Statistical Process Control – Part II

Eric J. Samp, Ph.D., CQE, CQM, 6σMBB
MillerCoors

Summary

Process control is the next frontier in a QA/QC program. Understanding variation in your process from a raw material, people, process, and equipment sources can provide insight into issues before they happen. This workshop is geared toward those already with a basic knowledge of process control and statistics. In this advanced session, industry specialists will showcase the tools and methods that are aimed at understanding how well your process meets requirements and methods to improve your processes. Speakers will cover how to execute process capability studies and statistics used to quantify this, normality testing, t-tests, analysis of variance (ANOVA), and correlation analysis. An introduction into measurements systems analysis will also be provided. This workshop will teach you how to effectively understand sources of variation in your processes so that these can be addressed.

Course Pre-Requisites

- SPC Fundamentals
  - X-Bar &R
  - IX & MR
  - Common cause vs Special Cause Variation
  - Basic Statistics

- Spreadsheet based techniques
  - Determining Grand Averages
  - Determining Standard Deviations

Course Overview

- Process Capability Studies & How to Improve

- Measurement Systems Analysis

Software Overview

	QIMacros Excel 2003-2013	Minitab® 17 + Quality Companion	CHARTrunner Lean®	SigmaXL® Excel 2010/2013	SPC XL
Users	No Belts to Black Belts	Statisticians & Black Belts	Statisticians & Black Belts	Green Belts & Black Belts	Green Belts & Black Belts
Price	$229	$1,495 + $1,195	$990+$295/yr	$249	$249
Type of Program	Excel Add-in PC & MAC	Stand Alone – Import Data from Excel	Stand Alone - Import Data from Excel	Excel Add-in PC & MAC	Excel Add-in PC only
Ease of Use & Sharing	Graphical point and click interface (Excel)	Multi-form interface Proprietary file format	Multi-form interface Proprietary Files	Multi-form interface Excel	Multi-form interface Excel
Training Required	Minutes Free online	Days/weeks Instructor-based	Days Instructor-based	Days Instructor-based	Days Instructor-based
Choosing the right chart or statistic	Wizards with built-in rules choose the right chart or statistic for you	Users have to know the rules to choose the right chart or statistic. No Wizard	Users have to know the rules to choose the right chart or statistic. No Wizard.	Users have to know the rules to choose the right chart or statistic. No Wizard.	Users have to know the rules to choose the right chart or statistic. No Wizard.
Lean Six Sigma Tools	44 Charts, 28 Stats, 100s of Document and Tool Templates	Common charts and uncommon statistics Lean tools cost extra	Limited options – DOE and GageR&R cost $1,000s more.	Common charts and templates	Common charts and templates DOE additional
Chart Quality & Tailoring	Clean, crisp, color one-click chart	B/W and Color (v17) Hard to change	Missing features	Missing features Excel changes	Missing features Excel changes
Data Mining & Analysis	PivotTable and Stat Wizards automate data mining/analysis	CrossTab & Stat tools No wizards, Decision-trees only.	Excel PivotTable No Wizard	Excel PivotTable No Wizard; Chart selection menu	Excel PivotTable No Wizards; Chart selection menu

Process Capability Analysis Overview

- Normal Distribution basics
- Lyapunov's Theorum
- Y=F(X) & DMAIC
- Assumptions on Process Capability Studies
  X~iid $N(\mu,\sigma^2)$
- Determination of indices (Pp, Ppk, Ppm)
- Capability 6 Six Pack
- DMAIC Road Map to Process Improvement
  SPC, t-test, ANOVA, Correlation
- Deriving Functional Limits on the X-vars

Normal Distribution Basics

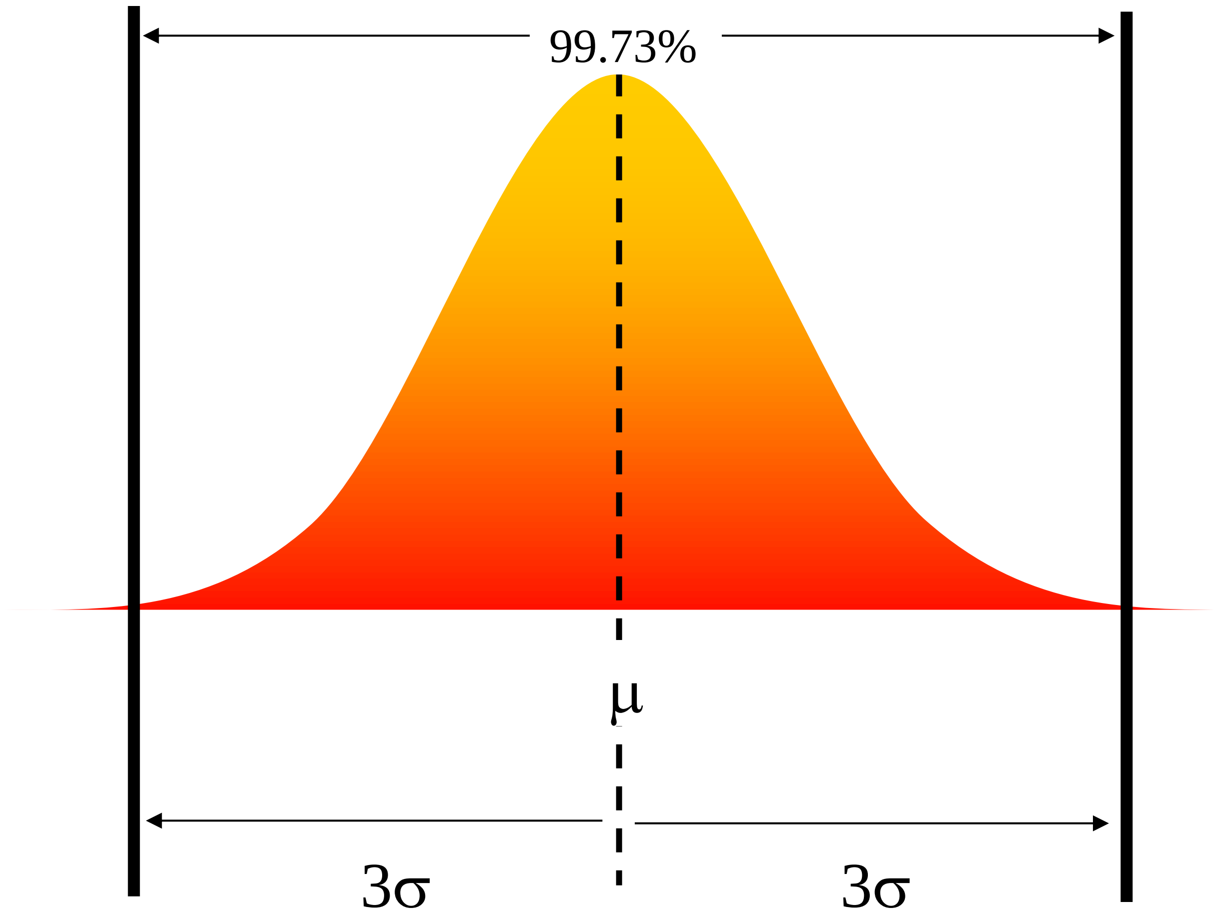
- Gaussian Distribution

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

Carl Friedrich Gauss

- Two Parameters:
  - $\mu \rightarrow$ Central Tendency
  - $\sigma \rightarrow$ Spread (Variation)

Normal Distribution Basics



Normal Distribution Basics

- Estimation of μ

$$\hat{\mu} = \overline{\overline{X}} = \sum_{i}^{p} \sum_{j=1}^{n} \frac{X_{ij}}{np}$$

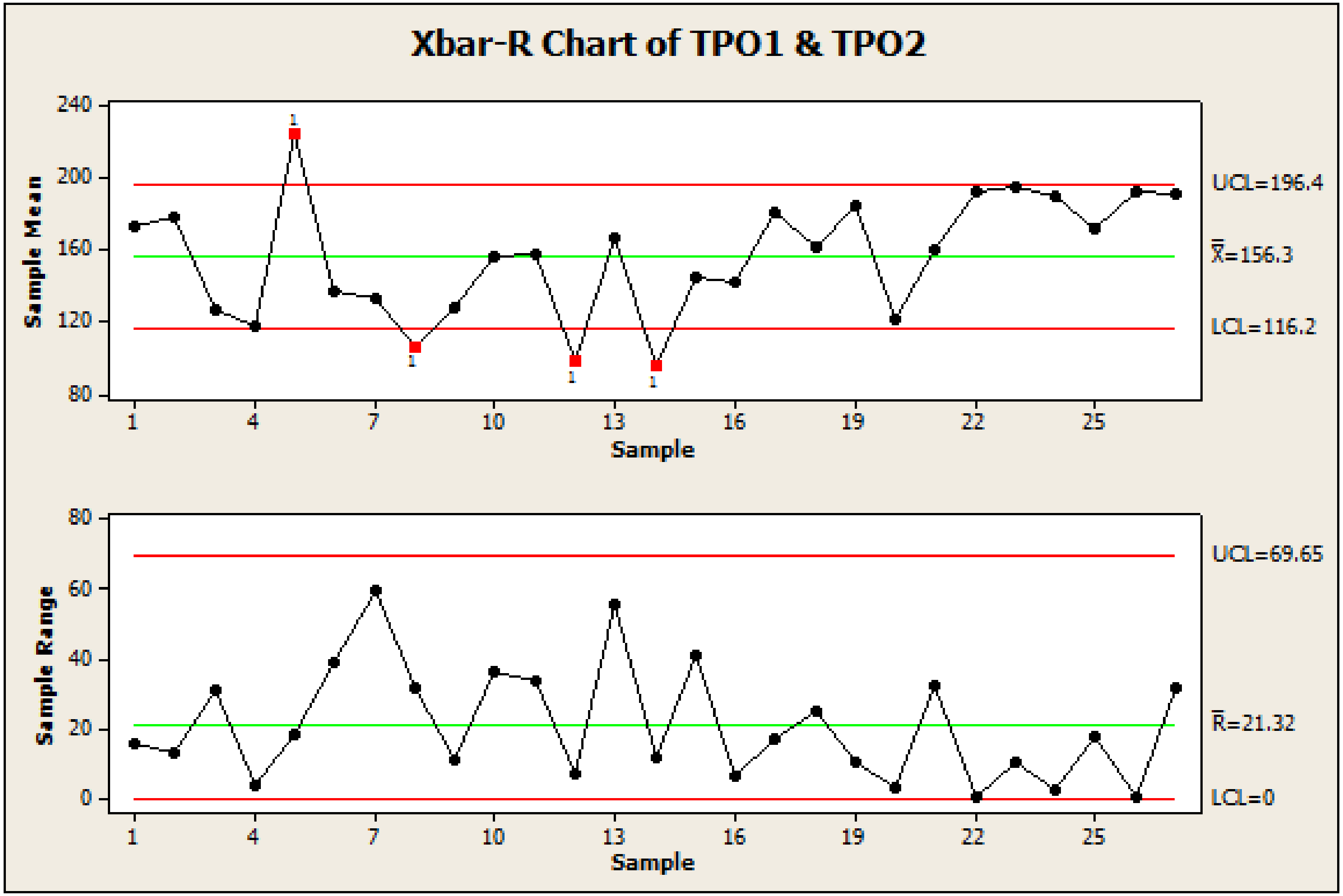
- Estimation of σ

$$\hat{\sigma}_{shortterm} = \frac{\overline{R}}{d_2} =$$

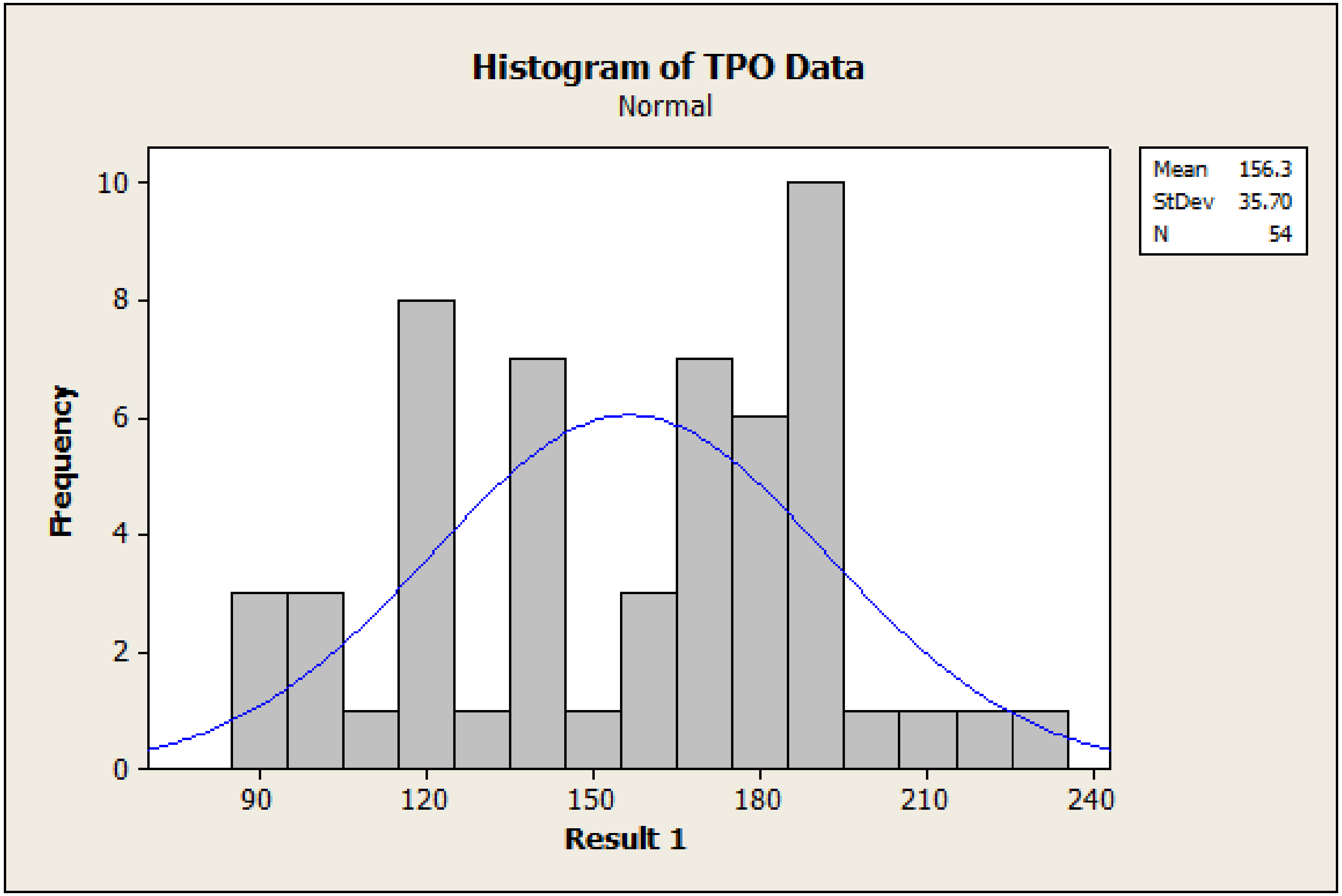
n	d2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704

$$\hat{\sigma}_{longterm} = S = \sqrt{\sum_{i}^{p} \sum_{j} \frac{\left(X_{ij} - \overline{\overline{X}}\right)^2}{p(n-1)}}$$

Normal Distribution Basics



Normal Distribution Basics



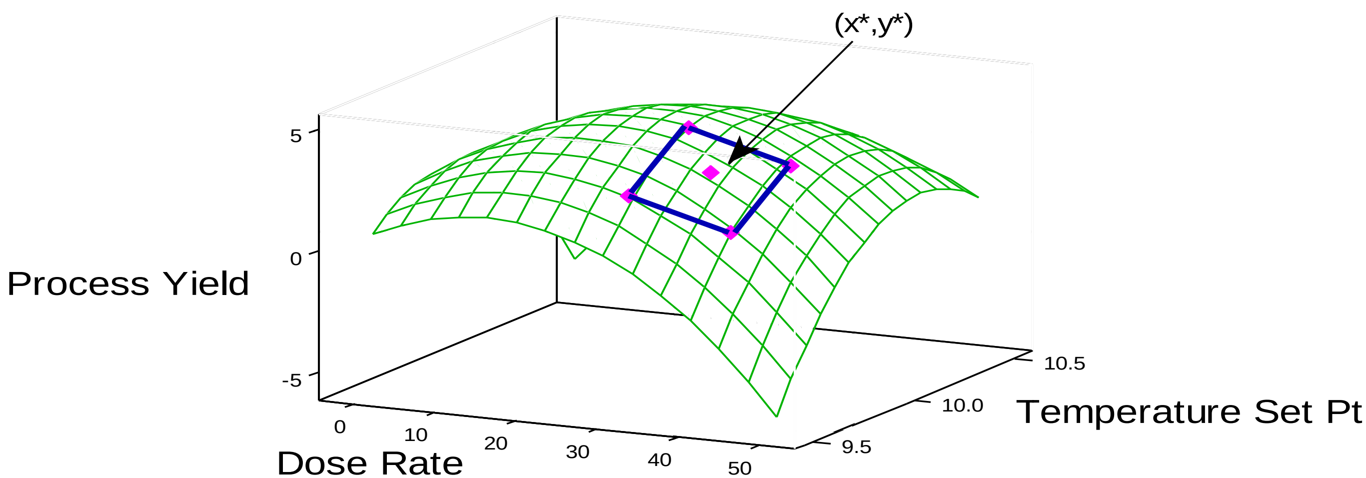
Y=F(X): Underlying principle of 6 Sigma

Consider the Taylor Series Expansion of a function $f(x_1,x_2)$ about a point (x^*,y^*)

$$f(x_1,x_2) \cong f(x_1^*,x_2^*) + \left.\frac{\partial f(x_1,x_2)}{\partial x}\right|_{(x_1^*,x_2^*)}(x_1 - x_1^*) + \left.\frac{\partial f(x_1,x_2)}{\partial x_2}\right|_{(x_1^*,x_x^*)}(x_2 - x_2^*) + \left.\frac{\partial^2 f(x_1,x_2)}{\partial x_1 \partial x_2}\right|_{(x_1^*,x_2^*)}(x_1 - x_1^*)(x_2 - x_2^*)$$

$$+ \left.\frac{\partial^2 f(x_1,x_2)}{\partial^2 x_1}\right|_{(x_1^*,x_2^*)}(x_1 - x_1^*)^2 + \left.\frac{\partial^2 f(x_1,x_2)}{\partial^2 x_2}\right|_{(x_1^*,x_2^*)}(x_2 - x_2^*)^2 + \Re_n$$

Taylor Series Expansion Approximation
to the Response Surface

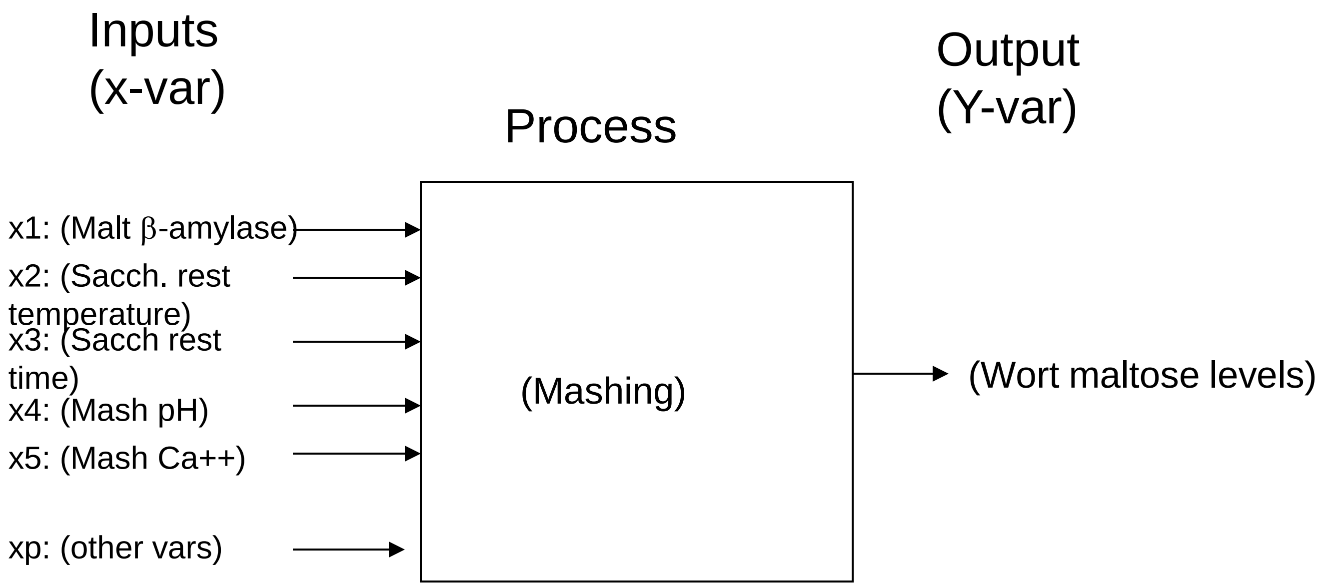


Any function f(x) can be reasonably approximated by the first few terms of the Taylor Series expansion about a point in space within a close region around "x"

$$Y_p = \beta_0 + \beta_1 X_{doserate} + \beta_2 X_{temp} + B_{12} X_{dosextemp}$$

The Science of Beer

Systems Thinking (SIPOC Model)



- In general, the output variable is a complex function of many input variables (x-vars). Some input variables we know quite well from brewing science and some we don't.
- The x-vars are not necessarily fixed (e.g. mash pH) and they also have a random component to them which may not exhibit Gaussian behavior
- So why is it justified to assume that the output variable (Y-var) are normally distributed?

Lyapunov's Theorem (1901)

Lyapunov CLT.[6] Suppose $\{X_1, X_2, \ldots\}$ is a sequence of independent random variables, each with finite expected value μ_i and variance σ_i^2. Define

$$s_n^2 = \sum_{i=1}^{n} \sigma_i^2$$

If for some $\delta > 0$, the *Lyapunov's condition*

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{i=1}^{n} \mathrm{E}\left[\, |X_i - \mu_i|^{2+\delta} \right] = 0$$

is satisfied, then a sum of $(X_i - \mu_i)/s_n$ converges in distribution to a standard normal random variable, as n goes to infinity:

$$\frac{1}{s_n} \sum_{i=1}^{n} (X_i - \mu_i) \xrightarrow{d} \mathcal{N}(0,\ 1).$$

- Brewer's condensed version:

  No matter what the underlying distribution of the input variables are, provided that they remain stable, the output variables will tend to have a Gaussian or bell shaped curve associated with behavior.

