CHAPTER 15

Quality Control

Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.

H.G.Wells (1866–1946)

CONTENTS

We introduce control charts, which are used to determine when a process that produces items has gone out of control. This is done both where the measurable value of an item is a continuous variable and where it is a discrete variable. In this latter case, each item has two possible values, with the value indicating whether the item is acceptable or not.

Introductory Statistics, DOI: 10.1016/B978-0-12-374388-6.00015-6

15.1 INTRODUCTION

Almost every industrial system—whether it involves the manufacturing of products or the servicing of customers—results in some random variation in the items it processes. That is, no matter how stringently the system is being controlled, there is always going to be some variation in the items processed. For instance, the successive items produced by a manufacturing process will not all be identical, and the successive times that it takes to service customers will often be different, even when the underlying operation is performing as required. This type of variation, called *chance variation*, is considered inherent to the system. However, there is another type of variation that sometimes appears. This variation, far from being inherent to the system, is due to some assignable cause, and it usually has an adverse effect on the quality of the industrial operation. For instance, in a manufacturing context, this latter variation may be caused by a faulty machine setting, or by poor quality of the raw materials being employed, or by incorrect software, or by human error, or by any of a large number of possibilities. When the only variation present is due to chance and not to any assignable cause, we say that the process is *in control*; a key problem in quality control is to determine when a process is in control and when it is out of control.

In this chapter we study control charts, which can be used to indicate when a process has gone out of control. The types of control charts we consider are determined by two numbers, called the *upper control limit* (*UCL*) and the *lower control limit* (*LCL*). To utilize these charts, first we divide the data generated by the industrial concern into subgroups. Then we compute the subgroup averages, and when one of these does not fall within the upper and lower control limits, we conclude that the process is out of control.

In Sec. 15.2 we suppose that the successive items processed have measurable characteristics—this could refer to their quality level in a manufacturing context or to their service time in a service industry—whose mean and variance are known when the process is operating in control. We show how to construct control charts that are useful for detecting a change in the mean of the in-control distribution. In Sec. 15.3 we construct a control chart for situations in which each item, rather than having some measurable characteristic, is classified as being either satisfactory or unsatisfactory. In Secs. 15.4 and 15.5 we introduce two types of control charts that are particularly effective in detecting small shifts in the mean value of a process: *Exponentially Weighted Moving-Average control charts* in Sec. 15.4, and *Cumulative Sum control charts* in Sec. 15.5.

15.2 THE $\overline{X}$ CONTROL CHART FOR DETECTING A SHIFT IN THE MEAN

Suppose that when an industrial system is in control, the successive items it processes have measurable values that are independent, normal random variables with mean μ and variance σ^2. However, due to unforeseeable circumstances, suppose that the system may go out of control and, as a result, begin to process items having values from a different distribution. We want to be able to recognize when this occurs so as to stop the system, learn what is wrong, and fix it.

Let the measurable values of the successive items processed by the system be denoted by $X_1, X_2, \ldots$. In our attempt to determine when the process goes out of control, we will find it convenient first to break up the data into subgroups of some fixed size—call this size n. Among other things, this value of n should be chosen so as to yield uniformity of data values within individual subgroups. That is, we should attempt to choose n so that it is reasonable, when a shift in distribution occurs, that it will occur between and not within subgroups. Thus, in practice n is often chosen so that all the data within a subgroup relate to items processed on the same day, or on the same shift, or with the same settings, etc.

Let $\overline{X}_i$, $i = 1, 2, \ldots$, denote the average of the ith subgroup. Since, when in control, all the data values are normal with mean μ and variance σ^2, it follows that $\overline{X}_i$, the sample mean of n of them, is normally distributed with mean and variance given, respectively, by

$$E\left[\overline{X}_i\right] = \mu$$

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

Hence, it follows that when the process is in control,

$$Z = \frac{\overline{X}_i - \mu}{\sqrt{\sigma^2/n}}$$

is a standard normal random variable. That is, if the process remains in control throughout the processing of subgroup i, then $\sqrt{n}(\overline{X}_i - \mu)/\sigma$ has a standard normal distribution. Now, a standard normal random variable Z will almost always be between -3 and $+3$. Indeed, from Table 6.1 we see that $P\{-3 < Z < 3\} = 0.9973$. Hence, if the process remains in control throughout the processing of subgroup i, then we would certainly expect that

$$-3 < \frac{\sqrt{n}\left(\overline{X}_i - \mu\right)}{\sigma} < 3$$

or, equivalently, that

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

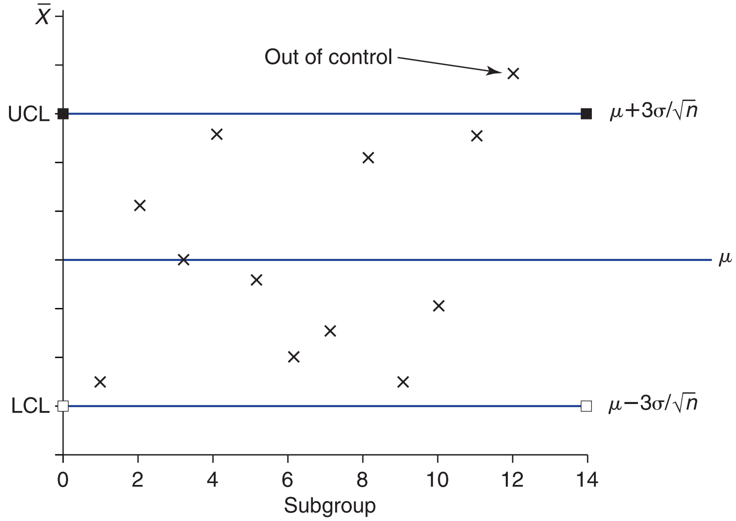

FIGURE 15.1

Control chart for $\overline{X}$; n = size of subgroup.

