Chapter 13

QUALITY CONTROL

13.1 INTRODUCTION

Almost every manufacturing process results in some random variation in the items it produces. That is, no matter how stringently the process is being controlled, there is always going to be some variation between the items produced. This variation is called *chance variation* and is considered to be inherent to the process. However, there is another type of variation that sometimes appears. This variation, far from being inherent to the process, is due to some *assignable cause* and usually results in an adverse effect on the quality of the items produced. For instance, this latter variation may be caused by a faulty machine setting, or by poor quality of the raw materials presently being used, or by incorrect software, or human error, or any other of a large number of possibilities. When the only variation present is due to chance, and not to assignable cause, we say that the process is in control, and a key problem is to determine whether a process is in or is *out of control*.

The determination of whether a process is in or out of control is greatly facilitated by the use of *control charts*, which are determined by two numbers — the upper and lower control limits. To employ such a chart, the data generated by the manufacturing process are divided into subgroups and subgroup statistics — such as the subgroup average and subgroup standard deviation — are computed. When the subgroup statistic does not fall within the upper and lower control limit, we conclude that the process is out of control.

In Sections 13.2 and 13.3, we suppose that the successive items produced have measurable characteristics, whose mean and variance are fixed when the process is in control. We show how to construct control charts based on subgroup averages (in Section 13.2) and on subgroup standard deviations (in Section 13.3). In Section 13.4, we suppose that rather than having a measurable characteristic, each item is judged by an *attribute* — that is, it is classified as either acceptable or unacceptable. Then we show how to construct control charts that can be used to indicate a change in the quality of the items produced. In Section 13.5, we consider control charts in situations where each item produced has a random number of defects. Finally, in Section 13.6 we consider more sophisticated types of control charts — ones that don't consider each subgroup value in isolation but rather take into account the values of other subgroups. Three different control charts of this type — known as moving average, exponential weighted moving average, and cumulative sum control charts — are presented in Section 13.6.

13.2 CONTROL CHARTS FOR AVERAGE VALUES: THE $\overline{X}$ CONTROL CHART

Suppose that when the process is in control the successive items produced have measurable characteristics that are independent, normal random variables with mean μ and variance σ^2. However, due to special circumstances, suppose that the process may go out of control and start producing items having a different distribution. We would like to be able to recognize when this occurs so as to stop the process, find out what is wrong, and fix it.

Let $X_1, X_2, \ldots$ denote the measurable characteristics of the successive items produced. To determine when the process goes out of control, we start by breaking the data up into subgroups of some fixed size — call it n. The value of n is chosen so as to yield uniformity within subgroups. That is, n may be chosen so that all data items within a subgroup were produced on the same day, or on the same shift, or using the same settings, and so on. In other words, the value of n is chosen so that it is reasonable that a shift in distribution would occur between and not within subgroups. Typical values of n are 4, 5, or 6.

Let $\overline{X}_i$, $i = 1, 2, \ldots$ denote the average of the ith subgroup. That is,

$$\overline{X}_1 = \frac{X_1 + \cdots + X_n}{n}$$

$$\overline{X}_2 = \frac{X_{n+1} + \cdots + X_{2n}}{n}$$

$$\overline{X}_3 = \frac{X_{2n+1} + \cdots + X_{3n}}{n}$$

and so on. Since, when in control, each of the X_i have mean μ and variance σ^2, it follows that

$$E(\overline{X}_i) = \mu$$

$$\text{Var}(\overline{X}_i) = \frac{\sigma^2}{n}$$

and so

$$\frac{\overline{X}_i - \mu}{\sqrt{\dfrac{\sigma^2}{n}}} \sim \mathcal{N}(0, 1)$$

That is, if the process is in control throughout the production of subgroup i, then $\sqrt{n}(\overline{X}_i - \mu)/\sigma$ has a standard normal distribution. Now it follows that a standard normal random variable Z will almost always be between -3 and $+3$. (Indeed, $P\{-3 < Z < 3\} = .9973$.) Hence, if the process is in control throughout the production of the items in subgroup i, then we would certainly expect that

$$-3 < \sqrt{n}\,\frac{\overline{X}_i - \mu}{\sigma} < 3$$

or, equivalently, that

$$\mu - \frac{3\sigma}{\sqrt{n}} < \overline{X}_i < \mu + \frac{3\sigma}{\sqrt{n}}$$

The values

$$\text{UCL} \equiv \mu + \frac{3\sigma}{\sqrt{n}}$$

and

$$\text{LCL} \equiv \mu - \frac{3\sigma}{\sqrt{n}}$$

are called, respectively, the *upper* and *lower control limits*.

The $\overline{X}$ control chart, which is designed to detect a change in the average value of an item produced, is obtained by plotting the successive subgroup averages $\overline{X}_i$ and declaring that the process is out of control the first time $\overline{X}_i$ does not fall between LCL and UCL (see Figure 13.1).

EXAMPLE 13.2a A manufacturer produces steel shafts having diameters that should be normally distributed with mean 3 mm and standard deviation .1 mm. Successive samples of four shafts have yielded the following sample averages in millimeters.

Sample	$\overline{X}$	Sample	$\overline{X}$
1	3.01	6	3.02
2	2.97	7	3.10
3	3.12	8	3.14
4	2.99	9	3.09
5	3.03	10	3.20

What conclusion should be drawn?

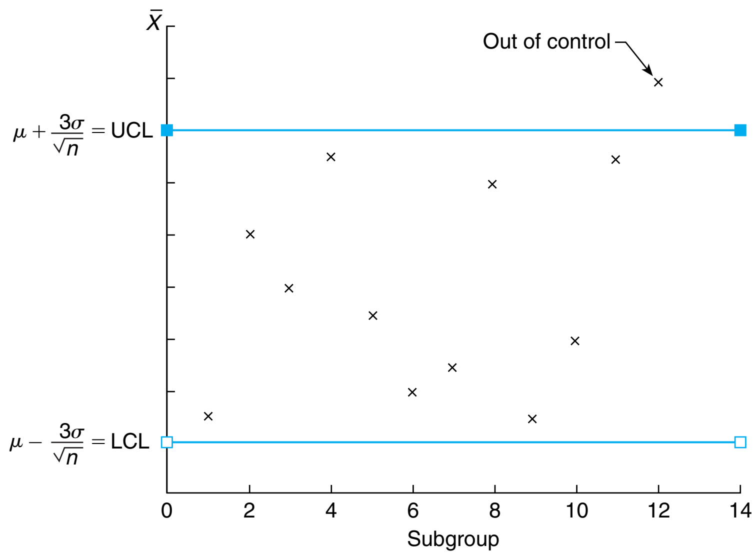

FIGURE 13.1 *Control chart for $\overline{X}$, n = size of subgroup.*

SOLUTION When in control the successive diameters have mean $\mu = 3$ and standard deviation $\sigma = .1$, and so with $n = 4$ the control limits are

$$\text{LCL} = 3 - \frac{3(.1)}{\sqrt{4}} = 2.85, \qquad \text{UCL} = 3 + \frac{3(.1)}{\sqrt{4}} = 3.15$$

Because sample number 10 falls above the upper control limit, it appears that there is reason to suspect that the mean diameter of shafts now differs from 3. (Clearly, judging from the results of Samples 6 through 10 it appears to have increased beyond 3.) ■

REMARKS

(a) The foregoing supposes that when the process is in control the underlying distribution is normal. However, even if this is not the case, by the central limit theorem it follows that the subgroup averages should have a distribution that is roughly normal and so would be unlikely to differ from its mean by more than 3 standard deviations.
(b) It is frequently the case that we do not determine the measurable qualities of all the items produced but only those of a randomly chosen subset of items. If this is so then it is natural to select, as a subgroup, items that are produced at roughly the same time.

It is important to note that even when the process is in control there is a chance — namely, .0027 — that a subgroup average will fall outside the control limit and so one would incorrectly stop the process and hunt for the nonexistent source of trouble.

