Why the number 3

Suppose there is a shift in the mean such that

$\mu = \mu_0 + r\sigma$

Let $\beta = P(LCL \le \bar{X} \le UCL \mid \mu = \mu_0 + r\sigma)$

$LCL = \mu_0 - k\frac{\sigma}{\sqrt{n}}$, $UCL = \mu_0 + k\frac{\sigma}{\sqrt{n}}$

This gives $\beta = P(\bar{X} \le \mu_0 + k\frac{\sigma}{\sqrt{n}} \mid \mu = \mu_0 + r\sigma) - P(\bar{X} \le \mu_0 - \frac{k\sigma}{\sqrt{n}} \mid \mu = \mu_0 + r\sigma)$

$= P(\bar{X} - (\mu_0 + r\sigma) \le \mu_0 - \mu_0 - r\sigma + \frac{k\sigma}{\sqrt{n}} \mid \mu = \mu_0 + r\sigma) - P(\bar{X} - (\mu_0 + r\sigma) \le \mu_0 - \mu_0 - r\sigma - k\frac{\sigma}{\sqrt{n}} \mid \mu = \mu_0 + r\sigma)$

$= P\left(\frac{\bar{X} - (\mu_0 + r\sigma)}{\frac{\sigma}{\sqrt{n}}} \le k - r\sqrt{n} \mid \mu = \mu_0 + r\sigma \right) - P\left(\frac{\bar{X} - (\mu_0 + r\sigma)}{\frac{\sigma}{\sqrt{n}}} \le -k - r\sqrt{n} \mid \mu = \mu_0 + r\sigma \right)$

$= \Phi(k - r\sqrt{n}) - \Phi(-k - r\sqrt{n})$

β is increasing with k and decreases when n and r increases.

β = type II error = P(not detecting a shift / given that there is one)

Let T be the time from the shift until the shift is detected

T is geometric distributed with probability $p = 1 - \beta$

such that $E[T] = \frac{1}{1-\beta}$

β plotted against r for varies values n is called and operating characteristic (OC) curve.

How to sample

A sample is representing a subgroup and is often called a rational subgroup.

To detect a shift in location - Advantage with small variation within the subgroups and the sample units should be taken to be close together in time.

To control variation - An advantage to have as large within sample variability as possible.

Control charts on variability should be established before control charts on the mean.

Construction of Control Charts (parameters unknown)

$\bar{X}$ - chart.

From $k \geq 20$ samples, each with averages

$\bar{X}_i = \frac{1}{n} \sum_{j=1}^{n} X_{ji}$, $i = 1, 2, \ldots, k$ we calculate the overall average $\bar{\bar{X}} = \frac{1}{k} \sum_{i=1}^{k} \bar{X}_i$ to obtain the centerline. n is normally 4, 5 or 6.

From historical reasons, to estimate the standard deviations we often use the range $R_i = X_{max, i} - X_{min, i}$ $i = 1, 2, \ldots, k$ and the overall range is estimated by $\bar{R} = \frac{1}{k} \sum_{i=1}^{k} R_i$

An estimate for σ is obtained by $\hat{\sigma} = \frac{\bar{R}}{d_2(n)}$ where $d_2(n)$ depends on the sample size. Values can be found in Table A.25.

Thereby the $\bar{x}$ chart is constructed with.

$$UCL = \bar{\bar{x}} + \frac{3\bar{R}}{d_2\sqrt{n}}, \quad \text{centerline} = \bar{\bar{x}}, \quad LCL = \bar{\bar{x}} - \frac{3\bar{R}}{d_2\sqrt{n}}$$

With $A_2 = \frac{3}{d_2\sqrt{n}}$ this can be written as

$UCL = \bar{\bar{x}} + A_2\bar{R}$ and $LCL = \bar{\bar{x}} - A_2\bar{R}$. A_2 is ~~tabulat~~ tabulated i

Table A.22

Control of variation

The relative range is defined as $W = \frac{R}{\sigma}$.

$$E[W] = \frac{E[R]}{\sigma} = d_2(n)$$

$SD[W] = d_3(n)$. Hence $SD(R) = \sigma d_3$. Replacing σ by $\hat{\sigma} = \frac{\bar{R}}{d_2}$ we have: $\hat{\sigma}_R = \frac{\bar{R}d_3}{d_2}$

The R-chart becomes

$$UCL = \bar{R} + 3\bar{R}\frac{d_3}{d_2}, \quad \text{centerline } \bar{R}, \quad LCL = \bar{R} \mp 3\bar{R}\frac{d_3}{2}$$