The values

$$\text{LCL} = \mu - \frac{3\sigma}{\sqrt{n}}$$
$$\text{UCL} = \mu + \frac{3\sigma}{\sqrt{n}}$$

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

The $\overline{X}$ control chart, which is primarily designed to detect a change in the average value of an item processed, is obtained by plotting the successive subgroup averages $\overline{X}_i$ and declaring that the process is out of control the first time that $\overline{X}_i$ does not fall between LCL and UCL (Fig. 15.1).

Since an $\overline{X}$ control chart will declare a process out of control only when a subgroup average falls outside the control limits, it is important that the subgroups be chosen so that it becomes highly likely that any shift in distribution that occurs is between subgroups. This is so because it is easier to detect a shift in a subgroup having all, rather than only some of, its values out of control.

■ Example 15.1

The time it takes a computer servicing firm to install a hard disk along with some sophisticated software for its use is a random variable having mean 25 minutes and standard deviation 6 minutes. The company has two employees who work on this operation. To monitor the efficiency of these employees, the company has plotted the successive average times that it takes them to complete four jobs. The odd-number subgroups refer to the first employee, and the even-number subgroups refer to the second. Suppose the first 20 of these successive subgroup averages are as follows (see Fig. 15.2 for a plot):

Subgroup	$\overline{X}$	Subgroup	$\overline{X}$	Subgroup	$\overline{X}$	Subgroup	$\overline{X}$
1	23.6	6	24.6	11	29.4	16	32.8
2	20.8	7	22.6	12	27.8	17	23.3
3	25.5	8	24.4	13	26.8	18	30.5
4	26.2	9	24.7	14	27.2	29	25.3
5	23.3	10	26.0	15	24.0	20	34.1

What conclusions can be drawn?

Solution

Since the subgroup size is 4 and since the successive data values have mean $\mu = 25$ and standard deviation $\sigma = 6$ when the process is in control, it follows that the control limits are given by

$$\mathrm{LCL} = 25 - \frac{18}{\sqrt{4}} = 16 \quad \mathrm{UCL} = 25 + \frac{18}{\sqrt{4}} = 34$$

Since the average of subgroup 20 is greater than UCL, it appears that the system is no longer in control. Indeed, since all the last six even-number subgroup averages are larger than the in-control mean of 25 (with the final three being significantly larger), it seems probable that the second employee has been out of control for some time. ■

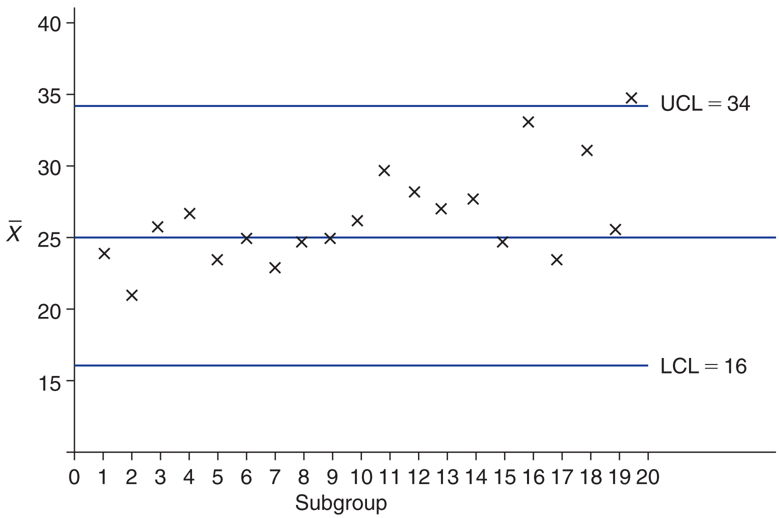

FIGURE 15.2

Control chart for data of Example 15.1.

We have assumed here that when the process is in control, the underlying distribution of an item's measurable characteristic is a normal distribution; and, in fact, this is often the case in manufacturing processes. In addition, as long as this distribution is somewhat close to being normal, the subgroup averages will, because of the central limit theorem, be approximately normal and so would be unlikely to differ from their mean by more than 3 standard deviations. As a result, by utilizing subgroup averages it is not necessary for us to know the entire in-control distribution, only its mean and variance. This is the most important reason for utilizing subgroup averages rather than individual data points. Typical values of the subgroup size are 4, 5, and 6. The reason for this is that if one uses a smaller subgroup size, then the approximate normality of the subgroup average might come into question. On the other hand, since one cannot detect an out-of-control system without processing all the items of at least one subgroup, the subgroup size should not be too large.

Because the $\overline{X}$ control chart was first suggested by Walter Shewhart, it is often referred to as the *Shewhart control chart*. In essence, it works by using each subgroup to test the statistical hypothesis

$$\mathrm{H}_0\text{: mean} = \mu \text{ against the alternative } \mathrm{H}_1\text{: mean} \neq \mu$$

at the significance level $\alpha = P\{|Z| > 3\} = 0.0027$. Whenever the null hypothesis is rejected, the process is declared to be out of control.

Although we usually talk about control charts in the context of a manufacturing process, they can be utilized in a variety of situations, as indicated by the next example.

■ Example 15.2

Consider a small video rental store in which the daily rentals on each of the weekdays Monday through Thursday have mean 52 and standard deviation 10. If the numbers of videos rented daily (Monday to Thursday) in the past week are

$$32, 38, 28, 30$$

can we conclude that a change in distribution has occurred?

Solution

Let a subgroup consist of the number of rentals on the four specified weekdays of each week. Since the in-control mean and standard deviation are $\mu = 52$ and $\sigma = 10$, respectively, it follows that the control limits are

$$\mathrm{LCL} = 52 - \frac{3(10)}{\sqrt{4}} = 37$$

$$\mathrm{UCL} = 52 + \frac{3(10)}{\sqrt{4}} = 67$$

Since the most recent subgroup average is $(32 + 38 + 28 + 30)/4 = 32$, which is outside these limits, we can declare that the average number of videos rented daily is no longer equal to 52.

