2.830J / 6.780J / ESD.63J Control of Manufacturing Processes (SMA 6303)
Spring 2008

Control of Manufacturing Processes

Subject 2.830/6.780/ESD.63

Spring 2008

Lecture #7

Shewhart SPC & Process Capability

February 28, 2008

Applying Statistics to Manufacturing: The Shewhart Approach

Text removed due to copyright restrictions. Please see the Abstract of Shewhart, W. A. “The Applications of Statistics as an Aid in Maintaining Quality of a Manufactured Product.” *Journal of the American Statistical Association* 20 (December 1925): 546-548.

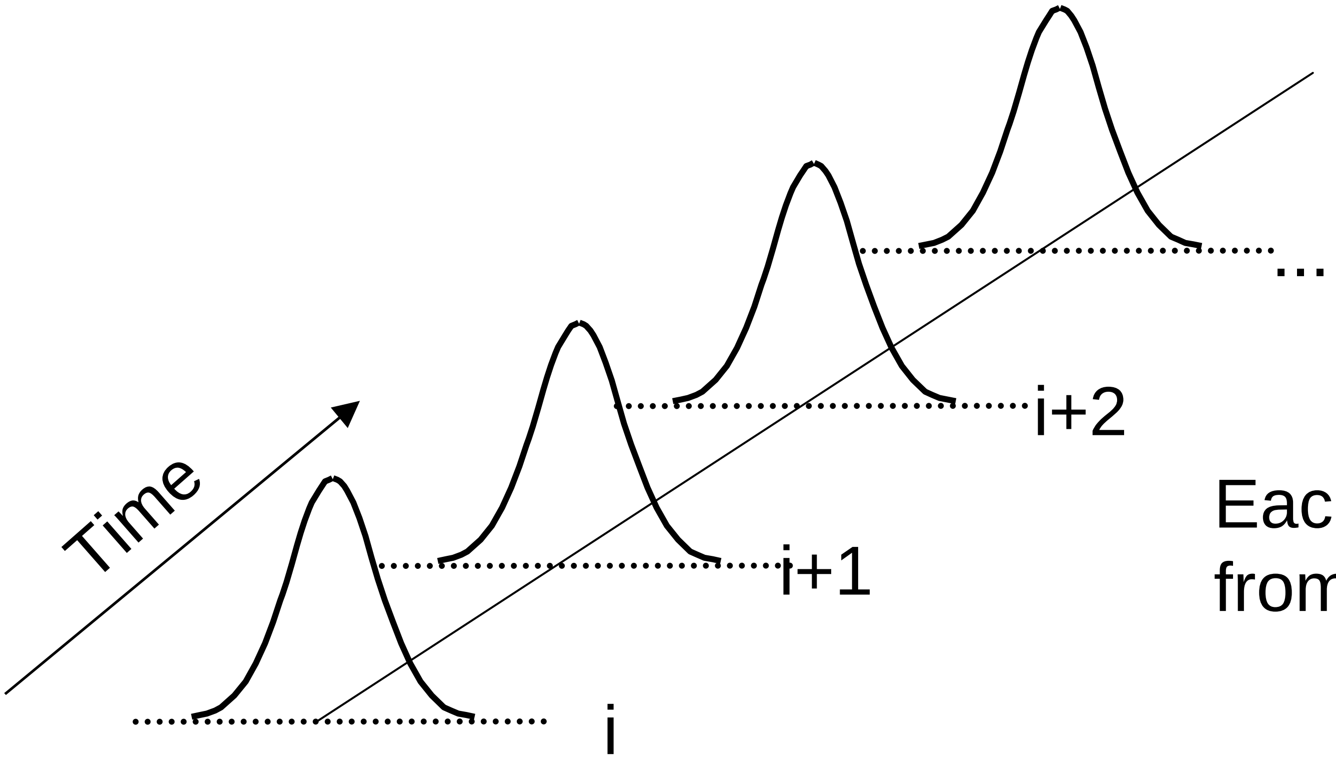
Applying Statistics to Manufacturing: The Shewhart Approach

Text removed due to copyright restrictions. Please see the Abstract of Shewhart, W. A. “The Applications of Statistics as an Aid in Maintaining Quality of a Manufactured Product.” *Journal of the American Statistical Association* 20 (December 1925): 546-548.

Applying Statistics to Manufacturing: The Shewhart Approach (circa 1925)*

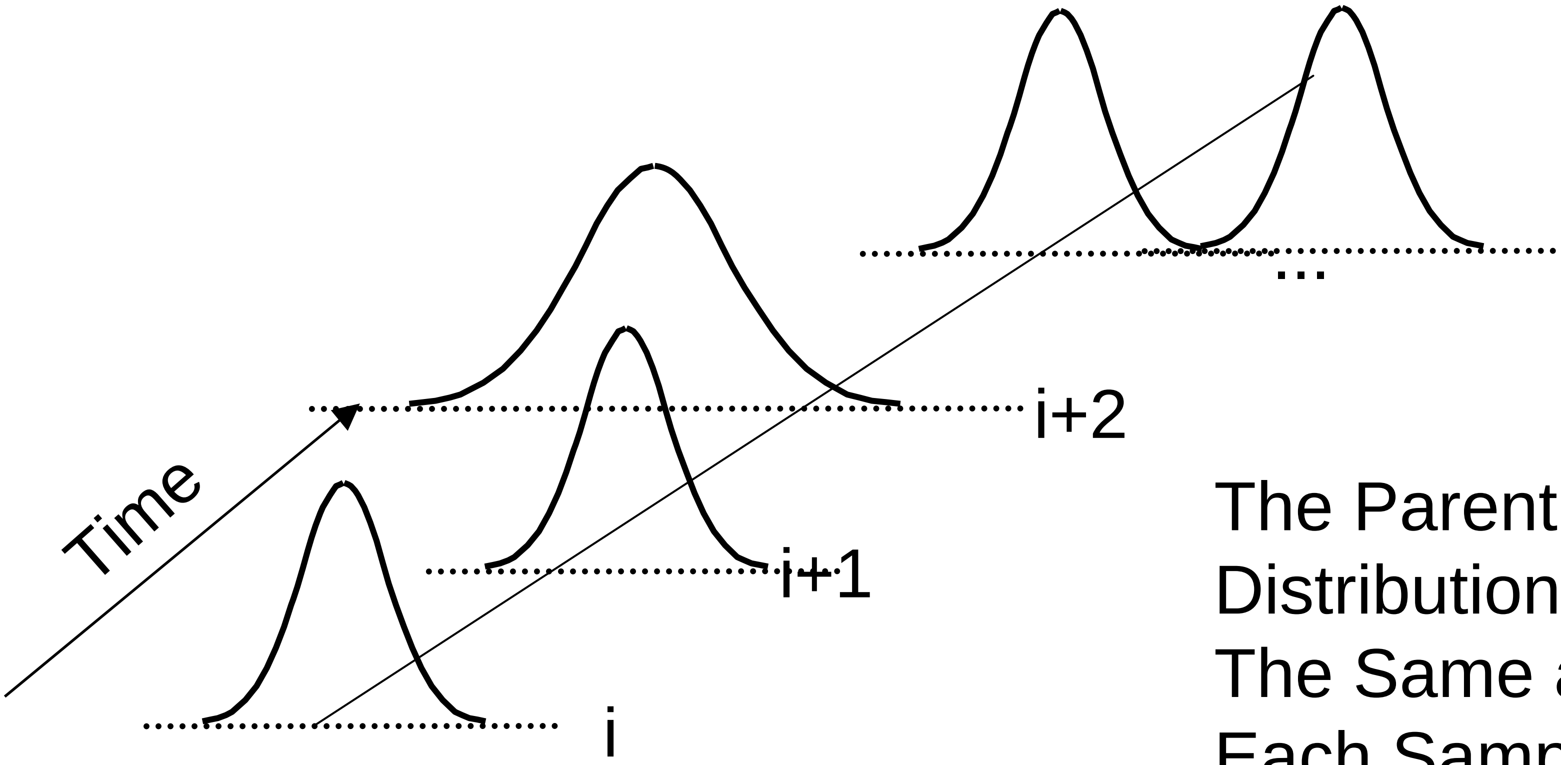
- All Physical Processes Have a Degree of Natural Randomness
- A Manufacturing Process is a Random Process if all "Assignable Causes" (identifiable disturbances) are eliminated
- A Process is "In Statistical Control" if only "Common Causes" (Purely Random Effects) are present.

“In-Control”



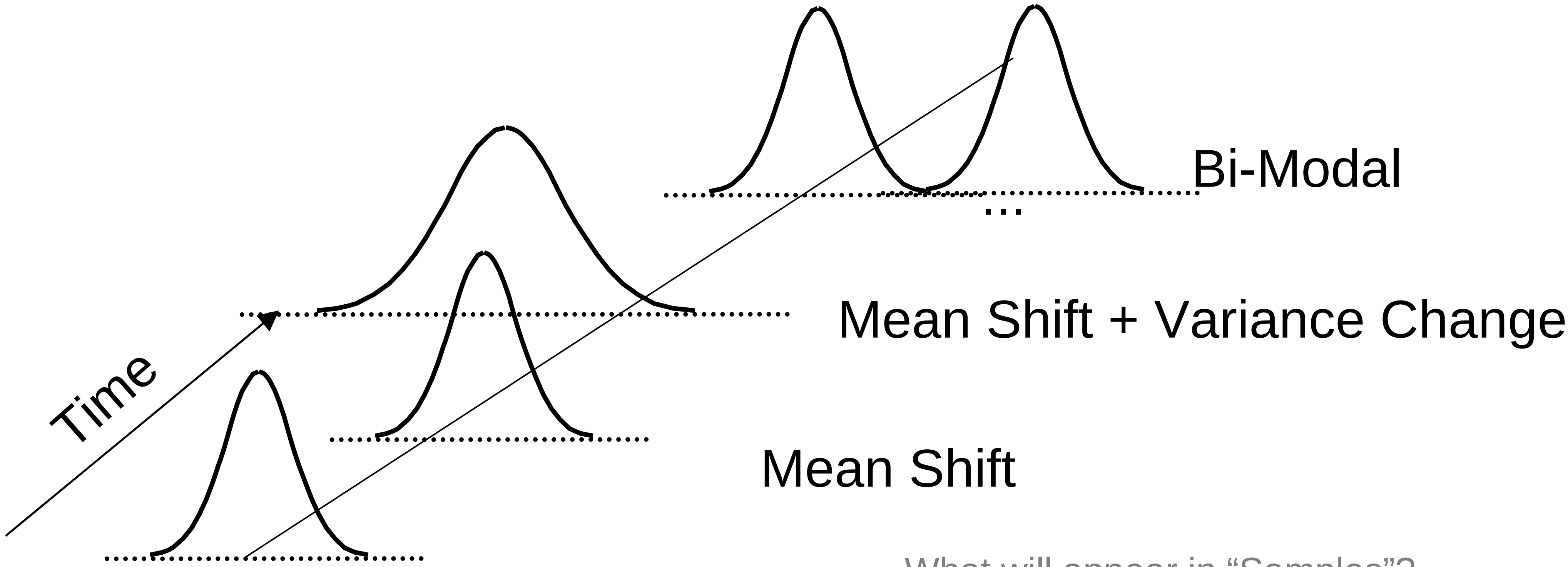
Each Sample is from Same Parent

"Not In-Control"



The Parent Distribution is Not The Same at Each Sample

“Not In-Control”



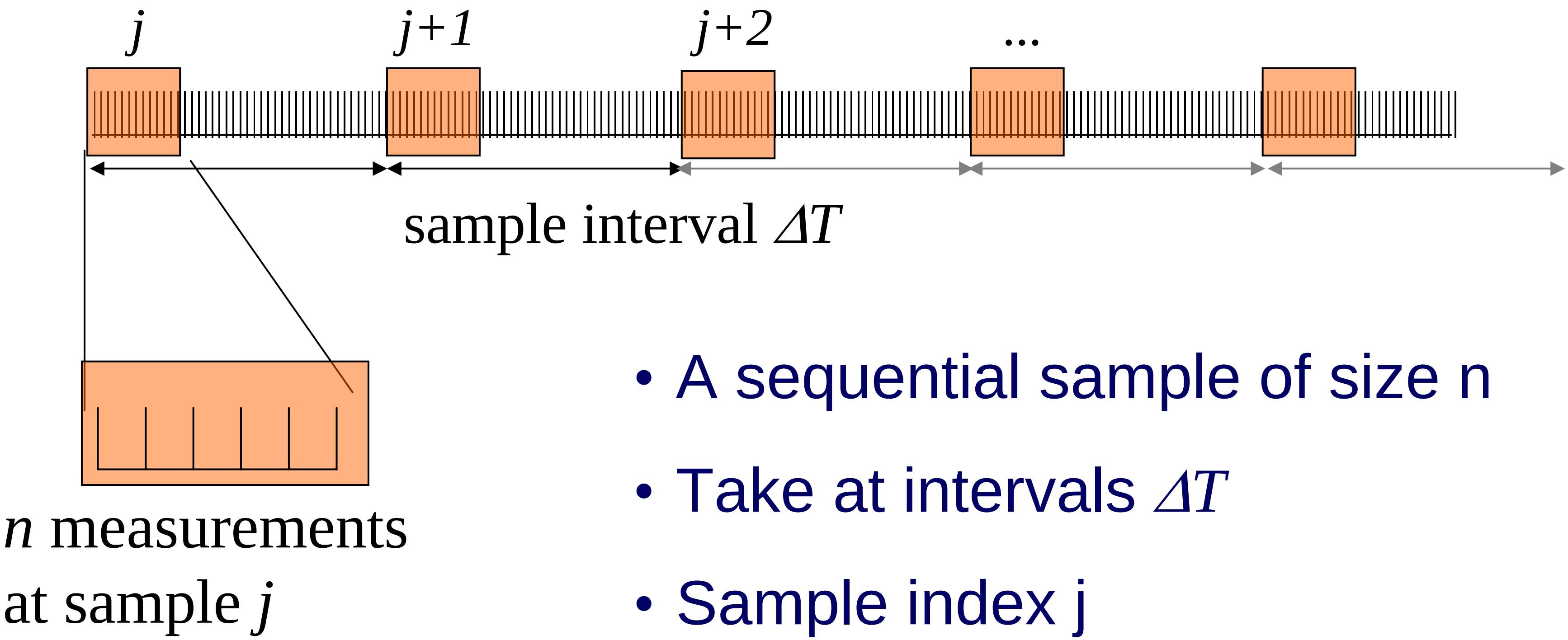
What will appear in “Samples”?

Xbar and S Charts

- Shewhart:
  - Plot *sequential average* of process
    - Xbar chart
    - Distribution?
  - Plot sequential sample standard deviation
    - S chart
    - Distribution?

