MAT2377

Catalin Rada

Version 2009/06/30

Statistical Quality Control

A production process is often subject to *variability*. There are basically two types of variability:

- The process works properly, and the variability is the cumulative effect of many small, essentially unavoidable causes. A process that is operating with only chance causes (or non-assignable causes) is said to be in statistical control.

- The other type of variability is due to assignable causes like improperly adjusted machines, operator errors, defective materials etc. Such variability is typically much larger than in the previous case. In this context we say that the process is out of control.

The aim of statistical quality control is to identify occurrence of assignable causes.

Control charts

— a tool for statistical quality control; a tool to identify out-of-control systems;

We assume that a quality is defined in terms of a random variable X. We suppose that $X \sim \mathcal{N}(\mu, \sigma^2)$. For example, in manufacturing automobile engine piston rings, the inside diameter is 74 millimeters and it is known that the standard deviation of ring diameter is 0.01. Then $X =$ ring diameter.

Terminology: Let μ and σ be the mean and standard deviation of X. In statistical quality control:

— μ is called the **mean quality**;

— σ is called the **variability of the process**.

At particular equally-spaced time points $t_1, \dots, t_N$ (for example, every hour) we take a sample of size n and $\bar{X}_i$ is the sample mean for ith sample (collected at time t_i). If the process is in control than $\bar{X}_i \sim \mathcal{N}(\mu, \sigma^2/n)$ and with the probability $1-\alpha$ the observed sample mean should be between $\mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$ and $\mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$.

We will denote by $\bar{x}_i$ the observed sample mean in the ith sample, $i = 1, \dots, N$.

Thus, $\bar{X}$ control chart is defined by

$$\begin{aligned} UCL &= \mu + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \text{Upper Control Limit} \\ LCL &= \mu - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \text{Lower Control Limit} \\ CL &= \mu \quad \text{Central Line} \end{aligned}$$

For such defined control chart, if we observe $\bar{x}_i > UCL$ or $\bar{x}_i < LCL$ then it indicates ***instability of the process***, and we say that we have observed **a warning signal of instability**. The parameter α is interpreted as

$$\alpha = P(\text{signal of instability when the process is stable}).$$

or

$$\alpha = P(\text{signal of instability when the process is under control}).$$

Typically we consider $z_{\alpha/2} = 3$, and this means that $\alpha = 0.0027$, or 0.27%:

$$UCL = \mu + 3\frac{\sigma}{\sqrt{n}}, \quad LCL = \mu - 3\frac{\sigma}{\sqrt{n}}.$$

This means that if the system is under control, then we will observe a warning signal of instability about every $1/.0027 \cong 370$ samples on average.

When μ and σ unknown

— that is the case in many problems;

When the parameters μ and σ are unknown we estimate them on the basis of preliminary m samples, *taken when the process is in control*. We estimate mean μ by the grand mean:

$$\bar{\bar{X}} = \frac{1}{m} \sum_{i=1}^{m} \bar{X}_i.$$

The grand mean is an unbiased estimator of μ. Then $\bar{\bar{x}}$, the observed grand mean, is taken as the center line of the control chart.

We may estimate σ either from the standard deviation of each sample or from the range of observations within each sample. If

$$\begin{aligned} R_i &= \text{Range of the preliminary sample number } i \\ &= \text{maximum of the } i\text{th sample} - \text{minimum of the } i\text{th sample} \end{aligned}$$

then

$$\mathrm{E}(R_i) = d_2\sigma \qquad \sqrt{Var(R_i)} = d_3\sigma,$$

where d_2 and d_3 are (tabulated) values which depend on the sample size n. Consequently,

$$\hat{\sigma} = \frac{\bar{R}}{d_2},$$

where $\bar{R} = \frac{1}{m}\sum_{i=1}^{m} R_i$. Of course $\hat{\sigma}$ is an unbiased estimator for σ.