At this juncture the store's manager should try to discover (1) the cause of the change in mean and (2) whether it appears to be a temporary or a more permanent change. For instance, he or she might discover that there were some particularly interesting television programs in the past week, such as a World Series or the Olympics, or a political convention, which would lead to the belief that the change is of a short-term nature. Or else it might be discovered that the change in mean was caused by a new competitor in the neighborhood, and this might result in a more permanent change. ■

Sometimes not all the measurable values of the items produced are noted, only those of a randomly selected subset of items. When this is the situation, it is natural to let subgroups consist of items that are produced at roughly the same time.

PROBLEMS

1. When it is in control, a process produces items having mean 100 and standard deviation 10. Determine upper and lower control limits for the subgroup average when the subgroup size is
   (a) 4 (b) 5 (c) 6 (d) 10
2. When a process is working properly, it produces items that have mean 35 and standard deviation 4. To monitor this process, subgroups of size 4 are sampled. If the following table represents the averages of the first 20 subgroups, does it appear that the process was in control throughout?

Subgroup	$\overline{X}$	Subgroup	$\overline{X}$
1	31.2	11	36.4
2	38.4	12	31.1
3	35.0	13	32.3
4	33.3	14	37.8
5	34.7	15	36.6
6	31.1	16	40.4
7	35.8	17	41.2
8	34.4	18	35.9
9	37.1	19	40.4
10	34.2	20	32.5

3. When a manufacturing process is in control it produces wire whose diameter has mean 80 with standard deviation 10 (in units of 1/10,000 inch). The following data represent the sample mean of subgroups of size 5:

$$85, 88, 90, 77, 79, 83, 90, 75, 94, 80, 84, 86, 88$$

Does the process appear to have been in control?

4. A control chart is maintained on the time it takes workers to perform a certain task. The time that it should take is normal with mean 26 minutes and standard deviation 4.2 minutes. The following are $\overline{X}$ values for 10 subgroups of size 4:

$$28.2, 28.4, 31.1, 27.3, 33.2, 31.4, 27.9, 30.4, 31.3, 30.4$$

(a) Determine the upper and lower control limits.
(b) Does it appear that the process was in control?

5. When a seam welding process is working correctly, the distance from the weld to the center of the seam is normally distributed with mean 0 and standard deviation 0.005 inches. If the following values are the average of the distances from the weld to the center in eight subgroups of size 5, does it appear that the process was in control during their processing?

Subgroup	Average distance
1	0.0023
2	−0.0012
3	−0.0015
4	0.0031
5	0.0038
6	0.0051
7	0.0022
8	−0.0033

6. Prior to 1995 the number of murders committed yearly in the United States per 100,000 population was normally distributed with mean 9.0 and standard deviation 1.1. The following are the rates from 1995 to 2002:

$$8.4, 7.4, 6.8, 6.3, 5.7, 5.5, 5.6, 5.4$$

Can we conclude that the murder rate has changed from its historical value? Use subgroups of size 2.

15.2.1 When the Mean and Variance Are Unknown

If one is just starting up a control chart and does not have reliable historical data, then the mean μ and the standard deviation σ have to be estimated. To do so, one employs k of the subgroups, where if possible k should be such that $k \geq 20$ and $nk \geq 100$, where n is the size of a subgroup. If $\overline{X}_i$ is the average of subgroup i, then μ can be estimated by $\overline{\overline{X}}$, the average of the subgroup averages. That is,

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

Since $\overline{\overline{X}}$ is the average of all the nk data values, it is the "natural" estimator of the population mean μ.

To estimate σ, let S_i be the sample standard deviation of the data of subgroup $i, i = 1, \ldots, k$, and let $\overline{S}$ be the average of these subgroup standard deviations. That is,

$$\overline{S} = \frac{S_1 + \cdots + S_k}{k}$$

Because the expected value of $\overline{S}$ is not equal to σ, we divide it by a constant $c(n)$ that depends on the subgroup sample size n, to obtain an estimator whose mean is σ. That is, we use the estimator $\overline{S}/c(n)$, which is such that

$$E\left[\overline{S}/c(n)\right] = \sigma$$

The values of $c(n)$ for n ranging between 3 and 9 are presented in the following table.

Values of $c(n)$
$c(3) = 0.8862266$
$c(4) = 0.9213181$
$c(5) = 0.9399851$
$c(6) = 0.9515332$
$c(7) = 0.9593684$
$c(8) = 0.9650309$
$c(9) = 0.9693103$

The preceding estimators of μ and σ make use of all of the k subgroups and are thus reasonable only if the system has remained in control throughout their processing. To verify this, we compute the control limits based on these estimators:

$$\text{LCL} = \overline{\overline{X}} - \frac{3\overline{S}}{\sqrt{n}c(n)}$$

$$\text{UCL} = \overline{\overline{X}} + \frac{3\overline{S}}{\sqrt{n}c(n)}$$

Then we check that each of the k subgroup averages falls within these limits. If any of them fall outside, then we must decide if it is reasonable to suppose that the system was temporarily out of control when the items in that subgroup were processed. If such is the case, or if the reason for the out-of-control data values is discovered and fixed, then those subgroup averages should be removed, additional ones obtained, and the estimates of μ and σ recomputed. This continues until all the subgroup averages fall within the estimated control limits. Of course, if many subgroup averages along the way do not fall within the limits and have to be removed, then it is clear that no control has yet been established.

■ Example 15.3

Consider a new plant set up to manufacture automobile air conditioners that release only minimal amounts of harmful chlorofluorocarbons. After each air conditioner is produced, it is checked to determine the amount of chlorofluorocarbons, suitably measured, that it releases in a 1-hour run period. The following data give the sample averages and sample standard deviations of 50 air conditioners that have been divided into 10 subgroups of size 5:

i	$\overline{X}_i$	S_i
1	30.1	1.22
2	29.7	1.40
3	31.2	0.81
4	29.9	1.10
5	30.3	0.93
6	30.2	0.82
7	31.0	1.54
8	31.4	1.58
9	30.9	1.26
10	32.0	1.60

The values of the estimators of μ and σ are

$$\overline{\overline{X}} = 30.670, \qquad \frac{\overline{S}}{c(5)} = \frac{1.226}{0.9399851} = 1.304$$

Since $3(1.304)/\sqrt{5} = 1.750$, it follows that the estimated control limits are

$$\text{LCL} = 30.670 - 1.750 = 28.920$$
$$\text{UCL} = 30.670 + 1.750 = 32.420$$

Because all of the subgroup averages fall within these limits, we can suppose that the process is in control with $\mu = 30.670$ and $\sigma = 1.304$.