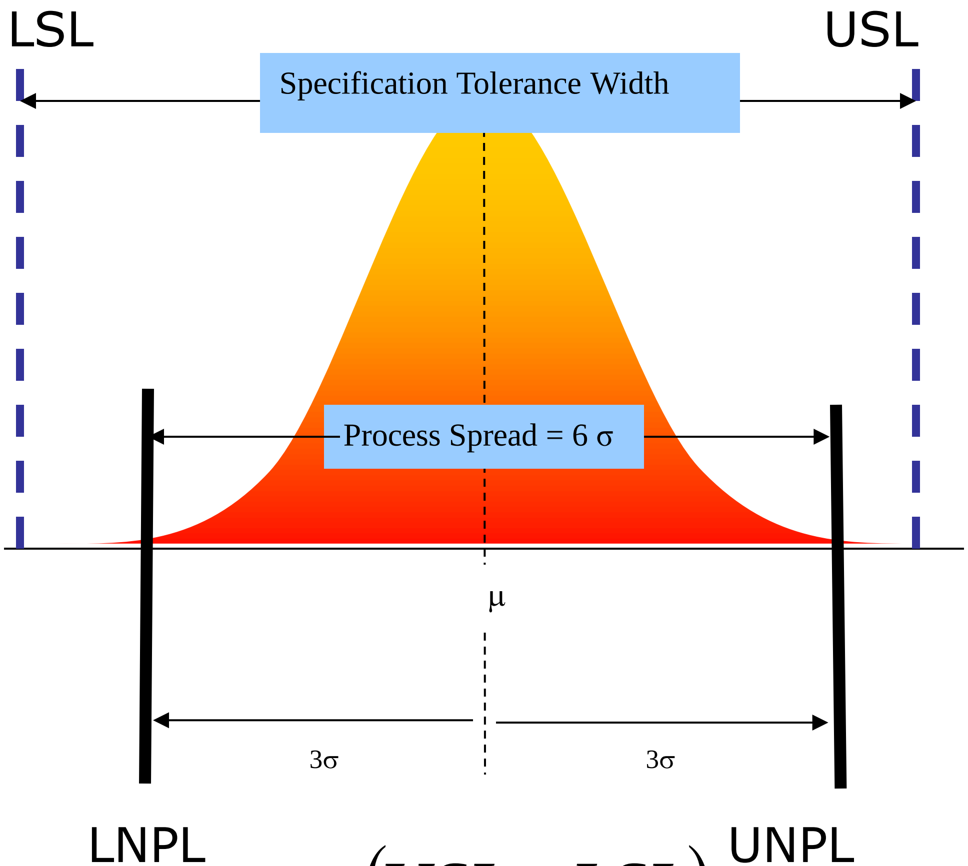
Assumptions

- Y ~ iid N(μ,σ^2)
- Y – output variable we want to study
- ~ - "is distributed as"
- First 'i': **I**ndependence
- 'id': **I**dentically **D**istributed
- N(μ,σ^2): data comes from the same normal population that has mean μ and standard deviation σ

Assumptions

Assumption	Validated by	Comment
Independence	Correlation Analysis within a subgroup and time series analysis	Not typically performed but heavily violated in our industry (TPOs, Fills, CO_2)
Identically Distributed	SPC Charts (X-bar & R) (IX & MR)	If SPC charts exhibit out of control conditions, technically a Process Capability Study is invalid; however, there is still merit in generating the Capability Histogram / 6 Pack
Normal Random Variables	Normality Tests: a) Anderson-Darling b) Ryan-Joiner c) Kolmogorov-Smirnoff	If processes are in-control but exhibit non-normal behaviour it is likely additional sources of variation are present (filler-valve to valve)

First Generation Index: Process Potential (Cp)



$$Cp = \frac{(USL - LSL)}{6\sigma}$$

Calculating Indices – Short & Long Term

- Cp (Short Term)

$$\hat{\sigma}_{shortterm} = \frac{\overline{R}}{d_2}$$

$$Cp(ShortTerm) = \frac{(USL - LSL)}{6\sigma_{shortterm}}$$

n	d2
2	1.128
3	1.693
4	2.059 ←
5	2.326
6	2.534
7	2.704

- Pp (Long Term)

$$\hat{\sigma}_{longterm} = S = \sqrt{\sum_{i}^{p}\sum_{j}\frac{\left(X_{ij} - \overline{\overline{X}}\right)^2}{p(n-1)}}$$

$$Pp(LongTerm) = \frac{(USL - LSL)}{6S}$$

Calculating Indices - Example

- Fills Data (12oz (355.0 ml) Cans) n=4
- LSL = 350.0 mls   USL = 360.0 mls
- Short Term and Long Term Stdevs next page

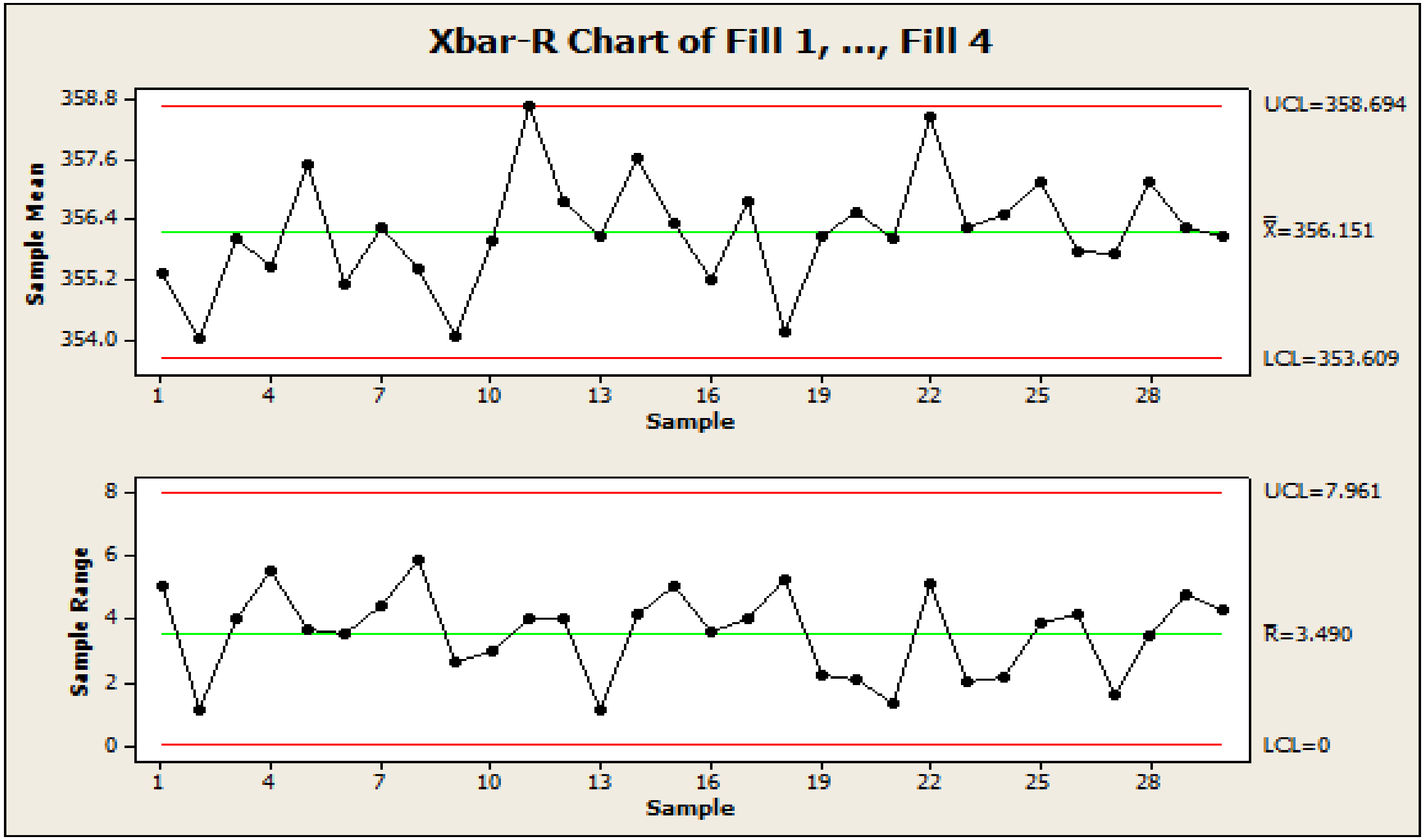
$$\hat{\sigma}_{shortterm} = \frac{\overline{R}}{d_2}$$

$$\hat{\sigma}_{longterm} = S =$$

$$Cp(ShortTerm) = \frac{(USL - LSL)}{6\sigma_{shortterm}}$$

$$Pp(LongTerm) = \frac{(USL - LSL)}{6S}$$

Calculating Indices



Given S = 1.838 mls

Calculating Indices - Example

- Fills Data (12oz (355.0 ml) Cans)
- LSL = 350.0 mls   USL = 360.0 mls
- Short Term and Long Term Stdevs next page

$$\hat{\sigma}_{shortterm} = \frac{\overline{R}}{d_2} = \frac{3.49}{2.059} = 1.40$$

$$\hat{\sigma}_{longterm} = S = 1.838$$

$$Cp(ShortTerm) = \frac{(USL - LSL)}{6\sigma_{shortterm}} = \frac{(360.0 - 350.0)}{6*1.40} = 1.19$$

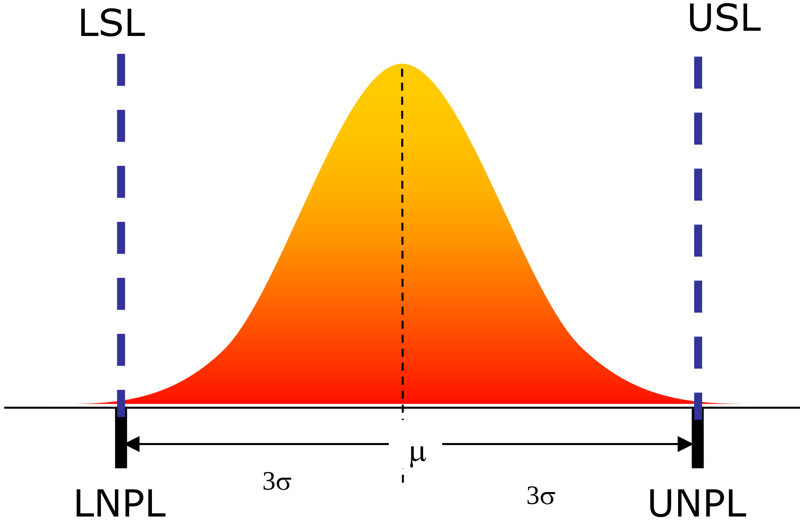
$$Pp(LongTerm) = \frac{(USL - LSL)}{6S} = 0.91$$

What is a Good Pp ?

$$if\ \ Pp = 1$$

$$\Rightarrow 6\sigma_{longterm} = (USL - LSL)$$

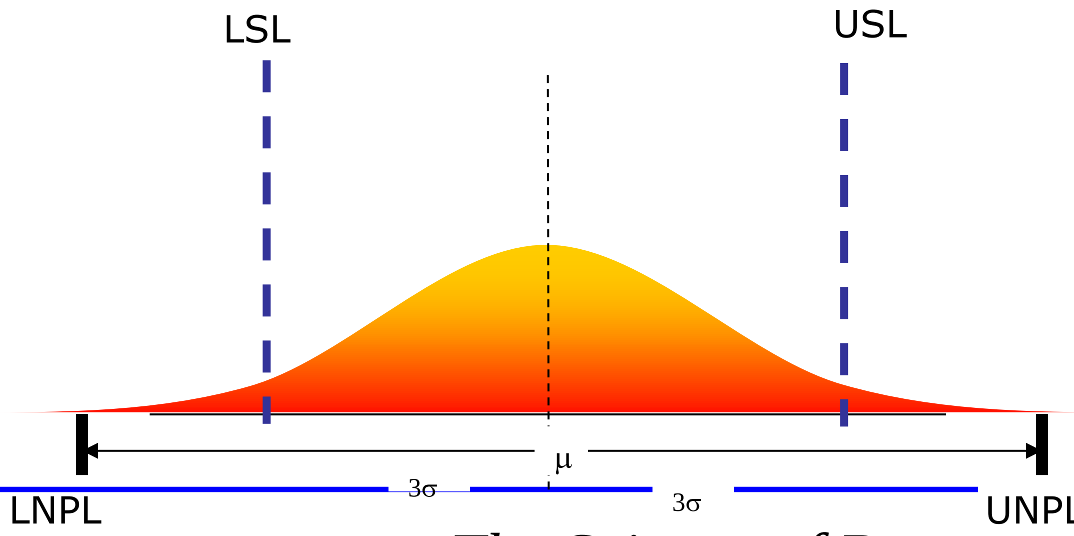
$$\Rightarrow$$



$$if\ \ Pp = 0.5$$

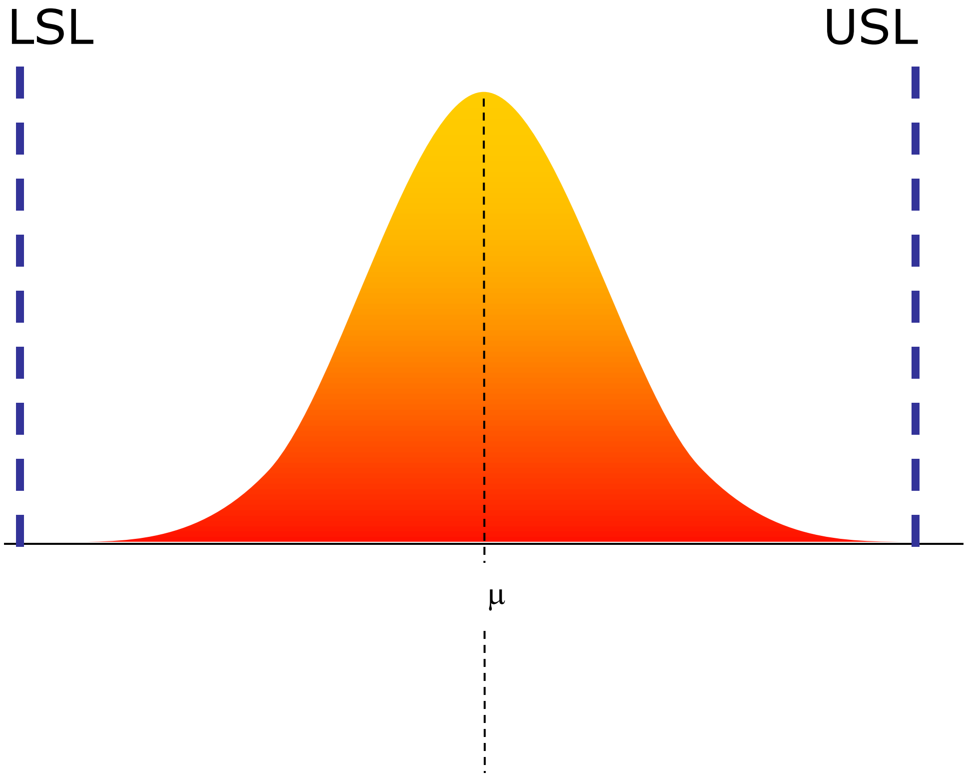
$$\Rightarrow 6\sigma_{longterm} = 2*(USL - LSL)$$

$$\Rightarrow$$



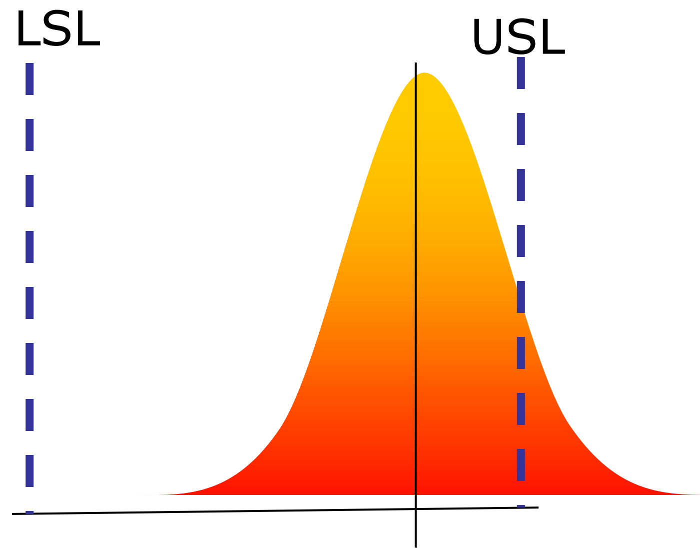
Calculating Indices

- Pp/Cp only measures “Potential”



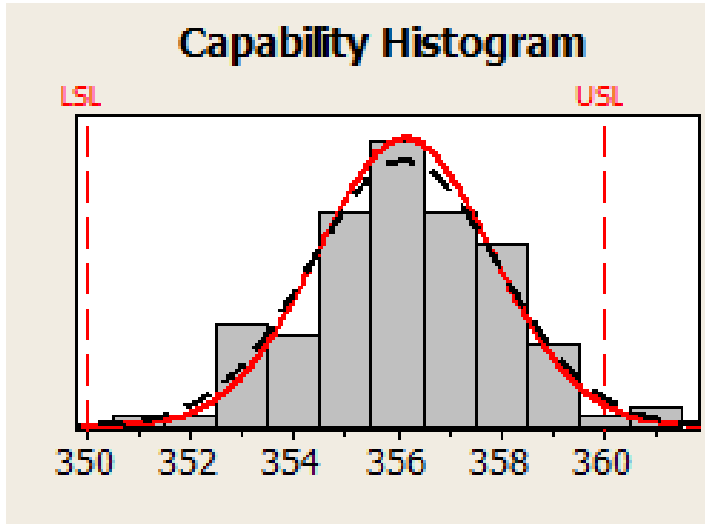
Calculating Indices – Ppk/Cpk

- Compare the distance from the mean to each specification, normalized to 3 standard deviations



Short Term	Long Term
$\hat{C}_{pl} = \left(\dfrac{\overline{\overline{X}} - LSL}{3\hat{\sigma}_{shortterm}} \right)$ $\quad \hat{C}_{pu} = \left(\dfrac{USL - \overline{\overline{X}}}{3\hat{\sigma}_{shortterm}} \right)$	$\hat{P}_{pl} = \left(\dfrac{\overline{X} - LSL}{3S} \right)$ $\quad \hat{P}_{pu} = \left(\dfrac{USL - \overline{X}}{3S} \right)$
$\hat{C}pk = \min\left\{ \hat{C}_{pl}, \hat{C}_{pu} \right\}$	$\hat{P}pk = \min\left\{ \hat{P}_{pl}, \hat{P}_{pu} \right\} =$

2nd Generation Index: Ppk & Cpk



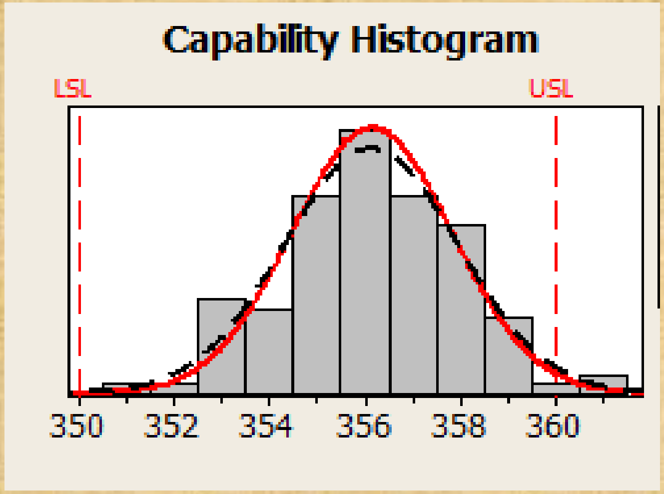
$$\hat{C}_{pl} = \left(\frac{\overline{\overline{X}} - LSL}{3\hat{\sigma}_{shortterm}} \right) \qquad \hat{C}_{pu} = \left(\frac{USL - \overline{\overline{X}}}{3\hat{\sigma}_{shortterm}} \right)$$

$$\hat{P}_{pl} = \left(\frac{\overline{\overline{X}} - LSL}{3S} \right) \qquad \hat{P}_{pu} = \left(\frac{USL - \overline{\overline{X}}}{3S} \right)$$

$$\hat{C}pk = \min\left\{ \hat{C}_{pl}, \hat{C}_{pu} \right\}$$

$$\hat{P}pk = \min\left\{ \hat{P}_{pl}, \hat{P}_{pu} \right\} =$$

2nd Generation Index: Ppk & Cpk



$$\hat{C}_{pl} = \left(\frac{\overline{\overline{X}} - LSL}{3\hat{\sigma}_{shortterm}} \right) = \left(\frac{356.151 - 350.00}{3*1.40} \right) = 1.46$$

$$\hat{C}_{pu} = \left(\frac{USL - \overline{\overline{X}}}{3\hat{\sigma}_{shortterm}} \right) = \left(\frac{360.00 - 356.151}{3*1.40} \right) = 0.92$$

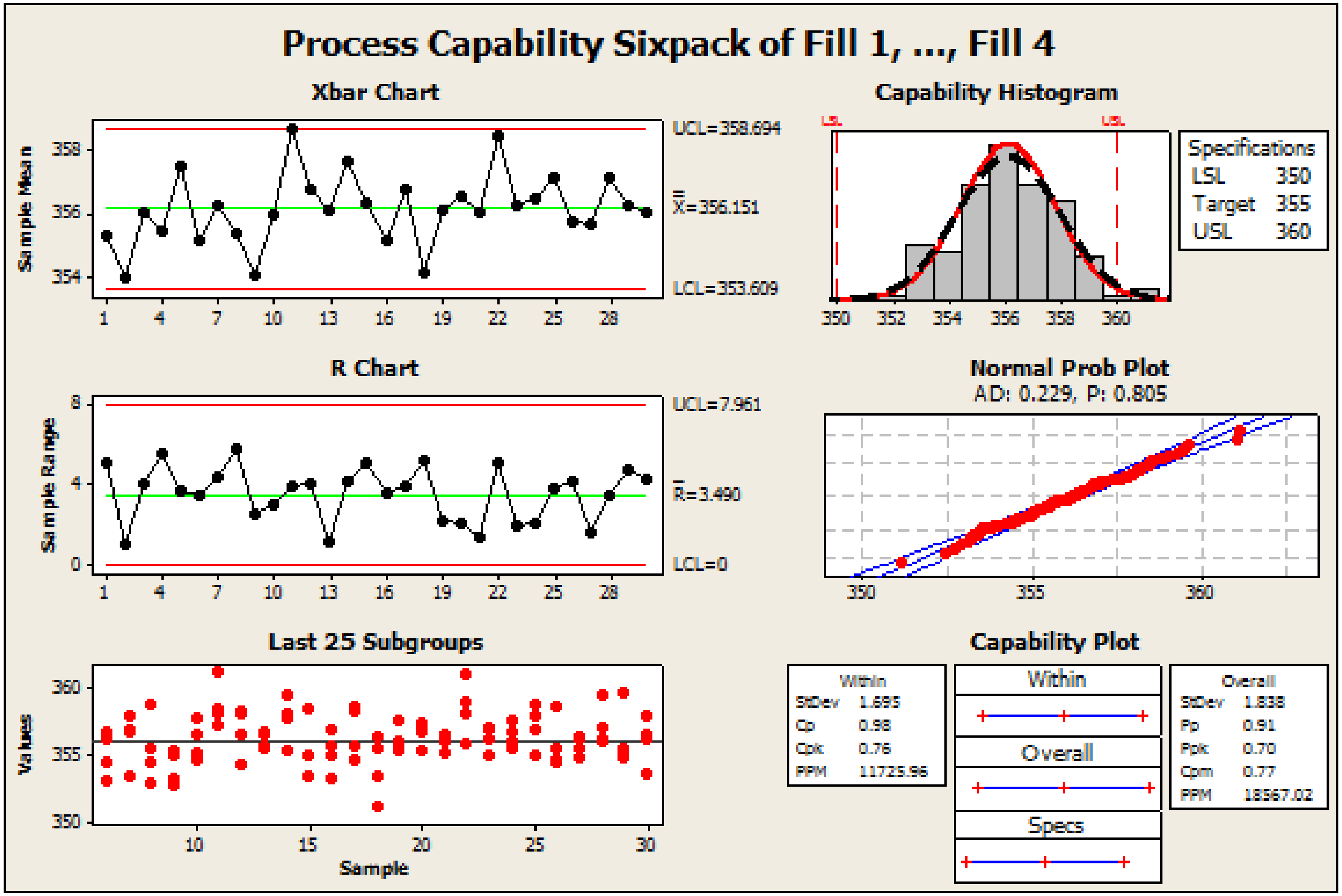
$$\hat{C}pk = \min\left\{ 1.46, 0.92 \right\} = 0.92$$

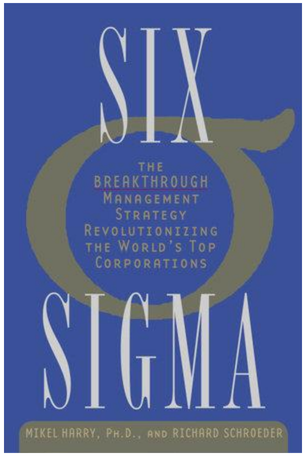
$$\hat{P}_{pl} = \left(\frac{\overline{\overline{X}} - LSL}{3S} \right) = \frac{356.151 - 350.0}{3*1.838} = 1.12$$

$$\hat{P}_{pu} = \left(\frac{USL - \overline{\overline{X}}}{3S} \right) = \left(\frac{360.00 - 356.151}{3*1.838} \right) = 0.70$$

$$\hat{P}pk = \min\left\{ \hat{P}_{pl}, \hat{P}_{pu} \right\} = 0.70$$

Introduction to the Capability 6 Pack





Testing Normality

The Anderson-Darling test is defined as:

H_0: The data follow a specified distribution.