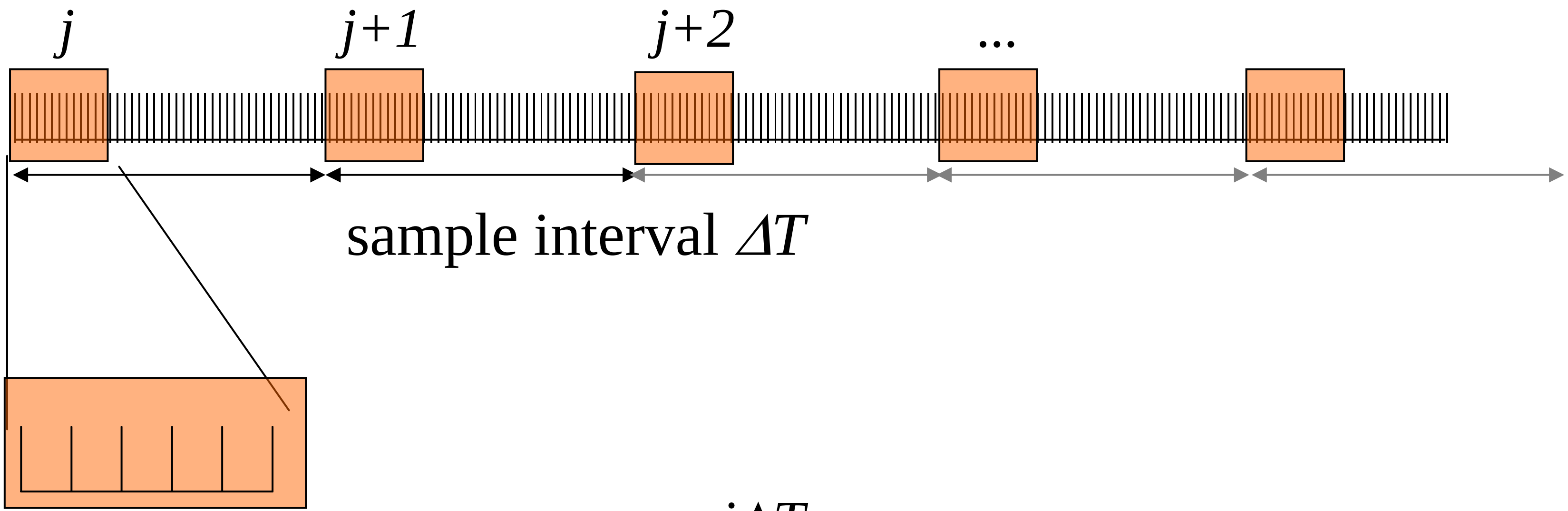
Data Sampling and Sequential Averages

- Given a sequence of process outputs x_i:



- A sequential sample of size n
- Take at intervals ΔT
- Sample index j

Data Sampling

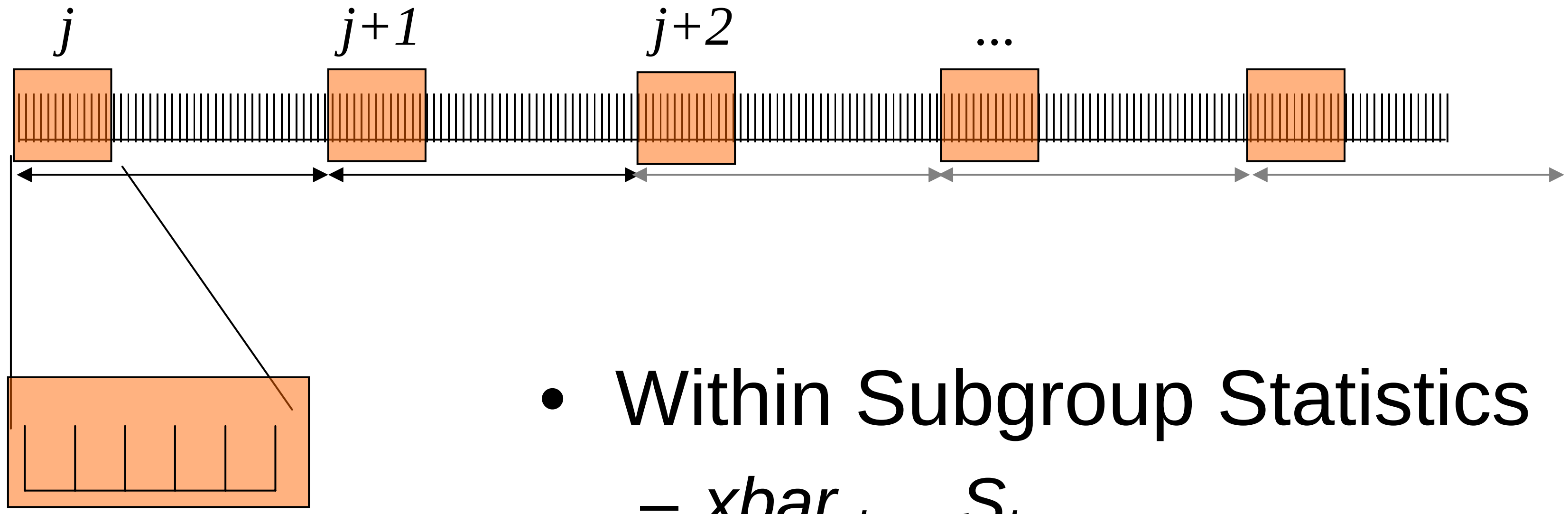


n measurements at sample *j*

$$\bar{x}_j = \frac{1}{n} \sum_{i=(j-1)\Delta T+1}^{j\Delta T+n} x_i \qquad \text{sample } j \text{ mean}$$

$$S_j^{\,2} = \frac{1}{n-1} \sum_{i=(j-1)\Delta T}^{j\Delta T+n} (x_i - \bar{x}_j)^2 \qquad \text{sample } j \text{ variance}$$

Subgroups

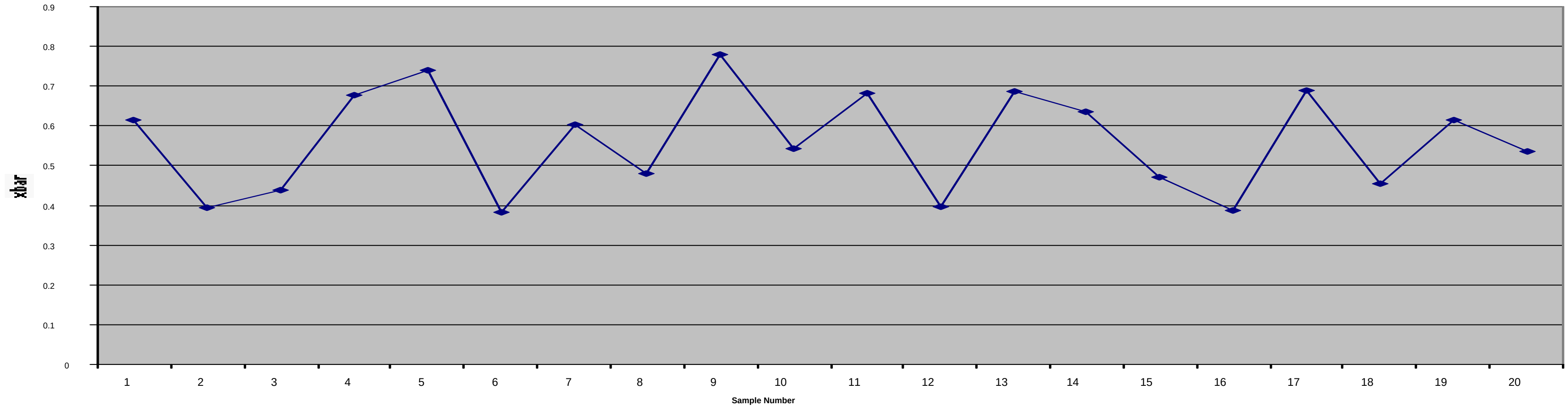


n measurements at sample *j*

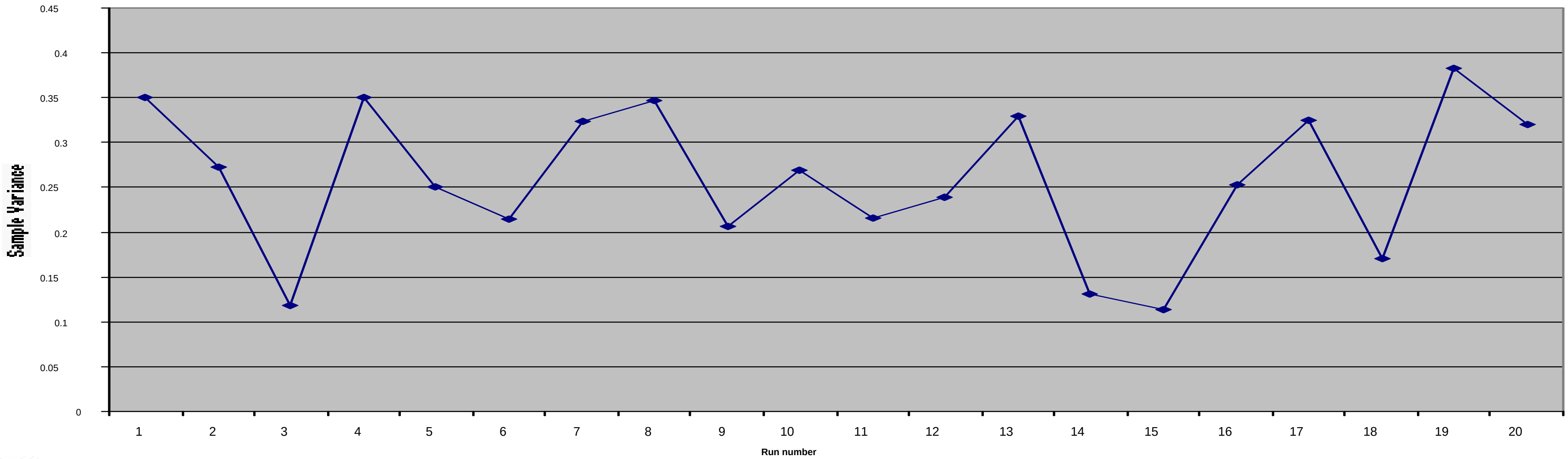
- Within Subgroup Statistics
  - $xbar_j$, S_j
- Between Subgroup Statistics
  - Average of $xbar_j$
  - Variance of xbar(j)

Plot of xbar and S
Random Data n=5

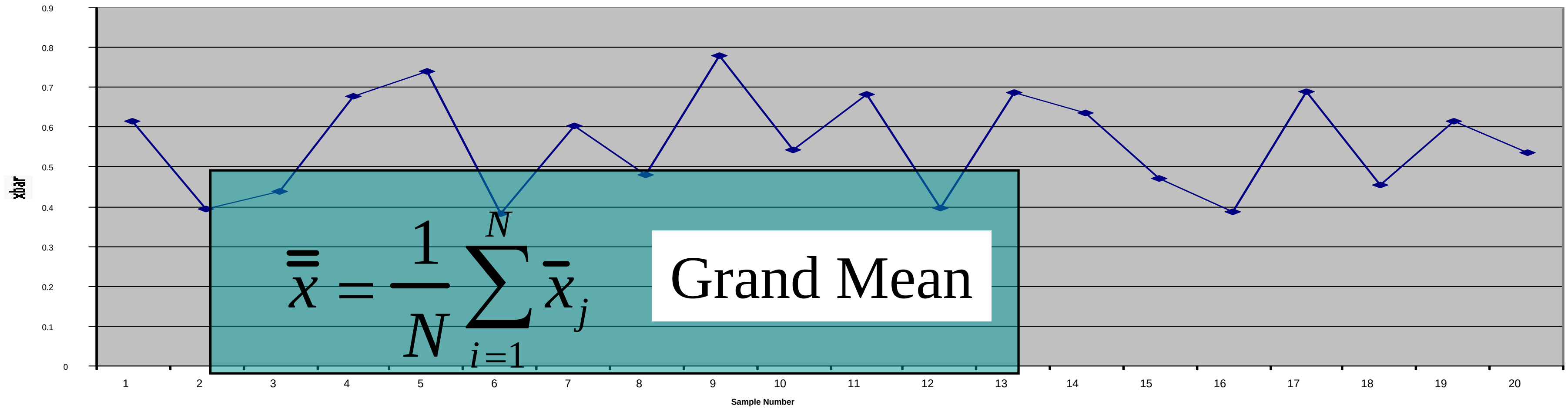
xbar



S

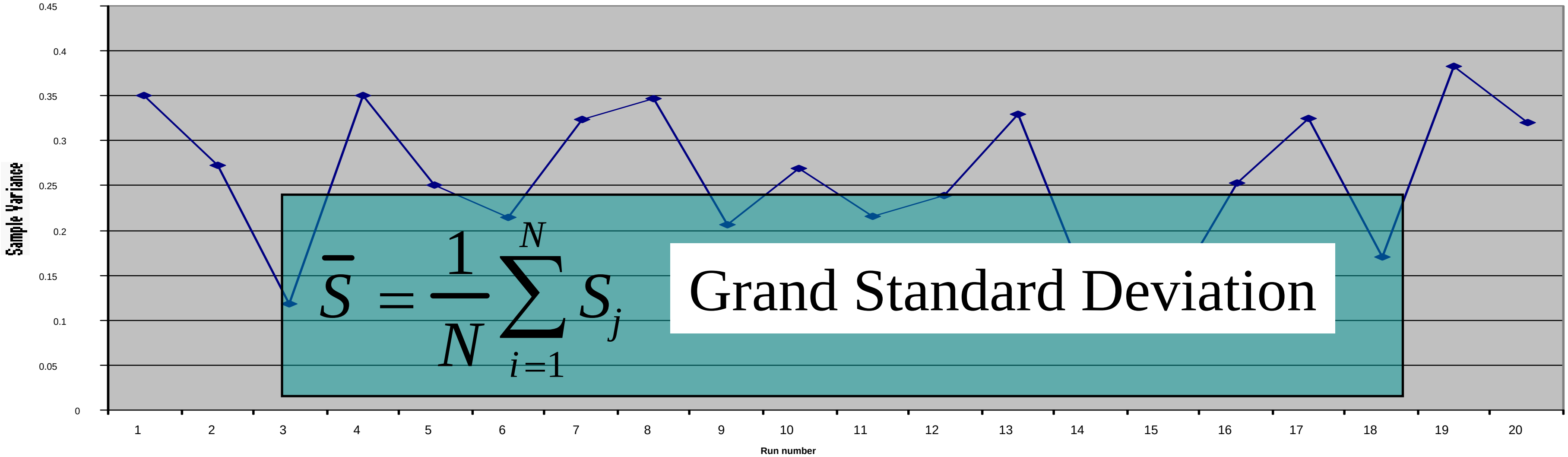


Overall Statistics



$$\bar{\bar{x}} = \frac{1}{N}\sum_{i=1}^{N} \bar{x}_j$$

Grand Mean



$$\bar{S} = \frac{1}{N}\sum_{i=1}^{N} S_j$$

Grand Standard Deviation

Setting Chart Limits

- Expected Ranges
  - Grand mean and Variance
    - (based on what data and how many data points?)

