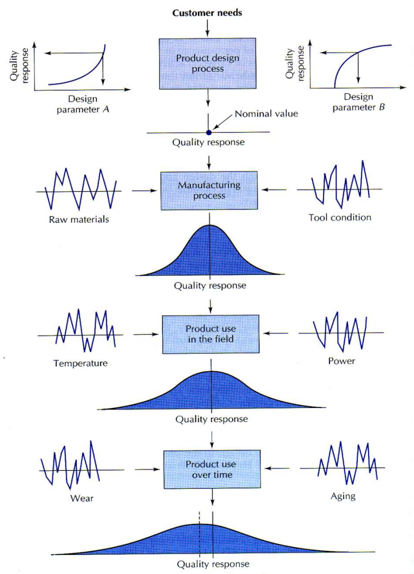
Statistical Process Control

Argon Chen (陳正剛)
Graduate Institute of Industrial Engineering
National Taiwan University

Impact of Variation on Quality Performance



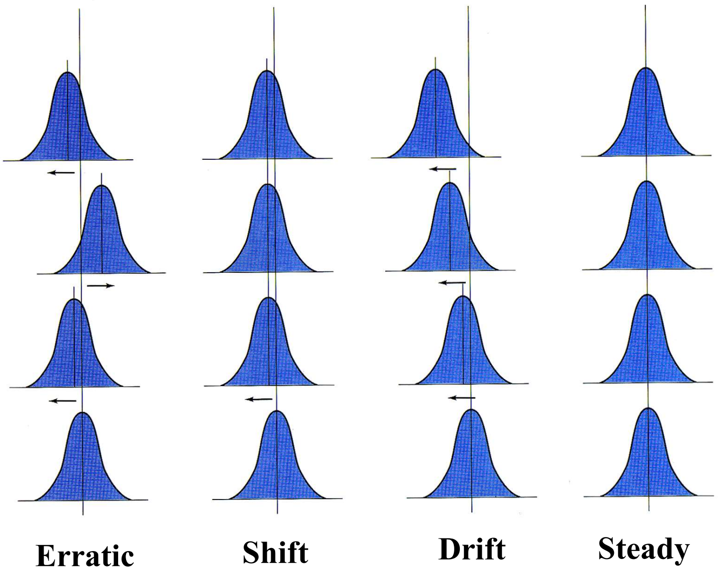
Variation Causes and Behavior

- Shewhart's Variation Causes
  – Chance (Common) Causes
    - the variation sources are under the normal conditions that are commonly observed in processes.
  – Special (Assignable) Causes
    - some variation source(s) goes beyond the normal conditions
- Process shift
  – mean level shift
  – process variation shift
  – mean and variation shift

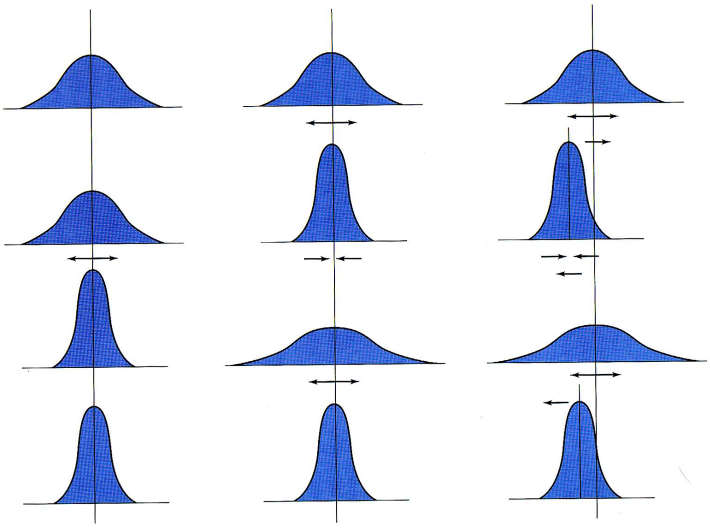
The Nature of Faults in the Process

Two fundamental types of faults or problems			
Faults	Local faults Special causes Sporadic problems Assignable causes :	versus	System faults Common causes Chronic problems Chance causes :
Examples	Broken tools Jammed machine Material contamination Human errors Accidents :	versus	Wrong specification Inappropriate method Poor supervision Poor training Poor design :
Action/ by Whom	Correctable locally (at the machine level) by the individual (operator or the first level of supervision)	versus	Requires a change in the system – only management can specify and implement the change

