Sample Size n	Relative Efficiency
2	1.000
3	0.992
4	0.975
5	0.955
6	0.930
10	0.850

For moderate values of n—say, $n \geq 10$—the range loses efficiency rapidly, as it ignores all of the information in the sample between the extremes. However, for small sample sizes—say, $n \leq 6$—it works very well and is entirely satisfactory. We will use the range method to estimate the standard deviation for certain types of control charts in Chapter 5. The **supplemental text material** contains more information about using the range to estimate variability. Also see Woodall and Montgomery (2000–01).

3-3 STATISTICAL INFERENCE FOR A SINGLE SAMPLE

The techniques of statistical inference can be classified into two broad categories: **parameter estimation** and **hypothesis testing.** We have already briefly introduced the general idea of **point estimation** of process parameters.

A **statistical hypothesis** is a statement about the values of the parameters of a probability distribution. For example, suppose we think that the mean inside diameter of a bearing is 1.500 in. We may express this statement in a formal manner as

$$\begin{aligned} H_0&: \mu = 1.500 \\ H_1&: \mu \neq 1.500 \end{aligned} \tag{3-21}$$

The statement H_0: $\mu = 1.500$ in equation 3-21 is called the **null hypothesis,** and H_1: $\mu \neq 1.500$ is called the **alternative hypothesis.** In our example, H_1 specifies values of the mean diameter that are either greater than 1.500 or less than 1.500, and is called a **two-sided alternative hypothesis.** Depending on the problem, various one-sided alternative hypotheses may be appropriate.

Hypothesis testing procedures are quite useful in many types of statistical quality-control problems. They also form the basis for most of the statistical process-control techniques to be described in Parts II and III of this textbook. An important part of any hypothesis testing problem is determining the parameter values specified in the null and alternative hypotheses. Generally, this is done in one of three ways. First, the values may result from past evidence or knowledge. This happens frequently in statistical quality control, where we use past information to specify values for a parameter corresponding to a state of control, and then periodically test the hypothesis that the parameter value has not changed. Second, the values may result from some theory or model of the process. Finally, the values chosen for the parameter may be the result of contractual or design specifications, a situation that occurs frequently. Statistical hypothesis testing procedures may be used to check the conformity of the process parameters to their specified values, or to assist in modifying the process until the desired values are obtained.

To test a hypothesis, we take a random sample from the population under study, compute an appropriate **test statistic,** and then either reject or fail to reject the null hypothesis

H_0. The set of values of the test statistic leading to rejection of H_0 is called the **critical region** or **rejection region** for the test.

Two kinds of errors may be committed when testing hypotheses. If the null hypothesis is rejected when it is true, then a type I error has occurred. If the null hypothesis is not rejected when it is false, then a type II error has been made. The probabilities of these two types of errors are denoted as

$$\alpha = P\{\text{type I error}\} = P\{\text{reject } H_0 \mid H_0 \text{ is true}\}$$

$$\beta = P\{\text{type II error}\} = P\{\text{fail to reject } H_0 \mid H_0 \text{ is false}\}$$

Sometimes it is more convenient to work with the **power** of the test, where

$$\text{Power} = 1 - \beta = P\{\text{reject } H_0 \mid H_0 \text{ is false}\}$$

Thus, the power is the probability of *correctly* rejecting H_0. In quality control work, α is sometimes called the **producer's risk,** because it denotes the probability that a good lot will be rejected, or the probability that a process producing acceptable values of a particular quality characteristic will be rejected as performing unsatisfactorily. In addition, β is sometimes called the **consumer's risk,** because it denotes the probability of accepting a lot of poor quality, or allowing a process that is operating in an unsatisfactory manner relative to some quality characteristic to continue in operation.

The general procedure in hypothesis testing is to specify a value of the probability of type I error α, and then to design a test procedure so that a small value of the probability of type II error β is obtained. Thus, we speak of directly controlling or choosing the α risk. The β risk is generally a function of sample size and is controlled indirectly. The larger is the sample size(s) used in the test, the smaller is the β risk.

In this section we will review hypothesis testing procedures when a **single sample** of n observations has been taken from the process. We will also show how the information about the values of the process parameters that is in this sample can be expressed in terms of an interval estimate called a **confidence interval.** In Section 3-4 we will consider statistical inference for two samples from two possibly different processes.

