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 #8

Process Capability & Alternative SPC Methods

March 4, 2008

Agenda

- Control Chart Review
  - hypothesis tests: α, β and n
  - control charts: α, β, n, and average run length (ARL)

- Process Capability

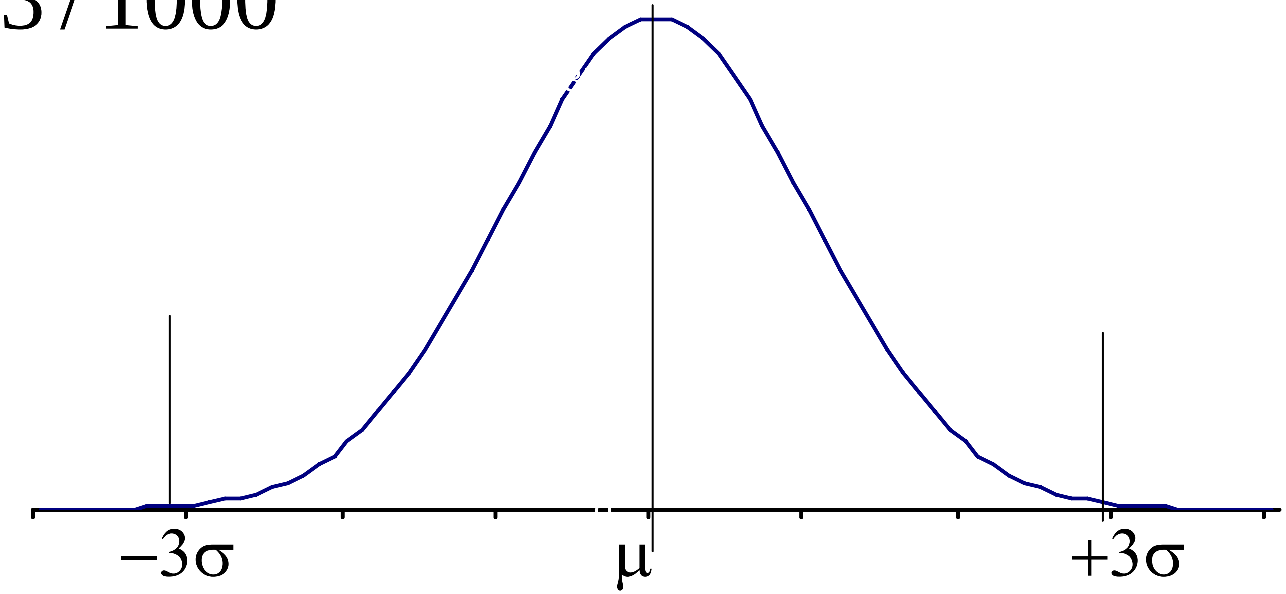
- Advanced Control Chart Concepts

Average Run Length

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

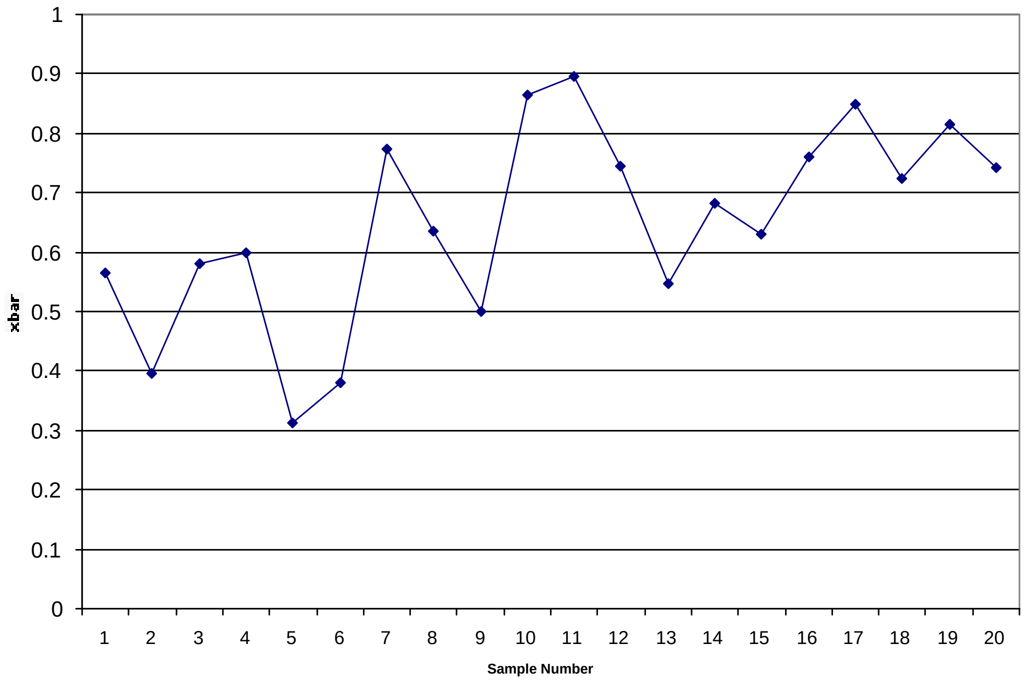
$$\mathrm{Prob}(x > \mu_x + 3\sigma_{\bar{x}}) + \mathrm{Prob}(x < \mu_x - 3\sigma_{\bar{x}})$$

$$= 3/1000$$



Detecting Mean Shifts: Chart Sensitivity

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



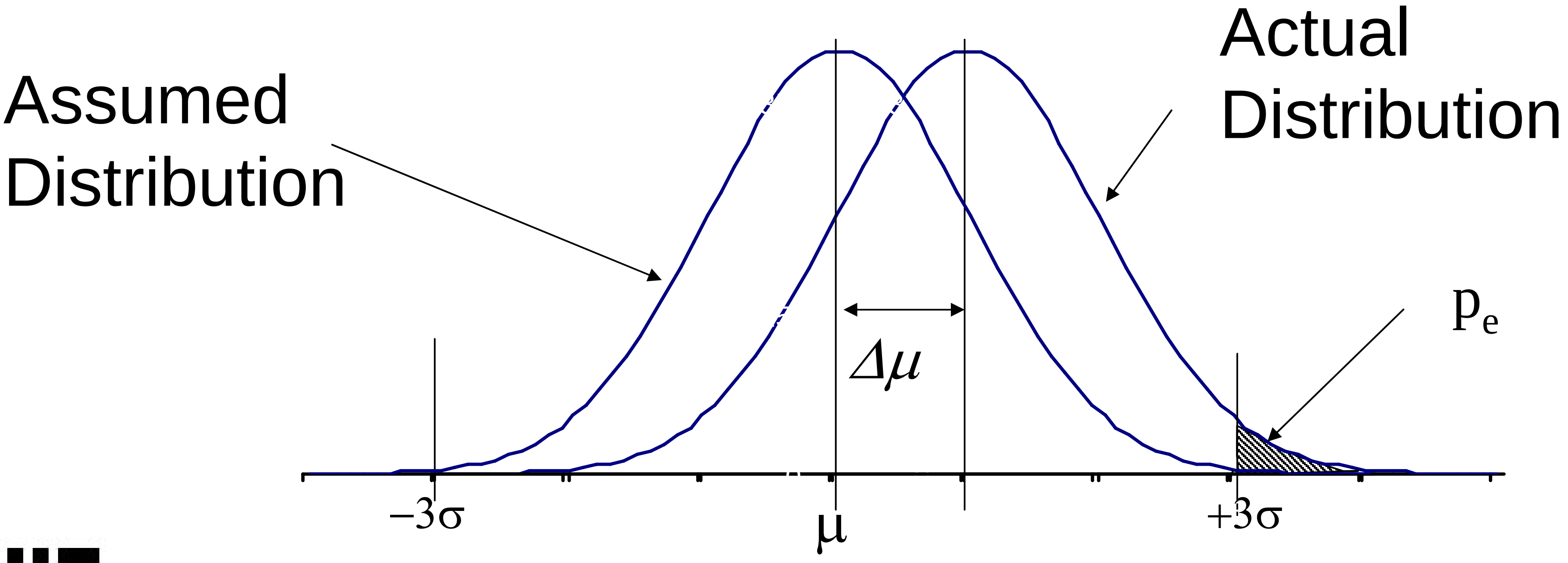
- How many samples before we can expect to detect the shift on the xbar chart?

Average Run Length

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

$$\text{Prob}(x > \mu_x + 2\sigma_{\bar{x}}) + \text{Prob}(x < \mu_x - 4\sigma_{\bar{x}})$$

$$= 0.023 + 0.001 = 24/1000$$



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

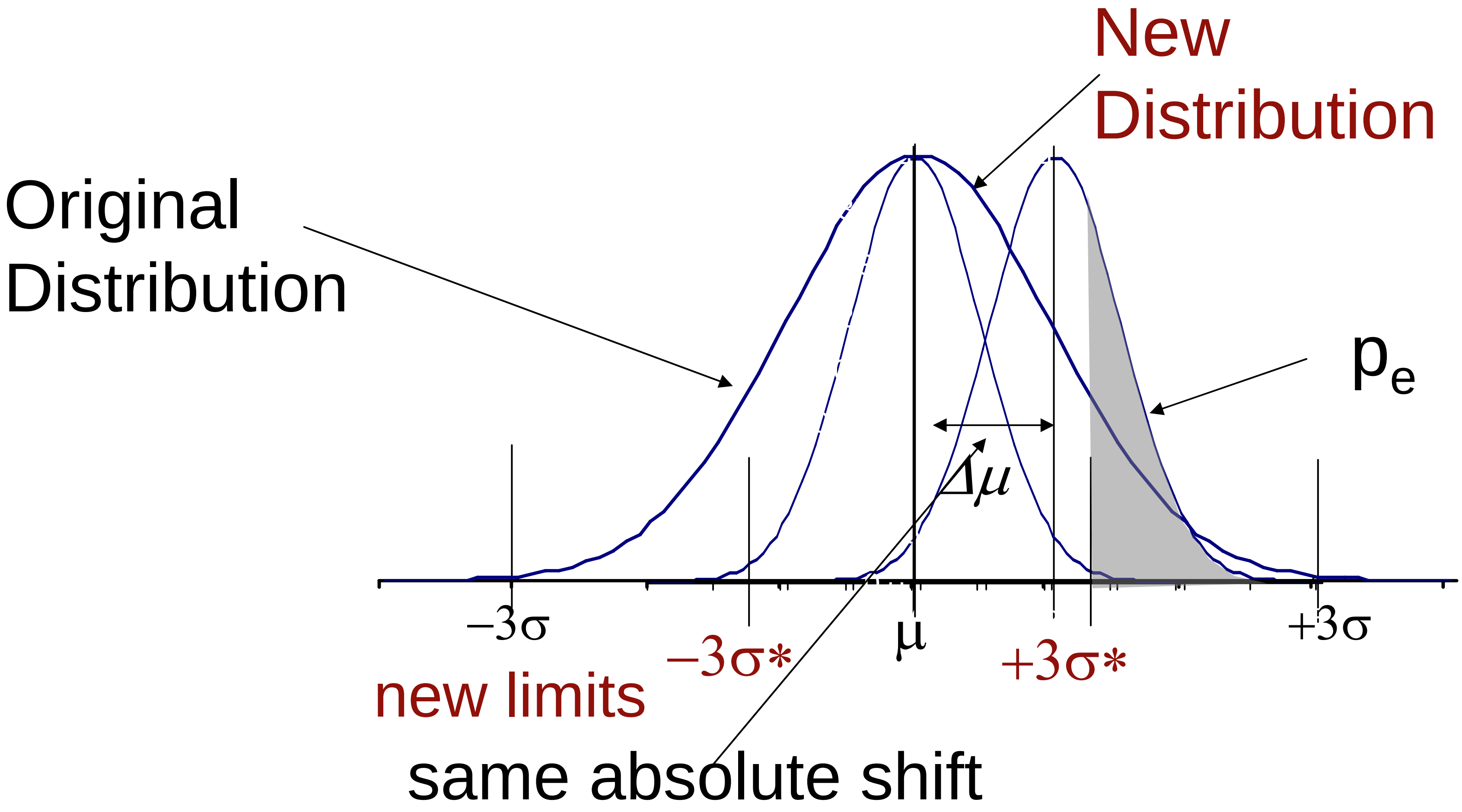
- Assume the same Δμ = 1σ
  - Note that Δμ is an absolute value

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

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

- So our ± 3σ limits move closer together

ARL Example



As *n* increases p_e increases so ARL decreases

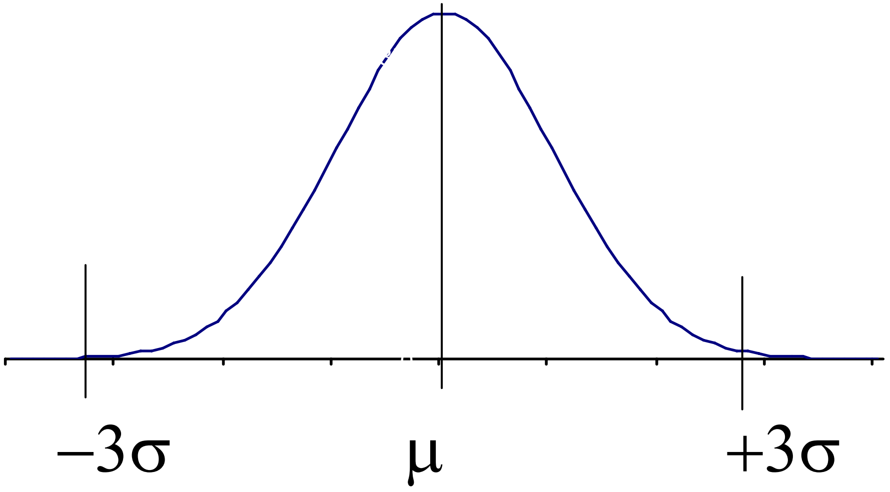
Another Use of the Statistical Process Model:

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?

