Chapter 2

Control charts

In this chapter we present the mathematical background of techniques to detect deviations in a production process that would lead to non-conforming items. Such deviations may lead to reduction of production costs or rework costs. The key factor to success is to accurately assess variability in production processes. If we accurately know the variability of a production process that is in control, then we are able to detect observations that indicate that the process has gone out of control. These procedures are known under the name control charts. The simplest and most widely used control chart is the Shewhart $\overline{X}$-chart, which should be used together with an R-chart, an S-chart or an S^2-chart. Shewhart introduced these charts in 1924 as simple tools to be used by workers in production lines. In spite of their simplicity, these charts turned out to be highly effective in practice. When properly set up, these charts quickly detect large changes in process means. In the 1950's CUSUM (Cumulative Sum) control charts were introduced to detect small process changes. Although they are mathematically optimal in a certain sense, it is hard to set them up correctly since they are very sensitive to small changes in the parameters. A convenient alternative is the EWMA chart (Exponentially Weighted Moving Average) which has its roots in time series analysis. The EWMA chart is easy to implement and not very sensitive to both parameter changes and non-normality of data, while it performs almost as good as a CUSUM chart for detecting small process changes.

It is important to note that there are two different uses of control charts in general. In retrospective or Phase I use (sometimes called Initial Study), observations are analyzed after they have all been collected. Usually this is done during a pilot study of a production process, when one needs to estimate in-control process behaviour. This is input to a capability analysis. The other use is on-line or Phase II (sometimes called Control to Standard), in which in-control process parameters are assumed to be known or estimated from Phase I. In this use analysis of data is performed sequentially, that is repeatedly after each observation.

This chapter is organized as follows. In Section 2.1 we discuss Shewhart $\overline{X}$ control charts, including extensions with warning zones and runs rules. Shewhart charts for the variance are discussed in Section 2.2. In Section 2.3 we show a simple way to calculate run length distributions, as well as a general Markov chain approach by Brook and Evans. CUSUM charts and EWMA charts are discussed in Sections 2.4 and 2.5, respectively.

2.1 The Shewhart $\overline{X}$ chart

The basic data collection scheme for a Shewhart chart is as follows. At equidistant points in time one takes a small sample (usually of size 4 or 5) of a product characteristic. Such samples are called rational subgroups. We denote the observations in rational subgroups by X_{ij}, $i = 1, 2, \ldots,$ and $j = 1, \ldots, n$, where n is the size of the rational subgroups. The rational subgroup should be chosen in such a way that the observations are independent and represent the short-term variability. The Shewhart $\overline{X}$ control chart basically is a time sequence plot of the averages $\overline{X_i}$ of the rational

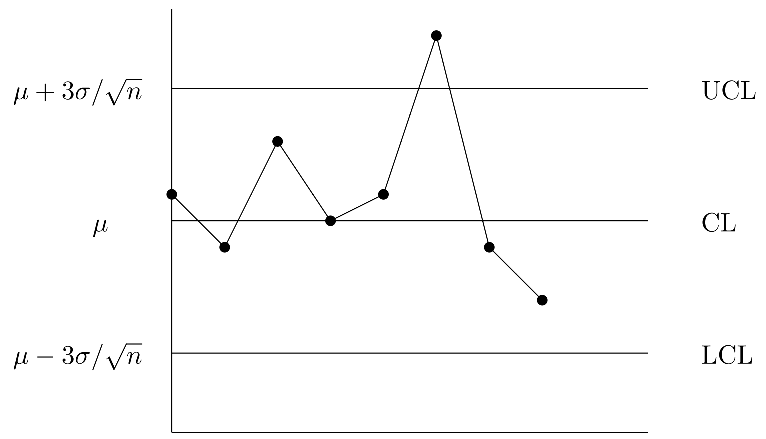

Figure 2.1: Shewhart $\overline{X}$-chart with control lines.

subgroups, together with the following 3 horizontal lines that indicate the process location and spread. Assume that μ is the mean $\mathrm{E}(X)$ of the quality characteristic and that σ^2 is the variance $\mathrm{V}(X)$ of the quality characteristic. Since $\mathrm{E}(\overline{X}_i) = \mu$ and $\mathrm{V}(\overline{X}_i) = \sigma^2/n$ for all i, the standard deviation of $\overline{X}_i$ equals $\sigma/\sqrt{n}$. The centre line (CL) is placed at the process mean μ of the quality characteristic. The other two lines are placed at distance $3\sigma/\sqrt{n}$ of the centre line. These lines are called the upper control limit (UCL) and the lower control limit (LCL), respectively. The control charts signals an alarm if an observation falls outside the region $(\mu - 3\sigma/\sqrt{n}, \mu + 3\sigma/\sqrt{n})$.

