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 #9

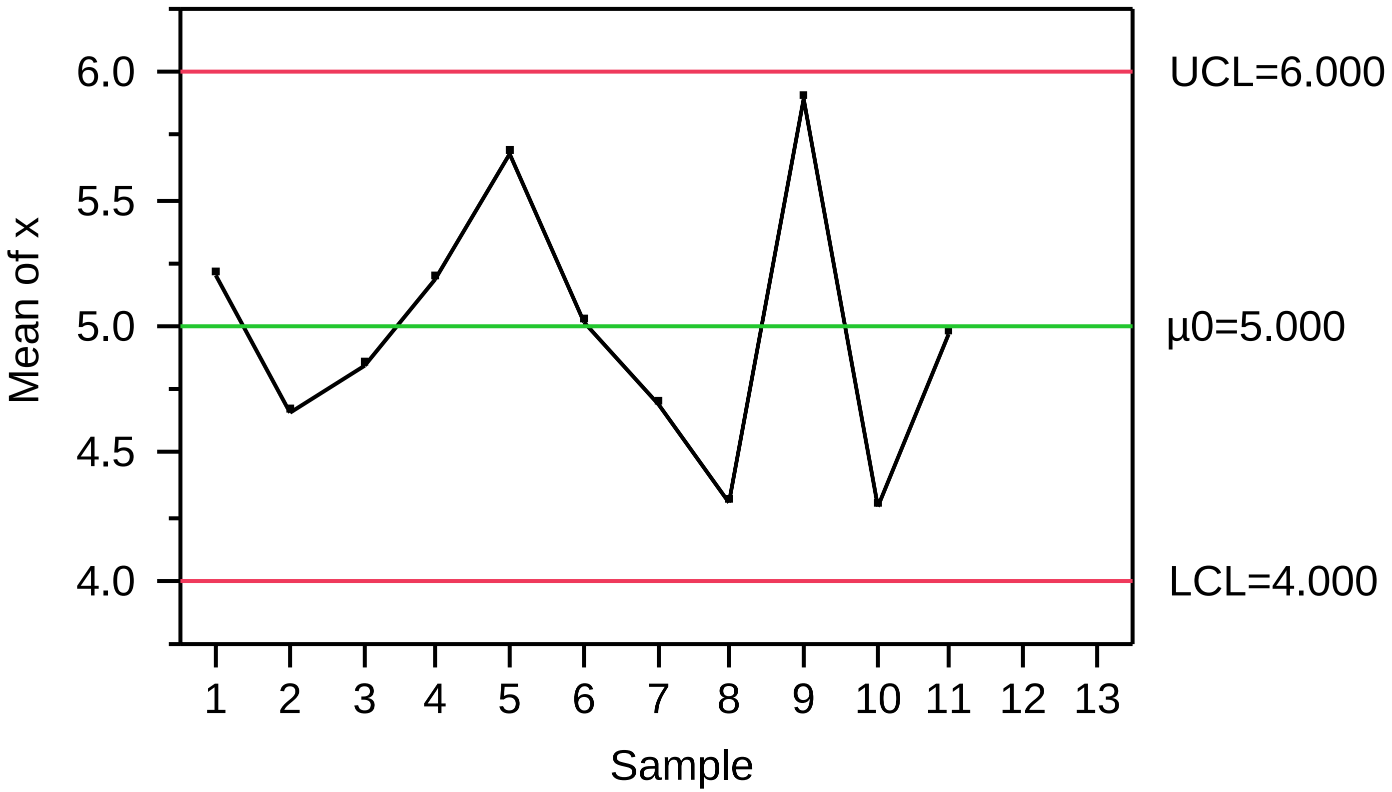
Advanced and Multivariate SPC

March 6, 2008

Agenda

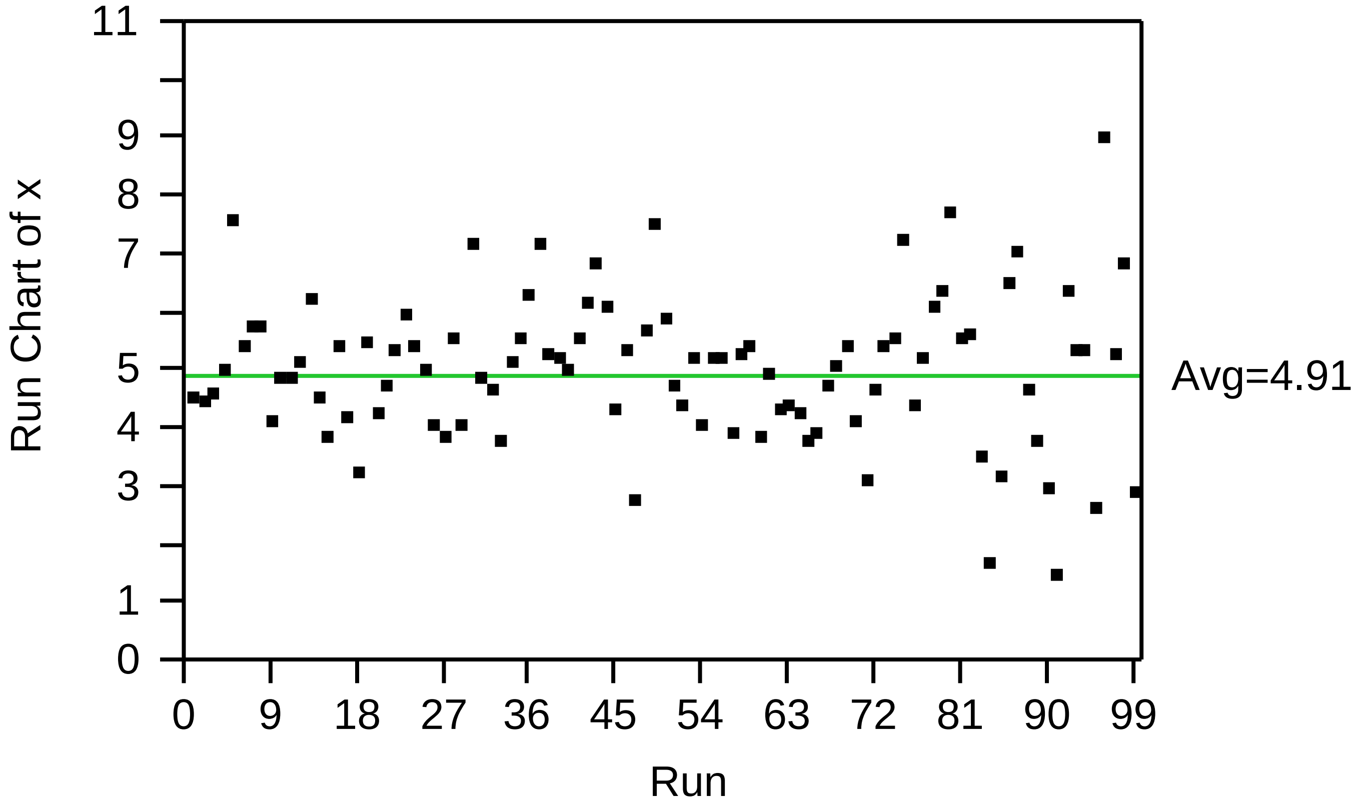
- Conventional Control Charts
  - Xbar and S
- Alternative Control Charts
  - Moving average
  - EWMA
  - CUSUM
- Multivariate SPC

Xbar Chart
Process Model: x ~ N(5,1), n = 9



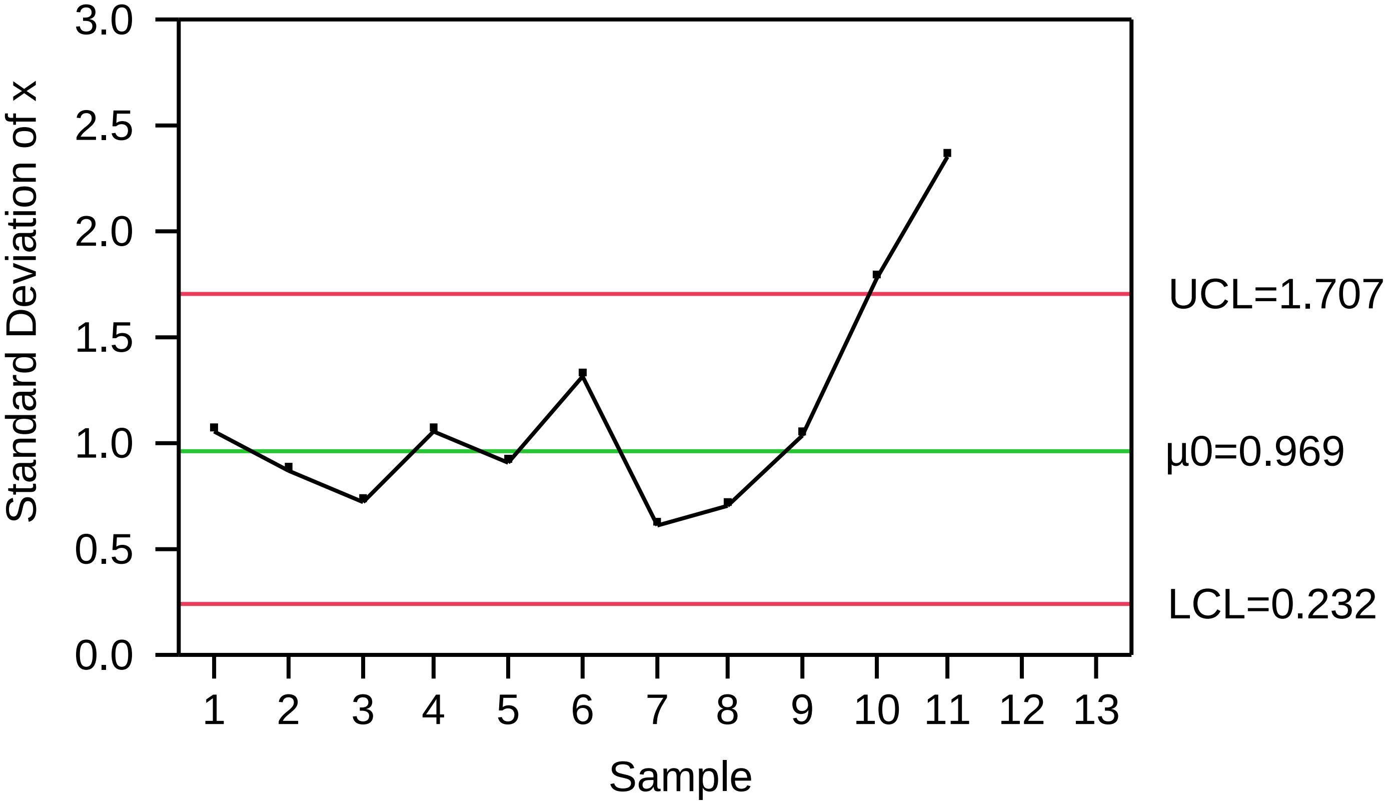
- Is process in control?

Run Data (n=9 sample size)



- Is process in control?

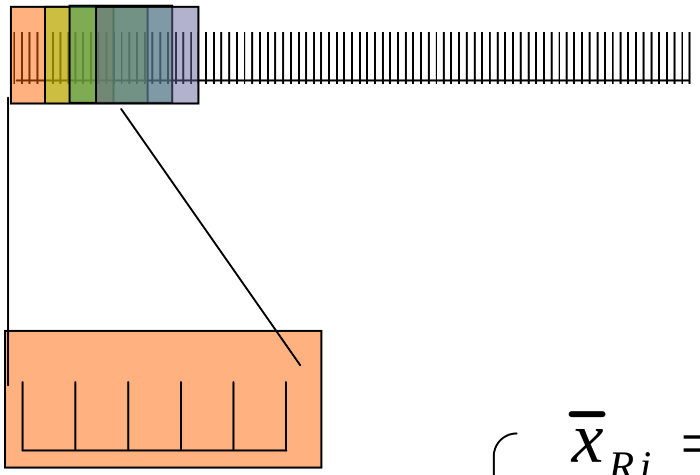
S Chart



- Is process in control?

