Six Sigma: Week 4

John Fico, Adjunct Professor
Fall 2020

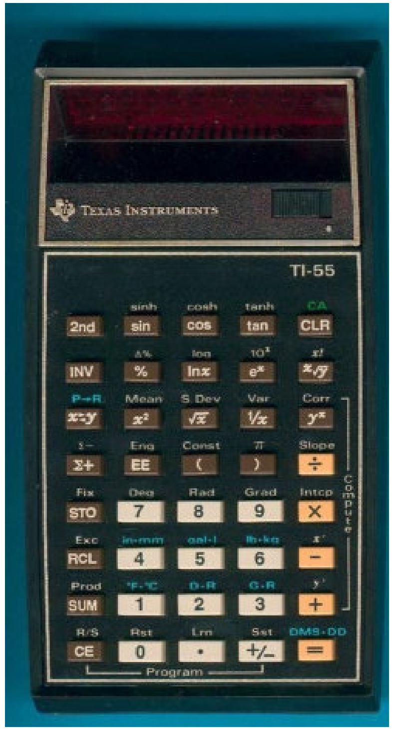
Agenda: Week 4

- ✓ Lean Sigma text now at Florida Poly Bookstore
- ✓ Schedule makeup class session for Tuesday, 9/3
  - ✓ To be held on an upcoming Friday or Saturday
  - ✓ Will discuss Article 3 assignment in class and have time to work on it
- ✓ Continue Six Sigma graphical tools introduction using Minitab
  - ✓ Box Plot demo with TI-Nspire calculator
- Assignment Review: Minitab Logistics Application (due 9/19)
  - Hint: Two Six Sigma/Minitab tools needed
- P-Values – introduction to hypothesis testing tools
  - Video (Khan Academy)
  - Additional examples
- Tools for Project Discovery (LS Text Chapter 5)
- Kaizen (LS Text Chapter 4)

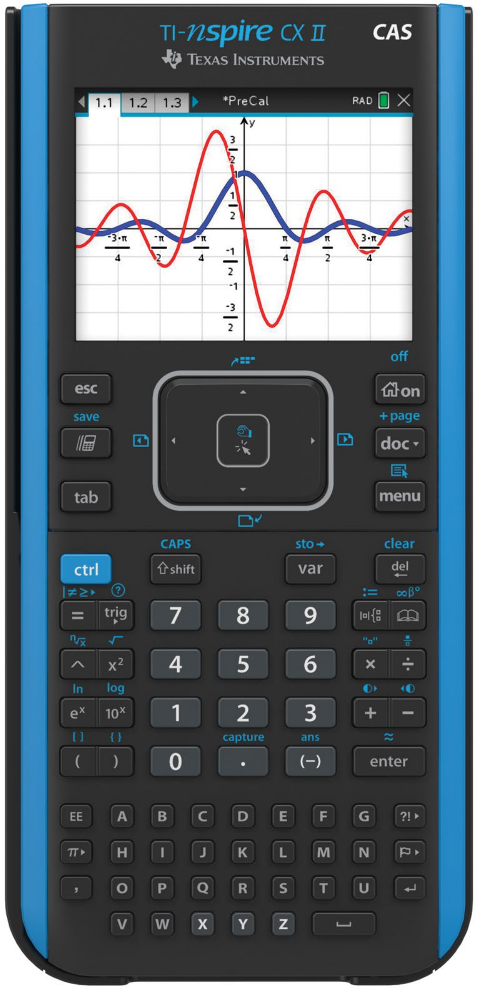
Measure: Six Sigma Graphical Tools

- ✓ Summary of graphical tools (page 118)
- ✓ Graphical Summary (page 126)
- ✓ Histograms (page 123)
- ✓ Pareto Analysis (page 136)
- ✓ Box Plots (page 140)
- ✓ Dot Plots (page 125)
- ✓ Time Series Plots (page 130)
- ✓ Scatter Plots (page 145)
- ✓ Fitted Line Plots (page 198)
- ✓ p-value – statistical tool (page 156)

Demonstration & in-class practice of Six Sigma analysis tools using Minitab



TI-55 (1977)

