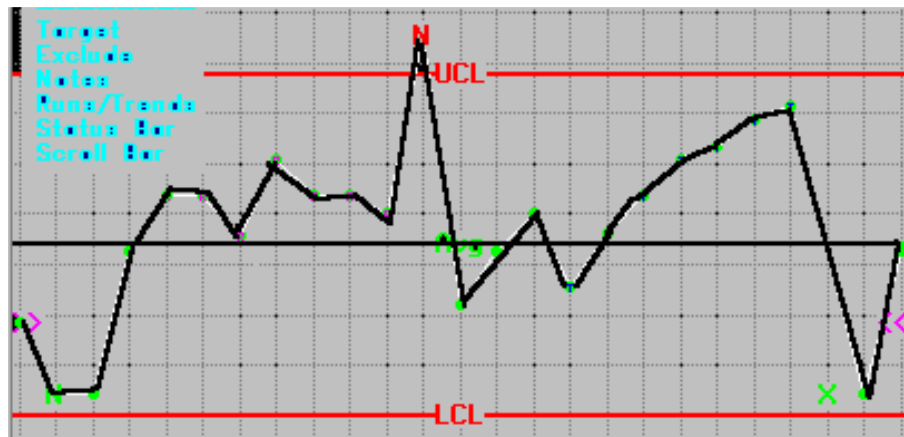


# Statistical Process Control (SPC)



## NOTICE

This is a shareware version of our seminar notes on the above topic. This document is placed in the public domain for free download & distribution by practitioners of quality improvement tools and techniques (QITT). You are allowed to make use of this documents as is and without any support.

**Nutek, Inc.** (Since 1987)

*Quality Engineering Seminar and Software*  
Bloomfield Hills, MI, USA. 2014



## Table of Contents

Introduction to Continuous Improvement and Statistical Process Control .....	Module 1
Basic Statistics for Statistical Process Control .....	Module 2
Sources of Variations: Common and Special Causes .....	Module 3
Distribution and Statistics for Normally Distributed Data.....	Module 4
Evaluation of Capability for a Stable Process .....	Module 5
Average and Range Charts .....	Module 6
Interpretation of Control Charts .....	Module 7
Strategies for Sampling Data .....	Module 8
Average and Standard Deviation Charts .....	Module 9
Median and Range Charts .....	Module 10
Individual and Moving Range Chart .....	Module 11
Proportion and Number of Nonconforming Charts .....	Module 12
Number of Nonconformities and Nonconformities per unit Charts .....	Module 13
Application Guidelines .....	Module 14
Application Examples .....	Module 15
Appendices .....	Module 16

## Short List of Notations

$x$  = a number (data, sample reading).

$\bar{X}$  = the average (mean) of a set of  $x$ .

$\bar{\bar{X}}$  = the average of the averages  $\bar{X}$ , also called grand average.

$\tilde{X}$  = the median of the numbers  $x$ . (Symbol  $\approx$  used for median)

$\bar{\tilde{X}}$  = the average of the medians

$\sum X$  = the sum of numbers  $x$

$n$  = the total number of data

$K$  = the number of samples (if 100 parts are grouped into 20 groups of 5 each for measurement purposes, then  $K = 20$ ,  $n = 5$  for group,  $n = 100$  for all data.)

$M$  = the mode of numbers

$S$  = the standard deviation of numbers  
( $S_X$ ,  $S_{\bar{X}}$ ,  $S_R$ , etc.).

$\bar{S}$  = the average of the Standard deviation ( $S$ )

$R$  = the range of numbers

$\bar{R}$  = the average of the ranges ( $R$ ).

$\mu_X$  = The average of all possible number (even of parts that are not included in measurement)

$\sigma_X$  = the standard deviation of all possible numbers

$<$  = less than

$\leq$  = less than or equal to

$>$  = greater than

$\geq$  = greater than or equal to

$\sqrt{\quad}$  = square root of quantity under the radical sign

## Course Overview

Industrial production often involves mass production of the same part. For such parts to properly assemble and function in the final product, it is necessary to keep the variation in quality characteristic to a minimum. The variation in quality characteristics are caused mainly by two sources; known as common and special causes. Statistical process control (SPC) is used to study the process performance and understand sources of variation with the intention of making corrective actions to reduce variation.

This brief session covers the basic concepts of statistical analysis and their application to practical problems in process control. It will deal with such standard tools as histograms, X-bar and R charts, process capability studies and sampling plans. It is intended to help attendees decide if SPC will be helpful their activities and whether further training should be sought before applications. Attendees to this session learn how to use histograms, Pareto charts, scatter diagrams, etc. The primary focus of the course will be to help attendees develop a working level understanding of the normal distribution and how to construct control charts for variable and attribute data. The attendees are expected to gather understanding of how to use DOE results for SPC, calculate process capabilities, and learn how to communicate with design engineering in terms of process capabilities.

### Instructor

Ranjit K. Roy, Ph.D., P.E. (M.E.) is an engineering consultant specializing in Taguchi approach of quality improvement. Dr. Roy has achieved international recognition as a consultant and trainer for his down-to-earth teaching style of the **Taguchi experimental design technique**, **project management**, and several **other quality engineering topics**. Dr. Roy began his career as senior design engineer with Burroughs Corporation following completion of graduate studies in engineering at the University of Missouri-Rolla in 1972. He then worked for General Motors Corp. (1976-1987) assuming various engineering responsibilities with his last position as that of reliability manager. Dr. Roy is a *fellow* of the American Society of Quality.



**SPC  
Learning Steps**

**STEP 4**

Learn what are:  
 - In control  
 - Stable  
 - In specification  
 - Capable

**STEP 5**

Understand the logic behind standardization of data collection and simplification of formula for calculations of limits for control charts.

**STEP 6**

Learn the rules of out-of-control detection based on probability of occurrence. Examine control charts carefully

**STEP 3**

Understand that process variation is the result of COMMON-CAUSE and SPECIAL-CAUSE variations (Fundamental Equation of SPC)

**STEP 2**

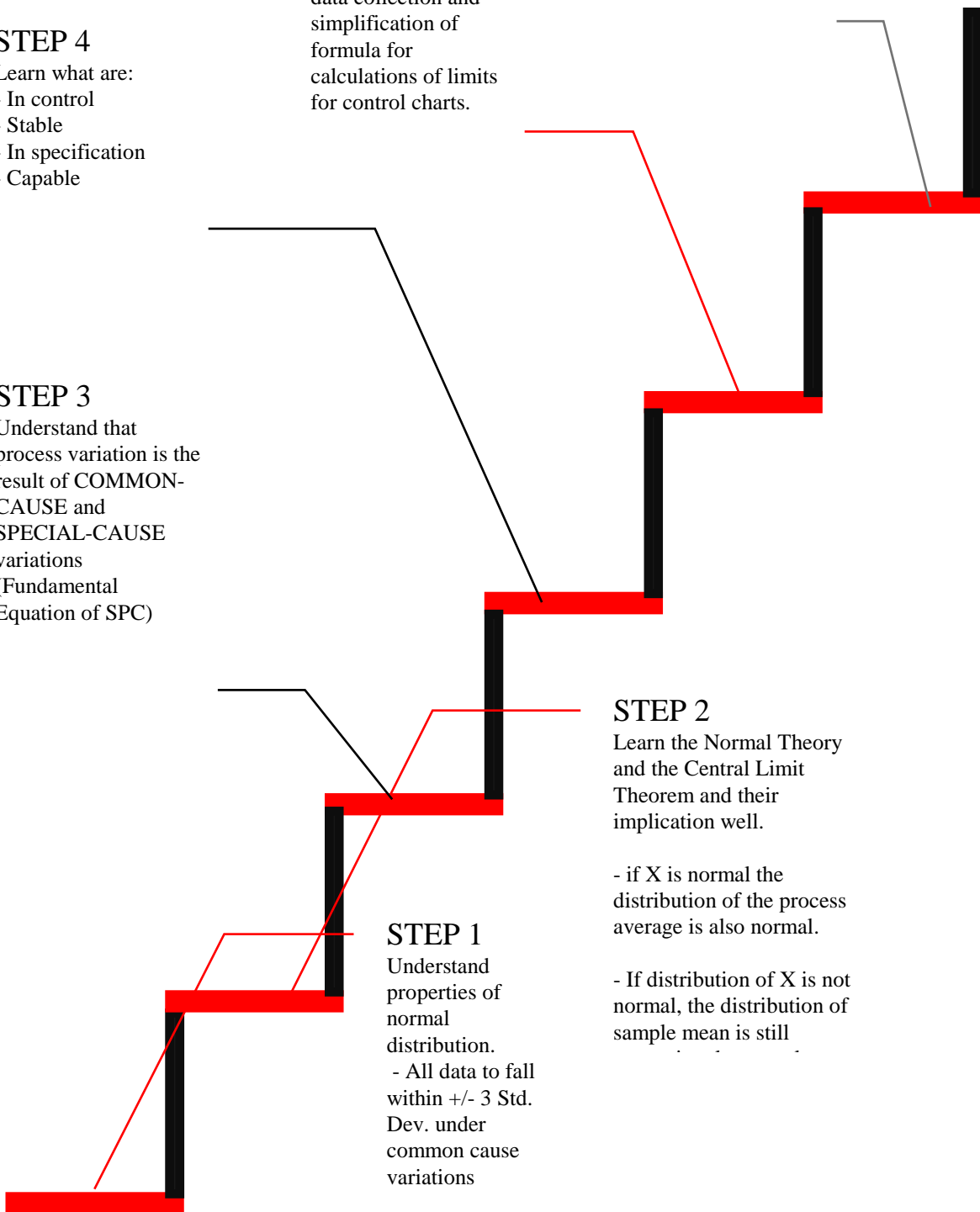
Learn the Normal Theory and the Central Limit Theorem and their implication well.

- if X is normal the distribution of the process average is also normal.

- If distribution of X is not normal, the distribution of sample mean is still

**STEP 1**

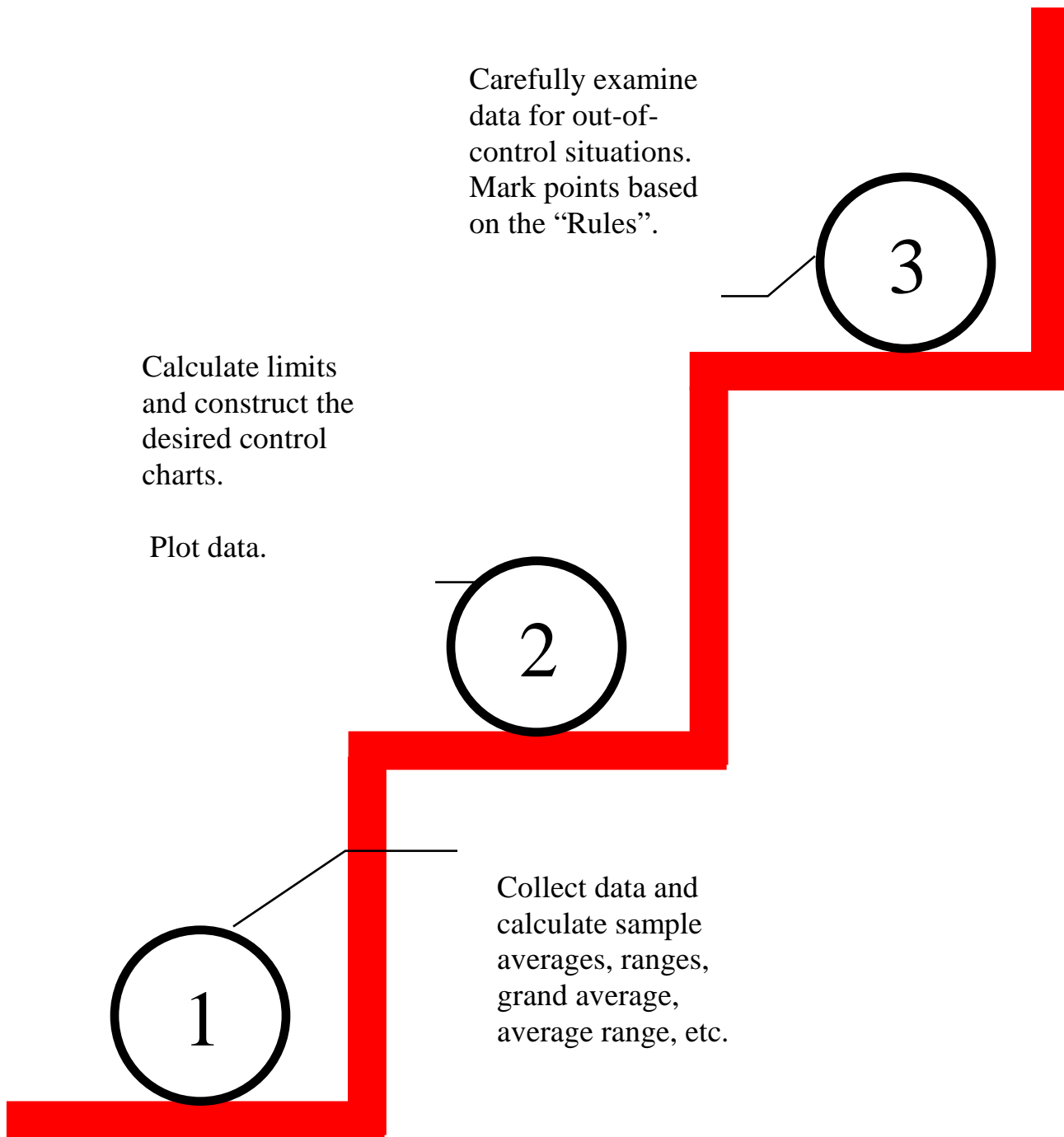
Understand properties of normal distribution.  
 - All data to fall within +/- 3 Std. Dev. under common cause variations



## SPC

### Application Steps

(Observe the Process - Define Standard -Compare Performance)



*What is*  
Statistical Process Control ?

## Statistical

What does it mean?

***Inexact:***

It is never the exact value.

***Predicted value based on more than one observed data:***

Data used for calculating any statistic (numeric value) are more than one.

***Estimate based on a small data sample from the whole:***

***Valid for average performance:***

***Conclusions based on past observations:***

## Process

What does it mean?

***Any task that affects cost and performance of products or services:***

- Incoming order processing
- Payroll checks
- Accounts payable
- Part fabrication (milling, turning, welding, molding, etc)
- Typing/data entry
- Assembly of parts
- Soldering

***Any operation or sequence of operation with measurable output:***

## Control

What does it mean?

***To correct:***

- Correct to what?
- How do would we know that it is not right?
- How bad are we?
- How far off is it from what it should be?
- How fast do we need to correct the situation?

***To keep it where it should be:***

- How do we find out where should it be?
- What can we do to bring it where it should be?
- How do we know it is not where it should be?

# **Module 1**

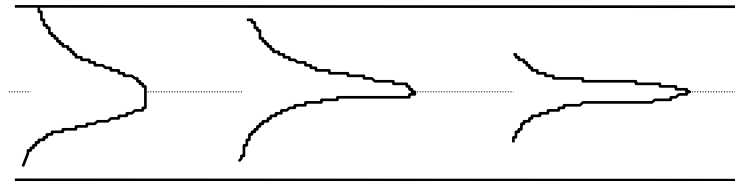
## **Introduction to Continuous Improvement and Statistical Process Control**



## Introduction to Continuous Improvement and Statistical Process Control

### What is Quality?

- Meeting Specifications
- Preventing Defects
- Practicing Continuous Improvement



“Quality is meeting or exceeding customer expectations at all times in regards to performance, durability, reliability, and delivery at a cost that represents value.”

“Quality is producing a reliable & durable product to exact standard at the lowest cost for the customers.”

### How?

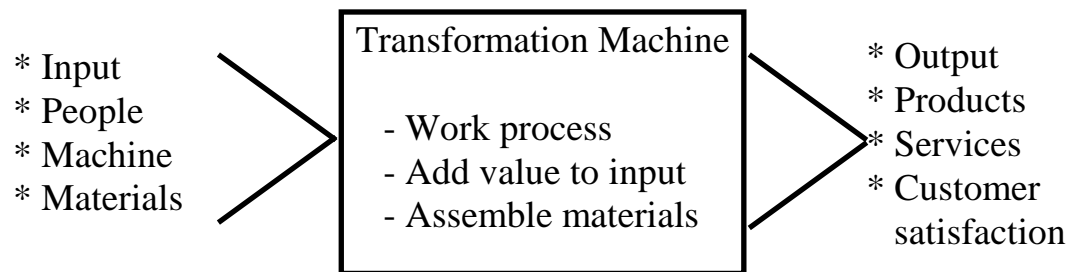
- **Product Control**
- **Process Control**

In 1970’s W. Deming, J. Juran and others emphasized the need to shift focus from product control to process control.

Stop depending on inspection to achieve quality. Eliminate the need for extensive inspection programs by designing and building quality into the product or service in the first place.

- W. Edward Deming

### A Sample Process



### Product Control Vs. Process Control

Items/Activities	How it is done	
	Product Control	Process Control
Practical Objectives	Keep Product Characteristics within specifications	Try to be on target with minimum variation satisfying the quality standard
Ideal Objectives	Detect and eliminate non-conforming products	Assure that non-conforming products are not produced
Improvement Strategy	Inspect and maintain outgoing product quality	Institute and maintain continuous quality improvement procedures
Statistical Techniques	Acceptance Sampling	Control Charts (SPC)

# **Module 2**

## **Basic Statistics for Statistical Process Control**

## Types of Data (Measured Data)

- **Variable Data**

- Diameter (mm)
- Length (inches)
- Weight (pounds)
- Torque (Inch-Pound)
- Tension (pound)
- Hardness, etc.

