Section III

Measure

Agenda

- ✓Pre-Measure Considerations and Tools
- ✓Types of Data and Measurement Scales
- ✓Central Tendency and Dispersion
- ✓Measurement System Analysis – Variables GAGE RR and Attribute RR
- ✓Stability Conditions
- ✓Capability Metrics (Cp, Cpk, Cpkm, Pp, Ppk)
- ✓Variations, Variability, Capability, and Process Conditions
- ✓Data Distributions (Normal, Binomial, Poisson, Exponential)
- ✓Sigma Shift, Mean Shift, and Reducing Variation
- ✓Baseline Data

Section III, Lesson 1

Pre-Measure Considerations and Tools

Agenda

- ✓ Define Phase Toll Gate Review
- ✓ DFMEA
- ✓ Cause and Effect Matrix

Define Phase Toll Gate Review

The define phase toll gate review:

- ✓ Post completion of the Define phase, a toll gate review is conducted with the Six Sigma team and management.
- ✓ The Sponsor presides over the toll gate review meeting. The Black Belt conducts the review meeting and presents the project in front of everybody.

The problem statement and the key project area opportunity are discussed upon.

- ✓ The project deadline and the possible allocation of funds for the project are finalized by the Sponsor.
- ✓ The Six Sigma team should prepare additional tools like the DFMEA (Design Failure Mode And Effects Analysis) and the project storyboard before the toll gate review meeting.

DFMEA

- ✓Stands for design failure mode and effects analysis (a.k.a Design FMEA).
- ✓Other types of FMEA include process FMEA and system FMEA.
- ✓At the measure stage, the Six Sigma team must update only the key process steps in the FMEA document. These process steps then feed on to KPOVs (key process output variables) that eventually drive the project.
- ✓The DFMEA template is ideal when updated post the completion of cause and effect matrix, which will be covered as the next tool.

The DFMEA template is provided as part of the toolkit titled in a separate excel worksheet design FMEA template.

DFMEA (Contd.)

How to use the DFMEA tool?

✓The DFMEA tool contains three main worksheets:

- Prepare FMEA;
- FMEA; and
- Data drawings and pictures.

✓Prepare FMEA: Provides a checklist for all the necessary items to do an FMEA.

✓FMEA: The core DFMEA template.

✓Data drawings and pictures: Helps provide supplementary drawings to the FMEA template.

Snapshot of prepare FMEA worksheet

	Failure Mode and effect analysis IN PROJECT **Prepare FMEA**	Date : yyyy-mm-dd
Concerned project		

Main background documents for this analysis

- [ ] Illustration material (hardware) - test part, prototype
- [ ] A principle layout of the system design and all drawings included
- [x] Lists of parts/components, material and function specifications
- [ ] Available data from suppliers/Other relevant (similar) reliability analyses
- [ ] Available failure reports from the field/Results from market surveys/ Campaigns related to the product/system
- [ ] Customer complaints (Features, failures/warranty claims) on the current product
- [ ] Available failure reports from current production/Campaigns related to the product/system
- [ ] Previously made FMEAs
- [ ] Other, eg. Data from any other experience banks, such as environmental effects in connection with failure.

Limits of analysis	
Meeting appointment :	Date - Time - Place
Mandatory attendees	☐ System FMEA ☐ Design FMEA ☐ Process FMEA
FMEA Facilitator:	
Design Engineer:	
Test Engineer:	
Reliability/Quality Control Engineer:	
Environmentally competent person:	

DFMEA (Contd.)

Increase page length	FMEA - FAILURE MODE AND EFFECT ANALYSIS SYSTEM/DESIGN/PROCESS					
Main system / Sub-system	Part/Operation name.	Drawing-Part No.	Machine No.	Functional Specification-Technical Regulation (link if possible)	Supplier	
Main Function / Operation	Date Performed	Date Time Updated	Status - Hardware/Digital modell	Issued by	Project	Issue

PART		CHARACTERISTICS OF FAILURE				1st RATING				ACTION-STATUS				2nd RATING				
No	Function/Part / Operation	Failure mode	Causes of failure	Undesirable customer effects	Testing - Simulation	Po	S	Pd	RPN	Recommended action	Planned action	Responsible	Follow up date	Po	S	Pd	RPN	Verifications-
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		

DFMEA (Contd.)

How to update the DFMEA template?

- ✓Function – Update the key process step.
- ✓Failure mode – Update the point or the key issue that could happen.
- ✓Causes of Failure – Update what caused the key issue to happen.
- ✓Understand customer effects – Update the key issue to the customer due to failure mode.
- ✓1st rating – Provide severity, occurrence, and detection ratings as per guidelines.
- ✓Recommended action – Add recommended action (updated in the Improve phase).
- ✓Planned action – Add planned action post first stage validation of Improve pilot run.
- ✓2nd rating – Provide revised ratings post control phase deliverables done.

Cause and Effect Matrix (CE Matrix)

- ✓ CE matrix is a prioritization tool that helps in prioritizing causes/input variables on the basis of the relationship of input variables with CTQs of the process.
- ✓ Look into the CE matrix template provided as part of the toolkit in the worksheet titled CE matrix.
- ✓ The CTQs are placed on the top of the matrix.
- ✓ The input variables are placed on the vertical side of the matrix.
- ✓ The correlation scores for the input variables and the CTQs are provided in the matrix.
- ✓ The highest overall scores are prioritized using Pareto charts.

Cause and Effect Matrix (Contd.)

How to update CE matrix?

✓ List CTQs across the top of the matrix.

✓ Rank and assign scores to each CTQ according to the importance to customer. Scores are provided on a scale of 1-9. The CTQ that has least importance is given a score of 1.

✓ List all input variables on the left side of the matrix.

✓ Determine correlation scores in the matrix. 1 – weak correlation, 3 – moderate and 9 – strong correlation.

✓ Prioritize correlation scores as per the overall scores.

Cause and Effect Matrix (CE Matrix) (Contd.)

- Snapshot of CE matrix

Process Output									To add a column, select the last 'Process Output' column, right-click and select insert. To add a row, copy and paste from any 'Process Step / Input' row.
Importance to Customer									
Process Step / Input	**Correlation Scores**								**Total**
									0
									0
									0
									0
									0
									0
									0
									0
									0
									0

Cause and Effect Matrix (CE Matrix) (Contd.)

Cause and Effect Matrix

				Key Process Output Variables														
			Customer Importance															
			Customer Rank															
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
		Process Step	KPIV	KPOV1	KPOV2	KPOV3	KPOV4	KPOV5										
Variables	1	Step 1	KPIV1															
	2		KPIV2															
	3	Step 2	KPIV3															
	4		KPIV4															
	5																	
	6																	

Cause and Effect Matrix (CE Matrix) (Contd.)

- ✓A CE matrix example: A supply chain company works on CTQs of delivery hours, packaging defects, and delivery cost. These would be the KPOVs (key process output variables).
- ✓The input variables that form integral portion of the process are listed below:

a) Delivery allocation time

b) Delivery units

c) Number of resources for packaging

d) Experience used in packaging

Let us look at the CE Matrix and see how it is updated.

Cause and Effect Matrix

				Key Process Output Variables																	
			Customer Importance																		
			Customer Rank	10	5	3															
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15			
		Process Step	KPIV	Delivery Hours	Packaging Defects	Delivery Cost													Rank	Rating	Total
put Variables	1	Step 1	Delivery allocation time	9	1	1													2		98
	2		Delivery units	5	9	9													1		122
	3	Step 2	# Resources	1	9	9													3		82
	4		Experience	1	9	5													4		70

Summary

- ✓Define phase toll gate review
- ✓Design failure mode and effects analysis
- ✓Cause and effect matrix

Section III, Lesson 2

Types of Data and Measurement Scales

Agenda

- ✓Objectives of Measure Phase
- ✓Process, Flowcharts, SIPOC Map
- ✓Metrics and Measurement Scales
- ✓Types of Data

Objectives of Measure Phase

- ✓Process definition – Ensure specific process under investigation is clearly defined.
- ✓Metric definition – Define reliable means of measuring the process relative to project deliverables.
- ✓Establish project baseline – Establish current operating metrics performance.
- ✓Evaluate measurement system – Validate reliability of data to ensure meaningful conclusions are drawn.

What is a Process

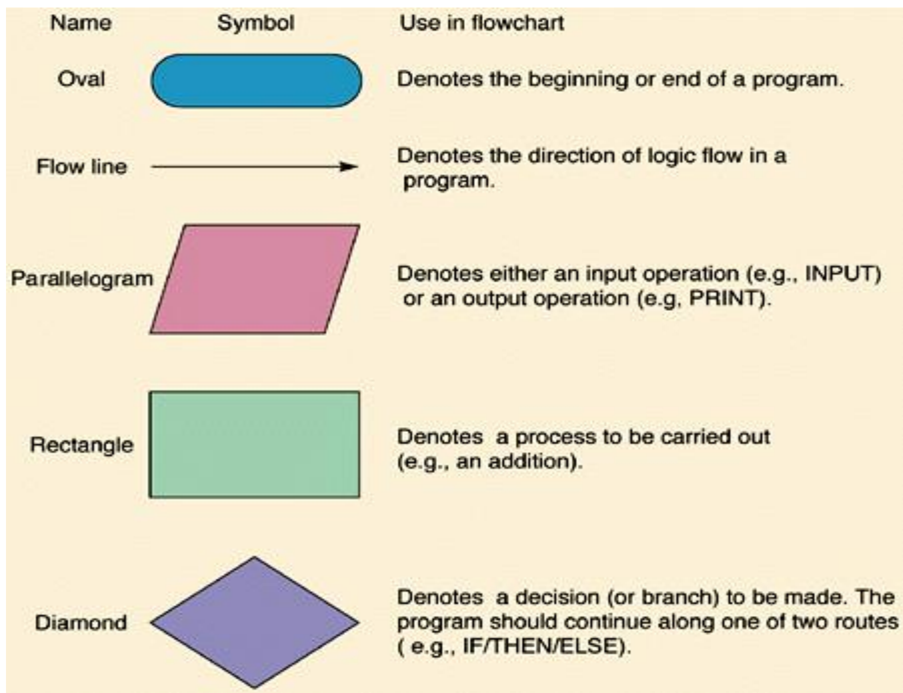
- ✓A process has a series of repeatable tasks carried out in a specific order.
- ✓If a process cannot be defined by the people who are using the process, it probably means multiple processes exist or a well defined process doesn't exist.
- ✓A well defined process helps in making the Six Sigma efforts successful.
- ✓Process can be defined by flowcharts, SIPOC, and process maps.

Flowcharts

- ✓A flowchart is a basic graphical format of how the current process is, as on date.

Important – Such flowcharts called As-Is flowcharts are often drawn in the measure stage.

- ✓A flowchart is a pictorial way of presenting process, its logic flow, decision points, external, and internal interfaces, etc.
- ✓Flowcharts are often drawn with the help of ANSI (American National Standards Institute) symbols.



Source: http://www.wiley.com/college/busin/icmis/oakman/outline/chap05/slides/symbols.htm

SIPOC

- ✓The Define phase helps to identify projects that can be implemented enterprise wide, yet most of these projects fail due to a large scope.
- ✓A SIPOC map is a broad-level map that allows to identify the scope of the project, by setting project boundaries.
- ✓A SIPOC map is a mandatory requirement for all Six Sigma projects, because without this map the scope of the project may not be determined.
- ✓Scoping can be done on the process as well as at the customer's end.

SIPOC (Contd.)

- **Questions to be asked before preparing SIPOC are:**
  - Who is the stakeholder for the process?
  - What is the value for the process? What value is created for the customers?
  - Who is the owner for the process?
  - Who provides input for the process?
  - Are the inputs consistently of good quality?
  - What are the inputs?
  - What resources does the process use?
  - What steps create value?

- SIPOC stands for suppliers, inputs, process, output, and customers. This is always conducted by starting with a focused brainstorming session.

SIPOC (Contd.)

A snapshot of the SIPOC map is given here. SIPOC is also provided in the form of an excel spreadsheet in the toolkit.

SIPOC					
Suppliers	**Inputs**	**Process**	**Outputs**	**Customers**	
(resource provider)	(process)	(high level process flow)	(from the process)	(receives an output from the process)	
Coffeemaker purchased - on countertop	>5 cup capacity coffee maker	**Making coffee**	heating to keep coffee warm for 1 hour after brewing	All of us enjoy the same brand of coffee with varying condiments	
city water supply into faucet	water supply	Add water	enough coffee to serve all of us within 15 minutes of start time.	1 cup of coffee	wife
purchase from XYZ company	1 filter	Add filter & ground coffee	one filter to prevent overflow	1 teaspoon of french vanilla creamer	
Use ABC brand beans	4 tablespoons of coffee grinds		correct amount of grinds		
Electric company	120V GFCI outlet	Plug-in and turn on		1 cup of coffee	husband
Upper left drawer next to refrigerator	measuring spoons		Source to heat water to temperature	Honey on the table	
			Pump to move water up through filter.	Dash of cinnamon	
Mugs purchased - in upper left cabinet	coffee mugs	Pour into mug	Hot coffee filled near the top of the mug.		
refrigerator and pantry	condiments and containers for sugar, creamer, honey, cinnamon.	Add condiments		1 small cup of coffee for each	
ABC brand for sugar & creamer.			coffee served in spouses favorite mug	1 teaspoon sugar	
XYZ brand for honey and cinnamon		Stir	coffee served in husbands mug	1 tablespoon french vanilla creamer	
pantry	stirrers, lids	Serve	coffee served in personalized kids mugs	Let sit for 5 minutes before serving	2 teenagers

Source: http://www.six-sigma-material.com/images/SIPOCExample.GIF

Metrics

- ✓Metrics should be considered in the Measure stage, as it quantifies the project deliverable.
- ✓Metrics are typically defined with the help of **process expertise.**
- ✓Metrics in some cases could also be process end-point measurements. For example, in chemical and manufacturing sectors, the customers may define their end-product requirements and not really process metrics.
- ✓Measurement is just a numerical value assigned to an observation.

Measurement Scales

- ✓Broadly, we have four types of measurement scales – Nominal, ordinal, interval, and ratio.
- ✓Nominal measurements are purely indicative.
- ✓Nominal measurements are also considered the weakest form of measurements or the weakest type of measurement scales, as they reveal very less information about the data itself.
- ✓Nominal scales define or classify only the presence or absence of a certain attribute. With nominal scales only the items can be counted; and the statistics used are percentage, proportion, and chi-square tests.

Measurement Scales (Contd.)

Ordinal scale

✓ Defines one item having more or less the attribute than another item. In simple words, a comparative scale of measurement.

✓ Example – Taste and attractiveness. The food tastes better than the drink.

✓ As a statistical measure, one can use rank order correlation to quantify the comparative data between the two.

Measurement Scales (Contd.)

Interval scale

✓Defines the difference between any two successive points. The difference may or may not be equal for the data to be treated on interval scale.

- Example – Calendar time and temperature.

✓Correlations, f-tests, t-tests, and regression can be used with interval scales.

Ratio scale

✓Defines all true zero value points. One can perform addition, subtraction, division, and multiplication with such data points.

- Example – In calculating elapsed time, distance etc., one can do mathematical calculations like addition, subtraction, division, and multiplication with the data.

Types of Data

- ✓Understanding the type of data being dealt with, provides us with a more intuitive way of understanding the data.
- ✓In other words, measurement scales form the foundation, while the types of data form the 'front-end' of the understanding.
- ✓We can categorize data in three divisions – continuous or variables, discrete, and attribute data.
- ✓Data collected on nominal scale are attribute data. The only possible relation with attribute data on a nominal scale is equality {=} and inequality {≠}.
- ✓Ordinal attribute data is often converted to nominal data and then further analyzed using binomial and Poisson distribution.

Types of Data (Contd.)

- ✓Discrete data –These would be data values that can be counted and take the form of non-negative integers.
  - For example, the number of defective parts in a shipment, the number of failed calls in a process and so on.
- ✓Continuous data – These would be data values that can be measured using a measurement system. Such data values can take any whole number in the finite or infinite set.
  - For example, the time taken for a shipment to reach the customer.

A quick assignment – Know what the customer wants in terms of deliverables. Map the project and find out the type of data that would be worked with.

Summary

In this lesson, we have learned:

- ✓ The objectives of measure phase
- ✓ Process, flowcharts, SIPOC map
- ✓ Measurement scales
- ✓ Types of data

Section III, Lesson 3

Central Tendency and Dispersion

Agenda

- ✓ Central Tendency and Dispersion – Introduction
- ✓ Mean
- ✓ Median
- ✓ Mode
- ✓ Range
- ✓ Variance
- ✓ Standard Deviation
- ✓ Mean Deviation

Central Tendency and Dispersion – Introduction

- ✓ A sample set of data can be easily represented with the help of measures of central tendency and dispersion.
- ✓ Central tendency represents the sample average while dispersion broadly represents spread or variation in the data.
  - For example, if one would wish to report the scores of 50 students in a class, it will be cumbersome to represent individual scores. The best method is to find out the class average and also, the dispersion in the class scores. Put together, the data would represent the class effectively. Such as, the class average is 62 and the range of the scores is 50.

Mean

- ✓The popular and often used measure of central tendency in Six Sigma applications.
- ✓Mean is the arithmetic average of all the data points and is popularly known as arithmetic mean.
- ✓Mean is extremely sensitive to the values in the data set. Extreme values can skew the result of the mean, rendering it unusable. In the case of extreme values, also known as outliers, one should use median instead of mean.
  - Data set example: 2,3,4,3.5,4.5,5.5,6,6.5,7.

Mean (Contd.)

- ✓Use the function, =AVERAGE() in Microsoft Excel. The mean for this data list is 4.66.
- ✓Consider a modification. Add a new number 100 to the list. The number 100 here is considered an extreme value (outlier).
- ✓Use the same function, =AVERAGE() in Microsoft Excel.
- ✓The mean now changes to 14.2. Interpreting this mean gives us an interesting insight on why one shouldn't use mean when data has outliers. Manually check the data. It can be found that, out of 10 values only 1 lies above 14.2 and the rest lie below 14.2.
- ✓Clearly, this cannot be a measure of central tendency.

Median

- ✓ Median, also known as positional mean, is the central location or position of the data list.
- ✓ With the same data list taken in the example of the mean and the results are as below, using the function = MEDIAN().
  - Median (Data without outlier) = 4.5
  - Median (Data with outlier) = 5
- ✓ The median doesn't change much even with the introduction of outliers.

Imp – Whenever extreme values or outliers in the data set is seen, try, and use median instead of mean.

