For example, suppose the population size is 10,000 and its variance is 10. If we want the precision of our estimate to be +/- 2, then the sample mean variance must be less than or equal to 1 (2/2 =1)[1]. The minimum sample size required to meet this level of precision is then:

$$n = \frac{10}{1^2 + 10/10{,}000} = 9.99 = 10$$

If we wish to narrow the precision to +/- 0.5, then the sample mean variance must be less than or equal to 0.25. The minimum sample size here is:

$$n = \frac{10}{0.25^2 + 10/10{,}000} = 157.5 \approx 160$$

[1]This calculation assumes that the precision is to be established at a 95% confidence level, i.e. +/- 2 standard deviations.

9.3.3 Simple Random Sampling

Purpose

As its name implies, *Simple Random Sampling* is the simplest method of collecting a portion of the population data. A Simple Random Sample can be obtained by itself, or as part of sampling methods such as *Two-Stage, Stratified,* and *Cluster*. Simple Random Sampling belongs to the class of *probability samples*. In a probability sample, the probability of selecting any item is *known* prior to the sample being collected. For Simple Random Sampling, each item in the population has an *equal* chance of being selected.

Procedure

1. Create a numbering system for the population to be sampled. Each element must be given a unique number. This is the *frame.*

2. Select an appropriate sample size. The tool or analysis you employ will often guide you here. A sample size procedure is included below.

3. Select random numbers that can range from 1 to the highest number in your numbering system. A random number table can be used, or a random number generator (found on many calculators). For example, if the highest number in your frame is 980, then you'll want to select three-digit random numbers. Select as many random numbers as you need to meet your sample size of Step 2. If duplicate random numbers appear, or numbers higher than the highest number in your system (i.e. 995), just pick another.

4. Associate the random numbers with the items' numbers. Pick these items and measure the characteristics of interest to you.

5. Estimate the mean and its variance, or proportion population parameters. Develop confidence intervals for these parameters (see Unit 9.2 – *Hypothesis Testing*). A statement of the results should read like the following:

> *Based on a simple random sample of 400 Bills of Materials, the point estimate of the error rate is 12%, plus/minus 3% at 95% confidence.*

For Continuous Characteristics:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s_{\bar{x}}^2 = \frac{s^2}{n}$$

where:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

For Proportions (Discrete Characteristics):

$$p = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s_p^2 = \frac{p(1-p)}{n-1}$$

where:

$x_i = 1$ *if the element displays the characteristic of interest*

$x_i = 0$ *if the element does not display the characteristic of interest*

n - *sample size*

Sample Size Determination for Simple Random Samples

The purpose of sampling is to estimate some population statistic or parameter. In Unit 9.2, we introduced *confidence intervals.* This interval quantifies our uncertainty about the "true" value of the statistic or parameter.

We found that there are three factors that affect the width of the interval:

1) our desired confidence level,

2) the population's variance and

3) the size of the sample.

We'll use this information to develop a general approach to determining the minimum sample size.

As part of planning the sample, you will determine the maximum desired *sampling error (AKA precision, or error of the estimate).* For example, you may want to know a certain weight within +/- 1 gram. Or, you may want to know a certain proportion within +/- 3%. You will also set the confidence level associated with your estimate (i.e. α = 0.05, 0.01, etc.).

These first two "ingredients" are in your control. The third "ingredient" is the population parameter's standard deviation. Unfortunately, this is out of your control and may not be known when planning the sample (OK, so we'll have to estimate it somehow!).

The relationship among these three ingredients is very simple. The product of the parameter's standard deviation and the "K" value associated with the desired confidence level is equal to your desired sampling error:

$$E = K_{\alpha/2}\sigma_p$$

$$where:$$

$$E \text{ - } Sampling\ Error$$

$$K_{\alpha/2} \text{ - } Appropriate\ Table\ Value\ for\ Parameter\ \&\ Confidence\ Level$$

$$\sigma_p \text{ - } Population\ Parameter's\ Standard\ Deviation$$

Recall that the population parameter's standard deviation is usually some function of the population's standard deviation and the sample size. We'll use this relationship to determine the sample size as a function of the remaining "ingredients."

For example, if we wish to estimate the mean of a population, the standard deviation of the mean is simply:

$$\sigma_\mu = \sigma / \sqrt{n}$$

If we insert this equation into the general formula shown above and solve for *n*, we have:

$$n = K_{\alpha/2}^2 \sigma^2 / E^2$$

For example, if our desired sampling error is +/1 gram and we have some information that allows us to estimate the population standard deviation at 5 grams, and we wish to be 95% confident in our estimate, then the sample size required is:

$$n = (1.96)^2 (5)^2 / (1)^2 = 96$$

We'll employ this basic approach to estimate sample sizes for the different types of sampling described in the remainder of this section. The following table may be used to determine a minimum sample size for a simple random sample:

	Discrete Data	Continuous Data
Large Population Relative to Sample Size (n/N < 5%)	$n = \frac{K^2_{\alpha/2} p(1-p)}{E^2}$	$n = \frac{K^2_{\alpha/2} s^2}{E^2}$
Large Sample Size Relative to Population (n/N > 5%)	$n = \frac{K^2_{\alpha/2} p(1-p)}{E^2 + \frac{K^2_{\alpha/2} p(1-p)}{N}}$	$n = \frac{K^2_{\alpha/2} s^2}{E^2 + \frac{K^2_{\alpha/2} s^2}{N}}$

Where:

- n - Sample Size
- N - Population Size
- p - Fraction or percent occurrence in population (estimated)
- E - Sampling Error (precision), expressed in same measurement scale as *s* or *p*
- $K_{\alpha/2}$ - Appropriate value from normal table for confidence level $\alpha/2$
- s - Sample standard deviation

You set the confidence level α, and the desired sampling error (E). The only unknowns then are the population standard deviation (s) and the population proportion. If there is no prior knowledge of these, the following strategies may be employed:

- Take a small pilot sample (perhaps *n = 30* for continuous data, *n = 50 - 75* for discrete data) and obtain estimates of *s* or *p.*
- If a pilot sample is not feasible, for discrete data, set *p* equal to 0.5, this will maximize the sample size.
- Estimate the standard deviation from other sources. For instance, if control charts of a production process are available, $\overline{R} / d_2$ provides an estimate of the process standard deviation.

9.3.4 Interval (Systematic) Sampling

Purpose

Interval Sampling is a process by which items are selected for the sample at some regular interval. Interval Sampling is a kind of "hybrid" sampling technique. Like Simple Random Sampling, it can be used when a *lot* of items is sampled (here, *lot* refers to a number of units of a part or product). But Interval Sampling can also be used to collect data from an ongoing process. Every tenth tubing coil that is received from a vendor can be included in the Interval Sample.

Procedure

1. Identify the number of items from which a sample will be taken (N).

2. Determine the size of the sample desired (n).

3. Determine the sampling interval (k) by dividing the number of items by the sample size (k = N/n) and rounding *up.*

Note: This procedure applies when the "bunch" already exists. It can be modified slightly for collecting process data by estimating the number of items (N) to be "produced" that day, week or whatever time period is of interest.

4. Randomly select the first item in the sample between 1 and k. Call this item "j."

5. Pick items *j, (j + k), (j + 2k), (j + 3k), etc.* until you've obtained your sample size. You'll have to "cycle back" to the beginning of the item numbers sometimes to get the last sample item.

Note: Interval Sampling can lead to a distorted picture if there is any "periodicity" in your data. If the interval equals the period of the data, then the data will not be random. For example: You measure the outside temperature each day at 8:00 am. This would not be representative of the "average" daily temperature.

6. Calculate the mean, its standard deviation and/or proportion (using the same formulas employed for simple random sampling).

9.3.5 Two-Stage Sampling

Purpose

Suppose you are responsible for receiving inspection of the incoming material at your plant. A truck has just delivered 100 boxes each containing 1000 sheet metal screws and you wish to determine if they meet your specifications. In this case, you can employ a *two-stage sampling* plan. Randomly select a number (*m*) of boxes from the 100 (*M*), and then randomly select a number of screws (*n*) from each box of 1000 (*N*). You will then have measured a total of *m x n* screws. In two-stage sampling (in general, *multi-stage* sampling), the boxes are termed the *primary sampling units;* the syringes are termed the *secondary sampling units.*

Procedure

1. Identify the primary and secondary sampling units.

2. Obtain a simple random sample (size *m*) of the *M* primary units. From each of the sampled primary units, obtain a simple random sample (size *n*) of the *N* secondary units (see reference 5 if you wish to draw unequal sample sizes from each of the *M* primary units).

3. Estimate the population mean, its standard deviation and/or proportion and associated interval estimates:

For Continuous Characteristics:

$$\bar{x} = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}$$

$$s_{\bar{X}} = \sqrt{\left(1 - \frac{m}{M}\right)\frac{s_b^2}{m} + \left(1 - \frac{n}{N}\right)\frac{s_w^2}{m \times n}} \quad \textit{(FPCF Included)}$$

Here, the overall variation is broken down into the dispersion among primary units (σ_b) and the dispersion among secondary units (σ_w). Estimates of these variation components are provided below:

$$s_b^2 = \frac{1}{m-1}\sum_{i=1}^{m}(\bar{x}_i - \bar{x})^2$$

and:

$$s_w^2 = \frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2$$

The population proportion parameter and its variance are estimated as follows:

For Proportions (Discrete Characteristics):

$$p = \frac{1}{m \times n}\sum_{j=1}^{m}\sum_{i=1}^{n}x_{ij}$$

where:

$x_{ij} = 1$ *if the element displays the characteristic of interest,* $x_{ij} = 0$ *otherwise*

$$s_p^2 = \frac{N}{m(m-1)}\sum_{i=1}^{m}(p_i - \bar{p})^2 + \frac{1}{mM(n-1)}\sum_{i=1}^{m}p_i(1-p_i) \quad \textit{(FPCF Included)}$$

where:

M - *number of primary sampling units in population*

m - *number of primary units sampled*

N - *number of secondary sampling units in each primary unit*

n - *number of secondary units sampled*

p_i - *proportion associated with "ith" primary sampling unit*

Sample Size Determination for Two-Stage Sampling

How should we determine the number of primary and secondary sampling units to examine? One approach is to constrain the *total cost* of the sampling process (T). Suppose that estimates of the primary (σ_b) and secondary units' (σ_w) variation are available.

Generally, the costs associated with sampling can be divided into three components:

C_0 *- fixed cost of sampling*
C_1 *- sampling cost per primary unit*
C_2 *- sampling cost per secondary unit*

The total cost of sampling (T) is then:

$$T = C_0 + m \times C_1 + m \times n \times C_2$$

The number of secondary sampling units can be found from the following:

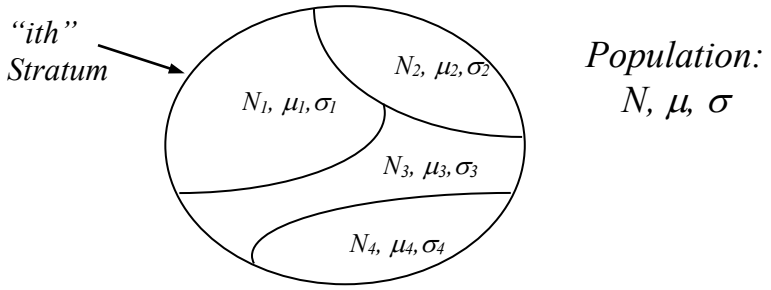
$$n = \sqrt{\frac{C_1}{C_2}} \frac{s_w}{s_b}$$

Once *n* is calculated, this value can be used in the total cost equation to determine *m.* If you wish to set the sampling error at a fixed value, first find *n* from the above equation and then use the general formula (see *Simple Random Sampling*) and the equation for the appropriate parameter's variance to determine *m.*

9.3.6 Stratified Sampling

Purpose

Often, a given population is composed of several *sub-populations* or *strata.* For example, the compressors built in one year could be divided into many different strata (type, time in service, customer, etc.). In general, we have the following picture:



Sampling from each stratum will generally provide us with a higher level of precision around our parameter estimates than would a simple random sample. There are three types of stratified samples: Proportionate, Neyman and Deming. These differ in the way the *strata* sample sizes are obtained. For each type, simple random samples are obtained from each stratum.

Proportionate Allocation Sampling

For *proportionate sampling,* the overall sample size *n* is determined and the *proportion* of the overall population calculated (*n/N*). The strata sample sizes are found (i.e. allocated) by multiplying this proportion by the number of elements in the strata:

$$n_i = \frac{n}{N} N_i$$

For example, to obtain a sample of size 100 from 1,000 compressors, the proportion applied to each strata would be 100/1,000 = 0.1. The following strata sample sizes (n_i) would then be:

Strata	Screw	Reciprocating	Total
N_i	600	400	1,000
n_i	60	40	100

Estimates of the population mean and its standard deviation are obtained as follows:

$$\bar{x} = \sum_{i=1}^{k} \bar{x}_1$$

$$s_{\bar{x}} = \sqrt{\frac{1}{n \times N}\left(1 - \frac{n}{N}\right)\sum_{i=1}^{k} N_i s_i^2} \quad \text{(FPCF Included)}$$

$$where:$$

$$\bar{x}_i \text{ - "ith" strata mean}$$

$$s_i \text{ - "ith" strata standard deviation}$$

$$n = \sum_{i=1}^{k} n_i \text{ and } N = \sum_{i=1}^{k} N_i$$

$$k \text{ - number of strata}$$

Estimates of the population proportion and its variance are provided below:

$$p = \frac{1}{N}\sum_{i=1}^{k} N_i p_i$$

$$s_p^2 = \frac{1}{N^2}\sum_{i=1}^{k} N_i^2 \frac{p_i(1-p_i)}{n_i - 1}$$

Proportionate sampling is sometimes known as *quota* sampling and there is a certain intuitive appeal to the method. However, proportionate sampling can provide biased results, especially when employed in survey sampling. One of the contributors to the "misprediction" of the 1948 election (Dewey vs. Truman) was the use of proportionate sampling by the Gallup and Roper polls.

Neyman Allocation Sampling

Neyman Sampling determines sample sizes for each stratum in such a way as to maximize the precision of the population mean's estimate. Each stratum's sample size is proportional to the product of the stratum size and the stratum's standard deviation:

$$n_i = \frac{N_i \sigma_i}{\sum_{i=1}^{k} N_i \sigma_i}$$

This method assumes that information is available to estimate the strata standard deviations. Estimates of the population mean and standard deviation are obtained the same as proportionate sampling.

Deming Allocation Sampling

The *Deming Sampling* method minimizes the cost of sampling associated with attaining a given level of precision (E). It considers the cost of investigating the individual strata (C_i):

$$n_i = \frac{\sum_{i=1}^{k} N_i \sqrt{C_i} \sigma_i}{N^2 E + \sum_{i=1}^{k} N_i \sigma_i^2} \times \frac{N_i \sigma_i}{C_i}$$

This method also assumes that information is available to estimate the strata standard deviations. Estimates of the population mean and standard deviation are obtained the same as proportionate sampling.

Sample Size Determination for Stratified Sampling

For continuous characteristics, the minimum sample size may be determined from the following:

$$n = \frac{\sum_{i=1}^{k} N_i^2 \sigma_i^2 / w_i}{\frac{N^2 E^2}{K_{\alpha/2}^2} + \sum_{i=1}^{k} N_i \sigma_i^2}$$

where:

$w_i = n_i / n$ *(this depends on the allocation procedure)*

For discrete characteristics, the minimum sample size may be determined as follows:

$$n = \frac{\sum_{i=1}^{k} N_i^2 p_i (1 - p_i) / w_i}{\frac{N^2 E^2}{K_{\alpha/2}^2} + \sum_{i=1}^{k} N_i p_i (1 - p_i)}$$

where:

$w_i = n_i / n$ *(this depends on the allocation procedure)*

9.3.7 Cluster Sampling

Purpose

Suppose part of the area served by a health care system is rural in nature, consisting of many small towns. To save sampling costs, *cluster sampling* picks a few of the towns at random and then each element (i.e. person or family) of the towns is sampled. Note that this is like two-stage sampling, with 100% sampling of the second stage.

Process

First identify all the clusters in the population. Then, select the desired number of clusters using simple random sampling. For each cluster selected, measure the desired characteristic(s) for each element of the cluster. The mean and standard deviation parameter estimates for data collected through a cluster sample are as follows:

$$\bar{x} = \sum_{i=1}^{m}\sum_{j=1}^{N_i} x_{ij} \Big/ \sum_{i=1}^{m} N_i$$

$$s_{\bar{x}} = \sqrt{\left(\frac{M-m}{M \times m \times \bar{N}^2}\right)\frac{1}{m-1}\sum_{i=1}^{m}\left(\sum_{j=1}^{N_i} x_{ij} - \bar{x} \times N_i\right)} \quad \text{(FPCF Included)}$$

where :

x_{ij} - jth element of the ith cluster

N_i - number of elements in the ith cluster

m - number of clusters sampled

M - number of clusters in the population

$\bar{N}$ - average cluster size $= \frac{1}{m}\sum_{i=1}^{m} N_i$

The proportion parameter and its variance estimates are provided below:

$$p = \sum_{i=1}^{m} a_i \Big/ \sum_{i=1}^{m} N_i$$

$$s_p^2 = \left(\frac{M-m}{Mm\overline{N}^2}\right)\frac{1}{m-1}\sum_{i=1}^{m}(a_i - pN_i)^2 \quad \text{(FPCFIncluded)}$$

$$where:$$

a_i - total of elements in ith cluster with the characteristic of interest

Sample Size Determination for Cluster Sampling

Determining the sample size here translates into determining the number of clusters to be sampled. Since every element of a cluster is measured, you will need to have some estimate of the number of elements in each cluster prior to determining the minimum sample size. In addition, some estimate of the variance will be necessary, perhaps obtained from a previous sample. For continuous characteristics, the minimum sample size (i.e. number of clusters) may be determined as follows:

$$n = \frac{M\sigma_c^2}{\frac{ME^2\overline{N}^2}{K_{\alpha/2}^2} + \sigma_c^2}$$

$$where:$$

$$\sigma_c^2 \approx s_c^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \overline{y}n_i)^2$$

$\overline{N}$ is the average cluster size

For discrete characteristics, the minimum number of clusters is found using the same formula as for continuous characteristics, however the variance is estimated as follows:

$$\sigma_c^2 \approx s_c^2 = \frac{1}{m-1}\sum_{i=1}^{m}(a_i - pn_i)^2$$

Table of Random Numbers

16408	81899	04153	53381	79401	21438	83025	92350	36693	31238	59649	91754	82772	02338	79556	04142	15387	66227	03229
18629	81953	15520	91962	04739	13092	97662	24822	94730	06496	35090	04822	86772	98289	92608	27072	17075	98204	79626
73115	35101	74798	87637	99160	71060	88824	71013	17835	20286	23153	72924	35165	43040	23982	40055	12293	14827	85636
57491	16703	23167	49323	45021	33121	12544	41035	80780	45393	44812	12515	98931	91202	09915	05908	28386	00821	18038
30405	83945	23792	14422	15059	45799	22716	19792	90983	74353	68668	30429	70735	25499	50937	46695	69927	50842	33258
16631	35006	89500	98275	32388	52390	16815	69298	82732	38480	73817	32523	41961	44437	42488	69882	34135	97525	01221
96773	20206	42559	78985	05300	22164	24269	54224	35083	19687	11052	91491	60383	19746	46764	63003	31204	40202	06486
38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318	29686	03387	59846	03237	55417	90816	88298	68335
34624	76384	17403	63365	44167	64486	64758	75366	76554	31601	12614	33072	60332	92325	86591	52667	14972	89534	14367
78919	19474	23620	27889	47917	02584	37680	20801	72152	39339	34806	09830	85001	87820	38534	94964	65680	39556	21261
03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833	25570	38818	46920	29068	16268	12856	38358	05418
74426	33278	43972	10119	89917	15665	52872	72923	73144	88662	88970	74492	51805	99378	82674	32534	27698	63863	03574
09066	00903	20795	95451	92648	45455	09552	88815	16553	51125	79375	97596	16296	66092	25835	67006	02753	22235	47539
42238	12465	87025	14367	20979	04508	64535	31355	86064	29472	47689	05974	52468	16834	96306	97901	14186	80703	61337
16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510	70625	28725	34191	33300	62247	76123	43834	99447
91227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867	14951	91696	85065	78077	61657	79180	43092	38982
50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405	56087	82790	70925	86273	93017	36692	35275	17668
65390	05224	72958	28609	81406	39147	25549	48542	42627	45223	57205	94617	23372	07896	45430	63282	17349	90183	03129
27504	96131	83944	41575	10573	08619	64482	73923	36152	05184	94142	25299	84887	34925	81482	61582	90053	76036	06177
37169	94851	39117	89632	00959	16487	65536	49071	39878	17095	02330	74301	00275	48280	01715	87288	43772	12918	29624
21457	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240	15957	16572	06004	22478	73373	88732	09443	55758
24851	57802	02050	89728	17937	37621	47057	42080	94703	48628	68995	46805	33386	21597	11951	34648	88022	56148	07785
55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88525	42786	05269	92532	35071	99704	37543	11601	65661
44657	66999	99324	51281	84463	60563	79312	94356	68876	54271	93911	25650	12682	73572	70426	75647	76310	88717	12143
91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203	71795	99533	50501	86651	70959	79425	93872	64708
82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870	52689	51275	83556	04098	73571	80799	76536	92237
21885	32906	92431	09060	62497	51642	64126	62516	26123	05125	59194	52799	28225	85762	57306	55543	53203	18096	76020
60336	98782	07408	53458	13564	59089	26445	29789	52802	41001	12535	12133	14645	23541	36600	78406	06216	95798	11977
49337	46891	24010	25560	86355	33941	25786	54990	71899	15475	98545	98227	21824	19585	49199	43716	94757	04379	46850
79556	63175	89303	16275	07100	92063	21942	18611	47346	20203	18534	03862	78095	50134	86537	62738	19636	51132	76609
11508	70225	51111	38351	19444	66499	71945	05442	13442	78675	84091	66938	93654	59894	82558	05250	09362	14556	26728
37449	30362	06694	54660	04052	53115	62757	95348	78662	11163	81651	50245	34971	52924	34925	57031	60627	53412	79924
46515	70331	85922	38329	57015	15765	97161	17869	45349	61796	66345	81073	49106	79860	35503	85171	36478	09013	02510
30986	81223	42416	58353	21532	30502	32305	86482	05174	07901	54339	58861	74818	46942	37890	40129	56548	07832	32989
63798	64995	46583	09765	44160	78128	89391	42865	92950	83531	80377	35909	81250	54238	28117	19233	16764	45174	85347

To use this table, close your eyes and put your pencil down anywhere on the table. Say you need random digits of size two. Pick the two digits next to your pencil and pick additional digits by going down, left, right, up, diagonally, any way you want.

9.4 Exercises

Probability and Statistics Foundations

Exercise - Bill of Materials Errors

A sample of 100 Bills of Materials is inspected for errors, with 12 errors being detected. What is your best estimate of the error probability?

Exercise - Heat Pump Warranties

For home heat pumps manufactured by one plant, the probability of a heat pump being produced by the first shift is 0.74. For these heat pumps, the probability of warranty failure is 0.13. For the heat pumps produced by the second shift, the probability of warranty failure is 0.08. What is the probability that a heat pump produced by the plant will have a warranty failure?

Exercise - Mutually Exclusive/Independent Events

Which of the following events are mutually exclusive? Independent?

_____ A reciprocating compressor *and* a reciprocating compressor over 10 years old.

_____ A lab report was completed within 20 minutes *and* a lab report was completed in over 40 minutes.

_____ Two brazing operations that take more than 25 minutes.

_____ Two shaft leaks in screw compressors.

_____ Two evaporators with leak rates greater than 100 cc's/hr.

_____ Sales in Europe, Asia, South America and North America

Exercise - Late Meetings

During reviews of annual plans, the probability of the first meeting of the day starting on time is 0.30. If the first meeting is late, the second meeting has a 70% chance of starting late. If the first meeting is on time, the second meeting has a 90% chance of starting on time. What is the probability that the second meeting will start *on time?*

Exercise - Populations and Samples

Which of the following describe populations? Samples? Why?

_____ 30 Bills of Materials generated each week.

_____ The 50 fan motors which failed last year.

_____ The tube sheets brazed by a worker last month.

_____ The staff trained in quality improvement methods in the last six months.

_____ The customers called for input to last month’s satisfaction survey.

_____ The Black Belts in General Electric.

Exercise - Successful Brazing

Of a sample of 100 impellers brazed with a new process, 35 passed inspection. What is the expected value of the proportion of good impellers? What is the expected value of the standard deviation?

Exercise - Set-up Time

The time to set up a particular job can be modeled using a lognormal distribution, with mean 2.3 and standard deviation 1.0. What proportion of jobs can be expected to be set up within 15 minutes?

Exercise - Chiller Repair Time

How many standard deviations away from the mean is a repair time of 10 days, given the Average repair time is 5 days and the variance is 9 days^2? What is the probability of a chiller's repair time being longer than 10 days?

Exercise - Welding Procedures

A certain welding procedure requires an average of 45 minutes to perform, with a variance of 8 minutes (normally distributed). What's the shortest time you'd expect the procedure to be performed? The longest? Answer the question using risk levels of 5% and 1%.

Exercise - Staffing Process

The Manager of a call center would like to explore some possible changes to her staffing process. Over the past year, the center has had an average call volume of 2300 calls/day, with a variance of 100 (normally distributed). Her staffing ratio is 400 calls/staff/day. Help her answer the following questions:

If she's willing to accept a risk of 5%, what's the lowest call volume she could expect? How many staff (in FTE's) does she need to cover this lowest volume?

At the same 5% risk, what's the highest call volume she could expect? How many staff (in FTE's) does she need to cover this highest volume?

A new bill scheduling system promises to cut the call volume variance in half to 50. What effect would this have on the minimum and maximum number of staff she'd need?

Exercise - Sales Order Processing

The time to complete a sales order is normally distributed with a mean of 40 minutes and a standard deviation of 8 minutes.

- What is the probability that any given sales order will require more than 50 minutes to process?
- What is the probability that it will require less than 20 minutes?
- What is the probability that it will take no more than 48 minutes?
- What is the probability that it will take between 20 and 50 minutes?

Exercise - Chiller "DOAs"

The probability of a chiller being "Dead-on-Arrival" is 20%. In a one-month time, 50 chillers were installed. What is the probability that exactly 10 "DOA's" occurred? What is the probability that less than 3 "DOA's" occurred?

Exercise - Employee Injuries

Employee injuries occur at the rate of 2/month at a certain facility. What is the probability that less than 1 employee will be injured in any given month? If 6 or more injuries occurred in one month, would this be considered “unusual?”

Exercise - Refrigerant Weight

In a manufacturing process, a fixed amount of liquid refrigerant is fed automatically into a metal container. The average weight of the bottles is 20.0 kilograms with a standard deviation of 0.05 kg. The average weight of the filled bottles is 105.0 kg, with a standard deviation of 0.5 kg. What are the mean and standard deviation of the refrigerant?

Exercise - Binomial and Poisson

Place an “B” next to the *binomial* random variables and an “P” next to the *Poisson* random variables on the list below:

_____ Phone calls to the service center that were not answered in 15 seconds.

_____ Time to complete the annual planning process.

_____ Number of scrap items each month.

_____ Errors on an expense account.

_____ Engineering reports not completed in 2 days.

_____ Eye injuries in a plant each quarter.

Exercise - Change Requests

Ten percent of change requests are not forwarded to Engineering within 24 hours of receipt. In a sample of 10 change requests, what is the probability that 3 or more are not forwarded within 24 hours?

Exercise - Service Center Calls

A service center receives calls at the rate of 10 calls per hour. What is the probability that 2 or fewer calls will be received in an hour?

Exercise - Oil Changes

A quality improvement team is investigating whether they can extend the interval between oil changes for a fleet of service trucks. They track oil condition in 17 trucks and find that, on average, the oil fouls after 6600 miles, with a standard deviation of 200 miles. If they wanted to set a “fixed” time to change the oil, what would be the maximum number of miles you’d recommend?

Exercise - Pump Failure Rate

The failure rate of water pumps can be modeled with an exponential distribution, with a λ = 0.001/hr. What is the probability that a new pump will operate for 1500 hours without failure?

Exercise - Drug Testing

A company involved in sensitive military work employs a random drug-testing program. Their program is set up at a 100% sampling rate/year. The following data indicates the number of times employees were selected for drug testing in a one-year period:

Number of times an Individual was selected	Frequency of Occurrence
0	771
1	902
2	375
3	131
4	31
5	9
6	1
Total	2220

Develop a frequency chart for this data. What is the average number of times an individual was selected? Why are so many employees not selected at all if the sampling rate is 100%/year? Is there something wrong with the random selection process?

Detecting Changes and Differences

Exercise - Failed Lights

Of 100 fluorescent lights in a factory, seven were found to be not working. Construct a 95% confidence interval for the percentage of all burned out fluorescent lights in the plant.

Exercise - Dimensions Measurement

A new, less expensive method for QC to check dimensions has been developed. The QC manager is willing to recommend the new method if it can be shown that the new method is as accurate as the more expensive one. Fifteen components are picked to compare the methods, with the data shown below. At a 2% level of significance, what recommendation should the QC manager make?

Component	Current Method Dimension (mm)	New Method Dimension (mm)
1	110	113
2	105	102
3	120	123
4	114	114
5	118	121
6	104	106
7	95	93
8	98	95
9	102	105
10	96	101
11	114	117
12	112	114
13	99	96
14	106	109
15	116	121

Exercise - Fire Department Accidents

A fire department has kept records of the number of accidents their fire trucks are involved in while responding to a fire alarm. One of the firemen thinks that the color of the fire truck makes a difference in the accident rate. Test the following data at the 5% level of significance:

	Red Fire Trucks	**Yellow Fire Trucks**
Number of Accidents	20	4
Number of Trips	153348	135035

Exercise - Difference of Proportions

Changes were made to the process of expanding tubes in tube sheets in hopes of reducing leaks. The previous leak rate was 3%. Over the next 100 tubes, only one leaked. At the 5% level of significance, has there been a change in the leak rate?

Exercise - Difference of Standard Deviations

A manufacturer of compressor shaft bearings claims that the standard deviation in the expected life is 0.5 years and thus it is easy to plan for maintenance and replacement. A random sample of 20 bearings was tested, and the standard deviation of the sample is 0.65 years. At a 5% level of significance, test whether the standard deviation of this product is greater than 0.5 years.

Exercise - Difference of Proportions

Accounts Payable tracks on time and late invoices. The quality department was informed that most of their bills (150 of 200) were paid late last month. Quality investigated the payment process and made several process changes. This month's results were 26 of 130 invoices. Has the proportion of late invoices decreased?

Exercise - Proportions Confidence Interval

In a sample of 100 electronic boards shipped by a vendor, seven were found to be not working. Construct a 95% confidence interval for the percentage of all failed boards from the vendor.

Exercise - Mean Confidence Interval

The average manufacturing cost of a compressor built last year was $8,200. To estimate the average cost this year, an analyst randomly selects 100 jobs and calculates an average of $9,000, with a standard deviation of $200. Calculate a 90% confidence interval for the average cost of the procedure.

Exercise - Difference of Means, Standard Deviation Unknown

An Education Director is interested in whether a computerized training program will improve the test scores for Black Belts (BB's). 25 BBs receive the traditional classroom training and 25 are trained through the computer approach. Given the following test results, is there a difference in the two approaches?

	Classroom Training	Computer Training
Average -	78	85
Standard Deviation -	6	5
Number of NAs -	25	25

Objective: To have fun with hypothesis tests.

Instructions:

1. The following data were collected from a torquing operation. Six spindles of a multi-driver tighten bolts on an assembly. The assembly is then dehydrated and the torque again measured. What questions could you generate about this process? What answers might come from performing hypothesis tests of the data? Try to answer your questions.

2. On the next page, similar data is listed. This data was collected from a machine being evaluated to replace the current multi-driver. Repeat question one on this data.

Time: 20 minutes

H29BU Thick Valve Plate Gasket, .034"

Torque Target: 25 ft-lb.

Machine A Line Multi-Driver																															
Assembly	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***	***11***	***12***	***13***	***14***	***15***	***16***	***17***	***18***	***19***	***20***	***21***	***22***	***23***	***24***	***25***	***26***	***27***	***28***	***29***	***30***	***31***	***32***
Before Dehydration																																
Spindle-1	21.0	21.0	21.0	22.0	21.0	21.0	20.0	20.0	20.0	19.0	20.0	22.0	20.0	20.0	21.0	22.0	18.0	21.0	21.0	21.0	20.0	20.0	21.0	18.0	20.0	21.0	20.0	21.0	20.0	25.0	23.0	20.0
Spindle-2	23.0	22.0	24.0	23.0	22.0	22.0	20.0	21.0	20.0	24.0	23.0	23.0	24.0	20.0	23.0	24.0	21.0	18.0	26.0	26.0	20.0	22.0	19.0	22.0	25.0	22.0	26.0	24.0	20.0	25.0	22.0	21.0
Spindle-3	24.0	21.0	23.0	21.0	23.0	20.0	20.0	21.0	20.0	20.0	21.0	22.0	23.0	20.0	23.0	22.0	20.0	20.0	22.0	22.0	19.0	22.0	22.0	22.0	22.0	23.0	21.0	23.0	20.0	25.0	22.0	22.0
Spindle-4	21.0	22.0	22.0	20.0	21.0	20.0	20.0	23.0	21.0	19.0	21.0	21.0	20.0	19.0	22.0	21.0	19.0	19.0	21.0	21.0	20.0	21.0	20.0	21.0	20.0	22.0	22.0	20.0	20.0	25.0	20.0	23.0
Spindle-5	21.0	23.0	23.0	21.0	23.0	20.0	23.0	19.0	22.0	20.0	21.0	21.0	20.0	21.0	22.0	19.0	21.0	19.0	21.0	20.0	19.0	18.0	20.0	20.0	20.0	19.0	20.0	15.0	22.0	25.0	21.0	20.0
Spindle-6	21.0	22.0	23.0	19.0	21.0	20.0	19.0	22.0	20.0	20.0	21.0	20.0	20.0	10.0	21.0	18.0	19.0	19.0	18.0	28.0	19.0	24.0	20.0	19.0	20.0	23.0	25.0	24.0	21.0	25.0	18.0	25.0
After Dehydration																																
Spindle-1	10.0	10.0	10.0	10.0	10.0	10.0	8.0	11.0	8.0	10.0	9.0	10.0	9.0	10.0	10.0	12.0	9.0	10.0	10.0	12.0	10.0	8.0	10.0	8.0	9.0	10.0	10.0	10.0	10.0	13.0	12.0	10.0
Spindle-2	13.0	12.0	13.0	12.0	13.0	11.0	7.0	12.0	11.0	14.0	16.0	13.0	16.0	9.0	12.0	15.0	10.0	6.0	19.0	13.0	13.0	10.0	8.0	12.0	18.0	12.0	15.0	14.0	8.0	14.0	13.0	10.0
Spindle-3	15.0	8.0	11.0	7.0	16.0	8.0	6.0	14.0	11.0	9.0	6.0	10.0	8.0	8.0	12.0	10.0	9.0	10.0	10.0	11.0	9.0	10.0	13.0	10.0	9.0	10.0	9.0	10.0	10.0	12.0	12.0	11.0
Spindle-4	11.0	15.0	13.0	10.0	10.0	11.0	10.0	15.0	11.0	7.0	10.0	11.0	10.0	10.0	14.0	14.0	11.0	12.0	8.0	11.0	13.0	12.0	12.0	14.0	10.0	14.0	15.0	10.0	11.0	17.0	10.0	15.0
Spindle-5	14.0	17.0	15.0	11.0	16.0	11.0	18.0	9.0	16.0	10.0	13.0	12.0	11.0	12.0	14.0	10.0	14.0	12.0	14.0	12.0	9.0	7.0	11.0	10.0	11.0	10.0	12.0	6.0	15.0	14.0	15.0	12.0
Spindle-6	10.0	12.0	15.0	8.0	10.0	8.0	8.0	12.0	8.0	8.0	11.0	8.0	9.0	7.0	12.0	8.0	10.0	8.0	7.0	15.0	8.0	16.0	11.0	8.0	11.0	15.0	15.0	14.0	14.0	16.0	10.0	20.0

Machine AAG #800864-05

Assembly	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*	*26*	*27*	*28*	*29*	*30*	*31*	*32*
Before Dehydration																																
Spindle-1	25.0	25.0	25.0	24.0	25.0	25.0	26.0	25.5	26.0	24.5	25.0	25.5	25.0	26.5	25.0	25.0	25.5	26.0	25.0	25.0	25.5	26.0	25.5	25.0	26.0	26.0	26.0	25.5	26.0	23.5	25.5	25.5
Spindle-2	26.5	26.5	26.0	26.0	26.0	25.5	26.0	26.5	26.0	26.0	27.0	27.0	24.5	27.0	28.0	26.5	27.5	25.5	26.5	27.5	28.0	27.0	27.0	26.0	27.0	28.0	27.5	26.0	28.0	27.0	28.0	26.5
Spindle-3	24.0	27.0	26.0	24.5	25.0	25.5	25.5	25.5	25.5	24.0	25.5	25.5	25.5	25.5	25.0	26.0	26.0	25.0	25.0	26.5	26.0	25.0	26.0	25.5	26.5	26.0	26.5	26.5	26.5	25.0	26.0	25.0
Spindle-4	24.5	25.0	25.0	25.0	24.5	24.5	25.0	25.5	25.5	25.0	25.5	26.0	25.5	25.0	25.0	25.5	26.0	24.5	26.0	26.0	26.0	25.5	26.0	26.0	25.0	26.0	27.0	25.5	26.0	25.5	26.0	26.5
Spindle-5	24.5	24.0	25.0	24.0	25.0	24.5	25.0	25.5	24.0	24.5	24.5	24.0	24.0	24.0	24.0	26.0	25.0	25.0	25.0	25.0	25.0	25.5	25.0	25.0	25.0	25.5	24.0	25.5	26.0	25.0	25.0	24.5
Spindle-6	24.0	25.0	25.5	25.0	25.0	25.0	25.0	25.0	25.0	25.0	25.5	24.5	25.0	25.0	25.0	25.0	25.5	26.0	26.0	25.0	25.5	25.0	25.5	25.5	26.0	26.0	25.0	26.0	24.5	26.0	26.0	25.5
After Dehydration																																
Spindle-1	18.0	20.0	18.0	23.0	17.0	19.0	20.0	15.0	22.0	23.0	15.0	20.0	21.0	20.0	22.0	19.0	18.0	19.0	17.0	21.0	20.0	19.0	23.0	22.0	18.0	21.0	17.0	18.0	19.0	24.0	16.0	18.0
Spindle-2	20.0	21.0	25.0	21.0	22.0	23.0	23.0	18.0	18.0	22.0	18.0	22.0	22.0	22.0	24.0	21.0	21.0	24.0	18.0	22.0	23.0	24.0	24.0	22.0	19.0	20.0	18.0	23.0	23.0	23.0	20.0	23.0
Spindle-3	22.0	20.0	20.0	23.0	21.0	18.0	15.0	16.0	18.0	21.0	18.0	20.0	19.0	18.0	16.0	18.0	19.0	20.0	17.0	21.0	19.0	16.0	18.0	20.0	16.0	18.0	17.0	19.0	19.0	22.0	18.0	17.0
Spindle-4	20.0	20.0	20.0	19.0	20.0	19.0	20.0	18.0	21.0	22.0	21.0	22.0	21.0	20.0	19.0	23.0	20.0	18.0	21.0	22.0	21.0	21.0	21.0	21.0	19.0	20.0	21.0	19.0	23.0	21.0	20.0	22.0
Spindle-5	20.0	20.0	21.0	19.0	22.0	20.0	18.0	19.0	16.0	18.0	20.0	21.0	20.0	20.0	19.0	20.0	20.0	19.0	18.0	20.0	20.0	22.0	28.0	19.0	20.0	18.0	20.0	20.0	22.0	21.0	20.0	20.0
Spindle-6	23.0	20.0	21.0	23.0	23.0	21.0	20.0	19.0	22.0	21.0	25.0	22.0	21.0	22.0	22.0	20.0	23.0	20.0	23.0	22.0	22.0	21.0	22.0	20.0	21.0	22.0	21.0	20.0	23.0	23.0	23.0	22.0

Objective: To have fun with hypothesis tests.

Instructions:

1. The following data were collected from a brazing operation. Two models are built on the line, the BU's and the B's. The plant runs four shifts (A – D). What questions could you generate about this process? What answers might come from performing hypothesis tests on the data? Try to answer your questions.

Time: 20 minutes

Models Built				*Leaks per Shift*					*Total*		*Leaks per Model*			
BU's	*B's*	*Total*		*A*	*B*	*C*	*D*	*Total*	*%*		*BU's*	*%*	*B's*	*%*
4158	11180	15338		39	9	0	0	48	0.31%		21	0.51%	27	0.24%
13850	33918	47768		13	5	25	21	64	0.13%		32	0.23%	32	0.09%
2062	40588	42650		1	5	22	19	47	0.11%		4	0.19%	43	0.11%
2076	36281	38357		13	17	8	4	42	0.11%		11	0.53%	31	0.09%
9563	32883	42446		41	25	26	14	106	0.25%		49	0.51%	57	0.17%
21054	25721	46775		33	2	35	27	97	0.21%		73	0.35%	24	0.09%
33026	8675	41701		26	3	73	46	148	0.35%		143	0.43%	5	0.06%
32388	9334	41722		6	3	80	26	115	0.28%		99	0.31%	16	0.17%
14040	25209	39249		42	15	7	3	67	0.17%		30	0.21%	37	0.15%
8807	34242	43049		44	9	12	3	68	0.16%		33	0.37%	35	0.10%
11092	35921	47013		24	0	15	19	58	0.12%		22	0.20%	36	0.10%
15614	28132	43746		14	19	27	29	89	0.20%		50	0.32%	39	0.14%
18267	20769	39036		72	15	5	13	105	0.27%		83	0.45%	22	0.11%
19659	18938	38597		56	36	1	6	99	0.26%		96	0.49%	3	0.02%
19566	23045	42611		58	25	6	14	103	0.24%		75	0.38%	28	0.12%
21278	27357	48635		8	1	78	25	112	0.23%		84	0.39%	28	0.10%
10408	30585	40993		24	3	9	13	49	0.12%		27	0.26%	22	0.07%
6200	33422	39622		42	16	16	11	85	0.21%		64	1.03%	21	0.06%
8865	33403	42268		25	12	9	8	54	0.13%		31	0.35%	23	0.07%

Models Built				Leaks per Shift							Leaks per Model			
BU's	B's	Total		A	B	C	D	Total	%		BU's	%	B's	%
8506	36660	45166		1	3	55	19	78	0.17%		40	0.47%	38	0.10%
14180	31391	45571		11	5	36	27	79	0.17%		45	0.32%	34	0.11%
16351	27049	43400		25	6	47	36	114	0.26%		54	0.33%	60	0.22%
4888	26321	31209		13	11	0	19	43	0.14%		8	0.16%	35	0.13%
9401	31168	40569		20	22	20	37	99	0.24%		25	0.27%	74	0.24%
13647	30139	43786		15	2	97	22	136	0.31%		68	0.50%	68	0.23%
12761	35060	47821		10	0	54	43	107	0.22%		52	0.41%	55	0.16%
12187	29000	41187		40	7	19	5	71	0.17%		30	0.25%	41	0.14%
12807	26991	39798		84	34	24	15	157	0.39%		93	0.73%	54	0.20%
10984	18146	29130		51	16	12	0	79	0.27%		41	0.37%	38	0.21%
3078	9197	12275		0	0	16	0	16	0.13%		10	0.32%	6	0.07%
237	12973	13210		0	0	21	5	26	0.20%		0	0.00%	26	0.20%
0	10355	10355		10	8	0	0	18	0.17%		0	#DIV/0!	18	0.17%
9	10066	10075		0	11	0	0	11	0.11%		0	0.00%	11	0.11%
5437	7429	12866		0	0	7	21	28	0.22%		25	0.46%	3	0.04%
5779	5950	11729		0	0	11	9	20	0.17%		15	0.26%	5	0.08%

Non-Parametric Tests

One Sample Runs Test

A shipment of pole-mounted transformers has been delivered to a district office of an electric utility. Core loss measurements were made and the data appears below. Does the data appear to have arisen from a random (or common cause) system?

Transformer	*Core Loss*		*Transformer*	*Core Loss*
1	114		11	117
2	114		12	122
3	116		13	122
4	112		14	123
5	109		15	119
6	112		16	117
7	110		17	118
8	110		18	121
9	115		19	113
10	114		20	108

Sign Tests

Here are some sign test scenarios for you to evaluate:

1. A marketing representative interviewed 10 homeowners who recently installed a new attic insulation system. They were asked if their utility bills were lower after the installation. Seven said “less,” two said “higher,” and one said, “no change.” Is the insulation effective in reducing utility bills?

2. A new skin cream is being tested by consumers. Fifty female consumers between the ages of 50 and 60 were asked to apply the cream in a mall study. Thirty-five said the cream made their skin feel smoother, 10 reported no difference and 5 reported the cream made their skin feel rougher. Is the new skin cream effective at providing a “smoother” feeling?

3. A new “cafeteria” style medical benefits plan was implemented in a company. One hundred employees were surveyed to see if their perceived out-of-pocket expenses were reduced with the new plan. Seventy-five said their expenses were reduced, 10 said there was no change and 15 reported a perceived increase in expenses. Is the benefits plan effective at reducing employee health expenses?

Wilcoxon Signed Rank Test

Here are some Signed Rank Test scenarios for you to evaluate:

1. Steel reinforcing rods are known to have a historical median length of 10 meters. A sample of 10 rods delivered to a job site yields the lengths below. Is the rod vendor "shorting" the customer?

9.83	10.09	9.72	9.87	10.04	9.95	9.82	9.73	9.79	9.90

2. A movie producer is testing the "ending" for a new action film. A group watches the first part of the movie. Ending 1 is then shown and their preference is recorded (1 – 10 Scale – 10 being "Two-Thumbs Up"). Ending 2 is then shown and their preference is again recorded. Which ending should the producer use?

Ending 1	Ending 2
5	8
6	6
6	9
6	8
6	7
8	8
6	7
6	8
8	8
5	7
6	6
7	7
6	8
8	7
8	8
8	9
6	7
7	6
7	9
5	9

Mann-Whitney Test

A researcher wishes to determine if there is any difference in attitudes toward a senior drug-benefit program, where party affiliation is a factor of interest. A random sample of Democrats and Republicans is obtained; they are asked how much they favor the program (1 – 10 scale – 1 being "not favor at all," 10 being "totally in favor of"). The researcher assumes that Republicans will be less inclined to favor the program. Evaluate this data.

Republican	Democrat
3	4
2	5
7	6
2	8
6	9
2	5
6	5
3	6
4	5
2	7
3	8
5	5
4	7
6	6
4	7
3	6
6	6
7	6
5	6
6	7
6	8
3	9
5	7
7	5
4	6

Mann Whitney Test

A consumer products company is testing new combinations of product & service to reduce sensitivity to stress situations. Twenty-two subjects are independently and randomly sorted into two groups, the first of size n_a = 11 and the second of size n_b = 11. The members of the first group each individually receive Combination A over a period of 15 weeks, while those of the second group receive Combination B. The investigators' directional hypothesis is that Treatment A will prove the more effective.

At the end of the experimental treatment period, the subjects are individually placed in a series of stress test situations, knowing that their reactions to these situations are being recorded on videotape. Subsequently three clinical experts, uninvolved in the experimental treatment and not knowing which subject received which treatment, independently view the videotapes and rate each subject according to the degree of stress tendency shown in the test situations. Each expert's rating takes place along a 10-point scale, with 1="very low" and 10="very high"; and the final measure for each subject is the simple average of the ratings of the three experts for that subject. The following table shows the average ratings for each subject in each of the two groups. Is there a difference in stress level?

Group A			Group B		
4.6	4.7	4.9	5.2	5.3	5.4
5.1	5.2	5.5	5.6	6.2	6.3
5.8	6.1	6.5	6.8	7.7	8.0
6.5	7.2		8.1	7.6	

Kruskal-Wallis Test

The internal auditing department is investigating unauthorized use of the Internet by employees. They select a random sample of seven employees from three departments. They review computer logs of web sites visited and classify the web sites as “authorized” or “un-authorized.” The following represents one week’s worth of data collected:

	Un-Authorized Web Site Hits		
Employee	**Department A**	**Department B**	**Department C**
1	40	73	51
2	54	60	46
3	44	63	44
4	47	73	62
5	62	64	55
6	42	57	45
7	60	65	62

Is there a difference in unauthorized web site use by department?

Spearman's rho

An amusement park manager wondered if the satisfaction score on customer surveys was related to the number of rides a customer took during the day. The following data was collected:

Rides	Satisfaction
14	71
8	66
7	65
12	71
12	77
14	78
8	67
10	50
12	74
12	65
10	71
12	68
9	56
13	78
15	75
9	68
9	49
13	83
15	71
13	78

Sampling Methods

Exercise - Enumerative and Analytic

What is the difference between an *enumerative* study and an *analytic* study?

Exercise - Black Belt Survey

Instructors of the Black Belt Tools course are designing a survey to be completed by class participants to evaluate the course content and instruction. List some possible subgroups that may be of interest when the survey results are analyzed.

Exercise - Budget Prep Time

The Strategic Planning VP is planning a survey to assess the time spent to prepare the business units' annual budget. From the BU's phone book, the VP randomly selects 20 departments and sends a survey to the manager of each department. What are the element, sampling unit, universe and frame for this survey?

Exercise - Survey Comparison

Which is the better of the following two surveys? Why?

	Survey # 1	Survey # 2
Total # Customers -	300,000	300,000
# Surveys Sent Out -	1,000	100
# Surveys Returned -	100	80

Exercise - Customer Sample

You send the following survey to a sample of 100 customers:

Will you order air conditioning equipment next year? _____ *Yes* _____ *No*

Everyone responds to your survey. Forty customers respond “yes,” sixty respond “no.” Estimate the % of all customers who will order new equipment next year. Calculate the sampling error at 95% confidence level. Interpret the sampling error.

Exercise - Neighborhood Sample

Describe how you would obtain a simple random sample of people living in your neighborhood. Define the population and a frame for the population. Suppose this random sample was associated with a questionnaire survey. How would you obtain the elements of the sample?

Exercise - Employee Opinions

Suppose you wish to identify employee opinions about your company's Black Belt Process. You are interested in identifying differences in opinion by *department* and whether the employee is *management* or *staff.* Develop a method of sampling the employees that will achieve these goals.

Exercise - Counting "F's"

How many "f's" appear in the following paragraph? You have 15 seconds to count the "f's." Have several other people count the "f's" and compare notes. What does this tell you about the ability of 100% inspection to detect defects?

> THE NECESSITY OF TRAINING FARM HANDS FOR FIRST CLASS FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK IS FOREMOST IN THE MINDS OF FARM OWNERS. SINCE THE FOREFATHERS OF THE FARM OWNERS TRAINED THE FARM HANDS FOR FIRST CLASS FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK, THE FARM OWNERS FEEL THEY SHOULD CARRY ON WITH THE FAMILY TRADITION OF TRAINING FARM HANDS OF FIRST CLASS FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK BECAUSE THEY BELIEVE IT IS THE BASIS OF GOOD FUNDAMENTAL FARM MANAGEMENT. FURTHERMORE, THE FUTURE FARMERS OF FINLAND PROVIDE FREE AND FIRST CLASS TRAINING FOR FUTURE FARM HANDS IN THE FINICKY FANTASTIC PECADILLOS OF FUTURE FARM ANIMALS.

Exercise - Receiving Report Errors

The supervisor of the Materials Management department suspects that the clerk who processes receiving reports has an extremely high error rate. The clerk has processed 2000 receiving reports in the past year. They are all in a file cabinet in date order. Using a simple random sampling technique, select a sample of 100 reports to be examined. How would you select a sample using interval sampling? Which method would you choose for this problem?

Exercise - Fitness Center Usage

In a particular facility, 62% of the employees are non-management males, 31% are non-management females and the remaining 7% are at the level of supervisor or above. Your task is to estimate the percentage that would use a proposed fitness center. An “educated” guess is that 40-50% of the non-management females, 20-30% of the non-management males and 5-10% of the management would use the facilities. The total employee population is 3000.

What overall sample size would you need to provide a 95% confidence level with a precision of 5%?

How would you allocate that sample among the strata, using proportional allocation?

Exercise - Control Circuits

A manufacturer of control circuits ships your company 100 boxes of circuit boards, each containing 100 boards. Suppose you are interested in measuring the input resistance of the boards (in ohms). Develop an approach to sample from these boxes that will provide you with the average resistance.

Exercise - Warranty Records

A set of 20,000 Warranty Records is stored in 400 file drawers, each containing 50 records. In a two-stage sample, five records are drawn at random from each of 80 randomly selected drawers. The between cluster variance is 362 and the within-cluster variance is 805. Compute the standard error of the mean per record.

Exercise - Wrong Addresses

A quality team suspects that delays in receiving inter-office mail are due to wrong addresses on the envelopes. The team randomly samples 300 pieces of mail out of 900 delivered in a week. Of the 300, 75 are incorrectly addressed. From this data, calculate the percent occurrence. Calculate the standard deviation of the estimate of *p.* (Use the finite population correction factor)

Exercise - Lifestyle Survey

Suppose you wish to conduct a survey to determine the percentage of employees currently enrolled in a particular health plan. You suspect that "lifestyle status" is an important stratification variable; that is single employees are expected to participate less than are married employees with two or more children. You collect the following stratified sample, using proportionate sampling:

Stratum	N_h	W_h	n_h	r_h	p_h
Single	3,000	0.21	150	30	0.20
Married, 0 or 1 child	6,000	0.43	300	135	0.45
Married, 2 + children	5,000	0.36	250	200	0.80
Total	14,000	1.00	700	365	

The $\mathbf{n_h}$ column indicates the number of employees sampled, the $\mathbf{r_h}$ column is the number of employees in that stratum participating in the health plan and the $\mathbf{p_h}$ column are the proportion in each stratum participating.

Calculate a) the overall sample proportion of employees participating in the health plan, b) the variance of the proportion, c) the standard error of the proportion, and d) a 95% confidence interval for the proportion.

Exercise - Community Programs

In the last year, there have been 78 community groups (each composed of 24 members) who have listened to talks by the company's public relations director. The director would like to assess, through personal interview, the attitude of these clubs towards the company's community programs. The primary variable of interest is the *proportion of those with a favorable attitude toward corporate community programs.* She decides on a cluster sample, since she realizes that each of the community groups are scattered geographically, and she wishes to reduce the cost of travel and field work time.

She wishes to establish a 95% confidence interval on the proportion, with a 9% error of the estimate. She decides to sample 11 groups. Do you agree with this number (Assume a between cluster variance (s_b^2) of 15.55)? She collects the following 11 clusters:

Group	A	B	C	D	E	F	G	H	I	J	K
# Favorable	9	11	13	15	16	17	18	20	20	21	16
# in Group	24	24	24	24	24	24	24	24	24	24	24

Calculate a) the overall proportion favorable, b) the variance of this estimate, and c) the 95% confidence interval.

Exercise – Pick the "Best" Hypothesis Test

Determine the best hypothesis test to employ for the following situations. Use the decision flowchart in Appendix C to help work through the decision.

Scenario	***H-Test***
1. A political candidate commissioned a survey to determine which of two tax plans is preferred by voters.	
2. A pharmaceutical company tested a new drug to see if it could reduce the occurrence of acne in teenagers better than the current product on the market.	
3. A business tested two processes to see which would result in the lower average cycle time.	
4. A manufacturer of air conditioners developed a new procedure to reduce brazing leaks in condensers. They want to determine if the new procedure is better than the current one.	
5. A power plant monitored buildup of corrosion products on the inside of heat exchanger tubes. They measure a sample of tubes each outage.	
6. A car painting process is trying to reduce the number of dust particles trapped in the paint. After changing the air filtration system, they wish to determine if the change has improved the situation.	
7. An agricultural chemicals company experienced problems with variation in the content of raw materials. Their supplier agreed to tighten material specifications. The chemicals company wants to determine if the content variation has been reduced.	
8. A bandage manufacturer is designing a new bandage that is designed to require less force to peal away from the skin. They want to compare the current prototype to the best-in-class competition.	
9. A hospital is working to reduce the proportion of medications administered in error. A robot has been installed in the pharmacy to improve the process of filling prescriptions without error.	
10. An insurance company has been trying to reduce the rate of fraudulent claims. They have changed the claims process and are trying to determine if this has improved the situation.	
11. A manufacturer of anti-fouling paint for boats claims that their paint lasts longer than competitors. How could they demonstrate this claim?	

10.0 Relationships Between Variables

Unit	*Description*	*Page*
10.1	Scatter Diagrams and Correlation Analysis	10.1 - 1
10.2	Regression Analysis	10.2 - 1
10.3	Analysis of Variance (ANOVA)	10.3 – 1
10.4	Exercises	10.4 - 1

We often wish to examine the *relationship* between variables. Establishing that a process factor or variable has a relationship to the quality characteristic is a key piece of a Cause-and-Effect Analysis.

Depending on the *type* of variables, there are different tools and analysis methods employed:

Effect / Cause	Discrete	Continuous
Discrete	Tool: *Contingency Table* Analysis: χ^2 *or Z Test, Signal-to-Noise*	Tool: *Histograms* Analysis: *Z, t, Tests (Tests of Means), ANOVA, Signal-to-Noise*
Continuous	Tool: *Histograms* Analysis: *Logistic Regression, Signal-to-Noise*	Tool: *Scatter Diagram* Analysis: *Correlation/ Regression Analysis, ANOVA, Signal-to-Noise*

Section 7, *Detecting Differences,* describes several Tests of Means and the Contingency Table is presented in Section 6. These tests allow you to determine if a change in a *discrete* factor results in a change in the quality characteristic or effect. They are also used to determine if a "step" change in a *continuous* cause results in a change in a *continuous* effect.

This section will explore the use of *Correlation/Regression Analysis,* and *Analysis of Variance (ANOVA).* For these methods, we will present the concept, process and calculations for "simple" cases. As your problems and analyses become more complex, the calculations become more challenging. Fortunately, today we have statistical analysis programs available, such as Minitab that can make short work of the "number-crunching" required. We'll survey the more complex cases and leave the calculations to the computer.

The *Signal-to-Noise* analysis method will be presented in Section 10 as part of the discussion on Taguchi Methods.

One note on the terms used in this section. In the quality literature, you'll find that several different terms are applied to essentially the same concept:

Concept	**Statistical Term**	**Quality Term**
The variable that we think *influences* the effect or output and that can be changed or adjusted.	Independent Variable	Cause, Factor, Process Variable
The output or effect of the process or the aspect of the population we are examining; it is not possible to change or adjust directly.	Dependent Variable	Effect, Quality Characteristic

The statistical terms are applied a bit more broadly than the quality terms. For example, two variables may be *correlated*, but not related through a cause-and-effect relationship. We would still set one of the variables to be the *independent* variable and the other to be the *dependent* variable.

10.1 Scatter Diagrams & Correlation Analysis

Learning Objectives

- Understand the Nature of Correlation
- Be able to construct and interpret a Scatter Diagram
- Be able to calculate and interpret a Correlation Coefficient

Unit Contents

- Correlation/Scatter Analysis
- The Scatter Diagram
- The Correlation Coefficient
- Additional Correlation Topics

10.1.1 Correlation/Scatter Analysis

Relationships between Variables

In all quality improvement efforts, we are interested in discovering which *factors* or *variables* influence the *effect* that we're trying to improve:

- *What factors will help us increase the life of this bearing?*
- *What factors will help us reduce the cost of maintaining these chillers?*
- *What factors will help us reduce the time to respond to our customers' orders?*
- *What factors will help us reduce the defects observed in our manufacturing process?*
- *What factors will help us increase the yield of this chemical reaction?*

In many cases, the *type* of relationship that we're looking for can be described as *absence/presence.* We will create a *hypothesis* that if the factor is present, the product or service performs better than it does when the factor is not present:

- *Changing the bearing material will improve the life of this bearing.*
- *Eliminating these unnecessary tests will reduce the cost of maintenance.*
- *Eliminating these unnecessary approvals will reduce the customer response time.*

For the *absence/presence* relationships, tools such as line graphs, Pareto Charts and Histograms can be used to examine if our hypotheses are to be accepted or rejected.

In other cases, though, the type of relationship we're investigating can be described as *variable.* These hypotheses are stated a bit differently. *Changing* the *value* of one or more factors will result in a change in product or service performance:

- *Increasing the pressure in the plastic injection molds reduces the number of defective parts.*

- *Increasing temperature and catalyst concentration improves the yield of the chemical reaction.*

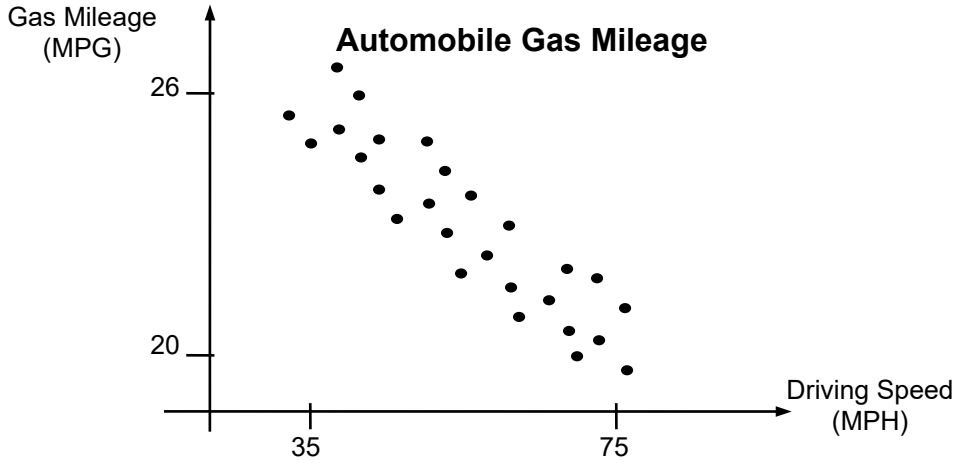
To test these hypotheses, a different sort of tool is required. Here, we'll introduce the Scatter Diagram as the basic (yet powerful!) method of testing these *variable* relationships.

We'll also introduce some rather natural extensions of the Scatter Diagram, providing you with the ability to *measure* correlation and to build a model of relationships through *linear regression.*

10.1.2 The Scatter Diagram

Purpose

The Scatter Diagram shows the relationship between two variables, usually of the *measurement* type. These two variables are often suspected of being tied together through a *cause & effect* relationship.



Here, the car's *driving speed* is the *cause* (or *independent*) variable. The *gas mileage* (number of miles per gallon of gasoline) is the *effect* (or *dependent* variable). Each point on the scatter diagram represents an *observation* - for a given driving speed, what gas mileage was observed? From the diagram, you can see that there is a *negative* relationship[1] (or *correlation*) between these variables - as driving speed increases, the gas mileage decreases.

Notice that the points do not fall on a straight line. There are other sources of variability at work in this process. Very rarely will a "real world" process display perfect correlation between the variables. The "Punch line" of the Scatter Diagram is important. If there exists a correlation between two variables, then you should be able to change the performance of the *effect* (perhaps this is some important quality characteristic), by changing the *independent variable.* The gas mileage example shows us that we could increase our gas mileage by decreasing our driving speed. Driving speed is something we can *control.*

[1]The term *negative* refers to the kind of relationship, as the independent variable *increases*, the dependent variable *decreases*. "Negative" doesn't mean it's a *bad* or undesirable relationship.

The Scatter Diagram also shows you *how much* benefit you'll get from changing the *independent variable.* In the gas mileage example, it appears that we could gain an additional 6 miles per gallon if we could *control* our speed at 35 MPH instead of 75 MPH.

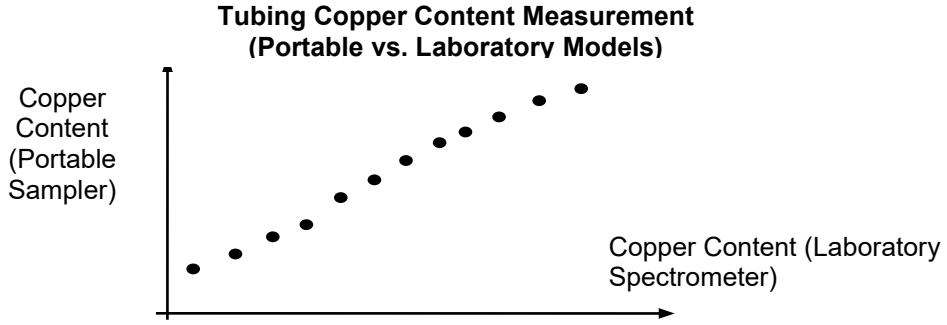
Finally, note that *driving speed* may not be the best independent variable – *wind resistance, road friction,* and others may be better engineering variables to begin exploring how to improve gas mileage.

Application

The Scatter Diagram is typically used in the *Analysis* step of a DMAIEC improvement project. Suspected cause and effect relationships can be verified through use of this tool.

In *Improve*, the Scatter Diagram can be used to determine where the *independent variable* should be controlled. At what level should you set this variable to achieve the desired improvement in the *effect.*

The Scatter Diagram is also used when it's necessary or desirable to develop a *surrogate indicator,* one that is perhaps cheaper or easier to obtain.



In this example, we are testing to see if a portable sampler (the *surrogate indicator*) can perform as well as the laboratory spectrometer. Copper samples of different concentrations are prepared, and both the laboratory and portable samplers are used to measure a given solution. If the portable sampler displays close correlation to the laboratory spectrometer, then we can feel confident in using this as the surrogate indicator. The example scatter diagram shows that there is very

good correlation in the "middle" range of copper samples, but toward the "tails" (very high and low content), the correlation is not as good.

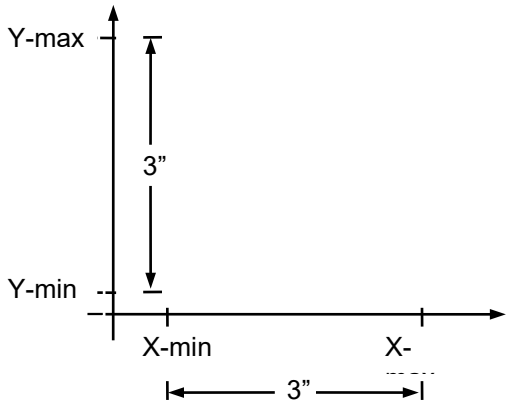
Construction

1. Define the two variables that you think are correlated. Determine which is the *causative* (or *independent*) variable and which is the *effect* (or *dependent* variable). If there is some confusion over this, ask the question, "Which variable will *move* the other variable?"

2. Collect the data in pairs. For some processes, the *independent* variable may vary "naturally." Air temperature, humidity, rainfall and sunlight are examples of naturally varying *independent* variables. For some processes, you may have control of the independent variable. Here, you will have to adjust the values of the independent variable to obtain the scatter diagram points. You will want to define some adjustment range for this variable. The correlation you'll be examining will be limited to this adjustment range. Chemical concentrations, furnace or processing temperatures, part dimensions and processing times are examples of "adjustable" variables.

3. Draw horizontal and vertical axes. Label the horizontal axis with the name of the *independent* (or *causative*) variable, the vertical axis with the *dependent* variable (or effect).

4. Scale these axes so that both the *independent* and *dependent* variables' *range* is about the same *distance.* This ensures that you get the best picture of the correlation between the variables:



5. Plot the data pairs as points on the scatter diagram. If there are identical points, draw a circle around the first one plotted.

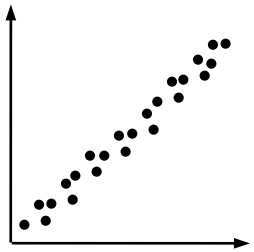
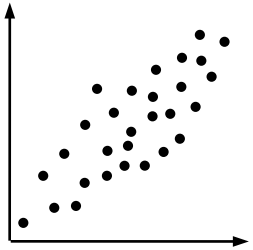
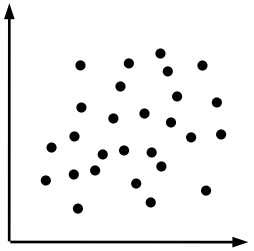
6. Make sure you put a title and label on the scatter diagram. Include the date(s) the data was collected and who prepared the diagram.

7. Interpret the scatter diagram for possible correlation between the variables.

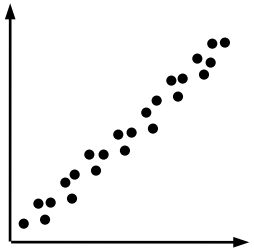
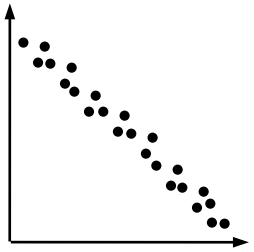
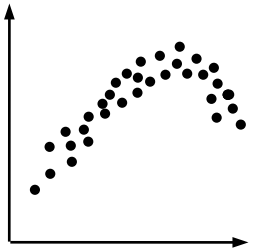
Interpretation

There are two dimensions to scatter diagram interpretation. First, you want to determine if there is any "worthwhile" correlation between the two variables. Second, you want to determine the *type* of relationship between the variables.

Correlation

Strong	Medium	Weak - None
		
The pattern of dots is very tight. You can almost "see" a line that could be drawn through the dots. Carefully controlled process experiments may display this correlation.	There is still some relationship between the variables, but it appears that there are other variables at work, influencing the variation in the *effect.* This is typical of many processes.	No relationship exists between these two variables.

Relationship Type

Positive	Negative	Non-Linear
		
Here, the dependent variable *increases* as the independent variable increases.	Here, the dependent variable *decreases* as the independent variable increases.	Here, the dependent variable first increases, but then begins to decrease as the independent variable increases.

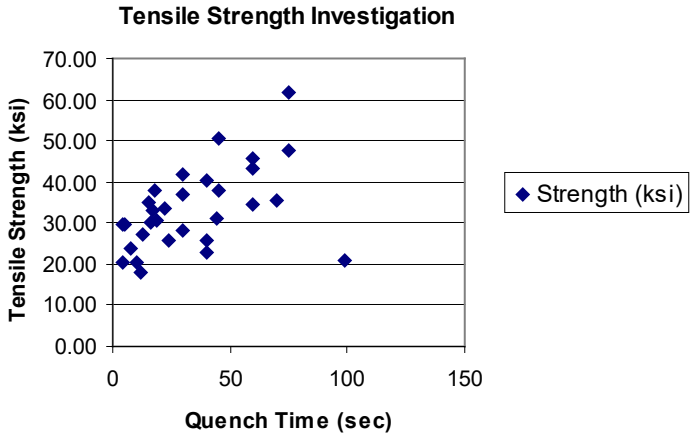
Example – Scatter Diagram & Interpretation

A Black Belt is investigating if the temperature of a quench tank influences the tensile strength of steel. The Black Belt obtains 30 steel coupons, heats them in an induction furnace and then quenches them for a given number of seconds. She then measured the tensile strength of each coupon. The data appear below:

Coupon	Seconds	Strength (ksi)		Coupon	Seconds	Strength (ksi)
1	8	23.61		16	30	28.11
2	75	61.50		17	60	34.50
3	44	31.20		18	30	41.88
4	75	47.70		19	10	20.40
5	30	36.93		20	60	43.20
6	4	29.52		21	24	25.74
7	45	50.70		22	70	35.40
8	16	30.12		23	12	17.94
9	45	37.86		24	60	45.72
10	40	22.86		25	40	25.62
11	22	33.36		26	17	33.27
12	40	40.29		27	15	35.10
13	13	27.21		28	18	38.13
14	**99**	**20.82**		29	4	20.34
15	19	30.84		30	5	29.46

Outlier

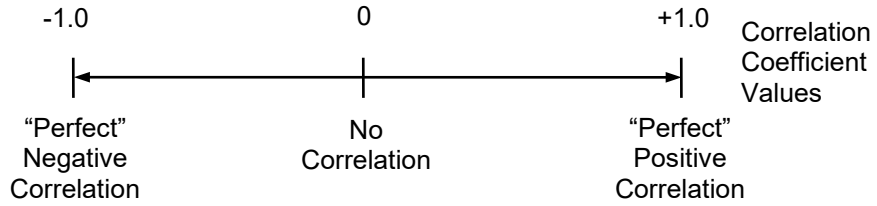
She then prepared a scatter diagram to investigate the relationship. There appears to be a positive correlation between the quench time and the tensile strength. She should investigate the outlier (99 seconds and 20.82 ksi). Perhaps the data was incorrectly recorded on the checksheet:



10.1.3 The Correlation Coefficient

Purpose

The Scatter Diagram gives you a picture of the correlation between two variables. Often, this is all that is needed. The Correlation Coefficient provides you with a *measure* of the *strength* of the relationship between two variables.



The term "Perfect" Correlation means that all the data fall on a straight line, in essence, there is no variation from this line. Negative correlation indicates that the dependent variable *decreases* as the independent variable increases; positive correlation indicates that the dependent variable *increases* as the independent variable increases.

Application

The Correlation Coefficient is used in support of a Scatter Diagram. But make sure that you ***always*** draw the picture! It's very easy to obtain correlation coefficients that indicate one thing, with the Scatter Diagram giving you another picture.

Calculation

Most quality software packages today will calculate one or more correlation coefficients. Here is the "simple" correlation coefficient calculation:

$$r = \frac{SS(xy)}{\sqrt{SS(x) \times SS(y)}}$$

$$where:$$

$$r \text{ - Simple Correlation Coefficient}$$

$$SS(xy) \text{ - Sum of Squares of xy, } SS(x) \text{ - Sum of Squares of x, and}$$

$$SS(y) \text{ - Sum of Squares of y}$$

The Sums of Squares are calculated as follows:

$$SS(xy) = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}$$

$$SS(x) = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x\right)^2}{n}$$

$$SS(y) = \sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y\right)^2}{n}$$

$$where: \bar{x} \text{ - average of x's, } \bar{y} \text{ - average of y's, and } n \text{ - number of data}$$

Since this is the first time we encounter *Sums of Squares* "in the wild," let's explain this important concept. We know that the mean tells us where the *center* of the data lies. The individual data points are found on either side of the mean. Intuitively, to estimate the *dispersion* in the data, we could subtract the mean from the individual points – the larger these differences are, the larger the data's dispersion. However, we want a single measure that provides an estimate of the data's dispersion. If we simply added all the differences, then the total sum of the differences would be zero (the "positive" differences from data above the mean would cancel out the "negative" differences from data below the mean). If we *square* each value, though, they will all be positive and we get a meaningful measure of dispersion, the *Sum of Squared Differences* (abbreviated as the *Sum of Squares*). If we take this calculation one step further and divide by the number of data, we will have an average of the sum of squared differences – this is called the *Variance.*

Interpretation

As mentioned above, the closer *r* is to -1 or +1, the stronger the correlation between the variables. If the absolute value of the correlation coefficient (|*r*|) is greater than about 0.8, the *independent variable* can be considered a good *predictor* of the *dependent variable.*

If the correlation coefficient is less than 0.8, then there are likely to be other variables at work influencing the dependent variable. An *r*-value close to zero indicates that there is little to no relationship.

If you suspect that other variables are at work, it may be profitable for you to examine the correlation of multiple factors to the dependent variable through a *designed experiment (see Section 11).*

Example – Correlation Coefficient

The Black Belt investigating the tensile strength/quench time relationship now takes her data and develops the correlation coefficient (the outlier point was removed since a mistake in data collection was confirmed):

Coupon	Seconds	Strength (ksi)
1	8	23.61
2	75	61.50
3	44	31.20
4	75	47.70
12	40	40.29
13	13	27.21
~~14~~	~~99~~	~~20.82~~
15	19	30.84

Coupon	Seconds	Strength (ksi)
16	30	28.11
17	60	34.50
18	30	41.88
19	10	20.40
27	15	35.10
28	18	38.13
29	4	20.34
30	5	29.46

Calculations:

$$Average\,of\ Seconds = \overline{X} = 32.1$$

$$SS(x) = 13596.7$$

$$Average\,of\ Strength = \overline{Y} = 33.7$$

$$SS(y) = 2743.8$$

$$SS(xy) = 35696.2 - \frac{931 \times 978.5}{29} = 4282.7$$

$$r = \frac{4282.7}{\sqrt{13596.7 \times 2743.8}} = 0.70$$

Conclusion: There is a good correlation between the quench time and tensile strength.

10.1.4 Additional Correlation Topics

Test of Significance for ρ

A correlation coefficient (r) calculated from a sample of data is an *estimate* of the population's correlation coefficient (ρ). The following hypothesis test allows you to determine if the evidence obtained from the sample is strong enough to conclude that the *population* correlation coefficient is significantly different from 0 (i.e. no correlation).

1. Calculate the correlation coefficient from the sample data (r).

2. Calculate the degrees of freedom, f, by subtracting 2 from your sample size ($f = n - 2$).

3. Using the degrees of freedom f, look up the corresponding ρ-value on the following table:

f	ρ	f	ρ	f	ρ
1	.997	16	.468	35	.325
2	.950	17	.456	40	.304
3	.878	18	.444	45	.288
4	.811	19	.433	50	.273
5	.754	20	.423	60	.250
6	.707	21	.413	70	.232
7	.666	22	.404	80	.217
8	.632	23	.396	90	.205
9	.602	24	.388	100	.195
10	.576	25	.381	125	.174
11	.553	26	.371	150	.159
12	.532	27	.367	200	.138
13	.514	28	.361	300	.113
14	.497	29	.355	400	.098
15	.482	30	.349	500	.088

4. If the absolute value of r (i.e. $|r|$) is greater than the table value, then reject the hypothesis that the population correlation coefficient is zero.

5. State the conclusion.

6. If the Null Hypothesis is rejected, then determine the confidence interval for the population correlation coefficient (note that this procedure is valid for sample sizes > 25). First, use the Fisher ***z*** transformation:

$$z = \frac{1}{2}\ln\left(\frac{1+r}{1-r}\right)$$

$$where: \ln \text{ - Natural (base e) logarithm}$$

Then, construct the interval:

$$z \pm K_{\alpha/2}\sqrt{1/(n-3)}$$

$$where:$$

$$K_{\alpha/2} \text{ - Normal Distribution percentile for } (a/2)100\%$$

$$\text{n - Sample Size}$$

Finally, "un-transform" the interval values to obtain the confidence bound for the population correlation coefficient (note that you will calculate an upper bound, lower bound and point estimate using the following formula):

$$r = \frac{e^{2z}-1}{e^{2z}+1}$$

Coefficient of Determination

The *square* of *r* is called the *Coefficient of Determination.* This coefficient's values range from 0 to 1, and there is a special interpretation you can make of r^2*:*

r^2 is the relative proportion of the variation in the dependent variable that can be explained by the independent variable.

The closer r^2 is to 1, the larger the proportion of the variation is explained. For example, an r^2 of 0.81 (equivalent to an r of 0.9) indicates that over 80% of the variation in the dependent variable is explained by the independent variable.

A confidence bound for the coefficient of determination may be found by squaring the values obtained for the correlation coefficient, if these bounds are both the same sign.

Correlation versus Cause & Effect

There is an old saying, attributed to Mark Twain:

There are liars, damn liars, and then there are statisticians!

In today's world of instant media, we hear frequent reports of "correlations" that have been discovered by this or that research group. Some of these are relatively harmless ("laboratory rats who eat more than 1000 Twinkies a day get fat"), some of these have major implications for the human race ("increases in fossil fuel burning are leading to global warming"), and some of these are irresponsible, raising fears among people who don't have the experimental or statistical knowledge necessary to distinguish the "true" from the "not-so-true" correlations.

> *A few years ago, when our daughter and her playmates were about two years old, the "Alar" pesticide scare was reported. There was little we could do to comfort the mothers who were terrified that the apple juice they had given their children could cause cancer. Responsible scientists later proved this correlation to be false.*

Although there are several issues associated with "responsible correlation" analysis, we will only tackle one important issue here. We'll assume that you are like Diogenes, holding up a lantern looking for truths that will help you improve the quality of your products and services.

Not all correlations correspond to a cause-and-effect relationship.

Consider the following examples:

- *The population of Paris was positively correlated to the number of stork nests in the city.*

- *The homicide rate in Chicago is positively correlated to the sales of ice cream by street vendors.*

- *At one time, the Dow Jones stock market index was positively correlated to the height of women's hemlines.*

These are all examples of "nonsense" correlations. There is no cause-and-effect relationship between any of these variables.

For the first two examples, though, there is a *third* variable that drives both supposedly correlated variables. In Paris, as the population increased, more houses were built. Guess where storks like to build their nests? In chimneys!

You've probably already identified the second example's "hidden" variable - *outside temperature*.

When a correlation appears on your Scatter Diagram, make use of your knowledge of your business' technology and science. Does the correlation make sense from a cause-and-effect standpoint? The ultimate measure, though, is *repeatability.* If a "true" cause and effect relationship exists, you'll be able to adjust the independent variable and see the corresponding change in the dependent variable.



10.2 Regression Analysis

Learning Objectives

- Be Able to Develop and Interpret a Linear Regression Model
- Be Able to Develop and Interpret Regression Confidence Bounds
- Be Able to Develop and Interpret Multiple Linear and Non-Linear Regression Models
- Be Able to Develop and Interpret Regression Models for Binomial (Logistic) Responses

Unit Contents

- Linear Regression
- Checking the Regression Model & Additional Topics
- Multiple Linear Regression Models
- Transformations & Non-Linear Regression Models
- Binary Logistic Regression

10.2.1 Linear Regression

Purpose

If two variables are correlated, the Scatter Diagram can be used to determine the *impact* of changing the independent variable (i.e. process factor) on the dependent variable (i.e. process outcome or effect). If you simply draw a line through the Scatter Diagram's points, then you are on the way to creating a model of how one variable affects another. You have taken on the role of a scientist, discovering physical "laws" that become part of your organization's *technology.* But there are many lines that could be drawn through the data. How do we judge the "best" line? *Linear Regression* is the method used to calculate the parameters of a line that best fits your data.

Application

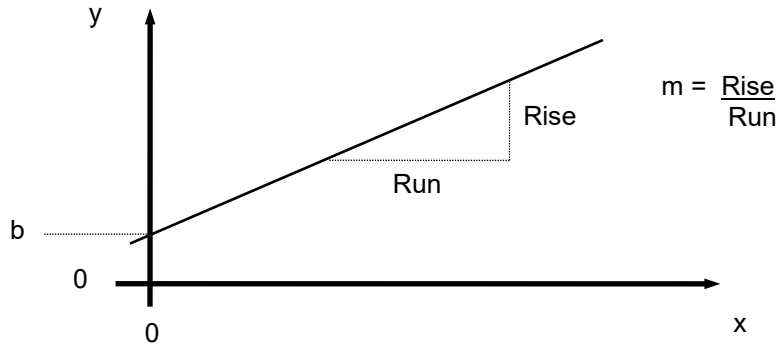
Linear Regression creates a model that predicts changes in the dependent variable (process output) as a function of the independent variable (process factor). In *Countermeasures,* you can use this model to determine the best "setting" or level of this factor.

Linear Regression Theory

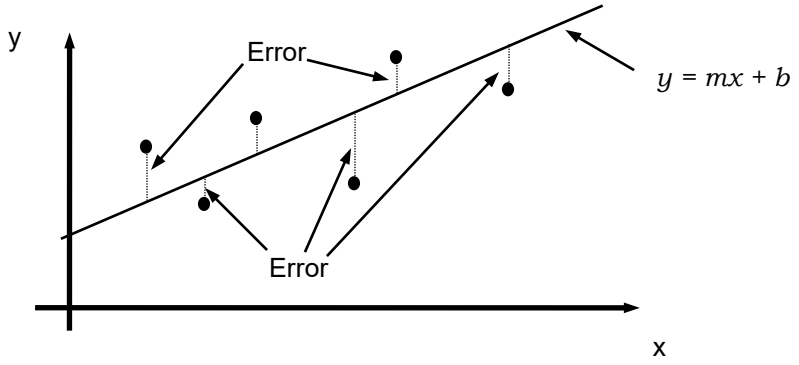
There are several different regression analysis methods; we will employ the method of *least squares.* From basic algebra, we recall the equation of a line:

$$y = mx + b$$

A picture of this line shows the purpose of the two parameters, *m* and *b:*



The slope of the line is *m,* found by dividing the rise over the run. The y-intercept is *b,* where the line passes through the y-axis (equivalent to an x-value of zero). The Scatter Diagram consists of points that are, of course, "scattered" on the x - y plane. Let's pass an arbitrary line through the data. From any point, then we can draw a vertical line to our line. We'll call the vertical line's distance the "error" (aka a *residual*):



Now we can calculate the error (e_i) for any point, x_i, y_i:

$$e_i = y_i - (mx_i - b)$$

The method of *least squares* determines estimates of the line's parameters (*m* and *b*) by minimizing the sum of squares of the errors:

$$SSe = \sum_{i=1}^{n} (e_i)^2$$

where:

SSe - Sum of Squares of Error Terms

Calculations

1. Collect the data pairs or obtain them from the raw data used to create the Scatter Diagram.

2. Calculate the average of the dependent variable and independent variable:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ and } \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

where :

y_i - "ith" value of independent variable

$\bar{y}$ - independent variable average

x_i - "ith" value of dependent variable

$\bar{x}$ - dependent variable average

3. Calculate the estimate of the slope of the regression line $\hat{m}$ (note that we use the "hat" symbol ($\wedge$) to show that we are estimating the true population slope):

$$\hat{m} = \frac{SS(xy)}{SS(x)}$$

where:

$SS(x)$ - *Sum of Squares of x*

$SS(xy)$ - *Sum of Squares of xy*

If you've already performed the correlation analysis, these Sums of Squares are the same as those calculated for the correlation coefficient:

$$SS(x) = \sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x\right)^2}{n}$$

$$SS(xy) = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}$$

where :

$\bar{x}$ - average of x's, $\bar{y}$ - average of y's, and n - number of data

4. Finally, calculate the estimate of the y-intercept, $\hat{b}$:

$$\hat{b} = \bar{y} - \hat{m}\bar{x}$$

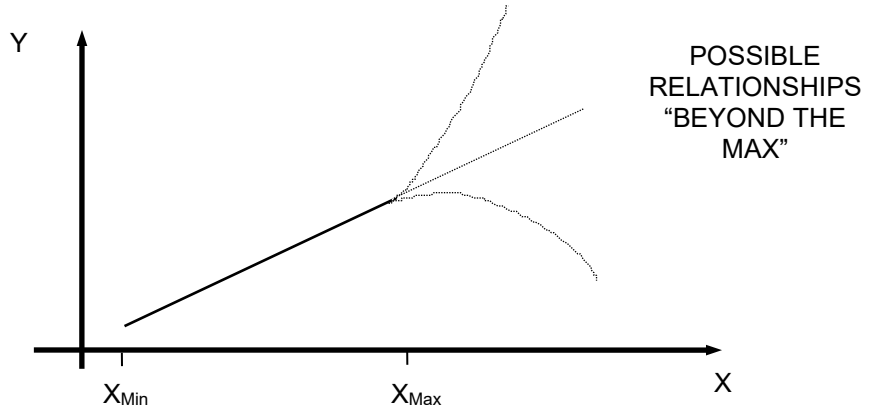
5. Plot the regression line on the scatter diagram. Visually check the line to see if it is a good fit and that no calculation errors were made.

Regression Cautions

Now that you've built the basic regression model, there are several other analyses that will help you both validate and use your model. Once these are complete (See 10.2.2, *Checking the Regression Model*), you can "throw away" the scatter diagram and rely on the model to predict the change in the dependent variable as a function of the dependent variable.

Most statisticians will caution you, though, to use the regression model only within the bounds of the data used to build the model. Engineers, on the other hand, will often want to *extrapolate* the regression model beyond these bounds.

The basic problem with extrapolation is that we do not know whether the relationship remains *linear* or, perhaps, assumes some new function:



When we collect data from an existing population or process, the bounds of the data will be dictated by whatever our sample happens to produce. In many cases, though, it is possible to deliberately alter the value of the dependent variable and, through experimentation, cover the range of values of interest to us.

Example – Linear Regression of Tensile Strength vs. Quench Time

Let's return to the strength example started in 10.1. Since both the Scatter Diagram and the Correlation Coefficient calculation indicated that quench time is a factor affecting tensile strength, the Black Belt now takes her data and develops the regression equation (again, the outlier point was removed since a mistake in data collection was confirmed):

Coupon	Seconds	Strength (ksi)
1	8	23.61
2	75	61.50
3	44	31.20
11	22	33.36
12	40	40.29
13	13	27.21
~~14~~	~~99~~	~~20.82~~
15	19	30.84

Coupon	Seconds	Strength (ksi)
16	30	28.11
17	60	34.50
18	30	41.88
26	17	33.27
27	15	35.10
28	18	38.13
29	4	20.34
30	5	29.46

Calculations:

$$Average\ of\ Seconds = \overline{X} = 32.1$$

$$SS(x) = 13596.7$$

$$Average\ of\ Strength = \overline{Y} = 33.7$$

$$SS(y) = 2743.8$$

$$SS(xy) = 35696.2 - \frac{931 \times 978.5}{29} = 4282.7$$

$$\tilde{m} = 4282.7/13596.7 = 0.31$$

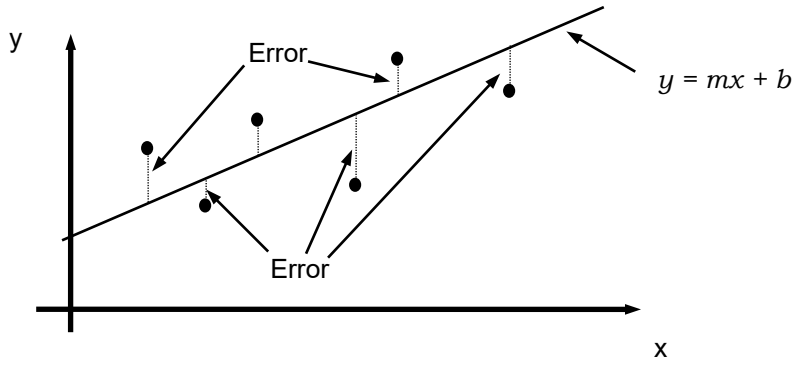
$$\tilde{b} = 33.7 - 0.31 \times 32.1 = 23.8$$

$$\therefore Y = 0.31X + 23.8 \text{ is the regression equation}$$

10.2.2 Checking the Regression Model & Additional Topics

Regression Residuals

The main assumption of *linear* regression is that the error terms, e_i (known as the *residuals*) are normally distributed with mean equal to 0 and variance equal to σ^2. Now that the regression model is constructed, we can test this assumption. The *residuals* are simply the vertical distance (i.e. error) from each point to the regression line:



Knowing the slope, *m,* and the y-intercept, *b,* we can calculate these residuals easily for each data pair (x_i, y_i):

$$e_i = y_i - (mx_i - b)$$

Three simple analyses are recommended for the residuals:

1. If the scatter diagram data were obtained in time sequence, construct an X, mR (Individuals) Control Chart of the residuals. Examine the chart for assignable causes. If these are found, try to understand what other variables may be acting on the process or population.

If you suspect there are other important variables, develop a residuals' control chart using these variables as subgroups. This information may help you to refine your model (i.e. by adding additional variables to your model, see *Multiple Linear Regression*, below).

2. Construct a histogram, or normal probability plot of the residuals. Examine the shape of these graphs and/or perform a test of normality (e.g. the Kolmogorov-Smirnov test) for the residuals. If the hypothesis of normality is rejected, the model will have to be modified.

3. Construct a Scatter diagram of the *residuals vs. the independent variable*. This can help reveal if the assumption of *constant variance* is met (i.e. *homoscedasticity*). If the Scatter broadens in either direction (< or >), then this is evidence of a *heteroscedastic* process.

4. Plot the residuals against the Fitted Values – for each "x," calculate the associated "y" using the regression equation. Plot these values on the horizontal axis and the associated residual on the vertical axis. Again, look for patterns in the residuals.

Following the discussion on confidence bounds, we will present possible remedies for these situations.

Confidence Bounds on the Regression Model

When we perform hypothesis tests, the last step is to construct a confidence interval for the population parameter. This confidence interval is centered at the point estimate of the parameter. The confidence level (set by us) and the standard deviation of the population parameter) determine its width.

This same concept applies to regression; we simply must adjust our thinking from a one-dimensional point to a two-dimensional line.

Testing the Slope, m for Significance

Here, we will determine if there is enough evidence to conclude that the slope of the line is not equal to zero (if it *is* zero, then changes in the independent variable do not result in changes in the dependent variable). The hypothesis test procedure is as follows:

1. Establish the Hypothesis
    a) Null Hypothesis (H_o) - $m = 0$
    b) Alternative Hypothesis (H_a) - $m \neq 0$

2. Choose a Significance Level (α - "alpha")

3. Plan the Test:
   a) The Test Statistic is:

$$t = \frac{\hat{m}\sum_{i=1}^{n}(X_i - \overline{X})^2}{MSe}$$

where:

MSe - Mean Square Error

$$MSe = \frac{1}{n-2}\left(\sum Y_i^2 - b\sum Y_i - m\sum X_i Y_i\right)$$

or

$$MSe = \frac{\sum_{i=1}^{n} e_i^2}{n-2}$$

   b) Determine the Rejection Region. From a t-distribution table, find the two-sided $K_{\alpha/2}$ for n-2 degrees of freedom.

4. Collect data and calculate the Test Statistic.
   a) Calculate the Mean Square Error and other components of the test statistic.
   b) Calculate the test statistic.

5. Draw a conclusion.

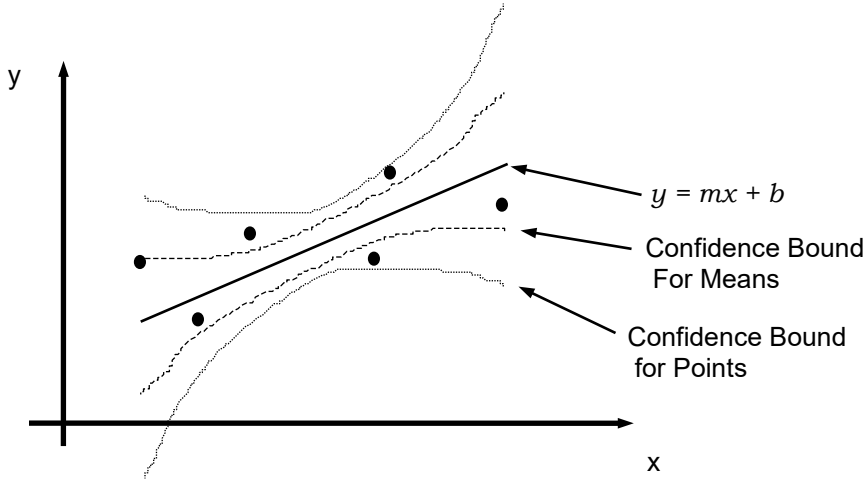
6. Estimate the Slope Parameter:

$$m \pm K_{\alpha/2}\frac{MSe}{\sum(X_i - \overline{X})^2}$$

Developing Confidence Bounds for the Regression Line & Points

Given that the Null Hypothesis, *m = 0* is rejected, confidence bounds for the regression *line* (indicating the uncertainty in both slope and y-intercept) and *data points* (indicating the dispersion in the points around the line) can be calculated.

The equations and calculations are tedious; we will present the picture of these bounds and leave the computations for the statistical software:



Confidence bounds for the Means – plugging in an "x" into the regression equation gives you the expected value of the mean "y" associated with that "x." If you draw a vertical line from the "x," you'll cross the lower-bound, the expected value and the upper-bound of the "y's" mean for that "x." You can state with, say, 95% confidence that the interval contains the true value of the mean for that "x."

Confidence bounds for the Points – Similarly, these bounds can be interpreted as "95% of the points will fall within the confidence bounds for the points for this "x."

Confidence bands for the Line (not shown above) – some software (not Minitab) will display a confidence band, here the interpretation is that there's a 95% chance that the "true line" falls within that band.

Note that the bounds get wider as they approach the minimum and maximum values. This reflects the increasing uncertainty we have at its ends.

We once had to investigate the relationship between centrifugal pump flows (gallons per minute) and the pump's motor power (in Kilowatts). The study required a conservative estimate of the power required to produce a certain flow. In this case, the upper confidence bound on the means was employed, instead of the actual line.

Auto-Correlated Time Series Data

Regression models are often generated for time series data. These models help us understand historical trends and may form the basis for predictions about the future. The linear regression model assumes that the error terms are independent and normally distributed. In many cases, time series data are *not* independent, but rather correlated to each other (termed *autocorrelation*). The most obvious type of autocorrelation is where this time period's data is correlated to the last period (*1 Slip* type (or lag-1)).

For example, suppose a person begins an exercise program in January. At the beginning of the program, the person's body mass index (BMI) and aerobic ability are measured. After each month, these measurements are repeated. Here, we would expect April's results to be related to the measurements of the preceding month.

Some processes may "skip" one or more time periods. April's results may be correlated to February, May's to March's and so on (*2 Slip* type (or lag-2)).

Autocorrelation affects the regression model in two significant ways:
1. Although the regression coefficients (*m* and *b*) are still unbiased, they will not have the property of minimum variance and will be inefficient.
2. The Mean Square Error (MSE) will underestimate the error variance. This impacts the standard deviation of the slope estimator (underestimating it - this then underestimates the confidence interval for the slope).

How to Detect Auto-Correlation

If you suspect the time series data to be auto-correlated, first decide the most likely type of autocorrelation (1 Slip, 2 Slip, 3 Slip, etc.). Then construct a Scatter diagram of the error term pairs and calculate the correlation coefficient.

How to Address Auto Correlation

If the Scatter diagram/correlation analysis provides evidence of autocorrelation, then estimate the autocorrelation factor, ρ_{AC} by performing a regression on the error pairs. The slope of the regression line estimates the autocorrelation factor.

Relationship between ρ and β

As part of the correlation/regression analysis process, we develop estimates of two parameters: ρ and β. Suppose that the correlation coefficient is very close to one, say 0.96. Does this mean that the independent variable has a very *strong* effect on the dependent variable? The answer here depends on how we define *strong.* The correlation coefficient does tell you that the relationship is very linear, as the independent variable changes, the dependent variable tracks right along.

To determine how *much* of an effect a unit change in the independent variable will have on the dependent variable, we look at the regression model and its β estimate. This is the *slope* of the line and answers the above question. For example, if a unit change in independent variable "A" results in the dependent variable "Y" changing by 4, this effect is greater than variable "B" whose unit change only changes the "Y" by 2. The slope of the Y vs. A line is twice that of the Y vs. B line.

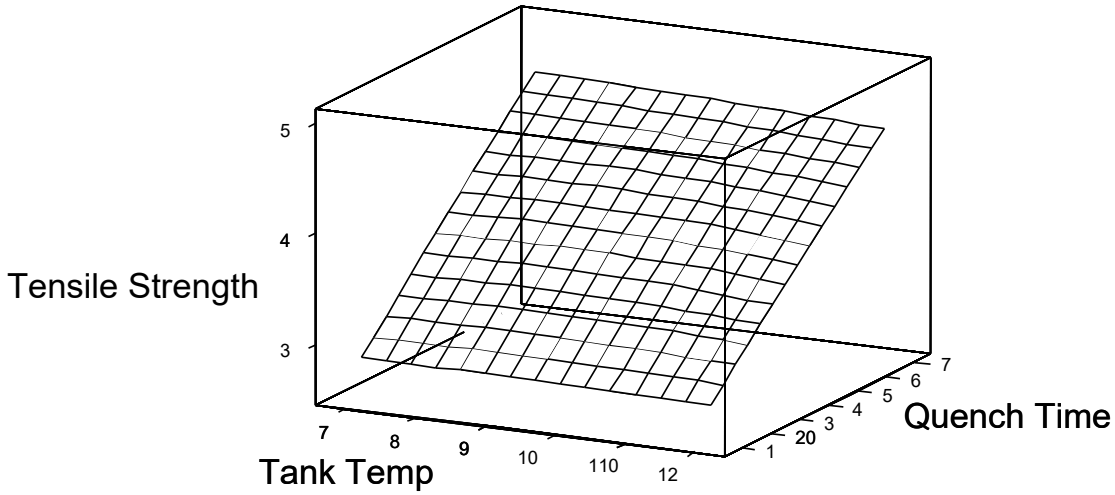
10.2.3 Multiple Linear Regression Models

Multiple Linear Regression

The concept of two-variable linear regression can be extended. We can build linear models that relate *multiple* independent variables to a dependent variable and perform multiple linear regression. This approach is often very valuable when there are multiple variables that are suspected of influencing the effect. The general model becomes:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + \ldots + b_nX_n + e$$

Notice that we've changed the notation for the model's parameters. The y-intercept is still *b,* but the slope parameters have been changed from *m* to b_0, b_1, b_2, etc. Instead of a regression *line,* we now have a regression *plane* or higher dimensional *hyperplane*. Instead of data pairs being needed to "feed" the regression model, data *vectors* are used $(Y_i, X_1, X_2, X_3, \ldots . X_n)$.



The procedure of least squares is still used to determine estimates of the parameters, residual analysis and confidence bounds for the parameters and the hyperplane are also obtained. Most "full-featured" statistical packages (such as Minitab) will perform multiple linear regression.

Historical data may be available to support the multiple linear regression model. For example, medical studies often gather data on lifestyle factors such as diet, exercise, genetic information, and others to attempt to predict the occurrence of heart disease or stroke. Large amounts of data are needed to ensure that the "X's" are seen across their entire ranges. An alternative to historical data (especially for industrial work) is to *design an experiment (see Section 11.1)*. Here, the "X's" are systematically varied over ranges of interest to see how they affect the "Y."

Example – Suppose our Black Belt is still investigating tensile strength. She has now identified the additional variables of *quench tank temperature* and *metal temperature at time of quench*. Experiments have been conducted and the data appear below. We will use Minitab to do the "heavy lifting" for us here:

Tensile Strength Data

Coupon	Quench Time	Tank Temp	Metal Temp	Tensile Strength
1	10	70	800	29.3
2	70	70	800	46.0
3	10	120	800	26.5
4	70	120	800	43.9
5	10	70	1000	26.9
6	70	70	1000	48.5
7	10	120	1000	28.5
8	70	120	1000	43.9
9	10	70	800	28.6
10	70	70	800	44.5
11	10	120	800	26.7
12	70	120	800	45.9
13	10	70	1000	28.3
14	70	70	1000	46.2
15	10	120	1000	27.7
16	70	120	1000	44.2
17	10	70	800	29.0
18	70	70	800	46.7
19	10	120	800	28.9
20	70	120	800	46.4
21	10	70	1000	28.3
22	70	70	1000	47.0
23	10	120	1000	29.8
24	70	120	1000	47.1
25	40	95	900	37.7
26	40	95	900	37.7
27	40	95	900	37.6
28	40	95	900	38.0

Minitab develops the regression model for us:

Regression Analysis: Tensile Strength versus Quench Time, Tank Temp, ...

1

```
The regression equation is
Tensile Strength = 25.4 + 0.294 Quench Time - 0.0163 Tank Temp
           + 0.00167 Metal Temp
```

2

Predictor	Coef	SE Coef	T	P
Constant	25.421	2.358	10.78	0.000
Quench T	0.294167	0.007937	37.06	0.000
Tank Tem	-0.016333	0.009525	-1.71	0.099
Metal Te	0.001667	0.002381	0.70	0.491

3

```
S = 1.167       R-Sq = 98.3%     R-Sq(adj) = 98.1%
```

Analysis of Variance

4

Source	DF	SS	MS	F	P
Regression	3	1873.80	624.60	458.97	0.000
Residual Error	24	32.66	1.36		
Lack of Fit	5	10.22	2.04	1.73	0.176
Pure Error	19	22.44	1.18		
Total	27	1906.46			

5

Source	DF	Seq SS
Quench T	1	1869.13
Tank Tem	1	4.00
Metal Te	1	0.67

Interpretation:

1. First, Minitab provides us with the “raw” regression equation. Tensile strength increases as quench time increases (positive coefficient of 0.294), decreases with increasing tank temperature (negative coefficient of -0.0163 and increases with increasing metal temperature.
2. This table shows which factors are significant and must be used to “clean-up” the “raw” regression equation. For an alpha of 0.1, the two factors – quench time and tank temperature are significant (p values < 0.1), but metal temperature is not significant (p = 0.491 > 0.1). Metal temperature must then be removed from the equation.
3. The R-Squared value of 98.3% tells us that the equation explains a good deal of the variation in tensile strength. Including more terms tends to increase the R-Squared. The R-Squared (adjusted) accounts for and “penalizes” us

for including all the terms in the equation (R-Squared (adjusted < R-Squared). "S" is the standard deviation associated with the error term.

4. Minitab conducts an Analysis of Variance (ANOVA – See Section 10.3). The p-value of 0.000 tells us that the regression equation is significant in explaining variation in tensile strength. Note the "lack of fit" line under the Residual Error. The null hypothesis is that "there is no lack of fit" – i.e. the regression model fits the data well. The p-value of 0.176 tells us we can't reject the null – in this case that is good news!
5. The Sequential Sum of Squares (Seq SS) measures the increase in the Sum of Squares that occurs by adding a factor (X) to the model, given that all the preceding terms are in the model. The Sequential Sum of Squares values depend on both the contribution of the factor and the order of the factors in the model (as entered into Minitab).

Cleaning Up the Regression Equation - The analysis is rerun in Minitab, with the Metal Temperature factor removed:

Regression Analysis: Tensile Strength versus Quench Time, Tank Temp

```
The regression equation is Tensile Strength = 26.9 + 0.294 Quench Time - 0.0163 Tank Temp

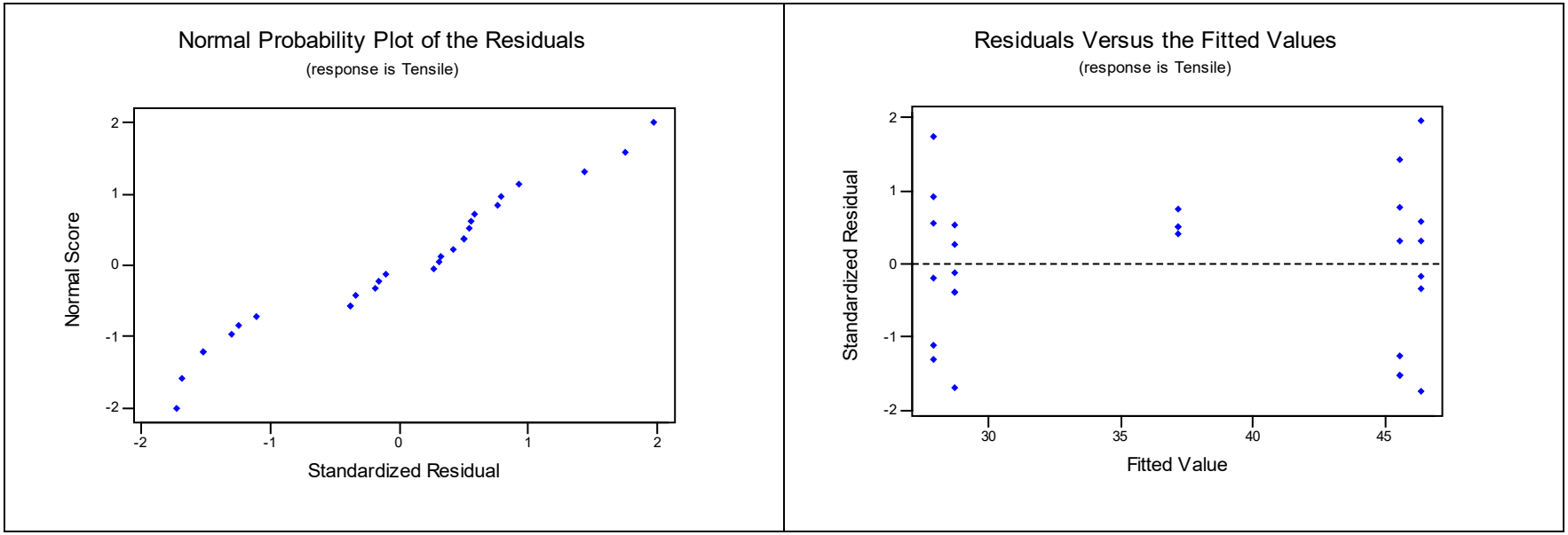
Predictor        Coef     SE Coef          T        P
Constant      26.9207      0.9739      27.64    0.000
Quench T     0.294167    0.007856      37.44    0.000
Tank Tem    -0.016333    0.009427      -1.73    0.095

S = 1.155       R-Sq = 98.3%     R-Sq(adj) = 98.1%

Analysis of Variance
Source            DF          SS          MS         F        P
Regression         2     1873.14      936.57    702.55    0.000
Residual Error    25       33.33        1.33
  Lack of Fit      2        2.89        1.44      1.09    0.353
  Pure Error      23       30.44        1.32
Total             27     1906.46

Source       DF      Seq SS
Quench T      1     1869.13
Tank Tem      1        4.00
```

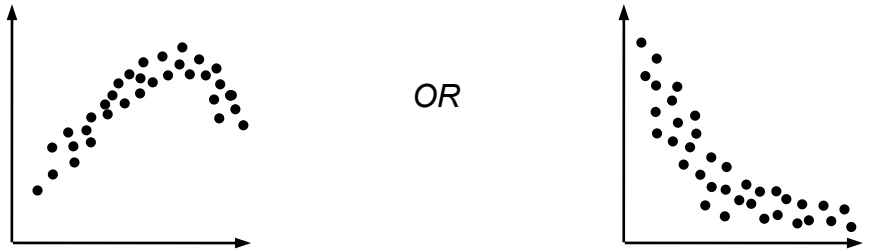
To ensure that the residual assumptions (random and normally distributed) are met, the residual graphs are generated. Neither graph indicates any problem with the residuals:



The final regression equation is: ***Strength = 26.9 + 0.294 Quench Time - 0.0163 Tank Temp***. This model explains about 98% of the variation in tensile strength. The residual error (ε) is normally distributed with a mean of 0 and a standard deviation of 1.155 (N(0, 1.155)).

10.2.4 Transformations & Non-Linear Regression Models

The discussion so far has focused on *linear* models, where a straight line describes the relationship between the variables. Fortunately, many "real world" relationships fall into this category. But suppose the Scatter diagram of data pairs appears as one of the following:



Here, a straight line would not provide a good "fit" for these data. There are two basic options we can pursue.

Transformations

In some cases, we may be able to *transform* a model that we think fits the data into a linear model. Here's a simple example of a linear transformation. Suppose the "true" relationship of X_1 and X_2 to Y involves an exponential function. A simple logarithmic transformation will enable us to express the ln(Y) as a linear function of the X's
:

$$\mathrm{Y} = \mathrm{b}_0 \exp(b_1 X_1 + b_2 X_2 + a)$$

$$\text{Transformation by taking natural log of both sides :}$$

$$\ln \mathrm{Y} = \ln \mathrm{b}_0 + b_1 X_1 + b_2 X_2 + a$$

The latter equation can be fit by multiple linear regression. The Y_i are first transformed by taking their natural logarithms. We would then perform the regression analysis using the $\ln(Y_i)$ as the dependent variable and the X_i's as the independent variables. If the model fits (don't forget the residuals' analysis!), then we would then generate the final model by raising the linear function to the power of *e*.

Non-Linear Regression Models

A higher-order model – i.e. a quadratic or cubic model may also be employed to fit the data. A Taylor Series expansion of many functions shows us that a quadratic equation may often provide a good approximation to the function. In this case, the regression model would look like this:

$$Y = b_0 + b_1 X_1 + b_2 X_1^2$$

or, for more than one X :

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1^2 + b_4 X_2^2$$

Advanced Regression Approaches

In other cases, there is no transformation that will transform the model into a linear form:

$$Y = b_0 10^{b_1 X_1} + e$$

There are several approaches for these models, including the *Solution of Normal Models, Gauss-Newton Method* and the *Method of Steepest Descent.* These are described in reference 7 (see Bibliography). Again, your statistical software package will likely have one or more of these approaches included.

Example – Transformation

A new design of small "neighborhood" nuclear reactor incorporates a power control system that, in part, depends on the position of control rods in the reactor. As part of the test program, the engineers want to validate the relationship between rod position (inches out of the core) and power level. Their simulations predict the relationship to be:

$$Power = 1400 \times e^{0.25 \times h}$$

where

Power - thermal output in kilowatts (kw)

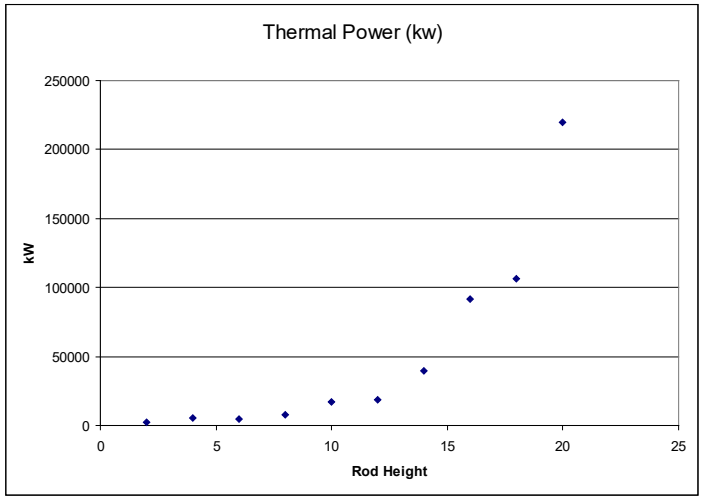
h – Rod Height (in.)

Data from the prototype reactor is collected, with the results appearing below:

Rod Height	Thermal Power (kw)
2	2,220
4	5,779
6	4,436
8	7,793
10	16,979
12	18,659
14	39,424
16	91,694
18	106,493
20	220,053

Approach – First, let's plot the data on a scatter chart. The relationship does not appear to be linear (as expected by the engineers). Since theory provides a "clue" that an exponential relationship exists, the natural logarithm of the thermal power data is taken (next page).

Scatter Plot of Data:



Raw and Transformed Data:

Rod Height	Thermal Power (kw)	Log Transform
2	2220	7.7051
4	5779	8.6619
6	4436	8.3976
8	7793	8.9610
10	16979	9.7397
12	18659	9.8341
14	39424	10.5821
16	91694	11.4262
18	106493	11.5758
20	220053	12.3016

We will use Minitab to perform the regression on the transformed Y.

Regression Analysis: LogTrans versus Rod

```
The regression equation is
LogTrans = 7.18961 + 0.248083 Rod

S = 0.268581      R-Sq = 97.2 %      R-Sq(adj) = 96.9 %

Analysis of Variance

Source            DF         SS         MS          F      P
Regression         1    20.3098    20.3098    281.549  0.000
Error              8     0.5771     0.0721
Total              9    20.8869
```

We see that the regression is significant and that the R-Squared is high (the equation explains a large amount of the variation in the Log Transformation of the data. To determine how good the engineers' prediction is, we raise the regression equation to the power of *e:*

$$Power = e^{7.19} \times e^{0.248 \times h} = 1326 \times e^{0.248 \times h}$$

This compares well with the results of the engineers' simulation:

$$Power = 1400 \times e^{0.25 \times h}$$

where

Power - thermal output in kilowatts (kw)

h – Rod Height (in.)

Example – Quadratic Regression

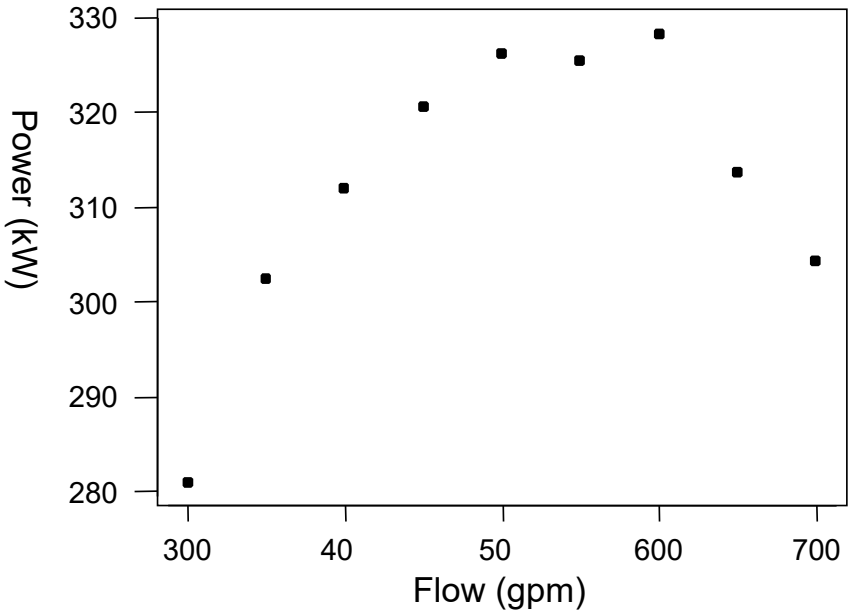
This example was one of the author's first "real" regression analyses, so permit us to reminisce. In the mid-1980's, a nuclear plant discovered that the two emergency diesel generators could be overloaded under certain accident conditions. The design process of controlling how additional loads were added to the generators over the years was weak. As part of the engineering analysis of the situation, the question arose: "*How much load could large safety-related pumps put on the diesel generators?"* The key variable affecting power was the flow produced by the pumps (this does seem backwards – power input to the pump determines output flow. However, the response of the systems during accident conditions meant that *power as a function of flow* was the more interesting question). Data available from start-up testing were available; these indicated that the relationship between flow and power was non-linear; some pump theory calculations indicated that a quadratic model might be a good fit.

The data below are typical of one of the safety pumps. The scatter diagram demonstrates the non-linearity of the response:

High Pressure Safety Injection Pump Data:

Flow (gpm)	Power(kw)
300	280.9
350	302.5
400	312.0
450	320.6
500	326.3
550	325.5
600	328.3
650	313.7
700	304.4

Scatter Diagram (Power vs. Flow):



Minitab is used to develop a quadratic regression model for this data:

Polynomial Regression Analysis: Power(kw) versus Flow (gpm)

1

```
The regression equation is
Power(kw) = 90.1924 + 0.890661 Flow (gpm) - 0.0008356 Flow (gpm)**2
```

2

```
S = 2.63793      R-Sq = 97.7 %      R-Sq(adj) = 97.0 %
```

```
Analysis of Variance
```

3

Source	DF	SS	MS	F	P
Regression	2	1798.48	899.238	129.226	0.000
Error	6	41.75	6.959		
Total	8	1840.23			

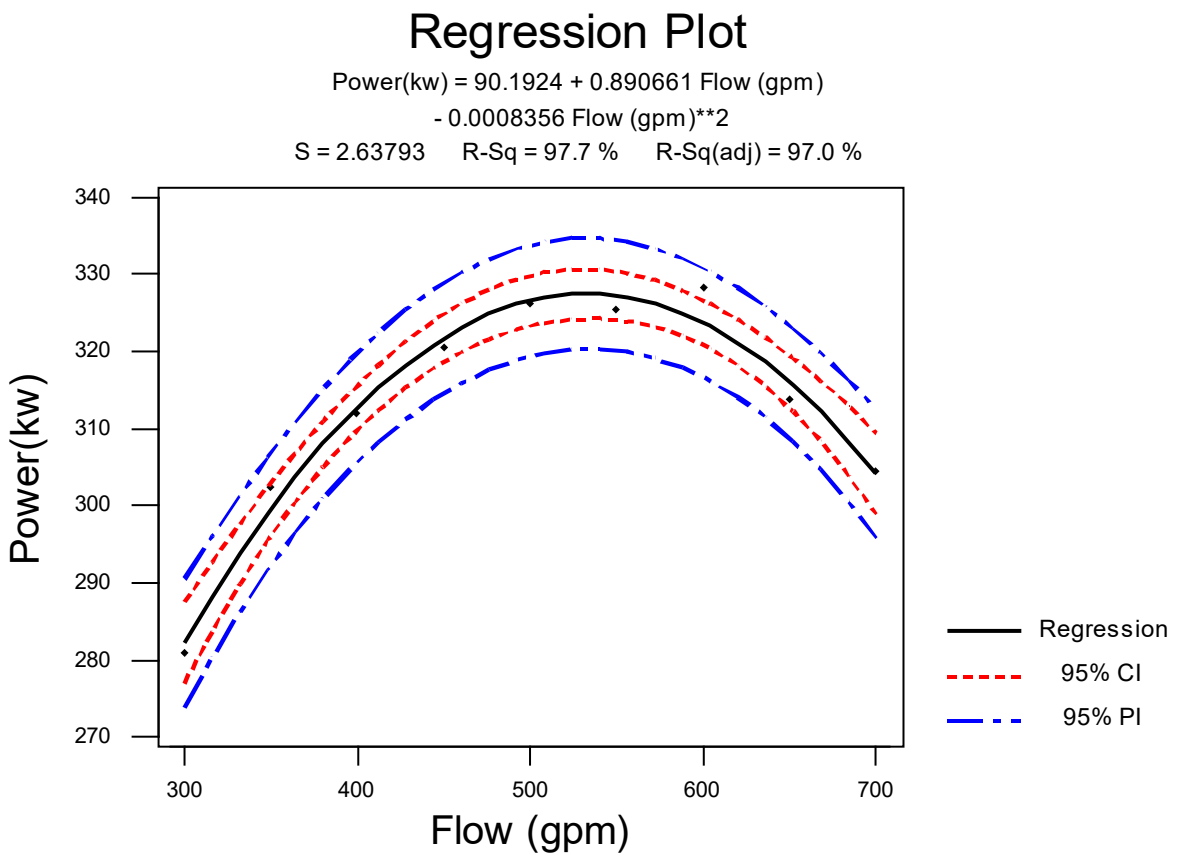
4

Source	DF	Seq SS	F	P
Linear	1	454.30	2.295	0.174
Quadratic	1	1344.18	193.166	0.000

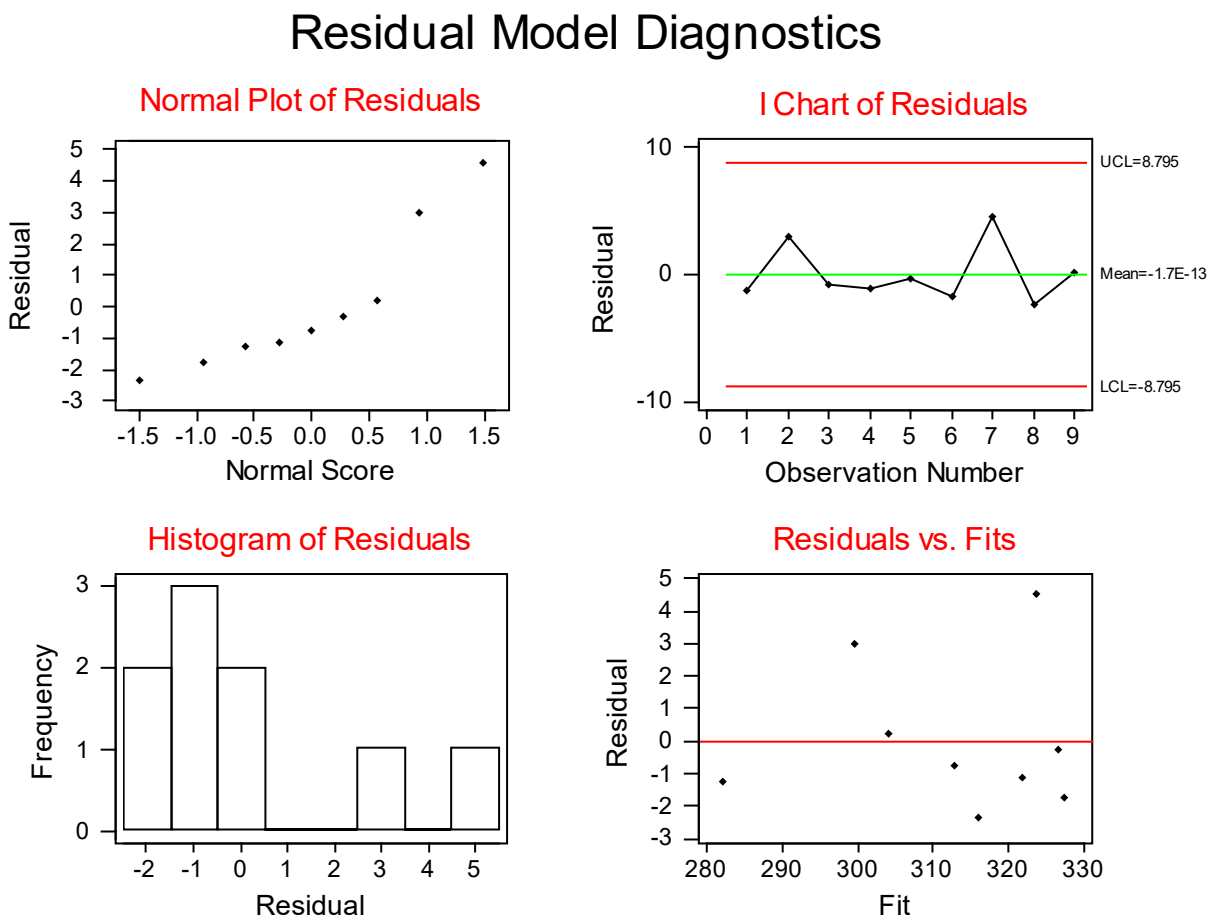
Interpretation:

1. The regression model is listed here. Note that the general form of the quadratic equation includes a) the constant, b) the linear term and c) the quadratic term.
2. The R-Squared is very good; the model explains 97% of the variation in Pump Power as a function of Flow.
3. The ANOVA table (see Section 10.3), shows the regression model is significant.
4. The Sequential Sum of Squares (Seq SS) measures the increase in the Sum of Squares that occurs by adding a factor (X) to the model, given that all the preceding terms are in the model. The Sequential Sum of Squares values depend on both the contribution of the factor and the order of the factors in the model (as entered into Minitab).

The fitted line plot appears below. The plot includes confidence bounds (95%) for the means (red curves) as well as prediction intervals (we expect 95% of the points to fall inside the blue prediction curves).



Finally, to ensure that the assumptions are met regarding normality and randomness of the residuals, these plots are included on the next page:



The final model is: **Power(kw) = 90.1924 + 0.890661 Flow (gpm) - 0.0008356 Flow (gpm)**2**. The model explains 97% of the variation in Pump Power (kW) as a function of Flow (gpm).

Post-Script – These models were used to develop load profiles for the diesel generators during design-basis accidents and were one (small!) part of a large effort to justify operation of the nuclear plant until two more diesel generators could be added to the plant's emergency power system.

10.2.5 Binary Logistic Regression

Introduction

So far, we have treated dependent variables that can be measured on a *continuous* scale. Linear regression, multiple regression, transformations and non-linear regression serve us well in these cases. What about dependent variables that can only be measured with a Go/No-Go, or On-Off scale? For example, welds may be good or defective, employees are either male or female, a customer has bought a product or not. There may be one or more independent variables; these may either be *continuous* or *discrete*.

For example, the quality of a weld (defective or not) may depend on the number of passes to complete the weld (*discrete*), the voltage/current of the welding circuit (*continuous*) and the welder's skill (years of welding experience (*continuous*)).

An ordinary linear regression model will not work for Go/No-Go independent variables. However, the *logistic regression model* can come to our rescue. The outcome of the experiment or trial (e.g. the *y*) is either a 0 or 1 (you should assign the "desired" outcome the value of 1 – interpretation is then easier). Independent variables may be either *continuous* or *discrete.* For a one factor (X) situation, the regression model takes the form:

$$Y_i = \beta_o + \beta_o \times X_i + \varepsilon_i$$

$$Where:$$

$$Y_i = 0, 1$$

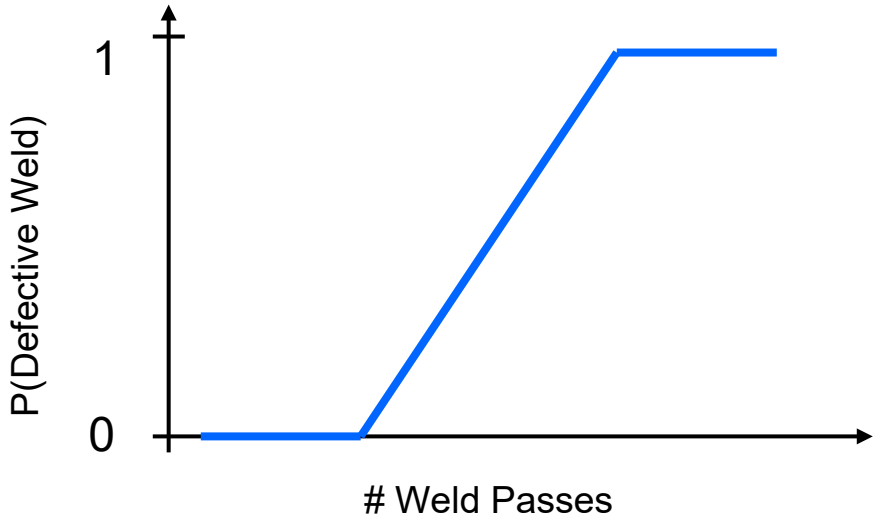
How should we interpret such a model? Since the response, *Y*, has only two possible outcomes, we have a *binomial* random variable (review the *Binomial Distribution* in Unit 9.1). Here, the probability of $Y_i = 1$ is designated by π; the probability of $Y_i = 0$ by $1 - \pi$. Returning to the regression equation above, we can develop the *expected value* of Y_i (see also Unit 9.1 for discussion of *Expected Values*):

$$E(Y_i) = \beta_o + \beta_o \times X_i \text{ but also: } E(Y_i) = 1 \times \pi_i + 0 \times (1 - \pi_i) = \pi_i$$

$$Therefore:$$

$$E(Y_i) = \beta_o + \beta_o \times X_i = \pi_i$$

The mean response $E(Y_i)$ then is simply the probability that $Y_i = 1$ when the independent variable is equal to X_i. A graph of this relationship appears on the next page. Here, the graph shows how the probability of having a defective weld increases as the number of passes required to complete the weld increases.

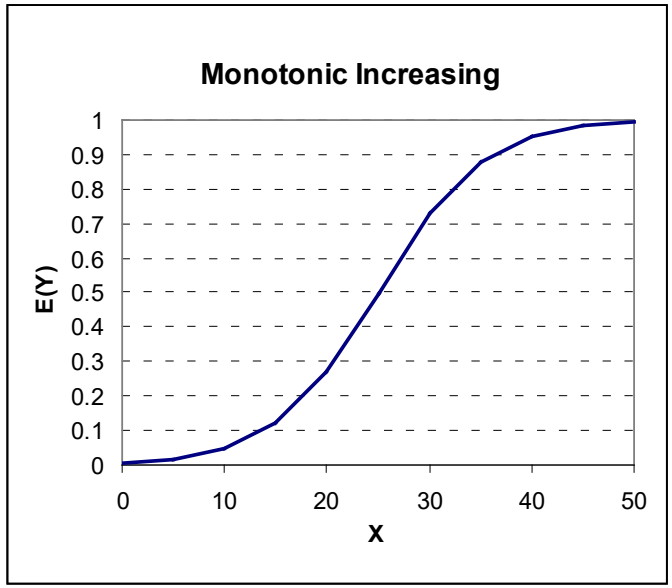


The Logistic Model

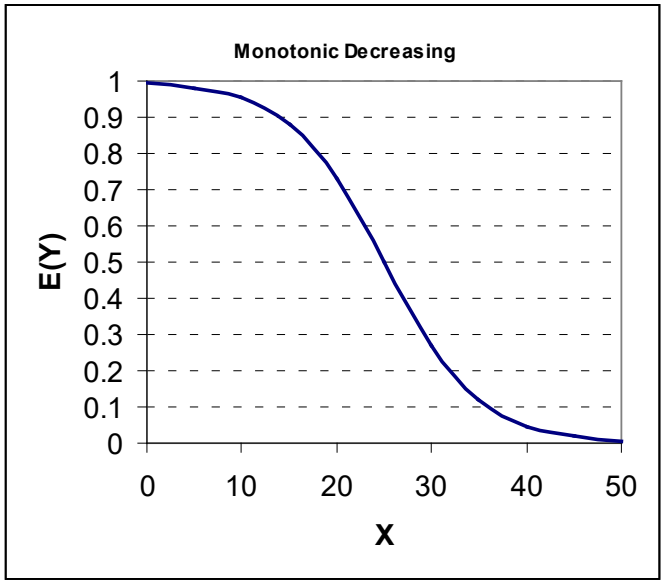
The graph above suggests that the Y_i response will *not* be a linear function of the X_i, but rather some curvilinear form. Desired properties of the function include a) that it is approximately linear, except at the ends and b) that it asymptotic to the 0 and 1 values of Y_i. *Sigmoidal* functions that fit these criteria have been developed and are known as *logistic response functions*. Such a function for one independent variable (X_i) is shown below. The values of β_0 and β_1 determine the direction of the function (increasing/decreasing) and its slope. Example graphs are shown on the next page.

$$E(Y) = \frac{\exp(\beta_0 + \beta_1 \times X)}{1 + \exp(\beta_0 + \beta_1 \times X)}$$

Logistic Response Functions



Equation: $E(Y) = \dfrac{\exp(-5 + (0.2) \times X)}{1 + \exp(-5 + (0.2) \times X)}$



Equation: $E(Y) = \dfrac{\exp(5 + (-0.2) \times X)}{1 + \exp(5 + (-.02) \times X)}$

As noted in the introduction to this section, there may be several independent variables that explain variation in the dependent variable. A general form of the logistic regression function is then:

$$E(Y) = \frac{\exp(\beta_0 + \sum_{i=1}^{n} \beta_i \times X_i)}{1 + \exp(\beta_0 + \sum_{i=1}^{n} \beta_i \times X_i)}$$

The Logistic Model

Earlier in this unit, we discussed the possibility of transforming data when the Y = f(X) relationship was not linear. For the logistic model, if we take the natural logarithm of both sides (known as a *logit transform*) and, by doing some rearranging, the following linear relationship results:

$$\ln(E(Y)/(1-E(Y)) = \beta_0 + \sum_{i=1}^{n} \beta_i \times X_i$$

Thus, the *logit transform* brings the rather complicated logistic regression model into our familiar world of linear regression. The logistic regression model is very robust in that it can handle a wide variety of data inputs that can be used as independent variables.

Fitting the Logistic Model

***Note**: Since the instructions are a bit complex, we'll intersperse them with a running example – weld defects.*

To develop a logistic regression model, we start by clarifying the structure of the data:

- There is a single dependent variable that can assume one of two possible outcomes (good/bad, yes/no, leaks/doesn't leak, etc.)
- A value of "1" is assigned to one of the outcomes (generally, the "desirable" outcome or the one of interest) and the value of "0" to the other. This assignment will help make the interpretation of the model more obvious.
- In logistic regression, the independent variables can either be measured on a **continuous scale,** or they can be **discrete variables.**
- The logistic regression model should *not* include independent variables measured on an **ordinal scale** (rank order). For these independent variables it is difficult to interpret or attribute values to the model coefficients (β_i).
- Data collection and organization occurs the same as we did for previous regressions (dependent variable outcome, associated values of independent variables).

Weld Defects Example

The dependent variable is the result of a visual inspection of a butt weld. The inspector classifies the weld as pass (assigned a 0) or fail (assigned a 1, since this is the variable of interest). For this example, we will use one variable – the number of passes required to complete the weld (a discrete variable). An example data pair – (1, 4) indicates that the weld was defective (1) and four (4) passes were needed to complete the weld.

Building the logistic regression model follows three steps (assuming there are multiple "X's" being considered):

1. Conduct a univariate analysis of all independent variables.

2. Run the Logistic Regression and include all independent variables with a p-value of $\leq$ 0.25.

3. Subtract variables until the model has only significant independent variables.

Step 1: Conduct a univariate analysis of all independent variables - Plot each independent variable by itself compared to the response variable to check for stratification, unusual observations, relationships, etc. For continuous independent variables use descriptive statistics, box plots, dot plots, charts, etc. For *continuous* independent variables, then fit the logistic regression model and get the estimated coefficient, estimated standard error, and univariate p-value. For discrete variables use a frequencies, tables, histograms, pie chart, etc.

Weld Defects Example

The data (collected from butt welds made over a one-week period) is organized by the number of passes required for each weld (recall – 0 is a good weld, 1 – a defective weld):

Weld Passes									
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
0	0	0	0	0	1	1	0	1	1
0	0	0	0	1	0	0	1	1	1
0	0	0	0	1	0	1	1	1	1
0	0	0	0	1	1	1	1	1	1
0	1	0	0	0	1	0	1	1	1
0	0	0	0	1	1	0	1	1	1
0	0	0	1	0	0	1	1	1	1
0	1	0	1	1	0	0	1	1	1
0	0	0	0	1	1	1	1	1	1
0	0	0	0	0	1	1	1	1	1

Here, simply organizing the data in the table above indicates that the more passes, the higher the probability of a weld being defective (this gives us some confidence that a logistic model will work).

Step 2: Run the Logistic Regression –To develop the final model, consider any variable whose p-value was less than 0.25 in the uni-variate analysis. Develop the regression model (we will use Minitab for this). Check the overall p-value. If it is not significant, you probably have the wrong independent variables in the model. Begin with a model containing all selected variables.

Step 3: Subtract and add independent variables one at a time - Remove insignificant variables from the model one by one. Note: You should not build the model using automatic variable selection procedures (a feature available in some statistical packages). The automatic variable selection sometimes makes the wrong selection. Verify the importance of each variable in the final model. Compare the coefficients in the model with those in the univariate model.

If the sign has changed for a variable, look for multi-collinearity problems. Multi-collinearity implies that the independent variables are not independent. Use the p-value for each independent variable to determine which are more likely to have an impact on the dependent variable. Usually, the variable with the highest p-value above 0.05 should be removed first. Eliminate non-contributing variables.

Refit the model and compare it with the old model using the log-likelihood ratio test (the log-likelihood ratio test will be explained soon.). Check Goodness-of-Fit to ensure that there is no lack of fit in any test (look for high (> alpha) p-values here. Check the *concordance* of each independent variable. If concordance is low, then look for other independent variables. Delete, refit, and verify until you are satisfied with the final model. Carefully examine the variables in the model and see if everything is logical based on process knowledge.

Weld Defects Example

In this case, there is only one independent variable (number of weld passes). See the later, "multiple-X" example for illustration of these instructions.

Interpreting the Logistic Regression Model

Interpretation of Coefficients – For logistic regression interpreting the coefficients is a bit more difficult than with simple or multiple regression. The estimated slope coefficient, β, gives the change in the natural logarithm of the estimated logit for a unit change in x. If the coefficient is significant in the model, then mainly look at the sign of the coefficient. This tells you about the directional relationship with the variable of interest.

Weld Defects Example

***Minitab Output* - Binary Logistic Regression: Fail versus Passes**

```
Link Function:  Logit

Response Information

Variable  Value       Count
Fail      1              51  (Event)
          0              49
          Total         100

Logistic Regression Table
                                                    Odds        95% CI
Predictor       Coef    SE Coef        Z      P    Ratio    Lower    Upper
Constant     -4.3017     0.8451    -5.09 0.000
Passes        0.7971     0.1469     5.43 0.000     2.22     1.66     2.96

Log-Likelihood = -38.543
Test that all slopes are zero: G = 61.504, DF = 1, P-Value = 0.000

Goodness-of-Fit Tests

Method                Chi-Square    DF      P
Pearson                    9.223     8  0.324
Deviance                  10.187     8  0.252
Hosmer-Lemeshow            9.223     8  0.324
```

1
2
3
4

Interpretation:

1. The observation counts are listed here. There were 51 “Fail” events (defective welds, assigned a value of 1) and 49 successful (good welds, assigned a value of 0).

2. The coefficients of the regression equation are shown in the Logistic Regression table above. The form of the regression equation is $\ln(E(Y)/(1-E(Y)) = \beta_0 + \beta_1 X$. For this data, β_0 = -4.30 (constant) and β_1 = 0.797 (passes). A coefficient associated with an independent variable represents the estimated change in the $\ln(E(Y)/(1-E(Y))$ (where the event is failed weld = 1). The odds ratio is $E(Y)/(1-E(Y)$. One way to look at this is to remember that

E(Y) is the probability of the event (P(event)), *weld defective,* or π. Thus $\ln(E(Y)/(1-E(Y))$ is equivalent to P(event)/(1 – P(event) or $\pi/(1 - \pi)$. This is the ratio of the probability that the event occurs to the probability that it doesn't occur, or its "odds" of occurring. For logistic regression models where multiple "X" factors are considered, higher odds ratios indicate more important factors.

3. The null hypothesis for the regression is that the slopes of the regression line = 0 (i.e. all coefficients = 0), the alternative is that they are not equal to 0. The p-value of 0.000 indicates the null hypothesis is rejected – the factor of *weld pass* explains the variation in the Y (leakage).

4. These statistics show how well the model fits the data. The null hypothesis here is that the model fits the data; the alternate is that it does not. All three goodness of fit tests (Pearson, Deviance & Hosmer-Lemeshow) do not reject the null (p-values > 0.05). Hence, the factor of *weld pass* is significant and the overall model fits the data.

***Minitab Output* - Binary Logistic Regression: Fail versus Passes (Continued)**

Table of Observed and Expected Frequencies:
(See Hosmer-Lemeshow Test for the Pearson Chi-Square Statistic)

5

					Group						
Value		1	2	3	4	5	6	7	8	9	10	Total
1												
	Obs	0	2	0	2	6	6	6	9	10	10	51
	Exp	0.3	0.6	1.3	2.5	4.2	6.2	7.8	8.9	9.5	9.8	
0												
	Obs	10	8	10	8	4	4	4	1	0	0	49
	Exp	9.7	9.4	8.7	7.5	5.8	3.8	2.2	1.1	0.5	0.2	
	Total	10	10	10	10	10	10	10	10	10	10	100

6

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	2208	88.4%	Somers' D	0.81
Discordant	178	7.1%	Goodman-Kruskal Gamma	0.85
Ties	113	4.5%	Kendall's Tau-a	0.41
Total	2499	100.0%		

5. The Observed and Expected Frequencies table calculates the expected number of observations for events 1 (defective weld) and 0 (good weld) and then compares these to the actual (observed) frequencies) for ten groups of

the data. Here, the observed data is not that far off from the expected frequencies (i.e. the model is good). This is an "informal" goodness of fit test.

6. Measures of Association display a table of the number and percentage of concordant, discordant, and tied pairs, as well as common rank correlation statistics. These values measure the association between the observed responses and the predicted probabilities.

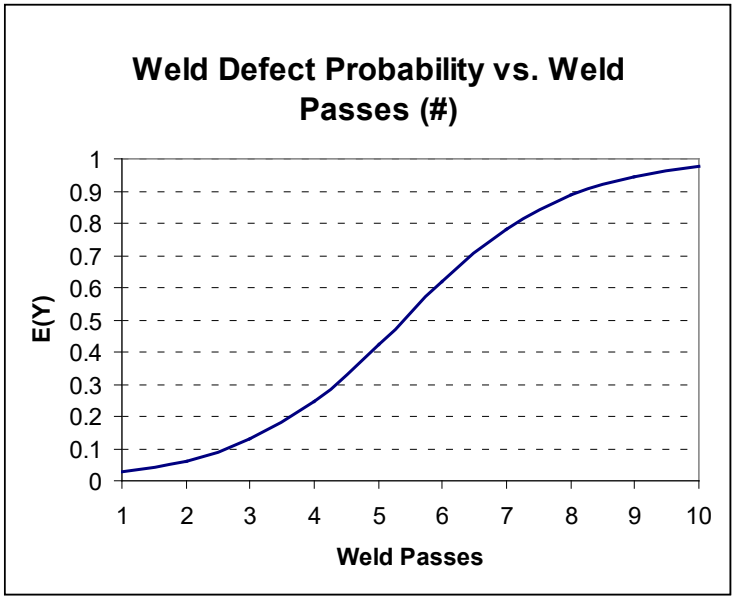
   The table of concordant, discordant, and tied pairs is calculated by pairing the observations with different response values. In this example, 88.4% of pairs are concordant, 7.1% are discordant and 4.5% are ties. You can use these values as a comparative measure of prediction, for example in comparing fits with different sets of independent variables or with different link functions.

   Somers' D, Goodman-Kruskal Gamma, and Kendall's Tau-a are summaries of the table of concordant and discordant pairs. These measures most likely lie between 0 and 1. Larger values indicate that the model has a better predictive ability. In this example, the measure range from 0.41 to 0.81 which implies good predictive ability.

We can now present the logistic regression model and associated graph for *defective butt welds* as a function of *weld passes.* The probability of a defective weld increases with number of weld passes. From the graph below, a weld requiring 6 passes has about a 60% chance that it will be declared defective.

Equation:

$$E(Y) = \frac{\exp(-4.30 + 0.797 \times X)}{1 + \exp(-4.30 + 0.797 \times X)}$$



Logistic Regression Cautions

Logistic regression is not appropriate in the following situations:

- A simpler test will provide the necessary verification (i.e. a contingency table/chi-Square analysis).
- If, when the observed outcome data is plotted against an independent variable, the graph appears like a 'horseshoe" or a rounded peak.
- The data does not appear to be increasing or decreasing around the area of interest for the independent variable.
    - E.g. fertilizer on crops; None - low yield; some - good yield; Too much - low yield

Possible solutions in these cases may be to:

- Pick the region for analysis,
- Remove that independent variable from the model, or
- Select a different model.

Logistic Regression Example – Multiple "X's"

A manufacturer of feminine products received a certain number of complaints about sanitary napkins that leaked when used. The company commissioned a Black Belt & Six Sigma team to study factors that contribute to leakage. One study focused on female characteristics – such as Body Mass Index (BMI), Size of Pants/Slacks worn, Thigh "Gap" (distance between thighs), and the age of the woman wearing the napkin. The response (Y) is whether the napkin leaked. The team decided to collect data and thought that a logistic regression would be appropriate to analyze the data.

Data Structure

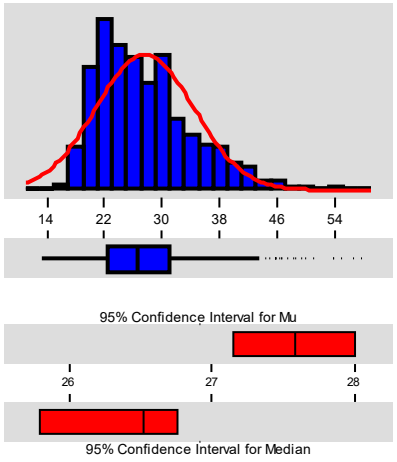
Variable	***Definition/Measure***
Leakage (Y)	*Women were asked to record (in a diary) whether the napkin leaked (assigned a 1) or not (assigned a 0)*
Body Mass Index (BMI)	*The Body Mass Index is a measure of weight to height (defined as BMI = kg/m^2). The range of BMI's in the study was 13 – 57)*
Size Pants/Slacks	*The size of the woman's pants/slacks – measured using the "standard" fashion industry scale. For this study, slack sizes ranges from 2 – 48.*
Thigh Gap	*The distance between the woman's inner thigh, measured at the top of the leg. For this study the gap ranged from 47 to 0 mm.*
Age	*The woman's age. For this study, the ages ranged from 16 – 49.*

The data gathered appears at the end of this section. One thousand women were randomly sampled and measured (*BMI, Size Pants/Slacks, Thigh Gap* and *Age*) and given 10 napkins to wear. If at least one of the napkins leaked during the test, the *leakage* variable was scored a 1, if no leakage, 0.

Univariate Analysis

First, basic statistics and histograms of the "X" independent variables were obtained (see below). Compare the distributions of the variables to the ranges noted in the table above. It's obvious that the random sampling process resulted in unequal numbers of women across the variable's range. This may influence the logistic model in that it will be difficult to estimate the probability of leakage at the extremes of the distributions (there may not be enough leakage events occurring at the extremes, simply because there aren't enough women and the probability of having a leakage event may not be high enough).

Descriptive Statistics



Variable: BMI

Anderson-Darling Normality Test

A-Squared:	15.974
P-Value:	0.000
Mean	27.5787
StDev	6.7929
Variance	46.1429
Skewness	1.01511
Kurtosis	1.13777
N	1000
Minimum	13.3100
1st Quartile	22.3100
Median	26.5200
3rd Quartile	31.0100
Maximum	57.4100

95% Confidence Interval for Mu

27.1572	28.0002

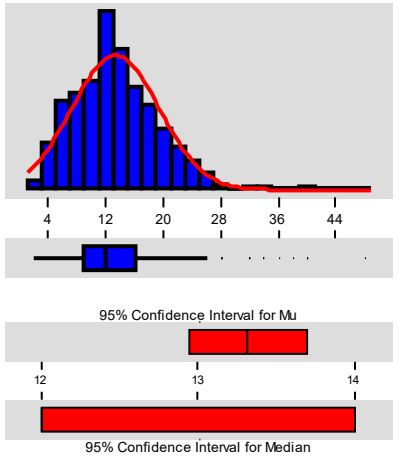
95% Confidence Interval for Sigma

6.5076	7.1044

95% Confidence Interval for Median

25.7900	26.7533

Descriptive Statistics



Variable: SlackSize

Anderson-Darling Normality Test

A-Squared:	7.774
P-Value:	0.000
Mean	13.3200
StDev	6.0690
Variance	36.8324
Skewness	0.836379
Kurtosis	1.97982
N	1000
Minimum	2.0000
1st Quartile	9.0000
Median	12.0000
3rd Quartile	16.0000
Maximum	48.0000

95% Confidence Interval for Mu

12.9434	13.6966

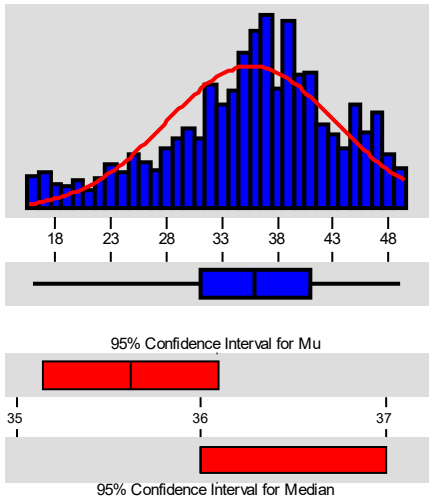
95% Confidence Interval for Sigma

5.8141	6.3473

95% Confidence Interval for Median

12.0000	14.0000

Descriptive Statistics



Variable: Age

Anderson-Darling Normality Test

A-Squared:	7.742
P-Value:	0.000
Mean	35.6220
StDev	7.6617
Variance	58.7018
Skewness	-5.4E-01
Kurtosis	-1.2E-01
N	1000
Minimum	16.0000
1st Quartile	31.0000
Median	36.0000
3rd Quartile	41.0000
Maximum	49.0000

95% Confidence Interval for Mu

35.1466	36.0974

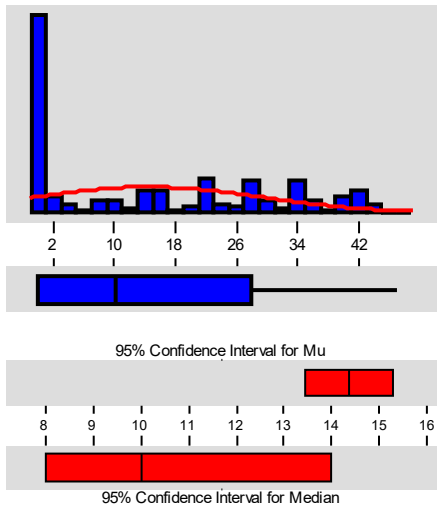
95% Confidence Interval for Sigma

7.3400	8.0131

95% Confidence Interval for Median

36.0000	37.0000

Descriptive Statistics



Variable: ThighGap

Anderson-Darling Normality Test

A-Squared:	64.433
P-Value:	0.000
Mean	14.3640
StDev	15.0070
Variance	225.211
Skewness	0.492853
Kurtosis	-1.25199
N	1000
Minimum	0.0000
1st Quartile	0.0000
Median	10.0000
3rd Quartile	28.0000
Maximum	47.0000

95% Confidence Interval for Mu

13.4327	15.2953

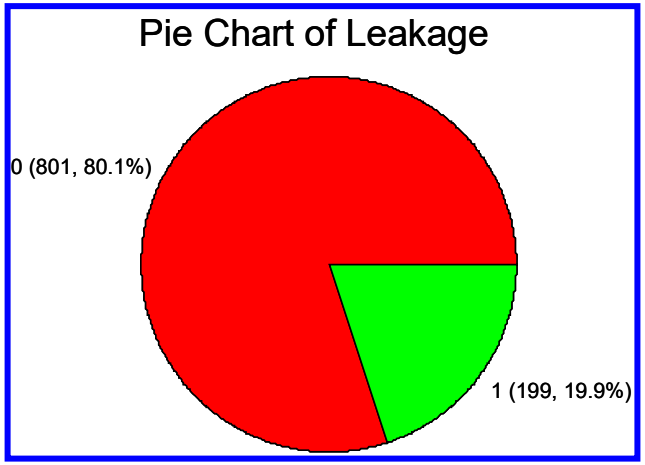
95% Confidence Interval for Sigma

14.3769	15.6953

95% Confidence Interval for Median

8.0000	14.0000

Next, a pie chart of leakage was developed. About 20% of the women experienced leakage during the test.



Logistic regression is now run for each of the variables by itself. Each factor is significant by itself, so it will be included in the “draft” final model (see following pages – Minitab Output).

Binary Logistic Regression: Leakage versus BMI

```
Response Information
Variable  Value       Count
Leakage   1             199  (Event)
          0             801
          Total        1000

Logistic Regression Table
                                                    Odds         95% CI
Predictor       Coef    SE Coef         Z      P    Ratio    Lower    Upper
Constant     -3.5516     0.3402   -10.44 0.000
BMI          0.07557    0.01122     6.74 0.000     1.08     1.06     1.10

Log-Likelihood = -475.934
Test that all slopes are zero: G = 46.157, DF = 1, P-Value = 0.000

Goodness-of-Fit Tests
Method                Chi-Square    DF      P
Pearson                  364.177   374  0.632
Deviance                 398.804   374  0.181
Hosmer-Lemeshow           33.269     8  0.000

Table of Observed and Expected Frequencies:

                                    Group
Value      1     2     3     4     5     6     7     8     9    10   Total
1
  Obs      2     7    11    14    24    28    29    29    35    20     199
  Exp     10.7  13.1  13.7  14.9  18.4  18.8  21.4  24.6  30.8  32.6
0
  Obs     98    99    90    87    86    72    73    73    67    56     801
  Exp     89.3  92.9  87.3  86.1  91.6  81.2  80.6  77.4  71.2  43.4

  Total  100   106   101   101   110   100   102   102   102    76    1000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)
Pairs           Number  Percent     Summary Measures
Concordant      107641    67.5%     Somers' D               0.36
Discordant       49760    31.2%     Goodman-Kruskal Gamma   0.37
Ties              1998     1.3%     Kendall's Tau-a         0.12
Total           159399   100.0%
```

Binary Logistic Regression: Leakage versus SlackSize

```
Response Information
Variable  Value       Count
Leakage   1             199  (Event)
          0             801
          Total        1000

Logistic Regression Table
                                                    Odds         95% CI
Predictor       Coef    SE Coef         Z      P    Ratio    Lower    Upper
Constant     -2.8170     0.2201   -12.80 0.000
SlackSiz     0.09952    0.01350     7.37 0.000     1.10     1.08     1.13

Log-Likelihood = -469.337
Test that all slopes are zero: G = 59.351, DF = 1, P-Value = 0.000

Goodness-of-Fit Tests
Method                Chi-Square    DF      P
Pearson                   43.072    26  0.019
Deviance                  48.567    26  0.005
Hosmer-Lemeshow           24.211     7  0.001

Table of Observed and Expected Frequencies:

                                  Group
Value      1     2     3     4     5     6     7     8     9   Total
1
  Obs      3    10    11    32    34    28    50    30     1     199
  Exp     12.6  12.6  14.5  26.9  26.7  22.7  40.1  42.0   0.9
0
  Obs    139    99    92   131   104    72    93    71     0     801
  Exp    129.4  96.4  88.5 136.1 111.3  77.3 102.9  59.0   0.1

  Total  142   109   103   163   138   100   143   101     1    1000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)
Pairs           Number  Percent     Summary Measures
Concordant      103336    64.8%     Somers' D               0.38
Discordant       42379    26.6%     Goodman-Kruskal Gamma   0.42
Ties             13684     8.6%     Kendall's Tau-a         0.12
Total           159399   100.0%
```

Binary Logistic Regression: Leakage versus Age

```
Response Information
Variable  Value       Count
Leakage   1             199  (Event)
          0             801
          Total        1000

Logistic Regression Table
                                                    Odds         95% CI
Predictor       Coef    SE Coef        Z      P    Ratio    Lower    Upper
Constant     -2.8979     0.4205    -6.89 0.000
Age          0.04142    0.01115     3.72 0.000     1.04     1.02     1.07

Log-Likelihood = -491.698
Test that all slopes are zero: G = 14.629, DF = 1, P-Value = 0.000

Goodness-of-Fit Tests
Method                Chi-Square    DF      P
Pearson                   31.485    32  0.492
Deviance                  35.242    32  0.317
Hosmer-Lemeshow            6.635     7  0.468

Table of Observed and Expected Frequencies:

                                  Group
Value      1     2     3     4     5     6     7     8     9   Total
1
  Obs     12    14    18    24    31    27    22    32    19     199
  Exp     13.9  16.4  18.8  18.8  27.2  24.2  30.0  29.8  19.9
0
  Obs    107    94    91    76   105    86   108    83    51     801
  Exp    105.1  91.6  90.2  81.2 108.8  88.8 100.0  85.2  50.1

  Total  119   108   109   100   136   113   130   115    70    1000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)
Pairs           Number  Percent     Summary Measures
Concordant       89576    56.2%     Somers' D               0.16
Discordant       63347    39.7%     Goodman-Kruskal Gamma   0.17
Ties              6476     4.1%     Kendall's Tau-a         0.05
Total           159399   100.0%
```

Binary Logistic Regression: Leakage versus ThighGap

```
Response Information
Variable  Value       Count
Leakage   1             199  (Event)
          0             801
          Total        1000

Logistic Regression Table
                                                    Odds         95% CI
Predictor       Coef    SE Coef        Z      P    Ratio    Lower    Upper
Constant    -0.79484    0.09818    -8.10 0.000
ThighGap    -0.055221   0.007031   -7.85 0.000     0.95     0.93     0.96

Log-Likelihood = -459.387
Test that all slopes are zero: G = 79.252, DF = 1, P-Value = 0.000

Goodness-of-Fit Tests
Method                 Chi-Square    DF      P
Pearson                    65.893    41  0.008
Deviance                   47.315    41  0.230
Hosmer-Lemeshow             3.695     4  0.449

Table of Observed and Expected Frequencies:

                       Group
Value      1     2     3     4     5     6   Total
1
  Obs      2     9    11    23    30   124     199
  Exp      4.7   8.2  12.6  19.3  26.1 128.1
0
  Obs    102   110   114   106    81   288     801
  Exp     99.3 110.8 112.4 109.7  84.9 283.9

  Total  104   119   125   129   111   412    1000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)
Pairs           Number  Percent     Summary Measures
Concordant       91251    57.2%     Somers' D               0.37
Discordant       32872    20.6%     Goodman-Kruskal Gamma   0.47
Ties             35276    22.1%     Kendall's Tau-a         0.12
Total           159399   100.0%
```

"Full" Logistic Regression Model

The logistic regression is now run with all independent variables included (since the univariate analysis indicated that they are all significant). The Minitab output appears below:

Binary Logistic Regression: Leakage versus BMI, SlackSize, Age, ThighGap

```
Response Information
Variable  Value       Count
Leakage   1             199  (Event)
          0             801
          Total        1000

Logistic Regression Table
                                                    Odds        95% CI
Predictor       Coef    SE Coef        Z      P    Ratio    Lower    Upper
Constant     -0.9625     0.8249    -1.17  0.243
BMI          -0.04825    0.02720   -1.77  0.076     0.95     0.90     1.01
SlackSiz      0.06699    0.02833    2.36  0.018     1.07     1.01     1.13
Age           0.01474    0.01203    1.22  0.221     1.01     0.99     1.04
ThighGap     -0.05068    0.01063   -4.77  0.000     0.95     0.93     0.97

Log-Likelihood = -455.057
Test that all slopes are zero: G = 87.912, DF = 4, P-Value = 0.000

Goodness-of-Fit Tests
Method                 Chi-Square    DF      P
Pearson                   832.717   864  0.772
Deviance                  785.835   864  0.973
Hosmer-Lemeshow            11.186     8  0.191

Table of Observed and Expected Frequencies:
                                   Group
Value      1     2     3     4     5     6     7     8     9    10   Total
1
  Obs      1     8     6    13    26    21    29    32    27    36     199
  Exp    4.2   6.4   9.0  12.7  17.3  23.3  27.4  29.9  32.1  36.9
0
  Obs    100    93    94    88    74    79    71    69    73    60     801
  Exp   96.8  94.6  91.0  88.3  82.7  76.7  72.6  71.1  67.9  59.1

  Total  101   101   100   101   100   100   100   101   100    96    1000
```

With the Full-Model, two factors (Slack Size and Thigh Gap) are significant. The other factors may be removed from the model (suggestion: do this one at a time)

Goodness of Fit Tests indicate that leakage is explained by the factors in the model.

Following the above suggestion, the model is then re-run with Age factor removed (Age had the highest p-value):

Binary Logistic Regression: Leakage versus BMI, SlackSize, ThighGap

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI Lower	95% CI Upper
Constant	-0.3342	0.6446	-0.52	0.604			
BMI	**-0.05377**	**0.02685**	**-2.00**	**0.045**	**0.95**	**0.90**	**1.00**
SlackSiz	**0.07314**	**0.02786**	**2.63**	**0.009**	**1.08**	**1.02**	**1.14**
ThighGap	**-0.05269**	**0.01049**	**-5.02**	**0.000**	**0.95**	**0.93**	**0.97**

Log-Likelihood = -455.817
Test that all slopes are zero: G = 86.392, DF = 3, P-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	**728.467**	**749**	**0.698**
Deviance	**673.907**	**749**	**0.977**
Hosmer-Lemeshow	**10.768**	**8**	**0.215**

Table of Observed and Expected Frequencies:

	Group										
Value	1	2	3	4	5	6	7	8	9	10	Total
1											
Obs	1	9	4	15	23	26	26	27	31	37	199
Exp	4.2	6.3	8.9	12.5	16.9	23.8	28.1	29.6	31.6	37.2	
0											
Obs	99	91	96	85	77	74	75	73	69	62	801
Exp	95.8	93.7	91.1	87.5	83.1	76.2	72.9	70.4	68.4	61.8	
Total	100	100	100	100	100	100	101	100	100	99	1000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures	
Concordant	110991	69.6%	Somers' D	0.40
Discordant	47127	29.6%	Goodman-Kruskal Gamma	0.40
Ties	1281	0.8%	Kendall's Tau-a	0.13
Total	159399	100.0%		

All Factors are now significant and the model "fits" the data

BMI has "switched signs" – its univariate coefficient was 0.076, now it is -0.054 – this may indicate a multi-collinearity problem – for example, Slack Size and Thigh Gap may be related to BMI (factors are not really independent of each other).

Final Model & Conclusion

The final model for leakage may be expressed as follows:

$$E(Y) = \frac{\exp(-0.334 + (-0.054) \times BMI + 0.073 \times SlackSize + (-0.053) \times ThighGap)}{1 + \exp(-0.334 + (-0.054) \times BMI + 0.073 \times SlackSize + (-0.053) \times ThighGap)}$$

The odds ratios give us a clue as to the importance of each predictor. Slack Size has the highest odds ratio – the probability of leakage/probability of no leakage is highest. This appears to be the most important factor affecting leakage. Body Mass Index and Thigh Gap are about the same in importance, but the confidence interval for BMI is wider than Thigh Gap. As noted in the univariate analysis, the "tails" of the independent variable distributions are not well represented – the team should consider gathering additional data from these "tails." One of the factors may then arise as a stronger predictor.

Predictor	Odds Ratio	95% CI	
		Lower	Upper
BMI	0.95	0.90	1.00
Slack Size	1.08	1.02	1.14
Thigh Gap	0.95	0.93	0.97

Leakage Data

BMI	SlackSize	Age	ThighGap	Leakage
13.31	8	36	47	1
16.31	1	22	42	0
16.60	12	24	43	0
16.73	4	23	44	0
16.83	2	35	44	0
17.14	4	18	44	0
17.14	4	18	42	0
17.14	4	18	42	0
17.14	4	18	41	0
17.22	6	31	43	0
17.46	12	39	42	0
17.74	2	32	41	0
17.79	4	35	42	0
17.81	5	42	42	0
17.94	4	46	44	0
17.94	5	32	41	0
17.95	1	22	44	0
18.01	4	29	42	0
18.01	5	24	43	0
18.02	1	17	42	0
18.02	3	29	42	0
18.02	4	29	41	0
18.29	4	30	40	0
18.29	3	43	43	0
18.30	7	17	42	0
18.30	3	17	41	0
18.30	4	25	40	0
18.30	7	17	42	0
18.46	8	41	43	0
18.46	6	32	43	0
18.48	6	29	42	0
18.49	9	22	40	0
18.55	2	37	42	0
18.55	2	37	41	0
18.60	5	19	42	0
18.60	5	19	42	0
18.79	9	47	40	0
18.88	3	37	42	0
18.88	6	35	40	0
18.88	6	35	42	0
18.88	6	35	41	0
18.88	4	29	41	0
19.01	4	32	41	0
19.11	8	37	40	0
19.13	2	44	43	0
19.14	6	36	42	0
19.14	3	19	41	0
19.19	1	36	43	0
19.20	5	33	40	0
19.20	5	25	40	0
19.20	2	35	39	0
19.20	4	20	42	0
19.20	12	40	43	0
19.21	7	21	40	0
19.22	10	46	42	0
19.37	6	42	41	0
19.37	5	32	41	0
19.37	5	22	39	0
19.37	10	47	41	0
19.40	3	27	43	0
19.46	9	16	40	0
19.47	6	29	40	0
19.53	6	39	41	0
19.53	6	30	39	0
19.53	6	30	41	0
19.58	6	36	40	0
19.58	6	36	40	0
19.58	6	36	41	0
19.69	7	36	39	0
19.73	0	22	40	0
19.74	6	42	40	0
19.74	6	42	40	0
19.74	5	16	40	0
19.74	6	42	42	0
19.74	4	28	38	0
19.74	5	16	41	0
19.74	3	20	40	0
19.74	3	20	40	0
19.74	5	16	39	0
19.79	2	36	42	0
19.84	7	37	41	0
19.84	4	32	38	0
19.84	7	37	40	0
19.84	4	32	41	0
19.85	6	26	40	0
19.92	2	29	40	0
19.94	10	38	41	0
19.97	7	16	40	0
19.97	6	40	40	0
19.97	18	18	41	0
19.97	7	43	40	0
20.05	6	36	39	0
20.09	7	31	39	0
20.22	10	40	41	0
20.25	4	36	40	0
20.30	6	36	38	0
20.30	6	36	40	0
20.30	6	36	43	0
20.30	10	32	34	1
20.30	6	36	34	0
20.36	7	32	34	0
20.36	6	25	34	0
20.36	14	17	34	0
20.36	7	32	34	0
20.36	7	32	35	0
20.36	8	41	35	0
20.36	6	28	36	1
20.36	8	42	33	0
20.36	7	32	36	0
20.37	3	36	32	0
20.37	4	30	34	0
20.37	3	36	34	0
20.37	8	22	34	0
20.37	3	36	34	0
20.37	6	38	34	0
20.41	8	30	34	0
20.47	6	43	34	0
20.51	4	35	34	0
20.53	8	32	34	0
20.55	7	45	35	0
20.55	8	33	35	0
20.55	7	45	34	0
20.55	8	33	34	0
20.55	8	33	34	0
20.60	8	37	34	0
20.60	8	38	35	0
20.60	10	25	35	0
20.60	10	25	34	0
20.66	8	39	33	0
20.68	9	32	35	0
20.70	8	32	34	0
20.70	6	21	34	0

BMI	SlackSize	Age	ThighGap	Leakage
20.73	8	49	34	0
20.80	4	34	35	0
20.80	12	45	34	0
20.82	12	21	37	0
20.82	8	26	35	0
20.83	6	22	33	0
20.90	4	35	34	0
20.90	6	32	34	0
20.90	4	27	33	0
20.90	5	30	34	0
20.97	10	26	33	0
20.97	10	26	36	0
20.97	10	26	35	0
20.97	10	26	35	1
20.98	8	34	34	1
20.98	8	31	33	0
20.99	10	29	35	0
21.03	7	29	34	1
21.03	4	32	34	0
21.03	7	29	33	0
21.09	4	36	35	0
21.11	7	42	35	0
21.14	10	34	34	0
21.14	10	34	34	0
21.14	8	23	35	0
21.14	12	47	33	0
21.14	10	45	34	0
21.14	10	38	35	0
21.14	10	38	33	0
21.22	6	39	34	0
21.26	6	22	34	0
21.26	6	34	32	0
21.26	6	34	33	0
21.26	6	34	34	0
21.29	8	39	34	0
21.29	12	32	36	0
21.29	10	40	35	0
21.29	10	40	34	0
21.30	8	40	32	0
21.30	8	40	33	1
21.41	12	25	34	0
21.41	2	45	33	0
21.46	8	41	33	0
21.46	8	30	32	0
21.46	8	41	31	0
21.46	8	37	32	1
21.46	6	25	36	0

BMI	SlackSize	Age	ThighGap	Leakage
21.46	3	22	33	0
21.46	8	37	33	0
21.48	6	26	34	0
21.48	4	38	35	0
21.48	6	26	34	0
21.52	12	30	34	0
21.52	12	45	33	0
21.58	6	35	34	0
21.58	6	35	33	0
21.61	8	35	34	0
21.61	10	33	32	0
21.61	4	32	35	0
21.63	6	32	33	0
21.63	8	41	34	0
21.70	14	48	34	0
21.73	7	30	32	0
21.73	6	40	34	0
21.73	6	40	32	0
21.73	8	32	33	0
21.79	9	27	33	0
21.79	8	43	28	0
21.79	6	27	29	0
21.79	10	17	28	0
21.79	6	27	30	0
21.79	12	34	30	0
21.79	6	22	29	0
21.79	12	34	27	1
21.80	7	33	28	0
21.80	8	36	28	0
21.83	12	21	29	0
21.83	12	21	28	0
21.93	8	23	29	0
21.93	10	41	28	0
21.93	12	36	28	0
21.93	10	35	28	0
21.93	11	20	29	0
21.93	12	36	28	0
21.93	10	41	27	0
21.95	6	28	29	0
21.95	5	44	28	0
21.95	3	23	28	0
21.95	6	33	29	0
21.97	6	38	29	0
21.97	6	27	28	0
21.97	10	38	26	0
22.05	8	19	28	0
22.05	10	25	28	0

BMI	SlackSize	Age	ThighGap	Leakage
22.05	10	25	27	0
22.05	10	25	27	0
22.13	5	40	28	0
22.14	9	39	28	0
22.14	5	32	26	0
22.14	5	32	29	0
22.14	7	39	28	0
22.14	12	32	28	0
22.14	6	35	28	0
22.22	4	38	29	0
22.22	4	45	29	0
22.31	5	17	28	0
22.31	8	35	29	0
22.31	10	31	29	1
22.31	5	17	27	0
22.31	5	17	29	0
22.31	11	40	28	0
22.31	16	49	28	0
22.31	6	34	28	0
22.31	10	35	26	0
22.31	16	49	28	0
22.31	6	27	29	1
22.31	12	41	29	0
22.31	8	45	28	0
22.31	10	31	26	0
22.32	6	35	29	0
22.46	14	38	28	1
22.46	9	20	27	0
22.46	9	20	30	0
22.46	7	46	27	1
22.50	5	37	27	0
22.50	5	37	27	1
22.50	9	45	27	1
22.50	9	45	28	0
22.50	9	40	26	0
22.50	5	37	26	0
22.50	9	45	27	0
22.60	10	46	29	0
22.60	10	44	28	0
22.60	12	26	27	0
22.60	15	19	28	1
22.60	6	29	28	0
22.60	12	26	28	0
22.66	8	39	28	0
22.67	7	33	27	0
22.67	8	20	28	0
22.67	7	48	27	0

BMI	SlackSize	Age	ThighGap	Leakage
22.67	10	41	27	0
22.67	7	48	27	0
22.68	7	44	29	1
22.71	14	34	27	0
22.71	14	34	28	1
22.71	12	38	27	0
22.71	12	41	28	0
22.71	12	38	27	0
22.71	14	41	25	0
22.71	14	34	27	0
22.74	11	16	28	0
22.80	10	41	28	0
22.81	11	39	26	0
22.81	11	39	28	0
22.81	8	36	27	0
22.86	8	48	28	0
22.86	10	38	27	0
22.86	10	38	27	0
22.89	14	27	27	0
22.92	10	42	30	0
22.92	11	49	27	0
22.96	8	36	29	0
22.96	15	20	28	0
22.96	8	36	27	0
23.02	11	38	26	1
23.02	11	38	26	1
23.03	8	28	22	0
23.03	8	28	22	0
23.03	10	33	23	0
23.03	8	31	21	0
23.03	8	28	22	0
23.03	10	43	21	0
23.05	6	34	22	0
23.05	6	44	23	0
23.17	7	35	22	0
23.17	9	23	22	0
23.17	14	41	22	0
23.17	12	48	22	0
23.17	8	32	22	0
23.17	10	40	22	0
23.18	12	41	23	0
23.18	12	41	22	0
23.18	12	41	22	0
23.24	12	47	21	1
23.30	12	47	22	0
23.30	14	29	22	0
23.30	12	28	20	0

BMI	SlackSize	Age	ThighGap	Leakage
23.30	8	38	22	0
23.30	12	28	21	0
23.34	10	30	22	1
23.34	10	30	21	0
23.38	10	31	23	0
23.38	10	31	21	0
23.40	12	36	21	1
23.40	12	41	22	1
23.40	12	47	21	0
23.43	10	42	22	0
23.44	8	40	21	0
23.44	12	41	21	0
23.44	5	33	21	0
23.44	12	41	21	0
23.44	8	40	22	0
23.49	12	31	22	0
23.49	10	17	21	0
23.49	12	33	21	0
23.49	10	23	23	0
23.49	12	37	23	0
23.49	10	17	25	0
23.57	14	34	21	0
23.57	12	37	20	0
23.57	14	43	23	0
23.57	12	46	21	0
23.57	12	37	21	0
23.59	3	18	22	0
23.62	8	30	21	0
23.63	12	30	23	0
23.63	12	40	20	1
23.67	12	37	24	0
23.67	12	37	21	1
23.69	10	35	21	0
23.71	12	25	21	0
23.73	14	34	21	0
23.73	12	41	20	0
23.74	16	42	20	0
23.74	16	42	23	0
23.78	7	36	21	0
23.78	10	38	22	0
23.78	8	32	22	0
23.78	8	45	22	0
23.86	10	35	22	0
23.91	8	40	21	0
23.91	12	37	23	0
23.91	6	27	23	0
23.91	8	38	21	0

BMI	SlackSize	Age	ThighGap	Leakage
24.03	12	30	21	0
24.03	12	25	21	0
24.03	12	30	22	0
24.03	12	33	20	0
24.03	10	40	23	0
24.03	10	25	24	0
24.03	12	33	21	0
24.03	10	47	20	0
24.03	8	36	20	1
24.03	8	40	21	0
24.09	10	47	20	0
24.12	10	45	20	0
24.13	12	39	21	0
24.13	10	37	21	0
24.13	12	47	22	0
24.13	12	19	22	0
24.13	12	28	22	0
24.13	12	47	21	1
24.13	12	39	21	1
24.13	12	19	23	0
24.27	10	39	20	0
24.27	10	39	23	0
24.28	10	28	22	0
24.28	10	28	22	0
24.28	12	29	23	0
24.33	10	32	20	0
24.33	10	32	21	0
24.37	12	29	21	0
24.41	8	32	21	0
24.55	14	40	20	0
24.56	6	24	21	0
24.63	10	22	20	0
24.66	12	47	15	0
24.69	12	44	17	1
24.69	16	42	14	0
24.69	10	41	14	1
24.69	12	44	14	1
24.69	12	44	17	0
24.80	7	45	16	0
24.80	14	45	16	1
24.80	14	45	15	1
24.86	12	36	15	0
24.87	12	28	15	0
24.87	10	38	15	0
24.87	12	28	16	0
24.87	10	38	15	0
24.89	10	38	14	0

BMI	SlackSize	Age	ThighGap	Leakage
24.89	14	48	14	0
24.89	14	16	15	0
24.89	10	31	15	0
24.89	10	35	15	0
24.96	14	40	16	0
24.96	18	23	15	0
24.96	8	32	14	0
24.96	10	35	15	0
25.02	10	36	15	0
25.02	12	37	15	0
25.02	12	44	14	1
25.02	12	37	14	0
25.02	12	40	14	0
25.02	12	37	15	0
25.02	12	37	16	0
25.06	14	26	17	0
25.06	14	26	16	1
25.06	12	24	16	1
25.06	12	37	13	0
25.06	12	37	15	0
25.06	12	37	14	0
25.09	12	30	16	0
25.09	14	47	16	0
25.09	14	35	16	1
25.10	12	33	15	0
25.23	8	32	14	0
25.39	8	40	15	0
25.39	12	46	14	1
25.51	12	32	13	0
25.61	14	37	13	1
25.61	9	23	15	0
25.61	14	37	13	1
25.61	9	24	15	0
25.61	10	37	17	0
25.61	14	37	16	0
25.61	10	37	14	0
25.69	8	39	15	1
25.69	12	35	14	1
25.70	10	44	14	0
25.75	14	37	13	0
25.75	12	36	16	1
25.75	10	25	14	0
25.75	10	25	14	0
25.75	12	39	15	0
25.75	12	36	13	1
25.75	14	37	15	0
25.75	10	25	14	0
25.75	12	30	14	0
25.75	10	25	15	0
25.75	10	36	15	0
25.75	12	30	15	0
25.75	14	43	14	0
25.75	10	36	16	0
25.75	12	39	14	0
25.79	14	39	14	1
25.79	10	35	13	0
25.82	12	47	13	0
25.82	14	41	14	1
25.82	12	47	14	0
25.82	12	33	14	0
25.82	12	33	13	0
25.82	14	41	15	0
25.82	12	37	14	0
25.82	10	21	12	0
25.82	12	33	15	0
25.84	10	23	15	0
25.84	12	39	12	0
25.85	12	41	14	0
25.85	14	36	13	0
25.85	14	46	15	0
26.09	12	39	14	0
26.26	10	26	12	0
26.26	12	47	12	1
26.26	12	47	13	1
26.29	11	36	14	1
26.31	14	46	15	0
26.31	14	39	11	0
26.34	8	34	13	0
26.37	10	34	13	1
26.45	10	41	14	0
26.45	12	32	12	0
26.45	11	32	12	1
26.45	16	16	12	1
26.45	12	40	14	0
26.50	14	35	9	0
26.52	12	37	9	1
26.52	12	41	8	0
26.52	10	34	11	0
26.54	16	24	10	1
26.54	16	24	10	0
26.57	12	43	8	0
26.57	14	49	8	0
26.57	12	39	9	0
26.61	14	43	8	0
26.61	12	34	9	1
26.61	14	39	10	0
26.61	14	40	9	0
26.61	14	36	8	0
26.61	14	39	7	0
26.61	12	38	8	1
26.61	14	36	10	1
26.61	14	39	10	0
26.61	14	25	8	0
26.61	12	34	8	0
26.63	14	44	8	0
26.63	12	32	9	0
26.63	12	48	9	0
26.63	18	46	10	1
26.63	12	32	7	0
26.63	18	37	8	1
26.63	12	36	9	0
26.63	14	43	9	0
26.63	14	43	8	1
26.63	12	45	9	1
26.63	12	28	9	0
26.63	16	47	7	0
26.63	12	20	8	0
26.89	11	30	9	0
26.89	11	30	9	0
26.91	14	25	6	1
26.93	12	35	6	0
26.93	12	46	9	0
26.94	14	36	9	0
27.10	12	44	10	0
27.12	12	37	8	0
27.12	12	37	8	0
27.17	7	37	9	0
27.26	16	35	7	0
27.28	12	35	9	0
27.28	12	35	6	0
27.34	18	32	7	1
27.34	14	35	8	0
27.34	13	38	9	0
27.34	12	36	7	0
27.34	13	38	8	0
27.34	10	41	7	1
27.37	12	40	1	0
27.37	14	28	2	0
27.37	20	47	2	0
27.37	16	35	3	1
27.37	16	31	3	1

BMI	SlackSize	Age	ThighGap	Leakage
27.37	14	28	3	0
27.37	20	44	2	1
27.40	12	45	3	0
27.44	12	34	2	0
27.44	12	47	2	0
27.44	12	47	3	1
27.44	14	16	4	0
27.44	14	16	2	0
27.44	14	35	2	1
27.46	13	19	2	0
27.46	10	46	4	0
27.46	12	31	2	0
27.46	14	33	2	0
27.46	14	34	3	1
27.46	12	34	2	0
27.46	12	34	3	1
27.46	12	26	3	0
27.46	14	36	4	0
27.46	12	34	1	1
27.46	16	46	2	0
27.46	12	34	2	0
27.52	16	43	3	0
27.61	16	29	2	0
27.76	14	45	2	0
27.88	14	39	2	0
27.98	14	38	2	1
27.99	12	32	1	0
28.13	14	35	2	0
28.13	14	41	1	0
28.13	14	35	2	1
28.19	18	33	0	1
28.19	16	33	2	1
28.19	18	48	2	0
28.19	12	26	1	0
28.19	16	42	1	0
28.19	14	37	1	1
28.19	16	39	2	0
28.19	16	46	3	0
28.21	14	43	1	1
28.29	14	35	3	0
28.29	14	35	1	0
28.29	16	48	1	0
28.29	16	48	2	0
28.29	16	48	3	0
28.29	16	45	0	1
28.32	16	38	0	1
28.32	12	37	0	0

BMI	SlackSize	Age	ThighGap	Leakage
28.32	12	32	0	0
28.32	12	37	0	0
28.32	12	32	0	1
28.32	14	44	0	1
28.34	14	28	0	0
28.34	12	49	0	0
28.34	14	41	0	0
28.34	12	36	0	0
28.34	14	40	0	1
28.34	14	38	0	1
28.34	14	34	0	1
28.34	14	41	0	0
28.34	14	28	0	0
28.34	14	40	0	0
28.41	16	46	0	0
28.43	12	45	0	0
28.70	22	44	0	0
28.70	18	45	0	0
28.70	18	45	0	0
28.80	18	47	0	1
28.84	12	39	0	1
28.89	18	29	0	0
28.90	18	49	0	0
28.97	16	40	0	0
28.97	16	40	0	0
28.97	18	44	0	0
28.97	16	40	0	0
28.97	18	39	0	0
29.05	12	30	0	0
29.05	14	44	0	0
29.05	12	30	0	0
29.05	14	40	0	0
29.05	14	40	0	0
29.05	16	45	0	1
29.05	16	45	0	1
29.08	14	39	0	0
29.08	14	39	0	0
29.10	12	34	0	0
29.12	14	46	0	0
29.12	14	31	0	0
29.12	14	40	0	1
29.18	12	21	0	0
29.18	14	30	0	0
29.18	12	29	0	0
29.18	14	45	0	0
29.18	16	29	0	0
29.18	16	36	0	1

BMI	SlackSize	Age	ThighGap	Leakage
29.18	16	36	0	1
29.18	18	18	0	1
29.18	12	47	0	1
29.18	18	18	0	1
29.18	16	35	0	0
29.18	16	31	0	0
29.23	14	40	0	0
29.23	18	41	0	0
29.23	14	39	0	0
29.23	14	39	0	0
29.23	18	41	0	0
29.23	16	34	0	1
29.23	16	34	0	1
29.23	14	40	0	0
29.23	14	39	0	1
29.23	18	41	0	0
29.23	14	39	0	0
29.23	14	16	0	1
29.23	14	39	0	0
29.26	14	45	0	0
29.26	14	39	0	0
29.26	12	44	0	0
29.26	12	40	0	1
29.29	12	23	0	1
29.29	12	31	0	1
29.35	12	42	0	0
29.35	12	34	0	0
29.52	14	34	0	0
29.52	16	45	0	1
29.53	20	38	0	1
29.53	22	27	0	0
29.65	14	32	0	0
29.65	14	28	0	0
29.65	18	30	0	1
29.76	18	49	0	1
29.76	16	36	0	0
29.76	18	49	0	0
29.86	14	38	0	0
29.86	18	48	0	0
29.86	14	38	0	0
29.86	16	35	0	0
29.86	14	34	0	0
29.86	14	38	0	1
29.86	16	42	0	0
29.86	16	42	0	0
29.86	16	35	0	1
29.95	14	39	0	1

BMI	SlackSize	Age	ThighGap	Leakage
29.95	16	39	0	1
29.95	16	39	0	0
29.95	14	31	0	0
29.95	16	37	0	1
29.95	14	39	0	0
29.95	16	38	0	0
29.95	16	39	0	0
29.95	14	31	0	0
29.95	14	32	0	0
29.95	16	42	0	0
29.95	12	20	0	0
29.95	16	27	0	0
29.95	16	42	0	1
30.02	14	24	0	0
30.04	14	47	0	0
30.04	16	37	0	0
30.04	14	48	0	1
30.04	16	46	0	0
30.04	16	24	0	0
30.04	14	47	0	0
30.04	16	40	0	0
30.08	14	37	0	1
30.08	14	37	0	0
30.08	14	37	0	0
30.11	14	47	0	1
30.11	14	41	0	0
30.18	16	41	0	1
30.18	14	33	0	0
30.23	18	40	0	0
30.27	20	34	0	0
30.30	14	36	0	0
30.30	20	46	0	0
30.38	16	39	0	0
30.41	16	30	0	0
30.41	20	42	0	1
30.54	20	36	0	0
30.62	16	44	0	0
30.67	14	37	0	0
30.68	18	31	0	0
30.79	16	37	0	1
30.79	14	48	0	0
30.79	16	37	0	0
30.79	16	37	0	1
30.82	14	45	0	0
30.85	18	36	0	0
30.90	18	43	0	0
30.90	18	43	0	0

BMI	SlackSize	Age	ThighGap	Leakage
30.90	18	42	0	0
30.90	16	35	0	0
30.90	18	43	0	1
31.00	12	31	0	0
31.00	14	29	0	0
31.00	14	28	0	0
31.01	18	35	0	1
31.01	16	37	0	0
31.01	16	37	0	1
31.01	20	35	0	0
31.09	16	29	0	0
31.09	16	44	0	0
31.09	14	38	0	1
31.17	24	41	0	0
31.19	18	26	0	1
31.25	12	36	0	0
31.25	12	36	0	1
31.25	16	38	0	1
31.32	22	45	0	0
31.32	17	33	0	0
31.47	16	38	0	0
31.75	16	37	0	0
31.75	14	33	0	0
31.75	16	39	0	0
31.75	14	33	0	0
31.89	16	34	0	0
31.95	20	42	0	0
31.95	16	35	0	0
32.01	16	39	0	0
32.01	16	39	0	1
32.01	18	41	0	0
32.01	18	45	0	0
32.01	16	46	0	0
32.08	22	23	0	0
32.11	20	46	0	0
32.11	20	46	0	0
32.12	16	34	0	0
32.27	18	38	0	1
32.28	16	32	0	0
32.28	18	23	0	0
32.28	18	33	0	0
32.28	20	39	0	1
32.28	18	46	0	0
32.28	20	37	0	1
32.45	14	33	0	1
32.49	24	33	0	1
32.49	20	37	0	0

BMI	SlackSize	Age	ThighGap	Leakage
32.49	18	30	0	1
32.60	22	45	0	0
32.61	14	41	0	1
32.61	14	47	0	0
32.61	18	35	0	0
32.61	22	39	0	1
32.61	16	25	0	0
32.61	16	49	0	0
32.61	14	47	0	0
32.69	22	42	0	1
32.71	18	40	0	0
32.77	20	35	0	1
32.77	22	43	0	0
32.84	18	38	0	0
32.89	20	43	0	0
32.89	16	36	0	0
32.92	18	27	0	0
32.92	18	27	0	1
33.00	22	46	0	1
33.07	16	36	0	0
33.07	16	36	0	1
33.09	16	35	0	0
33.13	18	39	0	0
33.20	16	40	0	0
33.20	16	37	0	0
33.20	16	40	0	0
33.28	18	30	0	0
33.28	18	30	0	1
33.28	20	34	0	0
33.28	18	42	0	1
33.28	18	42	0	1
33.28	18	45	0	0
33.30	16	29	0	1
33.45	18	48	0	0
33.45	18	48	0	1
33.45	16	40	0	0
33.47	18	39	0	0
33.47	18	36	0	0
33.47	18	41	0	0
33.66	18	37	0	1
33.67	18	47	0	0
33.84	14	28	0	0
33.99	20	39	0	1
33.99	38	42	0	0
34.01	18	36	0	0
34.01	16	43	0	0
34.11	20	43	0	1

BMI	SlackSize	Age	ThighGap	Leakage
34.33	16	34	0	1
34.33	18	47	0	1
34.33	18	29	0	0
34.33	18	34	0	1
34.33	18	32	0	0
34.44	20	41	0	0
34.44	20	41	0	0
34.45	20	46	0	0
34.45	20	46	0	1
34.54	16	18	0	0
34.61	18	39	0	0
34.70	16	31	0	0
34.75	18	33	0	0
34.90	18	35	0	1
34.95	26	45	0	1
34.95	26	45	0	1
34.97	20	37	0	0
34.97	20	37	0	0
35.12	16	41	0	0
35.15	18	35	0	0
35.15	20	41	0	0
35.28	20	33	0	0
35.35	22	42	0	0
35.35	22	42	0	0
35.40	22	48	0	1
35.51	20	43	0	0
35.51	18	34	0	1
35.51	20	43	0	0
35.67	18	36	0	1
35.67	18	36	0	0
35.71	16	17	0	0
35.78	18	35	0	1
35.78	0	33	0	1
36.05	22	49	0	0
36.05	22	49	0	0
36.05	22	49	0	1
36.05	22	49	0	1
36.13	16	34	0	0
36.31	18	35	0	0
36.32	18	33	0	1
36.32	22	41	0	1
36.57	18	39	0	0
36.58	18	39	0	0
36.58	16	39	0	0
36.58	24	47	0	1
36.58	40	45	0	0
36.58	40	45	0	0

BMI	SlackSize	Age	ThighGap	Leakage
36.58	40	45	0	1
36.58	16	39	0	0
36.61	20	48	0	0
36.61	20	48	0	0
36.61	22	17	0	0
36.61	18	35	0	0
36.81	18	29	0	0
36.90	22	42	0	0
36.90	18	29	0	0
36.96	22	41	0	1
37.11	20	33	0	1
37.11	20	31	0	0
37.12	20	46	0	1
37.12	20	33	0	0
37.12	20	46	0	0
37.12	24	39	0	0
37.12	20	46	0	0
37.20	18	37	0	1
37.20	20	35	0	0
37.20	20	35	0	0
37.31	24	47	0	0
37.38	20	38	0	0
37.38	20	38	0	0
37.38	20	38	0	0
37.38	20	38	0	1
37.44	20	33	0	1
37.59	18	40	0	1
37.76	20	39	0	0
37.76	18	39	0	1
37.76	20	43	0	0
37.76	20	43	0	1
37.76	18	39	0	1
37.76	22	48	0	1
37.76	22	47	0	0
37.76	22	47	0	1
38.01	20	41	0	0
38.01	22	39	0	0
38.01	22	41	0	0
38.01	26	39	0	0
38.01	20	24	0	0
38.25	20	37	0	0
38.27	22	36	0	0
38.27	22	28	0	0
38.62	20	41	0	1
38.62	18	39	0	0
38.62	22	35	0	0
38.79	20	38	0	0

BMI	SlackSize	Age	ThighGap	Leakage
39.06	24	39	0	1
39.06	24	39	0	0
39.11	22	42	0	0
39.16	22	43	0	0
39.16	22	37	0	0
39.16	22	43	0	1
39.16	22	43	0	1
39.94	20	24	0	0
40.09	26	37	0	0
40.09	26	37	0	0
40.09	26	37	0	1
40.24	22	30	0	0
40.24	22	27	0	1
40.35	24	39	0	0
40.35	24	24	0	0
40.35	24	39	0	0
40.35	22	29	0	0
40.41	24	35	0	0
40.41	24	35	0	0
40.74	22	40	0	0
40.74	24	26	0	0
40.77	26	46	0	1
40.77	24	23	0	0
41.60	24	47	0	0
41.96	22	47	0	0
41.96	24	36	0	0
41.96	24	36	0	0
41.96	24	36	0	0
42.07	22	33	0	0
42.07	22	33	0	1
42.07	20	44	0	1
42.09	22	23	0	0
42.29	20	40	0	0
42.29	20	40	0	0
42.33	26	33	0	0
42.57	22	33	0	0
42.57	24	36	0	0
42.57	24	36	0	0
42.91	24	41	0	0
42.91	24	44	0	0
42.91	24	41	0	0
42.97	20	37	0	0
43.07	26	34	0	0
43.07	28	36	0	0
43.24	32	31	0	1
43.24	32	31	0	0
43.42	24	23	0	1

BMI	SlackSize	Age	ThighGap	Leakage
43.42	24	23	0	1
44.30	48	49	0	1
44.85	24	39	0	0
45.61	36	42	0	0
45.76	24	40	0	0
45.76	26	37	0	0
45.76	26	37	0	0
45.76	26	16	0	1
46.26	20	29	0	1
46.59	26	31	0	0
46.59	26	31	0	0
47.30	28	37	0	1
48.09	26	37	0	0
48.42	28	38	0	0
49.26	26	24	0	0
49.75	34	35	0	0
50.90	32	36	0	1
53.56	28	39	0	0
53.59	20	42	0	1
54.74	34	31	0	1
56.49	26	32	0	0
57.41	28	40	0	1

10.3 Analysis of Variance (ANOVA)

Learning Objectives

- Be Able to Detect Factor Differences Through ANOVA

Unit Contents

- Analysis of Variance (ANOVA)
- One-Way ANOVA
- Two-Way ANOVA
- Two-Way ANOVA With Repetition
- Additional ANOVA Topics

10.3.1 Analysis of Variance (ANOVA)

ANOVA Background

Purpose and Application

In Unit 7.2 *Hypothesis Tests,* we developed methods of determining if there was a difference in the means of two populations. If you are interested in seeing if the average paint thickness was different for two different spray nozzles, you could use these tests. But what if there were several different nozzles available? You could do hypothesis tests of all possible pairs to see if there were any differences. With only four different nozzles though, you'd have to do six different hypothesis tests!