Suppose now that it is required that at least 99 percent of all air conditioners release no more than 33.40 units of chlorofluorocarbons per hour. Assuming that the production process remains in control with the values of its mean and standard deviation exactly as estimated in the preceding, will this requirement be met? To answer this, note that if X, the amount of chlorofluorocarbon released by an air conditioner in an hour, is a normal random variable with mean 30.670 and standard deviation 1.304, then

$$\begin{aligned} P\{X > 33.40\} &= P\left\{\frac{X - 30.670}{1.304} > \frac{33.40 - 30.670}{1.304}\right\} \\ &= P\{Z > 2.094\} \\ &= 0.018 \end{aligned}$$

Hence, 1.8 percent of the air conditioners will have an hourly release above 33.4, and so the requirement that at least 99 percent of them satisfy this condition will not be met. ■

Remark *The estimator $\overline{\overline{X}}$ is the average of all the nk data values and is thus the obvious estimator of the mean value μ. On the other hand, it may not be immediately apparent why we do not utilize the sample standard deviation of all the nk data values, namely, $\sqrt{S^2}$, where*

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

to obtain an estimator of σ. This is not done because the system may not have been in control throughout the processing of the first k subgroups, and thus this latter estimator may be very far from the actual in-control value of σ. On the other hand, even if there had been a change in the mean value during the processing of the k subgroups, provided that the standard deviation remained the same, the estimator $\overline{S}/c(n)$ will still be a good estimator of σ since it only requires that the data values in each subgroup have the same mean (which may differ for the different subgroups).

15.2.2 S Control Charts

The $\overline{X}$ control chart is designed to pick up any change in the population mean. In cases where we are also concerned about possible changes in the population standard deviation, we can utilize an S control chart.

As in the previous sections, suppose that the items produced by a process that is in control have values that are normally distributed with mean μ and standard deviation σ. If S_i is the sample standard deviation for subgroup i, then, as noted in Sec. 15.2.1,

$$E\left[\frac{S_i}{c(n)}\right] = \sigma$$

implying that

$$E[S_i] = c(n)\sigma$$

where n is the subgroup size. In addition, using the identity

$$\text{Var}(S_i) = E\left[S_i^2\right] - (E[S_i])^2$$

we obtain from the preceding, and the fact that the expected value of the sample variance is the population variance, that

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

That is, provided the process is in control, S_i is a random variable with mean $c(n)\sigma$ and standard deviation $\sigma\sqrt{1-c^2(n)}$. Hence, since a random variable is unlikely to be more than 3 standard deviations from its mean, it is reasonable to set the lower and upper control limits of the S control chart as follows:

$$\begin{aligned}\text{LCL} &= c(n)\sigma - 3\sigma\sqrt{1-c^2(n)} \\ \text{UCL} &= c(n)\sigma + 3\sigma\sqrt{1-c^2(n)}\end{aligned}$$

The successive values of S_i should be plotted to make certain they fall within the 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 $\overline{S}/c(n)$, where $\overline{S}$ is the average of k subgroup standard deviations. Using the foregoing, the estimated control limits become

$$\text{LCL} = \overline{S}\left[1 - 3\sqrt{\frac{1}{c^2(n)} - 1}\right]$$
$$\text{UCL} = \overline{S}\left[1 + 3\sqrt{\frac{1}{c^2(n)} - 1}\right]$$

As when an $\overline{X}$ control chart is begun, it should then be checked that the k subgroup standard deviations fall within these control limits. Those that fall outside should be discarded, and the estimate of σ should be recomputed (possibly using additional data).

■ Example 15.4

The following are the subgroup averages and subgroup standard deviations (in minutes) of 20 subgroups of size 5 from a newly established manufacturing facility that produces steel shafts. The data refer to the production time.

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

Since

$$\overline{\overline{X}} = 35.94, \quad \overline{S} = 4.35, \quad c(5) = 0.9400$$

we obtain that the preliminary lower and upper control limits for $\overline{X}$ and S are as follows:

$$\text{LCL}(\overline{X}) = 35.94 - \frac{3(4.35)}{0.94\sqrt{5}} = 29.731$$
$$\text{UCL}(\overline{X}) = 35.94 + \frac{3(4.35)}{0.94\sqrt{5}} = 42.149$$
$$\text{LCL}(S) = 4.35\left[1 - 3\sqrt{\frac{1}{(0.94)^2 - 1}}\right] = -0.386$$
$$\text{UCL}(S) = 4.35\left[1 + 3\sqrt{\frac{1}{(0.94)^2} - 1}\right] = 9.087$$

The control charts for and $\overline{X}$ and S, with the preceding control limits, are shown in Figs. 15.3 and 15.4. Since the sample means of subgroups 10 and 15 fall outside the control limits, these subgroups should be eliminated and the control limits recomputed. We leave the necessary computations as an exercise. ■

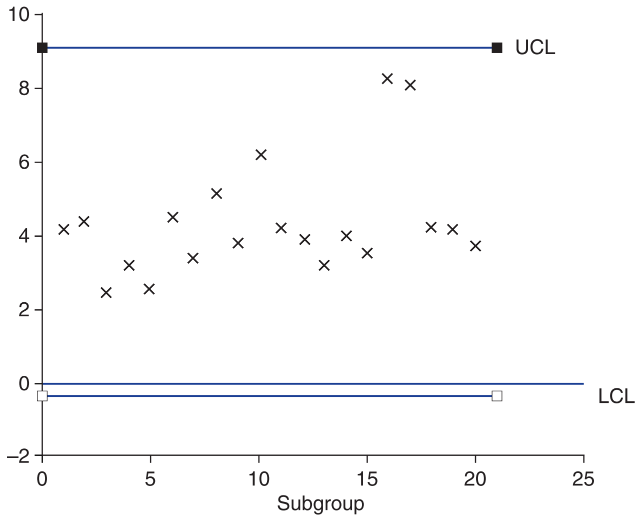

FIGURE 15.3

Control chart for S.