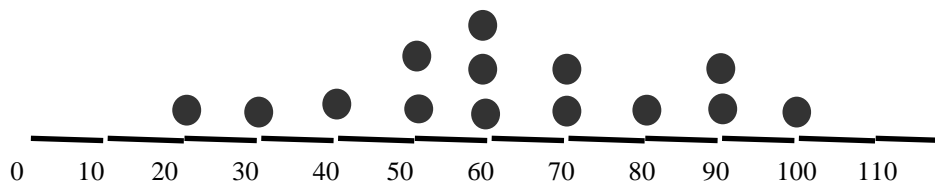
- **Attribute Data**

- Go/No go (% of no go)
- Conforming/Non-conforming (number non-conforming)

**Example:** Defects found by inspecting 1000 parts.

Day 1	$X_1 = 50$	Day 6	$X_6 = 70$	Day 11	$X_{11} = 90$
Day 2	$X_2 = 60$	Day 7	$X_7 = 60$	Day 12	$X_{12} = 30$
Day 3	$X_3 = 48$	Day 8	$X_8 = 90$	Day 13	$X_{13} = 42$
Day 4	$X_4 = 20$	Day 9	$X_9 = 80$	Day 14	$X_{14} = 60$
Day 5	$X_5 = 71$	Day 10	$X_{10} = 60$	Day 15	$X_{15} = 100$

### Data plot in a number scale



## Histogram

### Shape of the Histogram - Distribution

Let us use the following data to define the tools that measure a distribution.

**Example:** Sample readings of a few used batteries were as follows:

Battery no. 1	$X_1 = 6$	Volts
Battery no. 2	$X_2 = 3$	Volts
Battery no. 3	$X_3 = 5$	Volts
Battery no. 4	$X_4 = 4$	Volts
Battery no. 5	$X_5 = 9$	Volts
Battery no. 6	$X_6 = 6$	Volts
Battery no. 7	$X_7 = 11$	Volts

Notice the notations used to indicate a specific data. For example, the third data point or the third **observation**, 5, is described using the notation

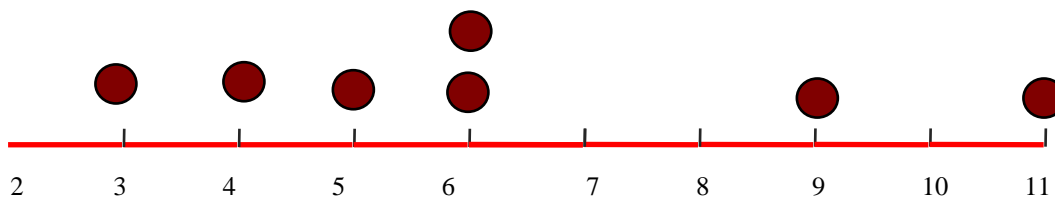
$$X_3 = 5$$

Notation is useful in explaining methods used and communicating ideas without actually using a specific number. However, most often we write data set without using notations like in this case the data set is:

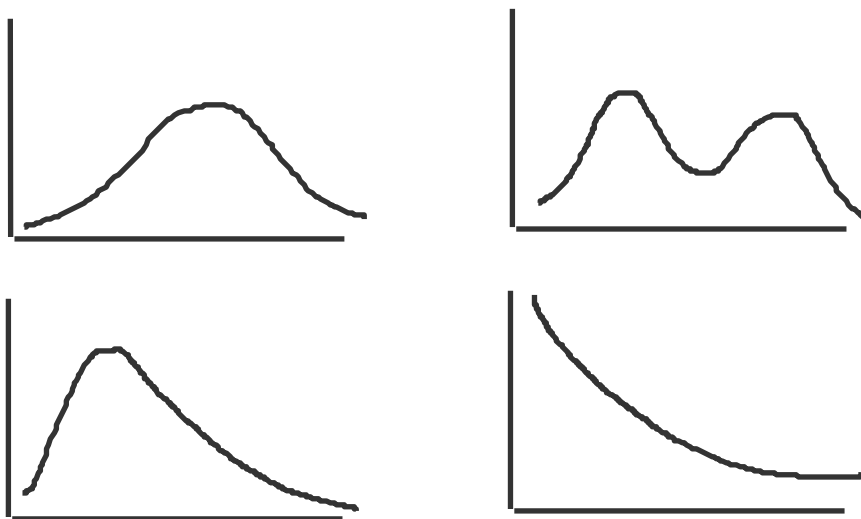
6 3 5 4 9 6 11

### Shape of Data

The shape of the data is the shape of its distribution and is obtained by plotting the data points in a histogram, which can be drawn by placing one ball for each data point on the corresponding value along the x-axis.



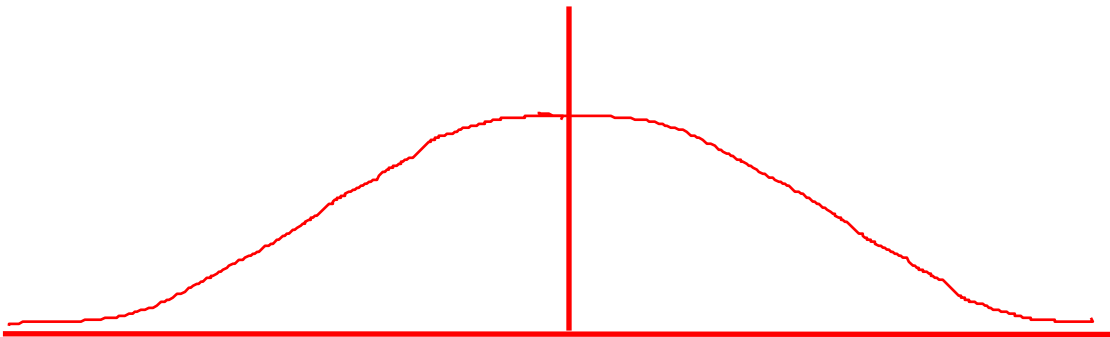
### Shapes of Distributions



### Normal Distribution (ND - Bell shaped curve)

- There are a large number of situations that exhibit ND
- In absence of any special causes many data will show ND
- A vast majority of industrial product and process data follow ND

Example: Human heart beat, Diameter of a machined shaft, Voltage of a particular brand of battery, etc.



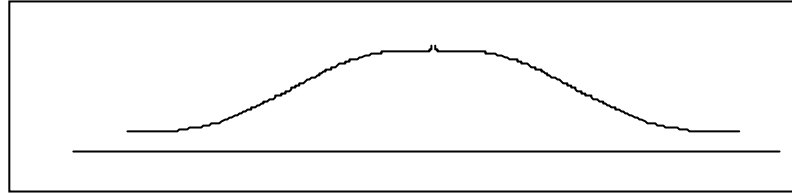
Predicting Shapes of Distributions ...

- \* Weights of apples in a bad year
- \* Grade point averages of freshman year in reputed schools
- \* Ages of people attending meetings of retirement planning



## Statistical Tools for analysis of data sets

There are a number of statistical tools are used for extracting information hidden in a set of numbers.



Questions which may arise while reviewing a Normal Distribution

Where is the stem of the “umbrella”?

How wide is it?

What shape is it?

Statistically speaking, how do we measure:

- **Central Tendency or location of the center**
  - Mean or Average (\*)
  - Median
  - Mode
  
- **Variation or Spread or closeness of data to the center**
  - Range (\*)
  - Variance
  - Standard Deviation (\*)
  
- **The Shape**
  - Symmetry
  - Skewness
  - Flatness
  - Similarity to normal distribution (\*)

(\*) More popular in SPC studies

1. Location of the center -

=

measured by Average (  $\bar{X}$  ), Median (x), and Mode

### Average, Mean or Balance point of data

Mean of the data is obtained by simply adding all data and dividing by the total number of data.

$$\text{Mean or Average } \bar{X} = (6 + 3 + 5 + 4 + 9 + 6 + 11) / 7 = 6.285$$

### Median

The median of the data is the observation, which is in the middle of the data set when the entire observation is placed in ascending order, i.e., from smallest to the largest.

data: 3      4      5      6      6      9      11

The observation, 6, which is in the 4th place of the total of 7 observations, is the median of the data.

$$\tilde{X} = 6 \quad (\text{notation uses curl over the character X})$$

When there are even number of observations, the median will be the average of the two closest numbers.

For a set of data: 9 10 12 5 19 20, median is the number between 12 and 15.

$$\tilde{X} = (12 + 15) / 2 = 13.5$$

### Mode

Mode is the number that occurs most frequently in a set of numbers. In this case the number 6 occurs twice, whereas the other numbers occur only once. The mode in this case is 6.

### Exercise:

Data: 5      7      2      7      3      9

Find:

$$\text{Average} = \text{Ans: } 5.5$$

$$\text{Median} = \text{Ans: } (5+7)/2 = 6$$

$$\text{Mode} = \text{Ans: } 7$$



**2. Closeness of data to the center -**  
measured by Range (R), Variance ( $S^2$ ), and Standard Deviation (S)

**Range, Variance, and Std. Deviation**

data:            2            5            3            7            4

**Range**

$$\begin{aligned} R &= \text{Difference between the largest and smallest data} \\ &= 7 - 2 = 5 \end{aligned}$$

**Variance**

$$S^2 = \frac{\text{Sum of Squares of distance from the average}}{\text{Number of observations minus 1}}$$

$$\text{average} = (2 + 5 + 3 + 7 + 4)/5 = 4.2$$

$$\begin{aligned} S^2 &= [(2 - 4.2)^2 + (5 - 4.2)^2 + (3 - 4.2)^2 + (7 - 4.2)^2 + (4 - 4.2)^2]/(5 - 1) \\ &= (4.84 + 10.24 + 1.44 + 10.24 + 1.44)/(5-1) \\ &= 7.05 \end{aligned}$$

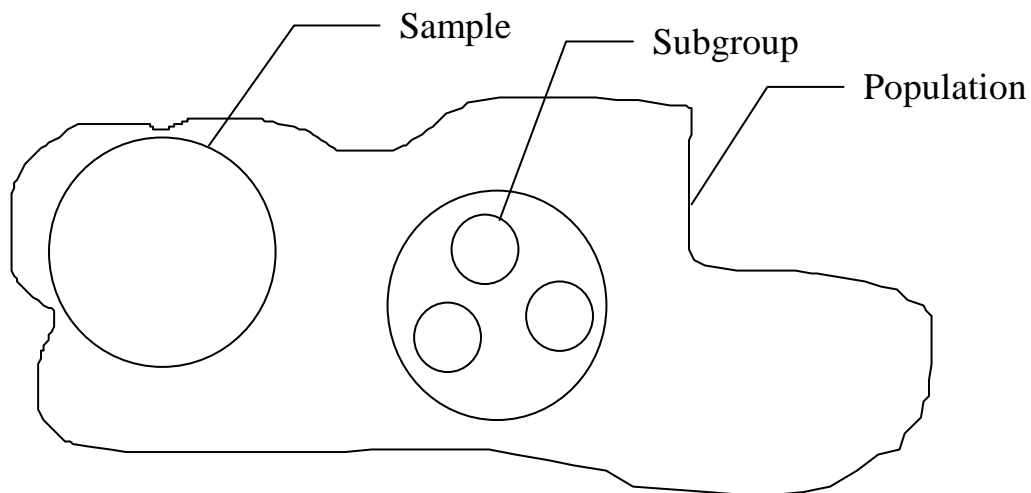
$$\text{Standard Deviation}(s) = \sqrt{\text{variance}}$$

$$S = \sqrt{S^2} = \sqrt{7.05} = 2.655$$

**3. Shape of the curve -**  
measured by Symmetry, Skewness, Flatness, and Similarity to ND

In SPC studies data is assumed to be normally distributed.

## Subgroup, Sample, and Population



**Sample** - a selected set of data from the total number of possible data. Consider a milling machine producing a total of 60,000 parts. If only 20 sets of 5 parts ( $5 \times 20 = 100$ ) out of the total are inspected, then:

**Sample size** = 100,

**Subgroup size** = 5, subgroups are smaller sample (20 subgroups)

**Population** is 60,000.

Generally for statistical calculations, the sample size is subdivided into smaller groups. The subgroup size of 5 or 10 data are common. With **5 data in a subgroup**, 100 data sample is divided into 20 subgroups.

### Subgroup, Sample and Population Statistics

Data:

Subgroup 1:            12 14 11 16 12 Avg = 13

Subgroup 2:            12 14 12 15 16 Avg = 13.8

Subgroup 3:            11 16 13 15 12 Avg = 13.4

### Subgroup averages

$$\bar{x}_1 = 13, \bar{x}_2 = 13.8, \bar{x}_3 = 13.4$$

$$\text{Sample average } \bar{X} = (13 + 13.8 + 13.4)/3 = 13.4$$

(Grand average)

Also, the average of all data, that is

$$(12 + 14 + 11 + 16 + 12 + 12 + 14 + 12 + 15 + 16 + 11 + 16 + 13 + 15 + 12)/15$$

$$= 201/15 = 13.4 \text{ (same as the average of all averages)}$$

### Subgroup Ranges

$$R_1 = 16 - 11 = 5$$

$$R_2 = 16 - 13 = 4$$

$$R_3 = 16 - 11 = 5$$

$$\text{Sample Range} = (R_1 + R_2 + R_3)/3 = (5 + 4 + 5)/3 = 4.666$$

(average range)

Likewise standard deviation for the sample can be calculated.

### Practice Problems (Fill in the missing answers)

1. Data:	7	4	3	8	6	Avg. = 5.599	Std. Dev.= 2.073	Range =
2. Data:	12	10	11	9	8	Avg. = 10	Std. Dev.=	Range = 4
3. Data:	2	4	3	5	5	Avg. = 3.8	Std. Dev.=	Range = 3
4. Data:	14	18	15	12	23	Avg. =	Std. Dev.=4.277	Range =
5. Data:	9	11	13	14	12	Avg. =	Std. Dev.= 1.923	Range =
6. Data:	6	9	12	11	12	Avg. = 10	Std. Dev.=	Range = 6

# **Module 3**

## Sources of Variations: Common and Special Causes

Sources of Variations: Common and Special Causes

Variation is natural and is expected in every process or production output.

Sources of variations

- **People**
  - Operator fatigues
  - Operator experience on the job
  - Training
- **Equipment**
  - Setup
  - Old/New machine
  - Maintenance
- **Production method**
  - Rate of production
  - Inspection & downtime
- **Materials**
  - Uniformity
  - Change of property
- **Environment**
  - Temperature
  - Humidity
  - Dust in the air

Depending on the source, the causes of variations are classified into two groups.

- **Common-cause variation**
  - Difficult to reduce by ordinary approach of fine-tuning. It usually require significant changes to the process
  - The primary responsibility in reducing this variation rests with the management
- **Special-cause variation**
  - Often possible to reduce by making adjustment at the point of production
  - Process operator and immediate management usually have the primary responsibility

## Fundamental Equation of Statistical Process Control

$$\text{Total variation of production output} = \text{Special-cause Variation} + \text{Common-cause Variation}$$

One of the objectives of SPC is to reduce variation in production output.

- Identify common-cause variation and special-cause variation
- Remove special-cause variation
- Reduce common-cause variation

“I should estimate that in my experience most troubles and most possibilities for improvement add up to proportions something like this:

- **94% are common-cause: they belong to the system (responsibility of management)**
- **6% are special-cause variation (responsibility of operators) ”**

### **Control chart** (to be studied later)

A control chart shows variation in a production process and it contains information, which can be used to detect the presence of special-cause variation.

The control chart for a process utilizes information from the process when it is only subjected to common-cause variation. When actual performance is monitored and data plotted in the control chart, information about the possible presence of special-cause variation in the process is revealed.

## Process stability

- A stable process has only common-cause variation and no special-cause variation.
- Most processes that are stable show distributions that are normal or close to being normal.
- If a process is stable and shows normal distribution for individual values of the output, or for the subgroup averages and ranges, the process is considered to be in control
- Special-cause variation causes instability in the process and must be corrected before the SPC data is collected

## In Control

A process is in control if over time there is no change in the distribution of the process output i.e., there is no change in the process center and there is no change in the process variation.