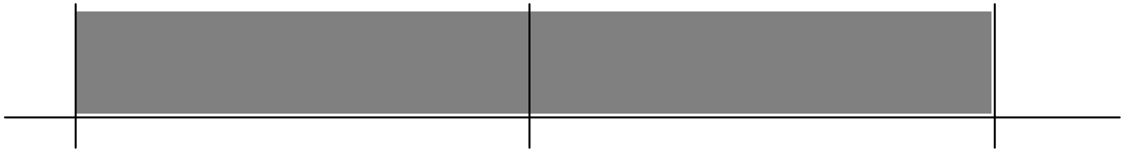


Process Capability

- 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



Characteristic Dimension

Lower Specification Limit

LSL

Target x*

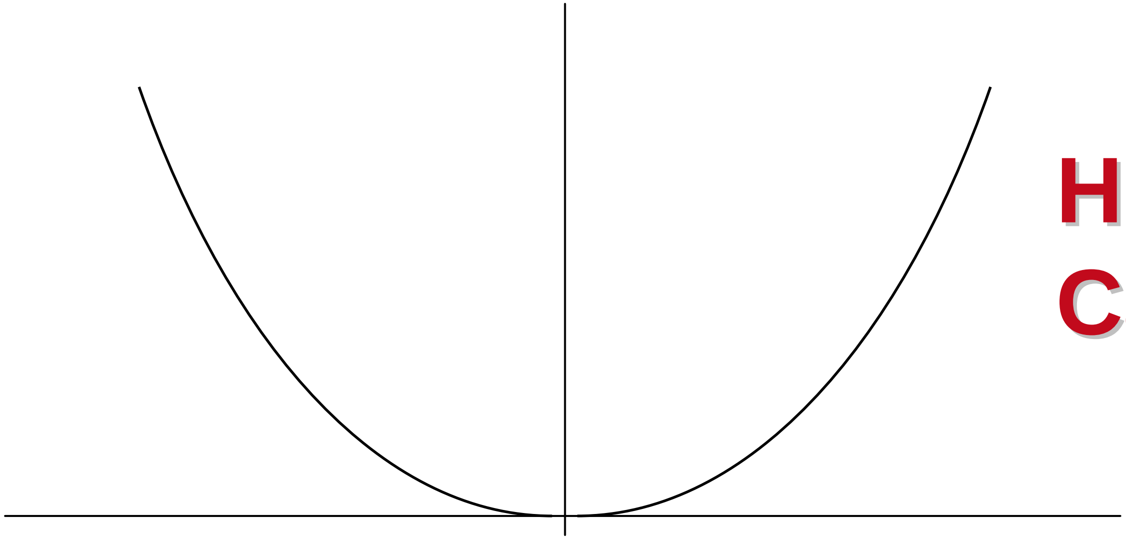
Upper Specification Limit

USL

Design Specifications

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

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

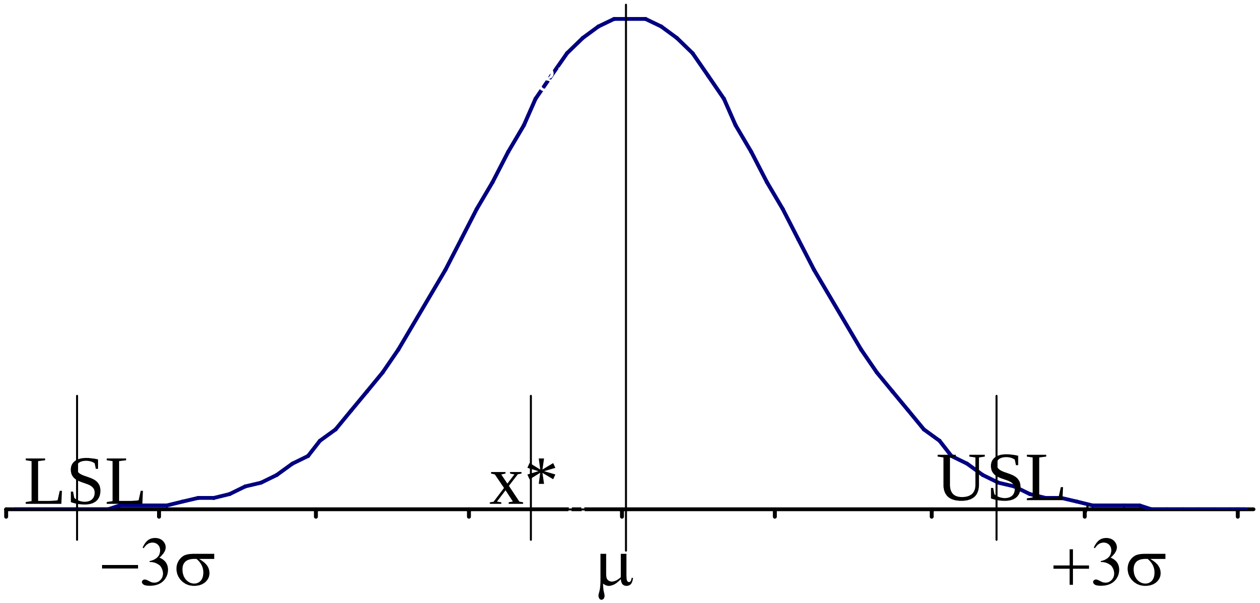


x^*=*target*

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

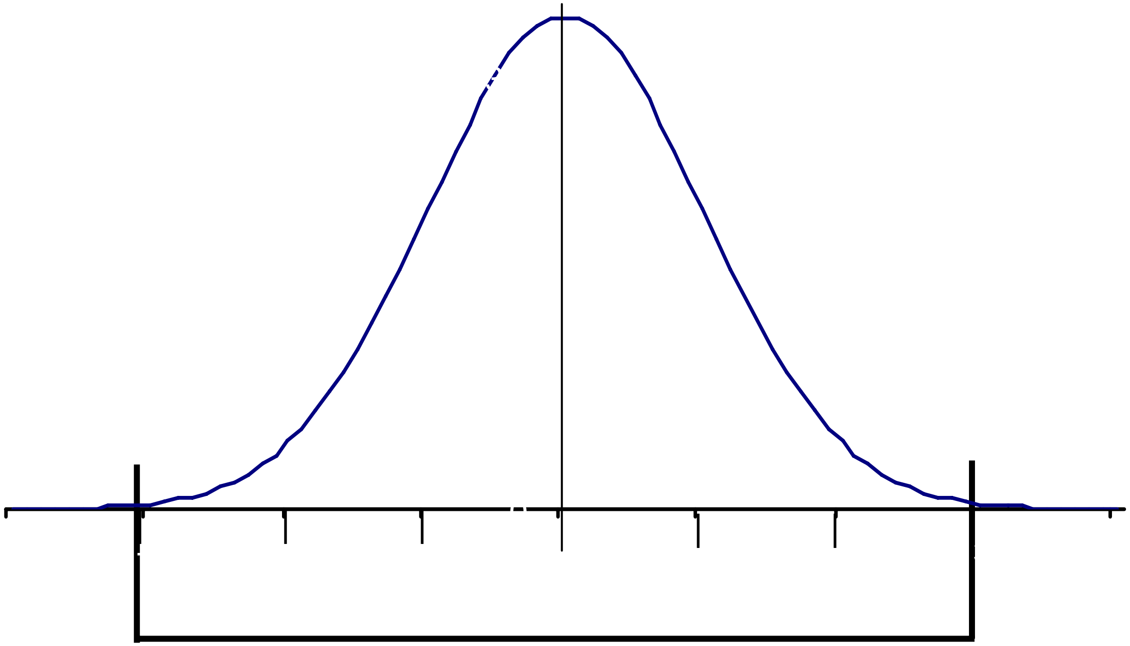
Process Capability: C_{pk}

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

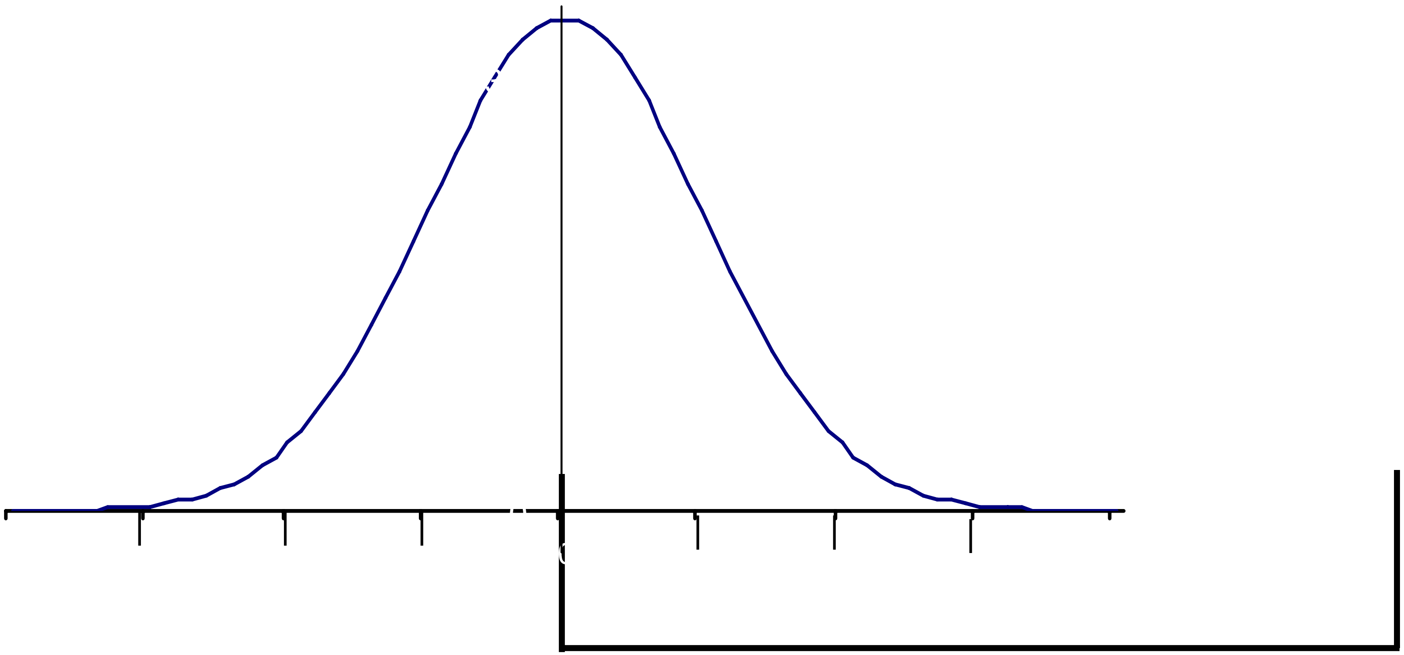
= Minimum of the normalized deviation from the mean