Mode

- ✓Mode, also known as frequency mean, is one of the least used measures of central tendency in Six Sigma applications.
- ✓Mode represents the value in a data list which repeats itself the most.
- ✓The data list given in the example of mean calculations is used. Use the function, =MODE() to find the mode of the data list.
- ✓The result says #N/A, which means this data list doesn't have any mode.
- ✓**Interpretations –** Presence of one mode in a data list is a must to proceed any further in a Six Sigma project case. Presence of two modes (bi-modal data) means the reasons of two modes should be looked into and one should be screened out.

One cannot proceed ahead in the Measure phase with bi-modal data.

Range

- ✓Range is a measure of dispersion or it is the difference between the maximum value and the minimum value of the data list.
- ✓Range doesn't need the presence of all the values in the data list to move ahead with calculations. One only needs the largest and the smallest values in the data list.
- ✓Range doesn't work well for sample sizes greater than 10 (statistically proven).
- ✓The data list for calculating measures of dispersion is 2,3,4,3.5,4.5,5.5,6,6.5,7.
- ✓Use =MAX() to find the largest value and = MIN() to find the smallest value.
- ✓Subtract the smallest value from the largest. In this case, the range is 7-2=5.

Variance

- ✓Variance is a measure of dispersion and is defined as the average of squared mean differences.
- ✓With the same data list in Microsoft Excel, one can use the function = VAR(). Please note that = VAR() is used to calculate sample variance and = VARP() is used to calculate population variance.
- ✓In Six Sigma project applications, sample variance is used more than population variance.
- ✓Using the formula =VAR(), the sample variance for the data list is 2.875. Variance is unit-less and 2.875 represents the difference or dispersion in the data.

Standard Deviation

- ✓ Standard deviation is a measure of dispersion and is technically defined as the square root of variance. In simple words, standard deviation is the value of difference from the mean.
- ✓ It is represented by the Greek alphabet, σ, popularly called as sigma.
- ✓ Standard deviation is a very popular measure of dispersion.
- ✓ Use the formula, =STDEV() to calculate sample standard deviation and =STDEVP() to calculate population standard deviation. Work with sample standard deviation in Six Sigma project applications most of the times.

Mean Deviation

- Mean deviation is not a frequently used measure of dispersion. Technically, Mean deviation is the mean of absolute deviations from the mean.

- Using the data list, 2,3,4,3.5,4.5,5.5,6,6.5,7 and the mean of 4.66, the mean deviation is 1.41. 1.41 indicates that the data is spread 1.41 from the mean of the data list.

	Data list	**Absolute Deviation from Mean**
	2	2.67
	3	1.67
	4	0.67
	3.5	1.17
	4.5	0.17
	5.5	0.83
	6	1.33
	6.5	1.83
	7	2.33
Mean	4.666667	
Mean Deviation	1.41	

Summary

In this lesson we have learned:

- ✓Mean
- ✓Median
- ✓Mode
- ✓Range
- ✓Variance
- ✓Standard deviation
- ✓Mean deviation

Section III, Lesson 4

Measurement System Analysis

Agenda

- ✓Purpose of Measurement System Analysis
- ✓Measurement System Errors
- ✓Properties of Good Measurement Systems
- ✓Measurement System Errors Illustrated
- ✓Measurement System Discrimination
- ✓Measurement System Analysis – Process Flow
- ✓Part Variation
- ✓Measurements Systems Analysis Formulas
- ✓Measurement Systems Analysis Example
- ✓Measurement Systems Analysis Graphs
- ✓Attribute RR
- ✓When to Do Measurement System Analysis
- ✓Data Collection Plan

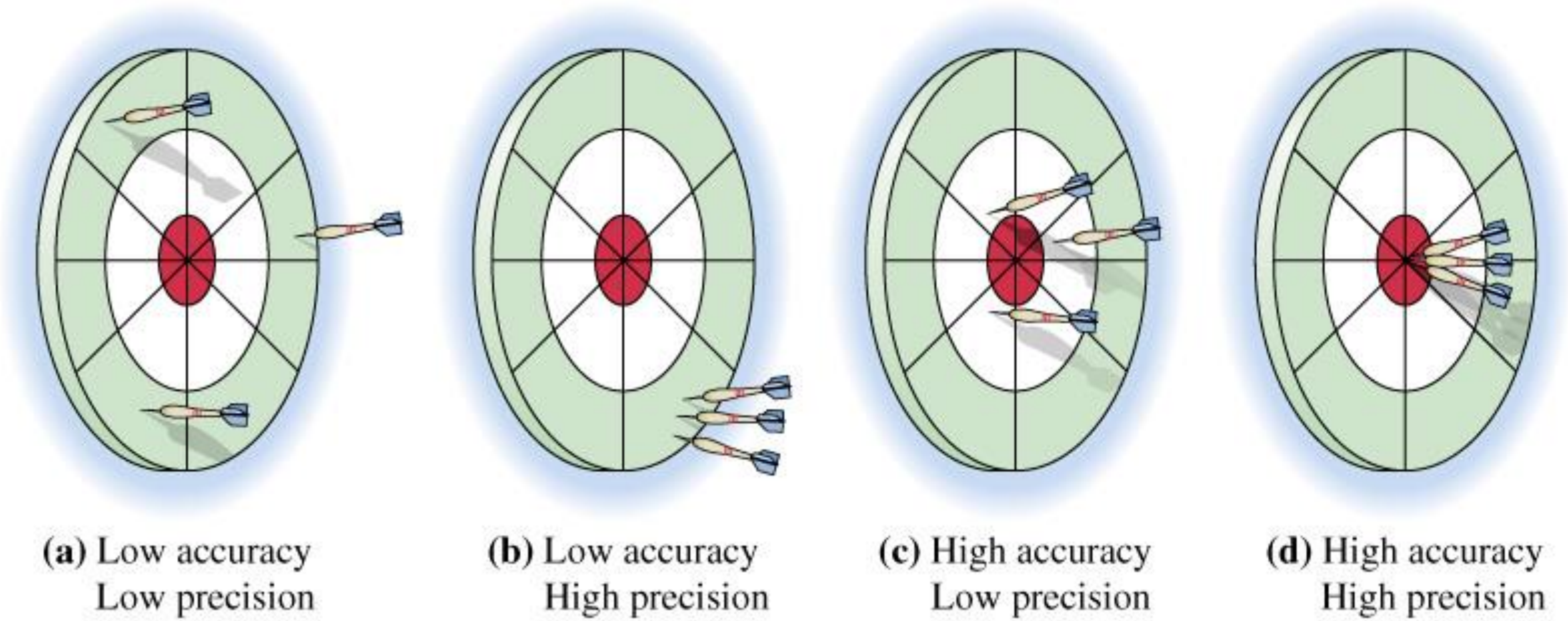
Purpose of Measurement System Analysis

- ✓ Measurement system analysis (MSA) or gage repeatability and reproducibility (Gage RR). It is a tool that is used to validate the measurement system to see if the measurement system is able to give the data that can be relied upon.
- ✓ If measurement system is not validated, the data coming out of the measurement system would be questionable.
- ✓ The main purpose of MSA is to verify if the measurement system is able to provide:
  - Representative values of characteristic being measured;
  - Unbiased results; and
  - Minimal variability.

Measurement System Errors

- ✓ Resolution – It is the lowest increment of measure, the measurement system can display.
- ✓ Accuracy – It is the difference between observed measurement and actual measurement.
- ✓ Precision – It is the difference observed when measuring the same part with the same equipment or instrument. Precision is further categorized into repeatability and reproducibility.
  - Repeatability = When the same operator measures the same part with the same instrument.
  - Reproducibility = When different operators measure the same part with the same instrument.

Measurement System Errors (Contd.)



Source: http://extensionengine.com/accuracy-precision/

Measurement System Errors (Contd.)

✓Accuracy further has three components to be studied in the order:

- Stability – Consistency and predictability of measurements over time.
- Bias – Tendency of measurement to provide different value than the actual value is called Bias. Bias can be positive or negative.
- Linearity – A measure of bias values through range of measurements.

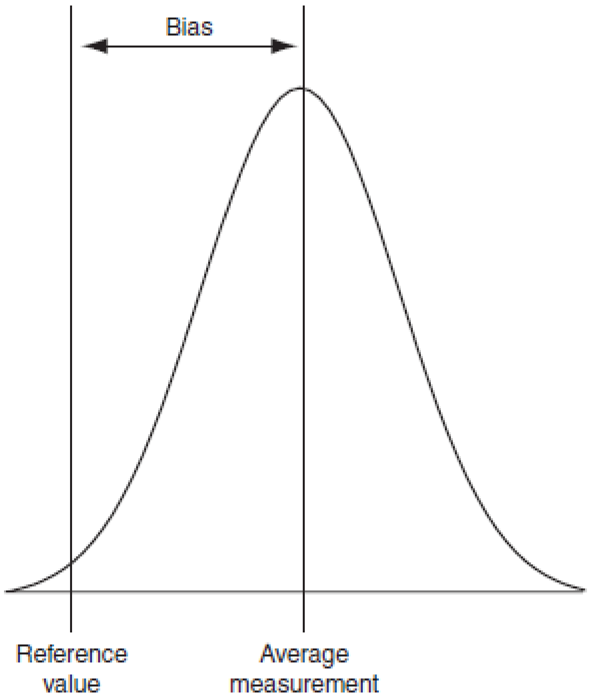
✓Precision has two components need to be studied:

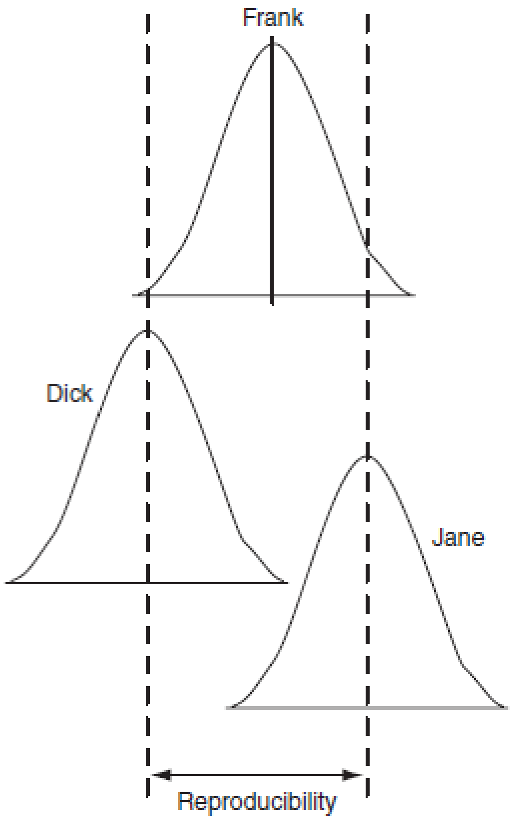
- Repeatability, also known as equipment variation.
- Reproducibility, also known as appraiser variation.

Properties of Good Measurement Systems

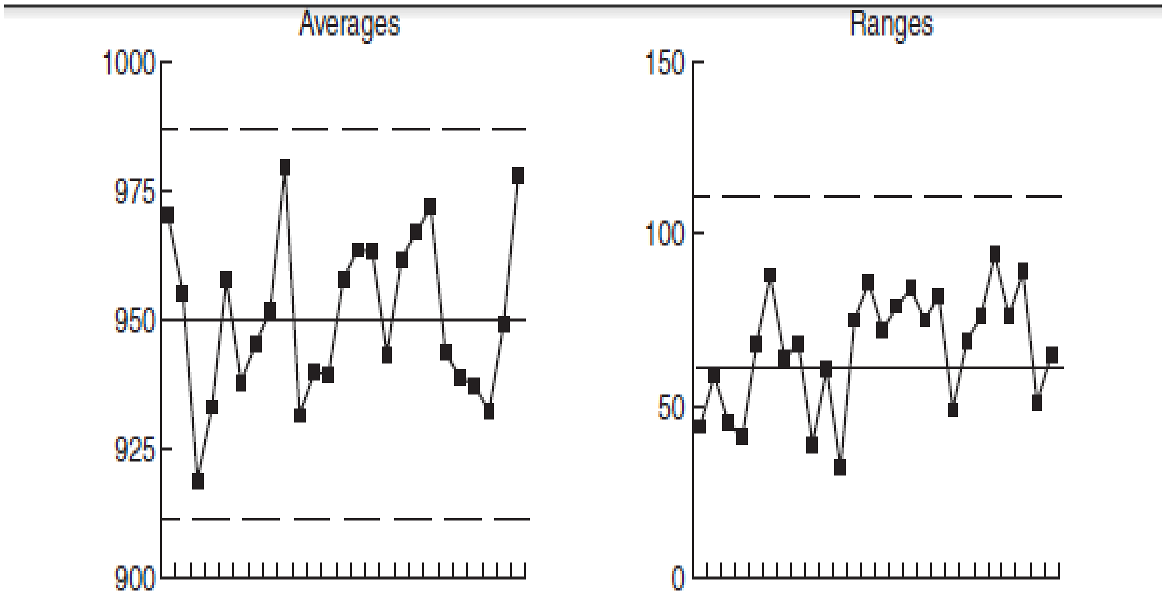
- ✓ Should produce a number that is close to actual property being measured – **Accuracy.**
- ✓ Should produce the same measurements when measured again by same equipment, person, and under same condition – **Repeatability.**
- ✓ Should produce consistent results in an expected range – **Linearity.**
- ✓ Should produce almost the same measurements (or the variation should be smaller than the agreed limit) when measured by different individuals or different equipment or at different location – **Reproducibility.**
- ✓ Should produce same results in the future, as it did in the past – **Stability.**

Measurement System Errors Illustrated





Measurement System Discrimination



Adequate Gage Discrimination as plotted on Xbar – R Chart.

Bias

- ✓ A standard with a known value (reference value) of 25.4 mm is checked 10 times by one mechanical inspector using a dial caliper with a resolution of 0.025 mm. The readings obtained are:
  - 25.425 25.425 25.400 25.400 25.375
  - 25.400 25.425 25.400 25.425 25.375

- ✓ The tolerance value of this gage is ± 0.25 mm. Calculate the bias value and provide interpretation.

Bias (Contd.)

Solution:

1. Find the average of measurements (Xbar) = 25.405 mm
2. Bias = Average – Reference value = 25.405 – 25.4 mm = 0.005 mm
3. % Bias = 100 * (Bias/Tolerance) % = 100 * (0.005/0.25) % = 2%

✓**Interpretation:**

On an average, the measurement system will produce readings 0.005 mm greater than the actual reading. This shows 2% of the variation is contributed due to bias in the measurement system. This concludes that the measurement is in acceptable range.

Measurement System Analysis – Process Flow

1. Prepare for study
2. Evaluate stability
3. Evaluate discrimination or resolution
4. Determine accuracy
5. Calibrate if necessary
6. Evaluate linearity
7. Determine repeatability and reproducibility
8. Determine part variation

Part Variation

- ✓The Xbar or the averages chart are plotted to capture part variation, or part to part variation and is also known as time series variation.

- ✓If the measurement system is adequate, most parts would fall outside the control limits of the averages chart.

- ✓If very few parts fall outside the control limits of the averages chart, the measurement system is not adequate.

- ✓Part to part variation is determined only when measurement system has:
  - Adequate discrimination;
  - Stable;
  - Accurate;
  - Linear; and
  - Consistent repeatability and reproducibility.

Measurement System Analysis Formulas

σ_m = Gage variation (Gage repeatability and reproducibility – GRR)

σ_p = Part variation (PV)

σ_t = Total process variation = $(\sigma_m{}^2 + \sigma_p{}^2)^{1/2}$

$\sigma_m = (\sigma_r{}^2 + \sigma_o{}^2)^{1/2}$

where σ_r = Repeatability or equipment variation

σ_o = Reproducibility or appraiser variation

% GRR = 100 * σ_m /tolerance %

% PV = 100 * σ_p /tolerance %

ndc = 1.41 * PV/GRR

If %GRR > 30%, gage needs to be replaced.

If ndc < 5, gage needs to be replaced.

Measurement Systems Analysis Example

A measurement system study is conducted on 10 parts. Three operators measure the 10 parts on 3 trials. The specification levels provided are 9 and 11. The operators measure all the 10 parts on three trials. Conduct a variables GAGE RR study and interpret the results.

Solution:

1. Use the GAGE RR worksheet provided in the toolkit.
2. Find the operator readings in the worksheet under columns highlighted Part 1, Part 2, and so on.
3. Check the %GRR value. It is 5.57%. Therefore GAGE is ok.
4. Check the ndc. It is 3. Thus, GAGE discrimination needs improvement.

Measurement Systems Analysis Example (Contd.)