If we assume that $\overline{X}_i$ is normally distributed with mean μ and variance σ^2/n, then the probability of a false alarm equals

$$1 - P(LCL < \overline{X_i} < UCL \mid \mu = \mu_0) = \Phi(-3) + (1 - \Phi(3)) = 2(1 - \Phi(3)) \approx 0.00270 \approx \frac{1}{370}.$$

If however the process mean shifts over a distance $\sigma/\sqrt{n}$, then (see Exercise 2.2)

$$1 - P(LCL < \overline{X_i} < UCL \mid \mu = \mu_0 \pm \frac{\sigma}{\sqrt{n}}) \approx 0.02278 \approx \frac{1}{44}.$$

In Phase I use of control charts, the centre line is placed at the overall mean of the observations. The control limits are computed using (pooled) estimates of the variance from rational subgroups, often in terms of the range rather than the sample standard deviation. The relation between the range and the standard deviation is explored in detail in Section 2.2. In Phase II use of control charts, the values μ and σ are either known from historical data or estimated from a capability study. In the latter case, it is not correct to apply the above calculation, although this is usually being ignored. For the effect of estimated parameters on the false alarm probability we refer to Exercise 2.3. Shewhart based his choice for placing UCL and LCL at distance 3 times the standard deviation of group averages on practical experience. It provides a good compromise to a low false alarm rate and an quick detection of out-of-control situations. Explicit calculations will be discussed in Section 2.3.

2.1.1 Additional stopping rules for the $\overline{X}$ control chart

The decision rule to signal an alarm when a group average falls below the LCL or above the UCL is based on the current rational subgroup and ignores information from previous rational subgroups. In order to increase the detection performance of an $\overline{X}$ control chart, several additional stopping rules have been proposed.

Sometimes warning limits are added at distance $2\sigma/\sqrt{n}$ of centre line. The idea of basing control charts on different regions of observation values is also used in so-called zone control charts; for an example see Exercise 2.1 One of the proposed stopping rules using warning limits is to signal an alarm if either a single group average falls outside the control limits or two successive group averages both fall either above the upper warning limit or below the lower warning limit. Such a stopping rule is known in the SPC literature as a runs rule. Examples of other runs rules include eight successive group averages below or above the centre line or eight increasing or decreasing group averages in a row. Often a combination of such rules is used. It is obvious that such combinations should be used with care, because they lead to an increase in the false alarm rate. Explicit computations can be found in Section 2.3.

2.2 Shewhart charts for the variance

Since the normal distribution is completely specified by the parameters μ (mean) and σ^2 (variance), it does not suffice to monitor the mean of a production process. Therefore it is good practice to set up a control chart for the process variance in addition to the $\overline{X}$ control chart. Both charts use the same rational subgroups. One would expect that the sample variance is the natural choice for the statistic be monitored. However, because control charts were devised by Shewhart in the pre-computer era, a simpler statistic, the range, was and is the standard choice. The range R of a sample $Y_1, \ldots, Y_n$ is defined as

$$R := \max_{1 \le i \le n} Y_i - \min_{1 \le i \le n} Y_i = Y_{(n)} - Y_{(1)},$$

where $Y_{(i)}$ denotes the ith order statistic. It can be shown that for small sample sizes, the range is performing almost as good as the standard deviation (see Exercise 2.6 which requires results from Subsection 2.2.1). Before we define control charts for the variance, we need to study the range and standard deviation.

2.2.1 The mean and variance of the standard deviation and the range

In this subsection we derive expression for the mean and variance of the standard deviation and the range. These expressions are needed to set up control limits. If the observations X_{ij} are independent and normally distributed with mean μ and variance σ^2, then the statistic $(n-1)S_i^2/\sigma^2$ is χ^2_{n-1} distributed. The expectation of $\sqrt{n-1}\; S_i/\sigma$ is therefore equal to

$$\int_0^\infty \sqrt{t}\; \frac{1}{2^{(n-1)/2}\,\Gamma((n-1)/2)} t^{(n-1)/2-1} e^{-t/2} dt = \sqrt{2}\, \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}.$$

Hence,

$$\mathrm{E}\,(S_i) = \frac{\sqrt{2}}{\sqrt{n-1}}\, \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}\, \sigma.$$