- Compares effect of offsets

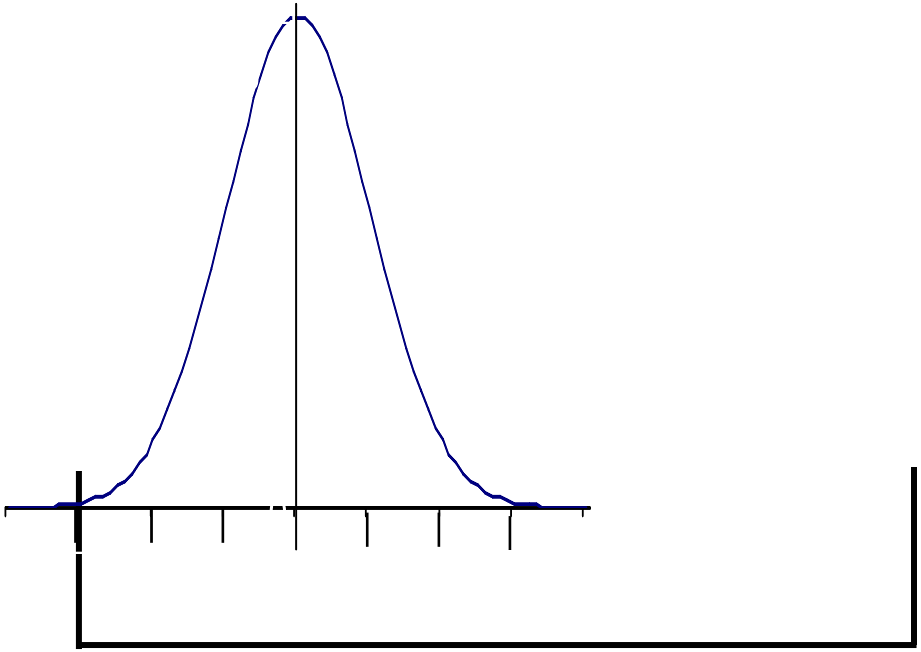
Cp = 1; Cpk = 1



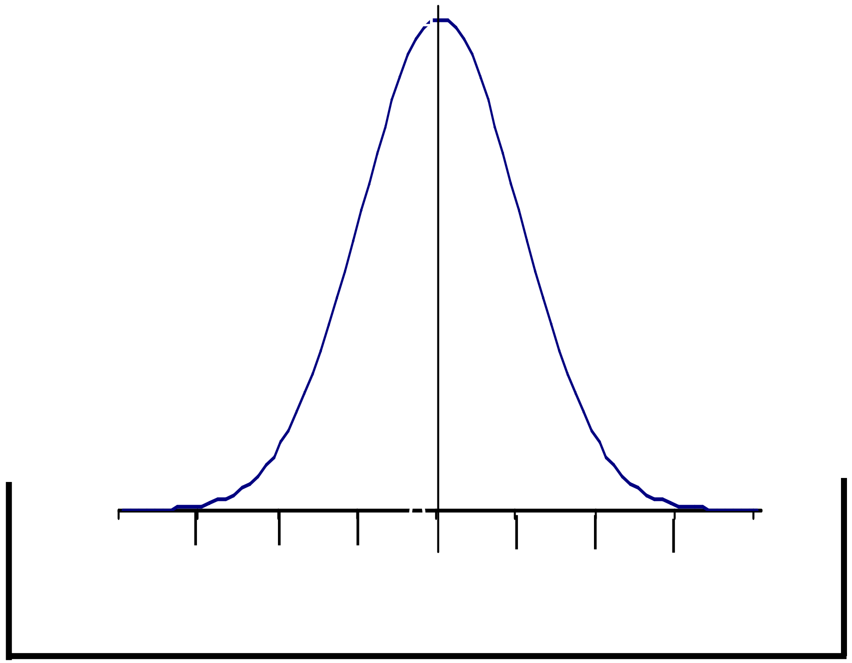
Cp = 1; Cpk = 0



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

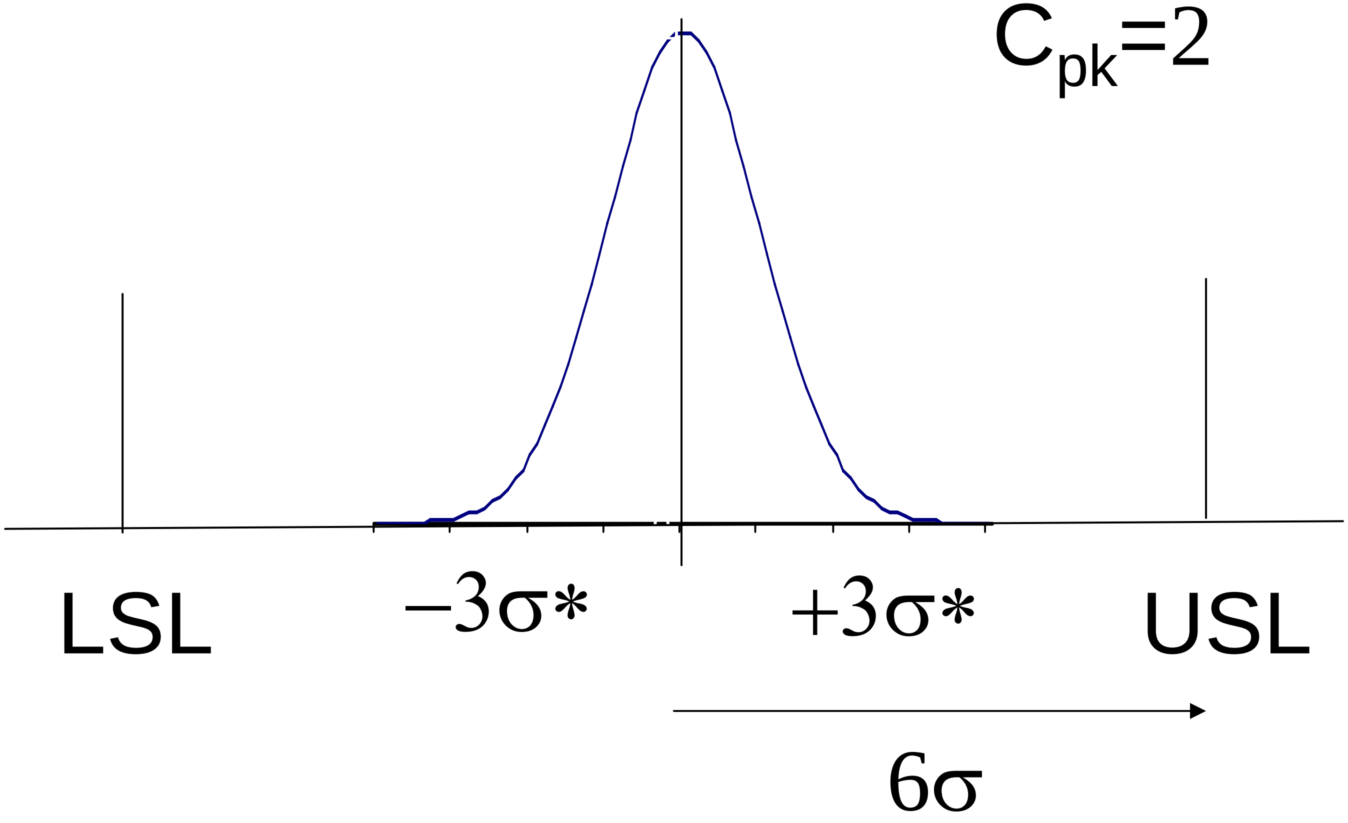
Cpk	z	P<LS or P>USL
1	3	1E-03
1.33	4	3E-05
1.67	5	3E-07
2	6	1E-09

The “6 Sigma” problem

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

$C_p=2$

$C_{pk}=2$



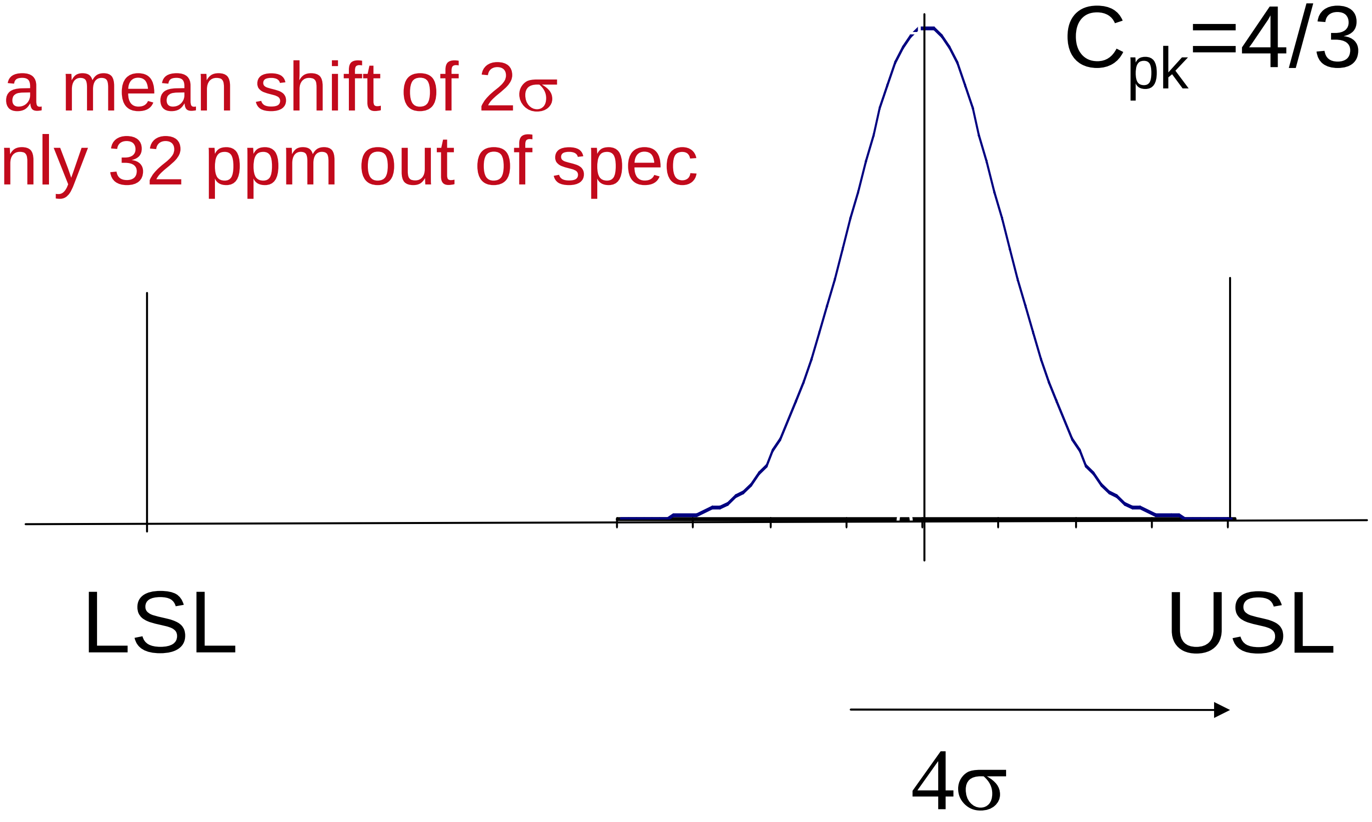
The 6 σ problem: Mean Shifts

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

$C_p=2$

$C_{pk}=4/3$

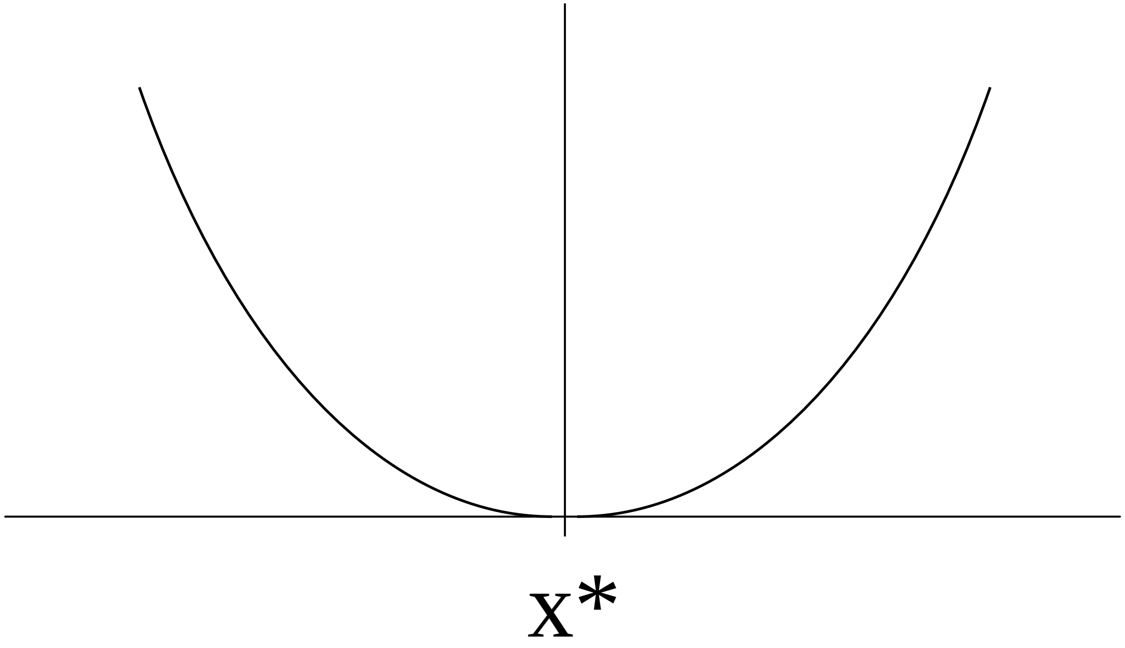
Even with a mean shift of 2σ
we have only 32 ppm out of spec