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*

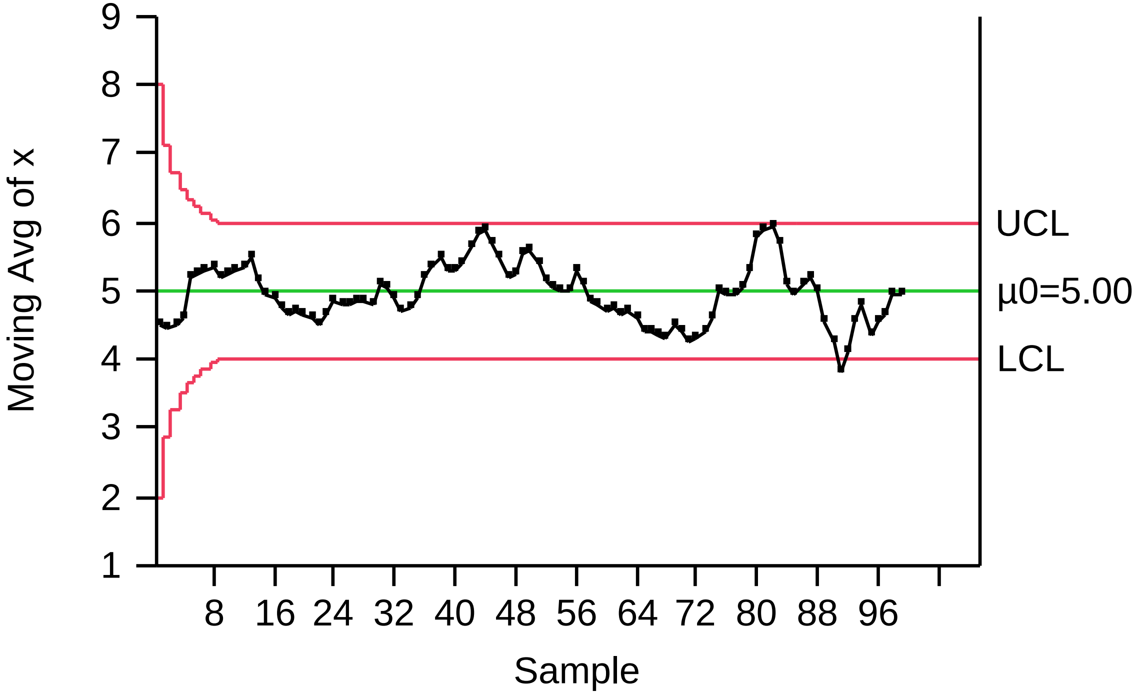
$$\left\{ \begin{aligned} \bar{x}_{Rj} &= \frac{1}{n}\sum_{i=j}^{j+n} x_i && \text{Running Average} \\ S_{Rj}^{\;2} &= \frac{1}{n-1}\sum_{i=j}^{j+n} (x_i - \bar{x}_{Rj})^2 && \text{Running Variance} \end{aligned} \right\}$$

Simplest Case: Moving Average

- Pick window size (e.g., w = 9)

$$M_i = \frac{x_i + x_{i-1} + \cdots + x_{i-w+1}}{w}$$

$$V(M_i) = \frac{1}{w^2}\sum_{j=i-w+1}^{i} V(x_j) = \frac{1}{w^2}\sum_{j=i-w+1}^{i} \sigma^2 = \frac{\sigma^2}{w}$$



General Case: Weighted Averages

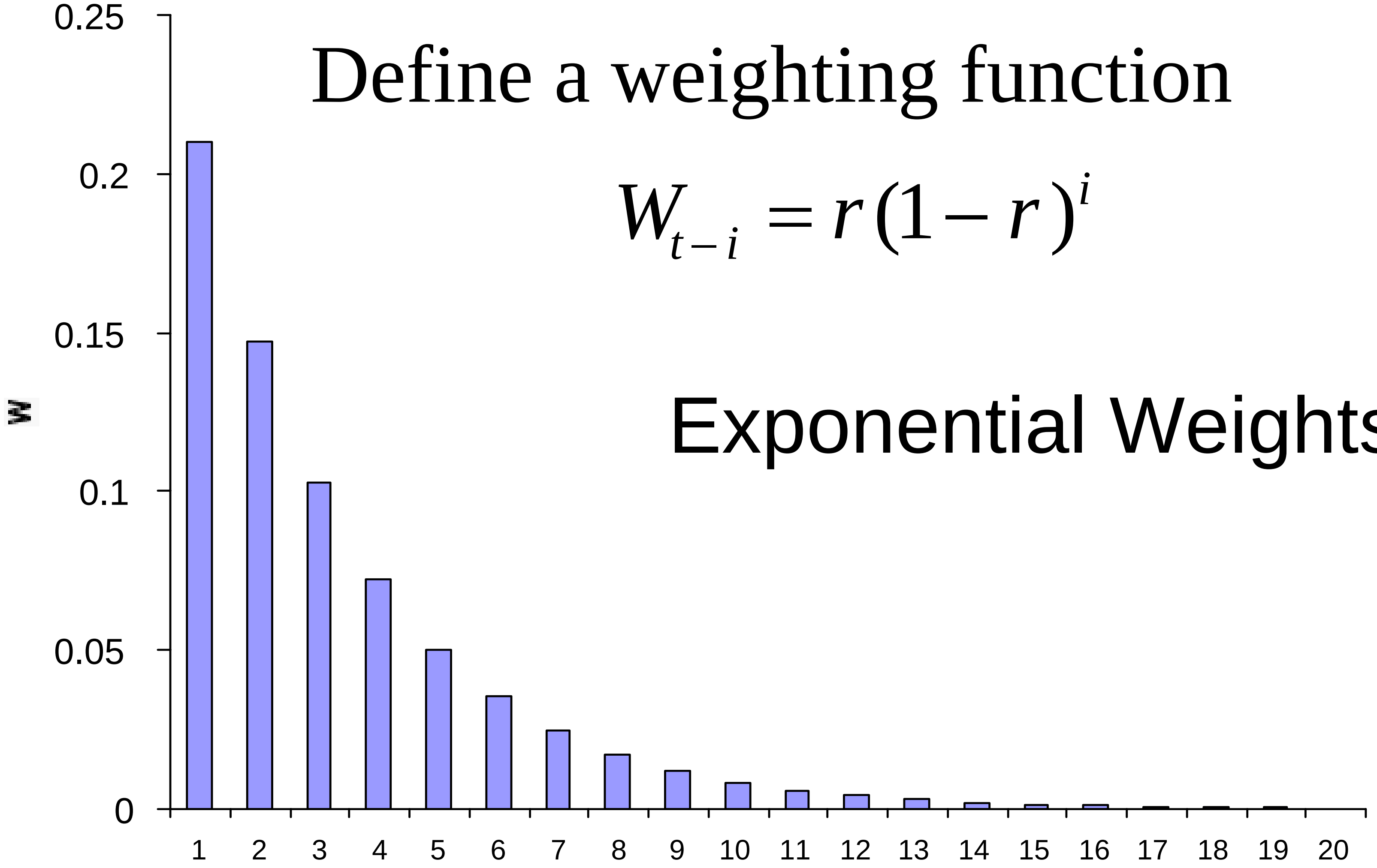
$$y_j = a_1 x_{j-1} + a_2 x_{j-2} + a_3 x_{j-3} + \ldots$$

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

Consider an Exponential Weighted Average

Define a weighting function

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



Exponentially Weighted Moving Average: (EWMA)

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

Recursive EWMA

time

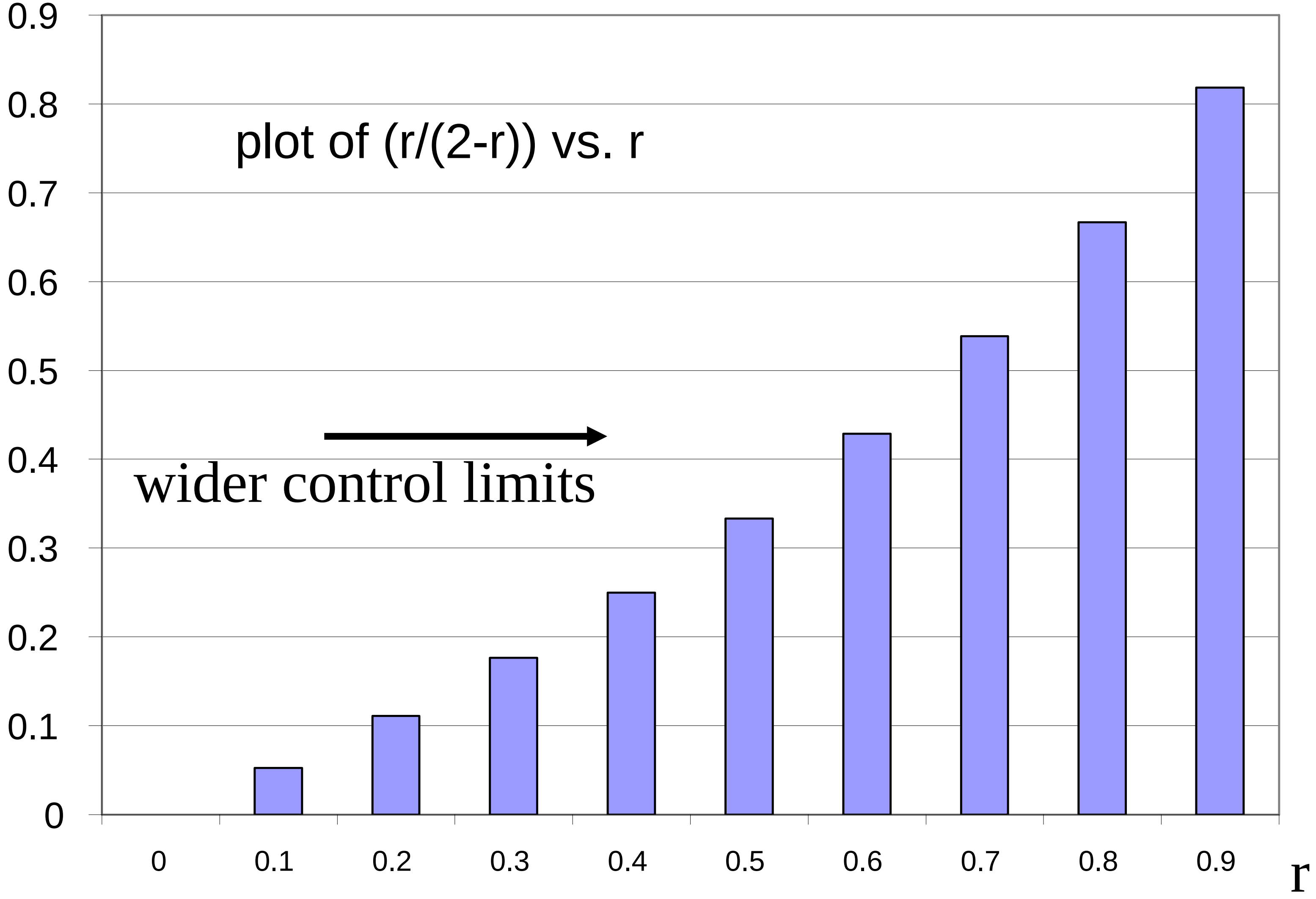
$$\sigma_A = \sqrt{\left(\frac{\sigma_x^{\,2}}{n}\right)\left(\frac{r}{2-r}\right)\left[1-(1-r)^{2t}\right]}$$

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

for large t

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

Effect of r on σ multiplier



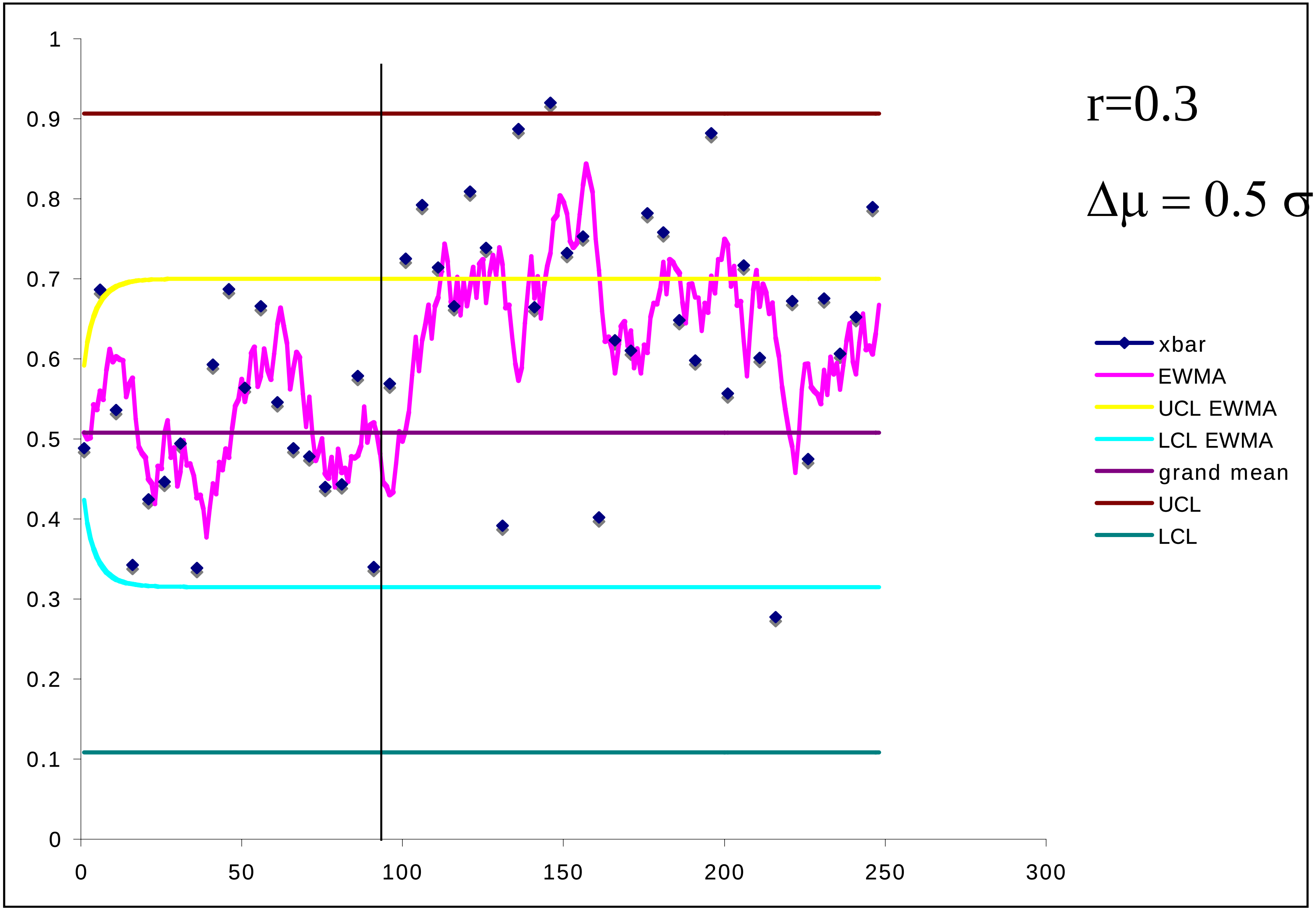
SO WHAT?

