Quality Engineering Statistics

Mathews Malnar and Bailey, Inc.
171 Casement Avenue
Painesville, OH 44077
440-350-0911
www.mmbstatistical.com
paul@mmbstatistical.com

Quality Engineering Statistics

by Mathews Malnar and Bailey, Inc.

Table of Contents

This page is blank

Quality Management

Quality Environment and Culture

- The methods of quality engineering and applied statistics you will use are determined/affected by the environment or culture that you work in.
- The quality environment or culture is determined by one or more of:
  - Team
  - Department or Company
  - Industry or Technical Community
  - Standards Agency (e.g. ASQ, ANSI, ASTM, ISO, NIST, IEEE, ...)
  - Regulatory Environment (e.g. FDA, EPA, NRC, ...)
  - External Consultant
- References:
  - Defeo, *Juran's Quality Handbook*
  - Juran Trilogy: Quality Planning, Quality Control, Quality Improvement
  - Juran, *Quality Planning and Analysis*
  - Juran, *Quality by Design*

Quality Engineering Tools

- All quality environments/cultures share a common tool set with their own individual specialized tools.
- Total Quality Management (TQM) is a loosely defined collection of well known quality tools and techniques.

  Pareto Chart
  Cause and Effect Diagram
  Histograms
  Scatter Diagrams
  Run Charts
  Control Charts
  Flow Charts
  ...
- Reference: Tague, *The Quality Toolbox*

Quality Management Systems

- TQM lacks a clear procedure for using the tools, consequently organizations with weak quality management systems often fail.
- Quality Management Systems (QMS) provide the structure/procedures for using the tools.
  - ISO 9000
  - TS 16949 (formerly QS 9000)
  - AS 9100
  - ISO 17025
  - ISO 14000
  - ISO 13485
  - Six Sigma
  - Lean
  - Lean Six Sigma

ISO9000 Requirements

1. Customer Focus
2. Engagement
3. Leadership
4. Process
5. Evidence-based Decision Making
6. Relationship Management
7. Continuous Improvement

ISO TS 16949

- ISO TS 16949 (formerly QS 9000) is a hybrid quality management system based on ISO9000.
- It was designed for and by automotive manufacturers (USA) who recognized the value of ISO 9000 but saw the need to add additional methods required by the industry.
- Was originally administrated by the Automotive Industry Action Group (AIAG), later by ISO, and today by IATF

TS 16949 Requirements

- ISO 9001
- PPAP - Production Part Approval Process
- APQP - Advanced Product Quality Planning
- FMEA - Failure Modes and Effects Analysis
- SPC - Statistical Process Control
- MSA - Measurement Systems Analysis
- QSA - Quality Systems Assessment (i.e. quality auditing)
- Reference: QS 9000 Supplier 7 Pack

Six Sigma

- Conceived by Motorola in the late 1980s
- Adopted by Jack Bossidy of Allied Signal
- Made popular by GE under Jack Welch in the late 1990s
- Integrated with Lean to make Lean Six Sigma
- Standardized in
  - ISO 13053
  - ASQ Certification

Six Sigma

- DMAIC: Used to improve existing processes
- DMADOV: Used to develop new products and processes

The Six Sigma DMAIC Process

DMAIC - the Six Sigma process used to improve existing processes:

1. **Define**: Define the problem in terms of relevant customer CTQs.
2. **Measure**: Measure actual CTQ performance and compare to goal.
3. **Analyze**: Experiment to 1) identify the KPIVs among the many PIVs that could affect the CTQs and 2) quantify their effect on the CTQs.
4. **Improve**: Leverage your understanding of the relationship between the KPIVs and the CTQs to improve the process.
5. **Control**: Put controls in place to make the improvements permanent.

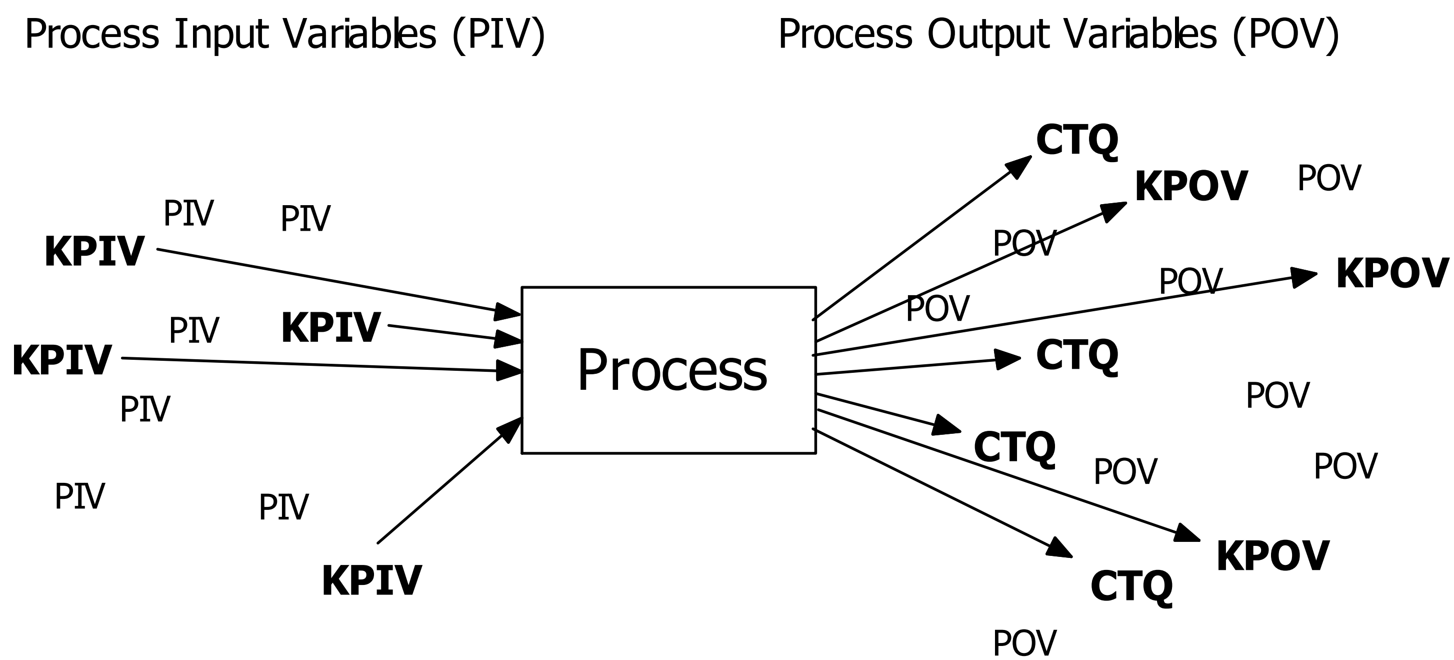
The Six Sigma DMADOV Process

DMADOV - the Six Sigma process used to develop new products and processes:

1. **Define**: **D**efine design goals that meet customer requirements
2. **Measure**: Identify and measure the CTQs
3. **Analyze**: Develop design concepts
4. **Design**: Choose a design concept
5. **Optimize**: Optimize the design
6. **Verify**: Verify the design, its implementation, and that it meets the customer's wants and needs

Six Sigma Process Improvement

- Determine which input variables are the most important.
- Quantify the relationships between the y and the x_i.
- Build a mathematical model for $y = f(x_1, x_2, \cdots)$.
- Set tolerances on y to meet the customer's requirements.
- Set and hold tolerances on x to guarantee the y.



KPIV = Key Process Input Variable

CTQ = Critical To Quality
KPOV = Key Process Output Variable

Lean

- Lean came from the Toyota Production System (TPS)
- Lean is focused on eliminating waste: Anything that adds cost to the process without adding value.
- References:
  - Womack and Jones, *The Machine That Changed the World*
  - Womack and Jones, *Lean Thinking*
  - Hammer and Champy, *Reengineering the Corporation*
- "Simplify, simplify." - Emerson
- "Simplify." - Emerson after Lean

Lean

Lean reduces waste in manufacturing, design, and administration:

- Overproduction, e.g. making more than needed, uncompleted product launches, excessive reporting
- Waiting, e.g. people or product waiting, waiting for approvals or data
- Transportation, e.g. moving work product or activities to another site
- Inventory, e.g. excess or incorrect product, outdated designs, backlogs
- Overprocessing, e.g. unnecessary process steps, approval routings, or analysis
- Duplication, e.g. repeating the same work at different steps in the process
- Unclear communication
- Motion, e.g. of people, obtaining forms, processing paperwork
- Defects, e.g. making defects, fixing defects, incorrect drawings or data, missing data
- Lost opportunities

The Lean Process

1. Define customer value (what will he/she pay for?)
2. Study the value stream
3. Make value flow
4. Establish a "pull" based system, e.g. kanban.
5. Improve to perfection (Kaizen)

The Integration of Lean and Six Sigma

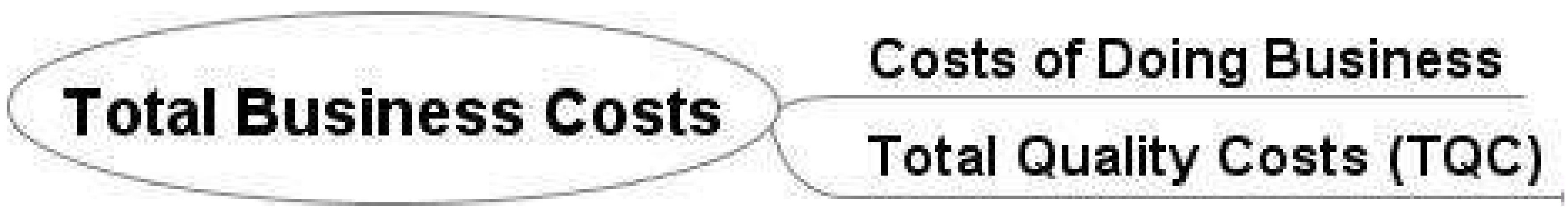
- Lean and Six Sigma are both improvement oriented.
- Lean involves reorganizing existing processes.
- Six Sigma is oriented towards making incremental improvements to existing processes.
- Both methods require continuous improvement.
- Lean and Six Sigma projects both use the DMAIC method.

Quality Cost

Why Are You Here #1?

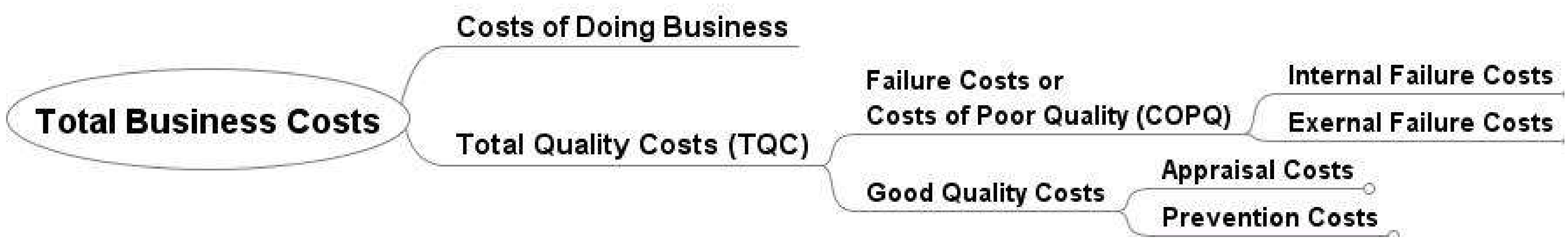
Total Business Costs

- Total business costs can be broken into two categories: costs of doing business and quality costs.
- Assign each cost item to one or the other category by considering the answer to the question, "If our product and/or process quality *were perfect*, would this cost item vanish?"
  - If the answer is "no", then the cost item is a Cost of Doing Business
  - If the answer is "yes," then the cost item is a Quality Cost.

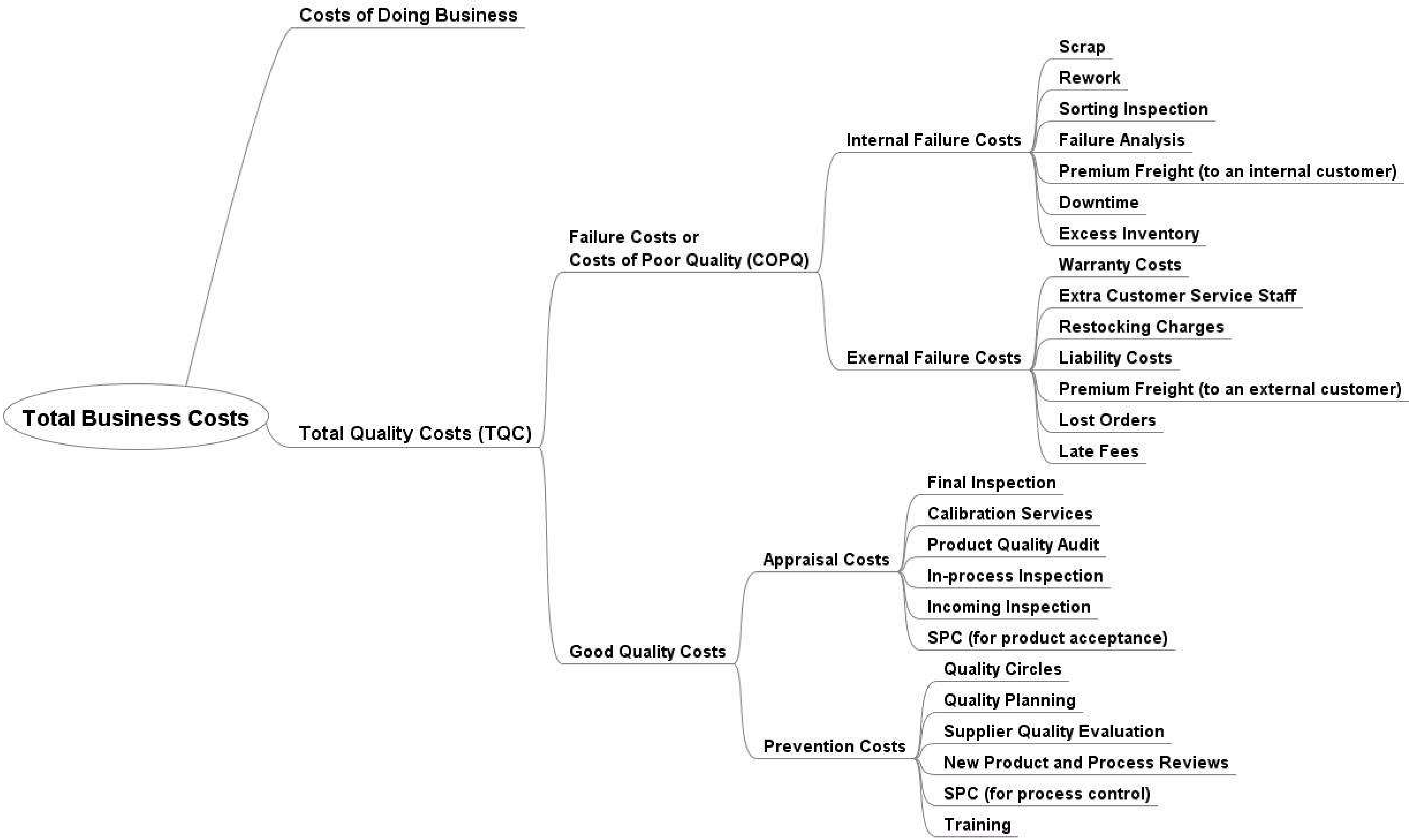


Quality Cost Categories

Quality costs can be further subdivided into good and bad cost categories:



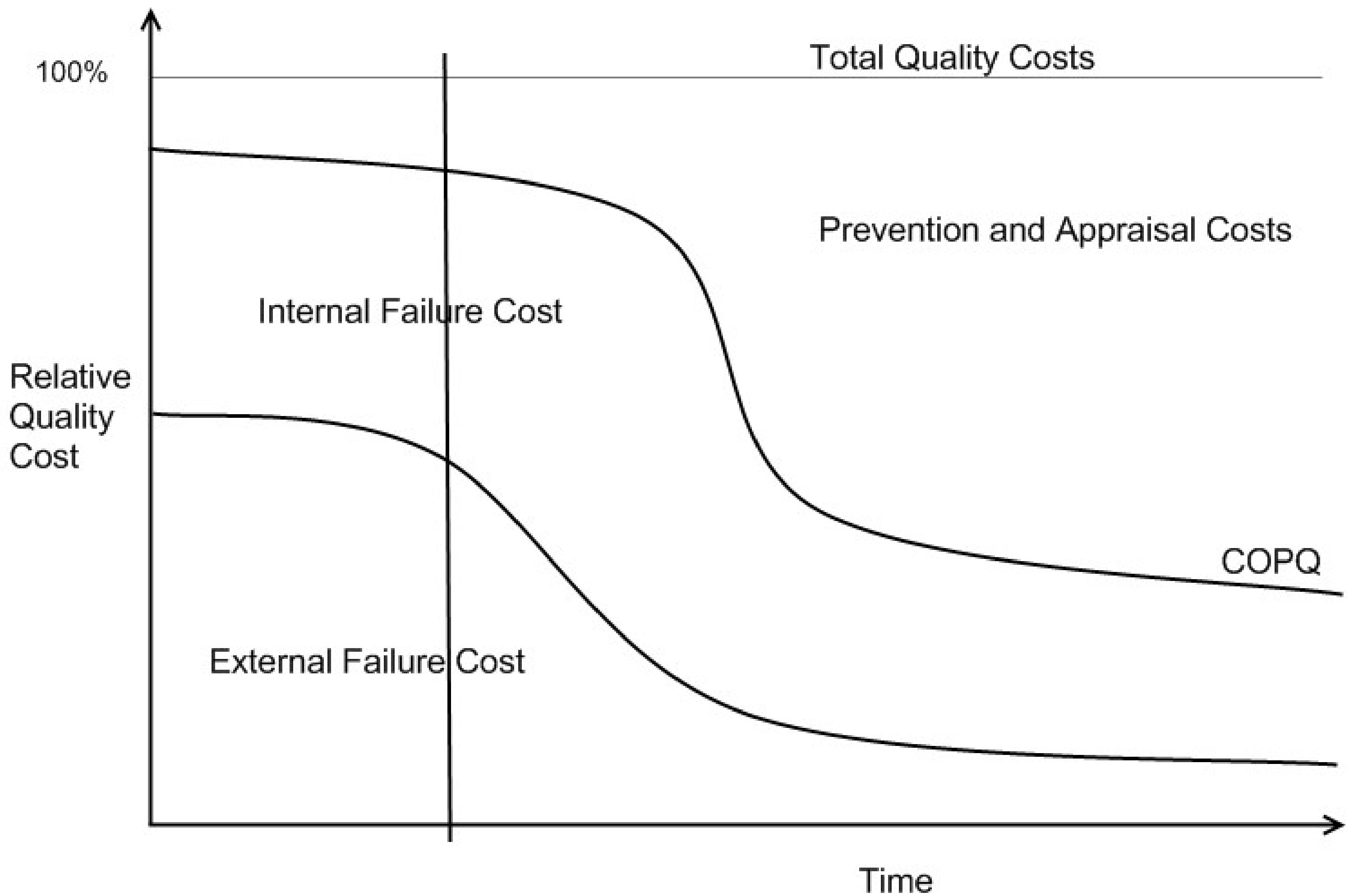
Quality Cost Details



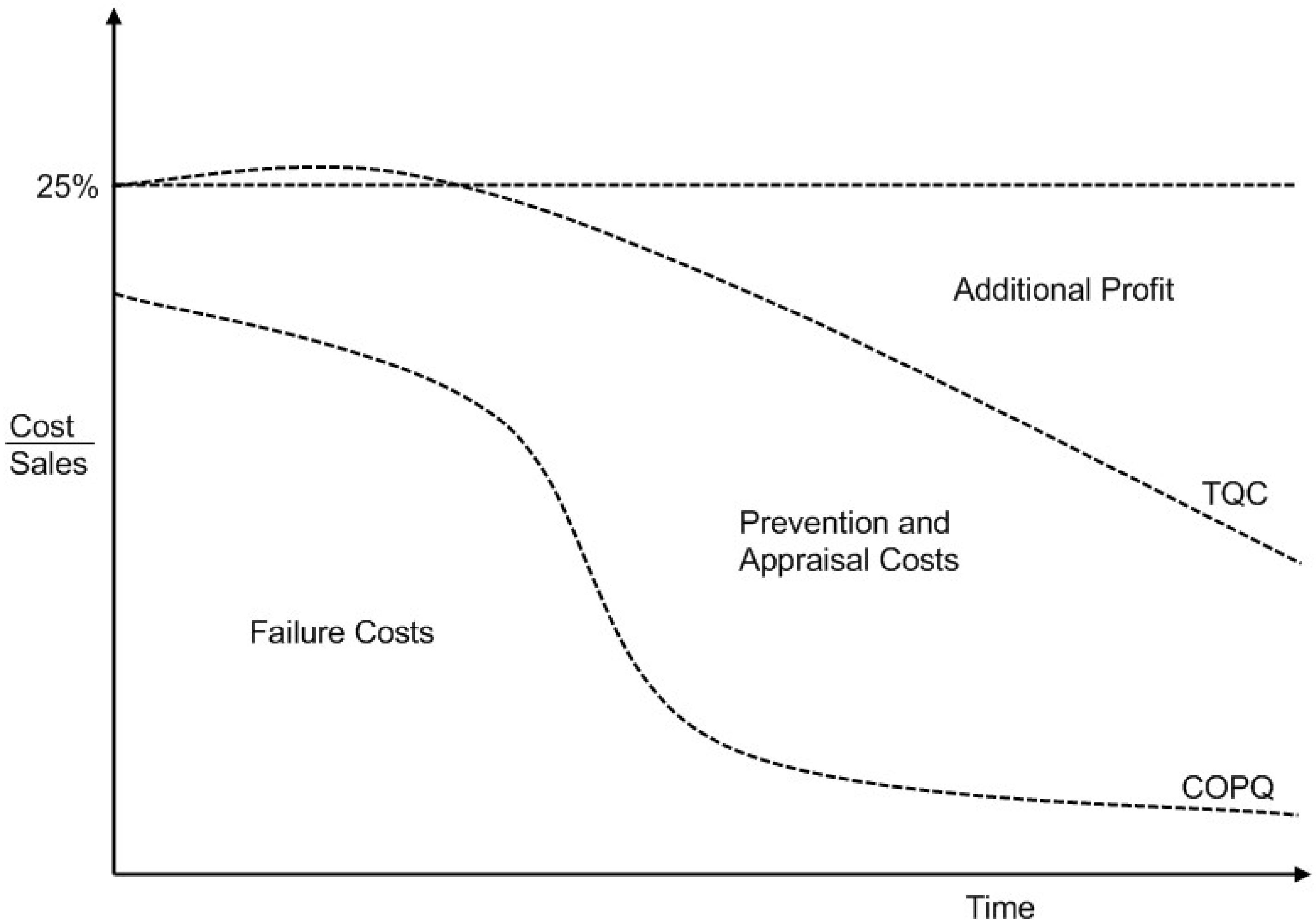
Cost of Quality Reports

- Costs are always reported in dollars.
- Costs are reported as percentages of:
  - Direct labor hour
  - Direct labor dollars
  - Standard manufacturing cost dollars
  - Value-added dollars
  - Sales dollars
  - Product units

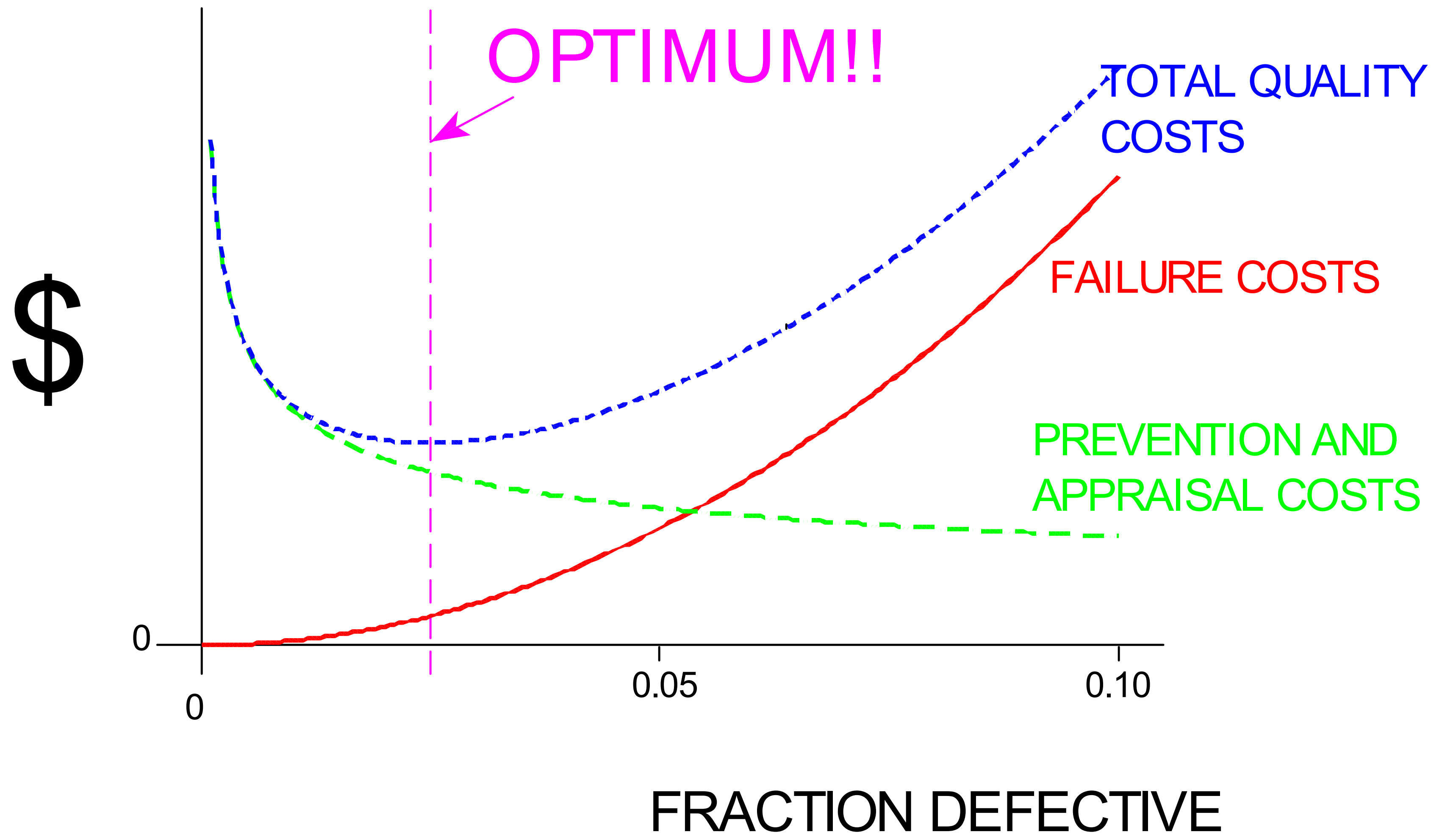
The Quality Cost Breakdown Changes During A Quality Improvement Program



Failure Costs Fall Faster Than Prevention And Appraisal Costs Increase



Optimum Quality Costs



Why Are You Here #2?

- How many of you have to deal with specifications in your jobs, either setting them or working to them?
- How many credit hours was the class that you had in college on how to set specifications?
- How many hours of class time did you spend talking about how to set specifications?
- What training have you had on how to set specifications?
- So why are we so surprised that we have difficulty managing specifications?

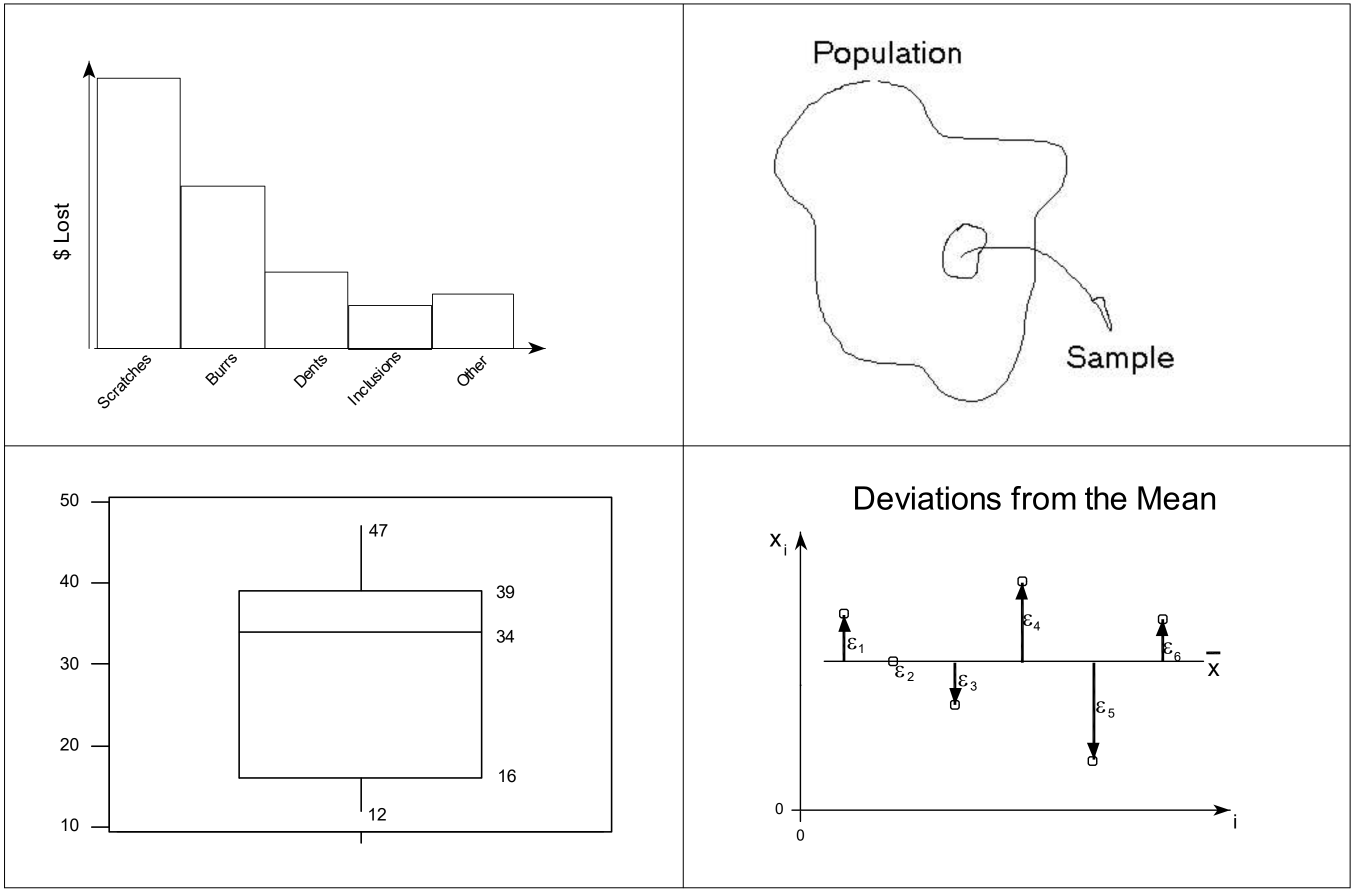
Strategies for Setting Specifications

- **CTQs and KPOVs**: Use voice of the customer analysis to determine what's important to the customer and set specifications to keep them happy.
- **KPIVs**: Identify and set specifications on the KPIVs that guarantee that the CTQs and KPOVs will meet the customer's requirements.
- **Ordinary POVs and PIVs**: Set specifications on ordinary POVs and PIVs to keep from getting into trouble, i.e. to the safe historical extremes.

References

- https://en.wikipedia.org/wiki/Quality_costs
- Campanella, *Principles of Quality Cost*
- Feigenbaum, *Total Quality Control*
- Crosby, *Quality is Free*
- https://en.wikipedia.org/wiki/Taguchi_loss_function

Introduction to Statistics

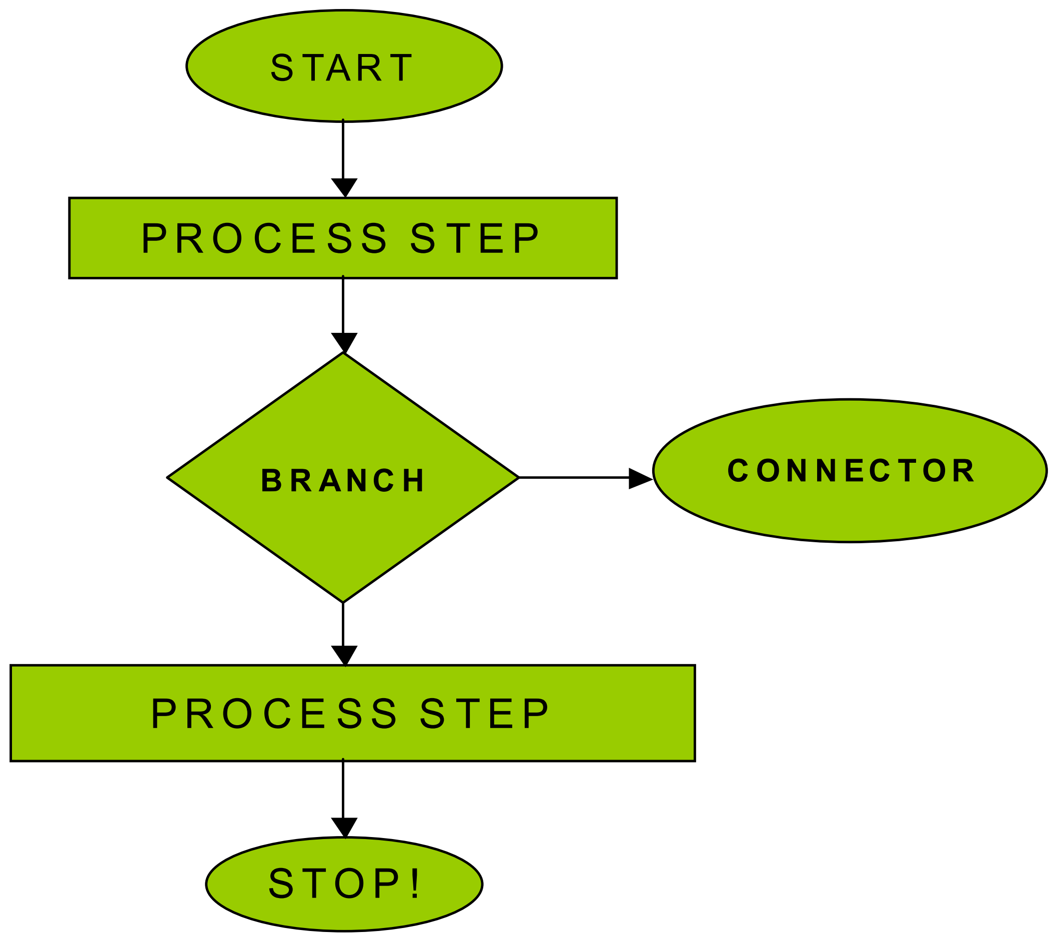


A Simple Process

Process - An activity that acts on or transforms an input to produce an output.

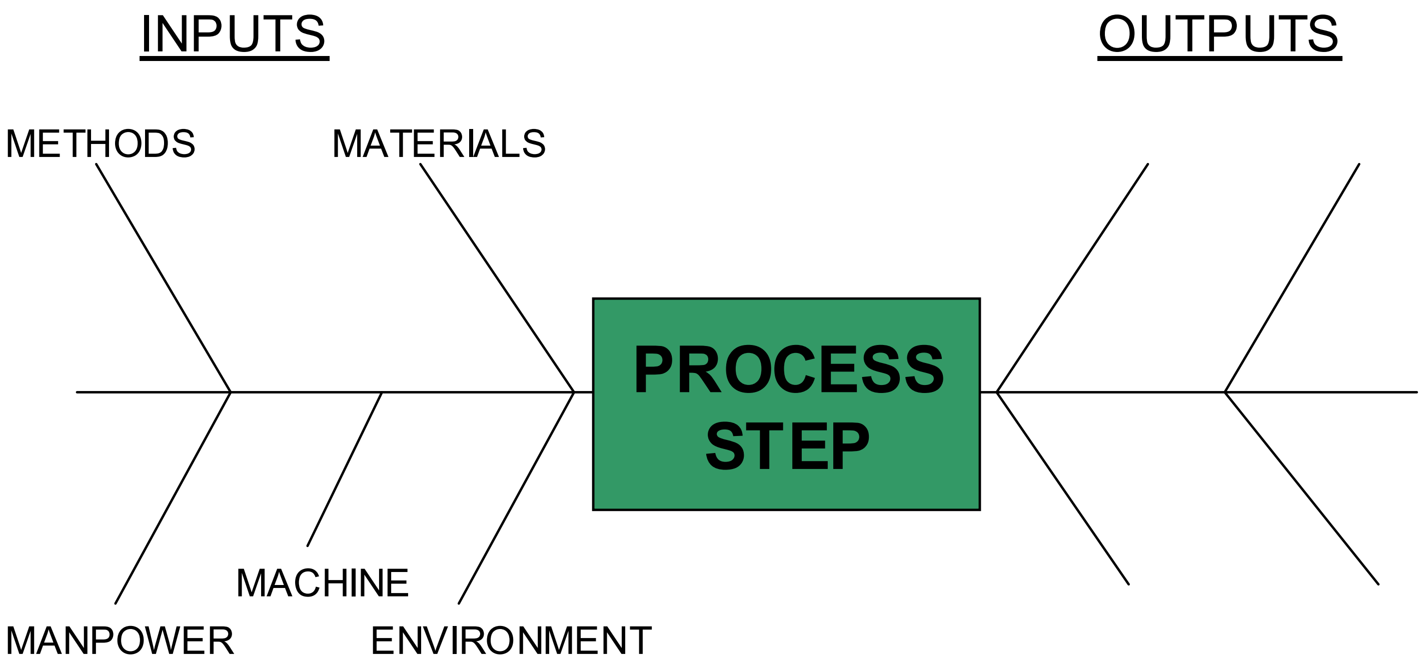
Flow Charts

Complex processes are documented in flow charts or process maps.



Input-Process-Output Diagrams

Each step in a process can be complex.

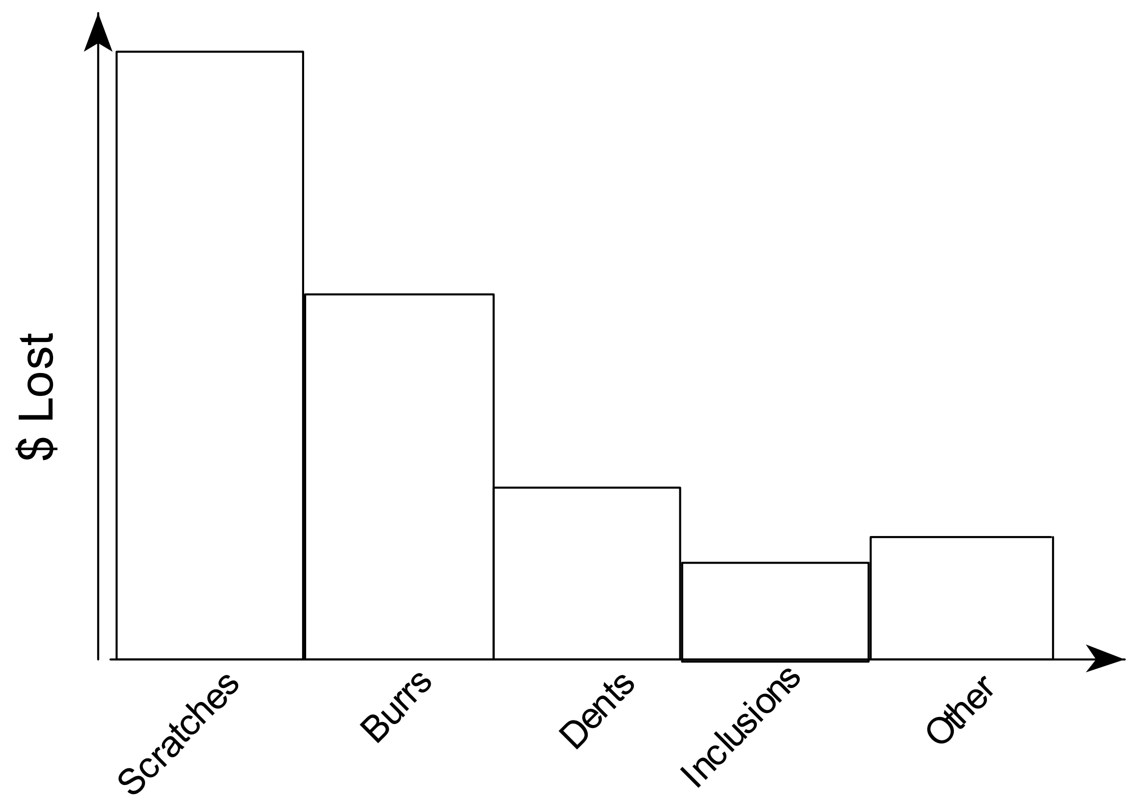


"The novice sees many possibilities. The expert sees few." - Shunryu Suzuki

Pareto Principle

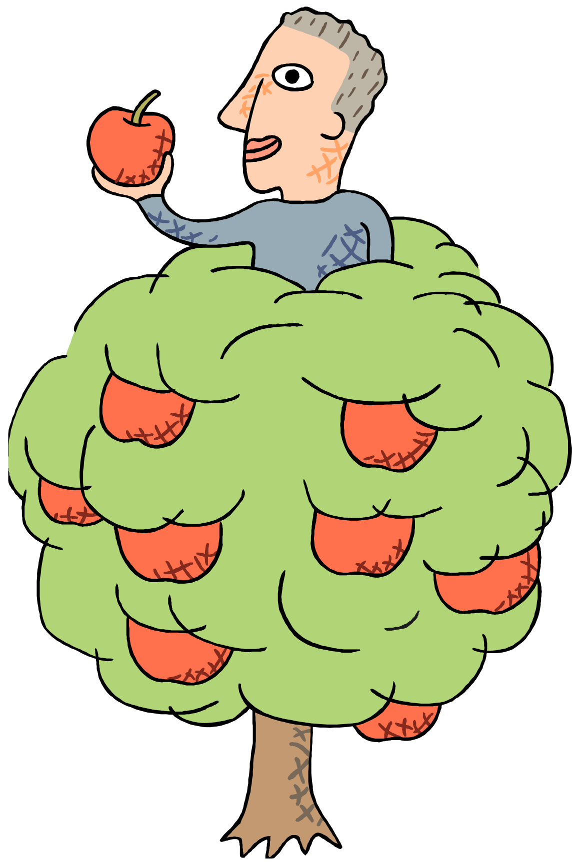
Pareto Principle

- A few factors tend to be much more important then the myriad others.
- 80% of the problems are caused by 20% of the causes.



Low-hanging Fruit

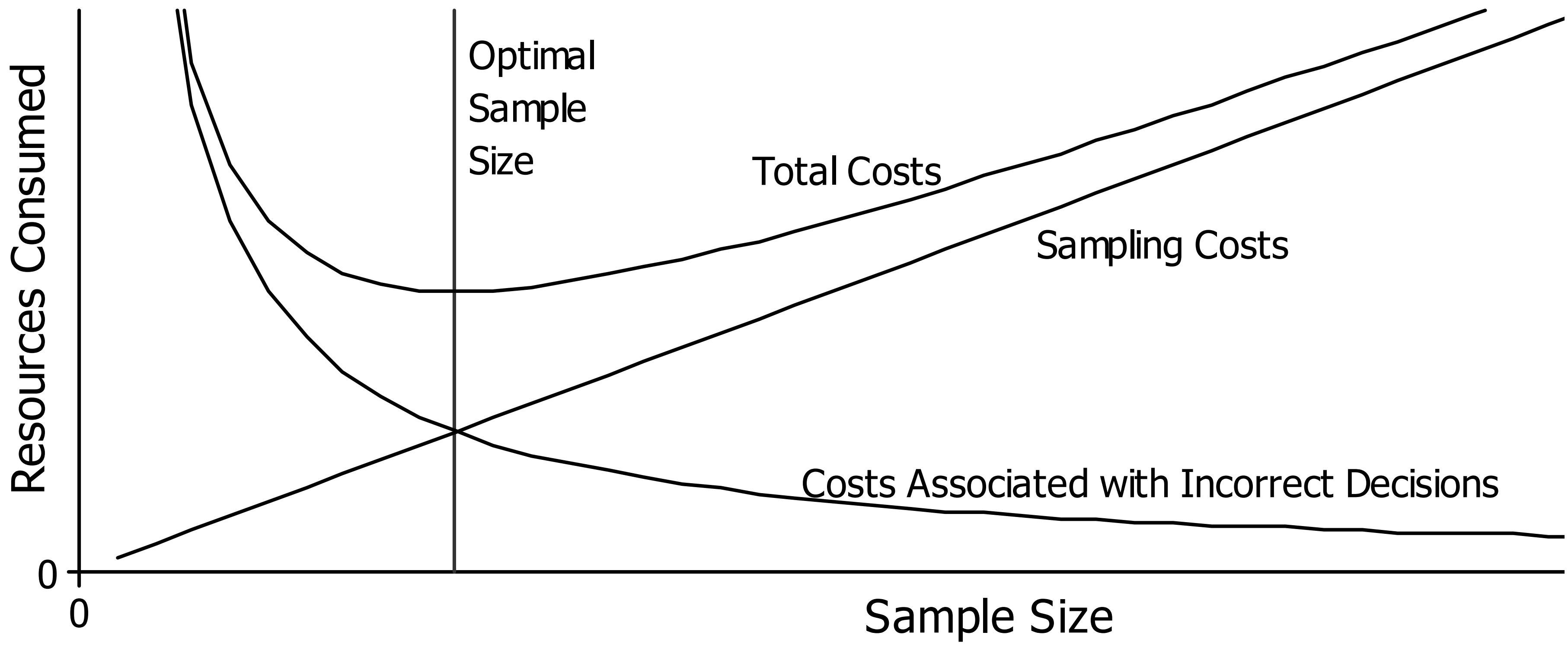
Also consider picking the low-hanging fruit first, that is, those projects that are easy to do.



Economic Need for Statistical Methods

- The need for statistical methods is driven by economic considerations.
- You must measure all of the individuals in a population to determine its parameters.
- Statistics provide practical estimates of these parameters.
- Better estimates are less risky but require larger sample sizes.

Economic Need for Statistical Methods



The Value of Data

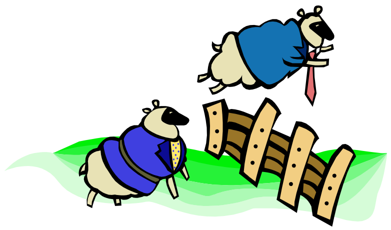
- Data are your sword and your shield.

- The person with the most, or the best, but usually the only data wins.
- Show me the data! Without data you only have an opinion.
- Always take the side of the data. The data will win.
- When in doubt, take more data!
- In God we trust. All others bring data.
- Theory plus supporting data trumps either theory or data alone.
- "The object of collecting data is to provide a basis for action." - Deming

Types of Data

In quality engineering we generall deal with two types of data:

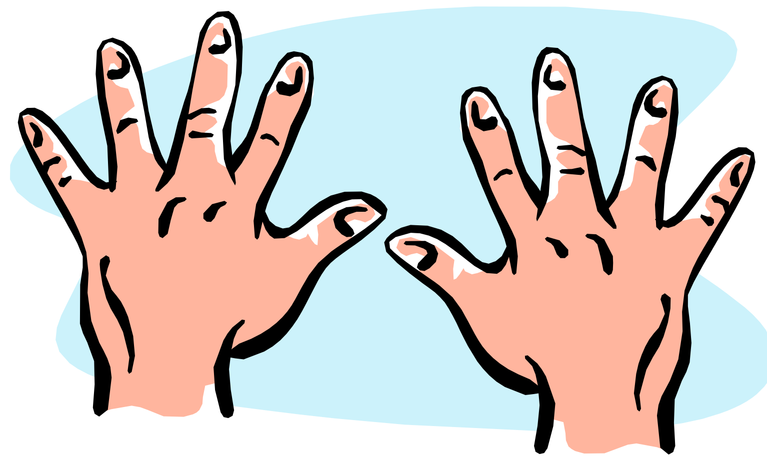
- Attribute data, also called count or qualitative data



- Variables data, also called measurement or quantitative data

Attribute Data

- Attribute data are used to count the number of times an event occurs, such as:
  - The number of defective (or nonconforming) parts found in a given sample of parts.
  - The number of part defects (or nonconformities) found on a given part or collection of parts.
- Attribute data are inherently discrete values and can take on only the counting numbers 0, 1, 2, 3, ...



Variables Data

- Variables data are used to determine the size of things, such as: length, area, volume, temperature, mass, time, angle, …
- Variables data are inherently continuous values.
- Variables data can contain as many significant digits as provided by the measuring instrument.

Measurement Scales

- The classification of data into attribute and variable types is oversimplified but often sufficient.
- A more detailed description of measurement types:
  - Observations made on a **nominal** scale are sorted into two or more qualitative categories. Nominal scales with two categories are also referred to as **binary** or **dichotomous** scales.
  - Observations made on an **ordinal** scale can be ordered by size, i.e., from smallest to largest.
  - Observations made on an **interval** scale are assigned values that can be ordered by size but also include information about how far apart they are, i.e. a measurement scale with an arbitrary origin and a fixed unit of measure.
  - Observations made on a **ratio** scale have an interval scale with meaningful values of the ratios of pairs of observations, e.g. *this observation is twice as big as that one*. Ratio scales have a natural (i.e. non-arbitrary) origin.

Stem and Leaf Plots

68	75	93	82
82	94	87	38
73	81	88	97
76	84	92	90
100	61	65	100

What is the smallest value?

What is the largest value?

What's a good estimate for the average value?

How much variability is there in the data?

Are there any outliers?

The Stem and Leaf Plot

We need a better way of viewing the data. One method is to use a stem-and-leaf plot:

3	8					
4						
5						
6	8	1	5			
7	5	3	6			
8	2	2	7	1	8	4
9	3	4	5	2	0	
10	0	0				

Visualizing Data:
The Stem and Leaf Plot

- The advantages of the stem-and-leaf plot are:
  - The original data values are preserved.
  - The final plot gives a histogram of the data.
  - Recording data in this form as they are taken helps prevent mistakes.
- The primary disadvantage of the stem-and-leaf plot is that the order in which the data are taken is lost.

Visualizing Data: The Stem and Leaf Plot

Problem: What were the original data values that were used to construct the following stem and leaf plots?

3.6	8					
3.7	5	3				
3.8	3	7	7	2	1	
3.9	0	7				
4.0	3	8	9	0	2	4
4.1	9	6	2			
4.2	6					

The Stem and Leaf Plot

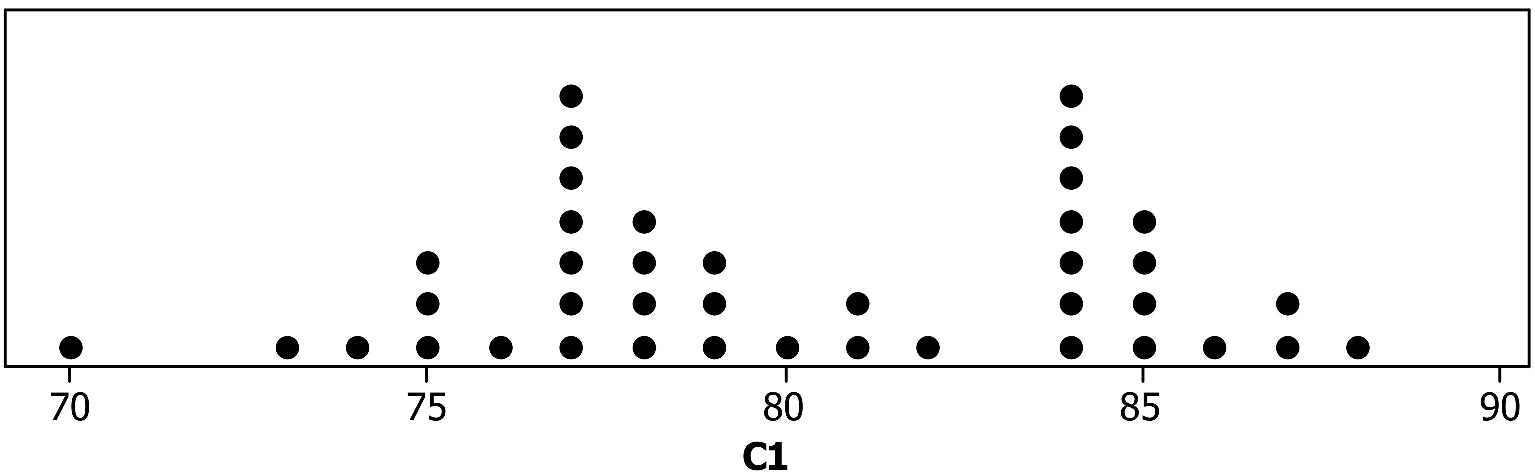
65-	4	1		
65+	5	9	6	
66-	1	2		
66+	8	7	8	6
67-	4			

The Stem and Leaf Plot

4a	1			
4b	2	2	3	
4c	5			
4d	7	6		
4e	9	9	8	9
5	4			
6				
7	5			

Dotplots

Dotplots are useful for graphical representations of small to moderate size data sets. Each dot usually corresponds to one observation but for larger samples each dot may represent more than one observation.



Visualizing Data: Histograms

Histograms are constructed using the same process as was used for stem-and-leaf plots:

The Process:

1. Choose the classes (interval or categories).
2. Sort or tally the data into these classes.
3. Count the number of items per class, i.e. the class frequency.
4. Plot the class frequencies against the classes.

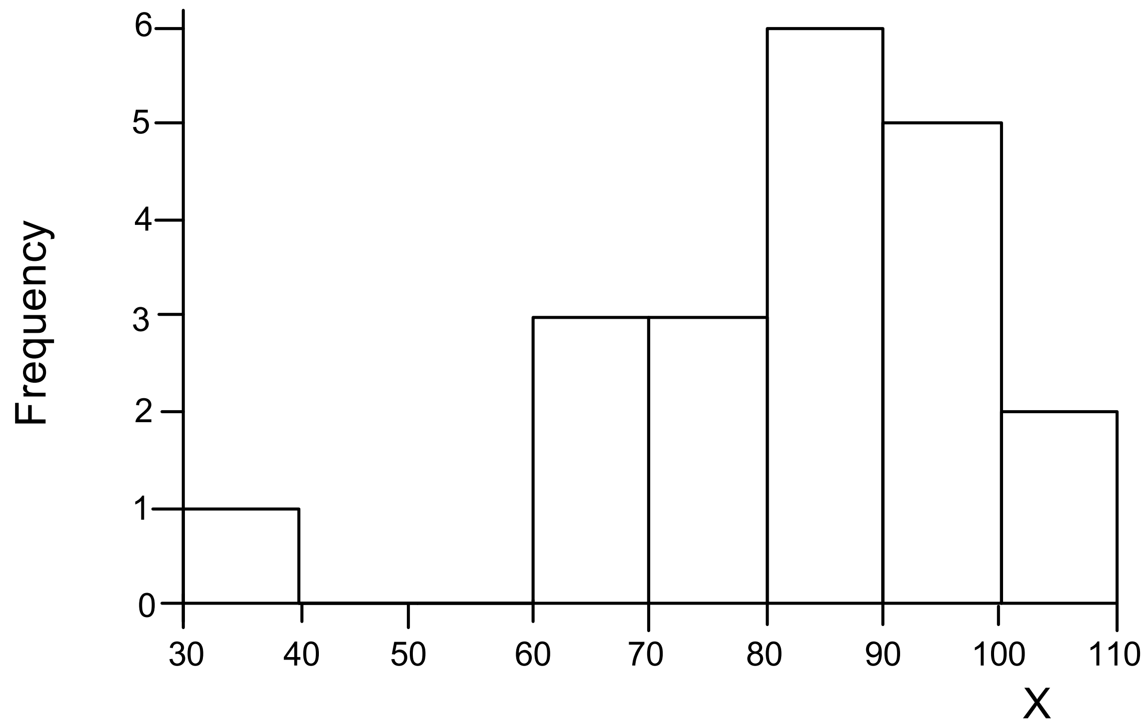
Rules for class selection:

- The number of classes should fall between 6 and 15.
- The exact number of classes depends on how many measurements or observations have been taken.
- Each value will go into one and only one class.
- Smallest and largest values must fall within the classifications.
- Classes can not overlap.
- Values can not fall into gaps between successive classes.
- Make sure successive classes have no values in common.
- Classes should cover equal ranges of values.

Visualizing Data: Histograms

Example: Construct a histogram for the data {68, 75, 93, 82, 82, 94, 87, 38, 73, 81, 88, 97, 76, 84, 92, 90, 100, 61, 65, 100}

Solution:



Visualizing Data

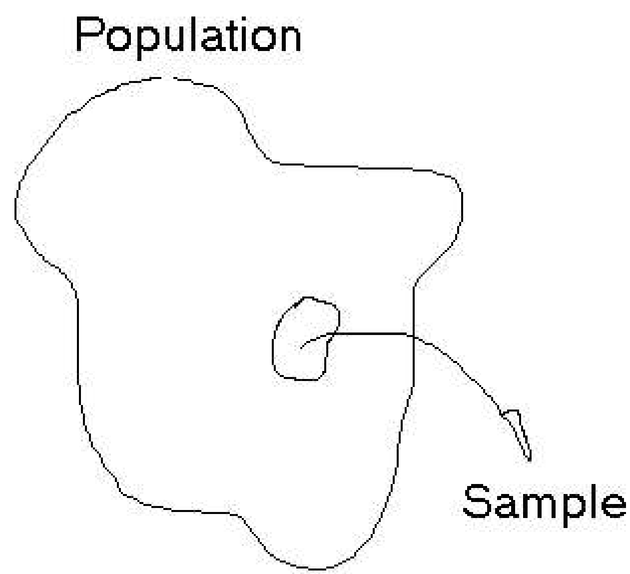
When we look at a stem-and-leaf plot or histogram the characteristics of the data set that we are interested in are:

1. Location (or Central Tendency)
2. Variation (or Dispersion, Scatter, Noise, ...)
3. Shape

We need to identify quantitative measures of location and variation.

Populations and Samples

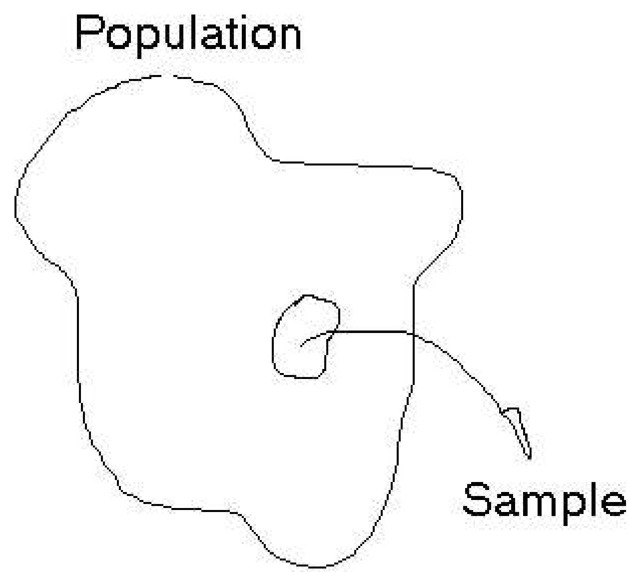
- A population is the collection of all possible values of a particular measurement.
- A sample is a subset taken from a population.



- Does the picture show a good way to take the sample?

Parameters and Statistics

- Parameters are numeric measures of the location or variation of a population.
- Statistics are numeric measures of the location or variation of a sample.
- Statistics determined from samples are used to estimate the parameters of a population.
- We want to know the parameters of a population, but we can only afford to estimate them with the statistics determined from samples.



Sampling Methods

- For a statistic to be a good estimate of a parameter, the sample must be representative of its population.
- The method of drawing a sample determines how well it represents its population.
- The specific sampling method used depends on patterns of variation within the population.
- The most commonly used methods of sampling in engineering/manufacturing statistics are random and periodic sampling.

Sampling Methods

- Acceptable methods of drawing samples are:
  - Random sampling - units are drawn at random from the entire population so that each unit has an equal chance of being chosen for the sample.
  - Periodic sampling - units are drawn for the sample in evenly spaced units of time, order, or space.
  - Stratified sampling - units are drawn from homogeneous subsets or *strata* within the population.
    - The sample may be drawn by random sampling from each strata a number of units in proportion to the relative size of the strata in the population.
    - If the same number of units are drawn from each strata but the strata are different in size, then the units may be weighted based on the relative size of the strata in the population.
  - Cluster sampling - units are drawn from randomly chosen clusters within the population.

Measures of Location: The Mean

- The mean of a sample ($\bar{x}$) is given by:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

where the subscripts 1, 2, 3, ..., n are used to identify the individual data values and n is the sample size.

- We need a shorthand notation for the summation operation:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- The mean of a population (μ) is given by:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i$$

where N is the population size.

- n and $\bar{x}$ are statistics. μ and N are parameters.

Example: For 8 months a police department reported 15, 16, 10, 9, 18, 16, 8, and 16 car thefts. Find the mean.

Solution: The total for the 8 months is

$$\begin{aligned}\sum_{i=1}^{8} x_i &= 15 + 16 + 10 + 9 + 18 + 16 + 8 + 16 \\ &= 108.0\end{aligned}$$

therefore the mean is $\bar{x} = \frac{108}{8} = 13.5$.

Example: A clerk lost one of 10 sales slips. The mean value of all 10 slips was \$7.20 and the remaining nine slips had the values: \$4.80, \$7.10, \$7.90, \$9.55, \$4.45, \$5.72, \$7.54, \$8.34, \$9.70. What is the value of the lost slip?

Solution: The total of the 10 slips must be

$$\sum_{i=1}^{10} x_i = n\bar{x} = 10 \times 7.20 = 72.00$$

The total of the nine slips is

$$\sum_{i=1}^{9} x_i = 4.80 + 7.10 + 7.90 + 9.55 + 4.45 + 5.72 + 7.54 + 8.34 + 9.70$$

$$= 65.10$$

The value of the missing slip is:

$$x_{10} = \$72.00 - \$65.10 = \$6.90$$

Measures of Location: The Median

- The median ($\tilde{x}$) is the middle value of a data set when the set is ordered from the smallest to the largest value.
- The median position is given by $\frac{n+1}{2}$
- The median of a data set with an odd number of points is a point of the data set.
- The median of a data set with an even number of points falls between two points of the data set and is determined from the average of the two points.
- The median is a noisier measure of location than the mean.

Example: Find the median of the following data set:

5	2	8			
6	3	5	5	6	9
7	2	4	5	7	7
8	3	5	7	8	
9	0	1	4		
10	1				

Solution: The median position will be:

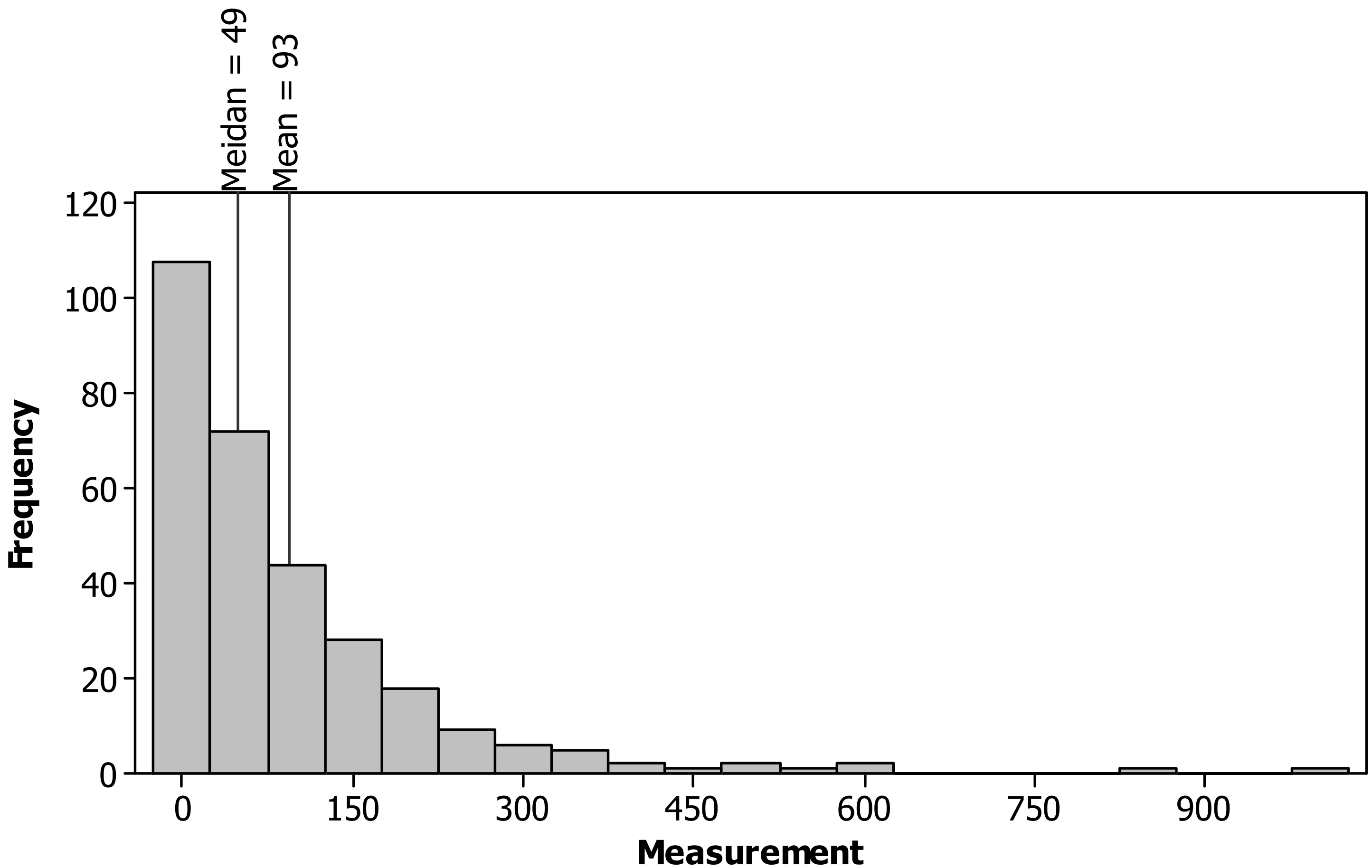
$$\frac{n+1}{2} = \frac{20+1}{2} = 10.5$$

The median is the average of the 10th and 11th values:

$$\tilde{x} = \frac{75+77}{2} = 76.0$$

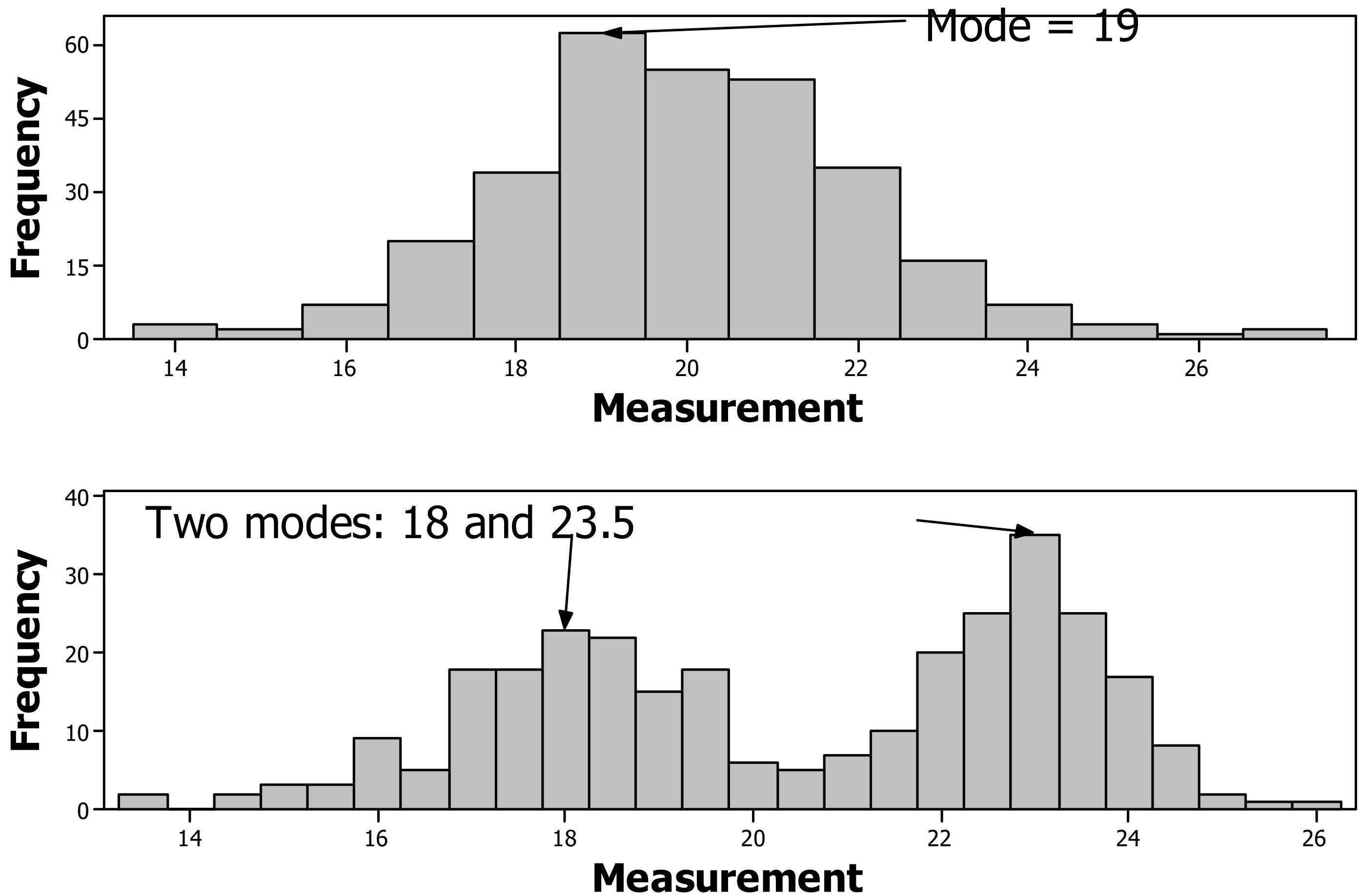
Mean versus Median

For highly skewed (asymmetric) data sets, the median is a better indicator of location than the mean.



Mode

The mode is the measurement value that has the highest frequency. A single, stable population is usually unimodal. Mixtures of two or more populations are described as bimodal or multi-modal, respectively.



Quartiles

- The median divides an ordered data set into two halves containing the same number of observations.
- Quartiles are the medians of half of the data values taken at a time.
- The lower quartile (Q_1) is the median of all values less than the median.
- The upper quartile (Q_3) is the median of all values greater than the median.
- $\tilde{x} = Q_2$
- The median and the two quartiles divide the data into four sets of points containing equal numbers of points.

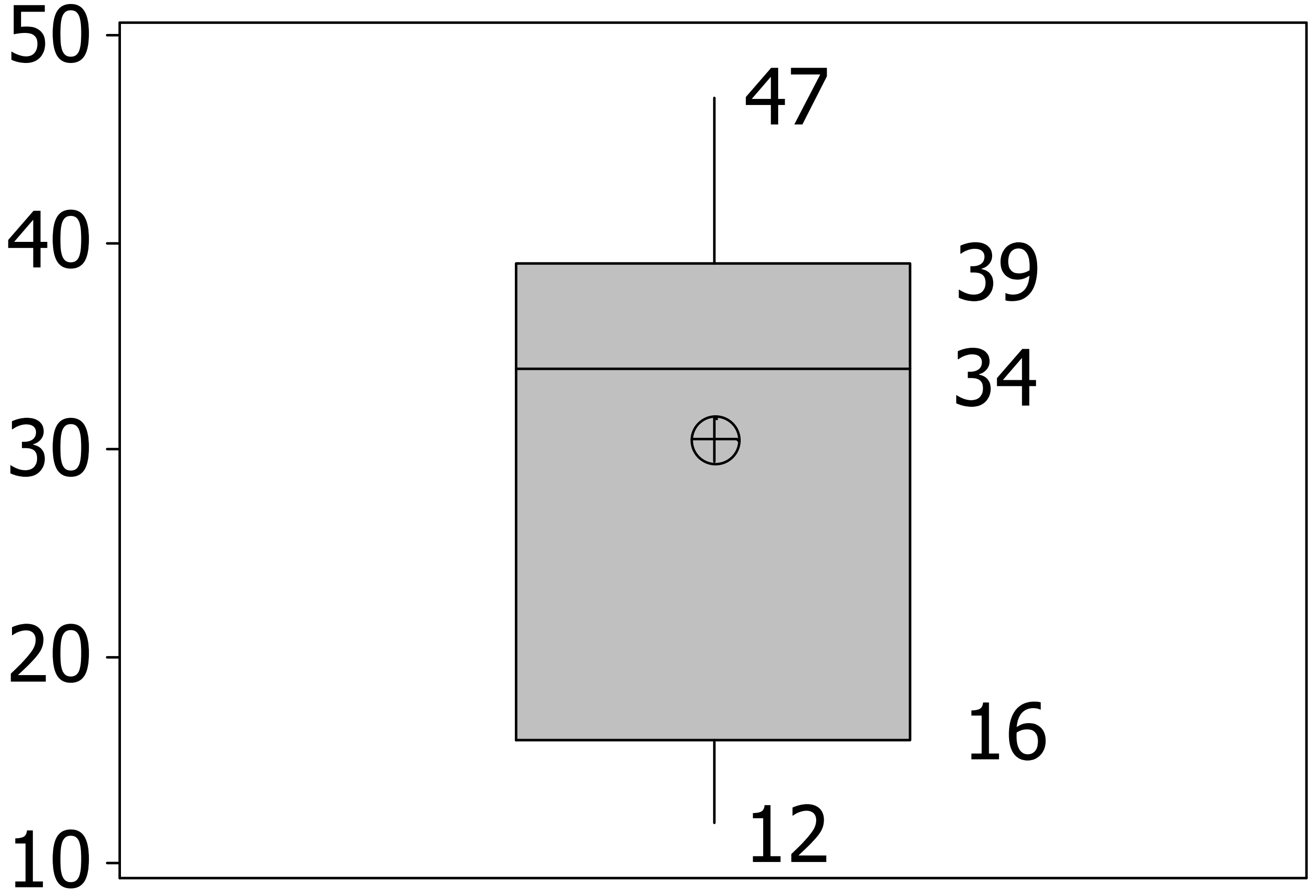
Box and Whisker Plots

Box and whisker plots are a very important application of quartiles.

Example: Construct the box and whisker plot for the following data set:

1	2	2	6		
2	5				
3	3	4	4	8	9
4	6	7			

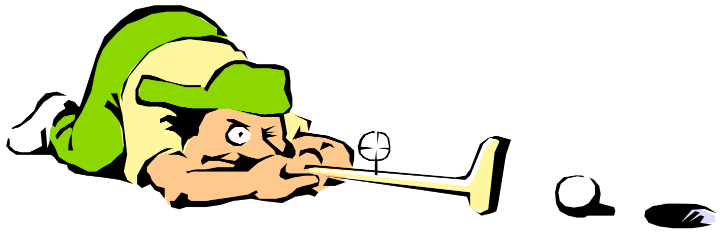
Solution: We have $Q_1 = 16$, $Q_3 = 39$, $\tilde{x} = 34$, $x_{\min} = 12$, and $x_{\max} = 47$.



On Average The Goose Is Dead?

Measures of Location Provide Incomplete Information

- Putts that are, on average, just the right length.



- Spinnaker takedowns that are, on average, at just the right time.
- Parts that are, on average, just the right size.
- A service that is, on average, delivered at just the right time.

All of these activities suffer severe consequences because of variation, i.e. location does not tell the whole story.

Measures of Variation: The Range

The simplest measure of variation in a data set is the Range (R):

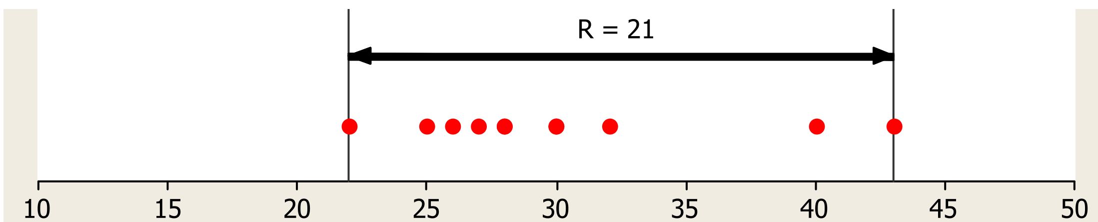
$$R = x_{\max} - x_{\min}$$

where $x_{\max}$ and $x_{\min}$ are the largest and smallest values in the data set. By definition the range is always positive.

Example: Find the range of the following data set (28, 32, 26, 40, 28, 27, 22, 30, 43, 25, 30)

Solution: **We have** $x_{\max} = 43$ and $x_{\min} = 22$, therefore:

$$R = 43 - 22 = 21$$

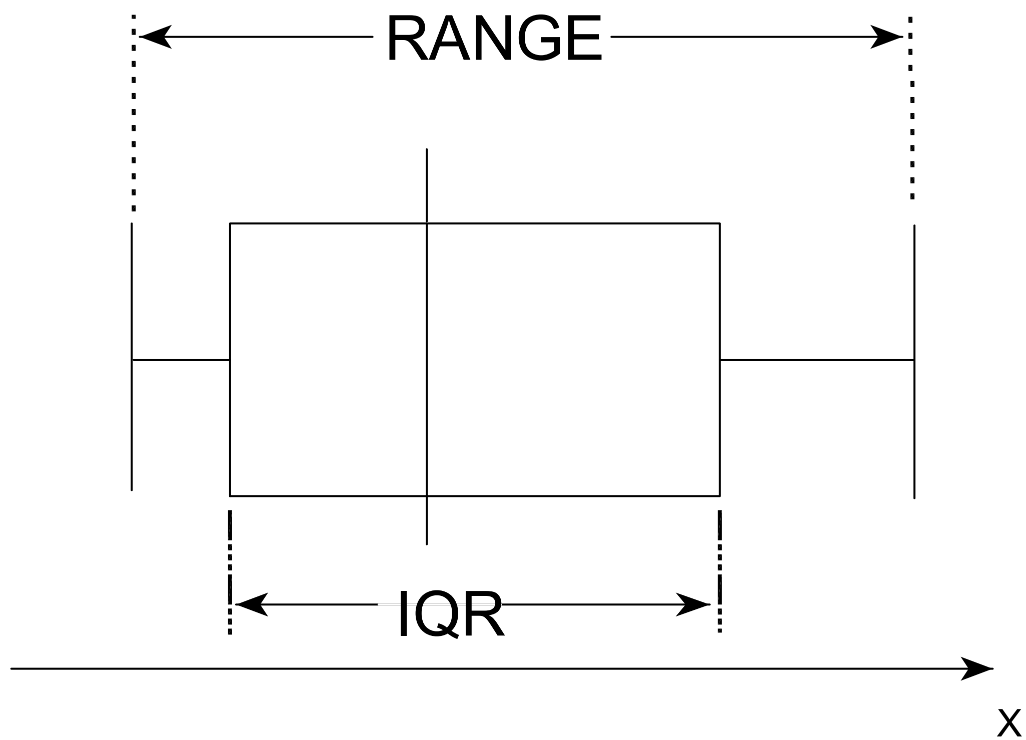


Another Measure of Variation: Inner Quartile Range (IQR)

- The inner quartile range is given by:

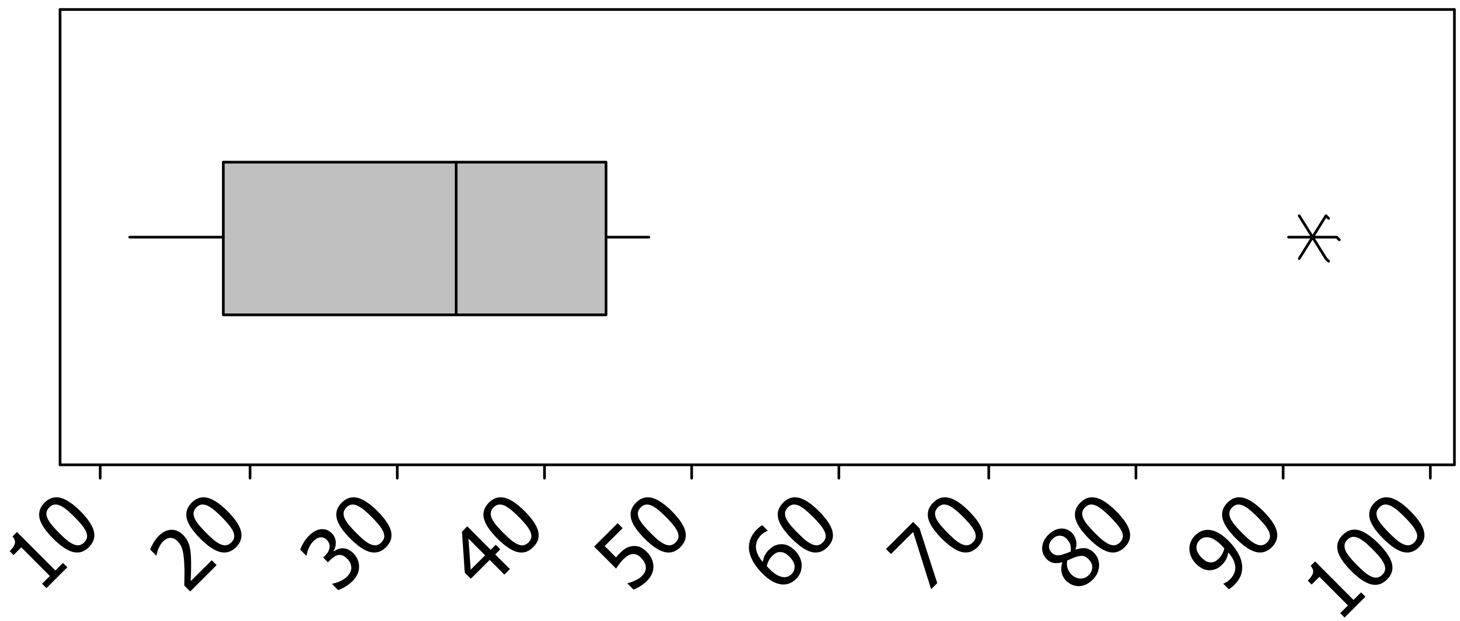
$$IQR = Q_3 - Q_1$$

- The *IQR* is the length of the box in a box and whisker plot.



Test for Potential Outliers

John Tukey showed that the maximum length of a boxplot's whiskers is about $1.5 \times IQR$ so observations beyond that range are usually indicated with special symbols to indicate that those observations might be outliers.



Measures of Variation

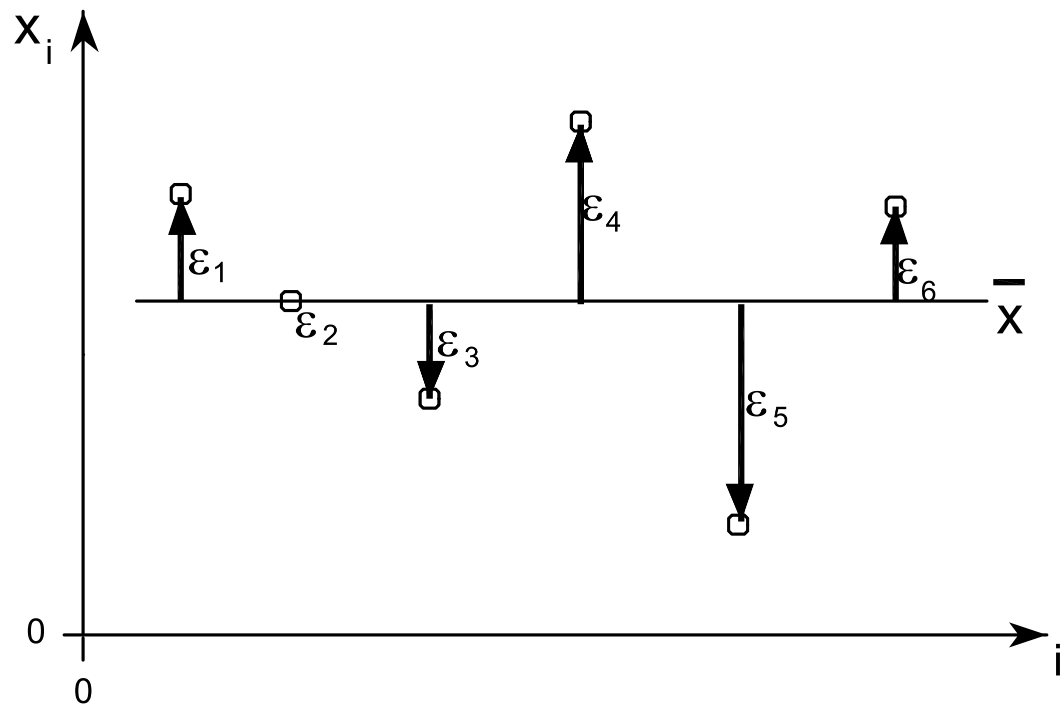
The range is easy to determine but because it only uses two values from the data set it tends to be noisy. We require a better measure of variation that uses all of the data values. It should have the form:

$$\begin{pmatrix} \text{Measure of} \\ \text{Variation} \end{pmatrix} = \frac{\begin{pmatrix} \text{Total Variation} \\ \text{in Data Set} \end{pmatrix}}{\begin{pmatrix} \text{Number of Values} \\ \text{to Measure Variation} \end{pmatrix}}$$

Our new measure of variation should also have the same units as the measurement units.

Measures of Variation

- Quantify the deviation of a data point from the sample mean as $\epsilon_i = x_i - \bar{x}$:



- Consider the ϵ_i as small errors relative to the sample mean.
- $\bar{x}$ is the signal and the ϵ_i are the noise.
- The Hungarian mathematician Paul Erdos referred to small children as epsilons.

Measures of Variation

We can determine a new measure of variation from:

$$\bar{\epsilon} = \frac{1}{n}\sum_{i=1}^{n}\epsilon_i$$

except $\bar{\epsilon} = 0$. Why?

We could fix the problem by taking the absolute value of the ϵ_is:

$$\bar{\epsilon} = \frac{1}{n}\sum_{i=1}^{n}|\epsilon_i|$$

This quantity is called the mean deviation but has no (or limited) physical meaning or applications.

Measures of Variation: The Variance

Another way to separate the sign from the size of a number is to square it, that is, we could consider the summation of the ϵ_i^2 values:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} \epsilon_i^2$$

where $\epsilon_i = x_i - \bar{x}$. This measure of variation called the *variance* (s^2). Variances are VERY important in statistics because:

- They are physically meaningful.
- They have nice mathematical properties:
  - Variances are additive: $s_{total}^2 = s_1^2 + s_2^2 + \cdots$
  - Variance ratios are useful: $F = s_1^2 / s_2^2$

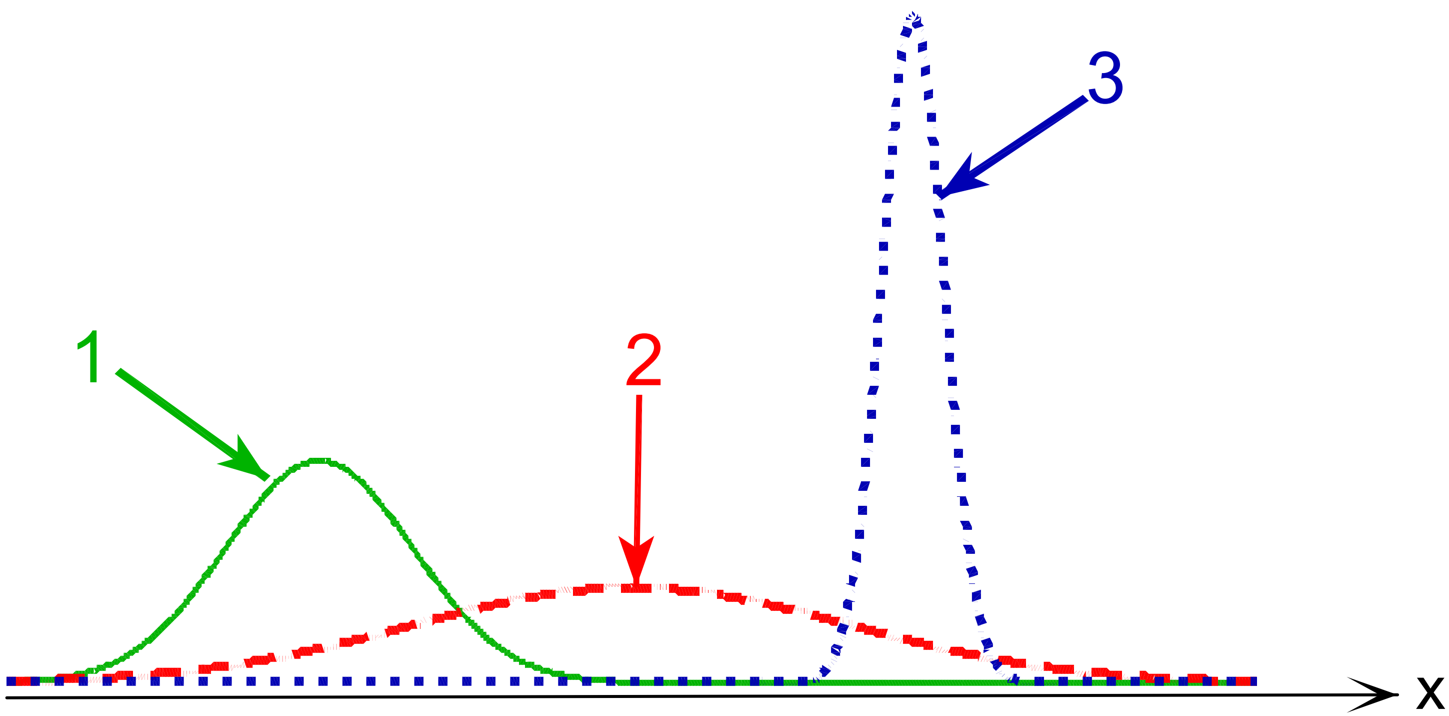
Measures of Variation: The Standard Deviation

- A problem with the variance is that it has units of the measurement units squared, that is, if we measure in inches the variance is in square inches.
- To keep discussions of variation expressed in terms of the measurement units, we use the square root of the variance called the standard deviation (s):

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\epsilon_i^2}$$

- If the x_i are in inches then s is in inches, too.
- When the standard deviation is small, the distribution of the data is tight. When the standard deviation is large, the data are all spread out.

Problem: Which distribution has the highest mean? Standard deviation?



The Standard Deviation

Example: Find the sample standard deviation of the following data set:

i	1	2	3	4	5
x_i	14	18	17	12	14

Solution: The mean of the data set is:

$$\begin{aligned} \bar{x} &= \tfrac{1}{5}\sum_{i=1}^{5} x_i \\ &= \tfrac{1}{5}(75) \\ &= 15 \end{aligned}$$

The deviations from the mean and their squares are:

i	1	2	3	4	5
x_i	14	18	17	12	14
ϵ_i	-1	3	2	-3	-1
ϵ_i^2	1	9	4	9	1

The standard deviation is:

$$\begin{aligned} s &= \sqrt{\tfrac{1}{n-1}\sum_{i=1}^{n}\epsilon_i^2} \\ &= \sqrt{\tfrac{1}{4}(24)} \\ &= 2.45 \end{aligned}$$

The Calculating Form of the Standard Deviation

The defining form of the standard deviation:

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\epsilon_i^2}$$

where $\epsilon_i = x_i - \bar{x}$ can be expanded to obtain:

$$s = \sqrt{\frac{n\sum x_i^2 - \left(\sum x_i\right)^2}{n(n-1)}}$$

This formula is easy to use in spreadsheets and calculators because you only need to keep track of n, $\sum x_i$, and $\sum x_i^2$.

The Calculating Form of the Standard Deviation

Example: Find the sample standard deviation of the following data set:

i	1	2	3	4	5
x_i	14	18	17	12	14

Solution: We have $\sum_{i=1}^{5} x_i = 75$ and $\sum_{i=1}^{5} x_i^2 = 1149$. The value of the standard deviation is:

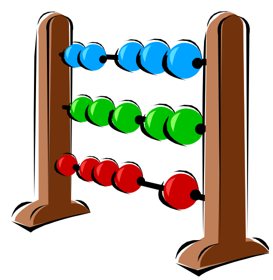
$$
\begin{aligned}
s &= \sqrt{\frac{5(1149)-(75)^2}{5(5-1)}} \\
&= 2.45
\end{aligned}
$$

The Population Standard Deviation

- The population standard deviation is given by:

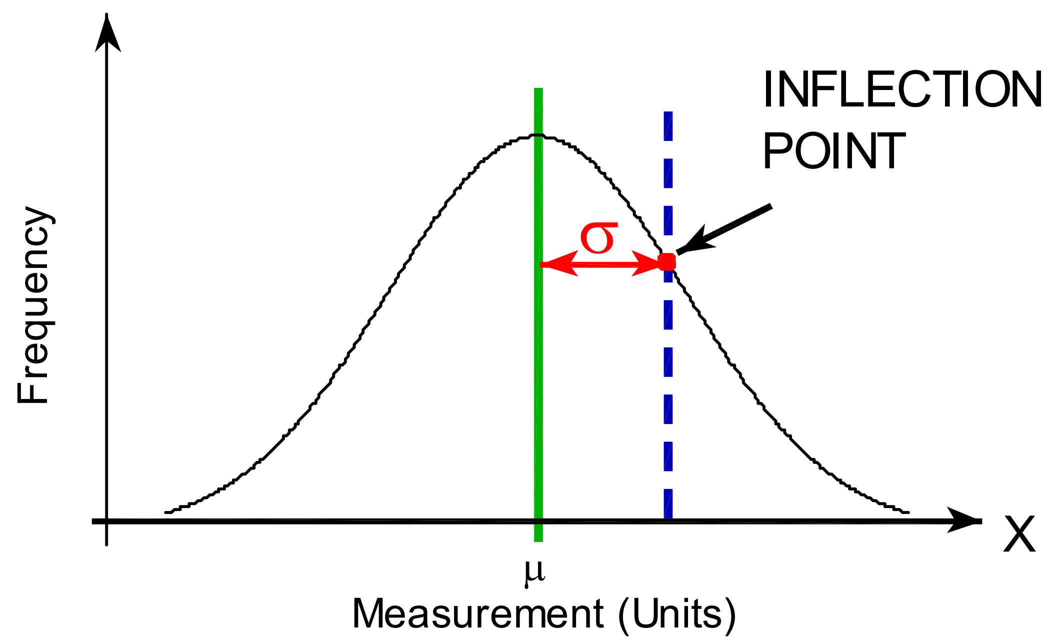
$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\epsilon_i^2}$$

  where $\epsilon_i = x_i - \mu$. It is very rare to use this form of the standard deviation since we are almost always dealing with sample data.
- Many calculators have two keys for standard deviation calculations, one for σ and the other for s. You will rarely, if ever, use the σ key on your calculator.



Measures of Variation: The Standard Deviation

The standard deviation is the most common quantity used to describe the amount of variation in a data set, but just what does it correspond to physically?



Estimating σ from the Range

The population standard deviation can be estimated from the range by:

$$\sigma \simeq \frac{R}{d_2}$$

where d_2 are fudge factors that depend on the sample size used to determine R.

n	2	3	4	5	6	7	8	9	10	15
d_2	1.128	1.693	2.059	2.326	2.534	2.704	2.847	2.970	3.078	3.472

Estimating σ from the Range

Example: The largest and smallest values from a sample of size $n = 8$ are 88 and 65. Estimate the population standard deviation from the range.

Solution: The range is $R = 88 - 65 = 23$. The d_2 value for a sample size of $n = 8$ is $d_2 = 2.847$ so the estimate for σ is:

$$\begin{aligned}\sigma &\simeq \frac{R}{d_2} \\ &\simeq \frac{23}{2.847} \\ &\simeq 8.1\end{aligned}$$

Estimating σ from the Range

Example: An SPC range chart plots ranges determined from samples of size of size $n = 5$ versus time. There are 50 ranges plotted on the chart and the mean range is $\bar{R} = \frac{1}{50}\sum_{i=1}^{50} R_i = 4.8$. Estimate the population standard deviation.

Solution: In order to proceed we must assume that the ranges are in control, i.e. that the population standard deviation was constant while the 50 samples of size $n = 5$ were drawn. With $d_2(n = 5) = 2.326$, the estimate for the population standard deviation is:

$$\begin{aligned}\sigma &\simeq \frac{\bar{R}}{d_2} \\ &\simeq \frac{4.8}{2.326} \\ &\simeq 2.1\end{aligned}$$

Applications of the Standard Deviation: Chebyshev's Theorem

The way data falls about the mean is constrained by the standard deviation. Chebyshev's Theorem is one of the most fundamental expressions of this constraint. Chebyshev's (also written Tschebychev) Theorem states that the fraction of the data which falls within k standard deviation of the mean is at least $1 - 1/k^2$:

$$f(\bar{x} - ks < x < \bar{x} + ks) \geq 1 - \frac{1}{k^2}$$

Chebyshev's Theorem is very important because it always works. All data sets taken from single processes must comply with Chebyshev. Chebyshev is the safety net of distributions. See *Appendix 1: Chebyshev's Theorem* for details.

Chebyshev's Theorem

Example: What fraction of the population of xs should fall in the interval $54 < x < 66$ if $\bar{x} = 60$ and $s = 2$?

Solution:

$$f(54 < x < 66)$$

$$f(60 - 6 < x < 60 + 6)$$

$$f(60 - 3(2) < x < 60 + 3(2)) \quad \rightarrow k = 3$$

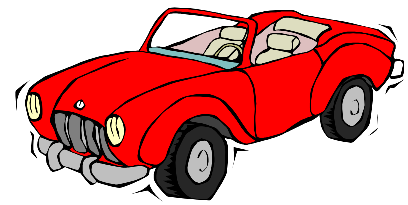
$$\begin{aligned} f(60 - 3(2) < x < 60 + 3(2)) &\geq 1 - \tfrac{1}{k^2} \\ &\geq 1 - \tfrac{1}{3^2} \\ &\geq \tfrac{8}{9} \end{aligned}$$

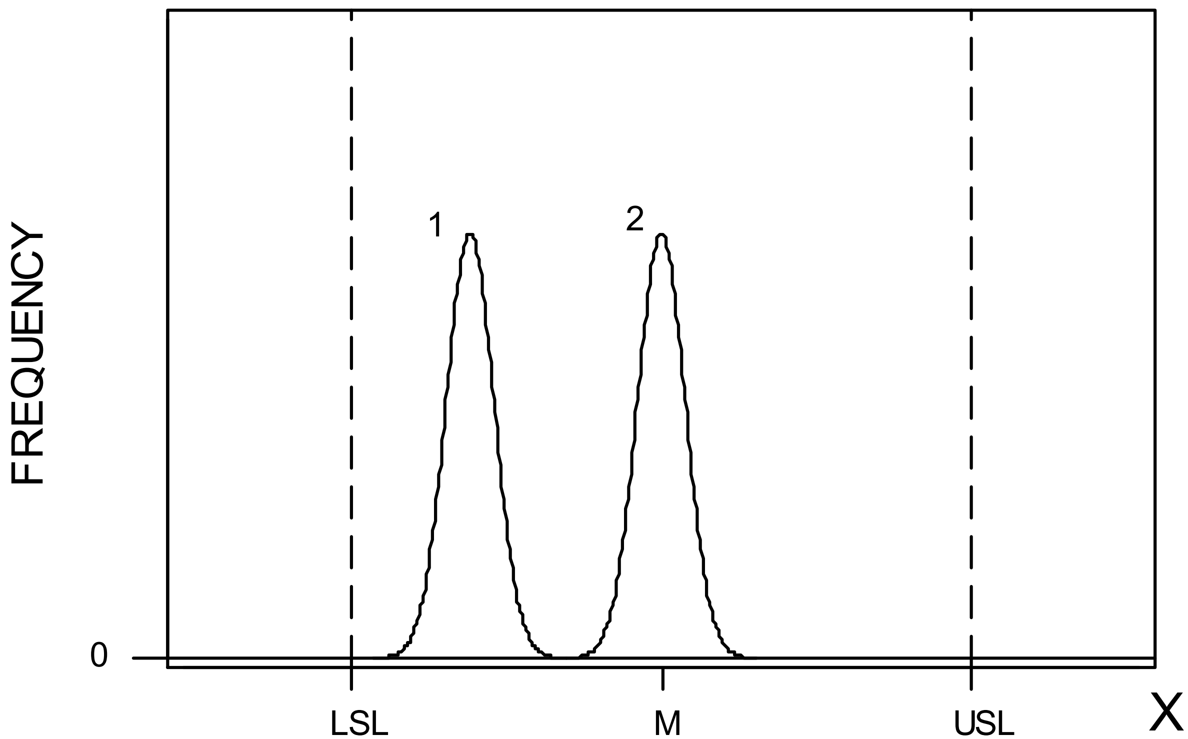
Applications of the Standard Deviation: **Taguchi's Loss Function**

Taguchi's Definition of Quality:

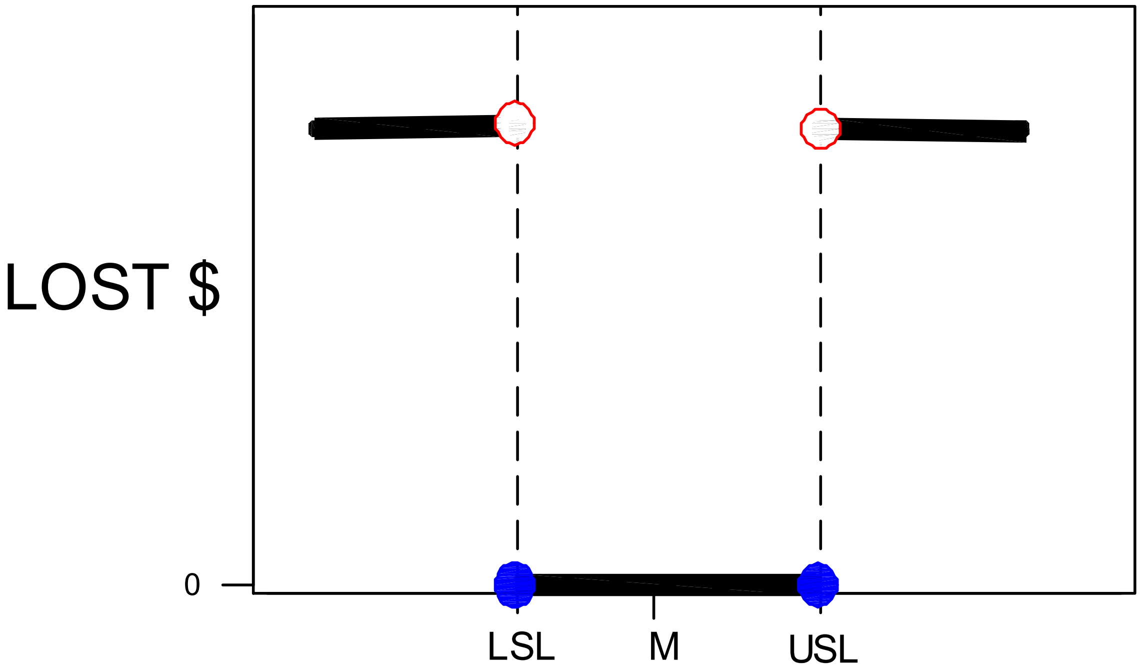
Quality is the value of the goods and services lost to society.

Which Car Do You Want?