Let us now suppose that the process has just gone out of control by a change in the mean value of an item from μ to $\mu + a$ where $a > 0$. How long will it take (assuming things do not change again) until the chart will indicate that the process is now out of control? To answer this, note that a subgroup average will be within the control limits if

$$-3 < \sqrt{n}\,\frac{\overline{X}-\mu}{\sigma} < 3$$

or, equivalently, if

$$-3 - \frac{a\sqrt{n}}{\sigma} < \sqrt{n}\,\frac{\overline{X}-\mu}{\sigma} - \frac{a\sqrt{n}}{\sigma} < 3 - \frac{a\sqrt{n}}{\sigma}$$

or

$$-3 - \frac{a\sqrt{n}}{\sigma} < \sqrt{n}\,\frac{\overline{X}-\mu-a}{\sigma} < 3 - \frac{a\sqrt{n}}{\sigma}$$

Hence, since $\overline{X}$ is normal with mean $\mu + a$ and variance σ^2/n — and so $\sqrt{n}(\overline{X}-\mu-a)/\sigma$ has a standard normal distribution — the probability that it will fall within the control limits is

$$\begin{aligned} P\left\{-3 - \frac{a\sqrt{n}}{\sigma} < Z < 3 - \frac{a\sqrt{n}}{\sigma}\right\} &= \Phi\left(3 - \frac{a\sqrt{n}}{\sigma}\right) - \Phi\left(-3 - \frac{a\sqrt{n}}{\sigma}\right) \\ &\approx \Phi\left(3 - \frac{a\sqrt{n}}{\sigma}\right) \end{aligned}$$

and so the probability that it falls outside is approximately $1 - \Phi(3 - a\sqrt{n}/\sigma)$. For instance, if the subgroup size is $n = 4$, then an increase in the mean value of 1 standard deviation — that is, $a = \sigma$ — will result in the subgroup average falling outside of the control limits with probability $1 - \Phi(1) = .1587$. Because each subgroup average will independently fall outside the control limits with probability $1 - \Phi(3 - a\sqrt{n}/\sigma)$, it follows that the number of subgroups that will be needed to detect this shift has a geometric distribution with mean $\{1 - \Phi(3 - a\sqrt{n}/\sigma)\}^{-1}$. (In the case mentioned before with $n = 4$, the number of subgroups one would have to chart to detect a change in the mean of 1 standard deviation has a geometric distribution with mean $1/.158 \approx 6.3$.)

13.2.1 Case of Unknown μ and σ

If one is just starting up a control chart and does not have reliable historical data, then μ and σ would not be known and would have to be estimated. To do so, we employ k of the subgroups where k should be chosen so that $k \geq 20$ and $nk \geq 100$. If $\overline{X}_i$, $i = 1, \ldots, k$ is the average of the ith subgroup, then it is natural to estimate μ by $\overline{\overline{X}}$ the average of these subgroup averages. That is,

$$\overline{\overline{X}} = \frac{\overline{X}_1 + \cdots + \overline{X}_k}{k}$$

To estimate σ, let S_i denote the sample standard deviation of the ith subgroup, $i = 1, \ldots, k$. That is,

$$S_1 = \sqrt{\sum_{i=1}^{n} \frac{(X_i - \overline{X}_1)^2}{n-1}}$$

$$S_2 = \sqrt{\sum_{i=1}^{n} \frac{(X_{n+i} - \overline{X}_2)^2}{n-1}}$$

$$\vdots$$

$$S_k = \sqrt{\sum_{i=1}^{n} \frac{(X_{(k-1)n+i} - \overline{X}_k)^2}{n-1}}$$

Let

$$\overline{S} = (S_1 + \cdots + S_k)/k$$

The statistic $\overline{S}$ will not be an unbiased estimator of σ — that is, $E[\overline{S}] \neq \sigma$. To transform it into an unbiased estimator, we must first compute $E[\overline{S}]$, which is accomplished as follows:

$$\begin{aligned} E[\overline{S}] &= \frac{E[S_1] + \cdots + E[S_k]}{k} \\ &= E[S_1] \end{aligned} \tag{13.2.1}$$

where the last equality follows since $S_1, \ldots, S_k$ are independent and identically distributed (and thus have the same mean). To compute $E[S_1]$, we make use of the following fundamental result about normal samples — namely, that

$$\frac{(n-1)S_1^2}{\sigma^2} = \sum_{i=1}^{n} \frac{(X_i - \overline{X}_1)^2}{\sigma^2} \sim \chi^2_{n-1} \tag{13.2.2}$$

Now it is not difficult to show (see Problem 3) that

$$E[\sqrt{Y}] = \frac{\sqrt{2}\,\Gamma(n/2)}{\Gamma(\frac{n-1}{2})} \quad \text{when } Y \sim \chi^2_{n-1} \tag{13.2.3}$$

Since

$$E[\sqrt{(n-1)S^2/\sigma^2}] = \sqrt{n-1}E[S_1]/\sigma$$

we see from Equations 13.2.2 and 13.2.3 that

$$E[S_1] = \frac{\sqrt{2}\Gamma(n/2)\sigma}{\sqrt{n-1}\Gamma(\frac{n-1}{2})}$$

Hence, if we set

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

then it follows from Equation 13.2.1 that $\overline{S}/c(n)$ is an unbiased estimator of σ.

Table 13.1 presents the values of $c(n)$ for $n = 2$ through $n = 10$.

TECHNICAL REMARK

In determining the values in Table 13.1, the computation of $\Gamma(n/2)$ and $\Gamma(n - \frac{1}{2})$ was based on the recursive formula

$$\Gamma(a) = (a-1)\Gamma(a-1)$$

TABLE 13.1 *Values of c(n)*

$c\,(2)$	=	.7978849
$c\,(3)$	=	.8862266
$c\,(4)$	=	.9213181
$c\,(5)$	=	.9399851
$c\,(6)$	=	.9515332
$c\,(7)$	=	.9593684
$c\,(8)$	=	.9650309
$c\,(9)$	=	.9693103
$c\,(10)$	=	.9726596

which was established in Section 5.7. This recursion yields that, for integer n,

$$\begin{aligned}\Gamma(n) &= (n-1)(n-2)\cdots 3\cdot 2\cdot 1\cdot \Gamma(1)\\ &= (n-1)! \quad \text{since } \Gamma(1) = \int_0^\infty e^{-x}\,dx = 1\end{aligned}$$

The recursion also yields that

$$\Gamma\left(\frac{n+1}{2}\right) = \left(\frac{n-1}{2}\right)\left(n-\frac{3}{2}\right)\cdots\frac{3}{2}\cdot\frac{1}{2}\cdot\Gamma\left(\frac{1}{2}\right)$$

with

$$
\begin{aligned}
\Gamma\left(\frac{1}{2}\right) &= \int_0^\infty e^{-x}x^{-1/2}\,dx \\
&= \int_0^\infty e^{-y^2/2}\frac{\sqrt{2}}{y}y\,dy \quad \text{by } x = \frac{y^2}{2} \quad dx = y\,dy \\
&= \sqrt{2}\int_0^\infty e^{-y^2/2}\,dy \\
&= 2\sqrt{\pi}\,\frac{1}{\sqrt{2\pi}}\int_0^\infty e^{-y^2/2}\,dy \\
&= 2\sqrt{\pi}\,P[N(0,1) > 0] \\
&= \sqrt{\pi}
\end{aligned}
$$

The preceding estimates for μ and σ make use of all k subgroups and thus are reasonable only if the process has remained in control throughout. To check this, we compute the control limits based on these estimates of μ and σ, namely,

$$
\begin{aligned}
\text{LCL} &= \overline{\overline{X}} - \frac{3\overline{S}}{c(n)\sqrt{n}} \qquad (13.2.4) \\
\text{UCL} &= \overline{\overline{X}} + \frac{3\overline{S}}{c(n)\sqrt{n}}
\end{aligned}
$$

We now check that each of the subgroup averages $\overline{X}_i$ falls within these lower and upper limits. Any subgroup whose average value does not fall within the limits is removed (we suppose that the process was temporarily out of control) and the estimates are recomputed. We then again check that all the remaining subgroup averages fall within the control limits. If not, then they are removed, and so on. Of course, if too many of the subgroup averages fall outside the control limits, then it is clear that no control has yet been established.

EXAMPLE 13.2b Let us reconsider Example 13.2a under the new supposition that the process is just beginning and so μ and σ are unknown. Also suppose that the sample standard deviations were as follows:

	$\overline{X}$	S		$\overline{X}$	S
1	3.01	.12	6	3.02	.08
2	2.97	.14	7	3.10	.15
3	3.12	.08	8	3.14	.16
4	2.99	.11	9	3.09	.13
5	3.03	.09	10	3.20	.16

Since $\overline{\overline{X}} = 3.067$, $\overline{S} = .122$, $c(4) = .9213$, the control limits are

$$\text{LCL} = 3.067 - \frac{3(.122)}{2 \times .9213} = 2.868$$

$$\text{UCL} = 3.067 + \frac{3(.122)}{2 \times .9213} = 3.266$$

Since all the $\overline{X}_i$ fall within these limits, we suppose that the process is in control with $\mu = 3.067$ and $\sigma = \overline{S}/c(4) = .1324$.

Suppose now that the values of the items produced are supposed to fall within the specifications $3 \pm .1$. Assuming that the process remains in control and that the foregoing are accurate estimates of the true mean and standard deviation, what proportion of the items will meet the desired specifications?