- The variance will be less than with xbar,

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

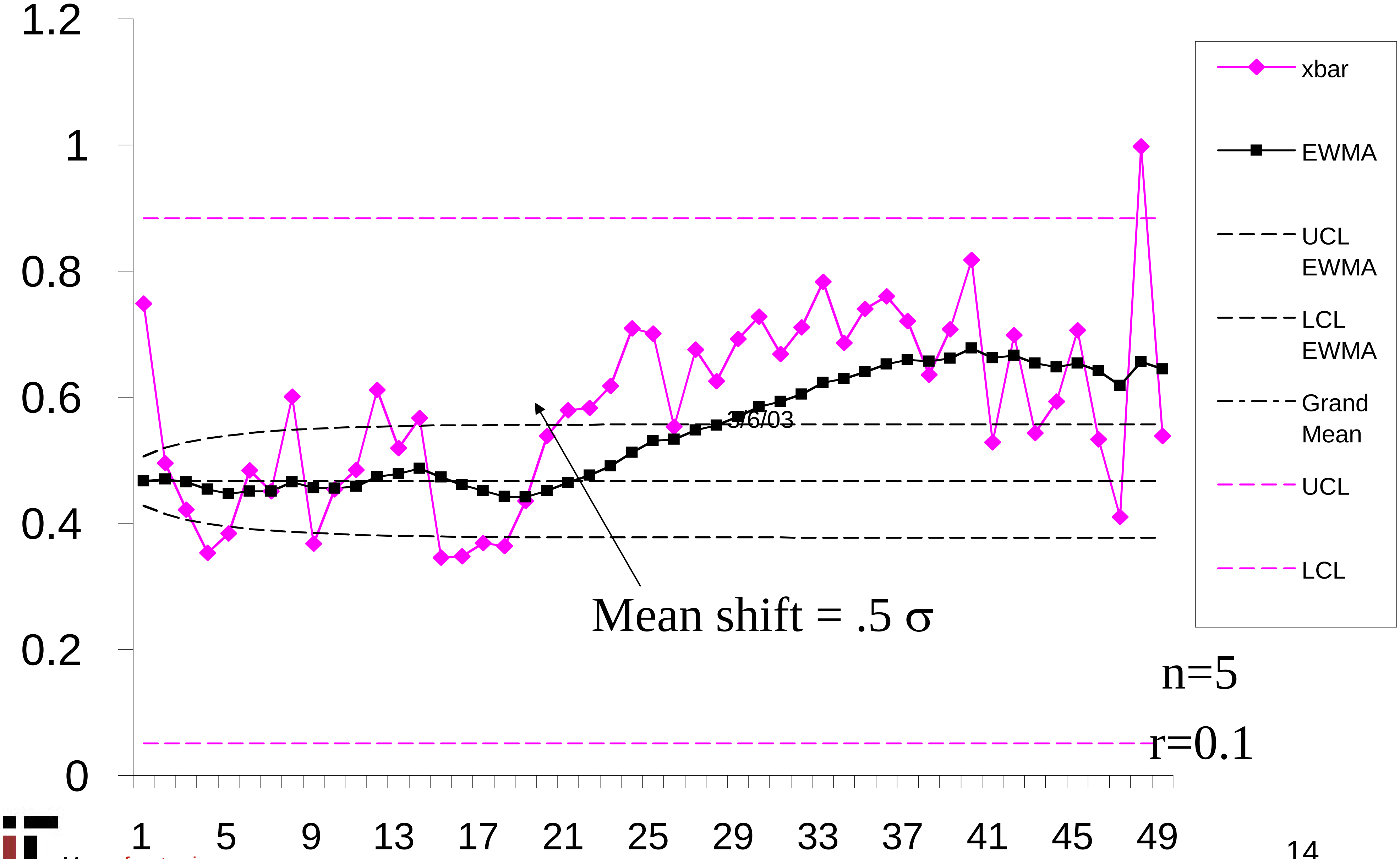
- n=1 case is valid
- If r=1 we have "unfiltered" data
  - Run data stays run data
  - Sequential averages remain
- If r<<1 we get long weighting and long delays
  - "Stronger" filter; longer response time

EWMA vs. Xbar

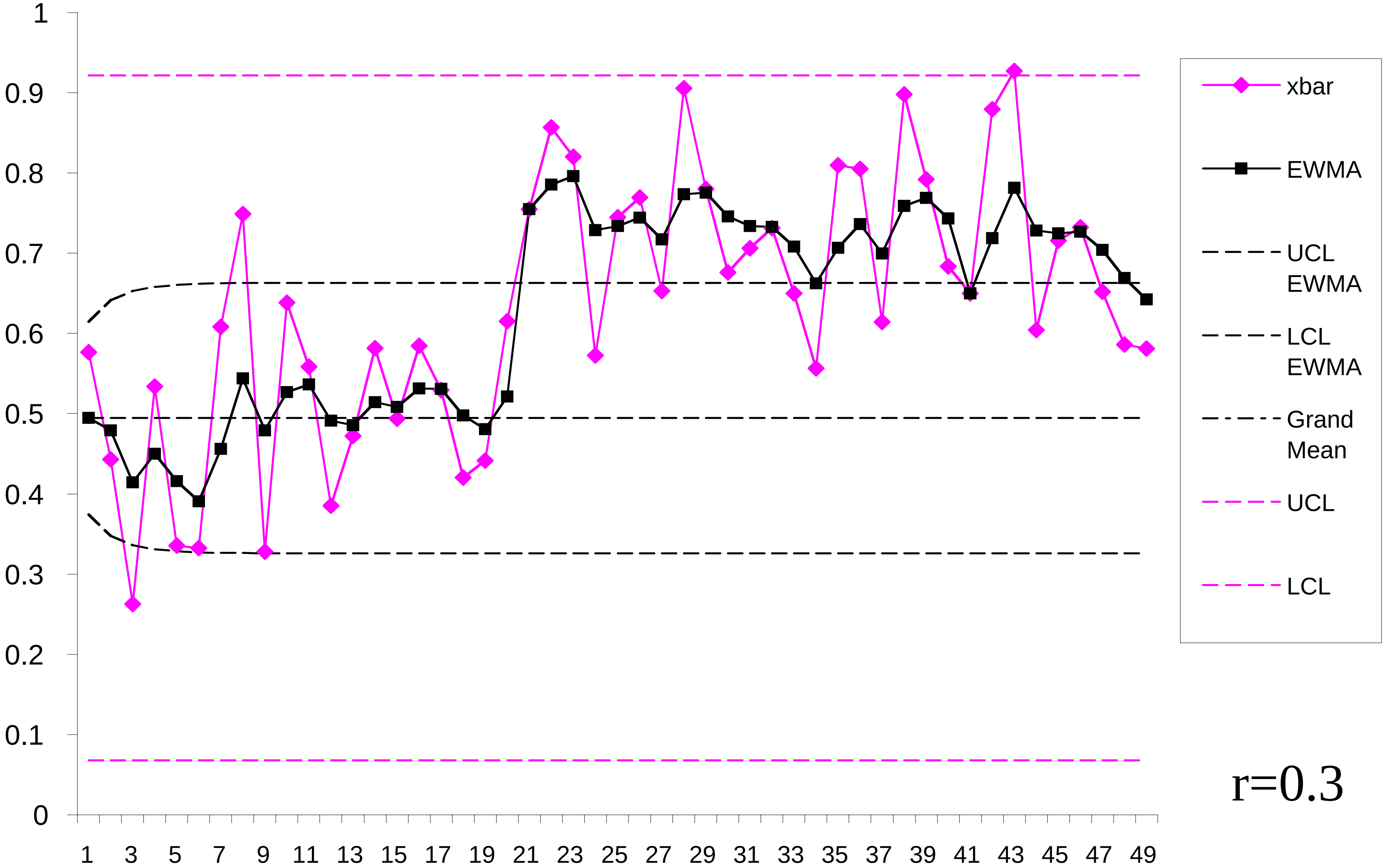


Manufacturing

Mean Shift Sensitivity
EWMA and Xbar comparison



Effect of r



r=0.3

Small Mean Shifts

- What if $\Delta\mu_x$ is small with respect to σ_x ?
- But it is "persistent"
- How could we detect?
  - ARL for xbar would be too large

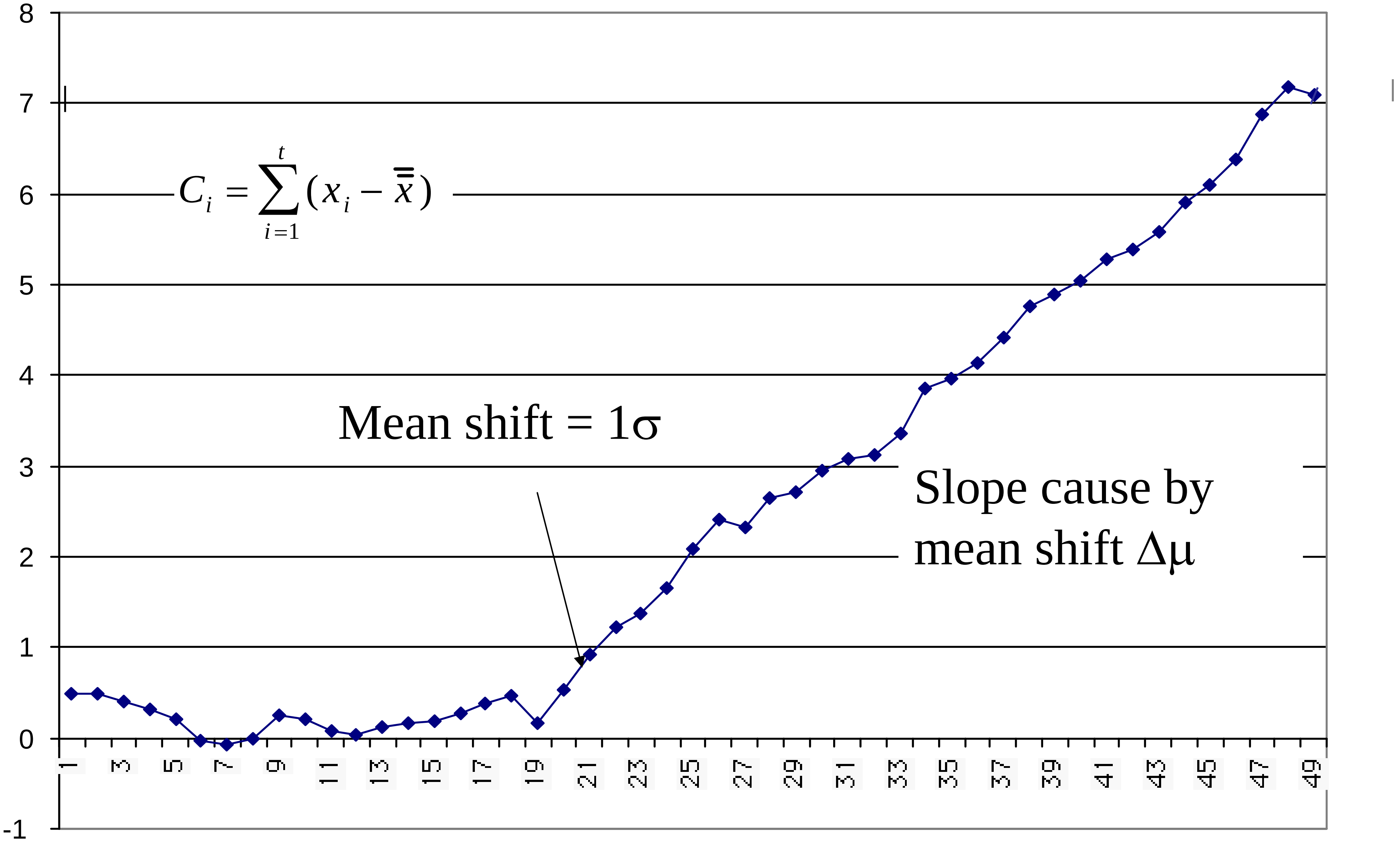
Another Approach: Cumulative Sums

- 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 when in control
- Any bias (mean shift) in *x* will show as a trend