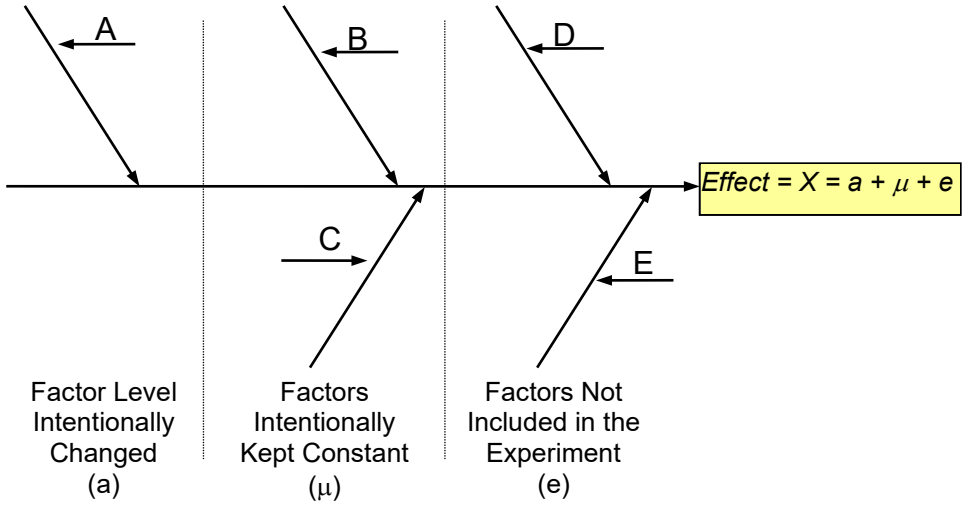
Analysis of Variance (*ANOVA*) is a technique you can use to test many means at once. In other words, you can test if the effects (i.e. paint thickness) are different depending upon the treatment (i.e. spray nozzle). ANOVA also gives you a method of examining the influence of changing several factors on the effect. For instance, suppose we think that *spray nozzle, paint type* and *booth temperature* are factors that affect paint thickness. Our simple hypothesis tests have a hard time handling these situations. ANOVA is related to regression analysis. For example, we may start a cause-and-effect relationship by identifying many different factors that *could* affect the process output.

By identifying widely separate *levels* for these factors (i.e. a high and low level), we can first establish which factors are significant through ANOVA. Then, we could build a regression model based on the significant factors. This process will be discussed further in *Design of Experiments,* (see Section 11).

Theory & Terminology

Consider the cause-and-effect diagram below. To determine the influence of one factor, we will change it and measure the output or effect. But other variables influence the effect. We will deliberately hold some of these constant and some will not be included in the experiment but will be "allowed" to vary (the "error").

These three elements combine in an additive model of the effect:



Let's expand this model one more step. Suppose that we identify three *levels* for factor *A*; we'll call these A_1, A_2, and A_3 (these could be three different spray nozzles; each is referred to as the *ith level* of the factor). We then measure the *effect* on paint thickness by running three experiments at these three different *factor levels* for a total of nine experiments. We record the results in the following table:

	Experiment Number			
Factor Level	1	2	3	**Level Mean**
A_1	X_{11}	X_{12}	X_{13}	μ_1
A_2	X_{21}	X_{22}	X_{23}	μ_2
A_3	X_{31}	X_{32}	X_{33}	μ_3

We now have a mean for each of the levels of factor A (μ_i), which represents the average effect of each nozzle on paint thickness. We can also calculate the *general mean*:

$$\mu = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij}$$

where:

m - *number of factor levels*

n - *number of experiment repetitions*

μ - *General Mean*

We now define the *main effect* of the ith level of A as:

$$a_i = \mu_i - \mu$$

This is the amount we must *add* to the general mean to get the mean associated with the i-th factor level (but note that a_i can be either positive, negative or zero). Now it would be nice if we could simply compare the factor level means or the main effects and pick the one that gives us the best result (i.e. the highest paint thickness). But we learned in Unit 9.2 – *Hypothesis Testing* that just because two numbers are different, it doesn't mean that they are *different!* We need to consider the dispersion of the data and ask the question "are these differences *significant?*"

ANOVA begins to answer the "significant" question by breaking down the *total* variation into components that correspond to the *elements* of the model described above. For our simple case, we will "decompose" the total variation into that due to factor A and that due to the error. If the portion of the total variation due to factor A is significantly greater than that due to the error, then we will conclude that the factor levels *do* have an influence on the effect. This is the basic idea of ANOVA. No matter how complicated the model becomes (i.e. in terms of numbers of factors and associated levels), the ANOVA process is one of decomposing the variation into appropriate components and comparing these components to the error variation.

Our last "theory" aspect is to explore *how* ANOVA decomposes variation. We start with a familiar measure of dispersion, the *Sum of Squares*. The total Sum of Squares (SST) is calculated as follows:

$$SST = \sum_{i} \sum_{j} (X_{ij} - \overline{X})^2$$

where:

$\overline{X}$ - *General Mean Estimate from Sample Data*

The term inside the parentheses is the deviation of each data point from the general mean. But we can break this term down into two components. The first component will be the deviation of the factor level means ($\overline{X}_i$) from the overall mean (this will provide us with the variation component due to the factor levels). The second component will be the deviation of each factor level's data from the factor level's mean (this provides us with the variation due to all other factors, i.e. the *error* term). Here's the math:

$$SST = \sum_i \sum_j (X_{ij} - \overline{X})^2$$

Add and subtract the ith factor level mean:

$$SST = \sum_i \sum_j (X_{ij} - \overline{X} + \overline{X}_i - \overline{X}_i)^2$$

Rearranging and multiplying the terms:

$$SST = \sum_i \sum_j ((\overline{X}_i - \overline{X}) + (X_{ij} - \overline{X}_i))^2$$

$$SST = \sum_i \sum_j (\overline{X}_i - \overline{X})^2 + \sum_i \sum_j (X_{ij} - \overline{X}_i)^2 - 2\sum_i \sum_j (\overline{X}_i - \overline{X})(X_{ij} - \overline{X}_i)$$

Definition of Terms:

SSA - Sum of Squares due to Factor A

$$SSA = \sum_i \sum_j (\overline{X}_i - \overline{X})^2$$

SSe - Sum of Squares due to error

$$SSe = \sum_i \sum_j (X_{ij} - \overline{X}_i)^2$$

$$Covariance = \sum_i \sum_j (\overline{X}_i - \overline{X})(X_{ij} - \overline{X}_i)$$

(but the Covariance equals 0 if Factor A and the "error" factors are independent - this is an assumption we'll make)

This shows that the total sum of squares can be decomposed into two components, the component due to factor A and the component due to error:

$$SST = SSA + SSE$$

Although we've decomposed the total variation, these components can't be directly compared. We can, though, compare the *mean squares* (i.e. the Sum of Squares divided by its associated degrees of freedom - the *mean squares* are estimates of variance). To make this comparison, we'll employ the F-test for equality of two population variances. The test statistic is the ratio of variances and so we'll calculate the following:

$$F = \frac{MSA}{MSe} = \frac{SSA / f_A}{SSe / f_e}$$

where:

MSA - *Mean Square of Factor A*

MSe - *Mean Square Error*

f_i - *degrees of freedom*

If the null hypothesis of this test is rejected, then we have established evidence that factor A has a significant influence on the effect we're measuring. Of course, we'll also be able to establish the parameter interval (confidence bound) and other "interesting" facts, such as which factor level is the "best."

As you can see from the above, although the calculations "only" involve addition, subtraction, multiplication and division, the ANOVA equations can get complicated. To simplify the equations, we'll employ the same notation that was introduced in Contingency Analysis:

Notation	Definition
X_{ij}	Measurement (i.e. individual data point) - for *One-Way ANOVA*, denotes "jth" repetition at "ith" factor level, for *Two-Way ANOVA*, denotes "ith" factor level of A and "jth" factor level of B. For *Two-Way ANOVA with repetition*, another subscript will be added for the repetitions: X_{ijk}.
$X_{i.}$	Summation over all "j's" in the "ith" row - $X_{i.} = \sum_j X_{ij}$ For Two-Way ANOVA, this is the total of measurements for the "ith" factor level of A.
$X_{.j}$	Summation over all "i's" in the "jth" column - $X_{.j} = \sum_i X_{ij}$. For Two-Way ANOVA, this is the total of measurements for the "jth" factor level of B.
$X_{..}$	Summation over all "i's" and "j's" - $X_{..} = \sum_i \sum_j X_{ij}$ This is simply the total of all measurements.
$\overline{X}_{i.}$ *or* $\overline{X}_{.j}$	The mean of the row or column. For n_A factor levels for A and n_B factor levels for B, $\overline{X}_{i.} = \sum_j X_{ij} / n_B \ \ and \ \overline{X}_{.j} = \sum_i X_{ij} / n_A$
$\overline{X}_{..}$	The overall mean of all measurements - $\overline{X}_{..} = X_{..} / (n_A \times n_B)$

The rest of this section will describe three common types of ANOVA, *One-Way ANOVA, Two-Way ANOVA,* and *Two-Way ANOVA with repetition.* It's not too difficult to perform these calculations by hand or spreadsheet. If we need to go beyond *Two-Way ANOVA,* we'll rely on statistical software such as MiniTab.

10.3.2 One-Way ANOVA

Purpose

Comparing spray nozzle type and paint thickness is an example of an experiment that can be analyzed through *One-Way ANOVA.* Only *one* factor varies at different *levels*. Notice that the term *level* could have different meanings:

- We could consider only one nozzle but vary the spray pressure. The different pressures would then be our factor levels. Note that this is like a regression analysis, where the independent variable is controlled at specific values, rather than being allowed to vary over a range.

- We could consider different nozzles; each nozzle would then be a factor level. This is much different than a regression analysis; although the nozzles are the independent variables, there is no meaning to the order of the nozzles.

Each measurement (i.e. each experiment of the process output at a given factor level) is called a *repetition.* We will assume that all factor levels have the same sample size, that is, the same number of repetitions.

One step in planning an experiment is to determine the minimum number of repetitions. The factors that determine this number include the variance of the process output, the minimum difference we wish to detect and the α level we set (See *Additional ANOVA Topics*).

One-Way ANOVA Process

1. Determine the characteristic (effect), the factor to be investigated and the levels of the factor. Establish the Hypothesis:
   a) Null Hypothesis (H_o) - The factor does not influence the characteristic.
   b) Alternative Hypothesis (H_a) - The factor does influence the characteristic.

2. Choose a Significance Level (α - “alpha”).

3. Determine the number of repetitions required.

4. Collect process data or run experiments to generate the measurements.

5. Calculate the Total Sum of Squares (SST), the Factor Sum of Squares (SSA) and the Error Sum of Squares (SSe):

a) For a small number of factor levels and repetitions, the table on the next page can help you organize the inputs to the Sums of Squares. First, enter the data for each repetition at each level (fill in the rows). Then, square these values. In the **Sums** column, add each row of data (i.e. $\sum X_{1j} =$) and then add the squares (i.e. $\sum X_{1j}^2 =$). In the **Sum Squared** column, square the sum of the data for each factor level. Fill in the number of repetitions for each factor level (i.e. n_1). In the two rows at the bottom of the table, first sum the row totals (data and squares) and sum the total number of repetitions (n_T). Then divide the "sum of sums" squared ($X..^2$) by the total number of repetitions.

		Repetitions						
Factor Level		**1**	**2**	**3**	**4**	**Sums**	**Sum Squared**	**n_i**
1	X_{1j}					$\sum X_{1j} = X_1. =$	$(X_1.)^2 =$	$n_1 =$
	X_{1j}^2					$\sum X_{1j}^2 =$		
2	X_{2j}					$\sum X_{2j} = X_2. =$	$(X_2.)^2 =$	$n_2 =$
	X_{2j}^2					$\sum X_{2j}^2 =$		
3	X_{3j}					$\sum X_{3j} = X_3. =$	$(X_3.)^2 =$	$n_3 =$
	X_{3j}^2					$\sum X_{3j}^2 =$		
4	X_{4j}					$\sum X_{4j} = X_4. =$	$(X_4.)^2 =$	$n_4 =$
	X_{4j}^2					$\sum X_{4j}^2 =$		
						$\sum\sum X_{ij} = X.. =$ $\sum\sum X_{ij}^2 =$		$n_T =$
						$X..^2 / n_T =$		

b) Calculate the Sums of Squares, using the input from above:

$$SST = \sum\sum X_{ij}^2 - X_{..}^2/n_T$$

$$SSA = \sum\left(\left(X_{i.}\right)^2 / n_i\right) - X_{..}^2/n_T$$

$$SSe = SST - SSA$$

Where:

$X_{i.}$ *is the sum of the repetitions of each factor level*

and n_i *is the number of repetitions of each level*

6. Develop the ANOVA Table:

Source	Sum of Squares	Degrees of freedom (f)	Mean Square	F (test statistic)
Factor A	SSA =	f_A = m - 1 =	MSA = SSA/f_A =	F = MSA/MSe =
Error	SSe =	f_e = n_T - m =	MSe = SSe/f_e =	
Total	SST =	f_T = n_T - 1 =		

m - Number of factor levels for A

7. Perform the F-Test:
   a) Look up the table value for F(α/2, f_A, f_e).
   b) If the F test statistic is greater than the table value, reject the Null Hypothesis; if not, do not reject the Null.

8. Examine the residuals (difference between the x_{ij} and the associated factor level mean. The residuals should not exhibit any significant departure from normality and there should be no unusual patterns (residuals vs. order of data and vs. fitted values).

9. Draw the conclusion.

One-Way ANOVA Example – Operating Costs by Manufacturer

Scenario – A consumer-testing laboratory wanted to compare the annual operating costs of air conditioners from several different manufacturers. They obtained four air conditioners from each of three manufacturers and measured the operating costs.

1. Determine the Hypothesis:
   a) Null Hypothesis (H_o) – The manufacturer does not influence operating cost.
   b) Alternative Hypothesis (H_a) - The manufacturer does influence the operating cost.

2. Significance Level - $\alpha = 0.1$.

3. Four repetitions were chosen for this experiment

4. Operating Cost Data:

Sample	Manufacturer		
	CoolAir	**WilliChilli**	**QuietCool**
1	$80	$85	$77
2	$75	$88	$74
3	$84	$83	$75
4	$82	$86	$72

5. Calculate the Total Sum of Squares (SST), the Factor Sum of Squares (SSA) and the Error Sum of Squares (SSe):

		Repetitions						
Factor Level		**1**	**2**	**3**	**4**	**Sums**	**Sum Squared**	**n_i**
Cool-Air	X_{1j}	80	75	84	82	$\sum X_{1j} = X_{1\cdot} = 321$	$(X_{1\cdot})^2 =$ 103041	$n_1 = 4$
	X_{1j}^2	6400	5625	7056	6724	$\sum X_{1j}^2 = 25805$		
Willi-Chilli	X_{2j}	85	88	83	86	$\sum X_{2j} = X_{2\cdot} = 342$	$(X_{2\cdot})^2 =$ 116964	$n_2 = 4$
	X_{2j}^2	7225	7744	6889	7396	$\sum X_{2j}^2 = 29254$		
Quiet-Cool	X_{3j}	77	74	75	72	$\sum X_{3j} = X_{3\cdot} = 298$	$(X_{3\cdot})^2 =$ 88804	$n_3 = 4$
	X_{3j}^2	5929	5476	5625	5184	$\sum X_{3j}^2 = 22214$		
						$\sum\sum X_{ij} = X_{\cdot\cdot} = 961$ $\sum\sum X_{ij}^2 = 77273$		$n_T = 12$
						$X_{\cdot\cdot}^2 / n_T = 923521/12 = 76960.08$		

6. Develop the ANOVA Table:

Source	Sum of Squares	Degrees of freedom (f)	Mean Square	F (test statistic)
Manu-facturer	*SSA = ((103041/4) + (116964/4) + (88804/4)) – 76960.08= 242.17*	*f_A = m - 1 = 3 – 1 = 2*	*MSA = SA/f_A = 242.17/2 = 121.08*	*F = MSA/MSe = 121.08/7.86 = 15.4*
Error	*SSe = 312.92 -242.17 = 70.75 70.75*	*f_e = n_T - m = 12 – 3 = 9*	*MSe = SSe/f_e = 70.75/9 = 7.86*	
Total	*SST = 77273-76960.08= 312.92*	*f_T = n_T - 1 = 12 – 1 = 11*		

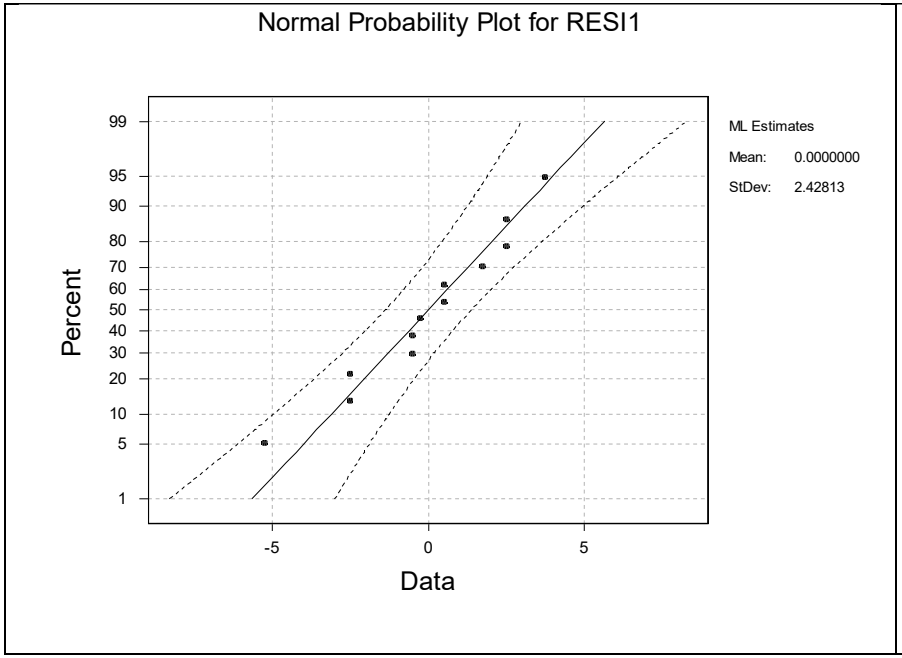
3 - Number of factor levels for A

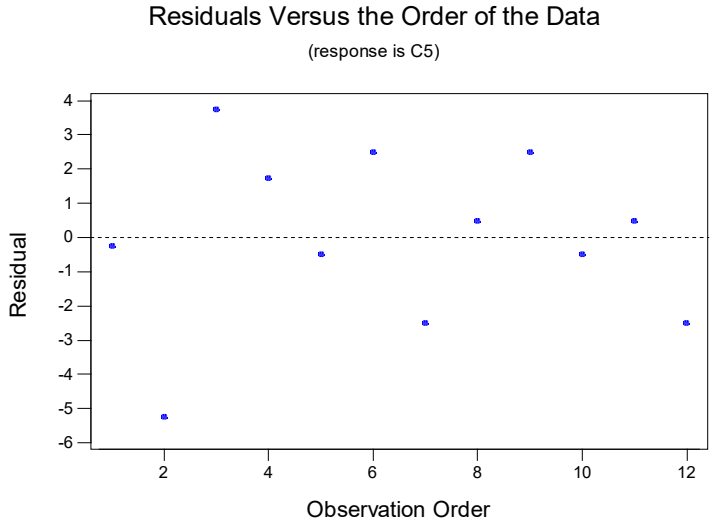
7. Perform the F-Test:

   a) F(0.1/2, 2, 9) = 4.26

   b) Since 15.4 is greater than 4.26, reject the Null Hypothesis in favor of the alternate.

8. The normal probability plot for the residuals appears below. No significant departure has been observed. Likewise, a plot of the residuals vs. order of the data does not show any unusual patterns:





9. Conclusion: Operating costs are different by manufacturer, with the conclusion significant at the 0.1 level.

10.3.3 Two-Way ANOVA

Purpose

Most output variables (or *effects*) are affected by many input variables (or *causes*). Suppose that we would like to investigate the effects of both *spray nozzle* and *spray pressure* on paint thickness. Now we have *two* factors and their associated levels. To analyze the data from this type of experiment, we employ *Two-Way ANOVA.*

This process decomposes the total variation into that due to factors A, B and the error. We will set up two test statistics: one the ratio of the Mean Square - A vs. the Mean Square error (MSA/MSe) and the other the ratio of the Mean Square - B vs. the Mean Square error (MSB/MSe).

Here is an example of a Two-Way ANOVA experiment:

		Spray Pressure		
		B_1	**B_2**	**B_3**
Nozzle Types	**A_1**	X_{11}	X_{12}	X_{13}
	A_2	X_{21}	X_{22}	X_{23}
	A_3	X_{31}	X_{32}	X_{33}
	A_4	X_{41}	X_{42}	X_{43}

Here, the X_{ij} represent the measurement associated with the *ith* level of factor A and the *jth* level of factor B. This experimental "layout" is a *full factorial experiment,* since all combinations of factor levels are tested.

As the number of factors and levels increase, so will the number of measurements. For example, if there are six factors at two levels each, a *full factorial experiment* will require $2^6 = 64$ experiments. These can become very time-consuming and costly. Fortunately, more efficient experimental layouts have been developed to avoid this problem.

Two Way ANOVA Process

1. Determine the characteristic (effect), the two factors to be investigated and the levels of the factors. Establish the Hypotheses:

a1) Null Hypothesis - A (H_o) - Factor A does not influence the characteristic.

b1) Alternative Hypothesis - A (H_a) - Factor A does influence the characteristic.

a2) Null Hypothesis - B (H_o) - Factor B does not influence the characteristic.

b2) Alternative Hypothesis - B (H_a) - Factor B does influence the characteristic.

2. Choose a Significance Level (α - "alpha").

3. Collect process data or run experiments.

4. Calculate the Total Sum of Squares (SST), the Factors' Sums of Squares (SSA & SSB) and the Error Sum of Squares (SSe):

$$SST = \sum_i \sum_j X_{ij}^2 - (X_{..})^2 / n_T$$

$$SSA = \sum_j \left\{ (X_{i.})^2 / n_i \right\} - (X_{..})^2 / n_T$$

$$SSB = \sum_i \left\{ (X_{.j})^2 / n_j \right\} - (X_{..})^2 / n_T$$

$$SSe = SST - SSA - SSB$$

where:

n_T – *Total Number of Data* $(n_A \times n_B)$

5. Develop the ANOVA Table:

Source	Sum of Squares	Degrees of freedom (f)	Mean Square	F (test statistic)
Factor A	SSA =	$f_A = n_A - 1$ =	MSA = SSA/f_A =	F = MSA/MSe =
Factor B	SSB =	$f_B = n_B - 1$ =	MSB = SSB/f_B =	F = MSB/MSe =
Error	SSe =	$f_e = (n_A - 1)(n_B - 1)$ =	MSe = SSe/f_e =	
Total	SST =	$f_T = n_T - 1$ =		

n_i - *Number of factor levels for ith factor*

6. Perform the F-Test:

a) Look up the table values for $F(\alpha/2, f_A, f_e)$ and $F(\alpha/2, f_B, f_e)$.

b) Determine if either or both null hypotheses are rejected. Note that either A or B or both factors may have a significant influence on the effect.

7. Examine the residuals (difference between the x_{ij} and the associated factor level mean. The residuals should not exhibit any significant departure from normality and there should be no unusual patterns (residuals vs. order of data and vs. fitted values).

8. Draw the conclusion.

Two-Way ANOVA Example – Heat Treat Time and Temperature

A Black Belt was trying to determine if the temperature and time gears were heat-treated could improve their material hardness. She ran an experiment with temperature set at two levels (500 and 550 F) and varied the time at three levels (4, 6 and 8 hours).

1. Establish the Hypotheses:

   a1) Null Hypothesis - Temperature (H_o) - Temperature does not influence the characteristic.

   b1) Alternative Hypothesis - Temperature (H_a) - Temperature does influence the characteristic.

   a2) Null Hypothesis - Time (H_o) - Time does not influence the characteristic.

   b2) Alternative Hypothesis - Time (H_a) - Time does influence the characteristic.

2. Significance Level $\alpha = 0.1$.

3. Results from the experiments:

4. Calculate the Total Sum of Squares (SST), the Factors' Sums of Squares (SSA & SSB) and the Error Sum of Squares (SSe) and prepare ANOVA table:

Gear Heat Treatment Experiments

		Time (Hours)					
Temperature (F)		***4***	***6***	***8***	***Sums***	***Sums Squared***	***Total Sums Squared - Temperature***
	500	15	18	22	55	3025	6625
	550	14	20	26	60	3600	2208.33
Sums		29	38	48	115	***X..***	***Sums Squared/3***
Sums Squared		841	1444	2304	13225	***X..^2***	
Total Sums Squared - Time		4589			2204.17	***X..^2/nt***	
Sums Squared/2		2294.50					

Squares

225	324	484	***Total:***
196	400	676	2305

n-t	6
n-Temp	2
n-Time	3

ANOVA Table	***SS***	***DoF***	***MSS***	***F-Statistics***
Total	100.83	5		
Temperature	4.17	1	4.17	1.32
Time	90.33	2	45.17	14.26
Error	6.33	2	3.17	

To see the details of the calculations presented above, double click on the icon below:



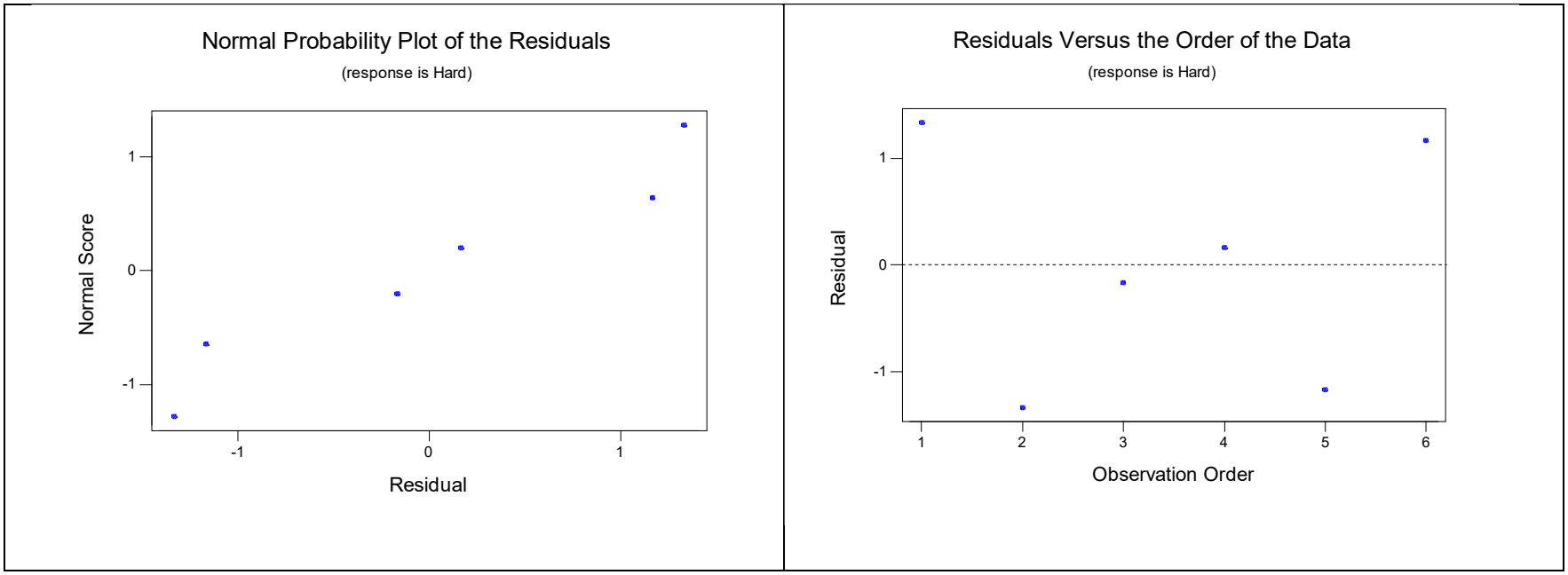
"Two-Way ANOVA.xls"

6. Perform the F-Test:

a) Table values for F(0.1/2, 1, 2) = 18.5 and F(0.1/2, 2, 2) = 19.0.

b) Neither of the F-Statistics are greater that the critical values; do not reject the null hypotheses.

7. The normal probability plot for the residuals appears below. No significant departure is observed. Likewise, a plot of the residuals vs. order of the data does not show any unusual patterns:



8. Neither of the factors can be claimed to have an impact on the gear hardness. However, there is some evidence that heat treat time increases hardness. The Black Belt may want to run the experiments again (replicating the experiment – See Two-Way ANOVA with Replication, next!).

10.3.4 Two-Way ANOVA With Repetition

Two-Way ANOVA without repetition allows us to decompose the variation into three components, that due to the factors A and B and that due to the error.

In some cases, there is the possibility of an *interaction* between the factors. The combined effect of a certain nozzle and spray pressure can cause an effect that is greater than the sum of the individual influences, or the factors may interact to cancel each other.

Repeating the experiments (i.e. conducting more than one experiment at each factor and level combination) allows us to investigate the presence of interactions. The total sum of squares is decomposed into four components: that due to factors A and B, that due to the interaction, AxB, and finally, that due to the error.

Three test statistics will be calculated: MSA/MSe, MSB/MSe and MSAxB/MSe.

Here is an example of a Two-Way ANOVA with repetition:

		Spray Pressure		
		B_1	**B_2**	**B_3**
Nozzle Types	**A_1**	X_{111}, X_{112}	X_{121}, X_{122}	X_{131}, X_{132}
	A_2	X_{211}, X_{212}	X_{221}, X_{222}	X_{231}, X_{232}
	A_3	X_{311}, X_{312}	X_{321}, X_{322}	X_{331}, X_{332}
	A_4	X_{411}, X_{412}	X_{421}, X_{422}	X_{431}, X_{432}

Here, the X_{ijk} represent the measurement associated with the *ith* level of factor A, the *jth* level of factor B and the *kth* repetition of the factor level combination.

Two Way ANOVA Process

1. Determine the characteristic (effect), the two factors to be investigated and the levels of the factors. Establish the Hypotheses:

a1) Null Hypothesis - A (H_o) - Factor A does not influence the characteristic.

b1) Alternative Hypothesis - A (H_a) - Factor A does influence the characteristic.

a2) Null Hypothesis - B (H_o) - Factor B does not influence the characteristic.
b2) Alternative Hypothesis - B (H_a) - Factor B does influence the characteristic.

a3) Null Hypothesis - AxB (H_o) - Interaction AxB does not influence the characteristic.
b3) Alternative Hypothesis - AxB (H_a) - Interaction AxB does influence the characteristic.

2. Choose a Significance Level (α - "alpha").

3. Determine the number of repetitions required.

4. Collect process data or run experiments to generate the measurements.

5. Calculate the Total Sum of Squares (SST), the Factors' Sums of Squares (SSA & SSB), the Interaction Sum of Squares and the Error Sum of Squares (SSe). To help us calculate the latter two terms, we'll calculate a Sum of Squares of AB (SSAB):

$$SST = \sum_i \sum_j \sum_k X_{ijk}^2 - (X_{...})^2 / n_T$$

$$SSA = \sum_i \left\{ (X_{i..})^2 / n_i \right\} - (X_{...})^2 / n_T$$

$$SSB = \sum_j \left\{ (X_{.j.})^2 / n_j \right\} - (X_{...})^2 / n_T$$

$$SSAB = \left(\sum_i \sum_j X_{ij.}^2 \right) / n_r - (X_{...})^2 / n_T$$

$$SSAxB = SSAB - SSA - SSB \textbf{ and } SSe = SST - SSAB$$

where :

n_T – **Total Number of Data** ($\mathbf{n_A \times n_B \times n_r}$)

n_r **- Number of Repetitions**

6. Develop the ANOVA Table:

Source	Sum of Squares	Degrees of freedom (f)	Mean Square	F (test statistic)
Factor A	SSA =	$f_A = n_A - 1$ =	MSA = SSA/f_A =	F = MSA/MSe =
Factor B	SSB =	$f_B = n_B - 1$ =	MSB = SSB/f_B =	F = MSB/MSe =
Interaction AxB	SSAxB =	$f_{AxB} = (n_A - 1)(n_B - 1)$ =	MSAxB = SSAxB/f_{AxB} =	F = MSAxB/MSe =
Error	SSe =	$f_e = n_A\, n_B\, (n_r - 1)$ =	MSe = SSe/f_e =	
Total	SST =	$f_T = n_T - 1$ =		

n_i - *Number of factor levels for ith factor (n_r - # of repetitions)*

7. Perform the F-Test:

   a) Look up the table values for $F(\alpha/2, f_A, f_e)$, $F(\alpha/2, f_B, f_e)$ and $F(\alpha/2, f_{AxB}, f_e)$.

   b) Determine if any or all the null hypotheses are rejected. Note that any of the factors/interaction may have a significant influence on the effect.

8. Examine the residuals (difference between the x_{ij} and the associated factor level mean. The residuals should not exhibit any significant departure from normality and there should be no unusual patterns (residuals vs. order of data and vs. fitted values).

9. Draw the conclusion.

Two-Way ANOVA With Repetition Example –

The Black Belt took our advice in the last example and repeated the experiments (with the factors of temperature set at 500 and 550F and heat treat time at 4, 6 and 8 hours.

1. Establish the Hypotheses:

   a1) Null Hypothesis - Temperature (H_o) - Temperature does not influence the characteristic.
   b1) Alternative Hypothesis - Temperature (H_a) - Temperature does influence the characteristic.
   a2) Null Hypothesis - Time (H_o) - Time does not influence the characteristic.
   b2) Alternative Hypothesis - Time (H_a) - Time does influence the characteristic.
   a3) Null Hypothesis – Time x Temperature (H_o) - Interaction Time x Temperature does not influence the characteristic.
   b3) Alternative Hypothesis – Time x Temperature (H_a) - Interaction Time x Temperature does influence the characteristic.

2. Choose a Significance Level: $\alpha = 0.1$.

3. Two repetitions of the experiment were run.

4. Collect process data:

	Time = 4 hours		***Time = 6 hours***		***Time = 8 hours***	
Temp = 500	15	*16*	18	*17*	22	*23*
Temp = 550	14	*16*	20	*19*	26	*17*

Data in italics came from the second replication.

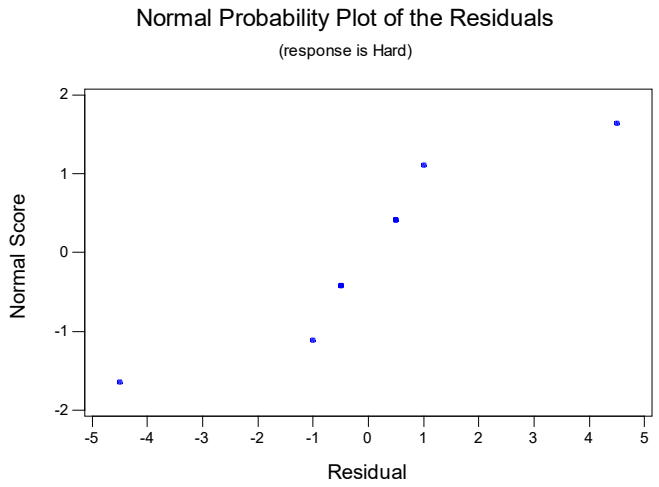
5, 6. Calculate the Total Sum of Squares (SST), the Factors' Sums of Squares (SSA & SSB), the Interaction Sum of Squares and the Error Sum of Squares (SSe). Develop the ANOVA table. Here, we'll rely on Minitab to perform the calculations. Choosing Two-Way ANOVA from the Stats Menu, the following ANOVA table was calculated. Note that Minitab reports the *source of variation,* then the *degrees of freedom (DF),* then the *Sum of Squares (SS)*, the *Mean Sum of Squares (MS),* the *F-ratio,* and finally the *p-value*:

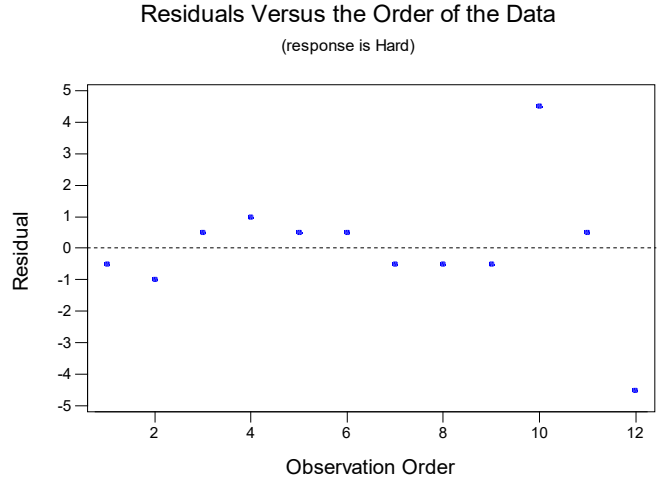
Analysis of Variance for Hardness

Source	DF	SS	MS	F	P
Temp	1	0.08	0.08	0.01	0.919
Time	2	91.17	45.58	6.15	0.035
Interaction	2	5.17	2.58	0.35	0.719
Error	6	44.50	7.42		
Total	11	140.92			

7. Perform the F-Test: For Temperature: F(0.1/2, 1, 6) = 5.99, for Time: F(0.1/2, 2, 6) = 5.14 and for the TimeXTemperature Interaction: F(0.1/2, 2, 6) = 5.14. The Time F-Statistic is the only one that is greater than the critical value. Note that Minitab provides a "P" value. If you don't want to look up the critical value of the F distribution, simply compare the "P" to your α value. If the "P" is less than your α, you can conclude that the factor is significant.

8. The normal probability plot for the residuals appears below. In this case, the normal plot does not appear to be a good fit and there seems to be a cyclic pattern in the residuals versus order of the data.





9. By repeating the experiment, the Black Belt can conclude that the time the gears spend in the heat treatment does influence hardness. However, the residuals analysis casts doubt on whether the ANOVA assumptions are met – she should investigate the non-normality and the residuals pattern.

10.3.5 Additional ANOVA Topics

Residuals

Always examine the residuals for patterns as part of your ANOVA. The residuals (difference between the experimental data and the fitted values (e.g. the factor level means) should not display a significant departure from normality. A histogram and/or normal probability plot of the data will help you analyze this requirement. Scatter Plots of the residuals versus the fitted values and a plot of the residuals versus order of the data should also be examined to see if unusual patterns exist. These diagnostics will help you determine if your fitted model is good or not. If this analysis indicates problems, one option is to transform the data. A logarithmic transformation of the response, for example, may improve the model fit.

Parameter Estimation

Performing an ANOVA is a "GO/NO-GO" analysis. This factor has either a significant influence on the effect, or it does not. The ANOVA establishes *statistical significance.*

For those significant factors, we want to go one step further and determine *which* level of the factor is different. We will employ the usual approach of developing confidence bounds for the factor level parameters to answer this question. This will help us understand the *operational significance* of the factor.

There are three types of confidence interval that may be calculated from our ANOVA data:

- *Individual Factor Level - Confidence Interval for the Mean,*
- *Paired Factor Levels - Confidence Interval for the Difference of Means, and*
- *Variance - Confidence Interval for the Error Variance.*

The process and calculations associated with the three types of ANOVA covered here are presented below:

Calculation Process

1. Obtain point estimates for the parameters/differences:

$$(\overline{X}_{i.}, \overline{X}_{i.} - \overline{X}_{i'.}, \sigma_e^2).$$

2. Using appropriate *t,* or χ^2 value, calculate the confidence interval (see tables below).

3. If the confidence interval for the individual factor level mean, μ_i does not cover the general mean μ, then the factor mean, μ_i and the general mean, μ are significantly different.

4. If the confidence interval for μ_i - μ_j does not cover 0, then μ_i and μ_j are significantly different.

One-Way Layout - Here, we can estimate the means of the individual factor levels and differences between the factor levels:

Parameter	Point Estimate	Interval Width
ith factor level mean (μ_i)	$\overline{X}_{i}.$	$\pm t(f_e, \alpha/2)\sqrt{\frac{1}{n_i} MSe}$
Difference between factor level means (μ_i - μ_i')	$\overline{X}_{i}. - \overline{X}_{i}.'$	$\pm t(f_e, \alpha/2)\sqrt{\left(\frac{1}{n_i} + \frac{1}{n_i'}\right) MSe}$

Two-Way Layout - (no interaction present) - Here, we can estimate the means of individual factor levels and differences between factor levels:

Parameter	Point Estimate	Interval Width
Factor A's *ith* level mean ($\mu_{i.}$)	$\overline{X}_{i}.$	$\pm t(f_e, \alpha/2)\sqrt{\frac{MSe}{n}}$ n - *number of levels of factor B*
Factor B's *jth* level mean ($\mu_{.j}$)	$\overline{X}_{.j}$	$\pm t(f_e, \alpha/2)\sqrt{\frac{MSe}{m}}$ m - *number of levels of factor A*

Parameter	Point Estimate	Interval Width
Difference between factor A level means ($\mu_{i.}$ - $\mu_{i.}'$)	$\overline{X}_{i.} - \overline{X}_{i.}'$	$\pm t(f_e, \alpha/2)\sqrt{\frac{2}{n} MSe}$
Difference between factor B level means ($\mu_{.j}$ - $\mu_{.j}'$)	$\overline{X}_{.i} - \overline{X}_{.i}'$	$\pm t(f_e, \alpha/2)\sqrt{\frac{2}{m} MSe}$

Two-Way Layout With Repetition (interaction present) - With interactions, the factor level means are not meaningful. Instead, we estimate the mean of the combinations of factor levels and differences between these combinations:

Parameter	Point Estimate	Interval Width
Individual mean (*ith* and *jth* combination of factors A & B - μ_{ij})	$\overline{X}_{ij.}$	$\pm t(f_e, \alpha/2)\sqrt{\frac{1}{r} MSe}$ r - *number of repetitions*
μ_{ij} - μ_{ij}''	$\overline{X}_{ij.} - \overline{X}_{ij.}''$	$\pm t(f_e, \alpha/2)\sqrt{\frac{2}{r} MSe}$

All Layouts - If a point estimate and/or confidence bound on the error variance is needed, the following estimates are used:

Parameter	Point Estimate	Interval Width
σ_e^2	MSe	$\sigma_L^2 = \frac{SSe}{\chi^2(f_e, \alpha/2)}$ $\sigma_U^2 = \frac{SSe}{\chi^2(f_e, 1-\alpha/2)}$

Minimum Repetitions for ANOVA

The ANOVA process helps us determine if the means of the different factor levels are the same or not. The number of repetitions influences the error variance - the larger the number of repetitions, the smaller the error variance. However, economy must be factored into the situation - the larger the number of repetitions, the higher the cost of the analysis.

One method for establishing the minimum number of repetitions relies on our knowing the *minimum* difference in means we wish to detect (Δ) and the *standard deviation* of the factor level populations (σ - assumed to be equal for all factor levels).

The number of repetitions is found by calculating the ratio of these values and determining desired risks of Type I and Type II errors. The following table provides the minimum number of repetitions for typical α values and for $\beta = 0.1$:

	$\Delta/\sigma = 1.0$			$\Delta/\sigma = 1.25$			$\Delta/\sigma = 1.5$			$\Delta/\sigma = 1.75$			$\Delta/\sigma = 2.0$			$\Delta/\sigma = 2.5$			$\Delta/\sigma = 3.0$		
	α			α			α			α			α			α			α		
m, n*	.1	.05	.01	.1	.05	.01	.1	.05	.01	.1	.05	.01	.1	.05	.01	.1	.05	.01	.1	.05	.01
2	18	23	32	12	15	21	9	11	15	7	8	12	6	7	10	4	5	7	3	4	6
3	22	27	37	15	18	24	11	13	18	8	10	13	7	8	11	5	6	8	4	5	6
4	25	30	40	16	20	27	12	14	19	9	11	15	7	9	12	5	6	8	4	5	6
5	27	32	43	18	21	28	13	15	20	10	12	15	8	9	12	5	6	9	4	5	7
6	29	34	46	19	23	30	14	16	21	10	12	16	8	10	13	6	7	9	4	5	7
7	31	36	48	20	24	31	14	17	22	11	13	17	9	10	13	6	7	9	5	5	7
8	32	38	50	21	25	33	15	18	23	11	13	17	9	11	14	6	7	9	5	6	7
9	33	40	52	22	26	34	16	18	24	12	14	18	9	11	14	6	8	10	5	6	7
10	35	41	54	23	27	35	16	19	25	12	14	19	10	11	15	7	8	10	5	6	7

*m and n are the number of factor levels for A and B. Use the larger of the numbers to determine the minimum number of repetitions.

For example, suppose we wish to compare four different impeller designs on compressor discharge pressure. If the standard deviation of the discharge pressures times is 1 psig and the minimum difference we wish to detect is 3 psig, then Δ/σ = 3 psig/1 psig = 3.0. For an α of 0.05, the minimum number of repetitions would be 5.

Types of Models

The ANOVA models considered in this manual are of the *Fixed Effects* type. Here, we have a special or intrinsic interest in the factor levels (i.e. drugs). The conclusions of a Fixed Effects analysis will be applied to the factor levels identified.

We could "construct" a different kind of model. Suppose sales response time was to be studied in a region. We could sample the population of salespeople, or sample from the population of sales offices. Here, the sampled salespeople or

offices become the factor levels. We are interested, though, in more than just the sampled "units." In this *Random Effects Model,* our interest is in making inferences about the entire population.

It is possible to create a model where some of the factors are *fixed* and some of the factors are *random;* this is called a *Mixed Effects Model.* The major difference in the types of model lies in the objective of our analysis. For Random Effects models, we are interested in the overall mean and variability of the population, not in differences in the factor level means. There are also differences in a few of the calculations, although in general, the ANOVA process is the same.

10.3.6 Contrasts

Contrast Analysis - Basics

ANOVA tests the null hypothesis that *all means are equal* versus the alternative hypothesis of *at least one mean is different*. We may also want to know if mean A is different from mean B or C. Contrasts are a non-parametric means of performing this additional analysis (we'll also use contrasts to analyze the results of a Designed Experiment in Unit 11.1).

A *comparison* creates an order relationship between two things or groups of things. If A = 5 and B = 6, we easily make the comparison that *B > A*.

A *contrast* compares two things or groups of things by comparing their *difference with zero:*

- *"A is greater than B as (A – B) is greater than zero"*
- *"A is less than B as (A – B) is less than zero"*
- *"A = B becomes (A – B) = zero"*

The advantage of Contrasts is that it reduces the comparison problem to one of computing a difference.

Suppose we have measured tensile strengths of four different materials (A, B, C, D). We've measured 5 strength values for each material and calculated the average and standard deviation for each material. We'll call the data collected for each material a *subgroup*:

Material	A	B	C	D
	55.9	58.7	60.7	62.7
	56.1	61.4	60.3	64.5
	57.3	60.9	60.9	63.1
	55.2	59.1	61.4	59.2
	58.1	58.2	62.3	60.3
Average	56.52	59.66	61.12	61.96
s^2	1.352	1.983	0.592	4.668

For contrasts, we'll use *k* as the count of subgroups (4 here), with *n* designating the size of the subgroups (5 here).

Definition: Starting with k subgroups, a contrast is defined on the set of k objects by any set of k coefficients (c1, c2, c3, c4, etc.) which sum to zero:

$$C = c_1\mu_1 + c_2\mu_2 + c_3\mu_3 + c_4\mu_4 + \cdots c_k\mu_k \Rightarrow C = \sum_i c_i\mu_i \qquad \sum_i c_i = 0$$

The **Signs** of the coefficients will define the comparisons, the **Magnitude** of the **Coefficients** is chosen to ensure sum of coefficients = 0 (Note: if a factor level is not included in the comparison, its coefficient is assigned to be 0).

Suppose we want to compare materials A versus B. We assign +1 to A and -1 to B as coefficients. This is a valid contrast since 1 + (-1) = 0.

We could also compare materials C and D. Again, we'll assign +1 to C and -1 to D as coefficients. The sum of the coefficients here is 0, making it a valid contrast.

Suppose that materials A & B came from one vendor and materials C and D came from another vendor. We could create an AB vs CD contrast by assigning coefficients + 1 to A and B and -1 to C and D. This is a valid contrast since (1 + 1) + (-1 + -1) = 0.

These three contrasts are shown below in tabular form. Note that if a material is not included in the contrast, its coefficient is 0:

Contrasts	A	B	C	D
A v B	1	-1	0	0
C v D	0	0	1	-1
AB v CD	1	1	-1	-1

Contrasts are very general; we can create many different comparisons. Suppose we wanted to compare materials A & B with material C? We just need to assign coefficients so that their sum = 0 (you should check this below):

Contrast	A	B	C	D
AB v C	1	1	-2	0

Evaluating Contrasts

To evaluate a specific contrast, we'll start by calculating the sum of the product of the coefficients and the subgroup averages. The form of a contrast is:

$$C = \sum_{i=1}^{k} c_i \, x \, \mu_i$$

Since we are estimating the means with averages, the contrast estimate is:

$$C = \sum_{i=1}^{k} c_i \, x \, \bar{x}_i$$

The contrast, *C*, for comparing materials A and B is:

	A	B	C	D	
	55.9	58.7	60.7	62.7	
	56.1	61.4	60.3	64.5	
	57.3	60.9	60.9	63.1	
	55.2	59.1	61.4	59.2	
	58.1	58.2	62.3	60.3	
Average	56.52	59.66	61.12	61.96	
Contrasts					
A v B	1	-1	0	0	$\bar{c}$
$c_i \, x \, \bar{x}_i$	56.52	-59.66	0	0	-3.14

The sign of the Contrast indicates that the average of material B is greater than the average of material A (obvious here, but more complicated contrasts can be trickier to interpret!). The Contrasts for the comparisons described above are shown below:

	A	B	C	D	
	55.9	58.7	60.7	62.7	
	56.1	61.4	60.3	64.5	
	57.3	60.9	60.9	63.1	
	55.2	59.1	61.4	59.2	
	58.1	58.2	62.3	60.3	
Average	56.52	59.66	61.12	61.96	
Contrasts					$\bar{c}$
A v B	1	-1	0	0	-3.14
C v D	0	0	1	-1	-0.84
AB v CD	1	1	-1	-1	-6.9

In Hypothesis testing, we've learned that just because two numbers are different, the question is, "*are they "really" different?"* We'll calculate an F-Ratio to answer this question, this is simply the ratio of the Mean Sum of Squares due to the Contrast (MS(C)) and the within-subgroup Mean Sum of Squares (MSW):

$$F = \frac{MS(C)}{MSW}$$

The MS(C) is the ratio of the Sum of Squares of the Contrast (SS(C)) and its degrees of freedom, which in this case is 1. Peeling the onion, the SS(C) is the ratio of C-hat squared, and an adjustment term as shown below:

$$SS(C) = \frac{(\bar{C})^2}{Adj.Term} \; with \; Adj.Term = \frac{1}{n}\sum_{j=1}^{k}(c_j)^2$$

For the A v B materials, the SS(C) is:

$$(\bar{C})^2 = (-3.14)^2 = 9.86$$

$$with\ Adj.Term = \frac{1}{n}\sum_{j=1}^{k}(c_j)^2 = \frac{1}{5}\left(1^2 + -1^2 + 0^2 + 0^2\right) = 0.4$$

$$SS(C) = \frac{9.86}{0.4} = 24.649$$

The MSW is the ratio of the within-group Sum of Squares divided by its degrees of freedom. In case you are confused by the double-summation, we take each sub-group value, subtract that sub-group's average and square it (the inner summation). We then repeat for all sub-groups (the outer summation):

$$MSW = \frac{SSW}{dof} = \frac{\sum_{j=1}^{k}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2}{k(n-1)}$$

For our material examples, the MSW is:

	A	B	C	D			
	55.9	58.7	60.7	62.7			
	56.1	61.4	60.3	64.5			
	57.3	60.9	60.9	63.1			
	55.2	59.1	61.4	59.2			
	58.1	58.2	62.3	60.3			
Average	56.52	59.66	61.12	61.96	**Total SS**	**k(n-1)**	**MSW**
Subgroup SS	5.408	7.932	2.368	18.672	34.38	16	**2.14875**

Note that the degrees of freedom = 4(5 – 1) or 16. If you are using Excel to calculate the Subgroup Sum of Squares, the function DEVSQ() is handy.

Finally, we can calculate the F-Ratio and determine if it is significant:

$$F = \frac{MS(C)}{MSW} = \frac{24.649}{2.15} = 11.47$$

We look up the critical F-value for 1, 16 degrees of freedom (α = 0.05) and find it is 4.49 (see Table A.4, Appendix A or use this equation in Excel: =F.INV.RT(α, DF1, DF2)). Therefore, the A vs. B contrast is significant. We can repeat the calculations for the other 2 contrasts, finding that the **AB vs. CD** contrast is also significant:

	A	**B**	**C**	**D**				
	55.9	58.7	60.7	62.7				
	56.1	61.4	60.3	64.5				
	57.3	60.9	60.9	63.1				
	55.2	59.1	61.4	59.2				
	58.1	58.2	62.3	60.3				
Average	56.52	59.66	61.12	61.96	**Total SS**	**k(n-1)**	**MSW**	
Contrasts					$\bar{C}$	**Adj Term**	**SS(C)**	**F-Ratio**
A v B	1	-1	0	0	-3.14	0.4	24.649	**11.47132**
	56.52	-59.66	0	0				
	1	1	0	0				
C v D	0	0	1	-1	-0.84	0.4	1.764	0.820942
	0	0	61.12	-61.96				
	0	0	1	1				
AB v CD	1	1	-1	-1	-6.9	0.8	59.5125	**27.69634**
	56.52	59.66	-61.12	-61.96				
	1	1	1	1				Critical Value = 4.49

The last step in Contrast Analysis is to estimate the size of the difference between the Sub-group averages. This is called the Estimated Contrast Effect (l-hat) and is calculated by dividing the Contrast value by the sum of the positive coefficients:

$$\bar{l} = \frac{\hat{C}}{\sum(+c_i's)}$$

For the A vs. B comparison, this is:

$$\bar{l} = \frac{-3.14}{1} = -3.14$$

These "l-hats" are only calculated for the significant contrasts:

	A	**B**	**C**	**D**						
Average	56.52	59.66	61.12	61.96	**Total SS**	**k(n-1)**	**MSW**			
Contrasts					$\bar{c}$	**Adj Term**	**SS(C)**	**F-Ratio**	**Sum$^+$ c_i**	$\hat{l}$
A v B	1	-1	0	0	-3.14	0.4	24.649	**11.47132**	1	-3.14
	56.52	-59.66	0	0						
	1	1	0	0						
C v D	0	0	1	-1	-0.84	0.4	1.764	0.820942		N/A
	0	0	61.12	-61.96						
	0	0	1	1						
AB v CD	1	1	-1	-1	-6.9	0.8	59.5125	**27.69634**	2	-3.45
	56.52	59.66	-61.12	-61.96						
	1	1	1	1						

The sign of the l-hat indicates which group has the highest average response.

A note on Alpha (α): In the example above, we performed three F-tests. For ANOVA, we only performed one F-Test. Here, the chosen alpha applies to the overall test, whereas for contrasts, the alpha is associated with each individual test. The overall alpha for a set of contrasts can be calculated using the Bonferroni inequality:

$$1 - (1 - \alpha)^p \leq \alpha_{overall} \leq p\alpha$$

For the three contrast F-tests performed above (using the individual α = 0.05), this inequality calculates to:

$$0.143 \leq \alpha_{overall} \leq 0.150$$

Since α (Type I error) is the probability of concluding that an effect exists when it does not, by using α = 0.05 for the individual F-tests, we increase the overall false positive probability to between 0.14 and 0.15. The experimenter might consider using lower alpha values for the individual tests to manage the overall alpha. Of course, this will make it harder to detect significant individual contrasts.

Contrasts in Excel

The Contrast calculations can be easily set up in Excel. Excel AVERAGE(), STDEV.S(), ANOVA (requires Analysis Toolpak to be activated), DEVSQ() (for the sums of squares) are helpful functions.

	Vendor XYZ		Vendor ABC	
Material	**A**	**B**	**C**	**D**
Values	55.9	58.7	60.7	62.7
	56.1	61.4	60.3	64.5
	57.3	60.9	60.9	63.1
	55.2	59.1	61.4	59.2
	58.1	58.2	62.3	60.3
Average	56.52	59.66	61.12	61.96
s*2	1.352	1.983	0.592	4.668

Contrasts

Internal Labs (A vs. B)

External Labs (C vs. D)

Internal and External (AB vs. CD)

	Internal		External	
Lab	**A**	**B**	**C**	**D**
Contrast	**c1**	**c2**	**c3**	**c4**
A vs. B	1	-1	0	0
C vs. D	0	0	1	-1
AB vs. CD	1	1	-1	-1

Contrast	**C**	**C^2**	**SS(C)**	**F-Ratio**
A vs. B	-3.14	9.8596	24.649	11.47132
C vs. D	-0.84	0.7056	1.764	0.820942
AB vs. CD	-6.9	47.61	59.5125	27.69634

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Column 1	5	282.6	56.52	1.352
Column 2	5	298.3	59.66	1.983
Column 3	5	305.6	61.12	0.592
Column 4	5	309.8	61.96	4.668

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	85.9255	3	28.641833	13.32953	0.000128	3.238872
Within Groups	34.38	16	**2.14875**			
Total	120.3055	19				

Dof **Mean Square**

20 obs - 4 groups

Adjustment Terms

0.4

0.4

0.8



Each F-Ratio has 1 & 16 dof

Critical value = 4.49

-3.14

-3.45

Summary – Contrast Process Steps

1. Collect the data by k Subgroups of size n
2. Calculate the Average and Std Dev^2 for each subgroup (Excel – stdev.s())
3. Determine the Contrasts of Interest (A vs. B, etc.)
4. Develop the Contrast Table – each row's sum = 0
5. Calculate the Adjustment Terms
6. Perform the ANOVA to get the MSW and degrees of freedom (Within Group dof)
7. Contrast Analysis:
    1. Calculate the Contrast Column (sum of products – contrast * Sub-Group Average)
    2. Calculate C^2
    3. Calculate the SS(C) – C^2/Adjustment Term for that Contrast
    4. Calculate the F-Ratio (SS(C)/MSW)
    5. Using 1, and dof from ANOVA, find the critical F-value from the table
    6. Determine which contrasts are significant
    7. Calculate the Estimated Contrast Effect ($\hat{l}$) and make appropriate comparisons

10.4 Exercises

Exercise - QC Workload

A QC supervisor is trying to predict the workload her inspectors will experience each day. She thinks that the number of units being built each day will influence the number of inspection requests they receive each day. Perform correlation/regression analyses for the data shown below. Can she use the number of units as a workload predictor?

Gettysburg Plant (#Units/# Requests)						Durango Plant (# Units/# Requests)				
55	323		50	301		52	269		35	224
61	191		48	157		63	281		41	224
52	353		52	276		71	247		42	285
58	316		51	291		66	235		37	291
54	312		60	277		70	228		44	261
61	321		54	315		66	277		46	225
65	368		60	351		71	258		43	252
54	308		64	375		67	258		49	282
52	260		55	338		63	259		39	255
50	369		53	230		57	187		41	227
55	346		61	343		65	255		39	191
52	234		58	281		78	232		45	283
52	267		52	251		69	218		52	273
49	263		65	412		60	248		54	281
48	218		53	255		66	344		52	308
46	245		53	228		73	316		48	245
47	167		56	274		80	308		41	295
40	137		51	404		60	255		47	250
41	245		63	356		34	93		48	199
52	329		65	322		46	234		48	228
57	385					41	295		41	131

Exercise - Inspection Time

The same QC supervisor is investigating the number of hours spent completing inspections. The following data was obtained by one of the analysts working on the problem:

Day	# Requests	# Hours
1	215	16.0
2	210	11.6
3	230	27.2
4	240	35.6
5	220	19.0
6	225	23.2
7	230	28.6
8	215	15.4
9	235	31.2
10	230	28.0

If you want to predict the number of hours required to perform a certain number of inspections, what is the independent variable? What is the dependent variable?

If you want to predict the number of requests that can be completed in a given number of hours, what is the independent variable? What is the dependent variable?

Which is a better prediction statement?

Perform a correlation/regression analysis of this data. Interpret the analysis.

Objective: To explore relationships between variables

Instructions: The following data represents production data for HSG compressor bodies. Operations have long thought that the dowel position (dowel pos A, opp) and parallelism impact their ability to hold tolerances on the rotor bores (Male and Female RTR Bore). Does the data support their theory?

			TOLERANCE	TOLERANCE	Tolerance	Tolerance	Tolerance	Tolerance	Tolerance	Tolerance	Tolerance
			0.001	0.001	0.0005	0.0018	0.0008	0.0008	0.0008	0.0008	0.0008
HSG NO	DATE	MACH.NO	M RTR BORE	FEM RTR BORE	SLD VAL	Parallel A	M dowel pos A	M dowel pos opp	F dowel pos A	F dowel pos opp	PosSld val
1393	11/23/1999	G&L3	-0.0001	0.0003	0.0012	0.0005	0.0007	0.0005	0.0004	0.0005	0.001
1394	11/24/1999	G&L3	0.0002	0.0002	-0.0005	0.0009	0.0005	0.0004	0.0002	0.0005	0.0012
1397	11/26/1999	G&L3	0.0001	0.0004	-0.0002	0.0012	0.0005	0.0003	0.0006	0.0009	0.0008
1398	11/27/1999	G&L3	0	0.0002	-0.001	0.0006	0.001	0.0005	0.0001	0.0007	0.0014
1399	11/27/1999	G&L3	0.0003	0.0003	0.0008	0.0015	0.0006	0.0003	0.0001	0.001	0.0008
1440	3/23/2000	Orion 1	0.0002	-0.0007	0.0006	0.0016	0.0011	0.0014	0.0011	0.0008	0.0008
1443	4/4/2000	Orion 1	-0.0003	-0.0005	0.0006	0.0011	0.001	0.001	0.0001	0.0005	0.0006
1444	4/4/2000	Orion 1	0	-0.0004	0.0006	0.0013	0.0008	0.0008	0.0001	0	0.0008
1488	8/8/2000	G&L3	0.001	0.0007	0.0001	0.0009	0.0001	0.0002	0.0001	0.0001	0.0001
1492	8/10/2000	G&L3	0.0013	0.0011	0	0.001	0.0003	0.0003	0.0007	0.0001	0.0011
1493	8/10/2000	G&L3	0.0008	0.0004	0.0001	0.0009	0.0006	0.0005	0.0006	0.0004	0.0012
1504	9/12/2000	Orion 1	-0.0002	-0.001	-0.0014	0.0011	0.0005	0.001	0.0002	0.0009	0.0003
1506	9/12/2000	Orion 1	0.0002	0	0.0002	0.0013	0.0005	0.0005	0.0002	0.0012	0.0002
1507	9/6/2000	Orion 1	-0.0011	-0.0018	-0.0003	0.0019	0.0003	0.0007	0.0001	0.0008	0.0011
1508	9/6/2000	Orion 1	0.0003	-0.0002	-0.0003	0.0032	0.0004	0.0005	0.0002	0.0009	0.0004
1509	9/11/2000	Orion 1	0.0004	-0.0001	-0.0003	0.0013	0.0004	0.0009	0.0001	0.001	0.0004
1510	9/12/2000	Orion 1	0.0005	-0.0002	-0.0011	0.0013	0.0004	0.0007	0.0003	0.0009	0.0004
1512	9/12/2000	Orion 1	0	-0.0001	0	0.0004	0.0002	0.0006	0.0007	0.0002	0.0012
1514	9/13/2000	G&L3	0.0014	0.0013	-0.0018	0.0002	0.0002	0.0002	0.0003	0.0001	0.0014
1516	9/14/2000	G&L3	0	-0.0003	-0.0013	0.0003	0.0002	0.0002	0.0003	0.0001	0.0007
1517	9/15/2000	G&L3	-0.0003	-0.0006	-0.0006	0.0005	0.0004	0.0001	0.0001	0.0006	0.0008

Objective: To explore relationships between variables

Instructions: Is there a relationship between the usage of sheet metal (sheets per day) and the number of Non-Conformance Reports (NCRs)?

Day	Sheet Metal Usage	NCR's/Day
1	130	4
2	145	7
3	223	12
4	300	17
5	110	5
6	155	9
7	170	11
8	140	6
9	123	3
10	154	4

Objective: To explore relationships between multiple independent variables and a single dependent variable.

Instructions: A utility is trying to develop a model that predicts electricity usage by household. Through discussions with engineering staff, the Black Belt has selected *house square footage, house age, % tree shading,* and *number of family members in residence* as potential explanatory variables, with kilowatt-hours consumed per month as the dependent variable.

Develop a multiple linear regression model for the data below. How 'good' is this model in explaining energy usage variation?

Sq. Footage	Age	% Shading	No. Family	KW-Hrs.
2000	5	20	4	1294
3000	10	10	1	1952
1000	40	50	2	989
1500	50	20	5	1976
1750	25	40	3	1306
2250	5	10	2	1438
3500	2	0	4	2325
2450	8	20	5	1650
2700	2	10	3	1711
1100	35	70	2	794
1050	40	60	2	919
2200	15	40	4	1420
2250	12	30	6	1620
750	40	80	1	482
1800	22	50	3	1167
1950	18	30	4	1430
2400	6	25	2	1387
3500	2	10	2	2143
2000	13	35	5	1355
1900	19	70	2	942
2100	10	50	3	1064
3300	5	25	1	1870
900	35	90	2	437
1400	25	70	6	950

Objective: To explore non-linear relationships between independent and dependent variables.

Instructions: Formulators of shampoos and other consumer products know that the concentration of sodium affects the viscosity of the solution. Given the following data, develop the "best" model of the relationship between sodium and viscosity (*Hint: Start with linear, proceed to quadratic and maybe even cubic*).

%NaCl	Viscosity
0.00	3140
0.05	4020
0.10	4440
0.15	4640
0.20	4720
0.25	4700
0.30	4740
0.35	4400
0.40	4600
0.50	4300
0.60	3920
0.65	3480
0.75	3060
0.90	2300
1.00	1920
1.20	1240
1.50	720

Objective: To explore non-linear relationships between independent and dependent variables (data transformation required).

Instructions: Chief Engineer Scott is working on improving the Enterprise's warp drive system. He is investigating how the matter/anti-matter mixture rate and the energy flow rate through the di-lithium crystals affect the warp speed of the ship. Using the data below, build a regression model for the Chief Engineer. (*Hint: develop scatter diagrams of each of the independent variables against the dependent variable first*).

Matter-Anti-Matter Flow Rate (gm/sec)	Dilithium Energy Flow (TeraWatts)	Warp Speed (x C)
10	2	3.78
10	2	3.74
10	3	3.54
10	4	5.98
15	2	2.91
15	3	10.89
15	4	7.66
15	5	9.84
20	3	8.93
20	4	12.56
20	5	14.56
20	6	23.93
25	4	26.53
25	5	37.17
25	6	47.45

Matter-Anti-Matter Flow Rate (gm/sec)	Dilithium Energy Flow (TeraWatts)	Warp Speed (x C)
25	7	39.27
30	5	24.97
30	6	42.24
30	7	57.23
30	8	100.11
35	6	85.91
35	7	109.51
35	8	158.99
35	9	232.16
40	7	220.60
40	8	149.64
40	9	403.00
40	10	363.01

Objective: To explore relationships between continuous independent variables and discrete dependent variables.

Instructions: A nuclear power plant had experienced failures of welds that held the end caps on fuel rods. The team assigned to the project thought that the length of time the rod had been in the reactor might affect the possibility of weld failure. In the table below, 0 = no failure, 1 = failed weld. Perform a logistic regression of the data below to help the team answer their question.

Months in Reactor Core									
6	**12**	**18**	**24**	**30**	**36**	**42**	**48**	**54**	**60**
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	1	0
0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	1	0
0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1	1	0

Months in Reactor Core									
6	**12**	**18**	**24**	**30**	**36**	**42**	**48**	**54**	**60**
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1
0	0	0	0	1	0	0	0	1	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	1	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	1	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	1	0	0	0
6	12	18	24	30	36	42	48	54	60

Objective: To explore relationships between continuous independent variables and discrete dependent variables.

Instructions: A company was designing a new instrument to seal blood vessels during surgery. A good seal (hemostasis) is difficult to measure. The surgeon, though, can provide a judgment of whether hemostasis has occurred or not. Design parameters being investigated include *sealing pressure, dwell time,* and *sealing power.* The parameters were varied as follows: *sealing pressure – 70 to 90, dwell time – 1 to 3* and *sealing temperature – 150 – 250.* Experiments performed on pigs produced the logistic equation shown below. Determine a few possible combinations of design parameters that could minimize the probability of hemostasis failure.

Logistic Regression Equation:

$$E(Hemostasis\ Failure) = \frac{\exp(55 - 0.7SP - 3DT - .03ST)}{1 + \exp(55 - 0.7SP - 3DT - .03ST)}$$

SP - Seal Pressure

DT - Dwell Time

ST - Seal Temperature

Objective: To explore relationships between discrete independent and dependent variables.

Instructions: At a local coffee shop, we found a questionnaire that professed to help us choose our "best bean" (below). What kind of regression model would you use to develop this prediction? Develop a data collection plan that would provide the data to allow you to calculate the parameters of the regression model.

Question	**Response**			
The "X's"				
What flavors do you prefer?	Nutty	Fruity	Spicy	Sweet
How do you prefer steak?	Rare	Medium rare	Medium well	Well done
Which fruit do you prefer?	Grapefruit	Orange	Melon	Banana
How do you like marshmallows?	White	Golden brown	Some charcoal	Burnt
Which chocolate do you prefer?	Milk		Dark	
Which grapes do you prefer?	White		Red	
What wine do you prefer?	Dry white	Sweet white	Light-body red	Heavy-body red
How to you prefer food?	Sautéed		Barbecued	
The "Y"				
Coffee Bean Preference	Very light	Light	Dark	Very dark

Exercise - One-Way ANOVA

Do a One-Way ANOVA on the following data, just to get used to the calculations:

	Factor Level				
	A_1	A_2	A_3	A_4	A_5
Measurement	16	8	2	8	20
	12	-4	4	12	16
	14	0	0	10	14
	10	-4	-2	10	8
	16	6	-6	8	18

Create and analyze the following contrasts:

- A1 vs A2
- A2 vs A3
- A1, A4 vs A5

Exercise - Two-Way ANOVA

Using the following data, perform a Two-Way ANOVA:

		Factor Level				
		A_1	**A_2**	**A_3**	**A_4**	**A_5**
Factor Level	**B_1**	16	8	4	18	20
	B_2	12	-4	2	12	16
	B_3	14	0	0	10	12
	B_4	10	-4	-2	6	8
	B_5	6	-6	-6	8	8

Exercise - Two-Way ANOVA (with Repetition)

Using the following data, perform a Two-Way ANOVA, with repetition:

		Factor Level				
		A_1	**A_2**	**A_3**	**A_4**	**A_5**
Factor Level	**B_1**	16	8	4	18	20
	B_2	12	-4	2	12	16
	B_3	14	0	0	10	12
	B_4	10	-4	-2	6	8
	B_5	6	-6	-6	8	8
	B_1	14	11	6	21	23
	B_2	11	-2	5	15	13
	B_3	12	3	-2	8	14
	B_4	11	-1	0	9	11
	B_5	8	-3	-8	11	9

Exercise – Weather Rocket Range

An experiment has been designed to determine the effect of fuel and nozzle configuration on the peak altitude of a weather rocket being designed for NOAA. Four fuels and three nozzle configurations were tested. Perform an ANOVA to determine if either of these factors makes a difference. Test at $\alpha = 0.10$.

		Nozzle Configuration		
		A	**B**	**C**
Rocket Fuel	**1**	158.2	156.2	165.3
	2	149.1	154.1	151.6
	3	160.1	170.9	139.2
	4	175.8	158.2	148.7

Note: Measurements are in meters.

The above experiment doesn't allow for estimation of the potential interaction between fuel and nozzle configuration. The experiment was replicated (data below). Reanalyze the data, but this time include the interaction in your model.

		Nozzle Configuration		
		A	**B**	**C**
Rocket Fuel	**1**	152.6	141.2	160.8
	2	142.8	150.5	148.4
	3	158.3	173.2	140.7
	4	171.5	151.0	141.4

Note: Measurements are in meters.

Exercise - Switchboard Answer Time

A call center supervisor is trying to improve service at the center. The staffing level at the center is varied and the average time to answer a call recorded for three days at each staffing level (three repetitions). Perform an ANOVA to determine if staffing level makes a difference. Test at $\alpha = 0.05$.

		Staffing Level (#)		
		2	**4**	**6**
Repetition	**1**	30	18	19
	2	19	21	16
	3	28	17	15

Measurements are Average Time to Answer Call in Seconds

Create and analyze the following contrasts:

- Staffing Level 2 vs Staffing Level 4
- Staffing Level 4 vs. Staffing Level 6

Exercise - Late Payments

The following Accounting Department data represents the number of late payments per month (dependent variable) and the number of follow-up actions per month (independent variable). Draw a Scatter Diagram of these two variables and perform a correlation/regression analysis.

Month	# Late Payments	# Follow-up Actions
JAN	3500	1400
FEB	4050	700
MAR	3700	650
APR	4375	800
MAY	4500	1250
JUN	3000	550
JUL	2250	4250
AUG	1600	4075
SEP	1875	4000
OCT	1525	3950
NOV	2000	2900
DEC	1750	3500

Exercise - Scale Removal

Three methods for removing scale in cooling tubes are being investigated. The methods' effectiveness is measured by the before & after tube relative cross-section. For example, if a clogged tube had only 5% of its normal area available for water flow before the method was used, and 50% of its area available after the treatment, the measurement is 50% - 5% = 45%. The experimental data is shown below. Perform an ANOVA to determine if there is a significant difference in method (at an $\alpha = 0.05$). Develop an interval estimate for the best method.

		Scale Removal Method		
		A	**B**	**C**
Repetition	**1**	27	15	48
	2	33	18	45
	3	36	24	57
	4	30	21	54

Create and analyze the following contrasts:

- Method A vs Method B
- Method A vs Method C

Exercise – Welding Process Improvement

A team is attempting to improve an arc welding process for pressure vessels. Two factors are being explored – the *type of weld electrode* and the *weld current.* Four different weld rods were examined and three values of weld current. Two repetitions of the experiments were run, with the *tensile strength* of the weld sample being measured. Perform an ANOVA to determine if there is a significant difference in either rod or current (at an $\alpha = 0.05$).

		Weld Current Method		
		A	**B**	**C**
Weld Electrode	**1**	71, 69	69, 70	71, 79
	2	66, 64	62, 65	76, 72
	3	61, 58	56, 59	74, 70
	4	68, 62	62, 64	77, 78

Note: Measurements are in PSIG/1000.

11.0 Experimentation

Unit	Description	Page
11.1	Designing and Running Experiments	11.1-1
11.2	Exercises	11.2-1

11.1 Designing and Running Experiments

Learning Objectives

- Be able to set up, run and analyze single and multiple factor experiments.

Unit Contents

- Experimentation – One Factor at a Time
- Multiple Factor Experiments
- Miscellaneous Experimental Design Topics
- Orthogonal Arrays

11.1.1 Introduction

Many improvement efforts rely on *observing* an existing process to obtain performance data or cause and effect information. For example, we can observe two different workers as they solder copper tubing for condensers. We may observe differences in the *outcomes* obtained by the workers (i.e. leak rates, rejects, etc.). We can then perform a study of their processes to try to determine which factors are responsible for the better outcomes and encourage the other workers to adopt these practice patterns.

This approach has much merit. But the observational approach fails us when we strive to go beyond the best current method. Here, we must adopt the role of the scientist and engage in *experimentation.* The experimental approach encourages us to deliberately vary factors and/or their levels to further improve the process. We hope to learn something new at the end of an experiment.

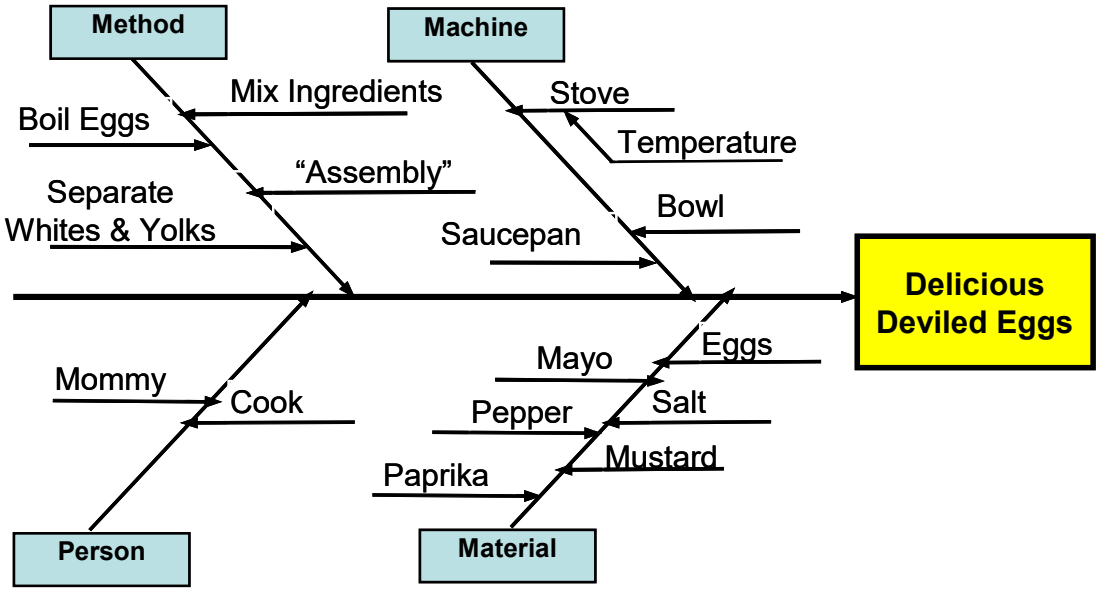
This section will describe several approaches to experimentation. Modern statistical theory has developed powerful, efficient experimental methods for the experimenter.

11.1.2 Experimentation – One Factor at a Time

A "Fair Test" on Deviled Eggs

A while ago, my daughter came to me with a question from her fourth-grade science class - *"Daddy, what's a fair test?"* I looked through her science book to get an idea of what the authors considered to be a fair test and then our discussion began. I talked about how we learned new things through experimentation and how we had to be careful about *how* we performed our experiments. Her eyes started glazing over and I realized that I needed to get out of my *theoretical* mode and into a *practical* discussion very quickly or I'd lose her.

I knew she liked to make deviled eggs for our family - it's her favorite recipe. I asked her how she made the deviled eggs, and she regained interest since she knew the recipe by heart. Very quickly, we developed the *Cause-and-Effect System* for her deviled eggs:



Now we had something that she could relate to a "fair test." I asked her what she would do if our family (her "customers") thought the eggs were too spicy. She said that she could probably put less salt in the eggs. We began to plan a *single*

factor experiment. I asked her how she would know that the eggs were better, and she said that she would ask us to tell her how they *tasted* the next time she made them - now we had established the *outcome measure* of the experiment.

Then, we talked about the other factors in the experiment. If she changed the amount of salt in the recipe, she would need to keep the other important factors *constant*. She'd need to use the same number of eggs, the same amount of mayo, etc. Finally, I asked her about other "extraneous" factors, such as whether she made the eggs in the morning or evening. She said that it didn't make any difference. So, we'd also identified factors that could vary *randomly.*

We summed up the *fair test* discussion. She now understood the notion of *experimentation*, the need to *measure* the experiment's outcome*,* the idea of changing *one factor at a time,* the need to hold other important variables *constant*, and that other, unimportant variables could be left to vary *randomly*.

Experimental Conditions

We face the same issues as my daughter when we attempt to improve *our* production processes. The *max-min-con* concept is a helpful way to remember these:

Maximize the effect of the factor of interest. Often, this means deliberately setting the factor of interest at two *extreme* values. We try to ensure that if the factor *does* affect the output, then we will be able to note and measure its effect. This principle may be limited by our machines' abilities to run at extreme values

Minimize the variation caused by other factors that can affect the output, but that we do not believe to have a major impact. To do this, we will *randomize* the order of the experiments so that these other sources of variation "average out" over all the experiments, rather than showing up in only a few experiments. Randomization will also help if there is an unknown factor that can affect the results. Suppose my daughter was wrong about no difference between morning or afternoon and there is an effect (maybe ambient temperature affects the process). Randomization will distribute this variation across the experiments.

Control the variation introduced into the experiment by factors which do have a major impact on the output, but which are not of interest to us in this experiment. In manufacturing, control means setting machine values at some constant number. In service, we can control the number of staff, availability of materials, computers, etc.

Single Factor Experiments - Advantages & Disadvantages

The single factor experiment is taught in many engineering and science college classes. In the early part of the 20th century, it was considered the *only* "scientific" way to experiment. It does have the advantage of *simplicity* and is generally easy to plan. There are, however, some serious disadvantages to the single factor experiment.

First, consider the results of the experiment. All we can really say is that if we change this one factor so much (with the other important factors held constant at some level), that it has a certain effect. We do not know whether this same effect will be achieved if the other factors vary. This often accounts for the difference between "laboratory" and "real-world" results. Under carefully controlled laboratory conditions, a factor may be shown to improve results. In the real world, however, other factors will either "drown out" the impact of the studied factor, or interactions will negate its impact.

Second, what if there are several factors that we wish to investigate. How do we plan such a *series* of experiments? Do we run one experiment with all the factors at their "low" level and then another experiment at their "high" level? Or should we run a *series* of single factor experiments? After the first experiment, do we set its factor at the "best" level, or at its "original" level? What if there are interactions between these factors? How can we detect these?

These and other problems motivated an English statistician, Sr. R. A. Fisher to develop a different experimental philosophy, one based on understanding the effects of multiple changes to a process. He first employed his methods in the agricultural field. It took a bit longer for other fields such as medicine and industry to apply his methods. Unfortunately, many college and graduate engineering curricula today still don't teach experimental design and analysis.

Single Factor Experiments - Analyzing the Results

In previous sections, we discussed ways of analyzing the results of a single factor experiment, including comparative histograms, hypothesis tests, and *analysis* control charts. These are simple, robust analysis methods and are highly recommended.

For multiple factor experiments, though, we will employ the more powerful techniques of Analysis of Variance (ANOVA) and Contrasts described in Unit 10.3.

11.1.3 Multiple Factor Experiments – Factorial Designs

The Factorial Experiment

Let's help our daughter improve her deviled eggs' "production process." Suppose she's interested in improving the *taste* of her deviled eggs. Based on her cause-and-effect diagram, she identifies two factors that she thinks could improve the eggs' taste, *salt* and *paprika*. Here, Dad recommends that she develop a *full factorial* experiment to test her hypothesis.

To conduct this kind of experiment, we'll perform the following 11 steps:

1. Determine the purpose of the experiment.

Why are we performing the experiment? What aspect(s) or quality characteristic(s) of the product or service are we trying to improve? In our case, our daughter wants to improve the ***taste*** *of the eggs.*

2. Determine how the experiment's outcome is to be measured.

Develop indicators or measures associated with the quality characteristic(s). Operationally define these indicators. For the deviled eggs, she decides to have Mom rate the eggs' taste on a 1-9 Likert scale (1 = yucky, 9 = delicious).

3. Identify the factors to be investigated and their *levels*.

Develop a Cause & Effect diagram of the production system; identify factors that are suspected of impacting the effect. Our little girl has already identified ***salt*** *and* ***paprika*** *as the factors, now she needs to identify appropriate levels.*

Often, a "high/low" strategy is employed. For example, if she thinks that more paprika will improve the eggs' taste, the low level of that factor will be the current amount she adds, the high level will be an increased amount. We could also investigate three levels: high/current/low. The low level is simply an amount that is lower than the current amount added. The current level will be called a "center point."

For her situation, Dad recommends that she investigate two levels. For the Salt factor, she sets the High level equal to the current amount she uses, and the Low level is one half that amount. For the Paprika factor, the Low level is the current amount, and the High level is twice that amount.

Note that the usual convention is to use current settings as the Low values and the experimental values as the High values (we'll see how her settings affect the results shortly):

Experiment Ingredients	High	Low
Salt	¼ teaspoon (current)	½ teaspoon (new)
Paprika	1 ½ teaspoon (new)	¾ teaspoon (current)

4. Determine how to control/address other variables

There are other variables that can affect the taste. These include both ingredients, process variables and noise variables. She identifies these variables and determines their values for the experiment. Since she's not changing these, they are the current values from her recipe:

Other Variables	How Addressed
Hard-boiled Eggs	6 (constant)
Mayonnaise	¼ cup (constant)
Ground Mustard	½ teaspoon (constant)
Pepper	¼ teaspoon (constant)
Water Temperature	212F (constant)
Boiling Time	12 minutes (constant)
Ice Bath Temperature	32F (constant)
Ice Bath Time	5 minutes (constant)
Stirring Time	2 minutes (constant)
Refrigeration Time	60 minutes (constant)
Time of Day	Any time (noise)
Ambient Temperature	Indoor – 72F (noise)
Ambient Humidity	Indoor – 60% RH (noise)

Notice that there are three noise variables, Time of Day, Ambient Temperature and Ambient Humidity. She'll randomize the experimental order to ensure that these don't creep into her experiment as "Lurking Variables."

5. Develop an experimental layout.

The full factorial experimental layout includes all combinations of factors and levels. The full factorial layout allows us to determine the effect of each factor and any interactions that may exist between the factors.

For this experiment, we have two factors, each with two levels. This results in a layout of four combinations:

		Salt (S)	
		Low (S_1)	High (S_2)
Paprika (P)	Low (P_1)	P_1 S_1	P_1 S_2
	High (P_2)	P_2 S_1	P_2 S_2

There is another way to display the experimental conditions. We assign a -1 to the low level and +1 to the high level of each factor. We can "generate" the full factorial design as follows:

Run	Paprika Factor	Salt Factor	Result
1	1	1	
2	1	-1	
3	-1	1	
4	-1	-1	

*We'll call these "coded units" (see **Notation and Analysis** below). These also are contrasts (Unit 10.3) – the sums of the +1's and -1's add to zero for each column. We can use Contrast ANOVA to analyze the experiments. Note that adding factors quickly increases the number of required trials. If we wanted to investigate six factors, each at two different levels, the number of combinations is 2 x 2 x 2 x 2 x 2 = 64. Don't worry, there's a solution to this!*

6. Decide how many repetitions of the experiment are needed.

How many times should the experiment be repeated at each factor level combination? The experimental layout shown above looks suspiciously like a Two-Way ANOVA, without repetition. Recall that we introduced repetition in ANOVA for two reasons:

a) To be able to detect interactions in a two-way layout (i.e. two factors A and B at m and n levels respectively).

b) To be able to improve our ability to detect differences between factor levels (as the number of repetitions is increased, the variation in the mean decreases).

For our daughter's experiment, we'll recommend that she use 2 repetitions per combination. Here we had to be practical. We'd like to be able to detect interactions, so that means we need to have at least 2 repetitions, but we don't want to force our daughter to make batches and batches of eggs - she'll lose interest.

Here, she'll need to make a total of 4 x 2 = 8 batches of deviled eggs. The number of repetitions also increases the number of experiments to be run. For factors A, B, C, D .... at a, b, c, d, ... levels and r repetitions, you will have to run a x b x c x d x . . . x r experiments.

7. Randomize the order of the experiments.

The max-min-con concept encourages us to minimize the effect of variables that may impact the result by randomizing the experiments' order.

To do this, simply number the experiments and then establish their order by drawing these numbers at random. Written numbers may be placed in a hat and drawn at random, or a random number table may be employed. If you are using statistical software like Minitab, it will have a randomize runs option.

Numbered Experiments (Two Repetitions/Combination):

		Salt (S)	
		Low (S_1)	**High (S_2)**
Paprika (P)	**Low (P_1)**	1,2	3,4
	High (P_2)	5,6	7,8

Randomized Experimental Order:

		Salt (S)	
		Low (S_1)	**High (S_2)**
Paprika (P)	**Low (P_1)**	*6,4*	*2,5*
	High (P_2)	*1,7*	*3,8*

8. Conduct the experiments and measure their outcomes.

*Conduct **all** the experiments, in the randomized order determined in step 6. Measure the quality characteristic(s) of interest. Our daughter made 8 batches of deviled eggs in one month. She obtained the following results from the taste tests:*

		Salt (S)	
		Low (S_1)	**High (S_2)**
Paprika (P)	**Low (P_1)**	*5.4, 5.8*	*7.8, 8.3*
	High (P_2)	*5.5, 6.2*	*8.2, 8.8*

9. Analyze the experimental results.

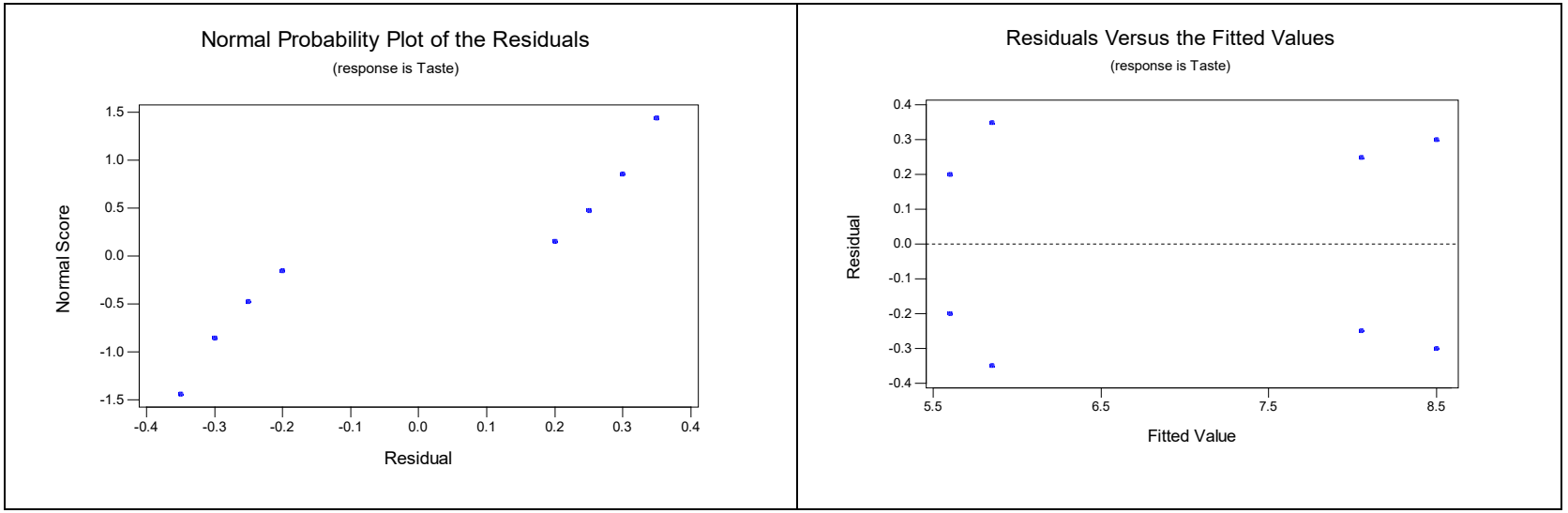
Dad helped perform an ANOVA of the experimental results. Here, we have a two-way ANOVA with Repetition. Determine which factors/interactions are significant, examine the ANOVA residuals, calculate confidence bounds for factor levels, etc.

*The ANOVA Table for the deviled egg experiments appears below. The critical F-value for 1, 4 degrees of freedom (dof) for $\alpha = 0.05$ is 7.71. The only significant term is **Salt**:*

Term	Sum of Squares	Degrees of freedom (f)	Mean Square	F (test statistic)
Salt (S)	SSS = 13.005	$f_A = n_A - 1 = 2 - 1 = 1$	MSA = SSA/f_A = 13.005	F = MSA/MSe = **81.8**
Paprika (P)	SSP = 0.245	$f_B = n_B - 1 = 2 - 1 = 1$	MSB = SSB/f_B = 0.245	F = MSB/MSe = 1.54
Interaction (SxP)	SSSxP = 0.020	$f_{AxB} = (n_A - 1)(n_B - 1) = (2 - 1)(2 - 1) = 1$	MSAxB = SSAxB/f_{AxB} = 0.020	F = MSAxB/MSe = 0.13
Error	SSe = 0.630	$f_e = n_A\, n_B\, (n_r - 1) = 2 \times 2 \times (2 - 1) = 4$	MSe = SSe/f_e = 0.159	
Total	SST = 13.900	$f_T = n_T - 1 = 8 - 1 = 7$		

As discussed in Unit 10.3, Analysis of Variance, you should examine the residuals to ensure that the ANOVA model's assumptions are met. A normal probability plot of the residuals and the residuals versus fits appears below. The residuals versus fits plot appears OK, but the residuals do not appear to fit a normal distribution.

One way of dealing with this problem is to transform either the response variable or the factors. For example, taking the logarithm of the response may improve the fit here.



10. Draw conclusions.

Which factors/interactions are significant? Which factor levels are "best?" Could additional experimentation help you develop even better performance?

From the ANOVA table, the only significant factor is the amount of salt in the eggs. Paprika doesn't affect the taste (at least over the range used in the experiment), nor is there an interaction between salt and paprika.

*A regression analysis of taste versus salt amount provides a predictor equation for our daughter (***Taste = 10.8 - 2.55 Salt***). She could then "manage" the taste response as a function of salt. Here, salt has a negative coefficient but remember that we set the High to ¼ teaspoon (current) Low to ½ teaspoon (experimental) – the more salt, the worse the taste.*

11. Perform confirmatory experiments.

Set the factors at their "best" levels. Run a few additional experiments at these levels to ensure that the experimental results are repeatable. If your experimentation was performed as a pilot study, it's now time to move up to the "full scale" process or you may continue experimentation with the significant factor(s).

Some Notes on Factorial Experiments

Interactions - The 2^2 experiment (2 factors at two levels - the notation is L^f - where L is the number of levels and f is the number of factors) cannot detect interactions between the A and B factors *without repetition.* This was discussed in Unit 10.3 - ANOVA. Experiments of order 2^3 and higher can detect interactions without repetitions. The need for repetitions at these higher order experiments should be determined by the need to be able to detect a certain difference (Δ) in means given a certain population variance (σ^2).

Interactions - Another Point - Except for 2^2 layouts without repetition, the factorial design can detect *almost* all order interactions. For example, a 2^4 (4 factors at 2 levels) layout will detect

- *first order* (or two factor - AxB, AxC, AxD, BxC, BxD, & CxD), and
- *second order* (or three factor - AxBxC, AxBxD, AxCxD, & BxCxD) interactions.

However, to detect the *third order* (or four factor - AxBxCxD) interaction requires a repetition of the experiment. Practically, these high-level interactions are often not important and the effort to detect them is unnecessary. More efficient experimental designs that deliberately ignore some or all interactions will be discussed below.

Interactions - Yet Another Point - Some experimental design texts give the impression that it is desirable to detect interactions. The Taguchi system of quality engineering takes another view. First, Taguchi attempts to identify quality characteristics that are *additive* in nature, that is, since the factors' effects are added to get the result (X = A + B + C + . .) there are no interactions. Second, Taguchi's use of orthogonal arrays is intended to identify significant factors *despite* interactions. See Unit 14.4 for more on this.

Randomization - The completely randomized experiment is the "gold standard" for designed experiments. In cases where it is expensive to change factor levels, the experimenter may decide to set the experimentation order, recognizing the risk involved. For example, if factor A is the most expensive to change, then the experimental order would run all experiments with A at its low level, then the remaining experiments at A's high level. This is a "factor difficult to change" situation and may result in a *nested* experiment.

11.1.4 Multiple Factor Experiments – Analyzing Large Factor Factorial Designs

Experiments with Large Numbers of Factors

The deviled egg experiment described above is full-factorial, with two factors set at two levels. You may wish to include additional factors and perhaps additional levels for the factors. For example, suppose that our daughter wanted to see if the amount of mustard, or the time the eggs were cooked affected the taste. Including the salt and paprika, we now have four factors.

If each factor is assigned two levels, one replication of the experiment will require 16 trials (2^4). The experimental layout appears below. The **-1** stands for the *low* level of the factor; the **1** stands for the high level (see **Notation & Analysis** below).

Trial	Salt	Paprika	Mustard	Cook Time	Taste
1	-1	-1	-1	-1	
2	1	-1	-1	-1	
3	-1	1	-1	-1	
4	1	1	-1	-1	
5	-1	-1	1	-1	
6	1	-1	1	-1	
7	-1	1	1	-1	
8	1	1	1	-1	
9	-1	-1	-1	1	
10	1	-1	-1	1	
11	-1	1	-1	1	
12	1	1	-1	1	
13	-1	-1	1	1	
14	1	-1	1	1	
15	-1	1	1	1	
16	1	1	1	1	

This experiment (if performed with at least one replication – i.e. a minimum of 32 trials) will detect the *main effects* of the four factors, and all *interactions* (two level such as *Salt x Paprika*, three level such as *Salt x Mustard x Cook Time* and the four level interaction of *Salt x Paprika x Mustard x Cook Time*). Note that the replication is required to generate an error term (Mean Sum of Squares – Error, or MSE). If we really want to be able to detect these interactions, the number of

trials required grows rapidly. Fortunately, if we are willing to give up some of this capability, there are alternatives to the full factorial experiment, either *fractional factorial* or *screening* designs as described in the next section.

Analyzing this experiment is no different from the two-factor full-factorial. We will (or Minitab will!) perform an ANOVA on the results. Minitab will also perform t-tests on the terms by taking the average of the low results and the high results and comparing them with a two-sample t-test. Minitab will also give you the option to create a variety of residual plots to analyze. Finally, Minitab will perform a regression analysis of the factors against the response. For factors/interactions that are not significant, you'll usually have the option to remove those terms and rerun the analysis with only the significant factors.

Notation & Analysis – The -1 and 1's are called *coded* units in Minitab. Each column is a *contrast* – the sum of the 1's and -1's will add to zero (recall Unit 10.3). If we substitute the actual value (e.g. Salt Low – ¼ teaspoon, High – ½ teaspoon) for the -1 and 1's, these are called *uncoded* units.

Using coded factors (−1, 0 (this is a "center point"), +1) puts all factors on the same scale. That gives you:

- Orthogonal columns (especially for fractional factorials)[1]
- Cleaner ANOVA sums of squares
- Coefficients you can compare directly

With coded units, a coefficient answers a very nice question: *"If I move this factor from low to high, how much does the response change?"*

Also, with coded units, effect sizes are directly comparable. For the deviled egg example, comparing the amount (teaspoon) of salt with cook time (minutes) is apples and oranges. But with coded units, the bigger coefficient, the bigger the effect. Main effects and interactions are immediately rankable. This is especially important for:

- Plackett-Burman designs
- Fractional factorial designs

Finally, with coded units, interactions behave nicely. Coded units preserve orthogonality so main effects don't "steal" variance from interactions and alias structures behave exactly as the design intended. In uncoded units, scaling differences can make interactions look distorted or harder to interpret.

[1] **Orthogonality** means that the effects you're trying to estimate are **statistically independent of one another**.

On the other hand, uncoded units still matter, mainly because engineers don't think in ±1's. They want to know "if we increase cook time by 5 minutes, the taste improves by 1.4 points. Graphics like contour plots are easier to interpret with uncoded units.

Coding for Minitab:

1. Define factors with uncoded values (in "engineering units")
2. Analyze the DOE, Minitab-22 will tell you what factors are significant using coded units (low p-values):

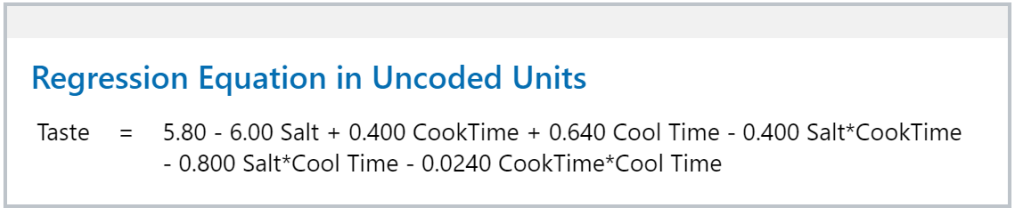
Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		6.625	0.200	33.13	0.019	
Salt	-3.750	-1.875	0.200	-9.38	0.068	1.00
CookTime	0.350	0.175	0.200	0.87	0.542	1.00
Cool Time	0.800	0.400	0.200	2.00	0.295	1.00
Salt*CookTime	-0.250	-0.125	0.200	-0.63	0.644	1.00
Salt*Cool Time	-0.500	-0.250	0.200	-1.25	0.430	1.00
CookTime*Cool Time	-0.300	-0.150	0.200	-0.75	0.590	1.00

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	6	30.4550	5.0758	15.86	0.190
Linear	3	29.6500	9.8833	30.89	0.131
Salt	1	28.1250	28.1250	87.89	0.068
CookTime	1	0.2450	0.2450	0.77	0.542
Cool Time	1	1.2800	1.2800	4.00	0.295
2-Way Interactions	3	0.8050	0.2683	0.84	0.645
Salt*CookTime	1	0.1250	0.1250	0.39	0.644
Salt*Cool Time	1	0.5000	0.5000	1.56	0.430
CookTime*Cool Time	1	0.1800	0.1800	0.56	0.590
Error	1	0.3200	0.3200		
Total	7	30.7750			

3. Interpret the model using uncoded units – Minitab generates the regression equation in uncoded units:



Regression Equation in Uncoded Units

Taste = 5.80 - 6.00 Salt + 0.400 CookTime + 0.640 Cool Time - 0.400 Salt*CookTime - 0.800 Salt*Cool Time - 0.0240 CookTime*Cool Time

You get the best of both worlds.

11.1.5 Multiple Factor Experiments – Fractional Designs

Reducing the Number of Experiments

The full-factorial experiment allows you to detect the *main effects* associated with factors as well as all *interactions between factors.* For a two-factor, two-level experiment, the number of runs required (for one replication) is four. As the number of factors increases, the number of experiments required by the full factorial can become large. For example, if you are dealing with 5 factors (again, each at two levels), one replication will require 32 experiments.

Your budget may not allow the cost of this experiment. In addition, you may not be worried about detecting higher-order interactions – i.e. the *AxBxCxDxE* interaction may not be high on your priority list. The *physics* or *chemistry* of your world should be considered. For mechanical devices, main effects and two-way interactions are generally important, with higher-order interactions being not significant. For chemical formula, two and three-way interactions are often seen, but four and five-way interactions will not be significant. Social science experiments will likely see two-way interactions.

If we are dealing with a quality characteristic which is *additive* in terms of its factors (resulting in no interactions), or if we are willing to give up the ability to detect some or all interactions, then we can examine alternatives to the full-factorial experiment that *reduce* the number of experiments. The same 11-step process applies, we'll principally be changing step 4, *Develop an Experimental Layout.*

In the rest of this unit, we'll cover:

- Fractional Factorial Designs
- Screening Designs
- Orthogonal Arrays

The Fractional Factorial Experiment

Fractional factorial experiments, as their name implies, only conduct a fraction of the experiments associated with a *full factorial* layout. The basic idea behind a fractional factorial layout is to *start* with the full factorial and then choose only those factor level combinations that detect the effects of interest. For example, we may choose only those combinations that allow us to detect *main* effects (the effect of the factor itself) and *first order* interactions (AxB, BxC, etc.).

All higher order interactions are then *confounded*, that is, their variation is either mixed in with the main or first order effects or with the error term (SSe). We then give up the ability to detect these higher order interactions. Suppose our daughter rebels at conducting the 16 experiments required by the full factorial design to test the four factors of salt, paprika, mustard and cooking time. A *one-half* fractional factorial experiment requires only 8 runs:

Trial	Salt	Paprika	Mustard	Cook Time	Taste
1	1	-1	-1	-1	
2	2	-1	-1	1	
3	1	1	-1	1	
4	2	1	-1	-1	
5	1	-1	1	1	
6	2	-1	1	-1	
7	1	1	1	-1	
8	2	1	1	1	

For this design, the main effects are clear of each other and clear of all two-way interactions. So, if two-way interactions are present, they will not be confounded with the main effects. The two-way interactions are confounded with each other, and the main effects are confounded with three-way interactions. This is referred to as a *Resolution IV* design.

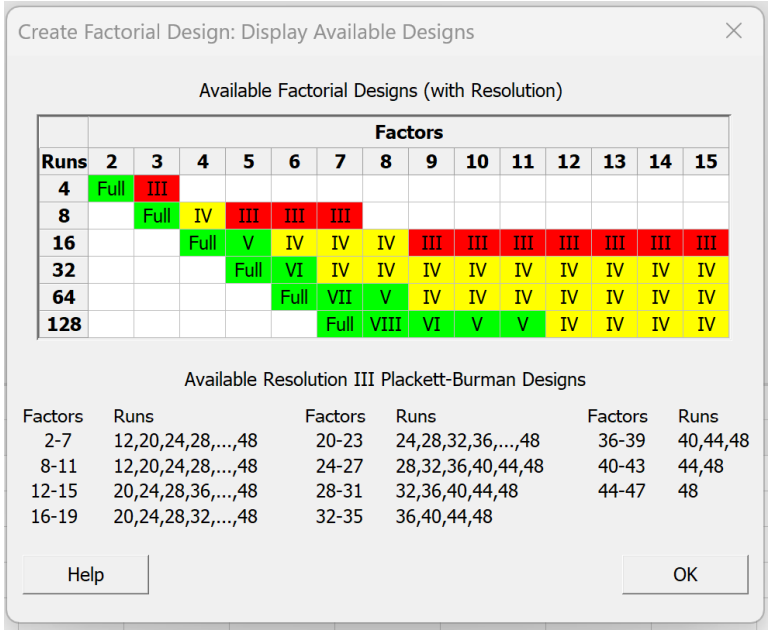
The ***Alias Structure*** lists which main effects and interactions are confounded with each other. Each fractional factorial design will have a unique alias structure. The alias structure for our half-fractional factorial experiment appears below:

Alias Structure	Description
A + BCD B + ACD C + ABD D + ABC	Shows how the Main Effects are confounded with three-way interactions
AB + CD AC + BD AD + BC	Shows how two-way interactions are confounded with each other

To relate this table to our deviled egg experiment, if A = Salt, B = Paprika, C = Mustard and D = Cook Time, then this half fractional design could detect all the main effects without confounding with two-way interactions. If there was a *Paprika x Mustard x Cook Time* interaction (BCD - three-way), though, this would be confounded with the main effect *Salt.* If the two two-way interactions of *Salt x Paprika* and *Mustard x Cook Time* existed, we could not separate these since they are

confounded (AB + CD). As you can see, the economy of fewer experiments must be balanced with the loss of ability to detect interactions. Practically, though, in many cases, three-way and higher-order interactions simply don't exist.

You might ask, "*Since we reduced the full factorial to a half-fractional design, is it possible to reduce the number of experiments even further*?" A *quarter-fractional* design, as its name implies, requires only ¼ of the trials of a full factorial, but with even more confounding. Minitab will tell you what designs are available for the number of factors in your experiment. From the table below, if you have 3 factors, you can create a full or half-fractional design. If you have 5 factors, you can create a full, half-fractional and quarter fractional design:



	Factors													
Runs	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
4	Full	III												
8		Full	IV	III	III	III								
16			Full	V	IV	IV	IV	III	III	III	III	III	III	III
32				Full	VI	IV	IV	IV	IV	IV	IV	IV	IV	IV
64					Full	VII	V	IV	IV	IV	IV	IV	IV	IV
128						Full	VIII	VI	V	V	IV	IV	IV	IV

Factors	Runs	Factors	Runs	Factors	Runs
2-7	12,20,24,28,...,48	20-23	24,28,32,36,...,48	36-39	40,44,48
8-11	12,20,24,28,...,48	24-27	28,32,36,40,44,48	40-43	44,48
12-15	20,24,28,36,...,48	28-31	32,36,40,44,48	44-47	48
16-19	20,24,28,32,...,48	32-35	36,40,44,48		

For example, adding *Vinegar* as a fifth factor, we could chose a quarter factorial experiment that results in an 8-trial design per replication. This design allows us to detect each of the main effects (clear of each other), but the main effects are confounded with two-way and higher order interactions.

All two-way and higher order interactions are either confounded with main effects or with each other. This is known as a *Resolution III* design. We will discuss this type of design below as an *Orthogonal Array*:

Trial	Salt (A)	Paprika (B)	Mustard (C)	Cook Time (D)	Vinegar (E)	Taste
1	-1	-1	-1	1	1	
2	1	-1	-1	-1	-1	
3	-1	1	-1	-1	1	
4	1	1	-1	1	-1	
5	-1	-1	1	1	-1	
6	1	-1	1	-1	1	
7	-1	1	1	-1	-1	
8	1	1	1	1	1	

The Alias Structure for this design is as follows:

Alias Structure	Description
A + BD + CE + ABCDE B + AD + CDE + ABCE C + AE + BDE + ABCD D + AB + BCE + ACDE E + AC + BCD + ABDE	Shows how the Main Effects are confounded with three-way interactions
BC + DE + ABE + ACD BE + CD + ABC + ADE	Shows how two-way interactions are confounded with each other and higher order interactions

Although tables of fractional factorial experiments are published, software packages such as Minitab have the capability to create your experimental design. You simply specify the number of factors, levels, and type of design. The software will then develop the appropriate design. To analyze a fractional factorial experiment, we still employ **Analysis of Variance (ANOVA)**. The *model* employed is based on the terms (main effects and interactions) that can be detected by the resolution of the experiment.

Example Half-Fractional Design

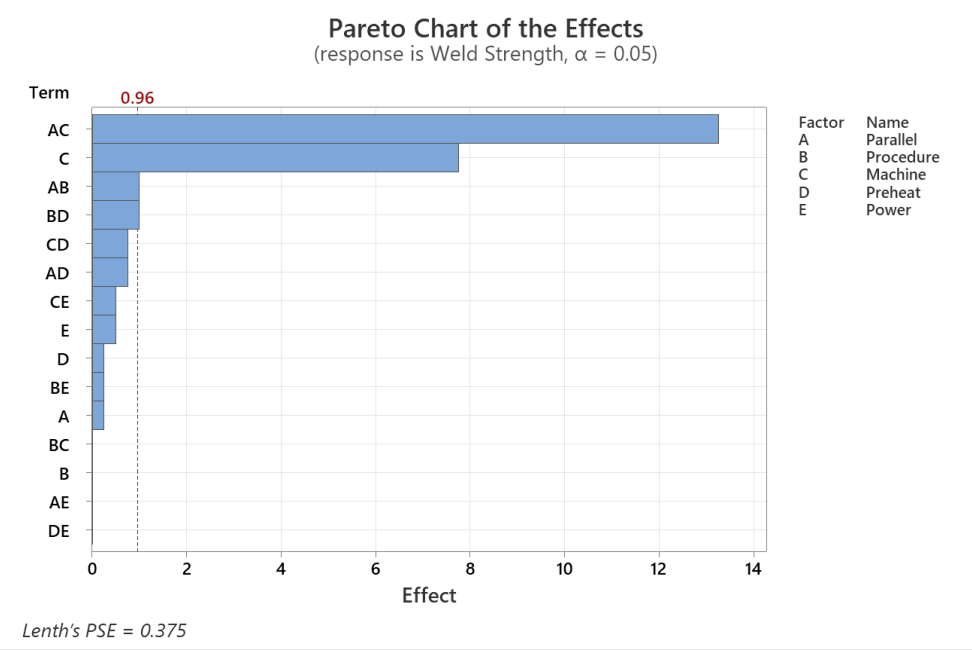
A team performed a half-fractional experiment to determine the effect of five factors (Parallel, Procedure, Machine, Preheat and Power) on the Weld Strength of a welded joint. A full-factorial design would require 32 runs, but here their design only required 16 runs. The Minitab design appears below:

	C1	C2	C3	C4	C5-T	C6-T	C7-T	C8-T	C9	C10
	StdOrder	RunOrder	CenterPt	Blocks	Parallel	Procedure	Machine	Preheat	Power	Weld Strength
1	1	1	1	1	Yes	Std	A	No	3	49
2	2	2	1	1	No	New	A	No	3	37
3	3	3	1	1	No	Std	B	No	3	44
4	4	4	1	1	Yes	New	B	No	3	31
5	5	5	1	1	No	Std	A	Yes	3	38
6	6	6	1	1	Yes	New	A	Yes	3	51
7	7	7	1	1	Yes	Std	B	Yes	3	30
8	8	8	1	1	No	New	B	Yes	3	41
9	9	9	1	1	No	Std	A	No	5	38
10	10	10	1	1	Yes	New	A	No	5	52
11	11	11	1	1	Yes	Std	B	No	5	29
12	12	12	1	1	No	New	B	No	5	44
13	13	13	1	1	Yes	Std	A	Yes	5	52
14	14	14	1	1	No	New	A	Yes	5	37
15	15	15	1	1	No	Std	B	Yes	5	43
16	16	16	1	1	Yes	New	B	Yes	5	30

If we attempt to analyze this design using Minitab, we'll find that there are no degrees of freedom available to calculate a Mean Square Error. We even have to remove some terms just to get a Pareto chart of the effects. We remove the higher-order (four-way and higher-order, not likely significant) interactions and re-run the analysis (see Pareto chart, next page).

The chart tells us that terms AC, C, AB & BD may be significant (effect sizes greater than the critical value of 0.96).

We'll re-run the analysis, leaving just these terms in the model.



After re-running with just significant factors (note that if a main effect is involved in an interaction, it must be left in the model). The t-tests, ANOVA and regression models are shown below:

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		40.375	0.230	175.82	0.000	
Parallel	0.250	0.125	0.230	0.54	0.601	1.00
Procedure	-0.000	-0.000	0.230	-0.00	1.000	1.00
Machine	-7.750	-3.875	0.230	-16.87	0.000	1.00
Preheat	-0.250	-0.125	0.230	-0.54	0.601	1.00
Parallel*Procedure	1.000	0.500	0.230	2.18	0.061	1.00
Parallel*Machine	-13.250	-6.625	0.230	-28.85	0.000	1.00
Procedure*Preheat	-1.000	-0.500	0.230	-2.18	0.061	1.00

Analysis of Variance

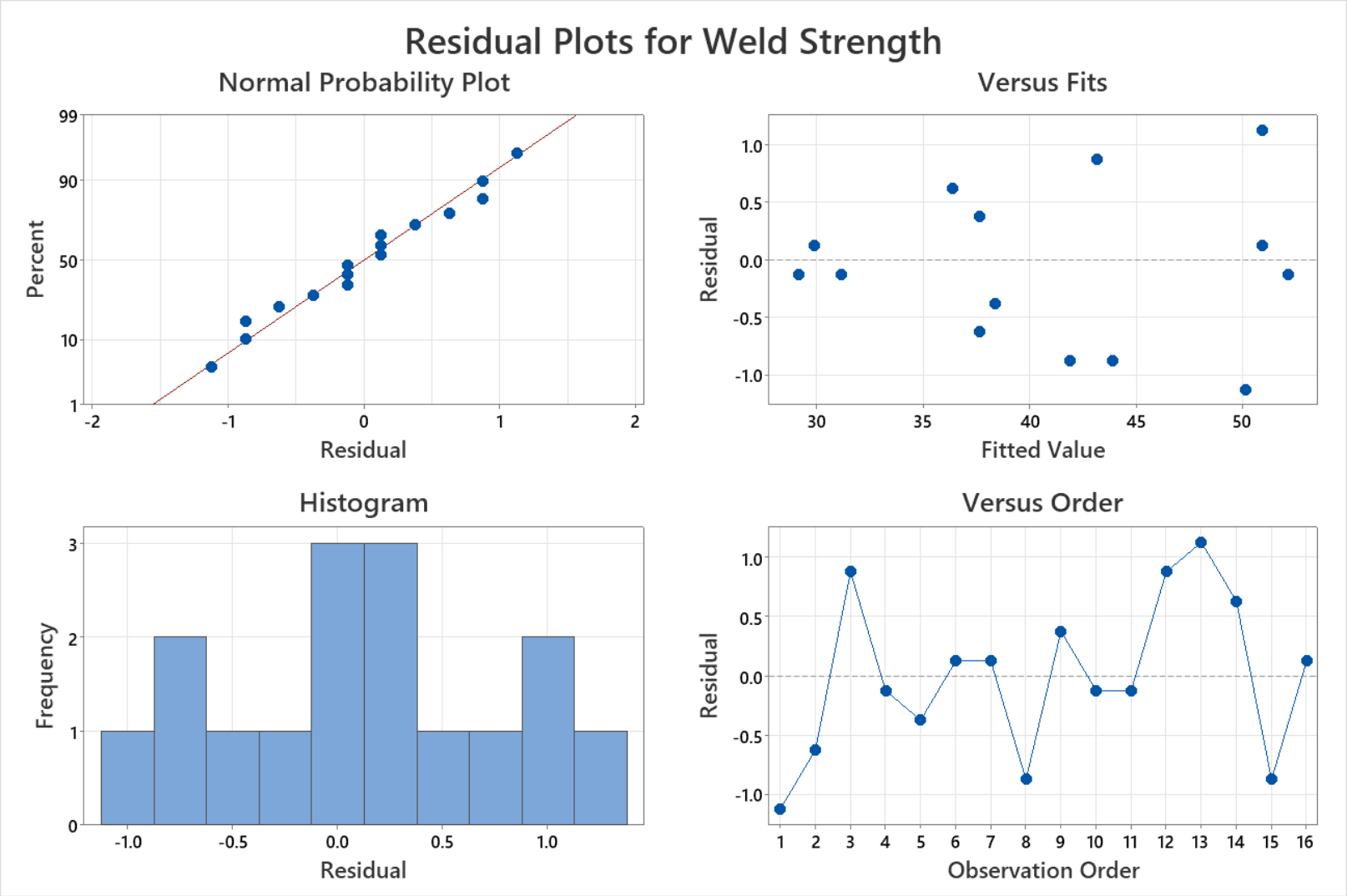
Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	7	951.000	135.857	161.02	0.000
Linear	4	240.750	60.187	71.33	0.000
Parallel	1	0.250	0.250	0.30	0.601
Procedure	1	0.000	0.000	0.00	1.000
Machine	1	240.250	240.250	284.74	0.000
Preheat	1	0.250	0.250	0.30	0.601
2-Way Interactions	3	710.250	236.750	280.59	0.000
Parallel*Procedure	1	4.000	4.000	4.74	0.061
Parallel*Machine	1	702.250	702.250	832.30	0.000
Procedure*Preheat	1	4.000	4.000	4.74	0.061
Error	8	6.750	0.844		
Total	15	957.750			

Regression Equation in Uncoded Units

Weld Strength = 40.375 + 0.125 Parallel - 0.000 Procedure - 3.875 Machine - 0.125 Preheat
+ 0.500 Parallel*Procedure - 6.625 Parallel*Machine - 0.500 Procedure*Preheat

The residual plots appear below.

- The normal probability plot does not show significant departure from normality
- The Histogram mirros the normal probability plot – no issues here
- Residuals vs. Fits does seem to indicate that the variance of the residuals isn't constant (heteroscedasticity)
- Residuals vs. Order does not show any issues.


Residual Plots for Weld Strength
Normal Probability Plot
Percent
99
90
50
10
1
-2
-1
0
1
2
Residual
Versus Fits
Residual
1.0
0.5
0.0
-0.5
-1.0
30
35
40
45
50
Fitted Value
Histogram
Frequency
3
2
1
0
-1.0
-0.5
0.0
0.5
1.0
Residual
Versus Order
Residual
1.0
0.5
0.0
-0.5
-1.0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Observation Order

Screening Experiments

If you have many factors, even a fractional factorial experiment may still need many trials. This case usually arises at the beginning of an experimental program, where little is known about the important factors affecting some outcome. You are generally interested in just understanding which factors may influence the effect through their *main effects.* In this case, you can employ *screening designs*, such as the Plackett & Burman (PB) designs.

The Plackett & Burman designs are *saturated* and allow you to detect main effects of many factors. The experimental layout appearing below can detect the main effects of twenty factors and requires (with one replication) twenty-four trials. Statistical software such as Minitab has the capability of designing the layout quickly – you need merely specify the number of factors in the experiment.

Although these experiments are *very* efficient, if interactions exist, they will be confounded with the main effects. You will not be able to determine if you are seeing the effect of a single factor or an interaction between (or among) factors. Nevertheless, at the beginning of an experimental program, screening designs can be used to quickly determine which factors are or are not important in influencing the effect. Once this is determined, optimization or further experimentation may occur using Fractional or Full-Factorial experiments.

Plackett-Burman Experimental Layout

Trial	A	B	C	D	E	F	G	H	J	K	L	M	N	O	P	Q	R	S	T	U
1	1	-1	1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1
2	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1
3	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1
4	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1
5	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	1	1	1	1	-1	1	1
6	-1	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	1	1	1	-1	-1
7	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	-1	-1	1	1
8	-1	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	-1	1	-1
9	1	-1	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	1	1
10	1	1	-1	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1	1
11	-1	1	1	-1	1	-1	1	1	1	1	1	-1	-1	-1	-1	1	-1	1	-1	-1
12	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	-1	-1	-1	-1	1	-1	1	-1
13	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	-1	1	-1	-1	1	-1	1
14	1	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	-1	-1	-1	-1	1	-1
15	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	-1	-1	-1	-1	1
16	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	-1	-1	-1	-1
17	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	1	1	1
18	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1	-1	1	1	-1
19	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	-1	1	-1	1	1
20	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1	1	1
21	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1	1
22	-1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1	1
23	-1	-1	-1	-1	1	-1	1	-1	-1	1	1	-1	-1	1	1	-1	1	-1	1	1
24	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

Twenty (20) Factors in the Experimental Layout

Each Factor Run at Two (2) Levels

Twenty-four (24) Trials Required

Screening Design Example

A team is trying to improve **adhesive bond strength** (response Y, in MPa). They suspect many variables matter, so they want a **fast screening design** to identify the "vital few" main effects. Here, they choose a **PB 12-run design** to screen **11 factors** (PB designs use Nruns to estimate up to $N-1$main effects).

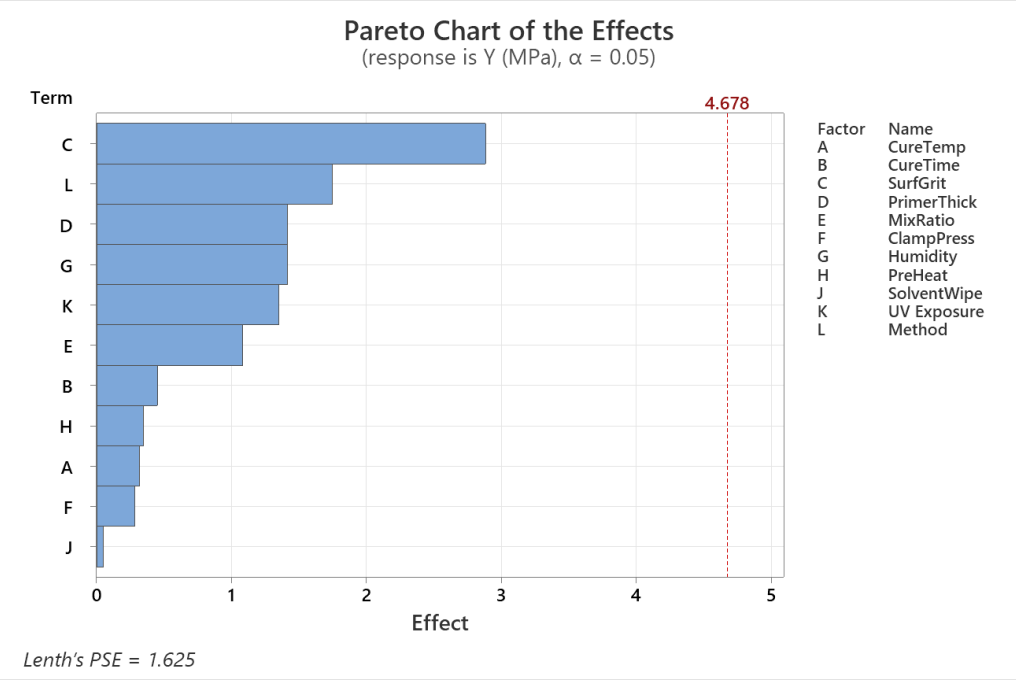
The Factors and levels are shown below:

Factor	Description	Low (−1)	High (+1)
A	Cure temperature (°C)	120	150
B	Cure time (min)	20	40
C	Surface grit (ANSI)	120	220
D	Primer thickness (µm)	5	15
E	Mix ratio (resin:hardener)	95:5	90:10
F	Clamp pressure (kPa)	50	150
G	Humidity (%)	30	70
H	Part preheat (°C)	20	60
I	Solvent wipe (Y/N)	No	Yes
J	UV exposure pre-cure (s)	0	30
K	Operator method	Method 1	Method 2

Here's the experimental layout in Minitab, with response (Y (MPa)) recorded:

C1	C2	C3	C4	C5	C6	C7	C8	C9-T	C10	C11	C12	C13-T	C14	C15-T	C16
StdOrder	RunOrder	PtType	Blocks	CureTemp	CureTime	SurfGrit	PrimerThick	MixRatio	ClampPress	Humidity	PreHeat	SolventWipe	UV Exposure	Method	Y (MPa)
5	1	1	1	150	40	120	15	90:10	50	70	20	No	0	Method2	25.6
4	2	1	1	150	20	220	15	95:5	150	30	20	No	30	Method2	18.7
9	3	1	1	120	20	120	15	90:10	150	30	60	Yes	0	Method2	24.1
8	4	1	1	120	20	220	15	90:10	50	70	60	No	30	Method1	19.2
12	5	1	1	120	20	120	5	95:5	50	30	20	No	0	Method1	22.0
2	6	1	1	150	40	120	15	95:5	50	30	60	Yes	30	Method1	20.4
3	7	1	1	120	40	220	5	90:10	50	30	20	Yes	30	Method2	21.1
10	8	1	1	150	20	120	5	90:10	150	70	20	Yes	30	Method1	23.8
11	9	1	1	120	40	120	5	95:5	150	70	60	No	30	Method2	24.9
7	10	1	1	120	40	220	15	95:5	150	70	20	Yes	0	Method1	19.9
1	11	1	1	150	20	220	5	95:5	50	70	60	Yes	0	Method2	23.0
6	12	1	1	150	40	220	5	90:10	150	30	60	No	0	Method1	21.6

The initial analysis does not have sufficient degrees of freedom to estimate a Mean Square Error, so we look to the Pareto diagram to see what factors may be removed – B, H, A, F and J have small effects and will be removed:



Rerunning the analysis, we see all the remaining terms are significant. Here are the two-sample t-tests:

Coded Coefficients

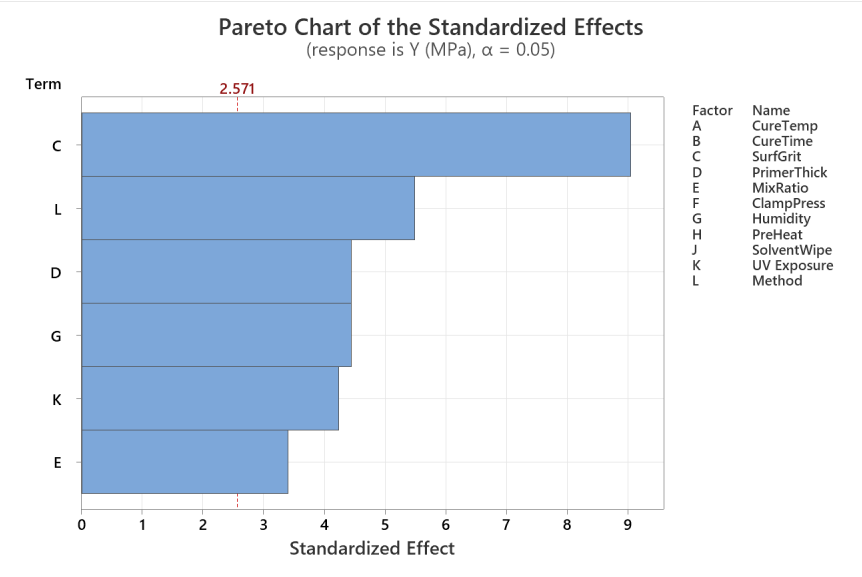
Term	Effect	Coef	SE Coef	T-Value	P-Value	VIF
Constant		22.025	0.159	138.19	0.000	
SurfGrit	-2.883	-1.442	0.159	-9.05	0.000	1.00
PrimerThick	-1.417	-0.708	0.159	-4.44	0.007	1.00
MixRatio	1.083	0.542	0.159	3.40	0.019	1.00
Humidity	1.417	0.708	0.159	4.44	0.007	1.00
UV Exposure	-1.350	-0.675	0.159	-4.24	0.008	1.00
Method	1.750	0.875	0.159	5.49	0.003	1.00

Here is the ANOVA table:

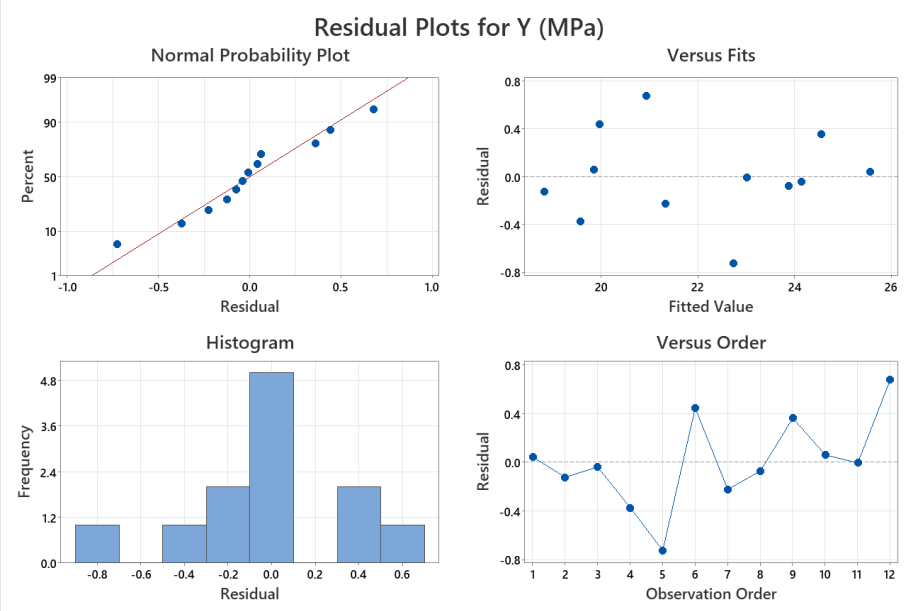
Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	6	55.158	9.1931	30.16	0.001
Linear	6	55.158	9.1931	30.16	0.001
SurfGrit	1	24.941	24.9408	81.82	0.000
PrimerThick	1	6.021	6.0208	19.75	0.007
MixRatio	1	3.521	3.5208	11.55	0.019
Humidity	1	6.021	6.0208	19.75	0.007
UV Exposure	1	5.468	5.4675	17.94	0.008
Method	1	9.188	9.1875	30.14	0.003
Error	5	1.524	0.3048		
Total	11	56.683			

The Pareto Chart shows that Surface Grit has the largest effect, with Operator Method second. Primer Thickness, Humidity (is this a noise variable?) and UV exposure all have about the same effect with Mix Ratio the smallest.:



The residuals appear to be well-behaved:



Finally, Minitab provides the regression equation:

Regression Equation in Uncoded Units

Y (MPa) = 27.248 - 0.02883 SurfGrit - 0.1417 PrimerThick + 0.542 MixRatio + 0.03542 Humidity - 0.0450 UV Exposure + 0.875 Method

Remember that the Plackett-Burman only considers main effects (and these are confounded with interactions), so the regression equation, while valid, has limited meaning. For example, Surface Grit (Factor C) is aliased with these terms (only up to order two). We don't know if Surface Grit is significant by itself or is the effect due to an interaction:

C	- 0.33 AB	- 0.33 AD	- 0.33 AE	+ 0.33 AF	- 0.33 AG	+ 0.33 AH	- 0.33 AJ	- 0.33 AK	+ 0.33 AL
	- 0.33 BD	+ 0.33 BE	+ 0.33 BF	- 0.33 BG	- 0.33 BH	+ 0.33 BJ	- 0.33 BK	- 0.33 BL	- 0.33 DE
	+ 0.33 DF	+ 0.33 DG	- 0.33 DH	- 0.33 DJ	+ 0.33 DK	- 0.33 DL	- 0.33 EF	- 0.33 EG	+ 0.33 EH
	- 0.33 EJ	+ 0.33 EK	- 0.33 EL	- 0.33 FG	- 0.33 FH	- 0.33 FJ	- 0.33 FK	- 0.33 FL	+ 0.33 GH
	+ 0.33 GJ - 0.33 GK - 0.33 GL - 0.33 HJ - 0.33 HK - 0.33 HL - 0.33 JK + 0.33 JL + 0.33 KL								

Here's what we can say about a significant factor:

A significant PB factor indicates that the factor is involved in at least one important effect—either as a main effect, as part of one or more interactions, or both. Changing this factor from its low level to its high level produces a consistent, non-random change in the response across the sampled conditions.

Here's four "Take-aways:"

1. You've learned the factor *matters in practice*, so a significant factor tells you:
   - Turning this "knob" *does something reproducible*
   - The effect survives randomization
   - It is large relative to noise and other effects
   - From an engineering standpoint, that's already powerful.

2. The effect is *directionally reliable*
   - If factor A has a positive effect:
     - High A → higher Y *on average*
     - Low A → lower Y *on average*
   - Even if part of that comes from interactions, the direction is real in the region you tested.
   - This is why PB designs are excellent for decision-making, not explanation.

3. Insignificant factors are usually safe to drop
   - PB designs are asymmetric in risk:
     - False positives can happen (interaction masquerading as a main effect)
     - False negatives are less common unless effects are very small
   - So when a factor looks unimportant, it's often genuinely unimportant—at least at the tested levels.

4. When a PB factor is "significant," you have not learned:
   - That the effect is purely a main effect
   - That interactions are negligible
   - That the regression coefficient is unbiased
   - That the factor will behave the same outside the tested range

PB cannot answer *why* the factor matters—only *that it does*.

The team's next steps should be to plan their next round of experiments with the significant factors. For example, the team could progress their design using one of the following options:

- Resolution IV fractional factorial → separate main effects from 2FIs
- Foldover → break aliasing
- Response surface → quantify curvature and optimization

Fractional Factorial Designs - Resolution

This table provides you with a guide to the ability of your experimental design to detect main and interaction effects. Note that *Resolution V* is usually good enough for practical experimentation, since three-way interactions are rare. Instead of conducting an experiment with a higher resolution, you should consider putting these "extra" experimental resources into replication.

Resolution	Interpretation Effect . . .	 is Confounded With	Comments
III	Main Effects +	2-way and higher order interactions	Main effects are clear of each other, but confounded with two-way interactions
	1 +	**2 = 3**	
IV	Main Effects +	3-way and higher order interactions	Main effects are clear of each other and clear of all two-way interactions (interpretation of main effects is not affected by presence of two-way interactions
	1 +	**3 = 4**	
	2-way interactions +	Other 2-way and higher order interactions	
	2 +	**2 = 4**	
V	Main Effects +	4-way and higher interactions	Main effects are clear of each other, of all two-way interactions, and of all three-way interactions. Two-way interactions are clear of each other.
	1 +	**4 = 5**	
	2-way interactions +	3-way and higher order interactions	
	2 +	**3 = 5**	

One way to remember this table is the "hand-rule." For a Resolution V experiment, hold up 5 fingers. Now separate one finger (the main effects) from the other four fingers. This tells you that the main effects are confounded with 4 way interactions. Separating two fingers tells you that the 2-way interactions are confounded with 3 way interactions. Try this for Resolution IV and III experiments.

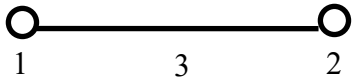
11.1.6 Multiple Factor Experiments – Orthogonal Arrays

Orthogonal Arrays

Orthogonal Arrays are a special class of the fractional factorial experiments. These experimental layouts help you determine main effects and some interactions in only a very few experiments. Let's introduce the Orthogonal Array through its simplest case, the L_4 (the subscript refers to the number of experiments or trials that are required). The L_4 is essentially a full factorial layout for two factors but examining it will help us understand the Orthogonal Array's (OA) notation and use. We'll present the theory here and an example in Unit 14.4 – Taguchi Design Approach.

The L_4 layout consists of two components, an experimental layout and a linear graph. The experimental layout tells you what combinations of factor levels to run as experiments, the linear graph helps you *design* your experiment - identifying the combinations of factor levels that will be run as experiments. Here's the L_4 Orthogonal Array and its linear graph:

	Column (Factor/Interaction)		
Trial #	**1**	**2**	**3**
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1



To design an experiment using the L_4 OA, we first assign our factors/ interactions to the *columns* of the OA. The linear graph helps us figure out what factors to assign to which columns. The numbers on the linear graph correspond to the OA's columns. The *factors* are assigned to the *circles* of the graph. The *lines* joining the circles then tell us which column to assign to the interaction between these factors. Using the L_4 OA's linear graph, we assign factor A to column 1 and factor B to column 2. Column 3 then picks up the AxB interaction:

	Column (Factor/Interaction)		
Trial #	**1 (A)**	**2 (B)**	**3 (AxB)**
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

The left-hand column of the L_4 OA shows that we will have to perform four experiments or trials. The rows then show you the combination of factor levels to run for that trial. Although you can do the reverse, it's easy to associate "1" with the low level of the factor and "2" with the high level.

For this L_4 OA, trial #1 is run with both factors at their low levels. Trial #2 is run with factor A at its low level and factor B at its high level. Trial #3 is run with factor A at its high level and B at its low level, etc. The *order* of the experiments should be randomized, of course.

Notice that we don't worry about "setting" the levels of the AxB interaction. The "1's" and "2's" in the interaction column (3) are used to calculate the sum of squares for the interaction.

Suppose that we were not concerned about the AxB interaction. Could we "squeeze in" another factor, C, to the L_4 OA? The answer is yes, and the layout would appear as follows:

	Column (Factor/Interaction)		
Trial #	**1 (A)**	**2 (B)**	**3 (C)**
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

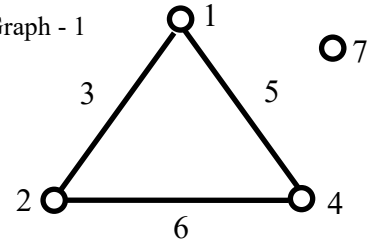
Any AxB interaction is *confounded* with the main effect of factor C; that is, we cannot separate these two effects. We also cannot detect the AxC, BxC and AxBxC interactions.

There is another, not so obvious concern with this layout. We cannot calculate the *error term* (the Sum of Squares - error). This prevents us from performing an ANOVA, since the test statistic is the ratio of the factor's Mean Sum of Squares to the error's Mean Sum of Squares. There are other techniques that can be used to analyze this situation. For example, Taguchi employs the *Signal-to-Noise* ratio as the principal means of analyzing the results of experiments.

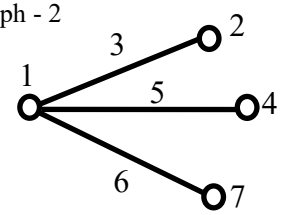
Recall, though, that a full factorial design for three factors at two levels (2^3) would require us to run eight experiments. At the "expense" of losing the ability to detect interactions, we have cut the number of required experiments in half! Now that we've explored the concept of the Orthogonal Array, let's introduce the L_8 OA, one often used for experimentation:

	Column (Factor/Interaction)						
Trial	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

L_8 Linear Graph - 1

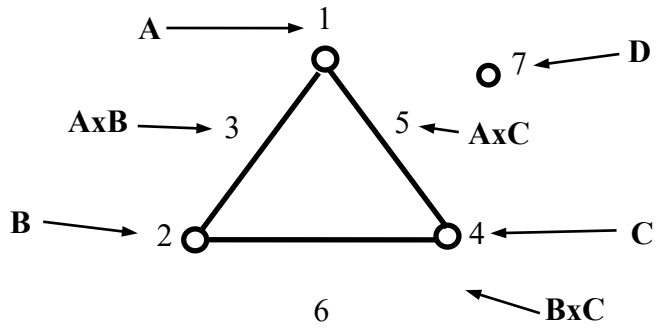


L_8 Linear Graph - 2



Up to seven factors may be assigned to the L_8 OA. Eight trials are required (as opposed to the 128 trials required for a full factorial!). Two linear graphs are provided to help you assign factors to columns.

Suppose you wish to investigate four factors (A, B, C, & D), and you are concerned about three possible interactions for these factors - AxB, BxC and AxC. The first linear graph is used to assign factors to columns:



The Orthogonal Array columns would then appear as follows:

	Column (Factor/Interaction)						
Trial	**1 (A)**	**2 (B)**	**3 (AxB)**	**4 (C)**	**5 (AxC)**	**6 (BxC)**	**7 (D)**
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

The eight experiments would be run with the factors at the following “low/high” levels. The trials are run in a randomized order:

	Factor				
Trial	**A**	**B**	**C**	**D**	**Result**
1	1	1	1	1	X_1
2	1	1	2	2	X_2
3	1	2	1	2	X_3
4	1	2	2	1	X_4
5	2	1	1	2	X_5
6	2	1	2	1	X_6
7	2	2	1	1	X_7
8	2	2	2	2	X_8

Analyzing an Orthogonal Array's Experiment's Results

Analysis of Variance is used to determine which factors/interactions are significant. The ANOVA is performed as follows:

1. Calculate the Total Sum of Squares:

$$SST = \sum_{i=1}^{N} X_i^2 - \frac{\left(\sum_{i=1}^{N} X_i\right)^2}{N}$$

where:

N - *Number of Trials*

2. Calculate the Sum of Squares for each column (i.e. this will correspond to either a factor or an interaction):

$$SSi = \frac{\left(\sum_{i=1}^{N/2} X_{i1} - \sum_{i=1}^{N/2} X_{i2}\right)^2}{N}$$

where:

SSi - *Sum of Squares for ith column*

X_{i1} - *ith value of the result for factor level 1*

X_{i2} - *ith value of the result for factor level 2*

Notes: 1. Use the OA column assigned to that factor to obtain the 1's and 2's

2. This equation is good for the 2 - level Orthogonal Arrays only

3. Calculate the Sum of Squares - error. This is the sum of the Sum of Squares for any columns left unassigned to a factor or interaction.

4. Develop the ANOVA table. Each factor/interaction has one degree of freedom; the error has N-K-1 degrees of freedom (where K is the number of columns to which factors and/or interactions have been assigned) and the total sum of squares has N-1 degrees of freedom. Calculate the Mean Sum of Squares, develop the F test statistics and perform the F-tests on these ratios. Conclude which factors/interactions are significant.

11.1.7 Miscellaneous Experimental Design Topics

Blocking in Experiments

Consider the following experimental situation: A company wishes to determine which of four brands of tires provides the longest life for their service representatives' cars. How can we design an experiment to help the agency answer their question?

Essentially, this is a single factor experiment. Two experimental plans that could be considered include:

1. *Pick 4 cars at random from the company's fleet. Equip each car with 4 tires from each manufacturer, selecting the brands at random.*

2. *Pick 4 cars at random from the fleet. Take 4 sets of tires, one from each brand, and assign these randomly to the 16 wheels of the 4 cars.*

What problems do these plans generate? To answer this question, we must consider "where" the variation is going. For example, the first plan will allow us to detect car-to-car variation. But we don't know whether this variation is due to the tires, to the cars, to the service rep driving the car, or to the driving conditions encountered by the driver on her travels. To conclude that any significant variation is due to the tires is likely to be met with skepticism.

The second plan seems to get around this problem by assigning tires not to the cars, but to the wheel positions. But what if our random assignment produced the following experimental layout?

	Wheel			
Car	**Right Front**	**Left Front**	**Right Rear**	**Left Rear**
1	A	A	C	B
2	D	C	D	C
3	B	D	B	B
4	A	C	D	A

A, B, C, D - Tire Brands

Although the tires have been randomized, notice that car #1 does not have any Brand D tires, car #2 does not have any Brand A or B tires, etc. A significant difference in tire A's performance may be due to car-to-car differences (i.e. cars 1 and 4 vs. 2 and 3). Plan 2, then is also not a good design.

To address this issue, we can consider each car to be a *block.* If we randomly assign tires to each block, so that one tire from each brand is assigned to each block, then we avoid the problems discussed above. The following layout makes use of this blocking principle:

	Wheel			
Car	**Right Front**	**Left Front**	**Right Rear**	**Left Rear**
1	C	D	B	A
2	A	D	C	B
3	D	A	C	B
4	B	C	A	D

A, B, C, D - Tire Brands

Notice that each car (the *block*) has one of each brand of tire. The analysis of this experiment would be conducted via a Two-Way ANOVA, although we are only interested in the significance of the *tire* factor. You might argue that there are not only car-to-car differences but also tire *position* differences. We could develop a layout where each tire brand appears once on each car *and* appears only once in each position. This further refinement of the *blocking* principle is called a *Latin Square* layout:

	Wheel			
Car	**Right Front**	**Left Front**	**Right Rear**	**Left Rear**
1	D	A	B	C
2	C	D	A	B
3	B	C	D	A
4	A	B	C	D

A, B, C, D - Tire Brands

In this design, there is nothing *random* about the assignment of the tires to car/position. There are, however, different Latin squares for a given size, and the particular Latin square used may be selected at random.

Nesting in Experiments

Consider the following experimental situation: A company is investigating different joining methods for sheet metal. The two methods of interest are *welding* and *gluing.* Welding can be done using one of two suppliers' machines. Three different types of glue are also available for testing.

This is a two-factor experiment (A = joining method, B = mode within the method). The twist here is that we cannot "mix" certain levels of one factor with another. For example, the sheet metal can't be welded with glue "A."

This type of experiment is known as a *nested experiment*, since the levels of one factor are nested within, or are sub-samples of another factor. The overall variation can be decomposed into the between factor component (i.e. welding vs. gluing) and that within the factor (supplier A's vs. supplier B's machine). Interactions do not exist among the factors.

11.1.8 An Experimental Approach

At the beginning of this section, we described a ten-step approach to conducting an experiment. While these ten steps form the basis for a sound experimental approach, solving "real-world" problems will often require a series of experiments to be conducted. Designed experiments are often carried out in four phases: ***planning***, ***screening***, ***optimization***, and ***verification***.

Planning

You should carefully plan your experimentation. Consider the overall budget you have been provided to solve the problem. Develop the experimental approach with this in mind. Don't "blow your whole wad" in full-factorial experiments. Think about the value of replication versus resolution. Consider how the availability of personnel, equipment and material may affect your ability to complete the experiment. Will your experiments be conducted "on-line" or "off-line?" If the experimental units will be sold to customers ("on-line" experiments), your factor levels will have to fall within current specifications. If you can perform the experiments "off-line," you may generate scrap, but the experimental window will be larger. If your project has low priority, you may want to carry out small sequential experiments. That way, if you lose resources to a higher priority project, you will not have to discard the data you have already collected. When resources become available again, you can resume experimentation.

General steps to planning your approach include:

- *Define the problem*. Developing a good problem statement helps make sure you are studying the right variables.

- *Define the objective*. A well-defined objective will ensure that the experiment answers the right questions and yields practical, usable information.

- *Develop an experimental plan that will provide meaningful information*. Be sure to review relevant background information, such as theoretical principles, and knowledge gained through observation or previous experimentation. For example, you may need to identify which factors or process conditions affect process performance and contribute to process variability. Or, if the process is already established and the influential factors have been identified, you may be trying to determine optimal process conditions.

- Make sure the process and measurement systems are in control. Ideally, both the process and the measurements should be in statistical control as measured by an SPC system. Even if you do not have the process completely in control, you must be able to reproduce process settings. You also need to determine the variability in the measurement system. If the variability in your system is greater than the difference/effect that you consider important, experimentation will not yield useful results.

Screening

As discussed above, if little is known about the process, and many potentially important factors have been identified, a screening approach can be employed to determine the truly important factors. Screening reduces the number of variables by identifying the key factors that affect product quality. This reduction allows you to focus process improvement efforts on the "vital few" factors. Screening may also suggest the "best" or optimal settings for these factors, and indicate whether or not curvature exists in the responses. Then, full factorial or response surface experiments can be done to determine the best settings and define the nature of the curvature. The following methods are often used for screening:

- Two-level full and fractional factorial designs
- Plackett-Burman (PB) designs
- Full factorial designs (for small screening experiments)

Optimization

After you have identified the "vital few" by screening, you need to determine the "best" or optimal values for these experimental factors. Optimal factor values depend on the process objective. For example, you may want to maximize process yield or reduce the product variability.

Here you may employ general full factorial designs (designs with more than two-levels), response surface and mixture designs (beyond the current scope of this manual). A brief description follows:

Response Surface Designs are used when the effect responds non-linearly to changes in the factors. For example, the yield of a process may be influenced by temperature, pressure and the amount of catalyst employed. The yield may be a

linear function of temperature and catalyst, but a quadratic function of the pressure. Response surface methods first help you "climb the mountain" towards the peak yield and then understand the curvature at the peak.

Mixture Designs, as their name suggests, are employed when the effect is a function of mixing two or more components (factors) together. Here, the proportions of the components are important, and they must add to 1 or 100%.

Unit 14.4, Taguchi Design Approach, will describe an application of experimental design aimed at achieving both the product or process' targets and minimizing their variation in response to various noises expected to affect the product or process. Here, instead of randomizing the experiment to account for uncontrollable variation, an **inner-outer array design** approach will explicitly include the sources of variation in the experiment. Either a Signal-to-Noise ratio will be used as the response (this includes both mean and variance) or separate responses will be generated to understand the affect of the factors on the mean response and the variance response.

Verification

Verification involves performing a follow-up experiment at the predicted "best" processing conditions to confirm the optimization results. For example, you may perform a few verification runs at the optimal settings, and then obtain a confidence interval for the mean response. This is an essential part of all sound experimental plans.

A Guideline to Choosing Experimental Types

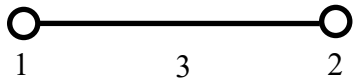
Type of Design	**Screening**	**Fractional Factorial**	**Full Factorial**	**Response Surface, Mixture, Taguchi**
Number of Factors	> 4	3 – 15	2 – 7	< 8
Purpose				
Identify	Most Important (Vital Few) Factors	Main Effects & Some Interactions	Relationships Among All Factors	Optimal Factor Settings
Estimate	Rough Direction for Improvement	All Main Effects & Some Interactions	All Main Effects & All Interactions	Curvature in Response, Empirical Models, Variation Response

11.1.9 Orthogonal Arrays

Here are the commonly used Orthogonal Arrays and their associated Linear Graphs. These first three are used for two-level factors:

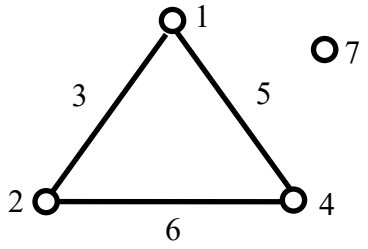
L_4 Orthogonal Array

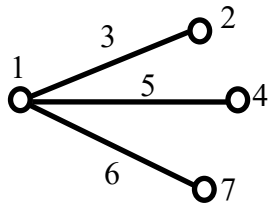
	Column (Factor/Interaction)		
Trial #	**1**	**2**	**3**
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1



L_8 Orthogonal Array

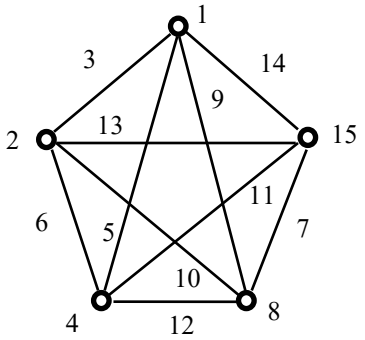
	Column (Factor/Interaction)						
Trial	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

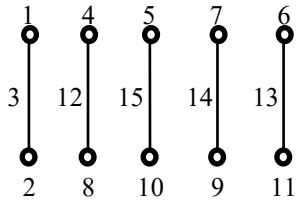


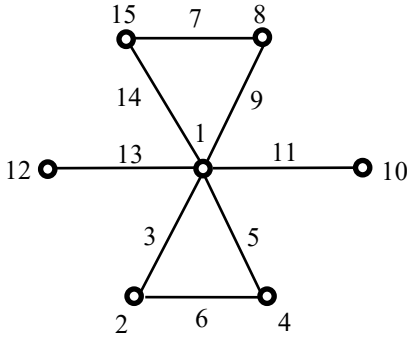


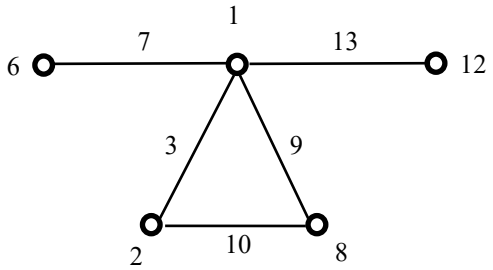
L_{16} Orthogonal Array

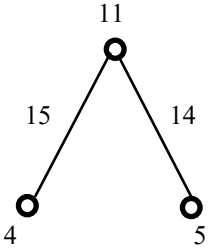
	Column (Factor/Interaction)														
Trial	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1







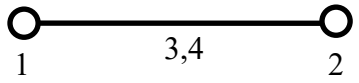




The following Orthogonal Arrays are used for factors with three levels:

L_9 Orthogonal Array

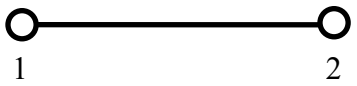
	Column (Factor/Interaction)			
Trial	**1**	**2**	**3**	**4**
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1



L_{18} Orthogonal Array

(Note that the first column has one factor at two levels, the remaining columns are used for three level factors. If the experiment involves more than one factor at two levels, a three level column can be used. Assign the more important level of that factor to the "3's" of that column).

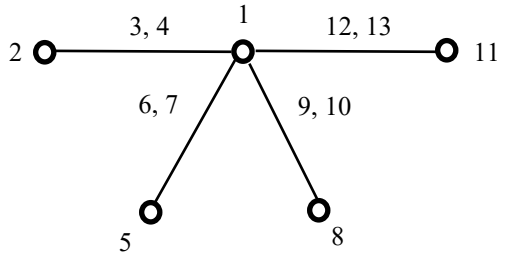
	Column (Factor/Interaction)							
Trial	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	1	1	1	1	1	1	1	1
2	1	1	2	2	2	2	2	2
3	1	1	3	3	3	3	3	3
4	1	2	1	1	2	2	3	3
5	1	2	2	2	3	3	1	1
6	1	2	3	3	1	1	2	2
7	1	3	1	2	3	3	2	3
8	1	3	2	3	1	1	3	1
9	1	3	3	1	2	2	1	2
10	2	1	1	3	2	2	2	1
11	2	1	2	1	3	3	3	2
12	2	1	3	2	1	1	1	3
13	2	2	1	2	1	1	3	2
14	2	2	2	3	2	2	1	3
15	2	2	3	1	3	3	2	1
16	2	3	1	3	3	3	1	2
17	2	3	2	1	1	1	2	3
18	2	3	3	2	2	2	3	1

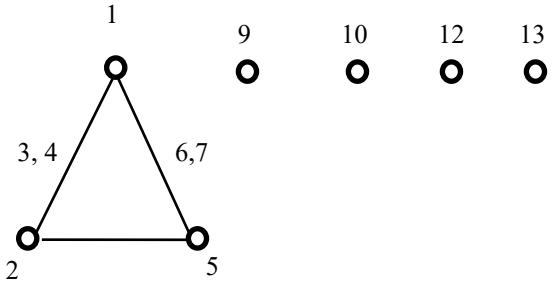


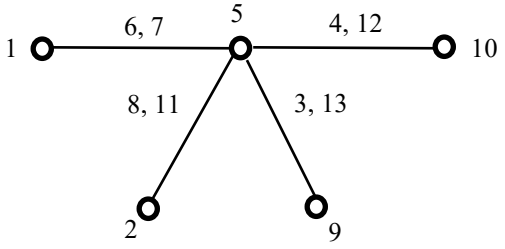
Interactions can be found by using the two-way layout of columns 1 and 2 (Two-way ANOVA of columns 1 & 2 for only 1 interaction which has 2 degrees of freedom). The interactions between the three level columns are partially confounded with the rest of the three level columns.

L_{27} Orthogonal Array

	Column (Factor/Interaction)												
Trial	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	2	2	2	2	2	2	2	2	2
3	1	1	1	1	3	3	3	3	3	3	3	3	3
4	1	2	2	2	1	1	1	2	2	2	3	3	3
5	1	2	2	2	2	2	2	3	3	3	1	1	1
6	1	2	2	2	3	3	3	1	1	1	2	2	2
7	1	3	3	3	1	1	1	3	3	3	2	2	2
8	1	3	3	3	2	2	2	1	1	1	3	3	3
9	1	3	3	3	3	3	3	2	2	2	1	1	1
10	2	1	2	3	1	2	3	1	2	3	1	2	3
11	2	1	2	3	2	3	1	2	3	1	2	3	1
12	2	1	2	3	3	1	2	3	1	2	3	1	2
13	2	2	3	1	1	2	3	2	3	1	3	1	2
14	2	2	3	1	2	3	1	3	1	2	1	2	3
15	2	2	3	1	3	1	2	1	2	3	2	3	1
16	2	3	1	2	1	2	3	3	1	2	2	3	1
17	2	3	1	2	2	3	1	1	2	3	3	1	2
18	2	3	1	2	3	1	2	2	3	1	1	2	3
19	3	1	3	2	1	3	2	1	3	2	1	3	2
20	3	1	3	2	2	1	3	2	1	3	2	1	3
21	3	1	3	2	3	2	1	3	2	1	3	2	1
22	3	2	1	3	1	3	2	2	1	3	3	2	1
23	3	2	1	3	2	1	3	3	2	1	1	3	2
24	3	2	1	3	3	2	1	1	3	2	2	1	3
25	3	3	2	1	1	3	2	3	2	1	2	1	3
26	3	3	2	1	2	1	3	1	3	2	3	2	1
27	3	3	2	1	3	2	1	2	1	3	1	3	2







11.2 Exercises

Objective: To practice setting up an OFAT (one-factor-at-a-time) experiment.

Instructions: Suppose you are a DNA biologist on a search for the gene that causes male-pattern baldness. Your hypothesis is that one gene is responsible for this debilitating condition in males. You have laboratory facilities that can compare DNA sequences from different people. Design an experiment that will (may!) answer this question. Include in your design considerations such as sample size and how the results will be analyzed (statistically).

Question: How will you implement the *min-max-con* principles?

Time: 20 minutes

Objective: To practice applying the *max-min-con* principles of experimentation.

Instructions:

1. Horatio Hornblower is trying to optimize ***range*** achieved by the new 16 lb. cannon delivered to his ship by the British Admiralty. Brainstorm factors that he could experiment with to increase the cannon's range (*Hint: Develop a Cause and Effect diagram)*.

2. Choose one of the factors that he could use in a *single factor experiment.* Apply the *min-max-con* principles to this and the other factors identified above.

Time: 20 minutes

Objective: To practice applying the *max-min-con* principles of experimentation.

Instructions: In 1747, James Lind (1716–1794) conducted his famous experiment on sailors afflicted with scurvy aboard *H.M.S. Salisbury.*

He took twelve patients, all with similar symptoms. They were kept on the same diet, and all remained in the forehold of the ship. He varied the treatment applied to the sailors as follows: 2 received a quart of cider a day, 2 took 25 gutts of *elixir vitriol* three times a day, two took 2 spoonfuls of vinegar three times a day, two were given a half-pint of seawater per day, two were given 2 oranges and a lemon each day, and the last two were given a complex electuary of garlic, mustard-seed, *rad. raphan*, balsam of Peru and gum myrrh, mixed with barley water "acidulated" with tamarinds (this treatment prepared by the ship's surgeon). He then observed their condition for six days.

Comment on the design of this experiment. What type of study is this? What kind of experiment was run? How were the *min-max-con* principles implemented? At the end of the six days, the two sailors given the citrus fruit were essentially cured of their scurvy. If a report of this experiment were published in a medical journal, do you think today's medical community would embrace the results?

Time: 20 minutes

Objective: To practice enumerating the combinations of experimental factors in a full-factorial design.

Instructions: A team is attempting to improve their protocol for treating patients with tachycardia (abnormal rapidity of heart action). The usual treatments include pressure on one or both carotid sinuses, pressure on the eyeballs, induction of gagging or vomiting, attempted expiration with glottis closed, lying down with feet in the air and bending over. The treatments are sometimes effective when administered singly. Often, though, two or more combinations of treatments are required to slow the patient's heart rate. List the combinations of two and three treatments the team would need to investigate in this process.

Time: 20 minutes

Objective: To practice planning an industrial experiment.

Instructions: A manufacturer is trying to improve the photoresist process for fabricating printed circuit boards. Photoresist is a light-sensitive material used to mask underlying layers from chemical etchants. Possible control factors include developer concentration, spray pressure, temperature and exposure light energy (wavelength). Line width on the circuit board is the critical customer requirement.

1. Selecting the "Y." Consider the following output measures, discuss which one would be "best" for the experiment:
   - σ Process Yield - Percentage of parts meeting the line width specification
   - σ A categorical scale: below spec, within spec, above spec
   - σ Actual line width measurement

2. When you begin to consult with the manufacturing engineers, they state that they are planning to change the control factors one at a time and see what improvements result. What would your response be? How would you convince them to adopt another path?

3. The current process settings produce an average yield of 85%. Discuss your first round of experimentation. Develop one or more possible designs that you could present to the engineers.

Time: 30 minutes

Objective: To practice planning an industrial experiment.

Instructions: You are in charge of developing the cooking instructions for a new microwave popcorn. The CTQs include a) number of unpopped kernels in the bag, b) fluffiness (measured by volume) of popped corn, c) color, d) taste & e) crispness.

Factors you can vary include:

- Oil type
- Oil amount
- Bag venting
- Cooking time
- Microwave setting
- Bag size

Develop both a factorial (possibly fractional) design and a Taguchi orthogonal array design for this scenario. What information is obtained by each experiment (i.e. effects & interactions)?

Time: 30 minutes

Objective: To practice using ANOVA to analyze the result of an experiment.

Instructions:

1. An Information Technology director is trying to improve service at the help desk. The staffing level at the help desk is varied and the average time to answer a call recorded for three days at each staffing level (three repetitions). Perform an ANOVA to determine if staffing level makes a difference. Test at $\alpha = 0.05$.

2. For each of the staffing levels, develop the confidence interval for the average answer time.

Time: 20 minutes

		Staffing Level		
		2	**4**	**6**
Repetition	**1**	30	18	19
	2	19	21	16
	3	28	17	15

Measurements are Average Time to Answer Call in Seconds

Objective: To practice using ANOVA to analyze the results of an experiment.

Instructions:

1. Three methods for removing scale on condenser tubing are being investigated. The methods' effectiveness is measured by the before & after tube cross-section. For example, if a clogged tube had only 5% of its normal area available for cooling flow before the method was used, and 50% of its area available after the treatment, the measurement is 50% - 5% = 45%. The experimental data is shown below. Perform an ANOVA to determine if there is a significant difference in method (at an $\alpha = 0.05$). Develop an interval estimate for the best method.

Time: 25 minutes

		Scale Removal Method		
		A	**B**	**C**
Repetition	**1**	27	15	48
	2	33	18	45
	3	36	24	57
	4	30	21	54

Objective: To practice setting up an experimental design, to contrast the full-factorial and fractional-factorial methods.

Instructions:

1. A team is trying to reduce brazing failures on impellers. They have identified three braze application factors (B_1, B_2 & B_3), two heat treat factors (H_1 & H_2) and three braze material factors (M_1, M_2 & M_3) that are suspected of affecting the failure rate. If each factor has two levels, design an experiment to investigate the effects of these factors. Only two factors, B_2 and M_2 are suspected of interaction. Compare the number of experimental combinations if a *full-factorial* design is employed, vs. a *fractional factorial* or an *orthogonal array.* What's lost when one of the *fractional factorial designs* is used?

Time: 20 minutes

Objective: To practice setting up an experimental design, to contrast the full-factorial and fractional-factorial methods.

Instructions: A shampoo manufacturer is interested in obtaining more uniform fill heights in bottles. The process engineer has identified three factors that are controllable: a) formula viscosity, b) filler operating pressure, and c) line speed (bottles produced per minute). She decides to run a full factorial experiment with two replicates. The response measured is *deviation from target fill height (in eights of an inch).* Analyze the results. Use an alpha = 0.1

Time: 20 minutes

	Operating Pressure			
	25 psi		***50 psi***	
	Line Speed		***Line Speed***	
Viscosity	***200***	***250***	***200***	***250***
100	-3, -1	-1, 0	-1, 0	1, 1
120	0, 1	2, 1	2, 3	6, 5

Objective: To understand interactions in orthogonal arrays.

Instructions:

1. If factor A is assigned to column 4 of an L8 Orthogonal Array and Factor B assigned to column 6, what column will estimate the AxB interaction?

2. Assign factors A, B, C, D & E as well as interactions CxD and CxE to an Orthogonal Array if all factors are using two levels.

3. Assign these factors and interactions to an Orthogonal Array:
   - A, B, C, D, E & F (two levels)
   - AxB, BxC, CxE
   - AxC, BxD, DxE
   - AxD, BxE, BxF

Time: 20 minutes

Objective: To practice designing a real world experiment.

Instructions: A building engineer is trying to reduce the air conditioning system noise in office areas of her building. She has four factors that can be varied, the airflow velocity (high vs. low), the ventilator design (Brand X vs. Brand Z), the use of white noise generators (Yes vs. No) and the air duct surface (regular vs. smooth). Design an experiment that will help her determine the optimum conditions for noise.

Time: 20 minutes

Objective: To practice analyzing the results of a real world experiment.

Instructions:

1. An orthogonal array was used to conduct an experiment. Five factors were included in the experiment (D, AS, E2, C, SS).

2. "Reverse engineer" the experiment – was the orthogonal array correctly employed (i.e. can the third column be used to detect the interaction of D and AS?)? If all factors are significant, can an ANOVA be performed?

3. Analyze the experiment. Which factors are significant? If the experiment was designed to *maximize* the response, which factor levels are best? Are the residuals "OK?"

Time: 20 minutes

Column		1	2	3	4	5	6	7	
StdOrder	**RunOrder**	**D**	**AS**	**D* AS**	**E2**	**C**	**SS**		**Response**
13	1	1	1	1	1	1	1	1	10
7	2	1	1	1	2	2	2	2	26
1	3	1	2	2	1	1	2	2	57
14	4	1	2	2	2	2	1	1	14
3	5	2	1	2	1	2	1	2	0
11	6	2	1	2	2	1	2	1	16
16	7	2	2	1	1	2	2	1	38
8	8	2	2	1	2	1	1	2	22
12	9	1	1	1	1	1	1	1	10
15	10	1	1	1	2	2	2	2	28
9	11	1	2	2	1	1	2	2	64
6	12	1	2	2	2	2	1	1	39
10	13	2	1	2	1	2	1	2	0
4	14	2	1	2	2	1	2	1	17
2	15	2	2	1	1	2	2	1	36
5	16	2	2	1	2	1	1	2	23

Objective: To practice analyzing the results of a real world experiment.

Instructions:

1. A full-factorial array was used to conduct an experiment on a brazing operation. Four factors were included in the experiment (Tube Cut, Skill, Flux and Cleaned). The output is fraction of rejected brazes from the process.

2. Analyze the experiment. Which factors are significant? If the experiment was designed to *minimize* the response, which factor levels are best? Are the residuals "OK?"

Time: 20 minutes

StdOrder	Tube Cut	Skill	Flux	Cleaned	Results
1	-1	-1	-1	-1	0.215
2	1	-1	-1	-1	0.269
3	-1	1	-1	-1	0.184
4	1	1	-1	-1	0.258
5	-1	-1	1	-1	0.146
6	1	-1	1	-1	0.344
7	-1	1	1	-1	0.107
8	1	1	1	-1	0.330
9	-1	-1	-1	1	0.200
10	1	-1	-1	1	0.430
11	-1	1	-1	1	0.123
12	1	1	-1	1	0.209
13	-1	-1	1	1	0.121
14	1	-1	1	1	0.311
15	-1	1	1	1	0.200
16	1	1	1	1	0.280

Objective: To practice analyzing the results of a real world experiment.

Instructions:

1. A full-factorial array was used to conduct an experiment on a compressor design. Three factors were included in the experiment (Eductor, Pump and O-Ring). Two outputs were measured: Temperature and Pressure.

2. Analyze the experiment. Which factors are significant? If the experiment was designed to *minimize* the temperature response, which factor levels are best? If the experiment was designed to *maximize* the pressure response, which factor levels are best? Are the residuals "OK?" Are there any conflicts in factor levels relative to the two responses?