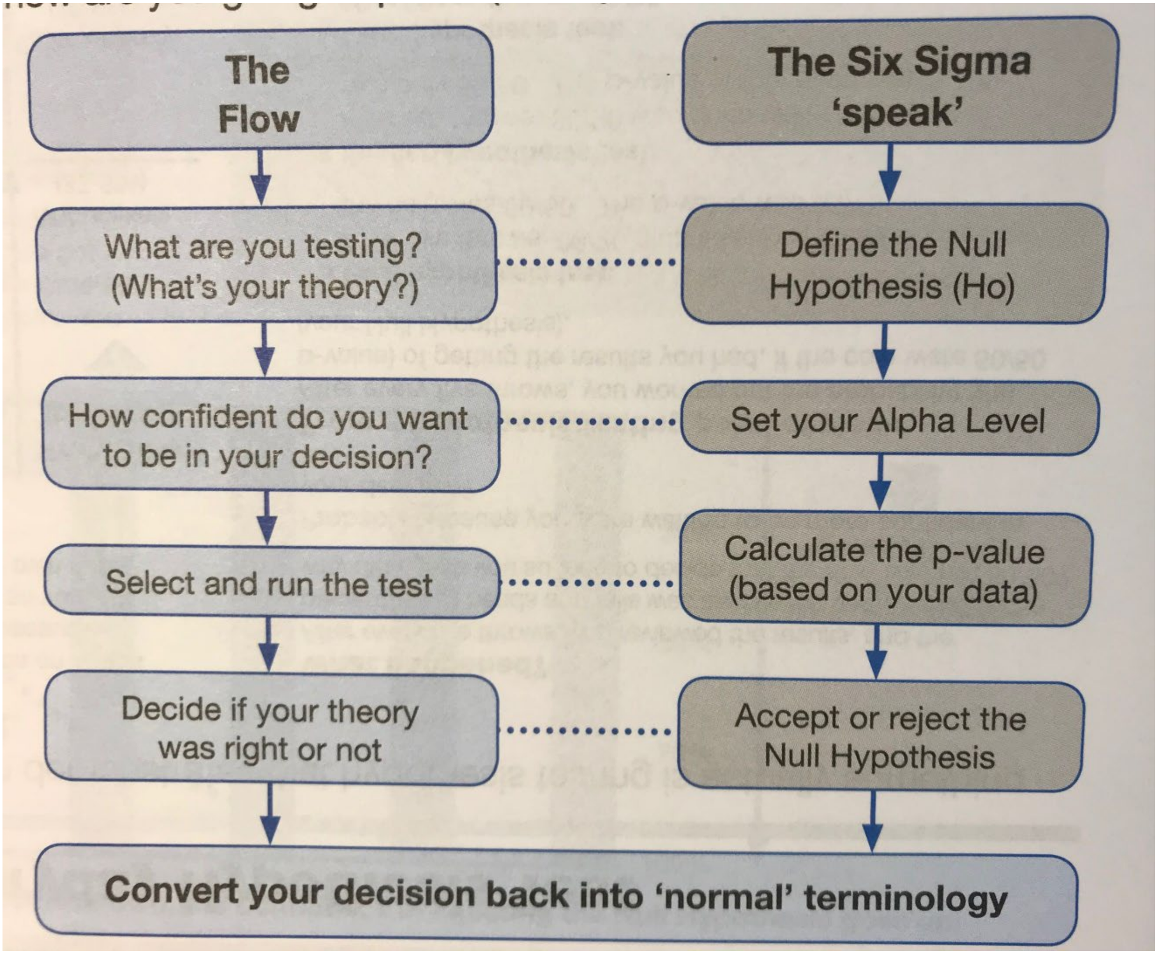


TI-Nspire (2019)

Introduction to Hypothesis Testing & P-Value

- Hypothesis testing refers to a set of tools that can tell us how certain we can be in making a specific decision or statement.
- By indicating how certain (confident) we can be in our decisions, hypothesis testing also tells us our risk of being wrong
  - This has rarely been quantified in business prior to Six Sigma deployment
  - Six Sigma is data-driven (our intuition matters too)
- P-value video (Khan Academy)
  - https://www.khanacademy.org/math/ap-statistics/tests-significance-ap/idea-significance-tests/v/p-values-and-significance-tests

Introduction to Hypothesis Testing & P-Value



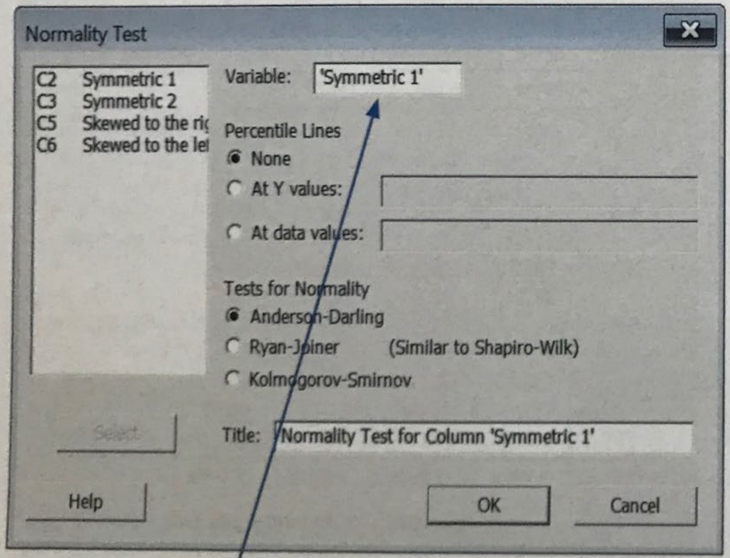
Data Normality

FLORIDA POLYTECHNIC UNIVERSITY

Anderson Darling Normality Test – Overview

Histograms with Normal curves (p123) and Probability Plots (p127) both provide **graphical** methods of assessing Normality. The Anderson Darling method provides a Hypothesis Test for assessing Normality.