Behavior of Mean Level



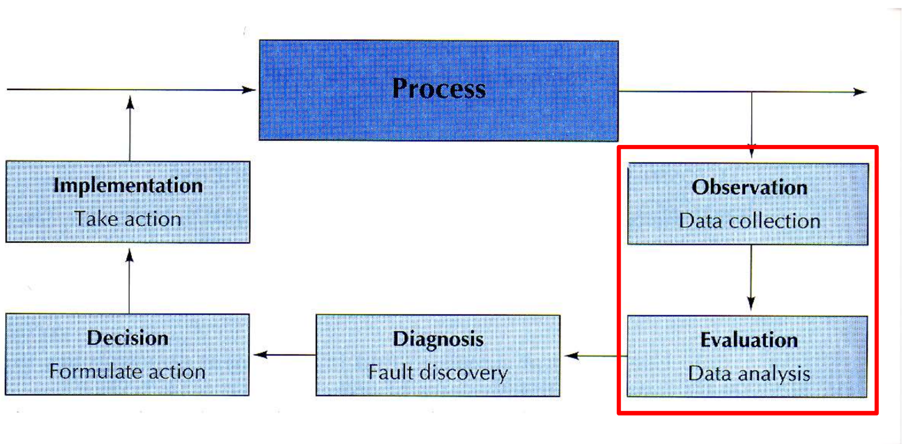
Behavior of the Variation Level



Statistical Process Control

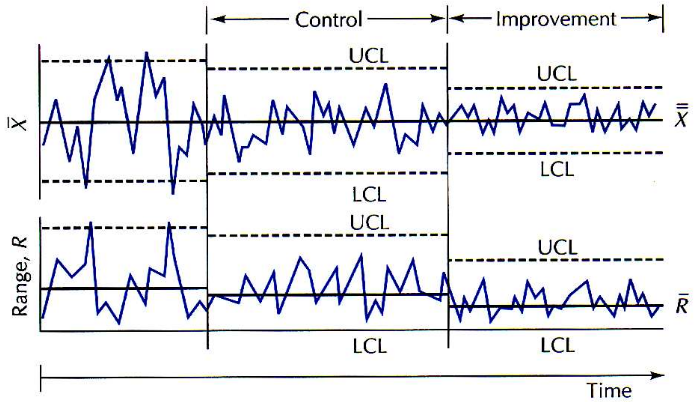
- A prime objective of a control chart is to detect **special (assignable) causes** of variation in a process.
- A control chart detects the presence of a special cause but does not "find" the cause.
- A process that is operating without special causes of variation (only under **common causes** of variation) is said to be "in a state of statistical control."

Classical Control System View of SPC Implementation



Moving from Instability to Control to Improvement

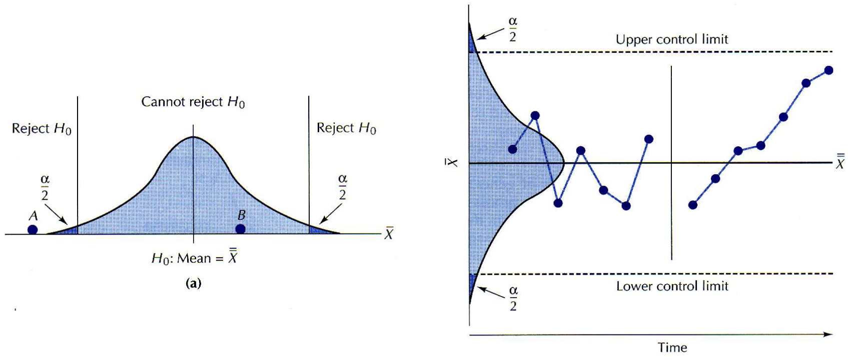
- **SPC**: detect and remove special cause of variation
- **Process improvement**: reduce common cause of variation



Statistical Process Control (SPC) Chart

- A scientific approach to detect process changes?
  - Hypothesis testing..…t-test? umh...too complicated...
  - Recall what hypothesis testing is
    - test statistic
    - rejection region
- SPC chart approach to detect process changes
  - sample statistic
  - out-of-control region
  - graphical trend chart

Hypothesis Testing vs. SPC Chart



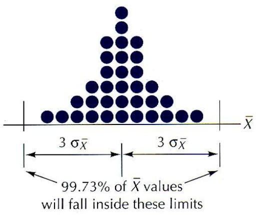
- We need a model to describe the variability behavior
- Is there any universal mathematical model that can describe the behavior? Yes, normal distribution (Central Limit Theorem)!

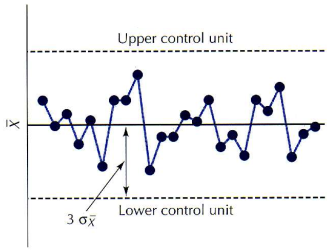
Shewhart Control Chart

- Sample (Test) statistic (*Y*):
  - $\bar{X}$ or *R*: variable characteristic (Ex: SiO_2 thickness, CD, etc.)
  - *p*: fraction defective (attribute characteristic) (Ex: number of defective wafers in a lot of wafers)
  - *c*: number of defects (attribute characteristic) (Ex: number of particles on a wafer)
- Out-of-control (Reject) region (Shewhart scheme):

  outside (***Expected Value*** **± 3×*****Standard Deviation***)
- Shewhart Control Chart
  - Central Line (CL): mean of *Y* (or target *T*)
  - Upper Control Limit (UCL): CL+3 $\hat{\sigma}_Y$
  - Lower Control Limit (LCL): CL–3 $\hat{\sigma}_Y$

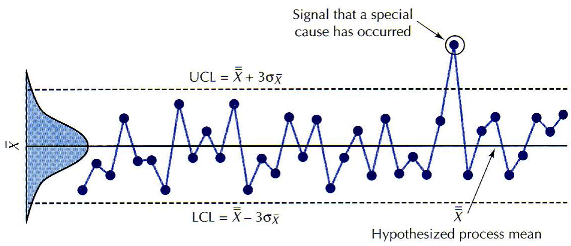
Control Chart for Central Tendency





$Y = \bar{X} = \sum_{i=1}^{n} X_i / n$

n : sample size



$CL = \bar{Y} = \bar{\bar{X}}$ (or T)

$UCL = \bar{\bar{X}}$ (or T) $+ 3\hat{\sigma}_{\bar{X}}$

$LCL = \bar{\bar{X}}$ (or T) $- 3\hat{\sigma}_{\bar{X}}$

Estimation of $\sigma_{\bar{X}}$

- Estimating $\sigma_{\bar{X}}$
  sample standard deviation

$$\hat{\sigma}_{\bar{X}} = S_{\bar{X}} = \sqrt{\frac{\sum_i (\bar{X}_i - \bar{\bar{X}})^2}{k-1}}$$

- **Do not use** $\sigma_{\bar{X}} = \dfrac{\sigma_X}{\sqrt{n}} \Rightarrow \cancel{\hat{\sigma}_{\bar{x}} = \dfrac{S_x}{\sqrt{n}}}$

(Variation sources of X and $\bar{X}$ are different)