Time: 20 minutes

RunOrder	Eductor	Pump	O-ring	Temp	Pressure
1	1	-1	-1	133.30	40.89
2	1	-1	1	142.00	37.66
3	-1	1	-1	135.15	67.50
4	-1	-1	1	134.20	51.20
5	-1	-1	-1	134.15	54.47
6	1	1	-1	133.50	49.23
7	-1	1	1	135.00	59.04
8	1	1	1	143.00	44.74

Objective: To practice analyzing the results of an experiment using contrasts

Instructions:

3. Two factors, Primer Viscosity and Coating Viscosity were varied to determine their effect on Paint Dry Time (minutes x 10). Determine the specified Contrasts (Primer 20 v 30, Primer 20 v 25) as shown below:

Time: 20 minutes

Treatment	**1**	**2**	**3**	**4**	**5**	
Primer Viscosity	20	20	25	30	30	
Coating Viscosity	15	25	15	25	15	
Averages	252.5	320	240	307.5	262.5	
Contrasts						Sum
Primer 20 v 30	-1	-1	0	1	1	0
Primer 20 v 25	-1	-1	2	0	0	0
Primer 25 v 30	0	0	-2	1	1	0
Coating 15 v 25	-2	3	-2	3	-2	0

12.0 Changing the Process

Unit	Description	Page
12.1	Selecting & Implementing Countermeasures	12.1-1
12.2	Financial Analysis of Changes	12.2-1
12.3	Exercises	12.3-1

In this section, we close the loop on the improvement cycle. The preceding sections have presented a wide variety of methods to study your processes and understand the variables that drive performance. Here, we'll discuss making changes to the process that are both *effective* and *efficient.*

12.1 Selecting & Implementing Countermeasures

Learning Objectives

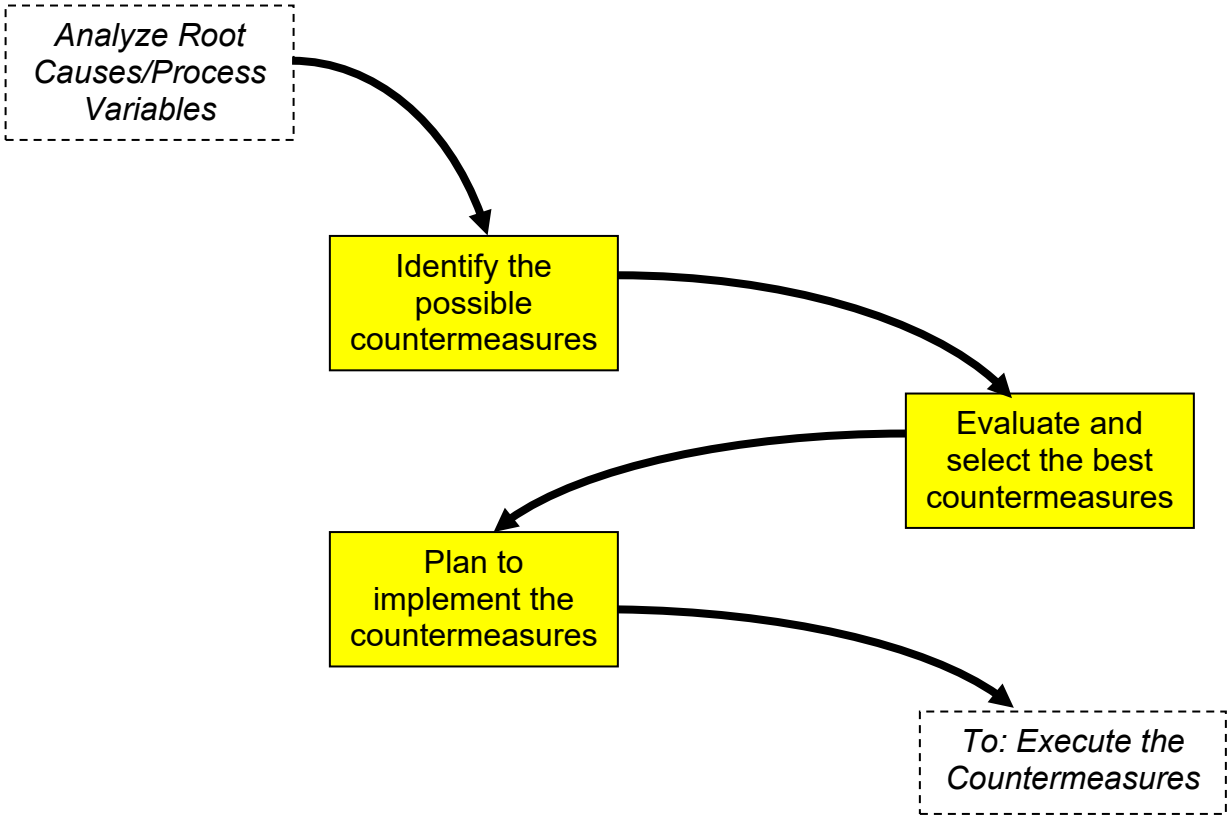
- Be able to select cost-effective process changes
- Be able to plan and implement process changes
- Be able to track results of changes

Unit Contents

- Identify/Select Countermeasures
- Execute Countermeasures/Evaluate Results
- Standardize/Replicate Improvements

12.1.1 Identify/Select Countermeasures

Road Map



12.1.2 Generating Countermeasures

Types of Countermeasures:

Attack the causes not the symptoms - Just in case we didn't mention it, make sure that you have identified causes and not symptoms. You can buy a remedy for your cold that attacks the symptoms of your runny nose, watery eyes, sore throat, etc. But the cold remedy does nothing to prevent the reoccurrence of the root cause of the problem, the cold virus.

Change the process to make a long-lasting improvement - The Countermeasures then, should be designed to prevent the reoccurrence of the problem you're addressing. They should focus on changing the production process that makes your products or services. This is quality improvement.

Consider temporary and long-term countermeasures - If the long-term solution to the root cause will take some time to implement (because of budget or other constraints) consider what might be temporary countermeasures to prevent the problem from occurring until the longer-term solution can be implemented.

Don't reinvent the wheel - Check around (with vendors, literature searches or other organizations) to see what has been done in other cases where the problem has occurred. There may already be proven, cost-effective solutions to your problem. Benchmarking methodologies can help here.

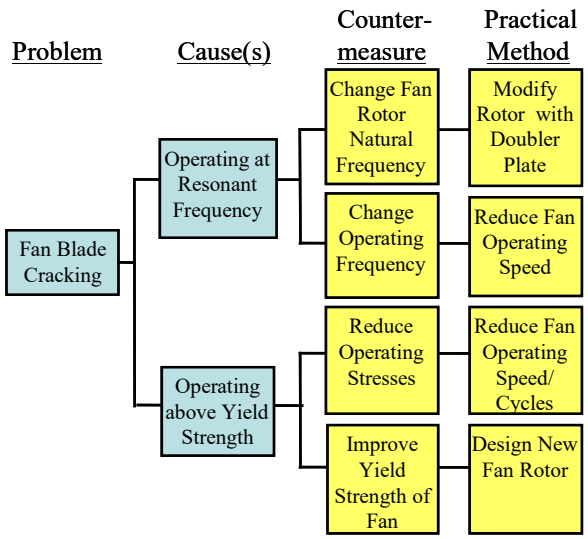
Phase in expensive countermeasures - Consider implementing in one location or product line first. If it is successful there the budget can usually be found.

Countermeasures for human-related causes - When a Human-Related Cause is identified, there is a tendency to select an "INSPECTION-type" or training countermeasure. For example, if frequent mistakes are identified in timekeeping or payroll, then a check (or double-check) of the reports may be considered. This adds cost and inspection never catches all the mistakes. Retraining may be needed but its effect will probably be short lived. Think of ways to make the improvement permanent. Consider ways to *error-proof* the process

Be creative - "If you always do what you've always done, you'll always get what you've always got. "

Selecting Countermeasures - The Countermeasure Matrix

The Countermeasures Matrix summarizes the problem, the causes and those countermeasures considered by the improvement project. It helps prioritize the countermeasures in terms of their costs, benefits, and feasibility and can be a simple graphic to display the recommendations of the improvement team.



	Evaluation			
Benefit	Cost	Feasible	Score	Action
8	3	7	168	Yes
2	5	4	40	No
2	4	4	32	No
6	3	8	144	Yes

This example summarizes recommendations to prevent the reoccurrence of a cracking problem in power plant boiler fan blades. Two causes were identified; each cause had two possible countermeasures. One countermeasure for each cause was evaluated as being the most cost-beneficial over the remaining life of the unit. Both countermeasures were combined into the new fan rotor design.

Notes on the Countermeasure Matrix

- By laying out the Problem, Cause, Countermeasure and Practical Method on the matrix, it's easy to see if there is a logical "linkage" among the issues.

- The Countermeasure is the "WHAT," the Practical Method is the "HOW."

- The Evaluation boxes are usually given scores based on the benefit of the countermeasure, its cost and its feasibility to implement (i.e. the practicality of implementing this countermeasure).

- For simple problems, the scores may be obtained by general agreement of the project team, for complex or expensive countermeasures, more detailed cost/benefit analysis is performed:

FACTOR	RANGE	
BENEFIT	1 - Low Benefit	10 - High Benefit
COST	1 - High Cost	10 - Low Cost
FEASIBILITY	1 - Not Practical to Implement	10 - Easy to Implement

- The individual scores are multiplied to get the overall score. Note how the COST factor is scored the reverse of the BENEFIT and FEASIBILITY factors (High Cost - Low Score).

- The Action box indicates the decision to adopt the countermeasure (YES) or not (NO).

Cost/Benefit Analysis

Purpose

Generally, there is more than one way to address the causes of your problem. The potential countermeasures you've identified should be evaluated for both their costs and benefits. Dr. Juran talks about the different "languages" spoken in organizations. Staff understands the language of technology, Senior Management understands the language of finance, and Middle Management must be "bilingual." Since Middle and Senior Management approval will often be required for process changes, the costs and benefits should be expressed in dollar amounts, if possible.

Quantifying Benefits

Changes that reduce or eliminate defects - Here, the challenge is to estimate the cost of the defect and to estimate the reduction in defects that will be achieved by your countermeasure(s). These costs are often slippery and can be inflated or deflated depending on the assumptions you make. Dr. Deming often said that the "most important figures are unknown and unknowable." Studies of government projects show that the costs are *extremely* under-estimated (California's high-speed rail project, repairing the Francis Scott Key bridge, Obamacare, the list goes on forever!).

> Example: The "Cost" of a rejected impeller is estimated at a few thousand dollars. The impact on revenue could be considerably more than that if the time spent repairing the impeller prevents the production of another unit.

Costs associated with failures should include, for instance, not only the part and labor costs, but also the costs of lost or replacement production due to the equipment downtime. For instance, when a chiller is delayed for re-testing, this impacts production. You may already have some clues as to the reduction in defects your countermeasure will achieve. For instance, a Pareto Chart of causes will provide you with the current frequency of defects by cause. If your countermeasure will eliminate one of the bars, then this becomes your reduction estimate.

Changes that reduce machine, material or labor costs - Here, the challenge generally is to estimate how much material, machine time or labor will be saved by the countermeasure, how often the resource is demanded, and how much the resource costs on a per hour or per demand basis. This can be multiplied by the expected demand over a year to identify the yearly savings:

Example: A team estimated that their process change would reduce the number of trips to get additional supplies by 10 per day. Each trip took 1 hour and was performed by a fitter earning $16.00 per hour. Estimated savings are then:

10 trips/day x 1 hr./trip x $16.00/hr = $160.00 per day

This was extrapolated over the year:

$160.00/day x 5 days/week x 52 weeks/year = $41,600.00 per year.

Of course, the natural "business" question is, "Does the plant now require one fewer full-time-employee, or will the fitter be assigned to other duties?"

Changes that improve customer satisfaction - These are the most difficult to quantify monetarily, but are usually the most beneficial changes. The immediate benefits of improved satisfaction are generally intangible, but these can translate into long-term repeat business or referrals from customers who "brag" about your organization. Many companies are developing "balanced scorecards," where customer satisfaction (as measured by a periodic survey) is one indicator tracked and targeted for improvement. Some organizations have also tracked customer satisfaction as a "negative" indicator, the number of complaints, and actively attempted to reduce complaints. The benefits of changes to improve satisfaction may be measured here in terms of expected reductions in complaints or improvements in satisfaction indices.

Quantifying Costs

If the change requires increased expenditure on materials, equipment or labor hours, the cost can be quantified just like the counterpart benefits. For purchases of capital equipment, though, the cost calculation is more complicated, since it will include the initial purchase price, setup and training costs, operating and maintenance costs. In addition, depreciation and salvage value of capital equipment usually enters the cost equation.

Each organization's accounting department has access to the cost equations that can be applied. For organizations wishing to include these figures in the cost/benefit evaluation, we recommend that a simple worksheet be prepared to help improvement project teams estimate these costs.

Selling the Countermeasures

After the proposed countermeasures have been selected, the job of "selling" them to management and co-workers begins (actually, it is a good idea to keep management and your co-workers involved during the selection process). The best way to "sell" countermeasures and gain needed support and approval is by explaining the facts and data that support your decisions.

Organize your analysis using the DMAIEC process improvement steps. Summarize the results of each step and be prepared to back up the summary with the more detailed work performed. This is a good time to check your entire analysis. Does it make sense? Is there "linkage" between the steps - have the causes been verified, do the countermeasures address the most probable causes, are the countermeasures the most cost-beneficial?

Explain the work to your co-workers (especially those who will be involved in implementing the countermeasures and who may have to change their work practices) and obtain any final input from them. Request a time for management to review the project work and recommendations. If you have a storybook of your work, send copies to the reviewers about one week in advance of the meeting.

During the review, present a summary of the project work and your recommendations. Answer questions that may arise. Before the meeting ends, ask for a management decision on your recommendations, or set a mutually agreed upon date when the decision will be communicated to you and your team. *Do your best to "close the sale."*



Implementation Planning

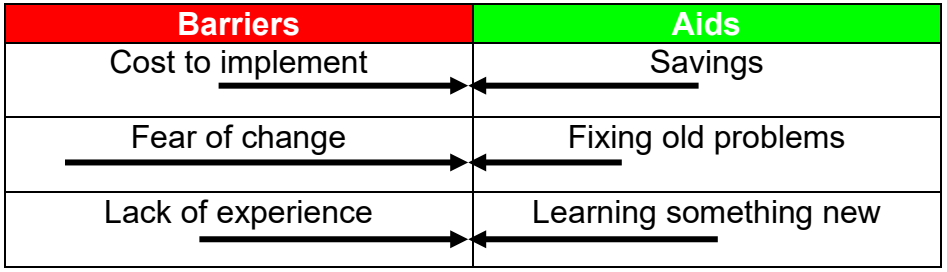
Now that the countermeasures have been determined and approved, they must be planned and implemented. An Action Plan may be helpful here. As part of your planning, consider what procedures or instructions must be revised, what material or tools must be ordered (or changed), what training must occur in the new method. In the example below, notice that the team planned both the changes to the process *and* how they intend to measure the *effectiveness of the changes.* They are thinking "*PLAN-DO-CHECK-ACT."* As the tasks are implemented, keep track of any problems or deviations from the Action Plan. Determine whether these problems are going to impact either the effectiveness or cost of the countermeasures.

ACTION PLAN

***Project:* Fan Rotor Redesign – KL Air Handler**			
What	***How***	***Who***	***Date***
Write technical specs for Fan	Develop specifications for fan which address problems found through cause & effect analysis	SL	9/XX
Evaluate current manufacturing processes	Visit several fan manufacturers to learn about fan design and evaluate quality control systems	JLB, RPF	6/XX
Develop Spec Package for Bids	Finalize specifications	TGK	10/XX
Select Vendor	Evaluate Bids for Compliance with specifications	ALL	12/XX
Ensure Vendor complies with specifications	Conduct design review at completion of detailed design Manufacturing engineering to qualify fabrication process	SL	TBD
Establish Schedule for Test Fan Installation	Work with vendor to formulate project schedule, Scheduling Department to determine next available air handler	SL, RPF	TBD
Install Test Fans on one unit for evaluation	Fabricate and ship fan to plant, train workers in installation procedure, install fan	RPF	TBD
Test Fan for compliance with specs	Gather fan performance data and evaluate	SL	TBD
Purchase Fans	Review fan performance evaluation and issue purchase order	ALL	TBD
Integrate into Production Process	Train workers in installation procedure, monitor initial installations	ALL	TBD
Monitor Performance and Warranty Cost	Gather/analyze fan warranty data – perform root causes of warranty claims	SL	TBD

Barriers and Aids table

One factor often overlooked is acceptance. A technically good countermeasure can fail if the people who must implement it do not accept the idea. A barriers and aids diagram can help identify acceptance factors that will need to be addressed in the plan.



Application:

This tool is often used with a team when planning the implementation of countermeasures that may not be an easy sell. It helps identify factors that can be barriers or aids to implementation.

Construction:

1. Gather people who know the issues that will impact implementation.

2. Explain the countermeasures that are being planned.

3. Draw a line down the middle of a flip chart and label the left side, aids and the right-side barriers.

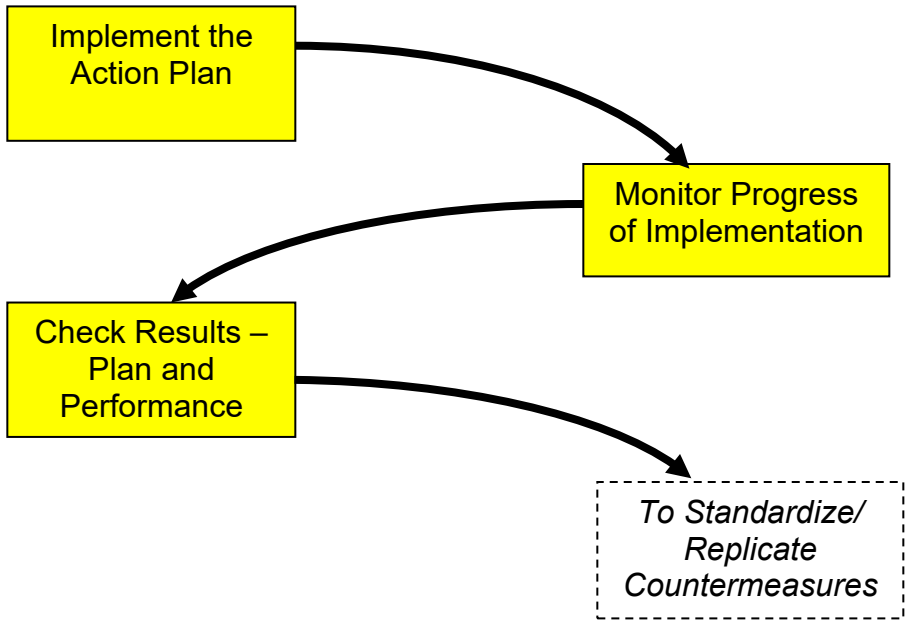
4. Have the group brainstorm barriers and aids. Try to line up barriers and aids that counter each other.

5. Discuss the strength of each barrier and aid with the group and display the relative strength by the length of an arrow pointing to the centerline.

6. Discuss actions to include in the plan to utilize the aids and overcome the barriers.

12.1.3 Execute Countermeasures/Evaluate Results

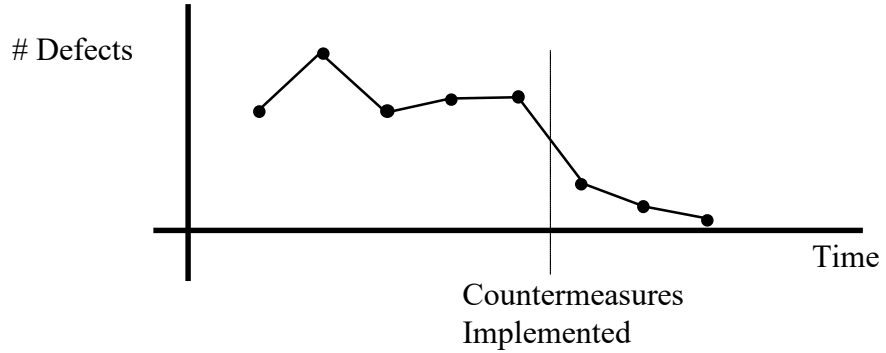
Road Map



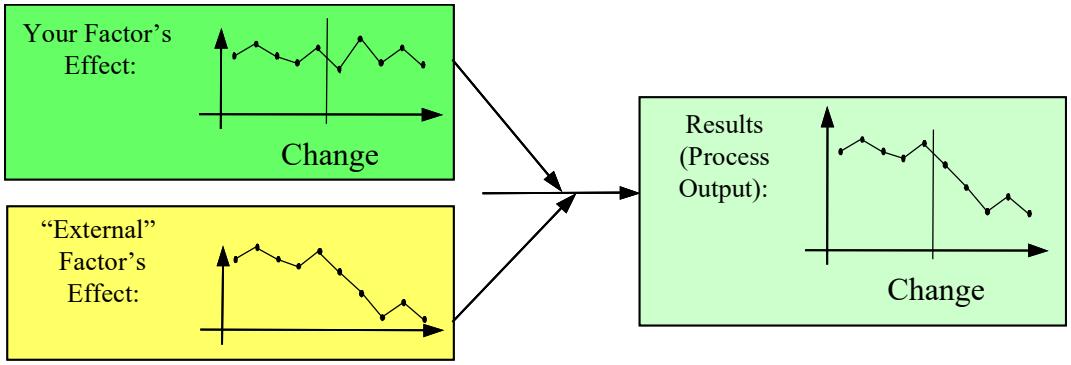
Tracking Results

Results and Effects

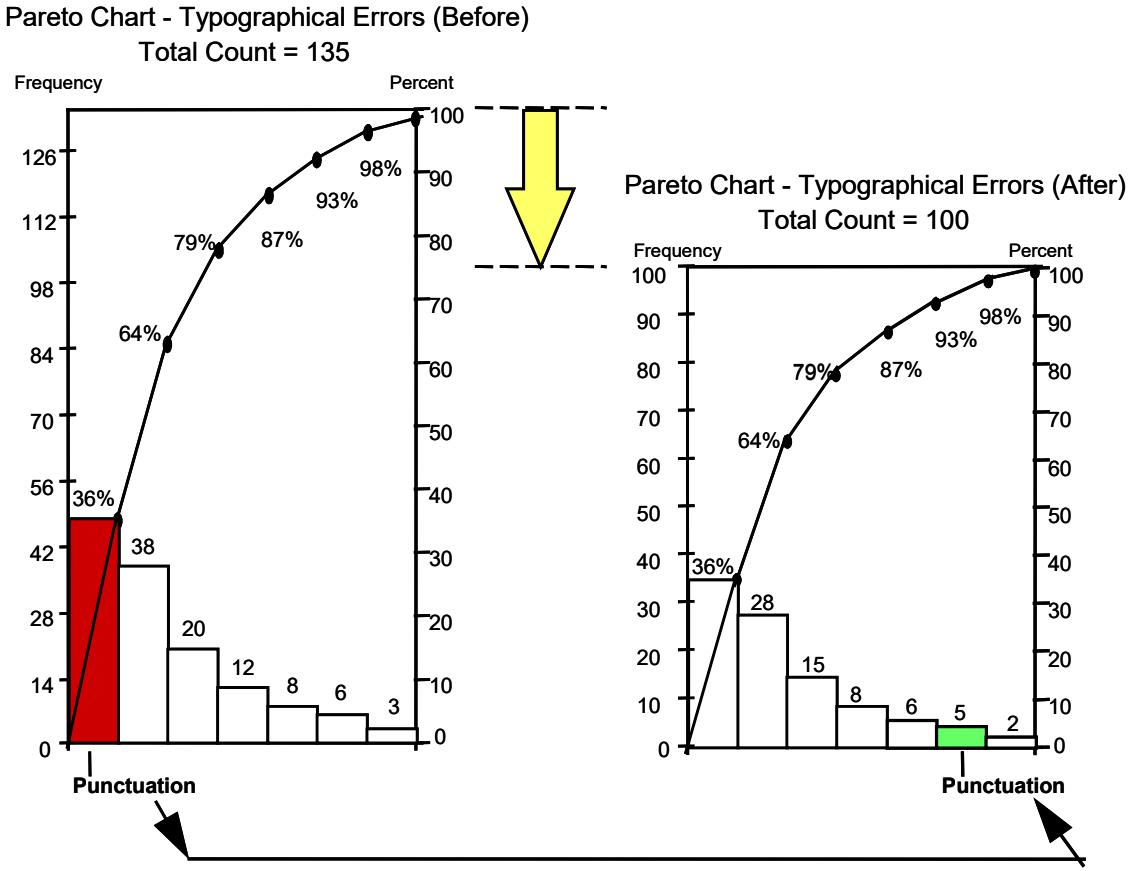
After the countermeasures are implemented, the problem should either decrease significantly in frequency or be eliminated:



The indicator you identified at the beginning of the project should be used to examine the results of your changes. Pareto analysis (see *Pareto - Before and After,* Unit 7.2) should be applied to show that the factor you addressed has, in fact, been decreased in frequency or eliminated. There is a difference between *results* and *effects. Results* are what you observe to happen after the changes are made. But you are interested in the *effects* of the changes *you've* made to the process. It's possible that other factors, either *internal* or *external* to your process, could have achieved the *results:*



This is why a *before and after Pareto* analysis is useful. You see both the change in the *output* and the change in the *process factor.* Here's a Pareto showing how both the total number of errors decreased (output) *and* that the factor that was addressed (punctuation) has moved to the "Trivial Many:"



Be careful of "side effects" that occur because of your change. One of our friends was discussing the issue of *Cesarean Section (C-Section) births* with an obstetrician. Insurance companies don't like C-Sections because they tend to cost more than natural deliveries. The obstetrician was curious to know whether his practice should try to reduce the number of C-Sections they performed. Our friend replied that, yes, that would be a good idea, but make sure that the other

important indicators such as *Complications, Infant & Maternal Mortality* were not adversely affected (i.e. "*Big-Quality"* is considered, see *Unit 4.1*).

We've got to keep the big picture in mind, even while we're focused on one aspect of improvement. Dr. Martin Luther King, Jr. said:

> *"All progress is precarious, and the solution of one problem brings us face to face with another problem."*

Quality Improvement "Rework"

In some cases, the countermeasures applied do not have the desired effect of eliminating the problems. Suggestions to address this issue generally include a review of the improvement analysis steps, although in reverse order:

Review the new problems - Are these the same type of problems experienced in the past? If so, why aren't the countermeasures preventing these problems? If not, do the countermeasures have any relationship (cause and effect) to the new problems?

Revisit the countermeasures - Were all the countermeasures implemented? Is everybody involved following the new process? Do the countermeasures selected address the identified causes of the problems? If not, these may need to be re-examined. If the countermeasures turn out to be faulty, develop an action plan to "un-implement" them.

Review the cause/effect analysis - Were the real causes identified? Use the new problem information to help answer this question.

External factors - Has something else changed in the process? This could be either internal or external to your organization.

12.1.4 Standardize/Replicate Improvements

Standardization

If the countermeasures are effective, then the final action is to make sure that they are incorporated into the daily work processes. Here, the "hand-off" occurs between *quality improvement* and *quality control.* How does your organization *control* or *manage* the quality of daily operations? Written instructions may need to be revised. If you make use of flowcharts to describe work processes, then these should be revised to show the new process. Make sure that training and education programs in your organization reflect the new work process.

Examples:

For a large US railroad, changes that improve the reliability of locomotives are captured in the following "standards:"

- *M-Sheets (M01 - 6, M60, M90 worksheets) - These define the scheduled maintenance program for the locomotives.*
- *Diesel Maintenance Procedures (DMP's) - These define how maintenance activities are conducted, by type of maintenance (e.g. engine removal and overhaul, power head reconditioning, etc.).*
- *Maintenance Instructions - These are vendor-supplied maintenance procedures for specific locomotive equipment.*

For a large US hospital, changes that improve quality or cost of patient care are captured in the following "standards:"

- *Clinical Pathways - These define the care process for patients, by type of diagnosis.*
- *Clinical Protocols - These define how specific patient care activities were performed (dialysis treatment, respirator management, etc.).*
- *Preference Cards - These were surgeon-supplied materials and equipment lists for specific surgery procedures.*

OK, we apologize! We're not implying that locomotive maintenance and human "maintenance" have many of the same characteristics. It's just that we get tired of hearing how every industry is *different!*[1]

[1]Of course, we've seen a lot of consultants make a lot of money by feeding on this idea. "Oh, yes, you need a quality improvement course just for *your* company. That manufacturing stuff just won't work in *your industry!*" Yeah, Right!! By the way, there's this bridge in Brooklyn that's for sale!

Replication

Often, countermeasures found to be successful in one area of an organization can be applied elsewhere. Each organization should consider how to address the *replication* issue. This is another reason the *Quality Improvement Story Summary* is useful. It can be placed in a "Quality Library," for review by other improvement efforts.

A database of quality improvement projects, indexed by keywords, is easy to set up today. With today's computer systems, the full text/graphic version of the story can be placed "on-line" for review by anybody in the organization. How many times has the same problem been "solved" in your organization? Let's see if we can't prevent this.

Replication outside your organization should also be considered. There may be competitive issues to address, but a prime method of increasing "general" technology is the system of standards (ISO, IEEE, etc.) that are maintained by various professional organizations. Submit a copy of your improvement to the appropriate standards committee if you think there is a general replication potential.

> *Example: The Nuclear Regulatory Commission prepares Information Notices and Bulletins that distribute information regarding issues and events that have occurred in the nuclear industry that may be applicable to all or some subset of operating reactors. Failures, operating events, and quality assurance system issues are the subject of these documents.*

12.2 Financial Analysis of Changes

Learning Objectives

- Be Able to Identify Cost Factors Associated With Changes
- Be Able to Predict Benefits of Changes
- Be Able to Develop an Economic Evaluation of a Change

Unit Contents

- Financial Analysis
- Economic Evaluation Process

12.2.1 Financial & Economic Analysis

Fundamental Goal

Teams often need a more sophisticated method of calculating the costs and benefits of a proposed change. For example, if the team's proposal is to close existing sites and move them to a lower labor cost area, management and other stakeholders will likely require a detailed cost evaluation. Regardless of the type of change, there are some basic goals associated with the change. The expected returns from the investment in the new product/service must exceed its costs. In economic terms, the *marginal* benefits must exceed the *marginal* costs. The following material lays out a general approach to support the financial/economic analysis of a change.

Objective of the Economic Evaluation

Your basic objective is to develop an estimate of the future "gozzinta's" and "goesoutta's." To formalize this, you will try to forecast the **cash flows** associated with the investment at the time they will be received (*cash benefits*) or disbursed *(cash payments*).

Sources of Cash Benefits:

- Cash received because of sales,
- Cost reductions (when a more efficient process replaces a less efficient one, or from tax benefits),
- Cash received when replaced equipment is sold, and
- Cash received from salvage value of equipment, plant or facility at end-of-life.

Sources of Cash Payments:

- Initial cost of the investment (capital cost),
- Capital improvements made during the life of the equipment/plant/facility,
- Operating costs, such as wages, materials, energy, taxes and maintenance.

12.2.2 Financial & Economic Evaluation Process

Cash Flows Example (Project with 5 Year Life)

Most changes have a certain economic "life." In addition, the change has a marginal impact on the company – that is, it will make a certain difference to either the company's costs or revenue. One solid method of evaluating one or more proposals is cash flow projection. To do this, you will itemize the benefits and costs of each change. Then you will estimate the marginal effect the investment will have on the firm (a new product/service may be evaluated alone, or in comparison to other related, independent or mutually exclusive products/services). In the example below, you can see that this proposal is projected to increase sales and decrease operating expenses. Tax, depreciation and other factors, though, will also affect the *net* cash flow associated with the proposal:

	Time Period (Years)					
	0	**1**	**2**	**3**	**4**	**5**
Income Statement Changes:						
Sales (A)	0	$7500	$8000	$8000	$7500	$7500
Operating Expenses (B)	(416)	(6972)	(6972)	(6972)	(6972)	(6972)
Depreciation (C)	0	(560)	(336)	(202)	(151)	(151)
Pretax Profit (D = A+ B + C)	(416)	(32)	692	826	377	377
Taxes (F = D x E) (E =tax rate = 0.34)	141	11	235	281	128	128
Profit after Taxes (G = D + F)	(275)	(21)	457	545	249	249
Noncash Charges:						
Depreciation (-C)	0	560	336	202	151	151
Capital Investments:						
Property, facility, equipment (H)	(1400)	0	0	0	0	0
Working Capital Changes (I)	(940)	0	0	0	0	900
Residual Net Cash Flow (J = G + H + I)	$(2615)	$539	$793	$747	$400	$1300

***Simple Methods of Valuing Different Proposals*:**

Fortunately (or unfortunately, depending on your perspective), accountants have developed a number of methods to evaluate proposals. Five that are commonly used are presented below:

Benefit Cost Ratio: Calculate the Ratio of Net Benefits to the Investment:

Benefit Cost Ratio = ($539 + 793 + 747 + 400 + 1300)/$2615
= $3779/2615 = 1.45

Payback Period: Calculate the time required for the annual net benefits exceed the initial investment cost:

Payback Period = Investment/Average Yearly Net Benefit
= $2615/756 = 3.5 Years

Annual Worth: Annual worth represents a uniform annual series of money for a certain period of time that is equivalent in amount to a particular schedule of cash benefits and/or disbursements under consideration (if only disbursements are considered, this term is referred to as annual cost). If the net annual worth is greater than zero, then the project can be expected to earn more than the minimum attractive interest rate, the interest that must be paid on invested capital

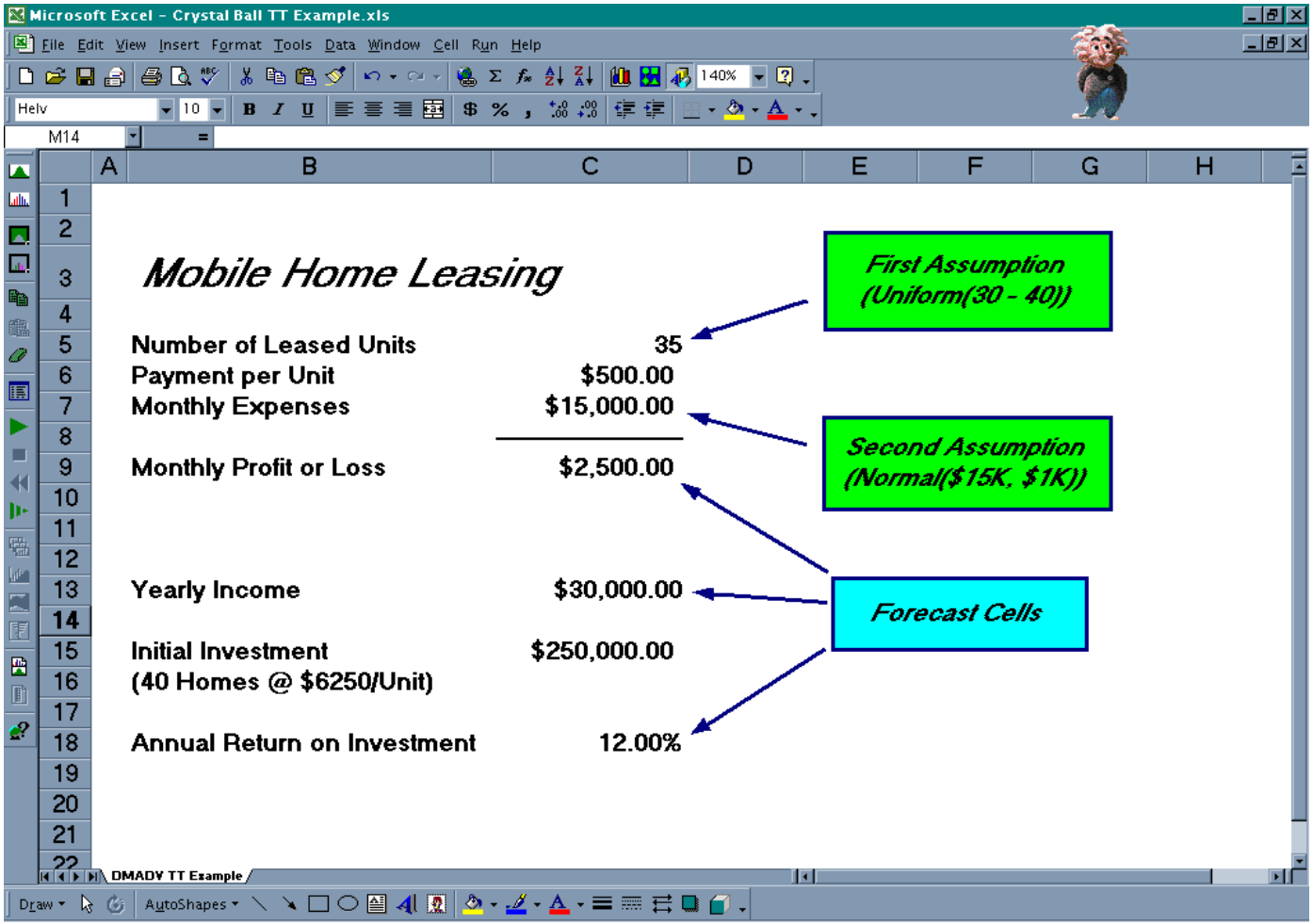
Present Worth: Present worth represents an amount at some beginning or base time that is equivalent to a particular schedule of cash benefits and/or disbursements under consideration. If the net present worth is greater than zero, then the project is economically desirable (using the same criteria as annual worth).

Rate of Return: This method calculates the rate of return on the investment and compares it with the minimum attractive rate of return. The rate of return for a single project involves finding the interest rate at which the present worth of the cash benefits (receipts or savings) equals the present worth of the cash disbursements.

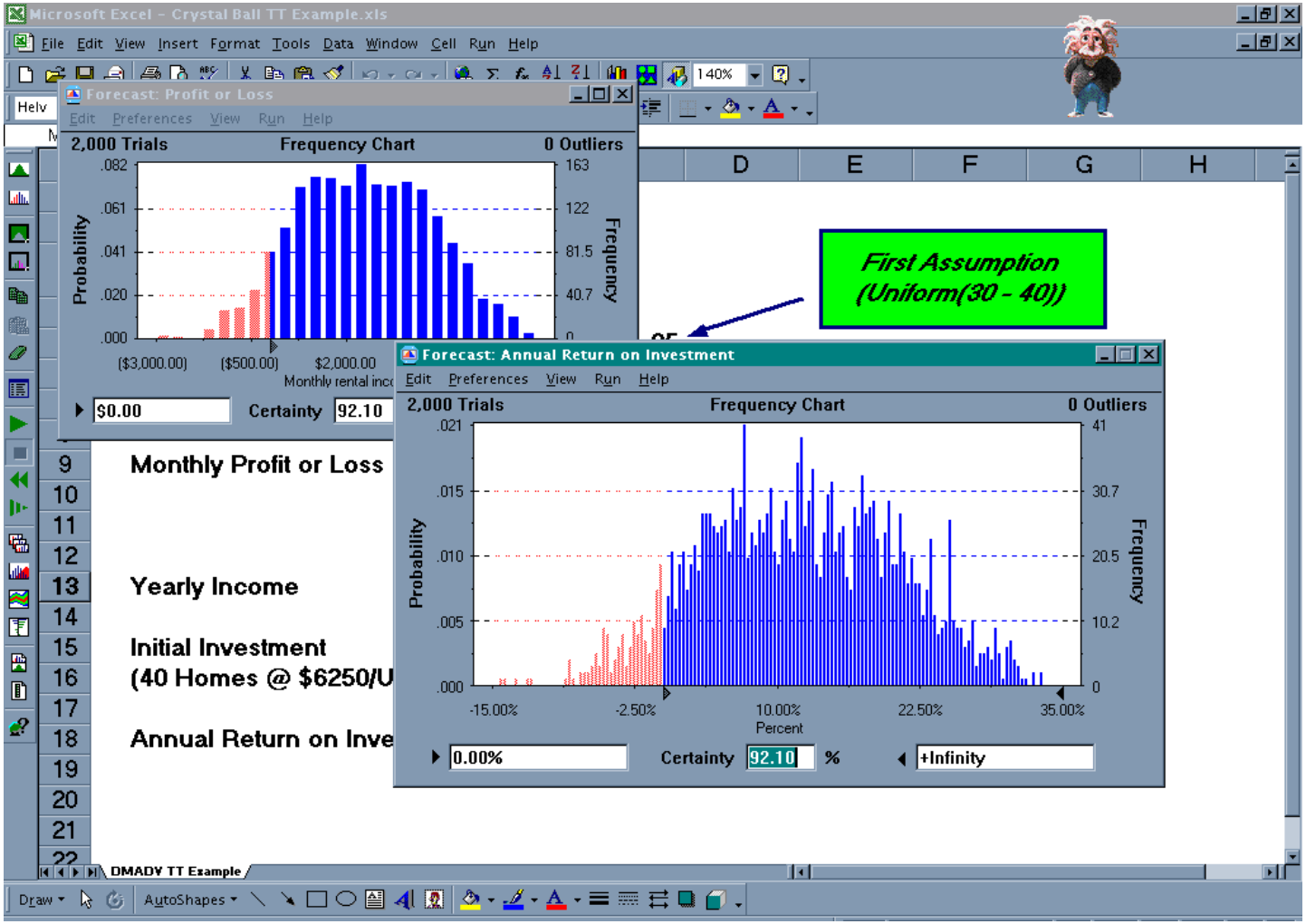
Uncertainty and Risk Assessment

Many economic evaluations treat the cash benefits and disbursements as *point estimates.* However, there is often a great deal of uncertainty about the actual value of these estimates. As Yogi Berra noted: *"Predicting the future is often an uncertain business."* "What-if" calculations can be performed when the analysis has been performed on a spreadsheet.

Monte Carlo simulation techniques may also be applied. Here, the key variables are modeled with probability distributions. The uncertainty in the variables is then propagated to the outcome through the Monte Carlo process. Crystal Ball is a software package that allows the user to easily perform simulations in an Excel environment. In this example, a mobile home leasing proposal is being evaluated. Two of the variables have significant uncertainty, the *number of leased units* and the *monthly expenses*. These are modeled with probability distributions as shown on the following screen shot:



The Monte Carlo process produces a distribution of the expected outcome (here *Yearly Income* and *Annual Return on Investment* are of main concern by repeatedly sampling from the probability distributions associated with the variables and then calculating the Income & ROI outcomes. Here we see the distribution of the two outcomes of interest and that the model predicts over a 90% chance of breaking even on the investment:



***Cost Factors for Economic Analysis* – Here is a "laundry list" of factors to consider in your economic analysis:**

Direct	Indirect	Indeterminate	Intangible
Fixed: *Depreciation* *Investment Interest* *Taxes* *Insurance* *Supervisory Personnel* *Clerical Help* *Maintenance Personnel* *Other*	***Equipment/Method:*** *Space Occupied* *Effect on Taxes* *Effect on Inventory Value* *Value of Repair Parts* *Demurrage Costs* *Downtime Charges* *Production Rate Changes*	***Equipment/Method:*** *Space lost/gained* *Changes in overhead* *Inventory control savings* *Inventory taking savings* *Production control savings* *Changes in product or material quality* *Life of job using equipment* *Reduction in physical effort*	***Equipment/Method:*** *Quality of equipment* *Durability of equipment* *Compatibility of equipment* *Standardization of equipment* *Flexibility* *Adaptability* *Complexity* *Safety* *Obsolescence Rate* *Manufacturer's reputation* *Equipment Availability* *Post-sale advice/service* *Service availability* *Repair parts availability* *Service quality*
Variable: *Operating personnel* *Material/Supplies* *Fuel, power* *Lubrication* *Maintenance, parts and supplies* *Maintenance, labor*	***Management:*** *Travel expenses incurred in investigation* *Cost of follow-up* *Relay out costs* *Training* *Overtime req'd to make up for lost production* *Volume of work-in-progress* *Charges to operation after full depreciation* *Handling returned goods*	***Management:*** *Lost production due to delay in installation* *Percentage of time equipment will be utilized* *Additional labor required for increased capacity* *Turnover of work in progress* *Changes in line balance* *Business volume trends* *Equipment costs trends* *Improved workflow* *Ease of supervision* *Reduction in paperwork*	***Management:*** *Financial policy* *Economic survival goals* *Effect of future changes* *Plans for expansion* *Labor relations aspects* *Effect on morale* *Increased salability of product* *Improved customer service* *Pride in installation*

Definition of Cost Factor Categories – Just in case we forgot any in the table above, here is what we are trying to capture in the different "buckets:"

Bucket	***Definition***
Direct	Associated with labor, materials used directly in production.
Indirect	Associated with labor, materials which cannot be tied directly to a product/service.
Indeterminate	Factors that can be reduced to a monetary value but are often omitted because of their seemingly vague relationship to costs.
Intangible	Factors that do not lend themselves to a cost quantification but can be important in determining the worth of a project.

Evaluating Costs

To evaluate the different cost factor categories, prepare tables or spreadsheets such as appear below:

Direct Costs - Estimate the dollar value of each factor:

Fixed Costs:	**Dollar Cost**
Depreciation	
Investment Interest	
Taxes	
Insurance	
Supervisory Personnel	
Clerical Help	
Maintenance Personnel	
Other	
Total:	

Variable Costs – Estimate the factors that contribute to the variable costs, the frequency of the variable cost and then develop the subtotals for each factor:

Variable Costs:	**Per Unit Cost**	**# Units/ Time**	**Subtotal**
Operating personnel			
Material/Supplies			
Fuel, power			
Lubrication			
Maintenance, parts and supplies			
Maintenance, labor			
Total:			

Indirect/Indeterminate Costs - Estimate the dollar value of each factor. See the directions below:

	Probable Effect on Analysis				Estimated Dollar Value of Factor		
Factor	Adv. (-)	Disadv. (+)	Cost Basis	Evaluation Basis	Adv. (-)	Disadv. (+)	Notes
Subtotals:							
Total Evaluation of Cost Factors:							

Probable Effect on Analysis: Check whether this cost factor will have a *negative* impact on cost (i.e. a cost savings) or a *positive* impact on cost. Make sure each factor is evaluated appropriately; it is easy to reverse the cost impact of a particular factor.

Cost Basis: Some factors' cost can be evaluated in terms of a cost/unit, e.g. dollars/unit, dollars/hr. or dollars/sq. ft. Estimate these costs here.

Evaluation Basis: Estimate the number of units associated with the cost. For factors that are not evaluated on a cost/unit basis, note the basis for the cost here.

Estimated Dollar Value of Factor: Estimate the dollar value of the cost and identify it as a negative or positive value.

Intangible Costs – When the factors are difficult to evaluate quantitatively, a "pseudo-quantitative" evaluation is still possible. The Pugh Matrix approach may be employed here (see Unit 14.2 for more information about the Pugh approach):

			Alternatives			
			1		2	
Factor	Importance	Adjusted Importance	Rating	Weighted Rating	Rating	Weighted Rating
Total						

Importance: Relative importance of each factor, the most important is given a rating of 100, less important factors are given lower values.

Adjusted Importance (Optional): Total the importance ratings; divide the individual ratings by the total.

Rating: How well each alternative satisfies the intangible factor (completely satisfies = 100)

Weighted Rating: Multiply the individual ratings by the importance or adjusted importance.

Total: Add up the individual weighted ratings for each alternative. If the adjusted importance is used, the total score will be between 0 and 100.

Combining Cost Components – There are several ways to combine the different cost categories (*direct, indirect, etc.*). The WEOAL approach is one useful method:

Suggested Method - Weighted evaluation of all factors (WEOAL):

WEOAL = *k* [Index of direct, indirect, indeterminate cost factors] + (*1-k*)[Total weighted evaluation of intangible factors]

k - relative importance of cost factors compared to intangible factors

Index = Lowest Cost Alternative/Alternative “X” Cost

12.3 Exercises

Exercise - Assembly Time

A team was working on improving the time needed to assemble an air handler. The team made three changes to the process. Did they improve cycle time?

Assembly Time

Unit	Time	Unit	Time	Unit	Time	Unit	Time
1	180	*15*	58	*29*	80	*43*	75
2	115	*16*	29	*30*	77	*44*	70
3	50	*17*	38	*31*	100	*45*	61
4	38	*18*	38	*32*	65	*46*	36
5	60	*19*	50	*33*	39	*47*	39
6	42	*20*	82 **(2)**	*34*	60	*48*	45
7	90	*21*	55	*35*	63	*49*	35
8	100	*22*	64	*36*	60	*50*	55
9	35	*23*	50	*37*	50	*51*	45 **(3)**
10	28**(1)**	*24*	32	*38*	40	*52*	84
11	55	*25*	88	*39*	60	*53*	17
12	97	*26*	50	*40*	60	*54*	43
13	40	*27*	115	*41*	95	*55*	42
14	95	*28*	60	*42*	95	*56*	37

Changes made at (1), (2), and (3).

Exercise - Engineering Drawings Preparation

The first quality improvement team chartered by an air conditioning company analyzed the problem of getting engineering drawings prepared in time to support manufacturing. They made several process changes, measured their results and showed that a significant improvement had been made.

Four months after the change was made, the process was measured again. The team was surprised to note that the "gains" made through their work had evaporated and the process was performing as before. Comment on what you think may have happened.

Exercise - Charging Pump Reliability

A nuclear power plant team worked on improving the reliability of *Charging Pumps,* which were a major source of maintenance time and expense. The team found that the operators were using some manual valves to control flow from the pumps and that this resulted in premature wear of pump components. The operators were notified in a memo from maintenance about the problem and, for a while, the pumps were more reliable.

Two years later, the pumps began breaking down again. The same problem was found to be the cause. Comment on why you think the countermeasures didn't "stick."

Exercise - Effective Problem Solving

Recall a few problems in your organization that have been “solved” more than once. Why haven’t these problems been *prevented from recurring?*

Exercise - That Same Ol' Problem

We've worked with some large organizations. We'll guarantee that multiple departments will be working on improving essentially the same process (i.e. engineering, manufacturing, set-up, defects, supplier issues, etc.). Even within one facility, several different areas or Black Belts will be working on the same process. Comment on this issue.

Exercise – Economic Evaluation

Here are a few “changes” that you might consider. Develop an economic evaluation of the “change.” Make assumptions as to costs and benefits.

- New motor for the bass boat vs. maintaining the existing one (still works – 6 years old)
- New refrigerator for the kitchen vs. keeping the current one (still works – 10 years old)
- New laptop for the Black Belt
- Purchasing an Exercise Machine vs. Joining the Exercise Club
- Outsourcing decision: doing your own laundry vs. having the dry cleaner do it

Prepare your analysis for presentation to the important “stakeholders” associated with the decision.

Exercise - Break-Even Analysis

A new hybrid SUV costs $38,775 and is advertised to get 31 mpg city, 36 mpg highway driving. A standard version of the SUV costs $30,440 (same features, except for the power train system) and is advertised to get 20 mpg city and 23 mpg highway driving. How many miles would you have to drive before the higher purchase cost of the hybrid is paid back in fuel savings? How much would a gallon of gasoline cost to reduce the payback time to 2 years?

Exercise - Cannibalization Analysis

When new products are introduced, their sales are forecasted. One aspect of the forecast that is often under-modeled is *cannibalization* of existing products.

Suppose you're launching Product B (unit price $100), which is expected to generate:

$$S_B^{\text{gross}} = 120{,}000 \text{ units/year}$$

You already sell Product A (unit price $120), which currently does:

$$S_A^{\text{baseline}} = 200{,}000 \text{ units/year}$$

If you ignore the potential for cannibalization – what is the forecast sales for Products A and B?

We define **cannibalization rate (CR)** as: Fraction of *new product sales* that come from customers who would have bought the old product.

Assume: Cannibalization rate (CR) = 30%

Considering cannibalization, how many Product A units will be sold after Product B's launch?

What would the revised forecast sales be for Product A and B together?

If you are the project leader for Product B, what would you report as *incremental* sales for your project?

Hint: Incremental Sales = (1 – CR) x Product B Sales

13.0 Changing

Unit	Description	Page
13.1	Change Management	13.1 - 1
13.3	Exercises	13.2 - 1

13.1 Change Management

Learning Objectives

- Be able to incorporate change management principles into process improvement efforts

Unit Contents

- Change Management
- From Current to Improved State
- Change Management Model and Activities

13.1.1 Change Management

This manual includes many techniques that support measurement and analysis of products and processes. We've included statistical methods to help you understand and tear apart the variation in your processes. We've also provided tools to help you understand and analyze the steps of your business processes. While these are excellent tools and methods, there is one last, *critical* dimension of process improvement that we will cover in this unit – *change management.*

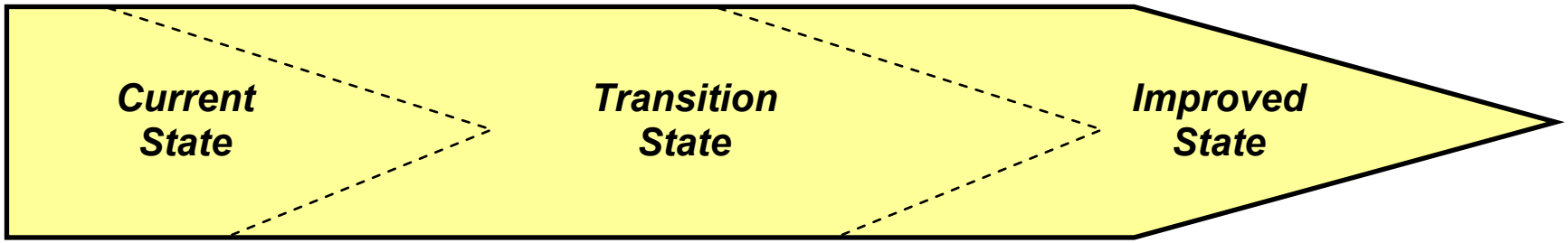
Periodically, someone will study the effects of programs such as ISO-9000, Total Quality Management, Reengineering and all the other "flavors" that litter the field of management improvement initiatives. Usually, the study will reach the conclusion that about 70% of all efforts fail to reach their goals. On one hand, there is no surprise in these figures. How many of us have started the New Year with wonderful goals to eat less, exercise more, spend more time with our families, etc. We had a conversation with one fitness center staff a few years ago and here's the picture she painted.

> "*Right after New Years, our center is filled. The regulars are there, but their ranks are swelled with all the New Years "resolutioners." They try, but by the end of March, almost all have dropped out and we're back to the regulars."*

Among the root causes of organizational failure to change include leadership support, resources applied, and inadequate attention to the *human* side of change. Whether you are planning the changes that result from your specific improvement project, or are planning how to improve the entire organization, the principles and practices on the next pages should be helpful.

13.1.2 From Current to Improved State

Rarely do people or organizations change overnight. Research has shown that to move from some current state to a new or improved state, we must experience a transition state. We first examine the change process; the change management activities applicable to the different states are then described.



Current State

Recall the last time you tried to learn something new. A few years ago, we decided to learn how to windsurf. In our "current state," we knew how to sail larger boats, and were intrigued by the idea of sailing *standing up*. There were, however, a few barriers to overcome in this journey. First, we faced the need to make a "capital investment." A good used windsurfer cost upwards of $1000 and the spouse wasn't willing to let us buy another "toy." Fortunately, a friend offered the use of his windsurfer while he was on a travel assignment. The next barrier involved time. With the current demands of work and family, we had to give something up. Although golf was a priority, we decided that game could take a hiatus while we learned how to windsurf. The last barrier was the "how-to." The friend that lent us the windsurfer gave us a couple of lessons to get started.

Transition State

The first couple of tries were frustrating. The first thing to learn is how to stand on the wobbling board. Next, you must pull the mast and sail out of the water while staying balanced on the board. It seemed to take a long time before these two acts could be accomplished without falling off the board. A lot of energy was expended without much in the way of results. However, persistence finally paid off. The magic moment occurred when we pulled the mast out of the water, grasped the boom and moved about 20 feet through the water before falling off again. The new skill was starting to take shape.

Improved State

Gradually, the distance traveled before falling off the board got greater. Soon, additional skills such as tacking and jibing the board were learned. Although we'll never be a 'world-class' windsurfer, we've gotten to the point where we can get up on the board and pretty much sail where we want. Oh, and the spouse did consent to us buying a used windsurfer when she saw that we were really interested in this new hobby!

Lessons Learned

A few lessons to "take away" from this experience:

- For any change to be successful, "*You've got to have the gottawanna!"* For teams or organizations, this includes *creating a shared need* for change and *shaping a vision* of the future.

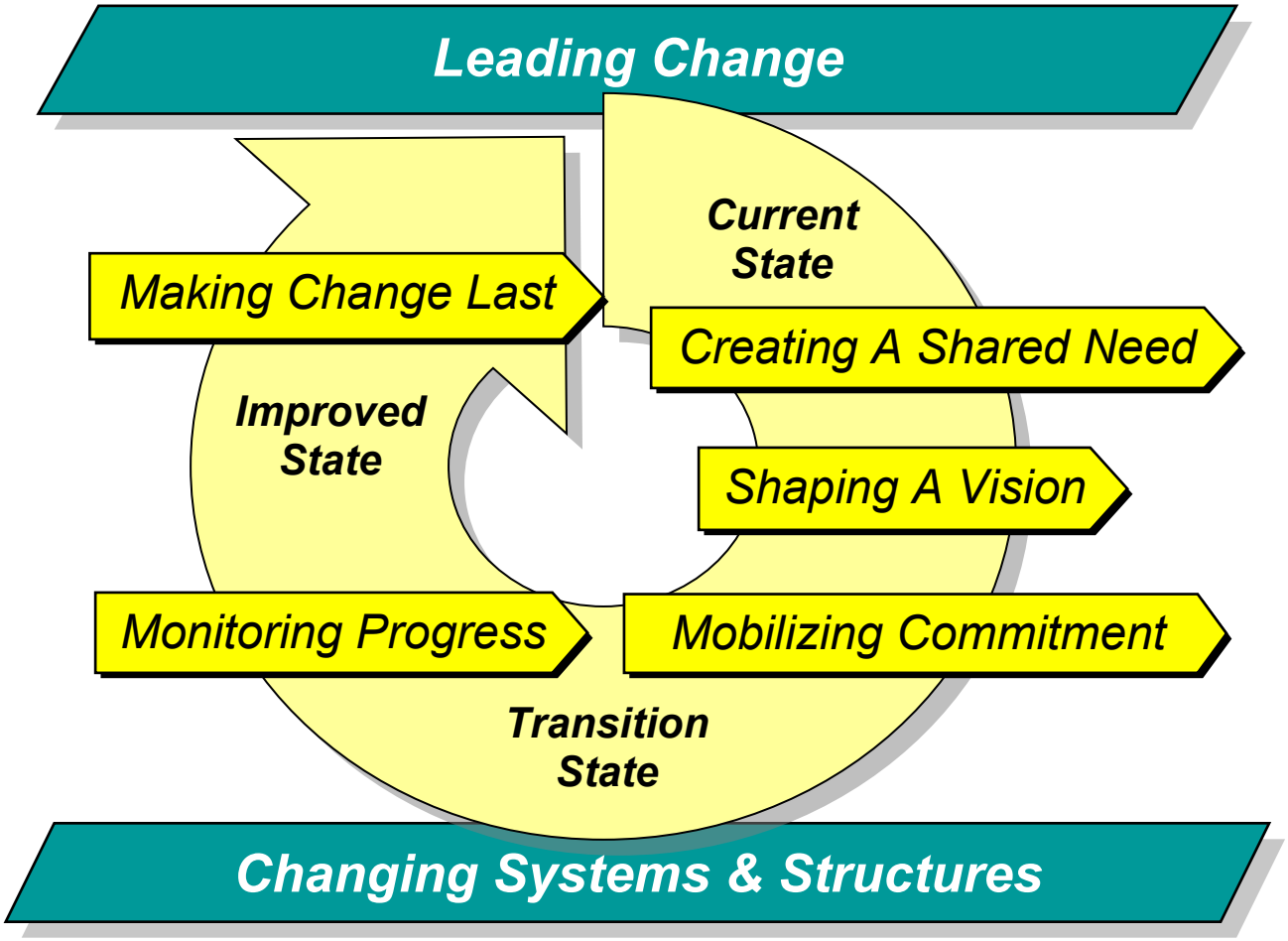
- Sufficient resources must be committed to help the desired change occur and support from key "stakeholders" must be secured. Some current activities must be put aside to learn and practice the new skills. In essence, we must *mobilize commitment* for the change.

- Allowance must be made for the "falls" that will inevitably occur during the learning phase. Both failures and successes should be analyzed to "get better at getting better." The change's *progress* must be *monitored.*

- Rewards for both trying and succeeding in the new "way" must be included. The new methods will require constant reinforcement and practice. *Systems and structures* must be changed to support the new way.

13.1.2 Change Management Model

From the windsurfing "change" example, we can derive some basic principles and processes of change. We'll use the following model to frame the change management discussion. The model includes specific activities (i.e. creating a Shared Need). The activities are mapped against the states of change that, as individuals or organizations, we generally go through in adopting a new way of doing business.



13.1.3 Leading Change

Key Principle: A Champion or Champions who sponsors the change

Why Bother?

- σ Strong, committed leadership is critical to accelerating change
- σ Leadership impacts all other change processes
- σ Leaders must play varied roles

What Are We After?

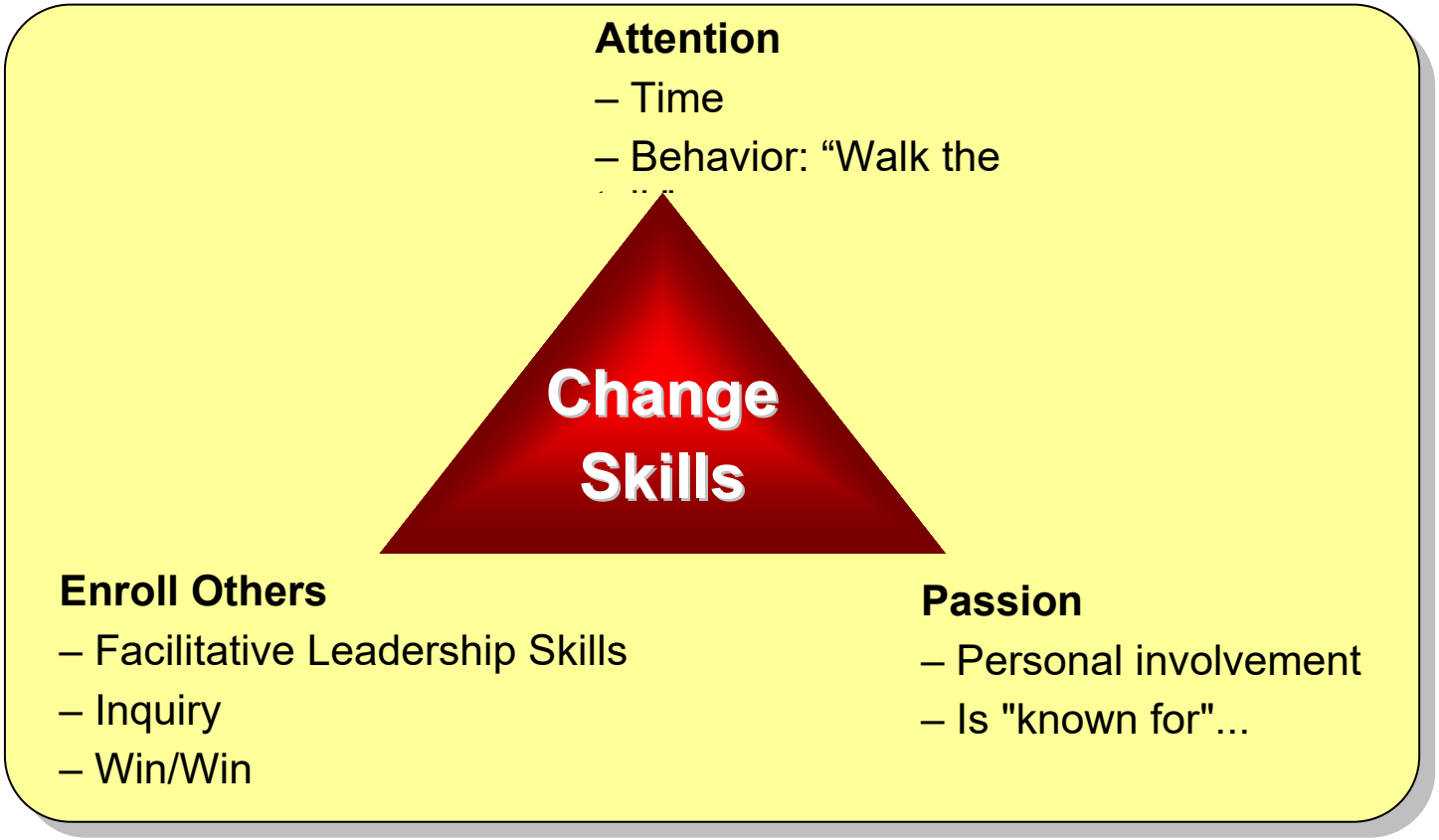
- σ Visible, active, and public commitment and support
- σ A willingness to take personal initiative and challenge status quo
- σ High levels of attention to the project through the time, passion and focus given to the project by leaders at all levels

Assessment

To what extent do change leaders:

- σ Seek and support process innovations that improve productivity?
- σ Clarify roles and responsibilities for accomplishing change?
- σ Vigorously challenge the status quo?
- σ Lead by example?
- σ Pay attention to change? (focus + time + passion)
- σ Demonstrate personal competencies across the seven CAP processes?

Leadership Change Skills



Tools and Techniques

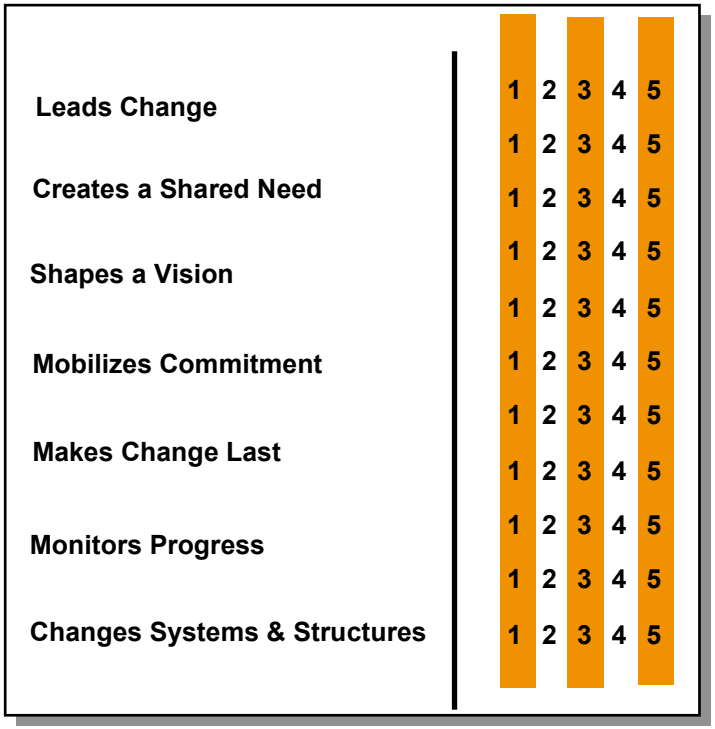
Calendar Test

1. **Identify 4-5 things you feel strongly about (at home or at work)**
2. **Check your calendar for the last 2-3 months to see what % of your time is spent on those things you say are important to you**

CAP Self-Assessment

After reviewing the material in this unit, try to honestly assess your skills relative to the key activities associated with change management. Ask a few co-workers to rate you – even though this may be a challenge!

Develop an action plan to address your key change management gaps.



Pitfalls - Teams can potentially derail when they:

- σ Fail to engage in behaviors necessary for change
- σ Are transferred too quickly before change has occurred
- σ Try to do it all alone without involving others
- σ Shift to other goals before completing the change initiative
- σ Fail to establish and clarify the key change roles of Champion, Agent, and Target
- σ Allow the change process to be diluted by too many competing initiatives and priorities

Leading Change - Summary

Outcomes of Leading Change

- σ Visible, active, public management, attention and behavior that support change
- σ Willing to take personal initiative and support others' initiatives for change
- σ Clarity around roles for accomplishing change
- σ Leaders are known as change advocates

Questions To Ask Assessing Leading Change

To what extent do our leaders:

- σ Seek and support process innovations that improve productivity?
- σ Clarify roles and responsibilities for accomplishing change?
- σ Vigorously question status quo?
- σ Lead by example?
- σ Find opportunities in change rather than excuses for avoidance?
- σ Pay attention (focus time, have passion) to change?
- σ Demonstrate personal competencies of a change advocate?
- σ Assign critical roles for change?

Actions to Lead Change

- σ Master the processes for accelerating change
- σ Manage, time, energy, and focus
- σ Demonstrate personal leadership competencies required for change
- σ Articulate roles of change sponsor, agent, and target

Pathologies Common to Leading Change

- σ Leaders fail to engage in behaviors necessary for change
- σ Leaders are transferred too quickly before change has occurred
- σ Leaders try to do it all alone without involving others
- σ Leaders shift to other goals before completing initiative

13.1.4 Creating a Shared Need

Key Principle: The reason to change whether driven by threat or opportunity is instilled within the organization and widely shared through data, demonstration, demand or diagnosis. The need for change must exceed its resistance

Why Bother?

- σ Forces any resistance or apathy to be addressed head-on
- σ Validates why the project is important and critical to do
- σ Builds the momentum needed to get the change initiative launched

What Are We After?

- σ A shared recognition, by both the team and key constituents, of the need and logic for change
- σ Dissatisfaction with the status quo (greater than the natural resistance to change)
- σ The ability to frame the need for change as both a threat and an opportunity

Assessment

- σ Are all members of the project team aligned in terms of the need for change?
- σ Have we framed the need for change in such a way to reflect the concerns of customers and key suppliers?
- σ Would each team member deliver essentially the same "message" regarding the need for change if asked by someone outside of the team?
- σ Who are the key constituencies affected by this initiative, and how much importance does each give to the initiative?
- σ How can we help others increase their sense of the need for change?

Tools and Techniques

SWOT Analysis

SWOT Analysis (Strengths, Weaknesses, Opportunities and Threats) is a classic tool to identify organizational gaps and to begin to build the case for change. A simple matrix can be employed to brainstorm these:

Strengths	Opportunities
Weaknesses	**Threats**

Teams Can Potentially Derail When They:

- σ Fail to check for alignment and build true consensus
- σ Assume the need for change is obvious
- σ Fail to frame the need for change in a meaningful way
- σ Assume that when others fail to appreciate the need for change it's "their" problem
- σ Fail to search beneath the surface for root causes
- σ Underestimate the resistance to change

Creating a Shared Need – Summary

Outcomes of Creating a Need

- σ Shared belief among key players that there is a need and logic for change, critical over time, and the need is greater than its resistance

Questions To Ask Assessing Creating a Need

- σ How well do we currently perform on the issue we want to change?

  - In the eyes of the customers?
  - In the eyes of the employees?
- How critical is improved performance on this issue for business results?
  - Because of threat?
  - Because of opportunity?
  - Short term or long term?
- How widely shared is the need for change?
- To what extent is the need for change greater than its resistance?

Actions to Creating a Need

- Demand: Set high expectations of others (e.g., devil advocates, bold statements)
- Data/Diagnosis: Generate external or internal data to induce change (e.g., external visits, reports, surveys, benchmarks)
- Demonstrate: By example, lead the change (being an example)

Pathologies Common to Creating a Need

- Assume need is obvious to all
- Not meaningfully characterized
- Crying wolf
- "They" just don't get it
- Fuzzy diagnosis (symptoms vs. causes)
- Underestimate resistance

13.1.5 Shaping a Vision

Key Principle: The desired outcome of change is clear, legitimate, widely understood and shared

Why Bother?

- σ Visions paint a picture that appeal to both the "head" and the "heart" and answer the question, "Why change?"
- σ Clear statement of the future state helps gain genuine commitment
- σ A well-articulated vision helps establish the milestones to monitor progress and change systems and structures

What Are We After?

- σ A clear statement about the outcomes of the change effort
- σ A view of the future state that is:
  - Customer focused – Not just one person's dream
  - Challenging – Evolving, not static
  - Easy to understand – Behavioral and actionable

Assessment
To What Extent:

- σ Has a vision been clearly articulated for the project?
- σ Is the vision simple and straightforward?
- σ Is the vision motivating and energizing?
- σ Is the vision shared and understood across the business?
- σ Is the vision actionable?

And Finally, What Are We After?

- σ How aligned is the team around the vision?

Tools and Techniques

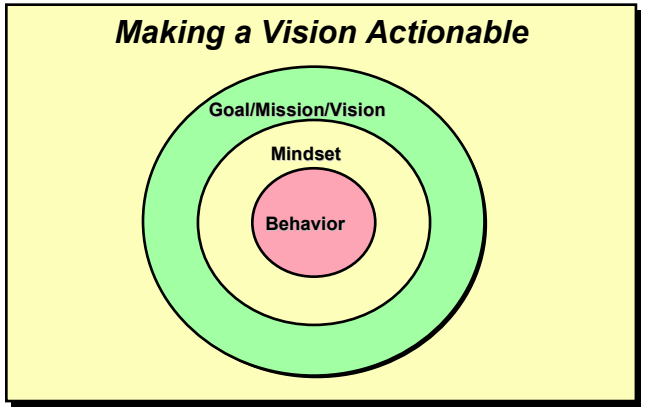
Looking Back from the Future

1. Imagine a point in the future when the project has been very successful.
2. Find words to describe what you would see, hear, feel as you observe key constituents functioning in the new, changed state.

3. Collate, debate, reach consensus, "test" on others and modify

Bull's eye Chart

Visions must be translated into concrete actions and behaviors. The Bulls eye Chart is simply a graphical way of starting with a clearly articulated vision and then exploring (i.e. through brainstorming) the *mindset* and *behaviors* that are necessary for the organization to achieve the vision.



Communicating the Vision – The Elevator Speech

1. Imagine a chance meeting of a project team member and a key stakeholder in an empty elevator with 90 seconds to ride.

2. Describe the need for change and the vision of the new state, as one might respond to the question, "Why are we doing this project?" by addressing the following elements:
- σ Problem/issue
- σ Benefit
- σ Where are we at
- σ What others can do

3. Team members practice this "speech" so they can convey a uniform message to others

Teams Can Potentially Derail When:

- σ Everyone has their own vision, and no effort is made to gain alignment
- σ Vision statements remain at such a "lofty" level that no one pushes back
- σ The vision changes too often, or conversely, is so rigid that others feel excluded
- σ The vision fails to reflect the interests and needs of customers and/or key suppliers
- σ The vision is too complex to be easily understood or translated into day-to-day behaviors
- σ A strong link between the Need for Change and the Vision of the desired state is missing or vague

Shaping a Vision – Summary

Outcomes of Shaping a Vision

- σ A clear statement about the outcome of the change effort will be articulated in both emotional (visual, enticing) and pragmatic (numerical) ways

Questions To Ask Assessing Shaping a Vision

To what extent has:

- σ A vision has been clearly articulated.
  - – What's in it for customers?
  - – What's in it for employees?
- σ Is the vision simple and understandable?
- σ Is the vision shared and known in the business?
- σ Is the vision motivating and energizing?

Actions to Shaping a Vision

- σ See the world from the customer's point of view (What would the customer like more/less of?)
- σ Articulate a vision that others can readily embrace
- σ Create a bold and clear sense of purpose that energizes others
- σ Create enthusiastic support for business objectives

Pathologies Common to Shaping a Vision

- σ No single statement of a vision; everyone has their own version
- σ No buy-in that this is the direction we want to move; everyone does not support the vision in private talks

- No continuity; the vision changes too often
- No tie with customers; the vision focuses too much on what we want, not what customer wants
- No simplicity of vision; the vision is too complex to be easily understood and translated to practice

13.1.6 Mobilizing Commitment

Key Principle: There is a strong commitment from key constituents to invest in the change, make it work, and demand and receive management attention

Why Bother?

- σ Need sufficient support and involvement from key stakeholders
- σ Critical mass must be won over
- σ Key difference between success and failure

What Are We After?

- σ Coalition of committed supporters
- σ Identification of potential resistance
- σ Conversion of key influencers

Assessment

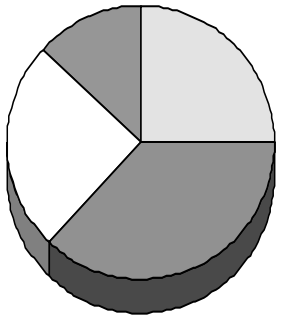
How well has the team:

- σ Identified key constituents?
- σ Analyzed sources of resistance?
- σ Maximized win/wins through conflict resolution?
- σ Developed problem-solving processes to build commitment?

Tools and Techniques

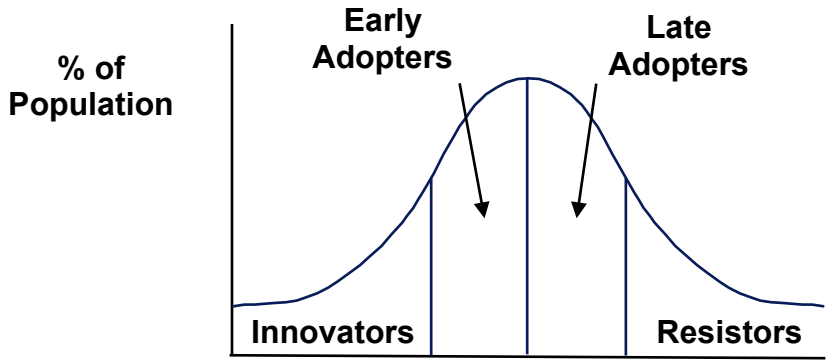
Key Stakeholders Map

Determine who the stakeholders are. One definition of a "stakeholder" is someone who can drive a stake through the heart of your project or change. Develop a pie chart of these stakeholders – the size of the stakeholder's slice is proportional either to the number of stakeholders in the group or to their influence.



Attitude Charting

Consider where the stakeholders fall on the attitude chart. Consider how to leverage the stakeholders' attitudes. You may start the change with a group of innovators, then, once the change has shown evidence of working, move to the early adopters. Whatever you do, don't start by trying to influence the resistors!



Stakeholder Analysis

The Stakeholder Analysis consists of three parts:

1. An assessment of the current "position" of each stakeholder relative to the change and a gap analysis of where they need to be for the change to be successful.
2. An analysis of why stakeholders who have large "gaps" are either against the change (or sometimes, too strongly for the change).
3. A plan to identify influence strategies to help close the stakeholder "gaps."

1. Stakeholder Assessment

Identify each important stakeholder. Discuss and note where they fall on the assessment matrix, using an "X." (below). Determine where the stakeholders should be for the change to be successful – note these with an "O." You may also

draw relationships between/among key stakeholders. For example, if stakeholder "A" is strongly against the change, but stakeholder "B" has an influence on "A," then draw a dotted arrow line from "B" to "A."

Stakeholder Assessment

Names	Strongly Against	Moderately Against	Neutral	Moderately Supportive	Strongly Supportive

2. TPCF Analysis (Technical-Political-Cultural-Financial)

Don't make the mistake of considering stakeholders who resist the change as either the "enemy" or as "dinosaurs." They are behaving rationally in their own minds. Try to understand *why* they are resisting the change. Sources of resistance usually fall into one of four categories – Technical (i.e. afraid to learn a new computer system because of potential to fail), Political (i.e. afraid since their power or prestige may suffer as a result of the change), Cultural (i.e. perhaps the change conflicts with their ethics, or other habits), and Financial (i.e. resistive due to the perceived cost of the change vs. perceived benefits).

If you can identify the source of resistance, then you will be in a much better position to identify influence strategies to address these sources.

3. Influence Strategy

Based on the analysis above, identify possible influence strategies to overcome resistance to your change. Use the matrix below to help you plan your actions.

Stakeholder	Issues/ Concerns	Identify "Wins"	Influence Strategy

Teams can derail when they:

- σ Have little political sensitivity
- σ Fail to recognize need to "share the glory"
- σ Assume technical solution is sufficient
- σ Don't involve others due to time constraints
- σ Use only 1 or 2 conflict resolution styles
- σ Fail to appreciate the human side of problem solving

Mobilizing Commitment – Summary

Outcomes of Mobilizing Commitment

- σ A coalition/network of relevant and committed individuals who embrace the change effort and visibly support it
- σ An ability to manage conflicts inherent in change and engage in appropriate problem solving
- σ An extended commitment to change throughout the organization

Questions To Ask Assessing Mobilizing Commitment

- σ To what extent have we identified the key individuals who must support and be involved with this change to guarantee success?
  - – Internally

  - Externally
- To what extent do we have extended buy-in for the change to happen?
  - Among employees
  - Among customers/suppliers
- To what extent have champions/sponsors been identified?

Actions to Mobilizing Commitment

- Form a coalition of key players who will be change advocates, sponsors, and agents
- Leverage sponsors to form a network of support
- Determine who will resist and the causes of resistance, so that resistance can be overcome
- Recognize ways of dealing with conflict to build commitment
- Engage in appropriate problem-solving activity

Pathologies of Mobilizing Commitment

- No political sensitivity to change
- Not sharing the glory of the success
- Assume that a technical solution is sufficient (e.g., I have the right answer, why isn't everyone else smart enough to see it)
- Not enough involvement and sharing responsibility
- It depends on conflict resolution style all the time

13.1.7 Making Change Last

Key Principle: Once change is started, it endures, flourishes and learnings are transferred throughout the organization

Why Bother?

- σ Experience shows that successful, sustained change is difficult to achieve without attention from the entire team
- σ Every change initiative will compete for time, resources and attention
- σ We often spend most available time on the launch of an initiative rather than its institutionalization

What Are We After?

- σ Consistent, visible, tangible reinforcement of the change initiative
- σ An integration of the new initiative with ongoing work patterns
- σ Changes to organizational systems and structures that help make the change a natural part of individual and team behavior

Assessment

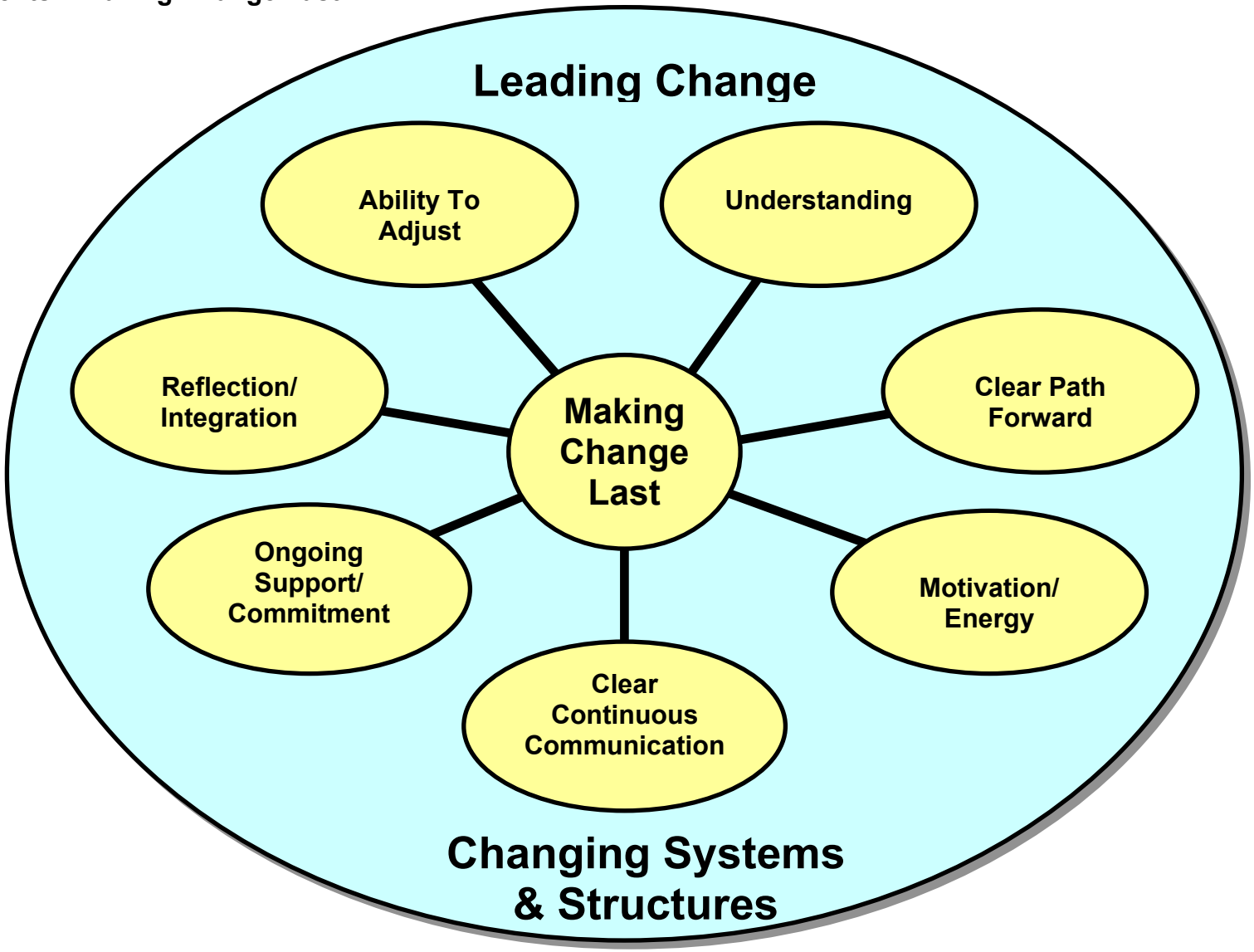
To What Extent Have We Accurately Estimated:

- σ The magnitude of the total change effort?
- σ The level of resistance this initiative will face?
- σ The amount of time required to implement the change?
- σ The level of clarity and alignment regarding the kind of implementation process required?

And Also…

- σ How has the change effort been integrated into other business initiatives?
- σ To what extent are needed resources made available?
- σ To what extent have we altered (or used) existing systems and structures as “levers for change”?

Elements – Making Change Last



Understanding

- σ What is being changed and what is not?
- σ What organizational context does the change take place in?
- σ Who/what will change affect and to what degree?

Clear Path Forward

- σ Are actions necessary for accomplishing the plan clear and logical?
- σ Do you have an action plan with responsibilities and dates assigned?
- σ Are there adequate resources available to accomplish actions proposed?

Motivation/Energy

- σ How will early successes be built into the project plan?
- σ Does the sponsor maintain a high level of personal enthusiasm?
- σ Does the team visibly show its excitement and enthusiasm?

Clear, Continuous Communication

- σ Is the organization aware of the progress of the change? (e.g. early successes; results of pilot projects; lessons learned; next steps)
- σ Was there a systematic effort to communicate the relatedness of the project to other ongoing initiatives?

Ongoing Support/Commitment

- σ Did there continue to be visible, impactful sponsorship?
- σ Did sufficient funds continue to be committed?
- σ Did sufficient employee time continue to be committed?
- σ Were appropriate deadlines honored? (or was the project completion date arbitrarily advanced, or were people prematurely diverted to other work?)

Reflection/Integration

- σ Were project learnings and best practices shared widely throughout the organization?
- σ Were the efforts on behalf of the project well integrated with other organizational initiatives?
- σ Was the desired change reflected in the organizations “systems and structures”?

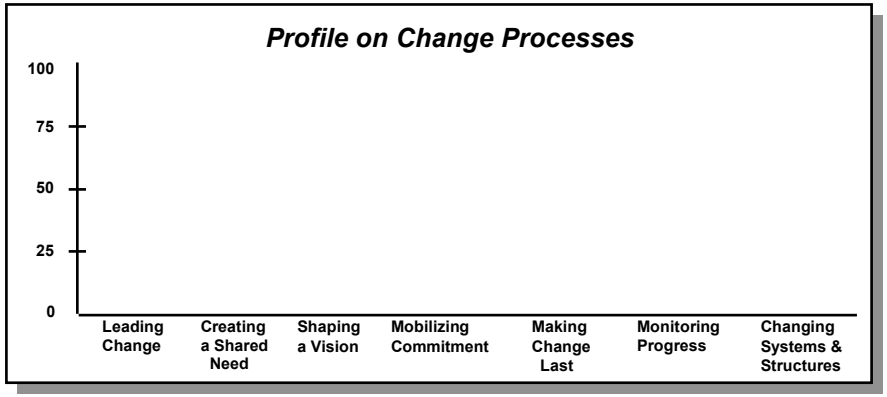
Ability To Adjust

- σ Did "downstream" project activities reflect and benefit from key learnings gained early on?
- σ Has our plan been adaptable to the actual needs of the organization?

Tools and Techniques

Change Profile

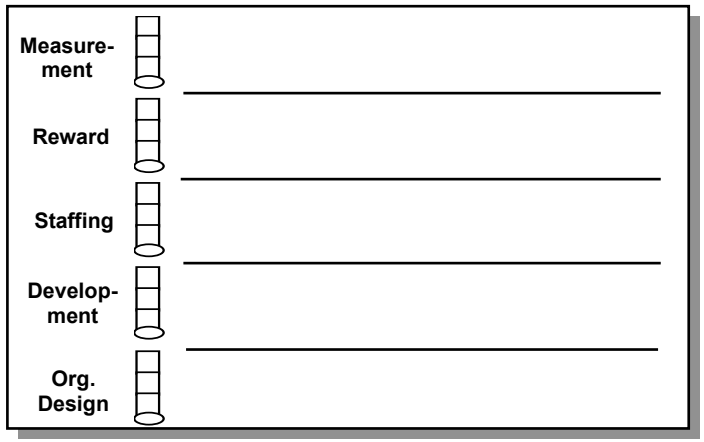
The Change Profile is a tool you can use to assess your teams/organizations' ability to execute on the key change activities. Have each member of your team assess the organization's skills/abilities for each of the change activities. Then develop a common profile (the discussion around why each team member scored a particular activity high or low is important, not the actual score).



Change Management Self-Assessment (See Leading the Change)

Systems and Structures Worksheet

The key "Systems and Structures" common to change include *measurement, rewards & recognition, staffing, development,* and *organization design.* Use the matrix below to plan what systems and structures will have to change and how these will be changed.



Teams can Derail When:

- σ Underestimating the time
- σ Unexpected problems
- σ Poorly coordinated activities
- σ Competing distractions
- σ Inadequate capabilities/skills of employees
- σ Lack of support for the initiative
- σ Unclear goals and objectives
- σ Lack of involvement of change targets
- σ Dismissing complaints outright
- σ Uncontrollable externalities (Life happens)

Making Change Last – Summary

Outcomes of Making Change Last

- σ Actions occur and changes are initiated that are visible, immediate, credible, integrated and lasting over time

Questions To Assess Making Change Last

- σ How has the change effort been integrated into other business initiatives?
- σ Are people willing to act without full plans and information?

- How effectively do we transfer learning across boundaries?
- Are the necessary resources made available?

Actions for Making Change Last

- Acts in ways that demonstrate public commitment to the change
- Transfers learning from one site to another; run lots of small experiments
- Assigns responsibility for making the change last
- Leverages symbols, language, and culture to support the change
- Encourages participative empowered leadership throughout the organization
- Integrates any one change to overall business process
- Drives results through change

Pathologies for Making Change Last

- Not institutionalizing the change process; seen as assignment
- Apathy as energy shifts to other projects
- “We have already done it” syndrome
- Trying to do all things at once and not making progress on any
- Waiting for the “perfect” solution before acting
- Resources required are not available
- Not sure where to start – too much needs to be done at once
- Focusing on activities not results
- We are unique, so it doesn’t apply

13.1.8 Monitoring Progress

Key Principle: Progress is real; benchmarks set and realized; indicators established to guarantee accountability

Why Bother?

- σ An accurate measure of the CAP project provides focus, direction and momentum
- σ Corrective action can only occur if you know you're off track
- σ Monitoring progress enhances your ability to reward key events and milestones, building momentum and commitment

What Are We After?

- σ Agreement and understanding on what the change effort will produce, in measurable and observable terms
- σ Baseline data and milestone results of the change process tracked and widely shared
- σ Increasing momentum as people begin to see progress and results are being realized

Assessment

- σ Have we stated our objectives in concrete terms?
- σ Have we translated these objectives to observable behaviors?
- σ Have we set milestones that we all understand and agree to?
- σ Are expected results tied to external and internal goals, and have we ensured that outcomes will be evident to stakeholders?
- σ Are individuals and teams accountable for results?
- σ Do we know which existing data will pick up progress toward our goal?
- σ Have we established new ways to gather data?
- σ Do we have accurate and timely baseline data to work from?

Teams Can Potentially Derail When They:

- σ Want results too soon and fail to look for long-term indicators of progress
- σ Assume all stakeholders know how things are going and fail to keep them informed
- σ Measure only against internal issues or goals, forgetting that customers are often impacted by the change initiative
- σ Don't see how the change project is connected to other initiatives and fail to measure impact
- σ Think some things are too "soft" to measure, only look at "hard" indicators of progress
- σ Simply get too busy to track progress

Monitoring Progress – Summary

Outcomes of Monitoring Progress

- Results of the change effort are tracked and shared widely

Questions To Ask Assessing Monitoring Progress

- Have desired results been articulated in concrete terms?
- Have milestones been set along the way?
- Are results tied to business goals?
- Are outcomes of the change evident to customers?
- Are outcomes related to speed/cycle time?
- Are results of the change effort widely shared?
- Are specific individuals accountable for results?

Actions for Monitoring Progress

- Develop measures of progress for change effort
- Track progress
- Share results widely
- Hold people accountable for results

Pathologies In Monitoring Progress

- No measures of success, so no indicators of progress
- Not tying change effort to business results
- Managers want results too soon and pre-judge the effort
- We assume that one change only affects one initiative; we miss the unanticipated consequences of change

13.1.9 Changing Systems and Structures

Key Principle: Making sure that the management practices are used to compliment and reinforce change

Why Bother?

- σ When the way we organize, train, develop, reward, compensate, promote, etc. is changed, we are likely to see individual behavior change
- σ Successful changes usually involve significant realignment of "organizational infrastructure"
- σ Need to develop the capacity to change, not just the ability to change – "Can we build this change into our ongoing systems?"

What Are We After?

- σ Identification of key System & Structure areas that must be addressed to assure long-lasting project completion and implementation
- σ Utilization of System & Structures Best Practices
- σ Alignment of Systems & Structures with desired behaviors

Six Aspects

Changing Systems & Structures Involves Using/Modifying:

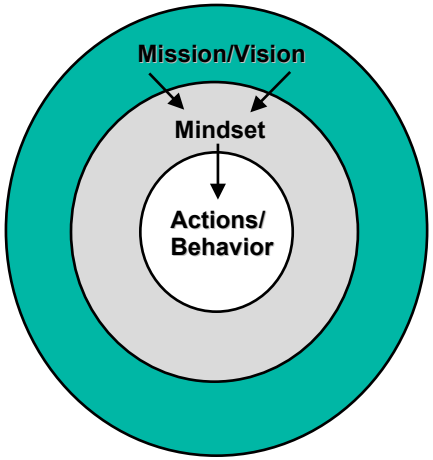
- σ Staffing (How we acquire/place talent)
- σ Development (How we build competence/capability)
- σ Measures (How we track performance)
- σ Rewards (How we recognize/reward desired behavior)
- σ Communication (How we use information to build and sustain momentum)
- σ Designing (How we organize to support the change initiative) Organizations

Tools and Techniques

Measurement Assessment

Use the following matrix to determine if the measures you are currently tracking will support or hinder the change you are trying to implement. People respond strongly to measures – especially if their compensation depends on achieving a

certain target or limit. Deming had some particularly harsh words to say about the practice of setting goals and targets in organizations. Read his book, *Out of the Crisis*, for his thoughts on performance goals.

Making a Vision Actionable 	**3** Existing Measures 1. 2. 3. 4. 5. 6. 7. 8. 9. 10.	**4** Which existing measures provide little or no information about any of the desired behaviors? 1. 2. 3. 4. 5. etc.	**5A** Which desired behaviors cannot be reliably measured by any existing measure? 1. 2. 3. 4. 5. etc. **5B** How could we measure these desired behaviors?	**6** If you achieve your desired changes, which existing measures will emit false signals indicating that performance is degrading? 1. 2. 3. 4. 5. etc.	**7** If an employee carries out each of the desired behaviors, what is likely to be the organization's reaction? A = Reward or Approval B = Punishment or Disapproval C = No Reaction D = Impossible to Predict More of: 1. 2. etc. Less of: 1. 2. etc.	**8A** Existing Rewards Financial 1. 2. 3. etc. Non-financial 1. 2. 3. etc. **8A** How could we reward the desired behaviors not now being rewarded?

Rewarding Performance Examples

The following "test" can help you understand the current state of your organization relative to changing systems and structures. Using the following response choices, indicate what would be most likely to happen if you were to do each of the actions described below:

1 = the action would usually bring reward or approval
2 = the action would probably bring neither approval nor disapproval
3 = the action would probably bring punishment or disapproval
4 = response to this action is unpredictable, ranging from approval to disapproval

Action		Choice (1 – 4)
a.	Coming up with and trying out new, untested ideas	
b.	Exceeding your authority when necessary to get the job done	
c.	Bending the rules when necessary to get the job done	
d.	Violating the chain of command when necessary to get the job done	
e.	Going along with the boss even when (s)he's wrong	
f.	Always going along with the majority view	
g.	Presenting your boss with an unpopular point of view	
h.	Sharing information with other units and departments	
i.	Keeping information from other units and departments	
j.	Achieving your group's goals at the expense of other groups	
l.	Setting very easy goals and then making them	
m.	Maximizing short-term bottom line at the expense of long-term bottom line	
n.	Spending money in the short term that will pay off later	
o.	Achieving your volume and budget objectives, but at the expense of product or service quality	

Teams Can Potentially Derail When:

- σ There is a failure to realize new competencies (acquired through staffing and development) may be required to fully implement their project
- σ Good intentions directed at refining the Reward System fail because of inattention to the critical Measurement System

- σ Early assessment of the organization's Measurement and Reward Systems (the Characteristics of Good Measurement and Reward Systems) is not conducted
- σ Effective communication about their project and its progress isn't started immediately and maintained throughout the effort
- σ Teams neglect to examine the organizational "infrastructure" for alignment with the new organizational goals and objectives required by their project

Changing Systems and Structures - Summary

Outcomes of Changing Systems & Structures

- σ A "tool kit" is available and the appropriate tools are identified and used. Tools include:
  - Acquiring/placing talent (staffing)
  - Building competence/capability (development)
  - Standards (measurements/rewards)
  - Designing organization (structure)
  - Communication
- σ Uses every available communication technique to learn, inform, persuade and inspire; helps others to do the same.

Questions To Ask Assessing Changing Systems & Structures

To what extent (has/have):

- σ A framework with multiple tools been articulated or used in making progress towards change?
- σ Tools have been prioritized and actions taken?
- σ Changes in one tool have been integrated with changes in another?
- σ Multiple tools have been used?

Actions for Changing Systems & Structures

- σ Selects the most talented people available
- σ Coaches others to improve
- σ Ensures employees have the competencies to do their job
- σ Establishes relevant, challenging objectives for self and unit
- σ Provides specific, frequent feedback that improves team performance
- σ Shares credit and recognition with others

- σ Shares and seeks information widely
- σ Listens effectively
- σ Communicates openly and candidly

Pathologies for Changing Systems & Structures

- σ Try one tool, then quit
- σ Impose one tool on everyone: “One size fits all”
- σ Misuse of the tool – not being competent in the design, delivery, or application of the tool

13.2 Exercises

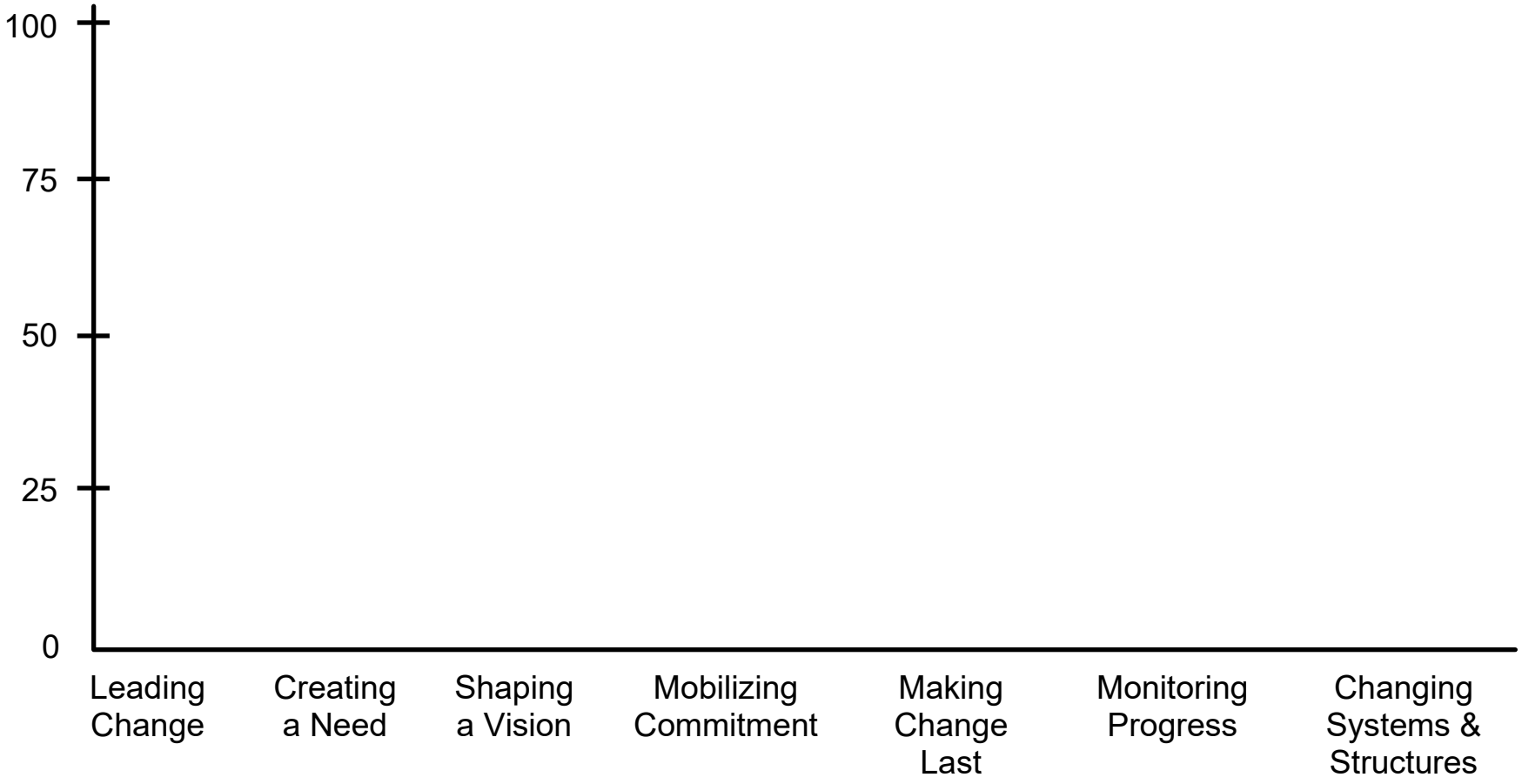
Change Profile

Individually:

- Using the worksheet below, “profile” your business’s typical approach to change (5 minutes)

Table Teams:

- Compare business change profiles (5 minutes)
- Identify elements of the profiles that are common



Need Alignment Test

- Each team member gets one card
- Each person writes a response to the following question (in 20-30 words):
  "*WHY are we doing this project?"*
- Ask each person to turn over their card and give a numerical response to the following questions (scale of 1 [not important] to 100 [critical]):

  "How important is the success of this project to your organization?"

  "How important is the success of this project to you, personally?"
- Have each person share their responses in small groups

Customer Focused Alignment Test

- Each team member gets three cards
- Each person writes one idea per card in response to the following question:

  "What do we (name of business) want to be known for by our customers?"

- Collect cards
- Separate into common piles
- Test: 75% of cards should be in one of three common piles

Stakeholder Analysis

Desired Outcomes:

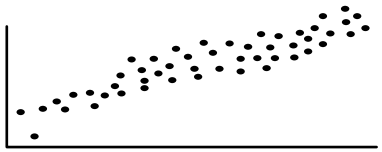
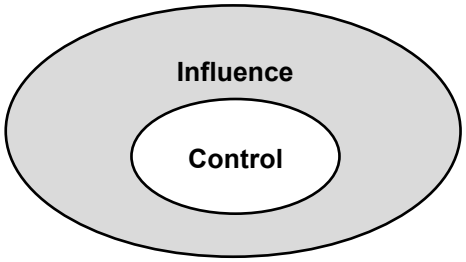
- σ Practice developing a Stakeholder Assessment to identify support and resistance for your project so that you can understand how to use this tool with your team at work.
- σ Practice identifying types of resistance (TPCF Analysis) for your project so that you can understand how to use this technique with your team at work.
- σ Discuss an Influence Strategy to overcome resistance so that you can understand how to use this approach with your team at work.

What	How	Who	Time
Team Preparation	■ Chose two team members to act as meeting facilitators ■ Facilitators prepare charts	Team Facilitators	10 minutes
Identify Support & Resistance for Change	■ Introduce the Stakeholders Assessment tool/approach ■ Lead group through brainstorming/discussion	1st Facilitator	15 minutes
Feedback	■ +/Δ : Use of tools Facilitation	All	5 minutes
Identify Types of Resistance for Project	■ Introduce TPCF Analysis tool/ approach ■ Lead group through brainstorming/discussion	2nd Facilitator	15 minutes
Feedback	■ +/Δ : Use of tools Facilitation	All	5 minutes
Develop Influence Strategy	■ Determine strategies/tactics to influence stakeholders resistive to your project	All	15 minutes
Feedback	■ +/Δ : Use of tools Facilitation	All	5 minutes

Changing Systems and Structures

Desired Outcome:

- σ Agreement on 2-3 systems & structures that your change project(s) most impacts.
- σ Agreement on aspects of systems & structures where you have control/influence/no control.
- σ Identification of 1-2 key stakeholders to involve for systems & structures for which you have no control.

What	How	Who	Time
Assess Project Impact And Agreement On 2-3 Systems & Structures Most Impacted	■ Propose ■ Advocate/evaluate ■ Check for agreement 		15 min.
Assess Control vs. No Control	■ Propose ■ Advocate/evaluate ■ Check for agreement 		10 min.
Identify 1-2 Key Stakeholders To Involve	■ Brainstorm ■ Prioritize (N/3) ■ Advocate/evaluate ■ Check for agreement		10 min.

14.0 Design Management

Unit	Description	Page
14.1	Defining Product/Service Requirements	14.1 - 1
14.2	Conceptual Design	14.2 - 1
14.3	Benchmarking	14.3 - 1
14.4	Taguchi Design Approach	14.4 - 1
14.5	Multi-Generation Planning	14.6 – 1
14.6	Exercises	14.7 - 1

These units support Unit 2.4 – Designing New Products and Services

14.1 Defining Product/Service Requirements

Learning Objectives

- Be able to develop requirements for new product/service designs
- Be able to deploy tolerances from high to low level requirements

Unit Contents

- Product/Service Requirements Development
- Quality Function Deployment
- Tolerance Deployment/Analysis

14.1.1 Product/Service Requirements

A critical first step in product/service design is the development of the Product/Service Requirements Definition (PRD). This doesn't state the product or service's *specifications* (these will be determined through the design process). Rather, the PRD specifies *what* the product or service is required to do. If the design team doesn't have adequate understanding of the product or service requirements, they are, in effect, set up for failure in their design activities. Dr. Stuart Pugh notes that many product/service requirements documents are "defective." This leads to defective products from the standpoint of meeting customer requirements. We've seen two major issues – a) the requirements *are* the specifications and b) the requirements are based on a chosen solution.

Product Requirements Definition	***Service Requirements Definition***

The following checklist provides an outline for a product requirements definition.
• Product performance
  ○ Overall systems
  ○ Subsystems
• Environment
• Life expectancy
• Product cost
• Quality & production life span
• Size and weight
• Appearance and styling
• Constraints and areas of flexibility
• Service and maintenance
  ○ Onsite maintenance
  ○ Offsite maintenance
• Auxiliaries and Interfaces
• Special features
• Regulatory standards and consideration
	The following checklist provides an outline for a service requirements definition.
	• Service performance
	  ○ Overall service
	  ○ Processes
	• Quality (i.e., sigma targets)
	• Service life expectancy
	• Service cost
	• Delivery (e.g., timeliness)
	• Quantity/production rates
	• Interfaces with existing systems/suppliers
	• Regulatory standards/considerations

What are the characteristics of "good" requirements for new product /service development? There are two questions we want to be able to answer at the end of product development – a) does the product meet the customers' needs and b)

does the product meet the design requirements that engineers set as guides to develop the product? In the DMEDVI method, the "V" stands for Verification and Validation. Design Verification is the act of checking the product against the design requirements and Design Validation is the act of ensuring the product meets customers' needs (There is also *Process Validation* – ensuring the process can reliably produce the product). To be able to accomplish these goals, we'll define two kinds of high-level requirements:

- **Customer Requirements** – These statements will be *validated* with the customer (e.g., via customer surveys after product use). We'll recommend the use of *functional* requirements – a statement that includes a Subject (the *who* generally the customer), a Verb (*what* function the customer is performing with the product) and a Modifier (the *how* - how well, how fast, how accurate, etc.) the customer wants the product to perform. This Subject-Verb-Modifier form is very useful as both a guide to design and when we present the product to the customer for Validation. The Customer Requirements should all have a rationale – why does this requirement exist (usually based in the team's Customer Research work). Example: *The surgeon can secure a clip on the targeted tissue with maximum accuracy.*

- **Product Requirements** – These statements will be *verified* with objective evidence (e.g., engineering testing). Like the Customer Requirements, we'll recommend product *functional* requirements – a statement that includes a Subject (the product), a Verb (describing the product's function) and a quantifiable Metric (a technical measure that will provide objective evidence the product meets the requirement). Associated with the Metric, we will define targets and limits, as appropriate. The Product Requirements are derived from the Customer Requirements and there should be a rationale for the requirement (based on the Customer Requirement). Example: *The clip applier can secure the clip to vessels < 5mm in diameter, with a pull-off force > 5 ft. lbs.*

Other than knowing we are developing a clip applier, the examples of Customer and Product Requirements are "solution open." They do not constrain the development team to a specific solution, allowing different design options to be considered in the Explore phase of DMEDVI. The first time you try to generate these kinds of requirements, it may seem frustrating (we know from experience!). We often recommend bringing an English Major into the session as we're doing a lot of language processing in the requirements generating sessions! You'll soon find that the process becomes natural and once you start thinking like this about requirements (and see their value), you'll be hooked!

14.1.2 Quality Function Deployment

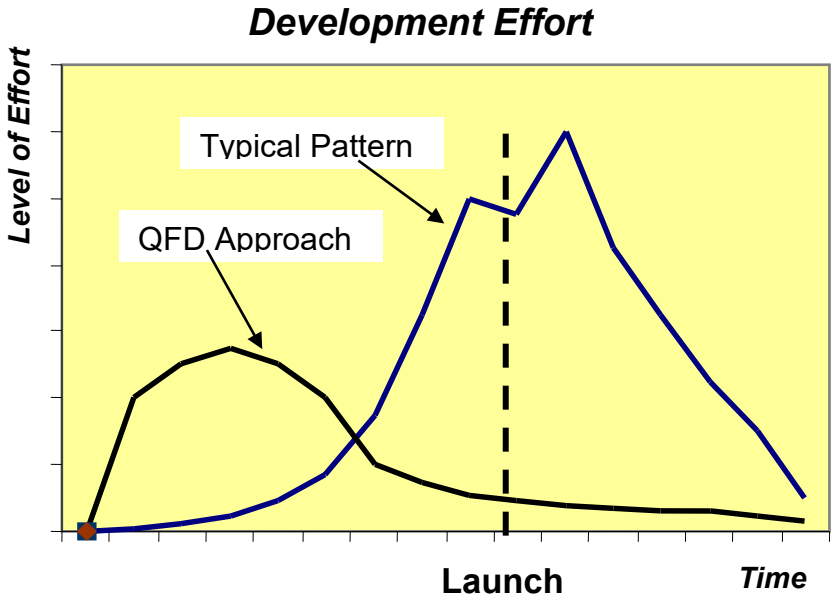
Quality Function Deployment (QFD) is a means of systematically deploying customers' requirements into the specifications, functional requirements, design requirements and process variables of a product or process. The systematic use of QFD is credited with reducing design cycle times by up to 50% and improving the chances of market success for new products and services. Practitioners describe QFD as:

- A means of letting the customer drive the design,
- A technique to keep us focused on the customer, and
- A process that forces discussion about customer requirements that either would not happen or would happen too late to do much about it.

Why QFD?

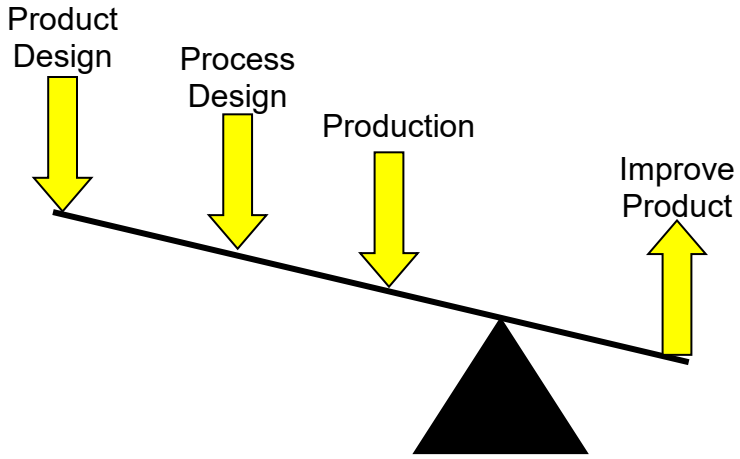
Companies that have integrated QFD into their product/service design processes have reported these benefits:

- Fewer design changes
- Design changes identified earlier.
- Less time in development.
- Fewer start up problems.
- Lower start up costs.
- Fewer field problems.
- More satisfied customers.



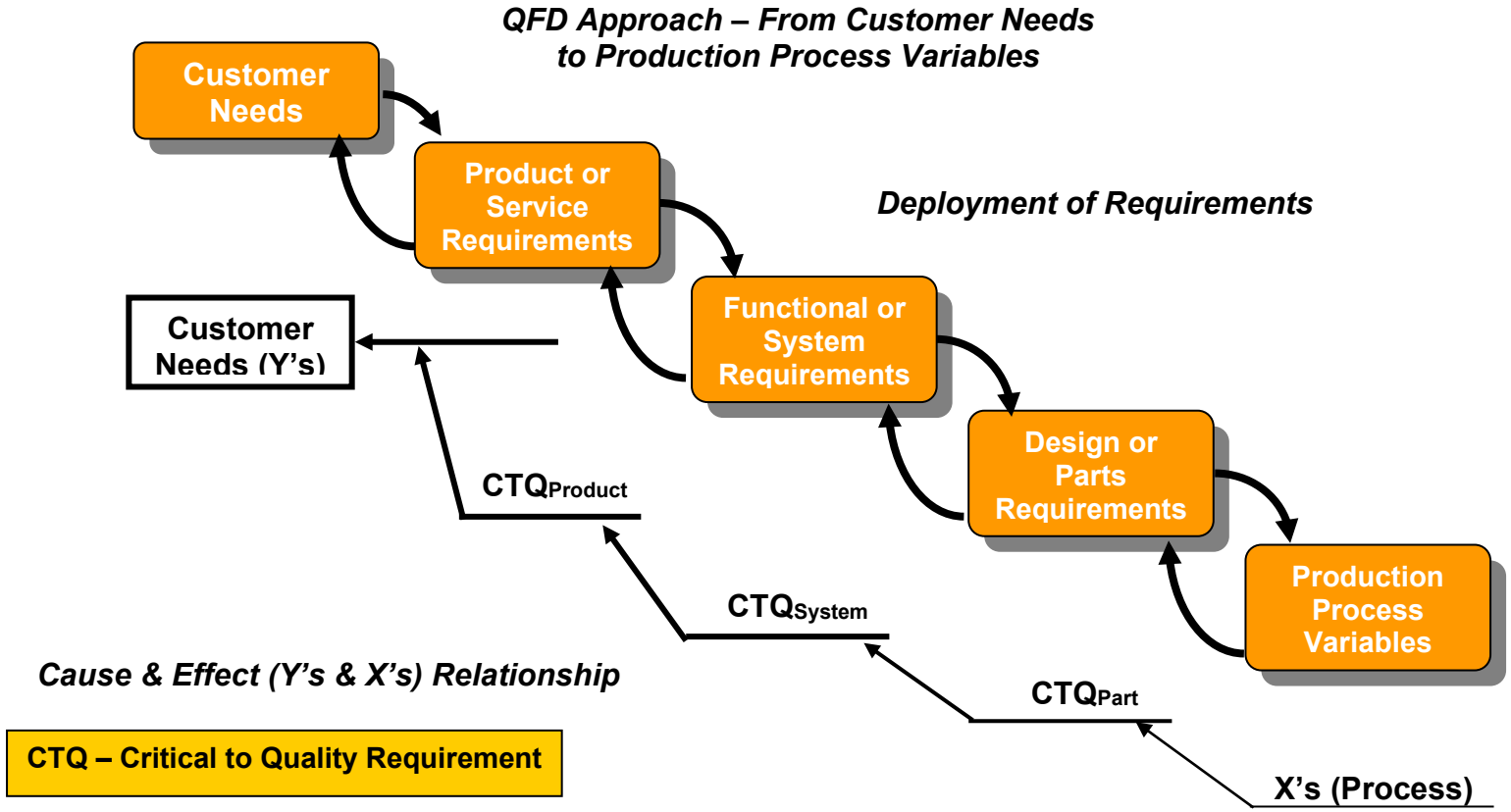
QFD – The Quality "Lever"

About 80% of the cost and quality of a product or service are determined by its design. QFD has been likened to a lever that acts to improve product/service quality "upstream" in the design process. The earlier QFD is used, the greater the potential impact of QFD.



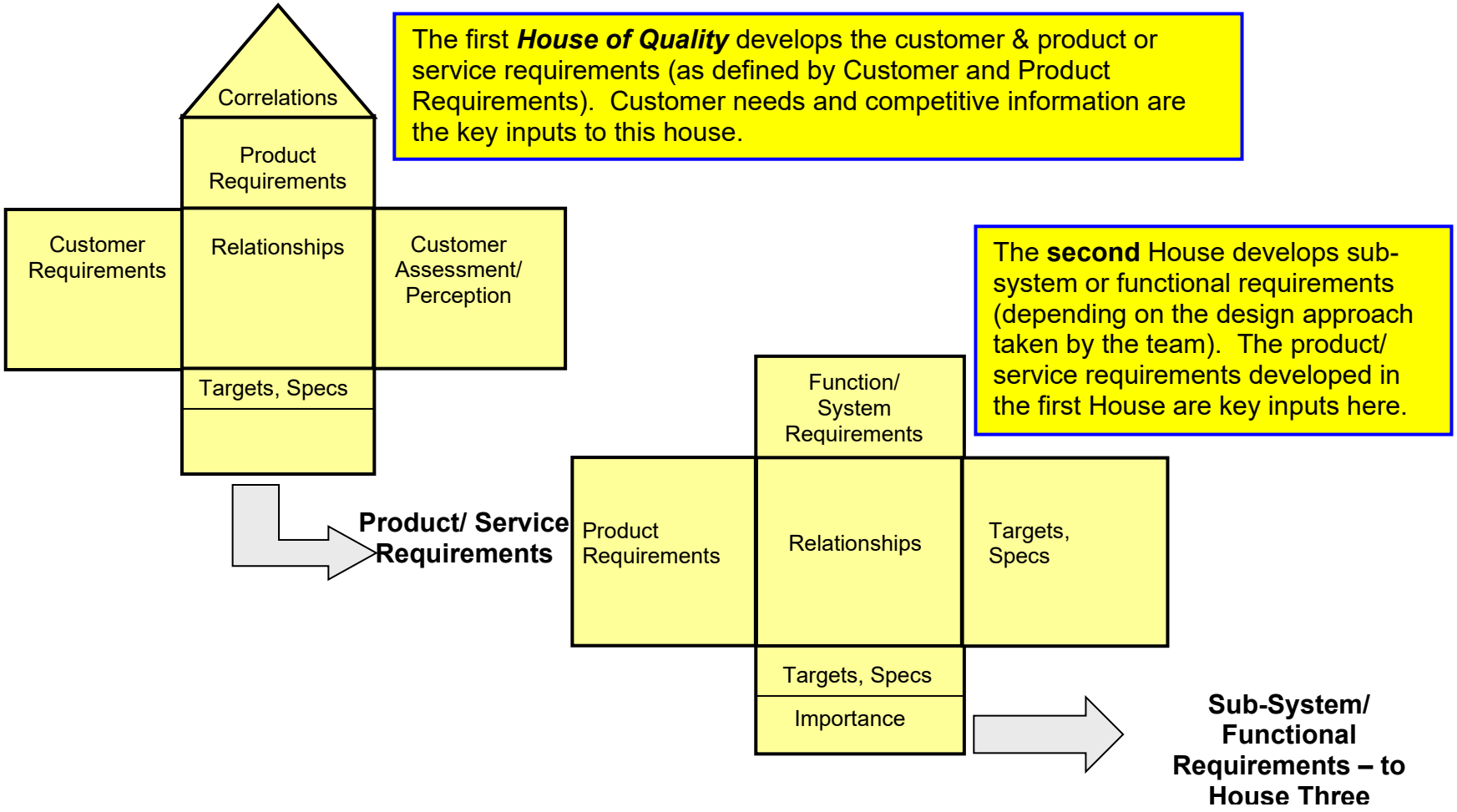
QFD Methodology – Broad Sense

QFD helps the design team *identify* and *deploy* the Voice of the Customer to the product or service and associated production processes. If you think of the customers' needs as the "Y's" or effects that you are trying to accomplish, then the product and process variables are the "X's" or factors that can be controlled to achieve the "Y's." In an **improvement** project, the team tries to *discover* and *improve* the important "X's." In a **design** project, however, the team *develops* and *specifies* the "X's." The development of these "X's" usually proceeds from a high (i.e. product requirements) to low level (i.e. production process variable). Intermediate requirements may be defined for product systems (or service processes) and the parts (or process steps) that comprise the product or service. To determine the cause-and-effect relationship, QFD will employ a series of analyses "punctuated" with "acts of design."



QFD – The "Houses of Quality"

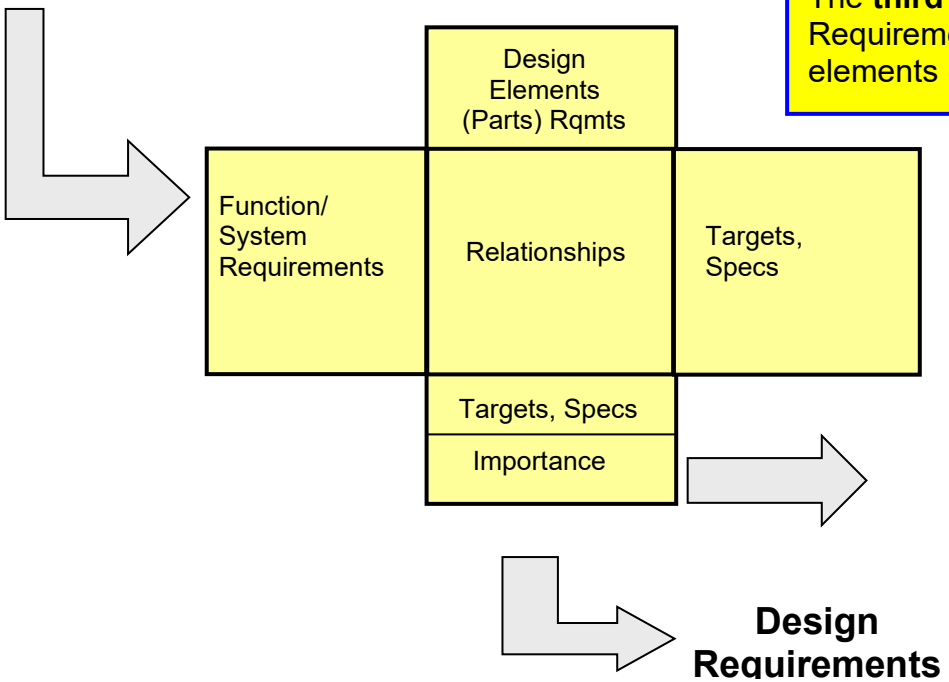
To implement the idea of *deploying* quality requirements, QFD employs a series of ***Houses of Quality***[1]***.*** Each "House" represents a design team analysis. The schematic below shows how, through four Houses, the needs of the customer may be deployed to specific product/service requirements and on to production process variable requirements. While this "Four House" approach is widely used, you must think about how to best integrate QFD into your specific design process



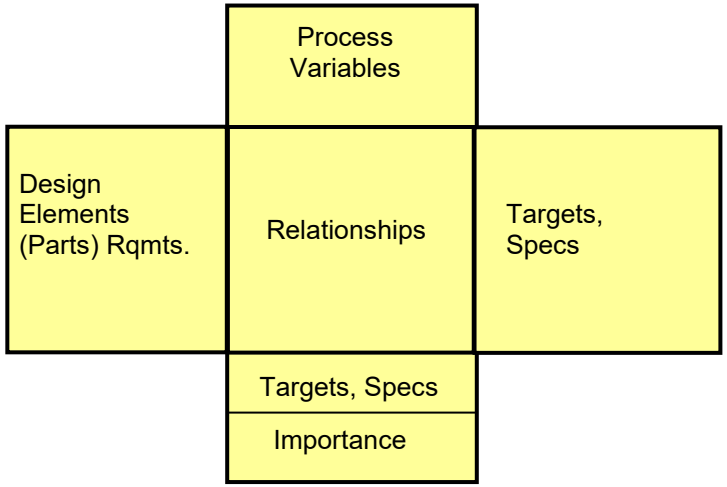
[1] Yeah, OK, so it "Takes a Village!"

Sub-System/ Functional Requirements – from House Two

The **third** House deploys Sub-System/ Functional Requirements into requirements for the design elements (or sub-assemblies/parts)



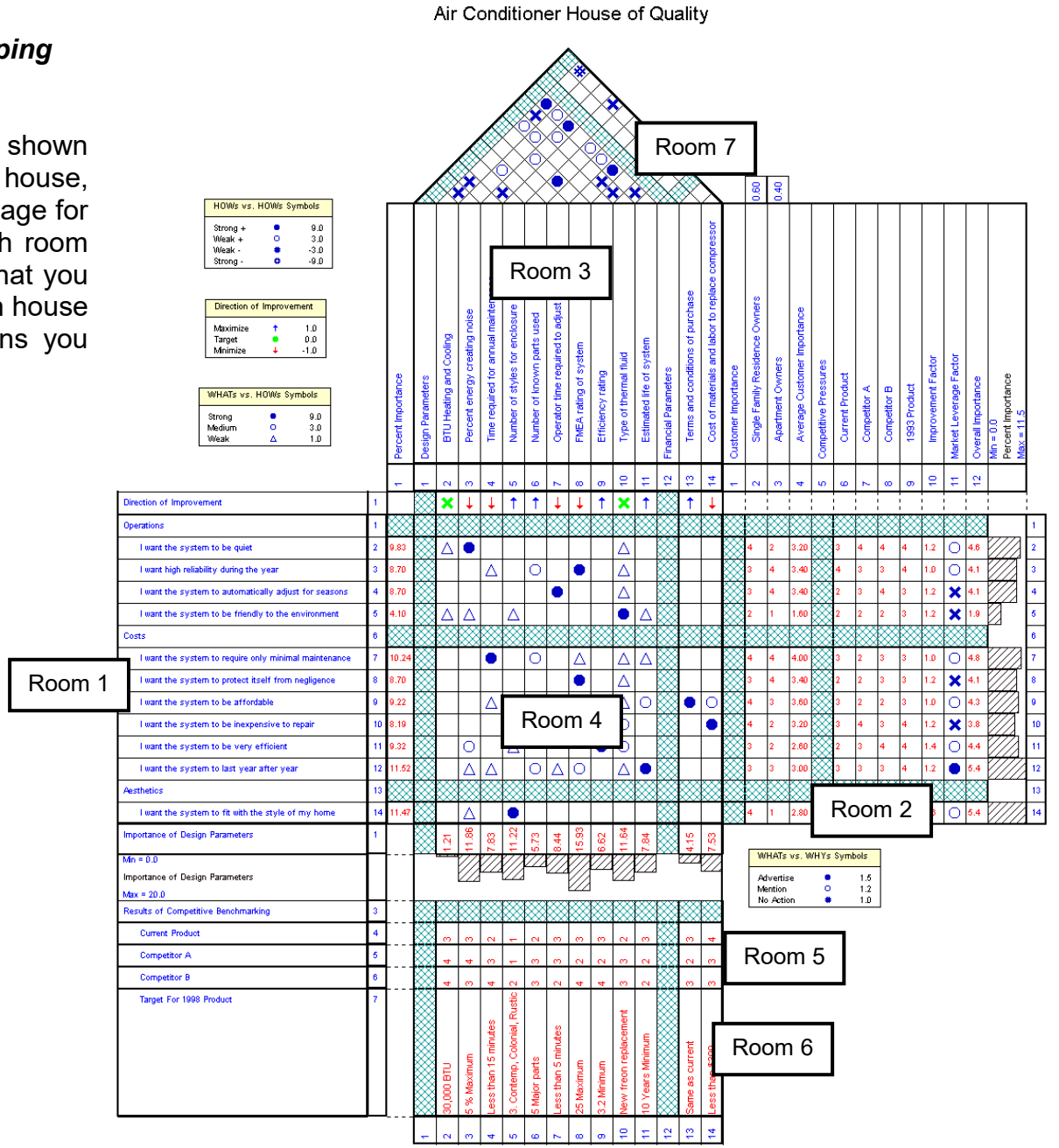
Design Requirements



The **fourth** House prioritizes the process variables based on their relationship to the design elements and sets control points for the production process.

The First House of Quality – Developing Product/Service Requirements

A complete, *first* House of Quality is shown here. The example is a "typical" house, comprised of 7 rooms. See the next page for a description of the "contents" of each room and how to "build" the rooms. Note that you have the flexibility to "design" your own house of quality; depending on the questions you need addressed.



The Rooms of the First House of Quality

1. **Customer Requirements** and their **importance**. This information is obtained from the customer through means such as observations, interviews or focus groups (see Section 4, Voice of the Customer). It is usually qualitative although some knowledgeable customers may give it in quantitative terms. The Customer Requirements are prioritized (by the customer) with a rating from 1 – low priority to 5 – high priority.

   Developing Room 1: The Quality Function Deployment process starts by "talking" to customers about their needs and desires for the new product and the priority they place on each need/desire. It's important to identify *all* the customers and conduct research on their needs. For example, an aircraft designer will seek the voice of customers such as the pilots, passengers, maintenance, flight attendants, and baggage handlers. Segmentation should be considered in defining customers. For example, passenger segments could include frequent business travelers and once-a-year family vacationers.

   Customer needs are generally qualitative in nature and must be obtained in enough detail to provide the design team with guidance for how to *respond* to the need. Need statements such as *good quality* or *easy to use* must be clarified and expanded into specific needs that can be *acted upon* by the design team. The structure tree approach can help organize customer needs from *broad* to *specific.* The team will then take the customer needs and *translate* them to Customer Requirements (Subject-Verb-Modifier language). The Customer Requirement statements are entered into Room 1 of the House.

High-Level Needs	Lower-Level Needs	Customer Requirement Statement (Room 1)
I want to feel comfortable during the flight	My legs should have enough room	The passenger should be able to sit with a maximum amount of legroom
	I don't want to feel too cold or hot	The passenger should experience a comfortable cabin temperature
	The plane shouldn't be too noisy	The passenger should experience the flight with a minimum of ambient noise
	The seat should be comfortable	The passenger should be able to sit with a maximum of comfort

Notice a couple of things about the Customer Requirements. They all start with the customer (passenger). They all include some function that the passenger is performing (or experiencing) during the flight. The modifiers are written a) so that they can be validated with the customer (we'll use a Likert scale), b) they all are either *maximizing, minimizing,* or *targeting* some scalable quantity.

2. **Customers' perceived performance rating**. This is a rating by the customer of your performance on the Customer Requirements as compared to your competitors. The rating employed is usually from 1 to 5 with 5 being the best performance.

   Developing Room 2: If the product or service is to be launched into an existing, competitive market, customers will have experience with similar products/services. They should be able to tell you how competitors are currently meeting their needs (and your performance, if you have a current product/service in the market). Surveys of your customers will help you quantify their *perceptions* of performance relative to meeting their needs. Typically, a 1 – 5 scale is employed, 1 being assigned to the poorest *perceptual* performance, 5 to the best. When comparing yourself to the competition, attempt to identify best-in-class companies; it will do you little good to know how you stack up against the bottom of the pack.

3. **Product/Services Requirements.** These will define what the product or service must do to support the Customer Requirements. They will have targets/limits set based on the analysis of competitors' performance, your own capability and the industry.

   Developing Room 3: Once the customer requirements are known, the *translation process* to product/service requirements begins. The team brainstorms *Subject-Verb-Metric* statements of the product or service that, if developed in the design, will result in meeting the customers' needs/customer requirements. In some cases, more than one product/service requirement will be needed to address the customer requirement. The team must determine how to *quantify* or *measure* each characteristic (the Metric). For example, several "candidate" product/service requirements are listed for the following customer needs:

Customer Requirement Statement (Room 1)	Product Requirement Statement (Room 3)
The passenger should be able to sit with a maximum amount of legroom	The back to front seat distance should be greater than A
The passenger should experience a comfortable cabin temperature	Cabin Temperature should be adjustable by the passenger (+/- B degrees F)
The passenger should experience the flight with a minimum of ambient noise	Cabin Ambient Noise should be less than C (db)
The passenger should be able to sit with a maximum of comfort	Seat cushion should have a softness greater than D (judgment scale 1-5)

4. **Relationships** of the product/service requirements to the customer requirements. This relationship signifies how well the "candidate" product/service requirement correlates to the customer requirement.

   Developing Room 4: Each "candidate" product/service requirement is correlated to each customer requirement. In each cell of the matrix, the team will score the correlation as either *strong* (score of 9, ◉ "bull's-eye" symbol), *medium* (score of 3, ○ "circle" symbol), *weak* (score of 1, △ "triangle" symbol) or *none* (blank cell). Make sure you correlate the ***product*** requirement to the ***customer*** requirement – many teams make the mistake of doing the reverse – this is not the cause-and-effect relationship we are trying to establish! The product of the customer's rating for the requirement and the strength of relationship, summed for the requirement's column is used to prioritize the product/service requirements. The priority scores can be normalized to a 1-5 scale and carried to the next house.

5. **Competitive performance**. This is a rating of your actual performance on the requirements and your competitors' performance.

   Developing Room 5: This information is obtained by benchmarking (see Unit 14.3). The performance is rated from 1 to 5 with 5 indicating the best performer. Generally, you will use the same best-in-class competitors as in Room 2. One of the valuable analyses now possible is a comparison of *customers' perception (Room 2)* with *actual performance (Room 5).* There may be valuable insights when there are differences here.

6. **Targets and Specifications**. The performance needed on the metric to meet your product/service requirements. It may be set to exceed these requirements and create a competitive advantage. Specifications and sigma targets may also be included, as the design is developed.

   Developing Room 6: Review your customer research data. Often, the customers will suggest a level of performance that is desirable. You may consider your competition – if the requirement under consideration is very important to the customer or is a selling point for your product or service, you will certainly want to set aggressive targets. The targets and specifications set here may be somewhat "mushy." You may not have a good idea at this point whether these can be achieved. Be prepared to revisit the targets and specifications as you proceed through the design. You may also set a "sigma" target here – how often are you willing to "fail" at meeting the target. This type of goal will become important when you begin to assess the capability of your design.

7. **Correlation of Measures**. This room correlates the product/service requirements to each other, looking for requirements that either reinforce each other or conflict with each other.

Developing Room 7: Room 7 is just one half of a correlation matrix. Here, you are correlating the requirements to each other, looking for reinforcing or conflicting requirements. A positive correlation is indicated with a +, a strong positive correlation is shown with ++. This would mean as one measure is improved so is the other, such as reducing weight and increasing fuel mileage. Negative correlations are indicated with a – or - - for a strong negative correlation.

Analyzing the House of Quality

Once the house is completed, analyze the information. Some typical indications to look for are listed below.

1. Blank or no strong relationship in a row. This means that you have not identified a requirement that correlates to a customer's need. Your action is to add a new requirement to the house.

2. Blank or no strong relationship in a column. This means that the requirement doesn't correlate to customer needs. This may be an "extra" requirement that can be dropped. Before dropping it, though, ask whether there is a customer or customer segment for which this requirement is important.

3. Low competitive comparisons (Room 2). There is a low perception of your product's performance by your customers. This will require improvement of the product and a marketing campaign to get the word out. If your current product performs better than the competitions' (Room 5) but is poorly perceived by customers, perhaps better marketing is all that is needed.

4. Low competitive benchmarks. Your product's performance is poor. A redesign is needed with aggressive targets on key measures that relate strongly to priority customer requirements.

5. Strong negative correlation of priority requirements (Room 7). You have conflicts in your design that will require resolution. If you cannot develop design concepts that address these conflicts, you will have to accept tradeoffs. "Inventive" methods such as TRIZ may help break these design bottlenecks. If these conflicts arise because of needs from different market segments, you may decide to design and offer two different products or services. The positive note is that QFD has helped you identify the design bottlenecks early in the design process.

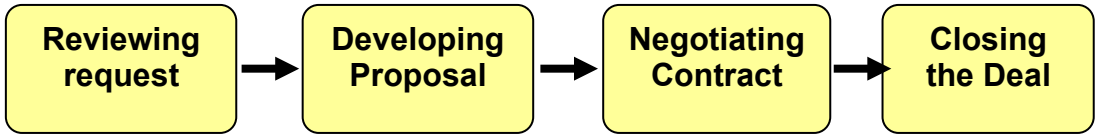
6. Requirements with strong relationships to several customer requirements. These are key requirements and meeting them should be the focus of your design effort. Set aggressive targets for these *Critical-to-Quality Requirements (CTQs)*.

7. Low priority for a requirement with a strong relationship to an important customer need. Be careful of the "mechanical" nature of the requirements prioritization. This requirement may still need to be included as a *CTQ,* even though it correlates to only one need.

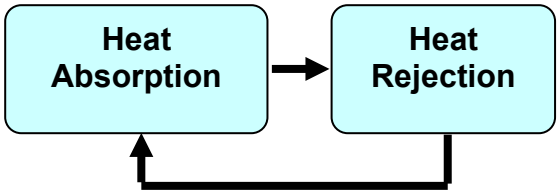
The Second House - Developing Functional/System Relationships/Requirements

The second QFD house relates the overall product/service requirements to the functions or sub-systems that are impacted. Note the difference between a *function* and a *sub-system.* The function tells us *what* must be accomplished, when we select a system (or a process, for service design), we are generally stating *how* the function will be performed. As discussed in Unit 14.2, if the design concept is *static,* the design team may begin with sub-systems, if *dynamic,* functions may be more appropriate since they do not constrain the *system concept* design at this point. In the matrix example below, the different systems of an airplane are depicted.

Functions of the Contracting Process



Functions of an Air Conditioner



Second House of Quality – Aircraft Systems to Product Requirements

Aircraft Systems

CTQs	**Importance**	Fuselage	Wings	Engines	Interior
Miles per hour	3	3	3	9	1
Size	4	9	1	3	3
Passengers	5	9	1	3	9
Gallons/Mile	5	3	9	9	1
Impact		105	63	99	65

Scale
9-Strong impact on the measure

3-Medium impact on the measure

1-Weak impact on the measure

Requirements are then generated for each CTQ impacted by this system, for example:

Requirement	**Target**
Weight	25,000 lb.
Drag	0.25
Volume	30,000 cu. Feet
Cross Section	20 ft. radius

Alternatively, a House of Quality may be developed for each sub-system to identify specific requirements.

The Third House - Design Elements to Functional Requirements

The third translation is to relate the design elements (or sub-assemblies/parts) to the functional/sub-system requirements. In the example below, the fuselage system has a weight requirement, a drag coefficient, and a volume and cross section requirement. Design elements are all the "pieces" of the design that can impact its function. They are the "parts" of the design. The output of this house includes the most important design elements. We then identify design requirements for the design elements, like developing the functional/system requirements.

Design Elements

Functional Req't	**Target**	Materials	Shape	Dimensions	Construction
Fuselage Weight	25000 lb	9		9	3
Fuselage Drag	.25	1	9	3	1
Fuselage Volume	30,000 cubic ft.	3	9	9	
Fuselage Cross Section	20 ft. radius	9	9	9	
Impact		22	27	30	4
Design Requirements		Light-weight	Drag coef.	Less than 20X20	Strong

Scale
9-Strong impact on the funct. req't.

3-Medium impact on the funct. req't.

1-Weak impact on the funct. req't.

The Fourth House - Production Variables to Design Elements

The fourth translation relates the production variables to design elements (or sub-assemblies/parts). These might be either fabrication or assembly requirements. In the example below, the fuselage system has a weight requirement, a drag coefficient, and a volume and cross section requirement. Production Variables are controllable factors in the plant that Design elements are all the variables of the design that can impact its function. They are the "knobs and dials" of the production process. We can now identify and specify requirements on the production process and suppliers – the "X's" of the overall product/service.

Production Variables

Parts Requirement	**Target**	Heat Treat Time	Jig Length	Nickel Content	Polishing Material
Material Strength	60,000 psig	9		9	
Surface Finish	20 Angstrom	1		3	9
Member Length	12"		9		
Material Hardness	8 (Rockwell Scale)	9		9	
Production Process Requirements		4 Hours	12"	0.3%	400 Grit

Scale
9-Strong impact on the parts reqm't.

3-Medium impact on the parts reqm't.

1-Weak impact on the parts reqm't.

The outcome of this house includes the critical production process variables. These will be translated into process management charts as important control variables (see Section 5).

14.1.3 Tolerance Deployment/Analysis

Through Quality Function Deployment, the design team establishes target values associated with product or service requirements. These also are *deployed* to lower-level requirements – sub-system, parts, process.

The team is also challenged with setting *specification or tolerance limits* for the requirements. At the product or service level, the tolerance limits are based on an understanding of the customer's willingness to tolerate deviation from the nominal or target value.

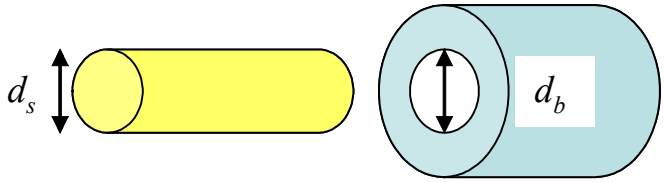
For example, a hand-operated surgical clip-applier will have a target for the "force-to-fire" requirement. The design team will also set an upper limit on force-to-fire – based on analysis of human strength and fatigue. Here, there may be no need for a lower limit. On the other hand, a rifle or pistol *will* have a lower limit on force-to-fire to prevent inadvertent weapon discharge.

At the other "end" of the design process, the design team must deal with the variability of the manufacturing process – i.e. how much variation is inherent in machining, casting, turning or other fabrication process.

This section provides some basic methods of evaluating the effect of manufacturing variability on the performance or dimensions of a manufactured product.

Linear Tolerancing

Let's start with the "classic" tolerance problem – a manufacturer wishes to produce a shaft and bearing that will form some part of a rotating or sliding assembly:



Here, the high level requirement is that the shaft be able to rotate freely (but not *too* freely!) inside the bearing. A lathe is used to turn the shaft, a drill is used to make the hole in the bearing stock. The shaft diameter is d_s, and the bearing diameter is d_b. The difference between the two diameters, the *interference,* will be termed *w.*

Several basic tolerancing methods are discussed below:

Worst Case Analysis

Suppose that the design team wants a "quick and dirty" estimate of the interference, *w*. The target value for the shaft diameter is 0.990" and the target value for the bearing inner diameter is 1.000". Further suppose that machine capability studies have identified that the lathe can produce shafts with a reported tolerance of +/-0.005" and the drill can produce holes with a tolerance of +/-0.007".

What *worst-case* interference may be expected? Here, we simply combine the dimensions at their worst-case maximum or minimum values. In this example, the worst-case interference is:

$$w_{WC} = d_{b-WC} - d_{s-WC} = (1.000" - 0.007") - (0.990" + 0.005")$$
$$= 0.993" - 0.995" = -0.002"$$

In the worst case, then, the shaft would *not* fit into the bearing. Worst case analysis should only be used when some assembly is made up of only a few parts and, because of some safety or customer issue, the parts can't interfere or be spaced too far apart. The functional relationship must be linear.

Note that we only did one calculation above – to predict if the shaft would not be able to be inserted into the bearing. We could also perform the opposite calculation – to predict if there was too much play between the shaft and the bearing. When performing worst case analysis, be careful to add or subtract the dimensions or characteristics in the appropriate manner. For example, some geometries will have dimensions where the parts double back on each other.

Root Sum of Squares Analysis – Expected variation in w

Suppose the design team wants to know the expected variation in *w,* the difference. We can use the *additivity of variances* property here – the standard deviation of the difference is the square root of the sum of the squares of the components (hence the term *root sum of squares*):

$$\sigma_w = \sqrt{\sigma_s^2 + \sigma_b^2}$$

In this case, we can estimate the standard deviations of the machines by dividing the tolerance half-width by 3 (this assumes a Cpk of 1 – the process fits in the tolerance band six standard deviations). For our machines, then the expected variation in *w* is:

$$\sigma_w = \sqrt{(0.005/3)^2 + (0.007/3)^2} = 0.0029"$$

The tolerance half-width for the interference is then 3 times the standard deviation or about 0.0086". Note that the combined tolerance half-width can be found by the following:

$$T_w = \sqrt{T_s^2 + T_b^2}$$

What useful information does this calculation provide to the design team? If the upper and lower specification limits for the assembly are known, then the probability of producing defective assemblies can be determined. If the specification limits for *w* are +0.020" and –0.000", then the inherent capability of the process can be calculated to be:

$$C_p = \frac{USL - LSL}{6\sigma} = \frac{0.020 - 0.000}{6 \times 0.0029} = 1.16$$

This corresponds to a process yield (see Unit 6.8 - *Process Capability Analysis*) of 99.95%, an expected fraction defective of 0.05% and a sigma (short-term) of about 3.2.

Root Sum of Squares Analysis – Required Variation in d_s *and* d_b

The case above asked what we could expect from a given production process. Let's turn the question around. Given that we have an allowable tolerance for *w,* what is the maximum variation allowable in the lathe and drilling processes? The same approach can be taken. However, we are starting with the output of the process and are seeking the inputs:

$$\sigma_w = \sqrt{\sigma_s^2 + \sigma_b^2}$$

If the allowable tolerance half-width for *w* is 0.010", and we evenly allocate the variability to the lathe and drill processes, then the process standard deviations must be no greater than:

$$\sigma_w = \sqrt{\sigma_s^2 + \sigma_b^2} \Rightarrow \sigma_w = \sqrt{2\sigma_p^2} \Rightarrow \sigma_w^2 = 2\sigma_p^2 \ and$$
$$\sigma_p = \sigma_w / \sqrt{2} = (0.010" / 2 \times 3) / \sqrt{2} = 0.0012"$$

The allowable tolerance half-width for the processes is then 3 X 0.0012" = +/-0.0036". Note that this is smaller than the current processes can deliver. Note also that we are employing the traditional relationship that "aims" for a process capability of 1.0 (by setting σ = T/3) that, over the short term, can be expected to yield 99.73% of its output within the specification limits. Long term shifts and drifts in the process will generally result in a much lower long-term process yield (M. Harry employs a "generic" 1.5 sigma shift to differentiate short- and long-term performance – in the example above, the long-term sigma would be 1.5).

Further, just because the product has been produced "within spec," degradation in reliability (time to failure) may occur as the actual dimension moves farther from its target (see Taguchi Design, Unit 14.4). Finally, you might challenge the equal allocation of variability to the drilling and lathe processes. Perhaps we know that one machine is inherently capable of machining to a smaller tolerance. We could then allocate less variability to this machine and more to the other.

Non-Linear Functional Relationships

The arguments above make use of the *Root Sum of Squares* technique, well known to most engineers. This approach is valid when the output variable is related to the input variables through an *additive* functional relationship (note that addition and subtraction are both "additive" functions). The general relationship for tolerance stack-up is shown below:

$$If\ Y = \sum_i X_i$$
$$Then\ \sigma_Y = \sqrt{\sum_i \sigma_i^2}$$

In many cases, though, the output and input variables are related through a *non-linear* relationship. For example, if resistors A and B are arranged in parallel and connected to a voltage source V, the resulting current in the circuit is:

$$I = V\left(\frac{1}{R_A} + \frac{1}{R_B}\right)$$

To address this tolerancing challenge, there are two common methods – *General Non-Linear Variance Decomposition* and *Monte Carlo Analysis.*

General Non-Linear Variance Decomposition

This method generalizes the Root Sum of Squares approach. For any "well-behaved" function:

$$Y = F(X_1, X_2, X_3, \ldots X_n)$$

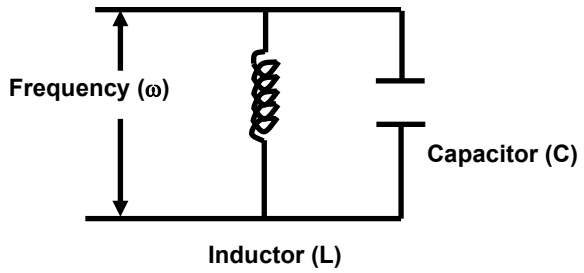
We can calculate the standard deviation of Y as:

$$s_Y = \left[\left(\frac{\partial f}{\partial x_1}\right)^2 s_{X_1}^2 + \left(\frac{\partial f}{\partial x_2}\right)^2 s_{X_2}^2 + \cdots \left(\frac{\partial f}{\partial x_n}\right)^2 s_{X_n}^2\right]^{1/2}$$

or:

$$s_Y = \left[\sum_i \left(\frac{\partial f}{\partial x_i}\right)^2 s_{X_i}^2\right]^{1/2} \quad where: \frac{\partial f}{\partial x_i} - \text{Partial Derivative of Y} = \text{f(x) wrt } x_i$$

Here's a simple example using an "L-C circuit." We want to calculate the variation in the frequency of this circuit as a function of the inductor (L) and the capacitor (C).



The equation for frequency is:

$$\omega = \sqrt{\frac{1}{LC}}$$

Where:
ω – Frequency
L – Inductance
C – Capacitance

We take the partial derivatives of frequency with respect to Inductance and Capacitance:

$$\frac{\partial \omega}{\partial L} = -\frac{1}{2}\sqrt{\frac{1}{CL^3}} \quad \text{and} \quad \frac{\partial \omega}{\partial C} = -\frac{1}{2}\sqrt{\frac{1}{C^3 L}}$$

Combining these into the equation for the frequency standard deviation, we obtain:

$$s_\omega = \left[\frac{s_L^2}{4CL^3} + \frac{s_C^2}{4C^3 L}\right]^{1/2}$$

To calculate the frequency standard deviation, we'd input the target values for Inductance (L) and Capacitance (C), along with estimates of their standard deviations. These could be obtained by dividing the tolerance half-width by three as we did above.

Using this method, our challenge is to develop the *partial deriviatives.* It's likely been a while since you had to practice these! Fortunately, our next method, *Monte Carlo Analysis* will avoid the need for this.

Monte Carlo Analysis

Although the idea of Monte Carlo analysis has been around for a while, it took the development of modern computers to make it a popular method for *simulating* variation in a "Y" as a function of variation in "X's." The basic idea is simple. We will define the functional relationship between the Y and its X's

$$Y = F(X_1, X_2, X_3, \ldots X_n)$$

We will then assign probability distributions (e.g., normal, log-normal, Weibull) to each X along with the necessary distribution parameters (means, standard deviations, etc.). Here's where the computer comes in. Monte Carlo will draw a sample from each input distribution. Suppose one of our input distributions is a Normal, with mean 10 and standard deviation 1. The samples will generally fall between the mean and +/- 3 standard deviations, so we'd expect our sample number to fall between 7 and 13.

We'll then take the samples from each distribution and calculate the Y value using the Y = F(X) function. This gives us one value of the Y. This process will be repeated many times. For a typical engineering prediction, 10,000 runs will occur. For estimation of rare events, more than 1,000,000 trials may be needed. The values of Y are stored as if they came from real experiments.

These results can then be analyzed – the mean and standard deviation can be calculated for the Y, a histogram can be generated and a probability distribution fit to the output. If the prediction is associated with a product/service requirement, the data can be used to predict the *capability* of the requirement (e.g., Cpk or Sigma-Level).

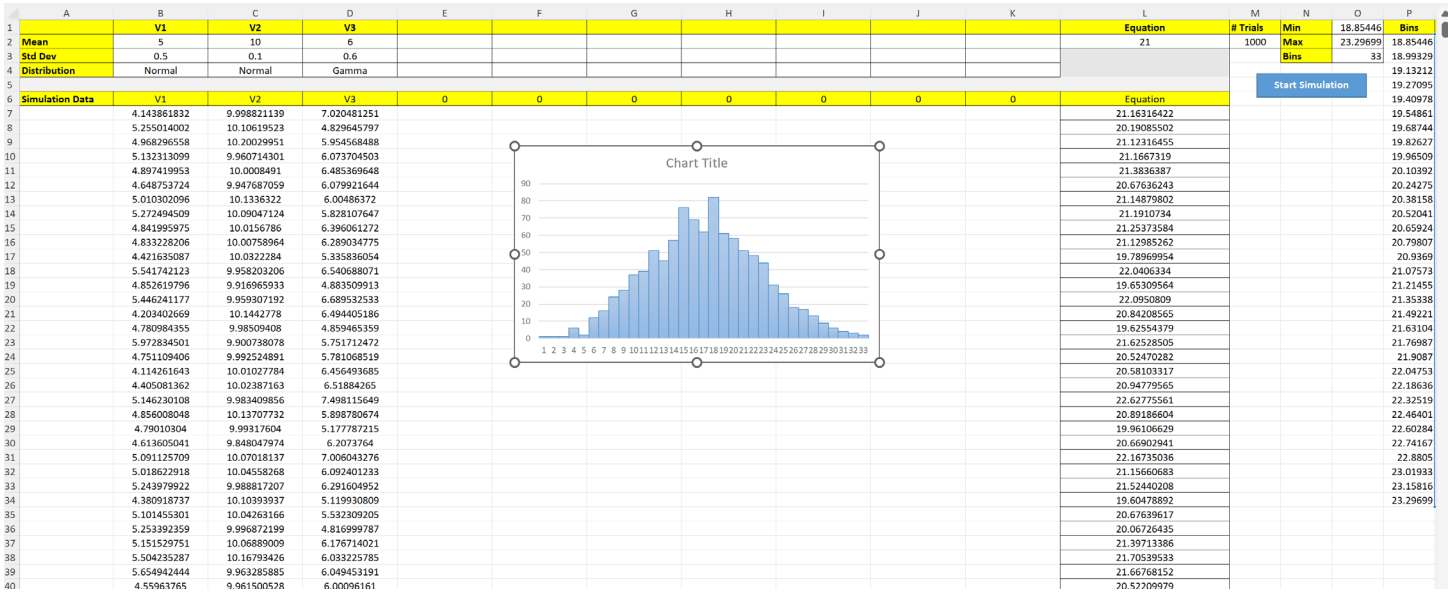
Although our focus here is on product requirements, Monte Carlo can be generally used to:

- Predict Variability In Overall Component Length As A Function Of Parts' Variability:
- Predict Variability In Revenues, Profit, ROI As A Function Of Variability In Sales, Market Conditions, Expenses, etc.
- Predict Reliability Of A Mechanical Device As A Function Of Variability In Stresses Imposed On The Device And The Strength Of Device's Materials
- *In General, Predict Probability Of A Desired Outcome As A Function Of Variability In Inputs (i.e., Dimensions, Force, Profit, Sales, Volumes, Risk)*

A few years ago, we were meeting with our financial advisor. He presented an analysis of our retirement portfolio. One of the outputs was the probability that we would outlive our portfolio (not a good thing!). His financial software used Monte Carlo analysis to make the prediction – including variation in market returns and variation in our longevity!

A Monte Carlo analysis can be performed in a number of ways:

- Excel – we've developed a simple Monte Carlo demonstrator in Excel (see *John's MonteCarlo-Distrs*) that allows you to enter up to 10 variables, assign either Normal, Log-Normal, Beta, Gamma or Uniform distributions, add a function, specify the number of runs and execute the simulation. The spreadsheet generates a histogram of the results:



	V1	V2	V3
Mean	5	10	6
Std Dev	0.5	0.1	0.6
Distribution	Normal	Normal	Gamma

Simulation Data	V1	V2	V3	0	0	0	0	0	0	0	Equation
	4.143861832	9.998821139	7.020481251								21.16316422
	5.255014002	10.10619523	4.829645797								20.19085502
	4.968296558	10.20029951	5.954568488								21.12316455
	5.132313099	9.960714301	6.073704503								21.1667319
	4.897419953	10.0008491	6.485369648								21.3836387
	4.648753724	9.947687059	6.079921644								20.67636243
	5.010302096	10.1336322	6.00486372								21.14879802
	5.272494509	10.09047124	5.828107647								21.1910734
	4.841995975	10.0156786	6.396061272								21.25373584
	4.833228206	10.00758964	6.289034775								21.12985262
	4.421635087	10.0322284	5.335836054								19.78969954
	5.541742123	9.958203206	6.540688071								22.0406334
	4.852619796	9.916965933	4.883509913								19.65309564
	5.446241177	9.959307192	6.689532533								22.0950809
	4.203402669	10.1442778	6.494405186								20.84208565
	4.780984355	9.98509408	4.859465359								19.62554379
	5.972834501	9.900738078	5.751712472								21.62528505
	4.751109406	9.992524891	5.781068519								20.52470282
	4.114261643	10.01027784	6.456493685								20.58103317
	4.405081362	10.02387163	6.51884265								20.94779565
	5.146230108	9.983409856	7.498115649								22.62775561
	4.856008048	10.13707732	5.898780674								20.89186604
	4.79010304	9.99317604	5.177787215								19.96106629
	4.613605041	9.848047974	6.2073764								20.66902941
	5.091125709	10.07018137	7.006043276								22.16735036
	5.018622918	10.04558268	6.092401233								21.15660683
	5.243979922	9.988817207	6.291604952								21.52440208
	4.380918737	10.10393937	5.119930809								19.60478892
	5.101455301	10.04263166	5.532309205								20.67639617
	5.253392359	9.996872199	4.816999787								20.06726435
	5.151529751	10.06889009	6.176714021								21.39713386
	5.504235287	10.16793426	6.033225785								21.70539533
	5.654942444	9.963285885	6.049453191								21.66768152
	4.55963765	9.961500528	6.00096161								20.52209979

# Trials	Min	18.85446	Bins
1000	Max	23.29699	18.85446
	Bins	33	18.99329
			19.13212
			19.27095
			19.40978
			19.54861
			19.68744
			19.82627
			19.96509
			20.10392
			20.24275
			20.38158
			20.52041
			20.65924
			20.79807
			20.9369
			21.07573
			21.21455
			21.35338
			21.49221
			21.63104
			21.76987
			21.9087
			22.04753
			22.18636
			22.32519
			22.46401
			22.60284
			22.74167
			22.8805
			23.01933
			23.15816
			23.29699

- Minitab – Minitab's Calc>Random Number function can be used to generate the input values – one per column. The Calculator function can then be used to calculate the Y values. You'll then have a column of numbers where you can calculate Basic Statistics, view a Histogram and perform Capability analyses if you know the specification limits.

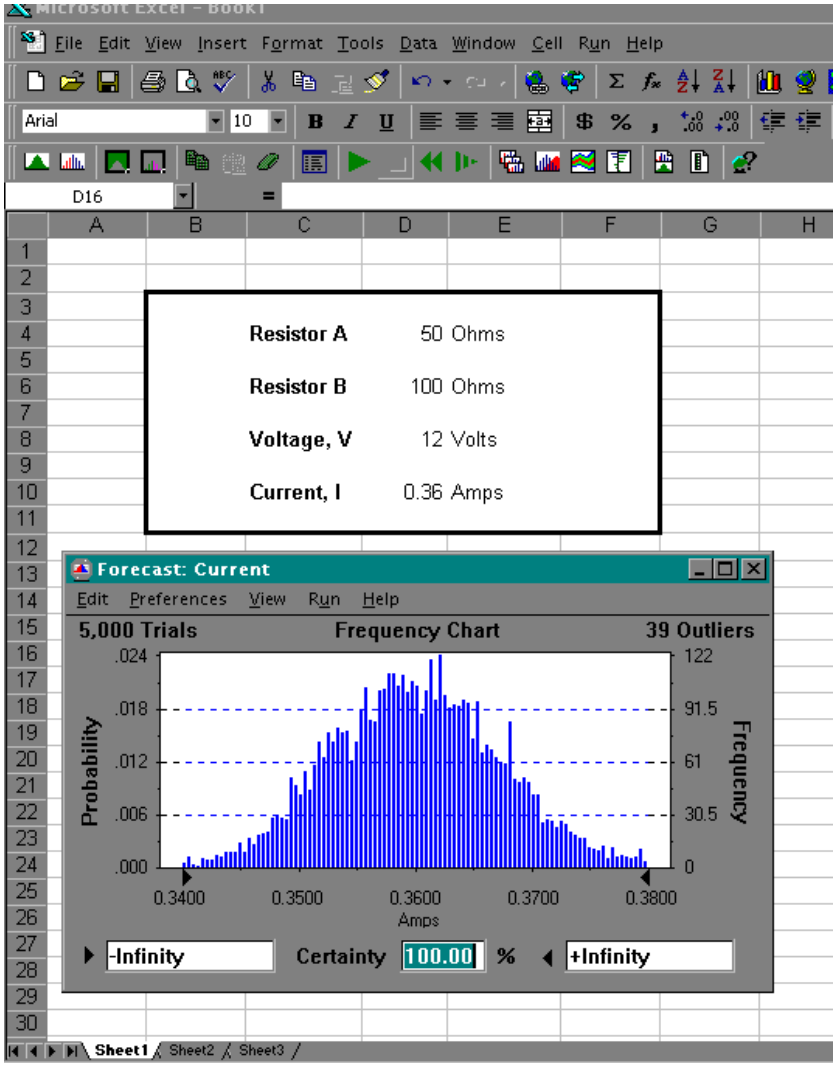
- Monte Carlo Software. There are a variety of Monte Carlo software packages for sale. For example, the screen shot below shows the output of the *Crystal Ball* Monte Carlo simulation package. Crystal Ball is an add-on to Microsoft Excel. It has the advantage that any engineering spreadsheet can be quickly converted to a Monte Carlo simulation. (Crystal Ball – add-on to Microsoft Excel). The example on the next page shows how a circuit with two resistors and a specified (but variable) input voltage combine to generate an output current.

  Here, 5000 trials were run, where each trial produces a calculated current as a function of random variables drawn from the input variables' probability distributions. Statistics are gathered, for example, the standard deviation of the current is predicted to be 0.0074 amps. If the specification limits for current were known, then the capability of the design could be predicted.

  The frequency chart shows the expected variation in the current, *I,* as a function of variation in the resistances and input voltage.

- Quality Loss Function - Genichi Taguchi makes use of a Quality Loss Function to specify both high level and low-level tolerances. Unit 14.4 will discuss his approach.

Crystal Ball Monte Carlo Simulation Example:



Design Scorecards

A product or service development effort will need to manage multiple requirements. Design scorecards are a tool to help the team assess how *capable* their design is predicted to be. The concept is simple – important requirements are listed on the left side of the scorecard, along with their targets and limits.

The team will include *design prediction* activities in their development process, consistent with their progress through design. For example, using the DMEDVI approach, during Explore, the team should attempt to predict the *average* performance of the design. In Design, when parts and process variability information becomes available, the team will predict both the *average* and *variability* expected of their design. To complete the right side of the scorecard, the team will need to perform either a *Root Sum of Squares* analysis for simple requirements, or a *Monte Carlo* analysis for non-linear Y = F(X) relationships. If the values are acceptable, the design can progress. If there are gaps, the team should work to address them. A typical Design Scorecard is shown below:

CTQ Scorecard

Product/System:	**Scarless Sutures**
Project Leader	**C. Miller**
Date	06/30/2000
Revision:	15

CTQ	Measure	Target/ Nominal	Upper Specification Limit	Lower Specification Limit	Sigma Target	Mean	Standard Deviation	Short/ Long Term?	DPU	DPMO	Predicted Sigma
Tensile Strength	Lbs/sq.in. breaking	150		100	6	160	10	L			6
Knot-Holding											
Bioburden	Microorganisms/ft.		0.01		N/A						
Disintegration Time	Days to Absorbtion	7	9	6	4						
Diameter Uniformity	mm	0.00005			4						
Tissue Reaction	Inflamation Rate		0.01%		N/A						
Tissue Pull Force	lbs.	0.5	0.7		5						
Fray Resistance	Frays/suture inch		1		N/A				0.5	500000	1.5

14.2 Conceptual Design

Learning Objectives

- Understand and apply the *static* and *dynamic* concept to design
- Apply creativity tools to new product/service design
- Be able to evaluate and improve conceptual designs

Unit Contents

- Concept Design
- Static vs. Dynamic Designs
- Concept Development
- Pugh Concept Selection Approach

14.2.1 Concept Design

Early in the design process, once the product or service's requirements have been developed, the design team is faced with the selection of the best **concept** to be developed further into a detailed design capable of being produced and sold successfully. Often, this important design phase is either overlooked or not given sufficient attention. Note that concept selection can occur at all levels in a design including the overall product or service concept, system or subsystem concept, equipment and component concepts.

Two situations are frequently observed in this phase of design. First, design teams and individual team members often come to the design with pre-conceived notions of what the new design should incorporate. Perhaps their experience biases them towards concepts that have worked in the past, or there are political pressures to employ a given concept. The use of computers in design offers many advantages, but they can suppress the creativity demanded of the design team here.

Second, even if the design team manages to develop several alternatives, concept selection is either done subjectively, or through overly complicated numerically based evaluation techniques (e.g. such as Kepner-Tregoe matrices). While it seems natural to want to quantify the "goodness" of a design alternative, design teams have been observed to select sub-optimal designs based on just the "numbers."

This unit focuses on conceptual design – both from the identification of design concepts as well as their evaluation. The work of the late Dr. Stuart Pugh forms the basis for the discussion and tools of this unit.

14.2.2 Static vs. Dynamic Designs

In his book, *Wonderful Life*, Stephen Jay Gould describes the Cambrian explosion of diverse animal and plant phyla that occurred over 500 million years ago on Earth. During that time, *diversity* was the theme – animals and plants evolved of many different structures and types. We would characterize the *design concepts* associated with these creatures as *dynamic.* Anything went – flat, round, cone-shape, cylinder-shape, worm-like, spiked, legged, segmented and other features were combined almost willy-nilly into many structural variations.

Today, although evolution continues, the occurrence of mass extinctions has resulted in variations around a few basic themes – mammals generally have a head connected to a body with four appendages; insects have eight legs attached to a three-segment body, etc. We would term such *design concepts* as *static.*

Products and services may also be categorized as *dynamic* or *static* in their evolution. Despite a wide variety of types of automobile, the basic design concept today (and expected for the next 20 – 50 years) includes a body, engine, four wheels, a passenger/luggage compartments. Similarly, although it represents a much younger concept, today's personal computer represents a *static* concept. Central processor, memory storage, keyboard, visual display and software (operating system and applications) are the key elements of this design concept.

Other products and services are *dynamic* in nature. Note that, although the basic personal computer seems to be *static* in design concept, many of its components are in a *dynamic* state. Visual displays and memory storage are *dynamic* in concept. Recall, for example, the introduction of the memory chip, the floppy disk, the hard drive, optical storage, CD-ROM and the DVD, with most storage today being solid-state. Perusing the computer trade literature will often reveal stories of new ways of storing the 1's and 0's we need for our machines being identified and developed.

Note that the emergence of a *dominant* design may lead to a concept becoming static. For example, although some different design concepts emerged early in the history of naval nuclear propulsion (e.g. sodium cooled reactors), Admiral Rickover quickly settled on the pressurized water reactor concept – a reactor system which transferred the heat of fission through steam generators to make steam to drive the propulsion turbine. Through his influence, this design concept is still employed today, and was also transferred to the commercial nuclear power industry via the prototype reactor at Shippingport, PA. Many competing design concepts were never given the chance to develop.

Relation to Product/Service Design

Static Concept Case

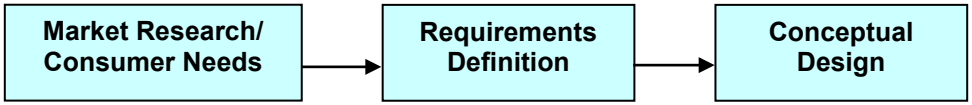
In the *static* case, the basic design concept is known at the beginning of the product or service development cycle. The product or service requirements can then be developed with this *a priori* knowledge. Using the nuclear propulsion example above, since the basic concept is known, requirements can be defined for the overall system as well as the subsystems (e.g. reactor system, turbine system, feedwater system, etc.). The relationship between requirements and conceptual design appears below:



There is an interesting relationship between computers and design that appears here. Many Computer-Aided-Design software packages have been developed to support mechanical, electrical and civil engineering design. Pugh notes that these packages tend to be most useful when the design concept has evolved to a static state. The effort of the design is to take an established concept and develop it more efficiently or effectively than previous or competitive designs. Here, the accumulated design experience contained in the software package is useful to the design team.

Dynamic Case

On the other hand, if the product or service is dynamic, then the design team should open themselves to truly different design concepts. Requirements should be defined for the entire system, but then the design process should include a vigorous search for different concepts. The early design sequence appears as below:



The opposite case may be stated for the use of CAD systems here. The CAD system is generally “tuned” for a given design concept, use in this situation can be detrimental to the design effort. Pugh relates the case of a design team developing a new dump truck. A new concept was developed for the truck bed lifting device – one design team member spent days trying to adapt an existing CAD system to model the new concept. With schedule pressure mounting, another team member created a model from cardboard in only a few hours.

Design Dangers – Dynamic vs. Static Approach

For many designers, acceptance that their product or service has reached a *static* state is the easy path to follow. The design process then becomes comfortable, predictable and manageable. The danger, of course, lies in what is happening in the competition's design shop. The harder, more uncertain path lies in the assumption of a *dynamic* design.

Pugh notes that, even when conventional "wisdom" has categorized the product or service as static, it is *essential* that the design team treat the design as *dynamic.* He has observed many cases where through iteration of alternative conceptual designs, a fundamentally different concept has emerged.

14.2.3 Concept Development

There are a variety of methods employed to generate conceptual design ideas a number of which are presented in ***Unit 3.2 – Idea Generation and Decision Making***. Interestingly, although *brainstorming* is often the first suggested, Pugh observed that this is one of the least effective methods for design concept generation. The methods he found most useful in developing conceptual designs include:

- Analogy (*See Forced Analogy*)
- Attribute Listing (*See Attribute Listing*)
- Inversion (*See Problem Reversal*)

Two methods that are specific to design, TRIZ and Parametric Analysis are presented below.

Theory of Inventive Problem Solving (TRIZ)

We tend to think of the inventor as somebody who is born with a knack for developing new solutions to problems. Beginning in 1946, a Russian named Genrich Altshuller developed a method of teaching scientists, engineers how to be inventive and creative. Under the sponsorship of the Soviet government, Altshuller conducted a patent search for new technologies that might have military or other use to the Soviets. As he performed this search, he noticed a pattern in the invention literature. He eventually identified forty "inventive principles;" a summary of these appears below.

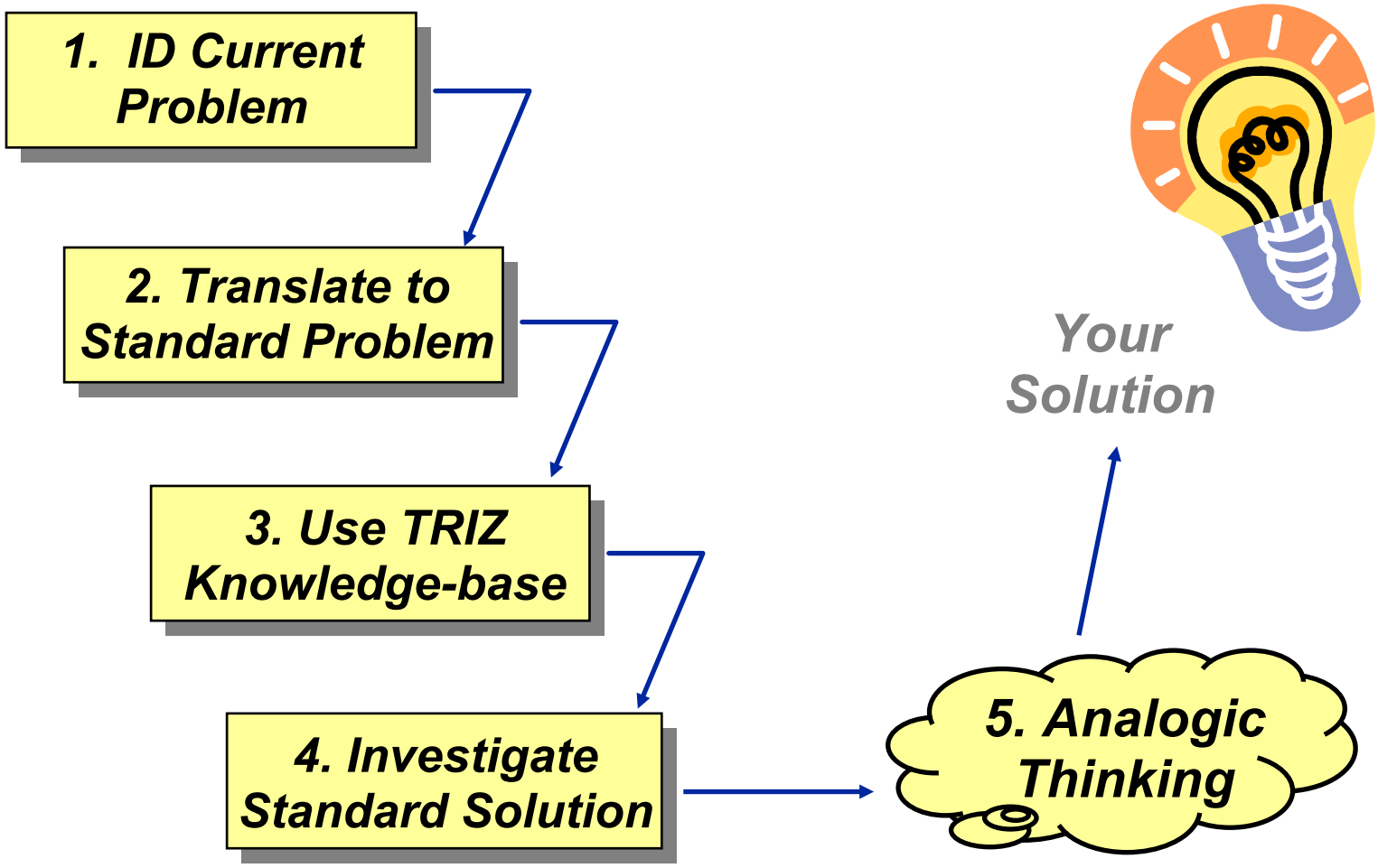
How to Get a Transformer Off a Pedestal (Without a Crane) - To illustrate his technique, Altshuller describes a problem he encountered in his youth. Workers were trying to remove an electric transformer from a pedestal about 3 feet high. They did not have access to a crane and could not figure out how to perform the task. A *bookkeeper* living in his apartment building caused the workers to build a pedestal of ice next to the transformer. The transformer was nudged to the ice pedestal and, over several days, the ice melted, lowering the transformer! Altshuller's insight was that the ice had been made to do something for which it was not intended.

His work led to not only the identification of the inventive principles, but a method (Algorithm for Inventive Problem Solving – given the acronym **ARIZ** (from the Russian)). TRIZ practitioners are trained to think about problems using this algorithm – some say that it takes almost a year to "rewire" your brain to consistently address problems using the ARIZ algorithm. The essence of the ARIZ approach is to understand the "real-world" problem but then convert it into a "standard" problem that can be addressed via the inventive principles.

For example, if you are trying to make a car door light and strong, ARIZ thinking would convert this into the standard problem of improving the *strength of a stationary object* without increasing its *weight.* The inventive principles that have been employed in the past to address such a problem (identified via a "Contradiction Matrix) include 1) Segmentation, 26) Copying, 27) Disposable: Cheap, Short-Living Object, and 40) Composite Materials.

Notice that the inventive principles don't solve your creative problem, but they help you identify *how* to solve the problem. We met a TRIZ enthusiast a while back. In his career as a chemical engineer, he applied for about 30 patents in his first 10 years. After "embracing" TRIZ, his patent volume rose to 30 in two years!

Continuing the car door example, you now employ "analogic thinking" to determine possible ways of applying the principles. You may decide that the use of composite materials, such as fiberglass reinforced plastic or carbon fiber composites are practical (principle 40 – Composite Materials), or you may explore ways of adding sheet metal strips to the car door panel (principle 1 – Segmentation) for your solution.