H_a: The data do not follow the specified distribution

Test Statistc: The Anderson-Darling test statistic is defined as

$$A^2 = -N - S$$

where

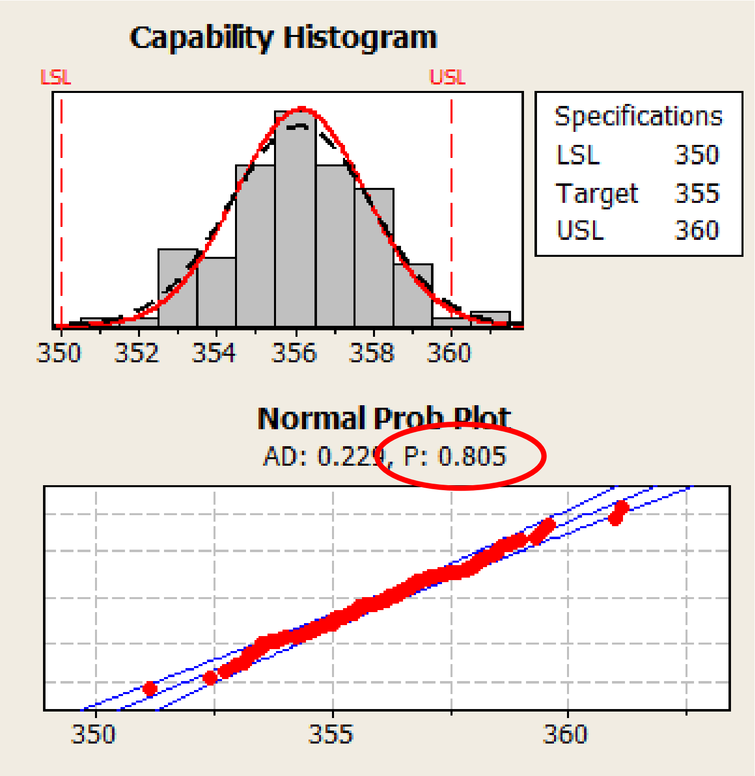
$$S = \sum_{i=1}^{N} \frac{(2i-1)}{N} [\ln F(Y_i) + \ln (1 - F(Y_{N+1-i}))]$$

F is the cumulative distribution function of the specified distribution. Note that the Y_i are the *ordered* data.

Blah Blah Blah Blah Blah ……………………..

Testing Normality

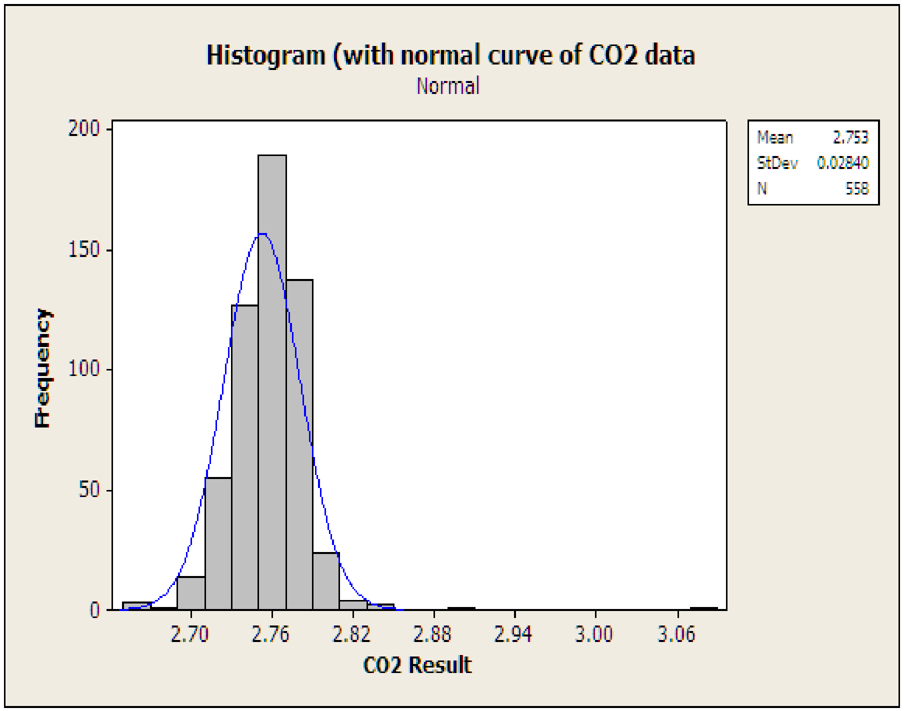
- P-value Is the only thing to worry about

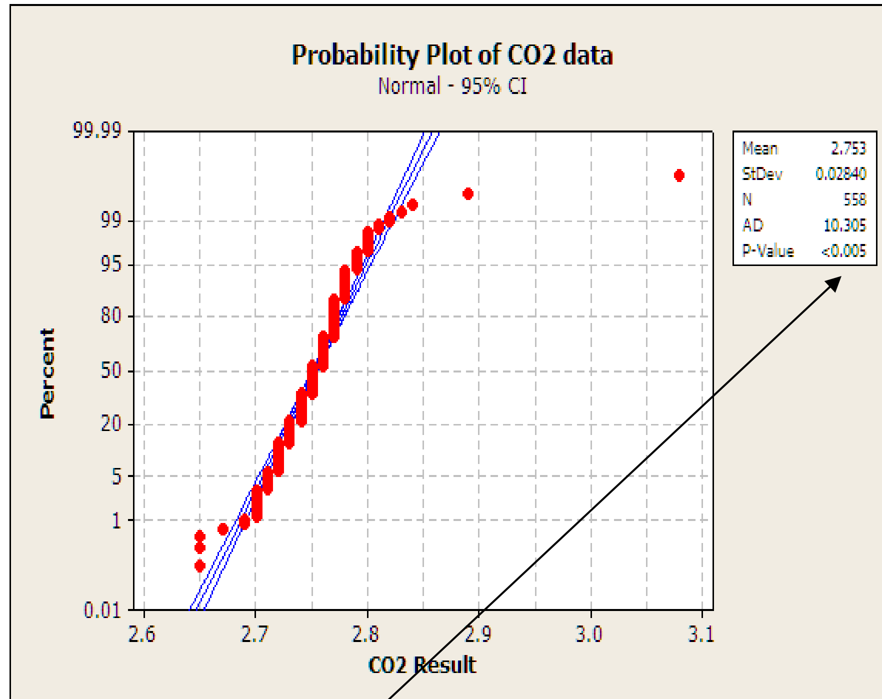


Rule: if p-val < then assume

Testing Normality

Watch out for Outliers !

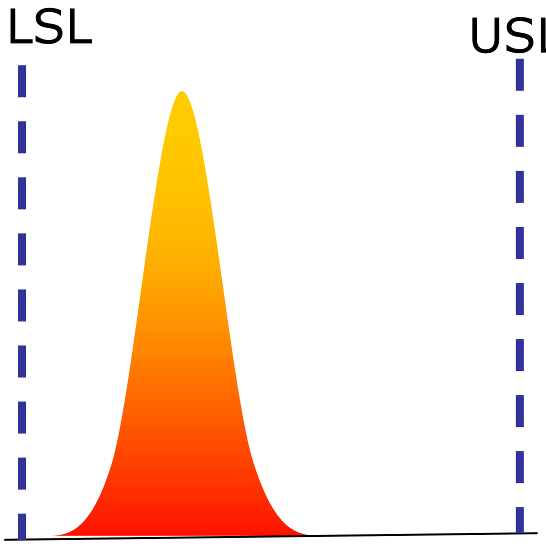




Data looks to be normally distributed with some outliers, and this most likely influenced the AD test statistic.

What are Good Ppk/Cpks

- If Ppk> 1.33 then it can be demonstrated that the process average is 4 standard deviations (long term) away to the nearest specification



- Ratio Ppk/Cpk:

One-Sided Specifications

- Common to have only one specification in brewing (Eg SO2, VDKs, TPOs)
- If a specification does not exist then do not calculate the one side index and assume it is very large (∞)

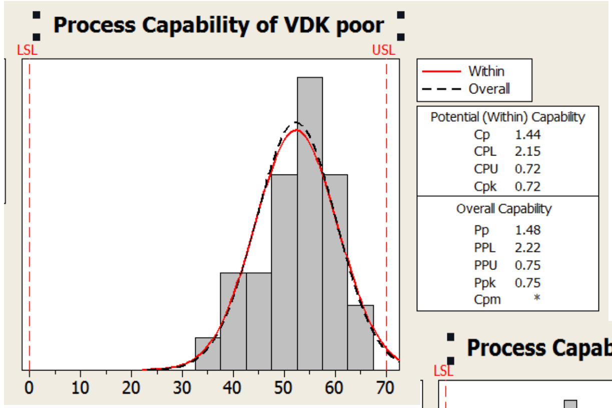
$$\infty$$

$$\hat{P}_{pl} = \left(\frac{\overline{X} - LSL}{3S} \right) \qquad \hat{P}_{pu} = \left(\frac{USL - \overline{X}}{3S} \right)$$

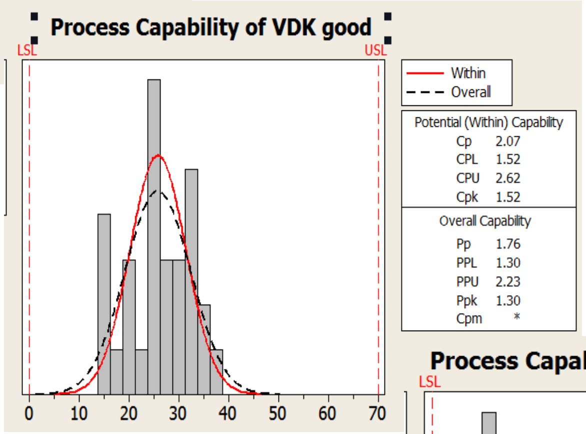
$$\hat{P}pk = \min\left\{ \infty, \hat{P}_{pu} \right\} = \hat{P}_{pu}$$

Minitab exercise: Set LSL = 0

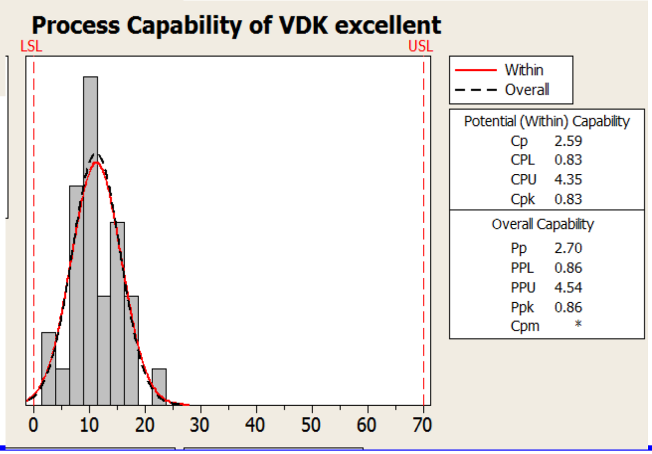
Example One Sided Specifications



Ppk = 0.75



Ppk = 1.33



Ppk = 0.86

Sample Size Requirements

- **Guirguis & Rodriguez (1992) derived exact lower confidence limits for Cpk. An exact 100 γ % lower confidence limit for Ppk is the solution for C_K that satisfies the solution to the following**

$$\gamma = Q\left(3\hat{C}_{pl}\sqrt{n},\, 3C_K\sqrt{n},\, \frac{6C_K\sqrt{n-1}}{3(\hat{C}_{pl}+\hat{C}_{pu})},\, \infty\right) - Q\left(-3\hat{C}_{pu}\sqrt{n},\, -3C_K\sqrt{n},\, \frac{6C_K\sqrt{n-1}}{3(\hat{C}_{pl}+\hat{C}_{pu})},\, \infty\right)$$

$$Q(t,\delta,a,b) = \frac{1}{\Gamma\left(\frac{n-1}{2}\right)2^{\frac{n-3}{2}}}\int_a^b \Phi\left(\frac{tx}{\sqrt{n-1}}-\delta\right)x^{n-2}e^{-\frac{x^2}{2}}\,dx$$

where $\Phi(*)$ *denotes the cumulative normal distr. function.*

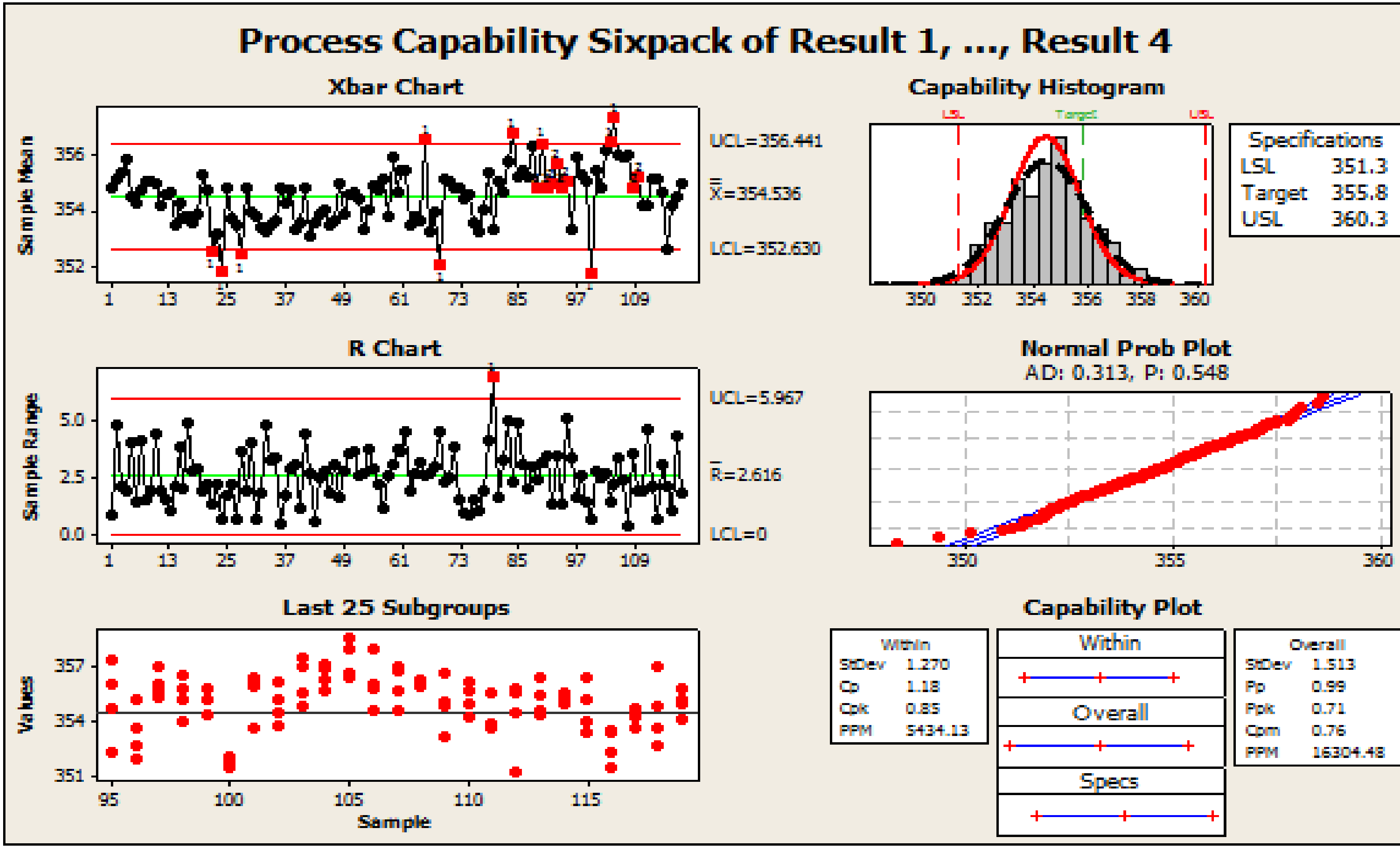
Improving Process Capability

- Examples

- Statistical Tools to Understand Sources of Variation

  - t-tests for mean centering
  - ANOVA
  - Correlation & Regression
  - Setting Functional Limits
  - I/O QFRs

Example 1: Unstable Process



Example 1: Unstable Process

- In this situation, the process is being influenced by special cause variation

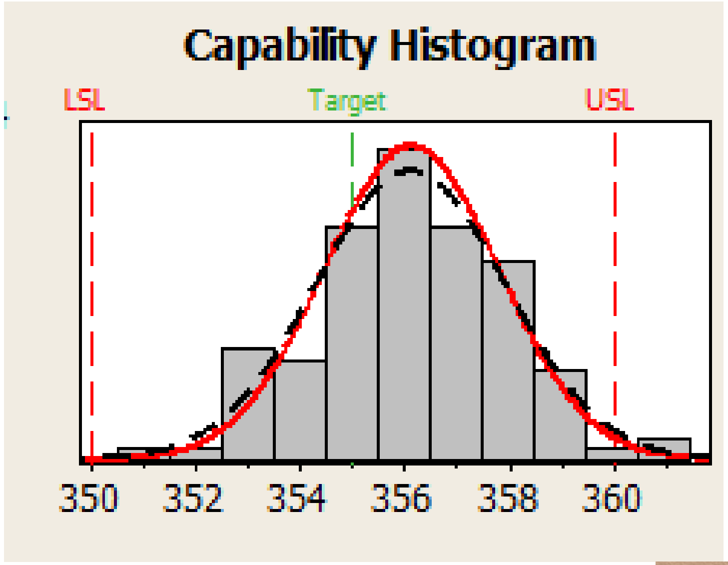
- Because of this determining any process capability indices or attempting to adjust the mean is meaningless
- $$\hat{P}_{pl} \downarrow = \left(\frac{\overline{X} - LSL}{3S \uparrow} \right) \qquad \hat{P}_{pu} \downarrow = \left(\frac{USL - \overline{X}}{3S \uparrow} \right)$$

$$\hat{P}pk = \min\left\{ \hat{P}_{pl}, \hat{P}_{pu} \right\} =$$

- You should always first understand what are the root causes of this excessive variation is and eliminate out

Example 2: Off Target

- It does appear the process is off center; however, do we know if this is statistically significant?