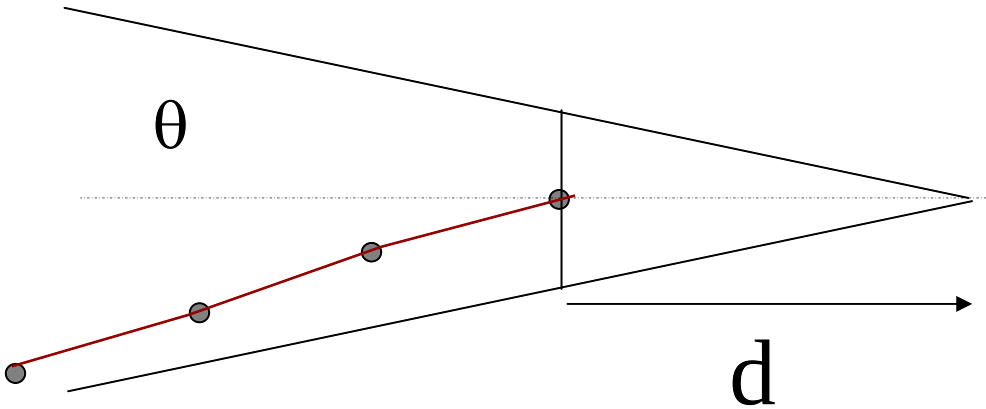
Mean Shift Sensitivity: CUSUM



Control Limits for CUSUM

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

Upper decision line



Lower decision line

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

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

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

where

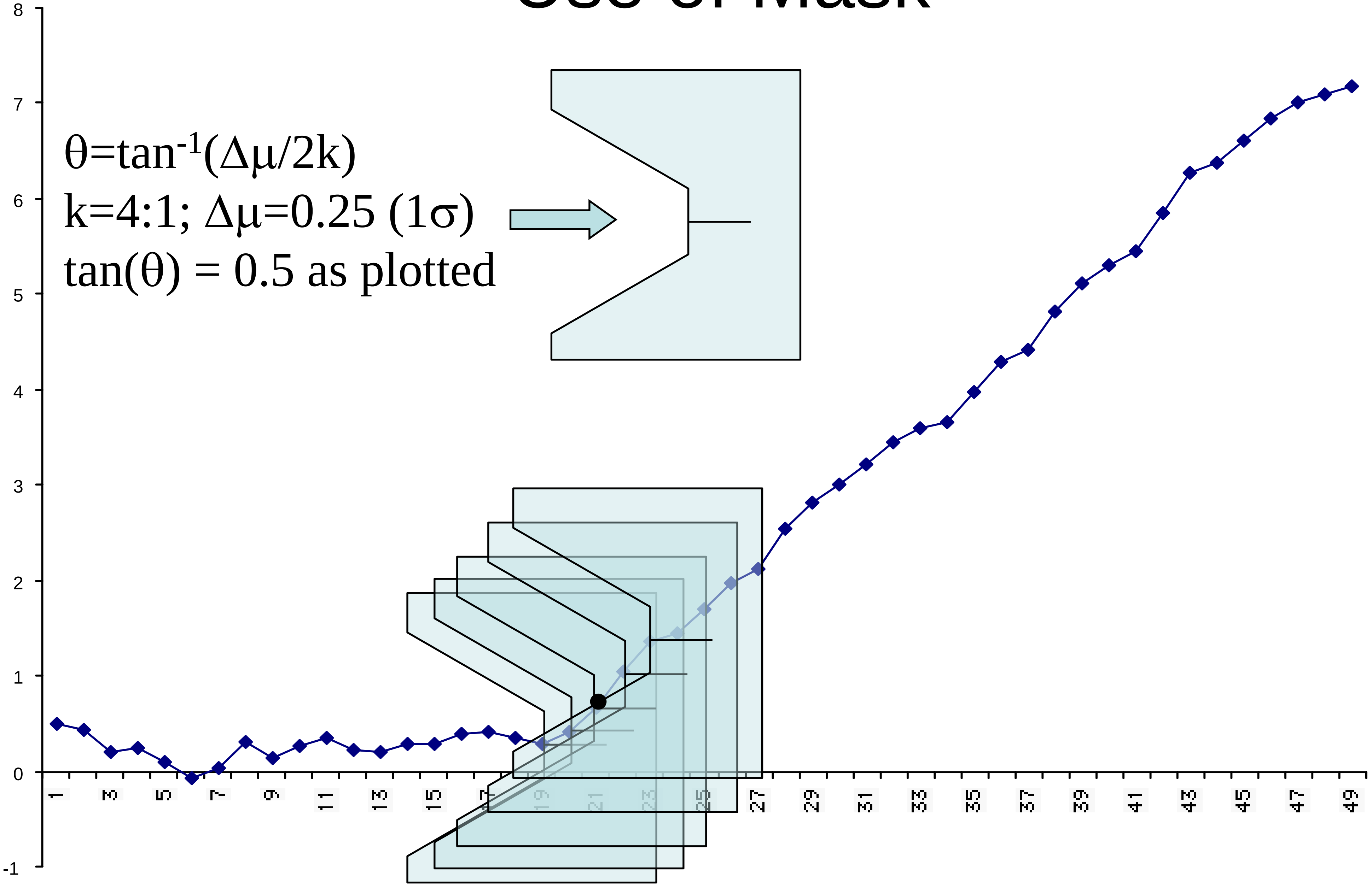
k = horizontal scale factor for plot

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

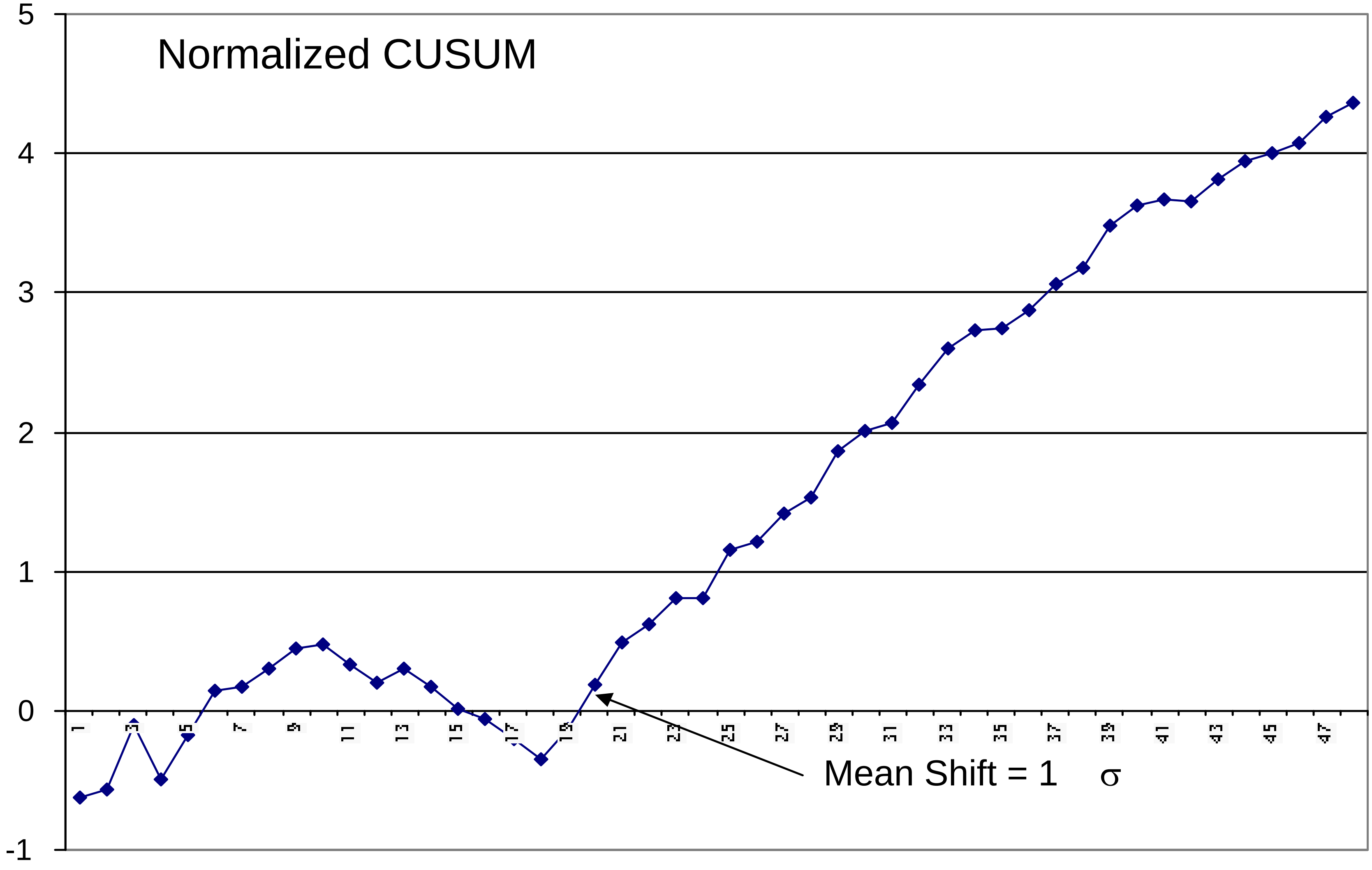
- And the CUSUM statistic

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

Which has an expected mean of 0 and variance of 1

Chart with Centerline =0 and Limits = ±3

Example for Mean Shift = 1σ



Manufacturing

Tabular CUSUM

- Create Threshold Variables:

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

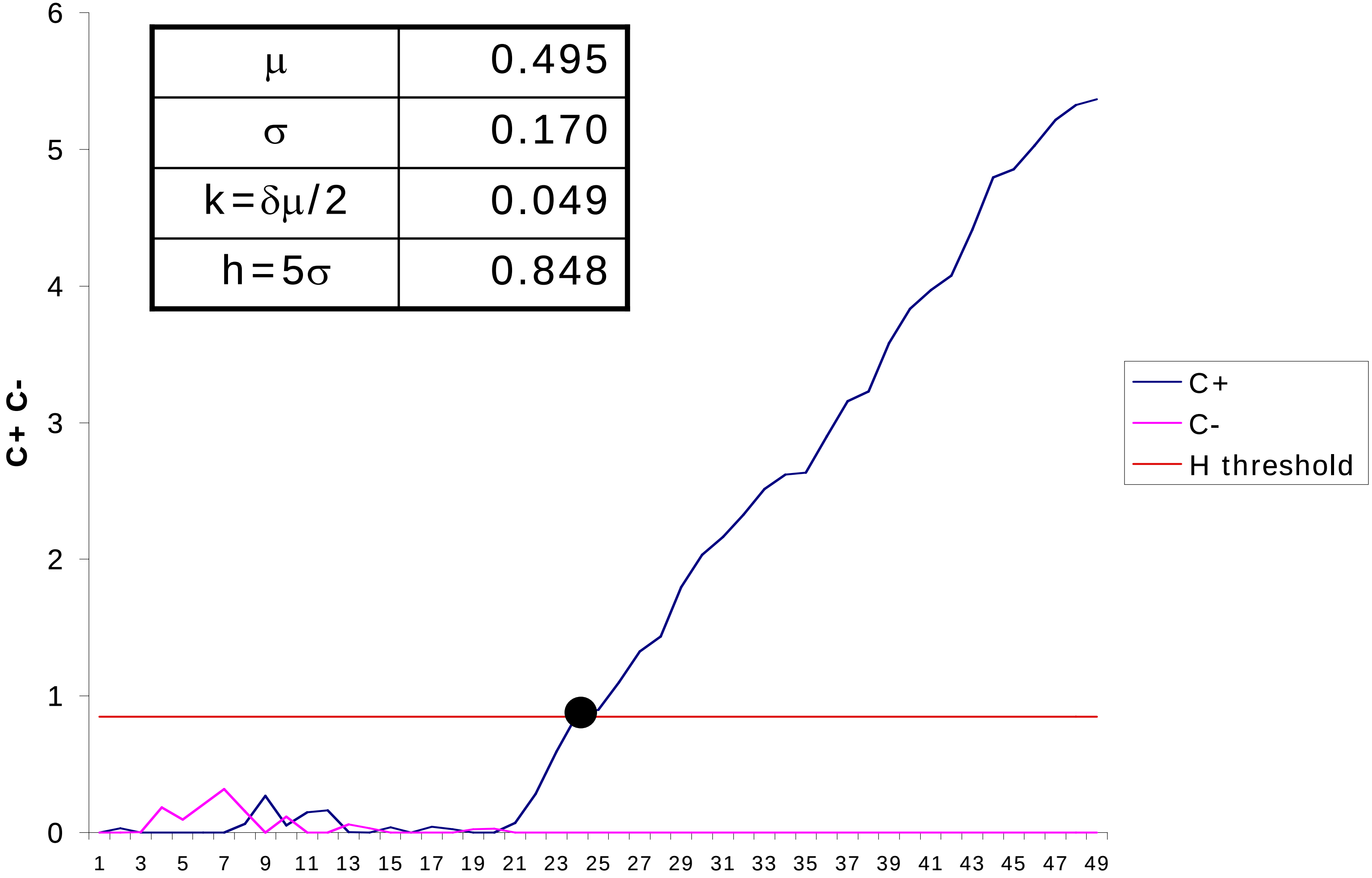
Accumulates deviations from the mean

K= threshold or slack value for accumulation

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

H : alarm level (typically 5σ)

Threshold Plot



μ	0.495
σ	0.170
$k=\delta\mu/2$	0.049
$h=5\sigma$	0.848

Alternative Charts Summary

- Noisy data need some filtering
- Sampling strategy can guarantee independence
- Linear discrete filters been proposed
  - EWMA
  - Running Integrator
- Choice depends on nature of process
- Noisy data need some filtering, BUT
  - Should generally monitor variance too!

Motivation: Multivariate Process Control

- More than one output of concern
  - many univariate control charts
  - many false alarms if not designed properly
  - common mistake #1
- Outputs may be coupled
  - exhibit covariance
  - independent probability models may not be appropriate
  - common mistake #2

Mistake #1 – Multiple Charts

- Multiple (independent) parameters being monitored at a process step
  - set control limits based on acceptable α = Pr(false alarm)
    - E.g., $\alpha = 0.0027$ (typical 3σ control limits), so 1/370 runs will be a false alarm
  - Consider *p* separate control charts
    - What is aggregate false alarm probability?

$$\alpha' = 1 - (1 - \alpha)^p$$
$$\alpha' \approx p\alpha$$

Mistake #1 – Multiple Tests for Significant Effects

- Multiple control charts are just a running hypothesis test – is process “in control” or has something *statistically significant* occurred (i.e., “unlikely to have occurred by chance”)?
- Same common mistake (testing for multiple significant effects and misinterpreting significance) applies to *many* uses of statistics – such as medical research!

The Economist (Feb. 22, 2007)

Text removed due to copyright restrictions. Please see *The Economist*, Science and Technology. “Signs of the times.” February 22, 2007.

Approximate Corrections for Multiple (Independent) Charts

- Approach: fixed α'
  - Decide aggregate acceptable false alarm rate, α'
  - Set individual chart α to compensate

$$\alpha = \alpha'/p$$

  - Expand individual control chart limits to match

$$UCL, LCL = \mu \pm z_{\alpha/2} \cdot \sigma_n$$

Mistake #2: Assuming Independent Parameters

- Performance related to many variables
- Outputs are often interrelated
  - e.g., two dimensions that make up a fit
  - thickness and strength
  - depth and width of a feature (e.g.. micro embossing)
  - multiple dimensions of body in white (BIW)
  - multiple characteristics on a wafer
- Why are independent charts deceiving?