Control chart for $\bar{X}$ from $\bar{R}$

If $r_1, \ldots, r_m$ are observed values for R_i and $\bar{r}$ is the observed value of $\bar{R}$, based on preliminary samples, then $\hat{\sigma}$ replaces σ in the control chart for $\bar{X}$:

$$\begin{aligned} CL &= \hat{\mu} = \bar{\bar{x}} \quad \text{Central Line} \\ UCL &= \bar{\bar{x}} + 3\frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} + \frac{3}{d_2\sqrt{n}}\bar{r} = \bar{\bar{x}} + A_2\bar{r} \\ LCL &= \bar{\bar{x}} - 3\frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} - \frac{3}{d_2\sqrt{n}}\bar{r} = \bar{\bar{x}} - A_2\bar{r}. \end{aligned}$$

The values A_2 are tabulated in Table XI.

Control chart for $\bar{X}$ from $\bar{S}$

If S_i is a sample standard deviation of the preliminary ith sample, then

$$\mathrm{E}(S_i) = c_4\sigma \qquad \sqrt{Var(S_i)} = \sigma^2(1 - c_4^2),$$

where c_4 is the (tabulated) value which depends on the sample size n. Consequently,

$$\hat{\sigma} = \frac{\bar{S}}{c_4},$$

where $\bar{S} = \frac{1}{m}\sum_{i=1}^{m} S_i$. Of course $\hat{\sigma}$ is an unbiased estimator for σ, the process variability. The control chart (from $\bar{S}$):

If now $s_1, \ldots, s_m$ are observed sample standard deviations and $\bar{s} = \frac{1}{m}\sum_{i=1}^{m} s_i$, then

$$\begin{aligned} CL &= \hat{\mu} = \bar{\bar{x}} \quad \text{Central Line} \\ UCL &= \bar{\bar{x}} + 3\frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} + \frac{3}{c_4\sqrt{n}}\bar{s} \\ LCL &= \bar{\bar{x}} - 3\frac{\hat{\sigma}}{\sqrt{n}} = \bar{\bar{x}} - \frac{3}{c_4\sqrt{n}}\bar{s}. \end{aligned}$$

Example

We have $m = 5$ preliminary samples of size $n = 3$:

i				$\bar{x}_i$	r_i	s_i
1	27.1	29.4	27.2	27.9	2.3	1.3
2	30.6	32.5	32.4	31.83	1.9	1.07
3	25.7	35.5	30	30.4	9.8	4.91
4	31.1	23.2	25	26.43	7.9	4.14
5	24.1	34.2	27.4	28.57	10.1	5.15
		total		145.13	32	16.57

- Indeed, note that $r_1 = 29.4 - 27.1$,

$$s_1 = \sqrt{\left((27.1 - 27.9)^2 + (29.4 - 27.9)^2 + (27.2 - 27.9)^2\right)/2}.$$

- The estimated grand mean is $145.13/5 = 29.026$; $\bar{r} = 32/5 = 6.4$ and $\bar{s} = 3.314$.

Thus, we estimated parameters based on preliminary samples. We will use them to construct the control charts. The Control chart for $\bar{X}$ from $\bar{R}$ and from $\bar{S}$ are, respectively:

$$
\begin{aligned}
CL &= \hat{\mu} = \bar{\bar{x}} = 29.026 \\
UCL &= \bar{\bar{x}} + A_2\bar{r} = 29.026 + 1.023 \times 6.4 = 35.5732 \\
LCL &= \bar{\bar{x}} - A_2\bar{r} = 22.4788.
\end{aligned}
$$

$$
\begin{aligned}
CL &= \hat{\mu} = \bar{\bar{x}} = 29.026 \\
UCL &= \bar{\bar{x}} + \frac{3}{c_4\sqrt{n}}\bar{s} = 29.026 + \frac{3}{0.8862 \times \sqrt{3}} \times 3.314 = 35.5031 \\
LCL &= \bar{\bar{x}} - \frac{3}{c_4\sqrt{n}}\bar{s} = 29.026 - \frac{3}{0.8862 \times \sqrt{3}} \times 3.314 = 22.5490.
\end{aligned}
$$

Control Charts for Variability

Before, we learned how to estimate σ using R_i's (ranges) or S_i's (sample standard deviations) from preliminary samples. Those estimates where used to construct the control charts for the mean.