Minitab: Stat > Basic Statistics > Normality Test



How to enter the data into Minitab:
The single column containing the data that you want to test is entered under **Variable**, and all other options should be left as standard.

Distributions.mpj

The Anderson Darling Normality Test is slightly different from other hypothesis tests and therefore requires careful interpretation. The hypotheses are as follows:

- The Null Hypothesis (Ho) is that the data **is** Normally distributed.
- The Alternative Hypothesis (Ha) is that the data **is not** Normally distributed.

Whereas often in hypothesis testing we are looking to **reject** the Null Hypothesis, in the case of Normality testing we are generally looking to **accept** it.

For this reason, the interpretation of the p-value is the reverse of what we are used to, as follows (for an Alpha Level of 0.05):

- If the p-value is less than 0.05, we can be confident that the data is **not** Normally distributed.
- If the p-value is greater than 0.05, there is a reasonable chance that it **could** be Normally distributed, and so we usually assume it is.

Alpha level of 0.05 translates to a 95% confidence level

Where to find the Anderson Darling Normality test:

- Minitab's **Normality Test** function (shown on the left) produces a Probability Plot (p127) and a p-value (next page).
- Minitab's **Graphical Summary** (p126) also produces the p-value result of the Anderson Darling Normality test.

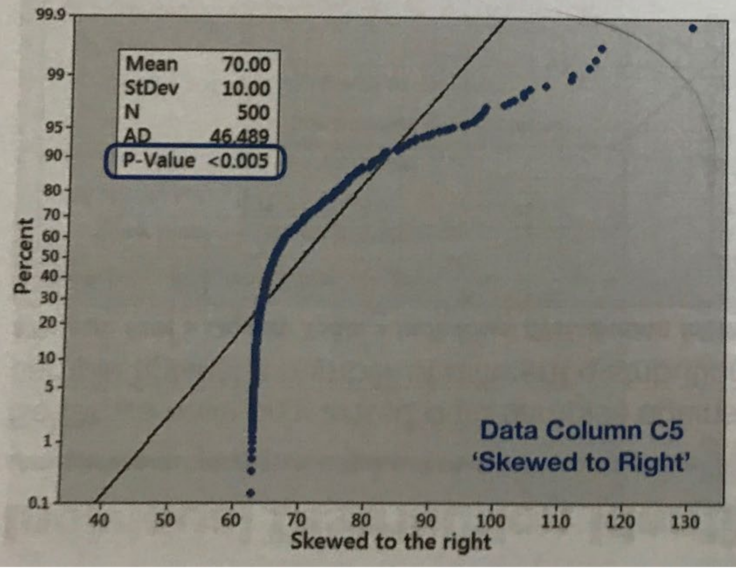
Data Normality

Anderson Darling Normality Test – Interpreting the Results

Minitab's Normality Test produces a probability plot with a summary of the data statistics and a p-value for the Anderson Darling Normality Test, as shown in the two examples below.

Example 1:
The probability plot below shows a clearly curved line, which indicates that the data set is definitely **not** Normally distributed.

In addition, the Anderson Darling p-value is very low (<0.005) which fits with the probability plot, and means that we can be very confident that the data is **not** Normally distributed.

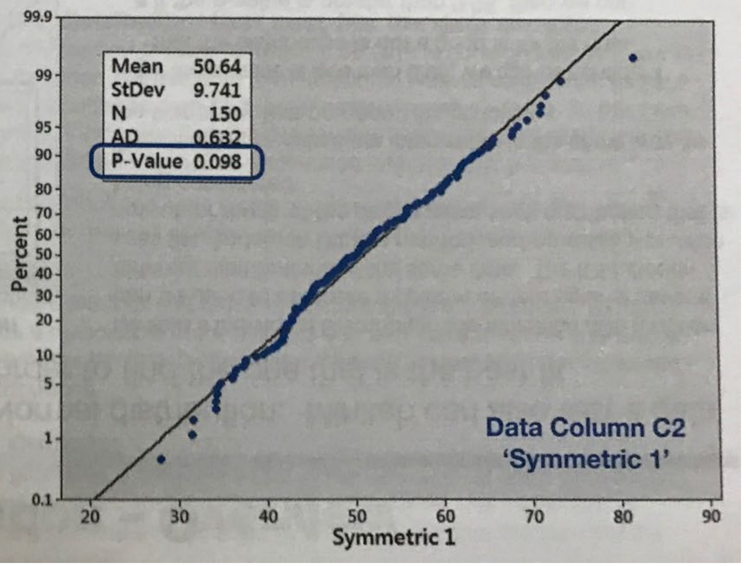


The two examples shown here are taken from the following data file:

Distributions.mpj

Example 2:
The probability plot below shows a (relatively) straight line, which indicates that the data **may** be Normally distributed.