- Use Gosset's t-test

Example 2: Off Target

- t-Statistic

$$t^* = \left(\frac{\bar{\bar{X}} - \mu_0}{S / \sqrt{n}} \right)$$

- μ_0 = Target value we desire

- Software p-value

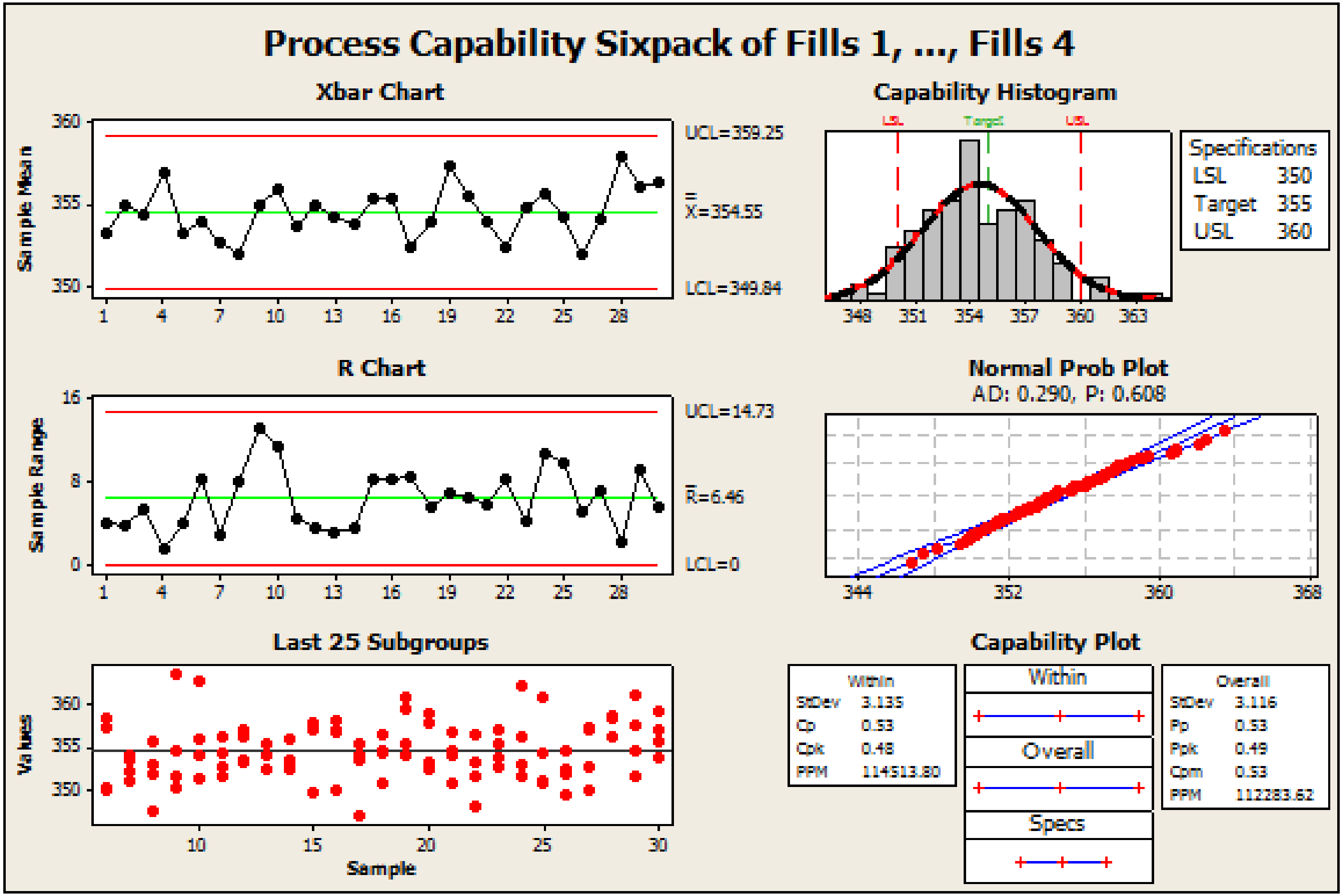
Example 2: Off Target

- μ_0 = 355.00 mls
- S = 1.838 mls. $\bar{\bar{X}} = 356.15 mls$
- n=120
- $$t^* = \left(\frac{\bar{\bar{X}} - \mu_0}{S / \sqrt{n}} \right) = \left(\frac{356.15 - 355.00}{1.838 / \sqrt{120}} \right) =$$
- Excel QI Macro Demo
- Δ: $\Delta_{shift} = \bar{\bar{X}} - \mu_0 = 356.15 - 355.00 =$
- Software p-value

Example 2: Off Target

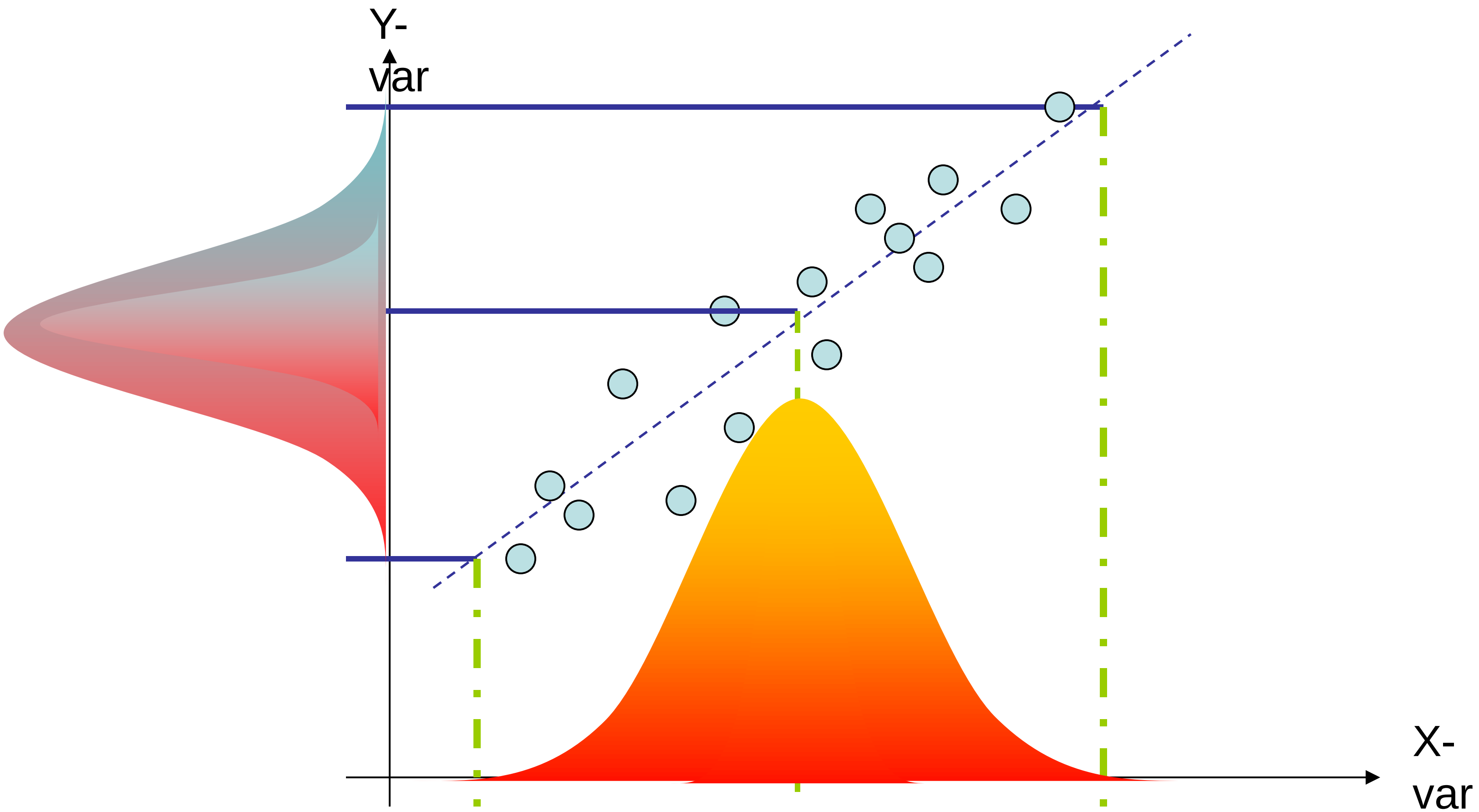
- μ_0 = 355.00 mls
- S = 1.838 mls. $\bar{\bar{X}} = 356.15 mls$
- n=120
- $$t^* = \left(\frac{\bar{\bar{X}} - \mu_0}{S / \sqrt{n}} \right) = \left(\frac{356.15 - 355.00}{1.838 / \sqrt{120}} \right) = 6.85$$
- Excel QI Macro Demo
- Δ: $\Delta_{shift} = \bar{\bar{X}} - \mu_0 = 356.15 - 355.00 = 1.15$
- Software p-value

Example 3: High Variability



Example 3: High Variability

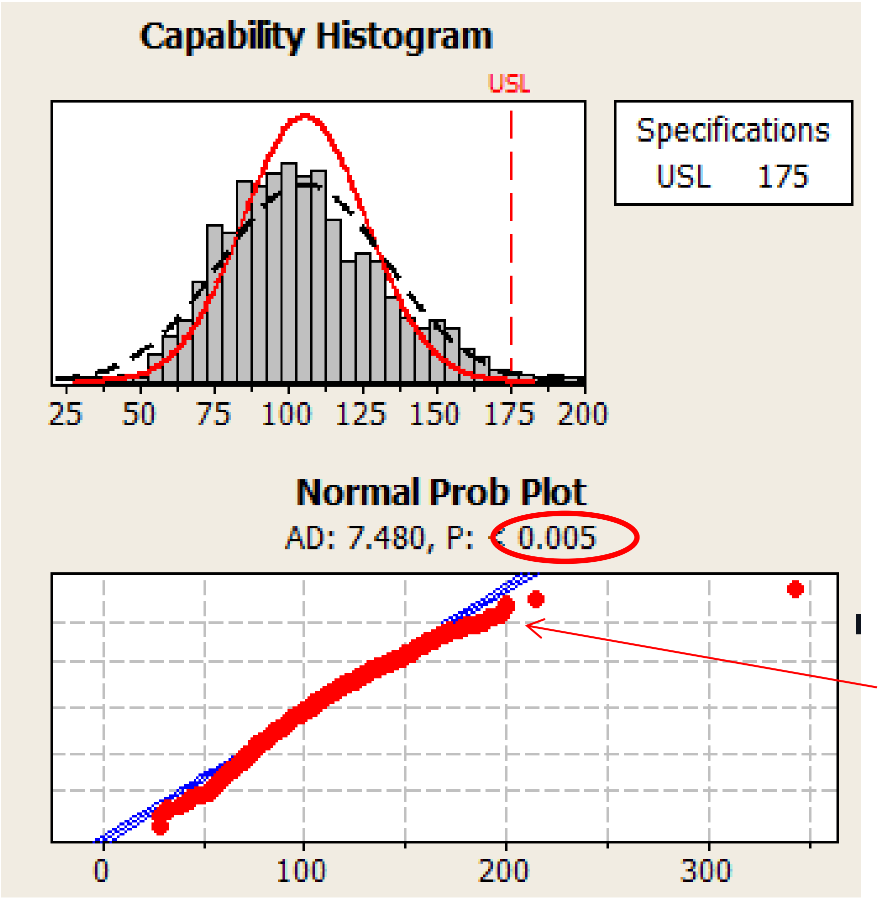
- Critical X-var Error Propagation



Example 3: High Variability

- X-variables that exhibit high degree of variability and known to influence the output of the process need to be controlled tighter:
  a) PID Controls
  b) SOPs
  c) Raw material specifications
  d) consolidation of suppliers

Example 4: Valve to Valve Variability



Upper tail appears to have humps (multi-modal) and a skew to the right side of the distribution

Example 4: Valve to Valve Variability

- When multiple process streams feed one general process, tools such as One Way **AN**alysis **O**f **Va**riance (**ANOVA**) can be used to data mine if certain streams are sources of variability

  Eg: Fermentation Vessel Type, Capper Elements, Filler Valves, Seamer Heads

$$H_0 : \mu_i = \mu_j, \quad i \neq j$$

$$H_A : at\ least\ one\ \mu\ differs$$

- Individual Plots by Head
- Example Using Minitab *TPO by Valve.MPJ*

Example 4: ANOVA

- F-Test

$$F^* = \frac{Variance\ between\ treatments}{Variance\ within\ treatments}$$

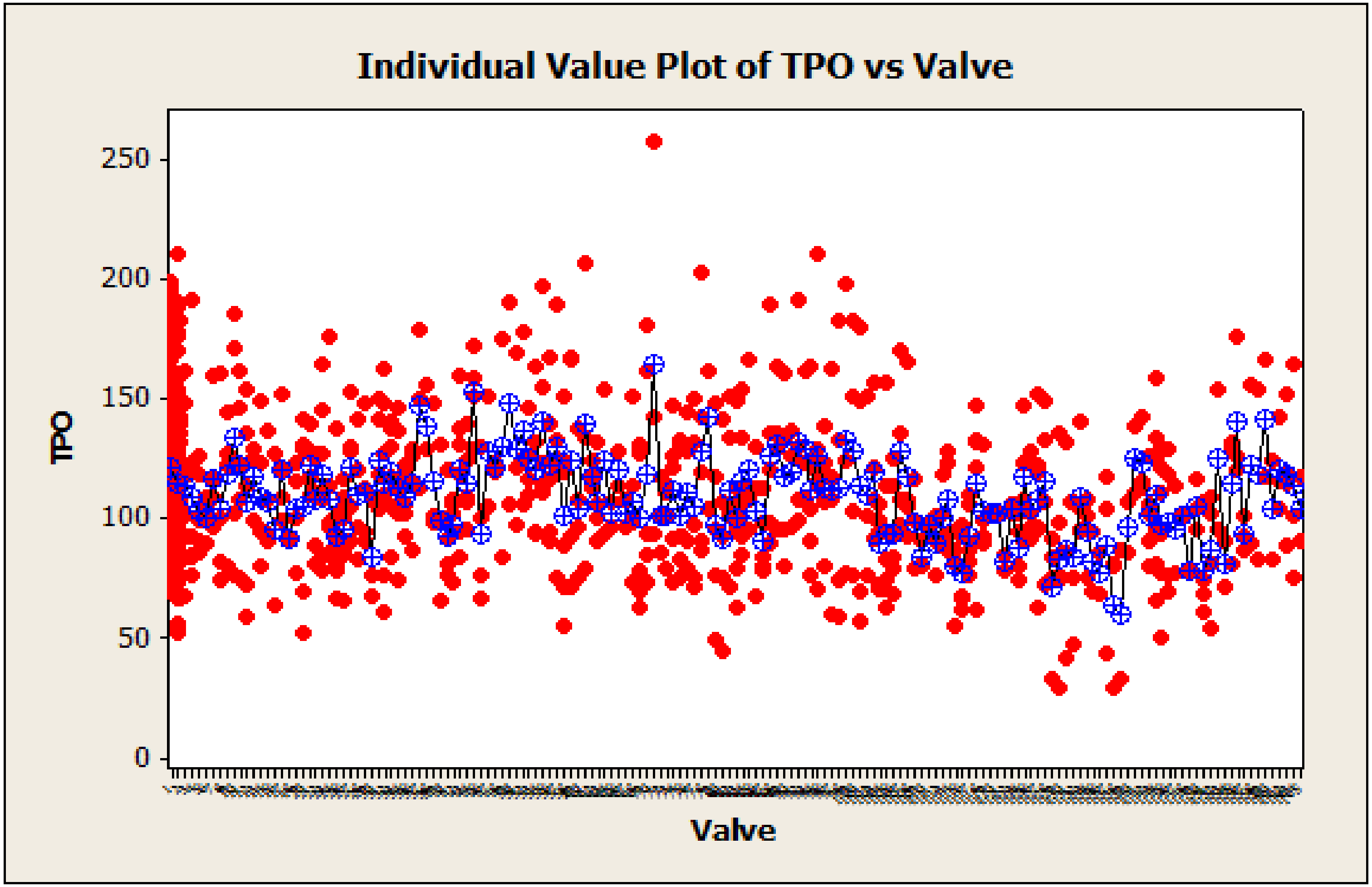
$$= \frac{\dfrac{\sum_{j}^{J}\left(\overline{X}_j - \overline{\overline{X}}\right)^2}{J-1}}{\sum_{j}^{J}\sum_{i}^{n_j}\dfrac{\left(X_{ij} - \overline{X}_j\right)^2}{J(n_j-1)}}$$

- p-value

One-way ANOVA: TPO versus Valve

Source	DF	SS	MS	F	P
Valve	164	184390	1124	2.33	0.0001
Error	1382	667131	483		
Total	855	851521			

Example 4: Valve to Valve Variability



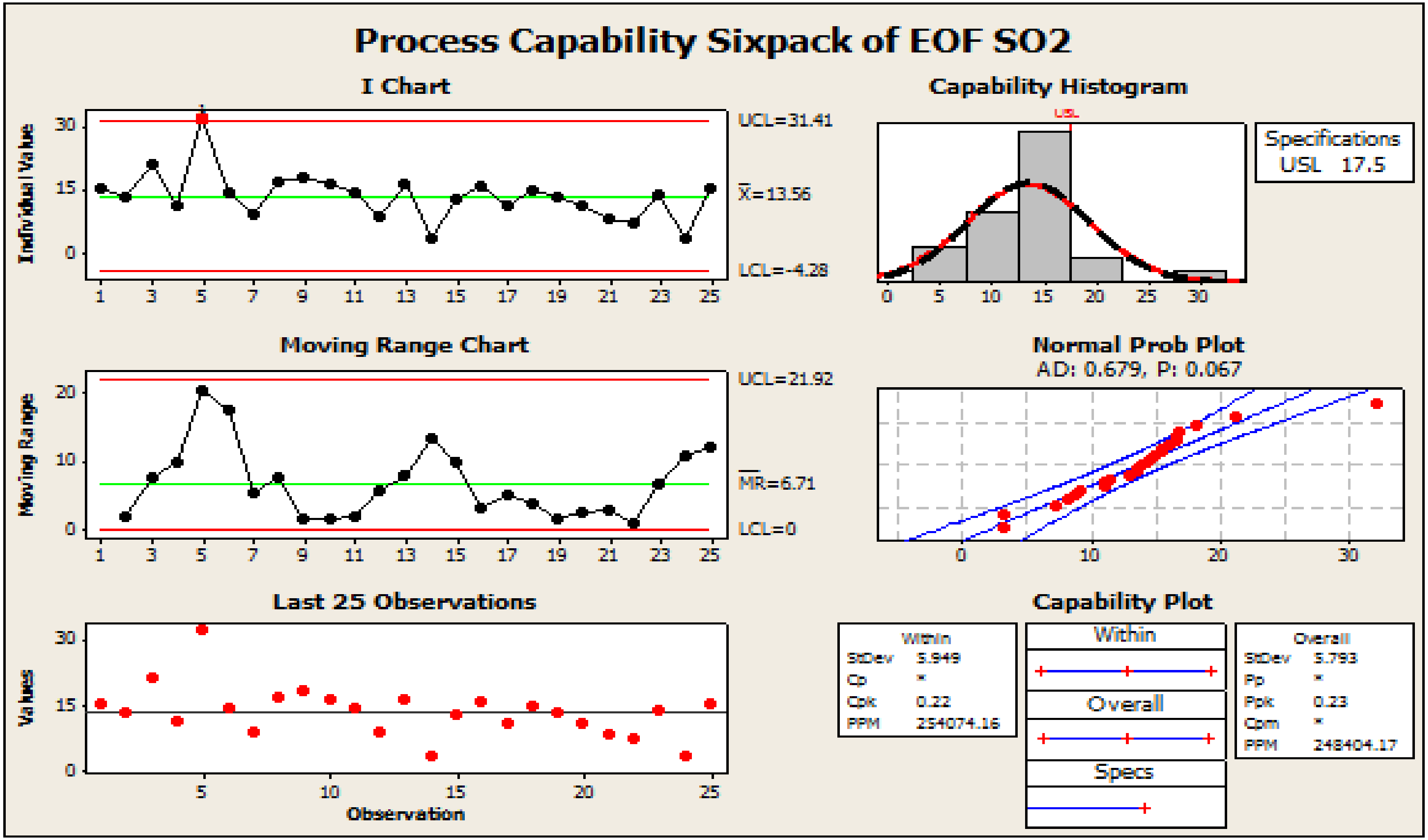
Worst Valves:
71
45
50
37

Best Valves:
139
138
129
116

Example 5: Y=F(X)

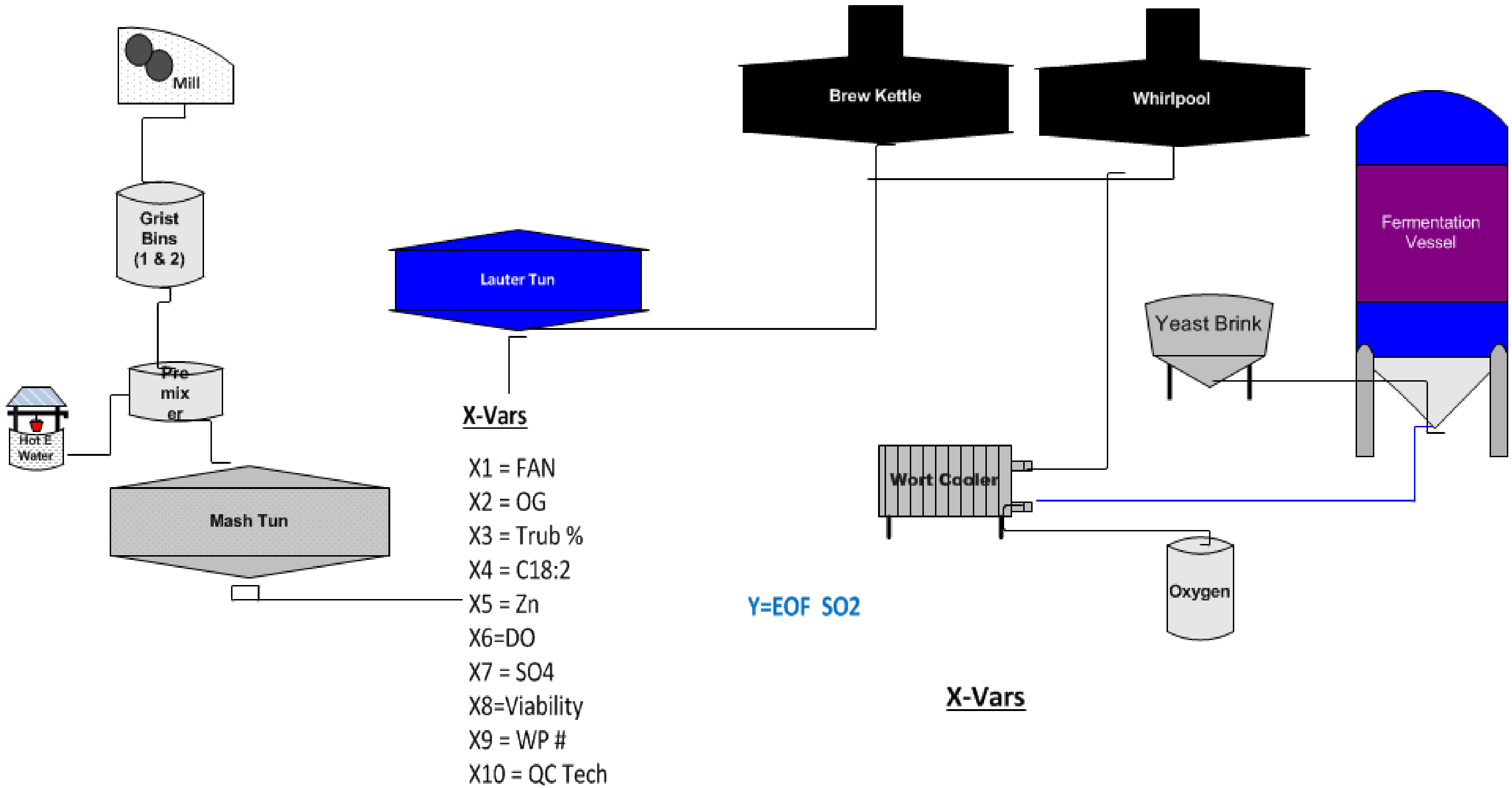
- EOF SO_2
- Preliminary Capability Study on Y
- Collect data on suspect X's
- Study Correlations
- Optimization
- Setting Functional Limits
- Gage Errors

Example 5: Y=F(X)



Example 5: Y = F(X)

- What is required to get this process capability improved?