3-3.1 Inference on the Mean of a Population, Variance Known

Hypothesis Testing

Suppose that x is a random variable with unknown mean μ and known variance σ^2. We wish to test the hypothesis that the mean is equal to a standard value—say, μ_0. The hypothesis may be formally stated as

$$\begin{aligned} H_0&: \ \mu = \mu_0 \\ H_1&: \ \mu \neq \mu_0 \end{aligned} \qquad (3\text{-}22)$$

The procedure for testing this hypothesis is to take a random sample of n observations on the random variable x, compute the test statistic

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \qquad (3\text{-}23)$$

3-3.6 The Probability of Type II Error and Sample Size Decisions

In most hypothesis testing situations, it is important to determine the probability of type II error associated with the test. Equivalently, we may elect to evaluate the power of the test. To illustrate how this may be done, we will find the probability of type II error associated with the test of

$$H_0: \quad \mu = \mu_0$$
$$H_1: \quad \mu \neq \mu_0$$

where the variance σ^2 is known. The test procedure was discussed in Section 3-3.1.

The test statistic for this hypothesis is

$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

and under the null hypothesis the distribution of Z_0 is $N(0, 1)$. To find the probability of type II error, we must assume that the null hypothesis H_0: $\mu = \mu_0$ is false and then find the distribution of Z_0. Suppose that the mean of the distribution is really $\mu_1 = \mu_0 + \delta$, where $\delta > 0$. Thus, the alternative hypothesis H_1: $\mu \neq \mu_0$ is true, and under this assumption the distribution of the test statistic Z_0 is

$$Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma},\ 1\right) \tag{3-45}$$

The distribution of the test statistic Z_0 under both hypotheses H_0 and H_1 is shown in Fig. 3-6. We note that the probability of type II error is the probability that Z_0 will fall between $-Z_{\alpha/2}$ and $Z_{\alpha/2}$ given that the alternative hypothesis H_1 is true. To evaluate this probability, we must find $F(Z_{\alpha/2}) - F(-Z_{\alpha/2})$, where F denotes the cumulative distribution function of the $N(\delta\sqrt{n}/\sigma, 1)$ distribution. In terms of the standard normal cumulative distribution, we then have

$$\beta = \Phi\left(Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \tag{3-46}$$

as the probability of type II error. This equation will also work when $\delta < 0$.

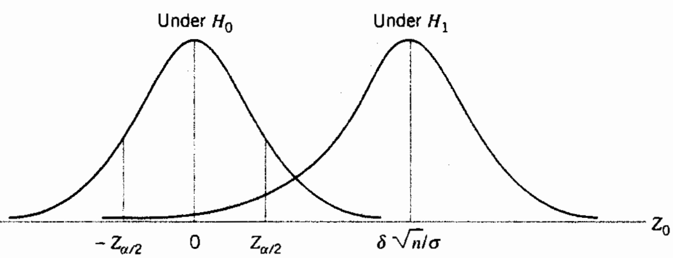

Figure 3-6 The distribution of Z_0 under H_0 and H_1.

EXAMPLE 3-7

The mean contents of coffee cans filled on a particular production line are being studied. Standards specify that the mean contents must be 16.0 oz, and from past experience it is known that the standard deviation of the can contents is 0.1 oz. The hypotheses are

$$H_0: \quad \mu = 16.0$$
$$H_1: \quad \mu \neq 16.0$$

A random sample of nine cans is to be used, and the type I error probability is specified as $\alpha = 0.05$. Therefore, the test statistic is

$$Z_0 = \frac{\bar{x} - 16.0}{0.1\sqrt{9}}$$

and H_0 is rejected if $|Z_0| > Z_{0.025} = 1.96$. Suppose that we wish to find the probability of type II error if the true mean contents are $\mu_1 = 16.1$ oz. Since this implies that $\delta = \mu_1 - \mu_0 = 16.1 - 16.0 = 0.1$, we have

$$\begin{aligned}
\beta &= \Phi\left(Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \\
&= \Phi\left(1.96 - \frac{(0.1)(3)}{0.1}\right) - \Phi\left(-1.96 - \frac{(0.1)(3)}{0.1}\right) \\
&= \Phi(-1.04) - \Phi(-4.96) \\
&= 0.1492
\end{aligned}$$

That is, the probability that we will incorrectly fail to reject H_0 if the true mean contents are 16.1 oz is 0.1492. Equivalently, we can say that the power of the test is $1 - \beta = 1 - 0.1492 = 0.8508$.