Manufacturing

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

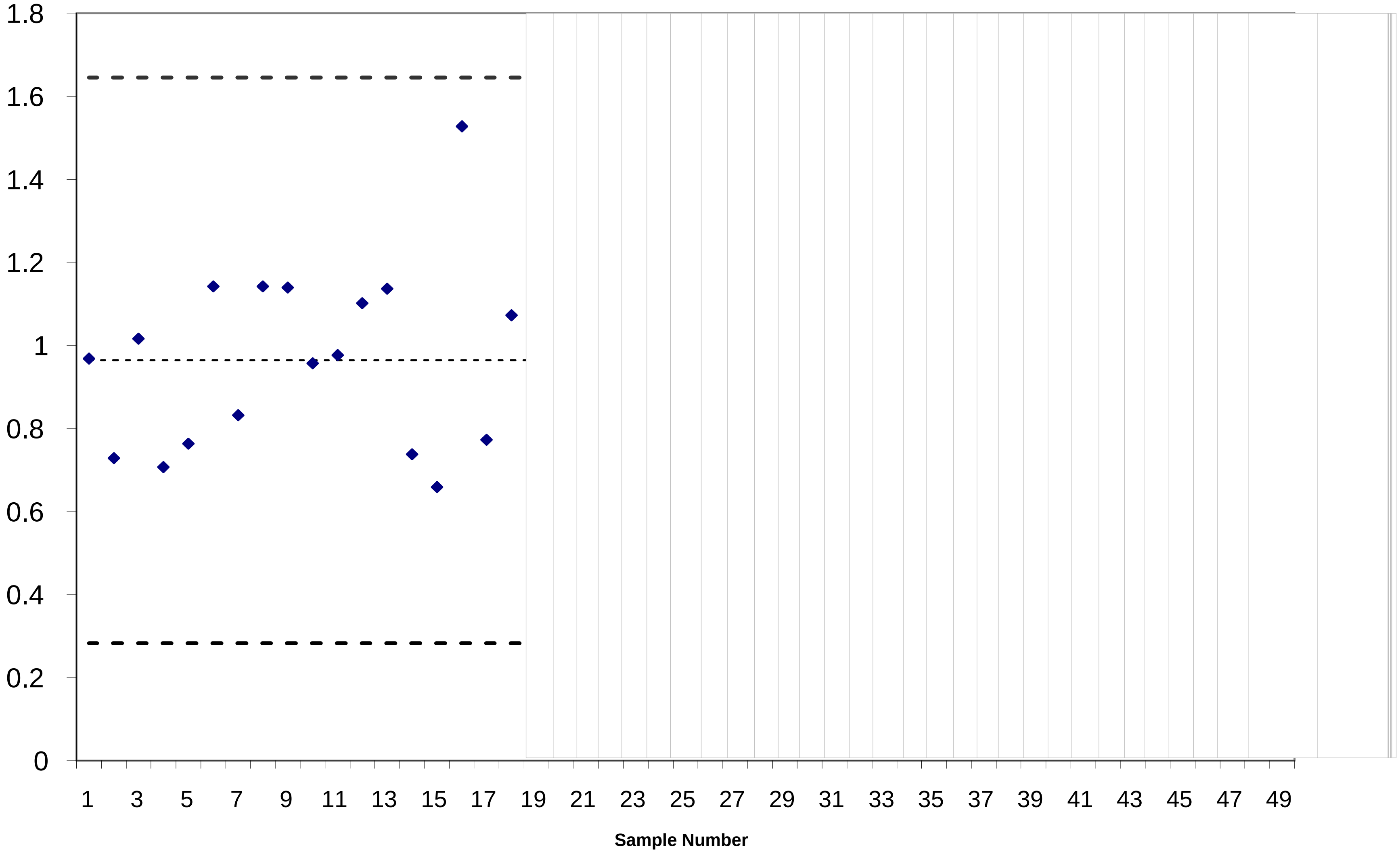
Process Capability

- 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

Xbar Chart Recap

- xbar - S (or R) charts
  - plot of sequential sample statistics
  - compare to assumptions
    - normal
    - stationary
- Interpretation
  - hypothesis tests on μ and σ
  - confidence intervals
  - “randomness”
- Application
  - Real-time decision making

Real-Time


1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
31
33
35
37
39
41
43
45
47
49
Sample Number

Beyond Xbar

- Good Points
  - Simple and “transparent”
  - Enforces Assumptions
    - Normality (via Central Limit)
    - Independent (via long sampling times)

- Limitations
  - n>1 to get Xbar and S
  - ARL is typically large
    - Not very sensitive to small changes
  - Slow time response

Beyond Xbar

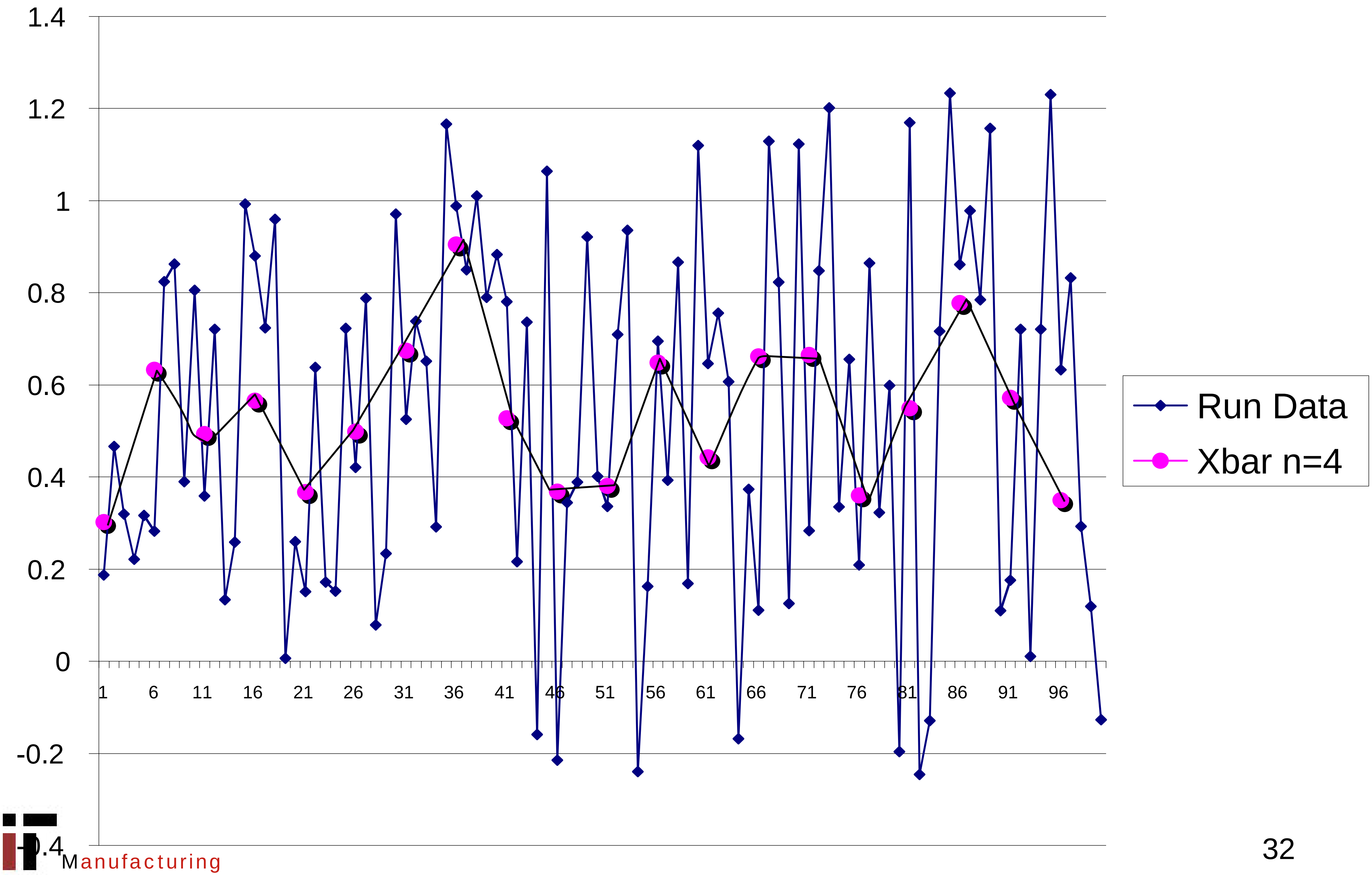
- What if n=1?
  - Have a Lot of Data
  - Want Fast Response to Changes
- How to Compute Control Chart Statistics?
  - Running Chart and Running Variance?
  - Running Average and Running Variance?
  - Running Average with Forgetting Factor
- How to Increase Sensitivity to Small, Persistent Mean Shift?
  - Integrate the Error

Chart Design:

n=1 Designs - Running Averages

- Sensitivity: Ability to detect small changes (e.g. mean shifts)
- Time Response: Ability to Catch Changes Quickly
- Noise Rejection?: Higher Variance

Xbar “Filtering”



Filtering

- Reduced Peaks
- Hides intermediate data
- Reduces the “frequency content” of the output

Independence and Correlation

- Independence: Current output does not depend on prior
- Correlation: Measure of Independence
  - e.g. auto correlation function

$$R_{xx}(\tau) = E[x(t)x(t+\tau)]$$

Correlation

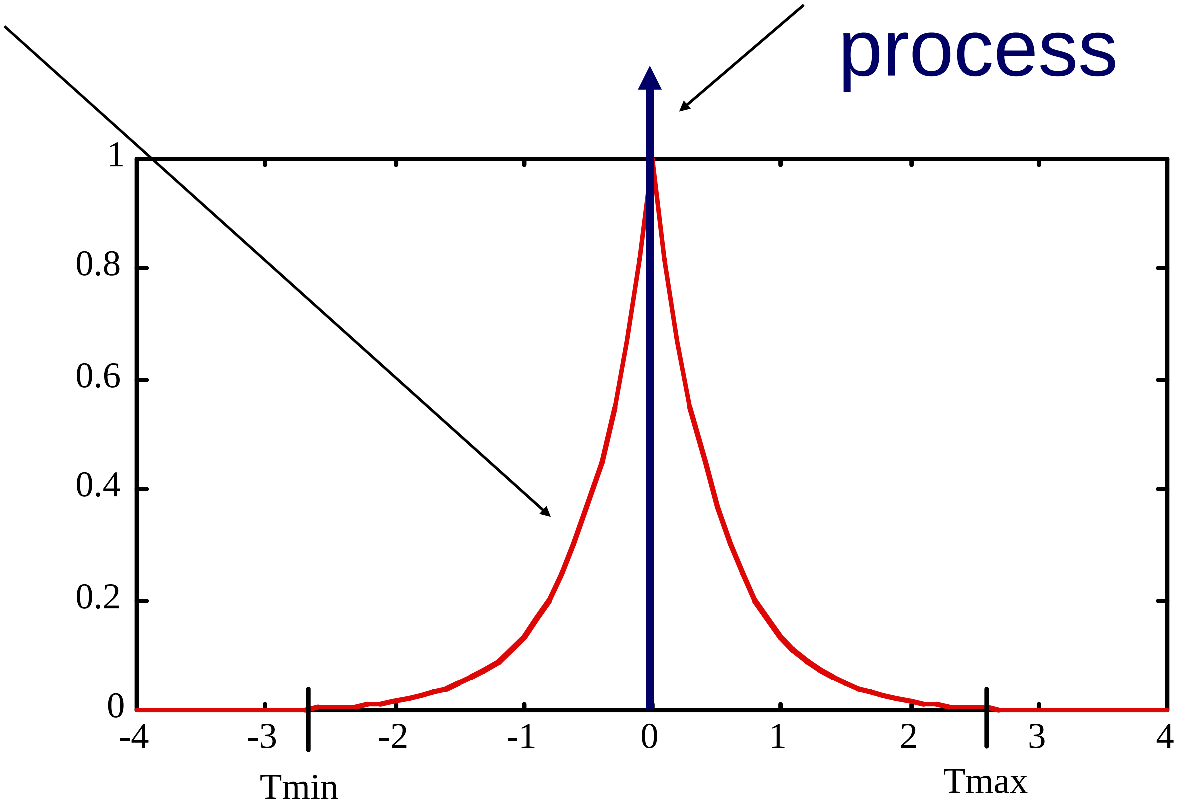
$$R_{xx}(\tau) = E[x(t)x(t+\tau)]$$

For a linear 1st order system

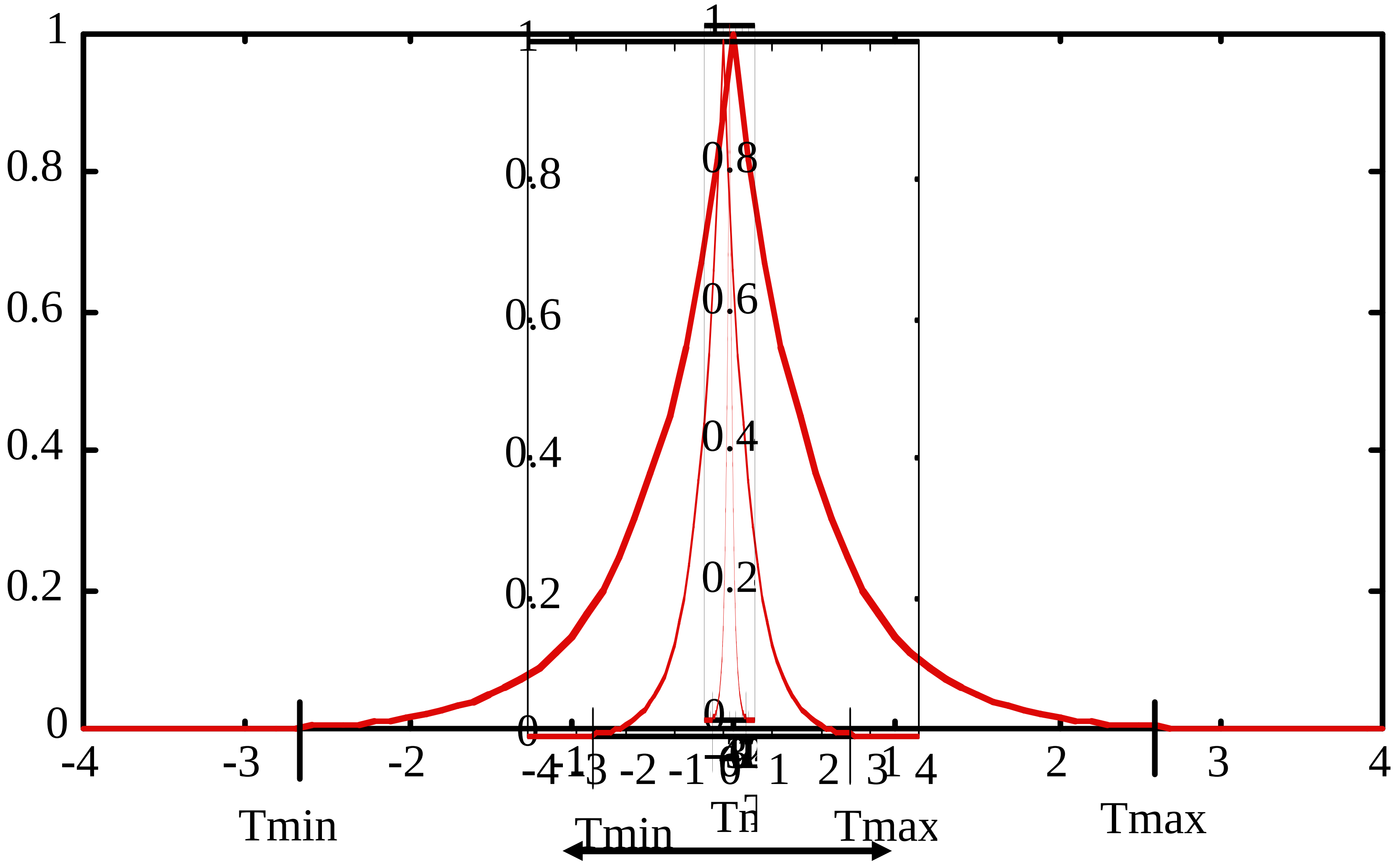
τ~ 1 sec:

For an uncorrelated

process



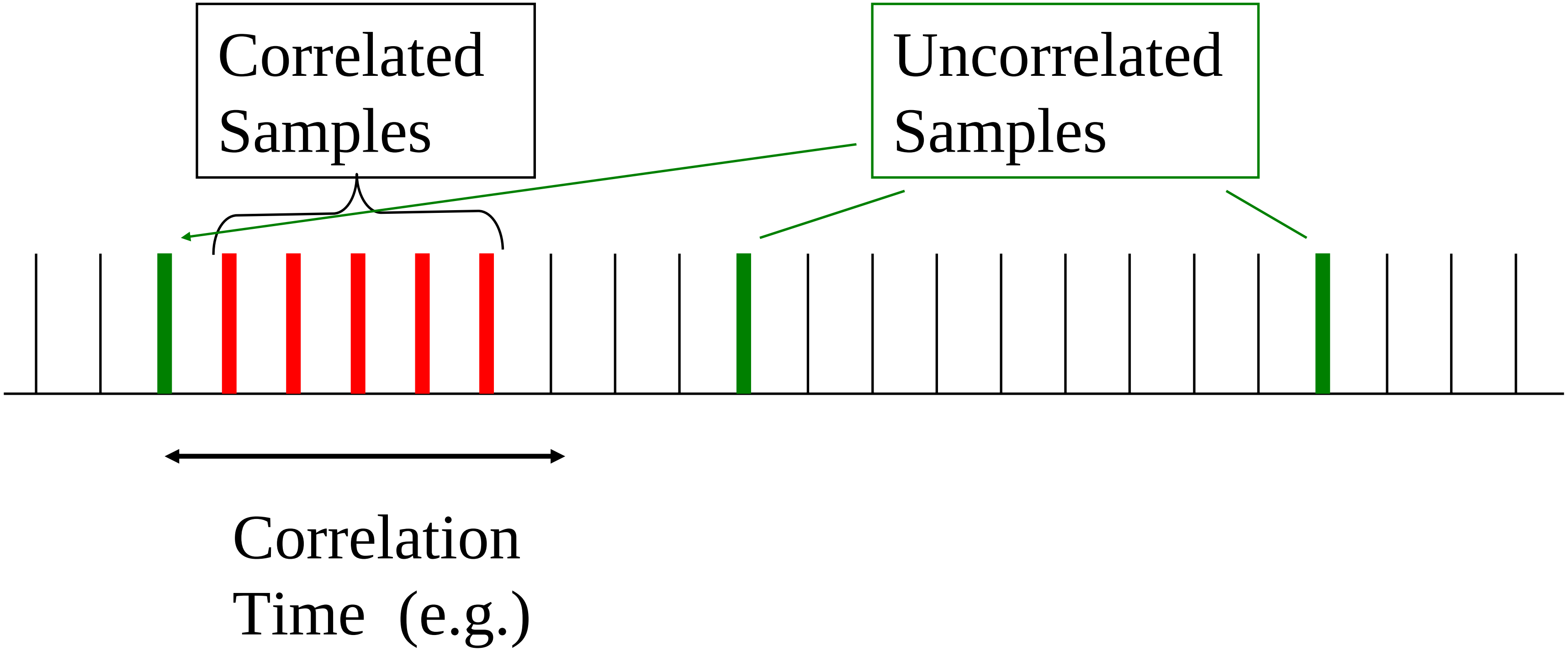
Sampling: Frequency and Distribution of Samples



SAMPLE TIME

Manufacturing

Correlation and Sampling



Correlated Samples

Uncorrelated Samples

Correlation Time (e.g.)

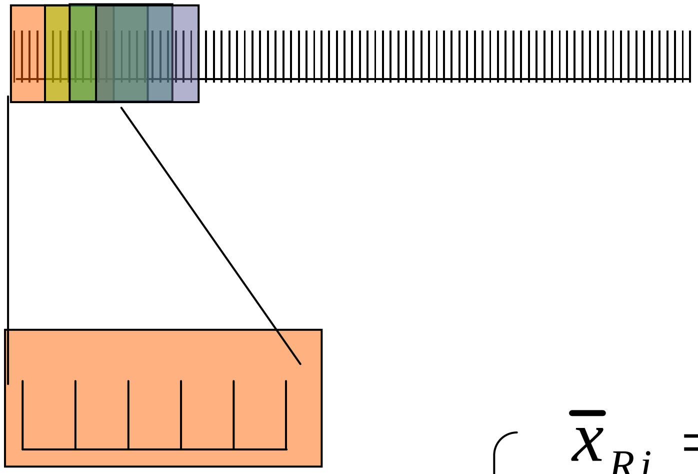
- Taking samples beyond correlation time guarantees independence

Sampling and Averaging

- Sampling Frequency Affects
  - Time Response
  - Correlation
- Averaging
  - Filters Data
  - Slows Response

Alternative Charts: Running Averages

- More averages/Data
- Can use run data alone and average for S only
- Can use to improve resolution of mean shift



n measurements at sample *j*

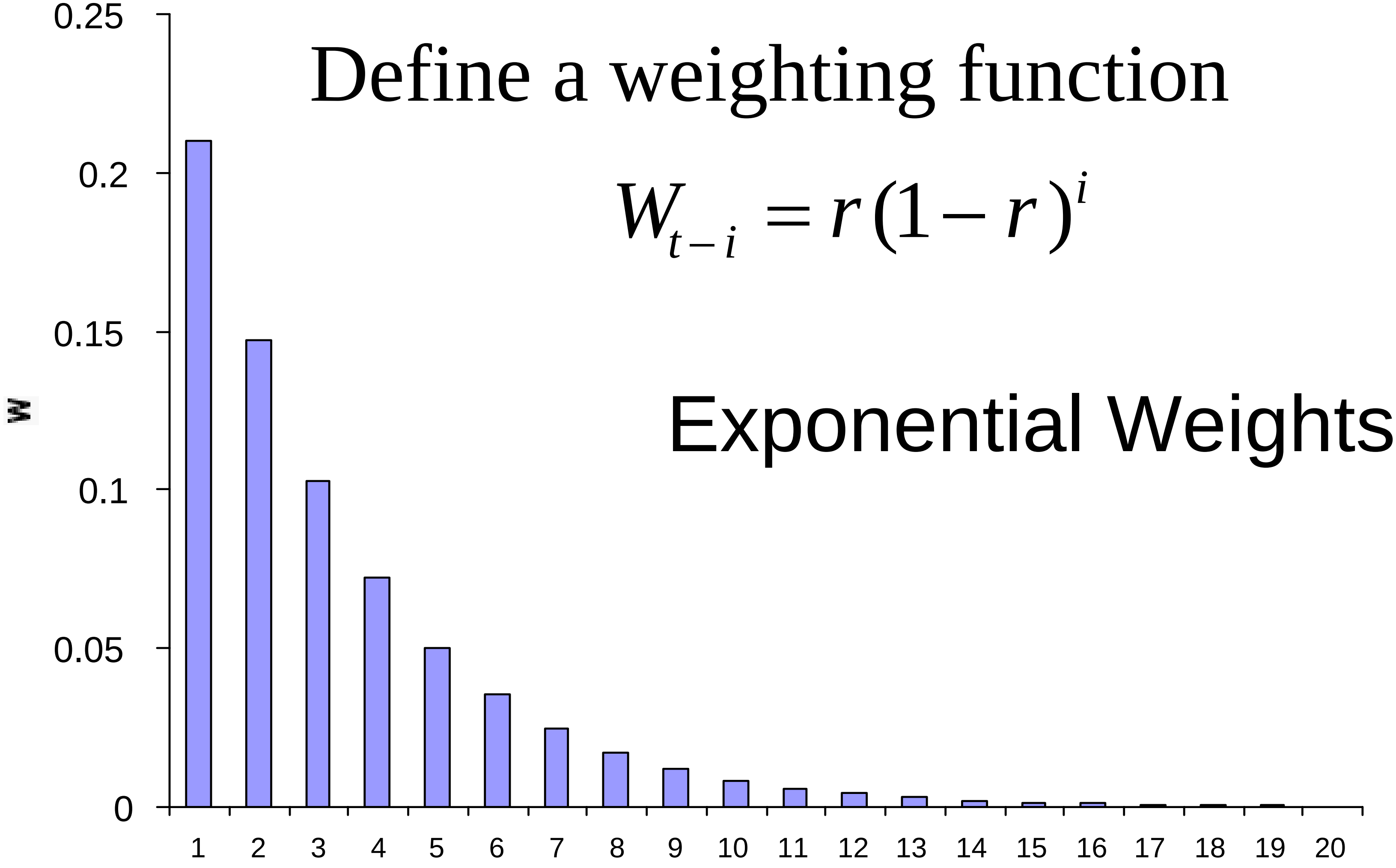
$$\left\{\begin{array}{ll} \bar{x}_{Rj} = \dfrac{1}{n}\displaystyle\sum_{i=j}^{j+n} x_i & \text{Running Average} \\ S_{Rj}^{\;2} = \dfrac{1}{n-1}\displaystyle\sum_{i=j}^{j+n}(x_i - \bar{x}_{Rj})^2 & \text{Running Variance} \end{array}\right\}$$

Specific Case: Weighted Averages

$$y_j = a_1 x_{j-1} + a_2 x_{j-2} + a_3 x_{j-3} + \ldots$$

- How should we weight measurements??
  - All equally? (as with Running Average)
  - Based on how recent?
    - e.g. Most recent are more relevant than less recent?

Consider an Exponential Weighted Average



Define a weighting function

$$W_{t-i} = r(1-r)^{i}$$

Exponential Weights

Exponentially Weighted Moving Average: (EWMA)

$$A_i = r x_i + (1-r) A_{i-1}$$

Recursive EWMA

$$\sigma_A = \sqrt{\left(\frac{\sigma_x{}^2}{n}\right)\left(\frac{r}{2-r}\right)\left[1-(1-r)^{2t}\right]}$$