Examples

- Body in White (BIW) assembly
  – Multiple individual dimensions measured
  – All could be OK and yet BIW be out of spec

- Injection molding part with multiple key dimensions

- Numerous critical dimensions on a semiconductor wafer or microfluidic chip

LFM Application

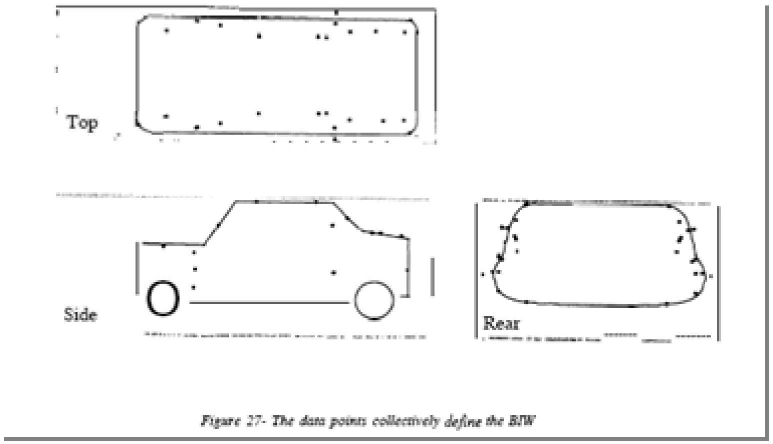


Figure 27- The data points collectively define the BIW

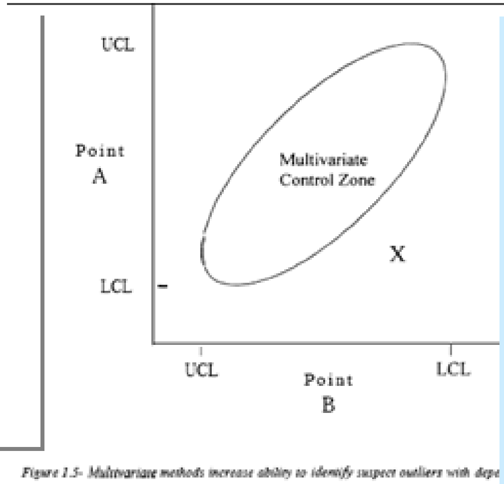


Figure 1.5- Multivariate methods increase ability to identify suspect outliers with depe

Rob York '95

"Distributed Gaging Methodologies for Variation Reduction in and Automotive Body Shop"

relationship	meaning	+	
1	% width	too wide	too narrow
2	% width	too wide	too narrow
3	% A or V gap		
4	% skew	O	

Body in White "Indicator"

Measurements: A, B,C,D

Design: A',B',C',D'

$$y_1 = \frac{|A-B| - |A'-B'|}{A'-B'} \qquad y_2 = \frac{|C-D| - |C'-D'|}{C'-D'}$$

$$y_3 = \frac{\left[|A-B| - |A'-B'|\right] - \left[|C-D| - |C'-D'|\right]}{\frac{1}{2}\left[|A'-B'| + |C'-D'|\right]}$$

$$y_4 = \frac{\left[|A+B| - |A'+B'|\right] - \left[|C+D| - |C'+D'|\right]}{|A'-B'| + |C'-D'|}$$

Figure 28- Four equations capture the fundamental relationships

A Methodology For Process Improvement Applied to Ultrasonic Welding

by

Michael J. Mihelick

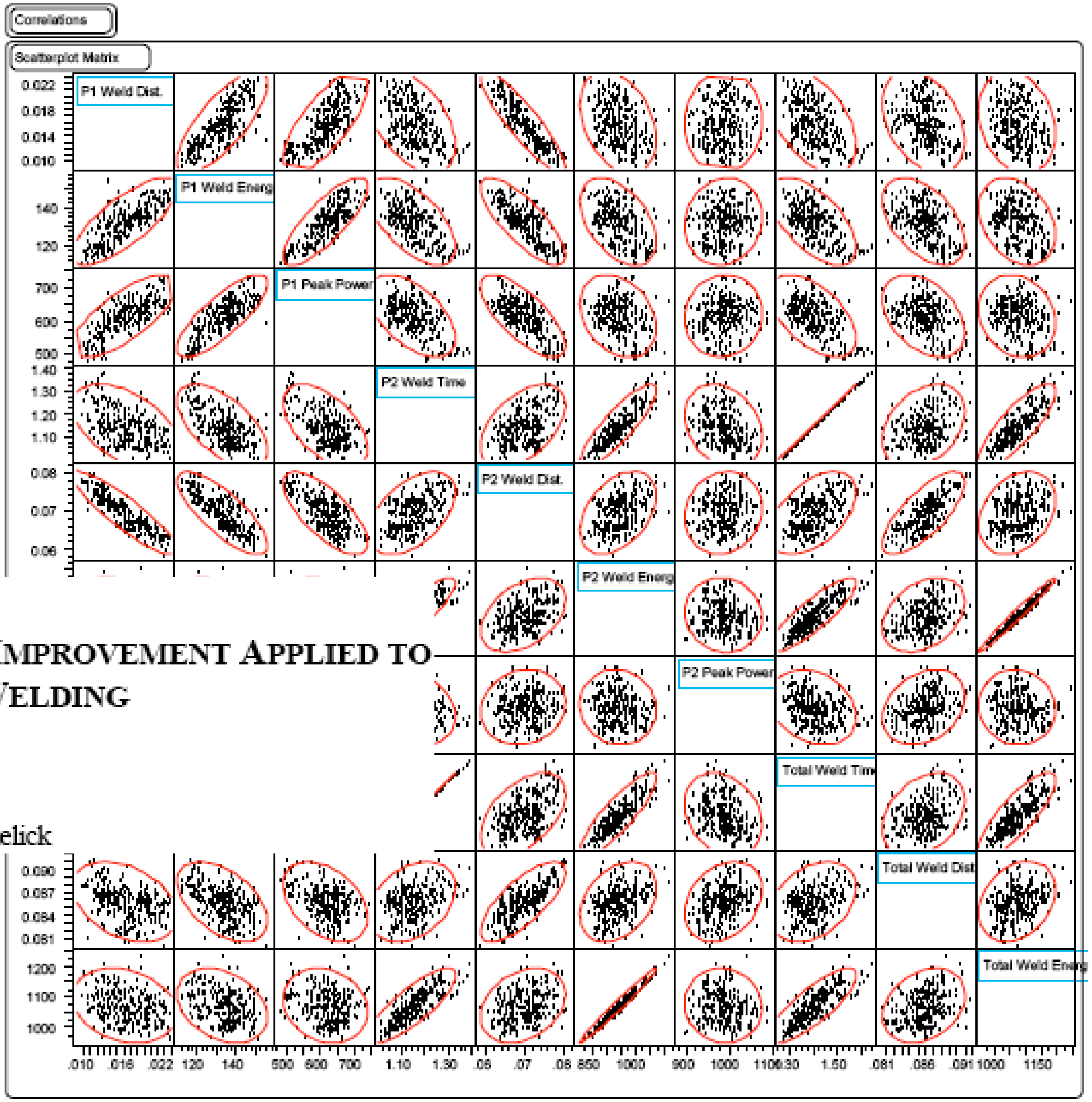
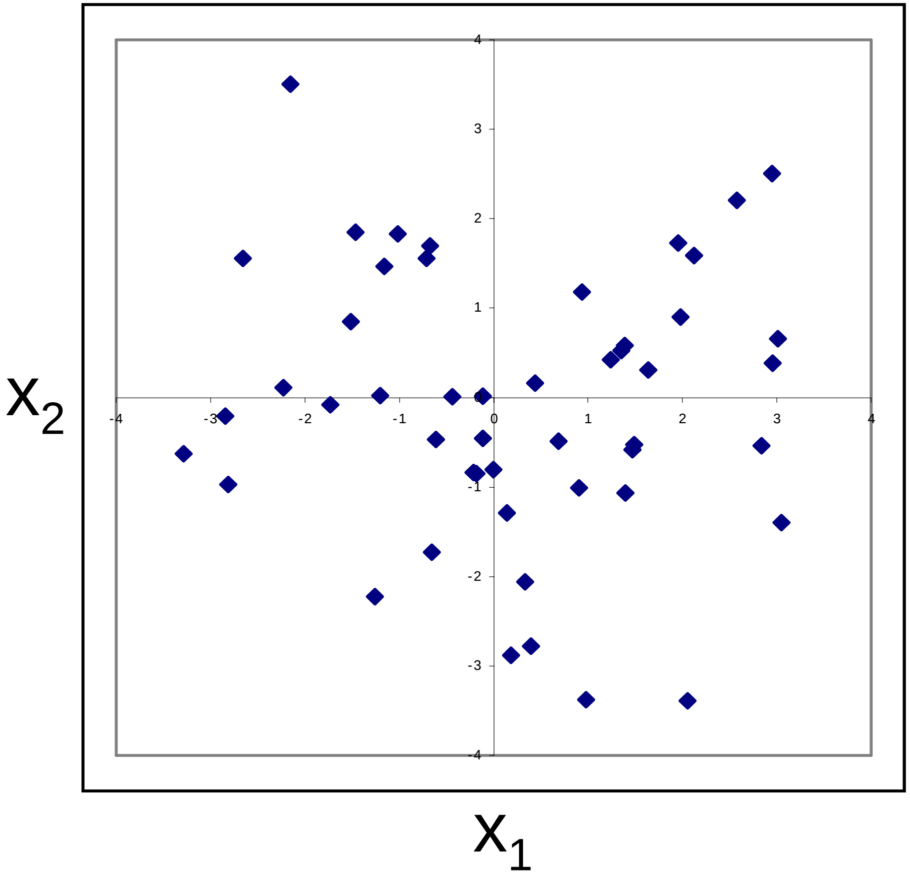
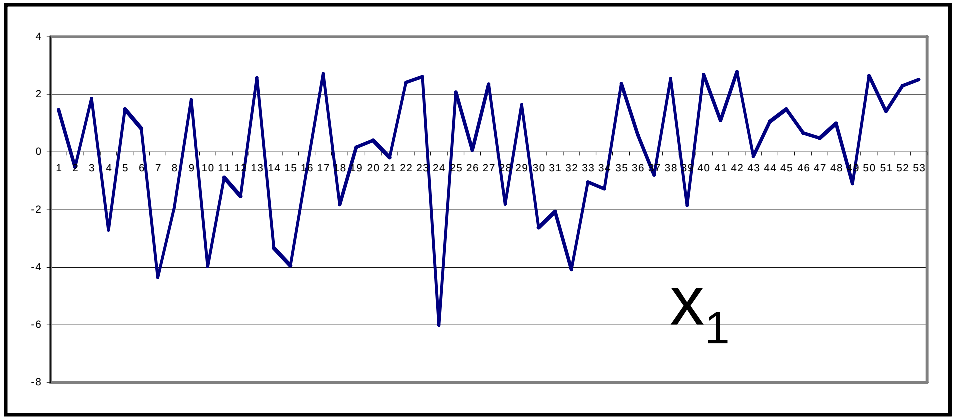
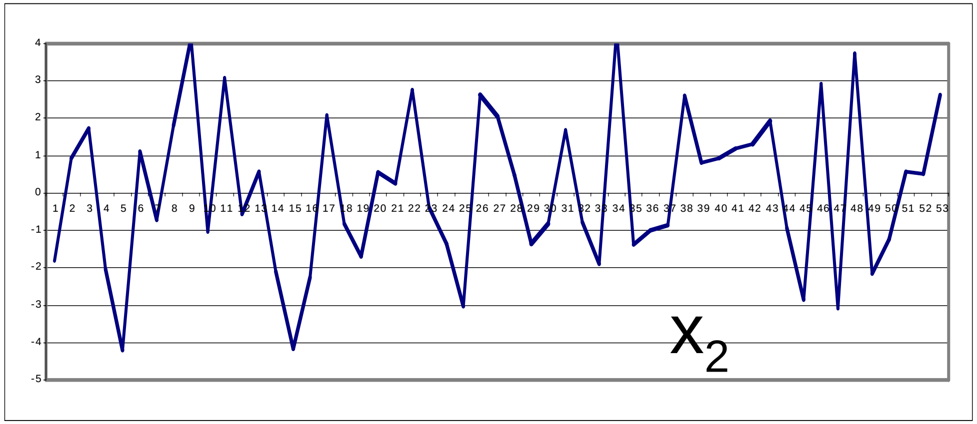