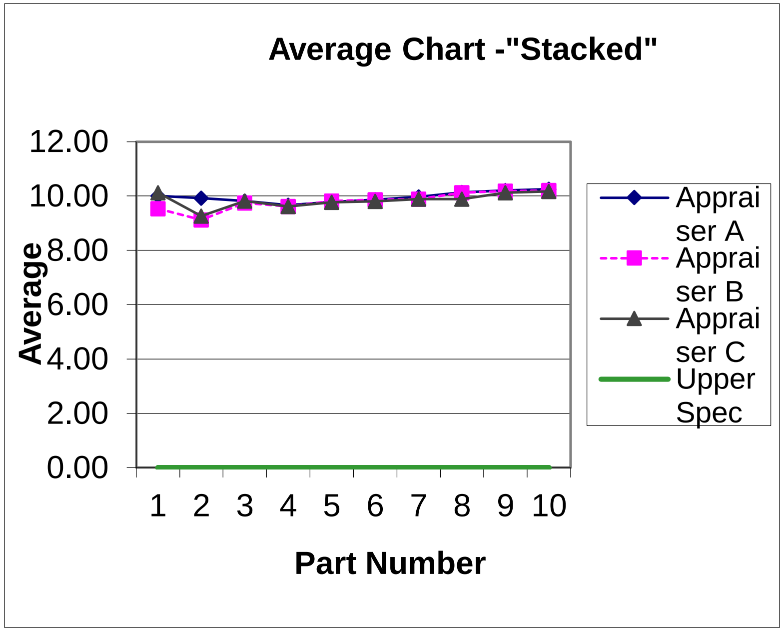
APPRAISER	TRIAL #	PART 1	2	3	4	5	6	7	8	9	10	AVERAGE	
1. A	1	10.00	9.90	9.80	9.70	9.80	9.90	10.00	10.10	10.20	10.30		9.970
2.	2	10.00	9.95	9.84	9.60	9.75	9.80	9.90	10.20	10.20	10.15		9.939
3.	3	10.00	9.90	9.80	9.70	9.80	9.90	10.00	10.10	10.20	10.30		9.970
4.	AVE	10.00	9.92	9.81	9.67	9.78	9.87	9.97	10.13	10.20	10.25	$\bar{X}_a$=	9.960
5.	R	0.00	0.05	0.04	0.10	0.05	0.10	0.10	0.10	0.00	0.15	$\bar{r}_a$=	0.069
6. B	1	9.90	9.10	9.70	9.65	9.76	9.80	9.95	10.05	10.18	10.20		9.829
7.	2	9.00	9.10	9.75	9.50	9.90	9.90	9.90	10.20	10.22	10.30		9.777
8.	3	9.70	9.15	9.80	9.70	9.80	9.90	9.80	10.10	10.15	10.10		9.820
9.	AVE	9.53	9.12	9.75	9.62	9.82	9.87	9.88	10.12	10.18	10.20	$\bar{X}_b$=	9.809
10.	R	0.90	0.05	0.10	0.20	0.14	0.10	0.15	0.15	0.07	0.20	$\bar{r}_b$=	0.206
11. C	1	10.20	9.20	9.90	9.60	9.70	9.80	9.95	9.95	10.10	10.20		9.860
12.	2	10.10	9.30	9.95	9.55	9.80	9.80	9.90	9.90	10.10	10.20		9.860
13.	3	10.05	9.30	9.60	9.70	9.80	9.80	9.80	9.80	10.15	10.10		9.810
14.	AVE	10.12	9.27	9.82	9.62	9.77	9.80	9.88	9.88	10.12	10.17	$\bar{X}_c$=	9.843
15.	R	0.15	0.10	0.35	0.15	0.10	0.00	0.15	0.15	0.05	0.10	$\bar{r}_c$=	0.130
16. PART												$\bar{\bar{X}}$=	9.871
AVERAGE		9.88	9.43	9.79	9.63	9.79	9.84	9.91	10.04	10.17	10.21	R_p=	0.772
17.	$(\bar{r}_a + \bar{r}_b + \bar{r}_c)$ / (# OF APPRAISERS) =											$\bar{\bar{R}}$=	0.135
18.	$\bar{X}_{DIFF}$ = (Max $\bar{X}$ - Min $\bar{X}$) =											$\bar{X}_{DIFF}$	0.151
19.	* $UCL_R = \bar{\bar{R}} \times D_4$ = APPRAISER					B		C	OUT OF CONTROL			UCL	0.348

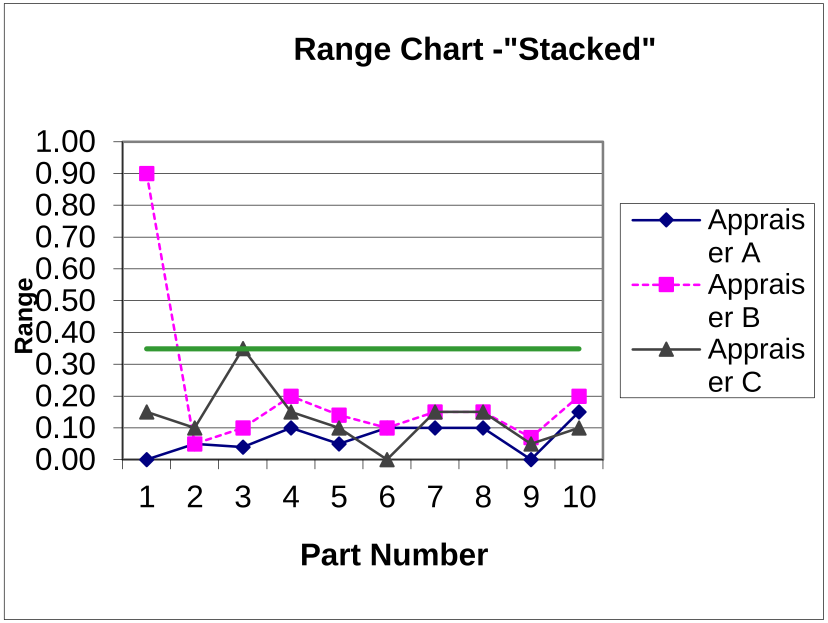
* D_4 = 3.27 for 2 trials and 2.58 for 3 trials. UCL_R represents the limit of individual R's. Circle those that are beyond this limit. Identify the cause and correct. Repeat these readings using the same appraiser and unit as originally used or discard values and re-average and recompute $\bar{\bar{R}}$ and the limiting value from the remaining observations.

Measurement Unit Analysis						% Tolerance (Tol)		
Repeatability - Equipment Variation (EV)								
EV	=	$\bar{\bar{R}} \times K_1$	Trials	K1		% EV	=	100 (EV/Tol)
	=	0.135 x 0.5907	2	0.8865			=	100(0.080/2.000)
	=	0.080	3	0.5907			=	3.99
Reproducibility - Appraiser Variation (AV)								
AV	=	$\{(\bar{X}_{DIFF} \times K_2)^2 - (EV^2/nr)\}^{1/2}$				% AV	=	100 (AV/Tol)
	=	{(0.151 x 0.5236)^2 - (0.080 ^2/(10 x 3))}^1/2					=	100(0.078/2.000)
	=	0.078					=	3.89
			Appraisers	2	3			
		n = parts r = trials	K_2	0.709	0.5236			
Repeatability & Reproducibility (GRR)						% GRR	=	100 (GRR/Tol)
GRR	=	$\{(EV^2 + AV^2)\}^{1/2}$	Parts	K_3			=	100(0.111/2.000)
	=	{(0.080^2 + 0.078^2)}^1/2	2	0.7087			=	5.57
	=	0.111	3	0.5236		*Gage system O.K*		
Part Variation (PV)			4	0.4464				
PV	=	$R_p \times K_3$	5	0.4032		% PV	=	100 (PV/Tol)
	=	0.772 x 0.3145	6	0.3745			=	100(0.243/2.000)
	=	0.243	7	0.3534			=	12.14
Tolerance (Tol)			8	0.3378				
Tol	=	Upper - Lower	9	0.3247		ndc	=	1.41(PV/GRR)
	=	2.00	10	0.3145			=	1.41(0.243/0.111)
	=	2.000					=	3
						Gage discrimination low		

For information on the theory and constants used in the form see *MSA Reference*.

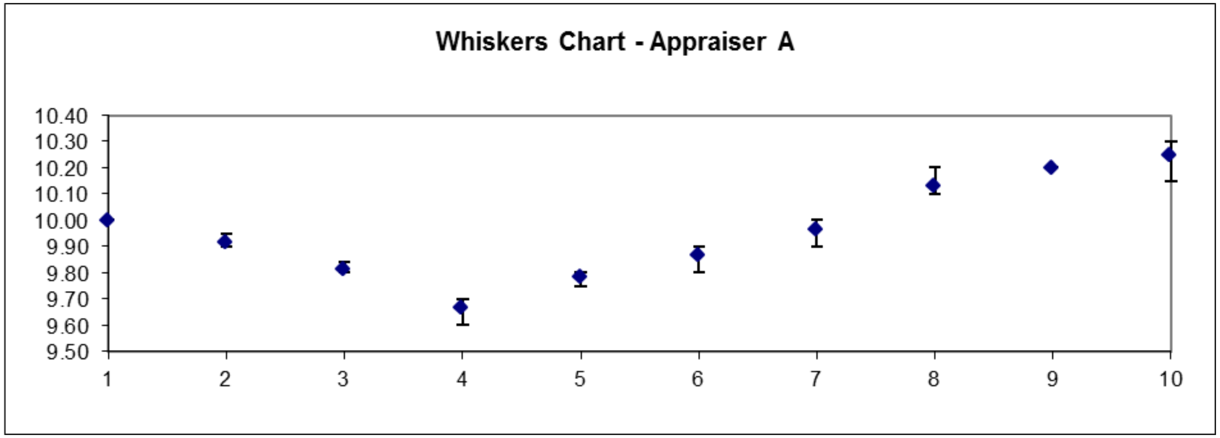
✓ The worksheet GRR graphical is provided in the GRR variable tool. A snapshot of the graph has also been provided here.

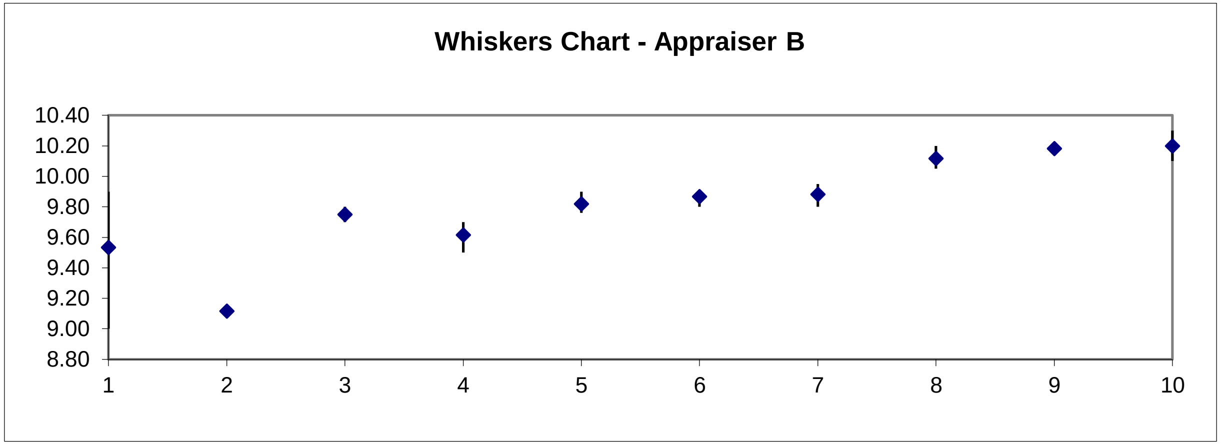




Measurement Systems Analysis Graphs (Contd.)

Measurement Systems Analysis Graphs (Contd.)





Assignment

- ✓Conduct a measurement systems analysis for 8 test parts. 2 Operators measure these 8 test parts and conduct 3 trials for each part. Use the GAGE RR – variable data graphical analysis and the GAGE RR blank worksheet to write your readings.
- ✓ For training purposes, dummy or assumed readings can be used.
- ✓ The specification limits provided are 8.5 and 12.5.
- ✓ Interpret the readings and discuss with facilitator.

Attribute RR

- ✓Measuring attribute data is fraught with plenty of variations as these types of data rely on human judgment.
- ✓Attribute data examples are good/bad, pass/fail, etc.
- ✓Measuring attribute data may result in some variation because tester's interpretation for operational definition might differ on what is good and what is bad.
- ✓A lot of industries rate their products and internal measurements, hence it is important to work on a technique that eliminates these measurement variations.
- ✓The technique used is known as attribute RR also known as attribute GAGE RR.

How to conduct an attribute RR study?

✓ **Step 1:**

- Select a set of 20 samples. This should be a mix of okay and defective samples. Ideally a 30-70 mix of good and defective is considered acceptable.

✓ **Step 2:**

- Let a process expert do the rating for each of the 20 samples. This process expert is called a master appraiser. Snapshot of the findings can be found in the next slide.

(Use the attribute RR excel worksheet provided in the toolkit to do attribute RR study.)

Attribute RR (Contd.)

Test Samples	Master Appraiser Rating
1	Okay
2	Okay
3	Defective
4	Defective
5	Okay
6	Defective
7	Defective
8	Defective
9	Defective
10	Okay
11	Okay
12	Okay
13	Defective
14	Defective
15	Okay
16	Defective
17	Defective
18	Defective
19	Defective
20	Okay

✓ **Step 3:** Select minimum two appraisers - Operator 1 (op1) and operator 2 (op2), and ask them to rate the product, not knowing the master appraiser's rating.

Test Samples	Master Appraiser Rating	Op1_Trial 1	Op1_ Trial 2	Op2_ Trial1	Op2_ Trial 2
1	Okay	Okay	Okay	Okay	Defective
2	Okay	Okay	Defective	Defective	Okay
3	Defective	Defective	Defective	Defective	Defective
4	Defective	Defective	Defective	Defective	Defective
5	Okay	Okay	Defective	Okay	Okay
6	Defective	Defective	Defective	Defective	Defective
7	Defective	Defective	Defective	Defective	Defective
8	Defective	Defective	Defective	Defective	Defective
9	Defective	Defective	Defective	Defective	Defective
10	Okay	Okay	Okay	Okay	Okay
11	Okay	Okay	Defective	Okay	Okay
12	Okay	Okay	Okay	Defective	Defective
13	Defective	Okay	Okay	Defective	Defective
14	Defective	Defective	Defective	Defective	Defective
15	Okay	Okay	Okay	Okay	Okay
16	Defective	Defective	Defective	Defective	Defective
17	Defective	Defective	Defective	Defective	Defective
18	Defective	Defective	Defective	Defective	Defective
19	Defective	Defective	Defective	Defective	Defective
20	Okay	Okay	Okay	Okay	Okay

✓ **Step 4:** Using the IF function in Excel check if the measurements/ratings of the individual operator match across both his trials, i.e., if one trial is scored okay then the other is scored okay or not.

Test Samples	Master Appraiser Rating	Op1_Trial 1	Op1_ Trial 2	Within Op1	Op2_ Trial1	Op2_ Trial 2
1	Okay	Okay	Okay	1	Okay	Defective
2	Okay	Okay	Defective	0	Defective	Okay
3	Defective	Defective	Defective	1	Defective	Defective
4	Defective	Defective	Defective	1	Defective	Defective
5	Okay	Okay	Defective	0	Okay	Okay
6	Defective	Defective	Defective	1	Defective	Defective
7	Defective	Defective	Defective	1	Defective	Defective
8	Defective	Defective	Defective	1	Defective	Defective
9	Defective	Defective	Defective	1	Defective	Defective
10	Okay	Okay	Okay	1	Okay	Okay
11	Okay	Okay	Defective	0	Okay	Okay
12	Okay	Okay	Okay	1	Defective	Defective
13	Defective	Okay	Okay	1	Defective	Defective
14	Defective	Defective	Defective	1	Defective	Defective
15	Okay	Okay	Okay	1	Okay	Okay
16	Defective	Defective	Defective	1	Defective	Defective
17	Defective	Defective	Defective	1	Defective	Defective
18	Defective	Defective	Defective	1	Defective	Defective
19	Defective	Defective	Defective	1	Defective	Defective
20	Okay	Okay	Okay	1	Okay	Okay
			# Matched	17		
			# Inspected	20		
			% Agree	85.00%		

The individual repeatability score of Op1 is 85%. This may not be acceptable to the company. Therefore, operator 1 needs retraining.

Attribute RR (Contd.)

- ✓ **Step 5**: Check if the individual measurements of the operators tally up to the standards set by the master appraiser.
- ✓ From the excel tool attribute RR snapshot, we can see that the individual effectiveness of the operator 1 is 80%.
- ✓ **This means only 80% of the time he could match up to the readings of the master appraiser, and he needs retraining.**

Attribute RR (Contd.)

Test Samples	Master Appraiser	Op1_Trial 1	Op1_ Trial 2	Within Op1	Op1 with standard	Op2_ Trial1	Op2_ Trial 2
1	Okay	Okay	Okay	1	1	Okay	Defective
2	Okay	Okay	Defective	0	0	Defective	Okay
3	Defective	Defective	Defective	1	1	Defective	Defective
4	Defective	Defective	Defective	1	1	Defective	Defective
5	Okay	Okay	Defective	0	0	Okay	Okay
6	Defective	Defective	Defective	1	1	Defective	Defective
7	Defective	Defective	Defective	1	1	Defective	Defective
8	Defective	Defective	Defective	1	1	Defective	Defective
9	Defective	Defective	Defective	1	1	Defective	Defective
10	Okay	Okay	Okay	1	1	Okay	Okay
11	Okay	Okay	Defective	0	0	Okay	Okay
12	Okay	Okay	Okay	1	1	Defective	Defective
13	Defective	Okay	Okay	1	0	Defective	Defective
14	Defective	Defective	Defective	1	1	Defective	Defective
15	Okay	Okay	Okay	1	1	Okay	Okay
16	Defective	Defective	Defective	1	1	Defective	Defective
17	Defective	Defective	Defective	1	1	Defective	Defective
18	Defective	Defective	Defective	1	1	Defective	Defective
19	Defective	Defective	Defective	1	1	Defective	Defective
20	Okay	Okay	Okay	1	1	Okay	Okay
			# Matched	17	16		
			# Inspected	20	20		
			% Agree	85.00%	80.00%		

Attribute RR (Contd.)

- Step 6: Calculate the 'between operators match', by finding the readings all the operators agree upon. Snapshot from toolkit as below:

Test Samples	Master Appraiser	Op1_Trial 1	Op1_ Trial 2	Within Op1	Op1 with standard	Op2_ Trial1	Op2_ Trial 2	Within Op2	Op2 with standard	Between Ops Match
1	Okay	Okay	Okay	1	1	Okay	Defective	0	0	0
2	Okay	Okay	Defective	0	0	Defective	Okay	0	0	1
3	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
4	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
5	Okay	Okay	Defective	0	0	Okay	Okay	1	1	0
6	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
7	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
8	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
9	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
10	Okay	Okay	Okay	1	1	Okay	Okay	1	1	1
11	Okay	Okay	Defective	0	0	Okay	Okay	1	1	0
12	Okay	Okay	Okay	1	1	Defective	Defective	1	0	0
13	Defective	Okay	Okay	1	0	Defective	Defective	1	1	0
14	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
15	Okay	Okay	Okay	1	1	Okay	Okay	1	1	1
16	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
17	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
18	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
19	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1
20	Okay	Okay	Okay	1	1	Okay	Okay	1	1	1
			# Matched	17	16			18	17	15
			# Inspected	20	20			20	20	20
			% Agree	85.00%	80.00%			90.00%	85.00%	75.00%

✓Step 7: Calculate overall effectiveness of the measurement system.