ARIZ – Algorithm for Inventive Problem-Solving

TRIZ Inventive Principles

Principle	Sub-Principle	• Example(s)
1. Segmentation	Divide an object into independent parts.	• *Replace mainframe computer by personal computers.* • *Replace a large truck by a truck and trailer.* • *Use a work breakdown structure for a large project.*
	Make an object easy to disassemble.	• *Modular furniture* • *Quick disconnect joints in plumbing*
	Increase the degree of fragmentation or segmentation.	• *Replace solid shades with Venetian blinds.* • *Use powdered welding metal instead of foil or rod to get better penetration of the joint.*
2. Taking out	Separate an interfering part or property from an object, or single out the only necessary part (or property) of an object.	• *Locate a noisy compressor outside the building where compressed air is used.* • *Use fiber optics or a light pipe to separate the hot light source from the location where light is needed.* • *Use the sound of a barking dog, without the dog, as a burglar alarm.*
3. Local quality	Change an object's structure from uniform to non-uniform; change an external environment (or external influence) from uniform to non-uniform.	• *Use a temperature, density, or pressure gradient instead of constant temperature, density or pressure.*
	Make each part of an object function in conditions most suitable for its operation	• *Lunch box with special compartments for hot and cold solid foods and for liquids*
	Make each part of an object fulfill a different and useful function.	• *Pencil with eraser* • *Hammer with nail puller* • *Multi-function tool that scales fish, acts as a pliers, a wire stripper, a flat-blade screwdriver, a Phillips screwdriver, manicure set, etc.*
4. Asymmetry	Change the shape of an object from symmetrical to asymmetrical.	• *Asymmetrical mixing vessels or asymmetrical vanes in symmetrical vessels improve mixing (cement trucks, cake mixers, blenders).* • *Put a flat spot on a cylindrical shaft to attach a knob securely.*
	If an object is asymmetrical, increase its degree of asymmetry.	• *Change from circular O-rings to oval cross-section to specialized shapes to improve sealing.* • *Use astigmatic optics to merge colors.*
5. Merging	Bring closer together (or merge) identical or similar objects; assemble identical or similar parts to perform parallel operations.	• *Personal computers in a network* • *Thousands of microprocessors in a parallel processor computer* • *Vanes in a ventilation system* • *Electronic chips mounted on both sides of a circuit board or subassembly*

Principle	*Sub-Principle*	• *Example(s)*
	Make operations contiguous or parallel; bring them together in time.	• *Link slats together in Venetian or vertical blinds.* • *Medical diagnostic instruments that analyze multiple blood parameters simultaneously* • *Mulching lawnmower*
6. Universality	Make a part or object perform multiple functions; eliminate the need for other parts	• *Handle of a toothbrush contains toothpaste* • *Child's car safety seat converts to a stroller* • *Mulching lawnmower (Yes, it demonstrates both Principles 5 and 6, Merging and Universality.)* • *Team leader acts as recorder and timekeeper.* • *CCD (Charge coupled device) with micro-lenses formed on the surface*
7. "Nested doll"	Place one object inside another; place each object, in turn, inside the other.	• *Measuring cups or spoons* • *Russian dolls* • *Portable audio system (microphone fits inside transmitter, which fits inside amplifier case)*
	Make one part pass through a cavity in the other.	• *Extending radio antenna* • *Extending pointer* • *Zoom lens* • *Seat belt retraction mechanism* • *Retractable aircraft landing gear stow inside the fuselage (also demonstrates Principle 15, Dynamism).*
8. Anti-weight	To compensate for the weight of an object, merge it with other objects that provide lift.	• *Inject foaming agent into a bundle of logs, to make it float better.* • *Use helium balloon to support advertising signs.*
	To compensate for the weight of an object, make it interact with the environment (e.g. use aerodynamic, hydrodynamic, buoyancy and other forces).	• *Aircraft wing shape reduces air density above the wing, increases density below wing, to create lift. (This also demonstrates Principle 4, Asymmetry.)* • *Vortex strips improve lift of aircraft wings.* • *Hydrofoils lift ship out of the water to reduce drag.*
9. Preliminary anti-action	If it will be necessary to do an action with both harmful and useful effects, this action should be replaced with anti-actions to control harmful effects.	• *Buffer a solution to prevent harm from extremes of pH.*
	Create beforehand stresses in an object that will oppose known undesirable working stresses later on.	• *Pre-stress rebar before pouring concrete.* • *Masking anything before harmful exposure: Use a lead apron on parts of the body not being exposed to X-rays. Use masking tape to protect the part of an object not being painted*
10. Preliminary action	Perform, before it is needed, the required change of an object (either fully or partially).	• *Pre-pasted wallpaper* • *Sterilize all instruments needed for a surgical procedure on a sealed tray.*

Principle	*Sub-Principle*	• *Example(s)*
	Pre-arrange objects such that they can come into action from the most convenient place and without losing time for their delivery.	• *Kanban arrangements in a Just-In-Time factory* • *Flexible manufacturing cell*
11. Beforehand cushioning	Prepare emergency means beforehand to compensate for the relatively low reliability of an object.	• *Magnetic strip on photographic film that directs the developer to compensate for poor exposure* • *Back-up parachute* • *Alternate air system for aircraft instruments*
12. Equipotentiality	In a potential field, limit position changes (e.g. change operating conditions to eliminate the need to raise or lower objects in a gravity field).	• *Spring loaded parts delivery system in a factory* • *Locks in a channel between 2 bodies of water (Panama Canal)* • *"Skillets" in an automobile plant that bring all tools to the right position (also demonstrates Principle 10, Preliminary Action)*
13. 'The other way round'	Invert the action(s) used to solve the problem (e.g. instead of cooling an object, heat it).	• *To loosen stuck parts, cool the inner part instead of heating the outer part.* • *Bring the mountain to Mohammed, instead of bringing Mohammed to the mountain.*
	Make movable parts (or the external environment) fixed, and fixed parts movable).	• *Rotate the part instead of the tool.* • *Moving sidewalk with standing people* • *Treadmill (for walking or running in place)*
	Turn the object (or process) 'upside down'.	• *Turn an assembly upside down to insert fasteners (especially screws).* • *Empty grain from containers (ship or railroad) by inverting them.*
14. Spheroidality - Curvature	Instead of using rectilinear parts, surfaces, or forms, use curvilinear ones; move from flat surfaces to spherical ones; from parts shaped as a cube (parallelepiped) to ball-shaped structures.	• *Use arches and domes for strength in architecture.*
	Use rollers, balls, spirals, and domes.	• *Spiral gear (Nautilus) produces continuous resistance for weight lifting.* • *Ball point and roller point pens for smooth ink distribution*
	Go from linear to rotary motion; use centrifugal forces.	• *Produce linear motion of the cursor on the computer screen using a mouse or a trackball.* • *Replace wringing clothes to remove water with spinning clothes in a washing machine.* • *Use spherical casters instead of cylindrical wheels to move furniture.*
15. Dynamics	Allow (or design) the characteristics of an object, external environment, or process to change to be optimal or to find an optimal operating condition.	• *Adjustable steering wheel (or seat, or back support, or mirror position...)*
	Divide an object into parts capable of movement relative to each other.	• *The "butterfly" computer keyboard, (also demonstrates Principle 7, "Nested doll".)*

Principle	Sub-Principle	• Example(s)
	If an object (or process) is rigid or inflexible, make it movable or adaptive.	• *The flexible boroscope for examining engines* • *The flexible sigmoidoscope, for medical examination*
16. Partial or excessive actions	If 100 percent of an object is hard to achieve using a given solution method then, by using 'slightly less' or 'slightly more' of the same method, the problem may be considerably easier to solve.	• *Over spray when painting, then remove excess. (Or, use a stencil--this is an application of Principle 3, Local Quality and Principle 9, Preliminary anti-action).* • *Fill, then "top off" when filling the gas tank of your car.*
17. Another dimension	To move an object in two- or three-dimensional space.	• *Infrared computer mouse moves in space, instead of on a surface, for presentations.* • *Five-axis cutting tool can be positioned where needed.*
	Use a multi-story arrangement of objects instead of a single-story arrangement.	• *Cassette with 6 CD's to increase music time and variety* • *Electronic chips on both sides of a printed circuit board* • *Employees "disappear" from the customers in a theme park, descend into a tunnel, and walk to their next assignment, where they return to the surface and "magically" reappear.*
	Tilt or re-orient the object, lay it on its side	• *Dump truck*
	Use 'another side' of a given area.	• *Stack microelectronic hybrid circuits to improve density.*
18. Mechanical vibration	Cause an object to oscillate or vibrate.	• *Electric carving knife with vibrating blades*
	Increase its frequency (even up to the ultrasonic).	• *Distribute powder with vibration.*
	Use an object's resonant frequency.	• *Destroy gallstones or kidney stones using ultrasonic resonance.*
	Use piezoelectric vibrators instead of mechanical ones.	• *Quartz crystal oscillations drive high accuracy clocks.*
	Use combined ultrasonic and electromagnetic field oscillations.	• *Mixing alloys in an induction furnace*
19. Periodic action	Instead of continuous action, use periodic or pulsating actions.	• *Hitting something repeatedly with a hammer* • *Replace a continuous siren with a pulsed sound.*
	If an action is already periodic, change the periodic magnitude or frequency.	• *Use Frequency Modulation to convey information, instead of Morse code.* • *Replace a continuous siren with sound that changes amplitude and frequency.*
	Use pauses between impulses to perform a different action.	• *In cardio-pulmonary respiration (CPR) breathe after every 5 chest compressions.*
20. Continuity of useful action	Carry on work continuously; make all parts of an object work at full load, all the time.	• *Flywheel (or hydraulic system) stores energy when a vehicle stops, so the motor can keep running at optimum power.* • *Run the bottleneck operations in a factory continuously, to reach the optimum pace. (From theory of constraints, or takt time operations)*

Principle	*Sub-Principle*	• *Example(s)*
	Eliminate all idle or intermittent actions or work.	• *Print during the return of a printer carriage--dot matrix printer, daisy wheel printers, inkjet printers.*
21. Skipping	Conduct a process, or certain stages (e.g. destructible, harmful or hazardous operations) at high speed.	• *Use a high-speed dentist's drill to avoid heating tissue.* • *Cut plastic faster than heat can propagate in the material, to avoid deforming the shape.* • *Run across hot sand on the beach to avoid burning bare feet.*
22. "Blessing in disguise"	Use harmful factors (particularly, harmful effects of the environment or surroundings) to achieve a positive effect.	• *Use waste heat to generate electric power.* • *Recycle waste (scrap) material from one process as raw materials for another.*
	Eliminate the primary harmful action by adding it to another harmful action to resolve the problem.	• *Add a buffering material to a corrosive solution.* • *Use a helium-oxygen mix for diving, to eliminate both nitrogen narcosis and oxygen poisoning from air and other nitrox mixes.*
	Amplify a harmful factor to such a degree that it is no longer harmful.	• *Use a backfire to eliminate the fuel from a forest fire.*
23. Feedback	Introduce feedback (referring back, cross-checking) to improve a process or action.	• *Automatic volume control in audio circuits* • *Signal from gyrocompass is used to control simple aircraft autopilots.* • *Statistical Process Control (SPC) -- Measurements are used to decide when to modify a process. (Not all feedback systems are automated!)* • *Budgets --Measurements are used to decide when to modify a process.*
	If feedback is already used, change its magnitude or influence.	• *Change sensitivity of an autopilot when within 5 miles of an airport.* • *Change sensitivity of a thermostat when cooling vs. heating, since it uses energy less efficiently when cooling.* • *Change a management measure from budget variance to customer satisfaction.*
24. 'Intermediary'	Use an intermediary carrier article or intermediary process.	• *Carpenter's nail set, used between the hammer and the nail*
	Merge one object temporarily with another (which can be easily removed).	• *Pot holder to carry hot dishes to the table*

Principle	Sub-Principle	• Example(s)
25. Self-service	Make an object serve itself by performing auxiliary helpful functions	• *A soda fountain pump that runs on the pressure of the carbon dioxide that is used to "fizz" the drinks. This assures that drinks will not be flat, and eliminates the need for sensors.* • *Halogen lamps regenerate the filament during use--evaporated material is redeposited.* • *To weld steel to aluminum, create an interface from alternating thin strips of the 2 materials. Cold weld the surface into a single unit with steel on one face and copper on the other, then use normal welding techniques to attach the steel object to the interface, and the interface to the aluminum. (This concept also has elements of Principle 24, Intermediary, and Principle 4, Asymmetry.)*
	Use waste resources, energy, or substances.	• *Use heat from a process to generate electricity: "Co-generation".* • *Use animal waste as fertilizer.* • *Use food and lawn waste to create compost.*
26. Copying	Instead of an unavailable, expensive, fragile object, use simpler and inexpensive copies.	• *Virtual reality via computer instead of an expensive vacation* • *Listen to an audiotape instead of attending a seminar.*
	Replace an object, or process with optical copies.	• *Do surveying from space photographs instead of on the ground.* • *Measure an object by measuring the photograph.* • *Make sonograms to evaluate the health of a fetus, instead of risking damage by direct testing.*
	If visible optical copies are already used, move to infrared or ultraviolet copies.	• *Make images in infrared to detect heat sources, such as diseases in crops, or intruders in a security system.*
27. Cheap short-living objects	Replace an inexpensive object with a multiple of inexpensive objects, comprising certain qualities (such as service life, for instance).	• *Use disposable paper objects to avoid the cost of cleaning and storing durable objects. Plastic cups in motels, disposable diapers, many kinds of medical supplies.*
28. Mechanics substitution	Replace a mechanical means with a sensory (optical, acoustic, taste or smell) means.	• *Replace a physical fence to confine a dog or cat with an acoustic "fence" (signal audible to the animal).* • *Use a bad smelling compound in natural gas to alert users to leakage, instead of a mechanical or electrical sensor.*
	Use electric, magnetic and electromagnetic fields to interact with the object.	• *To mix 2 powders, electrostatically charge one positive and the other negative. Either use fields to direct them, or mix them mechanically and let their acquired fields cause the grains of powder to pair up.*
	Change from static to movable fields, from unstructured fields to those having structure.	• *Early communications used omnidirectional broadcasting. We now use antennas with very detailed structure of the pattern of radiation.*
	Use fields in conjunction with field-activated (e.g. ferromagnetic) particles.	• *Heat a substance containing ferromagnetic material by using varying magnetic field. When the temperature exceeds the Curie point, the material becomes paramagnetic, and no longer absorbs heat.*

Principle	*Sub-Principle*	• *Example(s)*
29. Pneumatics and hydraulics	Use gas and liquid parts of an object instead of solid parts (e.g. inflatable, filled with liquids, air cushion, hydrostatic, hydro-reactive).	• *Comfortable shoe sole inserts filled with gel* • *Store energy from decelerating a vehicle in a hydraulic system, and then use the stored energy to accelerate later.*
30. Flexible shells and thin films	Use flexible shells and thin films instead of three-dimensional structures	• *Use inflatable (thin film) structures as winter covers on tennis courts.*
	Isolate the object from the external environment using flexible shells and thin films.	• *Float a film of bipolar material (one end hydrophilic, one end hydrophobic) on a reservoir to limit evaporation.*
31. Porous materials	Make an object porous or add porous elements (inserts, coatings, etc.).	• *Drill holes in a structure to reduce the weight.*
	If an object is already porous, use the pores to introduce a useful substance or function.	• *Use a porous metal mesh to wick excess solder away from a joint.* • *Store hydrogen in the pores of a palladium sponge. (Fuel "tank" for the hydrogen car--much safer than storing hydrogen gas)*
32. Color changes	Change the color of an object or its external environment.	• *Use safe lights in a photographic darkroom.*
	Change the transparency of an object or its external environment.	• *Use photolithography to change transparent material to a solid mask for semiconductor processing. Similarly, change mask material from transparent to opaque for silkscreen processing.*
33. Homogeneity	Make objects interacting with a given object of the same material (or material with identical properties).	• *Make the container out of the same material as the contents, to reduce chemical reactions.* • *Make a diamond-cutting tool out of diamonds.*
34. Discarding and recovering	Make portions of an object that have fulfilled their functions go away (discard by dissolving, evaporating, etc.) or modify these directly during operation.	• *Use a dissolving capsule for medicine.* • *Sprinkle water on cornstarch-based packaging and watch it reduce its volume by more than 1000X!* • *Ice structures: use water ice or carbon dioxide (dry ice) to make a template for a rammed earth structure, such as a temporary dam. Fill with earth, then, let the ice melt or sublime to leave the final structure.*
	Conversely, restore consumable parts of an object directly in operation.	• *Self-sharpening lawn mower blades* • *Automobile engines that give themselves a "tune up" while running (the ones that say "100,000 miles between tune ups")*
35. Parameter changes	Change an object's physical state (e.g. to a gas, liquid, or solid.)	• *Freeze the liquid centers of filled candies, and then dip in melted chocolate, instead of handling the messy, gooey, hot liquid.* • *Transport oxygen or nitrogen or petroleum gas as a liquid, instead of a gas, to reduce volume.*
	Change the concentration or consistency.	• *Liquid hand soap is concentrated and more viscous than bar soap at the point of use, making it easier to dispense in the correct amount and more sanitary when shared by several people.*

Principle	Sub-Principle	• Example(s)
	Change the degree of flexibility.	• *Use adjustable dampers to reduce the noise of parts falling into a container by restricting the motion of the walls of the container.* • *Vulcanize rubber to change its flexibility and durability.*
	Change the temperature.	• *Raise the temperature above the Curie point to change a ferromagnetic substance to a paramagnetic substance.* • *Raise the temperature of food to cook it. (Changes taste, aroma, texture, chemical properties, etc.)* • *Lower the temperature of medical specimens to preserve them for later analysis.*
36. Phase transitions	Use phenomena occurring during phase transitions (e.g. volume changes, loss or absorption of heat, etc.).	• *Water expands when frozen, unlike most other liquids. Hannibal is reputed to have used this when marching on Rome a few thousand years ago. Large rocks blocked passages in the Alps. He poured water on them at night. The overnight cold froze the water, and the expansion split the rocks into small pieces that could be pushed aside.* • *Heat pumps use the heat of vaporization and heat of condensation of a closed thermodynamic cycle to do useful work.*
37. Thermal expansion	Use thermal expansion (or contraction) of materials.	• *Fit a tight joint together by cooling the inner part to contract, heating the outer part to expand, putting the joint together, and returning to equilibrium*
	If thermal expansion is being used, use multiple materials with different coefficients of thermal expansion.	• *The basic leaf spring thermostat: (2 metals with different coefficients of expansion are linked so that it bends one way when warmer than nominal and the opposite way when cooler.)*
38. Strong oxidants	Replace common air with oxygen-enriched air.	• *Scuba diving with Nitrox or other non-air mixtures for extended endurance*
	Replace enriched air with pure oxygen.	• *Cut at a higher temperature using an oxy-acetylene torch.* • *Treat wounds in a high-pressure oxygen environment to kill anaerobic bacteria and aid healing.*
	Expose air or oxygen to ionizing radiation.	• *Nuclear reactor radiation creates hydrogen and oxygen from water; helps control corrosion in cooling system.*
	Use ionized oxygen.	• *Ionize air to trap pollutants in an air cleaner.*
	Replace ozonized (or ionized) oxygen with ozone.	• *Speed up chemical reactions by ionizing the gas before use.*
39. Inert atmosphere	Replace a normal environment with an inert one.	• *Prevent degradation of a hot metal filament by using an argon atmosphere.*
	Add neutral parts, or inert additives to an object.	• *Increase the volume of powdered detergent by adding inert ingredients. This makes it easier to measure with conventional tools.*
40. Composite	Change from uniform to composite (multiple)	• *Composite epoxy resin/carbon fiber golf club shafts are lighter, stronger,*

Principle	*Sub-Principle*	• *Example(s)*
materials	materials.	*and more flexible than metal. Same for airplane parts.* • *Fiberglass surfboards are lighter and more controllable and easier to form into a variety of shapes than wooden ones.*

See *And Suddenly the Inventor Appeared* by G. Altshuller for a fascinating introduction to the TRIZ approach.

Parametric Analysis

Benchmarking (see Unit 14.3) is an important input to concept identification. Pugh recommends that analyses of many competitive products across a variety of industries be done. These analyses should, of course, include cost of product or service, market share of product or service, consumer reaction (e.g. satisfaction) with product or service. To provide useful input to the design team, though, additional data about the performance and/or features of competitive products and services should be collected. Pugh regards this additional information as critical to developing the product or service requirements definition – market research generally not providing all-important information to the designers.

Parametric Analysis

Parametric Analysis is the process of seeking out relationships between parameters for a particular product area under consideration. "Catalog" data can be the starting point for this analysis – the design team first identifies parameters of interest, then they determine parameter values from competitive products, followed by a free-form cross-plotting search to find overall and "interesting" relationships between parameters.

Parametric Analysis Process

1. Gather as much information as possible on your own and your competitions' products/services (including products/services from outside your industry).

2. *Using a free-form approach,* cross-plot the data (one parameter against another) and look for patterns and relationships between parameters.

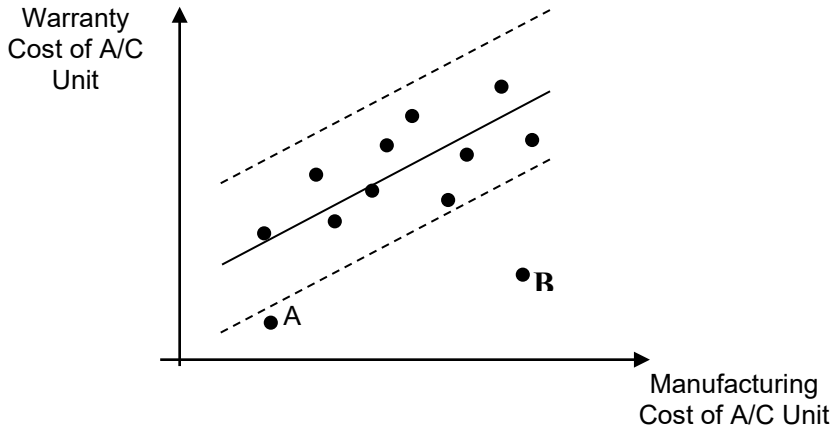
3. Start with the logical relationships (e.g. car weight and fuel consumption, air conditioner cooling capacity and energy consumption), but be willing to explore other, unusual relationships – this is generally where the "Ah-ha's" will be found.

4. Make sure that it's apples to apples you are comparing.

5. Look for patterns in the cross-plots and look for exceptions to the patterns. Regression analysis (including confidence bounds on the plots) can help identify both patterns and outliers/exceptions.

6. Draw conclusions from the plots – Ask "Why" for the exceptions noted.

Example Cross-Plot for Air-Conditioning Units:



14.2.4 Pugh Concept Selection Approach

One of Dr. Pugh's best-known contributions to design is his approach to evaluation and selection of the "best" design concept. Many evaluation techniques are available that allow the comparison of multiple alternatives against multiple criteria. For example, many companies have employed the Kepner-Tregoe (KT) matrix method to evaluate potential bidders in purchasing processes. While methods such as KT have their place, Dr. Pugh's experience identified two important disadvantages:

- **Numerically Based Decision Process** – Engineers tend to gravitate to methods where numbers are employed to make decisions. While this is a natural tendency, at this point in design numbers can often be misleading. Judgement based on a numerical rating can and has often led to a sub-optimal concept selection.

- **Either/Or vs. Both/And Thinking** – Methods such as KT analysis are designed to force an "either/or" approach – of all the alternatives, which is the best? At this point in the design process, Pugh found that the concept of *controlled convergence* – iterating to identify a superior design concept that meets or exceeds the product/service specification – to be a more effective approach.

Pugh Concept Selection Process Steps:

Note that these concept evaluations should be done using a large black or white board, so all can see the work-in-progress:

Phase One

1. Establish many embryonic solutions – sketch these to the same level of detail.

2. Establish the criteria for evaluation, weight if appropriate (using CTQs and others from the product specification, if there are many criteria, the evaluation may start with the 10 – 15 most critical criteria).

3. Create the evaluation matrix – concept sketches are concepts, criteria are rows.

4. Choose a datum for comparison – if a design exists, include this as the datum.

5. Evaluate the concepts against the datum:

+	Better than, less costly than, easier than, more reliable than
-	Worse than, more costly than, harder than, less reliable than
S	Same as datum

6. Assess individual concept scores. Look for strong concepts.

7. Rerun the matrix as follows: If one concept is strong, change the datum and repeat. If this is still strong, change the datum to the strong concept and rerun – reassess the result. If all appear to be uniformly strong (an unusual situation), change the datum and rerun the evaluation. If there is still no strong concept(s), one of two things may occur: a) the criteria may be ambiguous (leading to different interpretations by the team) or one criteria incorporates others, or b) one or more concepts are subsets of the others and the matrix can't help distinguish among them.

8. If additional concepts emerge, rerun the matrix.

Phase Two

Develop the strongest concepts emerging from Phase One to a higher level of detail and reevaluate using the same approach as described above.

Notes:
1. The matrix does not make the decision – the participants do.
2. Concept selection can occur at all levels of design – overall product/service, system, equipment, component, part
3. The iterative nature of Pugh Concept Selection is one of its key strengths.

Pugh Concept Evaluation Matrix Format

	Concept							
Criteria	1	2	3	4	5	6	7	8
	D							
	A							
	T							
	U							
	M							
Sum of Positives								
Sum of Negatives								
Sum of Sames								

Example Pugh Evaluation Matrix for Course Delivery Design Concepts – In this case, a body of quality skills was being delivered in a traditional classroom style. A team identified and evaluated three additional delivery concepts. Eventually, a CD-ROM concept was selected that included audio/visual instruction, plus the traditional classroom binders:

Criteria	**Rating**	**A- Current Classroom (Datum)**	**B - CD-ROM Only**	**C – CD-ROM Plus Classroom**	**D - Self-Paced Training Manual**
Quality (Learning Effectiveness)		***D***			
Builds Knowledge in Subject	5	***A***	S	S	-
Builds Skill in Subject	5	***T***	-	S	S
Allows Learning Application to Own Project (Immediate)	3	***U***	S	+	-
Compatible with Different Learning Styles (e.g. via Multiple Media)	3	***M***	S	S	-
Stimulates Learner Interest and Engagement:					
· Effective Presentation (e.g. Live, Text, Graphics, Video)	3		S	S	-
· Ability to Interact during Learning	3		-	S	-
· Ability to Practice Knowledge and Skills	4		-	S	-
· Ability to Ask Questions/Get Answers during Learning	4		-	S	-
· Availability of Expert Advice	3		-	S	-
· Delivery Matches Participant Expertise	4		S	S	S
Ability to Simulate Group Work	2		-	S	-
Evaluation of Knowledge and Skills	3		S	S	S
Learner Self-Evaluation	4		S	S	+
Consistent Quality of Instruction	4		+	+	+
Use as Reference Tool Throughout Project	4		+	+	+
Delivery (Access, Time to Obtain K&S)					
Learning Efficiency (Time to Learn)	5		+	-	+
Search Capability	4		+	+	S
Self-Paced Learning Opportunity	4		+	+	+
Hardware/Software Platform Compatibility	4		S	S	+
Cost of Service					
Course Costs	5		+	S	+
Travel and Living	4		+	S	+
Time Away from Work	3		+	S	+
Sum of Positives		0	8	5	9
Sum of Negatives		0	6	1	9
Sum of Sames		22	8	16	4
Weighted Sum of Positives		0	33	19	37
Weighted Sum of Negatives		0	21	5	30

14.3 Benchmarking

Learning Objectives

- Understand the purpose and process of benchmarking

Unit Contents

- Benchmarking Purpose and Definition
- Benchmarking Method

14.3.1 Benchmarking Purpose and Definition

Since the dawn of time, people have learned how to do things better by watching and adopting methods of friends, families and even enemies. You can imagine some early tribe of humans watching how another group rubbed sticks together to create a fire. Although this behavior is a part of everybody's life, when it comes to organizations, sometimes their behavior more resembles that of "Not-Invented-Here." If we didn't invent it, it must not be good.

Premier companies, though, have learned that benchmarking is an effective business tool. Jack Welch, CEO of General Electric, encourages his people to *"steal shamelessly."* In his view, it's not who invented the idea, it's who puts it into effective practice that counts.

We'll define benchmarking as simply:

Learning From Others Who Do Things Well And Applying That Learning To Our Business

Some benchmarking do's and don'ts include:

- Benchmarking should be externally, not internally, focused – look outside your company or department for ideas.
- The focus should be on improving performance, not dwelling on current poor performance or past failures – it's too easy to see how well others are doing things and get discouraged.
- Benchmarking should complement other improvement efforts – we include benchmarking in this section on design management since it can be an effective method of identifying best concepts for new designs or to aid in a product or process redesign.
- Benchmarking will help uncover best practices and shouldn't be performed to identify and criticize those in your organization who have not been quick to adopt new practices.
- Benchmarking may start off as a quantitative analysis, looking for *who* is the best performer, but must then transition to a qualitative analysis – what practices or processes are employed by best performers?

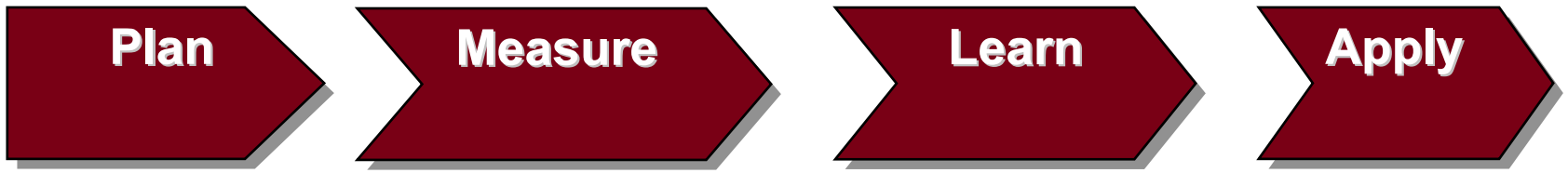
- Understanding best practices can take a good deal of study – many companies practice "*Industrial Tourism,"* a superficial analysis of best practices.

- The temptation in benchmarking is to go and see how our competition is doing something. While valuable lessons may be learned there, often looking at how completely different companies are performing similar *functions* is a more effective approach.

- Benchmarking is not simply a fad – premier companies such as Xerox have made benchmarking a key corporate process and part of their culture.

14.3.2 Benchmarking Methodology

Benchmarking is essentially a data collection effort. Therefore, the benchmarking process will follow similar principles and practice. A four-phase approach to benchmarking appears below:



Types of Benchmarking

There are two “different” types of benchmarking. The first purpose is to understand *who is doing what well* and is termed performance benchmarking.

Performance Benchmarking - Assists In Assessing Your Competitive Position. Examples:

- Elements Of Price
- Technical Quality
- Product Or Service Features
- Speed To Market
- Reliability

Performance Benchmarking is Widely Practiced In Diverse Industries:

- Automotive
- Computer
- Financial Services

The next type of benchmarking has as its purpose the discovery of best practices that lead to superior performance. This second type is actually a combination of identifying who is doing what well and then focusing on how they do it well.

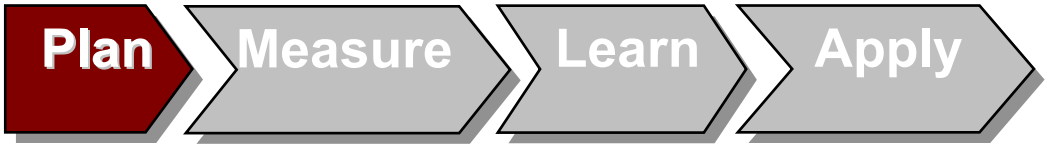
Process Benchmarking - Focuses On Discrete Work Processes And Operating Systems. Examples:

- Customer Complaint Process
- Billing Process
- Order And Fulfillment
- Recruitment
- Strategic Planning

When should benchmarking be considered? Benchmarking opportunities may include:

- Management Changes
- New Or Changing Operations
- Changing Industry Dynamics
- Planning Budget Processes
- Improvement Programs
- Changing Or New Processes, Products Or Services
- Process Redesign Initiatives
- Survival Or Crisis Situations

Benchmarking – Planning

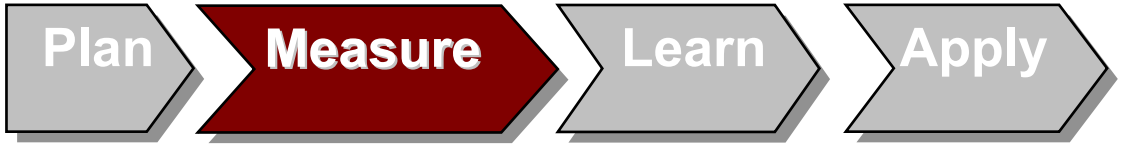


Two key activities are performed in this phase:

Define the Scope – What is the purpose of the benchmarking effort. Remember that this is a data collection effort. Is there a current business problem that has resisted solution? Is the benchmarking effort supporting the design/redesign of a product or service? Or is this an environmental scan to see what's going on in the marketplace? If some improvement is desired, is there a particular focus for the effort? During the 1980's, *reliability of service* was part of one utility's corporate strategy. Benchmarking efforts were conducted to discover best practices in design, operation and maintenance that led to high service reliability.

Plan the Project/Allocate Resources – A formal benchmarking effort may require several staff working for a few months. A project plan should be developed for the effort (and blended with the "higher" level project, if benchmarking is to support some improvement or design effort).

Benchmarking - Measure



The two key activities here include:

Establishing a Current Baseline – For an existing product or service, determine how well the product/service is currently working (performance to Key Characteristics). For a new product or service, determine the requirements, including targets and goals.

Steps	Key Issues	Tool
Set Baseline • For New Product/Service A. Review Functional/Process Targets/Goals • For Redesign, Use Current Process As Baseline A. Identify The Customer And Supplier B. Document The Process C. Understand Current Process Capability D. Assess Capability Gaps	• How Do We Do It? • How Are We Doing? • How Well Do We Need To Do?	• Process Mapping • KC Matrix • Quality Functional Deployment Matrix

Baseline data may come from the following sources:

Baseline Element	Source
Process Description	Process Participants
Process Flow Maps	Process Participants
Organization Chart (Who Reports To Whom)	Management Interviews/Process Participants
Deployment Chart (Who Does What In The Process: Matrix Of Process Steps And Process Participants)	Process Participants
Process Measures: Current Performance Goals And Levels, Cycle Times, Costs, etc.	Systems, Measurement Reports
Summary Of Enablers: Technology/Systems/Functionality	Process Participants
Personnel/Staffing Issues: Hiring/Training/Turnover	Management Interviews/HR
Personnel Measurement: Performance Planning And Evaluation/Incentives	Management Interviews/HR
Strategic Issues: Role Of Senior Management/Visibility Of The Process Or Function/Cultural Issues Change Initiatives/Objectives	Management Interviews Management Interviews/HR
Success Factors/Customer Requirements	Customer Interviews
Supplier Practices	Process Participants, Supplier Interviews

Selecting Benchmarking Partners – First, let's define the term *partner.* In any benchmarking effort, you will be asking someone for help. Be prepared from the start to *give* as well as *take.* Second, be systematic in your search for benchmarking partners. It's often tempting to just ask around the office to see if someone knows who is doing well. Although this may yield a few ideas, go beyond this casual research to identify partners.

Steps	Key Issues	Tool
Select Partners **A.** Identify Criteria For Comparability • Critical Quality Characteristics (Four Or Five) That Define The Product/Service • CTQs Identified through Voice of Customer **B.** Sample Criteria… Business Or Process • Regional vs. National Or International • Highly Trained Workers vs. Non-Trained Workers • High-Transaction Volume vs. Low Transaction • Standardized vs. One-Time Activities **C.** Criteria Are Not… • Costs • Customer Needs Or Requirements • Technology • Time Constraints	• Who Does It Well? • Who Is Comparable?	• Best-In-Class Matrix

Best Practices Research & Analysis

Here are a few sources you may research to discover possible benchmarking partners:

- Benchmarking Databases
  - External Organizations
  - American Productivity And Quality Center (www.ibc.apqc.org)
  - SPI Council On Benchmarking

- Professional And Trade Associations

- Companies Recognized For Performance Excellence
  - Winners Of The Malcolm Baldrige National Quality Award
  - American Operations Of Deming Award Winners

- Company Resources
  - Employees
  - Staff Units
  - Business Units
  - Business Laboratories

- Company Intranet

- Internet

- Personal Contacts

Specific Sources Include:

Source Type	Where To Go	Specific Resources
Computer Search	• Local Public And University Libraries • U.S. Department Of Commerce - Current Industrial Reports • Washington Service Bureau - GAO Studies • Congressional Reports	• Key-Word Literature Search • Access Data Sources - Dialog - Dun And Bradstreet - Standard & Poor's Value Line • Moody's Investment Service
Professional/Trade Associations	• APQC Benchmarking Clearinghouse • Association For Manufacturing Excellence • American Marketing Association • The Strategic Planning Institute • Encyclopedia Of Associations	• Industry-Specific Reports • Industry-Wide Awards • Industry Expert Recommendations • Membership Lists • Trade Show Exhibitors
Consultants/ Process Experts	• Referral By Colleague • Professional Association • Trade Publication Listing • Bradford's Directory Of Market Research Agencies And Management Consultants • Who's Who In Consulting • Directory Of Management Consultants	• Third Party Industry Surveys • References To Individuals Or Data Sources • Evaluation Of Search Results

Assess Potential Benchmarking Partners:

- Develop a set of Comparability Criteria (e.g. reputation for excellence, similar organization or function, similar market conditions, etc.)
- Match Potential Partners With The Comparability Criteria
- Assess Potential Partners' Performance And Reputation
- Explore Accessibility to Their "Shop"
- Identify Examples Of Innovation

Benchmarking Protocol And Ethics

A Benchmarking Policy should be communicated to all employees prior to contacting external organizations. Guidelines should address the following areas:

- Misrepresentation – Do Not Misrepresent Your Identity In Order To Gather Information
- Information Requests – A Request Should Be Made Only For Information Your Organization Would Be Willing To Share With Another Company
- Sensitive/Proprietary Information – Avoid Direct Benchmarking Of Sensitive Or Proprietary Information
- Confidentiality – Treat All Information As Confidential
- Keep It Legal - Avoid Discussion Or Actions That Might Lead To Or Imply Restraint Of Trade (e.g. Bribery, Dealing Arrangements, Price Fixing, Misappropriation)
- Be Willing To Give What You Get - Be Willing To Provide The Same Level Of Information That You Request
- Respect Confidentiality - Information Obtained Must Not Be Communicated Outside The Partnering Organizations Without Consent Of Participating Benchmarking Partners
- Keep Information Internal - Use Information Obtained Only For The Purpose Of Improving Processes; External Use Of The Partners Name Or Data Requires Permission
- Don’t Refer Without Permission - Obtain An Individual’s Permission Before Providing Their Name In Response To A Request
- Be Prepared At Initial Contact - Demonstrate Commitment To The Efficiency And Effectiveness Of The Benchmarking Process With Adequate Preparation At Each Process Step

Collecting Benchmarking Information

Regardless of what method is employed (see below), an interview guide should be prepared once the potential partner has agreed to participate in your study

Three Roles Of The Interview Guide:

- To Gain Commitment To Participate From Potential Partners
- To Facilitate Data Collection
- To Serve As A Framework For Analysis

Methods Of Information Collection

- Telephone Interviews
- Site Visits
- Surveys
- Publications/Media

Method - Telephone Interviews	
Advantages Easy to Plan And Conduct Enables Contact With Large Number Of Resources Can Be Conducted At Almost Any Time Relatively Inexpensive	**Disadvantages** "Cold Calling" Can Be Time Consuming Difficult To Get Return Calls May Be Interruptions People Are Less Likely To Spend A Lot Of Time On The Telephone

Method - Meeting/Site Visit	
Advantages Establishes Relationship Provides Greater Quality Time Likely To Produce A Good Deal Of Information Permits Better Insight Into Processes Through Tour, Office Layout, etc.	**Disadvantages** Expensive (Travel Cost) Time Consuming Scheduling Difficulties

Method - Surveys (e.g., SurveyMonkey)	
Advantages	**Disadvantages**
Ability To Collect Information From A Large Population Easy To Construct Relatively Inexpensive Easy Transfer Of Information For Analysis	Likelihood Of Low Return Rate Impersonal Difficult To Ask Follow-Up Questions Questionable Validity Of Some Information Must Be Relatively Brief Little Possibility For Detailed Response

Method - Publications/Media	
Advantages	**Disadvantages**
Information At Your Fingertips Variety Of Resources Inexpensive To Collect Large Quantities Of Information Produced For Many Types Of Industries	Need To Validate Sources/Statistics Many Obscure References Time-Consuming, Particularly When Overabundance of Information Exists

When Planning A Benchmarking Effort, Consider The Following Guidelines:

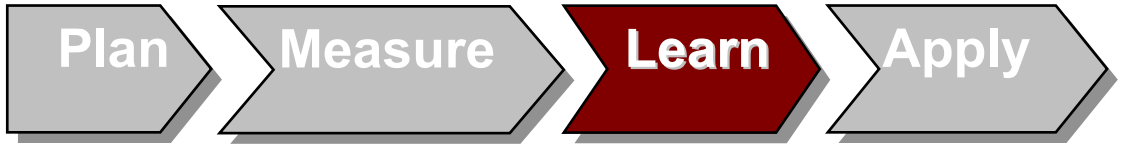
- Be Prepared To Explain Your Specific Objectives
- Allow The Partner Sufficient Time To Collect The Information
- Schedule Interviews Realistically
- Prepare A Presentation Package
- Pre-Screen Benchmarking Partners
- Use A Top-Down Contact Strategy
- Stick To Your Outline
- Limit The Size Of Your Site Visit Team To Two Or Three
- Provide The Questionnaire To Participants In Advance

- Interview In Pairs Or Trios
- Provide Opportunities For Participants' Personal Insights
- Allow One To Two Hours Per Telephone Interview Or 1/2 To One Day Per Visit
- Probe With Open-Ended Questions To Identify Core Issues
- Note And Record Observations During The Interview Or Site Tour
- Follow Up With A Thank You Letter And Request For Any Additional Information

Measure Phase Deliverables

Activities	Deliverables	Communication
Document process and issues	Complete process map or detailed flowchart	Confirm stakeholders obtain input from process owners
Identify, define and develop performance measures and drivers	List of performance measures and list/grid of drivers	Confirm with stakeholders; obtain input from process owners and customers
Assess current performance	Compile preliminary/existing data	Inform key stakeholders and process owners of summary information
Identify performance gaps	Begin gap analysis	Begin to consolidate gap analysis for stakeholders and process owners
Conduct Benchmarking partner research	Data/research package	Share key findings of research as appropriate
Identify partner candidates	List of candidates	Confirm list with management or champion
Set contact priorities	Prioritized list of partner candidates	Confirm list with management or champion
Benchmarking Research Performed	Benchmarking Interviews complete	Share key findings of research as appropriate

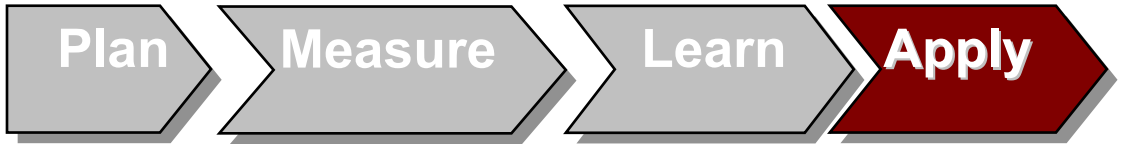
Benchmarking – Learn



Learning from the Benchmarking visits involves analyzing the data, understanding the *cause and effect* relationships between practices and performance and then developing necessary reports to communicate the results of the study.

Activities	Deliverables	Communication
• Arraying Raw Data • Testing Performance Data • Revising Measures • Ranking Partners • Reconciling Data And Anecdotal Information • Identify Best Practices • Gap Analysis	• Data Matrixes • List Of Outliers And Rationale For Each One • Revised Measures; Restatement Of Performance Data • Performance Ranking Matrix • Restatement Of Data Or Revision Of Other Information • Two-By-Two Matrixes • Best Practice Matrixes; Descriptions Of Individual Best Practices • Graphical Depictions Of Performance Gaps; Comparison Of Best Practices To Our Practices	• Oral And Written Reports On Results Of Analysis

Benchmarking – Apply Learning



Now is the time to apply what you've learned through your benchmarking efforts. The steps to perform here include:

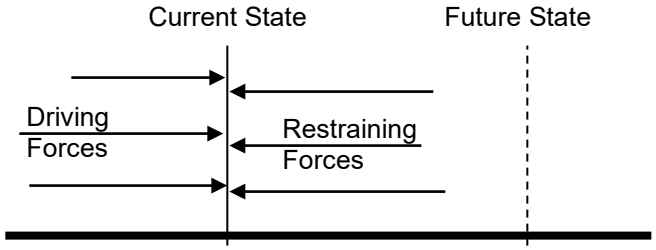
- Revisit Your Initial Hypotheses
- Review Baseline Assessment of Current Product/Service
- Based On Data and Analyses, Make Recommendations That Close Gaps And Assist The Organization In Achieving Goals

Solution Selection & Planning:

- What Will the Best Solutions Look Like?
- Which of the Best Solutions Will Be the Most Economical to Implement?
- Which Solutions Will Be the Most Dramatic? Least Disruptive?
- Which Solutions Will Be the Most Visible?
- Which Solutions Will Show the Fastest Results?
- Which Solutions Will Meet the Least Resistance and Be the Easiest to Put Into Place?
- What Technical Limitations Must Be Considered?
- Are Any Organizational "Sacred Cows" Involved?
- How Will the Solutions Blend With Existing Organizational Culture and Norms?

A Force Field Analysis May Help Understand Organizational Barriers To Change:

- Force Field Analysis Is A Technique To Generate, Organize And Assess The Consequences Of An Idea
- Driving Forces Encourage Implementation Of An Idea
- Restraining Forces Discourage Implementation Of The Idea.
- A Key To Successful Implementation Is To Reduce Those Forces Which Resist Implementation And Utilize Those Forces That Favor It



Some Elements of the Implementation Plan:

- Scope, Strategies, Goals And Objectives
- Anticipated Deliverables
- Test Sites
- Resource Needs
- Roles And Responsibilities
- Milestones And Tracking Procedures
- Training Schedules, If Needed
- Measurements And Standards
- Recognition Plan For Individuals And Teams
- Communication Strategies
- Analysis Of Critical Success Factors
- Back-Up Plans

A Few Benchmarking "Failure Modes:"

- Disruption Of Organizational Political Structures
- Inadequate Management Support
- Lack Of Stakeholder Buy-In To Process, Partners
- Failure To Include Key People In Decision Making
- Misunderstanding Of Change And Its Implications
- Low Tolerance For Change In General

The Role of the Benchmarking Champion/Sponsor

- Select Implementation Team
- Approve Implementation Plans
- Allocate Resources
- Communicate Vision
- Gain Stakeholder Support And Overcome/Manage Resistance
- Maintain Political Support And Remove Obstacles
- Reward And Recognize Progress And Achievement
- Formalize Change; Integrate With Existing Policies And Procedures

14.4 Taguchi Design Approach

Learning Objectives

- Be able to assess the Quality Loss associated with product/service variation.
- Be able to set up and run experiments to determine best parameter values for a design.

Unit Contents

- Taguchi Quality Engineering Methods
- Quality Engineering Process
- Sources of Product/Service Variability
- Quality Characteristics & Their Loss Functions
- Focus on Parameter Design
- Designed Experiment/Orthogonal Array and its Evaluation
- Pump Life Example Application
- Parameter Design Applications

14.4.1 Taguchi Quality Engineering Methods

This unit describes Genichi Taguchi's basic approach to Quality Engineering tasks. Taguchi uses the *Loss Function* concept as a prime measure of a product or service's quality. To minimize variation from the target value, he encourages engineers to design experiments using orthogonal arrays, mainly in the *Parameter Design* phase of product/production process design.

The Signal-to-Noise (S/N) Ratio plays an important role in evaluating the results of these designed experiments. The S/N ratio is primarily important because it is an *additive* quality characteristic, one for which interactions are often not important. The S/N ratios are also analyzed differently than the "traditional" ANOVA of the data. Use of the S/N ratio simplifies the analysis for the engineer who is not a statistician.

14.4.2 Quality Engineering Process

Taguchi's Quality Engineering process is aimed at improving the design, development and production of products and services. Let's examine his approach first from the "50,000 feet" level.

Definition and Evaluation of Quality

Various *definitions* of quality can be found, both in the dictionary and, depending on the "school" of quality to which one subscribes, from various quality "gurus." For example, Juran has long used the "fitness for use" definition, Crosby has focused on ”Zero Defects," many companies define quality as "doing the right thing right." Likewise, various methods are used to *evaluate* quality. The "voice of the process" is often determined through Shewhart control charts, which allow the production person to distinguish common versus assignable causes of variation in the product's quality characteristics. Process capability indices (C_p, C_{pk}, etc.) are used to compare the variation in quality characteristics to the customers' tolerance. Fraction defectives, defect rates, or sigma calculations are used to describe the degree of non-conformance to customer requirements.

Taguchi has introduced both a “new”[1] definition of quality and a means of evaluating quality. He *defines* quality as follows:

> "Quality is the **loss** that a product imparts to society."

Within the domain of quality *engineering*, the following qualifiers limit this “loss”:

> a) ". . . (loss to) society after the product has been shipped."
>
> b) “However, loss caused by the function itself is excluded."

Taguchi does consider the traditional concept of ***value*** (benefit vs. cost) but assigns responsibility for this to the *Marketing* activity of the Quality Assurance system, not to the design and production of products and services (see below). Practically, Taguchi assigns responsibility to the *quality engineering* function for losses due to deviations from the intended functions of the product/service and to "ill effects" (i.e. safety, environmental damage, costs of operation & maintenance, etc.), which result from use of the product or service.

[1] “New” still to many companies – Taguchi’s early work was published in the 1960’s.

To *evaluate* quality, Taguchi uses the **Quality Loss Function (QLF)**. He derives a general QLF as a quadratic function increasing as the actual value of the product's quality characteristic moves away from the target in either direction. Taguchi also uses the **Signal-to-Noise (S/N) Ratio** to evaluate the results of experiments designed to improve quality (i.e. minimize the loss to society). Both his definition of quality and his means of evaluating quality take some time to get used to and to apply in the "daily" work of the engineer.

Quality Assurance Systems for Products & Services

Consciously or unconsciously all companies makes use of a Quality Assurance system to develop, produce and sell products and services. A generalized Quality Assurance system for any product or service includes the following activities (note: these functions are embedded in DMEDVI):

Market Research - A company must understand from the market what are the needs of consumers and the *level* of such needs (i.e. is there the potential for sufficient sales volume for the product/service?).

Product/Service Planning - The Company must then determine which products/services it wishes to develop to meet the needs of its customers. Here, product/service functions, price and the designed life (for products) are determined.

Product/Service Design - The functions selected in the *Planning* phase are then developed into actual products or services.

Production Process Design - This phase determines how the product or service will be produced.

Production - The production processes are used to produce products or services.

Sales - The designed, produced product or service is offered for sale to consumers.

After-Sales Service - For products, there may be after-sales maintenance, warranty claims or operations/technical support.

Taguchi's Quality Engineering methods focus on three of these major activities: *Product/Service Design, Production Process Design,* and *Production.* For the first two, he provides efficient methods for *off-line* improvement of quality and cost; for the third, his methods address *on-line* quality control. The "off-line" component is the main focus of this section.

"Off-Line" Quality and Cost Improvement

To efficiently achieve the "off-line" improvement of quality and cost, Taguchi advocates the following system:

Product/Service Design Phases - The traditional design process consists of two basic steps: 1) design the system and its parts (focusing on the *average* response) and 2) set *tolerances* around these average values.

Taguchi takes this two-step process and adds another phase to the design effort:

> *System Design* - Once the "objective" functions are identified (Product Planning's responsibility), the system is selected that matches these objectives (e.g. *Concept Selection*).
>
> *Parameter Design* - In this phase, the system's parameters or design constants are determined, but effort is also expended to reduce the influence of sources of variation. Attempts are made to exploit the "non-linearity" of input to output variation.
>
> A key reason for *Parameter Design* comes from the shift in consumer demand to a high variety of products and services. In the "old days," product variety was limited and mass production in large quantities occurred. Under these conditions, the production forces could initiate quality improvement activities to correct design deficiencies and reduce defects. If the market now requires a high variety of products and low production volumes per product, this approach will not work very well, if the goal is to produce high quality products.
>
> *Tolerance Design* - Here, the grade of parts and materials is determined, using a trade-off between quality and cost. Taguchi states that, while the Parameter Design phase improves quality at no additional cost, this phase may improve quality at increased cost. The Loss Function is applied to evaluate these trade-offs.
>
> *Determine Tolerances* - Here, the tolerances to be stated on drawings, specifications, contracts, etc. are determined. This is the traditional “geometric” tolerancing process.

Production Process Design - The same three phases are applied to the design of the production processes:

> *System Selection* - Determine the production systems by applying available technology or developing new technology. Feedback and feed forward philosophies are applied to determine how control of quality will occur.
>
> *Parameter Design* - Determine the optimum operating conditions for each production "unit" (i.e. stamping, forging, machining, heat treatment, etc.). This is similar to the product parameter design in that the purpose is to minimize or eliminate the effects of production process sources of variability.
>
> *Tolerance Design* - Determine the functional tolerance limits for each production "unit." These will translate to the actual working limits applied to production processes.

Practical, Efficient Methods - To achieve the objectives of these phases (especially the *Parameter* and *Tolerance* design phases), Taguchi employs *Designed Experiments* (using both *Single Factor* and *Orthogonal Arrays* (although the latter is emphasized). The *Signal-to-Noise Ratio* is used to evaluate the results of these experiments in the Parameter phase; the Quality Loss Function is used in Tolerance Design.

Taguchi emphasizes the selection of *appropriate* quality characteristics (he states that this often takes up to 80% of his consulting time, with the other 20% going to the actual experimental design). His basic philosophy is to treat the function as an *energy transformation* and identify the appropriate outputs of this transformation. For example, he takes a dim view of using *product life* or *reliability* as a quality characteristic, preferring to identify a measure of the underlying physical process that determines life/reliability (e.g. wear, corrosion, etc.).

He strongly recommends selecting quality characteristics that are *additive* eliminating interactions between factors. One example of this thinking involves administering drugs to patients and recording their response. The drugs by themselves are observed to improve the patients' health, but, when administered in combination, worsen health. Here is a clear example of *interaction* at work. However, suppose the three drugs are all insulin compounds. By themselves, they work fine. When combined, though, they result in an overdose of insulin. In this case, the quality characteristic of *patients' health* is *not* additive. The characteristic of *insulin level* in the body *is additive and* should be used to measure the response of the experiments.

14.4.3 Sources of Product/Service Variability

A product is exposed to different environmental conditions, such as pressure, temperature and radiation. Variability in the product's function due to these sources is referred to as *External Noise.* Over time, a product's functions may degrade, and, at some point, there will be a loss of function. Several years ago, we had to replace the clutch in our car's manual transmission due to loss of function. Variability due to these sources is referred to as *Internal Noise.* Finally, some products produced according to specifications may not function at all. From our experience, about 10% of our Fourth of July bottle rockets fail to operate or explode on launch. This is referred to as *Between-Product Noise.*

This classification has two main uses in Quality Engineering. The first relates to responsibility. The engineering function controls virtually all these sources of noise. The production function, on the other hand, can basically control only the *Between-Product Noise* component. The table below summarizes the possibilities. The second reason this classification is important is found in Taguchi's approach to Design of Experiments using Orthogonal Arrays. Historically, Design of Experiments (DOE) arose in the agricultural field and has also been applied extensively in the medical arena.

In these applications, the experiments are conducted in a "sea of noise." For instance, a drug manufacturer attempts to determine the effects of a new treatment for cancer or HIV. Trial experiments may be conducted on humans, which are highly variable in their response. Here, Fisher's principles of *Replication, Randomization* and *Control* are necessarily applied to the design of the experiment and *tests of significance* are used to separate the signal (i.e. improvement in outcome) from the test subjects' noise.

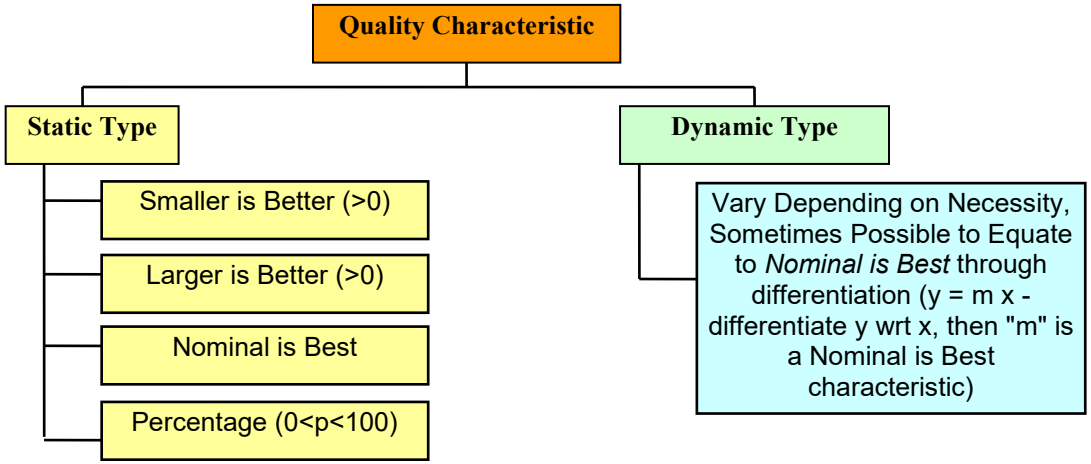
Taguchi has decided that industrial experimentation, especially in the Parameter Design phase, does not occur under these conditions. He has found that it is often possible to identify the extreme conditions of noise and includes these conditions explicitly in the experimental design. His approach, though, treats *external noise* through an **outer orthogonal array**, separate from the **inner orthogonal array** where the control factors are being examined. He finds it important, then, to distinguish these types of noise.

Function	Activity	Phase	Type of Noise		
			External	**Internal**	**Between - Product**
Engineering	***Product Development***	***System***	Possible to Control	Possible to Control	Possible to Control
		Parameter	Possible to Control	Possible to Control	Possible to Control
		Tolerance	Possible to Control, but Avoid	Possible to Control	Possible to Control
	Production Process Development	***System***	Not Possible to Control	Not Possible to Control	Possible to Control
		Parameter	Not Possible to Control	Not Possible to Control	Possible to Control
		Tolerance	Not Possible to Control	Not Possible to Control	Possible to Control
Production	***Production***	***Process Control***	Not Possible to Control	Not Possible to Control	Possible to Control
		Product Control	Not Possible to Control	Not Possible to Control	Possible to Control
Sales	***Sales***	***After Sales Service***	Not Possible to Control	Possible Through Preventive Maintenance	Possible Through Preventive Maintenance

14.4.4 Quality Characteristics & Their Loss Functions

Prior to designing experiments to improve quality,[2] the objective functions of the product or service are determined in the *Product Planning* activity of the Quality Assurance System. The *Quality Characteristics/Product/Service Requirements* associated with these functions will also have been identified. This is a key point. An item will have many different quality characteristics, but most do not relate to the *objective function* of the item. For example, a resistor's quality characteristics include its *resistance, amperage rating, precision, length, diameter,* and *weight.* But when it is applied in an electrical circuit, its objective function is to convert voltage to current (or vice versa). The *current* then is the output quality characteristic of interest. Taguchi defines an objective function as an *energy transformation.* This should be kept in mind when selecting appropriate quality characteristics.

Taguchi's quality engineering methods may be applied at all levels of product design: System, Subsystem, Equipment, Components and Parts. As the design effort focuses on the higher levels of indenture, though, the "product" and its associated production process will tend to be studied together (e.g. integrated circuits and their production process). Taguchi classifies quality characteristics and develops Quality Loss Functions for each major type.[3] This hierarchy appears below.



[2] Unless otherwise noted, "quality" is defined per Taguchi's concept of Loss to Society.
[3] This also applies to the associated Signal-to-Noise Ratio.

Nominal is Best Loss Function - The Loss Function associated with quality characteristics whose "best" condition is a target or nominal value is given by the following equation:

$$L(y) = \frac{A_0}{\Delta_0^2}(y-m)^2$$

$$where:$$

$$L(y) - Loss$$

$$A_0 - Monetary\ Loss\ at\ Function\ Limit\ \Delta_0$$

$$\Delta_0 - Function\ Limit\ (Point\ at\ which\ function\ is\ lost)$$

$$y - Value\ of\ Quality\ Characteristic$$

$$m - Target\ Value$$

Figure 14.4-1a) provides a graphic interpretation of this function. The function is derived by assuming that at the target "m," there is "zero" loss and that the first derivative of Loss is also 0. A Taylor's Series expansion provides the quadratic function, consideration of the dollar (or yen!) value of the Loss when the function is lost provides the A_0 and Δ_0 constants.

There is nothing "statistical" about the Loss Function (i.e. a Loss can be calculated for a single item). For many items though, the Loss Function can be evaluated as a function of the items' variance:

$$L(y) = \frac{A_0}{\Delta_0^2}\sigma^2$$

$$where:$$

$$\sigma^2 - Average\ of\ (y-m)^2\ \ (i.e.\ the\ Variance)$$

From this latter equation, the basic strategy of quality engineering appears:

Minimize A_0 - This is generally accomplished at the *System Design* phase and addresses the *Safety Loss* of the system.

Increase Δ_0 - To accomplish this, we must design the product to be *robust* against variation, that is, our goal is to increase the range in which the product can operate without loss of function. Parameter design will play an important role here.

Decrease σ - This results from both design and the control of production processes (on-line quality control).

Larger is Better Loss Function - For quality characteristics whose condition is "better" at larger values, the Loss Function is:

$$L(y) = A_0 \Delta_0^2 \frac{1}{y^2}$$

where :

A_0 – Monetary Loss at Function Limit Δ_0

Δ_0 – Function Limit (here, it is usually a lower specification limit)

y – Value of Quality Characteristic

Figure 14.4-1b) depicts this relationship. For many items, the *Larger is Better Loss Function* is:

$$L(y) = A_0 \Delta_0^2 \sigma^2$$

where :

$$\sigma^2 - Average\ of\ Inverse\ of\ y's\ Squared = 1/n \left(\sum_i 1/y_i^2 \right)$$

Smaller is Better Loss Function - For quality characteristics whose condition is "better" at smaller values, the Loss Function is:

$$L(y) = \frac{A_0}{\Delta_0^2} y^2$$

where :

A_0 – Monetary Loss at Function Limit Δ_0

Δ_0 – Function Limit (usually an upper specification limit,
sometimes based on the 50% Lethal Dose concept or LD50)

y – Value of Quality Characteristic

Figure 14.4-1c) depicts this relationship. For many items, the *Smaller is Better Loss Function* is:

$$L(y) = \frac{A_0}{\Delta_0^2}\sigma^2$$

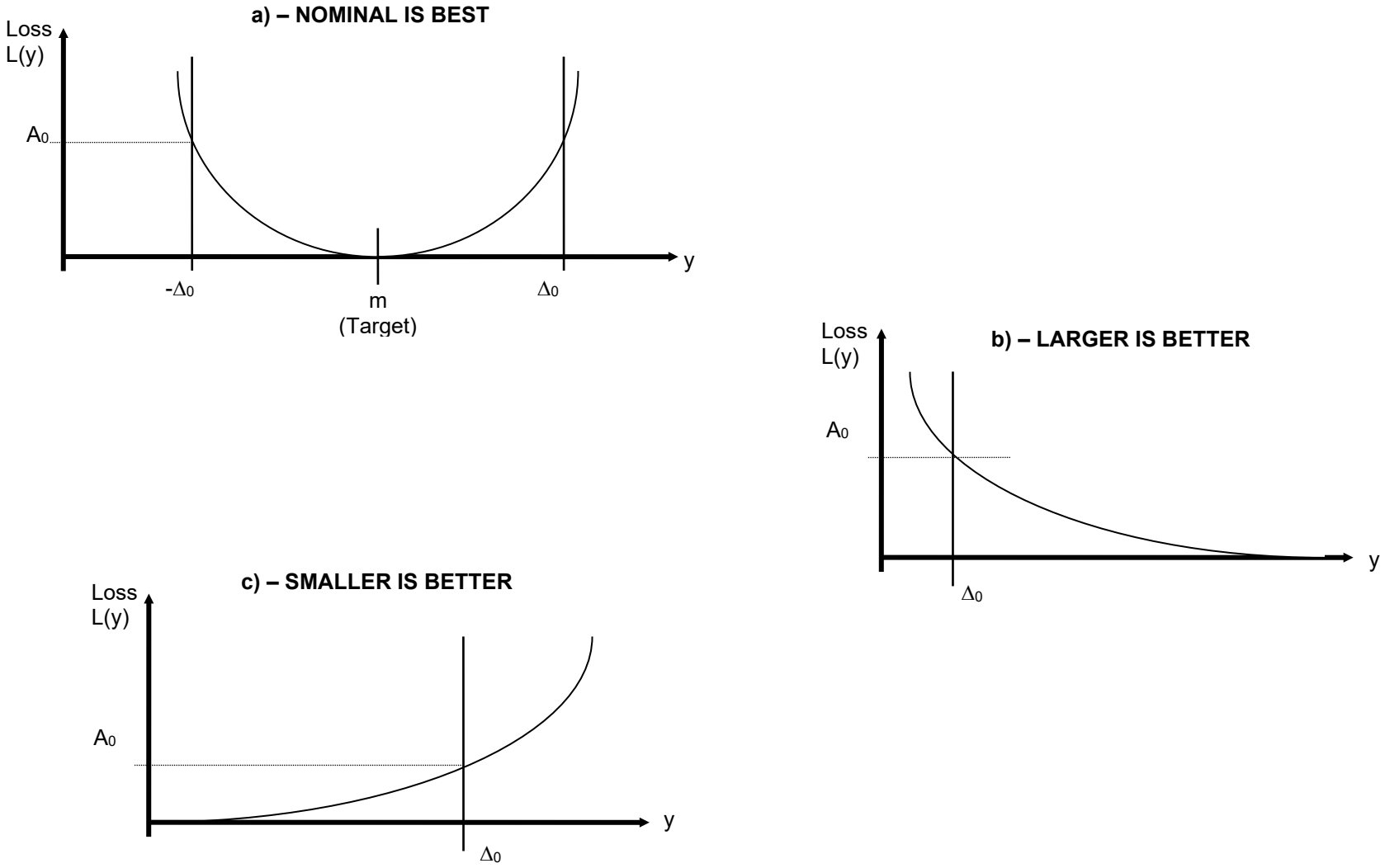
where:

$$\sigma^2 - \text{Average of } y^2 = {}^{1}\!/\!_{\text{n}} \sum_i y_i^2$$

(usually obtained from the Signal - to - Noise Ratio)

Percentage Loss Function - This Loss Function would be applied to situations such as efficiencies, concentrations, etc. The *Smaller is Better Loss Function* can often be applied here.

Figure 14.4-1 - Loss Functions



14.4.5 Focus on Parameter Design

The *Parameter Design* phase of system design is key to Taguchi's quality engineering method. This phase's aim is to identify the key *control factors* and their optimum levels so that the goal of "on target with minimum variation" is met.

Parameter Design Process

Parameter Design occurs after a system[4] has been "selected" to meet the required objective functions (see Unit 14.2 for concept design methods). The Designed Experiment using Orthogonal Arrays (DE/OA) is employed to examine the effects of varying the inputs (control factors) on the output (objective function's quality characteristic). The *Signal-to-Noise Ratio* is used to evaluate the results of the DE/OA. The *Loss Function* can then be calculated to estimate the improvement in quality achieved by the experiment (the Loss Function and Signal-to-Noise Ratio are inversely proportional).

The process employed in the Parameter Design phase is similar, regardless of the type of quality characteristic being considered. Variation occurs, for instance, in the specific Signal-to-Noise Ratio selected to evaluate the results of the DE/OA, or in the specific Orthogonal Array employed.

The basic steps followed in any Parameter Design effort include:

1. Clarify the Output - Understand the objective function (using the *energy transformation* concept) and develop quality characteristic(s) that measure this function (QFD can be employed here). The resistor's function is to transform voltage into current - *current* then is the quality characteristic that measures the objective function.

2. Determine Inputs that may affect the Output - Based on the technology, identify possible control factors whose level may affect both the average and variation of the output. Also, identify possible noise factors (based on the concept of *External, Internal,* and *Between-Product* noises).

3. Design the Experiment - Assign control factors to the columns of an OA. Assign as many control factors as possible to the columns. Only if there are "empty" columns should interactions be assigned to columns (See the discussion below). If there are noise factors to consider, these should be assigned to the outer array and *compounded* if possible (combining all the noise factors/levels at two extremes).

[4] "System" is used here loosely, since Parameter Design may (and should!) occur at various indenture levels.

4. Run the Experiment - Perform the experiments indicated by the Orthogonal Array. The experiments can be run in a manner that is "most convenient or practical." Taguchi does not worry much about randomization (as in "traditional" DOE). No adjustments should be made during the experiments; other than to change the control factors' levels based on the Orthogonal Array. Collect the experimental data.

5. Evaluate the Results - Based on the desired output (i.e. Smaller is Better, etc.), calculate appropriate Signal-to-Noise (S/N) Ratios see below). The Signal-to-Noise Ratio combines both the average and variation response of the output to the control factor levels.

The results of the experiments are evaluated **without regard for the target value.** This is one of the key features of the Parameter Design approach. The approach is to determine the combination of factor levels that **maximizes** the S/N Ratio. A *Sensitivity* value is also calculated to determine which factors will be the best *adjustment* factors (those that can be used to move the average output to the target value). Adjustment occurs after the best factor level combination has been determined.

The S/N Ratio & Sensitivity are calculated for each trial (row of the OA):

a) For each level of a factor, identify all the trials run at that level (i.e. all the "1's," or "2's," or "3's"). Average the Signal-to-Noise Ratios for that level.

b) Prepare a *Supplementary Table* that records the average S/N Ratios for each factor/level:

	Levels		
Factor	**1**	**2**	**3**
A	AVG_{A1}	AVG_{A2}	AVG_{A3}
B	AVG_{B1}	AVG_{B2}	AVG_{B3}
C	AVG_{C1}	AVG_{C2}	AVG_{C3}

c) Determine the optimum conditions by identifying the level for each factor that has the highest S/N Ratio (note that this will often be the *least negative* number). If interactions were considered in the OA and they exist as measured by the S/N Ratio, some thought must be given to picking the optimum level. For example, suppose A_1 and B_2 are optimum, based on

their S/N Ratio (and the A_1 S/N Ratio is greater than the B_2 S/N Ratio). Further, suppose that there is an A_1xB_1 interaction. If the A_1xB_1 S/N Ratio were greater than the B_2 S/N Ratio, the optimum condition selected would be A_1 B_1.

d) Calculate the predicted process average based on the following equation:

$$\overline{\mu} = \sum \overline{SN}_{F_i} - (n-1)\overline{T}$$

where:

$\overline{\mu}$ – *Predicted Process Average*

$\overline{SN}_{F_i}$ – *Average Signal - to - Noise Ratio for Optimum Level of Factor* F_i

n – *Number of Factors*

$\overline{T}$ – *Sum of S / N Ratios across Levels of a Factor (from Supplementary Table) divided by the total number of experiments*

If all factors are included in this equation, the process average tends to be over-predicted. Often, only the most important factors are included in this prediction (if this is done, "n" equals the number of factors included). A rational basis for including important factors is to take the difference between the S/N Ratio at the factor's optimum level and the *average* of the S/N Ratios for that factor. The Pareto Principle is then applied to pick the factors that have the largest differences.[5]

6. Perform a Confirmatory Experiment - Run an experiment at the optimum conditions found above. If an outer array was used in the DE/OA, run the confirmatory experiment at these noise levels. Calculate the average response (**not** the S/N Ratio) at these noise levels and compare to the target value (for Nominal is Best cases). If the response is close to the target, the design is optimized. If not, then *adjustment* must occur. Look for a factor that does not affect the S/N Ratio, but does affect the Sensitivity. If there are several control factors, this may be easy.

7. Evaluate the Loss - The *anti-log* of the Signal-to-Noise Ratio at the optimum condition can be input directly to the Loss Function and the average Loss per unit can be estimated. If this Loss is relatively small compared to the price of the product, the study may be concluded at this point. Otherwise, the process may then proceed to *Tolerance design* phase.

[5] This procedure is much simpler from the calculation viewpoint. An ANOVA could be used to identify the "Significant" Factors.

Signal-to-Noise Ratios

The Signal-to-Noise (S/N) Ratio is used in quality engineering to evaluate the results of Designed Experiments. The S/N Ratio depends on the type of quality characteristic (Larger is Better, Smaller is Better, Nominal is Best, etc.). If the results of an experiment yield data $y_1, y_2, y_3, \ldots y_n$, then the S/N Ratios ($\eta$) are calculated as follows:

Smaller is Better η:	$\eta = -10\log\sigma^2$ *where*: η – Signal - to - Noise Ratio $\sigma^2 = \frac{1}{n}\sum y_i^2 \ , i = 1 \ to \ n$
Larger is Better η:	$\eta = -10\log\sigma^2$ *where*: η – Signal - to - Noise Ratio $\sigma^2 = \frac{1}{n}\sum \frac{1}{y_i^2} \ , i = 1 \ to \ n$
Nominal is Best η:	$\eta = 10\log \frac{\frac{1}{n}(S_m - V_e)}{V_e}$ *where*: $S_m = \frac{1}{n}\left(\sum y_i\right)^2 , i = 1 to n$ *and* $V_e = \frac{1}{n-1}\left(\sum y_i^2 - S_m\right)$

All these S/N Ratios combine the effects of the average response as well as the variation. "Better" response occurs when the S/N Ratio is larger. For the *Smaller is Better* S/N Ratio, it is easy to see that this translates into a smaller average *and* variation.

For the *Larger is Better* S/N Ratio, the larger the average, the larger the ratio (the logarithms of smaller fractions are larger negative numbers). The opposite relationship holds for the variation - the smaller the variation, the larger the S/N Ratio. For example, suppose factor level-1 gives results 2, 4 and factor level-2 gives results 3, 3. Although the average of the factor levels is the same, the variation of level-1 is greater than level-2 and the respective S/N ratios are 8.1 and 9.5.

The *Nominal is Best* S/N Ratio is derived from the inverse of the coefficient of variation (σ/μ). This ratio is larger when the variation is small relative to a given average.

One useful property of the S/N Ratio is that it is inversely proportional to the Loss Function. If, for instance, the S/N Ratio increases by a factor of 5 (optimum to initial conditions), then the Loss due to variability is reduced by the same factor. This property holds for all the S/N ratios, regardless of their specific calculation.

Why is the Signal-to-Noise Ratio used to evaluate the results of the experiment? There are at least three key reasons explained by Taguchi:

Relation to Function - The S/N Ratio applied in other engineering fields (i.e. communications) is related to *energy* (sound pressure squared, current squared, etc.). Taguchi views the system's function as an energy transformation; therefore, the S/N Ratio is a useful measure of the function.

Interactions - When a simple output characteristic (i.e. a "*y*" value) is analyzed, interactions between factors can be significant. The Signal-to-Noise (S/N) Ratio is a measure of *stability* (i.e. variability). According to Taguchi, the interactions between control factors are usually small.

Taguchi is somewhat ambiguous about this latter point. In one instance, he says that this feature of the S/N Ratio obviates the need to use an OA (in favor of one-factor-at-a-time experiments). On the other hand, he notes that since there is no guarantee that interactions do **not** exist, the OA should be used to test for reproducibility.

Approaching the Target Value - When the objective characteristics are far from the target, the “conventional” characteristic values can be used to get close to the target (additivity exists). As the parameter design gets closer to the target, the effects of approaching the target value can change to the effects of departing from the target, even though the main effect remains the same. These do not also contribute to study of stability design and the optimum condition.

14.4.6 Some Comments on Taguchi's Approach

For those used to a "traditional" Design of Experiment approach (e.g. based on Fisher's methods), the Taguchi approach appears to have some different thinking incorporated. Here is a summary of some of the perceived and actual differences.

Fisher/Taguchi Experimental Approaches

Sir R. A. Fisher is recognized as the father of modern experimental design. His application field began in the agricultural world. Here, the experimenter deals with a large source of experimental error (or noise) that is *Nature.* The purpose of a *Fisher-based* experiment is to be able to detect a small signal component (either from main effects, interactions, or both) from the large noise.

Fisher's three principles of *Randomization, Replication,* and *Control* are necessary to achieve the purpose of this kind of experiment. *Significance tests* are used to provide us with a measure of confidence that an actual signal exists. The *Fisher-based* design of experiments was extended to industrial problems as well as to medical research (where the basic large error/small signal issue still applies!).

In the world of Fisher-type experiments, the *fractional factorial design* is viewed as a more efficient experiment than the *full factorial* (all combinations of factors & levels).

Taguchi's industrial world is different from Fisher's agricultural world. "Nature" does not intrude into the laboratory or, in many cases into the production facility, i.e. Taguchi's world does not involve the large experimental error component. Rather than the Fisher principles, the following conditions must be met for the OA:

1. The experiment must be conducted under different combinations of factor levels,

2. The number of experiments must be within "practical" limits, and

3. The comparison of factor levels must be *balanced* (i.e. for factors B, C, D, E, . . ., their factor levels must appear the same number of times for the factor levels of A).

Rather than a more efficient experiment, Taguchi views the Orthogonal Array as a means of checking the *reproducibility of results downstream.* When the confirmatory experiment is run, the predicted value for optimum conditions and the confirmatory value are compared.

If they disagree, then the experiment (and its results, the "optimum" design) must be scrapped. The parameter design should be continued using different control factors/conditions. This is a significant difference between the traditional and Taguchi approach to experimentation.

Experimental Setup Principles

Whole texts have been written on this topic; this section is intended to illustrate some of the unique features of the Taguchi approach.

Interactions - Unlike the Fisher-type experiment, the aim of the Taguchi experiment is not to identify interactions in the DE/OA, but to examine this in the confirmatory experiment (the *reproducibility* issue). Therefore, interactions are only assigned to columns if there are empty columns. Two of the OA's that are strongly recommended (L12 for 2-level factors, L18 for "mixed" 2- and 3-level factors) have the interactions confounded in each column. Taguchi recommends that control factors be assigned to all of these columns, if possible.[6]

Noise Factors - To develop a product that is robust to *external* (environment) and *inner* (deterioration) noise, the error variance should be calculated, which indicates the magnitude of variability due to these conditions. In an OA, as *A* is varied over two levels, the other control factors play the role of noise factors. This is another advantage to using the OA over single factor at a time experimentation.

Inner and Outer Arrays - Again, unlike the Fisher-type experiment (where "Mother Nature" provides the experimenter with all the noise he/she can stand!), Taguchi asks the experimenter to consider the different types of possible noise (*External, Internal, & Between-Product*) and address their impact on the S/N Ratio. These are treated through an *outer array* that may, itself, be an OA. However, noise levels may be *compounded* into "extreme" conditions to simplify the experiment. In the experiment below, three factors (temperature, length and flow) are controllable by the design and production process; two factors (size and layers) are not – these are treated through the external array:

[6]There is an "ulterior" motive behind this. With a large number of factors, the temptation to adjust the factors' levels during the experimentation is dampened, since it is generally difficult to predict the results. The experimentation has been observed to proceed quicker, since the results calculations can't occur until all experiments have been run.

				External Sources of Noise					
Internal/Between Product Noise				Size:	N	L	H	L	H
Test	Temp.	Length	Flow	Layers:	N	L	L	H	H
1	L	L	L		3.2	9.0	1.8	1.6	7.8
2	H	L	L		3.1	2.6	2.0	2.3	4.8
3	L	H	L		1.9	2.7	2.0	1.5	3.4
4	H	H	L		2.4	2.2	1.5	1.7	1.9
5	L	L	H		1.6	2.4	1.6	1.5	2.9
6	H	L	H		1.6	1.9	1.7	1.7	1.8
7	L	H	H		3.3	3.3	1.6	1.6	3.3
8	H	H	H		2.2	2.6	1.8	1.6	1.9

Running the Experiment

Since Taguchi abandons Fisher's Principles (especially *randomization*), the engineer may perform the experiments in the most convenient manner. In fact, since the trials are not randomized, a factor that is difficult to change may be assigned to column no. 1 of the OA. This results in all the level 1 experiments being performed first, followed by the level 2 experiments.

The secondary objective of the Orthogonal Array is to maintain the validity of the evaluation of stability by preventing adjustments (during the experimentation). Taguchi is aware of engineers' tendencies to "tamper" with the control factors during experiments.

14.4.7 Pump Life Example Application

In this case, a pump design was investigated for improvement of *time to failure.* This case is not strictly one that illustrates the *parameter design* approach, but it does demonstrate the Designed Experiment using Orthogonal Array method. In addition, it demonstrates the application of the *function* as *energy transformation* concept in the choice of output.

The pump in question has a "sliding part;" during the experiment measurements were taken on eight different positions of this part. Pumps with "sliding parts" typically are of the positive displacement type, with a piston/cylinder arrangement used to pump the fluid. The "sliding part" may very well be the piston ring(s) of the pump.

One possible approach to this improvement effort would be to design an experiment where *time to failure* was the quality characteristic, perhaps under accelerated conditions. In this example, the *function* of the piston rings is to provide a seal between the piston and the cylinder. *Wear* occurs during operation; this is the measure of the *energy transformation* occurring.

The current Loss associated with the pump was evaluated based on a repair/replace cost of $900.00 when the wear reached a 200 μM level. The design life of the pump is 20 years; the current wear rate of 28 μM/year results in an average loss of $84/pump (calculated via the Loss Function).

Factors (and their levels) that are thought to affect *wear* were identified:

Factor	Letter	Levels
Materials	A	2
Load	B	2
Surface Roughness	C	2
Clearance	D	2
Wall Materials	E	2

With only 5 factors (at 2 levels each), an L8 Orthogonal Array can be used, with two columns available for interactions (AxB and AxC were chosen). The column assignments are as follows:

	Columns' Assignment							
Trial	**A**	**B**	**AxB**	**C**	**AxC**	**D**	**E**	**Result**
1	1	1	1	1	1	1	1	**(8 Points**
2	1	1	1	2	2	2	2	**Taken for**
3	1	2	2	1	1	2	2	**each Part -**
4	1	2	2	2	2	1	1	**R1, R2, .. . R8)**
5	2	1	2	1	2	1	2	
6	2	1	2	2	1	2	1	
7	2	2	1	1	2	2	1	
8	2	2	1	2	1	1	2	

In this case eight trials were run (one pump per trial) and the wear measured on eight positions per sliding part. Signal-to-Noise Ratios were calculated for each trial as follows:

$$\eta = -10\log\left(1/8\sum R_i^2\right)$$

The results of the experiments and associated Signal-to-Noise Ratios are:

Trial	R1	R2	R3	R4	R5	R6	R7	R8	S/N
1	12	12	10	13	3	4	16	20	-21.90
2	6	10	3	5	3	4	20	18	-20.60
3	9	10	5	4	2	1	3	2	-14.77
4	8	8	5	4	3	4	9	9	-16.48
5	16	14	8	8	3	2	20	33	-24.15
6	18	26	4	2	3	3	7	10	-21.71
7	14	22	7	5	3	4	19	21	-22.96
8	16	13	5	4	11	4	14	30	-23.27
							Total		-165.85

The simplicity of the Signal-to-Noise Ratio evaluation procedure appears here. The Supplementary Table is prepared by adding up all the S/N Ratios for each factor (and interaction) at each level and dividing by the number of times a trial was run at that level (i.e. the average S/N Ratio for each factor level). The "trial" optimum levels are bolded:

Factor	Level 1	Level 2
A	**-18.44**	-23.02
B	-22.09	**-19.37**
AxB	-22.18	**-19.28**
C	-20.94	**-20.52**
AxC	**-20.41**	-21.05
D	-21.45	**-20.01**
E	-20.76	**-20.70**

Since interactions exist, they must be considered in establishing the optimum factor levels. For the AxB interaction, Level 2 is "optimum." The Orthogonal Array shows that, when AxB is at Level 2, the A and B factors are at *opposite* levels. Therefore, picking Level 1 for factor A and Level 2 for Factor B is consistent with the Level 2 AxB interaction.

For the AxC interaction, Level 1 is "optimum." The Orthogonal Array shows that, when AxC is at Level 1, the A and C factors are at the *same* levels. Since Level 1 has been picked for factor A (with a large S/N Ratio), Level 1 for factor C will be selected (*Note that factor C does not have a significant effect on wear; this exercise is almost "academic."*).

The optimum condition to minimize pump wear is given by A1, B2, C1, D2 and E2. The next step is to *predict* the *process average.* If all the optimum factor levels are averaged together, this tends to over-predict the average. An ANOVA of the S/N Ratios could be performed with only the significant factors included in the average. Taguchi recommends picking about half the factors, whose S/N Ratios differ "greatly" from the average effect (obtained by averaging any two factors). The bolded "delta's are the important factors:

Factor/Level	S/N	Average	Delta
A1	-18.44	-20.75	**2.31**
B2	-19.37	-20.75	**1.38**
(AxB)2	-19.28	-20.75	**1.47**
C1	-20.94	-20.75	0.19
(AxC)1	-20.41	-20.75	0.34
D2	-20.01	-20.75	**0.74**
E2	-20.70	-20.75	0.05

The predicted process average is now:

$$\hat{\mu} = (-18.44 - 19.37 - 19.28 - 20.01) - 3 \times \frac{-165.85}{8}$$

$$\hat{\mu} = -14.9db$$

A confirmatory experiment should be run at the optimum conditions. If the S/N Ratio obtained from this experiment agrees with the predicted process average, then the experiment can claim success from two standpoints. First, reproducibility has been demonstrated. Second, there has been a significant reduction in Loss (by calculating the optimum to current variance ratio through the *gain* (the difference between current and optimum S/N Ratio), the Loss is reduced by a factor of 1/8.28 or to $10.14 from $84).

In addition, a pump should be fabricated at the *worst* conditions and the wear measured (this is referred to as the *standard condition*). The quality level can then be estimated by comparing the optimum to worst conditions (i.e. the gain described above).

14.4.8 Parameter Design Applications

The Parameter Design approach can be applied to many different design problems. These include:

Case	Description	Example
Larger is Better	Loss is decreased when the quality characteristic is larger	Strength of Material
Nominal is Best	A target value exists, with loss of function possible on either side of the target	Current, Voltage, Thickness
Attributes	Qualitative assessment of the quality characteristic	Appearance judgments (bad, OK, good),
Dynamic Characteristics	Systems where the output is changed as a result of signal factors	Radio volume, gas pedal, throttle valves
Passive Functions	Systems which receive signals	Antenna, radiation, temperature, pressure, flow instruments
On-Off Systems	Systems which turn on or off depending on the input signal	Thermostatic control, automatic actuation system, relays (also results of "GO/NO GO tests")
Digital Systems	Systems which output in “1’s” and “0’s”	Computer systems
Vector Quantities	Systems whose output is multi-dimensional	Color copiers, printers, numerically controlled machinery

14.5 Multi-Generational Planning

Learning Objectives

- Be able to develop a Multi-Generational Plan to support sequential, competitive product release

Unit Contents

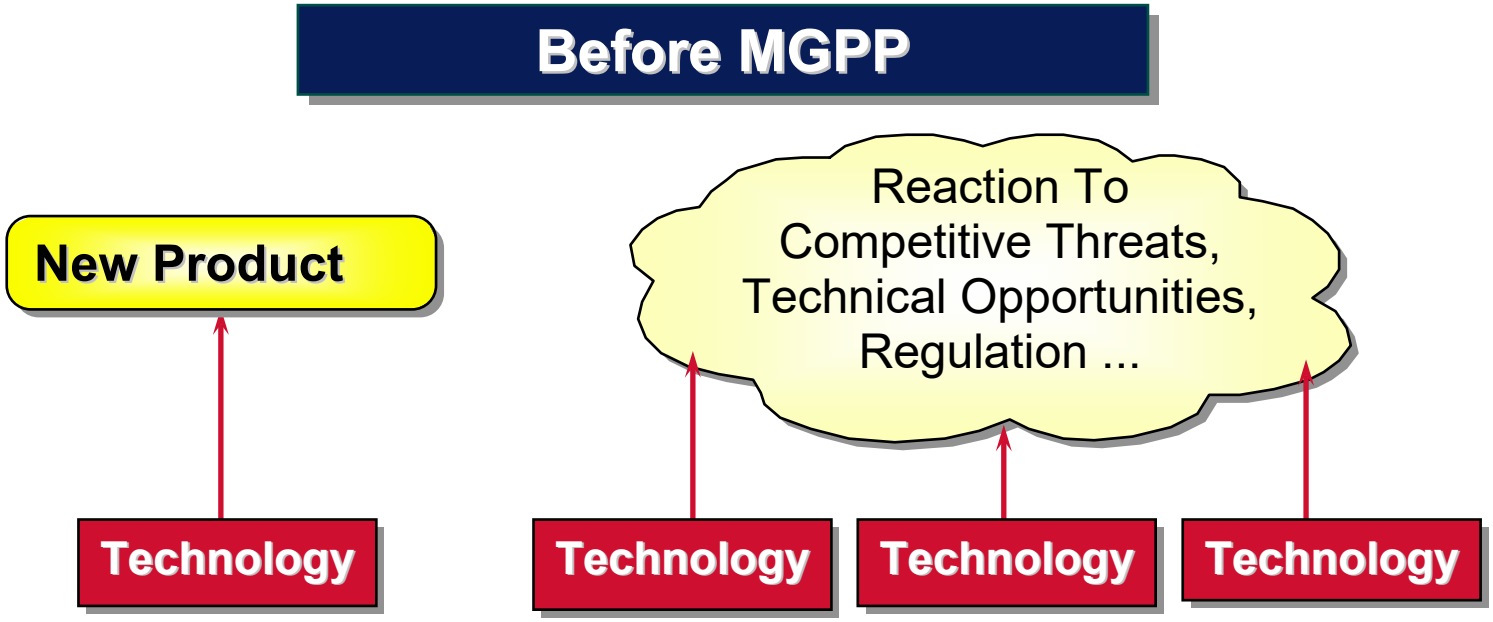
- Multi-Generational Product Planning (MGPP)
- MGPP Characteristics
- MGPP Development Process

14.5.1 Multi-Generational Product Planning (MGPP)

Very few companies achieve long-term success with just one product. Competitive response and increasing customer expectations can quickly make a successful product obsolete and force the company into a reactive, catch-up development mode. In addition, the company is continually faced with the need to identify, develop (and/or acquire) new technologies that can be employed in new or existing products. The ***proactive*** Product Development process, then, should continually look to the future – where should we go with our products (business goals), where are our competitors going (competitive strengths and weaknesses) and where would our customers like us to go (market demand)? What technologies will we need to accomplish these goals?

The key tool to accomplish this proactive approach is the Multi-Generational Product (or process, or service) Plan (MGPP). The Generational Plan is a:

- σ Sequenced, Multi-Year Product Development Which Is ...
- σ Driven By Anticipated Market Evolution And Is ...
- σ Built Around Core Technologies, Core Competencies, Or Core Platforms



Benefits of Proactive MGPP Approach

. . . to the Customer

- σ Saves The Customer $
- σ Gives Customer New Capability
- σ Makes It Easier For Customer
- σ Improves Customer's Skills

. . . to the Business

Product Platforms

- σ Provides A Product Or Process Base Which Can Be Leveraged Over Several Years
- σ Establishes The Basic Product Architecture For Follow-On Derivatives (Hardware Or Software)
- σ Is Designed For Expansion Into The Next Generation Product
- σ Is Designed For Adaptability For Custom Solutions
- σ Enables Base-Plus-Feature Up Pricing

Resulting Competitive Advantage

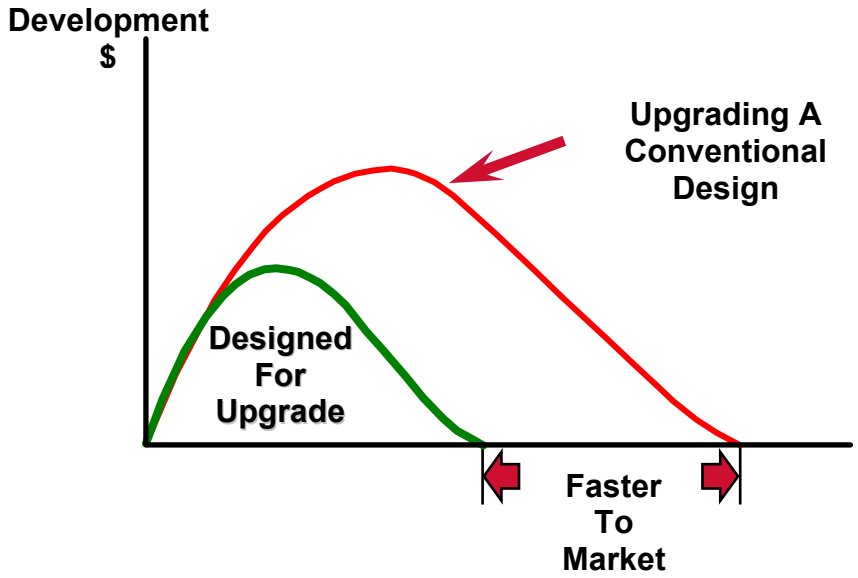
- σ Lower Cost Of Developing Future Generations
- σ Faster Introduction Of Product Upgrades
- σ Lower Risk Of Cost Or Time Overruns
- σ Cost Effective Customization Of Functionality
- σ Increased Customer Responsiveness While Maintaining Margins

Effects Of No MGPP

- σ Short Term Mentality – No Leverage: Every Product A New Product
- σ Reactive Product Development – Often Forced Into A "Me Too" Strategy
- σ Longer Time To Development With Increased Risk Of Vulnerability To Market Changes
- σ Higher Risk Of Consumer Rejection
- σ Reduced Ability To Keep Up With Technology Evolution
- σ Increased Difficulty And Cost Of Product Development, Product Variety, And Customizing Products
- σ Inefficient Resource Allocation And Usage

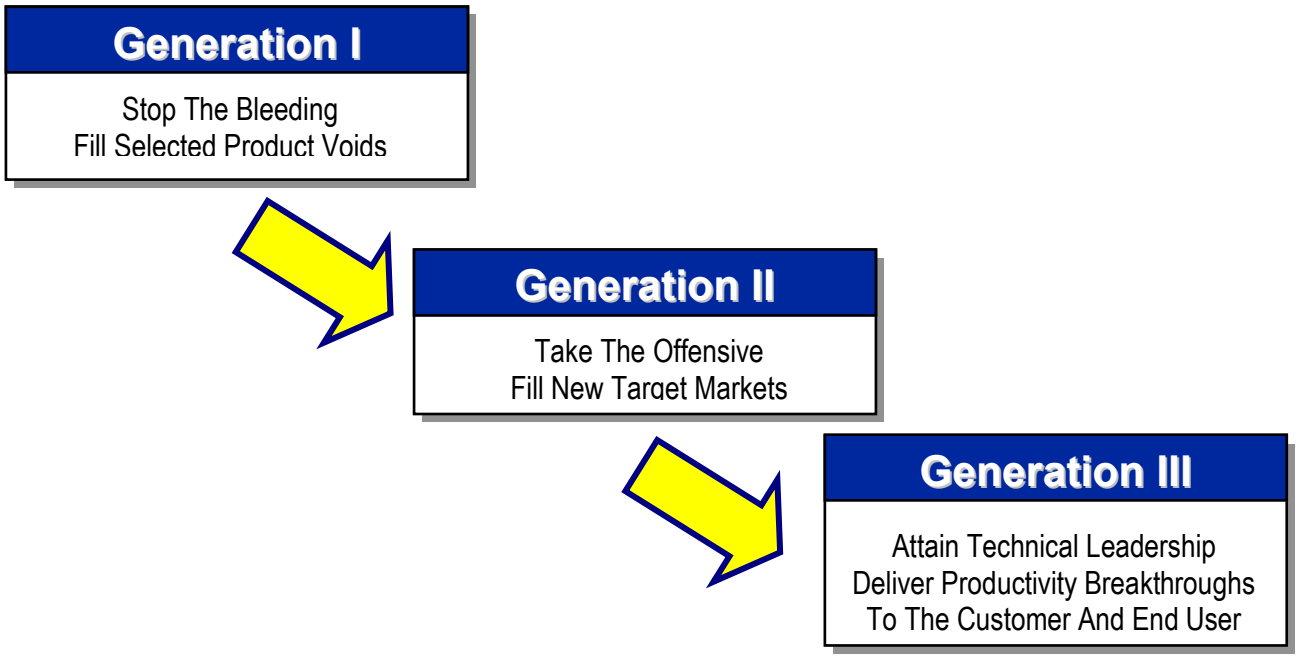
14.5.2 MGPP Characteristics

- ***A Vision Of Long-Term Direction*** - For Each Key Product Line Based On Anticipated Evolution Of The Market And Competitor Products Beyond Current Applications
- ***Product Generations*** – Series Of Releases; Each Release Characterized By Distinct Combinations Of Features Or Technology (At Least 3 Generations Planned)
- ***Assessment Of Technology Capability And Risk*** – Ability To Execute The MGPP With Current Technology And The Risk Associated With Developing The Plan
- ***Platforms*** – Product/Process Technology Competencies Which Can Be Leveraged To Introduce A Number Of Product Generations Quickly And To Reduce Cost Of New Product Development
- ***Technology Plans*** – Identification And Execution Of Technological Developments, Including Platforms, Needed For Each Generation

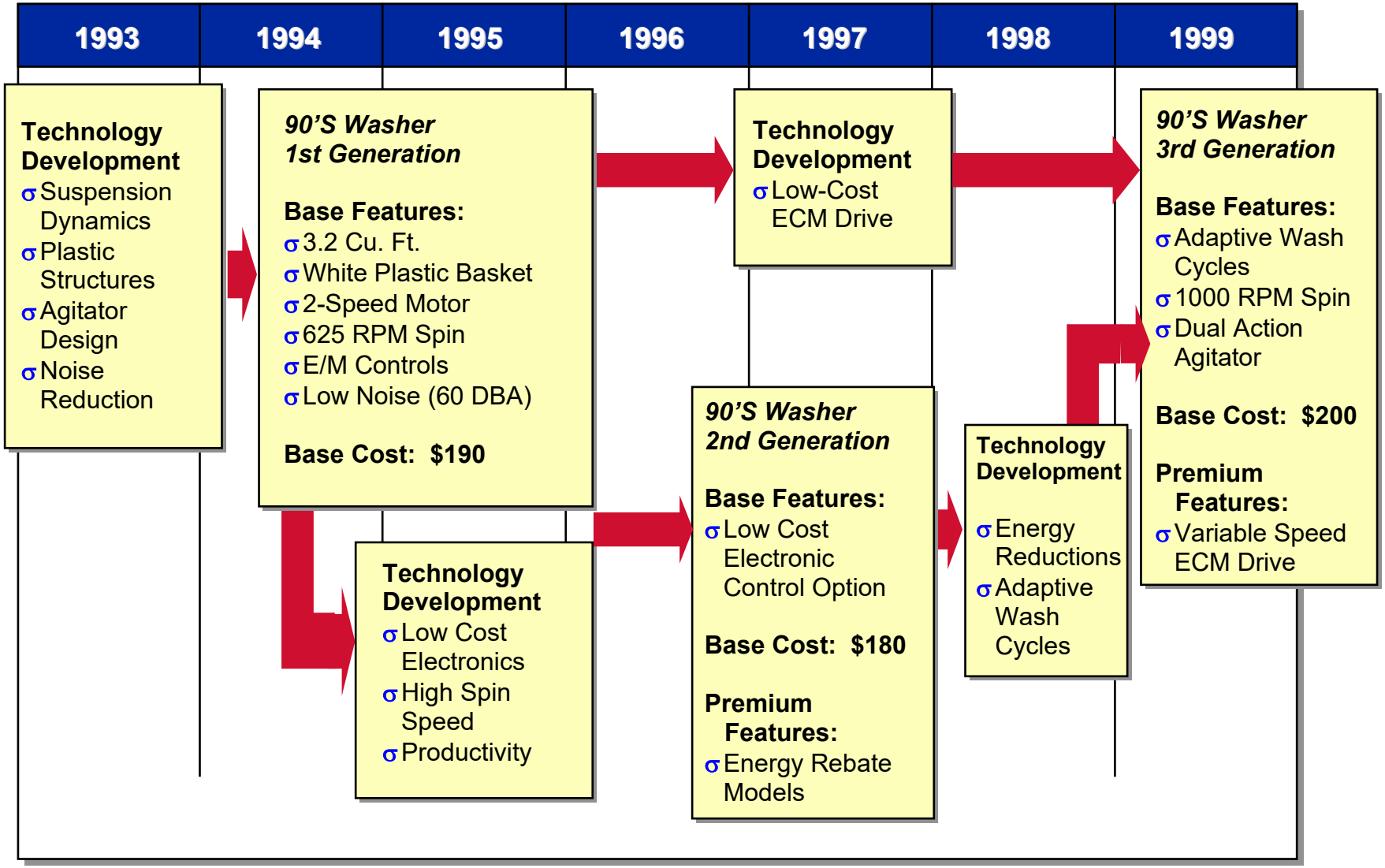


14.5.3 MGPP Development Process

Step	Tools & Methods
1. Identify Customer Driven Attributes	σ Voice of the Customer (Unit 4.1), σ Quality Function Deployment (Unit 14.1)
2. Define Current Industry Standard For Each Attribute	σ Benchmarking (Unit 14.3)
3. Envision Level/Status Of Each Attribute 3–5 Years Out	σ Benchmarking (Unit 14.3), σ Voice of Customer (Unit 4.1), σ Brainstorming (Unit 3.2)
4. Plan Intermediate Product Releases	σ Product Planning (Unit 2.4)
5. Identify Technology Required To Support Each Product Introduction	σ Concept Design (Unit 14.2) σ Technology Plan (Below, Unit 14.5)
6. Assess Risk Of Each Technology Development, Recognizing Time/Resources Constraints	σ Technology Risk Assessment (Below, Unit 14.5)
7. Prepare Resource Plan To Implement	σ Project Planning (Unit 3.1)



Example: MGPP for Clothes Washers



MGPP Format

Multi-Generational Plans may take a variety of formats; a typical MGP (*for an **Airline Check-In Service**)* is shown below. Note that the three major elements of vision, series of generations and supporting platforms/technologies are included.

	Generation 1	Generation 2	Generation 3
Vision	• Exceed desired "Customer Service" score per survey (would recommend/use again) • Clean passenger arrival traffic flow - No confusion on where to check in • No customer without a seat	• Separate walk-up and non-guaranteed passenger queue from reserved and guaranteed • Initiate Pre-Arrival check in • Update customer survey	• "Hertz Gold" type check-in available via phone, internet and airport kiosks • Real-time seat availability and readiness
Service Generation	• Clear signage/visual queues to guide passengers to check-in • Cross training of all staff on handling of arrivals • "AirTeam" incentives tied to their customer sat. score	• Photo ID/Express check-in for those with guaranteed reservations and no changes (e.g. seat or upgrade requests) • "Upgrade" features for frequent passengers and Express passengers	• All airline staff can perform all check-in procedures
Platforms/ Technology	• Integrated interface for check in staff - instant paperwork & pre-test key • Preassign seats in system until reservation (guaranteed only) cancelled	• Passenger names called in from curb-side luggage stand • Partner with area hotels for pre-arrival check-in kiosks at hotels	• Link reservation database to travel agent's and airlines/rail/hotels for passenger arrival information and changes (corporate & meetings • "Hertz-like" mobile check in devices (will also generate key) available for airline staff

Technology Planning

The MGPP brings together the following three product development processes/functions:

1. Product Plan…….....Sequence Of Product Introductions

2. Technology Plan.....……To Support Product Plan

3. Resource Plan.....……To Implement Technology

Many companies only focus on # 1 – the Product Plan. Advanced Development's work (i.e. identifying and developing robust new Technologies) is often disconnected from the needs of product development. The example below shows how the MGPP helps drive the identification and prioritization of technology development.

Medical Imaging Devices – Impact of MGPP on Technology Development Priorities

To support future generations of an imaging device sold to hospitals, the company identified and prioritized their technology development activities. Note that these priorities did not coincide with what the advanced development group "thought" were the most important (translate: *most fun*) technologies to work on.

1. Development Of 1.5D With 128 Channels Is On The Critical Path For Rad 2 And Card 1, To Be Introduced 3Q20XX, And Should Have The ***Highest Priority*** In Resourcing

2. Automatic Boundary Detection And Profusion Imaging Are On The Critical Path For Card 2, To Be Introduced In 199X. Coded Excitation And >80% Bandwidth Are On The Critical Path For Rad 3, Also To Be Introduced In 20XX. These Four Developments Should Have The ***Second Highest Priority*** In Resourcing

3. The Following "Quantum Leaps" Are Important But Have ***Third Highest Priority*** In Near-Term Development Resourcing

- 256 Channels
- 2D Transducers
- Phased Aberration Correction
- Vector Flow

Technology Developments Required Prior To Each Generation

Here, the technologies needed to support each product generation are identified. These become the priorities of advanced development (or technology acquisition).

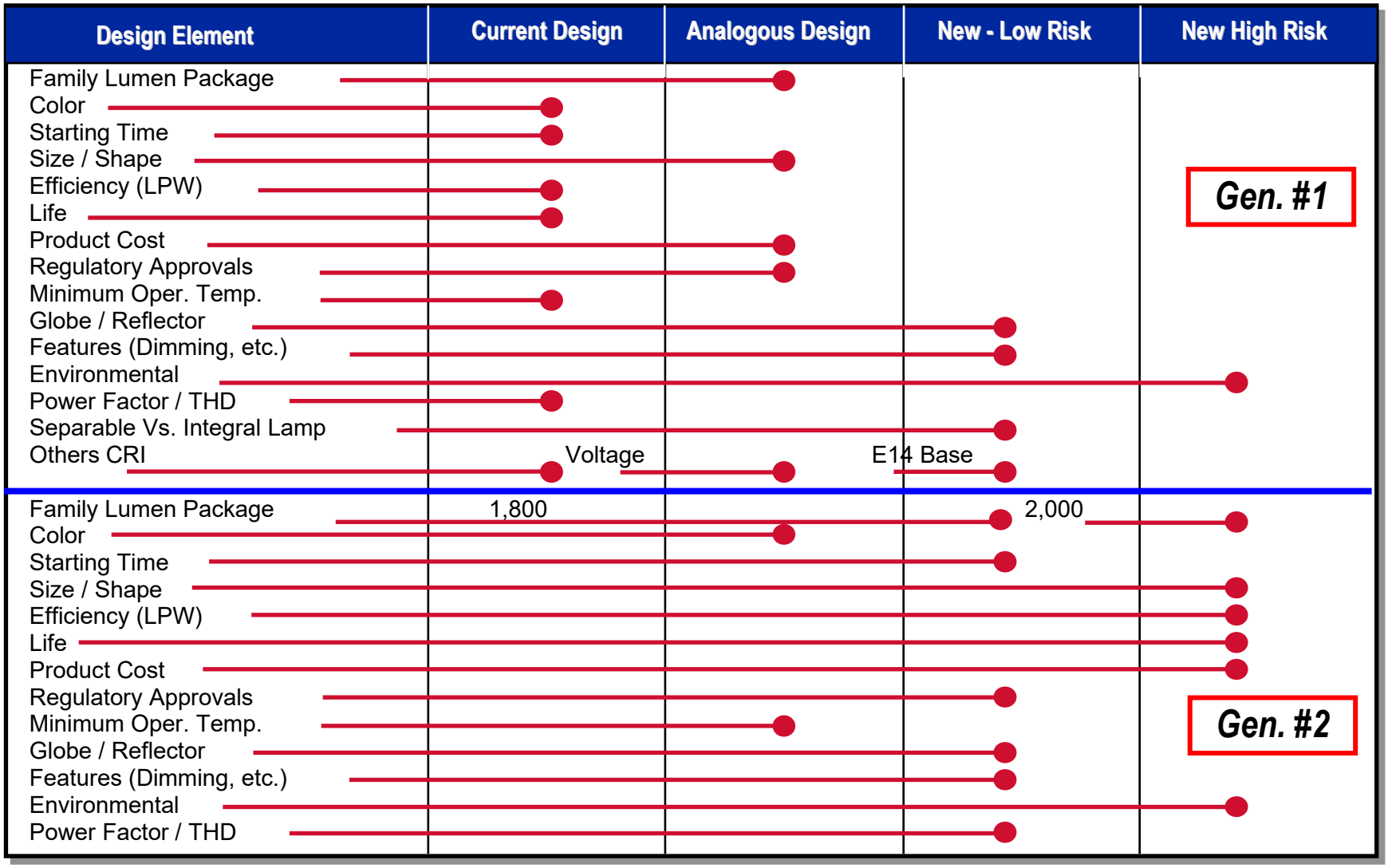
Example: Compact Fluorescent Light

Prior To Generation I	Prior To Generation II	Prior To Generation III
■ New Cathode Process ■ New Manufacturing Process And Equipment ■ New Lamp Design ■ Delta Repackaging ■ Low Dose	■ New Bending Process ■ New Packaging ■ Upgraded Phosphor ■ New Electronics Design/ Features ■ Bridging Process Modified ■ IC Development ■ Instant Start ■ Planar Magnetics ■ New Glass For High Wattages ■ New Base Change For Cap Change	■ New Glass Process And Glass Composition ■ New Lamp Design For High Lumens ■ New Cathode ■ New Packaging ■ New Electronics Design/ Features ■ DLC Capacitor ■ High Density Power Supply ■ Power Groove With Cold Spot Control ■ Fingerless Design

Technology Risk Assessment

In addition to identifying the needed technologies, the MGPP should include an assessment of the risk associated with the development effort as shown in the example below:

Compact Fluorescent Light - Example Technology Risk Assessment



14.6 Exercises

Objective: Motivate the need to improve your company's Design and Delivery System.

Instructions:

1. Brainstorm a list of product and service problems your company has experienced over the last two years. Prioritize these and pick the top 5 – 10 (*Suggestion: use N/3 to reduce your brainstormed list*). (10 min.)

2. Perform a "quick" root cause analysis of these – try to identify which step in the Design & Delivery system was responsible for the problem (and why). (10 min.)

3. Brainstorm countermeasures – what could have been done differently to *prevent* the occurrence of the problem – e.g. what changes to the Design & Delivery system would you suggest. (10 min.)

Time: 30 minutes

Objective: Distinguish between customer needs, specifications, and product features.

Instructions: 1. Review the description of the laptop computer and battery shown below

2. Classify the list as follows:
   - *Customer Need* – A function required by the user of the product
   - *Specification* – A performance requirement placed on the design
   - *Product Feature*– A specific choice of component or part applied in the design

3. For the specifications and product features, what customer need(s) do you think Dell is trying to meet?

30 minutes

Description of Dell Computer Corporation Laptop Computer:

- Intel® Pentium® II Mobile Module microprocessor with MMX™ technology with 32 KB of internal cache
- 512 KB of pipelined-burst SRAM external cache
- Hardware-accelerated PCI and AGP bus architecture that increases system performance; particularly video and hard disk drive performance
- 32 MB of SDRAM system memory minimum, with support for a maximum of 192 MB
- Ultra DMA/33 data transfer protocol for ATA /IDE hard-disk drive interface. Ultra DMA/33 allows data transfer rates of up to 33 MB/sec
- A combination module that contains a CD-ROM drive and a diskette drive
- Built-in stereo speakers and microphone
- Jacks for connecting external speakers, headphones, or an external microphone to the computer
- S-video TV-out connector and composite TV-out adaptor cable that allows you to connect a television to your computer
- An ATI video controller with an AGP 2X, 4 or 8 MB of video memory, 3D assist, dual-screen video, and flicker-free TV out
- A lithium-ion main battery and an optional secondary battery to double battery life
- Two power management modes ¾ standby mode and save-to-disk suspend mode ¾ that help conserve battery power
- A 14.1-inch active-matrix XGA color display

- A built-in keyboard that includes two special keys that support the Microsoft® Windows® 98 operating system
- A PS/2-compatible touch pad that provides the computer full mouse functionality
- USB capability, that simplifies connecting peripheral devices such as mice, printers, and computer speakers. The USB connector on your computer's back panel provides a single connection point for multiple USB-compliant devices. USB-compliant devices can also be connected and disconnected while the system is running.
- An options bay in which you can use a variety of combination modules, including the CD-ROM drive and diskette drive module, or the DVD-ROM drive and diskette drive module. In addition, you can use the options bay for a second battery.
- An infrared port that permits file transfers without using cable connections.
- PC Card slots with connectors for two 3.3- or 5-V cards. Both PC Card slots support CardBUS technology. In addition, a ZV Port is available from the lower slot (slot 0).

Detailed Description of Laptop Battery:

- Type - lithium ion
- Dimensions:
    - Height 23.8 mm (0.94 inch)
    - Depth 148.2 mm (5.83 inches)
    - Width 116.2 mm (4.58 inches)
    - Weight 0.60 kg (1.32 lb)
- Voltage 14.4 VDC
- Capacity - 60 WH
- Charge time (approximate):
    - Computer on 2.5 hours
    - Computer off 2.0 hours
- Operating time (approximate, with no power management features enabled)
    - 2.5 to 3.5 hours with one battery
    - 5 to 7 hours with two batteries
- Life span (approximate) 500 discharge/charge cycles
- Temperature range:
    - Charge 0° to 40°C (32° to 104°F)
    - Discharge 0° to 60°C (32° to 140°F)
    - Storage -20° to 50°C (-4° to 122°F)

Objective: Practice Planning to Gather Customer Needs for a New Product/Service.

Instructions: 1. Marketing needs your help designing a survey to be administered to potential customers of a new home heat pump. They have 20 questions regarding the importance of product features to the customer. They are wondering how large a sample size they need to obtain to feel confident in the results of their survey. You have a meeting with them in 15 minutes. What questions will you bring to the meeting?

Time: 15 minutes

Objective: Practice Planning to Analyze Customer Needs for a New Product/Service.

Instructions: As part of the survey design, you decided to use a 1 – 7 Likert scale to assess the preferences of the customers for the heat pump's features (1 – strongly negative, 7 – strongly positive). Based on your recommendation, Marketing sent out 1200 surveys to households across the country.

1. Suppose only 150 surveys were returned. What concerns would you have? What actions would you take?

2. Suppose that 1100 completed surveys were returned. How would you analyze the results of the individual questions? Show the equations you would employ.

Time: 15 minutes

Objective: Develop a House of Quality for a product or service.

Instructions:

1. Identify a product or service with which you are familiar.
2. Identify some customer requirements (*Subject-Verb-Modifer*) for the product or service. Use the structure tree to make them specific.
3. Rate the importance of each requirement from 1 to 5.
4. Identify product/service requirements (*Subject-Verb-Metric*) for each customer requirement (*Note: there may be multiple requirements per customer requirements*).
5. Put the customer and product/service requirements in a House of Quality matrix and complete the relations matrix (*Note: Use the Excel spreadsheet provided*).
6. Calculate the priority score from the importance and the strength of the relationships.
7. Set targets for the critical product/service characteristics you've identified.

Time: 60 minutes

Objective: To practice translating product/service level requirements to functional or system requirements.

Instructions:

1. Develop a functional block diagram for the product/service you identified in the previous exercise

2. Deploy the product/service level requirements to the functions. First, qualitatively evaluate which functions/systems will impact which product/service requirements.

3. Attempt to *allocate* the product or service requirements (CTQs) to the functions or systems *quantitatively.* (*Note: you will have to develop an understanding of how the function or system's performance relates to that of the overall product or service to do this.*) You can use the template below or the Excel spreadsheet provided.

Time: 40 minutes

Functional Diagram

Functional or System Level Impact on Product/Service Requirements (CTQs)
Functional Components

CTQs	Importance							
Impact								
Functional req'ts								

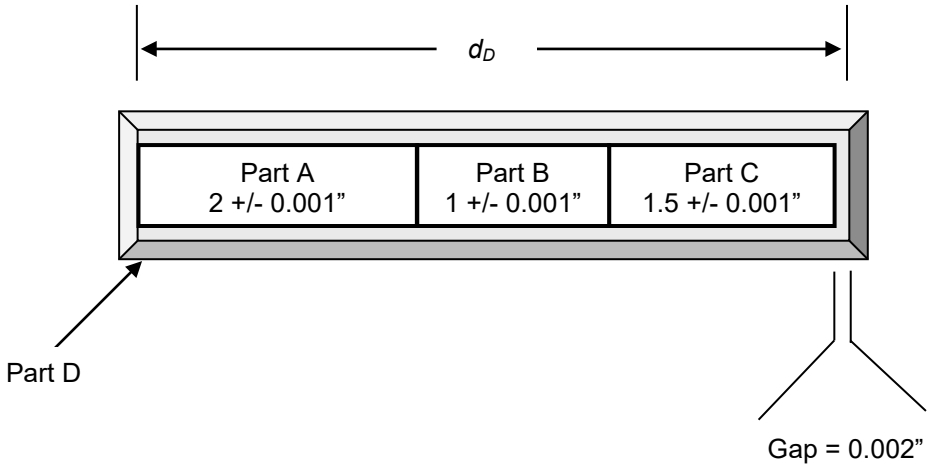
Objective: To practice setting tolerances for low level part characteristics.

Instructions:

1. *Worst Case Analysis* – If a gap of 0.002” is desired between the parts A, B and C and Part D, what should the **nominal** value be for dimension d_D? Part D can be fabricated to a tolerance of 0.001”.

2. *Root Sum of Squares* – If the assembly is made from parts randomly selected from the production line, what fraction of the assemblies will be defective because of excess interference (assume that Part D’s dimension d_D’s nominal value is set based on the worst-case analysis above)?

Time: 30 minutes



Objective: To “practice” creativity through Attribute Listing.

Instructions:

1. Pick some product that is available to you in the classroom.
2. Identify its parts and features.
3. Identify the current attribute that describes the feature or how that part is constructed (materials, size, etc.).
4. Brainstorm ideas to “improve” that part/feature.

Time: 30 minutes

Objective: To “practice” creativity through Forced Analogy.

Instructions: *(Stolen shamelessly from a Home Improvement episode)*

1. Pick a familiar home workshop tool or kitchen appliance.
2. List its attributes and/or features.
3. Force analogies from the tool/appliance to your spouse or significant other
4. Take this exercise home and show the results to them (just kidding!).

Time: 30 minutes

Objective: To practice identifying parameters to be compared in a parametric analysis.

Instructions:

1. Identify (e.g. brainstorm or identify by inspection) the parameters of a laptop personal computer.
2. Which sets of parameters would you expect to relate “logically?”
3. Which sets of parameters, if related, might provide unusual insights into competitive laptops?

Time: 25 minutes

Objective: To practice evaluating design concepts using the Pugh Concept Selection process.

Instructions:

1. You are trying to identify the best concept for cleaning your clothes. Your CTQs are as follows:
   - σ Residual Dirt (5)
   - σ Clean Scent (5)
   - σ Minimal Clothes Damage (4)
   - σ Minimal Color Bleeding (4)
   - σ Cycle Time (3)
   - σ Wrinkle Free Clothes (3)
   - σ Cost/Cycle (3)
   - σ Environmental Impact (4)

2. You have brainstormed the following concepts (feel free to add more ideas):
   - σ Beat clothes with a stick at the river (use this as the Datum)
   - σ Spin-cycle washing machine with detergent
   - σ Home dry-cleaning machine
   - σ Take your clothes to mom's house

3. Perform a Pugh Concept evaluation.

Time: 30 minutes

Objective: To practice evaluating design concepts using the Pugh Concept Selection process.

Instructions:

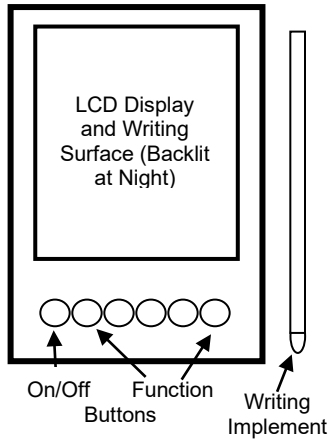
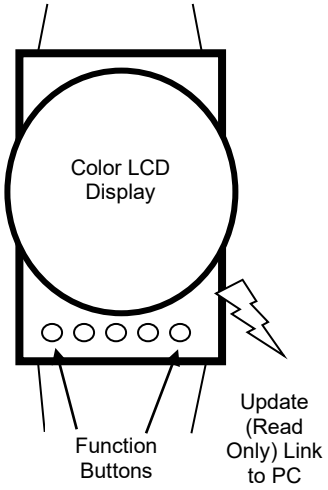
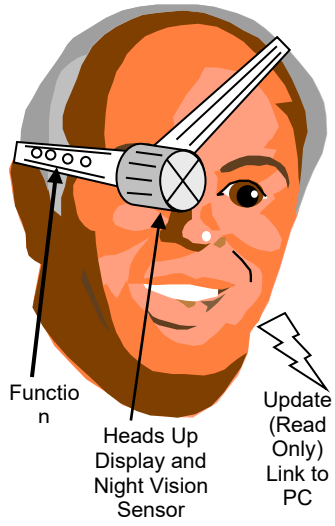
2. Examine the three conceptual designs shown on the next page.

3. Review the evaluation criteria shown below (if you are unclear about any of the criteria, ask your instructor for clarification).

4. Perform the first iteration of a Pugh Concept Selection process. Make assumptions where necessary. Which concept do you think is superior at this stage of the evaluation? (*Note: Use the Shirt Pocket Concept as the Datum and the Excel spreadsheet or a flipchart for the evaluation.*)

Time: 30 minutes

Personal Organizer Product Concepts:

Shirt Pocket (*The Sherlock Holmes*)	**Wrist-Watch** (*The Dick Tracy*)	**Strap-on Eyepiece** (*The Jean-Luc Picard*)
		

Evaluation Criteria:

Criteria	Wt.		Criteria	Wt.
Attractive	2		Makes Me Look “Techy”	4
Stores Addresses, Calendar, To-Do’s	5		Easy to Read	4
Hard to Lose	2		Quick to Reference Data	4
Light Weight	3		Easy to Hold/Wear	5
Easy to See in all Lights	3		Low Cost	1
Easy to Update	4		Easy to Design – Hardware	2
Easy to Use	4		Easy to Design - Software	2

Test yourself on these situations. Which is the best analogy to the process being benchmarked?

Process	Key Characteristics/Criteria	Possible Analogies
1. Brewing Beer	Chemical Process	**A.** Making Soda Syrup **B.** Manufacturing Shampoo **C.** Packaging Juice
2. Forming Uranium Pellets	Pressure Molding, Powder Molding	**A.** Extruding Pet Food Pellets **B.** Running X-Ray Equipment **C.** Manufacturing Aspirin
3. Retooling Production Lines In Real Time	Quick Reconfiguration	**A.** Changing Costumes Between Acts **B.** Servicing A Race Car During Pit stops **C.** Preparing Fast Food
4. Applying A Finish To Firearm Cartridges	Surface Smoothing	**A.** Polishing A Stainless Steel Knife **B.** Lacquering A Brass Lipstick Case **C.** Varnishing The Heel Of A Rifle

For the answers, delete the gray box below:

Objective: To practice calculating Loss Functions

Instructions: Develop Loss Functions for individual products and for the "population" of products for the scenarios shown below:

Time: 20 minutes

Product/ Characteristic	Specification or Function Limit	Loss at Function Limit	Current Population
Valve Stem Length	12" +/- 0.05"	$300.00	σ = 0.02"
Submarine Hull Steel - Tensile Strength	> 10,000 psi	$2 billion	σ^2 = 1.56E-10
Printed Circuit Board Chemical Contamination	< 0.1%	$1.00	σ^2 = 6.4E-3

Objective: To practice calculating Loss Functions

Instructions: A process that makes frictional clutch plates for automatic transmissions has an average scrap rate of 0.25%. The drawing specifies an overall thickness of 0.125 in. +/- 0.10 in. The cost to scrap the part is $2.50. What is the k value in the loss function?

What would the average loss per part be if the above process was centered at 0.120 inch, with a standard deviation of 0.00133 inch? Repeat for a process centered at 0.125 inch, with standard deviation of 0.00133 inch and one centered at 0.125 inch with standard deviation of 0.0033 inch.

Time: 20 minutes

Objective: To practice calculating Signal-to-Noise Ratios

Instructions: Calculate Signal-to-Noise Ratios for the scenarios shown below:

Valve Stem Length (inches)	Submarine Hull Steel Tensile Strength (psig)	Printed Circuit Board Chemical Contamination (% Contamination)
11.9848	77500	0.0754
12.0202	74456	0.0553
12.0089	80340	0.0996
11.9813	79876	0.0837
12.0054	78636	0.0853
11.9952	83191	0.0963
11.9948	82626	0.0829
11.9912	77770	0.0965
12.0079	78675	0.0751
11.9688	79868	0.0632
11.9989	77656	0.1041
11.9931	79215	0.0782
12.0054	77038	0.0764
12.0094	84754	0.0559
12.0151	80044	0.083
11.9906	80822	0.0742
12.005	79381	0.0786
11.966	78669	0.0676
11.9811	79049	0.0725
11.9788	79981	0.0825

Objective: To practice planning an industrial experiment.

Instructions: A screening experiment for an electronic circuit includes 8 factors (at two levels each). You would like to determine if the factors are significant and you suspect that interactions AxC, BxE, DxF & FxG exist.

1. Select an orthogonal array for the experiment.
2. Assign the factors to the columns in the array.
3. If any columns remain unassigned, what options are available for use of those columns.

Time: 30 minutes

Objective: To practice applying Taguchi's Parameter Design Technique.

Instructions:

1. Using the Card Drop Shop, design an experiment to optimize (e.g. *minimize*) the distance the card lands from the target. The only "production" constraints are that the card must be dropped from a standing position and that it must be aligned perpendicular to the ground. Factors you should consider in your experiment include: *weight (paper or binder clip), orientation (long side vs. short side down), stabilized (with and without a Post-it™ flag)* and *height (arm +/- 8 inches from horizontal).*

2. Once the experiment is designed, run 20 replications of each trial and analyze the results using Signal-to-Noise ratios.

3. Based on the optimum conditions discovered in the experiment, perform a confirmatory experiment.

4. Prior to starting the exercise, "predict" which set of conditions will result in the optimum conditions.

Time: 60 minutes

Prediction (Circle Your Best Guess):

Factor	**"Low" Level**	**"High" Level**
Weight	Paper Clip	Binder Clip
Orientation	Horizontal	Vertical
Stabilization	None	Post-it™ flag
Drop Height	+ 8 Inches	- 8 Inches

Loss Information: When the card drops a distance of more than 15" from the target, the customer experiences a loss of $1.50.

Objective: To practice applying Taguchi's Parameter Design Technique.

Instructions: The following experiments were run to determine factors important to reducing leaks in condenser tubing. An L-16 experiment was run with four factors being varied – tube cut, skill level, flux type and cleaning.

"Reverse engineer" this experiment. Take a blank L-16 array and determine which factors were assigned to which columns. What interactions will this design detect? (Hint: try to figure out which Linear Graph was used first).

Next, determine which of these factors are important. Analyze the results using the Signal-to-Noise Ratio approach described in Unit 14.3.

Time: 60 minutes

Tube Cut	Skill	Flux	Cleaned	Results
-1	-1	-1	-1	0.215
1	-1	-1	-1	0.269
-1	1	-1	-1	0.184
1	1	-1	-1	0.258
-1	-1	1	-1	0.146
1	-1	1	-1	0.344
-1	1	1	-1	0.107
1	1	1	-1	0.330
-1	-1	-1	1	0.200
1	-1	-1	1	0.430
-1	1	-1	1	0.123
1	1	-1	1	0.209
-1	-1	1	1	0.121
1	-1	1	1	0.311
-1	1	1	1	0.200
1	1	1	1	0.280

Objective: To practice applying Taguchi's Parameter Design Technique.

Instructions: The following experiments were run to determine factors important to improving temperature and pressure output of a J-compressor. An L-8 experiment was run with three factors being varied – eductor, pump and o-ring.

Determine which of these factors are important. Analyze the results using the Signal-to-Noise Ratio approach described in Unit 14.3.

Time: 60 minutes

RunOrder	Eductor	Pump	O-ring	Temp	Pressure
1	1	-1	-1	133.3	40.89
2	1	-1	1	142	37.66
3	-1	1	-1	135.15	67.5
4	-1	-1	1	134.2	51.2
5	-1	-1	-1	134.15	54.47
6	1	1	-1	133.5	49.23
7	-1	1	1	135	59.04
8	1	1	1	143	44.74

Objective: To practice developing a Multi-Generational Plan.

Instructions: Develop a Multi-Generational Product Plan (MGPP) for your product or one of the examples below. Use the format shown on page 14.5.7 for the MGPP. Include the development priority of new technologies and a technology risk assessment for same.

Examples:

- Your Product
- Computer Display
- Personal Digital Assistant (PDA)
- Mobile Telephone
- Refrigerator
- Power Tools: Drill, Circular Saw, Electric Screwdriver, Router
- Vacuum Cleaner
- Toothbrush

Time: 60 minutes

15.0 Reliability Management

Unit	Description	Page
15.1	Reliability Concepts and Management	15.1 - 1
15.2	Failure/Error Modes & Effects Analysis	15.2 - 1
15.3	Fault Tree Analysis	15.3 - 1
15.4	Quantifying Reliability	15.4 - 1
15.5	Root Cause Analysis	15.5 - 1
15.6	Exercises	15.6 - 1

15.1 Reliability Concepts and Management

Learning Objectives

- Understand core reliability concepts and definitions
- Understand how reliability activities are incorporated into the design & delivery process

Unit Contents

- Reliability Concepts and Definitions
- Reliability Tools and Methods “Fit” in the Design/Delivery System

15.1.1 Introduction

In this unit, we will focus on one Critical-to-Quality Requirement: *Reliability.* For many companies, reliability is a *critical* CTQ that can make the difference between corporate survival and failure. Extreme cases such as the space shuttle Challenger crash, the Bhopal chemical disaster, and Chernobyl are obvious reliability failures. But also consider how much of your company's profits are consumed by warranty costs (this is often a double-edged sword – your service staff are busy fixing the problems and thus have little time to pursue more profitable service work). We just received a notice that a small part of our home water heater may be defective – millions of these are potentially involved.

This unit introduces reliability terms and definitions; we'll replace the "common" meaning of reliability with operational definitions of *reliability, availability, maintainability, etc.* Then, we'll overview the "big picture" – describing the *management* of reliability and how the *tools* of reliability engineering can support your company's reliability goals. As engineers and managers of other companies and organizations have wrestled with the problem of reliability, they began to understand that a framework was needed to help the tools do their job. Over the last 30 years, the activities and events associated with planning, controlling and improving reliability have evolved into a logical, scientific system of reliability management.

We can manage reliability in different ways. One philosophy places the "responsibility" for reliability in the hands of reliability engineers or specialists. Here, reliability is treated as a special engineering discipline, and reliability-related activities must be integrated with those of other engineering disciplines (the Department of Defense *System Engineering* approach has used this method).

Another philosophy, promoted by the Total Quality Management "school," still makes use of specialists, but recognizes that *the best people to be using the basic reliability tools are the "line" engineers, technicians and managers closest to the systems*. Reliability is simply considered as one of many quality characteristics to assure in a product or service.

Both reliability management styles have resulted in successful products and services. The two styles have different characteristics, though, mainly in the areas of *responsibility* and *training.* An organization should consciously choose which style it will apply to its situation.

15.1.2 Terms and Definitions

Reliability Management has its own set of definitions and concepts. Some of the terms don't have *the same* meaning as they do in everyday English. Most of the terms are easy to understand, but we want to make sure that the subtleties are clear.

This section covers the *basic* concepts and definitions. As you begin to apply these concepts to your work, you may find that your business or field has developed specific definitions to describe and measure reliability. For example, in the utility industry, groups such as the Edison Electric Institute (EEI) and the Electric Power Research Institute (EPRI) have developed standard definitions for many different reliability measures. Military weapons programs have also developed specific definitions to describe reliability attributes of their systems.

Some of these definitions are useful to industry because they allow common *reporting* of reliability performance. They may or may not be useful to you as definitions that allow you to *analyze* the reliability of your particular system.

Reliability Building Blocks

Let's start at the beginning. Here are some terms that we will use as building blocks for our Reliability definitions.

TERM	*DEFINITION*	*EXAMPLES*
ITEM	Some "thing" for which reliability is a quality characteristic.	Nuts, bolts, relays, power plants
FUNCTION	The purpose of an item. What is the item supposed to do? The definition is simple, but determining the *specific* functions of an item may take some thought.	a) A power plant generates 910 MW of electricity. b) A fuse is expected to open at 10 amps of current.
TIME	The length of time the item is supposed to perform its function. Time can be expressed as clock time, operating time, cycles, number of demands, or even linear distance. In general, the "best" unit of time should correlate to the *stress* that is placed on the item that will eventually be expected to fail the item.	a) A diesel generator must start and run for a period of seven days. b) A valve is expected to cycle open and closed at least 10,000 times. c) A pen is expected to write at least 1 million feet.

TERM	*DEFINITION*	*EXAMPLES*
ENVIRON-MENT	The aggregate of all external and internal conditions either natural, man-made, or self-induced, that influences the performance, reliability or survival of an item.	Temperature, humidity, radiation, magnetic and electric fields, shock, vibration, etc.
SYSTEM	A composite of items, equipment, personnel and their associated skills and techniques capable of performing some function or mission.	Lateral, feeder, substation, personal computer, power plant, locomotive, air conditioning system
PROBABILITY	For this text, the notion of *relative frequency* will be adopted as the definition of probability. That is, by performing an "experiment" many times, the relative frequency of some event (number of times the event occurs divided by the number of experiments) will converge to the *probability* of the event.	a) The *probability* of rolling a *one* on a single, fair die is 1/6. b) The *probability* of picking the *ace of hearts* from a deck of cards is 1/52.
STATED CONDITIONS	The environmental conditions in which the item is expected to perform its function. Most generally, the conditions under which the customer will use the item.	a) A pump seal is expected to perform its function when the ambient temperature is above 33°F. b) An electrical device (i.e., Relay, Terminal Board, or Circuit Breaker) is expected to work in 100% humidity at 104°F. c) A laptop PC will be used in an office environment 80% of the time, will be transported in a briefcase (carried baggage) approximately 30 times a year and will be used outdoors 20% of the time.

TERM	*DEFINITION*	*EXAMPLES*
MISSION	A specific task or operation for which the system's function, time and stated conditions are clearly defined. Note: Some systems will not have a "finite" mission. Utility distribution systems, for example, are expected to continuously supply electricity to the customer. It may be more useful to understand what constitutes a *service interruption* to the customer. This is usually a function of both *power quality* and *continuity.*	a) The mission of a commercial aircraft is to take off, fly from the origination city to the destination city, and land, all in a safe manner (for both the passengers and the equipment), under "normal" weather conditions. b) The mission of an Auxiliary Feedwater System is to provide at least 325 gpm of water to one of three Steam Generators at 1000 psig for a period of 18 hours under a spectrum of South Florida weather conditions, including hurricanes with wind speeds up to 225 mph.
MISSION PHASES	When a system's configuration does not stay the same for the duration of the mission, *phases* are defined to help aid the reliability evaluation. A mission *phase* is a portion of the mission where the system's function is constant.	The aircraft *mission* in the previous example can be divided into three distinct phases: *Takeoff, Flight* and *Landing.* Think of how the aircraft "system" changes to respond to these different functional requirements. What equipment is needed for the *takeoff* phase that is not necessary during *flight*?

15.1.3 Reliability Concepts

Reliability Defined

Reliability in our everyday language has come to include the quality of a product, how strong the product is, how long it lasts, etc. There is, however, a definition of reliability that helps to differentiate it from other quality characteristics:

"The probability that an item will perform its intended function under stated conditions for at least a given period of time."

As a probability, *reliability* (applying this definition) is expressed as a number between 0 and 1. If a product is designed to work between 30°F and 120°F, and it fails when you use it at 150°F - the product may still be "technically" *reliable* because you used it under conditions other than those stated. As a customer of this product, though, you will not be happy if you need the product to operate at a higher temperature.

Reliability Measures

- **Reliability (R)** is, of course, one way to measure this concept. We can express the *starting* reliability of a Diesel Generator as 0.99. That is, we would expect the diesel to start about 99 out of 100 attempts. We could also describe its *operating* reliability as 0.98. That is, for a given mission time of 8 hours, the diesel will operate successfully at about 98 of 100 attempts (Thought question: Is this the same as saying "the diesel has a 98% chance of running for 8 hours?"). We also use the term *Unreliability* or *Failure Probability.* These quantities are simply the reliability subtracted from one (1).

- **Mean Time Between Failure (MTBF)** is the average time between failures of *repairable* equipment or systems. The time required to repair the equipment is not included in the calculation.

$$MTBF = \frac{1}{n}\sum_{i=1}^{n} t_i$$

where:

t_i - *time between failures*

n - *number of failures*

Note: This equation assumes all items in the sample have failed. If there are items still working, their operating times are included in the numerator of the MTBF equation.

- **Mean Time To Failure (MTTF)** is the average time to failure of *non-repairable* equipment or systems. The calculation is the same as for the MTBF:

$$MTTF = \frac{1}{n}\sum_{i=1}^{n} t_i$$

where:

t_i - *time to failure*

n - *number of failures*

Note: This equation also assumes all items in the sample have failed. If there are items still working, their operating times are included in the numerator of the MTTF equation. Also, time *may be measured in hours, cycles, etc.*

Maintainability Defined

Where *reliability* measures the equipment's time *between* or *to* failure, the time required to *repair* or *restore* the equipment to service is also important. This introduces the concept of **Maintainability** and its associated definition:

"The probability that a failed item will be repaired in a given time under stated conditions."

If the repair time is "acceptable" (within the expected time) then the product is said to have good maintainability. This definition, although generally accepted, does not give the reader a picture of all that is being addressed. After a system is designed and constructed, it is turned over to the operating and maintenance forces. As components begin to fail, the demand for maintenance on the system will rise. Although the maintainability *definition* only includes the time of repair, the *concept* of maintainability includes how many pieces of equipment in the system require repair (i.e., simplicity of design), commonality of equipment (i.e., how many different types, sizes and manufacturers of valves, transformers, insulators are in the system), how easy is it to detect, diagnose, repair and test the equipment, the skills of the maintenance workers, the available repair tools, available spare parts, accuracy of trouble-shooting and repair manuals

and all other support required for repair. Notice that some of these maintainability elements are the responsibility of the system's *designer*, in fact, the system design and construction play a major role in determining the maintainability of a system.

Maintainability Measures

Although the specific quantity of *Maintainability* can be expressed mathematically, it is rarely calculated. The measures listed below are commonly used for maintainability.

- **Mean Time to Repair (MTTR)** - The average time required to restore failed equipment to an operable condition. MTTR generally only includes the time spent detecting, diagnosing, repairing and testing the failed equipment. MTTR is often a subject addressed in contractual documents.

- **Mean Down Time (MDT)** - The average time a system or equipment remains in the failed state. The MDT includes the MTTR but also includes such times as administrative activities (ordering or prioritizing the failure), transport time (time required to travel to the failed equipment) and logistics time (time required to obtain spare parts). Contractual documents (i.e., Purchase Orders, Bid Specifications) do not usually specify an MDT since these additional times are generally outside the control of the equipment vendor. *Service Restoration* measures these components of maintainability.

- **Maximum Repair Time (MRT)** - The mean values listed above do not give us information about the dispersion of times to repair. If a histogram of repair times is available, then we can look for the time at which 95 or 99 percent of the repairs will be completed. This would be our MRT.

These measures can be used for both *corrective* (CM) and *preventive* maintenance (PM) activities.

Availability Defined

Our overall ability to make use of a system or equipment is given the term **Availability**.

*Availability (A) is the probability that a system is in an operable condition. In its simplest form, availability is expressed as the ratio of the total **Uptime** of a system to the total period time.*

Availability Measures

Availability is measured as a function of both reliability *and* maintainability. In fact, one way of calculating the *Average, Steady State System Availability* is shown below:

$$A = \frac{MTBF}{MTBF + MDT}$$

If a product fails (on average) every day and it takes (on average) 1 hour to get it back in service, how would you compare it to one that fails once a month and takes 30 hours to repair? The product that fails once a month has a *higher MTBF* (i.e. higher reliability) than the one that fails every day. This latter product also has a *higher MDT* (i.e. lower maintainability). When you consider the two products' *availabilities*, though, they are about the same.

Sometimes companies calculate an availability ratio and call it "reliability." At one electric utility, "Service Reliability" was defined as the average number of minutes/year the customer was without power – this is really a measure of availability. With your new knowledge of the "technical" definitions, you will now be able to tell the difference between *reliability* and *maintainability* and *availability*.

Failure Defined

All along we've been talking about **failures** without defining the term. A very broad definition of failure is:

A failure is an unsatisfactory condition.

This is a very "customer-oriented" definition of failure. If an item performs satisfactorily when first placed into service, *any identifiable deviation from the original condition that is unsatisfactory to a particular user* is a failure. Does this mean that anybody can walk up to a piece of equipment and declare it to be failed? Can Operations say that an item has failed and Maintenance disagree and say that it is working? This could lead to a rather chaotic situation!

An organization must come to agreement on how to operationally define failures in clear terms. This agreement will be based on the function of the item, the conditions under which the item is used, and how long the item must operate. We could try to use some of the terms we have already defined to make this definition a little clearer:

A failure occurs when a product ceases to perform its expected function, under stated conditions, for an expected period of time.

Consider two failure cases: In the first, a light bulb is expected to provide a certain (i.e. 100 watts) output. When the filament burns out, the light bulb "obviously" can no longer perform its expected function, and it has failed.

In the second case, a pump's expected function is to provide 325 gpm of water at 1000 psig discharge pressure. According to our failure definition, if the pump could only produce 324 gpm at the 1000 psig discharge pressure, we would consider the pump to be in a failed state. But the pump is still running, still moving water!

Has the pump failed? It has if we have agreed on the need for the pump to perform to a certain standard of performance.

As you can see, this situation can become complex, especially for components that degrade over time, rather than failing outright as our light bulb does (When do you consider your car tires to have “failed?”). There may be some "engineering calculations" that divide the component's world into success and failure states.

We will frequently use the term **failure mode**. This refers to the *consequence* or *effect* of the item's failure, such as a *wire open, short, valve fails to open, close, throttle, insulator arcing or flashovers*. This differs from the **failure cause** that is the physical, chemical, electrical, thermal or other process that results in the failure.

Customer Valid Requirements

We've just covered the major terms of Reliability. There are two other words that should be mentioned here - mainly because they have different everyday and technical meanings.

Dependability - Maytag washers are *dependable*. To the appliance owner, this implies that Maytags have a high reliability and need little maintenance.

Dependability is a measure of how successful a system is at achieving its mission. Dependability combines *mission reliability* with *mission maintainability*. That is, if our mission is to drive from Miami to Los Angeles in three days, *dependability* allows our car to break down along the way, if we can quickly repair it and complete the mission.

Durability - This is the layperson's "synonym" for reliability. In everyday usage, it implies some physical strength or something that lasts a long time. If gorillas jump on your Samsonite™ luggage after it's thrown from a moving train, and it doesn't pop open, it's considered durable. The Toyota™ that is driven for over 200,000 miles may be thought of as durable.

Technically, *Durability* is used as a measure of the useful life of a system. We might say that a car is designed for a durability of 20 or more years. Of course, there will be failures along the way, but if we repair these, the car will remain "useful" to us for that period of time.

Notice that these terms have one definition when used by our customer and another when used in a "technical" sense. Regardless of what term our customer uses, it is up to us to determine how they translate into the more concrete terms of reliability, maintainability and availability.

15.1.4 Reliability Management

The Challenge

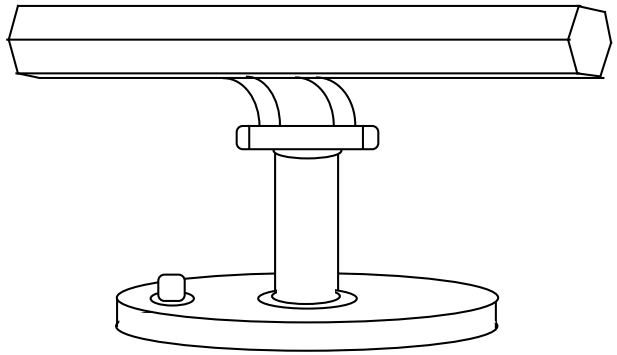
As we try to manage our product reliability, there are two major **elements** to consider. The first element is the *system* itself. An existing system will, by its design, exhibit an inherent reliability. It will age, degrade, experience stresses and it will ultimately fail. The system, though, is not just a collection of hardware and software. The system's reliability is influenced by how we operate and maintain it.

The second element of reliability management considers the *actions* that we take to operate maintain and upgrade or redesign the system.

A somewhat formal definition of reliability management can be stated as:

> *"The systematic activities that occur to ensure that the reliability required by the consumer is met or exceeded."*

Let's start with a simple system, a typical lamp that may be found in your office. What do we mean by the *reliability* or *availability* or *maintainability* (RAM) of the lamp? How can we *manage* its reliability (or maintainability or availability)?



DESK LAMP

Suppose the lamp is already sitting on our desk and we've been using it for some time. How is the lamp doing, reliability-wise? Well, every time we turn it on in the morning, it lights up our desk, and when we turn it off at night, it stops lighting our desk. Occasionally, though, a bulb burns out that we promptly replace from a box of bulbs in the storeroom. There have been one or two occasions when we've run out of bulbs and have had to go to the store to replenish our supply. It is rather difficult to replace the bulb, though. The tubular bulbs that fit this lamp are difficult to get out of the lamp and to replace. With the size of the bulb and the narrow slot of the light fixture, it's hard to get a grip on the bulb.

Managing Reliability (PDCA)

What can we say about the *reliability* of the desk lamp? Well, the first thing we could do is **measure** its reliability. Based on the definition of reliability (introduced, we would define a *function*, a *time* associated with performing the function, and the *environment* in which the lamp will perform its function.

Desk Lamp Reliability Requirement: The desk lamp must provide light for a period of 10 hours a day in a "benign, ground" environment. The reliability target for this lamp is 0.999 (Question: where did this requirement come from?).

Knowing this requirement, we could then count the number of times this function is not met. We could also measure the time to or between failures for the various components that failed. We would then report (to ourselves) the reliability of the lamp every month or quarter. This data collection is the beginning of our reliability management system.

Here's our failure report for the last two quarters:

Failure	**Date**	***Action Taken***
Light Bulb Burned Out (TTF = 1000 hours)	6/5	Replaced Light Bulb from Storeroom
Plug Loosened by Cleaning Personnel (time - N/A)	8/20	Pushed Plug Back in Socket
Light Bulb Burned Out (TTF = 1200 hours)	10/20	Replaced Light Bulb - Had to go to Pareto's Hardware Store for replacements

Based on this report, we can compare the *actual* reliability to the *targeted* reliability and see if there is a gap. If there is, we can do a **root cause analysis** of the failures and try to improve the reliability. Part of our root cause analysis would include understanding the failure patterns exhibited by the components. What is the probability distribution of the failures? What is the Mean Time To (or Between) Failures?

What would some improvement strategies look like? Would improving the life of the light bulbs help? Could we somehow predict the life of the bulbs and replace them *before* failure (say, at night, when the lamp is not required to be operational)? Notice that the first option would require either an improvement in the **design** or **manufacture** of the bulbs. The second option would involve changing the **maintenance** plan for the bulbs (presumably, we are now only performing *corrective* maintenance on the lamp).

Managing Maintainability (PDCA)

What about the *maintainability* of the lamp? Again, we could start by measuring the current maintainability, comparing it to a target value and looking for improvements. The maintainability record of the lamp for the last two quarters looks like this:

Failure	Date	Action Taken
Light Bulb Burned Out (TTF = 1000 hours)	6/5	Replaced Light Bulb from Storeroom **(Total Down Time = 15 minutes, Active Repair Time = 10 minutes)**
Plug Loosened by Cleaning Personnel (time - N/A)	8/20	Pushed Plug Back in Socket **(Total Down Time = Active Repair Time = 1 minute)**
Light Bulb Burned Out (TTF = 1200 hours)	10/20	Replaced Light Bulb - Had to go to Pareto Hardware Store for replacements **(Total Down Time = 65 minutes, Active Repair Time = 11 Minutes)**

When we compare the Active Repair Time of this lamp to others, it might seem excessive. What could we do about this? The lamp **design** makes it difficult to maintain. Could we change the design, or maybe build a tool to help us change the bulb easier?

In the second bulb failure, we had to incur a rather large Total Down Time, because our storeroom ran out of bulbs. Is there some way we could manage this (i.e. a min-max system in the storeroom - based on an understanding of bulb reliability and usage)?

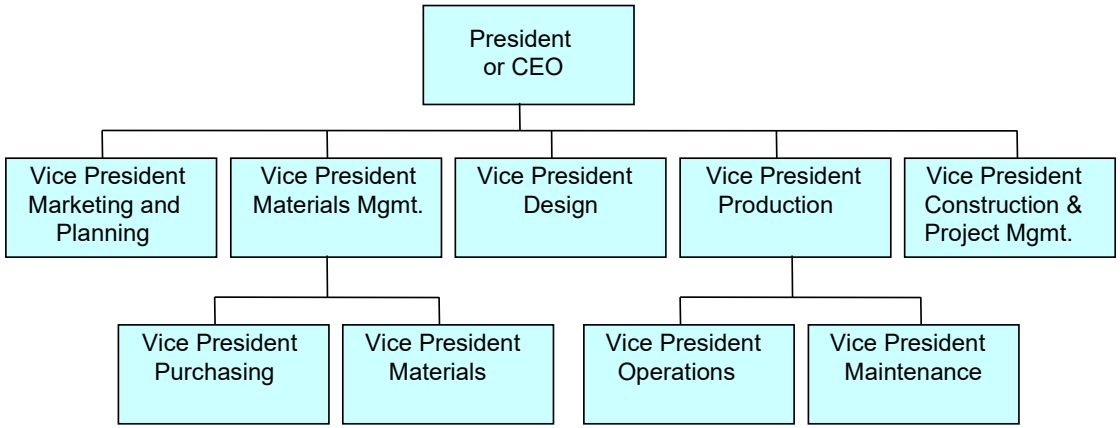
Returning to the original point, reliability management includes both the *physical system* as well as the *processes* that we use to *design, operate and maintain* the system. With a desk lamp, it is relatively easy to manage the reliability. We own the lamp; we manage it.

With complex systems and complex organizations divided by functions such as Planning, Engineering, Project Management, Operations, Maintenance, Stores, Purchasing, etc., reliability management is correspondingly complex. Often the very simple question of "Who is accountable for reliability?" is difficult to answer.

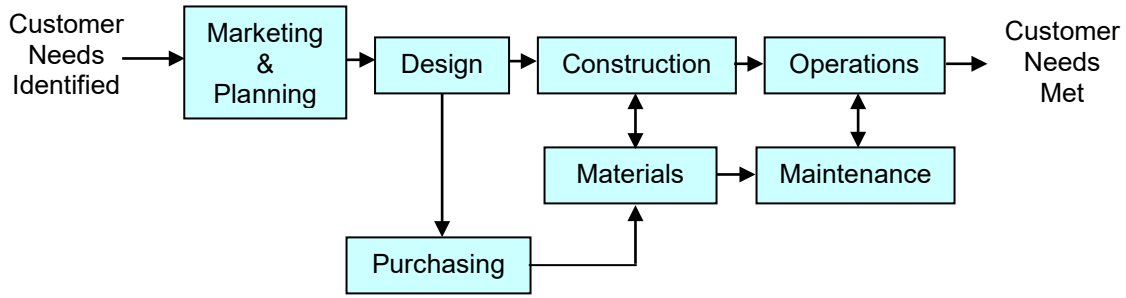
Reliability Management Processes

Cross-Functional Management

Most corporations are divided and *organized* into functional units. Here is a typical organization structure:



Notice that the organization has a *vertical* reporting structure. The VP-Operations reports to the VP Production who reports to the President or CEO. We have discovered, though, that the assurance of reliability requires activities that occur both *within* each function and *across* each function. If we were to line up the organization to show how they relate from a reliability perspective and include the customer, the picture could look like this:



There are two basic issues this type of organization must wrestle with. One, to assure reliability, the organization must assign reliability *accountabilities* to the functional activities (i.e. what are the responsibilities of the *Engineering* department) and second, how do these activities relate *across* the departments. This latter issue is known as **cross-functional management** and is generally the more difficult of the two to solve.

If you are in the Engineering department, how often have you blamed the Operations or Maintenance people for "screwing up" your design? If you are in Operations, when was the last time you complained about the crummy design of some system? In Maintenance, you have it worst of all; you're angry at Operations for breaking things and at Engineering for making it hard to fix what Operations "broke."

If a company has a reputation for consistently producing products of high reliability, you can almost be sure that they have come to terms with cross-functional management.

There is a third issue, learned over the years. A dollar spent in the Planning or Engineering phase to prevent a failure mode is worth many, many dollars "downstream" in Operations or Maintenance fixing the problem. Although this may seem obvious, this lesson has yet to be learned by many, many companies.

The remaining Sections will describe some of the reliability activities that can and should occur in a company interested in providing customers with reliable products and services. For this review, we will start at the beginning of PDCA, in the *Market Research* phase.

Market Research

In this stage of reliability management, the **customer needs** are determined and translated into **product requirements**. **Customer surveys** and **market analysis** are used to determine the needs of the market for the product or improvement to the product and the environment in which the product will be used.

When determining customer needs it is important to understand both the *visible* and *hidden* needs. For example, many utility customers who are on a fixed budget have a hard time paying their electric bill when they reach their peak usage months but have no problem making the payment during the low usage months. There was a visible need to remove this variation in the billing amount. To meet this need, one utility developed *budget billing* – a way of leveling out bills over the year so the customer received essentially a constant bill amount each month.

One tool for analyzing this data and translating the customer requirements into quality characteristics is **Quality Function Deployment (QFD – see Unit 14.1)**. QFD is both a tool and a method. As a tool (known as a “House of Quality”), it provides a design team with a systematic method of organizing customer needs, translating these into quality characteristics (with associated targets and specification limits). The “House” can include “rooms” which consider how the competition compares to your company (both quantitatively and in the perception of the customer). As a method, QFD is used to continue the deployment of quality requirements into the system or process being designed.

An assessment of the *environment* in which the product will be used and the conditions under which it will be used needs to be part of the market research. A few questions to address include:

- Will the product be used in a salt spray environment?
- Will it be expected to cycle?
- Will it be exposed to sunlight?

All other environmental or usage aspects that influence the product's strength or probability of failure need to be considered.

With the customers' needs understood and the environmental and usage conditions known, product planning can begin. Here's where a cross-functional "hand-off" occurs. The information gathered in this phase should be communicated to the other organizations involved with the product. If misinterpretation of the customer needs occurs the overall design may prove flawed, and/or the design time will be greatly increased.

Planning

The quality requirements determined through market research are often expressed in the customer's language, which is general and imprecise. The planning stage takes these requirements and translates them into a set of **product requirements**. These may be initially categorized into *quality, cost, delivery, safety, and corporate responsibility.*

At this point, further stratification takes place to identify specific performance requirements and the general system level reliability requirements. The overall reliability, maintainability and availability requirements need to be identified. **QFD** or **quality charts** are used to expand the quality requirements into a set of technical specifications.

These should be expressed in numerical terms and not in generalities. The statement "the pump should be highly reliable" does not adequately define the requirement. The requirement would be better defined by "the pump needs to be 95% available and should have a mean down time of 2 hours operating in outside ambient temperature limits of -40^{O}F to +115^{O}F with a pumping capacity of... etc."

Planning isn't just defining the quality and reliability aspects of the system you are going to produce. Planning also includes defining the activities we will undertake to assure the quality of the system. It is here that the schedules for design and implementation are developed. The steps for assuring that a quality product is designed, constructed, operated and maintained to meet the customer's requirements are defined.

For example, the planning schedule may require that an **FMEA** and an **FTA** be performed at a very early stage of the product design as an early warning system to detect possible problems. If the subsystem is going to be contracted out to a vendor, the planning schedule needs to include the steps to assure that the vendor is going to supply a product that meets the customers' requirements.

The result of the planning phase is a concept described in a planning schedule with reliability, maintainability and availability goals. This concept should be understood by all organizations involved with the product. A **preliminary design review** should be conducted and approved before proceeding to the design and development of the product.

Design and Development

At this point, the system level technical and reliability requirements will be known. Concept design will rigorously and creatively identify alternatives through the **Pugh Concept Evaluation Process.** As the design process proceeds, the system will be developed into subsystems, equipment and components. The system’s reliability requirement must be **allocated** to the lower levels. These can then be communicated to the hardware and software designers (or vendors).

Without specific reliability requirements for design, reliability becomes a vague and poorly defined objective for which no one can be and is held responsible. Even though the subsystem reliability requirements may be rough and subject to re-negotiation, they still give a picture of what is needed to meet the customer's requirements.

As the design work continues, efforts should be made to **predict** whether the system will meet the reliability requirements. Reliability predictions can be performed by the

- Parts Count Method,
- Reliability Block Diagram/Fault Tree Analysis
- Stress Analysis Method,
- Intersecting Stress and Strength Method or
- previously generated data.

The predictions identify design, application and environmental problem areas.

Predictions should be made as early as possible and updated whenever changes to the design occur. If the shortfall in reliability is found at an early stage it may be fixed by simplifying the design, de-rating a particular part or by various other means at a much lower cost than if the problem is uncovered after the system is constructed.

Allocation and **prediction** are interrelated tasks. The prediction is performed and compared with what was allocated for that system/subsystem/component. Shortfalls may indicate a need to repeat the reliability allocation. This becomes an iterative task. Predictions should not be considered a basis for determining that the reliability requirements have been met. Predictions are an early warning system. System testing will determine whether the reliability requirements have been met.

Fault Tree Analysis (FTA) and **Reliability Block Diagrams** can assist the designer in performing reliability predictions as well as providing the operator/maintainer with the knowledge of the critical paths. FTA is also one of the principle methods for performing safety analysis.

A **criticality assessment** needs to be conducted to determine which items are critical to the success of the system's functions. The primary method for doing this is a Failure Mode and Effects Analysis (FMEA). The results of an FMEA can be used as inputs to design trade-offs, safety engineering, maintenance planning, reliability testing, design of experiments, et cetera.

Once the drawing for the proposed design is completed an **intermediate design review** should be performed. This provides a forum for review of the design to ensure that it will be performed successfully, that it can be constructed at a reasonable cost, and that it is maintainable. Critical items for construction, operation and maintenance are identified. Not every designer has specialized knowledge in reliability, maintainability, safety, construction, and day-to-day system operation. The design review helps to overcome this.

After the intermediate design review is approved, an engineering prototype model may be constructed. **Environmental tests** and **accelerated life test** are performed to ensure that the system or subsystem can operate reliably under the specified usage environments and are in compliance with quantified reliability requirements. The statistical **Design of Experiment** may be used to determine the optimal operating configuration. This data is analyzed and a reliability assessment is performed. Critical components and processes are identified for control during construction, operation and maintenance. Once the test results are acceptable a final design review is conducted.

For many companies, much of the equipment and components that make up products and systems are purchased from vendors. Thus, the overall program for assuring that the customers' requirements are met needs to extend to the vendor. Policies regarding vendors need to be specific and provide guidance both inside and outside the company. Systems for vendor evaluation to determine if they can produce a reliable product need to be in place.

The same requirements hold true for a product being produced in-house as for a product being produced by a vendor. The vendor must be given full understanding of the use of the product. An early warning system for evaluating the design, and a system for testing to assure that the requirements are met are essential. The vendor should provide information on critical items, operation and maintenance. The design review teams should consist of members from the purchaser and the vendor. Warranties may also be included as part of the contract.

Manufacturing/Construction

During manufacturing (or construction), the design drawing is translated into hardware and software of the system. **Quality controls** are used to ensure that the delivered product will be in conformance to the design. Critical items defined in the design and development phase are highlighted for control. The design drawing is base lined. If appropriate, *statistical* quality control methods are employed to manage variability in manufacturing processes.

Field change control is used to ensure that changes required during the construction receive review and concurrence by Engineering, Operation and Maintenance. A **configuration management** system is necessary to document the as-built

system and control any changes that occur thereafter. **Qualification testing** is performed to demonstrate that the system can meet the reliability and maintainability requirements. If burn-in is required, it is performed before the system is turned over for operation.

Operation and Maintenance

The Operations and Maintenance departments provide input, during the design phase, to the development of the operation and maintenance procedures. This typically occurs in the design review process. **Error Mode and Effects Analysis (EMEA)** is used to ensure that potential, high-probability errors are eliminated from critical procedures.

Process control systems are used to monitor parameters that were identified during the design process as being critical.

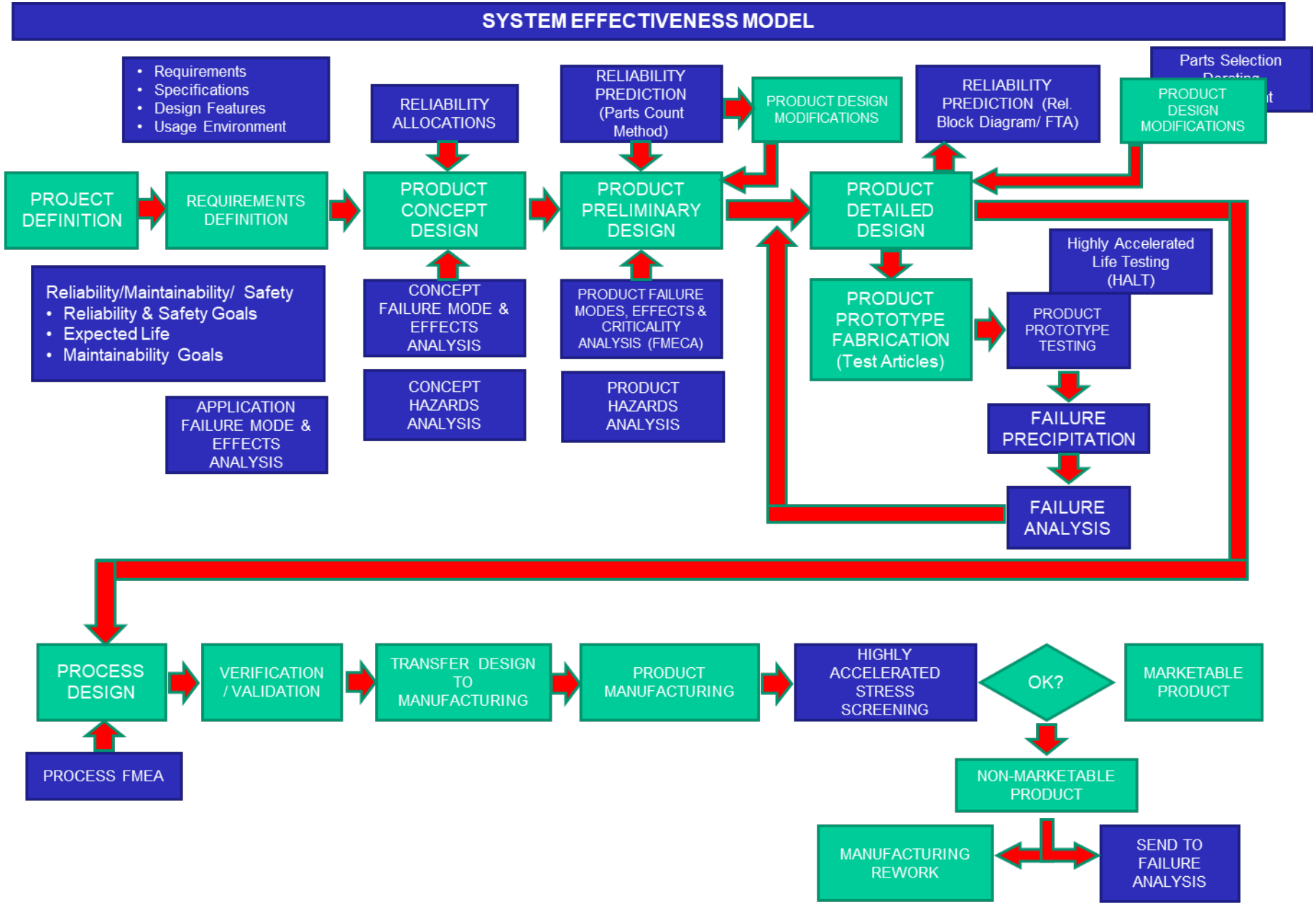
The Design Review, Failure Mode and Effects Analysis and Fault Tree Analysis are the tools that the designer uses to transmit the understanding of the system operation and failure characteristics to the operations and maintenance departments. These tools provide input to a **Reliability Centered Maintenance (RCM)** analysis to aid in determining what parts will fail, how often they will fail, what parts are critical to operation, and what kinds of spare parts need to be stocked. The maintenance plan is determined from the RCM analysis.

As the system is operated, failures will occur. Operations and maintenance prepare a **Failure Data Collection System** to capture the important reliability data. These failures are analyzed to determine their **root cause**. The Failure Mode and Effects Analysis and the Fault Tree Analysis are helpful tools in determining the root cause of a failure.

The failure data is fed back to the engineering group responsible for the design. If a design change should occur, Configuration Management is used to ensure that all critical documents (drawings, specifications, procedures, training material, etc.) are updated. A Design Review is used to ensure that all organizations understand and concur with the change.

Weibull Analysis can be used to analyze the failure data. Weibull analysis describes the failure distribution, and it can be used for setting preventive maintenance schedules, performing risk analysis, determining the need for redesign, and determining the required spare parts.

The PDCA cycle continues for the life of the system. For an example of one company's quality & reliability activities performed as part of their New Product Development process, see the next page.


SYSTEM EFFECTIVENESS MODEL
• Requirements
• Specifications
• Design Features
• Usage Environment
RELIABILITY ALLOCATIONS
RELIABILITY PREDICTION (Parts Count Method)
PRODUCT DESIGN MODIFICATIONS
RELIABILITY PREDICTION (Rel. Block Diagram/ FTA)
Parts Selection
PRODUCT DESIGN MODIFICATIONS
PROJECT DEFINITION
REQUIREMENTS DEFINITION
PRODUCT CONCEPT DESIGN
PRODUCT PRELIMINARY DESIGN
PRODUCT DETAILED DESIGN
Reliability/Maintainability/ Safety
• Reliability & Safety Goals
• Expected Life
• Maintainability Goals
CONCEPT FAILURE MODE & EFFECTS ANALYSIS
PRODUCT FAILURE MODES, EFFECTS & CRITICALITY ANALYSIS (FMECA)
Highly Accelerated Life Testing (HALT)
PRODUCT PROTOTYPE FABRICATION (Test Articles)
PRODUCT PROTOTYPE TESTING
APPLICATION FAILURE MODE & EFFECTS ANALYSIS
CONCEPT HAZARDS ANALYSIS
PRODUCT HAZARDS ANALYSIS
FAILURE PRECIPITATION
FAILURE ANALYSIS
PROCESS DESIGN
VERIFICATION / VALIDATION
TRANSFER DESIGN TO MANUFACTURING
PRODUCT MANUFACTURING
HIGHLY ACCELERATED STRESS SCREENING
OK?
MARKETABLE PRODUCT
PROCESS FMEA
NON-MARKETABLE PRODUCT
MANUFACTURING REWORK
SEND TO FAILURE ANALYSIS

15.2 Failure/Error Modes & Effects Analysis

Learning Objectives

- Be able to perform Failure Modes and Effects Analysis on products and processes.

Unit Contents

- Failure Modes Effects and Analysis (FMEA)
- Process/Error Failure Modes and Effects Analysis

15.2.1 Failure Modes & Effects Analysis

This unit will introduce one of the most useful tools of reliability management, the Failure Modes & Effects Analysis (FMEA)[1]. Here is a simple, but key insight into reliability management:

*To ensure the **reliability** of a system, we must manage the system's **failures**.*

For simple systems (e.g. a *simple system* could be a wire or cable, an insulator, a fuse, or a manually operated valve), the failures that can occur are generally easy to list. The wire can fail through an *open* circuit or through a *short* circuit. The important or *critical* components of the wire are its conductor and its insulation.

As systems become more complex, the critical components or operations often cannot be identified through simple inspection. The design process may split up the system into sub-systems, with the design responsibility being given to different groups (disciplines or vendors). The sub-system engineer may have a difficult time assessing the effect of failure of his components on the whole system.

Originally developed in the 1950's, Failure Modes & Effects Analysis (FMEA) has evolved over the years as the basic method of tying the system's pieces together by examining the effects of each component failure on the ultimate system functions. Here are some of the FMEA's many applications:

- Checking (and improving) the system's design as it exists on paper.
- Identifying those critical components that required extensive testing to determine their reliability.
- Planning the system's maintenance program to determine preventive and corrective maintenance actions.
- Improving the reliability of existing systems by identifying, ranking, and improving critical components of systems or processes.

[1] Note: Per Mil-Std-1629, the FMEA "Bible," a Failure Modes and Effects Analysis (FMEA) is a failure analysis *without* prioritization of the failures. A Failure Modes, Effects and Criticality Analysis (FMECA) includes such a prioritization via frequency, severity and/or other scoring factors.

FMEA Logic

The Failure Modes & Effects Analysis is an *inductive* logic analysis.

FMEA asks the basic question: What if this happens? The "this" refers to the specific failure mode (and cause) of a component of the system. The "what" refers to the effect of the failure mode on the performance of the system's functions. This basic question is asked of each component in the system and is the "Failure Modes, Effects" portion of the FMEA.

If information related to the frequency of failure, the time to repair, and the ability to detect the failure (either by work crews during construction or the system operators/maintainers) is available, then the component failures can be prioritized by their importance.

Management can then assign resources (for system redesign or preventive maintenance) based on the importance of the failure to the system. This is the "Criticality" portion of the FMEA.

An example of an FMEA worksheet follows. There is no "standard" format for the FMEA.

Example FMEA Worksheet

SYSTEM: Autotransformer (Xmer)	**SUBSYSTEM:** High Pressure Trip	**DRAWING:** 5610-E-1333, sheet 2

ITEM	FUNCTION	FAILURE MODE	CAUSE OF FAILURE	EFFECT OF FAILURE	FREQ.	SEVERITY	DETECTION	RPN	ACTIONS
Pressure Switch	Detect High Oil Pressure	Fails Low	Miscalibration Corroded Contacts Drift Plugged Sensing Line	Local: Failure to send trip Signal to Xmer protective Circuitry System: Failure to trip Xmer on High Pressure Conditions - possible Xmer Case rupture	Very Low (1)	High (8)	Quarterly Circuit Check (2)	Low (16)	None
		Spurious Trip	Miscalibration Drift	Local: Spurious Trip Signal to Xmer protective circuitry System: Loss of Power to 3 Feeders, 2500 Customers Affected.	Low (2)	Medium (5)	On Trip (10)	High (100)	Redesign Circuit, Use 2 out of 3 trip logic
Trip Relay	Trip Xmer on High oil pressure	Fail to transfer	Coil open Coil Short Circuit Contacts Corroded	Local: Failure to send signal to lockout Relay System: Failure to trip Xmer on High Pressure Conditions - possible Xmer Case rupture	Medium (4)	High (8)	Quarterly Circuit Check (2)	Medium (64)	Inspect Relay Contacts Monthly, Coil Continuity Check
		Spurious transfer	Relay jarred during maintenance Contact Spring failure	Local: Spurious Trip Signal to Xmer Lockout Relay System: Loss of Power to 3 Feeders, 2500 Customers Affected.	Medium (4)	High (10)	On Trip (10)	Very High (400)	Remount Relay - relocate away from high maintenance areas

RPN – Risk Priority Number (Product of Frequency, Severity, Detection)

FMEA Definitions and Building Blocks

Definitions

Here are the terms you should understand to perform a FMEA.

TERM	*DEFINITION*
Mission Profile	A description of the system's functions and the environments it experiences in all phases (start-up, operate at load, shut down), including criteria for performing the functions successfully.
Indenture Level	The level at which the system is being analyzed. *Higher* indenture levels represent a *higher* level of detail for the FMEA (i.e. to "lower" levels of the system). For a power plant, the indenture levels would proceed from *unit (lowest),* to *system,* to *subsystem,* to *equipment,* to *component,* to *part (highest)*.
Failure Mode:	The ways in which the item can fail. For example, a relay's functions are to open and close upon demand (and stay open or closed). Failure modes could be: "fail to open", "fail to close," or "spurious transfer."
Failure Causes	The possible physical, chemical, electrical (and other) causes associated with each failure mode. For the relay's failure mode of "failure to open upon demand," the possible causes could include magnetized contacts, burred contacts, coil insulation failure, broken wire in coil, etc.
Failure Effect	The consequences of the failure mode on the operation or function of an item. Failure effects are classified as: • **Local effect,** the consequence of the failure mode on the item being analyzed, • **Next higher-level effect**, the consequence the failure mode has on the next *lower* indenture level, • **End effect**, the consequence the failure mode has on the *lowest* indenture level (i.e. system).
Failure Frequency	The expected rate of occurrence of the individual failure modes (how often the failure mode is predicted to occur).
Failure Severity	A quantitative measure of the effect of the failure mode. It is a relative measure of the worst potential effect of the failure mode.
Failure Detection	The means or methods by which a failure can be detected or predicted by an operator during normal system operation or can be detected by a maintenance crew by some diagnostic action.
Criticality Assessment	A procedure by which each failure mode is ranked with respect to the other failure modes to determine its criticality in the system. This is typically determined as a function of failure severity and failure frequency, but other factors, such as ease of detection, can also be included.

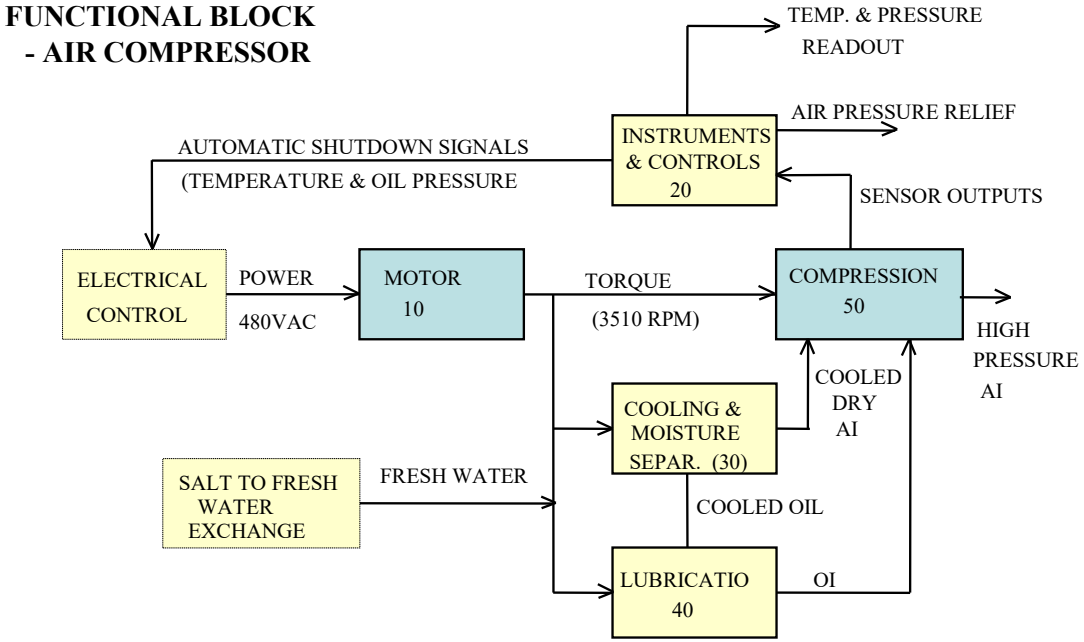
Block Diagrams

Block diagrams are often prepared to support a FMEA. They provide the analyst with the ability to trace the effects of failure modes through all levels of the system. For all but the simplest systems, it is difficult to accurately model a system *without* the use of block diagrams.

Functional block diagrams depict the functional flow of "energies" through the system. Reliability block diagrams depict the series/parallel success configuration of the system.

Functional Block Diagrams

Functional block diagrams (FBD) show how the system's components relate to each other based on their *functions.* Blocks are used to represent the different components. Arrows ***out*** of the block are used to show what is *provided* by the component. Arrows ***into*** the block show what is *supplied* to the component by other functions.

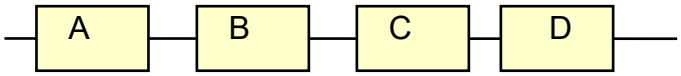


Reliability Block Diagrams

Reliability block diagrams (RBD) graphically depict the set of components (and their configuration) required for system *success*. RBDs help the analyst understand how a failure of a component or subsystem can affect the overall system.

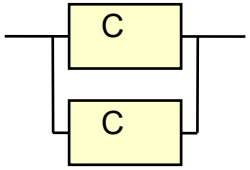
For many systems, *all* components must function for the system to succeed in its mission. The reliability block diagram would show this relationship by joining the components together in ***series.***

SERIES SYSTEM



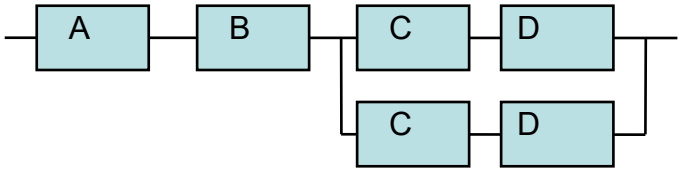
In some systems, redundancy (or backup) is provided for critical devices. The reliability block diagram shows this relationship by joining the components together in ***parallel.***

PARALLEL SYSTEM



Often, a system will incorporate a mixture of series components and redundant components. Redundancy may be employed for those components whose individual reliabilities are not judged adequate to meet the reliability requirements of the system

SERIES - PARALLEL SYSTEM



Keeping Track of Components & Failure Modes

Some method of coding or naming the components of the system and their associated failure modes should be adopted for the FMEA. For example, electric utilities generally adopt "tagging" systems for power plants, transmission and distribution systems. If the analysis is performed on an existing system, then the existing tagging scheme probably makes the most sense to adopt for the FMEA.

The tagging systems typically name equipment or component-level items in the system (e.g. Motor-Operated Valve MOV-749). If the FMEA will penetrate to a lower indenture level, then some additional coding system will be required.

This coding system needs to be consistent throughout the FMEA to provide a logical method of identifying each item with its next higher assembly. One system for doing this is by assigning a number to each item in the functional and reliability block diagrams at the initial indenture level and expanding on these numbers as the FMEA progresses through different levels of the system.

The numbered reliability block diagram appears below for the high-pressure air compressor.

RELIABILITY BLOCK DIAGRAM
- AIR COMPRESSOR



Analyzing the FMEA at this level would produce failure modes for each of the subsystems. For example, if the motor has three failure modes, they could be identified as:

10-1 FAIL TO START
10-2 FAIL TO RUN
10-3 INSUFFICIENT TORQUE

The *higher* indenture level parts are also identified so that the item being analyzed can easily be traced up through its associated *lower* indenture levels. Again, a systematic method for coding should be adopted. For the High Pressure Air Compressor example, the lubrication subsystem could be broken down into equipment as follows:

EQUIPMENT	LABEL
OIL RESERVOIR	41
OIL HEATERS	42
MAIN PUMP	43
FILTER "A"	44A
FILTER "B"	44B
OIL COOLER	45
OIL PIPING	46

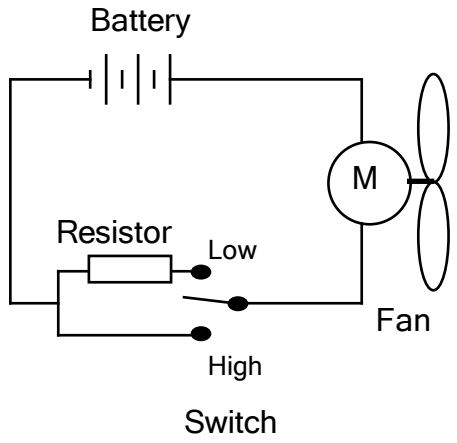
In many companies' *equipment* is generally "tagged," but components and parts are often not. The FMEA analyst must then develop a coding system for these *higher* indenture level items.

Ground Rules and Assumptions

As the analysis is planned and implemented, various ground rules and assumptions will be made. These will include some of the items already discussed, such as indenture level, functions or missions analyzed, definition of system failure (especially for equipment which degrade over time), environmental conditions as well as others. As a matter of good FMEA (and engineering!) practice, these assumptions should be documented. They will help guide the analysis and facilitate the review of the FMEA.

FMEA Procedure

The basic FMEA process is straightforward. Each single item failure is analyzed to determine its effect on the system. The failure modes are ranked in order of criticality, and action items are assigned. The process steps for completing an FMEA will be explained by working through a *two-speed, portable fan* example:



1. UNDERSTAND THE SYSTEM FUNCTIONS

The first step of the FMEA is the description of the system to be analyzed and a definition of its mission. Recall that the definition of Reliability is *the probability that an item will perform its intended functions for a given period under stated conditions.* The functions, time, and the stated conditions must be well understood to analyze the effect of each item's failure on the system. Developing a mission profile can help here. If necessary, the mission profile should separate the mission into *phases*. Generally, a phase should be an operating mode in which the following conditions apply:

- The performance functions required of the system do not change.
- The system configuration required for satisfactory performance does not change.
- The environment in which the system operates does not change.
- The definition and consequences of system failure do not change.

For example, a steam turbine's mission may be partitioned into several phases (start-up, operate at constant load operate at variable load, shutdown). The turbine's equipment performs different functions in each of these phases. The turbine's turning gear *is* required for the start-up and shutdown phases, but not for the operating phases.

If the FMEA is to be performed on the system's entire mission, each phase should be analyzed separately (a failure of an item may have a different consequence in the start-up phase than it does in the operate at constant load phase).

Sometimes, systems have a **standby** mode. The FMEA may also consider failures occurring in this mode.

We can't emphasize enough how important this step is. Experience has proven time and again that the need for clear understanding of the system's function(s) is vital to the success of any reliability analysis.

Example - Functional Analysis of a *Two-Speed Fan:*

SYSTEM FUNCTIONS: To provide a portable means of cooling electronic equipment when the normal cooling system is inoperable.

SYSTEM SUCCESS CRITERIA: Provide up to 200 CFM portable supply of air, able to be directed towards operating electronic equipment cabinets. Cooling to be established within 30 minutes of loss of normal supply and must operate for two hours at maximum air flow (Note: cabinet heat load calculation C4587-1, Rev. 2 is basis for this criteria. Duration time based on time to repair normal cooling study N3209-2, Rev. 0).

SYSTEM ENVIRONMENT: Bounding Conditions: 30° to 120° F, 100% Relative Humidity, Normal Atmospheric Pressure Range, Benign Ground Environment.

SUBSYSTEM FUNCTIONAL ANALYSIS: Although the actual cooling requirements may be less, the functional analysis is performed for the "bounding" conditions:

Subsystem	*Function(s)*	*Success Criteria*	*Function Time*	*Environment*
Fan	Provide directable cooling air to electronic equipment cabinets	No less than 200 cfm cooling air	No less than 2 hours - all functions	Bounding conditions: 30F 120F, 100% RH,
Motor	Transforms electrical energy into shaft torque to turn fan blade	"x" torque at "y" speed		Normal atmospheric pressure, benign
	Fan blade mounting system	Able to support fan weight, startup & operating thrusts		Ground environment
Battery	Provide electrical power to fan circuit	Provide "m" amps at "n" volts for a minimum of 2 hr.		
Switch	Provides electrical continuity (on demand)	Conduct "m" amps at "n" volts		
	Interrupts circuit (on demand)	Interrupt "m" amps at "n" volts		
	Allows selection of “high” or “low” speed or “off”	Switch circuit from “off” to “low” to “high” and back to “off”		
Wires	Provide electrical continuity, transmit electrical energy	Conduct "m" amps at "n" volts		

2. DETERMINE THE ANALYSIS LEVEL OF DETAIL

Systems are composed of equipment, which are made up of components that are built from parts. The person or group performing the FMEA must decide at what level of detail (indenture level) the analysis is to be performed. The *HIGHER* the indenture level, the *LOWER* will be the level of detail. The main purpose for performing a FMEA is to identify critical items; so, in general, the FMEA should be performed to the level necessary to support this objective. If a vendor provides the system, the level of indenture may need to be specified in the contract. The following guidelines can be used:

- Set the *highest* indenture level at the replaceable component level. For example, a hermetically sealed relay is replaced as a unit when it fails. Other types of relays may be repaired, and data may be available at lower levels of the component.

- Perform the FMEA to the level for which data exist.

- Perform the FMEA to the level for which information is available to establish definitions and descriptions of the item's functions.

- The level of detail may be influenced by the reliability history of the system. A lesser level of detail can be justified for systems having a good reliability record, whereas a greater level of detail may be required for systems having questionable reliability history or for unproved designs.

FMEA's are often performed in an iterative fashion. As the analysis is completed at the initial indenture level, a decision may be taken to proceed to lower levels of analysis only on those items that are determined to be critical. This occurs at each level of indenture to form a stratified FMEA.

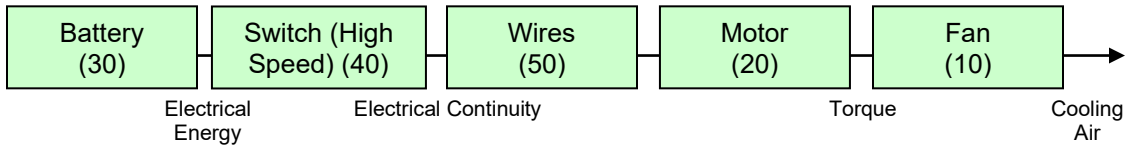
Example - *Two Speed Portable Fan* Indenture Level Breakdown

SYSTEM	EQUIPMENT	COMPONENTS	
Two Speed Portable Fan	10. Fan	11. Fan Blades 12. Fan Hub	13. Shaft Spline
	20. Motor	21. Case 22. End Bells 23. Shaft 24. Bearings 25. Stator	26. Rotor 27. Collector Rings 28. Brushes 29. Leads
	30. Battery	31. Case 32. Plates	33. Electrolyte 34. Terminals
	40. Switch*	41. Contacts - High 42. Switch Pole 43. Terminals	44. Case 45. Contacts - Low 46. Resistor
	50. Wires	51. Wire Segments	52. Wire Joints

* Note that both the "High" and "Low" speed elements of the switch are included here. Although the analysis is being directed at the bounding mission of 200 CFM cooling air for 2 hours, failures of the "Low" speed components may affect this function.

3. DEVELOP THE FUNCTIONAL BLOCK DIAGRAM

It may not be easy to determine the direct impact of a component failure on the sub-system or system function. For this reason, a functional block diagram is used to show the operation, interrelationships and functional flow between sub-systems. Generally, one block diagram is required for each phase of the mission. All inputs, outputs and functions of each element or block should be shown.

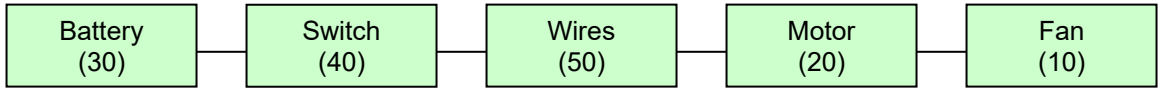


4. DEVELOP THE RELIABILITY BLOCK DIAGRAM

Drawing the reliability block diagram of the system is useful for two main reasons:

1) The series/parallel relationships are made clear, both to the group performing the FMEA and to anyone reviewing it. The effect of a parallel component's failure is documented on the FMEA as "Fails system (redundant)".
2) The reliability block diagram can be used to develop an estimate of the system's reliability through the combination of failure probabilities.

In the *Two-Speed Portable Fan* example, the reliability block diagram shows that the entire system is a *series system*; any subsystem failure will result in the system's failure:



NOTE: The following steps are repeated for each item analyzed in the FMEA.

5. LIST THE FUNCTION OF THE ITEM BEING ANALYZED.

This is a concise statement of the function performed by the item being analyzed. It should include the function of the item and its relationship to other items in the system, and the mission phase for which the analysis is being conducted. This information will be needed when determining the item's possible failure modes and their effects on the system.

6. LIST THE FAILURE MODE(S) OF THE ITEM BEING ANALYZED

How can the item fail to accomplish its function? List all the failure modes of the item. When entering a failure mode on the worksheet a serial number or some sort of coding system should be used for identification and trace-ability.

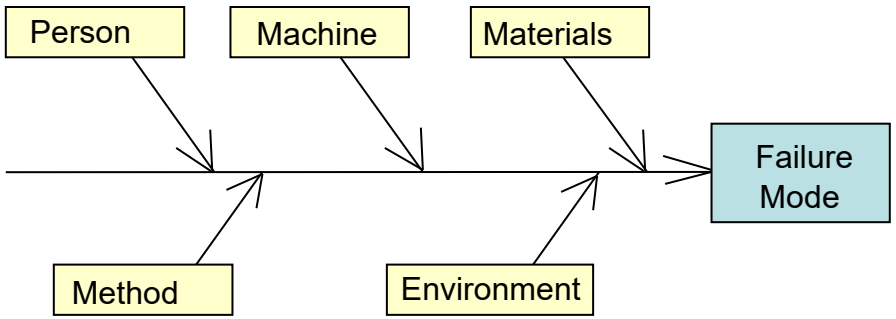
Here are some typical failure modes that should be explored for each item:

- Premature Operation
- Failure to Operate at a Prescribed Time/Condition
- Intermittent Operation
- Failure to Cease Operation at a Prescribed Time/Condition
- Loss of Output or Failure during Operation
- Degraded Output or Operational Capability

7. LIST THE POSSIBLE CAUSES OF THE FAILURE MODE

There may be more than one possible cause for the failure mode. Try to list all probable independent causes for each failure mode identified. The possible causes of each failure mode are identified to estimate the probability of occurrence and to formulate recommended corrective actions.