The Anderson Darling p-value (of 0.098) fits the probability plot because it is **above 0.05**, which means that the data **can** be assumed to be Normally distributed.



Hypothesis Testing*

- Branch of statistics that determines whether a particular value of interest is contained within a calculated range (a confidence interval)
- Calculates the probability that your conclusion is wrong
- Common application of hypothesis testing is to see if two means are equal
  - Due to variation, no two data sets will ever be exactly the same even if they come from the same population

*Source: The Lean Six Sigma Pocket Tool Book, Michael George, et. Al., 2005.

Hypothesis testing will indicate if observed differences are likely due to *true* differences within the populations or due to random variation.

Hypothesis Testing

Table 3.12 The general scheme of hypothesis testing.

Hidden truth \ Decision	Accepting H_0	Rejecting H_0
H_0 is true	Correct decision	Type I error
H_0 is false	Type II error	Correct decision

Type I error: *Conclude there is a difference when there really isn't*

Type II error: *Conclude there is no difference when there really is*

Impact of Type I error: Incorrect conclusions are reached & will likely lead to ineffective solutions.
Impact of Type II errors: Solution options are treated identically even though they are not the same.

Hypothesis Testing Terms & Concepts*

- *Null hypothesis (H_0)*
  - a statement being tested to determine whether or not it is true
  - Expressed as an equation H_0: $\mu_1 = \mu_2$ or H_0: $\mu_1 - \mu_2 = 0$
    - Equations mean that within null hypothesis, the means from two sets of data are the same
      - If true, then subtracting one mean from another = 0
  - Null hypothesis is assumed to be true, unless enough evidence proves otherwise
    - If proven otherwise, null hypothesis (H_0) is rejected

*Source: The Lean Six Sigma Pocket Tool Book, Michael George, et. Al., 2005.

Hypothesis testing will indicate if observed differences are likely due to *true* differences within the populations or due to random variation.

Hypothesis Testing Terms & Concepts*

- *Alternate hypothesis (H_a)*
  - a statement that represents reality if there is enough evidence to reject H_0
  - Expressed as an equation H_a: $\mu_1 \neq \mu_2$ or H_a: $\mu_1 - \mu_2 \neq 0$
    - Equations mean that within alternate hypothesis, the means from two sets of data are not the same
  - If null hypothesis is rejected, the alternate hypothesis is accepted

*Source: The Lean Six Sigma Pocket Tool Book, Michael George, et. Al., 2005.

Hypothesis testing will indicate if observed differences are likely due to *true* differences within the populations or due to random variation.

Hypothesis Testing*

- *p-values*
  - If null hypothesis is rejected, p-value is the probability of being wrong
    - p-value is the probability of making a Type I error
  - It is the critical alpha risk (Type I) value at which the null hypothesis is rejected
  - If alpha risk should not be more than 0.05, then the null hypothesis is rejected when p-value is <=0.05.
  - Also refer to page 156 of LSSM – Minitab book for review of p-values

*Source: The Lean Six Sigma Pocket Tool Book, Michael George, et. Al., 2005.

In Six Sigma projects where entities such as means from data sets are compared, general practice is that if p-value <0.05, the means are not statistically the same.

FLORIDA POLYTECHNIC UNIVERSITY

Interpreting p-values

Interpreting p-values (also see Appendix B, p291)

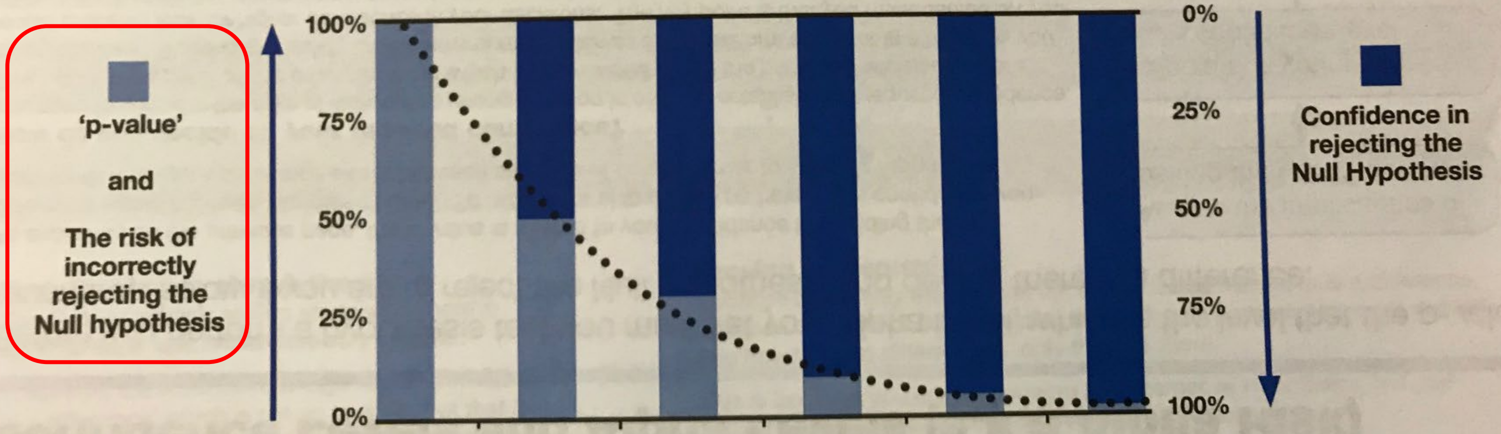
Understanding p-values is critical to interpreting hypothesis results. The best way is to find and learn a definition that works for you, and repeat it back every time you need to interpret a p-value.