Test Samples	Master Appraiser Rating	Op1_Trial 1	Op1_ Trial 2	Within Op1	Op1 with standard	Op2_ Trial1	Op2_ Trial 2	Within Op2	Op2 with standard	Between Ops Match	All Ops vs Standard
1	Okay	Okay	Okay	1	1	Okay	Defective	0	0	0	0
2	Okay	Okay	Defective	0	0	Defective	Okay	0	0	1	0
3	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
4	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
5	Okay	Okay	Defective	0	0	Okay	Okay	1	1	0	0
6	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
7	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
8	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
9	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
10	Okay	Okay	Okay	1	1	Okay	Okay	1	1	1	1
11	Okay	Okay	Defective	0	0	Okay	Okay	1	1	0	0
12	Okay	Okay	Okay	1	1	Defective	Defective	1	0	0	0
13	Defective	Okay	Okay	1	0	Defective	Defective	1	1	0	0
14	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
15	Okay	Okay	Okay	1	1	Okay	Okay	1	1	1	1
16	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
17	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
18	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
19	Defective	Defective	Defective	1	1	Defective	Defective	1	1	1	1
20	Okay	Okay	Okay	1	1	Okay	Okay	1	1	1	1
			# Matched	17	16			18	17	15	14
			# Inspected	20	20			20	20	20	20
			% Agree	85.00%	80.00%			90.00%	85.00%	75.00%	70.00%

Attribute RR (Contd.)

Interpretations

- ✓ Operator 1 effectiveness is 85% and operator 1 effectiveness compared to the standard is 80%. This operator needs training with respect to standard procedures so that the effectiveness compared to standard can be taken beyond 90%.

- ✓ Operator 2 effectiveness is 90% and operator 2 effectiveness compared to the standard is 85%. This operator needs training but not to the extent of operator 1.

- ✓ Both the operators' measurement accuracy to the standards set is only 70%. Therefore, the overall measurement system is ineffective and needs improvement.

(Use the attribute RR tool provided in the toolkit to conduct an attribute RR study.)

When to Do Measurement System Analysis?

- ✓Measurement system analysis is the first activity in the measure stage, post conducting the first baseline study. Although the baseline study may still have some measurement system variations, the 1^{st} baseline is an unbiased estimate of what happened in the define phase.
- ✓So, conduct a baseline → Perform measurement system analysis → Repeat a baseline.

Data Collection Plan

- ✓To develop a written strategy for collecting the data for the project.
- ✓To define a clear strategy for collecting reliable data efficiently.
- ✓Focuses on the project's output (Y), its operational definition and performance characteristics.
- ✓Helps in ensuring that resources are used effectively to collect only the data that is critical to the success of the project.
- ✓Considers potential input variables (X) for the selected project output variables (Y).

Data Collection Plan Template and Example

Metric	Operational Definition	Data Source and Location	Sample Size	Who Will Collect the Data	Dates When Data Will be Collected	Other Data to be Collected at the Same Time
Cycle time for hiring process	Cycle time will be measured in days rounding up to nearest whole day. Start: date stamp when on-line application was received Stop: date stamp on email firm-offer letter. When letter was sent.	Start date will be collected from HR on-line web tool. Stop date will be collected from HR email offer letters.	280 new hires (all new hires from last 6 months)	XXX Assoc Recruiter Xxx Project Lead	Data will be collected from all applications dated from Jan 1, 2008-Jun 30, 2008. This data will be collected during the week of Oct 2, 2008.	• Recruiter • Assoc Recruiter • Hiring Mgr • Organization • Campus • Position type • Hard-to-fill (Y/N) • Day of Week application received • Was their compensation review (Y/N) • Was there screening interview (Y/N) • Application open & filed cycle time (days) • Interview cycle time (days) • Background check cycle time (days)

Summary

In this lesson, we have learned:

- ✓ Purpose of measurement system analysis
- ✓ Measurement system errors – accuracy and precision
- ✓ Linearity, resolution, bias, and stability
- ✓ Repeatability and reproducibility
- ✓ How to conduct a variables GAGE RR and attribute RR, and how to interpret them.
- ✓ The data collection plan on who will do the measurement, where, and how

Section III, Lesson 5

Stability Conditions

Agenda

- ✓Controlled Process and Variation
- ✓Special Causes of Variation
- ✓Common Causes of Variation
- ✓Stability Introduction and SPC
- ✓Stability Check with Minitab
- ✓Stability Conditions
- ✓Central Limit Theorem

Controlled Process and Variation

- By Walter Shewhart : "A phenomenon is said to be controlled when through the past experiences, we can predict within limits how the phenomenon can vary in the future."
- It is here that Walter Shewhart calls a phenomenon controlled.
- A controlled phenomenon doesn't mean absence of variation. It just means predictable variation.
- A controlled process would normally not deliver an out-of-specifications or a non-conforming product.

Special Causes of Variation

- ✓ Special causes are unusual, not previously observed, non-quantifiable or, unanticipated variation. These can bring an emergent change in the system or previously neglected inherent phenomena within the system.
- ✓ Variations are usually unpredictable and outside the historical values and range.
- ✓ Presence of special causes of variations in a process leads to instability and uncontrolled process.
  - Example: You normally take 35-40 minutes to reach office. However, it took 65 minutes to reach office and you are late to work today because of an accident on the road. This is a special cause of variation, something that does not happen everyday, unpredictable, and outside historical value.

Common Causes of Variation

- ✓Any unknown random cause of variations in the process is known as chance cause or random cause of variation. As these unknown random causes of variation keep happening regularly, they are referred to as common causes of variation.
- ✓Common-cause variation is the noise within the system. They are the usual, historical, quantifiable variation in a system.
- ✓If the effect of a common cause of variation is small and the presence of common cause effects in number is large, the variation becomes predictable and within limits.
- ✓If an process has only common cause of variation and no other variation, one can call the process as controlled.
  - Example: You normally take 35-40 minutes to reach office. The range 35 to 40 minutes is considered as common cause of variation. It takes more or less time due to traffic signals, variations in the number of cars on the road, etc. Here, the difference is small and within expected limits of variation.

Common Causes of Variation (Contd.)

- ✓In short, all special causes of variation need to be identified and eliminated; and common causes of variation should be minimized, if possible.
- ✓Does that mean we leave common causes of variation unattended?
  - The answer is No – As per Walter Shewhart, these variations are attended to in the process of long term improvement.
- ✓Most of the times distance between control limits indicate common causes of variation. Sometimes, special causes happen due to trends, oscillations, and shifts. All of these special cause variations are detected with the help of control charts.

Stability Introduction and SPC

- ✓ A process is said to be stable when all the causes of variations are known. All the special causes of variation are eliminated. And, the process is governed by common causes of variations to bring the process in statistical control. This results in the **process output to be fairly predictable and consistent**.

- ✓ A stable process will have no special causes of variation. Even if special causes do manage to exist, they will be desirable special causes of variation.

- ✓ As per SPC or statistical process control, every measurable phenomenon is a **statistical distribution**, i.e., every set of observed data is a set of effects of **unknown common causes.**

- ✓ Even after special causes of variation are eliminated, some variability will exist which is explained because of common causes of variation.

- ✓ Six Sigma Black Belts should seek to reduce common cause of variation. They may not be able to achieve complete elimination.

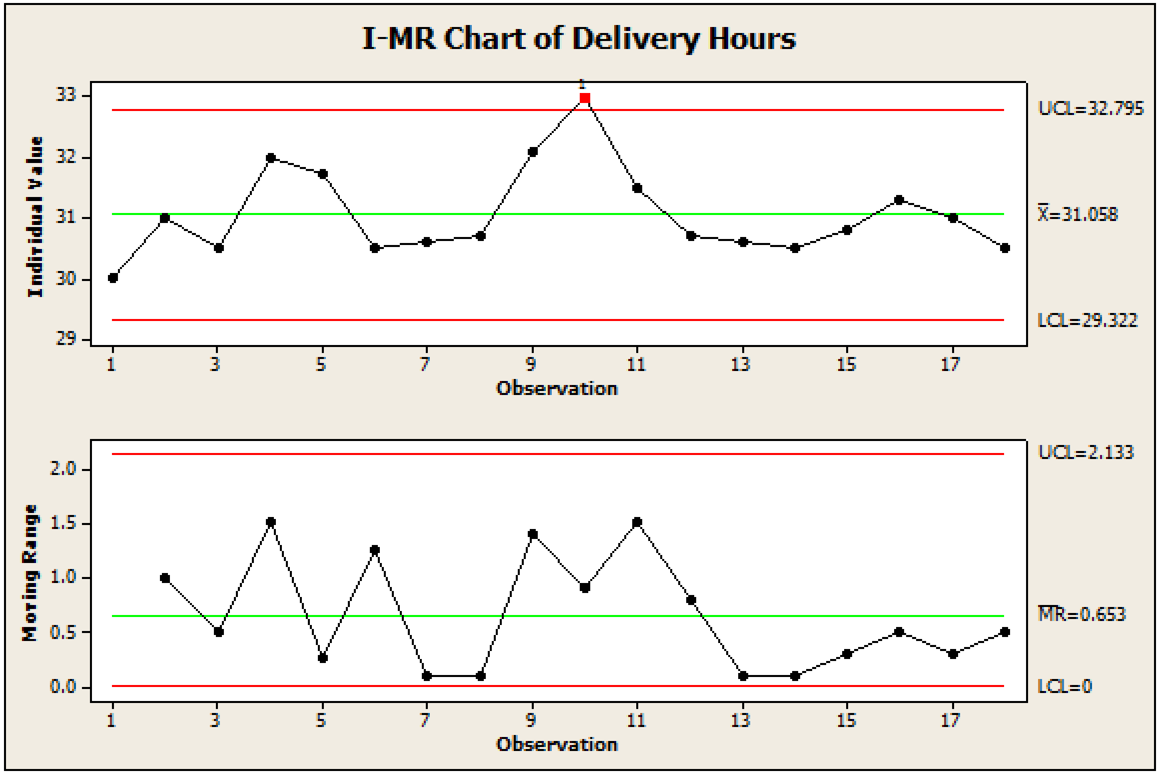
- ✓ In a stable process the location of the measurement distribution curve, spread, and shape will not change significantly over a period of time.

Data for checking stability

- ✓ A process measures delivery hours as the project Y. Measurements recorded day wise is shown below. Conduct a stability check and interpret.

Day	Delivery Hours	Day	Delivery Hours
1	30	10	33
2	31	11	31.5
3	30.5	12	30.7
4	32	13	30.6
5	31.75	14	30.5
6	30.5	15	30.8
7	30.6	16	31.3
8	30.7	17	31
9	32.1	18	30.5

- ✓For convenience sake, the I-MR chart in Minitab is used.
- ✓To select I-MR, Click on Stat → Control charts → Variables charts for individuals → I-MR



Please note: Reasoning for the use of I-MR chart will be revealed in section VI.

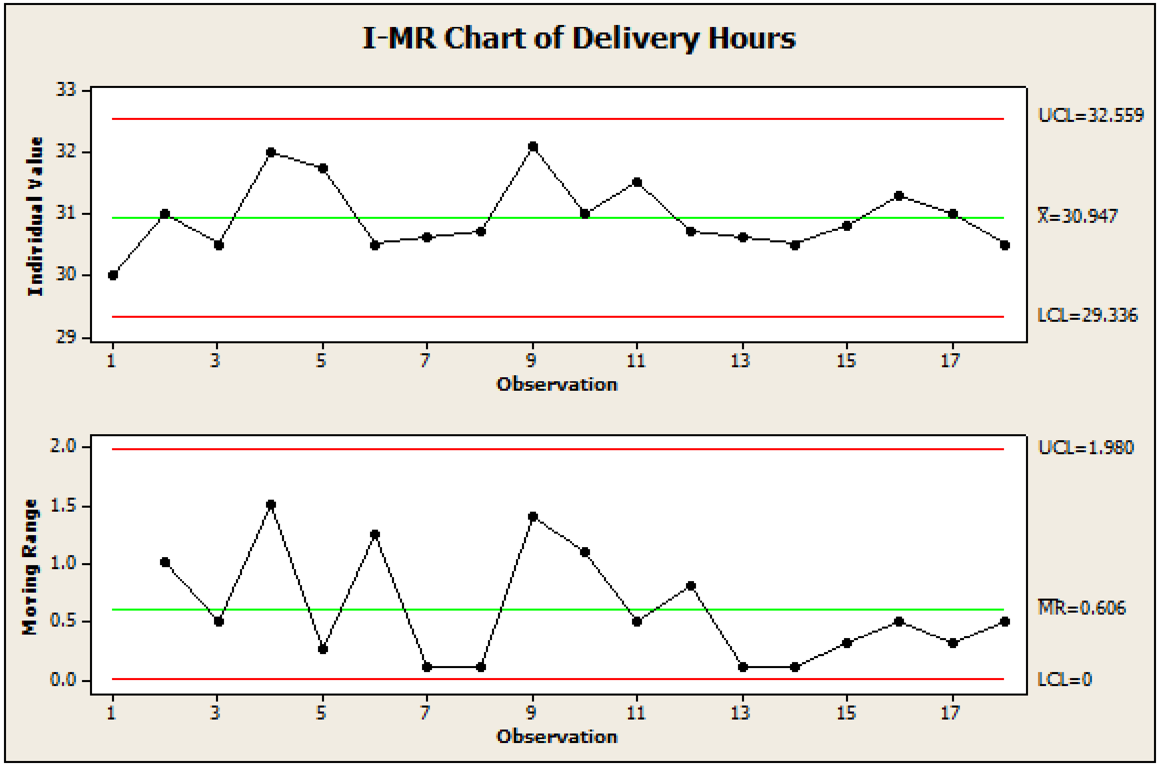
Stability Check with Minitab (Contd.)

- ✓Day 10 shows a point out of control. Clearly indicates the process being unstable with respect to its control limits.
- ✓The observation of 33 could clearly be due to special cause of variation (SCV), which would trigger off a series of actions to eliminate SCV.

Imp: On identification of SCV, an investigation needs to be triggered on what caused the variation and how to eliminate the SCV. The SCV needs to be eliminated and the process needs to be brought back in control.

Please note: - Details on how to eliminate special causes of variation will be discussed in section V.

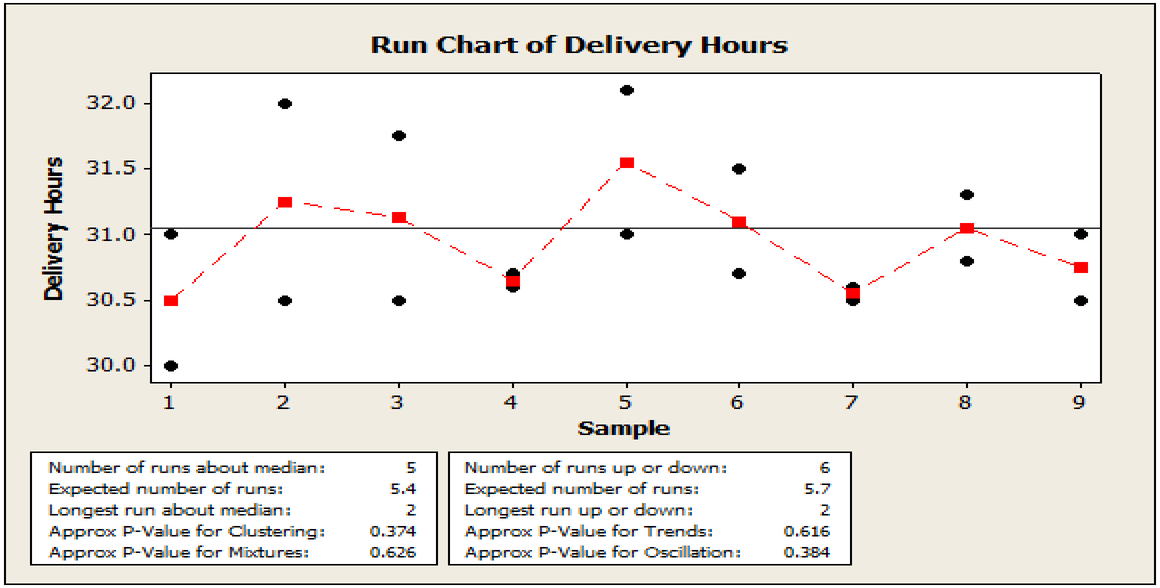
Stability Check with Minitab (Contd.)



Process is stable. We can proceed with our Measure phase activities.

Stability Check using Run Charts

✓In Minitab, Click on Stat → Quality tools → Run chart.



Check the 4 p-values. If any of the p-value is less than 0.05, the process is unstable due to special causes of variation.

Stability Conditions

- ✓The process having output (Y) as delivery hours. Check to see if the process is stable.
- ✓Let us consider the customer having given the target of 29 hours, with lower specification limit = 26 and upper specification limit = 32.
- ✓As we can see from the graph, the mean is working at 31 with a moving range of approximately 2 hours from the mean.
- ✓**Interpretation:** The process could be stable today, but the degree of confidence on stability for future is subjectively low.

Central Limit Theorem

- ✓Irrespective of the shape of the distribution of the population or universe, the distribution of average values of samples drawn from that universe approaches a normal distribution as the sample size increases.
- ✓The average of sample averages or mean of sample means equals the mean of the universe or population.
- ✓The standard deviation of the averages denoted by SEM or standard error of mean is represented by standard deviation of population divided by square root of sample size.
- ✓SEM = $\sigma /(n)^{1/2}$

Imp: The central limit theorem is the most powerful basis for Walter Shewhart's SPC charts or control charts.

Summary

In this lesson, we have learned:

- ✓Controlled process and variation
- ✓Special causes of variation
- ✓Common causes of variation
- ✓Stability introduction and SPC
- ✓Stability check with Minitab
- ✓Stability conditions
- ✓Central limit theorem

Section III, Lesson 6

Capability Metrics

Agenda

- ✓ Process Capability Pre-Considerations
- ✓ Process Capability Indices for Continuous Data
- ✓ Process Capability Indices for Discrete Data
- ✓ Non-Normal Capability Analysis

Process Capability Pre-Considerations

- ✓Calculating process capability (Cp/Cpk or DPMO/PPM) is not correct without determining if the process is in statistical control or not.

- ✓It is important the process data follows a normal distribution. Without checking for normality, it is futile to do a Cp/Cpk calculation, as Minitab allows a non-normal and between normal capability analysis.

- ✓**Example:**An improvement team uses Lean techniques to reduce the time to process an order. A random sample of 25 orders before the change had an average time of 31/2 hours to process. A random sample of 25 orders after the change has an average time of 2 hours to process. The team asserts they have decreased the order processing time by more than 40%.

Is that credible? Would it be more credible if the confidence interval for the order processing time was calculated for the original 25 orders, and the improvement asserted only if the new average of 2 hours fell outside the original interval?

Process Capability Pre-Considerations (Contd.)