Control Chart for Dispersion

- Control chart (hypothesis test) for dispersion
  - test statistic: range (R)
  - rejection region: above or below 3 standard deviation of R

$$\overline{R} \pm 3 \cdot \hat{\sigma}_R$$

- Estimating σ_R :

$$\hat{\sigma}_R = S_R = \sqrt{\frac{\sum_i (R_i - \overline{R})^2}{k-1}}$$

Control Limit Calculations

Given $\hat{\sigma}_{\bar{x}} = S_{\bar{x}}$ and $\hat{\sigma}_R = S_R$

Average chart

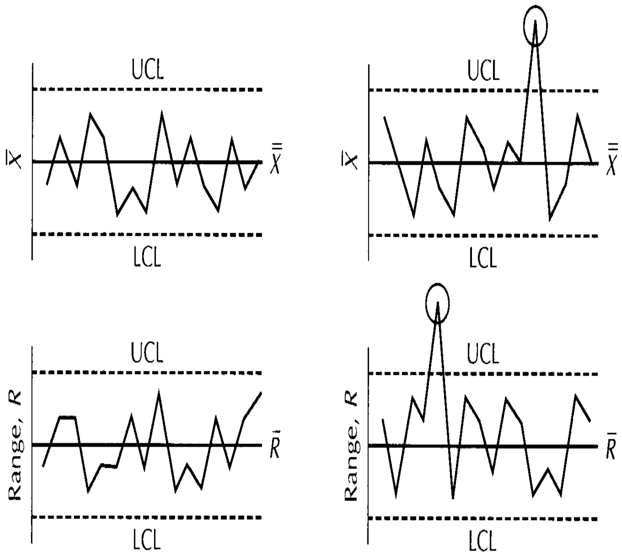
$$UCL(LCL) = \bar{\bar{X}} + (-)3\hat{\sigma}_{\bar{x}} = \bar{\bar{X}} + (-)3S_{\bar{x}}$$

$$\cancel{UCL(LCL) = \bar{\bar{X}} + (-)3\frac{\hat{\sigma}_x}{\sqrt{n}} = \bar{\bar{X}} + (-)3\frac{S_x}{\sqrt{n}}}$$

Range chart

$$UCL(LCL) = \overline{R} + (-)3\sigma_R = \overline{R} + (-)3S_R$$
$$= \overline{R} + 3S_R(\max\left(0, \overline{R} - 3S\right))$$

$\bar{X}$ -R Control Chart



$$Y = \overline{X} = \sum_{i=1}^{n} X_i / n$$

n : sample size

$$CL = \overline{Y} = \overline{\overline{X}} \text{ (or } T\text{)}$$

$$UCL = \overline{\overline{X}} \text{ (or } T\text{)} + 3S_{\overline{X}}$$

$$LCL = \overline{\overline{X}} \text{ (or } T\text{)} - 3S_{\overline{X}}$$

$$Y = R = X_{\max} - X_{\min}$$

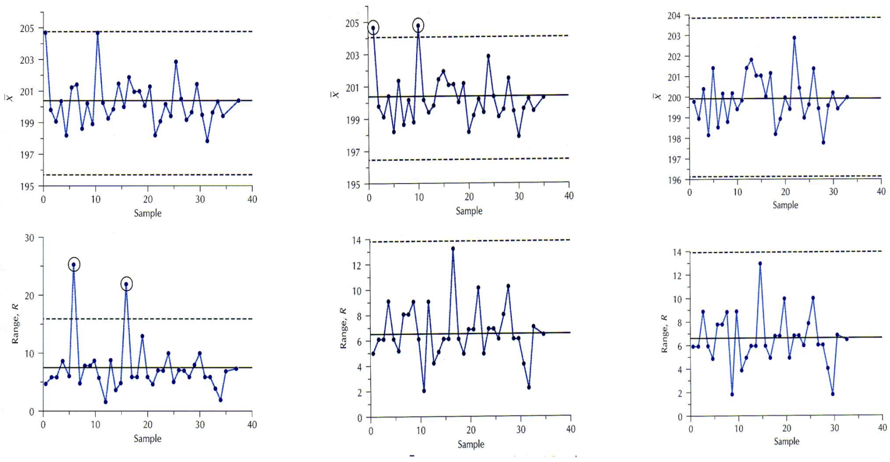
$$CL = \overline{Y} = \overline{R}$$

$$UCL = \overline{R} + 3S_R$$

$$LCL = \max\left[0, \overline{R} - 3S_R\right]$$

Constructing $\bar{X}$ and R Control Charts – making sure data used are under only common causes

initial 1st revision 2nd revision



Two Types of Errors and Chart Performance

- Two errors of $\bar{X}$ hypothesis testing:
  - type I error?
    chart signals but the process is still in control
  - type II error?
    chart does not signal but the process is actually out of control
- Calculating the probabilities of type I and type II errors (α and β)
  - assumptions: normal and independence

Example: $\bar{X}$-R control chart

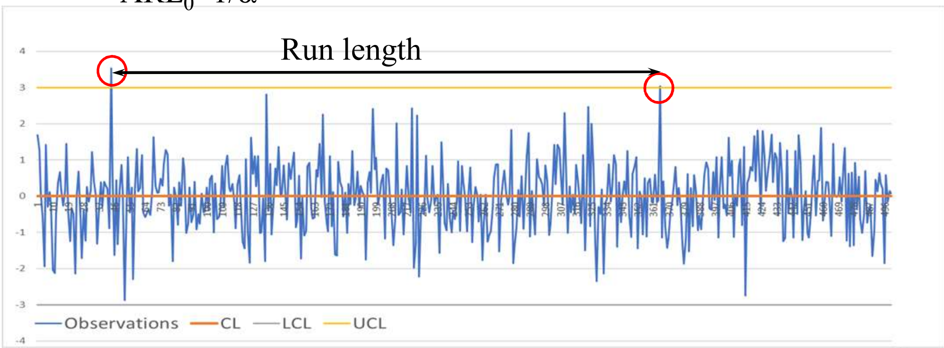
Two Types of Errors and Average Run Length

- Average Run Length (ARL): The average time length a control chart takes to give an alarm.
- ARL_0: The average time length a control chart takes to give an false alarm when the process is actually in-control. (measuring Type I error probability, α)
- ARL_1: The average time length a control chart takes to give an out-of-control alarm given the process is indeed out-of-control (measuring Type II error probability, β)