We note from examining equation 3-46 and Fig. 3-6 that β is a function of n, δ, and α. It is customary to plot curves illustrating the relationship between these parameters. Such a set of curves is shown in Fig. 3-7 for $\alpha = 0.05$. Graphs such as these are usually called **operating-characteristic (OC) curves.** The parameter on the vertical axis of these curves is β, and the parameter on the horizontal axis is $d = |\delta|/\sigma$. From examining the operating-characteristic curves, we see that

1. The further the true mean μ_1 is from the hypothesized value μ_0 (i.e., the larger the value of δ), the smaller is the probability of type II error for a given n and α. That is, for a specified sample size and α, the test will detect large differences more easily than small ones.
2. As the sample size n increases, the probability of type II error gets smaller for a specified δ and α. That is, to detect a specified difference we may make the test more powerful by increasing the sample size.

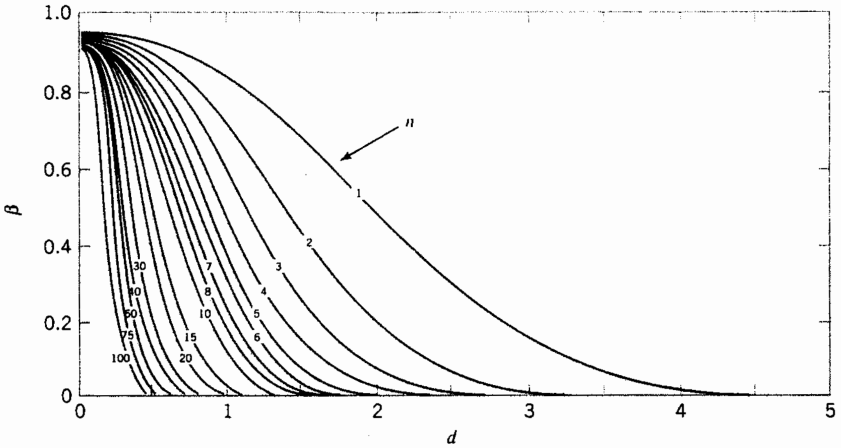

Figure 3-7 Operating-characteristic curves for the two-sided normal test with $\alpha = 0.05$. (Reproduced with permission from C. L. Ferris, F. E. Grubbs, and C. L. Weaver, "Operating Characteristic Curves for the Common Statistical Tests of Significance," *Annals of Mathematical Statistics*, June 1946.)

Operating-characteristic curves are useful in determining how large a sample is required to detect a specified difference with a particular probability. As an illustration, suppose that in Example 3-7 we wish to determine how large a sample will be necessary to have a 0.90 probability of rejecting H_0: $\mu = 16.0$ if the true mean is $\mu = 16.05$. Since $\delta = 16.05 - 16.0 = 0.05$, we have $d = |\delta|/\sigma = |0.05|/0.1 = 0.5$. From Fig. 3-7 with $\beta = 0.10$ and $d = 0.5$, we find $n = 45$, approximately. That is, 45 observations must be taken to ensure that the test has the desired probability of type II error.

Operating-characteristic curves are available for most of the standard statistical tests discussed in this chapter. For a detailed discussion of the use of operating-characteristic curves, refer to Montgomery and Runger (2003).

Minitab can also perform power and sample size calculations for several hypothesis testing problems. The following Minitab display reproduces the power calculations from the coffee can–filling problem in Example 3-7.

Power and Sample Size

```
1-Sample Z Test

Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05 Sigma = 0.1

                 Sample
  Difference       Size       Power
         0.1          9      0.8508
```

The following display shows several sample size and power calculations based on the rubberized asphalt problem in Example 3-3.

```
Power and Sample Size

1-Sample t Test
Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05 Sigma = 117.61

                Sample
 Difference      Size       Power
         50        15      0.3354

1-Sample t Test
Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05 Sigma = 117.61

                Sample      Target     Actual
 Difference      Size       Power      Power
         50        46      0.8000     0.8055

1-Sample t Test
Testing mean = null (versus not = null)
Calculating power for mean = null + difference
Alpha = 0.05 Sigma = 117.61

                Sample
 Difference      Size       Power
        100        15      0.8644
```