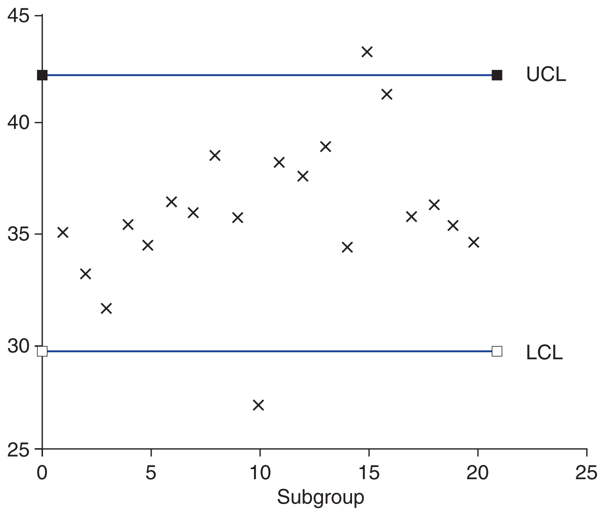

FIGURE 15.4

Control chart for $\overline{X}$.

PROBLEMS

1. Data from 150 items produced by a newly started production process are collected. The data are divided into 25 subgroups of size 6 each, and the sample mean and sample standard deviation of each subgroup are determined. Suppose that

$$\sum_{i=1}^{25} \overline{X}_i = 357.3, \quad \sum_{i=1}^{25} S_i = 5.44$$

   (a) Using the preceding data, determine the trial control limits for an $\overline{X}$ control chart. Assume that each of the 25 subgroup averages falls within these limits.
   (b) Suppose that the requirement of a produced item is that it should have a measurable value in the range 14.3 ± 0.50. Assuming that the process remains in control with a mean and variance equal to the estimates derived from the preceding data, what percentage of items will have values within these specification limits?

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

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

   (a) Determine trial control limits for an $\overline{X}$ control chart.
   (b) Does it appear that the process was in control throughout?
   (c) Estimate the percentage of the produced items that have values between 25 and 45.

3. The following are the successive subgroup averages and sample standard deviations of data relating to an electrical characteristic (measured in decibels) of ceramic strips. The subgroups are of size 4.

$\overline{X}$: 16.1, 15.7, 16.6, 16.0, 14.7, 15.8, 16.4, 14.5, 15.8, 17.2

S: 2.7, 2.9, 2.2, 1.0, 1.3, 2.6, 3.1, 2.5, 5.3, 4.4

(a) Use the preceding to estimate the population mean and standard deviation.
(b) Does the process appear to have been in control?

4. Historical data indicate that the amount of time it should take to perform a certain job has mean 26 minutes and standard deviation 8.3 minutes.
(a) Determine the control limits for an $\overline{X}$ control chart using subgroups of size 5.
(b) Determine the control limits for an S control chart using subgroups of size 5.

5. The following data refer to the amounts by which the diameters of pieces of wire, measured in units of 0.001 inches, exceed a specified value.

Subgroup	Data values				
1	2.5	0.5	2.0	−1.2	1.4
2	0.2	0.3	0.5	1.1	1.5
3	1.5	1.3	1.2	−1.0	0.7
4	0.2	0.5	−2.0	0.0	−1.3
5	−0.2	0.1	0.3	−0.6	0.5
6	1.1	−0.5	0.6	0.5	0.2
7	1.1	−1.0	−1.2	1.3	0.1
8	0.2	−1.5	−0.5	1.5	0.3
9	−2.0	−1.5	1.6	1.4	0.1
10	−0.5	3.2	−0.1	−1.0	−1.5
11	0.1	1.5	−0.2	0.3	2.1
12	0.0	−2.0	−0.5	0.6	−0.5
13	−1.0	−0.5	−0.5	−1.0	0.2
14	0.5	1.3	−1.2	−0.5	−2.7
15	1.1	0.8	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?
(c) If the answer to (b) is no, construct revised control limits.

6. The following are $\overline{X}$ and S values for the initial 10 subgroups, each of size 5, of a new production facility:

$\overline{X}$: 20.2, 28.4, 31.1, 27.3, 33.2, 31.4, 27.9, 30.4, 31.3, 30.4

S: 7.2, 2.8, 3.4, 4.1, 4.0, 3.3, 4.5, 3.0, 2.7, 2.1

(a) Use the preceding to determine trial control limits for an $\overline{X}$ control chart and an S control chart.

(b) Does the process appear to have been in control throughout?
(c) If no additional data are available, what control limits would you suggest be used for future data?

7. Complete Example 15.4.
8. For Prob. 1, determine the trial control limits of an S control chart.
9. For Prob. 2, determine the trial control limits of an S control chart.

15.3 CONTROL CHARTS FOR FRACTION DEFECTIVE

The $\overline{X}$ control chart can be used when the data are measurements whose values can vary continuously over a region. However, in some situations the items processed have values that are classified as either acceptable or unacceptable. For instance, a manufactured item can be classified as defective or not; or the service of a customer could be rated (by the customer) as acceptable or not. In this section we show how to construct control charts for these situations.