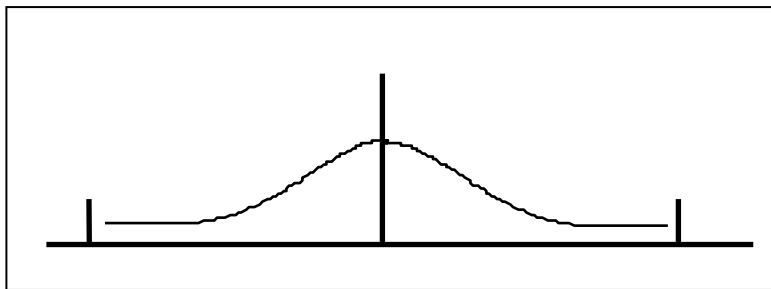
Mean/Average	Variation	
No change	No change	In control and Stable
Change	No Change	Out of control and Unstable
No Change	Change	“
Change	Change	“

## In Specification

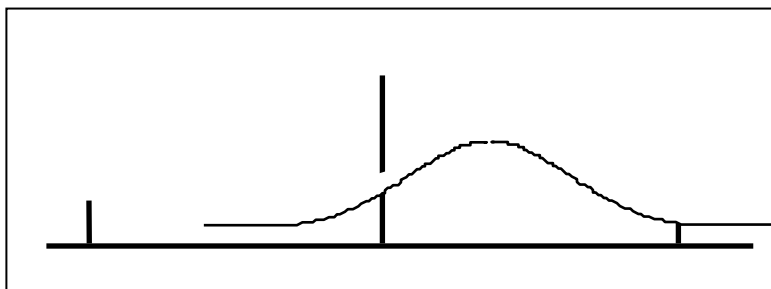
A process is considered in specifications when the distribution is within the lower (LCL) and upper (UCL) control limits.

A stable process can be:

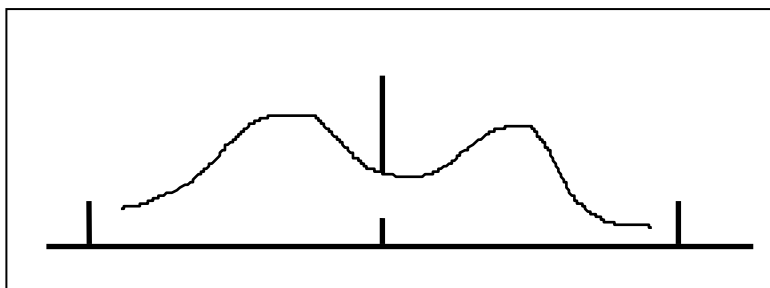
- **In control and in spec.**



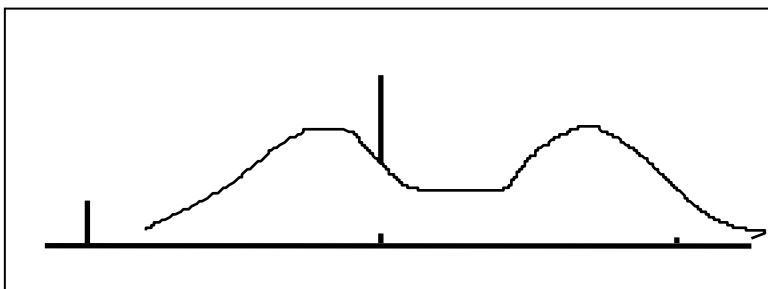
- **In control but not in spec.**



- **Out of control but in spec.**



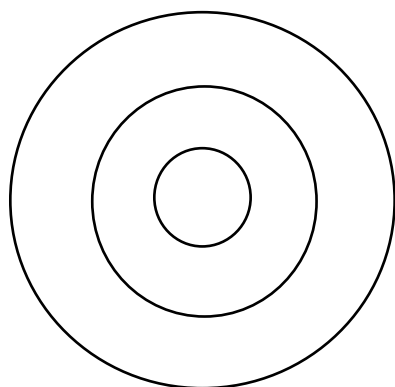
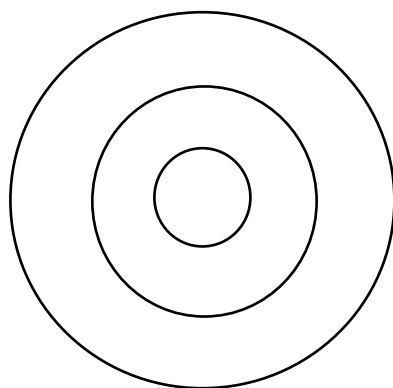
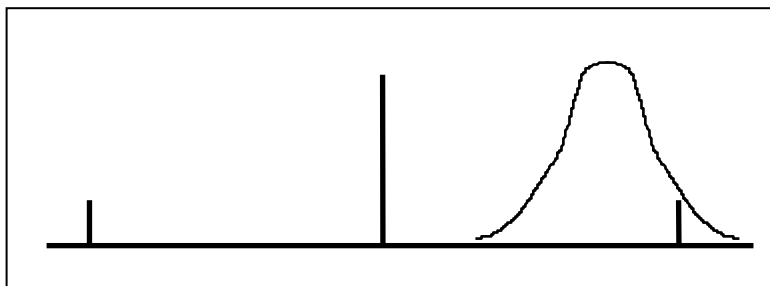
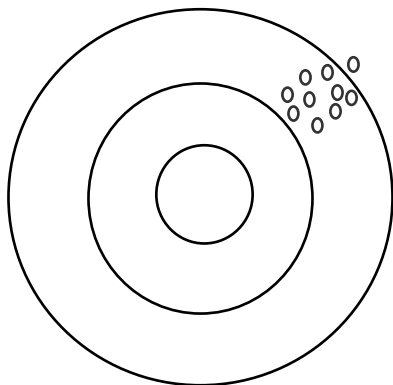
- **Out of control and out of spec.**





Process capability (to be studied later)

A process can be in control but not in spec but capable.



# **Module 4**

## Statistics for Normally Distributed Data

## Statistics for Normally Distributed Data

### Normal Distribution (ND)

Distribution - refers to the shape of the histogram created from a set of numbers (data).

Data - diameter of a shaft, ages of people, delivery times, number of faults, etc.

Histogram - a bar graph made using the frequency of occurrence of data.

Frequency - a number that represents how many times a particular number occurs in a data set. In a data set if number 5 occurs three times then frequency of occurrence of number 5 is 3.

Normal - a natural or usual behavior.

Data collected from process under normal operating condition will follow a specific type of distribution. Mathematicians using different mathematical equations have modeled a few common distribution shapes: Normal, Binomial, Exponential, etc. The data appear to follow close similarities with normal distribution (ND).

#### Characteristics of ND:

It looks like a bell.

It is symmetrical about the center.

Ends are almost touching the x-axis (never actually touching though).

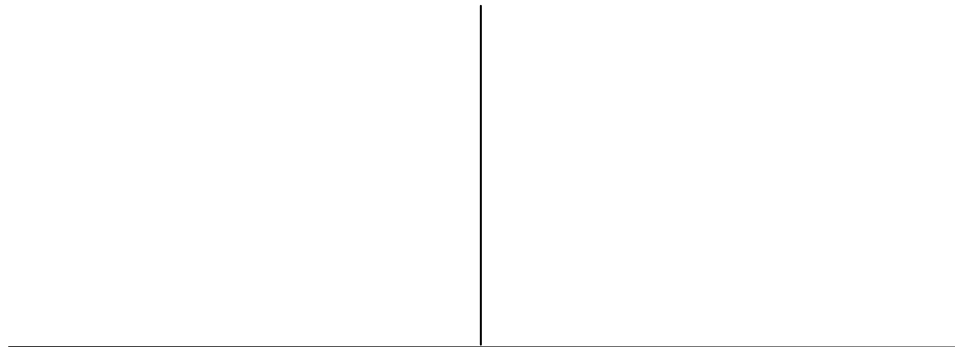
Highest point is at the center.

#### Practice drawing ND:

Draw ND with mean at 5.

Draw a narrower ND at 5.

Draw a flat ND at 7.



### Where does the ND shape come from?

The shape of the normal distribution curve (word curve is generally dropped) comes from the shape of the HISTOGRAM created from a type of data that fits a normal distribution. The shape can be easily seen by drawing the histogram.

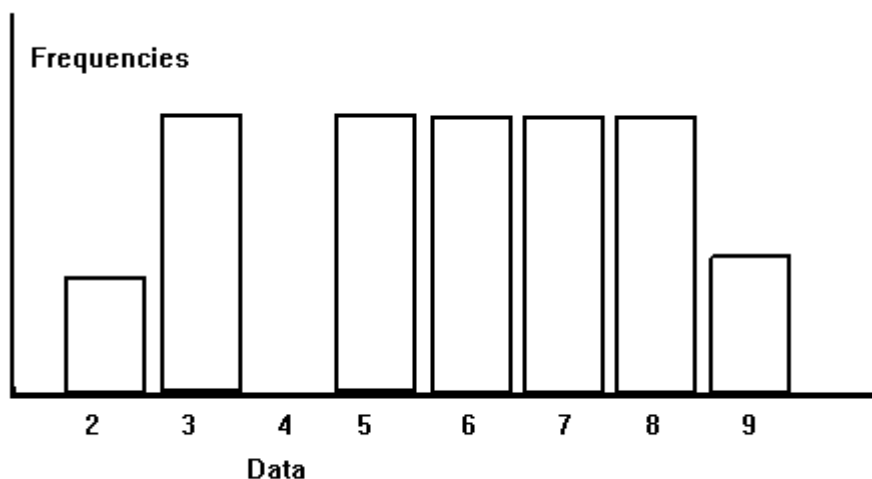
**Histogram** - a bar graph of the frequency (number of times a particular data occurs) of the data set. Consider a data set:

5    7    8    3    2    6    5    8    9    7    and    6

This data set has

	One(1) 2	frequency of occurrence of number 2 is 1
	One(1) 3	frequency of occurrence is 1
	Two(2) 5	frequency of occurrence is 2
	Two(2) 6	etc.
	Two (2) 7	
	Two (2) 8	
and	One (1) 9	

If we were to plot the frequencies in a bar graph, it will look like the following graph.



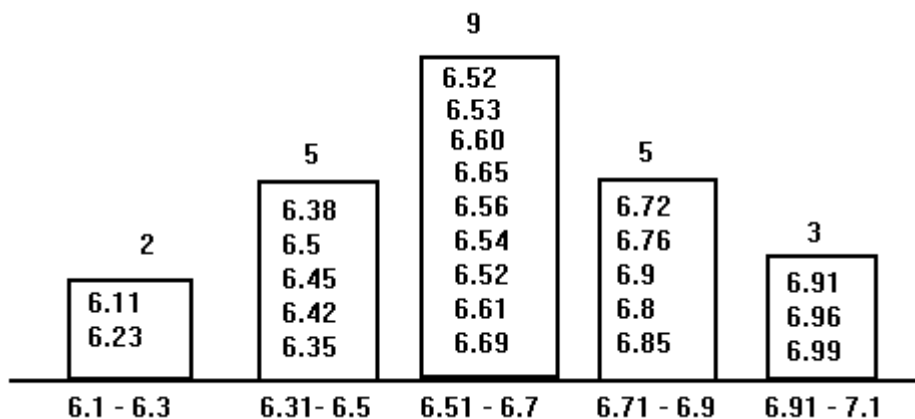
For Plotting

purposes data may be grouped in to fixed ranges.

Data:

6.23 6.35 6.69 6.85 6.99 6.11 6.42 6.61 6.8 6.96 6.45 6.52 6.9 6.91  
6.5 6.54 6.76 6.38 6.56 6.72 6.65 6.60 6.53 6.52

The data is grouped in ranges between 6.1 -6.3, 6.31 - 6.5, 6.51 -6.7, etc.  
There are 2 data that falls in the first group, i.e., between 6.1 - 6.3, etc.

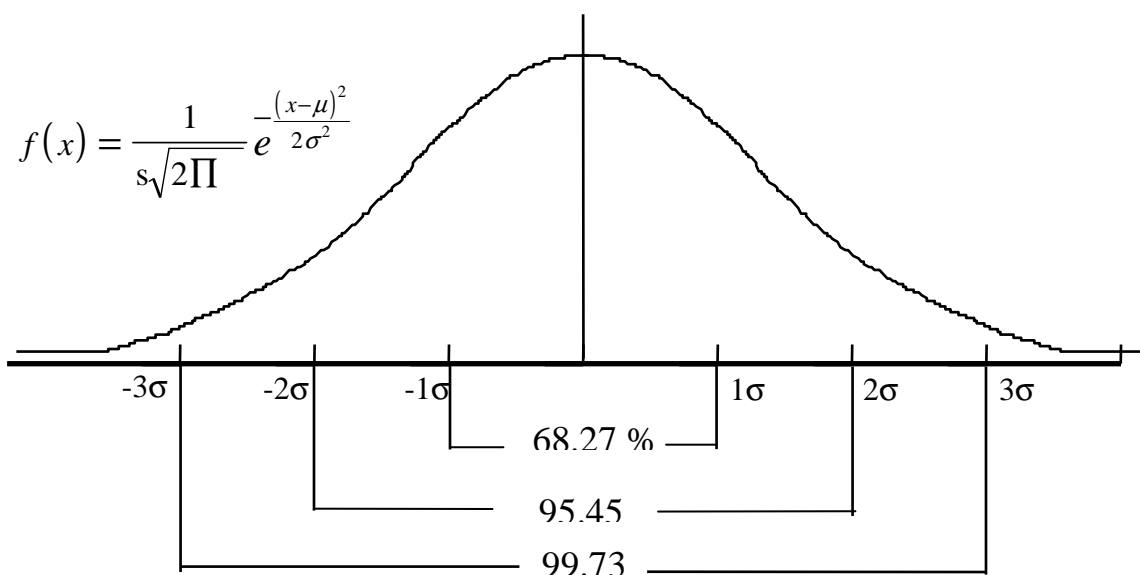


Depending on the nature of the data, the histogram might exhibit many other shapes.

But the most common data are found to exhibit normal distribution.

### More facts about normal distribution:

Research has determined that the ND is centered about the average and the entire area under the curve is covered less than 3 times the standard deviation from the center.



Area under Normal Distribution

### Percentage of Total under Normal Distribution Curve

Within	1 Std. Dev. from Average	68.268949%
	2 “ “	95.449974%
	3 “ “	99.730020%
	4 “ “	99.993666%

Area under the “tails” of the ND:

at 2 St. Deviation:  $100 - 95.449974 = 4.55003\%$

### What if Distribution is not normal?

Not all data will show normal distribution and there is no guarantee the data collected will follow a normal curve.

However, whether the collected data (from product, process or service characteristic) is normal or not, the sample averages  $\bar{x}$  fall in a normal distribution. (Central Limit Theorem).

**Normal Theory:** If the distribution of the measurement, X is normal then the distribution of  $\bar{x}$  is normal.

.

**Central Limit Theorem:** whether x is normal or not, if sample size n (number of subgroups) is large,  $\bar{x}$  is still normal.

Facts:

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

Average of sample averages is equal to the average of all data (population).

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}} \quad S_x, S_{\bar{x}} \text{ are estimates of standard deviations}$$

Standard deviation of normal distribution plot of sample averages is equal to standard deviation of population (all data) divided by square root of sample size.

If the distribution is normal, then 99.73% of x will be within  $\mu \pm 3\sigma$ . This means that if limits are set at 3 standard deviations apart from the mean, only .27% of the observations. In

other words, no more than 3 (2.7) observations out of a thousand are expected to be outside the limits.

Data and data reduced to any other form such as average, range, standard deviation, etc. all can be plotted and distributions obtained. When the distributions are normal, its characteristics can be used to identify normal behavior.

**Example:**

data:            2    5    3    7    4

Range = Difference between the largest and smallest data

$$= 7 - 2 = 5$$

Variance =  $S^2 = \frac{\text{Sum of Squares of distance from average}}{(\text{Number of observations} - 1)}$

$$\text{average} = (2 + 5 + 3 + 7 + 4)/5 = 4.2$$

$$S^2 = [(2 - 4.2)^2 + (5 - 4.2)^2 + (3 - 4.2)^2 + (7 - 4.2)^2 + (4 - 4.2)^2]/(5 - 1)$$

$$= (4.84 + 10.24 + 1.44 + 10.24 + 1.44) / 4$$

$$= 7.05$$

$$\text{Standard Deviation}(s) = \sqrt{7.05} = 2.655$$

**Sample, Subgroup, and Population**

**Sample** - a selected set of data from the total number of possible data. Consider a milling machine producing a total of 60,000 parts. If only 20 sets of 5 parts (5 x 20 = 100) out of the total are inspected, then:

Sample size = 100, and

subgroup size = 20

**Subgroups** are smaller sample

Population is 60,000.

Generally for statistical calculations, the sample size is subdivided into smaller groups. The subgroup size of 5 or 10 data are common. With 5 data in a subgroup, 100 data sample is divided into 20 subgroups.

**Subgroup, Sample and Population Statistics**

Data:

Subgroup 1:	12	14	11	16	12	Avg. = 13
Subgroup 2:	12	14	12	15	16	Avg. = 13.8
Subgroup 3:	11	16	13	15	12	Avg. = 13.4

Sample average  $\bar{\bar{X}} = (13 + 13.8 + 13.4) / 3 = 13.4$   
(Grand average)

Subgroup Ranges

$$R_1 = 16 - 11 = 5$$

$$R_2 = 16 - 13 = 4$$

$$R_3 = 16 - 11 = 5$$

Sample Range  $\bar{R} = (R_1 + R_2 + R_3) / 3 = (5 + 4 + 5) / 3 = 4.666$   
(average range)

Likewise standard deviation for sample can be calculated.

The Standard deviation of sample averages (three averages)

$$S_{\bar{X}} = [ (13 - 13.4)^2 + (13.8 - 13.4)^2 + (13.4 - 13.4)^2 / (3-1) ]^{0.5}$$

$$= (0.16)^{0.5} = 0.40$$

The standard deviation for the entire data (15 numbers) will require the use of all 15 numbers. This process is generally complicated and not done. Instead, the standard deviation is estimated from the value of average range ( $\bar{R}$ ).