In the first portion of the display, Minitab calculates the power of the test in Example 3-3, assuming that the engineer would wish to reject the null hypothesis if the true mean stabilized viscosity differed from 3200 by as much as 50, using $s = 117.61$ as an estimate of the true standard deviation. The power is 0.3354, which is low. The next calculation determines the sample size that would be required to produce a power of 0.8, a much better value. Minitab reports that a considerably larger sample size, $n = 46$, would be required. The final calculation determines the power with $n = 15$ if a larger difference between the true mean stabilized viscosity and the hypothesized value is of interest. For a difference of 100, Minitab reports the power to be 0.8644.

3-4 STATISTICAL INFERENCE FOR TWO SAMPLES

The previous section presented hypothesis tests and confidence intervals for a single population parameter (the mean μ, the variance σ^2, or a proportion p). This section extends those results to the case of two independent populations.

4 Methods and Philosophy of Statistical Process Control

CHAPTER OUTLINE

The supplemental material is on the textbook website www.wiley.com/college/montgomery.

CHAPTER OVERVIEW AND LEARNING OBJECTIVES

This chapter has three objectives. The first is to present the basic SPC problem-solving tools, called the "magnificent seven," and to illustrate how these tools form a cohesive, practical framework for quality improvement. The second objective is to describe the statistical basis of the Shewhart control chart. The reader will see how decisions about sample size, sampling interval, and placement of control limits affect the performance of a control chart. Other key concepts include the idea of rational subgroups, interpretation of control chart signals and patterns, and the average run length as a measure of control chart performance. The third objective is to discuss and illustrate some practical issues in the implementation of SPC.

After careful study of this chapter you should be able to do the following:

1. Understand chance and assignable causes of variability in a process
2. Explain the statistical basis of the Shewhart control chart, including choice of sample size, control limits, and sampling interval
3. Explain the rational subgroup concept
4. Understand the basic tools of SPC; the histogram or stem-and-leaf plot, the check sheet, the Pareto chart, the cause-and-effect diagram, the defect concentration diagram, the scatter diagram, and the control chart

5. Explain phase I and phase II use of control charts
6. Explain how average run length is used as a performance measure for a control chart
7. Explain how sensitizing rules and pattern recognition are used in conjunction with control charts

4-1 INTRODUCTION

If a product is to meet or exceed customer expectations, generally it should be produced by a process that is stable or repeatable. More precisely, the process must be capable of operating with little variability around the target or nominal dimensions of the product's quality characteristics. **Statistical process control (SPC)** is a powerful collection of problem-solving tools useful in achieving process stability and improving capability through the reduction of variability.

SPC can be applied to *any* process. Its seven major tools are

1. Histogram or stem-and-leaf plot
2. Check sheet
3. Pareto chart
4. Cause-and-effect diagram
5. Defect concentration diagram
6. Scatter diagram
7. Control chart

Although these tools, often called **"the magnificent seven,"** are an important part of SPC, they comprise only its technical aspects. SPC builds an environment in which all individuals in an organization seek continuous improvement in quality and productivity. This environment is best developed when management becomes involved in the process. Once this environment is established, routine application of the magnificent seven becomes part of the usual manner of doing business, and the organization is well on its way to achieving its quality improvement objectives.

In this chapter we will present an overview of the magnificent seven. Of these tools, the Shewhart control chart is probably the most technically sophisticated. It was developed in the 1920s by Walter A. Shewhart of the Bell Telephone Laboratories. To understand the statistical concepts that form the basis of SPC, we must first describe Shewhart's theory of variability.

4-2 CHANCE AND ASSIGNABLE CAUSES OF QUALITY VARIATION

In any production process, regardless of how well designed or carefully maintained it is, a certain amount of inherent or natural variability will always exist. This natural variability or "background noise" is the cumulative effect of many small, essentially unavoidable causes. In the framework of statistical quality control, this natural variability is often called a "stable system of chance causes." A process that is operating with only **chance causes of variation** present is said to be **in statistical control.** In other words, the chance causes are an inherent part of the process.