As you move to this step of the FMEA, remember the major categories of causes used on the Ishikawa diagram:



The following table lists just a few typical causes of equipment failure:

Defective Material	Defective Processing	Defective Assembly	Fatigue
Assembly Flaws	Welding Flaws	Construction Flaws	High Temperature
Low Temperature	Humidity	Vibration	External Force
Contamination	Chemical Spray	Salt Spray	Mechanical Stress
Temperature Change	Pressure Change	Radiation (UV, Gamma, etc.)	Other Environmental Factor

Sometimes, the most important factors that result in each failure mode have little to do with the actual operation of the equipment. The laptop PC used to write part of this manual is usually operated in an air-conditioned, low-humidity environment. Since it is portable, though, it is frequently transported from place to place, and sometimes experiences extremes of heat and cold, low and high humidity and shock. The FMEA should consider these types of causal factors as well.

The important point here is that we seek out the physical or process causes that can result in the failure mode. We generally can't act directly to prevent a given failure mode. We can act on the causes of the failure mode, through improvements in materials, training, processes, control of the component's environment, etc. If the FMEA is to be useful to us as a reliability management tool, we must identify and manage the causes of failure.

***FMEA Development Strategy**: **Listing** all possible causes of the many failure modes of a system can prove very tedious. One alternate approach is to first identify and prioritize (via the criticality analysis described below) all the failure modes. Then, return to the important failure modes and identify their causes. This may save you some analysis time and effort.*

8. DETERMINE THE EFFECT OF EACH FAILURE MODE ON THE SUB-SYSTEM AND SYSTEM

Each failure mode is analyzed to determine its effect on the item being analyzed, its effect on the next higher indenture level (i.e. the subsystem), and its end effect on the total system. Show how the failure affects the item's operations, functions, status, and outputs.

The functional and reliability block diagrams are very useful in tracking the failure's effect through the system. In some instances, the local effect may be no different from the failure mode.

9. DETERMINE HOW THE FAILURE MODE CAN BE DETECTED.

This has several meanings depending on the FMEA's purpose. If the system is in operation, it relates to how the failure is discovered by the operations or maintenance organizations. It can either be a symptom indicating that a malfunction is about to occur, in which case the failure may be preempted, or that a malfunction has occurred resulting in a need for repair. If the item is critical to system performance special design provisions should be made to detect the malfunction before it occurs.

In another sense, how a failure is detected can be related to construction of the system. Can the conditions or factors that cause the failure be detected during production of the system? You should define what you mean by “detect-ability” as part of the FMEA assumptions list.

10. PERFORM A CRITICALITY ASSESSMENT OF EACH FAILURE MODE

The criticality assessment ranks each failure mode and its associated causes according to a combination of influence. The influences are usually failure severity and failure frequency. Other parameters such as failure detect-ability and mean time to repair may be included as part of the criticality assessment. Generally, scales are developed to rank each of the influences being considered for the assessment. The product of each of these rankings, known as a ***Risk Priority Number (RPN)*** is used to develop an overall ranking of the failure modes. Page 18 depicts typical scales/rankings.

11. CORRECTIVE ACTION RECOMMENDATIONS

The relative importance of each failure mode has now been established. Root cause analysis needs to be performed on the possible causes identified in step 7 so that effective corrective action can be implemented.

Corrective actions should be formulated to mitigate the effects or reduce the probability of occurrence of the key failure modes. Design changes to eliminate the critical failure modes should be considered; however, if a design change is not possible, then preventive maintenance and condition monitoring should be considered. The disposition of "no action" represents management's acceptance of the failure mode into the system design. The important thing here is that the FMEA is ***used***. Simply going through the motions of the FMEA without acting is a waste. In the corrective action column of the FMEA, list which department is responsible along with when the action will be completed. In this way, the FMEA becomes a reliability "punch-list." Each time a corrective action is implemented, the FMEA is updated to reflect the new criticality ranking of the failure modes.

EXAMPLE: FAILURE MODE RANKING/CRITICALITY ASSESSMENT

FAILURE PROBABILITY	***SCORE***
Failure is very unlikely to occur	1 - 2
Failure probability is low based on data from similar components	3 - 4
Possibility exists for failure to occur	5 - 6
Failure probability is high based on data from similar components	7 - 8
Failure is very likely to occur	9 - 10

NOTE: If failure data is available, the scale should be based on actual failure frequencies or probabilities.

FAILURE SEVERITY	***SCORE***
Minor failure that may pass unnoticed by the user	1 - 2
Minor failure that deteriorates item's appearance	3 - 4
Medium failure that causes functional deterioration during operation/use	5 - 6
Major failure that disables operation/use	7 - 8
Critical safety failure resulting in injury or death, or damage to property	9 - 10

PROBABILITY OF FAILURE DETECTION	***SCORE***
Possible to detect almost every failure in-house prior to construction/assembly (i.e. at receipt inspection)	1 - 2
Mostly detectable prior to completion of construction/ maintenance. Usually detected during assembly/fit-up	3 - 4
Undetectable until turnover to operations. Most failures detected during startup or post-maintenance testing	5 - 6
Undetectable until use by operations/end-user. Most failures detected during operations/periodic testing	7 - 8
Undetectable until demanded by abnormal condition/ accident, not detectable during operations/periodic testing.	9 - 10

Example FMEA – Two Speed Portable Fan

ITEM	FUNCTION	FAILURE MODE	FAILURE CAUSE	FAILURE EFFECT	CRITICALITY			RECOMMENDED ACTION
					FREQ.	SEVER.	RPN	
11. Fan Blades	Increase air momentum	Separates from fan hub	Cracks, fatigue Bolt failure Bolts loosen (not torqued properly	Local: decreased air flow System: decreased flow to cabinets	2	6	12	
12. Fan Hub	Transmit shaft torque to blades	Slipping Separates from shaft	Shaft spline works loose Cracks, fatigue Shaft spline works loose	Local: decreased or loss of air flow System: insufficient flow to cabinets	3 2	6 8	18 16	Investigate design change – secure shaft spline with spot weld or glue
13. Shaft Spline	Secures hub & shaft	Works loose Cracks	Improper installation Fatigue, improper installation	Local: hub slips or separates from shaft	4 2	6 6	24 12	Check/ revise installation procedure
21. Motor Case	Structural support for motor components (end bells, stator leads)	Structural failure	Cracks, fatigue Bolts over-torqued Screws loosen due to vibration Bolts loosen	Local: motor failure System: insufficient flow to cabinets	1	9	9	
22. End Bells	Stator positioning	Separation from case	Cracks, fatigue Bolts degraded Bolts loosen	Local: motor failure System: Insufficient flow to cabinets	1	9	9	

Example FMEA – Two Speed Portable Fan

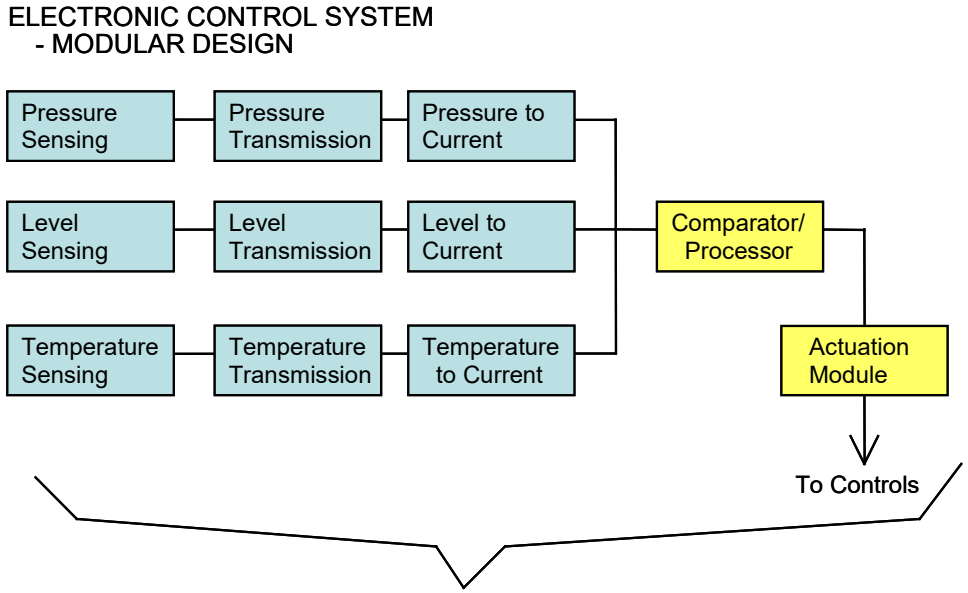
ITEM	FUNCTION	FAILURE MODE	FAILURE CAUSE	FAILURE EFFECT	CRITICALITY			REC'D ACTION
					FREQ.	SEVER.	CRIT.	
23. Shaft	Rotor Support & Positioning	Shears	Cracks, fatigue Bearing failure	Local: motor failure System: Insufficient flow to cabinets	1 6	9 9	9 54	 (see bearings)
	Fan hub torque transmission	Shears Slipping	(See Above) (See #12, #13, fan hub, spline)	Local: decreased or loss of air flow System: insufficient flow to cabinets	1 3	9 6	9 18	
24. Bearings	Axial Support for shaft (thrust)	Axial wobble	Bearing Wear	Local: Stator contact with motor components	3	5	15	
	Radial Support for shaft	Radial wobble	Bearing Wear	System: Motor failure, loss of flow to cabinets	3	5	15	
	Low friction surface for shaft rotation	Seize	Lack of lubrication	Local: shaft rotation stopped or slowed System: short circuits, loss of flow to cabinets	6	9	54	Develop periodic greasing PM

Types of FMEA

The procedure described above focused on the "typical" FMEA; one performed on hardware and focused on the parts of the hardware. The *Reliability* of the system was assumed to be the subject of the FMEA. There are other types and purposes of FMEA. A few will be mentioned here.

Functional FMEA

If a FMEA is being performed when the design is at a *conceptual* stage, there may not be any "hardware" or software to consider. As the mission is defined, the functions that will be required to perform the mission will begin to take shape. In this case, a functional FMEA can be performed. For example, an electronic control system may be designed with a modular philosophy, to facilitate repair and upgrading of the system. At an early phase of the system's design, the modules and their functional relationships (inputs, outputs and module functions) may be identified. The actual circuitry associated with the modules will not be available for analysis. The FMEA can be used here to identify weaknesses in the functional design.



FUNCTIONAL FMECA PERFORMED ON MODULES

Safety Hazard Analysis

FMEA can be used to determine how system functions relate to personnel (employee, contractor and public) safety. Here, a modified FMEA process known as Hazards Analysis is employed. The FMEA addresses the system's hazardous elements, their failure modes that could cause a safety hazard, potential causes, effects, hazard classification and preventive measures. The same general process for performing FMEA is then followed. Instead of a severity scale that reflects the impact of the failure on the system's function(s), the hazard analysis makes use of a scale such as the following:

Description	*Category*	*Mishap Definition*
Catastrophic	I	Death or System Loss
Critical	II	Severe Injury or Major System Damage
Marginal	III	Minor Injury or Minor System Damage
Negligible	IV	Less than Minor Injury or System Damage

Managing Reliability Through FMEA

Now that we've seen how to *do* a FMEA, let's talk about *when* to do a FMEA and *how* to use the FMEA. We mentioned earlier that FMEA has been traditionally used as a design tool and that its use has spread from this function to other areas such as maintenance, spare parts analysis, safety analysis, etc. We'll go through these "traditional" uses of FMEA first.

Many companies, though, have a great deal of installed equipment. A major concern is how to maximize the reliability, availability and safety of this existing inventory. We'll explore how FMEA has been helpful in these areas, too.

Traditional FMEA Process/Use

Let's say we are planning how to ensure the quality of a new refrigeration system. The system's reliability is a quality characteristic that is of interest to management. The refrigeration system will be constructed from equipment that has a previous history of operation and maintenance. Some new equipment will be installed, though, and the instrumentation and controls are now electronic.

FMEA IN THE DESIGN PROCESS - The most important decision to make is to include the FMEA as *an integral part* of the design process. For instance, if there are reviews of the design scheduled, the FMEA should be specified as a *required* input to these reviews. The design should not be allowed to proceed past certain checkpoints unless the FMEA has been completed, reviewed and comments resolved.

Two electric utility experiences emphasize this:

- During the procurement process for a large piece of rotating machinery, utility personnel were reviewing the quality and reliability assurance activities of the vendor. The vendor stated that they performed FMEA on the machine's design. When utility staff visited the vendor's engineering offices, they found that 1) there was no formal FMEA done, 2) a simple failure analysis of the machine had identified failure modes that had been experienced at other utilities' plants, but no action had been taken by the vendor to reduce the probability of occurrence of these failure modes.

- In the process of managing a large modification to an emergency power system, the project manager decided that it would be a good idea to perform a reliability analysis of the modification. Unfortunately, the timing was poor. The FMEA was requested as the equipment was being fabricated (much too late in the design cycle). A FMEA was performed, several important failure modes were identified, but by the time the FMEA was complete, the equipment

was complete, schedules and budget were tight and the recommendations from the FMEA were not incorporated into the design (They were put into a list of "post-installation" modifications to be installed "sometime.").

PRIORITIZING THE FMEA - The second decision in planning the FMEA would be how to concentrate the FMEA efforts. Refrigeration systems are built from similar components, replicated in many different locations. Valves, pipe, motors, cable, supports, circuit breakers, etc. are applied all throughout the design.

We could make the FMEA process more efficient by developing "generic" FMEA for these components and then checking to see if there was something unique about specific components based on their function or environmental conditions. For major equipment with a history of experience, we should seek out data on the failure modes that have occurred as well as their causes. In the FMEA process, these would be the "high criticality" failure modes. Depending on the causes of failure, the review of the FMEA would result in actions to either improve the design, operation or maintenance of the equipment.

We could then concentrate our detailed FMEA efforts on the new equipment and the new instrumentation and control system. This FMEA will point out areas to improve in the design as well as areas that must be thoroughly tested before installation in the field.

The FMEA efforts will also identify reliability-critical items whose quality must be controlled during fabrication, assembly, construction and start-up testing.

FMEA'S USE AFTER DESIGN - The refrigeration system has been designed and testing is underway. Is this the end of our need for the FMEA? No, the next stage of planning the maintenance program is now underway and the FMEA provides valuable input in two ways.

First, the maintenance performed on the plant will be based on the principles of Reliability Centered Maintenance, or RCM. The FMEA is the basic input to this analysis (this is why the FMEA must cover all critical systems in the system). By understanding the effects of failures at the component or part level, we can decide what the appropriate maintenance activity is for that component.

- Should we let the device fail and then replace/repair it?
- Should we institute some usage or schedule-based preventive maintenance activity?

- Should we develop some condition monitoring activity to check the degradation of the component?

The FMEA is a basic decision-making input to help us answer these questions.

Second, the customer is looking to us to decide what spare parts are needed. Which ones should they keep on-site, how many do they need of each? Budgets are tight, they must keep their costs down, but reliability is important, too. They can't afford to have the refrigeration system sit idle, waiting for a part to be flown in from the Europe or the West Coast.

Here the FMEA is a prime input to the spare parts analysis, sometimes called a Logistics Analysis. Again, by understanding the effects of component failure and their frequencies, we can systematically plan the necessary spare parts.

Finally, as the system operates, we will begin to experience failures. Some will be expected, some the FMEA will not have predicted. We will rotate the Plan-Do-Check-Act cycle and develop design changes for the system. Here the FMEA will help us ensure that the design change does not introduce any new failures in our attempt to eliminate an existing failure.

Does this sound like a fairy-tale to you? Could you ever see this happening at your company? These uses of FMEA have been in place in several industries for years. This is the way some companies manage reliability. In many cases, these companies are moving ahead to automate the FMEA, relationally linking this information to the design, and bringing operational and failure data into the picture as well.

Unique FMEA Applications

Even though a FMEA was not prepared as part of the system's design, this basic reliability tool can be applied wherever it is needed. Some examples follow:

- As part of determining those *Reliability-Centered* maintenance activities appropriate for a system, FMEA has been performed to understand the component failure modes and their impact on the system. This information is then analyzed to develop applicable and effective maintenance tasks for a system.

- FMEA has been applied to *Reliability improvement* efforts. Radio failures in a utility's distribution maintenance department were analyzed by performing a FMEA of the radio system, collecting data to prioritize the critical failures and then develop root causes and corrective actions.

- FMEA has been performed in support of *Design change* efforts. In redesigning a nuclear plant's 4160V power supply, FMEA was performed on the major components of the system. FMEA was also performed on the major redesign of the plant's emergency power system. In these cases, compliance with regulatory requirements was a major objective of the FMEA.

FMEA Summary

The Failure Modes & Effects Analysis is a simple tool for analyzing the effects of individual failure modes on the system's function. The knowledge gained through performing FMEA is very useful when making decisions pertaining to system redesign, critical operation control, maintenance planning, etc. The following lists some of the many uses of FMEA:

- As an “early warning system” for design problems.
- An input to the design review.
- Provides information for selecting critical items for reliability tests.
- Determines critical control points for construction, operation and maintenance.
- Provides input to maintenance planning and the Reliability Centered Maintenance process.

The concept of FMEA is not limited to equipment failures. It may include software and firmware failures. Next, we’ll apply this same concept to process failures and errors made by humans involved in business processes.

15.2.2 Process Failure Modes & Effects Analysis

In 15.2.1, we learned how to use the FMEA to analyze the effects of equipment failures on the system. But what if the "system" is a *production*[2] process that involves a mixture of humans and machines, or primarily involves humans performing some activities? Can we use this same thought process to understand the types and effects of failures and human errors on the process?

It turns out that the technique of Process Failure Modes & Effects Analysis (PFMEA) can help you focus on problems in a process. A subset of PFMEA, known as Error Mode and Effects Analysis (EMEA), is used to analyze processes where human errors are the main concern. We'll combine both techniques into the PF/EMEA.

PF/EMEA can be used to analyze the process "up-front" when you are designing the process or as a method to reduce frequent errors that are the fault of a process in need of improvement.

A Process Failure/Error Modes & Effects Analysis may be applied as part of "building" a process management system to help assure the quality, timeliness and cost of your products and services. At a more basic level, PF/EMEA may be used as part of the preparation or revision of procedures that govern day-to-day activities or "special" circumstances (abnormal conditions, accidents).

[2] Here, a *production* process is defined as any set of activities used to transform some inputs into outputs. Production processes include manufacturing and service processes.

PF/EMEA Logic and Philosophy

PF/EMEA is an inductive logic tool just like FMEA (bottom-up approach). The process is examined step by step. At each step, potential errors are identified and assessed according to their impact on the function of the process. A criticality assessment like that of the FMEA is performed; however, for human errors, this is not as easy as the same assessment done on equipment failures. If a series piece of equipment fails, the system does not generally respond to recover from the failure. Human errors can be detected, corrected and/or compensated for by the person who committed the error, by someone downstream in the process or by automated checking functions. This aspect of error detection and correction should be considered when performing the criticality assessment.

Once the process failures and errors have been ranked by importance, priorities can be placed on where to improve the process. Changes made in response to errors often focus on adding layers of inspections and checks. This approach adds cost and dilutes responsibility, though. A better approach is to develop means of reducing the probability of the error occurring or mitigating its effects. The concepts of ERROR PROOFING should be employed when considering potential process improvements.

This last paragraph points out a major difference in approach between "traditional" quality assurance and more modern thinking. How many times, when a problem occurs, are we quick to blame the *person* involved in the problem? It's true; ultimately some department should accept responsibility for analyzing the problem and developing a solution.

But do you know the "default" values for the major bones of an Ishikawa diagram? Aren't there three or four other categories besides "People?" Things like the *Method*, the *Machine*, the *Material* and the *Environment* can play an important role in determining the successful outcome of a process. In fact, experts in quality management, such as the late Dr. W. Edwards Deming, believe that about 90 to 95 percent of problems are due to these factors, with only 5 to 10 percent due to the "People."

In the late 1980's, an important valve was left closed by a technician at a southeastern US nuclear plant. The technician was given three days off without pay for "his" error. An investigation of the causal factors behind this error revealed that the procedure he was following did not include a step telling him to reopen the critical valve. Plant management had been emphasizing "Verbatim Compliance" with procedures at the time - i.e. no deviation from written procedures. How would you criticize the management of this event?

PF/EMEA Definitions and Building Blocks

Definitions

The following new terms will be used during the description of the PF/EMEA tool:

TERM	DEFINITION
Process	Generally, any set of actions that transform *inputs* into *outputs*. Elements of a process include methods, materials, equipment, staff, information, and environment.
Process Boundary	That portion of the process for which the analysis is being performed.
Error Mode	The ways in which the step or task under analysis can be performed incorrectly. These are typically placed into two categories: errors of omission (process step not performed) and errors of commission (process step performed incorrectly or performed out of sequence). *Note: a third category - the conscious error - is somewhat difficult to predict using PF/EMEA.*
Error Causes	The possible causes associated with each error mode.
Error Effect	The consequence the error mode has on the function of the process.
Error Prevention	The principles of error proofing applied to decrease the possibility of the error occurring. Error prevention can be broken up into three categories: a) Error Elimination, b) Decision Delegation and c) Task Facilitation.
Error Mitigation	The principles of error proofing applied to minimize the effect once an error has occurred. Error mitigation can be broken down into two steps: a) How to detect and b) How to mitigate.

Process Flowchart

The process for which the PF/EMEA is being performed must first be understood. Processes start somewhere, end somewhere, and follow some sort of sequence. One activity comes first, then a second and so on. In some cases the step may be performed concurrent with another step. Understanding just how the process works can be a challenge.

Flowcharting the process helps us to understand how the process works. In fact, the flowchart will take the place of the drawings, functional and reliability block diagrams we constructed for the FMEA. See Unit 5.2 for flowchart development instructions.

As you flowchart the process, make sure that it is the process that is happening (not necessarily the one that is described in the procedures manual!). People who work in the process are the best ones to consult to gain this knowledge.

For an existing process, while you are flowcharting the process, you can also collect data and information on failures and errors that have occurred and their causes. This can give you a real “leg up” on the PF/EMEA.

General PF/EMEA Procedure

The procedure for performing a PF/EMEA is very similar to that of an FMEA. Instead of analyzing equipment failure modes to determine the effect on the system, PF/EMEA analyzes process step failure and error modes to determine the effect on the process. The failure and error modes are ranked in order of criticality and principles of reliability improvement and error proofing are used to improve the process. The process steps for completing a PF/EMEA are as follows:

1. UNDERSTAND THE PROCESS FUNCTIONS

In this step you determine what the process' functions are. The inputs and outputs are defined, along with success and failure criteria.

2. IDENTIFY THE LEVEL AT WHICH THE PROCESS IS TO BE ANALYZED

The process may be very large and "process boundaries" may need to be used to focus the analysis on a particular portion of the process. In addition, the "indenture level" should be defined for the analysis. We broke a system down into subsystems, equipment, components and parts. Equivalently, we may consider a system of processes, processes, activities and tasks/steps.

Note: Level 1 processes are the highest in your organization (e.g. Marketing, Sales, Manufacturing, Shipping, Operations, Maintenance, Finance). Level 2 processes are the next layer down – Maintenance' Level 2 processes include Plan, Schedule, Conduct, Evaluate. For some businesses, you will have to go to Level 4 or 5 to get to the task level.

3. DEVELOP THE PROCESS FLOWCHART

The process flowchart is like the functional block diagram in that it shows you the functional flow sequence and the inputs and outputs of each step of the process. The flowchart will be useful when determining how an error in one step of the process affects the process function. The flowchart can be verified by walking through the process with those responsible for it, or with the operator in a man-machine interface system.

Verification may not be possible for a new process being developed; however, useful information can be gained by walking through a similar, established process.

NOTE: The following steps are performed for each step of the process being analyzed.

4. LIST THE PROCESS STEP/TASK BEING ANALYZED

This is a statement describing the function of the step/task being analyzed. Start points, stop points, inputs, outputs and success criteria are included. If the step/task is a man/machine interface, a description of the machine, instrument or control panel should be included.

5. IDENTIFY THE FAILURE/ERROR MODES OF EACH PROCESS STEP/ TASK

List all the possible failure and error modes for each step/task of the process. For “hardware” failures, the “usual” failure modes are identified. For human errors, the PF/EMEA should consider both errors of omission and errors of commission. Here’s the distinction:

Errors of Omission - Look for ways that the process step or task could be missed. The technician who left the valve in the wrong position committed an error of *omission*. When you forget your wallet, as you are getting ready for work, you most likely commit an error of omission.

Errors of Commission - These are deliberate actions taken in response to a given set of inputs that produce an incorrect output. During a filling procedure, an operator looks at a tank gauge, incorrectly judges that the tank is not full and continues the fill until the tank overflows. This is an error of *commission*. During a switching operation, an operator opens the wrong circuit breaker, cutting power to a feeder. Here, again, we have an error of *commission*.

Errors of *omission* are generally easier to identify than those of *commission*. Discussions with personnel who have performed the process in the past (or similar processes) may reveal potentials for both types of errors. Ask about the "near misses," those instances where an error almost occurred, but was prevented by the operator recognizing the error or by detection through another person's alertness.

6. DETERMINE THE EFFECT OF THE FAILURE/ERROR ON THE PROCESS

What are the consequences given that the failure or error mode has occurred? The process flowchart can be a very useful tool to aid in analyzing the effect of each failure/error mode. Failure/Error modes can be analyzed for their

"immediate effect" and their "ultimate effect". In the case of a meter being read incorrectly the immediate effect may be an incorrect bill being delivered to a customer. The ultimate effect may be a dissatisfied customer.

7. LIST THE POSSIBLE CAUSES FOR EACH FAILURE/ERROR MODE

In this step the potential root causes for each failure/error mode are identified. Whenever possible, the identified root causes should be verified using historical data to determine which root causes are most likely to result in the failure/error mode.

For errors, the physical layout of the process may give some clues to the possible causes. Will the worker have to transport a sensitive piece of equipment through narrow walkways? Is the floor smooth, rough or are there any steps? Is the assembly area a small, cramped, hot space?

Along with the physical layout, keep the environment in mind. Environmental factors such temperature, humidity, light, noise, etc., may influence human performance.

The training and the procedures given to the person performing the step/task should be examined for their adequacy.

If the process has been in existence for some time, there may be historical data on the causes of errors. Use this data when analyzing the process or a new process that may be similar.

Be careful when attempting to analyze the *causes* of human errors. Often, this will take you down the path of psychology and human behavior. Sometimes, it's better to jump to error proofing methods rather than try to understand the "true" root cause of a human error. See later in this unit for a set of error proofing principles and examples.

8. DETERMINE HOW THE FAILURE/ERROR CAN BE DETECTED

Given that the failure or error has occurred, how can it be detected? For hardware failures, the "usual" detection criteria apply. For existing processes, there may be some means of identifying errors when they occur, whether it be a QC hold point, a sign off, a design review, etc. For a new process, the PF/EMEA may be used to identify necessary detection methods or steps.

For processes that combine the person and machines, there may be some alarm or instrumentation that will clue the operator (or other personnel) to the error. For these situations, some estimate of the *recovery* potential associated with the error should be documented.

A ranking of how easy and how far through the process the error is detected can be established and used as part of the criticality assessment.

9. DETERMINE THE CRITICALITY OF EACH ERROR MODE

Numerical scales are set up to rank each failure/error mode and its associated causes with respect to severity, likelihood of occurrence, and ease of detection. A Risk Priority Number (RPN) is developed for the failure/error modes by taking the product of these rankings. The RPN can be used to prioritize steps/tasks for process improvement. Make sure you tailor the criticality assessment to your project. For example, one team assessed the potential for a manufacturing process to produce defects. The frequency scale they used was like that presented earlier in this unit. However, their severity scale was set up based on how long it took to rework the defect (minutes and hours).

10. PROCESS IMPROVEMENT RECOMMENDATIONS

Once the critical failure/error modes of the process are identified, improvement recommendations can be made. For human errors, the principles of *mistake proofing* can be used to help develop these recommendations. Methods of error prevention and error mitigation should be considered.

When a process improvement is recommended the PF/EMEA can be updated to show the overall effect of the change. If several different changes are being considered the PF/EMEA can be used to determine which change or combination of changes will have the greatest impact.

Principles of Error proofing

We have observed numerous occasions where the response to an error has been the addition of some *inspection step.* While inspection is necessary for certain processes, it adds cost and time to the process and tends to dilute responsibility for quality. Inspection is also a “downstream” countermeasure; the inspection “accepts” the occurrence of error and does nothing to prevent the error.

An alternative to inspection is the idea of *error proofing* a process – identifying ways of preventing/mitigating errors without inspection. Dr. Hitoshi Kume developed the following *error proofing* principles by examining many hundreds of human errors and the actions taken by organizations to prevent their reoccurrence. Whenever possible, adopting one of the "error prevention" strategies is preferred, since the error is prevented from happening in the first place. The "error mitigation" strategies are more "downstream" in nature, and they try to detect the error before it has a chance to result in an undesirable effect.

I. Error Prevention

Error Elimination - Example: Instead of a "Do Not Touch" sign, to prevent burns, insulate pipe carrying high temperature fluids.

Decision Delegation - Example: Instead of requiring the operator to read a Torque Value, provide a torque wrench with an audible "click" when desired torque is reached. For cashiers or salesclerks, program the function of making change into the cash register.

Task Facilitation -

1. Match the task to the operator's abilities - Example: Attach carrying handles to bulky objects; develop easy-to-read instructions (note the use of cartoons/diagrams in many of today’s setup instructions for VCRs, PCs, etc.).

2. Stratify and specialize tasks – Example: Separate similar tasks spatially (performed in different locations) or timewise (Task "A" this week, Task "B" next week); place similar parts in different bins.

3. Distinguish tasks - Example: Color-coding of similar parts (e.g. wiring), different color identification tags for equipment on adjacent power plants.

II. Error Mitigation (To employ this strategy, first the error must be detected, then some means of mitigation employed)

How to detect the Error

1. Based on Physical Movement - Example: An assembly process required 10 drillings. If the proper number of holes is not drilled, an automatic alarm will sound.

2. Based on Pre-Set Criteria – Example: Use of spell checkers on word processors, use of validation techniques in data entry

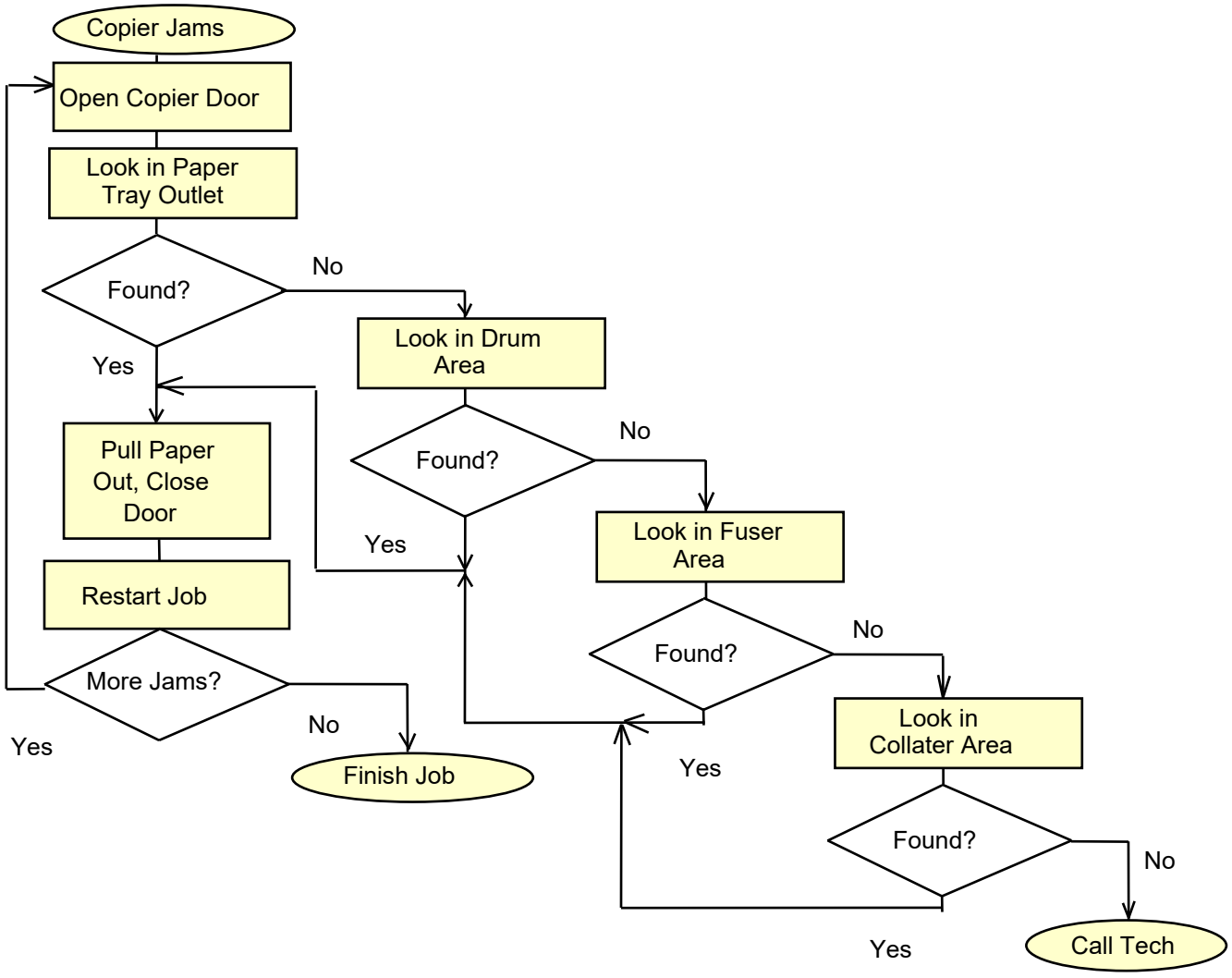
3. Post-Action Detection - Example: Use of different bolt patterns (e.g. exhaust manifold to engine block).

Mitigation - Example: If driver error causes the automobile to crash, safety belts or air bags mitigate the effects of the error. If a spelling error is made, provide an auto-correct feature.

Example PF/EMEA

Here is a simple example of a PF/EMEA performed on the process of finding and clearing a paper jam in a copier.

PAPER JAM PROCESS FLOWCHART



EXAMPLE PF/EMEA FOR COPIER JAM CLEARING PROCESS

TASK	ERROR MODE	EFFECT	ERROR CAUSE	DETECTION METHOD	CRITICALITY				REC'D ACTIONS
					FREQ.	SEVER.	DETECT.	RPN	
Open copier door	Can't find door latch	Can't remove paper jam	Labeling	By User/ Technician	3	7	3	63	Label door latch
	Excess force opening door	Breaks copier door, copier disabled	Sticking latch	By User	1	8	2	16	-
Look in paper tray outlet	Can't find paper tray outlet	Can't remove paper jam	Labeling	By User/ Technician	4	7	3	84	Label tray outlet
Look in drum area	Can't find drum area	Can't remove paper jam	Labeling	By User/ Technician	5	7	3	105	Label drum area
	Scratch drum during search	Poor copies in future	Unaware of drum sensitivity	By next user	3	5	7	105	Caution label
	Damage corotron during search	Poor copies in future	Unaware of corotron sensitivity	By next user	2	5	7	70	Caution label, protective shield
	Disturb toner dispenser	Dirty hands, poor copies	Unaware of dispenser location	By user, or next user	6	4	6	144	Add retractor to toner dispenser
Look in fuser area	Can't find fuser area	Can't remove paper jam	Labeling	By user/ Technician	3	7	3	63	Label fuser area
	Burn fingers on hot fuser	Injury to user	Unaware of high temp., labeling	By user	3	9	9	243	Caution label, interlock
	Electrical shock	Injury to user	Exposed wires, energized	By user	1	10	10	100	Door interlock shuts down fuser

PF/EMEA Summary

Process Failure/Error Mode and Effects Analysis is a method of identifying the important events that can result in the failure of a process to perform its functions. PF/EMEA is identical in concept and very similar in application to FMEA. PF/EMEA may be used in a task team project to analyze the causes of process failure from a "bottom-up" approach. Some of the characteristics and uses of PF/EMEA are as follows:

1. There is a natural link between PF/EMEA and process control systems:
   - PF/EMEA can assist in diagnosing present systems.
   - PF/EMEA can assist in designing control systems for new processes.
2. PF/EMEA only describes the effects of one type of failure/error at a time. Some failures are the result of two or more actions working in concert.
3. Remember that in a process or interface system, environmental factors can play a significant role in error occurrence.

15.3 Fault Tree Analysis

Learning Objectives

- Be able to create a Fault Tree for some “undesired event”
- Be able to determine minimal cutsets of a Fault Tree
- Be able to quantify the top event probability/frequency of a Fault Tree

Unit Contents

- Fault Tree Analysis Logic
- Fault Tree Definitions and Building Blocks
- Fault Tree Analysis Procedure
- Fault Trees and Reliability Block Diagrams
- Fault Tree Computer Codes

15.3.1 Introduction to Fault Tree Analysis

In Unit 15.2, we discussed two tools that help examine the effects of component failures or human errors on the system/process. The FMEA and PF/EMEA use a "bottom-up" approach to failure identification and management. What if we took the opposite approach to this problem? Let's start at the "top" (i.e. system level) and work our way "down" to the components whose failure will cause the system to fail. This approach can be generalized. In this analysis any "event" can be included that results in the system's failure, regardless of whether it was due to a component failure *or* a human error. In the 1960's, this approach was taken by reliability engineers and developed into the Fault Tree. Fault Tree Analysis (FTA) focuses on determining the causes of an identified undesired event. It falls into the same category of management tools as the Cause and Effect (Ishikawa) Diagram. A few of the Fault Tree's many features and uses include:

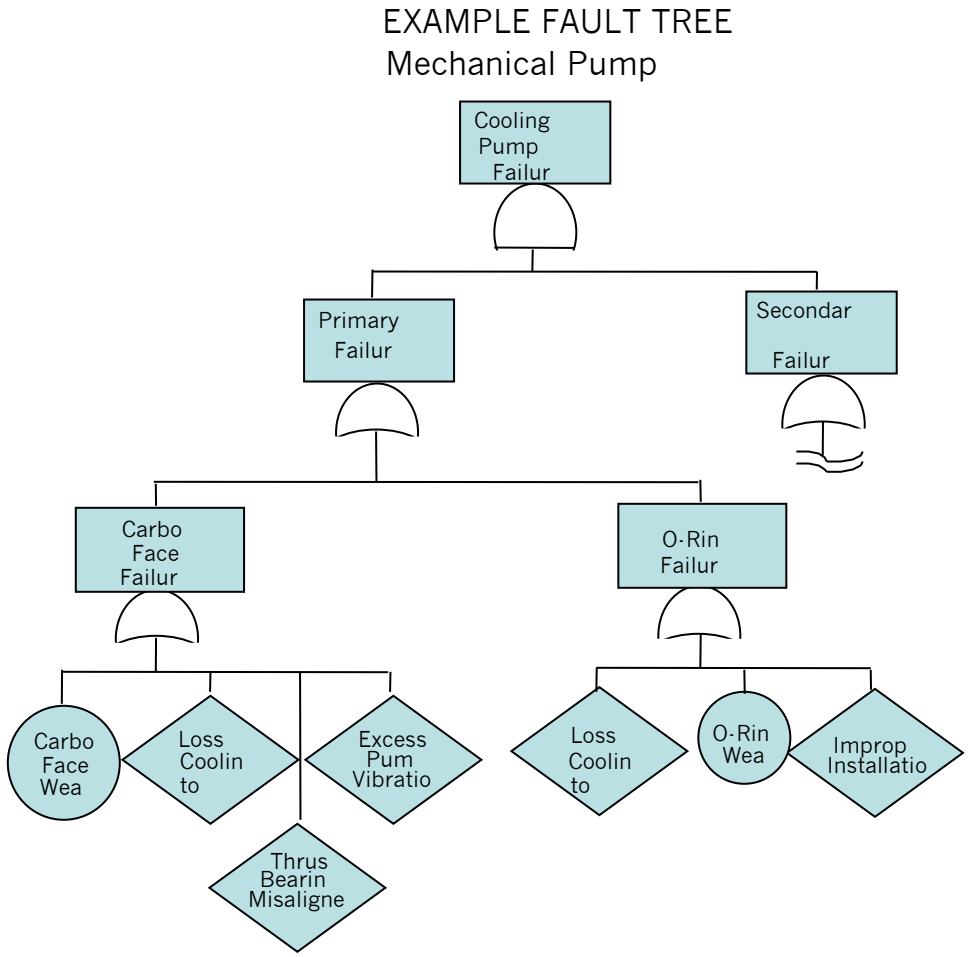
- Fault Tree analysis (FTA) can yield both *qualitative* and *quantitative* information about the system.
- FTA provides a better understanding of potential design problem areas and gives the designer a chance to improve the design at a less costly stage of the product's lifecycle.
- Fault Tree Analysis has been used to aid in uncovering the *root cause(s)* of important failures.
- Fault Trees prepared ahead of time have been used as diagnostic tools, enabling the operations and maintenance forces to quickly zero-in on the cause of a problem.
- Fault Trees have also been used to evaluate existing systems (typically safety systems) to determine if these systems are vulnerable to "hidden" failures not found through other design checks. This versatile tool has also been used to examine the impact of proposed modifications to a system.

There are two other features that separate FTA from FMECA and EMEA. First, the latter two tools examine failures one at a time. If our system is primarily a "series" system, then this is OK. If our system incorporates any degree of redundancy, or if there are human backups to automatic actions, the FMECA-PF/EMEA approach becomes somewhat clumsy. The Fault Tree identifies all possible combinations of component failures and operational errors result in the undesired event. Second, the Fault Tree can be **quantified**. The logic expressed by the Fault Tree can be transformed into a probability expression. If the probability of occurrence is known for each failure/error, then the probability of the undesired event occurring can be calculated. This latter feature can be useful when attempting to predict the reliability or unavailability of a system in either its design stage, or when the system incorporates redundancy and has never been "observed" to be in a failed state.

15.3.2 Fault Tree Analysis Logic

Fault Tree Analysis is a *deductive* logic tool. The basic question here is "Why or how could (or did) this happen?" The Fault Tree is a logic diagram, developed through deductive analysis. Fault Trees graphically portray the sequence of events required to cause the undesired event to occur.

The analysis starts at an undesired event (called the top event) and works downward, examining the system in increasing detail to determine all possible causes of the top event. It does this by asking "why did this happen?" The "this" refers to the failure of a system or process. The "why" goes back to the root cause of the effect (or the "this"). This question is asked for the top event and for each event under the top event until the root causes are reached. By doing this, the top event branches out to many lower events (sub events); thus, the name fault "tree".



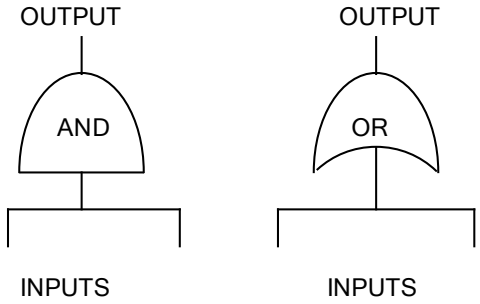
15.3.3 Fault Tree Definitions and Building Blocks

Fault Tree Definitions and Symbols

Fault Trees are a graphical representation of the combination of events required for the top event to occur. The graphical symbols used in Fault Trees fall into two categories, *Logic* and *Event*.

Logic Symbols:

The logic symbols, or logic gates are used to show the interrelationship of events that lead to a "higher" event. Logic gates serve to permit or inhibit the passage of fault logic up the tree. The "higher" event is the output of the gate, and the "lower" events are the input to the gate. The logic gates that are most frequently used to develop a Fault Tree are the AND and OR gates.



The **AND** gate provides an output if and only if all input events occur concurrently. This equates to the intersection of the events. A Truth Table is shown below which displays the logic of the AND gate.

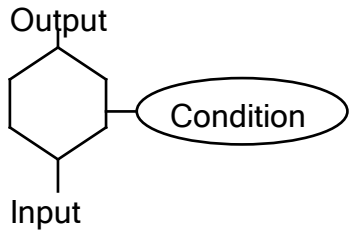
INPUTS		*OUTPUT*
A	**B**	**C**
Occurs	Occurs	Occurs
Occurs	Does Not Occur	Does Not Occur
Does Not Occur	Occurs	Does Not Occur
Does Not Occur	Does Not Occur	Does Not Occur

The **OR** gate provides an output if one or more of the input events occur. This equates to the union of the events. A Truth Table for the OR gate appears below.

INPUTS		***OUTPUT***
A	**B**	**C**
Occurs	Occurs	Occurs
Occurs	Does Not Occur	Occurs
Does Not Occur	Occurs	Occurs
Does Not Occur	Does Not Occur	Does Not Occur

Another gate that is useful is the **INHIBIT** gate. Here, the output occurs if *both* the input(s) occurs and a certain *condition* is met. For instance, a *fire* (output) occurs if a gas leak (input) and a spark (another input) occur. The fire, though, will not occur without the presence of *oxygen*. The *oxygen* (i.e. air) is the *condition* that must be met for the fire event. The symbol for the *inhibit gate* is:

INHIBIT GATE



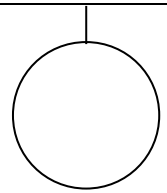
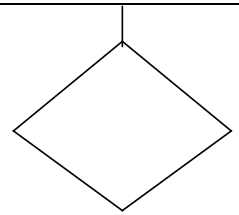
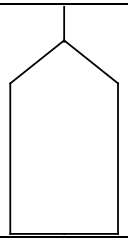
Event Symbols

Events appear as inputs and outputs to logic gates. The symbols for events can be classified into two categories: *primary* and *intermediate*.

Primary Events

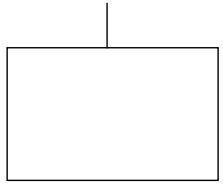
The *primary events* of a Fault Tree are those events that have simply not been further developed. It could be that the analysis has reached its intended level of detail or there is insufficient data available to continue a further breakdown. A Fault Tree Analysis of a system may be conducted to the equipment level, the component level or the part level or even a mixture of these levels. If one of the analysis' purposes is to quantify the top or undesired event, these primary events will

also need to be quantified. The **Basic Event**, the **Undeveloped Event** and the **External** or **House Events** are most frequently used.

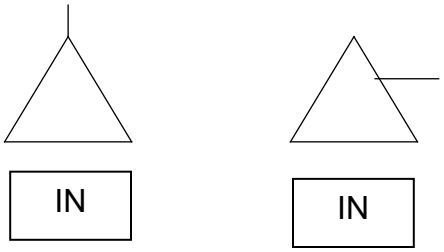
The **Basic Event**, represented by a circle, describes a basic initiating fault event that cannot be developed further or a basic event in the lowest level whose probability of occurrence can be obtained independently.	
The **Undeveloped Event**, represented by a diamond, describes fault events that are not further developed, either because the event is of insufficient consequence or because of the lack of necessary information. To obtain a solution for the Fault Tree, both circles and diamonds must be represented by probabilities of failure.	
The **External or House Event** is used to describe events that are normally expected to occur and are generally not “faults.” For example, the presence of oxygen in the analysis of a fire could be a House Event.	

Intermediate Events

An **Intermediate Event**, represented by a rectangle, is a fault event that results from the combination of two or more "lower-level" faults acting through logic gates. The Top Event and all Sub-Events down to (but not including) the Primary Events fit into this category.



Transfer symbols, represented by triangles, are used to indicate the transfer from one part of the Fault Tree to another. These symbols are used for "bookkeeping" purposes. Often, a Fault Tree will extend over more than one page. These symbols are used to connect the logic from page to page. Sometimes, the same "chunk" of logic will appear in multiple places in the Fault Tree. These symbols allow the analyst to reference this repeated logic.



The Fault Tree and Boolean Algebra

Here, we'll provide some of the background behind the Fault Tree. If you are going to be using the Fault Tree as a system analysis tool or to estimate the probability of the top event's occurrence (i.e. system unreliability or unavailability), this is important. On the other hand, if you are using the Fault Tree only as a root cause analysis tool, this information may not be as critical.

If you ever sat through a course in basic probability, you probably remember that there were a lot of exercises dealing with finding the probability of some combination of events (i.e. what is the probability of drawing three green balls and four black balls from a Greek urn containing . . . . ?). Well, it turns out that there is an interesting relationship between these kinds of problems and the Fault Tree. The Fault Tree's logic gates serve to permit or inhibit the passage of fault logic up the tree. You saw the *Truth Tables* associated with the AND and OR gates above. The underlying body of knowledge that describes this logic is called **Boolean Algebra**.

Boolean Algebra allows us to consider the top event of the Fault Tree as a *function* of all the basic, intermediate, undeveloped and other events that appear below the top. Let's stretch our memories even further. In "regular" algebra, we were often given some equation: (remember "quadratic" equations) that expressed some "Y" as a function of some "X's." We were asked to solve the roots of the algebraic equation.

Find the "roots" of this equation:

$$y = f(x) = x^2 - 3x + 2 = 0$$

$$y = f(x) = (x-1)(x-2) = 0$$

$$Roots\ of\ y : x\ equals\ 1,\ 2$$

When we apply Boolean Algebra to our Fault Tree, we can also solve the "roots" of the equation graphically displayed by the Fault Tree. These "roots" turn out to be the fundamental combinations of "low-level" events (basic, undeveloped, house, etc.) that result in the top event. These roots are given a specific name: **Minimal Cutsets**.

There are two important reasons to "solve" a Fault Tree. First, in studying a complex system (one that incorporates redundancy), one of the main purposes in preparing a Fault Tree is to understand the *combinations* of equipment failures and operator errors that can result in system failure.

The Minimal Cutsets tell us the *minimum* number of failures that must occur for the system to fail. No more, no less. For example, a system that incorporates one level of redundancy (i.e. a system that is "single-failure proof") should have at least **two** equipment failures per minimal cutset.

Second, we can use this "algebraic" expression to estimate the probability of the top event occurring if we know the probabilities of the low-level events. Let's start with two basic logic combinations and see how this works.

The OR gate is represented by the union of two or more events, i.e. either event A OR event B occurs for the next higher event to occur. To calculate the probability of the union of two events, the laws of probability state that we add the probabilities of the two events and then subtract their product. So, when combining failure probabilities for events that enter an OR gate, these same rules will apply.

$$P(C) = P(A) + P(B) - P(A)xP(B)$$

*Note: If the two events, A and B, are **mutually exclusive**, that is, they cannot occur at the same time, the last term of this equation is zero.*

If the probability of failure for the input events to the OR gate are small (less than or equal to .05), then subtracting the probabilities' *product* can be neglected. The probability of the output event occurring is reduced to:

$$P(C) = P(A) + P(B)$$

This simplified calculation is called the *"rare event approximation."* For example, if the P(A) = 0.05 and the P(B) = 0.01, we can calculate both the exact probability and the rare event approximation:

EXACT PROBABILITY: P(C) = 0.05 + 0.01 - (0.05 x 0.01) = 0.0595

RARE-EVENT APPROXIMATION: P(C) = 0.05 + 0.01 = 0.06

You can see that there is only a small difference. Notice that the rare event approximation tends to *over predict* the probability of event "C" occurring. Since we construct and quantify a Fault Tree from the failure perspective, this approximation tends to predict a somewhat conservative estimate of the system's failure probability and *under predicts* the reliability or availability of the system.

Let's turn our attention to the AND gate. The AND gate is represented by the intersection of two or more events, i.e., event A and B must occur for the output event to occur. To calculate the probability of the intersection of two or more events we take the product of the events. The probability of the output event (C) occurring where there are two events (A&B) entering an AND gate is:

$$P(C) = P(A)\ x\ P(B)$$

For example, if the P(A) = 0.05 and the P(B) = 0.01, then P(C) = 0.05 x 0.01 = 0.0005.

Since the AND gate is used when the system incorporates redundancy, we can start to see the *quantitative* advantages to this design approach.

The discussion above assumes that the individual events are *independent* of each other. That is, the occurrence of one event does not affect the probability of the other event occurring. But what if the failure of one component increases the probability of failure of the others?

For example, consider the situation where there are two pumps operating in parallel and any one pump operating is sufficient to handle the load. If one pump fails, the probability of the second pump failing is increased due to the increase in load. To calculate the probability of *both* pumps failing in this case, we must use the conditional probability laws:

$$P(A \times B) = P(A / B) \times P(B)$$

or

$$P(A \times B) = P(B / A) \times P(A)$$

Practically, these kinds of relationships are very important to discover. The *design review* process should include steps to examine the design for these dependencies. They can be difficult to quantify, though.

A "real-life" example may help. Two redundant pumps were installed in a power plant backup system. Both pumps discharge into a common pipe. When the need arose, they were both signaled to start. Because of differences in their respective discharge pressures, one of the pumps would dominate the system. The other (lower discharge pressure) pump's flow would be shutoff from the system. Although there was a recirculation line present, the pump was predicted to eventually overheat and fail as it continually recirculated the same water. The dependency here is that the second pump will fail *given that the other pump operates.*

The INHIBIT-gate can also be used to model these conditional probabilities.

We've just examined how the graphical tool of Fault Tree Analysis is connected to the world of probability and event logic. Practically speaking, if you are in the business of solving large Fault Trees, you will be using one or more computer programs that perform the tedious calculations involved.

There are several Fault Tree analysis codes available. The Electric Power Research Institute has developed the CAFTA series of codes for use within the industry (currently available through Science Applications International Corporation (SAIC). EPRI has also developed several reliability-applications code packages (i.e. UNIRAM) that include a Fault Tree analysis code.

These codes help you build and edit the Fault Tree, solve for Minimal Cutsets and calculate the probability of the top event.

Fault Trees and FMECA

FTA uses a "top down" approach and only focuses on causes of the top event. It does not model all the possible system failures or all the possible causes of system failure. The Fault Tree is tailored to a particular failure mode, which is the top event, and shows only the faults that contribute to this failure mode occurring.

FMECA is a "bottom up" approach, and it looks at all failure modes of all items in the system and ranks each one with respect to their criticality.

The Fault Tree begins with a particular top event, typically a failure mode. Systems may have many failure modes, so determining the top event that will gain the most benefit from analysis can sometimes be difficult. The FMECA can lead the analyst to the failure modes that have the greatest impact on the system. These failure modes could be likely candidates for Fault Tree analysis.

There are several differences between the FTA and the FMECA analysis. A list of some of them follows:

- FTA can consider two or more concurrent events leading to the undesired event. FMECA looks at only one failure at a time.

- FTA can be quantified to determine the probability of the undesired event occurring. FMECA is basically qualitative in nature; it does not quantify the probability of the system failing.

- FTA can easily become large and frequently requires a computer to obtain the solution. The criticality assessment of the FMECA can be solved by simple arithmetic.

15.3.4 Fault Tree Analysis Procedure

The Fault Tree Analysis process is straightforward. As the size or complexity of the system being analyzed increases, the analysis can quickly become very complicated, though. Fault Trees developed for a system by different people can also *look* very different. Analyzing a complex system incorporating mechanical, electrical and electronic devices often requires a diverse group of engineers, operations and maintenance personnel.

1. EXAMINE THE SYSTEM TO BE ANALYZED

The first step in performing an FTA is to gain an understanding of the system's functions. If a FMECA has been performed on the system to determine the critical failure mode for analysis much of the information required to perform an FTA has already been developed (mission profile, reliability and functional block diagrams). For a typical analysis, the information required to perform the FTA could include the following:

- A mission profile for the system describing the system's functions, operating times, usage conditions, and mission phases.
- Functional or reliability block diagrams
- Assembly drawings and physical location drawings.
- Operation and maintenance procedures.
- Failure data and component/part reliability data.

2. DEFINE THE UNDESIRED (TOP) EVENT

The next step is to define the undesired event, i.e., what event must not happen. The top event is always stated as some "bad thing." Even if we are using the Fault Tree in a *reliability* study, we express the top event in terms of *unreliability.* The selection of this event is driven by the purpose of the analysis. Top events may focus on the unreliability/unavailability of the system or may focus on some damage that could be caused by system failure.

***Example**: Let's consider the humble lawn mower. If we are concerned about the reliability of the lawn mower, then the undesired event might be "Failure of the lawn mower to start and run." If we are concerned about the safety aspects of the mower, then undesirable events could include "Body part contacts rotating cutting blade," or "Injury due to lawn mower induced projectile." In all three cases, we are dealing with the same system, the lawn mower, but there are three very different Fault Trees that would be constructed based on these undesirable events.*

Here's an example of a top event associated with a relay failure mode:



3. ESTABLISH THE ANALYSIS BOUNDARIES

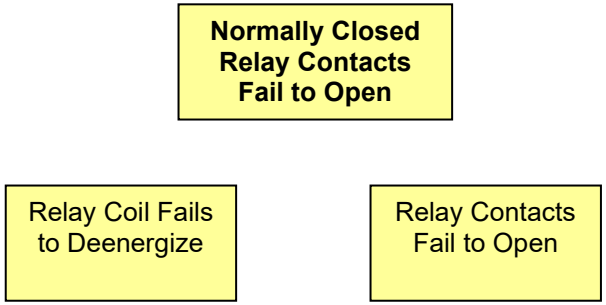
In this step, *boundaries* are established for the analysis. The *boundaries* term has several meanings. Many systems require “support” from other systems. For example, for a lateral to function, its associated feeder must function. In turn, the feeder’s substation must function, the transmission system must function, etc. The fault tree analyst, though, may only be interested in the lateral. The analyst would then set the system boundary at the feeder.

Since the fault tree starts with a specific top event, another boundary issue will involve the configuration of the system. A feedwater system being analyzed for “*failure to control steam generator level during startup”* will be in a different configuration than one at full power operation. In this case the term "boundaries" does not just refer to physical boundaries but includes the system's configuration, functions, operating and maintenance states.

4. LIST THE SUB-EVENTS

Once the top event is identified and system boundaries are established, the construction of the Fault Tree can begin. The top event is analyzed to determine the next level of faults that could cause the top event. These next level faults are known as sub-events.

Example:



There are several strategies that people have employed to break down a system:

Functional - The top event is usually stated in terms of the loss of some system function. But this system function is achieved by "sub-functions." In Unit 15.2, we presented a functional block diagram of an air compressor. The sub-functions needed for this air compressor to work included Compression, Cooling and Moisture Separation, Lubrication, Motive Force, and Instrumentation/ Monitoring.

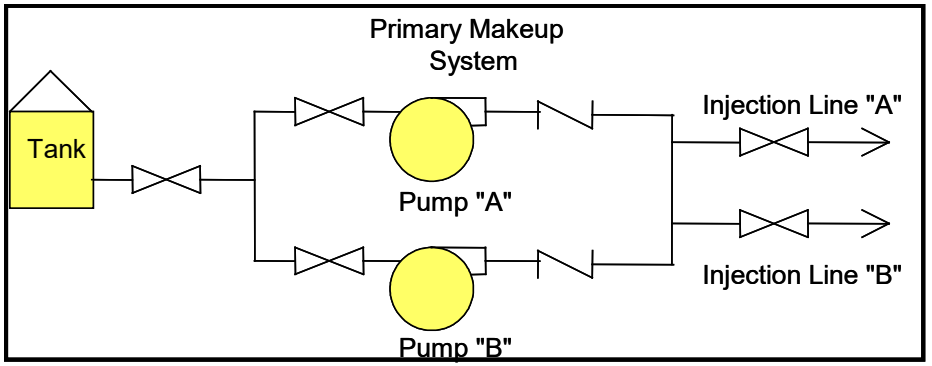
The top event can be broken down in terms of these sub-functions. As the analysis proceeds deeper, the same question is asked of the sub-functions. What are the "sub-sub-functions" that are necessary? At some point, these functions will be performed by discrete hardware components. This is where the breakdown will switch from functional to hardware.

System Segment/Nodal - For many mechanical and electrical systems, we can state the failure of the system in terms of an inability to provide flow to some point or cooling to equipment or to provide power at some point such as a bus, breaker or load. If we have a piping drawing or electrical "one-line" diagram, we can use these diagrams to break the system up into segments.

Nodal points are the junction of system segments. A "skeleton" fault tree can be constructed, simply by considering how the segments contribute to the failure of the system. Once this logic has been established, the failures of the components within the segments are included, putting the "meat" on the logic skeleton.

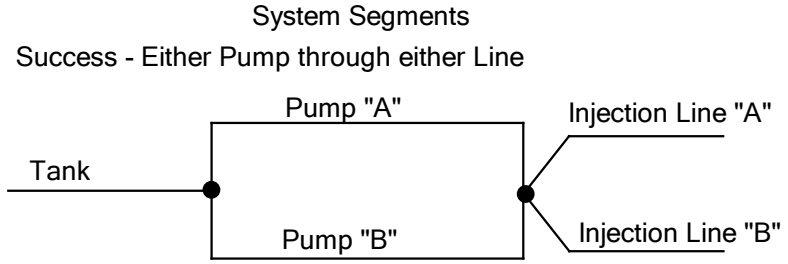
There are two factors that help you decide which pieces of the system are segments. The first is an understanding of the success criteria for the system. The second is the "independence" of the components.

Examples: Here's a simple makeup system designed to pump water from a tank to either a boiler, or primary coolant system:

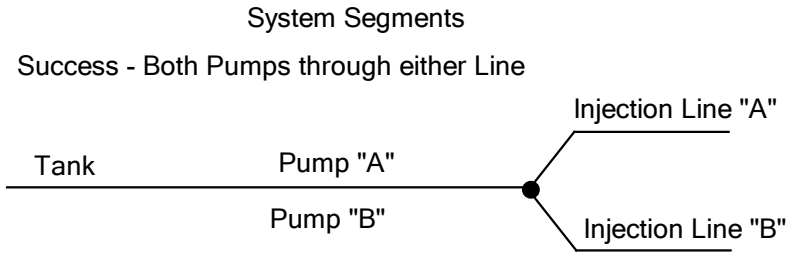


In the first example below, success is defined to be *either* pump injecting through *either* line. The injection lines are each segments since either line is sufficient (both must fail for the system to fail) and either pump (and associated valves) is sufficient. The tank segment is required no matter which pump or injection

line is successful.

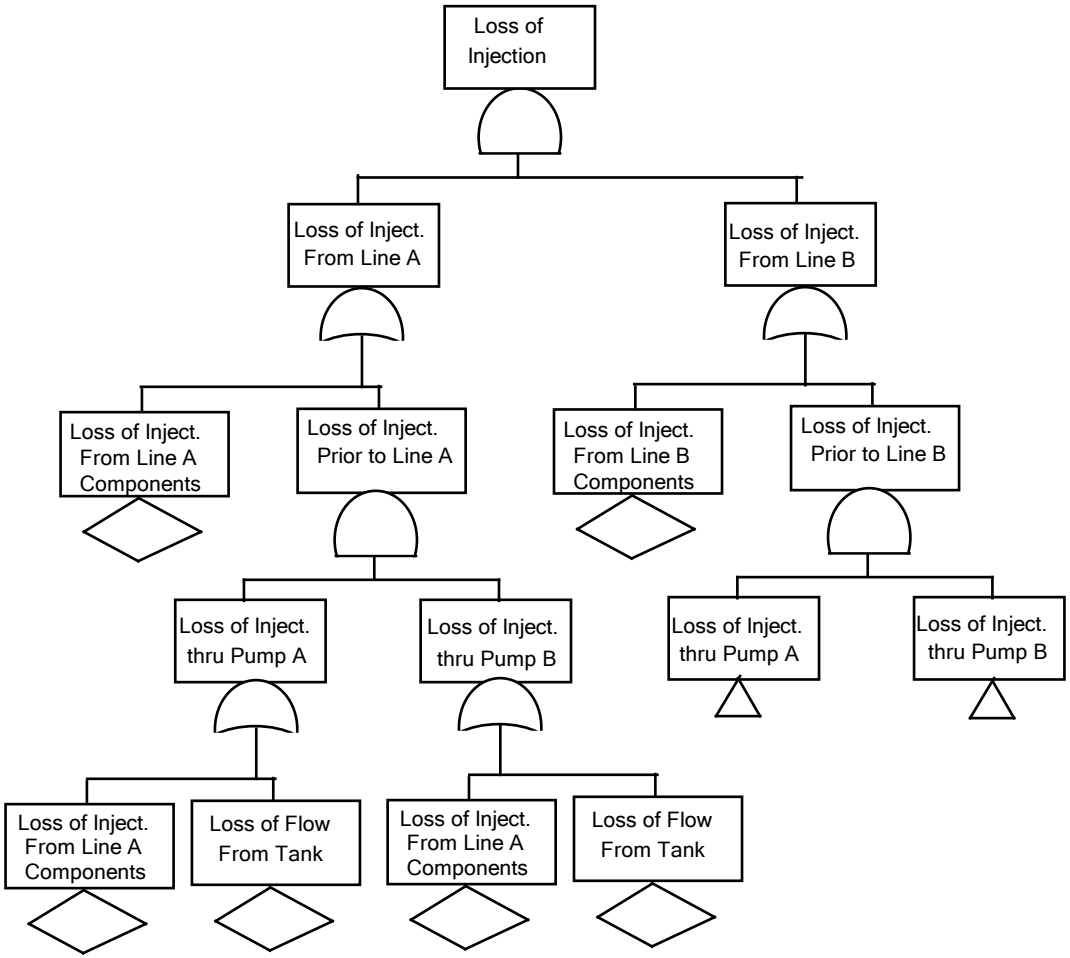


In the second example, *both* pumps are required, but *either* injection line is sufficient. In essence, we need all components functioning from the injection line node back to the tank for this configuration to be successful.



The *segment-level* fault tree for the first (either line, either pump) system success criteria follows:

EXAMPLE FAULT TREE
Injection System with Success =
Either Pump through either Injection Line

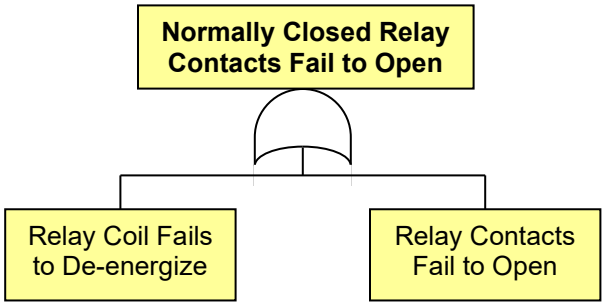


Notice how the logic events input to "Loss of Inject. Prior to Line A" are identical to those inputting to the event "Loss of Inject. Prior to Line B."

5. CONNECT THE SUB-EVENTS TO THE TOP EVENT

In this step, the sub-events are connected to the top event by logic gates. If any one of the sub-events can result in the top event occurring, then an OR gate is used. If all the sub-events must occur for the top event to occur, then an AND gate is used.

Example:



6. CONTINUE THE ANALYSIS

The same procedure is followed for each sub-event as was described for the top event. Each sub-event is analyzed to determine the next lower level of fault that could contribute to the sub-event occurring. These next lower-level sub-events are connected through the same use of logic gates to the next higher level of sub-events. Continue to analyze and break down the sub-events until the primary event level (basic, undeveloped or house) is reached. For all but the simplest systems this requires many levels of analysis.

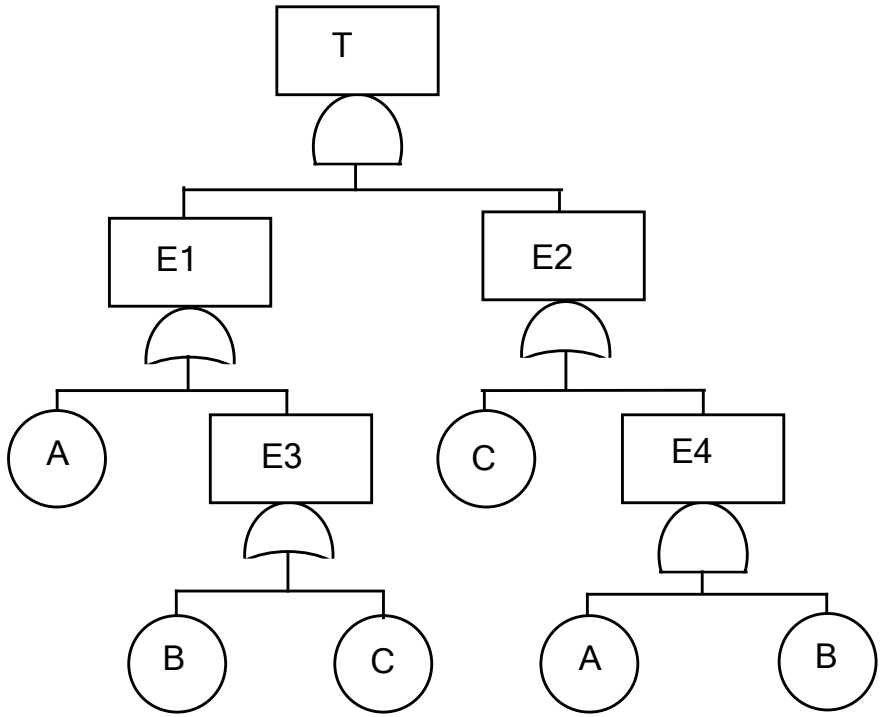
***NOTE**: Depending on the purpose of your Fault Tree, these next steps may or may not be necessary. If you have used the Fault Tree to aid in a root cause analysis, your attention will now turn to examining which basic events are the real root causes (i.e. verification of root causes). You will begin to look for evidence confirming or eliminating the "suspects" listed on your Fault Tree. If you are using the Fault Tree to aid in a system analysis, then these next steps are necessary. As mentioned above, computer codes are available which have automated these steps.*

Steps 7 through 9 describe the evaluation of a Fault Tree. This evaluation can be performed in two ways. First, an expression based on Boolean logic is developed using AND and OR gates to describe the combinations of the primary events required to result in the top event. This is referred to as *qualitative* analysis. If probabilities of the primary events

occurring are known, this expression can be used to calculate the probability of the top event occurring. This is referred to as *quantitative* analysis.

7. QUALITATIVE ANALYSIS

The qualitative analysis is performed in a two-step process. First, the Fault Tree is "expanded," that is, the intermediate gates are eliminated, and the top event is expressed as a function of the basic, undeveloped and house events. The two methods of obtaining this expression are the *top-down* and the *bottom-up* approach. We'll use the following fault tree to illustrate these approaches:



TOP-DOWN - To obtain the expression for the Fault Tree using the top-down approach you start at the top event and work your way down through the lower levels of the Fault Tree replacing the AND and OR gates with intersection and union symbols:

NOTE: This text will use the "engineering" symbols convention of ***x*** *(AND) and* ***+*** *(OR). Some books use the "mathematical" symbols of* $\cap$ *(AND) and* $\cup$ *(OR).*

Expanding:

$$T = E1 \times E2$$

$$E1 = A + E3$$

$$E2 = C + E4$$

$$E3 = B + C$$

$$E4 = A \times B$$

Substituting:

$$T = (A + (B + C) \times (C + (A \times B))$$

BOTTOM-UP - The bottom-up approach develops the expression by replacing the gates at the bottom of the Fault Tree with the intersection or union symbol and continues upward to the top event.

Starting with Lowest Gates:

$$E3 = B + C$$

$$E4 = A \times B$$

Then, Proceeding Upwards:

$$E1 = A + E3 \text{ and } E2 = C + E4$$

Substituting in for E3 and E4:

$$E1 = A + (B + C) \text{ and } E2 = C + (A \times B)$$

$$T = E1 \times E2$$

$$T = (A + (B + C)) \times (C + (A \times B))$$

8. FAULT TREE REDUCTION/SIMPLIFICATION

The next step in the qualitative analysis is to simplify or reduce the expanded expression. The Boolean algebra expressions developed above are not in their simplest form. Generally, this reduction step will produce the set of **Minimal**

Cutsets associated with the top event. The reduction of the Fault Tree to Minimal Cutsets is useful for both qualitative and quantitative evaluation of the Tree. The rules of Boolean algebra, shown below, are necessary to perform this simplification. Notice that many of these rules are similar to those of numerical algebra, although there are some notable differences (e.g. Idempotent Law and Law of Absorption).

Boolean Algebra "Rules"

Designation	***Symbolism***
Commutative Law	$X \times Y = Y \times X$ $X + Y = Y + X$
Associative Law	$X \times (Y \times Z) = (X \times Y) \times Z$
Distributive Law	$X \times (Y + Z) = X \times Y + X \times Z$
Idempotent Law	$X \times X = X$ $X + X = X$
Law of Absorption	$X \times (X + Y) = X$ $X + X \times Y = X$
Complementation	$X \times X' = \Phi$ *(The Null Set)* $X + X' = \Omega$ *(The Universe)*
deMorgan's Theorem	$(X \times Y)' = X' + Y'$ $(X + Y)' = X' \times Y'$
Unnamed, but frequently used relationships	$X + X' \times Y = X + Y$ $X' \times (X + Y') = X' \times Y' = (X + Y)'$

A *minimal cutset* is the smallest combination of primary events which all must occur for the top event to occur. The minimal cutsets represent the modes by which the top event can occur. For example, the minimal cutset A, B means that both primary events A and B must occur for the top event to occur. If either A or B does not occur, then the top event will not occur. If all the minimal cutsets for a system are found, the top event can be expressed as the *union* of the sets. For *N* minimal cutsets,

$$\text{TOP EVENT} = M_1 \text{ or } M_2 \text{ or } .... \text{ or } M_N$$

Each minimal cutset (M_i) then consists of the intersection of the minimal number of primary events required to produce the top event. From a *qualitative* standpoint, knowing the minimal cutsets for a Fault Tree can provide valuable information about potential problem areas of a system. The qualitative approach can be useful in ranking the minimal cutsets by the number of primary events in each set, in depicting which failures must be removed in order to remove the top event; in assessing the importance of particular component failures' effects on the system; and in assessing the system's susceptibility to common-mode failures (failures affecting multiple equipment due to the same *mode*, i.e. assembly, maintenance or design).

Minimal cutsets are typically referred to as *singles, doubles, triples* and so on, depending on the number of primary events in the cutset. This allows the emphasis to be placed on the cutset with the smallest number of primary events.

Designing a Nuclear Power Plant requires that no *single* component failure cause system failure. This is equivalent to saying that all single cutsets must be removed from the Fault Tree. This same logic can be applied when assessing qualitatively the importance of a component. If a component appears in one or more single or double cutset, it is likely to have a large effect on the system.

Expanded Equation of Top Event :

$T = (A + (B + C) \times (C + (A \times B))$

First, simplify using the associative, then commutative laws :

$A + (B + C) = (A + B) + C = C + (A + B)$

$T = (C + (A + B)) \times (C + (A \times B))$

Next, use the distributive law :

$T = C + ((A + B) \times (A \times B))$

Finally, use the law of absorption :

$T = C + (A \times B)$

The terms "C" and "A × B" are the minimal cutsets for this Fault Tree.

9. QUANTITATIVE ANALYSIS

Once the Fault Tree's top event has been expressed in terms of its Minimal Cutsets, the tree can be quantified. There are several methods developed for Fault Tree quantification.

An "exact" solution of a large fault tree requires tremendous computation power and time. All the practical methods are approximations of the exact solution; they generally "over predict" the probability of the top event. Since the top event is expressed as a "bad thing," this is conservative. The Minimal Cutset strategy is often employed in computer codes, and since it is the simplest, it will be described here.

The first step is to obtain failure probabilities for the primary events (exactly *how* these probabilities are obtained is discussed in Unit 15.4). The individual Minimal Cutset probabilities are then found by multiplying the probabilities of the cutset basic events (i.e. failure probabilities). The individual cutset probabilities are then summed to obtain an approximation of the top event probability.

Example:

The minimal cutsets for the example Fault Tree are C and A × B.

$$If\ P(A) = 0.02,\ P(B) = 0.07\ and\ P(C) = 0.05,\ then,$$

$$P(T) = P(C) + P(A) \times P(B) = 0.05 + 0.02 \times 0.07$$

$$P(T) = 0.05 + 0.001 = 0.051$$

If events A, B, and C represented equipment failure probabilities and T was the failure of a system, we would estimate the system's failure probability to be 0.051 (the system success probability or reliability would be 1 - 0.051 = 0.949).

If we were not satisfied with this level of reliability, the quantified minimal cutsets help us to prioritize our improvement efforts. Here, improvement in the reliability of equipment "C" is our best choice, even though equipment "B" has a higher failure probability.

15.3.5 Fault Trees and Reliability Block Diagrams

In Unit 15.2, we discussed Reliability Block Diagrams and their use in depicting the system's success paths. Fault Trees and Reliability Block Diagrams can be equivalent logic models. When the top event described for the Fault Tree is the failure of the mission described for the reliability block diagram, they are the same. Calculations involving the probability of failure for the Fault Tree will result in one minus the probability of success calculated in the reliability block diagram.

The reliability block diagram is a success model while the Fault Tree is a failure model. Some reliability engineers prefer the Fault Tree to the Reliability Block Diagram or vice versa. Drawing success models for a complex system is an expensive and time-consuming operation. Many success models may be required to cover the various definitions of success. When failures are considered, it may be necessary to construct only one or two failure models, which cover all the significant failure modes.

15.3.6 Fault Tree Computer Codes

Even for simple systems a Fault Tree can quickly become large, and the evaluation can be very difficult. Computer codes are available to aid in the analysis. Most of the codes deal with computing the minimal cutsets of the Fault Tree. Some of the codes/software available are listed below:

CAFTA: Available through Electric Power Research Institute (EPRI), SAIC, IBM PC compatible

GRAFTER: Available through Westinghouse, IBM PC compatible.

UNIRAM: Available though EPRI. Performs unavailability calculations.

SETS: Developed by Richard Worrell (formerly of Sandia National Laboratories), it is very useful for large Fault Trees (SETS has been installed on a CRAY computer). SETS is available in a PC version through Richard Worrell, Albuquerque, NM.

15.3.6 Fault Tree Summary

The Fault Tree is a tool useful for developing a Cause-and-Effect logic relationship. This relationship can be explored for many different applications:

- Performing a Root Cause Analysis of a system, equipment or component.
- Developing a reliability/availability model of a system to:
  - Examine a new or existing design for potential improvements, or
  - Evaluate the reliability impact of changes proposed to an existing design
- Examine a system from a safety perspective.
- Integrate the effect of equipment failures and human errors on system reliability.
- Assist in the diagnosis of degrading conditions that can impact system reliability.

For a simple system or equipment, Fault Trees can be drawn and evaluated by hand (Berol RapiDesign Templates R-555 and R-542 may be obtained at any good office supply store). For large, complex systems, a computer code such as EPRI's CAFTA is necessary to both develop and evaluate the Fault Tree.

15.4 Quantifying Reliability

Learning Objectives

- Understand and be able to calculate core RAM measures
- Be able to develop and interpret a failure model for a product or part

Unit Contents

- Quantifying Reliability Characteristics
- Gathering Failure Data
- Developing a Failure Model
- Quantifying Maintainability Characteristics
- Quantifying Availability Characteristics
- Applications
- Weibull Analysis
- Why The Traditional Reliability Prediction Models Do Not Work

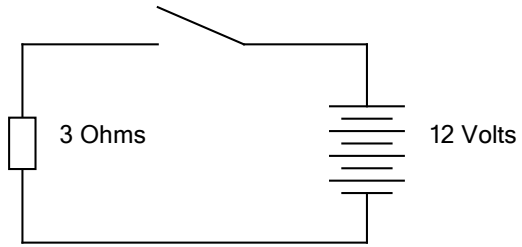
15.4.1 Introduction

In this Unit, we will introduce the statistical nature of Reliability, Maintainability and Availability. Basically, the Reliability discipline occupies itself with understanding the nature of failures. FMECA, PF/EMEA and FTA help us organize our knowledge about **how** and **why** an item can fail and its **effect** on the system or process. Although we have mentioned that these tools may use frequencies of failure and probabilities of failure, we have not indicated *how* these quantities were obtained. To obtain the answer to the questions **when** or **how often** will the item fail, we need to *boldly* enter the world of probability and statistics.

Here are some simple reliability questions that might arise in our work and for which a *numerical* answer is required:

- What's the probability a Chiller Unit will start the next time it's needed?
- How often do we expect an ice-skating cooling system to have a forced shutdown?
- How many spare parts (and what kind) do we need to stock for these circuit breakers?
- Will the copier make it through my job without breaking down?
- How long will it take to restore service to a customer?
- Can we skip maintenance on this valve until the next outage?
- How many home air conditioner failures will consumers experience this year?
- Is it economical to repair or replace this piece of equipment?

To answer these questions (using facts and not guesses!), we often need to gather some data, do some *statistical* analysis of the data and then project the results of that analysis into the future (now we're dealing with the *Probability* part of the equation). Reliability generally deals with models of the world that are *probabilistic* rather than many engineering models you are used to working with that are *deterministic*. What's the difference? Here is a simple circuit composed of a battery, a switch and a resistor:



If we close the switch, a current (I) will flow through this circuit. There is a physical relationship between the circuit's voltage, its resistance and the current. We can predict that the current in this circuit will be 12 Volts/3 Ohms or 4 Amps. No matter how many times we close this switch (given that the battery stays fresh), four Amps of current will flow through the circuit.

This type of physical situation is *deterministic* in nature. The characteristic of *electrical current* is “completely” determined by the battery, resistor, wire, and our ammeter.

Now let's focus on a different characteristic, the *time to failure* of the resistor. We will close the switch and start a clock. After several weeks or months, we may note a puff of smoke and discover the resistor has burned out.

If we do this experiment again, will we see the resistor fail at the *same* time (or even within a few hours of the first one)? Or will we see that perhaps the time to failure varies by several days or even weeks? Is there a "formula" that allows us to predict the resistor's time to failure based on the circuit's current and voltage and some set of factors peculiar to the resistor?

The reliability characteristic has not (in general) been kind enough to allow us to develop exact formula and equations to predict its values. Rather, we have had to rely on *probabilistic* models to describe this characteristic. This is a major distinction between reliability engineering and other "brands" of engineering you may have studied.

For those unfamiliar with the methods of probability, it often represents a significant barrier to their understanding of reliability management. As you become familiar with reliability methods, prepare yourself to communicate with people who do not have your knowledge. From our experience, these can often be interesting and amusing conversations.

One of the major challenges faced by Reliability engineers is in understanding the *time* and *life-behavior* of systems, equipment, components and parts. In many cases, the paucity of data will force us to try and get a great deal of information out of a very few failures. In these cases, the "art" of reliability becomes a combination of statistical methods with detailed physical investigations of the failure phenomena.

15.4.2 Quantifying Reliability Characteristics

In Unit 15.1, we defined reliability as a *probability*, so our first step is to learn how to *quantify* this characteristic. Without quantification, it will not be possible to compare our reliability to customer specifications, or to require a certain reliability from a vendor, or, in general, to make decisions about the reliability of our systems, equipment and components. Before we jump into the mathematics, let's fit this quantification effort into a "bigger picture." When and where will we find the need to do some of these calcs?

The Plan-Do-Check-Act (PDCA) cycle helps put this issue into perspective. Many companies, such as utilities, railroads and telecommunications have extensive, *existing* systems. The most frequent type of reliability analysis here will start with a "Check" of existing system (or equipment, component, or part) reliability characteristics. We will use the PDCA cycle in its "CAP-DO" form. Here, the reliability calcs will be part of a broader analysis that will help us make decisions leading to system reliability improvement.

***Example**: Valve failures were causing outages in a refrigeration system. As part of the "Check-Act," the maintenance department gathered data on previous failures, performed a Weibull analysis of the failures as part of a root cause analysis and concluded that the valve's design was not adequate to reliably support the valve's mission. The valve's design was modified (the "Plan" was revised), and operation of the system ("Do" phase) will be monitored to determine if the design change is successful in preventing the valve's failure (golly, we're all the way back to the "Check" phase!).*

When we are designing a *new* system, then we start in the "Plan" phase of PDCA. Here, reliability calculations can help us determine if the proposed system can meet the reliability goalst. In fact, using this approach, we are rotating the PDCA cycle *within* the "Plan" phase. If the "paper" system does not meet the reliability requirements, then we must revise the design until it does.

***Example**: A combined-cycle generating unit was being designed with an Equivalent Forced Outage Rate (EFOR) target of 0.02. A reliability model of the initial design was constructed and quantified. This model predicted an EFOR of 0.04. The major contributors to this EFOR (as predicted by the model) were the controls for the combustion and steam turbines. The controls design was then subjected to detailed FMEA and FTA; numerous design improvements were identified and implemented. Based on these improvements, a revised prediction calculated unit EFOR to be 0.015, within the design target. The completed controls systems were also subjected to reliability life testing, including environmental stress screening. Here, a PDCA cycle occurred completely within the "P" phase of this combined-cycle unit's life.*

This last example sounds simple, but to accomplish the "prediction" task requires that we know both the critical components in the unit (those whose failure could result in a Forced Outage) *and* their failure probabilities.

In the following sections, we will present the calculations, as they will occur in a "CAP-DO" cycle of existing systems. Unit 15.5 - Root Cause Analysis describes the activities that will occur to understand the physical phenomena resulting in the failures that will lead us to potential reliability improvements.

15.4.3 Gathering Failure Data

The process of quantifying the **reliability** of some equipment, component or part starts with gathering some data that describes the item's failures. Here are the data elements that comprise a *minimum* set of information needed for reliability calculations:

- Identification of the equipment included in the population of interest (by tag number, serial number, manufacturer, etc.)
- Modes by which equipment has failed
- Cause of failure
- Operating time between or to failure (time can be operating hours, cycles, etc.)
- Operating time for items that have not yet failed
- Number of failures

Generally, a company interested in managing the reliability of its systems will set up some sort of data collection system. For example, at a western railroad, a Microsoft Access database was created to track failures of locomotives leaving a repair facility. Japanese and French nuclear utilities have extensive failure databases for their plants. Military weapons system managers collect and analyze failure data.

Example (Time-related failures): *At a Japanese nuclear plant, reliability data is maintained for plant equipment. An engineer submitted a request for a reliability report on the Main Feedwater Pumps. The computer printout included "engineering" data on the three 60% capacity pumps in the plant, such as their tag numbers, pump, motor and coupling type, manufacturer, rating and installation dates. Operating times were identified in terms of hours of operation, each time the pump was started and stopped the date and time (to the hour) was recorded. Reasons for stopping the pump were listed (including normal plant shutdown as well as pump planned and corrective maintenance). For each corrective maintenance action, a root cause analysis summary was included, as well as a reference to a more detailed, written report. For each planned maintenance action, a list of the actions performed, pump parameters inspected (as found and as left values) and parts replaced or refurbished was included.*

Example (Cycle-related failures): *Emergency Diesel Generators (EDG) are tested by starting and running the machine for a given time. If the EDG does not achieve rated speed and voltage within 10 seconds of the start signal, the "start" is a failure. The number of start failures (of the last 100 valid start attempts) is tracked and reported to a regulatory agency quarterly.*

15.4.4 Developing a Failure Model

Next, we will take our raw failure data and begin to apply the statistical process. You will calculate the mean (or some other central tendency measure) and the variance (or some other dispersion measure). You will develop a "picture" of the data such as a histogram of times to failure or a probability plot (i.e. Weibull chart). You can determine the uncertainty in your estimates of these reliability characteristics by calculating a confidence interval for the population parameters. From this analysis, we will be able to apply the definition of reliability and calculate part, component or equipment reliabilities associated with given missions.

Let's demonstrate this with an example: Suppose you have 90 "items" placed in service at time "0" and the following failure times (in hours) were observed. The 90 items are a sample of size "n" from the population of items we are studying:

0.83	0.56	0.27	0.04	1.20	1.75	1.83	1.95	1.68	1.63
1.22	1.73	1.49	1.77	1.70	1.29	1.48	1.82	1.92	1.83
1.64	1.97	1.51	1.88	1.99	2.40	2.76	2.54	2.84	2.39
2.22	2.91	2.47	2.28	2.24	2.75	2.21	2.26	2.68	2.30
2.94	2.90	2.22	2.33	2.96	2.53	2.27	2.38	2.26	2.38
2.51	2.18	2.18	2.57	2.30	3.34	3.56	3.85	3.99	3.14
3.15	3.13	3.98	3.59	3.51	3.92	3.96	3.66	3.33	3.41
3.71	3.74	3.30	3.46	3.40	3.26	3.92	3.39	3.44	3.55
4.64	4.48	4.42	4.85	4.40	4.39	4.43	4.59	5.45	5.27

Mean Time To Failure (MTTF)

In our example, we observed the items from "birth to death." If these items could not be repaired, like light bulbs, we call their average life the **Mean Time To Failure (MTTF)**.

$$MTTF = \frac{\sum \text{failure times}}{\text{number of failures}} = \frac{\sum t_i}{n}$$

NOTE: If, in addition to failure times, operating times of un-failed equipment are available, these are included with the times to failure. The MTTF equation becomes:

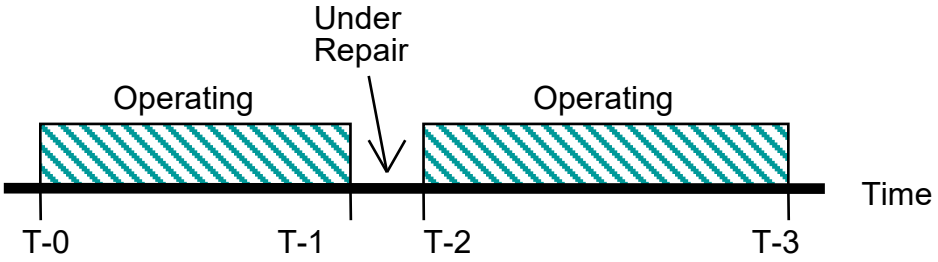
$$MTTF = \frac{\text{Total Operating Time}}{\text{number of failures}}$$

For the 90 times to failure listed above, this calculation would appear as below:

$$MTTF = \frac{0.83 + 0.56 + 0.27 + \ldots\ldots + 4.59 + 5.45 + 5.27}{90} = 2.74\, hours$$

Mean Time Between Failure (MTBF)

If the items in our example *could* be repaired, then we use the term, **Mean Time Between Failures (MTBF)**. The calculation of MTBF is identical to MTTF with this important note: the failure times *do not* include time needed to repair the item. In the timeline below, we would take T-1 minus T-0 as the first time to failure, and T-3 minus T-2 as the second time to failure. The repair time T-2 minus T-1 is not included in the MTBF calculation.



Dispersion in Times to Failure

While the *dispersion* in times to failure is important, no special term is applied to any one measure of dispersion. The *range, variance* or its square root, the *standard deviation* are useful measures of dispersion. Unit 6.3 includes the formula to calculate these statistics. For example, the *range* and *standard deviation* of the 90 times to failure are:

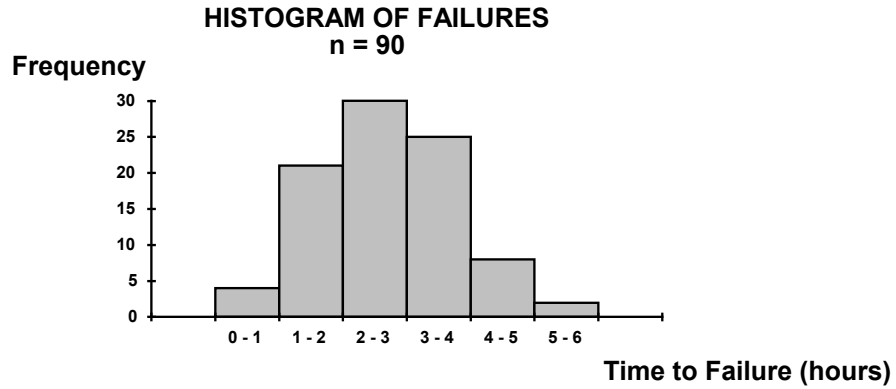
Range = 5.45 hours - 0.04 hours = 5.41 hours

Standard Deviation (s) = 1.10 hours

Knowing *both* the mean and standard deviation of the data can provide us with useful information about the variability of the components' life. A useful thumb rule to remember is that generally over 95% of the data will fall within plus or minus *three* standard deviations of the mean. For our data, 2.74 +/- (3 x 1.10) gives us a range of 0 to 6.04 hours.

The "Picture" of Failure Times & Probability Models

There are essentially two ways to get a "picture" of the sample of data. If there is enough data (i.e. at least 20 - 30 failures), then we can construct a **histogram** of the data. The 90 failures described above are shown in histogram form below:



Notice that this histogram looks somewhat like a *probability density function*. By examining the shape of this histogram, we can "guess" at which probability distribution may fit the failure data. Does the histogram look normal, exponential, log-normal?

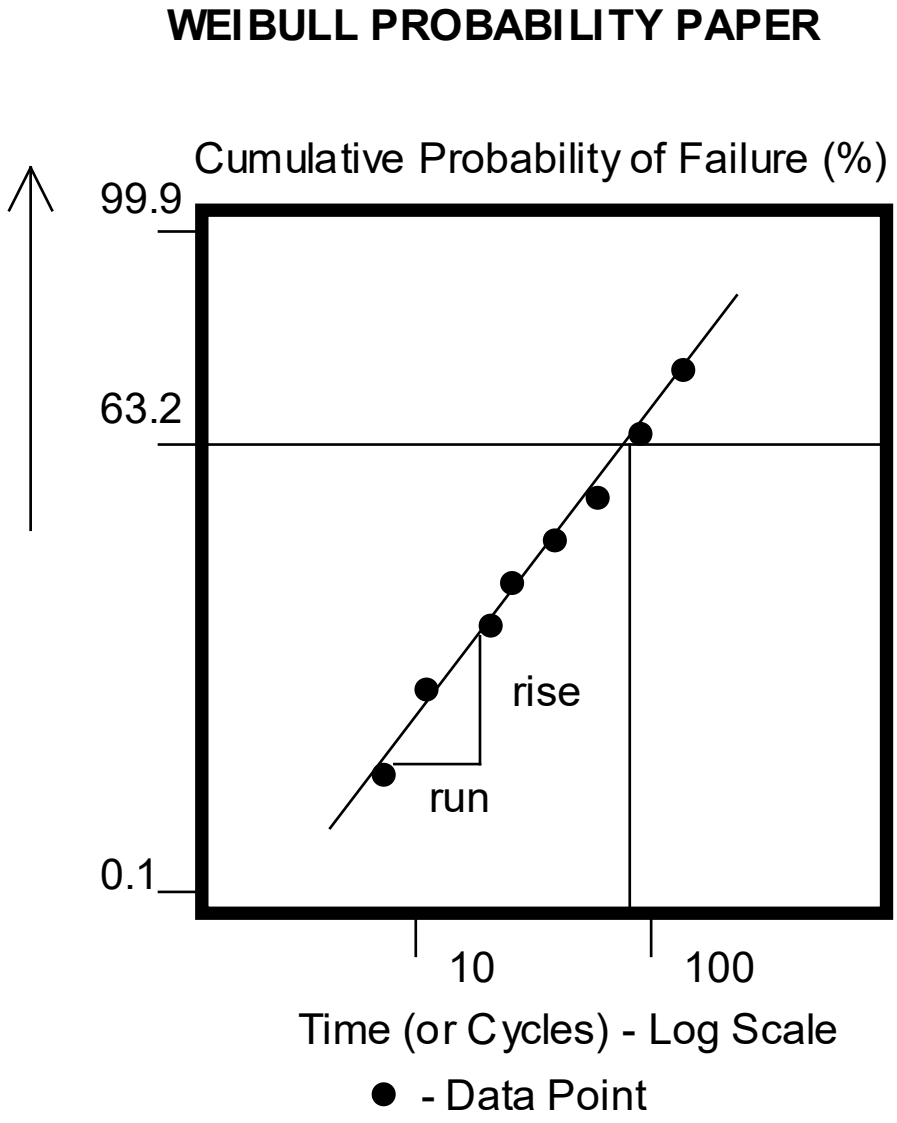
Here, we may hypothesize that the **normal** distribution could be used to model this item's failure pattern. The data is symmetric around the mean and “falls off” quickly as we move away from the mean on either side. The histogram approach works well when we have "enough" failure data to get a clear picture of what's happening. An alternative approach is to make use of **probability paper**, such as is used in Weibull Analysis. This is a very powerful approach and has been widely used.

The probability paper approach plots the data in the form of a *cumulative density function*, the integral of the probability density function. Probability paper is superior to the histogram approach when you are working with only a few failures (less than the 20 to 30 needed for a histogram). Although we must be aware of the large uncertainties in dealing with small sample sizes, probability paper does give us a picture to work with.

When we use probability paper, the *population parameter* estimates are often read directly from the plot of the data. For example, the Weibull distribution parameters β, η, and t_0 are obtained directly from the graph. There may not be a need to calculate the *mean* and *standard deviation* of the data in this case, since the plot tells us all we need to know. Basic Weibull Analysis methods will be presented below.

Relationships between Statistics and Parameters

We can take any set of data and calculate *statistics* such as the mean, variance, etc. As shown in Unit 9.1, probability distributions are described by *parameters* such as λ, β, or η which may be neither the mean nor the variance of the data. Unit 9.1 provides you with formulae that relate these distribution parameters to the *mean* and *variance* statistics. It takes a little algebra, but if you



calculate the mean and standard deviation from your sample data, then you can obtain the equivalent estimates of the distribution parameters.

Example: The simplest case is the exponential distribution. If the failure data can be described by this distribution, then the failure rate, λ, is simply the reciprocal of the MTTF or MTBF.

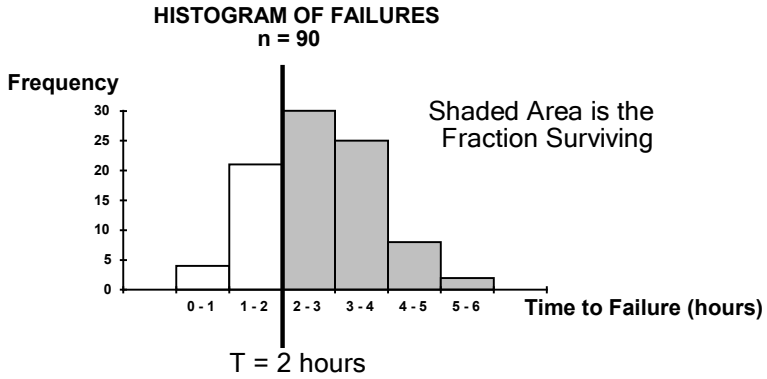
Interpreting the Failure Model

So far, we have collected data on times to failure, developed estimates of the central tendency and dispersion, constructed a picture of the data and made a decision regarding which probability model works best with our data. Now we can interpret this information.

Obtaining Reliability Estimates from Times to Failure

Once again, the definition of *Reliability* includes the *probability* an item will perform its function for at least time "T," where "T" is the mission time. We can examine our histogram at time "T" and find the fraction of items that failed past time "T." This fraction is a pretty good estimate of the item's reliability.

For example, what is the reliability of our "item" at T = 2 hours?



The fraction surviving is the quotient of the number surviving (30 + 25 + 8 + 2 = 65) and the total number of items (90):

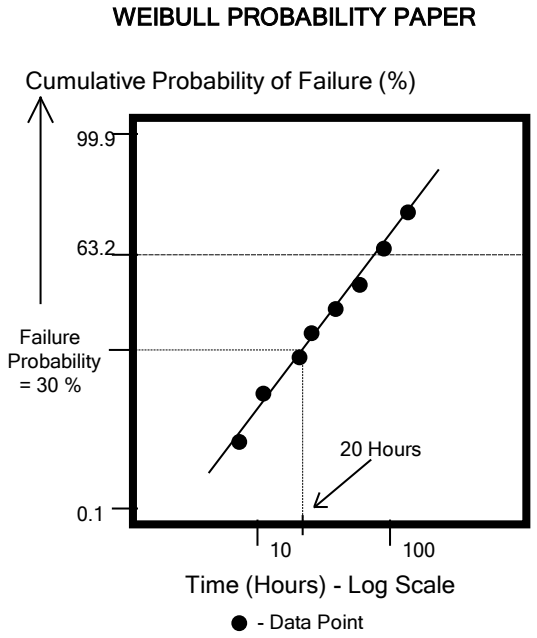
$$\text{Fraction Surviving} = \text{Reliability} = \frac{65}{90} = 0.722$$

Is this value of reliability OK? Does it need improvement? Why do the items fail in the pattern we've observed? If we change the mission time of this item, what will the reliability impact be? What can we do about this situation?

In general, the Reliability of an item is expressed in terms of either the probability density function or the cumulative density function:

$$\text{Reliability} = R(T) = 1 - \int_0^T f(t)dt \;\; or \;\; R(T) = 1 - F(T)$$

When we use the histogram of failures, we are, in essence, integrating the probability density function, *f(t)*, as in the first equation. When we use the probability paper approach, as in Weibull Analysis, we obtain the *un*reliability by examining the cumulative density function as shown by the second equation.



Of course, the reliability of the item is simply 1 - F(T).

Obtaining Reliability Estimates from Success/Failure Attempts

The preceding discussions developed reliability estimates from the *times to failure* observed and knowledge of how long the item must perform its function. This approach will be applicable to many types of equipment, components and parts.

Some devices, though, are of the "go/no go" type. When the item's function is demanded, either the item works or it doesn't. When current or voltage reaches a certain value, a fuse either opens or it doesn't. When a motor-operated valve is demanded to open, it does or doesn't.

Some equipment's failure modes will include those that are related to time of operation *and* those related to a go/no go situation. Our Diesel Generator can fail to run (time-related) or it can fail to start when demanded (go/no go).

Reliability estimates are obtained in a straightforward manner for these go/no go situations:

$$\text{Reliability} = \text{R} = \frac{\text{Number of Successes}}{\text{Number of Demands}}$$

If our Diesel Generator failed to start once in the last 100 demands, then an estimate of its reliability is 0.99.

The Binomial distribution can often be used to model these success/failure situations. "R," the reliability, is equal to *(1 - p)*, if *p* is taken to be the "fraction defective" or fraction of failures.

Obtaining Interval Estimates of Reliability Parameters and Reliability

Unit 9.1 discusses the uncertainty in our estimates of the population parameters introduced by our dealing with a limited sample of data. Since the item's *Reliability* is estimated from this same sample of data, there will be uncertainty regarding this characteristic, too. In some cases, it will be sufficient to understand the uncertainty associated with one of the reliability parameters.

For instance, if a vendor claims that, based on historical data, the MTBF of a variable speed motor is 10,000 hours, we should ask the vendor for the confidence interval associated with the estimate. What is the error of the estimate and at what confidence level is this error stated?

The procedures listed in Unit 9.1 may be employed to obtain this information if the motor time to failures can be modeled with a normal distribution.

When we are dealing with the times between failures of equipment (composed of many parts and, hence, many failure modes), we can often use the *exponential* distribution as a model of this process. When this is the case, a confidence interval can be constructed for the MTTF or MTBF through use of the *chi-square* distribution. We must introduce two new terms here, since there are two calculations of this confidence interval that are possible:

Failure Truncated "Tests" - If our data only includes times between failures, then each "test" was ended or *truncated* by failure. In this case, the "1 - α" confidence interval (for a 95% confidence interval, α is 5% or 0.05) is:

$$\frac{2\sum t_i}{\chi^2(1-\alpha/2,2r)} \geq MTBF \geq \frac{2\sum t_i}{\chi^2(\alpha/2,2r)}$$

$$where:$$

$$t_i \text{ are the times to failure}$$

$$\text{r is the number of failures}$$

Time Truncated "Tests" - If our data includes a mixture of times between failure *and* operating times not associated with failures, then we have truncated the "test" based on time. In this case, the "1 - α" confidence interval is:

$$\frac{2\sum t_i}{\chi^2(1-\alpha/2,2r+2)} \geq MTBF \geq \frac{2\sum t_i}{\chi^2(\alpha/2,2r+2)}$$

$$where:$$

$$t_i \text{ are the operating times (both failed and non - failed)}$$

$$\text{r is the number of failures}$$

Notice that we can obtain a confidence interval for time truncated tests even if no failures have occurred in the time interval. If the data can be modeled with the exponential distribution, then λ (the exponential parameter) is the reciprocal of the MTBF. For this special situation, we can obtain a confidence interval for the equipment Reliability. We are generally interested only in a *lower* confidence bound on reliability. The "1 - α" upper bound on the failure rate λ is:

$$\lambda_{upper} = \frac{\chi^2(1-\alpha,2r)}{2\sum t_i}$$

$$\text{and the lower bound on Reliability is :}$$

$$R_{lower}(T) = e^{-\lambda_{upper} T}$$

$$\text{where :}$$

$$T \text{ is the mission time}$$

We can do the same thing if we are calculating the reliability of a "go/no go" device. Here, an approximation for the lower bound on reliability is given by:

$$R_{lower} = \hat{R} - z_{\alpha}\sqrt{\frac{\hat{R}(1-\hat{R})}{n}}$$

$$where:$$

$$\hat{R} \text{ is the point estimate of Reliability,}$$

$$z_{\alpha} \text{ is the standard normal deviate,}$$

$$\alpha \text{ is one minus the confidence level, and}$$

$$n \text{ is the number of go/no go attempts}$$

Clues from the Picture - the Bathtub Curve

In many cases, we will be doing these reliability calculations to support a root cause analysis of our item. You will be interested in discovering the physical mechanism of the failures; detailed inspections and laboratory analysis may be needed here. You will also ask yourself *why* this item failed:

- Was the design inadequate?
- Could periodic maintenance have prevented these failures?
- Was the manufacturing process flawed in some way?

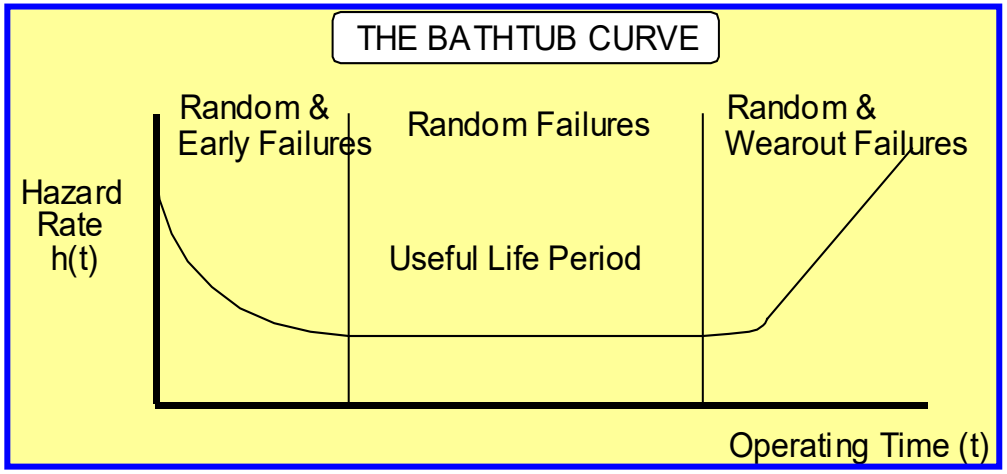
The picture of the data can be examined to help point you in the right direction. We will introduce a concept known as the **hazard rate** to demonstrate this concept. The **hazard rate** is the conditional or instantaneous failure rate of an item, given that the item has survived to time "t." It is defined as:

$$h(t) = \frac{f(t)}{R(t)} = \frac{f(t)}{1 - F(t)}$$

If the device is still working at time "t," then its probability of failure in the next time interval, "Δt," is $h(t)\Delta t$. It's often more useful as a measure of what's happening to the population than the failure rate.

For example, if you took 1000 people and looked at their "failure rate," you would simply calculate the number of failures in a given time – e.g. one year. As the population ages, the number of people remaining decreases, so the failure rate for age 70 will be lower than the failure rate for age 40. However, if you consider the probability of survival (R(t) to age 70, that is lower than that for age 40. The hazard rate for people of age 70 is then the failure rate modified by the probability that they survive to that age (e.g. hazard rate at 70 will be higher than the failure rate at the same age).

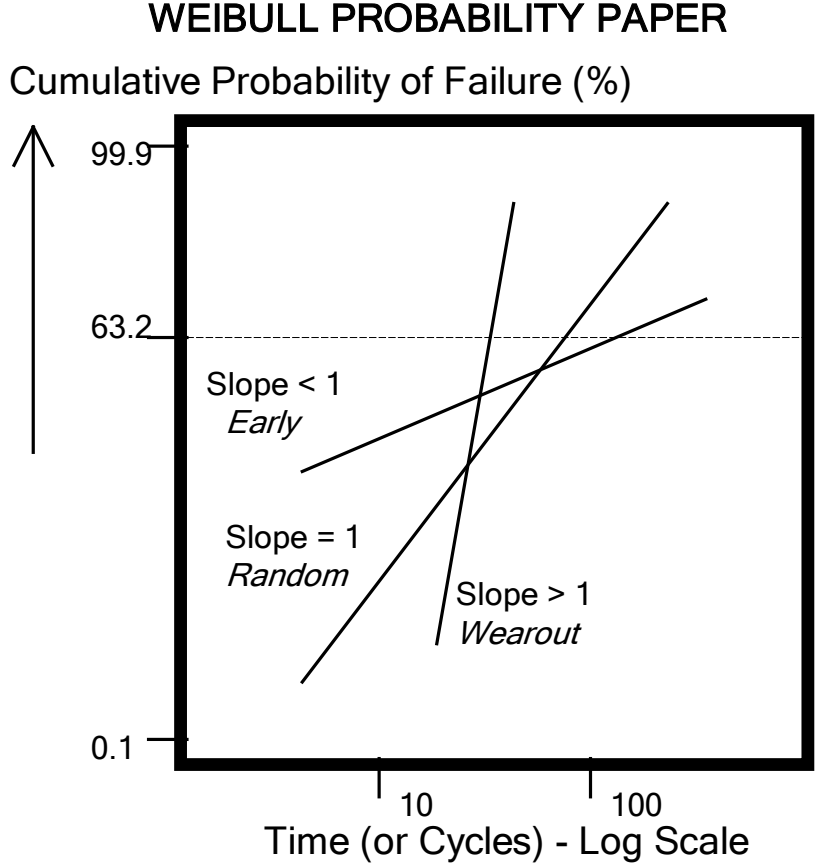
Each of the probability models described above has an associated *hazard function.* These functions fall into three classes, *increasing*, *constant* and *decreasing* hazards (over time). For instance, the exponential distribution's hazard function is constant and $h(t) = \lambda$, the exponential distribution's parameter.



The next several pages describe the characteristics of each type of hazard rate and our improvement options. The "Bathtub Curve" shown to the right is an idealized model of a system's hazard rate over its entire "life." Systems will often exhibit an early failure period, with a relatively high number of failures. As these "bugs" are worked out of the system, it settles into a period of low failure rate, i.e. its useful life. As equipment begins to wear, the hazard rate increases; this is the "end of life" for the system.

Ideally, a system should complete the early failure (sometimes called *burn-in* period at the factory (or during start-up testing) and be repaired, replaced or refurbished before the wear-out period begins. Then, its entire life will be spent in the flat portion of the bathtub curve. For many mechanical devices, the hazard rate is practically zero in this portion of the curve.

The Weibull probability plot can help us identify which type of failure the item is experiencing. The *slope* of the Weibull line is our estimate of the Weibull's β parameter. The value of β corresponds to different failure types:



Decreasing Hazard Rate

This failure pattern is known as *infant mortality* or *burn-in.* Apparently, the item is improving in reliability with the passage of time. This seems to violate the entropy principle, but there is usually a more logical explanation.

Suppose an initial population of items is placed in service that do not all have the same level of quality. We could consider some of the items to have a failure rate, $\lambda 1$, and others to have failure rate, $\lambda 2$, where $\lambda 2$ is greater than $\lambda 1$. As opportunities for failure to occur, more of those items having $\lambda 2$ fail than those having $\lambda 1$. As failures accumulate, the surviving population becomes "richer" in $\lambda 1$ items, "poorer" in $\lambda 2$ items. An observer who assumes the population to be uniform may conclude that the hazard rate is decreasing and the devices are somehow improving with the passage of time. Here, an apparently decreasing hazard rate is evidence of uneven item quality.

Complex electronic components (RAM and ROM memory chips, central processing units) have been observed to display an apparently decreasing hazard rate.

There is some evidence that a complex system will display a decreasing hazard rate. At one utility, the number of forced shutdowns of nuclear units was plotted as a function of time after their refueling outages. The plot showed a high number of shutdowns in the first few months after the outage, this tended to decrease as the "bugs" were worked out. This may be evidence of failure modes introduced during the maintenance activities that occur during the outage, or from inadequately tested design changes made to the plant during the outage.

How to Detect: The time to or between failures will be described by Weibull distribution. In the latter case, the Weibull's shape parameter, β, will take on a value less than 1.

Actions to Take: For an item which displays a decreasing hazard, decreasing the variation in the item's quality (during manufacturing), debugging or operating the items for a period of time prior to their installation or shipping (i.e. burn-in) are options. For example, assembled computers are generally turned on and operated for a few hours before they are packed.

Constant Hazard Rate

Experience indicates that some simple electronic devices (resistors, transistors, potentiometers) and complex, repairable systems often exhibit a constant hazard rate. For the "simple" devices, failure often occurs because of a sudden stress, greater than the strength of the device. Since these stresses may occur "randomly" throughout the life of the device, the hazard rate for the population of devices is constant.

On the other hand, a complex system is composed of many parts, having different ages, strengths and failure distributions, and is repaired when it fails. Here, the various failure modes "combine" to display an essentially constant hazard rate. The system exhibits the "memory less" property of the exponential distribution, that is, it will be as reliable for the next hour of operation as when it was initially placed in service.

How to Detect: The time to or between failures will be best described by an exponential distribution or Weibull distribution. In the latter case, the Weibull's shape parameter, β, will take on a value close to 1.

Actions to Take: For the simple device with a constant hazard rate, the strength of the item must be improved if reliability is inadequate. *De-rating* the device, improved cooling (so the device operates at lower temperatures), or redundancy are options to explore.

For the complex system, the answer is not so simple. Since many different failure modes contribute to the system's failure, each with their own pattern, a combination of design, operation and maintenance changes may be necessary. To improve the reliability of a Southeastern US' generating units, this strategy was found to be necessary.

Increasing Hazard Rate

For some items, the accumulation of stress (operating or environmental) causes them to exhibit an increasing hazard rate. Mechanical items will wear out and fail, with many different physical phenomena being possible culprits. Thermal or mechanical stresses, vibration, corrosion, erosion, fatigue, are among these phenomena.

Chemical devices eventually experience end-of-life effects analogous to wear out, as the chemical reactions upon which they depend go to completion. Batteries lose their ability to hold a charge, lubricating oils break down, and resin filters lose their potency.

How to Detect: The time to or between failures will be described by a normal or lognormal distribution or Weibull distribution. In the latter case, the Weibull's shape parameter, β, will take on a value exceeding 1.

Actions to Take: If the reliability analysis reveals a failure probability that is too high for the mission time, redesign is necessary. Preventive maintenance (either time-based or condition monitoring) is also a very effective method of managing and improving the reliability of the device. Here, reliability is improved by replacing or refurbishing the item before the wear out phenomena results in high failure probabilities.

Sometimes a trade-off is necessary. Preventive maintenance can be a very effective reliability management tool, but it results in increased costs. A lifecycle cost analysis can be helpful in deciding whether to improve the item's design or implement preventive maintenance.

15.4.5 Quantifying Maintainability Characteristics

The issue of maintainability only starts with its quantification. Although *maintainability* was defined as a probability in Unit 15.1, it is rarely quantified as such. We will first show how maintainability is broken down, since the designer influences certain factors that affect maintainability and the operator/maintainer influences others. We will also describe some of the common measures used for maintainability, such as Mean Time to Repair, Mean Down Time and Maximum Repair Time.

The Elements of Maintainability

Let's consider a simple maintenance action, like changing a light bulb in a desk lamp (no, this is not leading to a joke about how many reliability engineers does it take!). We first need to *detect* the lamp has failed. Then we need to *diagnose* that the light bulb is the problem (we've checked that the lamp is plugged in!).

We will go find a new light bulb, one may be in a drawer, or we may have to go to a store and buy one. Then we will *repair* the lamp by removing the old bulb and screwing in the new one. We then turn on the lamp to *test* our repair efforts. We could measure the time it takes us to complete the entire repair as well as the individual activities.

As we have mentioned in several places, these times are only the tip of the iceberg when it comes to maintainability. Two key questions arise: 1) **What** are the important factors that impact these times and 2) **Who** is responsible for managing these factors?

The design of the lamp (or any system) plays a very important role in assisting the detection, diagnosis, repair, and testing activities. The author has a desk lamp with two tubular bulbs. The bulb is one and *one* quarter inches in diameter; the opening in the bulb holder is one and *three* quarter inches. The bulb is very hard to install, by design.

So, the *design* organization has accountability for maintainability. Our decision, though, to stock spare light bulbs in the kitchen drawer may be *influenced* by the designer (i.e. through a recommended spare parts list), but the accountability clearly lies in the *maintenance* organization implementing these recommendations. Responsibility for the skills and training of the maintenance forces also lies in this organization.

Suppose our lamp had multiple bulbs and somebody was reading from its light. That somebody may not want us to turn the lamp off until they have finished reading. We may need an *operations* clearance to begin the maintenance activities. This would influence time to repair.

Finally (the example is starting to stretch a little thin!), suppose we must complete paperwork (OK, we charge the lamp at the store to a credit card), or the lamp fails during our lunch break. Here, there is an *administrative* component to the maintenance activity.

These elements are generally grouped into three major areas:

Active Repair Time: The time spent detecting, diagnosing, actively working on the repair and testing the effects of the repair. The design and maintenance organizations most strongly influence this component.

Logistics Time: The time spent obtaining or waiting for parts that are needed to repair the item. The maintenance organization most strongly influences this component.

Administrative Time: The time spent doing paperwork, obtaining clearances, breaks, lunch, shift turnover, etc., generally any time not included in active repair or logistics time (the "mosey" factor fits here!).

Mean Time To Repair (MTTR)

The Mean Time To Repair (MTTR) is a useful measure of maintainability. MTTR is simply the average time to bring a system from a failed state to an operational state. The MTTR is calculated as follows:

$$MTTR = \frac{\sum t_i}{n}$$

$$where:$$

$$t_i \text{ are the active repair times, and}$$

$$n \text{ is the number of repairs}$$

The MTTR only includes the active repair times, and not logistic or administrative time. The reason for this is that MTTR is usually specified in contractual documents.

We can specify that equipment or a system be designed so that the active repair time averages a certain time, and tests can be conducted to demonstrate this specification. The equipment vendor, on the other hand, is not held responsible for the logistics and administrative time components of maintainability.

Mean Down Time (MDT)

The Mean Down Time (MDT) is the average time a system or equipment remains in the failed state. The MDT includes the MTTR, but also logistic and administrative time components. From the perspective of the organization operating and maintaining the system, and from the customer's perspective, the MDT is the important maintainability characteristic.

Maximum Repair Time (MRT)

We are not only concerned with the *average* time to repair, but also in the dispersion of these times. The details will not be repeated here, but the process of calculating the variance or standard deviation, fitting a probability model and describing uncertainty in the form of a confidence interval on the MTTR or MDT is the same as described in Unit 9.1.

A useful measure of the dispersion in maintenance times is the Maximum Repair Time or MRT. From the histogram or probability plot of the repair times, we look for the time at which 95 or 99 percent of the repairs are completed. Note that this MRT bounds the *individual* repair times, not a confidence bound on the average or mean time to repair.

There are several other maintainability measures in use. Some, like the *Mean Maintenance Man-hours* (total number of hours needed to keep a system operating), *Mean Maintenance Hours/Operating Hours* (average number of maintenance hours required to support one hour of system operation) and *Preventive Maintenance* (total number of hours of preventive maintenance required per year) not only measure the maintainability of the system, but also the cost of corrective and/or preventive maintenance.

Example: Repair times for tube leaks were found normally distributed, with an estimated mean of 70 minutes, and a standard deviation of 10 minutes. Find the MDT and the 95% Maximum Repair Time.

Answers:

The MDT is simply the mean of the repair times, or 70 minutes.

The 95% MRT is calculated as follows:

$$F(t \le T; 70\text{min}, 10\text{min}) = 0.95$$

$$\text{Forming the standard normal deviate :}$$

$$F\left(\frac{t-70}{10} \le 1.64\right) = 0.95$$

$$\text{The } 1.64 \text{ is the z - value for a cumulative probability of } 0.95,$$

$$\text{Solving for t :}$$

$$F(t \le 70 + 16.4) = 0.95$$

$$F(t \le 86.4\text{min}.) = 0.95$$

For the tube oil leaks then, 95% of the outages will be repaired within about 86 minutes.

15.4.6 Quantifying Availability Characteristics

Availability is a measure of all the time that a component is available to perform its required function if needed. Availability is a rather broad concept, encompassing as it does both the *reliability* of an item as well as its *maintainability.* The concept and measure of Availability is applied mainly to *repairable* devices. The Availability of a *non*-repairable device is simply equal to its reliability.

Availability Measures

There are many and various ways of measuring availability. However, each of these tends to be a variation on the basic theme of:

$$Availability = A = \frac{Uptime}{TotalTime} = \frac{Uptime}{Uptime + Downtime + Freetime}$$

For example, in electric utilities, each power plant is measured by the *Equivalent Availability Factor (EAF),* composed of Forced Outage Factor, Maintenance Outage Factor, Planned Outage Factor and Derating Factor – when the unit is producing less than its rated power. On the distribution side of the utility, though, the measure of availability is *Customer Minutes Interrupted* – the average time a customer is without power each year.

In the basic equation, *Freetime* is the time that the system is not needed. For example, the utility customer is assumed to require power continuously – no free time is included in the *Customer Minutes Interrupted* measure.

For our purposes, though, we will relate availability to reliability and maintainability. If we know the MTBF and MDT of a component, then the *Average State Availability* is given by:

$$A = \frac{MTBF}{MTBF + MDT}$$

More complex availability measures, such as time-related availability functions will not be covered in this introductory text. The problem with availability measurement is that it always involves two random variables, one to account for the reliability, one to account for the maintainability. Functions of random variables become complicated very rapidly.

While the treatment of a single item's availability is not that difficult, we are usually interested in the availability of systems of components. Computer simulations, such as Markov modeling, are used to study availability problems at a system level.

Availability Improvement

Many equipment or system users are interested in availability, such as electric power availability or computer system availability. Since availability has two components, we must stratify the contributors to *un*reliability as well as understand the components of maintainability.

As usual, we will search for the improvements that will give us the biggest bang for our buck, but we should understand what we are trying to accomplish, improvement in *reliability* or improvement in *maintainability*. Customer knowledge will often help us focus on one or the other.

15.4.7 Applications

Let's return to the questions posed at the beginning of this Unit. How would we answer at least these problems, with the information contained in this and other units? Let's at least examine some suggested approaches to these problems:

- **What's the probability a Chiller Unit will start the next time it's needed?**
  Approach 1 - Estimate the reliability of the unit based on its past failure/success (go/no go) history (or of a family of similar units). Consider eliminating failures from the calculation for which corrective action has been taken (i.e. the failure types are not likely to recur).
  Approach 2 - Develop a reliability model (Reliability Block Diagram or Fault Tree) and predict the starting reliability as a function of individual component reliabilities. EPRI's UNIRAM program, CAFTA code or GO code can be used to build such a model.

- **How often do we expect an ice-skating cooling system to have a forced shutdown?**
  Approach 1 - Based on past data of the system (or a family of similar systems), calculate a MTBF (or its inverse, the *average failure rate,* λ).
  Approach 2 - Develop a reliability model and predict the failure rate of the system as a function of the individual, *critical* component reliabilities.

- **How many spare parts (and what kind) do we need to stock for these circuit breakers?**
  Approach 1 - If there is a history of maintenance on these breakers, develop an estimate of the average parts usage/per time unit (and the *dispersion*) for the frequently replaced parts. If a *min/max* spares program is used, base the *minimum* value on the criteria of not exhausting the supply during the time it takes to reorder the parts *and* considering the average usage rate *plus* 2 or 3 standard deviations.
  Approach 2 - Obtain the circuit breaker FMECA from the manufacturer (if they have one!) and stock spares for the highest criticality failure modes. Use the estimates of *failure frequency* as a basis for the stocking levels.

- **Will the copier make it through my job without breaking down?**
  Approach 1 - Develop a distribution of the times (or copying cycles) between failures for the copier (either histogram or probability paper method). Given that the copier is functioning at the start of the job, predict the reliability associated with your copying "mission" (i.e. number of copies).

- **How long will it take to restore service to a customer?**
  Approach 1 - If a history of repair times is available, then develop a distribution of these times. You can then report to the customer the *mean down time* (including *active, logistic* and *administrative* components) or, the *maximum repair time* (based on a 90 or 95% confidence bound).
  Approach 2 - For a new system, perform maintainability testing (either simulate failures or, during reliability testing, measure the maintainability of the actual failures). As this will generally give information regarding the *active* component of down time, simulate or estimate the *logistic* and *administrative* components of the down time.

- **Can we skip maintenance on this valve until the next outage?**
  Approach 1 - Identify the functions and significant failure modes of the valve. If failure/maintenance history is available, develop distributions of times to failure. Identify when the valve was last maintained. Perform a reliability prediction based on the total operating time (operating time up until this outage *plus* expected operating time until the next outage).

- **How many home air conditioner failures will consumers see this year?**
  Approach 1 - If history is available, develop a frequency distribution of failures per time interval. Report to management the *average* frequency as well as an *upper bound* on the frequency.
  Approach 2 - For a new air conditioner design, obtain the engineering drawings of the system. Develop a reliability model (Reliability Block Diagram or Fault Tree) and predict the system failure rate as a function of the component failure rates.

- **Is it economical to repair or replace this piece of equipment?**
  Approach 1 - Develop a cost model that includes these factors: a) Cost of repair, b) Cost of Replacement, c) Expected remaining life of the equipment (a failure distribution may be needed to support this factor). This model will generally be of the form that estimates the *risk* of making a decision.

15.4.8 Weibull Analysis

Weibull Analysis Process

Weibull Analysis has several basic steps, one of which is to develop the Weibull model described below.

Step 1 - CLARIFY THE QUESTION. Before we perform a Weibull Analysis, we must somehow understand that we *need* to use this tool. This is "jumping the gun" ahead to the applications of Weibull, but here are some typical questions that may lead us to the Weibull tool:

- How often does this fail?
- How many failures can we expect in the next few months?
- What are our chances of operating this item until the next outage?
- When should we do maintenance on this thing?
- Is it economical to repair or replace this item?
- Should we do maintenance or do we need to redesign this thing?

The common element in these questions is that we need to understand something about the *time to failure pattern* of the item in question. What is the average life of the item, what is the dispersion of the item's life? There are other ways of answering these questions, but, *if there is a history of the items' failures available,* then **Weibull Analysis** can be a viable option.

Step 2 - COLLECT THE DATA To do a Weibull Analysis, the absolute, bare minimum information needed is simply the **times to failure** of the items in question. These times may be in calendar hours, operating hours, cycles, or other measure that correlates to the stress placed on the item.

Failure can be defined by the needs of the analysis. Failure could mean the rupture of a boiler tube, or it could mean the loss of function of a relay, or it could mean the drift of an instrument out of calibration. There are other data that may be useful to support the Weibull Analysis. The *failure mode* of the item should be identified. Most items can fail by several different modes and these often exhibit different time patterns. The Weibull plot will be helpful in identifying *mixed* failure modes, pointing to the need to *stratify* the data. If the items are labeled with serial numbers, this data may help track failures back to a particular batch of items produced by the manufacturer.

Step 3 - BUILD THE WEIBULL MODEL This step involves fitting a Weibull distribution to the failure data. Two methods of building this model will be described below. The probability density function of the Weibull Distribution is

$$f(t:\beta,\eta,t_o)=\frac{\beta}{\eta}\left(\frac{t-t_o}{\eta}\right)^{\beta-1}\exp-\left(\frac{t-t_o}{\eta}\right)^{\beta}$$

The modeling procedure builds both a graphic display of the failure distribution and it provides us with estimates of the Weibull parameters β, η, and t_0).

Step 4 - INTERPRET THE MODEL Here, we learn what the model is telling us about the behavior of our item. Depending on the question we are asking, we may be interested in understanding the probability of failure at a given time in the item's life, we may try to predict the number of failures that will occur in the future, we may try to understand whether maintenance can help manage the item's reliability or whether redesign is necessary. One very important result of building our Weibull model is found in the estimates of the three Weibull distribution parameters. These provide us with "clues" about the failures we're experiencing:

BETA (β) - β is called the shape parameter. β is interpreted as a measure of the *uniformity* of the item whose times to failure are being modeled. The value of this parameter has been "translated" into three general areas:

1. **β < 1**: If the model reveals a beta in this range, the item is said to exhibit a "burn-in" or "infant mortality" failure pattern. This means that we are seeing a high early rate of failure. As the early failures occur, the remaining items fail a "lower" rate. This is typical of manufacturing problems or maintenance errors.

2. **β = 1**: At this value of beta, the Weibull collapses to the exponential distribution. As discussed above, we are seeing a constant rate of failure. This is the so-called "chance" or "random" failure pattern. Physically, this has been explained as the behavior of a system of components that are repaired after they fail, and by items who are subject to random stresses that exceed their strengths.

3. **β > 1**: Items that fall into this range are in the "wear out" pattern. There may be physical degradation or chemical breakdowns occurring that result in failure. Many mechanical devices and their associated failure modes have been observed to fall into this pattern. Boiler tubes, bearings, valve packing, lubricating oils, spark plugs and cable insulation have been observed to fail in wear out modes.

Recognizing these different patterns can help us understand how to manage the failures we are seeing. For instance, one type of valve failure was thought to be the result of difficulties during maintenance, since the valve was located in a hot, cramped, high-radiation area of a power plant. Careful root cause analysis, with major support from a Weibull analysis revealed that one aspect of the valve's design was inadequate to meet the valve's mission.

ETA (η) - η is called the scale parameter. It is related to the life we can expect from the item under study. When $\beta = 1$, the Weibull distribution collapses to the exponential. In this case, η is equal to the mean time to failure, or to the inverse of λ. η shares the same characteristic as the inverse of λ; at this point, thirty seven percent of the original population will have survived; sixty-three (actually, 63.2%) of the population has failed.

TEE-ZERO (t_0) - t_0 is called the location parameter. Under certain situations, the Weibull distribution will exhibit a curved lower "tail" when plotted on probability paper. Physically, this behavior can exist when a manufacturer has "screened" items; that is, operate the items for some time to identify and remove items that fail early in life. The t_0 parameter is then used to correct the Weibull distribution to account for this screening.

Step 5 - TAKE ACTION BASED ON THE ANSWER - This step is self-explanatory, but critical.

Weibull Modeling – Failures Only

1. Organize the failure information into a table such as below. The *Median Rank* column will be completed in Step 3:

Item Number (Tag, Serial)	Failure Time (Hours, Cycles)	Median Rank	Remarks (Failure Mode, other data)
1001	87		Short
1002	68		Short
1003	72		Short
1004	82		Short
1005	65		Short

2. Order the failures in ascending times to failure (known as *rank order)*.

Item Number (Tag, Serial)	Failure Time (Hours, Cycles)	Median Rank	Remarks (Failure Mode, other data)
1005	65		Short
1002	68		Short
1003	72		Short
1004	82		Short
1001	87		Short

3. Obtain the *Median Ranks* from Table 15.4.1. The number of failures in your data set is the ***Sample Size*** on this table:

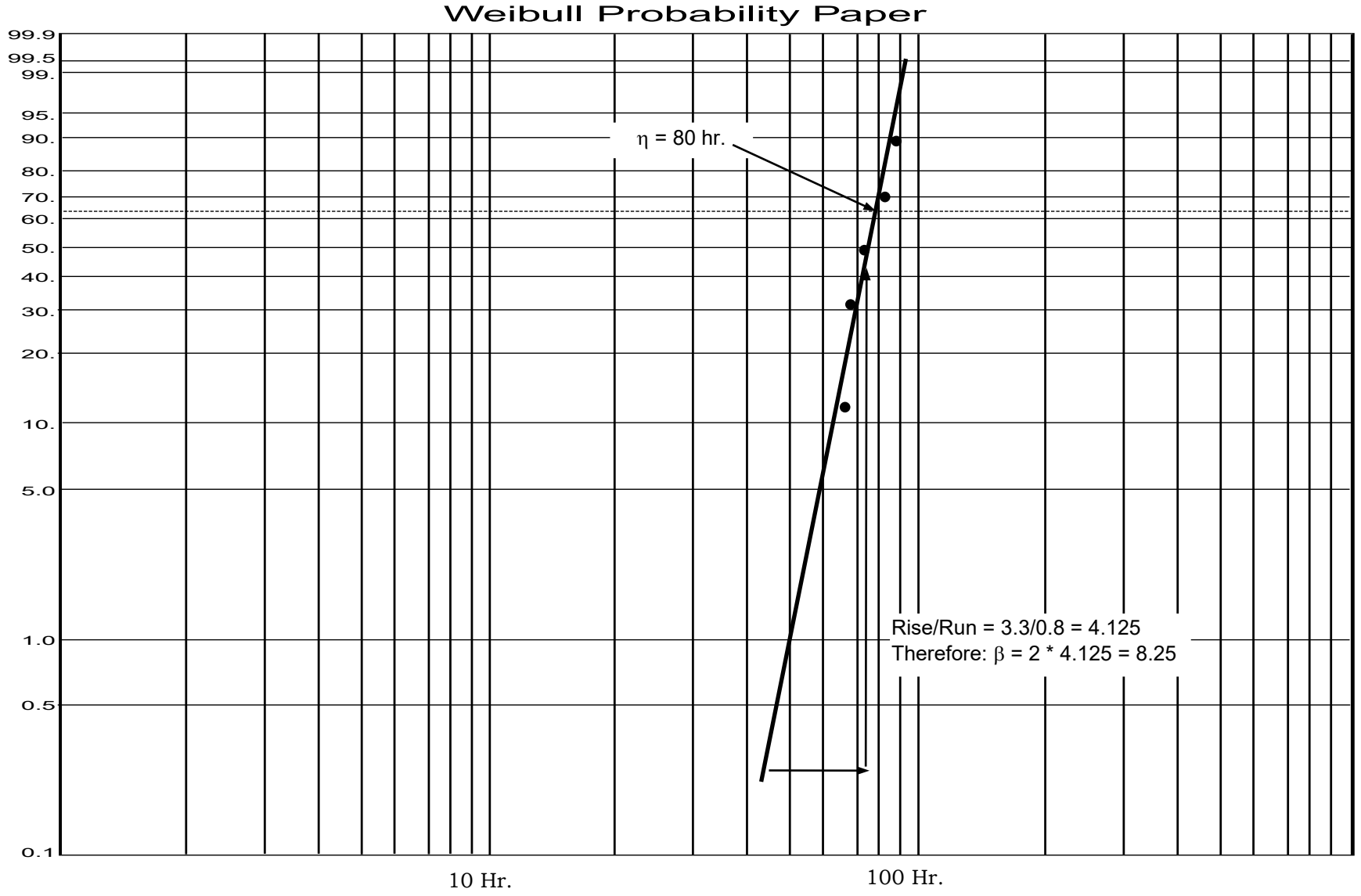
Item Number (Tag, Serial)	Failure Time (Hours, Cycles)	Median Rank	Remarks (Failure Mode, other data)
1005	65	12.9	Short
1002	68	31.3	Short
1003	72	50.0	Short
1004	82	68.6	Short
1001	87	87.0	Short

4. Plot the Failure Time/Median Rank points on Weibull Paper, “Eyeball" and draw a line that best fits the points on the Weibull Paper (see next page).

5. Determine the estimate of the item's characteristic life (η). This is the intersection of the Weibull data line and the pre-printed, dotted line at 63.2% (see next page).

6. Determine the estimate of the shape parameter (β). Use a ruler to obtain the rise and the run of the Weibull data line. For the Weibull paper in this section, β is two times the ratio of rise to run (see next page). When you use the Weibull paper included in the appendix, β will be equal to the rise/run ratio (it depends how the Weibull paper is constructed).

7. Label the plot. The Weibull model is now complete and ready for interpretation.


Weibull Probability Paper
99.9
99.5
99.
95.
90.
80.
70.
60.
50.
40.
30.
20.
10.
5.0
1.0
0.5
0.1
η = 80 hr.
Rise/Run = 3.3/0.8 = 4.125
Therefore: β = 2 * 4.125 = 8.25
10 Hr.
100 Hr.

Weibull Modeling – Failures & Suspensions

1. Organize the failure information into a table such as below.

Item Number	Failure Time	Rank Increment	Adjusted Rank	Median Rank	Remarks (Failure Mode, other data)
2001	87				Short
2002	84				No Failure
2003	68				Short
2004	91				Open Circuit
2005	72				Short
2006	82				Short
2007	67				No Failure
2008	65				Short

2. Order the failures *and* suspensions in ascending times:

Item Number	Failure Time	Rank Increment	Adjusted Rank	Median Rank	Remarks (Failure Mode, other data)
2008	65				Short
2007	67				No Failure
2003	68				Short
2005	72				Short
2006	82				Short
2002	84				No Failure
2001	87				Short
2004	91				Open Circuit

These next two steps develop the Y-coordinates that we need to plot our failure points. The "suspended" items in our sample will not be plotted on the Weibull plot. However, we must "adjust" for the fact that these components have not failed in the mode we are studying.

3.a Calculate *Rank Increments* for the **failed items**. This is done for the *first* failure in a series and whenever a *suspension* interrupts the series of failures. Use the following formula:

$$Rank Increment = \frac{(N+1) - (\text{Previous Adjusted Rank})}{1 + \text{Number of Items Beyond Present Suspended Item}}$$

where :

N - Number of items in the sample (failures and suspensions)

Previous Adjusted Rank - Calculated in Step 3b, 0 if this is the first failure in the series

Number of Items Beyond Present Suspended Item - Simply count the number of items in the series beyond (but not including) the suspended item.

For example, the first **failed item** is Serial Number 2008 (failure time = 65). The *Rank Increment* is:

$$Rank\ Increment = \frac{((8+1)-0)}{(1+8)} = 1.0$$

Since there are no suspensions "before" this failure, the sample size is used here.

3b. The *Adjusted Rank* is the previous adjusted rank, plus the rank increment. For this first failed item, the previous adjusted rank is 0, so the adjusted rank is simply 0 + 1 = 1.0. Since this first failed item is the "trivial case," let's examine the second failure in the series, Serial Number 2003. At this point, our table would look like this:

Item Number	Failure Time	Rank Increment	Adjusted Rank	Median Rank	Remarks (Failure Mode, other data)
2008	65	1.0	1.0		Short
2007	67	-	-		No Failure
2003	68				Short
2005	72				Short
2006	82				Short
2002	84	-	-		No Failure
2001	87				Short
2004	91	-	-		Open Circuit

For S/N 2003, we would first calculate the *Rank Increment:*

$$Rank\ Increment = \frac{\left((8+1)-1.0\right)}{(1+6)} = 1.143$$

The *Adjusted Rank* is now 1.0 + 1.143 = 2.143 (previous rank plus rank increment).

For the next failures (S/N's 2005 and 2006) we use the same *rank increment* and simply add the previous rank to the rank increment to obtain the adjusted rank. The table now looks like this:

Item Number	Failure Time	Rank Increment	Adjusted Rank	Median Rank	Remarks (Failure Mode, other data)
2008	65	1.0	1.0		Short
2007	67	-	-		No Failure
2003	68	1.143	2.143		Short
2005	72	1.143	3.286		Short
2006	82	1.143	4.429		Short
2002	84	-	-		No Failure
2001	87				Short
2004	91	-	-		Open Circuit

The calculation of the last rank increment and adjusted rank is left as an exercise for the student.

4. Obtain the *Median Ranks*. Although there are several methods for doing this, *Benard's Formula* has been found to be a good approximation.

$$P_{0.5} = \frac{i-0.3}{N+0.4} \times 100\%$$

$P_{0.5}$ - Median Rank, i - Adjusted Rank, N - Sample Size

Examples of this calculation for the first and second failures are shown below"

S/N 2008: $$P_{0.5} = \frac{(1 - 0.3)}{(8 + 0.4)} x\ 100\% = 8.33\%$$

S/N 2003: $$P_{0.5} = \frac{(2.143 - 0.3)}{(8 + 0.4)} x\ 100\% = 21.94\%$$

The completed table is shown below:

Item Number	Failure Time	Rank Increment	Adjusted Rank (i)	Median Rank (%)	Remarks
2008	65	1.0	1.0	8.3	Short
2007	67	-	-		No Failure
2003	68	1.143	2.143	21.9	Short
2005	72	1.143	3.286	35.6	Short
2006	82	1.143	4.429	49.2	Short
2002	84	-	-		No Failure
2001	87	1.524	5.953	67.3	Short
2004	91	-	-		Open Circuit

4. Plot the Failure Time/Median Rank points on Weibull Paper, “Eyeball" and draw a line that best fits the points on the Weibull Paper (see page 41).

5. Determine the estimate of the item's characteristic life (η). This is the intersection of the Weibull data line and the pre-printed, dotted line at 63.2% (see next page).

6. Determine the estimate of the shape parameter (β). Use a ruler to obtain the rise and the run of the Weibull data line. For the Weibull paper in this section, β is two times the ratio of rise to run (see next page). When you use the Weibull paper included in the appendix, β will be equal to the rise/run ratio (it depends how the Weibull paper is constructed).

7. Label the plot. The Weibull model is now complete and ready for interpretation.

After performing this procedure, you can see why a computer package like WeibullSmith or MiniTab is handy to have around!

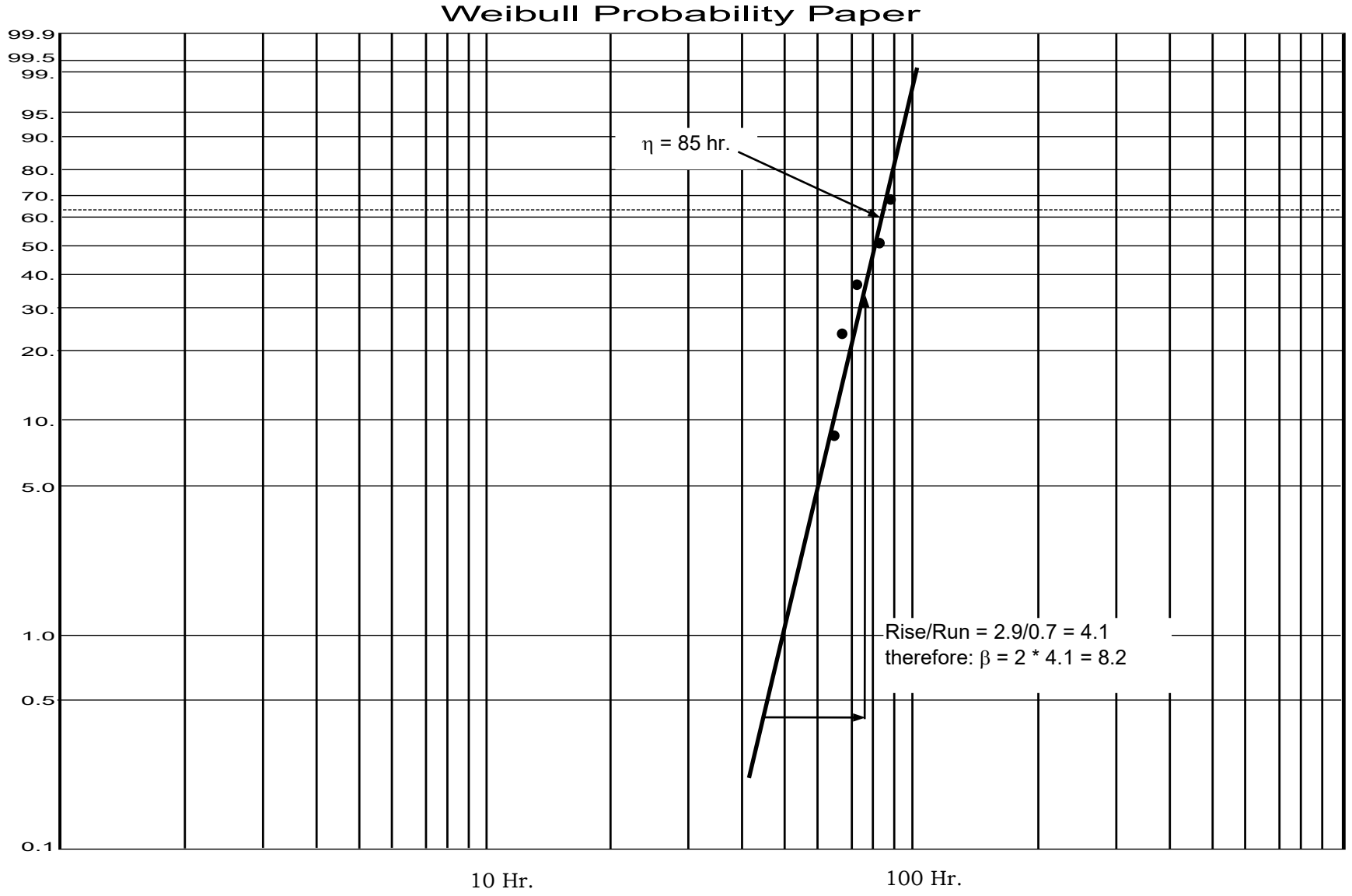

Weibull Probability Paper
99.9
99.5
99.
95.
90.
80.
70.
60.
50.
40.
30.
20.
10.
5.0
1.0
0.5
0.1
η = 85 hr.
Rise/Run = 2.9/0.7 = 4.1
therefore: β = 2 * 4.1 = 8.2
10 Hr.
100 Hr.

Table 15.4-1 - Median Ranks

	Sample Size																			
Rank Order	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	50.0	29.2	20.6	15.9	12.9	10.9	9.4	8.3	7.4	6.6	6.1	5.6	5.1	4.8	4.5	4.2	3.9	3.7	3.5	3.4
2		70.7	50.0	38.5	31.3	26.4	22.8	20.1	17.9	16.2	14.7	13.5	12.5	11.7	10.9	10.2	9.6	9.1	8.6	8.2
3			79.3	61.4	50.0	42.1	36.4	32.0	28.6	25.8	23.5	21.6	20.0	18.6	17.4	16.3	15.4	14.5	13.8	13.1
4				84.0	68.6	57.8	50.0	44.0	39.3	35.5	32.3	29.7	27.5	25.6	23.9	22.4	21.1	20.0	18.9	18.0
5					87.0	73.5	63.5	55.9	50.0	45.1	41.1	37.8	35.0	32.5	30.4	28.5	26.9	25.4	24.1	22.9
6						89.0	77.1	67.9	60.6	54.8	50.0	45.9	42.5	39.5	36.9	34.7	32.7	30.9	29.3	27.8
7							90.5	79.8	71.3	64.4	58.8	54.0	50.0	46.5	43.4	40.8	38.4	36.3	34.4	32.7
8								91.7	82.0	74.1	67.6	62.1	57.4	53.4	50.0	46.9	44.2	41.8	39.6	37.7
9									92.5	83.7	76.4	70.2	64.9	60.4	56.5	53.0	50.0	47.2	44.8	42.6
10										93.3	85.2	78.3	72.4	67.4	63.0	59.1	55.7	52.7	50.0	47.5
11											93.8	86.4	79.9	74.3	69.5	65.2	61.5	58.1	55.1	52.4
12												94.3	87.4	81.3	76.0	71.4	67.2	63.6	60.3	57.3
13													94.8	88.2	82.5	77.5	73.0	69.0	65.5	62.2
14														95.1	89.0	83.6	78.8	74.5	70.6	67.2
15															95.4	89.7	84.5	79.9	75.8	72.1
16																95.7	90.3	85.4	81.0	77.0
17																	96.0	90.8	86.1	81.9
18																		96.2	91.3	86.8
19																			96.4	91.7
20																				96.5

If larger sample sizes are encountered, Benard's formula may be used in a spreadsheet program to calculate the median ranks:

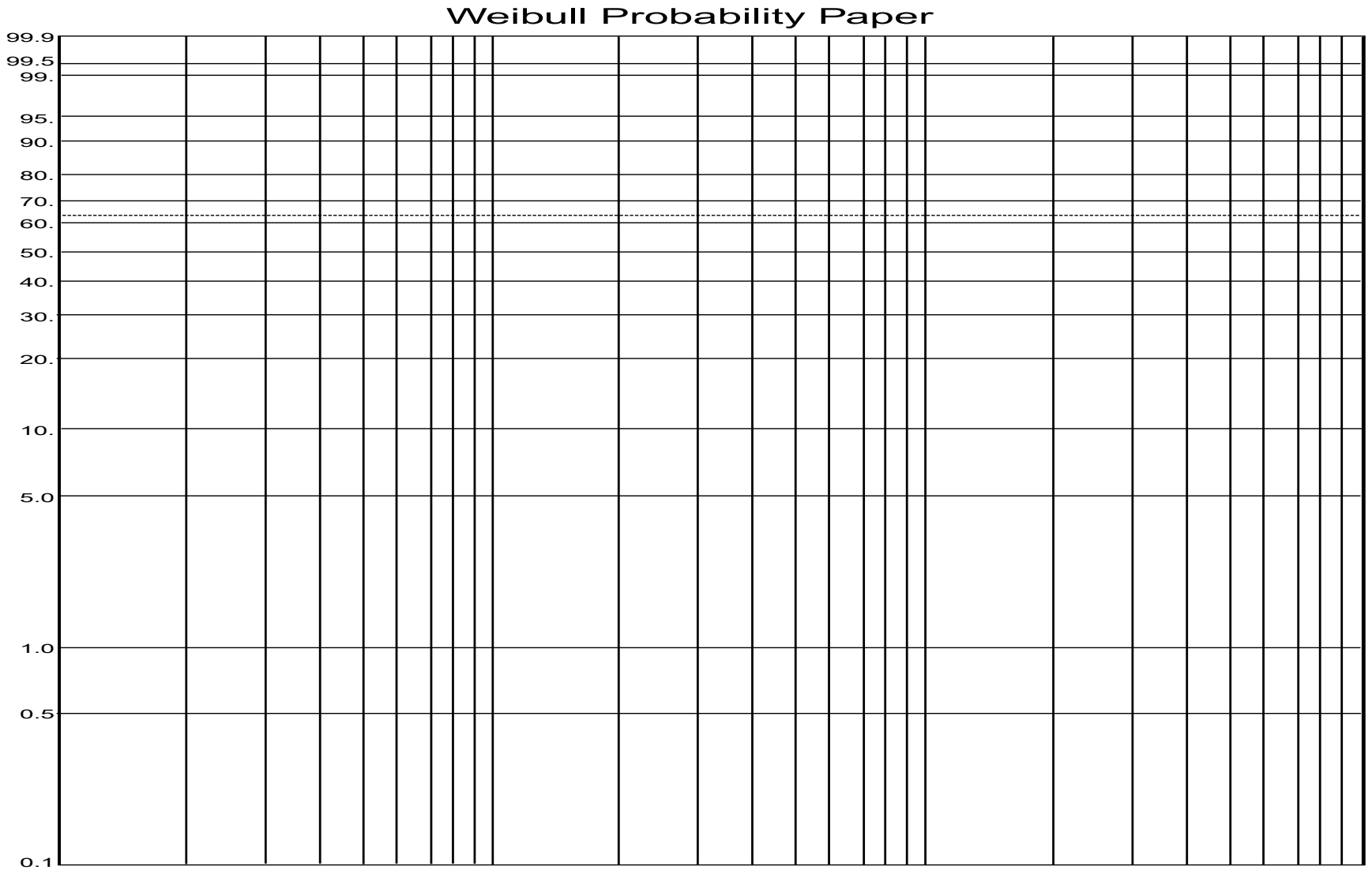
$$Median\,Rank = \frac{i - 0.3}{N + 0.4}$$

where:

i - *rank order*

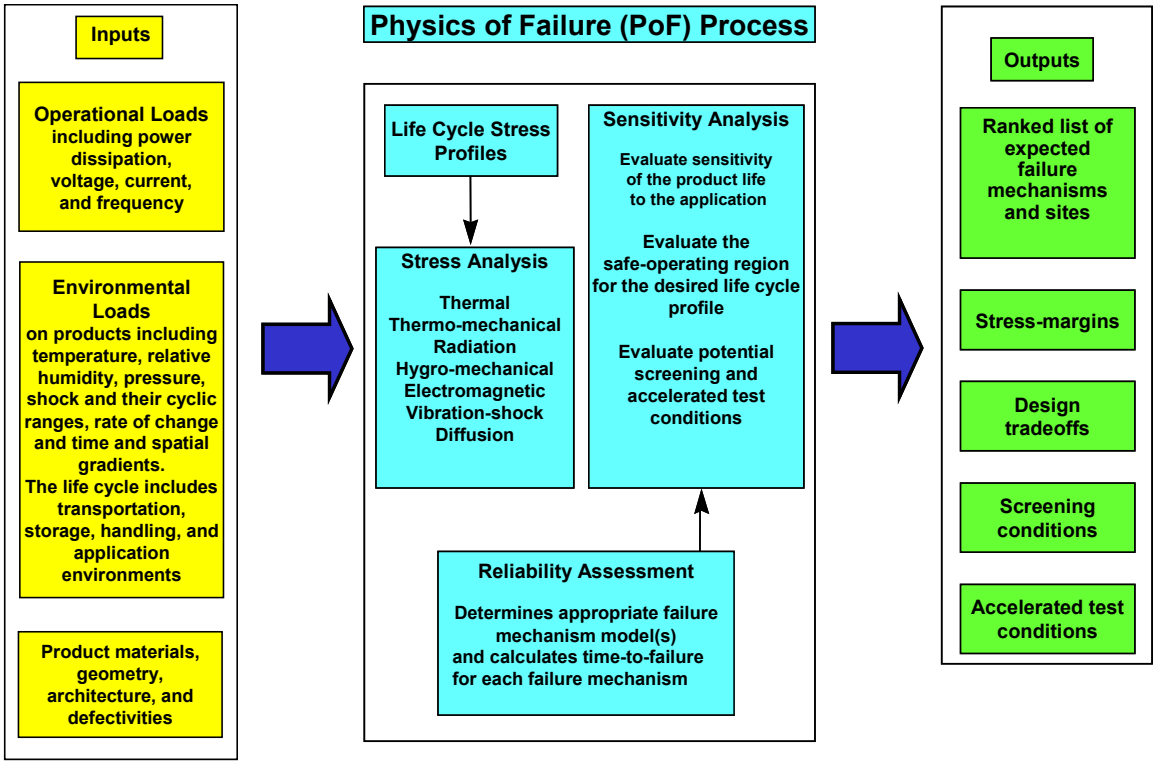
N - *sample size*

Note that this is "2 times" Weibull Paper – to find beta, multiply the slope of the line by 2.



15.4.9 Why The Traditional Reliability Prediction Models Do Not Work - Is There An Alternative?[1]

While it is generally believed that reliability prediction methods should be used to aid in product design and product development, the integrity and auditability of the traditional prediction methods have been found to be questionable; in that the models do not predict field failures, cannot be used for comparative purposes, and present misleading trends and relations. This paper presents a historical overview of reliability predictions for electronics, discusses the traditional reliability prediction approaches, and then presents an effective alternative that is becoming widely accepted.



1. Historical perspective to the traditional reliability prediction models

Stemming from a perceived need to place a figure of merit on a system's reliability during World War II the U.S. government procurement agencies sought standardization of requirement specifications and a prediction process. The view was that without standardization, each supplier could develop their own predictions based on their own data, and it would be difficult to evaluate system predictions against requirements based on components from different suppliers or to compare competitive designs for the same component or system.

Reliability prediction and assessment specifications can be traced to November 1956, with publication of the RCA release TR-1100, "Reliability Stress Analysis for Electronic Equipment," which presented models for computing rates of

[1] From a Paper written by Michael Pecht - CALCE Electronic Packaging Research Center, University of Maryland, College Park, MD 20742

component failures. This was followed by the publication of the "RADC Reliability Notebook" in October 1959, and a military reliability prediction handbook format known as MIL-HDBK-217.

In MIL-HDBK-217A, a single-point failure rate of 0.4 failures per million hours was listed for all monolithic integrated circuits, regardless of the environment, the application, the materials, the device power, the manufacturing processes, or the manufacturer. This single-valued failure rate was indicative of the fact that accuracy and science were of less concern than standardization.

The advent of more complex microelectronic devices continuously pushed the MIL-HDBK-217 beyond reasonable limits, as was seen for example in the inability of MIL-HDBK-217B models to address 64K or 256K RAM. In fact, when the RAM model was used for the 64K capability common at that time, the resulting computed mean time between failures was 13 seconds, many orders of magnitude from the MTBF experienced in real life. As a result of such incidents, new versions of MIL-HDBK-217 appeared around every seven years to "band-aid" the problems.

Today, the U.S. government and the military, as well as various U.S. and European manufacturers of electronic components, printed wiring and circuit boards, and electronic equipment and systems, still subscribe to the traditional reliability prediction techniques (e.g. MIL-HDBK-217 and progeny)[2] in some manner; although sometimes unknowingly. But with studies conducted by the National Institute of Standards and Technology (NIST), Bell Northern Research, the U.S. Army, Boeing, Honeywell, Delco and Ford Motor Co., it is now clear that the approach has been damaging to the industry and a change is needed.

2. Problems with the traditional approach to reliability prediction

Problems that arise with the traditional reliability prediction methods, and some of the reasons for these problems are described below.

(1) Up-to-date collection of the pertinent reliability data needed for the traditional reliability prediction approaches is a major undertaking, especially when manufacturers make yearly improvements. Most of the data used by the traditional models are out-of-date. For example, the connector models in MIL-HDBK-217 have not been up-dated for at least 10

[2] The traditional approach to predicting reliability is common to various international handbooks [MIL-HDBK-217 1991; TR-TSY-000332 1988; HRD5 1995; SRT 1985; CNET 1983; SN29500 1986]; all derived from some predecessor of MIL-HDBK-217.

years and were formulated based on data 20 years old.

Nevertheless, reliance on even a single outdated or poorly conceived reliability prediction approach can prove costly for systems design and development. For example, the use of military allocation documents (JIAWG), which utilizes the MIL-HDBK-217 approach upfront in the design process, initially led to design decisions maximizing the junction temperature in the F-22 advanced tactical fighter electronics to 60C and in the Comanche light helicopter to 65C. Boeing noted that:

> "The System Segment Specification normal cooling requirements were in place due to military electronic packaging reliability allocations and the backup temperature limits to provide stable electronic component performance. The validity of the junction temperature relationship to reliability is constantly in question and under attacks as it lacks solid foundational data."

For the Comanche, cooling temperatures as low as -40C at the electronics rails were at one time required to obtain the specified junction temperatures; even though the resulting temperature cycles were known to precipitate standing water as well as many unique failure mechanisms. Slight changes have been made in these programs when these problems surfaced, but scheduling costs cannot be recovered.

(2) In general, equipment removals and part failures are not equal. Often field-removed parts are re-tested as operational (called re-test OK, or fault-not-found, or could-not duplicate) and the true cause of "failure" is never determined. As the focus of reliability engineering has been on probabilistic assessment of field data, rather than on failure analysis, it has generally been perceived to be cheaper for a supplier to replace a failed subsystem (such as a circuit card) and ignore how the card failed.

(3) Many assembly failures are *not* component-related but due to an error in socketing, calibration or instrument reading or due to the improper interconnection of components during a higher-level assembly process. Today, reliability limiting items are much more likely to be in the system design (such as misapplication of a component, inadequate timing analysis, lack of transient control, stress-margins oversights), than in a manufacturing or design defect in the device.

(4) Failure of the component is *not* always due to a component-intrinsic mechanism but can be caused by: (I) an inadvertent overstress event after installation; (ii) latent damage during storage, handling or installation after shipment; (iii) improper assembly into a system; or (iv) choice of the wrong component for use in the system by either the installer or designer. Variable stress environments can also make a model inadequate in predicting field failures. For example, one Westinghouse fire control radar model has been applied in fighter aircraft, a bomber, and on the topmast of a ship. Each

application has its unique configuration, packaging, reliability and maintenance requirements.

(5) Electronics do not fail at a constant rate, as predicted by the models. The models were originally used to characterize device reliability because earlier data was tainted by equipment accidents, repair blunders, inadequate failure reporting, reporting of mixed age equipment, defective records of equipment operating times, mixed operational environmental conditions. The totality of these effects conspired to produce what appeared to be an approximately constant hazard rate. Further, earlier devices had several intrinsic failure mechanisms that manifested themselves as several sub-populations of infant mortality and wear out failures resulting in a constant failure rate. The above assumptions of constant failure rate do not hold true for present day devices.

(6) The reliability prediction models are based upon industry-average values of failure rates, which are neither vendor- nor device-specific. For example, failures may come from defects caused by uncontrolled fabrication methods, some of which were unknown and some of which were simply too expensive to control (i.e. the manufacturer took a yield loss rather than spend more money to control fabrication. In such cases, the failure was *not* representative of the field failures upon which the reliability prediction was based.

(7) The reliability prediction was based upon an inappropriate statistical model. For example, a failure in a lot of radio-frequency amplifiers was detected at Westinghouse in which the insulation of a wire was rubbed off against the package during thermal cycling. This resulted in an amplifier short. X-ray inspection of the amplifier during failure analysis confirmed this problem. The fact that a pattern failure (as opposed to a random failure) existed under the given conditions proved that the original MIL-HDBK-217 modeling assumptions were in error, and that either an improvement in design, improved quality, or inspection was required.

(8) The traditional reliability prediction approaches can produce what are likely to be highly variable assessments. As one example, the predicted reliability, using different prediction handbooks, for a memory board with 70 64k DRAMS in a "ground benign" environment at 40□C, varied from 700 FITS to 4,240,460 FITS. Overly optimistic predictions may prove fatal. Overly pessimistic predictions can increase the cost of a system (e.g., through excessive testing, or a redundancy requirement), or delay or even terminate deployment. Thus, these methods should not be used for preliminary assessments, baselining or initial design tradeoffs.

3. An alternative approach: physics-of-failure

In Japan, Taiwan, Singapore, Malaysia, the U.K. Ministry of Defense and many of the leading U.S. commercial electronics

companies, the traditional methods of reliability prediction have been abandoned. Instead, they use reliability assessment techniques based on the root-cause analysis of failure mechanism, failure modes and failures causing stresses. This approach, called physics-of-failure, has proven to be effective in the prevention, detection, and correction, of failures associated with design, manufacture and operation of a product.

The physics-of-failure (PoF) approach to electronics products is founded on the fact that fundamental mechanical, electrical, thermal, and chemical processes govern failure mechanisms. By understanding the possible failure mechanisms, potential problems in new and existing technologies can be identified and solved before they occur.

The PoF approach begins within the first stages of design (see Figure 1). A designer defines the product requirements, based on the customer's needs and the supplier's capabilities. These requirements can include the product's functional, physical, testability, maintainability, safety, and serviceability characteristics. At the same time, the service environment is identified, first broadly as aerospace, automotive, business office, storage, or the like, and then more specifically as a series of defined temperature, humidity, vibration, shock, or other conditions. The conditions are either measured or specified by the customer. From this information, the designer, usually with the aid of a computer, can model the thermal, mechanical, electrical, and electrochemical stresses acting on the product.

Next, stress analysis is combined with knowledge about the stress response of the chosen materials and structures to identify where failure might occur (failure sites), what form it might take (failure modes), and how it might take place (failure mechanisms). Failure is generally caused by one of the four following types of stresses: mechanical, electrical, thermal, or chemical, and it generally results either from the application of a single overstress, or by the accumulation of damage over time from lower-level stresses. Once the potential failure mechanisms have been identified, the specific failure mechanism model is employed.

The reliability assessment consists of calculating the time to failure for each potential failure mechanism, and then, using the principle that a chain is only as strong as its weakest link, choosing the dominant failure sites and mechanisms as those resulting in the least time to failure. The information from this assessment can be used to determine whether a product will survive for its intended application life, or it can be used to redesign a product for increased robustness against the dominant failure mechanisms. The physics-of-failure approach is also used to qualify design and manufacturing processes to ensure that the nominal design and manufacturing specifications meet or exceed reliability targets.

Computer software has been developed by organizations such as Phillips and the CALCE EPRC at University of

Maryland to conduct a physics-of-failure analysis at the component level. Numerous organizations have PoF software that is used at the circuit card level. These software tools make design, qualification planning and reliability assessment, manageable and timely.

4. Summary comments

The physics-of-failure approach has been used quite successfully for decades in the design of mechanical, civil, and aerospace structures. This approach is almost mandatory for buildings and bridges, because the sample size is usually one, affording little opportunity for testing the completed product, or for reliability growth. Instead, the product must work properly the first time, even though it often relies on unique materials and architectures placed in unique environments.

Today, the PoF approach is being demanded by (1) suppliers to measure how well they are doing and to determine what kind of reliability assurances they can give to a customer and (2) by customers to determine that the suppliers know what they are doing and that they are likely to deliver what is desired. In addition, both groups use PoF to assess and minimize risks. This knowledge is essential, because the supplier of a product that fails in the field loses the customers' confidence and often his repeat business, while the customer who buys a faulty product endangers his business and possibly the safety of his customers.

In terms of the U.S. military, the U.S. Army has discovered that the problems with the traditional reliability prediction techniques are enormous and have canceled the use of MIL-HDBK-217 in Army specifications. Instead they have developed Military Acquisition Handbook-179A that recommends best commercial practices, including physics-of-failure.

15.4.10 Summary

Reliability, Maintainability and *Availability* are characteristics that can and should be measured for equipment and systems. Both *average* and *dispersion* values of these characteristics are important, and we must adopt *probabilistic* models to describe them.

Different failure modes will require different models. Several probability distributions were presented that have been helpful in describing the failure modes. Clues to the underlying failure mechanism can often be gathered by understanding this time-based behavior.

15.5 Root Cause Analysis

Learning Objectives

- Be able to conduct a root cause analysis of equipment or process failures.

Unit Contents

- Root Cause Program and Elements
- Root Cause Analysis – Process
- Root Cause Analysis Reporting
- Troubleshooting & Failure Analysis
- Electrical/Electronic Failure Analysis
- Mechanical Failure Analysis Process
- Civil/Structural Failure Analysis Process
- Other Failure Diagnostic/Analysis Methods

15.5.1 Root Cause Program and Elements

What Is a Root Cause Analysis?

Generally, failures are much more expensive than is realized. Labor and material costs for repair are only the tip of the iceberg. When the costs of responding to the failure are added, together with lost time, lost revenue, direct and consequential damages, customer and regulatory impacts, and other indirect consequential costs, the total cost of a supposedly simple equipment failure may be incredibly high.

Despite requirements for failure corrective action (either externally imposed, such as by the Nuclear Regulatory Agency through 10CFR50, Appendix B, or internally imposed, such as by commitment to ISO-9000 or similar quality assurance systems), companies often do not understand the *costs* of failure and therefore do not devote energies to understanding the root causes of failure, beyond a superficial level.

We suspect that this behavior is an example of "*What we don't know won't hurt us"* thinking. If the true cost of failures was realized, this knowledge would create an overwhelming imperative to reduce the number of failures. Progressive companies have realized that an effective Root Cause Analysis Program is an essential component of their quality management systems.

The major purpose of a Root Cause Analysis program is simple: to reduce the **cost** of failures by preventing their *reoccurrence.* These failures may occur during design and development of a new product or service, or they may be operational failures. There are two parts to a root cause analysis: 1) the *troubleshooting* or *immediate remedy* phase, where we look for the component or part which failed and 2) the *recurrence prevention* phase where we look for the management system that "failed." The success of the root cause analysis lies in determining *both* these failures. In this unit, we'll provide you with both a process to follow in performing a root cause analysis as well as tools and methods to support your analysis efforts.

Root Cause Program Elements

To be effective, a root cause analysis program needs support by management and continuing direct participation by operating and maintenance groups, procurement groups and engineers from design engineering, quality control and reliability engineering. The essential features of a root cause analysis program are:

1. Responsibility for performing root cause analysis and specifying corrective actions clearly defined,
2. Adequate reporting of all failures and the conditions surrounding them (practical for “in-house” failures, more difficult for customer/field failures),
3. A uniform failure screening process applied to all failures,
4. A high level of engineering competence in root cause analysis,
5. Laboratory support for failure analysis as necessary,
6. Authority to require corrective actions to be implemented,
7. Adequate documentation of root cause analysis, documentation, follow-up of corrective actions and determination of corrective actions’ effectiveness.

Some comments follow on these elements:

Responsibilities

If the root cause analysis program has been centralized, upon receipt of a failure report, an engineer should be assigned to coordinate and perform the root cause analysis. The first decision should be whether the cause of the failure is evident. Depending on this decision, a root cause analysis team may or may not be necessary. If root cause analysis is decentralized, responsibility for root cause analysis must still be clarified and a lead person assigned to the work.

Failure Identification and Troubleshooting

Root cause analysis starts with the identification of the failure symptoms and relies on the *troubleshooting* process to identify the failed part(s). The difficulty of determining which part failed should not be understated. For complex systems or failures, a team of system “experts” may be needed to determine the cause of failure. This team may then proceed to determine the root cause of the failure and identify necessary corrective actions.

Failure Reporting & Information Flow

Root cause analysis must be based on an adequate failure documentation system that reports all failures promptly and in sufficient detail. A failure report should be completed and forwarded to the central point of contact, usually a reliability or quality or system engineering group. The failure, root cause and corrective action information flow must be well defined within the organization conducting root cause analysis. The elements that must be incorporated into this information flow include:

- *Failure Reporting* - responsibilities and forms for those personnel who detect and report the failure must be clarified,

- *Failure Analysis Reporting* - this covers the documents generated by the root cause analysis team,

- *Failure Trending* - responsibilities and forms for those personnel who receive the completed reports and perform trend analysis,

- *Corrective Action Assignments* - the tracking system for corrective actions committed to by engineering, operations, maintenance, etc.

With today's widespread use of networked computers, much of this reporting can and should be automated. This facilitates the root cause analysis process as well as the long-term retention of data for failure trending, to establish the effectiveness of corrective actions, and to provide input to design engineering for the next "generation" design.

Screening Process

An engineer reviews each failure report to determine if the root cause is apparent or directly discernible. If practical, the engineer should first visit the failure site to examine the failed item and talk with the person who initiated the failure report. The engineer should not concentrate on the failed item to the exclusion of its surroundings. The local failure may have been triggered by other events.

> *Example: Locomotive Axle Bearing Failures - Cracking of locomotive drive axle shrink-fit bearings was occurring. Root cause analysis was proceeding slowly, with little results, until a bearing in the initial stages of cracking was obtained (the author is proud to be the one who "dug" it out of an oil & grease filled dumpster!).*
>
> *The mechanical maintenance foreman observed that the crack pattern appeared to be the result of welding performed on the locomotive where the weld machine's ground was attached to the rails and not the locomotive. High current flowing through this resistive path (the bearing/axle interface) could produce the unusual burn pattern observed around the small, axially oriented cracks. This was later verified to be the root cause.*

Operating personnel should supply any relevant information or insights to the failure reporting system. They are closest to the failure and usually have pertinent information not available elsewhere. But operating personnel should not make the decision as to what the root cause is, or whether root cause analysis is necessary. Experience has shown that, while operating personnel may believe they understand the root causes of many failures, subsequent investigation often shows their judgment to be in error.

> *Example: At one plant, a particular pump had a history of failures. The pump failed so often that the plant had designed and installed a redundant "backup" pump. The operations personnel were convinced the design of the original pump was poor and rarely used it, relying on the "backup." During a period when the backup pump was unavailable, the original pump was used. High vibration caused the pump's vent plug to back off, resulting in a spill of several thousand gallons of contaminated water. The pump's shaft eventually sheared, initiating a low flow alarm in the system through which the leak was detected.*
>
> *During the root cause investigation that followed the incident, it was determined that the pump's oil level sight glass had been displaced several inches downward (most likely from someone using the sight glass as a "step"). When the sight glass indicated a "full" oil level, there was not enough oil in the sump to be picked up by the pump's internal slinger ring. No wonder the pump would burn up each time it was used!*

Guesses and easy conclusions should be avoided; the engineer should collect all the facts and then try to eliminate unnecessary information. He/she should rely on objective evidence to the maximum possible extent. If all of the following are evident, then a root cause analysis team is not required:

1. Mechanism of Failure,
2. Root Cause of Failure
3. Organizational Responsibility for the Failure, and
4. Necessary Corrective Action to Prevent Reoccurrence.

In such cases, the engineer contacts the logical action-item assignee and secures agreement to take the necessary corrective action by a specified date and to document the action. Then the engineer prepares a root cause analysis report, issues it and tracks corrective action in the same manner as if a team had been formed. Recurrence prevention will be noted by the absence of future failures of the type noted.

Budget or other corporate considerations[1] may affect the timing of the corrective actions' implementation. If a fleet of equipment is affected by the corrective actions, a planned implementation may try to predict the failure potential for equipment with different service lives (i.e. through Weibull analysis, if the failure is a wear-out type).

In-Depth Analysis

When the root cause of a failure is not evident, engineering analysis is required. If the number and frequency of failures exceeds the analysis resources available, the responsible engineer may have to rank failures in order of importance and analyze only the most important. The corporate system cannot achieve its long-term goals under those conditions, so management commitment to staff the failure analysis function adequately to handle the full workload is essential.

When a company is starting a root cause program, though, it may be difficult to staff a program to meet the backlog of failures. Under these conditions, it's better to "Pareto" the failures and only analyze the most important, rather than attempting to provide superficial coverage of all the failures.

An effective root cause program requires that analysis be undertaken and completed promptly after the failure event. Like criminal investigations, the "trail" of the root cause becomes cold as time passes and undocumented information is forgotten.

> *Example: Locomotive Failure Analysis "Startup" - A major Western US railroad began its root cause analysis effort in 1992. Failures to be analyzed were limited to those "produced" through the 60 and 90-month overhaul process at their Denver Heavy Overhaul Facility and that of a Kentucky-based contractor to the railroad. Even here, the failure backlog was excessive. Plant management decided to limit analysis to only those failures that were occurring within 48 hours of the locomotive's release from the overhaul shop.*

The actual root cause analysis is often conducted as a team activity.[2] If a centralized root cause organization is used, the responsible engineer convenes and chairs the Root Cause Analysis Team (RCAT). Specific membership varies from one analysis to another. Participants should include the design or system engineer(s), representatives from operations or maintenance, plus discipline-specific representatives - materials specialist, stress analyst, laboratory analyst, etc.

[1]For power plant hardware-related root causes, a planned outage may be required to implement design changes resulting from a root cause analysis.
[2]As part of an organization's quality improvement efforts, this is often a prime "early" use of teams. They can achieve two purposes: 1) give teams practice in problem-solving techniques and 2) reduce the backlog of failures that do not require sophisticated engineering analysis, since many failures fall into the "low-hanging fruit" category - easy to discover the root cause and to solve.

The team's charter is to examine the failure and attempt to understand what happened and why. The group is responsible for determining the mode and mechanism of failure, functional responsibility for the failure (design, quality, mishandling, shipping, accident, vendor, maintenance, operation, etc.) and the immediate, contributory and ultimate root causes (*see the lawn mower example, above*). This group also determines the corrective action(s) necessary and secures commitment from the responsible groups to implement these actions.

Corrective Actions

Corrective action is the key to effective failure reduction. Root cause responsibility may lie in the design, operations, or maintenance functions. These functions may lie inside the organization or be provided through vendors.

Responsibility for corrective action must be clarified. In some organizations, a central point of responsibility is established, or the corrective action status is reported through a cross-functional committee or at a level high enough to encompass all functions.

Completion of all action items emerging from the analysis (engineering change notice or design package, operating or maintenance procedure change, procurement specification revision, etc.) should be documented in the Root Cause Analysis Report.

The Root Cause Analysis Team may "close-out" their activities at this point, but the failure is not closed-out until the PDCA cycle is truly rotated. Failure trends should be monitored to determine if the corrective action has achieved its intended results of preventing or minimizing the frequency of failures.

Vendor-Supplied Equipment Failures

If a vendor's equipment has failed, the vendor should be informed at once and invited to participate on the team. However, failed items should not be returned to the vendor for analysis. That practice, which is sometimes used to save money or avoid committing in-house resources to root cause analysis, virtually guarantees that root causes will not be identified.

Vendors have vested interests in not admitting to design or quality defects affecting their products, often out of fear of legal liability or warranty claims. Moreover, the vendor often lacks the detailed application information, equipment history and surrounding circumstances needed to identify root causes.

Root Cause Data Requirements

We've often walked into companies interested in preventing failures and find that they are relying on non-existent or inadequate failure data collection systems. Frequently the company's "formal" failure data collection system is designed primarily for accounting the cost of spare parts utilized and maintenance man-hours expended.

Often the early activities to establish a root cause analysis program will center on developing an effective data collection system. While this is being developed, analysis teams have to develop creative means of determining the priority of failures, their frequency, causes, etc.

A company that establishes a failure reporting system must define:

1. Forms (input screens) to be used in reporting failures of hardware and software,
2. Information to be entered in the forms,
3. Approval requirements, report routing, processing and retention requirements,
4. Form requirements for collecting data generated by subsequent diagnostic and repair activities.

Hardware Failure Data

The minimum information content for hardware failures must include the following:

Identification of Failed Equipment - The failed equipment must be identified; drawing numbers or tag numbers are commonly used. If the equipment is serialized, its serial number must also be reported, so that its previous service, failure and maintenance history can be reviewed. This is also important if the failure analysis points to the potential for a batch problem in manufacturing.[3]

Operating Configuration - The system or plant configuration in which the failed item was operating at failure should be reported. This is especially important if the configuration is variable (full vs. partial power, operating vs. standby) or has been the subject of a recent configuration change, or if there is doubt that without such information the configuration will be completely evident to the failure analysis team.

[3]In many applications, both tag number and serial number can be important. For example, a pump motor may be associated with a given tag number - locating the pump within a system. If the motors are periodically removed, refurbished and then reinstalled, the serial number is necessary to determine pump tag numbers where the motor has seen service.

Operating Time - Accumulated operating time (hours, cycles, takeoffs/landings, etc.) is an essential data element for failure analysis. Often this is difficult to determine, especially when the root cause analysis program is in its infancy. The time reported may be cumulative operating time from installation, or from last scheduled maintenance or from the preceding failure. In some cases, for example, to perform Weibull analysis of times to failure, all the above data may be required.

Preferably, operating time should be taken from an elapsed time meter in the equipment itself. If there is none, it may be necessary to estimate cumulative time from knowledge of duty cycle and elapsed calendar time. Clearly, accurate records of installation, maintenance and previous failure dates must be available to personnel who initiate failure reports.

Downtime and Classification of Time Elements - Calendar downtime and active downtime (when the failure is being worked on) should be recorded. Separating operating and down times can be difficult in practice.

If the hardware's Mean Time between Failure (MTBF) is much longer than the Mean Time to Repair (MTTR), this may not be an important issue. For complex systems whose MTBF and MTTR are similar (i.e. MTBF of 100 hours, MTTR of 10 hours), though, failing to account for up and down times can make the operating time estimates and subsequent analysis difficult.

This issue is complicated by the various components of up and down time. For instance, equipment may be "up," but in standby, or partial load. Downtime components include active repair time, but also logistic and administrative delays, plus post-maintenance test/debugging.

> *Example: Nuclear power plant safety equipment is a challenge to classify, since much of its time is spent in a standby mode, "waiting" for the accident. Periodic tests of the equipment may be the only operating time accumulated by the equipment. Often these tests occur at reduced or throttled loads, or in system configurations that do not completely match accident conditions.*

Failure Data - The time and date of the failure (occurred and/or discovered), symptoms of failure, circumstances and environmental conditions must also be captured. The failure symptoms should be described in narrative form, coding into failure modes should occur as part of the failure analysis, not as part of failure reporting. Personnel who report failures

should be trained formally and feedback on the reports should be given frequently.[4] These latter are essential, yet often ignored aspects of a root cause analysis program.

Diagnostic and Repair Data - The data system must be designed to allow appending necessary diagnostic and repair data, which may be obtained days or sometimes weeks after the failure occurred/was observed.

> *Example: Printed Circuit Modules used in locomotive controls were "bad-ordered" (declared failed) by shop repair personnel. The modules were then sent to the locomotive vendor for repair. The component failures were never captured by the railroad's information system. Root cause analysis and improvement were virtually impossible under these circumstances.*
>
> *On the other hand, similar failed modules used in nuclear reactor protection circuits were diagnosed and repaired by the plant's Instrumentation and Control Maintenance Department. This allowed specific component failures to be trended and resulted in many root cause determinations and subsequent corrective actions. As an example, Weibull analysis of capacitor failures in Hagan plant protection modules led to initiating a 4-year replacement policy.*

Software Failures

Software does not break down or fail in the same sense as hardware. Software is correct or incorrect. If it is incorrect, all copies of it will be incorrect; therefore, redundancy does not improve software reliability in the conventional sense. However, software can be treated exactly like hardware for purposes of failure reporting, root cause analysis and corrective action.

The potential for software error is virtually unbounded. Even simple programs have so many paths, loops and decision points that exhaustive testing is not practical. Process plant control and protection software depends on numerical algorithms to solve linear or non-linear integral-differential equations of dynamics to filter redundant and error-contaminated multiple-sensor data, etc. These algorithms are often only approximations of physical laws, which in turn, approximate nature.

[4]Feedback takes two forms here: 1) to improve the quality of reports received and 2) to show personnel collecting the data that something is actually being done with the data, i.e. failure analyses and corrective actions taken.

In many cases, real-time computational speed constraints require the programmer to use simplified algorithms, which then become "approximations of approximations." While these algorithms may be valid over a specified range of inputs, as the extremes of the range are approached, the algorithm may break down and software errors occur.

Data that should be reported concerning software failures included error category, error source, the module and program affected, whether program termination was normal or abnormal (e.g. infinite loop, pointer reference out of bounds, program crash, lockup, etc.), severity classification and timing data (cumulative or run time at failure). Analysis information and correction data should be added to the failure report as soon as possible, similar to the hardware diagnostic and repair data.

Human Failures

In many industries, human error dominates the probability of system failure. Medical processes such as drug administration, diagnostic testing, disease diagnosis[5] and treatment are examples of systems whose success is almost totally dependent on human performance. Although the same principles for data collection can be applied to human error as to hardware/software failures, the "human element" must be considered carefully in designing a data collection program. Detection of human error is often difficult, relying in many cases on self-detection. When a co-worker detects an error, there is a reluctance to "tell on" the worker committing the error.