- ✓The answer to the previous question is stability.
- ✓Every distribution has three properties that need to be determined: Location, spread, and shape. Location and spread are determined by stability, while shape is determined by doing a normality check.
- ✓Mean shows the location of the distribution and standard deviation is used to show dispersion or spread.
- ✓The following questions are to be asked before calculating process capability:
  1. Is the process central tendency stable over time or is it shifting?
  2. Is the process dispersion stable?
  3. Is the process distribution consistent over a period of time?

Process Capability Pre-Considerations (Contd.)

✓ If the answer is **NO** to any of the questions asked in the previous slide, we cannot and should not calculate capability of the process.

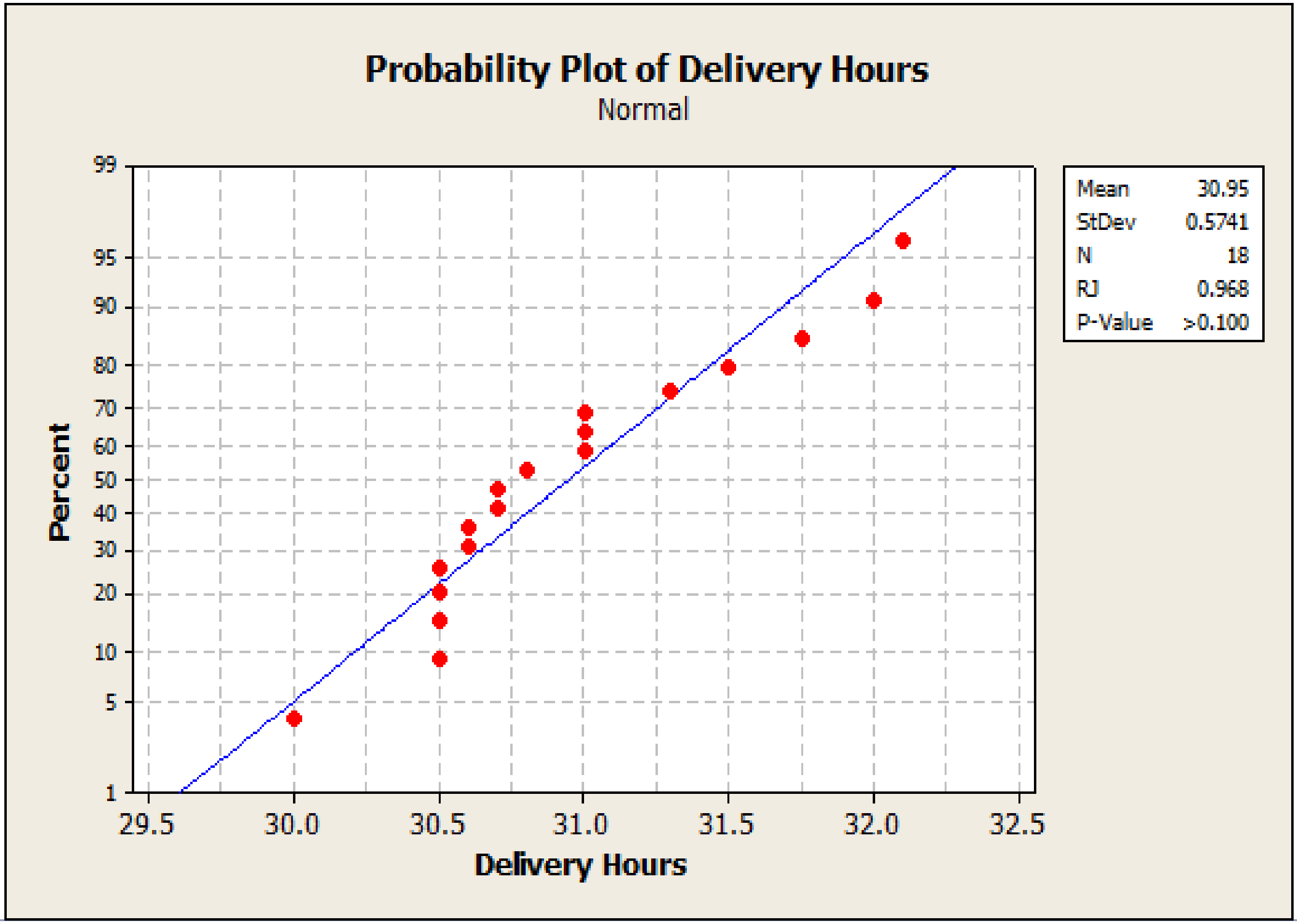
✓ If the answer is **YES** to all the three questions, we can proceed to ask questions on the capability of the process.

1. Is the process meeting requirements?
2. Is the process capable of meeting requirements?
3. Can re-centering the process help improve its performance?
4. How can we reduce variations in the process?

✓ We can ask these questions by calculating Cp, Cpk, Pp, and Ppk indices, but before that check if the data is normal.

Process Capability Pre-Considerations (Contd.)

✓ With the same data for stability post stability check being used, the Ryan Joiner test in Minitab is done. The normality test results shown here indicate that the data is normal and so, capability calculations can be done.

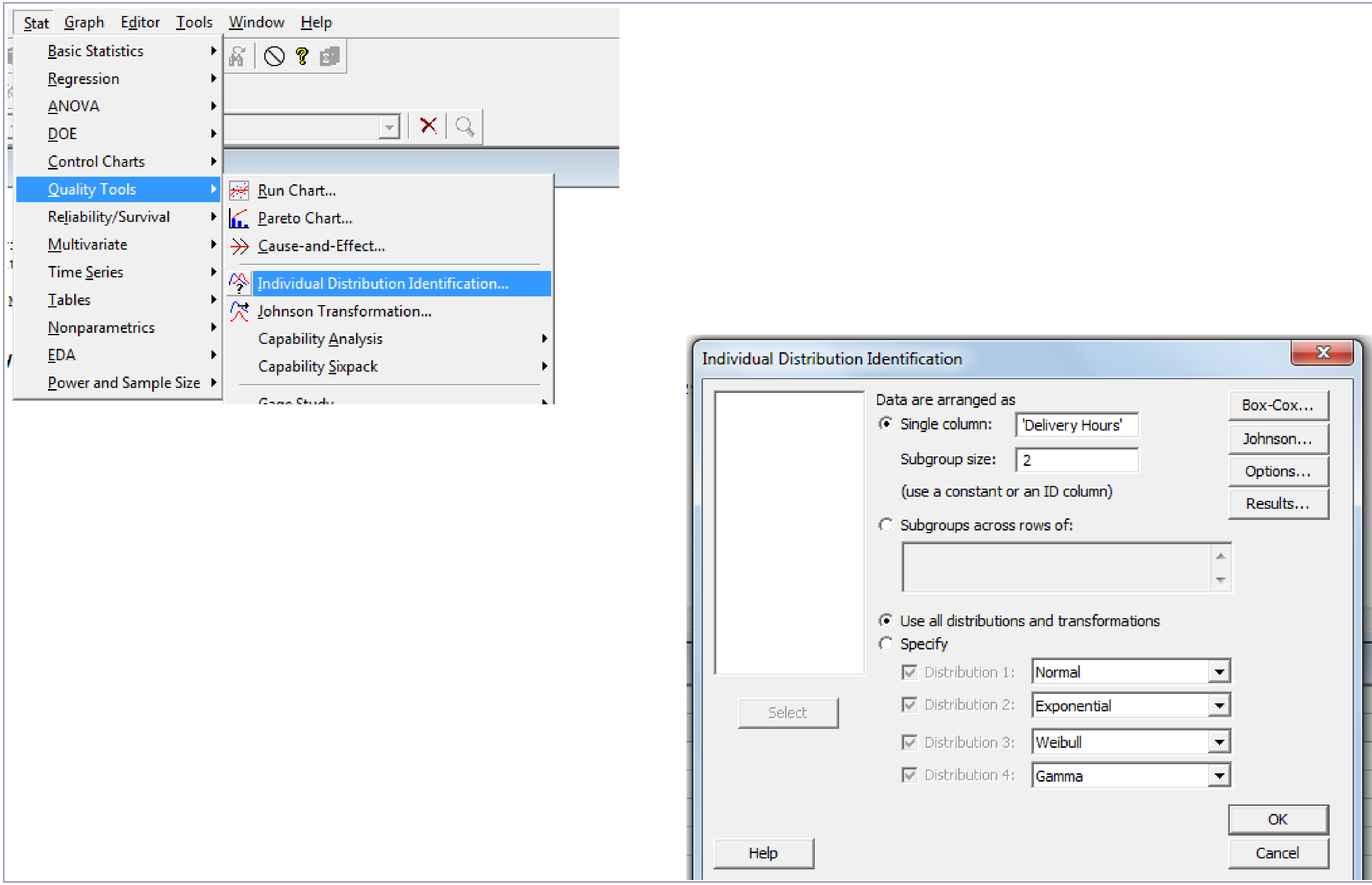


Process Capability Pre-Considerations (Contd.)

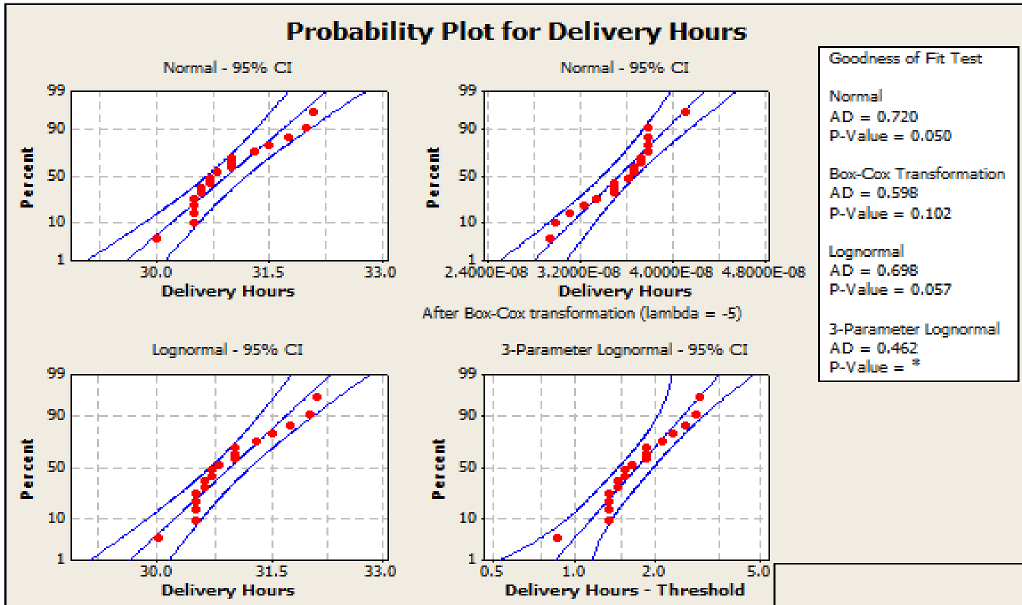
✓ With the same set of data, if normality failed, the below process should be followed:

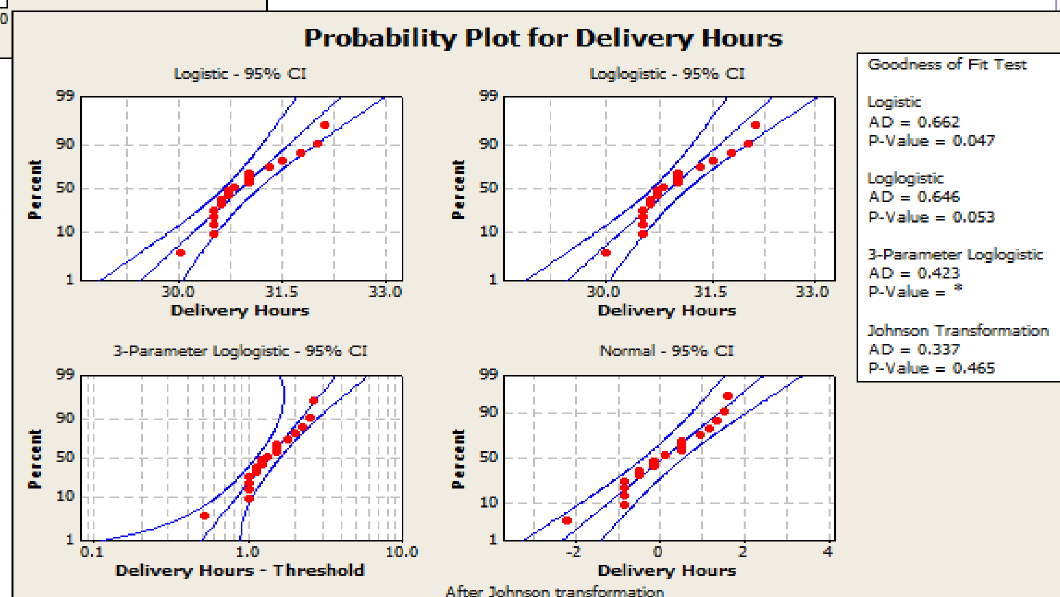
1. Click on Stat → Quality tools → Individual distribution identification → Select the variable and click Ok.
2. This tool will show the probability plots for all possible data distribution models.
3. The Anderson-Darling test for the set of data returns a p-value of 0.050. The Black Belt is not confident if normal distribution is the best fit so decides to go for a re-check.

Process Capability Pre-Considerations (Contd.)

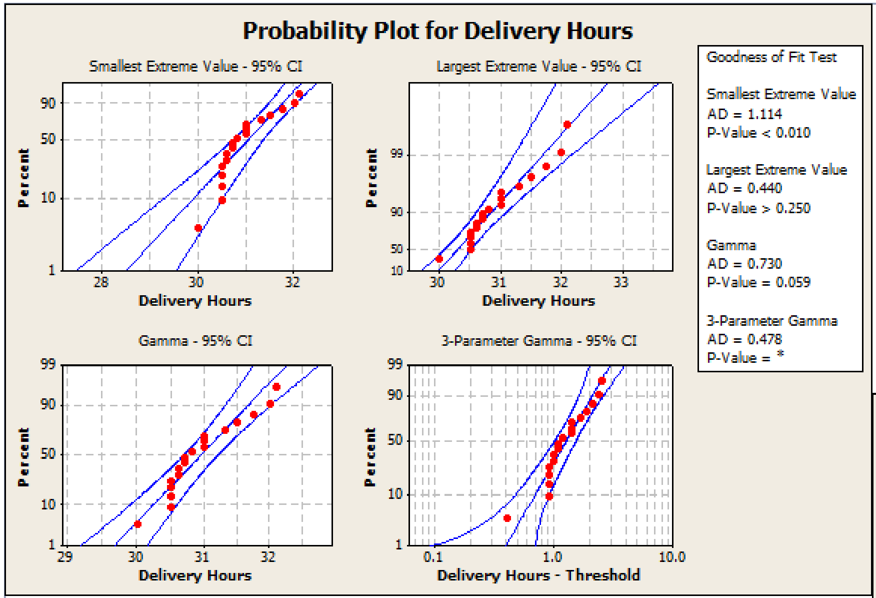


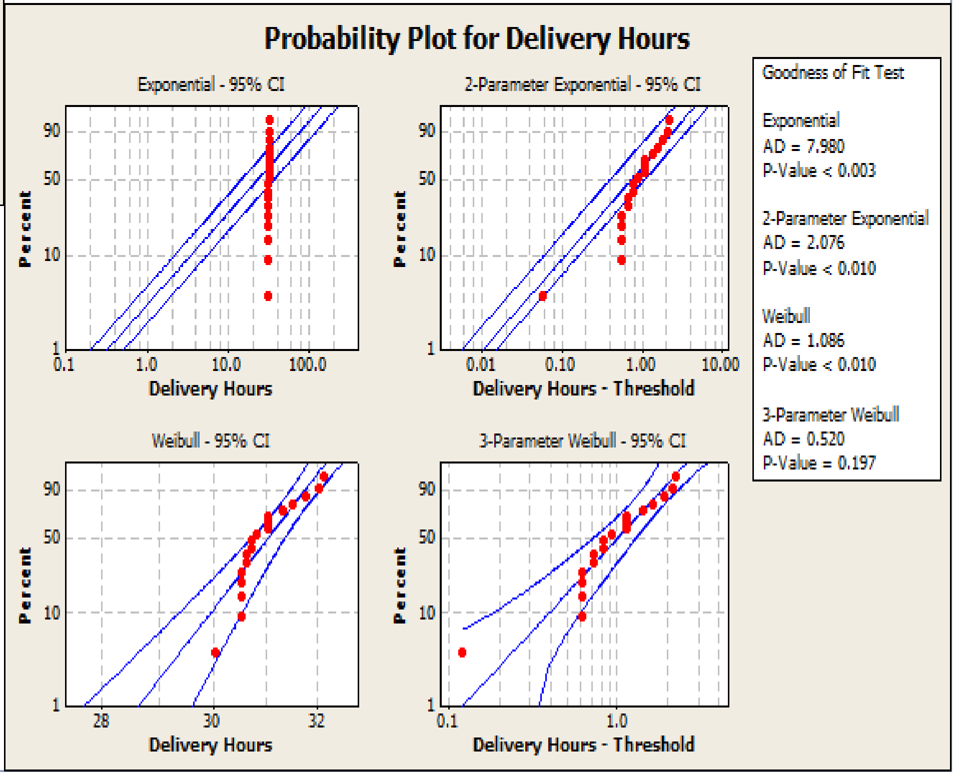
Process Capability Pre-Considerations (Contd.)





Process Capability Pre-Considerations (Contd.)





Process Capability Pre-Considerations (Contd.)

✓ Reviewing the p-values for all the distributions, the noteworthy ones are mentioned below in tabular format.

Weibull	1.086	<0.010	
3-Parameter Weibull	0.520	0.197	0.009
Smallest Extreme Value	1.114	<0.010	
Largest Extreme Value	0.440	>0.250	
Gamma	0.730	0.059	
3-Parameter Gamma	0.478	*	0.160
Logistic	0.662	0.047	
Loglogistic	0.646	0.053	
3-Parameter Loglogistic	0.423	*	0.143
Johnson Transformation	0.337	0.465	

Largest Extreme Value seems to be the best fit distribution for this data, and thus should be considered as the data distribution model. Some Black Belts may also go with 3-Parameter Weibull for simplicity considerations.

Process Capability Indices for Continuous Data

- ✓Cp = Process capability → Shows how good the process is in delivering what the customer wants with respect to specifications.
- ✓Cpk = Process capability index → Shows how good the process is in delivering what the customer wants with respect to mean.
- ✓Pp = Process potential → Shows how good the process can perform in the long run with respect to specifications.
- ✓Ppk = Process performance index → Shows how good the process is in delivering what the customer wants in terms of performance to mean in the long run.