or,

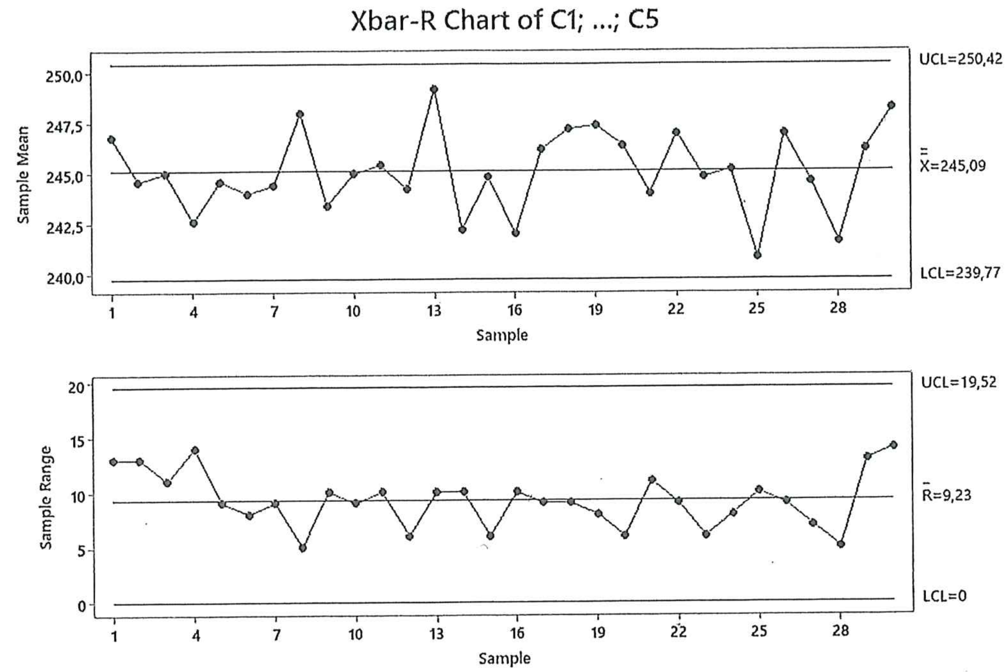
$UCL = \bar{R}D_4$ and $LCL = \bar{R}D_3$ where

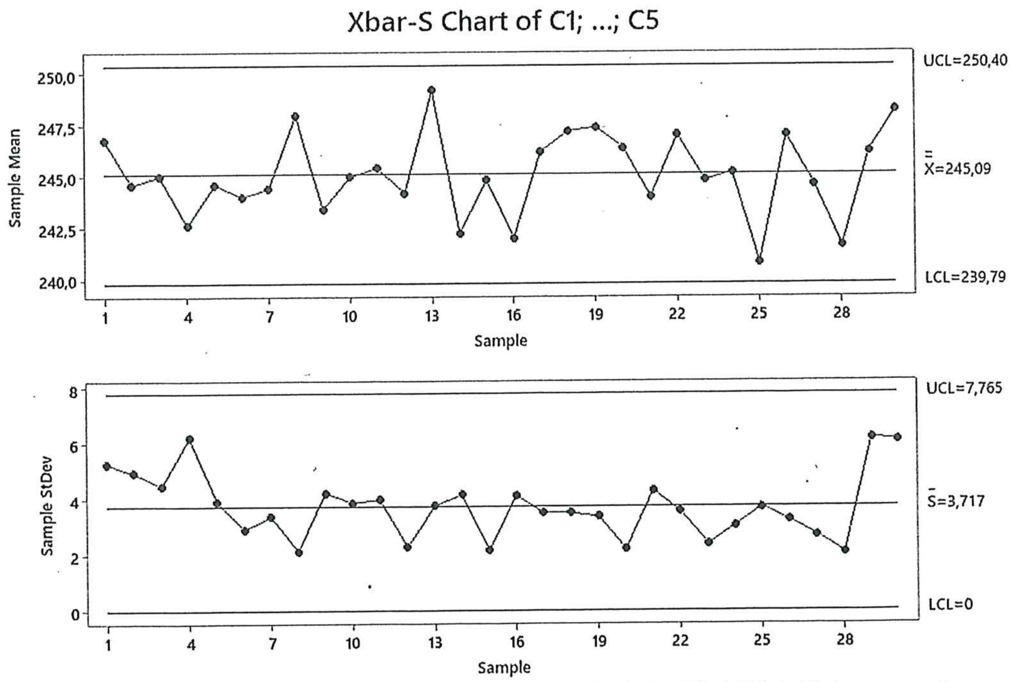
$$D_4 = 1 + 3\frac{d_3}{d_2}, \quad D_3 = 1 - 3\frac{d_3}{d_2}$$ Table A22.

Control Chart for Silicon Wafer Data

Data: We have k=30 and n=5

Row	C1	C2	C3	C4	C5
1	240	243	250	253	248
2	238	242	245	251	247
3	239	242	246	250	248
4	235	237	246	249	246
5	240	241	246	247	249
6	240	243	244	248	245
7	240	243	244	249	246
8	245	250	250	247	248
9	238	240	245	248	246
10	240	242	246	249	248
11	240	243	246	250	248
12	241	245	243	247	245
13	247	245	255	250	249
14	237	239	243	247	245
15	242	244	245	248	245
16	237	239	242	247	245
17	242	244	246	251	248
18	243	245	247	252	249
19	243	245	248	251	250
20	244	246	246	250	246
21	241	239	244	250	246
22	242	245	248	251	249
23	242	245	248	243	246
24	241	244	245	249	247
25	236	239	241	246	242
26	243	246	247	252	247
27	241	243	245	248	246
28	239	240	242	243	244
29	239	240	250	252	250
30	241	243	249	255	253


Xbar-R Chart of C1; ...; C5
Sample Mean
250,0
247,5
245,0
242,5
240,0
UCL=250,42
$\bar{\bar{X}}$=245,09
LCL=239,77
1
4
7
10
13
16
19
22
25
28
Sample
Sample Range
20
15
10
5
0
UCL=19,52
$\bar{R}$=9,23
LCL=0
Sample


Xbar-S Chart of C1; ...; C5
Sample Mean
250,0
247,5
245,0
242,5
240,0
UCL=250,40
$\bar{\bar{X}}$=245,09
LCL=239,79
1
4
7
10
13
16
19
22
25
28
Sample
Sample StDev
8
6
4
2
0
UCL=7,765
$\bar{S}$=3,717
LCL=0
Sample