Human error rates are often under-reported, due to fear of punishment or legal liability. In fact, the Joint Commission on Accreditation of Hospital Organizations' (JCAHO) surveyors will often look at a hospital's error records and claim that they are too low and must be raised. Of course, they are referring to the under-reporting of errors and are not implying that the hospital should commit more errors. Human error analysis is often more difficult than corresponding hardware/software failure analysis. The Institute for Nuclear Plant Operations (INPO) has developed several toolkits for use by nuclear plant operators to discover the causes of errors.

> *Example: Barrier analysis is used to analyze how an error could occur when there are multiple checkpoints, inspections, holdpoints, etc. included in a plant evolution. The management installed "barriers" are first identified, sketched and then the analysis focuses on which barriers failed and tries to discover why.*

Error-proofing techniques (error-prevention and error-mitigation) have been developed and can be employed instead of adding inspection layers to a process.

[5]One Medical Director confided to the author that he believes up to 30% of diagnoses made by doctors are incorrect.

15.5.2 Root Cause Analysis - Process

The “boundaries” of the root cause analysis process start with the system failure event and end with confirming the results of the changes made to prevent the reoccurrence of the failure. The organization should adopt a standard method of root cause analysis, with the specific activities occurring within each step of the analysis customized for the root cause being addressed. This is especially helpful when a team is performing the root cause analysis, as it allows the team to talk the “same language” as they proceed through their work.

Cause determination should follow the “ask WHY five times” philosophy discussed earlier. Superficial analysis may result in immediate remedies but generally do not address the root causes that will lead to reoccurrence prevention. Management reviews of root cause team output should also serve to ensure that the “WHY” question was asked enough times. The organization should not assume that the root cause will be identified for every failure. Typically, industrial failure analysis programs fail to identify root causes in up to 25% of failures investigated.

Root Cause Analysis Steps

The DMAIIC improvement method described in Unit 2.2 can be easily adapted to support root cause analysis. Each step should be “customized” for the specific failure(s) under investigation and whether the analysis is being done on an “active” failure or for “historical” failures. The matrix on the following page shows how this process “maps” to the DMAIIC process and the following pages describe each step in more detail.

DMAIIC Tailored to Root Cause Analysis

DMAIIC Method	Root Cause Analysis Description
Define the Problem, Measure the Current Situation	*On Occurrence of Failure:* Define the initiating event/failure and its circumstances/ consequences. *Preventing "Historical" Failures:* Prioritize past failures to determine the "most important" to address currently. Gather information and data relating to the event/failure. This includes physical evidence, interviews, records, etc. & documents needed to support the root cause analysis (drawings, procedures, manuals, design basis documents, etc.). Clarify the statement of the problem, if necessary.
Analyze Causes	Troubleshoot the system failure and determine what parts/actions caused the failure. Perform a failure analysis to determine why the problem occurred. Analyze the current system and/or perform experiments to determine the causes. Determine the *management system* factors that "produced" the failure.
Identify Potential Improvements,	Based on the "why" understanding, take immediate remedies to restore the system to operability, and then develop practical methods of attacking the causes, predict their benefits, costs, side-effects. Do the countermeasures *prevent reoccurrence* of the problem; are they directed at the system that "produced" the failure? Which one(s) should be implemented? Experiments may be necessary here.
Implement Changes – the "Do and Check"	Implement the countermeasures; pilot them, if necessary. What are the effects of the countermeasures? Has recurrence of the failure been prevented? Did they achieve what was expected? What other effects did the countermeasures have? Good or bad?
Control the Improved Process	"Build" the countermeasures into plant processes - make them part of the PLAN (the Who, What, When, Where, Why and How). Conduct the necessary training, revise procedures, etc. Make sure action responsibilities are clear. Evaluate the analysis. What went well, what could have gone better? What problems remain to be addressed? What plans are needed to address these?

Define the Problem, Measure the Current Situation

PURPOSE:	*On Occurrence of Failure:* Define the initiating event/failure and its circumstances/consequences. *Preventing "Historical" Failures:* Prioritize past failures to determine the "most important" to address at this time. Gather information and data relating to the event(s)/ failure(s). This includes physical evidence, interviews, records, etc. & documents needed to support the root cause analysis (drawings, procedures, manuals, design basis documents, etc.). Clarify the statement of the problem, if necessary.
POSSIBLE METHODS:	*On Occurrence of Failure:* Develop a clear understanding of the failure that has occurred. Create/develop a problem statement (i.e. *During monthly surveillance testing (OP-81-013), Emergency Diesel Generator - 01 failed to start when demanded by the control room manual start switch).* *Preventing "Historical" Failures:* Determine the most important failures which have occurred and which have not had recurrence prevention measures developed. Create/develop a problem statement (i.e. *During the last 5 years, 15 failures of Intake Cooling Water Pump motors have occurred during plant operation.).* Identify who needs to work on this project, obtain their support and develop the initial root cause analysis plan and schedule for the effort. Gather information relevant to the failure(s): • *Operating Logs* • *Maintenance, Equipment History & Inspection Records* • *Procedures, Instructions & Vendor Manuals* • *Drawings and Specifications* • *Strip Chart & Trend Chart Recordings* • *Sequence of Events Recorders* • *Sample Analysis and Results* • *Interviews/Personnel Statements* • *Correspondence* • *Design Basis Information* • *Photographs/Sketches of Failure Site* • *Industry Bulletins* • *Plant Parameter Readings* Create a Timeline of the events leading to the failure. *Clarify the Problem Statement.* Often the initial problem statement may be very broad. A Pareto analysis of historical failures may help focus the root cause investigation.
TOOLS:	• Pareto Analysis • Project Planning Worksheet • Event Timeline

Analyze Causes

<table>
<tr><td>PURPOSE:</td><td colspan="2">This is the “why” or diagnostic step. Where the Gather and Organize Data step helped us gather “clues” about the failure, here we will understand a) the part(s) that failed, b) the physical cause of failure, and c) the management system that “produced” the failure.</td></tr>
<tr><td>METHODS:</td><td colspan="2">There are always two parts to this step:

Develop Hypotheses about why the problem occurs. These may be Material, Machine/Equipment, Method, People, Measurement and Environment factors.

Confirm or Refute the Hypotheses. Gather evidence to establish the “guilt” or “innocence” of the different factors. This may be done through analysis of physical failure evidence, or experiments performed that deliberately attempt to reproduce the failure conditions. The following three criteria must be met:

a) The failure event would not have occurred had the causes not been present,
b) The failure event will not reoccur due to the same causal factors if the causes are corrected or eliminated, and
c) Correction or elimination of the causes will prevent reoccurrence of similar events/conditions.

Troubleshooting/Failure Analysis steps include:
• Determine the Failure Sequence/Circumstances
• Develop a Troubleshooting Plan
• Identify the Failed Part
• Confirm Failure of the Part
• Develop a Failure Analysis Plan
• Analyze the Part’s Failure Causes
• Determine the Sources of these Causes
• Develop a Conclusion & Recommendations
• Generate the Failure Analysis Report

Understand the Statistical Behavior of Time to Failure. Understanding the distribution of times or cycles to failure can provide clues as to the type of failure experienced.

Determine the System Factors that “Produced” the Failure. The root cause analysis should usually result in a conclusion that points to the design, operation, maintenance, procurement, or other management system as the system factor(s) responsible for the failure.</td></tr>
<tr><td>TOOLS:</td><td>• Cause and Effect Analysis
• Fault Tree Analysis
• FMEA, PF/EMEA
• Barrier Analysis
• Change Analysis</td><td>• Pareto Analysis of Causes
• Histograms
• Scatter Diagram
• Weibull Analysis</td></tr>
</table>

Identify Improvements

PURPOSE:	Here, changes will be identified that impact the root causes discovered during *Analysis* and that we think will prevent failure recurrence. The changes should be evaluated for their benefits, costs and possible side effects. The changes must be "sold" and planned. Short-term countermeasures may be implemented until longer-term fixes can be accomplished.
POSSIBLE METHODS:	*Brainstorm* possible countermeasures to address the root causes of failure. Select one or more that have the highest likelihood (and lowest cost) of impacting the root causes. *Benchmark* "best" practices and select the aspects of these that address your situation. *Experiments* may be performed to determine the best "level" for the important factors. *Factor levels* could mean anything from pressure sensing circuit trip set point to maintenance frequency for packing replacement. *Redesign* - In some cases, either the existing component or equipment is not able to meet its mission or is obsolete and must be redesigned. Once the countermeasures have been selected, they must be "sold" to the stakeholders (operations, maintenance, management, etc.). Then, detailed planning and implementation follow. A pilot or demonstration effort may occur prior to "full-scale" implementation. Also, plan to collect data to measure the before and after performance.
TOOLS:	• Root Cause/Countermeasure Matrix • Benchmarking • Cost/Benefit Analysis • Design Process • Action Plan/Project Planning Worksheet

Implement Improvements – the "Do & Check" of PDCA

PURPOSE:	Here, the changes are made to the product or process. After the changes are made, what effect have they had on performance - have the failures been eliminated? Do we understand that the changes we made caused the change in performance?
METHODS:	*Collect and Analyze Performance Data* to determine if the change has had a measurable impact on the failure frequency. Has the failure mode(s) addressed by the countermeasures been eliminated or significantly reduced in frequency? Determine if the *results* (observed changes in performance) are due to the *effects* of the changes you made to the process (sometimes other variables may be acting on the process that are outside your control). Three outcomes are possible here: 1) the results are due to our changes and performance is as expected. Here, move to *Standardization* step. 2) the results are much less than expected. Here, go back and understand why. 3) the results are much better than expected. Here, too, go back and understand why.
TOOLS:	• Line Graphs, Run Charts, Control Charts • Pareto Analysis • Histograms • Weibull Analysis
NOTES:	Nuclear plant countermeasures may take some time a) before they are installed and b) before results are seen in the form of reduced failures.

Control the Improved Process

PURPOSE:	The changes may have been performed on a pilot basis, or under temporary procedures. If the changes improved system reliability, then we must ensure that they are repeated each time the system is operated or maintained. The changes must be built into the PLAN, training & education performed and responsibilities clarified. Monitoring tools may be put in place. By this step, initial failure that triggered the root cause analysis should be "solved." But there are additional problems, maybe even additional aspects of the original problem to address. What are the plans to continue improvement work? Also, take time to evaluate the root cause effort.
METHODS:	*Revise the appropriate "Standards"* in your organization to build the improvements into the work processes. "Standards" may be procedures, drawings, technical standards, specifications, etc. *Train and educate* staff in the new methods, materials, machines, etc. *Monitor performance* to ensure that the changes aren't *Teflon-coated,* i.e. that they don't "stick." *Consider Replicating the Improvements.* If the changes worked in one system or plant, are they applicable to other systems, other plants? *Review the Root Cause Effort* with the team. Celebrate and recognize what was accomplished. Determine what went well and what could have been done better. Look at the team's use of the root cause *method,* and *tools.* Also examine how the group worked together, how much time the effort took and resources it required. *Review progress made toward the performance targets/goals* with this root cause cycle. What problems remain, what should be done about these problems? Plan how to address these remaining problems or consider the next project to address.
TOOLS:	• Procedures, Standards • Training • Line Graph, Run Chart, Control Chart • Root Cause Analysis Review Form • Project Planning Worksheet
NOTES:	One of management's review responsibilities is to examine not only the *results* obtained by the project and team, but also the *process* used to achieve the results. One company executive used to spend most of his questioning time asking the teams about why they used this or that tool, how they learned the tool or method, and what management could do to support the teams better in the future.

15.5.3 Root Cause Analysis Reporting

Improvements made to prevent the reoccurrence of failures should be documented. There are four important reasons why:

1. If a team is working on the root cause analysis, they need to keep track of where they are in the effort; what progress has been made, what's left to do.

2. When presenting the root cause analysis work to management and/or colleagues, some common method of communication has been found to be helpful.

3. Improvements generally result in an increase in the company's *technology inventory.* We now know how to prevent this failure in the future because. . . . The root cause analysis report is an important way of capturing this increase in technology.

4. In some cases, the initial root cause analysis will not discover the cause of failure and repeat failures will occur. The initial root cause analysis report should be reviewed to determine where the analysis process went awry and a determination of how to improve the analysis process made.

The Root Cause Analysis Report

As a working document, the leader and members of the project can keep a binder with the following tabs:

1. *Problem Identification*
2. *Cause Analysis*
3. *Countermeasures*
4. *Implementation Plans*
5. *Results*
6. *Standardization*
7. *Replication*
8. *Next Steps*
9. *Team Meeting Minutes & Notes (including Improvement Project Plan)*

The team leader often takes responsibility for keeping the root cause analysis report up to date and for distributing copies of important documents, charts, graphs, etc. Today's networking & communication software offer the opportunity for the team to keep a "virtual report" without the need for paper copies.

For presentation purposes, the important aspects of the improvement story should be prepared in a "storybook" form. There's a lot of confusion about these storybooks. The essential purpose of the storybook is just that, to tell a story. Think of the customers of your story - generally, they will include management and your co-workers.

If you only had 10 - 15 minutes to tell what you've done, what would you put into your story? You would keep it simple, remembering that not everybody is as familiar with the details of the analysis as you are, you would stick to the main points and not divert into all the dead-end paths that the team might have followed, and you try to convince your listeners that you understand the failure, its causes and that your countermeasures are going to (or have) solve the problem.

We prefer very little text in root cause analysis reports. You can tell your analysis "story" almost completely with the graphs and charts prepared during the analysis work. Figure 15.5 - 1 shows a root cause analysis on one page, mainly communicated through pictures.

The Root Cause Analysis Storyboard

The Root Cause Analysis Storyboard is merely a large version of the report, suitable for hanging in the hallway or meeting room. Sometimes the Storyboard is used for presentation purposes, as a working document, it supplements the notebook.

One purpose of displaying the team's progress in public is to solicit ideas from other members of the department or organization. If, for instance, the team's Cause and Effect Diagram appears on the storyboard, others can note additional causes that the team may have missed. This helps give the rest of the department a chance to "participate" in the improvement work. Figure 15.5 - 2 is a possible format for a Root Cause Analysis Storyboard.

Figure 15.5 -1 FUEL OIL SYSTEM LEAK REDUCTION

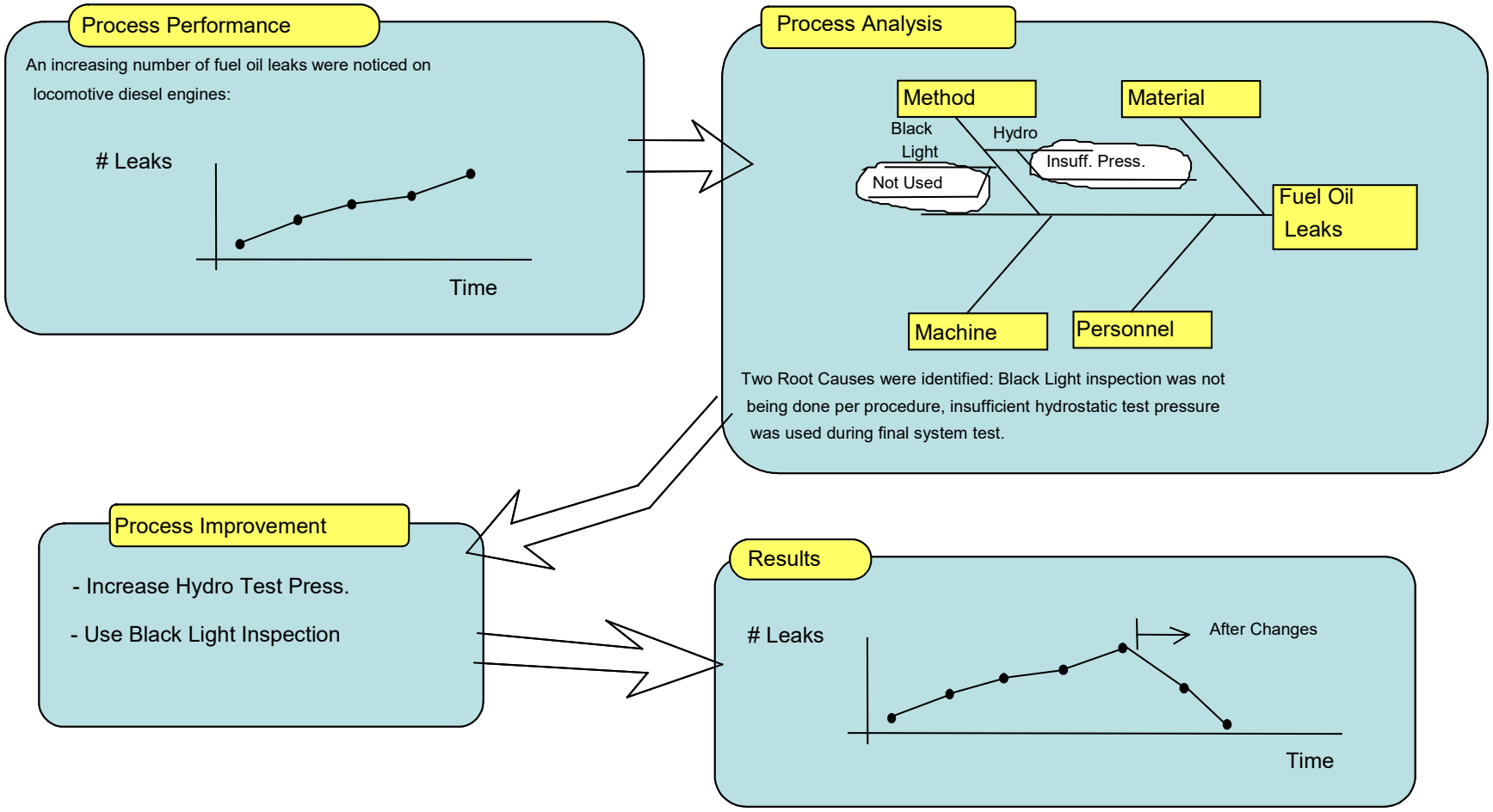


Figure 15.5 - 2 - Root Cause Analysis Storyboard

Suggested Dimensions: 36” High x 48” Wide

<table>
<tr><td colspan="4">IDENTIFY THE PROBLEM</td></tr>
<tr><td colspan="2">DEFINE THE EVENT/PRIORITIZE FAILURES</td><td colspan="2">GATHER AND ORGANIZE DATA</td></tr>
<tr><td colspan="3">ANALYSIS</td><td>COUNTERMEASURES</td></tr>
<tr><td>RESULTS</td><td>STANDARDIZATION</td><td>FUTURE PLANS</td><td>TEAM MEETINGS/NOTES</td></tr>
</table>

15.5.4 Troubleshooting & Failure Analysis

The aim of the Root Cause Analysis is to determine the *management system* that failed, so that recurrence measures may be taken. When the failure occurs, though, we may be a long way from determining which management system requires improvement.

We must follow a process that first helps us understand **why** the system failed. What component(s) failed, what were the causes of failure, and how did those causes come to be present in the system? This is the goal of **Troubleshooting/Failure Analysis.**

Troubleshooting can be defined as identifying the part or parts that failed, resulting in the observed system failure. *Failure Analysis* then attempts to discover *why* the part or parts failed. For example, troubleshooting a turbine control system may identify a failed pressure switch or flow controller. Failure analysis then proceeds to determine why the switch or controller failed.

The difference between these stages of root cause analysis lies mainly in the ability of the organization to take action. Once the failed parts have been identified, they can be replaced and the system returned to service. Failure analysis can then proceed as an “off-line” activity.

The *inputs* to the troubleshooting process include the system failure event, the circumstances surrounding the event and the history of the system. The *outputs* of the failure analysis will include the specific physical failure cause(s) and why they were present. These outputs will be used to determine possible corrective actions for the failure. If the root cause team concludes “the pipe had a crack in a longitudinal weld,” corrective action beyond the “immediate remedy” of welding the crack is not possible.

Troubleshooting/Failure Analysis may take many different paths, depending on the type and complexity of the system (mechanical, electrical/electronic, civil) being analyzed. The failure analysis may also proceed to different “depths,” depending often on corporate goals or plans or on the possible corrective actions available to the system owner.

> *Example: “Shallow-level” Failure Analysis - Valve leaks were a major source of nuclear plant unavailability. The troubleshooting history quickly pointed to packing wear-out as the component/failure mode causing the leaks. No additional failure analysis of the packing was pursued, since “new & improved” packing designs had been developed since the original packing was installed.*

Example: "Many-level" Failure Analysis - Boiler tube failures are a major source of fossil plant unavailability. There are several different types of boiler tube failures, occurring in different locations in the boiler (waterwall, superheater, economizer, etc.). The boiler is a system that is installed in many power plants, and is not currently in danger of technological obsolescence. Consequently, the Electric Power Research Institute has expended much effort into understanding the causes of tube failures and on advising utilities how to manage boilers to minimize the frequency of tube failures.

The next section will describe failure analysis processes for electrical/electronic, mechanical and civil/structural systems. The root cause analysis team should approach this discussion as a guideline, not a step-by-step procedure. It is the team's responsibility to select the most appropriate analysis approach and techniques. In this part of the root cause analysis, the team's goal is to find the root cause(s) with a minimum expenditure of time and cost.

15.5.5 Electrical/Electronic Failure Analysis

There are three general stages for electrical/electronic failure analysis. *Preliminary troubleshooting's* goal is to confirm the presence of a failure and isolate the failure to one section or circuit of the system. The next phase *isolates the failure to the part level.* At this point, we will know which part has failed and its failure mode. The final phase requires *failure analysis inside the part.* At the end of this phase, we will understand the physical cause of failure and may have clues that will lead us to understand why the physical cause occurred.

Preliminary Troubleshooting

Basic Troubleshooting Process

Effective troubleshooting for electronic circuits is a process of checking out **expectations**. The expectation may be general, such as "If pin 2 of the IC is shorted to pin 1, then pin 3 should go high for about one millisecond." The analyst may not be able to form detailed expectations about every part in the circuit, but neither should he/she be completely without expectations at the outset of the work.

Troubleshooting is primarily a process of checking the circuit's behavior against expectations, one by one, until one is found that doesn't match. At that point, there are only three possibilities:

- *That part of the circuit is malfunctioning, or*
- *The analyst's expectation was in error, or*
- *The measurement was in error.*

If the first is true, then the problem will be narrowed down to a small area relative to the initial circuit being investigated. If the second is true, the analyst will have learned something about circuit behavior that may be useful to her later. If the third is true, the analyst will likely be frustrated but will have learned something about measurement technique or will have discovered a faulty measurement device. From the author's experience in electrical circuit troubleshooting, it takes a certain mental discipline to perform this type of analysis. There's often a temptation to say, "I don't know if this is the right reading or not" and maybe move on to the next stage in the troubleshooting process. An opposite tendency is to "read what we need" and accept a reading without taking the time to see if the circuit and measuring equipment are set up correctly. Both tendencies lead to frustration for the analyst and, generally, difficulty finding the failed part. Here are

some general rules that aid the troubleshooting process. They apply to all three phases of the electrical/electronic failure analysis process:

Define the Problem - Verify the failure report and amplify it if necessary. If the failure report is general (i.e. *Unit doesn't work*), a more specific statement is necessary (i.e. *Locomotive SP6538 Dynamic Brakes are not generating required braking amps when placed in Test Mode*). The first benefit of defining the failure accurately is that it helps to concentrate your attention on the defective area. Many electronic circuits are too extensive to be taken in at one bite, so "divide and conquer" is a good rule. The second benefit is that it gives the failure analyst a clear definition to discuss the problem with plant personnel, a vendor, or others who may help point toward the cause.

Get the Schematic - Electrical analysis without a schematic is like working a crossword puzzle blindfolded. The analyst should get and use the schematic. Obtain a copy that can be marked up as the analysis proceeds. Also, make sure the schematic is the latest, "as-built" version. Many companies do not make drawing update a corporate priority and you may have to ask "Charlie" for a copy of the marked-up drawings he keeps in the lower right drawer of his shop desk.

Operating and service manuals may include block diagrams, circuit descriptions, logic diagrams, and troubleshooting procedures. These should be used in conjunction with the schematic diagram, not in place of it.

Work Systematically - There are several different approaches to troubleshooting - the key is to pick the one that fits the failure you're working to understand.

The *Fault Tree Approach* starts with loss of output or function and proceeds "backwards" through the circuit until the failed part is identified. Anybody who's opened a car repair manual is familiar with this approach. The symptoms are first identified - "Car won't start." Then the high-level functions that are required to start the car are tested - Gas, spark, starter, etc. This approach continues through the sub-functions until a failed component is identified. It may be helpful to actually draw the Fault Tree and use it as a diagnostic tool.

Another approach could be characterized as *See if it's Plugged in, First.* This approach starts with the inputs to the circuit - power supply, signals, etc. and works forward through the circuit until the failure is identified.

Both approaches break the circuit down into functional blocks, such as power supplies, amplifiers, oscillators, sensors, etc. The components will be functionally arranged on a "good" schematic; often dotted lines are drawn around these

functional blocks. This helps the troubleshooting effort, since the functional inputs and outputs can be identified. These can then be checked using the "divide and conquer" philosophy.

Board Swapping - Many electronic systems are assembled in modular fashion, with a "plug-in, pull-out" ability. Often, these modules will contain a distinct function. Swapping a known good module with a suspected failed one can often isolate the failure very quickly.

Keep a Record - In the middle of a two-day troubleshooting effort, it's tough to remember whether you have checked *Overcurrent Protection module KR12* or not, when there are sixteen other modules you've played with. A notebook divided in half - one column for expectations, one for measurements can be an effective troubleshooting aid. If parts are removed or wires/cables de-terminated, this notebook can help keep track of how the re-assembly should proceed.

Don't Redesign the Circuit - The analyst may find that when resistor R542, which appears on the schematic and in the circuit as 820 Ω, is replaced with a 560 Ω resistor, the circuit suddenly works. While this is interesting, it doesn't justify redesigning the circuit. If the circuit worked once with an 820 Ω resistor, it should work again with the same resistor. The analyst should look further, perhaps a leaky capacitor is present, or there's a transistor whose beta has changed, or there is corrosion on the board.

Measure Correctly - Often, the measurement instrument is not set up correctly. Some common measurement errors include:

- *Forgetting x10 scales*
- *Failing to check zero and calibrate controls*
- *Substituting an rms value for a peak-to-peak expectation*
- *Measuring line noise picked up on an open probe and calling it a signal*
- *Failing to notice that meter resistance or capacitance is loading the circuit*
- *Attempting to measure AC voltages (or other signals) above the max rated frequency of the instrument*

Keep it Simple but Consider Multiple Failures - There's two parts to this suggestion. First, the majority of circuit failures are due to single parts giving up the ghost. Remember this. Occasionally, though, a voltage spike or other stress may cause multiple part failures, or the combined effect of two degraded components may result in a failed circuit.

Example: An old sailboat once owned by the author had a "cranky" outboard engine. The engine could start with slightly fouled spark plugs, or a slightly discharged battery, but not with both components in degraded condition.

Second, when a part failure occurs due to some external stress, it may be the "weak link" in the chain. When the failed part is identified, it may be a good idea to connect a replacement, turn on the power and look for smoke or other signs of malfunction. Keep the circuit on for at least five minutes to see if the replacement fails or other part failures occur. If possible, let the circuit operate for a longer period, giving the other components that may have been weakened by the stress a "chance" to fail.

Sources of Expectations

Returning to the failure analysis process, when the analyst finds a discrepancy between the expectation and the circuit's performance, she must stop at that point until the conflict is resolved. But where do these expectations come from? The following apply to all three phases of the electrical/electronic failure analysis process:

Experience - When "working on the railroad," we'd sometimes run up against some mysterious failures that seemed impossible to track down. That's when we'd go see "Jake," or "Shorty," or "Smitty." They'd worked locomotives so long, that just from hearing the failure symptoms, they'd often pinpoint the exact component or module that failed. Or, they'd climb up in the cab, pull out their VOM, and find the trouble with a few continuity and voltage checks without regard to any schematic or manual. While this is impressive, we have to recognize that the first time *they* experienced the failure, it probably took some time for them to track it down. This kind of troubleshooting is very effective with mass-produced items, but typical industrial failure analysis is seldom so repetitive.

Visual/Auditory Checks - Visual and auditory checks lead directly to the trouble often enough to make this the first avenue of attack. Fuse elements can be inspected, and transformer hum can be telegraphed up a large screwdriver to see if power is reaching the supply. Resistors may have a black charred ring or may be puffed up and cracked in the center. Switch contacts may be bent or corroded. Wire terminations may be loose, corroded or broken. Wires may have cracked insulation, may be touching, or metal "bits and pieces" may be wedged between two wires or terminals. Printed circuit tracks may be broken from overstress or shorted by corrosion from water, battery electrolyte or other foreign material.

While these checks may help isolate the failed component, they are generally still symptoms of a deeper problem. Fuses, resistors and transformers seldom fail unless they are overstressed - by a leaky filter capacitor, perhaps. When it is

observed that a fuse has blown for no apparent reason, it is a good idea to measure the line current with the fuse out and an ammeter across the fuse holder terminals, to see if the new fuse is also being stressed near its limits. Similarly, the voltage across a newly replaced resistor should be measured and the power calculated (*Power = V^2 /R, 5 volts across a 60 ohm resistor equals (5 volts)^2 /60 ohms = 0.42Watts*) to see if it is necessary to look further for the real trouble cause.

> *Example: A house owned by the author had air conditioning "back-fitted" by the previous owner. Periodically, one of the fuses in the main panel would blow during A/C operation. The root cause analysis first identified that the fuses were very warm when the A/C unit was running, further investigation revealed that both the fuses and the power cable from the fuse box to the A/C unit were too small for the current drawn by the unit. The compressor's startup current was a significant stress on the fuses.*
>
> *The author replaced the fuse panel with a modern breaker panel and ran larger gauge wire to the A/C unit, solving the problem. This scenario is often seen on a larger scale in older facilities whose electrical loads have grown incrementally over time.*

General Expectations - Simple checks of circuits and components can be effective means of isolating the problem:

Test Points - Many industrial circuits are designed with test points. Troubleshooting procedures provide tests that can be run to determine if the circuit or function is operating correctly. External on-off or variable signals can be inserted to check basic circuit functionality, such as amplification, relay or switching transistor pickup or dropout.

Continuity Tracing - Many problems can be identified by checking out the simple expectation that a conductive path should have nearly zero resistance between its ends.[6] Similarly, if a conductor connects two points, they will measure the same voltage with respect to ground. The humble Volt-Ohm-Meter is the electrical troubleshooter's best friend. Some likely places to check for broken continuity include:

- *Two ends of a cable or wire,*
- *The actual pin of an IC and the printed circuit track leading to it,*
- *The two ends of a long, thin printed circuit track, and*
- *The fixed and moving contacts of a switch or relay.*

[6]Remember to turn power off when performing resistance continuity checks!

Shorts and Opens - In a series circuit, a *short* in the circuit will deprive the shorted elements of any voltage and throw added voltage to the remaining circuit elements. An *open* circuit deprives all of the circuit elements of voltage, since the entire source voltage appears across the break in the circuit. The VOM can detect these kinds of circuit problems.

Component Specific Expectations - Some simple checks can be applied to components in electrical circuits:

- Batteries - Voltage across terminals, electrolyte specific gravity.
- Switches and Relays - Zero resistance across contacts when closed, infinite when open.
- Relays - Contact pickup or dropout when signal to coil is applied.
- Transistors - Should have a forward base-emitter drop of no more than 1.1 volts, also, should show at least a few tens of millivolts from collector to emitter. Zero volts may indicate a shorted transistor.
- IC's - DC supply voltage should be present on their V_{cc} pins, specific IC's can be checked for pin-to-pin voltages.
- Silicon Power Diodes - Should drop no more than 1.1 volts in the forward direction.
- Power Resistors and Transistors - Should be moderately warm when under full signal.
- Pressure Test - Gentle pressure with a plastic probe on components, wire bundles and circuit boards should not cause abrupt changes in circuit operation.

Schematics - Although this was described above, it's worth repeating here. The schematic provides the most valuable and complete source of expectations about the operation of a circuit. Learn to read and interpret the schematic. Often, the diagram will include expected waveforms (for oscilloscope readings) and voltmeter readings at specified points on the circuit. Remember to think functionally, though. Fit these readings into the larger picture of how the circuit is supposed to work.

Failure Isolation to Part Level

When a malfunction has been isolated to a certain section of a circuit, it becomes necessary to identify the individual faulty component. Checking components once they have been removed from the circuit is often simple. The art of analysis at this stage is to locate the defective component without tearing the circuit down.

Continuity Testing/Opens and Shorts - The discussion *supra* on these failures applies to this phase of the failure analysis, too.

Analysis Logic - Often, it will be impossible to isolate a problem to a specific component by simply taking measurements on a circuit as it stands. Then, it becomes necessary to form some expectations of what should happen if the circuit is deliberately altered. This is almost always a sequential process. Some initial troubleshooting will lead perhaps to a range of possibilities regarding the failed components. The analyst may then study the schematic, looking for a point where a lifted lead or unsoldered connection will help narrow the possibilities. Again, “divide and conquer” should be the philosophy used here.

Intermittent Failures - These are among the most frustrating and challenging to analyze. It is not uncommon for a device to appear bad in the field but work perfectly on the bench. Procedures for isolating intermittent failures include:

Check it on-site - If possible, check intermittent failures at the site where they are occurring. Often, wiring or connector problems will appear by simply moving the suspected conductors. The author “suffers” with an intermittent failure to print from his computer that goes away when the parallel cable connector at the back of his PC is “wiggled.” Often, several months will go by without this failure’s occurrence.

Try to induce the failure - Testing a circuit board outside its normal cabinet environment may have two, opposite effects. If the bench environment is cooler than the cabinet, temperature related failures might not appear. A heat gun may help simulate the cabinet’s environment. On the other hand, forced circulation through a cabinet generally depends on the cabinet doors being closed. Operating the equipment with the doors open may cause it to run hotter, because the normal air velocity is lost.

One trick for isolating a component failure that is temperature dependent is to obtain an aerosol can of spray coolant. The circuit is warmed up until the intermittent problem appears. Then each component in the functional area is sprayed until one is found that causes the circuit to perform correctly. Varying the line voltage through use of an autotransformer may induce some intermittent failures. Other intermittent failures may be caused by interference from nearby electric motors or other sources of electromagnetic energy.

Power Quality Problems - The presence of electronics in virtually every industrial control, protection and computing application has given rise to a new class of failures collectively termed *power quality.* These include power surges, spikes, sags and harmonics. Power quality problems manifest themselves in several different ways. Fuses can blow, lights flicker, equipment shutdowns may occur, and computer displays can “shrink.”

Harmonics can burn out a motor or transformer by inflicting incremental damage that is not noticed until the first symptoms of burnout are noticed. They can overload neutral current conductors and bus bars, cause circuit breakers to trip thermally, interference on telecommunications lines (when the telecommunications cable is run next to power cabling and vibration/buzzing sounds from electrical panels.

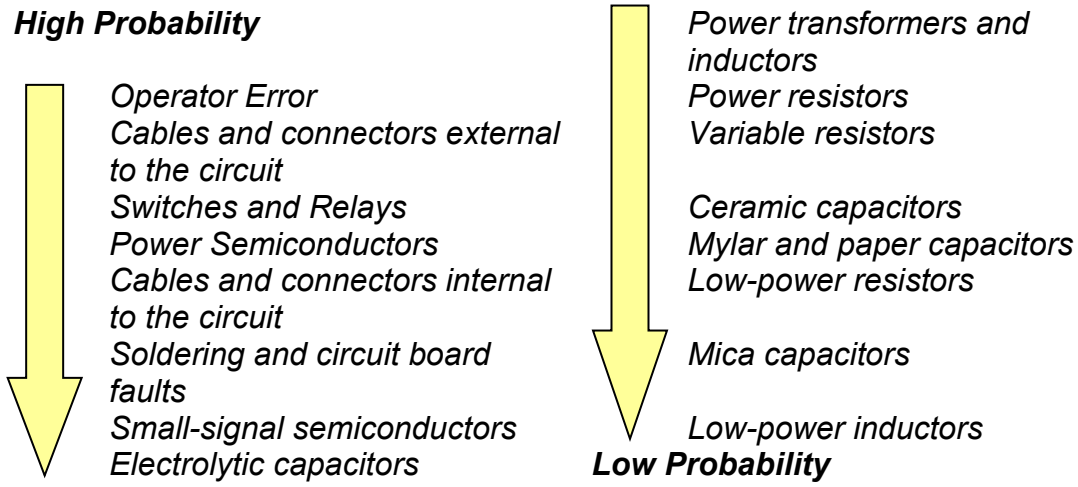
Experience has shown that many power-related disturbances occur because basic wiring and grounding practices are ignored or allowed to deteriorate. Loose connections in transformers, buses, distribution panels, junction boxes, etc. can be a source of power quality problems. In other cases, recommended wiring practices are defeated, such as allowing vacuum cleaners to be plugged into isolated ground circuits and frequently started induction motors to be served from the same transformer as sensitive computer power supplies.

Nonlinear loads such as PC's, fluorescent lights, adjustable speed motors, battery chargers and other solid-state equipment cause harmonics. Power supplies in these loads draw current in abrupt pulses rather than in a smooth sinusoidal shape. The result is distorted current wave shapes that contain harmonic currents that flow back into other parts of the power system. Harmonics troubleshooting may be done by first *inventorying the loads* supplied. If several of the equipment listed above appear on the inventory, then chances are good that harmonics exist. Locate the transformer supplying the loads and *check for excessive heating.* Make sure cooling vents are unobstructed.

Transformer secondary currents are then checked for phase imbalance (high neutral current *measured* vs. *calculated* from the vector sum of the phase currents) and harmonic frequencies measured on the neutral. *Panels feeding non-linear loads* are then checked, measuring the neutral current and comparing this to the rating for the wire size used. Thermography may help detect overheating here. Finally, *receptacle neutral to ground voltage* can be measured with the loads on. Two volts or less is normal, if higher, measure the frequency. A reading of 180 Hz strongly suggests the presence of harmonics, 60 Hz suggests the phases are out of balance.

Failure Probabilities

Certain types of components are observed to cause equipment failure more often than others. Knowledge of this hierarchy can be employed in failure analysis. The following list, gleaned from experience, proceeds from the most likely to least likely cause of the problem:



Some comments about the list:

Many failures are due to operators or maintenance personnel who make errors using the equipment. The analyst must first verify that the reported failure is real before proceeding with the analysis.

External wires and connectors are vulnerable to abuse and are often a source of trouble. Merely examining them or flexing them while the unit is operating will often show up a short, open or high resistance. Three examples of this include:

- *Emergency Diesel Generator control-related failures, due to cracked wiring insulation,*
- *Locomotive power wiring connector failures, due to their location above the hot diesel engine exhaust, and*
- *Nuclear Plant corroded control and protection wire terminal boards, due to poorly maintained junction box doors and seals.*

Switches and relays often become corroded, pitted or lose contact pressure. Slow operation of the switch handle, or a gentle pressure on the relay armature, may indicate an intermittent contact. Cleaning the contacts may help.

Power transistors and diodes fail frequently enough to make it profitable to disconnect the emitter and base leads, or either diode end, so that junctions can be checked with a VOM.

Failure Analysis Inside the Part

Now that the failed part and the failure mode have been identified, the failure analysis proceeds to its last stage, determining what went wrong inside. This phase should also be planned and conducted in a systematic manner, proceeding from non-destructive visual and micrographic examination to disassembly and perhaps cross-sectioning or other destructive techniques. The key point: *Don't start by cutting the part into pieces!*

At this point, too, the part has been removed from the system or circuit, and the analysis will occur in a laboratory equipped with the necessary instruments, sectioning equipment, etc. Several points are important here:

The scope of the analysis, observations from the previous failure analysis work, etc. should be communicated to the person(s) performing this analysis. There's nothing worse than to be the recipient of a failed electronic device with a work order that reads "*Analyze this by Tuesday and report back to J. Smith, ext. 3145."*

When removing the part from its location in the system or circuit, care should be taken to avoid damage, impact and vibration. The rule is *"Don't do anything that could damage the failed component or mask the failure mechanism!"* If the device is sensitive to electrostatic discharge, then it should be placed immediately in a clean, static-protective envelope or should be completely wrapped in aluminum foil before further packing.

A small, failed device should be placed in a clean plastic bag; larger devices may be wrapped in plastic sheet, with joints and overlaps taped. The packaging should be designed to prevent further damage, loss of critical parts and contamination. Final packaging should include a box or crate, with adequate packing material.

Here's an example that illustrates the process of determining the cause of failure within a part. Note how the process proceeds from "macro" to "micro," collecting evidence all along the way:

> ***The Case of the Spurious Trip:*** *A three-phase molded case circuit breaker, unexpectedly tripped at 66 amps, its load current. This breaker's normal trip setting was 100 amps and was of a type widely used in nuclear plant medium-voltage electrical distribution systems. Plant personnel observed that the breaker could not be reset for several minutes after tripping.*

The breaker was removed from its distribution panel and shipped to the failure analysis laboratory in its tripped condition;[7] this setting was not changed until interior visual inspection was completed.

Exterior inspection did not reveal any damage to the breaker housing; however, a black, tar-like substance had been used to insulate screw heads that held the breaker mechanism in place on the rear face of the housing. The substance was pliable everywhere except in the center phase (B phase) area, where it was brittle.

Radiographic examination of the breaker showed the interior structure, but no defects or damage was noted from this test. At this point, a systematic, step-by-step disassembly began. Plugs of white silicone rubber were removed from screw heads securing the breaker cover; traces of the black, tar-like substance were noted in these areas.

The interior mechanism appeared to be in good condition; photographs were taken and the breaker operated. All trip, reset, and closing functions were found to be operational. Resistances of the thermal trip heaters were measured, focusing on the interconnection point between the heater and the moving contact. Phase A and C resistances were approximately 1 milliohm, phase B was 5 milliohms.

Further disassembly (shunt-trip solenoid, aux contact switches, arc chutes, trip shaft) revealed a metallic discoloration of the phase B components. Additional disassembly allowed the mechanism to be lifted from its housing.

At this point, a close inspection of the phase B components occurred. Corrosion deposits were noted on both sides of the thermal-trip-heater mounting foot. The mounting foot and a wire lug (connected to the breaker's phase B movable contact arm) are screwed to the case. No such deposits were noted on the feet of the other phases.

Energy dispersive X-ray analysis (EDX) determined that potassium, silicon and bromine were present. As a minimum, it was suspected that the corrosive material was potassium bromate, potassium bromide and potassium silicate. Because EDX does not detect the first ten elements of the periodic table, it is possible that other potassium compounds were present, such as the oxide, hydroxide, borate, sulfide, etc.

[7]Resetting merely allows the breaker to be reclosed, but does not cause the breaker to close.

By this time, the analysis could conclude that the corrosion present on the current carrying mounting foot was creating a high resistance path. The subsequent heating caused the thermal-trip-heater's bimetal strip to deflect, thus tripping the circuit breaker.

The only question to resolve was how and when the corrosive substance was introduced to the breaker mechanism. This question proved intractable, however. The black, tar-like substance was evidence that the mechanism had been disassembled after its initial manufacturing.

The available manufacturing and plant records could not pin this point down, though. The corrosive material could have come from silver-plating residues on the mounting foot. It was equally possible that a corrosive material spilled on the mounting foot while it was awaiting assembly, or during the assembly process. Of course, if the breaker had been disassembled post-manufacturing, the corrosive substance could have been introduced then.

The only definitive conclusion as to root cause was that the observed failure was an isolated event, rather than one with generic implications for that class of molded case circuit breakers. This was based on the location of the corrosion in a tight metal-to-metal interface and because of the evidence that the breaker may have been opened after assembly.

Electrical/Electronic Failures – Diagnostic & Analysis Tools

Volt-Ohm-Meter - The electrical failure analyst's best friend is still the "humble" VOM. Voltages, both AC and DC, resistances and currents may be determined through this device. Needle probes will be useful for small circuit work. Although **not** recommended, many electricians still use the "two-wet fingers" test for establishing the presence of low voltages.

Megger - Motor and generator stator and rotor windings may be checked through use of a meggering device. Cable insulation is also checked using this device.

Oscilloscope - To measure waveforms and other time dependent signals, the oscilloscope is used. Some electrical schematics will include expected waveforms, aiding the analyst's diagnosis effort.

Curve Tracer - This is an extremely useful device for checking electronic components “out-of-circuit.” Transistors, diodes, etc. are all easily checked for shorts, opens and performance (gain). The curve tracer can also be used to perform pin-to-pin analysis of integrated circuits - open circuits show up as flat lines, shorts as vertical lines.

Leak Tester - Various methods of detecting leakage in hermetic packaging are available. The helium leak tester is one of the more common devices for detecting fine leaks. Care must be taken during the pressurization phase, especially for larger packages. Following pressurization, the device is placed in a spectrometer where the helium now leaking out of the package can be detected and the leak rate measured. Gross leakage may be detected by dye penetrant, bubble tests or gross leak weight gain tests. These are used for larger parts and larger holes.

Microscopes - Various types of microscope are available to examine circuit components. Low magnification stereoscopic viewing (x20 or so) is needed for overall inspection and lead examination. Pastelling that indicates overheating can be easily observed. Higher (x50) magnification is required for die interconnection pattern inspection. Bright-field vs. dark-field lighting may show different details better. For example, electrostatic discharge “punch-through” appears easier on dark-field. Magnifications up to x200 are needed to find details such as loss of continuity of aluminum interconnections at oxide step edges and photolithography faults. At this level, the search for pinholes and other defects is not too tedious. Microscopy beyond about x1000 is generally not practical for electronic failure analysis. The surface must be extremely level to be in focus. Higher magnification will require the electron microscope.

De-lidding Equipment - To examine the interior of a packaged chip or other electronic device, some means of removing the lid or hermetic seal must be available. Transistor cans may be removed by filing, ceramic flat-packs can be opened by lateral blows at the lid/body junction, grinding, or by bowing and cracking the seal in a vise.

Sectioning and Polishing Equipment - Failure analysis may benefit from cutting and polishing a section of the microcircuit, for analysis by optical or electron microscope. Electronic parts are typically encapsulated into an epoxy mold and then sectioned and polished.

Software Circuit Testers - Specialized circuit testers have been developed and used to determine the “health” of complicated integrated circuits. These testers have a software test package written for a specific IC, the IC is placed in the tester and exercised to determine if an internal failure exists.

Electron Beam Induced Current (EBIC) - This technique is used for electronic components and involves energizing the circuit (or a portion of the circuit) while examining the device in an electron microscope. For instance, if the circuit

contains a p-n junction and this can be reverse-biased, secondary electrons generated by the microscope's beam will be collected and information may be gained about the quality of the junction. Through this method, it is possible to discover breaks in electrical continuity (not otherwise apparent), junctions not able to hold reverse potentials, and the presence and dimensions of surface inversion layers under the passivating oxide.

Particle Impact Noise Detector (PIN-D) **-** Packaged electronics can fail due to particles that are introduced during the manufacturing process. These particles may cause shorts across conductors or otherwise damage the chip and its leads. The *Particle Impact Noise Detector (PIN-D)* is a method whereby the electronics package is shaken and acoustic monitoring tries to "listen" for the presence of a small particle as it bounces around in the package.

Other equipment that may prove useful in the failure analysis process includes:

- *DC Power Supplies*
- *Pulse Generators*
- *Electrometer (nanoamp measurement capability)*
- *Oven*
- *Infra-red Microscope*
- *Multiple Probe Unit*
- *Surgical Tools for Dissection*
- *Wire Bond Tester (Hook Test)*

MIL-STD-883 describes many electronic test procedures, which may be used for post-manufacturing and acceptance testing, as well as in the failure analysis process.

15.5.6 Mechanical Failure Analysis Process

Mechanical failure analysis differs from electrical failure analysis in very few aspects. Many of the same techniques (micrography, sectioning) are used in both analysis types. One difference, though, lies in the ease of detecting the failed part. In typical industrial mechanical systems, the component failure is relatively easy to identify. Visual or auditory inspection is often all that is required. A "terminology" note is useful here. Circuit breakers, relays, switches and generators are classified as "electrical" equipment. However, their operation depends on electro-mechanical principles. In many cases, their failure analysis may be based more on the discussion contained in this section than in the previous section. Similarly, IC wire bonding failures are often "mechanical" in nature. The boundary between electrical and mechanical failure analysis is not crystal clear. The general process consists of the following steps:

Collect and Document Pertinent Information

This is like the electrical failure analysis. The failed component or part must be identified and removed from its system or equipment (if possible - some large equipment failures must be analyzed *in-situ*). Fracture surfaces should be preserved and parts susceptible to corrosion protected. Background information on the component/part should be collected. Operating and maintenance histories, manufacturing records, and design information should be gathered. A sequence of events leading to the failure should be assembled. When was degradation first noticed? How often has this component failed in the past? Has industry experienced similar failures?

Perform Analysis

The same philosophy described in electrical failure analysis applies here. Review the molded case circuit breaker failure analysis example described above. Proceed with a *macroscopic analysis* (visual, low-power microscopic examination) first. Examine the fracture surface; look for discoloration due to overheating, corrosion, oxidation or contamination.

Examine the surface markings for clues. Shear lips indicate that a ductile fracture occurred; a combination of smooth, velvety surface and a coarse, grainy surface are indicative of elastic fracture. The specific mechanism sections below describe some failure indications. If necessary, perform a microscopic examination, using an optical or electron microscope. This may be necessary to confirm the results of the macroscopic examination. Material tests may also be necessary. Metal structure, hardness, tensile strength, fracture toughness may be in question. Material composition, contaminants and corrosion products may be determined via chemical analysis, spectrography or X-ray diffraction techniques.

Perform Stress or Other Engineering Analysis/Tests

If a design problem is suspected, a stress analysis can be performed to determine if the component has been adequately sized, or if there are stress raisers or concentrations. Many industrial failures occur simply because only corrective maintenance is performed. A Weibull analysis of similar failures can provide a basis for a preventive maintenance schedule or identify when it is economic to begin periodic equipment inspections. For some failures, to understand cause, the failure must be recreated or simulated.

Develop Conclusions & Prepare Failure Report

This step is the same as the other failure analyses. The ultimate root cause analysis report should be used both to prevent the reoccurrence of the current failure, but also be fed back into the design, manufacturing, assembly or operation process that "produced" the failure.

> *"Bad" Example: During the mid-1980s, a nuclear plant was experiencing a high number of Charging Pump failures. The root cause analysis identified an operating procedure as contributing to several of the failure modes. Valves downstream of the Charging Pumps were being used to throttle charging flow, instead of the variable speed positive displacement pumps. The procedure was changed and the failure frequency decreased. In the early 1990's, though, corporate "memory" had faded, the valves were once more being used to throttle flow and pump failures were on the rise.*

Mechanical Failure Classifications

Mechanical failures may be broadly classified into those that result in a *fracture* and those that do not. These, in turn, may be classified according to chemical, mechanical and thermal causes. This classification scheme may progress several more layers.

Fatigue may result in a fracture failure and is due to fluctuating stresses that tension, torsion or bend the material. *Corrosion* may be either *trans-* or *inter-crystalline.* *Stress-corrosion* may result in a fracture failure and is due to a combination of corrosion and fatigue mechanisms. *Creep* is a non-fracture failure resulting from operation at elevated temperatures. Metals may change properties during service; *neutron embrittlement* is of concern for nuclear reactor

pressure vessels. Here, the constant fast neutron bombardment causes the inner surface of the vessel to become less ductile and more brittle.

Extending this failure classification to the many non-metals incorporated in modern components would involve including the various sources of instability (apart from buckling): elastic, plastic, dynamic strength, together with more detailed consideration of non-metallic joining methods failures and their causes: faulty joining by glue, trapped air, glue instabilities, softening by humidity, embrittlement, overheating, etc. New materials result in new failure modes: fiberglass-reinforced plastic boats were plagued by a "blistering" phenomena, where water absorbed into the plastic caused localized delaminating of the structure.

The failures experienced most frequently in machinery are fracture, wear, excessive deformation and surface failure, particularly corrosion deterioration. Fluid systems may experience the above failures, as well as blockages from various causes. The next sections describe the failure mechanisms, their effects, causes and detection.

Fatigue Failure Mechanisms

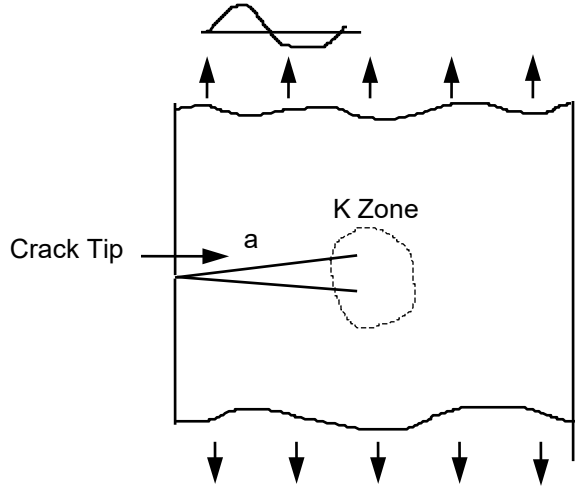
Fatigue occurs under the action of cyclic loading, when a crack initiates and grows. Understanding the *rate* of crack growth is of great practical significance, since the ultimate fracture failure can be prevented through periodic inspection and repair/ replacement. Fatigue starts with the formation of surface microcracks (whether by surface roughening, grain boundary cracking, corrosion-produced pits, or cracking around hard inclusions), which then extend across and penetrate the body of the metal.

The surface microcracks initially develop in the same direction as the operative slip plains. This continues until the microcrack is large enough to affect a significant volume of material. Continued cyclic stresses (the *normal* component of these stresses to the crack provides the open/close force) then cause the crack to grow as if it was in a continuum.

The growth process is now associated with the magnitude of the tensile strain range in the volume of material just ahead of the crack edge, and the growth direction becomes normal to the direction of the maximum cyclic tensile stress. By this time, the crack may be detected either visually or through various test/inspection methods and is a *macrocrack.* The figure below is a cartoon description of the key aspects of this phenomenon.

Cyclic Stress (σ)

Figure - Fatigue Cracking



K = Stress Intensity Factor (units - psi $in.^{1/2}$)
"K" is a unique, singular term that describes the magnitude of the intensification of elastic stresses in the region of the crack tip. "K" is dependent on:
o Externally applied load (σ),
o Length of crack (a), and
o Geometry of cracked body and load application method (f)

Unstable fracture occurs if: $K > K_{IC}$
(Applied Stress Intensity) > (Material Fracture Toughness)

While the end of component life is based on the *Applied Stress Intensity* exceeding the *Material Fracture Toughness,* the useful life of a component depends on the rate of growth of flaws from the microcrack region to the critical size ($K = K_{IC}$). For fatigue, crack growth rate (expressed as *da/dN* - change in crack size, *a,* with respect to elapsed loading cycles, *N*) has been found to depend primarily on the cyclic range of the applied stress intensity factor, ΔK. ΔK is analogous to the $\Delta\sigma$ used in conventional fatigue analyses. But, from the Figure below, K depends on the crack length, stress, loading method and geometry. An example of the relationship between K and its factors for simple tensile opening of a crack is approximated by:

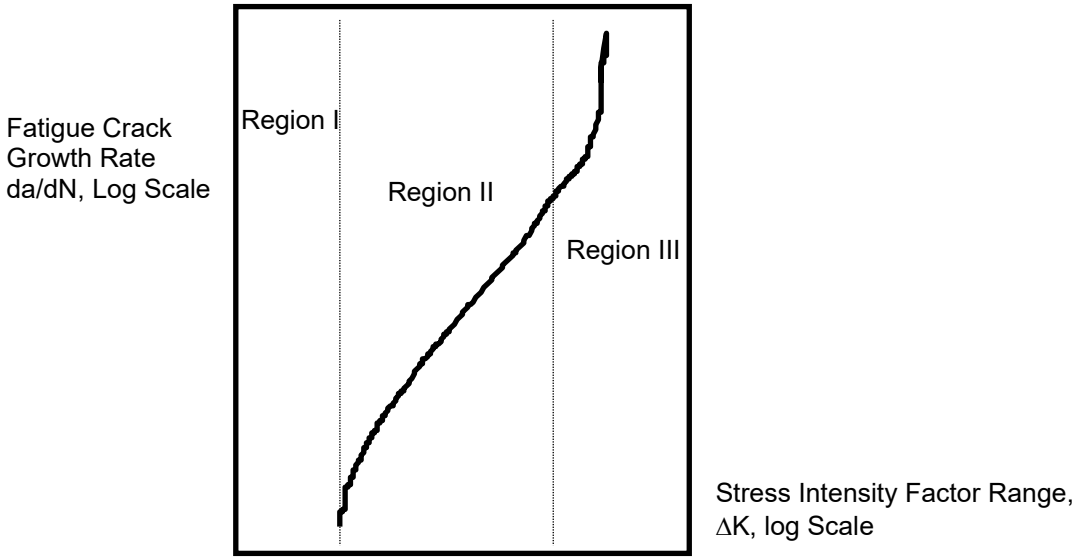
$$K = \sigma(\pi \times a)^{1/2}$$

and, for fatigue analysis,

$$\Delta K = \Delta\sigma(\pi \times a)^{1/2}$$

As the crack grows during cyclical loading under constant $\Delta\sigma$, the increase in crack length, *a,* results in a corresponding increase in ΔK. When ΔK reaches the material fracture toughness, brittle fracture will occur.

Figure - Fatigue Crack Growth Behavior



The figure above is a schematic of general crack growth behavior. Three regions exist. Region I is where very low ΔK ranges do not cause preexisting cracks to grow under cyclic loading. Region II shows a linear relationship between the log da/dN and log ΔK over a wide range of ΔK levels. This linear relationship allows the fatigue crack growth behavior to be characterized in terms of a general model:

$$da / dN = C_o(\Delta K)^n$$

where:

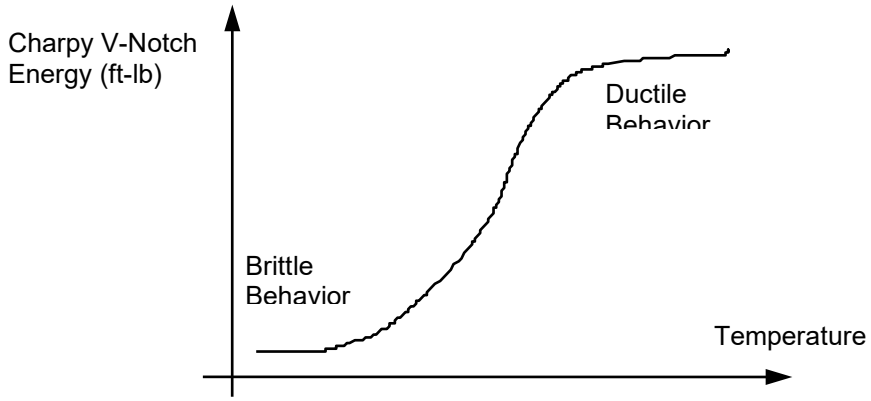
C_o - empirical intercept constant

n - slope of log da / dN vs. log ΔK line

In Region III, the ΔK level is so close to the material fracture toughness, some unstable crack growth occurs during each cycle; final fracture is soon reached. The Region II linear behavior allows us to predict crack growth, and schedule periodic inspections to monitor the condition of the component or part. The fracture toughness of a material is dependent on several factors, including crack location and its orientation to the material's processing direction and, especially, material temperature. As the material's temperature is decreased, fracture behavior changes from ductile to brittle, with a

corresponding decrease in fracture toughness. The Charpy V-Notch test can be used to obtain a picture of this behavior. Impact Energy is plotted against material temperature as shown below.

Brittle-Ductile Behavior vs. Temperature



Types of Fatigue

There are two basic classifications of fatigue – *Low* and *High Cycle* fatigue. In addition, fatigue may act in combination with other failure mechanisms (primarily corrosion) to cause component failure through fracture.

Low-Cycle Fatigue - Under stress, a metal will initially *elastically* deform. Upon removal of the stress, the metal will return to its previous shape. If the stress becomes too high, though, the metal will undergo *plastic deformation.* It is under these conditions that *low-cycle fatigue* occurs. Each stress cycle results in plastic deformation of the material, failure typically occurs within a few tens or hundreds of cycles, hence the term *low-cycle.* Bending a paper clip back and forth will result in its fracture within 10 – 30 cycles and is a simple example of low-cycle fatigue. This failure mechanism will be discussed under *Excessive Deformation;* the remaining fatigue discussion concerns *High Cycle Fatigue.*

High-Cycle Fatigue - High cycle fatigue occurs when the stresses do not exceed the *yield point* of the material, that is, all bending, torsion or tension results in *elastic deformation*. Historically, high-cycle fatigue was a surprise to engineers, since the prevailing theories held that the material would not fail if it was only elastically deformed. As machinery was developed, which ran at higher speeds under load (and hence, accumulated more cycles), high-cycle fatigue failures became common.

Subsurface Fatigue - Although most fatigue failures initiate from surface defects, a number may originate from below the surface. The common sources of subsurface-originated failures include:

Inclusions – The presence of discontinuities in the material introduces possible sites for stress raising. Investigations of the effects of biaxial or triaxial stress concentrations suggest that high subsurface shear stresses play an important role. The effects of inclusions are particularly critical if they consist of hard and nonductile materials and are located in areas of high stress.

Hardened Surfaces – The transition areas between a case-hardened surface and softer inner core have been frequently observed to be the sites from which fatigue may develop. The change in microscopic structure, together with residual stresses is believed to be responsible here. Bearing surfaces are often case-hardened and may be subject to this type of fatigue cracking.

Rolling Loads – Ball bearings and rail tracks experience fluctuating shear stresses under normal working conditions. Cracks may form under the surface of these components. For ball bearings, pitting, spalling and flaking are the result, for rails, “shelling” is observed.

Causes and Accelerators of Fatigue Failure

Although any cyclically stressed metal piece can fail through fatigue, there are several design, assembly and operational factors that can accelerate fatigue failures:

Geometrical Shape Changes – The geometry of a metal component can cause local intense stress concentrations. Finite element analysis is used today to predict these stress concentrations.

Notches – Any change of section that alters the local stress distribution is termed a “notch.” Keyways, circumferential grooves, holes, contour changes, scratches, threads, etc. are all “notches.” Hard and high-tensile strength materials are particularly sensitive.

Corrosion – Pits on the surface act as notches with the same damaging effects; the effects are exacerbated when, during the tensile loading, the crack is opened and exposed to the corrosive environment.

Fretting Corrosion – When two parts are clamped, press-fitted or shrink-fitted and subjected to alternating flexure, an oxide is produced at the inter-surface, causing failure like a corrosion product. In many cases, the initial crack is formed, undetected, in the press fit.

Internal Stresses – Residual stresses combined with those arising from the cyclic stress can result in high local material stresses. Heat treatment, quenching, drawing, rolling, cold working, welds not stress-relieved are sources of residual stresses.

Decarburization – Carbon lost from the surface of a ferrous alloy following heating in a reactive medium reduces the fatigue resistance on the surface.

Fatigue Detection and Fatigue Fracture Identification

Following fatigue failure, the break surface will have the following characteristic appearance:

- *Little permanent deformation,*
- *Break marks showing the progression of the crack can be seen under a microscope,*
- *The break marks are smooth because of rubbing, and*
- *The fractures propagate in a direction normal to the principal tensile axis (where combined stresses are involved).*

The fracture surface consists of two distinct regions, one smooth and velvety (the fatigue zone) and the other coarse and crystalline (the instantaneous zone). Rubbing of the mating surfaces as the crack opens and closes causes the smooth velvety appearance of the fatigue zone (this is often mistaken for “good” metal). The coarse appearance of the other surface is the fatigue fracture that occurs as the stress-intensity exceeds the inherent material fracture toughness. This surface is a brittle-type fracture. “Atlases” of fracture surfaces exist to help the failure analyst quickly identify the type of failure mechanism she is observing.

Wear Failures

Most deterioration of machines is a consequence of wear between two surfaces moving relative to each other. Wear is a process of surface damage producing wear debris (which may have secondary effects.). When two surfaces are in rubbing contact, the surface layer processes may involve one or more of the following:

Microcutting is where abrasive wear particles gouge out tiny particles of metal or cause surface deformations under plastic flow.

Plastic and Elastic-Plastic Deformation, which occurs to specific areas of the surface as a result of high (Hertzian) local pressures. This is the result of contact between surface irregularities when the surfaces come into contact.

Surface fatigue, which is the result of repeated elastic deformations to the surface as a result of cyclic forces.

Local Heating, which is the result of insufficient heat transfer due to friction (usually the combination of high speed and pressure). Local temperature rises can be so high as to cause phase changes and melting of the surfaces (local weld junctions).

Oxidation, which is the formation of films of solid solutions and oxides that are removed.

Molecular interaction occurs when surfaces become bound together under intense pressures and low speeds. Cold welding results and metal particles are transferred from one surface to another.

Rehbinder effect occurs when lubricant fills up microcracks in the material surface, resulting in an increase in pressure that leads to damage to the surface layers.

These mechanisms may occur simultaneously. The actual forms of wear may be classified as follows:

Abrasive Wear - Plowing or gouging of hard particles against a relatively smooth working surface causes abrasive wear. This is probably the most serious single cause of wear in engineering practice. Lubricant filtration and efficient sealing of bearings are important corrective actions to prevent this source of deterioration.

Scuffing - Scuffing occurs when the sliding surfaces come into contact and the “ideal” hydrodynamic lubrication cannot be maintained. Scuffing is a cumulative process that can lead to seizure of bearings or to excessive wear in piston rings or cylinder bores. This phenomenon is difficult to interpret, since the cumulative action destroys its initial stages. Compatibility of contacting materials is important, as are the choice of surface treatments and lubricants. Extreme pressure and anti-wear additives may protect against scuffing, but their possible adverse effect on some bearing materials must be noted.

Fatigue (Pitting) Wear - Fatigue wear occurs usually in rolling friction and is caused by fatigue of the surface layers. If there is also relative sliding of the surfaces, wear resulting from fatigue of the micro-surface roughnesses is also possible.

G*alling/Adhesive Wear* - Galling is characterized by the development of local metal joints. Particles may be removed from or adhere to the rubbing surfaces. This type of wear occurs with high pressure and, as a rule, develops at a rapid rate. If there is considerable heating in the sliding zone (perhaps due to high sliding velocities), this is termed thermal wear.

Mechanical-Corrosion Wear - This occurs if oxidation processes are important on the surfaces. The plastically deformed and oxygen-saturated surface layer of the component fractures because of **repeated loading and fresh sublayers of the metal become exposed.**

Cavitation Wear - Cavitation wear is usually associated with high-speed propellers, or pumps operated without adequate suction head. As the surface moves through the fluid, bubbles may form on the low-pressure side of the propeller, if the pressure is below saturation for the fluid's temperature. The bubbles' collapse results in shock waves that erode the surface through impingement and/or chemical surface activity. The surface becomes pitted and fatigue cracks are likely to arise from these pits.

Excessive Deformation Failures

Loads that impose stresses in excess of the elastic limit may result in functional failure. Failure via fracture is usual, but the component may functionally fail before fracture occurs. These loads may be applied *statically, cyclically,* or *dynamically.*

Static Loads may be applied gradually, so that at any time instant, all parts are essentially in equilibrium. Such static loading is typical when the load slowly and progressively increases to its maximum service value for some short or long period of time.

Cyclic Loads are associated with Low Cycle fatigue failures. Plastic crack progression is different than elastic. They both start with a sharp microcrack, but the plastic crack initially blunts, followed by a period of stable, ductile crack growth. The final failure occurs via a rapid, unstable ductile fracture.

Dynamic Loads are either sudden or impact loads. There is no static equilibrium in the load, except for the terminal positions. Sudden loading occurs when a mass or “dead load” not in motion is suddenly applied to a body. This can create a stress approximately twice as great as if the mass were applied gently.

Impact loads are associated with motion as one body strikes another. High stresses are generated as the kinetic energy of the body is transferred to strain energy. Materials that ordinarily fail in a ductile manner under static loads often fail in a brittle manner if the impact loading rate is high.

Effects of these load types include the following failures:

Indentation - Dents, pits, grooves and distortions commonly result from the application of excessive stress. If one material is significantly harder than the other, the softer material will undergo greater deformation.

Cleavage Fracture - This is a failure mode that takes place along well-defined crystal planes with the grains. The cleavage fracture surface contains large areas that are relatively smooth and featureless, separated by characteristic features such as cleavage steps, feathers and tongues and river markings.

Ductile Failure - Here, the material is overloaded. The metal pulls apart under the applied stress at various discontinuities (inclusions, precipitates and grain boundaries). As stress increases, these “microvoids” grow and coalesce to form a continuous fracture surface. The failure surface has a dimpled appearance from this process. If the failure was due to purely tensile stresses, the dimples will be circular. If failure was due to shear stresses, the dimples will assume a parabolic form.

Static Load Fractures - Here, a single load causes stresses in excess of the ultimate strength of the material. The fracture will involve permanent deformation of the part.

Tension Fractures - This produces a local deformation or “necking” of the part. The fracture surface is made up of planes of separation inclined at about 45 degrees to the direction of the load. Two parts of a rod broken by axial tension may resemble a “cup and core” with 45-degree beveled edges. Pure tension fractures part cleanly with no rubbing between the pieces.

Compression Failures - Compression failures come in two varieties: Block compression occurs in short, heavy sections which separate on oblique planes (as in tension), except that there is rubbing of the halves during separation. Buckling occurs in long sections and typically results in bending of the part.

Bending Failure - When a part is subjected to a bending moment, it resists through both tensile and compressive stresses (the cantilever is a simple example). Failures through bending moments will then display a tension-type fracture on the outside of the bend and a compression-type fracture on the inside.

Shear Failures - Two types of failure under shear can be identified: With Block shear, the two halves of the fracture slide one across the other and the surface will appear rubbed, polished or scored - the direction of the scoring indicates the direction of the shear force. Buckling occurs in metal sheets, usually in a diagonal of the sheared panel. When rivets, screws or bolts fail in shear, the hole elongates and an open, moon-shaped space appears behind the fastener.

Torsion - This is a form of shear. The two halves of the fractured metal specimen retain some permanent twist - the fracture surface often exhibits tensile fracture surfaces oblique to the angle of twist.

Impacts - Collisions and explosions create stresses under impulsive action because of stress waves. A central cavity may be formed; the opposite face of the material may have spalled due to reflected stress waves.

Corrosion Failures

As with wear, corrosion is a major source of mechanical failures (as well as civil/structural failures). Corrosive deterioration arises because of electrochemical or chemical erosion attack due to environmental conditions such as salt atmosphere. Corrosion also results from the presence of anodic (more noble) bodies in an electrolytic environment, such as occur when dissimilar metals are immersed in seawater. Biological corrosion can result from the deterioration of marine growths (barnacles, etc.) that attach themselves to underwater or tidal metal structures. Several common types of corrosion experienced in power generation equipment are presented below:

General Corrosion **-** Power plants are often located near the ocean or brackish bays and rivers. Constant exposure to salt atmosphere can cause general corrosion of metal surfaces. Corrosion management is a constant battle, with protective coatings relied on to seal the metal surfaces from the salt atmosphere. Attention must be paid to a variety of components. For example, electrical junction and terminal boxes must be sealed to prevent salt and moisture build up across connections.

Galvanic Corrosion **-** Power plant systems are constructed from different metals, which are more or less noble. Fluids pumped through the systems can act as electrolytes, with the result being galvanic corrosion. Normally, the plant is equipped with a galvanic protection system; however, this may not be maintained or operated well.

Stress Corrosion **-** This is cracking of the material under the action of a steady stress in a corrosive environment. The cracks tend to lie perpendicular to the principal tensile stress without plastic deformation. The cracks follow an irregular route and result in a coarsely textured fracture face. *Transgranular* and *intergranular* stress corrosion is possible depending on the material and corrosive agents involved. Water chemistry control is critical to managing stress corrosion cracking. One critical example of this involves condenser tube leaks. If the raw water is obtained from the sea or any salt-containing water, these leaks will introduce chlorides into the feedwater system, with subsequent corrosion. Nuclear plant steam generators are sensitive to this phenomenon.

Corrosion Fatigue **-** Corrosion fatigue arises under the combined influence of a fluctuating tensile stress and a corrosive atmosphere. The cracking process is like that described in the *Fatigue* section, however, the presence of corrosive fluid in the crack can accelerate the failure.

Sulfur Corrosion **-** Most fossil fuels contain some amount of sulfur as an impurity. Burning the fuel in a boiler produces water and sulfur trioxide, which can combine to form sulfuric acid. This acid condenses on cold surfaces and causes corrosion of boiler metal. In gas turbines, problems have occurred when sulfur in the fuel combined with sodium from airborne sea salt to form molten sodium sulfate. The sodium sulfate collected on turbine nozzle guide vanes with damaging results. The resulting corrosion process destroyed over 90% of the guide vanes. In diesel engines, an opposite problem has been noted. When low sulfur fuels are used, a “designed” reaction between the sulfur and high alkaline lubricants does not occur. The unreacted alkaline material has caused high wear rates of engine liner and piston rings.

Vanadium Deposits **-** Diesel and boiler fuels often contain vanadium, present in the metallic compound of vanadium porphyrin. Following combustion, various low-melting point vanadium compounds are found in ash particles that collect

around fuel injectors. These compounds adhere to metal surfaces, resulting in corrosion. In boilers, the compounds also reduce heat transfer efficiency.

Lead Deposits **-** Lead deposits on exhaust valves can, at high temperature (>800 C), react with protective chromium-rich oxide layers and destroy the oxidation resistance of the valves. Colored compounds seen on exhaust valves indicate that the corrosion reaction is proceeding.

Blockages

Fluid systems can fail due to various types of blockage. Examples of common failure types are described below:

Cooling and Water System Blockages **-** Scale and mineral deposits can form in heat exchanger tubes when water chemistry is not or cannot be controlled. Condensers and Raw water heat exchangers typically have the cooling water pumped through their tube sides. Marine organisms (barnacles, clams, mussels, etc.) find these environments a wonderful place to live, since a source of nutrients is constantly flowing past them.

Degraded thermal performance is the immediate effect of these deposits; however, the mineral deposits may induce tube corrosion if not directly, then through the wear accumulated by the tube cleaning processes employed. Blockage of cooling systems in equipment cooling heat exchangers (pump seal and bearing coolers, turbine lube oil coolers, air compressor intercoolers and dryers, diesel engine block cooling passages) can lead to equipment failure through overheating and thermal stress. Hard water will deposit scale on water supply pipes, eventually restricting flow.

Fuel System Blockages **-** Fuel often contains impurities that, if not properly filtered, can clog injector ports. Algae can grow in diesel fuel tanks; standby diesel generator fuel supplies are particularly susceptible. Fuel additives are available to prevent algae growth.

Lubrication and Control Oil System Blockages - In older turbine generator systems, lubrication and turbine control oil are supplied through a common pumping and filtration system. Unless the filtration is maintained in good condition, wear particles from the turbine bearings can be carried into the control system piping and valves. The control system valves are especially susceptible to blockage from particles; poor turbine control or turbine trip-initiated shutdowns may result from this blockage. Control oils must be matched carefully to the other materials in the hydraulic control system. In one case, new fire-retardant oil was introduced into the St. Lucie Nuclear Plant turbine control system as part of an NRC commitment. The new oil dissolved valve-sealing material; the dissolved particles were then carried to the control system

valves where they caused the valves to stick. During a routine shutdown, the control room attempted to trip the turbine, with no success. An operator had to be dispatched to the turbine control stand (north side of the high-pressure turbine) to manually shutdown the turbine.

Compressed Air System Blockages - Industrial compressed air systems draw moisture-laden outside air into the compressor where it is compressed, cooled and dried. If the drying function is not maintained in good condition, water carryover will occur in the air system. This water may collect in solenoid actuator valves or air-control valve diaphragms, degrading or preventing their operation.

Design, Manufacturing and Assembly Failures

Faults "produced" by the design, manufacturing or assembly processes cause a variety of mechanical failures. A non-exhaustive list appears below:

Process	Failure Cause
Design	*Bad Geometry, Sudden Change of Section, Sharp Corners*
	Bad Choice of Materials, Dissimilar Metals in Contact, Differential Coefficient of Expansion
	Underestimate of Stress, Failure to consider Fluctuating Stress
Manufacturing	*Error in Material Used, Mixed Stores, Parts Control Problems*
	Unsuitable Chemical Composition, Beta Structure in Brass and Bronze, Poor Physical Properties
	Poor Casting Technique, Cavities, Segregation, Local Shrinkage, Micro-porosity, Displaced Cores
	Non-metallic Inclusions, Roaks, Laps, Lamination, Blistering, Oxide Films
	Fabrication defects, Wrong Grain Disposition, Excessive Cold Work, Surface Cracks, High Surface Stress, Surface Contamination, Die Marks
	Faulty Heat Treatment, Over/Under Heating, Oxidation, Lack of Stress Relieving, Unsuitable Quenching
	Faulty Machining, Roughness, Grooving, Undercutting, Stepped Fillets, Sharp Corners
	Faulty Protection, Inadequate Cleaning, Hydrogen Embrittlement, Thin Coatings, Pinholes
Assembly	*Faulty Joining, Riveting, Bad Drilling, Over-riveting, Welding, Brazing, Soldering, Poor Fusion, Entrapped Flux, Oxidation, Over-heating, Grain penetration by molten metal, Cracking*
	Faulty Fitting, Damage, Tool Marks, Foreign material in fluid systems, Screw threads, Over/Under tightening, Different/Crossed Threads, Bad Fits, Cold working, Mispositioned Component

Mechanical Failures – Diagnostic & Analysis Methods

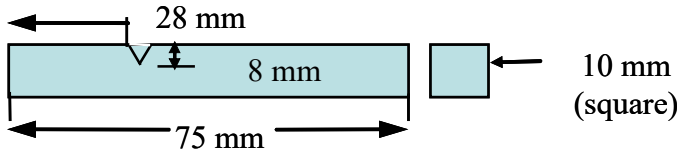
Dimensioning/Measuring Tools - Calipers, micrometers and even rulers are required to obtain basic dimensions of failed components. An "instant-type" camera is almost indispensable for recording the *in-situ* failure conditions. Video cameras are likewise valuable.

General Tools for Disassembly - Screwdrivers, wrenches, pliers, hammers, punches, etc. are needed to disassemble mechanical equipment. Often, the equipment vendor will provide special tools to facilitate disassembly.

Microscopes - Visual examination of failed components is an important failure analysis tool. Microscopes can help reveal the type of fracture (i.e. brittle vs. ductile) and other surface/subsurface features that provide clues as to the origin of failure (See electrical/electronic equipment).

Acoustic microscopy can be employed in the non-destructive examination of electronic parts or other materials where these interfaces can be examined. Cracking or fractures in IC's, delamination of the chip from the substrate, inclusions in material are examples of failures which can be detected through this method.

Charpy V-Notch Test - If the failed material's properties are in question, the Charpy V-Notch Test (CVNT) can be used to provide an estimate of the material's fracture toughness. The CVNT measures the energy required to fracture a material by swinging a hammer against a notched sample of the material (standard dimensions shown below).



If the CVNT is being used to determine if the material fractured due to low toughness, it is important that the factors affecting toughness be taken into consideration during the test. For instance, if the material failed at low temperature, the test sample should be cooled and tested at that temperature. *Fatigue Failure Mechanisms* describes the temperature effect on fracture toughness and how, at low temperatures, the toughness may be significantly lower than at elevated temperatures.

Load Tests - The failed material's tensile, compressive, shear and/or torsion strengths may be in question. A load cell coupled with proper measurement instrumentation can be used to establish these properties for samples of the failed material.

Hardness Testers - The failed material's hardness may also be in question and may be determined through a hardness tester. The Rockwell scale is most used for hardness comparisons.

Sectioning and Polishing Equipment - Like electrical/electronic parts, examination of material sections may give clues as to failure cause. Cross-sections of the fracture surface can help reveal the ductile and brittle fracture details; microscopic examination can detect the presence of pores or inclusions in the metal structure. Sample preparation typically involves *cutting* the material perpendicular to the fracture surface, *embedding* (typically in Bakelite), *rough* and *fine grinding,* and *polishing.*

Spectrograph - Material composition may be determined through a spectrograph. *Oil Sample Analysis* is often used to determine the presence of wear particles in lubricating oil; this can be used in both failure analysis and to predict failure. Often, determining the material composition can be the conclusion of the failure analysis. Experience has shown that many failures are the result of wrong material used versus specified.

Acoustic Monitors/Ultrasonics - Fluid passing through a pipe, tube, valve or heat exchanger creates noise. Acoustic monitors can help determine the *condition* of the component. For example, acoustic monitoring of power plant check valves has been employed to determine back-leakage volumetric rates, as well as the condition of the flapper during forward flow conditions. Whereas acoustic monitoring is a passive, "listening" approach, ultrasonic testing of materials consists of transmitting sound waves through the material. The sound waves are reflected and/or diffracted at discontinuities in the material that generally indicate the presence of flaws or defects inside the material.

Vibration Detection Equipment - Rotating machinery emits a characteristic spectrum of vibration - amplitude vs. vibration frequency. If a baseline spectrum is established during machinery startup, changes in this spectrum can be indications of a failure in progress. Recent advances in vibration monitoring equipment have caused most monitoring programs to shift from the old, amplitude-based vibration monitoring (outboard shaft bearing reading 4 mils vibration) to the vibration spectrum approach.

Eddy current testing is widely used in applications where the condition of heat exchanger tubes is of importance. For example, nuclear plant steam generator tubes are subjected to eddy current testing periodically, to monitor the

progression of tube cracks. A probe is passed up into the tube; data collection equipment then monitors and maps the condition of the tube. Flaws as small as 10% of through-wall thickness can be detected.

NDE Tests (Dye Penetrant, Magnetic Particle, X-Ray, Others) - A variety of tests are available to detect the presence of surface cracks and other flaws in materials. These are used to establish a baseline when the equipment is first placed into service, as well as to monitor the progression of failures during the life of the equipment and as part of the failure analysis effort.

15.5.7 Civil/Structural Failure Analysis Process

There is even less difference between the process of civic/structural and mechanical failure analysis than there is between electrical and mechanical. Many civil/structural failures result from fatigue, corrosion or wear. Civil/structural failures, though, will often involve materials other than metals and involve "passive" (i.e. non-rotating, or moving) devices. Examples of "devices" that fall into this category include:

- Buildings
- Storage Tanks
- Bridges
- Roadways
- Railways
- Chimneys
- Dams
- Piers
- Mines
- TV/Radio Transmitting Antenna
- Ships and Boats
- Transmission Line Structures
- Utility Poles
- Automobiles (Crash-worthiness)
- Structural Supports

The timing of a civil/structural failure analysis may differ slightly from its electrical/mechanical counterparts. If a diesel engine failure occurs due to a faulty injector, the injector may be quickly replaced and the engine placed back in service while the failure analysis proceeds. For these failures, the failure analysis must often be completed prior to repairs being initiated, since the cause of failure must be understood to develop an effective repair that prevents failure recurrence.

> *Example: Freeze/thaw damage of pier structures was visible on the concrete surfaces. Initial repair recommendations were to patch the loose and spalled concrete. Core samples from the piers, though, revealed that the damage was more extensive, running over a foot deep in some areas. The repair required the extensive removal of deteriorated concrete to sound concrete, placement of rebar dowels, anchors and additional reinforcing bars, followed by concrete poured into forms installed to recreate the original pier shapes and sizes.*

The general analysis steps are discussed below:

Collect and Document Pertinent Information

This is like mechanical failure analysis. If possible, the failed component or part must be identified and removed from its system or equipment. In many civil failure analyses, this is not practical or even desirable. Photography and measurements are often collected at the failure site instead. Core samples of concrete or other structural material may be the equivalent of "parts." Fracture surfaces should be preserved and parts susceptible to corrosion protected.

Background information on the device should be collected. Operating and maintenance histories, manufacturing and construction records, and design information should be gathered. This may be difficult since many civil structures are not well documented when constructed and changes/repairs are likewise poorly documented. A sequence of events leading to the failure should be assembled. When was degradation first noticed? How often has this device failed in the past? Has industry experienced similar failures? Many civil structure failures are due to the cumulative effects of environmental factors (chemicals, saltwater, freeze/thaw cycles, insects, etc.). It may be necessary to obtain weather data or other environmental information.

Perform Analysis

The same philosophy described in mechanical failure analysis applies here. Proceed with a *macroscopic analysis* (visual, low-power microscopic examination) first. If necessary, perform a microscopic examination. This may be necessary to confirm the results of the macroscopic examination. Material tests may also be necessary. Metal structure, hardness, tensile strength, fracture toughness may be in question. Concrete, fiberglass, or wood compressive/tensile strength may be of interest. Material composition, contaminants and corrosion products may be determined via chemical analysis, spectrography or X-Ray diffraction techniques.

Perform Stress or Other Engineering Analysis/Tests

If a design problem is suspected, a stress analysis can be performed to determine if the device has been adequately sized, or if there are stress raisers or concentrations. As with mechanical failures, many civil/structural failures occur simply because only corrective maintenance is performed. A Weibull analysis of similar failures can provide a basis for a preventive maintenance schedule or identify when it is economic to begin periodic equipment inspections.

> *Example: Distribution pole failures were occurring for years at a central-US electric utility. The utility has simply followed a corrective maintenance philosophy - when a wooden pole fell or one of the guy wires*

rusted through, it was replaced. Initial investigation of the failures identified two primary reasons for failure: Butt rot and woodpecker damage.

For some failures, to understand cause, the failure must be recreated or simulated.

Develop Conclusions & Prepare Failure Report

This step is the same as in other failure analyses. The ultimate root cause analysis report should be used both to prevent the reoccurrence of the current failure, but also be fed back into the design, manufacturing, construction, assembly or operation process that "produced" the failure.

Civil/Structural Failures

Some common materials used in the construction of structures and their failures are presented below:

Concrete/Reinforced Concrete Failures

Concrete and reinforced concrete are widely used in buildings, bridges, roadways, piers, etc. Failures of this material can result in surface imperfections (spalling, surface cracking), leakage (through joints or cracks in monolithic structures) or outright structural failure (collapse). Causes generally fall into one or more of the following categories:

Freeze/Thaw Cycles - If the structure exists in a climate where freezes are common, accumulation of water in on the structure's surface or joints prior to a cold spell can damage the structure. Since water expands when it freezes, this can cause cracks and joints in the structure to expand, leading to failure. Even Mt. Rushmore is monitored every year for cracks in the rock; these are filled with caulking to prevent the intrusion of water prior to the winter season.

Chemical Attack - Water treatment plants, chemical processing plants and other facilities use chemicals that can attack concrete. Acids will dissolve concrete. In one case, at a power plant water treatment plant, acid runoff over years created large voids in the ground underneath the plant, since the ground was composed mainly of crushed limestone.

Deicing Salts - In the winter season, deicing salts are often applied to roadways and other concrete surfaces used for travel. The solution formed from melting snow/ice and the salts then attack the concrete and its reinforcing steel. Residents of the "Snow Belt" are familiar with road cracks and potholes formed through this process

Physical Wear - As with any surface, concrete is subject to wear, either from mechanical sources (i.e. traffic on a concrete roadway) or fluid (water flow through concrete-lined pipes).

Corrosion of Reinforcing Steel - If a corrosive chemical is allowed to come in contact with the reinforcing steel, then the steel will corrode, producing iron oxide (rust). Two problems occur here. First, the reinforcing steel swells as it rusts. This results in a "debonding" of the steel/concrete structure, with consequent loss of strength. Second, loss of reinforcing steel as the rust process continues contributes to further loss of strength.

Crack/Joint Leakage - Expansion and contraction joints are designed to allow for structural movement. Greater than anticipated movement can lead to failure of these joints and result in leaks. Differential pressure across the joint can also lead to failure.

Structure Overloads - Every structure will bend, vibrate, flex and settle in service. Three primary failure modes occur: Shear capacity, compressive capacity, and tensile capacity. Overload failures can be the result of design or construction errors, operational overload, or outside factors such as weather (wind, snow, ice).

Design Error - A variety of design errors can contribute to structural failure. They may be as simple as not following the local or national code for "simple" structures, to inadequate stress analysis for more complex structures.

Construction Error - Errors can occur in mixing concrete (wrong ratios of water, cement and aggregate) - visual examination (indications of porosity, scaling and shrinkage) and concrete compressive strength will indicate problems in this area. Errors can also occur in pouring/curing the concrete/reinforced concrete structure. Voids are common in pouring concrete; for multiple pours, poor preparation of the previously poured surface can lead to inadequate bonding (cold joints). Attempts at "economy" can result in surfaces thinner than required by design/code. Reinforcing steel (rebar) may be positioned too close to the concrete surface - this can lead to corrosion and spalling of the surface. Some reinforcing systems require the rebar to be mechanically joined or welded. This may not have occurred or may have been done inadequately. In some cases, placing too much reinforcing steel can result in voids in the concrete. Poor site preparation can lead to floor settlement and cracking.

Of course, the construction forces may simply not have implemented the design as shown on the engineering drawings.

Example: At a Nuclear Plant, over 40 walls in the Auxiliary and Control Buildings were supposed to be constructed of concrete block, reinforced and grouted. In the early 1980's inspection of the walls revealed all to be defective - some had rebar but no grout, some had grout but no rebar, some had nothing at all.

Metals

See 15.5.6 for a discussion of metals and their failure characteristics.

Wood and its Composites

Today, wood and its composites (plywood, pressed board, particle board) are mainly used for "low-tech" applications, where ease of cutting or shaping is important. Wood is, however, often used in combination with other materials to achieve strength or stiffness design goals. For example, balsa and plywood are frequently used as the core of a fiberglass "sandwich" for boat hulls and decks. Wood is subject to the following failures:

Rot - Since wood is an organic material, it is considered food by many organisms. Molds and fungi can consume wood, especially if the wood is in a damp, warm location. Wood used below the waterline of boats, or buried in the ground (utility poles, other structures) is susceptible to this failure. Even though utility poles are treated (creosote or pressure-treated), *butt rot* is one of the most common failure modes.

Insect Attack - Termites, toredo worms and other insects also treat wood as a source of nourishment and can lead to its failure in service.

Animal and Bird Attack - Since insects often make their homes in the natural or weather cracks of wooden structures, woodpeckers will damage wood as they seek their food. This is another leading cause of utility pole failures.

Delamination - Plywood and other wood composites can delaminate or debond. Indoor-rated plywood's laminating glue is not waterproof; exposure to the elements will quickly result in delamination of the plies.

Structural Overload - As with all other materials, wood has strength limits that can be exceeded. With the many different varieties of wood available, there is wide variation in tensile and compressive strength of wood. Sapwood is inferior to wood from the heart of the tree; dry wood is superior to wet wood. In the "old days" of wooden boat building, a significant inventory of oak, mahogany, pine, etc. would be seen in the yards drying, often for up to 10 years before it was deemed

ready to be used. Today, kiln drying is the substitute for this longer process, especially for the less-expensive grades of wood used in construction.

Other Composites

Thirty years ago, virtually all tennis racket frames were made of wood. Today, carbon composite or *graphite* rackets are common. Pleasure and workboats previously made of wood are now constructed of *glass-reinforced plastic*, more commonly known as *fiberglass.* Today, *Kevlar* is replacing fiberglass in high-load areas of hulls. These materials are all classed as *composites,* since they consist generally of a *matrix* (epoxy, polyester resin, metals and carbon, for example) combined with a *reinforcement* (fibers, whiskers, flakes, and particles, for example). The matrix spaces and protects the reinforcements from the environment and, most importantly, transfers the load to the reinforcements. The reinforcement provides strength and stiffness to the structure. In addition to the sport applications described above, composites have found their way into a wide variety of “high-tech” structures, such as helicopter blades, aircraft wings and fuselages, automobile body panels and medical prostheses.

Composites fail in the same modes as other materials (tensile, compressive, shear, etc.); however, their *directional strength* properties will result in different failure characteristics.

The *lay-up schedule* for the composite (number, orientation and sequence of reinforcement layers - determined in the design of the structure) in effect customizes the strength and stiffness properties of the structure.

For example, a lay-up schedule that orients all reinforcements along one direction will produce a structure with very little strength to resist loads perpendicular to the reinforcements. Adding reinforcements perpendicular and at various angles to the original layer will radically change strength and stiffness characteristics. The choice of matrix will impact the chemical and temperature-related failure behavior. *Thermosetting* matrices are stiffer (and hence, more brittle) and will resist temperature changes better than the *thermoplastic* matrices. These latter are more ductile, however they will begin to flow at elevated temperatures.

Composites respond to fatigue cycles differently than metals. Crack initiation and growth through metals to a critical crack size has been modeled successfully through fracture mechanics analyses for years. The failure process for composites is likened to that of a rope - failure does not occur when any one strand or fiber is broken, but only occurs when there is damage to many strands in a localized area. The ability to predict failure as a function of fatigue cycles is less well understood, however, composites have been shown to be more tolerant of fatigue than metals, resulting in a longer life.

Helicopter blades when fabricated from metal have fatigue lives measured in the tens of hours. Composite blades will essentially outlive the helicopter.

Failures of composites due to manufacturing problems occur. For fiberglass boats, two common failures are related to the lay-up process. In hand lay-up work, the fiberglass is first dipped or brushed with resin, then set in place in the mold and "squeegeed" in place. As the fiberglass are applied, it is important to remove air bubbles from between the layers. If this does not occur, these form voids that become delamination crack sites. The resin to fiberglass ratio is important to control. Since resin is expensive, there is a tendency to go light on this ratio. Inadequate impregnation of the fiberglass can result, which significantly reduces the strength of the composite.

Design problems are also common in composites. Powerboats designed for planing experience high hull stresses in wave conditions. Several years ago, one manufacturer's boats were experiencing stress cracks in the bottom portion of their V-Hulls. The hulls were flexing severely even in moderate wave conditions, producing cracks parallel to the keels. The design of these large, flat panels did not provide adequate stiffness against the wave stresses.

Civil/Structural Failures – Diagnostic & Analysis Methods

Virtually all the equipment discussed in the electrical/mechanical discussion will find a use in civil/structural failure diagnosis. *Ultrasonic* detection equipment is particularly useful for locating voids in concrete.

Core Sampling - Concrete core samples can be the equivalent of "parts" when performing a civil/structural failure diagnosis. Material properties (chemical composition, compressive and tensile strength) can be determined and compared to design specifications.

15.6.8 Other Failure Diagnostic/Analysis Methods

Thermography makes use of the *differences* in temperature between bodies to detect both normal and abnormal conditions. Just as a normal photograph shows visible wavelength differences as colors, a *thermograph* will show different *infrared* wavelengths as colors. For example, a current carrying resistor will dissipate energy as heat and will appear red, compared to the cooler printed circuit board that will appear blue. Thermography is very useful when attempting to identify a failed or degraded component. High-resistance connectors, shorts and opens can be detected through thermography. If a baseline thermograph is established when the system is operating properly, comparison thermographs can help pinpoint the failed component. *Video thermography* may prove to be a good investment if the system is to be used for both equipment *condition monitoring* as well as failure analysis.

X-Ray analysis is widely used to determine the interior structure and condition of components, parts and materials. In a failure analysis of a molded case circuit breaker, one of the first steps employed in the failure analysis was to X-Ray the interior to determine if any damage was present. If your laboratory cannot afford to purchase an X-Ray machine, remember that virtually every hospital is equipped with medical X-Ray machines. The radiology department may be persuaded to take X-Ray images of your component.

Neutron Radiography - While X-Rays are good for producing images of materials with high atomic numbers; they are not very good at developing images of low atomic number materials, such as plastics. “Slow” neutrons (room temperature neutrons of about 0.025 ev), though, are scattered easily by low atomic number materials and can be used to develop images of these materials.

15.6 Exercises

Objective: To ensure understanding of RAM concepts and definitions.

Instructions:
1. Here are some reliability situations and questions to answer. Your instructor will assign one or more of these to your table.
2. Use the definitions and concepts presented in this unit to help answer these questions.
3. Record your ideas on the flipcharts provided and be prepared to report your findings to the group:

Time: 30 minutes

1. How would you measure the reliability, maintainability and availability of your personal car?

2. In response to a regulator's concern about the reliability of safety system equipment, a power plant developed a new series of preventive maintenance procedures. Comment on this plan.

3. Recall some item that you purchased or product you helped design for which reliability is an important characteristic. What steps did you take to make sure the item was reliable "enough?"

4. A nuclear plant's three emergency containment coolers are *required* to operate for at least a year (8760 hours) after an accident. Due to failures of the fan bearings (about every 1000 hours), the mechanical maintenance department replaces the bearings every refueling outage (about every 18 months). Each fan runs about 600 hours between outages. No failures have been experienced for the last two refueling cycles. Should plant management be happy with this performance?

5. Divide your table into two groups - *customers* and *HVAC designers.* Come to agreement on the reliability requirements for the HVAC system:

    A. Multi-use commercial office building,
    B. Outpatient medical clinic with surgical facilities,
    C. New sports arena,
    D. Citrus processing plant with co-generating unit.
    E. Residential home located in Nashville, TN.

6. Microsoft has just updated Windows operating system with new AI features. Your applications now run much slower than before; it takes a minute to print a document that used to take 10 seconds. Has the computer failed?

7. What is the mission of a cell phone?

8. A 4160V AC switchgear bus is being installed in a factory located in south Texas near the coast. Due to cost overruns on the project, the air conditioning system originally designed for the switchgear rooms has been deferred one year. Temporary blowers have been set up instead. What impact might this decision have?

Objective: To practice developing an FMEA on a real-world system.

Instructions:

1. Perform an FMEA on the *Home Air Conditioning System* design (see the *Reliability Exercises Material* file). Make any assumptions you think are appropriate regarding its mission, environment, etc. Use the FMEA Excel spreadsheet or a flipchart. (Hint: Consider the functions necessary to perform the *system's* mission - preparing a functional block diagram may be helpful.)

   Subsystems:
   - Ductwork
   - Chiller/Air Handler
   - Electrical/Controls

   Summarize your analysis on the flipchart. What are the important failure modes you uncovered? What actions do you recommend? What did you learn because of this exercise?

Time: 40 minutes

Objective: To practice performing an EMEA on a real-world system.

Instructions:

1. Perform an EMEA on the *Home Air Conditioning System's* installation instructions (see the *Reliability Exercises Material* file). Make any assumptions you think are appropriate regarding the installer's ability to follow the instructions. Use the FMEA Excel spreadsheet or a flipchart. (Hint: Flowcharting the process/ segments may be a helpful first step in developing the EMEA).

   Process Segments:
   - Site Preparation
   - Foundation
   - Unit Installation (Mechanical, Electrical, Controls)
   - System Purge
   - System Charging/Testing

   Summarize your analysis on the flipchart provided. What are the important error modes you uncovered? What actions do you recommend? What did you learn because of this exercise?

Time: 40 minutes

Objective: To practice predicting the reliability of a complex system.

Instructions: 1. Develop a Fault Tree or Reliability Block Diagram of the *Home Air Conditioning System* for the failure event: *Failure to Provide Adequate Home Cooling* (see the *Reliability Exercises Material* file). Use the Fault Tree to quantify the failure *frequency* of the system. Express this frequency as a yearly rate (8760 hrs. per year). Some assumptions you may find helpful include:

- All dampers are required to be functional for system success.
- Any leak will cause sufficient loss of Freon to fail the system.
- The system operates at an average 10% duty cycle. Each operating cycle lasts about 20 minutes for a total of about 2600 start/stop cycles/year.
- Primarily, corrective maintenance is performed on the system. The air filter is changed 4 times a year (system operation can continue) and the fan motors are lubricated once a year (system is taken out of service for about 2 hours).
- Upon loss of electric power, the failure has to be detected (average time of 4 hours to detect). System restart takes 5 minutes.

Time: 60 minutes

Note: Failure rate information you may find helpful is listed on the table on the following page:

Component	Failure Mode	Failure Rate
Motor Compressor	Leak Fail to Start Fail to Run	1E-5/hr. 1E-4/d 1E-5/hr.
Heat Exchangers	Leak Plug	1E-6/hr. 3E-7/hr.
Expansion Valve	Plug	3E-6/hr.
Service Valves	Plug	3E-6/hr.
Freon Filter/Drier	Plug	3E-6/hr.
Freon Lines (Field Inst'd.)	Leak Plug	7E-6/hr. 3E-7/hr.
Fan Motor	Fail to Start Fail to Run	6E-5/d 3E-6/hr.
Dampers	Plug	1E-7/hr.
Ductwork	Collapse Leak	3E-7/hr. 1E-6/hr.
Air Filter	Plug	5E-3/hr.
Contactor	Fail to Operate	5E-5/d.
Relays	Fail to Operate	5E-5/d.
High Pressure Switch	Spurious Open	3E-6/hr.
Capacitors	Short Open	6E-6/hr. 3E-6/hr.
24V Transformer	Short Open	3E-6/hr. 1E-6/hr.
Thermostat	Fail to Operate	2E-5/d.
Disconnect Switch	Open	3E-7/hr.
Fuses/Main Panel	Open	3E-7/hr.
Wiring - Power	Short Open	6E-6/hr. 3E-6/hr.
Wiring - Control	Short Open	4E-6/hr. 2E-6/hr.
Utility Supply	Loss of Supply	15/year

Objective: To understand and build skills in basic reliability calculations methods.

Instructions: 1. Here are some more reliability calculation exercises. Try these to test your understanding of the RAM calculation methods.

1. An electronic device used in an air conditioning control circuit has a constant failure rate λ = 0.00125 failures per hour. What is its MTBF? What is its reliability for a 10- hour mission? For a 100-hour mission?

2. A compressor requires an average of 4 hours to repair. What MTBF is needed if steady-state availability is to be at least 0.99?

3. A sample of five constant hazard rate devices is tested without replacement until the fourth failure, at which time the test is terminated. Times to failure are t_1 = 800 hours, t_2 = 1800 hours, t_3 = 2125 hours, and t_4 = 2812 hours. The time on the fifth test unit is t_5 = 2812 hours. Make a point estimate of reliability for a 100-hour mission for this device. Find a 95% upper confidence limit on the failure rate and a 95% lower confidence limit on reliability for this same mission.

4. A manufacturer claims that his device has a mean time to failure of 15000 hours. His maintenance manual requires semi-annual lubrication and monthly filter changes. Comments?

5. For the following descriptions, determine which probability distribution would most likely fit the situation:

- An Instrumentation and Controls supervisor is testing integrated circuits to see if they meet company standards for use. Each time they are delivered, he selects 10 and checks to see how many fail an inspection test.
- The number of interruptions to the transmission system is measured each month.
- Samples of 5 transformers are gathered each month and the average core loss is measured for each sample.
- A mechanical coupling is observed to behave with a constant hazard rate.
- Tube leaks in a Component Cooling Water heat exchanger were found to be the result of a fatigue failure mechanism.

6. A turbine-driven compressor has exhibited a starting reliability of 0.97 for the last two years of operation. Management asks you to predict how many starting failures are expected this year if the compressor is to be started 60 times.

7. A company block purchased replacement Volt-Ohm-Meters for electrical troubleshooting. The battery packs on these devices have been failing, and the electrical maintenance department has been complaining. Here is the failure data collected by the electrical maintenance supervisor:

Time to Failure (Months)	**Number of Failures**
0 - 3	21
3 - 6	10
6 - 9	7
9 - 12	9
12 - 15	2
15 - 18	1
Total	50

Develop a histogram of the failures. What distribution seems to best fit this data? Calculate the associated distribution parameters. What is this distribution telling you about the nature of the failures?

The problem you addressed above only included *times to failure.* How would your conclusions vary if the following information were included regarding VOM's that had *not* failed?

Operating Time (Months)	**Number of Meters**
0 - 3	2
3 - 6	4
6 - 9	16
9 - 12	12
12 - 15	9
15 - 18	6
18- 22	4

8. The following data was obtained on the times to "failure" of a Cooling Water Basket Strainer ("failure" occurs when the differential pressure across the basket strainer exceeds 3 psi):

Time of Failure (Days)	Time Repair Complete
12.8	13.0
14.2	14.8
25.4	25.8
31.4	33.3
35.3	35.6
56.4	57.3
62.7	63.0
131.2	134.9
146.7	150.0
177.0	177.4

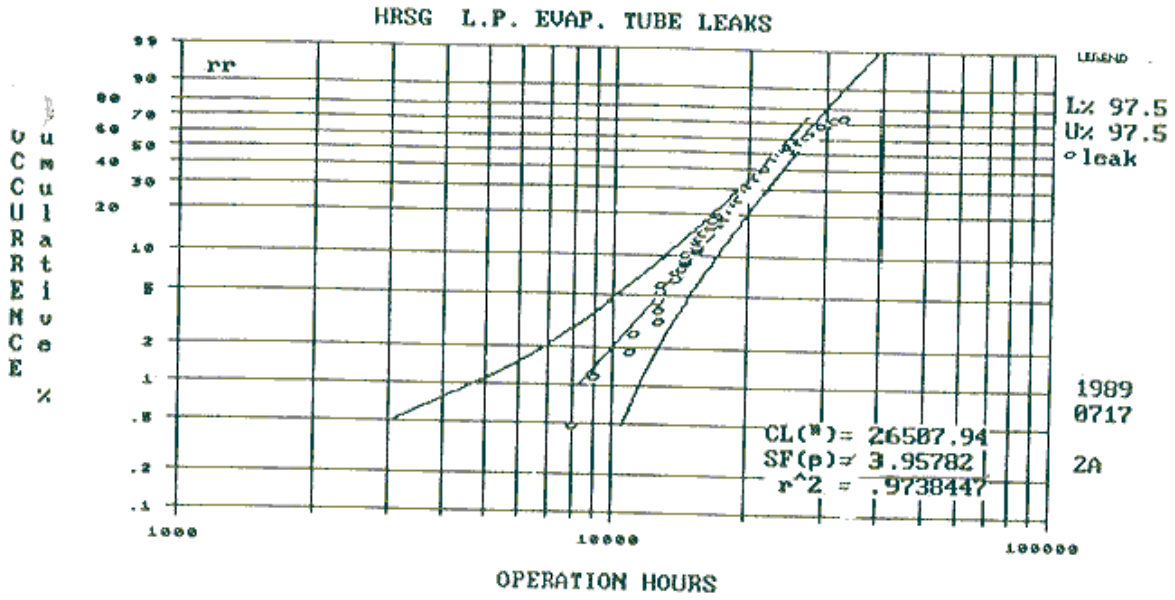
Calculate the MTBF, MTTF, MDT and Availability. Comment on the use of *mean values* as characterizations of the basket strainer's RAM. What does the dispersion of the data suggest to you? Why might this dispersion exist? Perform a Weibull Analysis of these times to failure. What recommendations would you make as a result of this analysis?

9. Locomotives are intended to haul freight and passenger trains from one destination to another. Between trips, they may be stored in a yard or maintained. How would you measure their reliability? How would you measure their availability?

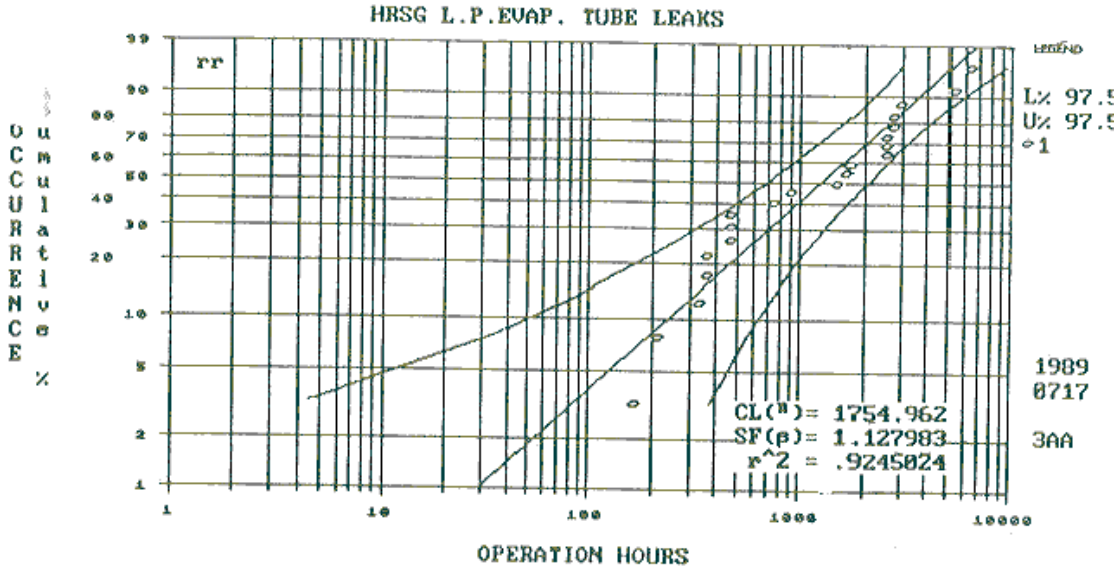
10. How do you define and measure reliability and availability for your products and/or systems? Obtain data for at least one product or system. How does it compare to company or customer targets?

11. The following Weibull Plots were obtained from an analysis of tube leaks occurring in a Low Pressure Evaporator of a combined cycle power plant. Interpret these plots. What can you learn about the failures from these charts?

Tube Failures Occurring on 5 Outer Rows of the Evaporator (at Bend Area):



Tube Failures Occurring Across the Evaporator (at Bend Area):



12. The following valve repair data was collected. Note that in some cases, upon overhaul, the valve was found to be in a failed state, in others, the valve was still operable. Perform a Weibull Analysis on this data. What can you learn from this analysis? If the valve's mission time is 18 months, what are the chances of it meeting that mission without failure (i.e. what is the reliability of the valve)?

Time Between Repairs (Months)	Valve Condition
47	OK
23	Failed
10	Failed
11	OK
28	OK
32	Failed
4	OK
13	Failed
17	Failed
47	Failed
25	Failed
5	OK
20	Failed
18	Failed
27	OK
36	Failed

13. Take 7 – 10 paper clips and straighten them. Grasping them with both hands, bend them back and forth. Count the number of cycles (one bend is a cycle) until the wire breaks. Perform and interpret a Weibull Analysis of this data.

14. The following failure data was collected for a new design of turbine blade to be applied in an advanced liquid fuel rocket engine. Perform and interpret a Weibull Analysis of the data. Does the Weibull appear to fit this data well? Try the lognormal distribution as an alternative.

Time to Failure (x 10^{-3} sec)	Failure Type (1 = censored)		Time to Failure (x 10^{-3} sec)	Failure Type (1 = censored)
10.1348	2		10.3482	2
5.9175	2		14.4255	2
7.0802	2		9.1499	2
17.7740	2		10.3604	2
14.2184	2		19.1694	2
7.8660	2		12.2707	2
9.7598	2		17.0686	2
8.8241	2		20.0000	1
14.1468	2		20.0000	1
11.5319	2		20.0000	1
12.6105	2		20.0000	1
12.6800	2		20.0000	1
17.9078	2		20.0000	1

15. A bearing manufacturer has run samples to failure at loads of 7000 lb. and 30,000 lb. The life test data is shown below. An order has come in for this bearing with a specified load of 15000 lb. Estimate the B10 life of the bearing at this load.

Times to Failure (Hours)	
7000 lb.	**30000 lb.**
1910	650
2450	905
2940	1150
3400	1390
4000	1760
4700	2180

16. Three components in series comprise a system. For a 10,000-hour mission, their actual (field) reliabilities are:

- Component A – 0.996
- Component B – 0.983
- Component C – 0.960

A new client is ordering "mass quantities" of the system; however, her specifications call for a system reliability of 0.980. How would you allocate the component reliabilities to meet the new target?

17. Accelerometer devices are installed on shipping containers containing your company's LED screens. These accelerometers record **peak** acceleration and are used to monitor for possible shipping damage. The readings on 10 successive shipments appear below:

Shipment	Max. Accel. (G)		Shipment	Max. Accel. (G)
1	8.6		6	10.2
2	7.3		7	8.1
3	5.8		8	6.6
4	7.0		9	7.7
5	6.2		10	9.2

The shipping container is designed to withstand 12 g's without damage. If an accelerometer reads more than 12 g's, the entire box of LED screens is returned. You ship 10,000 players a month (10 per box). How many would you expect to be returned each month because of excessive shock loads?

18. A reliability test was performed for bearings at loads L1 and L2. Ten bearings were tested at each load with the following times to failure (hours):

- L1 – 1936, 5125, 1650, 3170, 810, 3780, 1108, 2580, 1395, 2130
- L2 – 2018, 610, 885, 2860, 1202, 440, 1048, 1428, 750, 1706

Estimate the B-1 and B-10 life of the bearing at each load.

Objective: To practice uncovering the potential root causes of a system failure.

Instructions:

1. The owner of the house has had trouble with the air conditioning system (see the *Reliability Exercise Material* file). About every two - three weeks during the summer, one of the fuses on the main panel blows, causing loss of power to and failure of the system. Develop a Fault Tree of this system failure and a series of verification tests to determine the ***root cause***(s) of the problem. Consider both "usual" and "unusual" faults.

   Summarize your analysis on the flipchart provided. What are the important faults you uncovered? Did you learn anything from this analysis that was not apparent from the FMEA or EMEA? What did you learn because of this exercise?

Time: 45 minutes

16.0 Planning & Feedback Tools

Unit	*Description*	*Page*
16.1	Seven Planning Tools	16.1 - 1
16.2	Operating Reviews	16.2 - 1
16.3	Exercises	16.2 - 1

Many of the sections in this many deal with tools to analyze *data*. Quite a few planning type problems involve *ideas*. The Seven Planning Tools help you organize ideas from several viewpoints. Operating reviews are an essential part of any business feedback system. The types and conduct of reviews are described.

16.1 Seven Planning Tools

Learning Objectives

- To understand and apply a number of “language” based planning tools

Unit Contents

- Affinity Diagram
- Arrow (PERT) Diagram
- Idea Matrix
- Matrix Data Analysis
- Process Decision Program Chart
- Relations Diagram
- Structure Tree

Introduction to the Planning Tools

Much of this manual focuses on data analysis tools. In many planning applications, though, the use of data is limited, although *ideas* are important. The following seven tools represent a collection of ways you can organize and relate ideas to each other.

You have already encountered applications of these tools. For example, Unit 4.1 describes the use of the affinity and structure trees to help organize many customer needs. Unit 14.1 presented Quality Function Deployment, which makes use of the Idea Matrix.

Affinity Diagram

Purpose

The affinity diagram organizes seemingly unrelated ideas. For example, affinity sorting can be used to take customer inputs (needs and expectations) and organize them into groupings as input to a House of Quality.

Construction

1. Determine the purpose of the Affinity session. For example, a group of senior managers conducted an Affinity session to identify possible major corporate initiatives as part of their strategic planning effort. A team working to plan the implementation of a new information system conducted an Affinity session to identify and group the barriers and aids to implementation.

2. Brainstorm a list of ideas relating to the purpose. Record these on note cards or “sticky” notes.

3. As a group, silently take the cards or notes and begin to group them into related categories (use a table, wall or other large flat surface). It’s OK to take a card and move it from one group to another. If this happens several times, include the idea in both groups.

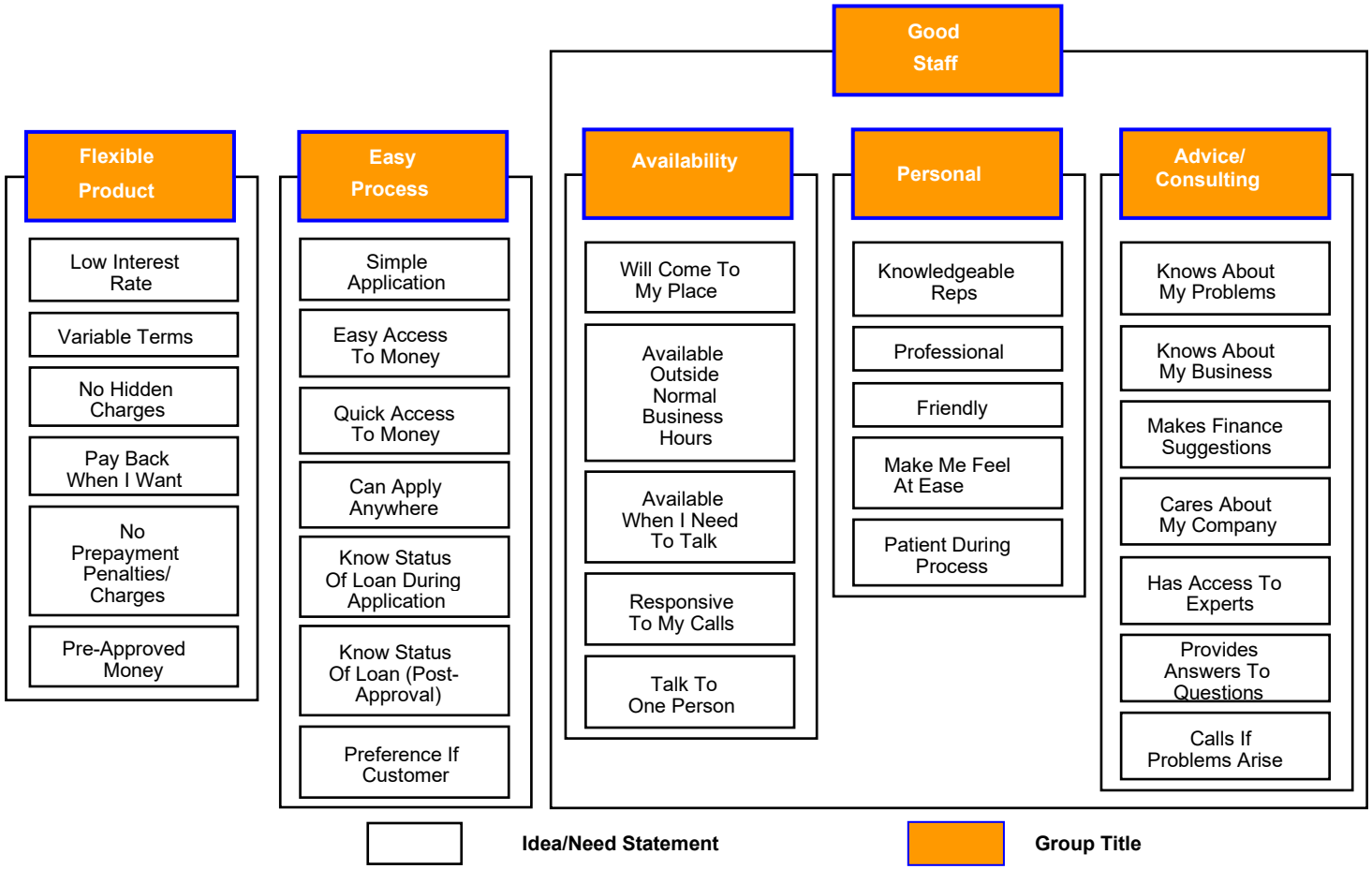
4. Step back and review the groupings. Determine a heading/title for each of the major groups identified. In the strategic planning example, these headings became the major corporate initiatives.

Uses

Earlier, we described various methods of obtaining the customer information that will be input into the design process. Interviews, focus groups, surveys and customer observations may be conducted, and research into how customers use products and services performed.

This can result in a large body of ideas that is difficult to pass "unfiltered" to the design team. The affinity process can take these seemingly unrelated ideas and, without losing the detailed content important for the design work, organize them into "natural" groupings. These natural groupings may relate to the key functions to be performed by the new or redesigned product or service.

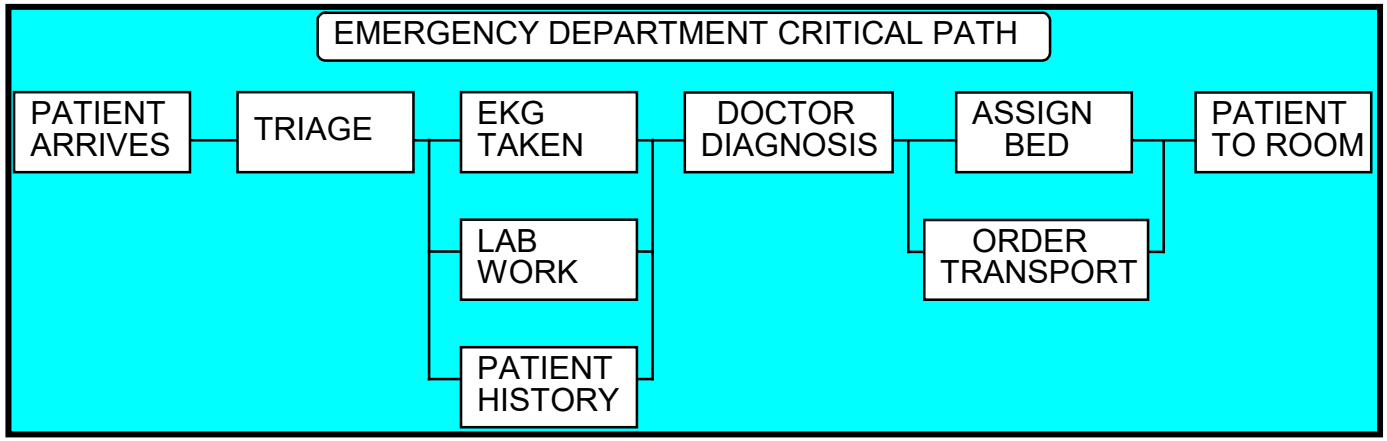
Example Affinity Diagram – Loan Product



Arrow (PERT) Diagram

Purpose – The Arrow diagram helps identify the time-sequence of events necessary to achieve an objective. When used in project management the Arrow diagram helps understand the specific sequence of events that drives the time required to achieve the objective and to manage the implementation of a project requiring multiple activities.

This analysis has its origins in engineering and construction project management (PERT stands for *Program Evaluation Review Technique*) but has been applied widely. The *Critical Pathway* method of identifying and standardizing patient, hospital and doctor activities draws heavily on this concept.



Construction

1. Identify the objective to be achieved.

2. Identify the activities that must be performed to achieve the objective. Write these on note cards.

3. Arrange the activities in time sequence, beginning with the **start** of the project until the completion or **end.** Draw lines between the tasks to indicate their relationships. Be aware of the types of activity relationships (*it may be helpful to label the relationships)*:

Finish to Start (FS) - Here one activity cannot begin until the preceding activity is finished. For example, triage cannot begin until the patient arrives.

Finish to Finish (FF) - Here, one activity cannot finish until a preceding activity is also finished. In the example shown above, the Doctor cannot *finish* his/her diagnosis until the EKG, Lab Work and Patient History steps are completed.

These are the most common relationships that constrain the completion time for a project. In some cases, though, the following relationship is important as a constraint:

Start to Start (SS) - Here, one activity cannot start until a preceding activity has also started. For example, in the Design Process, the *Design Product/Service* phase cannot begin until the *Determine Customer/Product Requirements* phase has begun.

4. Some measures of the resources (time, materials, manpower) required to perform the individual activities are assigned to each box representing that activity.

Uses

At this point, various analyses of the *network* created above may be performed. For instance, the network may be analyzed to identify opportunities to reduce the overall time required to accomplish the objective. The network may be examined from the resource perspective to identify and optimize the demand on available resources (either manpower or cash flow or both).

Once the network is created, a Critical Path is often calculated, this is the **longest** time that will be required to complete the project. To shorten the time required for the process, the Critical Path shows the important steps of the process to analyze.

The Critical Path is often monitored as part of project management activities. As activities are completed, a periodic update of the Critical Path is performed to predict the remaining project completion time and resources. Activities "on the critical path" are reported in update meetings, if problems are anticipated with their completion, extra resources or attention will be devoted to these.

Idea Matrix

Purpose

The Idea Matrix shows the cause-and-effect relationships or correlations between a set of factors and a set of effects.

Construction

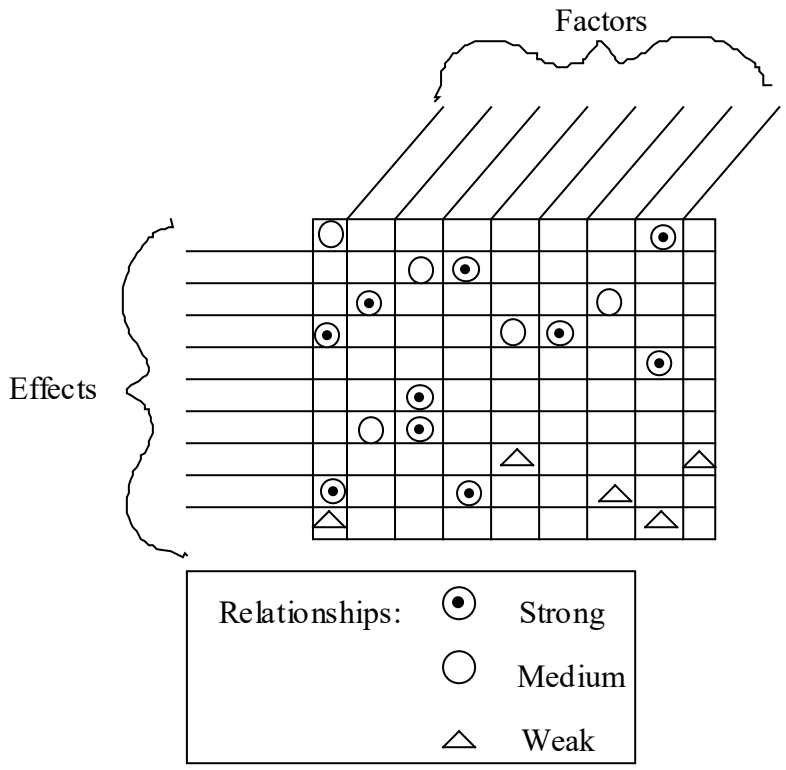
1. Identify the effects in which you are interested. For instance, the customer needs/expectations may be the effects (note that these may be identified through a structure tree analysis).

2. Develop hypotheses regarding the factors that will result in these effects. For instance, the product or service quality characteristics will be the factors that will result in the customer needs/expectations being satisfied (again, a structure tree may be used to start with high level factors and break them down into lower-level characteristics.

3. Write the effects in a column on the left side of the paper. Write the factors across the top of the paper. Draw a matrix that allows each factor to be evaluated against each effect:

4. Evaluate the cause-and-effect relationship between the factors and the effects. The symbols shown above may be used to summarize this relationship, or numerical scores (5 - strong, 1 weak or none) may be employed.

5. Based on the results of the analysis, act. In the example above, some product characteristics that we wish to incorporate may be found to have little relationship with customer needs. These characteristics may be eliminated. Some customer needs may not be "covered" by product or



service characteristics. Additional development of the product or service is necessary here.

Uses

The idea matrix has many uses in product or service design. This matrix is the foundation of *quality function deployment.* Besides the "square" matrix shown above, there are "T," "L," "Y," and many other configurations.

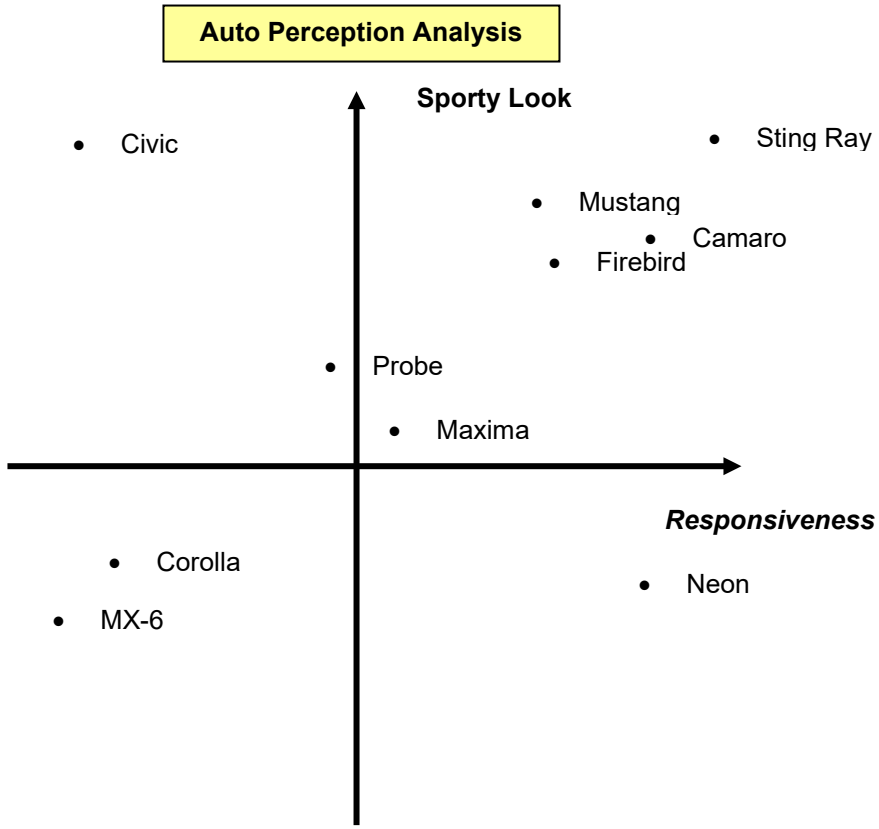
Customer Needs to Product or Service Characteristics - The example in the construction steps described this use. Here, the cause-and-effect relationship explored helps translate the needs of the customer into the quality characteristics that will be designed into the product or service.

Product/Service Characteristics to Product/Service Features - The next cause and effect relationship that needs to be explored is that between the *quality characteristics* and the *methods* of achieving these characteristics (product or service *features*). The idea matrix is useful here when the methods relate to multiple characteristics. Otherwise, the structure tree may be a simpler approach.

Product/Service Features to Process Characteristics - Here, the cause-and-effect relationship being explored is what characteristics of the *production process* are responsible for the product/service's features. This cause-and-effect relationship helps establish the *control points* of the production process, i.e. what are the important process variables that must be managed to ensure the quality of the product or service?

Matrix Data Analysis Chart

Purpose - To portray correlations between variables, using an X-Y plot. In Quality Function Deployment, this chart is often used to analyze market segments, determine what markets a company should address and how they stack up against the competition. We saw a similar application in *Parametric Analysis* (see Unit 14.2, Conceptual Design).



Construction

1. Identify the question to be addressed. Some examples appear below:

- Are there important customer segments based on product characteristics?
- How do customers perceive the competition's products meet their needs?

2. Gather data and prepare the matrix chart.

3. Analyze the information. Look for correlations, groupings, and clusters that may help you understand the relationships.

Uses

Matrix Data Analysis is used in market research studies (e.g. the segmentation analysis discussed above), and in defining product/service requirements.

Process Decision Program Chart (PDPC)

Purpose - The PDP chart helps you identify what can go wrong or obstacles in the performance of a process. This chart is like Fault Tree Analysis (See Unit 15.3); however, it is simpler and adds a time-dimension to the analysis. The PDPC is a preventive tool; in essence, it allows the team to perform a dry run of a process.

One caution is in order – the PDPC development can "explode" – in many cases there are many possible alternative paths that can occur in a process. The team developing the chart must learn to manage this potential.

Construction

1. Identify the objective (or undesirable event in the case of using PDPC to prevent failure or liability issues).

2. Develop the structure of the objective – this is like developing a Structure Tree that breaks down the overall objective into pieces.

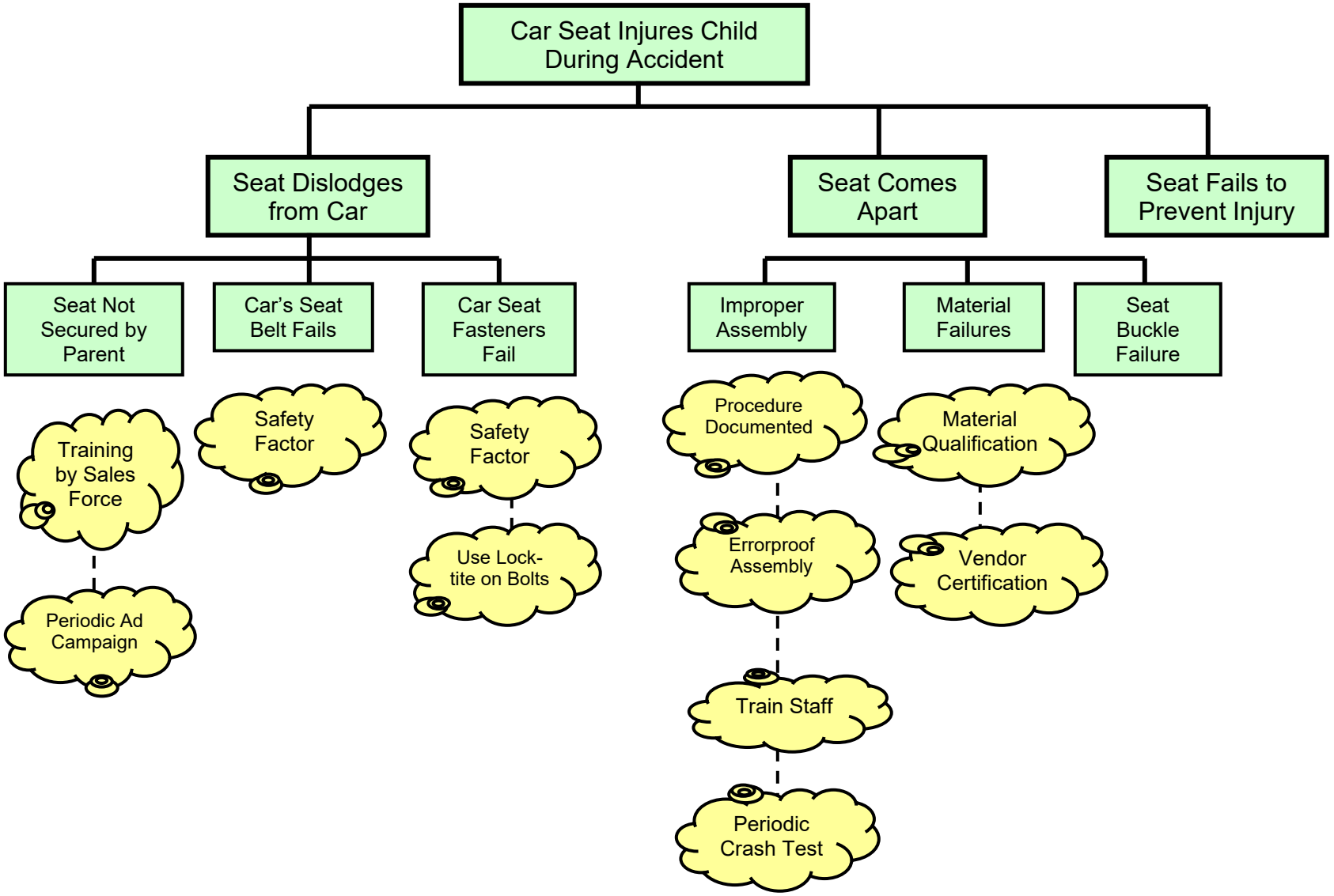
3. Take one of the lower-level pieces. Ask, "what could go wrong here" or "what other path could this process take?" Identify countermeasures that will address the failures. Write these down and circle them like clouds on the chart.

4. Continue this process until all or at least the major failure points have been addressed.

Uses

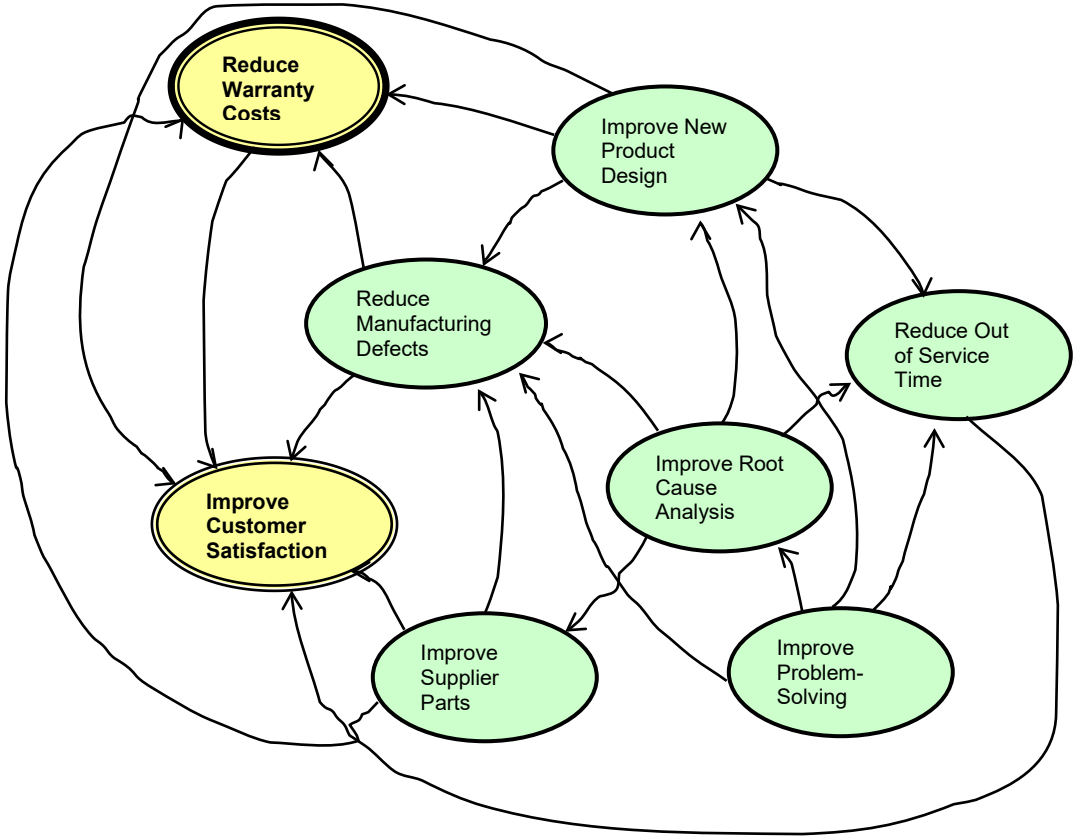
PDPC is popular in reliability work, or safety analysis activities. See the next page for an example PDPC diagram.

Example PDPC Diagram



Relations Diagram

Purpose – The Relations Diagram helps you understand the relations between ideas. The basic question asked is "which idea influences (or drives) another idea). Often used to find strategic drivers – what are the key issues that an organization should focus on to achieve several objectives. The Relations Diagram may be used in combination with the affinity diagram. For example, a husband and wife identified several "issues" that they were having trouble dealing with. They affinitized the issues and then took the themes and developed a relations diagram. This helped them see how the issues related to each other and to identify ways to improve their relationship.



Construction

1. Identify one or two key issues to be analyzed.

2. Record the idea(s) on sticky notes and place them on a flipchart or board.

3. Identify/brainstorm related issues or problems. Record these on notes and place them on the flipchart/board.

4. For each idea, consider how it relates to all other ideas. For ideas where a relationship exists, draw a line between the ideas. Draw an arrowhead indicating the direction of the relationship (e.g. causality).

5. Examine the completed diagram. Look for *key ideas.* These are the ones that have many arrows leading away from them and few coming into them.

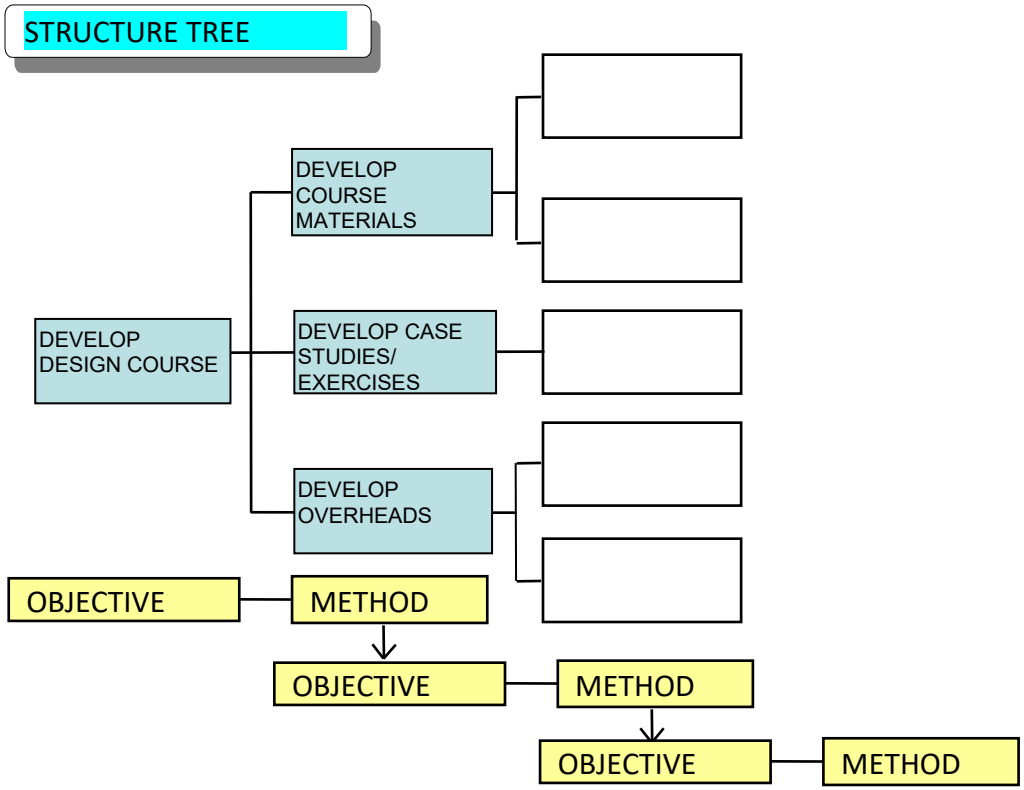
Uses

Relations Diagrams are useful when there are several ideas whose relationships are not obvious. Strategic planning workouts often make use of relations diagrams – the key issues the company is facing are identified, then possible strategies to address the issues developed. The relations diagram approach helps identify the key driving strategies that will enable the company to focus on resolving their key issues.

Structure Tree

Purpose

The structure tree shows the relationship between some item and its elements. The tree helps breakdown an objective in terms of the methods required to achieve the objective.



Construction

Note: This analysis can be done on a sheet of paper, or, for large group development, several flipchart pages may be taped together and "sticky" notes used.

1. Determine the objective that is to be achieved. State this in the form of a sentence that is understood by everybody participating. Write this in a box on the left side of the paper.

2. Develop the "high-level" means of achieving the objective. For instance, if you wish to automate a particular process, the high-level means will include computer hardware, software and operators.

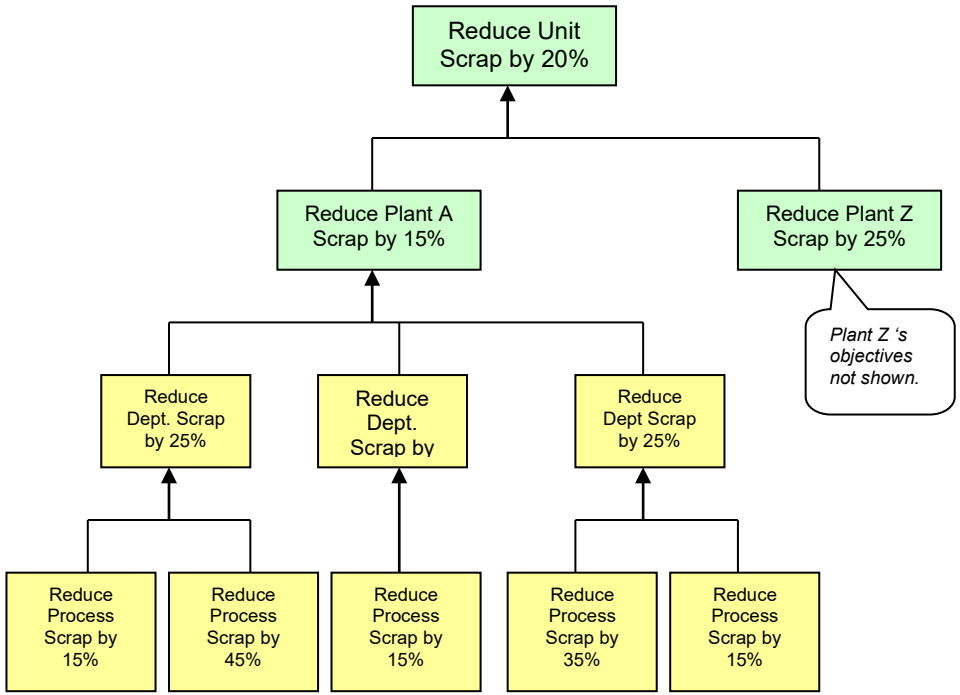
3. Continue the breakdown until specific means have been discovered to achieve the lower-level objectives. For instance, identifying the specific solid-state drive (manufacture and model number) as the *method* to achieve a required mass storage *objective* would complete the breakdown.

Uses

The structure tree is a very general diagram. It has many different uses and is often combined with other analysis tools:

Cause and Effect - The structure tree may be thought of as an organized cause and effect diagram. Here, the purpose would be to identify hypotheses regarding which elements of the structure tree were key factors affecting the cause. Methods of verifying the hypotheses should then be attached to the factors.

Objective/Methods - The construction steps listed above are based on this use. If we want to break down a high-level objective into its components, the structure tree is an excellent tool. For some "R&D" efforts, the low-level methods may be a mixture of methods that we know can be accomplished with methods that we are not sure are practical or achievable. Here, the structure tree's elements should be marked accordingly and a plan developed to



determine the "achievability" of the latter elements.

Functional Analysis - In product or service design, we can think in terms of the *functions* that need to be performed by the product or service. These functions, in turn, can be broken down into sub-functions, etc. For instance, a system may need to perform various functions, these, in turn, are provided by subsystems, then by equipment, components and, finally parts. The structure tree can be used to develop and display these relationships.

16.2 Operating Reviews

Learning Objectives

- Participate in and lead effective Operating Reviews.
- Provide feedback to others leading and participating in Operating Reviews.
- Recognize various levels and types of Operating Reviews.
- Clarify the outcomes expected from three types of Operating Reviews.
- Conduct a system review

Unit Contents

- Operating Reviews
- Conducting a Review
- Four Types of Reviews

16.2.1 Operating Reviews

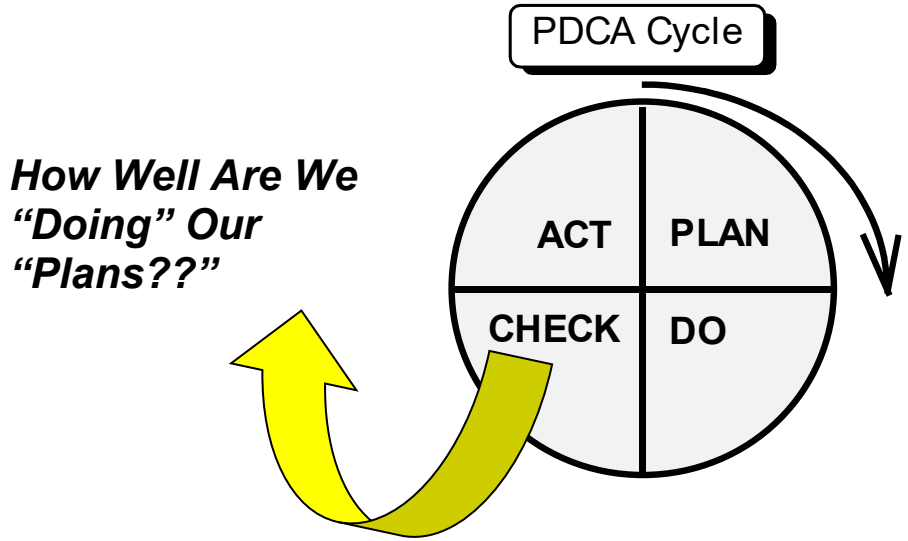
Reviews are one level of management seeking to determine how well supporting levels of management are executing an agreed upon plan. They also provide an opportunity to offer guidance and support.

Reviews are used to:

- Establish strategic direction (Strategic Review).
- Confirm alignment with Corporate-level Key Focus Areas (Operating Plan Review).
- Determine the extent to which the supporting levels are meeting their commitments relative to these KFAs – Department /Unit-level Key Focus Areas (Operating Plan Review).

Conducted by:

Reviews are a critical element in the achievement of stated goals, strategies, and key focus areas. They are conducted at every level of the organization starting with the CEO and help the organization rotate the PDCA wheel through the "Check" function.



Purpose and Benefits of a Review

A constructive Review is one where one level of leadership follows a line of questioning with others to:

- Accurately determine current progress,
- Identify strengths and weaknesses in actions already taken,
- Break down barriers and foster cross-functional interdependence,
- Provide guidance,
- Support or alter a course of action based on data, and
- Ensure future actions are tied to a schedule.

Characteristics of Operating Reviews

Operating Reviews can be viewed from different organizational levels.

Business Unit Level	Department/Functional Level
○ Drive Performance at all Business Unit Levels	○ Drive Specific Functional Performance Needs
○ Held at Least Monthly	○ Will Vary Between Function
○ Business Unit Leadership Drives Reviews	○ Numerous at this Level
○ Focus on Detailed Action Plans and Project Activities	○ Complement Other Reviews

Formal or Informal Operating Reviews?

The setting for an Operating Review may be formal or informal. The discussion may take place over lunch or in a conference room, *but the content of the Review is always formal in nature.*

Operating Reviews Sound Threatening

"Feedback is the breakfast of champions," says Dr. Ken Blanchard, co-author of the *One Minute Manager,* and feedback is at the heart of every Operating Review. Giving and receiving constructive feedback requires that all participants treat each other with dignity and respect.

Destructive feedback occurs when a leader uses the Operating Review to vent frustration, play out personal agendas, exercise inappropriate power of position, or otherwise treat the other person in a manner that is demeaning or disrespectful.

This distinction is critical to understanding the Operating Review process. It can be the difference between success and failure.



Operating Reviews vs. Written Reports

Arguments made for traditional written monthly reports become less convincing as leaders demonstrate the power and effectiveness of face-to-face Reviews.

Operating Reviews:	Written Reports:
♦ Focus on the customer ♦ Are simple (charts tell the story) ♦ Focus on how performance gaps are eliminated ♦ Establish a supporting team environment ♦ Challenge the team to work on the real issues ♦ Search for customer assigned value ♦ Focus on results data ♦ Allow the presenters to easily clarify their positions ♦ Create a process for staying (or getting back on) track ♦ Focus on solving problems	♦ Perceive the reader as customer ♦ Must be “pretty” ♦ Focus on selling ideas ♦ Distance levels of management ♦ Frequently avoid dealing with the real issues ♦ Ask the reader to assign value ♦ Usually hide results data in verbiage ♦ Make it difficult to ask for and provide clarification ♦ Are geared toward a good grade from reader ♦ Focus on explaining problems

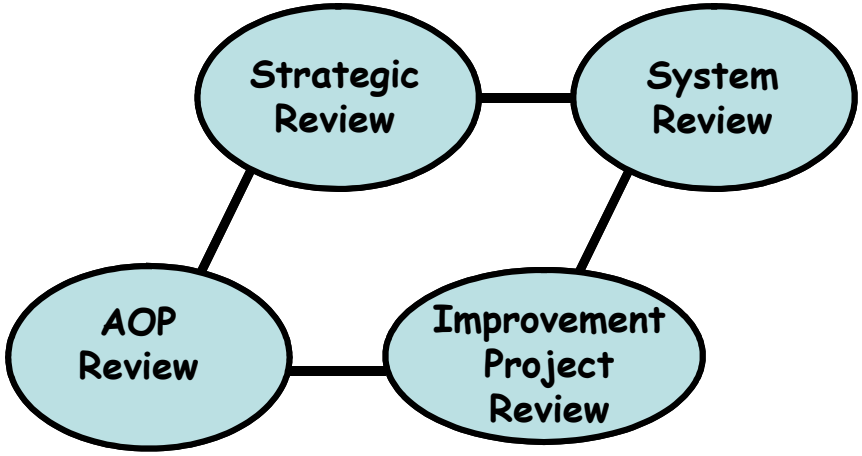
A Brief Story: One of our friends was an Operating Leader in a major US hospital, reporting to the Chief Operating Officer (COO). He required monthly department reports be submitted. Occasionally, she would slip a curious sentence into the middle of her report (i.e. a *mild* sentence might read: “Dr. Pierce and Nurse Houlihan were found drinking champagne and eating caviar in the rehabilitation department’s Jacuzzi three times last month.”). If she did *not* get a reaction from the COO, then she would not send her next month’s report. Inevitably, he would ask her where it was, and she would tell him that since he obviously did not read *last* month’s, it was a waste of her time to prepare *this* month’s!

16.2.2 Four Types of Operating Reviews

The Operating Review process is closely tied to a company's strategic or business planning process. The model presented here assumes that a Strategic Plan guides the company's long-term improvement efforts and that this is translated into an Annual Operating Plan (AOP) with goals and targets for the current year. Based on this model, four types of Reviews can be conducted:

- *Strategic Review* – assessment of overall organizational performance; customers suppliers, and markets; products and services; competitors; and economic/financial environment.
- *Operating Plan Reviews* – to assess alignment, performance, plus offer support and guidance for activities and projects that are part of the AOP.
- *Improvement Project Reviews* – to determine progress and offer support and guidance in applying tools and techniques to improve processes and solve problems.
- *System* Reviews – an evaluation of the effectiveness of the business planning process.

Types of Reviews



Involvement in Reviews

Involvement in Reviews is different for each type:

Type of Review	Who Conducts?	Frequency
Strategic Review	Corporate and Department/Unit leadership with staff support.	Monthly/ Quarterly (as appropriate)
Operating Plan Review	Every level of the organization beginning with the CEO.	Monthly
Improvement Project Review	Local management conducts these reviews. Other leaders may also choose to participate because of a mutual interest in an improvement project.	Monthly
System Review	All levels of management provide input on the effectiveness of the process. The unit president should summarize and develop plans to enhance the system based on the system review.	Quarterly

Strategic Review

The approach used for this review includes:

- Gather information about customers and markets, the competitive environment, organizational performance, Department/Unit capabilities, Supplier/Partner capabilities, and the financial/economic environment.
- Analyze this information.
- Identify strategic opportunities
- Confirm existing strategic opportunities – Key Focus Areas (KFAs)
- Evaluate the linkage and alignment of proposed KFAs

Reasons for a Strategic Review

Strategic Reviews define the direction your company will be taking during the "long-term."[1] It is the intent of this review to:

- Ensure you understand your customer requirements
- Analyze your operating environment
- Establish a common direction – Key Focus Areas
- Coordinate annual Key Focus Areas with long-term plans

[1] "Long-Term" is deliberately left vague. For some industries, such as an electric utility, long-term may translate to the next 10 – 15 years. For others, such as the microelectronics industry, "long-term" may be the next 10 – 15 months.

Annual Operating Plan Review

Annual Operating Plan (AOP) Reviews have a three-fold purpose:

- Focus on the *alignment of the organization* toward common **business objectives**, and *performance* against agreed upon targets and budget dollars.

- Provide **people** - everyone involved in the achievement of the stated plan - with ongoing performance feedback.

  Evaluate the *training* plan in place to ensure that each employee has the appropriate skills to support the future health of the organization. The strategy for providing employee *recognition* is also considered.

- Ensure that local management is implementing a **Local Leadership System** in place that is consistent with the following principles:

  - ✓ Customer-focused,
  - ✓ Data driven, and
  - ✓ Committed to continuous improvement.



Reasons for Annual Operating Plan Reviews

- **Leaders must be actively engaged in the process of getting results.** Reviews create the expectation that progress is being monitored. There is demonstrated interest in the achievement of stated plans. People are held accountable with a shared focus on continuous improvement to ensure success at the end of the effort rather than hoping things will turn out well.

- **Address customer requirements -** Reviews help us focus on the customer. Maintaining a customer perspective is sometimes difficult in the day-to-day operation of Department/Unit. Like someone has said*, "It's hard to remember*

your objective was to drain the swamp when you're up to your neck in alligators." There must be a shared and clear focus on what is critical to the success of the company; what will help maintain a competitive advantage.

- **Pay attention to the details -** Efforts to ensure success are not left to chance. They are the product of business basics like careful planning, comprehensive analysis, and project management.

- **Determine true causal factors -** Comprehensive analysis creates the potential for getting to the actual causes of a problem. Operating Plan Reviews encourage this approach and reinforce the practice, thus greatly increasing the probability that the correct improvements will be identified and implemented.

- **Promote ongoing commitment -** AOP Reviews maintain what W. Edwards Deming called, *"A constancy of purpose."* When conducted on a pre-scheduled basis, Reviews help coordinate and align activities toward a common purpose. The ability of an organization to accomplish improvements consistently is far more dramatic than a practice of sporadic interventions. Conducted on a monthly or quarterly basis, Operating Plan Reviews are the difference between one-time, stand-alone improvements and integrated cumulative improvements, which result in a dramatic change.

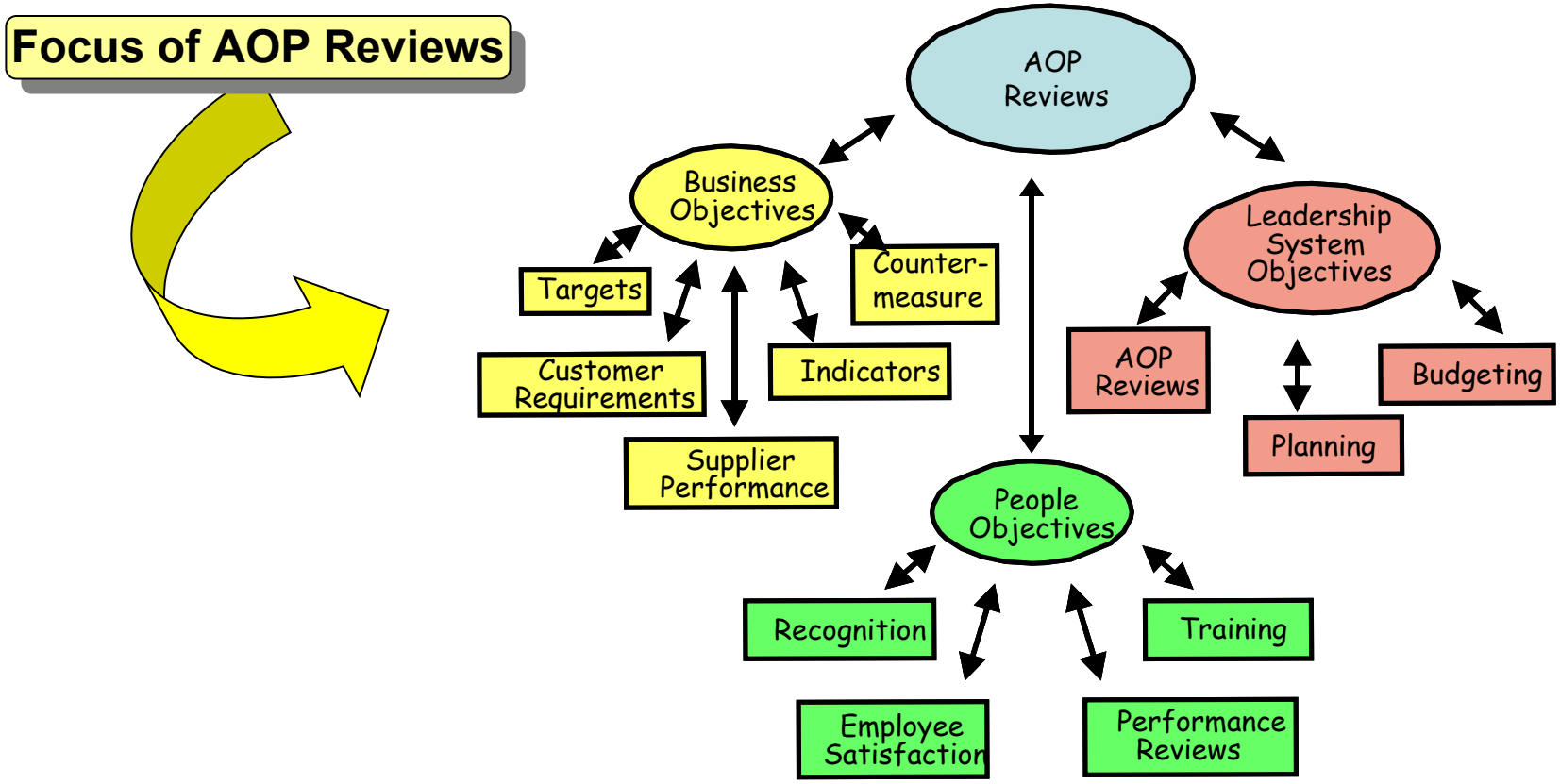
- **Champion continuous improvement -** During Reviews we ask, *"What can be? What is possible over time?"* This is fundamental to empowerment because it moves the perspective from hand wringing about the present to that of gaining control over what causes the present situation and improving them.

- **Focus on Work Processes -** Operating Plan Reviews focus attention on the work processes that must be improved to achieve the targets as stated in the business plan. Questions are asked such as, *"What are the processes? What is the performance of these processes? What is being done to improve them?"* This is a refreshing and powerful approach when contrasted with seeking out someone to blame. The spirit of the Review is to focus on processes, not people.

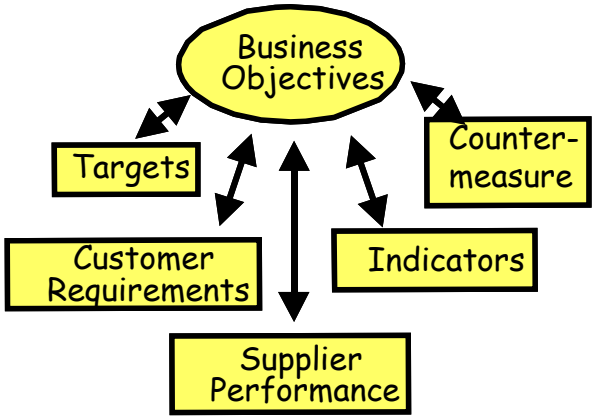
Questions to ask at AOP Reviews

The core objectives of an AOP Review are addressed with questions that:

- Ensure alignment with **business objectives**, including budget performance,
- Provide performance feedback to **people**, with particular attention to training and recognition, and
- Ensure that a **leadership system** is in place to maintain the ongoing health of the organization.



Typical Business-related Questions:



1. Do your plan and activities **link** with the company's Key Focus Areas?

2. Do the **indicators** make sense?

   - ❑ Is this the best measure of performance relative to the **customer requirements**?
   - ❑ Is data presented in such a way that progress is easily determined?
   - ❑ Is the **target** challenging and aggressive?
   - ❑ How was the target determined?
   - ❑ Is the benchmark such that it reflects "Premier?"

3. What is the current performance?

   - ❑ Is there a gap between current performance and the target?
   - ❑ What are root causes of the gap?

4. Is there a **countermeasure** to close identified gaps?

   - ❑ Are these projects the result of data driven analysis?
   - ❑ Are the projects described at a level of detail where someone can implement them?
   - ❑ Are the projects pushed down to the lowest possible and most appropriate level for implementation?
   - ❑ Is it clear who is accountable for implementation and on what schedule?
   - ❑ Are you confident that these actions will eliminate the gap between where you are and where you want to be?
   - ❑ When will the gap be eliminated?

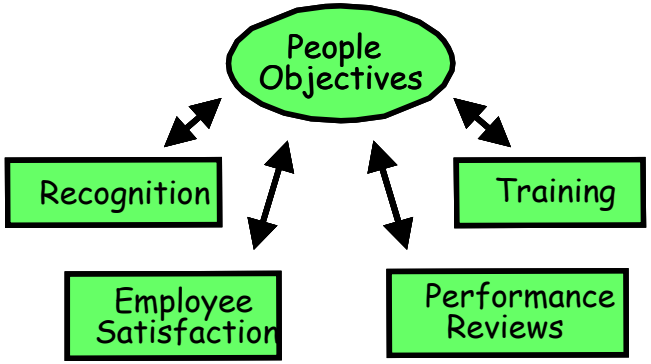
4. What steps have been taken due to not achieving the target?

   - ❑ Show me your analysis to determine the root cause of what is preventing you from reaching the target.
   - ❑ What countermeasures / action plans are you implementing to assure that you get back on track?
   - ❑ When do you expect to see improvement?
   - ❑ What changes has been made to assure that this doesn't happen again?

5. How are you ensuring that your **suppliers** are providing products and services that meet your specifications?

   - ❑ What has been done to improve supplier performance?

Typical "People" Questions asked During Operating Reviews:



1. Performance Reviews / LRP

- What is the schedule for Appraisals and the LRP?
- Show your approach for training managers and supervisors to conduct performance reviews and LRP.
- Any common patterns in the feedback from this process?

2. Training

- What skills/competences need to be improved? Why?
- What is your training plan? What are the assumptions that drive the training plan?
- What is the performance vs. target?
- What are some examples of differences that training has made?

3. Recognition

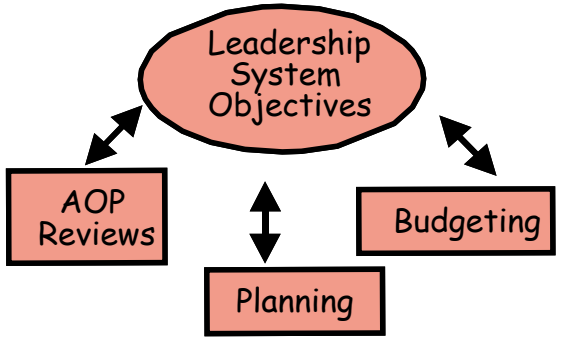
- What is your organization's plan for providing recognition?
- What are some examples of how recognition has been given?

4. Employee Satisfaction

- What are the three areas of greatest employee dissatisfaction?
- What is the target for improvement?
- Share the best example of an improvement in this area.

Typical Leadership System Questions asked during Operating Reviews:

The objective in this part of the Review is to ensure that a leadership system is in place that will maintain the gains and generate additional improvements.

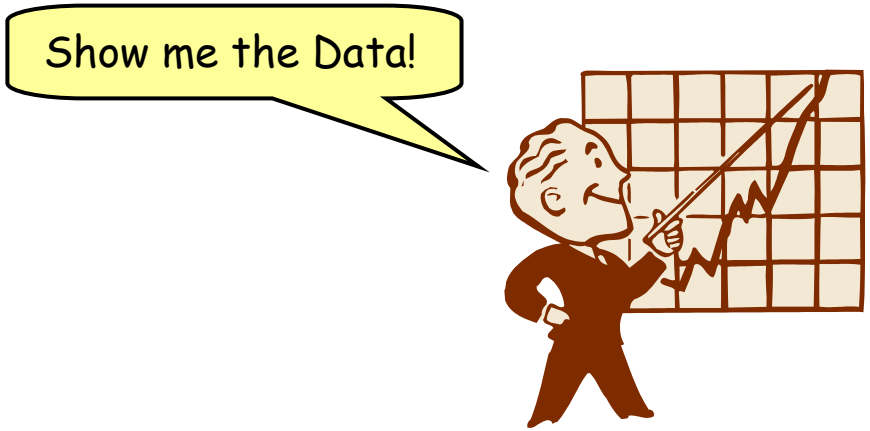


1. What is your schedule for **Operating Reviews**?
2. What organization **levels** are involved?
3. How are you involving your leadership team in business **planning**?
4. How are you involving your management team in **budgeting**?
5. Show **examples** of identified improvements and how they have been incorporated into the system.

Communicating with Data

Communicating with Data is key behavior to be developed during Annual Operating Reviews.

- Those conducting reviews can develop this behavior by asking to see supporting documents whenever a question is answered.
- Encourage the use of tools.
- Commend others for having documentation that supports their response.



Format for Review Presentations

Each level of management should adopt its own approach for conducting an Operating Review. Care should be taken though, to ensure consistency to enhance communication between those involved in the review.

The format for Operating Review presentations generally includes two major components.

- ***Improvement Story Summary***

  The Improvement Story Summary is a synopsis or roll-up of the information detailed in the Improvement Story.

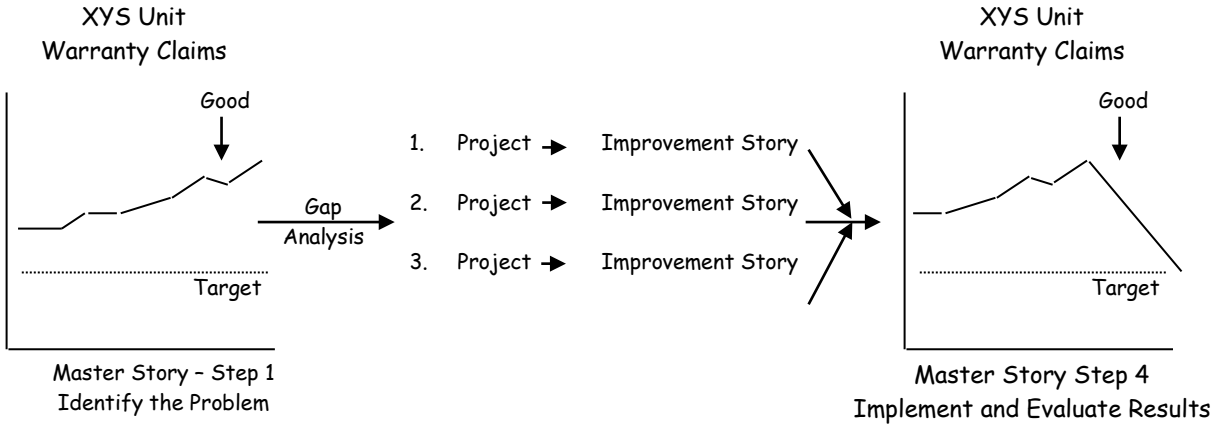
- ***Improvement Story***

  An Improvement Story details only the steps, actions, and tools a team or individual used to improve a process or condition.

Master Improvement Story

A Master Improvement Story is a variation of an Improvement Story. It is also presented using the *Improvement Story Summary* to roll up the detail. A Master Improvement Story is different because it contains the outcome of several projects focused on one problem area.

For example, a Geographic Unit adopts a KFA to reduce warranty claims occurring within 90 days of delivery by 50%. Three root causes are identified, and each is assigned as a project to different work groups. Each team completes an Improvement Story for their project. The combined impact of the three Stories is described in a Master Improvement Story.



Improvement Project Reviews

Local management conducts Improvement Story reviews to monitor the progress of local process improvement efforts. In addition, they provide an opportunity for coaching on the use of the basic problem-solving tools, such as cause-and-effect diagram, check sheet, line graph, Pareto chart, etc.

When to Conduct Improvement Project Reviews

Reviews of Improvement Stories may be conducted at any time, but should be scheduled monthly at a regularly scheduled time, e.g., the last Thursday of the month at 1 P. M. This helps ensure that the review will become part of the way business is normally conducted. Regularly scheduled reviews create a *"predictable event"* that facilitates preparation planning for both the presenting team and leadership conducting the review.

Improvement Project Review Agenda

A typical Improvement Story Review lasts 30-45 minutes.

Sample Team Review Agenda		
Who	**What**	**Time**
Team	A. Discuss process improvement progress and application of tools and techniques.	10-15 min.
Team	B. Address any open action items.	5 min.
Reviewers	C. Ask the presenting team questions.	5-10 min.
Team	D. Respond to questions.	5 min.
Team	E. Discuss next steps.	5 min.
Reviewers	F. Summarize review and feedback.	5 min.

Guidelines for Conducting an Improvement Project Review

- Copies of up-to-date Improvement Stories that are scheduled for review should be provided to the reviewers ahead of time to ensure that questions or feedback can be prepared.
- Teams should be allowed to complete their presentation before questions are asked or comments offered.
- When asking questions or offering comments, refer to specific pages or tools in the Improvement Story.
- Concentrate on the logic flow. Does the data support the conclusion in each step?
- Pay attention to detail. Have the tools been applied correctly?
- Involve all the team members. If only one or two people are presenting, ask questions of those not presenting.
- Consider the development level of the team, i.e., new team, experienced team, cross-functional team, etc.
- Be generous with praise and encouragement. Highlight things done well. Use this opportunity to recognize team accomplishments.
- Address areas in need of improvement. Make constructive comments to the Team Leader, Sponsor, and local supervisor as needed after the team presents its Story. This helps maintain team confidence and reinforces these important roles.
- Keep an open mind. There is usually more than one way to approach a problem.
- Thank the team members for their commitment and hard work. Encourage them to stay involved in their improvement activities.
- Utilize the Process Improvement Checkpoints forms (see next page) to guide feedback.

Process Improvement Checkpoints Process: _____ Date: _____

Steps	Questions	Commendable	Opportunity
Define Process and Determine Ownership	1.1 What is the process name? 1.2 What level one objective does the process contribute to? 1.3 Is the process owner responsible for process results? 1.4 What is the purpose of the process? 1.5 Have the process boundaries been clearly defined?		
Identify Customers and Establish Valid Requirements	2.1 Have all customers, external & internal been accurately identified? 2.2 What are the customer requirements? 2.3 How were the requirements verified? 2.4 Have the requirements been formally communicated to the suppliers?		
Document Process Flow	3.1 Does the process flowchart make sense? 3.2 Does the process flow include responsible people and steps? 3.3 Are the process boundaries consistent with the process profile? 3.4 What ATypical Process Deficiencies were identified that allow for immediate improvement? 3.5 Have the process changes been communicated to everyone involved in the process?		
Identify and Implement Indicators and Targets	4.1 What is the outcome indicator? 4.2 Does it specifically measure performance against customer requirements? 4.3 What are the process indicators? 4.4 Are the indicators consistent, controllable, and customer oriented? 4.5 What types of graphs/charts were used?		
Identify Problem Areas	5.1 What is the gap in current performance? 5.2 Has a prioritization matrix been developed? 5.3 How has the data been stratified?		
Resolve Problem(s)/ Redesign Process	6.1 How were the root causes verified? 6.2 Which countermeasures were implemented and why? 6.3 Can each countermeasure be linked to the improvement? 6.4 Was the flowchart revised? 6.5 How did you replicate the process improvements?		
PDCA Process	7.1 What is your plan to continually monitor performance? 7.2 What is your plan to move towards the ultimate design?		

The System Review

The Business Planning process needs to be improved as it is implemented. Typically, the process will go through several iterations of learning, becoming more effective with each cycle. The mechanism for this improvement is a review. The objective is to diagnose problems in the business planning process itself and the capability of the organization to implement it. Areas to be reviewed are:

- Identification of projects,
- Problem solving skills,
- Follow through on projects, and
- Achievement of targeted levels of performance.

The diagnosis should lead to changes that will improve upon each of these items.

Activity:

Perform a diagnosis of your previous experiences with strategic planning by answering the following questions.

1. Were projects identified that aligned with the corporate/unit key focus areas? If not, why?
2. Were the projects completed on schedule? If not, why?
3. Did the projects result in the targeted level of performance? If not, why?
4. Based on your diagnosis, what improvements would you recommend?

Getting Operating Reviews Wrong!

This system of business planning is a big opportunity for any company. Similar systems have been implemented in other companies and some reasons why they fail are shown below.

- **Not keyed to specific results** – a clear linkage must be established between improvement activities and measurable.
- **Too large a scale and diffused** – management must be able to determine which initiatives are delivering results to the organization.
- **Not results focused** – improvement activities must produce financial and operational improvements in both the short and long term.
- **Delusional measurements** – financial and operational performance measures that link P3 and the company's business goals are necessary.
- **Staff and Consultant Driven -** ownership of leadership system and its component parts must reside with everyone in the organization.
- **Overly focused on details rather than the process and results** – an understanding of the cause and effect relationship between improvement activities and financial and operational results will yield lessons learned from both successes and failures.

16.2.3 Summary

Organizational change is driven in part by a demonstrated commitment on the part of leaders. Following the guidelines provided in this unit will help ensure that Operating Reviews provide a powerful stimulus for continuous improvement.

Expect the following outputs from conducting effective Operating Reviews:

- Alignment toward a common set of objectives.
- Results are achieved on schedule.
- Performance feedback to everyone involved.
- Training needs are assessed.
- Recognition is freely given for results that improve performance toward stated objectives.
- Your company's Values are upheld.

Can you think of others?

16.3 Exercises

Exercise – Affinity Diagram

Brainstorm or generate a list of ideas. Suggestions for topics include:

- Customer needs for a pizza parlor.
- Plan Elements to allow you to retire in the next 5-10 years.
- The skills and behaviors of a Black Belt.
- The features you'd like in your next car (or boat, house, etc.)

Create an Affinity diagram for these ideas.

Exercise – Arrow (PERT) Diagram

Brainstorm a list of activities necessary to perform one of the following processes:

- Performing a frontal lobotomy
- Preparing for vacation
- Building a tree house for your children
- Preparing for a backyard bar-b-q
- Overhauling your car's transmission
- Putting up a basketball hoop over the garage

Organize them into an Arrow Diagram.

Exercise – Idea Matrix

Generate lists of ideas for the following exploratory relationships.

- Needs of your spouse *and* your behaviors or characteristics
- Company indicators *and* strategies
- Customer needs for your product or service *and* associated characteristics (e.g. this is a QFD exercise).
- Core processes (or products) *and* company departments

Develop an Idea Matrix to show how these ideas are related.

Exercise – Process Decision Program Chart

Take one of the processes developed as part of the Arrow diagram exercise. Develop a Process Decision Program Chart to identify what can go wrong and what can be done to prevent or mitigate the problems.

Exercise – Structure Tree

Brainstorm or generate ideas to accomplish one of the following goals. Arrange them into the first level of a Structure tree and then proceed to develop one or two more layers (or until activities are identified):

- A current company strategy.
- Plan Elements to allow you to retire in the next 5-10 years.
- The skills and behaviors of an effective Black Belt.
- Convert the US to the metric system.
- A current community goal.

Exercise – Relations Diagram

Using one of the topics developed in the Structure tree exercise, develop a Relations diagram to determine the relationships and drivers of these goals.

Exercise - Operating Reviews

This activity will help you practice effective Operating Review behaviors.

Instructions for Reviewing Team:

1. Meet with others to ***plan for an operating review*** of another team. Detail your plan on a flipchart. Your plan should reflect the information provided in this module. Your plan should include the following:
   - Why the review is being conducted.
   - What level is being reviewed.
   - When, where, and how long the review will take.
   - Who is involved.

2. ***Develop a set of questions*** to be asked during the review. Note, there may not be any documentation to study prior to this review, how will you handle this situation?

3. ***Establish an agenda and any guidelines*** for the review you would like the team being reviewed to follow. Provide this information to the team being reviewed prior to the review.

4. ***Conduct the review***. Stick to the agenda during the review session.

Instructions for Team being reviewed:

1. Prepare for the review by identifying a project or objective about which to report your current situation. You may chose something real or make something up. A sample case study is provided on the next page that your team may elect to use.

2. The reviewing team is establishing review guidelines, and prior to the review should be communicating this information to your team.

3. Actively participate in the review.

Instructions for Observers:

1. Review the guidelines presented in this module.

2. Observe the review and make note of the behaviors of both the reviewing team and the team being reviewed. Be prepared to offer feedback during a discussion following the review.

After the Review

Debrief the review by discussing the following questions:

a. What did the reviewing team like about the way they conducted the review?

b. What would they do differently if they could do it all over again?

c. How did the reviewed team feel about the review?

d. What is their level of motivation and enthusiasm to continue work on the project or objective?

e. What behaviors did the observers notice during the review?

f. Summarize any “lessons learned” from this activity.

Sample Case Situation

Your Business Unit Quality Lead Team attended Champions Training several months ago. As a result, your team identified two candidates to become Team Training instructors, and one person to become a certified Black Belt. Results from projects in this year’s AOP are scattered, but you are confident in the direction your unit is heading. Sales are up slightly, a warranty reduction team has met 3-4 times, and another team is working on a cycle-time project to reduce response time for customer service complaints.

Appendices

Appendix	*Description*	*Page*
A	Probability Distributions – Cumulative Values	A-1
B	Sigma Conversion Table	B-1
C	Forms and Templates	C-1
D	Answers to Selected Exercises	D-1
E	Reliability Exercise	E-1

Appendix A – Probability Distributions

Four commonly used tables of probability distributions and a table of critical values for an important non-parametric test are included here:

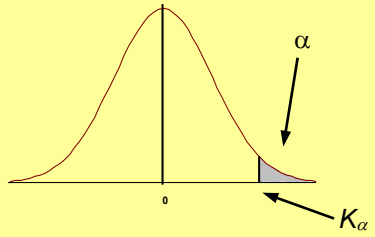
- *Table A.1 - The Standard Normal Distribution*
- *Table A.2 - The student's t Distribution*
- *Table A.3 - The Chi-Square Distribution*
- *Table A.4 - The F-Distribution*
- *Table A.5 – Mann-Whitney Test Critical Values*

A summary table of common probability distributions and their properties/parameters is also included.

Table A.1 - The Standard Normal Distribution

K_α	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

This table provides values of one minus the cumulative standard normal distribution (α). If you want to find α from K_α, enter the table at the appropriate value of K_α and read the α.



For example, the value of α for K_α = 2.55 is found at the intersection of 2.5 and 0.05 which is 0.0054. Since the normal distribution is symmetric, α's for negative K_α's are the same as for positive K_α's. For hypothesis testing, you generally want to find K_α from α. The following table provides a quick reference for this situation:

α	0.001	0.005	0.010	0.025	0.05	0.10	0.20	0.30	0.40
K_α	3.090	2.576	2.326	1.960	1.645	1.282	0.842	0.524	0.253

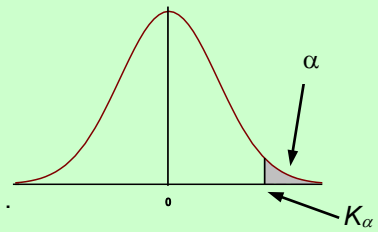
For example, for an $\alpha = 0.05$, the associated value of K_α would be 1.645.

If the hypothesis test is *two-sided*, you should divide α by 2. Read off the plus or minus $K_{\alpha/2}$ values. For example, for an $\alpha = 0.05$ *two-sided test,* read off the $K_{\alpha/2}$ values for $\alpha/2 = 0.025$ - these are +/- 1.960.