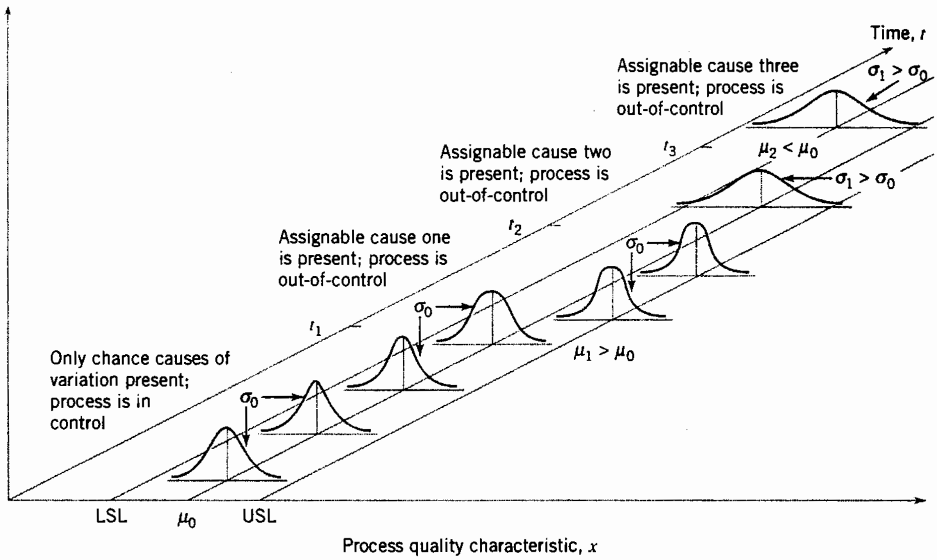

Figure 4-1 Chance and assignable causes of variation.

Other kinds of variability may occasionally be present in the output of a process. This variability in key quality characteristics usually arises from three sources: improperly adjusted or controlled machines, operator errors, or defective raw material. Such variability is generally large when compared to the background noise, and it usually represents an unacceptable level of process performance. We refer to these sources of variability that are not part of the chance cause pattern as "assignable causes." A process that is operating in the presence of **assignable causes** is said to be **out of control.**[1]

These chance and assignable causes of variation are illustrated in Fig. 4-1. Until time t_1 the process shown in this figure is in control; that is, only chance causes of variation are present. As a result, both the mean and standard deviation of the process are at their in-control values (say, μ_0 and σ_0). At time t_1 an assignable cause occurs. As shown in Fig. 4-1, the effect of this assignable cause is to shift the process mean to a new value $\mu_1 > \mu_0$. At time t_2 another assignable cause occurs, resulting in $\mu = \mu_0$, but now the process standard deviation has shifted to a larger value $\sigma_1 > \sigma_0$. At time t_3 there is another assignable cause present, resulting in both the process mean and standard deviation taking on out-of-control values. From time t_1 forward, the presence of assignable causes has resulted in an out-of-control process.

Processes will often operate in the in-control state for relatively long periods of time. However, no process is truly stable forever, and, eventually, assignable causes will occur, seemingly at random, resulting in a "shift" to an out-of-control state where a larger proportion of the process output does not conform to requirements. For example, note from Fig. 4-1 that when the process is in control, most of the production will fall between the lower and upper specification limits (LSL and USL, respectively). When the process is out of control, a higher proportion of the process lies outside of these specifications.

[1]The terminology **chance** and **assignable causes** was developed by Shewhart. Today, some writers use the terminology **common cause** instead of **chance cause** and **special cause** instead of **assignable cause.**

A major objective of statistical process control is to quickly detect the occurrence of assignable causes of process shifts so that investigation of the process and corrective action may be undertaken before many nonconforming units are manufactured. The control chart is an on-line process-monitoring technique widely used for this purpose. Control charts may also be used to estimate the parameters of a production process, and, through this information, to determine process capability. The control chart may also provide information useful in improving the process. Finally, remember that the eventual goal of statistical process control is the **elimination of variability in the process.** It may not be possible to completely eliminate variability, but the control chart is an effective tool in reducing variability as much as possible.

We now present the statistical concepts that form the basis of control charts. Chapters 5 and 6 develop the details of construction and use of the standard types of control charts.