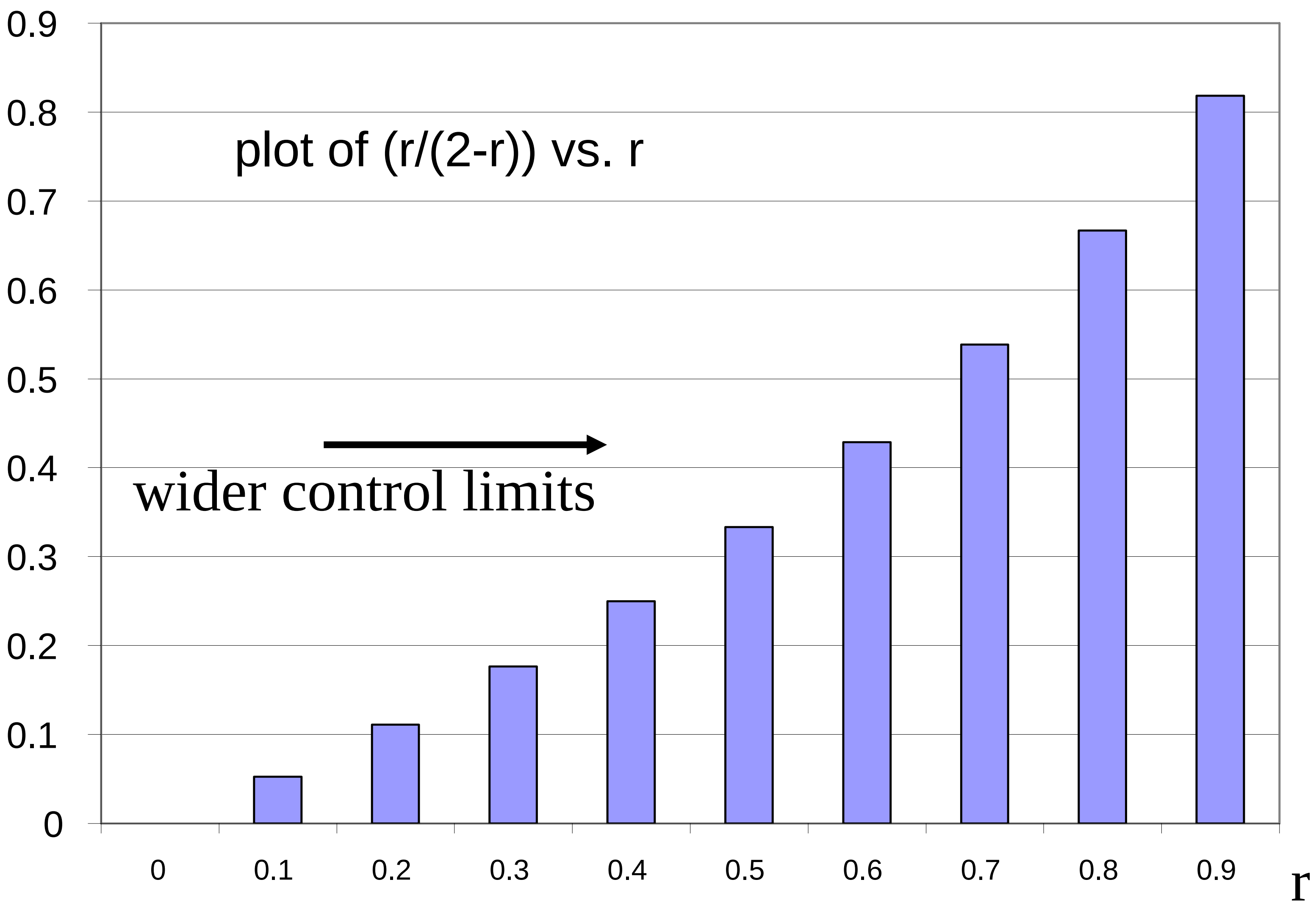
time

$$UCL, LCL = \bar{\bar{x}} \pm 3\sigma_A$$

$$\sigma_A = \sqrt{\frac{\sigma_x{}^2}{n}\left(\frac{r}{2-r}\right)}$$

for large t

Effect of r on σ multiplier



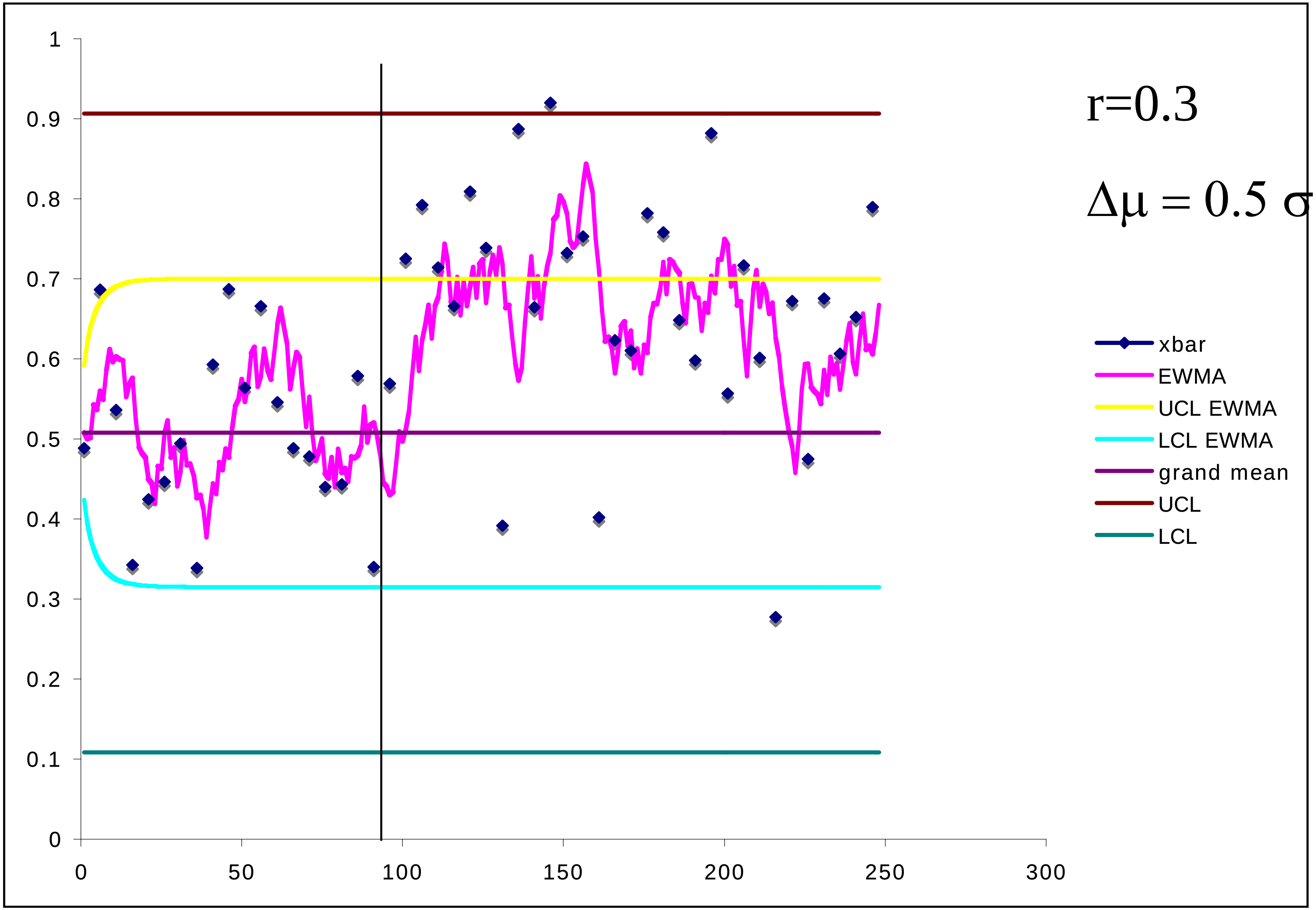
SO WHAT?

- The variance will be less than with xbar,

$$\sigma_A = \frac{\sigma_x}{\sqrt{n}} \sqrt{\left(\frac{r}{2-r}\right)} \qquad = \sigma_{\bar{x}} \sqrt{\left(\frac{r}{2-r}\right)}$$

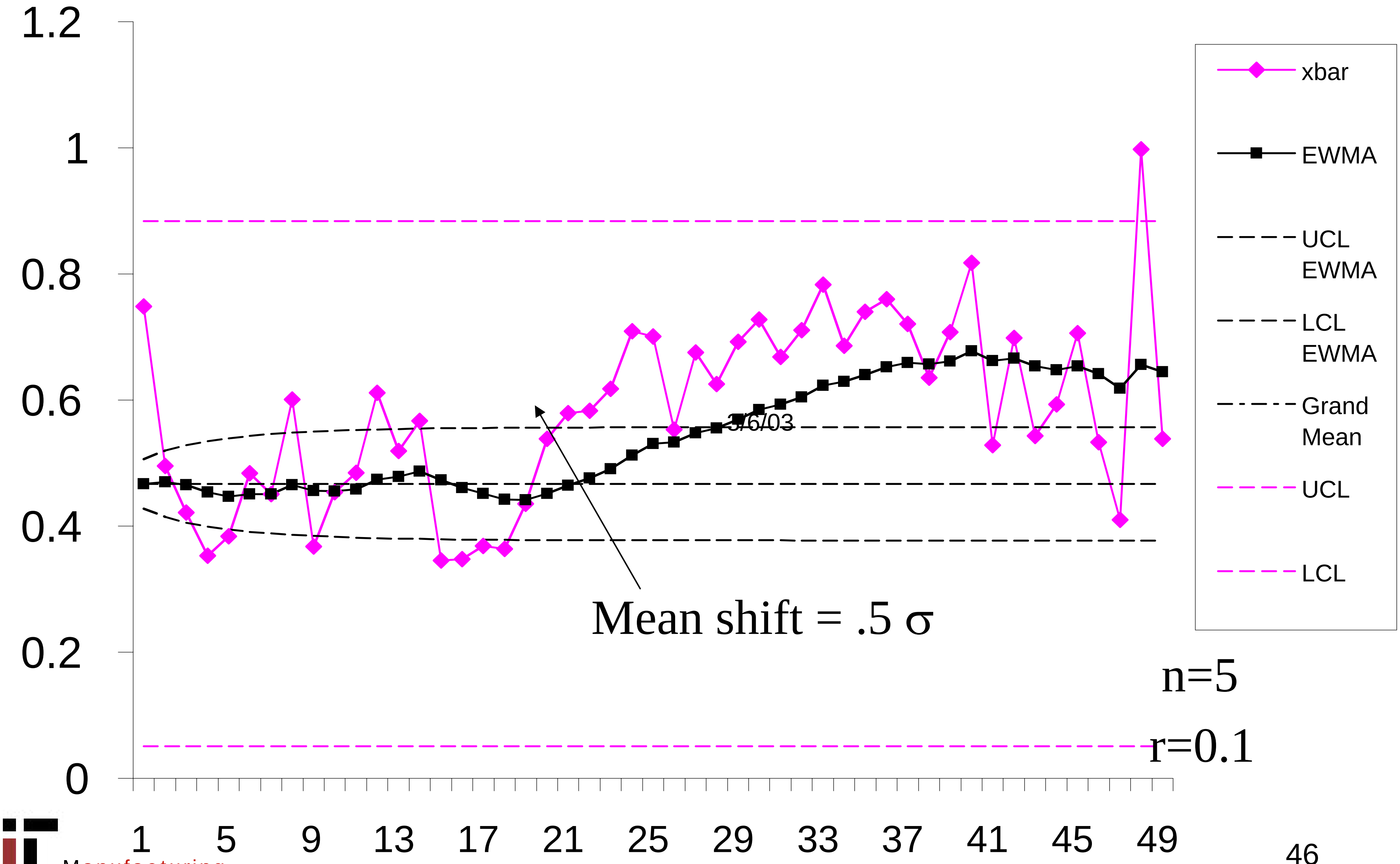
- n=1 case is valid
- If r=1 we have “unfiltered” data
  - Run data stays run data
  - Sequential averages remain
- If r<<1 we get long weighting and long delays
  - “Stronger” filter; longer response time

EWMA vs. Xbar



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
50
100
150
200
250
300
r=0.3
Δμ = 0.5 σ
xbar
EWMA
UCL EWMA
LCL EWMA
grand mean
UCL
LCL

Mean Shift Sensitivity
EWMA and Xbar comparison

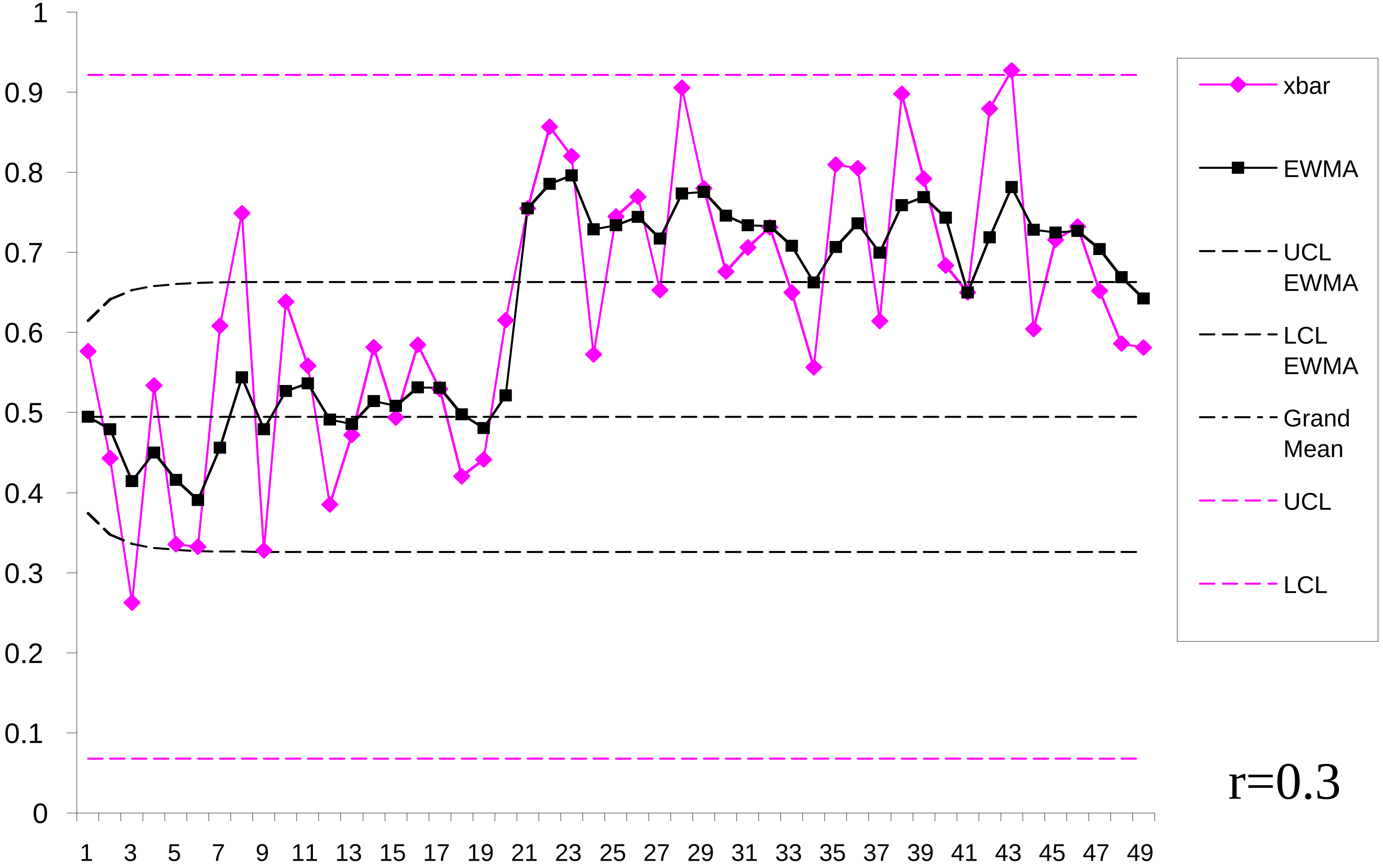


Mean shift = .5 σ

n=5

r=0.1

Effect of r



r=0.3

Small Mean Shifts

- What if $\Delta\mu_x$ is small wrt σ_x ?
- But it is “persistent”
- How could we detect?
  - ARL for xbar would be too large