✓Cp = (USL-LSL)/(6 * σ), with USL - LSL representing the process tolerance.

- USL – Upper specification limit
- LSL – Lower specification limit

Note that Cp doesn't work for a single specification limit. The customer should give both the specification limits for the value of Cp to be calculated.

✓Cpk = min (Cpl, Cpu)

- Cpl = (Xbar – LSL)/(3 * σ)
- Cpu = (USL – Xbar)/(3 * σ)

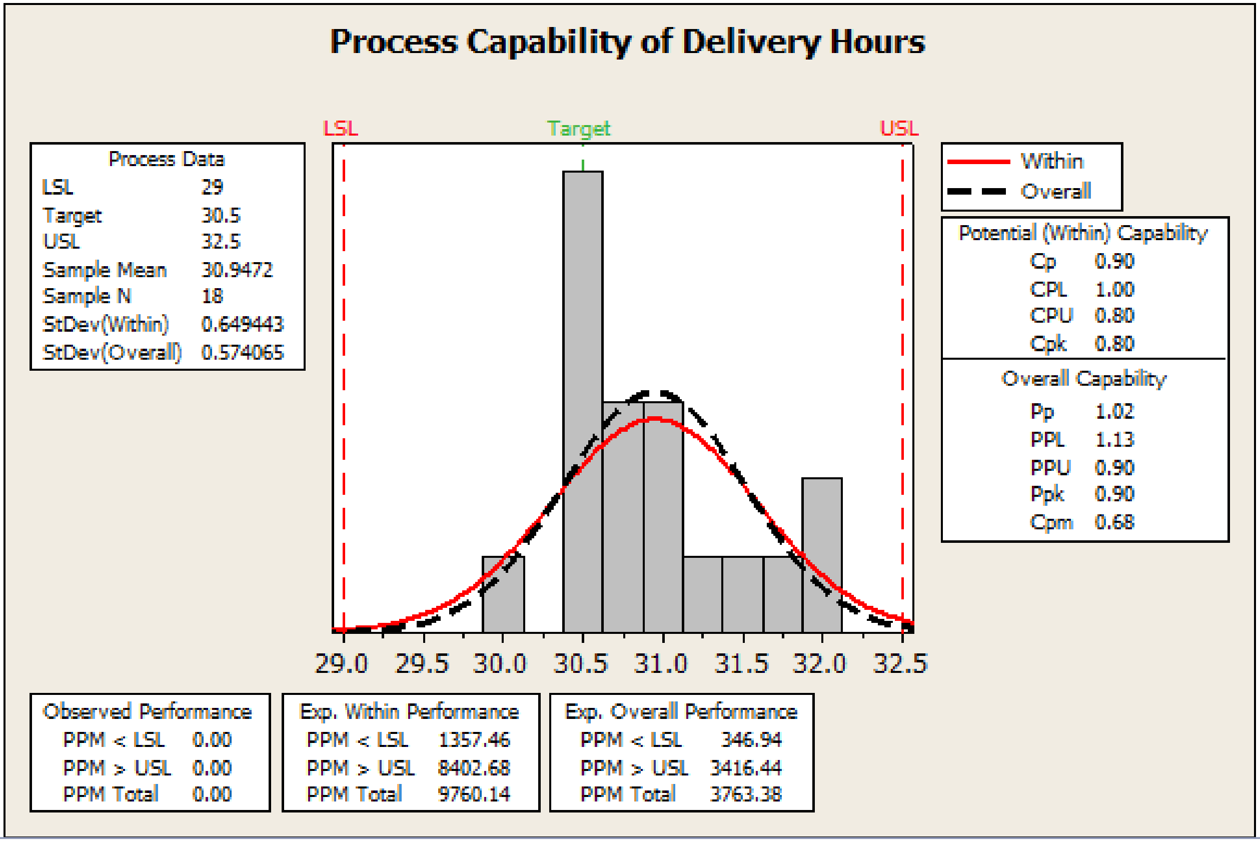
✓If Cpk < Cp, mean is not centered.

✓If Cpk = Cp, mean is centered and process is accurate, but one needs to know how much is the variation.

Process Capability Indices for Continuous Data (Contd.)

✓Using Minitab, let us calculate the process capability indices for the data we have for the Project Y = Delivery hours.

✓Click on Stat → Click on Quality Tools → Click on Capability Analysis → Click on Normal.



Process Capability Indices Interpretation

- ✓The LSL was defined at 29, and USL at 32.5.
- ✓The Cpk value of 0.80 is less than Cp value of 0.90, indicating mean is not centered yet, the Cp value indicates that the variations are high resulting in capability being low.
- ✓The same data was checked against control charts and it was found to be stable. It is due to high common cause of variation that we find the Cp value being so low.
- ✓Centering the mean in this case wouldn't do much of good. This is a clear case of reducing variation.

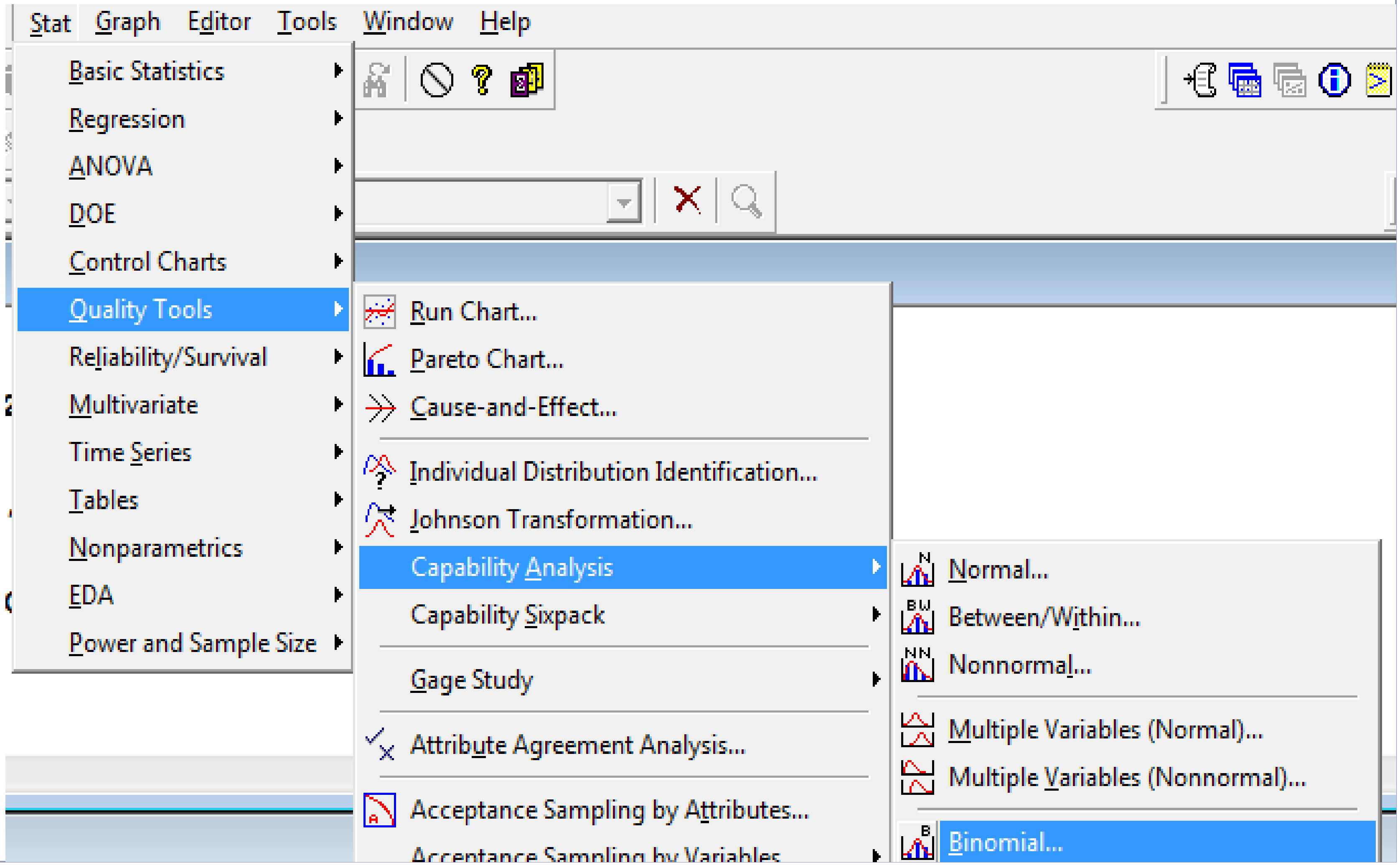
Process Capability for Discrete Data

- ✓Cp, Cpk, Pp and Ppk cannot be used for discrete data. Discrete data is countable data and in a sense, either count defects or defectives.
- ✓Use DPMO or PPM calculations in finding process capability when project Y is discrete.
- ✓**Example:** On a measure of 20 days, the number of defective items reported in a company are as mentioned below. It has been checked that the defective items are stable in the way they come out. Calculate process capability.

20,21,25,22,23,24,25,22,23,20,21,20,19,20,21,22,23,23,24,20

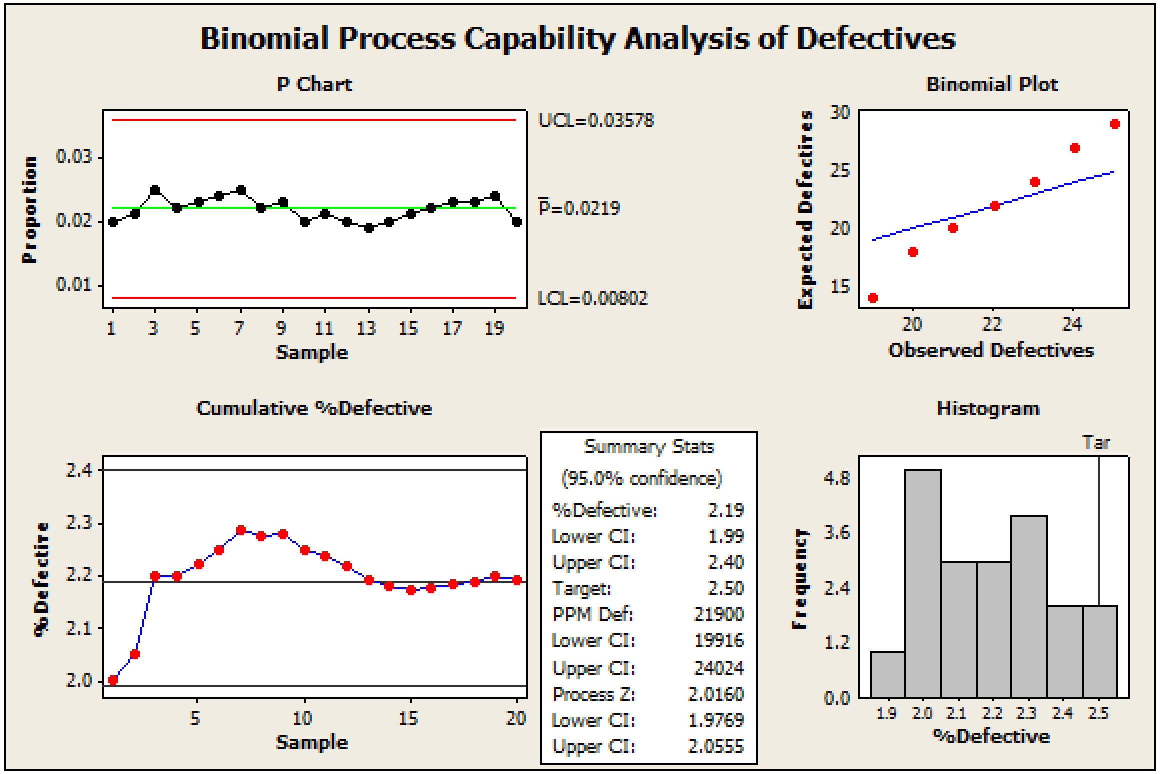
Process Capability for Discrete Data (Contd.)

✓ Use Minitab and Click on Stat → Quality Tools → Capability Analysis → Binomial.



Process Capability for Discrete Data (Contd.)

✓Process capability analysis of binomial data (Discrete)



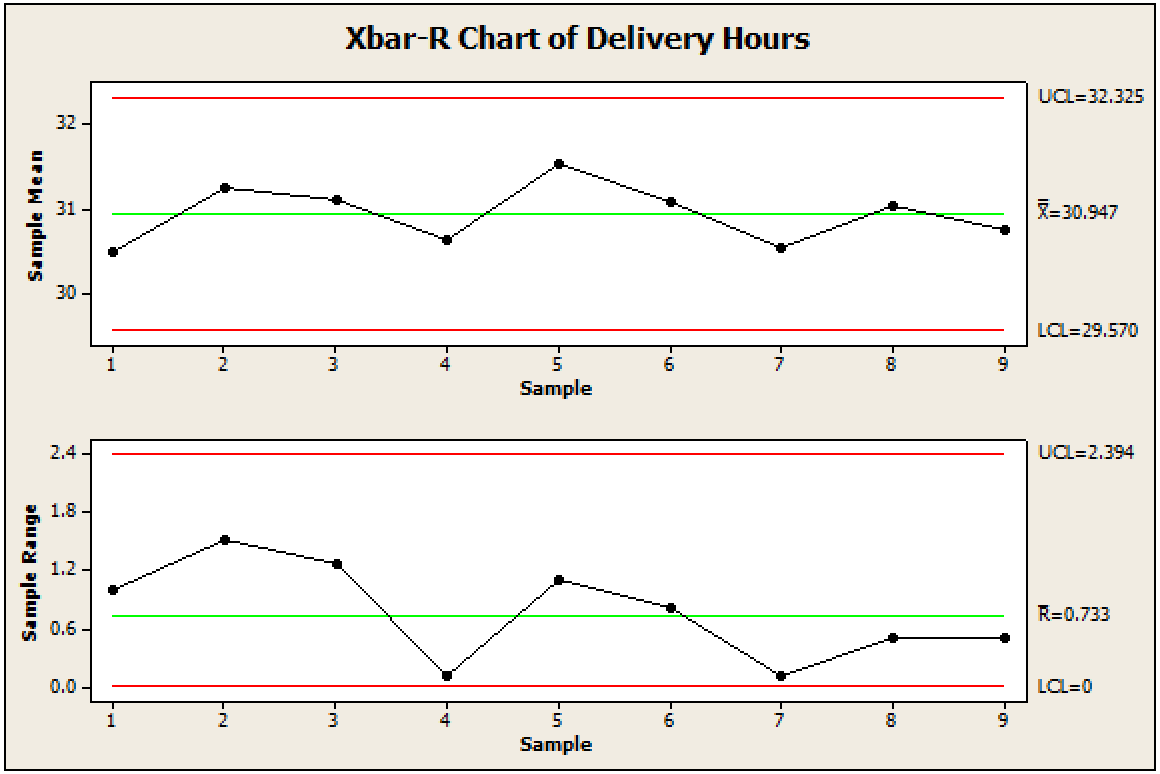
1) The % defective show what percent of parts are defective.
2) Take the PPM number from here, and place it in the DPMO-Sig conversion tool to get process sigma level.

Non-Normal Capability Analysis

- ✓Normality along with stability is one of the important pre-requisites to do a capability analysis on the data.
- ✓The delivery hours data collected as part of project Y is first determined if it is in statistical control. It is then checked for normality and finally, for capability.
- ✓A process in statistical control will produce a normal distribution most of the times; therefore ,determining stability first is important.

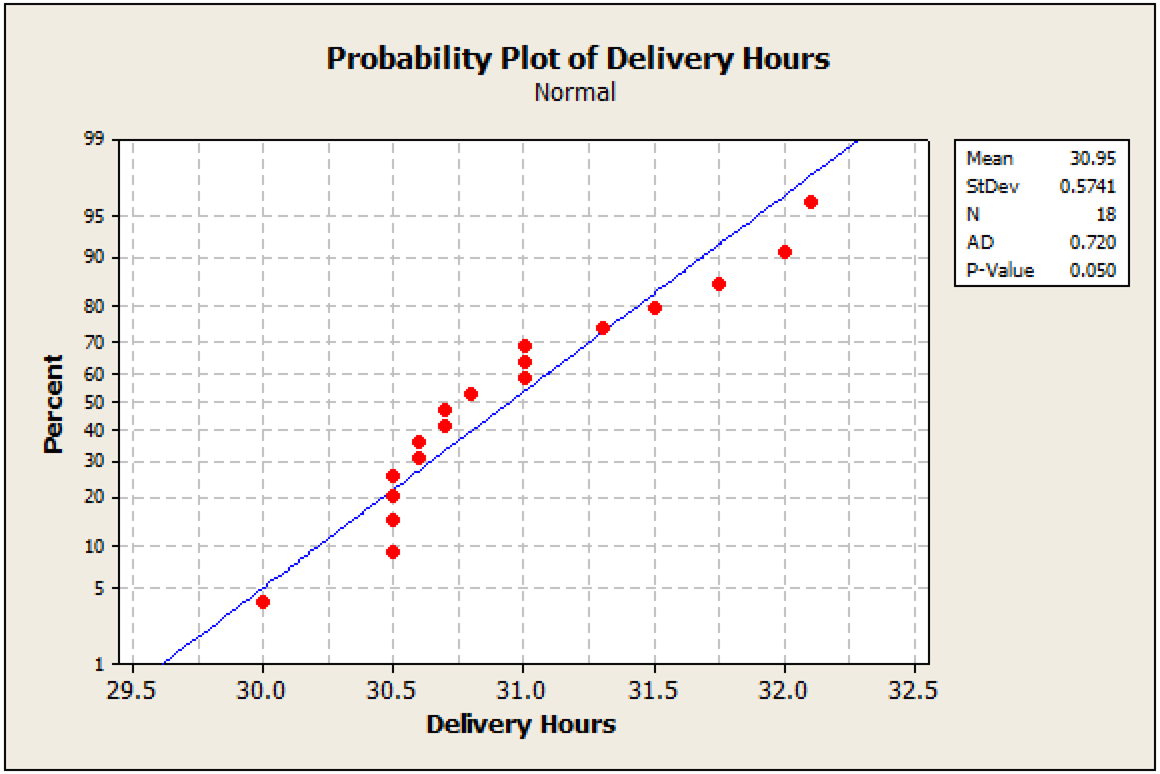
Non-Normal Capability Analysis (Contd.)