SOLUTION To answer the foregoing, we note that when $\mu = 3.067$ and $\sigma = .1324$,

$$\begin{aligned} P\{2.9 \leq X \leq 3.1\} &= P\left\{\frac{2.9 - 3.067}{.1324} \leq \frac{X - 3.067}{.1324} \leq \frac{3.1 - 3.067}{.1324}\right\} \\ &= \Phi(.2492) - \Phi(-1.2613) \\ &= .5984 - (1 - .8964) \\ &= .4948 \end{aligned}$$

Hence, 49 percent of the items produced will meet the specifications. ■

REMARKS

(a) The estimator $\overline{\overline{X}}$ is equal to the average of all nk measurements and is thus the obvious estimator of μ. However, it may not immediately be clear why the sample standard deviation of all the nk measurements, namely,

$$S \equiv \sqrt{\sum_{i=1}^{nk} \frac{(X_i - \overline{\overline{X}})^2}{nk - 1}}$$

is not used as the initial estimator of σ. The reason it is not is that the process may not have been in control throughout the first k subgroups, and thus this latter estimator could be far away from the true value. Also, it often happens that a process goes out of control by an occurrence that results in a change of its mean value μ while leaving its standard deviation unchanged. In such a case, the subgroup sample deviations would still be estimators of σ, whereas the entire sample standard deviation would not. Indeed, even in the case where the process appears to be in control throughout, the estimator of σ presented is preferred over the sample standard deviation S. The reason for this is that we cannot be certain that the mean has not changed throughout this time. That is, even though all the subgroup averages fall within the control limits, and so we have concluded that the process is in control, there is no assurance that there are no assignable causes of variation present (which might have resulted in a change in the mean that has not yet been picked up by the chart). It merely means that for practical purposes it pays to act as if the process was in control and let it continue to produce items. However, since we realize that some assignable cause of variation might be present, it has been argued that $\overline{S}/c(n)$ is a "safer" estimator than the sample standard deviation. That is, although it is not quite as good when the process has really been in control throughout, it could be a lot better if there had been some small shifts in the mean.

(b) In the past, an estimator of σ based on subgroup ranges — defined as the difference between the largest and smallest value in the subgroup — has been employed. This was done to keep the necessary computations simple (it is clearly much easier to compute the range than it is to compute the subgroup's sample standard deviation). However, with modern-day computational power this should no longer be a consideration, and since the standard deviation estimator both has smaller variance than the range estimator and is more robust (in the sense that it would still yield a reasonable estimate of the population standard deviation even when the underlying distribution is not normal), we will not consider the latter estimator in this text.

13.3 *S*-CONTROL CHARTS

The $\overline{X}$ control charts presented in the previous section are designed to pick up changes in the population mean. In cases where one is also concerned about possible changes in the population variance, we can utilize an S-control chart.

As before, suppose that, when in control, the items produced have a measurable characteristic that is normally distributed with mean μ and variance σ^2. If S_i is the sample standard deviation for the ith subgroup, that is,

$$S_i = \sqrt{\sum_{j=1}^{n} \frac{(X_{(i-1)n+j} - \overline{X}_i)^2}{(n-1)}}$$

then, as was shown in Section 13.2.1,

$$E[S_i] = c(n)\sigma \tag{13.3.1}$$

In addition,

$$\begin{aligned} \text{Var}(S_i) &= E[S_i^2] - (E[S_i])^2 \\ &= \sigma^2 - c^2(n)\sigma^2 \\ &= \sigma^2[1 - c^2(n)] \end{aligned} \tag{13.3.2}$$

where the next to last equality follows from Equation 13.2.2 and the fact that the expected value of a chi-square random variable is equal to its degrees of freedom parameter.

On using the fact that, when in control, S_i has the distribution of a constant (equal to $\sigma/\sqrt{n-1}$) times the square root of a chi-square random variable with $n-1$ degrees of freedom, it can be shown that S_i will, with probability near to 1, be within 3 standard deviations of its mean. That is,

$$P\{E[S_i] - 3\sqrt{\text{Var}(S_i)} < S_i < E[S_i] + 3\sqrt{\text{Var}(S_i)}\} \approx .99$$

Thus, using the formulas 13.3.1 and 13.3.2 for $E[S_i]$ and $\text{Var}(S_i)$, it is natural to set the upper and lower control limits for the S chart by

$$\text{UCL} = \sigma[c(n) + 3\sqrt{1 - c^2(n)}] \tag{13.3.3}$$

$$\text{LCL} = \sigma[c(n) - 3\sqrt{1 - c^2(n)}]$$

The successive values of S_i should be plotted to make certain they fall within the upper and lower control limits. When a value falls outside, the process should be stopped and declared to be out of control.

When one is just starting up a control chart and σ is unknown, it can be estimated from $\overline{S}/c(n)$. Using the foregoing, the estimated control limits would then be

$$\text{UCL} = \overline{S}[1 + 3\sqrt{1/c^2(n) - 1}] \tag{13.3.4}$$

$$\text{LCL} = \overline{S}[1 - 3\sqrt{1/c^2(n) - 1}]$$

As in the case of starting up an $\overline{X}$ control chart, it should then be checked that the k subgroup standard deviations $S_1, S_2, \ldots, S_k$ all fall within these control limits. If any of them falls outside, then those subgroups should be discarded and $\overline{S}$ recomputed.

EXAMPLE 13.3a The following are the $\overline{X}$ and S values for 20 subgroups of size 5 for a recently started process.

Subgroup	$\overline{X}$	S	Subgroup	$\overline{X}$	S	Subgroup	$\overline{X}$	S	Subgroup	$\overline{X}$	S
1	35.1	4.2	6	36.4	4.5	11	38.1	4.2	16	41.3	8.2
2	33.2	4.4	7	35.9	3.4	12	37.6	3.9	17	35.7	8.1
3	31.7	2.5	8	38.4	5.1	13	38.8	3.2	18	36.3	4.2
4	35.4	3.2	9	35.7	3.8	14	34.3	4.0	19	35.4	4.1
5	34.5	2.6	10	27.2	6.2	15	43.2	3.5	20	34.6	3.7

Since $\overline{\overline{X}} = 35.94$, $\overline{S} = 4.35$, $c(5) = .9400$, we see from Equations 13.2.4 and 13.3.4 that the preliminary upper and lower control limits for $\overline{X}$ and S are

$$\mathrm{UCL}(\overline{X}) = 42.149$$
$$\mathrm{LCL}(\overline{X}) = 29.731$$
$$\mathrm{UCL}(S) = 9.087$$
$$\mathrm{LCL}(S) = -.386$$

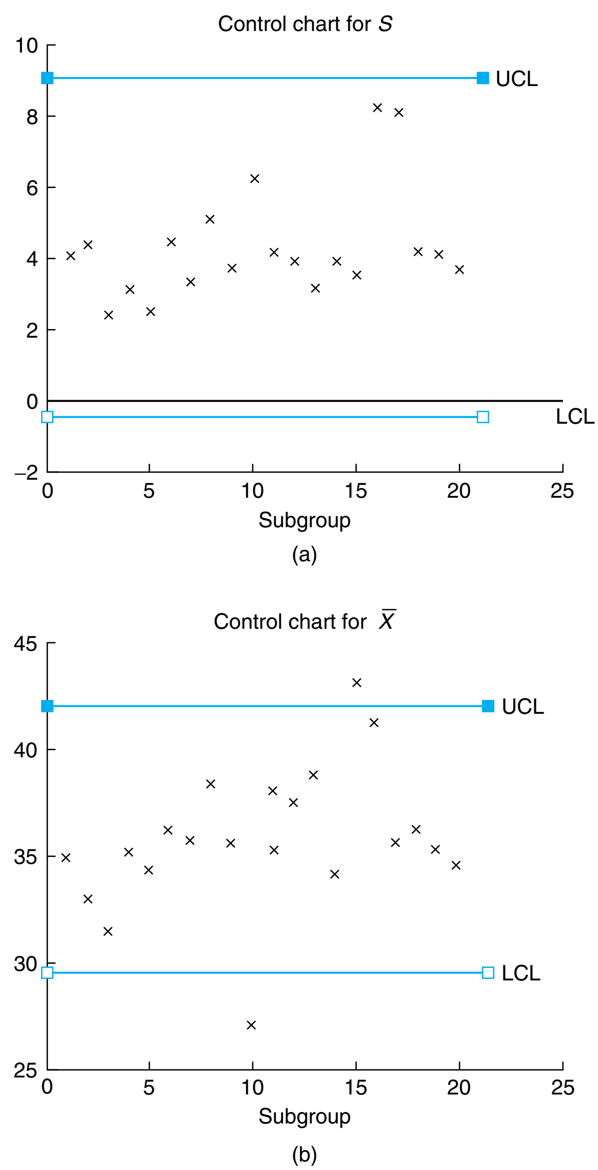

FIGURE 13.2

The control charts for $\overline{X}$ and S with the preceding control limits are shown in Figures 13.2a and 13.2b. Since $\overline{X}_{10}$ and $\overline{X}_{15}$ fall outside the $\overline{X}$ control limits, these subgroups must be eliminated and the control limits recomputed. We leave the necessary computations as an exercise. ■

13.4 CONTROL CHARTS FOR THE FRACTION DEFECTIVE

The $\overline{X}$ and S-control charts can be used when the data are measurements whose values can vary continuously over a region. There are also situations in which the items produced have quality characteristics that are classified as either being defective or nondefective. Control charts can also be constructed in this latter situation.