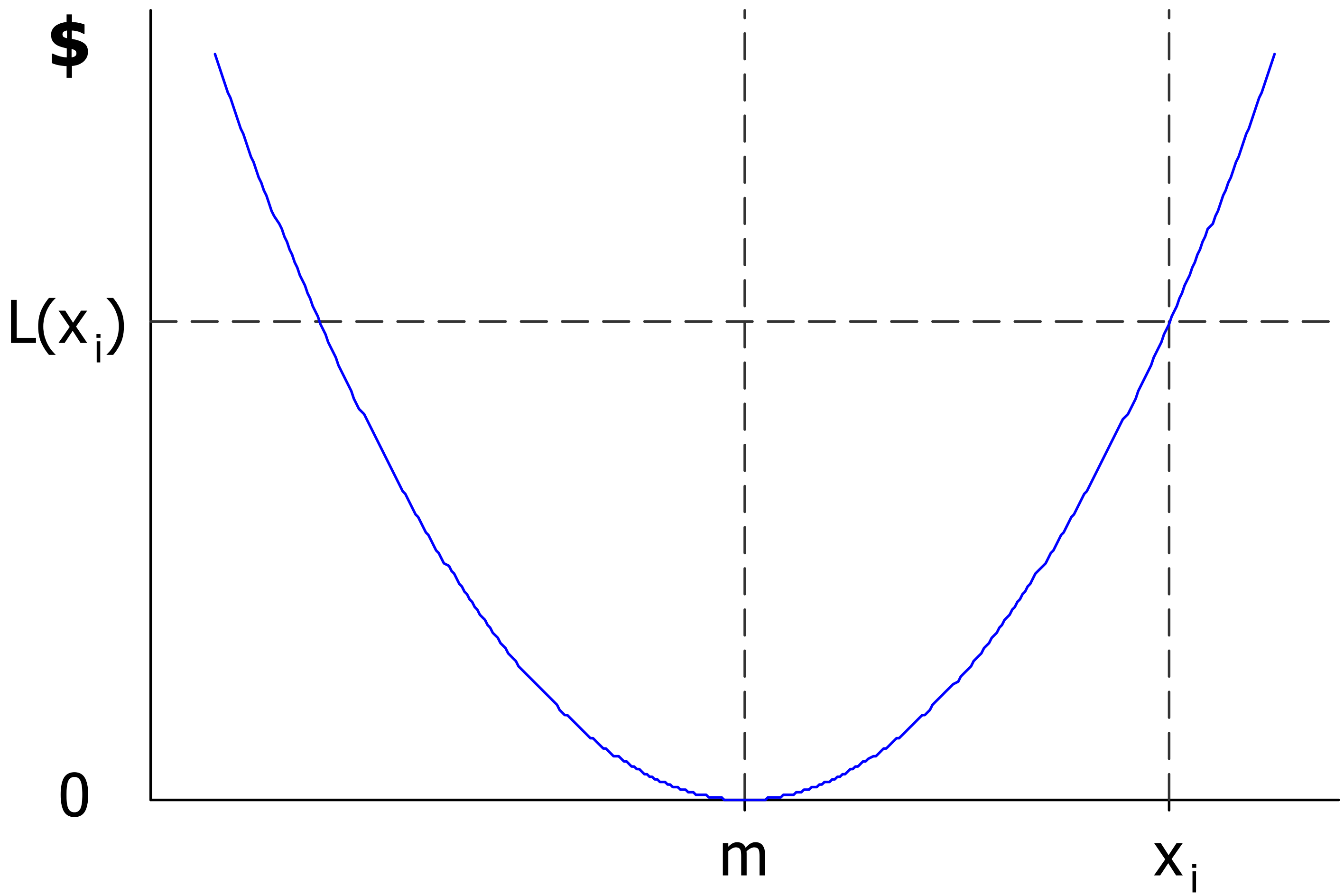




The Goalpost Philosophy of Lost Value



Taguchi's View of Lost Value



Taguchi has shown that:

- The goalpost philosophy of lost value is incorrect.
- The size of the loss (L) increases quadratically with the amount of deviation from the designer's intended target value (m), i.e.

$$L(x_i) = k(x_i - m)^2$$

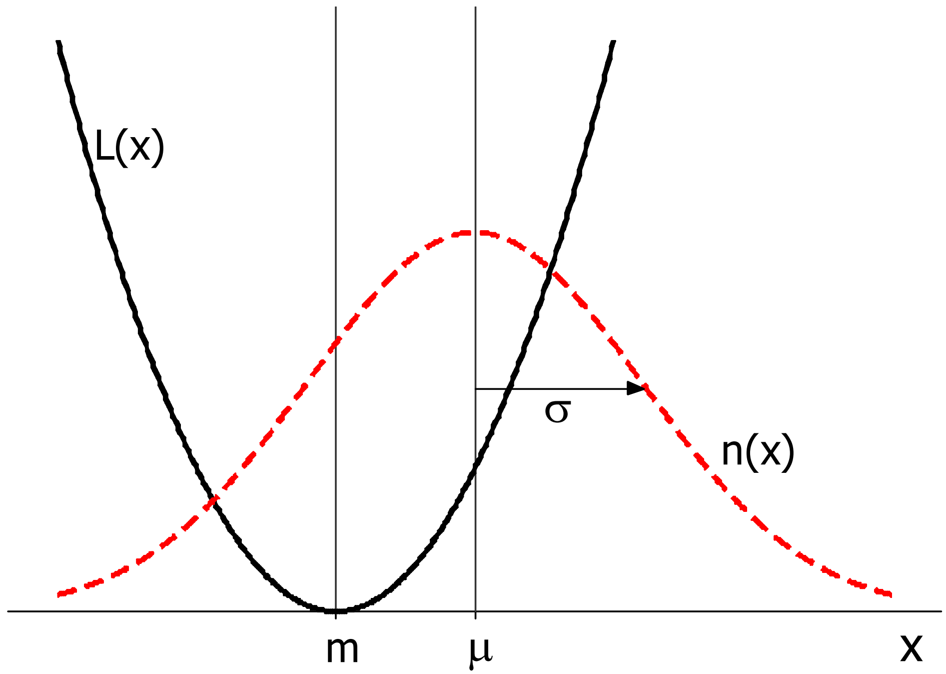
where k is in dollars per measurement unit squared and is determined from quality cost accounting data.

Taguchi has shown that:

- The average loss per part produced ($\bar{L}$) is given by:

$$\bar{L} = \frac{\sum n(x_i)L(x_i)}{\sum n(x_i)} = k\left[\sigma^2 + (\mu - m)^2\right]$$

where $n(x_i)$ is the number of parts produced with dimension x_i.



This provides a mandate for SPC and DOE!

Ford Taurus Transmissions

- Ford introduced the Taurus in 1984.
- The product was very popular so Ford supplemented US-made transmissions with Mazda-made transmissions.
- The Ford and Mazda manufacturing processes were making the same transmissions to the same specs using comparable equipment and staff.
- Ford observed that the US-made transmissions came back at their expected warranty return rate but none of the Mazda-made transmissions came back.
- Upon investigating the situation, Ford determined that the US-made transmissions used about 80% of the allowed tolerance while the Mazda-made transmissions used only about 30% of the allowed tolerance on the KPIVs and kept them centered in their specs.
- Ford published their observations.

GM versus Honda Connecting Rods

- A connecting rod (shaped like a dog bone with holes in the ends) in an engine connects the piston to the crankshaft.
- GM mass produces connecting rods.
- GM connecting rods have raised cylindrical pads on their ends. Each set of connecting rods must be dynamically balanced on a special machine by grinding material off the pads.
- Honda connecting rods don't have pads on their ends and there is no balancing/grinding operation.
- Honda collects and wires together 4 or 6 consecutive connecting rods coming off of the machine to install together in an engine. This strategy works because consecutive rods are almost identical; however, rods made at different times can be wildly different.
- GM's process is more expensive to operate because of the extra operation; however, that cost is negligible compared to the cost of ...
- GM's cars are 300 pounds heavier than Honda's because the crankcase has to be bigger to provide clearance for the connecting rod pads.
- GM didn't publish.

Counting

- Multiplication of choices
- Factorials
- Permutations
- Combinations
- Definition of Probability

Counting: Multiplication of Choices

If a series of k decisions must be made and the first can be made in n_1 ways, the second in n_2 ways, and so on, then the total number of different ways that all k decisions can be made, n_{total}, is:

$$n_{total} = n_1 n_2 \cdots n_k$$

This is the multiplication of choices rule.

Example: If an arc lamp experiment is going to be constructed and there are five arctube designs, three mount designs, two bulb types, and four bases, how many unique configurations can be constructed?

Solution: Since $n_{total} = 5 \times 3 \times 2 \times 4 = 120$ there are 120 unique lamp configurations.

Note: This experiment design is called a factorial design.

Counting: Factorials

If there are n distinct objects in a set and all n of them must be picked then the total number of different ways they can be picked is:

$$\text{Number of ways} = n(n-1)(n-2)(n-3)\cdots(3)(2)(1) = n!$$

Counting: Permutations

- If there are n distinct objects in a set and r of them are to be picked where *the order in which they are picked is important*, then there are ${}_nP_r$ ways to make the selections where:

$$
\begin{aligned}
{}_nP_r &= n(n-1)(n-2)\ldots(n-r+1) \\
&= \frac{n!}{(n-r)!}
\end{aligned}
$$

- Derivation:

$$n! = \underbrace{\overbrace{n(n-1)(n-2)\ldots(n-r+1)}}_{{}_nP_r}\underbrace{(n-r)\ldots 3\cdot 2\cdot 1}_{(n-r)!}$$

$${}_nP_r = \frac{n!}{(n-r)!}$$

Example: How many different ways can a salesman fly to 5 different cities if there are 8 cities in his territory?

Solution: The number of five-city flight plans is:

$$
\begin{aligned}
{}_8P_5 &= \frac{8!}{(8-5)!} \\
&= \frac{8!}{3!} \\
&= \frac{8\times 7\times 6\times 5\times 4\times 3!}{3!} \\
&= 6720
\end{aligned}
$$

Counting: Combinations

- In many situations we do not care about the order that the objects are obtained, only how many different sets of selections are possible. In these cases the permutation over-counts by a factor of ${}_rP_r$.
- If there are n objects in a set and r of them are to be picked and the order in which the picked objects are received is not important then there are ${}_nC_r$ ways to make the selections where:

$$ {}_nC_r = \binom{n}{r} = \frac{{}_nP_r}{{}_rP_r} = \frac{n!}{r!(n-r)!} $$

Example (**revisiting the air-travelling salesman**): How many different sets of five cities can the salesman visit if there are 8 cities in his territory?

Solution: The number of sets of five cities he has to select from is:

$$\binom{8}{5} = \frac{8!}{5!(8-5)!}$$

$$= \frac{8\times7\times6\times5!}{5!3!}$$

$$= 56$$

Counting: Combination Examples

Example: Product supplied from five different vendors is to be tested and compared for differences in location. If each vendor's mean is compared to every other vendor's mean then how many tests have to be performed?

Solution:

$$\binom{5}{2} = \frac{5!}{2!3!} = \frac{5 \times 4 \times 3!}{2!3!} = 10$$

If the numbers 1 through 5 are used to indicate the five vendors, then the two-vendor *multiple comparisons tests* that must be performed are: 12, 13, 14, 15, 23, 24, 25, 34, 35, 45.

Counting: Combination Examples

Example: An experiment with six variables is to be performed. If we are concerned about the possibility of interactions between variables, then how many two-factor and three-factor interactions are there?

Solution:

$$\binom{6}{2} = \frac{6!}{2!4!} = \frac{6 \times 5 \times 4!}{2!4!} = 15$$

$$\binom{6}{3} = \frac{6!}{3!3!} = \frac{6 \times 5 \times 4 \times 3!}{3!3!} = 20$$

The two-factor interactions are: 12, 13, 14, 15, 16, 23, 24 25, 26, 34, 35, 36, 45, 46, 56 and the three-factor interactions are: 123, 124, 125, 126, 134, 135, 136, 145, 146, 156, 234, 235, 236, 245, 246, 256, 345, 346, 356, 456.

Special Combinations

$$\binom{n}{0} = 1$$

$$\binom{n}{1} = n$$

$$\binom{n}{2} = \ldots = \frac{n(n-1)}{2}$$

$$\binom{n}{r} = \binom{n}{n-r}$$

Combinations

Example: A person is on 10 different medications. In addition to the good and bad effects of each medication there is a risk of interactions between drugs. How many different two drug interactions must the doctor be aware of in treating this person? Three drug interactions?

Solution: There are $\binom{10}{2} = 45$ possible two drug interactions and $\binom{10}{3} = 120$ possible three drug interactions.

Events, Sample Spaces, and Probability

- The outcome of a statistical experiment is referred to as an event.

  Example: Obtaining a 3 when rolling a single die.
- The set of all possible outcomes or events is called the experiment's sample space.

  Example: $\{1, 2, 3, 4, 5, 6\}$ is the sample space for rolling a single die.
- If a certain number of the events in an experiment's sample space are successes and the remaining events are failures, then the probability of obtaining a success on any given attempt or trial is:

$$P = \frac{\text{Number of Events that are Successes}}{\text{Number of Events in the Sample Space}}$$

Probability Problems:

Let’s try a game of cards. Assume the cards are pulled from a well shuffled deck of 52 cards.

Example: What is the probability of drawing...

1) an ace?

2) a heart?

3) a four?

4) a face card?

5) a six of diamonds?

6) two sevens?

7) two hearts?

8) two face cards?

9) two cards: a seven then a king?

10) two cards: a seven and a king?

Solutions:

1) $P = \frac{s}{n} = \frac{4}{52} = \frac{1}{13}$

2) $P = \frac{s}{n} = \frac{13}{52} = \frac{1}{4}$

3) $P = \frac{s}{n} = \frac{4}{52} = \frac{1}{13}$

4) $P = \frac{s}{n} = \frac{12}{52} = \frac{3}{13}$

5) $P = \frac{s}{n} = \frac{1}{52}$

6) $P = \frac{s}{n} = \frac{4}{52} \times \frac{3}{51} = \frac{1}{221}$

7) $P = \frac{s}{n} = \frac{13}{52} \times \frac{12}{51} = \frac{1}{17}$

8) $P = \frac{s}{n} = \frac{12}{52} \times \frac{11}{51} = \frac{11}{221}$

9) $P = \frac{s}{n} = \frac{4}{52} \times \frac{4}{51} = \frac{4}{663}$

10) $P = \frac{s}{n} = \frac{4}{52} \times \frac{4}{51} \times 2 = \frac{8}{663}$

More Probability:

Example: Find the probability of drawing 5 cards from a standard randomized deck and getting a full house, kings over queens.

Solution 1:

$$
\begin{aligned}
P &= \binom{5}{3}\left(\frac{4}{52}\right)\left(\frac{3}{51}\right)\left(\frac{2}{50}\right)\left(\frac{4}{49}\right)\left(\frac{3}{48}\right) \\
&= (10)\left(\frac{4}{52}\right)\left(\frac{3}{51}\right)\left(\frac{2}{50}\right)\left(\frac{4}{49}\right)\left(\frac{3}{48}\right) \\
&= \frac{1}{108\,290} \\
&= 9.2 \times 10^{-6}
\end{aligned}
$$

$$\binom{5}{3} = \frac{5!}{3! * 2!}$$

$$= \frac{5 \times 4 \times 3!}{3! \times 2!}$$

$$= \frac{5 \times 4}{2!}$$

$$= 5 \times 2$$

$$= 10$$

KKKQQ

KKQQK

KQQKK

QQKKK

QKQKK

QKKQK

QKKKQ

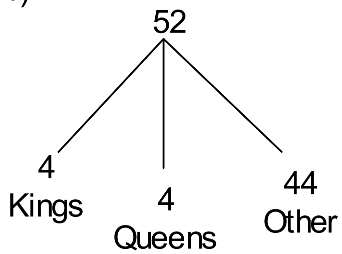
KQKKQ

KKQKQ

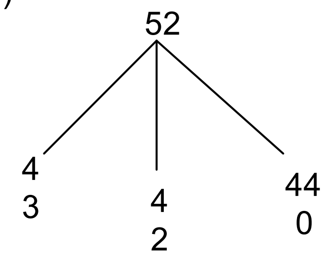
KQKQK

Solution 2:

1)



2)



$$
\begin{aligned}
P &= \frac{\binom{4}{3}\binom{4}{2}\binom{44}{0}}{\binom{52}{5}} \\
&= \frac{(4)(6)(1)}{(2598960)} \\
&= \frac{1}{108\,290} \\
&= 9.\,2 \times 10^{-6}
\end{aligned}
$$

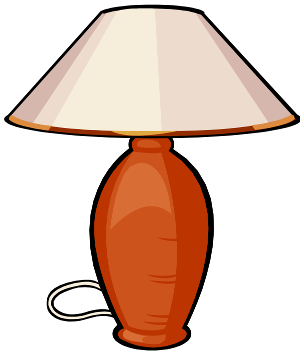
Example: Find the probability of drawing 2 kings, 2 queens, and 1 jack.

Solution 1:

$$(?)(\frac{4}{52})(\frac{3}{51})(\frac{4}{50})(\frac{3}{49})(\frac{4}{48})$$

Solution 2:

$$\begin{aligned} P &= \frac{\binom{4}{2}\binom{4}{2}\binom{4}{1}\binom{40}{0}}{\binom{52}{5}} \\ &= \frac{(6)(6)(4)(1)}{(2.6*10^6)} \\ &= \frac{(144)}{(2.6*10^6)} \\ &=.000055 \end{aligned}$$



Example: There are 21 good lamps and 3 bad lamps in a package. What is the probability of randomly picking 4 lamps and getting exactly one bad lamp?

Solution:

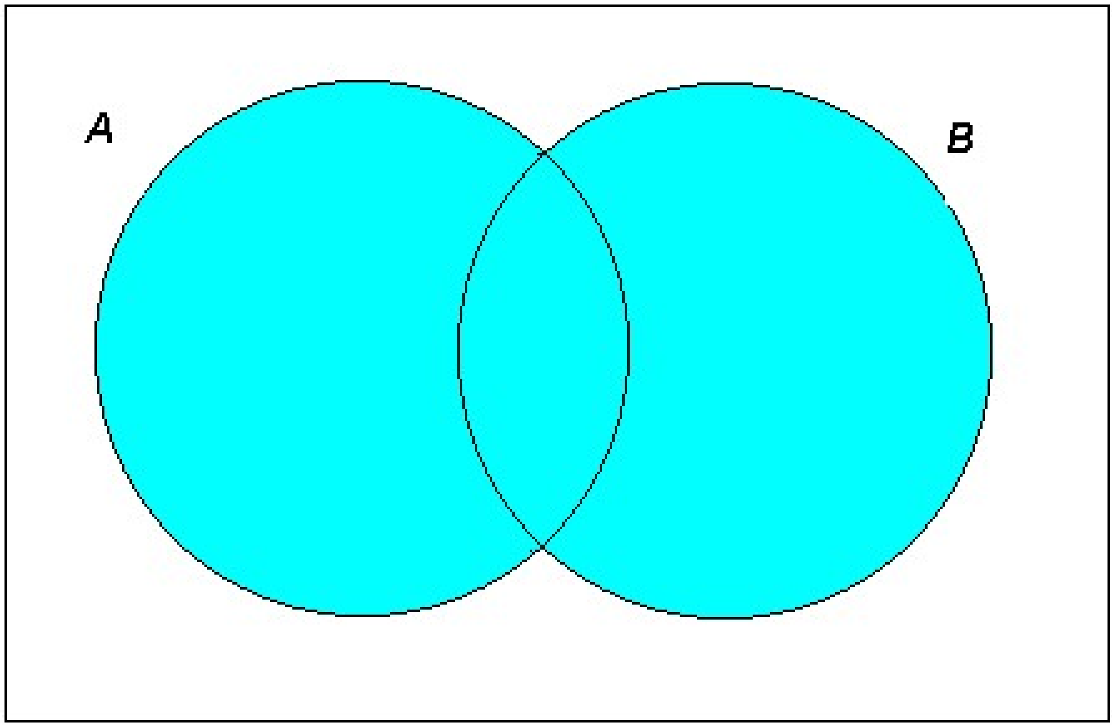
$$\frac{\binom{3}{1}\binom{21}{3}}{\binom{24}{4}} = 0.375$$

Operations Used to Relate Events:

- Unions
- Intersections
- Complements
- These operations may be performed graphically using a Venn diagram or symbolically.

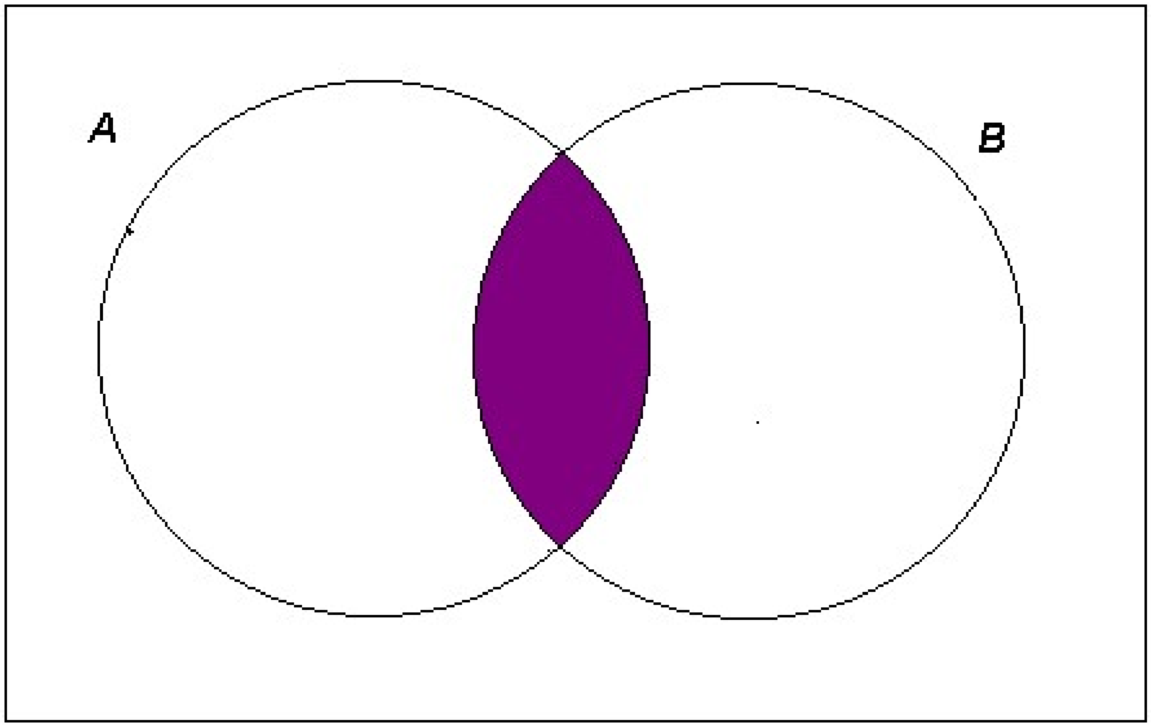
Unions

If A and B are events in the sample space S, then the set of events that comes from A OR B is given by the union of A and B. The union of A and B is given by $A \cup B$.



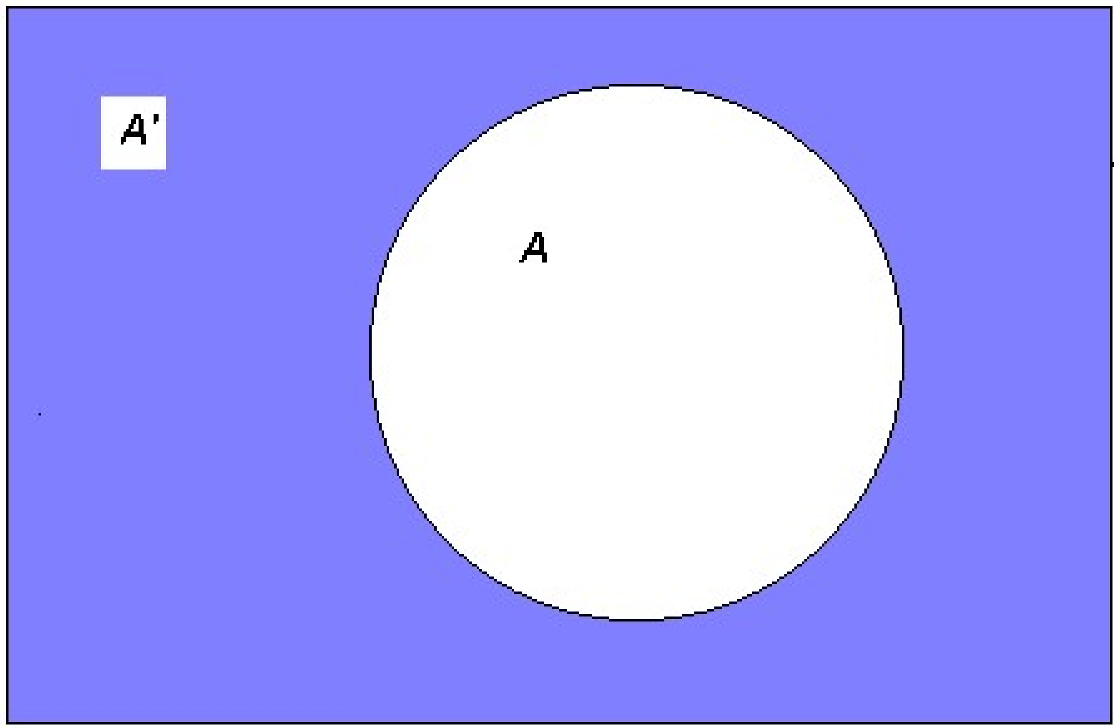
Intersections

If A and B are events in the sample space S, then the set of events that is common to both A AND B is given by the intersection of A and B. The intersection of A and B is given by $A \cap B$.



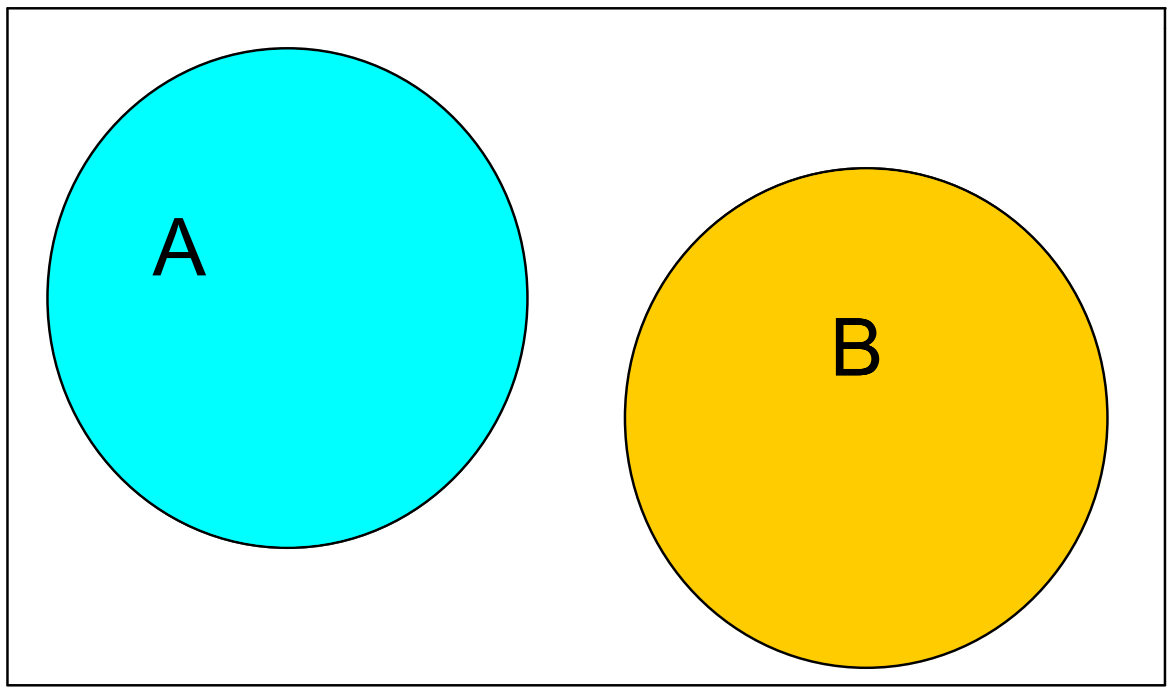
Complements

If A is an event in S, then those events in S that are not in A $\left(\text{NOT } A\right)$ are given by the complement of A. The complement operation is indicated by A'.



Mutually Exclusive

- Definition: A set is said to be null if it is empty. The null set is indicated by $\{\}$ or $\emptyset$.
- If $A \cap B = \{\}$ then events A and B are mutually exclusive.



- Examples: military intelligence, efficient government, jumbo shrimp, basic statistics.

Problem: Event A corresponds to rolling a single die and obtaining an even number. Event B corresponds to rolling a single die and obtaining a 1, 2, or 3. Find the events that correspond to:

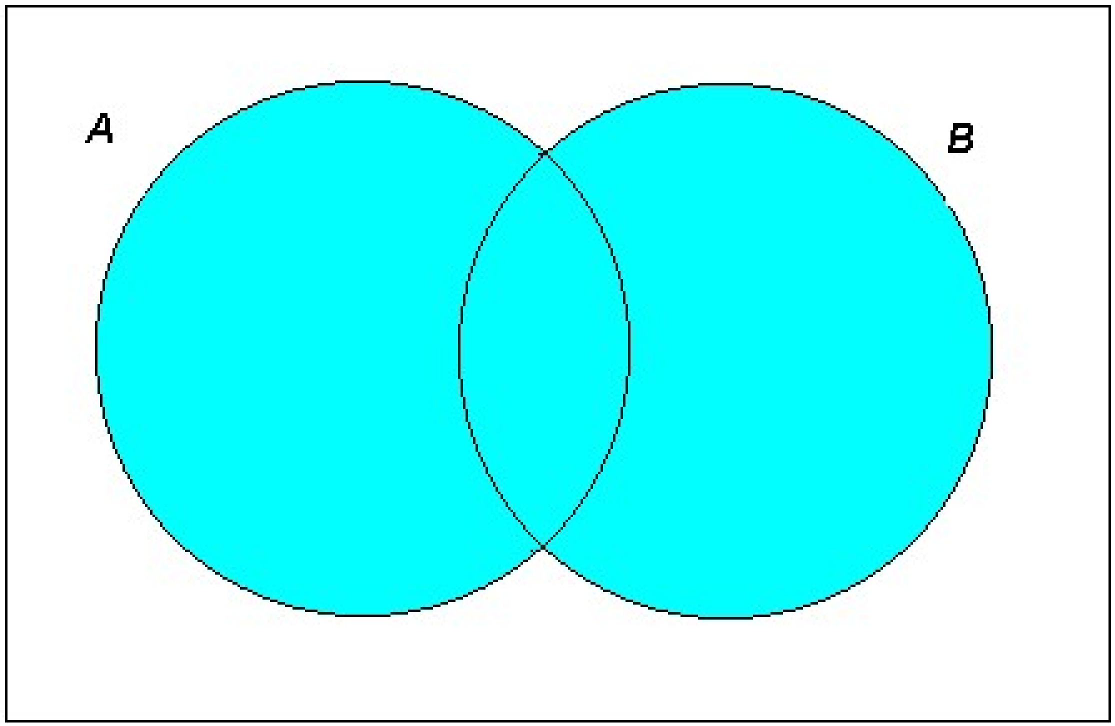
1. $A \cup B$
2. $A \cap B$
3. A'

Solution:

1. $A \cup B = \{2,4,6\} \cup \{1,2,3\} = \{1,2,3,4,6\}$
2. $A \cap B = \{2,4,6\} \cap \{1,2,3\} = \{2\}$
3. $A' = \{2,4,6\}' = \{1,3,5\}$

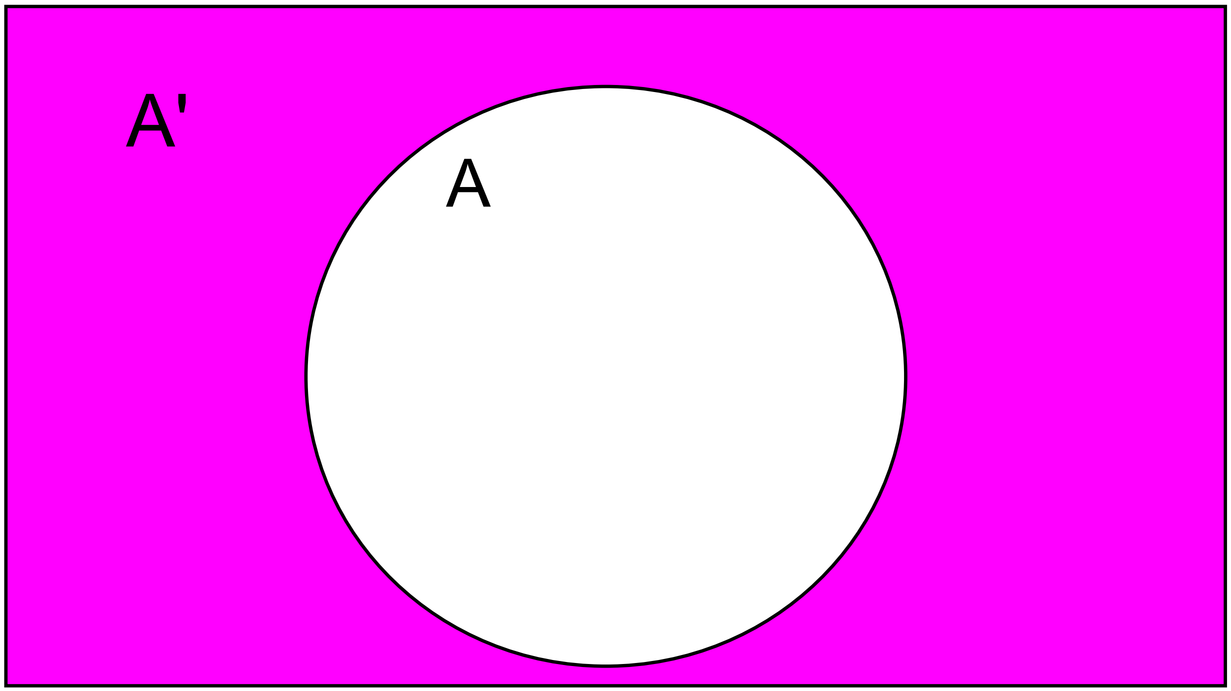
Probability Rules

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$



Probability Rules

$$P(A') = 1 - P(A)$$

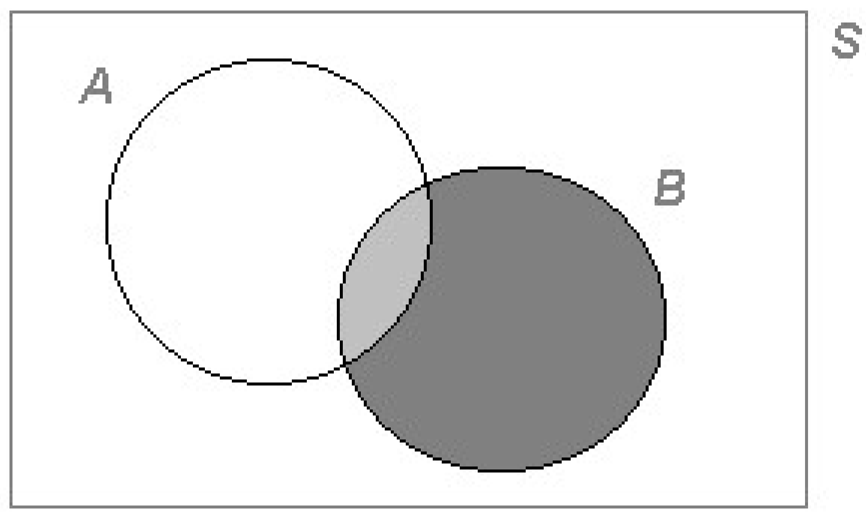


Conditional Probability

The probability of event A occurring given that event B has occurred is given by the conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Event B determines the subset of S that the problem is limited to.

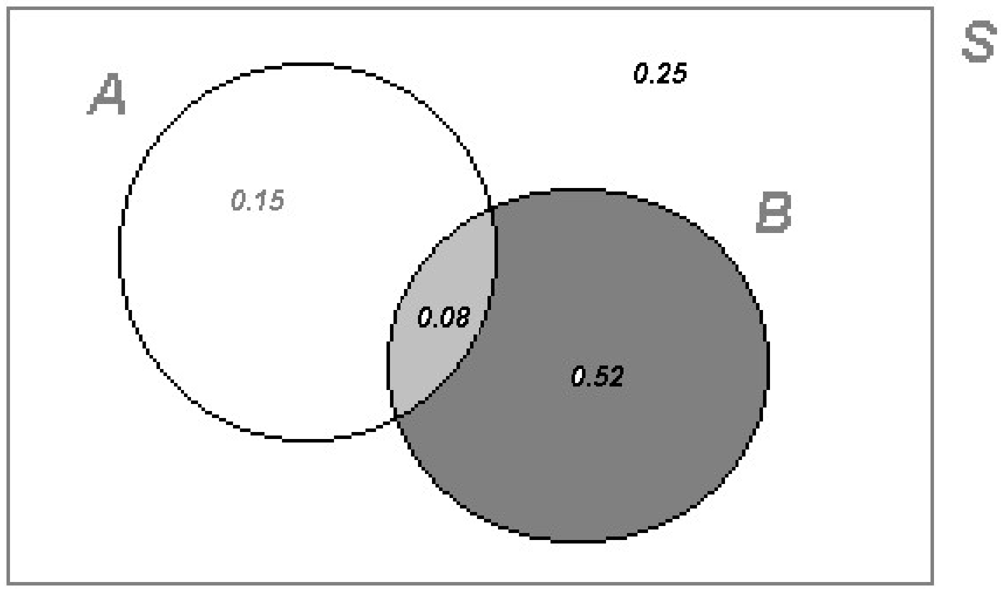


Problem: For the situation in the Venn diagram find the following probabilities:

a) $P(A)$ d) $P(A \cap B)$

b) $P(A')$ e) $P(A|B)$

c) $P(A \cup B)$ f) $P(B|A)$



Solution:

a) $P(A) = 0.15 + 0.08 = 0.23$

b) $P(A') = 1 - P(A) = 1 - 0.23 = 0.77$

c) $P(A \cup B) = 0.15 + 0.08 + 0.52 = 0.75$

d) $P(A \cap B) = 0.08$

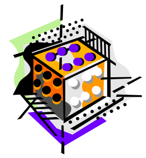
e) $P(A|B) = \frac{0.08}{0.08+0.52} = 0.133$

f) $P(B|A) = \frac{0.08}{0.08+0.15} = 0.348$

Expectation Value

If a game of chance with n possible outcomes pays amounts a_1, a_2, ..., a_n and the probabilities of winning these amounts are p_1, p_2, ..., p_n then in a long series of plays the expected earnings per play are given by:

$$E = \sum_{i=1}^{n} a_i p_i$$

Expectation Value 

Example: In a game of chance players roll a single die. If they roll an odd number they get that many dollars (e.g. rolling a 3 pays $3). If they roll an even number they lose. What are the expected earnings? Should you play to win if it costs $1 to play? $2?

Solution: The expected earnings are:

$$
\begin{aligned}
E &= \sum_{i=1}^{n} a_i p_i \\
&= \left(\$1 \times \frac{1}{6}\right) + \left(\$0 \times \frac{1}{6}\right) + \left(\$3 \times \frac{1}{6}\right) + \left(\$0 \times \frac{1}{6}\right) + \left(\$5 \times \frac{1}{6}\right) + \left(\$0 \times \frac{1}{6}\right. \\
&= \frac{1}{6} + 0 + \frac{3}{6} + 0 + \frac{5}{6} + 0 \\
&= \$1.50
\end{aligned}
$$

At $1 the game would be profitable, but at $2 you’re going to lose!

Net Expectation Value

Example: A card game is played by drawing a single card from a standard deck. The game pays \$1 if you draw a face card and \$5 if you draw an ace. Should you play to win if it costs \$1 to play?

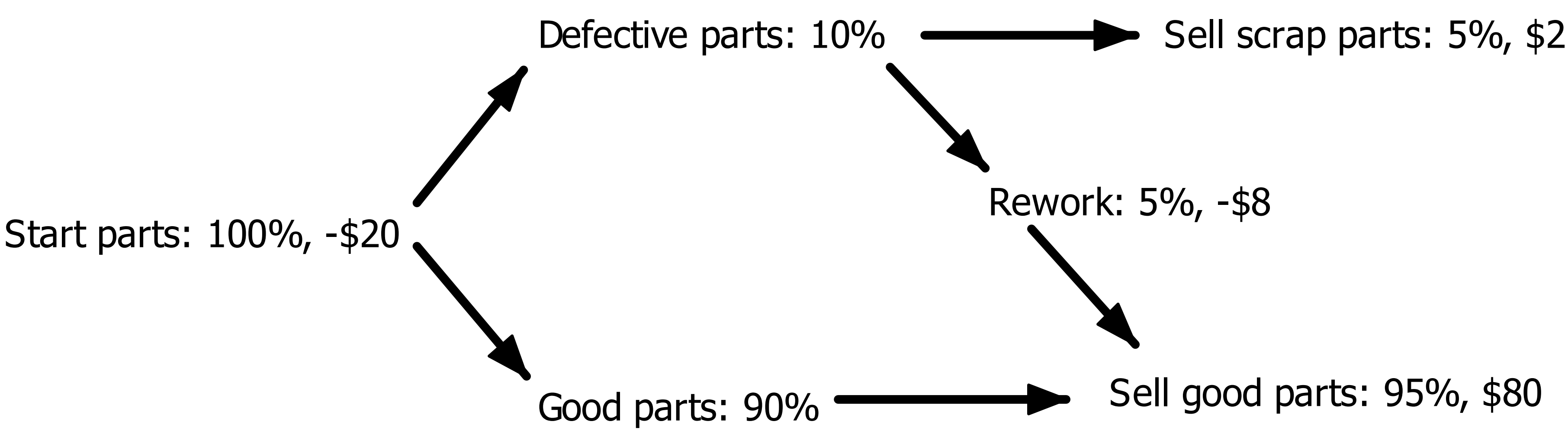
$$
\begin{aligned}
E_{Net} &= E_{Expense} + E_{Income} \\
&= (-\$1 \times 1) + (\$1 \times \frac{12}{52}) + (\$5 \times \frac{4}{52}) \\
&= -1 + \frac{12}{52} + \frac{20}{52} \\
&= -\$0.385
\end{aligned}
$$

On average you will lose $\$0.385$ each time you play.

Expectation Value

Example: A manufacturing business produces parts that use $20 in material and labor. 90% of the parts produced are good and can be sold for full price at $80 each. Of the 10% of the parts that are bad, half can be reworked at a cost of $8 each and sold for full price. The remaining 5% must be scrapped. Scrapped parts can be sold for salvage at $2 each. Find the expected earnings per part and determine the cost of poor quality (COPQ).

Solution:



$$
\begin{aligned}
E &= \sum_{i=1}^{n} a_i p_i \\
&= material + rework + salvage + sales \\
&= (-\$20 \times 1.0) + (-\$8 \times 0.05) + (\$2 \times 0.05) + (\$80 \times 0.95) \\
&= (-20.00) + (-0.40) + (0.10) + (76.00) \\
&= \$55.70
\end{aligned}
$$

Expected earnings are $\$55.70$ per part manufactured. If there were no quality problems, then the expected earnings per part would be:

$$E = (-\$20 \times 1.0) + (\$80 \times 1.0) = \$60.00$$

The COPQ is the difference between the two expected earnings:

$$COPQ = \$60.00 - \$55.70 = \$4.30$$

Problem: Determine if the rework operation is cost effective assuming that parts that would have been reworked will just be sold for scrap instead.

Solution: The new model is:

$$
\begin{aligned}
E &= \sum_{i=1}^{n} a_i p_i \\
&= material + salvage + sales \\
&= (-\$20 \times 1.0) + (\$2 \times 0.10) + (\$80 \times 0.90) \\
&= (-20.00) + (0.20) + (72.00) \\
&= \$52.20
\end{aligned}
$$

so the cost of poor quality without the rework operation is:

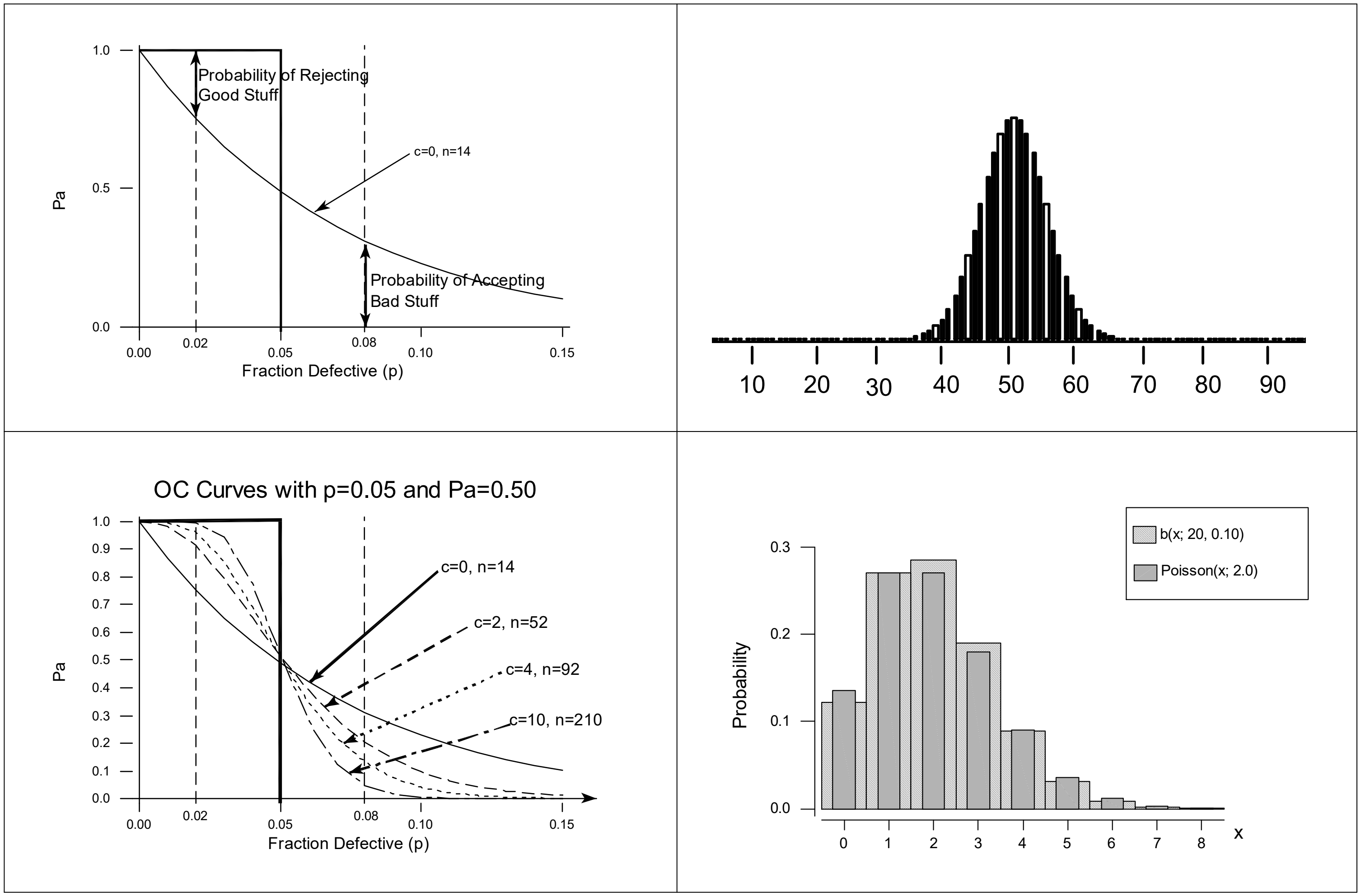
$$COPQ = \$60.00 - \$52.20 = \$7.80$$

The rework operation is cost effective because it saves about $\$3.50$ per part started.

MINITAB Commands

- Stat> Quality Tools> Pareto Chart
- Stat> Quality Tools> Cause-and-Effect
- Graph> Stem-and-Leaf
- Graph> Histogram
- Graph> Dotplot
- Graph> Boxplot
- Stat> Basic Statistics> Display Descriptive Statitistics
- Cal> Column Statistics
- Cal> Row Statistics
- Calc> Calculator

Probability Distributions



Random Variables

Definition: A random variable is a variable (usually x) that can take on the values from an experiment's sample space.

Example 1: $S = \{H, T\}$, $x = H$

Example 2: $S = \{1, 2, 3, 4, 5, 6\}$, $x = 4$

Probability Distributions

Definition: A probability distribution is a function that assigns probabilities to each value of a random variable.

Example 1: $f(x) = 1/2$ for $x = H, T$

Example 2: $f(x) = 1/6$ for $x = 1, 2, 3, 4, 5, 6$

Rules for Probability Distributions

- $0 \leq f(x) \leq 1$
- $f(x) = 0$ for all x not in S
- $\sum_{\text{all } x} f(x) = 1$

Probability Distributions

There are many different probability distributions but only a handful of them are really important to quality work (Pareto principle). The ones we are concerned about for Basic Probability are:

- Hypergeometric Distribution
- Binomial Distribution
- Poisson Distribution
- Normal Distribution
- Student's t Distribution
- Chi-square Distribution
- F Distribution

Distribution Characteristics

- The hypergeometric, binomial, and Poisson distributions characterize discrete random variables (e.g. defectives and defects).
- The normal distribution characterizes many types of measurement data.
- The normal, Student's t, chi-square, and F distributions characterize statistics determined from sample data.

The Hypergeometric Distribution

The functional form of the hypergeometric distribution may be written:

$$h(x; a, b, n) = \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}$$

where a is the number of successes in the lot, b is the number of failures in the lot, n is the sample size, and x is the random variable. Note $N = a + b$ is the lot size.

The hypergeometric distribution is used when:

- Single unique lots are being considered.
- The lot size is small.
- Samples are taken without replacement.

The Hypergeometric Distribution

Example: A lot of 40 parts contains 5 bad parts. If a random sample of 8 parts is drawn from the lot, what is the probability of obtaining 7 good parts and one bad part?

Solution:

40

$$\frac{\binom{5}{1}\binom{35}{7}}{\binom{40}{8}}$$

$$\frac{\binom{5}{1}\binom{35}{7}}{\binom{40}{8}} = \frac{(5)(6.7\times10^{6})}{(7.7\times10^{7})}$$

$$= \frac{(3.4\times10^{7})}{(7.7\times10^{7})}$$

$$= 0.437$$

The Hypergeometric Distribution

Example: An incoming receiving operation inspects boxes of 24 lamps by randomly inspecting 4 lamps. If all 4 lamps pass the inspection the box is accepted, otherwise it is returned to the supplier (Sylvania). If a box of 24 lamps contains 4 bad lamps, what is the probability that the box will be accepted? Rejected?

Solution: We have: $a = 4$, $b = 20$, $n = 4$, and $x = 0$ so:

$$\begin{aligned} h(x; a, b, n) &= \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}} \\ h(0; 4, 20, 4) &= \frac{\binom{4}{0}\binom{20}{4-0}}{\binom{4+20}{4}} \\ &= 0.456 \end{aligned}$$

The probability that the lot will be accepted is $P_A = 0.456$. The probability that the lot will be rejected is $P_R = 1 - 0.456 = 0.544$.

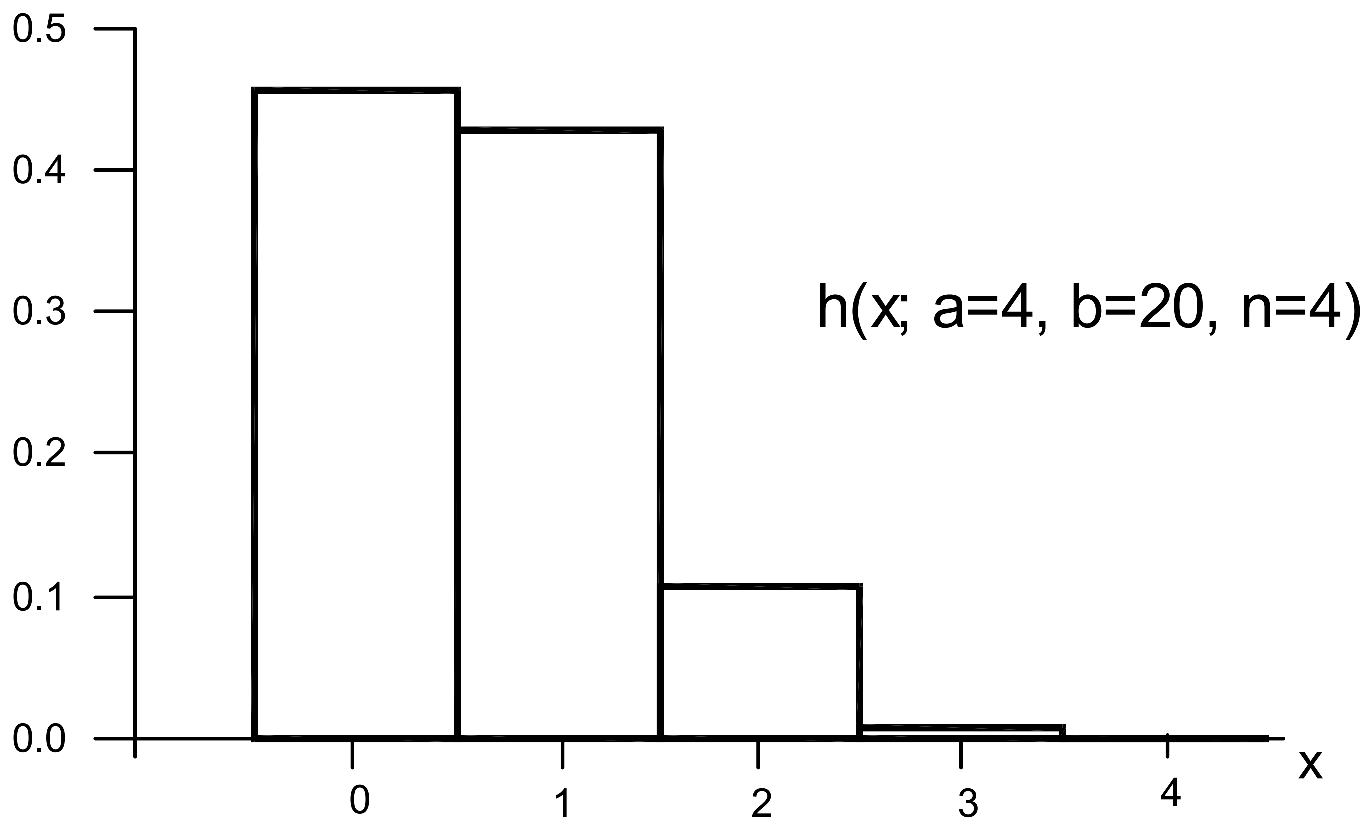
The Hypergeometric Distribution

Example: An incoming receiving operation inspects boxes of 24 lamps by randomly inspecting 4 lamps. If a box of 24 lamps contains 4 bad lamps what is the probability of getting 0, 1, 2, 3, 4, bad lamps? Plot the histogram.

Solution: We have: $a = 4$, $b = 20$, $n = 4$, and $x = 0$ so:

$$
\begin{aligned}
h(x;a,b,n) &= \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}} \\
h(x=0;4,20,4) &= \frac{\binom{4}{0}\binom{20}{4}}{\binom{24}{4}} &= 0.456 \\
h(x=1;4,20,4) &= \frac{\binom{4}{1}\binom{20}{3}}{\binom{24}{4}} &= 0.429 \\
h(x=2;4,20,4) &= \frac{\binom{4}{2}\binom{20}{2}}{\binom{24}{4}} &= 0.107 \\
h(x=3;4,20,4) &= \frac{\binom{4}{3}\binom{20}{1}}{\binom{24}{4}} &= 0.007 \\
h(x=4;4,20,4) &= \frac{\binom{4}{4}\binom{20}{0}}{\binom{24}{4}} &= 0.00009
\end{aligned}
$$

The Hypergeometric Histogram



The Hypergeometric Distribution

Example: Lots of size $N = 40$ are submitted for inspection for defectives. The process is tolerant of a few defectives in lots, but lots with $a \geq 4$ defectives are definitely objectionable. Find the sample size n for the $c = 0$ sampling plan that delivers a 95% probability of rejecting lots with $a = 4$.

Solution: The probability of accepting a lot that contains four defectives corresponds to the hypergeometric probability of finding zero defectives in the sample of size n:

$$\begin{aligned} P_A &= h(x = 0; a = 4, b = 36, n) \\ &= \frac{\binom{4}{0}\binom{36}{n}}{\binom{40}{n}} \end{aligned}$$

The value of n must be chosen to meet the condition:

$$P_A \leq 0.05$$

to deliver a 95% probability of rejecting such a lot. By iterating through several values of n:

$$h(0; 4, 36, 21) = 0.042$$

the smallest value of n that meets the sample size conditions is $n = 21$.

The Binomial Distribution

In an experiment that results in only successes and failures, the probability of obtaining x successes in n trials when the probability of a success p is constant for all trials is:

$$b(x; n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

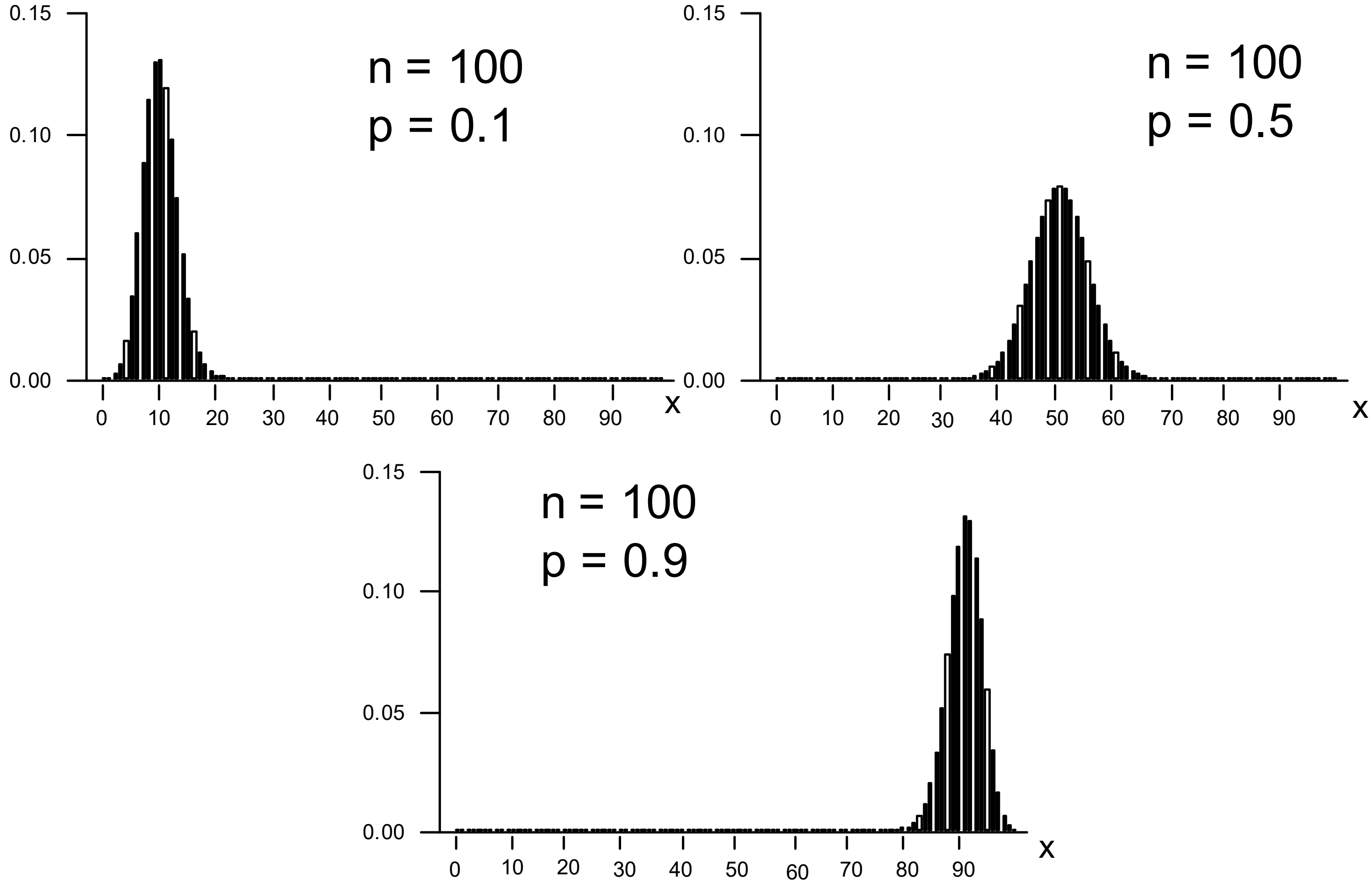
The binomial distribution is used to track the number of defectives produced by a continuous process.

The Binomial Distribution

How It Works:

1. The probability of a success on any one trial is p so the probability of getting x successes must be p^x.
2. The probability of a failure on any one trial is $(1-p)$ so the probability of getting $n-x$ failures is $(1-p)^{n-x}$.
3. The probability $p^x(1-p)^{n-x}$ corresponds to getting x successes followed by $n-x$ failures in *exactly* that order. Since we do not care about the order, only getting the required number of successes and failures, we must consider the number of ways this can happen. The number of ways that we can get x successes in n trials is $\binom{n}{x}$.

The Binomial Histograms



A Note About Successes and Failures

Statisticians classify events as successes or failures, but it is up to the user to define a success. For example, an inspector's purpose in life is to find defective parts. If he's not finding defective parts then he's not doing his job. So a success for an inspector is a bad part, i.e. a failure.

The point of view taken universally in the Quality field is that of the inspector.

The Binomial Distribution

Example: A sample of 100 pieces is taken each hour from a continuous manufacturing process. Find the probability of getting exactly 2 defective parts if the process is producing 3% bad parts.

Solution:

$$
\begin{aligned}
b(x;n,p) &= \binom{n}{x} p^x (1-p)^{n-x} \\
b(2;100,0.03) &= \binom{100}{2} (0.03)^2 (1-0.03)^{100-2} \\
&= (4950)(0.00090)(0.97)^{98} \\
&= (4950)(0.00090)(0.0505) \\
&= 0.225
\end{aligned}
$$

The Binomial Distribution

The Binomial Distribution is used when:

- Outcomes are binary - only successes or failures.
- The sample size is known and fixed.
- The probability of a success on any one trial is constant.
- The outcomes are independent.

The Cumulative Binomial Distribution

In many QC situations the number of successes ranges from $x = 0$ to some upper limit $x = c$. Since this problem is so common, it is useful to define the cumulative binomial distribution:

$$b(c; n, p) = \sum_{x=0}^{c} b(x; n, p)$$

Cumulative Binomial Probability Tables

Most binomial probability tables display cumulative probabilities. For example, the following table gives the values of $b(c; n = 10, p)$:

c	$p = 0.01$	$p = 0.05$	$p = 0.10$
0	0.9044	0.5987	0.3487
1	0.9957	0.9139	0.7361
2	0.9999	0.9885	0.9298
3	1	0.9990	0.9872
4	1	0.9999	0.9984
5	1	1	0.9999
6	1	1	1

The Cumulative Binomial Distribution

Example: A sampling plan for controlling defectives from a continuous process uses samples of size 80 taken and inspected each hour. If two or fewer of the parts are found to be defective the process is continued. If three or more defectives are found in a sample the process is shut down until the problem can be found. What's the probability of letting the process run if it is producing 3% bad parts? What's the probability of shutting the process down?

Solution: Probability of letting the process run is:

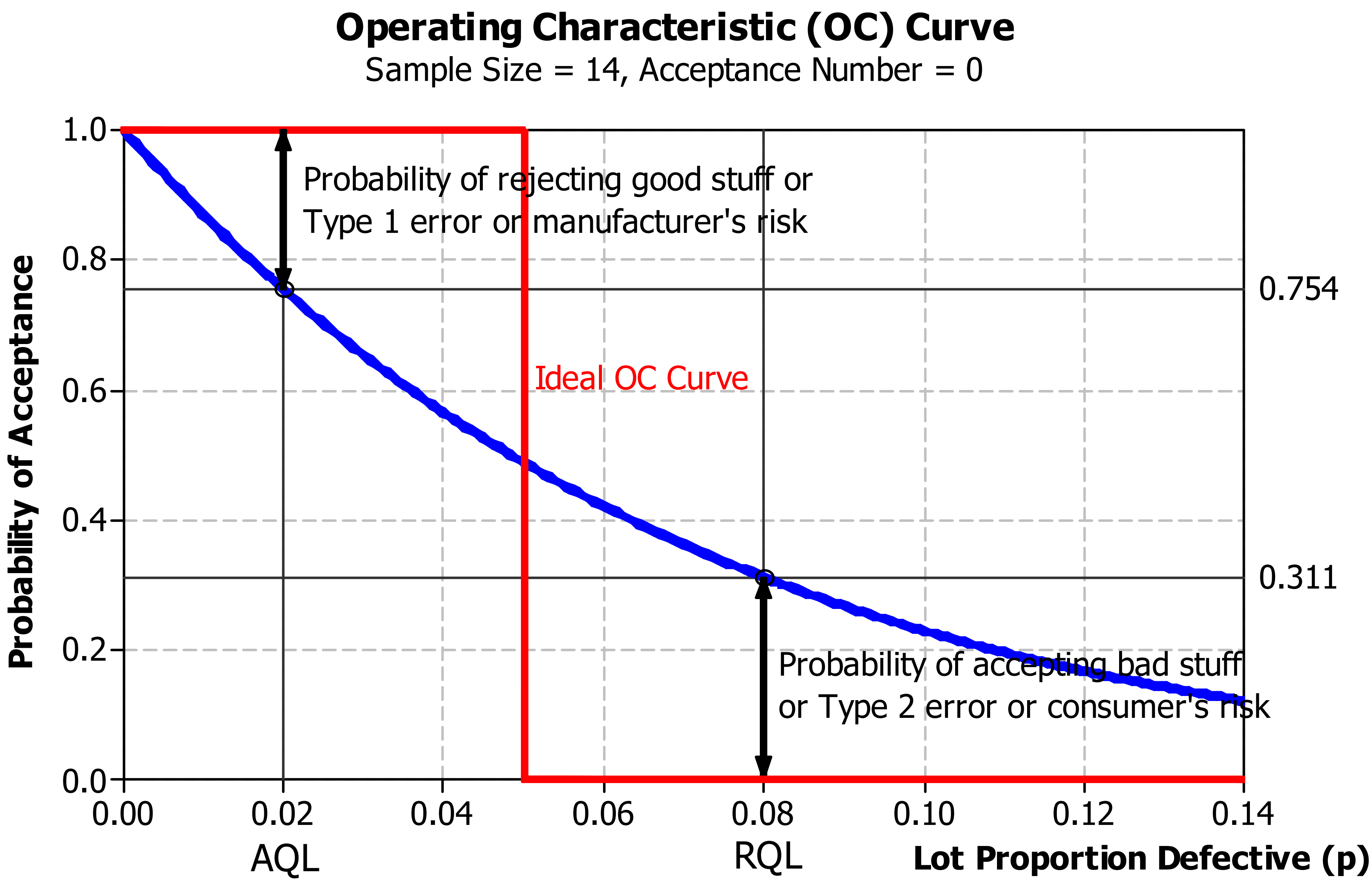
$$
\begin{aligned}
b(c = 2; n = 80, p = 0.03) &= \sum_{x=0}^{c} b(x; n, p) \\
&= \binom{80}{0} 0.03^0 (1 - 0.03)^{80-0} \\
&\quad + \binom{80}{1} 0.03^1 (1 - 0.03)^{80-1} \\
&\quad + \binom{80}{2} 0.03^2 (1 - 0.03)^{80-2} \\
&= 0.0875 + 0.2164 + 0.2643 \\
&= 0.568
\end{aligned}
$$

The probability of shutting the process down is $P = 1 - 0.568 = 0.432$.

Acceptance Sampling for Defectives

- Acceptance sampling is the practice of deciding to accept or reject batches or lots of material on the basis of sample data.
- A single sampling plan for defectives (see the document in Appendix A) is characterized by a sample size (n) and acceptance number (c). If the number of defectives (D) is less than or equal to the acceptance number (c), then the lot is accepted. If $D > c$ then the lot is rejected.
- The performance of a sampling plan for defectives is determined by its *operating characteristic curve* or *OC curve* - a plot of the probability of accepting (P_A) the claim that a lot is good versus the actual process fraction defective (p).
- The OC curve for a sampling plan is given by $P_A = b(c;n,p)$.

Operating Characteristic Curves



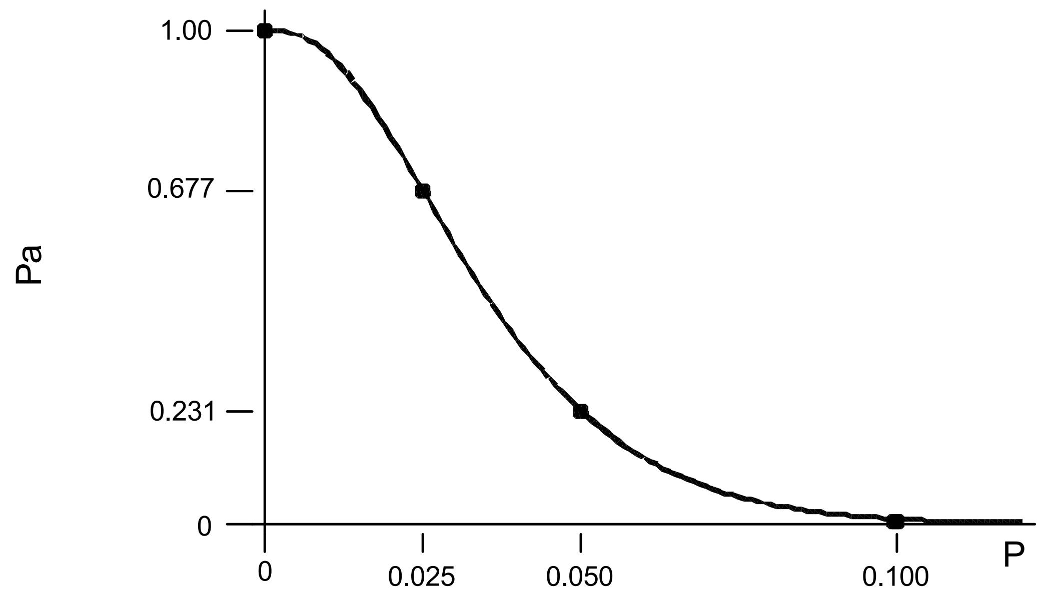
OC Curves

Example: Plot the OC curve for the sampling plan for defectives with $n = 100$ and $c = 3$.

Solution:

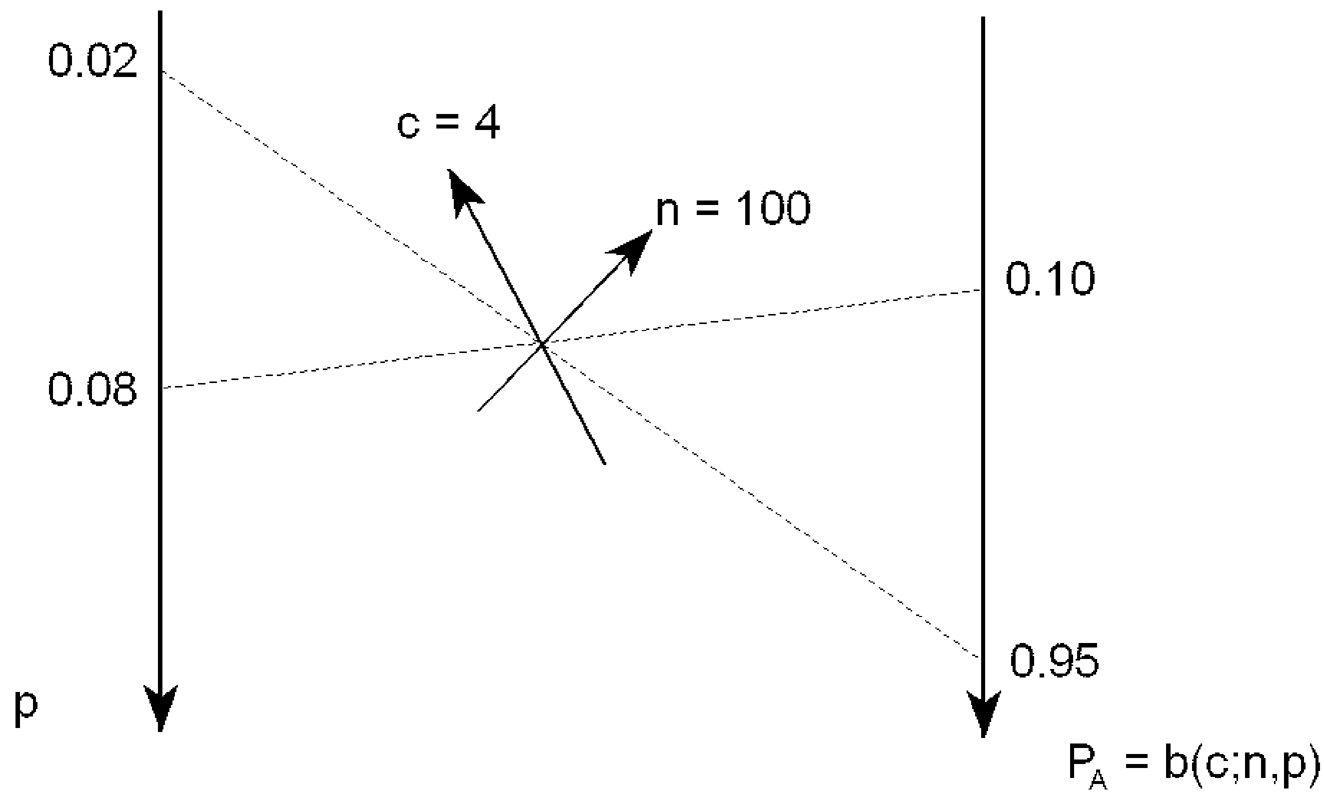
$Pa = b(c = 3; n = 100, p)$:	OC Curve for $n = 100$ and $c = 3$

p	Pa
0.0	1.00
0.025	0.677
0.050	0.231
0.100	0.008
0.145	0.0002

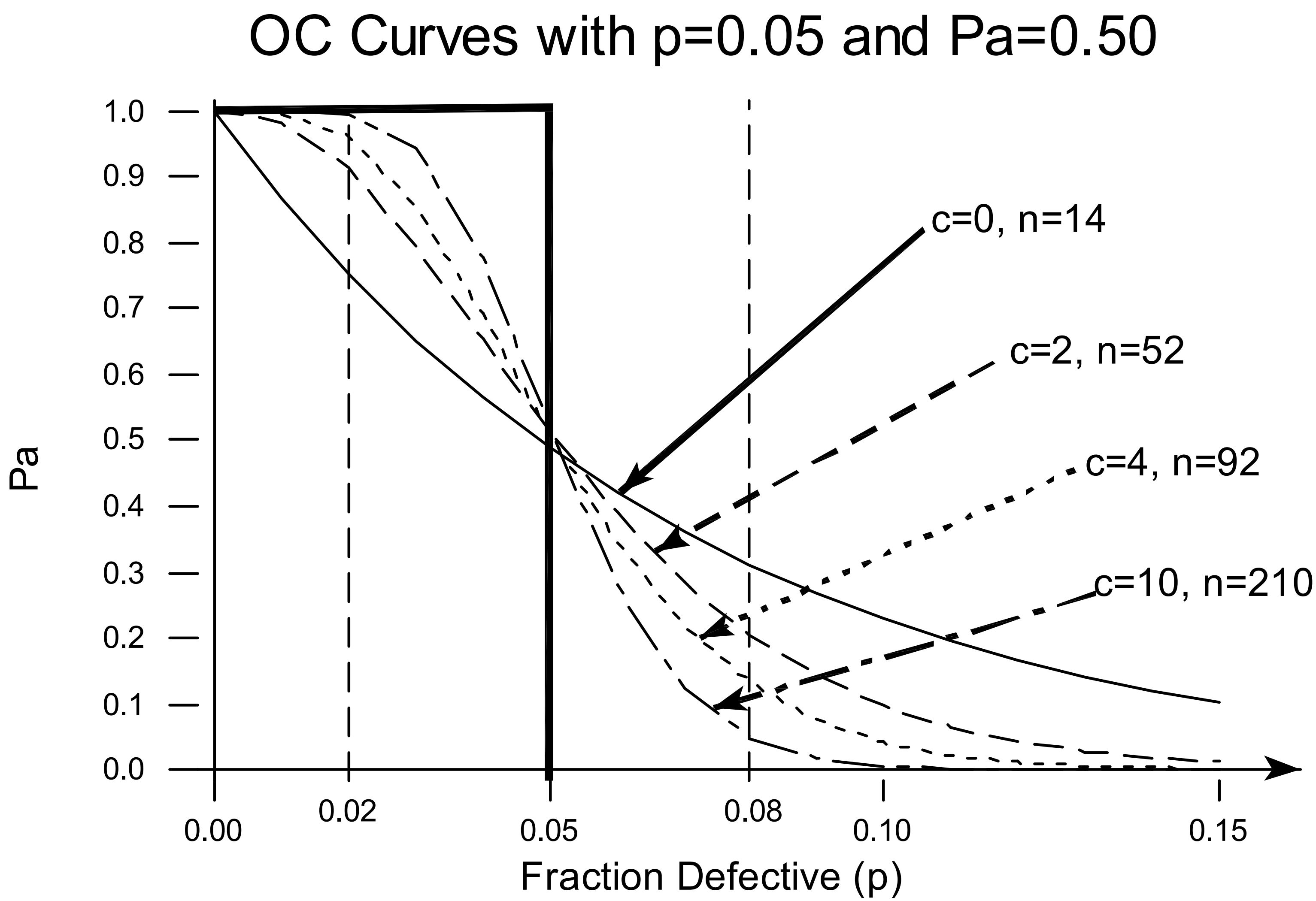


OC Curves

Use Larson's nomogram for the cumulative binomial distribution $b(c;n,p)$ to design sampling plans. See Appendix A: Design of Single Sampling Plans for Defectives.



OC Curves



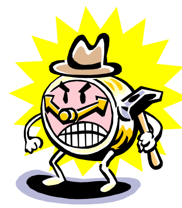
OC Curves with p=0.05 and Pa=0.50
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
Pa
c=0, n=14
c=2, n=52
c=4, n=92
c=10, n=210
0.00
0.02
0.05
0.08
0.10
0.15
Fraction Defective (p)

The Poisson Distribution

The Poisson distribution governs the number of successes that occur per inspection unit. The inspection unit, often referred to as the sampling unit, is defined by the user.

Examples of Sampling Units:

- A unit of time



- A unit of length
- A unit of area
- A unit of volume
- A car door

- 0.1 square meters of car door surface area
- 5 computer cases

The Poisson Distribution

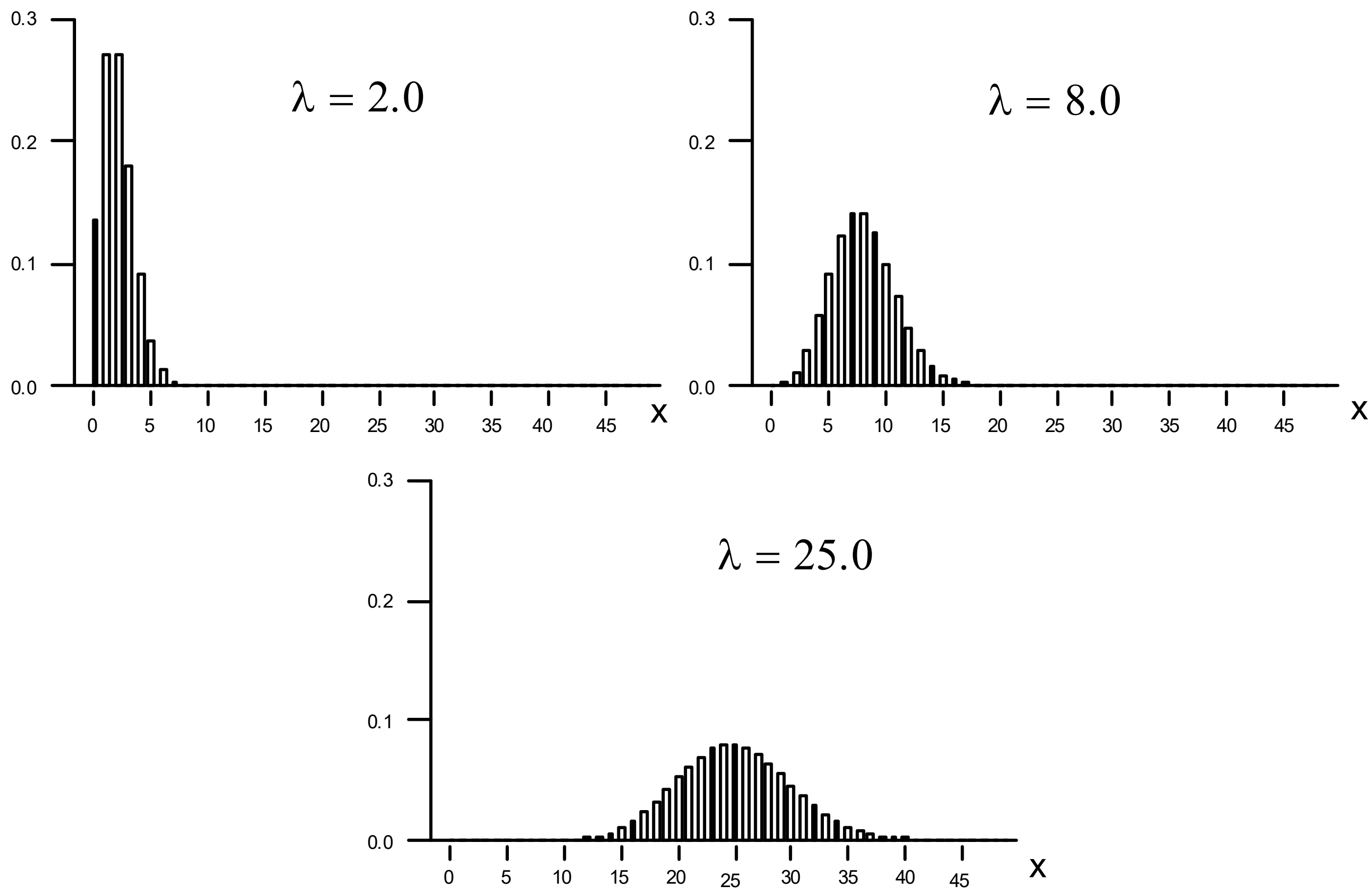
The probability of observing x events per sampling unit when the mean number of events per sampling unit is λ is given by:

$$Poisson(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$$

where $x = 0, 1, 2, 3, \ldots$

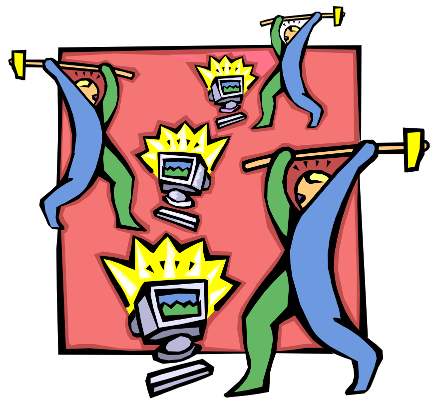
The Poisson distribution is used to track the number of defects that occur per unit something. The possible number of defects is unlimited.

The Poisson Histograms



The Poisson Distribution

Example: A certain Microsoft NT system crashes on average 2.4 times per month. What is the probability of the system running a month without crashing?



Solution: We have $x = 0$ and $\lambda = 2.4$ so:

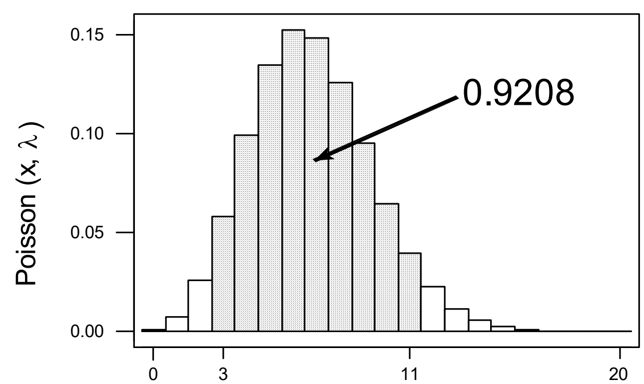
$$\begin{aligned} Poisson(x;\lambda) &= \frac{\lambda^x e^{-\lambda}}{x!} \\ Poisson(0;2.4) &= \frac{2.4^0 e^{-2.4}}{0!} = 0.091 \end{aligned}$$

The Poisson Distribution

Example: Samples of 3 lamps are drawn each hour and inspected for defects. The mean number of defects is known to be 6.8 per 3 lamp sampling unit. What is the probability of obtaining between 3 and 11 defects inclusive, (i.e. $3 \le x \le 11$) from this process?

Solution:

$$\begin{aligned} Poisson(3 \le x \le 11; \lambda = 6.8) &= Poisson(c = 11; 6.8) \\ &\quad - Poisson(c = 2; 6.8) \\ &= 0.9552 - 0.0344 = 0.9208 \end{aligned}$$



Approximations

- In many cases a probability distribution can be approximated by another distribution.
- The benefits of an approximation are:
  - The approximation is usually easier to calculate or the answer can be looked up in a table
  - The physical insight provided by an approximation can aid in understanding a problem.
- Approximations work when the two distributions are almost identical. This happens only when the parameters of the original distribution meet certain conditions.
- A table summarizing the approximations is included in the Appendix.

Binomial Approximation to the Hypergeometric Distribution

If a problem is fundamentally hypergeometric and the sample size is less than about 10% of the lot size then:

$$h(x; a, b, n) \simeq b\left(x; n, p = \frac{a}{a+b}\right)$$

Example: Calculate the exact and approximate solutions to $h(1;5,25,3)$.

Solution: The exact solution is:

$$h(x;a,b,n) = \frac{\binom{a}{x}\binom{b}{n-x}}{\binom{a+b}{n}}$$

$$h(1;5,25,3) = \frac{\binom{5}{1}\binom{25}{2}}{\binom{30}{3}}$$

$$= \frac{(5)(300)}{(4060)}$$

$$= 0.369$$

Since the sample size $n = 3$ is exactly 10% of the population size $N = 30$ we can use the binomial approximation to the hypergeometric distribution:

$$
\begin{aligned}
h(x; a, b, n) &\simeq b\left(x; n, p = \tfrac{a}{a+b}\right) \\
h(1; 5, 25, 3) &\simeq b(1; 3, \tfrac{5}{30} = 0.167) \\
&\simeq \left(\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right)(0.167)^1(1 - 0.167)^{3-1} \\
&\simeq (3)(0.167)(0.833)^2 \\
&\simeq 0.347
\end{aligned}
$$

Poisson Approximation to the Binomial Distribution

If a problem is fundamentally binomial and n is large and p is small then:

$$b(x; n, p) \simeq Poisson(x; \lambda = np)$$

Common guidelines for n and p are:

$$n \geq 100 \text{ AND } p \leq 0.10$$

Example: Calculate the exact and approximate solutions to $b(c = 4; 200, 0.05)$.

Solution: The exact solution from MINITAB is

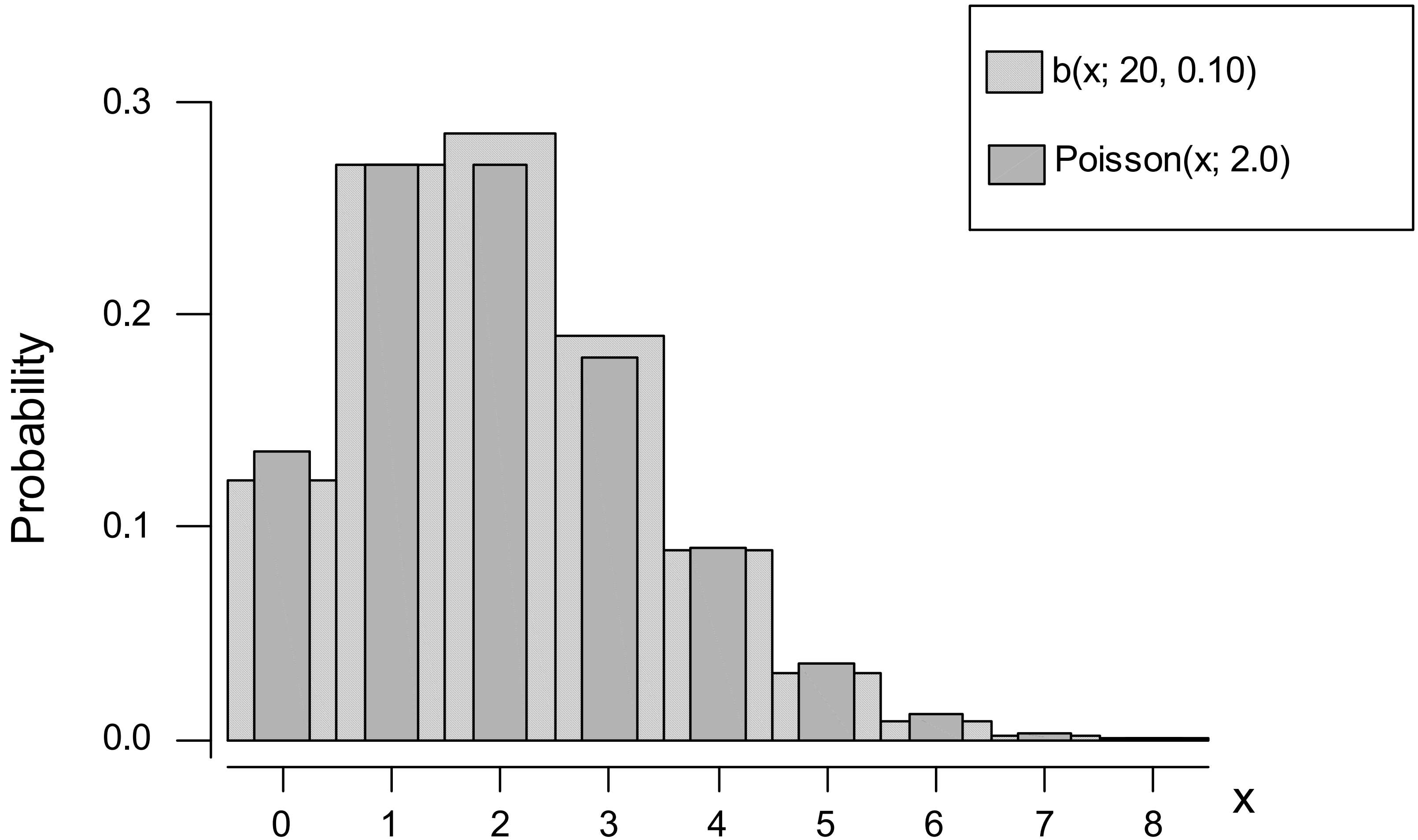
```
MTB > CDF 4;
SUBC> Binomial 200 0.05.
```

Cumulative Distribution Function

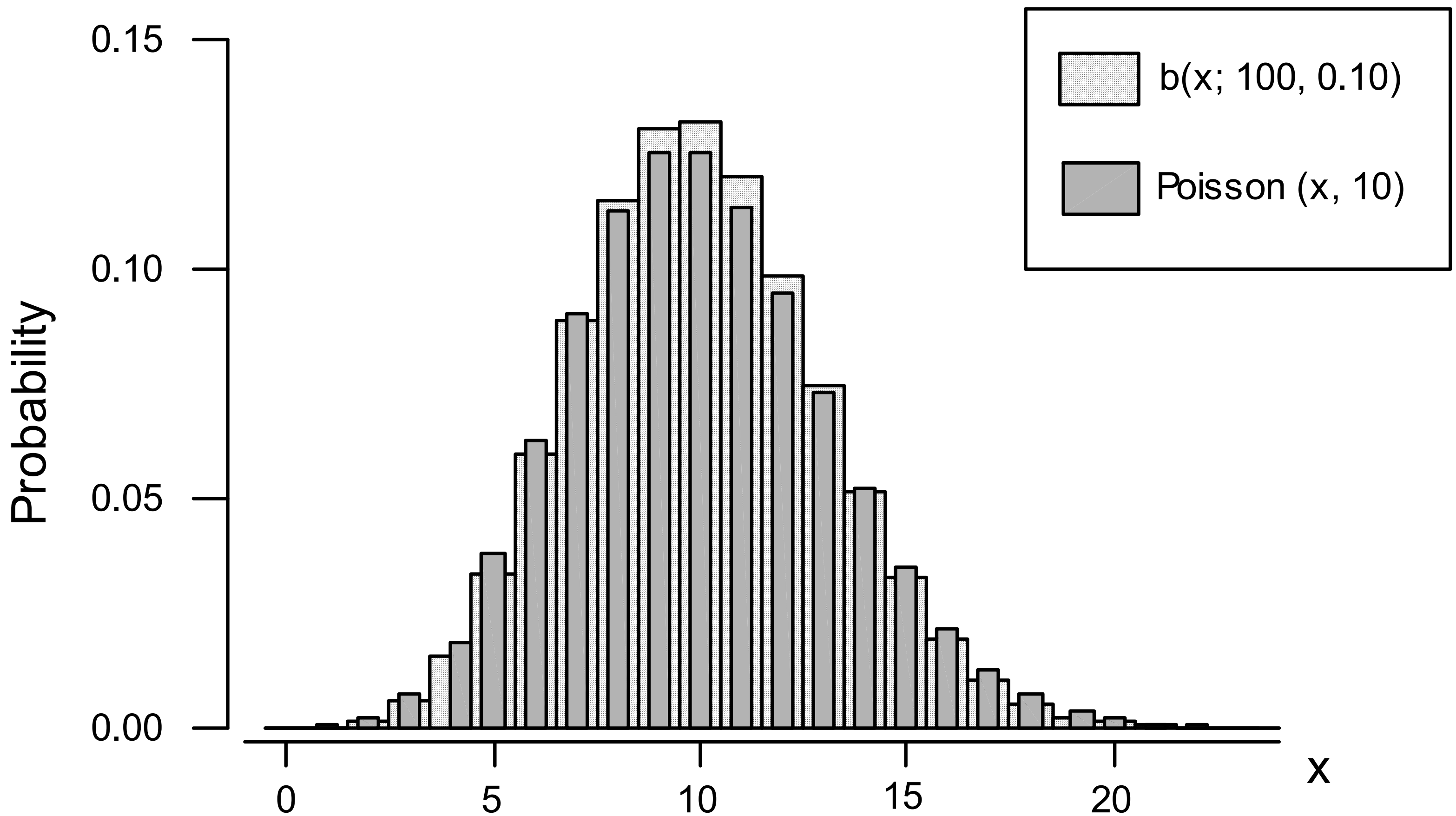
```
Binomial with n = 200 and p = 0.0500000

        x     P( X <= x )
     4.00          0.0264
```

Since n is large and p is small the Poisson approximation to the binomial distribution is:

$$
\begin{aligned}
b(c; n, p) &\simeq Poisson(c; \lambda = np) \\
b(c = 4; n = 200, p = 0.05) &\simeq Poisson(c = 4; \lambda = 10) \\
&\simeq 0.0293
\end{aligned}
$$



b(x; 20, 0.10)
Poisson(x; 2.0)
0.3
0.2
0.1
0.0
Probability
0
1
2
3
4
5
6
7
8
x


0.15
0.10
0.05
0.00
Probability
b(x; 100, 0.10)
Poisson (x, 10)
0
5
10
15
20
x

Where are the Different Distributions Used?

Hypergeometric:

- Defectives (or nonconforming units)
- Small lots

Binomial:

- Defectives (or nonconforming units)
- Continuous processes
- p and np charts
- As approximation to hypergeometric distribution
- MIL-STD-105 (ANSI/ASQ Z1.4)

Poisson:

- Defects
- c and u charts
- Anytime there are ”events per unit something”
- As approximation to binomial distribution
- Rare events

Attribute Distribution Quiz

Determine which distribution, binomial or Poisson, applies:

- Number of heads obtained in coin tosses
- Number of ones obtained rolling six dice
- Number of cavalry officers killed by horse kicks
- The number of employees late to work in the morning
- The number of cars in the parking lot that won't start at the end of the day
- The number of phone calls processed daily by customer service
- Number of radioactive decays
- The number of days that a delivery is late
- Number of mistakes on purchase orders
- Unresolved calls for customer service
- Incomplete patient admissions forms
- Number of colony forming units on a petri dish

Attribute Distribution Quiz

Determine which distribution, binomial or Poisson, applies:

- Number of stones thrown by the truck in front of you on the freeway
- Daily counts of golf balls lost by all players on the golf course
- Number of golfers who lose their temper each day
- Number of mutations on a length of DNA
- Number of people in line at the grocery store
- Number of TV channel changes per hour
- Number of pieces of junk mail relative to total mail on a daily basis.
- Number of red stop lights that you hit using your normal route home.
- Number of flat tires over the life of the car.
- Number of phone calls per day.
- Number of phone calls from family members per day.

The Mean of a Probability Distribution

The mean (μ) of a probability distribution is defined as the expected x value $(E(x))$:

$$\mu = E(x) = \sum_{\text{all } x} xf(x)$$

where $f(x)$ is the probability distribution of x.

Example: Find the mean of the binomial distribution $b(x;n,p)$.

Solution: From the definition above:

$$\mu = \sum_{x=0}^{n} x b(x;n,p) = np$$

Example: Find the mean of the distribution $b(x;20,0.10)$.

Solution: The required mean is:

$$\begin{aligned} \mu &= np \\ &= 20 \times 0.10 \\ &= 2.0 \end{aligned}$$

The Variance of a Probability Distribution

The variance (σ^2) of a probability distribution is defined as the expected ϵ^2 value $(E(\epsilon^2))$ where $\epsilon = x - \mu$:

$$\sigma^2 = E(\epsilon^2) = \sum_{\text{all } x} \epsilon^2 f(x)$$

Example: Find the standard deviation of the binomial distribution $b(x; n, p)$.

Solution: From the definition above:

$$\sigma^2 = \sum_{x=0}^{n} (x - \mu)^2 b(x; n, p) = np(1 - p)$$

where $\mu = np$ so the standard deviation of the binomial distribution is:

$$\sigma = \sqrt{np(1 - p)}$$

The Standard Deviation of a Probability Distribution

Example: Find the standard deviation of $b(x; 20, 0.10)$.

Solution:

$$\begin{aligned} \sigma &= \sqrt{np(1-p)} \\ &= \sqrt{20 \times 0.10 \times (1-0.10)} \\ &= 1.34 \end{aligned}$$

Summary of Means and Standard Deviations for Probability Distributions

Distribution	Function	Mean (μ)	Standard Deviation (σ)
Hypergeometric	$h(x;a,N,n)$	$\frac{na}{N}$	$\sqrt{n(\frac{a}{N})(1-\frac{a}{N})(1-\frac{n-1}{N-1})}$
Binomial	$b(x;n,p)$	np	$\sqrt{np(1-p)}$
Poisson	$Poisson(x,\lambda)$	λ	$\sqrt{\lambda}$

MINITAB Commands

- Calc> Probability Distributions> Hypergeometric
- Calc> Probability Distributions> Binomial
- Calc> Probability Distributions> Poisson
- Stat> Quality Tools> Acceptance Sampling by Attributes
- Graph> Probability Distribution Plot

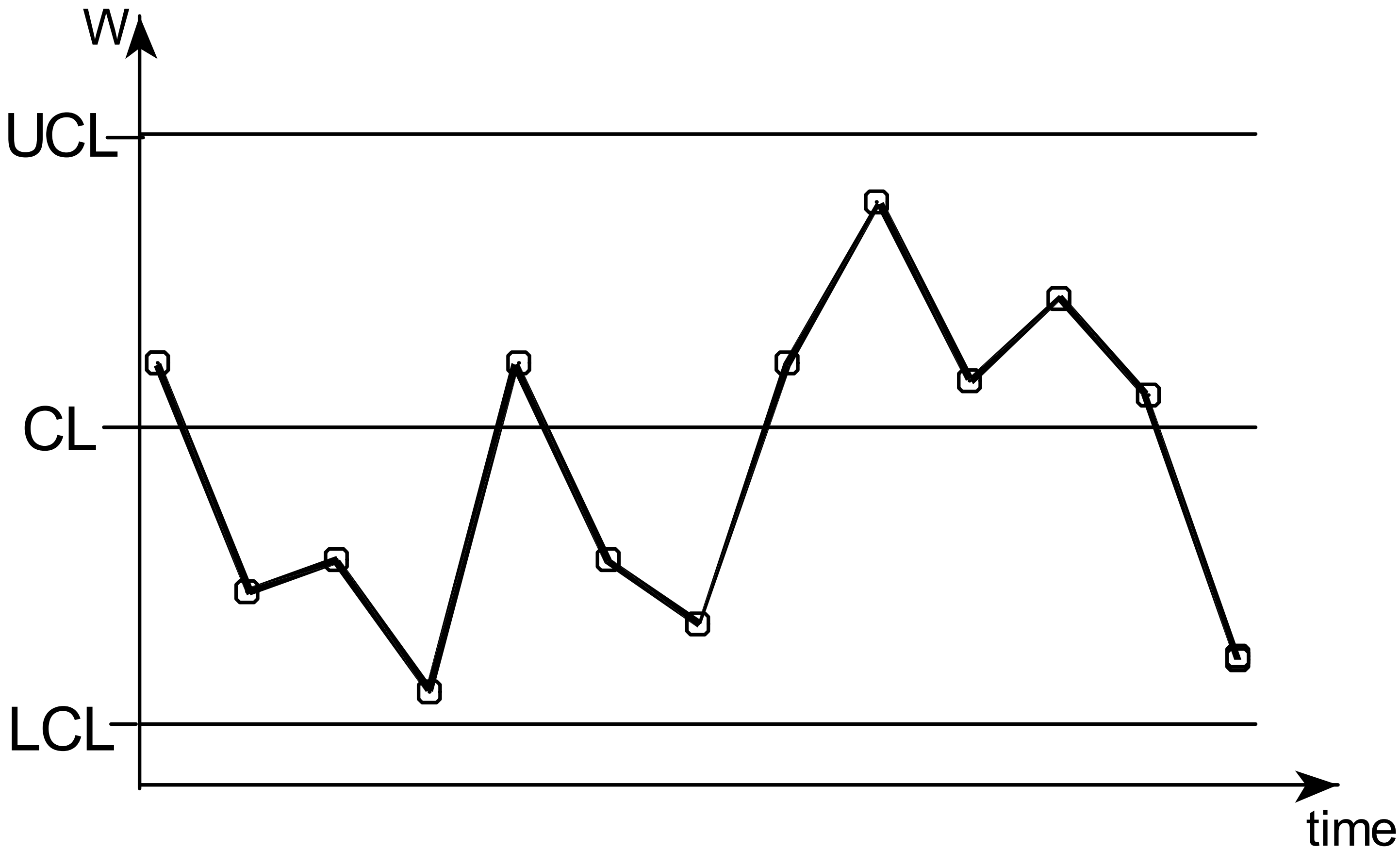
An Introduction to Statistical Process Control Charts for Defectives and Defects

Chart For:	n Constant	n Variable
Defectives	np	p
Defects	c	u

Shewhart Charts

- All of the traditional SPC control charts (np, p, c, u, $\bar{x}$ and R, $\bar{x}$ and s, x and MR) are classified as Shewhart charts.
- All Shewhart charts plot a process statistic versus time.
- All Shewhart charts use a center line (CL), an upper control limit (UCL), and a lower control limit (LCL).
- If the process statistic being measured is w, then the center line is positioned at $CL_w = \mu_w$.
- The upper and lower control limits are positioned at $(UCL/LCL)_w = \mu_w \pm 3\sigma_w$.

Shewhart Control Chart



Operating a Shewhart Chart

Assuming that the control limits have already been established...

1. Collect the required sample from the process. (This is called a subgroup. Subgroups are drawn at regular time intervals.)
2. Inspect the subgroup and determine w.
3. Add the new value of w to the chart and **note any changes that might affect the process**.
4. Connect successive w points on the chart with straight lines.
5. Inspect the chart to see if w falls outside the control limits or if any special patterns of points (called run rules) are present.

Operating a Shewhart Chart

6. If w falls within the control limits and none of the run rules are observed the process must be left alone.

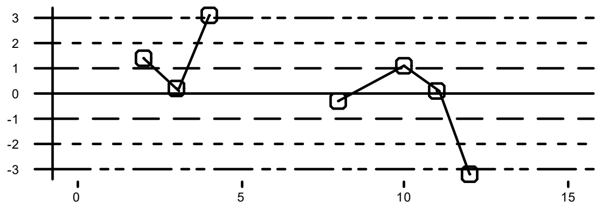
7. If w falls beyond one of the control limits or if one or more of the run rules are found then the process must be stopped or adjusted and the special cause must be identified.

   a. If a special cause is bad the process should be changed to prevent it from happening again.

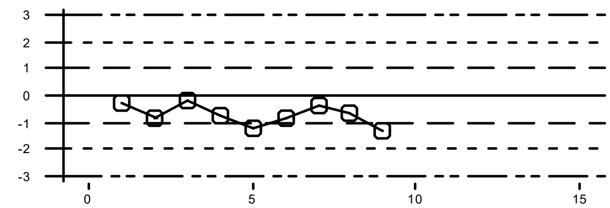
   b. If a special cause is good the process should be changed to make it permanent in the process.