# **Module 5**

## Evaluation of Capability for a Stable Process

## Process Capability

Process capability describes the relationship between actual production output and acceptable production output. (Actual production output is compared with the specification limits)

### Process Control

Control charts are used to compare values from a sample production output relates to control limits. (Sample data are compared with the control limits)

## Common Terms

- Upper specification limit (USL) the maximum allowable measurement of a product or service characteristics
- Lower specification limit (LSL) the minimum allowable measurement of a product or service characteristics
- Target value - value of the measured value of the product or service characteristic that the customer ideally wants at all time

### Process Capability Analysis

The distribution of process output to the specification limits is done when:

- the process is in statistical control
- it is a single product or service characteristic
- capability of the process is dependent on the distribution of characteristic measured
- capability of the process is related to the target value and the percentage of production parts outside the specification limits

## Desirable Characteristics of Measure of Capability (Indices)

- Should be simple
- Pure number
- Relate to measured characteristic

There are two common indices used.

C<sub>p</sub> - The process potential index

It represents the ability of the process to satisfy engineering or manufacturing specification.

$$C_p = \frac{|USL - LSL|}{6S_x}$$

$C_p$  does not depend on the location of the center of output and the target.

### $C_{pk}$ - The Process Capability Index

It indicates the actual performance of the process relative to engineering or manufacturing specification.

$$C_{pk} = \frac{|\bar{\bar{X}} - CSL|}{3S_x}$$

depends both on the specification limits and the target.

where the estimated Standard deviation from the population ( $S_x$ ),

$$S_x = \frac{\bar{R}}{d_2}$$

and

CSL = USL or LSL depending on whichever is closer to the grand average

$\bar{\bar{X}} = \mu_x$  (The mean of sample mean is always the same as the population mean)

$C_p = (\text{Difference between the specification limits}) / 6\sigma$

$C_{pk} = (\text{Population Mean} - \text{Nearest specification limit}) / 3\sigma$

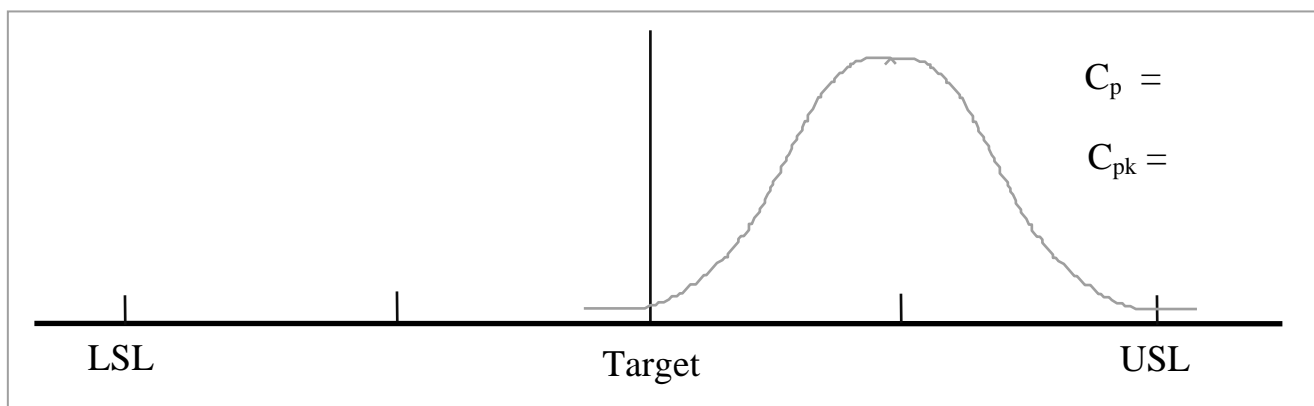
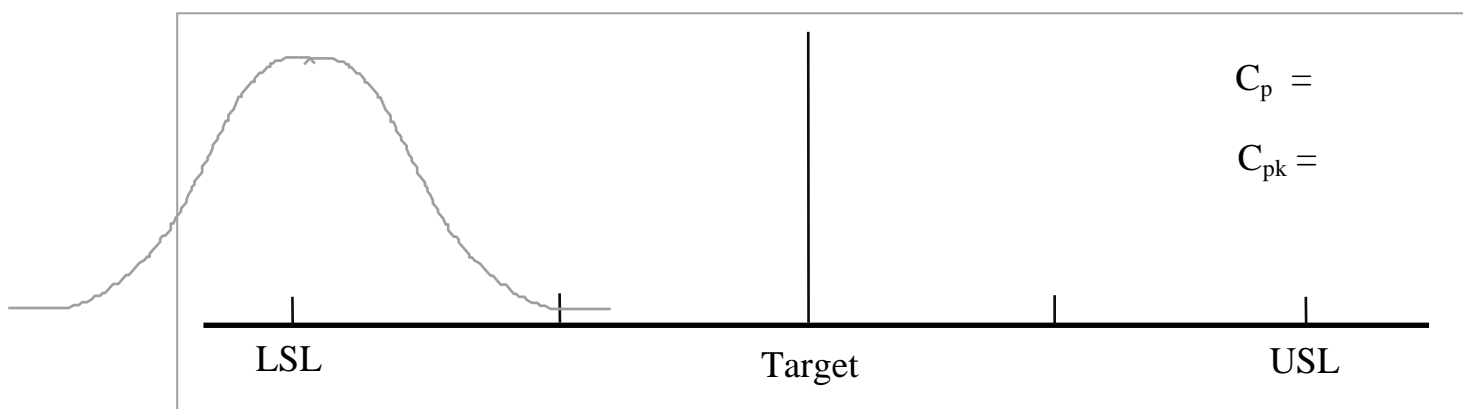
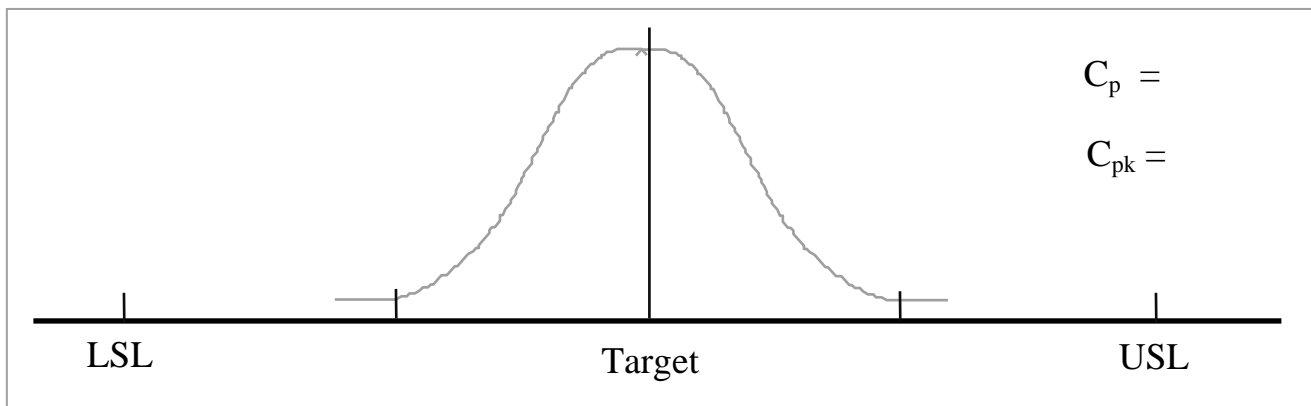
Note:  $C_p$  is what  $C_{pk}$  can be.

$C_{pk} = 1$  represents  $3\sigma$  quality

$C_{pk} = 1.33$  represents  $4\sigma$  quality

$C_{pk} = 1.66$  represents  $5\sigma$  quality

$C_{pk} = 2.00$  represents  $6\sigma$  quality



$C_p$  is the same regardless of the mean performance.  $C_{pk}$  accounts for product center (preferred index). It can be viewed as a distance measurement of the product distribution mean from the nearest specification limit in terms of a “measuring yardstick” which is three standard deviation long.

$C_{pk}$  - Characteristics

- Increases as the mean of the product output fall close to the target
- Decreases as the variation (std. Deviation) of the product performance increases

### Product and Process Capabilities

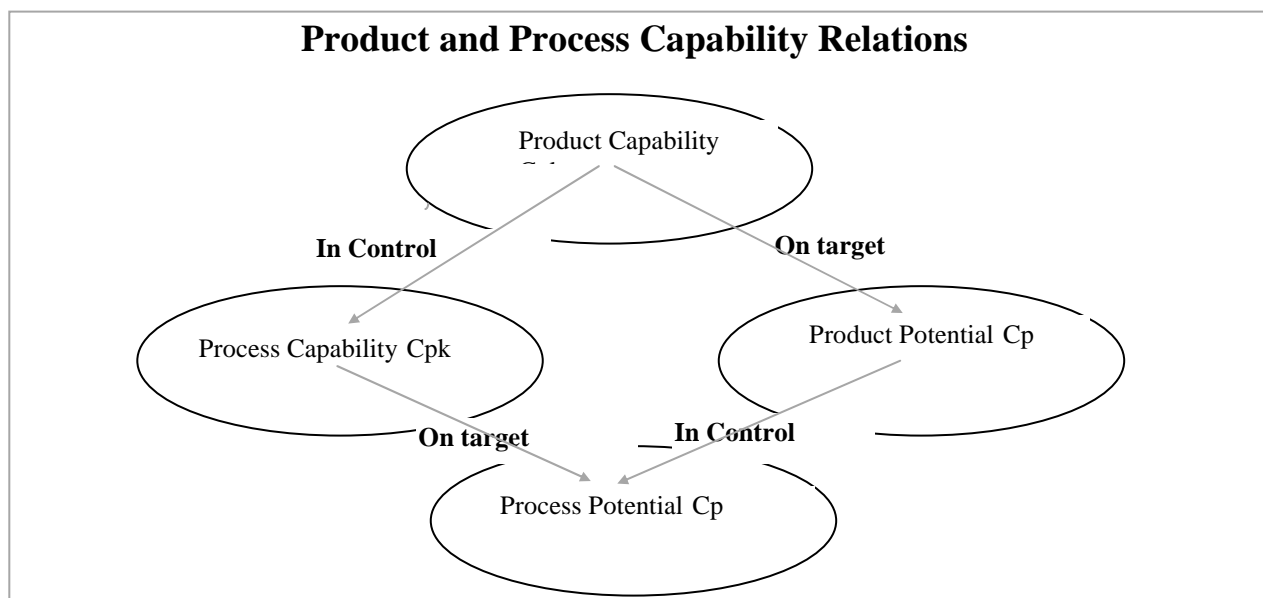
**Product Capability**  $\propto 1/\sigma_T$   
 (when Cp and Cpk are calculated using total variation)

**Process Capability**  $\propto 1/\sigma_{CC}$   
 (when Cp and Cpk are calculated using only common cause variation)

$$\sigma_T^2 = \sigma_{CC}^2 + \sigma_{SC}^2$$

When the process is in control and all variation is due to common cause, then

$$\sigma_{SC} = 0 \text{ and } \sigma_T^2 = \sigma_{CC}^2$$



Review of Capability Indices

Figure 5.1

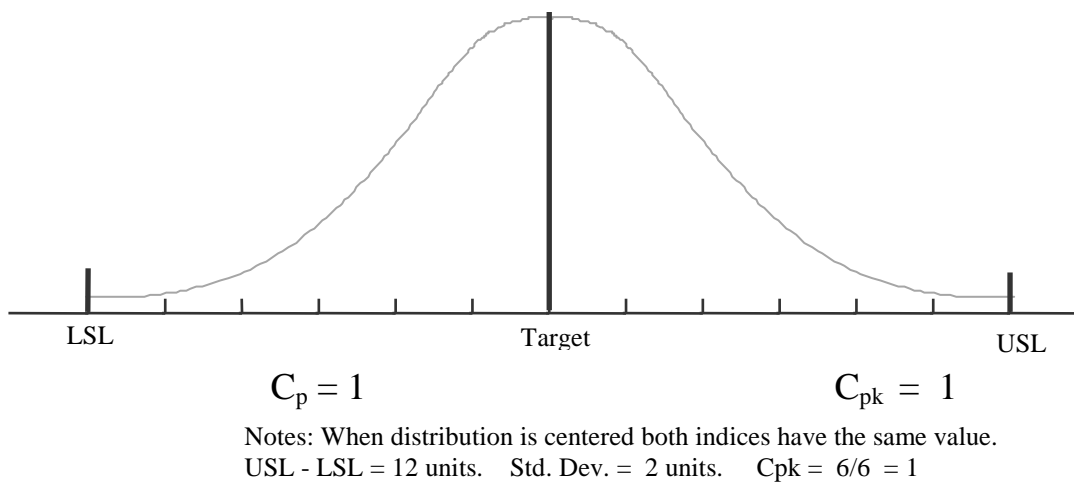


Figure 5.2

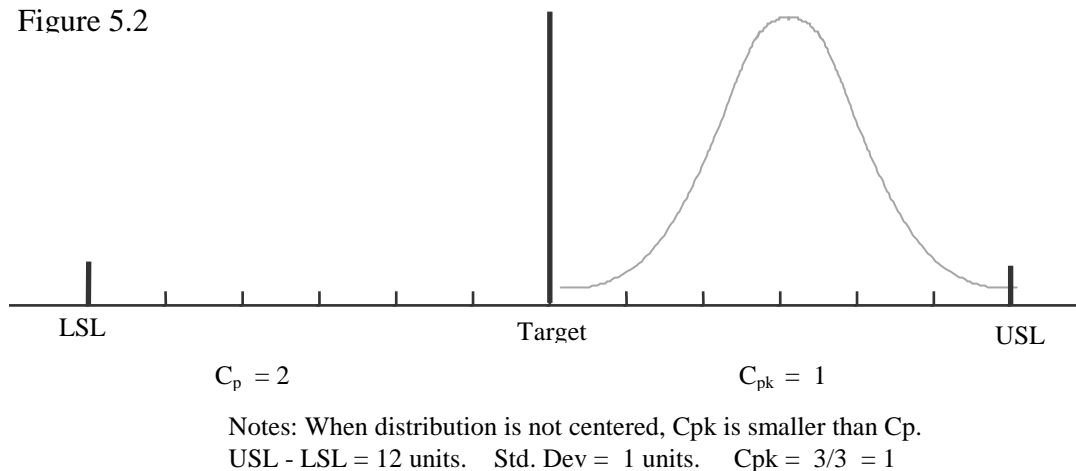


Figure 5.3

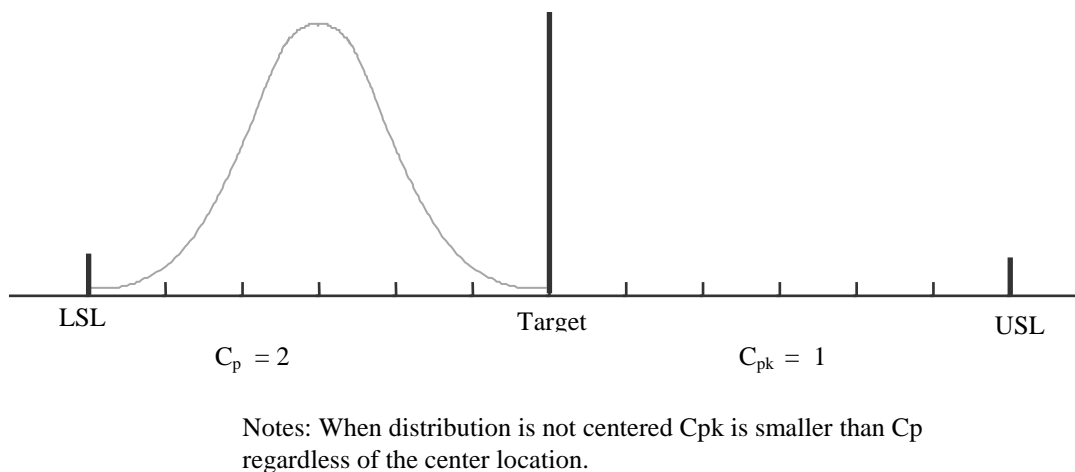
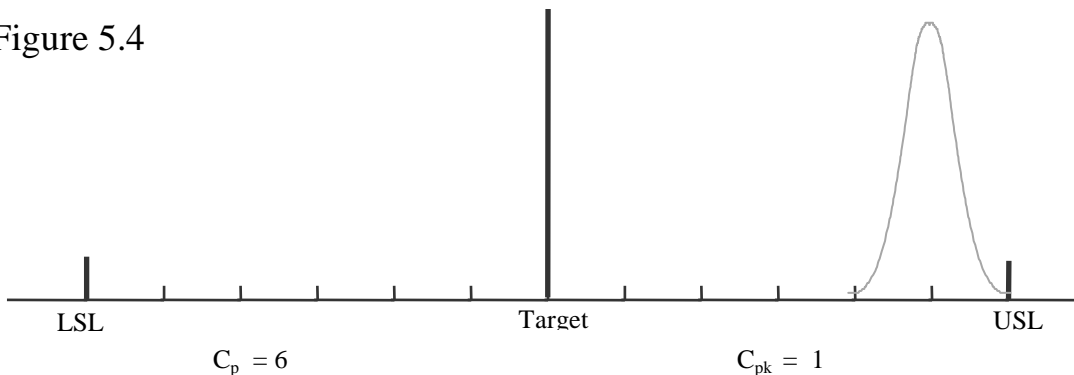
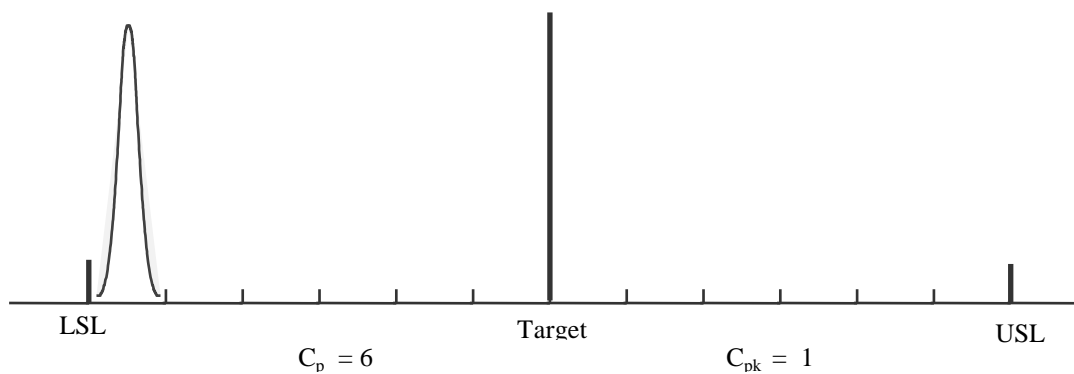