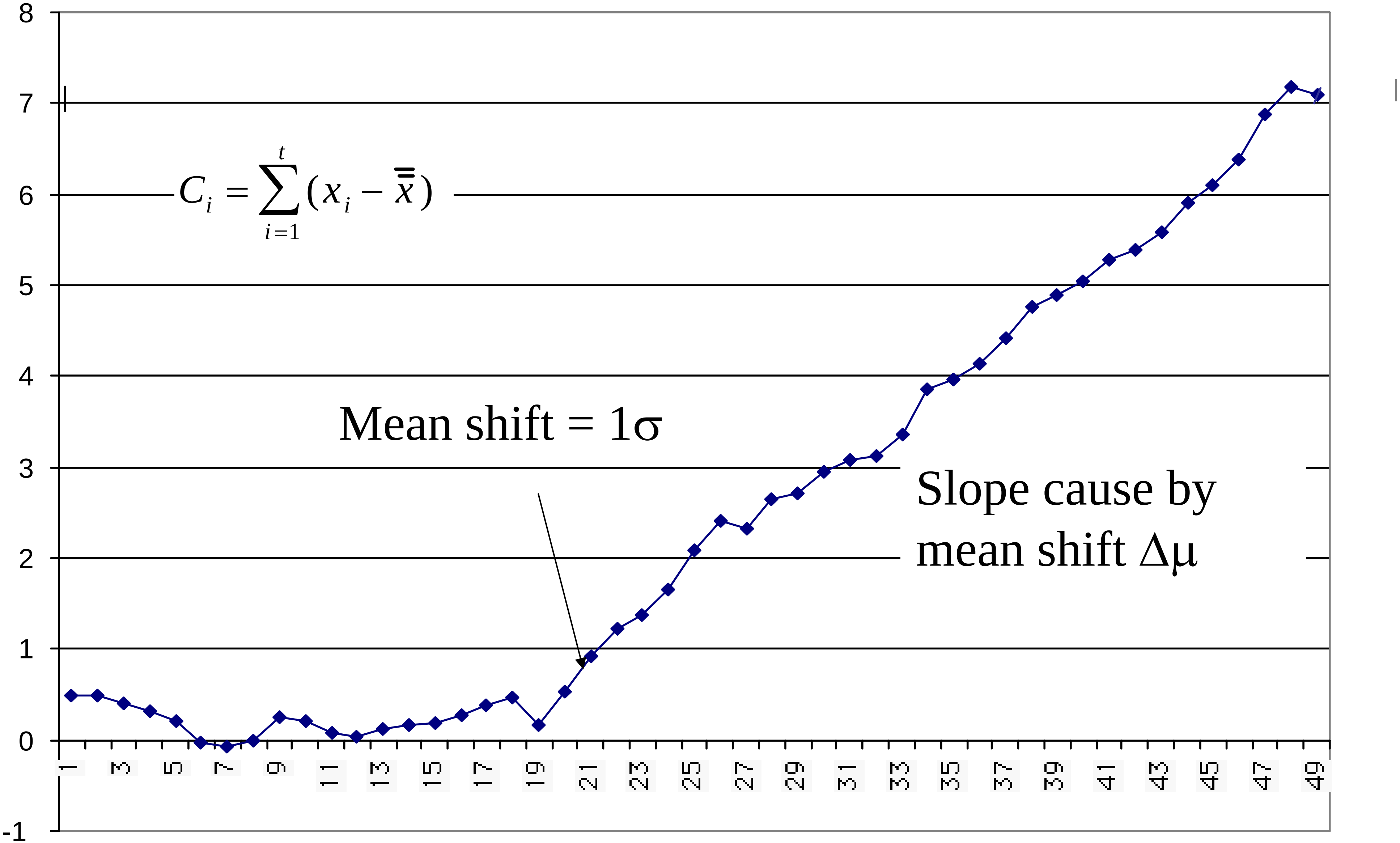
Another Approach: Cumulative Sums

- Add up deviations from mean
  - A Discrete Time Integrator

$$C_j = \sum_{i=1}^{j} (x_i - \bar{\bar{x}})$$

- Since E{x-μ}=0 this sum should stay near zero
- Any bias in *x* will show as a trend

Mean Shift Sensitivity: CUSUM



$$C_i = \sum_{i=1}^{t} (x_i - \bar{\bar{x}})$$

Mean shift = 1σ

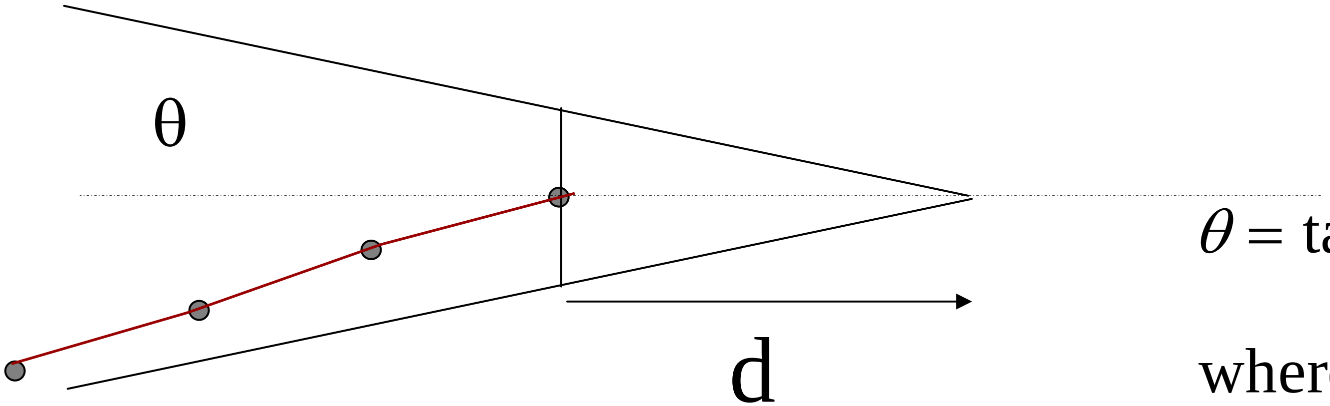
Slope cause by mean shift Δμ

Control Limits for CUSUM

- Significance of Slope Changes?
  - Detecting Mean Shifts
- Use of v-mask
  - Slope Test with Deadband

$$d = \frac{2}{\delta} \ln\left(\frac{1-\beta}{\alpha}\right)$$

Upper decision line



Lower decision line

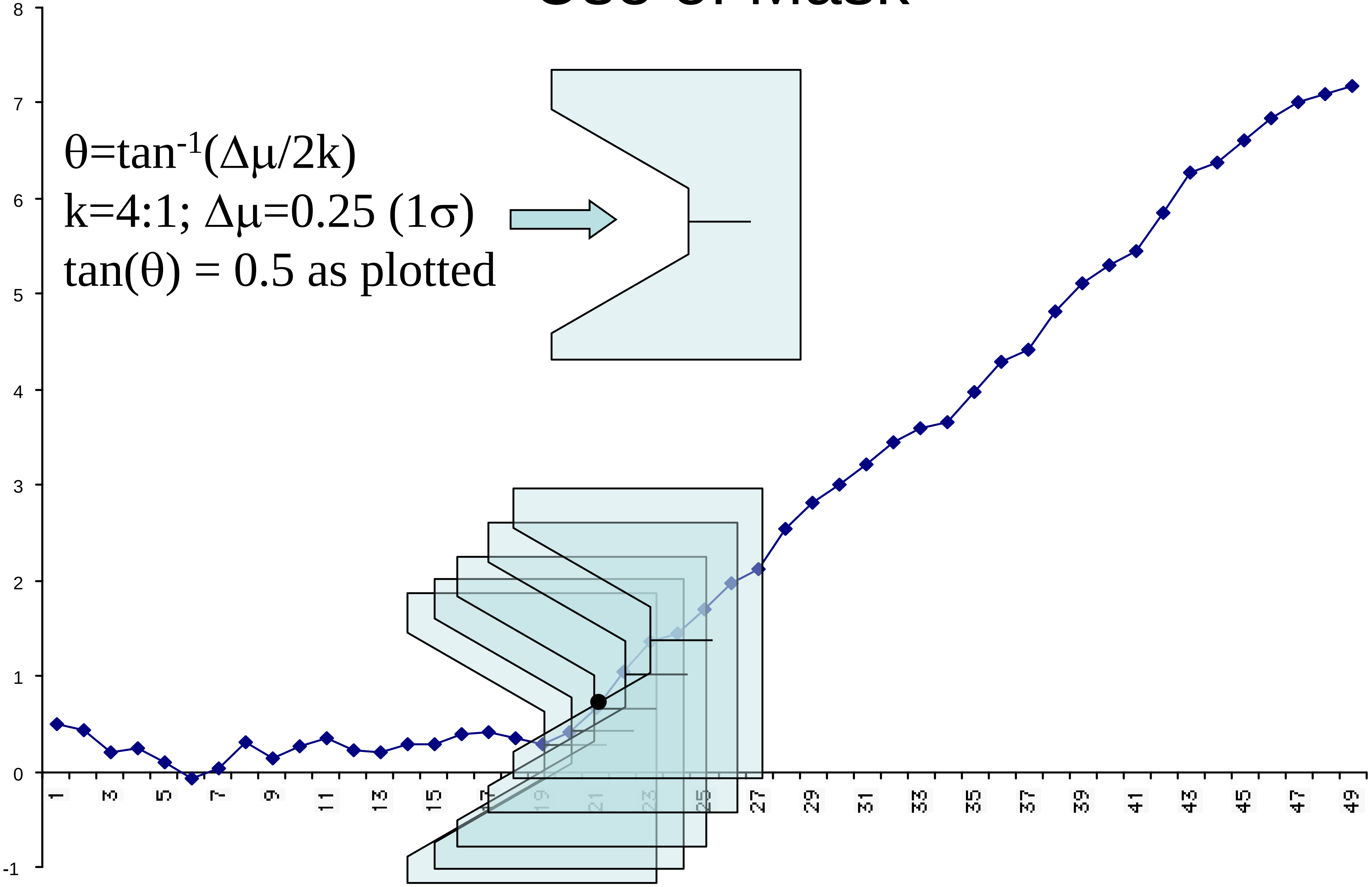
$$\delta = \frac{\Delta \bar{x}}{\sigma_{\bar{x}}}$$

$$\theta = \tan^{-1}\left(\frac{\Delta \bar{x}}{2k}\right)$$

where

k = horizontal scale factor for plot

Use of Mask



$\theta = \tan^{-1}(\Delta\mu / 2k)$

$k = 4:1$; $\Delta\mu = 0.25$ (1σ)