Example 5: Y = F(X)

WP	Zn	DO	Trub	FAN	Viability	SO4	OG	C18:2	QC Tech	EOF SO2
A1	170.78	14.52	1.37	199.26	87.48	86.91	15.95	2.50	Betty Jean	15.35
A1	146.35	16.40	2.09	229.50	91.85	75.04	15.90	1.98	Betty Jean	13.55
A1	123.97	15.44	1.10	222.52	84.88	70.23	15.99	1.96	Betty Jean	21.21
A1	168.59	12.30	0.73	222.21	92.93	75.01	16.02	3.66	Betty Jean	11.50
A1	151.68	9.93	0.04	228.00	86.12	75.37	16.06	0.84	Betty Jean	31.98
A1	149.28	17.32	1.18	213.54	87.75	68.44	16.16	2.20	Betty Jean	14.37
B1	145.40	18.00	3.28	243.02	89.13	81.81	16.01	1.97	Betty Jean	9.07
A1	164.73	11.74	1.70	219.52	87.87	79.68	15.98	2.52	Betty Jean	16.73
A1	163.97	16.19	0.98	225.87	92.62	88.21	15.94	1.86	Betty Jean	18.15
A1	166.13	14.30	1.20	217.99	89.60	93.41	16.09	0.86	Tim Wood	16.50
A1	130.32	19.62	0.70	226.41	86.31	72.37	16.04	1.72	Betty Jean	14.53
B1	161.30	17.19	1.23	220.58	88.11	78.44	15.92	1.69	Betty Jean	8.72
B1	175.50	14.13	1.26	217.68	91.73	82.56	15.85	0.84	Betty Jean	16.55
B1	156.73	17.28	1.73	238.27	92.11	80.53	15.83	2.41	Tim Wood	3.20
B1	165.19	12.50	1.74	206.03	86.70	79.14	16.18	1.84	Betty Jean	12.97
B1	150.92	15.07	1.30	210.54	89.08	79.80	15.88	0.75	Tim Wood	16.06
B1	168.88	15.86	1.98	224.46	88.11	82.13	16.26	1.32	Tim Wood	11.07
B1	156.17	13.36	1.97	222.67	88.17	75.11	15.89	1.89	Tim Wood	14.86
B2	164.59	14.43	1.63	206.19	91.50	76.85	16.00	1.98	Betty Jean	13.45
B2	138.29	14.88	1.43	230.53	87.68	75.36	15.89	2.22	Tim Wood	11.08
B2	171.97	17.91	2.96	201.94	84.90	71.72	16.02	1.42	Tim Wood	8.14
B2	178.31	15.44	2.64	206.19	87.32	77.02	15.81	2.41	Tim Wood	7.32
B2	129.61	13.79	1.18	219.57	88.82	71.05	15.97	2.37	Tim Wood	13.99
B2	148.75	15.21	2.92	218.03	90.09	75.11	16.01	3.23	Tim Wood	3.26
B2	150.57	13.36	1.36	201.45	90.09	76.82	15.94	1.54	Tim Wood	15.48

Example: Y = F(X)

- Correlation Analysis using Software
- Graph Y (SO2) versus **Continuous** X-vars
- Correlation Coefficient

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)\left(\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}\right)}}$$

- Test Statistic

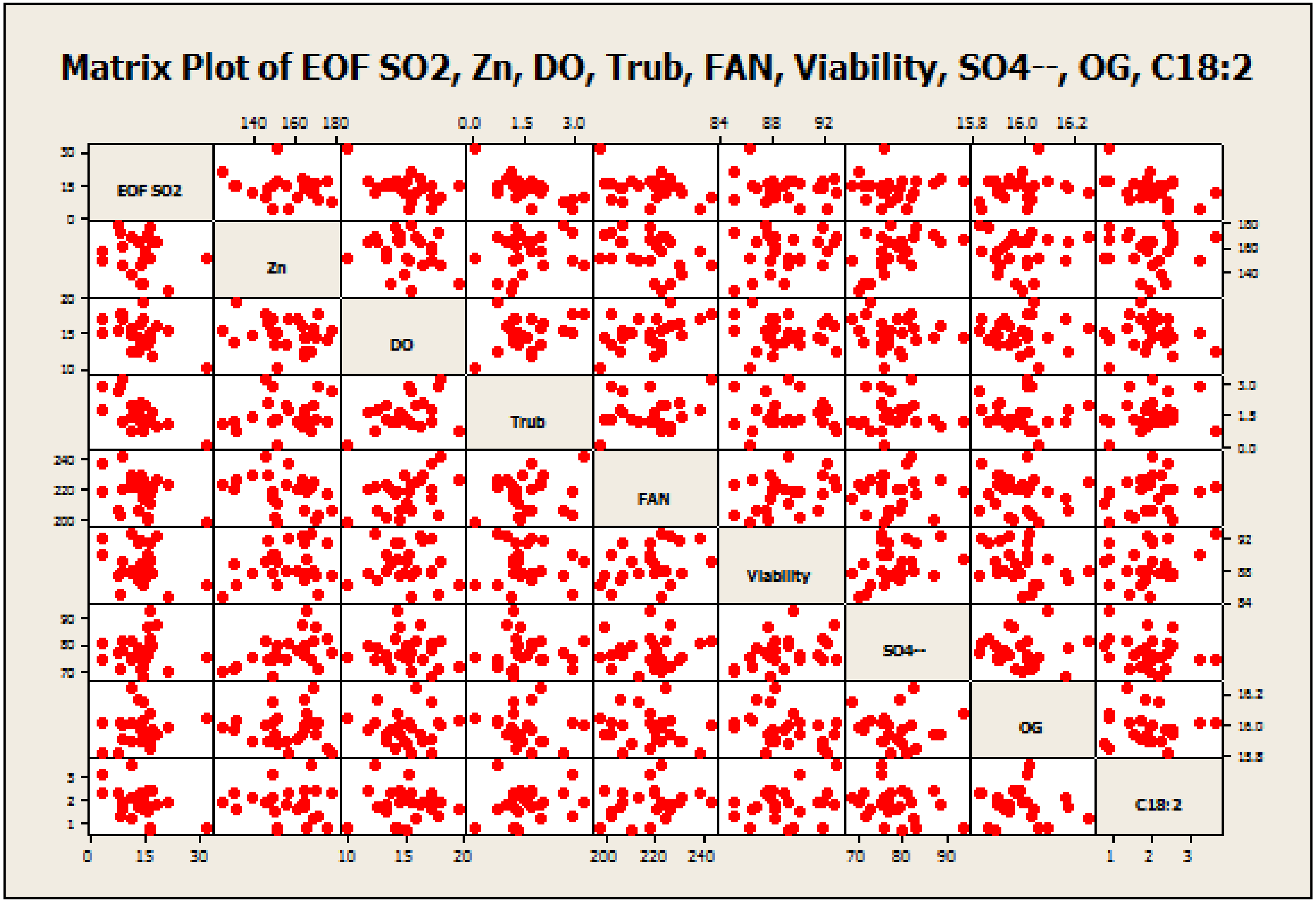
$$T = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$$

Example 5: Y = F(X)

- Exercises: QI Macros Correlation

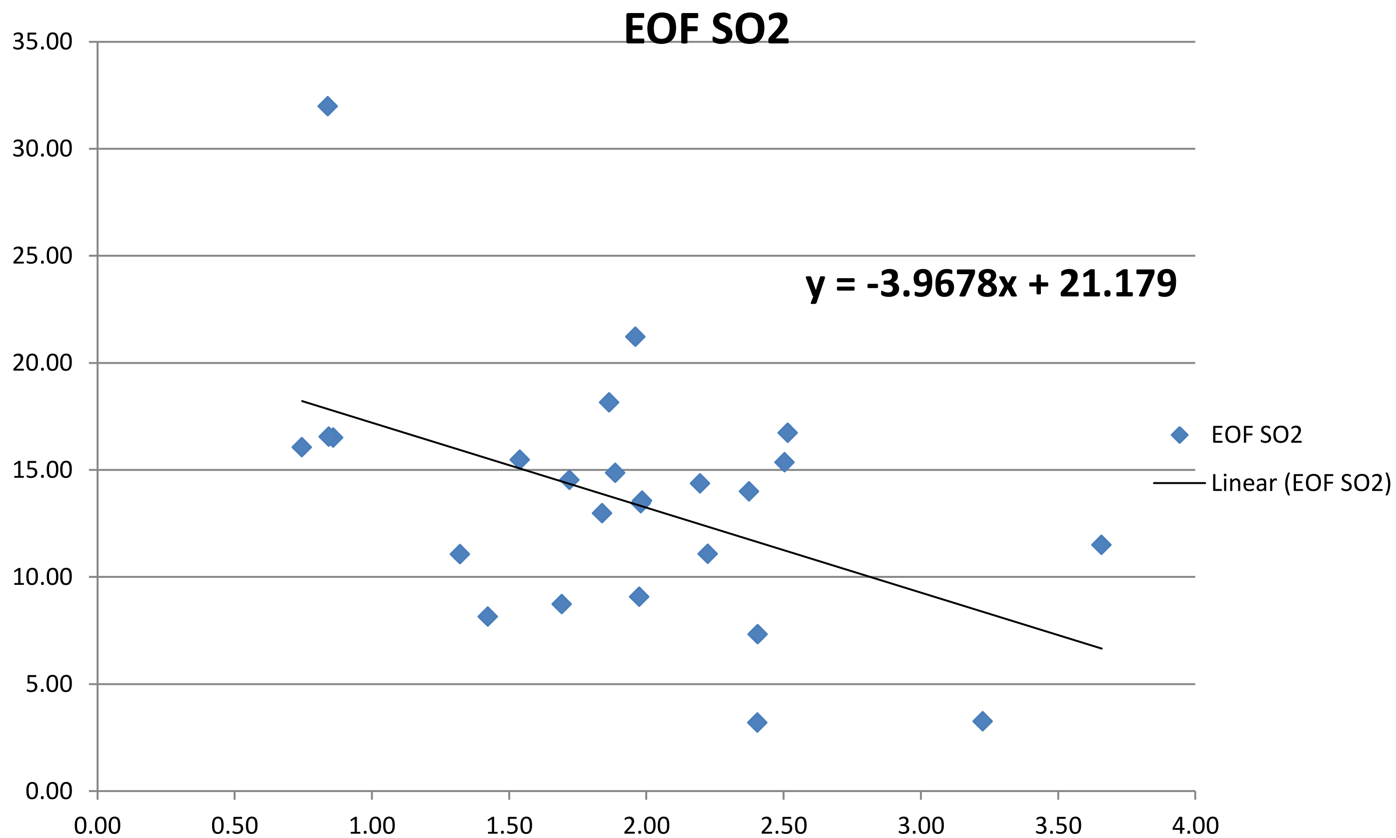
CORREL	Zn	DO	Trub	FAN	Viability	SO4	OG	C18:2	EOF SO2
Zn	1.000	-0.208	0.201	-0.418	0.239	0.525	-0.001	-0.063	-0.180
DO	-0.208	1.000	0.367	0.218	-0.052	-0.100	-0.046	-0.014	-0.530
Trub	0.201	0.367	1.000	-0.029	-0.038	-0.008	-0.077	0.214	-0.697
FAN	-0.418	0.218	-0.029	1.000	0.233	0.025	-0.079	0.069	-0.023
Viability	0.239	-0.052	-0.038	0.233	1.000	0.367	-0.314	0.231	-0.234
SO4	0.525	-0.100	-0.008	0.025	0.367	1.000	-0.037	-0.284	0.035
OG	-0.001	-0.046	-0.077	-0.079	-0.314	-0.037	1.000	-0.092	0.165
C18:2	-0.063	-0.014	0.214	0.069	0.231	-0.284	-0.092	1.000	-0.484
EOF SO2	-0.180	-0.530	-0.697	-0.023	-0.234	0.035	0.165	-0.484	1.000
p Values	Zn	DO	Trub	FAN	Viability	SO4	OG	C18:2	EOF SO2
DO	0.318		0.071	0.295	0.805	0.633	0.826	0.946	0.006
Trub	0.334	0.071		0.889	0.858	0.971	0.713	0.304	0.000
FAN	0.037	0.295	0.889		0.263	0.906	0.707	0.744	0.913
Viability	0.250	0.805	0.858	0.263		0.071	0.127	0.266	0.259
SO4	0.007	0.633	0.971	0.906	0.071		0.861	0.169	0.870
OG	0.996	0.826	0.713	0.707	0.127	0.861		0.661	0.430
C18:2	0.765	0.946	0.304	0.744	0.266	0.169	0.661		0.014
EOF SO2	0.389	0.006	0.000	0.913	0.259	0.870	0.430	0.014	

Example 5: Y = F(X)

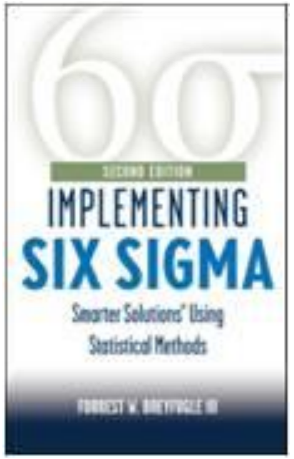


Example 5: Y = F(X)

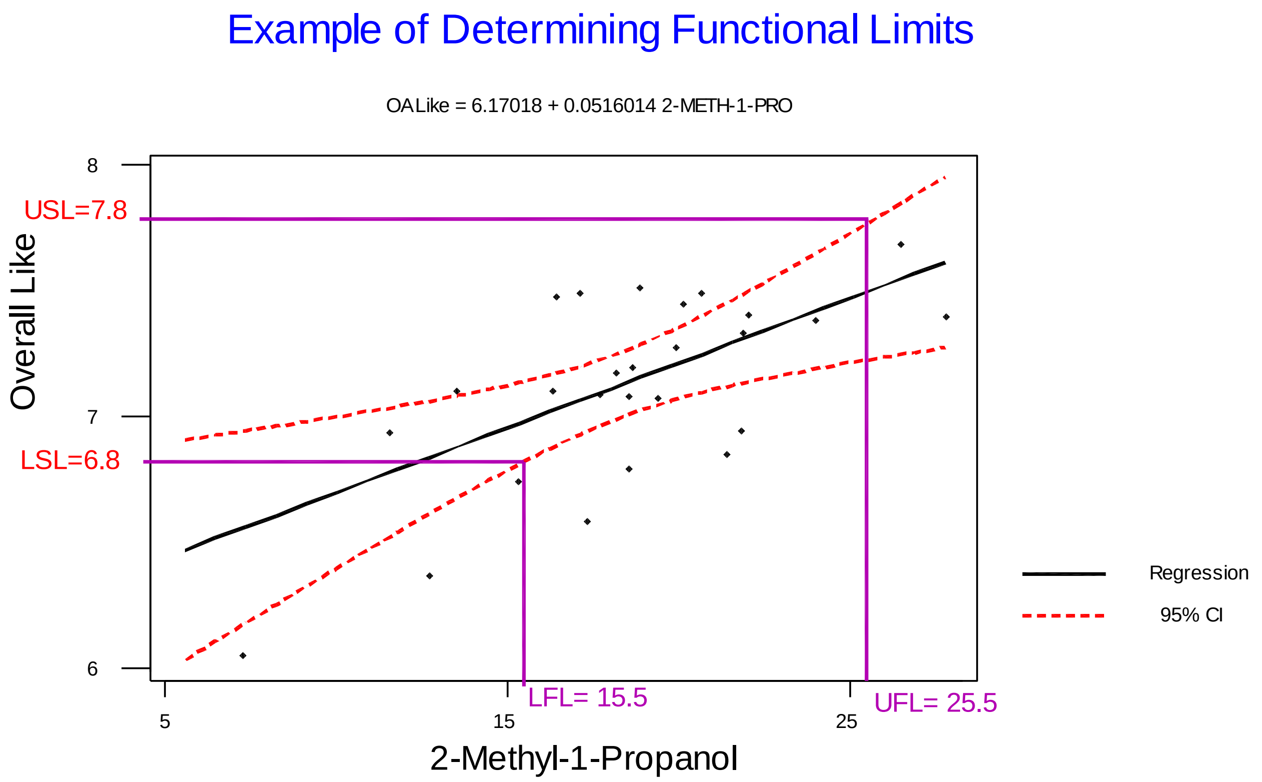
- SO2 vs C18:2



Example 5: Y = F(X)



- Breyfoggle's Approach



Setting Functional Limits

- Simple Linear regression models yield a basic relationship

  y = mx + b

- If we know the tolerances for y, we can solve for x. For example

$$if \quad y = mx + b$$

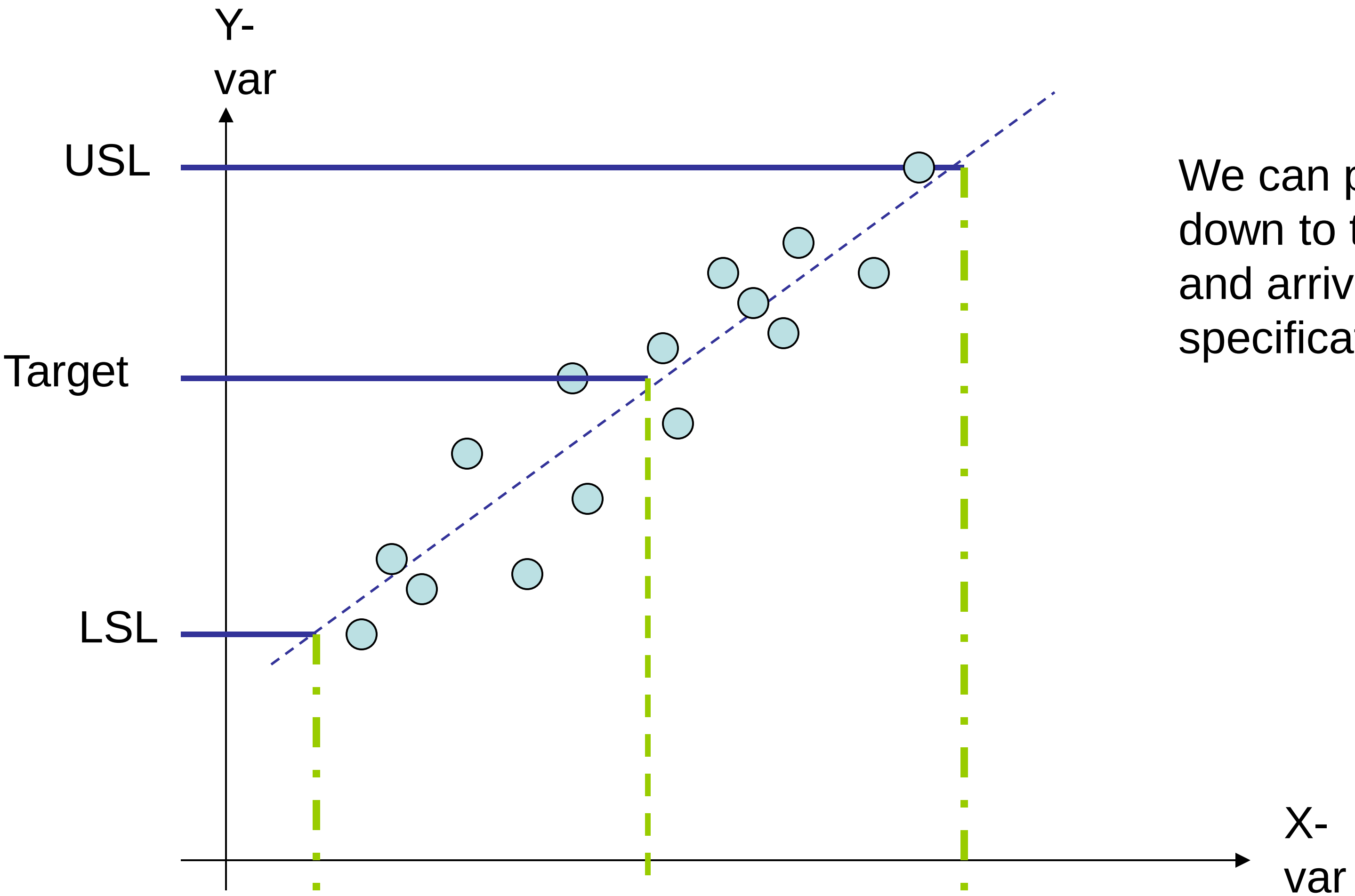
$$\Rightarrow x = \frac{(y-b)}{m}$$

$$\Rightarrow \tau_x = \left(\frac{\tau_y - b}{\beta_x}\right)$$

Key Point:
If we plug in the USL_Y we can get a specification for X

If we have a target for Y then we can derive a target for X

Setting Functional Limits



We can project down to the x-axis and arrive at specifications

Example 5: Y = F(X)

- Mathematical Approach

$$\sigma_x \leq \frac{\left(USL_y - \tau_y\right)}{4\hat{\beta}_x}$$

- Tolerances: $\tau_x \pm 3\sigma_x$

Example 5: Y = F(X)

- **Exercise**

  $USL_Y = 17.5$

  $\tau_Y = 10.0$

  $\beta_x = -3.97$

  $b = 21.18$

$$\tau_x = \left(\frac{\tau_y - b}{\beta_x} \right) =$$

Determine Tolerances:

$$\sigma_x \leq \frac{\left(USL_y - \tau_y\right)}{4\left|\hat{\beta}_x\right|} =$$

$$\tau_x \pm 3\sigma_x$$

Example 5: Y = F(X)

- **Exercise**

  $USL_Y = 17.5$

  $\tau_Y = 10.0$

  $\beta_x = -3.97$

  $b = 21.18$

$$\tau_x = \left(\frac{\tau_y - b}{\beta_x} \right) = \frac{10 - 21.18}{-3.97} = 2.816$$

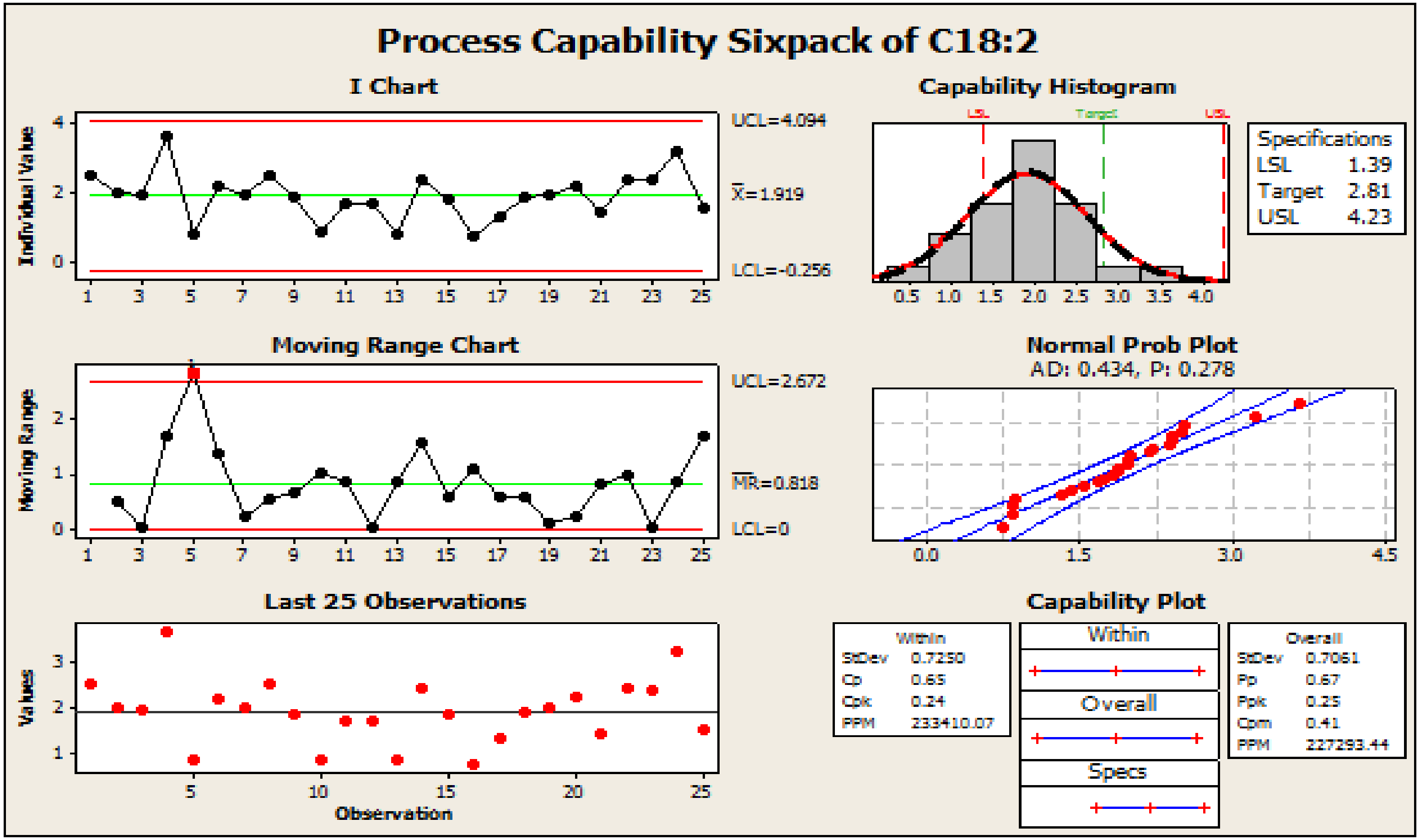
Determine Tolerances:

$$\sigma_x \leq \frac{\left(USL_y - \tau_y\right)}{4\left|\hat{\beta}_x\right|} = \left(\frac{17.5 - 10}{4 * 3.97} \right) = 0.472$$

$$\tau_x \pm 3\sigma_x = 2.81 \pm 1.42 = 1.39, 4.23$$

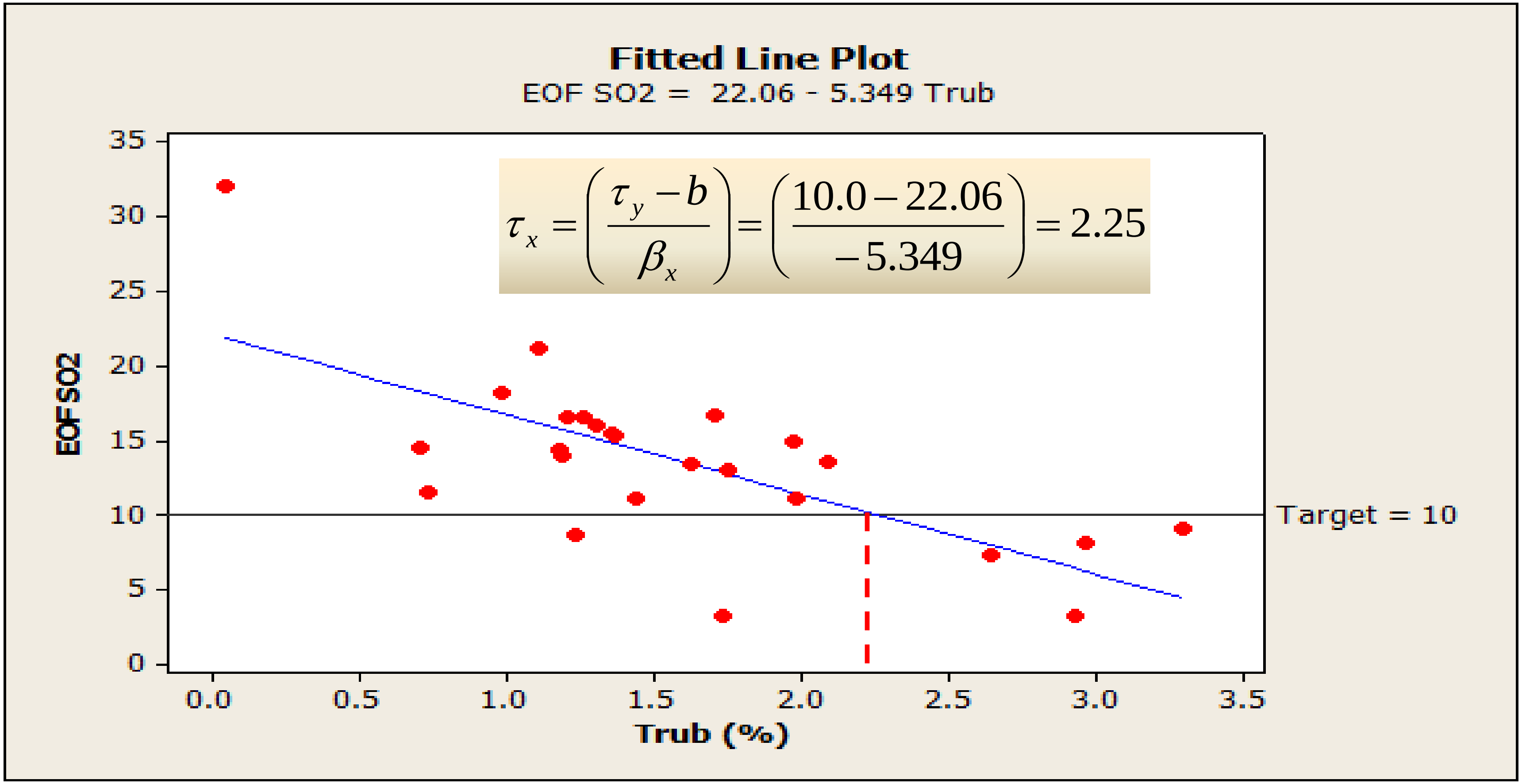
Example 5: Y = F(X)

- Review Process Capability of Critical X-vars



Example 5: Y = F(X)

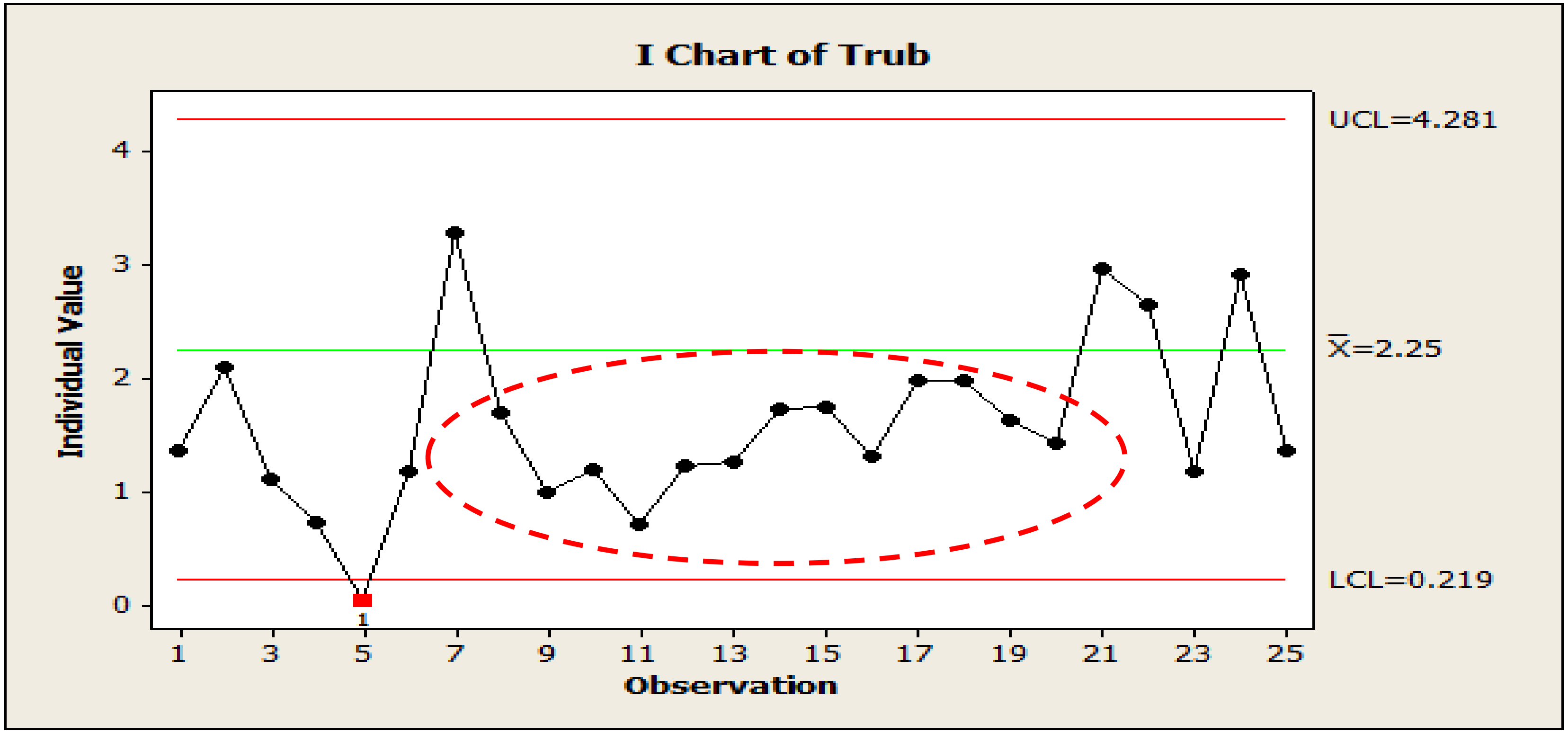
- Trub Functional Limits



$$\tau_x = \left(\frac{\tau_y - b}{\beta_x}\right) = \left(\frac{10.0 - 22.06}{-5.349}\right) = 2.25$$

Example 5: Y = F(X)

- X-variable SPC
- Center Line set as Target



Example 5: Y = F(X)

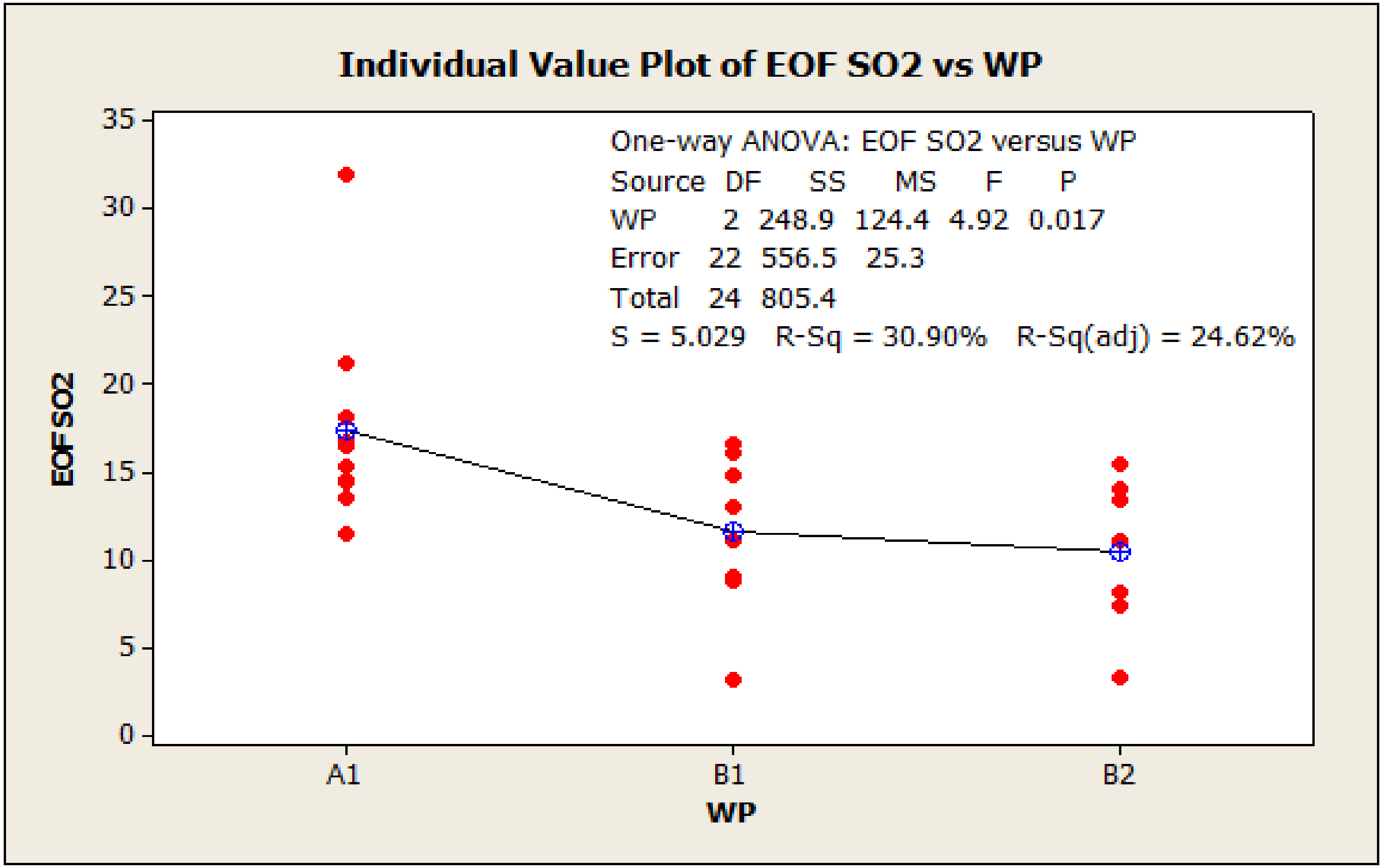
- Two categorical variables
  Whirlpool   &   QC Tech

- To understand if there are potential relationships we apply _______________

- Minitab exercise

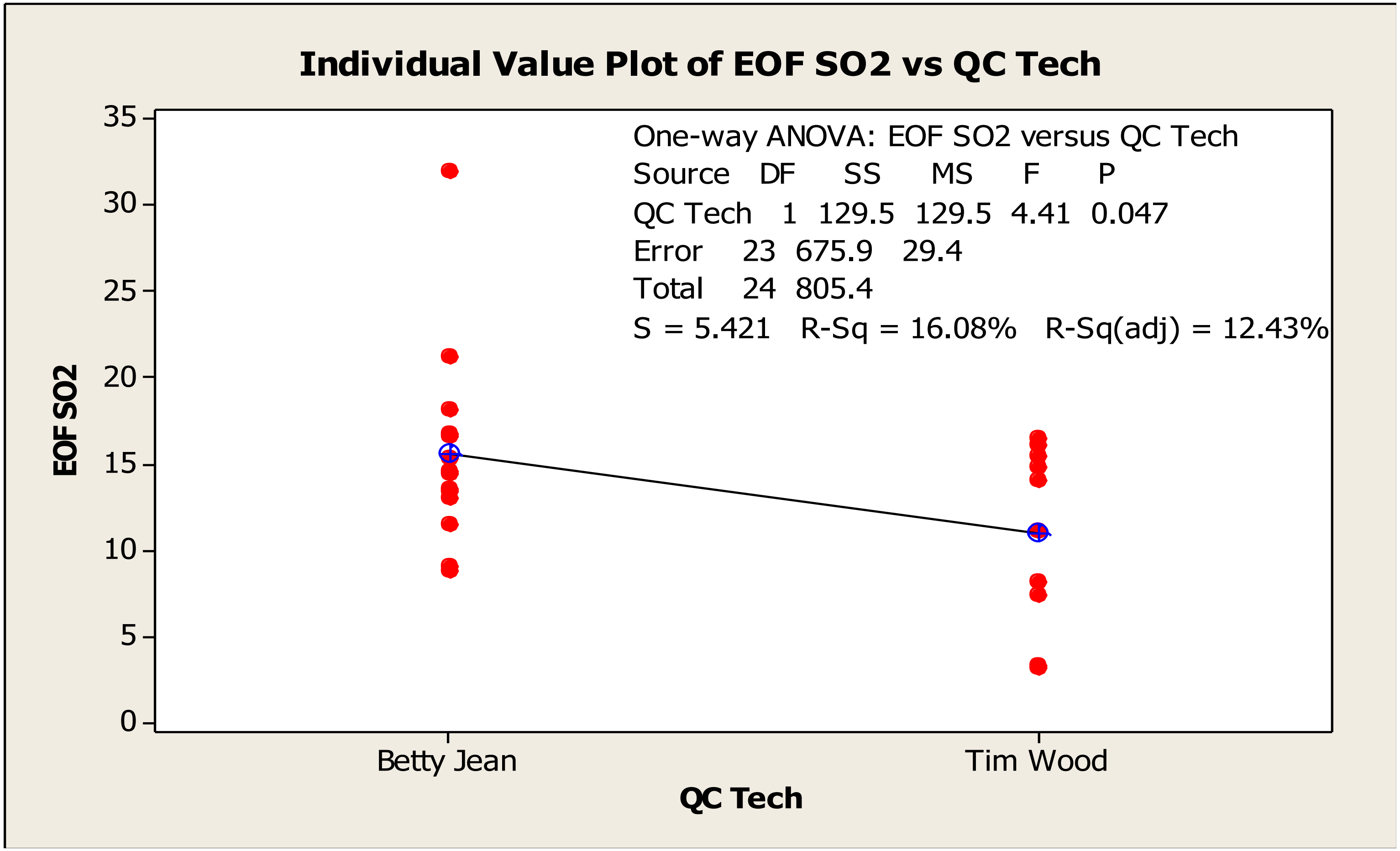
Example 5: Y = F(X)

- Whirlpool Effect



Example 5: Y = F(X)

Technician Effect



Input / Output Matrix - Example

		Process Outputs							
		Boilers	Fill heights	TPO	CO_2	Damaged/ missing crowns	Torques	Crimps	Rejects
Process Inputs	**Jetter pressure/stream**	x	x	**X**	x				x
	Jetter position	x	x	**X**	x				x
	Crown chute twist / vibrator					x			x
	Lift cylinder lube								x
	Lift Cylinder Pressure	x	x	**X**	x				x
	Air control pressure								x
	Filler speed	x	x	**X**	x				x
	C02 counter pressure	x	x	**X**	x				x
	Crown air jets					x			x
	Beer temp.	x	x	**X**	x				x
	Purge Cam	x	x	**X**	x				x
	Spreader Rubber	x	x						x
	Bowl levels	x	x						x
	$C0_2$ volume	x	x	**X**	x				x
	Center Cup Seal	x	x						x
	Inches of vacuum	x	x	**X**	x				x
	Valve performance	x	x						x
	Beer leaks			**X**					
	Beer Pressure	x	x	**X**					x
	Crown extractor position					x			x
	Crowner height						x	x	x

Input / Output Matrix- Exampe

PROCESS INPUT MONITORING SHEET (PIMS)

INPUTS	Standard / [Trigger]	Frequency	Impacted Outputs
Jetter pressure	**CQS 0-211-348 & 0-211-385** 12oz = 80-100 psi [<80 or >100] 22oz = 50-60psi [<50 or >60]	During samples for fills/TPO (every 4 hours)	Boilers, Fill Height, TPO, CO2, Rejects
Jetter stream color	Clear [Non clear] Straight stream [Fanning or stream at angle]	During samples for fills/TPO (every 4 hours)	Boilers, Fill Height, TPO, CO2, Rejects
Jetter position	**CQS 0-211-385** Centered side to side & vertically above bottle. No fanning of jetter spray [TPO out of spec.; Position incorrect]	During samples for fills/TPO (every 4 hours)	Boilers, Fill Height, TPO, CO2, Rejects
Crown chute twist / vibrator	Zero jammed crowns; 100% proper orientation [1 jammed crown; improper orientation]	Once/shift; After changeover	Damaged-missing crowns, Rejects
Lift cylinder lube	Oil/lube present in plastic sight glass/tube [Minimal oil present]	Start up	Rejects
Lift Cylinder Pressure	40-60 psi [<40 or >60 psi]	Once/shift; After c/o; After maint. day	Boilers, Fill Height, TPO, CO2, Rejects
Air control pressure	60 - 70 psi [<60 or >70psi]	Once/shift; After c/o; After maint. day	Rejects
Filler speed	12oz = 1050 BPM [<1050 or >1050 BPM] / 22oz = 560 BPM [<560 or >560 BPM]	Start-up; After changeover, ongoing observation	Boilers, Fill Height, TPO, CO2, Rejects
CO_2 counter pressure	30 psi +/- 2 psi [<28 psi or >32 psi]	Start-up; After changeover, ongoing observation	Boilers, Fill Height, TPO, CO2, Rejects
Crown air jets	Air jets functioning & holding crowns in place [1 blocked air jet]	During start-up shift; ongoing observation	Damaged-missing crowns, Rejects
Beer temp.	34 to 36 degrees [<32 or >40 degrees]	Start-up; After changeovers; Exception basis	Boilers, Fill Height, TPO, CO2, Rejects
Purge Cam	Full cylinder stroke [Partial cylinder stroke]	Once/shift; After c/o; After maint. day	Boilers, TPO, CO2, Rejects
Spreader Rubber	No tears or damage and set at correct height [Visual inspection reveals damage or being out of position]	Start up; During vent tube change	Boilers, Fill heights, Rejects
Bowl levels	At 55% on HMI - Should be 50% covering sight glass [Auto stops bottle stop at 35%]	Start-up; After changeover, ongoing observation	Boilers, Fill heights, Rejects
CO_2 volume	40-80 scfm [<40 scfm]	Once per shift; exception basis	Boilers, Fill Height, TPO, CO2, Rejects
Center Cup Seal	No damage or foreign objects [Visual inspection reveals damage or foreign objects]	Maintenance day; During vent tube change	Boilers, Fill heights, Rejects
Inches of vacuum	22 inches or higher [lower than 22 inches]	Once per shift	Boilers, Fill Height, TPO, Rejects
Valve performance	Zero boilers [1 boiler / revolution]	Once per shift; exception basis	Boilers, Fill heights, Rejects
Beer leaks	**CQS 0-211-380** Zero leaks [1 leak]		TPO
Beer Pressure	85 psi [< 55 psi - > 95 psi]	Start-up; After changeover, ongoing observation	Boilers, Fill Height, TPO, CO2, Rejects
Crown extractor position	Zero damaged crowns [1 damaged crown / 100]	Once per shift	Damaged-missing crowns, Rejects
Crowner height	Set at 0.125" compensation with 0.125" piece of key stock and bottle with crown [<1.125" or >1.145"]	After crowner height change; Exception basis	Torques, Crimps, Rejects

The Science of Beer

Measurement Systems Analysis

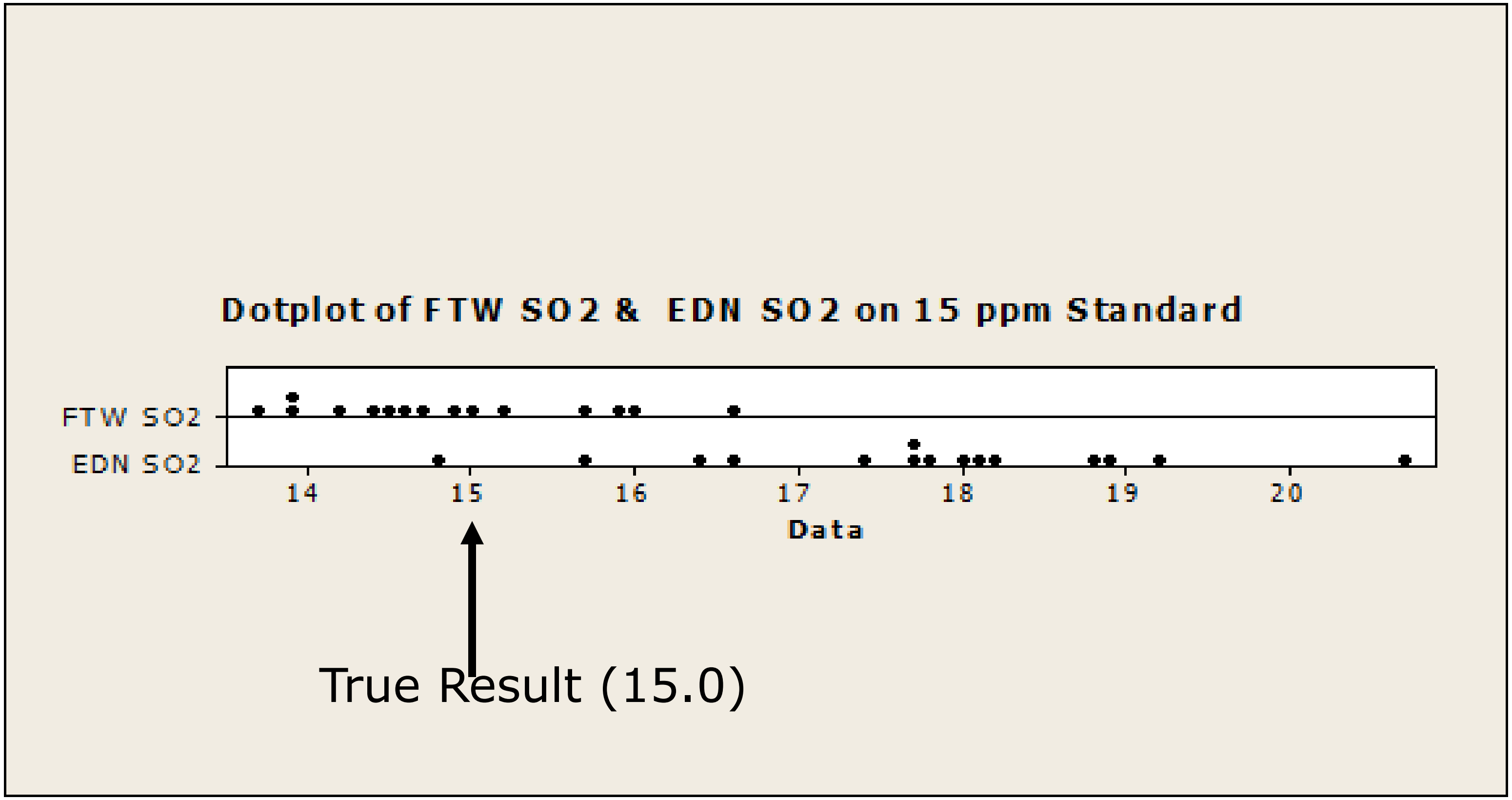
- Gage *Repeatability* and *Reproducibility* Studies