Let us suppose that when the system is in control, each item processed 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 the system has been in control, X will be a binomial random variable with parameters n and p; and thus

$$E[X] = np$$
$$\text{Var}(X) = np(1-p)$$

Hence, when the system is in control, the number of defectives in a subgroup of size n should be, with high probability, between the lower and upper limits

$$\text{LCL} = np - 3\sqrt{np(1-p)}$$
$$\text{UCL} = np + 3\sqrt{np(1-p)}$$

The subgroup size n is usually much larger than the typical values of between 4 and 10 used in $\overline{X}$ control charts. The main reason is that if p is small (as is typically the case) 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 to detect this out-of-control situation than it would if n were chosen so that np was not too small. A secondary reason for using a larger value of n is that when np is of moderate size, X will have an approximately normal distribution, and so when in control, each subgroup statistic will fall within the control limits with probability approximately equal to $1 - 0.0027 = 0.9973$.

■ Example 15.5

Successive samples of 200 screws are drawn from the production of an automatic screw machine, with each screw being rated as acceptable or defective. Suppose that it is known from historical data that when the process is in control, each screw is independently defective with probability 0.07. If the following values represent the number of defective screws in each of 20 samples, would the process have been declared out of control at any time during the collection of these samples?

Subgroup	Defectives	Subgroup	Defectives
1	23	11	4
2	22	12	13
3	12	13	17
4	13	14	5
5	15	15	9
6	11	16	5
7	25	17	19
8	16	18	7
9	23	19	22
10	14	20	17

Solution

Since $n = 200$ and $p = 0.07$, we have

$$np = 14 \quad 3\sqrt{np(1-p)} = 10.825$$

and so

$$\text{LCL} = 14 - 10.825 = 3.175$$

$$\text{UCL} = 14 + 10.825 = 24.825$$

Since the number of defectives in subgroup 7 falls outside the range from LCL to UCL, the process would have been declared out of control at that point. ■

Remark *Note that we are attempting to detect any change in quality even when this change results in a quality improvement. 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 recognize any change in quality, for either better or worse, so as 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 determine the reason for this improvement (what are we doing right?).*

PROBLEMS

1. The following data represent the number of defective bearing and seal assemblies in samples of size 200.

Sample number	Number of defectives	Sample number	Number of defectives
1	7	11	4
2	3	12	10
3	2	13	0
4	6	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

Suppose that when the process is in control, each assembly is defective with probability 0.03. Does it appear that the process was in control throughout?

2. Suppose that when a process is in control, each item produced is defective with probability 0.04. If the control chart takes daily samples of size 500, compute the upper and lower control limits.

3. Historically, 4 percent of fiber containers are defective due to contamination from gluing. The following data represent the number of defective containers in successive samples of size 100.

$$3, 5, 1, 0, 4, 7, 8, 9, 5, 7, 1, 3, 0, 5, 3, 6, 4, 8, 3, 6$$

(a) Does the process appear to have been in control?
(b) What are the control limits?

15.4 EXPONENTIALLY WEIGHTED MOVING-AVERAGE CONTROL CHARTS

Whereas the $\overline{X}$ control chart is powerful for detecting large changes in the mean value that may occur temporarily, it is not as efficient for detecting a smaller change that tends to persist. For instance, consider a process whose subgroup averages have mean $\mu = 100$ and standard deviation $\sigma/\sqrt{n} = 2$ when it is in control.

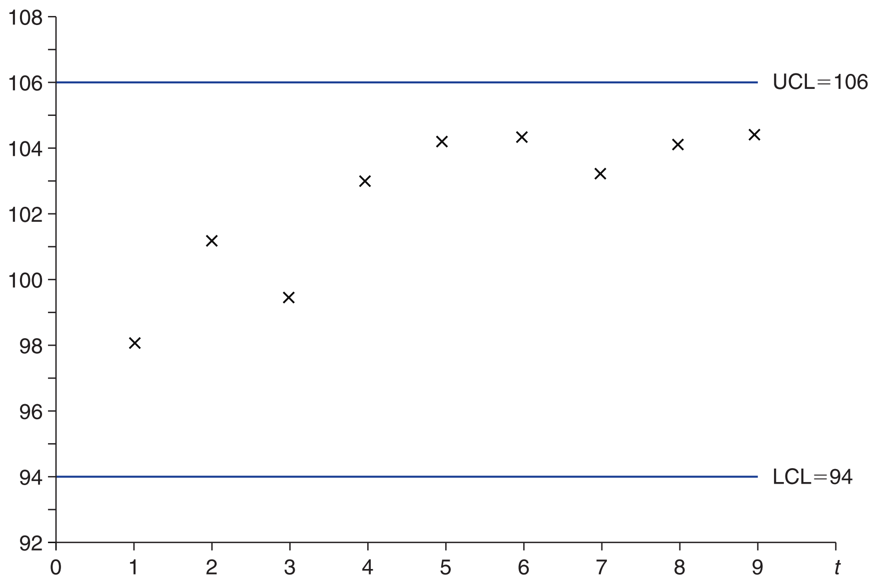

FIGURE 15.5
Successive subgroup averages.

The $\overline{X}$ control chart would thus have control limits 100 ± 6. Suppose that the successive subgroup averages are (see Fig. 15.5):

$$98, 101.2, 99.4, 103, 104.2, 104.3, 103.2, 104.1, 104.4$$

Although it is fairly obvious that the process is out of control (since, among other things, four of the last five subgroup averages exceed μ by more than 2 subgroup standard deviations), the $\overline{X}$ control chart would not have detected this result. The foregoing illustrates a fundamental weakness of the $\overline{X}$ control chart, namely that it considers each subgroup average in isolation and not in relation to the values of nearby subgroup averages. Because of this, it requires a subgroup average to be far away from μ before it declares the process out of control.