- Confidence Intervals
  - Intervals of $\pm$ n Standard Deviations
  - Most Typical is $\pm$ 3σ (US) or 0.1% (Europe)

Chart Limits - Xbar

- If we knew σ_x then:

$$\sigma_{\bar{x}} = \sqrt{\frac{1}{n}}\sigma_x$$

- But Since we *Estimate* the Sample Standard Deviation, then

$$E(S_j) = C_4 \sigma_{\bar{x}}$$ (S$_j$ is a biased estimator)

where $$C_4 = \left(\frac{2}{n-1}\right)^{1/2} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}$$

Chart Limits xbar chart

With this "correction" we can set limit at $\pm 3\sigma_{xbar}$

Or set a confidence interval of 99.7%

Or a test significance of 0.3%

$$UCL = \bar{\bar{x}} + 3\frac{\bar{S}}{C_4\sqrt{n}} \qquad LCL = \bar{\bar{x}} - 3\frac{\bar{S}}{C_4\sqrt{n}}$$

For the example *n*=5 $\quad C_4 = (0.5)^{1/2}\frac{\Gamma(2.5)}{\Gamma(2)} = 0.707\frac{1.33}{1} = 0.94$

Chart Limits S

The variance of the estimate of S can be shown to be:

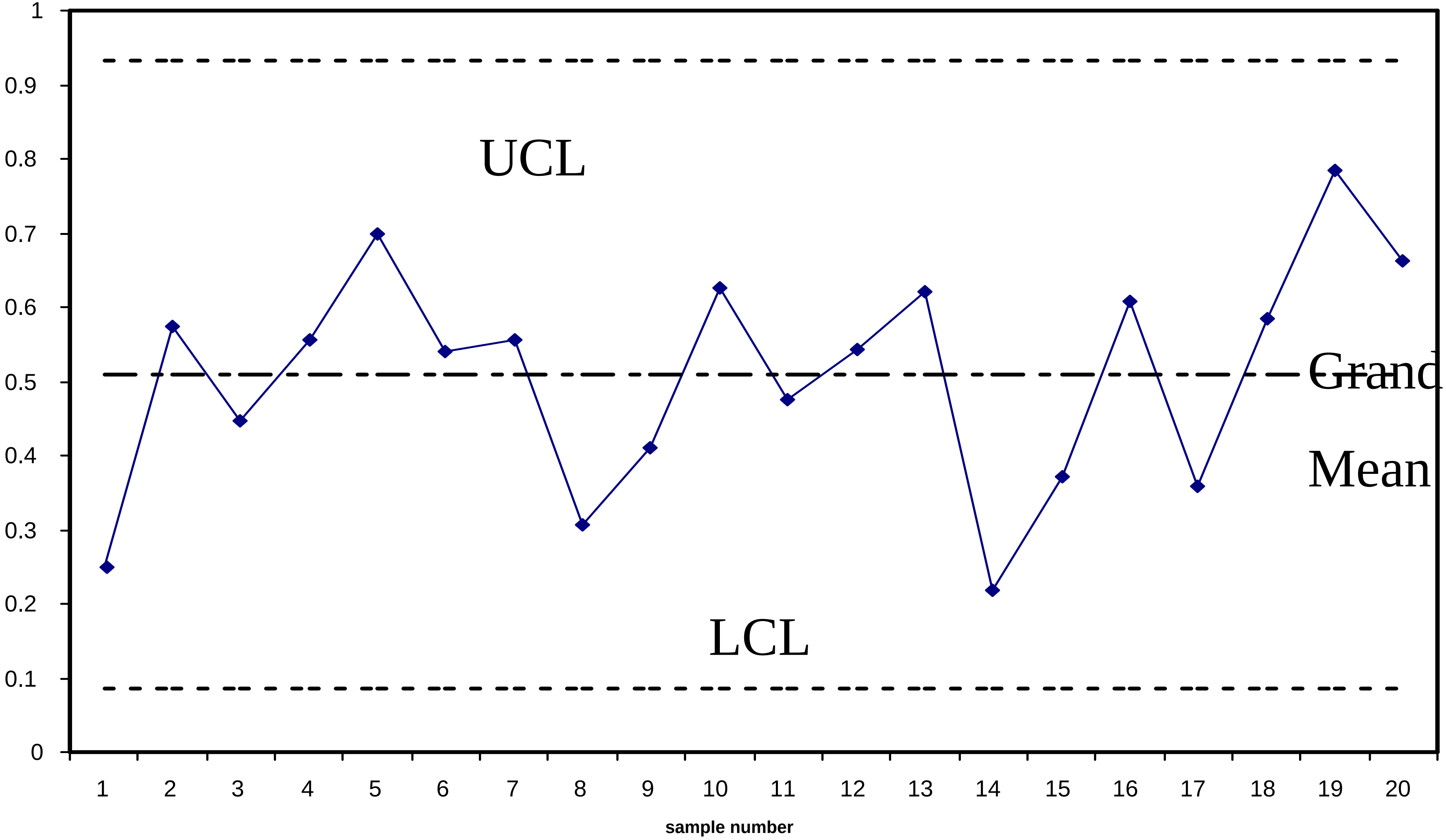
$$\sigma_S = \sigma\sqrt{1 - C_4{}^2}$$

So we get the chart limits:

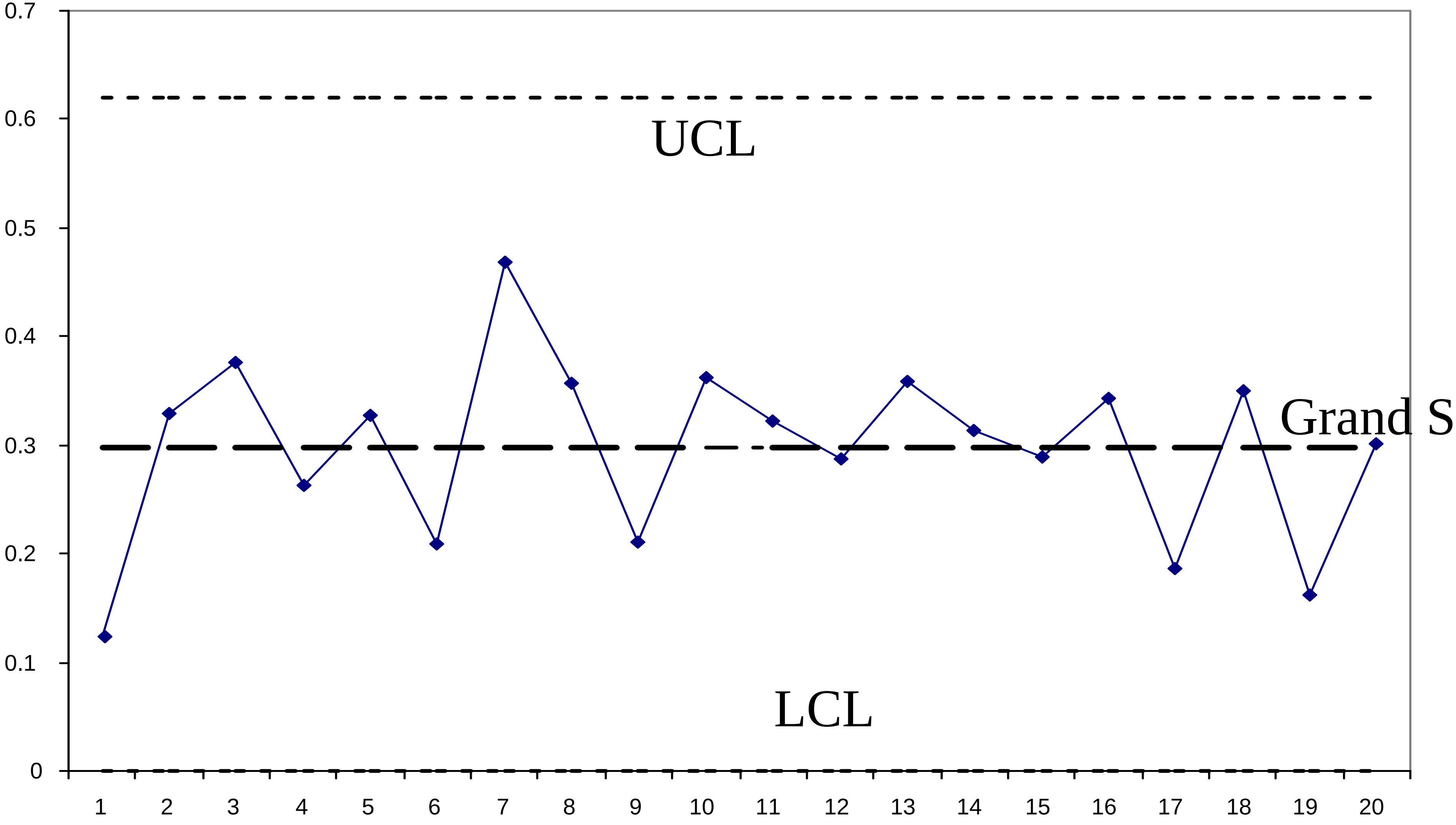
$$UCL = \bar{S} + 3\frac{\bar{S}}{C_4}\sqrt{1 - C_4{}^2}$$

$$LCL = \bar{S} - 3\frac{\bar{S}}{C_4}\sqrt{1 - C_4{}^2}$$

Example xbar



Example S



Manufacturing

Detecting Problems from Running Data

- Appearance of data
  - Confidence Intervals
  - Frequency of extremes
  - Trends

The 8 rules from Devor et al

(Based on Confidence Intervals)

- Prob. of data in a band
- Based on Periodicity
- Based on Linear Trends
- Based on Mean Shift

Test for "Out of Control"

- Extreme Points
  - Outside $\pm 3\sigma$
- Improbable Points
  - 2 of 3 $>\pm 2\sigma$
  - 4 of 5 $>\pm 1\sigma$
  - All points inside $\pm 1\sigma$

Tests for "Out of Control"

- Consistently above or below centerline
  - Runs of 8 or more
- Linear Trends
  - 6 or more points in consistent direction
- Bi-Modal Data
  - 8 successive points outside $\pm 1\sigma$

Applying Shewhart Charting

- Find a run of 25-50 points that are "in-control"
- Compute chart centerlines and limits
- Begin Plotting subsequent $xbar_j$ and S_j
- Apply the 8 rules, or look for trends, improbable events or extremes.
- If these occur, process is "out of control"

Out of Control

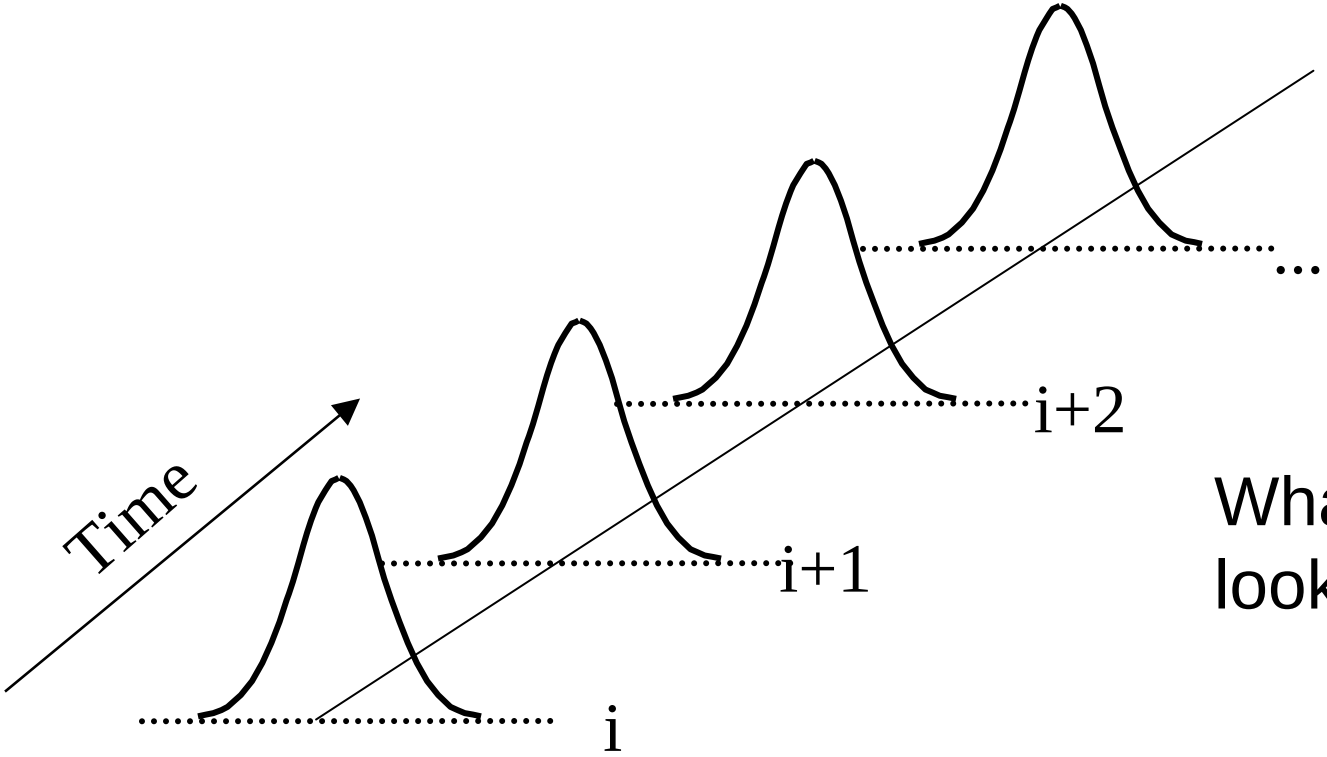
- Data is not Stationary

  (μ or σ are not constant)
- Process Output is being “caused” by a disturbance (common cause)
- This disturbance can be identified and eliminated
  – Trends indicate certain types
  – Correlation with know events
    - shift changes
    - material changes

Western Electric Rules
(See Table 4-1)

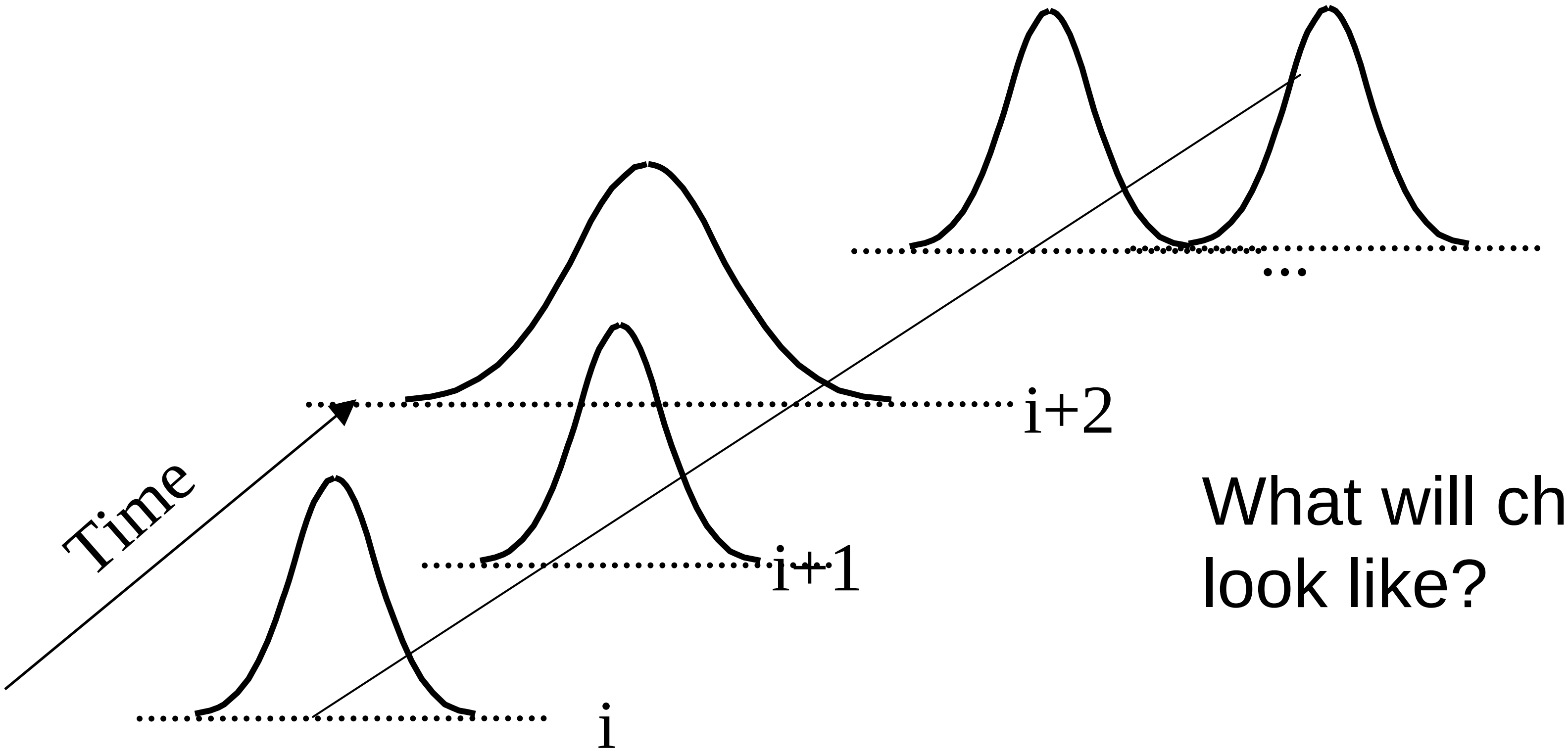
- Points outside limits
- 2-3 consecutive points outside 2 sigma
- Four of five consecutive points beyond 1 sigma
- Run of 8 consecutive points on one side of center

“In-Control”



What will chart look like?