8. Update the CL and UCL/LCL as the process improves.

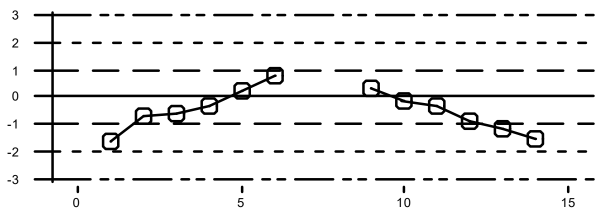
TEST 1: One point beyond $3\ \sigma$



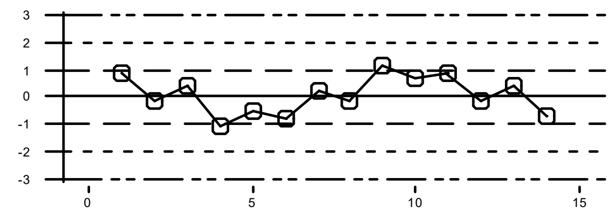
TEST 2: 9 points to one side of CL



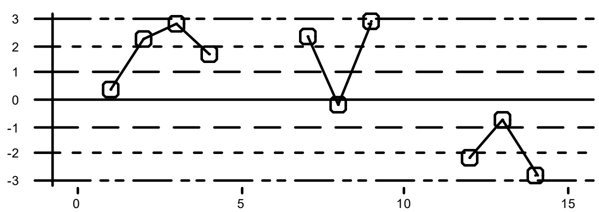
TEST 3: 6 point trend up or down



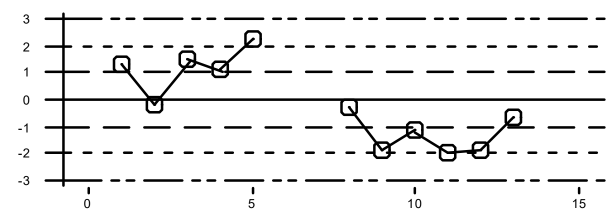
TEST 4: 14 points alternating up/down



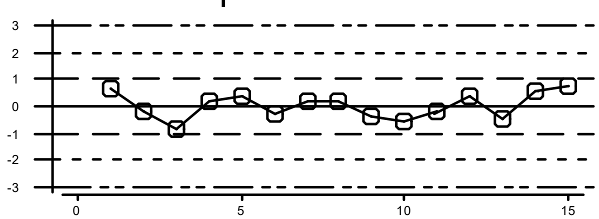
TEST 5: 2 of 3 points beyond 2σ



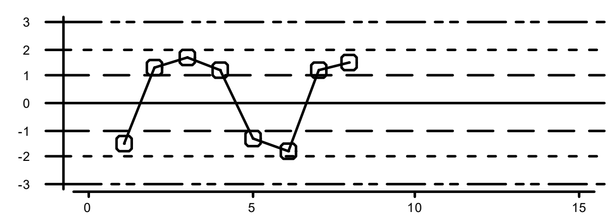
TEST 6: 4 out of 5 points beyond 1σ



TEST 7: 15 points within +/-1σ



TEST 8: 8 points beyond +/-$1\ \sigma$

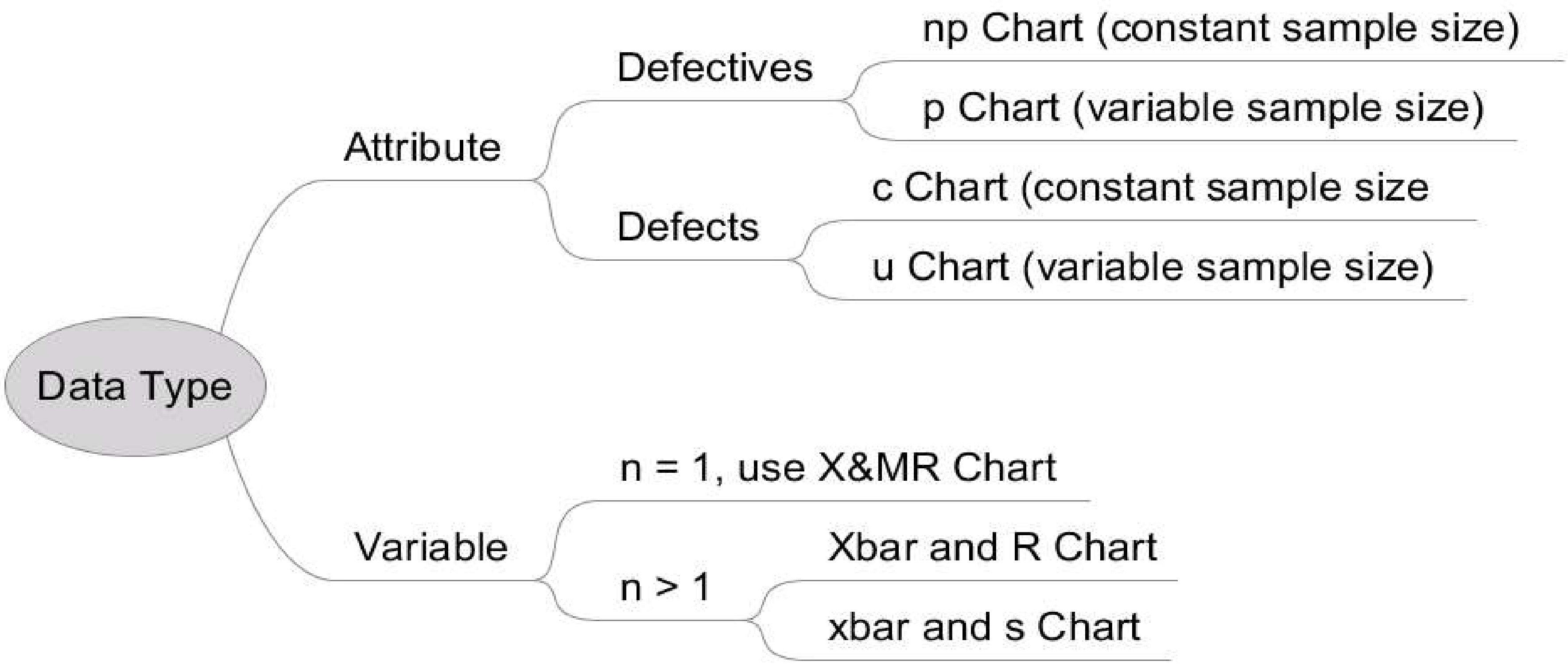


Created by: Rebecca Malnar 9/12/99

Defectives vs. Defects

- A part is either defective or it is not defective.
- A part can have many defects.
- A part with defects might not be a defective part.
- One critical defect can make a part defective.
- If the sample drawn is of size n, then the largest number of defectives possible is n.
- If the sample drawn is of size n, then the largest number of defects possible is ∞.

How to Choose a Control Chart



Defectives (np) Charts

- Defectives (np) charts use a fixed sample size n.
- The number of defectives (D) are plotted versus time.
- The binomial distribution governs how the data behave. Recall that $\mu = np$ and $\sigma = \sqrt{np(1-p)}$.
- If the process has fraction defective p then the chart's center line will be:

$$CL_{np} = \mu = np$$

with control limits:

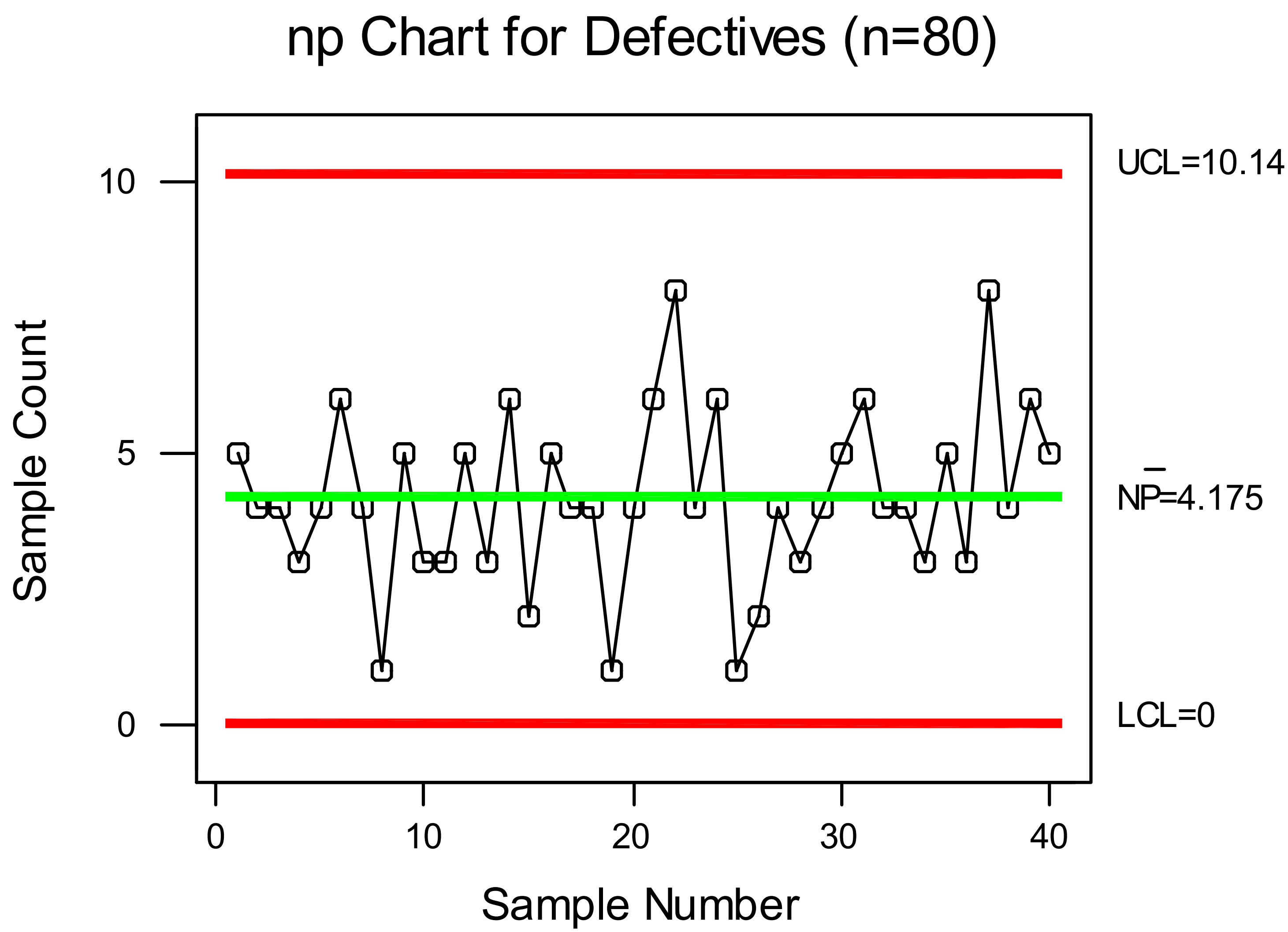
$$\begin{aligned} (UCL/LCL)_{np} &= \mu \pm 3\sigma \\ &= np \pm 3\sqrt{np(1-p)} \end{aligned}$$

Example: Find the center line and control limits for a defectives chart which has $n = 100$ and $p = 0.10$.

Solution:

$$\begin{aligned} CL_{np} &= np \\ &= 100 \times 0.10 \\ &= 10 \end{aligned}$$

$$\begin{aligned} (UCL/LCL)_{np} &= \mu \pm 3\sigma \\ &= np \pm 3\sqrt{np(1-p)} \\ &= 100(0.10) \pm 3\sqrt{100(0.10)(1-0.10)} \\ &= 10 \pm 9 \\ &= 19/1 \end{aligned}$$


np Chart for Defectives (n=80)
UCL=10.14
NP=4.175
LCL=0
Sample Count
Sample Number
10
5
0
0
10
20
30
40

Specification Limits versus Control Limits

- Specification limits apply to individual parts.
- Control limits apply to statistics calculated from random samples.
- Specification limits generally fall outside of control limits.
- Specification limits and control limits **never** belong on the same graph except in the case of an individuals chart.

Understanding Variation

All variation is caused:

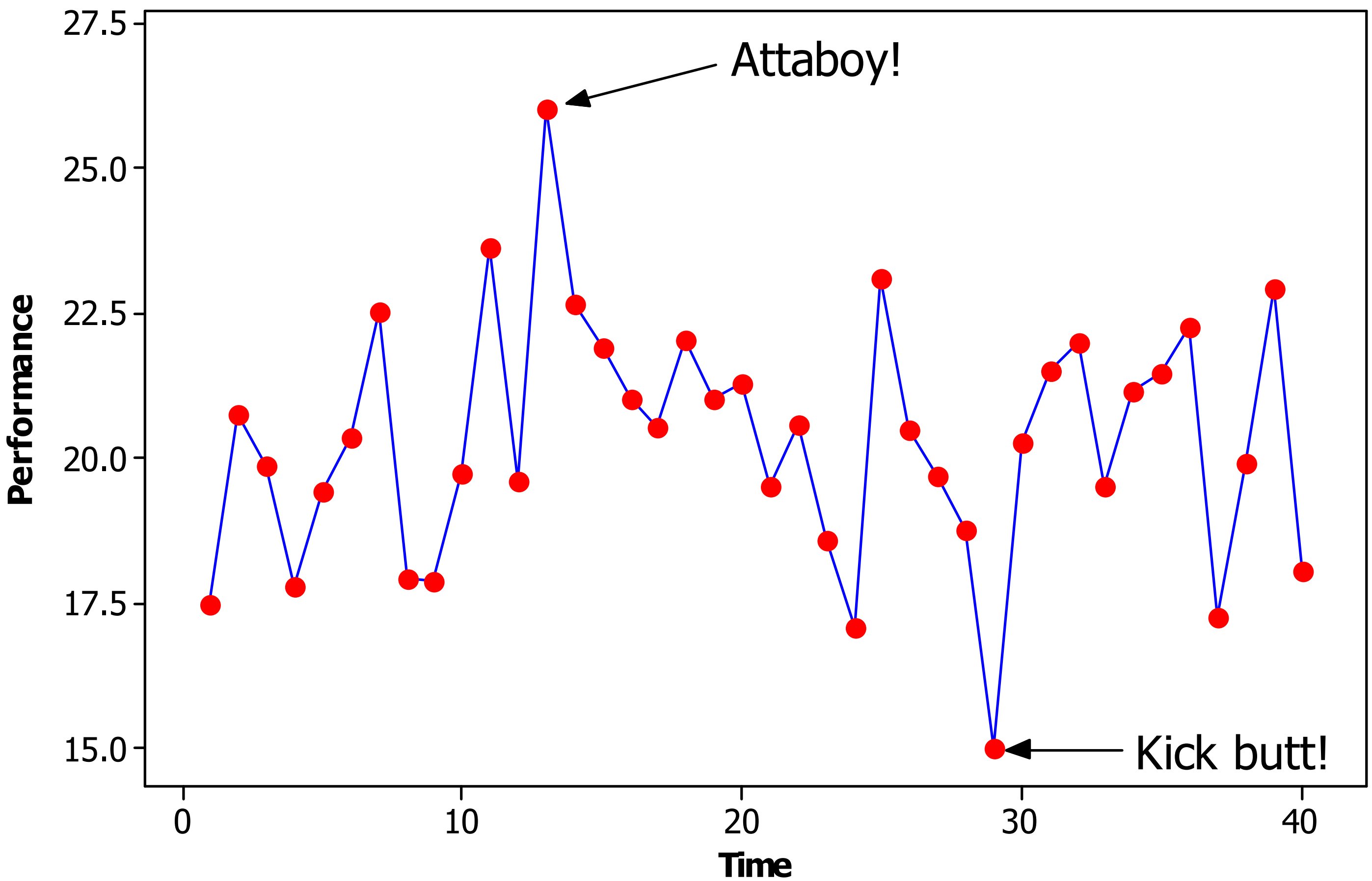
- Common cause variation - many small generally uncontrollable and individually insignificant sources of variation
- Special cause variation- those few causes with very large, obvious, and usually traumatic effects on the process
- Tampering variation - variation induced by inappropriate attempts to control common cause variation
- Structural variation - variation caused by predictable sources, such as seasonal fluctuations, arbitrary deadlines, etc.

Understanding Variation

The different types of variation are managed in different ways:

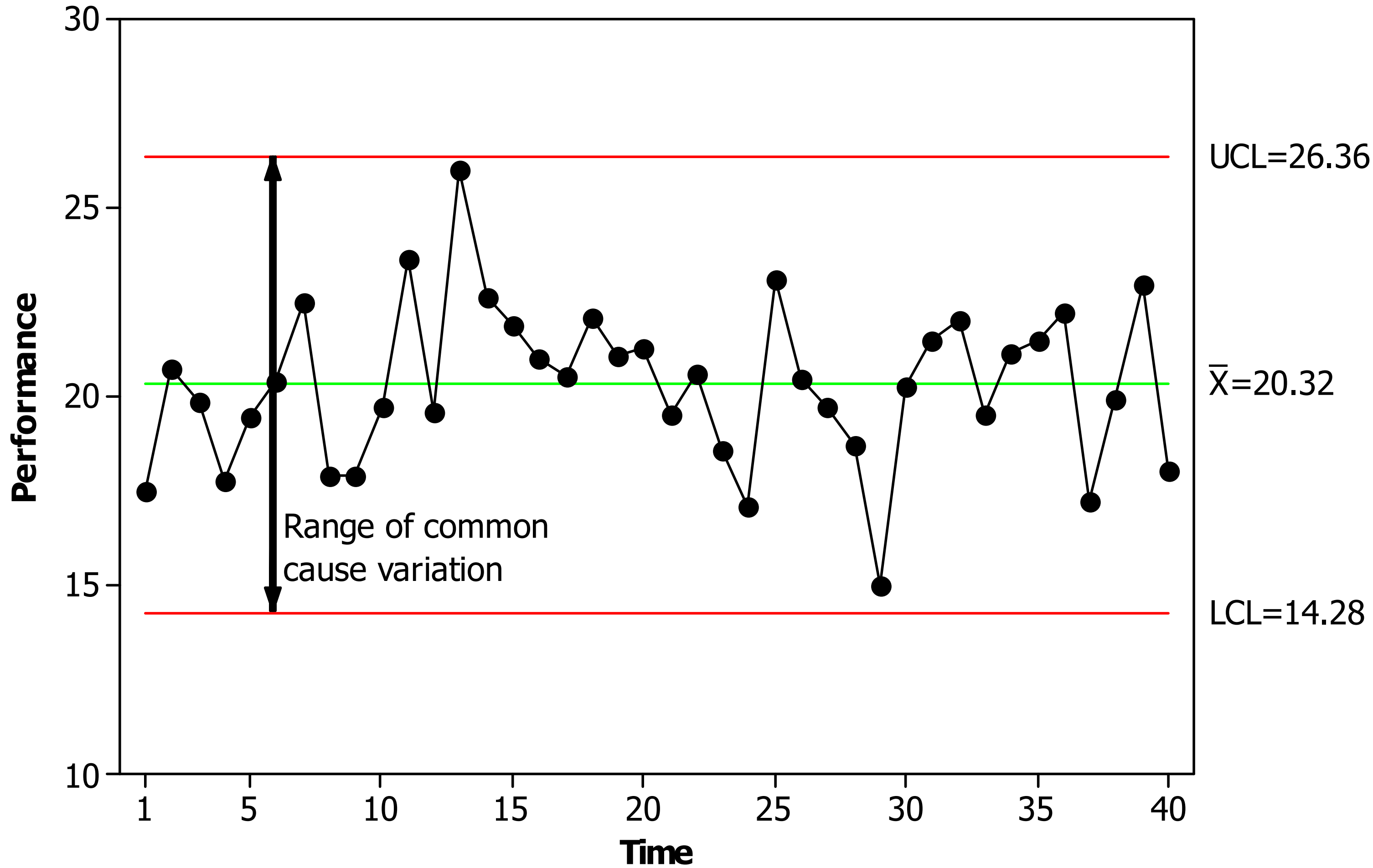
- Common cause variation:
  - Can’t do much about it.
  - Hands off!
- Special cause variation:
  - Take immediate action.
  - Identify the special cause and
    - If it is bad, change the process so that it never happens again.
    - If it is good, make the special cause a permanent part of the process.
- Tampering variation:
  - Recognized it.
  - Eliminate it.
- Structural variation:
  - Normalize the data.

Why Does My Boss Always Yell And Scream And Never Give Me Praise?



The effect is called *regression to the mean*.

Why Does My Boss Always Yell And Scream And Never Give Me Praise?



What Control Charts Do

- Control charts permit the process owner to decide if:
  - The process is in control, i.e. only common causes are present.
  - The process is out of control, i.e. a special cause is present.
- When a data point falls within the control limits and none of the run rules are turned on, the process owner accepts the first statement as true and leaves the process alone.
- When a data point falls outside the control limits or if one of the run rules is turned on, the process owner accepts the second statement as true and takes appropriate action to resolve the problem.

p Charts

- When defectives data are being collected but the sample size is variable you cannot use an *np* chart.
- Use a *fraction defective* or *p* chart instead. The *p* chart tracks the sample fraction defective:

$$p_i = \frac{D_i}{n_i}$$

  where D_i is the number of defective parts found on the *i*th sample and n_i is the sample size.
- Be careful to distinguish between the parameter p and the statistic p_i.

p Charts

- The p chart is constructed by plotting the p_i vs. time.
- The center line of the chart will be:

$$CL_p = p$$

with control limits:

$$(UCL/LCL)_p = p \pm 3\sqrt{\frac{p(1-p)}{n_i}}$$

where n_i is the size of the ith sample.

p Charts

- Notice that the center line of the *p* chart is fixed but the *UCL* and *LCL* move in towards the *CL* as the sample size increases. This makes it difficult to interpret *p* charts and there are special techniques and training to do this.
- Whenever possible try to use a fixed sample size so that you can use a defectives chart (*np*) instead of the *p* chart. Defectives charts are much easier for operators to maintain and interpret than *p* charts.

Example: Find the control limits and create the control chart for the following defectives data.

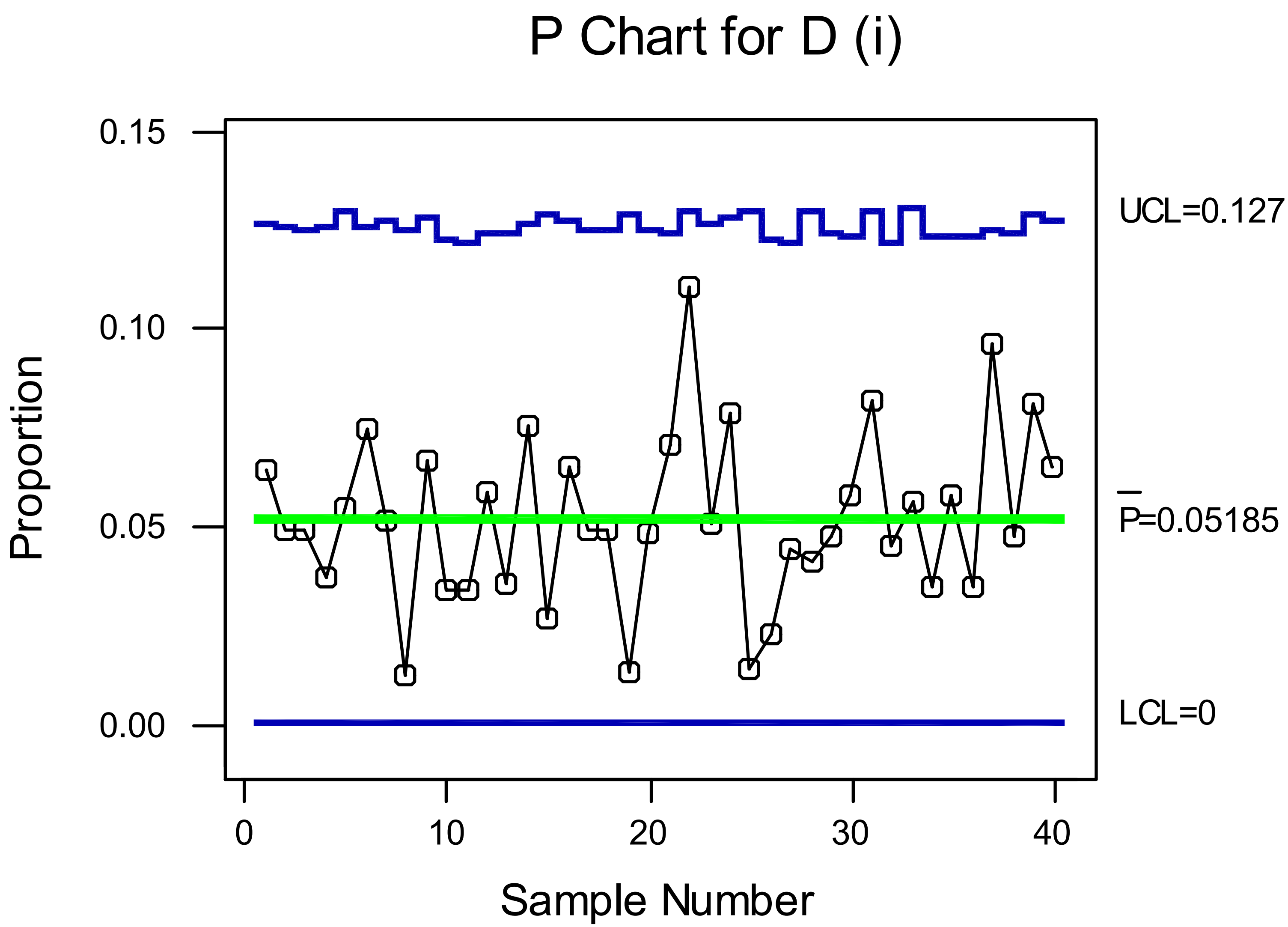
Solution:

time (i)	N(i)	D (i)	P(i)
1	78	5	0.064
2	81	4	0.049
3	82	4	0.049
4	81	3	0.037
5	73	4	0.055
6	80	6	0.075
7	77	4	0.052
8	83	1	0.012
9	75	5	0.067
10	88	3	0.034
38	84	4	0.048
39	74	6	0.081
40	77	5	0.065

$$\sum_{i=1}^{40} n = 3221$$

$$\sum_{i=1}^{40} D = 167$$

$$\bar{p} = \frac{\sum D}{\sum n} = \frac{167}{3221} = 0.0519$$



P Chart for D (i)
0.15
0.10
0.05
0.00
Proportion
UCL=0.127
P=0.05185
LCL=0
0
10
20
30
40
Sample Number

p Charts

There are special methods for operating p charts when the sample size is variable:

- Use individual limits for each p_i.
- Use approximate limits determined from $\bar{n}$ or $n_{\max}$.
- Use multiple limits, e.g. for $n = 100, 200, 300$.
- Standardize the chart by plotting:

$$z_i = \frac{p_i - p}{\sqrt{\frac{p(1-p)}{n_i}}}$$

with $CL = 0$ and $UCL/LCL = \pm 3$.

What If I Don't Know p?

If you are setting up a new np or p chart and don't know p:

1. Begin collecting and plotting the data (D_i vs. $time$).
2. Log any unusual circumstances or observations.
3. After 20 and preferably 30 subgroups have been inspected calculate the CL and UCL/LCL from $\bar{p}$ where:

$$\bar{p} = \frac{\sum D_i}{\sum n_i}$$

4. Use the control limits to check the original subgroups for special causes. If a point is out of control it is OK to eliminate it, **but only if you can identify its special cause**.
5. Use the revised control limits to decide how to manage future subgroups.
6. Revise the control limits as more data become available and process improvements are implemented.

Example: $m = 30$ subgroups of size $n = 80$ were collected in preparation for using a new defectives chart. The total number of defectives found in the 30 subgroups was:

$$\sum_{i=1}^{30} D_i = 144$$

Find the preliminary CL_{np} and $(UCL/LCL)_{np}$.

Solution: The total number of parts inspected was:

$$\begin{aligned} \textstyle\sum_{i=1}^{30} n_i &= m \times n \\ &= 30 \times 80 = 2400 \end{aligned}$$

From these preliminary data the mean fraction defective $\bar{p}$ is:

$$\begin{aligned} \bar{p} &= \frac{\sum D_i}{\sum n_i} \\ &= \frac{144}{2400} = 0.060 \end{aligned}$$

The center line for the defectives chart is:

$$
\begin{aligned}
CL_{np} &= n\bar{p} \\
&= 80 \times 0.060 = 4.8
\end{aligned}
$$

The $(UCL/LCL)_{np}$ are:

$$
\begin{aligned}
(UCL/LCL)_{np} &= n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})} \\
&= 80 \times 0.06 \pm 3\sqrt{80 \times 0.06(1-0.06)} \\
&= 4.8 \pm 6.4 \\
&= 11.2/0
\end{aligned}
$$

Sample Size for np and p Charts

Rules to use when determining sample size for np and p charts.

1. To obtain a lower control limit that is greater than or equal to zero:

$$n \geq \frac{9(1-p)}{p}$$

2. To detect a shift of the fraction defective from p to $p+\delta$ with 50% probability use:

$$n = \frac{9p(1-p)}{\delta^2}$$

3. To ensure that there won't be too many occurrences of 0 defectives ($\leq 5\%$) use:

$$n \geq \frac{3}{p}$$

The third rule is safe and is the easiest to use.

Defects Charts

- *Defects* charts or c charts are used to track the number of defects that are observed in sampling units of fixed size.
- Defects charts are analogous to defectives charts - both have constant sample size.
- Defects charts plot the number of defects found in each sampling unit (c_i) vs. time.
- Be careful to distinguish between the parameter c and the statistic c_i.
- The Poisson distribution governs how the data behave.
- The center line of a c chart is: $CL_c = c$
- The upper and lower control limits are:

$$(UCL/LCL)_c = c \pm 3\sqrt{c}$$

Example: Find the CL and UCL/LCL for a c chart when the process has an average of 14.4 defects per sampling unit.

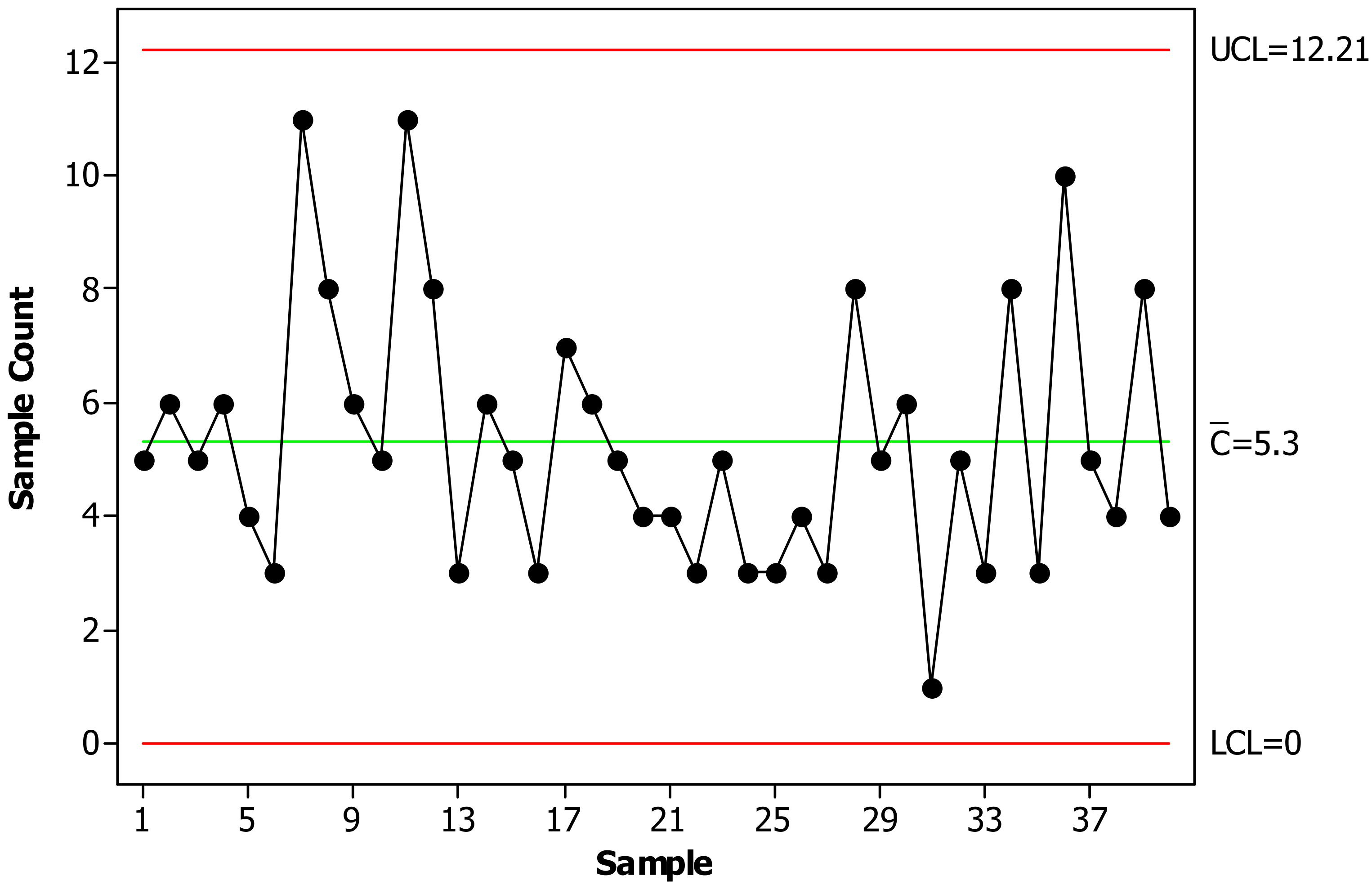
Solution: Since the mean number of defects per sampling unit is $c = 14.4$ we have:

$$CL_c = c = 14.4$$

and

$$\begin{aligned} UCL/LCL &= 14.4 \pm 3\sqrt{14.4} \\ &= 14.4 \pm 11.4 \\ &= 25.8/3.0 \end{aligned}$$

Because the number of defects is an integer response it is common to express these control limits as 25 and 3; however, it is wise to draw the control limits on the control chart at 25.5 and 2.5 to avoid confusion about what is in control and what is out of control.



UCL=12.21
$\bar{C}$=5.3
LCL=0
Sample Count
Sample
0
2
4
6
8
10
12
1
5
9
13
17
21
25
29
33
37

What if I Don't Know c?

If you are setting up a new c chart and don't know c:

1. Begin collecting and plotting the data (c_i vs. $time$).
2. Log any unusual circumstances or observations.
3. After $m = 20$ and preferably 30 subgroups have been inspected calculate the CL and UCL/LCL from $\bar{c}$ where:

$$\bar{c} = \frac{\sum c_i}{m}$$

4. Use the control limits to check the original subgroups for special causes. If a point is out of control it is OK to eliminate it, **but only if you can identify its special cause**.
5. Use the revised control limits to decide how to manage future subgroups.
6. Revise the control limits as more data become available and process improvements are implemented.

Example: $m = 25$ subgroups (or sampling units) of fixed size were collected in preparation for using a new defects chart. The total number of defects found in the 25 subgroups was:

$$\sum_{i=1}^{25} c_i = 80$$

Find the preliminary CL_c and $(UCL/LCL)_c$.

Solution: The mean number of defects per sampling unit from the preliminary data is:

$$\begin{aligned}\bar{c} &= \frac{\sum c_i}{m} \\ &= \frac{80}{25} \\ &= 3.2\end{aligned}$$

The center line for the defects charts is:

$$\begin{aligned} CL_c &= \bar{c} \\ &= 3.2 \end{aligned}$$

The $(UCL/LCL)_c$ are:

$$\begin{aligned} (UCL/LCL)_c &= \bar{c} \pm 3\sqrt{\bar{c}} \\ &= 3.2 \pm 3\sqrt{3.2} \\ &= 3.2 \pm 5.4 \\ &= 8.6/0 \end{aligned}$$

u Charts

- When defect data are being collected but the sampling unit size is variable you cannot use a c chart.
- Instead use a u chart which tracks the number of defects per unit inspected. The sample statistic u_i is given by:

$$u_i = \frac{c_i}{n_i}$$

where c_i is the number of defects found on the ith subgroup and n_i is the size of the ith subgroup.

u Charts

- The u chart is constructed by plotting the u_i vs. time.
- The center line of the chart will be:

$$CL_u = u$$

with control limits:

$$(UCL/LCL)_u = u \pm 3\sqrt{\frac{u}{n_i}}$$

where n_i is the size of the subgroup.

u Charts

- Notice that the center line of the u chart is fixed but the UCL and LCL move in towards the CL as the sample size increases. This makes it difficult to interpret u charts and there are special techniques and training to do this.
- Whenever possible try to use a fixed sample size so that you can use a defects chart (c) instead of the u chart. Defects charts are much easier for operators to maintain and interpret than u charts.

Sample Size for c and u Charts

There are several rules for determining sample size for c and u charts:

1. To obtain a zero or positive lower control limit use a sampling unit that is large enough so that $c \geq 9$.

   Example: Determine the sample size that will give a zero or positive LCL if there are 2.4 defects per part inspected, on average.
   Solution: If the sampling unit consists of 4 parts then the mean number of defects per sampling unit will be $c = 4 \times 2.4 = 9.6$ and the LCL will be positive. $\left(LCL = 9.6 - 3\sqrt{9.6} = 0.3\right)$

Sample Size for c and u Charts

2. To detect a shift of the mean defect rate from c to $c + \delta$ with 50% probability use use a sampling unit that is large enough so that $c \geq \delta^2/9$..

Example: Determine the sample size that will detect a shift from c to $c + 18$ with 50% probability if there are 2.4 defects per part inspected, on average.
Solution: The c chart's mean will need to be $c \geq 18^2/9 = 36$. The number of parts required in the sampling unit will be $n = 36/2.4 = 15$.

Sample Size for c and u Charts

3. To ensure that there won't be too many occurrences of 0 defects (less than 1 of 20 points plotted) use a sampling unit that is large enough so that $c \geq 3$.

Example: Determine the sample size that will ensure that no more than about 5% of the points plotted on a c chart will be 0s if there are 2.5 defects per part.
Solution: In order to obtain $c \geq 3$ it is necessary to use a sampling unit of 2 parts. This gives $c = 2 \times 2.5 = 5.0$.

Sensitizing Rules for Control Charts

- Predefined patterns or *run rules* on the control chart are used to determine when the process is out of control.
- The presence of a single run rule is sufficient to indicate that the process is in control.
- The first rule is to check where individual points fall with respect to the $\pm 3\sigma$ upper and lower control limits.
  - If a point falls between the $\pm 3\sigma$ limits the process is said to be in control.
  - If a point falls beyond the $\pm 3\sigma$ limits (either above or below) the process is said to be out of control.
- Other special patterns of points can indicate that the process is out of control.
- Many people think that the sensitizing rules are sacred - they are not.

Sensitizing Rules for Control Charts

A good run rule can be any pattern of points that meets the following conditions:

1. The pattern must be easy to recognize.
2. The pattern must have a low probability of occurring when the process is actually in control.
3. The pattern must have a high probability of occurring when the process is out of control.

The Western Electric Rules

One set of commonly used sensitizing rules that meets these conditions was developed by Walter Shewhart who invented control charts. These rules are called the Western Electric rules. To make the Western Electric rules easier to recognize it is common to draw $\pm 1\sigma$ and $\pm 2\sigma$ lines on the chart in addition to the $\pm 3\sigma$ control limits.

The Western Electric rules that indicate that a process is out of control are:

1. One point beyond the $\pm 3\sigma$ control limits.
2. At least two out of three consecutive points beyond the same 2σ limit.
3. At least four out of five consecutive points beyond the same 1σ limit.
4. Eight consecutive points to one side of the center line.

Special Notes about Sensitizing Rules

1. Decisions about the control state of the process can be incorrect.
   - **a**. It is possible that the process could be in control and one of the rules could be triggered by accident.
   - **b**. It is possible that the process could be out of control and none of the rules will be triggered.
   - **c**. The purpose of the rules is to make the operation of the chart as sensitive as possible without permitting too many of these false signals.

2. Often the addition of one data point to a control chart triggers several different sensitizing rules. The fact that several rules are triggered provides strong evidence that the process is really out of control.

Special Notes about Sensitizing Rules

3. It is dangerous to use too many sensitizing rules. Each rule has its own error rate and the error rates for the different rules are roughly additive. For example, if each rule in a set of 10 rules gives a false out of control signal 0.5% of the time, then the overall rate of false out of control signals will be about 5% or about 1 in 20 points plotted on the chart.

4. It is dangerous to use too many charts. An operator can maintain only 3-5 charts at a time, but a computer can keep hundreds of charts easily. Since each chart provides an opportunity for false out of control signals, the use of many charts will give rise to frequent false signals, especially if too many rules are being used on each chart.

SPC Chart Design

The ability of a chart to detect shifts in the process mean is determined by:

1. The values of the upper and lower control limits.
2. The sample size.
3. The sampling interval.
4. The runs rules used to identify out of control conditions.

These characteristics must be chosen so that the chart is sensitive enough to detect small shifts in the process mean before the size of the shifts become objectionable to the customer. (Remember that out-of-control doesn’t necessarily mean out-of-spec.) If these conditions are met and if the operator really uses the chart to control the process then the SPC chart is truly a prevention tool.

SPC Chart Lifetime

- When Shewhart wrote the rules of SPC, charts were kept using pencil and paper.
- Operators could only use a few run rules on a few charts.
- Charts were expected to have a finite lifetime:
    - A new chart would be created in reaction to a new problem.
    - The chart’s design and use would be standardized while it was in regular use.
    - Out of control signals on the chart would be used to find and eliminate their special causes.
    - The sample size and sampling frequency would be reduced as the problem was solved.
    - The chart would be eliminated from use after the improved process eliminated the problem.

Electronic versus Paper Control Charts

- My preference is for paper charts.
  - Operators feel more ownership of paper charts.
  - Operators are more likely to use the charts.
  - More likely to annotate with process changes.
  - Less likely to use too many charts.
  - Less likely to use too many run runs.
  - More likely to keep the charts up to date.
  - Easier to post more than one important chart at a time so that they can all be seen easily.
  - More likely to detect an out of control event early.
- Use electronic charts if the time required to maintain paper charts is hurting productivity.

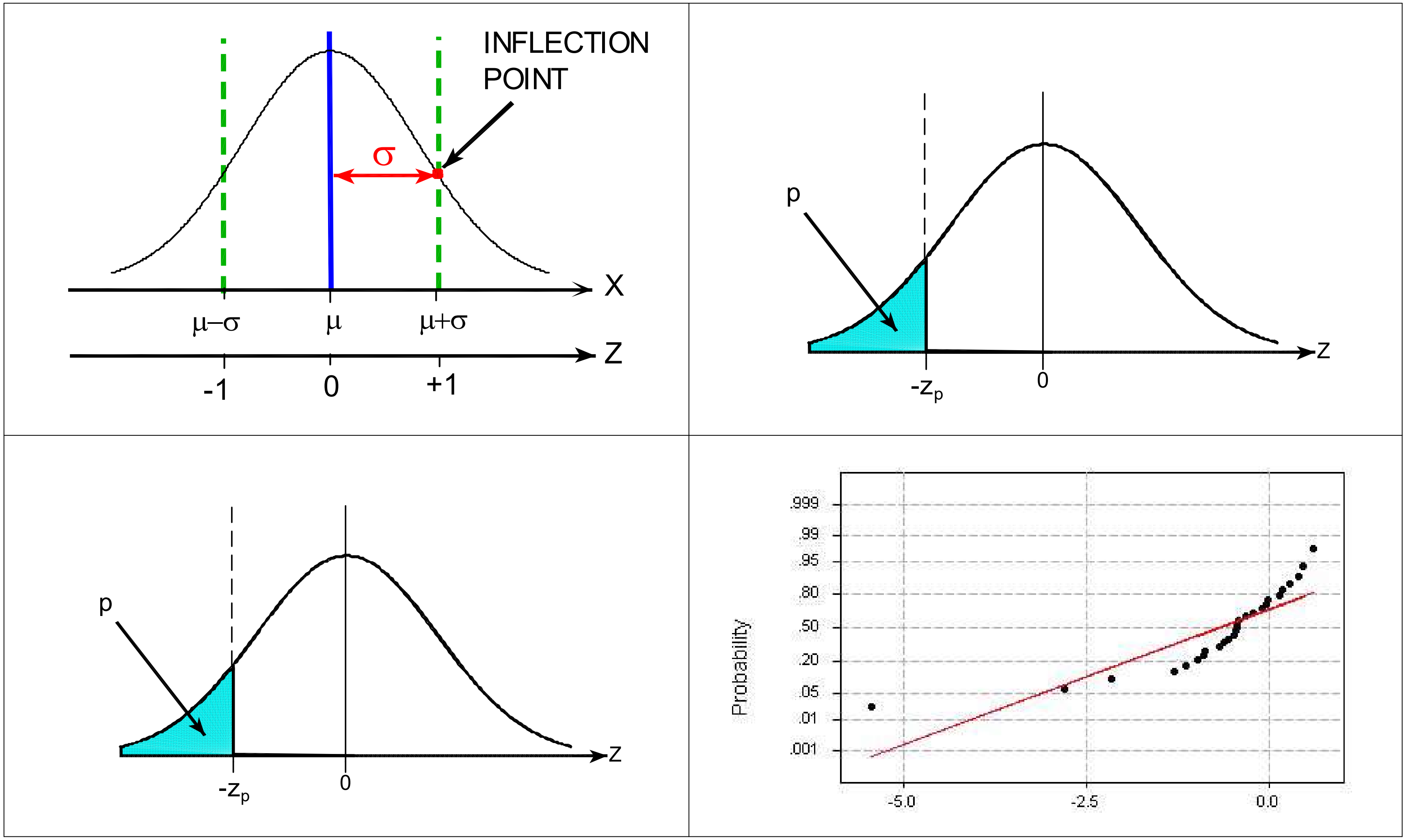
Are We Keeping the Right Charts?

- Do the operators use the charts to diagnose a problem when the process goes out of control?
- Are there problems with the process that don't show up on one or more charts?
- Are out of control events detected in a timely manner?
- Can the operator run the process without the charts?
- Are there charts that never get looked at?
- If both paper and electronic charts are being kept, which format do the operators look at when the process goes out of control?

MINITAB Commands

- Stat> Control Charts> Attributes Charts> NP
- Stat> Control Charts> Attributes Charts> P
- Stat> Control Charts> Attributes Charts> C
- Stat> Control Charts> Attributes Charts> U

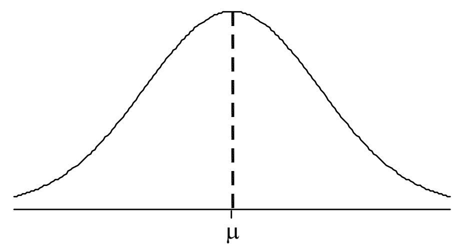
The Normal Distribution



The Normal Distribution

Aliases:

- The Bell Curve
- The Normal Distribution
- The Gaussian Distribution
- The Error Function



Also:

- There's nothing "normal" about it. The normal distribution is very special and very specific.
- "Normal" is just its name.

The Normal Distribution

- Characterizes many types of measurement or variables data, i.e. continuous variables, e.g. length, time, mass, temperature, ...
- Characterizes many statistics, e.g. the distribution of sample means ($\bar{x}$)
- Used to approximate many discrete probability distributions, e.g. the binomial and hypergeometric distributions
- Characterizes many types of measurement data after application of a transformation, e.g. square-root, square, logarithm, ...

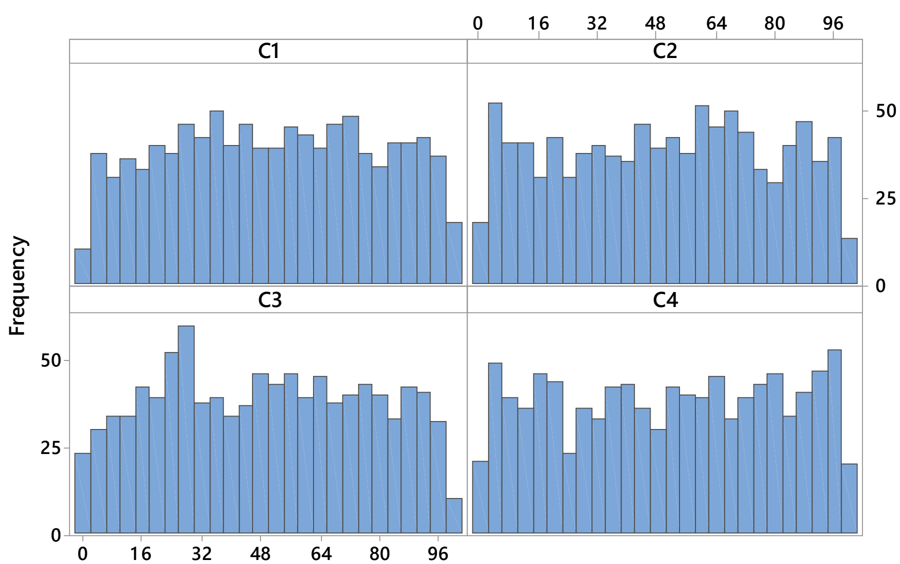
Why Does Normal Matter?

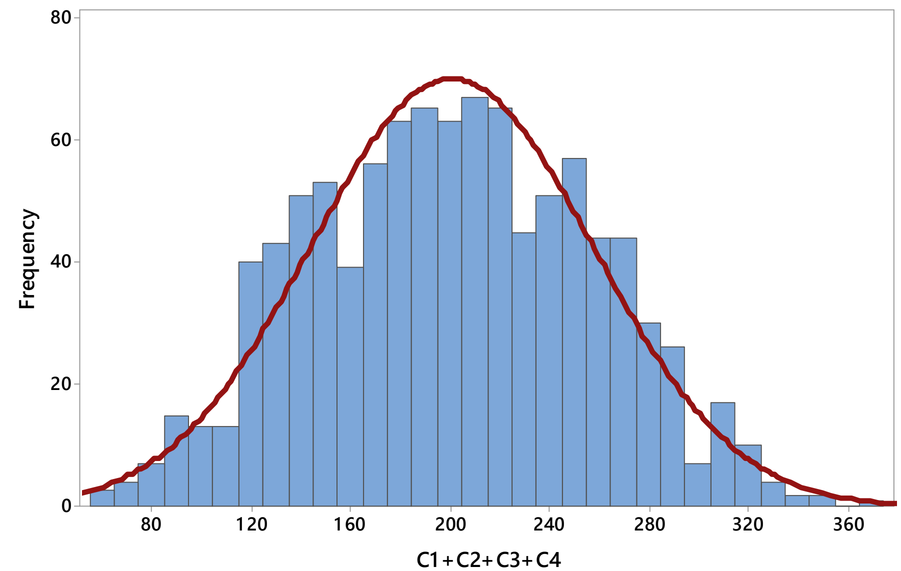
- Nature makes normal distributions
- It’s one of the most common
- There are powerful analysis methods for normal data
- If you assume normal and it’s not there could be big discrepancies

Why Does Normal Matter?

Example: Nature makes normal distributions

- The sum of two or more normals is normal
- The sum of two or more nonnormals is normal





The Normal Distribution

- The function that describes the bell curve or the normal *probability density function* (*pdf*) is:

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

- μ is the mean and is the value of x that corresponds to $\frac{\partial f}{\partial x} = 0$, i.e. the peak of $f(x;\mu,\sigma)$. The distribution is symmetric about μ.
- σ is the standard deviation and is the x distance from the mean to the inflection point corresponding to $\frac{\partial^2 f}{\partial x^2} = 0$.
- The term involving x suggests the standard normal transformation:

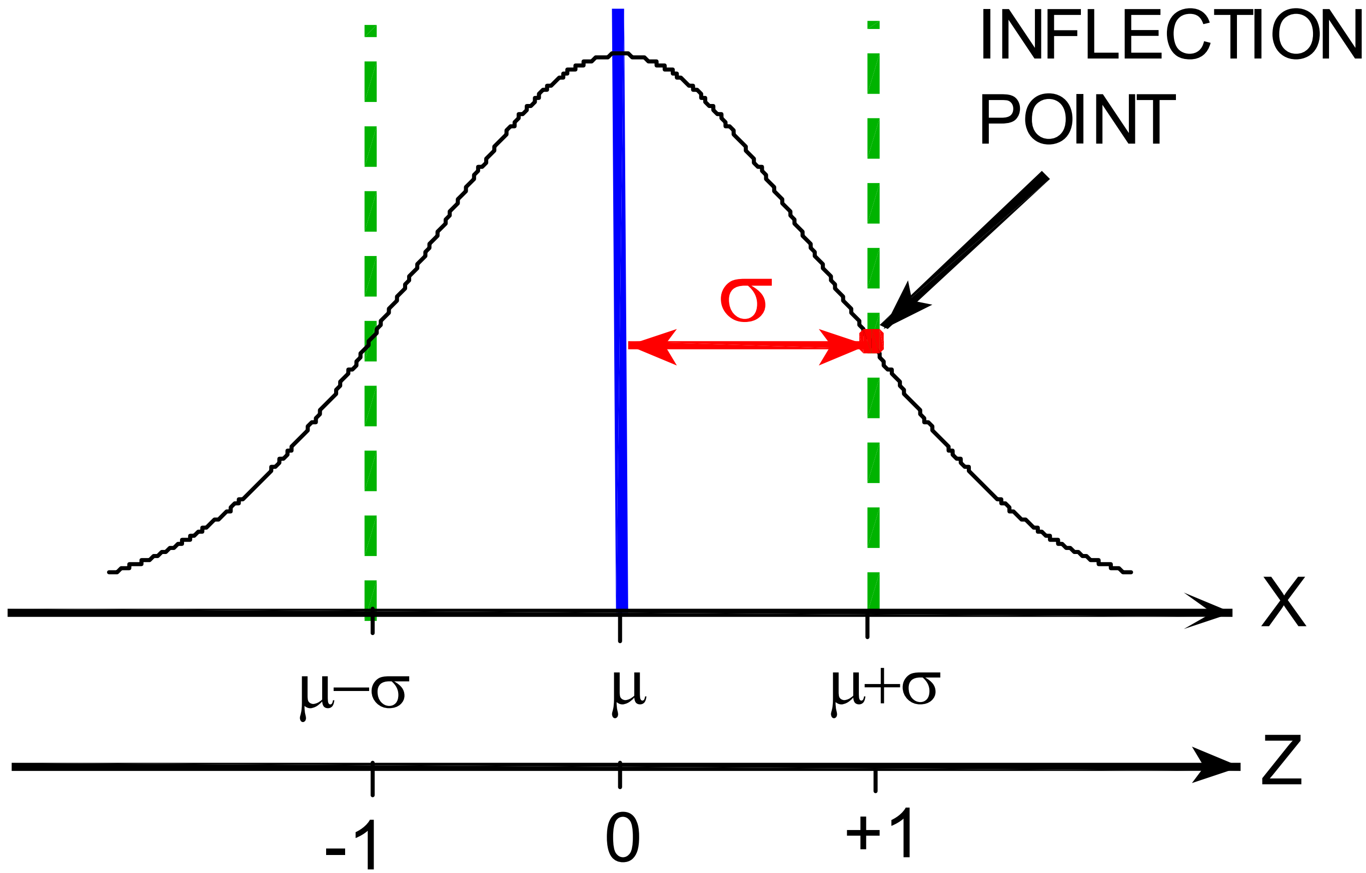
$$z = \frac{x-\mu}{\sigma}$$

  with inverse transformation:

$$x = \mu + z\sigma$$

- The standard normal z distribution's mean is $\mu_z = 0$ and its standard deviation is $\sigma_z = 1$.

The Normal Distribution: μ and σ



The Normal Distribution

We are concerned with the probability that x will fall within a specified range of values. This probability corresponds to the area under the $f(x;\mu,\sigma)$ curve:

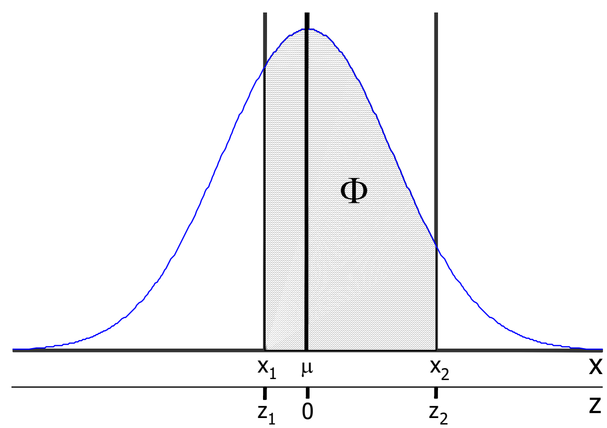
$$\Phi(x_1 < x < x_2;\mu,\sigma) = \int_{x_1}^{x_2} f(x;\mu,\sigma)dx$$

With application of the standardizing transform $z = \frac{x-\mu}{\sigma}$:

$$\Phi(x_1 < x < x_2;\mu,\sigma) = \Phi(z_1 < z < z_2)$$

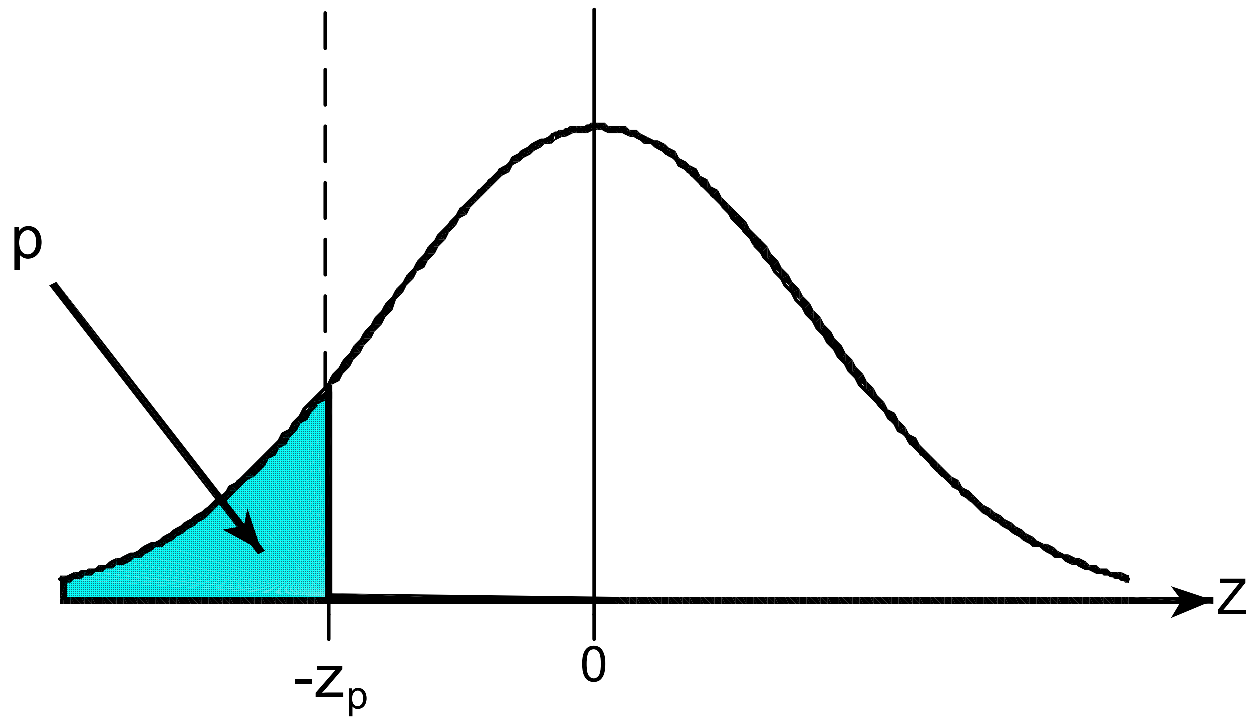
where

$$z_1 = \frac{x_1 - \mu}{\sigma} \text{ and } z_2 = \frac{x_2 - \mu}{\sigma}$$



Finding Normal Probabilities

Left tail normal probabiltiy values $p = \Phi(-\infty < z < z_p)$ are given in the table on p. 21 of Appendix B.



Finding Normal Probabilities Using Tables

Standard normal probability tables found in textbooks, such as p. 21 of Appendix B, have the following format where the z value is the sum of the row and column header and the left tail normal probability is given at their intersection in the body of the table:

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.0										
-2.9										
-2.8			0.0024							
-2.7										
⋮										
+2.9										
+3.0										

For example, the table gives $\Phi(-\infty < z < -2.82) = 0.0024$.

Finding Normal Probabilities in MINITAB

- Use **Graph**> **Probability Distribution Plot**> **View Probability** to create a normal curve with specified shaded area. The shaded area can be specified in terms of
  - x, μ, and σ (or use $\mu = 0$ and $\sigma = 1$ for the standard normal z distribution)
  - a left, right, two-tailed, or interval probability
- Use **Calc**> **Probability Distributions**> **Normal**
  - **Probability Density** to find the amplitude of the normal curve as a function of z
  - **Cumulative Probability** to find the left tail probability for a specified z value
  - **Inverse Cumulative Probability** to find the z value for a specified left tail probability
  - Input types:
    - **Input Column**
    - **Input Constant**

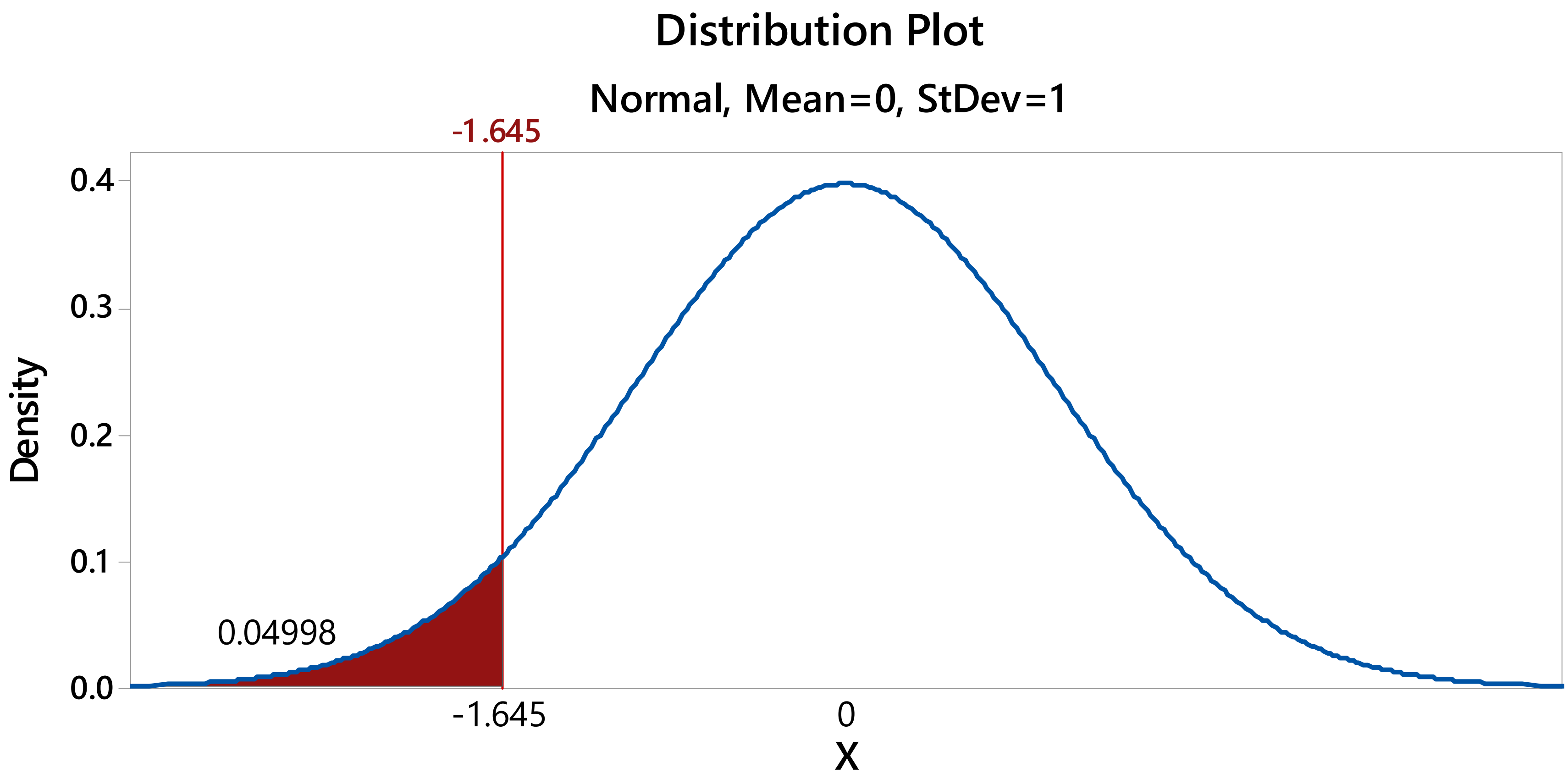
Finding Normal Probabilities in MINITAB

- At the MINITAB command prompt or command line editor (Control + L):
  - Use the *cdf* function to find the left tail probability for a given x or z value
  - Use the *invcdf* function to find the x or z value for a given left tail probability
  - Use the *pdf* function to find the amplitude of the probability distribution for a given x or z value

Finding Normal Probabilities in MINITAB

Problem: Use MINITAB to find $\Phi(-\infty < z < -1.645)$.

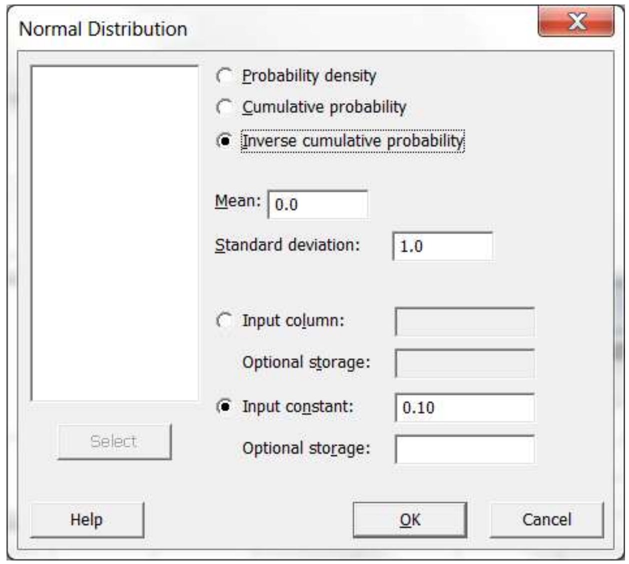
Solution: Using **Graph**> **Probability Distribution Plot**> **View Probability**> **Shaded Area**> **Left Tail**> **X**



Finding Normal Probabilities in MINITAB

Problem: Use MINITAB to find z' for $\Phi(-\infty < z < z') = 0.10$.

Solution: Using **Calc**> **Probability Distributions**> **Normal**> **Inverse Cumulative Probability**> **Input Constant**:




```
Inverse Cumulative Distribution Function
Normal with mean = 0 and standard deviation = 1
P( X \U{2264} x)   x
0.1         -1.28155
```

Finding Normal Probabilities in MINITAB

Problem: Use MINITAB to find z' for $\Phi(-\infty < z < z') = 0.025$.

Solution: Using the *invcdf* function at the command prompt:

```
MTB > invcdf 0.05

Inverse Cumulative Distribution Function
Normal with mean = 0 and standard deviation = 1
P( X \U{2264} x)    x
0.1          -1.64485
```

Normal Distribution Problems

1. Find $\Phi(-\infty < z < -2.52)$

2. Find $\Phi(-\infty < z < 2.52)$

3. Find $\Phi(-2.52 < z < 2.52)$

4. Find $\Phi(-1.18 < z < -0.06)$

5. Find $\Phi(1.73 < z < 3.08)$

Normal Distribution Problems

6. Find $\Phi(-1.00 < z < 1.00)$

7. Find $\Phi(-2.00 < z < 2.00)$

8. Find $\Phi(-3.00 < z < 3.00)$

Normal Distribution Problems

9. Find a symmetric interval on z that contains 80% of a normal population.

10. Find a symmetric interval on z that contains 99% of a normal population.

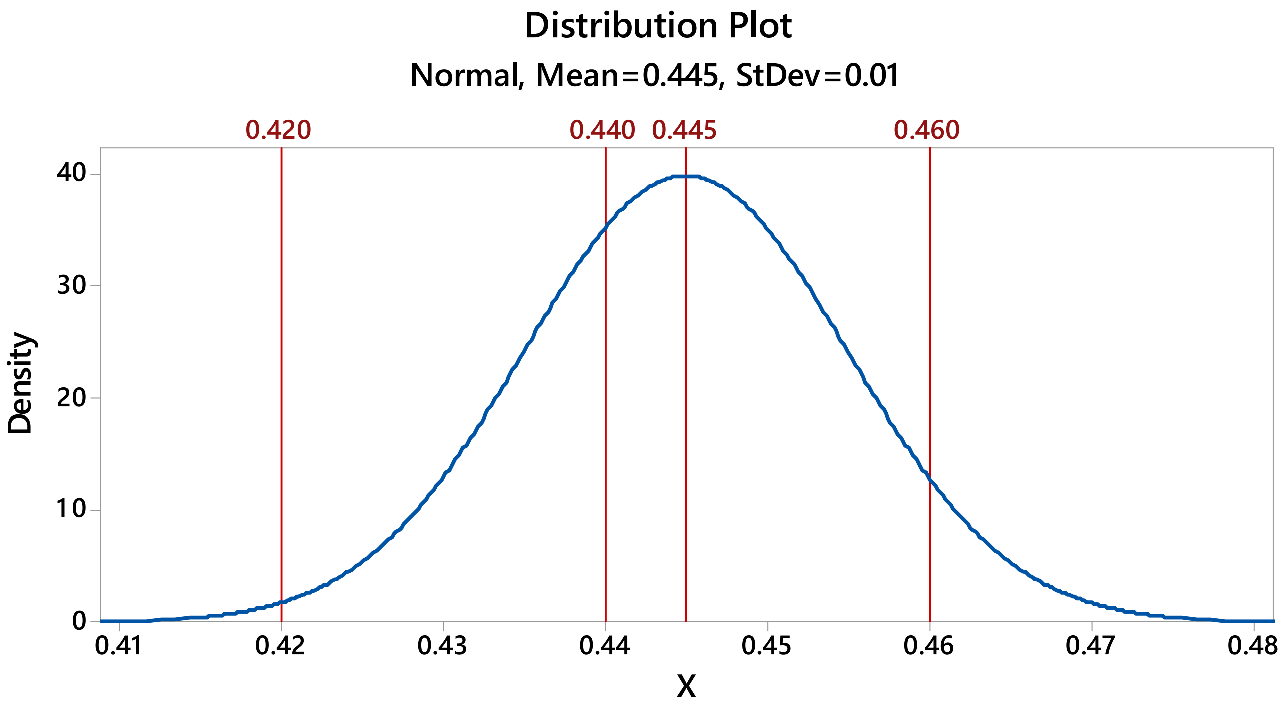
11. Find the value of z' that satisfies $\Phi(0.00 < z < z') = 0.4222$.

12. Find the value of z' that satisfies $\Phi(z' < z < 1.84) = 0.3222$.

The Standard (z) Transform

Example: Find the fraction defective produced by a process to specification $USL/LSL = 0.440 \pm 0.020$ inches if the mean of the process is $\mu = 0.445$ inches and the standard deviation is $\sigma = 0.010$ inches. Assume that the distribution is normal.

Solution: Start with DTDP by using **Graph**> **Probability Distribution Plot**> **View Single** and right clicking to add reference lines:



Solution: We need to find:

$$\Phi(0.420 < x < 0.460; \mu = 0.445, \sigma = 0.010)$$

If we apply the standardizing transformation to the LSL:

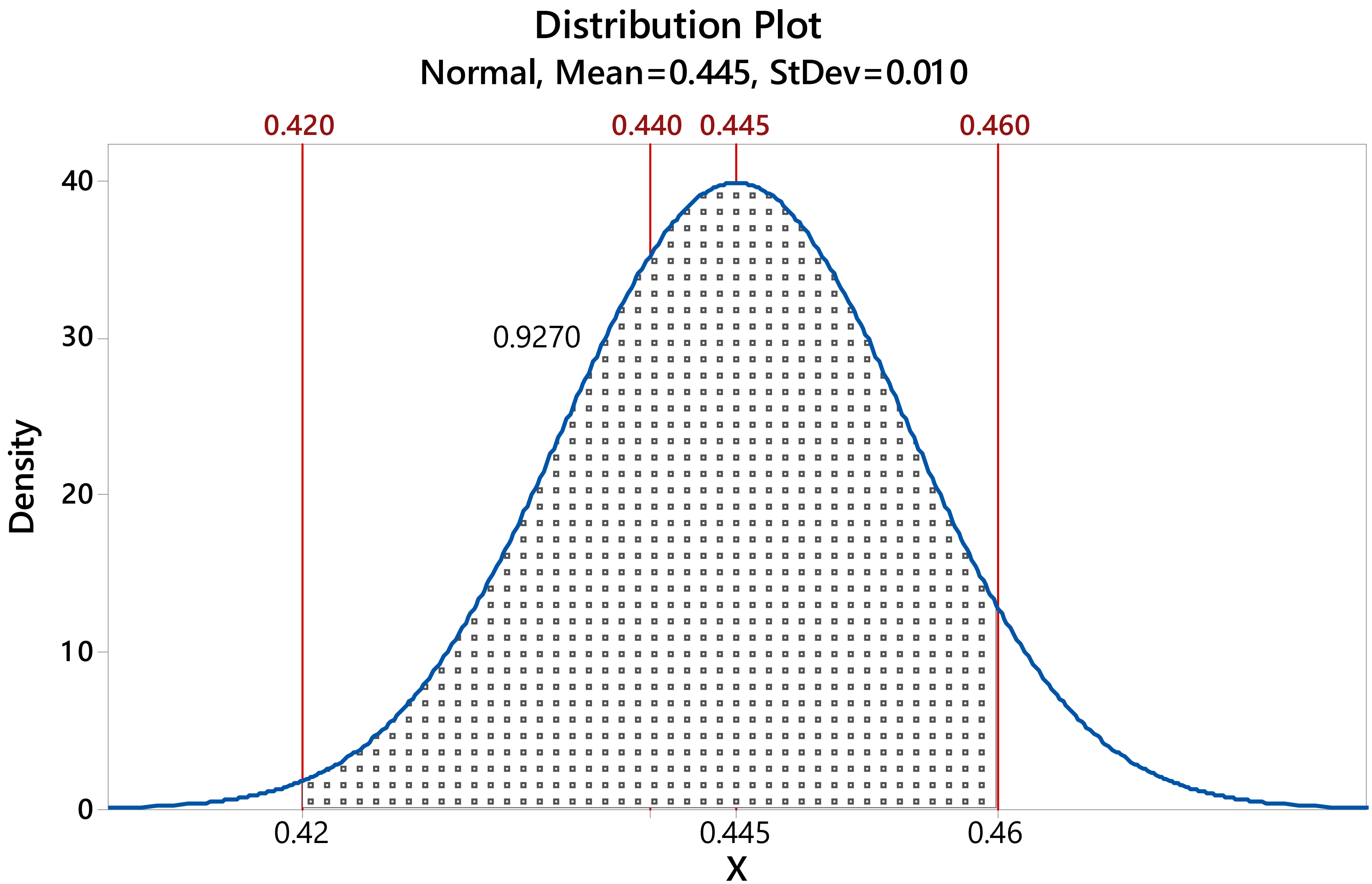
$$z_{LSL} = \frac{LSL - \mu}{\sigma}$$
$$= \frac{0.420 - 0.445}{0.010}$$
$$= -2.50$$

Similarly the z value of the USL is $z_{USL} = \frac{0.460-0.445}{0.010} = 1.50$.

So the required probability is

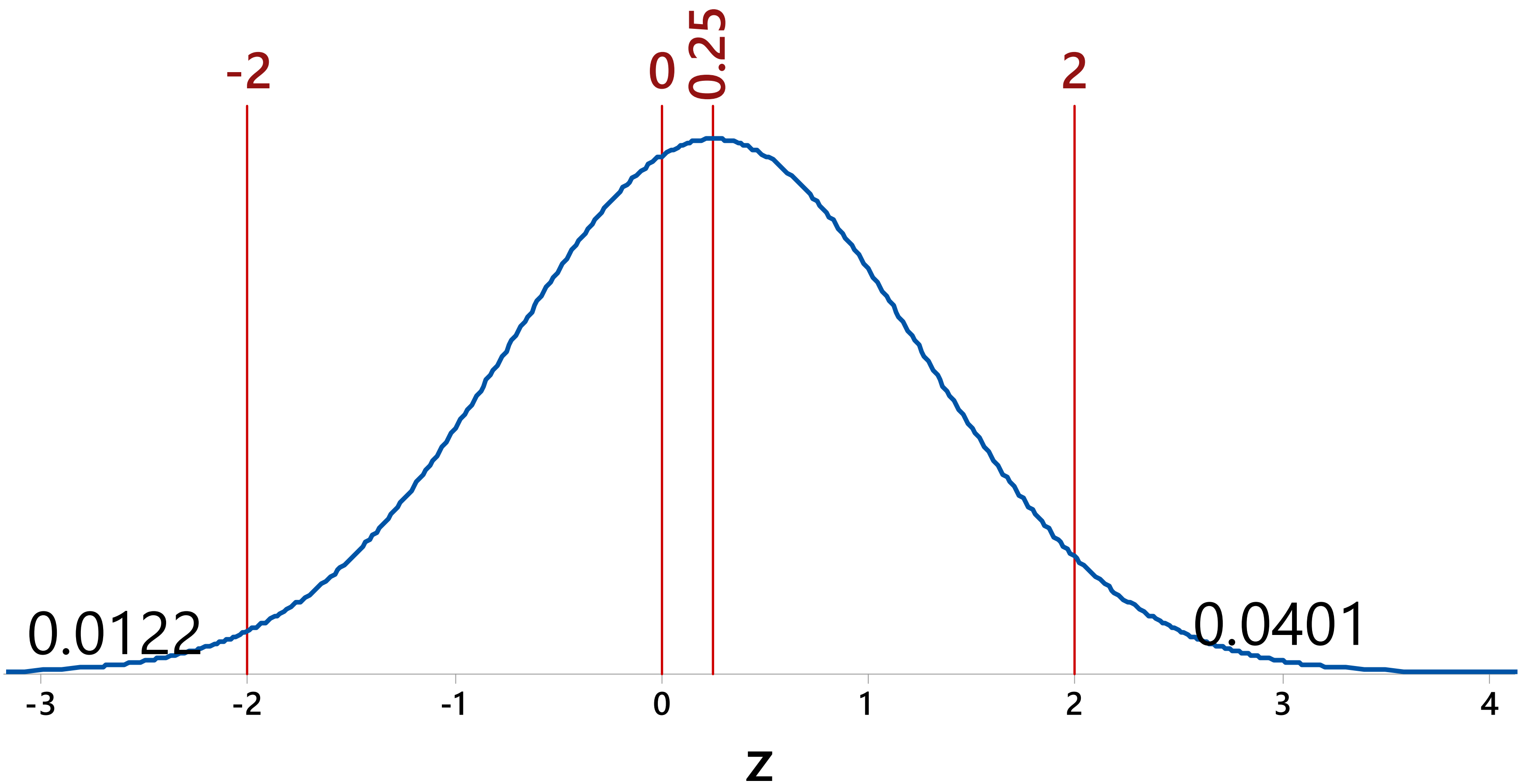
$$p = \Phi(0.420 < x < 0.460; \mu = 0.445, \sigma = 0.010)$$
$$= \Phi(-2.50 < z < 1.50)$$
$$= 0.9332 - 0.0062$$
$$= 0.9270 = 1 - 0.0730$$

This means that 92.7% of the product is in spec and 7.3% of the product is out of spec. Most off the defectives are coming from the right tail.



Tip

The normal distribution's tails fall off very quickly so if a distribution is shifted in its two-sided spec then most of the defectives come from one tail and it's usually safe to ignore the contribution from the other tail.



Special Areas Under the Normal Curve

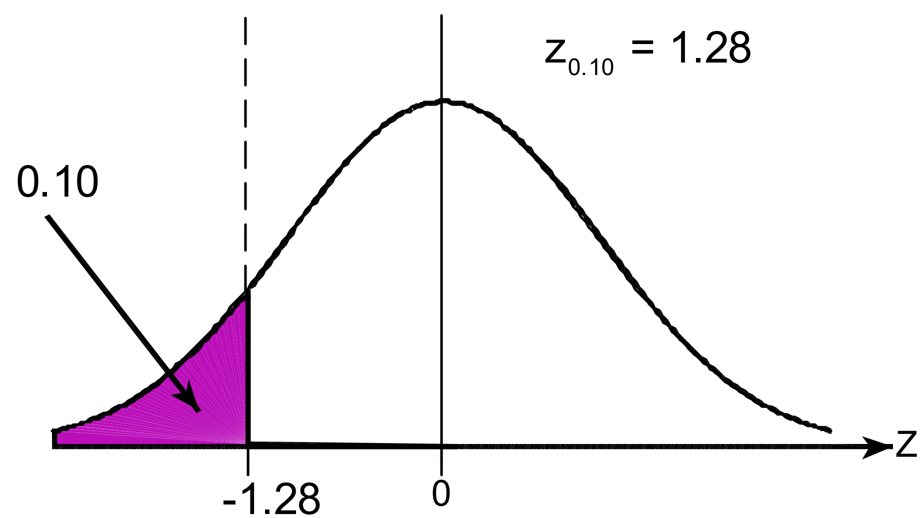
A subscript on the symbol z indicates a tail area under the normal curve, i.e. z_p is the z value that has a tail area of p:

$$p = \Phi(-\infty < z < -z_p)$$

The $+$ or $-$ sign on z_p corresponding to the left or right tail, respectively, is added by context.

Example: Find $z_{0.10}$.

Solution:



Special Areas Under the Normal Curve

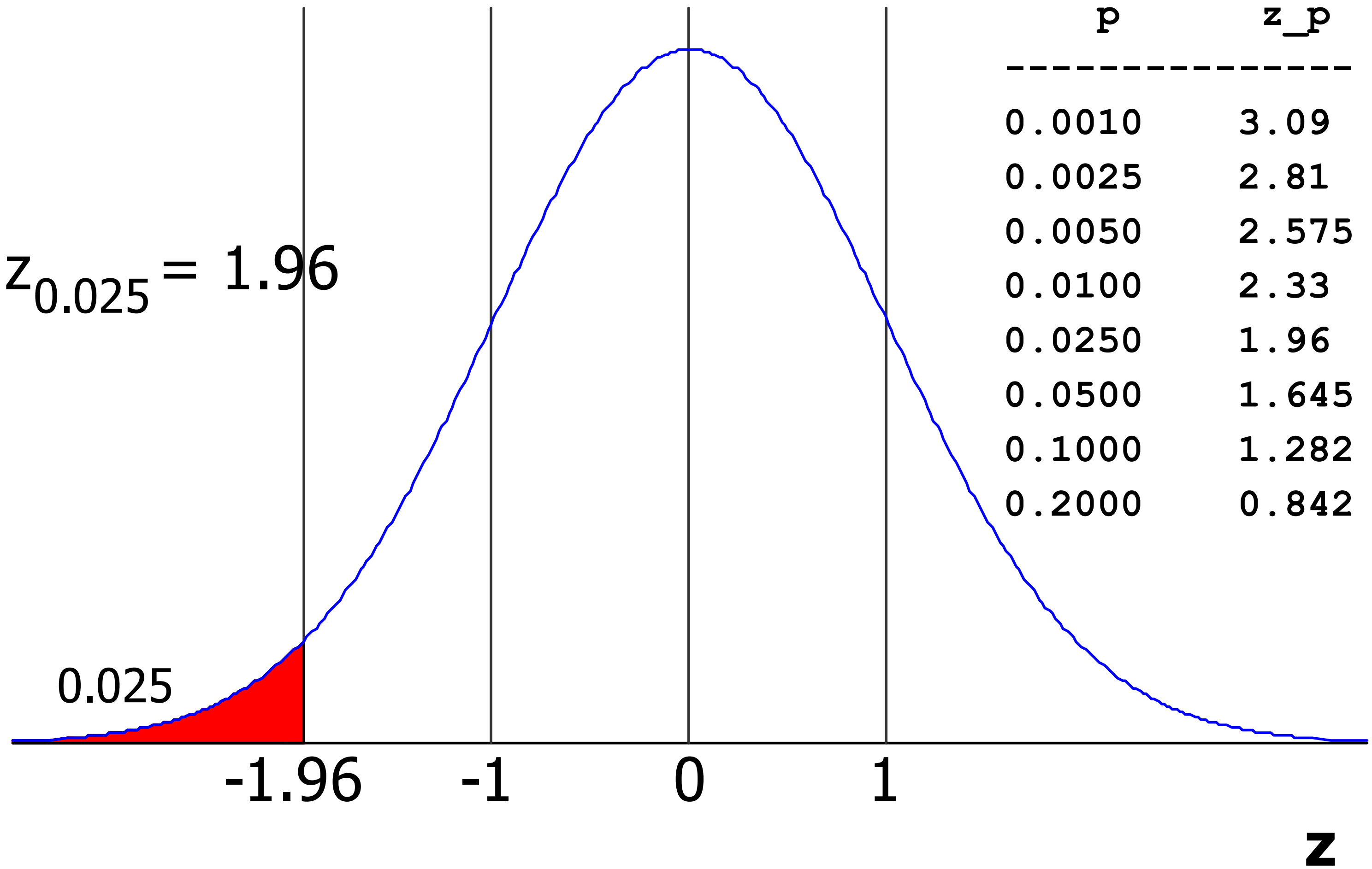
1. Find $z_{0.05}$.

2. Find $z_{0.025}$.

3. Find $z_{0.01}$.

4. Find $z_{0.005}$.

Special Areas Under the Normal Curve

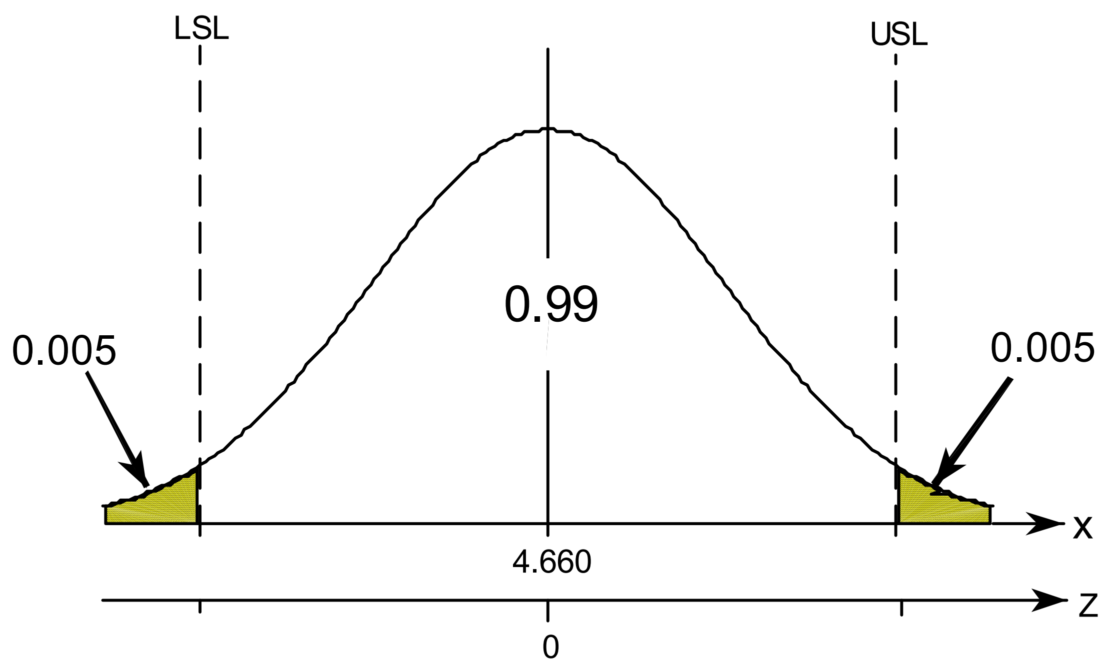


p	z_p
0.0010	3.09
0.0025	2.81
0.0050	2.575
0.0100	2.33
0.0250	1.96
0.0500	1.645
0.1000	1.282
0.2000	0.842

The Standardizing Transform

Example: Determine a two-sided specification for a process that has $\mu = 4.660$ and $\sigma = 0.008$ if the specification must contain 99% of the population. Assume that the distribution is normal.

Solution:



If 99% of the product must be in the symmetric two-sided specification then there will be 0.5% of the product out of spec on the high and low ends of the distribution. Since $z_{0.005} = 2.575$ the required specification is:

$$\Phi(LSL < x < USL; 4.660, 0.008) = 0.99$$

where

$$\begin{aligned} LSL &= \mu - z_{0.005}\sigma \\ &= 4.660 - 2.575 \times 0.008 \\ &= 4.639 \end{aligned}$$

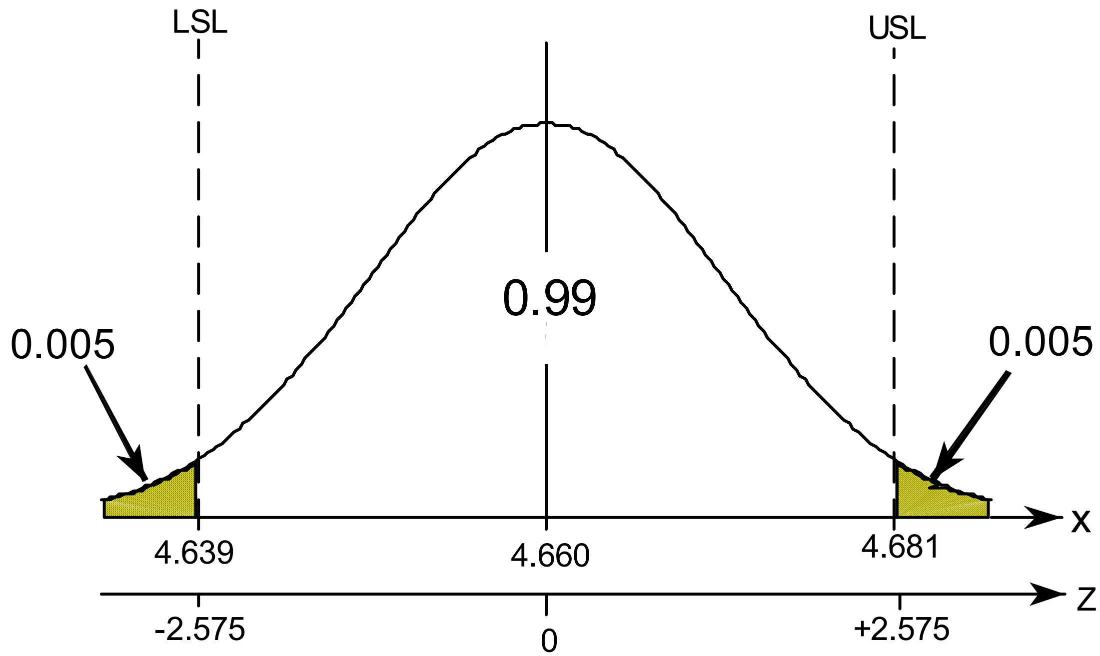
and

$$\begin{aligned} USL &= \mu + z_{0.005}\sigma \\ &= 4.660 + 2.575 \times 0.008 \\ &= 4.681 \end{aligned}$$

Finally we have:

$$\Phi(4.639 < x < 4.681; 4.660, 0.008) = 0.99$$

so our spec of $USL/LSL = 4.681/4.639$ will contain 99% of the population.



Stack Ups

When two or more components are stacked up in one dimension the stack up's mean μ_{total} is

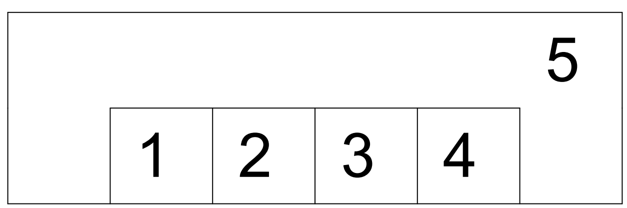
$$\mu_{total} = \sum_{i=1}^{n} \mu_i$$

and its standard deviation σ_{total} is

$$\sigma_{total} = \sqrt{\sum_{i=1}^{n} \sigma_i^2} \ .$$

The μ_i may be signed. If the individual components are normally distributed then the stack up will also be normally distributed.

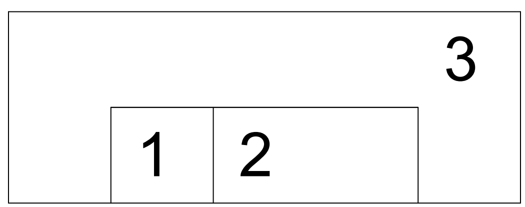
Here's an example of a five component stack up:



Stack Ups

Example: An assembly has two components $(1,2)$ that are stacked and then inserted inside of the gap in a third component (3). The component means and standard deviations are shown in the table below and all three component distributions are normal. Determine the distribution of the gap in the assembly and the fraction of the assemblies that will suffer an interference failure.

i	μ_i	σ_i
1	0.250	0.003
2	0.500	0.004
3	0.762	0.005



Stack Ups

Solution: The mean gap will be

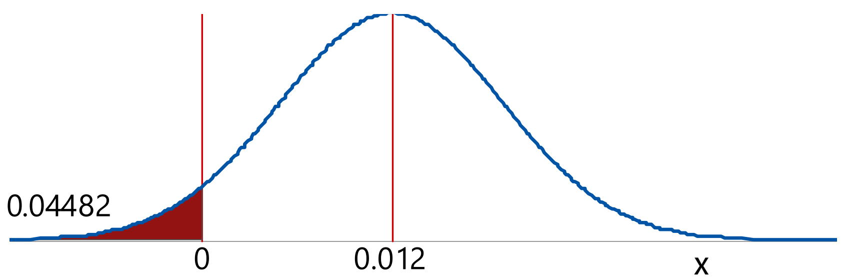
$$\mu_{Gap} = 0.762 - (0.250 + 0.500) = 0.012$$

and the standard deviation of the gap will be

$$\sigma_{Gap} = \sqrt{0.003^2 + 0.004^2 + 0.005^2} = 0.00707$$

An interference fit will occur when $Gap < 0$ so

$$\begin{aligned} p &= \Phi(-\infty < Gap < 0; \mu = 0.012, \sigma = 0.00707) \\ &= \Phi\left(-\infty < z_{Gap} < \frac{0 - 0.012}{0.00707}\right) \\ &= \Phi(-\infty < z_{Gap} < -1.70) \\ &= 0.0446 \end{aligned}$$



Normal Approximation to Discrete Probability Distributions

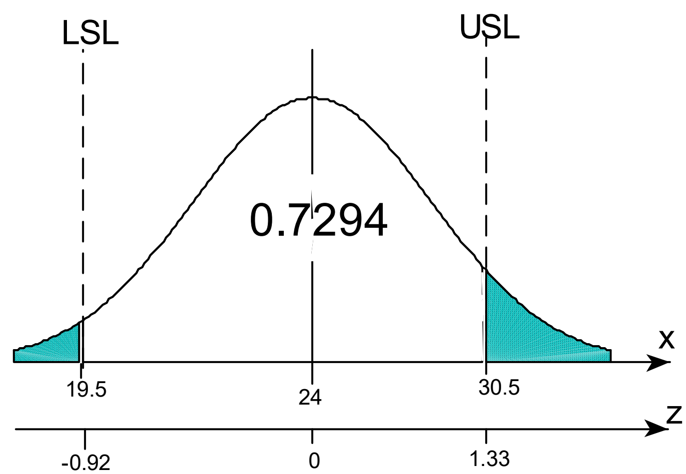
- Under certain conditions the normal distribution can be used to approximate the binomial and Poisson distributions:
  - $b(x';n,p) \simeq \Phi\left(x' - \frac{1}{2} < x < x' + \frac{1}{2}; \mu = np, \sigma = \sqrt{np(1-p)}\right)$
  - $Poisson(x';\lambda) \simeq \Phi\left(x' - \frac{1}{2} < x < x' + \frac{1}{2}; \mu = \lambda, \sigma = \sqrt{\lambda}\right)$
- This requires the use of the *continuity correction* which meshes the discrete random variable of the binomial and Poisson distribution into the continuous random variable of the normal distribution.
- Instructions for these approximations are given in the table of approximations in Appendix B.

Normal Approximation

Example: Approximate $Poisson(20 \leq x \leq 30; \lambda = 24)$ and determine the error of the estimate.

Solution: Since λ is greater than 20 we can use the normal approximation to the Poisson distribution:

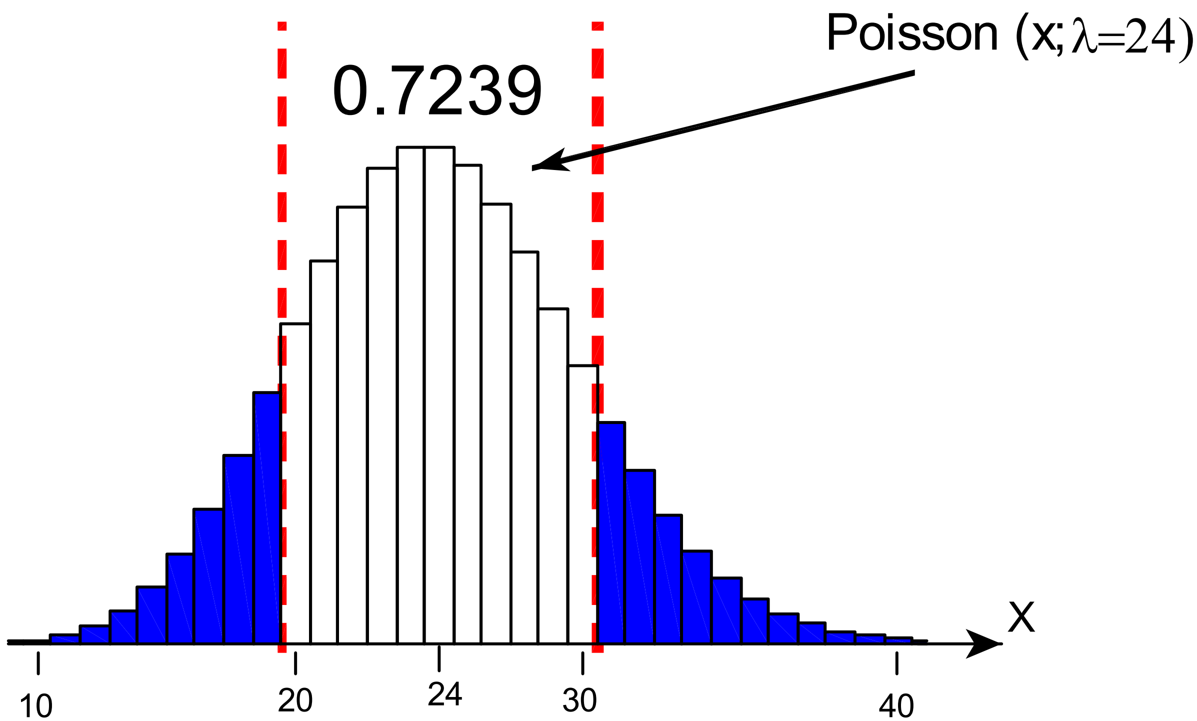
$$\begin{aligned} Poisson(20 \leq x \leq 30; \lambda = 24) &\simeq \Phi(19.5 < x < 30.5; 24, 4.90) \\ &\simeq \Phi(-0.92 < z < 1.33) \\ &\simeq 0.9082 - 0.1788 \\ &\simeq 0.7294 \end{aligned}$$



The exact answer (by HP32C calculator) is:

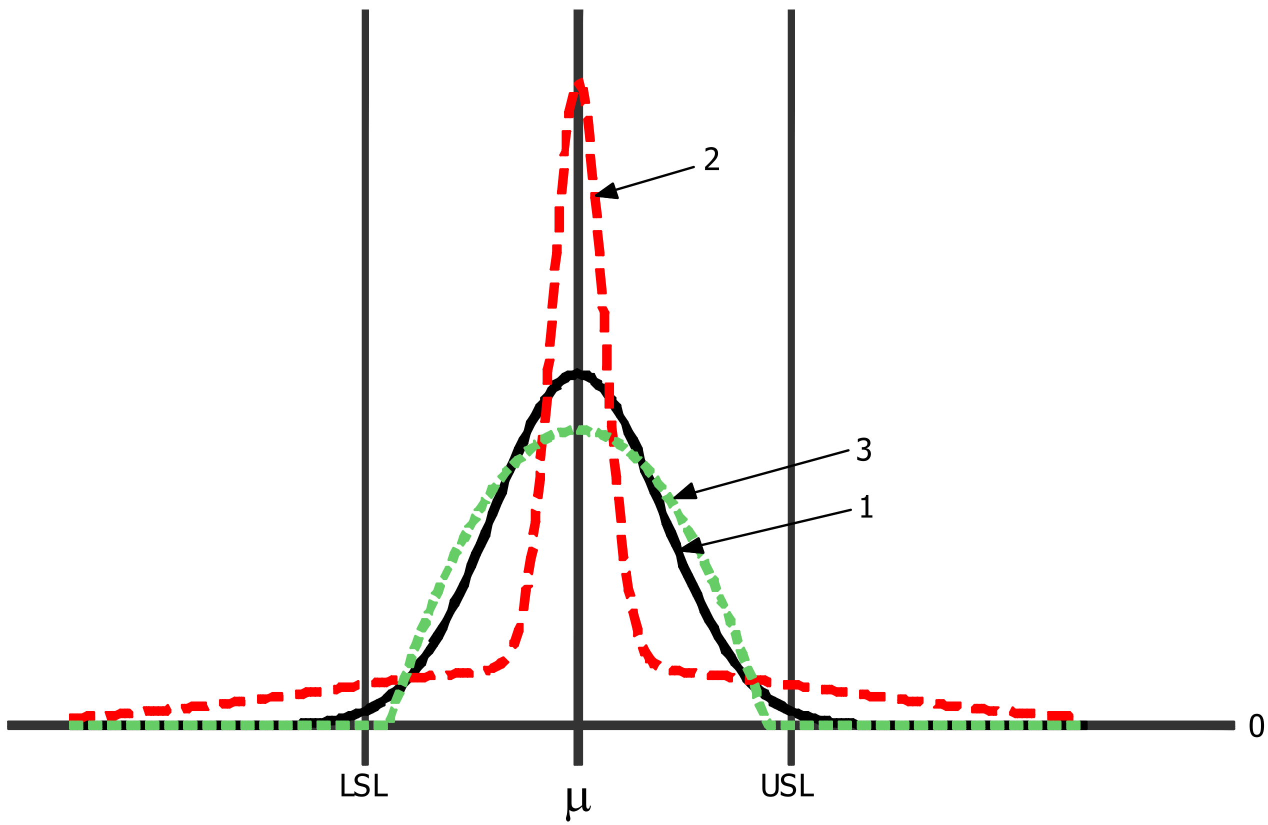
$$
\begin{aligned}
Poisson(20 \leq x \leq 30; \lambda = 24) &= 0.9042 - 0.1803 \\
&= 0.7239
\end{aligned}
$$

so the error of the estimate is $\frac{0.7294 - 0.7239}{0.7239} = 0.0076.$



Testing Data for Normality

These three distributions have the same mean and standard deviation but their defective rates relative to the specification limits are very different. We need to test for normality.