Figure 5.4



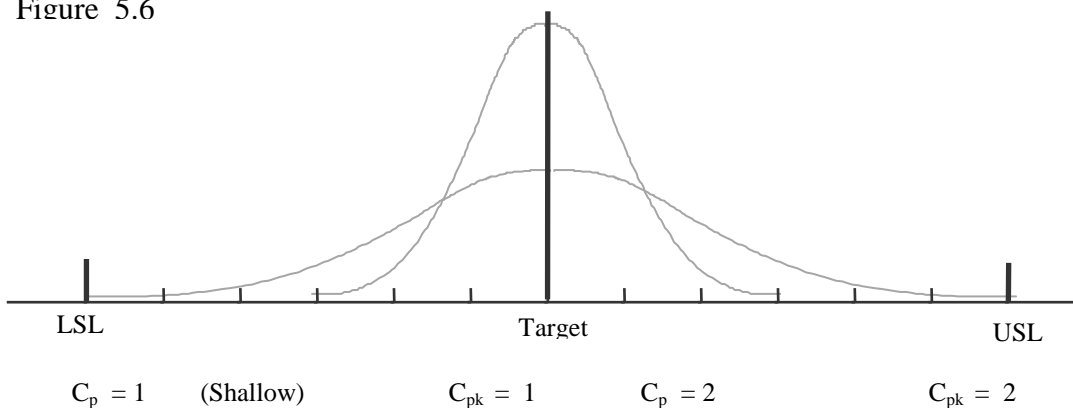
Notes: Extremely off-centered distribution for the same Cpk value.  
 USL - LSL = 12 units. Std. Dev. = 1/3 units. Cpk = 1/1 = 1

Figure 5.5



Notes: Extremely off-centered distribution, close to the LSL.  
 USL - LSL = 12 units. Std. Dev. = 1/6 units. Cpk = 0.50 / .050 = 1

Figure 5.6



Notes: Two centered distributions, one is twice as good as the other.

Figure 5.7

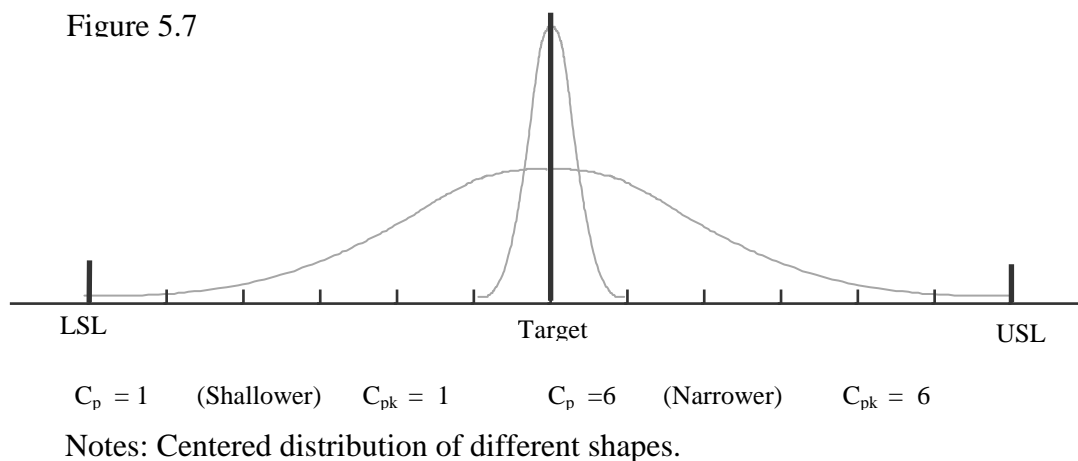


Figure 5.8

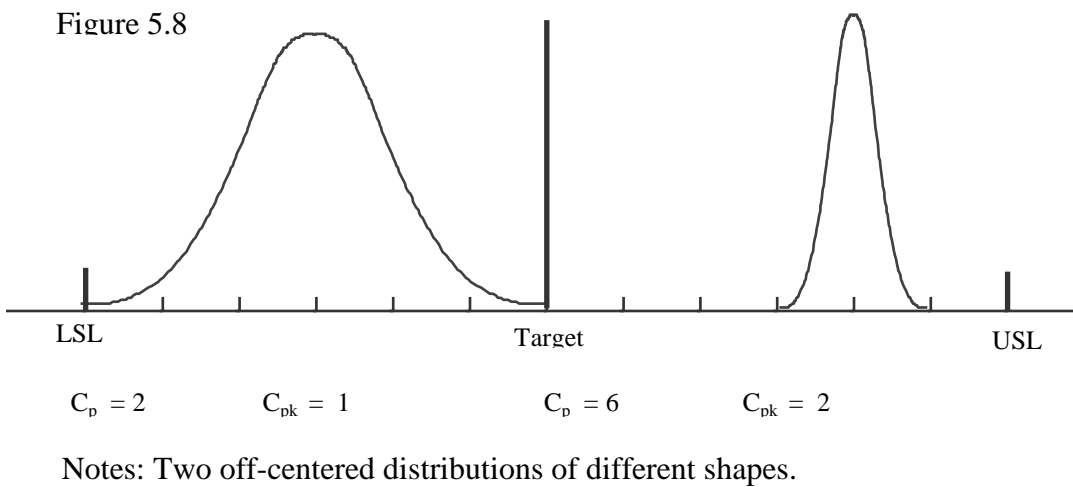


Figure 5.9

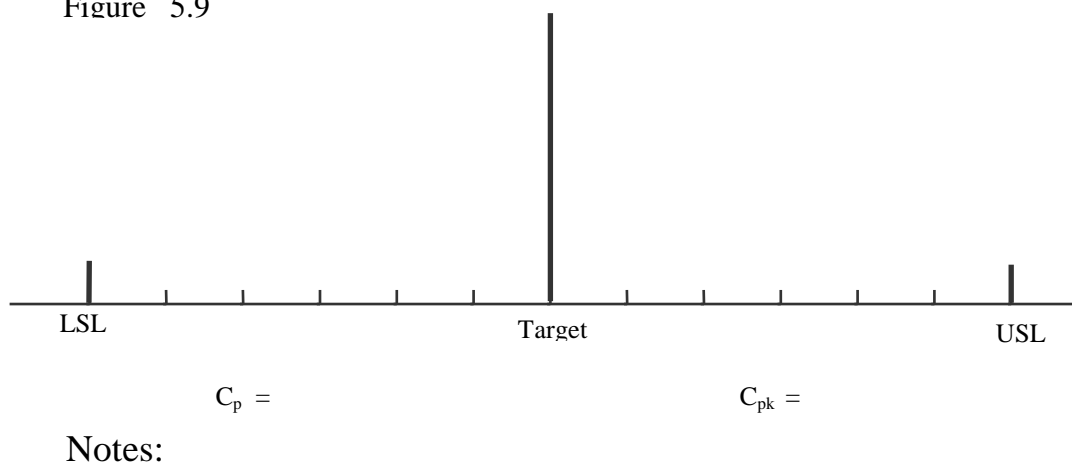
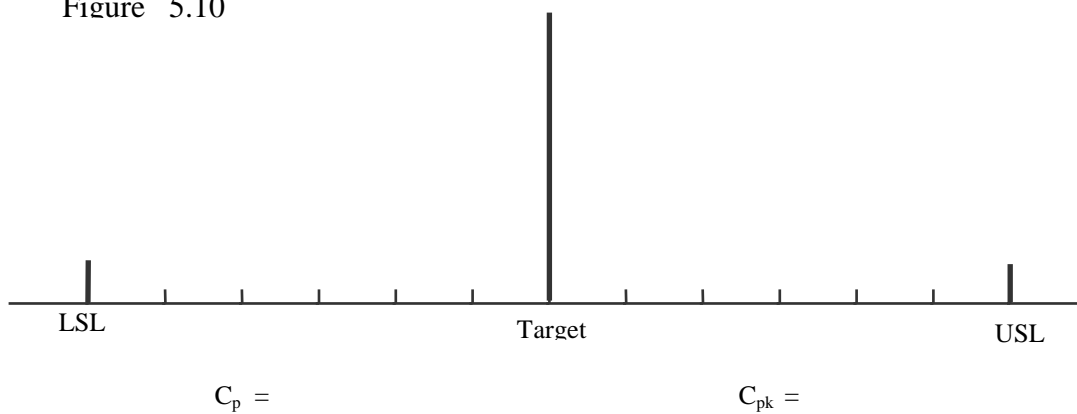


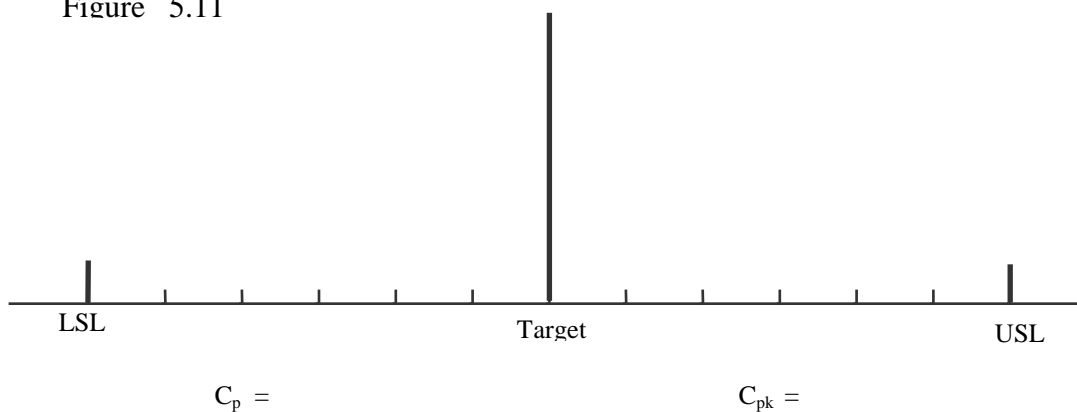


Figure 5.10



Notes:

Figure 5.11



Notes:

Dr. Genechi Taguchi proposed a modified Capability Index ( $C_{pm}$ ) which makes use of the average deviation of data ( $X$ ) from the target ( $X_0$ ) to compute the standard deviation ( $\sigma_T$ ).

$$C_{pm} = \frac{USL - LSL}{6 \sigma_T} \quad \text{where} \quad \sigma_T = \sqrt{\frac{\sum (X - X_0)^2}{N}}$$

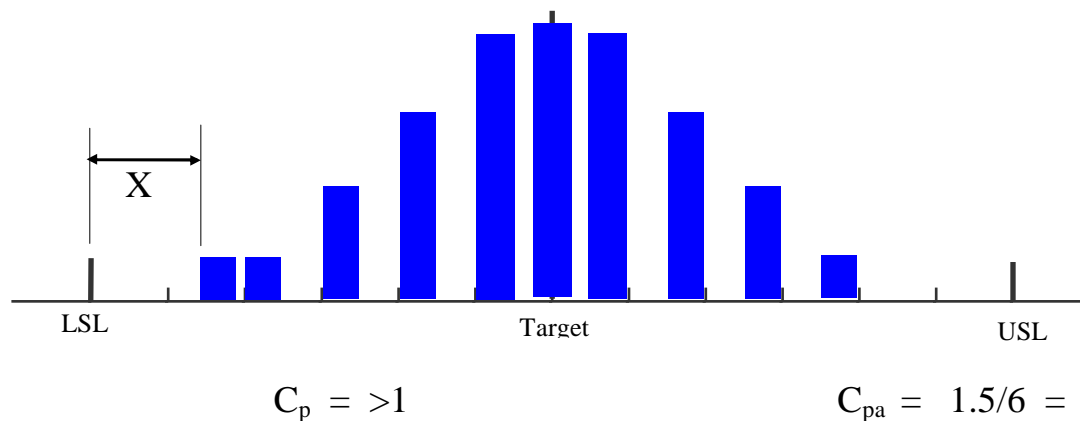
$$\text{also } \sigma_T = \sqrt{\frac{(N-1)}{N} \sigma_c^2 + (X_0 - \bar{X})^2}$$

Notice that  $\sigma_T$ , is same as the Mean Squared Deviation (MSD) defined for the Nominal is the Best quality characteristics by Taguchi.

Yet another index proposed by Dr. Carol Vale, Ford Motor Company, makes use of the data spread to compute the capability index ( $C_{pa}$ ). In this formulation, the distance ( $X_c$ ) of the data group (histogram bar) from the nearest specification limit is divided by the distance of the specification limit from the target.

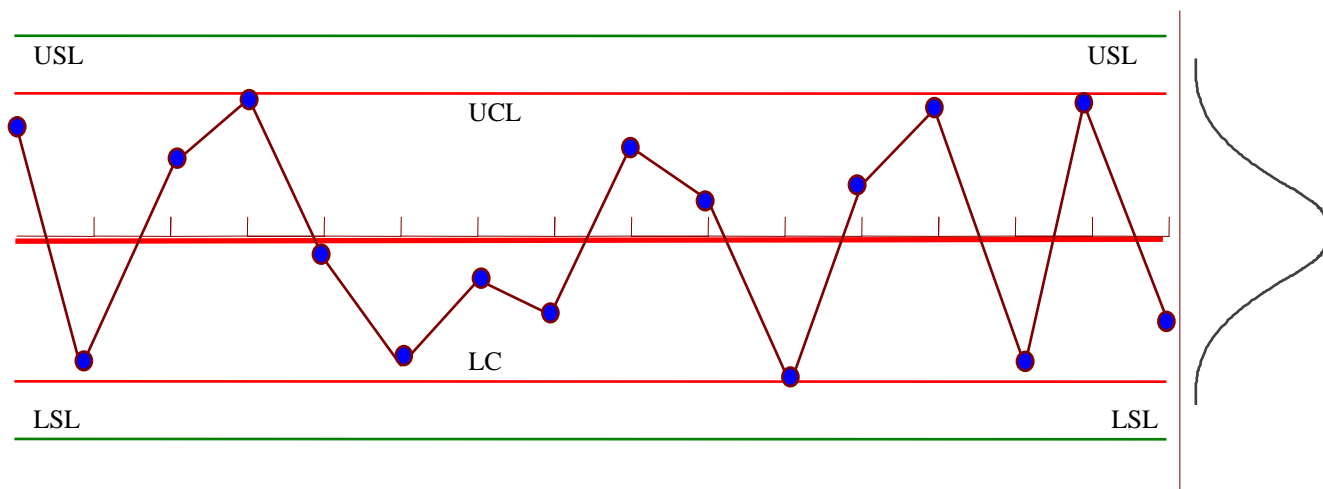
$$C_{pa} = \frac{\text{Target} - LSL}{X_c}$$

Figure 5.12

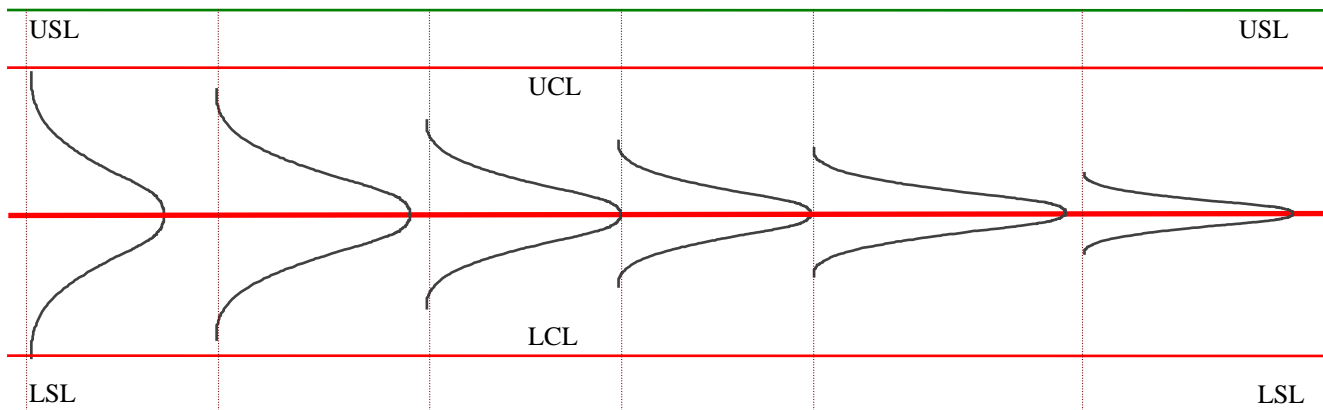


Notes:

Control chart and the resulting distribution.