- *Repeatability*:
  That component of gage error that is the direct result of instrument variability

- *Reproducibility*
  That component of gage error that is the direct result of technician to technician differences

Measurement Systems Analysis

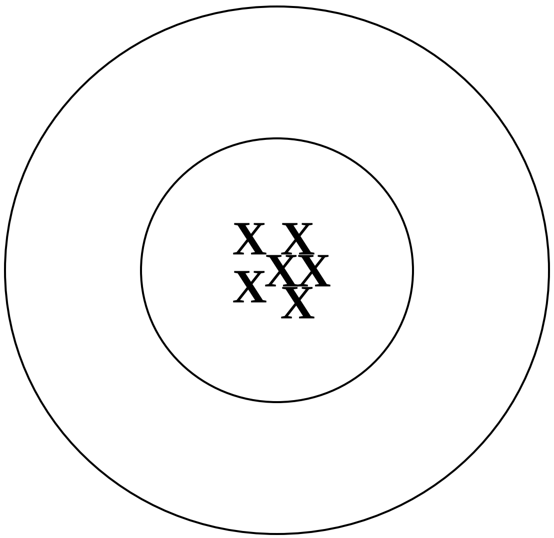
- True Case Study



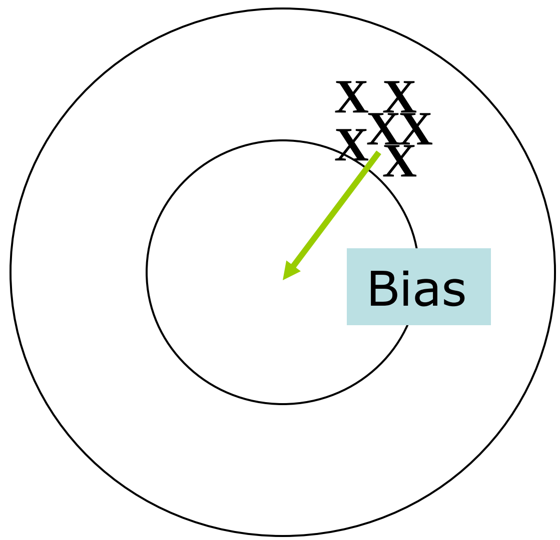
Measurement Systems Analysis

Accuracy

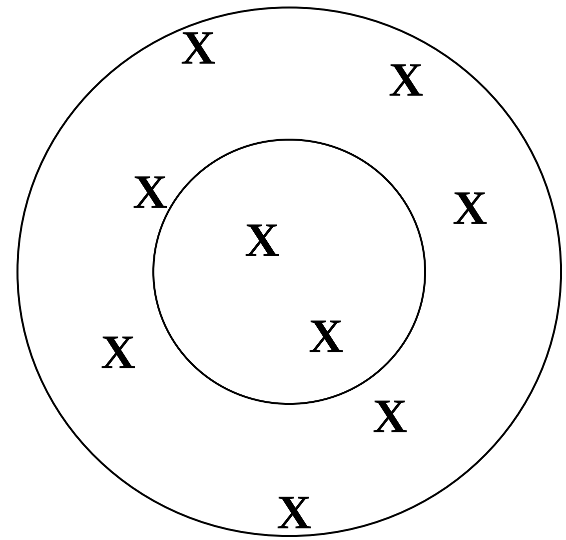
Key Point: Accuracy deals with the Measurement Systems ability to be close, on average, to the actual value



Accurate?



Accurate?

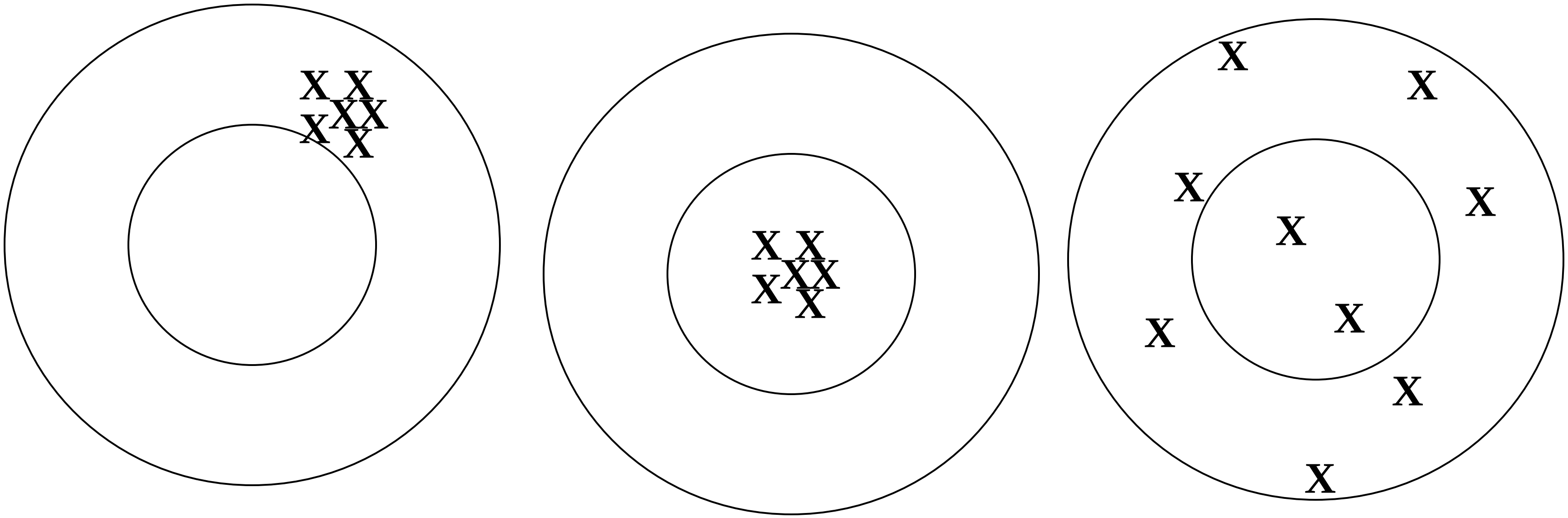


Accurate?

Measurement Systems Analysis

- Repeatable

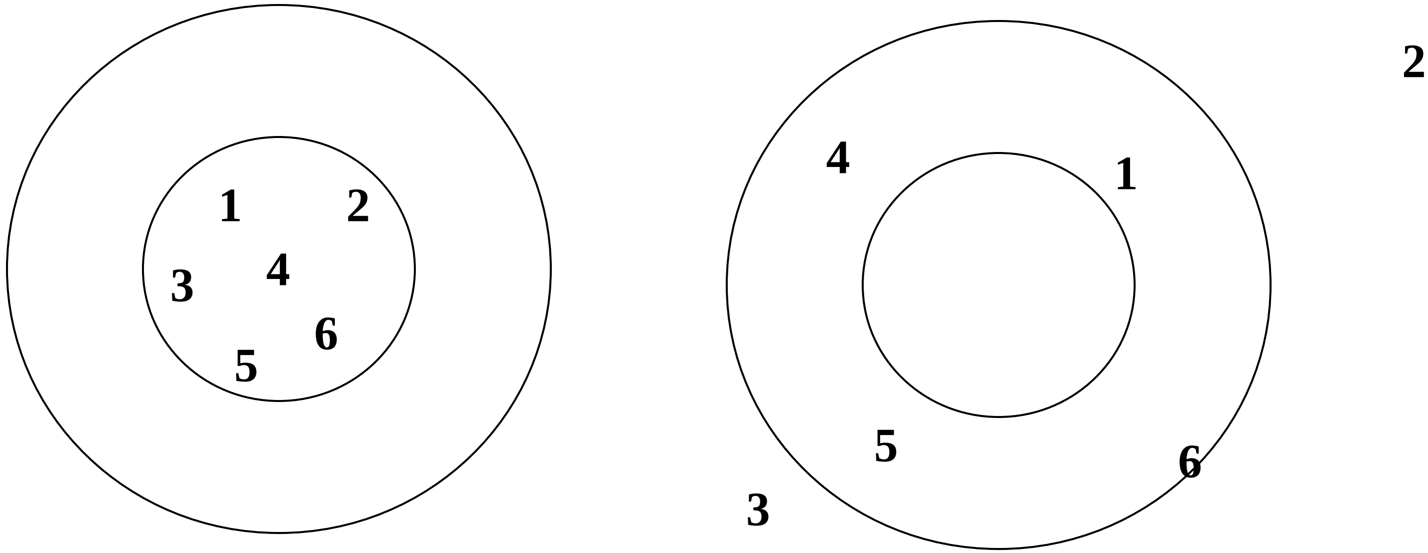
Key Point: Repeatability deals with the Measurement System's ability to measure over numerous trials within a limited range of variation



Measurement Systems Analysis

- Reproducibility

Key Point: Reproducibility deals with the Measurement Systems ability to reproduce results between labs, instruments, or analysts (people)



Note: each # represents a QC Technician

Measurement Systems Analysis

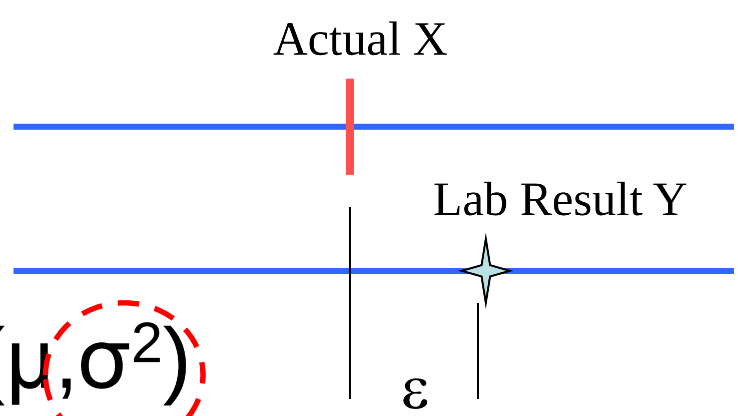
- Mathematical Model

$$Y = X + \varepsilon$$

Lab Result

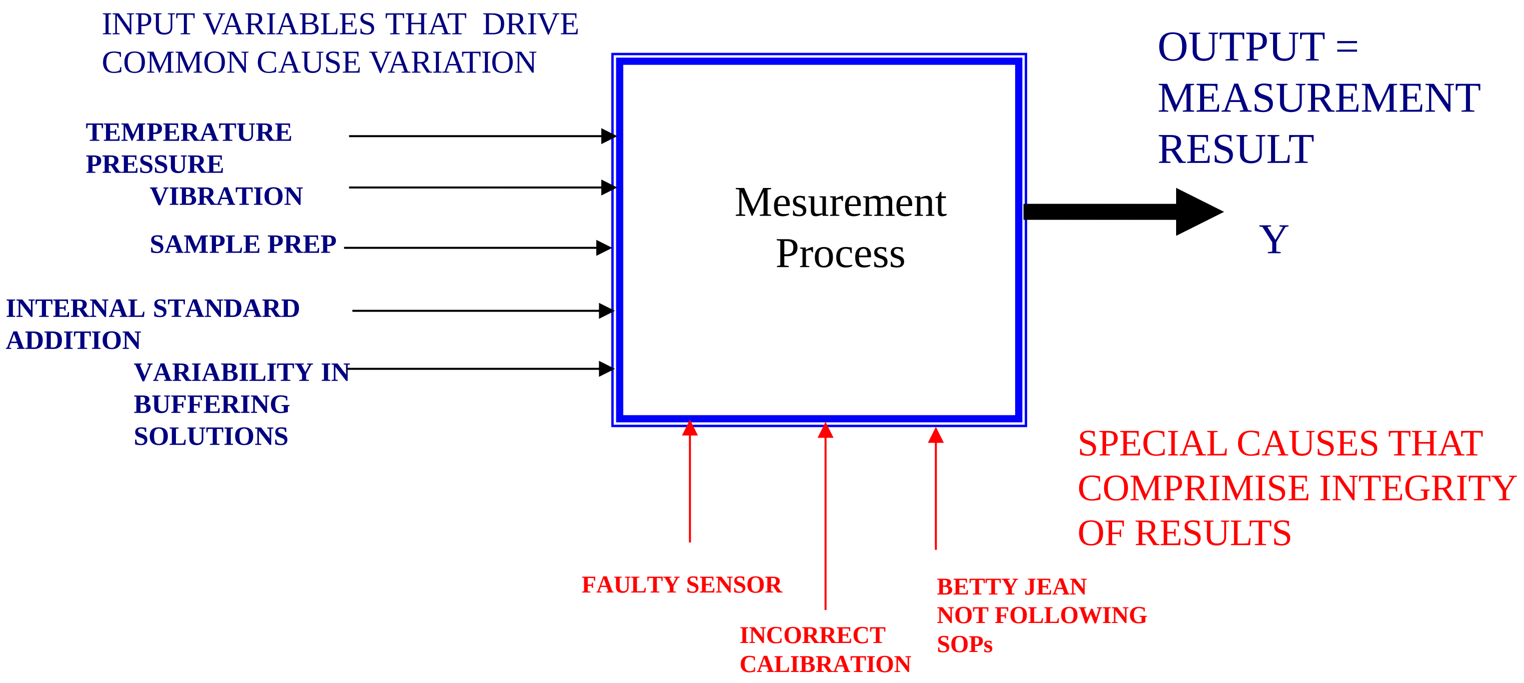
Actual Unknown Value from the process

Error introduced by the measurement system



- Assume
  $\varepsilon \sim \text{IID } N(\mu, \sigma^2)$

A Measurement SIPOC System Model



Measurement Systems Analysis

- Variance Components

$$\sigma^2_{gage} = \sigma^2_{repeatability} + \sigma^2_{reproducibility}$$

Instrument (repeatability) Analyst (reproducibility)

- Experimental Design: Gage R&R Study estimates these components of error by having the same sample measured a few times each by a few different operators, and conducting that for a representative set of samples that would encompass the natural process variation

Measurement Systems Analysis

- Methods of statistical analysis

  A) X-Bar & R Method (Traditional)

  B) ANOVA GLM - Method (Montgomery)

- Standard Deviation Components are determined by the R-bar/d_2 method

n	d2
2	1.128
3	1.693
4	2.059
5	2.326
6	2.534
7	2.704

Measurement Systems Analysis

- X-bar & R Method

Betty Jean

Sample	Trial 1	Trial2	Trial3	Average-Appraiser 1	Range1
P1	10.2	10.3	10.1	10.20	0.2
P2	7.6	7.9	8.2	7.90	0.6
P3	12.3	12.7	12.4	12.47	0.4
P4	11.1	11.3	11.4	11.27	0.3
P5	9.5	9.7	9.5	9.57	0.2
P6	10.6	10.4	10.3	10.43	0.3
P7	8.3	8.2	8.5	8.33	0.3
P8	10.3	10.2	10.5	10.33	0.3
P9	11.3	11.2	11	11.17	0.3
P10	15.8	15.4	15.2	15.47	0.6
				X-bar 1	R-bar1
				10.71	0.35

Tim Wood

Trial 1	Trial2	Trial3	Average-Appraiser 2	Range2
9.8	9.9	10.1	9.93	0.3
8.3	8.1	7.8	8.07	0.5
11.8	11.5	11.9	11.73	0.4
10.9	10.8	10.6	10.77	0.3
9.3	9.1	9.4	9.27	0.3
10.3	10.2	10.4	10.30	0.2
7.8	8.1	8	7.97	0.3
10.3	9.9	10.1	10.10	0.4
10.9	10.8	11.1	10.93	0.3
14.1	14.3	14.4	14.27	0.3
			Xbar2	Rbar2
			10.33	0.33

Eric Samp

Trial 1	Trial2	Trial3	Average-Appraiser 3	Range3
9.7	9.9	10.2	9.93	0.5
8	8.2	7.9	8.03	0.3
11.7	11.6	12	11.77	0.4
10.6	10.9	10.7	10.73	0.3
9.3	9.4	9.5	9.40	0.2
10.2	10.4	10.2	10.27	0.2
8	8.2	7.8	8.00	0.4
9.8	10	10.2	10.00	0.4
11	11.2	10.8	11.00	0.4
14.1	14.2	14.5	14.27	0.4
			Xbar3	Rbar3
			10.34	0.35

$$\sigma_{repeatabilty} = \frac{\overline{\overline{R}}}{d_2} = \left(\frac{\frac{\overline{R}_1 + \overline{R}_2 + \overline{R}_3}{3}}{1.693} \right) = \frac{(0.35 + 0.33 + 0.35)}{3*1.693} = 0.2028$$

$$R_{operator} = \max\left\{\overline{X}_1, \overline{X}_2, \overline{X}_3\right\} - \min\left\{\overline{X}_1, \overline{X}_2, \overline{X}_3\right\} = 10.71 - 10.33 = 0.38$$

Measurement Systems Analysis

- Determining the components of Error

$$\hat{\sigma}_{reproducibility} = \sqrt{\left(\frac{R_{operator}}{d_2}\right)^2 - \left(\frac{\hat{\sigma}^2_{repeatability}}{p\,r}\right)} = 0.1953$$

$$\hat{\sigma}_{repeatability} = 0.2028$$

- Determining Gage Error

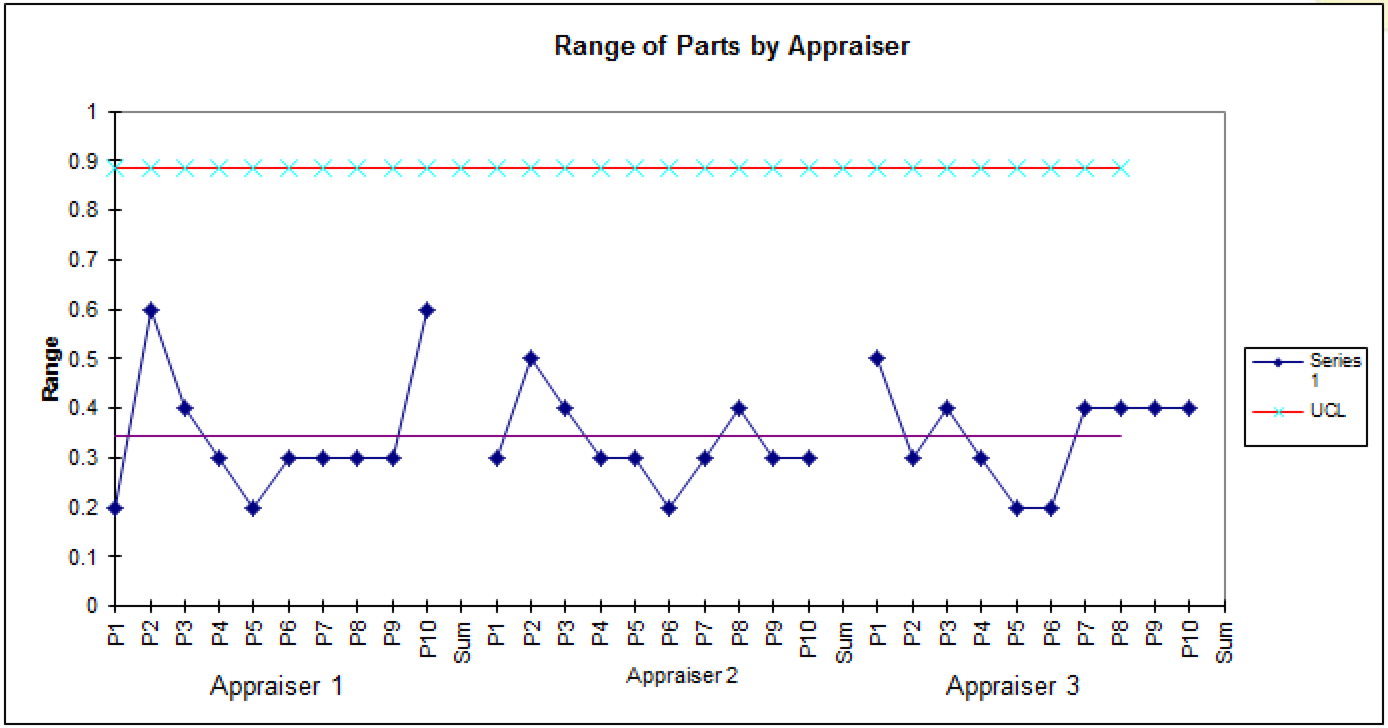
$$\hat{\sigma}_{gage} = \sqrt{\hat{\sigma}^2_{repeatability} + \hat{\sigma}^2_{reproducibility}} = \sqrt{0.2028^2 + 0.1953^2} = 0.2816$$

Measurement Systems Analysis

- Guarantee: $\pm 2 * \sigma_{gage} = \pm 0.56$

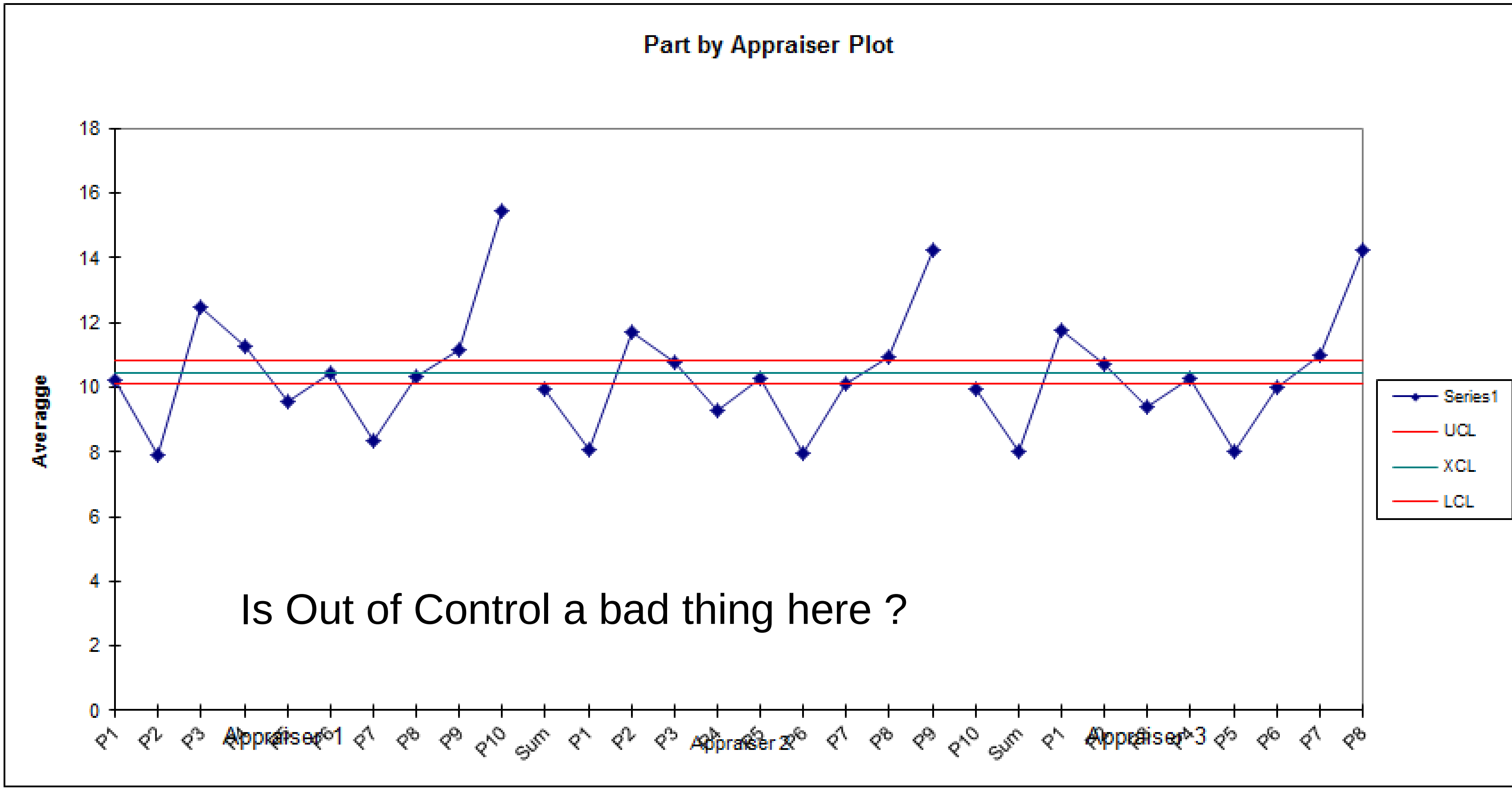
If the R Chart is out of control you must stabilize this first!