Now, we want to construct control charts for σ.

First method

If $r_1, \ldots, r_m$ are observed ranges from the preliminary samples, then **there is a signal of instability of the process if** $r_i < LCL$ or $r_i > UCL$, where

$$LCL = \mu_R - 3\sigma_R, \quad UCL = \mu_R + 3\sigma_R,$$

where $\mu_R := \mathrm{E}(R_i) = d_2\sigma$, $\sigma_R := \sqrt{\mathrm{Var}(R_i)} = d_3\sigma$. Of course, this formula makes sense if σ^2, the variance of the process is known. If this is not the case, then we have to estimate μ_R and σ_R from m preliminary samples of size n. We take:

$$\hat{\mu}_R = \frac{1}{m}\sum_{i=1}^{m} R_i = \bar{R}, \qquad \hat{\sigma}_R = \frac{d_3}{d_2}\bar{R}.$$

Control Chart for R

$$\begin{aligned} CL &= \hat{\mu}_R = \bar{r} \\ UCL &= \hat{\mu}_R + 3\hat{\sigma}_R = \bar{r}\left(1 + 3\frac{d_3}{d_2}\right) =: D_4\bar{r} \\ LCL &= \hat{\mu}_R - 3\hat{\sigma}_R = \bar{r}\left(1 - 3\frac{d_3}{d_2}\right) =: D_3\bar{r} \end{aligned}$$

Note: If $\left(1 - 3\frac{d_3}{d_2}\right) < 0$ then we set $D_3 \equiv 0$ since a variability cannot be negative.

Second method

If $s_1, \ldots, s_m$ are observed ranges from the preliminary) samples, then there is a signal of instability of the process if $s_i < LCL$ or $s_i > UCL$, where

$$LCL = \mu_S - 3\sigma_S, \quad UCL = \mu_S + 3\sigma_S,$$

where $\mu_S := \mathrm{E}(S_i) = c_4\sigma$, $\sigma_S := \sqrt{\mathrm{Var}(S_i)} = \sigma\sqrt{1 - c_4^2}$. Of course, this formula makes sense if σ^2, the variance of the process is known. If this is not the case, then we have to estimate σ from m preliminary samples of size n. We take:

$$\hat{\sigma} = \frac{\bar{S}}{c_4}$$

Control Chart for S

$$
\begin{aligned}
CL &= \hat{\mu}_S = \bar{s} \\
UCL &= \hat{\mu}_S + 3\hat{\sigma}_S = \bar{s}\left(1 + \frac{3}{c_4}\sqrt{1 - c_4^2}\right) \\
LCL &= \hat{\mu}_S - 3\hat{\sigma}_S = \bar{s}\left(1 - \frac{3}{c_4}\sqrt{1 - c_4^2}\right)
\end{aligned}
$$

Note: If $\left(1 - \frac{3}{c_4}\sqrt{1 - c_4^2}\right) < 0$ then we set $LCL \equiv 0$ since a variability cannot be negative.