Testing Data for Normality

- The normal distribution applies in many situations.
- It is necessary to check that a data set follows the normal distribution before using it.
- Testing for normality:
  - Histogram with superimposed normal curve - can identify significant deviations from normality but incapable of identifying smaller ones
  - Normal probability plots are much preferred
  - Quantitative tests for normality: e.g. Anderson-Darling and many others

Normal Probability Plot Procedure

- See Appendix A. *Normal Probability Plots*
- Procedure
  1. Collect the data
  2. Sort from smallest to largest
  3. Calculate the relative position: $(i - 0.5)/n$
  4. Calculate the z scores: $z_{(i-0.5)/n}$
  5. Plot the z scores versus the sorted data values
  6. Draw a straight line through the data
  7. If the plotted points fall along the straight line without any curves, hooks, jumps, or other deviations then conclude that the distribution being sampled is normal
  8. If the plotted points follow a smooth curve try
    a. Using a variable transform like log, sqrt, square, reciprocal, or a power function
    b. Using a different probability distribution. e.g. Weibull

Normal Probability Plots

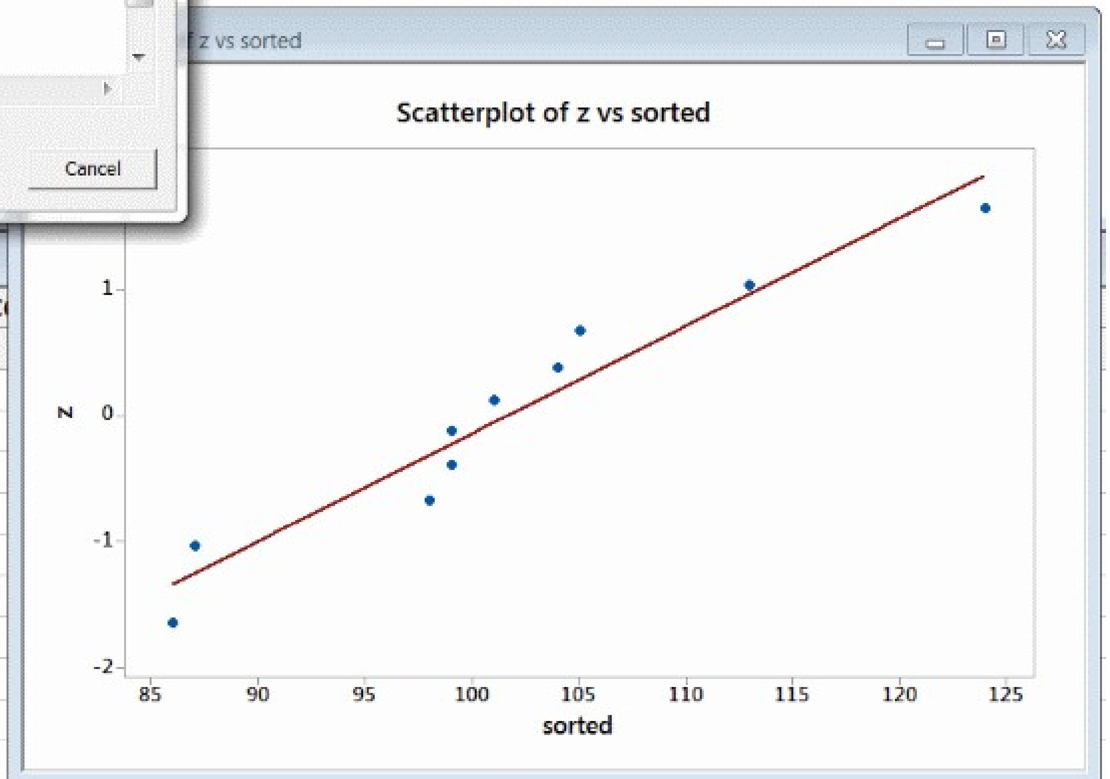
Command Line Editor

```
rand 10 c100;
  normal 100 10.
let c1 = round(c100)
sort c1 c2
set c3
  1:10
end
let c4 = (i - 0.5)/10
invcdf c4 c5
Plot 'z'*'sorted';
 Symbol;
```

Submit Commands | Cancel

Worksheet 1 ***

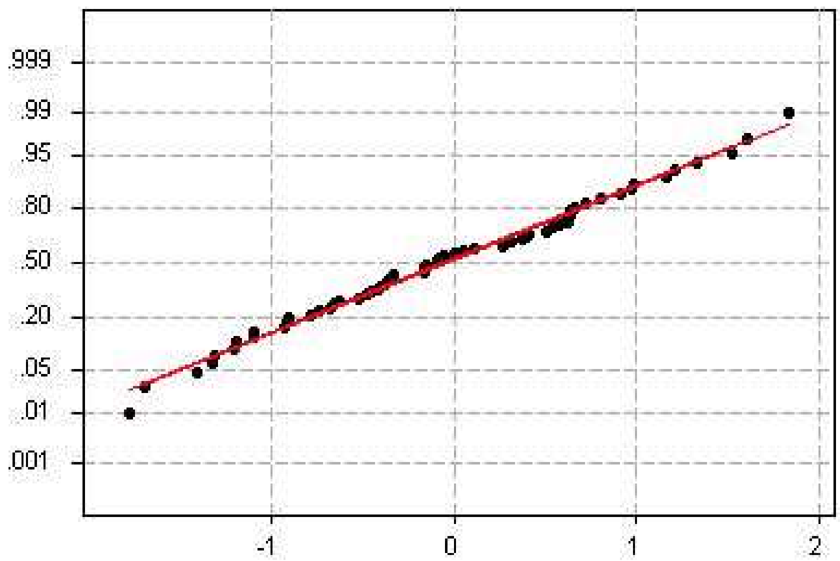
↓	C1	C2	C3	C4	C5
	x	sorted	i	(i - 0.5)/n	z
1	113	86	1	0.05	-1.64485
2	87	87	2	0.15	-1.03643
3	104	98	3	0.25	-0.67449
4	98	99	4	0.35	-0.38532
5	105	99	5	0.45	-0.12566
6	99	101	6	0.55	0.12566
7	99	104	7	0.65	0.38532
8	86	105	8	0.75	0.67449
9	101	113	9	0.85	1.03643
10	124	124	10	0.95	1.64485

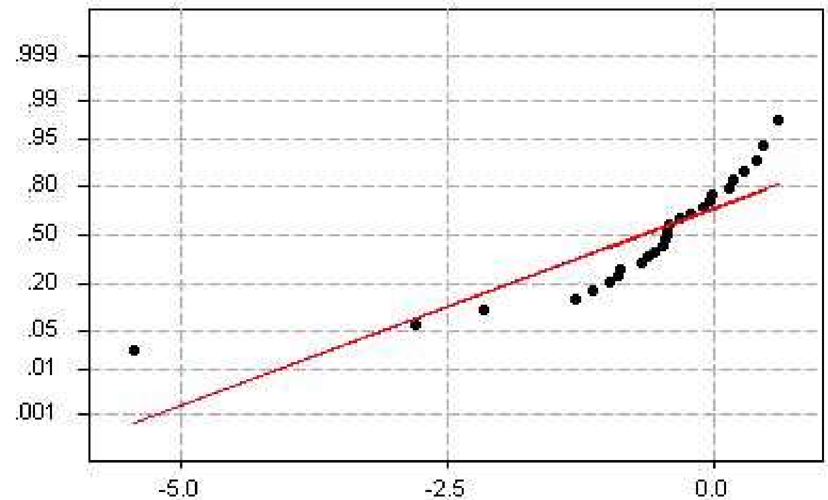


Normal Probability Plots

Normal plots in MINITAB:

- **Stat> Basic Statistics> Normality Test**
- **Graph> Probability Plot> Normal**





Anderson-Darling Test

- There are many quantitative tests for normality
  - Chi-square goodness of fit test
  - Kolmogorov
  - Lillifors
  - Shapiro-Wilk
- The Anderson-Darling (AD) is the current favorite
  - The AD test statistic is a measure of the discrepancy between the plotted points on the normal plot and the straight line fit to the data
  - The AD test statistic's value is sample size dependent which makes it hard to interpret so we use its *p* value instead

$0 \leq p \leq 0.01$	strong evidence against normality
$0.01 < p \leq 0.05$	some evidence against normality
$0.05 < p \leq 0.10$	some evidence for normality
$0.10 < p \leq 1$	strong evidence for normality

Normality Requirements

Statistical methods have different sensitivities to nonnormal data.

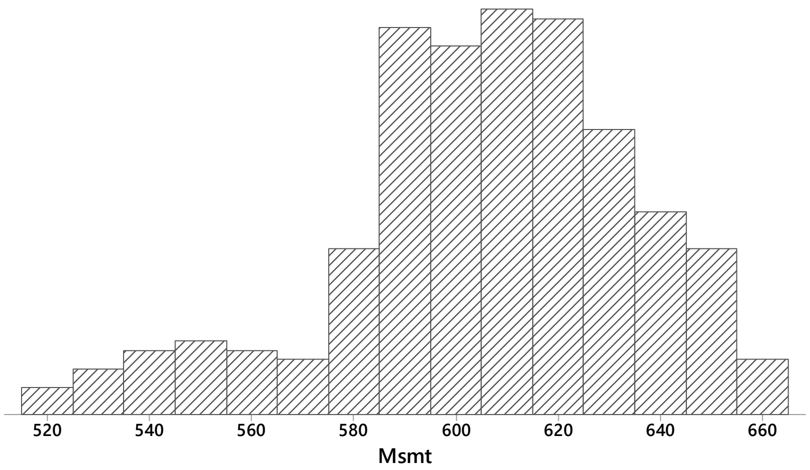
- Some require strong normality:
  - Methods that estimate standard deviations, distribution tails
  - Process capability statistics
  - Statistical tolerance intervals
  - Reliability
- Some tolerate weak normality:
  - Methods that estimate means
  - $\bar{x}$ charts
  - ANOVA
  - Linear Regression

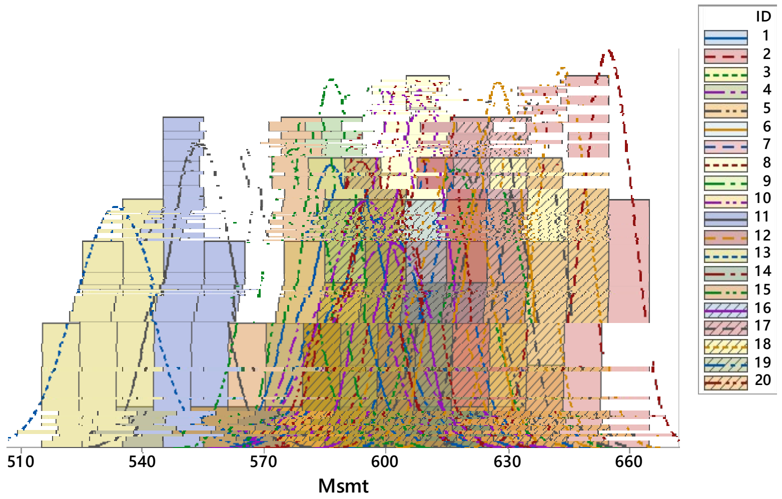
Where to Expect Non-normal Data

- Life test failure times - exponential, lognormal, Weibull, or other
- Response covers wide range of values, e.g. factor of 10 or more
- Run-out from a target position in one dimension - half- or folded-normal
- Run-out from a target position in two dimensions - circular normal
- Values near a boundary, e.g. defective rate and yield, spec limit
- Roughness average *Ra*
- Particle size
- Most economic responses, e.g. income, net worth, etc
- Extreme values, e.g. manifold burst testing
- Running within spec limits with no attempt to control to the nominal
- Waiting times - exponential
- Truncated distributions, e.g. with the out-of-spec units removed
- Coarse data, poor measurement resolution
- Count data, e.g. defectives and defects
- **Mixture of populations**
- **Time-varying parameters, lumpy**

Mixture of Populations

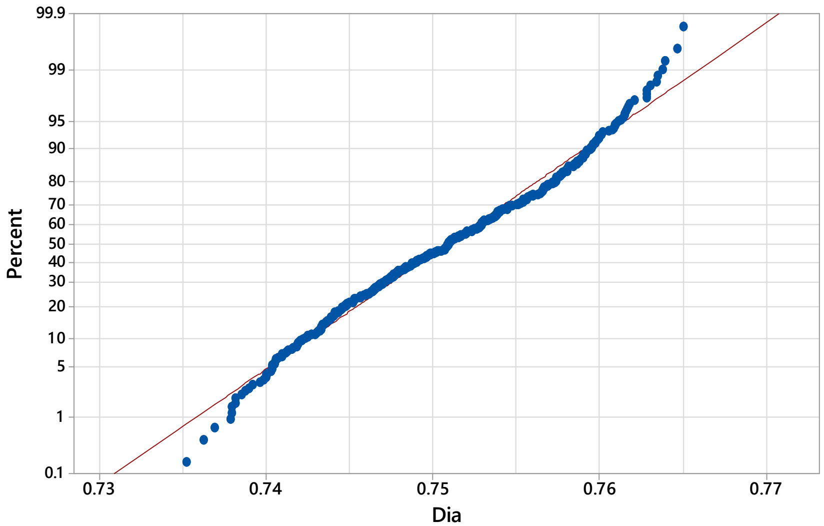
Identify the subgroups and separate the within- and between-subgroup variabilities

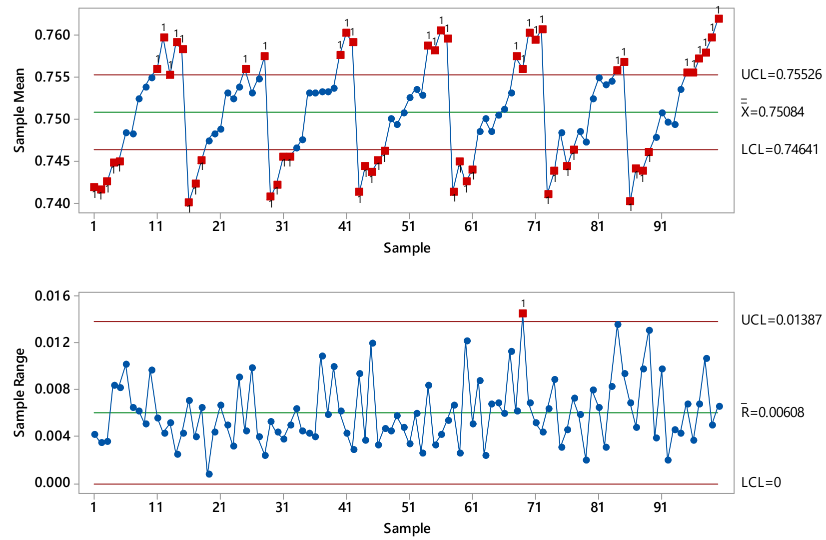




Time-varying Parameters

Identify and account for the time dependency





Time-varying Parameters

- Autocorrelated data:
  - Slowly time varying data so that each observation is well predicted by the previous one
  - Examples
    - Trend, e.g. tool wear
    - Random walk
    - Oscillation - a repeating cyclic pattern of shifts in the mean
  - Detection methods
    - Time series plot
    - Lag-1 plot: Plot x_i versus x_{i-1} using the same scale on both axes
    - Calculate the lag-1 correlation
- Analysis Tools
  - **Stat**> **Quality Tools**> **Run Chart** does tests for clustering, mixtures, trends, and oscillations
  - **Stat**> **Time Series**> **Time Series Plot**

MINITAB Methods for Fitting a Distribution

- **Stat> Basic Statistics> Normality Test**
- **Graph> Probability Plot**
- **Stat> Quality Tools> Individual Distribution Identification**
- **Stat> Reliability/Survival> Distribution ID Plot**
- **Stat> Quality Tools> Johnson Transformation**
- **Stat> Quality Tools> Capability Analysis (Normal)> Transform**
- Some methods also offer Box-Cox and/or Johnson transforms

Guidance for Fitting a Distribution

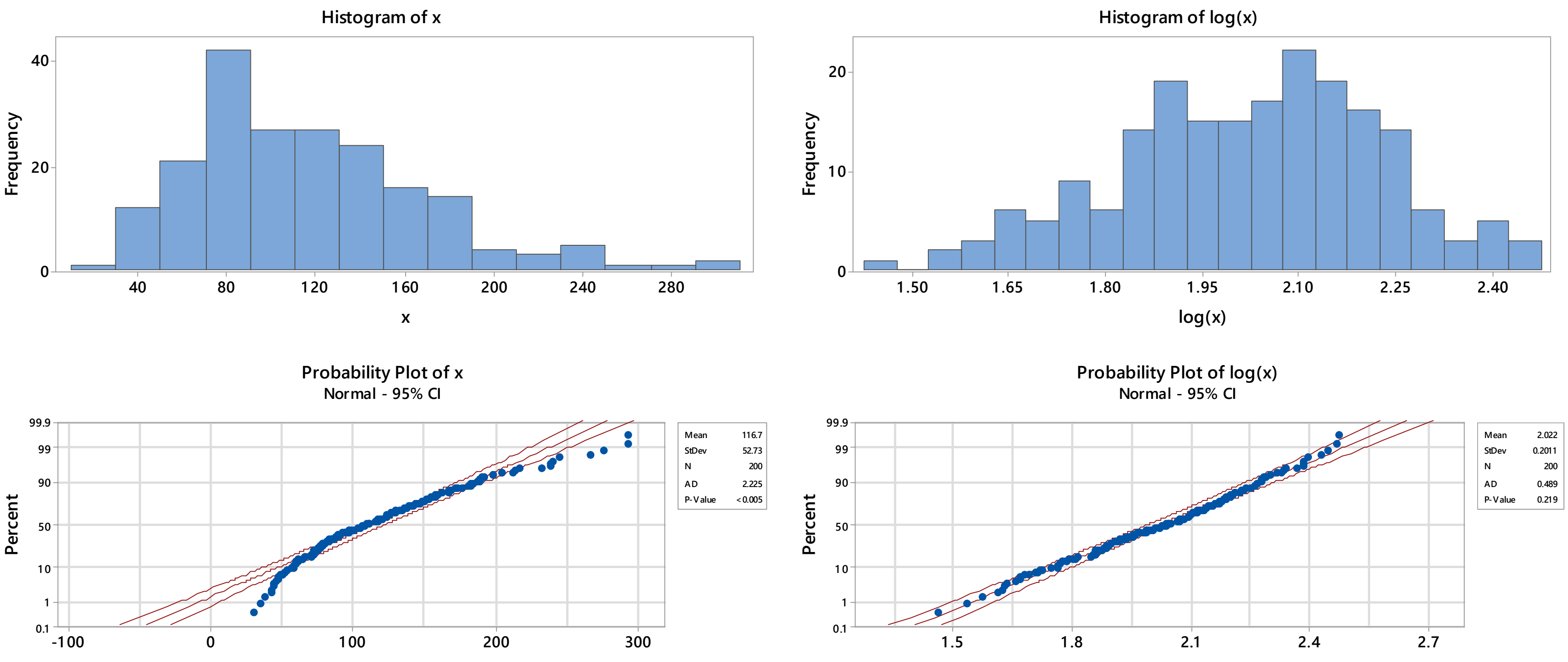
"All models are wrong. Some are useful." - George Box

Use the following order when attempting to fit a distribution model:

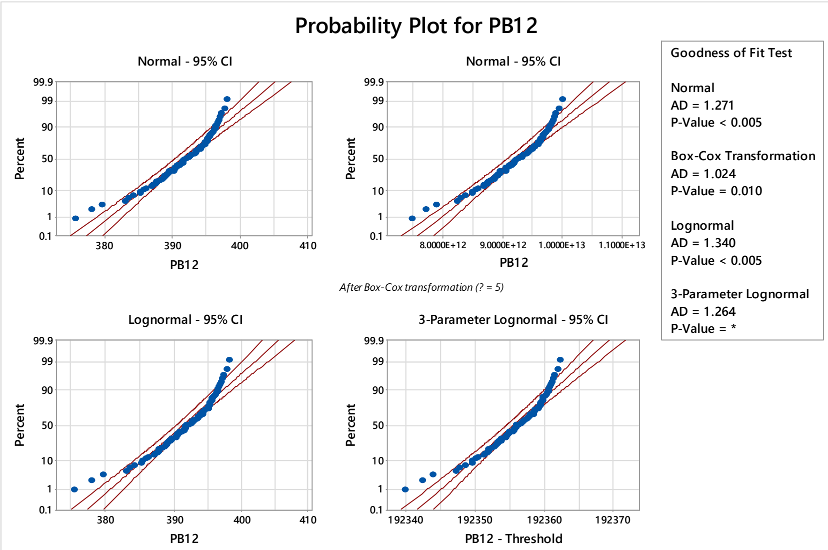
1. Make sure that you're dealing with a single population
2. Use the appropriate first principles distribution
3. Try the most common distributions, e.g. normal, lognormal, exponential, Weibull, uniform
4. Try simple variable transforms: $\sqrt{y}$, $\ln(y)$, y^{-1}, y^{λ}. Try to use first principles to inform the choice.
5. Try Box-Cox transform
6. Consider three-parameter distributions, e.g. Weibull with a threshold parameter
7. Try Johnson transforms (three families with four parameters)

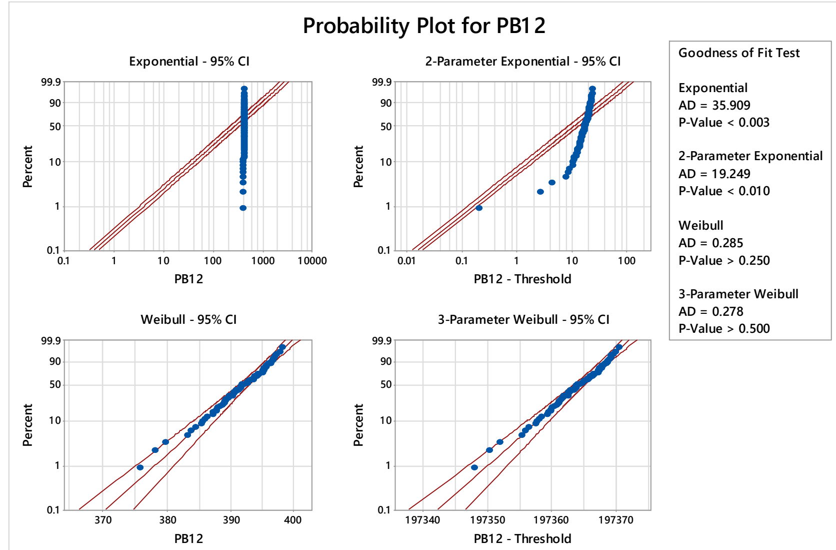
Variable Transform

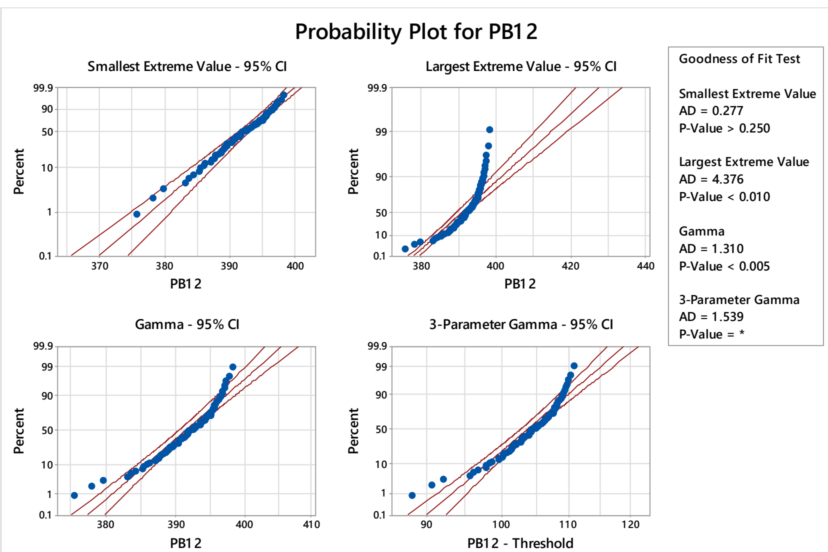
The histogram and normal plot show that the data are not normal. A log transform fixed the problem.

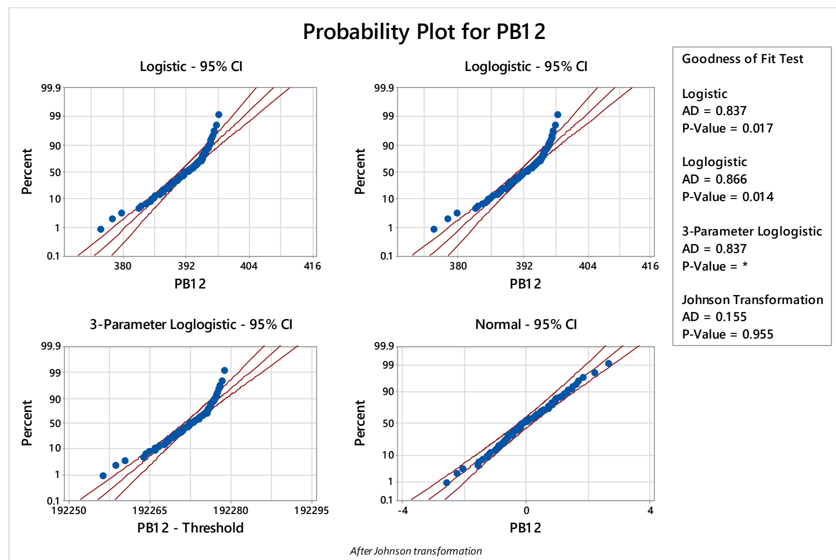


Stat> Quality Tools> Individual Distribution Identification









Box-Cox and Johnson Transforms

- Empirical transforms not based on first principles
- The Box-Cox transform is a power transform where the exponent is a parameter to be estimated:

$$y' = \begin{cases} y^{\lambda} \text{ for } \lambda \neq 0 \\ \ln(y) \text{ for } \lambda = 0 \end{cases}$$

- Johnson transforms, where γ, η, λ, and ϵ are parameters to be estimated:

Bounded	S_B	$y' = \gamma + \eta \ln\left(\frac{y-\epsilon}{\lambda+\epsilon-y}\right)$
Lognormal	S_L	$y' = \gamma + \eta \ln(y - \epsilon)$
Unbounded	S_U	$y' = \gamma + \eta \sinh^{-1}\left(\frac{y-\epsilon}{\lambda}\right)$

  - *"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk,"* John von Neumann
  - Johnson transforms reek of desperation.

Count Data

- Defectives
  - Transform binomially-distributed fraction defective data using

$$p_i' = \sin^{-1}\left(\sqrt{p_i}\right)$$

  - This transform is implemented in MINITAB as the Freeman-Tukey Proportion *ftp* function available in the **Calc**> **Calculator** menu.
- Defects
  - Transform Poisson-distributed count data using

$$x' = \sqrt{x}$$

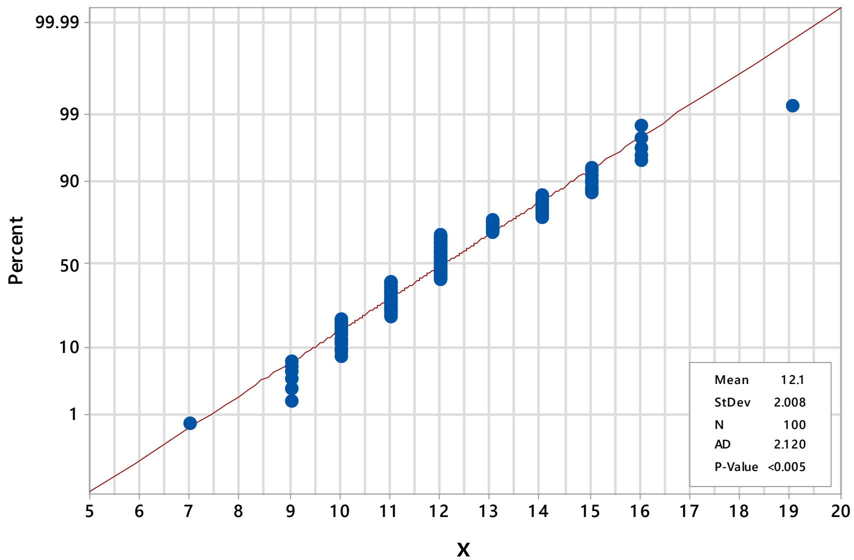
  - A variant of this transform is implemented in MINITAB as the Freeman-Tukey Count *ftc* function.

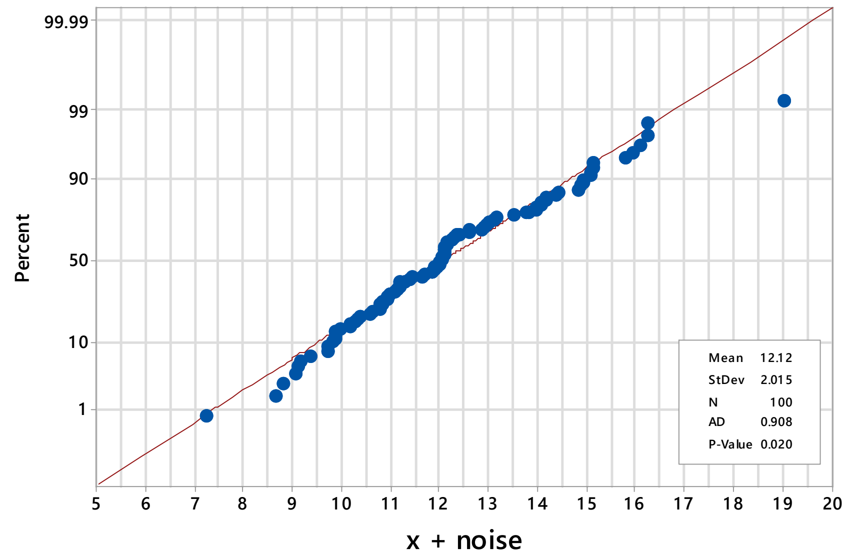
Coarse Data

- If the measurement scale is too coarse and the range of the observed values is too small then the observations may only occupy a few bins.
- The Anderson-Darling statistic will reject the claim that the data are normal with a small p value less than 0.05. Don't worry yet. The A-D test may be reacting to the coarseness of the data and not to a nonnormal deviation in the data.
- Focus on the big picture in the normal plot.
- If you want to feel better about the normal plot, *jitter* the data by adding random normal noise to the measurement values with standard deviation equal to 1/4 of the measurement scale resolution and normal plot the jittered data.

Coarse Data

- The coarse data in the normal plot on the left upsets the A-D test.
- The normal plot on the right has random normal noise with $\sigma = 0.25$ added to the measurement values.



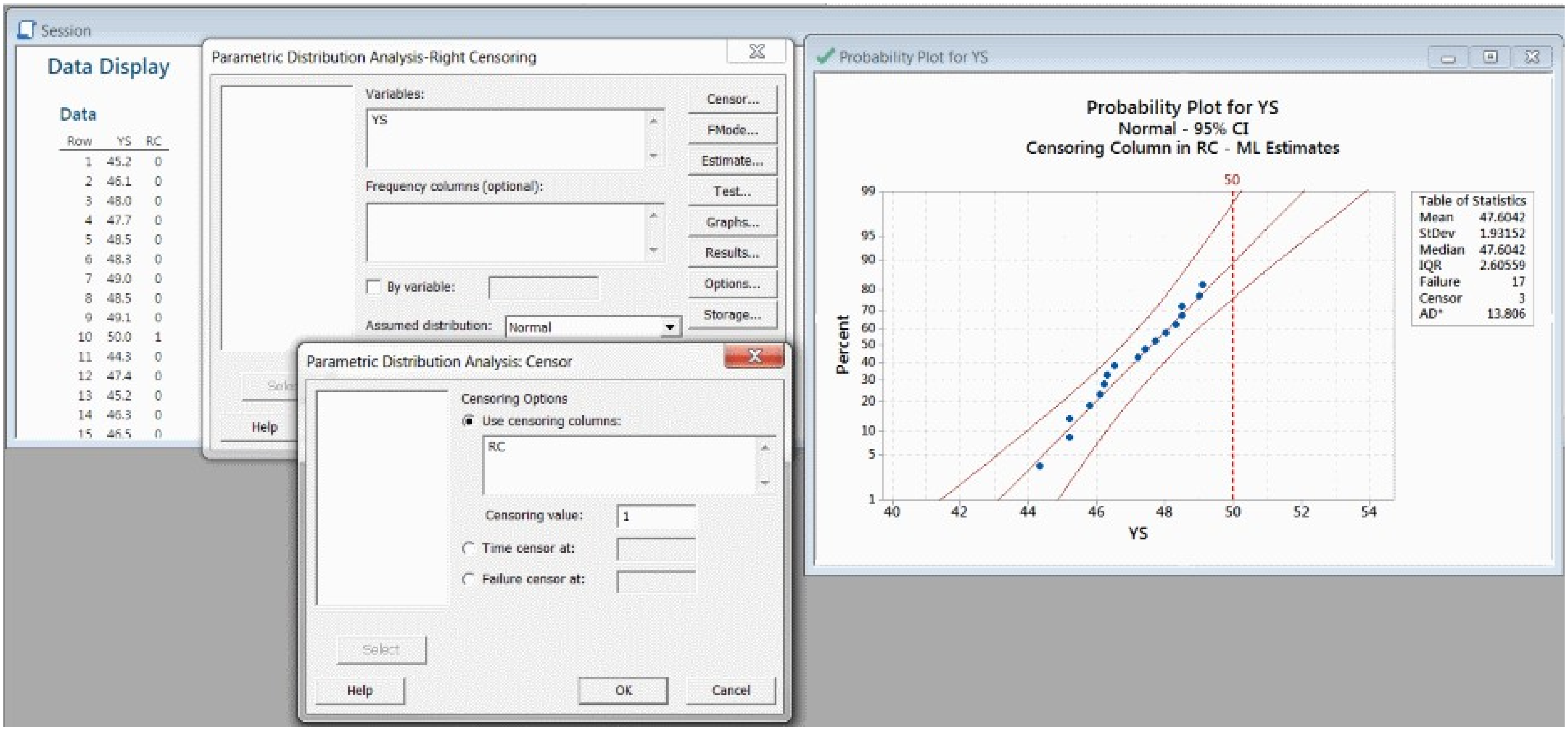


Censored Data

- If the measured response has an upper bound beyond which it cannot measure and observations could fall above that upper bound, then those observations are *right-censored*. Common examples of right-censored data are:
  - Product life data where the response is time or number of cycles to failure and the study has been suspended before all of the units have failed
  - Post-inspection data where the units that exceed *USL* have been culled
- Special reliability analysis methods must be used to analyze right-censored data. In MINITAB use the **Stat**> **Reliability**/**Survival**> **Distribution Analysis** (**Right Censoring**) menu.
- Observations can also be left-censored or interval-censored. Analyze those using the **Stat**> **Reliability**/**Survival**> **Distribution Analysis** (**Arbitrary Censoring**) menu.
- We'll talk more about censored data in Chapter 11 (Reliability).

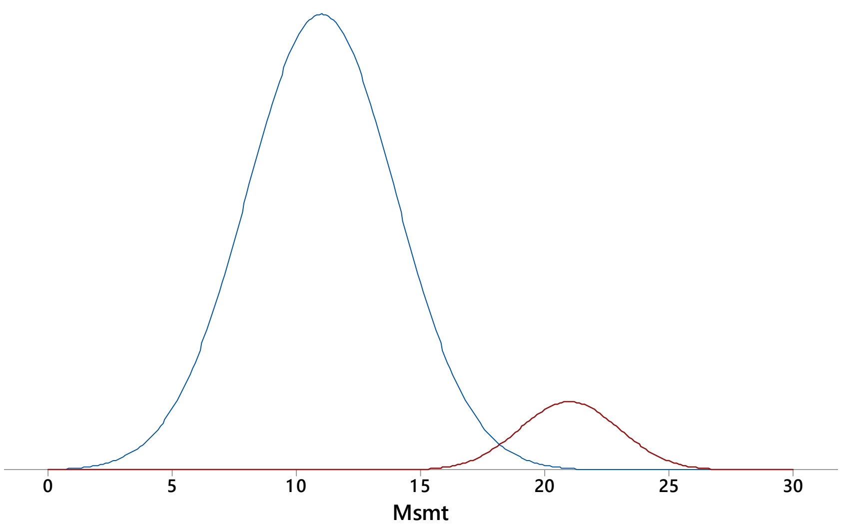
Right-Censored Data

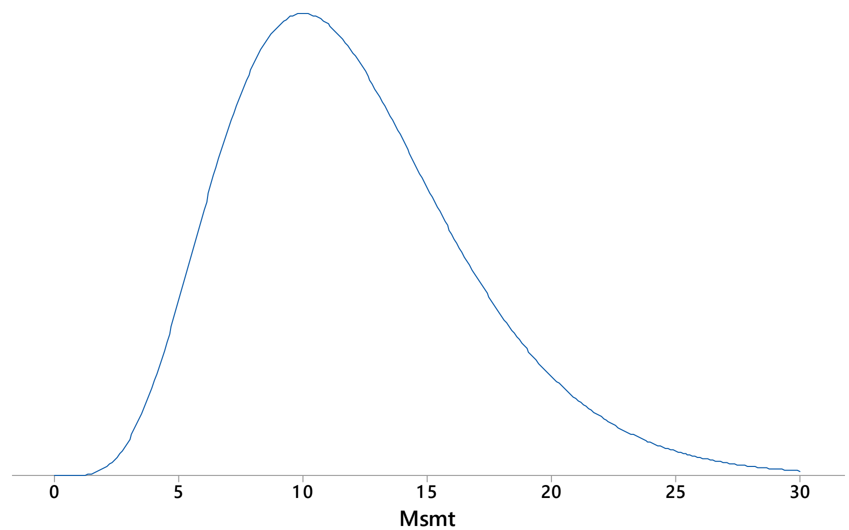
Example: A tensile test was performed to determine the yield strength of a brass alloy. Some of the units in the sample exceded the tensile tester's upper limit of 50kpsi.



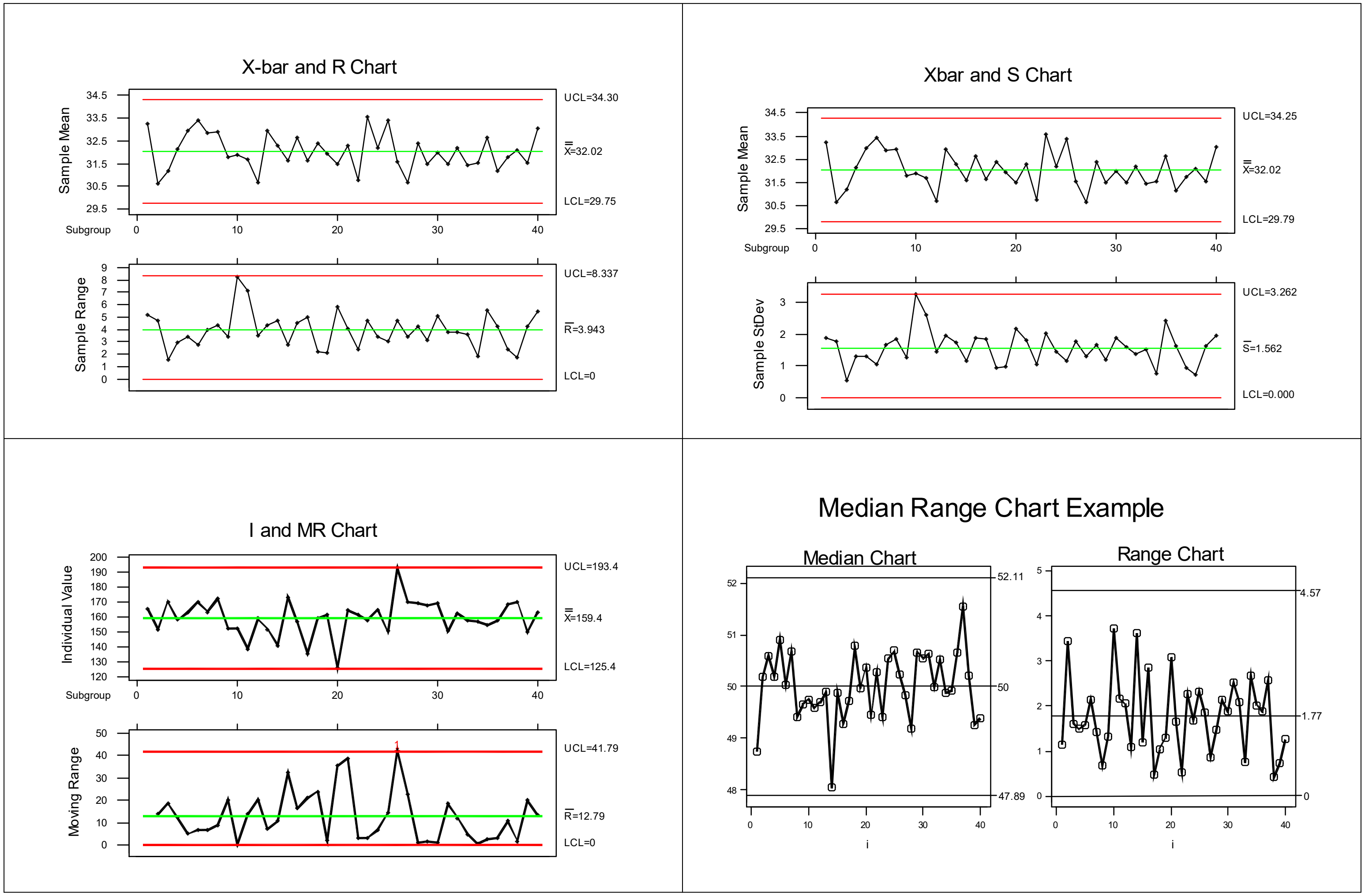
Outliers

- Causes of outliers
  - Data error, e.g. transposed digits
  - Botched run
  - Different population than the rest
- Every outlier is associated with a special cause. The rest suffer common cause variation.
- Is it really an outlier?
  - Does the outlier come from a different population?
  - Or is it just evidence for a long tail?





Statistical Process Control



SPC Variables Charts

- The most common charts used to control measurement data are $\bar{x}$ and R charts.
- The $\bar{x}$ chart tracks changes in location.
- The R chart tracks changes in variation.

$\bar{x}$ and R Chart Procedure

1. Samples of fixed size are drawn from the process at regular intervals. A typical sample size is $n = 5$.

2. The sample mean $\bar{x}$ and range R are determined for each subgroup.

3. $\bar{x}$ and R are plotted on separate charts vs. time.

4. After $m \geq 20$ subgroups have been collected the center lines and control limits for the charts are determined from:

$$CL_{\bar{x}} = \bar{\bar{x}} = \frac{1}{m}\sum_{j=1}^{m}\bar{x}_j$$

$$CL_R = \bar{R} = \frac{1}{m}\sum_{j=1}^{m}R_j$$

$\bar{x}$ and R Chart Procedure

$$(UCL/LCL)_R = \bar{R} \pm 3\sigma_R$$
$$= D_4\bar{R}/D_3\bar{R}$$

$$(UCL/LCL)_{\bar{x}} = \bar{\bar{x}} \pm 3\sigma_{\bar{x}}$$
$$= \bar{\bar{x}} \pm A_2\bar{R}$$

where D_4, D_3, and A_2 can be found in an appropriate table.

5. A detailed process log **must** be kept.

6. The process is managed by keeping hands off when the process is in control and by taking action when it is out of control.

7. The control limits should be revised whenever the process has improved significantly.

Control Limits for $\bar{x}$ Charts

$$
\begin{aligned}
(UCL/LCL)_{\bar{x}} &= \bar{\bar{x}} \pm 3\sigma_{\bar{x}} \\
&= \bar{\bar{x}} \pm 3\frac{\sigma_x}{\sqrt{n}} \\
&= \bar{\bar{x}} \pm 3\left(\frac{\bar{R}}{d_2\sqrt{n}}\right) \\
&= \bar{\bar{x}} \pm \left(\frac{3}{d_2\sqrt{n}}\right)\bar{R} \\
&= \bar{\bar{x}} \pm A_2\bar{R}
\end{aligned}
$$

where $\sigma_x \simeq \frac{\bar{R}}{d_2}$ and $A_2 = \frac{3}{d_2\sqrt{n}}$.

Control Limits for R Charts

$$
\begin{aligned}
(UCL/LCL)_R &= \bar{R} \pm 3\sigma_R \\
&= \bar{R} \pm 3(d_3\sigma_x) \\
&= \bar{R} \pm 3d_3\left(\frac{\bar{R}}{d_2}\right) \\
&= \bar{R} \pm \left(\frac{3d_3}{d_2}\right)\bar{R} \\
&= \bar{R}\left(1 \pm \left(\frac{3d_3}{d_2}\right)\right) \\
&= D_4\bar{R}/D_3\bar{R}
\end{aligned}
$$

where $\sigma_R \simeq d_3\sigma_x$, $D_4 = 1 + \frac{3d_3}{d_2}$, and $D_3 = 1 - \frac{3d_3}{d_2}$.

Example: Create the $\bar{x}$ and R control chart for the following data:

Subgroups	x1	x2	x3	x4	x5	Mean	Range
1	33.6	35.8	30.5	33.6	32.6		
2	31.3	33.0	28.3	31.0	29.5		
3	30.4	31.1	32.0	31.2	31.0		
4	30.6	33.5	31.6	33.4	31.5		
5	33.4	34.5	31.0	32.4	33.5		
6	34.1	33.4	34.7	31.9	32.9		
7	32.2	31.6	33.4	31.6	35.5		
8	33.5	34.2	30.5	31.5	34.9		
9	30.4	31.8	31.9	31.0	33.7		
10	29.8	35.7	27.5	33.6	32.9		
35	30.8	34.9	34.0	29.3	34.2		
36	33.5	29.3	31.6	31.2	30.1		
37	30.6	32.4	33.0	31.5	31.3		
38	32.8	31.1	32.5	32.4	31.7		
39	31.7	33.4	32.6	29.1	30.9		
40	30.5	33.3	32.6	32.8	36.0		
						x-bar-bar =	R-bar =

Solution:

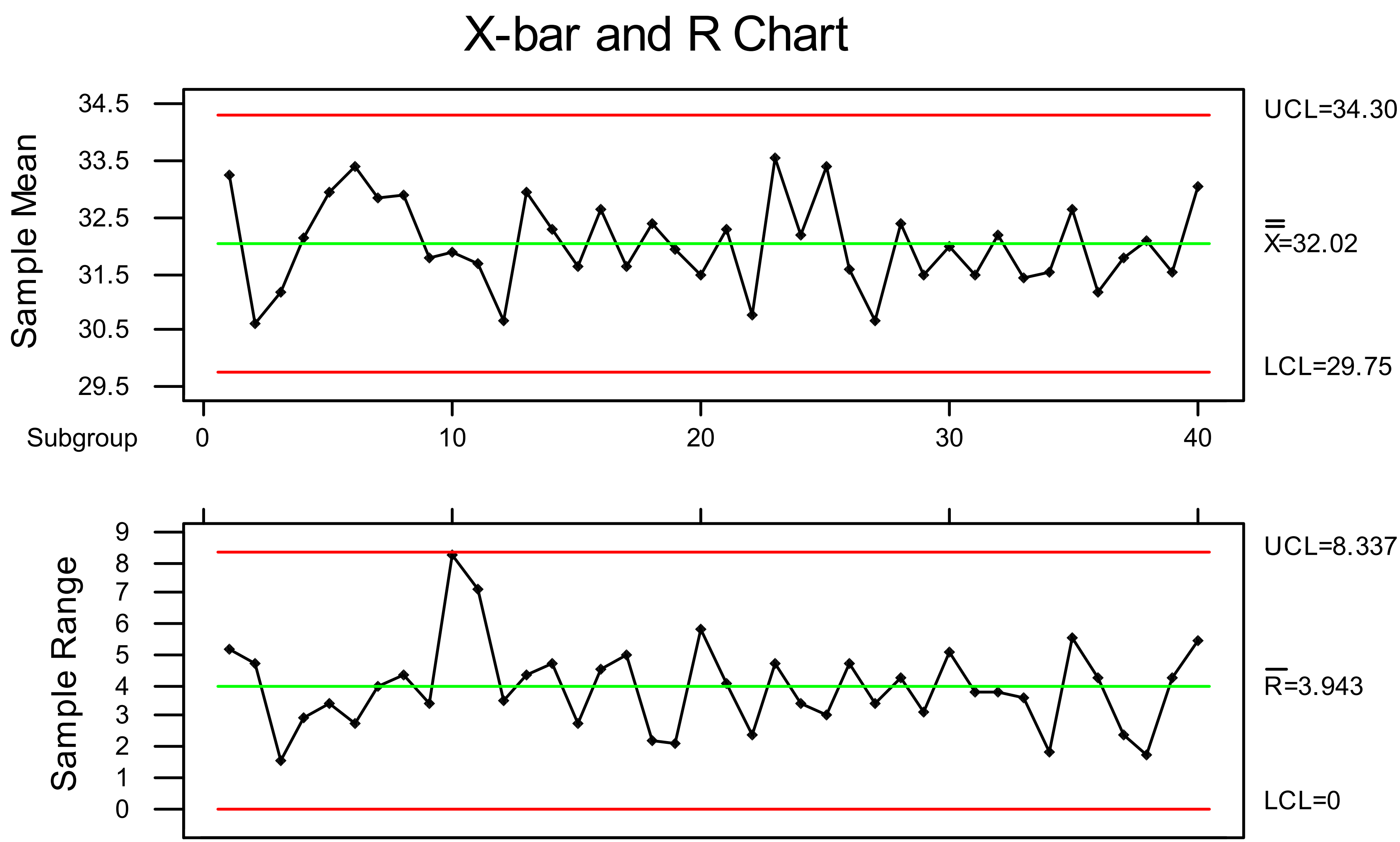
Subgroups	x1	x2	x3	x4	x5	Mean	Range
1	33.6	35.8	30.5	33.6	32.6	33.2	5.2
2	31.3	33.0	28.3	31.0	29.5	30.6	4.7
3	30.4	31.1	32.0	31.2	31.0	31.2	1.5
4	30.6	33.5	31.6	33.4	31.5	32.1	2.9
5	33.4	34.5	31.0	32.4	33.5	33.0	3.5
6	34.1	33.4	34.7	31.9	32.9	33.4	2.7
7	32.2	31.6	33.4	31.6	35.5	32.9	4.0
8	33.5	34.2	30.5	31.5	34.9	32.9	4.4
9	30.4	31.8	31.9	31.0	33.7	31.8	3.4
10	29.8	35.7	27.5	33.6	32.9	31.9	8.3
35	30.8	34.9	34.0	29.3	34.2	32.6	5.6
36	33.5	29.3	31.6	31.2	30.1	31.1	4.3
37	30.6	32.4	33.0	31.5	31.3	31.7	2.4
38	32.8	31.1	32.5	32.4	31.7	32.1	1.7
39	31.7	33.4	32.6	29.1	30.9	31.5	4.3
40	30.5	33.3	32.6	32.8	36.0	33.0	5.4
						x-bar-bar = 32.0	R-bar = 3.94

From the spreadsheet we have $\bar{\bar{x}} = 32.0$ and $\bar{R} = 3.94$. The control limits are:

$$
\begin{aligned}
(UCL/LCL)_{\bar{x}} &= \bar{\bar{x}} \pm A_2\bar{R} \\
&= 32.0 \pm (0.577 \times 3.94) \\
&= 32.0 \pm 2.3 \\
&= 34.3/29.7
\end{aligned}
$$

$$
\begin{aligned}
(UCL/LCL)_{R} &= D_4\bar{R}/D_3\bar{R} \\
&= (2.115 \times 3.94)/(0 \times 3.94) \\
&= 8.34/0
\end{aligned}
$$

where A_2, D_4 and D_3 values come from the Table of *Factors for Constructing Variables Control Charts*.


X-bar and R Chart
Sample Mean
34.5
33.5
32.5
31.5
30.5
29.5
UCL=34.30
X=32.02
LCL=29.75
Subgroup
0
10
20
30
40
Sample Range
9
8
7
6
5
4
3
2
1
0
UCL=8.337
R=3.943
LCL=0

$\bar{x}$ and s Charts

- The center line for the $\bar{x}$ chart is:

$$CL_{\bar{x}} = \bar{\bar{x}}$$

- The center line for the s chart is:

$$CL_s = \bar{s}$$

- The $(UCL/LCL)_{\bar{x}}$ for the $\bar{x}$ chart are:

$$(UCL/LCL)_{\bar{x}} = \bar{\bar{x}} \pm A_3\bar{s}$$

- The $(UCL/LCL)_s$ for the s chart are:

$$(UCL/LCL)_s = B_4\bar{s}/B_3\bar{s}$$

Example: Create the $\bar{x}$ and s control chart for the following data:

Subgroups	x1	x2	x3	x4	x5	x-bar	s
1	33.6	35.8	30.5	33.6	32.6	33.2	1.898
2	31.3	33.0	28.3	31.0	29.5	30.6	1.792
3	30.4	31.1	32.0	31.2	31.0	31.2	0.548
4	30.6	33.5	31.6	33.4	31.5	32.1	1.286
5	33.4	34.5	31.0	32.4	33.5	33.0	1.306
6	34.1	33.4	34.7	31.9	32.9	33.4	1.062
7	32.2	31.6	33.4	31.6	35.5	32.9	1.672
8	33.5	34.2	30.5	31.5	34.9	32.9	1.836
9	30.4	31.8	31.9	31.0	33.7	31.8	1.269
10	29.8	35.7	27.5	33.6	32.9	31.9	3.259
35	30.8	34.9	34.0	29.3	34.2	32.6	2.430
36	33.5	29.3	31.6	31.2	30.1	31.1	1.624
37	30.6	32.4	33.0	31.5	31.3	31.7	0.950
38	32.8	31.1	32.5	32.4	31.7	32.1	0.698
39	31.7	33.4	32.6	29.1	30.9	31.5	1.639
40	30.5	33.3	32.6	32.8	36.0	33.0	1.947
						32.2	1.576

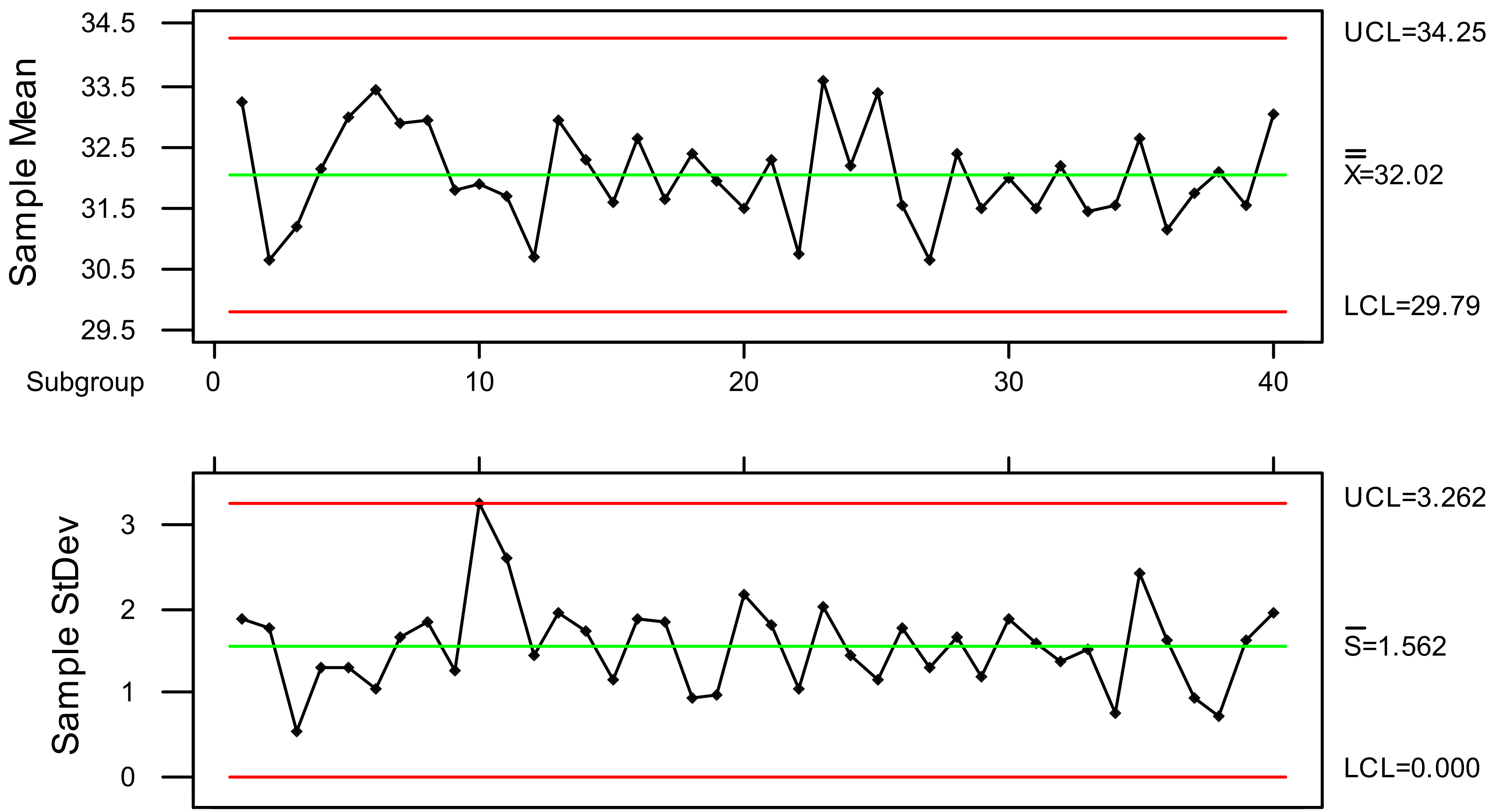
From the spreadsheet we have $\bar{\bar{x}} = 32.0$ and $\bar{s} = 1.56$. The control limits are:

$$
\begin{aligned}
(UCL/LCL)_{\bar{x}} &= \bar{\bar{x}} \pm A_3\bar{s} \\
&= 32.0 \pm (1.427 \times 1.56) \\
&= 32.0 \pm 2.23 \\
&= 34.2/29.8
\end{aligned}
$$

$$
\begin{aligned}
(UCL/LCL)_{s} &= B_4\bar{s}/B_3\bar{s} \\
&= (2.089 \times 1.562)/(0 \times 1.562) \\
&= 3.26/0
\end{aligned}
$$

where A_3, B_4, and B_3 values come from the Table of *Factors for Constructing Variables Control Charts*.

Xbar and S Chart



$\tilde{x}$ and R Charts

- The center line for the Median chart is:

$$CL_{\tilde{x}} = \bar{\tilde{x}}$$

- The center line for the Range chart is:

$$CL_R = \bar{R}$$

- The $(UCL/LCL)_{\tilde{x}}$ for the Median chart are:

$$(UCL/LCL)_{\tilde{x}} = \bar{\tilde{x}} \pm \tilde{A}_2\bar{R}$$

where

$$\tilde{A}_2(n = 3) = 1.19$$
$$\tilde{A}_2(n = 5) = 0.69$$

- The $(UCL/LCL)_R$ for the Range chart are:

$$(UCL/LCL)_R = D_4\bar{R}/0$$

Example: Construct the $\tilde{x}$ and R chart for the following data:

Subgroup	x1	x2	x3	Median	Range
1	48.7	49.1	48.0		
2	48.9	50.2	52.3		
3	51.1	49.5	50.6		
4	50.2	49.4	50.9		
5	50.9	51.3	49.7		
6	51.8	49.6	50.0		
7	50.7	50.0	51.5		
8	49.2	49.4	49.8		
9	49.5	50.8	49.7		
10	48.6	49.8	52.3		
35	50.7	48.7	49.9		
36	50.7	49.8	51.7		
37	51.6	52.6	50.0		
38	50.2	50.0	50.4		
39	49.9	49.2	49.3		
40	49.0	49.4	50.3		

Solution:

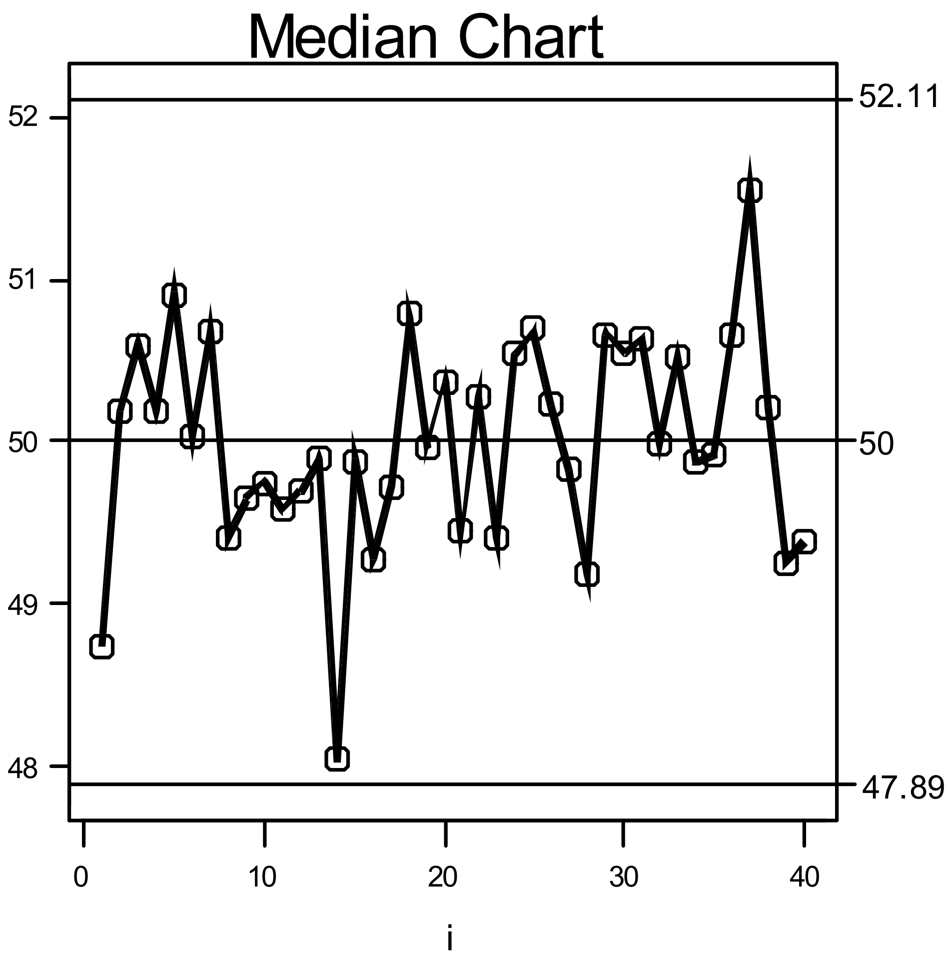
Subgroup	x1	x2	x3	Median	Range
1	48.7	49.1	48.0	48.7	1.1
2	48.9	50.2	52.3	50.2	3.5
3	51.1	49.5	50.6	50.6	1.6
4	50.2	49.4	50.9	50.2	1.5
5	50.9	51.3	49.7	50.9	1.6
6	51.8	49.6	50.0	50.0	2.1
7	50.7	50.0	51.5	50.7	1.4
8	49.2	49.4	49.8	49.4	0.7
9	49.5	50.8	49.7	49.7	1.3
10	48.6	49.8	52.3	49.8	3.7
35	50.7	48.7	49.9	49.9	2.0
36	50.7	49.8	51.7	50.7	1.9
37	51.6	52.6	50.0	51.6	2.6
38	50.2	50.0	50.4	50.2	0.4
39	49.9	49.2	49.3	49.3	0.7
40	49.0	49.4	50.3	49.4	1.3
				50.1	1.7

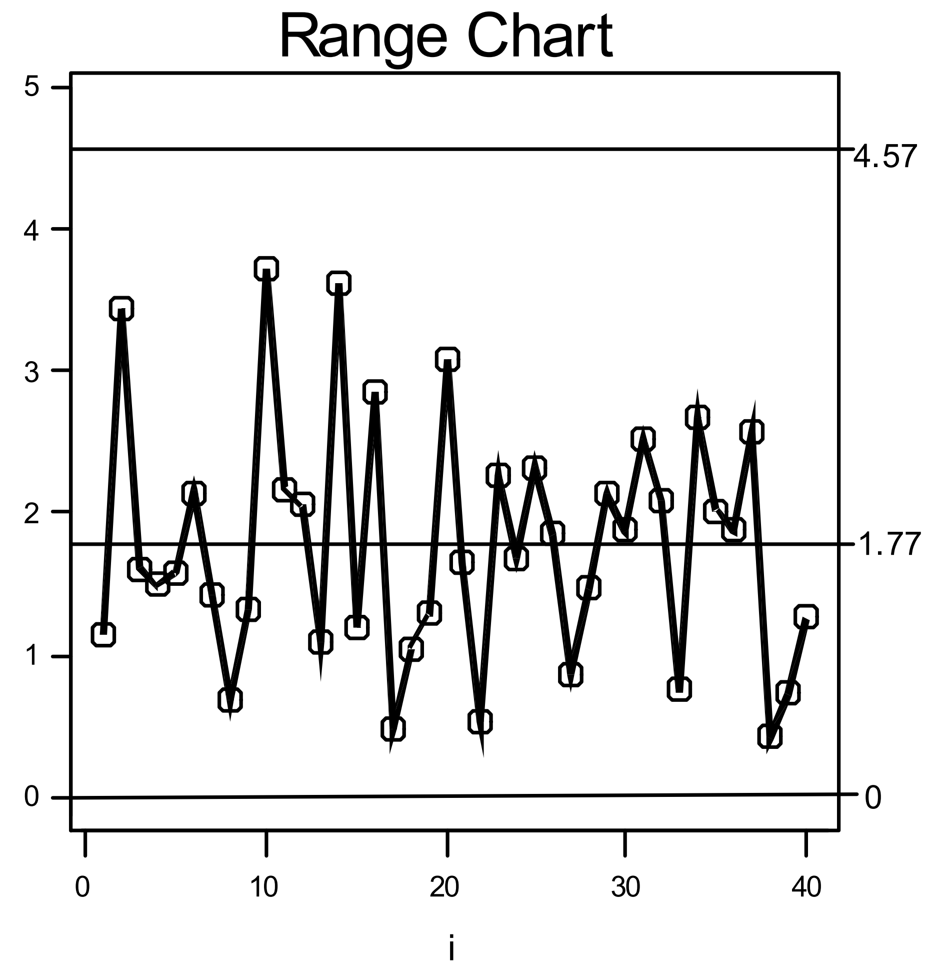
From the spreadsheet we have $\bar{\tilde{x}} = 50.0$ and the $\bar{R} = 1.7$. The control limits are:

$$
\begin{aligned}
(UCL/LCL)_{\tilde{x}} &= \bar{\tilde{x}} \pm \tilde{A}_2\bar{R} \\
&= 50.0 \pm (1.19 \times 1.7) \\
&= 50.0 \pm 2.11 \\
&= 52.11/47.89
\end{aligned}
$$

$$
\begin{aligned}
(UCL/LCL)_R &= D_4\bar{R}/0 \\
&= (2.575 \times 1.77)/0 \\
&= 4.57/0
\end{aligned}
$$

Median Range Chart Example





x and MR Charts

- x and MR charts are used when $n = 1$
- MR is the moving range determined from the difference between successive values:

$$MR_j = |x_j - x_{j-1}|$$

- x and MR charts are also called I and MR charts where I stands for individuals

x and MR Charts

- The center line for the x chart is given by:

$$CL_x = \bar{x} = \frac{1}{m}\sum_{j=1}^{m} x_j$$

where m is the number of subgroups (or points in this case).

- The center line for the MR chart is given by:

$$CL_{MR} = \overline{MR} = \frac{1}{m-1}\sum_{j=2}^{m} MR_j$$

where MR_1 does not exist.

x and MR Charts

- The $(UCL/LCL)_x$ are given by:

$$(UCL/LCL)_x = \bar{x} \pm 3\overline{MR}/d_2$$

  where d_2 is taken with $n = 2$ since two observations are used to determine each MR.

- The UCL_{MR} is given by:

$$UCL_{MR} = D_4\overline{MR}$$

  where D_4 is determined from $n = 2$ since two observations are used to determine each MR.

- The LCL_{MR} is given by:

$$\begin{aligned} LCL_{MR} &= D_3\overline{MR} \\ &= 0 \end{aligned}$$

  where $D_3 = 0$ for $n = 2$.

Example: Construct the *XMR* chart for the following data:

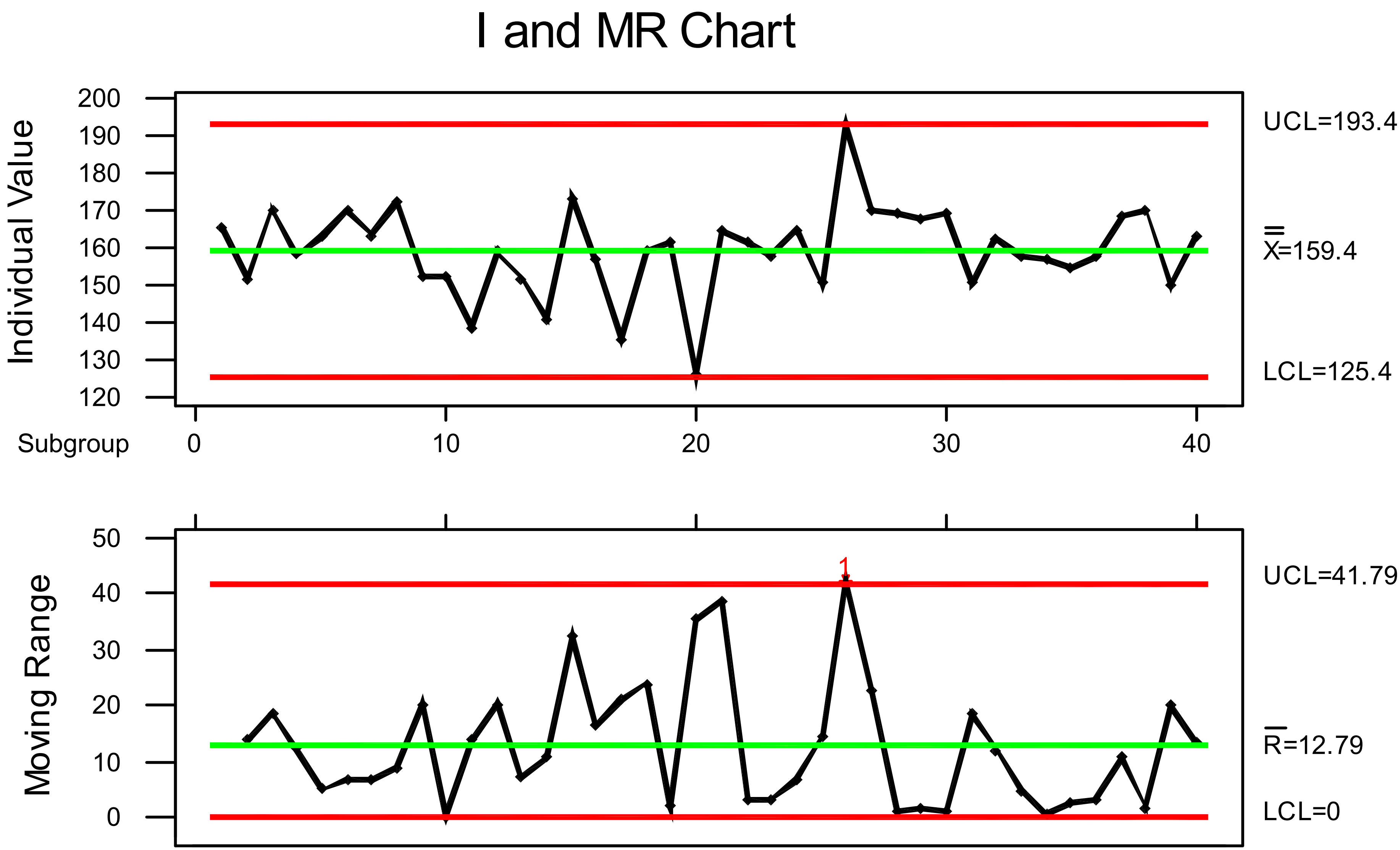
i	x	MR
1	165	*
2	151	14
3	170	19
4	158	12
5	163	5
35	154	3
36	157	3
37	168	11
38	170	2
39	150	20
40	163	13
Sum:	6373	499
Count:	40	39
Mean:	159.3	12.8

From the spreadsheet we have $\bar{x} = 159.3$ and $\overline{MR} = 12.8$. The control limits are:

$$
\begin{aligned}
(UCL/LCL)_x &= \bar{x} \pm 3\overline{MR}/d_2 \\
&= 159.3 \pm (3 \times 12.8)/1.128 \\
&= 159.3 \pm 34.0 \\
&= 193/125
\end{aligned}
$$

$$
\begin{aligned}
UCL/LCL_{MR} &= D_4\overline{MR}/0 \\
&= (3.267 \times 12.8)/0 \\
&= 42/0
\end{aligned}
$$

where $d_2$and D_4 values come from the Table of *Factors for Constructing Variables Control Charts*.



I and MR Chart
Individual Value
200
190
180
170
160
150
140
130
120
UCL=193.4
$\bar{\bar{X}}$=159.4
LCL=125.4
Subgroup
0
10
20
30
40
Moving Range
50
40
30
20
10
0
1
UCL=41.79
$\bar{R}$=12.79
LCL=0

Beyond the Shewhart Paradigm

- Short Run SPC Charts - SPC is statistical **process** control.
  - Use target charts when σ is constant over part types. Plot $x'_{ip} = x_{ip} - \mu_p$ versus time.
  - Use standardized charts when σ is not constant over part types. Plot $z_{ip} = (x_{ip} - \mu_p)/\sigma_p$ versus time.
- Modified and Acceptance Charts - *UAL*/*LAL* are calculated inward from *USL*/*LSL* rather than outward from μ:
  - $p = AQL$ is the fraction defective that will be accepted with probability $1 - \alpha$:

$$UAL/LAL = USL/LSL \mp \left(z_{AQL} - \frac{z_{\alpha/2}}{\sqrt{n}} \right) \sigma$$

  - $p = RQL$ is the fraction defective that will be accepted with probability β:

$$UAL/LAL = USL/LSL \mp \left(z_{RQL} + \frac{z_{\beta}}{\sqrt{n}} \right) \sigma$$

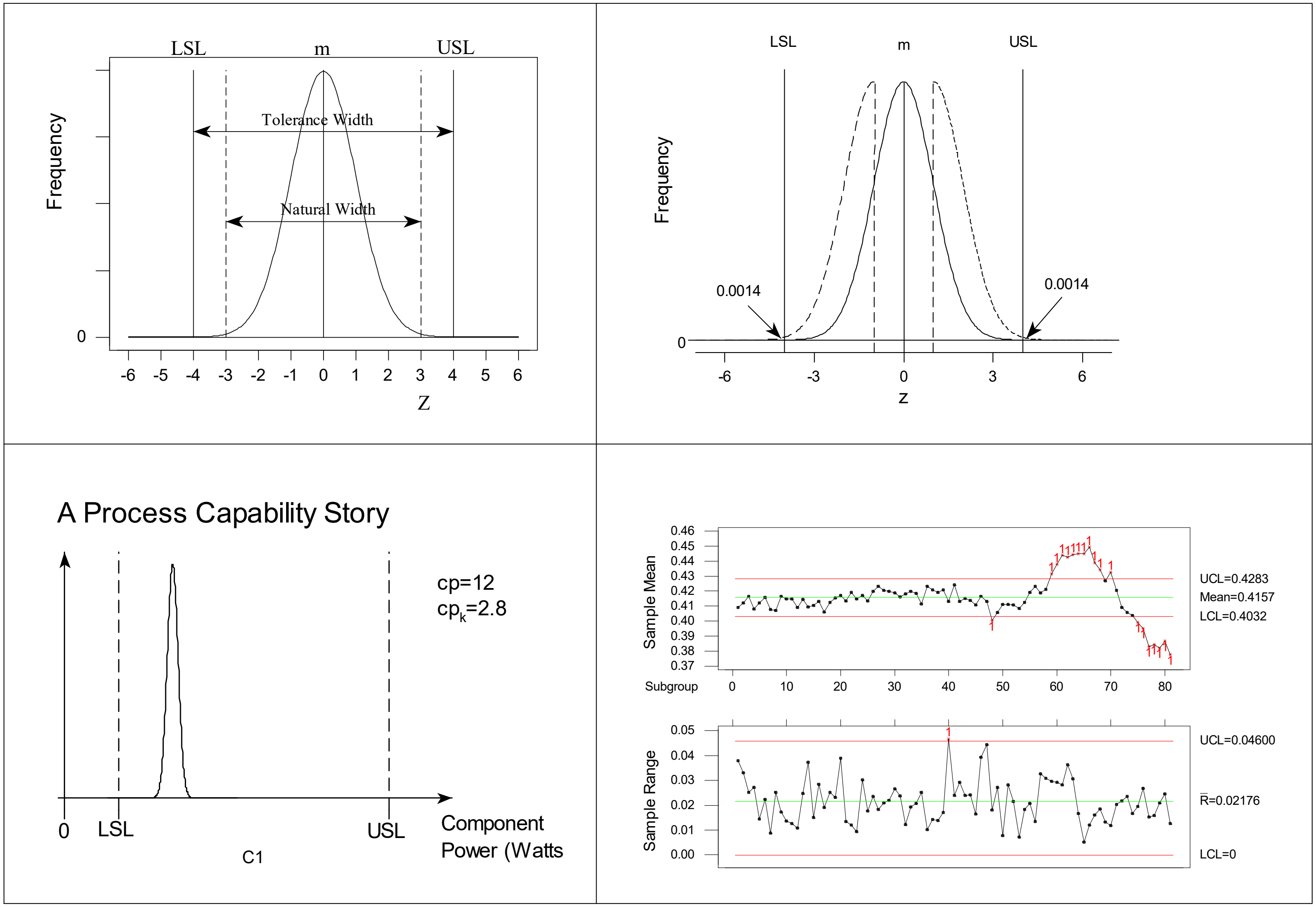
Beyond the Shewhart Paradigm

- Special Averaging Charts - These charts offer better sensitivity to small shifts in the process mean than that of Shewhart charts:
  - Cusum Charts: Plot $c_i = \sum_{j=1}^{i}(\bar{x}_j - m)$ vs. i where $x = m$ is the process target.
  - Exponentially Weighted Moving Average Charts (EWMA): Plot $z_i = \lambda x_i + (1 - \lambda)z_{i-1}$ vs. i where $0 < \lambda < 1$. ($\lambda = 0.2$ is popular.)
  - Moving Average (or Boxcar Average) Charts: Plot $m_i = \frac{1}{w}(x_i + x_{i-1} + \cdots + x_{i-w+1})$ vs. i where w is the number of most recent observations considered in the statistic.
- Group Charts for Multiple-Stream Processes - e.g. units from a multiple cavity mold.
- Multivariate Charts - for combining two or more related responses, e.g. x and y deviations from a target position.

Process Precontrol

- Little sister to SPC
- Can be used on long or short runs
- Starts in 100% inspection mode
- Switches to sampling mode when there is evidence that the process is stable.
- Also sensitive to process capability.
- If documented correctly, can be used to validate process capability.
- Very easy to use and can be infectious among operators.
- See QES Appendix A: Process Precontrol

Process Capability Statistics

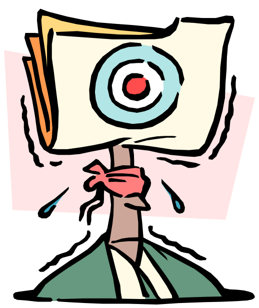


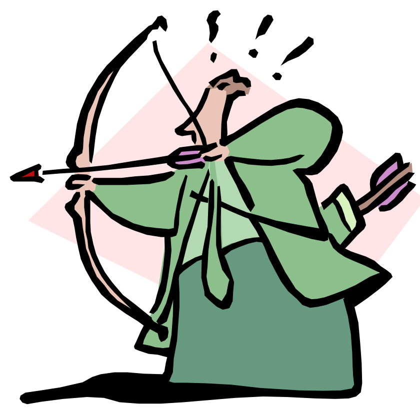
Process Capability Statistics

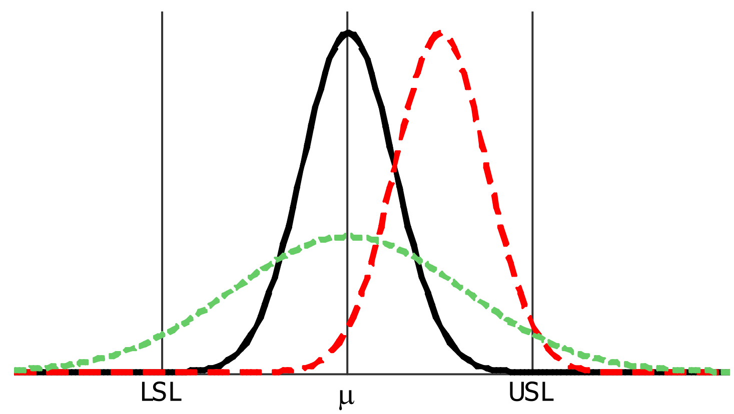
Processes can make excessive defective material if:

1. There is too much variation in the process.
2. The process mean shifts from the desired target value.

SPC charts are great for monitoring the location and variation of a process, but a technique called **process capability analysis** is necessary to compare the performance of the process to the specification limits. There are many measures of process capability but the most common ones are c_p and c_{pk}.







The c_p Statistic

The c_p statistic compares the tolerance width of a process to the natural width of the process:

$$c_p = \frac{\text{Tolerance Width}}{\text{Natural Width}}$$

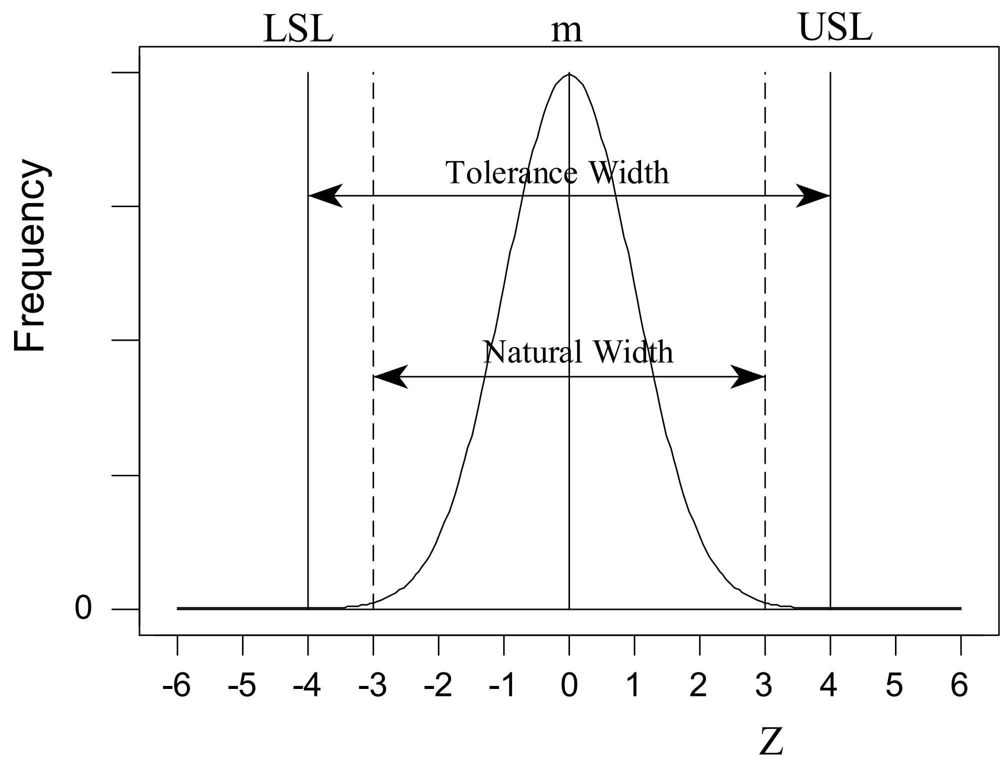
The tolerance width is $USL - LSL$ and the natural width is 6σ $(\pm 3\sigma)$ so c_p is given by:

$$c_p = \frac{USL - LSL}{6\sigma}$$

c_p is used to check if there is a *chance* of producing product within specification.

The c_p Statistic

$$c_p = \frac{\text{Tolerance Width}}{\text{Natural Width}} = \frac{USL-LSL}{6\sigma}$$



The c_p Statistic

- The sample size must be sufficient, typically $n \geq 200$.
- The distribution must be normal.
- Does not exist if the specification is one sided.
- Larger values are more desirable.
- Most companies work hard to obtain $c_p \geq 1.33$.
- A Six Sigma process has $c_p \geq 2.0$.
- c_p will be small (bad) if there is lots of variation in the process relative to the specification width.
- c_p will be large (good) if there is little variation in the process relative to the specification width.

The c_{pk} Statistic

The c_{pk} statistic is given by:

$$c_{pk} = \frac{|NSL - \mu|}{3\sigma}$$

where NSL is the nearest specification limit to the process mean μ. This might also be written:

$$c_{pk} = \min\left\{\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right\}$$

where the $\min$ function indicates that the minimum value from the list is taken.

c_{pk} takes into consideration how well centered the process is.

The c_{pk} Statistic

- The sample size must be sufficient, typically $n \geq 200$.
- The distribution must be normal.
- Exists when the specification is one sided.
- Is always less than or equal to c_p.
- If the process mean is centered within the specification limits then $c_{pk} = c_p$.
- $c_{pk} \geq 1.33$ is a common goal for many companies.
- A Six Sigma process has $c_{pk} \geq 1.5$.

Process Capability Example

Example: Find c_p and c_{pk} if a part has specification $USL/LSL = 0.380 \pm 0.020$ inches and the process runs with a historical process mean of $\bar{\bar{x}} = 0.383$ inches and a standard deviation of $\sigma = 0.004$ inches.

Solution:

$$c_p = \frac{0.400 - 0.360}{6 \times 0.004} = 1.67$$

$$c_{pk} = \frac{|0.400 - 0.383|}{3 \times 0.004} = 1.42$$

If the target process performance is to have c_p and c_{pk} greater than 1.33, then this process meets the goal. The product quality could be improved by centering the process so the new process mean was 0.380 inches. Then the new c_{pk} value would become $c_{pk} = 1.67$ and c_p would be unchanged. Further improvement would require decreasing the process standard deviation.

Process Capability Example

Example: Show that if a process's mean and standard deviation are known, and if a specification is set from $USL/LSL = \mu \pm 4\sigma$, then a value of $c_p = 1.33$ is obtained.

Solution: Since the tolerance width is 8σ then:

$$c_p = \frac{8\sigma}{6\sigma} = 1.33$$

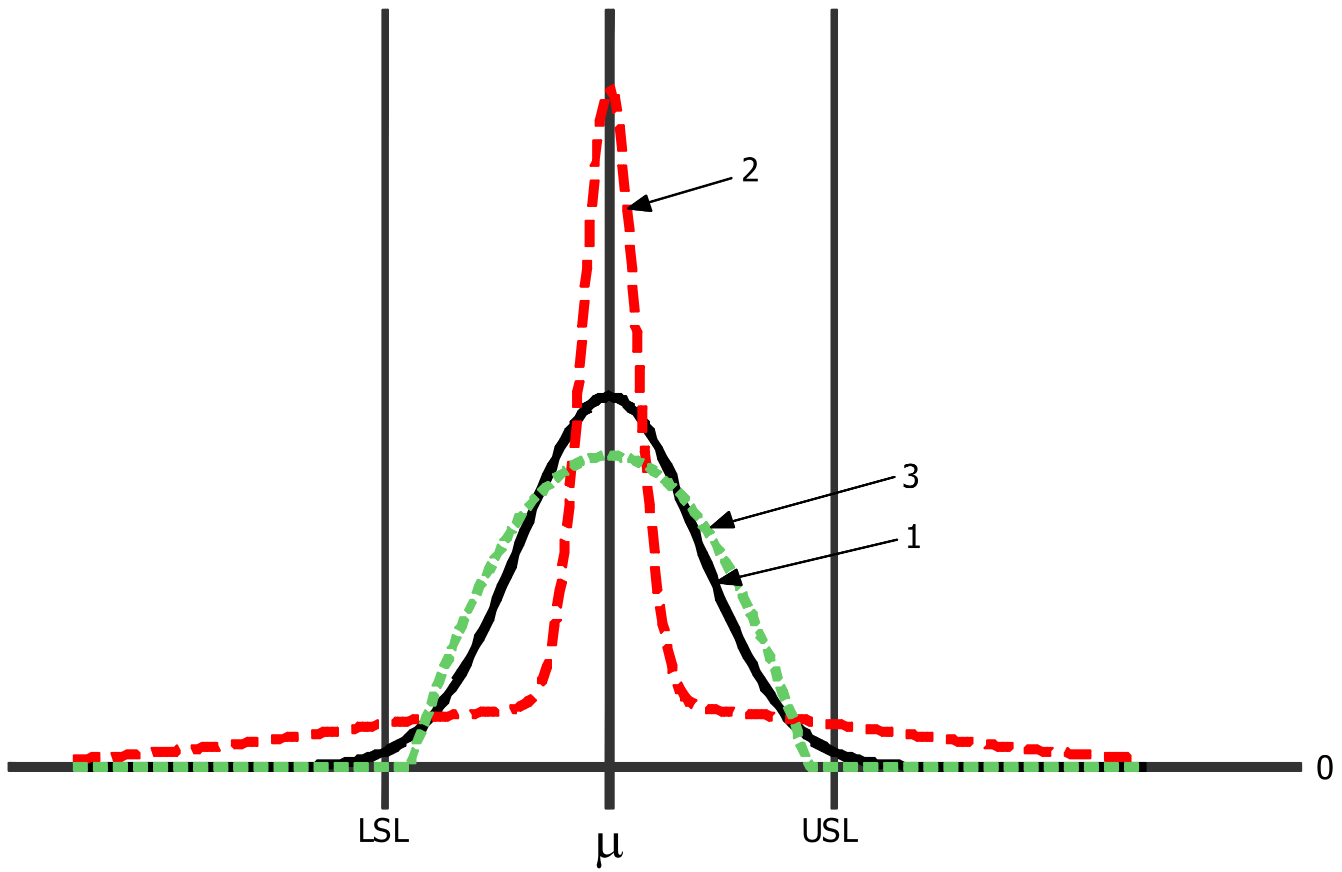
Problem: Determine the c_p value that corresponds to a specification set from $USL/LSL = \mu \pm 6\sigma$.

Solution: Since the tolerance width is 12σ then:

$$c_p = \frac{12\sigma}{6\sigma} = 2.0$$

Process Capability and Normality

c_p and c_{pk} are very sensitive to the normality requirement. If data are not normal then c_p and c_{pk} will over- or under-estimate the true capability of the process.



Process Capability and Normality

- If a distribution is not normal, attempt to transform it to approximate normality using an appropriate transformation, e.g. square root, log, reciprocal, etc., then perform the process capability analysis on the transformed data.
- If a suitable transform cannot be identified by inspection, try the more aggressive methods: the Box-Cox transform and the Johnson transform.
- If the data can't be transformed to approximate normality, special techniques should be used instead of c_p and c_{pk}, such as …
- Universally acceptable ways to state the capability of a process, independent of the normality assumption, are:
  - fraction defective
  - percentage defective
  - defectives per million (dpm)

Notes on Process Capability

- c_p and c_{pk} are very sensitive measures of capability. For example, 1.0 is bad and 2.0 is very very good (for most people).
- Small changes in the values of process capability statistic can cause very large changes in the defective rates.
- It is necessary to look at both c_p and c_{pk} to understand the situation.
- Process capability was created after SPC and to support it. Good SPC data on a process are required to obtain meaningful measures of c_p and c_{pk}.
- Process capability analysis has been so misused that it has been called "statistical terrorism".

Process Capability and SPC

Most people think that they would be happy if their process ran so that $USL/LSL = \mu \pm 3\sigma$. This would give a defective rate of:

$$p = 1 - \Phi(-3 < z < 3) = 0.0027$$

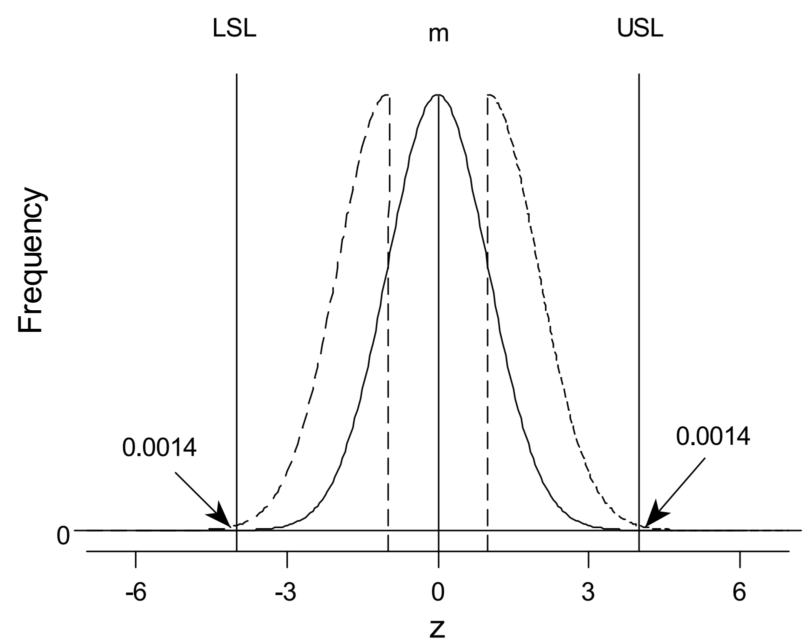
or 0.27% except that process centers (means) always wander and even a tool as powerful as an SPC chart cannot easily detect shifts in the mean as large as $\pm 1\sigma$. If a process with $USL/LSL = \mu \pm 3\sigma$ experienced a shift in the mean of $\pm 1\sigma$ then the defective rate would increase to:

$$p = 1 - \Phi(-2 < z < 4) = 0.023$$

or 2.3% which is usually unacceptable.

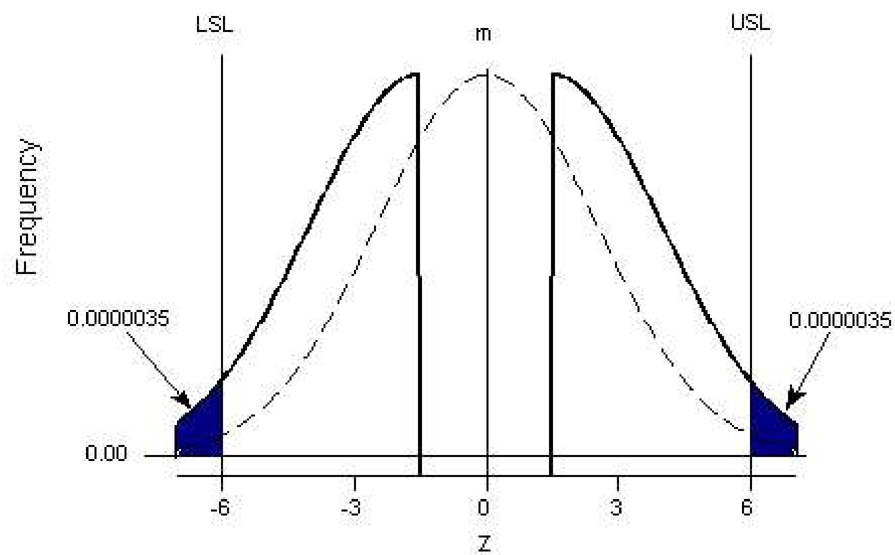
Process Capability and SPC

To compensate for the possibility of a shift in the process mean specifications are padded, usually by $\pm 1\sigma$. This means that a specification of $USL/LSL = \mu \pm 4\sigma$ can tolerate a shift of the mean by $\pm 1\sigma$ and still deliver product with less than 0.14% defective (half of 0.27% since only one tail of the distribution falls outside the specification limits at a time).



Process Capability and SPC

In a Six Sigma program specifications are set at $USL/LSL = \mu \pm 6\sigma$ to plan for shifts in the process mean of up to $\pm 1.5\sigma$. This means that under the worst condition of a $\pm 1.5\sigma$ shift the nearest specification limit will still be 4.5σ from the process mean which results in a defective rate of 0.0000035 or 3.5 defectives per million. However, even if a process has a capability of $c_p = 2.0$ it is necessary to keep track of it using SPC charts because of possible shifts of the process mean.



Example: What is the c_{pk} value for a process that has $c_p = 2.0$ and a process mean that is shifted by $\pm 1.5\sigma$ from its target value?

Solution: With $c_p = 2.0$ the specification limits must fall at $USL/LSL = \mu_0 \pm 6\sigma$ assuming that the process mean μ_0 is centered within the spec. If the mean is shifted to $\mu = \mu_0 \pm 1.5\sigma$ then the distance between the shifted process mean and the nearest spec limit is 4.5σ. This means that the c_{pk} value is:

$$c_{pk} = \frac{|NSL - \mu|}{3\sigma} = \frac{4.5\sigma}{3\sigma} = 1.5$$

This process, with $c_p = 2.0$ and $c_{pk} = 1.5$, will have fraction defective:

$$p = 1 - \Phi(-4.5 < z < 7.5) = 0.0000035$$

or $3.5dpm$.

Relationship To Fraction Defective

Specifying c_p and c_{pk} uniquely determines the fraction defective if: 1) the process is in control and 2) the distribution is normal. The fraction defective is given by:

$$
\begin{aligned}
p &= 1 - \Phi(LSL < x < USL; \mu, \sigma) \\
&= 1 - \Phi(-3c_{pk} < z < 6c_p - 3c_{pk})
\end{aligned}
$$

Example: Determine the fraction defective for a process with $c_p = 2.0$ and $c_{pk} = 1.5$.

Solution:

$$
\begin{aligned}
p &= 1 - \Phi(-3c_{pk} < z < 6c_p - 3c_{pk}) \\
&= 1 - \Phi(-3(1.5) < z < 6(2.0) - 3(1.5)) \\
&= 1 - \Phi(-4.5 < z < 7.5) \\
&= 0.0000035 \\
&= 3.5 dpm
\end{aligned}
$$

Process Capability Quiz

In each case describe the appropriate action. Assume that the goal is to achieve a Six Sigma process.

a. $c_p = 1.1, c_{pk} = 0.8$

b. $c_p = 1.8, c_{pk} = 1.4$

c. $c_p = 1.4, c_{pk} = 1.8$

d. $c_p = 2.8, c_{pk} = 1.4$

e. $c_p = 2.8, c_{pk} = 2.4$

f. $c_p = DNE, c_{pk} = 1.4$

g. $c_p = 12, c_{pk} = 2.6$

Rational Subgroups

From Shewhart: Samples should be chosen to minimize the variation within subgroups. Then variation within subgroups should come from common causes and differences between subgroups will most likely come from special causes. This will increase the sensitivity of control charts to special causes.

We often deviate from Shewhart's recommendation to use the rational subgroup to suit the purpose of the inspection operation.

From a practical standpoint

There are many ways to pick units for a subgroup:

- Select consecutive parts
  - Consecutive parts will be most alike and will have the least variation.
  - Use consecutive parts if the purpose is to determine the true capability of the process.
  - We almost always use consecutive parts!
- Take a random sample of parts from the time interval between samples
  - A random sample will include short-term and longer-term variation
  - Use a random sample of parts if the purpose is to validate the parts produced within a time interval.

Calculating Process Capability

- In practice we don't know σ and must be satisfied with an estimate ($\hat{\sigma}$)
- If an R chart is kept use:

$$\hat{\sigma} = \frac{\bar{R}}{d_2}$$

- If an s chart is kept use:

$$\hat{\sigma} = \frac{\bar{s}}{c_4}$$

- Whether the process is in control or not:

$$\hat{c}_p = \frac{USL - LSL}{6\hat{\sigma}}$$

$$\hat{c}_{pk} = \frac{|NSL - \bar{\bar{x}}|}{3\hat{\sigma}}$$

The c_{pm} Statistic

Whereas the c_{pk} statistic contrasts the specification limits to the process mean, it may make more sense to contrast them to the process target or nominal value m:

$$c_{pm} = \frac{USL - LSL}{6\sqrt{\sigma^2 + (\mu - m)^2}}$$

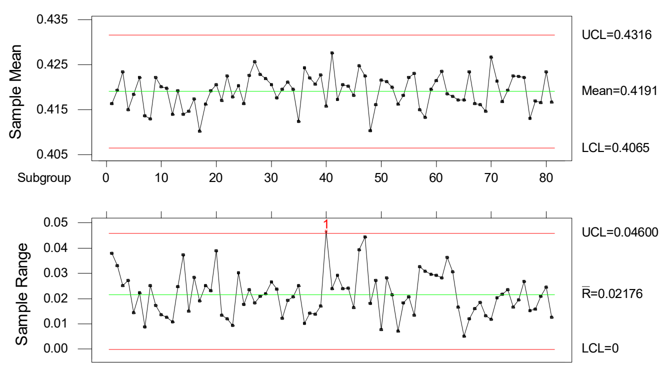
where the variation term in the denominator comes from Taguchi's loss function. When the process is running on target (i.e. $\mu = m$) then $c_{pm} = c_p$.

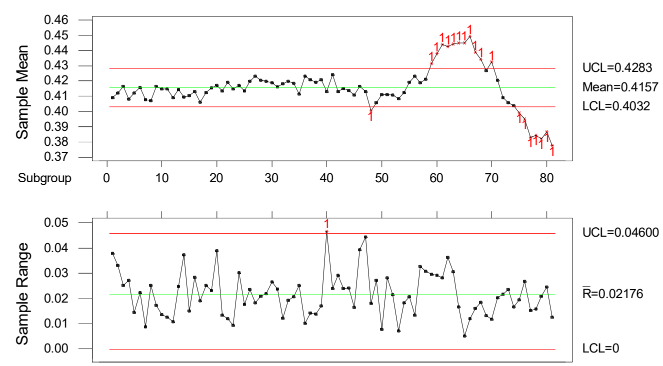
Process Performance Statistics

Suppose that the process is not in control, i.e. the process mean wanders around. Then c_p and c_{pk} tell you what is possible from the process - not what is actually being produced.

Example: Compare the c_p and c_{pk} values for the following two processes. Both processes have the same tolerances and are sampled with $n = 5$.

Solution: Since both of the processes have the exact same $\bar{R}$ value they have the exact same process standard deviation estimate so their c_p and c_{pk} values are identical.





Process Performance Statistics

- To find out what the process is actually delivering calculate:

$$\widehat{P}_p = \frac{USL - LSL}{6\hat{\sigma}_{total}}$$

$$\widehat{P}_{pk} = \frac{|NSL - \bar{\bar{x}}|}{3\hat{\sigma}_{total}}$$

where $\hat{\sigma}_{total}$ is estimated from the sample standard deviation taken over all measurement values:

$$\hat{\sigma}_{total} = \sqrt{\frac{\sum_{j=1}^{m} \sum_{i=1}^{n} (x_{ij} - \bar{\bar{x}})^2}{nm - 1}}$$

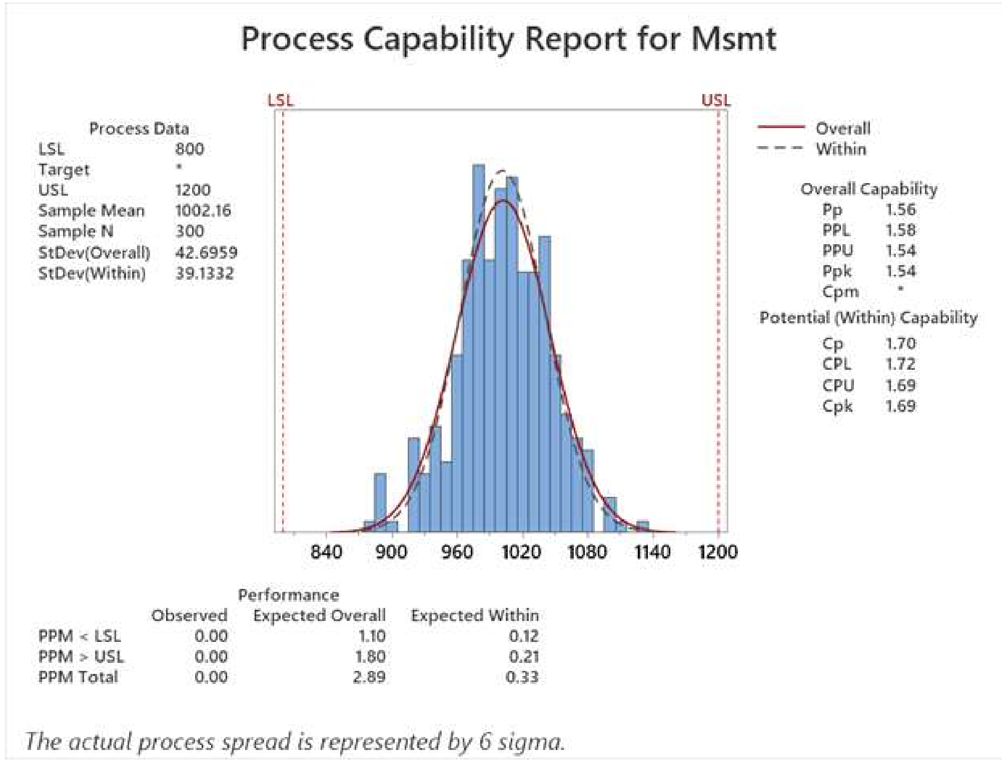
- $\hat{\sigma}_{total}$ includes variation within subgroups plus variation between subgroups

$$\sigma^2_{total} = \sigma^2_{within} + \sigma^2_{between}$$

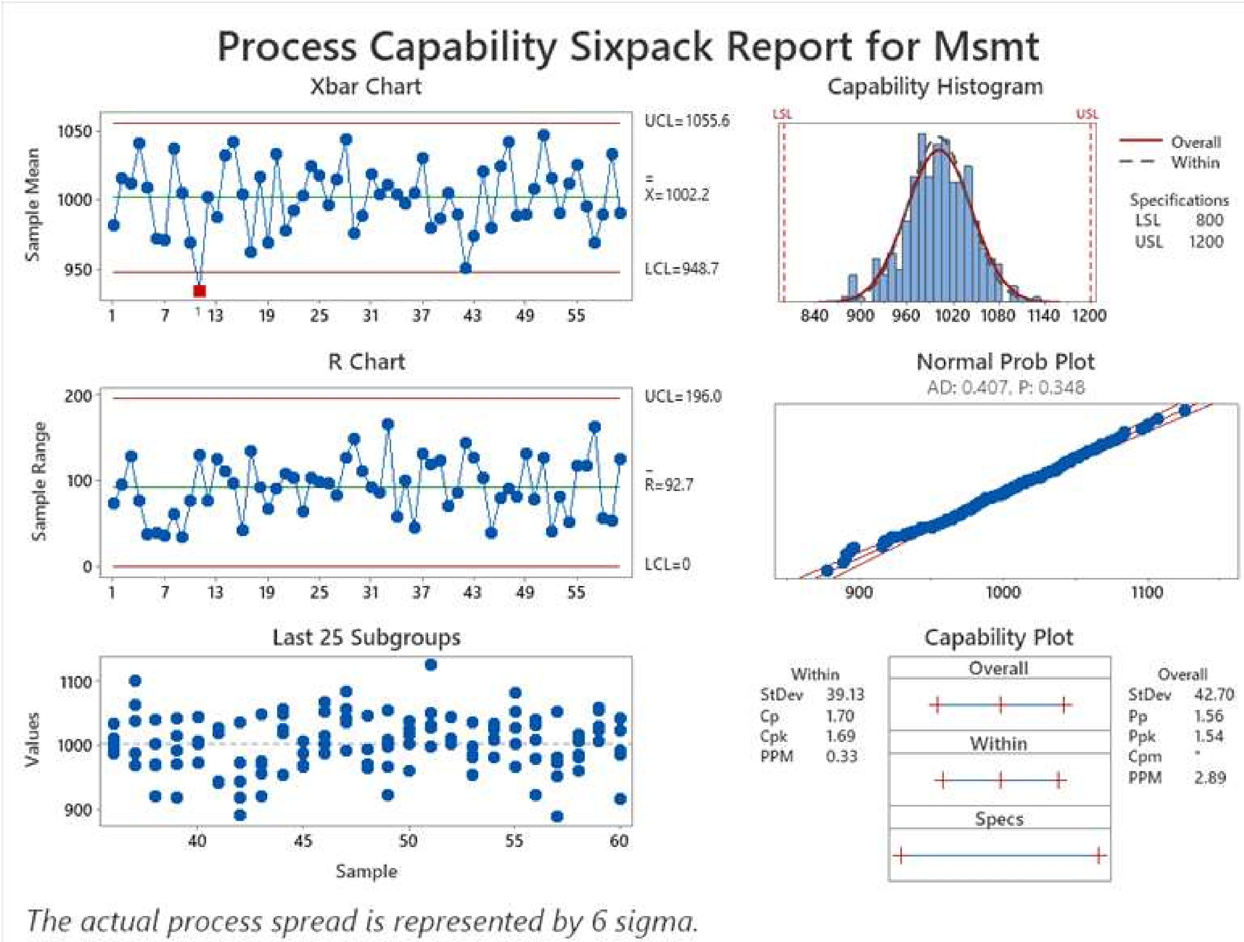
Process Performance Statistics

- c_p and c_{pk} measure short term process capability and are only valid indicators of product quality if the process is in control.
- P_p and P_{pk} consider both short and long term process variation so they are better indicates of actual product quality.
- P_p and P_{pk} will be less than c_p and c_{pk}, respectively, if the process mean wanders around.
- P_p and P_{pk} will be equal to c_p and c_{pk}, respectively, if the process mean is always centered.
- c_p and c_{pk} are the targets for P_p and P_{pk}, respectively.