Let us suppose that when the process is in control each item produced will independently be defective with probability p. If we let X denote the number of defective items in a subgroup of n items, then assuming control, X will be a binomial random variable with parameters (n, p). If $F = X/n$ is the fraction of the subgroup that is defective, then assuming the process is in control, its mean and standard deviation are given by

$$E[F] = \frac{E[X]}{n} = \frac{np}{n} = p$$

$$\sqrt{\text{Var}(F)} = \sqrt{\frac{\text{Var}(X)}{n^2}} = \sqrt{\frac{np(1-p)}{n^2}} = \sqrt{\frac{p(1-p)}{n}}$$

Hence, when the process is in control the fraction defective in a subgroup of size n should, with high probability, be between the limits

$$\text{LCL} = p - 3\sqrt{\frac{p(1-p)}{n}}, \quad \text{UCL} = p + 3\sqrt{\frac{p(1-p)}{n}}$$

The subgroup size n is usually much larger than the typical values of between 4 and 10 used in $\overline{X}$ and S charts. The main reason for this is that if p is small and n is not of reasonable size, then most of the subgroups will have zero defects even when the process goes out of control. Thus, it would take longer than it would if n were chosen so that np were not close to zero to detect a shift in quality.

To start such a control chart it is, of course, necessary first to estimate p. To do so, choose k of the subgroups, where again one should try to take $k \geq 20$, and let F_i denote the fraction of the ith subgroup that are defective. The estimate of p is given by $\overline{F}$ defined by

$$\overline{F} = \frac{F_1 + \cdots + F_k}{k}$$

Since nF_i is equal to the number of defectives in subgroup i, we see that F_k can also be expressed as

$$\overline{F} = \frac{nF_1 + \cdots + nF_k}{nk}$$
$$= \frac{\text{total number of defectives in all the subgroups}}{\text{number of items in the subgroups}}$$

In other words, the estimate of p is just the proportion of items inspected that are defective. The upper and lower control limits are now given by

$$\text{LCL} = \overline{F} - 3\sqrt{\frac{\overline{F}(1-\overline{F})}{n}}, \qquad \text{UCL} = \overline{F} + 3\sqrt{\frac{\overline{F}(1-\overline{F})}{n}}$$

We should now check whether the subgroup fractions $F_1, F_2, \ldots, F_k$ fall within these control limits. If some of them fall outside, then the corresponding subgroups should be eliminated and $\overline{F}$ recomputed.

EXAMPLE 13.4a Successive samples of 50 screws are drawn from the hourly production of an automatic screw machine, with each screw being rated as either acceptable or defective. This is done for 20 such samples with the following data resulting.

Subgroup	Defectives	F	Subgroup	Defectives	F
1	6	.12	11	1	.02
2	5	.10	12	3	.06
3	3	.06	13	2	.04
4	0	.00	14	0	.00
5	1	.02	15	1	.02
6	2	.04	16	1	.02
7	1	.02	17	0	.00
8	0	.00	18	2	.04
9	2	.04	19	1	.02
10	1	.02	20	2	.04

We can compute the trial control limits as follows:

$$\overline{F} = \frac{\text{total number defectives}}{\text{total number items}} = \frac{34}{1{,}000} = .034$$

and so

$$\text{UCL} = .034 + 3\sqrt{\frac{(.034)(.968)}{50}} = .1109$$

$$\text{LCL} = .034 - 3\sqrt{\frac{(.034)(.966)}{50}} = -.0429$$

Since the proportion of defectives in the first subgroup falls outside the upper control limit, we eliminate that subgroup and recompute $\overline{F}$ as

$$\overline{F} = \frac{34 - 6}{950} = .0295$$

The new upper and lower control limits are $.0295 \pm \sqrt{(.0295)(1 - .0295)/50}$, or

$$\text{LCL} = -.0423, \quad \text{UCL} = .1013$$

Since the remaining subgroups all have fraction defectives that fall within the control limits, we can accept that, when in control, the fraction of defective items in a subgroup should be below .1013. ■

REMARK

Note that we are attempting to detect any change in quality even when this change results in improved quality. That is, we regard the process as being "out of control" even when the probability of a defective item decreases. The reason for this is that it is important to notice any change in quality, for either better or worse, to be able to evaluate the reason for the change. In other words, if an improvement in product quality occurs, then it is important to analyze the production process to determine the reason for the improvement. (That is, what are we doing right?)

13.5 CONTROL CHARTS FOR NUMBER OF DEFECTS

In this section, we consider situations in which the data are the numbers of defects in units that consist of an item or group of items. For instance, it could be the number of defective rivets in an airplane wing, or the number of defective computer chips that are produced daily by a given company. Because it is often the case that there are a large number of possible things that can be defective, with each of these having a small probability of actually being defective, it is probably reasonable to assume that the resulting number of defects has a Poisson distribution.* So let us suppose that, when the process is in control, the number of defects per unit has a Poisson distribution with mean λ.

---

* See Section 5.2 for a theoretical explanation.

If we let X_i denote the number of defects in the ith unit, then, since the variance of a Poisson random variable is equal to its mean, when the process is in control

$$E[X_i] = \lambda, \qquad \text{Var}(X_i) = \lambda$$

Hence, when in control each X_i should with high probability be within $\lambda \pm 3\sqrt{\lambda}$, and so the upper and lower control limits are given by

$$\text{UCL} = \lambda + 3\sqrt{\lambda}, \qquad \text{LCL} = \lambda - 3\sqrt{\lambda}$$

As before, when the control chart is started and λ is unknown, a sample of k units should be used to estimate λ by

$$\overline{X} = (X_1 + \cdots + X_k)/k$$

This results in trial control limits

$$\overline{X} + 3\sqrt{\overline{X}} \quad \text{and} \quad \overline{X} - 3\sqrt{\overline{X}}$$

If all the $X_i, i = 1, \ldots, k$ fall within these limits, then we suppose that the process is in control with $\lambda = \overline{X}$. If some fall outside, then these points are eliminated and we recompute $\overline{X}$, and so on.

In situations where the mean number of defects per item (or per day) is small, one should combine items (days) and use as data the number of defects in a given number — say, n — of items (or days). Since the sum of independent Poisson random variables remains a Poisson random variable, the data values will be Poisson distributed with a larger mean value λ. Such combining of items is useful when the mean number of defects per item is less than 25.

To obtain a feel for the advantage in combining items, suppose that the mean number of defects per item is 4 when the process is under control, and suppose that something occurs that results in this value changing from 4 to 6, that is, an increase of 1 standard deviation occurs. Let us see how many items will be produced, on average, until the process is declared out of control when the successive data consist of the number of defects in n items.

Since the number of defects in a sample of n items is, when under control, Poisson distributed with mean and variance equal to $4n$, the control limits are $4n \pm 3\sqrt{4n}$ or $4n \pm 6\sqrt{n}$. Now if the mean number of defects per item changes to 6, then a data value will be Poisson with mean $6n$ and so the probability that it will fall outside the control limits — call it $p(n)$ — is given by

$$p(n) = P\{Y > 4n + 6\sqrt{n}\} + P\{Y < 4n - 6\sqrt{n}\}$$

when Y is Poisson with mean $6n$. Now

$$\begin{aligned}
p(n) &\approx P\{Y > 4n + 6\sqrt{n}\} \\
&= P\left\{\frac{Y-6n}{\sqrt{6n}} > \frac{6\sqrt{n}-2n}{\sqrt{6n}}\right\} \\
&\approx P\left\{Z > \frac{6\sqrt{n}-2n}{\sqrt{6n}}\right\} \quad \text{where } Z \sim N(0,1) \\
&= 1 - \Phi\left(\sqrt{6} - 2\sqrt{\frac{n}{6}}\right)
\end{aligned}$$

Because each data value will be outside the control limits with probability $p(n)$, it follows that the number of data values needed to obtain one outside the limits is a geometric random variable with parameter $p(n)$, and thus has mean $1/p(n)$. Finally, since there are n items for each data value, it follows that the number of items produced before the process is seen to be out of control has mean value $n/p(n)$:

$$\text{Average number of items produced while out of control} = n/(1 - \Phi(\sqrt{6} - \sqrt{\tfrac{2n}{3}}))$$

We plot this for various n in Table 13.2. Since larger values of n are better when the process is in control (because the average number of items produced before the process is incorrectly said to be out of control is approximately $n/.0027$), it is clear from Table 13.2 that one should combine at least 9 of the items. This would mean that each data value (equal to the number of defects in the combined set) would have mean at least $9 \times 4 = 36$.