4-3 STATISTICAL BASIS OF THE CONTROL CHART

4-3.1 Basic Principles

A typical control chart is shown in Fig. 4-2. The control chart is a graphical display of a quality characteristic that has been measured or computed from a sample versus the sample number or time. The chart contains a **center line** that represents the average value of the quality characteristic corresponding to the in-control state. (That is, only chance causes are present.) Two other horizontal lines, called the **upper control limit** (UCL) and the **lower control limit** (LCL), are also shown on the chart. These control limits are chosen so that if the process is in control, nearly all of the sample points will fall between them. As long as the points plot within the control limits, the process is assumed to be in control, and no action is necessary. However, a point that plots outside of the control limits is interpreted as evidence that the process is out of control, and investigation and corrective action are required to find and eliminate the assignable cause or causes responsible for this behavior. It is customary to connect the sample points on the control chart with straight-line segments, so that it is easier to visualize how the sequence of points has evolved over time.

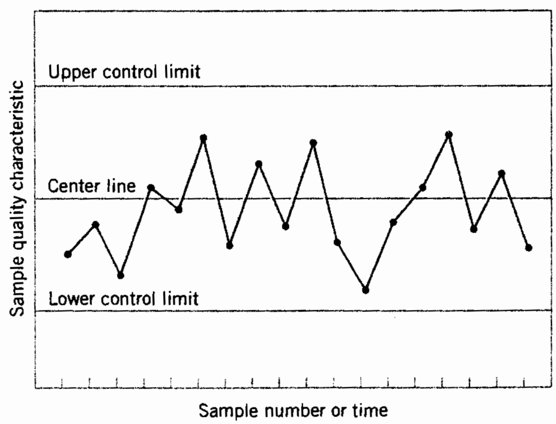

Figure 4-2 A typical control chart.

Even if all the points plot inside the control limits, if they behave in a systematic or nonrandom manner, then this could be an indication that the process is out of control. For example, if 18 of the last 20 points plotted above the center line but below the upper control limit and only two of these points plotted below the center line but above the lower control limit, we would be very suspicious that something was wrong. If the process is in control, all the plotted points should have an essentially random pattern. Methods for looking for sequences or nonrandom patterns can be applied to control charts as an aid in detecting out-of-control conditions. Usually, there is a reason why a particular nonrandom pattern appears on a control chart, and if it can be found and eliminated, process performance can be improved. This topic is discussed further in Sections 4-3.5 and 5-2.4.

There is a close connection between control charts and **hypothesis testing.** To illustrate this connection, suppose that the vertical axis in Fig. 4-2 is the sample average $\bar{x}$. Now, if the current value of $\bar{x}$ plots between the control limits, we conclude that the process mean is in control; that is, it is equal to some value μ_0. On the other hand, if $\bar{x}$ exceeds either control limit, we conclude that the process mean is out of control; that is, it is equal to some value $\mu_1 \neq \mu_0$. In a sense, then, the control chart is a test of the hypothesis that the process is in a state of statistical control. A point plotting within the control limits is equivalent to failing to reject the hypothesis of statistical control, and a point plotting outside the control limits is equivalent to rejecting the hypothesis of statistical control. This hypothesis testing framework is useful in many ways, but there are some differences in viewpoint between control charts and hypothesis tests. For example, when testing statistical hypotheses, we usually check the validity of assumptions, whereas control charts are used to detect departures from an assumed state of statistical control. In general, we should not worry too much about assumptions such as the form of the distribution or independence when we are applying control charts to a process to reduce variability and achieve statistical control. Furthermore, an assignable cause can result in many different types of shifts in the process parameters. For example, the mean could shift instantaneously to a new value and remain there (this is sometimes called a **sustained shift**); or it could shift abruptly; but the assignable cause could be short lived and the mean could then return to its nominal or in-control value; or the assignable cause could result in a steady drift or trend in the value of the mean. Only the sustained shift fits nicely within the usual statistical hypothesis testing model.

One place where the hypothesis testing framework is useful is in analyzing the **performance** of a control chart. For example, we may think of the probability of type I error of the control chart (concluding the process is out of control when it is really in control) and the probability of type II error of the control chart (concluding the process is in control when it is really out of control). It is occasionally helpful to use the operating-characteristic curve of a control chart to display its probability of type II error. This would be an indication of the ability of the control chart to detect process shifts of different magnitudes. This can be of value in determining which type of control chart to apply in certain situations. For more discussion of hypothesis testing, the role of statistical theory, and control charts, see Woodall (2000).