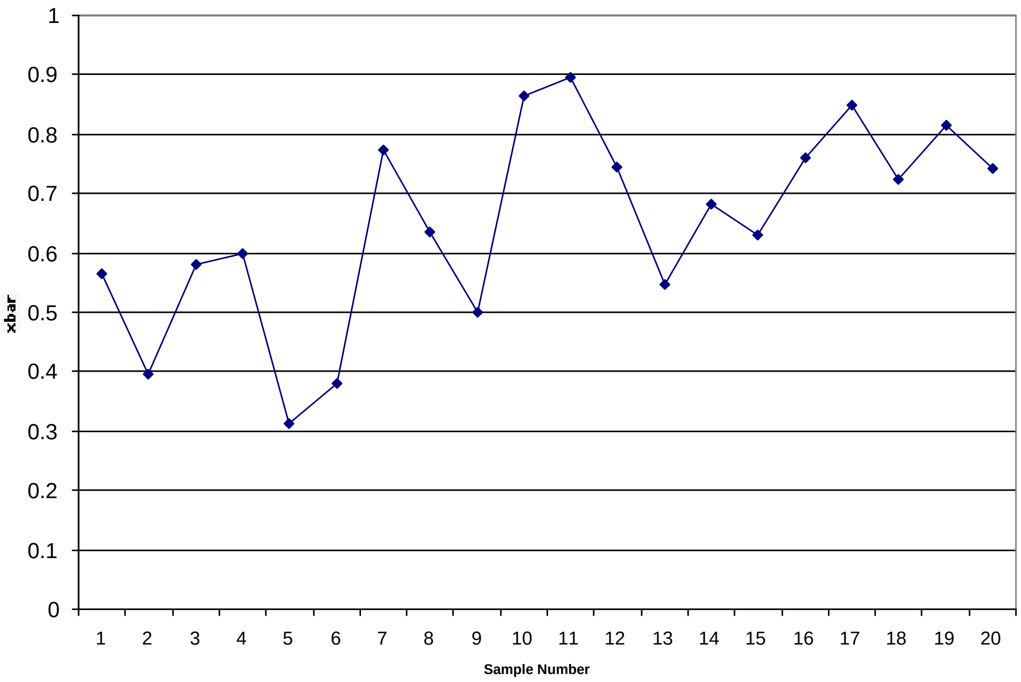
“Not In-Control”



What will chart look like?

Detecting Mean Shifts: Chart Sensitivity

- Consider a real shift of $\Delta\mu_x$:



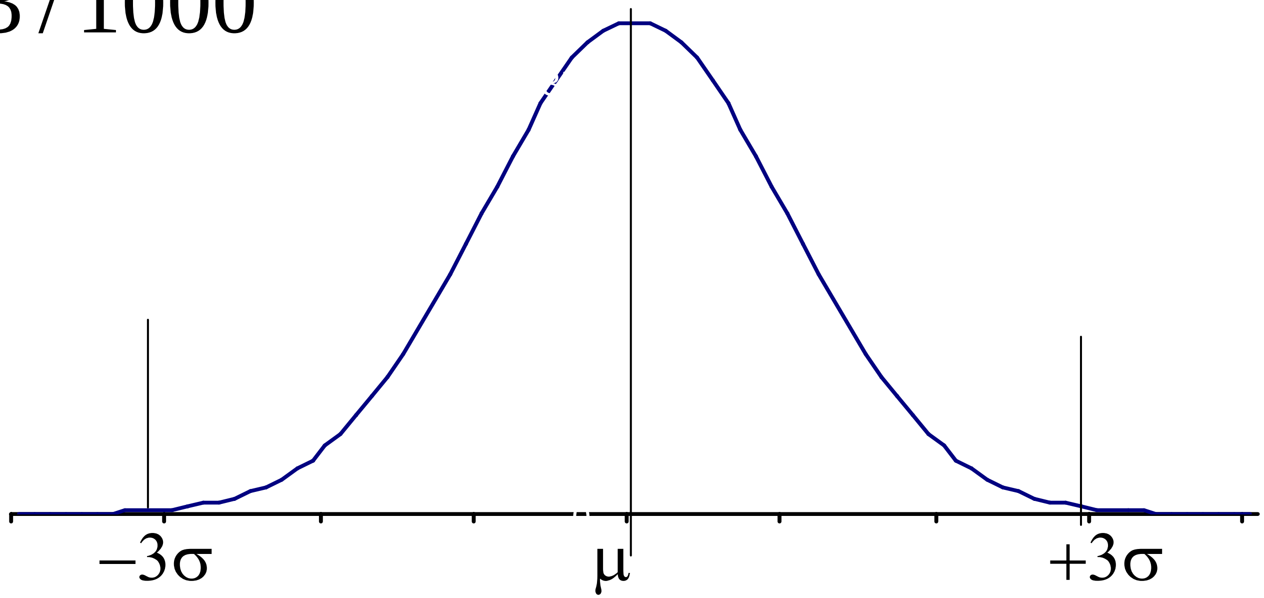
- How many samples before we can expect to detect the shift on the xbar chart?

Average Run Length

- How often will the data exceed the ±3σ limits if $\Delta\mu_x = 0$?

$$\text{Prob}(x > \mu_x + 3\sigma_{\bar{x}}) + \text{Prob}(x < \mu_x - 3\sigma_{\bar{x}})$$

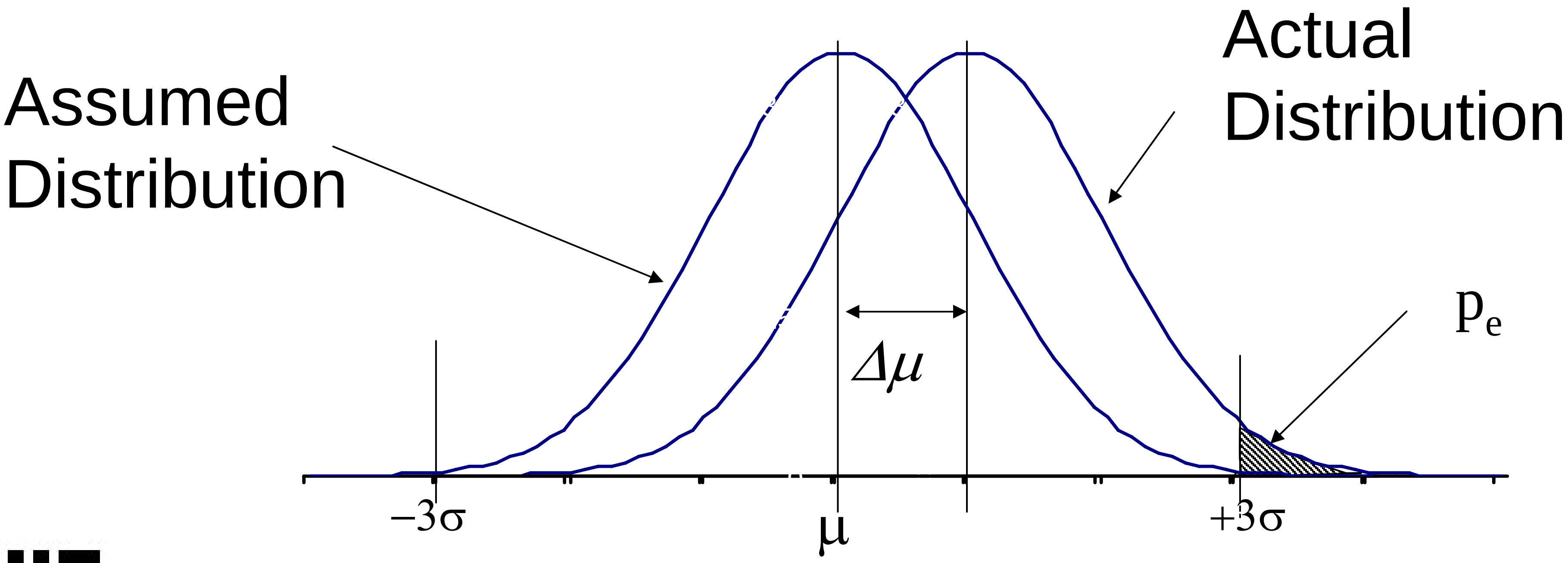
$$= 3/1000$$



Average Run Length

- How often will the data exceed the $\pm 3\sigma$ limits if $\Delta\mu_x = +1\sigma$?

$$\text{Prob}(x > \mu_x + 2\sigma_{\bar{x}}) + \text{Prob}(x < \mu_x - 4\sigma_{\bar{x}})$$
$$= 0.023 + 0.001 = 24/1000$$



Definition

- Average Run Length (arl): Number of runs (or samples) before we can expect a limit to be exceeded = $1/p_e$
  - for $\Delta\mu = 0$ arl = 3/1000 = 333 samples
  - for $\Delta\mu = 1\sigma$ arl = 24/1000 = 42 samples

Even with a mean shift as large as 1σ, it could take **42** samples before we know it!!!

Effect of Sample Size n on ARL

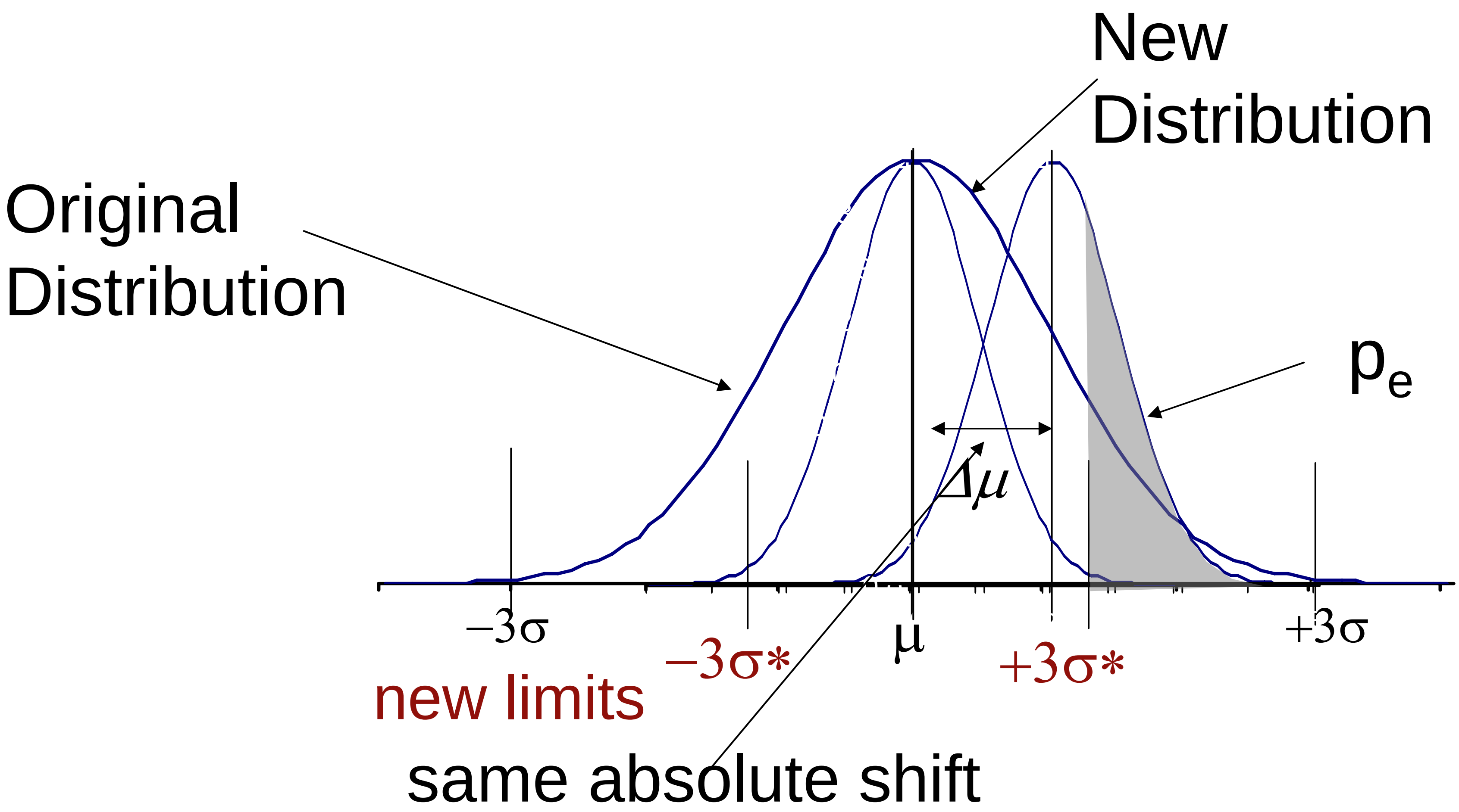
- Assume the same $\Delta\mu = 1\sigma$
  - Note that $\Delta\mu$ is an absolute value

- If we increase n, the Variance of xbar decreases:

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

- So our $\pm\ 3\sigma$ limits move closer together

ARL Example

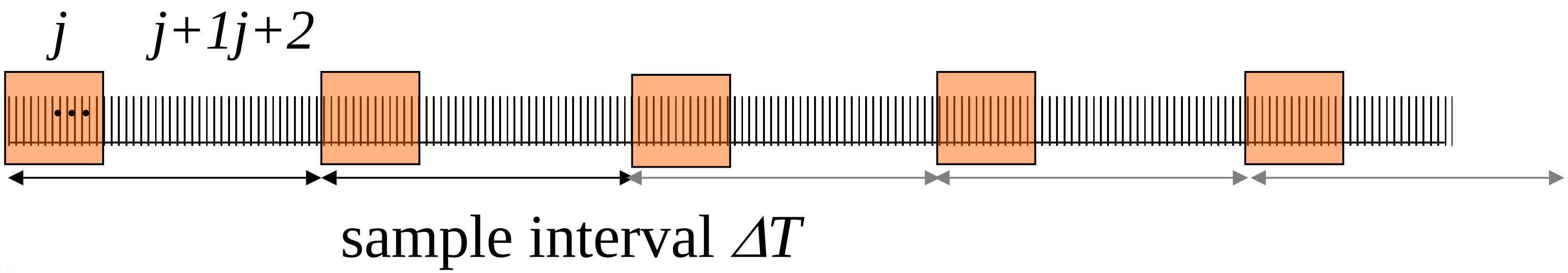


As n <u>increases</u> p_e increases so ARL <u>decreases</u>

Design of the Chart

- Sample size n
  - Central Limit theorem
  - ARL effects?
- Selection of Reference Data
  - Is S at a minimum ?
- Sample time ΔT
  - Cost of sampling
  - production without data
  - Rapid phenomena

Sample size and “filtering” versus response time to changes



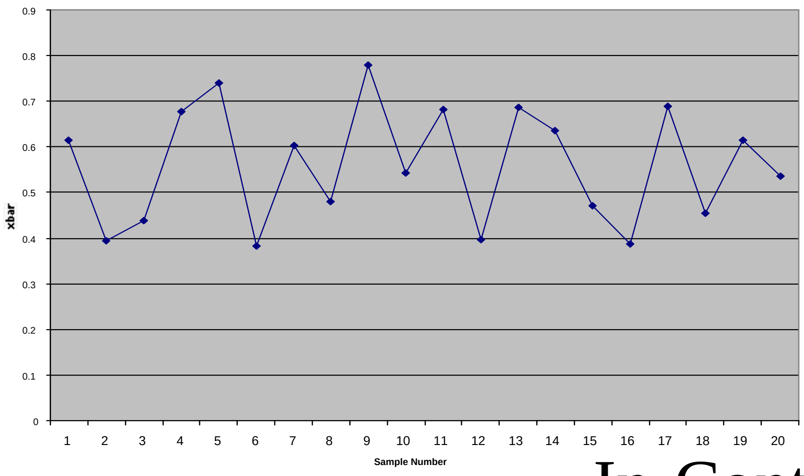
Limits and Extensions

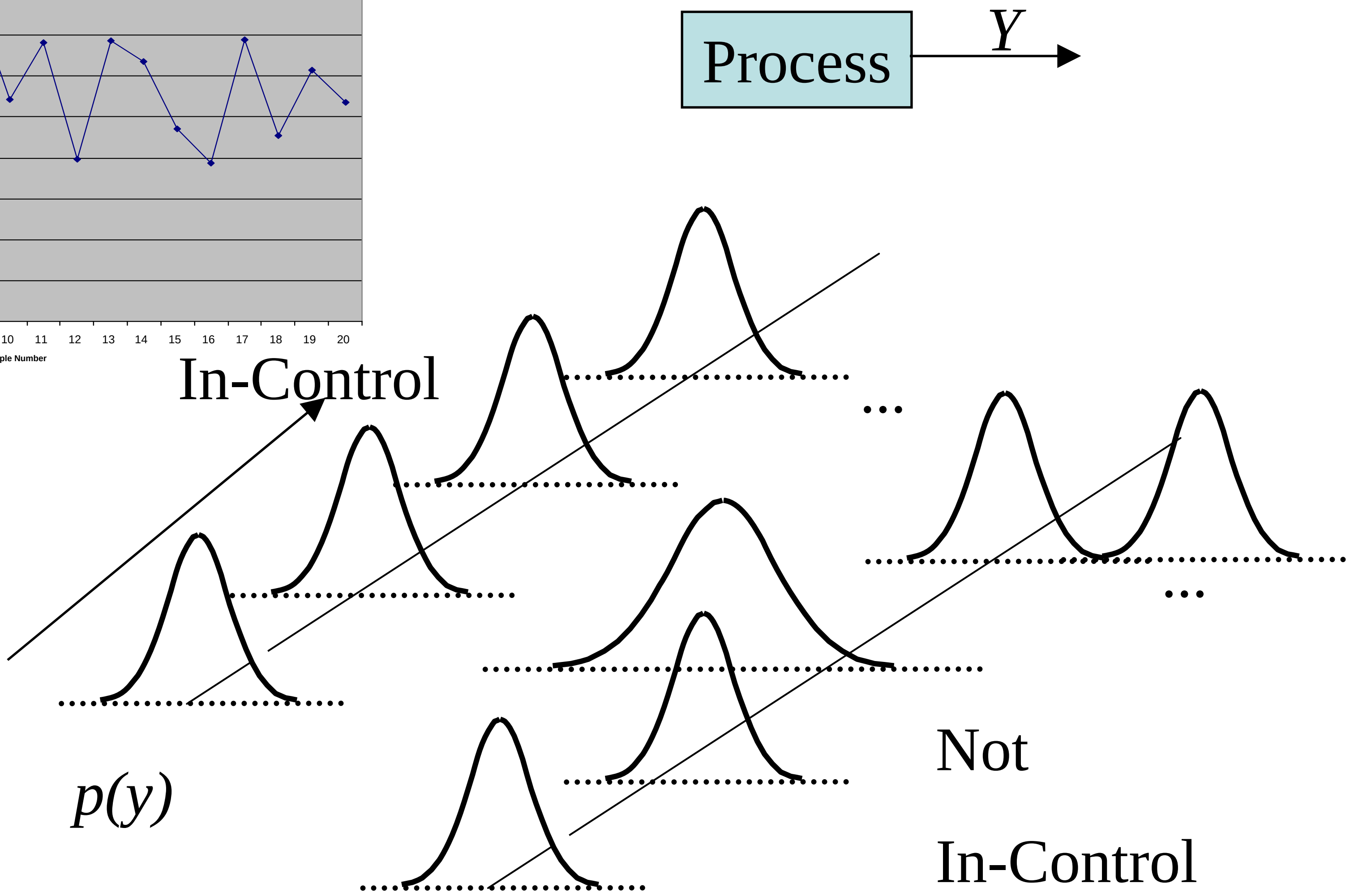
- Need for averaging
- Assumptions of Normality
- Assumption of independence
- Pitfalls
  - Misinterpretation of Data
  - Improper Sampling
- What are alternatives?
  - Different Sampling Schemes
  - Different Averaging Schemes
  - Continuous Update to Improve Statistics

Conclusions

- Hypothesis Testing
  - Use knowledge of PDFs to evaluate hypotheses
  - Quantify the degree of certainty (a and b)
  - Evaluate effect of sampling and sample size
- Shewhart Charts
  - Application of Statistics to Production
  - Plot Evolution of Sample Statistics $\bar{x}$ and S
  - Look for Deviations from Model

Detection : The SPC Hypothesis





Out of Control

- Data is not Stationary

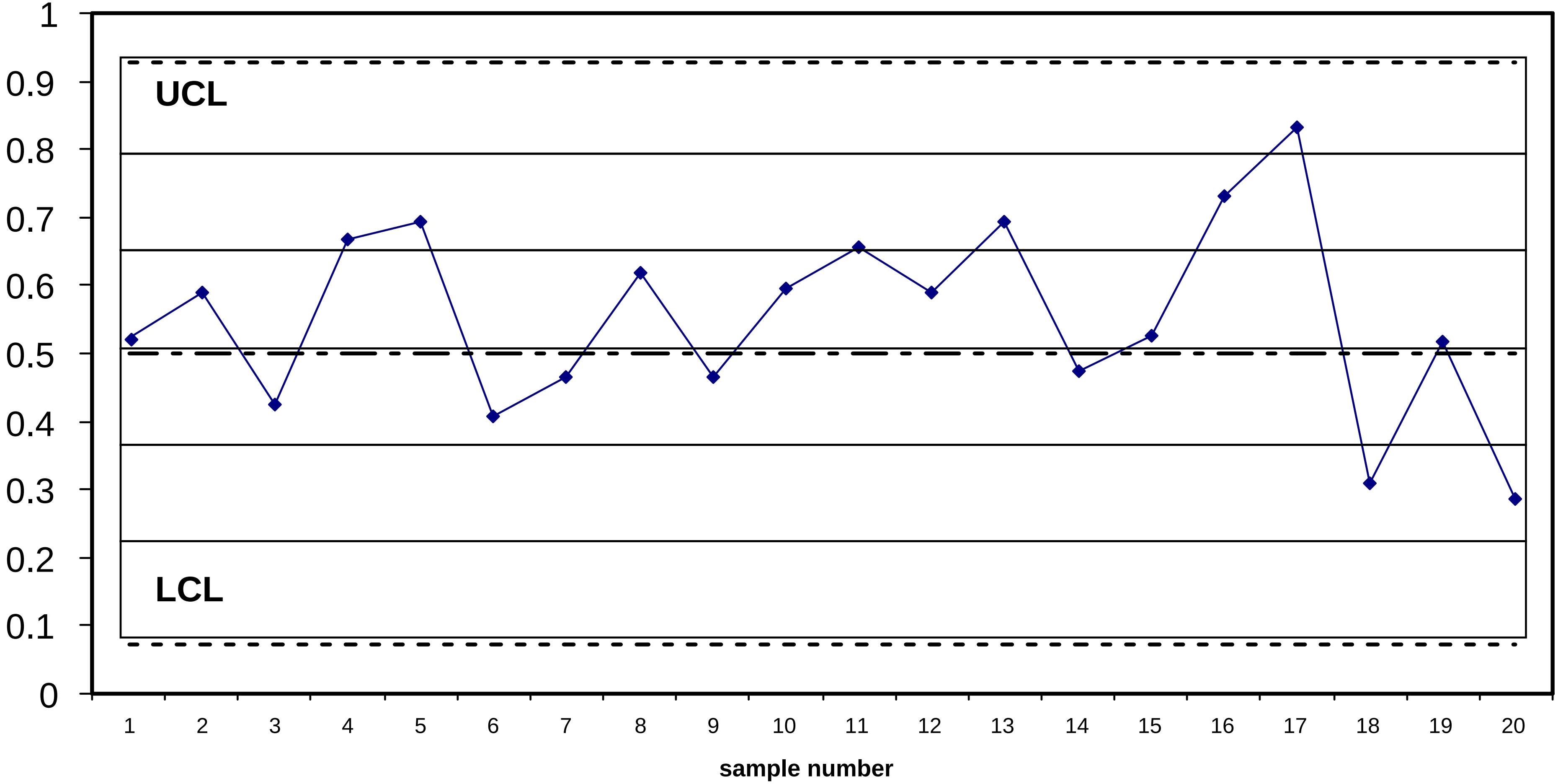
  (μ or σ are not constant)
- Process Output is being “caused” by a disturbance (assignable or special cause)
- This disturbance can be identified and eliminated
  - Trends indicate certain types
  - Correlation with know events
    - shift changes
    - material changes

Use of the S Chart

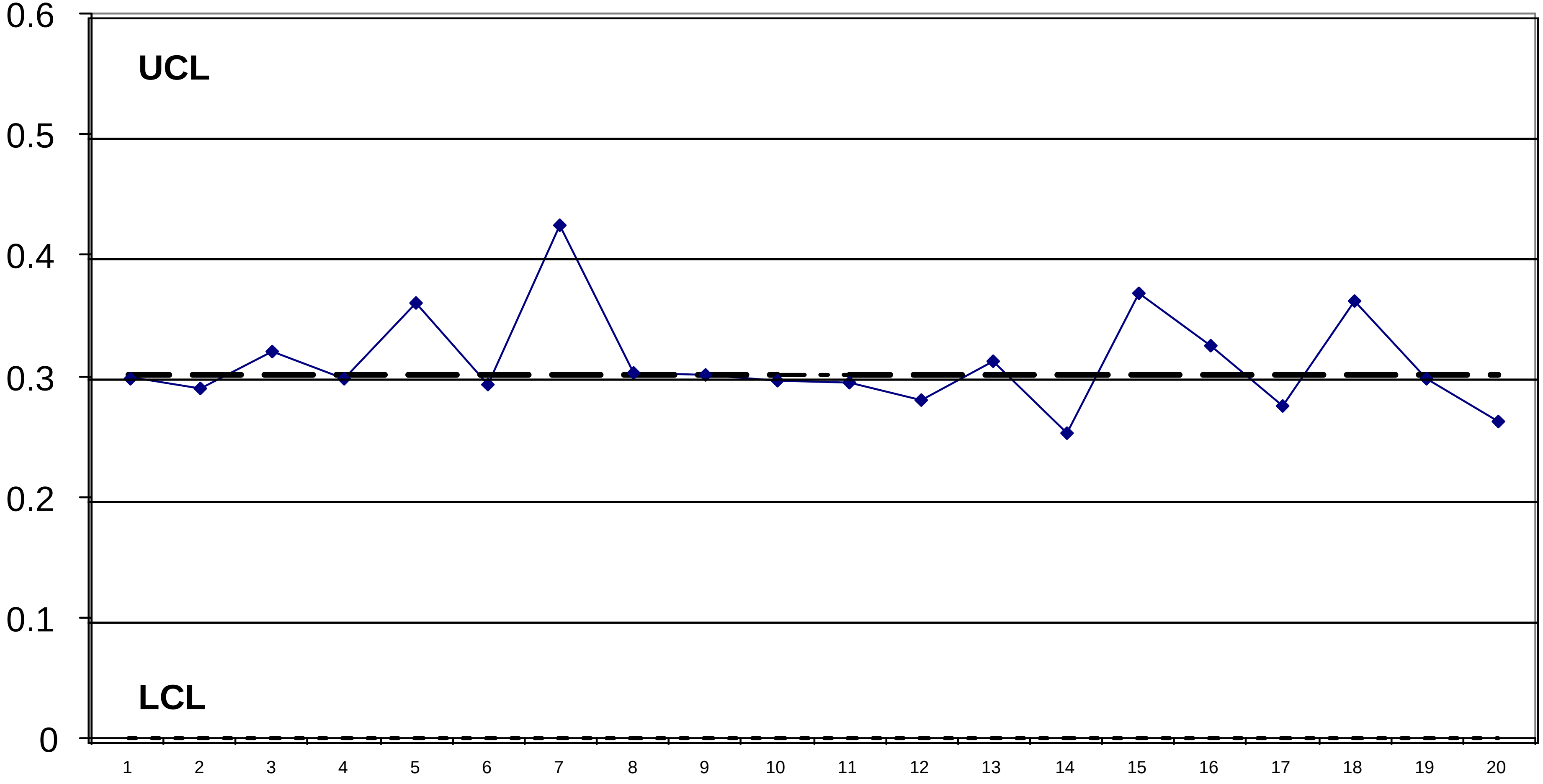
- Plot of sample Variance
  - Variance of the Mean for Shewhart xbar (n>1)

- What Does it Tell Us about State of Control?
  - It simply plots the “other” statistic