Figure 4.13 Scatterplot Matrix for Weld Parameters

Independent Random Variables



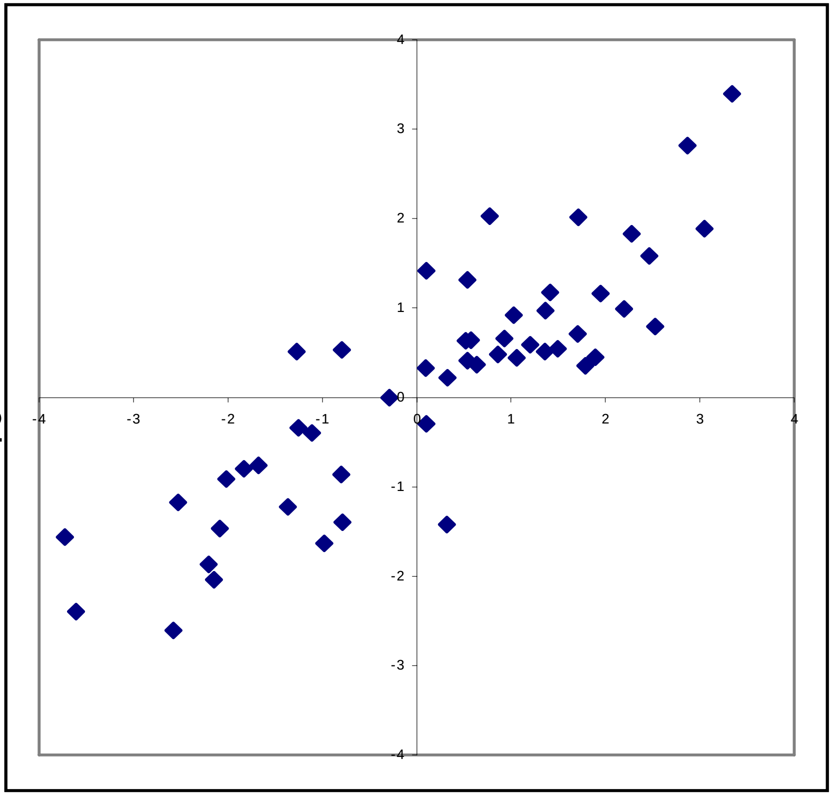




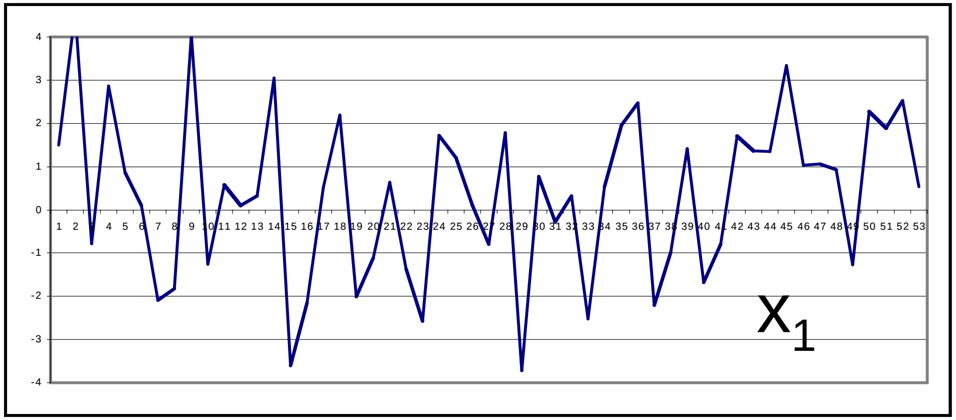
Proper Limits?

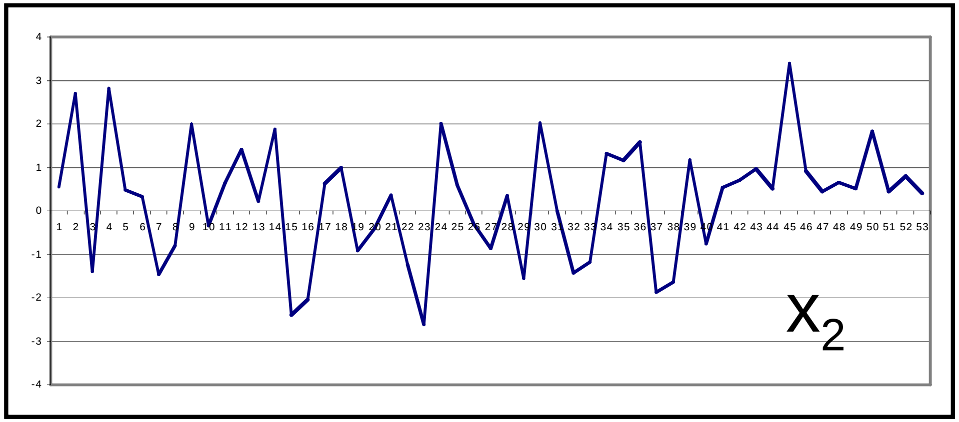
Correlated Random Variables

x_2



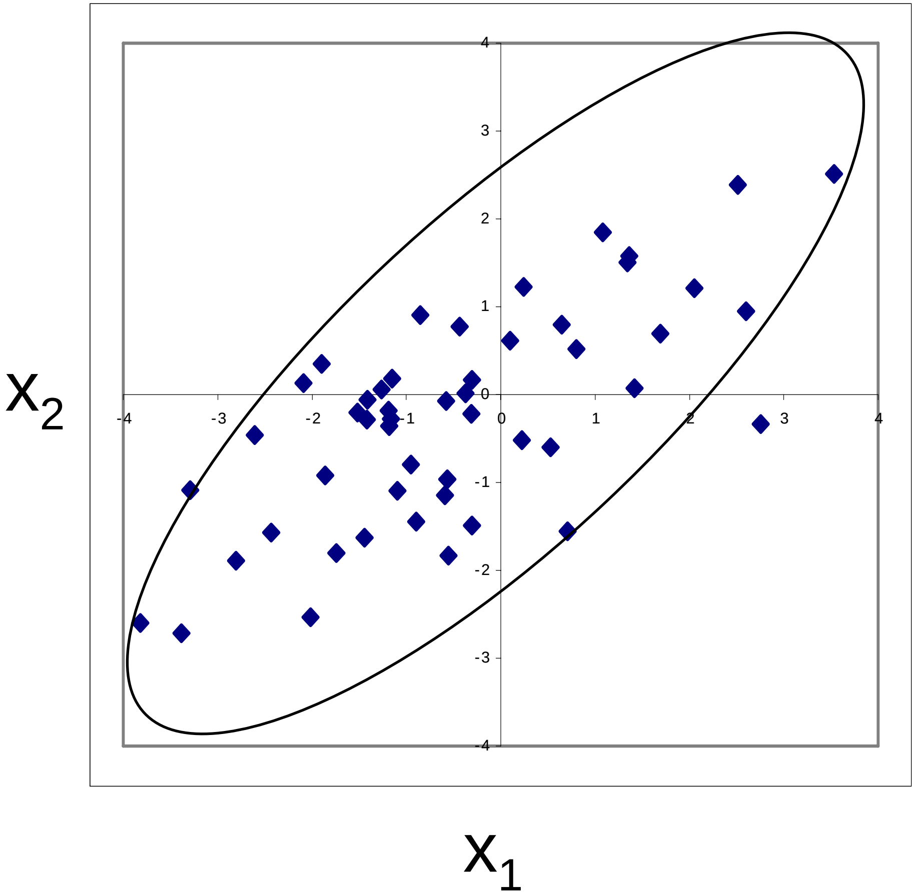
x_1

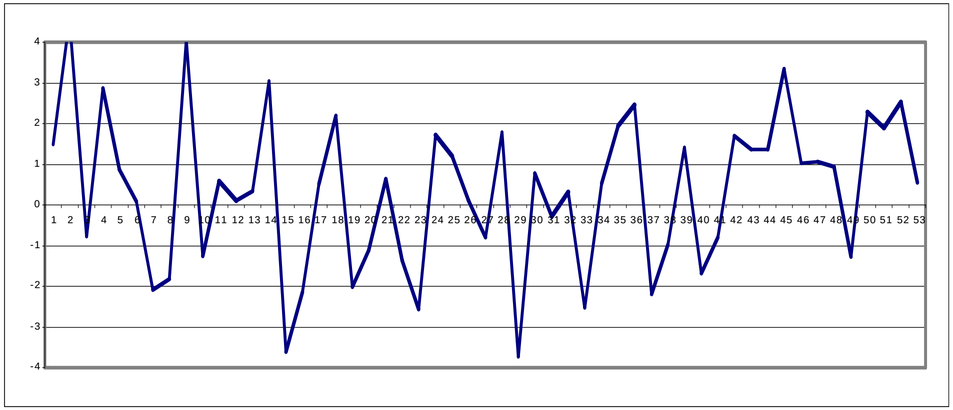


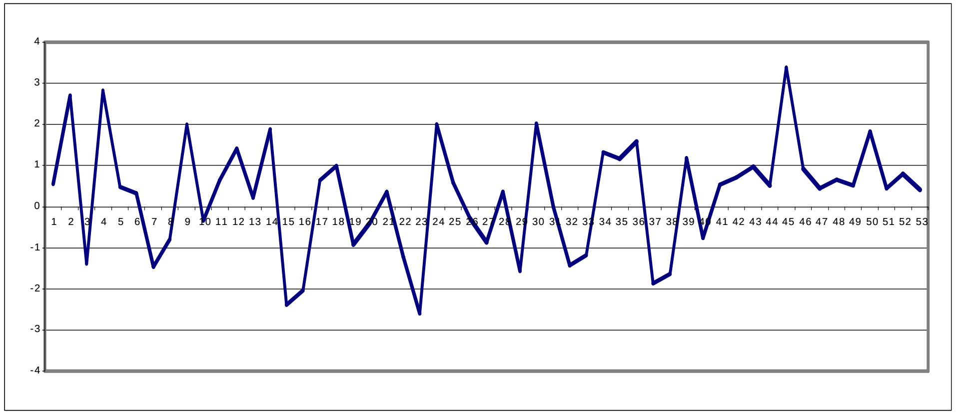


Proper Limits?

Outliers?







Multivariate Charts

- Create a single chart based on joint probability distribution
  - Using sample statistics: Hotelling T^2
- Set limits to detect mean shift based on α
- Find a way to back out the underlying causes
- EWMA and CUSUM extensions
  - MEWMA and MCUSUM

Background

- Joint Probability Distributions
- Development of a single scale control chart
  - Hotelling T^2
- Causality Detection
  - Which characteristic likely caused a problem
- Reduction of Large Dimension Problems
  - Principal Component Analysis (PCA)

Multivariate Elements

- Given a vector of measurements

$$\underline{x} = [x_1, x_2, x_3, \ldots, x_p]$$

- We can define vector of means:

$$\underline{\mu} = [\mu_1, \mu_2, \mu_3, \ldots, \mu_p]$$

where *p* = # parameters

- and covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots \\ \sigma_{21}^2 & \sigma_{22}^2 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}$$

variance

covariance

Joint Probability Distributions

- Single Variable Normal Distribution

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

squared standardized distance from mean

- Multivariable Normal Distribution

$$f(\underline{x}) = \frac{1}{(2\pi)^{p/2}\left|\Sigma\right|^{1/2}} e^{-\frac{1}{2}\underbrace{(\underline{x}-\underline{\mu})^T \Sigma^{-1} (\underline{x}-\underline{\mu})}_{\text{squared standardized distance from mean}}}$$

squared standardized distance from mean

Sample Statistics

- For a set of samples of the vector $\underline{x}$

$$X = [\underline{x}_1, \underline{x}_2, \underline{x}_3, \ldots, \underline{x}_n]$$

- Sample Mean

$$\underline{\bar{x}} = \frac{1}{n}\sum_{i=1}^{n} \underline{x}_i$$

- Sample Covariance

$$S = \frac{1}{n-1}\sum_{i=1}^{n} (\underline{x}_i - \underline{\bar{x}})(\underline{x}_i - \underline{\bar{x}})^T$$

Chi-Squared Example - True Distributions Known, Two Variables

- If we know $\underline{\mu}$ and Σ *a priori*:

$$\chi_0^2 = \frac{n}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2} \begin{array}{l}[\sigma_2^2(\bar{x}_1 - \mu_1)^2 + \sigma_1^2(\bar{x}_2 - \mu_2)^2 \\ -2\sigma_{12}(\bar{x}_1 - \mu_1)(\bar{x}_2 - \mu_2)]\end{array}$$

will be distributed as χ_2^2

  – sum of squares of two unit normals

- More generally, for *p* variables:

$$\chi_0^2 = n(\underline{x} - \underline{\mu})^T \Sigma^{-1} (\underline{x} - \underline{\mu})$$

distributed as χ_p^2 (and *n* = # samples)

Control Chart for χ^2?

- Assume an acceptable probability of Type I errors (upper α)
- $UCL = \chi^2_{\alpha, p}$
  - where p = order of the system
- If process means are μ_1 and μ_2 then $\chi^2_0 < UCL$

Univariate vs. χ^2 Chart

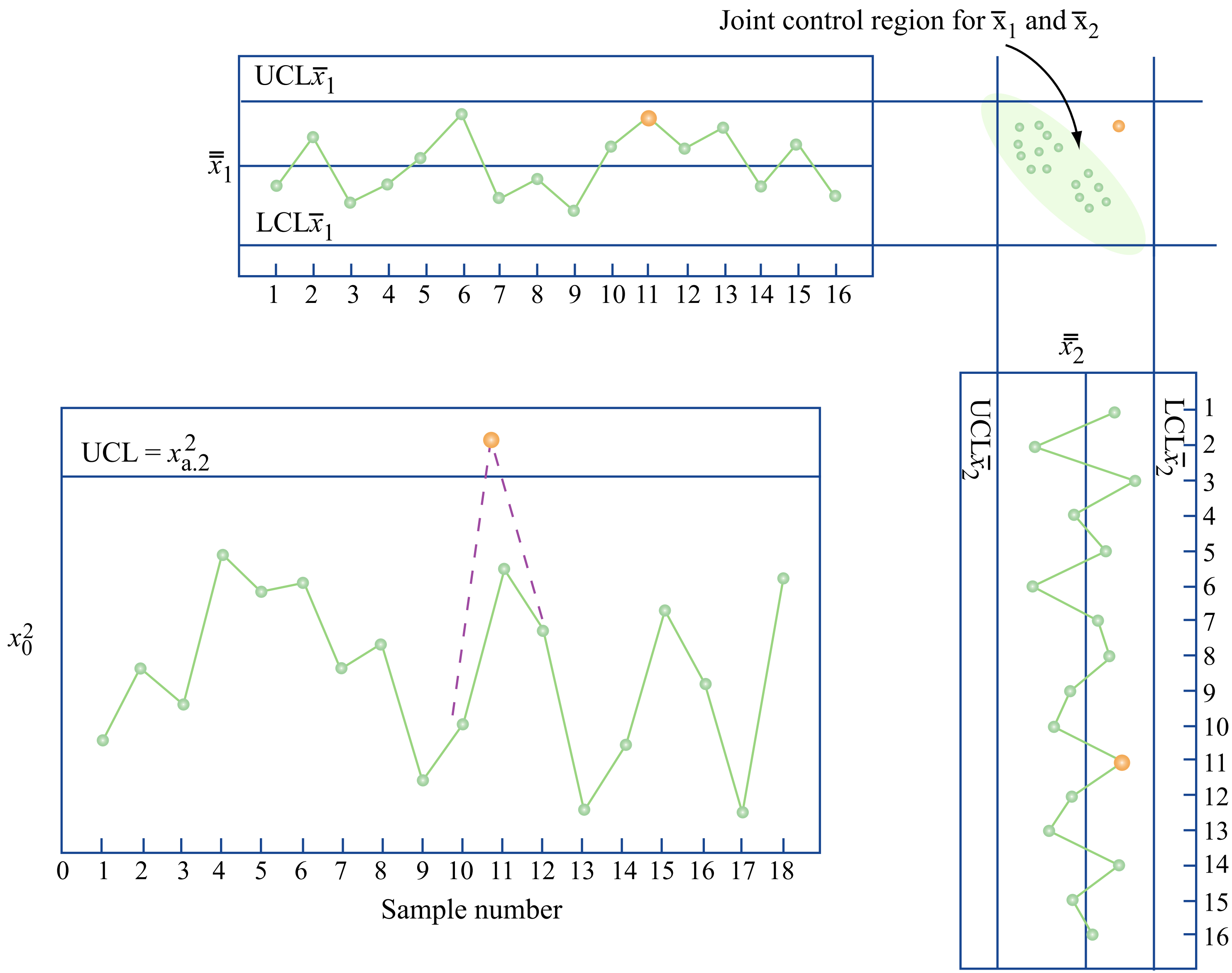


Figure by MIT OpenCourseWare.