A statistical definition of the p-value is:
'The probability of getting the same results (data) that you have got (or worse) if the Null hypothesis were true.'

A practical definition of the p-value is:
The p-value is your confidence in the Null Hypothesis…

*…so, when it's low, you **reject the Null** and when it's **not** low, you **keep the Null***

The definition above sounds a little abstract by itself, but take a few minutes to look back at the coin example. Every time you got another five results, you looked at **all** the results so far and estimated the probability of those results occurring if it were a 50/50 coin. You were estimating p-values.



The dotted line shows that as the p-value comes down, the confidence in rejecting the Null Hypothesis goes up.

Hypothesis Testing Uses*

- Determine statistically whether or not a value is cause for alarm
- Determine whether or not two sets of data are truly different (with a certain level of confidence)
- Identifies whether or not a statistical parameter (mean, standard deviation, etc. is different from a value of interest
- Assess the “strength” of a conclusion (probability of being correct or incorrect)

*Source: The Lean Six Sigma Pocket Tool Book, Michael George, et. Al., 2005.

Hypothesis testing is a verification, a check based on sample evidence.
This is followed by either accepting or rejecting the “null” hypothesis (H_0).

Introduction to Project Discovery

- A process is a sequence of activities with a definite beginning and end, including defined deliverables
- A “something” travels through a sequence of steps
- If a process does not have a start, and end, deliverables or an entity, it probably should not be selected as an improvement project
- For some processes, it is not obvious where the project should focus
  - Data collection/analysis and focus on customer needs helps to clarify this

.

General Guidelines for Project Selection*

- Any project should have identifiable process inputs and outputs.
- A good Six Sigma project should never have a pre-determined solution.
- If you already know the answer, then just go fix it!
- For projects that have operator or operator training as an input, focus on ways to reduce operator variation, therefore making your process more robust to different or untrained operators.
- All projects need to be approached from the perspective of understanding the variation in process inputs (inputs from SIPOC), controlling them, and eliminating the defects.

*Source: isixsigma.com

https://www.isixsigma.com/implementation/project-selection-tracking/finding-and-selecting-good-six-sigma-projects/

General Guidelines for Project Selection*

Example #1

- Problem: *We are experiencing slow cycle time at Station 30 because we are getting bad parts from Station 20 and need to rework them.*
- Non-Six Sigma Solution: Rebalance the line in order to do the rework and keep your cycle time below specifications while not spending extra labor cost.
- Six Sigma Solution: Investigate and control key inputs that contribute to making a bad part production at Station 20.

*Source: isixsigma.com

https://www.isixsigma.com/implementation/project-selection-tracking/finding-and-selecting-good-six-sigma-projects/

General Guidelines for Project Selection*

Example #2

- Problem: *We have had two quality related issues reported this year for missing armrest screws.*

- Non-Six Sigma Solution: Add sensors to detect screws further down the line. If screws are missing, operator manually fixes.
- Six Sigma Solution: Determine process inputs causing missing screws. For example, auto gun does not always feed correctly due to air pressure variation. Either study range required for 100 percent operation and control in that range, or find way to make gun more robust to compensate for air pressure variation.