©Argon Chen

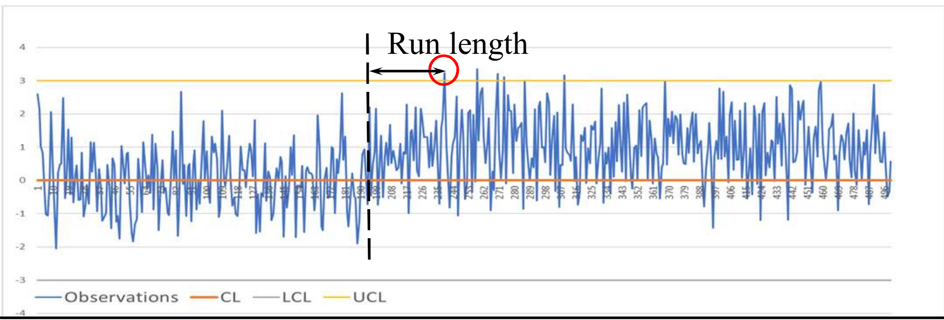
Average Run Length (ARL_0) under Control

- RL_0 : number of runs between false alarms (a random variable)
- Average run length (ARL_0)
  – Average of RL_0 (Geometric Distribution with $p=\alpha$)
  – Higher $ARL_0 \Rightarrow$ higher robustness
  – $ARL_0=1/\alpha$

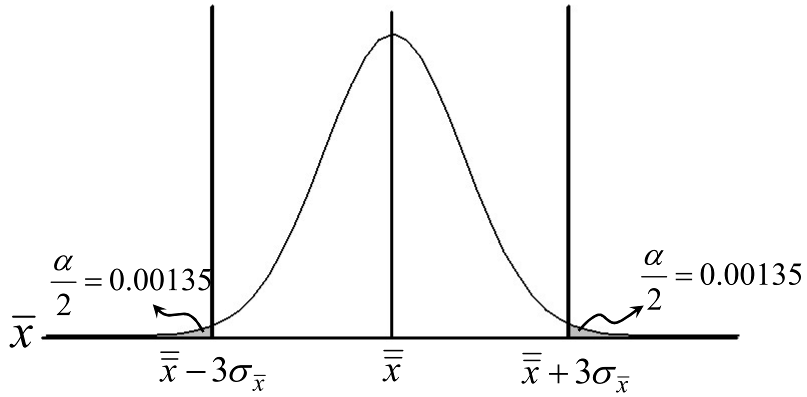


Average Run Length (ARL_1) under Out-of-Control

- RL_1 : the number of runs it takes to detect a shift (a random variable)
- Average run length (ARL_1)
  – Average of RL_1(Geometric Distribution with p=1-β)
  – Lower $ARL_1 \Rightarrow$ higher sensitivity
  – $ARL_1=1/(1-\beta)$ for Shewhart control chart

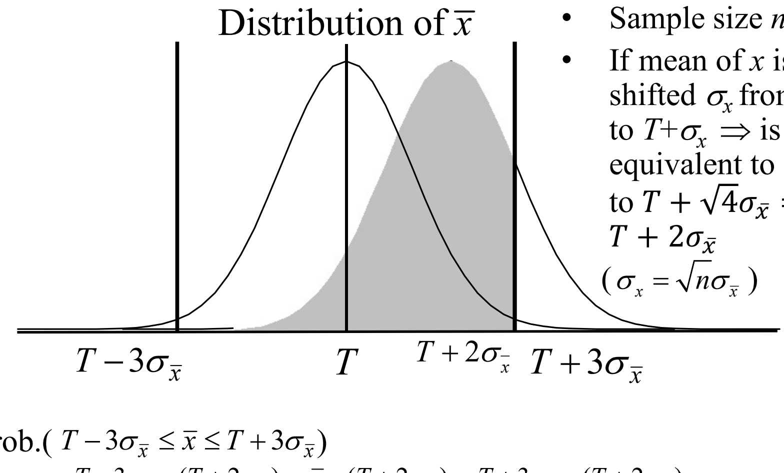


Type I Error Prob. of $\bar{X}$ Chart



$$\alpha = 0.0027 \Rightarrow ARL_0 = \frac{1}{\alpha} = 370.37$$

Type II Error Prob. of $\bar{X}$ Chart



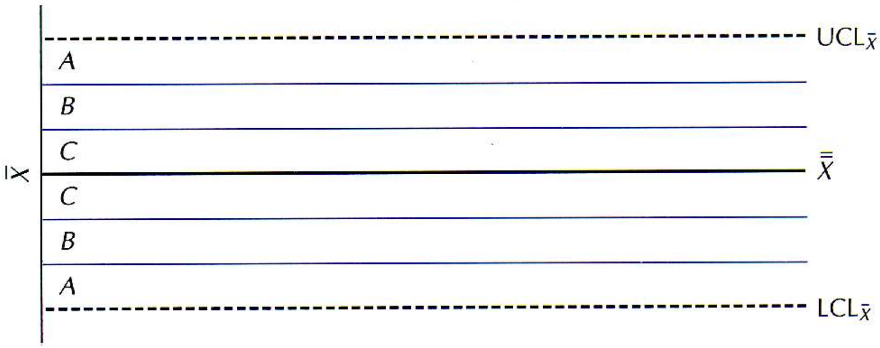
- Sample size n=4
- If mean of x is shifted σ_x from T to $T+\sigma_x$ $\Rightarrow$ is equivalent to shift to $T + \sqrt{4}\sigma_{\bar{x}} = T + 2\sigma_{\bar{x}}$ ($\sigma_x = \sqrt{n}\sigma_{\bar{x}}$)

β = Prob.($T - 3\sigma_{\bar{x}} \le \bar{x} \le T + 3\sigma_{\bar{x}}$)

= Prob.($\frac{T - 3\sigma_{\bar{x}} - (T + 2\sigma_{\bar{x}})}{\sigma_{\bar{x}}} \le \frac{\bar{x} - (T + 2\sigma_{\bar{x}})}{\sigma_{\bar{x}}} \le \frac{T + 3\sigma_{\bar{x}} - (T + 2\sigma_{\bar{x}})}{\sigma_{\bar{x}}}$)