In particular, we have

$$c_4(n) = \frac{\sqrt{2}}{\sqrt{n-1}} \frac{\Gamma(n/2)}{\Gamma((n-1)/2)}.$$

Recall that $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(x+1) = x\Gamma(x)$ for $x \neq 0, -1, -2, \ldots$. This immediately yields the following recursion for $c_4(n)$:

$$c_4(n+1) = \frac{\sqrt{n-1}}{\sqrt{n}\, c_4(n)},$$

with the initial condition $c_4(2) = \sqrt{2}/\sqrt{\pi}$. Note that the variance of S_i can also be expressed in terms of $c_4(n)$, because

$$\mathrm{V}\,(S_i) = \mathrm{E}\,(S_i^2) - (\mathrm{E}\, S_i)^2 = \sigma^2 - c_4^2(n)\sigma^2 = \left(1 - c_4^2(n)\right) \sigma^2. \tag{2.1}$$

We proceed by deriving an integral expression for $d_2(n)$.. Let F be the cumulative distribution function of X_{ij}. Then the distribution function of $\max_j(X_{ij})$ at y equals $F(y)^n$ and of $\min_j(X_{ij})$ at y equals $1-(1-F(y))^n$. We thus see that the expectation of $R_i = \max_j(X_{ij}) - \min_j(X_{ij})$ equals

$$\mathrm{E}\,(R_i) = \int_{-\infty}^{\infty} y\frac{\partial}{\partial y}(F(y)^n) - y\frac{\partial}{\partial y}(1-(1-F(y))^n)\,dy = \int_{-\infty}^{\infty} 1-F(y)^n-(1-F(y))^n\,dy,$$

where the second equality follows from integration by parts. Note that the cumulative distribution function F of a normal distribution with mean μ and variance σ^2 equals $F(y) = \Phi\left(\frac{y-\mu}{\sigma}\right)$. We thus obtain

$$\mathrm{E}\,(R_i) = \int_{-\infty}^{\infty} 1-F(y)^n-(1-F(y))^n\,dy = \int_{-\infty}^{\infty} 1-\Phi\left(\frac{y-\mu}{\sigma}\right)^n - \left(1-\Phi\left(\frac{y-\mu}{\sigma}\right)\right)^n\,dy.$$

Hence, a simple change of variables yields

$$d_2(n) = \int_{-\infty}^{\infty} 1-\Phi(y)^n-(1-\Phi(y))^n\,dy.$$

We conclude this subsection with the calculation of the variance of the range. It suffices to calculate $\mathrm{E}\,(R^2)$, because $\mathrm{V}\,(R) = \mathrm{E}\,(R^2) - \mathrm{E}\,(R)^2$. Therefore we need to determine the joint distribution G of $(\max_j(X_{ij}), \min_j(X_{ij}))$. As before we denote with $F(x)$ and $f(x)$, the cumulative distribution function and the density, respectively of the $X'_{ij}s$. We thus may write

$$\begin{aligned} G(x,y) &= P\{\max_j(X_{ij}) \le x \text{ and } \min_j(X_{ij}) \le y\} \\ &= P\{\max_j(X_{ij}) \le x\} - P\{\max_j(X_{ij}) \le x \text{ and} \min_j(X_{ij}) > y\} \\ &= F(x)^n - (F(x)-F(y))^n\, I_{\{y<x\}}, \end{aligned} \tag{2.2}$$

where $I_{\{y<x\}}$ denotes the indicator function of the set $\{y<x\}$, which equals1 if $y<x$ and 0 if $y \ge x$. The joint density thus equals

$$\frac{\partial^2}{\partial x \partial y}\,G(x,y) = n(n-1)(F(x)-F(y))^{n-2} f(x)f(y)\,I_{\{y<x\}}.$$

Hence,

$$\begin{aligned} \mathrm{E}\,(R^2) &= \int_{-\infty}^{\infty}\int_{y}^{\infty} (x-y)^2 n(n-1)(F(x)-F(y))^{n-2} f(x)f(y)\,dxdy \\ &= -\int_{-\infty}^{\infty}\int_{y}^{\infty} 2(x-y)n[(F(x)-F(y))^{n-1} - (1-F(y))^{n-1}]f(y)\,dxdy \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x} 2[(F(x)-F(y))^n - (1-F(y))^n + 1 - F(x)^n]\,dydx \\ &= 2\int_{-\infty}^{\infty}\int_{0}^{\infty} (F(y+r)-F(y))^n - (1-F(y))^n + 1 - F(y+r)^n\,drdy, \end{aligned}$$

where the second and third equalities follow from integration by parts (note the special choice of primitives, which causes some terms to vanish). If we take for F the cumulative distribution function Φ of the standard normal distribution, then we obtain that the expectation of R^2 is proportional to σ^2. One usually writes

$$\mathrm{E}\,(R^2) = \left(C(n)^2 + d_2(n)^2\right)\sigma^2,$$

so that $\mathrm{V}\,(R) = C(n)^2\sigma^2$. Numerical calculation of the double integral $\mathrm{E}\,(R^2)$ leads to values of $C(n)$.