✓Stability check results: Process is stable



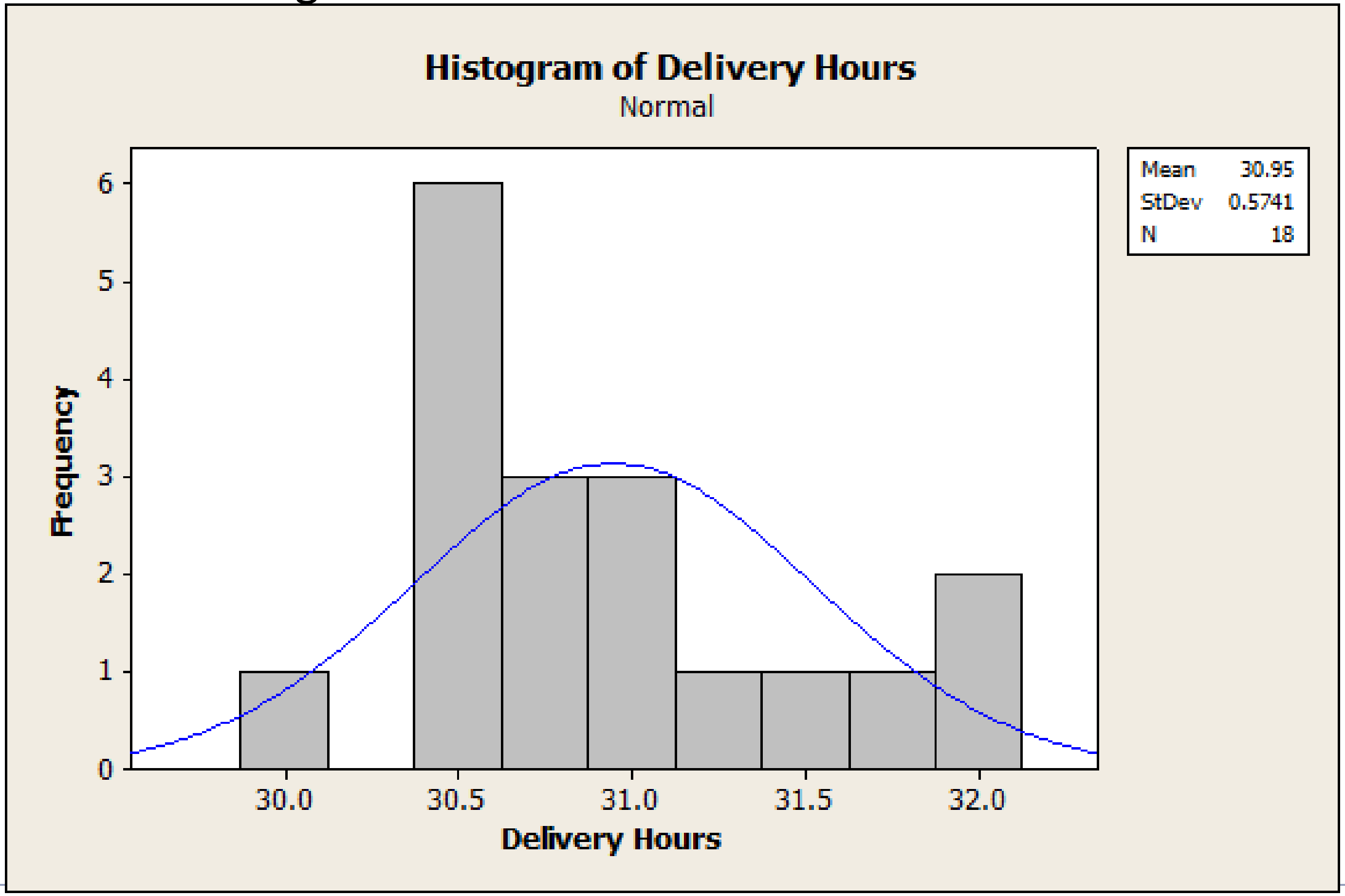
Non-Normal Capability Analysis (Contd.)

✓ Normality check results



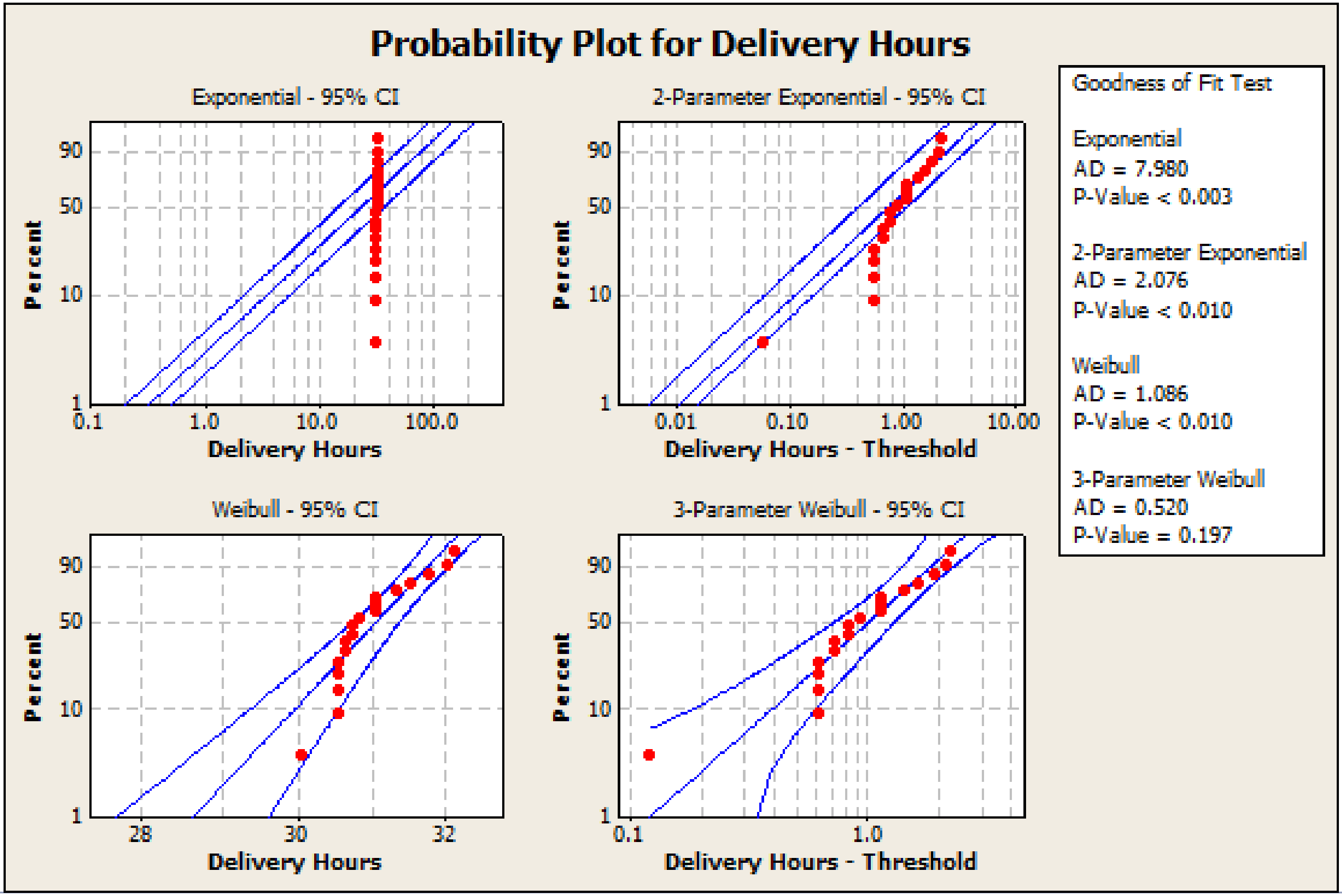
Non-Normal Capability Analysis (Contd.)

- ✓The problem is both of these (normality check and stability check graphs) reveal little on how the data fits a particular model. We use a histogram to study the fit. The histogram shown here with a normal fit shows normal distribution is not a good fit.



Non-Normal Capability Analysis (Contd.)

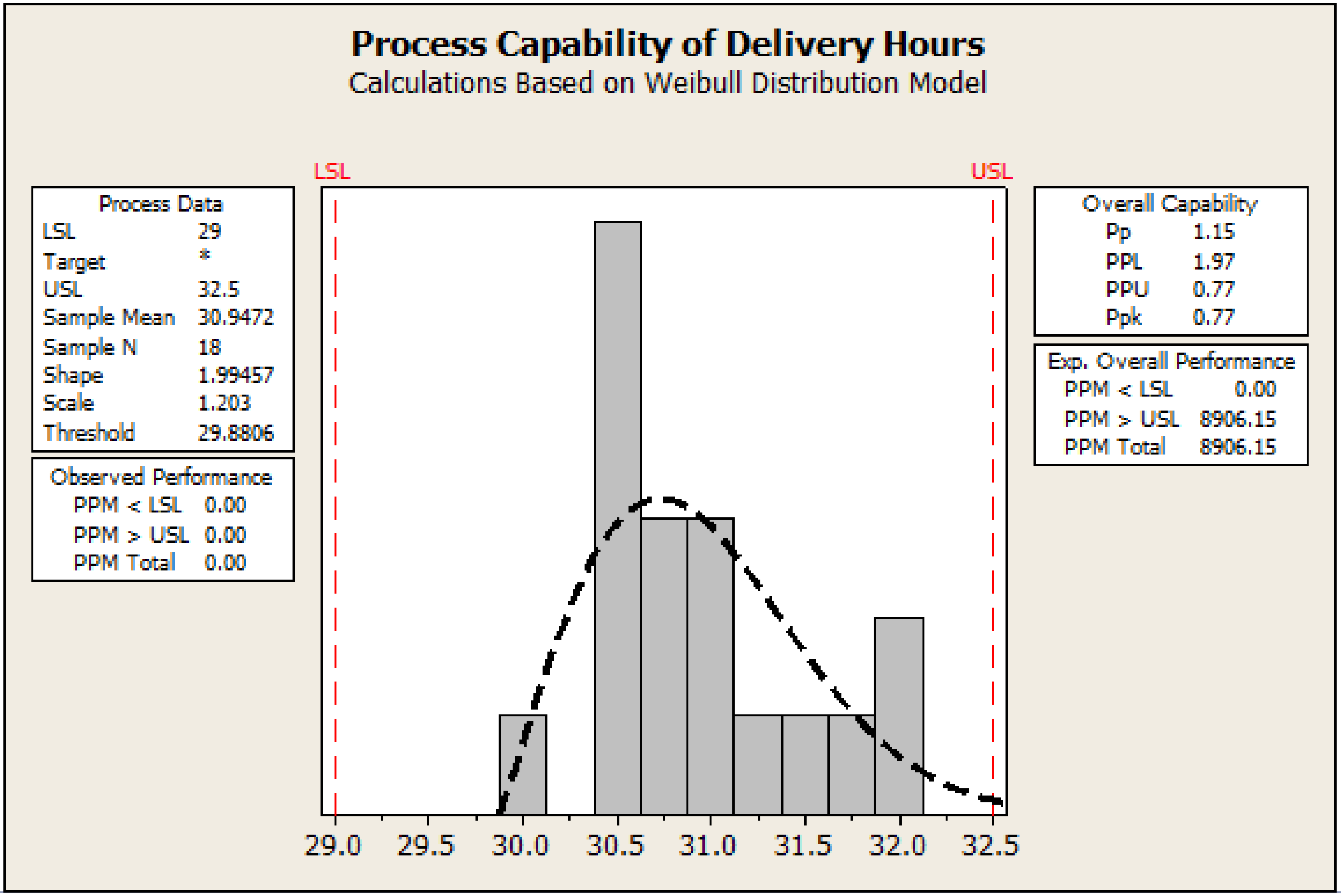
✓With the individual distribution identification function in Minitab results below, we conclude this data belongs to the 3-parameter Weibull function.



✓ Finally, using the non-normal capability analysis function in Minitab and the graph below, we can summarize capability as shown here:

Ppk = 0.77

PPM = 8906



Assignment

1. Plot your Project Y Data. Hypothetical value are okay. Assume a LSL and USL for continuous data and a target for discrete data.
2. Find if data is stable.
3. Find if data is normal.
4. Calculate capability indices appropriately and interpret.

Summary

✓In this lesson, we have learned:

- Process capability pre-considerations
- Process capability for continuous data
- Process capability for discrete data
- Non-normal capability analysis

Section III, Lesson 7

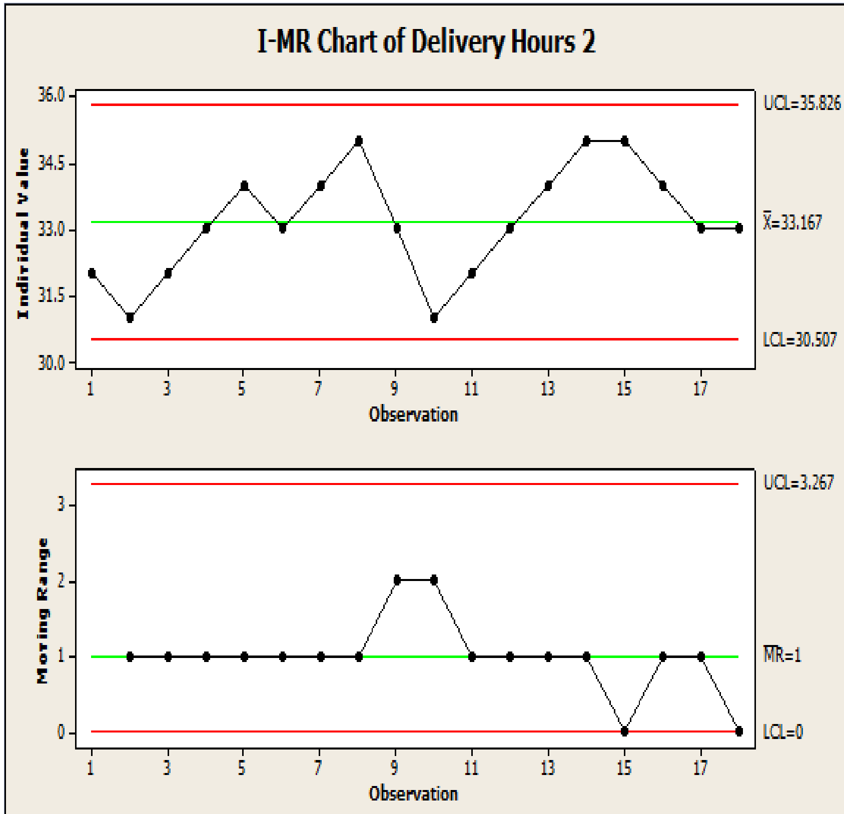
Variations, Variability, Capability, and Process Conditions

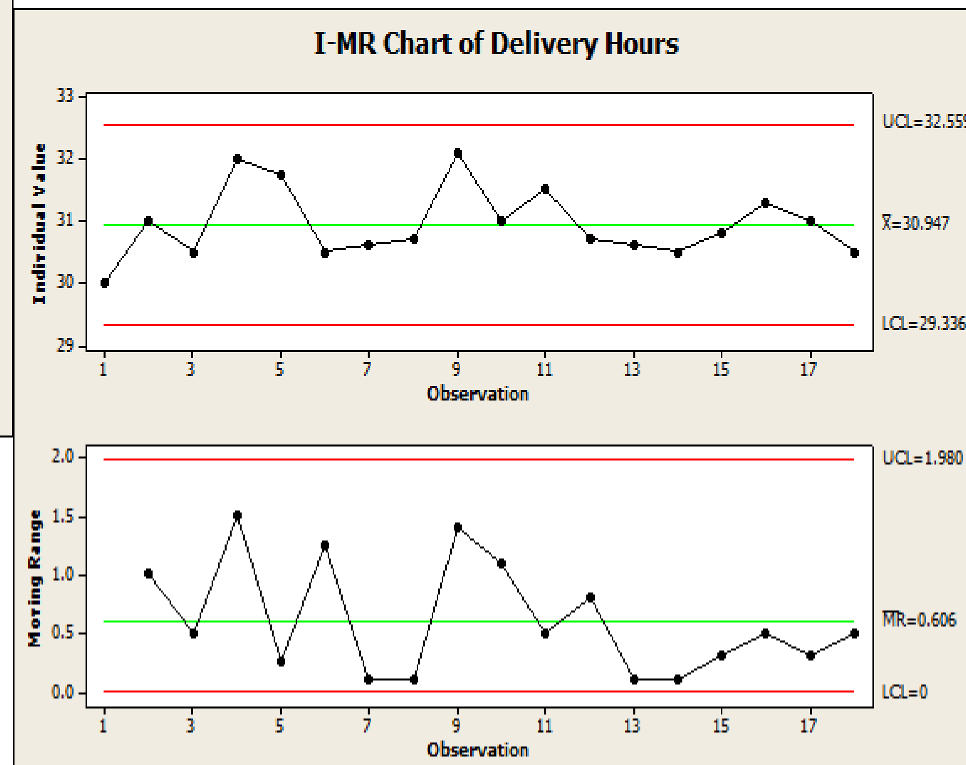
Agenda

- ✓ Variations and Variability
- ✓ Capability and Process Condition

Variations and Variability

- ✓Special cause is something that can occur anytime and has no pattern.
- ✓Special causes of variation, if any, need to be identified and eliminated before we could comment on variability, capability, and other process conditions.
- ✓Special causes of variation result in the process to go unstable, and hence need to be arrested, if undesirable for the process.
- ✓Common cause of variation is something that is inherent in the process and may follow some pattern.
- ✓Common causes of variation need to be controlled. Excessive effects of CCV could result in the process spread being high, resulting in overall capability being low.


I-MR Chart of Delivery Hours 2
Individual Value
36.0
34.5
33.0
31.5
30.0
UCL=35.826
X̄=33.167
LCL=30.507
1
3
5
7
9
11
13
15
17
Observation
Moving Range
3
2
1
0
UCL=3.267
MR=1
LCL=0
Observation


I-MR Chart of Delivery Hours
Individual Value
33
32
31
30
29
UCL=32.559
X̄=30.947
LCL=29.336
1
3
5
7
9
11
13
15
17
Observation
Moving Range
2.0
1.5
1.0
0.5
0.0
UCL=1.980
MR=0.606
LCL=0
Observation

Variations and Variability (Contd.)

- ✓ The comparison of the two charts on different situations with project Y of deliverable hours clearly shows the absence of SCV. The common causes of variation is lesser in delivery hours as compared to delivery hours 2.
- ✓ The control limits of delivery hours are 32.53 and 29.34, with the moving range being in control.
- ✓ If this was a time-series generated data, reduction in variation could have been attributed from delivery hours 2 to delivery hours, resulting in improved capability.