=Prob. ($-5 \le Z \le 1$)=? $\Rightarrow$ $ARL_1 = 1/(1-\beta)$

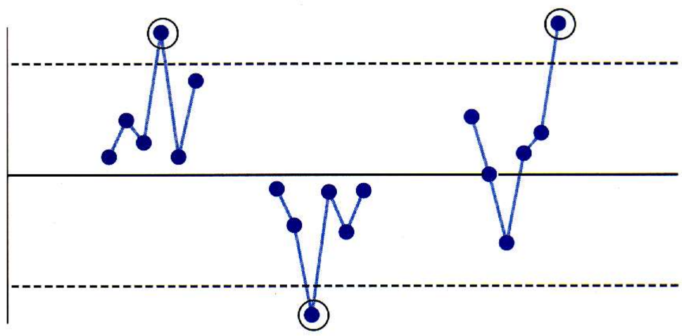
Control Chart Zones to Aid Control Chart Rules



- Western Electric Run Rules
- apply to $\bar{X}$ and R control charts

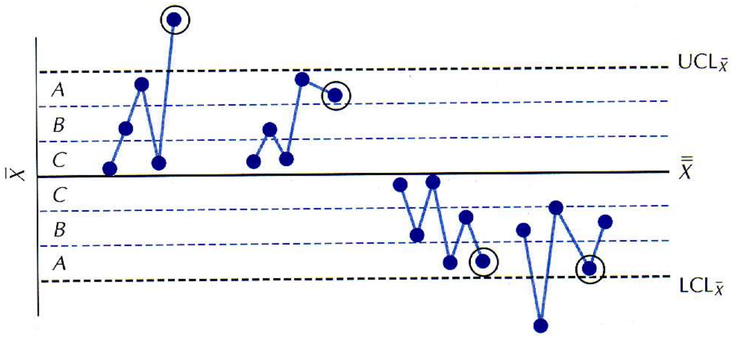
Test 1: Extreme Points

- one point beyond the control limit
- apply to $\bar{X}$ and R control charts
- Western Electric Rule 1



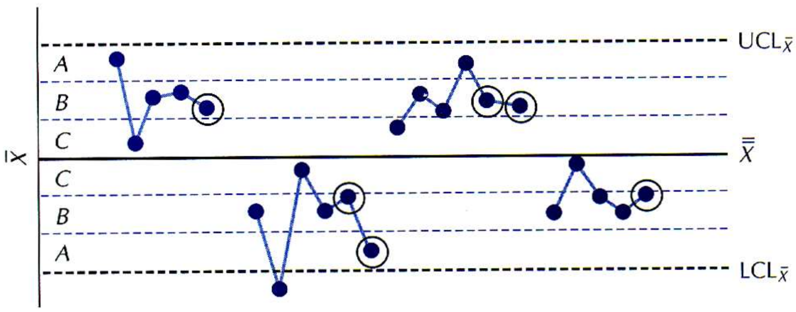
Test 2: Two-out-of-Three

- two out of three points in zone A or beyond
- apply to $\bar{X}$ control chart (not to R since distribution of R is more likely asymmetric)
- Western Electric Rule 2



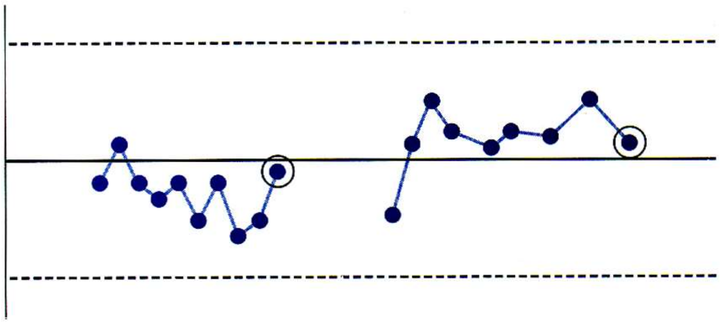
Test 3: Four-out-of-Five

- four out of five points in zone B or beyond
- apply to $\bar{X}$ control chart (not to R since distribution of R is more likely asymmetric)
- Western Electric Rule 3



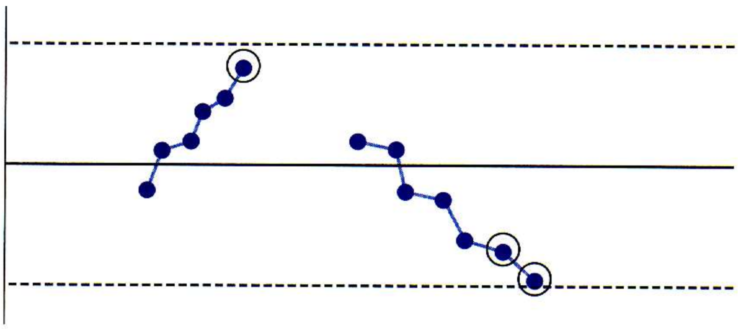
Test 4: Runs above or below the Centerline

- eight points above or below the centerline
- apply to $\bar{X}$ and R control charts
- Western Electric Rule 1



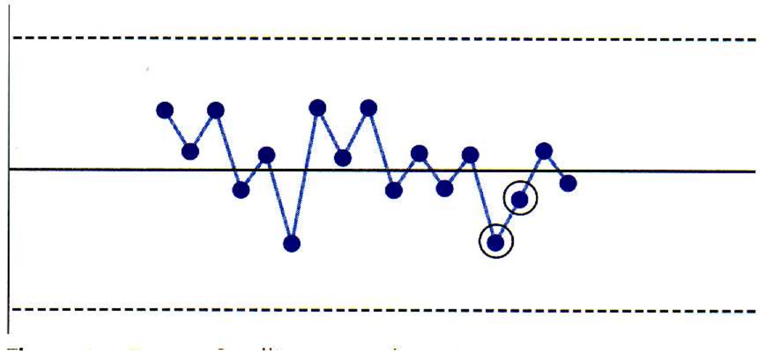
Test 5: Linear Trend Identification

- six points show a continuing increase or decrease
- apply to $\bar{X}$ and R control charts