n	c_4	d_2	A_2	D_3	D_4	$D_{.001}$	$D_{.999}$
2	0.7979	1.128	1.880	0.000	3.267		
3	0.8862	1.693	1.023	0.000	2.575		
4	0.9213	2.059	0.729	0.000	2.282	0.199	5.309
5	0.9400	2.326	0.577	0.000	2.115	0.367	5.484
6	0.9515	2.534	0.483	0.000	2.004	0.535	5.619
7	0.9594	2.704	0.419	0.076	1.924	0.691	5.730

Table 2.1: Control chart constants.

2.2.2 The R control chart

The R control chart contains 3 lines. The centre line is placed at the expectation R_0 of R, the range of a rational subgroup. If R_0 is unknown, then one uses the estimator $\overline{R}$ of a phase I study. In order to imitate the 3σ limits of the $\overline{X}$ control chart, it is convenient introduce some notation. We throughout assume that the individual observation are independent and normally distributed with mean μ and variance σ^2. We will prove in this subsection that both the mean and the standard deviation of the range are proportional to the standard deviation of individual observations. The proportionality constants are denoted by $d_2(n)$ and $C(n)$, respectively. I.e.,

$$\begin{aligned} \mathrm{E}\,(R) &= d_2(n)\sigma && (2.3)\\ \mathrm{V}\,(R) &= C(n)^2\sigma^2. && (2.4) \end{aligned}$$

The upper and lower control limits are placed at 3 times the standard deviation of R. The upper control limit, UCL, is placed at

$$R_0 + 3\sqrt{\mathrm{V}\,(R)} = R_0 + 3C(n)\sigma = \left(1 + \frac{C(n)}{d_2(n)}\right) R_0.$$

One often writes $D_4(n) = 1 + 3C(n)/d_2(n)$. The lower control limit, LCL, is similarly placed at $D_3(n)R_0$ with $D_3(n) = \max(1 - 3C(n)/d_2(n), 0)$. For an $\overline{X}$ control chart with control limits based on ranges one often uses $A_2(n) = 3/(\sqrt{n}\,d_2(n))$, so that the control limits are placed at $\mu \pm A_2(n)\overline{R}$. The range R is not normally distributed, so that alarm probabilities are not equal to 0.0027 as is the case for the $\overline{X}$ chart with known parameters. An alternative way to set up control limits is based on quantiles of the range. These control limit are called probabilistic control limits, defined as

$$\begin{aligned} UCL &= D_{.999}(n)\tfrac{\overline{R}}{d_2(n)} && (2.5)\\ LCL &= D_{.001}(n)\tfrac{\overline{R}}{d_2(n)}. && (2.6) \end{aligned}$$

These limits are set up in such a way that the upper control limit UCL will be exceeded with probability 0.001 The same holds for the lower control limit LCL. There exist tables of values for $D_\gamma(n)$. It would be logical to use the values D_γ and $D_{1-\gamma}$ with $\gamma = 1 - \Phi(3) \approx 0.00135$. The exceedance probabilities for both control limits are then equal. In practice this is not done, because tables for D_γ usually do not include these values. The control limits based on the .001 and .999 quantiles should be compared with 3.09 standard deviation control limits, because $\Phi(3.09) \approx 0.999$. We conclude this subsection with Table 2.1, which gives values of some control chart constants.

2.2.3 The S and S^2 control charts

Alternatively, one may use the statistics S_i^2 or S_i to set up an S^2 control chart or an S control chart, respectively. The control limits for the S^2 control chart follow easily because $(n-1)S^2/\sigma^2$ follows a χ^2_{n-1} distribution. One may set up control limits at distance 3 times the standard deviation of S^2 (cf. Exercise 2.7). An alternative is to use probabilistic control limits as for the R control chart (see Exercise 2.7).