Multivariate Chart with No Prior Statistics: T^2

- If we must use data to get $\underline{\bar{x}}$ and S
- Define a new statistic, Hotelling T^2

$$T^2 = n(\underline{\bar{x}} - \underline{\bar{\bar{x}}})^T S^{-1} (\underline{\bar{x}} - \underline{\bar{\bar{x}}})$$

- Where $\underline{\bar{\bar{x}}}$ is the vector of the averages for each variable over all measurements
- S is the matrix of sample *covariance* over all data

Similarity of T^2 and t^2

$$T^2 = n(\underline{\bar{x}} - \underline{\bar{\bar{x}}})^T S^{-1} (\underline{\bar{x}} - \underline{\bar{\bar{x}}})$$

vs.

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$t^2 = \frac{n(\bar{x} - \mu)^2}{s^2}$$

$$t^2 = n(\bar{x} - \mu) s^{-2} (\bar{x} - \mu)$$

Distribution for T^2

– Given by a scaled F distribution

$$LCL = 0$$
$$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,(mn-m-p+1)}$$

α is type I error probability

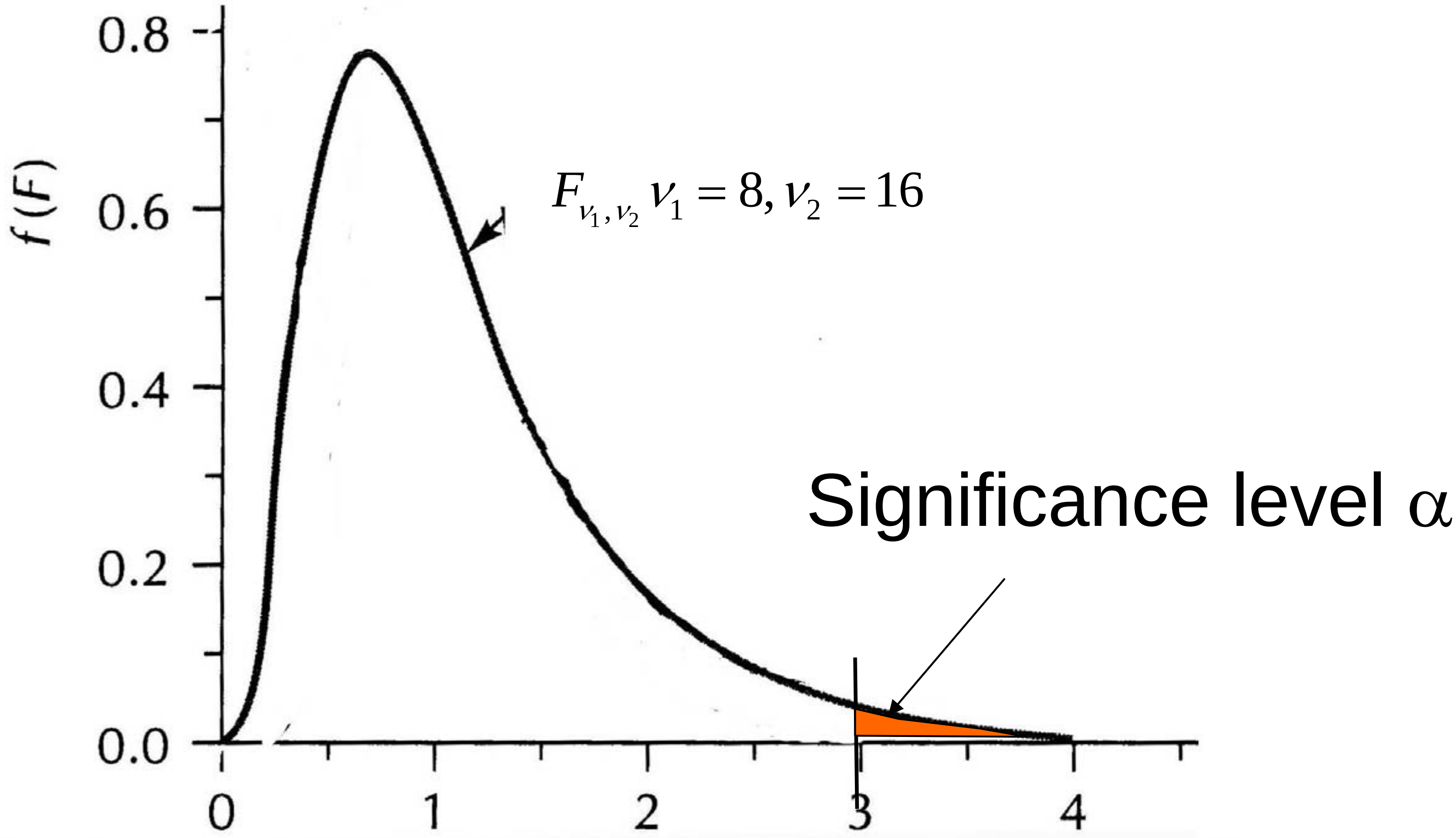
p and ($mn - m - p + 1$) are d.o.f. for the F distribution

n is the size of a given sample

m is the number of samples taken

p is the number of outputs

F - Distribution



Tabulated Values → $F_{\alpha,p,(mn-m-p+1)}$

Phase I and II?

- Phase I - Establishing Limits

$$UCL = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,(mn-m-p+1)}$$

- Phase II - Monitoring the Process

$$UCL = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha,p,(mn-m-p+1)}$$

NB if *m* used in phase 1 is large then they are nearly the same

Example

- Fiber production
- Outputs are strength and weight
- 20 samples of subgroups size 4
  - $m = 20$, $n = 4$
- Compare T^2 result to individual control charts

Data Set

		Output x1					Output x2		
Sample		Subgroup n=4			R1		Subgroup n=4		
1	80	82	78	85	7	19	22	20	20
2	75	78	84	81	9	24	21	18	21
3	83	86	84	87	4	19	24	21	22
4	79	84	80	83	5	18	20	17	16
5	82	81	78	86	8	23	21	18	22
6	86	84	85	87	3	21	20	23	21
7	84	88	82	85	6	19	23	19	22
8	76	84	78	82	8	22	17	19	18
9	85	88	85	87	3	18	16	20	16
10	80	78	81	83	5	18	19	20	18
11	86	84	85	86	2	23	20	24	22
12	81	81	83	82	2	22	21	23	21
13	81	86	82	79	7	16	18	20	19
14	75	78	82	80	7	22	21	23	22
15	77	84	78	85	8	22	19	21	18
16	86	82	84	84	4	19	23	18	22
17	84	85	78	79	7	17	22	18	19
18	82	86	79	83	7	20	19	23	21
19	79	88	85	83	9	21	23	20	18
20	80	84	82	85	5	18	22	19	20

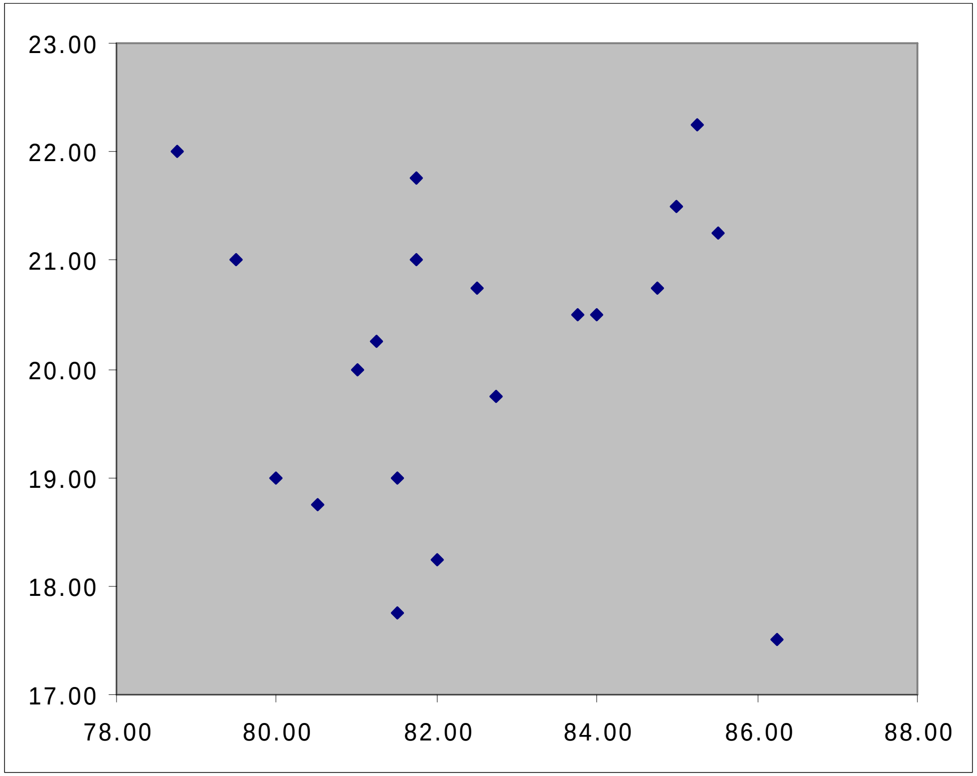
Mean Vector

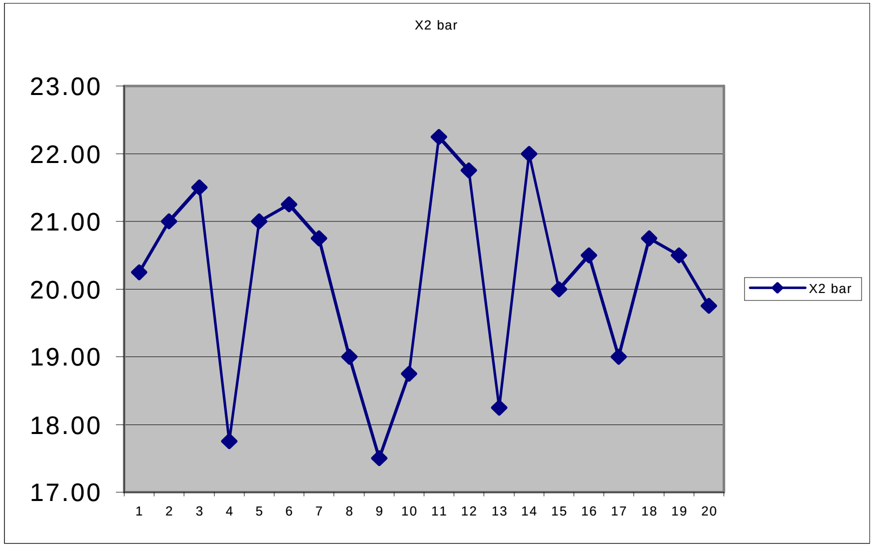
$$\begin{vmatrix} 82.46 \\ 20.18 \end{vmatrix}$$