TABLE 13.2

n	Average Number of Items
1	19.6
2	20.66
3	19.80
4	19.32
5	18.80
6	18.18
7	18.13
8	18.02
9	18
10	18.18
11	18.33
12	18.51

EXAMPLE 13.5a The following data represent the number of defects discovered at a factory on successive units of 10 cars each.

Cars	Defects	Cars	Defects	Cars	Defects	Cars	Defects
1	141	6	74	11	63	16	68
2	162	7	85	12	74	17	95
3	150	8	95	13	103	18	81
4	111	9	76	14	81	19	102
5	92	10	68	15	94	20	73

Does it appear that the production process was in control throughout?

SOLUTION Since $\overline{X} = 94.4$, it follows that the trial control limits are

$$\text{LCL} = 94.4 - 3\sqrt{94.4} = 65.25$$

$$\text{UCL} = 94.4 + 3\sqrt{94.4} = 123.55$$

Since the first three data values are larger than UCL, they are removed and the sample mean recomputed. This yields

$$\overline{X} = \frac{(94.4)20 - (141 + 162 + 150)}{17} = 84.41$$

and so the new trial control limits are

$$\text{LCL} = 84.41 - 3\sqrt{84.41} = 56.85$$

$$\text{UCL} = 84.41 + 3\sqrt{84.41} = 111.97$$

At this point since all remaining 17 data values fall within the limits, we could declare that the process is now in control with a mean value of 84.41. However, because it seems that the mean number of defects was initially high before settling into control, it seems quite plausible that the data value X_4 also originated before the process was in control. Thus, it would seem prudent in this situation to also eliminate X_4 and recompute. Based on the remaining 16 data values, we obtain that

$$\overline{X} = 82.56$$

$$\text{LCL} = 82.56 - 3\sqrt{82.56} = 55.30$$

$$\text{UCL} = 82.56 + 3\sqrt{82.56} = 109.82$$

and so it appears that the process is now in control with a mean value of 82.56. ■

13.6 OTHER CONTROL CHARTS FOR DETECTING CHANGES IN THE POPULATION MEAN

The major weakness of the $\overline{X}$ control chart presented in Section 13.2 is that it is relatively insensitive to small changes in the population mean. That is, when such a change occurs, since each plotted value is based on only a single subgroup and so tends to have a relatively large variance, it takes, on average, a large number of plotted values to detect the change. One way to remedy this weakness is to allow each plotted value to depend not only on the most recent subgroup average but on some of the other subgroup averages as well. Three approaches for doing this that have been found to be quite effective are based on (1) moving averages, (2) exponentially weighted moving averages, and (3) cumulative sum control charts.

13.6.1 Moving-Average Control Charts

The moving-average control chart of span size k is obtained by continually plotting the average of the k most recent subgroups. That is, the moving average at time t, call it M_t, is defined by

$$M_t = \frac{\overline{X}_t + \overline{X}_{t-1} + \cdots + \overline{X}_{t-k+1}}{k}$$

where $\overline{X}_i$ is the average of the values of subgroup i. The successive computations can be easily performed by noting that

$$kM_t = \overline{X}_t + \overline{X}_{t-1} + \cdots + \overline{X}_{t-k+1}$$

and, substituting $t+1$ for t,

$$kM_{t+1} = \overline{X}_{t+1} + \overline{X}_t + \cdots + \overline{X}_{t-k+2}$$

Subtraction now yields that

$$kM_{t+1} - kM_t = \overline{X}_{t+1} - \overline{X}_{t-k+1}$$

or

$$M_{t+1} = M_t + \frac{\overline{X}_{t+1} - \overline{X}_{t-k+1}}{k}$$

In words, the moving average at time $t+1$ is equal to the moving average at time t plus $1/k$ times the difference between the newly added and the deleted value in the moving average. For values of t less than k, M_t is defined as the average of the first t subgroups. That is,

$$M_t = \frac{\overline{X}_1 + \cdots + \overline{X}_t}{t} \qquad \text{if } t < k$$

Suppose now that when the process is in control the successive values come from a normal population with mean μ and variance σ^2. Therefore, if n is the subgroup size, it follows that $\overline{X}_i$ is normal with mean μ and variance σ^2/n. From this we see that the average of m of the $\overline{X}_i$ will be normal with mean μ and variance given by $\text{Var}(\overline{X}_i)/m = \sigma^2/nm$ and, therefore, when the process is in control

$$E[M_t] = \mu$$
$$\text{Var}(M_t) = \begin{cases} \sigma^2/nt & \text{if } t < k \\ \sigma^2/nk & \text{otherwise} \end{cases}$$

Because a normal random variable is almost always within 3 standard deviations of its mean, we have the following upper and lower control limits for M_t:

$$\text{UCL} = \begin{cases} \mu + 3\sigma/\sqrt{nt} & \text{if } t < k \\ \mu + 3\sigma/\sqrt{nk} & \text{otherwise} \end{cases}$$

$$\text{LCL} = \begin{cases} \mu - 3\sigma/\sqrt{nt} & \text{if } t < k \\ \mu - 3\sigma/\sqrt{nk} & \text{otherwise} \end{cases}$$

In other words, aside from the first $k - 1$ moving averages, the process will be declared out of control whenever a moving average differs from μ by more than $3\sigma/\sqrt{nk}$.

EXAMPLE 13.6a When a certain manufacturing process is in control, it produces items whose values are normally distributed with mean 10 and standard deviation 2. The following simulated data represent the values of 25 subgroup averages of size 5 from a normal population with mean 11 and standard deviation 2. That is, these data represent the subgroup averages after the process has gone out of control with its mean value increasing from 10 to 11. Table 13.3 presents these 25 values along with the moving averages based on span size $k = 8$ as well as the upper and lower control limits. The lower and upper control limits for $t > 8$ are 9.051318 and 10.94868.

As the reader can see, the first moving average to fall outside its control limits occurred at time 11, with other such occurrences at times 12, 13, 14, 16, and 25. (It is interesting to note that the usual control chart — that is, the moving average with $k = 1$ — would have declared the process out of control at time 7 since $\overline{X}_7$ was so large. However, this is the only point where this chart would have indicated a lack of control (see Figure 13.3).

TABLE 13.3

t	$\overline{X}_t$	M_t	LCL	UCL
1	9.617728	9.617728	7.316719	12.68328
2	10.25437	9.936049	8.102634	11.89737
3	9.876195	9.913098	8.450807	11.54919
4	10.79338	10.13317	8.658359	11.34164
5	10.60699	10.22793	8.8	11.2
6	10.48396	10.2706	8.904554	11.09545
7	13.33961	10.70903	8.95815	11.01419
8	9.462969	10.55328	9.051318	10.94868
			⋮	⋮
9	10.14556	10.61926		
10	11.66342	10.79539		
*11	11.55484	11.00634		
*12	11.26203	11.06492		
*13	12.31473	11.27839		
*14	9.220009	11.1204		
15	11.25206	10.85945		
16	10.48662	10.98741		
17	9.025091	10.84735		
18	9.693386	10.6011		
19	11.45989	10.58923		
20	12.44213	10.73674		
21	11.18981	10.59613		
22	11.56674	10.88947		
23	9.869849	10.71669		
24	12.11311	10.92		
*25	11.48656	11.22768		

* = *Out of control.*

There is an inverse relationship between the size of the change in the mean value that one wants to guard against and the appropriate moving-average span size k. That is, the smaller this change is, the larger k ought to be. ■

13.6.2 Exponentially Weighted Moving-Average Control Charts

The moving-average control chart of Section 13.6.1 considered at each time t a weighted average of all subgroup averages up to that time, with the k most recent values being given weight $1/k$ and the others given weight 0. Since this appears to be a most effective procedure for detecting small changes in the population mean, it raises the possibility that other sets of weights might also be successfully employed. One set of weights that is often utilized is obtained by decreasing the weight of each earlier subgroup average by a constant factor.