Distribution after continuous improvement.



# **Module 6**

## Average and Range Charts

## What are Average ( $\bar{X}$ ) and Range (R) Charts?

These charts are plots of Average ( $\bar{X}$ ) and Range(R) over time with two additional lines drawn on the plots. the two additional lines are known as the control limits.

Average and Range are typically done for measurement characteristics of type

- Thickness ( cm, inches, etc)
- Length (feet, mm, inches, etc)
- Diameter (inches, mm, etc)
- Tensile strength (pounds per sq. inch, kg per sq. cm. etc.)
- Time (minutes, seconds, etc)
- Number of defects (in units)

## How to Construct Average ( $\bar{X}$ ) Chart

1. Calculate  $\bar{X}$  and R for all samples
2. Calculate average of the sample averages ( $\bar{\bar{X}}$ ) and the sample Range average ( $\bar{\bar{R}}$ )
3. Calculate control limits

$$U/LCL_{\bar{X}} = \bar{\bar{X}} \pm 3 S_{\bar{X}}$$

If the process is in control, 99.73% of the sample averages will fall within the control limits.

$S_{\bar{X}}$  = “best estimate” of Standard Deviation of the sample averages( $\bar{X}$ )

$S_X$  = “best estimate” of the measured characteristic (X)

**Estimate of the Standard Deviation (of Population, X)**

$\sigma_x$  = The standard deviation of population of the measurements of the process characterizes  
(can be calculated by directly using all the readings)

$S, \bar{S}$  = Sample Standard deviation and average of the sample standard deviation.

$R, \bar{R}$  = Sample range and average of the sample ranges

$\sigma_x$  can be estimated as ( $S_x$ )

$$\text{using } \bar{R} \text{ as } S_x = \frac{\bar{R}}{d_2}$$

$$\text{using } \bar{S} \text{ as } S_x = \frac{\bar{S}}{C_4}$$

where constants  $d_2 = 2.326$  and  $C_4 = .9400$  for subgroup/sample size = 5

**Estimate of the Standard Deviation (of Sample Means)**

$S_{\bar{x}}$  = Standard deviation of sample means ( $\bar{X}$ )

$S_x$  = Standard deviation of sample population (x)

$n$  = number of measurements in each sample

$\bar{\bar{X}}, \bar{X}$  = sample mean and mean of the sample means

The standard deviation of the sample means is always smaller than the population mean and is inversely proportional to the square of the number of samples.

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}$$

Thus, from the definition of population standard deviation,  $S_x$ , above,

$$S_{\bar{x}} = \frac{\bar{R}}{d_2 \sqrt{n}} \quad \text{using } \bar{R}$$

$$S_{\bar{x}} = \frac{\bar{S}}{C_4 \sqrt{n}} \quad \text{using } \bar{S}$$

### Control limits for Sample Means

Limits =

$$UCL_{\bar{X}} = \bar{\bar{X}} \pm 3 S_{\bar{X}}$$

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \bar{R}$$

or

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \bar{R}$$

where

$$A_2 = \frac{3}{d_2 \sqrt{n}} = \frac{3}{2.326 \sqrt{5}} = 0.577$$

(for sample size  $n = 5$ ,  $d_2 = 2.326$ )

### Control Limits for Sample Range (R)

$$UCL_R = D_4 \bar{R} = 2.114 \bar{R}$$

$$LCL_R = 0$$

(for sample size  $n = 5$ , and  $D_4 = 2.114$ )

### Common Plots of X-bar and R Charts

Plot sample  $\bar{X}$  on  $\bar{X}$  control chart

Plot sample R on R control chart

### How to analyze data plots on the range charts

#### i) Points beyond the control limits

The occurrence of one or more points beyond either control limits primary evidence of non-control at that point. If the process is in a state of control and only variation due to common causes is present, any points beyond control limits would be rare.

A point above upper control limit for ranges is generally due to one of the following:

- The control limits calculated erroneously
- The piece-to-piece variability or the spread of the distribution has worsened, either at that point or as part of the trend

- The measurement system has changed.
- The measurement system lacks appropriate discrimination

A point below the lower control limit (for  $n > 6$ ) is generally due to one of the following:

- The control limit calculated is wrong
- The spread of the distribution has decreased
- The measurement system has changed

### **Patterns or Trends Within the Control Limits:**

The presence of unusual pattern or trends even when all ranges are within control limits, can be evidence of non-control. Unfavorable trends are subject of investigation and correction. Favorable trend  $\pm$  should be studied for possible permanent improvement of the process.

#### **ii) Runs – The following are signs of process shift**

- 7 points in a row on one side of the average
- 7 points in a row that are consistently increasing or decreasing

Possible causes:

- change in measurement system
- greater range indicates trouble and should be investigated
- reduced range should be studied for permanent process change.

#### **iii) Obvious nonrandom pattern -- presence of distinctive pattern offers clues to special causes.**

Distance from  $\bar{R}$  - for a process in a state of control, 2/3 of the data should be within 1/3 of the control limits and 1/3 of the data should be within the outer two thirds of the region.

If substantially more than 2/3 of the points lie close to the process average (for 25 subgroups if 90% are in the middle third of the control limits), look for the following causes.

- Control limits miscalculated
- Data have been edited
- Subgroups have been stratified and have different averages.

If significantly fewer than 2/3 of the data points lie close to the process average (40% or fewer), suspect one of the following.

- Control limits have been miscalculated.
- The sampling method used caused subgroups to contain measurements from two or more very different process streams. If several process streams are present they should be identified and tracked separately.



<b>Example 6.1</b>		<b>Data Type: Seat Deformation</b>						
Sample	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Average	Range	Std.Dev.
1	124	122	129	118	122	123.00	11	4.00
2	126	131	126	122	125	126.00	9	3.24
3	119	127	124	128	132	126.00	13	4.85
4	122	120	121	117	125	121.00	8	2.92
5	119	127	122	132	124	124.80	13	4.97
6	118	128	129	126	130	126.20	12	4.82
7	117	130	119	136	114	123.20	22	9.36
8	113	130	122	123	116	120.80	17	6.61
9	116	118	112	128	125	119.80	16	6.57
10	126	119	123	129	129	125.20	10	4.27
11	123	133	129	115	127	125.40	18	6.84
12	128	106	122	121	120	119.40	22	8.11
13	141	130	129	122	129	130.20	19	6.83
14	120	121	122	135	116	122.80	19	7.19
15	127	121	116	118	115	119.40	12	4.83
16	122	136	126	124	139	129.40	17	7.60
17	119	122	123	120	124	121.60	5	2.07
18	117	125	137	126	133	127.60	20	7.73
19	128	122	129	113	129	124.20	16	6.91
20	129	124	117	119	121	122.00	12	4.69

Grand Avg. = 123.900    Avg. Range = 14.550    Avg. Std.Dev. = 5.721

**Control limit for X-bar chart**

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \bar{R} \quad (\text{Upper limit}) = 123.9 + .577 \times 14.55 = 132.29$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \bar{R} \quad (\text{Lower limit}) = 123.9 - .577 \times 14.55 = 115.50$$

$$(\text{Constant } A_2 = 0.577 \text{ for sample size } = 5) \quad \text{One Std.Dev} = .577 \times 14.55 / 3 = 2.80$$

Calculation of  $C_P$  &  $C_{PK}$  (assuming USL = 150, LSL = 110, Target = 125)

$$S_X = \bar{R} / d_2 \quad (\text{estimate of } \sigma_X) = 14.55 / 2.326 = 6.25 \quad (d_2 = 2.326)$$

$$\bar{\bar{X}} = \mu_X \quad (\text{estimate of Avg.}) = 123.90$$

$$C_P = \frac{USL - LSL}{6S_X} = (150 - 110) / (6 \times 6.25) = 1.066$$

$$C_{PK} = \frac{|\bar{X} - CSL|}{3S_X} = (123.9 - 110) / (3 \times 6.25) = 0.741$$

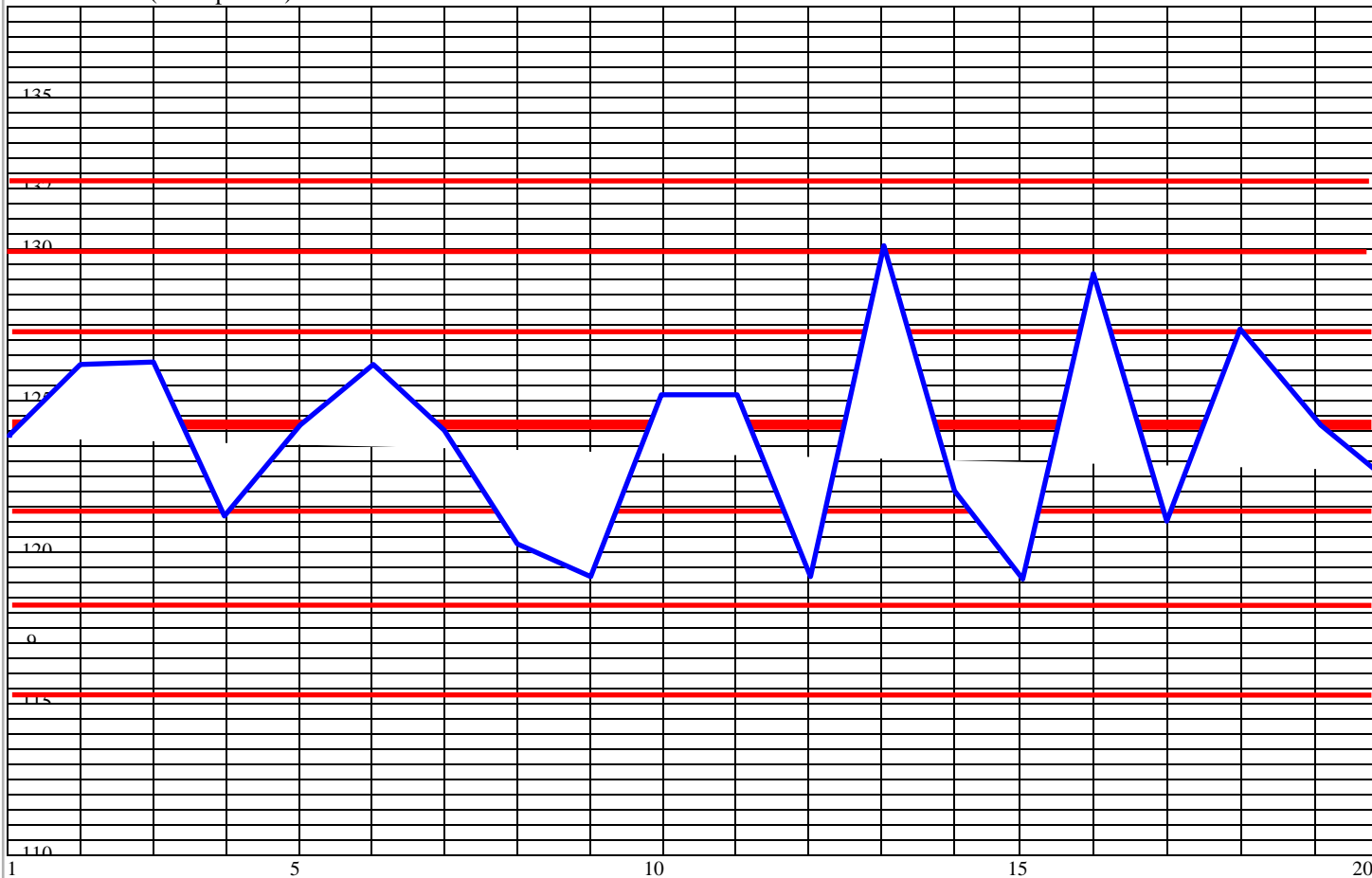
Plant Machine#      Dept. Date:      Operation Measure      =      Eng. Specs. Sample size      Part# Part Name

**X-bar Chart** (Example 6.1)

LCL = 115.50

$\bar{X}$  = 123.9

UCL = 132.29



Measured Data (Units of measurement: )

Date	Time	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
Rea di n g s	1																					
	2																					
	3																					
	4																					
	5																					
	$\bar{X}$																					
	S																					

LCL =

$\bar{X}$  =

UCL =

Example 6.2		Data Type: Pressure Plate Height						
Sample	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Average	Range	Std.Dev.
1	5.74	5.80	5.72	5.78	5.63	5.73	0.17	0.07
2	6.17	6.29	5.69	6.14	6.14	6.09	0.60	0.23
3	7.09	6.63	5.65	6.39	6.35	6.42	1.44	0.52
4	5.80	6.15	5.96	5.97	5.48	5.87	0.67	0.25
5	6.85	5.55	6.16	6.25	5.43	6.05	1.42	0.58
6	6.19	6.29	6.37	5.84	5.36	6.01	1.01	0.42
7	5.97	6.13	5.25	5.80	5.45	5.72	0.88	0.36
8	5.51	5.71	7.09	6.19	5.70	6.04	1.58	0.64
9	5.81	5.92	5.26	5.89	5.85	5.75	0.66	0.27
10	6.72	5.70	5.54	5.99	5.76	5.94	1.18	0.46
11	6.02	5.87	5.67	5.27	6.17	5.80	0.90	0.35
12	5.80	6.08	6.51	5.83	6.17	6.08	0.71	0.29
13	5.84	5.25	5.43	5.51	5.50	5.51	0.59	0.21
14	6.33	5.68	5.49	5.14	6.27	5.78	1.19	0.51
15	6.46	6.02	6.37	7.08	6.04	6.39	1.06	0.43
16	5.27	6.14	5.74	5.71	6.25	5.82	0.98	0.39
17	5.78	5.84	5.61	6.50	5.50	5.85	1.00	0.39
18	5.78	6.20	6.09	6.27	5.46	5.96	0.81	0.34
19	5.73	4.80	6.37	6.61	5.75	5.85	1.81	0.70
20	5.61	6.74	5.57	5.81	5.80	5.91	1.17	0.48
Grand Avg. =		5.928	Avg. Range =		0.991	Avg. Std. Dev. =		0.395

#### Control limit for X-bar chart

$$UCL_{\bar{X}} = \bar{\bar{X}} + A_2 \bar{R} \quad (\text{Upper limit}) = 5.928 + .577 \times .991 = > 6.5$$

$$LCL_{\bar{X}} = \bar{\bar{X}} - A_2 \bar{R} \quad (\text{Lower limit}) = 5.928 - .577 \times .991 = > 5.3$$

(Constant  $A_2 = 0.577$  for sample size = 5)    One Std.Dev =  $.577 \times 0.991 / 3 = > 0.20$