Control limits for the S control chart may be set up using (2.1). The centre line is placed at the expectation $\tilde{\sigma}$ of S_i, the range of a rational subgroup. If this expectation is unknown, then one uses the estimator $\overline{S}$ of a phase I study. The UCL is placed at $\tilde{\sigma} + 3\sqrt{1 - c_4(n)^2}\frac{\tilde{\sigma}}{c_4(n)} = B_4\tilde{\sigma}$. Likewise the LCL placed at $\tilde{\sigma} - 3\sqrt{1 - c_4(n)^2}\frac{\tilde{\sigma}}{c_4(n)} = B_3\tilde{\sigma}$. Probabilistic control limits may be obtained by taking square roots of the endpoints of two-sided confidence intervals for S_i^2 and replacing σ by $\tilde{\sigma}/c_4(n)$ (see Exercise 2.8).

By setting up appropriate control limits, all three types of control charts for the variance have comparable false alarm probabilities. The difference in performance between these charts is the ability to detect changes in the process variance. This is the topic of the next section.

2.3 Calculation of run lengths for Shewhart charts

A good control chart in Phase II should have a low false alarm rate and the ability to quickly detect out-of-control situations. The goal of this section is to put these properties into quantitative terms. We will present two ways to perform calculations.

The way control charts are used is a special way of hypothesis testing. The difference with ordinary one-sample testing is that control charts in phase II involve sequential testing, since each time a new rational subgroup has been observed we again check whether the control chart signals. In that sense the false alarm probability of an individual point on the control chart is related to the type I error in hypothesis testing. The ability to detect process changes is related to the power of the test and thus to the type II error. For detailed discussions on the differences between control charts and hypothesis testing, we refer to [6, 20, 23, 24]. In order to incorporate the sequential character of control charts in phase II, the notions of type I and II error are replaced by properties of the run length N, defined as the number of rational subgroups observed until the control chart signals an alarm. The most widely studied property of the run length is its mean. In the SPC literature one usually uses the term Average Run Length (ARL). One distinguishes between an in-control ARL (ARL_{in}) and an out-of-control (ARL (ARL_{out}). Obviously, ARL_{in} should be large (low false alarm rate) and ARL_{in} should be small (fast detection of changes).

It is obvious that if the process is in control (i.e., the distribution of the observations does not change), each group average of an $\overline{X}$-chart has equal probability to fall outside the control limits. We denote this probability by p. Obviously $P(N = n) = (1-p)^{n-1}p$ for $p = 1, 2, \ldots$. Thus in this case N has a geometric distribution. Note that $P(N > n) = (1 - p)^n$. Hence,

$$\mathrm{E}\, X = \sum_{n=0}^{\infty} P(N > n) = \sum_{n=0}^{\infty} (1-p)^n = \frac{1}{p}.$$

Thus ARL_{in} of an $\overline{X}$ control chart equals $1/p = (\Phi(-3)+1-\Phi(3))^{-1} \approx 370$. A similar calculation yields that $\mathrm{V}\, N = (1-p)/p^2$. Note that if p is small as is usual the case for control charts, then the standard deviation of N almost equals its mean. The practical implication is that it may be very misleading to judge the performance of a control chart by considering ARL's only. Run length distributions are usually highly skewed with large variability. The above calculation also holds for the R, S and S^2 control charts if we use the appropriate values for p. Since it is recommended to use the $\overline{X}$ chart in conjunction with one of these three control charts for the variance, this drastically changes the run length distribution (see Exercise 2.2 for an explicit calculation). Before we proceed by discussing run lengths of control charts with runs rules, we present a general lemma, that may be useful for calculating the probability generating function of non-negative lattice distributions.

Lemma 2.3.1 *Let X be a discrete random variable taking values on $1, 2, \ldots$ and define*

$$\widetilde{P}(z) := \sum_{j=1}^{\infty} P(X \geq j)\, z^j.$$

Then we have:

1. $P(z) = 1 + \dfrac{z-1}{z}\,\widetilde{P}(z)$ *and* $\widetilde{P}(z) = z\,\dfrac{P(z)-1}{z-1}$,
   where $P(z)$ *is the probability generating function of* X.

2. $\mathrm{E}\,X_{(m)} = P^{(m)}(1) = -\sum_{k=0}^{m-1} \dfrac{m!}{k!}(-1)^{m-k}\widetilde{P}^{(k)}(1)$,
   where $\mathrm{E}\,X_{(m)} = EX(X-1)(X-2)\cdots(X-m+1)$ *is the mth factorial moment of* X *and* $P^{(n)}$ *denotes the* n^{th} *derivative of* P.
   In particular, $\mathrm{E}\,X = P'(1) = \widetilde{P}(1)$ *and* $\mathrm{V}\,(X) = 2\widetilde{P}'(1) - \widetilde{P}(1) - (\widetilde{P}(1))^2$.

*P*roof.