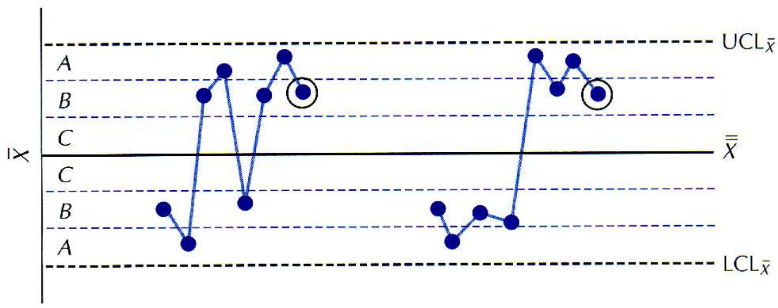
Test 6: Oscillatory Trend Identification

- 14 points oscillate up and down
- apply to $\bar{X}$ and R control charts



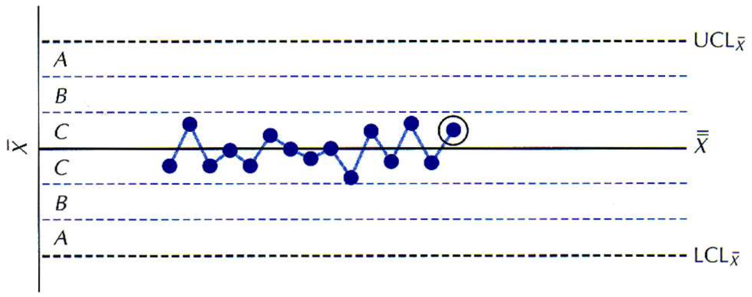
Test 7: Avoidance of Zone C

- eight points avoid zone C
- apply to $\bar{X}$ control chart



Test 8: Run in Zone C

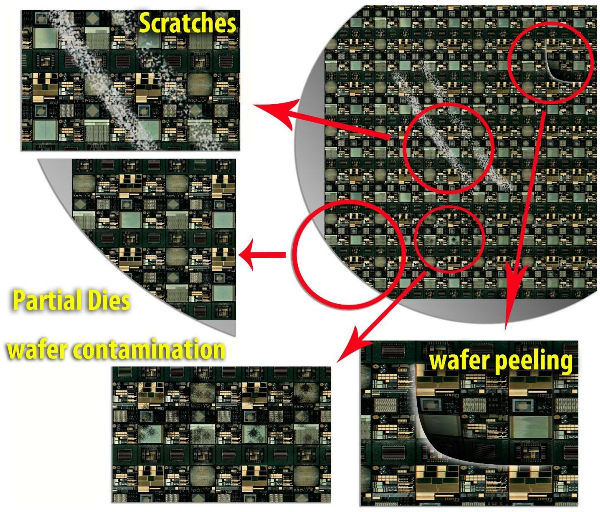
- 15 points fall in zone C only
- apply to $\bar{X}$ control chart



Notice: these tests do improve the sensitivity of the charts, but increase α-risk as well.

Control Chart for Attribute Variable

- Example: Defects on semiconductor wafer surface



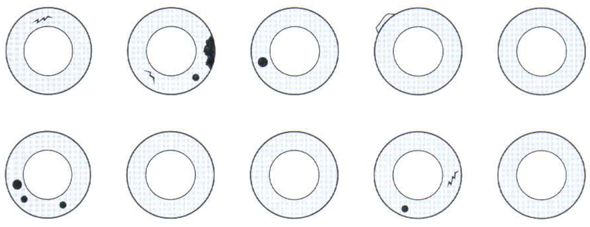
Source: https://www.rsipvision.com/automated-optical-inspection/

Shewhart Attribute Control Chart

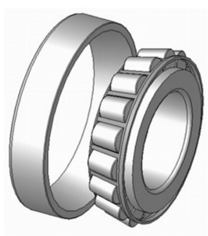
- *p* chart (fraction defective item)
- *np* (*d*) chart (number of defective item)
- *c* chart (number of defects)
- *u* chart (number of defects per unit)

Attribute Quality Characterization

- Example: defects of bearing cast



3 Cracks	Number of defects = 11
6 Holes	Number of defectives = 6
1 Flash	Fraction defective = $\frac{6}{10}$ = 0.6
1 Gate breakout	Number of defects/unit = $\frac{11}{10}$ = 1.1



Ideas of Number of Defectives *np* (*d*) Chart

- Number of Defectives: In a sample of *n* items, *d* is the number of bad items
- Assuming Binomial Distribution Model

$$P(d) = \binom{n}{d} p^d (1-p)^{n-d}$$

- Test statistic: *d*
- Control chart scheme: $E(d) \pm 3\sqrt{Var(d)}$
- Based on binomial model:

$$E(d) = np$$
$$Var(d) = np(1-p)$$

Number of Defectives (*np*) Chart

$$n\hat{p} \pm 3\sqrt{n\hat{p}(1-\hat{p})} \quad \text{where} \quad \hat{p} = \overline{p} = \frac{\sum_{i=1}^{k} d_i}{\sum_{i=1}^{k} n_i} \quad \text{(why not } \frac{\sum_{i=1}^{k} \frac{d_i}{n_i}}{k}\text{ ?)}$$

Average weighted by sample size

$$= \frac{n_1}{\sum_{i=1}^{k} n_i} \bullet \frac{d_1}{n_1} + \frac{n_2}{\sum_{i=1}^{k} n_i} \bullet \frac{d_2}{n_2} + \cdots + \frac{n_k}{\sum_{i=1}^{k} n_i} \bullet \frac{d_k}{n_k} = \frac{\sum_{i=1}^{k} d_i}{\sum_{i=1}^{k} n_i}$$

- n=100
- k=30

Sample	Number Defective	Fraction Defective
1	7	0.07
2	8	0.08
3	6	0.06
4	8	0.08
5	6	0.06
6	8	0.08
7	3	0.03
8	5	0.05
9	9	0.09
10	7	0.07
11	7	0.07
12	9	0.09
13	8	0.08
14	7	0.07
15	8	0.08
16	10	0.10
17	10	0.10
18	5	0.05
19	12	0.12
20	11	0.11
21	8	0.08
22	10	0.10
23	4	0.04
24	10	0.10
25	7	0.07
26	7	0.07
27	9	0.09
28	8	0.08
29	10	0.10
30	10	0.10
	237	