Process Capability Analysis with MINITAB



Process Capability Analysis with MINITAB



Process Capability/Performance z Scores

- Some organizations prefer to report process capability using normal distribution z scores instead of c_p, c_{pk}, P_p, and P_{pk}.
- z_{LSL} and z_{USL} are the z scores associated with the distances between the mean and the specification limits in standard deviation units:

$$z_{LSL} = \frac{\widehat{\mu} - LSL}{\widehat{\sigma}} \text{ and } z_{USL} = \frac{USL - \widehat{\mu}}{\widehat{\sigma}}$$

where $\widehat{\mu}$ and $\widehat{\sigma}$ may be determined using short term (ST) or long term (LT) variation.

Process Capability/Performance z Scores

- The fractions defective in the left and right tails of the distribution are given by the normal distribution tail areas relative to z_{LSL} and z_{USL}, that is:

$$p_{LSL} = \Phi(-\infty < z < -z_{LSL})$$
$$p_{USL} = \Phi(z_{USL} < z < \infty)$$

- The total fraction defective is the sum of the fractions defective from the left and right tails:

$$p_{total} = p_{LSL} + p_{USL}$$

- p_{total} is used to calculate z_{Bench} - the corresponding one-tailed z score:

$$p_{total} = \Phi(z_{Bench} < z < \infty)$$

Process Capability/Performance z Scores

z_{LSL}, z_{USL}, and z_{bench} are calculated two ways:

- In short term z calculations:
  - $\hat{\mu}$ is set equal to its target value (m or T) and $\hat{\sigma}$ includes short term (within-subgroup) variation but not long term (between subgroup) variation:

$$z_{ST} = \frac{|NSL - T|}{\hat{\sigma}_{ST}}$$

  - When the specification is two-sided z_{ST} and c_p are related by:

$$z_{ST} = 3c_p$$

  - z_{ST} is a measure of how good the process *could* be.

Process Capability/Performance z Scores

- In long term z calculations:
  - $\hat{\mu}$ is set equal to the long term process mean and $\hat{\sigma}$ includes both short term and long term variation:

$$z_{LT} = \frac{|NSL - \hat{\mu}_{LT}|}{\hat{\sigma}_{LT}}$$

  - z_{LT} and P_{pk} are related by:

$$z_{LT} = 3P_{pk}$$

  - z_{LT} is a measure of how good the process actually is.
- The difference between $z_{Bench-ST}$ and $z_{Bench-LT}$, called z_{Shift}, is a measure of potential versus actual performance:

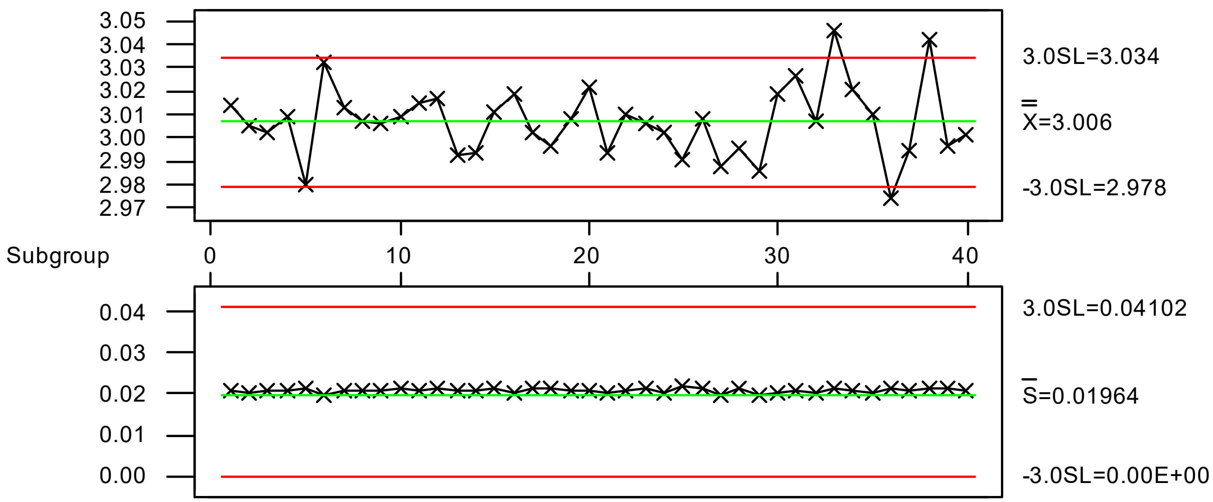
$$z_{Shift} = z_{Bench-ST} - z_{Bench-LT}$$

  When long term data aren't available, z_{Shift} is assumed to be $z_{Shift} = 1.5$ - the size of the shift that common SPC charts can detect.

Process Capability Analysis with MINITAB

Report 2: Process Capability for Msmt

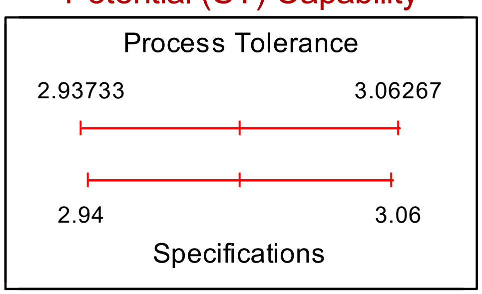
Xbar and S Chart



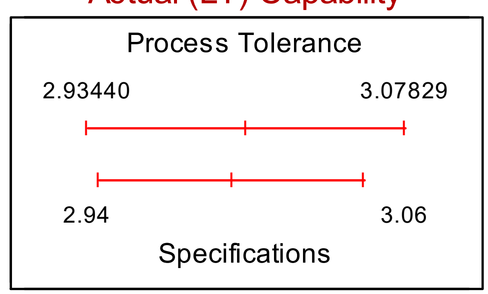
Capability Indices

	ST	LT
Mean	3.00000	3.00634
StDev	0.02086	0.02395
Z.USL	2.87649	2.24015
Z.LSL	2.87649	2.76994
Z.Bench	2.65028	2.16110
Z.Shift	0.48918	0.48918
P.USL	0.002011	0.012541
P.LSL	0.002011	0.002803
P.Total	0.004021	0.015344
Yield	99.5979	98.4656
PPM	4021.31	15343.8
Cp	0.96	
Cpk	0.86	
Pp		0.83
Ppk		0.75

Potential (ST) Capability



Actual (LT) Capability

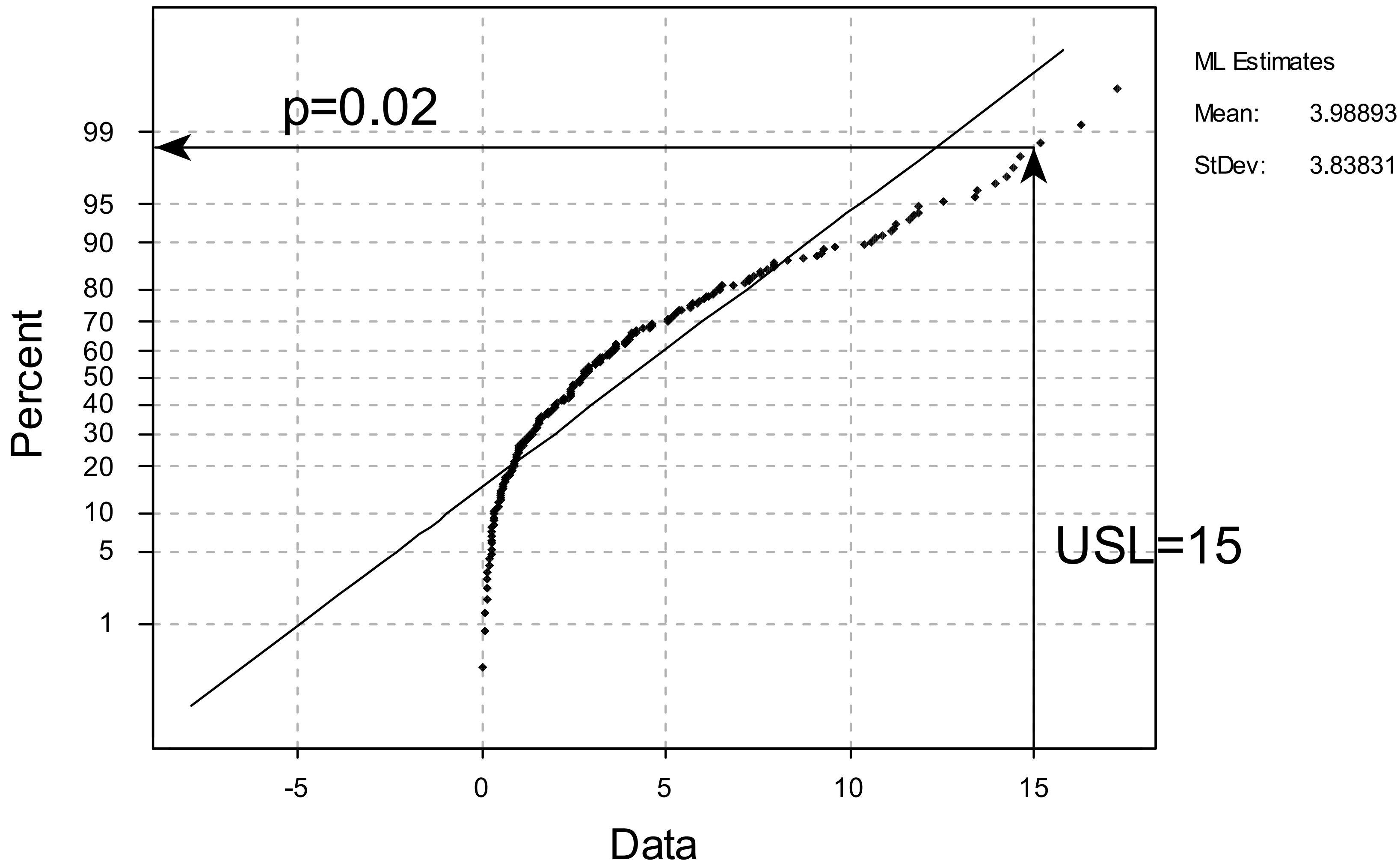


Data Source:
Time Span:
Data Trace:

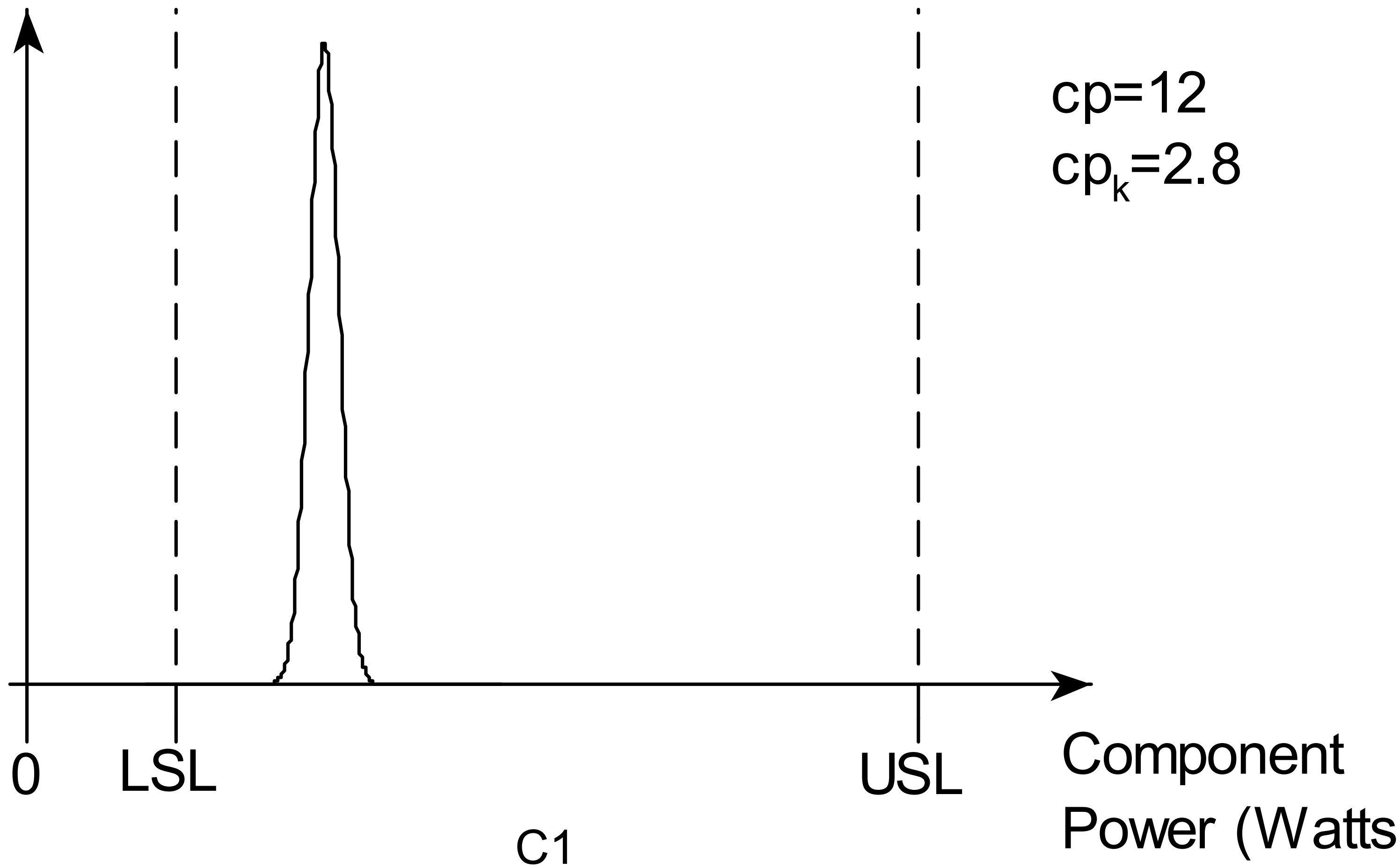
Process Capability for Nonnormal Data

- When the data are normal a pair of c_p and c_{pk} values can be translated into product fraction defective.
- When data are not normal:
  - c_p and c_{pk} are misleading.
  - Attempt to transform (Box-Cox or Johnson) them back to normality and use the usual analysis for normal data.
  - If the data can't be transformed to normality, specify the process capability as product fraction defective.

Problem: Find the capability of the process who's data is plotted below. A single sided spec of USL=15 applies.



A Process Capability Story



Strategies for Setting Specifications

- Quality characteristics that directly affect responses that are critical to quality must be set so that the responses meet the customer's requirements.
- Quality characteristics that are not critical to quality should be set based on the observed range of variation for which the product is acceptable. If only preliminary data are available and they span a narrow range then the spec might be tight. If, over time, it is discovered that the quality characteristic is acceptable over a wider range of values then the specification limits should be opened up accordingly.
- By focusing attention on the key quality characteristics that determine product quality other characteristics become less important and their tolerances can be relaxed.

Procedure for Process Capability Studies

1. Collect and plot data on an $\bar{x}$ and R chart using a sample size of about $n = 5$.
2. Collect enough subgroups that there are at least 200 and preferably 300 observations in total, e.g. $m = 60$ subgroups of size $n = 5$.
3. Use a normal probability plot of the individual measurements to determine if the population is normal.
4. Inspect the control chart to determine if the process is in control.
5. If the population is normal and the process is in control use $\bar{\bar{x}}$ and $\bar{R}$ to calculate $\hat{c}_p$ and $\hat{c}_{pk}$.
6. If the population is normal but the process is out of control use $\bar{\bar{x}}$ and $\hat{\sigma}_{total}$ to calculate $\hat{P}_p$ and $\hat{P}_{pk}$.
7. If the population is not normal use one of the methods discussed earlier or get help from your BB or MBB.

MINITAB Commands

- Stat> Control Charts> Variables Charts for Subgroups
- Stat> Control Charts> Variables Charts for Individuals
- Stat> Control Charts> Time-Weighted Charts
- Stat> Quality Tools> Capability Analysis
- Stat> Quality Tools> Capability SixPack

Confidence Intervals and Hypothesis Tests

- One-sample Z
- One-sample T
- Paired-sample T
- Two-sample T
- One variance
- Two variances
- One Porportion
- Two Proportion
- Chi-square Test for Association
- Chi-square Goodness of Fit Test

One Sample Z Test

One Sample Z Test

- The one-sample *z* test's scope includes:
  - Distribution characteristic: The mean of a normal population
  - Context/Experiment: A single population

- Methods:
  - Hypothesis Test
  - Confidence Interval

Limits on a Population

Example: A population (x) has $\mu_x = 320$, $\sigma_x = 20$, and is normally distributed. Find a symmetric interval on x that contains 95% of the population.

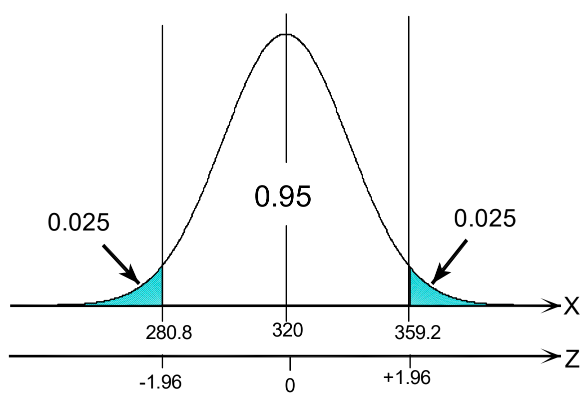
Solution: The required interval is given by:

$$\Phi(\mu_x - z_{\alpha/2}\sigma_x < x < \mu_x + z_{\alpha/2}\sigma_x) = 1 - \alpha$$

Since $1 - \alpha = 0.95$ we have $\alpha = 0.05$ and $z_{\alpha/2} = z_{0.025} = 1.96$. The required interval becomes:

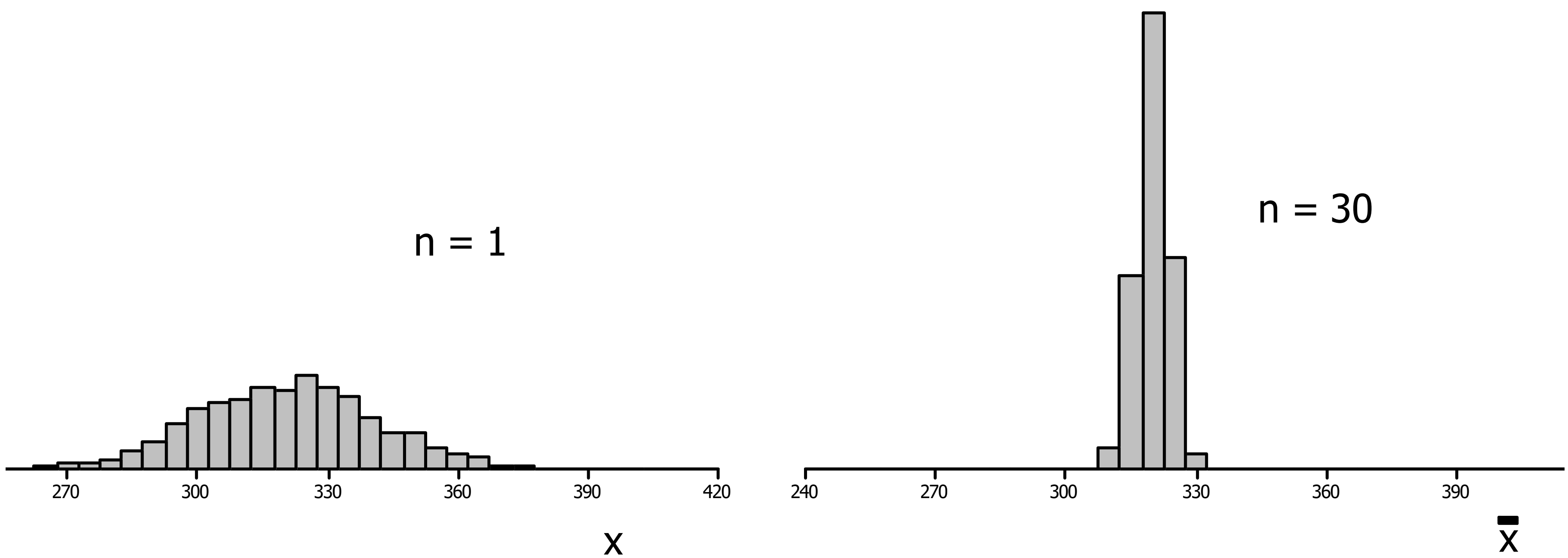
$$\Phi(320 - 1.96(20) < x < 320 + 1.96(20)) = 1 - 0.05$$

$$\Phi(280.8 < x < 359.2) = 0.95$$



Gedanken Experiment

Suppose that we compare the histogram of the measurements from 1000 samples taken from a normal distribution with $\mu = 320$ and $\sigma = 20$ to the histogram of the sample means for samples of size $n = 30$ taken from the same population:



The Central Limit Theorem

The distribution of sample means $(\bar{x})$ for samples of size n is normal (Φ) with mean:

$$\mu_{\bar{x}} = \mu_x$$

and standard deviation:

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

if the following conditions are met:

1. The population standard deviation σ_x is known or the sample size is very large $(n \geq 30)$ so that σ_x can be approximated with the sample standard deviation s.
2. The distribution of the population (x) is normal.

The central limit theorem is very robust to deviations from these conditions so the scope of its applications is very broad.

Using the Central Limit Theorem

An immediate application of the Central Limit Theorem is for the calculation of an interval that contains a specified fraction of the expected sample means. Given μ_x, σ_x, n, and α the interval that contains $(1-\alpha)100\%$ of the expected sample means is:

$$\Phi(\mu_x - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} < \mu_x + z_{\alpha/2}\sigma_{\bar{x}}) = 1 - \alpha$$

where

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

Limits on Sample Means

Example: Samples of size $n = 30$ are drawn from a population that has $\mu_x = 320$ and $\sigma_x = 20$. Find a symmetric interval that contains 95% of the sample means.

Solution: Since the sample size is large the Central Limit Theorem is valid. The required interval for $\bar{x}$s is given by:

$$\Phi(\mu_x - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} < \mu_x + z_{\alpha/2}\sigma_{\bar{x}}) = 1 - \alpha$$

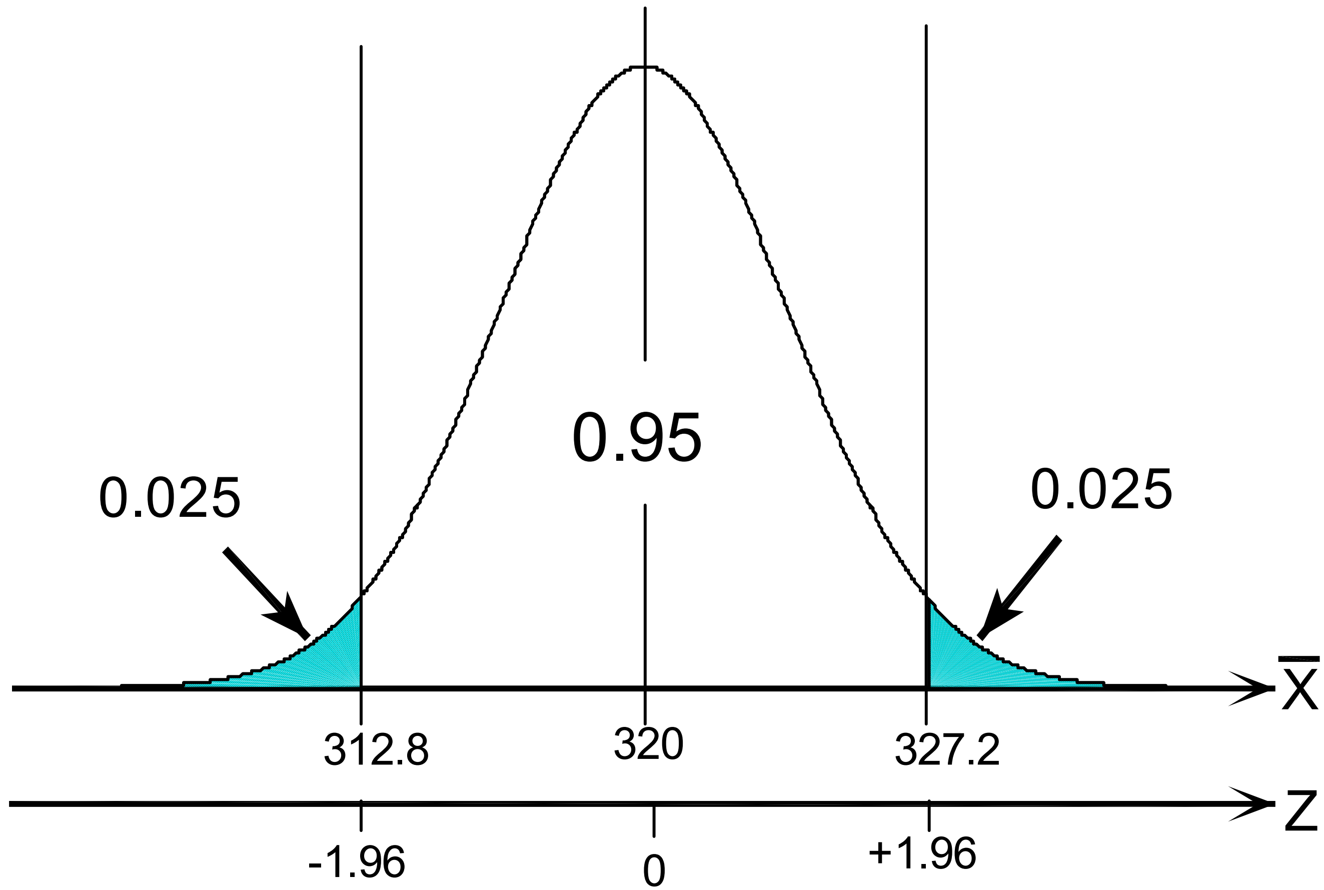
Since $1 - \alpha = 0.95$ we have $\alpha = 0.05$ and $z_{\alpha/2} = z_{0.025} = 1.96$. The standard deviation of the $\bar{x}$s is

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} = \frac{20}{\sqrt{30}} = 3.65$$

The required interval becomes:

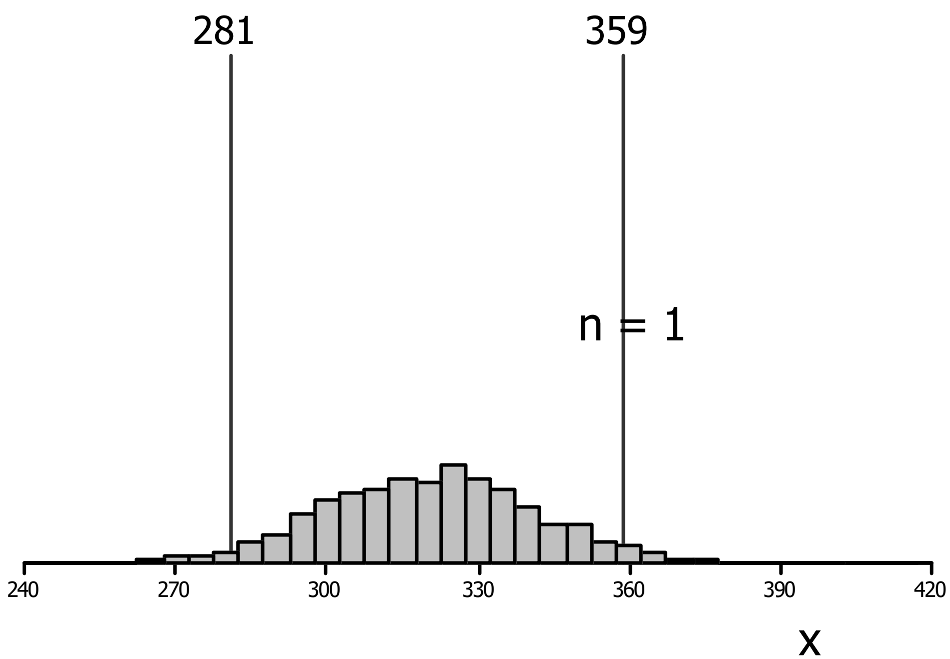
$$\Phi(320 - 1.96(3.65) < \bar{x} < 320 + 1.96(3.65)) = 1 - 0.05$$

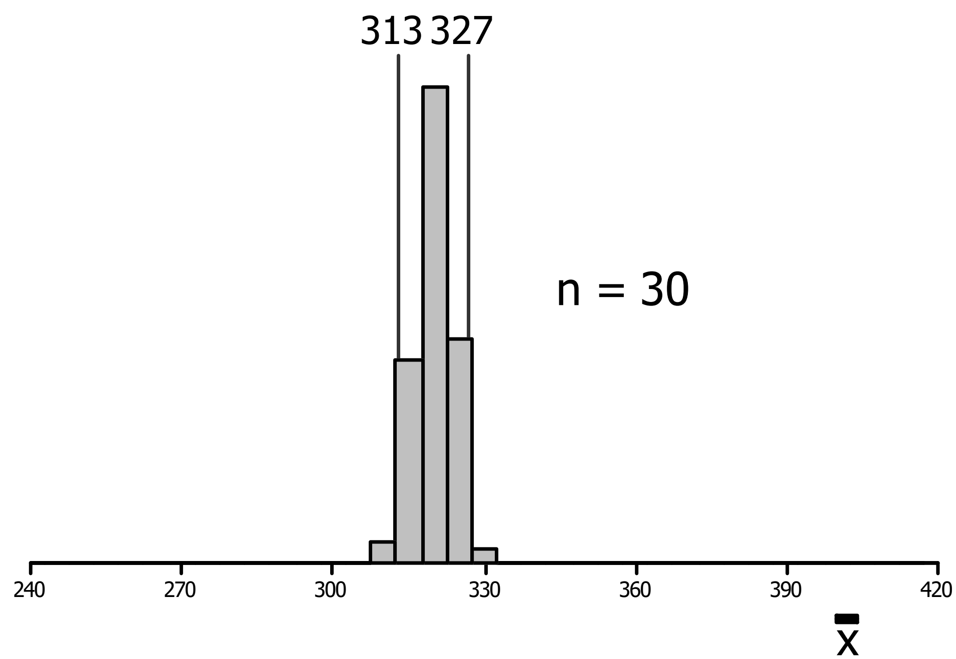
$$\Phi(312.8 < \bar{x} < 327.2) = 0.95$$



0.95
0.025
0.025
$\bar{X}$
312.8
320
327.2
Z
-1.96
0
+1.96

Comparing the Intervals





Confidence Interval for the Population Mean

The Central Limit Theorem gives us:

$$\Phi(\mu_x - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} < \mu_x + z_{\alpha/2}\sigma_{\bar{x}}) = 1 - \alpha$$

The random variable $\bar{x}$ is bounded on the lower and upper sides in two inequalities:

$$\mu_x - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} \quad \text{and} \quad \bar{x} < \mu_x + z_{\alpha/2}\sigma_{\bar{x}}$$

If we solve these inequalities for μ_x we obtain:

$$\mu_x < \bar{x} + z_{\alpha/2}\sigma_{\bar{x}} \quad \text{and} \quad \bar{x} - z_{\alpha/2}\sigma_{\bar{x}} < \mu_x$$

Now, if we put these two inequalities back together:

$$\Phi(\bar{x} - z_{\alpha/2}\sigma_{\bar{x}} < \mu_x < \bar{x} + z_{\alpha/2}\sigma_{\bar{x}}) = 1 - \alpha$$

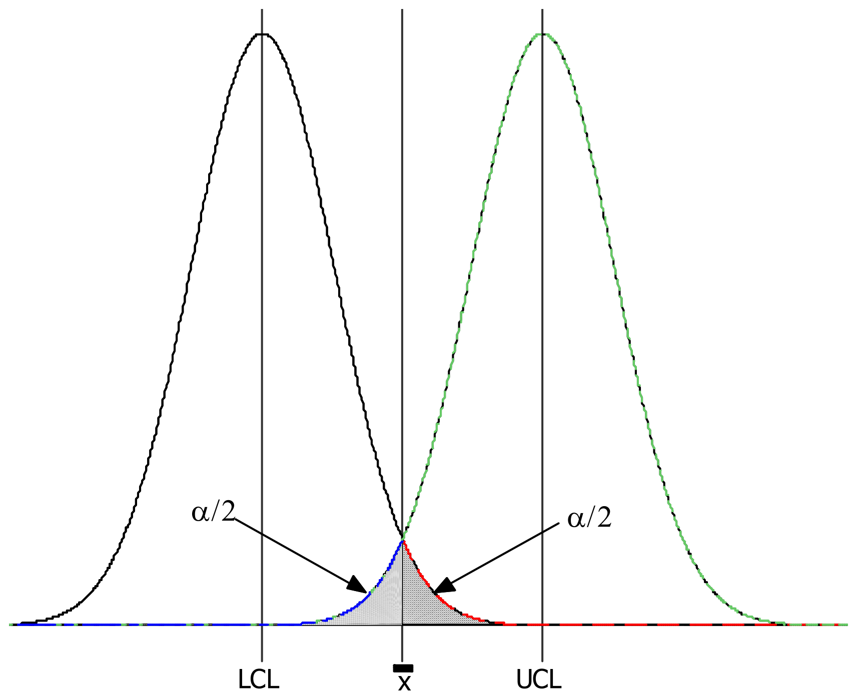
which is the two sided $(1-\alpha)100\%$ confidence interval for the unknown population mean μ_x based on a sample which has sample mean $\bar{x}$.

Graphical Interpretation

The upper and lower confidence limits given by:

$$UCL/LCL = \bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

represent the extreme high and low values of μ_x that could be expected to deliver the experimental $\bar{x}$ value.



Confidence Interval Example

Example: Construct a two-sided 95% confidence interval for the true population mean based on a sample of size $n = 30$ which yields $\bar{x} = 290$. The population standard deviation is $\sigma = 20$ and the distribution of the xs is normal.

Solution: Since the Central Limit Theorem is satisfied (distribution of x is normal and σ_x is known) the confidence interval is given by:

$$\Phi(\bar{x} - z_{\alpha/2}\sigma_{\bar{x}} < \mu_x < \bar{x} + z_{\alpha/2}\sigma_{\bar{x}}) = 1 - \alpha$$

Since $\alpha = 0.05$ we have $z_{\alpha/2} = z_{0.025} = 1.96$ so:

$$\Phi\left(290 - 1.96\left(\frac{20}{\sqrt{30}}\right) < \mu_x < 290 + 1.96\left(\frac{20}{\sqrt{30}}\right)\right) = 1 - 0.05$$

The required confidence interval is:

$$\Phi(282.8 < \mu_x < 297.2) = 0.95$$

That is, we can be 95% confident that the true but unknown value of the population mean lies between 282.8 and 297.2.

Confidence Interval Interpretation

- A two-sided confidence interval for the mean has the form

$$P(LCL < \mu < UCL) = 1 - \alpha$$

- The interval $LCL < \mu < UCL$ indicates the range of possible μ values that are statistically consistent with the observed value of $\bar{x}$.
- If the confidence interval is sufficiently narrow then the interval $LCL < \mu < UCL$ will indicate a single action. Take it.
- If the confidence interval is too wide then the interval will indicate two or more actions. More data will be required.
- Ask yourself:
  - What action would I take if $\mu = LCL$?
  - What action would I take if $\mu = UCL$?
  - If the two actions are the same then take the indicated action.
  - If the two actions are different then the confidence interval is too wide. When in doubt, take more data.

Hypothesis Tests

Definition: A hypothesis test is a statistical method for deciding which of two complementary statements about a population parameter or distribution is true on the basis of sample data. The two statements are called the null hypothesis (H_0) and the alternative hypothesis (H_A).

Hypothesis Tests

Examples (one population):

- $H_0 : \mu = 320$ versus $H_A : \mu \neq 320$ (two-tailed test)
- $H_0 : \mu = 320$ versus $H_A : \mu < 320$ (one- / left-tailed test)
- $H_0 : \mu = 320$ versus $H_A : \mu > 320$ (one- / right-tailed test)
- $H_0 : \sigma = 20$ versus $H_A : \sigma \neq 20$
- $H_0 : \sigma = 20$ versus $H_A : \sigma < 20$
- $H_0 : \sigma = 20$ versus $H_A : \sigma > 20$
- $H_0 : p = p_0$ versus $H_A : p \neq p_0$
- $H_0 : \lambda = \lambda_0$ versus $H_A : \lambda \neq \lambda_0$
- H_0 :The distribution of x is Φ versus H_A :The distribution of x is not Φ
- H_0 :The distribution of s^2 is χ^2 versus H_A :The distribution of s^2 is not χ^2

Hypothesis Tests

Examples (two populations):

- $H_0 : \mu_1 = \mu_2$ versus $H_A : \mu_1 \neq \mu_2$
- $H_0 : \sigma_1 = \sigma_2$ versus $H_A : \sigma_1 \neq \sigma_2$
- $H_0 : p_1 = p_2$ versus $H_A : p_1 \neq p_2$
- $H_0 : \lambda_1 = \lambda_2$ versus $H_A : \lambda_1 \neq \lambda_2$
- H_0 : *The distribution shape of* x_1 *is the same as the distribution shape of* x_2 versus H_A : *The distribution shape of* x_1 *is NOT the same as the distribution shape of* x_2.

- **Examples (many populations)**:
- $H_0 : \mu_1 = \mu_2 = \cdots$ versus $H_A : \mu_i \neq \mu_j$ for at least one i,j pair
- $H_0 : \sigma_1 = \sigma_2 = \cdots$ versus $H_A : \sigma_i \neq \sigma_j$ for at least one i,j pair
- $H_0 : p_1 = p_2 = \cdots$ versus $H_A : p_i \neq p_j$ for at least one i,j pair
- $H_0 : \lambda_1 = \lambda_2 = \cdots$ versus $H_A : \lambda_i \neq \lambda_j$ for at least one i,j pair

Which Test?

- What type of data?
  - Measurement/variable
  - Attribute
    - Binary/dichotomous, e.g. defectives
    - Count, e.g. defects
- How many populations?
- What population characteristic?
  - Location
  - Variation
  - Distribution Shape
  - Other
- Exact or approximate method?
- See Appendix B: Hypothesis Test Matrix

Understanding Hypotheses

- Statistical hypotheses have two forms, one stated mathematically and the other stated in the language of the context. For example, in SPC the hypothesis $H_0 : \mu = 25$ corresponds to the statement *the process is in control*.
- Sagan's Rule: *To test the hypotheses*

$$H_o\text{: Something ordinary happens}$$

versus

$$H_A\text{: Something extraordinary happens,}$$

the extraordinary claim requires extraordinary evidence.

- In quality engineering, sometimes the hypotheses are determined by historical choice:
  - SPC: H_0: *the process is in control* versus H_A: *the process is out of control*.
  - Acceptance sampling: H_0: *the lot is good* versus H_A: *the lot is bad*.

General Hypothesis Testing Procedure

1. Formulate the null (H_0) and alternative hypotheses (H_A). Put the desired conclusion in H_A.
2. Specify the significance level α (the risk of a Type 1 error).
3. Construct accept and reject criteria for the hypotheses based on the sampling distribution of an appropriate test statistic at the required significance level.
4. Collect the data and calculate the value of the test statistic.
5. Compare the test statistic to the acceptance interval and decide whether to accept or reject H_0. In practice, we never accept H_0. We either reject H_0 and accept H_A or we say that the test is inconclusive.

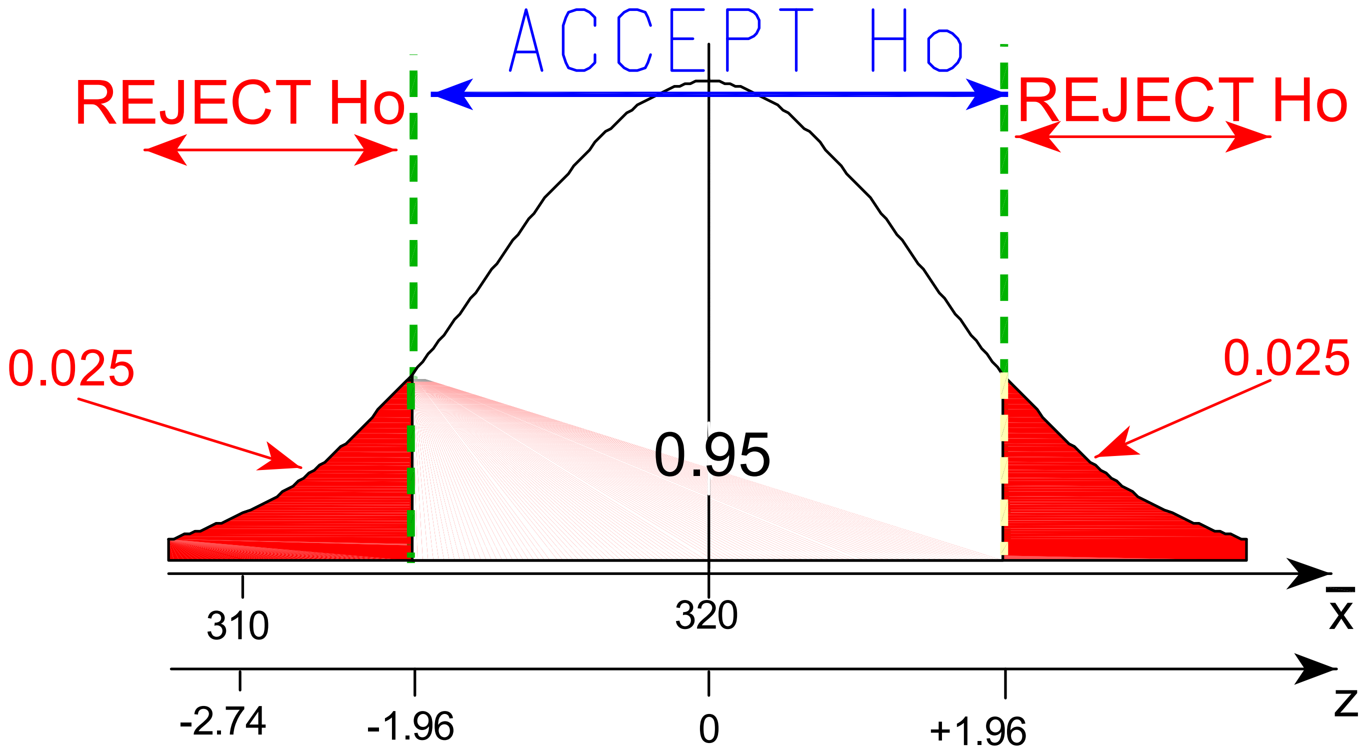
Hypothesis Test Example

Example A: Test the hypotheses $H_0 : \mu = 320$ vs. $H_A : \mu \neq 320$ on the basis of a sample of size $n = 30$ taken from a normal population with standard deviation $\sigma = 20$ which yields $\bar{x} = 310$. Use the 5% significance level.

Solution: The two hypotheses are already given to us. The appropriate statistic to test them is $\bar{x}$. If $\bar{x}$ falls very close to 320 then we will accept H_0, otherwise we will reject it. The Central Limit Theorem describes the distribution of the $\bar{x}$s and with $\alpha = 0.05$ we have a critical z value of $z_{0.025} = 1.96$. This means that we will accept H_0 if the test statistic falls in the interval $-1.96 < z < +1.96$. The z value that corresponds to $\bar{x}$ is given by:

$$
\begin{aligned}
z &= \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} \\
&= \frac{310 - 320}{20/\sqrt{30}} \\
&= -2.74
\end{aligned}
$$

Since $z = -2.74$ falls outside the acceptance interval $\bar{x}$ must be significantly different from the hypothesized mean of $H_0 : \mu = 320$ so we must reject H_0 in favor of $H_A : \mu \neq 320$.



Relationship Between Confidence Intervals and Hypothesis Tests

- The confidence interval and hypothesis test provide different ways of performing the same analysis but they both offer unique features that prohibit the exclusive use of one method or the other.
- The confidence interval for the mean is centered on the sample mean:

$$UCL/LCL = \bar{x} \pm \delta$$

where the confidence interval half-width is

$$\delta = z_{\alpha/2}\sigma_{\bar{x}}$$

- The accept/reject decision limits for the hypothesis test are centered on μ_0:

$$UDL/LDL = \mu_0 \pm \delta$$

where δ has the same value as the confidence interval half-width.
- The confidence interval is the set of all possible values of μ_0 for which we would accept H_0, so …
- If μ_0 falls inside of the confidence limits then we accept $H_0 : \mu = \mu_0$

and if μ_0 falls outside of the confidence limits then we reject H_0.

Example: Construct the confidence interval for the population mean in Example A and use it to test the hypotheses $H_0 : \mu = 320$ vs. $H_A : \mu \neq 320$.

Solution: The confidence interval is

$$\Phi\left(310 - 1.96\left(\frac{20}{\sqrt{30}}\right) < \mu_x < 310 + 1.96\left(\frac{20}{\sqrt{30}}\right)\right) = 0.95$$

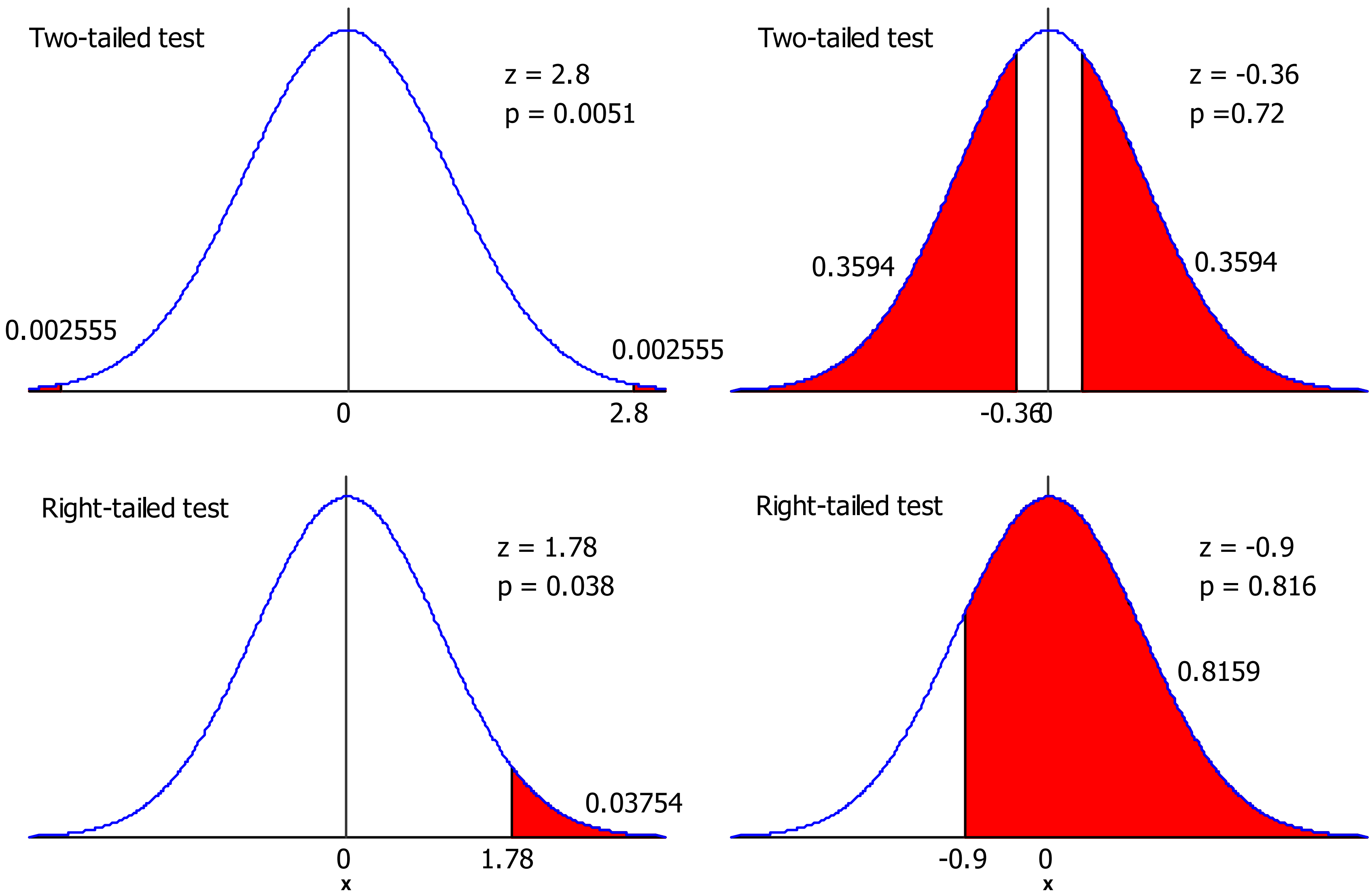
$$\Phi(302.8 < \mu_x < 317.2) = 0.95$$

The confidence interval does NOT contain $\mu = 320$ so we must reject $H_0 : \mu = 320$ in favor of $H_A : \mu \neq 320$.

p Values

- p values provide a concise and universal way of communicating statistical significance.
- The p value of a hypothesis test is the probability of obtaining the observed experimental result or something more extreme if the null hypothesis was true.
- p values are compared directly to α (typically $\alpha = 0.05$ or $\alpha = 0.01$) to make decisions about accepting or rejecting the null hypothesis.
  - If $p \geq \alpha$ accept H_0, that is, the data support the null hypothesis.
  - If $p < \alpha$ reject H_0, that is, the data don't support the null hypothesis.
- For two tailed hypothesis tests, the p value corresponds to the area in the two tails of the sampling distribution of the test statistic outside of the value obtained for the test statistic.
- For one tailed hypothesis tests, the p value corresponds to the area in one tail of the sampling distribution of the test statistic outside of the value obtained for the test statistic.

p Values

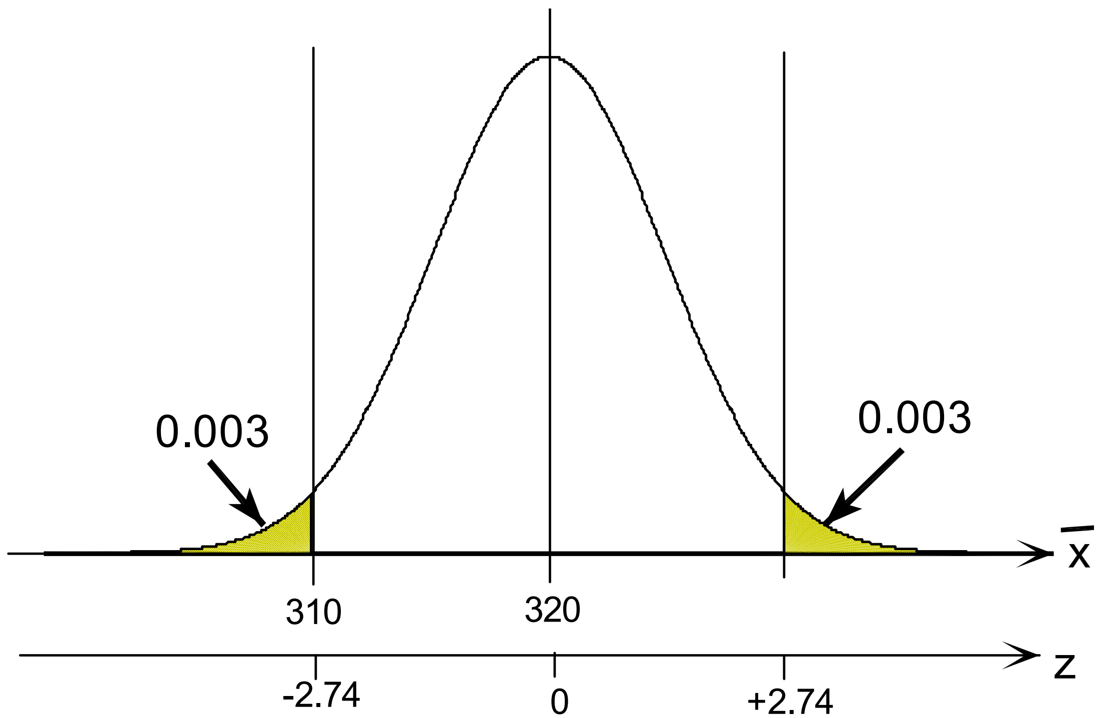


Two-tailed test
z = 2.8
p = 0.0051
0.002555
0.002555
0
2.8
Two-tailed test
z = -0.36
p =0.72
0.3594
0.3594
-0.36
0
Right-tailed test
z = 1.78
p = 0.038
0.03754
0
1.78
x
Right-tailed test
z = -0.9
p = 0.816
0.8159
-0.9
0
x

p Values

Example: Find the p value for Example A.

Solution: Since $z_{0.003} = 2.74$ the p value for this Example is $p = 2(0.003) = 0.006$. Because $(p = 0.006) < (\alpha = 0.05)$ we must reject the claim $H_0 : \mu = 320$.



The One-sample Z Test in MINITAB

- Use **Stat**> **Basic Statistics**> **1-Sample Z**
- The data can be observations in a column or summarized data ($\bar{x}$, s, n)
- By default MINITAB will report a two-sided 95% confidence interval
- Use the **Options** submenu to choose a one-sided test or to change the confidence level
- MINITAB will perform a hypothesis test if you specify a null hypothesis μ_0 value in $H_0 : \mu = \mu_0$

The One-sample Z Test in MINITAB

Example: Use MINITAB to confirm the hypothesis test and construct the confidence interval results for the case in Example A. The hypotheses were $H_0 : \mu = 320$ vs. $H_A : \mu \neq 320$ and the summarized data were $\bar{x} = 310$, $s = 20$, and $n = 30$.

Solution:

One-Sample Z

WORKSHEET 1

One-Sample Z

Descriptive Statistics

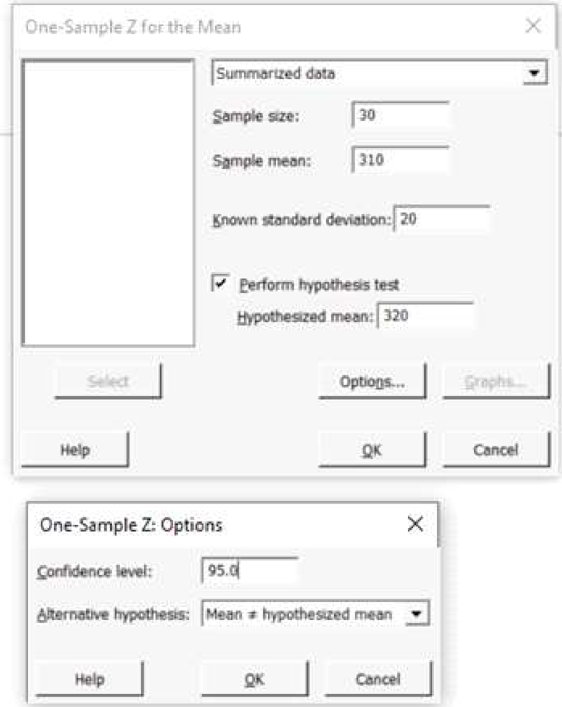
N	Mean	SE Mean	95% CI for μ
30	310.00	3.65	(302.84, 317.16)

μ: population mean of Sample
Known standard deviation = 20

Test

Null hypothesis	H_0: $\mu = 320$
Alternative hypothesis	H_1: $\mu \neq 320$

Z-Value	P-Value
-2.74	0.006

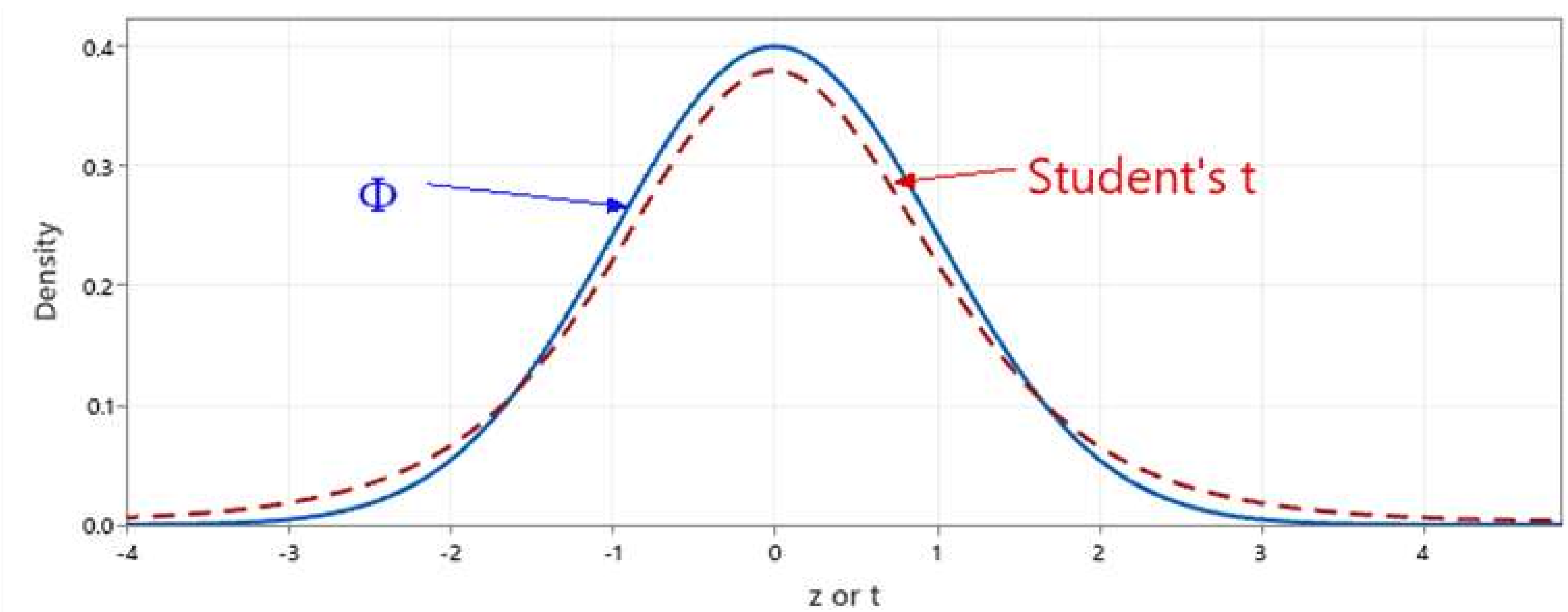


One-Sample T Test

One Sample t Test

Example B: Test the hypothesis $H_0 : \mu = 440$ vs. $H_A : \mu \neq 440$ if a sample of size $n = 10$ yields $\bar{x} = 442$ and $s = 5.1$. Assume that the distribution of x is normal and work at a 5% significance level.

Solution: This is a hypothesis test for one population mean but the central limit theorem doesn't apply because we don't know σ. So ... we have to use Student's t distribution instead.



Solution:

- If σ was known we would calculate the one-sample z statistic

$$z = \frac{\bar{x} - \mu_0}{\sigma_x / \sqrt{n}}$$

and compare it to the acceptance interval

$$-z_{0.025} \leq z \leq z_{0.025}$$

$$-1.96 \leq z \leq 1.96$$

- However, we don't know the true population standard deviation so we must estimate σ_x with s and use Student's t distribution to characterize the distribution of sample means.
- The value of the t statistic is:

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

$$= \frac{442 - 440}{5.1 / \sqrt{10}}$$

$$= 1.24$$

Solution:

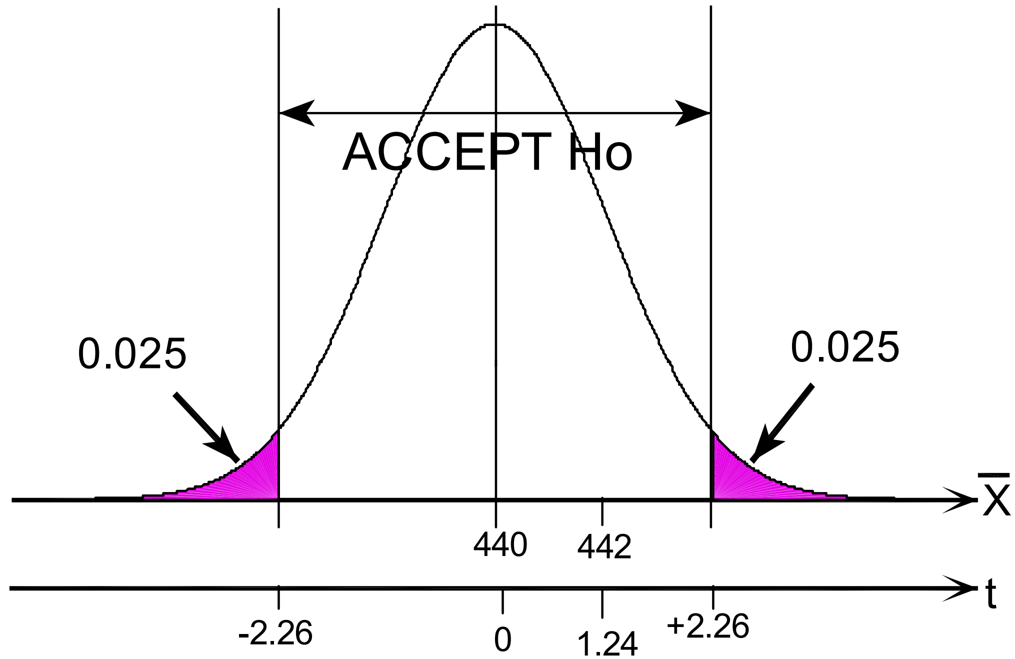
- The acceptance interval for $H_0 : \mu = 440$ is

$$- t_{0.025,9} \leq t \leq t_{0.025,9}$$

$$- 2.26 \leq t \leq 2.26$$

where Student's t distribution has $n - 1 = 9$ degrees of freedom.

- The test statistic $t = 1.24$ falls inside of the acceptance interval $-2.26 \leq t \leq 2.26$ so we must accept the null hypothesis $H_0 : \mu = 440$ or reserve judgement.


ACCEPT Ho
0.025
0.025
$\overline{X}$
440
442
t
-2.26
0
1.24
+2.26

Using Student t Tables

- Remember that $t \to z$ as $n \to \infty$ so always expect $t \simeq z$ unless n is very small
- When $n > 30$ it's common to assume $t \simeq z$
- For more discriminating cases:
  - Tables in the backs of textbooks given critical values for $t_{0.01}$, $t_{0.025}$, $t_{0.05}$, $t_{0.10}$ as a function of the degrees of freedom for determining s
  - MINITAB:
    - **Calc**> **Probability Distributions**> **t**
    - **Graph**> **Probability Distribution Plot**> **View Probability**> **t**

Example: Find the p value for Example B.

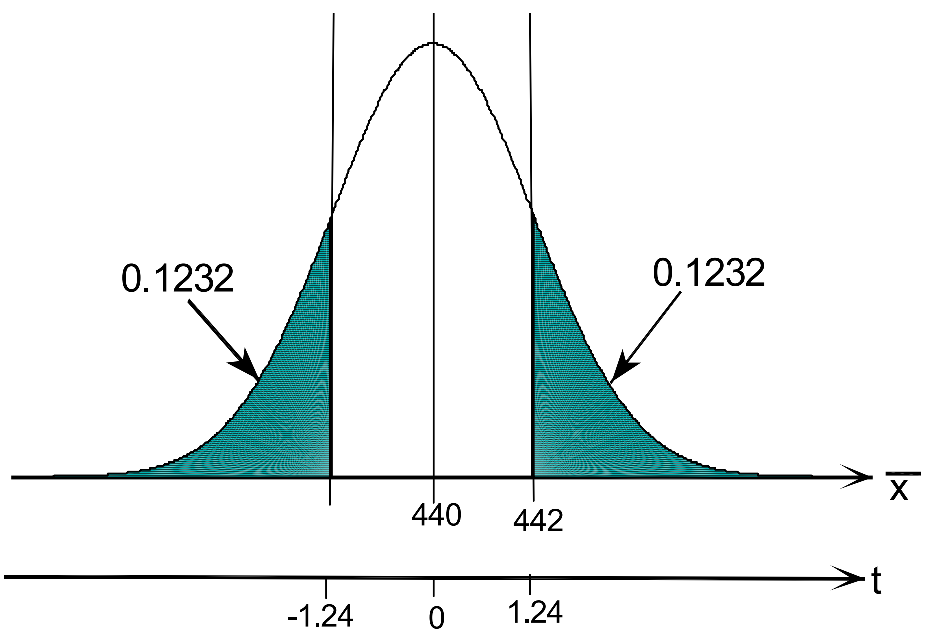
Solution: The p value is given by:

$$1 - p = P(-1.24 \le t \le 1.24)$$

where the Student's t distribution has $n - 1 = 9$ degrees of freedom. Generally it would be necessary to interpolate in a t table to estimate the true p value but MINITAB gives the exact p value:

$$p = 2(0.1232) = 0.246$$

Since $(p = 0.246) > (\alpha = 0.05)$ we must accept $H_0 : \mu = 440$.



We Never Accept H_0

- We are being lazy saying "Accept H_0" and there can be serious consequences.
- In hypothesis testing we **never** accept H_0. Hypothesis testing is all about H_A.
- Remember Sagan: *The extraordinary claim* (H_A) *requires extraordinary evidence.*
- Valid hypothesis test decisions are to either reject H_0 in favor of H_A or reserve judgement.
- Reserve judgement is easier to parse than saying "We can't reject H_0." The double negative is difficult to parse.
- So every time you see "Accept H_0" replace it with "reserve judgement."

Confidence Interval

- Use the Student t distribution when:
  - When the population standard deviation σ is unknown and must be estimated with the sample standard deviation s
  - The distribution of x is normal
  - The t distribution is robust to deviations from normality so only a weak normality requirement is necessary, especially when the sample size is large
- The confidence interval for the population mean based on a sample of size n taken from a normal population which yields $\bar{x}$ and s is given by:

$$P(\bar{x} - t_{\alpha/2}s/\sqrt{n} < \mu < \bar{x} + t_{\alpha/2}s/\sqrt{n}) = 1 - \alpha$$

  where $t_{\alpha/2}$ comes from Student's t distribution with $\nu = n - 1$ degrees of freedom.

Confidence Interval

Example: Construct the 95% confidence interval for the true population mean for the situation in Example B.

Solution: The confidence interval for μ is:

$$P(\bar{x} - t_{\alpha/2}s/\sqrt{n} < \mu < \bar{x} + t_{\alpha/2}s/\sqrt{n}) = 1 - \alpha$$

$$P\left(442 - 2.26 \times 5.1/\sqrt{10} < \mu < 442 + 2.26 \times 5.1/\sqrt{10}\right) = 0.95$$

$$P(438.4 < \mu < 445.6) = 0.95$$

That is, we can be 95% confident that the true population mean falls in the interval from 438.4 to 445.6.

This confidence interval demonstrates the relationship between confidence intervals and hypothesis tests: *a confidence interval for the mean is the set of population means for which the null hypothesis must be accepted*, so because the example's confidence interval contains $\mu = 440$ we know that we have to accept the null hypothesis $H_0 : \mu = 440$.

The One-sample t Test in MINITAB

- Use **Stat**> **Basic Statistics**> **1-Sample t**
- The data can be observations in a column or summarized data ($\bar{x}$, s, n)
- By default MINITAB will report a two-sided 95% confidence interval
- Use the **Options** submenu to choose a one-sided test or to change the confidence level
- MINITAB will perform a hypothesis test if you specify a null hypothesis μ_0 value in $H_0 : \mu = \mu_0$

The One-sample t Test in MINITAB

Example: Use MINITAB to confirm the hypothesis test and construct the confidence interval results for the case in Example B. The hypotheses were $H_0 : \mu = 440$ vs. $H_A : \mu \neq 440$ and the summarized data were $\bar{x} = 442$, $s = 5.1$, and $n = 10$.

Solution:

One-Sample T

WORKSHEET 1

One-Sample T

Descriptive Statistics

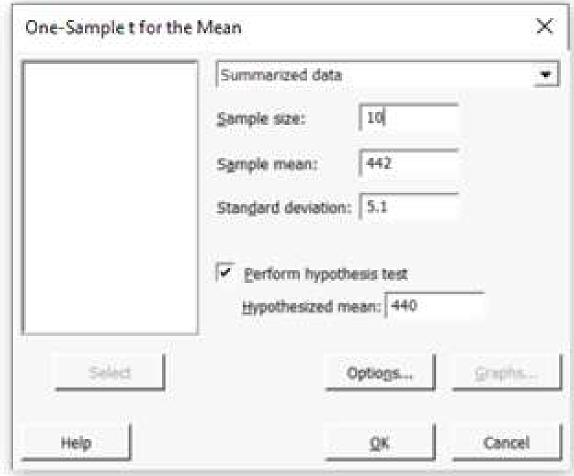
N	Mean	StDev	SE Mean	95% CI for μ
10	442.00	5.10	1.61	(438.35, 445.65)

μ: population mean of Sample

Test

Null hypothesis	H_0: $\mu = 440$
Alternative hypothesis	H_1: $\mu \neq 440$

T-Value	P-Value
1.24	0.246



Paired Sample t

Paired Sample t Test

Data: n paired samples (x_{1i}, x_{2i}) of measurement data taken from normal populations. The data pairs are ”before and after” type.

Test Statistic: The quantities of interest are the signed differences between the paired observations:

$$\Delta x_i = x_{1i} - x_{2i}$$

Note that this makes the paired-sample t test a special case of the one-sample t test when the one-sample t test is applied to the Δx_i values.

The sample mean and standard deviation of the Δx_i are required:

$$\overline{\Delta x} = \frac{1}{n}\sum_{i=1}^{n} \Delta x_i$$

and

$$s = \sqrt{\frac{\sum (\Delta x_i - \overline{\Delta x})^2}{n-1}}$$

The test statistic for $H_0 : \Delta\mu = 0$ is:

$$t = \frac{\overline{\Delta x} - 0}{s/\sqrt{n}} = \frac{\overline{\Delta x}}{s/\sqrt{n}}$$

Hypotheses Tested:

- $H_0 : \Delta\mu = 0$ vs. $H_A : \Delta\mu \neq 0$
- $H_0 : \Delta\mu = 0$ vs. $H_A : \Delta\mu < 0$
- $H_0 : \Delta\mu = 0$ vs. $H_A : \Delta\mu > 0$

Critical Values:

- For $H_0 : \Delta\mu = 0$ vs. $H_A : \Delta\mu \neq 0$ accept H_0 iff $-t_{\alpha/2,n-1} < t < t_{\alpha/2,n-1}$
- For $H_0 : \Delta\mu = 0$ vs. $H_A : \Delta\mu < 0$ accept H_0 iff $t > -t_{\alpha,n-1}$
- For $H_0 : \Delta\mu = 0$ vs. $H_A : \Delta\mu > 0$ accept H_0 iff $t < t_{\alpha,n-1}$

Paired Sample t Test

Example: The following table shows measurements taken by two operators on the same 10 parts. Determine if there is evidence of a bias between the operators at the 5% significance level.

Part Number	1	2	3	4	5	6	7	8	9	10
Operator 1	2.4	2.8	3.1	2.7	3.0	2.5	2.2	4.3	3.8	3.4
Operator 2	2.6	2.9	3.4	2.7	2.9	2.7	2.3	4.4	4.1	3.4

Solution: The differences between the paired readings are shown below:

Part Number	1	2	3	4	5	6	7	8	9	10
Operator 1	2.4	2.8	3.1	2.7	3.0	2.5	2.2	4.3	3.8	3.4
Operator 2	2.6	2.9	3.4	2.7	2.9	2.7	2.3	4.4	4.1	3.4
Δx_i	-0.2	-0.1	-0.3	0.0	0.1	-0.2	-0.1	-0.1	-0.3	0.0

Solution:

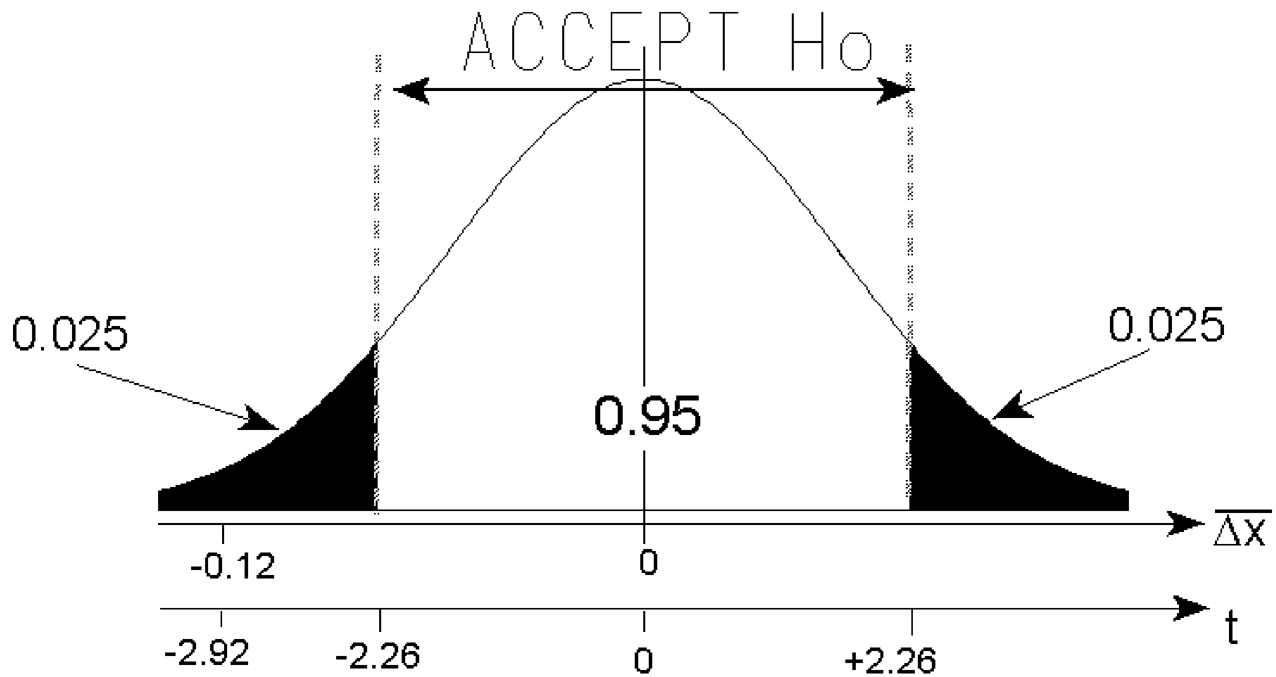
- The mean and standard deviation of the Δx_i are $\overline{\Delta x} = -0.12$ and $s = 0.13$, respectively.
- The test statistic for the hypotheses $H_0 : \Delta\mu = 0$ versus $\Delta H_A : \mu \neq 0$ is

$$t = \frac{-0.12 - 0}{0.13/\sqrt{10}}$$
$$= -2.92$$

- There are 10 paired observations so there are $df = 9$ degrees of freedom to estimate σ from s. Then the critical t value for $\alpha = 0.05$ is $t_{0.025,9} = 2.26$ and the acceptance interval for H_0 is $-2.26 < t < 2.26$.
- Because $t = -2.92$ falls outside this interval $-2.26 < t < 2.26$ we must reject H_0 and conclude that there is a statistically significant bias between the two operators.
- The p value for the test is

$$p = 2 \times P(-\infty < t < -2.92)$$
$$= 0.017$$

Solution:



MINITAB

There are two ways to do the paired-sample t test in MINITAB:

- By calculating the Δx_i values explicitly and applying the one-sample t test (**Stat**> **Basic Statistics**> **1-Sample t**) to them.
- By letting MINITAB calculate the Δx_i values from the x_{1i} and x_{2i} input columns using the **Stat**> **Basic Statistics**> **Paired t** menu.

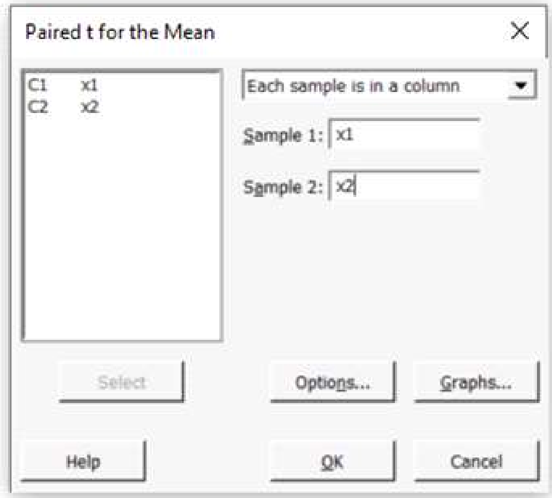
Example: Confirm the results from the paired-sample t test analysis using MINITAB.

Solution: The MINITAB output from the **Stat**> **Basic Statistics**> **Paired t** menu is shown on the next page. The hypothesis test t statistics and its p value are a bit different from the values obtained by manual calculation. The difference is due to round off errors in the manual calculations and do not affect the interpretation of the results or the conclusions.

Paired T-Test and CI: x1, x2

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
x1	10	3.020	0.656	0.208
x2	10	3.140	0.679	0.215



Estimation for Paired Difference

Mean	StDev	SE Mean	95% CI for μ_difference
-0.1200	0.1317	0.0416	(-0.2142, -0.0258)

μ_difference: population mean of (x1 - x2)

Test

Null hypothesis H_0: μ_difference = 0
Alternative hypothesis H_1: μ_difference ≠ 0

T-Value	P-Value
-2.88	0.018

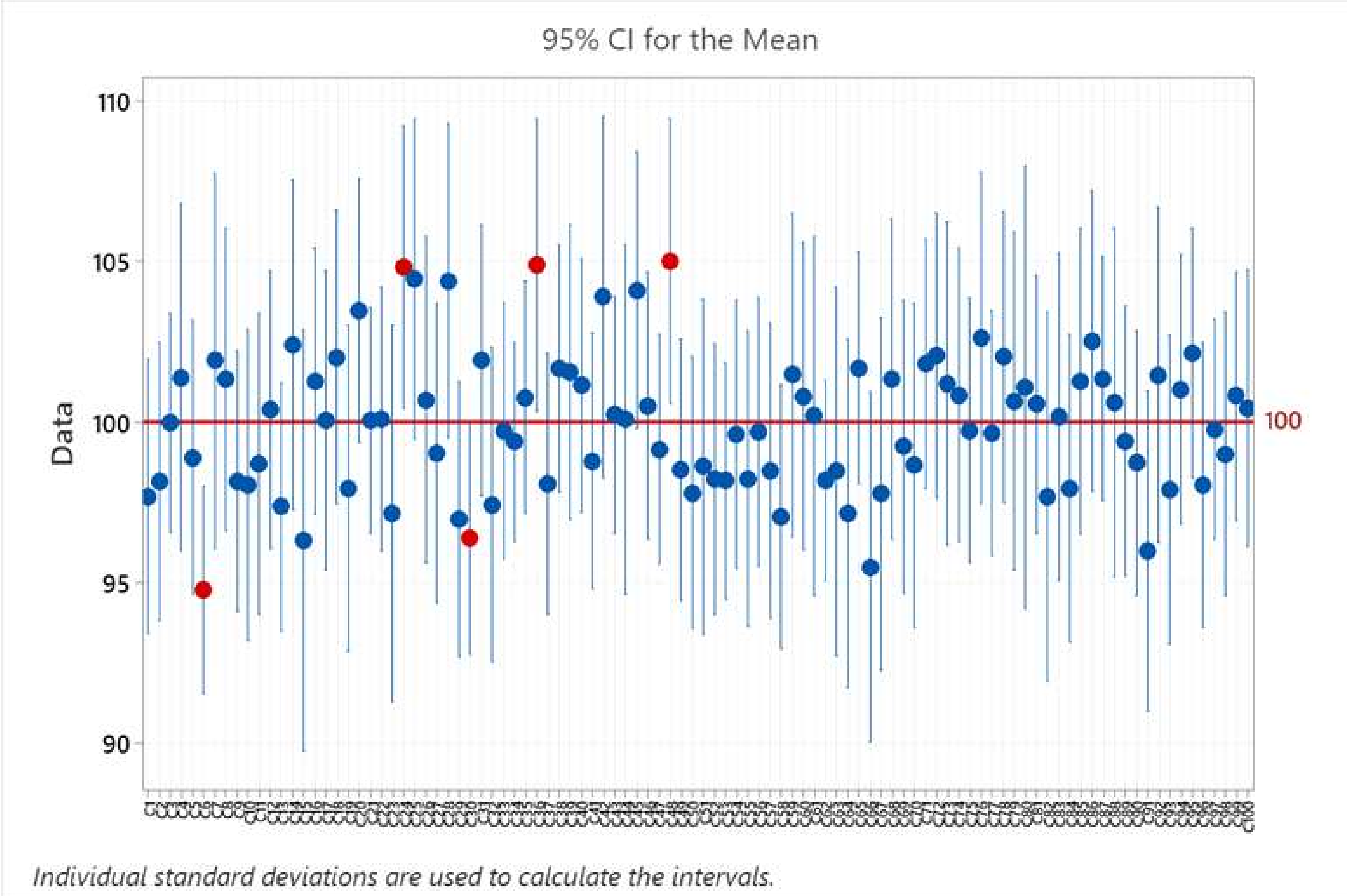
↓	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
	x1	x2									
1	2.4	2.6									
2	2.8	2.9									
3	3.1	3.4									
4	2.7	2.7									
5	3.0	2.9									
6	2.5	2.7									
7	2.2	2.3									
8	4.3	4.4									
9	3.8	4.1									
10	3.4	3.4									

Errors in Hypothesis Tests and Confidence Intervals

Errors in Confidence Intervals

- A confidence interval for the population mean has the form
$$P(LCL < \mu < UCL) = 1 - \alpha$$
where $1 - \alpha$ is the confidence level.
- The confidence level $1 - \alpha$ gives the probability that the confidence interval contains the true value of the population mean.
- The α value is the probability that the confidence interval does NOT contain the true value of the population mean.
- So given many 95% confidence intervals, 95% of them will contain the true population mean and 5% won't.
- Keep in mind that any one 95% confidence interval has a 5% chance of not containing the true population mean.

Errors in Confidence Intervals



Errors in Hypothesis Testing

There are two kinds of errors that can occur in hypothesis testing:

1. Type 1 Error: Reject the null hypothesis when it is really true.
2. Type 2 Error: Accept the null hypothesis when it is really false.

These potential error situations are summarized in the following table:

The truth is:→	H_0 is true	H_0 is false
The test says accept H_0	Correct Decision	Type 2 Error
The test says reject H_0	Type 1 Error	Correct Decision

Errors in the Legal System

- Hypotheses:
  - *H_0: The defendant is innocent*
  - *H_A: The defendant is guilty*
- Was the correct decision made and, if not, what type of error occurred?

Truth	Verdict	Error?
Innocent	Not Guilty	
Innocent	Guilty	
Guilty	Not Guilty	
Guilty	Guilty	

Type 1 and Type 2 Errors in Inspection

A final inspection operation is used to check each lot before it is shipped to the customer. The hypotheses tested are H_0: *The lot is good* versus H_A: *The lot is bad.* Lots determined to be bad are scrapped or 100% inspected (i.e. rectifying inspection).

- If a good lot (the truth) is tested and the test confirms that the lot is good:
  - The correct decision has been made
  - The good lot gets shipped to the customer.
- If a good lot (the truth) is tested and the test indicates that the lot is bad:
  - Then a Type 1 error has occurred
  - The good lot gets scrapped or 100% inspected.
  - The manufacturer scraps this good material (lost value for him) so the loss is called *manufacturer's risk*.
  - It may help to think about a Type 1 error as a false alarm or rejecting a good lot.

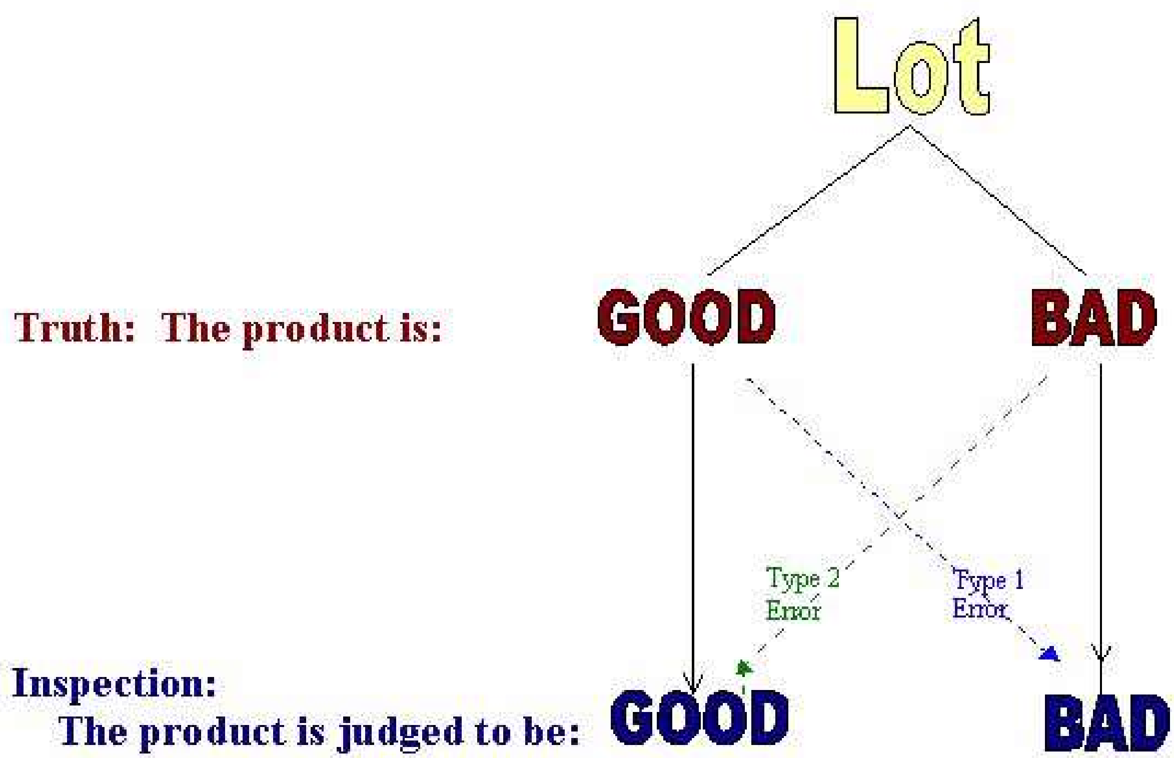
Type 1 and Type 2 Errors in Inspection

- If a bad lot (the truth) is tested and the test indicates that the lot is good:
  - Then a Type 2 error has occurred
  - The bad lot gets sent to the customer (lost value for him) so the loss is called *consumer's risk*.
- If a bad lot (the truth) is tested and the test confirms that the lot is bad:
  - The correct decision has been made
  - The bad lot gets scrapped or 100% inspected.
  - It may help to think about a Type 2 error as a missed alarm or accepting a bad lot.

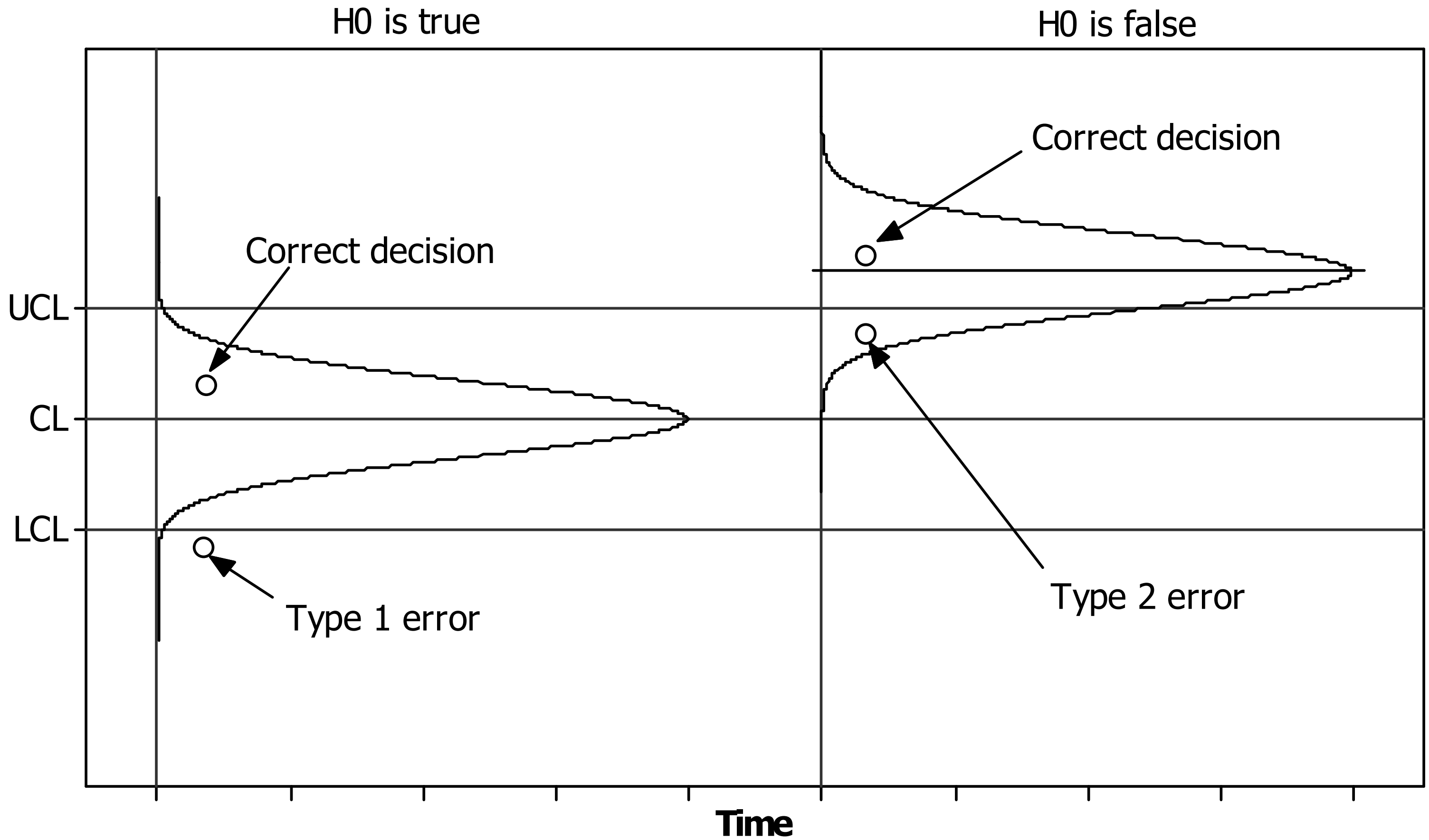
Decision Errors in Acceptance Sampling

The hypotheses are:

H_0: *the lot is good* versus H_A: *the lot is bad*



Decision Errors in SPC



Type 1 Error

Example: In a hypothesis test for $H_0 : \mu = 18$ vs. $H_A : \mu \neq 18$ the null hypothesis is accepted if the mean of a sample of size $n = 16$ falls within the interval $17.2 \leq \bar{x} \leq 18.8$. The population being sampled is normal and has $\sigma = 1.5$. Find the probability of committing a Type 1 error.

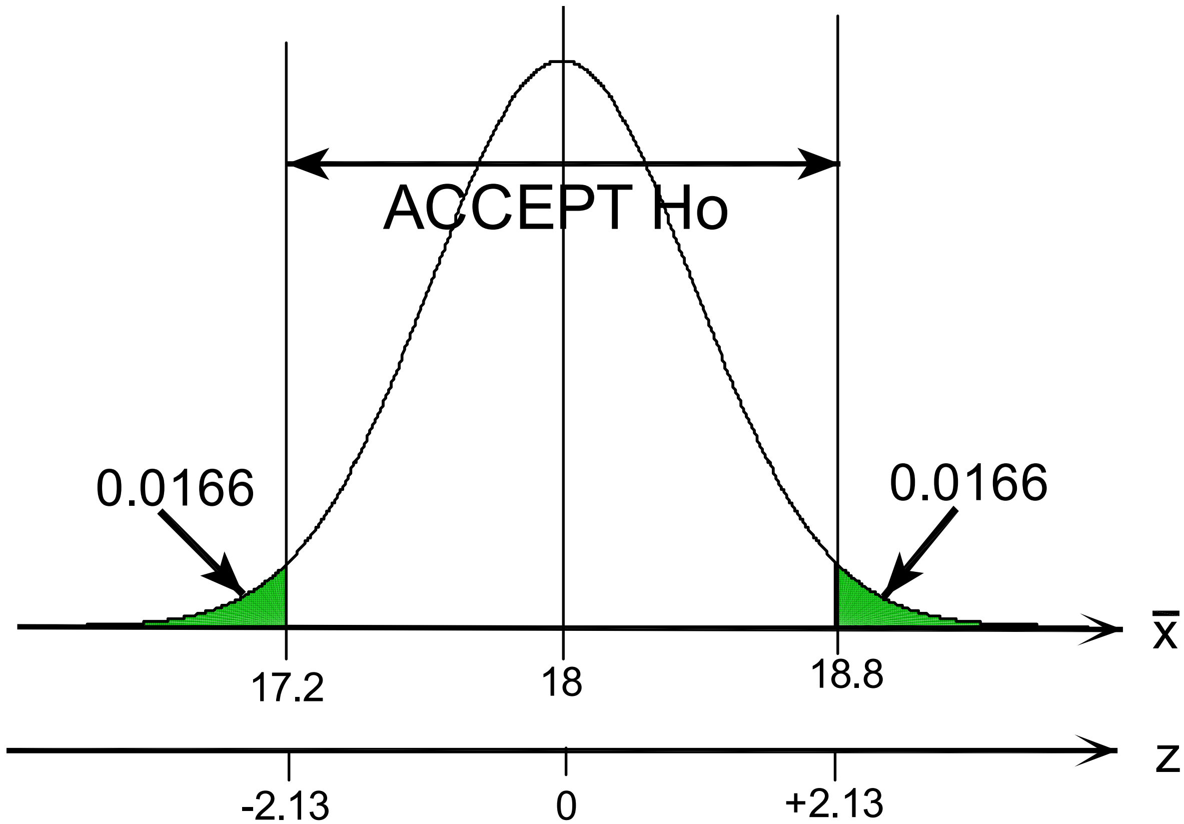
Solution: Type 1 errors occur when the null hypothesis is really true but a sample is obtained with a mean that falls outside of the acceptance interval. The probability of $\bar{x}$s falling inside the acceptance interval is:

$$\Phi(\mu - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} < \mu + z_{\alpha/2}\sigma_{\bar{x}}; \mu = \mu_0, \sigma_{\bar{x}}) = 1 - \alpha$$

where μ_0 is the hypothesized mean in the null hypothesis (i.e. $\mu_0 = 18$). If we check the upper decision limit ($UDL = 18.8$) we have $\mu + z_{\alpha/2}\sigma_{\bar{x}} = UDL$ and solving for $z_{\alpha/2}$:

$$z_{\alpha/2} = \frac{UDL - \mu}{\sigma_{\bar{x}}} = \frac{18.8 - 18.0}{1.5/\sqrt{16}} = 2.13$$

Similarly, the lower decision limit $(LDL = 17.2)$ corresponds to $-z_{0.0166} = -2.13$. Since $z_{0.0166} = 2.13$ the probability of committing a Type 1 error is $\alpha = 2(0.0166) = 0.033$.



Type 2 Error

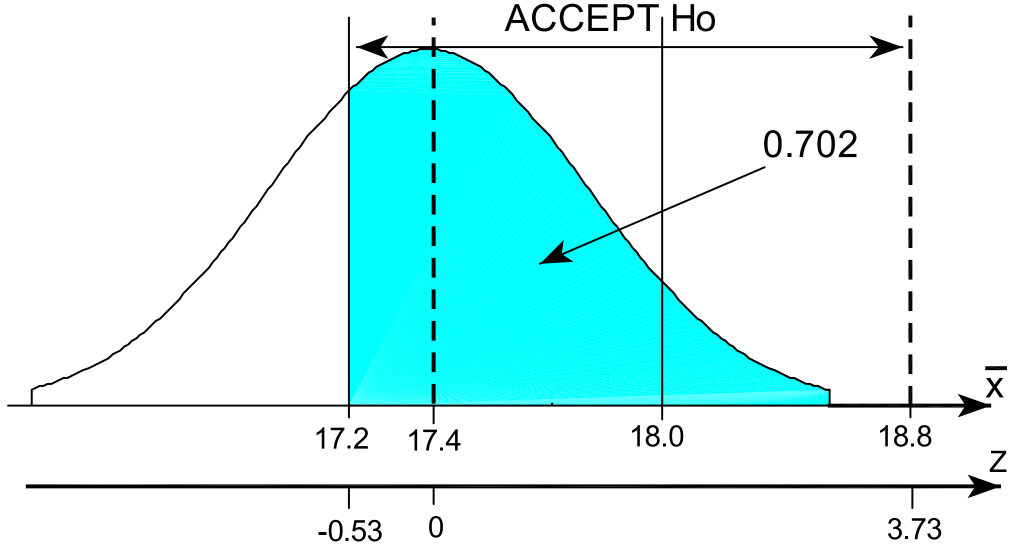
Example: In a hypothesis test for $H_0 : \mu = 18$ vs. $H_A : \mu \neq 18$ the null hypothesis is accepted if the mean of a sample of size $n = 16$ falls within the interval $17.2 \leq \bar{x} \leq 18.8$. The population being sampled is normal and has $\sigma = 1.5$. Find the probability of committing a Type 2 error when the true mean is $\mu = 17.4$.

Solution: Type 2 errors occur when the null hypothesis is really false but the test returns an erroneous *accept* H_0 result. The probability of committing a Type 2 error when the null hypothesis is really false is:

$$\beta = \Phi(\mu - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} < \mu + z_{\alpha/2}\sigma_{\bar{x}}; \mu \neq \mu_0; \sigma_{\bar{x}})$$

In this case we have:

$$
\begin{aligned}
\beta &= \Phi(\mu - z_{\alpha/2}\sigma_{\bar{x}} < \bar{x} < \mu + z_{\alpha/2}\sigma_{\bar{x}}; \mu \neq \mu_0; \sigma_{\bar{x}}) \\
&= \Phi(17.2 < \bar{x} < 18.8; \mu = 17.4; 0.375) \\
&= \Phi(-0.53 < z < 3.73) \\
&= 1.00 - 0.298 \\
&= 0.702
\end{aligned}
$$



Managing Errors

- At the time that we decide to accept H_0 (or better: reserve judgement) or reject H_0 in favor of H_A, we can't know if we've committed an error or not.
- However, we **can** control the rate at which the errors occur so that we have a chance of managing our risks.
- Setting the error rates to specified, controlled values effectively sets an experiment's sample size and acceptance criteria. We'll learn how to do this later.

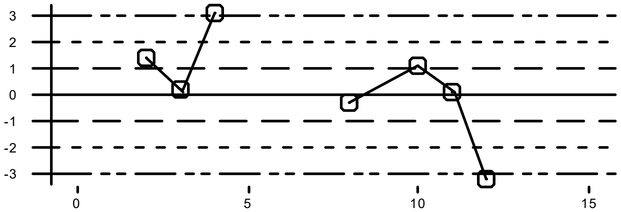
Return to SPC:
Good and Bad Run Rules

Run rules for SPC charts must meet the following conditions:

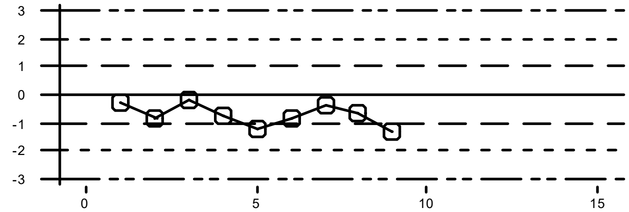
1. They must be easy to recognize on the chart.
2. They must not have a high probability of turning on when the process is really in control, i.e. α is small.
3. They must have a high probability of turning on when the process is out of control, i.e. $1-\beta$ is large.

If the first rule is satisfied then the third rule is usually OK, too. The second rule is the hardest of the three to meet.

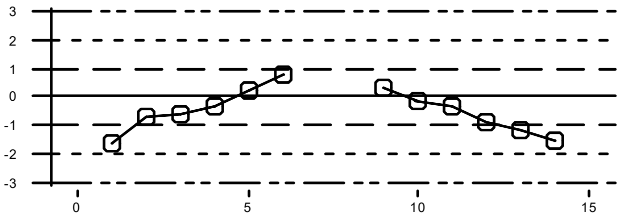
TEST 1: One point beyond 3 σ



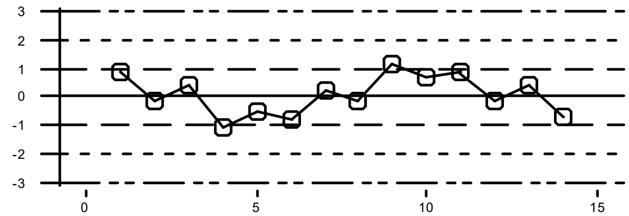
TEST 2: 9 points to one side of CL



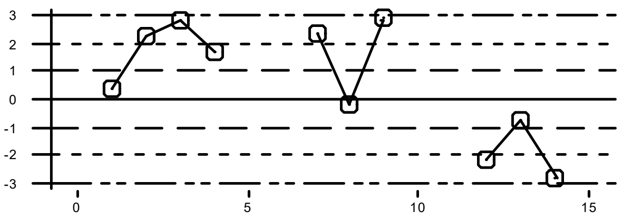
TEST 3: 6 point trend up or down



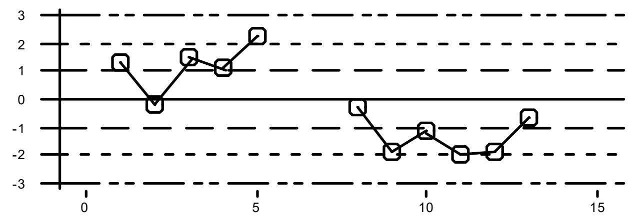
TEST 4: 14 points alternating up/down



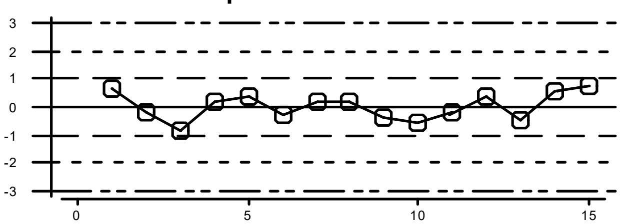
TEST 5: 2 of 3 points beyond 2σ



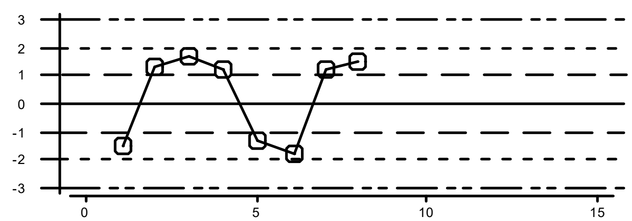
TEST 6: 4 out of 5 points beyond 1σ



TEST 7: 15 points within +/-1σ



TEST 8: 8 points beyond +/-1 σ



Created by: Rebecca Malnar 9/12/99

Example: Evaluate the run rule: *If at least four of five consecutive points fall beyond* 1σ *to the same side of the center line then the process is out of control.*

Solution: This pattern is easy to recognize on the chart so the first condition is met.