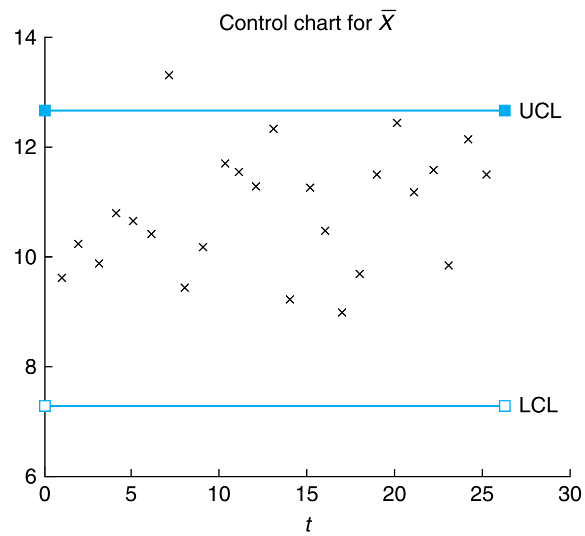

FIGURE 13.3

Let

$$W_t = \alpha \overline{X}_t + (1-\alpha)W_{t-1} \tag{13.6.1}$$

where α is a constant between 0 and 1, and where

$$W_0 = \mu$$

The sequence of values $W_t, t = 0, 1, 2, \ldots$ is called an *exponentially weighted moving average*. To understand why it has been given that name, note that if we continually substitute for the W term on the right side of Equation 13.6.1, we obtain that

$$\begin{aligned} W_t &= \alpha\overline{X}_t + (1-\alpha)[\alpha\overline{X}_{t-1} + (1-\alpha)W_{t-2}] \qquad (13.6.2)\\ &= \alpha\overline{X}_t + \alpha(1-\alpha)\overline{X}_{t-1} + (1-\alpha)^2 W_{t-2}\\ &= \alpha\overline{X}_t + \alpha(1-\alpha)\overline{X}_{t-1} + (1-\alpha)^2[\alpha\overline{X}_{t-2} + (1-\alpha)W_{t-3}]\\ &= \alpha\overline{X}_t + \alpha(1-\alpha)\overline{X}_{t-1} + \alpha(1-\alpha)^2\overline{X}_{t-2} + (1-\alpha)^3 W_{t-3}\\ &\vdots\\ &= \alpha\overline{X}_t + \alpha(1-\alpha)\overline{X}_{t-1} + \alpha(1-\alpha)^2\overline{X}_{t-2} + \cdots\\ &\quad + \alpha(1-\alpha)^{t-1}\overline{X}_1 + (1-\alpha)^t\mu \end{aligned}$$

where the foregoing used the fact that $W_0 = \mu$. Thus we see from Equation 13.6.2 that W_t is a weighted average of all the subgroup averages up to time t, giving weight α to the most recent subgroup and then successively decreasing the weight of earlier subgroup averages by the constant factor $1 - \alpha$, and then giving weight $(1-\alpha)^t$ to the in-control population mean.

The smaller the value of α, the more even the successive weights. For instance, if $\alpha = .1$ then the initial weight is .1 and the successive weights decrease by the factor .9; that is, the weights are .1, .09, .081, .073, .066, .059, and so on. On the other hand, if one chooses, say, $\alpha = .4$, then the successive weights are .4, .24, .144, .087, .052, ... Since the successive weights $\alpha(1-\alpha)^{i-1}$, $i = 1, 2, \ldots$, can be written as

$$\alpha(1-\alpha)^{i-1} = \overline{\alpha} e^{-\beta i}$$

where

$$\overline{\alpha} = \frac{\alpha}{1-\alpha}, \qquad \beta = -\log(1-\alpha)$$

we say that the successively older data values are "exponentially weighted" (see Figure 13.4).

To compute the mean and variance of the W_t, recall that, when in control, the subgroup averages $\overline{X}_i$ are independent normal random variables each having mean μ and variance σ^2/n. Therefore, using Equation 13.6.2, we see that

$$\begin{aligned}
E[W_t] &= \mu[\alpha + \alpha(1-\alpha) + \alpha(1-\alpha)^2 + \cdots + \alpha(1-\alpha)^{t-1} + (1-\alpha)^t] \\
&= \frac{\mu\alpha[1-(1-\alpha)^t]}{1-(1-\alpha)} + \mu(1-\alpha)^t \\
&= \mu
\end{aligned}$$

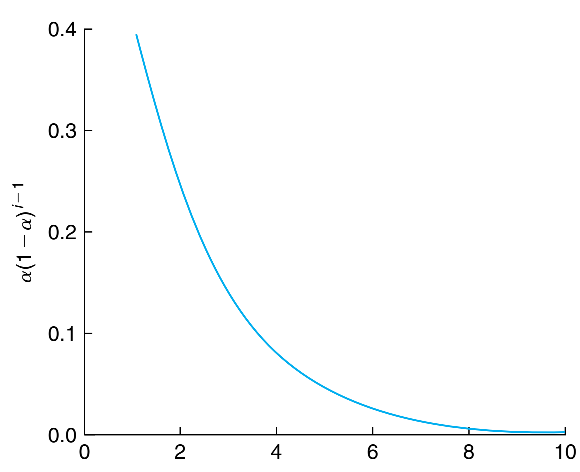

FIGURE 13.4 *Plot of* $\alpha(1-\alpha)^{i-1}$ *when* $\alpha = .4$.

To determine the variance, we again use Equation 13.6.2:

$$\begin{aligned}
\mathrm{Var}(W_t) &= \frac{\sigma^2}{n}\left\{\alpha^2 + [\alpha(1-\alpha)]^2 + [\alpha(1-\alpha)^2]^2 + \cdots + [\alpha(1-\alpha)^{t-1}]^2\right\} \\
&= \frac{\sigma^2}{n}\alpha^2[1+\beta+\beta^2+\cdots+\beta^{t-1}] \qquad \text{where } \beta = (1-\alpha)^2 \\
&= \frac{\sigma^2\alpha^2[1-(1-\alpha)^{2t}]}{n[1-(1-\alpha)^2]} \\
&= \frac{\sigma^2\alpha[1-(1-\alpha)^{2t}]}{n(2-\alpha)}
\end{aligned}$$

Hence, when t is large we see that, provided that the process has remained in control throughout,

$$\begin{aligned}
E[W_t] &= \mu \\
\mathrm{Var}(W_t) &\approx \frac{\sigma^2\alpha}{n(2-\alpha)} \qquad \text{since } (1-\alpha)^{2t} \approx 0
\end{aligned}$$

Thus, the upper and lower control limits for W_t are given by

$$\mathrm{UCL} = \mu + 3\sigma\sqrt{\frac{\alpha}{n(2-\alpha)}}$$

$$\mathrm{LCL} = \mu - 3\sigma\sqrt{\frac{\alpha}{n(2-\alpha)}}$$

Note that the preceding control limits are the same as those in a moving-average control chart with span k (after the initial k values) when

$$\frac{3\sigma}{\sqrt{nk}} = 3\sigma\sqrt{\frac{\alpha}{n(2-\alpha)}}$$

or, equivalently, when

$$k = \frac{2-\alpha}{\alpha} \qquad \text{or} \qquad \alpha = \frac{2}{k+1}$$

EXAMPLE 13.6b A repair shop will send a worker to a caller's home to repair electronic equipment. Upon receiving a request, it dispatches a worker who is instructed to call in when the job is completed. Historical data indicate that the time from when the server is dispatched until he or she calls is a normal random variable with mean 62 minutes and standard deviation 24 minutes. To keep aware of any changes in this distribution, the repair shop plots a standard exponentially weighted moving-average (EWMA) control chart with each data value being the average of 4 successive times, and with a weighting factor of $\alpha = .25$. If the present value of the chart is 60 and the following are the next 16 subgroup averages, what can we conclude?

$$48, 52, 70, 62, 57, 81, 56, 59, 77, 82, 78, 80, 74, 82, 68, 84$$

SOLUTION Starting with $W_0 = 60$, the successive values of $W_1, \ldots, W_{16}$ can be obtained from the formula

$$W_t = .25\overline{X}_t + .75W_{t-1}$$

This gives

$$W_1 = (.25)(48) + (.75)(60) = 57$$
$$W_2 = (.25)(52) + (.75)(57) = 55.75$$
$$W_3 = (.25)(70) + (.75)(55.75) = 59.31$$
$$W_4 = (.25)(62) + (.75)(59.31) = 59.98$$
$$W_5 = (.25)(57) + (.75)(59.98) = 59.24$$
$$W_6 = (.25)(81) + (.75)(59.24) = 64.68$$

and so on, with the following being the values of W_7 through W_{16}:

$$62.50, 61.61, 65.48, 69.60, 71.70, 73.78, 73.83, 75.87, 73.90, 76.43$$

Since

$$3\sqrt{\frac{.25}{1.75}}\frac{24}{\sqrt{4}} = 13.61$$

the control limits of the standard EWMA control chart with weighting factor $\alpha = .25$ are

$$\text{LCL} = 62 - 13.61 = 48.39$$
$$\text{UCL} = 62 + 13.61 = 75.61$$

Thus, the EWMA control chart would have declared the system out of control after determining W_{14} (and also after W_{16}). On the other hand, since a subgroup standard deviation is $\sigma/\sqrt{n} = 12$, it is interesting that no data value differed from $\mu = 62$ by even as much as 2 subgroup standard deviations, and so the standard $\overline{X}$ control chart would not have declared the system out of control. ■

EXAMPLE 13.6c Consider the data of Example 13.6a but now use an exponentially weighted moving-average control chart with $\alpha = 2/9$. This gives rise to the following data set.