Consider this Process
Xbar Chart



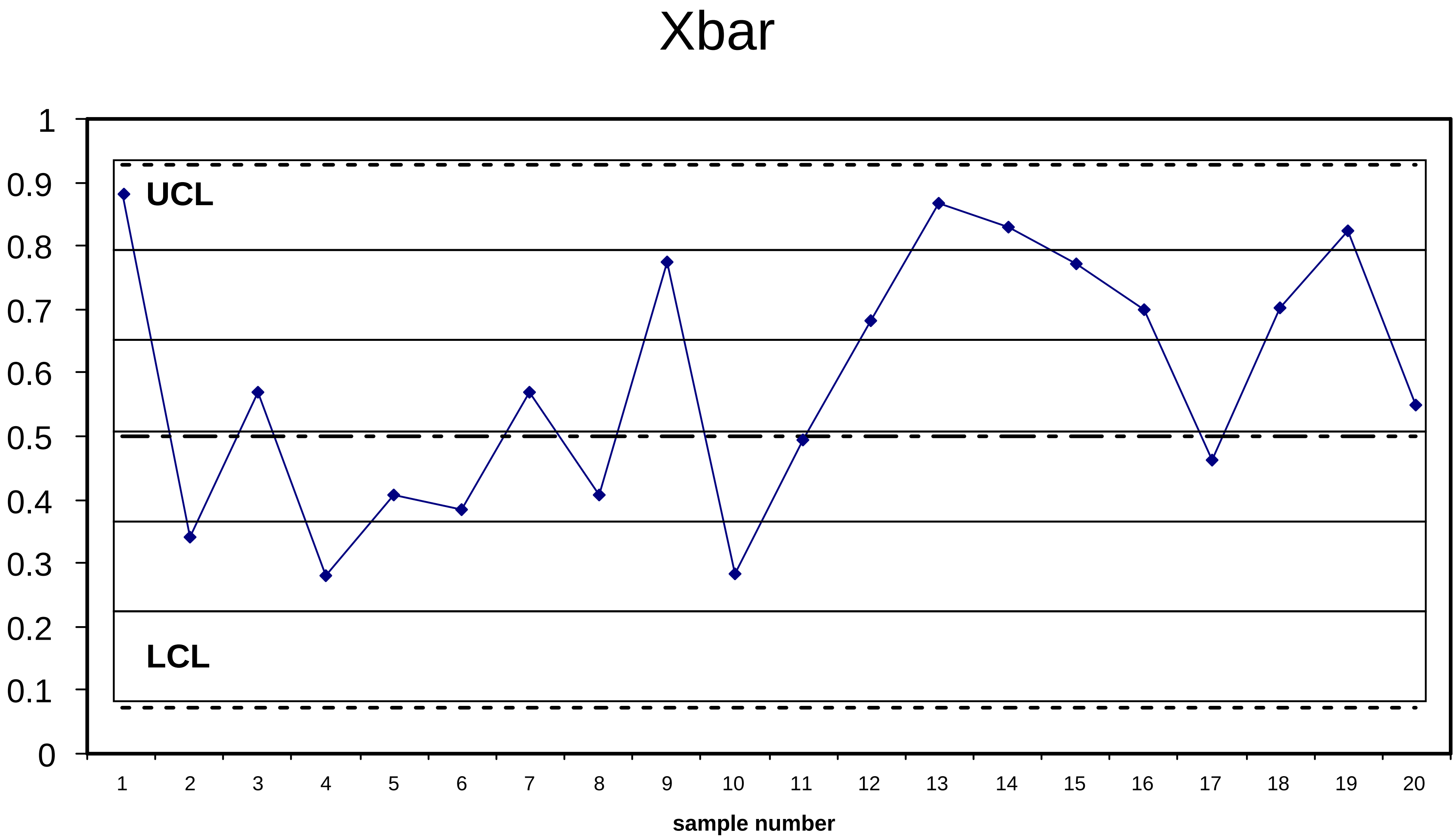
And the S Chart



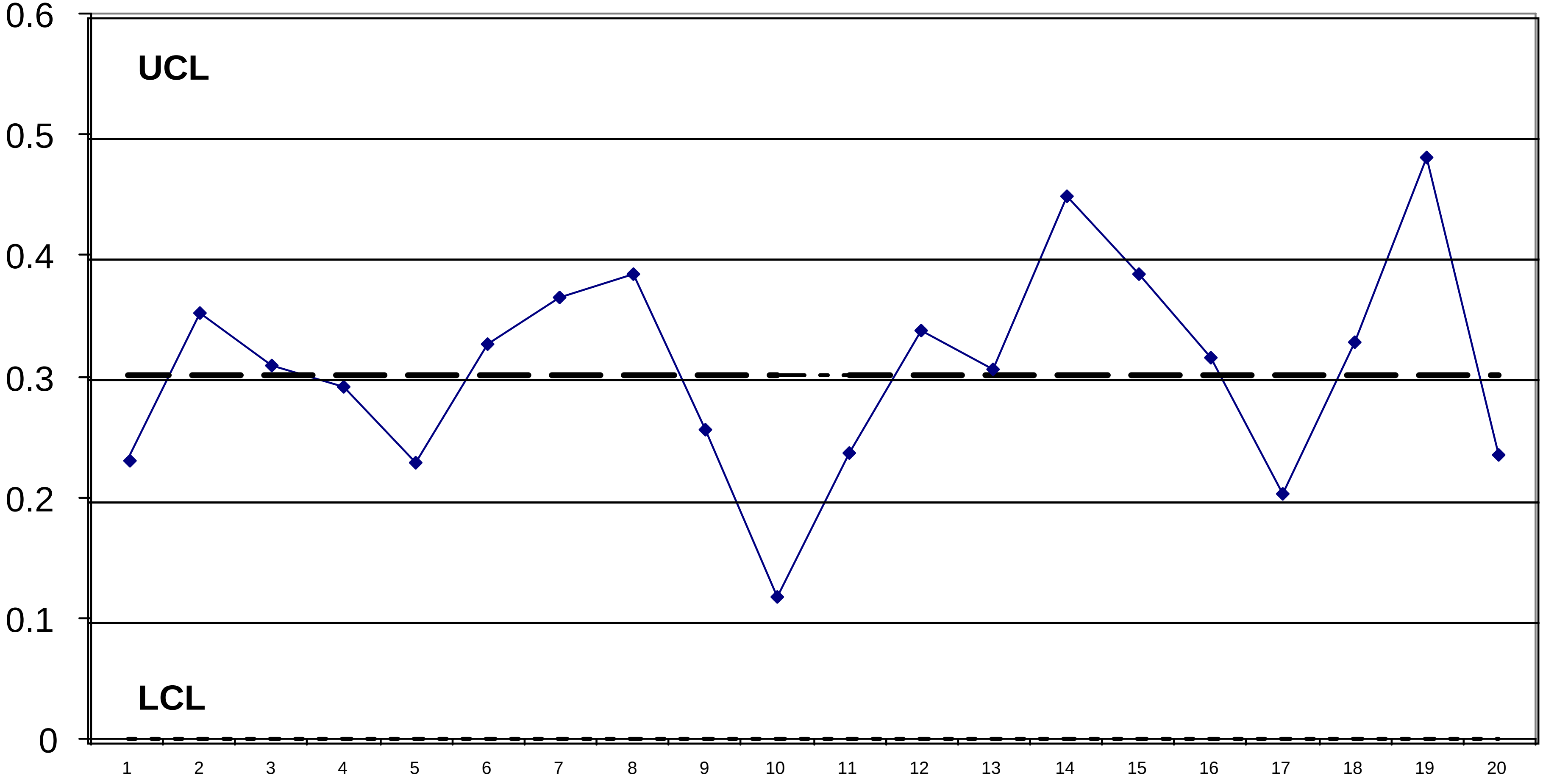
Manufacturing

In Control?

Same Process Later in Time



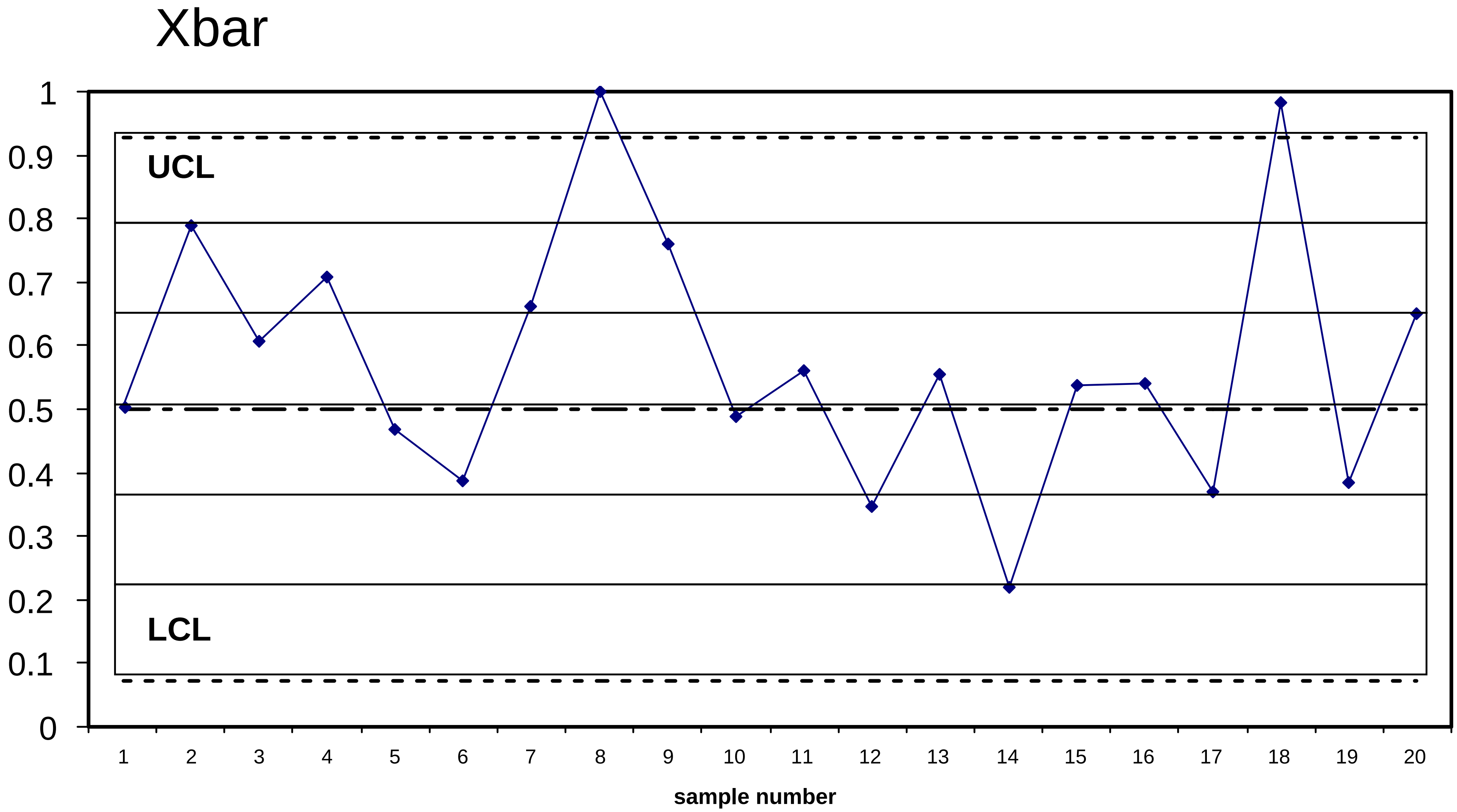
Later S Chart



Manufacturing

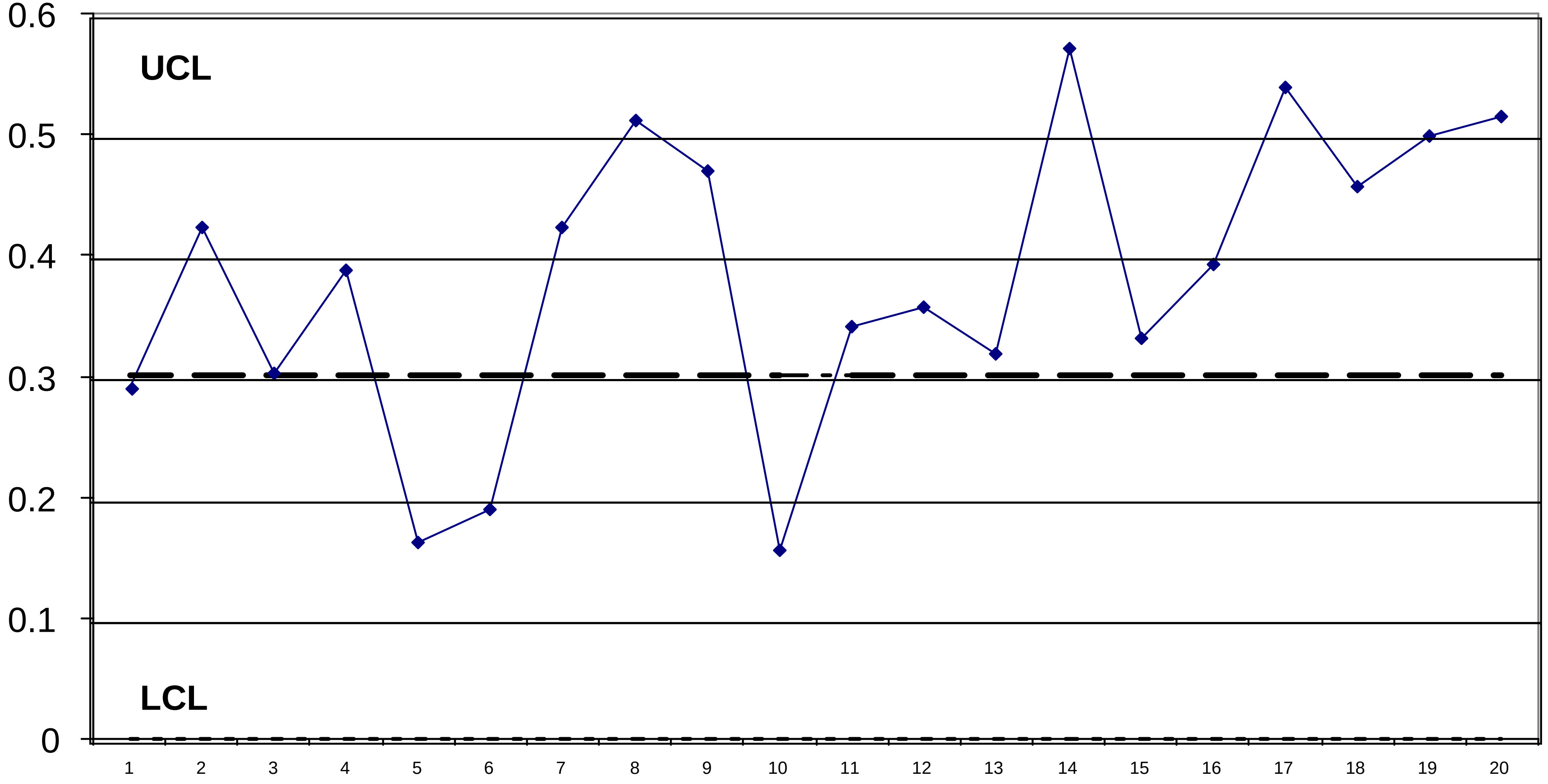
What Changed??