1. We have

$$\begin{aligned}
\widetilde{P}(z) &= \sum_{j=1}^{\infty}\sum_{k=j}^{\infty} P(X=k)z^j \\
&= \sum_{k=1}^{\infty}\sum_{j=1}^{k} P(X=k)z^j \\
&= \sum_{k=1}^{\infty} P(X=k)\frac{z^{k+1}-z}{z-1} \\
&= \frac{z}{z-1}\left(\sum_{k=1}^{\infty} P(X=k)\left(z^k-1\right)\right) \\
&= \frac{z}{z-1}(P(z)-1).
\end{aligned}$$

   Rewriting the above formula yields a).

2. The first equality is a basic property of probability generating functions. The second equality follows by applying Leibniz's formula for derivatives of products:

$$(f\cdot g)^{(m)} = \sum_{k=0}^{m}\binom{m}{k} f^{(k)}g^{(m-k)}$$

   to a) and evaluating at $z=1$. □

It has been mentioned in Subsection 2.1.1 that adding runs rules changes the ARL's of a control chart. We start with a simple example that can be analyzed directly using a recursion. Suppose we have an $\overline{X}$ chart with warning limits for which we agree that the chart signals if a point is outside the control limits or when there are two successive points beyond either of the warning limits. Let p be the probability that a point of the control charts falls outside the control limits, q be the probability that a point falls between a control limit and the nearby warning limit (the warning region), and r be the probability that a point is between the two warning limits (the "safe" region). Obviously $p+q+r=1$. Let L denote the remaining ARL after a point in the good region and L' the corresponding ARL after a point in the warning region. If we start with a point in the safe region, then there are three possibilities:

1. the next point is safe again, with probability r, and we expect L more points before an alarm
2. the next point is in the warning region, with probability q, and we expect L' more points before an alarm
3. the next point falls outside the control limits, with probability p, and we stop.

Note that we use that the points of a Shewhart chart are independent. Combining these three possibilities, we obtain

$$L = r(1+L) + q(1+L') + p.$$

In a similar way we obtain

$$L' = r(1+L) + q + p.$$

Combining these two equations, we obtain two linear equations with two unknowns. Hence, we may solve for L and L'. The ARL is then given by $p + q(L+1) + r(L'+1)$. Note that since we did not specify p, q and r the above derivation holds for both ARL_{in} and ARL_{out}. A similar way of reasoning yields ARL's for other runs rules.

In order to obtain more detailed results on run length distributions, a different approach is necessary. We present a method of Brook and Evans (see [2]), which was first presented in the context of run lengths for CUSUM charts. However, the method is very general and may also be applied to run lengths of Shewhart charts. The key in this approach is to represent the control chart as a simple Markov chain. In order to illustrate the method, we consider an $\overline{X}$ control chart with the runs rule "two successive points outside the same warning limit". The Markov chain has four states:

State 1 Current point is in the safe region.

State 2 Current point is in the lower warning region, and the previous point was not.

State 3 Current point is in the upper warning region, and the previous point was not.

State 4 Current point is outside the control limits, or is in the same warning region as the previous point (hence, the control chart signals).

Note that State 4 is an absorbing state. The probabilities of a point to fall into a certain region are denoted as follows: p_1: safe region, p_2: upper warning region, p_3: lower warning region, and $p_4 = 1 - p_1 - p_2 - p_3$ the action region. We thus have the following transition matrix, where the rows denote the state before observing a new rational subgroup and the columns after observing a new rational subgroup:

$$P = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \\ p_1 & 0 & p_3 & p_2 + p_4 \\ p_1 & p_2 & 0 & p_3 + p_4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Now let $S_k = (S_{k,1}, S_{k,2}, S_{k,3}, S_{k,4}$ be the distribution of the state space after n observed rational subgroups. Because of the Markov property, we have for$k \geq 1$ that

$$S_k = PS_{k-1} = P^{k-1}S_1,$$

where $S_1 = (p_1, p_2, p_3, p_4)$ is distribution after the first observed rational subgroup. Of course, other starting conditions can be easily incorporated. The run length distribution of N may now be computed from the relation

$$P(N \leq k) = S_{k,4}.$$

This relation can be easily programmed. It may be advantageous to partition the transition matrix P by singling out the absorbing state 4 as follows:

$$P = \begin{pmatrix} R & \mathbf{p} \\ \mathbf{0} & 1 \end{pmatrix},$$

where $\mathbf{0}$ denotes a vector of zeros. Hence, we have $P^k = \begin{pmatrix} R^k & \mathbf{c} \\ \mathbf{0} & 1 \end{pmatrix}$. Since the rows of a transition matrix sum to 1, it follows that $\mathbf{c} = \mathbf{1} - R^k\mathbf{1} = (I - R^k)\mathbf{1}$. Thus ith element of $(I - R^k)\mathbf{1}$ gives $P(N \leq k \mid \text{first point is in state i}$. This Markov chain can also be used to obtain expression for the moments of N, i.e. the moment of the run length distribution. Since N is a discrete random variable, it is convenient to study the factorial moments For more calculations of run length distributions for Shewhart charts with runs rules, we refer to [4, 7].