Capability and Process Conditions

- ✓ Assuming special causes of variation are eliminated,
  1) If variations are high, variability would be high, capability would be low and process would be incapable, i.e Cp or Cpk would be less than 1.
  2) If variations are low, variability would be low, capability would be high and process would be capable, i.e. Cp or Cpk would be more than 1.

- ✓ If process is found capable, a decision needs to be taken whether reduction in variation needs to be pursued or not. If only mean re-centering needs to be done by making fundamental changes to the process, classical DMAIC (Define, Measure, Analyze, Improve, Control) need not be pursued.

Summary

In this lesson we have learned:

✓Variations and variability

✓Capability and process condition

Section III, Lesson 8

Data Distributions

Agenda

- ✓ Permutations and Combinations
- ✓ Frequency Distributions
- ✓ Binomial Distribution
- ✓ Poisson Distribution
- ✓ Normal Distribution
- ✓ Exponential Distribution

- ✓Ordered arrangement of elements is known as permutation.
- ✓Example: You have 5 boxes and 5 objects to place in each box in such a way that every box contains only one object. How many total arrangements can you work on such a setting?
- ✓Solution: Object 1 can choose from any of the 5 boxes and be placed. Object 2 can only be placed in 4 boxes now. Object 3 in 3 and so on.
- ✓Thus, the total number of arrangements is 5*4*3*2*1 = 120
- ✓Generally for n positions to be filled with n objects, the equation is:
  - n * (n-1) * (n-2)..... This is also known as n!, read as n factorial.

- ✓Combination is about arranging n objects by taking groups of r objects at a time.
- ✓For example, if you have three items X, Y, and Z in groups of 2 items at a time, what could be the effective combinations of all three items?
- ✓The formula is ^nC_r = n!/(r!)*(n-r!). ^nC_r is also referred to as n choose r.
- ✓For this case, $^3C_2$ = 3!/(2!)(1!) = (3*2)/(2*1)(1) = 6/2 = 3.
- ✓In other words, there are three combinations here and they are XY, YZ, and ZX.

Frequency and Cumulative Distributions

- Frequency distribution is an empirical set of observations presented. Commonly used frequency distribution tools are histogram and stem and leaf plots.

- Frequency distributions show frequency of specific values or a group of values.

- Cumulative distributions show the cumulative value of frequency for that particular item or a group of items.

- If we can assume a state of SPC, we can answer the below questions:
  1. Probability that x will occur;
  2. Probability that a value less than x will occur;
  3. Probability that a value greater than x will occur; and
  4. Probability that a value between x and y will occur.

Binomial Distribution

- ✓ Binomial distribution helps in studying the probability of having x ***defectives*** in a sample of n units. In simple words, this distribution is used to study the probability of the occurrence of a defective item in a lot.
- ✓ $P(x) = {}^{n}C_{x} * p^{x} * (1-p)^{(n-x)}$ By standard, this is known as binomial probability distribution, where P(x) stands for probability of having x defectives in a sample.
- ✓ We will work with binomial distributions in three modes for deeper understanding:
  1. Manual calculations;
  2. Microsoft Excel; and
  3. Minitab.

- ✓ **Example:** A process produces bottles on a continuous basis. Past process performance indicates 2% chance of the bottles having one or more flaws. If we draw a sample of 20 units from the lot, what is the probability we will have 0 defectives?

Solution

- ✓ With the above information presented, n = 20, x = 0, p = 0.02
- ✓ Using the standard binomial distribution equation, ${}^{n}C_{x} * p^{x} * (1-p)^{(n-x)}$
  - ${}^{n}C_{x} = {}^{20}C_{0} = 20!/(0! * 20!) = 1$
  - $P^{x} = 1$, $(1-p)^{(n-x)} = 0.66$

- ✓ Multiplying all of them, the result is 66%. This means there is a 66% probability of accepting a process that on an average produces 2% defectives, to produce 0 defectives in a sample of 20 items.

In Microsoft Excel:

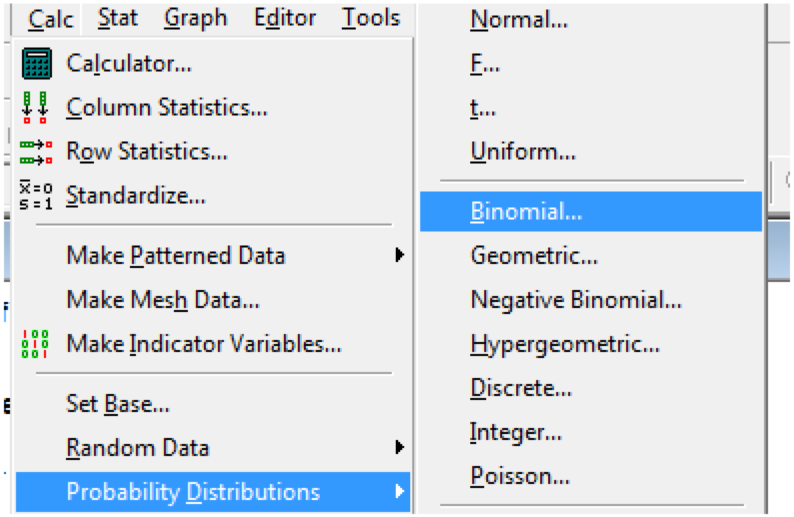
n	20
p	0.02
x	0
Probability	=binomdist(B3,B1,B2,"False")

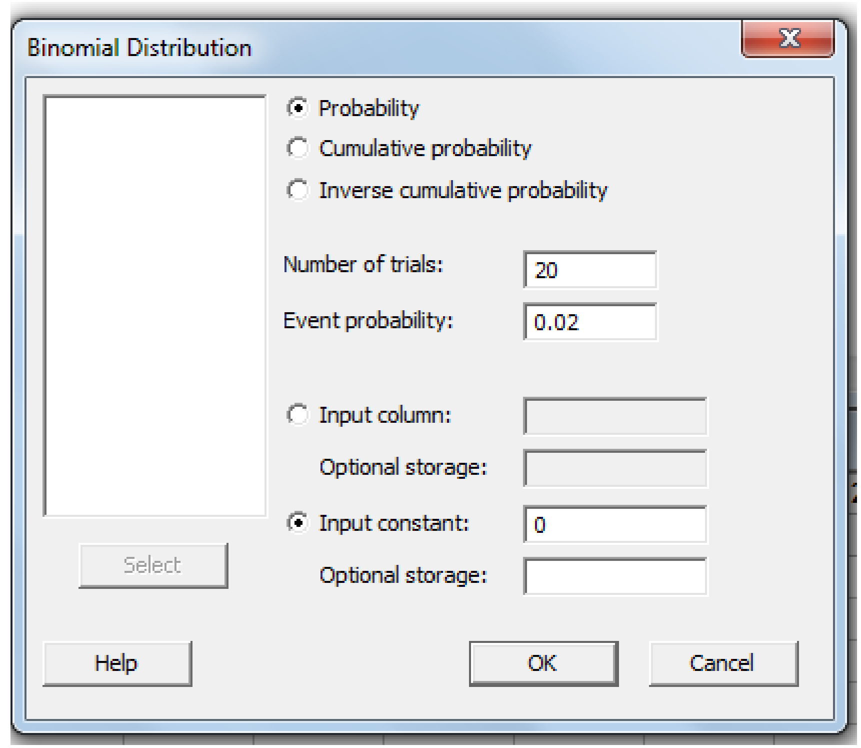
Use the formula BINOMDIST (upto Excel 2007).
BINOM.DIST (Excel 2010 onwards).
The formula template is BINOMDIST(Numbers, Trials, Probability, Cumulative)

The results returned by the BINOMDIST function is 0.667,
i.e., there is 66.7% probability of accepting such a process.

Binomial Distribution (Contd.)

In Minitab





Results returned from Minitab are shown below

x P(X = x)

0 0.667608

Poisson Distribution

- ✓ Poisson distribution is used when the subject of study is non-conformities and not **only non-conforming units**. Non-conforming units are referred to as defectives, which are studied using binomial distribution. Non-conformities in a unit are defects, which are studied using Poisson distribution.

- ✓ For example, we check the quality of 50 computers and find minor non-conformances in each computer. Applying the non-conforming units logic, we could find 100% non-conforming in this case, which really may not be the case.

- ✓ In such a scenario, we should use the Poisson distribution, with the probability distribution function given below:
  - $P(x) = \mu^x * e^{-\mu}/x!;$
  - μ = Average number of non-conformances per unit; and
  - x = Number of non-conformances in a sample.

- ✓ With P(x) we can find the exact number of non-conformances in the sample.

Poisson Distribution (Contd.)

- **Example:** A company producing marker pens conducts external audits once production of a lot is completed. Several audits on 50 pens revealed 3 minor non-conformances. What is the probability that the next audit will reveal 0 non-conformances?
- **Manual calculations:**
  - From the information, $\mu = 3$, $x = 0$
  - P (0 non-conformances) = $3^0 * e^{-3}/0! = 0.049$
- The probability of the pens having 0 non-conformances is 5% (approx.). That means probability of pens having at least one non-conformance is 95%.

Poisson Distribution (Contd.)

- Use the =POISSON() function in Microsoft Excel to calculate desired Poisson probability. (Excel 2010 and beyond use POISSON.DIST() function)

  POISSON.DIST(x, mean, cumulative)
  x: The number of events.
  mean: The expected numeric value.
  cumulative: A logical value that determines the form of the probability distribution returned.

- **In Minitab**
- Click on Calc → Probability distributions → Poisson

- Minitab results as below:

  Cumulative Distribution Function
  Poisson with mean = 3

X	P(X <= x)
0	0.0497871

- ✓Normal distribution, also known as bell curve, is a distribution most processes desire to have. This is not a mandatory requirement or compulsion for processes.
- ✓By now, it is clear how Myu (Mean) and Sigma (Standard Deviation) value should be calculated, using Excel as well as Minitab. With these two statistics, we can express a normal distribution.
- ✓z-score is an important metric, which can be used with the normal distribution:

  $z = (x - \mu)/ \sigma$

  Where "x" is a random variable from a normal distribution with mean μ and standard deviation, σ.

Normal Distribution (Contd.)

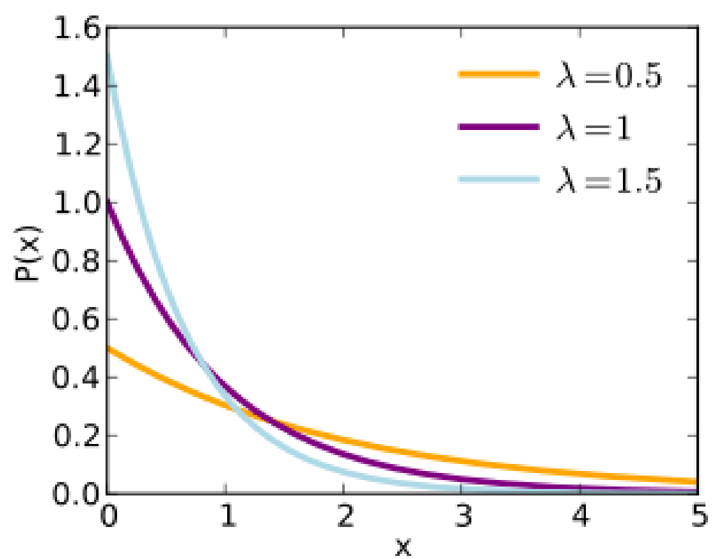
- **Example:** Assume we have checked the breaking strength of a gold wire bonding process used in microcircuit production, and we have found that the process average strength is 9 micron and the standard deviation is 4 micron. The process distribution is normal. If the engineering specification is 3 micron minimum, what percentage of the process will be below the low specification? (From The Six Sigma Quality Handbook)
- **Solution**
  - μ = 9 micron, σ = 4 micron
  - Z = (3-9)/4 = -1.5.
  - A Z-table has been provided with the toolkit. The corresponding probability value per Z table is 6.68%.

Exponential Distribution

- ✓Exponential distribution is extremely useful in analyzing reliability.
- ✓Probability density function for exponential distribution is:

$$f(x) = 1/\mu * e^{-x/\mu}$$

- ✓A typical exponential distribution function is heavily skewed. In fact, more than 63% of all the data falls below the mean.



- **Example solved using Excel:** A city water company averages 500 system leaks per year. What is the probability that the weekend crew, which works from 6 p.m. Friday to 6 a.m. Monday, will get no calls?
- **Solution:**
  - A typical year having 365 days will have a total of 8760 hours. Average number of leaks is 500. Thus, mean time between failures is 8760/500 = 17.52 hours.
  - Use the =EXPONDIST() function in Microsoft Excel (up to Excel 2007)
  - Use the =EXPON.DIST() function in Microsoft Excel (Excel 2010 and beyond)
  - Time between the ranges illustrated is 60 hours. The probability of the staff getting calls is 96.7%. The staff thus has free time worth 3.3% during weekends.

Summary

In this lesson, we've learned:

- ✓ Permutations and combinations
- ✓ Frequency distributions
- ✓ Binomial distribution
- ✓ Poisson distribution
- ✓ Normal distribution
- ✓ Exponential distribution

Section III, Lesson 9

Sigma Shift, **Mean Shift,** and Reducing Variations

Agenda

- ✓Sigma Shift
- ✓Mean Shift or Reducing Variations
- ✓Baseline Data

Sigma Shift

- ✓Conceptually, the long term sigma and short term sigma differ by 1.5, as proven by Motorola after conducting empirical studies.
- ✓This shift is known as 1.5 sigma shift.
- ✓The LT and ST sigma could differ by more or less than 1.5 sigma. But, over a period of time with data collected across multiple groups, this difference is observed to be 1.5 sigma.
- ✓The long term sigma is less than the short term sigma by 1.5. Example, if the short term sigma level is 6, the long term sigma level is 4.5.
- ✓This shift happens due to unexpected and unforeseen variations, because of which it is difficult to maintain statistical control at all times.

Mean Shift or Reducing Variations

- ✓Mean shift or reducing variations is one of the key decisions a Black Belt may take depending on how the Cp and Cpk values shape up.

- ✓Assuming the process is in statistical control.

- ✓Check the Cp value first, and then the Cpk value. Is Cp > 1? If it is greater than 1 and Cpk is also greater than 1, it indicates the process is capable of performing to customer expectations.

- ✓If both Cp and Cpk are greater than 1 and they are equal. Example, Cp = 1.3 and Cpk = 1.33, the process is considered accurate and capable. The Black Belt can then decide to move to the Control stage directly.

- ✓If both Cp and Cpk are greater than 1 and Cp > Cpk, the Black Belt should first find out the difference. If Cp = 1.3 and Cpk = 1.27, he may still want to move to the Control stage.

Mean Shift or Reducing Variations (Contd.)

- ✓ If Cp = 1.3 and Cpk = 1, then the mean needs to be centered. Variations are still in control, which means post the centering of the mean, the Black Belt can move to the Control stage.
- ✓ If Cp = 1.3 and Cpk = 0.8, mean has to be centered.
- ✓ If Cp = 0.8 and Cpk = 0.7, the process is not capable because of variations. **The project is a variation reduction project.**

Imp: These considerations should be thought of by the Black Belt only after he ascertains process stability and normality status.

Baseline Data

- ✓Initial collection of data which serves as a basis for comparison with the subsequently acquired data.
- ✓It is the current snapshot of data using current process and current measurement system / tools.
- ✓This data can also be used for calculating current sigma level or defect per million opportunities.

Summary

- ✓ Sigma shift
- ✓ Mean shift or reducing variations
- ✓ Baseline data

Measure Phase Summary

In this section we have learned:

- ✓ Pre-measure considerations and tools
- ✓ Types of data and measurement scales
- ✓ Central tendency and dispersion
- ✓ Measurement system analysis – Variables GAGE RR and attribute RR
- ✓ Stability conditions
- ✓ Capability metrics (Cp, Cpk, Cpkm, Pp, Ppk)
- ✓ Variations, variability, and capability
- ✓ Process conditions
- ✓ Data distribution (Normal, Binomial, Poisson, Exponential)
- ✓ Explaining 1.5 shift and sigma shift
- ✓ Mean shift or variations
- ✓ Baseline data

Measure: Activity Summary

- ✓ Understand types of data
- ✓ Measurement system analysis
- ✓ Sampling
- ✓ Stability check
- ✓ Descriptive statistics
- ✓ Normality check
- ✓ Capability check
- ✓ Understanding the process on distribution
- ✓ Location, shape, and spread determination

Measure Tools Summary

1. CE Matrix
2. GAGE RR (Variables for Continuous and Discrete) and attribute RR for attribute data
3. Control charts
4. Graphical summary in Minitab
5. Normality test with Anderson Darling
6. Capability analysis in Minitab
7. Reviewing current process conditions

Quiz

1. Which one of these activities should ideally come first in the implementation of a Measure stage?

a) Checking normality
b) Checking precision
c) Checking repeatability
d) Checking resolution

2. In a typical measurement system analysis activity, the GRR is 42% and the contribution is split equally amongst gage variation and part variation. What should be done?

a) Leave the GAGE. It is acceptable.
b) Calibrate the equipment.
c) Calibrate the equipment and check operator variation.
d) Calibrate the equipment, check operator variation and identify part variation.

Quiz

3. The tool that is typically used in the Measure stage to scope a project is:

a) DFMEA

b) PFMEA

c) CE Matrix

d) SIPOC

4. Which one of these is the most important pre-cursors for calculating capability?

a) Stability

b) Normality

c) No skew in data

d) None of the above

5. Which of these measures would you normally use to associate with process capability of discrete data?

a) DPMO/PPM

b) DPU

c) Cp

d) Cpk

6. In doing a normality check, the p-values for Weibull, Gamma, Laplace, and Normal tests show up 0.501, 0.423, 0.190 and 0.325. Which of these distributions should the Black Belt concur about his data?

a) Weibull

b) Gamma

c) Laplace

d) Normal

Answers

1. d) Checking resolution is the first activity in the Measure stage and thus it is the first activity amongst the list.
2. d) Presence of part variation in the study indicates that it needs to be studied.
3. d) SIPOC is the best possible scoping tool, so this is the correct answer.
4. a) Stability is the most important pre-cursor for calculating capability done before any calculations are done in measure.
5. a) With discrete data, DPMO/PPM are the best possible metrics showing capability.
6. d) If p-value of normal distribution is significant, we should select it irrespective of other distributions showing greater significant values.

Thank You