A Different Sequence



Manufacturing

S Chart


0.6
0.5
0.4
0.3
0.2
0.1
0
UCL
LCL
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

Use of S Chart

- Detect Changes in Variance of Parent Distribution
- Distinguish Between Mean and Variance Changes

Statistical Process Control

- Model Process as a Normal *Independent** Random Variable
- <u>Completely</u> described by μ and σ
- Estimate using *xbar* and *s*
- Enforce Stationary Conditions
- Look for Deviations in Either Statistic
- If so ………..?
- **Call an Engineer!**

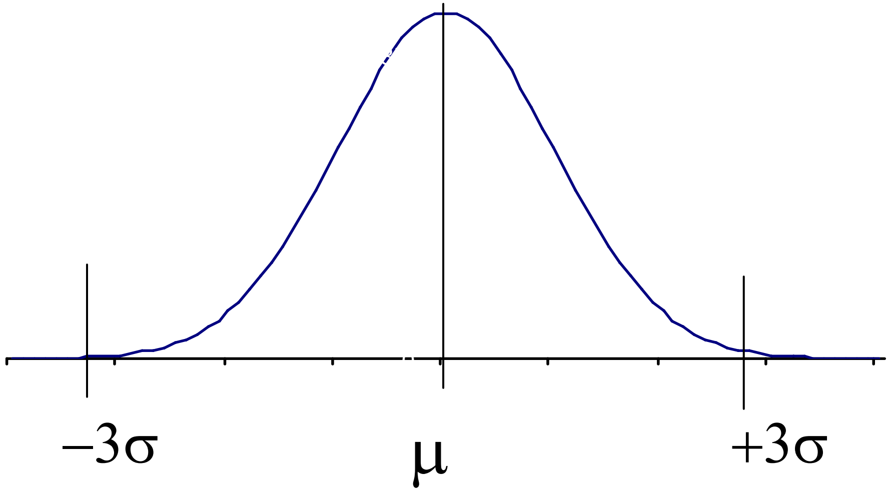
Another Use of the Statistical Process Model:

The Manufacturing -Design Interface

- We now have an empirical model of the process

How “good” is the process?

Is it capable of producing what we need?

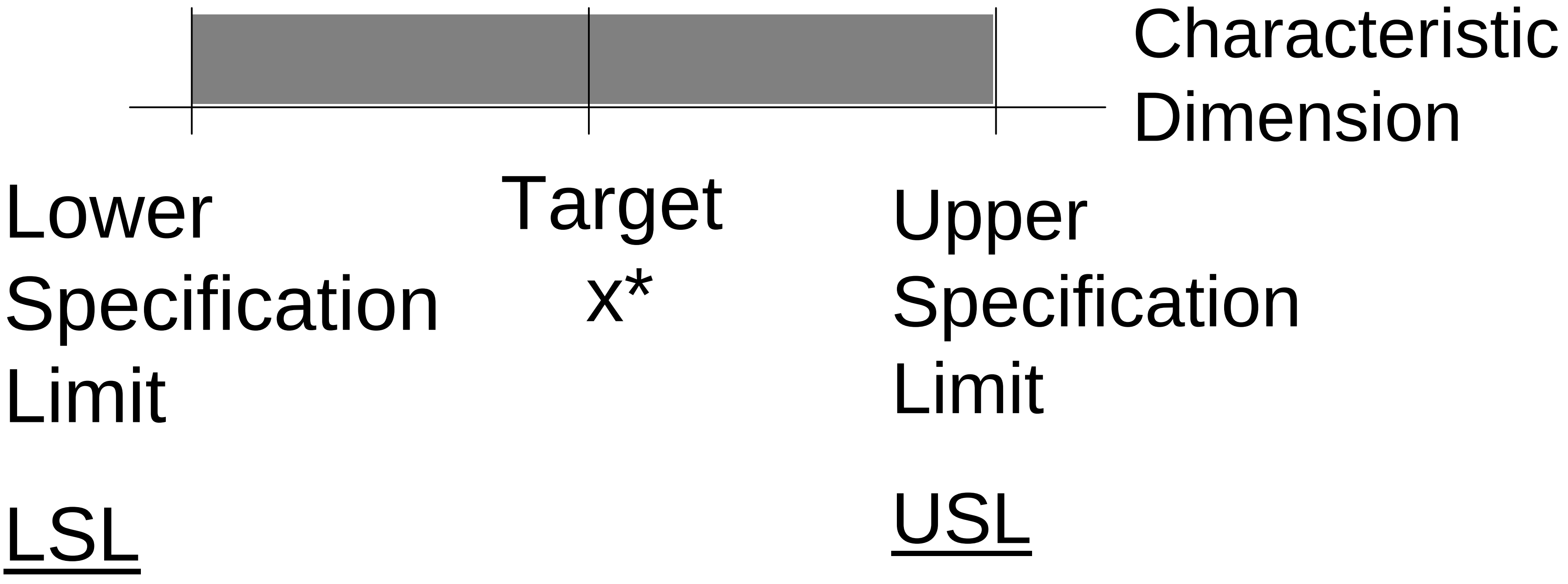


Process Capability

- Assume Process is In-control
- Described fully by *xbar* and *s*
- Compare to Design Specifications
  - Tolerances
  - Quality Loss

Design Specifications

- **Tolerances**: Upper and Lower Limits

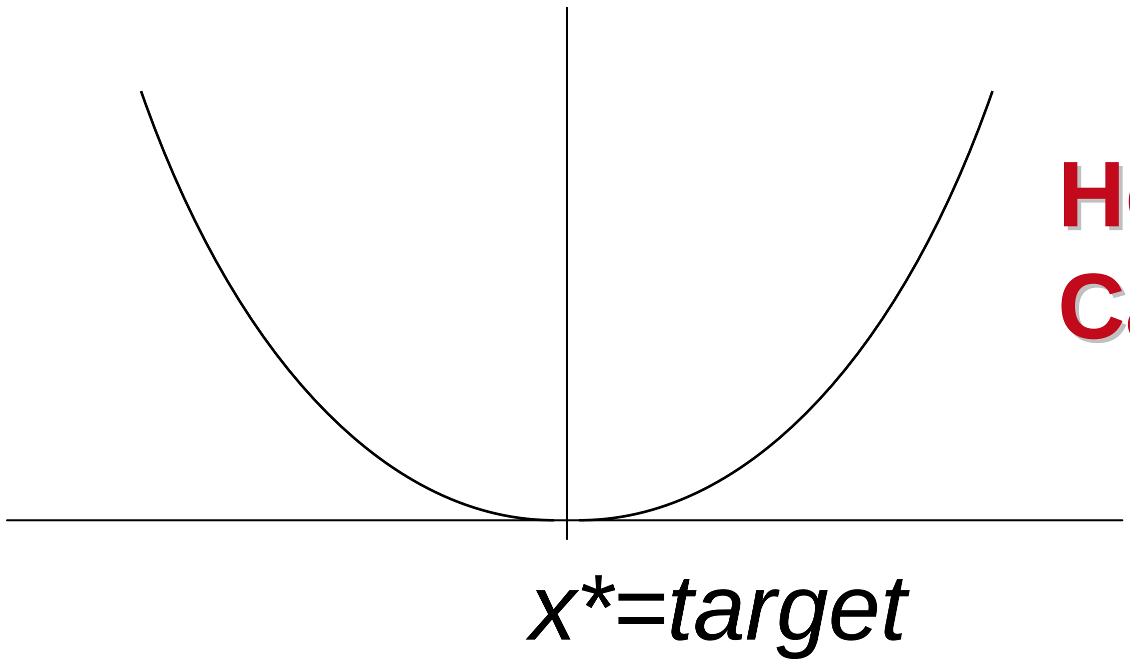


Manufacturing

Design Specifications

- **Quality Loss**: Penalty for Any Deviation from Target

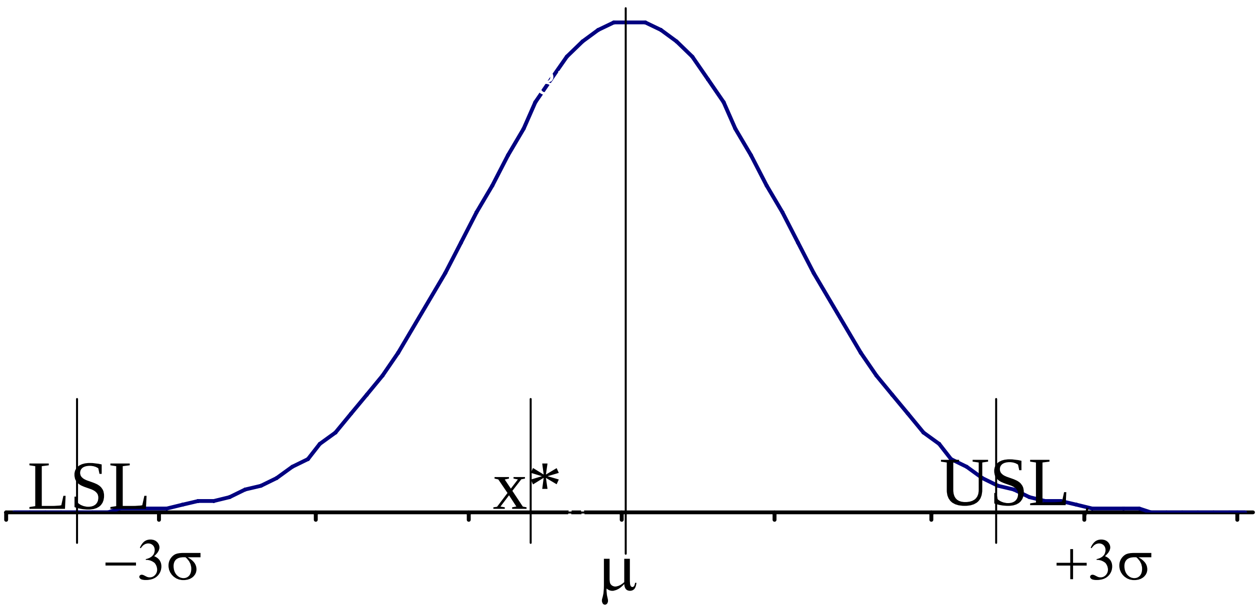
$$QLF = L^*(x-x^*)^2$$



How to Calibrate?

Use of Tolerances: Process Capability

- Define Process using a Normal Distribution
- Superimpose x*, LSL and USL
- Evaluate Expected Performance



Process Capability

- Definitions

$$C_p = \frac{(USL - LSL)}{6\sigma} = \frac{\text{tolerance range}}{\text{99.97\% confidence range}}$$

- Compares ranges only
- No effect of a mean shift:

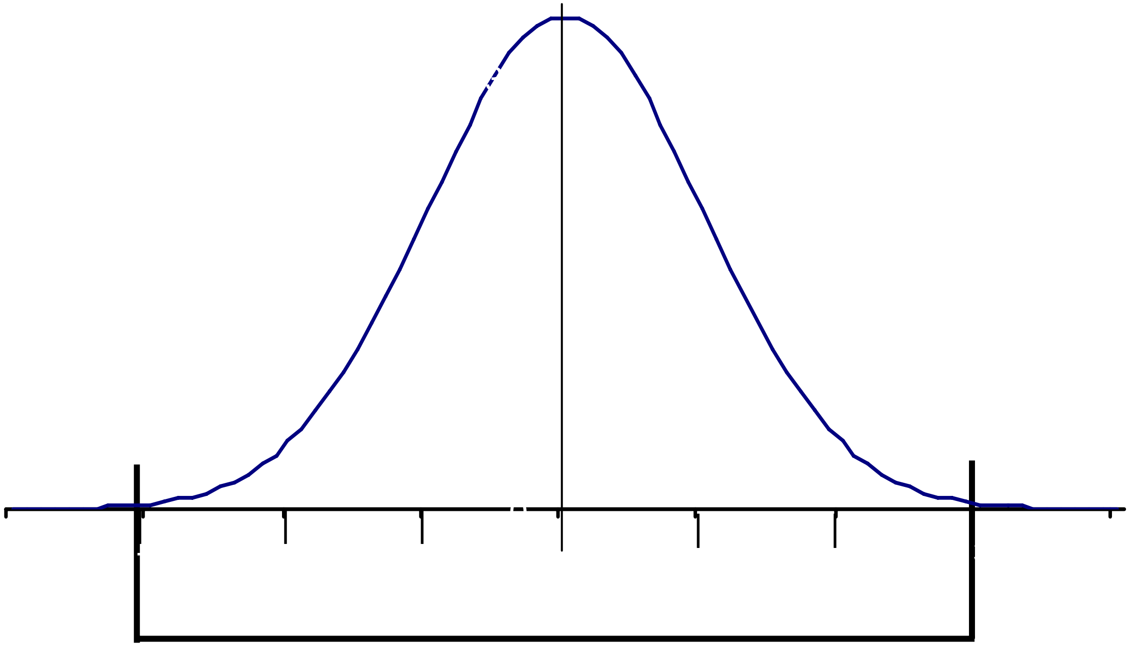
Process Capability: C_{pk}

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

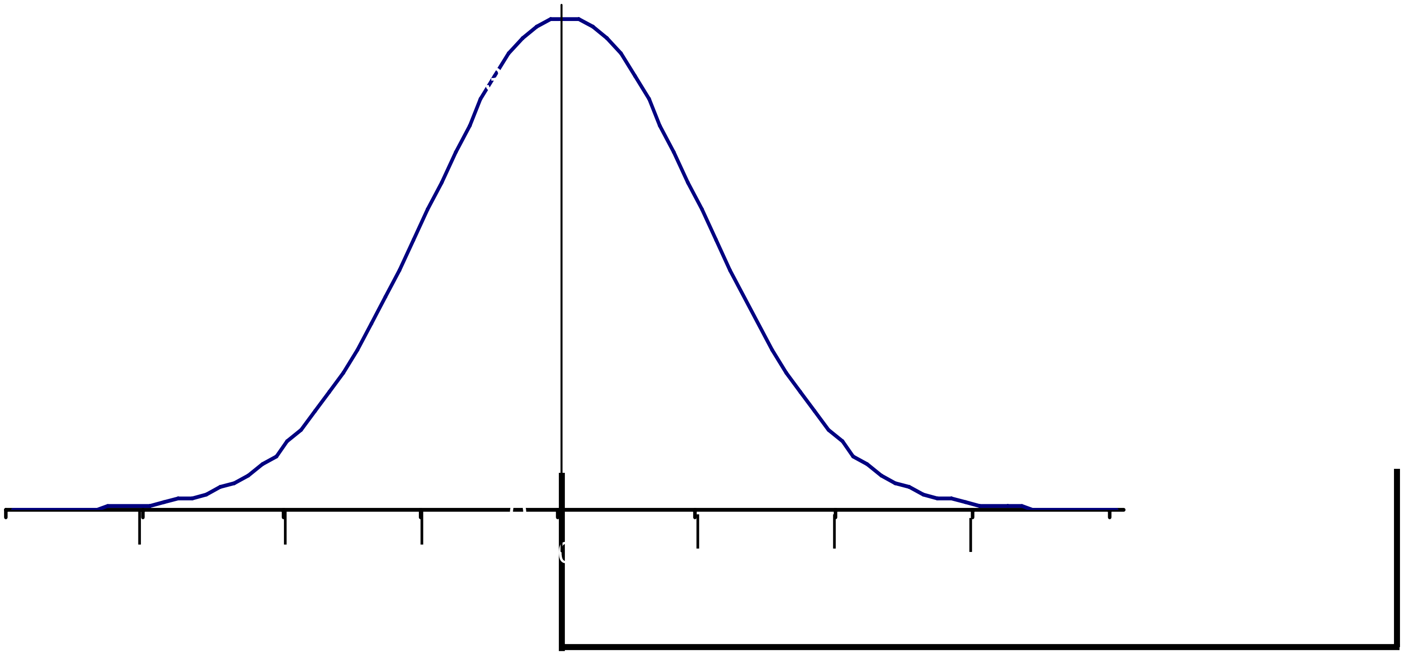
= Minimum of the normalized deviation from the mean