2.4 CUSUM Procedures

2.5 EWMA control charts

2.6 Exercises

Exercise 2.1 *Let $L < U$ be given real numbers. Assume that instead of the actual outcomes $X_1, \ldots, X_n$ of the sample we only have information on the numbers of outcomes that fall below L, fall between U and L, and fall above U. Find MLE's for μ and σ^2 based on this information only. This is used in so-called gauge control charts in statistical process control. The idea goes back to Stephens (see [19]; for modern accounts we refer to [10, 11]).*

Exercise 2.2 *Consider the standard assumptions of the Shewhart $\overline{X}$-chart with rational subgroups of size n.*

a) *What is the probability that 2 out of 3 consecutive observations fall in the region $(\mu_0 - 2\sigma/\sqrt{n}, \mu_0 + 2\sigma/\sqrt{n})$, if the individual observations are independent and normally distributed with expectation μ_0 and variance σ^2?*

b) *What is the alarm probability for a single group mean $\overline{X_i}$ if $\overline{X_i}$ is normally distributed with expectation $\mu_0 + \sigma/\sqrt{n}$ and variance σ^2/n?*

c) *We would like to change the control limits $\mu_0 \pm 3\sigma/\sqrt{n}$ into $\mu_0 \pm k\sigma/\sqrt{n}$. Which value should we choose for k in order to achieve that the probability of a single group mean $\overline{X}_i$ falling outside the control limits equals 0.01? Assume that $\overline{X}_i$ is normally distributed with expectation μ_0 and variance σ^2/n.*

Exercise 2.3 *Assume that a capability study has been performed with N observations, subdivided into m rational subgroups of size $n = N/m$. In the subsequent Phase II, a Shewhart $\overline{X}$ chart has been set up based on rational subgroups of size n. The control limits are set at set at a distance of $3\overline{S} = \sqrt{\frac{1}{N} \sum i = 1^m S_i^2}$.*

a) *Assume that the centre line is set at the known process mean. Compute the probability that a group average falls outside the control limits. Compare your result with the well-known probability 0.0027 when N ranges from 50 to 100.*

b) *Repeat your calculations for the case that the centre line is set at the overall mean from the capability study. Compare your results with a) and the standard 0.0027 probability for the same ranges of N.*

c) *The above results are independent of n. May we therefore conclude that the choice of n is unimportant?*

Exercise 2.4 *A standard Shewhart $\overline{X}$ chart has an in-control ARL equal to 370. What is the ARL of a combined Shewhart $\overline{X}$ and S^2-chart?*

Exercise 2.5 *Use a computer to produce of table of $d_2(n)$.*

Exercise 2.6 *Study the relative efficiency (i.e., the ratio of the variances) of the unbiased estimators $S/c_4(n)$ and $R/d_2(n)$ for σ when we sample from a normal distribution. For which values is $R/d_2(n)$ a reasonable alternative for $S/c_4(n)$?*

Exercise 2.7 a) *Show that $\mathrm{V}\, S^2 = \sigma^4/(n-1)^2$ when the observations form a sample of size n from a normal distribution. Use this to derive control limits placed at a distance of 3 times the standard deviation of the sample variance.*

b) *Derive probabilistic control limits for the S^2 control chart.*

Exercise 2.8 a) *Set up an S control chart by placing control limits at a distance of 3 times the standard deviation of the sample standard deviation.*

b) *Derive probabilistic control limits for the S control chart. Hint: use a two-sided confidence interval for the variance and take square roots of the end-points.*

Exercise 2.9 *Use a computer to produce of table of $D_{0.00135}$ and $D_{0.99875}$.*

Exercise 2.10 *Give a formal justification of the conditional arguments with average run lengths in the first halve of Section 2.3.*

Exercise 2.11 *Compute ARL_{in} for an $\overline{X}$ control chart with the extra runs rule "two out of three successive point outside either warning limit".*