We will now consider a more sophisticated control chart, known as an *exponential weighted moving-average chart*, that considers subgroup averages in relation to those around it. To begin, suppose as before that when the process is in control it produces items whose values are normally distributed with mean μ and standard deviation σ. Let the subgroup size be n, and let $\overline{X}_i$ denote the average of the values in subgroup i, $i \geq 1$. Now, let β denote a constant value between 0 and 1, and define the sequence of values $W_0, W_1, \ldots$ as follows:

$$W_0 = \mu$$
$$W_t = \beta \overline{X}_t + (1-\beta) W_{t-1}, \quad t = 1, 2, \ldots$$

That is, the initial value of the W sequence is μ, and each successive value is a weighted average of the next subgroup average and the previous W value. The sequence of values $W_0, W_1, W_2, \ldots$ is called an *exponentially weighted moving average (EWMA)*. (It is also sometimes called a *geometrically weighted moving average.*) It is called this because it can be shown that W_t, the value of the moving average at time t, can be expressed as

$$\begin{aligned} W_t = \beta\overline{X}_t + \beta(1-\beta)\overline{X}_{t-1} + \beta(1-\beta)^2\overline{X}_{t-2} \\ + \cdots + \beta(1-\beta)^{t-1}\overline{X}_1 + \beta(1-\beta)^t\mu \end{aligned}$$

In other words, W_t is a weighted average of all the subgroup averages up to time t, giving weight β to the most recent value and then successively decreasing the weights of earlier subgroup averages by the factor $1-\beta$, and finally giving weight $\beta(1-\beta)^t$ to the in-control mean μ.

The smaller the value chosen for β, the more alike are the successive weights. For instance, if we take $\beta = 0.2$, then the successive weights are 0.2, 0.16, 0.128, 0.1024, 0.08192, and so on; whereas if we take $\beta = 0.9$, then the successive weights are 0.9, 0.09, 0.009, and so on.

Since W_t can be expressed as the sum of independent normal random variables, it too is normal. Its expected value is

$$E[W_t] = \mu$$

and, for t of at least moderate size, its standard deviation is approximately given by

$$\mathrm{SD}(W_t) = \sqrt{\frac{\beta}{2-\beta}}\frac{\sigma}{\sqrt{n}}$$

The control chart that continually plots W_t and declares that the process is out of control the first time that it falls outside the control limits

$$\begin{aligned} \mathrm{LCL} &= \mu - 3\sqrt{\frac{\beta}{2-\beta}}\frac{\sigma}{\sqrt{n}} \\ \mathrm{UCL} &= \mu + 3\sqrt{\frac{\beta}{2-\beta}}\frac{\sigma}{\sqrt{n}} \end{aligned}$$

is called the *standard exponential weighted moving-average* control chart with weighting factor β.

■ Example 15.6

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 EWMA control chart, with each data value being the average of 4 successive times and with a weighting factor $\beta = 0.25$. 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

$$W_t = 0.25\overline{X}_t + 0.75W_{t-1}$$

This gives,

$$\begin{aligned}
W_1 &= (0.25)(48) + (0.75)(60) = 57 \\
W_2 &= (0.25)(52) + (0.75)(57) = 55.75 \\
W_3 &= (0.25)(70) + (0.75)(55.75) = 59.31 \\
W_4 &= (0.25)(62) + (0.75)(59.31) = 59.98 \\
W_5 &= (0.25)(57) + (0.75)(59.98) = 59.24 \\
W_6 &= (0.25)(81) + (0.75)(59.24) = 64.68
\end{aligned}$$

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

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

A graph of the successive values of the moving average is given in Fig. 15.6. Since

$$3\sqrt{\frac{0.25}{1.75}}\frac{24}{\sqrt{4}} = 13.61$$

the control limits of the standard EWMA control chart with weighting factor $\beta = 0.25$ are

$$\begin{aligned}
\text{LCL} &= 62 - 13.61 = 48.39 \\
\text{UCL} &= 62 + 13.61 = 75.61
\end{aligned}$$

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 as much as 2 subgroup standard deviations, and so the standard $\overline{X}$ control chart would not have declared the system out of control. ■

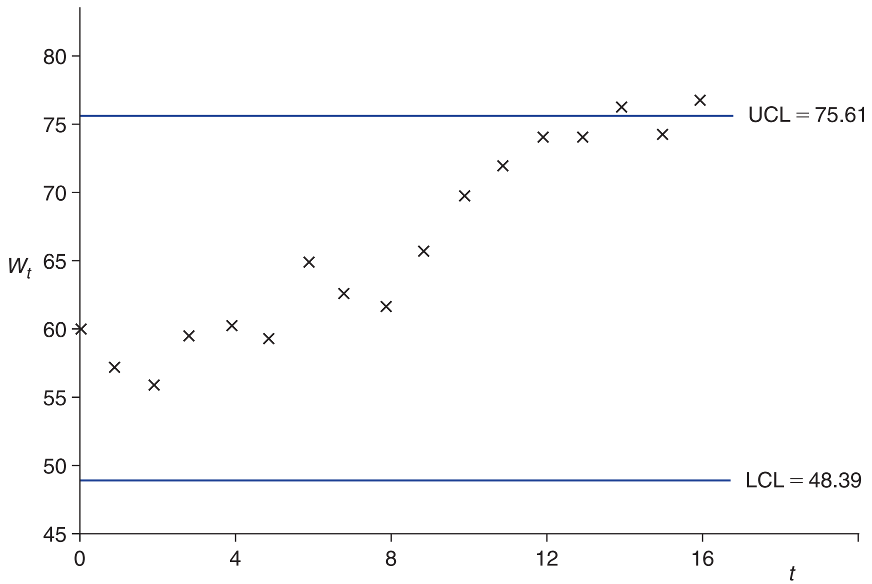

FIGURE 15.6
An EWMA control chart.

PROBLEMS

1. Consider a process whose in-control subgroup averages are normal with mean 50 and standard deviation 5. The following table represents 50 simulated values from a normal distribution with mean 54 and standard deviation 5. That is, the data represent subgroup averages after the process has gone out of control by an increase in the mean of 0.8 subgroup standard deviations.