To illustrate the preceding ideas, we give an example of a control chart. In semiconductor manufacturing, an important fabrication step is photolithography, in which a light sensitive photoresist material is applied to the silicon wafer, the circuit pattern is exposed on the resist typically through the use of high-intensity UV light, and the unwanted resist material removed through a developing process. After the resist pattern is defined, the underlying material is removed by either wet chemical or plasma etching. It is fairly typical to

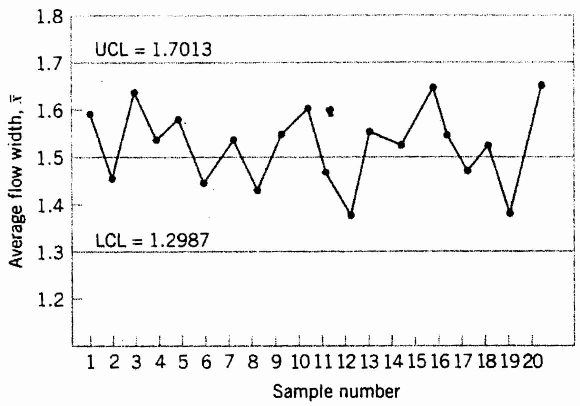

Figure 4-3 $\bar{x}$ control chart for flow width.

follow development with a hard-bake process to increase resist adherence and etch resistance. An important quality characteristic in hard bake is the flow width of the resist, a measure of how much it expands due to the baking process. Suppose that flow width can be controlled at a mean of 1.5 microns, and it is known that the standard deviation of flow width is 0.15 microns. A control chart for the average flow width is shown in Fig. 4-3. Every hour, a sample of five wafers is taken, the average flow width ($\bar{x}$) computed, and $\bar{x}$ plotted on the chart. Because this control chart utilizes the sample average $\bar{x}$ to monitor the process mean, it is usually called an $\bar{x}$ control chart. Note that all of the plotted points fall inside the control limits, so the chart indicates that the process is considered to be in statistical control.

To assist in understanding the statistical basis of this control chart, consider how the control limits were determined. The process mean is 1.5 microns, and the process standard deviation is $\sigma = 0.15$ microns. Now if samples of size $n = 5$ are taken, the standard deviation of the sample average $\bar{x}$ is

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{0.15}{\sqrt{5}} = 0.0671$$

Therefore, if the process is in control with a mean flow width of 1.5 microns, then by using the central limit theorem to assume that $\bar{x}$ is approximately normally distributed, we would expect $100(1 - \alpha)\%$ of the sample means $\bar{x}$ to fall between $1.5 + Z_{\alpha/2}(0.0671)$ and $1.5 - Z_{\alpha/2}\,(0.0671)$. We will arbitrarily choose the constant $Z_{\alpha/2}$ to be 3, so that the upper and lower control limits become

$$\text{UCL} = 1.5 + 3(0.0671) = 1.7013$$

and

$$\text{LCL} = 1.5 - 3(0.0671) = 1.2987$$

as shown on the control chart. These are typically called **"three-sigma"[2] control limits.**

[2]Note that "sigma" refers to the standard deviation of the statistic plotted on the chart (i.e., $\sigma_{\bar{x}}$), *not* the standard deviation of the quality characteristic.

The width of the control limits is inversely proportional to the sample size n for a given multiple of sigma. Note that choosing the control limits is equivalent to setting up the critical region for testing the hypothesis

$$H_0: \quad \mu = 1.5$$
$$H_1: \quad \mu \neq 1.5$$

where $\sigma = 0.15$ is known. Essentially, the control chart tests this hypothesis repeatedly at different points in time. The situation is illustrated graphically in Fig. 4-4.

We may give a general **model** for a control chart. Let w be a sample statistic that measures some quality characteristic of interest, and suppose that the mean of w is μ_w and the standard deviation of w is σ_w. Then the center line, the upper control limit, and the lower control limit become

$$\begin{aligned} \text{UCL} &= \mu_w + L\sigma_w \\ \text{Center line} &= \mu_w \\ \text{LCL} &= \mu_w - L\sigma_w \end{aligned} \tag{4-1}$$

where L is the "distance" of the control limits from the center line, expressed in standard deviation units. This general theory of control charts was first proposed by Walter A. Shewhart, and control charts developed according to these principles are often called **Shewhart control charts.**