- R Chart



Measurement Systems Analysis

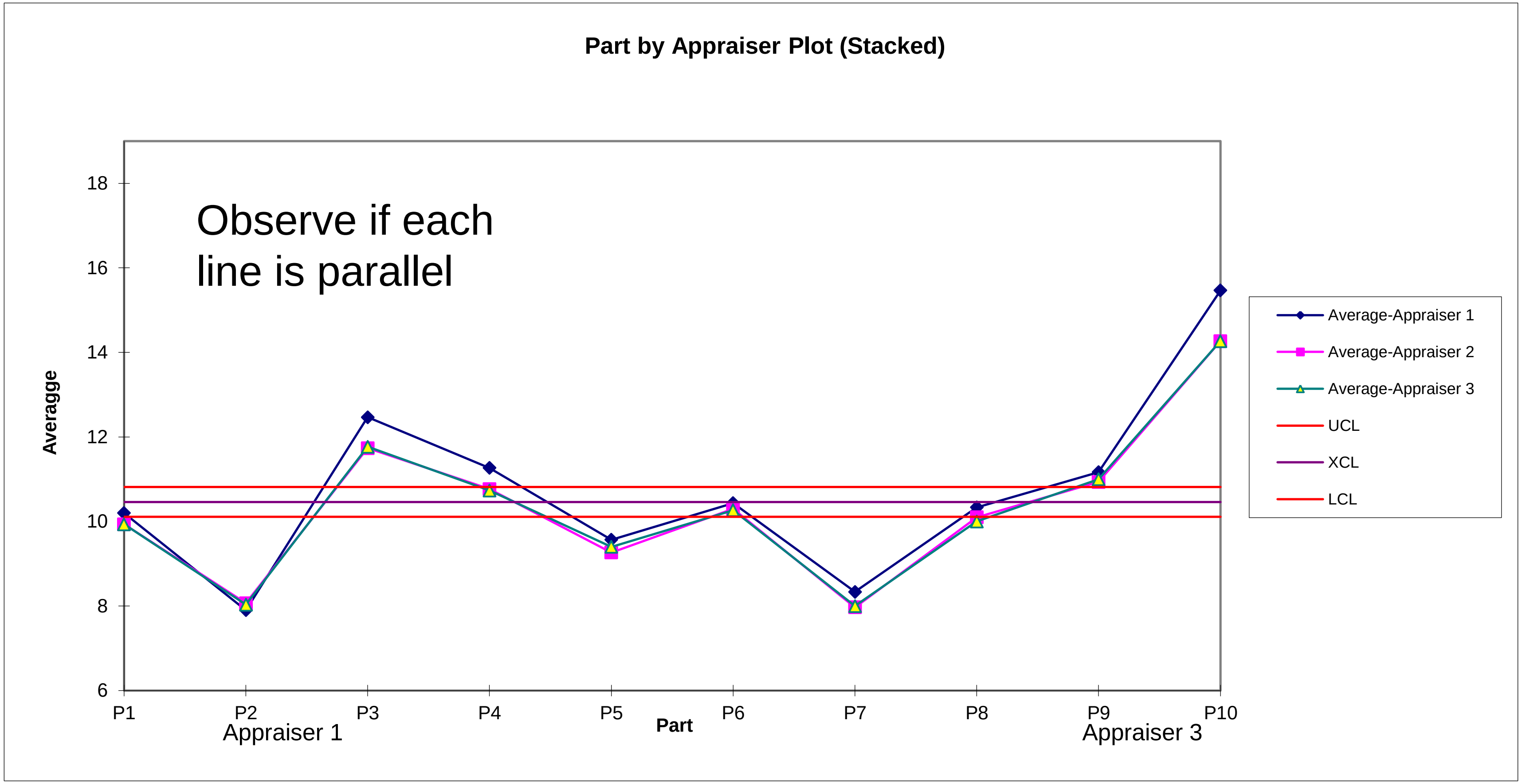
- X-bar Chart



Is Out of Control a bad thing here ?

Measurement Systems Analysis

- Interaction Check



Measurement Systems Analysis

- What do we do with the data?
  Q1: Is my measurement system adequate for the tolerances I am working to in the process?

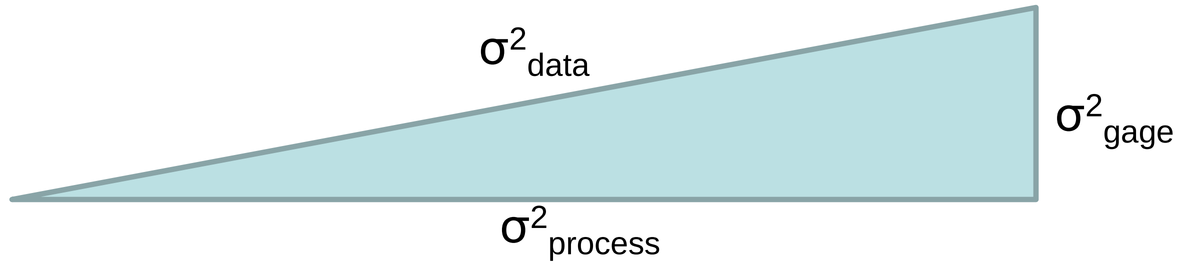
$$\frac{P}{T} = \frac{5.15\hat{\sigma}_{gage}}{USL - LSL} * 100\%$$

If P/T > 30% then ....

Note: Careful Review of the ratio $\frac{\sigma_{reprod}}{\sigma_{repeat}}$ will direct us towards where most of the effort is required to improve overall gage error

Measurement Systems Analysis

- Q2: How adequate is my instrument from discriminating between process variation and instrument variation?
  Variances (not standard deviations) are additive just like sides of a triangle



- Intra-class correlation coefficient $\rho = \dfrac{\sigma_{process}}{\sigma_{data}}$

As $\rho \to 1.0 \Rightarrow$ No Gage Error Exists !!!!

Measurement Systems Analysis

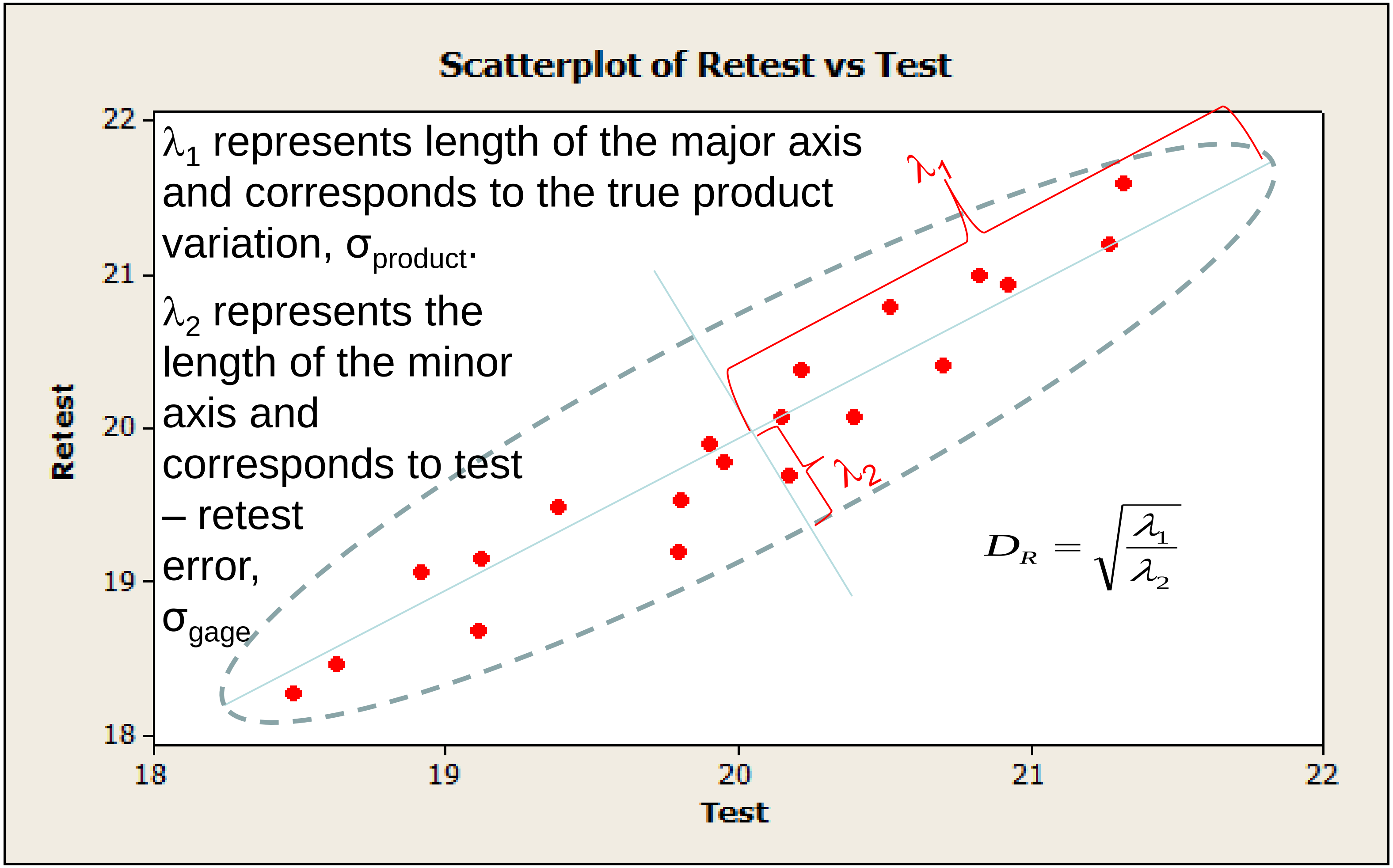
- Consider an example where we test a few samples from our process

- Lets suppose we retested each of these samples

Test	Retest
19.60	20.20
21.30	21.00
20.40	20.90
18.90	19.60
20.80	20.40
18.50	18.70
20.40	20.70
21.60	21.90
20.30	20.00
18.60	18.90

- Plotted Retest vs Test on an X-Y Scatter

Discrimination Ratio



Discrimination Ratio

$$D_R = \sqrt{\frac{\lambda_1}{\lambda_2}}$$

This ratio compares true product variation to the test-retest error

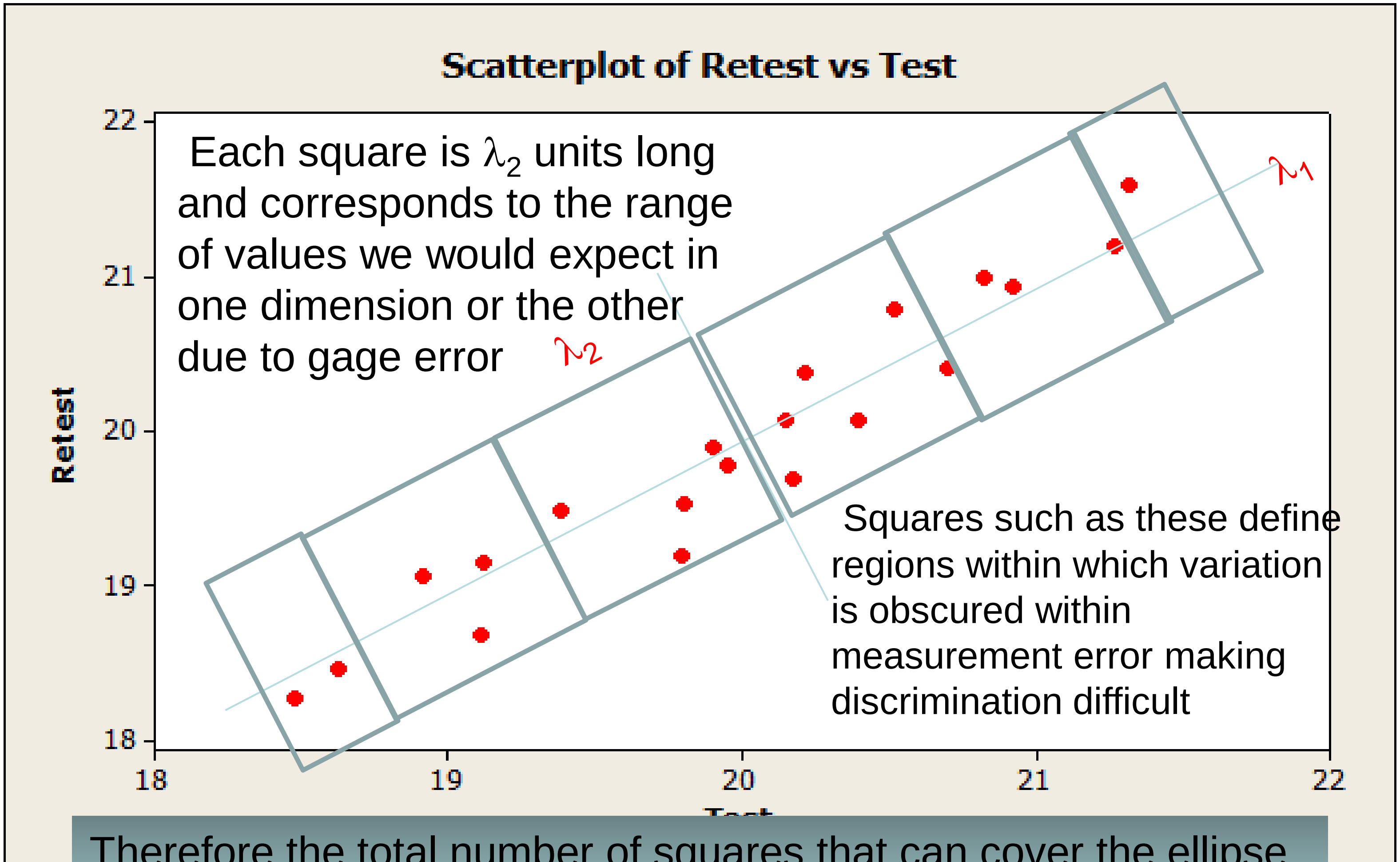
$$= \sqrt{\frac{\sigma^2_{process}(1+\rho)}{\sigma^2_{process}(1-\rho)}}, \quad \rho = \frac{\sigma^2_{process}}{\sigma^2_{data}}$$

$$= \sqrt{\frac{1+\frac{\sigma^2_{process}}{\sigma^2_{data}}}{1-\frac{\sigma^2_{process}}{\sigma^2_{data}}}}, \text{ A Few algebraic steps } \quad IOTTMCO$$

$$= \sqrt{2\frac{\sigma^2_{data}}{\sigma^2_{gage}} - 1}$$

In this form, we can take the information from both process capability studies and gage studies and calculate Dr directly ☺

Discrimination Ratio



Therefore the total number of squares that can cover the ellipse can be inteprreted as the number of product categories within the variability observed in the data from the process

Measurement Systems Analysis

$$D_R = \sqrt{\frac{2\hat{\sigma}^2_{process}}{\hat{\sigma}^2_{repeatabilty}} - 1}$$

If $D_R = 2.0 \Rightarrow$

If $D_R = 3.0 \Rightarrow$

If $D_R > 4.0 \Rightarrow$

If $D_R < 2.0 \Rightarrow$

Measurement Systems Analysis

- What can we do if our $D_R < 2.0$?

- ASBC Method 3: Ruggedness Testing
  - what are the critical X-vars using DOE

- Average of n-measurements $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$

$$D_R = \sqrt{\frac{2\hat{\sigma}^2_{data}}{\hat{\sigma}^2_{repeatabilty} / n} - 1} > 4.0$$

$$\Rightarrow n > \frac{17}{2}\left(\frac{\sigma^2_{repeatabilty}}{\sigma^2_{data}}\right)$$

Measurement Systems Analysis

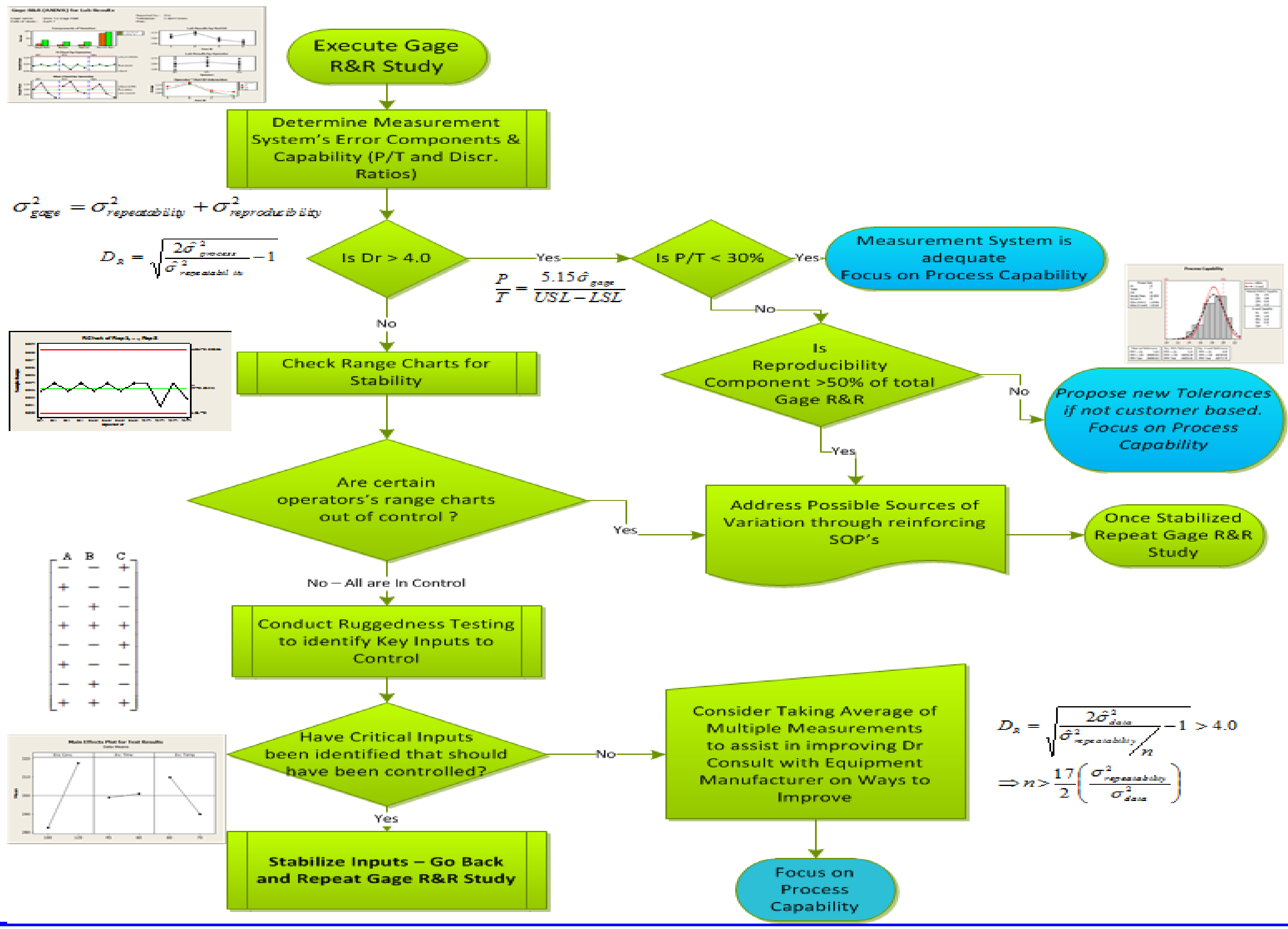
- Why is it important ?

$$\hat{P}_{pl} = \left(\frac{\overline{\overline{X}} - LSL}{3\hat{\sigma}_{longterm}} \right) \qquad \hat{P}_{pu} = \left(\frac{USL - \overline{\overline{X}}}{3\hat{\sigma}_{longterm}} \right)$$

- In reality there are two components of variability associated with our process data
  1. Actual process 2. Gage Error

$$\hat{\sigma}^2_{longterm} = \hat{\sigma}^2_{process} + \hat{\sigma}^2_{gage} = \hat{\sigma}^2_{process} + \hat{\sigma}^2_{repeatabilty} + \hat{\sigma}^2_{reproducibility}$$

Measurement Systems Analysis – Roadmap to DMAIC



SUMMARY

- Assumptions in a Process Capability Study
- How do we check for Identically Distributed data
- How to we check for Normality
- What is the difference between Ppk and Cpk
- What is considered a good Ppk number?

SUMMARY

- If your Ppk is not sufficient what is the first thing you must assess?
- How do you test if your process is on-target?
- How do you test if a process stream may have a different mean than the others
- Y=F(X) – what does this mean

SUMMARY

- If your Ppk is not sufficient what is the first thing you must assess?
- How do you test if your process is on-target?
- How do you test if a process stream may have a different mean than the others
- Y=F(X) – what does this mean

Summary

- How do we test if continuous X-variables influence the output of the process
- Functional Limits?
- How do we test if categorical X-variables influence the output of the process
- If we control the _______ we will control the _____

Summary

- Why is a gage capability study important?
- When should you conduct one?
- What are the components of gage error?
- What does the discrimination ratio tell us?
- What is a good Dr?

Thank You

eric.samp@millercoors.com