$\tan(\theta) = 0.5$ as plotted

An Alternative

- Define the Normalized Statistic

$$Z_i = \frac{X_i - \mu_x}{\sigma_x}$$

Which has an expected mean of 0 and variance of 1

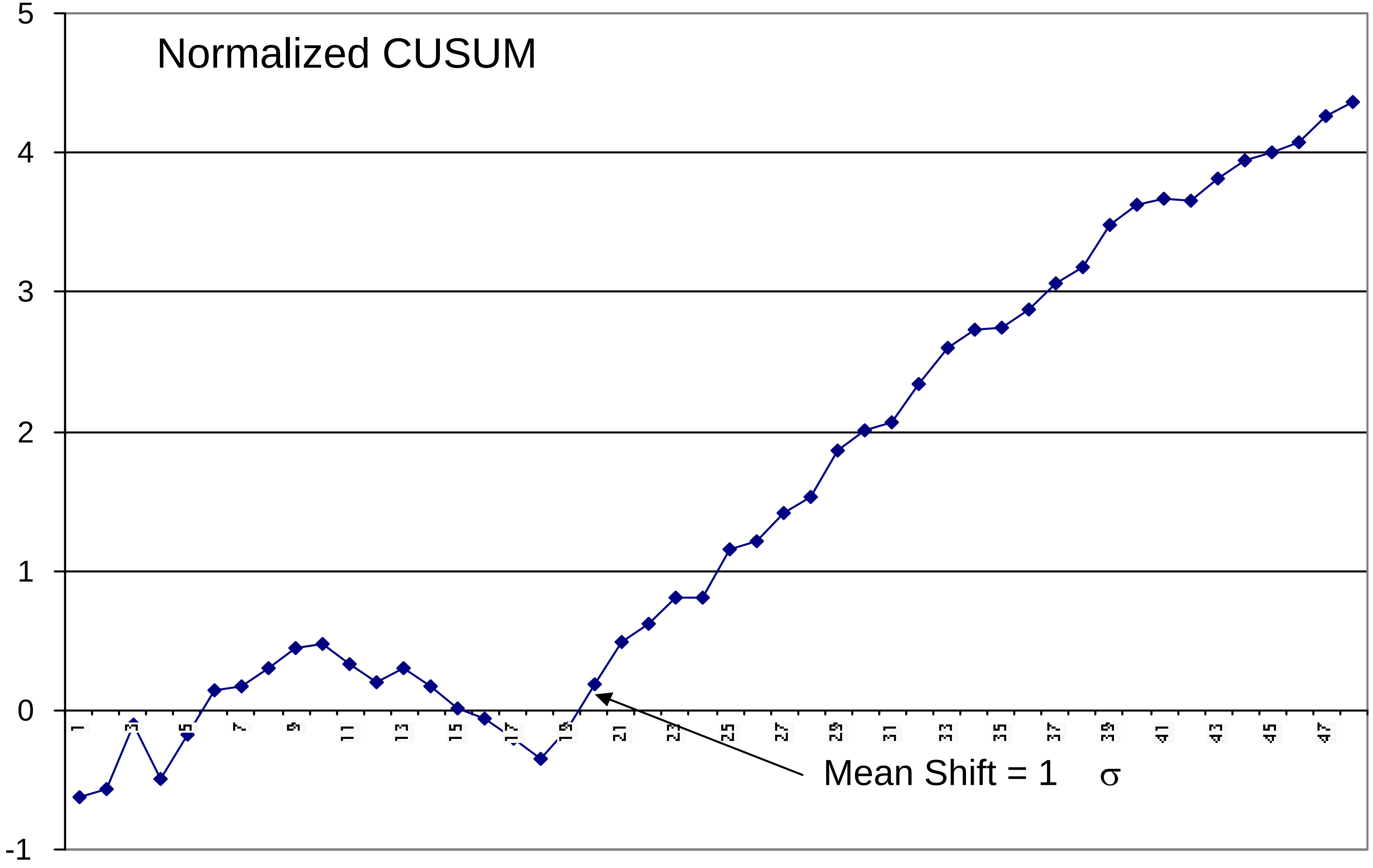
- And the CUSUM statistic

$$S_i = \frac{\sum_{i=1}^{t} Z_i}{\sqrt{t}}$$

Which has an expected mean of 0 and variance of 1

Chart with Centerline =0 and Limits = ±3

Example for Mean Shift = 1σ



Tabular CUSUM

- Create Threshold Variables:

$$C_i^+ = \max[0, x_i - (\mu_0 + K) + C_{i-1}^+]$$

$$C_i^- = \max[0, (\mu_0 - K) - x_i + C_{i-1}^-]$$

Accumulates deviations from the mean

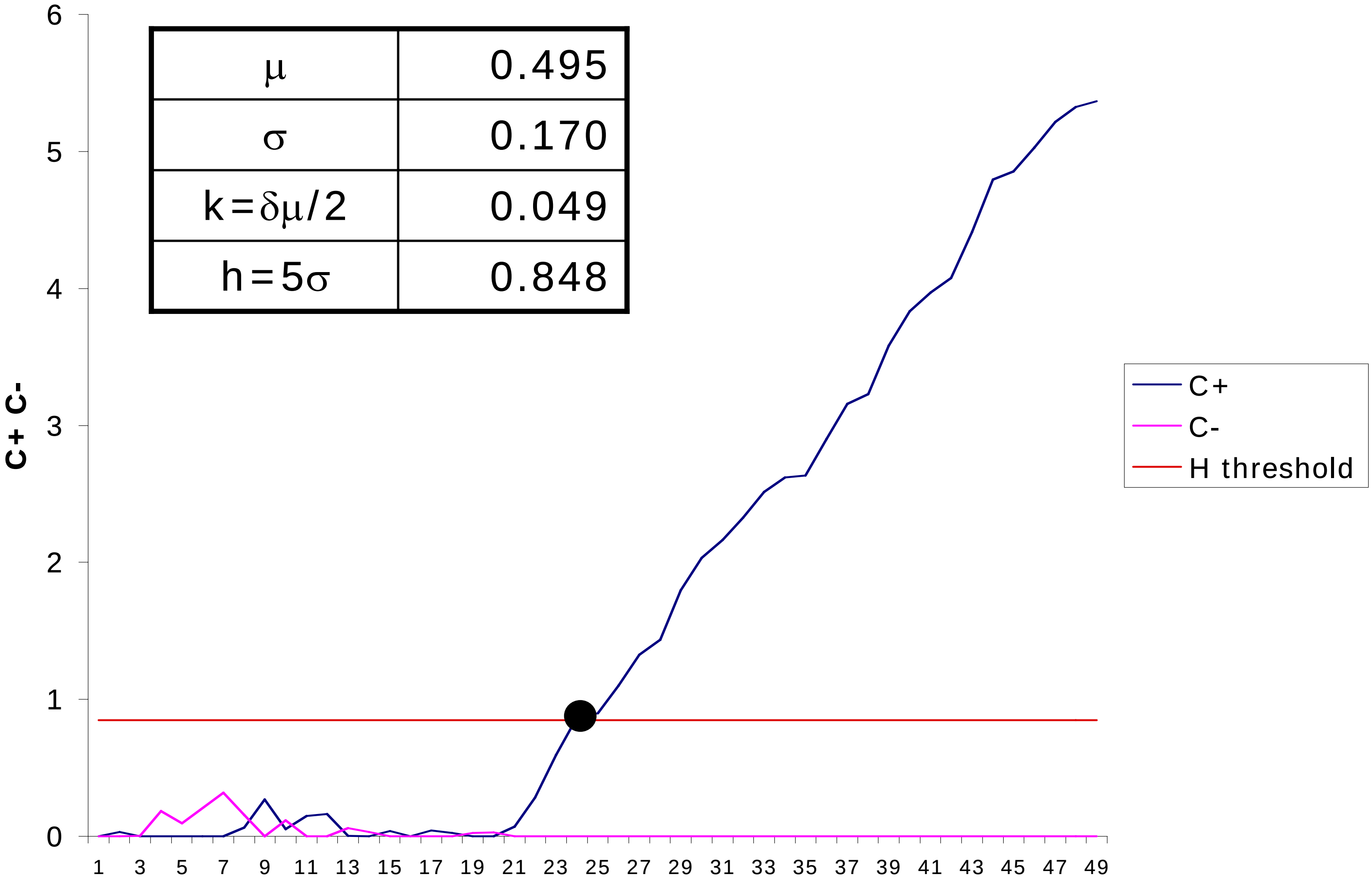
K= threshold or slack value for accumulation

$$K = \left|\frac{\Delta\mu}{2}\right| \qquad \Delta\mu = \text{mean shift to detect}$$

typical

$H:$ alarm level (typically 5σ)

Threshold Plot



μ	0.495
σ	0.170
k=δμ/2	0.049
h=5σ	0.848

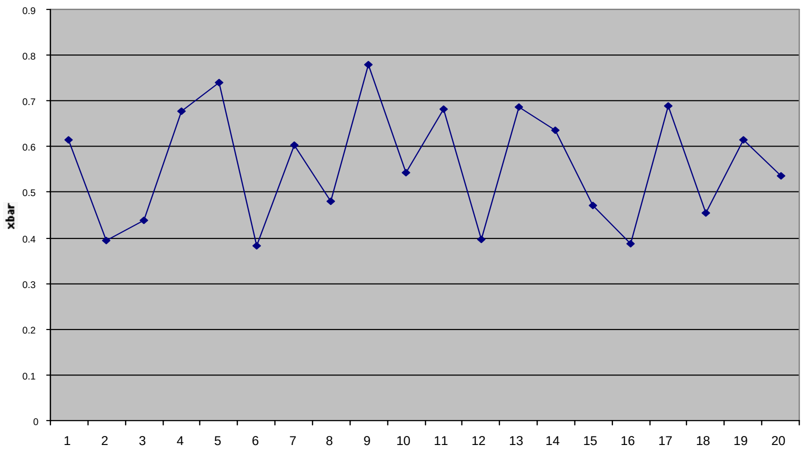
Alternative Charts Summary

- Noisy Data Need Some Filtering
- Sampling Strategy Can Guarantee Independence
- Linear Discrete Filters have Been Proposed
  - EWMA
  - Running Integrator
- Choice Depends on Nature of Process

Summary of SPC

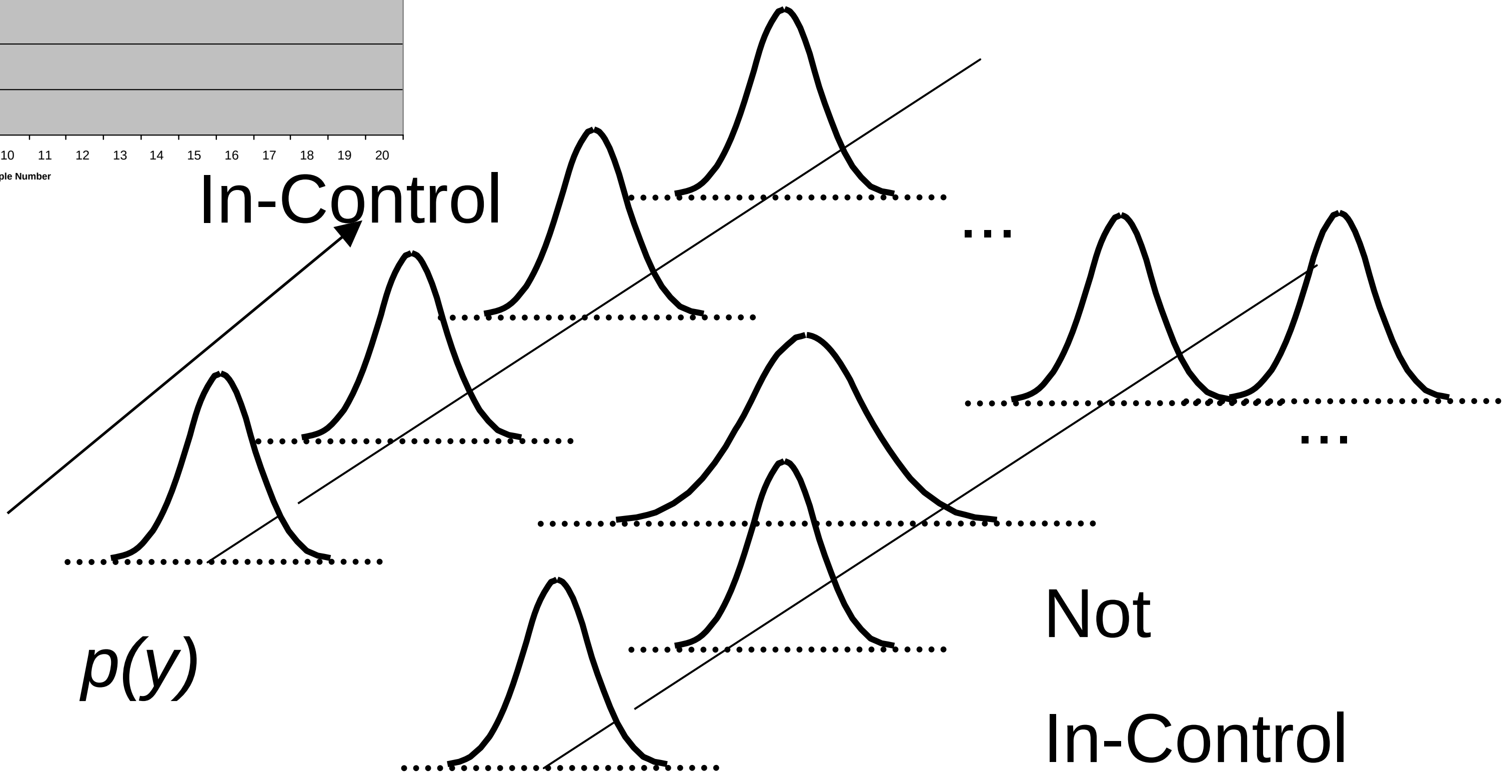
- Consider Process a Random Process
  - Can never predict precise value
- Model with $P(x)$ or $p(x)$
  - Assume $p(x,t) = p(x)$
- Shewhart Hypothesis
  - In-control = purely random output
    - Normal, independent stationary
    - “The best you can do!”
  - Not in-control
    - Non-random behavior
    - Source can be found and eliminated

The SPC Hypothesis



Process

Y



In-Control

…

…

$p(y)$

Not

In-Control