*Source: isixsigma.com

https://www.isixsigma.com/implementation/project-selection-tracking/finding-and-selecting-good-six-sigma-projects/

Tools to be Applied During Project Discovery

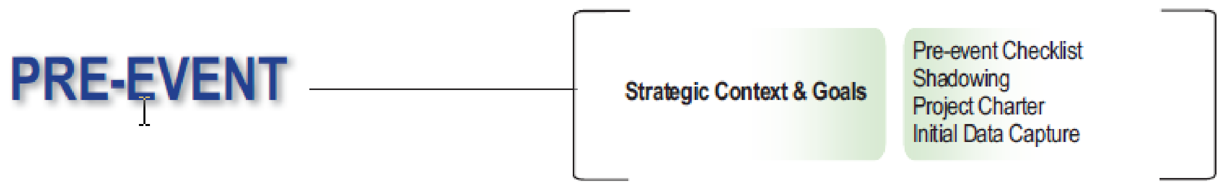
- Project Charter
- SIPOC
- Value Stream Map
- Core Process Map
- Process Scorecard – Key Process Output Variables (KPOVs)
  - Identify the “Y” and the “Xs”; Y is based on customer requirements
- Data Collection Plan
- Customer Interviewing & Surveys
- Affinity Diagramming
- Customer Requirements Trees / 5 Whys
- Gap Analysis

LS Text (Wedgwood) Chapter 5

Chapter 4: Kaizen – Class Discussion

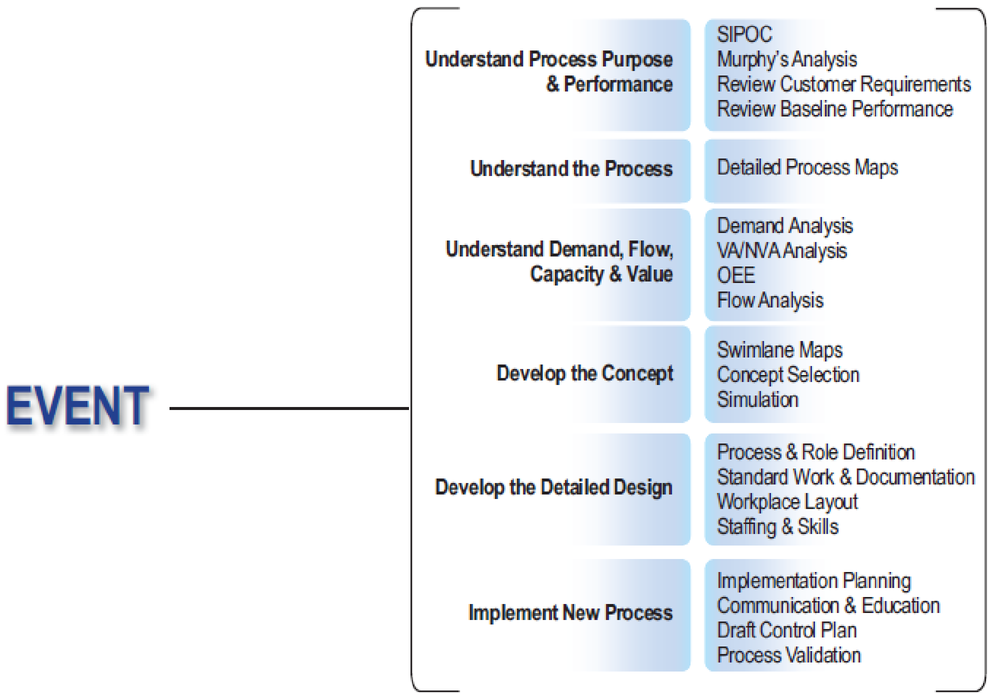
- Kaizen is Japanese for “improvement” or “change for the better”
- What are some of the key concepts of Kaizen events compared to a project?
- What is the Kaizen roadmap?
  - Next slide
- Author shows examples of problems that are
  - viable in a Kaizen
  - not viable in a Kaizen

Chapter 4: Kaizen – Class Discussion



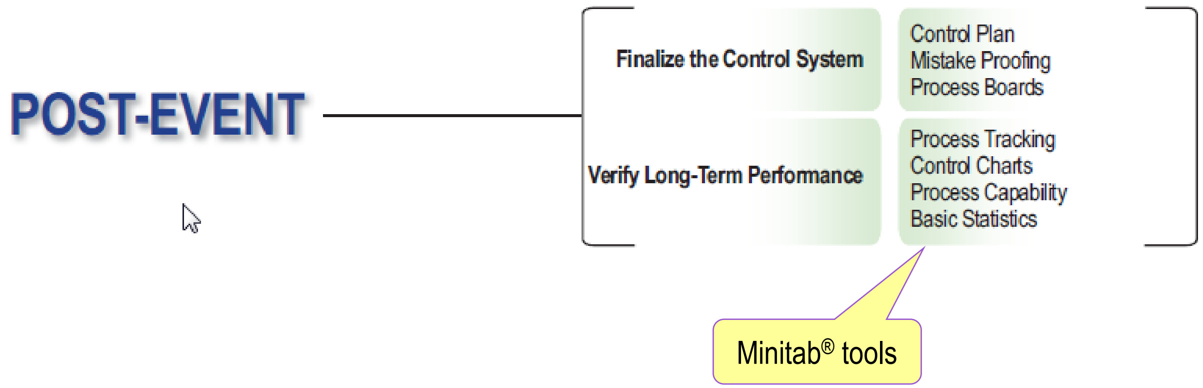
- Pre-Event (Pre-work from Belt, Champion, & Process Owner)
  - In planning a Kaizen, current state of process as defined in project charter must be baselined
  - Ensure a “Belt” is part of the team
  - “Why are we doing this and what are we trying to achieve?”
- Collect process data based on type of problem in problem statement