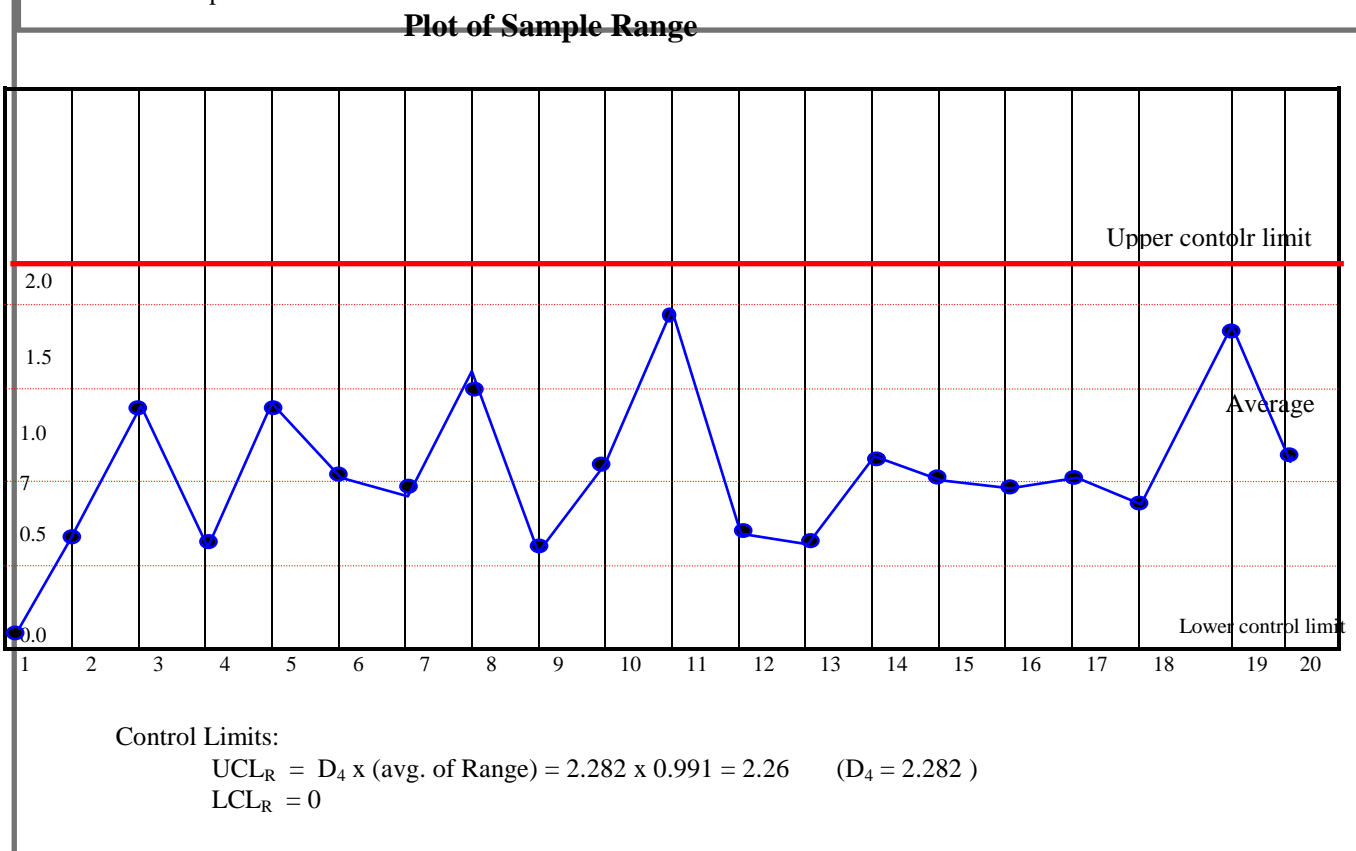
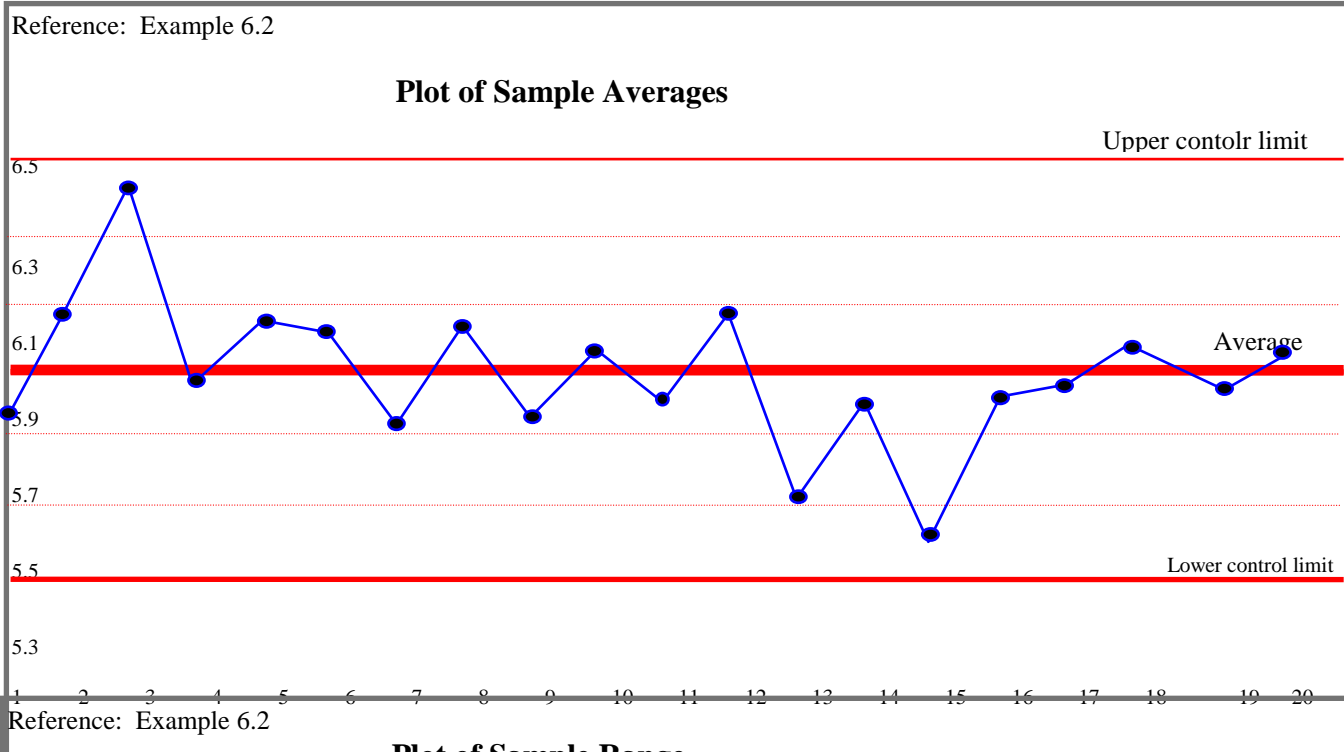
Calculation of  $C_P$  &  $C_{PK}$  (assuming USL = 7.5, LSL = 3.5, Target = 5.5)

$$S_X = \bar{R} / d_2 \quad (\text{estimate of } \sigma_X) = .991 / 2.326 = 0.426 \quad (d_2 = 2.326)$$

$$\bar{\bar{X}} = \mu_X \quad (\text{estimate of Avg.}) = 5.928$$

$$C_P = \frac{USL - LSL}{6S_X} = (7.5 - 3.5) / (6 \times 0.426) = 1.565$$

$$C_{PK} = \frac{|\bar{\bar{X}} - CSL|}{3S_X} = (7.5 - 5.928) / (3 \times 0.426) = 1.23$$



<b>Exercise 6.3</b>		<b>Data Type: Door Closing Effort</b>						
Sample	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Average	Range	Std.Dev.
1	38.59	44.24	47.19	47.44	48.16	45.12	9.57	3.95
2	42.11	44.38	48.70	48.24	43.56			
3	44.31	43.34	42.44	47.91	45.61	44.72	5.47	2.14
4	45.54	46.11	40.31	47.34	42.58			
5	39.41	42.84	50.68	43.01	46.44	44.48	11.27	4.27
6	43.55	43.63	43.01	42.19	42.15	42.91	1.48	0.71
7	43.74	45.97	42.33	46.98	46.25			
8	46.47	40.17	38.59	41.32	42.40	41.79	7.88	2.97
9	44.06	46.18	43.02	48.34	49.13	46.15	6.11	2.64
10	45.36	43.91	36.00	42.47	41.47	41.84	9.36	3.58
11	46.40	52.34	42.80	42.33	48.85	46.54	10.01	4.20
12	41.07	45.97	46.51	46.58	47.08	45.44	6.01	2.48
13	48.56	40.46	43.99	48.16	42.73	44.78	8.10	3.51
14	46.87	44.02	48.31	43.88	48.34	46.28	4.46	2.21
15	42.26	47.08	49.13	42.26	52.34			
16	51.37	48.74	45.32	46.47	44.20	47.22	7.17	2.86
17	41.36	46.36	45.93	48.60	43.59			
18	39.34	44.92	39.34	43.70	44.13	42.29	5.58	2.72
19	43.16	51.40	35.89	46.47	45.93	44.57	15.51	5.69
20	45.46	36.86	46.87	42.84	52.34			
Grand Avg. =		Avg. Range =			Avg. Std. Dev. =			

Reference:

### Plot of Sample Averages

Upper control limit

Average

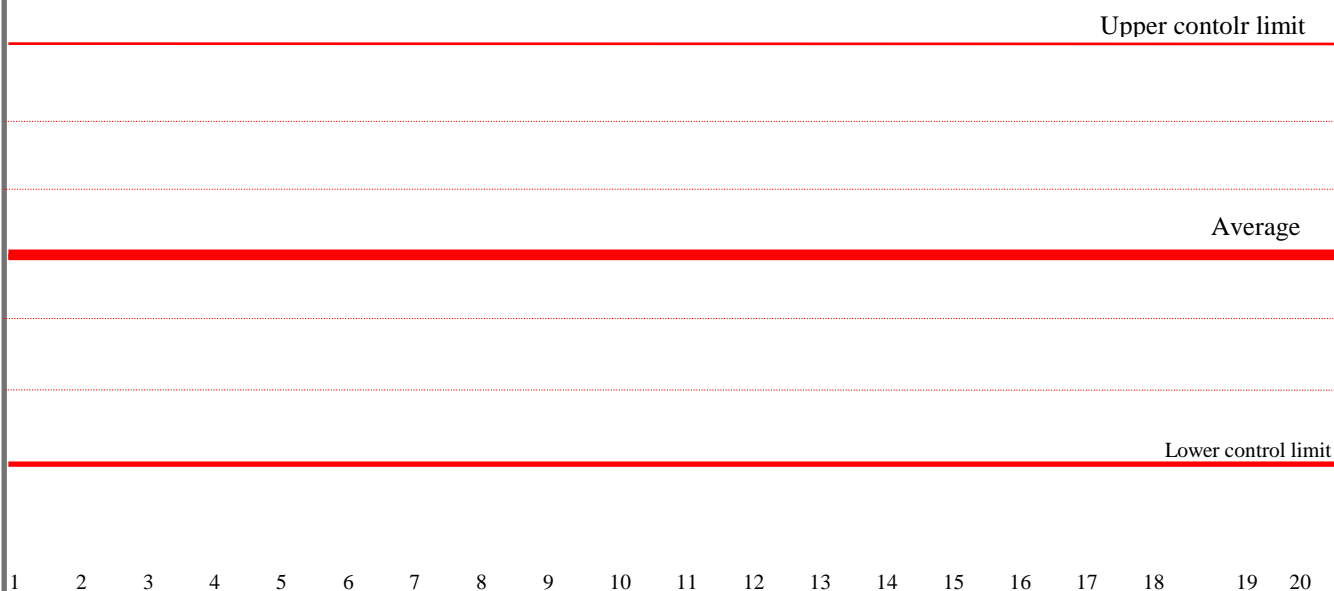
Lower control limit

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

**Exercise 6.4****Data Type: Pressure Plate Height**

Sample	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Average	Range	Std.Dev.
1	5.74	5.80	5.72	5.78	5.63	5.73	0.17	0.07
2	6.17	6.29	5.69	6.14	6.14	6.09	0.60	0.23
3	7.09	6.63	5.65	6.39	6.35			
4	5.80	6.15	5.96	5.97	5.48	5.87	0.67	0.25
5	6.85	5.55	6.16	6.25	5.43	6.05	1.42	0.58
6	6.19	6.29	6.37	5.84	9.60			
7	5.97	6.13	5.25	5.80	5.45	5.72	0.88	0.36
8	5.51	5.71	7.09	6.19	5.70	6.04	1.58	0.64
9	5.81	5.92	5.26	5.89	5.85	5.75	0.66	0.27
10	6.72	5.70	5.54	5.99	6.50			
11	6.02	5.87	5.67	5.27	6.17	5.80	0.90	0.35
12	5.80	6.08	6.51	5.83	6.17	6.08	0.71	0.29
13	5.84	5.25	5.43	5.51	5.50	5.51	0.59	0.21
14	6.33	5.68	5.49	5.14	6.27	5.78	1.19	0.51
15	6.46	6.02	6.37	7.08	6.04			
16	5.27	6.14	5.74	5.71	6.25	5.82	0.98	0.39
17	5.78	5.84	5.61	6.50	5.50	5.85	1.00	0.39
18	5.78	6.20	6.09	6.27	5.46	5.96	0.81	0.34
19	5.73	4.80	6.37	6.61	5.75			
20	5.61	6.74	5.57	5.81	5.80	5.91	1.17	0.48
Grand Avg. =	Avg. Range =				Avg. Std. Dev. =			

Reference:

**Plot of Sample Averages**

<b>Exercise 6.5</b>		<b>Data Type: Stopping Distance</b>						
Sample	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	Average	Range	Std.Dev.
1	151	149	130	160	134	144.80	30	12.48
2	167	131	146	126	115	137.00	52	20.14
3	150	139	149	146	145	145.80	11	4.32
4	153	112	147	123	126	132.20	41	17.20
5	175	151	167	125	112			
6	128	123	123	119	185			
7	139	149	116	123	139	133.20	33	13.39
8	132	126	148	125	163			
9	131	146	146	141	128	138.40	18	8.44
10	139	141	139	112	148	135.80	36	13.81
11	112	120	126	130	140			
12	149	112	151	124	112	129.60	39	19.27
13	146	105	119	141	125	127.20	41	16.65
14	148	123	123	134	139	133.40	25	10.74
15	141	117	156	130	128	134.40	39	14.78
16	106	151	133	128	154	134.40	48	19.42
17	124	138	109	147	132	130.00	38	14.44
18	137	129	120	143	123	130.40	23	9.58
19	123	126	129	143	118	127.80	25	9.42
20	130	147	176	145	127			
Grand Avg. =		Avg. Range =			Avg. Std. Dev. =			

Reference:

**Plot of Sample Averages**

Upper control limit

Average

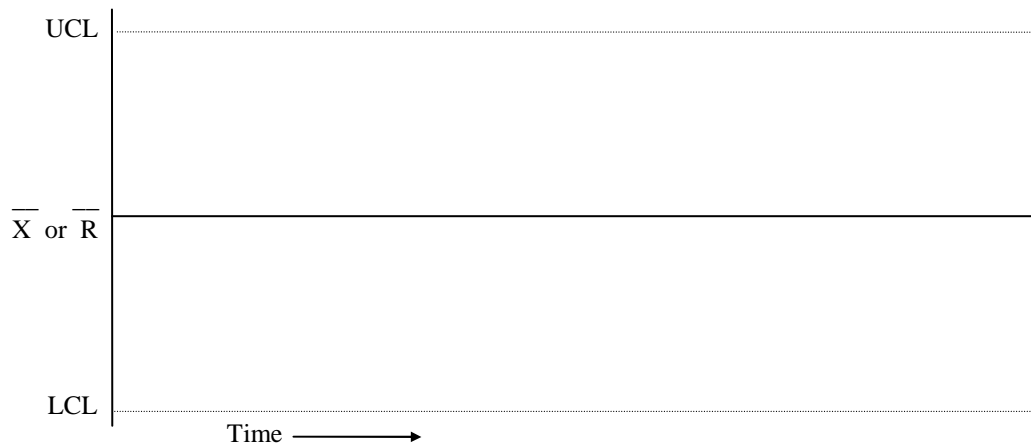
Lower control limit

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20



# **Module 7**

## Interpretation of Control Charts



### How to read the control chart?

The purpose of examining a control chart is to detect if there is an out of control situation during the time the data was collected.

If a process is in control, then, there will be no change in mean or range present over time. This means that the variations in the values of individual sample means and the ranges are due only to common causes (random behavior). If there is variation due to special causes (non-random behavior), the process will be considered out of control.

### How do we detect that there is variation due to special causes?

When the process output is affected by special causes of variation, it often produces distinct control chart (pattern). Detecting one or more such distinct behavior derives conclusions about a particular data plot.

Mathematics behind detection of unlikely (non-random) behavior.

Considering the normal distribution, the area under the  $\pm 3$  Standard Deviation is 99.73%. This means that the probability of a data point falling outside the upper or lower control limit is  $(100 - 99.73) = .27\%$ . or 2.7 parts per thousand (1.3 per 1000 on one side). This is a good number to remember for comparing with other probabilities.

Probability that a data point falls on one side of the mean is 0.5

Probability that two points in a row will fall on the same side of the mean is  $(.05)(.05) = 0.25$

Probability that 9 points in a row will fall on one side of the distribution is  $(0.5)^9 = .00195 = 2/1000$

Probability that head will show 10 times in a row is  $(0.5)^{10} = 0.097 = 0.1\% = 1/1000$

Probability that you will get a queen is  $4/52$

Probability that you will get two queens in a row is  $4/52 \times 4/52 = (16/2704) = 0.0059 = 6/1000$

Probability that a data point will be outside  $2\sigma$  limit is  $(100 - 95.45)/2 = 2.275\% = .02275$

Probability that two points in a row will be beyond the  $2\sigma$  limit is  $(.02275 \times .02275) = .0005 = .05\%$

Likewise, probability that 4 points in a row will be outside the  $1\sigma$  limit is  $(.159)^4 = .000064 = .06\% = .6/1000$

since, probability of one point outside is  $(100 - 68.27)/2 = 15.86\% = .159$

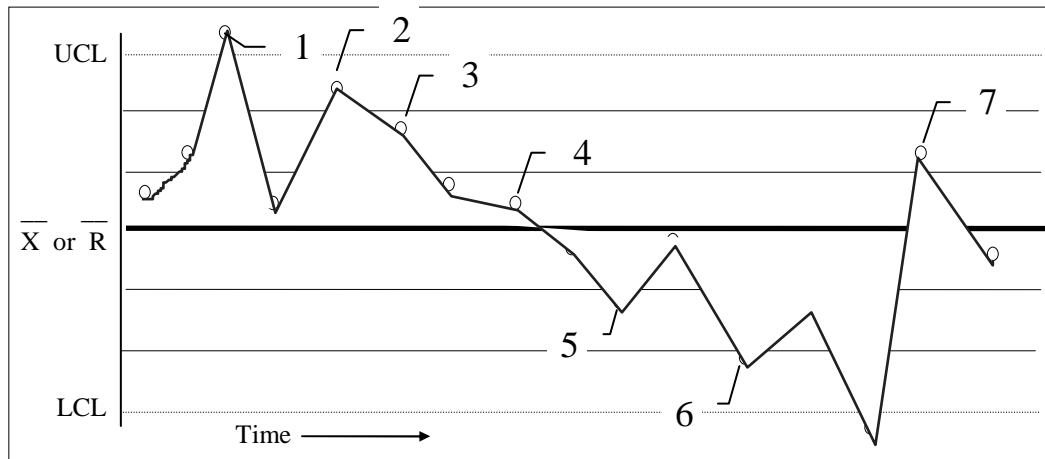
### Who makes the rules for out of control detection?