Table A.2 - The student's t Distribution

f \ α	0.1	0.05	0.025	0.01
1	3.078	6.314	12.71	31.82
2	1.886	2.920	4.303	6.965
3	1.638	2.353	3.182	4.541
4	1.533	2.132	2.776	3.747
5	1.476	2.015	2.571	3.365
6	1.440	1.943	2.447	3.143
7	1.415	1.895	2.365	2.998
8	1.397	1.860	2.306	2.896
9	1.383	1.833	2.262	2.821
10	1.372	1.812	2.228	2.764
11	1.363	1.796	2.201	2.718
12	1.356	1.782	2.179	2.681
13	1.350	1.771	2.160	2.650
14	1.345	1.761	2.145	2.624
15	1.341	1.753	2.131	2.602
16	1.337	1.746	2.120	2.583
17	1.333	1.740	2.110	2.567
18	1.330	1.734	2.101	2.552
19	1.328	1.729	2.093	2.539
20	1.325	1.725	2.086	2.528
21	1.323	1.721	2.080	2.518
22	1.321	1.717	2.074	2.508
23	1.319	1.714	2.069	2.500
24	1.318	1.711	2.064	2.492
25	1.316	1.708	2.060	2.485
26	1.315	1.706	2.056	2.479
27	1.314	1.703	2.052	2.473
28	1.313	1.701	2.048	2.467
29	1.311	1.699	2.045	2.462
30	1.310	1.697	2.042	2.457
40	1.303	1.684	2.021	2.423
50	1.298	1.676	2.009	2.403
60	1.296	1.671	2.000	2.390
80	1.292	1.664	1.990	2.374
100	1.290	1.660	1.984	2.365
200	1.286	1.653	1.972	2.345
500	1.283	1.648	1.965	2.334
∞	1.282	1.645	1.960	2.326

This table provides values of one minus the cumulative student's t distribution (α). If you want to find K_α from α, enter the table at the appropriate value of α and *f,* the degrees of freedom and read K_α



For example, for an α = 0.05 and *f* = 20 degrees of freedom, the value of K_α is equal to 1.725.

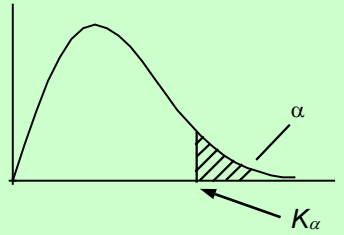
If you are performing a *two-sided* hypothesis test, you should divide α by 2. Read off the plus or minus $K_{\alpha/2}$ values. For example, for an α = 0.05 *two-sided test* with 30 degrees of freedom, read off the $K_{\alpha/2}$ values for α/2 = 0.025 - these are +/- 2.042.

Notice that the values of K_α for large (*f* = ∞) degrees of freedom are identical to those of the normal distribution. As the sample size approaches the size of the population the variance becomes "known" and there is no difference between the normal and the student's t.

Table A.3 - The Chi-Square Distribution

f α	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.00004	0.00016	0.00098	0.003	3.84	5.02	6.64	7.88
2	0.0100	0.0201	0.0506	0.103	5.99	7.38	9.21	10.60
3	0.0717	0.115	0.216	0.352	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.625	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.17	14.07	16.01	18.48	20.3
8	1.344	1.646	2.18	2.73	15.51	17.53	20.1	22.0
9	1.735	2.09	2.70	3.33	16.92	19.02	21.7	23.6
10	2.16	2.56	3.25	3.94	18.31	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	19.68	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	21.0	23.3	26.2	28.3
13	3.57	4.11	5.01	5.89	22.4	24.7	27.7	29.8
14	4.07	4.66	5.63	6.57	23.7	26.1	29.1	31.3
15	4.60	5.23	6.26	7.26	25.0	27.5	30.6	32.8
16	5.14	5.81	6.91	7.96	26.3	28.8	32.0	34.3
17	5.70	6.41	7.56	8.67	27.6	30.2	33.4	35.7
18	6.26	7.01	8.23	9.39	28.8	31.5	34.8	37.2
19	6.84	7.63	8.91	10.12	30.2	32.9	36.2	38.6
20	7.43	8.26	9.59	10.85	31.4	34.2	37.6	40.0
21	8.03	8.90	10.28	11.59	32.7	35.5	38.9	41.4
22	8.64	9.54	10.98	12.34	33.9	36.8	40.3	42.8
23	9.26	10.20	11.69	13.09	35.2	38.1	41.6	44.2
24	9.89	10.86	12.40	13.85	36.4	39.4	43.0	45.6
25	10.52	11.52	13.12	14.61	37.7	40.6	44.3	46.9
26	11.16	12.20	13.84	15.38	38.9	41.9	45.6	48.3
27	11.81	12.88	14.57	16.15	40.1	43.2	47.0	49.6
28	12.46	13.56	15.31	16.93	41.3	44.5	48.3	51.0
29	13.12	14.26	16.05	17.71	42.6	45.7	49.6	52.3
30	13.79	14.95	16.79	18.49	43.8	47.0	50.9	53.7
40	20.7	22.2	24.4	26.5	55.8	59.3	63.7	66.8
50	28.0	29.7	32.4	34.8	67.5	71.4	76.2	79.5
60	35.5	37.5	40.5	43.2	79.1	83.3	88.4	92.0
70	43.3	45.4	48.8	51.7	90.5	95.0	100.4	104.2
80	51.2	53.5	57.2	60.4	101.9	106.6	112.3	116.3
90	59.2	61.8	65.6	69.1	113.1	118.1	124.1	128.3
100	67.3	70.1	74.2	77.9	124.3	129.6	135.8	140.2
f α	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005

This table provides values of one minus the cumulative chi-square distribution (α). If you want to find K_α from α, enter the table at the appropriate value of α and *f,* the degrees of freedom and read K_α.

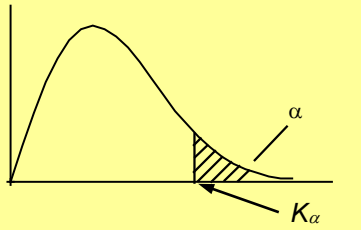


For example, for an α = 0.05, and 20 degrees of freedom, the K_α value is 31.4. If the hypothesis test is two-sided, divide α by 2 and find the associated $K_{\alpha/2}$ values. For example, for α = 0.05, two-sided $K_{\alpha/2}$ values for 27 degrees of freedom are 14.57 and 43.2.

Table A.4 - The F Distribution

d \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	n \ d
1	**161**	**200.**	**216.**	**225.**	**230.**	**234.**	**237.**	**239.**	**241.**	**242.**	**244.**	**246.**	**248.**	**249.**	**250.**	**251.**	**252.**	**253.**	**254.**	1
	4052	5000	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366	
2	**18.5**	**19.0**	**19.2**	**19.2**	**19.3**	**19.3**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.4**	**19.5**	**19.5**	**19.5**	**19.5**	**19.5**	**19.5**	2
	98.5	99.0	99.2	99.2	99.3	99.3	99.4	99.4	99.4	99.4	99.4	99.4	99.4	99.5	99.5	99.5	99.5	99.5	99.5	
3	**10.1**	**9.55**	**9.28**	**9.12**	**9.01**	**8.94**	**8.89**	**8.85**	**8.81**	**8.79**	**8.74**	**8.70**	**8.66**	**8.64**	**8.62**	**8.59**	**8.57**	**8.55**	**8.53**	3
	34.1	30.8	29.5	28.7	28.2	27.9	27.7	27.5	27.3	27.2	27.1	26.9	26.7	26.6	26.5	26.4	26.3	26.2	26.1	
4	**7.71**	**6.94**	**6.59**	**6.39**	**6.26**	**6.16**	**6.09**	**6.04**	**6.00**	**5.96**	**5.91**	**5.86**	**5.80**	**5.77**	**5.75**	**5.72**	**5.69**	**5.66**	**5.63**	4
	21.2	18.0	16.7	16.0	15.5	15.2	15.0	14.8	14.7	14.5	14.4	14.2	14.0	13.9	13.8	13.7	13.7	13.6	13.5	
5	**6.61**	**5.76**	**5.41**	**5.19**	**5.05**	**4.95**	**4.88**	**4.82**	**4.77**	**4.74**	**4.68**	**4.62**	**4.56**	**4.53**	**4.50**	**4.46**	**4.43**	**4.40**	**4.36**	5
	16.3	13.3	12.1	11.4	11.0	10.7	10.5	10.3	10.2	10.1	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02	
6	**5.99**	**5.14**	**4.76**	**4.53**	**4.39**	**4.28**	**4.21**	**4.15**	**4.10**	**4.06**	**4.00**	**3.94**	**3.87**	**3.84**	**3.81**	**3.77**	**3.74**	**3.70**	**3.67**	6
	13.7	10.9	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88	
7	**5.59**	**4.74**	**4.35**	**4.12**	**3.97**	**3.87**	**3.79**	**3.73**	**3.68**	**3.64**	**3.57**	**3.51**	**3.44**	**3.41**	**3.38**	**3.34**	**3.30**	**3.27**	**3.23**	7
	12.2	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65	
8	**5.32**	**4.46**	**4.07**	**3.84**	**3.69**	**3.58**	**3.50**	**3.44**	**3.39**	**3.35**	**3.28**	**3.22**	**3.15**	**3.12**	**3.08**	**3.04**	**3.01**	**2.97**	**2.93**	8
	11.3	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.86	
9	**5.12**	**4.26**	**3.86**	**3.63**	**3.48**	**3.37**	**3.29**	**3.23**	**3.18**	**3.14**	**3.07**	**3.01**	**2.94**	**2.90**	**2.86**	**2.83**	**2.79**	**2.75**	**2.71**	9
	10.6	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31	
10	**4.96**	**4.10**	**3.71**	**3.48**	**3.33**	**3.22**	**3.14**	**3.07**	**3.02**	**2.98**	**2.91**	**2.84**	**2.77**	**2.74**	**2.70**	**2.66**	**2.62**	**2.58**	**2.54**	10
	10.0	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91	
12	**4.75**	**3.89**	**3.49**	**3.26**	**3.11**	**3.00**	**2.91**	**2.85**	**2.80**	**2.75**	**2.69**	**2.62**	**2.54**	**2.51**	**2.47**	**2.43**	**2.38**	**2.34**	**2.30**	12
	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36	
15	**4.54**	**3.68**	**3.29**	**3.06**	**2.90**	**2.79**	**2.71**	**2.64**	**2.59**	**2.54**	**2.48**	**2.40**	**2.33**	**2.29**	**2.25**	**2.20**	**2.16**	**2.11**	**2.07**	15
	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87	
20	**4.35**	**3.49**	**3.10**	**2.87**	**2.71**	**2.60**	**2.51**	**2.45**	**2.39**	**2.35**	**2.28**	**2.20**	**2.12**	**2.08**	**2.04**	**1.99**	**1.95**	**1.90**	**1.84**	20
	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42	
24	**4.26**	**3.40**	**3.01**	**2.78**	**2.62**	**2.51**	**2.42**	**2.36**	**2.30**	**2.25**	**2.18**	**2.11**	**2.03**	**1.98**	**1.94**	**1.89**	**1.84**	**1.79**	**1.73**	24
	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2.40	2.31	2.21	
30	**4.17**	**3.32**	**2.92**	**2.69**	**2.53**	**2.42**	**2.33**	**2.27**	**2.21**	**2.16**	**2.09**	**2.01**	**1.93**	**1.89**	**1.84**	**1.79**	**1.74**	**1.68**	**1.62**	30
	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01	
40	**4.08**	**3.23**	**2.84**	**2.61**	**2.45**	**2.34**	**2.25**	**2.18**	**2.12**	**2.08**	**2.00**	**1.92**	**1.84**	**1.79**	**1.74**	**1.69**	**1.64**	**1.58**	**1.51**	40
	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80	
60	**4.00**	**3.15**	**2.76**	**2.53**	**2.37**	**2.25**	**2.17**	**2.10**	**2.04**	**1.99**	**1.92**	**1.84**	**1.75**	**1.70**	**1.65**	**1.59**	**1.53**	**1.47**	**1.39**	60
	7.08	7.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.78	1.60	
120	**3.92**	**3.07**	**2.68**	**2.45**	**2.29**	**2.18**	**2.09**	**2.02**	**1.96**	**1.91**	**1.83**	**1.75**	**1.66**	**1.61**	**1.55**	**1.50**	**1.43**	**1.35**	**1.25**	120
	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38	
∞	**3.84**	**3.00**	**2.60**	**2.37**	**2.21**	**2.10**	**2.01**	**1.94**	**1.88**	**1.83**	**1.75**	**1.67**	**1.57**	**1.52**	**1.46**	**1.39**	**1.32**	**1.22**	**1.00**	∞
	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00	
d \ n	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞	n / d

This table provides values of one minus the cumulative F distribution (α). If you want to find K_α from α, enter the table at the appropriate value of α and f_n & f_d, the degrees of freedom and read K_α. The bold values are for α = 0.05, the fine values for α = 0.01.



For example, for an α = 0.05, 24 degrees of freedom for the numerator and 15 degrees of freedom for the denominator, the K_α is 2.29.

n - Degrees of Freedom for Numerator (Table Columns), d - Degrees of Freedom for Denominator (Table Rows)

Table A.5 - Mann-Whitney Test Critical Values (W_α)

		n_1																	
n_2	α	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	--	0	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
	.01	--	0	0	0	0	0	0	0	0	1	1	1	2	2	2	2	3	3
4	.05	--	0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
	.01	--	--	0	0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5	.05	0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
	.01	--	--	0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6	.05	1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
	.01	--	0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7	.05	1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
	.01	--	0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8	.05	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
	.01	--	1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9	.05	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
	.01	0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10	.05	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
	.01	0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11	.05	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
	.01	0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12	.05	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
	.01	1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13	.05	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
	.01	1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	56	60
14	.05	5	9	13	17	22	26	31	36	40	45	50	55	59	64	67	74	78	83
	.01	1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15	.05	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
	.01	2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16	.05	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
	.01	2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17	.05	6	11	17	22	28	34	39	45	51	57	63	67	75	81	87	93	99	105
	.01	2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18	.05	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
	.01	2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	.05	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
	.01	3	7	12	17	22	28	33	39	45	51	56	63	69	74	81	87	93	99
20	.05	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127
	.01	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

Use This Table for Two-Tailed Testing

Table A.5 - Mann-Whitney Test Critical Values (W_α) (Continued)

		n_1																	
n_2	α	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	.05	0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
	.01	--	0	0	0	0	0	1	1	1	2	2	2	3	3	4	4	4	5
4	.05	0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
	.01	--	--	0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5	.05	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
	.01	--	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6	.05	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
	.01	--	1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7	.05	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
	.01	0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8	.05	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
	.01	0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9	.05	4	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
	.01	1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10	.05	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
	.01	1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11	.05	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
	.01	1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12	.05	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
	.01	2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	.05	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
	.01	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	.05	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
	.01	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	.05	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
	.01	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	.05	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
	.01	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	.05	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
	.01	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	.05	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
	.01	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	.05	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
	.01	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	.05	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138
	.01	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Use This Table for One-Tailed Testing

Some Common Probability Distributions

Distribution	Parameters	Density Function	Expected Value	Variance
Binomial	n $0 \le p \le 1$	$f(x) = \frac{n!}{x!(n-x)}$	np	$np(1-p)$
Multinomial	n – positive integer $p_i \ge 0, where$ $\sum p_i = 1$	$f(x) = \frac{n!}{x_1!x_2!x_3!...x_k!}$ $\sum x_i = n$	$\sum x_i = np_i$	$\sigma^2(x_i) = np_i(1-p_i)$
Poisson	$\lambda > 0$	$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$	λ	λ
Geometric	$0 \le p \le 1$	$f(x) = (1-p)^{x-1} p$	$1/p$	$(1-p)/p^2$
Hypergeometric	a, b, n (all > 0) $n \le a+b$	$f(x) = \frac{C_x^a C_{n-x}^b}{C_n^{a+b}}$ for x = 0, 1, 2, .. and x ≤ a and x ≤ n	$an/(a+b)$	$\frac{ab(a+b+n)n}{(a+b)^2(a+b-1)}$
Normal	μ, σ	$f(x) = \frac{1}{2\pi\sigma} \exp(-(x-\mu)^2 / 2\sigma^2$	μ	σ^2
Beta	$\alpha > 0, \beta > 0$	$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$	$\alpha/(\alpha+\beta)$	$\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$
Uniform	a, b where $a < b$	$f(x) = 1/(b-a)$	$(a+b)/2$	$(b-a)^2/12$
Gamma	$\alpha > 0, \beta > 0$	$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}$	$\alpha\beta$	$\alpha\beta^2$
Rayleigh	$\sigma > 0$	$f(x) = \frac{x}{\sigma^2} \exp(\frac{1}{2}\left(\frac{x}{\sigma}\right)^2$ for x > 0	$\sigma\left(\frac{\pi}{2}\right)^{1/2}$	$2\sigma^2\left(1-\frac{\pi}{4}\right)$
Extreme Value	$\sigma \ne 0, \mu$	$f(x) = \frac{1}{\sigma} \exp\left(-\frac{(x-\mu}{\sigma} - \exp\left(-\frac{x-\mu}{\sigma}\right)\right)$	$\mu + 0.577\sigma$	$\frac{(\pi\sigma)^2}{6}$

Distribution	Parameters	Density Function	Expected Value	Variance
Lognormal	μ, σ	$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} \exp\left(-\frac{(\log(x-\mu)^2}{2\sigma^2}\right)$	$\exp\left(\mu + \frac{\sigma^2}{2}\right)$	$\exp(2\mu+\sigma^2)(\exp(\sigma^2)-1)$
Exponential	$\lambda > 0$	$f(x) = \frac{\exp(-x/\lambda)}{\lambda}$	λ	λ^2
Chi-Square	γ	$f(\chi^2) = \frac{(\chi^2)^{\gamma/2-1}\exp(-\chi^2/2)}{2^{\gamma/2}(\gamma/2-1)}$	γ	2γ
F	γ_1, γ_2	$f(F) = \frac{\Gamma((\gamma_1+\gamma_2)/2)}{\Gamma(\gamma_1/2)\Gamma(\gamma_2/2)}\left(\frac{\gamma_1}{\gamma_2}\right)^{\gamma_1/2} \frac{F^{(\gamma_1/2-1)}}{\left(1+\frac{\gamma_1 F}{\gamma_2}\right)^{(\gamma_1+\gamma_2)/2}}$	$\frac{\gamma_1}{\gamma_2-2}$	$\frac{2\gamma_2^2(\gamma_1+\gamma_2-2)}{\gamma_1(\gamma_2-2)^2(\gamma_2-4)}$
Student's t	γ	$f(t) = \frac{(\gamma+1)/2}{\sqrt{k\pi}(\gamma/2)} \frac{1}{(1+(t^2/\gamma))^{(\gamma+1)/2}}$	0	$\gamma/(\gamma-2)$
Weibull	η, β	$f(x) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp(x/\eta)^{\beta}$	$\eta\Gamma\left(1+\frac{1}{\beta}\right)$	$\eta^2\left[\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)\right]$

Appendix B – Sigma Conversion Table

Table One – Yield/Sigma/DPMO Values

Long-Term Process Yield (%)	*Short-Term Sigma*	*DPMO*	*Long-Term Sigma*
6.68	0	933,193	Negative
8.08	0.1	919,243	Negative
9.68	0.2	903,199	Negative
11.51	0.3	884,930	Negative
13.57	0.4	864,334	Negative
15.87	0.5	841,345	Negative
18.41	0.6	815,940	Negative
21.19	0.7	788,145	Negative
24.20	0.8	758,036	Negative
27.43	0.9	725,747	Negative
30.85	**1**	**691,462**	**Negative**
34.46	1.1	655,422	Negative
38.21	1.2	617,911	Negative
42.07	1.3	579,260	Negative
46.02	1.4	539,828	Negative
50.00	1.5	500,000	0.0
53.98	1.6	460,172	0.1
57.93	1.7	420,740	0.2
61.79	1.8	382,089	0.3
65.54	1.9	344,578	0.4
69.15	**2**	**308,538**	**0.5**
72.57	2.1	274,253	0.6
75.80	2.2	241,964	0.7
78.81	2.3	211,855	0.8
81.59	2.4	184,060	0.9
84.13	2.5	158,655	1.0
86.43	2.6	135,666	1.1
88.49	2.7	115,070	1.2
90.32	2.8	96,801	1.3
91.92	2.9	80,757	1.4
93.32	**3**	**66,807**	**1.5**
94.52	3.1	54,799	1.6
95.54	3.2	44,565	1.7
96.41	3.3	35,930	1.8
97.13	3.4	28,716	1.9
97.73	3.5	22,750	2.0
98.21	3.6	17,864	2.1
98.61	3.7	13,903	2.2
98.93	3.8	10,724	2.3
99.18	3.9	8,198	2.4
99.38	4	6,210	2.5
99.53	4.1	4,661	2.6
99.65	4.2	3,467	2.7
99.74	4.3	2,555	2.8
99.81	4.4	1,866	2.9
99.87	4.5	1,350	3.0
99.90	4.6	968	3.1
99.93	4.7	687	3.2
99.95	4.8	483	3.3
99.97	4.9	337	3.4
99.98	**5**	**233**	**3.5**
99.98	5.1	159	3.6
99.989	5.2	108	3.7
99.9927	5.3	72	3.8
99.9951	5.4	48	3.9
99.9968	5.5	32	4.0
99.9979	5.6	21	4.1
99.99866	5.7	13	4.2
99.99914	5.8	8.5	4.3
99.99946	5.9	5.4	4.4
99.99966	**6**	**3.4**	**4.5**

Notes:

a) Yields lower than 6.68% are not included since both short and long-term Sigmas are negative.

b) The table is arranged so that Long-Term Process Yields and DPMO correspond to Long-Term Sigma values. For example, a Long-Term Yield of 99.99966% (or defect rate of 3.4 DPMO) corresponds to a Long-Term process Sigma of 4.5. Assuming a 1.5 Sigma "shift," the corresponding Short-Term Sigma would be 6.0.

Appendix C - Forms and Templates

Control Chart Forms

Even though there are many Statistical Software programs on the market today, there is a place for the “hand-drawn” control chart. We’ve had the pleasure of working with some hospital laboratory workers who plot their performance every day on a chart taped to a refrigerator door.

We’ve designed the control chart forms with space for 32 data points. Many applications that we see collect at least one subgroup of data daily, so each sheet could be your “chart-of-the-month.” If your subgroup frequency is different, well, we tried.

X-BAR, R & X-BAR, S CONTROL CHARTS	PROCESS	INDICATOR	
DEPARTMENT	RESPONSIBLE:	NOTES:	

DATE																															
SUBGROUP																															
SUM																															
AVERAGE																															
R or S																															

AVERAGE

RANGE/STD. DEV.

INDIVIDUAL (X,mR) CONTROL CHART	PROCESS	INDICATOR	
DEPARTMENT	RESPONSIBLE:	NOTES:	

DATE																														
SUBGROUP																														
RANGE																														

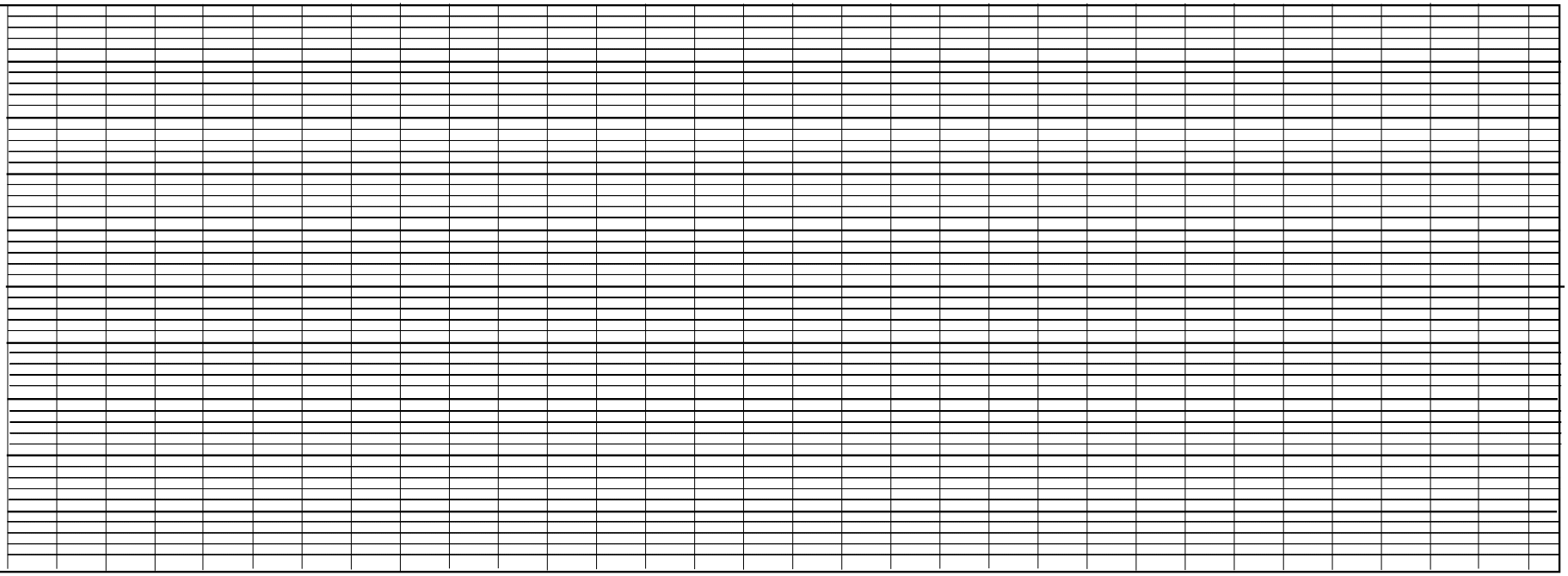
INDIVIDUAL

RANGE

ATTRIBUTE/COUNT DATA CONTROL CHART	PROCESS	INDICATOR	
DEPARTMENT	RESPONSIBLE:	NOTES:	

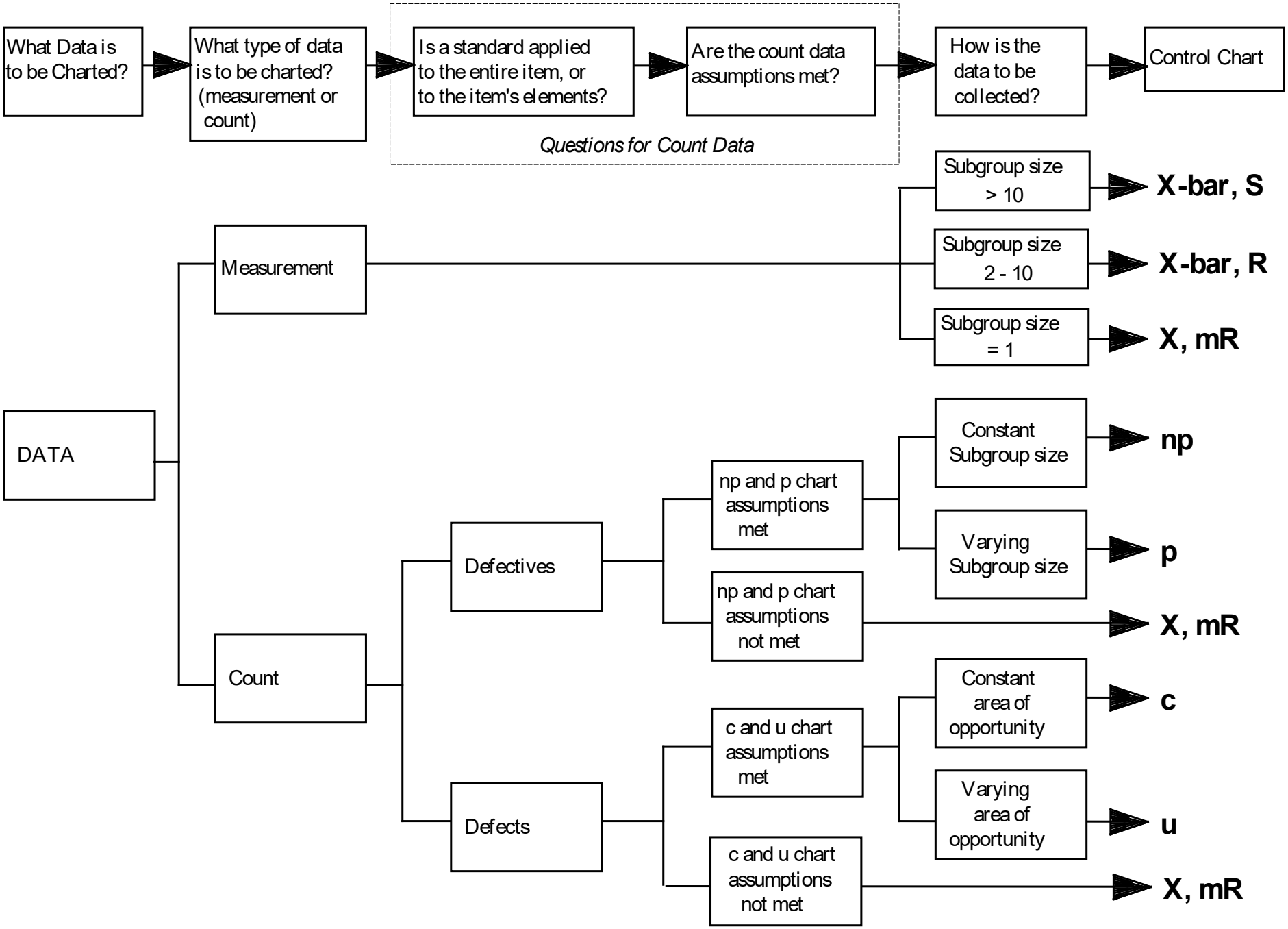
DATE																															
NUMBER DEFECTS/ DEFECTIVES																															
SUB-GROUP SIZE																															
P or U																															

NUMBER - FRACTION - RATE



NOTES:

CONTROL CHART SELECTION GUIDE



CONTROL CHART CALCULATIONS

X-bar, R Control Chart

Subgroup Ranges: $R_i = X_{\max} - X_{\min}$

Average Range: $\overline{R} = \frac{1}{k}\sum_{i=1}^{k} R_i$

Upper Control Limit$_{Range}$*:* $UCL_R = \overline{R} \times D_4$

Lower Control Limit$_{Range}$*:* $LCL_R = \overline{R} \times D_3$

Subgroup Averages: $\overline{X}_i = \frac{1}{m}\sum_{j=1}^{m} X_{ij}$

Grand Average: $\overline{\overline{X}} = \frac{1}{k}\sum_{i=1}^{k} \overline{X}_i$

Upper Control Limit$_{\overline{X}}$*:* $UCL_{\overline{X}} = \overline{\overline{X}} + (\overline{R} \times A_2)$

Lower Control Limit$_{\overline{X}}$*:* $UCL_{\overline{X}} = \overline{\overline{X}} - (\overline{R} \times A_2)$

X, mR Control Chart

Moving Ranges: $R_i = |X_i - X_{i-1}|$

Average Range: $\overline{R} = \frac{1}{k-1}\sum_{i=2}^{k} R_i$

Upper Control Limit$_{Range}$*:* $UCL_R = \overline{R} \times 3.268$

Individuals (X) Average: $\overline{X} = \frac{1}{k}\sum_{i=1}^{k} X_i$

Upper Control Limit$_X$*:* $UCL_X = \overline{X} + (\overline{R} \times 2.66)$

Lower Control Limit$_X$*:* $UCL_X = \overline{X} - (\overline{R} \times 2.66)$

np Control Chart

Defective Item Average : $\overline{np} = \frac{1}{k}\sum_{i=1}^{k} np_i$

Upper Control Limit : $UCL_{np} = \overline{np} + 3\sqrt{\overline{np}(1-\overline{np}/n)}$

Lower Control Limit : $LCL_{np} = \overline{np} - 3\sqrt{\overline{np}(1-\overline{np}/n)}$

p Control Chart

Average Fraction Defective: $\overline{p} = \sum_{i=1}^{k} np_i \Big/ \sum_{i=1}^{k} n_i$

Upper Control Limit: $UCL_p = \overline{p} + 3\sqrt{\overline{p}(1-\overline{p})/n_i}$

Lower Control Limit: $LCL_p = \overline{p} - 3\sqrt{\overline{p}(1-\overline{p})/n_i}$

c Control Chart

Average Number Defects: $\overline{c} = \frac{1}{k}\sum_{i=1}^{k} c_i$

Upper Control Limit: $UCL_c = \overline{c} + 3\sqrt{\overline{c}}$

Lower Control Limit: $LCL_c = \overline{c} - 3\sqrt{\overline{c}}$

u Control Chart

Subgroup Defect Rate: $u_i = c_i / n_i$

Average Defect Rate: $\overline{u} = \sum_{i=1}^{k} c_i \Big/ \sum_{i=1}^{k} n_i$

Upper Control Limit: $UCL_u = \overline{u} + 3\sqrt{\overline{u}/n_i}$

Lower Control Limit: $LCL_u = \overline{u} - 3\sqrt{\overline{u}/n_i}$

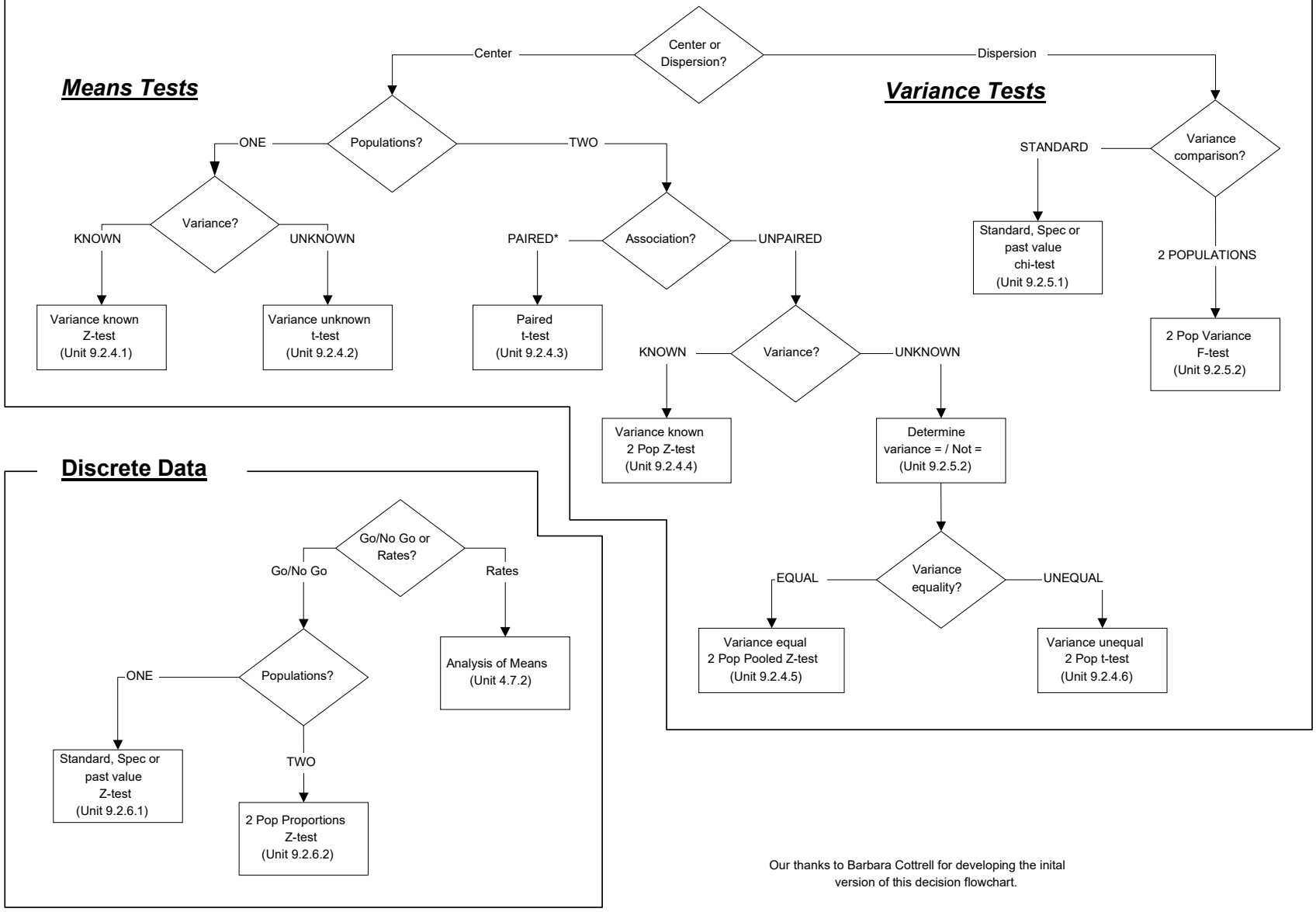
Constants for Control Limits

Subgroup Size	A2	D3	D4	d2
2	1.880	-	3.268	1.128
3	1.023	-	2.574	1.693
4	0.729	-	2.282	2.059
5	0.577	-	2.114	2.326
6	0.483	-	2.004	2.534
7	0.419	0.076	1.924	2.704
8	0.373	0.136	1.864	2.847
9	0.337	0.184	1.816	2.970
10	0.308	0.223	1.777	3.078

Interpretation Rules for Out-of-Control Conditions

1. Any single point outside the control limits.
2. Two of three consecutive points more than two sigma away from the center line.
3. Four of five points on the same side of and more than one sigma away from the center line.
4. A shift of seven or more consecutive points on either side of the center line.
5. A trend of seven consecutive points either increasing or decreasing.
6. Eight or more points that lie very close to the center line ("Hugging").
7. “Non-random" patterns that recur frequently.

Hypothesis Test Selection



Our thanks to Barbara Cottrell for developing the inital version of this decision flowchart.

DMAIEC Project Storyboard

Project:_____ **Team**:_____

Define

Measure

Analyze

Improve

Execute

Control

Appendix D – Answers to Selected Exercises

Here are answers to a few of the problems presented in the manual.

Section 3 – Team Facilitation & Management

Team Situations

In which of these following situations do you think a team should be formed to improve quality? If a team is needed, what *type* should it be? If you don't think a team is needed, how could or should the situation be addressed?

- A railroad has been experiencing water leaks on its locomotives' diesel engines. There are about 1000 locomotives in the railroad's fleet. The engineer's failure report includes where the leak is observed, but not why it occurred.

  Good opportunity for a team – problem known, root cause unknown, could occur because of design, operation, maintenance – cross-functional is recommended.

- An architectural firm has been receiving complaints from customers that "they are not responsive" to the customers' needs. The firm has four design groups, each acting as a design team for projects.

  Good opportunity for a team – problem known, root cause unknown, cross-functional (across design groups, functions) recommended.

- A small manufacturing company wishes to improve its employee safety record. The company president wants to form a team, but the Safety Officer tells him that he can solve the problem with a new training program for proper lifting techniques.

  Tough call – if Safety Officer has good data supporting the need for training, the president could authorize implementation.

- An unacceptably high defect rate of integrated circuits has plagued a small electronics firm for the last few weeks. The reliability engineer is working on a test plan, the design engineers are preparing changes to the IC design and manufacturing is changing their "clean room" procedures.

  Good opportunity for a team – problem known, root cause unknown, but countermeasures are being planned. Time to stop and find the real root causes – cross-functional team recommended.

- A hospital's case managers have identified one physician as being "high" on both patient Length of Stay and Cost per Case for a certain diagnosis.

  Not a good team situation – case manager could approach the physician with the data and offer to help identify practice patterns that impact these factors.

- Nurse managers have been complaining to the chief nurse executive about delays in receiving laboratory "stat" specimen reports. The lab director says the orders are only being sent to the lab twice a shift.

  Good opportunity for a team – problem known, root cause unknown (even the lab director has "data," the cross-department process could stand investigation – cross-functional team recommended.

- A physician called plant maintenance about dust blowing into one of her examining rooms from an air conditioning vent. The problem has existed for three days now.

  This is just a "do-it" situation – no need for a team!

- Two employees on the evening shift at a plastics plant are chronically late. The other shift members are angry at having to carry their "load" when they are late.

  Discussion with the employees and the supervisor is needed – not a team situation.

- A manufacturer of ceramics for hobbyists found that their product sales were declining. Projections indicated that the manufacturer would suffer a $10 million loss if the current trend continues.

  Good opportunity for a team – problem known, root cause unknown, could occur because of design, production, sales – cross-functional (high-level) team recommended.

Team Appropriateness

Comment on the following "team" situations described below. Was the use of a team appropriate? What issues do you see that may lead (or did lead) to the success or failure of these efforts?

- A manager of an engineering division told a group of engineers to investigate computerizing a certain reference document. He told them to make sure and "prove" that the computerization was necessary so the necessary budget approvals could be obtained.

  BAAAD team situation – the team became frustrated since they were being asked to "solve" a problem and also prove that the problem existed in the first place.

- A new chief engineer of a nuclear engineering department identified a "laundry list" of engineering practice problems. The chief assigned a group of engineering managers to form a team, prioritize the problems and start working on fixing them.

  Good team situation – this cross-functional team tackled these important problems and made significant improvements to the design processes.

- The senior managers of a bank had just been through quality improvement training and were excited to begin improvement efforts. They assigned 10 projects to branch office and "back office" staff. The branch and "back" office managers were not consulted before these assignments were made.

  Bad situation that occurs frequently – even though it seems to hold things up, middle management MUST be engaged before their people are assigned to teams – this slowed down improvement efforts across the company.

- Factory management assigned a group of maintenance workers, purchasing and receiving personnel to work on reducing the time to obtain "non-stocked" spare parts for plant equipment. Three weeks after the team began, they realized a new parts inventory database was being installed in the next month.

  Not good, but recoverable situation – here, the team worked on other parts of the process that needed improvement, then integrated their improvements with the new database.

- A manager of a nursing unit assigned a group of nurses and nurse assistants to improve morale and communication in the unit. She thought that would help reduce turnover in the unit, which was running at 40% annually.

  The manager sets the tone for morale and communication – delegating this responsibility to a team (in our view) is abrogating her duties.

- A president of a small consulting firm had decided to expand the company office space. He assembled a team of clerical support staff to determine the best strategy to "handle the increased need for product inventory space." The team came back to him with the recommendation to let the consultants "telecommute" from their homes and use their office space for the product inventory. The president disagreed and proceeded to lease additional space.

Bad situation – the team thought they were tackling a real problem, but the president had already identified the solution. The team was VERY frustrated at this outcome and said they would never work on a team in this company again!

- A hospital initiated a number of teams whose purpose was to improve the clinical quality of care for patients. Physicians were invited to participate on these teams. Although some were initially interested, the meetings were held during the day and, gradually, the doctors stopped coming to meetings.

  Good team situation, but bad meeting situation. The hospital had to rethink their strategy for engaging the doctors in the improvement efforts.

- Nurses and quality assurance/utilization review staff were assigned to develop "Clinical Pathways," standardized methods of patient care. After reviewing the first five Pathways developed, the physicians told the chief nurse executive that the "Pathways were worthless, they weren't going to practice 'cookbook' medicine."

  Here, the doctors are important stakeholders. The hospital had failed to "sell" the idea of pathways to the doctors upfront (e.g. by engaging a physician champion who would support development and use in his/her practice).

- The corporate quality improvement department told power plant managers that they needed to have a certain number of teams "running" by years end. Over 80% of plant personnel work shifts that only allow for short breaks and lunch. By the end of the year, the only "functioning" teams were composed of administrative clerks.

  Unfortunately, this occurs all too often. The company revisited their compensation plan and agreed to pay overtime to shift workers so that they could be paid for working on improvement efforts.

- A new car dealer assigned members of his sales force to develop "best practices" for selling cars to customers. After three months of meeting, the team had not made any progress. The sales personnel are paid based on commission.

  No incentive to share practices with the current compensation program. This situation also occurs often in companies. The fundamental factor that impacts performance is beyond the scope of the team to address – often management action is necessary to remove these barriers before improvement can occur.

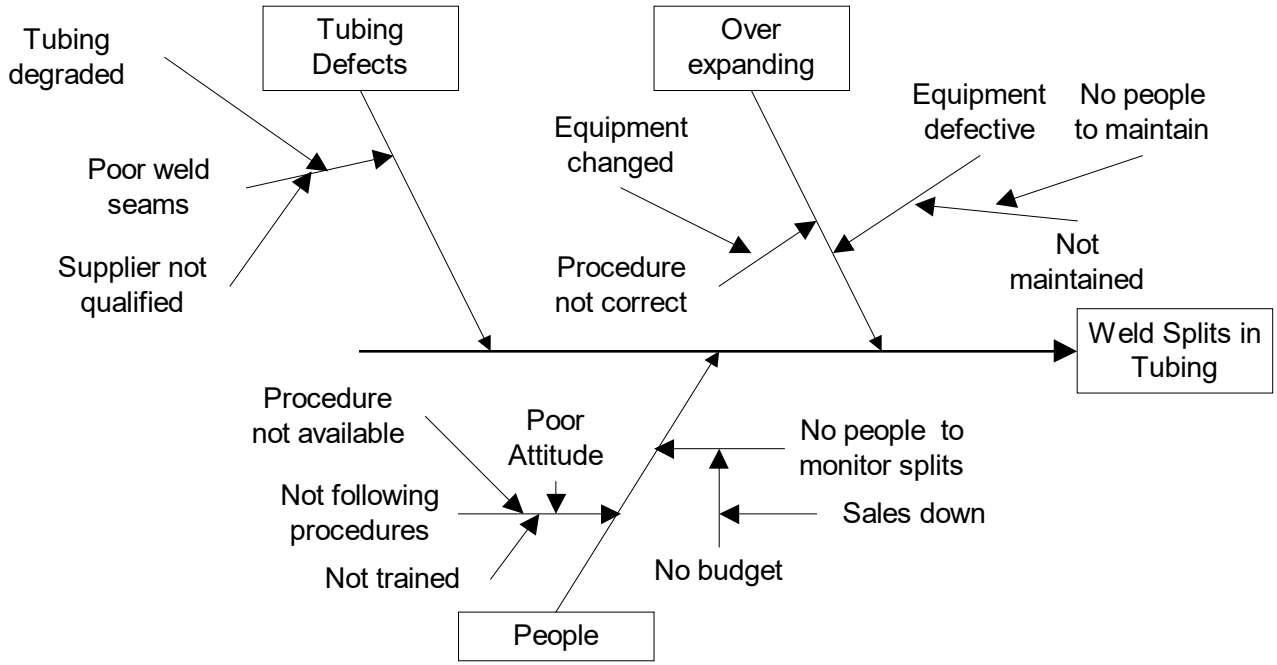
- A hospital's administration has decided to decentralize the respiratory therapy function to the patient care units. The leader of the team is the current department director. The patient care unit managers don't want the additional responsibility of respiratory therapy and the department director is reluctant to give up his "power."

  A case of lack of ownership of the solution. In this case, a senior hospital administrator had to step in to facilitate the change.

Section 5 – Process Management & Analysis

Exercise – Cause and Effect Diagram Critique

Review the cause & effect diagram below. Comment on the effect and the potential causes.



Effect: Short description of the problem. Could these splits have been stratified to a more refined problem statement? Are these occurring during fabrication or after some period of operation (the causes seem to indicate during production).

Potential Causes: Most of these causes should not appear on a good fishbone. Factors such as – *Not trained, not following procedures,* and *procedure not available* should have been addressed earlier in the process improvement effort. *Poor attitude* should also not appear on the diagram. *No people to monitor splits* is a solution – not a potential cause. *Equipment defective* is a good start, but the next "why's" should focus on the equipment failure modes/causes. *No people to maintain* is a solution – not a potential cause. The *Tubing Defects* bone could be filled in more with specific tubing defects (thin tube wall, weld cracking, etc.). Overall, this is a poor example of a cause-and-effect analysis – one typically found in a company's first few improvement efforts.

Exercise – Verification by Pareto

A team working on electrical splice failures analyzed 30 "pin and socket" splices that had failed in service. They developed the following Pareto Table and concluded that not meeting the clearance specification was the cause of the failures. They could only identify failed splices, as the remaining splices were buried underground. What could be wrong with their conclusion?

Pin and Socket Splice Failures
Pareto Table

Cause	Frequency	Cum. %
1/8" clearance specification not met	22	73.3
Unknown	3	83.3
Other	5	100.0

The team has taken a *continuous variable* and transformed it into a go/no-go variable. In doing so, they need to establish that when clearance specification is *not met*, the failure occurs (their data collection supports this). But they also need to establish that when the clearance specification is *met*, the failures *do not* occur (this is the essence of the absence/presence verification effort). They are not currently able to meet the latter requirement.

At a higher level, assuming the clearance is a continuous variable, they may have been able to develop a *correlation* between clearance and splice life. This could have provided them with a means of predicting how long a splice would survive, given a certain clearance. Correlation/regression analysis would have been the verification method.

Section 6– Measuring Performance & Variability

Objective: To develop definitions of Key Characteristics.

Instructions: 1. Determine one or more KC's for the products and services below:

Product/ Service	Customer need	Characteristic	Measure	Target	Specification(s)	Allowable Defect Rate
Fast Food	Quick Service	Service Time	Order to Food Delivery	2 min.	USL – 4 min	6 Sigma
Airline Travel	Luggage delivered to destination	Lost Luggage	% Not Arriving at Baggage Handling w/I 20 minutes of plane arrival	0%		6 Sigma
Air Conditioner	Reliability	Reliability	Time to First Failure	10 years	LSL – 5 years	4 Sigma
Hospital care	Correct medication	Medication Delivery	Correct Medication delivered on time	100%		8 Sigma
Tuna Sandwich	Taste	*Yours to Figure Out!*				

Objective: To practice calculating the Skewness and Kurtosis of a set of data.

Instructions:

1. For the set of data below, calculate the Skewness and Kurtosis values. Compare these values to those of a normal distribution.

2. Using Minitab or Excel; create a histogram of the data. Does the visual display agree with your interpretation?

18.92	18.63	16.08	16.15	20.64
18.76	16.43	19.63	16.33	26.78
18.24	14.38	17.58	22.34	17.71
18.89	22.41	15.02	16.14	17.59
24.13	22.83	17.05	17.39	39.17
19.38	10.41	23.96	12.50	19.59
19.02	11.59	23.20	26.61	19.07
20.21	21.94	28.95	13.77	22.77
10.85	11.26	17.81	16.07	15.55
10.62	18.52	16.91	16.13	21.22

Answer: From Excel (Tools>Data Analysis>Descriptive Statistics):

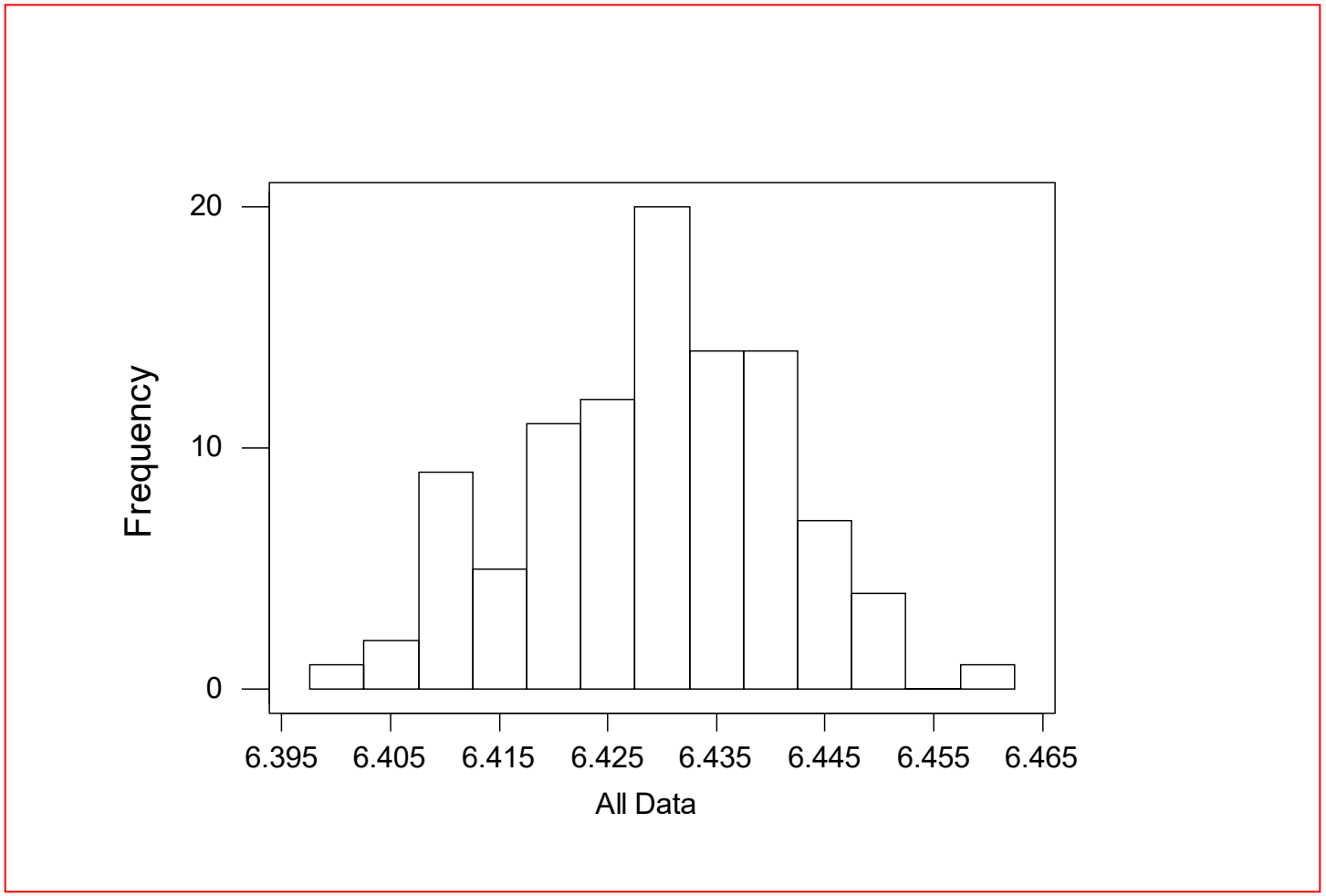
Mean	18.7426
Standard Error	0.722240551
Median	18.38
Mode	#N/A
Standard Deviation	5.10701191
Sample Variance	26.08157065
Kurtosis	**4.197544868**
Skewness	**1.31020281**
Range	28.76
Minimum	10.41
Maximum	39.17
Sum	937.13
Count	50
Confidence Level(95.0%)	1.451395846

Exercise – Histogram:

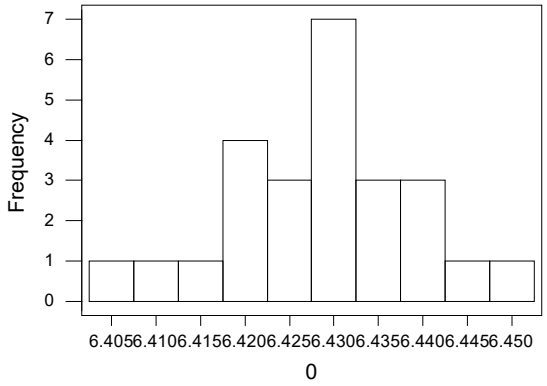
Piston Rings for Reciprocating Compressors are measured for width (in millimeters, outside diameter - inside diameter). Four measurements are taken, at 90-degree angles around the piston ring. Create a histogram of the entire data set. What does this tell you? Create histograms for each of the measurement positions. Are there any differences?

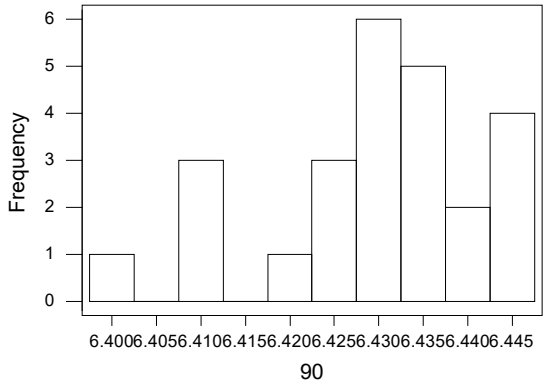
	Position (degrees)			
Ring	**0**	**90**	**180**	**270**
1	6.447	6.432	6.442	6.435
2	6.419	6.437	6.429	6.425
3	6.419	6.411	6.414	6.411
4	6.429	6.429	6.441	6.459
5	6.428	6.412	6.443	6.436
6	6.440	6.435	6.409	6.438
7	6.415	6.430	6.410	6.433
8	6.435	6.444	6.430	6.411
9	6.427	6.437	6.424	6.420
10	6.423	6.445	6.424	6.437
11	6.428	6.444	6.438	6.431
12	6.431	6.425	6.422	6.432
13	6.422	6.437	6.417	6.447
14	6.437	6.432	6.410	6.438
15	6.425	6.440	6.422	6.450
16	6.407	6.431	6.421	6.418
17	6.438	6.400	6.439	6.440
18	6.435	6.412	6.427	6.448
19	6.431	6.420	6.433	6.424
20	6.412	6.427	6.436	6.440
21	6.452	6.442	6.450	6.424
22	6.420	6.431	6.413	6.403
23	6.429	6.447	6.439	6.432
24	6.428	6.427	6.420	6.432
25	6.442	6.434	6.413	6.429

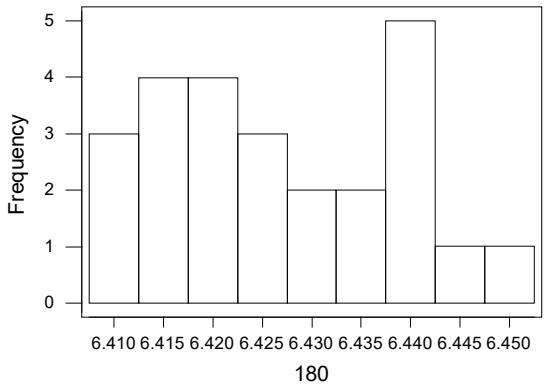
When all the data is plotted, the distribution appears to be normally distributed:

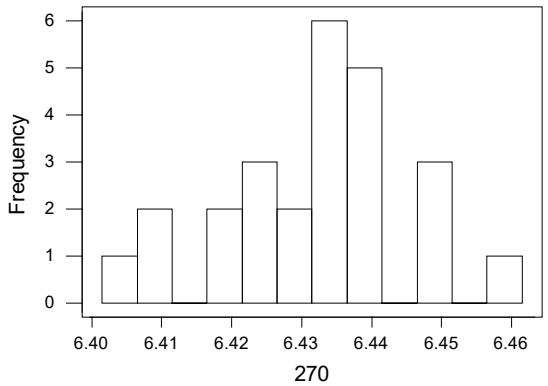


Stratifying the data reveals a slightly different picture. Positions “0” and “270” appear symmetric, however, positions “90” and “180” appear skewed (90 is skewed negatively; 180 is skewed positively). The process investigation may profit by focusing on these areas.









Exercise – Combining Variability:

Three components of a valve stem/gate assembly are produced. What is the expected length and standard deviation of the assembly? The three components are welded together in series:

Component	Mean	Std. Dev.
Valve Stem	18.00”	0.03”
Valve Disk	8.00”	0.02”
Valve Guide	4.00”	0.02”

If the specification calls for the assembly to be no longer than 30.10 inches, what is the current manufacturing process capability (C_p) of meeting the spec? (Note: typically, if the average plus/minus ***3*** times the standard deviation is within the spec limits, the process is considered OK).

Employing Additivity of Variances:

$$\overline{X}_{Valve} = 18.00 + 8.00 + 4.00 = 30.00"$$

$$s_{Valve} = \sqrt{0.03^2 + 0.02^2 + 0.02^2} = 0.041"$$

$$C_p = \frac{30.10 - 30.00}{3 \times 0.041} = 0.81 \text{ Not Very Good!}$$

Objective: To develop and interpret an X-Bar, R control chart

Instructions:

1. Develop an X-Bar, R control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form. Interpret the control chart – are there assignable causes present?

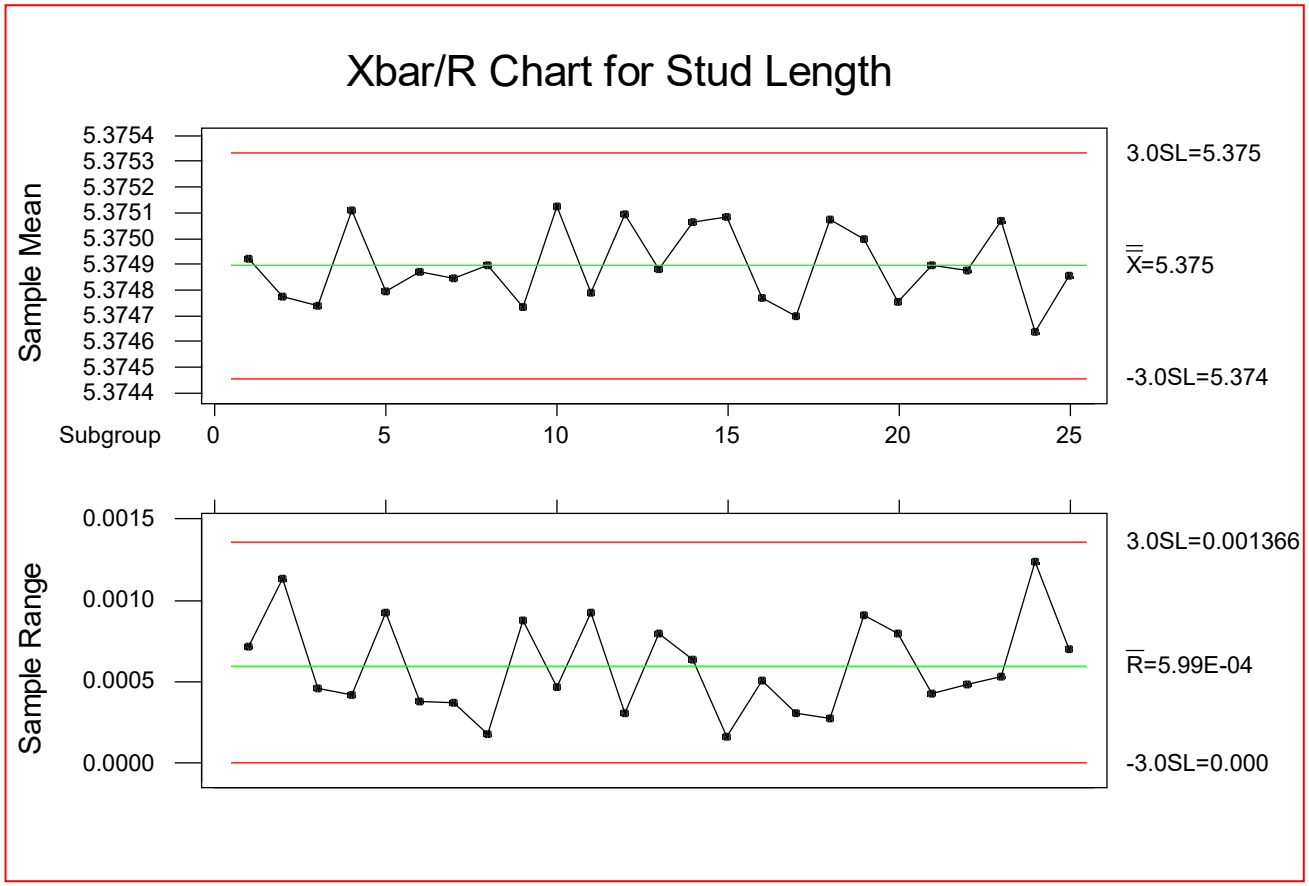
2. Open Mini-Tab on your PC. Create the X-Bar, R control chart using the STUDS data file. Compare results from Mini-Tab to your hand-drawn chart.

Time: 40 minutes

Compressor Stud Lengths - A supplier fabricates studs for critical compressor applications. One key quality characteristic of the studs is their length. Their customer specifications call for a nominal value of 5.3750" with a tolerance of +/- 0.0005". The supplier pulls a subgroup of four studs each hour from the fabrication process and measures their length with a calibrated micrometer. In the table below, each row is a subgroup.

Subgroup	Stud Length (in.)			
1	5.37526	5.37525	5.37454	5.37464
2	5.37478	5.37525	5.37495	5.37411
3	5.37446	5.37476	5.37482	5.37492
4	5.37525	5.37485	5.37527	5.37506
5	5.37463	5.37430	5.37502	5.37523
6	5.37511	5.37473	5.37486	5.37478
7	5.37473	5.37475	5.37510	5.37480
8	5.37484	5.37497	5.37480	5.37498
9	5.37520	5.37457	5.37432	5.37484
10	5.37534	5.37487	5.37511	5.37517
11	5.37472	5.37433	5.37526	5.37486
12	5.37502	5.37501	5.37532	5.37502
13	5.37475	5.37542	5.37462	5.37473
14	5.37482	5.37529	5.37539	5.37475
15	5.37499	5.37504	5.37515	5.37515
16	5.37464	5.37509	5.37458	5.37476
17	5.37465	5.37487	5.37456	5.37472
18	5.37515	5.37492	5.37504	5.37519
19	5.37440	5.37531	5.37504	5.37525
20	5.37436	5.37475	5.37516	5.37474
21	5.37493	5.37514	5.37471	5.37481
22	5.37463	5.37467	5.37511	5.37510

Subgroup	Stud Length (in.)			
23	5.37511	5.37510	5.37530	5.37477
24	5.37436	5.37401	5.37525	5.37493
25	5.37483	5.37493	5.37448	5.37518



Minitab was used to generate this control chart (Stat>Control Charts>X-Bar, R). No Assignable Causes noted for this process

Objective: To practice developing and interpreting an X-Bar, S Control Chart.

Instructions:

1. Develop an X-Bar, S control chart for the data on the following pages. Perform the calculations by hand; plot the points and limits on the control chart form.

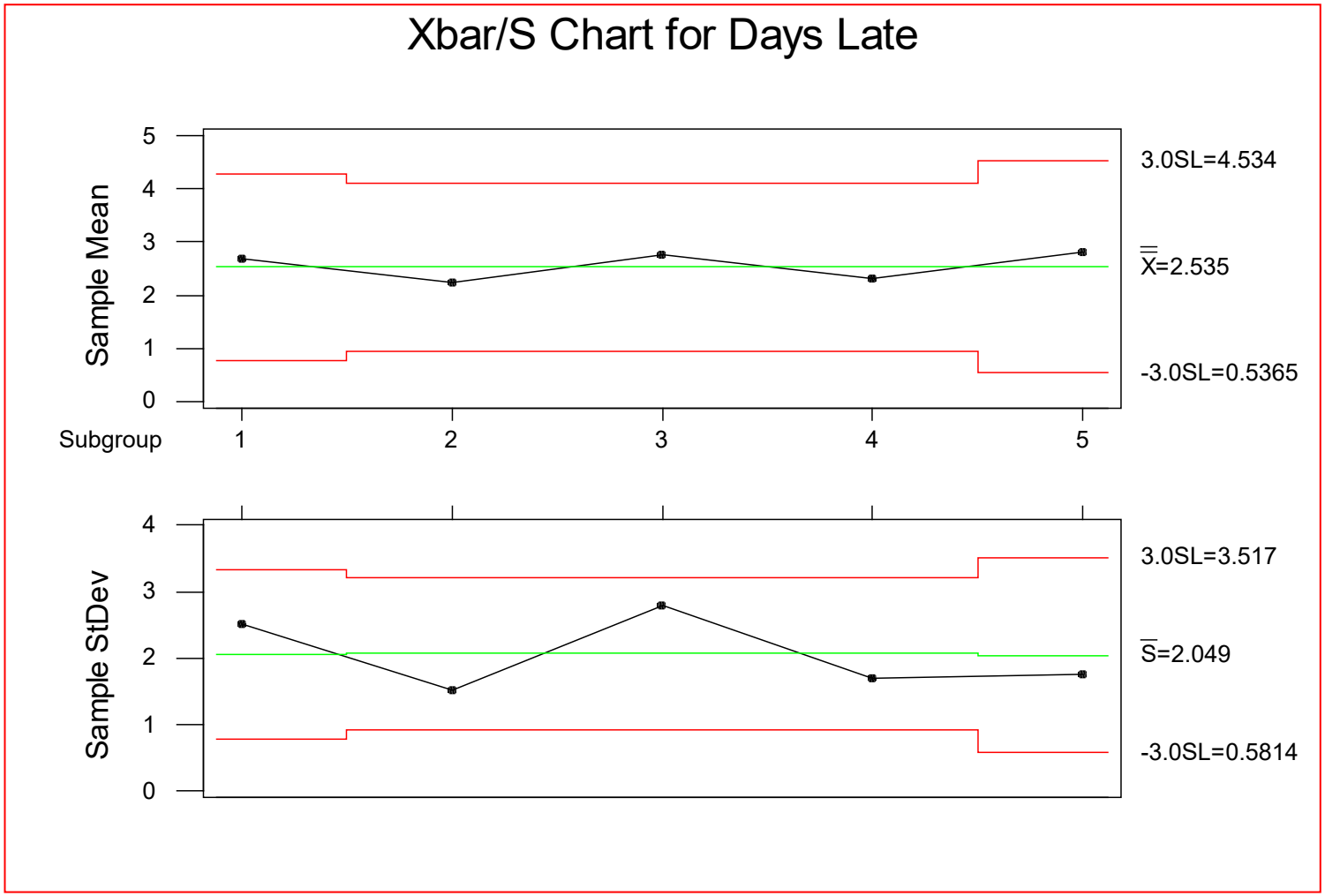
2. Open Mini-Tab on your PC. Create the X-Bar, S control chart using the DELIVERY data files. Compare results from Mini-Tab to your hand-drawn charts.

Time: 40 minutes

Delivery Schedule Data by Month - In order to provide better on time delivery and to increase throughput a plant is monitoring DAYS LATE FOR DELIVERY. Take the following delivery data for Unit 1075B and prepare an X-Bar, S Control Chart. Each data point is for a late unit. Interpret the data when the month is used as a subgroup strategy:

APR	MAY	JUN	JUL	AUG
1	1	1	1	2
4	1	3	1	3
4	1	1	1	2
10	1	2	6	2
3	2	12	2	7
1	4	1	5	3
2	2	4	2	3
1	4	1	2	1
1	2	1	1	1
2	1	1	1	4
4	3	5	4	
1	1	4	1	
1	1	2	1	
	4	1	4	
	2	3	4	
	6	2	1	

Minitab was used to generate this chart (Stat>Control Charts>X-Bar, S). The data from this process shows no assignable causes:



Objective: To practice developing and interpreting an X, mR Control Chart.

Instructions:

1. Develop an X, mR control chart for the data below. Perform the calculations by hand, plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the X, mR control chart using the VALVE data file. Compare results from Mini-Tab to your hand-drawn charts.

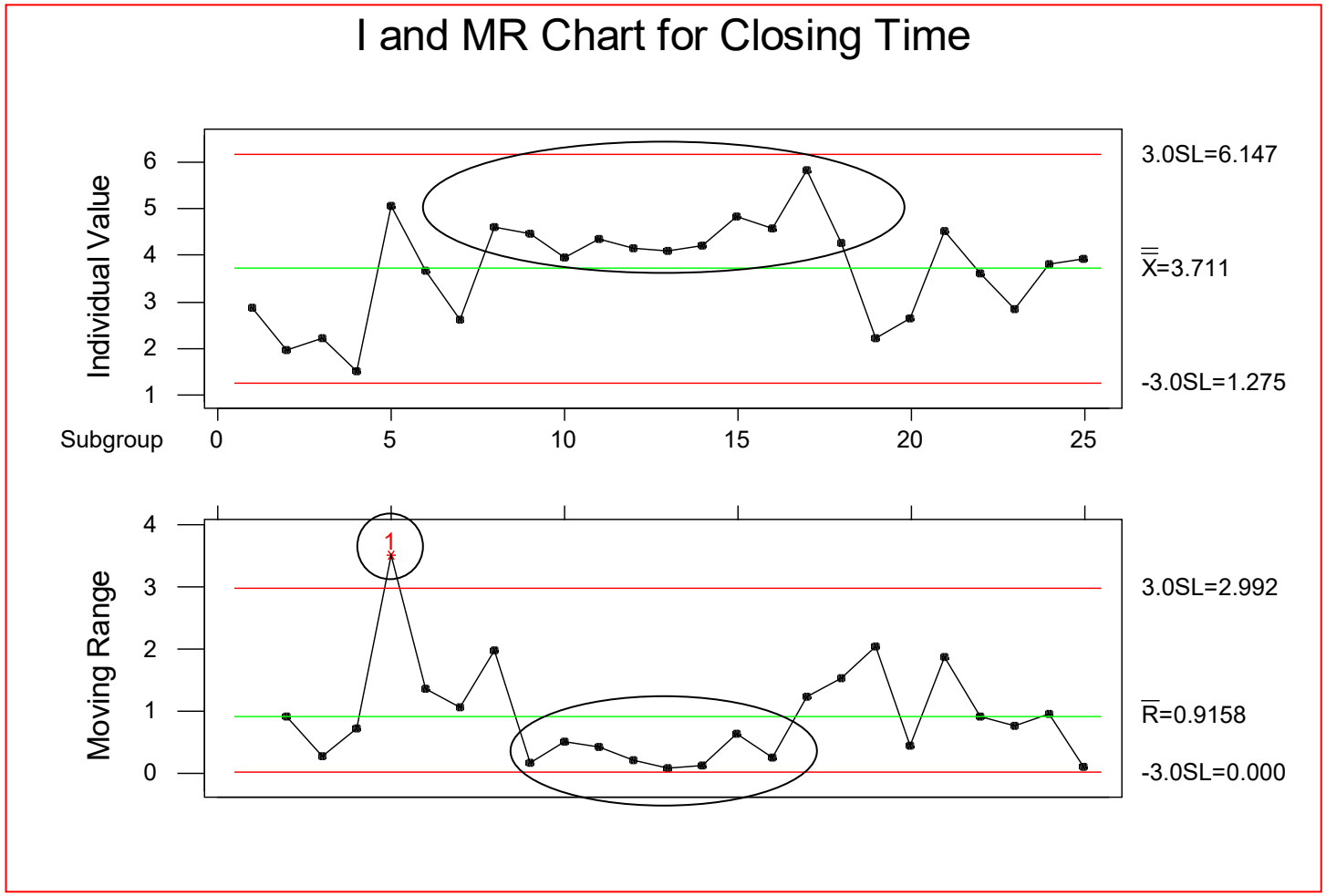
Time: 25 minutes

Butterfly Control Valve - An air-operated butterfly control valve is used to control cooling water flow to heat exchangers in an air conditioning unit. The valve must close within ten seconds of receipt of the signal from the unit's protective circuitry. The valve is tested monthly and maintenance technical personnel record its closing time (in seconds):

2.87	4.2
1.96	4.82
2.22	4.58
1.51	5.81
5.04	4.27
3.67	2.22
2.62	2.65
4.61	4.52
4.46	3.62
3.95	2.86
4.36	3.81
4.16	3.91
4.08	

Minitab was used to generate this chart (Stat>Control Charts>I, mR). The closing time for this process shows several assignable causes:

- Range Chart – Point outside of control (labeled 1)
- Range Chart – Eight points below the center line
- Individuals Chart – Eleven points above the center line



Objective: To develop and interpret an *np* control chart

Instructions:

1. Develop an *np* control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *np* control chart using the MOTOR data file. Compare results from Mini-Tab to your hand-drawn charts.

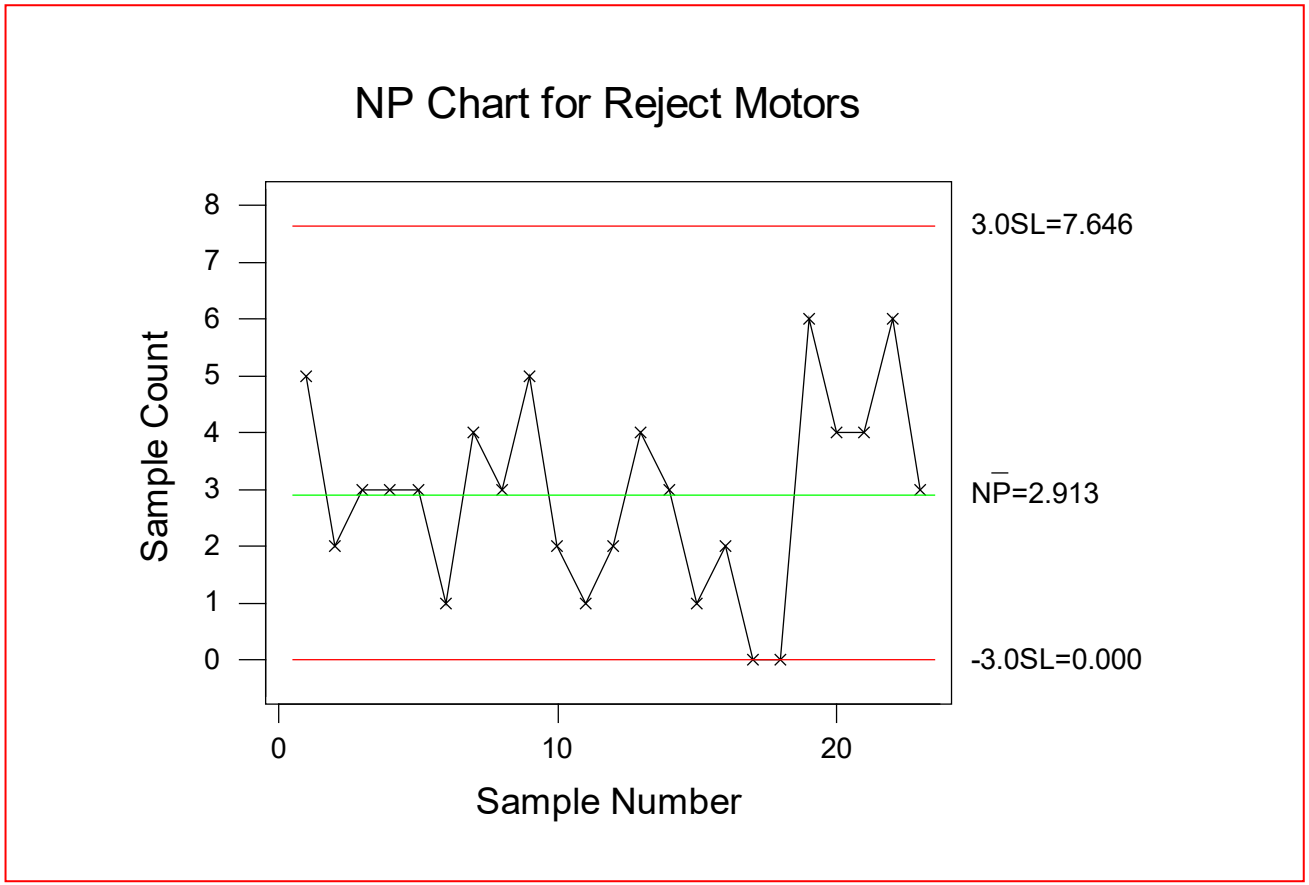
Time: 30 minutes

Motor Rejects - A company that produces air conditioning units orders batches of the motors from a supplier. Due to past quality problems, the company inspects 20 motors from each batch. Each motor is accepted or rejected. Based on the number of motors rejected, a decision is made to either inspect the remaining motors or return the batch to the supplier for rework.

Number of Rejected Motors:

5	4
2	3
3	1
3	2
3	0
1	0
4	6
3	4
5	4
2	6
1	3
2	

Minitab was used to generate this chart (Stat>Control Charts>np). No assignable causes are evident from this data.



Objective: To develop and interpret a *p* control chart

Instructions:

1. Develop a *p* control chart for the data on the following pages. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *p* control chart using the WELDS data file. Compare results from Mini-Tab to your hand-drawn charts.

Time: 30 minutes

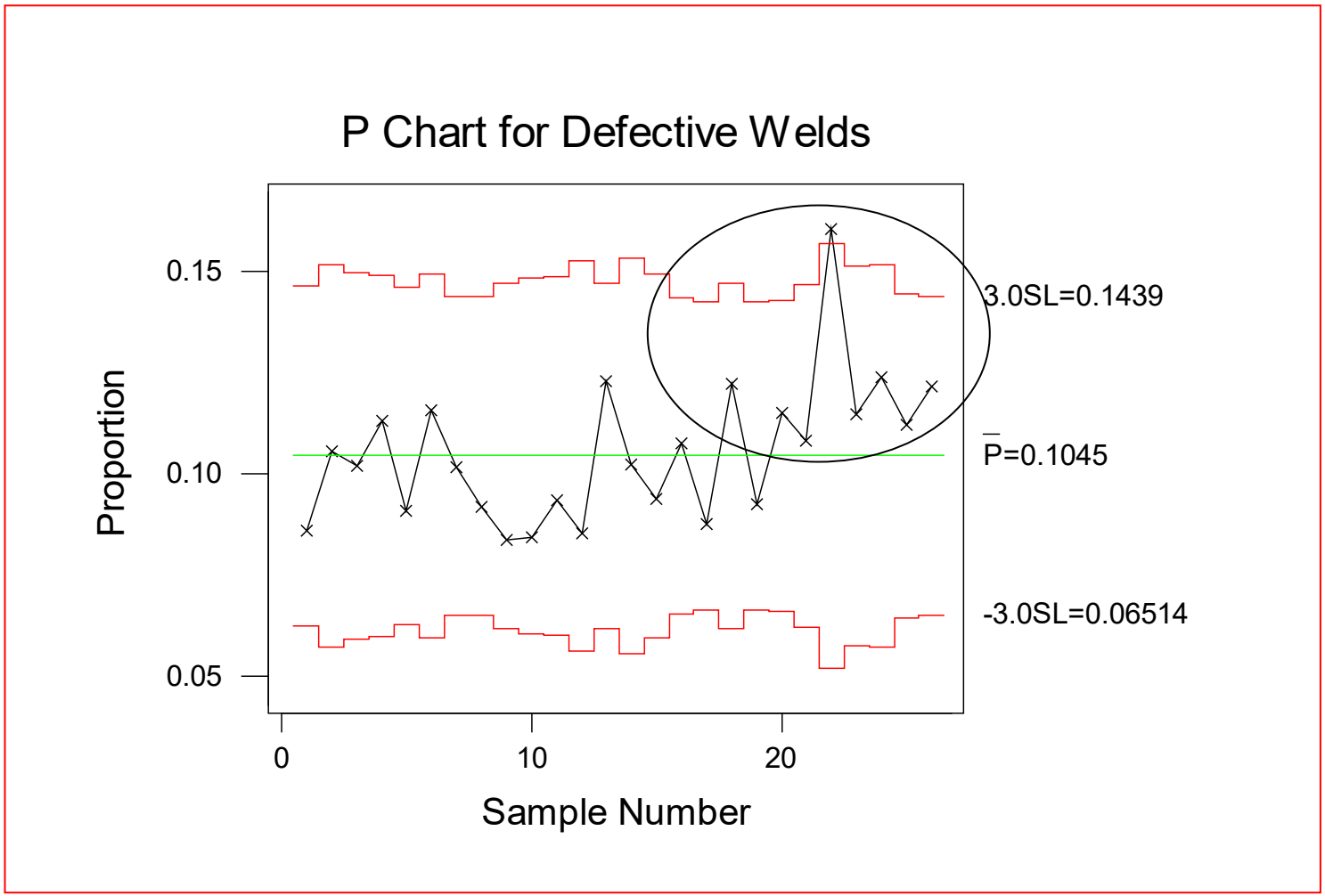
Defective Full-Penetration Welds

A welding supervisor receives inspection reports by the Quality Control Department. The QC supervisor has recently called his attention to a seemingly high number of rejected full-penetration welds on critical high pressure piping systems. The welding supervisor begins his analysis of the situation by preparing a p-chart of rejected welds for the past six months.

Week	# Welds	# Defective		Week	# Welds	# Defective
1	476	41		14	352	36
2	379	40		15	415	39
3	412	42		16	557	60
4	424	48		17	581	51
5	483	44		18	466	57
6	415	48		19	584	54
7	541	55		20	573	66
8	544	50		21	471	51
9	466	39		22	305	49
10	439	37		23	383	44
11	428	40		24	379	47
12	363	31		25	526	59
13	463	57		26	543	66

Minitab was used to generate this chart (Stat>Control Charts>p). Two assignable causes are noted:

- Point outside the upper control limit
- A run of seven points above the center line



Objective: To develop and interpret a *c* control chart

Instructions:

1. Develop a *c* control chart for the data on the following pages. Perform the calculations by hand, plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *np* control chart using the PINHOLES data file. Compare results from Mini-Tab to your hand-drawn charts.

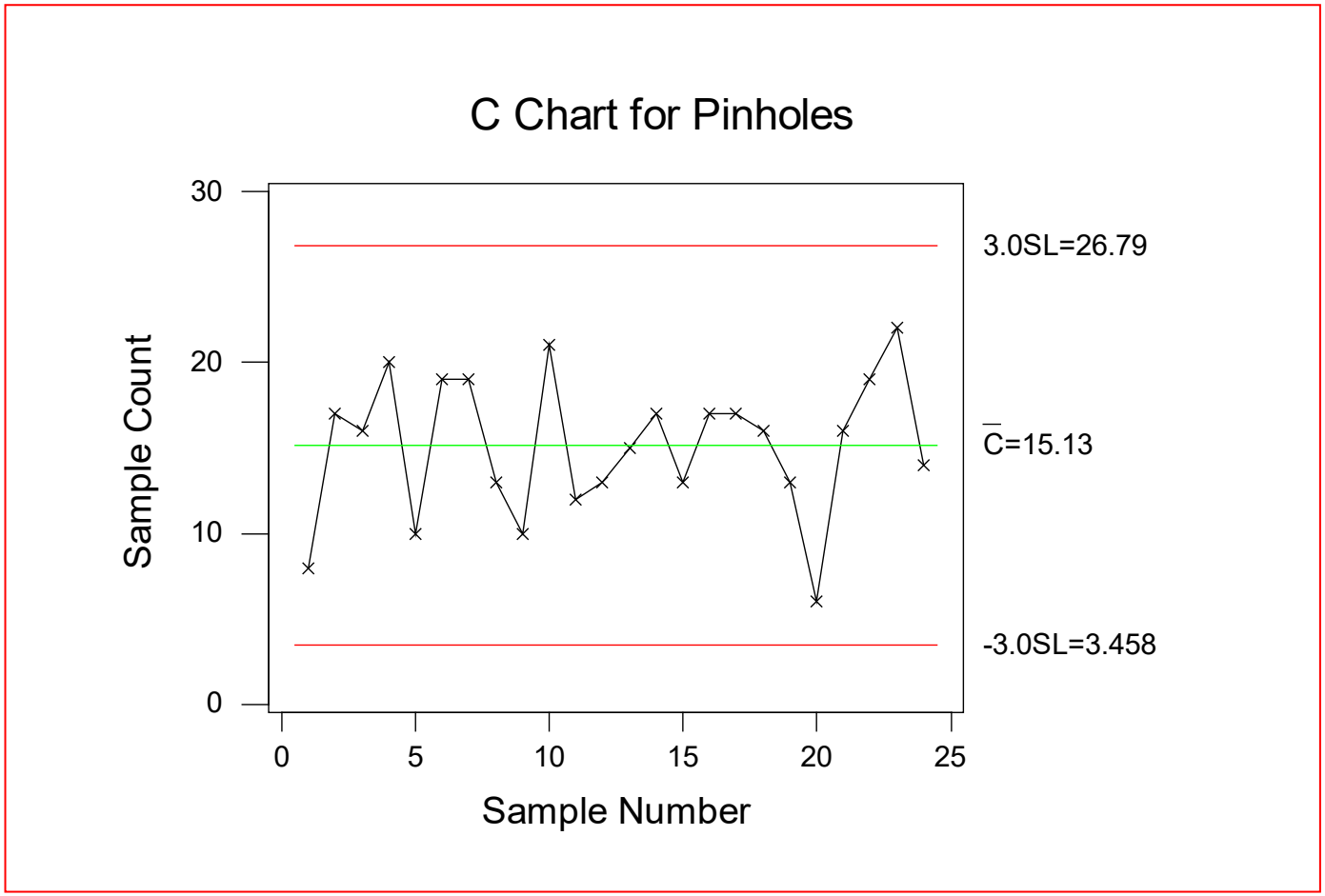
Time: 25 minutes

Ceramic Paint Pinholes - A paint manufacturing company, which produces special paints used by hobbyists on ceramics, tests samples of their paint daily. They apply the paint to unfired ceramic plates, fire the plates in a kiln and then inspect the finished plates. Among other defect categories, they count the number of pinholes in each sample. The test manager has recently begun to track the number of pinholes obtained from each sample on a c chart.

Number of Pinholes per Sample:

18	14
8	15
17	17
16	13
20	17
10	17
19	16
19	13
13	6
10	16
21	19
12	22
13	14

Minitab was used to generate this chart (Stat>Control Charts>c). No assignable causes are noted for this process.



Objective: To develop and interpret a *u* control chart

Instructions:

1. Develop a *u* control chart for the data below. Perform the calculations by hand; plot the points and limits on the control chart form.

2. Open Mini-Tab on your PC. Create the *u* control chart using the same data using the DCRs data file. Compare results from Mini-Tab to your hand-drawn charts.

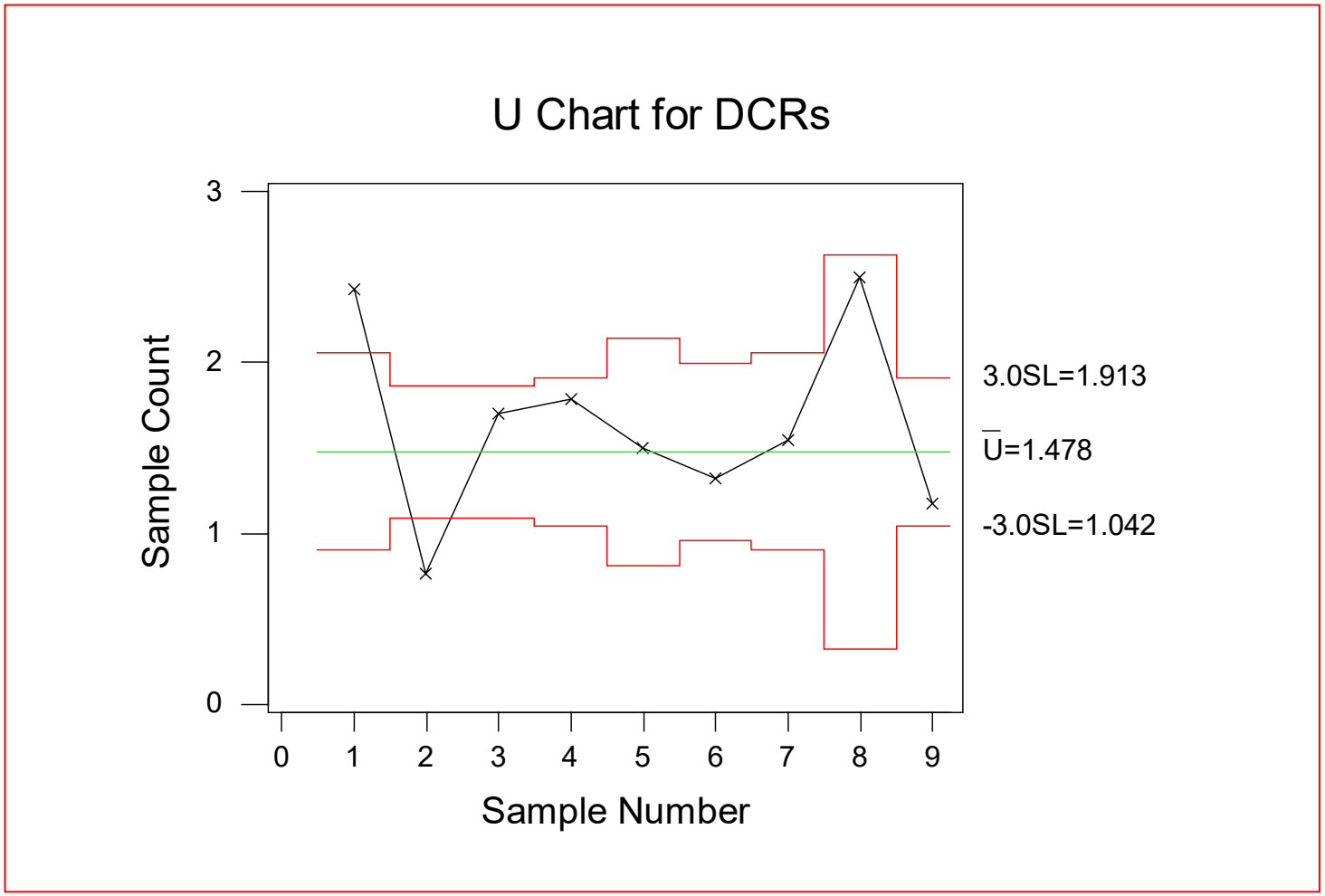
Time: 25 minutes

Design Change Requests (DCRs) - Engineers are responsible for developing custom designs of air conditioning systems. As they are built, manufacturing discovers problems with the designs and requests changes from Engineering (Design Change Request). The Engineering Manager was curious to see if there were significant differences between the engineers.

# Units	# DCRs	Engineer
40	97	Maynard
90	69	Kinney
90	153	Gibbs
70	125	Nichols
30	45	Fritz
50	66	Stone
40	62	Fielding
10	25	Adams
70	82	Pelham

Minitab was used to generate this chart (Stat>Control Charts>u). Two assignable causes are noted:

- Engineer 1 (Maynard) "produces" a high number of DCRs.
- Engineer 2 (Kinney) "produces" a low number of DCRs.



Pick-a-Chart (Control Chart Selection)

Often, one of the difficulties people face with control charts is the question of "Which is the right one?" This exercise is intended to give you some practice going through the logic of the Control Chart Selection Guide (Unit 4.4).

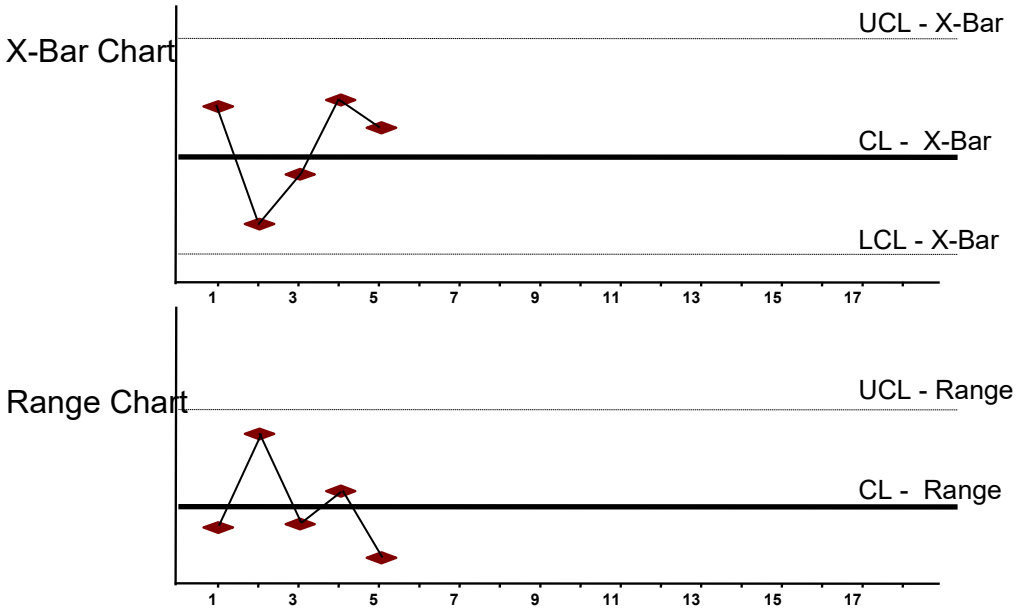
As you develop your answer, note your assumptions. There is more than one way many of these scenarios could be charted. We'll start out with some "warm-up" exercises, move on to more complex situations:

Scenario	Control Chart(s)
1. Each day, the number of units shipped is counted at 12:00 AM.	Xmr or c
2. The Sales department keeps track of the number of units sold each day, by type of unit.	Xmr or c
3. A laboratory gets a report, once a day, which provides the number of samples analyzed, the average processing time, and the standard deviation of the processing time.	Number – Xmr, or c Time – X-bar, S
4. Each day, a technician measures the time she takes tubing a condenser on her shift.	Xmr
5. At a ballpark, a hot dog vendor counts the number of wieners sold each game day.	Xmr or c
6. A factory worker measures the diameter of valve stems after machining. She takes four stems at random from each hour's production.	X-bar, R
7. An engineer measures the cycle time for engineering change orders weekly.	Xmr
8. An administrative assistant tracks the number of days it takes customers to pay their bills. She keeps a chart for each of the company's top 6 customers.	Xmr (for each customer)
9. A quality consultant tracks the number of days she is on the road each month.	Xmr or p
10. A Sales supervisor has developed control charts for her clerks - they track the number of line items entered each day.	Xmr or c
11.The power of a motor is measured and is subject to a purchase specification.	X-bar, R or Xmr
12. Coatings are purchased in tank car lots. The material is sampled and the chemical composition determined.	X-bar, R or Xmr
13. The procedures group has noticed an increase in the number of comments made on their draft procedures being circulated for review. They are wondering if something unusual is going on in the procedure drafting/ review process.	U
14. The LAN (Local Area Network) administrator has been trying to improve the reliability of the system. She is interested in seeing if the number of LAN "crashes" has decreased.	C
15. This same LAN administrator has also been working on trying to reduce the time required to restore the LAN after it crashes.	Xmr
16. The Production Manager is interested in employee absenteeism, measured in days/employee. The corporate staff supplies her with a monthly report, which breaks down this measure into weekly increments.	Xmr
17. The Financial Officer is concerned about the utilization of company cars; he suspects that there are too many cars. He begins tracking the number of hours the cars are utilized for business purposes each week.	Xmr
18. A production facility wishes to improve the set-up time required when products being produced are changed. They usually make the same product for about two days and then switch over to another product.	Xmr, X-bar, R or S
19. A bolt manufacturer must ensure that the tensile strength of stainless steel bolts meets the customers' specifications. About 5000 bolts are produced daily.	X-bar, R or S

Scenario	Control Chart
20. You have been troubled by the number of times your production facility has been stopped due to power interruptions by the local utility. You have records of all production stoppages for the last two years.	C
21. A certain vendor provides you with bolts for your product. Before the bolts are used in your production process, you sample 50 from each box of 1000 and inspect them for defects.	Np
22. A large consulting firm prepares about 30 proposals per week for prospective clients. The Sales Department manager is interested in the number of proposals that are not accepted by clients.	Np or p
23. An engineering department prepares design changes to improve the performance of a chemical processing plant. They are interested in the number of *field change requests*, those changes that are requested by construction engineering because the design change cannot be implemented in the field.	U
24. A Sales Manager tracks weekly sales volumes by number of items sold, dollar amount of sales and items sold per salesperson.	Number – Xmr or c $ Amount – Xmr Items – Xmr or c
25. An Automotive Manager is concerned about the quality of a particular brand of tire used on company cars. His primary concern is the possibility of a tire blowout. If the size of the company car fleet stays constant, how should he track this process?	Np or c
26. A Records department director is concerned about the errors made by her staff. She asks for help in determining the best chart to use. She tells you that the number of records varies significantly from week to week.	U
27. Each month you receive a departmental budget variance report that, among other things, provides the dollar amount you are over or under salary budget, supply expense budget and overtime hours.	Xmr
28. A physician thinks that the complications associated with a particular surgical procedure varies from surgeon to surgeon. Each surgeon does a different number of these procedures each year.	P
29. A company is interested in using a new vendor to supply control circuits that emit a specified signal for a specified time. They wish to determine if the process used to produce the circuits is in control. They are particularly interested in the signal's duration.	X-bar, R

Control Chart Application Example

Consider the following scenario. A manufacturing plant runs a two-shift operation. Ten parts are produced each shift. The process control plan calls for maintaining the current process settings until the control chart displays assignable cause signals. All parts are measured and plotted *real-time* on an X-Bar, R control chart, with subgroup size = 2. The following control chart shows the data from the last shift. The control limits are based on the previous 5 days of production, not including the last shift. For each of the three scenarios described below, discuss and predict what the data would look like on the chart.



Scenario 1 – The instrument is checked at the beginning of the second shift. Due to a bias noted against the plant's standard, the gauge is adjusted noticeably higher (e.g. for a part previously measured to be 0.900", the new reading would be 1.000"), prior to production. Sketch the next two shifts.

Scenario 2 – With the second shift, a new gauge is introduced. Compared to the old gauge, the bias is the same, but the gauge variation is much less. Sketch the next two shifts.

Scenario 3 – A new operator starts work on the second shift. He tends to read the gauge lower than the other operators, although no one knows this since a measurement system study has not been performed. Sketch the next three shifts.

- **Scenario 1** – If the production process doesn't change, the next subgroup (i.e. next two parts produced) is likely to be out-of-control (high) on the X-Bar chart (the Range chart shouldn't be affected). If the operator reacts to the assignable cause, the production process will be adjusted before the next parts are produced and the remaining subgroups will fall within the control limits (assuming the adjustment re-centers the process).

- **Scenario 2 –** The gauge variation appears in the range chart. If the gauge variation is much less, the measured variation between parts will be less. An out-of-control signal will appear (a run of points below the center line of the range chart). If a seven-in-a-row rule is employed, the out-of-control signal will be called after the parts 3 and 4 of the third shift are produced. If the assignable cause is identified to be the gauge, the control limits can then be recalculated to reflect the "reduced" process variation.

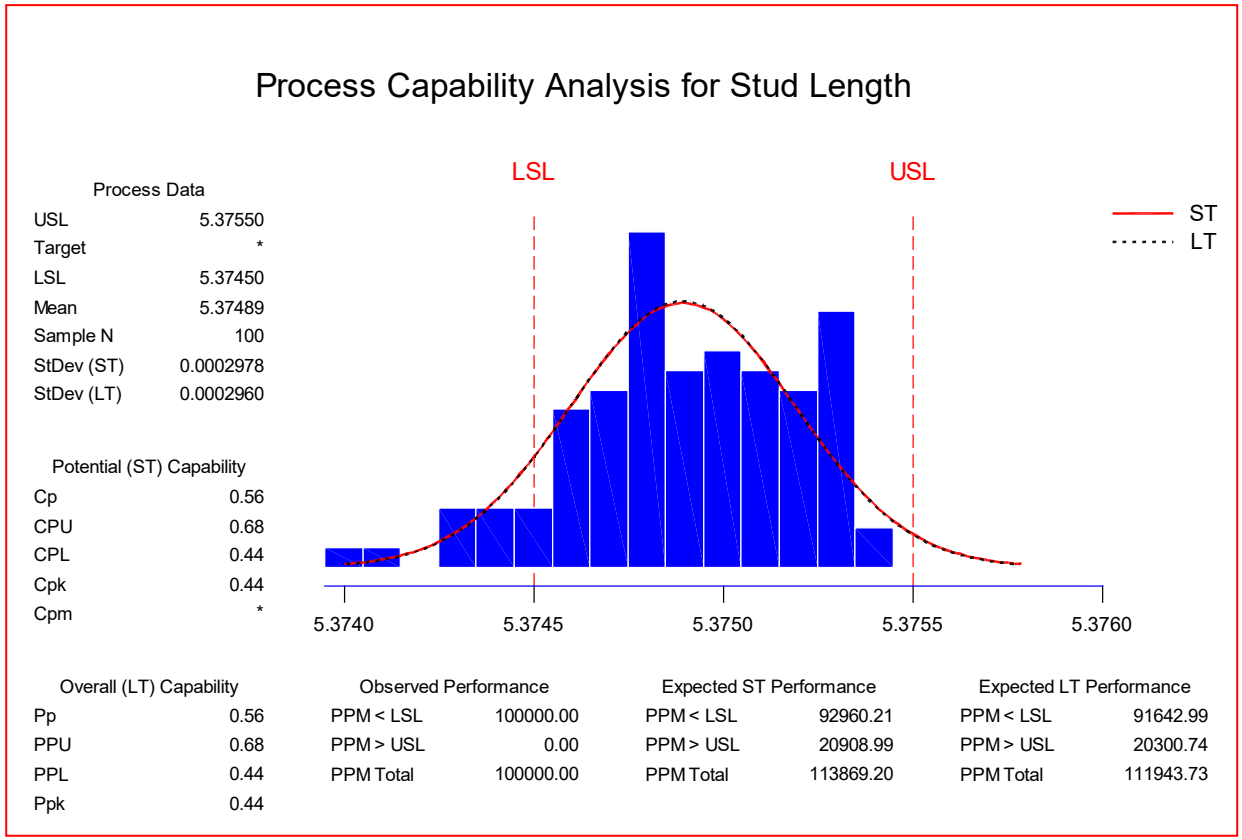
- **Scenario 3 –** This scenario is basically the opposite of the first. The first subgroup will likely show up as an out-of-control signal on the low side of the X-bar chart (note that it may show up as a run below the center line if the operator bias is not high). The search for the assignable cause should identify the operator as the key variable. The action to take should be to "recalibrate" the operator, not adjust the production process.

Objective: To perform basic capability calculations.

Instructions: 1. For the control chart exercises above, develop the picture of process capability, calculate Cp, Sigma and, if the picture indicates the need, Cpk.

Description	Specification Limits
Compressor Stud Length (page 13)	5.3750" +/- 0.0005"

Minitab: Stat>Quality Tools>Capability Analysis (Normal)



Process Data

USL	5.37550
Target	*
LSL	5.37450
Mean	5.37489
Sample N	100
StDev (ST)	0.0002978
StDev (LT)	0.0002960

Potential (ST) Capability

Cp	0.56
CPU	0.68
CPL	0.44
Cpk	0.44
Cpm	*

Overall (LT) Capability

Pp	0.56
PPU	0.68
PPL	0.44
Ppk	0.44

Observed Performance

PPM < LSL	100000.00
PPM > USL	0.00
PPM Total	100000.00

Expected ST Performance

PPM < LSL	92960.21
PPM > USL	20908.99
PPM Total	113869.20

Expected LT Performance

PPM < LSL	91642.99
PPM > USL	20300.74
PPM Total	111943.73

Objective: To understand the concept and calculation of process yield.

Instructions: 1. For the following process, calculate the First Pass Yields of the process steps, the Normalized Yield of the overall process and Rolled Through-put Yield of the process. Based on the number of defects detected through inspection, calculate the Final Pass Yield of the process. See the AXLE worksheet on your Excel file for the data.

Time: 20 minutes

Axle Production - The following steps are performed to manufacture this axle:



Process Step	# of Units Produced	# of Defect Opportunities	# of Defects Produced	First Pass Yield	# Detected Prior to Customer	Final Pass Yield
1. End Face Milling	10,000	1	82	0.9918	82	1.0000
2. Rough Machining (Lathe)	10,000	1	25	0.9975	25	1.0000
3. Finish Turning (Lathe)	10,000	1	650	0.9350	500	0.9850
4. Axle End Gear Cutting	10,000	6[1]	235	0.9961	120	0.9981
5. Cleaning	10,000	1	3	0.9997	3	1.0000
6. Heat Treat/Quenching	10,000	1	100	0.9900	10	0.9910
7. Axle Grinding	10,000	1	140	0.9860	20	0.9880
8. Flange Machining (Automatic Lathe)	10,000	1	5	0.9995	3	0.9998
9. Axle Flange Drilling (6 Holes)	10,000	6[1]	256	0.9957	30	0.9962

Notes: 1. These operations are applied six times to the axle.

First Pass Yield: $Y_{FP} = 1 - \frac{d}{n \times o}$ (See Table for Calculations)

Normalized Yield: $Y_{NORM} = 1 - \frac{\sum_i d_i}{n \times \sum_i o_i} = 0.9921$

Rolled Throughput Yield: $Y_{RTP} = \prod_i Y_{FPY-i} = 0.895$

Final Pass Yield: $Y_{FINALPASS} = 1 - \frac{d - d'}{n \times o}$ (See Table for Calculations)

Objective: To practice improving the information quality of data arising from sporadic events.

Instructions:
1. Chart the information below on a *c* chart (see the INJURY worksheet on your Excel file for the data).
2. Apply the lessons of the Sporadic Events special topic to improve the information "quality."

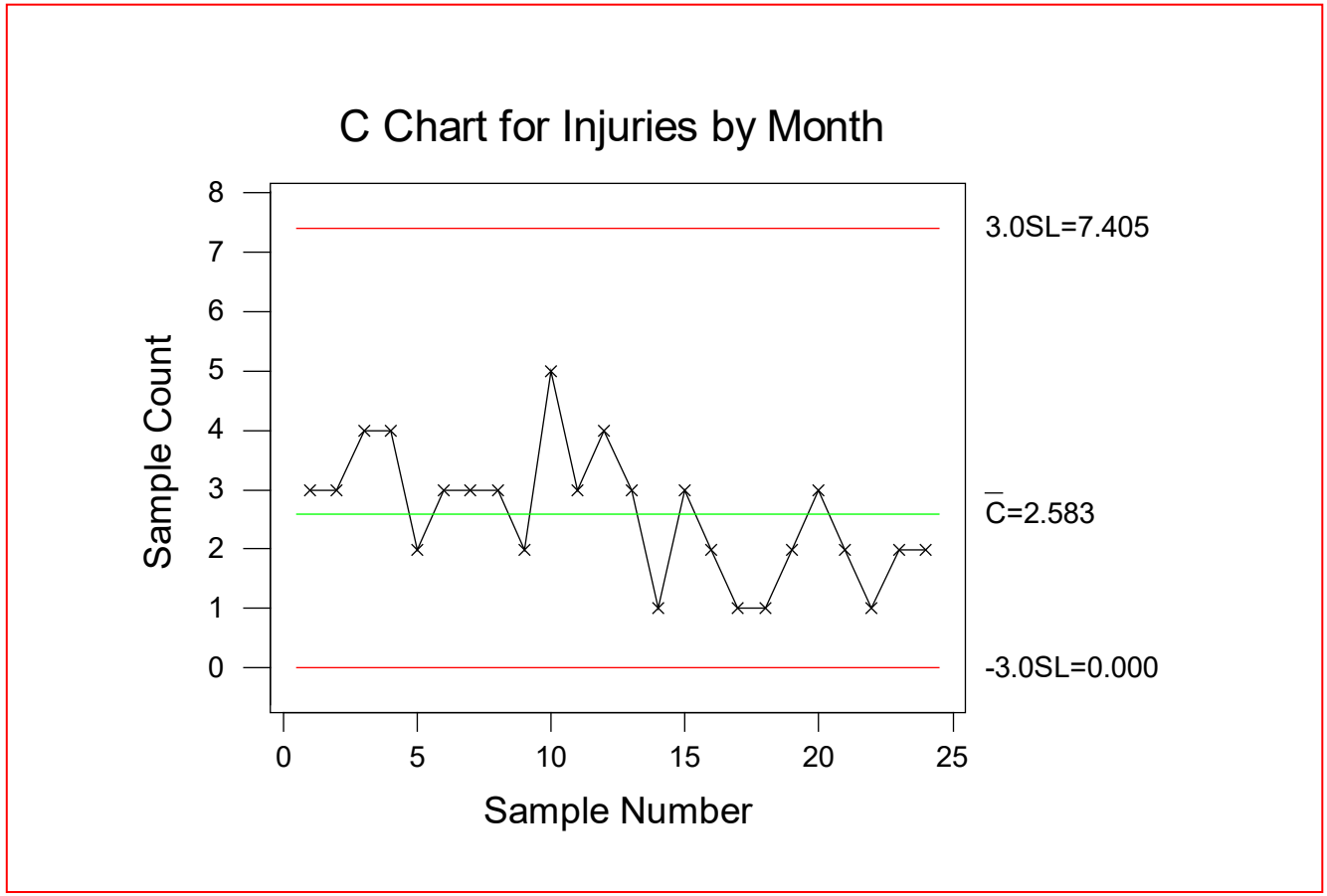
Time: 30 minutes

Employee Injuries - A manufacturing plant has been working on reducing employee injuries through root cause analysis and corrective actions on the processes "producing" injuries. At the beginning of the year (1998), the plant put several countermeasures in place to address back injuries. Have these helped reduce this class of injury?

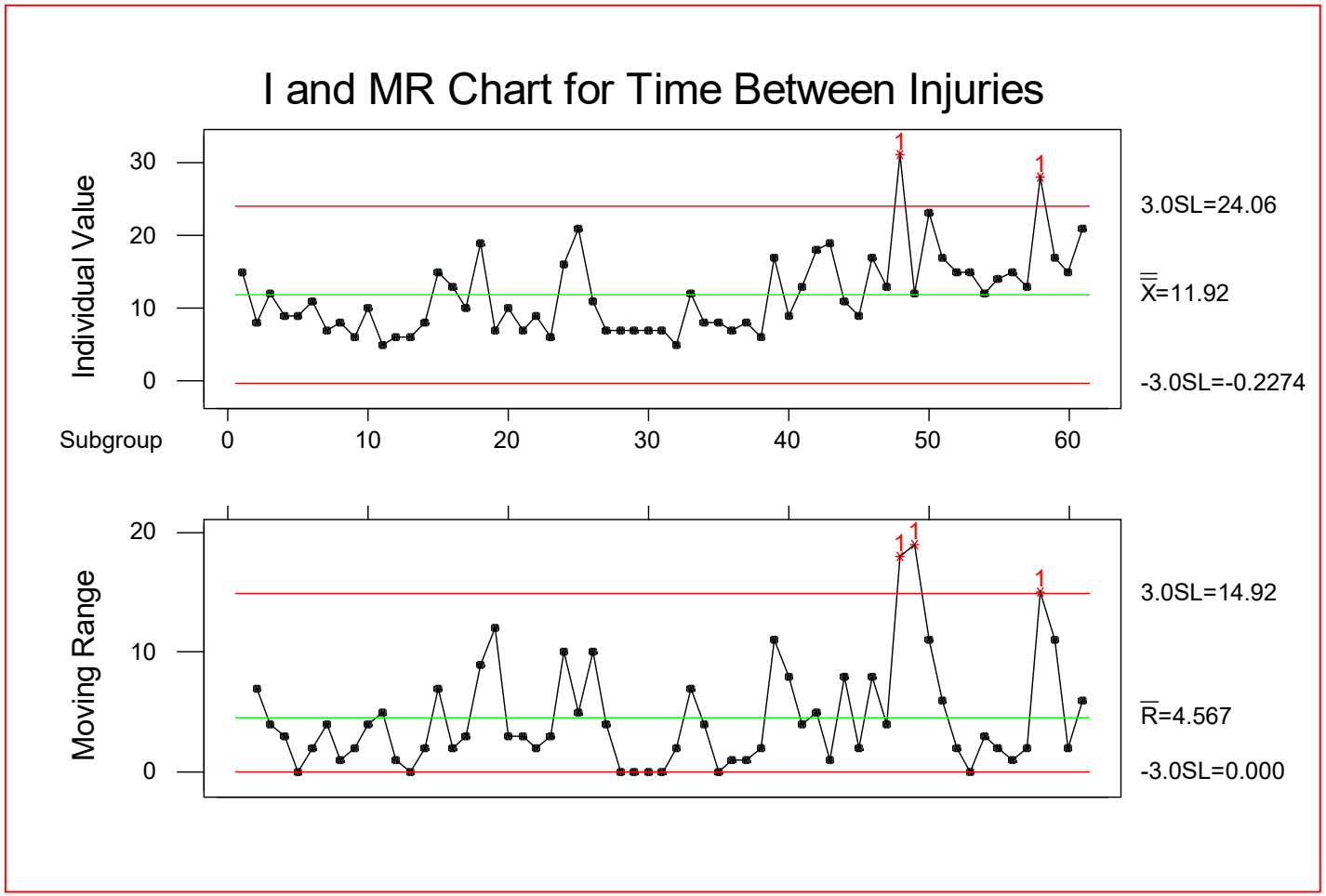
1998		
Month	**# of Injuries**	**Date(s) of Injury**
Jan	3	5-Jan, 20-Jan, 28-Jan
Feb	3	9-Feb, 18-Feb, 27-Feb
Mar	4	9-Mar, 16-Mar, 24-Mar, 30-Mar
Apr	4	9-Apr, 14-Apr, 20-Apr, 26-Apr
May	2	4-May, 19-May
Jun	3	1-Jun, 11-Jun, 30-Jun
Jul	3	7-Jul, 17-Jul, 24-Jul
Aug	3	2-Aug, 8-Aug, 24-Aug
Sep	2	14-Sep, 25-Sep
Oct	5	2-Oct, 9-Oct, 16-Oct, 23-Oct, 30-Oct
Nov	3	4-Nov, 16-Nov, 24-Nov
Dec	4	2-Dec, 9-Dec, 17-Dec, 23-Dec

1999		
Month	**# of Injuries**	**Date(s) of Injury**
Jan	3	9-Jan, 18-Jan, 31-Jan
Feb	1	18-Feb
Mar	3	8-Mar,19-Mar, 28-Mar
Apr	2	14-Apr, 27-Apr
May	1	28-May
Jun	1	9-Jun
Jul	2	2-Jul, 19-Jul
Aug	3	3-Aug, 18-Aug, 30-Aug
Sep	2	13-Sep, 28-Sep
Oct	1	11-Oct
Nov	2	8-Nov, 25-Nov
Dec	2	10-Dec, 31-Dec

C Chart for Injuries appears below. Note that there appears to be a reduction in injuries, but no assignable cause are noted according to our rules:



Individuals Chart for Time Between Injuries appears below. However, assignable causes (indicating an increased time between injuries – Good!) are noted on the chart:



Objective: To apply the ANOM method of detecting differences in experimental data.

Instructions:

1. Chart the following experimental data on an X-Bar, R Control Chart (see the COATING worksheet on your Excel file). Do any of the factor levels produce a "signal?"
2. Take the same data and perform an ANOM. What difference does this produce?

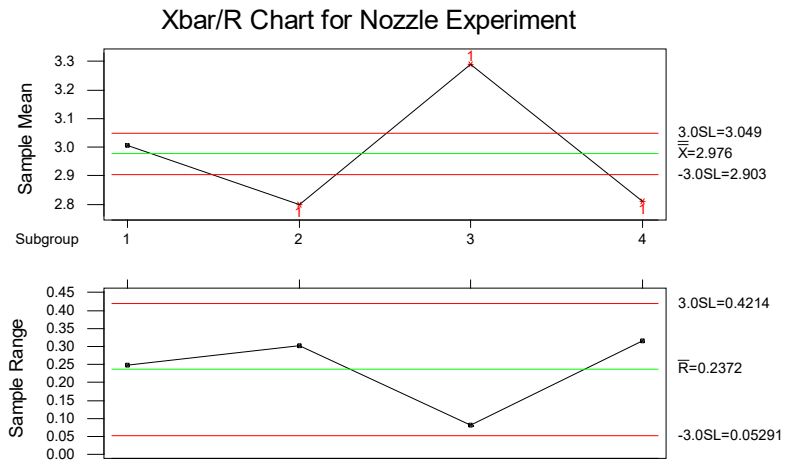
Time: 30 minutes

Coating Process Improvement

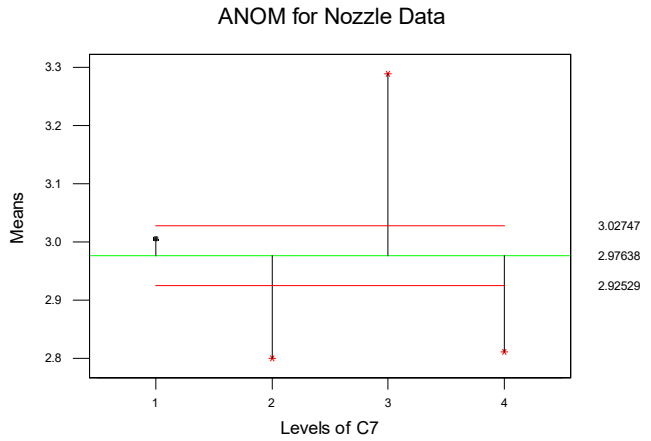
A Six Sigma team has run experiments to increase the coating thickness for air conditioner housings in an attempt to provide a more durable surface finish. The coatings are sprayed on the air conditioner housing and then the housings are baked. Four different spray nozzles (A – D) were used in the experiment with the following results:

Coating Thickness by Nozzle (mils)			
A	**B**	**C**	**D**
2.9364	2.9882	3.2488	2.8960
3.0135	2.6863	3.2824	2.7380
3.1551	2.7986	3.3301	2.8240
2.9543	2.8324	3.2620	2.7837
2.9839	2.7991	3.3198	2.8050
3.0006	2.8375	3.2669	2.6654
3.1108	2.7202	3.2788	2.9812
3.0059	2.7531	3.2703	2.8110
2.9054	2.8139	3.3224	2.8543
2.9897	2.7728	3.3029	2.7546

X-Bar, R Chart for Nozzle Data:



ANOM Procedure for Nozzle Data:



In this case, the X-Bar, R control chart and the ANOM procedure provide the same result – the B (2), C (3), and D (4) nozzles produce significantly different coating thicknesses.

Objective: To determine the relative advantage of the CUSUM chart over the X, mR control chart.

Instructions:

1. Chart the following data on an X, mR Control Chart (see the TOOL WEAR worksheet on your Excel file). Do any of the data produce a "signal?"
2. Take the same data and develop a CUSUM chart. What difference does this produce?

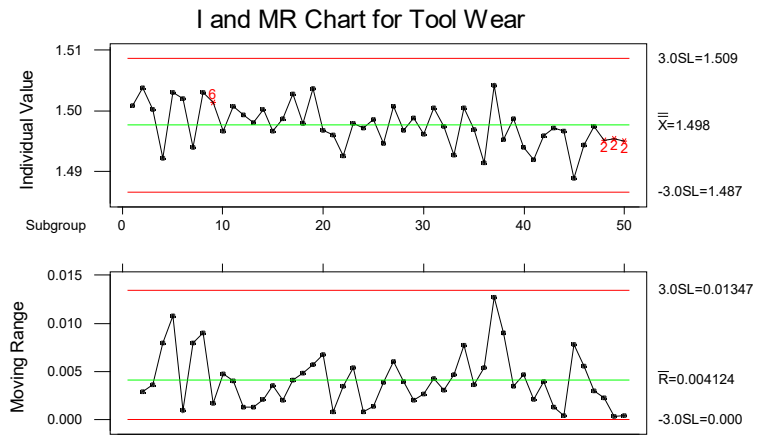
Time: 20 minutes

Suspected Tool Wear

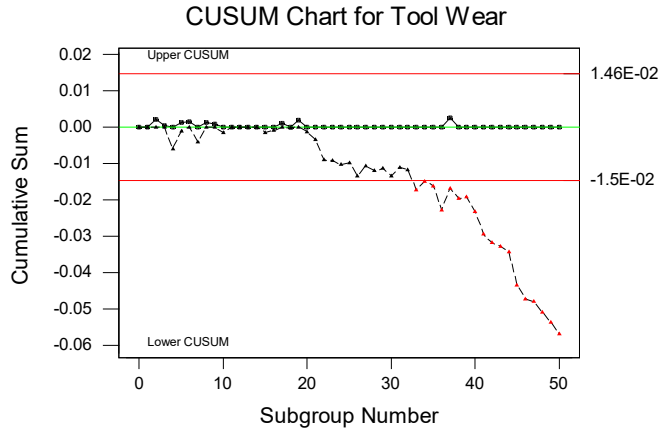
Manufacturing engineers are trying to determine if tool wear is affecting a particular cutting machine. They have collected the following data from the process (order of data proceeds down column 1 then to column 2 etc.):

Length	Length	Length	Length	Length
1.50095	1.50071	1.49602	1.50050	1.49195
1.50386	1.49940	1.49257	1.49744	1.49588
1.50019	1.49810	1.49799	1.49277	1.49715
1.49223	1.50023	1.49716	1.50050	1.49673
1.50306	1.49668	1.49855	1.49689	1.48889
1.50209	1.49867	1.49470	1.49146	1.49444
1.49407	1.50278	1.50078	1.50424	1.49743
1.50309	1.49790	1.49685	1.49524	1.49518
1.50139	1.50367	1.49884	1.49874	1.49548
1.49664	1.49687	1.49618	1.49406	1.49505

Individuals Chart for Tool Wear:



CUSUM Chart for Tool Wear:



Here, the Individuals chart takes a long time to detect the effects of tool wear; however, the CUSUM chart detects the degradation in the average very quickly.

Objective: To practice developing the Difference short run control chart.

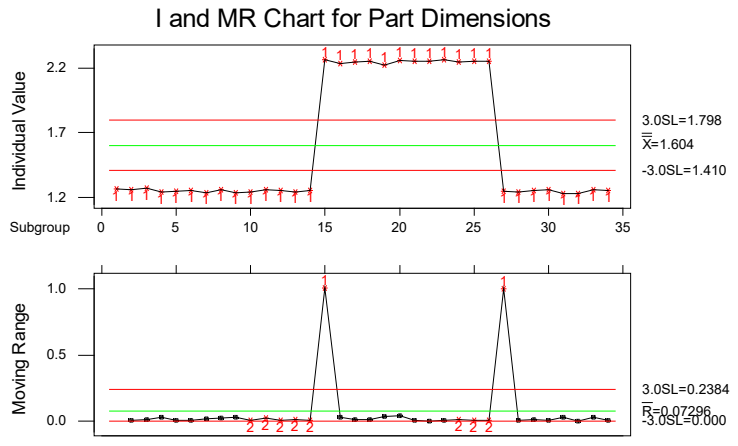
Instructions:
1. Chart the following data on an X, mR Control Chart (see the SHAFT worksheet on your Excel file).
2. Take the same data and develop a Difference control chart. How does this help you produce better information from the data?

Time: 20 minutes

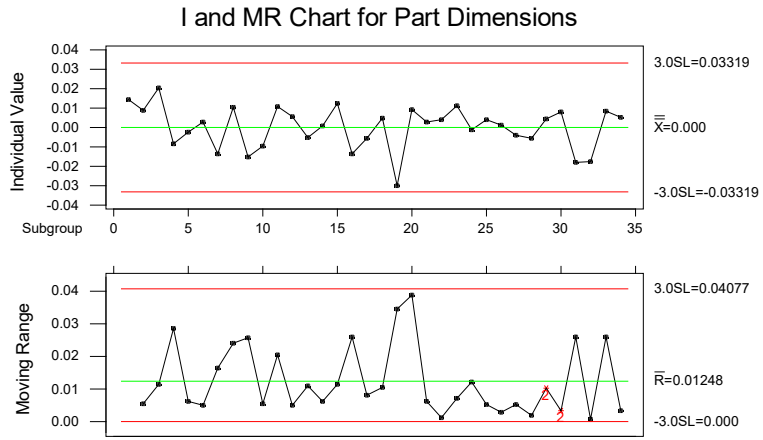
Short Run Motor Shafts - Compressor motor shafts are machined to support a Just-in-Time production operation. Manufacturing engineers believe that the machining process' variation doesn't change from shaft to shaft, however the shaft diameters differ (order of production proceeds down column 1 and then to column 2).

Part	Diameter	Part	Diameter
XB-4	1.2659	KJ-11	2.2544
XB-4	1.2604	KJ-11	2.2197
XB-4	1.2718	KJ-11	2.2586
XB-4	1.2431	KJ-11	2.2524
XB-4	1.2493	KJ-11	2.2536
XB-4	1.2543	KJ-11	2.2607
XB-4	1.2379	KJ-11	2.2485
XB-4	1.2621	KJ-11	2.2537
XB-4	1.2364	KJ-11	2.2508
XB-4	1.2418	XB-4	1.2477
XB-4	1.2622	XB-4	1.2458
XB-4	1.2573	XB-4	1.2561
XB-4	1.2464	XB-4	1.2595
XB-4	1.2525	XB-4	1.2334
KJ-11	2.2618	XB-4	1.2341
KJ-11	2.2359	XB-4	1.2600
KJ-11	2.2440	XB-4	1.2566

Individuals Chart for Shaft Machining Operation:



Difference Chart for Shaft Machining Operation:



The Individuals Chart for the Shafts simply shows the difference in average part size. Subtracting the mean from each of the part dimensions provides a better picture of the variability in the process.

Objective: To practice developing the ZED short run control chart.

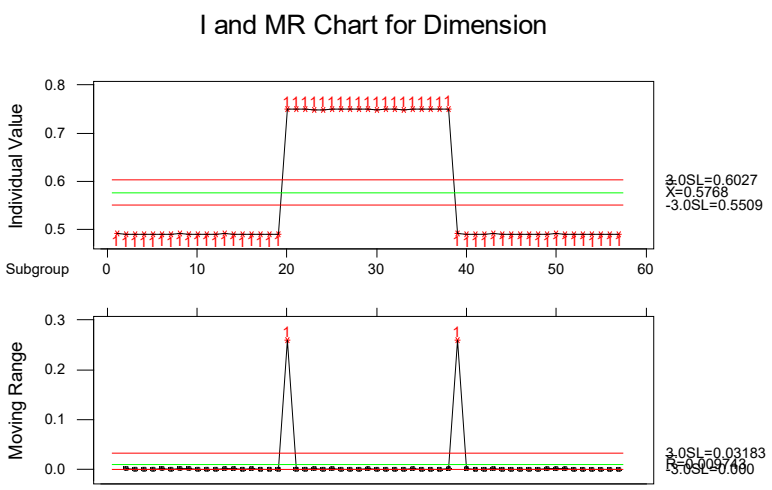
Instructions:
1. Chart the following data on an X, mR Control Chart (see the TUBE SHEET worksheet on your Excel file).
2. Take the same data and develop a ZED control chart. How does this help you produce better information from the data?

Time: 20 minutes

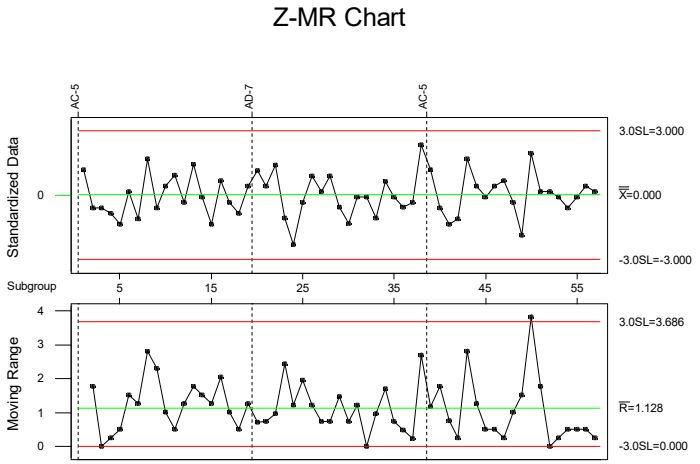
Tube Sheet Hole Drilling -The following data were collected from three commercial air conditioning tube sheets all drilled by the same machine. The tube diameters are different (Sheets 1 and 3 are drilled for 0.5 inch tubes, Sheet 2 is drilled for 0.75 inch tubes). Manufacturing engineering suspects that the variation in the tube diameters is different.

Tube Sheet 1	Diameter	Tube Sheet 2	Diameter	Tube Sheet 3	Diameter
AC-5	0.4905	AD-7	0.7507	AC-5	0.4905
AC-5	0.4898	AD-7	0.7504	AC-5	0.4898
AC-5	0.4898	AD-7	0.7508	AC-5	0.4895
AC-5	0.4897	AD-7	0.7498	AC-5	0.4896
AC-5	0.4895	AD-7	0.7493	AC-5	0.4907
AC-5	0.4901	AD-7	0.7501	AC-5	0.4902
AC-5	0.4896	AD-7	0.7506	AC-5	0.4900
AC-5	0.4907	AD-7	0.7503	AC-5	0.4902
AC-5	0.4898	AD-7	0.7506	AC-5	0.4903
AC-5	0.4902	AD-7	0.7500	AC-5	0.4899
AC-5	0.4904	AD-7	0.7497	AC-5	0.4893
AC-5	0.4899	AD-7	0.7502	AC-5	0.4908
AC-5	0.4906	AD-7	0.7502	AC-5	0.4901
AC-5	0.4900	AD-7	0.7498	AC-5	0.4901
AC-5	0.4895	AD-7	0.7505	AC-5	0.4900
AC-5	0.4903	AD-7	0.7502	AC-5	0.4898
AC-5	0.4899	AD-7	0.7500	AC-5	0.4900
AC-5	0.4897	AD-7	0.7501	AC-5	0.4902
AC-5	0.4902	AD-7	0.7512	AC-5	0.4901

Individuals Chart for Tube Diameter:



ZED Chart for Tube Diameters:



The Individuals Chart simply shows the differences in tube diameters. The ZED chart subtracts the differences due to the mean and the different process variation (Minitab was used to generate this chart (Stat>Control Chart>Z, mR). Two averages and standard deviations were calculated; one for each run of the AC-5, and the AC-7. The first two runs are similar in variation; an assignable cause is noted in the range of the second run of AC-5's.

Objective: To practice developing X-Bar, R control charts with varying subgroup sizes.

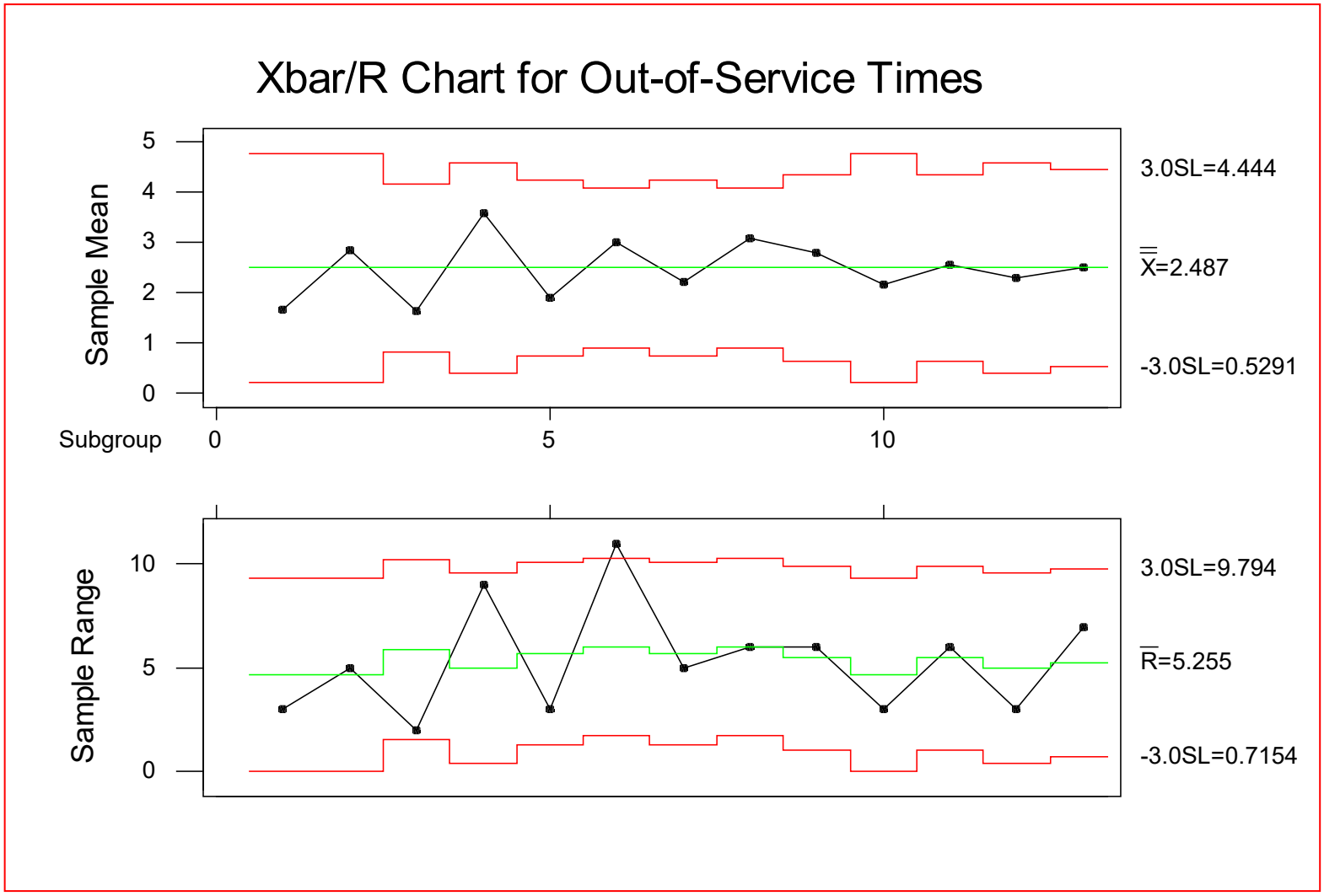
Instructions: 1. Chart the following data on an X-Bar, R Control Chart.

Time: 20 minutes

Lathe Out-of-Service Times - Take the following Lathe Out-of-Service times (hours) data and create an X-Bar, R Control Chart. To accommodate the varying subgroup sizes, you will have to calculate limits for each subgroup, using the "A" and "D" coefficients. Also, don't forget to use the raw data to calculate the grand average; you can't average the subgroup averages (without weighting them!):

JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	JAN
1	3	2	1	1	1	1	2	4	3	6	4	1
1	1	1	4	1	3	1	3	1	1	1	3	8
2	1	3	4	1	1	1	2	1	4	2	1	2
4	4	2	10	1	2	6	2	4	3	7	1	3
1	2	1	3	2	12	2	7	4	1	2	2	2
1	6	3	1	4	1	5	3	1	1	1	3	1
		1	2	2	4	2	3	7		1	2	2
		2		4	1	2	1	1		2		1
		1		2	1	1	1	2		1		
		1		1	1	1	4					
		1			5		3					
					4		6					

Note the variable control limits due to the varying subgroup sizes each month:



Section 7 – Stratification and Prioritization

Controller Failures

In the last six months, HVAC controllers from three manufacturers have failed while in service. As part of their improvement effort, a team identified how many controllers were installed (by manufacturer). They also counted the number of failures experienced:

Controller Failures

Manufacturer	# Installed	# Failed
Jonson	24	7
Airaid	32	9
BlowPulse	9	2

How would you display this data graphically? Do so. Do you think there is a difference in reliability by manufacturer? Perform a contingency table analysis of the data. What does this test tell you?

A stacked bar graph would be a good display. The contingency table analysis appears below (Minitab: Stat>Tables>Chi-Square Test):

Chi-Square Test

```
Expected counts are printed below observed counts

         Jonson   Airaid BlowPuls    Total
    1         7        9        2       18
           6.65     8.86     2.49

    2        17       23        7       47
          17.35    23.14     6.51

Total        24       32        9       65

Chi-Sq =  0.019 +  0.002 +  0.097 +
          0.007 +  0.001 +  0.037 = 0.164
DF = 2, P-Value = 0.921
1 cells with expected counts less than 5.0
```

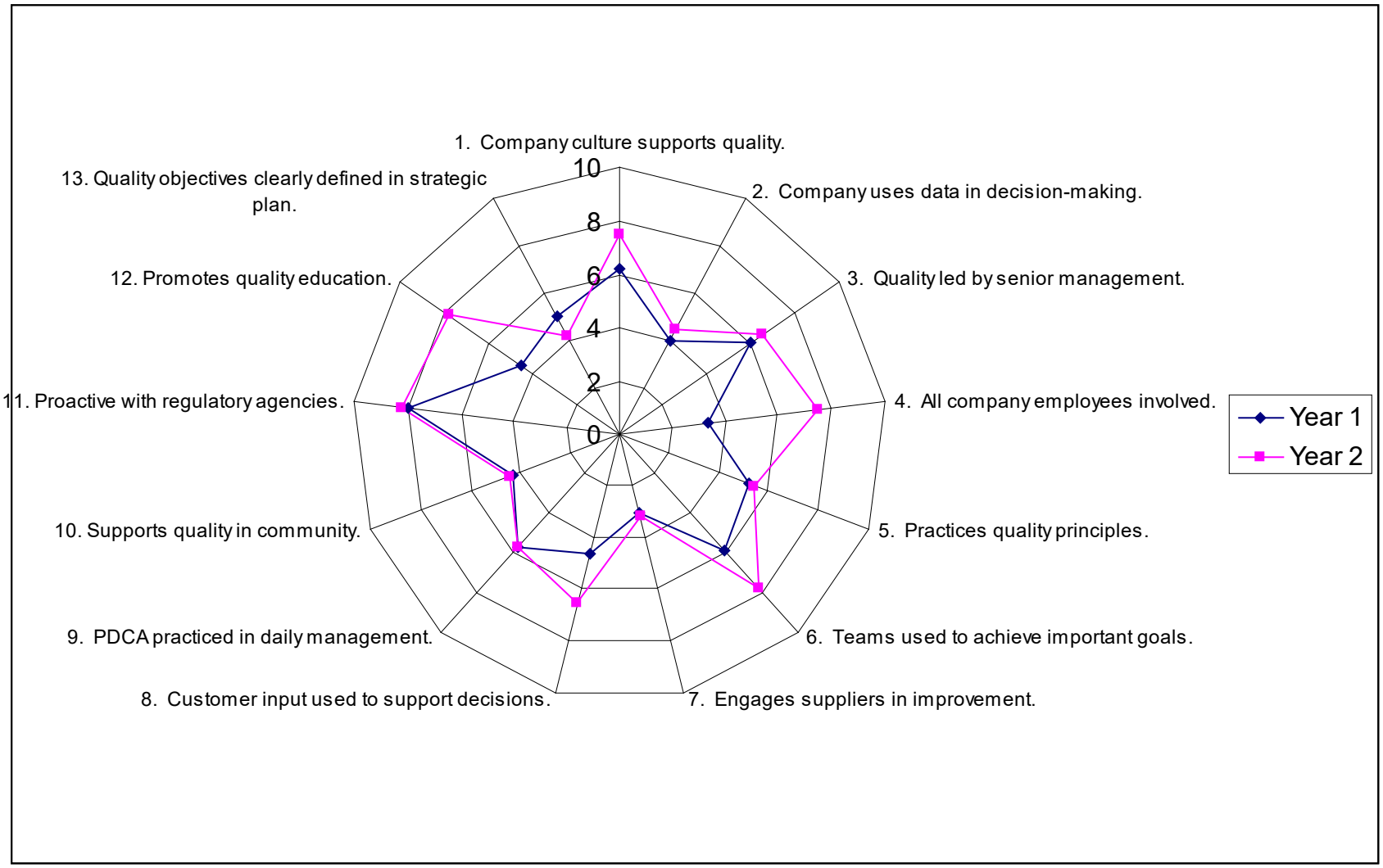
Conclusion: With a p value of 0.921, there is no evidence that there is a relationship between manufacturer and failure rate of controllers.

Employee Perception

A large engineering firm conducted a survey of employees one year and two years after introduction of their Total Quality Management system. The questions were designed to determine how employees perceived progress made by management in transforming their style and practice of management. Prepare a radar chart and plot both of these survey results on the same chart. What changed from year one to year two? Where is management strongest, weakest in TQM?

	Average Score	
Survey Question	**Year 1**	**Year 2**
1. Company culture supports quality.	6.2	7.5
2. Company uses data in decision-making.	4.0	4.5
3. Quality led by senior management.	6.0	6.5
4. All company employees involved.	3.3	7.5
5. Practices quality principles.	5.2	5.4
6. Teams used to achieve important goals.	5.8	7.8
7. Engages suppliers in improvement.	3.0	3.2
8. Customer input used to support decisions.	4.6	6.5
9. PDCA practiced in daily management.	5.7	5.7
10. Supports quality in community.	4.3	4.4
11. Proactive with regulatory agencies.	8.0	8.2
12. Promotes quality education.	4.5	7.8
13. Quality objectives clearly defined in strategic plan.	5.0	4.2

The Radar Chart appears below:



Treatment Costs

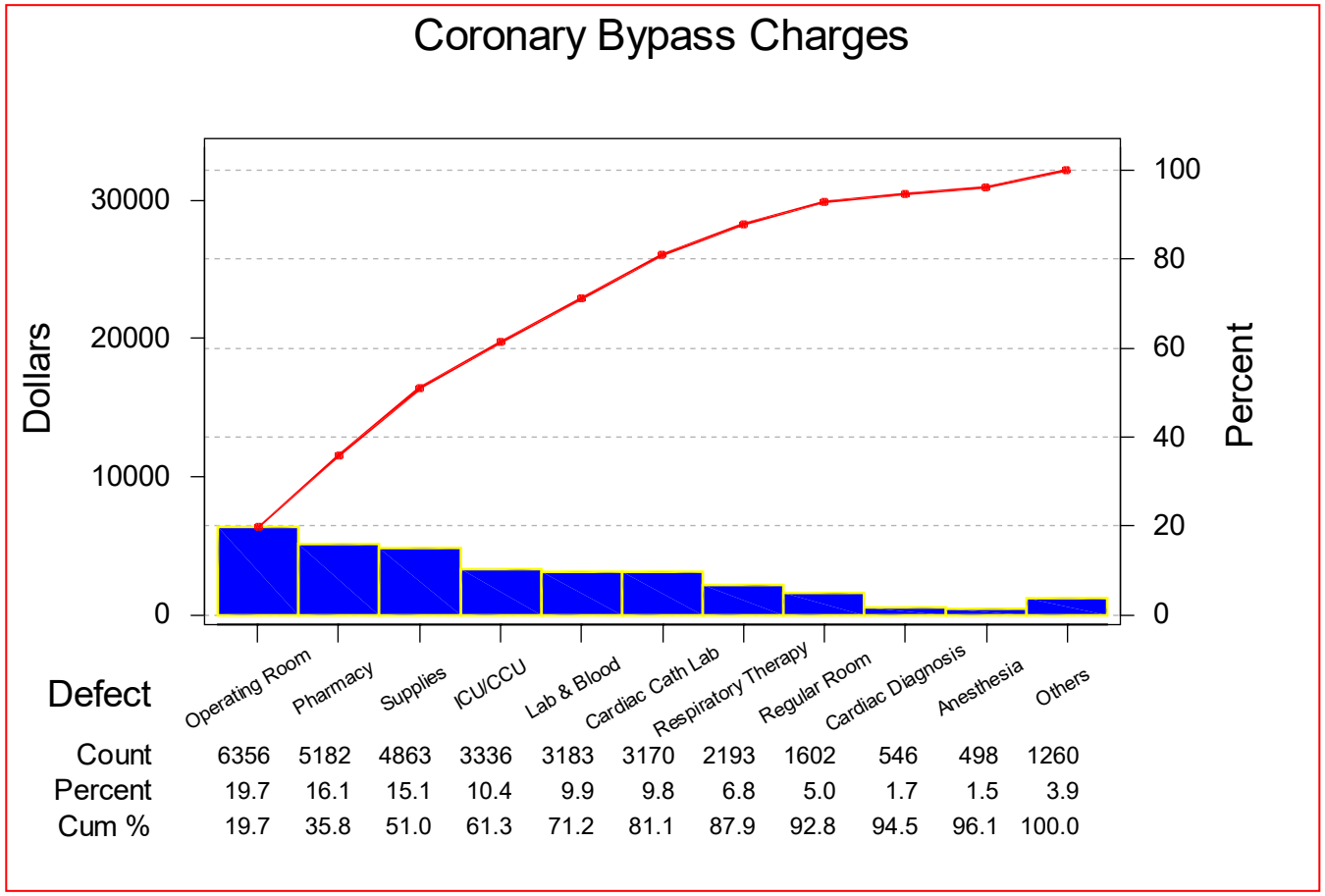
A hospital that tracked the Length of Stay data for the diagnosis, Coronary Bypass with Cardiac Catheterization, began an improvement effort to reduce the *unnecessary costs* of this diagnosis. They collected data on the charges associated with 13 patients who fell into this diagnosis. Prepare a Pareto Chart of this data.

Coronary Bypass with Cardiac Catheterization
DRG-106 (13 Patients)

Category	Charges ($)
Anesthesia	498
Cardiac Cath Lab	3170
Cardiac Diagnosis	546
ICU/CCU	3336
Lab & Blood	3183
Operating Room	6356
Other	347
Pharmacy	5182
Radiology	475
Regular Room	1602
Respiratory Therapy	2193
Step Down	438
Supplies	4863

Does this Pareto provide you with clues as to where to begin to reduce *unnecessary costs?* What's the problem with this Pareto? (Hint: what kind of "problem" is this?).

Here's the Pareto. The problem with this chart is that the "zero-problem" thinking associated with Pareto Analysis has not been applied. All we know are where the highest charges are incurred. A value-added/non-value added analysis would be an alternate way to attack this problem.



Labor and Delivery

A Labor & Delivery team is investigating the relationship between the mother's dilation when an epidural is administered and the C-Section rate. Four dilation ranges were identified and C-Section rates measured for two months. Perform a Contingency Table analysis of the data. Use α = 0.05. Is there a difference?

	Dilation (cm)			
Delivery	**0 - 2.5 cm**	**2.5 - 5.0**	**5.0 - 7.5**	**7.5 to 10**
C-Section	48	51	28	12
Vaginal	142	219	272	228
Total	190	270	300	240

Minitab was used to perform the analysis (Stat>Tables>Chi-Square Test):

Chi-Square Test

```
Expected counts are printed below observed counts

       0 - 2.5 2.5 - 5. 5.0 - 7. 7.5 to 1    Total
    1       48       51       28       12      139
         26.41    37.53    41.70    33.36

    2      142      219      272      228      861
        163.59   232.47   258.30   206.64

Total      190      270      300      240     1000

Chi-Sq = 17.650 +  4.835 +  4.501 + 13.677 +
          2.849 +  0.780 +  0.727 +  2.208 = 47.226
DF = 3, P-Value = 0.000
```

The p-value shows that there is a significant relationship between C-section/vaginal delivery and mother's dilation.

Section 9 – Detecting Differences

Probability and Statistics Foundations

Exercise - Bill of Materials Errors

A sample of 100 Bills of Materials is inspected for errors, with 12 errors being detected. What is your best estimate of the error probability?

Our best estimate here uses the frequentist definition of probability: 12/100 = 0.12

Exercise - Heat Pump Warranties

For home heat pumps manufactured by one plant, the probability of a heat pump being produced by the first shift is 0.74. For these heat pumps, the probability of warranty failure is 0.13. For the heat pumps produced by the second shift, the probability of warranty failure is 0.08. What is the probability that a heat pump produced by the plant will have a warranty failure?

The probability of warranty failure is calculated below:

$$P(Warranty\ Failure) = P(Failure \mid First\ Shift) \times P(First\ Shift) + P(Failure \mid Second\ Shift) \times P(Second\ Shift)$$
$$= 0.13 \times 0.74 + 0.08 \times (1 - 0.74) = 0.12$$

Exercise - Mutually Exclusive/Independent Events

Which of the following events are mutually exclusive? Independent?

Neither	A reciprocating compressor *and* a reciprocating compressor over 10 years old.
Mutually Exclusive, Independent	A lab report completed within 20 minutes *and* a lab report completed in over 40 minutes.
Independent	Two brazing operations that take more than 25 minutes.
Independent	Two shaft leaks in screw compressors.
Independent	Two evaporators with leak rates greater than 100 cc's/hr.
Mutually Exclusive, Independent	Sales in Europe, Asia, South America and North America

Exercise - Late Meetings

During reviews of annual plans, the probability of the first meeting of the day starting on time is 0.30. If the first meeting is late, the second meeting has a 70% chance of starting late. If the first meeting is on time, the second meeting has a 90% chance of starting on time. What is the probability that the second meeting will start *on time?*

This is an application of the law of total probability: $P(A) = P(A|B_1)\,P(B_1) + P(A|B_2)\,P(B_2) + \ldots + P(A|B_n)P(B_n)$

$$\begin{aligned} P(\textit{Second Meeting on time}) &= P(\textit{Second Meeting on time} \mid \textit{First Starts on time}) \times P(\textit{First Starts on time}) + \\ &\quad P(\textit{Second Meeting on time} \mid \textit{First Starts late}) \times P(\textit{First Starts late}) \\ &= 0.9 \times 0.3 + (1 - 0.7) \times (1 - 0.7) = 0.36 \end{aligned}$$

Exercise - Populations and Samples

Which of the following describe populations? Samples? Why?

The 30 Bills of Material are a sample from an ongoing process.	30 Bills of Materials generated each week.
Could be treated as a population, or as a sample of failed fans from a larger population.	The 50 fan motors which failed last year.
Could be treated as a population, or as a sample of failed fans from a larger population.	The tube sheets brazed by a worker last month.
Could be treated as a population, or as a sample of failed fans from a larger population.	The staff trained in quality improvement methods in the last six months.
Most likely a sample from all of last month's customers.	The customers called for input to last month's satisfaction survey.
Could be treated as a population, or as a sample of all Black Belts.	The Black Belts in General Electric.

Exercise - Successful Brazing

Of a sample of 100 impellers brazed with a new process, 35 passed inspection. What is the expected value of the proportion of good impellers? What is the expected value of the standard deviation?

This is a Binomial process, with proportion mean of *np/n* = *p* and the variance is *np(1 - p)/n* or *pq*:

$$p = 35/100 = 0.35$$

$$pq = 0.35(1 - .35) = 0.2275$$

The Standard Deviation is the square root of the variance:

$$s = \sqrt{0.2275} = 0.477$$

Exercise - Set-up Time

The time to set-up a particular job can be modeled using a lognormal distribution, with mean 2.3 and standard deviation 1.0. What proportion of jobs can be expected to be set-up within 15 minutes?

The cumulative lognormal distribution is given by $F(X:\mu,\sigma)=\frac{1}{\sqrt{2\pi}\sigma}\int_0^X e^{-\frac{1}{2}(\frac{\ln x-\mu}{\sigma})^2}\frac{dx}{x}$

The probability evaluation is then found by transforming the variable to a normal distribution:

$$F(X<15\,\text{min}.:2.3,1.0)=F((\ln 15-2.3)/1.0:0,1)=F(0.41:0,1)=1-0.3409=0.6591$$

The value of F is found in Appendix A.

Exercise - Chiller Repair Time

How many standard deviations away from the mean is a repair time of 10 days, given the Average repair time is 5 days and the variance is 9 days^2? What is the probability of a chiller's repair time being longer than 10 days?

For a variance of 9 days^2, the standard deviation is 3 days. The 10 days is then (10 – 5)/3 or 1.66 standard deviations from the mean.

The probability of a repair time greater than 10 days is:

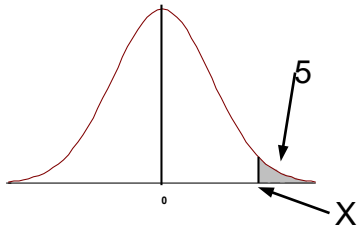
$$F(X > 10 : 5,3) = F((10 - 5)/3 : 0,1) = F(1.66 : 0,1) = 0.0485$$

The value of F is found from the Standard Normal Table in Appendix A.

Exercise - Welding Procedures

A certain welding procedure requires an average of 45 minutes to perform, with a variance of 8 minutes (normally distributed). What's the shortest time you'd expect the procedure to be performed? The longest? Answer the question for risks of 5% and 1%.

Here, we are looking for the "X" which corresponds to a probability of 5% for a normal distribution with mean 45 and standard deviation 2.83 (square root of the variance):



The probability expression is written:

$$\begin{aligned}
&Longest\ Time:\\
&P(X:45,2.83) = P(Z = (X-45)/2.83:0,1) = 0.95\\
&\therefore Z = 1.645\ and\ X = 45 + 1.645 \times 2.83 = 49.66\,\text{min}.\\
&Shortest\ Time:\\
&P(X:45,2.83) = 0.05\\
&\therefore Z = -1.645\ and\ X = 45 + (-1.645 \times 2.83) = 40.34\,\text{min}.
\end{aligned}$$

Exercise - Sales Order Processing

The time to complete a sales order is normally distributed with a mean of 40 minutes and a standard deviation of 8 minutes.

- What is the probability that any given sales order will require more than 50 minutes to process?

  Answer: $P(X > 50 : 40,8) = P((50 - 40)/8 : 0,1) = P(1.25 : 0,1) = 0.1056$

- What is the probability that it will require less than 20 minutes?

  Answer: $P(X < 20 : 40,8) = P((20 - 40)/8 : 0,1) = P(-2.5 : 0,1) = 0.0062$

- What is the probability that it will take no more than 48 minutes?

  Answer: $P(X < 48 : 40,8) = P((48 - 40)/8 : 0,1) = P(1.0 : 0,1) = 1 - 0.1587 = 0.8413$

- What is the probability that it will take between 20 and 50 minutes?

  Answer: You should be able to figure this out by now!

Note that you have to be careful how to obtain the value from Appendix A. We find that it is helpful to draw the picture of the normal distribution and shade in the probability we are attempting to find.

Exercise - Chiller "DOAs"

The probability of a chiller being "Dead on Arrival" is 20%. In a one-month time period, 50 chillers were installed. What is the probability that exactly 10 "DOA's" occurred? What is the probability that less than 3 "DOA's" occurred?

This is a Binomial Process:

For the first question, we need the probability mass function:

$$f(x:n,p)=\binom{n}{x}p^{x}q^{n-x} \Rightarrow f(10:50,0.20)=\binom{50}{10}0.2^{10}0.8^{50-10}=?$$

Now we could do this calculation, or let Minitab handle the details (Calc>Probability Distributions>Binomial):

Probability Density Function

```
Binomial with n = 50 and p = 0.200000

          x      P( X = x)
      10.00        0.1398
```

For the second question, we need to evaluate the cumulative probability function (sum of the probability of 0, 1 or 2 DOA's):

Cumulative Distribution Function

```
Binomial with n = 50 and p = 0.200000

          x     P( X <= x)
       2.00        0.0013
```

Note that Minitab gives you the probability of X being less than or equal to the value. That's why we input 2 instead of 3.

Exercise - Employee Injuries

Employee injuries occur at the rate of 2/month at a certain facility. What is the probability that less than 1 employee will be injured in any given month? If 6 or more injuries occurred in one month, would this be considered "unusual?"

This is a Poisson process. We'll use Minitab to calculate the probability of having zero injuries in one month:

Calc>Probability Distributions>Poisson

Probability Density Function

```
Poisson with mu = 2.00000

          x       P( X = x)
       0.00          0.1353
```

Note that we used the probability density function; here the cumulative and density functions provide the same answer.

For the second part of this question:

Cumulative Distribution Function

```
Poisson with mu = 2.00000

          x      P( X <= x)
       5.00          0.9834
```

Therefore, the probability of six or more injuries is 1 – 0.9834 = 0.0166. Most would judge that this a low enough probability to consider the number of injuries unusual.

Exercise - Refrigerant Weight

In a manufacturing process, a fixed amount of liquid refrigerant is fed automatically into a metal container. The average weight of the bottles is 20.0 kilograms with a standard deviation of 0.05 kg. The average weight of the filled bottles is 105.0 kg, with a standard deviation of 0.5 kg. What are the mean and standard deviation of the refrigerant?

The mean of the weights is simply 105.0 – 20.0 = 85.0 kilograms

To find the standard deviation, we employ additivity of variances:

$$s^2_{Filled} = 0.5^2 = s^2_{\text{Re}\,frigerant} + 0.05^2$$

$$\therefore s_{\text{Re}\,frigerant} = \sqrt{0.5^2 - 0.05^2} = 0.497kg$$

Exercise - Binomial and Poisson

Place an “X” next to the *binomial* random variables and an “O” next to the *Poisson* random variables on the list below:

X	Phone calls to the service center not answered in 45 seconds.
Neither	Time to complete the annual planning process.
O	Number of scrap items each month.
O	Errors on an expense account.
X	Engineering reports not completed in 2 days.
O	Eye injuries in a plant each quarter.

Exercise - Change Requests

Ten percent of change requests are not forwarded to Engineering within 24 hours of receipt. In a sample of 10 change requests, what is the probability that 3 or more are not forwarded within 24 hours?

This is a Binomial process. Minitab will be used to calculate the probability:

Cumulative Distribution Function

```
Binomial with n = 10 and p = 0.100000

         x      P( X <= x)
      2.00          0.9298
```

Therefore, 1 – 0.9298 = **0.0702** is the probability of three or more not being forwarded in 24 hours.

Exercise - Service Center Calls

A service center receives calls at the rate of 10 calls per hour. What is the probability that 2 or fewer calls will be received in an hour?

This is a Poisson process:

Cumulative Distribution Function

```
Poisson with mu = 10.0000

         x      P( X <= x)
      2.00          0.0028
```

Exercise - Oil Changes

A quality improvement team is investigating whether they can extend the interval between oil changes for a fleet of service trucks. They track oil condition in 17 trucks and find that, on average, the oil fouls after 6600 miles, with a standard deviation of 200 miles. If they wanted to set a "fixed" time to change the oil, what would be the maximum number of miles you'd recommend?

Here, the team might set an acceptable probability of having fouled oil. For example, if they wanted to keep this probability at 1% or less, they would set the following oil change interval:

$$P(Z:0,1)=0.01 \Rightarrow Z=-2.32 \Rightarrow (X-6600)/200=-2.32 \Rightarrow X=6600-2.32\times 200=6136 miles$$

Exercise - Pump Failure Rate

The failure rate of water pumps can be modeled with an exponential distribution, with a λ = 0.001/hr. What is the probability that a new pump will operate for 1500 hours without failure?

Since this is an exponential process, the survival probability is:

$$R = \exp(-0.001 \times 1500) = 0.22313$$

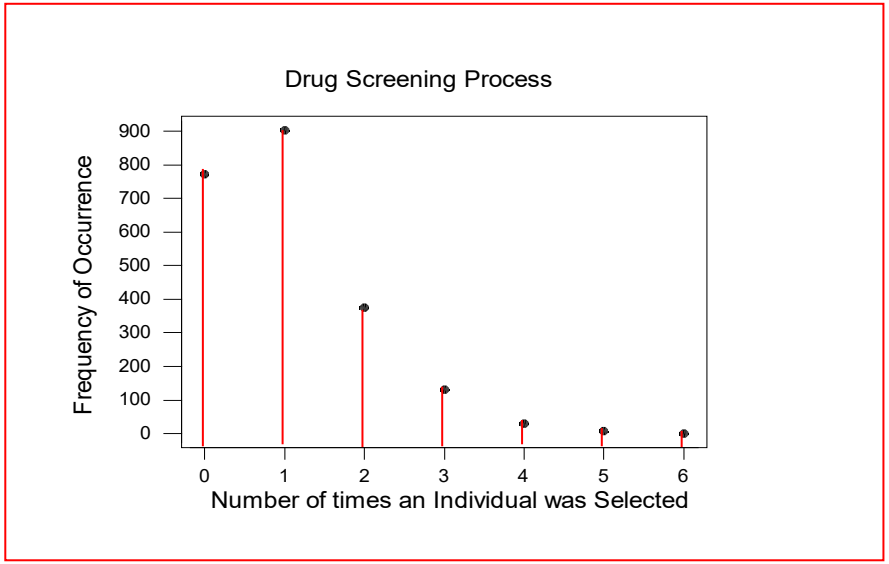
Exercise - Drug Testing

A company involved in sensitive military work employs a random drug-testing program. Their program is set up at a 100% sampling rate/year. The following data indicates the number of times employees were selected for drug testing in a one-year period:

Number of times an Individual was selected	Frequency of Occurrence
0	771
1	902
2	375
3	131
4	31
5	9
6	1
Total	2220

Develop a frequency chart for this data. What is the average number of times an individual was selected? Why are so many employees not selected at all if the sampling rate is 100%/year? Is there something wrong with the random selection process?

The average number of times an individual was sampled is 2200/2200 = 1.0/year or a 100% sampling rate. This is a Binomial process; each worker has a chance of being selected each time a sample is pulled. Some of the employees will be selected multiple times, some not at all.



Detecting Changes and Differences

Exercise - Failed Lights

In a sample of 100 fluorescent lights in a factory, seven were found to be not working. Construct a 95% confidence interval for the percentage of all burned out fluorescent lights in the plant.

The confidence interval for the percentage of burned out lights is:

$$p - K_{\alpha/2}\sqrt{p(1-p)/n} \le \mathbf{P} \le p + K_{\alpha/2}\sqrt{p(1-p)/n}$$

$$0.07 - 1.96\sqrt{0.07(1-0.07)/100} \le \mathbf{P} \le 0.07 + 1.96\sqrt{0.07(1-0.07)/100}$$

$$0.02 \le \mathbf{P} \le 0.12$$

Exercise - Dimensions Measurement

A new, less expensive method for QC to check dimensions has been developed. The QC manager is willing to recommend the new method if it can be shown that the new method is as accurate as the more expensive one. Fifteen components are picked to compare the methods, with the data shown below. At a 2% level of significance, what recommendation should the case manager make?

Component	Dimension (mm)	Dimension (mm)
1	110	113
2	105	102
3	120	123
4	114	114
5	118	121
6	104	106
7	95	93
8	98	95
9	102	105
10	96	101
11	114	117
12	112	114
13	99	96
14	106	109
15	116	121

This situation calls for a paired t-test (Minitab: Stat>Basic Statistics>Paired-t). The results of the test are:

Paired T-Test and Confidence Interval

```
Paired T for Old Method - New Method

                 N      Mean     StDev   SE Mean
Old Method      15    107.27      8.24      2.13
New Method      15    108.67      9.90      2.56
Difference      15    -1.400     2.849     0.735

95% CI for mean difference: (-2.977, 0.177)
T-Test of mean difference = 0 (vs not = 0): T-Value = -1.90  P-Value = 0.078
```

Conclusion: Since the p-value is greater than 0.02, there is not enough evidence to reject the null hypothesis that the methods are the same.

Exercise - Fire Department Accidents

A fire department has kept records of the number of accidents their fire trucks are involved in while responding to a fire alarm. One of the firemen thinks that the color of the fire truck makes a difference in the accident rate. Test the following data at the 5% level of significance:

	Red Fire Trucks	Yellow Fire Trucks
Number of Accidents	20	4
Number of Trips	153348	135035

Here, we'll use the book procedure to perform the hypothesis test:

$a)\ \mathrm{H_o}$: Population Proportions are Equal $(\mathrm{P_R} = \mathrm{P_Y})$

H_a : The Population Proportions are Not Equal $(\mathrm{P_R} \neq P_Y)$

$b)\ \alpha = 0.05$

$c)\ \text{Test Statistic}: Z = \dfrac{P_R - P_Y}{\sqrt{\mathrm{p(1-p)}\left(1/\mathrm{n_R} + 1/n_Y\right)}}$ Rejection Region : $\mathrm{K}_{0.05} = \pm 1.96$ (Normal)

d) Calculations : $p_R = 20/153{,}348 = 1.3E-4,\ p_Y = 4/135{,}035 = 2.96E-5,$

$$p = \frac{20+4}{153{,}348+135{,}035} = 8.32E-5$$

$$Z = \frac{(1.3E-4)-(2.96E-5)}{\sqrt{8.32E-5(1-08.32E-5)(1/153{,}348+1/135{,}035)}} = 2.96$$

$e)$ Conclusion : 2.96 is outside $\pm 1.645, \therefore$ Reject $\mathrm{H_0}$ in favor of $\mathrm{H_a}$ at the 5% significance level.

Exercise - Difference of Proportions

Changes were made to the process of expanding tubes in tube sheets in hopes of reducing leaks. The previous leak rate was 3%. Over the next 100 tubes, only one leaked. At the 5% level of significance, has there been a change in the leak rate?

Hypothesis Test:

a) H_o: *Population Proportion is 0.03* ($P_o = 0.03$.)

H_a: *Population Proportion is Less Than 0.03* ($P_o < 0.03$.)

b) $\alpha = 0.05$

c) Test Statistic: $Z = \dfrac{p - P_o}{\sqrt{P_o(1 - P_o)/n}}$ *Rejection Region:* $K_{0.05} = -1.645$ *(Normal)*

d) Calculations: $p = 1/100 = 0.01$

$$Z = \frac{0.01 - 0.03}{\sqrt{0.03(1 - 0.03)/100}} = -1.17$$

e) Conclusion: $-1.17 > -1.645$, $\therefore$ *Do not Reject* H_0 *in favor of* H_a *at the 5% significance level.*

f) Parameter Estimation: *Our best estimate of* P_o *is still 0.03.*

Note: Since *np* is less than 5, the assumption of normality may be challenged. An alternative procedure is to calculate the probability of having 1 or fewer leaks in a sample of 100, using the binomial distribution. This calculation yields a probability of 0.195, or about 20%. This agrees with our conclusion to *not reject* the null hypothesis.

Exercise - Difference of Standard Deviations

A manufacturer of compressor shaft bearings claims that the standard deviation in the expected life is 0.5 years and thus it is easy to plan for maintenance and replacement. A random sample of 20 bearings was tested and the standard deviation of the sample is 0.65 years. At a 5% level of significance, test whether the standard deviation of this product is greater than 0.5 years.

Hypothesis Test:

$a)\ \mathrm{H_o}$: Population Variance is 0.5^2 years^2 $(\sigma_o^2 = 0.025)$

H_a : Population Variance is Greater Than 0.5^2 years^2 $(\sigma_o^2 < 0.025)$

$b)\ \alpha = 0.05$

$c)$ Test Statistic : $\chi^2 = \dfrac{(n-1)s^2}{\sigma_o^2}$ Rejection Region : $\mathrm{K}_{0.05} = 30.2\ (\chi^2$ dist, f = 20 - 1 = 19)

d) Calculations : $\chi^2 = \dfrac{(20-1)0.65^2}{0.5^2} = 32.11$

$e)$ Conclusion : $32.11 > 30.2, \therefore$ Reject $\mathrm{H_0}$ in favor of $\mathrm{H_a}$ at the 5% significance level.

f) Parameter Estimation : $s = 0.65, \mathrm{s_{Upper}} = \sqrt{\dfrac{0.65^2(20-1)}{8.91}} = 0.95,$

$\mathrm{s_{Lower}} = \sqrt{\dfrac{0.65^2(20-1)}{32.9}} = 0.49$

Exercise - Difference of Proportions

Accounts Payable tracks on-time and late invoices. The quality department was informed that the majority of their bills (150 of 200) were paid late last month. Quality investigated the payment process and made several process changes. This month's results were 26 late of 130 invoices. Has the proportion of late invoices decreased?

Hypothesis Test:

a) H_o: *Population Proportions are Equal* ($P_B = P_A$.)

H_a: *The "After" Population Proportion is Less Than the "Before"* ($P_A < P_B$)

b) $\alpha = 0.05$

c) Test Statistic: $Z = \dfrac{p_A - p_B}{\sqrt{p(1-p)\left(1/n_A + 1/n_B\right)}}$ *Rejection Region:* $K_{0.05} = -1.645$ *(Normal)*

d) Calculations: $p_B = 150/200 = 0.75$, $p_A = 26/130 = 0.2$, $p = \dfrac{150+26}{200+130} = 0.53$

$$Z = \frac{0.2 - 0.75}{\sqrt{0.53(1-0.53)(1/200 + 1/130)}} = -9.78$$

e) Conclusion: $-9.78 < -1.645$, $\therefore$ *Reject* H_0 *in favor of* H_a *at the 5% significance level.*

f) Parameter Estimation: $p \pm 1.96\sqrt{p(1-p)/n} = 0.2 \pm 1.96\sqrt{0.2(0.8)/130} = 0.2 \pm 0.07$

Exercise - Mean Confidence Interval

The average manufacturing cost of a compressor built last year was $8,200. To estimate the average cost this year, an analyst randomly selects 100 jobs and calculates an average of $9,000, with a standard deviation of $200. Calculate a 90% confidence interval for the average cost of the procedure.

Confidence Interval:

$$\overline{X} \pm K_{\alpha/2} \times s/\sqrt{n}$$

$$\text{for } a/2 = 0.05, \mathrm{K}_{\alpha/2} = 1.66 \,(\text{t - dist, f} = 100 - 1 = 99)$$

$$CI: \ \$9{,}000 \pm 1.66 \times \$200/\sqrt{100} \Rightarrow \$9{,}000 \pm \$33.20$$

Exercise - Difference of Means, Standard Deviation Unknown

An Education Director is interested in whether a computerized training program will improve the test scores for Black Belts (BB's). 25 BBs receive the traditional classroom training and 25 are trained through the computer approach. Given the following test results, is there a difference in the two approaches?

	Classroom Training	Computer Training
Average -	78	85
Standard Deviation -	6	5
Number of NAs -	25	25

Hypothesis Test:

a) H_o: *Population Means are Equal* ($\mu_A = \mu_B$)

H_a: *Computer (" A") Population Mean is Greater Than Classroom (" B")* ($\mu_A > \mu_B$)

b) $\alpha = 0.05$

c) Test Statistic: $t = \dfrac{\overline{X}_A - \overline{X}_B}{\sqrt{(s_A^2 + s_B^2)/n}}$

Rejection Region: $K_{0.05} = 1.678$ *(t - dist.,* f *= 2(25 - 1) = 48)*

d) Calculations: $t = \dfrac{85 - 78}{\sqrt{(5^2 + 6^2)/25}} = 4.48$ *(Assume Equal Variances here)*

e) Conclusion: 4.48 > 1.678, $\therefore$ *Reject* H_0 *in favor of* H_a *at the 5% significance level.*

f) Parameter Estimation: $\overline{X} \pm K_{\alpha/2} s/\sqrt{2} = 85 \pm 2.011 \times 5/\sqrt{25} = 85 \pm 2.0$

Sampling Methods

Exercise - Black Belt Survey

Instructors of the Black Belt Tools course are designing a survey to be completed by class participants to evaluate the course content and instruction. List some possible subgroups that may be of interest when the survey results are analyzed.

Possible Subgroups:

- Age
- Gender
- Education
- Job Position
- Previous Quality Experience

Exercise - Budget Prep Time

The Strategic Planning VP is planning a survey to estimate the time spent to prepare the business units' annual budget. From the BU's phone book, the VP randomly selects 20 departments and sends a survey to the manager of each department. What are the element, sampling unit, universe and frame for this survey?

Element: The unit about which information is to be collected – the **business unit** is the subject of the data collection here.

Sampling Unit: In single stage sampling, a sampling unit is the same as an element – i.e. the **business unit**.

Universe: The entire population – **all business units in the company**.

Frame: A list of *every* element of the population – here, the **BU phone book** is the frame.

Exercise - Survey Comparison

Which is the better of the following two surveys? Why?

	Survey # 1	**Survey # 2**
Total # Customers -	300,000	300,000
# Surveys Sent Out -	1,000	100
# Surveys Returned -	100	80

Survey # 1 is superior in design – the larger sample size will enable the data collectors to measure customer attributes at a higher *precision* level. Survey # 2 is superior in execution – given that both surveys employed a random sampling technique, the higher response rate for Survey # 2 leads to a lower non-response error.

Exercise - Customer Sample

You send the following survey to a sample of 100 customers:

Will you order air conditioning equipment next year? _____ *Yes* _____ *No*

Everyone responds to your survey. Forty customers respond "yes," sixty respond "no." Estimate the % of all customers who will order new equipment next year. Calculate the sampling error at 95% confidence level. Interpret the sampling error.

Our best estimate of the percentage is 40/100 = 0.4. The sampling error is calculated from formula obtained from Unit 7.2:

$$p \pm K_{\alpha/2}\sqrt{p(1-p)/n}$$

$$\text{for } a = 0.05:$$

$$p = 40/100 = 0.40$$

$$CI:\ 0.40 \pm 1.96\sqrt{0.40(0.60)/100} \Rightarrow 0.40 \pm 0.049$$

Interpretation: We know the "true" value of the percentage of customers who will buy air conditioning equipment to within +/- 5%.

Exercise - Neighborhood Sample

Describe how you would obtain a simple random sample of people living in your neighborhood. Define the population and a frame for the population. Suppose this random sample was associated with a questionnaire survey. How would you obtain the elements of the sample?

Population: All human beings living in a defined geographic area (bounded, perhaps, by streets, or latitude/longitude).

Frame: A phone book is a natural place to start, but that will list the owners of phones (some may not own phones, some may have more than one listing). If the neighborhood has an association, perhaps there is a homeowner's listing, including parents and children. Question: Should visiting relatives be included in the population, or are you interested in permanent residents?

See *Simple Random Sampling* for the process you would use to obtain the sample elements.

Exercise - Employee Opinions

Suppose you wish to identify employee opinions about your company's Black Belt Process. You are interested in identifying differences in opinion by *department* and whether the employee is *management* or *staff.* Develop a method of sampling the employees that will achieve these goals.

The key here is that you are dealing with two *strata* in the population of employees. See Unit 7.3.6 for information on how to conduct this type of sampling.

Exercise - Counting "F's"

How many "f's" appear in the following paragraph? You have 15 seconds to count the "f's." Have several other people count the "f's" and compare notes. What does this tell you about the ability of 100% inspection to detect defects?

> THE NECESSITY OF TRAINING FARM HANDS FOR FIRST CLASS FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK IS FOREMOST IN THE MINDS OF FARM OWNERS. SINCE THE FOREFATHERS OF THE FARM OWNERS TRAINED THE FARM HANDS FOR FIRST CLASS FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK, THE FARM OWNERS FEEL THEY SHOULD CARRY ON WITH THE FAMILY TRADITION OF TRAINING FARM HANDS OF FIRST CLASS FARMS IN THE FATHERLY HANDLING OF FARM LIVESTOCK BECAUSE THEY BELIEVE IT IS THE BASIS OF GOOD FUNDAMENTAL FARM MANAGEMENT. FURTHERMORE, THE FUTURE FARMERS OF FINLAND PROVIDE FREE AND FIRST CLASS TRAINING FOR FUTURE FARM HANDS IN THE FINICKY FANTASTIC PECADILLOS OF FUTURE FARM ANIMALS.

We count 51 "F's."

Exercise - Fitness Center Usage

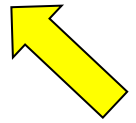
In a particular facility, 62% of the employees are non-management males, 31% are non-management females and the remaining 7% are at the level of supervisor or above. Your task is to estimate the percentage that would use a proposed fitness center. An "educated" guess is that 40-50% of the non-management females, 20-30% of the non-management males and 5-10% of the management would use the facilities. The total employee population is 3000.

What overall sample size would you need to provide a 95% confidence level with a precision of 5%?
How would you allocate that sample among the strata, using proportional allocation?

We set the answer up in Excel to calculate the overall sample size (n) and the strata (allocation column). See Unit 7.3.6 for relevant equations:

Sample Size Determination

Strata	Fraction (wi)	Number (Ni)	Will Use Center Est. Fraction (pi)	Numerator	Denominator	Allocation
Non-management Male	0.62	1860	0.45	1381050	460.35	**186**
Non-management Female	0.31	930	0.25	523125	174.375	**93**
Supervisory	0.07	210	0.075	43706.25	14.56875	**21**
	Sum	3000	**Sum**	1947881.25	649.29375	
	(N*E/K)^2	5856.93				
		n	**299.39**			





ni - sample size for ith strata
n - overall sample size
N - population size
Ni - size of ith strata
E - desired precision - 5%
K - 95% - 1.96
pi - estimate of proportion for ith strata
wi - Ni/N for proportionate sampling

Exercise - Control Circuits

A manufacturer of control circuits ships your pharmacy 100 boxes of circuit boards, each containing 100 boards. Suppose you are interested in measuring the input resistance of the boards (in ohms). Develop an approach to sample from these boxes that will provide you with the average resistance.

This is a two-stage sampling problem. The general approach is to take a random sample of the boxes, then a random sample of the boards within these boxes. See Unit 7.3.5 for details.

Exercise - Warranty Records

A set of 20,000 Warranty Records is stored in 400 file drawers, each containing 50 records. In a two-stage sample, five records are drawn at random from each of 80 randomly selected drawers. The between cluster variance is 362 and the within-cluster variance is 805. Compute the standard error of the mean per record.

We set this problem up in Excel:

The relevant equation from Unit 7.3 is:

For Continuous Characteristics

$$s_{\bar{X}} = \sqrt{\left(1 - \frac{m}{M}\right)\frac{s_b^2}{m} + \left(1 - \frac{n}{N}\right)\frac{s_w^2}{m \times n}} \quad \text{(FPCF Included)}$$

Note that we've performed the calculation with the Finite Population Correction Factor, despite the small sample size relative to the population size.

Two-Stage Sampling

m	80
M	400
n	5
N	50
s-b^2	362
s-w^2	805
s-X-bar	**2.330504**

Exercise - Lifestyle Survey

Suppose you wish to conduct a survey to determine the percentage of employees currently enrolled in a particular health plan. You suspect that "lifestyle status" is an important stratification variable; that is single employees are expected to participate less than are married employees with two or more children. You collect the following stratified sample, using proportionate sampling:

Stratum	**N_h**	**W_h**	**n_h**	**r_h**	**p_h**
Single	3,000	0.21	150	30	0.20
Married, 0 or 1 child	6,000	0.43	300	135	0.45
Married, 2 + children	5,000	0.36	250	200	0.80
Total	14,000	1.00	700	365	

The **n_h** column indicates the number of employees sampled, the **r_h** column is the number of employees in that stratum participating in the health plan and the **p_h** column are the proportion in each stratum participating. Calculate a) the overall sample proportion of employees participating in the health plan, b) the variance of the proportion, c) the standard error of the proportion, and d) a 95% confidence interval for the proportion.

Excel was used to set up the calculations for this problem. See Unit 7.3.6 for relevant equations. Given the sample proportion and the standard error of the proportion, the student should be able to construct the confidence interval by now.

Stratum	**N_h**	**W_h**	**n_h**	**r_h**	**p_h**		**Ni pi**	**Numerator**
Single	3,000	0.21	150	30	0.2		600	9664.43
Married, 0 or 1 child	6,000	0.43	300	135	0.45		2700	29799.33
Married, 2 + children	5,000	0.36	250	200	0.8		4000	16064.26
Totals	14,000	1	700	365		Sum	**7300**	**55528.02**

Overall Sample Proportion:

$$p = \frac{1}{N}\sum_{i=1}^{k} N_i p_i$$ **0.52**

Proportion Variance:

$$s_p^2 = \frac{1}{N^2}\sum_{i=1}^{k} N_i^2 \frac{p_i(1-p_i)}{n_i - 1}$$ **0.000283**

Standard Error of the Proportion: **0.016832**

Exercise - Community Programs

In the last year, there have been 78 community groups (each composed of 24 members) who have listened to talks by the company's public relations director. The director would like to assess, through personal interview, the attitude of these clubs towards the company's community programs. The primary variable of interest is the *proportion of those with a favorable attitude toward corporate community programs.* She decides on a cluster sample, since she realizes that each of the community groups are scattered geographically, and she wishes to reduce the cost of travel and field work time.

She wishes to establish a 95% confidence interval on the proportion, with a 9% error of the estimate. She decides to sample 11 groups. Do you agree with this number? (Assume a between cluster variance (s_b^2) of 15.55) She collects the following 11 clusters:

Group	A	B	C	D	E	F	G	H	I	J	K
# Favorable	9	11	13	15	16	17	18	20	20	21	16
# in Group	24	24	24	24	24	24	24	24	24	24	24

Calculate a) the overall proportion favorable, b) the variance of this estimate, and c) the 95% confidence interval.