Number of Defective (np) Chart

$$\bar{p} = \frac{237}{3000} = 0.079$$

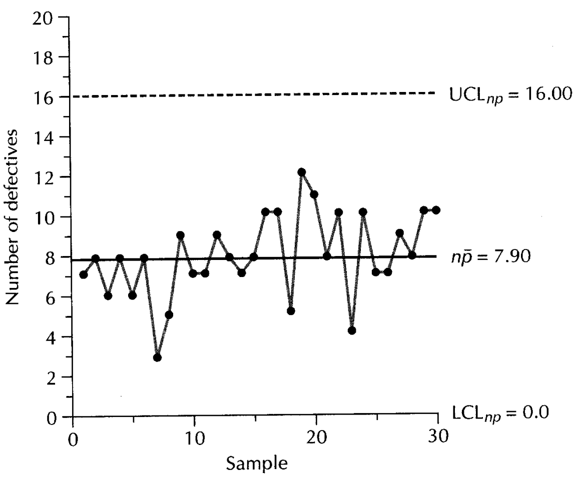
$$n\bar{p} = 100(0.079) = 7.9$$

Control limits:

$$n\bar{p} \pm 3\sqrt{n\bar{p}(1-\bar{p})}$$

$$7.9 \pm 3\sqrt{7.9(1-0.079)}$$

$$UCL_{np} = 16.00$$



$LCL_{np} = -0.20$, so a lower limit of 0.00 is used

Ideas of Fraction Defectives (p) Chart

- Fraction Defectives: In a sample of n items, d is the number of bad items
- Assuming Binomial Distribution Model

$$P(d) = \binom{n}{d} p^d (1-p)^{n-d}$$

- Test statistic: d/n
- Control chart scheme: $E(d/n) \pm 3\sqrt{Var(d/n)}$
- Based on binomial model:

$$E(d/n) = E(d)/n = np/n = p$$

$$Var(d/n) = Var(d)/n^2 = [np(1-p)]/n^2 = [p(1-p)]/n$$

Fraction Defectives (p) Chart

$$\hat{p} \pm 3\sqrt{[\hat{p}(1-\hat{p})]/n}$$

where $$\hat{p} = \overline{p} = \frac{\sum_{i=1}^{k} d_i}{\sum_{i=1}^{k} n_i}$$

p Chart Example

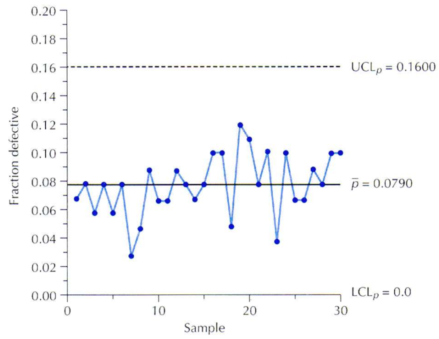
Sample	Number Defective	Fraction Defective
1	7	0.07
2	8	0.08
3	6	0.06
4	8	0.08
5	6	0.06
6	8	0.08
7	3	0.03
8	5	0.05
9	9	0.09
10	7	0.07
11	7	0.07
12	9	0.09
13	8	0.08
14	7	0.07
15	8	0.08
16	10	0.10
17	10	0.10
18	5	0.05
19	12	0.12
20	11	0.11
21	8	0.08
22	10	0.10
23	4	0.04
24	10	0.10
25	7	0.07
26	7	0.07
27	9	0.09
28	8	0.08
29	10	0.10
30	10	0.10
	237	

$$\bar{p} = \frac{\text{sum of all defectives}}{\text{total number of units}} = \frac{237}{3000} = 0.0790$$

For $\bar{p} = 0.0790$ and $n = 100$,

$$\hat{\sigma}_p = \sqrt{\frac{\bar{p}(1-\bar{p})}{n}} = \sqrt{\frac{(0.0790)(1-0.0790)}{100}} = 0.0270$$

$$\begin{aligned} \text{UCL} &= \bar{p} + 3\hat{\sigma}_p \\ &= 0.0790 + (3)(0.0270) \\ &= 0.1600 \\ \text{LCL} &= \bar{p} - 3\hat{\sigma}_p \\ &= 0.0790 - (3)(0.0270) \\ &= -0.002 \\ &= 0.0. \end{aligned}$$



Variable-Sample-Size p Chart

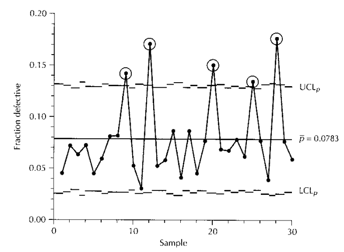
- compute separate limits for each subgroup
- use an average sample size
- use standarized p values

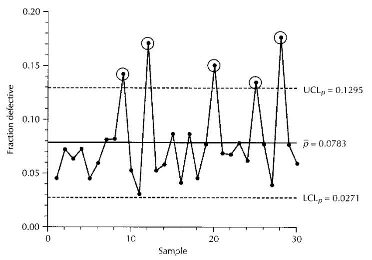
$$Z_i = \frac{\frac{d_i}{n_i} - \bar{p}}{\sqrt{\bar{p}(1-\bar{p})/n_i}}$$