It varies from author to author. You can create your own rule based on the probability of occurrence.

General approach: If some situation has less than 1 in a thousand chance (1.3 per 1000 beyond either control limits), it is a signal for unusual (non-random) circumstance.

General guidelines for detection of probable out-of-control situation. Each of the following situations will signal the presence of special cause of variations.

Conditions	Probability of occurrence
1. A point is beyond the upper or lower control limits	0.135 %
2. Two out of three consecutive points are beyond the two standard deviation away from the centerline.	0.153 %
3. Four out of five consecutive points are beyond one standard deviation away from the centerline.	0.277 %
4. Six consecutive points are in an upward trend.	0.139 %
5. Eight consecutive points are above the centerline.	0.391 %
6. Eight out of ten consecutive points, taken from left to right, shows upward/downward trend.	.069%
7. Second of two consecutive point is at least four standard deviation above/below the first one.	0.234



The number next to a point indicates the condition of out-of-control

### How to analyze data plots on the range charts

#### i) Points beyond the control limits

The occurrence of one or more points beyond either control limits primary evidence of non-control at that point. If the process is in a state of control and only variation due to common causes is present, any points beyond control limits would be rare.

A point above upper control limit for ranges is generally due to one of the following:

- The control limits calculated erroneously
- The piece-to-piece variability or the spread of the distribution has worsened, either at that point or as part of the trend.
- The measurement system has changed.
- The measurement system lacks appropriate discrimination.

A point below the lower control limit (for  $n > 6$ ) is generally due to one of the following:

- The control limit calculated is wrong.
- The spread of the distribution has decreased.
- The measurement system has changed.

Patterns or Trends Within the Control Limits:

The presence of unusual pattern or trends even when all ranges are within control limits can be evidence of non-control. An unfavorable trend should be corrected. A favorable trend  $\pm$  should be studied for possible permanent improvement of the process.

#### ii) Runs - following are signs of process shift:

- 7 points in a row on one side of the average
- 7 points in a row that are consistently increasing or decreasing

Possible causes:

-- change in measurement system

- greater range indicates trouble and should be investigated
- reduced range should be studied for permanent process change.

**iii) Obvious nonrandom pattern -- presence of distinctive pattern offers clues to special causes.**

Distance from  $\bar{R}$  - for a process in a state of control, 2/3 of the data should be within 1/3 of the control limits and 1/3 of the data should be within the outer two thirds of the region.

If substantially more than 2/3 of the points lie close to the process average (for 25 subgroups if 90% are in the middle third of the control limits), look for the following causes.

- Control limits miscalculated
- Data have been edited
- Subgroups have been stratified and have different averages.

If significantly fewer than 2/3 of the data points lie close to the process average (40% or fewer), suspect one of the following.

- Control limits have been miscalculated.
- The sampling method used caused subgroups to contain measurements from two or more very different process streams. If several process streams are present they should be identified and tracked separately.

(Course materials to be added later)

Module 8 – Module 16 Excluded from this document

**Module 17**

**Review Topics, Exercise, and Q&A**

Key Topics of Discussions (**Things you should understand from this overview sessions**)

- (1) **Distribution Functions**  
Normal, Binomial, Exponential, logarithmic
  - Preparation
  - Areas of application
  - Characteristics & special properties
- (2) **Capability Statistics**  
How to compare two populations

## Capability indices (calculations and implications)

**(3) Statistical Process Control (SPC)**

Purpose: Identify common and special causes

Technique: Create standard and compare performance with it.

- Use key properties of ND (Avg. & Std. Dev.)
- Calculate Std. Dev. Using arithmetic
- Create standard AVG. chart for stable performance
- Collect data and plot them
- Use std. Guidelines for causes of variation

***Review Problems and Exercises***

1. What is the common way to measure the central tendency of data?  
 Mean     Median     Mode
2. What is the most common characteristic used to represent dispersion of data?  
 Range     Standard Deviation     Variance
3. Calculate the average (a) and St. Dev. (s) of the following data and plot the population distribution.

7 5 5 4 6 and 3

(Ans: a = 5    s = 1.414 )

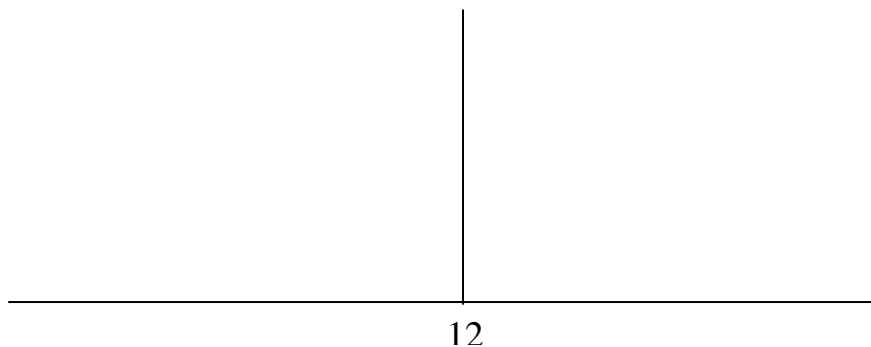
4. Calculate the average (a) and St. Dev. (s) of the following data and plot the population distribution.

9.5 8.3 8.6 9.2 9.6 10.0 8.8 and 9.3

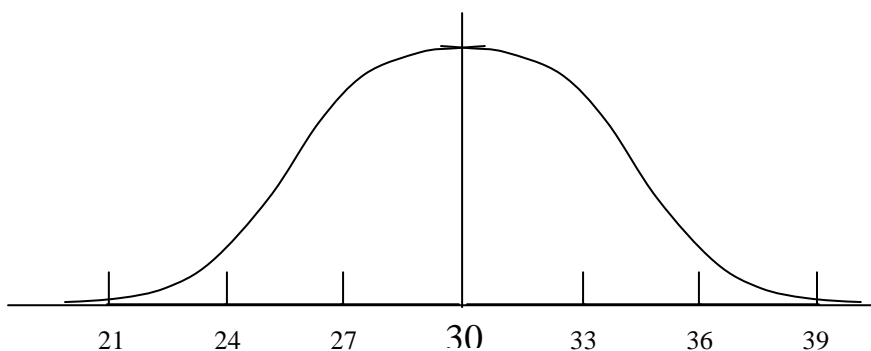
(Ans. a = 9.162    s = 0.562)

5. Plot the normal distribution for a distribution with a = 12 and s = 2. Show the distribution when the standard deviation, s = 1. How high would the peak of the second distribution be compared to the first one? Calculate maximum height of distributions (use expression for normal distribution) and label x and y-axes before plotting graph.

(Ans. Approximately twice as much)



6. Determine approximate value of the standard deviation of the distribution shown below.



7. What are the causes of variability? Which one is more common?  
(Ans. Common cases, Special causes, 94%, 6%)

8. When is a process considered to be in CONTROL (or STABLE)?

Plot a time history graph of the following four situations.

- No variation in Mean (a) and Spread (s)
- No variation in Mean (a)
- No in Spread (s)
- There are variations in Mean (a) and Spread (s)

9. Plot the graph for the following scenarios (use separate pages to draw 2-d and 3-D graph with time axis).

A process that is:

- (a) STABLE and within SPECS.
- (b) STABLE but out-of-SPECS.
- (c) STABLE, out-of-SPECS, but CAPABLE.
- (d) STABLE, out-of-SPECS, and INCAPABLE.
- (e) UNSTABLE but in-SPECS.

- (f) UNSTABLE and out-of-SPECS.
- (g) UNSTABLE but improving mean
- (h) STABLE but improving spread

10. Hotel operator Mr. Marriott routinely fills most of his 800-room hotel in a major metropolitan city. Analysis of the occupancy data (number of rooms sold each night) for past several years show average occupancy of 700 rooms and a standard deviation of 12 rooms. For last few weeks, the hotel's occupancy jumped to 725 (average occupancy). Should Mr. Marriott initiate a hotel expansion project?

### Process Capability Indices ( $C_p$ & $C_{pk}$ )

11. A newly installed electricity generating plant is found to operate with the following data: Average voltage = 112 volts and Standard Deviation = 2.93 volts. The target voltage is 110 and the specification limits are  $USL = 120$  and  $LSL = 110$  volts. Calculate the generator process potential ( $C_p$ ) and the process capability ( $C_{pk}$ ).

(Ans.  $C_p = 1.137$   $C_{pk} = 0.909$ )

12. Plot a typical data/distribution and label  $USL$ ,  $LSL$ , Target, and Mean of the data. Also, write the expressions for  $C_p$  and  $C_{pk}$ . Answer following questions.

- (a) When is  $C_{pk}$  same as  $C_p$ ?
- (b) How can  $C_p$  value be increased?
  - Increase ( $USL - LSL$ )
  - Decrease standard deviation ( $s$ )
  - Increase ( $USL - LSL$ ) and decrease  $s$
- (c) What is the desirable location of the mean for better  $C_{pk}$ ?
- (d) If you want over 99% good production part, what should the  $C_{pk}$  of the process be?
- (e) How are  $C_p$  and  $C_{pk}$  related? Can  $C_{pk}$  ever exceed  $C_p$ ?
- (f) How can  $C_{pk}$  value be increased?
  - Move mean closer to target
  - Reduce standard deviation
  - Move mean closer to target and reduce standard deviation
  - All the above plus, wide  $USL/LSL$  further

13. The foreman of a machine shop was to select the best process among the two competing production line producing *resistors* for TV-appliance. The following specifications and production data were available:

Target value = 75 Ohms,  $LSL = 55$        $USL = 95$

Measured data from sample production output:



Line 1: 66 76 83 80 73 75 84 and 69 ( $n = 8$ )

Line 2: 74 79 67 68 75 74 76 67 and 71 ( $n = 9$ )

(Ans:  $\bar{x} = 75.75$   $s = 6.408$   $\bar{x} = 72.333$   $s = 4.301$ ,  $cp=1.040$   $Cpk = 1.00$   $Cp = 1.55$   $Cpk = 1.343$  )

14. All things remaining the same, if the standard deviation of the process parameter is reduced to its 50% level, what will happen to the Cpk index?
15. What happens to the Cpk index when the specification limits are arbitrarily narrowed to half of the current range?

### Construction of Control Charts (SPC)

#### QA:

Q1. What are control charts?

Ans: Control charts are powerful tools that can be used to monitor the performance when appropriate measurements from a manufacturing or service delivery process are available.

Q2. What are typical measurement data for which control charts (X-bar and R charts are most common) are appropriate?

Ans: The data must be of variable type. Typical measurements are: Length (in centimeters, inches, etc.), Width, Diameter, Time, Strength, etc.

Q3. What are X-bar and R charts any way?

Ans: X-bar and R charts are nothing more than plots of process X-bar and R over time, with two additional lines (UCL & LCL) drawn on the plots. The two additional lines are called the **control Limits**.

Q4. What are benefits of the control charts (Say for X-bar and R charts)?

Ans: A variable data control chart is used to determine a process performance characteristic in terms of both its location (mean of the characteristic) and its spread (variability of the characteristic). For this reason, control charts for variable data are used in pairs – one chart (X-bar chart) to assess centrality and the other to evaluate process spread (R chart).

Q5. How do control charts indicate if the process is out of control?

Ans: Control chart represents the normal or standard process performance. Obviously, control charts are constructed by collecting data from the process when it is STABLE. Current and future performance data then are compared against the expected normal behavior. Observation about abnormal behavior then can be made by comparing the performance with statistically valid characteristics of normal distribution (99.7% within USL & LSL).

Q6. What is a common data sampling strategy?

Ans: For both construction of control chart and monitoring of performance, a common practice is to collect data 25 samples of data with 5 data in each sample (5x25 readings). Average of each

sample is  $\bar{X}$  and range of a sample is  $R$ . The average of all sample averages is called  $\bar{\bar{X}}$  and the average of sample ranges is called  $\bar{R}$ .

Q7. Is standard deviation for 125 samples needed for the control chart?

Ans: Because it is difficult to directly calculate standard deviation of the total data (125), approximate, but simpler method is used to calculate standard deviation. For control charts upper and lower control limits are usually calculated from  $\bar{R}$ .

Q8. What are plotted in the  $\bar{X}$ -bar chart?

Ans: After the control chart is constructed (with  $\bar{\bar{X}}$ , UCL, and LCL),  $\bar{X}$ -bar of each collected sample (of 5 data) are plotted in the  $\bar{X}$ -bar chart.

Q9. Why are the upper and lower control limits (LCL & UCL) of  $\bar{X}$ -bar chart set at three standard deviation of  $\bar{X}$ -bars apart.

Ans: Three standard deviations from the mean are expected to contain 99.7% of the data.

Q10. What are the steps in constructing  $\bar{X}$ -bar control chart?

Ans: Step 1 – Collect Production data (variable data, 25 sets of 5 data in a set.), Step 2 – Record data in a table form, Step 3 – Calculate  $\bar{X}$ -bar and  $R$  for all 25 samples, Step 4 – Calculate  $\bar{\bar{X}}$  and  $\bar{R}$ , Step 5 – Calculate UPPER and LOWER CONTROL LIMITS, Step 6 – Scale the vertical axis of the  $\bar{X}$ -bar chart, Step 7 – Plot the values of the  $\bar{X}$ -bar in the control chart, Step 8 – Draw the center line for  $\bar{\bar{X}}$ , and Step 9 – Draw the control limits.

Q11. How are control limits calculated?

Ans: Control limits are calculated as follows.

$$\text{UCL (for } \bar{X}\text{-bar)} = \bar{\bar{X}} + 3 \times \text{Std. Dev. Of } \bar{X}\text{-bars}$$

$$\text{LCL (for } \bar{X}\text{-bar)} = \bar{\bar{X}} - 3 \times \text{Std. Dev. Of } \bar{X}\text{-bars}$$

Where  $\text{Std. Dev. Of } \bar{X}\text{-bars} = (\text{Std. Dev. Of } X) / \text{Sq. Root of } n \quad (n = 5)$

Also  $\text{Std. Dev. Of } X = \bar{R} / d_2 \quad (d_2 = 2.326 \text{ for } n = 5)$

Or  $\text{Std. Dev. Of } \bar{X}\text{-bars} = \text{Constant} \times \bar{R}$

$$\text{UCL (for } \bar{X}\text{-bar)} = \bar{\bar{X}} + A_2 \times \bar{R} \quad (A_2 = 0.577 \text{ for } n = 5)$$

$$\text{LCL (for } \bar{X}\text{-bar)} = \bar{\bar{X}} - A_2 \times \bar{R}$$

(See class notes or seminar materials for formula notations)

Q12. What does it mean when we say a process is out of control?

Ans: There are a number of ways a process can be out-of control. **Out-of-control** means that over time there are **changes in either the center** of the process or the **spread** of the process.

Q13. What is the nature of common changes?

Ans:

*Gradual change* in the center or the spread (worker fatigue, change of working condition, ),

*Abrupt and persistent* ... (change of operator, reaction to new setup, ),

*Abrupt but temporary change...*,  
*Periodic movement* in ...(weekly environmental changes, operator rotation, ),  
*Erratic or systematic* changes in the center or the spread (due to special & specific causes).

Q14. What are some of the common rules for detecting out-of-control patterns?

Ans:

- a. Single point outside the control limits
- b. Three consecutive points between 2 & 3 Std. Dev.
- c. Seven points in a row on one side of the X-double-bar line (Run)
- d. Six points in a row either increasing or decreasing (Trend)

Ref: Normal Distribution

+/- s = 68.27%, +/- 2s = 95.45%, +/- 3s = 99.73%, +/- 4s = 99.9937%

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \dots -\infty \leq x \leq +\infty$$

## Program Evaluations

Program Title \_\_\_\_\_

Program completion date \_\_\_\_\_

Instructor Ranjit K. Roy & .... Training Location/host. \_\_\_\_\_

We appreciate your comments and suggestions. Please take a moment to let us know how we can improve and serve you better. (Please use the following numbers for evaluation purposes).

[6] Excellent [5] Very Good [4] Satisfactory [3] Poor [2] Unsatisfactory [1] Needs Improvement

1. Overall Reaction to Program: 6 5 4 3 2 1 (Circle one number)

Comments: \_\_\_\_\_

2. Reaction to Instructor: 6 5 4 3 2 1 (Circle one number)

Comments: \_\_\_\_\_

3. What should be added and/or deleted to improve this program? (Please be specific)

\_\_\_\_\_

4. What portions of this program do you feel will be most helpful at your work? How?

\_\_\_\_\_

5. Do you feel the training provided you enough understanding of the technique for you to be able to start applying it to your own projects if opportunities were available?

Yes       No       May be

6. Would you consider/recommend us for training at your facility?  yes  No  May be

Who should we contact for training at your facility \_\_\_\_\_ Ph: \_\_\_\_\_

7. Would you recommend that others attend this seminar?  YES  NO

8. (Optional)

Your Name \_\_\_\_\_ Your Ph# & E-mail: \_\_\_\_\_

THANK YOU for taking the time to complete this evaluation.