Exercise 2.12 *A quality characteristic of a certain production process is known to be normally distributed with $\sigma = 4$. In order to obtain a good quality of the products, it is required to detect changes in the mean of 4 units. Currently rational subgroups are taken every 15 minutes.*

a) *What is the necessary minimum size of the rational subgroups if the mean detection time is 1 hour?*

b) *What should be the sampling frequency if the producer wishes to be 90% certain of detecting changes in the mean of 4 units within 1 hour?*

Exercise 2.13 *An EWMA control charts is defined in terms of moving averages as follows: $Z_n = (1-\lambda)Z_{n-1} + \lambda X_n$, where $Z_0 = \mu$ and $0 < \lambda < 1$. Prove that*

$$Z_3 = (1-\lambda)^3\mu + \lambda(1-\lambda)^2 X_1 + \lambda(1-\lambda)X_2 + \lambda X_3.$$

Assume that the observations X_1, X_2, X_3 are independent with expectation μ and variance σ^2. Calculate the expectation and variance of Z_3.

Index

Bibliography

[1] V. Abel and E. Fuchs, On the run length of one-sided cusum procedures in the change-point case, Meth. Oper. Res. 50 (1985), 133–137.

[2] D. Brook and D.A. Evans, An approach to the probability distribution of CUSUM run length, Biometrika 59 (1972), 539–550.

[3] M.E. Calzada and S.M. Scariano, Reconciling the integral equation and Markov chain approaches for computing EWMA average run lengths Communications In Statistics - Simulation and Computation 2003, 32 (2), 591–604.

[4] C.W. Champ and W.H. Woodall, Exact results for Shewhart control chart withsupplementary runs rules, Technometrics 29 (1987), 393–401.

[5] S.V. Crowder, D.M. Hawkins, M.R. Reynolds, Jr., and E. Yashchin, Process Control and Statistical Inference, Journal of Quality Technology 29 (1997), 134–139.

[6] Crowder, Stephen V., Average Run Lengths of Exponentially Weighted Moving Average Control Charts, J. Qual. Techn. 19 (1987), 161–164.

[7] R.J.M.M. Does and B.F. Schriever (1992). Variables control chart limits and tests for special causes, Statistica Neerlandica 46 (1992) 229–245.

[8] C.S. van Dobben De Bruijn, Cumulative Sum Tests: Theory and Practice, Griffin, London, 1968.

[9] P.W.M. John, Statistical Methods in Engineering and Quality Assurance, Wiley, 1990.

[10] N.R. Farnum and L.W. Stanton, Using counts to monitor a process mean, Journal of Quality Technology 18 (1986), 22–28.

[11] N.R. Farnum and L.W. Stanton, Narrow limit gauging: Go or No Go?, Quality Engineering 3 (1991), 293–307.

[12] Fu, J.C., Spiring, F.A. and Xie, H.S., On the average run lengths of quality control schemes using a Markov chain approach, Statistics & Probability Letters 56 (2002), 369–380

[13] Gan, F.F., and Chang, T.C., Computing average run lengths of exponential EWMA charts, Journal Of Quality Technology, (32) 2000, 183–187.

[14] A.L. Goel and S.M. Wu, Determination of A.R.L. and a Contour Nomogram for Cusum Charts to Control Normal Mean, Technometrics 13 (1971), 221–230.

[15] M.S. Gold, The geometric approximation to the cusum run length, Biometrika 76 (1989), 725–733.

[16] D.M. Hawkins and D.M. Olwell, Cumulative Sum Charts and Charting for Quality Improvement, Springer, 1998.

[17] K.W. Kemp, The Average Run Length of the Cumulative Sum Chart when a V-mask is used, J. Royal Stat. Soc. 23 (1961), 149–153.

[18] J.M. Lucas, Average run lengths for CUSUM schemes when observations are exponentially distributed, Technometrics 27 (1985), 145-150.

[19] W.L. Stephens, Control by gauging, Journal of the Royal Statistical Society B 10 (1948), 54-108.

[20] Z. Stoumbos, M.R. Reynolds, Jr., T.P. Ryan and W.H. Woodall, The State of Statistical Process Control as We Proceed into the 21st Century Journal of the American Statistical Association 95 (2000), 992–998.

[21] G. B. Wetherill and D.W. Brown: Statistical Process Control : Theory and Practice, Chapman and Hall, 1991.

[22] W.H. William, On the Markov chain approach to the two-sided CUSUM procedure, Technometrics 26 (1984), 41–46

[23] W.H. Woodall, Controversies and Contradictions in Statistical Process Control, Journal of Quality Technology 32 (2000), 341–378.

[24] W.H. Woodall and D.C. Montgomery, Research Issues and Ideas in Statistical Process Control, Journal of Quality Technology 31 (1999) 376–386.