Chapter 4: Kaizen – Class Discussion



- Event (4-5 day session)
  - Develop the improved process and implement it
  - Day 1: Current state
    - Review process purpose and performance
    - Understand the process
  - Day 2: Develop the future state
    - Develop the concept and signoff
    - Developed detailed design
  - Days 3-5: Implement and stabilize new process

Chapter 4: Kaizen – Class Discussion



- Post-event
  - Process control mechanisms in place
  - Process is executed consistently
  - Process performance meets project goals and is sustainable
  - Process resources are maintained at appropriate competency level

Kaizen roadmap

PRE-EVENT	Strategic Context & Goals	Pre-event Checklist Shadowing Project Charter Initial Data Capture
EVENT	Understand Process Purpose & Performance	SIPOC Murphy's Analysis Review Customer Requirements Review Baseline Performance
	Understand the Process	Detailed Process Maps
	Understand Demand, Flow, Capacity & Value	Demand Analysis VA/NVA Analysis OEE Flow Analysis
	Develop the Concept	Swimlane Maps Concept Selection Simulation
	Develop the Detailed Design	Process & Role Definition Standard Work & Documentation Workplace Layout Staffing & Skills
	Implement New Process	Implementation Planning Communication & Education Draft Control Plan Process Validation
POST-EVENT	Finalize the Control System	Control Plan Mistake Proofing Process Boards
	Verify Long-Term Performance	Process Tracking Control Charts Process Capability Basic Statistics

FLORIDA POLYTECHNIC UNIVERSITY

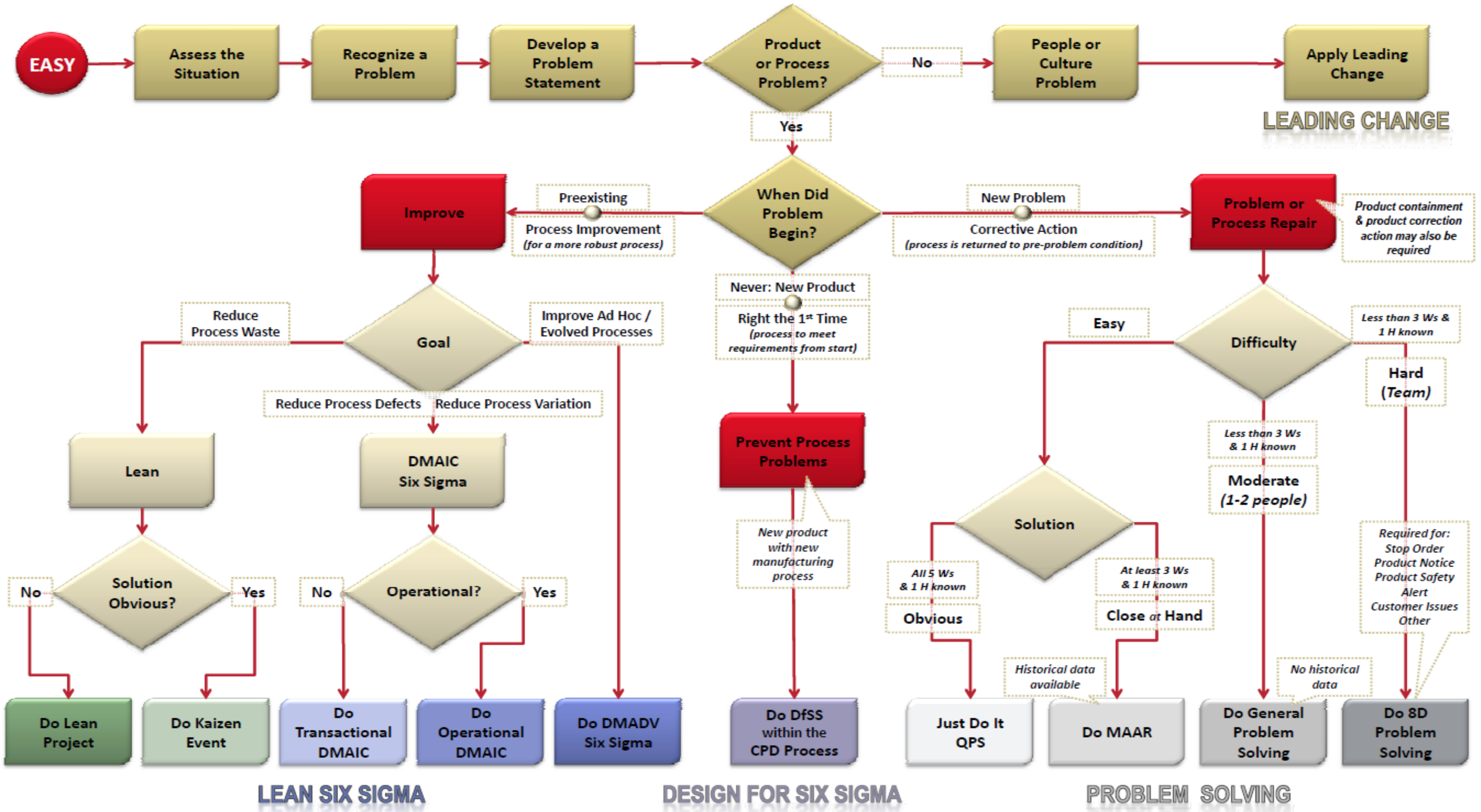
What are the Key Process Output Variables (KPOVs)?