- Compares effect of offsets

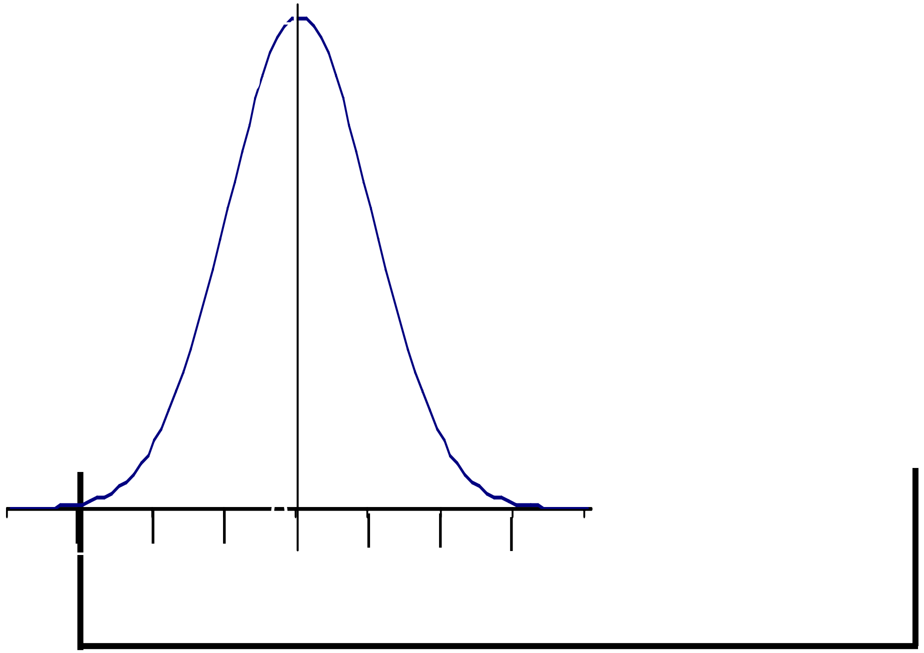
Cp = 1; Cpk = 1



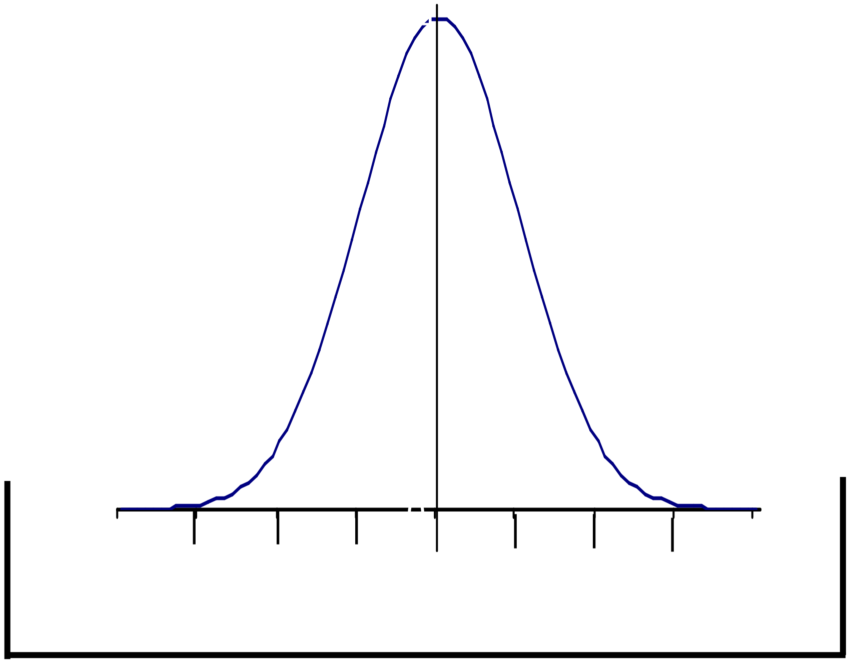
Cp = 1; Cpk = 0



Cp = 2; Cpk = 1



Cp = 2; Cpk = 2



Effect of Changes

- In Design Specs
- In Process Mean
- In Process Variance

- What are good values of Cp and Cpk?

Cpk Table

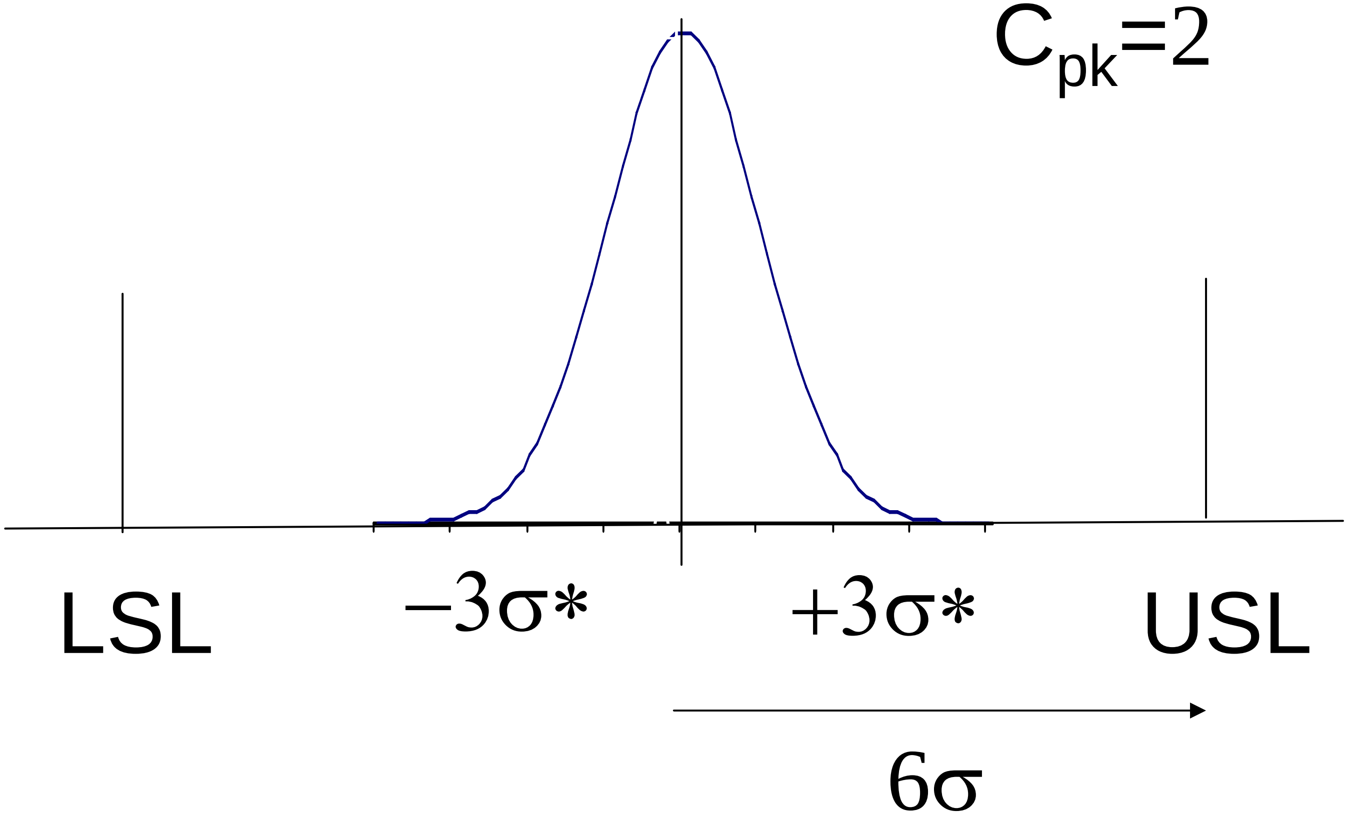
Cpk	z	P<LS or P>USL
1	3	1E-03
1.33	5	3E-07
1.67	4	3E-05
2	6	1E-09

The “6 Sigma” problem

$P(x > 6\sigma) = 18.8x10^{-10}$

$C_p=2$

$C_{pk}=2$



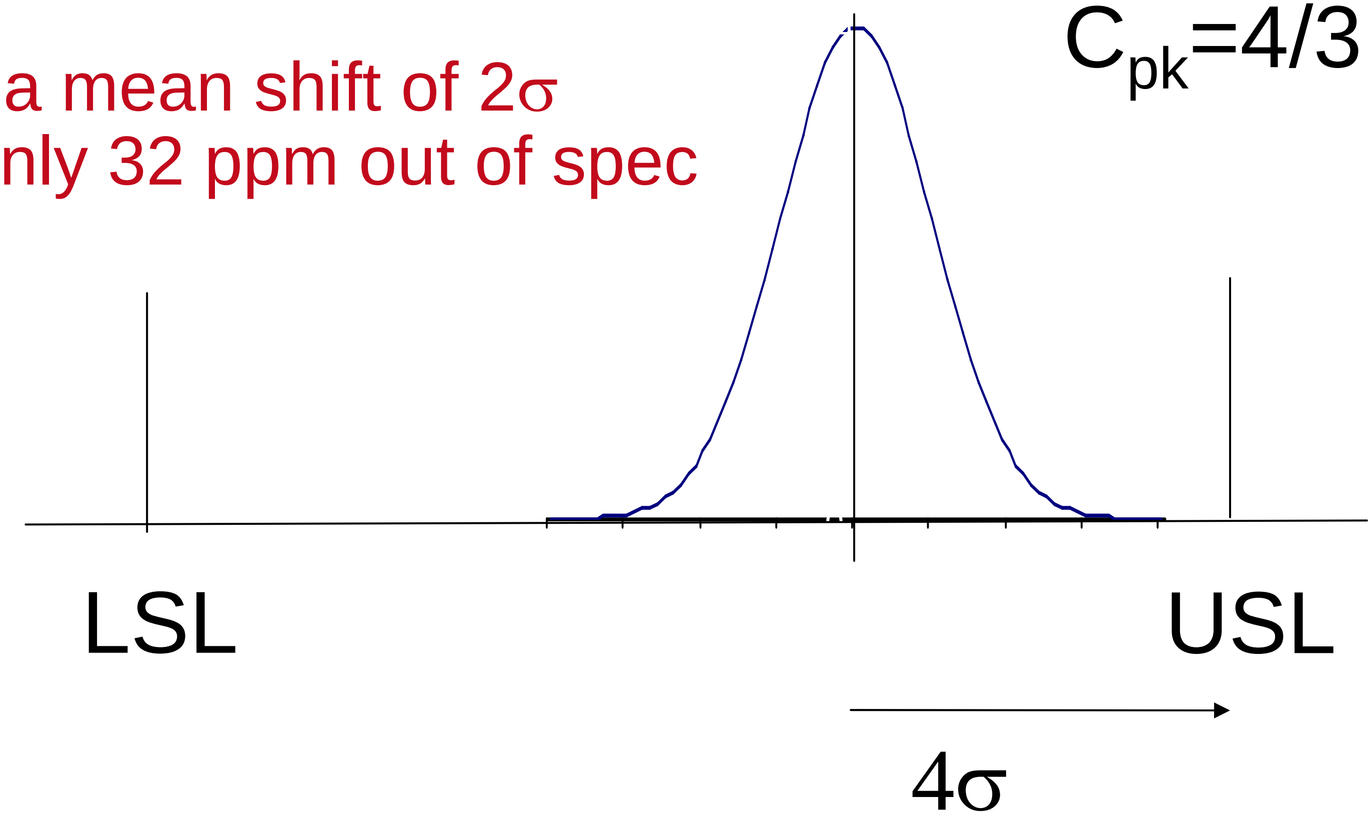
The 6 σ problem: Mean Shifts

$P(x>4\sigma) = 31.6\text{x}10^{-6}$

$C_p=2$

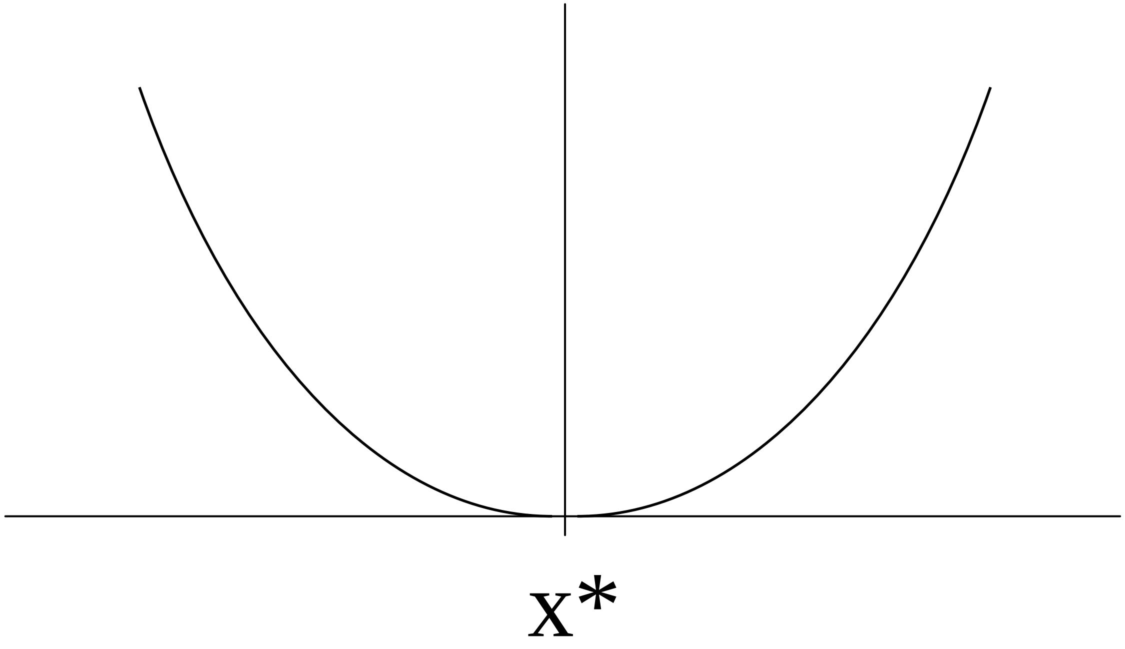
$C_{pk}=4/3$

Even with a mean shift of 2σ
we have only 32 ppm out of spec



Capability from the Quality Loss Function

$$QLF = L(x) = k*(x-x^*)^2$$



Given L(x) and p(x) what is E{L(x)}?

Expected Quality Loss

$$E\{L(x)\} = E\left[k(x - x^*)^2\right]$$

$$= k\left[E(x^2) - 2E(xx^*) + E(x^{*2})\right]$$

$$= k\sigma_x^2 + k(\mu_x - x^*)^2$$

Penalizes Variation

Penalizes Deviation

Process Capability

- The reality (the process statistics)
- The requirements (the design specs)
- Cp - a measure of variance vs. tolerance
- Cpk - a measure of variance from target
- Expected Loss- An overall measure of goodness