t	$\overline{X}_t$	W_t	t	$\overline{X}_t$	W_t
1	9.617728	9.915051	14	9.220009	10.84522
2	10.25437	9.990456	15	11.25206	10.93563
3	9.867195	9.963064	16	10.48662	10.83585
4	10.79338	10.14758	17	9.025091	10.43346
5	10.60699	10.24967	18	9.693386	10.269
6	10.48396	10.30174	19	11.45989	10.53364
*7	13.33961	10.97682	*20	12.44213	10.95775
8	9.462969	10.64041	*21	11.18981	11.00932
9	10.14556	10.53044	*22	11.56674	11.13319
10	11.66342	10.78221	23	9.869849	10.85245
*11	11.55484	10.95391	*24	12.11311	11.13259
*12	11.26203	11.02238	*25	11.48656	11.21125
*13	12.31473	11.30957			

* = *Out of control.*

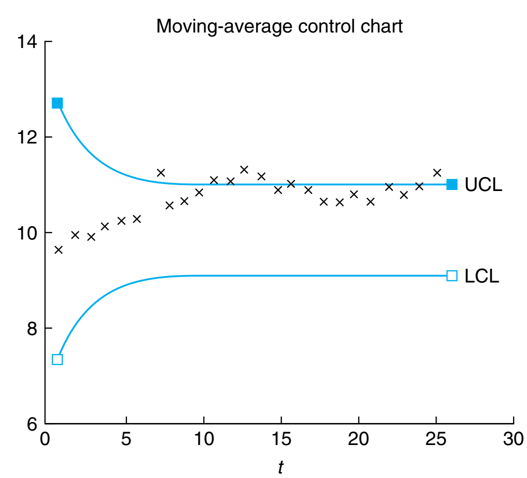

FIGURE 13.5

Since

$$\text{UCL} = 10.94868$$
$$\text{LCL} = 9.051318$$

we see that the process could be declared out of control as early as $t = 7$ (see Figure 13.5). ■

13.6.3 Cumulative Sum Control Charts

The major competitor to the moving-average type of control chart for detecting a small- to moderate-sized change in the mean is the cumulative sum (often reduced to cu-sum) control chart.

Suppose, as before, that $\overline{X}_1, \overline{X}_2, \ldots$ represent successive averages of subgroups of size n and that when the process is in control these random variables have mean μ and standard deviation $\sigma/\sqrt{n}$. Initially, suppose that we are only interested in determining when an increase in the mean value occurs. The (one-sided) cumulative sum control chart for detecting an increase in the mean operates as follows: Choose positive constants d and B, and let

$$Y_j = \overline{X}_j - \mu - d\sigma/\sqrt{n}, \quad j \geq 1$$

Note that when the process is in control, and so $E[\overline{X}_j] = \mu$,

$$E[Y_j] = -d\sigma/\sqrt{n} < 0$$

Now, let

$$S_0 = 0$$
$$S_{j+1} = \max\{S_j + Y_{j+1}, 0\}, \quad j \geq 0$$

The cumulative sum control chart having parameters d and B continually plots S_j, and declares that the mean value has increased at the first j such that

$$S_j > B\sigma/\sqrt{n}$$

To understand the rationale behind this control chart, suppose that we had decided to continually plot the sum of all the random variables Y_i that have been observed so far. That is, suppose we had decided to plot the successive values of P_j, where

$$P_j = \sum_{i=1}^{j} Y_i$$

which can also be written as

$$P_0 = 0$$
$$P_{j+1} = P_j + Y_{j+1}, \quad j \geq 0$$

Now, when the system has always been in control, all of the Y_i have a negative expected value, and thus we would expect their sum to be negative. Hence, if the value of P_j ever became large — say, greater than $B\sigma/\sqrt{n}$ — then this would be strong evidence that the process has gone out of control (by having an increase in the mean value of a produced item). The difficulty, however, is that if the system goes out of control only after some large time, then the value of P_j at that time will most likely be strongly negative (since up to then we would have been summing random variables having a negative mean), and thus it would take a long time for its value to exceed $B\sigma/\sqrt{n}$. Therefore, to keep the sum from becoming very negative while the process is in control, the cumulative sum control chart employs the simple trick of resetting its value to 0 whenever it becomes negative. That is, the quantity S_j is the cumulative sum of all of the Y_i up to time j, with the exception that any time this sum becomes negative its value is reset to 0.

EXAMPLE 13.6d Suppose that the mean and standard deviation of a subgroup average are $\mu = 30$ and $\sigma/\sqrt{n} = 8$, respectively, and consider the cumulative sum control chart with $d = .5, B = 5$. If the first eight subgroup averages are

$$29, 33, 35, 42, 36, 44, 43, 45$$

then the successive values of $Y_j = \overline{X}_j - 30 - 4 = \overline{X}_j - 34$ are

$$Y_1 = -5, Y_2 = -1, Y_3 = 1, Y_4 = 8, Y_5 = 2, Y_6 = 10, Y_7 = 9, Y_8 = 11$$

Therefore,

$$\begin{aligned}
S_1 &= \max\{-5, 0\} = 0 \\
S_2 &= \max\{-1, 0\} = 0 \\
S_3 &= \max\{1, 0\} = 1 \\
S_4 &= \max\{9, 0\} = 9 \\
S_5 &= \max\{11, 0\} = 11 \\
S_6 &= \max\{21, 0\} = 21 \\
S_7 &= \max\{30, 0\} = 30 \\
S_8 &= \max\{41, 0\} = 41
\end{aligned}$$

Since the control limit is

$$B\sigma/\sqrt{n} = 5(8) = 40$$

the cumulative sum chart would declare that the mean has increased after observing the eighth subgroup average. ■

To detect either a positive or a negative change in the mean, we employ two one-sided cumulative sum charts simultaneously. We begin by noting that a decrease in $E[X_i]$ is equivalent to an increase in $E[-X_i]$. Hence, we can detect a decrease in the mean value of an item by running a one-sided cumulative sum chart on the negatives of the subgroup averages. That is, for specified values d and B, not only do we plot the quantities S_j as before, but, in addition, we let

$$W_j = -\overline{X}_j - (-\mu) - d\sigma/\sqrt{n} = \mu - \overline{X}_j - d\sigma/\sqrt{n}$$

and then also plot the values T_j, where

$$T_0 = 0$$
$$T_{j+1} = \max\{T_j + W_{j+1}, 0\}, \qquad j \geq 0$$

The first time that either S_j or T_j exceeds $B\sigma/\sqrt{n}$, the process is said to be out of control.

Summing up, the following steps result in a cumulative sum control chart for detecting a change in the mean value of a produced item: Choose positive constants d and B; use the successive subgroup averages to determine the values of S_j and T_j; declare the process out of control the first time that either exceeds $B\sigma/\sqrt{n}$. Three common choices of the pair of values d and B are $d = .25, B = 8.00$, or $d = .50, B = 4.77$, or $d = 1, B = 2.49$. Any of these choices results in a control rule that has approximately the same false alarm rate as does the $\overline{X}$ control chart that declares the process out of control the first time a subgroup average differs from μ by more than $3\sigma/\sqrt{n}$. As a general rule of thumb, the smaller the change in mean that one wants to guard against, the smaller should be the chosen value of d.

Problems

1. Assume that items produced are supposed to be normally distributed with mean 35 and standard deviation 3. To monitor this process, subgroups of size 5 are sampled. If the following represents the averages of the first 20 subgroups, does it appear that the process was in control?

Subgroup No.	$\overline{X}$	Subgroup No.	$\overline{X}$
1	34.0	6	32.2
2	31.6	7	33.0
3	30.8	8	32.6
4	33.0	9	33.8
5	35.0	10	35.8

(continued)

Subgroup No.	$\overline{X}$	Subgroup No.	$\overline{X}$
11	35.8	16	31.6
12	35.8	17	33.0
13	34.0	18	33.2
14	35.0	19	31.8
15	33.8	20	35.6

2. Suppose that a process is in control with $\mu = 14$ and $\sigma = 2$. An $\overline{X}$ control chart based on subgroups of size 5 is employed. If a shift in the mean of 2.2 units occurs, what is the probability that the next subgroup average will fall outside the control limits? On average, how many subgroups will have to be looked at in order to detect this shift?

3. If Y has a chi-square distribution with $n-1$ degrees of freedom, show that

$$E[\sqrt{Y}] = \sqrt{2}\frac{\Gamma(n/2)}{\Gamma[(n-1)/2]}$$

(*Hint*: Write

$$\begin{aligned} E[\sqrt{Y}] &= \int_0^\infty \sqrt{y} f_{\chi^2_{n-1}}(y)\, dy \\ &= \int_0^\infty \sqrt{y}\, \frac{e^{-y/2} y^{(n-1)/2-1}\, dy}{2^{(n-1)/2}\Gamma\left[\dfrac{(n-1)}{2}\right]} \\ &= \int_0^\infty \frac{e^{-y/2} y^{n/2-1}\, dy}{2^{(n-1)/2}\Gamma\left[\dfrac{(n-1)}{2}\right]} \end{aligned}$$

Now make the transformation $x = y/2$.)