The first question is answered by calculating the minimum sample size using the formula from Unit 7.3.7:

$$n = \frac{M\sigma_c^2}{\frac{ME^2\overline{N}^2}{K_{\alpha/2}^2} + \sigma_c^2} = \frac{78 \times 15.55}{\frac{78 \times 0.09^2 \times 24^2}{1.96^2} + 15.55} = 11$$

It appears as though our director will be able to achieve the precision she desires. The proportion favorable is simply the quotient of the total number favorable and the total number sampled:

$$p = 176 / 264 = 0.67$$

The variance of this estimate is given by:

$$s_p^2 = \left(\frac{M-m}{Mm\overline{N}^2}\right)\frac{1}{m-1}\sum_{i=1}^{m}(a_i - pN_i)^2 = \left(\frac{78-11}{78 \times 11 \times 24^2}\right)\frac{1}{11-1}146 = 0.001979$$

The confidence interval calculation is left for the student.

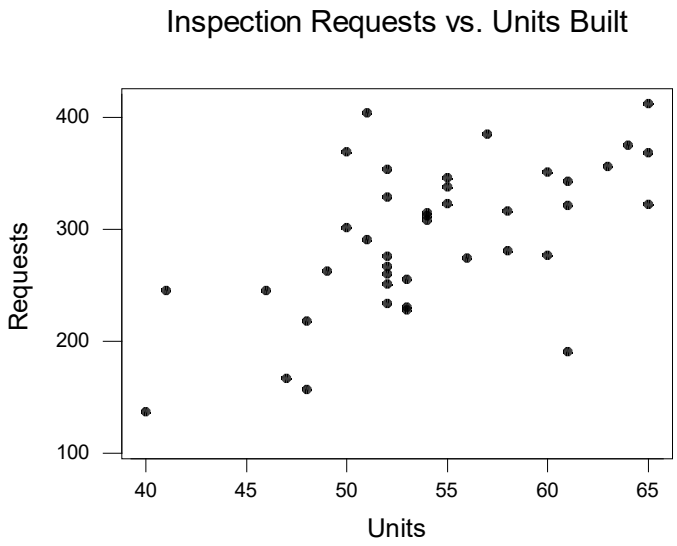
Section 10 – Relationships Between Variables

Exercise - QC Workload

A QC supervisor is trying to predict the workload her inspectors will experience each day. She thinks that the number of units being built each day will influence the number of inspection requests they receive each day. Perform correlation/regression analyses for the data shown below. Can she use the number of units as a workload predictor?

Gettysburg Plant						**Durango Plant**				
(#Units/# Requests)						**(# Units/# Requests)**				
55	323		50	301		52	269		35	224
61	191		48	157		63	281		41	224
52	353		52	276		71	247		42	285
58	316		51	291		66	235		37	291
54	312		60	277		70	228		44	261
61	321		54	315		66	277		46	225
65	368		60	351		71	258		43	252
54	308		64	375		67	258		49	282
52	260		55	338		63	259		39	255
50	369		53	230		57	187		41	227
55	346		61	343		65	255		39	191
52	234		58	281		78	232		45	283
52	267		52	251		69	218		52	273
49	263		65	412		60	248		54	281
48	218		53	255		66	344		52	308
46	245		53	228		73	316		48	245
47	167		56	274		80	308		41	295
40	137		51	404		60	255		47	250
41	245		63	356		34	93		48	199
52	329		65	322		46	234		48	228
57	385					41	295		41	131

The Scatter Plot for the Gettysburg Plant provides some evidence that a relationship exists:



The correlation/regression analysis from Minitab (Stat>Regression>Regression) appears below:

Regression Analysis

The regression equation is

Requests = - 59.4 + 6.48 Units

Predictor	Coef	StDev	T	P
Constant	-59.38	76.87	-0.77	0.444
Units	6.485	1.408	4.61	0.000

S = 53.87 R-Sq = 35.2% R-Sq(adj) = 33.6%

Analysis of Variance

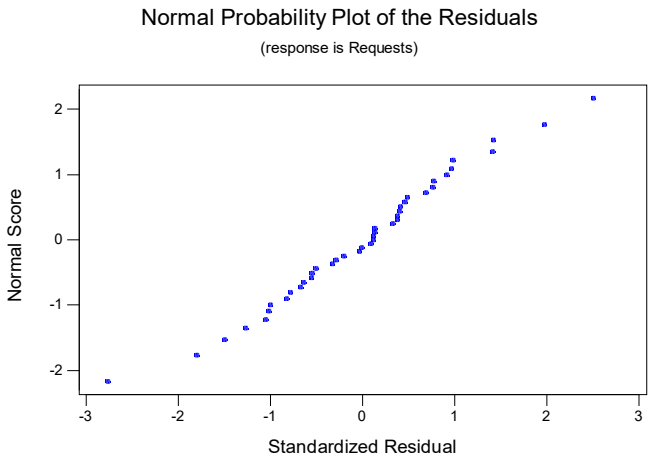
Source	DF	SS	MS	F	P
Regression	1	61568	61568	21.21	0.000
Residual Error	39	113193	2902		
Total	40	174760			

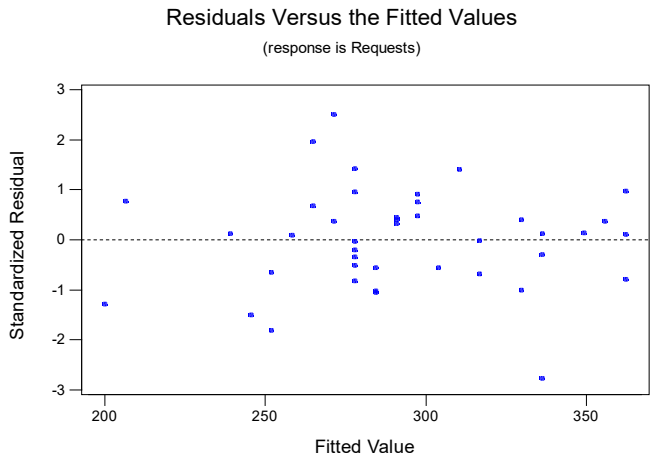
Unusual Observations

Obs	Units	Requests	Fit	StDev Fit	Residual	St Resid
2	61.0	191.00	336.19	12.67	-145.19	-2.77R
18	40.0	137.00	200.01	21.78	-63.01	-1.28 X
39	51.0	404.00	271.34	9.59	132.66	2.50R

R denotes an observation with a large standardized residual
X denotes an observation whose X value gives it large influence.

Interpretation: Minitab provides the **regression equation**. The R-Sq of 35.2% is the **coefficient of determination**; about 35% of the variation in the dependent variable (**requests**) can be explained by variation in the independent variable (**Units**). The **p-value** of 0.000 listed on the **Analysis of Variance** table indicates that the variation due to the regression is **significant** – that is; we reject the null hypothesis that there is *no relationship* between the dependent and independent variables. Minitab also lists **unusual observations** for you to examine. Finally, there do not appear to be any problems with the residuals. The standardized residuals fall along a fairly straight line on the **normal probability plot** and there are no unusual patterns in the plot of **residuals versus the fitted values**:





The student should take the data from the Durango Plant (and perhaps the combined plants) and repeat this analysis.

Exercise - Inspection Time

The same QC supervisor is investigating the number of hours expended completing inspections. The following data was obtained by one of the analysts working on the problem:

Day	# Requests	# Hours
1	215	16.0
2	210	11.6
3	230	27.2
4	240	35.6
5	220	19.0
6	225	23.2
7	230	28.6
8	215	15.4
9	235	31.2
10	230	28.0

If you want to predict the number of hours required to perform a certain number of inspections, what is the independent variable? What is the dependent variable? **Independent**: Number of Hours, **Dependent**: Number of Inspections

If you want to predict the number of requests that can be completed in a given number of hours, what is the independent variable? What is the dependent variable? **Independent:** Number of Inspections, **Dependent**: Number of Hours

Which is a better prediction statement? It seems that the number of inspections is a better **independent** variable – that drives the number of hours expended.

Perform a correlation/regression analysis of this data. Interpret the analysis.

Minitab was used to perform the regression analysis:

Regression Analysis

```
The regression equation is
Hours = - 157 + 0.802 Requests

Predictor        Coef       StDev          T        P
Constant     -156.949       4.042     -38.83    0.000
Requests      0.80235     0.01795      44.70    0.000
```

```
S = 0.5234      R-Sq = 99.6%     R-Sq(adj) = 99.6%

Analysis of Variance

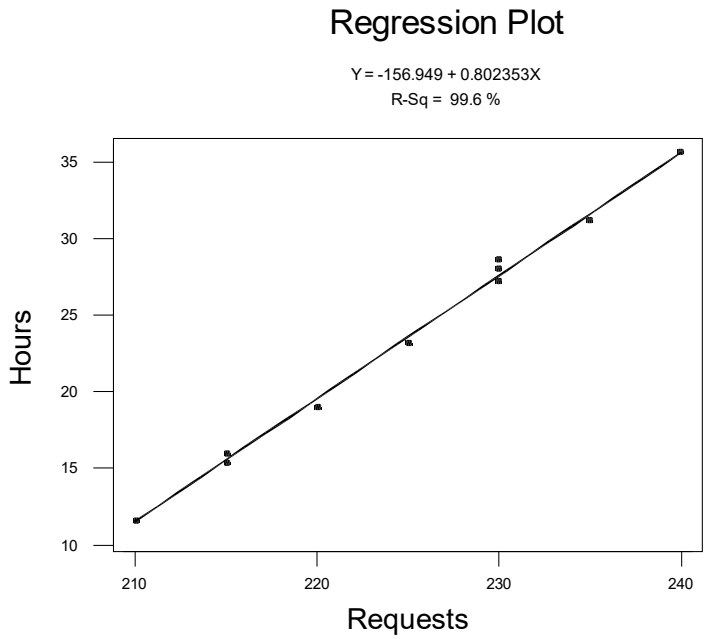
Source            DF          SS          MS          F        P
Regression         1      547.20      547.20    1997.74    0.000
Residual Error     8        2.19        0.27
Total              9      549.40

Unusual Observations
Obs    Requests      Hours         Fit    StDev Fit    Residual    St Resid
  7         230     28.600      27.592        0.188       1.008        2.06R

R denotes an observation with a large standardized residual
```

Interpretation: The r-squared value of 99.6% shows that the **number of requests** explains virtually all of the variation in the **number of hours** expended. The F-statistic and associated p-value on the Analysis of Variance table reflects this.

The fitted line plot appears to the right. Although we've done the regression first in this example, the scatter plot should really be examined first. This plot shows an unusually strong correlation between these two variables – what do you think could be happening here?



Exercise - One-Way ANOVA

Do a One-Way ANOVA on the following data, just to get used to the calculations:

	Factor Level				
Measurement	A_1	A_2	A_3	A_4	A_5
	16	8	2	8	20
	12	-4	4	12	16
	14	0	0	10	14
	10	-4	-2	10	8
	16	6	-6	8	18

Minitab was employed to perform the ANOVA (Stat>ANOVA>One-Way - Note that the data must be stacked to use this option; alternatively you can do a One-Way Unstacked ANOVA):

One-way Analysis of Variance

```
Analysis of Variance for Data
Source     DF        SS        MS         F        P
Factor L    4    1012.2     253.0     16.47    0.000
Error      20     307.2      15.4
Total      24    1319.4

                                 Individual 95% CIs For Mean
                                 Based on Pooled StDev
Level       N      Mean     StDev  ------+---------+---------+---------+
1           5    13.600     2.608                       (----*-----)
2           5     1.200     5.586    (-----*----)
3           5    -0.400     3.847  (----*-----)
4           5     9.600     1.673               (-----*----)
5           5    15.200     4.604                         (-----*----)
                                 ------+---------+---------+---------+
Pooled StDev =    3.919              0.0       7.0      14.0      21.0
```

Interpretation: The p-value of 0.000 allows us to reject the null hypothesis that the factor level does not affect the result. If we were trying to **maximize** the results of this experiment, factor levels 1 and 5 provide the best results. The student should also examine the residual plots (histogram, normal probability plot, etc.) to ensure there are no problems here.

Exercise - Two-Way ANOVA

Using the following data, perform a Two-Way ANOVA:

		Factor Level				
		A_1	A_2	A_3	A_4	A_5
Factor Level	B_1	16	8	4	18	20
	B_2	12	-4	2	12	16
	B_3	14	0	0	10	12
	B_4	10	-4	-2	6	8
	B_5	6	-6	-6	8	8

Minitab was used to perform the analysis (Stat>ANOVA>Two-Way):

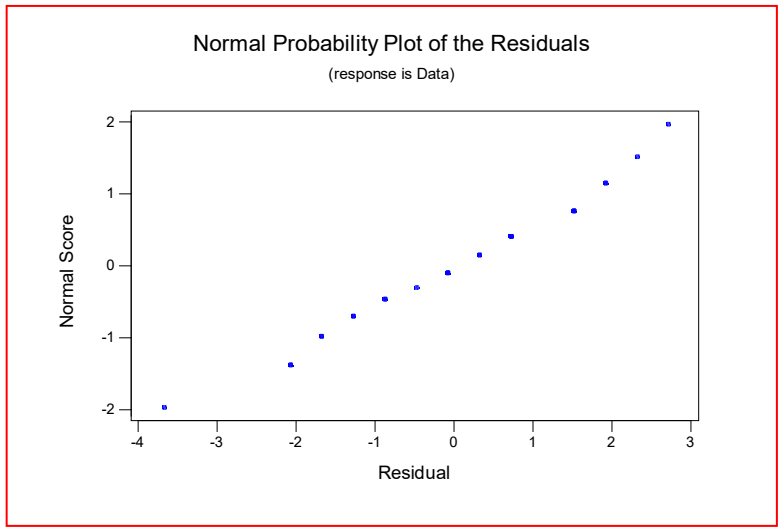
Two-way Analysis of Variance

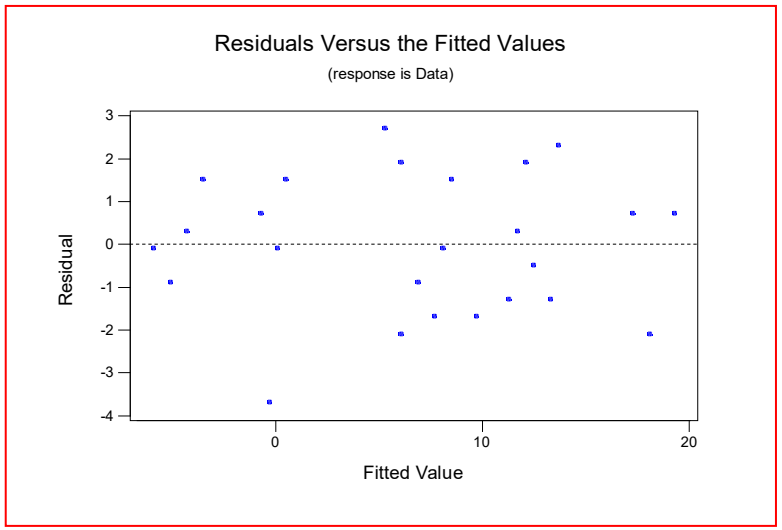
```
Analysis of Variance for Data
Source        DF        SS        MS        F        P
Row            4    375.04     93.76    24.29    0.000
Column         4    954.24    238.56    61.80    0.000
Error         16     61.76      3.86
Total         24   1391.04

                      Individual 95% CI
Row             Mean  ---------+---------+---------+---------+-
1              13.20                                (----*----)
2               7.60                    (----*----)
3               7.20                   (----*----)
4               3.60        (----*----)
5               2.00   (----*----)
                      ---------+---------+---------+---------+-
                             4.00      8.00     12.00     16.00
```

```
                      Individual 95% CI
Column          Mean  -------+---------+---------+---------+----
1              11.60                              (---*---)
2              -1.20    (---*--)
3              -0.40     (---*---)
4              10.80                          (---*--)
5              12.80                             (---*--)
                      -------+---------+---------+---------+----
                          0.00      5.00     10.00     15.00
```

Interpretation: Both factors are significant here, based on the p-values. The mean plots shows which factor levels are "best." A Two-Way ANOVA without repetitions cannot detect interactions between the factors. The residuals do not reveal any problem with the data:





Exercise - Two-Way ANOVA (with Repetition)

Using the following data, perform a Two-Way ANOVA, with repetition:

		Factor Level				
		A_1	A_2	A_3	A_4	A_5
Factor Level	B_1	16	8	4	18	20
	B_2	12	-4	2	12	16
	B_3	14	0	0	10	12
	B_4	10	-4	-2	6	8
	B_5	6	-6	-6	8	8
	B_1	14	11	6	21	23
	B_2	11	-2	5	15	13
	B_3	12	3	-2	8	14
	B_4	11	-1	0	9	11
	B_5	8	-3	-8	11	9

Minitab was used to analyze the data (Stat>ANOVA>Two-Way):

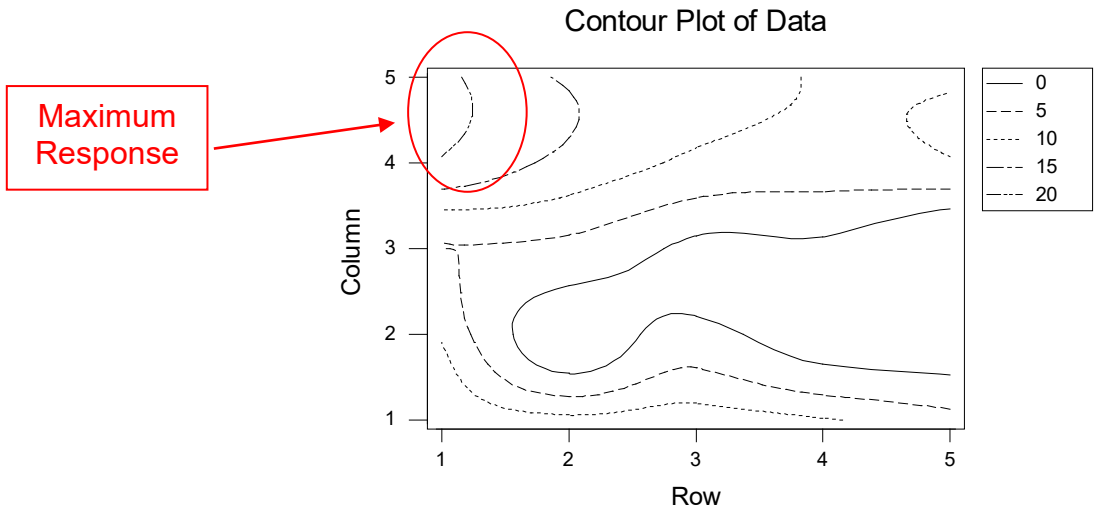
Two-way Analysis of Variance

```
Analysis of Variance for Data
Source         DF        SS        MS         F        P
Row             4    741.72    185.43    61.40    0.000
Column          4   1794.32    448.58   148.54    0.000
Interaction    16    161.68     10.11     3.35    0.003
Error          25     75.50      3.02
Total          49   2773.22

                        Individual 95% CI
Row             Mean    ------+---------+---------+---------+-----
1              14.10                                    (--*---)
2               8.00                    (--*--)
3               7.10                 (--*---)
4               4.80          (---*--)
5               2.70    (---*--)
                        ------+---------+---------+---------+-----
                           3.50      7.00     10.50     14.00
```

```
                      Individual 95% CI
Column          Mean  ----+---------+---------+---------+-------
1              11.40                                (--*-)
2               0.20     (-*--)
3              -0.10    (--*--)
4              11.80                                 (--*-)
5              13.40                                     (--*-)
                      ----+---------+---------+---------+-------
                       0.00      4.00      8.00     12.00
```

Interpretation: Both factors are significant *and* there is an interaction between the factors, based on the Analysis of Variance table. Based on the mean plots, level **one** of the **B** (row) factor and levels **one, four** and **five** of the **A** (column) factor will maximize response. A contour plot (Graph>Contour Plot) can help understand how the response variable is affected by the factors. The maximum response is obtained for factor B, level one and factor A, level 5:



Exercise - Switchboard Answer Time

A call center supervisor is trying to improve service at the center. The staffing level at the center is varied and the average time to answer a call recorded for three days at each staffing level (three repetitions). Perform an ANOVA to determine if staffing level makes a difference. Test at α = 0.05.

		Staffing Level (#)		
		2	**4**	**6**
Repetition	**1**	30	18	19
	2	19	21	16
	3	28	17	15

Measurements are Average Time to Answer Call in Seconds

The Minitab output (One-Way ANOVA) appears below:

One-way Analysis of Variance

```
Analysis of Variance for Answer T
Source     DF        SS        MS        F        P
Staff       2     134.0      67.0     4.67    0.060
Error       6      86.0      14.3
Total       8     220.0

                                 Individual 95% CIs For Mean
                                 Based on Pooled StDev
Level       N      Mean     StDev  --+---------+---------+---------+----
2           3    25.667     5.859                  (--------*--------)
4           3    18.667     2.082       (--------*--------)
6           3    16.667     2.082    (--------*--------)
                                 --+---------+---------+---------+----
Pooled StDev =    3.786           12.0      18.0      24.0      30.0
```

Interpretation: The p-value of 0.6 is not low enough to reject the null hypothesis. Based on this experiment, the supervisor cannot conclude that staffing level makes a difference in the average time to answer a call.

Exercise - Scale Removal

Three methods for removing scale in cooling tubes are being investigated. The methods' effectiveness is measured by the before & after tube relative cross-section. For example, if a clogged tube had only 5% of its normal area available for water flow before the method was used, and 50% of its area available after the treatment, the measurement is 50% - 5% = 45%. The experimental data is shown below. Perform an ANOVA to determine if there is a significant difference in method (at an $\alpha = 0.05$). Develop an interval estimate for the best method.

		Scale Removal Method		
		A	**B**	**C**
Repetition	**1**	27	15	48
	2	33	18	45
	3	36	24	57
	4	30	21	54

A One-Way ANOVA (Unstacked) of this data was performed in Minitab with the following results:

One-way Analysis of Variance

```
Analysis of Variance
Source      DF        SS        MS         F        P
Factor       2    2022.0    1011.0     50.55    0.000
Error        9     180.0      20.0
Total       11    2202.0

                                 Individual 95% CIs For Mean
                                 Based on Pooled StDev
Level        N      Mean     StDev  --------+---------+---------+--------
A            4    31.500     3.873            (---*---)
B            4    19.500     3.873  (---*---)
C            4    51.000     5.477                             (---*----)
                                     --------+---------+---------+--------
Pooled StDev =    4.472               24        36        48
```

Interpretation: The p-value of 0.000 allows us to reject the null hypothesis that the removal methods do not make a difference. The mean plot shows that method C provides the best scale removal.

Exercise – Welding Process Improvement

A team is attempting to improve an arc welding process for pressure vessels. Two factors are being explored – the *type of weld electrode* and the *weld current.* Four different weld rods were examined and three values of weld current. Two repetitions of the experiments were run, with the *tensile strength* of the weld sample being measured. Perform an ANOVA to determine if there is a significant difference in either rod or current (at an $\alpha = 0.05$).

		Weld Current Method		
		A	**B**	**C**
Weld Electrode	**1**	71, 69	69, 70	71, 79
	2	66, 64	62, 65	76, 72
	3	61, 58	56, 59	74, 70
	4	68, 62	62, 64	77, 78

Note: Measurements are in PSIG/1000.

Minitab was used to analyze this data:

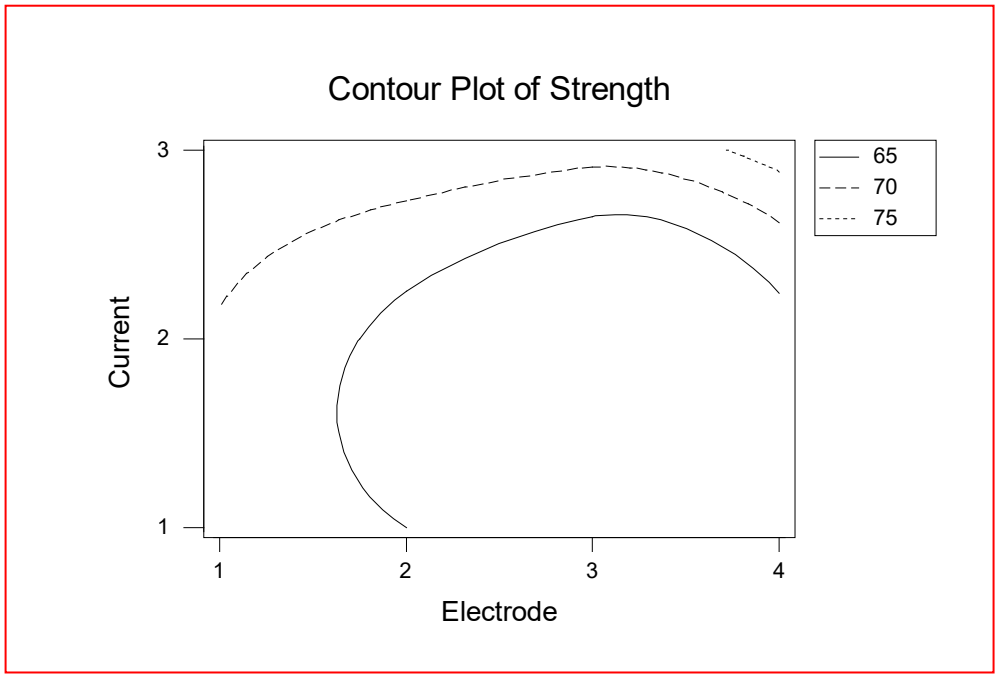
Two-way Analysis of Variance

```
Analysis of Variance for Strength
Source        DF        SS        MS        F        P
Electrod       3    223.13     74.38    10.32    0.001
Current        2    597.00    298.50    41.41    0.000
Interaction    6     63.00     10.50     1.46    0.272
Error         12     86.50      7.21
Total         23    969.63

                        Individual 95% CI
Electrod        Mean    -------+---------+---------+---------+----
1               71.5                              (------*------)
2               67.5                    (------*------)
3               63.0    (------*------)
4               68.5                         (------*------)
                        -------+---------+---------+---------+----
                            63.0      66.5      70.0      73.5
```

```
                      Individual 95% CI
Current         Mean  -------+---------+---------+---------+----
1              64.87       (----*----)
2              63.38  (----*-----)
3              74.62                              (-----*----)
                      -------+---------+---------+---------+----
                         64.00     68.00     72.00     76.00
```

Interpretation: Both factors are significant, based on the p-values of the ANOVA table. There is no evidence of an interaction between the factors, again, based on the interaction p-value of 0.272. The mean plots show Electrode **1** and Current **3** (Level **C**) provide the best tensile strength. The contour plot, though, provides a slightly different picture. The highest strength contour (75) occurs for Electrode **4** and Current 3 (Level **C**). It's always a good idea to examine the data from a number of angles:



Section 11 - Experimentation

Objective: To practice applying the *max-min-con* principles of experimentation.

Instructions:

1. Horatio Hornblower is trying to optimize ***range*** achieved by the new 16 lb. cannon delivered to his ship by the British Admiralty. Brainstorm factors that he could experiment with to increase the cannon's range (*Hint: Develop a Cause and Effect diagram)*.

2. Choose one of the factors that he could use in a *single factor experiment.* Apply the *min-max-con* principles to this and the other factors identified above.

Time: 20 minutes

For those not familiar with a 19th century sailing warship's armament, a few factors could include:

- Cannon angle
- Size of cannon ball
- Amount of gunpowder
- Amount of wadding
- Tamping process
- Bore cleanliness
- Height of cannon above waterline
- Gun recoil

Note that Horatio will have to deal with some *noise factors* over which he may have little control

- Ambient temperature
- Pitch, Yaw and Roll of the ship
- Wind speed and direction
- Wave height

He will also have to address the Measurement System to ensure this doesn't introduce to much variation in his experiments.

Objective: To practice enumerating the combinations of experimental factors in a full-factorial design.

Instructions: A team is attempting to improve their protocol for treating patients with tachycardia (abnormal rapidity of heart action). The usual treatments include pressure on one or both carotid sinuses, pressure on the eyeballs, induction of gagging or vomiting, attempted expiration with glottis closed, lying down with feet in the air and bending over. The treatments are sometimes effective when administered singly. Often, though, two or more combinations of treatments are required to slow the patient's heart rate. List the combinations of two and three treatments the team would need to investigate in this process.

Time: 20 minutes

Two Factor Treatments (assuming they are being performed *simultaneously*, not *sequentially*):

	Carotid Sinus Pressure	Eyeball Pressure	Gagging Induction	Glottis Closed Expiration	Lying Down	Bending Over
Carotid Sinus Pressure		P	P	P	P	P
Eyeball Pressure			P	P	P	P
Gagging Induction				P	P	P
Glottis Closed Expiration					P	P
Lying Down						N
Bending Over						

P – Possible
N – Not Possible

Objective: To practice setting up an experimental design, to contrast the full-factorial and fractional-factorial methods.

Instructions:

1. A team is trying to reduce brazing failures on impellers. They have identified three braze application factors (B_1, B_2 & B_3), two heat treat factors (H_1 & H_2) and three braze material factors (M_1, M_2 & M_3) that are suspected of affecting the failure rate. If each factor has two levels, design an experiment to investigate the effects of these factors. Only two factors, B_2 and M_2 are suspected of interaction. Compare the number of experimental combinations if a *full-factorial* design is employed, vs. an *orthogonal array (fractional factorial).* What's lost when a *fractional factorial design* is used?

Time: 20 minutes

Here, there are eight factors, each run at two levels. The number of runs for a full-factorial experiment (with only one replication) is $2^8 = 256$.

On the other hand, an L_{16} orthogonal array has 15 columns to which you can assign factors and interactions (if you are using ANOVA to analyze the results, you will need to leave at least one column unassigned). The L_{16} array required 16 trials (for one replication).

The key decision here involves the interactions expected among the factors. If there is evidence that no or only a few interactions are possible, then the fractional factorial (e.g. orthogonal array) can be employed without worrying about confounding.

Objective: To practice analyzing the results of a real world experiment.

Instructions:

1. An orthogonal array was used to conduct an experiment. Five factors were included in the experiment (D, AS, E2, C, SS).

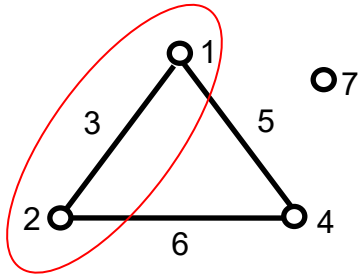
2. "Reverse engineer" the experiment – was the orthogonal array correctly employed (i.e. can the third column be used to detect the interaction of D and AS?)? If all factors are significant, can an ANOVA be performed?

3. Analyze the experiment. Which factors are significant? If the experiment was designed to *maximize* the response, which factor levels are best? Are the residuals "OK?"

Time: 20 minutes

Column		**1**	**2**	**3**	**4**	**5**	**6**	**7**	
StdOrder	**RunOrder**	**D**	**AS**	**D* AS**	**E2**	**C**	**SS**		**Response**
13	1	1	1	1	1	1	1	1	10
7	2	1	1	1	2	2	2	2	26
1	3	1	2	2	1	1	2	2	57
14	4	1	2	2	2	2	1	1	14
3	5	2	1	2	1	2	1	2	0
11	6	2	1	2	2	1	2	1	16
16	7	2	2	1	1	2	2	1	38
8	8	2	2	1	2	1	1	2	22
12	9	1	1	1	1	1	1	1	10
15	10	1	1	1	2	2	2	2	28
9	11	1	2	2	1	1	2	2	64
6	12	1	2	2	2	2	1	1	39
10	13	2	1	2	1	2	1	2	0
4	14	2	1	2	2	1	2	1	17
2	15	2	2	1	1	2	2	1	36
5	16	2	2	1	2	1	1	2	23

Using an L_8 orthogonal array allows the experimenter to employ 6 factors/interactions and leave one column available for the error term. Assigning column 1 to factor **D** and column 2 to factor **AS** allows the interaction **D*AS** to be detected in the third column:



Minitab was employed to analyze the data (Stat>DOE>General Linear Model):

General Linear Model

```
Factor     Type Levels Values
D         fixed      2 1 2
AS        fixed      2 1 2
D* AS     fixed      2 1 2
E2        fixed      2 1 2
C         fixed      2 1 2
SS        fixed      2 1 2

Analysis of Variance for Response, using Adjusted SS for Tests

Source     DF     Seq SS     Adj SS     Adj MS       F      P
D           1     576.00     576.00     576.00   11.73  0.008
AS          1    2162.25    2162.25    2162.25   44.03  0.000
D* AS       1      12.25      12.25      12.25    0.25  0.629
E2          1      56.25      56.25      56.25    1.15  0.312
C           1      90.25      90.25      90.25    1.84  0.208
SS          1    1681.00    1681.00    1681.00   34.23  0.000
Error       9     442.00     442.00      49.11
Total      15    5020.00

Unusual Observations for Response

Obs  Response       Fit   StDev Fit  Residual   St Resid
  4   14.0000   29.0000      4.6353  -15.0000     -2.85R

R denotes an observation with a large standardized residual.
```

Interpretation: Factors **D**, **AS**, and **SS** are significant, based on the p-values of the ANOVA table. A followup regression analysis shows that the best settings of these factors are **D – 1**, **AS – 2**, and **SS – 2** (assuming the intent is to maximize the response. The three variables account for over 85% of the variation in the response.

Regression Analysis

```
The regression equation is
Response = - 22.6 - 12.0 D + 23.2 AS + 20.5 SS
```

```
Predictor        Coef       StDev          T        P
Constant      -22.625       9.360      -2.42    0.032
D             -12.000       3.538      -3.39    0.005
AS             23.250       3.538       6.57    0.000
SS             20.500       3.538       5.79    0.000

S = 7.075       R-Sq = 88.0%     R-Sq(adj) = 85.0%
```

Objective: To practice analyzing the results of a real world experiment.

Instructions:

1. A full-factorial array was used to conduct an experiment on a brazing operation. Four factors were included in the experiment (Tube Cut, Skill, Flux and Cleaned). The output is fraction of rejected brazes from the process.

2. Analyze the experiment. Which factors are significant? If the experiment was designed to *minimize* the response, which factor levels are best? Are the residuals "OK?"

Time: 20 minutes

StdOrder	Tube Cut	Skill	Flux	Cleaned	Results
1	-1	-1	-1	-1	0.215
2	1	-1	-1	-1	0.269
3	-1	1	-1	-1	0.184
4	1	1	-1	-1	0.258
5	-1	-1	1	-1	0.146
6	1	-1	1	-1	0.344
7	-1	1	1	-1	0.107
8	1	1	1	-1	0.330
9	-1	-1	-1	1	0.200
10	1	-1	-1	1	0.430
11	-1	1	-1	1	0.123
12	1	1	-1	1	0.209
13	-1	-1	1	1	0.121
14	1	-1	1	1	0.311
15	-1	1	1	1	0.200
16	1	1	1	1	0.280

Minitab was used to analyze the results of this experiment.

Fractional Factorial Fit - All Factors & Interactions

```
Estimated Effects and Coefficients for Results (coded units)

Term          Effect      Coef
Constant                0.23294
A            0.14187    0.07094
B           -0.04312   -0.02156
```

```
C          -0.00612  -0.00306
D           0.00263   0.00131
A*B        -0.02613  -0.01306
A*C         0.03088   0.01544
A*D         0.00463   0.00231
B*C         0.04187   0.02094
B*D        -0.01937  -0.00969
C*D        -0.00638  -0.00319
A*B*C       0.00488   0.00244
A*B*D      -0.03738  -0.01869
A*C*D      -0.04238  -0.02119
B*C*D       0.04463   0.02231
A*B*C*D     0.00362   0.00181

Analysis of Variance for Results (coded units)

Source                DF      Seq SS      Adj SS      Adj MS      F      P
Main Effects           4    0.088131   0.0881307   0.0220327      *      *
2-Way Interactions     6    0.015307   0.0153069   0.0025511      *      *
3-Way Interactions     4    0.020831   0.0208308   0.0052077      *      *
4-Way Interactions     1    0.000053   0.0000526   0.0000526      *      *
Residual Error         0    0.000000   0.0000000   0.0000000
Total                 15    0.124321
```

F-statistics cannot be calculated – no error term available

Interpretation: In the analysis above, all factors and interactions were included in the model. Note that no error term can be calculated and hence no significance (i.e. p-values) assessed. In the second analysis, the four-factor interaction was removed from the model (under the assumption that this high-order interaction is not important). Here, the only significant factor (α = 0.05) is the **Tube Cut**, although some of the other factors/interactions are "close:" **Skill, Skill*Flux, Tube Cut*Skill*Cleaned, Skill*Flux*Cleaned** and **Tube Cut*Flux*Cleaned**. At this early stage in experimentation, a higher alpha may be set, recognizing the risk of claiming a factor is significant when it is not.

Fractional Factorial Fit – Four-Factor Interaction Removed

```
Estimated Effects and Coefficients for Response (coded units)

Term                    Effect      Coef  StDev Coef       T      P
Constant                         0.23294    0.001812  128.52  0.005
Tube Cut               0.14187   0.07094    0.001812   39.14  0.016
Skill                 -0.04312  -0.02156    0.001812  -11.90  0.053
Flux                  -0.00612  -0.00306    0.001812   -1.69  0.340
Cleaned                0.00263   0.00131    0.001812    0.72  0.601
```

Tube Cut*Skill	-0.02613	-0.01306	0.001812	-7.21	0.088
Tube Cut*Flux	0.03088	0.01544	0.001812	8.52	0.074
Tube Cut*Cleaned	0.00463	0.00231	0.001812	1.28	0.423
Skill*Flux	0.04187	0.02094	0.001812	11.55	**0.055**
Skill*Cleaned	-0.01937	-0.00969	0.001812	-5.34	0.118
Flux*Cleaned	-0.00638	-0.00319	0.001812	-1.76	0.329
Tube Cut*Skill*Flux	0.00488	0.00244	0.001812	1.34	0.407
Tube Cut*Skill*Cleaned	-0.03738	-0.01869	0.001812	-10.31	**0.062**
Tube Cut*Flux*Cleaned	-0.04238	-0.02119	0.001812	-11.69	**0.054**
Skill*Flux*Cleaned	0.04463	0.02231	0.001812	12.31	**0.052**

Analysis of Variance for Response (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	0.088131	0.0881307	0.0220327	419.17	0.037
2-Way Interactions	6	0.015307	0.0153069	0.0025511	48.54	0.109
3-Way Interactions	4	0.020831	0.0208308	0.0052077	99.08	0.075
Residual Error	1	0.000053	0.0000526	0.0000526		
Total	15	0.124321				

The best level of the **Tube Cut** factor (to minimize the response) Is **–1** (or low), for **Skill**, the **1** (or high) value is best.

Objective: To practice analyzing the results of a real world experiment.

Instructions:

1. A full-factorial array was used to conduct an experiment on a compressor design. Three factors were included in the experiment (Eductor, Pump and O-Ring). Two outputs were measured: Temperature and Pressure.

2. Analyze the experiment. Which factors are significant? If the experiment was designed to *minimize* the temperature response, which factor levels are best? If the experiment was designed to *maximize* the pressure response, which factor levels are best? Are the residuals "OK?" Are there any conflicts in factor levels relative to the two responses?

Time: 20 minutes

RunOrder	Eductor	Pump	O-ring	Temp	Pressure
1	1	-1	-1	133.30	40.89
2	1	-1	1	142.00	37.66
3	-1	1	-1	135.15	67.50
4	-1	-1	1	134.20	51.20
5	-1	-1	-1	134.15	54.47
6	1	1	-1	133.50	49.23
7	-1	1	1	135.00	59.04
8	1	1	1	143.00	44.74

Minitab was used to analyze the data. The three-way interaction was removed from the model to allow an error term to be calculated:

Fractional Factorial Fit

Estimated Effects and Coefficients for Temp (coded units)

Term	Effect	Coef	StDev Coef	T	P
Constant		136.287	0.1250	1090.30	0.001
Eductor	3.325	1.663	0.1250	13.30	**0.048**
Pump	0.750	0.375	0.1250	3.00	0.205
O-Ring	4.525	2.262	0.1250	18.10	**0.035**
Eductor*Pump	-0.150	-0.075	0.1250	-0.60	0.656
Eductor*O-Ring	4.575	2.288	0.1250	18.30	**0.035**
Pump*O-Ring	0.150	0.075	0.1250	0.60	0.656

Analysis of Variance for Temp (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	64.187	64.1875	21.3958	171.17	0.056
2-Way Interactions	3	41.951	41.9513	13.9838	111.87	0.069
Residual Error	1	0.125	0.1250	0.1250		
Total	7	106.264				

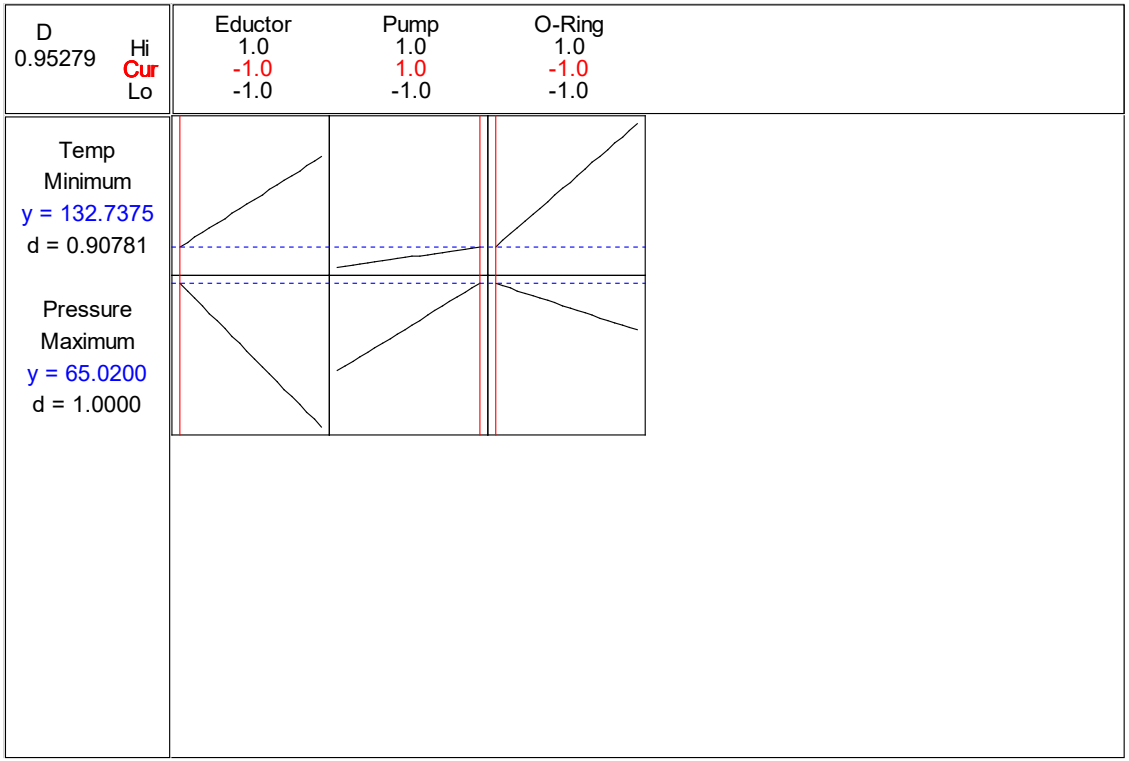
Estimated Effects and Coefficients for Pressure (coded units)

Term	Effect	Coef	StDev Coef	T	P
Constant		50.591	0.4913	102.98	0.006
Eductor	-14.923	-7.461	0.4913	-15.19	**0.042**
Pump	9.072	4.536	0.4913	9.23	0.069
O-Ring	-4.862	-2.431	0.4913	-4.95	0.127
Eductor*Pump	-1.362	-0.681	0.4913	-1.39	0.398
Eductor*O-Ring	1.002	0.501	0.4913	1.02	0.494
Pump*O-Ring	-1.612	-0.806	0.4913	-1.64	0.348

Analysis of Variance for Pressure (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	657.270	657.270	219.090	113.48	0.069
2-Way Interactions	3	10.923	10.923	3.641	1.89	0.481
Residual Error	1	1.931	1.931	1.931		
Total	7	670.124				

Interpretation: For **Temperature**, the Eductor and O-Ring are significant factors (as well as their interaction). To *minimize* temperature, pick the **–1** (low) values of these factors. For **Pressure**, the Eductor is significant, although the Pump is "close." To *maximize* pressure, pick the **–1** (low) value of the Eductor and the **1** (high) value of the Pump. Although this is a linear model, a feature of Minitab is shown below – the Multiple Response Optimizer. This feature allows you to explore the values of the factors that achieve the highest "desirability" – the "D" in the upper left corner of the graph.


D
0.95279
Hi
Cur
Lo
Eductor
1.0
-1.0
-1.0
Pump
1.0
1.0
-1.0
O-Ring
1.0
-1.0
-1.0
Temp
Minimum
y = 132.7375
d = 0.90781
Pressure
Maximum
y = 65.0200
d = 1.0000

Section 12 – Changing the Process

Exercise - Assembly Time

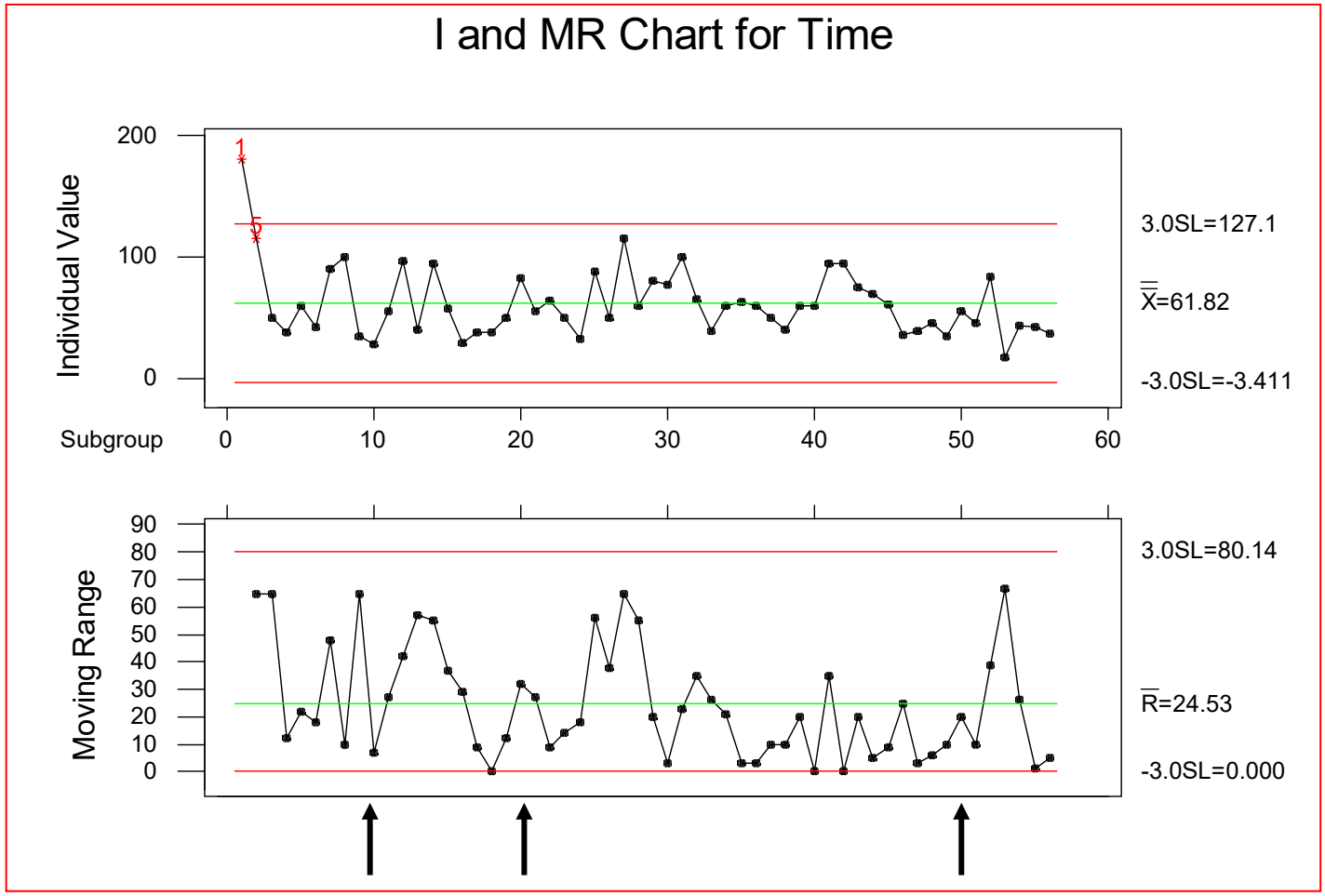
A team was working on improving the time needed to assemble an air handler. The team made three changes to the process. Did they improve the timeliness?

Assembly Time

Unit	Time	Unit	Time	Unit	Time	Unit	Time
1	180	*15*	58	*29*	80	*43*	75
2	115	*16*	29	*30*	77	*44*	70
3	50	*17*	38	*31*	100	*45*	61
4	38	*18*	38	*32*	65	*46*	36
5	60	*19*	50	*33*	39	*47*	39
6	42	*20*	82 **(2)**	*34*	60	*48*	45
7	90	*21*	55	*35*	63	*49*	35
8	100	*22*	64	*36*	60	*50*	55
9	35	*23*	50	*37*	50	*51*	45 **(3)**
10	28**(1)**	*24*	32	*38*	40	*52*	84
11	55	*25*	88	*39*	60	*53*	17
12	97	*26*	50	*40*	60	*54*	43
13	40	*27*	115	*41*	95	*55*	42
14	95	*28*	60	*42*	95	*56*	37

Changes made at (1), (2), and (3).

The Individuals Chart appears below. None of the changes resulted in an assignable cause (note there are a couple of "close" ones – the reduction in the moving range from data 40 to 50 and the last 11 points, 10 of which are below the center line. The team that collected this data was "sure" they had actually reduced the assembly time!



Section 14 – Design Management

Objective: Distinguish between customer needs, specifications, and product features.

Instructions:

1. Review the description of the laptop computer and battery shown below

2. Classify the list as follows:
   - *Customer Need* – A function required by the user of the product (N)
   - *Specification* – A performance requirement placed on the design (S)
   - *Product Feature*– A specific choice of component or part applied in the design (F)

3. For the specifications and product features, what customer need(s) do you think Dell is trying to meet?

30 minutes

Description of Dell Computer Corporation Laptop Computer:

F	Intel® Pentium® II Mobile Module microprocessor with MMX™ technology with 32 KB of internal cache
F	512 KB of pipelined-burst SRAM external cache
F	Hardware-accelerated PCI and AGP bus architecture that increases system performance; particularly video and hard disk drive performance
F	32 MB of SDRAM system memory minimum, with support for a maximum of 192 MB
F, S	Ultra DMA/33 data transfer protocol for ATA /IDE hard-disk drive interface. Ultra DMA/33 allows data transfer rates of up to 33 MB/sec
F	A combination module that contains a CD-ROM drive and a diskette drive
F	Built-in stereo speakers and microphone
F	Jacks for connecting external speakers, headphones, or an external microphone to the computer
F	S-video TV-out connector and composite TV-out adaptor cable that allows you to connect a television to your computer
F	An ATI video controller with an AGP 2X, 4 or 8 MB of video memory, 3D assist, dual-screen video, and flicker-free TV out
F, S	A lithium ion main battery and an optional secondary battery to double battery life
F, S	Two power management modes ¾ standby mode and save-to-disk suspend mode ¾ that help conserve battery power
F	A 14.1-inch active-matrix XGA color display
F	A built-in keyboard that includes two special keys that support the Microsoft® Windows® 98 operating system
F	A PS/2-compatible touch pad that provides the computer full mouse functionality
F, N	USB capability, that simplifies connecting peripheral devices such as mice, printers, and computer speakers. The USB connector on your computer's back panel provides a single connection point for multiple USB-compliant devices. USB-compliant devices can also be connected and disconnected while the system is running.
F, N	An options bay in which you can use a variety of combination modules, including the CD-ROM drive and diskette drive module, or the DVD-ROM drive and diskette drive module. In addition, you can use the options bay for a second battery.
F, N	An infrared port that permits file transfers without using cable connections.
F	PC Card slots with connectors for two 3.3- or 5-V cards. Both PC Card slots support CardBUS technology. In addition, a ZV Port is available from the lower slot (slot 0).

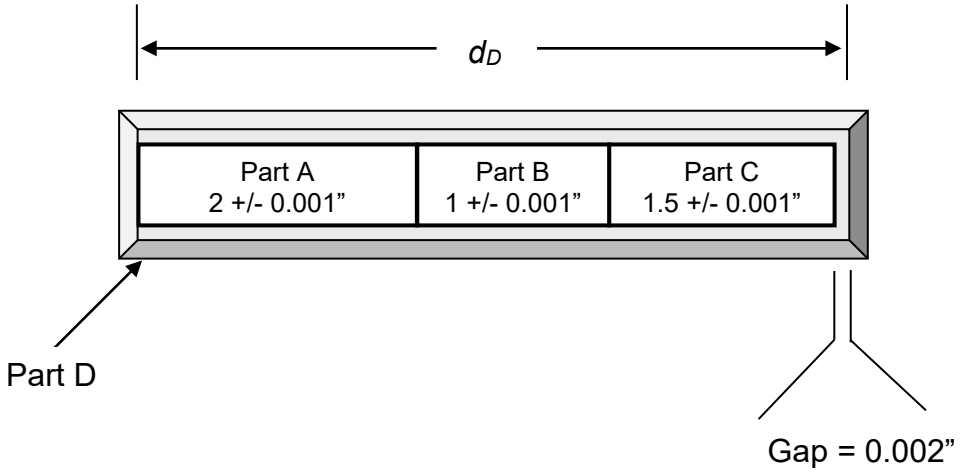
Objective: To practice setting tolerances for low level part characteristics.

Instructions:

1. *Worst Case Analysis* – If a gap of 0.002" is desired between the parts A, B and C and Part D, what should the **nominal** value be for dimension d_D? Part D can be fabricated to a tolerance of 0.001".

2. *Root Sum of Squares* – If the assembly is made from parts randomly selected from the production line, what fraction of the assemblies will be defective because of excess interference (assume that Part D's dimension d_D's nominal value is set based on the worst case analysis above)?

Time: 30 minutes



Worst Case Approach:

$$Gap = d_D - (d_A + d_B + d_c)$$

Worst Case Gap :

$$G_{WC} = (d_{D-Nom} - .001") - (2.001"+1.001"+1.501") = 0.002"$$

$$\therefore d_{D-Nom} = 4.506"$$

Root Sum of Squares Approach:

$$\overline{X}_{Gap} = 4.506 - (2.000 + 1.000 + 1.500) = 0.006"$$

$$s_{Gap} = \sqrt{s_A^2 + s_B^2 + s_C^2 + s_D^2} = \sqrt{4 \times 0.00033^2} = 0.00066"$$

$$P(X < 0.002 : 0.006, 0.00066) = P(Z < (0.002 - 0.006)/0.00066 : 0,1)$$

$$P(Z < -6.1 : 0,1) = 0.00053034 \times 10^{-6}$$ A rather small number!

Test yourself on these situations. Which is the best analogy to the process being benchmarked?

Process	Key Characteristics/Criteria	Possible Analogies
1. Brewing Beer	Chemical Process	**A.** Making Soda Syrup **B.** Manufacturing Shampoo **C.** Packaging Juice
2. Forming Uranium Pellets	Pressure Molding, Powder Molding	**A.** Extruding Pet Food Pellets **B.** Running X-Ray Equipment **C.** Manufacturing Aspirin
3. Retooling Production Lines In Real Time	Quick Reconfiguration	**A.** Changing Costumes Between Acts **B.** Servicing A Race Car During Pit stops **C.** Preparing Fast Food
4. Applying A Finish To Firearm Cartridges	Surface Smoothing	**A.** Polishing A Stainless Steel Knife **B.** Lacquering A Brass Lipstick Case **C.** Varnishing The Heel Of A Rifle

Answers: 1 – A or B, 2 – A, 3 – A, B, C, 4 - B

Objective: To practice calculating Loss Functions

Instructions: Develop Loss Functions for individual products and for the "population" of products for the scenarios shown below:

Time: 20 minutes

Product/ Characteristic	Specification or Function Limit	Loss at Function Limit	Current Population
Valve Stem Length	12" +/- 0.05"	$300.00	σ = 0.02"
Submarine Hull Steel - Tensile Strength	> 10,000 psi	$2 billion	σ^2 = 1.56E-10
Printed Circuit Board Chemical Contamination	< 0.1%	$1.00	σ^2 = 6.4E-3

Valve Stem – Nominal is Best

$$L(y) = \frac{A_0}{\Delta_0^2}\sigma^2 = \frac{\$300}{0.05^2} \times 0.02^2 = \$48$$

Tensile Strength – Larger is Better

$$L(y) = A_0\Delta_0^2\sigma^2 = \$2 \times 10^9 \times 10{,}000^2 \times 1.56 \times 10^{-9} = \$312{,}000{,}000$$

Chemical Contamination - Smaller is Better

$$L(y) = \frac{A_0}{\Delta_0^2}\sigma^2 = \frac{\$1.00}{.001^2} \times 6.4 \times 10^{-3} = \$6400$$

Objective: To practice calculating Signal-to-Noise Ratios

Instructions: Calculate Signal-to-Noise Ratios for the scenarios shown below:

Valve Stem Length (inches)	Submarine Hull Steel Tensile Strength (psig)	Printed Circuit Board Chemical Contamination (% Contamination)
11.9848	77500	0.0754
12.0202	74456	0.0553
12.0089	80340	0.0996
11.9813	79876	0.0837
12.0054	78636	0.0853
11.9952	83191	0.0963
11.9948	82626	0.0829
11.9912	77770	0.0965
12.0079	78675	0.0751
11.9688	79868	0.0632
11.9989	77656	0.1041
11.9931	79215	0.0782
12.0054	77038	0.0764
12.0094	84754	0.0559
12.0151	80044	0.083
11.9906	80822	0.0742
12.005	79381	0.0786
11.966	78669	0.0676
11.9811	79049	0.0725
11.9788	79981	0.0825

Valve Stem Length:

$$\eta = 10\log\frac{\frac{1}{n}(S_m - V_e)}{V_e} = 10\log\frac{\frac{1}{20}(2877.646081 - 0.00022)}{0.00022} = 58.1$$

where:

$$\mathrm{S_m} = \frac{1}{\mathrm{n}}\left(\sum y_i\right)^2 = \frac{1}{20}(239.9)^2 = 2877.646081$$

and

$$V_e = \frac{1}{n-1}\left(\sum y_i^2 - S_m\right) = \frac{1}{19}(2877.650331 - 2877.646081) = 0.00022$$

Tensile Strength:

$$\eta = -10\log\sigma^2 = -10\log(1.59\times10^{-10}) = 98$$

where:

$$\sigma^2 = \frac{1}{\mathrm{n}}\sum\frac{1}{y_i^2} = \frac{1}{20}(3.17\times10^{-9}) = 1.59\times10^{-10}$$

Chemical Contamination:

$$\eta = -10\log\sigma^2 = -10\log(0.0065) = 21.9$$

where:

$$\sigma^2 = \frac{1}{\mathrm{n}}\sum y_i^2 = \frac{1}{20}(0.129) = 0.0065$$

Objective: To practice applying Taguchi's Parameter Design Technique.

Instructions: The following experiments were run to determine factors important to reducing leaks in condenser tubing. An L-16 experiment was run with four factors being varied – tube cut, skill level, flux type and cleaning.

"Reverse engineer" this experiment. Take a blank L-16 array and determine which factors (below) were assigned to which columns. What interactions will this design detect? (Hint: try to figure out which Linear Graph was used first).

Next, determine which of these factors are important. Analyze the results using the Signal-to-Noise Ratio approach described in Unit 10.3.

Time: 60 minutes

Tube Cut	Skill	Flux	Cleaned	Results
-1	-1	-1	-1	0.215
1	-1	-1	-1	0.269
-1	1	-1	-1	0.184
1	1	-1	-1	0.258
-1	-1	1	-1	0.146
1	-1	1	-1	0.344
-1	1	1	-1	0.107
1	1	1	-1	0.330
-1	-1	-1	1	0.200
1	-1	-1	1	0.430
-1	1	-1	1	0.123
1	1	-1	1	0.209
-1	-1	1	1	0.121
1	-1	1	1	0.311
-1	1	1	1	0.200
1	1	1	1	0.280

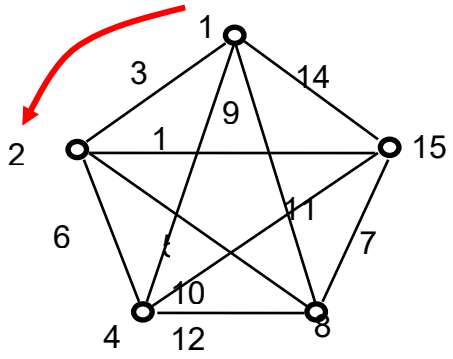
"Reverse Engineering" the Experiment – Here is the L-16 orthogonal array, with the factors assigned to columns. Follow the Linear Graph to see how the factors were assigned to detect interactions between factors. For example, given that **Cleaned** is assigned to **column 1** and **Flux** is assigned to **column 2**, the interaction between these factors (**CxF**) is detected in **column 3**:

Tube Cut	Skill	Flux	Cleaned
-1	-1	-1	-1
1	-1	-1	-1
-1	1	-1	-1
1	1	-1	-1
-1	-1	1	-1
1	-1	1	-1
-1	1	1	
1	1	1	
-1	-1	-1	1
1	-1	-1	1
-1	1	-1	1
1	1	-1	1
-1	-1	1	1
1	-1	1	1
-1	1	1	1
1	1	1	1

	Column (Factor/Interaction)															
Trial	1-Cleaned	2-Flux	CxF	4-Skill	CxS	FxS	7	8-Tube	CxT	FxT	11	SxT	13	14	15	Results
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.215
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	0.269
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2	0.184
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1	0.258
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2	0.146
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1	0.344
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1	0.107
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2	0.33
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	0.2
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1	0.43
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1	0.123
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2	0.209
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1	0.121
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2	0.311
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2	0.2
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1	0.28

Interactions Columns

	Clean	Flux	Skill	Tube
Clean	-	**3**	**5**	**9**
Flux		-	**6**	**10**
Skill			-	**12**
Tube				-

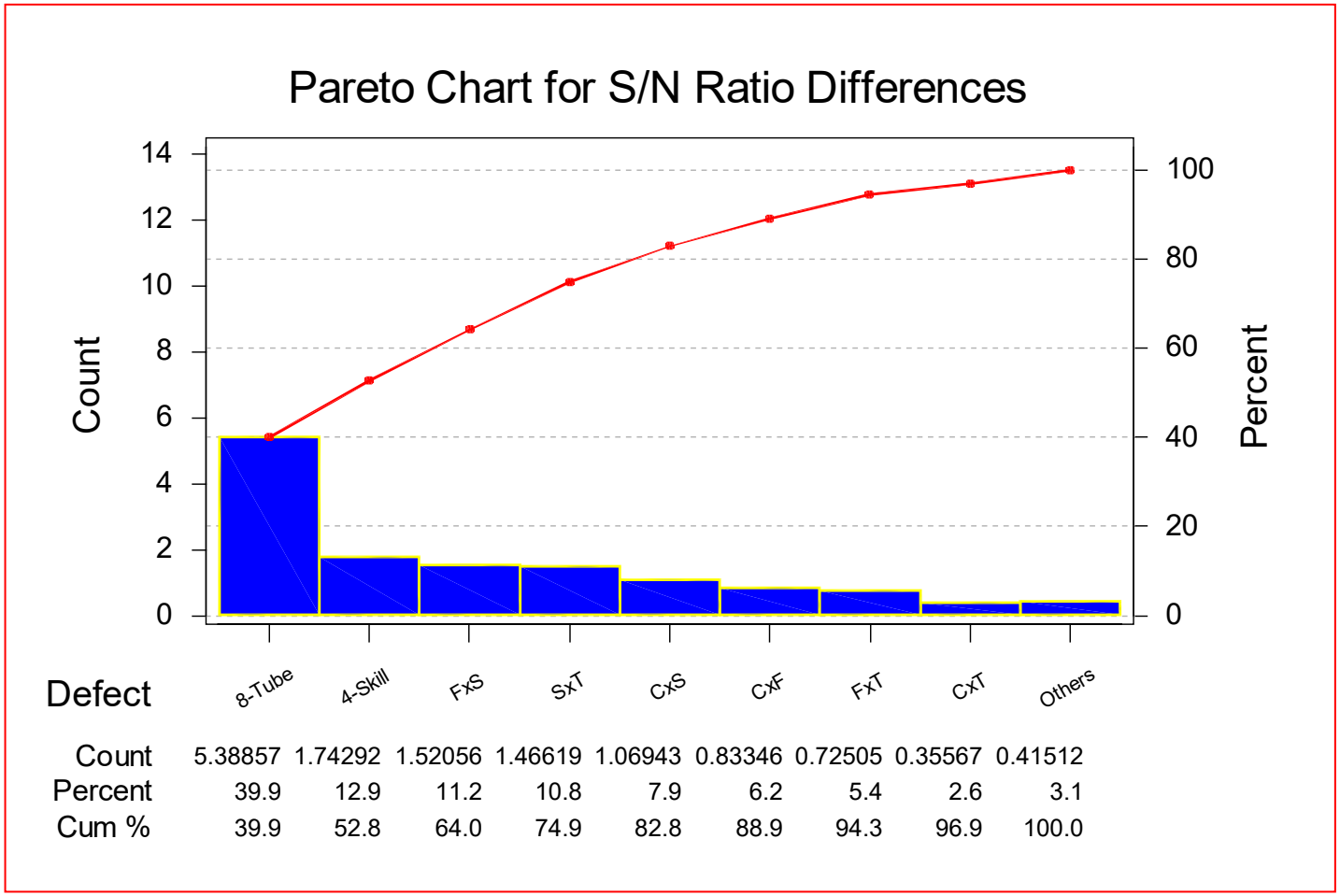


To analyze the data, we calculate the square of each response in order to calculate the Signal-to-Noise Ratio for each factor level. The calculations were performed in Excel and appear below:

	Column (Factor/Interaction)											
Trial	**1-Cleaned**	**2-Flux**	**CxF**	**4-Skill**	**CxS**	**FxS**	**8-Tube**	**CxT**	**FxT**	**SxT**	**Results**	**Results Squared**
1	1	1	1	1	1	1	1	1	1	1	0.215	0.046225
2	1	1	1	1	1	1	2	2	2	2	0.269	0.072361
3	1	1	1	2	2	2	1	1	1	2	0.184	0.033856
4	1	1	1	2	2	2	2	2	2	1	0.258	0.066564
5	1	2	2	1	1	2	1	1	2	1	0.146	0.021316
6	1	2	2	1	1	2	2	2	1	2	0.344	0.118336
7	1	2	2	2	2	1	1	1	2	2	0.107	0.011449
8	1	2	2	2	2	1	2	2	1	1	0.33	0.1089
9	2	1	2	1	2	1	1	2	1	1	0.2	0.04
10	2	1	2	1	2	1	2	1	2	2	0.43	0.1849
11	2	1	2	2	1	2	1	2	1	2	0.123	0.015129
12	2	1	2	2	1	2	2	1	2	1	0.209	0.043681
13	2	2	1	1	2	2	1	2	2	1	0.121	0.014641
14	2	2	1	1	2	2	2	1	1	2	0.311	0.096721
15	2	2	1	2	1	1	1	2	2	2	0.2	0.04
16	2	2	1	2	1	1	2	1	1	1	0.28	0.0784

The Supplementary Table is calculated for each factor/interaction. For each level, the first row shows the sigma-squared component of the Signal-to-Noise Ratio, the second row is the Signal-to-Noise ratio. The larger S/N ratio is bolded (the best level of the factor/interaction). The final row shows the absolute value of the difference between the two Ratios (see the Pareto of these values on the next page. Note that these results are consistent with those found through the ANOVA. The **tube cut** is the only significant factor.

Level 1	**σ^2 - 1**	0.05988	0.063	0.1	0.074	0.1	0.1	0.0278	0.06	0.07	0.051
	S/N - 1	**12.2275**	12.02	**13**	11.29	**13**	**13**	**15.555**	11.9	11.7	**12.92**
Level 2	**σ^2 – 2**	0.06418	0.061	0.1	0.05	0.1	0.1	0.0962	0.06	0.06	0.072
	S/N - 2	11.9257	**12.13**	12	**13.03**	12	11	10.167	**12.3**	**12.5**	11.45
Difference		0.30175	0.113	0.8	1.743	1.1	1.5	5.3886	0.36	0.73	1.466


Pareto Chart for S/N Ratio Differences
Count
Percent
14
12
10
8
6
4
2
0
100
80
60
40
20
0
Defect
8-Tube
4-Skill
FxS
SxT
CxS
CxF
FxT
CxT
Others
Count 5.38857 1.74292 1.52056 1.46619 1.06943 0.83346 0.72505 0.35567 0.41512
Percent 39.9 12.9 11.2 10.8 7.9 6.2 5.4 2.6 3.1
Cum % 39.9 52.8 64.0 74.9 82.8 88.9 94.3 96.9 100.0

Section 15 – Reliability Management

Objective: To ensure understanding of RAM concepts and definitions.

Instructions:
1. Here are some reliability situations and questions to answer. Your instructor will assign one or more of these to your table.
2. Use the definitions and concepts presented in this unit to help answer these questions.
3. Record your ideas on the flipcharts provided and be prepared to report your findings to the group:

Time: 30 minutes

2. In response to a regulator's concern about the reliability of safety system equipment, a power plant developed a new series of preventive maintenance procedures. Comment on this plan.

   Answer: The countermeasure (preventive maintenance procedures) may or may not have any relationship to the root causes of the equipment failures.

4. A nuclear plant's three emergency containment coolers are *required* to operate for at least a year (8760 hours) after an accident. Due to failures of the fan bearings (about every 1000 hours), the mechanical maintenance department replaces the bearings every refueling outage (about every 18 months). Each fan runs about 600 hours between outages. No failures have been experienced for the last two refueling cycles. Should plant management be happy with this performance?

   Answer: This is a mission-time issue. The fans will run (on average) about 1000 hours, however they are required to run for at least 8760 hours. Even though no failures have been observed, the fans have not demonstrated their required reliability.

6. You have just loaded the new "Doors2000" graphics environment on your MBI personal computer. Your applications now run much slower than before; it takes three minutes to print a document that used to take 30 seconds. Has the computer failed?

   Answer: It depends on your success requirement. If your requirement is, for example, to print the document in less than one minute (e.g. printing a hotel bill), then a failure has occurred.

Objective: To understand and build skills in basic reliability calculations methods.

Instructions: 1. Here are some more reliability calculation exercises. Try these to test your understanding of the RAM calculation methods.

1. An electronic device used in an air conditioning control circuit has a constant failure rate λ = 0.00125 failures per hour. What is its MTBF? What is its reliability for a 10- hour mission? For a 100-hour mission?

Answer: MTBF = 1/0.00125, $R(10hr) = \exp(-0.00125/hr \times 10hr) = 0.987578,\ R(100hr) = \exp(-0.00125/hr \times 100hr) = 0.882497$

2. A compressor requires an average of 4 hours to repair. What MTBF is needed if steady-state availability is to be at least 0.99?

Answer: $A = \dfrac{MTBF}{MTBF + MDT} \Rightarrow MTBF = \dfrac{A \times MDT}{1 - A} = \dfrac{0.99 \times 4}{1 - 0.99} = 396hr.$

3. A sample of five constant hazard rate devices is tested without replacement until the fourth failure, at which time the test is terminated. Times to failure are t_1 = 800 hours, t_2 = 1800 hours, t_3 = 2125 hours, and t_4 = 2812 hours. The time on the fifth test unit is t_5 = 2812 hours. Make a point estimate of reliability for a 100-hour mission for this device. Find a 95% upper confidence limit on the failure rate and a 95% lower confidence limit on reliability for this same mission.

Answer:

$$MTBF = \frac{800 + 1800 + 2125 + 2812 + 2812}{4} = 2587.25\text{hr.} \Rightarrow \text{R(100hr)} = \exp(\text{-100/2587.25}) = 0.962086$$

$$\frac{2\sum t_i}{\chi^2(1-\alpha, 2r+2)} \geq MTBF \Rightarrow \frac{10349hr.}{18.31} = 565.2hr. = MTBF_{Lower} \Rightarrow \lambda_{Upper} = 1/565.2hr = 0.001769/\text{hr}$$

$$\text{R}_{\text{Lower}}(100\text{hr}) = \exp(\text{-0.001769/hr} \times 100) = 0.837842$$

4. A manufacturer claims that his device has a mean time to failure of 15000 hours. His maintenance manual requires semi-annual lubrication and monthly filter changes. Comments?

Answer: A minor discrepancy – components of the device fail at intervals less than the published MTBF of 15000 hours.

5. For the following descriptions, determine which probability distribution would most likely fit the situation:

- An Instrumentation and Controls supervisor is testing integrated circuits to see if they meet company standards for use. Each time they are delivered, he selects 10 and checks to see how many fail an inspection test. Answer: Binomial
- The number of interruptions to the transmission system is measured each month. Answer: Poisson
- Samples of 5 transformers are gathered each month and the average core loss is measured for each sample. Answer: Normal
- A mechanical coupling is observed to behave with a constant hazard rate. Answer: Exponential, Weibull
- Tube leaks in a Component Cooling Water heat exchanger were found to be the result of a fatigue failure mechanism. Answer: Weibull

6. A turbine-driven compressor has exhibited a starting reliability of 0.97 for the last two years of operation. Management asks you to predict how many starting failures are expected this year if the compressor is to be started 60 times.

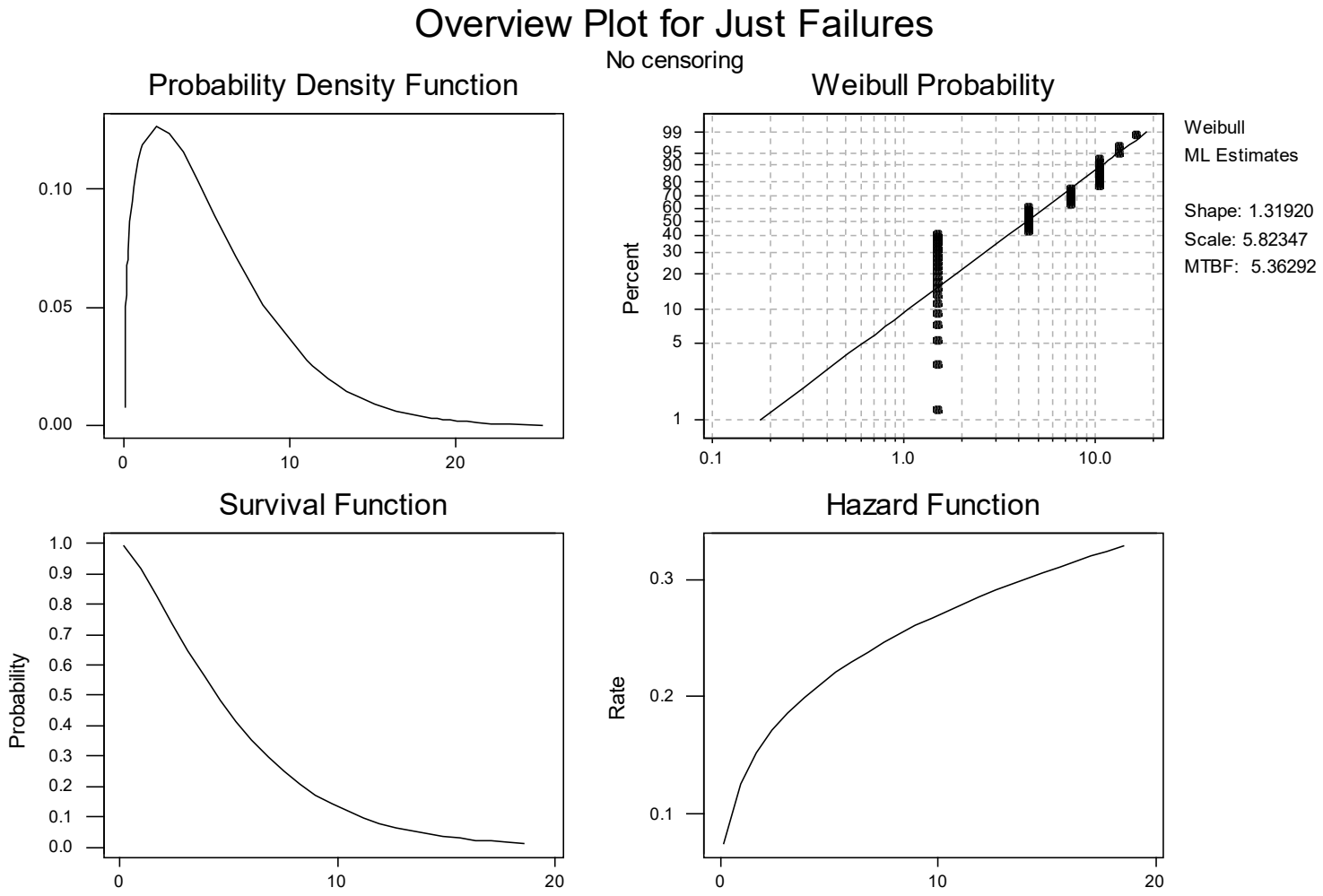
Answer: Expected Number of Failures = 60 (1 – 0.97) = 1.8 or 2 failures.

7. A company block purchased replacement Volt-Ohm-Meters for electrical troubleshooting. The battery packs on these devices have been failing and the electrical maintenance department has been complaining. Here is the failure data collected by the electrical maintenance supervisor:

Time to Failure (Months)	Number of Failures
0 - 3	21
3 - 6	10
6 - 9	7
9 - 12	9
12 - 15	2
15 - 18	1
Total	50

Develop a histogram of the failures. What distribution seems to best fit this data? Calculate the associated distribution parameters. What is this distribution telling you about the nature of the failures?

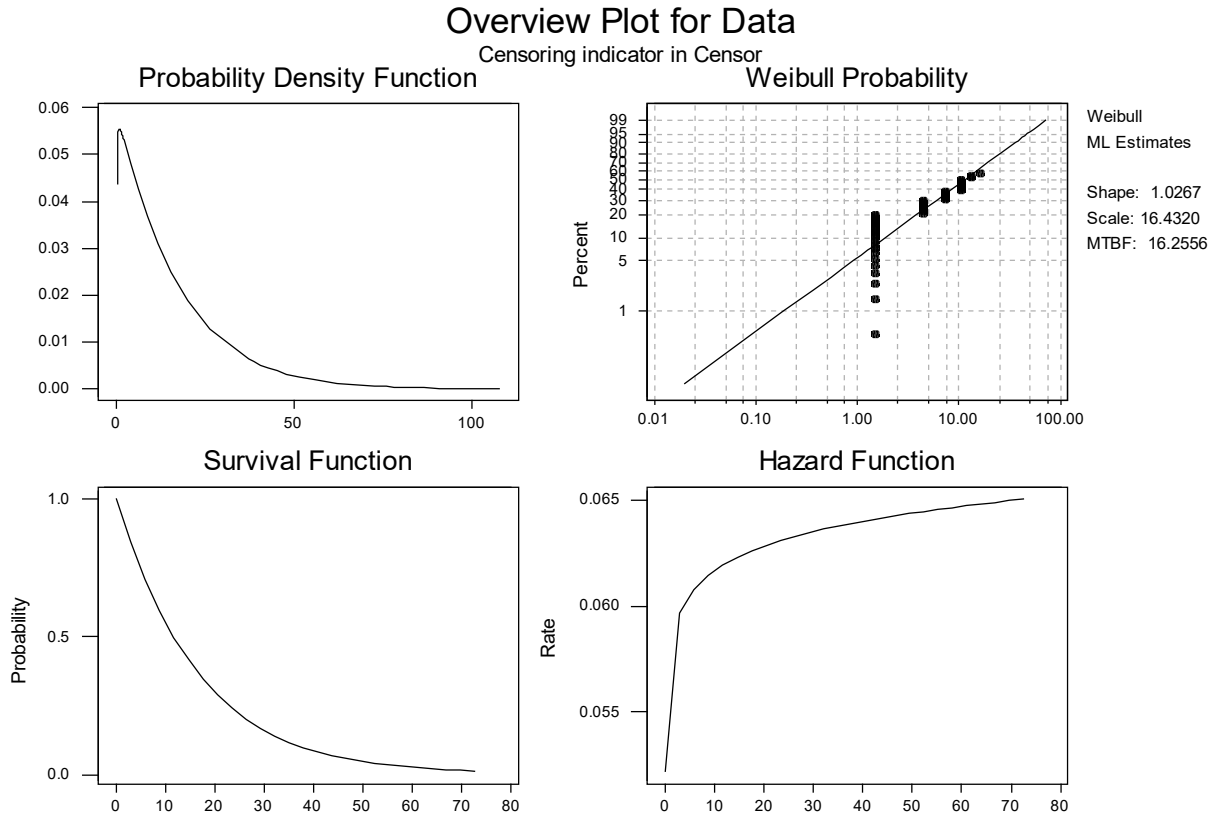
Answer: The histogram (left to the student) can be modeled with an exponential distribution – shows failures result from a constant hazard process (e.g. failures occur at random – perhaps a random stress that "shocks" the VOM battery). A Weibull Analysis appears below. The shape factor of 1.3 is "close" to 1 (especially since the failure times were obtained from the midpoints of the above intervals.



The problem you addressed above only included *times to failure.* How would your conclusions vary if the following information were included regarding VOM's that had *not* failed?

Operating Time (Months)	Number of Meters
0 - 3	2
3 - 6	4
6 - 9	16
9 - 12	12
12 - 15	9
15 - 18	6
18- 22	4

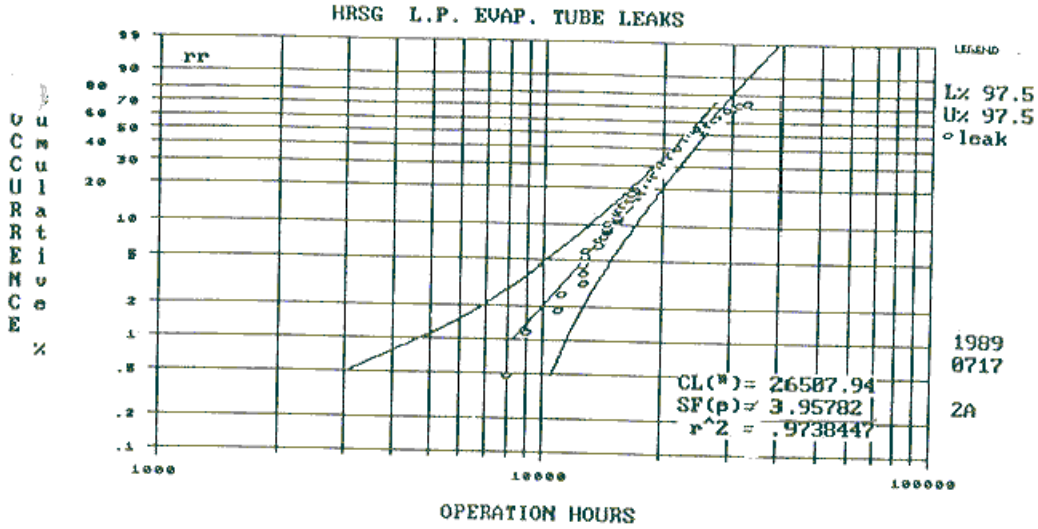
Answer: These are suspensions that should be included in the failure model. A Weibull analysis of the data appears below. The failure/suspension times were obtained from the midpoint of each of the time blocks above. Note that the Shape factor of 1.03 confirms the random nature of the failures. Note that the suspensions increase our estimate of the Scale factor (and MTBF).



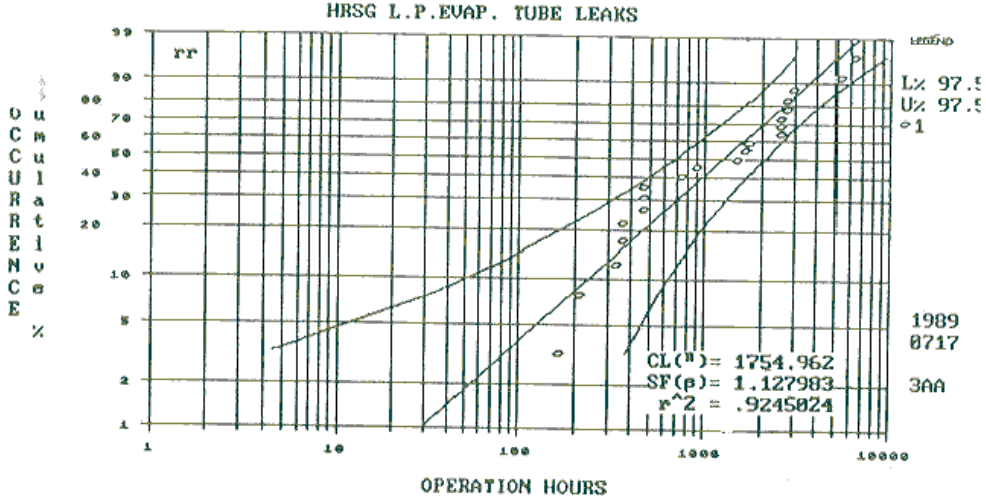
11. The following Weibull Plots were obtained from an analysis of tube leaks occurring in a Low Pressure Evaporator of a combined cycle power plant. Interpret these plots. What can you learn about the failures from these charts?

Answer: There are different failure mechanisms at work in the Evaporator. There is a wearout mode (shape factor (SF) =3.96) at work on the outer rows (top Weibull), with a scale factor (CL) of about 26,600 hours. Tube failure across the evaporator experience a constant hazard rate (shape factor (SF) = 1.13), with a scale factor of about 1750 hours (much lower life). These Weibulls were used to stratify the tube failures and provide clues to the failure causes occurring in the Evaporator.

Tube Failures Occurring on 5 Outer Rows of the Evaporator (at Bend Area):



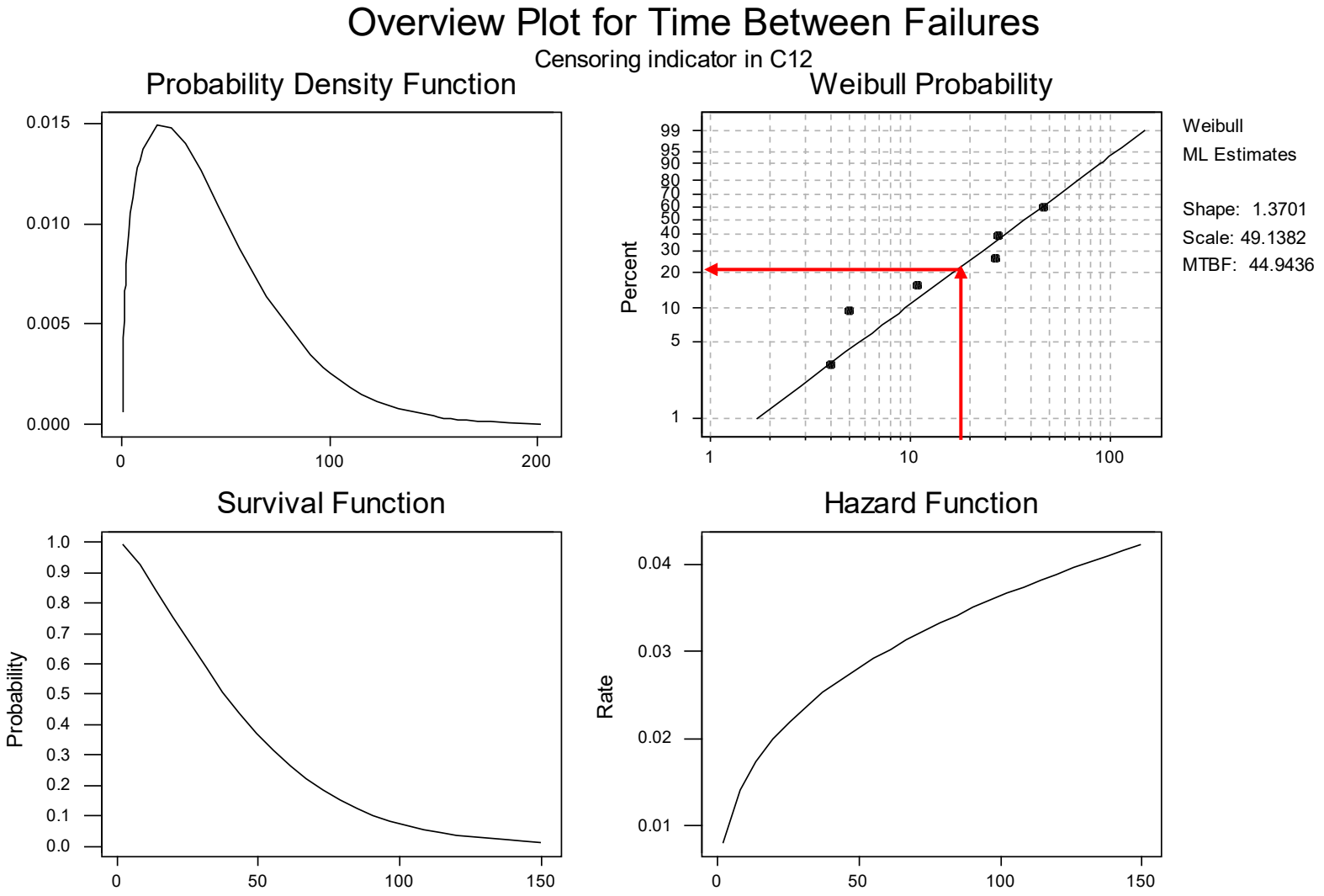
Tube Failures Occurring Across the Evaporator (at Bend Area):



10. The following valve repair data was collected. Note that in some cases, upon overhaul, the valve was found to be in a failed state, in others, the valve was still operable. Perform a Weibull Analysis on this data. What can you learn from this analysis? If the valve's mission time is 18 months, what are the chances of it meeting that mission without failure (i.e. what is the reliability of the valve)?

Time Between Repairs (Months)	Valve Condition
47	OK
23	Failed
10	Failed
11	OK
28	OK
32	Failed
4	OK
13	Failed
17	Failed
47	Failed
25	Failed
5	OK
20	Failed
18	Failed
27	OK
36	Failed

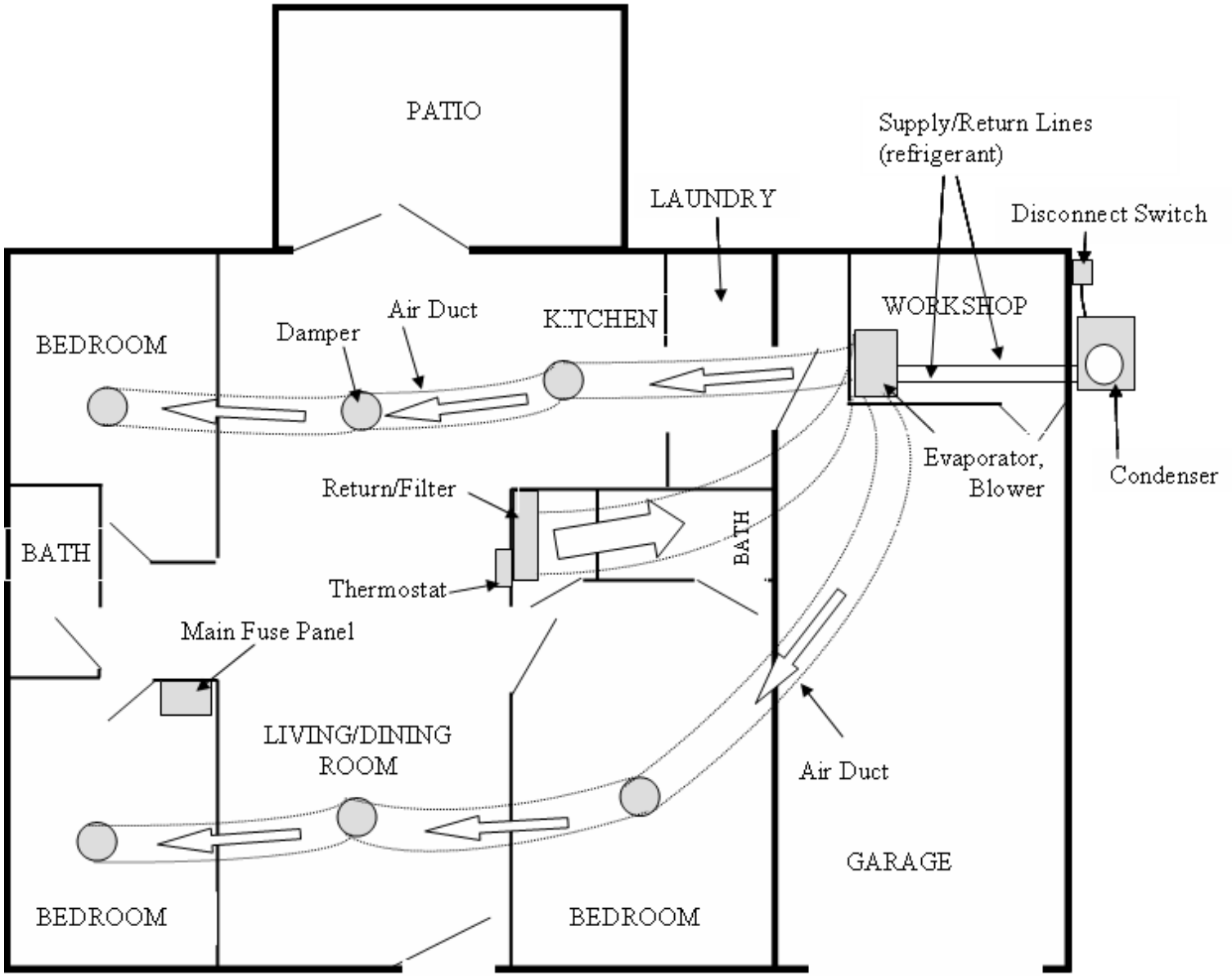
Answer: The Weibull Analysis appears below. Minitab was employed (Stat>Reliability/Survival>Distribution Overview Plot) for the analysis. The shape factor is 1.37, borderline between a random and wearout failure pattern. The reliability of the valve is about 80%, found by entering the Weibull plot at 18 months on the horizontal axis and finding the failure probability on the vertical axis.


Overview Plot for Time Between Failures
Censoring indicator in C12
Probability Density Function
0.015
0.010
0.005
0.000
0
100
200
Weibull Probability
Percent
99
95
90
80
70
60
50
40
30
20
10
5
1
1
10
100
Weibull
ML Estimates
Shape: 1.3701
Scale: 49.1382
MTBF: 44.9436
Survival Function
Probability
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
50
100
150
Hazard Function
Rate
0.04
0.03
0.02
0.01
0
50
100
150

Appendix E - Reliability Exercise Material

On the following pages, you will find the technical data you'll need to perform several of the exercises in this manual.

AIR CONDITIONING GENERAL ARRANGEMENT



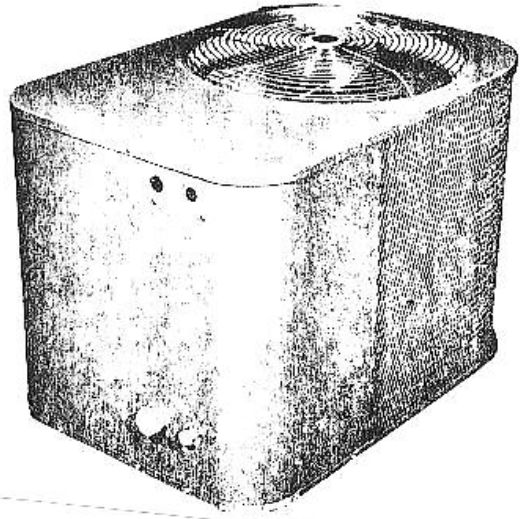
CONDENSOR – GENERAL DESCRIPTION

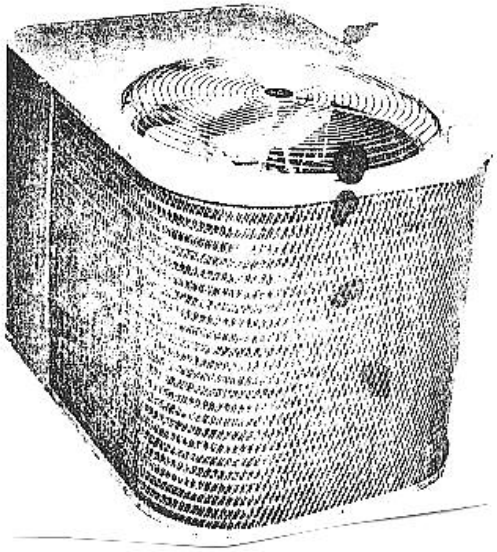
DESCRIPTION

Model SP036C Condensing Units are precharged air cooled condensing units complete with compressor, condenser coil, direct drive propeller fan and motor assembly and electrical control components. They are designed for installation with the matching evaporator coils indicated in the Nameplate Data Table. Models SP036CBW/-CBS may be used with other units of comparable size.

These units are designed with a single row U-shaped condenser coil. Airflow is draw-thru on 3 sides with vertical discharge. The air outlet in the top panel is covered by a grille to protect and prevent physical contact with the rotating fan blade.

The compressor, control components and internal tubing are located in the compressor compartment at the rear end of the unit. The compressor compartment is separated from the condenser coil and fan compartment by a sheet metal panel.





Refrigerant couplings and electrical entrance openings are located at the rear end of the unit.

CONDENSOR – SPECIFICATIONS

PHYSICAL DATA TABLE

MODEL NO.	SP036CAW	SP036CAE	SP036CAS	SP036CBW	SP036CBS
WESTINGHOUSE COMPRESSOR	CD036BAY	CD036BAS	CD036BAS	CD036BAY	CD036BAS
	3500 RPM (60 Cy.) 2920 RPM (50 Cy.) Internal High Pressure Relief — 410 PSI Differential Internal Line Break Overload Device Oil Charge — 3 Pts. (Suniso 3 Gs.)				
COMPRESSOR CAPACITOR	30-40 MFD 440 V.	—	—	30-40 MFD 440 V.	—
START ASSIST DEVICE	For Single Phase Units Model SLZSCXB Style 3962A94G01 (Sprague #305C2)				
FAN MOTOR	1/4 HP 48 Frame 1075 RPM (60 Cy.) Direct Drive				
FAN MOTOR CAPACITOR	4 MFD/370 V.				
FAN BLADE	18 In. Dia., 3 Blades, 34° Pitch (Vertical Outlet)				
STANDARD AIR	2700 CFM				
COIL FINNED TUBE TYPE	18 Fins/In., 3/8 In. Aluminum Tubing, 1 Row Deep, 24 Tubes High 24 In. Finned Height, 60 In. Finned Length Coil Face Area 10.00 Sq. Ft.				
REFRIGERANT CHARGE (R-22)	(60 Cy.).— 1 Lb. 11 Oz.		(50 Cy.) 1 Lb. 8 Oz.		
OVERALL DIMENSIONS	26 In. High, 24 In. Wide, 34 In. Long				
WEIGHT	168 Lb.				
REFRIGERANT COUPLINGS	5/8 - 18 Liquid (Aeroquip — 5750-S2-6-6) 1-7/16 - 16 Suction (Aeroquip — 5750-S2-12-12)				
REFRIGERANT LINES	Liquid Line — 3/8 In. O.D.		Suction Line — 3/4 In. O.D.		
EXTERNAL HIGH PRESSURE SWITCH	Not Used			Open 405 PSIG Manual Reset	

ACCESSORIES TABLE

UNIT MODEL	SP036CAW	SP036CAE	SP036CAS	SP036CBW	SP036CBS
START COMPONENT KIT	Model SPZSC3A (145 MFD/320 V.)	—	—	Model SPZSC3A (145 MFD/320 V.)	—
LIQUID COUPLING STUB KIT	Model SLZCL2A (3/8 O.D. Line) Aeroquip 5701-S5-6-6 Style No. 3953A72G02				
SUCTION COUPLING STUB KIT	Model SLZCS2A (3/4 O.D. Line) Aeroquip 5701-S5-12-12 Style No. 3953A72G06				
THERMOSTATS	MODEL NO. / STYLE NO. / DESCRIPTION ZT0H1CMSB1 / 4388843H01 / 1 Cool (T87F1933/Q539C1004) ZT1H1CMSB1 / 4388842H01 / 1 Heat, 1 Cool (T87F1933/Q539C1006) ZT1H1CMSD1 / 350A092H22 / 1 Heat, 1 Cool (T834C) ZT1H1CMSC1 / 350A092H35 / 1 Heat, 1 Cool (T199BL121) ZT1H2CCSA1 / 350A792H04 / 1 Heat, 2 Cool (T872B1039/Q672E1001)				
CRANKCASE HEATER	Model ZH00C07 (35 Watts 230 Volts) Style No. 353A667G06				

COOLING CIRCUIT SCHEMATIC AND VALVE DETAILS

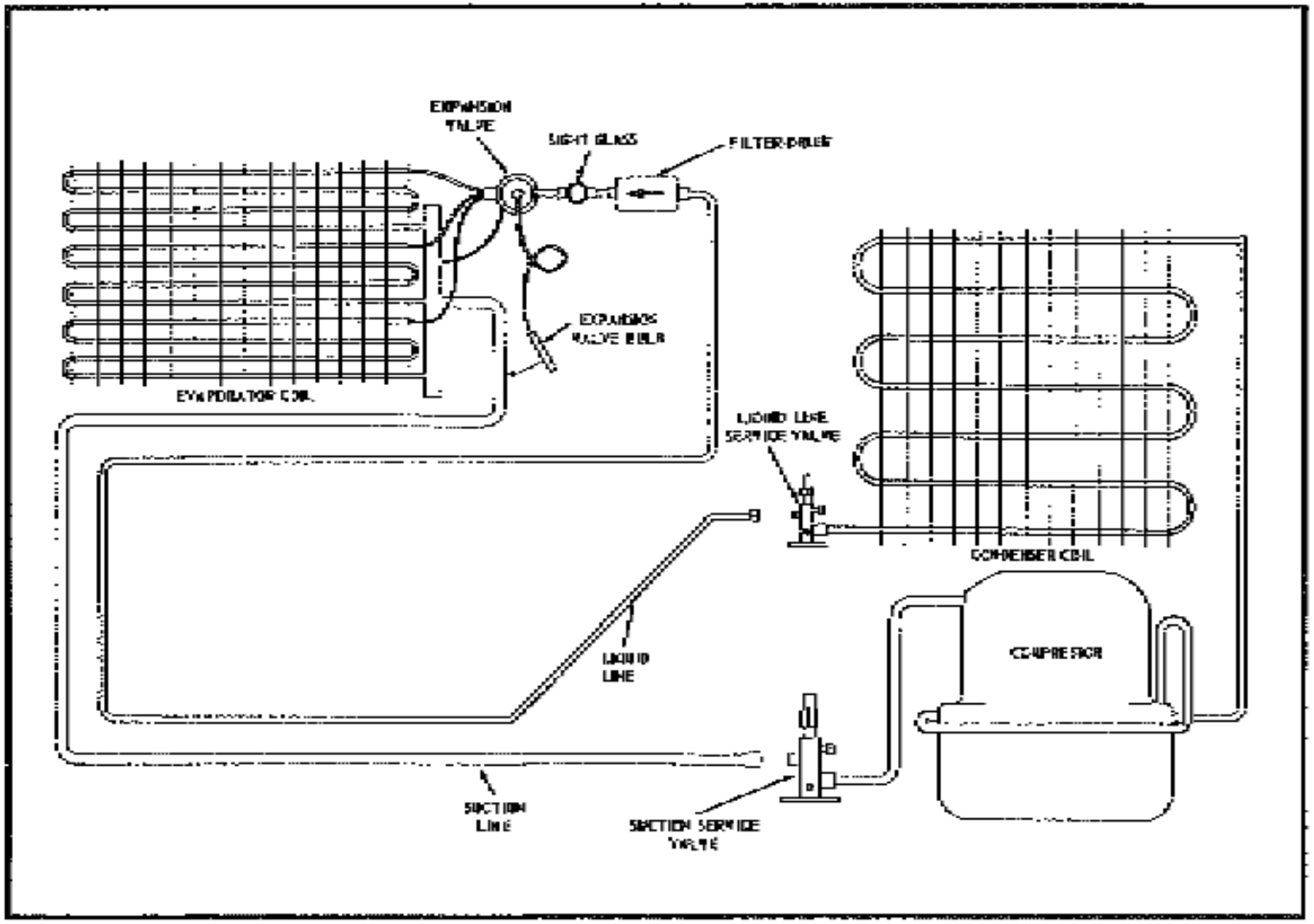


Fig. 1 – Schematic Drawing of Cooling Circuit

Three-way Service Valve positions are as follows: (See Figs. 8 and 9)

Front-Seat Position (Turned completely clockwise)	Isolates Condensing Unit	Gauge Port to Evaporator Section.....OPEN Gauge Port to Condensing Section.....CLOSED Condensing Section to Evaporator Section...CLOSED
Back-Seat Position (Turned completely counterclockwise)	Isolates Gauge Port	Condensing Section to Evaporator Section.....OPEN Gauge Port to Condensing Section.....CLOSED Gauge Port to Evaporator Section.....CLOSED
Mid-Seat Position (Turned two full turns off Front or Back-Seat)	Used to Charge	Condensing Section to Evaporator Section to Gauge Port .....ALL OPEN

Use only refrigeration wrenches on refrigerant valves.

It is easier to avoid contamination and moisture than to get rid of it. Avoid opening any part of circuit to atmosphere unnecessarily, and NEVER open tubing to atmosphere when temperature of the tubing is below dew point.

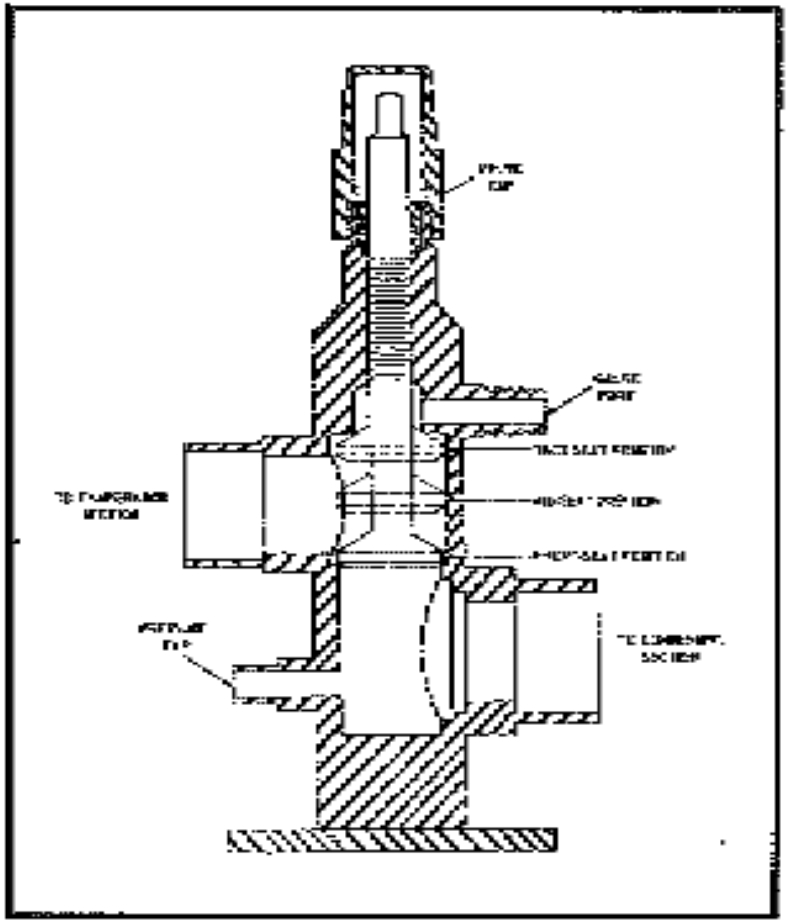


Fig. 2 – Service Valve Cross-Section

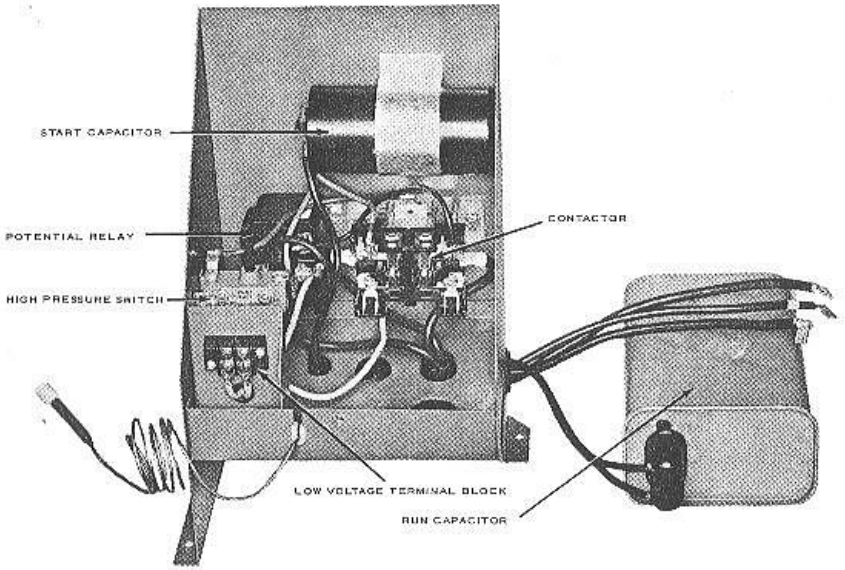
Use only NON-CORROSIVE type Flux when soldering. NEVER USE SOLDER CONTAINING LEAD

ELECTRICAL DATA & CONTROLS ARRANGEMENT

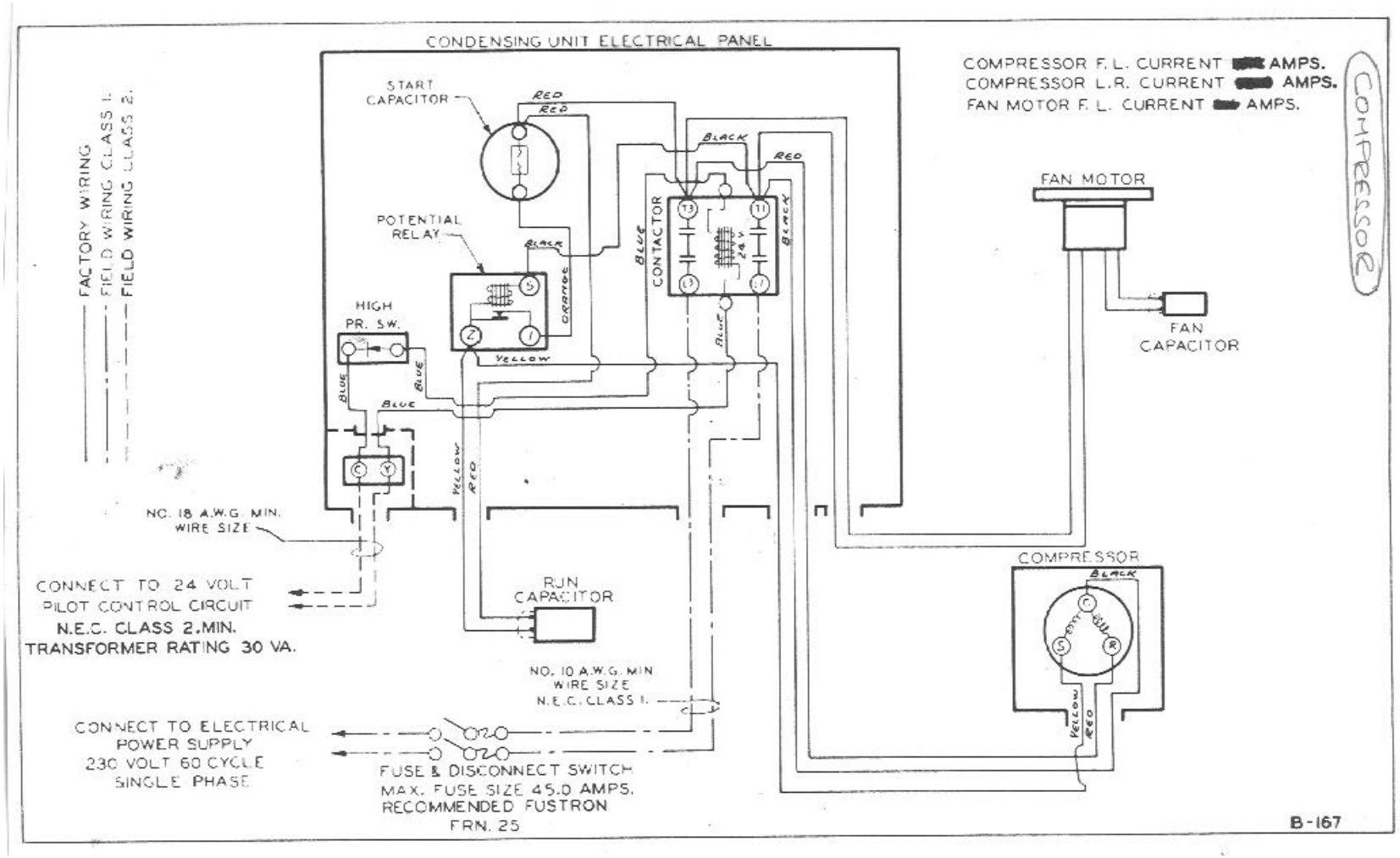
ELECTRICAL DATA

CURRENT – WIRE – FUSE – DISCONNECT SWITCH – CONDUIT

UNIT MODEL	SP036CAW	SP036CAE	SP036CAS	SP036CBW	SP036CBS
ELECTRICAL CHARACTERISTICS	1/60/208-230	3/50/220	3/60/208-230	1/60/208-230	3/60/208-230
COMPRESSOR FLA LRA	24.6/23.3 90/99	14.6 68	15.6 59/65	24.6/23.3 90/99	15.6 59/65
FAN MOTOR (1/4 HP) FLA	1.6	1.6	1.6	1.6	1.6
MINIMUM WIRE SIZE, AWG (COPPER)	#8 (2)	#10 (3)	#10 (3)	#8 (2)	#10 (3)
MAXIMUM FUSE - AMPS	50	35	35	50	35
MINIMUM AMPACITY	40	30	30	40	30
DISCONNECT SWITCH, HP	5	7.5	7.5	5	7.5
CONDUIT SIZE, IN. (FOR COPPER WIRE)	1/2	1/2	1/2	1/2	1/2
OPERATING VOLTAGES MIN./MAX.	197/253	197/242	197/253	197/253	197/253

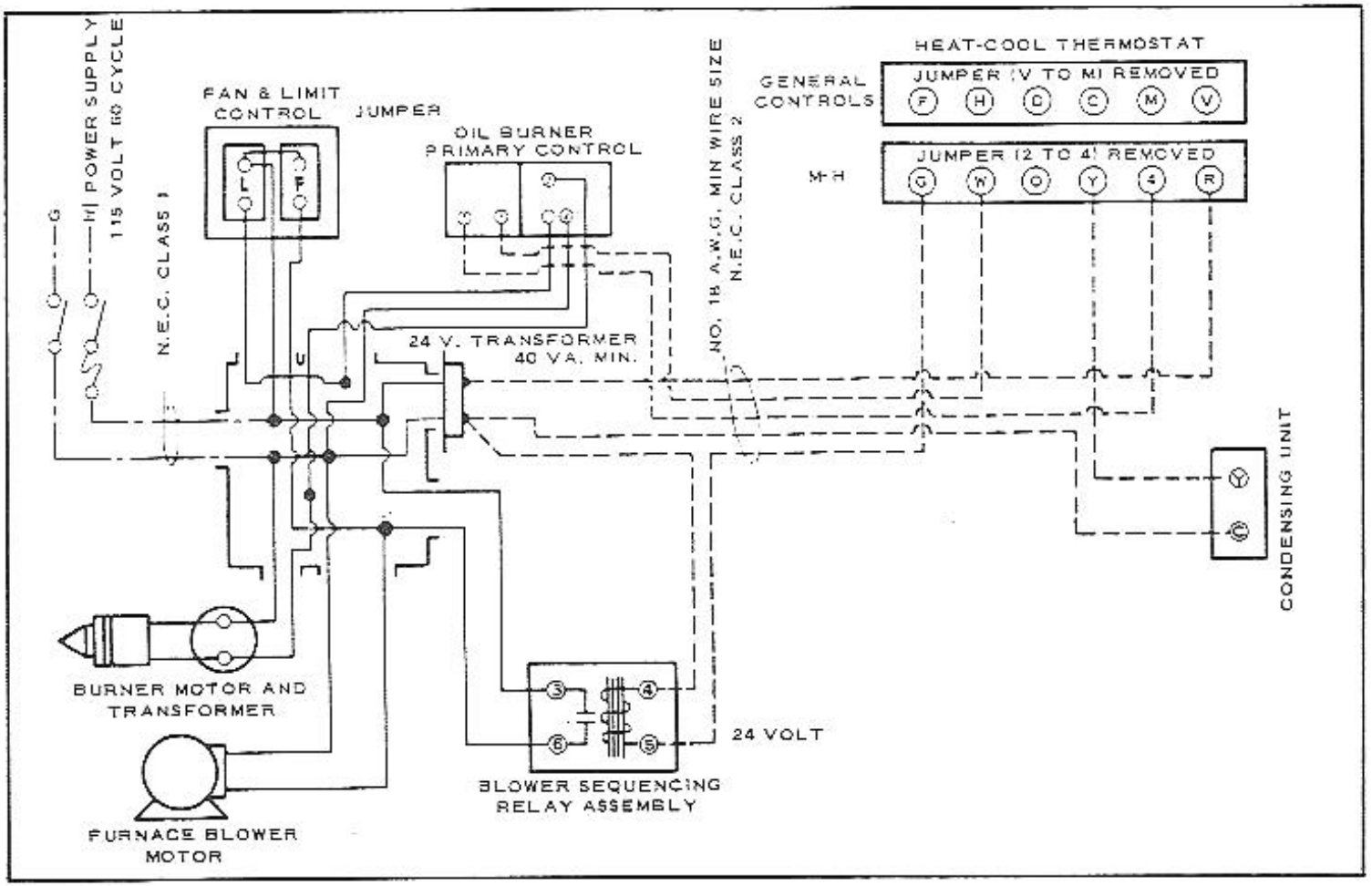


COMPRESSOR & CONDENSING UNIT – INTERNAL WIRING SCHEMATIC



A/C CONTROL CIRCUIT

APPLICATION WIRING DIAGRAMS



Typical Application With Oil Fired Furnace

INSTALLATION

GENERAL REQUIREMENTS

The SP036C Condensing Units are designed for outdoor installation. The unit should be mounted on a concrete pad or otherwise raised off the ground to provide adequate drainage of water thru drain holes in base pan. **Do not install unit where it will be subjected to heavy downpour of water from roof drainage or similar locations.**

RECOMMENDED CLEARANCES

Condenser air inlet to hard walls . . . . . 10 inches
Condenser air inlet to bushes . . . . . 16 inches
Rear end (compressor compartment) . . 36 inches
Top of unit (air discharge) . . . . . No Obstruction
See Figure 1.

NOTE: Although the SP condensing unit may be installed to within 10 inches of walls, it is recommended that it **not** be installed with walls this close on 3 sides due to air recirculation thru the condenser coil.

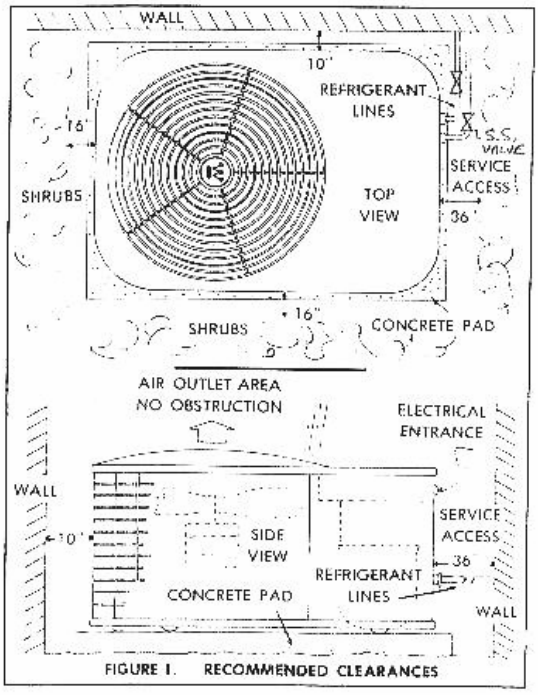


FIGURE 1. RECOMMENDED CLEARANCES

ROOF TOP INSTALLATION

It is recommended that the unit be raised 8 inches above the roof surface to reduce the hot surface air being drawn thru the condenser coil.

It is also recommended that resilient pads be placed under the unit to reduce the transmission of vibration and noise into the building structure.

REFRIGERANT LINES

Refrigerant lines of the sizes indicated in the Physical Data Table should be used.

Interconnecting lines should not exceed 50 feet.

Vertical runs (SP unit BELOW indoor coil) —

When vertical runs exceed 30 ft., liquid and suction line should be strapped or soldered together and insulated to prevent loss of sub-cooling. (Maximum practical lift is 50 ft.)

Vertical runs (SP unit ABOVE indoor coil) —

When vertical run exceeds 25 ft., an oil trap (minimum size 3 in.) should be located at the mid-point of the suction line to assist in proper oil return.

Refrigerant lines subjected to high temperature conditions such as across roof tops or thru attics must be insulated to prevent liquid flashing in liquid line. Liquid and suction lines should be strapped and wrapped together to help prevent liquid flashing.

Suction lines run thru building interiors must be well insulated to prevent condensate moisture drip and possible water damage to building interiors.

FIELD FABRICATED REFRIGERANT LINES

Observe the following:

1. Braze with inert gas in lines to prevent scale formation.
2. Install liquid line filter-drier in liquid line (Sporlan C-080 series or equal).
3. Evacuate lines and charge with amount of refrigerant indicated in refrigerant charging table.
4. Add an additional 8 ounces of refrigerant to system for the filter-drier.
5. Vertical runs exceeding 25 ft. with SP unit above evaporator coil must have a 3 inch oil trap at mid-point in suction line to aid in oil return. Traps may be fabricated using copper tubing street elbows.

→ SEE SPECIFIC INSTALLATION INST'S

CONNECTION OF REFRIGERANT LINES

Refrigerant couplings on the indoor and outdoor unit and the refrigerant lines are provided with protective caps. These should be left in place until the refrigerant lines are ready to be connected.

When ready to connect the lines, remove protective caps. Clean the couplings of any foreign matter. Using refrigerant oil, lubricate the threads, diaphragms and rubber seals.

Thread couplings together by hand to assure proper mating of threads. With wrenches on the body hex and union nut, tighten union nut until a definite resistance is felt. Tighten an additional 1/4 turn to seat fitting. If torque wrenches are used, the following torque values are applicable:

SP036	Suction Coupling	50-65 Ft. Lbs.
	Liquid Coupling	10-12 Ft. Lbs.

All connections should be leak tested.

REFRIGERANT

SP units used with matching indoor coils or air handlers and matching precharged refrigerant lines do not require additional field charging of refrigerant at start-up.

Any field charging of equipment should be in accordance with the unit nameplate of the System Refrigerant Charging Table below. If the charge is not weighed in, the system should be charged to the Temperature-pressure Performance Chart.

CAUTION: Charging the system until the compressor draws nameplate full load amps may result in serious system overcharge. This should not be done.

ELECTRICAL

Electrical power as indicated on the unit nameplate will be required. It is recommended that electrical wiring be run in metal conduit from the service entrance to the unit, which will serve as an electrical ground. If conduit is not used, a separate ground wire must be run. A ground connection screw is provided in the control box for grounding the unit. This unit complies with the N.E.C. requirements for electrical grounding of electrical components and panels.

Power wiring should be run from a fused disconnect switch (weatherproof if located outside) to the unit control box and connected to the terminals of the compressor contactor inside of the control box. A 24-volt control voltage power supply must be provided from a separate source. Connect thermostat control wiring to the control connection "pig tail" leads located in the low voltage compartment in the control box. Refer to the unit electrical data and wiring diagrams for detailed wiring instructions and for proper sizes of wire, fuses and disconnect switch. Comply with N.E.C. and/or local electrical code requirements.

SYSTEM REFRIGERANT CHARGING TABLE (R-22)

CONDENSING UNIT	SP UNIT CHARGE	UPFLOW COIL	HORIZONTAL COIL	COUNTERFLOW COIL	AIR HANDLER	DUAL CIRCUIT COIL
SP036CAW	1 lb. 11 oz.	14 oz.	9 oz.	13 oz.	16 oz.	14 oz. (each system)
SP036CAE	1 lb. 8 oz.					
SP036CAS	1 lb. 11 oz.					
SP036CBW	1 lb. 11 oz.					
SP036CBS	1 lb. 11 oz.					

Total System Charge With Allowance For Liquid Line
SP Unit — Indoor Coil + Allowance For Liquid Line
With 5/16" OD Liquid Line - 4 oz. for each 10 ft. of line
With 3/8" OD Liquid Line - 6 oz. for each 10 ft. of line

NOTE: Allow additional 8 oz. of refrigerant for filter-drier in liquid line (Sporlan C-080 series or equal).

TEMPERATURE PRESSURE CHART
Indoor Air 75 DB/62 WB
1200 Nom. CFM

AIR TEMP. ENT. OD COIL °F	LIQUID PRESS. LV. COND. PSIG	SUCT. PRESS. PSIG	LIQ. SUB-COOLING °F	SUCTION SUPERHEAT °F
75	211	59	19	22
80	228	61	18	18
85	245	63	16	14
90	262	64	15	10
95	279	66	14	5
100	296	68	13	4
105	315	69	11	4
110	332	71	10	3
115	351	72	9	3

E-306
October, 1962

SPECIFIC INSTALLATION INSTRUCTIONS
AIR COOLED CONDENSING UNITS

GENERAL

These units are designed for use as remote air-cooled condensing units in conjunction with blower-evaporator units or evaporator coils of similar capacity using Refrigerant-22 only.

UNPACKING

As soon as received, unit should be inspected for damage done in transit. If damage is found it should be noted on carrier's freight bill. Request for inspection by carrier's agent should be made in writing at once.

Check electrical characteristics (stencilled on carton) to make sure they are correct for the electrical supply.

CODES AND ORDINANCES

Cooling systems must be installed in conformance with local refrigeration codes.

Units must be installed in accordance with Regulations of the National Board of Fire Underwriters. Authorities having jurisdiction should be consulted before installation is made.

Air duct systems should be installed in accordance with Standards of the National Board of Fire Underwriters for the Installation of Air Conditioning, Warm Air Heating and Ventilating Systems, Pamphlets 90A and 90B.

ELECTRICAL WIRING

All wiring must conform to the provisions of National Electrical Code and/or local ordinances.

Wiring diagrams furnished show internal wiring and application. Diagram covering internal wiring of condensing unit is also furnished on inside of control box cover.

Control panel may be mounted in any convenient wall location. Wiring from condensing unit to panel may be Class 2.

Condensing unit casing must be connected to a suitable ground.

SAFETY PRECAUTIONS IN HANDLING REFRIGERANT

- Refrigerant drums should be stored upright in a reasonably cool place.
- Temperature above 125° should be avoided at all times.
- Never overfill Refrigerant Drums.
- Refrigerant, gas or liquid can seriously damage the eye by freezing moisture in the eyeball. Goggles or large-lens glasses are recommended when handling Refrigerant.
- When subjected to open flame, Refrigerant will produce Phosgene gas. Avoid danger by always bleeding or purging to outside air.
- When, during charging, it becomes necessary to warm the Refrigerant cylinder, partially immerse drum in warm water, but NEVER APPLY FLAME OR LIVE STEAM TO DRUM. Avoid immersing drum except when valve is open, and never immerse drum valve or fittings.

IMPORTANT NOTICE

Installer must be familiar with valve system of this circuit before attempting leak-test or charging. Trial and error can be very expensive. Service valves of these units are designed to permit leak-testing and evacuation of evaporator coil and all field-installed piping with condensing unit remaining isolated. Except to check holding charge, both service valves should be left in front-seat position (as shipped) until field-installed part of circuit has been leak-tested, evacuated (or purged), and partially charged (to approximately 50 PSI gauge).

INSTALLATION INSTRUCTIONS – SHEET FOUR

LOCATION

Condensing Unit should be located with consideration given to the following:

1. Unit is designed for outdoor installation without duct work.
2. Since air is used as a cooling medium, inlet air to coil must be drawn from as cool a location as possible. East or North sides of building usually are desirable as they are away from sun load.
3. Condenser intake should not be closer than 12" to any wall to provide adequate air circulation through unit.
4. Select location near evaporator coil that requires not more than 50 linear feet of refrigerant piping and a minimum number of turns. Vertical rise of liquid line must not exceed 20 feet. (Units over 5 tons: See instructions for specific unit).
5. Allow a minimum of 2 feet for service access at front of unit.
6. A 4" (minimum) concrete slab at least 6" wider and 6" longer than the box is recommended as a foundation for outdoor installation. Slab should not be tied to house foundation as noises will telegraph through.
7. If unit is installed on roof, be sure structure will support the weight. If roof has a parapet maintain a minimum clearance of 12" between parapet and Condenser intake grilles.
8. Guard condenser from damage if local activities are such that a possibility exists. When necessary, provide safeguards to prevent small children from burning or cutting themselves on condensing coil.

INSTALLING EVAPORATOR COIL

See manufacturers instructions packed with coil.

FILTERS

Filters must always be used ahead of evaporator coil and must be kept clean—absence of filters will result in clogged coils and frequent, costly cleaning. Dirty filters will reduce capacity and will result in frosted coils or safety shut down.

CONDENSATE PIPING

A condensate line is required from evaporator coil to an open drain. Line should be same pipe size as drain nipple and should pitch downward toward drain. It is suggested that line be installed with a trap, and and plugged tees used instead of elbows to facilitate cleaning.

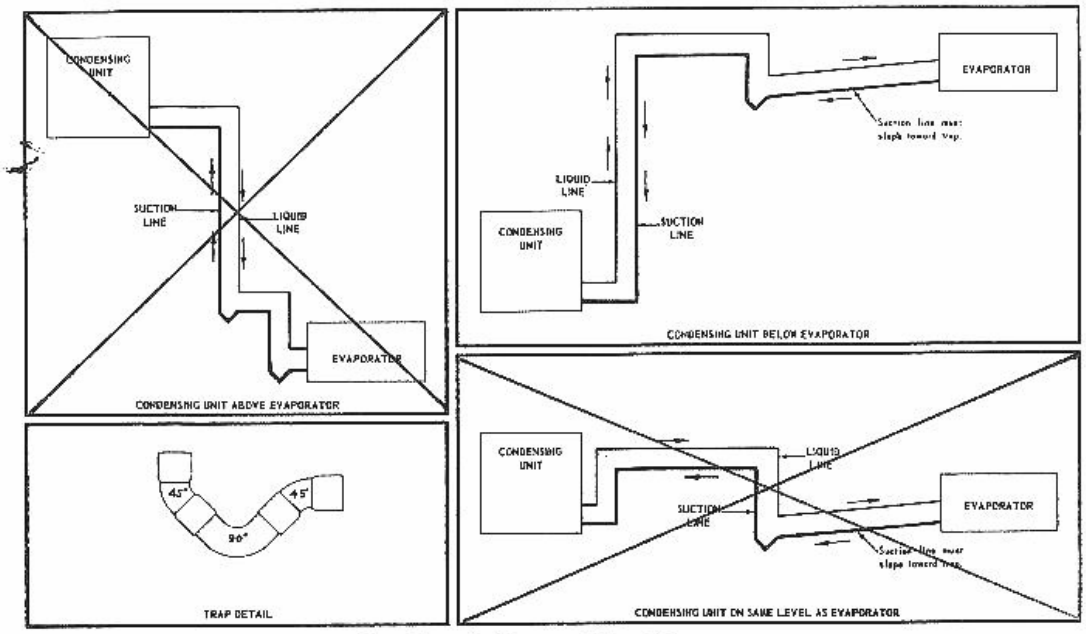


Fig. 5 – Refrigerant Piping Diagrams

REFRIGERANT PIPING

Hard or soft drawn copper tubing may be used at the convenience of installer within the limitation of available sizes. Soft drawn tubing is available only up to 7/8" OD. Tubing sizes required are given in instruction for the specific unit.

Tubing should be handled with a great deal of care to prevent contamination of the system by entry of dirt, moisture, metal filings, etc. The following precautions are suggested:

Soft Drawn Tubing: Use only new refrigerant tubing, which has been dehydrated and sealed by the manufacturer. Hard Drawn Tubing: Normally not available sealed; use extra care in handling and storage.

During installation, keep tubing sealed until adjacent to joint to be made. Never open tubing to atmosphere when temperature of tubing is below dew point. (Tubing which is "sweating" is below dew point and should not be opened until warmed above room temperature.) A cut length of soft tubing can be effectively sealed by hammering the end shut, while hard tubing should be kept closed with plastic plugs.

Cut ends must be reamed, but with tubing in such a position that filings will fall out of rather than into the tubing.

For soft drawn tubing, flare unions are usually as effective as soldered joints.

Soldered joints should be made with a low melting point silver solder or with 95-5 soft solder (95% tin, 5% antimony). NEVER USE SOLDER CONTAINING LEAD.

Always use a non-corrosive type flux, as even very small quantities of acid in the system can seriously damage compresser. Silver-soldered joints should be cleaned with a damp cloth to prevent corrosion.

PIPING SEQUENCE AND LEAK-TEST

1. Solder suction line to evaporator stub, and run line to stub of Suction Service Valve at condensing unit, soldering line to the stub. Be sure to provide necessary traps in suction line. (See Fig. 8).
2. Connect liquid line to Liquid Line Service Valve, and run line to filter-drier location near expansion valve. Leave enough separation between suction line and liquid line so that suction line can be insulated.
3. Connect sight glass (if used) to *outlet* end of filter drier. Connect other end of sight glass to line leading to the expansion valve, and other end of filter-drier to the liquid line coming from condensing unit. *Avoid exposing filter-drier to atmosphere for more than a few seconds.*
4. Purge air from evaporator and piping as follows:

NOTE: *USE REFRIGERANT-22 FOR ALL PURGING.*

a) Check to be sure that both service valves are front-seated (turned tightly clockwise).

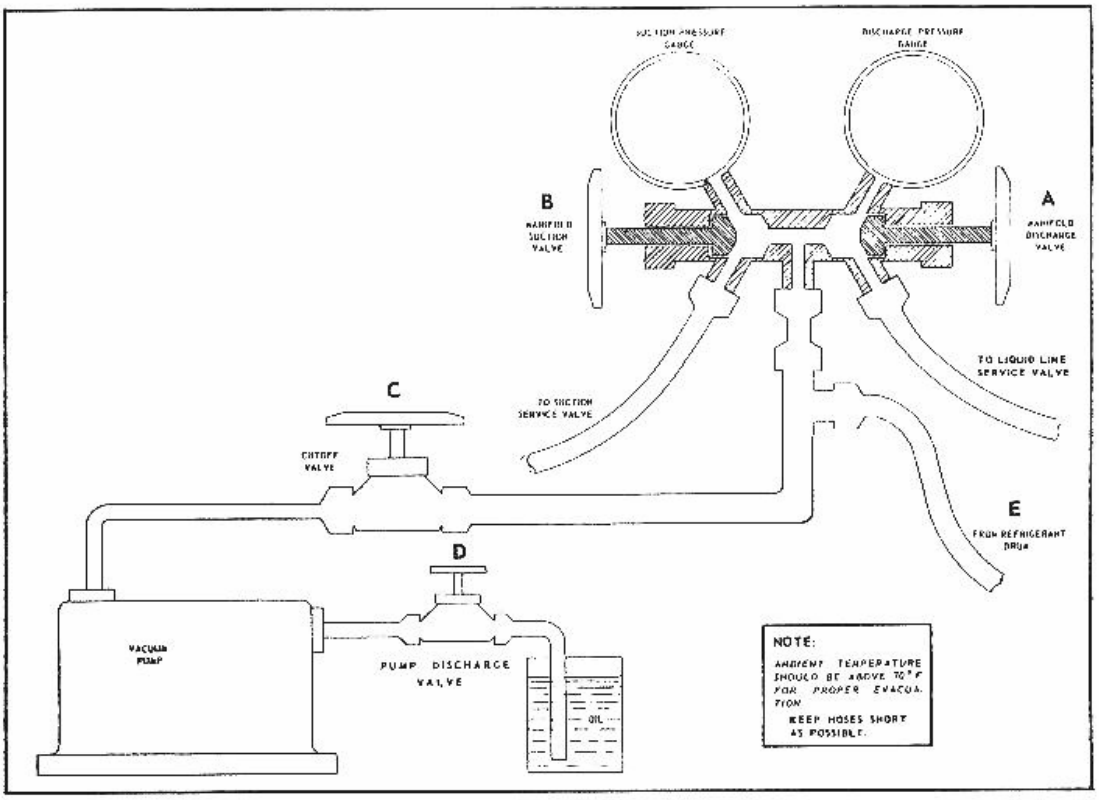


Fig. 6 – Evacuating and Charging through Manifold

b) Remove gauge port cap from SUCTION service valve, and if condensing unit is indoors, attach purging hose leading outdoors.

c) Remove gauge port cap from LIQUID LINE service valve and connect refrigerant drum through charging manifold.

d) Open refrigerant drum valve quickly, and purge for 15 seconds. Close refrigerant drum valve. Replace cap on SUCTION service valve.

e) Charge field-installed part of circuit to 75 PSI gauge for leak-test.

f) Check with Halide torch for leaks at flare connections, tightening flare fittings as necessary.

g) Check all soldered joints, including those on the evaporator coil, with Halide torch. If a leak is found which requires soldering, pressure in the circuit must be bled off since it is impossible to solder with unit pressurized. Be sure *all* leaks are located and marked before bleeding pressure from the circuit. If indoors, room should be thoroughly ventilated before soldering.

h) When leaks, if any, have been repaired, circuit is ready to be evacuated (or purged) and charged.

EVACUATION vs PURGING

Even with all possible precautions it is impossible to install a remote type cooling circuit without contamination. Soldering will cause oxidation unless a continuous flow of oil-pumped dry nitrogen is maintained in the circuit during soldering operation. Careful handling of tubing during installation will minimize but will not eliminate moisture in the circuit.

Purging (without evacuation) may or may not reduce this contamination to a point where it will not interfere with proper operation of the system. Proper evacuation with a good vacuum pump will insure that contamination will not interfere with proper operation. For that reason, evacuation of evaporator coil and all field piping is recommended in all cases, and evacuation of condensing unit is recommended whenever repair of that section has been necessary.

In recognition of the fact that a suitable vacuum pump may not always be available in the field, these instructions cover both the purging and evacuation methods.

Proper evacuation requires a good vacuum pump. A pump specified by the manufacturer to be capable of a vacuum of 1,000 microns absolute pressure will be adequate for installation work.

Avoid unnecessarily pulling refrigerant through a vacuum pump, as it will cause the pump to lose efficiency. Instructions call for refrigerant pressure to be bled off before connecting pump. It is good practice to let the pump operate in open air for a few minutes after it has been used to evacuate a circuit in order to purge refrigerant from pump.

EVACUATION AND REPLACING HOLDING CHARGE

An indicator to show when evacuation is complete may be made as follows (See Fig. 6).

Connect pipe or tubing with a good cut-off valve to the discharge side of vacuum pump and run tubing into a fruit jar or similar receptacle containing about two inches of light weight oil so that discharge is below oil level. When bubbles no longer appear in the oil, the pump has done all it can do and further evacuation is a waste of time.

WARNING: *SHUT OFF VALVE IN PUMP DISCHARGE LINE BEFORE SHUTTING PUMP OFF.*

To remove non-condensables and moisture the entire circuit should be double evacuated and the charge replaced. The double evacuation (evacuating, breaking vacuum to positive pressure with refrigerant, and then evacuating again) can best be done using a charging manifold with a Tee and cut off valve as shown in diagram, Fig. 6. The Tee connection eliminates need to disconnect pump in order to connect refrigerant drum, and may be omitted if care is taken to properly purge hoses each time connection is made.

1. Bleed pressure from circuit through discharge valve gauge port.
2. Connect vacuum pump as shown in diagram (Fig. 6) and completely evacuate circuit until no bubbles appear in oil in "indicator".
3. Close cut off valve (C) in line to vacuum pump.
4. Close valve (D) on vacuum pump discharged line and shut pump off.
5. Open refrigerant drum valve and charge to positive pressure (not more than 5 PSI gauge).
6. Close refrigerant drum valve.
7. Loosen hose connection at drum to bleed pressure from circuit. Then tighten hose connection.
8. Open pump discharge valve (D) and start pump.
9. Open cut off valve (C) in line to vacuum pump.
10. Evacuate circuit completely.
11. Close cut off valve (C) to vacuum pump.
12. Close valve (D) on vacuum pump discharge line and shut pump off.
13. Open refrigerant drum valve and charge until pressure in circuit balances pressure of drum. (Minimum 50 PSI).
14. Unless circuit is to be charged immediately, back-seat both service valves, disconnect hoses and replace both gauge port plugs. Otherwise, see "**CHARGING**" Page 7.

PURGING (Alternate Procedure)

If vacuum pump is not available, circuit should be thoroughly purged and charge replaced after leak-test is completed. Proceed as follows:

1. Remove cap from suction service valve gauge port to bleed pressure from field-installed part of circuit.
2. Connect refrigerant drum through charging manifold to liquid line service valve gauge port, purging hose before tightening connection at gauge port.
3. Open refrigerant drum valve quickly so that flow of gas through the lines will be fast enough to purge effectively. Purge for 15 seconds, then close drum valve and wait 3 to 4 minutes.
4. Open drum valve quickly, purge for 30 seconds, close drum valve, and wait 3 to 4 minutes.
5. Open drum valve quickly, purge 30 seconds, and replace cap on suction service valve gauge port while still purging.
6. Charge until pressure in circuit balances pressure of drum (Minimum 50 PSI).
7. If circuit is *not* to be charged immediately, back-seat liquid line service valve and replace gauge port cap.
8. If circuit is to be changed immediately, back-seat suction service valve, connect hose from suction side of manifold to gauge port (purging hose before tightening connection at gauge port) and follow instructions under "**CHARGING**" on next page.

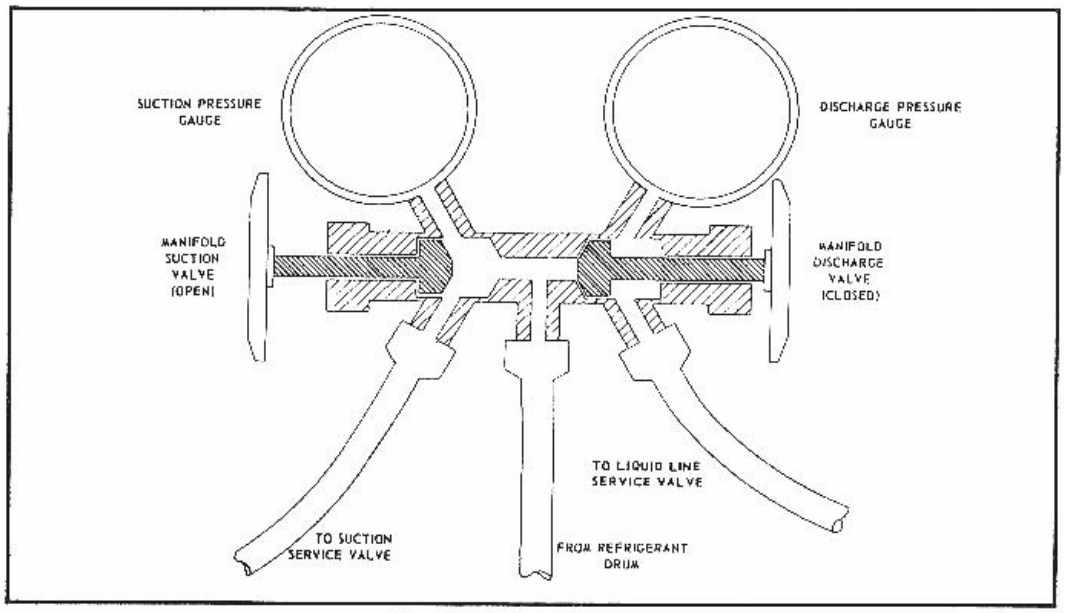


Fig. 7 – Charging Manifold

TABLE I — SATURATED TEMPERATURE — PRESSURE EQUIVALENTS

Temperature °F.	Refrigerant–22 PSI
30	55.2
32	57.8
34	60.5
36	63.3
38	66.1
40	69.0
45	76.6
50	84.7
55	93.3
60	102.5
65	112.2
70	122.5
75	133.4
80	145.0
85	157.2
90	170.1
95	183.7
100	197.9
105	212.9
110	228.7
115	245.3
120	262.6
125	280.7
130	299.3
135	319.6
140	341.3
145	364.0
150	387.2

CHARGING

Proceed as follows to charge circuit, keeping refrigerant drum upright at all times:

1. Back-seat service valves and remove caps from gauge ports.
2. Connect refrigerant drum through charging manifold to both gauge ports leaving connections at gauge ports loose. Purge each hose and tighten connections. Mid-seat both service valves.
3. With manifold suction valve **open** and manifold discharge valve **closed**, open refrigerant drum valve and allow pressure in circuit to balance pressure of drum.
4. Set system control on thermostat to "COOL" and set thermostat dial below room temperature.

   Set thermostat fan switch on "AUTOMATIC". Turn main disconnect switch "ON" to start compressor.
5. Warm refrigerant drum by partly immersing in warm water as necessary to expedite charging. Do not allow suction pressure to exceed 100 PSI gauge.
6. Continue charging until discharge pressure gauge shows a reading equal to ambient dry bulb temperature plus 32°, converted to pressure by means of Table I.

EXAMPLE: Ambient dry bulb temperature = 93° F.
− 32° F.
125° F.

INSTALLATION INSTRUCTIONS – SHEET NINE

7. From Table I, Refrigerant-22 pressure corresponding to 125° F. 280.7 PSI. Charge until discharge pressure gauge reads 280 PSI.
8. Close refrigerant drum valve and manifold suction valve. Turn main disconnect switch off.
9. Allow unit to run for 20 minutes, reading discharge pressure to see that it remains constant. If discharge pressure drops, more refrigerant must be added to bring it back to the computed pressure.

PUMP DOWN

If a leak develops in the evaporator section (anywhere between Liquid Line Service Valve and Suction Service Valve) most of the refrigerant in system can be saved by pump-down.

1. Front-seat LIQUID LINE SERVICE VALVE.
2. Attach gauge to suction service valve gauge port, and run compressor until gauge reads 5 lbs., then shut off compressor.

   WARNING: ***Never pump down to zero suction pressure.***
3. Front-seat SUCTION SERVICE VALVE to hold refrigerant in condenser.
4. Bleed refrigerant from evaporator section through SUCTION VALVE Gauge Port.
5. Complete repair of leak.
6. Follow instructions in this manual under "EVACUATION" (or "PURGING" if vacuum pump is not available.)
7. Follow instructions in this manual under "CHARGING"

INSULATION

DO NOT INSULATE SUCTION LINE AND LIQUID LINE TOGETHER.

Entire length of Suction line must be insulated with Prestite or equivalent insulation. In addition, all outdoor portions should receive two coats of asphaltic water-proofing. Insulation should be brought up tight against coil and unit casing to prevent sweating at these points. Pipe hangers and pipe clamps, if required, should be installed after applying insulation to prevent line pulsations from entering building structure. Where lines pass through walls or partitions, insulation should be used to isolate tubing from structure as well as to seal openings.

OPERATION INSTRUCTIONS – SHEET ONE

OPERATION

The operation of the system is controlled by the remote mounted thermostat. The thermostat turns on the compressor, outdoor fan motor and indoor unit fan motor.

The compressor and outdoor fan motor are protected by internal automatic reset protective devices. The devices are sensitive to thermal rise in the motor windings and are responsive to overload conditions.

Should the fan motor protector cut off the fan motor, the compressor will continue to operate until the high pressure relief opens and shuts the compressor down. Should the compressor be shut down by its protective device, the fan motor will continue to operate.

The compressor protective device will trip in approximately 3 to 5 seconds under locked rotor conditions. Reset time is 1/2 to 3 minutes. Should the condition continue, causing the compressor to cycle on and off, each period grows progressively longer due to thermal build-up in the motor windings. When this occurs, reset time may require 30 to 45 minutes.

When the compressor internal pressure relief opens, the compressor continues to operate for a short time. The relief valve will not reset until the high to low pressure differential drops below 100 PSI. Normally, this will not occur until the compressor stops.

With the pressure relief valve open, the compressor continues to operate under heavy overload conditions. With hot discharge gas being relieved inside the compressor shell, the compressor internal protective device will soon open to stop the compressor due to thermal rise and overload conditions. Several minutes may be required for reset.

SERVICE

Service access to the compressor compartment and control box is had by removal of the entire top panel and fan grille. (See Figure 2.) The fan grille is attached to the top panel and should not be removed from the panel for normal service. Four screws, one located on each side and each end of unit, secures top panel in place.

The panel adjacent to the compressor may be removed for additional accessibility to the compressor compartment.

The system may be run with the top panel removed for various operational checks. **Use caution with exposed high voltage components and rotating fan blade!**

The fan motor mounting bracket is secured to the fan orifice panel.

For service access to the fan motor and assembly:

1. Disconnect all electrical power.
2. Disconnect fan motor leads from control connections at contactor and fan capacitor. (See wiring diagram.) Also disconnect green grounding lead from the partition panel.
3. Remove the four screws securing the fan orifice panel in place, 2 at the compressor compartment end and 2 thru the panel at the opposite end. (See Figure 2.)
4. Lift out the entire orifice panel and fan assembly.

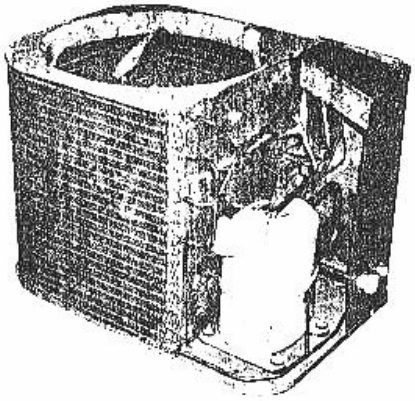


FIGURE 2
SP UNIT WITH TOP PANEL REMOVED

Reassemble by reversing the above procedure. Make sure the coil protective grille is inside the flange of the orifice panel.

Should it be necessary to remove the fan blade or fan motor, reassemble as indicated below:

1. Locate lower edge of motor rain shield approximately 1/4 in. from face of motor.
2. Locate upper wire of motor belly band approximately 7/8 in. below uppper edge of motor.
3. Position fan blade with top edge of blade 3/4 in. below flat top of orifice panel.
4. Fan blade set screws should be torqued to 50-60 in. lbs.

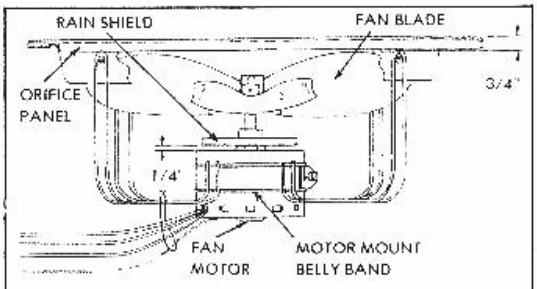


FIGURE 3. FAN ASSEMBLY POSITION

OPERATIONAL PREVENTIVE MAINTENANCE

1. Westinghouse fan motor with oiling port on lower bearing will require annual oiling. Use SAE No. 10 oil, 15-30 drops. Upper bearing is sealed and does not require oiling. If unit has alternate GE fan motor, the motor has sealed bearings (no oiling ports) and does not require annual oiling.
2. Periodically clean the inside of the unit to keep the weep holes in the base pan open to assure proper drainage of water from the unit.
3. Keep the condenser coil clean and free of anything that restricts free airflow.
4. Periodically check for looseness of cabinetry and other components, abnormal vibration and any undue rubbing of the unit tubing. Tighten or adjust as necessary.
5. Periodically check the system and tubing for indications of leaks. Unexplained oil may indicate a system leak.
6. Indoor air filters should be changed or cleaned on a regular basis depending upon operating conditions. If entering surfaces are coated or light cannot readily be seen thru the filter, it should be cleaned or replaced. Disposable filters should not be cleaned for re-use more than once.
7. Duct system should be checked periodically for loose insulation and air leaks. Correct as needed. All supply and return air openings should be fully open and unblocked to allow full airflow.
8. If used, indoor fan belt should be periodically checked for proper tension, wear and aging. A properly adjusted belt will allow approximately 1 inch depression at mid-point between pulleys for each 12 inches between shafts. Worn or cracked belts should be replaced.
9. Outdoor and indoor unit access panels properly in place and secure.

LIMITED WARRANTY

FIRST YEAR WARRANTY ON ENTIRE UNIT

If any part of your new Westinghouse Air Conditioning unit fails because of defective material or workmanship within twelve (12) months from the date of installation, Westinghouse will furnish a replacement part.

Any labor charges connected with the removal of the defective part and the installation of the replacement part are not included. Parts are available through Westinghouse Central Residential Air Conditioning outlets.

Transportation charges, if any, are not included. Filters and refrigerant are not included.

The servicing of the products covered by this warranty is the responsibility of the Dealer or Contractor who installs the equipment, and the installing Dealer or Contractor should be contacted directly by the Owner for service and in-warranty parts. Requests received at the Westinghouse factory for service, including orders for parts, will be referred to your nearest Westinghouse Distributor.

FOUR YEAR EXTENDED WARRANTY ON MOTOR-COMPRESSOR UNIT
(Applicable to units with sealed motor-compressors)

If the sealed motor-compressor should fail because of defective material or a manufacturing defect during the second through the fifth years after the date of installation, Westinghouse will furnish a replacement motor-compressor. Transportation charges from the factory will be the responsibility of the owner.

Labor charges, refrigerant, and any miscellaneous parts and/or material associated with the installation of the new motor compressor are the responsibility of the owner.

EXCLUSIONS

This warranty applies only to units sold and retained within continental U.S.A.

This warranty does not apply to the product or parts that have been damaged by accident, abuse, lack of proper maintenance, unauthorized alterations, misapplication, improper installation, fire, flood or acts of God.

IN NO CASE WILL WESTINGHOUSE BE RESPONSIBLE FOR CONSEQUENTIAL DAMAGES.

GENERAL

Furnishing of parts as described above shall constitute fulfillment of all Westinghouse obligations with respect to this warranty, and replacement parts will be warranted only for the unexpired portion of the original warranty.

Westinghouse

CENTRAL RESIDENTIAL AIR CONDITIONING DIVISION NORMAN, OKLAHOMA 73069

WORLDWIDE
WESTINGHOUSE ELECTRIC CORPORATION
5005 INTERSTATE DR. NORTH NORMAN, OKLAHOMA 73069
[illegible] Norman OK

IN CANADA
WESTINGHOUSE CANADA LIMITED
AIR CONDITIONING DEPARTMENT
692 SPEERS ROAD OAKVILLE, ONTARIO L6K 2G3

SPECIFICATIONS SUBJECT TO CHANGE WITHOUT NOTICE

Glossary of Statistical Terms

Statistical Terms Glossary

Term	Definition
α (alpha)	○ The probability of making a type I error. It is related to the confidence by confidence = (1 – α) x 100.
β (beta)	○ The probability of making a type II error. It is also called Power. This symbol is also sometimes used to indicate coefficients in a polynomial model.
σ (sigma)	○ The symbol for the true standard deviation. We can never know the true standard deviation because it is obscured by response variation, so it is a theoretical idea.
χ2 (chi-square)	○ A distribution used to make decisions about standard deviations.
χ2 Critical	○ A value from the χ2 distribution to compare with a computed χ2 value in a χ2 test.
χ2 Test	○ A test comparing a sample divided into at least three categories with a standard.
Antagonism	○ An interaction of factors producing a poorer response than anticipated by considering the factor levels independently.
Attribute Data	○ Data that are not continuous. Attribute data fit into categories that can be described by numbers or words. Examples are Pass/Fail and Number of people residing in a home.
Attributes Control Chart	○ A control chart for monitoring attribute data.
Average	○ A term used to indicate either the mean or the median depending on the situation. Typically it refers to the mean in industrial contexts.
b-Coefficient	○ A number that multiplies a term in a model to define the shape of the response surface.
Block	○ A group of experimental trials.
Blocking	○ A technique that breaks experiments into smaller groups of trials to eliminate the effect of a nuisance factor.
Blocking Factor	○ A factor used to divide experimental trials into blocks for blocking.
Box-Behnken Design	○ An experiment design for collecting data to fit a full quadratic model. Also called a “Face-Centered Cubic Design” or an “Interaction Plus Star Design.”
Box-Cox Plot	○ A plot of the log of the mean vs. the log of the standard deviation for replicate measurements. The slope of this line can help determine an appropriate transformation for data to make the standard deviation constant over the region of interest. The Box-Cox Transformation raises

Term	Definition
	each response to the power of 1 minus the slope of this line.
Box-Cox Transformation	○ A mathematical operation performed on responses to make the standard deviation of the transformed responses constant over the region of interest. The Box-Cox Transformation raises each response to the power of 1 minus the slope of the line in a Box-Cox Plot.
Cp	○ A measure of process potential, it compares the process variation to specification limits.
Cpk	○ A measure of process capability, it compares the process variation to the specification limits taking the process mean into account.
Cpl	○ A measure of the process capability, it compares the process variation below the process average with the specification limit below the process average.
Cpu	○ A measure of the process capability, it compares the process variation above the process average with the specification limit above the process mean.
c-Chart	○ An attribute control chart used to monitor the number of non-conformities in a unit.
Center Point	○ An experiment in which all factors are set to their middle value.
Central Composite Design	○ An experiment design to collect data to fit a full quadratic model. The region of interest is spherical for this design.
Central Limit Theorem	○ The theorem that states that the distribution of means approaches a normal distribution as the sample size approaches infinity. It allows us to assume Normal behavior for averages.
Coded Factors	○ Factors can be coded to a different scale to improve numerical accuracy in calculations. A common coding for process factors is to to code the high level as +1 and the low level as -1 with intermediate values coded linearly over the range [-1,1].
Coded Responses	○ Sometimes it is valuable to code responses to make it easier to find a Sweet Spot. For example, you can code your goal as 1, the unacceptable values as 0, and intermediate response values a linearly over the range [0, 1]. The Sweet Spot will have a maximum value for the coded response regardless of whether you are seeking a maximum, a minimum, or a number in a range.
Coding	○ Changing a value to a different value, such as coding a high level as +1 and a low level as -1.
Combined s	○ An estimate of the standard deviation for the region of interest formed by combining separate estimates of the standard deviation by a technique called pooling.
Common Cause	○ An intrinsic cause of variation in a process. Variation which is part of a process when it is operating normally.
Confidence	○ The probability that an estimate is correct. For example, the probability that a confidence

Term	Definition
	interval contains the long-term mean.
Confidence Interval	○ An interval between a lower and an upper limit in which the long-term mean should lie within the specified confidence.
Confidence Limits	○ The lower and upper limits for a confidence interval.
Consistent Standard Deviation Estimates	○ A term used to indicate that the standard deviation is constant over the region of interest.
Continuous	○ Refers to the property of a factor or response that a level can be found between any two other levels, no matter how small the difference between these other levels.
Contour Plot	○ A picture of a model that shows contour lines of constant response.
Control Chart	○ A run chart with limits indicating the normal operation of a process. Typically these limits are ± 3 times the standard deviation of the statistic being tracked.
Critical Value	○ A value from a distribution to which a calculated value is compared in a significance test.
Cube	○ A three-dimensional object that represents the region of interest for a 3-factor process factor experiment.
Cubic Model	○ A polynomial model containing, at a minimum, a constant term, a term for each main effect, a term for each two-factor interaction between main effects, a term for the square of each main effect, and a term for each 3-factor interaction of the main effects. The full cubic model also includes terms for the interactions of each quadratic main effect with each linear main effect and terms for the cube of each main effect.
Cycle	○ A periodic fluctuation, such as a temperature cycle.
Degrees of Freedom	○ The number of values which could be arbitrarily assigned. For example, if you want to assign data to the categories “pass,” “fail high,” and “fail low,” the degrees of freedom is 2, because once data has been assigned to any 2 categories the remainder must be assigned to the third – it cannot be assigned arbitrarily.
Design Cube	○ A three-dimensional object that represents the region for data collection for a 3-factor process factor experiment.
Desireability	○ An indication of how well a response meets a goal. Typically desireability uses coded responses with 0 indicating that the goal is not met at all, 1 indicating that the goal was met perfectly, and intermediate values indicating that the goal is partially met. For multiple goals the overall desireability is typically calculated as the geometric mean of the individual desireabilities.

Term	Definition
Discrete	○ The property of a factor or response that only certain levels are allowed.
Drift	○ An increase or decrease over time in a factor level or response.
Effect	○ The average difference in a response as a factor varies from its high to its low level. These can be ranked in magnitude to determine the relative importance of terms in a model.
Experiment Design	○ A plan for collecting and analyzing data.
Experimental Error	○ A term encompassing both Systematic Error and Random Response Variation.
Exponential Distribution	○ A distribution described by one parameter – "lambda." Often used to model times to failure for electronic devices, or the amount of material remaining in a radioactive decay process
External Factor	○ A factor not being studied in an experiment, but influencing one or more experimental responses.
F	○ The ratio of two variances, each divided by its degrees of freedom.
F Critical	○ A value from the F-distribution that is compared to a calculated F statistic in a significance test. This is used to test for statistically significant differences in variances (standard deviations).
F-distribution	○ A distribution used to describe variance ratios.
F-test	○ A test to determine whether a difference between two variance estimates is statistically significant.
Face-Centered-Cubic Design	○ An experiment design for collecting data to fit a full quadratic model. Also called an "Interaction Plus Star Design" or a "Box-Benkhen Design."
Factor	○ Something over which you have direct control in an experiment.
Factor Transformation	○ A mathematical operation performed on factors to simplify the shape of the response surface.
Factorial Design	○ An experiment design used to study factors and their interactions.
Fractional Factorial Design	○ A factorial design that studies only some of the possible factor interactions, usually just the 2-factor interactions.
Full Cubic Model	○ A polynomial model containing a constant term, a term for each main effect, a term for each two-factor interaction between main effects, a term for the square of each main effect, a term for each 3-factor interaction of the main effects, a term for the interactions of each quadratic main effect with each linear main effect, and a term for the cube of each main effect.
Full Factorial Design	○ A factorial design that studies all possible interactions.
Full Quadratic Model	○ A polynomial model containing a constant term, a term for each main effect, a term for each two-factor interaction between main effects, and a term for the square of each main effect.

Term	Definition
G-Optimal Design	○ An experiment design that minimizes the largest variance in the region of interest. While this may the best type of design for DOE work, it is not yet possible to generate these designs to order. The I-Optimal design is the next best choice in most situations.
Gaussian Distribution	○ A distribution of measurement error that is widely applicable to many types of continuous data. This distribution is bell-shaped (AKA – “normal distribution”)
Grid Search	○ A method used for finding a Sweet Spot in which the region of interest is divided into a grid, a response is predicted at each grid point, and the grid point or points meeting the goal are listed as potential Sweet Spots.
Homogeneous Variance	○ A term used to indicate that the variance (standard deviation) is constant over the region of interest.
Hypercube	○ A multi-dimensional object that represents the region of interest for a four or more factor experiment.
Optimal Design	○ An experiment design that minimizes the average variance of prediction over the region of interest. The goal is to make accurate and precise predictions.
In-Control	○ Describes a process which is only subject to common causes of variability.
Interactions	○ Changes in responses due to changes in factor levels that depend on other factor levels. The effect of factors changing levels together.
Interaction Design	○ An experiment design for collecting data to fit an interaction model. Also known as “Factorial” or “Fractional Factorial” designs.
Interaction Model	○ A polynomial model that includes, at a minimum, a constant term, terms for each main effect, and terms for each 2-factor interaction between main effects. It can also contain higher order interaction terms, such as 3-factor interactions among the main effects.
Interaction Order	○ The number of main effects included in an interaction term.
K	○ A constant used to calculate statistical tolerance limits.
Level	○ A value for a factor.
Main Effect	○ A reference to a factor in an experiment as it is represented in a model.
Main Effects Design	○ An experiment design for collecting data to fit a main effects model.
Main Effects Model	○ A model that includes only a constant term and a term for each main effect. This model is used for screening.
Mathematical Model	○ A formula that defines the surface that expresses a theory.
Mean	○ The arithmetic average.

Term	Definition
Mean Shift	○ A change in the mean value for a process.
Median	○ The number below which half of a data set fall and above which the other half of this data set fall.
Mode	○ The most frequently occurring number in a data set.
Model	○ A theory expressed as a surface.
N	○ Used to indicate the number of replicate measurements.
Np-Chart	○ An attribute control chart used to monitor the number of nonconforming units.
Noise	○ A colloquial term for Random Response Variation.
Normal Distribution	○ A distribution of measurement error that is widely applicable to many types of continuous data. This distribution is bell-shaped (AKA “Gaussian distribution”).
Normal Probability Plot	○ A plot that displays normally distributed data as a straight line.
Normal Quantile Plot	○ A plot that displays normally distributed data as a straight line.
One-Factor-At-a-Time (OFAT)	○ A method for performing experiments in which one factor is varied while all others are held constant.
Out-of-Control	○ Describes a process which is subject to a special cause or special causes of variability.
p-Chart	○ An attribute control chart used to monitor the proportion of nonconforming units.
p-Level	○ An estimate of the probability of making a type I error.
Pareto Analysis	○ A data driven approach to identify the biggest contribution to a problem or defect – relies on empirical observation that often 80% of the problems are caused by less than 20% of the process variables.
Ppk	○ Similar to CPK, but calculated using the process variability instead of the within sample variability. It measures process performance as opposed to capability.
Pooled s	○ An estimate of the standard deviation for the region of interest formed by combining separate estimates of the standard deviation by a technique called pooling.
Pooled Standard Deviation	○ An estimate of the standard deviation for the region of interest formed by combining separate estimates of the standard deviation by a technique called pooling.
Pooling	○ A technique which combines standard deviation estimates. Pooling calculates the square root of the degrees-of-freedom-weighted average of the variances.
Power	○ The confidence that two numbers are the same. 1-Power is the probability of making a type II error.

Term	Definition
Prediction Interval	○ The interval between a lower and an upper limit in which the next sample measurement is likely to lie with specified probability.
Prediction Limits	○ The lower and upper limits of a prediction interval.
Probability	○ The likelihood that something will occur. For example, the probability that a mean will lie within a 95% confidence interval is 0.95, or 95%.
Process Capability	○ A measure of the ability of a process to meet specifications (metrics include Cp, Pp, Sigma-Level).
Pure Error	○ An estimate of the standard deviation for the region of interest formed by combining separate estimates of the standard deviation by a technique called pooling.
Quadratic Design	○ An experiment design for collecting data to fit a full quadratic model.
Quadratic Model	○ A polynomial model containing, at a minimum, a constant term, a term for each main effect, and a term for each two-factor interaction between main effects (see interaction model). If it also contains terms for the squares of the main effects it is a “Full Quadratic Model.”
R-Chart	○ A control chart for monitoring ranges in continuous data.
Random Response Variation	○ A description of the naturally occurring phenomenon that introduces random variation into measured data.
Random Run Order	○ A randomly selected order for running experimental trials. The purpose is to protect against being mislead by the effects of external factors.
Range	○ The difference between the highest and lowest values in a set of replicate measurements.
Range Chart	○ A control chart for monitoring ranges in continuous data.
Region of Interest	○ The experimental region over which you would like to make predictions of experimental results.
Replicate	○ An experimental run.
Residual	○ The difference between an observed value for an experimental trial and the corresponding predicted value.
Response	○ Something over which you have no direct control, but which you need to control. You must control responses by adjusting appropriate factors.
Response Surface	○ A model of a response to factor level changes. It is a multidimensional surface, having a dimension of 1 + the number of factors.
Response Surface Model	○ A model of a response to factor level changes. It is a multi-dimensional surface, having a dimension of 1 + the number of factors.

Term	Definition
Response Transformation	○ A mathematical operation performed on responses to make the standard deviation of the transformed responses constant over the region of interest.
Response Variation	○ A description of the naturally occurring phenomenon that introduces random variation into measured data.
Run	○ To perform an experimental trial. Runs are not unique. One trial can be performed several times, each time counting as a run.
Run Chart	○ A chart plotting measurements in their run order.
Run Order	○ The order in which experimental trials are performed.
s-Chart	○ A control chart monitoring standard deviation for continuous data.
Sample Size	○ The number of replicate measurements performed or to be performed.
Scrambled Run Order	○ A random run order providing better protection against shifts, drifts, and cycles than other random run orders.
Screening	○ The act of reducing the number of factors in an experiment to a more manageable size.
Shift	○ An immediate change in a response value.
Significance	○ Importance of an observed difference.
Significance Test	○ A test to determine if an observed difference is important in terms of probability. In other words, if the probability of a difference being real is sufficiently high, this test will indicate that the difference is important.
Sigma	○ The symbol for the true standard deviation. We can never know the true standard deviation because it is obscured by response variation, so it is a theoretical idea.
Sigma-Level	○ A measure of the capability of a process. Essentially, the number of standard deviations the process mean is from the specification limits.
Significant	○ Important.
Six Sigma	○ A program for improving decisions using data and statistical analysis.
Special Cause	○ A cause of variation in a process that is not a part of normal operation.
Special Cubic Model	○ A full cubic model minus the terms for the cube of each main effect. This is often used to study mixtures and is presented in an equivalent form that has no constant term. Please see Math Options texts for more details.
Standard Deviation	○ A measure of the variation in a set of data. Calculated by squaring the differences between the data and the mean, dividing by the number of data (or n-1 for samples) and then taking the square of the result.

Term	Definition
Standard Deviation Chart	○ A control chart monitoring standard deviation for continuous data.
Star Points	○ Experimental trials added to an Interaction Design or a Factorial Design to allow a fit to a full quadratic model. They are located at the centers of edges in square regions of interest and in the centers of faces in cubical and hypercubical regions of interest.
Statistical Process Control (SPC)	○ A system for monitoring process performance and taking action when special cause signals are detected.
Statistical Control	○ Describes a process which is only subject to common causes of variability.
Statistically Significant	○ Refers to the probability that a difference is real is high.
Statistical Tolerance Interval	○ The interval in which a specified proportion of all future measurements should lie with a specified probability.
Statistical Tolerance Limits	○ The lower and upper limits of a tolerance interval.
Sweet Spot	○ A combination of factor levels that satisfies all response goals.
Student's t Value	○ A value from a t-distribution.
Synergism	○ An interaction between factors producing a response better than anticipated when considering the factors separately.
Systematic Error	○ An error which can, if identified, be eliminated, at least in theory.
t	○ A value from a t-distribution.
T critical	○ A value from a t-distribution that is compared to a calculated t-value in a significance test.
t-Test	○ A significance test that compares means.
Tolerance Interval	○ The interval in which a specified proportion of all future measurements should lie with a specified probability.
Tolerance Limits	○ The lower and upper limits of a tolerance interval.
Transformation	○ A mathematical operation performed on responses to make the standard deviation of the transformed responses constant over the region of interest, or on factors to simplify the shape of the response surface.
Twisted Plane	○ Interactions cause a plane to twist. An interaction between two factors manifests itself in a response surface as a twisted plane.
Trial	○ A set of factors and levels for an experiment.
Type I Error	○ Concluding that a difference is real when it is not.

Term	Definition
Type II Error	○ Concluding that a difference is not real when it is real.
u-Chart	○ An attribute control chart that monitors the proportion of non-conformities.
Uncertainty	○ The doubt about a conclusion due to Random Response Variation.
Uniform Shell Design	○ An experiment design for a spherical region of interest that is more efficient than the Central Composite Design but less efficient than IOptimal Designs can be.
Uniform Standard Deviation	○ A term used to indicate that the standard deviation is constant over the region of interest.
Variable Data	○ Continuous data.
Variables Control Chart	○ Any control chart monitoring continuous data.
Variance	○ The square of the standard deviation.
Weibull	○ A probability distribution described by three parameters – “beta, eta, and t-0” Often used to model times to failure for reliability analyses.
Width of a Pile of Data	○ The width of a histogram. For a normal distribution this is generally considered to be 6 times the standard deviation.
X-Bar Chart	○ A control chart monitoring averages for continuous data.

Bibliography

Advanced Topics in Statistical Process Control – Donald J. Wheeler, SPC Press, Knoxville, TN, 1995

Applied Linear Regression Models, Neter, Wasserman & Kutner, Irwin, 1989.

Applied Linear Statistical Models, Neter, Wasserman & Kutner, Irwin, 1985.

Company-Wide Total Quality Control, Shigeru Mizuno, Asian Productivity Organization, ISBN 92-833-1100-0, 1989.

Corporate Financial Analysis - Decisions in a Global Environment, Diana R. Harrington, Irwin, 1993.

Creating Innovative Products Using Total Design, Stuart Pugh, Edited by Don Clausing and Ron Andrade, Addison-Wesely, ISBN 0-201-63485-6.

Dave Barry Does Japan, Dave Barry, Fawcett-Columbine, ISBN 0-449-90810-0, 1992.

Deming Management at Work, Mary Walton, G. P. Putnam's Sons, ISBN 0-399-13557-X, 1990.

Design and Management of Service Processes, Rohit Ramaswamy, Addison-Wesely, ISBN-0-201-63383-3

Economic Control of Quality of Manufactured Product - Walter A. Shewhart, ASQC Quality Press, ISBN 0-87389-076-0, 1980.

Elementary Survey Sampling, Scheaffer, Mendenhall & Ott, PWS-Kent Publishing Company, 1990.

Evolutionary Operation, Box & Draper, Wiley, 1969.

Failure Analysis of Electrical Components, H. S. Silvus, Jr., Nuclear Plant Journal, May-June, 1992.

Fundamental Concepts in the Design of Experiments, Charles R. Hicks, Holt-Rinehart & Winston, 1982.

Intermediate Economic Analysis for Management and Engineers, John R. Canada, Prentice-Hall, 1971.

Introduction to Quality Control, Kaoru Ishikawa, 3A Corporation, ISBN 4-906224-61-X C0034, 1990.

Juran's Quality Control Handbook, J. M. Juran, Editor-in-Chief, McGraw-Hill Publishing Company, ISBN 0-07-033176-6, 4th Edition, 1988.

Maintenance Engineering Handbook, Lindley R. Higgins, 5th Edition, McGraw-Hill, Inc., ISBN 0-07-028811-9.

Modern Microelectronic Circuit Design, IC Applications, Fabrication Technology, Volumes I & II, Dr. M. Fogiel, Research and Education Association, ISBN 0-87891-520-6.

Out of the Crisis, W. Edwards Deming, MIT Center for Advanced Engineering Study, ISBN 0-911379-01-0, 1991 (14th Printing).

Primer: Fracture Mechanics in the Nuclear Power Industry, EPRI-NP-5792-SR, Wessel, Server & Kennedy, Electric Power Research Institute, 1990.

Practical Reliability Engineering, Patrick D. O'Connor, John Wiley & Sons, ISBN 0-471-95767-4

Product Design for Manufacture and Assembly, Boothroyd, Dewhurst & Knight, Marcel-Dekker, ISBN 0-8247-9176-2

Quality Control and Industrial Statistics, Acheson J. Duncan, Richard D. Irwin, Inc., ISBN 0-256-03535-0, 1986 (5th Edition).

Quality Engineering Series Volume 1 - Taguchi Methods, Research and Development, G. Taguchi, S. Konishi, Y. Wu, American Supplier Institute, 1992.

Quality Engineering Series Volume 4 - Taguchi Methods, Design of Experiments, G. Taguchi, S. Konishi, Y. Wu, American Supplier Institute, 1993.

Quality Engineering Series Volume 6 - Taguchi Methods, Case Studies from the U.S. and Europe, G. Taguchi, S. Konishi, Y. Wu, American Supplier Institute, 1989.

Quality Function Deployment – How to Make QFD Work for You, Lou Cohen, Addison-Wesely, ISBN 0-201-63330-2.

Short Run SPC - – Donald J. Wheeler, SPC Press, Knoxville, TN, 1995

Simulation Modeling & Analysis, Law & Kelton, McGraw-Hill, ISBN 0-07-036698-5

Special Report: Overcoming Power Quality Problems, Dick Troberg, CEE News, October, 1995.

Statistical Methods for Medical Investigations, Brian S. Everitt, Halsted Press, 1994.

Statistical Methods for Quality Improvement, Hitoshi Kume, Association for Overseas Technical Scholarship, ISBN 4-906224-34-2 C0034, 1985 (1st published).

Statistical Quality Control Handbook – AT&T Technologies, Inc., copyright renewed 1984.

Taguchi Methods, Genichi Taguchi, ASI, ISBN 0-941243-16-8.

The Gales of November - The Sinking of the Edmund Fitzgerald, Robert J. Hemming, Contemporary Books, Inc., ISBN 0-8092-5384-4.

The Guidebook to Concrete Repair, Structural Preservation Systems, Inc.

The New Weibull Handbook, Dr. Robert B. Abernethy, 1993 (Available through the author, write to 536 Oyster Road, North Palm Beach, FL 33408).

The Vision of Six Sigma - A Roadmap for Breakthrough - Dr. Mikel Harry, Sigma Publishing Co., 4th Edition.

Tolerance Design – A Handbook for Developing Optimal Specifications, C. M. Creveling, Addison-Wesely, ISBN 0-201-63473-2

Understanding Statistical Process Control – Donald J. Wheeler and David S. Chambers, SPC Press, Knoxville, TN, 1992

Black Belt Manual Index

Index of Tools & Methods

Tool or Method	Unit(s)		Tool or Method	Unit(s)
5-S's	5.4		DOIT (Creativity Technique)	3.2
Additivity of Variances	6.3		Dynamic Design	14.2
Affinity Diagram	16.1		Economic Evaluation	12.2
Analysis of Means (ANOM)	6.9		Error Modes and Effects Analysis	15.2
Analysis of Variance (ANOVA)	10.3		Experimental Design	11.1
Arrow (PERT) Chart	16.1		Failure Modes and Effects Analysis	15.2
Assumption Bashing	3.2		Fault Tree Analysis	15.3
Attribute Listing	3.2		Financial Analysis	12.2
Auto-Correlated Data & Control Charts	6.9		Flowchart	5.3
Bar Chart	7.1		Focus Groups	4.1
Barriers and Aids Table	12.1		Forced Analogy	3.2
Benchmarking	14.3		Frequency Charts	6.3
Brainstorming	3.2		Gage R&R Study	6.7
Business Planning Process	2.5		Gantt Chart	3.1
C Control Chart	6.6		Histogram	6.3
Cause and Effect Diagram	8.5		Hypothesis Testing – Means	9.2
Central Tendency Metrics	6.3		Hypothesis Testing – Proportions/Rates	9.2
Check Points	6.1		Hypothesis Testing – Variance	9.2
Checklists	3.2		Idea Matrix	16.1
Checksheets	6.2		Imitation (Creativity Technique)	3.2
Cluster Sampling	9.3		Implementation Planning	12.1
Common Cause Variation	6.3		Improvement Process	2.2
Comparative Process Analysis	5.4		Inherent Capability Index	6.8
Consensus	3.2		Inner/Outer Array	14.4
Contingency Table	7.2		Interval Sampling	6.2, 9.3
Control Points	6.1		Interviews	4.1
Correlation Analysis	10.1		Kanban System	5.5
Cost Benefit Analysis	12.1, 12.2		Kano Model	4.1
Count (Discrete) Data	6.3		Key Characteristics	6.1
Countermeasure Matrix	12.1		Kurtosis	6.3
Countermeasure Matrix	12.1		Layout Diagram	5.3
Critical Pathway	5.4		Lean Manufacturing	5.5
CUSUM Chart	6.9		Line Graph	6.3
Data Collection	6.2		Linear Regression	10.2
Data Collection Principles	6.2		Logistic Regression	10.2
Design Process	2.4		Loss Function	14.4

Tool or Method	Unit(s)		Tool or Method	Unit(s)
Low-Hanging Fruit	5.4		Random Sampling	6.2, 9.3
Matrix Data Analysis	16.1		Rank Ordering	3.2
Measurement Data	6.3		Relations Diagram	16.1
Measurement System Analysis	6.7		Reliability Definitions	15.1
Meeting Process	3.1		Responsibility Flowchart	5.3
Muda Elimination	5.5		Retain, Capture, Develop Model	4.1
Multiple Linear Regression	10.2		Root Cause Analysis	8.5, 15.5
MultiVoting	3.2		Root Cause Verification Matrix	8.5
Np Control Chart	6.6		Run Chart	6.3
Operating Reviews	16.2		Scatter Diagrams	10.1
Operational Capability Index	6.8		Short Run Control Charts	6.9
Organization as a System	5.3		Sigma Capability Metrics	6.8
Orthogonal Arrays	11.1		Signal-to-Noise Ratio	14.4
P Control Chart	6.6		Six Hats of DeBono	3.2
Parametric Analysis	14.2		Six Universal Questions	3.2
Pareto Chart	7.2		Skewness	6.3
Pie Chart	7.1		Solution Selection Matrix	12.1
Plan-Do-Check-Act	5.3		Special Cause Variation	6.3
Probability Distributions	9.1		Sporadic Events Control Chart	6.9
Problem Reversal	3.2		Standardization	12.1
Process Capability Index	6.8		Static Design	14.2
Process Capability Study	6.8		Stratified Samples	9.3
Process Decision Program Chart	16.1		Structure Tree	16.1
Process Inventory Sheet	5.1		Surveys	4.1
Process Levels	5.1		System of Indicators	6.1
Process Management	2.3, 5.3		Taguchi Design Approach	14.4
Process Management Book	5.3		Team Reviews	16.2, 3.1
Process Management Charts	5.3		Team Roles	3.1
Process Simulation Software	5.2		Teams	3.1
Process Storyboard	5.3		Tolerance Deployment/Analysis	14.1
Process Watch	5.2		TRIZ	14.2
Process Yield	6.8		Twenty Questions	5.4
Product/Service Requirements Document	14.1		Two-Stage Sampling	9.3
Project Charter	3.1		U Control Chart	6.6
Pugh Concept Evaluation Matrix	14.2		Value Added/Non-Value Added Analysis	5.4
Quality Function Deployment (QFD)	14.1		Variability Metrics	6.3
Radar Chart	7.2		Variable Control Limits	6.9

Tool or Method	Unit(s)		Tool or Method	Unit(s)
Voice of Customer Listening Process	4.1			
Weibull Analysis	15.4			
X, mR Control Chart	6.5			
X-Bar, R Control Chart	6.5			
X-Bar, S Control Chart	6.5			