Subgroups 1–10	Subgroups 11–20	Subgroups 21–30	Subgroups 31–40	Subgroups 41–50
59.81	52.79	50.04	58.56	54.95
51.38	56.52	51.49	50.79	53.22
52.87	54.98	56.93	54.19	49.31
62.38	58.00	50.13	58.65	47.08
53.59	56.91	52.29	57.41	53.31
60.39	54.26	56.74	56.79	51.42
46.64	58.43	48.62	50.86	61.57
55.84	53.41	44.58	51.14	57.33
55.78	48.43	57.46	61.26	60.34
53.27	52.65	60.98	56.68	49.97

(a) When would a standard $\overline{X}$ control chart have discovered the process to be out of control?

(b) Repeat part (a), this time using a standard EWMA chart with $\beta = 0.5$.

(c) Repeat part (a), this time using a standard EWMA chart with $\beta = 0.25$.

2. Repeat Problem 1, this time using the data in the reverse order. That is, let 49.97 be the first value, 60.34 the second, and so on.
3. Repeat Problem 2 of Section 15.2, this time employing a standard EWMA control chart with weighting factor $\beta = 0.7$.
4. Repeat Problem 3 of Section 15.2, this time employing a standard EWMA control chart with weighting factor $\beta = 0.5$.

15.5 CUMULATIVE-SUM CONTROL CHARTS

The major competitor to the EWMA type of control chart for detecting a small to moderate-sized change in the mean is the *cumulative-sum* (often reduced to as *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 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

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

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

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

Now let

$$\begin{aligned} S_0 &= 0 \\ S_{j+1} &= \max\left\{S_j + Y_{j+1}, 0\right\}, \quad j \geq 0 \end{aligned}$$

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

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

To understand the rationale behind this control chart, suppose we had decided to continually plot the sum of all the random variables $\boldsymbol{Y_i}$ 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

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

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 long 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 $\boldsymbol{S_j}$ is the cumulative sum of all of the Y_i up to time $\boldsymbol{j}$, with the exception that any time this sum becomes negative its value is reset to 0.

■ Example 15.7

Suppose that the mean and standard deviation of a subgroup average are, respectively, $\mu = 30$ and $\sigma/\sqrt{n} = 8$, and consider the cumulative-sum control chart with $d = 0.5, B = 5$. 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

$$\begin{aligned} &Y_1 = -5, \quad Y_2 = -1, \quad Y_3 = 1, \quad Y_4 = 8, \quad Y_5 = 2, \\ &\qquad Y_6 = 10, \quad Y_7 = 9, \quad Y_8 = 11 \end{aligned}$$

Therefore,

$$\begin{aligned} S_1 &= \max\{-5, 0\} = 0 \\ S_2 &= \max\{-1, 0\} = 0 \\ S_3 &= \max\{1, 0\} = 1 \end{aligned}$$

$$S_4 = \max\{9, 0\} = 9$$
$$S_5 = \max\{11, 0\} = 11$$
$$S_6 = \max\{21, 0\} = 21$$
$$S_7 = \max\{30, 0\} = 30$$
$$S_8 = \max\{41, 0\} = 41$$

Since the control limit is

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

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

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

and then also plot the values T_j, where

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

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: $d = 0.25, B = 8.00$; $d = 0.50, B = 4.77$; and $d = 1, B = 2.49$. 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 one wants to guard against, the smaller should be the chosen value of d.

PROBLEMS

1. Repeat Prob. 1 of Sec. 15.5, this time using a cumulative-sum chart with
   (a) $d = 0.25, B = 8$
   (b) $d = 0.5, B = 4.77$
2. Repeat Prob. 2 of Sec. 15.2, this time employing a cumulative-sum control chart with $d = 1$ and $B = 2.49$.
3. Repeat Prob. 3 of Sec. 15.2, this time employing a cumulative-sum control chart with $d = 0.5$ and $B = 4.77$.

KEY TERMS

Control chart: A graphical procedure for enabling one to detect when a production process has gone out of control.

SUMMARY

Suppose that a production process produces items, each of which has a measurable value that, when the process is in control, has mean μ and standard deviation σ. To detect any change, the items are put into subgroups of size n, and the subgroup averages $\overline{X}$ are plotted. Whenever a subgroup average is either less than the lower control limit

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

or greater than the upper control limit

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

then the process is declared to be out of control.

Sometimes rather than having a continuous value, each item is classified as being either acceptable or defective. Let p denote the probability that an item is defective when the process is in control. To determine when it goes out of control, items are again put into subgroups of size n. When the number of defectives in a subgroup falls outside the control limits

$$\text{LCL} = np - 3\sqrt{np(1-p)} \quad \text{and} \quad \text{UCL} = np + 3\sqrt{np(1-p)}$$

then the process is declared to be out of control.

Other types of control charts considered are the exponentially weighted moving-average control charts and cumulative-sum control charts. The former plots a weighted average of all the subgroup averages, with the weights decreasing exponentially the further in the past are data to which the subgroup average refers. The latter plots a cumulative sum of terms whose mean is negative when the process is in control, with this sum being reset to 0 when it becomes negative. The process is then called out of control when the cumulative sum becomes larger than some preset value.

REVIEW PROBLEMS

1. The distance between two adjacent pins of a memory chip for enhanced graphics adapters has, when the production process is in control, a mean of 1.5 millimeters and standard deviation of 0.001 millimeters. Determine the upper and lower control limits for an $\overline{X}$ control chart, using subgroups of size 4.
2. Prior to 1993 the number of burglaries committed yearly in the United States per 100,000 population was normally distributed with mean 1236 and standard deviation 120. The following are the rates from 1993 through 2001:

   1099.7, 1042.1, 987.0, 945.0, 918.8, 863.2, 770.4, 728.8, 740.8

   Can we conclude that the burglary rate changed from its historical value? Use subgroups of size 3.
3. When a process is performing correctly, 1.5 percent of the items produced do not conform to specifications. If items are grouped into subgroups of size 300, determine the upper and lower control limits for this control chart.
4. The numbers of defective switches in 12 samples of size 200 are as follows:

   4, 7, 2, 5, 9, 5, 7, 10, 8, 3, 12, 9

   Suppose that when the process is in control, each switch is defective with probability 0.03. Does the process appear to have been in control?