4. Samples of size 5 are taken at regular intervals from a production process, and the values of the sample averages and sample standard deviations are calculated. Suppose that the sum of the $\overline{X}$ and S values for the first 25 samples are given by

$$\sum \overline{X}_i = 357.2, \qquad \sum S_i = 4.88$$

(a) Assuming control, determine the control limits for an $\overline{X}$ control chart.
(b) Suppose that the measurable values of the items produced are supposed to be within the limits $14.3 \pm .45$. Assuming that the process remains in control with a mean and variance that is approximately equal to the estimates derived,

approximately what percentage of the items produced will fall within the specification limits?

5. Determine the revised $\overline{X}$ and S-control limits for the data in Example 13.3a.

6. In Problem 4, determine the control limits for an S-control chart.

7. The following are $\overline{X}$ and S values for 20 subgroups of size 5.

Subgroup	$\overline{X}$	S	Subgroup	$\overline{X}$	S	Subgroup	$\overline{X}$	S
1	33.8	5.1	8	36.1	4.1	15	35.6	4.8
2	37.2	5.4	9	38.2	7.3	16	36.4	4.6
3	40.4	6.1	10	32.4	6.6	17	37.2	6.1
4	39.3	5.5	11	29.7	5.1	18	31.3	5.7
5	41.1	5.2	12	31.6	5.3	19	33.6	5.5
6	40.4	4.8	13	38.4	5.8	20	36.7	4.2
7	35.0	5.0	14	40.2	6.4			

(a) Determine trial control limits for an $\overline{X}$ control chart.
(b) Determine trial control limits for an S-control chart.
(c) Does it appear that the process was in control throughout?
(d) If your answer in part (c) is no, suggest values for upper and lower control limits to be used with succeeding subgroups.
(e) If each item is supposed to have a value within 35 ± 10, what is your estimate of the percentage of items that will fall within this specification?

8. Control charts for $\overline{X}$ and S are maintained on the shear strength of spot welds. After 30 subgroups of size 4, $\sum \overline{X}_i = 12{,}660$ and $\sum S_i = 500$. Assume that the process is in control.

(a) What are the $\overline{X}$ control limits?
(b) What are the S-control limits?
(c) Estimate the standard deviation for the process.
(d) If the minimum specification for this weld is 400 pounds, what percentage of the welds will not meet the minimum specification?

9. Control charts for $\overline{X}$ and S are maintained on resistors (in ohms). The subgroup size is 4. The values of $\overline{X}$ and S are computed for each subgroup. After 20 subgroups, $\sum \overline{X}_i = 8{,}620$ and $\sum S_i = 450$.

(a) Compute the values of the limits for the $\overline{X}$ and S charts.
(b) Estimate the value of σ on the assumption that the process is in statistical control.
(c) If the specification limits are 430 ± 30, what conclusions can you draw regarding the ability of the process to produce items within these specifications?

(d) If μ is increased by 60, what is the probability of a subgroup average falling outside the control limits?

10. The following data refer to the amounts by which the diameters of $\frac{1}{4}$ inch ball bearings differ from $\frac{1}{4}$ inch in units of .001 inches. The subgroup size is $n = 5$.

Subgroup	Data Values				
1	2.5	.5	2.0	−1.2	1.4
2	.2	.3	.5	1.1	1.5
3	1.5	1.3	1.2	−1.0	.7
4	.2	.5	−2.0	.0	−1.3
5	−.2	.1	.3	−.6	.5
6	1.1	−.5	.6	.5	.2
7	1.1	−1.0	−1.2	1.3	.1
8	.2	−1.5	−.5	1.5	.3
9	−2.0	−1.5	1.6	1.4	.1
10	−.5	3.2	−.1	−1.0	−1.5
11	.1	1.5	−.2	.3	2.1
12	.0	−2.0	−.5	.6	−.5
13	−1.0	−.5	−.5	−1.0	.2
14	.5	1.3	−1.2	−.5	−2.7
15	1.1	.8	1.5	−1.5	1.2

(a) Set up trial control limits for $\overline{X}$ and S-control charts.
(b) Does the process appear to have been in control throughout the sampling?
(c) If the answer to part (b) is no, construct revised control limits.

11. Samples of $n = 6$ items are taken from a manufacturing process at regular intervals. A normally distributed quality characteristic is measured, and $\overline{X}$ and S values are calculated for each sample. After 50 subgroups have been analyzed, we have

$$\sum_{i=1}^{50} \overline{X}_i = 970 \qquad \text{and} \qquad \sum_{i=1}^{50} S_i = 85$$

(a) Compute the control limit for the $\overline{X}$ and S-control charts. Assume that all points on both charts plot within the control limits.
(b) If the specification limits are 19 ± 4.0, what are your conclusions regarding the ability of the process to produce items conforming to specifications?

12. The following data present the number of defective bearing and seal assemblies in samples of size 100.

Sample Number	Number of Defectives	Sample Number	Number of Defectives
1	5	11	4
2	2	12	10
3	1	13	0
4	5	14	8
5	9	15	3
6	4	16	6
7	3	17	2
8	3	18	1
9	2	19	6
10	5	20	10

Does it appear that the process was in control throughout? If not, determine revised control limits if possible.

13. The following data represent the results of inspecting all personal computers produced at a given plant during the past 12 days.

Day	Number of Units	Number Defective
1	80	5
2	110	7
3	90	4
4	80	9
5	100	12
6	90	10
7	80	4
8	70	3
9	80	5
10	90	6
11	90	5
12	110	7

Does the process appear to have been in control? Determine control limits for future production.

14. Suppose that when a process is in control each item will be defective with probability .04. Suppose that your control chart calls for taking daily samples of size 500. What is the probability that, if the probability of a defective item should suddenly shift to .08, your control chart would detect this shift on the next sample?

15. The following data represent the number of defective chips produced on the last 15 days: 121, 133, 98, 85, 101, 78, 66, 82, 90, 78, 85, 81, 100, 75, 89. Would you conclude that the process has been in control throughout these 15 days? What control limits would you advise using for future production?

16. Surface defects have been counted on 25 rectangular steel plates, and the data are shown below. Set up a control chart. Does the process producing the plates appear to be in statistical control?

Plate Number	Number of Defects	Plate Number	Number of Defects
1	2	14	10
2	3	15	2
3	4	16	2
4	3	17	6
5	1	18	5
6	2	19	4
7	5	20	6
8	0	21	3
9	2	22	7
10	5	23	0
11	1	24	2
12	7	25	4
13	8		

17. The following data represent 25 successive subgroup averages and moving averages of span size 5 of these subgroup averages. The data are generated by a process that, when in control, produces normally distributed items having mean 30 and variance 40. The subgroups are of size 4. Would you judge that the process has been in control throughout?

$\bar{X}_t$	M_t	$\bar{X}_t$	M_t
35.62938	35.62938	35.80945	32.34106
39.13018	37.37978	30.9136	33.1748
29.45974	34.73976	30.54829	32.47771
32.5872	34.20162	36.39414	33.17019
30.06041	33.37338	27.62703	32.2585
26.54353	31.55621	34.02624	31.90186
37.75199	31.28057	27.81629	31.2824
26.88128	30.76488	26.99926	30.57259
32.4807	30.74358	32.44703	29.78317
26.7449	30.08048	38.53433	31.96463
34.03377	31.57853	28.53698	30.86678
32.93174	30.61448	28.65725	31.03497
32.18547	31.67531		

18. The data shown below give subgroup averages and moving averages of the values from Problem 17. The span of the moving averages is $k = 8$. When in control the subgroup averages are normally distributed with mean 50 and variance 5. What can you conclude?

$\overline{X}_t$	M_t
50.79806	50.79806
46.21413	48.50609
51.85793	49.62337
50.27771	49.78696
53.81512	50.59259
50.67635	50.60655
51.39083	50.71859
51.65246	50.83533
52.15607	51.00508
54.57523	52.05022
53.08497	52.2036
55.02968	52.79759
54.25338	52.85237
50.48405	52.82834
50.34928	52.69814
50.86896	52.6002
52.03695	52.58531
53.23255	52.41748
48.12588	51.79759
52.23154	51.44783

19. Redo Problem 17 by employing an exponential weighted moving average control chart with $\alpha = \frac{1}{3}$.

20. Analyze the data of Problem 18 with an exponential weighted moving-average control chart having $\alpha = \frac{2}{9}$.

21. Explain why a moving-average control chart with span size k must use different control limits for the first $k - 1$ moving averages, whereas an exponentially weighted moving-average control chart can use the same control limits throughout. [*Hint*: Argue that $\text{Var}(M_t)$ decreases in t, whereas $\text{Var}(W_t)$ increases, and explain why this is relevant.]

22. Repeat Problem 17, this time using a cumulative sum control chart with

(a) $d = .25, B = 8$;
(b) $d = .5, B = 4.77$.

23. Repeat Problem 18, this time using a cumulative sum control chart with $d = 1$ and $B = 2.49$.