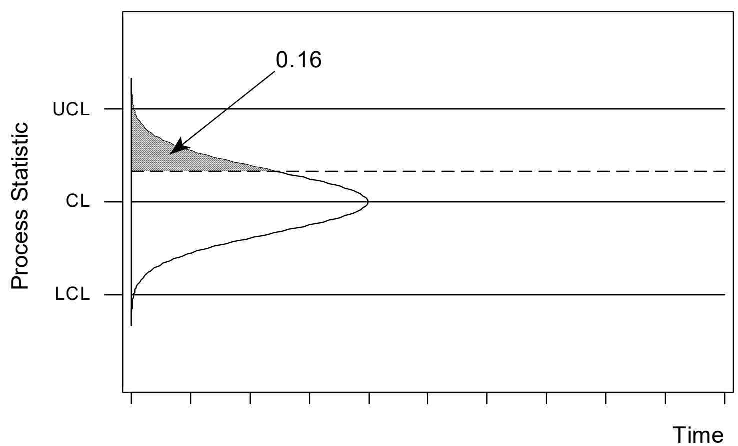
For the second condition, if the process is in control then the probability that any one point falls beyond 1σ of the centerline is:

$$\Phi(1 < z < \infty) = 0.16$$

Then the probability that at least four of five points fall beyond 1σ of the centerline is:

$$b(4 \leq x \leq 5; 5, 0.16) = 0.0029$$

Since this pattern can show up on either side of the chart we have $\alpha = 2(0.0029) = 0.0058$ which is acceptably low.



For the third condition, suppose the process shifts so the new process mean falls right on a control limit. Then the probability of a single point falling beyond 1σ of the center line is given by:

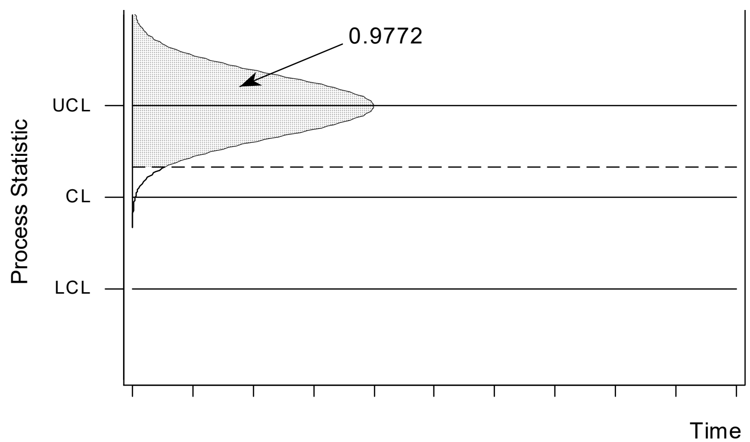
$$\Phi(-2 < z < \infty) = 0.9772$$

The probability that at least four of five consecutive points fall beyond 1σ of the center line is then:

$$b(4 \leq x \leq 5; 5, 0.9772) = 0.9950$$

This means that the probability of detecting the shift using this rule is about $1 - \beta = 0.9950$.

This rule meets all three conditions so it is a good run rule.



Example: Evaluate the run rule: *If at least two of three consecutive points fall beyond* 1σ *to the same side of the center line then the process is out of control.*

Solution: The rule is easy to identify so it meets the first condition.

Since the probability that one point will fall beyond 1σ of the center line is 16%, the probability that at least two of three points will meet this condition is:

$$b(2 \leq x \leq 3; 3, 0.16) = 0.069$$

Since this pattern can appear on either side of the chart the probability of the pattern occurring when the process is really in control is:

$$\alpha = 2(0.069) = 0.138$$

which is excessive. This is not a good run rule.

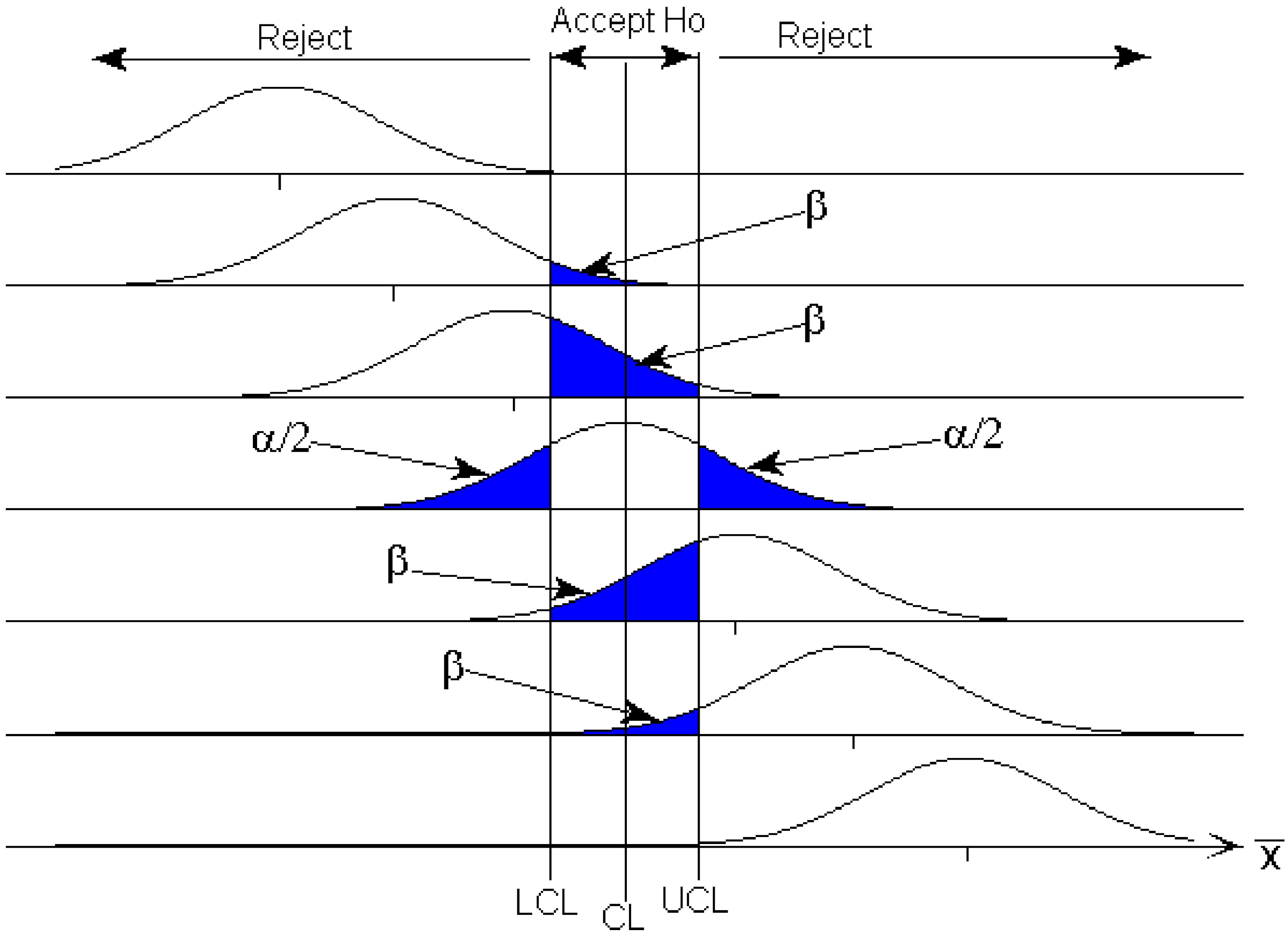
OC Curves and ARL Curves for $\bar{x}$ Charts

- The OC curve for the $\bar{x}$ chart is given by:

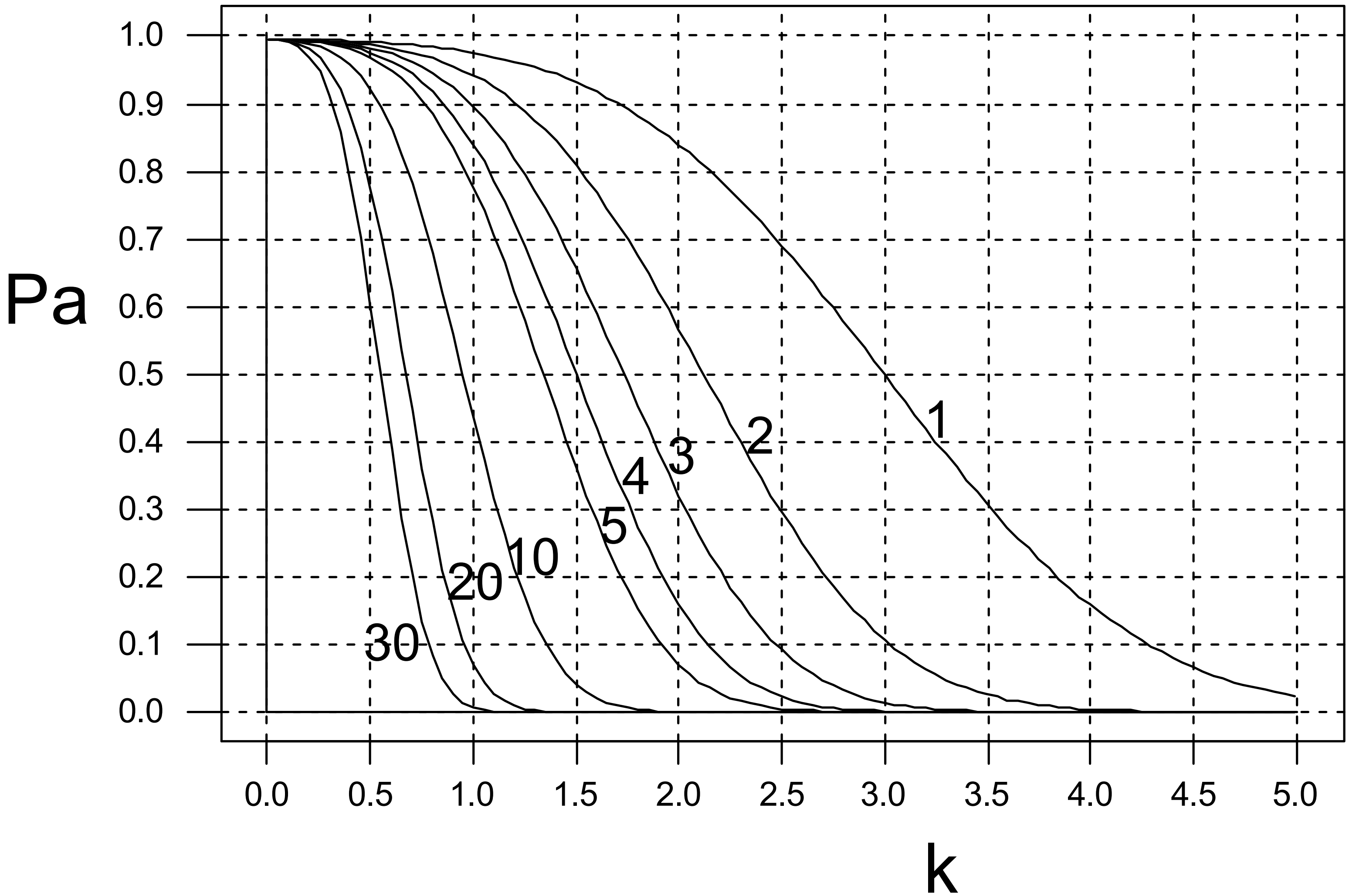
$$\Phi(LCL < \bar{x} < UCL; \mu \neq \mu_0, \sigma_{\bar{x}}) = \beta$$

- The Average Run Length is the expected number of subgroups that will be drawn in order to detect a shift. The *ARL* is given by:

$$ARL = \sum_{L=0}^{\infty} LP(L) = \sum_{L=0}^{\infty} L\beta^{L-1}(1-\beta) = \frac{1}{1-\beta}$$

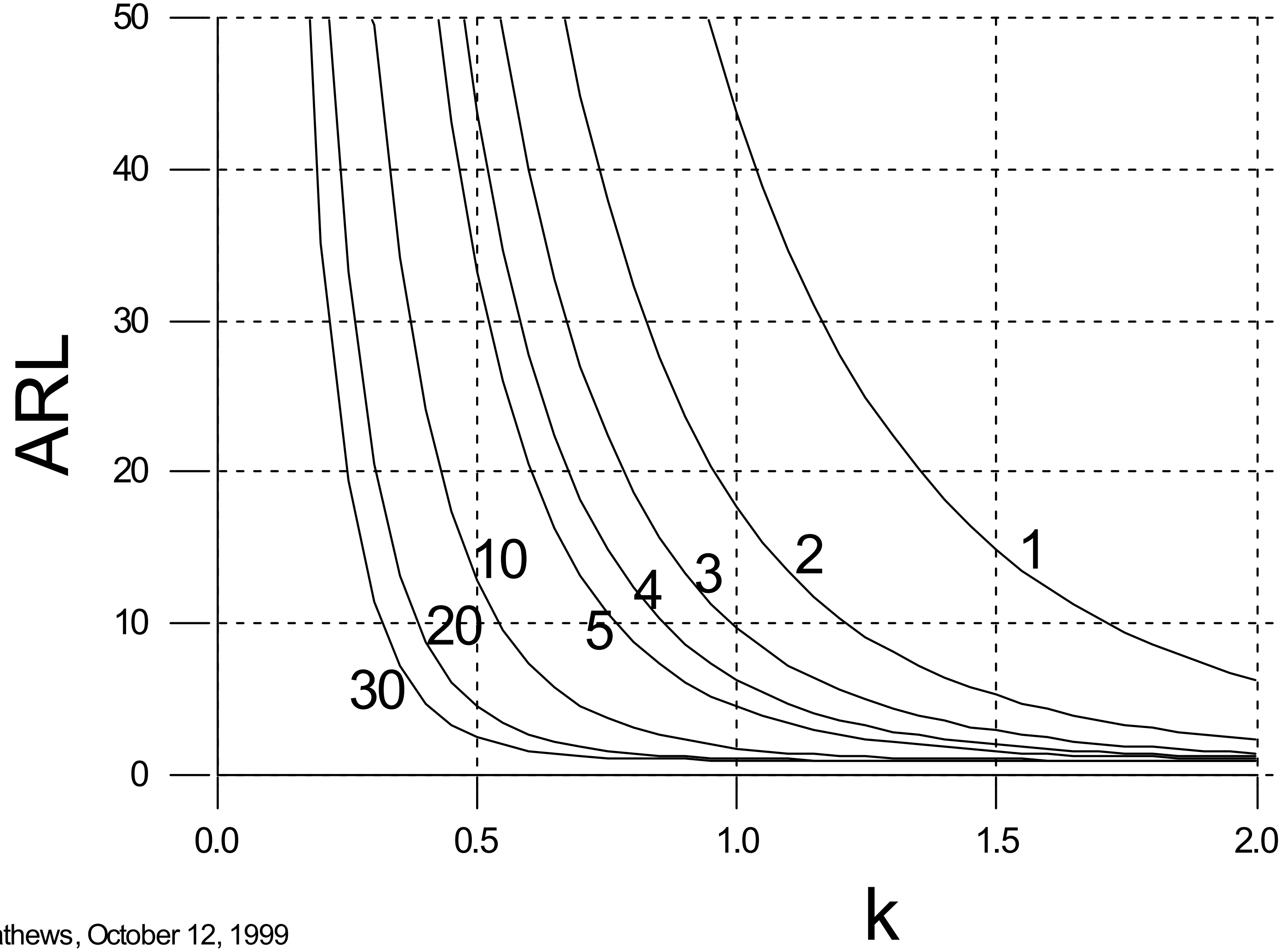

Accept Ho
Reject
Reject
β
β
α/2
α/2
β
β
LCL
CL
UCL
$\overline{x}$

OC Curves for X-bar Charts: Pa(k;n)



PGMathews, October 12, 1999

ARL Curves for X-bar Charts: ARL(k;n)



PGMathews, October 12, 1999

Two Sample t Test

Two Sample t Test

Data: Two independent samples of measurement data drawn from normal populations with equal variances ($\sigma_1^2 = \sigma_2^2$) and sample sizes n_1 and n_2. The sample sizes are allowed to be different.

Hypotheses Tested:

- $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$
- $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 < \mu_2$
- $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 > \mu_2$

Applications:

- Compare means of two treatment groups
- Compare mean of a treatment group to mean of a control group, e.g. drug versus placebo
- Compare three or more treatment groups with modifications to the method

Two Sample t Test

Test Statistic: The t test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where the *pooled* standard deviation or *standard error* is

$$\begin{aligned} s_{pooled} &= \sqrt{\frac{\sum \epsilon_{1i}^2 + \sum \epsilon_{2i}^2}{df_1 + df_2}} \\ &= \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{df_\epsilon}} \end{aligned}$$

and the error degrees of freedom are

$$\begin{aligned} df_\epsilon &= df_1 + df_2 \\ &= n_1 - 1 + n_2 - 1 \\ &= n_1 + n_2 - 2 \end{aligned}$$

Two Sample t Test

Critical Values: With test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

the critical values for the t test are:

H_0	H_A	Accept H_0	Reject H_0
$\mu_1 = \mu_2$	$\mu_1 \neq \mu_2$	$-t_{\alpha/2,df_\epsilon} < t < t_{\alpha/2,df_\epsilon}$	$t < -t_{\alpha/2,df_\epsilon}$ or $t > t_{\alpha/2,df_\epsilon}$
$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	$t > -t_{\alpha,df_\epsilon}$	$t < -t_{\alpha,df_\epsilon}$
$\mu_1 = \mu_2$	$\mu_1 > \mu_2$	$t < t_{\alpha,df_\epsilon}$	$t > t_{\alpha,df_\epsilon}$

Remember that we're stating the Accept H_0 case for convenience. We never accept H_0. We only ever reject H_0 or reserve judgement.

Two Sample t Test

Example: Samples are drawn from two processes to compare their means. The summary statistics are:

Sample	n	$\bar{x}$	s
1	10	278	4.4
2	12	280	5.9

Test the hypotheses $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ at the $\alpha = 0.05$ significance level.

Two Sample t Test

Solution: The pooled standard deviation is

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{(10 - 1)(4.4)^2 + (12 - 1)(5.9)^2}{10 + 12 - 2}}$$

$$= 5.28$$

The test statistic is:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_{pooled}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{278 - 280}{5.28\sqrt{\frac{1}{10} + \frac{1}{12}}}$$

$$= -0.88$$

Two Sample t Test

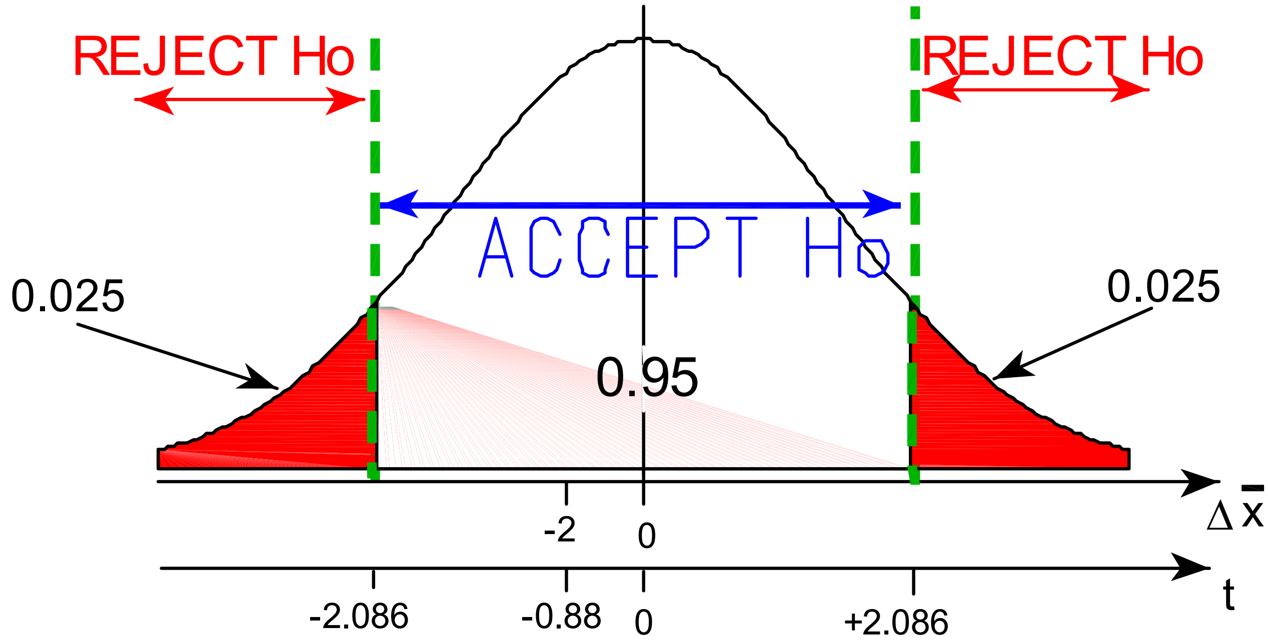
The critical $t_{\alpha/2}$ value that defines the accept/reject boundary is

$$t_{\alpha/2,df_\epsilon} = t_{0.025,20} = 2.086$$

so the acceptance interval for the null hypothesis is

$$\text{Accept } H_0 \text{ iff } -2.086 \le t \le 2.086$$

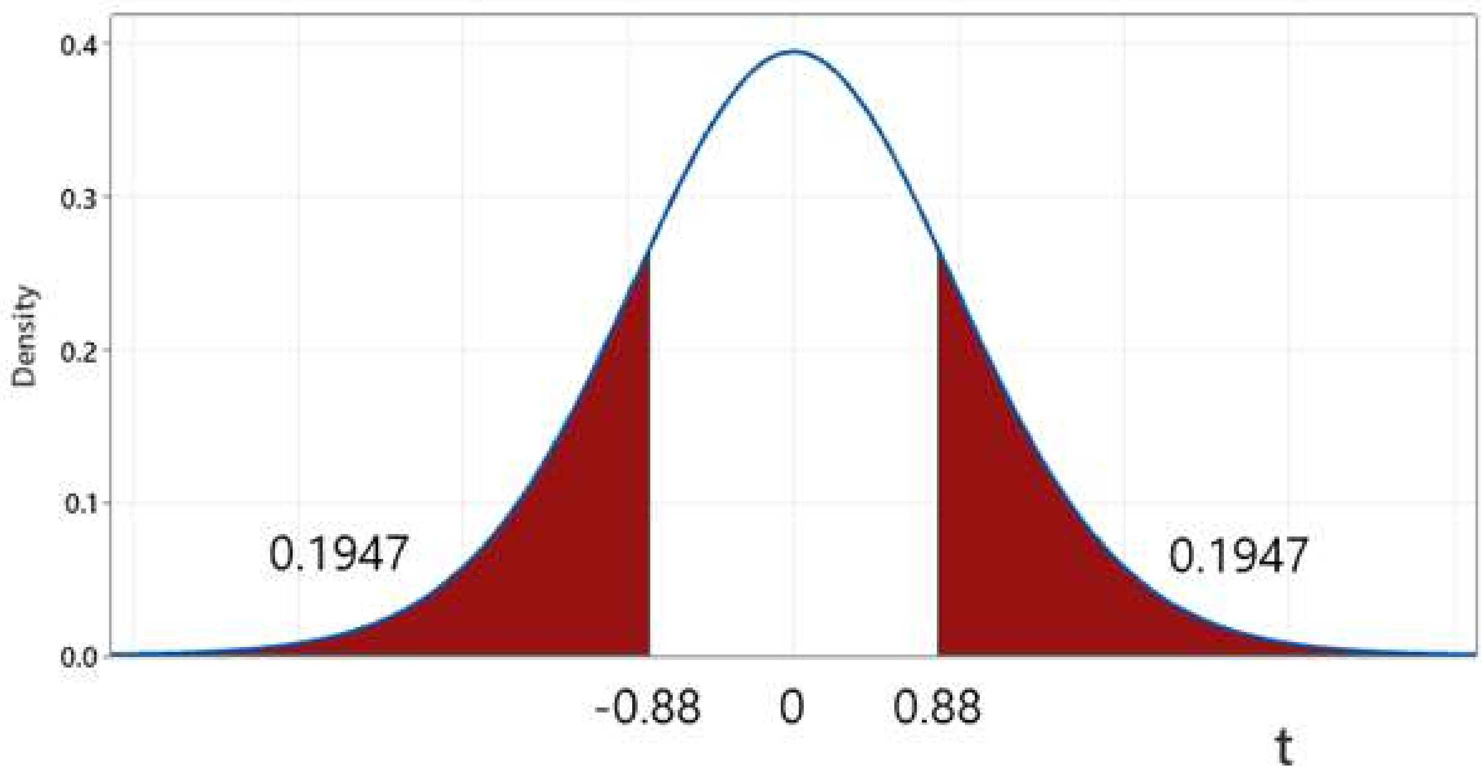
The test statistic $t = -0.88$ falls within this interval so we must accept the null hypothesis or reserve judgement.



REJECT Ho
REJECT Ho
ACCEPT Ho
0.025
0.025
0.95
$\Delta \bar{x}$
-2
0
t
-2.086
-0.88
0
+2.086

Two Sample t Test

With $t = -0.88$ and $df_\epsilon = 20$ the p value for the test is $p = 0.387$.



Confidence Interval

The confidence interval for the difference between two population means $\Delta\mu = \mu_1 - \mu_2$ is given by:

$$P(\Delta\bar{x} - \delta < \Delta\mu < \Delta\bar{x} + \delta) = 1 - \alpha$$

where

$$\Delta\bar{x} = \bar{x}_1 - \bar{x}_2$$

and the confidence interval half-width is

$$\delta = t_{\alpha/2} s_\epsilon \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

with standard error

$$s_\epsilon = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

where $t_{\alpha/2}$ has $df_\epsilon = n_1 + n_2 - 2$ degrees of freedom.

Confidence Interval

Example: Calculate the confidence interval for the example problem. The summary statistics were

Sample	n	$\bar{x}$	s
1	10	278	4.4
2	12	280	5.9

from which we determined

$$s_\epsilon = 5.28$$

$$\Delta\bar{x} = -2.0$$

$$df_\epsilon = 20$$

$$t_{0.025,20} = 2.086$$

Confidence Interval

Solution: The confidence interval half-width is

$$\begin{aligned}\delta &= t_{0.025} s_{\epsilon} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \\ &= 2.09 \times 5.28 \times \sqrt{\frac{1}{10} + \frac{1}{12}} \\ &= 4.7\end{aligned}$$

so the confidence interval is

$$\begin{aligned}P(\Delta\bar{x} - \delta < \Delta\mu < \Delta\bar{x} + \delta) &= 0.95 \\ P(-2.0 - 4.7 < \Delta\mu < -2.0 + 4.7) &= 0.95 \\ P(-6.7 < \Delta\mu < 2.7) &= 0.95\end{aligned}$$

Two Sample t Test

Behrens-Fisher Problem:

- Behrens and Fisher asked how to perform the two-sample *t* test when the two variances are not equal.
- The solution is called the **Satterthwaite** or **Welch** method.
- The Satterthwaite method applies a penalty to the *t* test's degrees of freedom based on the values of the sample variances.
- The Satterthwaite method is in excellent agreement with the assumed-equal-variances method when the variances are equal so we usually use the Satterthwaite method at all times.
- The Satterthwaite method is painful to calculate so it's usually done with software.
- The Satterthwaite method is often the default form of two-sample *t* test implemented in software.

MINITAB

- Use the **Stat**> **Basic Statistics**> **2-Sample t** menu
  - Data formats:
    - Both samples in one column with ID in another
    - Each sample in its own column
    - Summarized data
  - Use the **Graphs** submenu to choose a graphical display like boxplots or individual value plots to check for normality, equal variances (homoscedasticity), and outliers.
  - The **Options** submenu allows the choice of the equal variances *t* test or Satterthwaite's *t* test. The default is Satterthwaite.

Example: Confirm the results of the manual calculations for the two-sample t test for the example problem using MINITAB.

Solution: Use the **Stat**> **Basic Statistics**> **2-Sample t** menu with the equal variances method chosen in the **Options** submenu:

Two-Sample T-Test and CI

Method

μ_1: population mean of Sample 1
μ_2: population mean of Sample 2
Difference: $\mu_1 - \mu_2$

Equal variances are assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Sample 1	10	278.00	4.40	1.4
Sample 2	12	280.00	5.90	1.7

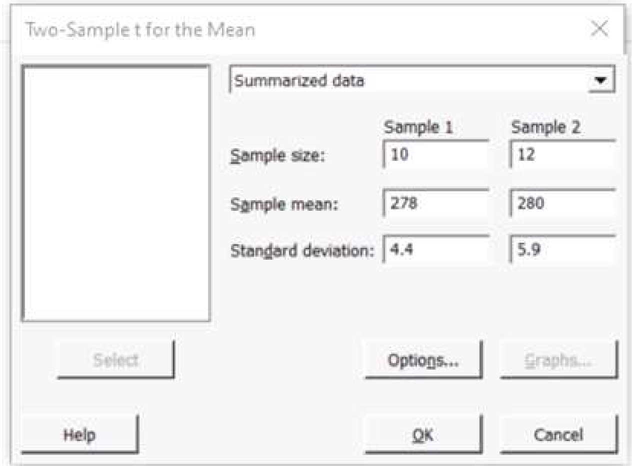
Estimation for Difference

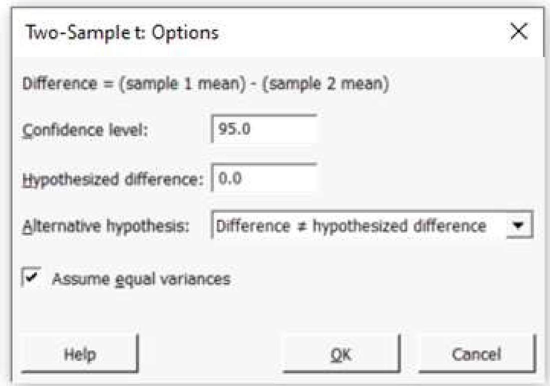
Difference	Pooled StDev	95% CI for Difference
-2.00	5.28	(-6.71, 2.71)

Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-0.88	20	0.387





Example: Reanalyze the two-sample t test for the example problem using the Satterthwaite method and compare the results between the two methods.

Solution: Use the **Stat**> **Basic Statistics**> **2-Sample t** menu with the Satterthwaite method chosen in the **Options** submenu:

Two-Sample T-Test and CI

Method

μ_1: population mean of Sample 1
μ_2: population mean of Sample 2
Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Sample 1	10	278.00	4.40	1.4
Sample 2	12	280.00	5.90	1.7

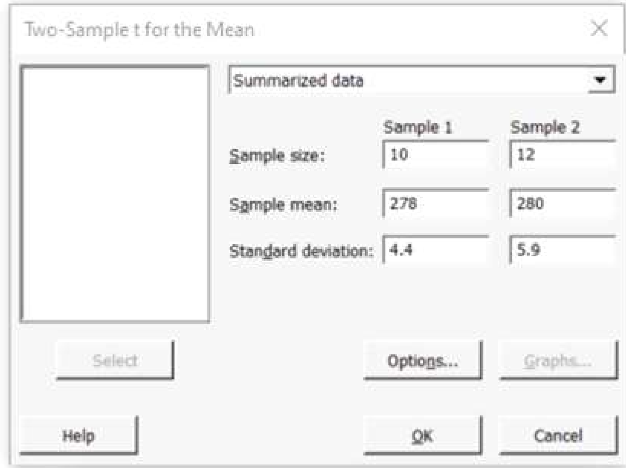
Estimation for Difference

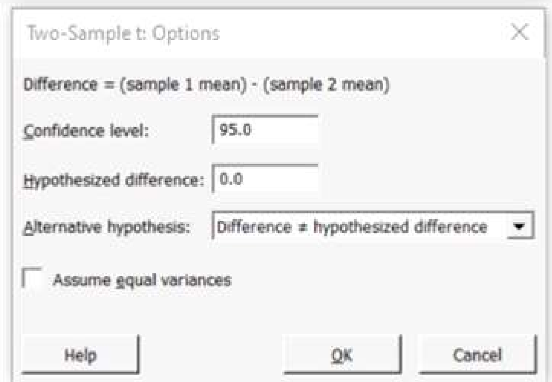
Difference	95% CI for Difference
-2.00	(-6.60, 2.60)

Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
-0.91	19	0.375





WARNING: Multiple Comparisons Tests

- The two-sample t test is used for testing for a difference between two population means but it is adapted for use under many other circumstances.
- One of the common adaptations is in the *multiple comparisons tests* situation:
  - When there are $k \geq 3$ treatment groups and we want to compare all possible pairs of treatment group means, there will be $\binom{k}{2}$ tests to perform.
  - If the Type 1 error rate for each test is set to $\alpha = 0.05$, then the overall Type 1 error rate for the family of $\binom{k}{2}$ tests will be much larger.
  - A correction to the Type 1 error rate for individual tests will be required to hold the Type 1 error rate for the family of tests to a reasonable value, e.g. $\alpha_{family} = 0.05$.

Bonferroni's Correction

- The simplest correction to the Type 1 error rate for the multiple comparisons test problem is given by Bonferroni's correction:

$$\alpha' = \frac{\alpha_{family}}{K}$$

  where α_{family} is the family error rate (usually $\alpha_{family} = 0.05$), α' is the Type 1 error rate for each planned test, and K is the number of planned tests.
- Another method of applying Bonferroni's correction is by calculating the usual two-sample t test p values and then multiplying them by K, i.e.

$$p_{corrected} = Kp$$

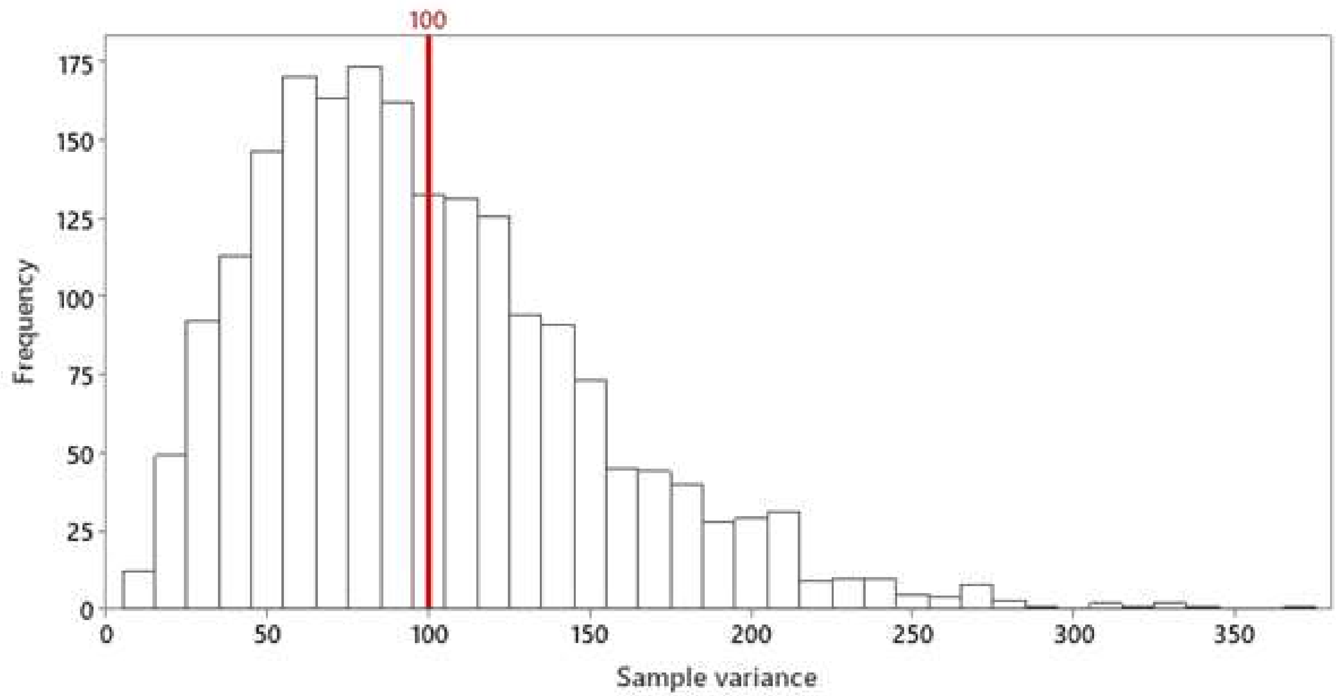
  and then comparing the $p_{corrected}$ values to the usual $\alpha = 0.05$ value.
- We will see the multiple comparisons test problem with Bonferroni's correction and other corrections in later modules.

One Variance

Distribution of Sample Variances

If repeated samples of size n are drawn from a normal population and the sample variances s^2 are calculated, then the distribution of sample variances is chi-square (χ^2) with $df = n - 1$ degrees of freedom.

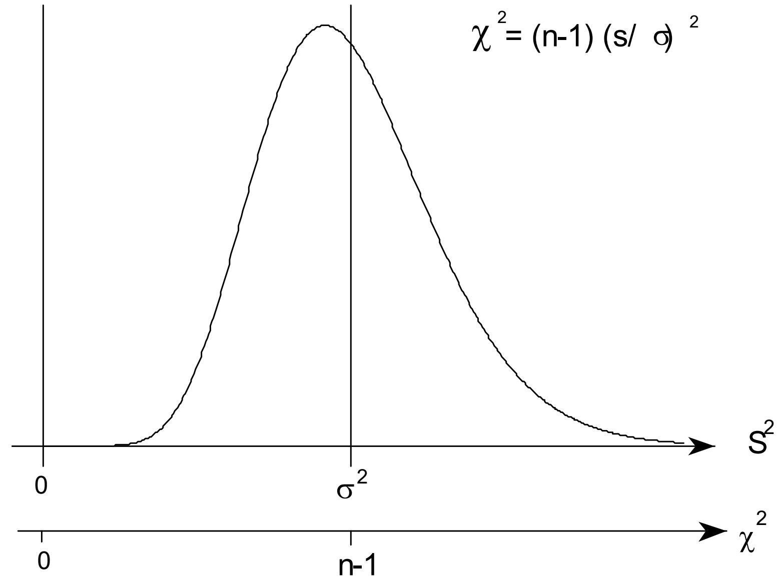
Example: The following histogram shows 2000 sample variances (s^2) calculated from random samples of sample size $n = 8$ drawn from a normal population with $\sigma^2 = 100$:


100
Frequency
175
150
125
100
75
50
25
0
0
50
100
150
200
250
300
350
Sample variance

Distribution of Sample Variances

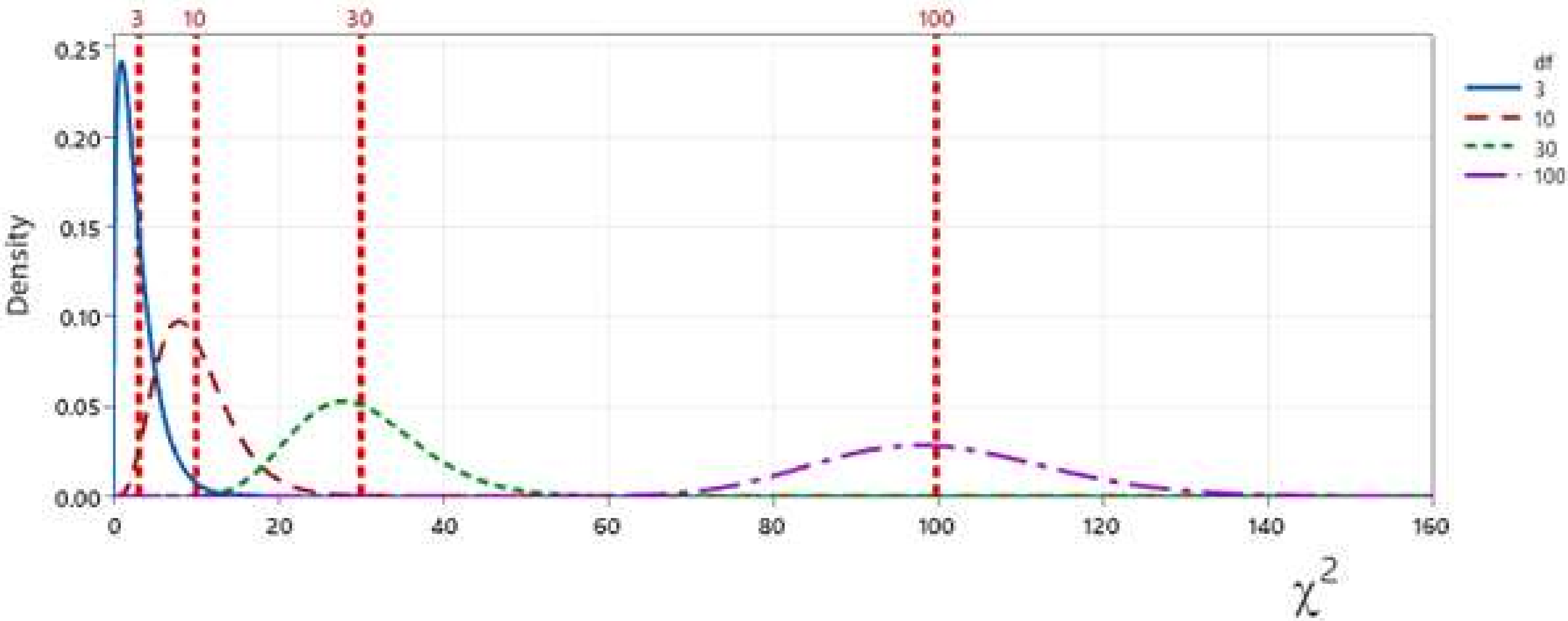
- Recall that when dealing with a normal distribution of observations (x) we transformed the x values to z values.
- Recall that when dealing with the distribution of sample means ($\bar{x}$) we transformed the $\bar{x}$ values to z values.
- When dealing with the distribution of sample variances (s^2) we transform the s^2 values to chi-square (χ^2) values where

$$\chi^2 = (n-1)\left(\frac{s}{\sigma}\right)^2$$


$\chi^2 = (n-1)(s/\sigma)^2$
S^2
0
σ^2
χ^2
0
n-1

Notes About the χ^2 Distribution

- Asymmetric: Always skewed right
- Mean is $\mu_{\chi^2} = df = n - 1$
- Changes shape as n changes
- Becomes normal (Φ) as $n \to \infty$
- Is VERY sensitive to the normality assumption
- Sample variances are very very noisy


3
10
30
100
0.25
0.20
0.15
0.10
0.05
0.00
Density
0
20
40
60
80
100
120
140
160
df
3
10
30
100
χ^2

Applications of the χ^2 Distribution

- Used to construct confidence intervals for the population variance

$$P(LCL < \sigma^2 < UCL) = 1 - \alpha$$

- Used to construct confidence intervals for the process capability c_p

$$P(LCL < c_p < UCL) = 1 - \alpha$$

- Used to determine accept/reject limits for hypothesis tests for one variance
  - $H_0 : \sigma^2 = \sigma_0^2$ versus $H_0 : \sigma^2 \neq \sigma_0^2$
  - $H_0 : \sigma^2 = \sigma_0^2$ versus $H_0 : \sigma^2 < \sigma_0^2$
  - $H_0 : \sigma^2 = \sigma_0^2$ versus $H_0 : \sigma^2 > \sigma_0^2$

Confidence Interval for σ^2

The two sided confidence interval for σ^2 determined from the sample variance s^2 with a sample of size n is given by:

$$P\left(\frac{n-1}{\chi^2_{1-\alpha/2}}s^2 < \sigma^2 < \frac{n-1}{\chi^2_{\alpha/2}}s^2\right) = 1 - \alpha$$

where the chi-square distribution has $df = n - 1$ degrees of freedom.

Notation warning: As used here, the subscript on χ^2 indicates the left tail area under the χ^2 distribution. Some texts and software index χ^2 tables by the right tail area instead. Remember that the χ^2 distribution's mean is its degrees of freedom $df = n - 1$ which you can use to check if a χ^2 table is indexed by left or right tail area.

Confidence Interval for σ^2

Example: A random sample of size $n = 18$ taken from a normal population yields a standard deviation of $s = 5.4$. Determine the 95% confidence interval for the population standard deviation.

Solution: The confidence interval is given by:

$$P\left(\frac{n-1}{\chi^2_{1-\alpha/2}}s^2 < \sigma^2 < \frac{n-1}{\chi^2_{\alpha/2}}s^2\right) = 1-\alpha$$

From the χ^2 tables we find

$$\chi^2_{0.025,17} = 7.56$$

$$\chi^2_{0.975,17} = 30.19$$

The required confidence interval for the population variance is:

$$P\left(\frac{17}{30.19}(5.4)^2 < \sigma^2 < \frac{17}{7.56}(5.4)^2\right) = 0.95$$

$$P(16.4 < \sigma^2 < 65.6) = 0.95$$

$$P(4.05 < \sigma < 8.10) = 0.95$$

Confidence Interval for c_p

Recall

$$c_p = \frac{USL - LSL}{6\sigma}$$

Starting from the confidence interval for σ^2:

$$P\left(\frac{n-1}{\chi^2_{1-\alpha/2}}s^2 < \sigma^2 < \frac{n-1}{\chi^2_{\alpha/2}}s^2\right) = 1 - \alpha,$$

we can manipulate the compound inequality in the parentheses ...

Confidence Interval for c_p

$$P\left(\frac{n-1}{\chi^2_{1-\alpha/2}}s^2 < \sigma^2 < \frac{n-1}{\chi^2_{\alpha/2}}s^2\right) = 1-\alpha$$

$$P\left(\sqrt{\frac{n-1}{\chi^2_{1-\alpha/2}}}\,s < \sigma < \sqrt{\frac{n-1}{\chi^2_{\alpha/2}}}\,s\right) = 1-\alpha$$

$$P\left(\sqrt{\frac{\chi^2_{\alpha/2}}{n-1}}\,\frac{1}{s} < \frac{1}{\sigma} < \sqrt{\frac{\chi^2_{1-\alpha/2}}{n-1}}\,\frac{1}{s}\right) = 1-\alpha$$

$$P\left(\sqrt{\frac{\chi^2_{\alpha/2}}{n-1}}\,\frac{USL-LSL}{6s} < \frac{USL-LSL}{6\sigma} < \sqrt{\frac{\chi^2_{1-\alpha/2}}{n-1}}\,\frac{USL-LSL}{6s}\right) = 1-\alpha$$

$$P\left(\sqrt{\frac{\chi^2_{\alpha/2}}{n-1}}\,\widehat{c}_p < c_p < \sqrt{\frac{\chi^2_{1-\alpha/2}}{n-1}}\,\widehat{c}_p\right) = 1-\alpha$$

Confidence Interval for c_p

- The $(1 - \alpha)100\%$ confidence interval for c_p based on a sample of size n is given by:

$$P\left(\hat{c}_p \sqrt{\frac{\chi^2_{\alpha/2}}{n-1}} < c_p < \hat{c}_p \sqrt{\frac{\chi^2_{1-\alpha/2}}{n-1}} \right) = 1 - \alpha$$

  where the χ^2 distribution has $n - 1$ degrees of freedom and is indexed by the left tail area.
- $\hat{c}_p$ should be determined from an estimate of σ such as:

$$\sigma \simeq \frac{\bar{R}}{d_2}$$

  where $\bar{R}$ is the center line from the control chart for the range.
- For the interval to be valid the process being studied must be in statistical control and its error distribution must be rigorously normal.

Example: Find the 95% confidence interval for c_p if a random sample of size $n = 50$ drawn from an in-control and normally distributed process yields $\widehat{c}_p = 1.60$.

Solution: The 95% confidence interval is:

$$P\left(1.60\sqrt{\frac{\chi^2_{0.025,49}}{49}} < c_p < 1.60\sqrt{\frac{\chi^2_{0.975,49}}{49}}\right) = 0.95$$

$$P\left(1.60\sqrt{\frac{31.6}{49}} < c_p < 1.60\sqrt{\frac{70.2}{49}}\right) = 0.95$$

$$P(1.28 < c_p < 1.92) = 0.95$$

Example: Find the 95% confidence interval for c_p if a random sample of size $n = 300$ drawn from an in-control and normally distributed process yields $\widehat{c}_p = 1.45$.

Solution: The 95% confidence interval is:

$$P\left(1.45\sqrt{\frac{\chi^2_{0.025,299}}{299}} < c_p < 1.45\sqrt{\frac{\chi^2_{0.975,299}}{299}}\right) = 0.95$$

$$P\left(1.45\sqrt{\frac{253.0}{299}} < c_p < 1.45\sqrt{\frac{348.8}{299}}\right) = 0.95$$

$$P(1.33 < c_p < 1.57) = 0.95$$

This indicates that a sample of size $n = 300$ must deliver an experimental c_p value of $\widehat{c}_p \geq 1.45$ to be able to safely claim that a minimum c_p condition of $c_p \geq 1.33$ has been achieved.

Hypothesis Test for One Variance

The hypotheses to be tested are $H_0 : \sigma^2 = \sigma_0^2$ vs. $H_A : \sigma^2 \neq \sigma_0^2$. The distribution of sample variances suggests the following form for the acceptance interval for H_0:

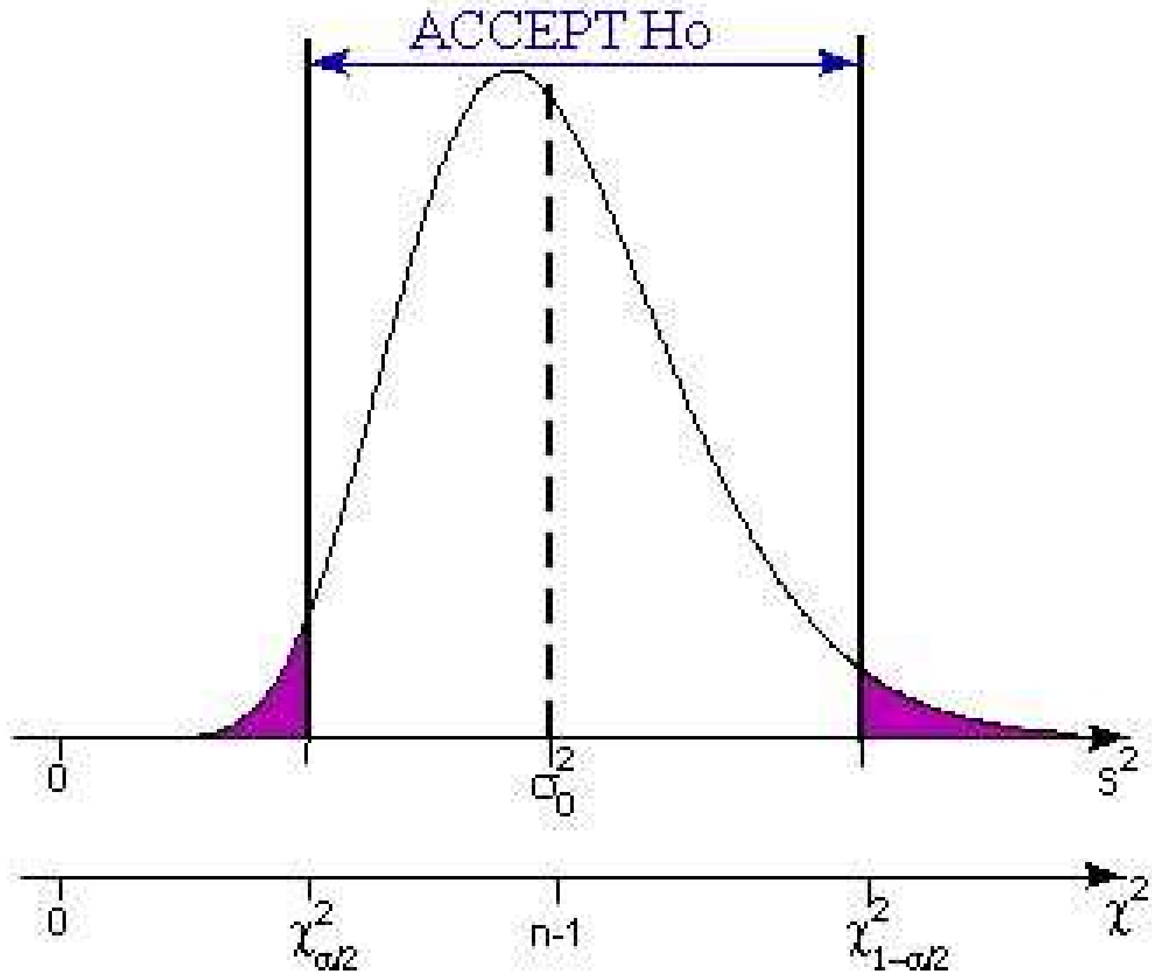
$$P\left(\frac{\chi^2_{\alpha/2}}{n-1}\sigma_0^2 < s^2 < \frac{\chi^2_{1-\alpha/2}}{n-1}\sigma_0^2\right) = 1-\alpha$$

However, it is generally easier to make the decision on the basis of the test statistic:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

with acceptance interval for the null hypothesis given by:

$$P(\chi^2_{\alpha/2} < \chi^2 < \chi^2_{1-\alpha/2}) = 1-\alpha$$


ACCEPT Ho
0
σ_0^2
S^2
0
$\chi^2_{\alpha/2}$
n-1
$\chi^2_{1-\alpha/2}$
χ^2

Hypothesis Test for One Variance

Example: Test the hypotheses $H_0 : \sigma^2 = 50$ vs. $H_A : \sigma^2 \neq 50$ at the $\alpha = 0.05$ significance level if a random sample of size $n = 25$ taken from a normal population yields $s^2 = 75$.

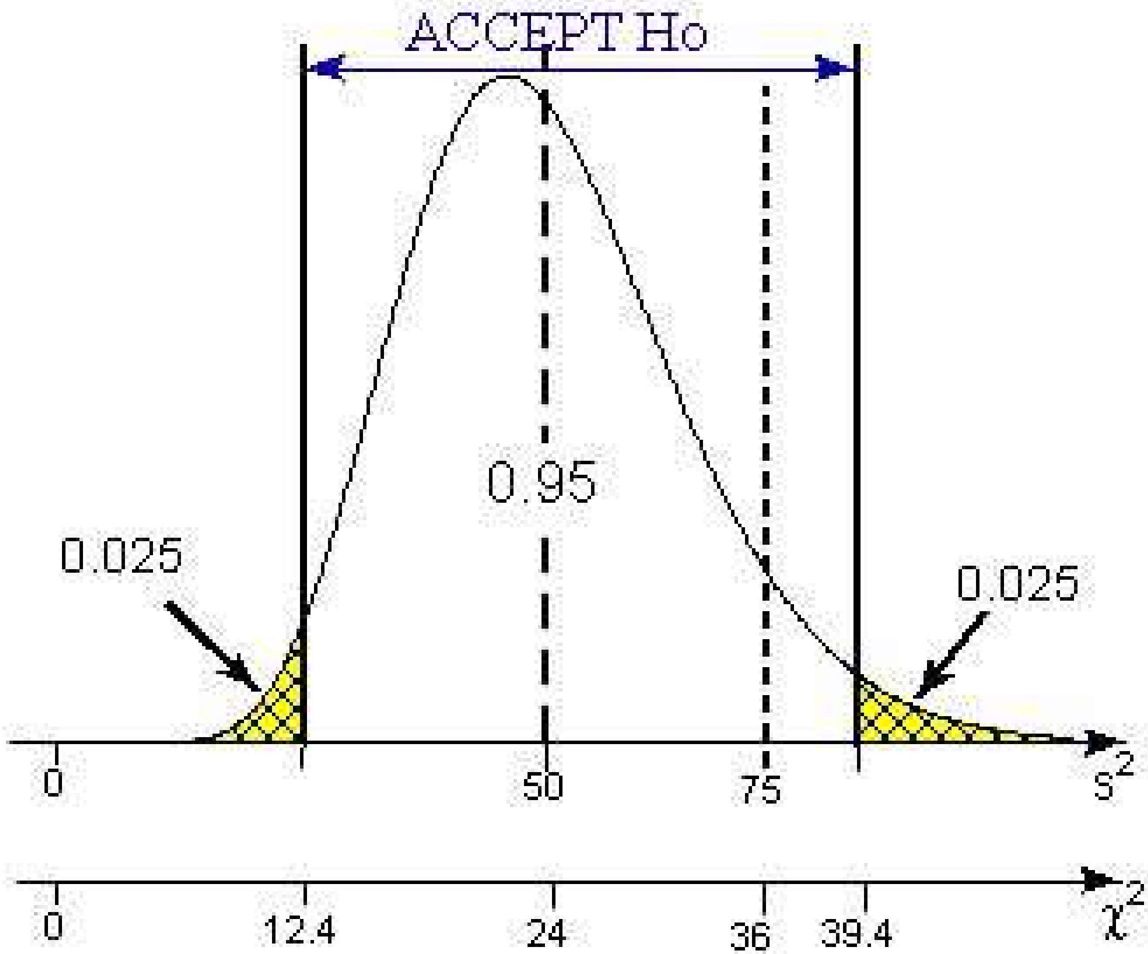
Solution: The χ^2 statistic is:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(24)75}{50} = 36$$

From the χ^2 table we have $\chi^2_{0.025,24} = 12.4$ and $\chi^2_{0.975,24} = 39.4$ so the acceptance interval for H_0 is:

$$\begin{aligned} P(\chi^2_{0.025} < \chi^2 < \chi^2_{0.975}) &= 0.95 \\ P(12.4 < \chi^2 < 39.4) &= 0.95 \end{aligned}$$

Since $\chi^2 = 36$ falls inside of the acceptance interval we must accept $H_0 : \sigma^2 = 50$ or reserve judgement.


ACCEPT Ho
0.95
0.025
0.025
0
50
75
S2
0
12.4
24
36
39.4
χ2

One Variance Problem in MINITAB

- Use the **Stat**> **Basic Statistics**> **1 Variance** menu to:
  - Perform a hypothesis test
  - Construct a confidence interval
- Data formats:
  - One sample in a column
  - Summary data:
    - Sample standard deviation (s)
    - Sample variance (s^2)

One Variance Problem in MINITAB

Example: Use MINITAB to confirm the results of the hypothesis test example problem where the hypotheses were $H_0 : \sigma^2 = 50$ versus $H_A : \sigma^2 \neq 50$ and the random sample of size $n = 25$ had sample variance $s^2 = 75$.

Solution:

Test and CI for One Variance

WORKSHEET 1

Test and CI for One Variance

Method

σ^2: variance of Sample
The Bonett method cannot be calculated for summarized data.
The chi-square method is valid only for the normal distribution.

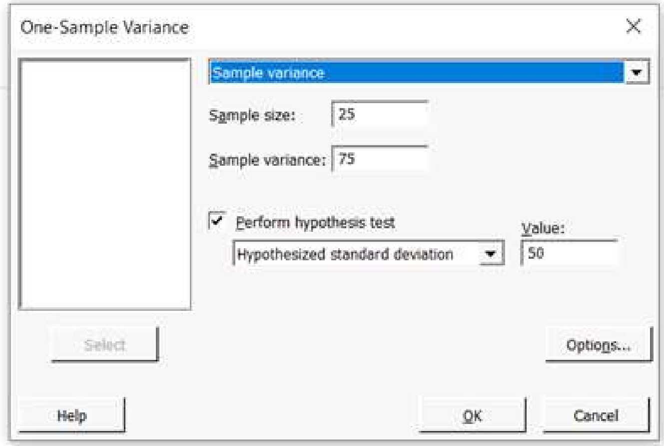
Descriptive Statistics

N	StDev	Variance	95% CI for σ^2 using Chi-Square
25	8.66	75.0	(45.73, 145.15)

Test

Null hypothesis H_0: $\sigma = 50$
Alternative hypothesis H_1: $\sigma \neq 50$

Method	Test Statistic	DF	P-Value
Chi-Square	0.72	24	0.000



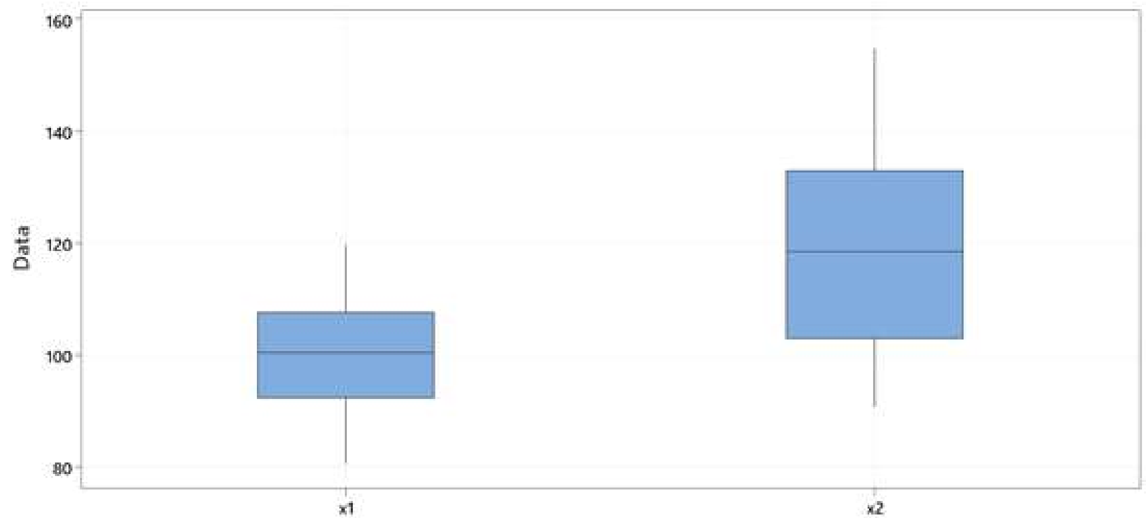
Bonett's Test

- The χ^2 distribution is very sensitive to the normality requirement so confidence intervals and hypotheses tests based on it may suffer significant errors if the population is not normal.
- Bonett's method is much less sensitive to the normality requirement so it is safer than the χ^2-based method. MINITAB implements Bonett's method in addition to the χ^2 method when it is available ($n > 20$). Use it.

Two Variances

Graphical Analysis

- Before rushing into a formal quantitative test to see if two populations have equal standard deviations or not always plot the data first.
- For example, compare two boxplots to see if there are significant differences between their ranges and IQRs.
- Whether differences in ranges and/or IQRs are sufficiently large to indicate that two populations have different standard deviations or not is a strong function of sample size.
- Follow up inspection of the boxplots with more rigorous quantitative methods.



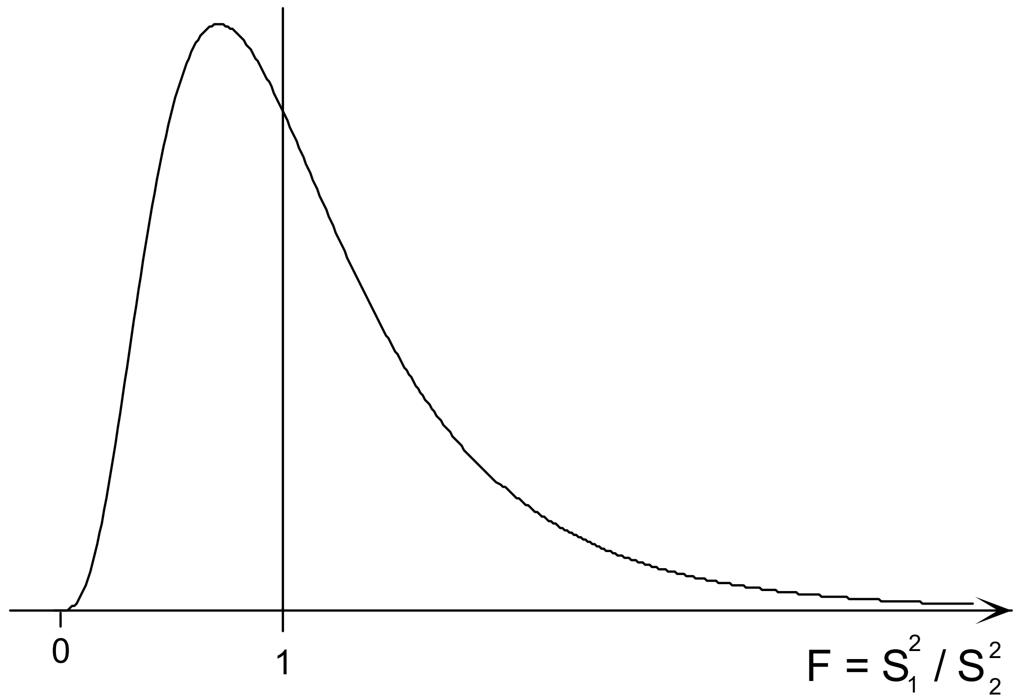
160
140
120
100
80
Data
x1
x2

Distribution of the Ratio of Two Sample Variances

If two samples of size n_1 and n_2 are drawn from normal populations that have equal population variances, then the ratio of their sample variances

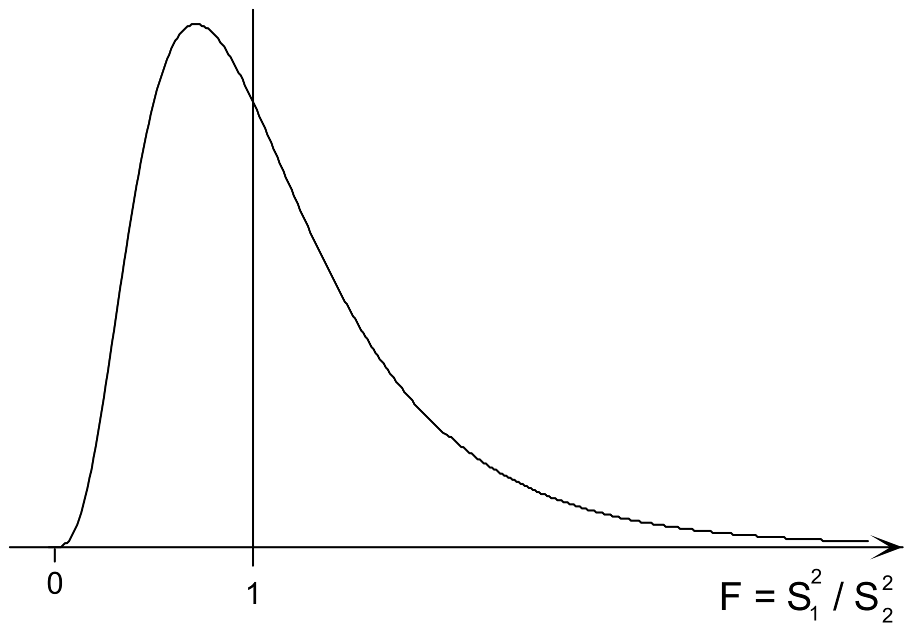
$$F = \left(\frac{s_1}{s_2}\right)^2$$

follows the F distribution with $n_1 - 1$ and $n_2 - 1$ numerator and denominator degrees of freedom, respectively.


0
1
$F = S_1^2 / S_2^2$

Notes About the F Distribution

- Always skewed right
- Mean is $\mu_F = 1$
- Changes shape as n_1 and n_2 change
- Used to determine accept/reject limits for hypothesis tests comparing two sample variances
- Plays a critical role in ANOVA, regression, and design of experiments



Notes About the F Distribution

- The F statistic

$$F = \left(\frac{s_1}{s_2}\right)^2$$

  is usually constructed with $s_1 > s_2$ so that $F > 1$ and only right tail F values are indexed in the tables, sometimes by right tail area and sometimes by left tail area.
- The reciprocal of the F statistic has the same meaning so is redundant.
- Be aware that some software (e.g. MINITAB) is not choosy about whether $F > 1$ or $F < 1$ so be prepared to see either value.
- Accessing tables of F values:
  - Use published tables to find critical values
  - Use MINITAB **Calc**> **Probability Distributions**> **F**
  - Use MINITAB **Graph**> **Probability Distribution Plot**> **View Probability**> **F**

Hypothesis Test for Two Variances

Example: Random samples of size $n_1 = 12$ and $n_2 = 16$ are drawn from two populations. The sample standard deviations are found to be $s_1 = 145$ and $s_2 = 82$. Test to see if there is evidence that the population variances are equal at the $\alpha = 0.05$ significance level.

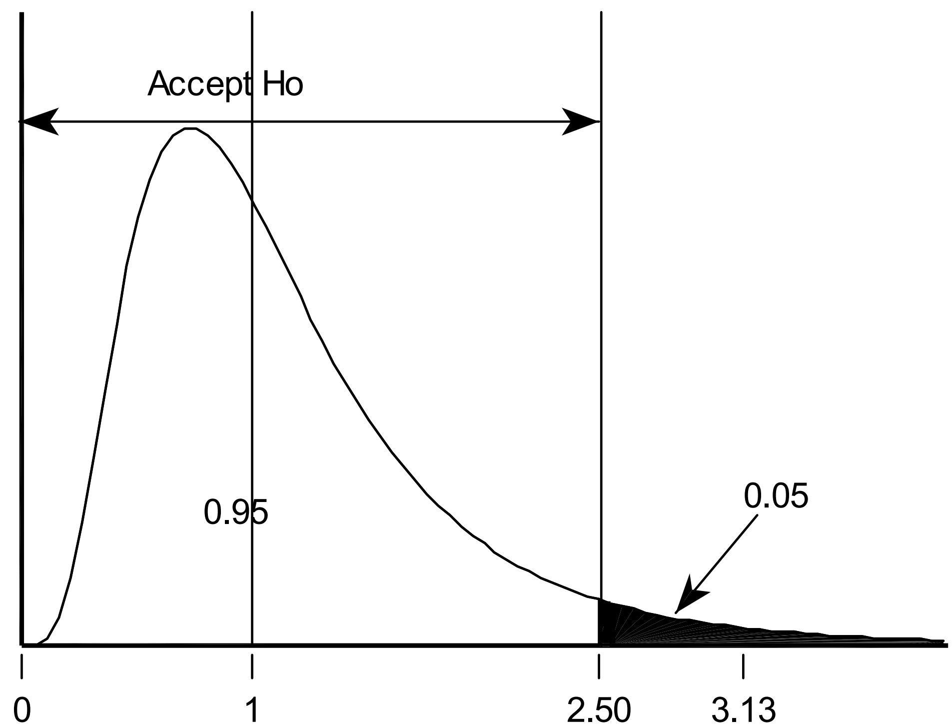
Solution: The hypotheses to be tested are $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_A : \sigma_1^2 > \sigma_2^2$. The acceptance interval for the null hypothesis is given by:

$$P\left(0 < \frac{s_1^2}{s_2^2} < F_{1-\alpha}\right) = 1 - \alpha$$

From the F tables with 11 numerator and 15 denominator degrees of freedom we find $F_{0.95} = 2.51$. The F statistic is given by:

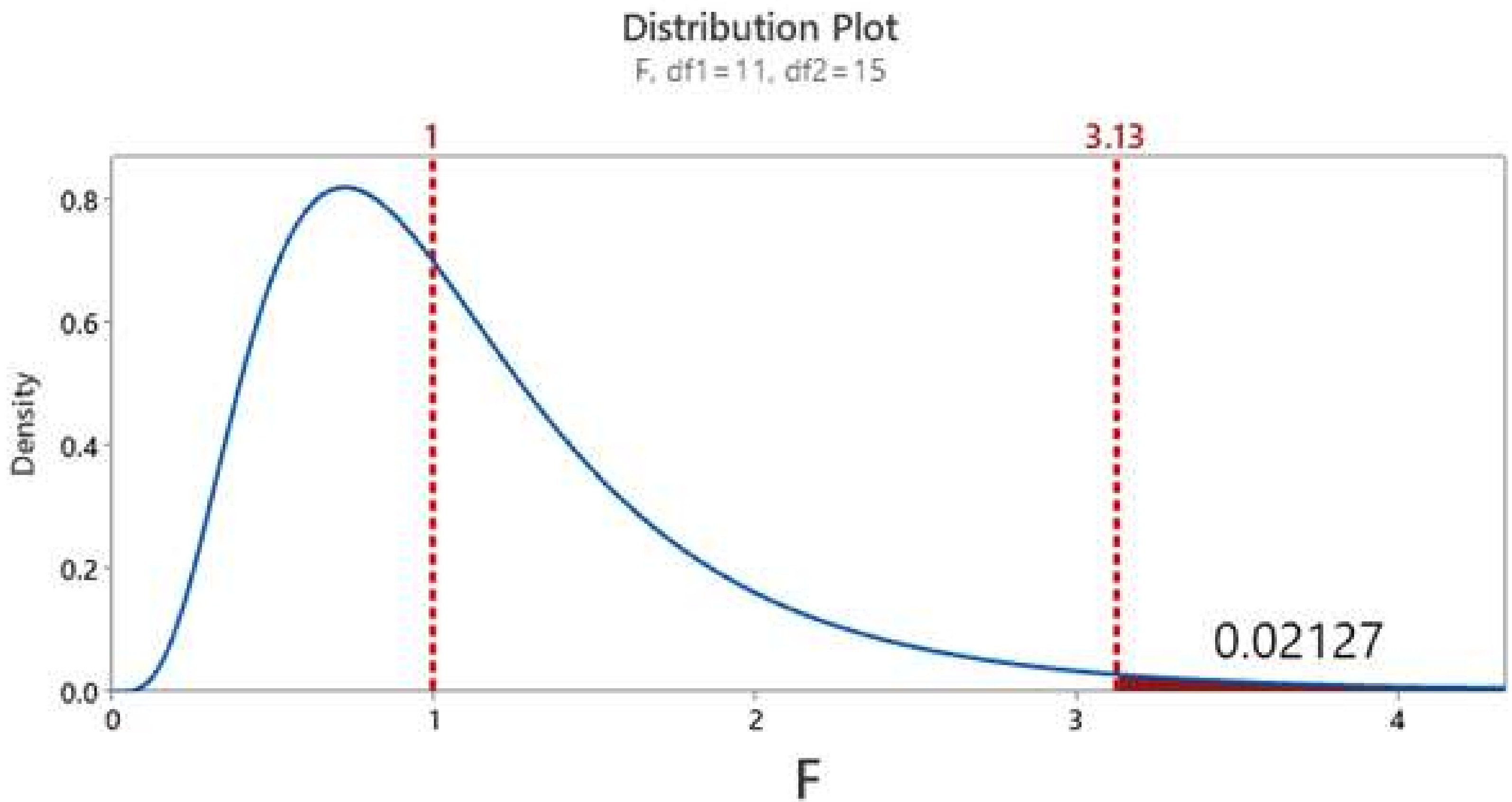
$$\begin{aligned} F &= \left(\frac{s_1}{s_2}\right)^2 \\ &= (145/82)^2 \\ &= 3.13 \end{aligned}$$

Since $F = 3.13$ falls outside the acceptance interval we must reject H_0 and conclude that there is evidence that the two populations being sampled have different variances.



F Test *p* Value

As usual, the F test's p value is the probability of obtaining the observed result or something more extreme under H_0, so the p value is the right tail area relative to the F statistic.



F Test for Two Variances in MINITAB

- Use the **Stat**> **Basic Statistics**> **2 Variances** menu to:
  - Perform a hypothesis test
  - Construct a confidence interval
- Use the **Options** submenu to choose a one-tailed or two-tailed test. The two-tailed test is MINITAB's default but a one-tailed test is often preferred.
- Data formats:
  - Both samples in one column with ID in another
  - Each sample in its own column
  - Summarized data:
    - Sample standard deviation (s)
    - Sample variance (s^2)

Example: Confirm the results from the example using MINITAB.

Solution:

Test and CI for Two Variances

WORKSHEET 1

Test and CI for Two Variances

Method

σ_1: standard deviation of Sample 1
σ_2: standard deviation of Sample 2
Ratio: σ_1/σ_2
F method was used. This method is accurate for normal data only.

Descriptive Statistics

Sample	N	StDev	Variance	95% Lower Bound for σ
Sample 1	12	145.000	21025.000	108.419
Sample 2	16	82.000	6724.000	63.522

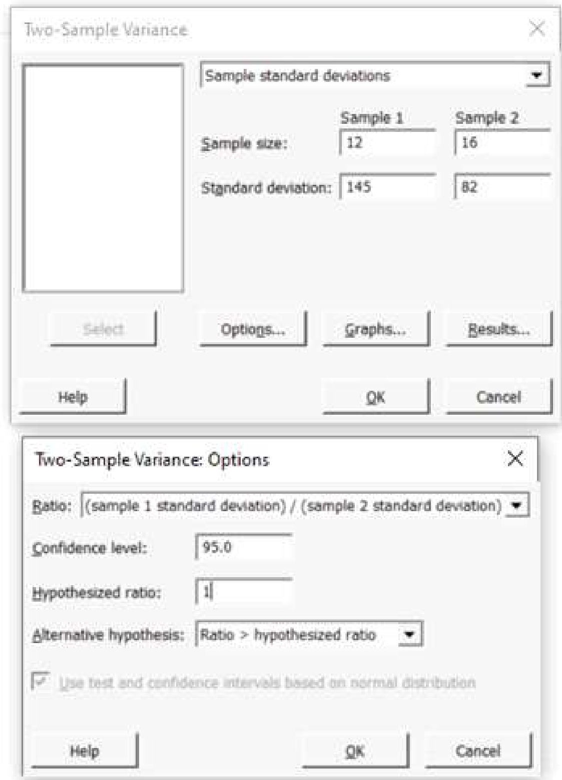
Ratio of Standard Deviations

Estimated Ratio	95% Lower Bound for Ratio using F
1.76829	1.117

Test

Null hypothesis	H_0: $\sigma_1 / \sigma_2 = 1$
Alternative hypothesis	H_1: $\sigma_1 / \sigma_2 > 1$
Significance level	$\alpha = 0.05$

Method	Test Statistic	DF1	DF2	P-Value
F	3.13	11	15	0.021



Levene's Test

- The F test for two variances is sensitive to the normality assumption
- Levene's test is less sensitive to deviations from normality so is preferred
- Levene's test is also applicable to three or more treatment groups
- Use Levene's test in MINITAB with **Stat**> **ANOVA**> **Test for Equal Variances**
  - If $p < 0.05$ reject $H_0 : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \cdots$

One Proportion

Test for One Proportion

Hypotheses:

- $H_0 : p = p_0$ versus $H_A : p \neq p_0$
- $H_0 : p = p_0$ versus $H_A : p > p_0$
- $H_0 : p = p_0$ versus $H_A : p < p_0$

Assumptions: The sample is drawn from a single stable binomial distribution.

Procedure:

1. Draw a random sample of size n from the population.
2. Inspect the sample and count the number of defectives D.
3. Reject H_0 if $p < \alpha$ where:
   - If the hypotheses are $H_0 : p = p_0$ versus $H_A : p < p_0$ then:

$$p = \sum_{x=0}^{D} b(x; n, p_0) = b(c = D; n, p_0)$$

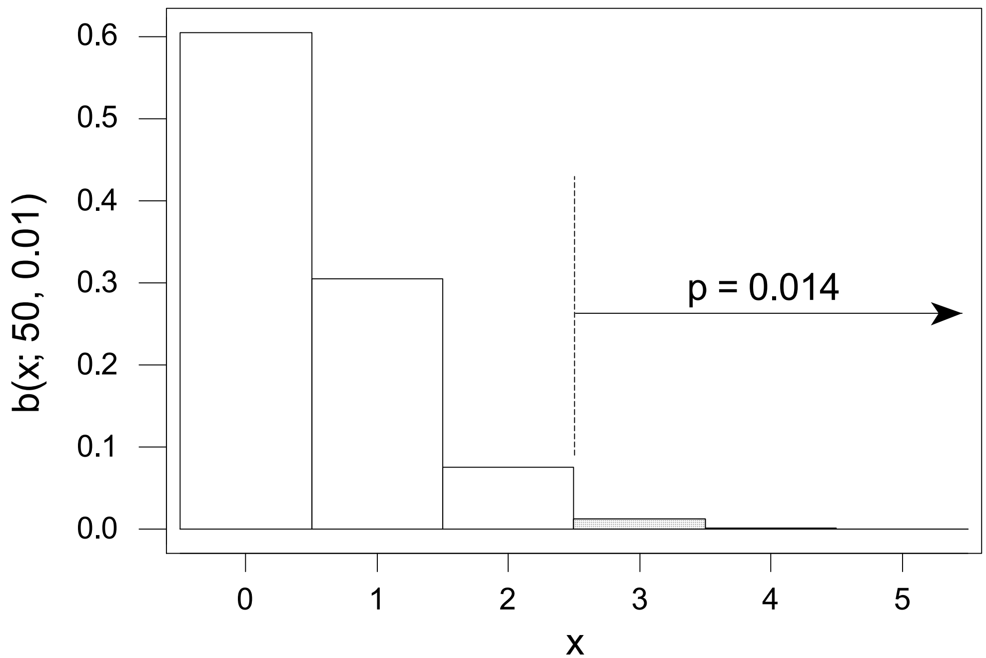
   - If the hypotheses are $H_0 : p = p_0$ versus $H_A : p > p_0$ then:

$$p = \sum_{x=D}^{n} b(x; n, p_0) = 1 - b(c = D - 1; n, p_0)$$

Example: A random sample of $n = 50$ parts was inspected and determined to have $D = 3$ defectives. Is there sufficient evidence to reject $H_0 : p = 0.01$ in favor of $H_A : p > 0.01$?
Solution: The p value for the hypothesis test is:

$$p = \sum_{x=3}^{50} b(x; 50, 0.01) = 0.014$$

and is shown in the Figure below. Since $(p = 0.014) < (\alpha = 0.05)$ there is sufficient evidence to reject H_0.



0.6
0.5
0.4
0.3
0.2
0.1
0.0
b(x; 50, 0.01)
p = 0.014
0
1
2
3
4
5
x

Example: Re-solve the example problem using MINITAB **Stat**> **Basic Statistics**> **1 Proportion**.

Solution:

Test and CI for One Propor...

WORKSHEET 1

Test and CI for One Proportion

Method

p: event proportion

Exact method is used for this analysis.

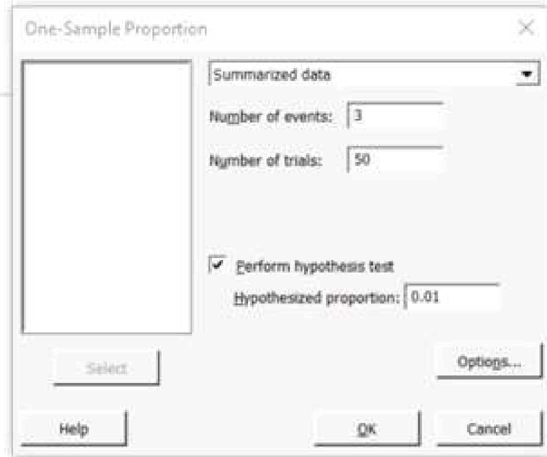
Descriptive Statistics

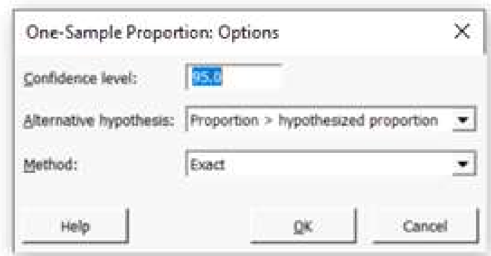
N	Event	Sample p	95% Lower Bound for p
50	3	0.060000	0.016552

Test

Null hypothesis	H_0: p = 0.01
Alternative hypothesis	H_1: p > 0.01

P-Value
0.014



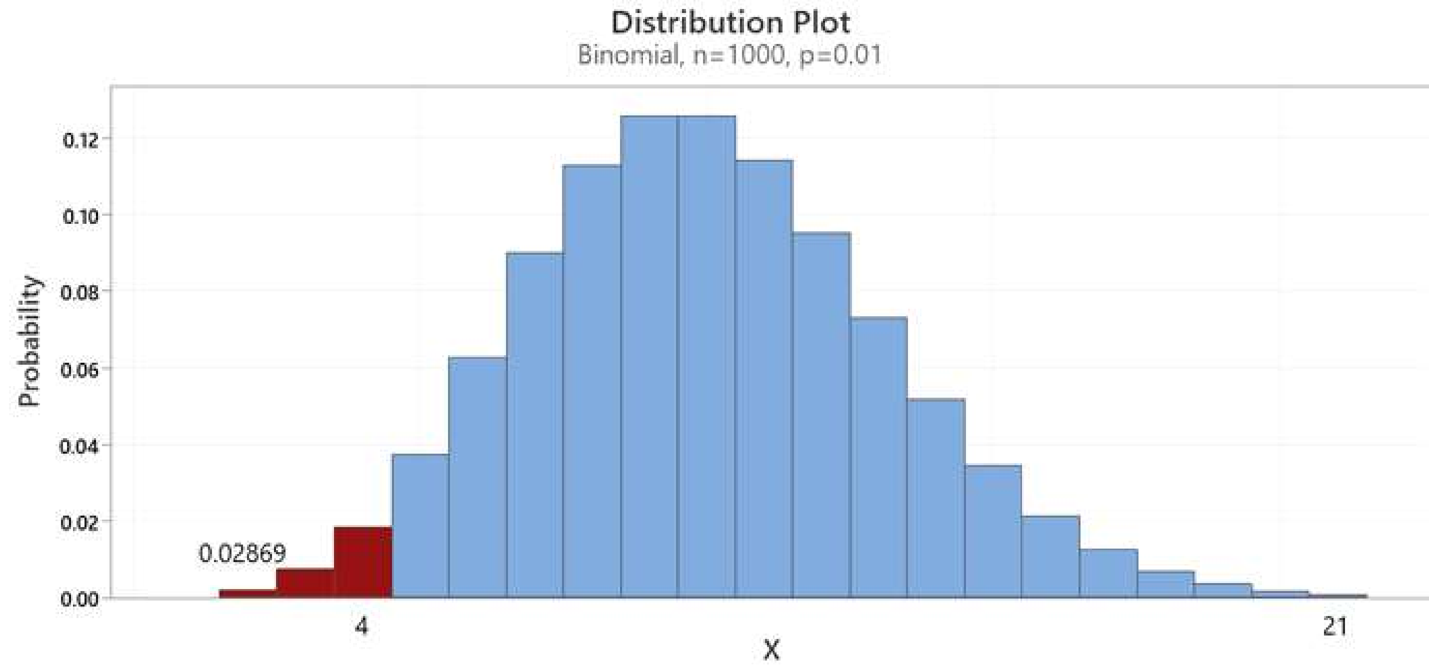


Example: A random sample of $n = 1000$ parts was inspected and determined to have $D = 4$ defectives. Is there sufficient evidence to reject $H_0 : p = 0.01$ in favor of $H_A : p < 0.01$?

Solution: The p value for the hypothesis test is:

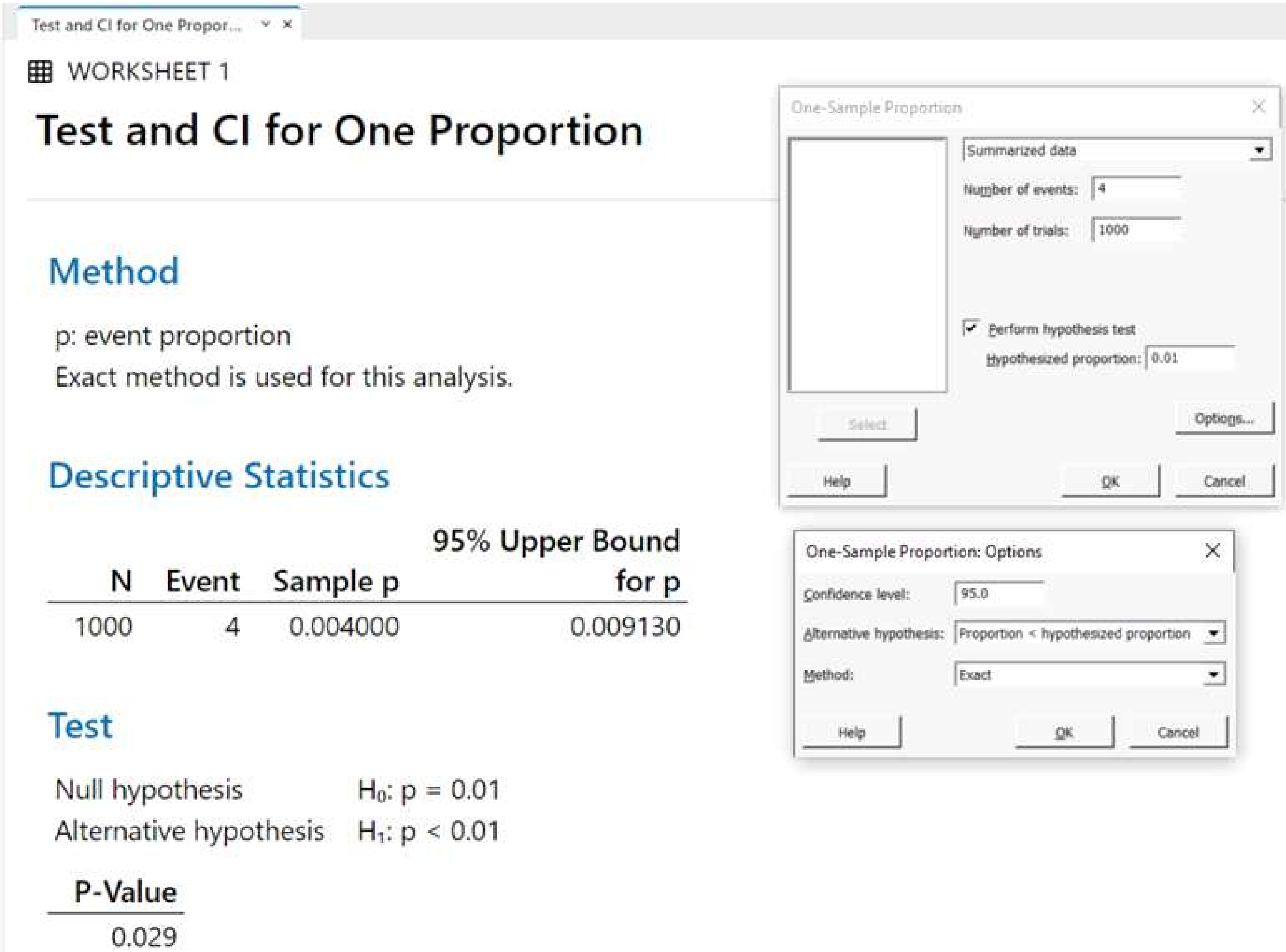
$$p = \sum_{x=0}^{4} b(x; 1000, 0.01) = 0.029$$

and is shown in the Figure below. Because $(p = 0.029) < (\alpha = 0.05)$ there is sufficient evidence to reject H_0.


Distribution Plot
Binomial, n=1000, p=0.01
Probability
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0.02869
4
21
X

Example: Re-solve the example problem using MINITAB **Stat**> **Basic Statistics**> **1 Proportion**.

Solution:


Test and CI for One Propor...

WORKSHEET 1

Test and CI for One Proportion

Method

p: event proportion

Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Upper Bound for p
1000	4	0.004000	0.009130

Test

Null hypothesis H_0: p = 0.01

Alternative hypothesis H_1: p < 0.01

P-Value
0.029

Other Applications of the One Proportion Test

- We've seen the test for one proportion before ...
- SPC defectives (np) and proportion defective (p) charts
- The hypothesis test of

$$H_0 : p = p_0 \text{ versus } H_A : p > p_0$$

  is identical to our coverage of single sampling plans for defectives and acceptance sampling by attributes where $p_0 = AQL$.
- The acceptance sampling by attributes topic solved the sample size and acceptance criterion problem for two points on the sampling plan's operating characteristic curve:

$$P_A = b(c;n,p = AQL) = 1 - \alpha$$

$$P_A = b(c;n,p = RQL) = \beta$$

Reliability Demonstration Tests

- The hypothesis test of

$$H_0 : p = p_0 \text{ versus } H_A : p < p_0$$

  or the confidence interval

$$P(0 < p < p_0) = 1 - \alpha$$

  are often used to demonstrate an upper limit on the proportion defective.
- The same problem can be redefined in terms of reliability where $R = 1 - p$, i.e.

$$H_0 : R = R_0 \text{ versus } H_A : R > R_0$$

  or

$$P(R_0 < R < 1) = 1 - \alpha$$

- These demonstration tests are often executed as simple attribute single sampling plans indexed by n and c where

$$b(c; n, p_0) \leq \alpha$$

- See Appendix B p. 19 or *Sample size to demonstrate the upper one-sided* *confidence limit for a proportion.xls.*

Example: The Excel document *Sample size to demonstrate the upper one-sided confidence limit for a proportion.xls* indicates that a sample of $n = 30$ with $D = 0$ defectives demonstrates that the proportion defective is less than $p = 0.10$ with 95% confidence, i.e.

$$P(0 < p < 0.01) = 0.95$$

The spreadsheet uses an approximation to the binomial distribution that slightly exaggerates the sample size. Use MINITAB to confirm that $n = 29$ is sufficient to make that demonstration.

Solution: Using the **Stat**> **Basic Statistics**> **1 Proportion** method:

Test and CI for One Propor...

WORKSHEET 1

Test and CI for One Proportion

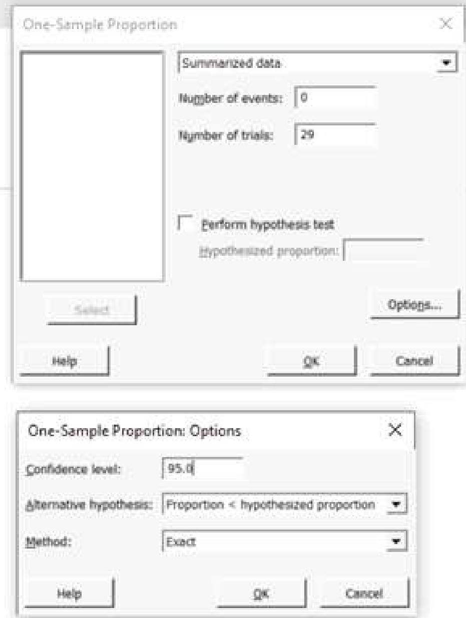
Method

p: event proportion

Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% Upper Bound for p
29	0	0.000000	0.098145



Two Proportions

Fisher's Exact Test for Two Proportions

Hypotheses:

$H_0 : p_1 = p_2$ versus $H_A : p_1 < p_2$

Assumptions:

- The samples are independent.
- The samples are random.
- The samples are drawn from binomial populations.

Procedure:

1. Draw random samples of size A from population 1 and B from population 2. The samples should be approximately the same size.
2. Inspect the samples of size A and B and determine the number of defectives in each sample, a and b, respectively. By arbitrary choice $\frac{a}{A} < \frac{b}{B}$.
3. Determine the p value for the test from:

$$p = \sum_{x=0}^{a} \frac{\binom{A}{x}\binom{B}{a+b-x}}{\binom{A+B}{a+b}}$$

4. If $p < \alpha$ then reject H_0 and conclude that $p_1 < p_2$.

Example: The fractions defective of two processes are to be compared. Two samples of the same size, $A = B = 100$, are drawn. Upon inspection A is found to have $a = 2$ defectives and B has $b = 8$ defectives. Is there evidence that the fraction defective of the first process (A) is smaller than the fraction defective of the second process (B)? That is, we want to test the hypotheses:

$$H_0 : p_1 = p_2$$

$$H_A : p_1 < p_2$$

Make the decision using a significance level of $\alpha = 0.05$.

Solution: The total number of parts inspected was $A + B = 200$ and the total number of defective parts found was $a + b = 10$. The list of all possible (a, b) pairs that sum to 10 and the corresponding hypergeometric probabilities $P(a) = h(a; 100, 100, 10)$ are:

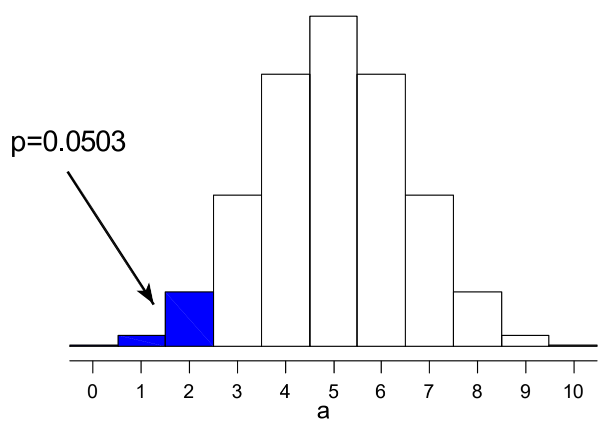
a	0	1	2	3	4	5	6	7	8	9	10
b	10	9	8	7	6	5	4	3	2	1	0
P	0.001	0.008	0.041	0.115	0.208	0.252	0.208	0.115	0.041	0.008	0.001

Since the sample sizes A and B are equal then if the two processes have the same fractions defective we can expect the number of defectives found in them to be about the same. For a total of $a + b = 10$ defectives in the two samples we would expect to find $a = b = 5$, or perhaps $a = 4$ and $b = 6$, or $a = 3$ and $b = 7$, and so on. But we don't expect to get something like $a = 0$ and $b = 10$. These cases $(a = 0, 1, \dots)$ would suggest that the two processes actually have different fractions defective, and specifically that $p_1 < p_2$.

The exact p value for the test is given by:

$$\begin{aligned} p &= \sum_{x=0}^{2} \frac{\binom{100}{x}\binom{100}{10-x}}{\binom{200}{10}} \\ &= 0.0008 + 0.0085 + 0.0410 \\ &= 0.0503 \end{aligned}$$

Since $(p = 0.0503) > (\alpha = 0.05)$ there is insufficient evidence to conclude that $p_1 < p_2$. Barely. If it is critical to get the correct answer then another larger pair of samples should be drawn from the two processes.



Example: Reanalyze the example problem using MINITAB **Stat**> **Basic Statistics**> **2 Proportions**.

Solution:

Test and CI for Two Propor...

WORKSHEET 1

Test and CI for Two Proportions

Method

p_1: proportion where Sample 1 = Event
p_2: proportion where Sample 2 = Event
Difference: $p_1 - p_2$

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	100	2	0.020000
Sample 2	100	8	0.080000

Estimation for Difference

Difference	95% Upper Bound for Difference
-0.06	-0.009785

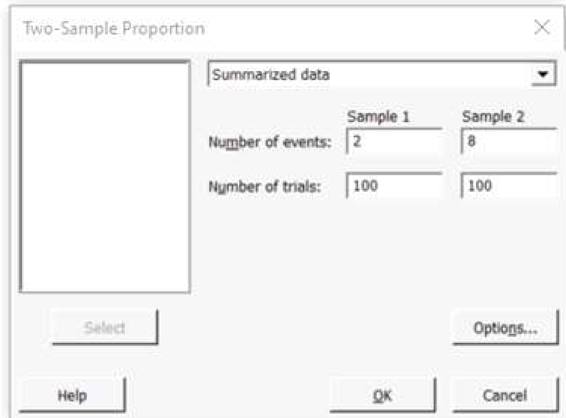
CI based on normal approximation

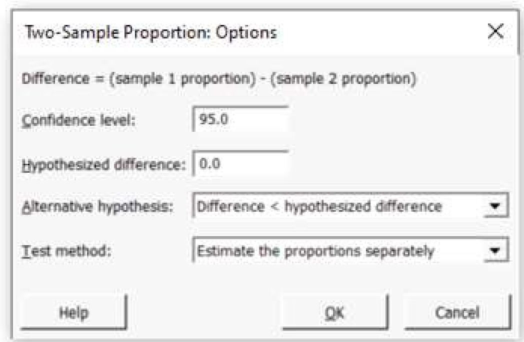
Test

Null hypothesis	H_0: $p_1 - p_2 = 0$
Alternative hypothesis	H_1: $p_1 - p_2 < 0$

Method	Z-Value	P-Value
Normal approximation	-1.97	0.0247
Fisher's exact		0.0503

The normal approximation may be inaccurate for small samples.





Normal Approximation Test for Two Proportions

Hypotheses:

$H_0 : p_1 = p_2$ versus $H_A : p_1 < p_2$

Assumptions: The populations being sampled are binomially distributed and may be approximated with the normal distribution.

Procedure:

1. Draw random samples of size A from population 1 and B from population 2. The samples should be approxiamately the same size.
2. Inspect the samples of size A and B and determine the number of defectives in each sample, a and b, respectively. By arbitrary choice, $a \leq b$.

3. Determine the fractions defective for the two samples from:

$$\widehat{p}_1 = \frac{a}{A} \text{ and } \widehat{p}_2 = \frac{b}{B}$$

4. Estimate the common population fraction defective from:

$$\widehat{p} = \frac{a+b}{A+B}$$

5. Calculate the z value for the test from:

$$z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})(\frac{1}{A} + \frac{1}{B})}}$$

6. If $z < -z_\alpha$ then reject H_0 and conclude that there is evidence that $p_1 < p_2$.

Example: The fractions defective of two processes are to be compared. A sample of $A = 300$ parts is found to have $a = 20$ defectives and a sample of $B = 200$ parts is found to have $b = 30$ defectives. Is there evidence that the fraction defective of the first process (from which A was drawn) is smaller than the fraction defective of the second process (from which B was drawn)?

Solution: The conditions for binomial sampling and the normal approximation appear to be met. The sample fractions defective are:

$$\hat{p}_1 = \frac{20}{300} = 0.067 \quad \text{and} \quad \hat{p}_2 = \frac{30}{200} = 0.150$$

The estimate for the common fraction defective is:

$$\hat{p} = \frac{20 + 30}{300 + 200} = 0.100$$

The z statistic is:

$$\begin{aligned} z &= \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{A}+\frac{1}{B}\right)}} \\ &= \frac{0.067 - 0.150}{\sqrt{(0.100)(1-0.100)\left(\frac{1}{300}+\frac{1}{200}\right)}} \\ &= -3.03 \end{aligned}$$

Since $(z = -3.03) < (-z_{0.01} = -2.32)$ we can conclude that the fraction defective of population 1 is less than that of population 2.

Example: Reanalyze the example problem using MINITAB **Stat**> **Basic Statistics**> **2 Proportions**.

Solution:

Test and CI for Two Propor...

WORKSHEET 1

Test and CI for Two Proportions

Method

p_1: proportion where Sample 1 = Event
p_2: proportion where Sample 2 = Event
Difference: $p_1 - p_2$

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	300	20	0.066667
Sample 2	200	30	0.150000

Estimation for Difference

Difference	95% Upper Bound for Difference
-0.0833333	-0.035522

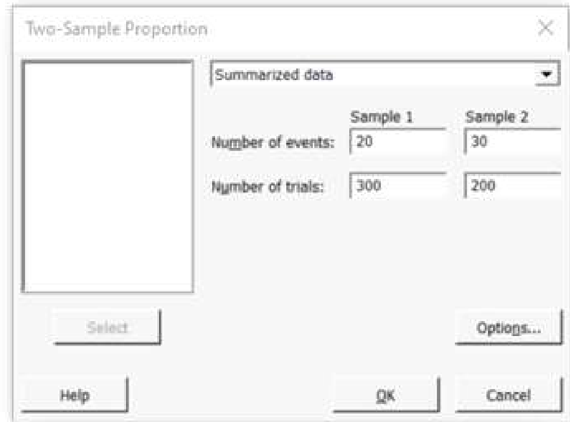
CI based on normal approximation

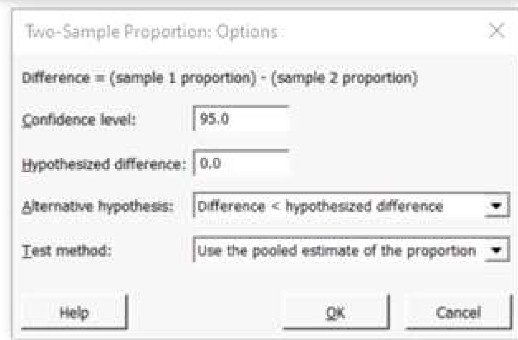
Test

Null hypothesis H_0: $p_1 - p_2 = 0$
Alternative hypothesis H_1: $p_1 - p_2 < 0$

Method	Z-Value	P-Value
Normal approximation	-3.04	0.001
Fisher's exact		0.002

The test based on the normal approximation uses the pooled estimate of the proportion (0.1).





Chisquare Test for Association

Chisquare Tests for Class Frequencies

When data are sorted into classes or categories in a one-way or two-way classification table such that the observed class frequencies (o_i) can be compared to the expected class frequencies (e_i), their agreement can be measured by:

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i} = \sum \frac{o_i^2}{e_i} - n$$

The decision criteria for: H_0*: the model used to calculate the expected frequencies matches the data* vs. H_A*: the model used to calculate the expected frequencies does not match the data* are:

$$\text{Accept } H_0: \quad \chi^2 < \chi^2_{1-\alpha,\nu}$$

$$\text{Reject } H_0: \quad \chi^2 > \chi^2_{1-\alpha,\nu}$$

where ν is the degrees a freedom for the test and is related to the number of categories. Categories must be designed or combined so that $e_i \geq 5$.

Chisquare Tests for Assocation

The chisquare test for class frequencies can be applied to a two-way classification table to test the hypotheses:

$$H_0 : \textit{The two variables are independent}$$

$$H_A : \textit{The two variables are dependent}$$

The condition defined by H_0 is often referred to as *homogeneity*.

Chisquare Tests for Assocation

Example: A company wants to determine if the product quality of three vendors is different. They inspect samples of the three vendors' products and find the following defect counts:

Vendor	Major	Minor	Good
A	12	23	89
B	8	12	62
C	21	30	119

Do the data indicate that there are any significant differences between the three vendors with respect to their product quality?