Looking Ahead

- Graphical Tools
  - Affinity diagram
  - Cause and Effect Matrix
  - Fishbone Diagram
  - Value Stream Map
- Statistical Tools
  - Test of Equal Variance
  - t-test - 1 Sample
  - t-test - 2 Sample
  - t-test – Paired
  - Analysis of Variance (ANOVA)

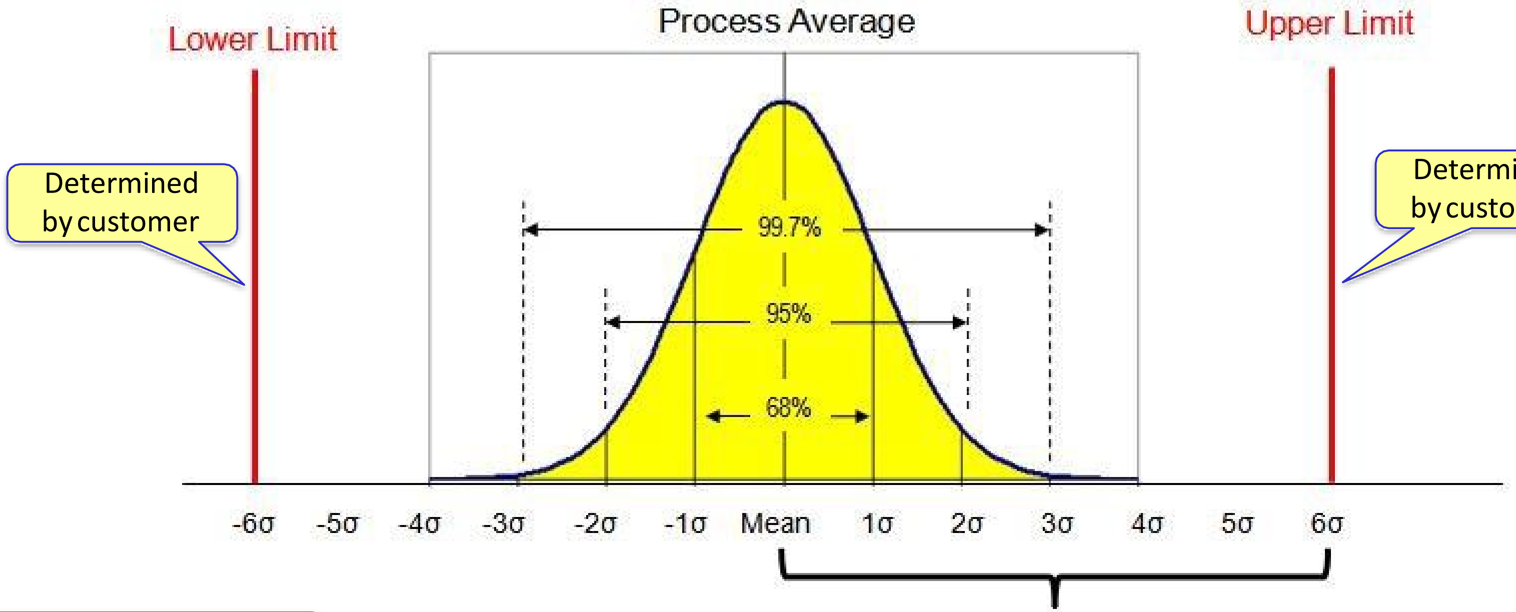
Problem Solvers Guide

FLORIDA POLYTECHNIC UNIVERSITY



Six Sigma Defined Visually (cont'd)

FLORIDA POLYTECHNIC UNIVERSITY



Initially deployed at Motorola in 1986. Adopted at GE at a global scale in 1990s; inspiring many other companies to follow.

Example: In a 2 Sigma process, 95% of the measured values taken in a process will be within two standard deviations from the process average.

Within in a standard normal distribution:
- 68% of the data points will fall within ± one standard deviation from the mean
- 95% will fall within ± two standard deviations
- 99.73% of the data points will fall within ± three standard deviations from the mean

σ = Standard Deviation