Example

Refer to the data from the previous example. Control Chart for R:

$$UCL = \hat{\mu}_R + 3\hat{\sigma}_R = \bar{r}\left(1 + 3\frac{d_3}{d_2}\right) =: 6.4 \times D_4 = 6.4 \times 2.574$$

$$LCL = \hat{\mu}_R - 3\hat{\sigma}_R = \bar{r}\left(1 - 3\frac{d_3}{d_2}\right) =: 6.4 \times D_3 = 0$$

Control Chart for S:

$$UCL = \bar{s}\left(1 + \frac{3}{c_4}\sqrt{1 - c_4^2}\right) = 3.314\left(1 + \frac{3}{0.8862}\sqrt{1 - 0.8862^2}\right) = 8.5117$$

$$LCL = \bar{s}\left(1 - \frac{3}{c_4}\sqrt{1 - c_4^2}\right) = 3.314\left(1 - \frac{3}{0.8862}\sqrt{1 - 0.8862^2}\right) = -1.88 \Rightarrow 0$$

Is the process under control?

— Plot control charts using preliminary data.

— If the data are within control limits, then the process is under control.

— Otherwise, if we have observations which are outside the limits, then $remove\ them$ and compute the revised control limits.

Do exc 16-1/page 662 $n = 5$, $\overline{x}$ and r are computed for each of 35 preliminary samples; the summary is: $\sum_{i=1}^{35} \overline{x_i} = 7805$; $\sum_{i=1}^{35} r_i = 1200$.

a) find trial control limits for $\overline{X}$ chart from R;

b) assume the process is under control. Estimate the process mean and standard deviation.

Sol: a) We have $\overline{\overline{x}} = \frac{7805}{35} = 223$ and $\overline{r} = \frac{1200}{35} = 34.286$. Hence the chart (from R) for $\overline{X}$ is:

$CL = \overline{\overline{x}} = 223$; $LCL = 223 - A_2\overline{r} = 223 - 0.577 \times 34.286 = 203.22$ and $LCL = 223 + A_2\overline{r} = 223 + 0.577 \times 34.286 = 242.78$;

b) Recall page 7 (and the table) and compute: $\hat{\sigma} = \frac{\overline{R}}{d_2} = \frac{34.286}{2.326} = 14.74$.

Do exc 16-2/page 662 The statement is telling us: 25 samples of size 5; $\sum_{i=1}^{25} \overline{x_i} = 362.75$; $\sum_{i=1}^{25} r_i = 8.60$, and $\sum_{i=1}^{25} s_i = 3.64$.

a) find the control chart for $\overline{X}$ from R;

b) find the control chart for $\overline{X}$ from S;

Sol: a) We compute $\overline{\overline{x}} = \frac{362.75}{25} = 14.51$ and $\overline{r} = \frac{8.60}{25} = 0.344$. Hence the chart (from R) for $\overline{X}$ is:

$CL = 14.51$, $LCL = CL - A_2\overline{r} = 14.51 - 0.577 \times 0.344 = 14.312$, and $LCL = CL + A_2\overline{r} = 14.51 + 0.577 \times 0.344 = 14.708$;

b) By the table on page 732 one gets $c_4 = 0.94$. Moreover one has that $\overline{s} = \frac{3.64}{25} = 0.1456$. Hence the chart (from S) for $\overline{X}$ is (see page 10):

$CL = 14.51$, $LCL = CL - \frac{3\overline{s}}{c_4\sqrt{n}} = 14.51 - \frac{3\times 0.1456}{0.94\times 5} = 14.417$, and $UCL = CL + \frac{3\overline{s}}{c_4\sqrt{n}} = 14.51 + \frac{3\times 0.1456}{0.94\times 5} = 14.603$.

READ 16-1 to 16-4 FOR MORE EXAMPLES AND MORE COMPUTATIONS!

Control Chart for Proportion

If the number of defective units is D and the sample size is n, then the fraction of defective is $P = D/n$. Let $p_1, \ldots, p_m$ be the observed proportions in m samples. Control Chart is

$$CL = \bar{p} = \frac{1}{m}\sum_{i=1}^{m} p_i,$$

$$LCL = \bar{p} - 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}, \qquad UCL = \bar{p} + 3\sqrt{\frac{\bar{p}(1-\bar{p})}{n}}.$$

If $LCL < 0$, then we set $LCL = 0$. If $UCL > 1$, then we set $UCL = 1$.