Solution: The hypotheses to be tested are H_0: *all of the vendors have the same defect rate* versus H_A : *one or more of the vendors has a different defect rate from the others.*

It is necessary to determine the expected frequency that corresponds to each observed frequency. The expected frequency for the first case (A-Major) is given by:

$$e_1 = \left(\frac{124}{376}\right)\left(\frac{41}{376}\right)376 = 13.5$$

and the other expected frequencies are calculated in the same manner:

$$e_2 = \left(\frac{82}{376}\right)\left(\frac{41}{376}\right)376 = 8.94$$

$$\vdots$$

$$e_9 = \left(\frac{170}{376}\right)\left(\frac{270}{376}\right)376 = 122.1$$

The following table shows the observed and expected frequencies (o_i/e_i) for each case:

Vendor	Major	Minor	Good	Row Totals
A	12/13.5	23/21.4	89/89.0	124
B	8/8.94	12/14.2	62/58.9	82
C	21/18.5	30/29.4	119/122.1	170
Column Totals:	41	65	270	376

In order to detect a difference between the vendors we must calculate the χ^2 statistic:

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = \sum \frac{o_i^2}{e_i} - n$$

For the data from the Problem we find:

$$\begin{aligned}
\chi^2 &= \frac{12^2}{13.5} + \frac{8^2}{8.94} + \cdots + \frac{119^2}{122.1} - 376 \\
&= 377.3 - 376 \\
&= 1.30
\end{aligned}$$

The χ^2 statistic follows a χ^2 distribution where the number of degrees of freedom is determined from the number of rows (n) and columns (m) in the data set:

$$\begin{aligned}
df &= (n-1)(m-1) \\
&= (3-1)(3-1) \\
&= 4
\end{aligned}$$

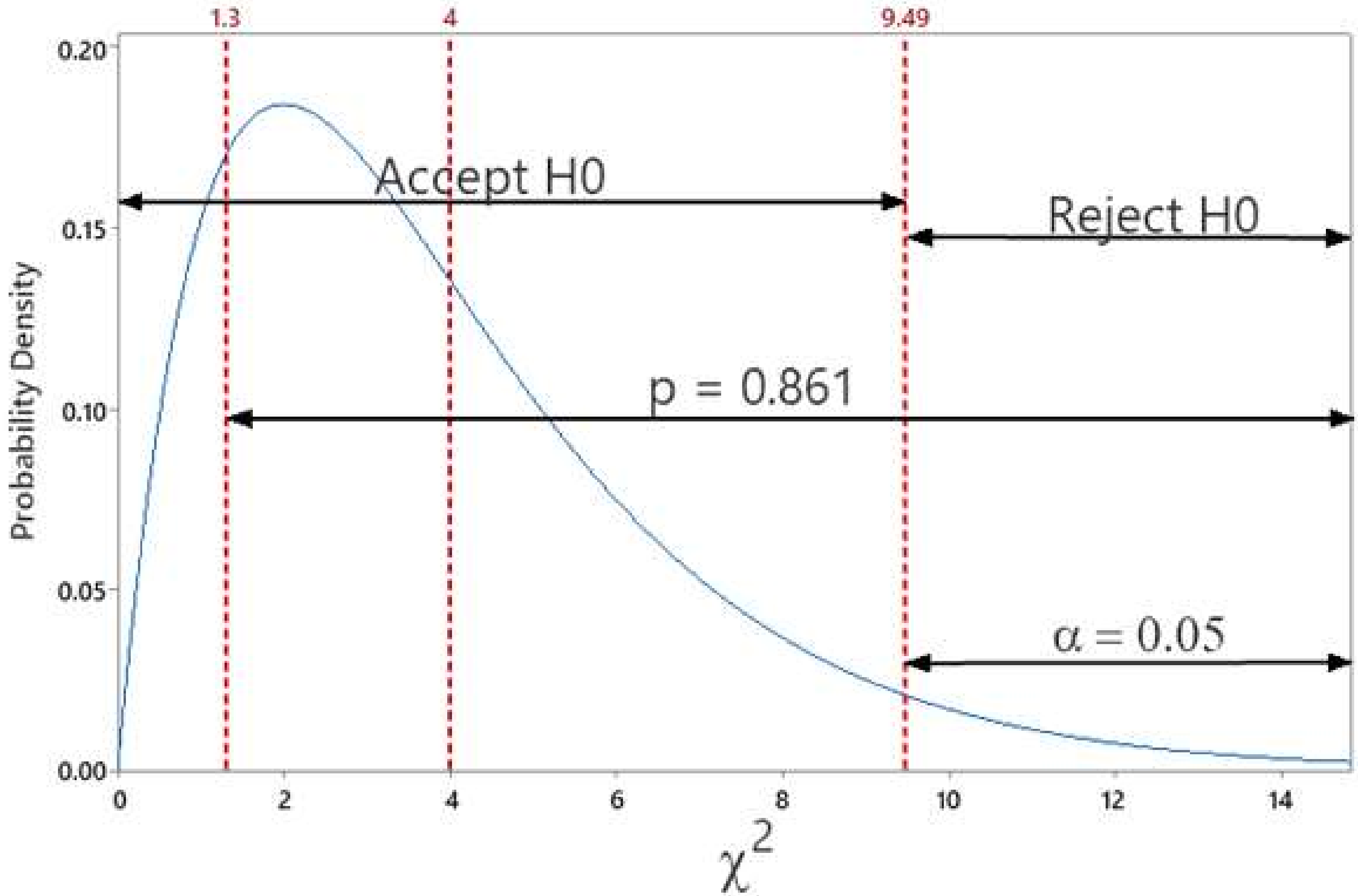
From the χ^2 table the critical value of the statistic with $\alpha = 0.05$ is:

$$\chi^2_{0.95} = 9.488$$

The p value for the test is

$$P(1.3 < \chi^2 < \infty; df = 4) = 0.861$$

We can clearly see that the test statistic $\chi^2 = 1.4$ falls within the accept region of the null hypothesis.


1.3
4
9.49
Accept H0
Reject H0
p = 0.861
α = 0.05
Probability Density
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10
12
14
χ2

Solution by Minitab

Using MINITAB **Stat**> **Tables**> **Cross Tabulation and Chi-square**:

```
MTB > print c1-c4
```

Data Display

```
 Row  Vendor  Major  Minor   Good

   1      A      12     23     89
   2      B       8     12     62
   3      C      21     30    119

MTB > ChiSquare c2-c4.
```

Chi-Square Test: Major, Minor, Good

```
Expected counts are printed below observed counts

         Major    Minor     Good    Total
    1       12       23       89      124
         13.52    21.44    89.04

    2        8       12       62       82
          8.94    14.18    58.88

    3       21       30      119      170
         18.54    29.39   122.07

Total       41       65      270      376

Chi-Sq =  0.171 +  0.114 +  0.000 +
          0.099 +  0.334 +  0.165 +
          0.327 +  0.013 +  0.077 = 1.301
DF = 4, P-Value = 0.861
```

Chisquare Test for 2×2 Tables

- A special case of the chisquare test for association is the 2×2 table
- The method is usually applied with a correction by Yates's method
- This method is considered to be archaic today because with good software we can use Fisher's Exact Test for two proportions instead
- The chisquare test method requires larger sample sizes than Fisher's Exact Test to meet the $e_{ij} \geq 5$ requirement for each cell in the 2×2 table
- When the sample sizes are sufficiently large, the p value for the chisquare test approaches the p value for Fisher's Exact Test

Chisquare Goodness of Fit Test

Chisquare Goodness of Fit Test

- The χ^2 goodness of fit (GOF) test is used to determine if observed sample frequencies (such as from a histogram) follow a specified distribution with specified parameter values.
- The parameter values may be stated independent of the data or may be estimated from the data with an associated penalty.
- The χ^2 GOF test requires at least 50 observations and the classes should be combined so that the minimum expected class frequency is $e_i \geq 5$.
- The χ^2 GOF test is not as sensitive as other methods for specific distributions; such as the Anderson-Darling test for normality
- The χ^2 GOF test is still very general and is used frequently
- The GOF test is also referred to as a lack of fit (LOF) test.

Chisquare Test for Class Frequencies

When data are sorted into classes or categories in a one-way or two-way classification table such that the observed class frequencies (o_i) can be compared to the expected class frequencies (e_i), their agreement can be measured by:

$$\chi^2 = \sum_i \frac{(o_i - e_i)^2}{e_i} = \sum \frac{o_i^2}{e_i} - n$$

The decision criteria for: H_0*: the model used to calculate the expected frequencies matches the data* vs. H_A*: the model used to calculate the expected frequencies does not match the data* are:

$$\text{Accept } H_0\text{:} \quad \chi^2 < \chi^2_{1-\alpha,\nu}$$

$$\text{Reject } H_0\text{:} \quad \chi^2 > \chi^2_{1-\alpha,\nu}$$

where ν is the degrees a freedom for the test and is related to the number of categories. Categories must be designed or combined so that $e_i \geq 5$.

Example: We need to determine if the number of car door paint defects follows a Poisson distribution with $\lambda = 2.4$. The defect count inspection data from 300 car doors are shown below. Are the data consistent with a Poisson distribution with $\lambda = 2.4$?

#Defects	*Frequency*
0	19
1	48
2	66
3	74
4	44
5	35
6	10
7	4

Solution: The hypotheses to be tested are:

$$H_0 : \quad \textit{The distribution is Poisson with } \lambda = 2.4$$

$$H_A : \quad \textit{The distribution is not Poisson with } \lambda = 2.4$$

The expected frequencies are given by the Poisson distribution:

$$e_i = n \times Poisson(x_i; \lambda)$$

For example, for the case of $x = 0$ defects:

$$\begin{aligned} e_1 &= 300 \times Poisson(x = 0; \lambda = 2.4) \\ &= 300 \times 0.091 \\ &= 27.2 \end{aligned}$$

The observed and expected frequencies for all of the classes are shown in the following table. Notice that the last two classes were merged to form one open class with $o_i \geq 6$. This was necessary to give the class the required 5 or more observations.

#Defects	o_i	p_i	e_i
0	19	0.091	27.2
1	48	0.218	65.3
2	66	0.261	78.4
3	74	0.209	62.7
4	44	0.125	37.6
5	35	0.060	18.1
≥ 6	14	0.036	10.8
Totals :	300	1.000	300

The χ^2 statistic for the data is:

$$\begin{aligned}\chi^2 &= \frac{19^2}{27.2} + \frac{48^2}{65.3} + \frac{66^2}{78.4} + \frac{74^2}{62.7} + \frac{44^2}{37.6} + \frac{35^2}{18.1} + \frac{14^2}{10.8} - 300 \\ &= 28.8\end{aligned}$$

The degrees of freedom are:

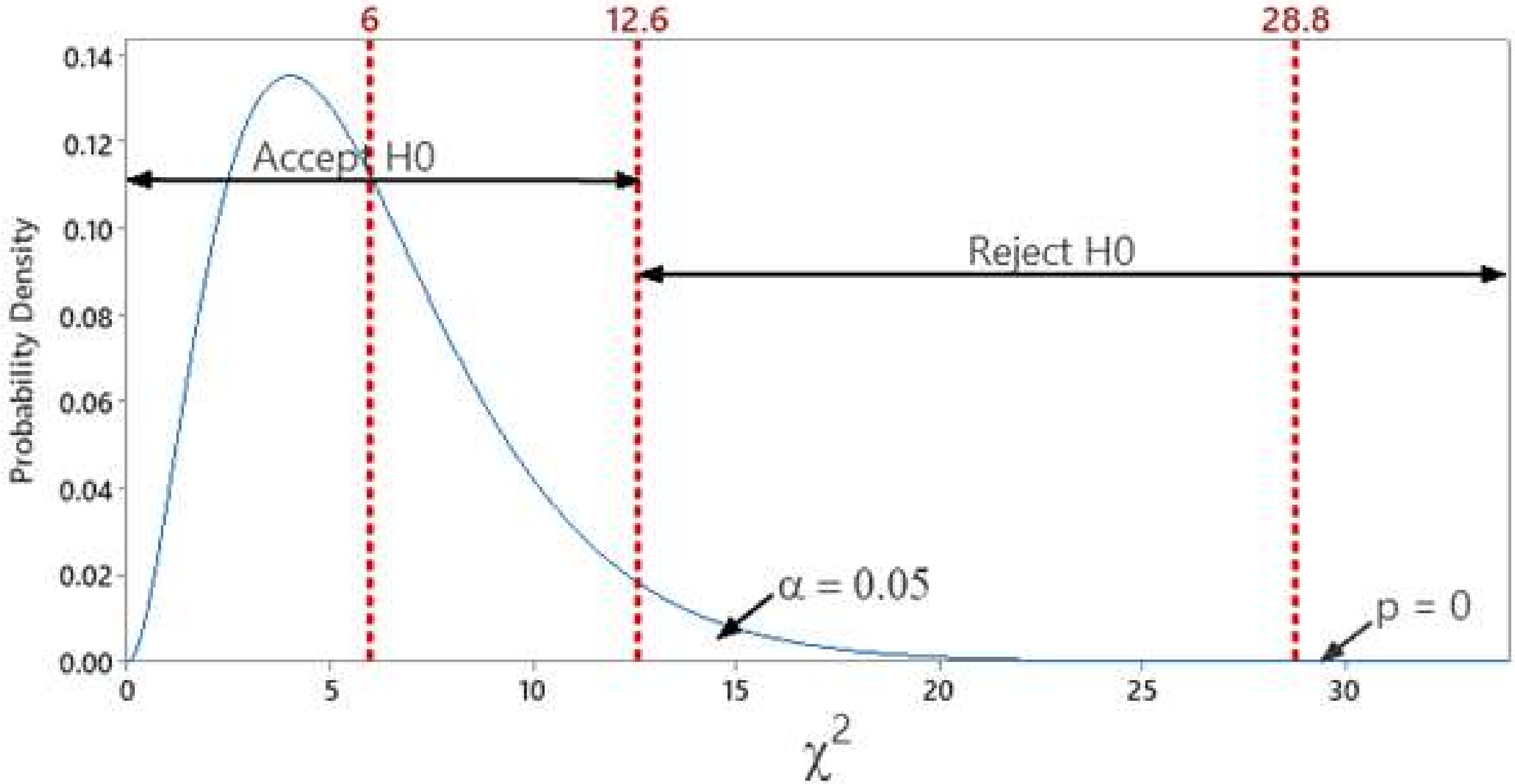
$$\begin{aligned}df &= k - m - 1 \\ &= 7 - 0 - 1 \\ &= 6\end{aligned}$$

where k is the number of classes and m is the number of parameters estimated from the data set to form the hypotheses. The extra -1 is lost because the grand total must be determined.

From the χ^2 table we find the critical value is:

$$\chi^2_{0.95,6} = 12.6$$

We can clearly see that $\chi^2 = 28.8$ exceeds the critical value of $\chi^2_{0.95,6} = 12.6$, therefore we must reject H_o. We can conclude that either the distribution is not Poisson or that the distribution mean is not $\lambda = 2.4$.



Determining m in the χ^2 Goodness of Fit Test:

The value of m in the χ^2 goodness of fit test frequently takes on the values $m = 0, 1,$ or 2.

- $m = 0$ when the distribution shape, location, and dispersion are all specified before data are collected.
- $m = 1$ when one parameter (e.g. μ, σ, p, or λ) is estimated from the sample data.
- $m = 2$ when two parameters (e.g. μ and σ) are estimated from the sample data.

Example: Use the χ^2 goodness of fit test to determine if the following data set (in stem and leaf plot format) comes from a normal population:

```
34| 1
35| 788
36| 2467
37| 07
38| 22789
39| 11233447779
40| 112222233347899
41| 2344555679
42| 001257
43| 022
```

There are 60 observations in the data set. The requirement that there are at least 5 observations in each class demands that we regroup the values. The regrouped class frequencies become:

```
340 to 369|  8
370 to 389|  7
390 to 399| 11
400 to 409| 15
410 to 419| 10
420 to 439|  9
```

The mean of the data set is $\bar{x} = 398.7$ and the standard deviation is $s = 20.7$. Using the sample statistics to estimate μ and σ, the expected frequency for the class of $-\infty < x < 369.5$ is:

$$\begin{aligned} e_1 &= 60 \times \Phi(-\infty < x < 369.5; 398.7, 20.7) \\ &= 60 \times 0.0792 \\ &= 4.74 \end{aligned}$$

The expected frequency for the class of $369.5 < x < 389.5$ is:

$$\begin{aligned} e_2 &= 60 \times \Phi(369.5 < x < 389.5; 398.7, 20.7) \\ &= 60 \times 0.2492 \\ &= 14.95 \end{aligned}$$

The other expected class frequencies are calculated in the same manner.

The following table summarizes the observed and expected class frequencies and each class's contribution to the χ^2 statistic:

Class	$\Phi(a < x < b)$	e_i	o_i	$\frac{(o_i - e_i)^2}{e_i}$
$-\infty < x < 369.5$	0.0792	4.75	8	2.22
$369.5 < x < 389.5$	0.2492	14.95	7	4.23
$389.5 < x < 399.5$	0.1870	11.22	11	0.00
$399.5 < x < 409.5$	0.1837	11.02	15	1.43
$409.5 < x < 419.5$	0.1434	8.60	10	0.23
$419.5 < x < \infty$	0.1575	9.45	9	0.02
$-\infty < x < \infty$	1.000	59.99	60.0	8.13

The test statistic is:

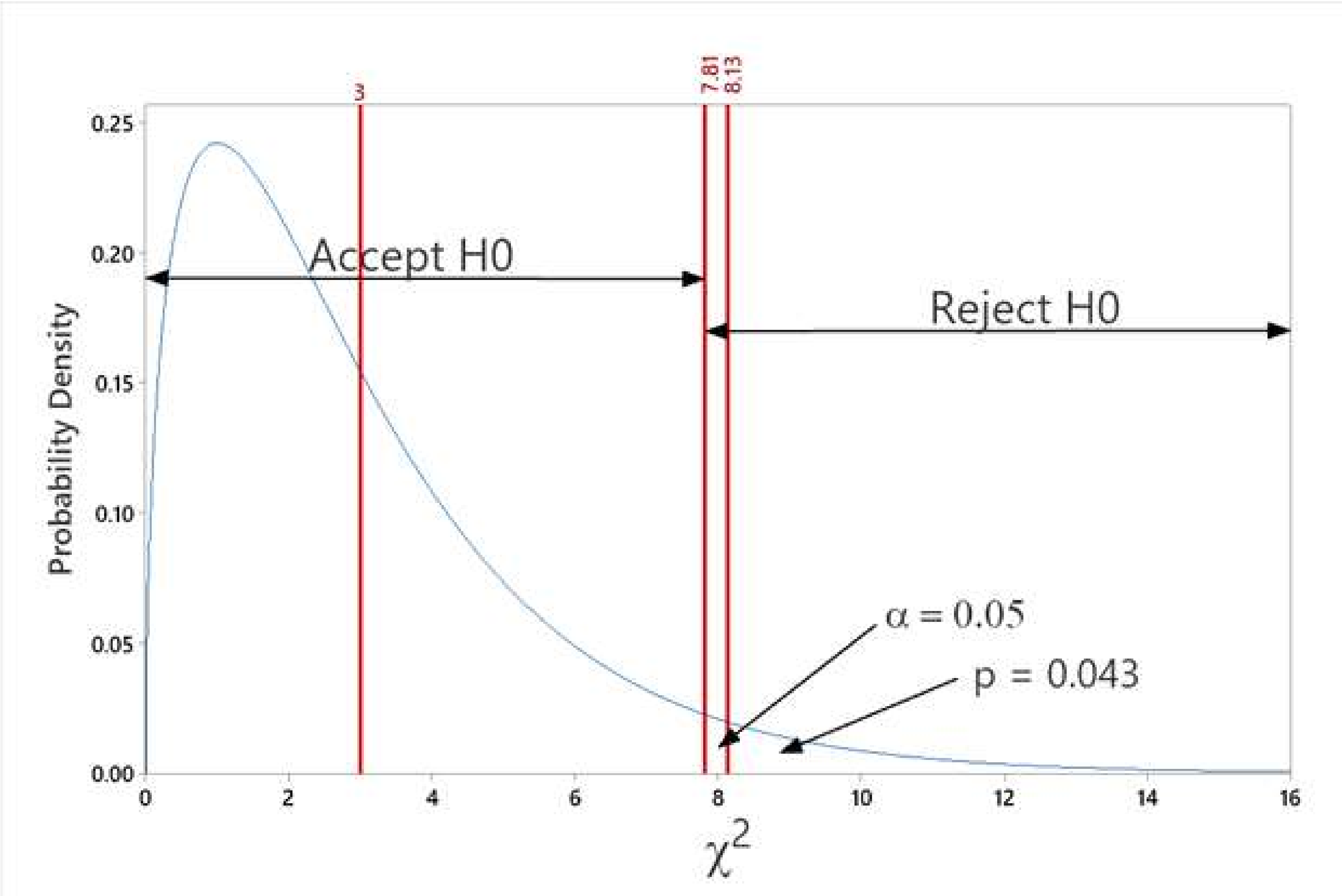
$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i} = 8.13$$

There were $m = 2$ parameters estimated from the sample data (μ and σ) so the χ^2 distribution has:

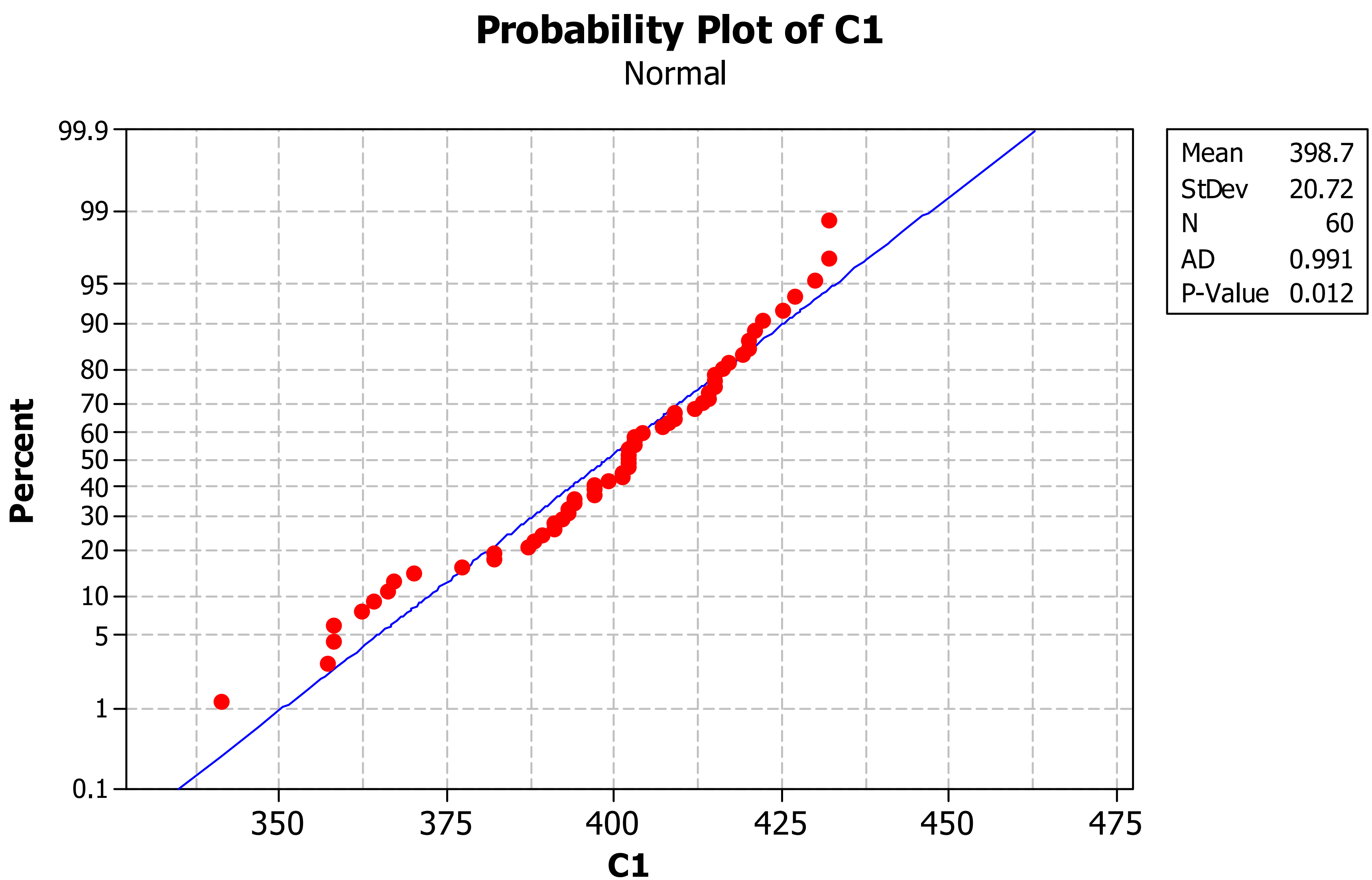
$$\begin{aligned} df &= k - m - 1 \\ &= 6 - 2 - 1 \\ &= 3 \end{aligned}$$

degrees of freedom and the critical value of χ^2 is $\chi^2_{0.95,3} = 7.81$. Since $(\chi^2 = 8.13) > (\chi^2_{0.95,3} = 7.81)$ we must reject the claim that the population being sampled is normally distributed. The test's p value is

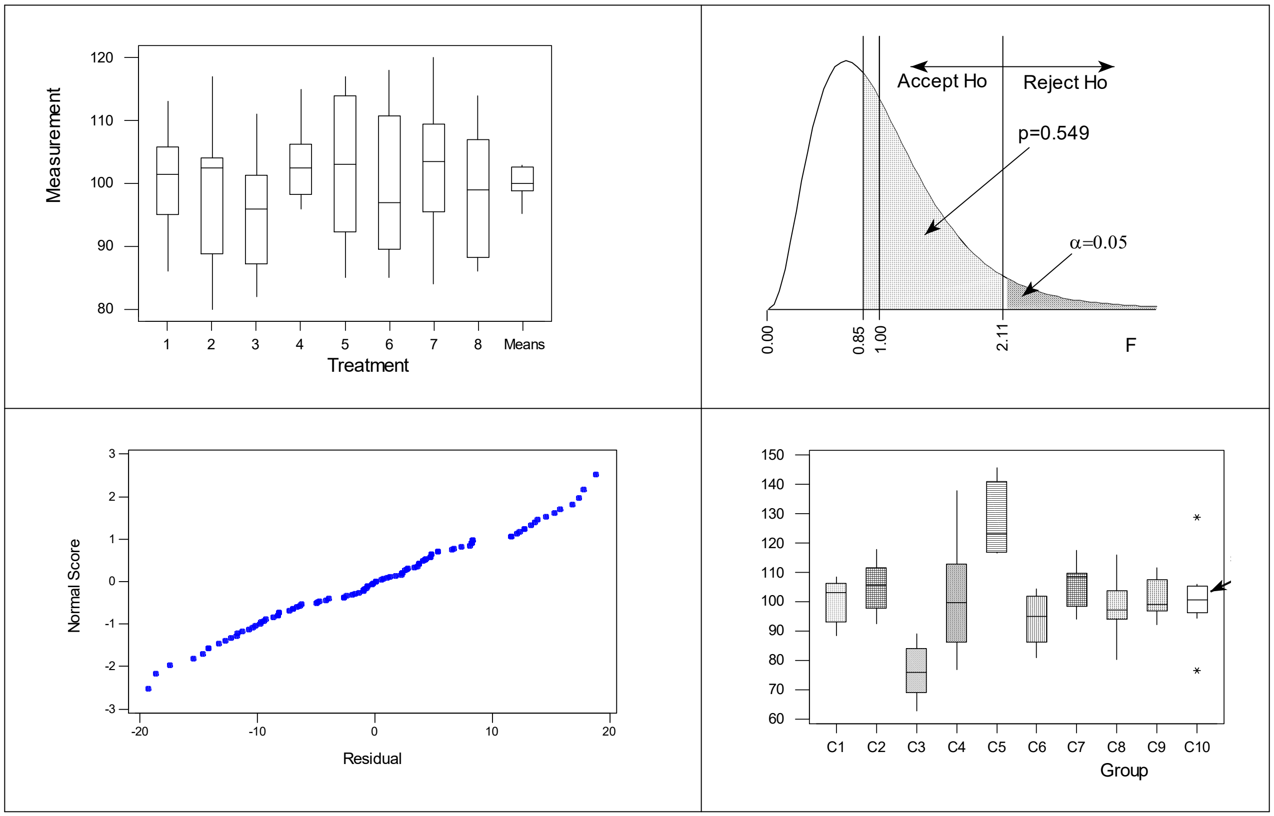
$$\begin{aligned} p &= P(8.13 < \chi^2 < \infty; df = 3) \\ &= 0.043 \end{aligned}$$


3
7.81
8.13
0.25
0.20
0.15
0.10
0.05
0.00
Probability Density
Accept H0
Reject H0
$\alpha = 0.05$
p = 0.043
0
2
4
6
8
10
12
14
16
χ^2

The chisquare GOF test's result agrees with the Anderson-Darling test which is actually preferred:

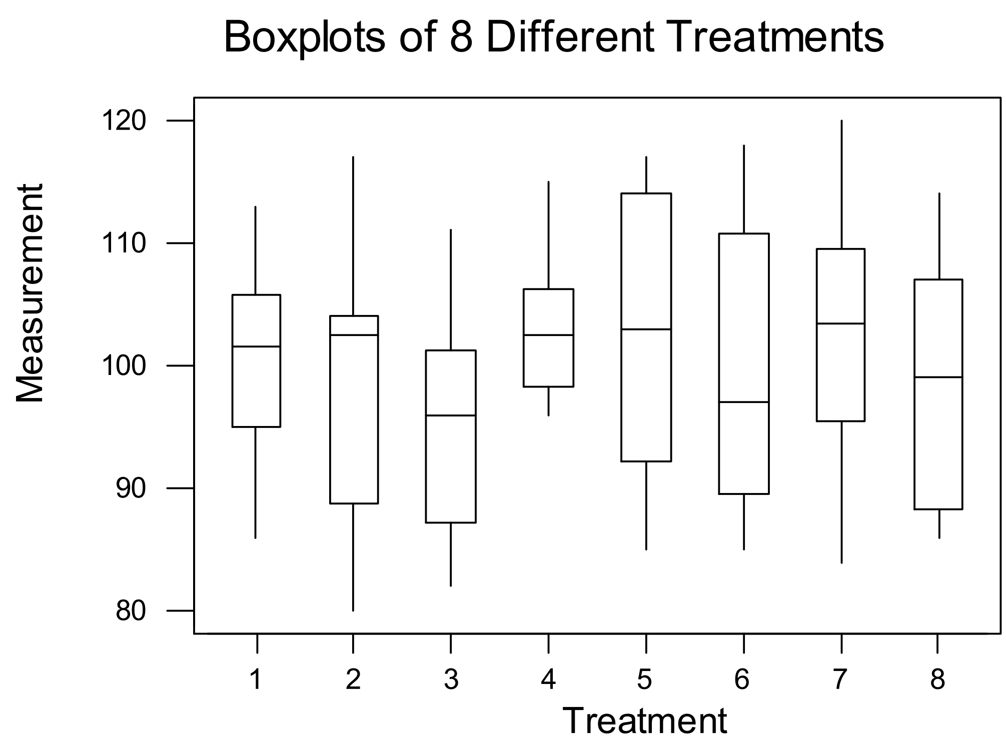


Introduction to Analysis of Variance



The Case for ANOVA

Suppose that $k = 8$ different treatments are to be tested for differences in their means. We could do hypothesis tests of all possible pairs of treatments using two sample z or t tests or the boxplot slippage tests, but there are a total of $\binom{8}{2} = 28$ pairwise tests that must be performed.



The Case for ANOVA

If a single test has an error rate of $\alpha = 0.05$ then the overall error rate will be determined from:

$$1 - \alpha_{total} = (1 - 0.05)^{28}$$

so

$$\alpha_{total} = 0.762$$

This means that the probability of committing one or more type 1 errors in the 28 pairwise tests is 76%. This discourages the use of pairwise tests in these situations. A single test that is capable of considering all 8 treatments simultaneously is required.

The ANOVA Rationale

Hypotheses:

$$H_0 \; : \; \mu_i = \mu_j \text{ for all pairs of treatments}$$

$$H_A \; : \; \mu_i \neq \mu_j \text{ for at least one pair of treatments}$$

Assumptions:

- All k treatments have the same variance (σ_y^2).
- All k treatments are normally distributed.
- All observations are independent.

The ANOVA Rationale

- Recall that the hypotheses $H_0 : \sigma_1^2 = \sigma_2^2$ versus $H_A : \sigma_1^2 > \sigma_2^2$ can be tested using the F test where:

$$F = \left(\frac{s_1}{s_2}\right)^2$$

  When H_0 is true we get $F \simeq 1$ and when H_0 is false we get $F \gg 1$.

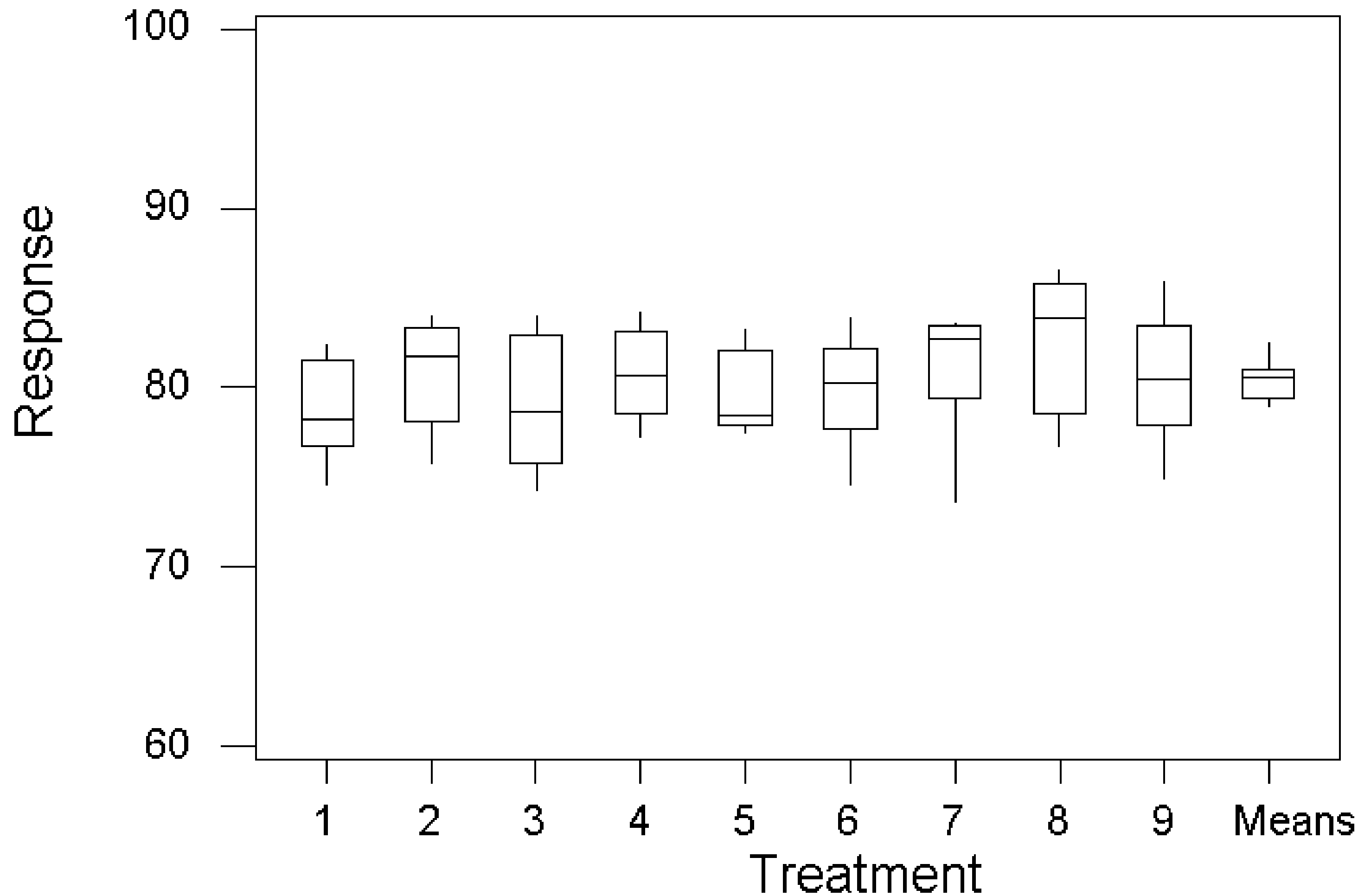
- ANOVA uses the F test to compare two independent estimates of the population standard deviation, one determined from differences between treatments and the other determined from variation within treatments:

$$F = \left(\frac{s_{between}}{s_{within}}\right)^2$$

  $s_{between}$ is constructed so that when H_0 is true, $s_{between} \simeq s_{within}$ so that $F \simeq 1$ and when H_0 is false, $s_{between} \gg s_{within}$ so that $F \gg 1$.

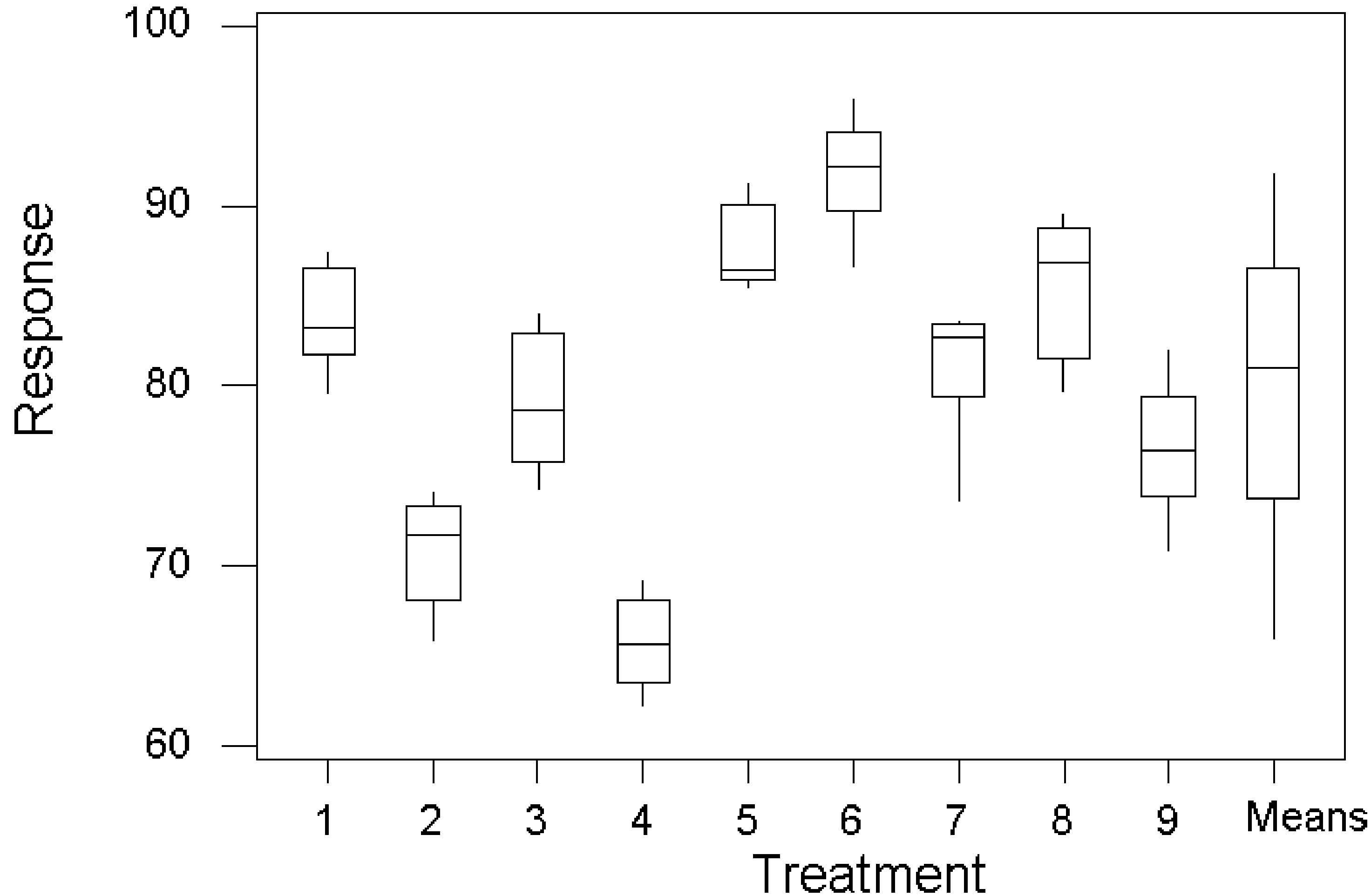
The ANOVA Rationale

If $H_0 : \mu_i = \mu_j$ *for all pairs of treatments* is true, then $\sigma_{\bar{y}} = \sigma_y / \sqrt{n}$:



The ANOVA Rationale

If $H_0 : \mu_i = \mu_j$ *for all pairs of treatments* is false, then $\sigma_{\bar{y}} \gg \sigma_y/\sqrt{n}$:



ANOVA Procedure:

1. Draw random samples of size n from each of the k treatments.
2. Calculate $\bar{y}_j$ and s_j^2 for each treatment.
3. Determine the pooled or error variance from:

$$s_\epsilon^2 = \frac{1}{k}\sum_{j=1}^{k} s_j^2 = \hat{\sigma}_y^2$$

   The error variance is the best estimate of the population variance determined from all of the treatments.

4. Determine $s_{\bar{y}}^2$ from:

$$s_{\bar{y}}^2 = \frac{1}{k-1}\sum_{j=1}^{k}(\bar{y}_j - \bar{\bar{y}})^2 = \frac{\hat{\sigma}_y^2}{n}$$

ANOVA Procedure:

5. Since s_ϵ^2 and $ns_{\bar{y}}^2$ are both estimators for σ_y^2 then their ratio:

$$F = \frac{\hat{\sigma}_y^2}{\hat{\sigma}_y^2} = \frac{ns_{\bar{y}}^2}{s_\epsilon^2}$$

should be close to $F = 1$ if H_0 is true. If one or more of the means are different from the others then $s_{\bar{y}}^2$ will inflate the numerator and F will be significantly larger than $F = 1$.

6. If $F \leq F_\alpha$ with $df_{model} = k - 1$ numerator and $df_\epsilon = k(n-1)$ denominator degrees of freedom then accept H_0. If $F > F_\alpha$ then reject H_0.

7. Check assumptions:

a. The treatments are homoscedastic.

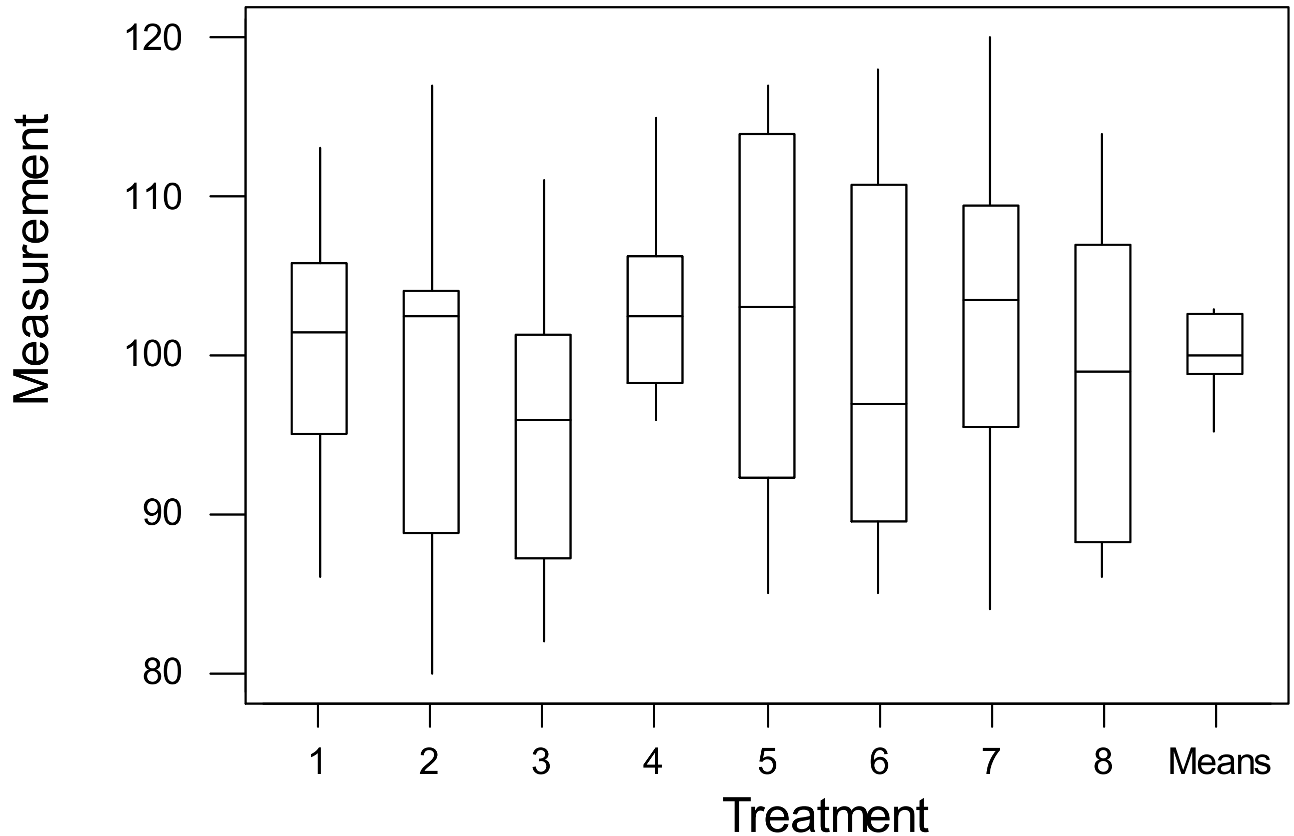
b. The residuals are normal.

c. The observations are independent. (Usually obtained by randomizing the order of the runs.)

Example: Samples of size $n = 12$ are drawn from each of $k = 8$ treatments. The data are shown in the attached table. Perform the ANOVA to determine if the mean of one or more of the treatments is different from the others.

A	B	C	D	E	F	G	H
91	88	99	104	92	85	111	112
105	87	89	98	114	104	94	87
106	104	102	115	93	118	102	94
103	104	95	103	116	104	84	103
94	103	82	96	85	85	100	100
103	91	87	103	106	91	120	86
100	112	111	107	100	89	110	98
113	102	108	96	109	95	106	107
98	80	86	99	114	116	105	89
86	117	88	111	87	99	101	107
109	100	99	101	117	113	91	114

Solution:



- Variation within treatments: $s_{\epsilon}^2 = \frac{1}{k}\sum_{j=1}^{k} s_j^2 = \hat{\sigma}_y^2$
- Variation between treatments: $s_{\bar{\bar{y}}}^2 = \frac{1}{k-1}\sum_{j=1}^{k} (\bar{y}_j - \bar{\bar{y}})^2 = \frac{\hat{\sigma}_y^2}{n}$

	A	B	C	D	E	F	G	H
	91	88	99	104	92	85	111	112
	105	87	89	98	114	104	94	87
	106	104	102	115	93	118	102	94
	103	104	95	103	116	104	84	103
	94	103	82	96	85	85	100	100
	103	91	87	103	106	91	120	86
	100	112	111	107	100	89	110	98
	113	102	108	96	109	95	106	107
	98	80	86	99	114	116	105	89
	86	117	88	111	87	99	101	107
	109	100	99	101	117	113	91	114
	100	103	97	102	96	91	108	88
x-bar	100.67	99.25	95.25	102.92	102.42	99.17	102.67	98.75
s	7.63	10.76	9.09	5.79	11.70	11.80	9.72	9.99

- Variation within treatments: $s_\epsilon^2 = \frac{1}{k}\sum_{j=1}^{k} s_j^2 = \widehat{\sigma}_y^2$
- Variation between treatments: $s_{\bar{y}}^2 = \frac{1}{k-1}\sum_{j=1}^{k}(\bar{y}_j - \bar{\bar{y}})^2 = \frac{\widehat{\sigma}_y^2}{n}$

From the table the error variance is:

$$s_\epsilon^2 = \frac{1}{8}(7.63^2 + 10.76^2 + 9.09^2 + \ldots) = 95.1$$

The grand mean is:

$$\bar{\bar{y}} = \frac{1}{8}(100.67 + 99.25 + 102.92 + \ldots) = 100.1$$

The variance of the $\bar{y}$s is:

$$s_{\bar{\bar{y}}}^2 = \left(\frac{1}{k-1}\right)\sum_{i=1}^{k}(\bar{y}_i - \bar{\bar{y}})^2$$

$$= \left(\frac{1}{8-1}\right)\left((100.67 - 100.1)^2 + (99.25 - 100.1)^2 + \ldots\right) = 6.76$$

The F ratio is:

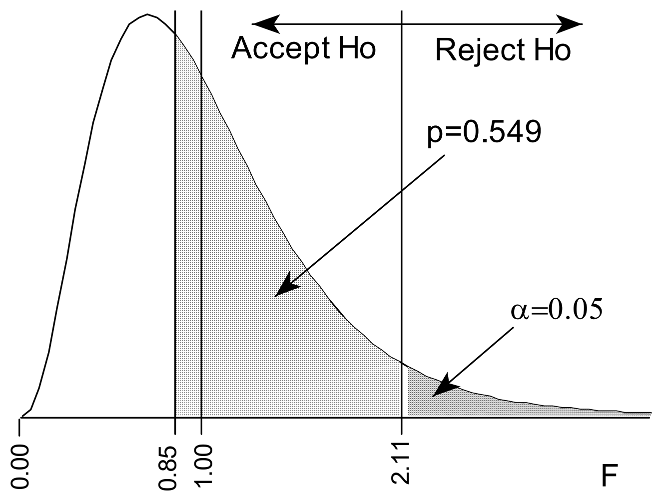
$$F = \frac{ns_{\bar{y}}^2}{s_\epsilon^2}$$

$$= \frac{12\times 6.75}{95.1} = \frac{81.0}{95.1} = 0.85$$

The critical value of F that defines the accept-reject bound has $8-1=7$ numerator degrees of freedom and $8(12-1)=88$ denominator degrees of freedom. From a look-up table we have $F_{0.95,7,88}=2.11$. Since $(F=0.85)<(F_{0.95}=2.11)$ we must accept H_0 and conclude that there are no differences among the treatment means.

One-way Analysis of Variance

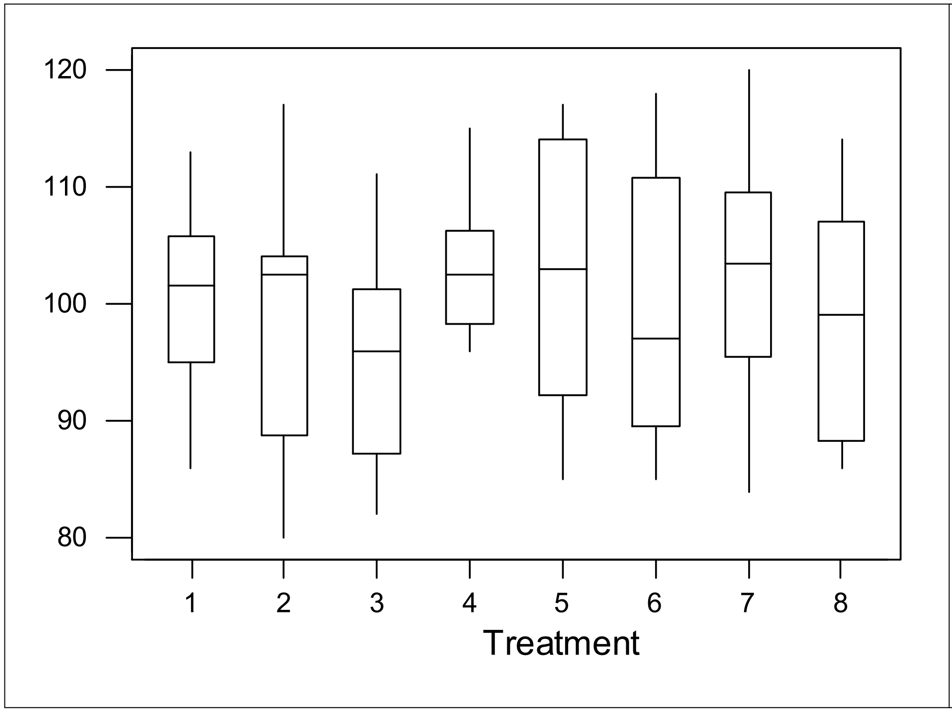
```
Source  DF      SS    MS     F      P
Factor   7   565.7  80.8  0.85  0.549
Error   88  8367.6  95.1
Total   95  8933.2

S = 9.751   R-Sq = 6.33%   R-Sq(adj) = 0.00%


                           Individual 95% CIs For Mean Based on Pooled StDev
Level   N    Mean  StDev    -+---------+---------+---------+--------
A      12  100.67   7.63               (----------*-----------)
B      12   99.25  10.76            (-----------*----------)
C      12   95.25   9.09    (-----------*----------)
D      12  102.92   5.79                    (----------*----------)
E      12  102.42  11.70                   (----------*----------)
F      12   99.17  11.80            (----------*-----------)
G      12  102.67   9.72                   (----------*-----------)
H      12   98.75   9.99           (-----------*----------)
                            -+---------+---------+---------+--------
                           90.0      95.0     100.0     105.0

Pooled StDev = 9.75
```

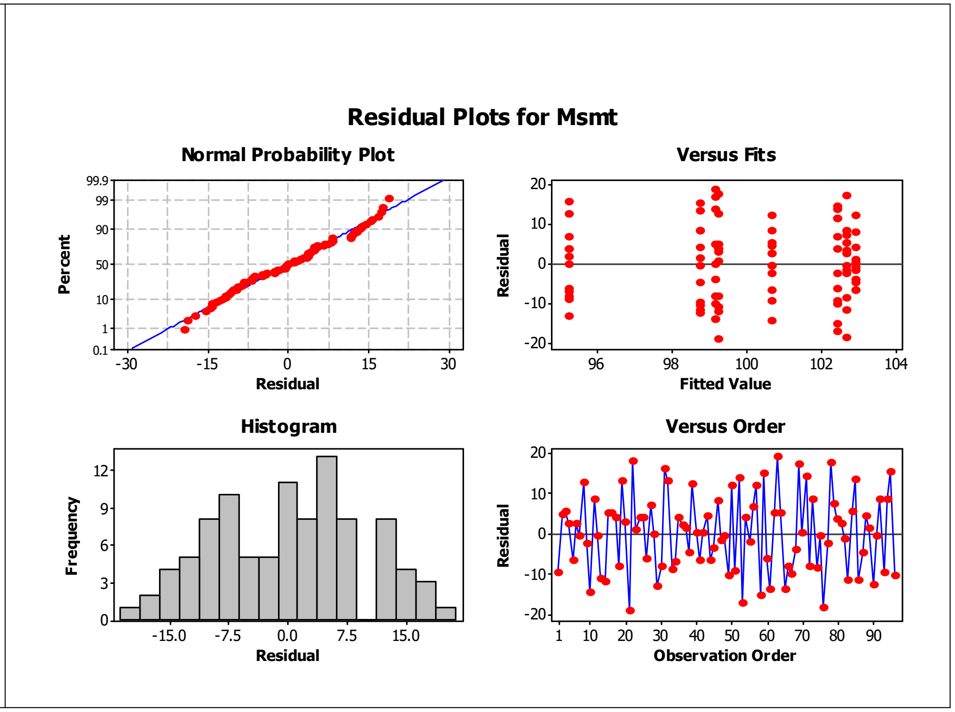


ANOVA Assumptions

- ANOVA assumptions can be checked graphically.
- The distributions of the model residuals ϵ_{ij} are normal:
  - Boxplot
    - Are the boxplots symmetric?
    - Are whisker lengths in correct proportion to the IQRs?
  - Normal plot of the residuals
- The residuals are homoscedastic, i.e. constant for all of the treatments:
  - Boxplot - Are the ranges and IQRs comparable?
  - Residuals versus fits plot
  - Normal plot of the residuals
- The residuals are independent:
  - Run chart of residuals in time order
  - Lag-one plot: ϵ_i vs. ϵ_{i-1}

ANOVA Assumptions

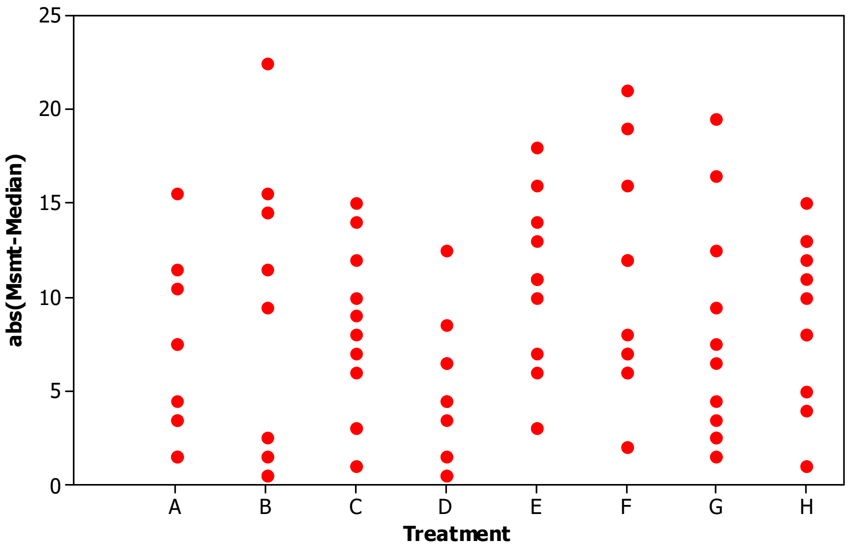




Levene's Test

- Levene's test is a nonparametric test for homoscedasticity (an assumption of ANOVA) that makes use of ANOVA.
- Hypotheses: H_0*: the treatments are homoscedastic* versus H_A*: the treatments are heteroscedastic*.
- Procedure:
  1. Median adjust the observations: $y'_{ij} = y_{ij} - \tilde{y}_i$
  2. Calculate the absolute values of the y'_{ij}, i.e. $\left|y'_{ij}\right|$
  3. Perform ANOVA on the $\left|y'_{ij}\right|$.
  4. Reject H_0 if the ANOVA F statistic's p value is less than 0.05.
- Levene's test is much preferred over other methods (e.g. the F test for two variances or Bartlett's test) because it is much more immune to deviations from normality.

Levene's Test

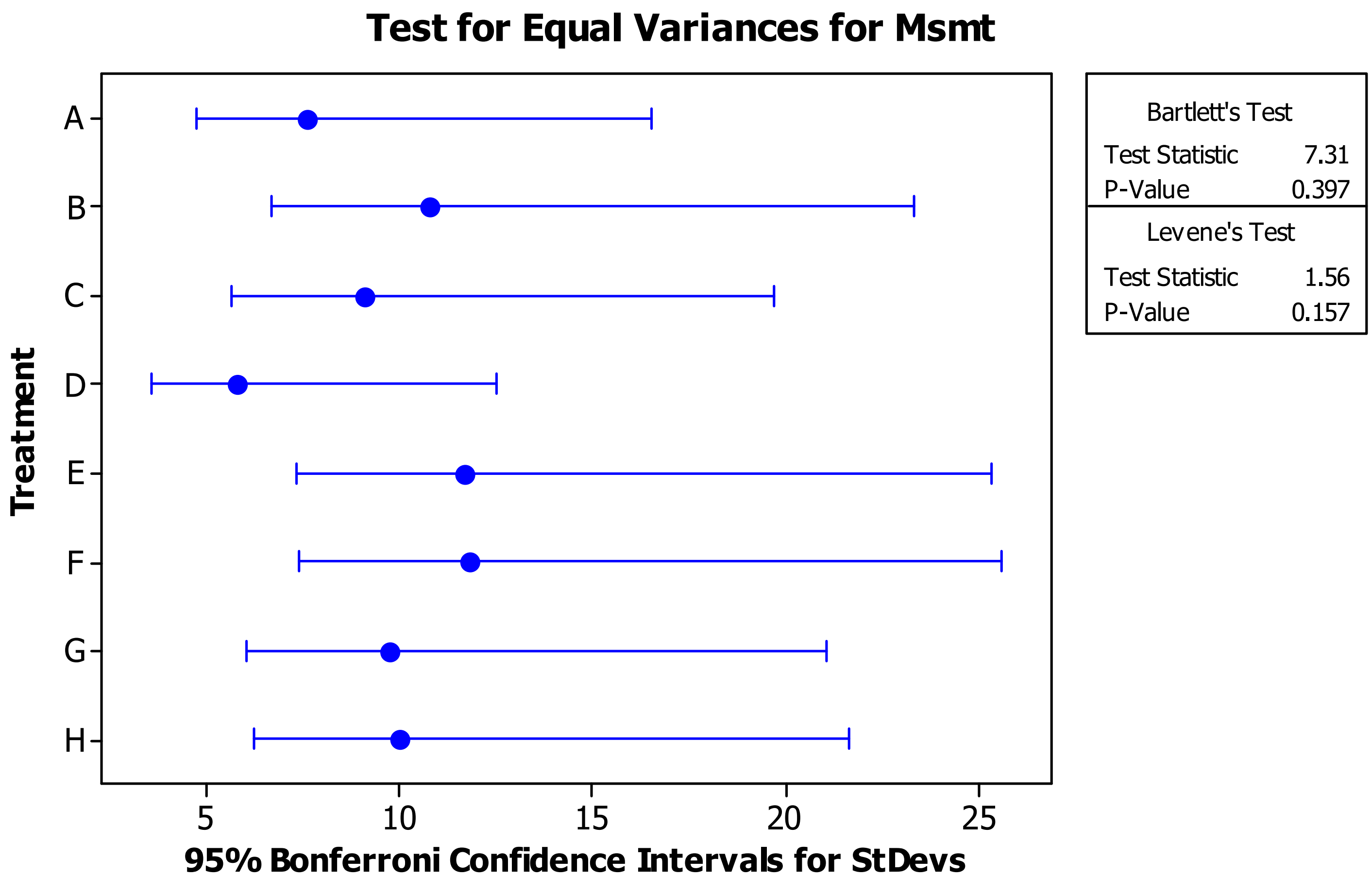


General Linear Model: abs(Msmt-Median) versus Treatment

```
Factor      Type   Levels  Values
Treatment   fixed       8  A, B, C, D, E, F, G, H


Analysis of Variance for abs(Msmt-Median),
using Adjusted SS for Tests

Source      DF    Seq SS    Adj SS   Adj MS      F      P
Treatment    7    326.41    326.41    46.63   1.56  0.157
Error       88   2627.25   2627.25    29.86
Total       95   2953.66
```

Levene's Test

Perform Levene's test in MINITAB using **Stat**> **ANOVA**> **Test for Equal Variances**:



ANOVA Sums of Squares

ANOVA separates the total variation in the data set into components attributed to different sources. The total amount of variation in the data set is:

$$SS_{total} = \sum_{j=1}^{k} \sum_{i=1}^{n} (y_{ij} - \bar{\bar{y}})^2$$

If the k treatment means are $\bar{y}_1, \bar{y}_2, ..., \bar{y}_k$, that is:

$$\bar{y}_j = \frac{1}{n} \sum_{i=1}^{n} y_{ij}$$

then

$$\begin{aligned} SS_{total} &= \sum_{j=1}^{k} \sum_{i=1}^{n} (y_{ij} - \bar{y}_j + \bar{y}_j - \bar{\bar{y}})^2 \\ &= \sum_{j=1}^{k} \sum_{i=1}^{n} (y_{ij} - \bar{y}_j)^2 + n \sum_{j=1}^{k} (\bar{y}_j - \bar{\bar{y}})^2 \\ &= SS_{\epsilon} + SS_{model} \end{aligned}$$

The degrees of freedom are also partitioned:

$$\begin{aligned} df_{total} &= df_{\mathit{model}} + df_{\epsilon} \\ kn - 1 &= (k-1) + k(n-1) \end{aligned}$$

The required variances, also called mean squares (MS), are given by:

$$MS_{\epsilon} = s^2_{pooled} = \frac{SS_{\epsilon}}{df_{\epsilon}}$$

and

$$MS_{model} = ns^2_{\bar{\bar{y}}} = \frac{SS_{model}}{df_{model}}$$

ANOVA Table

Source	*df*	*SS*	*MS*	*F*
Treatment	$k-1$	$n\sum_{j=1}^{k}(\bar{y}_j-\bar{\bar{y}})^2$	$\frac{SS_{Treatment}}{df_{Treatment}}$	$\frac{MS_{Treatment}}{MS_{\epsilon}}$
Error	$k(n-1)$	$\sum_{j=1}^{k}\sum_{i=1}^{n}(y_{ij}-\bar{y}_j)^2$	$\frac{SS_{\epsilon}}{df_{\epsilon}}$	
Total	$kn-1$	$\sum_{j=1}^{k}\sum_{i=1}^{n}(y_{ij}-\bar{\bar{y}})^2$		

ANOVA Summary Statistics

- Standard error:

$$s_\epsilon = \sqrt{MS_\epsilon}$$

- Coefficient of determination (r^2) measures the fraction of the variation in the y_{ij} relative to $\bar{\bar{y}}$ that is attributed to biases between the $\bar{y}_j$:

$$r^2 = 1 - \frac{SS_\epsilon}{SS_{total}}$$

- Correlation coefficient:

$$r = \sqrt{r^2}$$

- r^2 is optimistic. The adjusted coefficient of determination $\left(r^2_{adjusted}\right)$ corrects for the complexity of the model giving a more indicator of model performance:

$$r^2_{adjusted} = 1 - \left(\frac{df_{total}}{df_\epsilon}\right)\left(\frac{SS_\epsilon}{SS_{total}}\right)$$

- $r^2_{adjusted} \leq r^2$.

Multiple Comparison Tests

- If ANOVA indicates that there are significant differences among the means, then we have to resort to pairwise testing to find them.
- We must protect ourselves from the elevated risk of type 1 errors.
- **Bonferroni's correction** (two equivalent methods):
  - Reduce the α for individual tests so that the overall (family) error rate is tolerable. For k tests:

$$\alpha = \frac{\alpha_{family}}{k}$$

  - Multiply each p value by the number of tests:

$$p_{corrected} = kp$$

- Recommendations:
  - Use the Tukey-Kramer method for all possible comparisons.
  - Use Dunnett's method to compare treatments to a control.
  - Use Hsu's method to compare the best treatment to the others.

Grouping Information Using Tukey Method
and 95.0% Confidence

Level	N	Mean	Grouping
D	12	102.9	A
G	12	102.7	A
E	12	102.4	A
A	12	100.7	A
B	12	99.3	A
F	12	99.2	A
H	12	98.8	A
C	12	95.3	A

Tukey Simultaneous Tests
Response Variable Y
All Pairwise Comparisons among Levels of Level
Level = A subtracted from:

Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
B	-1.417	3.981	-0.356	1.0000
C	-5.417	3.981	-1.361	0.8723
D	2.250	3.981	0.565	0.9992
E	1.750	3.981	0.440	0.9998
F	-1.500	3.981	-0.377	0.9999
G	2.000	3.981	0.502	0.9996
H	-1.917	3.981	-0.481	0.9997

Level = B subtracted from:

Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
C	-4.000	3.981	-1.005	0.9726
D	3.667	3.981	0.921	0.9832
E	3.167	3.981	0.795	0.9930
F	-0.083	3.981	-0.021	1.0000
G	3.417	3.981	0.858	0.9889
H	-0.500	3.981	-0.126	1.0000

Level = C subtracted from:

Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
D	7.667	3.981	1.9259	0.5373
E	7.167	3.981	1.8003	0.6216
F	3.917	3.981	0.9839	0.9756
G	7.417	3.981	1.8631	0.5795
H	3.500	3.981	0.8792	0.9872

Level = D subtracted from:

Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
E	-0.500	3.981	-0.126	1.0000
F	-3.750	3.981	-0.942	0.9809
G	-0.250	3.981	-0.063	1.0000
H	-4.167	3.981	-1.047	0.9657

Level = E subtracted from:

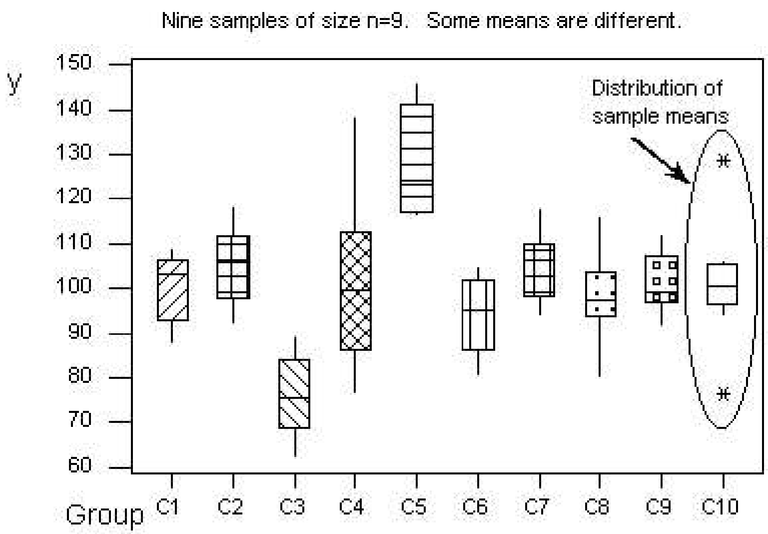
Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
F	-3.250	3.981	-0.8164	0.9918
G	0.250	3.981	0.0628	1.0000
H	-3.667	3.981	-0.9211	0.9832

Level = F subtracted from:

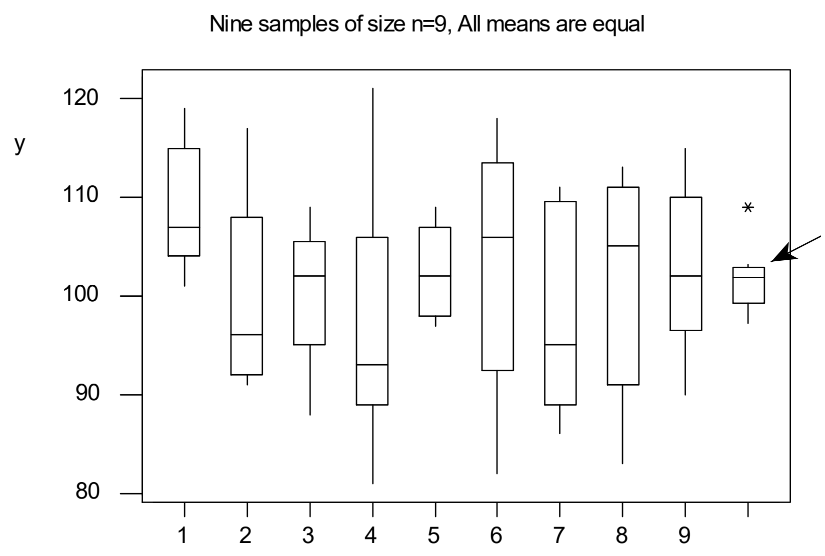
Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
G	3.5000	3.981	0.8792	0.9872
H	-0.4167	3.981	-0.1047	1.0000

Level = G subtracted from:

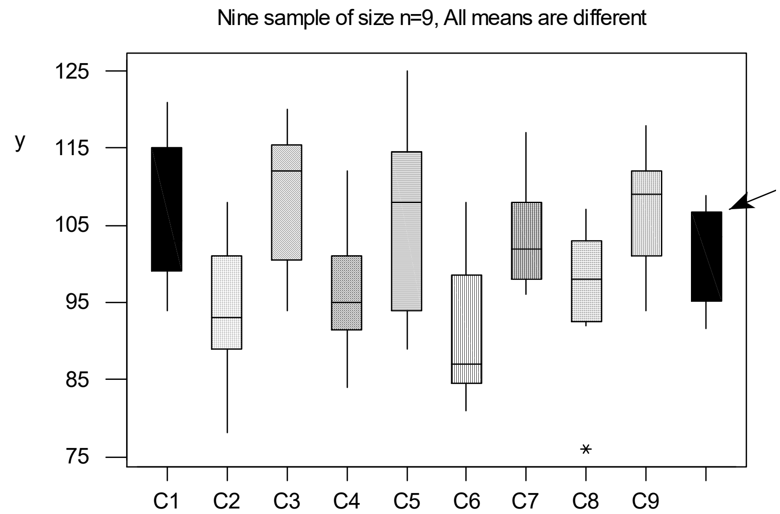
Level	Difference of Means	SE of Difference	T-Value	Adjusted P-Value
H	-3.917	3.981	-0.9839	0.9756




```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      8     13164      1645    15.49    0.000
Error      72      7650       106
Total      80     20814

s = 10.3   r2 = 0.632   r2adj = 0.592

                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev   ------+---------+---------+---------+
C1          9    100.39      7.49                (--*---)
C2          9    104.55      8.29                  (--*---)
C3          9     76.42      8.74    (--*---)
C4          9    100.64     18.71                (--*---)
C5          9    128.77     11.88                               (--*---)
C6          9     94.47      8.44             (--*---)
C7          9    105.93      7.58                   (--*--)
C8          9     98.30      9.76               (--*---)
C9          9    101.04      6.33                (---*--)
                                    ------+---------+---------+---------+
Pooled StDev =    10.31                    80       100       120       140
```




```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      8     820.2     102.5     1.15    0.341
Error      72    6418.7      89.1
Total      80    7238.9
                                    Individual 95% CIs For Mean
                                    Based on Pooled StDev
Level       N      Mean     StDev   -+---------+---------+---------+-----
C1          9    109.00      6.16                      (--------*--------)
C2          9     99.78      9.69        (--------*--------)
C3          9    100.56      6.67         (--------*--------)
C4          9     97.22     12.18    (--------*--------)
C5          9    102.33      4.74           (--------*--------)
C6          9    103.11     11.95            (--------*--------)
C7          9     98.67     10.57      (--------*--------)
C8          9    101.89     11.58           (--------*--------)
C9          9    102.78      8.14            (--------*--------)
                                    -+---------+---------+---------+-----
Pooled StDev =     9.44           91.0      98.0     105.0     112.0
```




```
Analysis of Variance
Source     DF        SS        MS        F        P
Factor      8    2961.2     370.2     4.65    0.000
Error      72    5731.8      79.6
Total      80    8693.0
                                   Individual 95% CIs For Mean
                                   Based on Pooled StDev
Level       N      Mean     StDev  ---+---------+---------+---------+---
C1          9    106.00      9.10                    (------*-------)
C2          9     93.78      8.80     (------*-------)
C3          9    108.78      9.07                        (------*------)
C4          9     96.67      7.91        (-------*------)
C5          9    105.00     12.01                  (------*-------)
C6          9     91.56      9.07  (------*-------)
C7          9    103.56      6.69                (------*-------)
C8          9     96.44      9.28        (-------*------)
C9          9    107.33      7.37                      (------*-------)
                                   ---+---------+---------+---------+---
Pooled StDev =       8.92          88.0      96.0     104.0     112.0
```

Randomization in Experiments

There are many incorrect ways and two correct ways of running a one-way classification design.

- An incorrect way would be to do all of the runs from treatment #1, then all of the runs from treatment #2, and so on until all of the runs are complete, such as:

$$x = \{1111222233334444\}$$

This approach *confounds* the treatments with unknown or *lurking* variables so that observed differences between treatments might actually be caused by the changing states of lurking variables.

$$\begin{Bmatrix} x \\ L \end{Bmatrix} = \begin{Bmatrix} 1111222233334444 \\ 0000001111222222 \end{Bmatrix}$$

- **Any study variable that you intend to test for differences between its treatment levels MUST have its treatment levels in random order**.

Randomization in Experiments

- An acceptable method of doing the runs is the *completely randomized design* (CRD); for example:

  $$\{231443132314142\}$$

  The randomized order of the runs provides some protection from the effects of lurking variables; however, the noise from any lurking variable is pooled with the model error reducing the sensitivity of the experiment.
- A better method of doing the runs is the *randomized block design* (RBD); for example:

  $$\{1342, 1243, 2134, 4312\}$$

  This RBD is run in four blocks where each block contains one *replicate* of the experiment design. Conditions within each block must be homogeneous but differences between blocks are allowed. The analysis is run using two-way ANOVA with terms for treatments and blocks. The effects of lurking variables are captured in the block term, reducing the error estimate, so the experiment has increased sensitivity for the study variable effect.

Analysis of Means

- ANOM is an alternative to ANOVA
- ANOM is user-friendly because its results are presented in the form of a control chart
- The hypotheses tested in ANOM are:

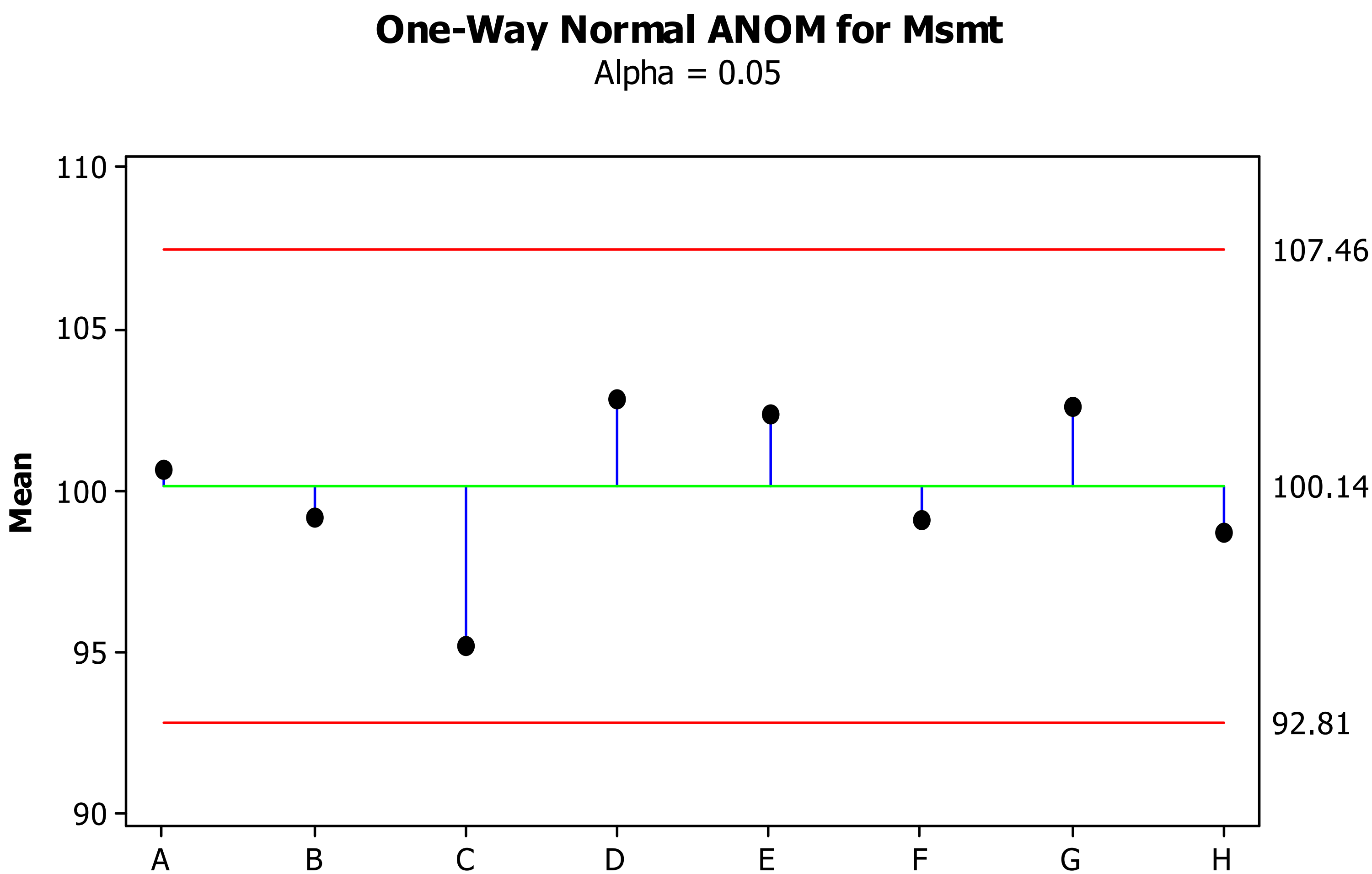
$$H_0 : \mu_i = \mu \text{ for all treatments}$$

$$H_A : \mu_i \neq \mu \text{ for at least one treatment}$$

- ANOM is not as versatile as ANOVA so it isn't used as often
- With appropriate variable transformations, ANOM can be used to analyze proportion and count data.

Analysis of Means

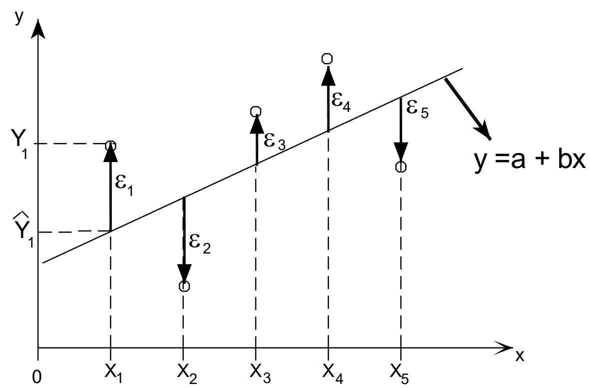
For the data from the original example:

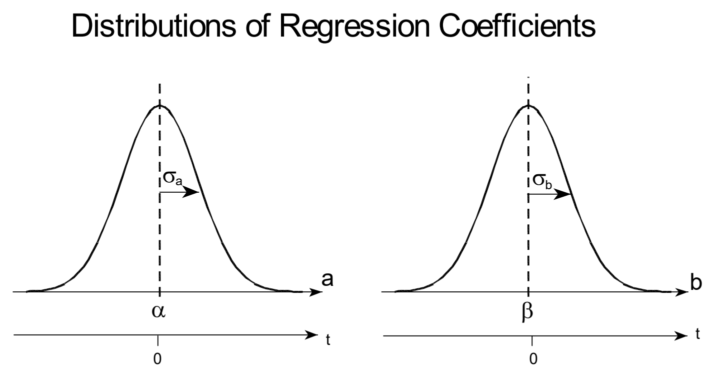


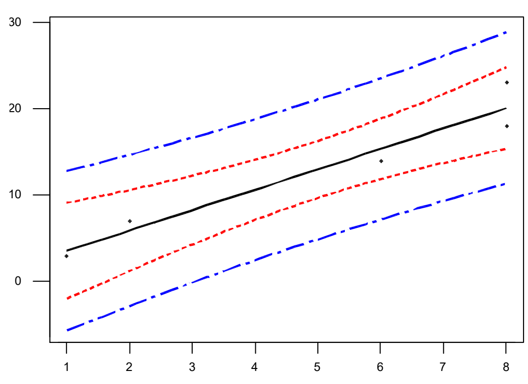
MINITAB Commands

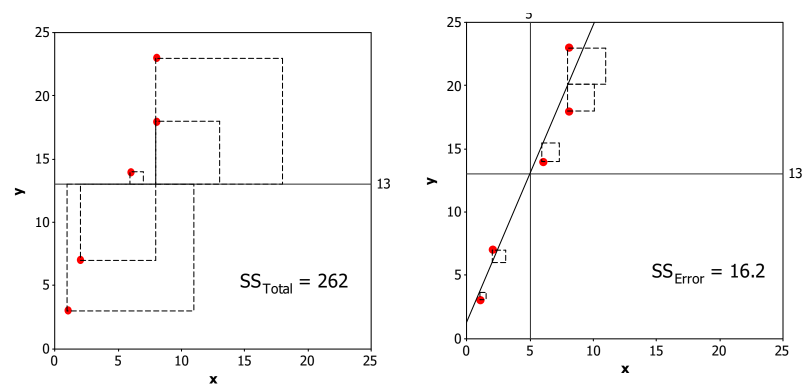
- Graph> Boxplot> With Groups
- Stat> ANOVA> General Linear Model
- Stat> ANOVA> Test for Equal Variances
- Stat> ANOVA> Analysis of Means

Introduction to Linear Regression







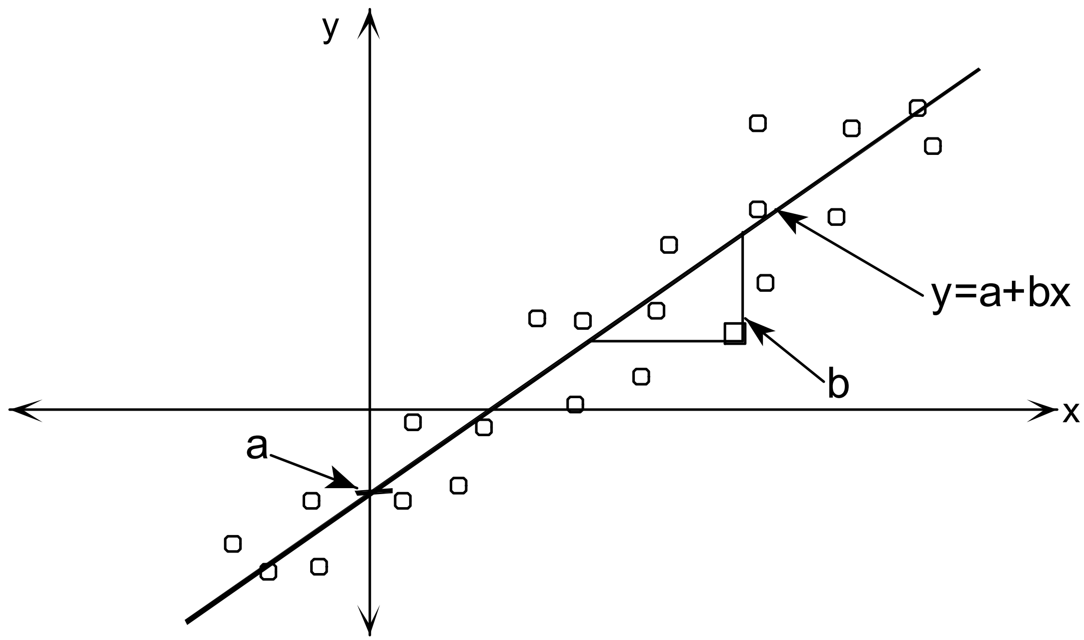


Linear Regression

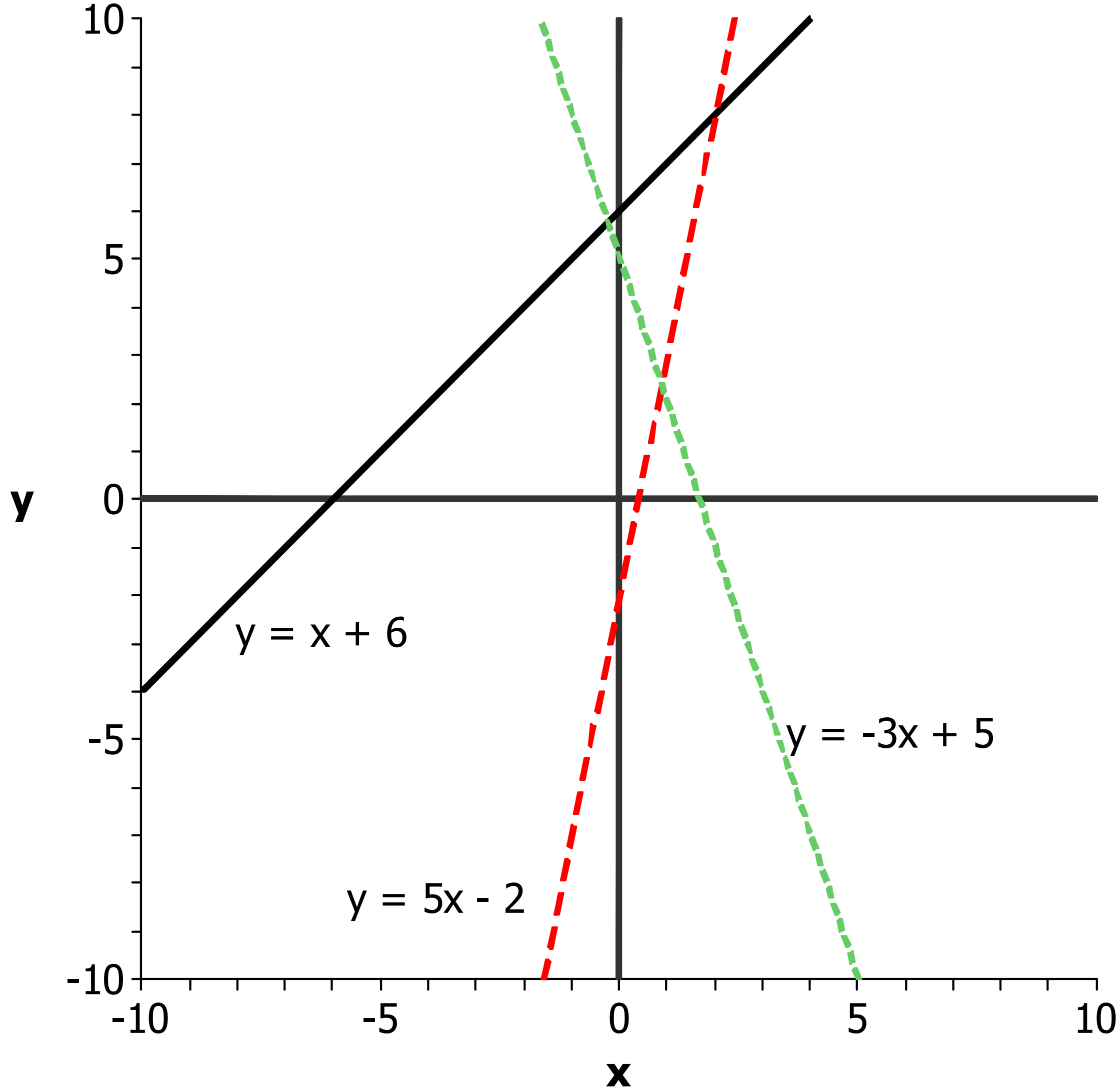
Linear regression relates a quantitative response y to a quantitative predictor variable x using a relationship of the form:

$$y = a + bx$$

where a is the y axis intercept and b is the slope.



Examples:



Linear Regression Notation

- The data are (x_i, y_i) pairs where x_i is the independent variable (the knob) and y_i is the response or output.
- Each observed value of y_i has two components - one described by the model and the other consisting of error:

$$y_i = a + bx_i + \epsilon_i$$

- The observed values may also be written as:

$$y_i = \widehat{y}_i + \epsilon_i$$

- ... where the model or predicted value of $\widehat{y}_i$ is given by:

$$\widehat{y}_i = a + bx_i$$

- The errors or discrepancies between the observed and predicted values are called the model residuals:

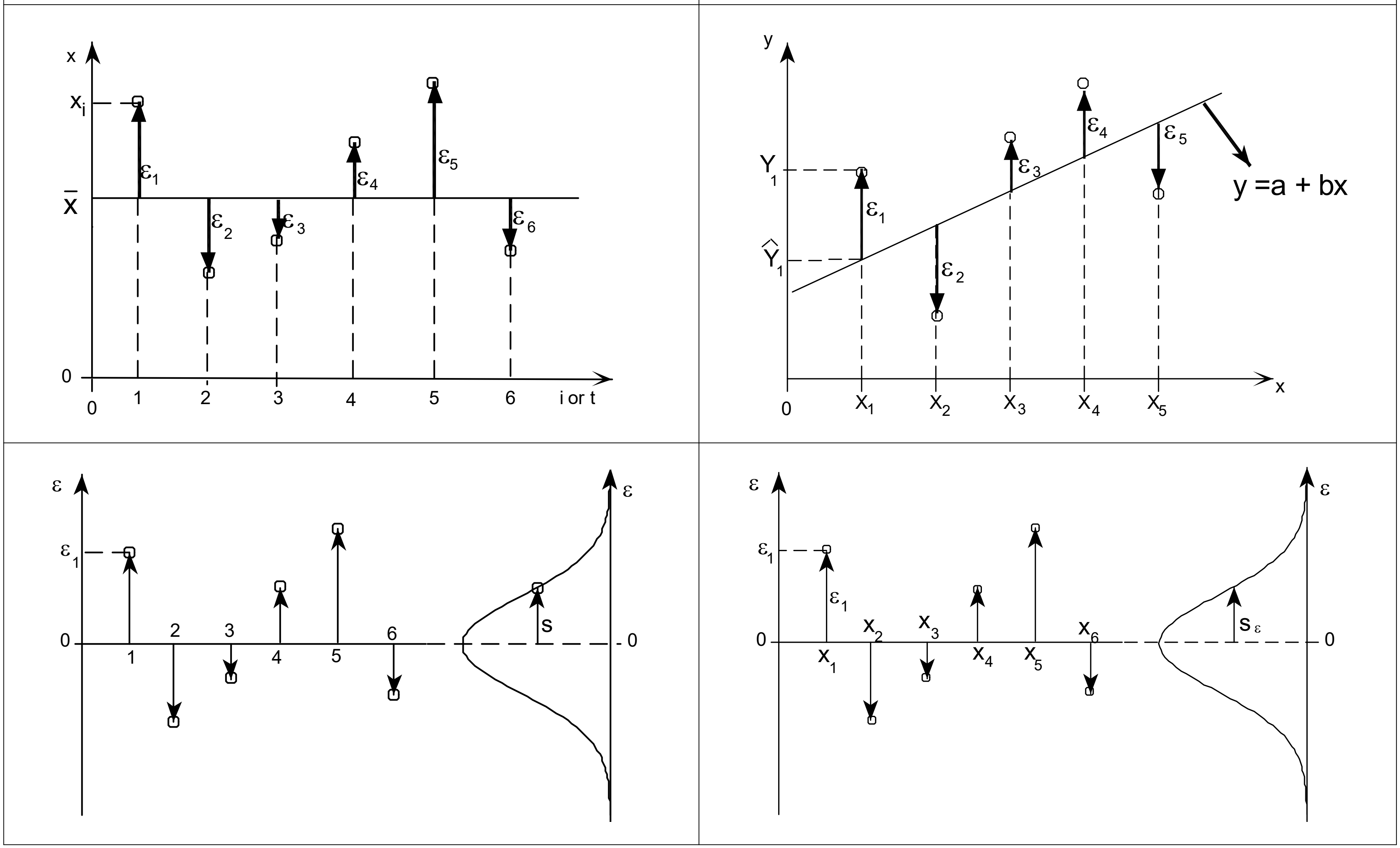
$$\begin{aligned} \epsilon_i &= y_i - \widehat{y}_i \\ &= y_i - (a + bx_i) \end{aligned}$$

Relationship Between Standard Deviation of a Sample and Standard Error of the Regression

Calculation of:	*Std. Deviation*	*Std. Error of Regression*
Data:	$x_i, i = 1$ to n	$(x_i, y_i), i = 1$ to n
Predicted Value:	$\widehat{x}_i = \bar{x} = \frac{\Sigma x_i}{n}$	$\widehat{y}_i = a + bx_i$
df consumed:	1	2
Deviation:	$\epsilon_i = x_i - \bar{x}$	$\epsilon_i = y_i - \widehat{y}_i$
Variation:	$s = \sqrt{\frac{\Sigma \epsilon_i^2}{n-1}}$	$s = \sqrt{\frac{\Sigma \epsilon_i^2}{n-2}}$

Graphical Relationship

Standard Deviation	*Std. Error of the Regression*



Sums of Squares

Special quantities called sums of squares (SS) make calculation of regression statistics easier:

$$SS_x = \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$SS_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$SS_{xy} = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \text{ and } \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

Sums of Squares

- These sums of squares are also used to determine the variances of x and y :

$$s_x^2 = \frac{SS_x}{n-1} \text{ and } s_y^2 = \frac{SS_y}{n-1}$$

- The quantity $\sum_{i=1}^{n} \epsilon_i^2$ shows up so often that we give it a special name: the error sum of squares:

$$SS_\epsilon = \sum_{i=1}^{n} \epsilon_i^2$$

- The total variation in the response is given by the total sum of squares indicated by SS_y or SS_{total}:

$$SS_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Least Squares Regression

The **linear least squares regression equation** that expresses y as:

$$y_i = a + bx_i + \epsilon_i$$

simultaneously satisfies the conditions:

$$\frac{\partial}{\partial a}\sum_{i=1}^{n}\epsilon_i^2 = 0 \text{ and } \frac{\partial}{\partial b}\sum_{i=1}^{n}\epsilon_i^2 = 0$$

These equations lead to the two normal equations for a and b:

$$b = \frac{SS_{xy}}{SS_x}$$

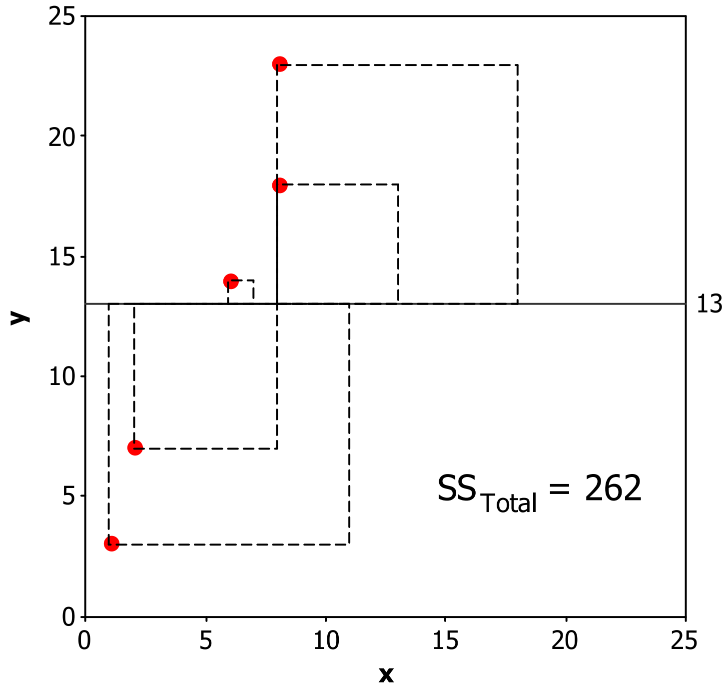
and

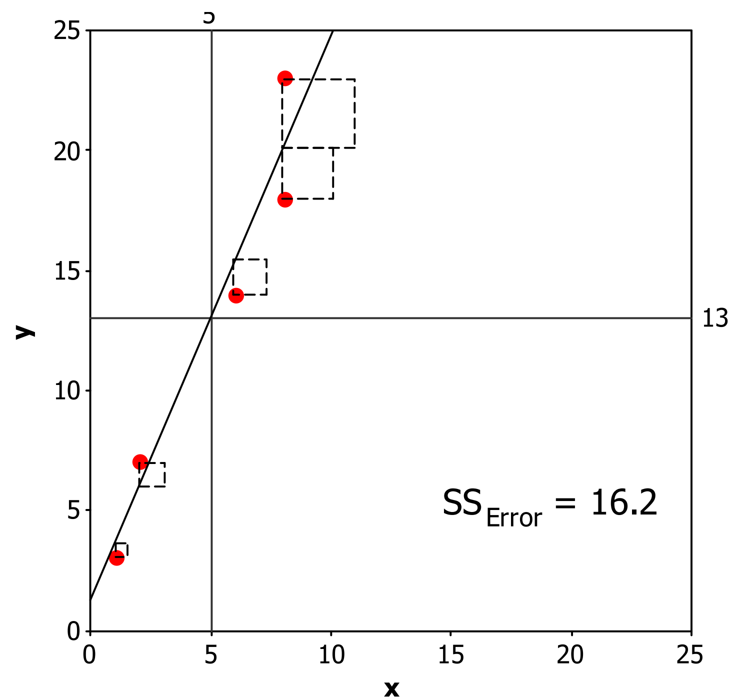
$$a = \bar{y} - b\bar{x}$$

Note that $(\bar{x},\bar{y})$ is on the regression line.

Least Squares Regression

- Total variation in the response y relative to the mean $\bar{y}$ is given by SS_y or SS_{Total}.
- Variation in the response relative to the least squares fitted line is given by SS_ϵ or SS_{Error}.
- Variation explained by the fitted line is given by $SS_{Regression} = SS_{Total} - SS_{Error}$





Example: Determine the linear regression equation that provides the best fit to the following data set using a calculator.

i	x_i	y_i
1	1	3
2	2	7
3	6	14
4	8	18
5	8	23

Solution: We need to find $\bar{x}$, $\bar{y}$, SS_x, SS_y, and SS_{xy}.

$$\bar{x} = \frac{1}{5}\sum_{i=1}^{5} x_i = \frac{1}{5}(1+2+6+8+8) = 5.0$$

$$\bar{y} = \frac{1}{5}\sum_{i=1}^{5} y_i = \frac{1}{5}(3+7+14+18+23) = 13.0$$

i	x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	y_i	$(y_i - \bar{y})$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
1	1	-4	16	3	-10	100	40
2	2	-3	9	7	-6	36	18
3	6	1	1	14	1	1	1
4	8	3	9	18	5	25	15
5	8	3	9	23	10	100	30
Totals	25	0	44	65	0	262	104

So, we have $SS_x = 44$, $SS_y = 262$, and $SS_{xy} = 104$.

- The value of the slope in the regression equation is:

$$b = \frac{SS_{xy}}{SS_x} = \frac{104}{44} = 2.36$$

- The value of the y axis intercept is:

$$\begin{aligned} a &= \bar{y} - b\bar{x} \\ &= 13.0 - 2.36 \times 5.0 \\ &= 1.18 \end{aligned}$$

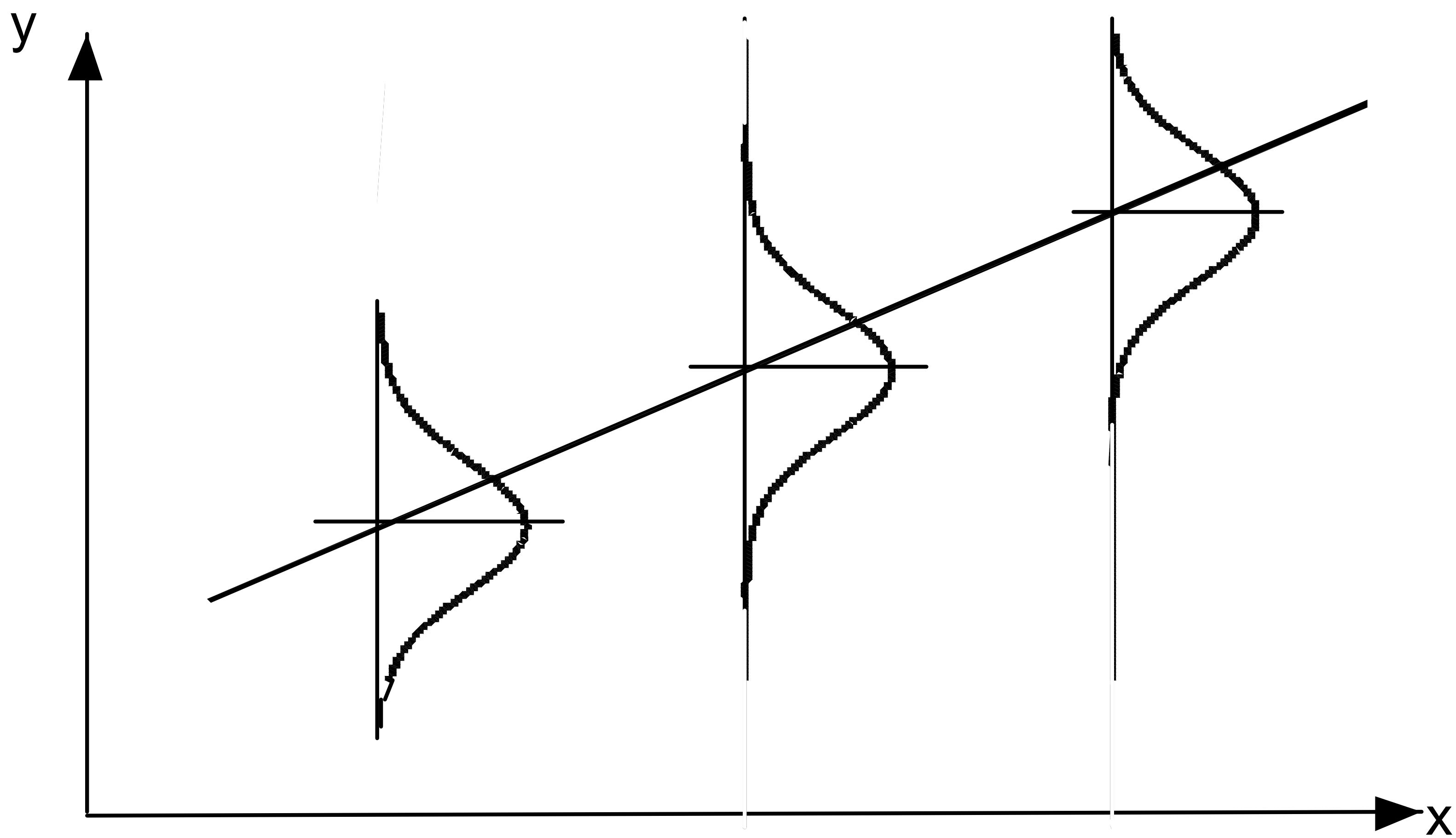
- The regression equation is:

$$y = 1.18 + 2.36x$$

Linear Regression Assumptions

1. The values of x_i are known exactly. If there is error in the x_i then the regression coefficients will be biased.
2. The distribution of the ϵ_i is normal. Check this assumption using a normal plot.
3. The error variance is constant along the fitted line. Check this assumption from a plot of ϵ_i vs. x_i.
4. The ϵ_i are independent of each other. Check this assumption from a run chart of the ϵ_i or from a lag-one chart (ϵ_i vs. ϵ_{i-1}).
5. A linear model is an appropriate model for the data. Check this assumption using a lack of fit (or goodness of fit) test.