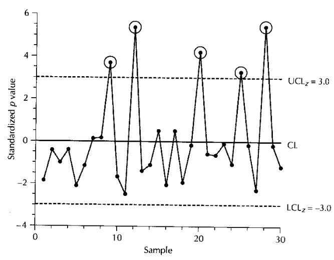
Variable-Sample-Size p Chart Example

Sample	n	d	p	LCL_p	UCL_p	z
1	238	11	0.046	0.026	0.131	−1.84
2	245	18	0.073	0.027	0.130	−0.28
3	270	17	0.063	0.029	0.127	−0.94
4	207	15	0.072	0.022	0.134	−0.31
5	251	11	0.044	0.027	0.129	−2.03
6	254	15	0.059	0.028	0.129	−1.14
7	236	19	0.081	0.026	0.131	0.13
8	245	20	0.082	0.027	0.130	0.19
9	246	35	0.142	0.027	0.130	3.74
10	269	14	0.052	0.029	0.127	−1.60
11	223	7	0.031	0.024	0.132	−2.61
12	246	42	0.171	0.027	0.130	5.40
13	262	14	0.053	0.029	0.128	−1.50
14	258	15	0.058	0.028	0.128	−1.21
15	232	20	0.086	0.025	0.131	0.45
16	219	9	0.041	0.024	0.133	−2.05
17	263	23	0.087	0.029	0.128	0.55
18	244	11	0.045	0.027	0.130	−1.93
19	274	21	0.077	0.030	0.127	−0.10
20	245	37	0.151	0.027	0.130	4.24
21	233	16	0.069	0.026	0.131	−0.55
22	267	18	0.067	0.029	0.128	−0.66
23	254	20	0.079	0.028	0.129	0.03
24	264	16	0.061	0.029	0.128	−1.07
25	253	34	0.134	0.028	0.129	3.32
26	290	22	0.076	0.031	0.126	−0.15
27	231	9	0.039	0.025	0.131	−2.23
28	227	40	0.176	0.025	0.132	5.49
29	234	18	0.077	0.026	0.131	−0.08
30	253	15	0.059	0.028	0.129	−1.13

Variable-Sample-Size p Chart Example







Ideas of Number of Defect Chart

- Assuming Poisson Distribution Model (why?)

$$P(\text{defect number} = c; \lambda) = \frac{e^{-\lambda}\lambda^{c}}{c!}$$

- Sample statistic: c
- Control chart scheme: $E(c) \pm 3\sqrt{Var(c)}$
- Base on Poisson model:

$$E(c) = Var(c) = \lambda$$

Number of Defect (c) Chart

$$\hat{\lambda} \pm 3\sqrt{\hat{\lambda}}$$

$$\text{where } \hat{\lambda} = \bar{c} = \frac{\sum_{i=1}^{k} c_i}{k}$$

Example of c Chart

Sample	Number of Defects
1	16
2	14
3	28
4	16
5	12
6	20
7	10
8	12
9	30
10	17
11	9
12	17
13	14
14	16
15	15
16	13
17	14
18	16
19	11
20	20
21	11
22	9
23	16
24	31
25	13

Centerline:

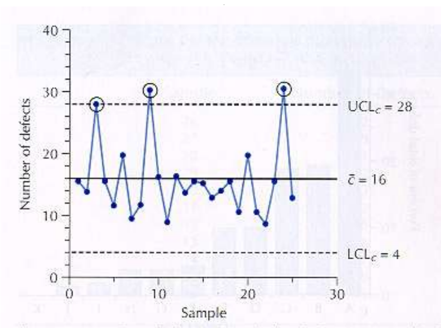
$$\bar{c} = \frac{\text{total number of defects}}{\text{total number of samples}} = \frac{400}{25} = 16.$$

Control limits:

$$\text{UCL}_c, \text{LCL}_c = \bar{c} \pm 3\sqrt{\bar{c}} = 16 \pm 3\sqrt{16}$$

$$\text{UCL}_c = 28$$

$$\text{LCL}_c = 4.$$



Ideas of Number of Defects per Unit Chart

- Assuming Poisson Distribution Model again

$$P(\text{defect number/unit } b\,;\lambda) = \frac{e^{-\lambda}\lambda^{b}}{b!}$$

- Sample statistic: $u = \sum_{i=1}^{n} b_i / n = \frac{c}{n} = \bar{b}$
- Control chart scheme: $E(\bar{b}) \pm 3\sqrt{Var(\bar{b})}$
- Base on Poisson model:

$$E(\bar{b}) = E(b) = \lambda \qquad Var(\bar{b}) = Var(b)/n = \lambda / n$$

Number of Defects per Unit (u) Chart

- Variable limits due to variable sample sizes

$$\hat{\lambda} \pm 3\sqrt{\hat{\lambda}/n}$$

$$\text{where } \hat{\lambda} = \frac{\sum_{i=1}^{k} c_i}{\sum_{i=1}^{k} n_i} \quad k: \text{number of samples (lots)}$$

Number of Defect per Unit (u) Chart Example

Sample	Sample Size, n	Number of Defects per Sample, c	Average Number of Defects per Unit, u	LCL_u	UCL_u
1	16	23	1.44	0.49	2.25
2	20	30	1.50	0.59	2.16
3	26	35	1.35	0.68	2.06
4	8	12	1.50	0.13	2.61
5	22	29	1.32	0.62	2.12
6	29	35	1.21	0.72	2.02
7	31	50	1.61	0.74	2.00
8	13	15	1.15	0.40	2.35
9	28	36	1.29	0.71	2.04
10	23	38	1.65	0.64	2.10
11	19	24	1.26	0.57	2.18
12	23	32	1.39	0.64	2.10
13	14	24	1.71	0.43	2.31
14	29	34	1.17	0.72	0.72
15	27	38	1.41	0.70	2.05
16	15	25	1.67	0.46	2.28
17	22	26	1.18	0.62	2.12
18	22	24	1.09	0.62	2.12
19	14	22	1.57	0.43	2.31
20	16	17	1.06	0.49	2.25
21	22	33	1.50	0.62	2.12
22	16	21	1.31	0.49	2.25
23	14	18	1.29	0.43	2.31
24	5	9	1.80	0.00	2.94
25	13	18	1.38	0.40	2.35
26	19	26	1.37	0.57	2.18
27	10	12	1.20	0.26	2.48

An example calculation for control limits follows.

$$\text{control limits} = \bar{u} \pm 3\sqrt{\frac{\bar{u}}{n}}$$

$$= 1.37 \pm 3\sqrt{\frac{1.37}{n}}.$$

For sample 1, $n_1 = 16$, and therefore,

$LCL_{u_1} = 0.49$

$UCL_{u_1} = 2.25.$