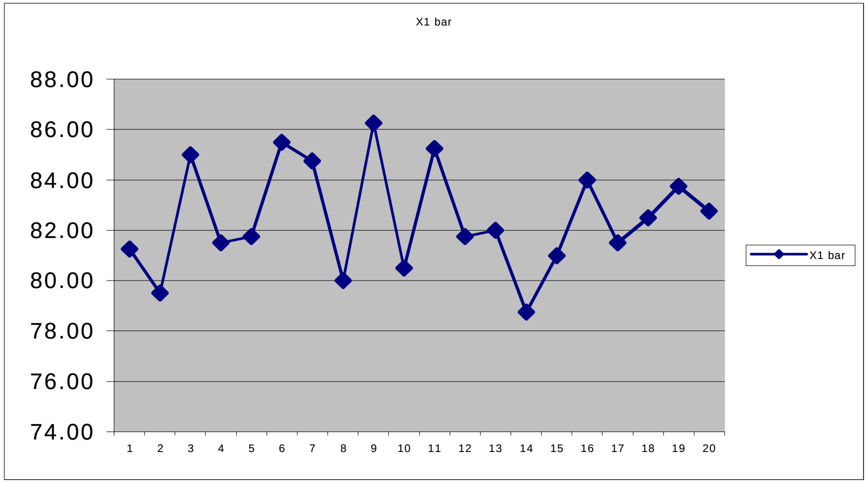
Covariance Matrix

$$\begin{vmatrix} 7.51 & -0.35 \\ -0.35 & 3.29 \end{vmatrix}$$

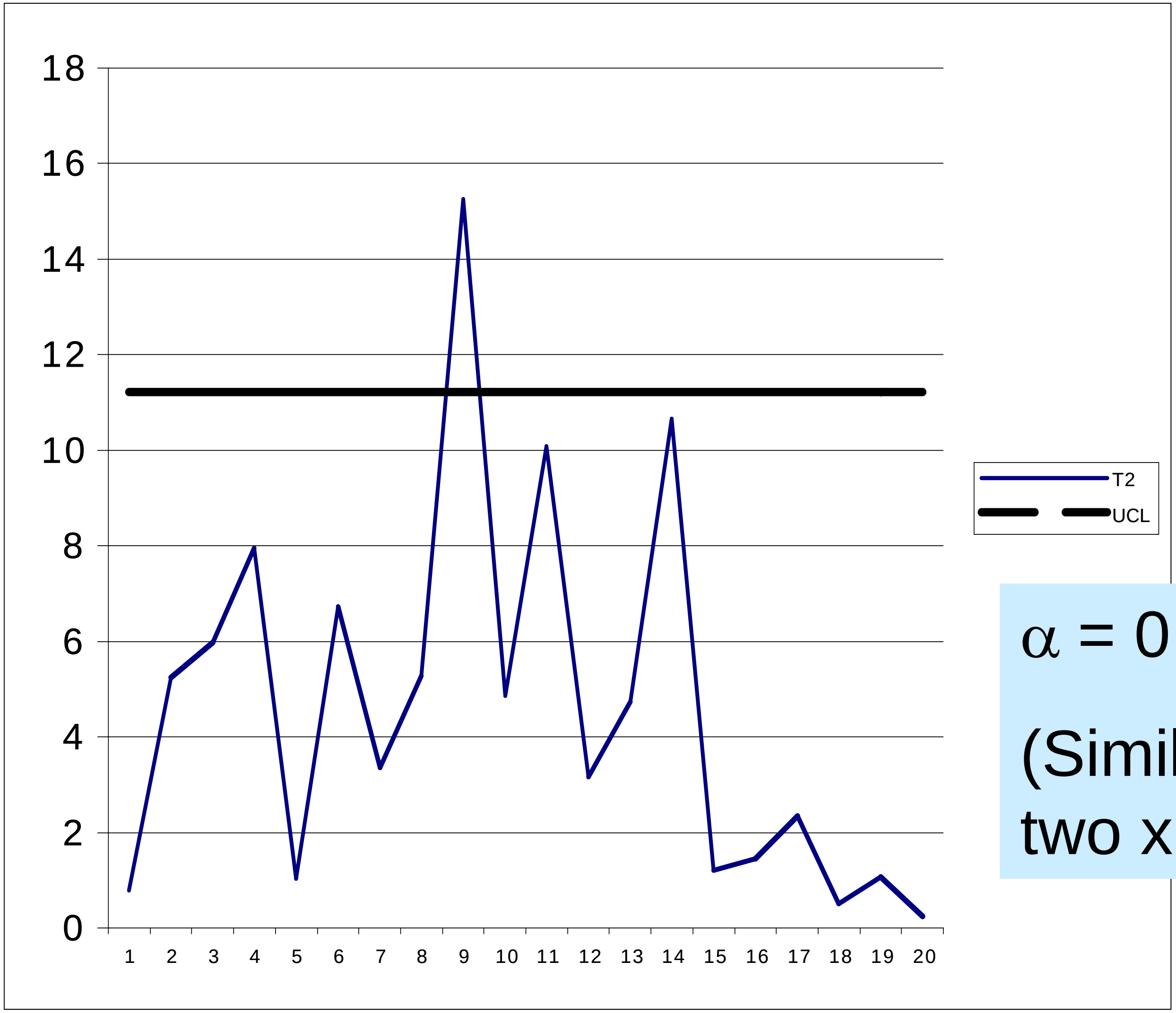
Cross Plot of Data







T^2 Chart

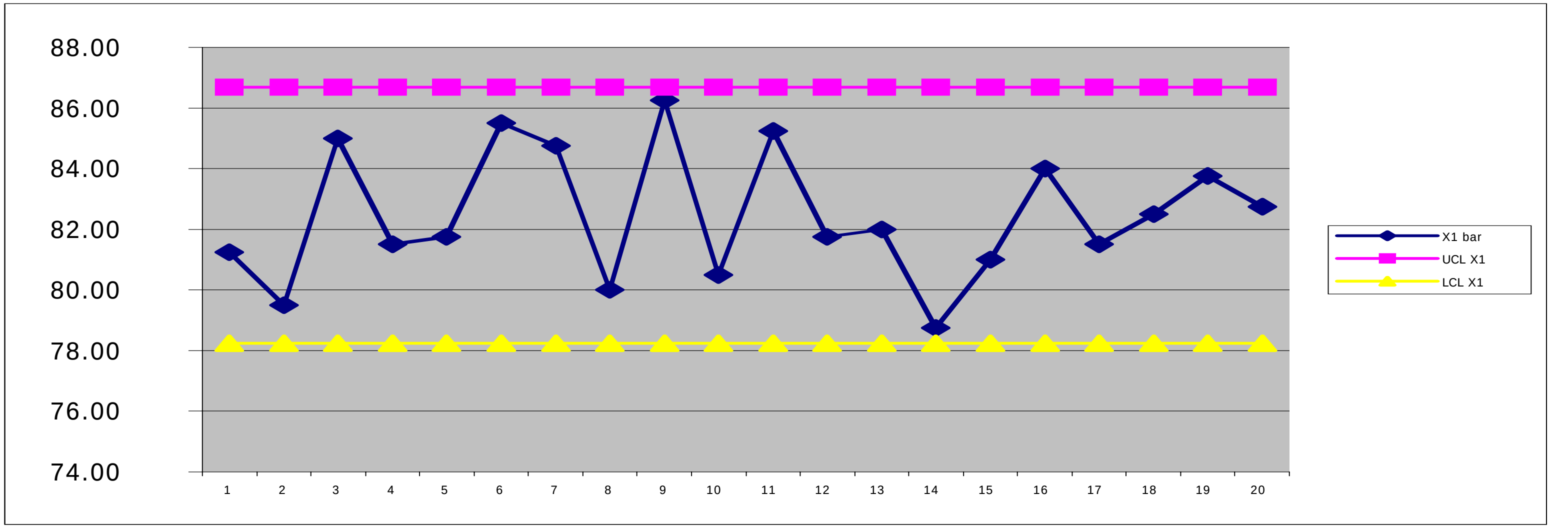


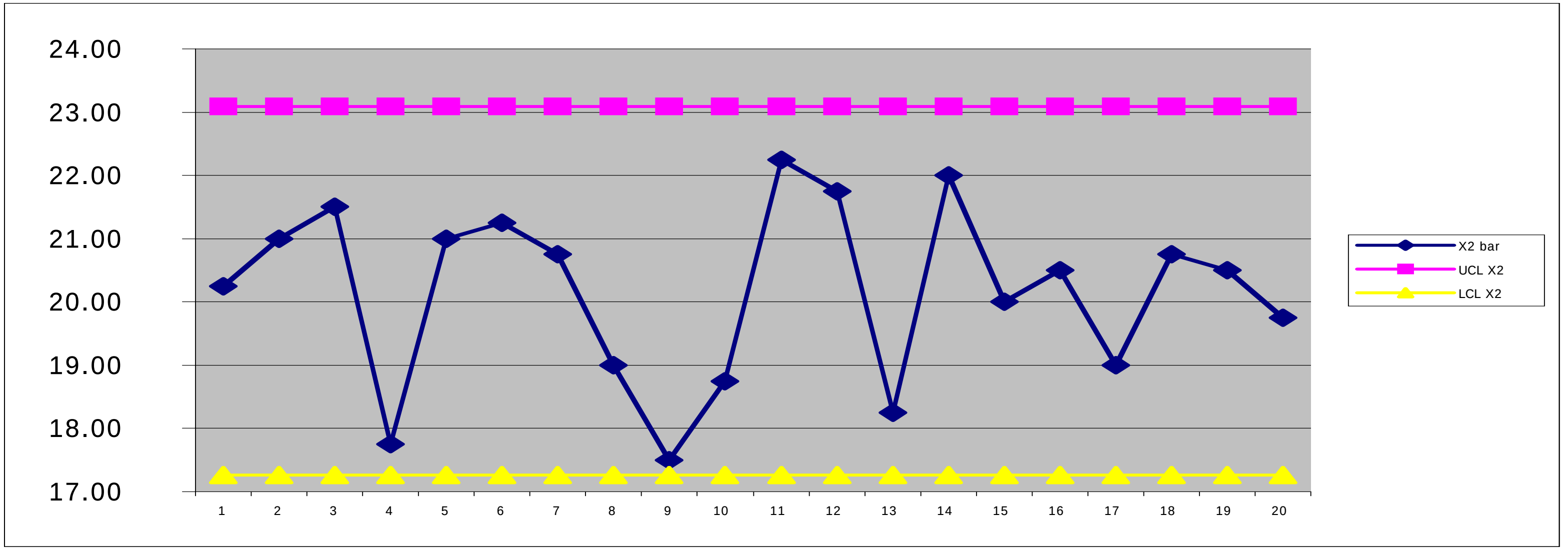
α = 0.0054

(Similar to $\pm 3\sigma$ limits on two xbar charts combined)

Manufacturing

Individual Xbar Charts





Finding Cause of Alarms

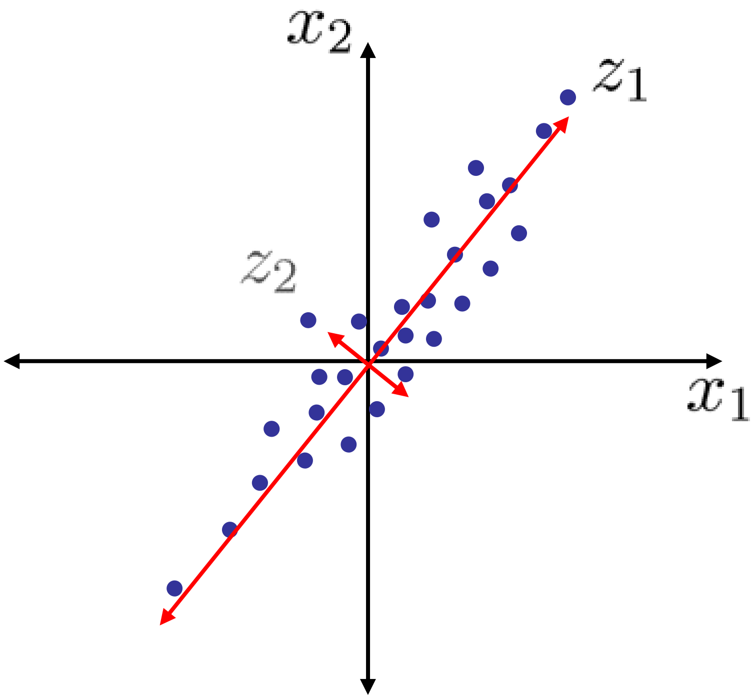
- With only one variable to plot, which variable(s) caused an alarm?
- Montgomery
  - Compute T^2
  - Compute $T^2_{(i)}$ where the i th variable is not included
  - Define the relative contribution of each variable as
    - $d_i = T^2 - T^2_{(i)}$

Principal Component Analysis

- Some systems may have *many* measured variables p
  - Often, strong correlation among these variables: actual degrees of freedom are fewer
- Approach: reduce order of system to track only $q \ll p$ variables
  - where each $z_1 \ldots z_q$ is a linear combination of the measured $x_1 \ldots x_p$ variables

Principal Component Analysis

- x_1 and x_2 are highly correlated in the z_1 direction
- Can define new axes z_i in order of decreasing variance in the data
- The z_i are *independent*
- May choose to neglect dimensions with only small contributions to total variance $\Rightarrow$ dimension reduction



$$
\begin{aligned}
z_1 &= c_{11}x_1 + c_{12}x_2 + \ldots + c_{1p}x_p \\
z_2 &= c_{21}x_1 + c_{22}x_2 + \ldots + c_{2p}x_p \\
&\vdots \\
z_q &= c_{q1}x_1 + c_{q2}x_2 + \ldots + c_{qp}x_p
\end{aligned}
$$

Truncate at $q < p$

Principal Component Analysis

- Finding the c_{ij} that define the principal components:
  - Find Σ covariance matrix for data x
  - Let eigenvalues of Σ be $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$
  - Then constants c_{ij} are the elements of the i^{th} eigenvector associated with eigenvalue λ_{is}
    - Let C be the matrix whose columns are the eigenvectors
    - Then $C^T \Sigma C = \Lambda$

      where Λ is a p x p diagonal matrix whose diagonals are the eigenvalues
    - Can find C efficiently by singular value decomposition (SVD)
  - The fraction of variability explained by the i^{th} principal components is $\dfrac{\lambda_i}{\lambda_1 + \lambda_2 + ... + \lambda_p}$

Extension to EWMA and CUSUM

- Define a vector EWMA

$$Z_i = r\underline{x}_i + (1 + r)Z_{i-1}$$

- And for the control chart plot

where $$T_i^2 = Z_i^T \Sigma_{Z_i}^{-1} Z_i$$

$$\Sigma_{Z_i i} = \frac{r}{2-r}[1-(1-r)^2]\Sigma$$

Conclusions

- Multivariate processes need multivariate statistical methods
- Complexity of approach mitigated by computer codes
- Requires understanding of underlying process to see if necessary
  - i.e. if there is correlation among the variables of interest