Regression Assumptions



The Standard Error of the Model

The standard error of the linear regression model $y = a + bx$ is:

$$s_\epsilon = \sqrt{\frac{\sum_{i=1}^{n} \epsilon_i^2}{n-2}}$$

where $n - 2$ degrees of freedom must be used because two statistics are calculated from the data (a and b). The standard error measures how far the normally distributed residuals fall from the regression line.

Example: Determine the standard error of the model for the example data set.

Solution: The standard error is found from the model residuals:

i	x_i	y_i	$\widehat{y}_i$	ϵ_i	ϵ_i^2
1	1	3	3.54	–0.54	0.29
2	2	7	5.90	1.10	1.21
3	6	14	15.34	–1.34	1.80
4	8	18	20.06	–2.06	4.24
5	8	23	20.06	2.94	8.64
			Totals	0.00	16.18

With $SS_\epsilon = \sum \epsilon_i^2 = 16.18$ the model standard error is:

$$s_\epsilon = \sqrt{\frac{SS_\epsilon}{n-2}} = \sqrt{\frac{16.18}{5-2}} = 2.32$$

Prediction Interval (Approximation)

The model and the standard error can be combined to form an approximate prediction interval - an interval that should contain a specified fraction of the population of observations. Calculate the approximate $(1-\alpha)100\%$ prediction interval for $y(x_i)$ a given x_i value from:

$$P(\hat{y} - t_{\alpha/2}s_\epsilon < y(x_i) < \hat{y} + t_{\alpha/2}s_\epsilon) = 1-\alpha$$

where $\hat{y} = a + bx_i$ and the t distribution has degrees of freedom equal to the error degrees of freedom.

Example: Use the model and standard error to predict the response for the example data set when $x = 4$ and construct the 95% prediction interval.

Solution: The predicted value of y is:

$$\begin{aligned} \hat{y}(x = 4) &= 1.18 + 2.36(4) \\ &= 10.6 \end{aligned}$$

The 95% prediction interval is given by

$$P(\hat{y} - t_{0.025} s_{\epsilon} < y < \hat{y} + t_{0.025} s_{\epsilon}) = 0.95$$

where the t distribution has $df_{\epsilon} = n - 2 = 5 - 2 = 3$ degrees of freedom. The required t value is $t_{0.025,3} = 3.18$ so the prediction interval is:

$$P(10.6 - 3.18 \times 2.32 < y < 10.6 + 3.18 \times 2.32) = 1 - 0.05$$

$$P(3.25 < y < 17.9) = 0.95$$

That is, 95% of the observations taken at $x = 4$ should yield a value of y that falls between 3.25 and 17.9.

Coefficient of Determination

- The coefficient of determination r^2 indicates the fraction of the total variation in y that is explained by its dependence on x:

$$
\begin{aligned}
r^2 &= \frac{SS_{regression}}{SS_{total}} \\
&= 1 - \frac{SS_{\epsilon}}{SS_{total}} \\
&= 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}
\end{aligned}
$$

- The calculating form of the coefficient of determination is:

$$
r^2 = \frac{SS_{xy}^2}{SS_x SS_y}
$$

- Correlation analysis requires less rigorous assumptions than regression so it may still be used in cases where regression is not valid, e.g. analysis of (y_1, y_2) where both variables contain error.

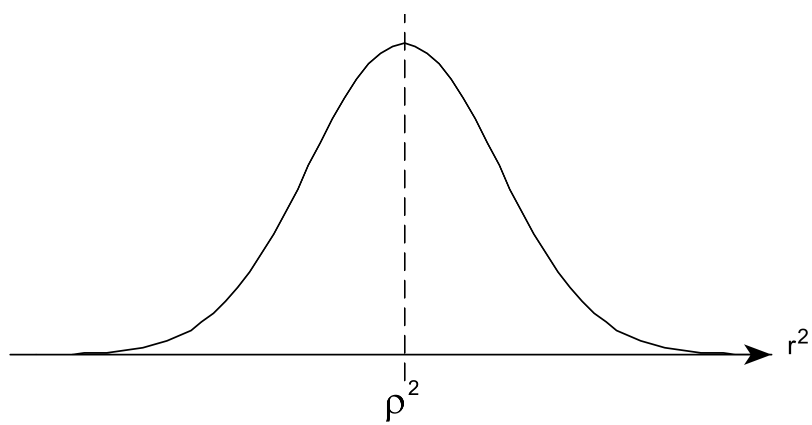
Population Coefficient of Determination

- The value of r^2 determined from the data is a statistic.
- r^2 is an estimate for the population coefficient of determination ρ^2 which is a parameter.
- The distribution of r is complicated, but the distribution of $Z = \frac{1}{2}\ln(\frac{1+r}{1-r})$ is approximately normal with mean

$$\mu_Z = \frac{1}{2}\ln\left(\frac{1+\rho}{1-\rho}\right)$$

and standard deviation

$$\sigma_Z = \sqrt{\frac{1}{n-3}}$$



Adjusted Coefficient of Determination

- The adjusted coefficient of determination $r^2_{adjusted}$ corrects for the degree of model complexity. For a linear model:

$$r^2_{adjusted} = 1 - \left(\frac{n-1}{n-2}\right)\left(\frac{SS_\epsilon}{SS_{total}}\right)$$

- $r^2_{adjusted} \leq r^2$
- $r^2_{adjusted}$ should always be used instead of r^2.

Example: Determine the coefficient of determination and the adjusted coefficient of determination for the example data set.

- The coefficient of determination is given by:

$$r^2 = 1 - \frac{SS_\epsilon}{SS_y} = 1 - \frac{16.18}{262} = 0.938$$

- Alternatively the coefficient of determination can also be determined from:

$$r^2 = \frac{SS_{xy}^2}{SS_x SS_y} = \frac{104^2}{44 \times 262} = 0.938$$

- The adjusted coefficient of determination is given by:

$$\begin{aligned} r^2_{adjusted} &= 1 - \frac{(5-1)SS_\epsilon}{(5-2)SS_y} \\ &= 1 - \frac{4\times16}{3\times262} \\ &= 0.919 \end{aligned}$$

Distributions of Sample Regression Coefficients

- a is an estimate for the true parameter α
- b is an estimate for the true parameter β
- The correct equation is:

$$y = \alpha + \beta x$$

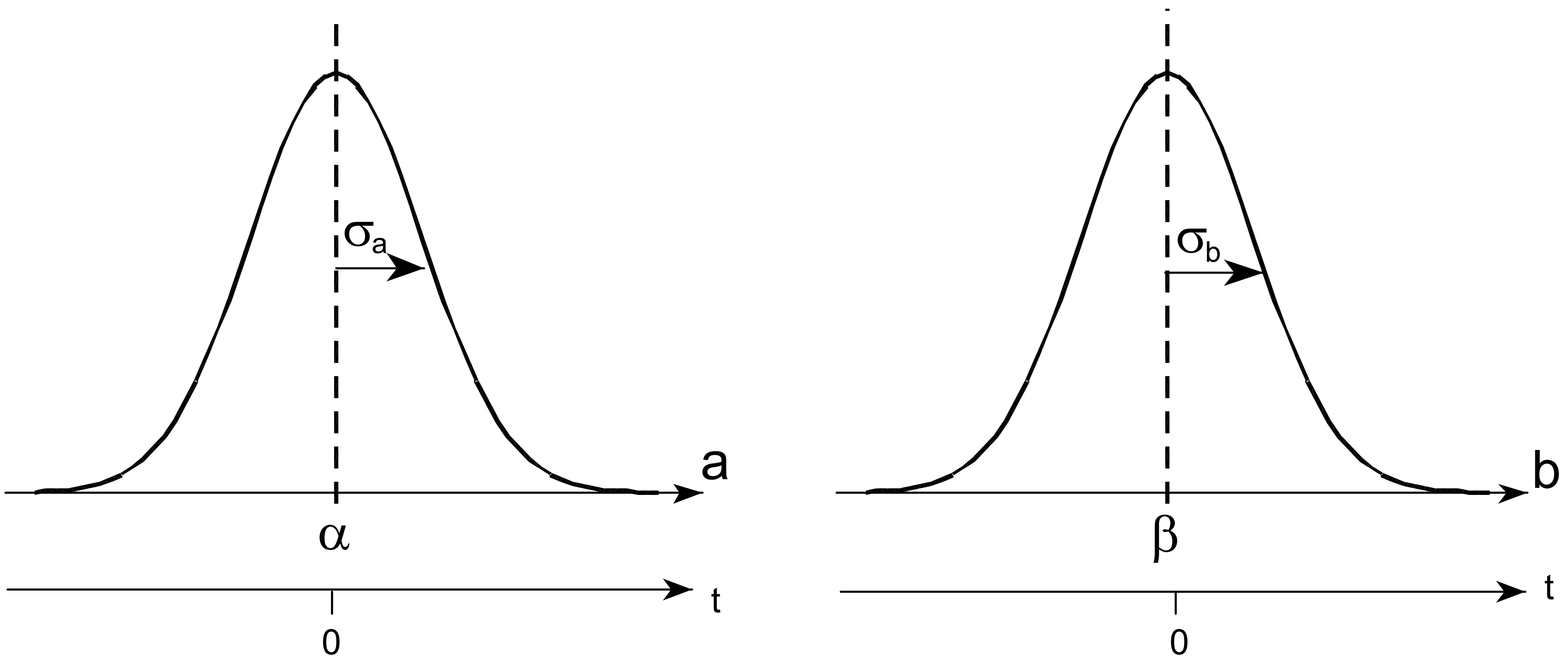
- The distribution of a values about α is Student's t with $n - 2$ degrees of freedom and:

$$s_a = s_\epsilon \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{SS_x}}$$

- The distribution of b values about β is Student's t with $n - 2$ degrees of freedom and:

$$s_b = s_\epsilon \sqrt{\frac{1}{SS_x}}$$

Distributions of Regression Coefficients



Hypothesis Tests for Regression Coefficients

- Test $H_0 : \alpha = 0$ vs. $H_A : \alpha \neq 0$ using the test statistic $t = a/s_a$. Accept H_0 iff $-t_{\alpha/2,n-2} < t < t_{\alpha/2,n-2}$.
- Test $H_0 : \beta = 0$ vs. $H_A : \beta \neq 0$ using the test statistic $t = b/s_b$. Accept H_0 iff $-t_{\alpha/2,n-2} < t < t_{\alpha/2,n-2}$.

Example: Determine the standard deviations of the regression coefficients from the example problem and test them at $\alpha = 0.05$ to see if they differ from 0.

Solution:

- The standard deviations of the regression coefficients are given by:

$$
\begin{aligned}
s_a &= s_\epsilon \sqrt{\frac{1}{n} + \frac{\bar{x}^2}{SS_x}} \\
&= 2.32\sqrt{\frac{1}{5} + \frac{5^2}{44}} \\
&= 2.03
\end{aligned}
$$

and

$$
\begin{aligned}
s_b &= \frac{s_\epsilon}{\sqrt{SS_x}} \\
&= \frac{2.32}{\sqrt{44}} \\
&= 0.35
\end{aligned}
$$

- The t value for the hypothesis test of $H_0 : \alpha = 0$ versus $H_A : \alpha \neq 0$ is given by:

$$t_a = \frac{a}{s_a} = \frac{1.18}{2.03} = 0.58$$

- Similarly, for $H_0 : \beta = 0$ versus $H_A : \beta \neq 0$:

$$t_b = \frac{b}{s_b} = \frac{2.36}{0.35} = 6.74$$

- The critical t value that these statistics must be compared to is $t_{0.025,3} = 3.18$ so the acceptance interval for H_0 is $P(-3.18 < t < 3.18) = 0.95$.
- Since $t_a = 0.58$ falls within the acceptance interval of H_0 we have to accept the null hypothesis $H_0 : \alpha = 0$.
- Since $t_b = 6.74$ falls outside of the acceptance interval for H_0 we must conclude that $b = 2.36$ is significantly different from zero, that is, we must reject $H_0 : \beta = 0$ in favor of $H_A : \beta \neq 0$.

Confidence Intervals for the Regression Parameters

- The $(1 - \eta)100\%$ confidence interval for the y-axis intercept parameter α is:

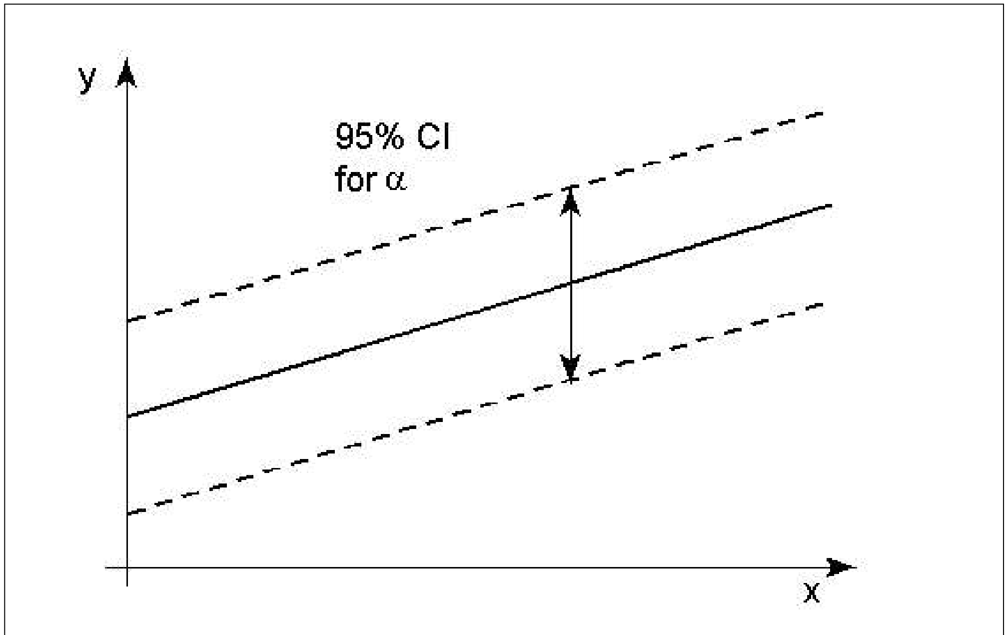
$$P(a - t_{\eta/2}s_a < \alpha < a + t_{\eta/2}s_a) = 1 - \eta$$

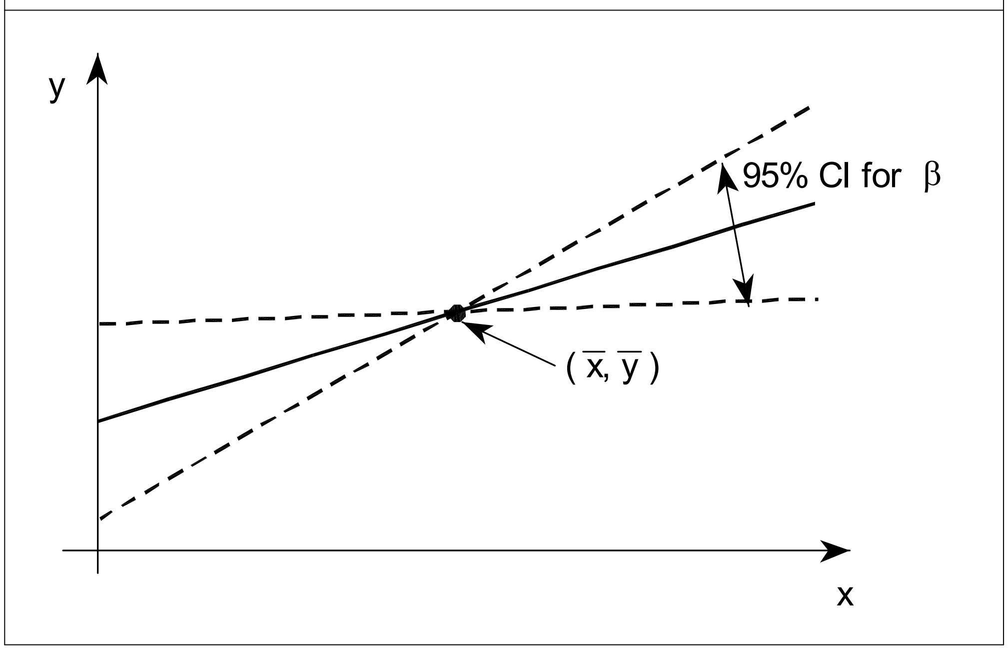
  where $t_{\eta/2}$ has $n - 2$ degrees of freedom.

- The $(1 - \eta)100\%$ confidence interval for the slope parameter β is:

$$P(b - t_{\eta/2}s_b < \beta < b + t_{\eta/2}s_b) = 1 - \eta$$

  where $t_{\eta/2}$ has $n - 2$ degrees of freedom.


y
95% CI
for α
x


y
95% CI for β
$(\bar{x}, \bar{y})$
x

Example: Construct the 95% confidence intervals for the true population coefficients α and β for the example data set.

Solution:

- The confidence interval for α is:

$$P(1.18 - 3.18 \times 2.03 < \alpha < 1.18 + 3.18 \times 2.03) = 1 - 0.05$$

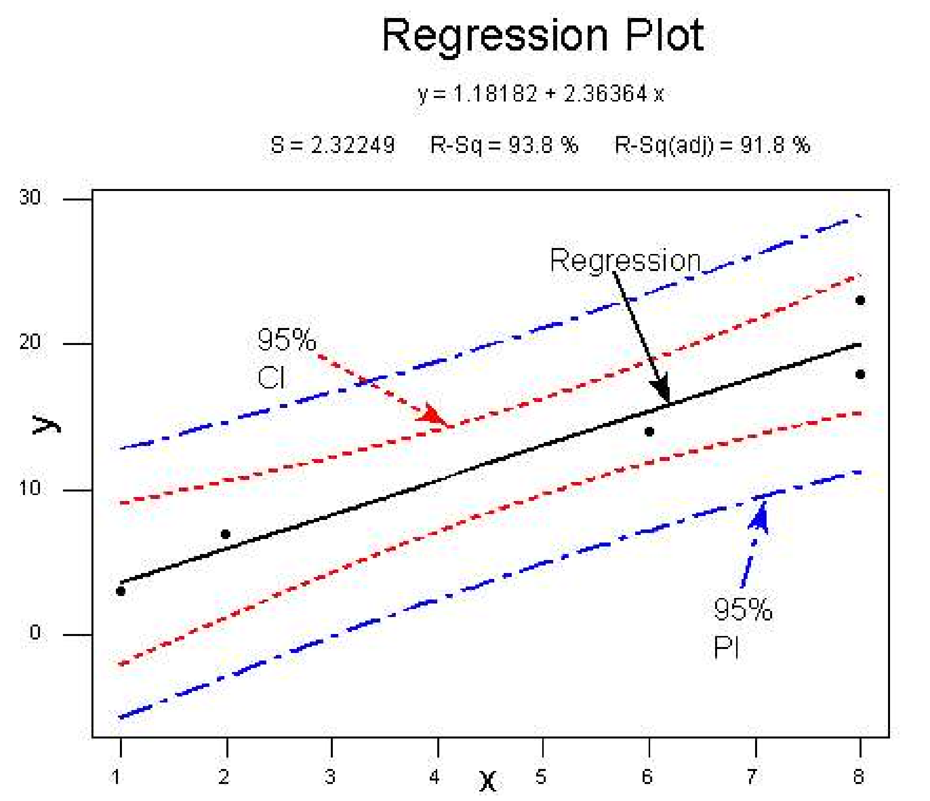
$$P(-5.2 < \alpha < 7.6) = 0.95$$

- The confidence interval for β is:

$$P(2.36 - 3.18 \times 0.35 < \beta < 2.36 + 3.18 \times 0.35) = 1 - 0.05$$

$$P(1.2 < \beta < 3.5) = 0.95$$

Notice that even though the t test indicates that β is significantly different from zero there is still a large degree of uncertainty in the value of β as estimated from $b = 2.36$.



Regression Analysis

```
The regression equation is
y = 1.18 + 2.36 x

Predictor        Coef       StDev          T        P
Constant        1.182       2.036       0.58    0.602
x              2.3636      0.3501       6.75    0.007

S = 2.322      R-Sq = 93.8%     R-Sq(adj) = 91.8%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         1      245.82      245.82     45.57    0.007
Residual Error     3       16.18        5.39
Total              4      262.00

MTB > print c1-c4
```

Data Display

Row	x	y	resid
1	1	3	-0.54545
2	2	7	1.09091
3	6	14	-1.36364
4	8	18	-2.09091
5	8	23	2.90909

Transformations

There are many situations in which y is a nonlinear function of x but a transformation on x and/or y will linearize the relationship. For example, the power function

$$y = ax^b$$

can be linearized and expressed as

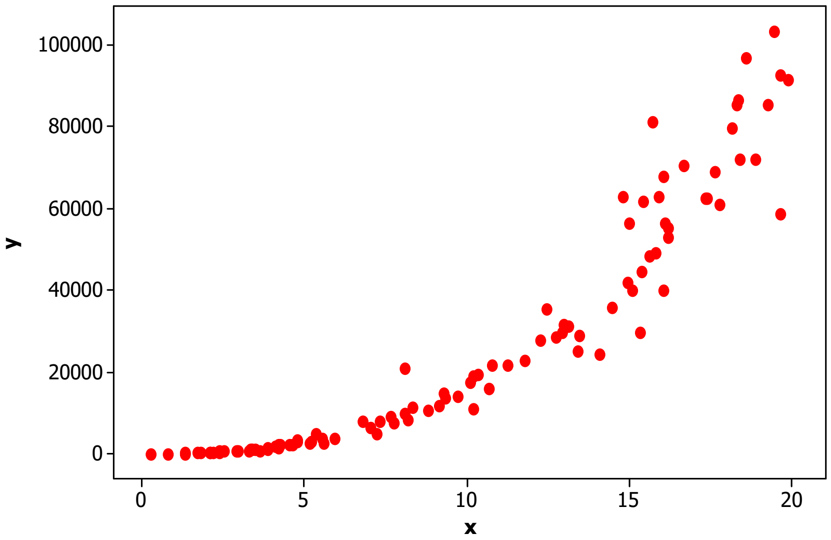
$$y' = a' + bx'$$

where $y' = \log(y)$, $x' = \log(x)$, and $a' = \log(a)$. A plot of the original (x_i, y_i) data would be highly nonlinear but the plots of the (x_i, y_i) data on log-log paper or $(log(x_i), log(y_i))$ on linear-linear paper would show a straight line.

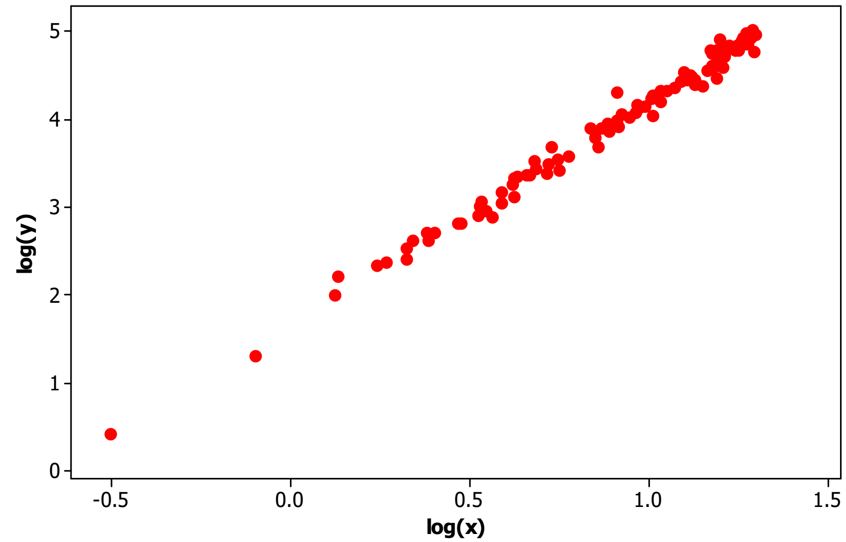
Transformation to Linear Form:

Function	y'	x'	a'	Linear Form
$y = ae^{bx}$	$\ln y$		$\ln a$	$y' = a' + bx$
$y = ax^b$	$\log y$	$\log x$	$\log a$	$y' = a' + bx'$
$y = a + \frac{b}{x}$		$\frac{1}{x}$		$y = a + bx'$
$y = \frac{1}{a+bx}$	$\frac{1}{y}$			$y' = a + bx$
$y = ae^{\frac{b}{x}}$	$\ln y$	$\frac{1}{x}$	$\ln a$	$y' = a' + bx'$
$y = ax^2e^{bx}$	$\ln\left(\frac{y}{x^2}\right)$		$\ln a$	$y' = a' + bx$
$n = n_o e^{\frac{-\varphi}{kT}}$	$\ln n$	$\frac{1}{kT}$	$\ln n_o$	$y' = a' - \varphi x'$
$j = AT^2 e^{\frac{-\varphi}{kT}}$	$\ln\left(\frac{j}{T^2}\right)$	$\frac{1}{kT}$	$\ln A$	$y' = a' - \varphi x'$
$f(y) = a + bf(x)$	$f(y)$	$f(x)$		$y' = a + bx'$

Example



100000
80000
60000
40000
20000
0
y
0
5
10
15
20
x



5
4
3
2
1
0
log(y)
-0.5
0.0
0.5
1.0
1.5
log(x)

Polynomial and Multivariable Models

- A polynomial model has one independent variable:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots$$

- A multivariable model has more than one independent variable:

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + \ldots$$

- The form of the polynomial and multivariable models are very similar so they are analyzed in a similar manner.
- We need a systematic way of analyzing these complex models.

Sums of Squares and Degrees of Freedom

- The total variation in the response is given by:

$$SS_{total} = SS_y = \sum_{i=1}^{n}(y_i - \bar{y})^2$$

- In a complex model the total variability is partitioned into components associated with each term in the model and with the error:

$$SS_{total} = SS_1 + SS_2 + \ldots + SS_\epsilon$$

- Just as the sum of squares can be partitioned, there is a corresponding partitioning of the degrees of freedom:

$$df_{total} = df_1 + df_2 + df_3 + \ldots + df_\epsilon$$

  where $df_1 = 1$, $df_2 = 1$, $df_3 = 1$, ...
- $df_{total} = n - 1$ always! (The 1 lost degree of freedom was used to determine $\bar{y}$.)
- $df_{model} = df_1 + df_2 + df_3 + \ldots$ is the number of terms in the model
- $df_\epsilon = df_{total} - df_{model}$

Adjusted Coefficient of Determination

- The adjusted coefficient of determination $r^2_{adjusted}$ corrects for the degree of model complexity:

$$r^2_{adjusted} = 1 - \left(\frac{df_{total}}{df_{\epsilon}}\right)\left(\frac{SS_{\epsilon}}{SS_{total}}\right)$$

- $r^2_{adjusted} \leq r^2$.

Polynomial Models

- Only used as a last resort if a first principles model or a model involving a variable transformation cannot be found.
- The general form is:

$$y = a_0 + a_1 x + a_2 x^2 + \ldots + a_p x^p$$

- $df_{model} = p$
- $df_{\epsilon} = n - p - 1$
- The model standard error will be:

$$s_{\epsilon} = \sqrt{\frac{\sum_{i=1}^{n} \epsilon_i^2}{df_{\epsilon}}}$$

Polynomial Model Example

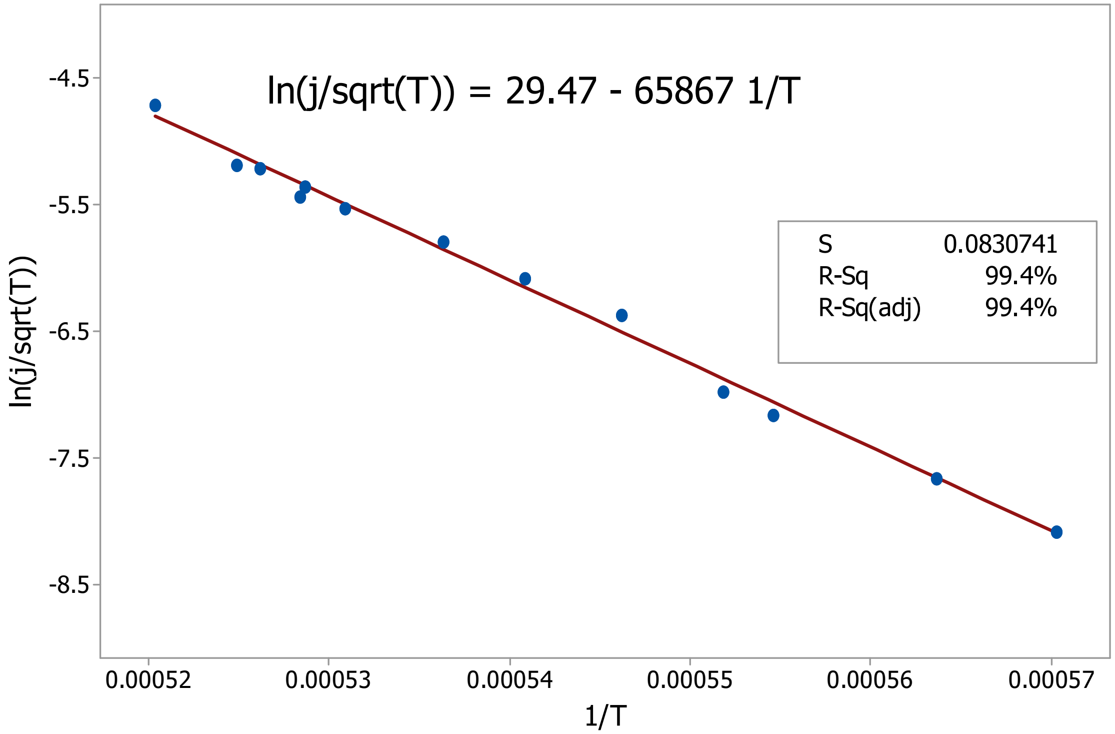
The total emissivity of tungsten is a nonlinear function of temperature. There is no known transformation or simple first principles model so a polynomial model is appropriate. The data set includes $n = 26$ points from the CRC Handbook of Chemistry and Physics (GE data!). All temperatures are in Kelvin.

- $\epsilon(T) = 0.459 - 388/T, s_\epsilon = 0.0067$
- $\epsilon(T) = -0.1055 + 2.5 \cdot 10^{-4} T - 3.43 \cdot 10^{-8} T^2, s_\epsilon = 0.0022$
- $\epsilon(T) = 0.510 - \frac{597}{T} + \frac{188320}{T^2}, s_\epsilon = 0.0029$

Warning!

- Distrust empirical models!
- Distrust the data sets that empirical models are derived from.
- Models should be based on accepted first principles.
- An OK first principles model is better than a good empirical model.
- All models are wrong. Some are useful. - George Box

Richardson's (Wrong) Empirical Model for Thermionic Emission

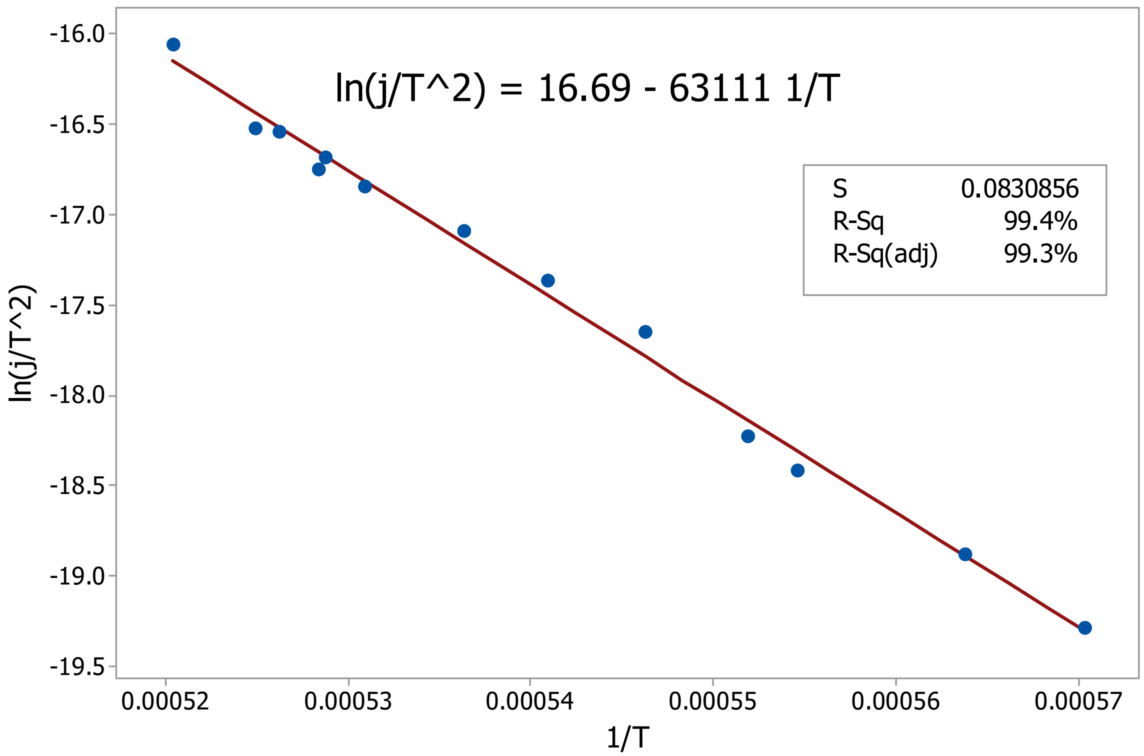


$$j = A\sqrt{T}\, e^{-\frac{\phi}{kT}}$$

$$\frac{j}{\sqrt{T}} = Ae^{-\frac{\phi}{kT}}$$

$$\ln\left(\frac{j}{\sqrt{T}}\right) = \ln(A) - \left(\frac{\phi}{kT}\right)\frac{1}{T}$$

Dushman's (Correct) First Principles Model

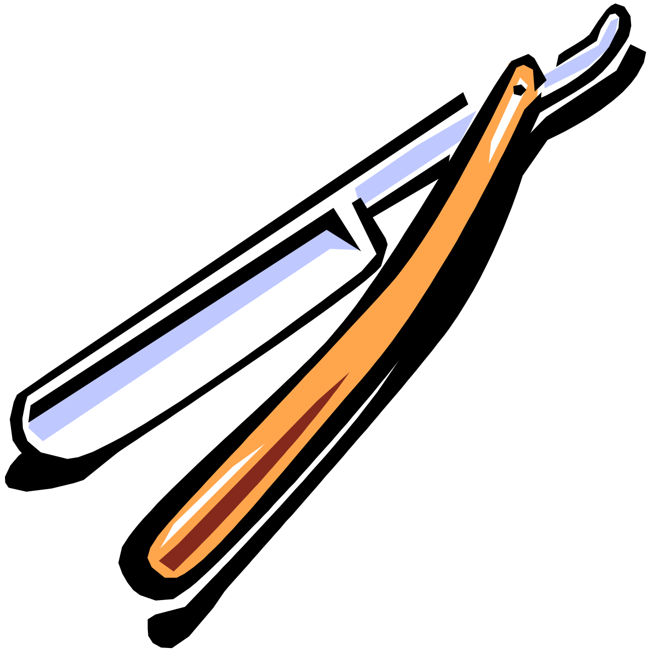


$$j = AT^2 e^{-\frac{\phi}{kT}}$$

$$\frac{j}{T^2} = Ae^{-\frac{\phi}{kT}}$$

$$\ln\left(\frac{j}{T^2}\right) = \ln(A) - \left(\frac{\phi}{kT}\right)\frac{1}{T}$$

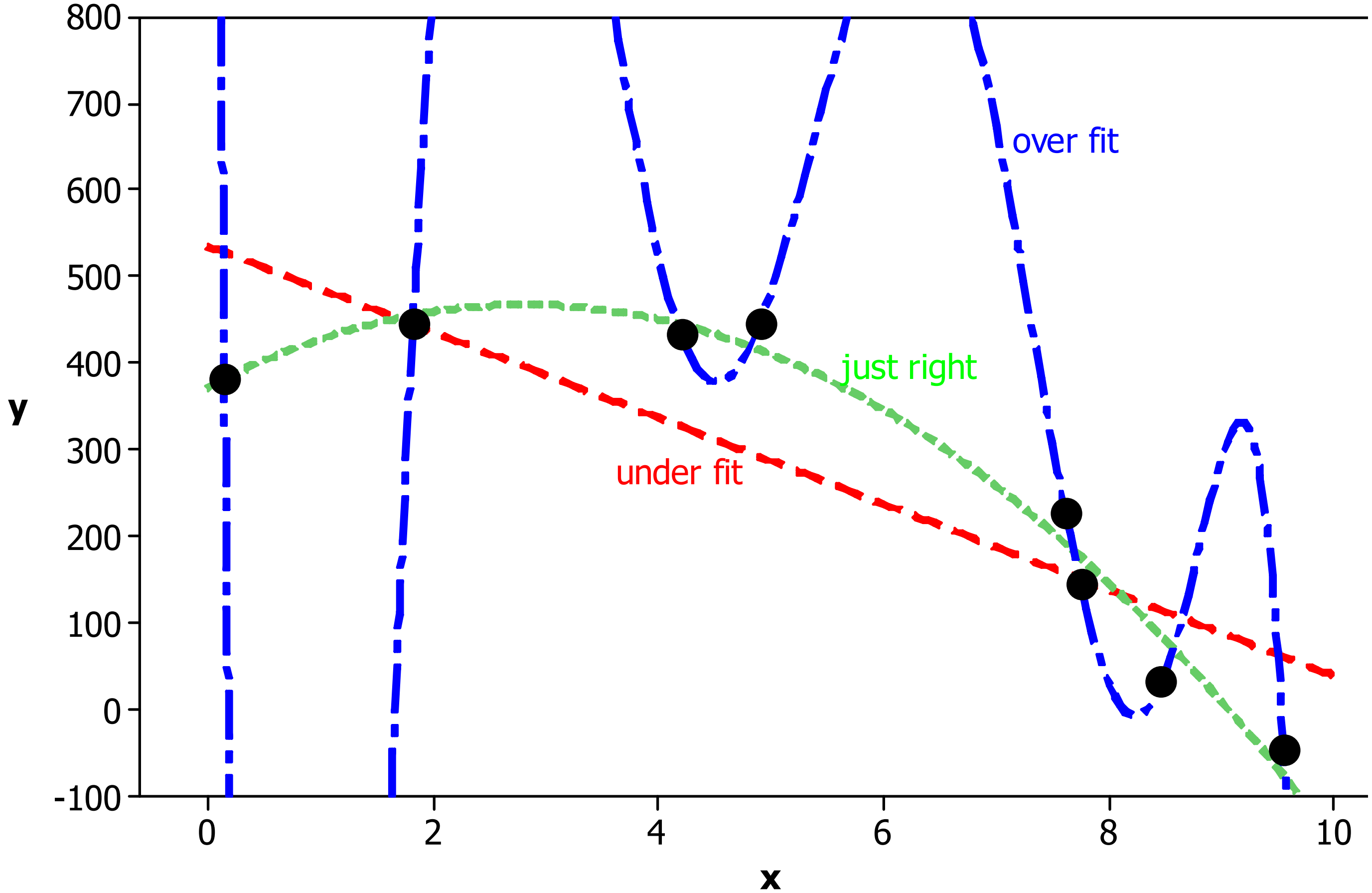
Occam's Razor



The simplest model that explains the data is the best model.

Occam's Razor

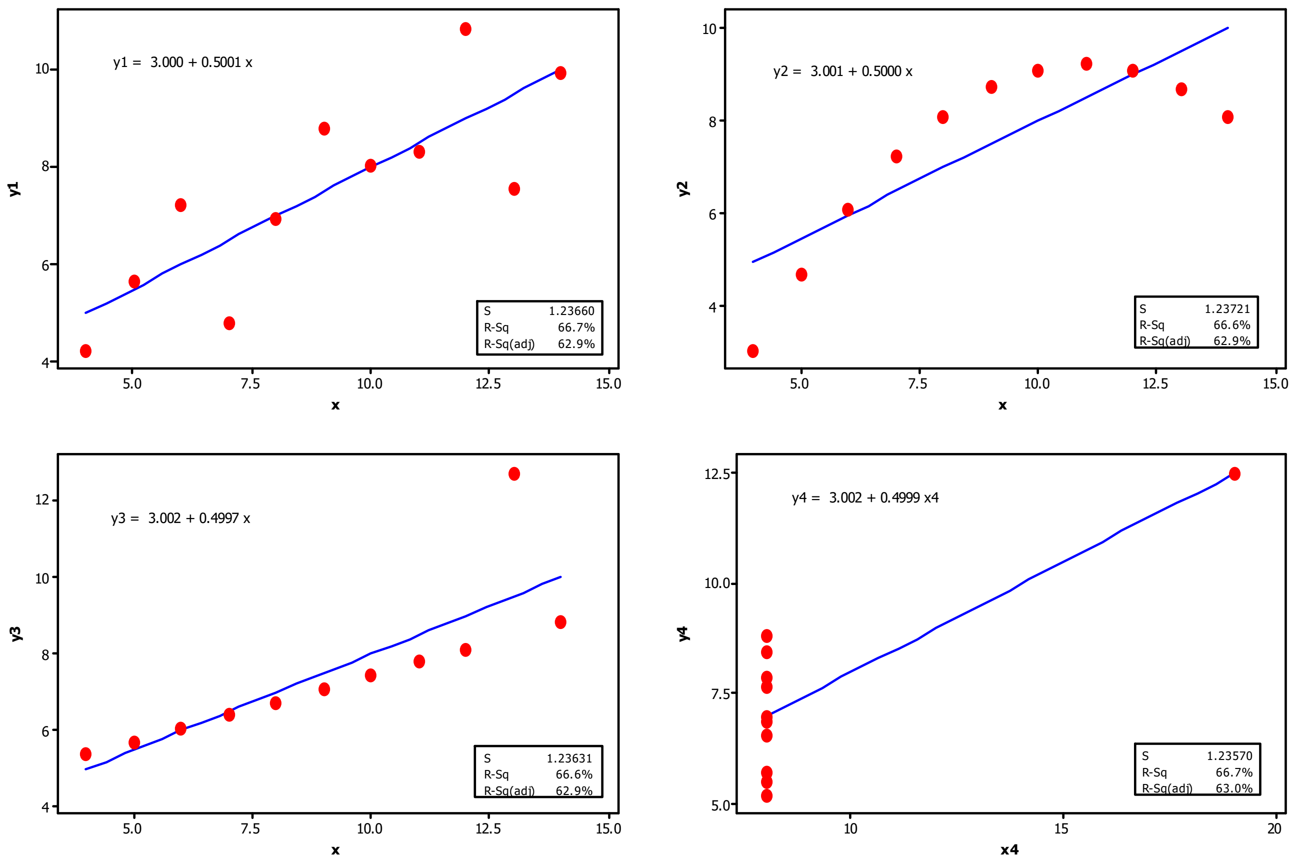
Choose the best model for the data:



"With four parameters I can fit an elephant, and with five I can make him wiggle his trunk." John von Neumann

Lack of Fit

- Don't rely on r^2 to assess lack of fit. r^2 tells how much of the total variation is explained by the model, not whether the model is any good.
- Inspection of a fitted line plot with an experienced eye provides an excellent test for lack of fit. The following data were created by Frank Anscombe to demonstrate this weakness of r^2:



Quantitative Linear Lack of Fit Test

One of the easiest lack of fit tests for a linear model is to fit a quadratic model:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2$$

and test the β_2 coefficient for statistical significance.

- Test $H_0 : \beta_2 = 0$ vs. $H_A : \beta_2 \neq 0$ using the test statistic $t = b_2 / s_{b_2}$. Accept $H_0 : \beta_2 = 0$ iff $-t_{\alpha/2,n-3} < t < t_{\alpha/2,n-3}$.
- If β_2 is not statistically different from 0 then there is no evidence of lack of fit.
- If β_2 is statistically different from 0 then there is evidence of lack of fit.

Problem: Evaluate the following model for lack of fit:

Data Display

Row	x	y
1	3	65
2	3	60
3	3	62
4	5	86
5	5	85
6	5	89
7	7	100
8	7	102
9	7	98
10	9	109
11	9	113
12	9	112
13	11	117
14	11	112
15	11	118

Regression Analysis

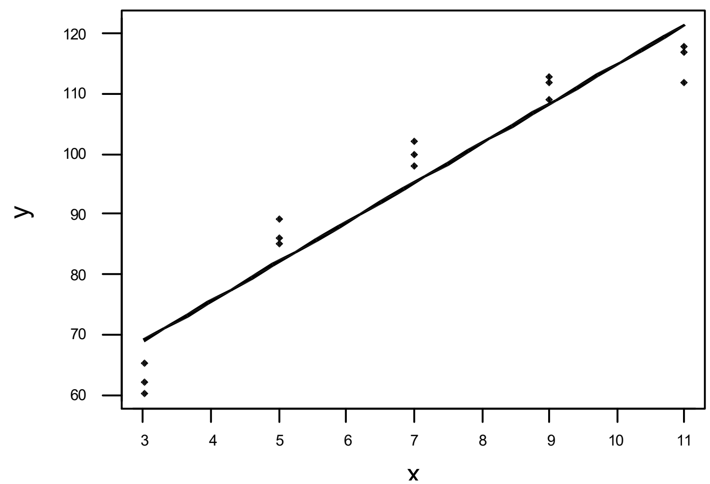
The regression equation is
y = 49.2 + 6.57 x

Predictor	Coef	StDev	T	P
Constant	49.233	4.054	12.14	0.000
x	6.5667	0.5370	12.23	0.000

S = 5.883   R-Sq = 92.0%   R-Sq(adj) = 91.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	5174.5	5174.5	149.53	0.000
Residual Error	13	449.9	34.6		
Total	14	5624.4			

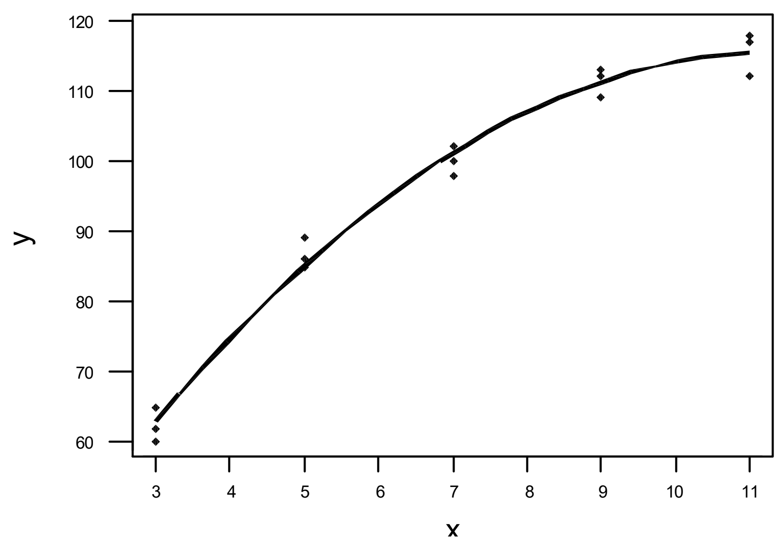


Data Display

Row	x	y	x2
1	3	65	9
2	3	60	9
3	3	62	9
4	5	86	25
5	5	85	25
6	5	89	25
7	7	100	49
8	7	102	49
9	7	98	49
10	9	109	81
11	9	113	81
12	9	112	81
13	11	117	121
14	11	112	121
15	11	118	121

Regression Analysis

```
The regression equation is
y = 18.5 + 17.1 x - 0.750 x2

Predictor        Coef       StDev          T        P
Constant       18.483       4.222       4.38    0.001
x              17.067       1.340      12.73    0.000
x2           -0.75000     0.09440      -7.94    0.000

S = 2.447       R-Sq = 98.7%     R-Sq(adj) = 98.5%

Analysis of Variance

Source            DF          SS          MS         F        P
Regression         2      5552.5      2776.3    463.57    0.000
Residual Error    12        71.9         6.0
Total             14      5624.4

Source       DF      Seq SS
x             1      5174.5
x2            1       378.0
```



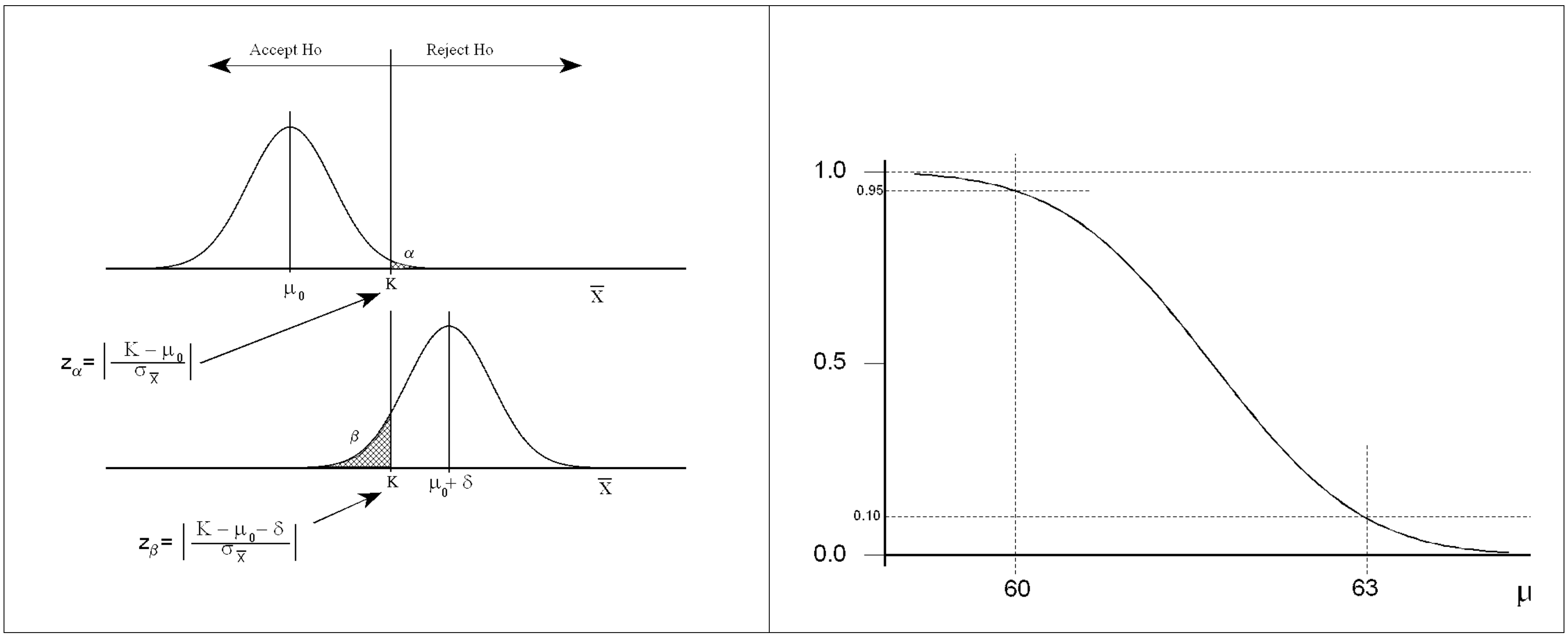
General Lack of Fit Strategy

- State the final model that you expect to claim.
- Propose a logical added complication to the model that would potentially fit the data better.
- Demonstrate that the added complication isn't statistically significant.
- Invoke Occam's Razor to remove the complication from the model.
- Accept the simplified model.

MINITAB Commands

- Stat> Regression and Correlation> Fitted Line Plot
- Graph> Scatterplot> With Regression
- Stat> Regression> Regression
- Stat> Regression> General Regression
- Stat> Basic Statistics> Correlation
- Calc> Calculator

Sample Size Calculations for Confidence Intervals and Hypothesis Tests



Sample Size Calculations

- Before you perform any experiment you should decide what statistical analysis you will perform on the data.
- Once the method of analysis has been identified a sample size calculation can be done to determine the unique number of observations required to obtain practically significant results.
- If too small a sample is used then your experiment may not detect an effect that is considered to be practically significant.
- If too large a sample is used then your experiment will be oversensitive and wasteful of resources.
- Round fractional sample sizes up to the nearest integer.

Sample Size for a Confidence Interval for the Population Mean

Conditions:

- σ known
- Distribution of x is Φ

Confidence Interval: The confidence interval will have the form:

$$\Phi(\bar{x} - \delta < \mu < \bar{x} + \delta) = 1 - \alpha$$

where

$$\delta = \frac{z_{\alpha/2}\sigma}{\sqrt{n}}$$

The value of δ should be chosen so that a single management action is indicated over the range of the confidence interval.

Sample Size for a Confidence Interval for the Population Mean

Sample Size: To be $(1-\alpha)100\%$ confident that the population mean μ is within $\pm\delta$ of the sample mean $\bar{x}$, the required sample size is:

$$n = \left(\frac{z_{\alpha/2}\sigma}{\delta}\right)^2$$

Example: Find the sample size required to estimate the population mean to within ± 0.8 with 95% confidence if measurements are normally distributed with standard deviation $\sigma = 2.3$.

Solution: The sample size required is:

$$\begin{aligned} n &= \left(\frac{z_{0.025}\sigma}{\delta}\right)^2 \\ &= \left(\frac{1.96 \times 2.3}{0.8}\right)^2 \\ &= 31.8 \rightarrow 32 \end{aligned}$$

Sample Size for the Confidence Interval for the Difference Between Two Population Means

Conditions:

- σ_1 and σ_2 are known and equal
- Distributions of x_1 and x_2 are Φ

Confidence Interval:

$$\Phi(\Delta\bar{x} - \delta < \Delta\mu < \Delta\bar{x} + \delta) = 1 - \alpha$$

where $\Delta\bar{x} = \bar{x}_1 - \bar{x}_2$ and $\Delta\mu = \mu_1 - \mu_2$.

Sample Size: To be $(1 - \alpha)100\%$ confident that the difference between two population means is within $\pm\delta$ of the difference in the sample means, the required sample size is:

$$n = 2\left(\frac{z_{\alpha/2}\sigma}{\delta}\right)^2$$

Example: What sample size should be used to determine the difference between two population means to within ± 6 of the estimated difference to 99% confidence. The populations are normal and both have standard deviation $\sigma = 12.5$.

Solution: The required sample size is:

$$
\begin{aligned}
n &= 2\left(\frac{z_{\alpha/2}\sigma}{\delta}\right)^2 \\
&= 2\left(\frac{2.575 \times 12.5}{6}\right)^2 \\
&= 57.6 \rightarrow 58
\end{aligned}
$$

Sample Size for a One-Sided Upper Confidence Interval for Fraction Defective

Conditions: The situation meets the requirements of the binomial distribution.

Confidence Interval:

$$P(0 < p < p_{\max}) = 1 - \alpha$$

Sample Size: To be $(1-\alpha)100\%$ confident that the population fraction defective is less than $p_{\max}$ the required sample size is:

$$n = \frac{\chi^2_{1-\alpha, 2(D+1)}}{2p_{\max}}$$

where D is the number of defectives found in the sample and $1-\alpha$ is the left tail area of the χ^2 distribution with $2(D+1)$ degrees of freedom. Sample sizes determined with this formula for 99% confidence intervals for $0 \le D \le 5$ and $0.0001 \le p \le 0.20$ are shown in the Appendix: *Number of Parts to Inspect with D Defectives to be 99% Confident That the True Process Fraction Defective is Less Than p*.

Example: What sample size must be inspected and found to be free of defectives to be 99% confident that the true population fraction defective is less than 3%?

Solution: The goal is to demonstrate:

$$P(0 < p < 0.03) = 0.99.$$

With $\alpha = 0.01$, $D = 0$, and $p_{\max} = 0.03$ the sample size is:

$$\begin{aligned} n &= \frac{\chi^2_{1-\alpha,2(D+1)}}{2p_{\max}} \\ &= \frac{\chi^2_{0.99,2}}{2(0.03)} \\ &= \frac{9.21}{2(0.03)} \\ &= 154 \end{aligned}$$

That is, if $n = 154$ parts are inspected and $D = 0$ are defective then we can be 99% confident that the true fraction defective is less than 3%.

Example: What sample size must be inspected if $D = 2$ defectives are found to be 95% confident that the true population fraction defective is less than 1%?

Solution: The goal is to demonstrate:

$$P(0 < p < 0.01) = 0.95.$$

With $\alpha = 0.05$, $D = 2$, and $p_{\text{max}} = 0.01$ the sample size is:

$$\begin{aligned} n &= \frac{\chi^2_{0.95,6}}{2(0.01)} \\ &= \frac{12.6}{2(0.01)} \\ &= 630 \end{aligned}$$

That is, if $n = 630$ parts are inspected and $D = 2$ are defective then we can be 95% confident that the true fraction defective is less than 1%.

Rule of Three

A very common special case of the one-sided upper confidence interval for the proportion defective is to demonstrate that

$$P(0 < p < p_{\max}) = 1 - \alpha$$

with a $c = 0$ sampling plan for which the sample size is given by

$1 - \alpha$	n
0.90	$2.3/p_{max}$
0.95	$3/p_{max}$
0.99	$4.6/p_{max}$

Example: What sample size for the $c = 0$ sampling plan is required to be 95% confident that the true population fraction defective is less than 1%?

Solution: By the rule of 3:

$$n = \frac{3}{p_{max}} = \frac{3}{0.01} = 300.$$

Sample Size for a Two-Sided Confidence Interval for a Proportion

Conditions: The situation meets the requirements of the normal approximation to the binomial distribution.

Confidence Interval: The confidence interval will have the form:

$$P(\hat{p} - \delta < p < \hat{p} + \delta) = 1 - \alpha$$

Sample Size: To be $(1 - \alpha)100\%$ confident that the population proportion is within $\pm\delta$ of the experimental proportion, the sample size must be:

$$n = \hat{p}(1 - \hat{p})\left(\frac{z_{\alpha/2}}{\delta}\right)^2$$

The largest sample size occurs when $\hat{p} = 0.5$ which may be taken as a worst case.

Example: Determine the sample size required to estimate the fraction of the voters who support one of two candidates in very close election to within 2% with 95% confidence.

Solution: With $\widehat{p} = 0.5$, $\delta = 0.02$, and $\alpha = 0.05$ the required sample size is:

$$\begin{aligned} n &= \widehat{p}(1 - \widehat{p})\left(\frac{z_{0.025}}{\delta}\right)^2 \\ &= 0.5(1 - 0.5)\left(\frac{1.96}{0.02}\right)^2 \\ &= 2401 \end{aligned}$$

Sample Size Calculations for Hypothesis Tests

Why bother?

- If the sample size is too small then a practically significant effect may go undetected.
- If the sample size is too large then the experiment may detect a statistically significant effect that is not practically significant.
- A sample size calculation effectively sets the statistical significance equal to the practical significance resulting in a unique sample size.
- An objective sample size calculation increases the likelihood that a statistically significant result is also practically significant.

Sample Size Calculations for Hypothesis Tests

- When determining sample size for hypothesis tests it is necessary to specify the conditions and probabilities associated with Type 1 and Type 2 errors.
- The *power* of a test given by:

$$\Pi = 1 - \beta$$

  is the probability of rejecting H_0 when H_A is true.
- A value of power is always associated with a corresponding value of *effect size* δ - the smallest practically significant difference between the population parameter under H_0 and H_A that the experiment should detect with probability Π.
- In all sample size calculations round n up to the nearest integer value.

Sample Size for a One-Sided Hypothesis Test of the Population Mean (σ_x known)

Conditions:

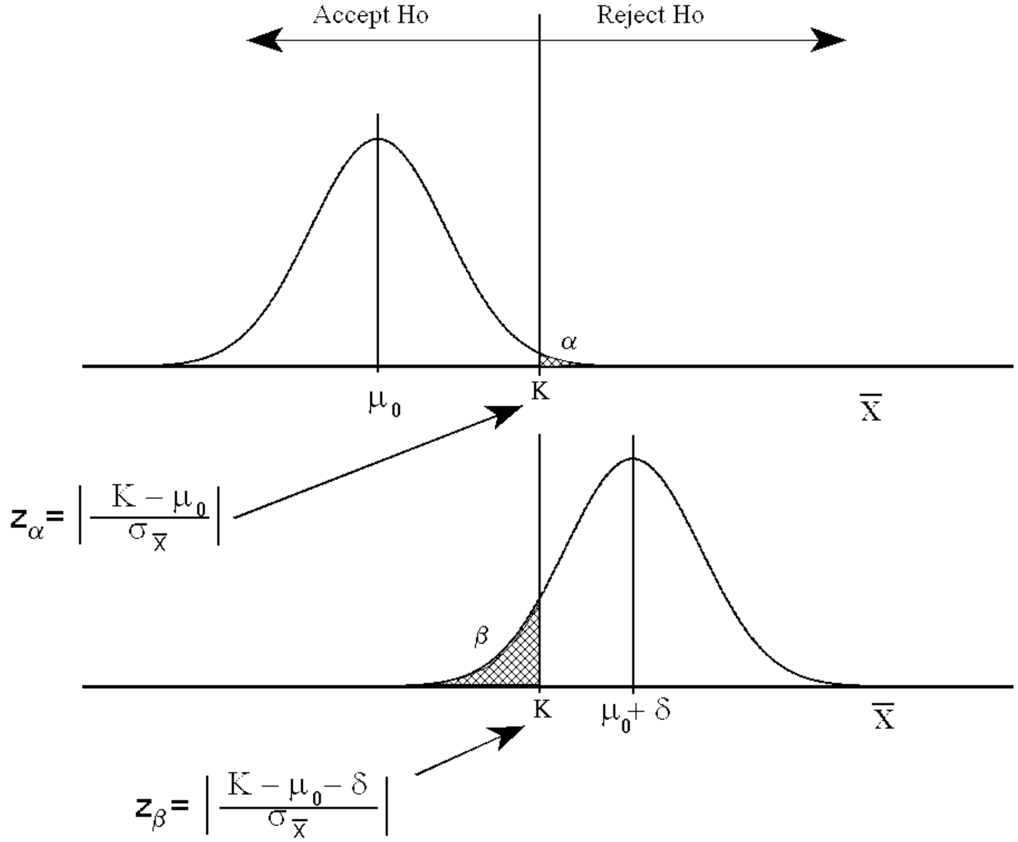
- σ_x is known
- x is normally distributed.

Hypotheses: $H_0 : \mu = \mu_0$ vs. $H_A : \mu > \mu_0$ or alternatively, $H_0 : \delta = 0$ vs. $H_A : \delta > 0$ where $\delta = \mu - \mu_0$.

Sample Size: The sample size required to obtain power $P = 1 - \beta$ for a shift from $\mu = \mu_0$ to $\mu = \mu_0 + \delta$ is given by:

$$n = \left(\frac{(z_\alpha + z_\beta)\sigma_x}{\delta} \right)^2$$

where z_α and z_β are both positive.



$$n = \left(\frac{(z_\alpha + z_\beta)\sigma_x}{\delta} \right)^2$$

$$K = \mu_0 + \delta \left(\frac{z_\alpha}{z_\alpha + z_\beta} \right)$$

Example: An experiment will be performed to determine if the burst pressure of a small pressure vessel is 60psi or if the burst pressure is greater than 60psi. The standard deviation of burst pressure is known to be 5psi and the experiment should reject $H_0 : \mu = 60$ with 90% probability if $\mu = 63$. Determine the sample size and acceptance condition for the experiment. The distribution of x is normal and use $\alpha = 0.05$.

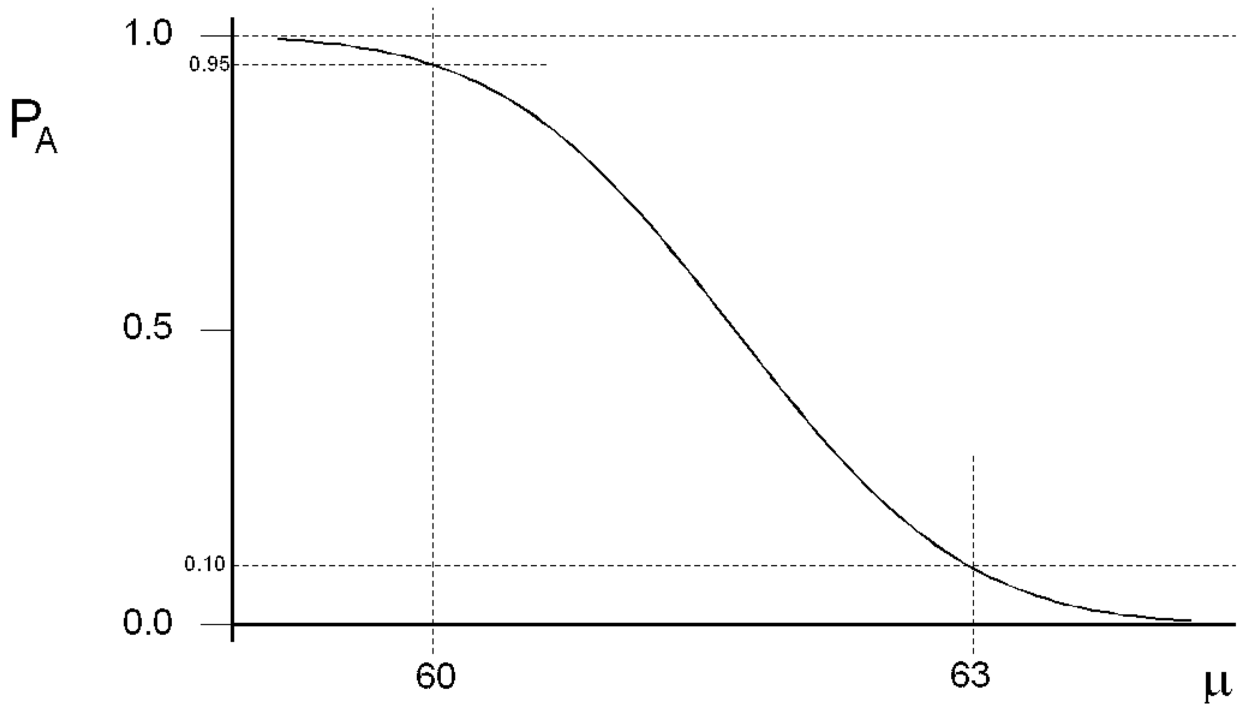
Solution: The hypotheses to be tested are $H_0 : \mu = 60$ vs. $H_A : \mu > 60$. The power of the experiment to reject H_0 when $\mu = 63$ or $\delta = 3$ is $P = 1 - \beta = 0.90$ so $\beta = 0.10$. The sample size is given by:

$$
\begin{aligned}
n &= \left(\frac{(z_{0.05}+z_{0.10})\sigma_x}{\delta}\right)^2 \\
&= \left(\frac{(1.645+1.282)5}{3}\right)^2 \\
&= 24
\end{aligned}
$$

The critical accept/reject value of $\bar{x}$ is given by:

$$\begin{aligned} K &= \mu_0 + \delta\left(\frac{z_{0.05}}{z_{0.05}+z_{0.10}}\right) \\ &= 60 + 3\left(\frac{1.645}{1.645+1.282}\right) \\ &= 61.69 \end{aligned}$$

The following graph shows the OC curve for the sampling plan:



Sample Size for a Two-Sided Hypothesis Test of the Population Mean (σ_x known)

Conditions:

- σ_x is known
- x is normally distributed.

Hypotheses: $H_0 : \mu = \mu_0$ vs. $H_A : \mu \neq \mu_0$ or alternatively, $H_0 : \delta = 0$ vs. $H_A : \delta \neq 0$ where $\delta = |\mu_0 - \mu|$.

Sample Size: The sample size required to reject $H_0 : \mu = \mu_0$ with probability $P = 1 - \beta$ for a shift from $\mu = \mu_0$ to $\mu = \mu_0 \pm \delta$ is given by:

$$n = \left(\frac{(z_{\alpha/2} + z_\beta)\sigma_x}{\delta} \right)^2$$

where $z_{\alpha/2}$ and z_β are both positive.

Example: Determine the sample size required to detect a shift from $\mu = 30$ to $\mu = 30 \pm 2$ with probability $P = 0.90$. Use $\alpha = 0.05$. The population standard deviation is $\sigma_x = 1.8$ and the distribution of x is Φ.

Solution: The hypotheses being tested are $H_0 : \mu = 30$ vs. $H_A : \mu \neq 30$. The size of the shift that we want to detect is $\delta = 2$ and we have $\sigma = 1.8$. Since $z_{\alpha/2} = z_{0.025} = 1.96$ and $z_\beta = z_{0.10} = 1.28$ the sample size required for the test is:

$$
\begin{aligned}
n &= \left(\frac{(z_{\alpha/2}+z_\beta)\sigma_x}{\delta}\right)^2 \\
&= \left(\frac{(1.96+1.28)1.8}{2}\right)^2 \\
&= 8.5 \rightarrow 9
\end{aligned}
$$

Sample Size Calculator Trick

- Compare the sample size calculations for the one-sample confidence interval and hypothesis test:

$$n = (z_{\alpha/2} \qquad)^2 \left(\frac{\sigma_\epsilon}{\delta} \right)^2$$

$$n = (z_{\alpha/2} + z_\beta)^2 \left(\frac{\sigma_\epsilon}{\delta} \right)^2$$

- Note that the sample size equation for the confidence interval can be obtained by setting $z_\beta = 0$ or $\beta = 0.5$ or $P = 0.5$.
- Many sample size calculators only provide methods for hypothesis tests because the corresponding calculations for confidence intervals can be obtained by setting $P = 0.5$.

Sample Size for Hypothesis Tests for the Difference Between Two Population Means

Conditions:

- σ_1 and σ_2 are both known and $\sigma_1 = \sigma_2$
- x_1 and x_2 are normally distributed

Hypotheses: $H_0 : \mu_1 = \mu_2$ vs. $H_A : \mu_1 \neq \mu_2$ or alternatively, $H_0 : \delta = 0$ vs. $H_A : \delta \neq 0$ where $\delta = |\mu_1 - \mu_2|$.

Sample Size: The sample size required to reject H_0 with probability $P = 1 - \beta$ for a difference between the means of $|\mu_1 - \mu_2| = \delta$ is given by:

$$n_1 = n_2 = 2\left(\frac{(z_{\alpha/2} + z_\beta)\sigma_x}{\delta} \right)^2$$

where $z_{\alpha/2}$ and z_β are both positive. For the one-sided tests replace $z_{\alpha/2}$ with z_α.

Example: Determine the common sample sizes required to detect a difference between two population means of $|\mu_1 - \mu_2| = \delta = 8$ with probability $P = 0.95$. Use $\alpha = 0.01$. The population standard deviation is $\sigma_x = 6.2$ and the distribution of x is Φ.

Solution: The hypotheses to be tested are $H_0 : \delta = 0$ vs. $H_A : \delta \neq 0$. We want to detect a difference between the two means of $\delta = 8$ with probability $P = 0.95$ so we have $\beta = 1 - P = 0.05$ so $z_\beta = z_{0.05} = 1.645$. For the two-tailed test we need $z_{\alpha/2} = z_{0.005} = 2.575$ so the required sample size is:

$$
\begin{aligned}
n_1 = n_2 &= 2\left(\frac{(z_{\alpha/2}+z_\beta)\sigma_x}{\delta}\right)^2 \\
&= 2\left(\frac{(2.575+1.645)6.2}{8}\right)^2 \\
&= 21.4 \rightarrow 22
\end{aligned}
$$

Sample Size for One-way ANOVA

Conditions: The treatments are normally distributed and homoscedastic.

Hypotheses: $H_0 : \mu_i = \mu_j$ for all i,j versus $H_A : \mu_i \neq \mu_j$ for at least one i,j pair

Sample Size: An approximate sample size for one-way ANOVA can be calculated by the two-sample t test method with Bonferroni's correction for multiple comparisons. The exact sample size calculation is complicated (requires use of the non-central F distribution).

Example: Determine the approximate number of replicates required for a one-way classification design to be analyzed by ANOVA with $k = 5$ treatments when the experiment should resolve a difference of $\delta = 30$ between two treatments with power $P = 0.90$ and the process standard deviation is expected to be $\sigma_\epsilon = 20$.

Solution: The number of multiple comparisons is $\binom{5}{2} = 10$. With family error rate $\alpha_{family} = 0.05$, the error rate for individual tests will be (by Bonferroni's correction) $\alpha = 0.05/10 = 0.005$. Then by the two-sample t test sample size method:

$$
\begin{aligned}
n &\simeq 2\left(\frac{(t_{\alpha/2} + t_{\beta})\sigma_{\epsilon}}{\delta}\right)^2 \\
&\simeq 2\left(\frac{(z_{0.0025} + z_{0.10})\sigma_{\epsilon}}{\delta}\right)^2 \\
&\simeq 2\left(\frac{(2.81 + 1.282)20}{30}\right)^2 \\
&\simeq 15
\end{aligned}
$$

There will be $df_{\epsilon} = 5(15) - 5 = 70$ error degrees of freedom so the approximation $t \simeq z$ is justified.

For comparison, the exact sample size (MINITAB: Stat> Power and Sample Size> One-way ANOVA) is also $n = 15$.

Single Sampling Plan Design for Defectives

Goal: Design the sampling plan (n, c) for defectives that accepts good lots and rejects bad ones.

Specify:

- p_0 - the low fraction defective that corresponds to acceptable lots
- p_1 - the high fraction defective that corresponds to rejectable lots
- $P_A(p_0)$ - the desired high probability (usually 95%) of accepting lots with fraction defective p_0
- $P_A(p_1)$ - the desired low probability (often 10%) of accepting lots with fraction defective p_1

Hypotheses Tested: $H_0 : p = p_0$ vs. $H_A : p > p_0$

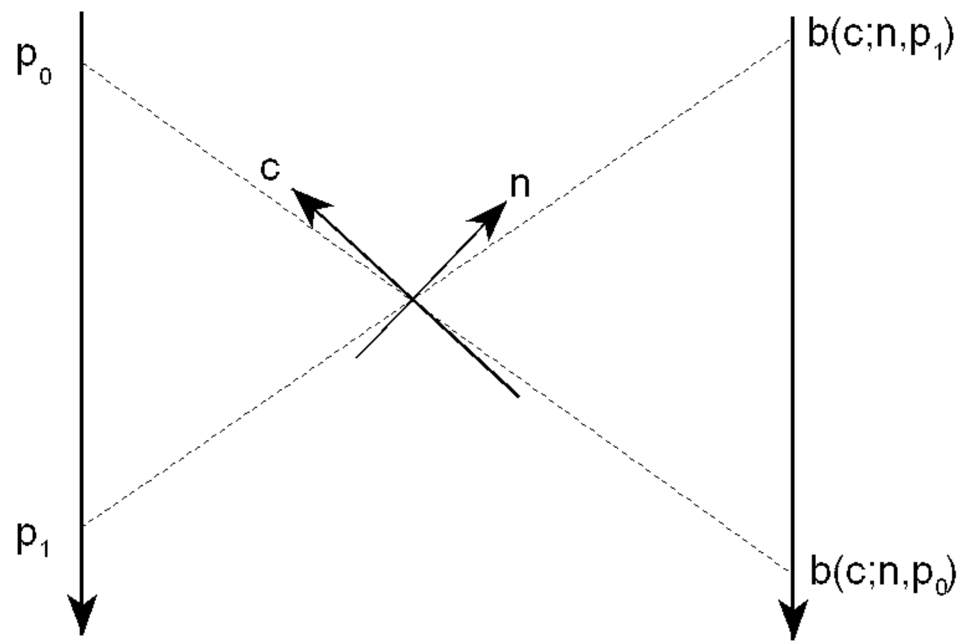
Single Sampling Plan Design for Defectives

Method: The distribution of defectives is binomial so the sampling plan must meet the conditions:

$$P_A(p_0) = b(c;n,p_0) = 1-\alpha$$

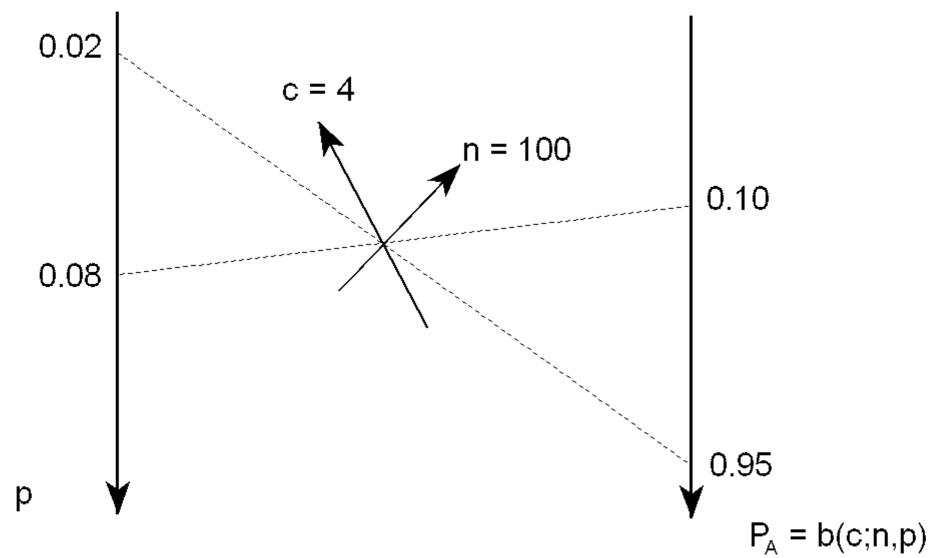
$$P_A(p_1) = b(c;n,p_1) = \beta$$

Find the solution using Larson's nomogram. (See Mathews, Design of a Single Sampling Plan for Defectives, in Appendix)



Example: Design the single sampling plan for defectives that will accept $(1-\alpha)100\% = 95\%$ of lots with acceptable fraction defective $p_0 = 0.02$ and will accept $(\beta)100\% = 10\%$ of lots with rejectable fraction defective $p_1 = 0.08$.

Solution:



The sampling plan is $(n, c) = (100, 4)$. Random samples of size $n = 100$ should be drawn from each lot and inspected. Lots with defectives $D \leq 4$ should be accepted and lots with $D > 4$ should be rejected.

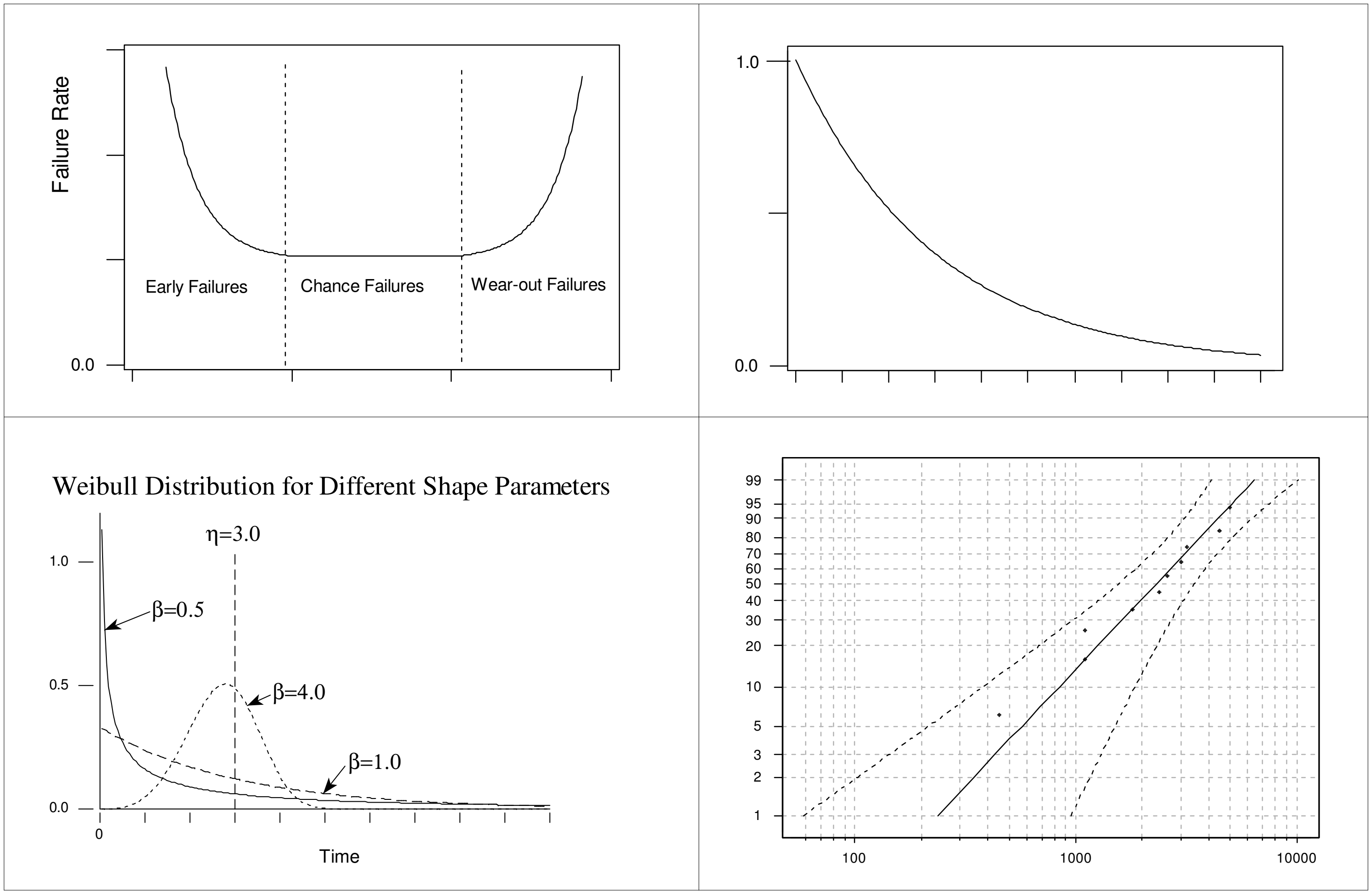
MINITAB Commands

Stat> Power and Sample Size

Stat> Quality Tools> Acceptance Sampling by Attributes

Stat> Quality Tools> Acceptance Sampling by Variables

Introduction to Reliability and Life Testing



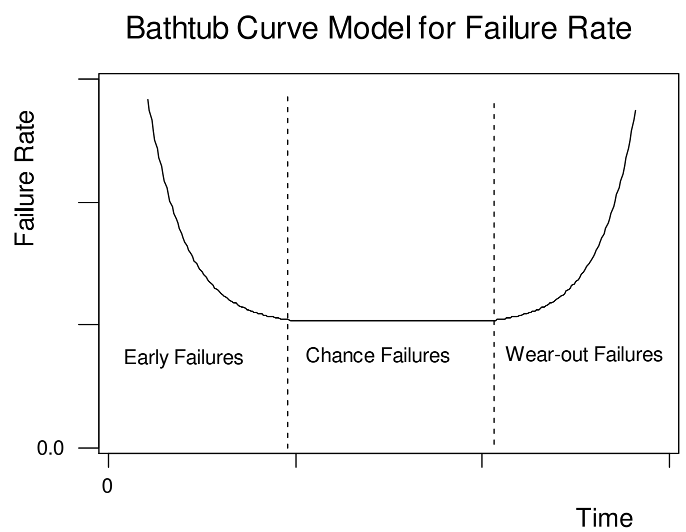
Definition of Reliability

The ability of an item to perform a required function under stated conditions for a stated period of time. The three characteristics that must be clearly identified in a reliability problem are:

1. The required function of the item.
2. The operating conditions.
3. The period of operation.

Failure Rate vs. Time

- The number of failures per unit time often follows a curve shaped like a bathtub.
- The high failure rate at the beginning is the early failure or infant mortality region of the bathtub curve.
- The relatively low and constant failure rate in the middle of the bathtub curve is the chance failure region.
- The high failure rate at the end of the bathtub curve is the wear-out failure or end-of-life failure region.



Probability Models for Reliability $R(t)$

The reliability $R(t)$ is the probability of survival to time t.

There are four common probability models for reliability:

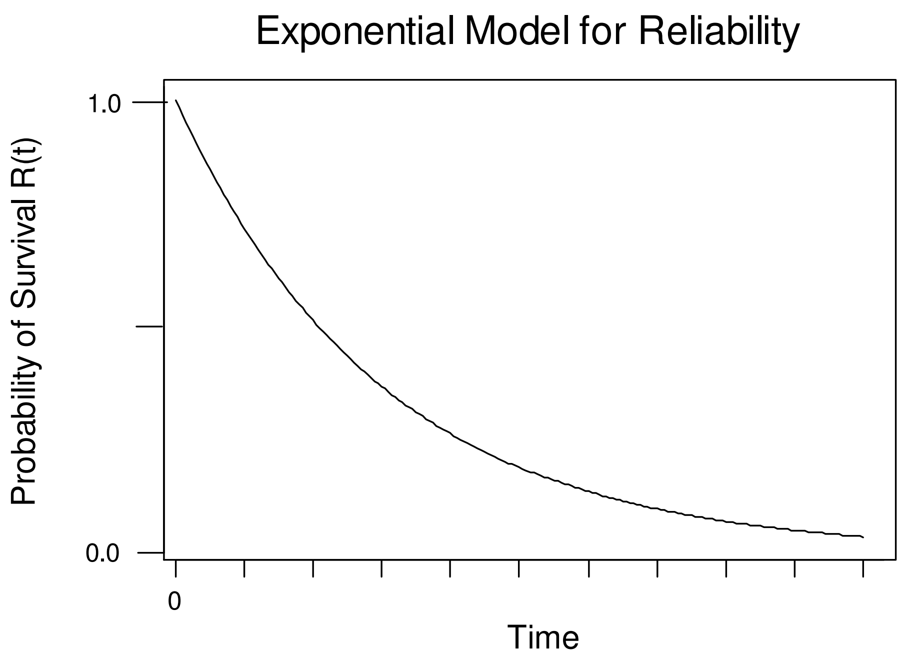
1. Exponential
2. Normal
3. Weibull
4. Log-Normal

The Weibull model is the most commonly used because it contains a shape factor that allows it to take on different shapes including exponential and approximately normal.

The Exponential Model for Reliability

$$R(t; \mu) = e^{-t/\mu}$$

where μ is the mean time to failure $(MTTF)$ and time t is in the chance failure region of the bathtub curve.



Example: Twenty components are operated to failure. They run a total of 14400 hours. Determine the 500 hour reliability of a component assuming an exponential model.

Solution: The mean time to failure is:

$$
\begin{aligned}
MTTF = \mu &= \frac{\sum t}{F} \\
&= \frac{14400}{20} \\
&= 720
\end{aligned}
$$

The 500 hour reliability (or the survival probability to 500 hours) is:

$$
\begin{aligned}
R(t = 500; \mu = 720) &= e^{-t/\mu} \\
&= e^{-500/720} \\
&= 0.50
\end{aligned}
$$

Notation: The notation Bxx is used to indicate the time at which $xx\%$ of the units failed. For this example, $B50 = 500$, or 50% of the units have failed at 500 hours.

Example: Twenty units are tested to failure for 100 hours each. When a unit fails it is repaired and put back into service. If a total of 4 failures are observed, find the *mean time between failures* ($MTBF$) and the 100 hour reliability assuming an exponential model.

Solution: The $MTBF$ is:

$$
\begin{aligned}
MTBF = \mu &= \frac{\sum t}{F} \\
&= \frac{20 \times 100}{4} \\
&= 500
\end{aligned}
$$

The 100 hour reliability is given by:

$$
\begin{aligned}
R(t = 100; \mu = 500) &= e^{-t/\mu} \\
&= e^{-100/500} \\
&= 0.819
\end{aligned}
$$

Example: Ten units are tested for 200 hours each except for 3 that fail at 80, 120, and 140 hours. Find the *MTTF* and the 300 hour reliability assuming an exponential model.

Solution: The *MTTF* is:

$$
\begin{aligned}
MTTF = \mu &= \frac{\sum t}{F} \\
&= \frac{80+120+140+7(200)}{3} \\
&= 580
\end{aligned}
$$

The 300 hour reliability is:

$$
\begin{aligned}
R(t = 300; \mu = 580) &= e^{-t/\mu} \\
&= e^{-300/580} \\
&= 0.596
\end{aligned}
$$

A Confidence Interval for the Mean Life

Let T_r be the accumulated life on test until the rth failure occurs. Then the $(1-\alpha)100\%$ confidence interval for the true population $MTTF$ is given by:

$$P\left(\frac{2T_r}{\chi^2_{1-\alpha/2}} < \mu < \frac{2T_r}{\chi^2_{\alpha/2}}\right) = 1-\alpha$$

where the χ^2 distribution has $2r$ degrees of freedom if the test is failure terminated or $2(r+1)$ degrees of freedom if the test is time terminated and the χ^2 distribution is indexed by its left tail area.

Example: Find a 95% confidence interval for the $MTTF$ if in a total of 2000 hours of life testing 4 failures occur. Assume the test is failure terminated.

Solution: We have $T_r = 2000$ and $r = 4$. From a χ^2 table with $2r = 2 \times 4 = 8$ degrees of freedom we have $\chi^2_{0.025,8} = 2.18$ and $\chi^2_{0.975,8} = 17.5$. The required confidence interval is:

$$
\begin{aligned}
P\left(\frac{2T_r}{\chi^2_{1-\alpha/2}} < \mu < \frac{2T_r}{\chi^2_{\alpha/2}} \right) &= 1 - \alpha \\
P\left(\frac{2 \times 2000}{17.5} < \mu < \frac{2 \times 2000}{2.18} \right) &= 0.95 \\
P(229 < \mu < 1835) &= 0.95
\end{aligned}
$$

Hypothesis Test for the Mean Life

Hypotheses Tested: $H_0 : \mu \leq \mu_0$ vs. $H_A : \mu > \mu_0$

Data Required: T_r and r

Decision: Calculate the test statistic:

$$\chi^2 = \frac{2T_r}{\mu_0}$$

Reject H_0 iff $\chi^2 > \chi^2_{1-\alpha,2r}$ if the test is failure terminated or iff $\chi^2 > \chi^2_{1-\alpha,2(r+1)}$ if the test is time terminated where α is the left tail area of the χ^2 distribution.

Example: Test to see if there is evidence that the *MTTF* is greater than 400 hours if in 2240 hours of life testing 3 failures occur. Use $\alpha = 0.05$, and assume the test was failure terminated.

Solution: The hypotheses being tested are $H_0 : \mu \leq 400$ vs. $H_A : \mu > 400$. We have $T_r = 2240$ and with $r = 3$ failures $\chi^2_{0.95,6} = 12.6$. The test statistic is:

$$\begin{aligned} \chi^2 &= \tfrac{2 \times 2240}{400} \\ &= 11.2 \end{aligned}$$

Since $(\chi^2 = 11.2) < (\chi^2_{0.95,6} = 12.6)$ we cannot reject $H_0 : \mu \leq 400$.

Demonstration Test - Rule of Three for Reliability

If the individual unit time on test will be equal to the endpoint time, then use the rule of three (see Ch. 10). This sample size is independent of the distribution choice and its parameter values.

Example: What sample size is required for the $c = 0$ sampling plan to demonstrate 98% reliability at 1000 hours with 95% confidence if the test units will be operated 1000 hours?

Solution: By the rule of 3:

$$n = \frac{3}{0.02} = 60.$$

That is, if $n = 60$ units are operated for 1000 hours each with zero failures then we can claim

$$P(0 < p < 0.02) = 0.95.$$

Demonstration Test - General

A life test experiment will be performed by putting n units on test for time t and observing the number of failures that occur within that time, i.e. the test is time-terminated. In order to demonstrate that the exponential mean life μ exceeds a specified value μ_0 with confidence $1 - \alpha$, that is:

$$P(\mu_0 < \mu < \infty) = 1 - \alpha$$

the test parameters must meet the condition:

$$b(c = r; n, p) \leq \alpha$$

where

$$p = 1 - R(t; \mu_0) = 1 - e^{-t/\mu_0}$$

Example: Determine the number of units that must be put on test for 200 hours without any failures to show that the $MTTF$ of a system exceeds 400 hours with 95% confidence. Assume that the life distribution is exponential and that the test is time terminated.

Solution: The goal of the experiment is to determine the value of n with $r = 0$ failures in $t = 200$ hours of testing such that:

$$P(400 < \mu < \infty) = 0.95$$

With $\mu_0 = 400$ the $t = 200$ hour reliability is:

$$R(t = 200; \mu_0 = 400) = e^{-\frac{200}{400}} = 0.6065$$

so the probability that a unit will fail before 200 hours is $p = 1 - 0.6065 = 0.3935$. With $r = 0$ and $\alpha = 0.05$ the smallest value of n that meets the condition:

$$b(0; n, 0.3935) \leq 0.05$$

is $n = 6$ since:

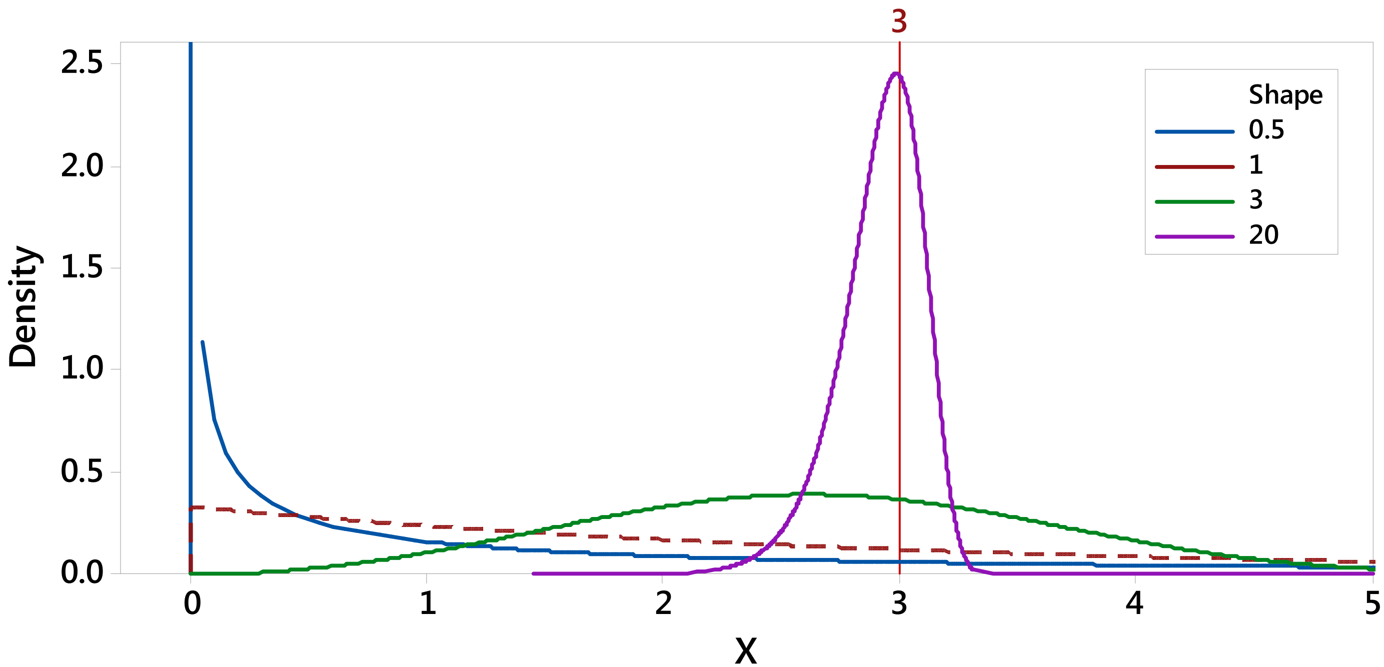
$$b(0; 6, 0.3935) = 0.04977$$

The Weibull Model for Reliability

- The probability of survival for Weibull process is:

$$R(t;\eta,\beta) = e^{-\left(\frac{t}{\eta}\right)^{\beta}}$$

  where η is the scale parameter and β is the shape factor.
- Weibull is exactly exponential for $\beta = 1$
- Weibull is approximately normal for $2.5 < \beta < 4$
- Weibull becomes asymmetric when β is very large



Weibull Analysis

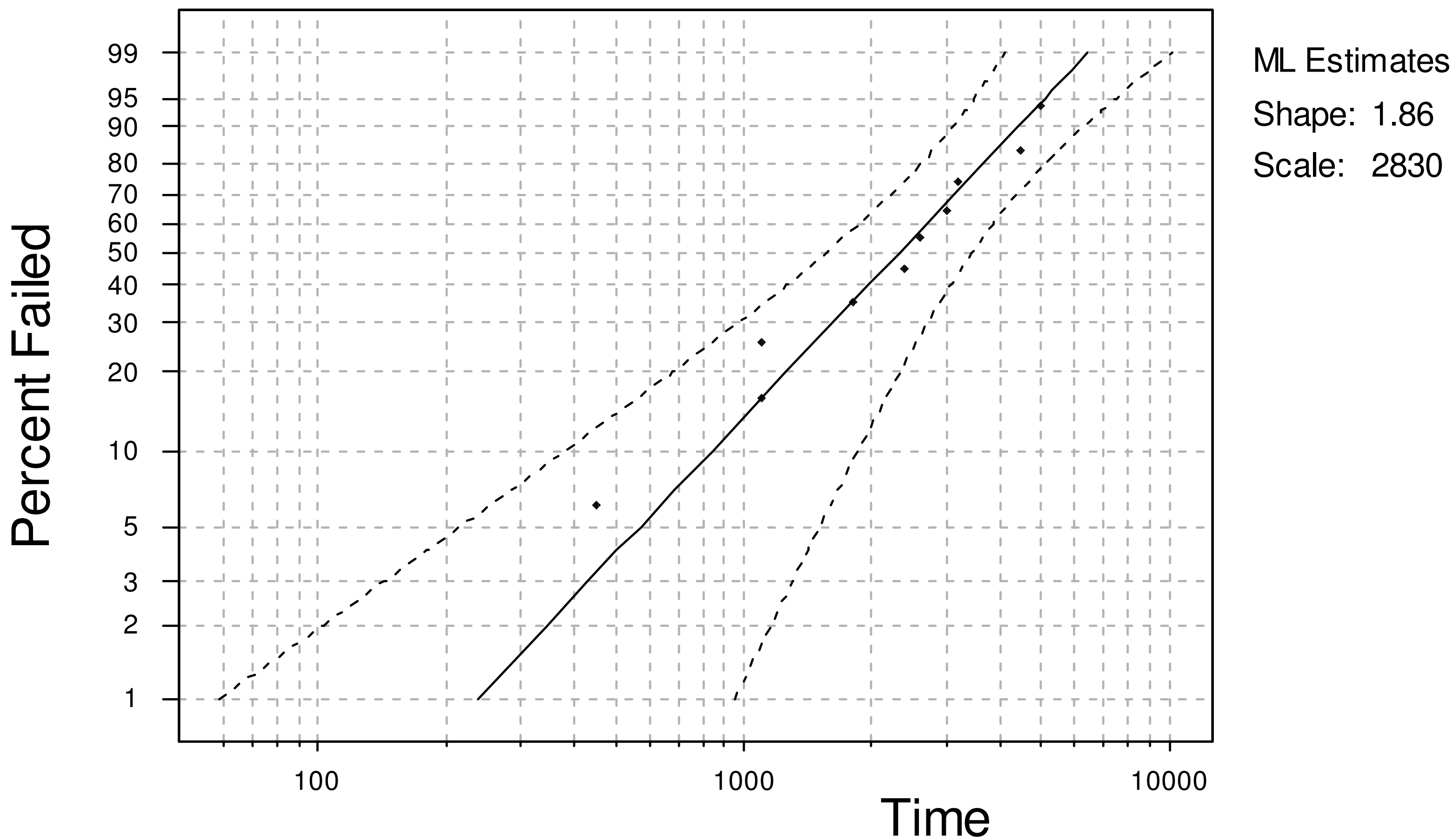
Analysis of Weibull life data is done graphically.

Example: Ten units were tested to failure. The failure times were $\{450, 1100, 1100, 1800, 2400, 2600, 3000, 3200, 4500, 5000\}$. Determine the 800 hour reliability.

Solution: The Weibull plot of the data is shown on the following page. From the analysis we have $\hat{\beta} = 1.86$ and $\hat{\eta} = 2830$. The 800 hour reliability will be:

$$
\begin{aligned}
R\left(t = 800; \hat{\beta} = 1.86, \hat{\eta} = 2830\right) &= e^{-\left(\frac{t}{\eta}\right)^{\beta}} \\
&= e^{-\left(\frac{800}{2830}\right)^{1.86}} \\
&= 0.91
\end{aligned}
$$

Weibull Probability Plot



Weibull Analysis with Censoring

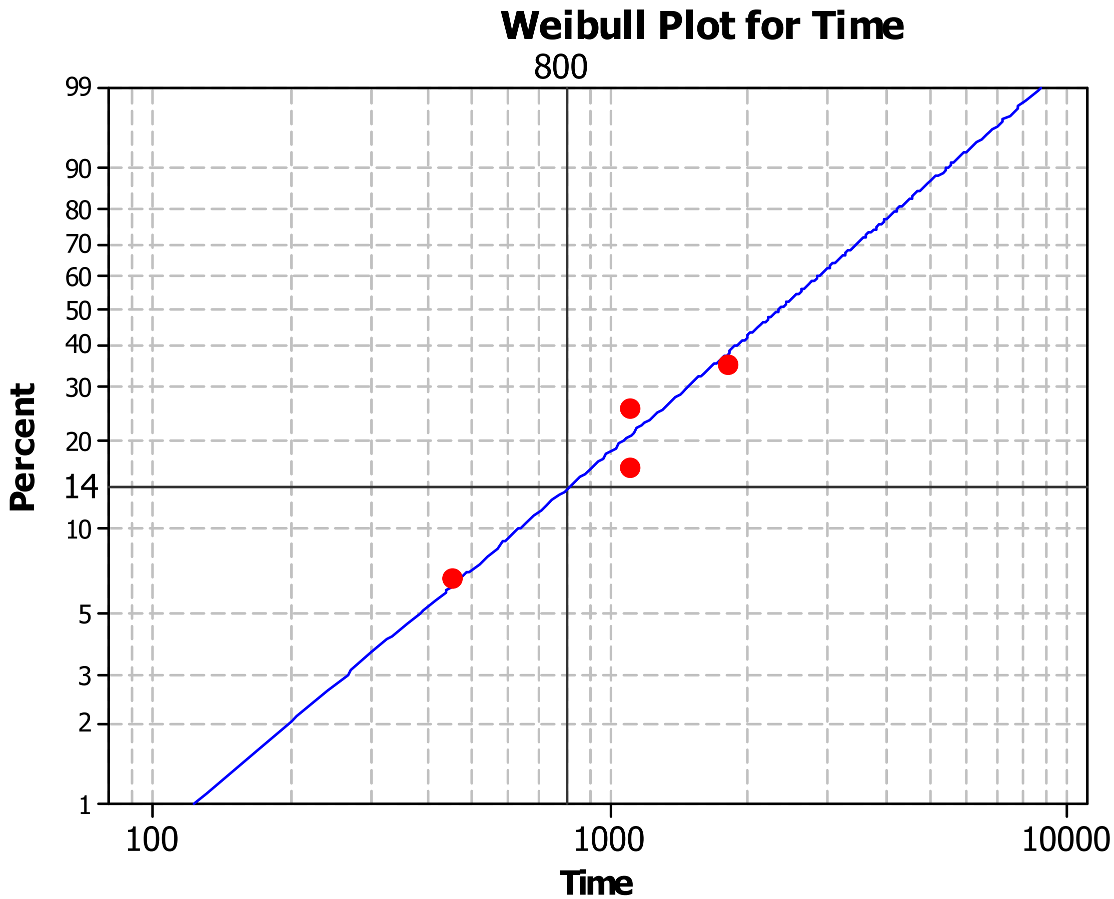
- The Weibull shape factor tends to be constant for a system. For example, most lamps in a single lamp family (e.g. metal halide, fluorescent, standard incandescent, ...) will have the same β.
- It is common to use the first couple of failures to estimate η and extrapolate the failure rate using the family β.
- You can analyze censored (incomplete) life data by plotting the percentage of the units that have failed for as many failures as have occurred.

Weibull Analysis with Censoring

Example: Suppose that in the preceding example the test was suspended at 2000 hours instead of being run until all ten units failed. Determine the 800 hour reliability for the right censored data.

Solution: The Weibull plot of the data is shown on the following page. From the analysis we have $\hat{\beta} = 1.86$ and $\hat{\eta} = 2830$. The 800 hour reliability will be:

$$\begin{aligned} R\left(t = 800; \hat{\beta} = 1.43, \hat{\eta} = 3043\right) &= e^{-\left(\frac{t}{\eta}\right)^{\beta}} \\ &= e^{-\left(\frac{800}{3043}\right)^{1.43}} \\ &= 0.86 \end{aligned}$$



Time	Censored
450	0
1100	0
1100	0
1800	0
2000	1
2000	1
2000	1
2000	1
2000	1
2000	1

Table of Statistics	
Shape	1.43358
Scale	3043.31
Mean	2763.87
StDev	1956.81
Median	2356.75
IQR	2545.90
Failure	4
Censor	6
AD*	43.132
Correlation	0.964

MINITAB Commands

- Stat> Reliability/Survival> Distribution Analysis (Right Censoring)
- Stat> Reliability/Survival> Distribution Analysis (Arbitrary Censoring)
- Graph> Probability Plot (complete data only)
- Graph> Probability Distribution Plot
- Stat> Reliability/Survival> Test Plans