The control chart is a device for describing in a precise manner exactly what is meant by statistical control; as such, it may be used in a variety of ways. In many applications, it is used for on-line process surveillance. That is, sample data are collected and used to construct the control chart, and if the sample values of $\bar{x}$ (say) fall within the control limits and do not exhibit any systematic pattern, we say the process is in control at the level indicated

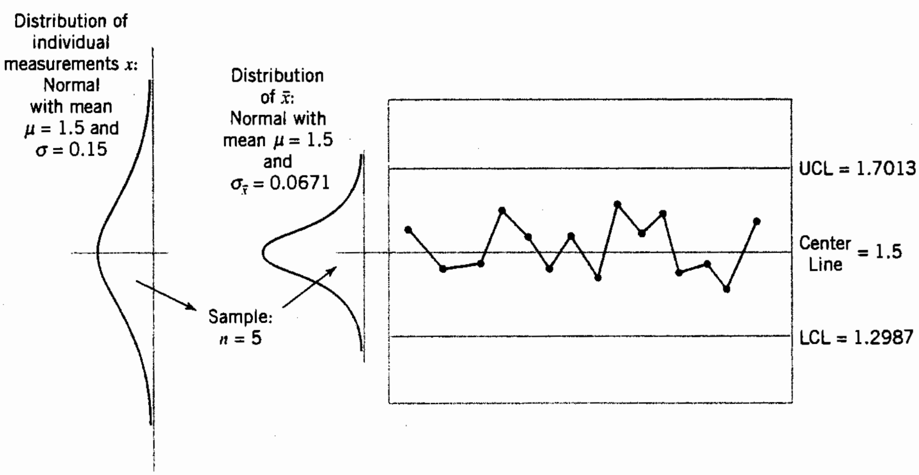

Figure 4-4 How the control chart works.

by the chart. Note that we may be interested here in determining *both* whether the past data came from a process that was in control and whether future samples from this process indicate statistical control.

The most important use of a control chart is to **improve** the process. We have found that, generally,

1. Most processes do not operate in a state of statistical control.
2. Consequently, the routine and attentive use of control charts will identify assignable causes. If these causes can be eliminated from the process, variability will be reduced and the process will be improved.

This process improvement activity using the control chart is illustrated in Fig. 4-5. Note that

3. The control chart will only **detect** assignable causes. Management, operator, and engineering **action** will usually be necessary to eliminate the assignable causes.

In identifying and eliminating assignable causes, it is important to find the underlying **root cause** of the problem and to attack it. A cosmetic solution will not result in any real, long-term process improvement. Developing an effective system for corrective action is an essential component of an effective SPC implementation.

A very important part of the corrective action process associated with control chart usage is the **Out-of-Control-Action Plan** or **OCAP.** An OCAP is a flow chart or text-based description of the sequence of activities that must take place following the occurrence of an *activating event.* These are usually out-of-control signals from the control chart. The OCAP consists of *checkpoints*, which are potential assignable causes, and *terminators*, which are actions taken to resolve the out-of-control condition, hopefully by eliminating the assignable cause. It is very important that the OCAP specify as complete a set as possible of checkpoints and terminators, and that these be arranged in an order that facilitates process diagnostic activities. Often, analysis of prior failure modes of the process and/or product can be helpful in designing this aspect of the OCAP. Furthermore, an OCAP is a *living document* in the sense that it will be modified over time as more knowledge and understanding of the process is gained. Consequently, when a control chart is introduced, an initial OCAP should accompany it. Control charts without an OCAP are not likely to be very useful as a process improvement tool.

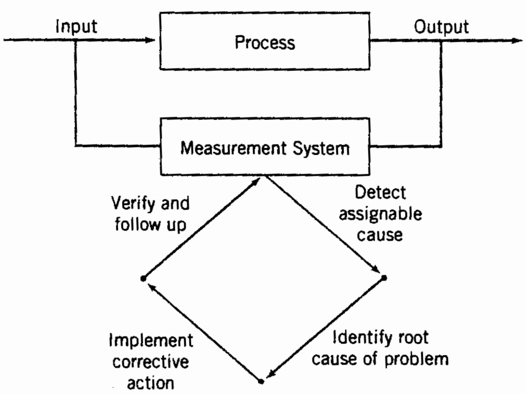

Figure 4-5 Process improvement using the control chart.

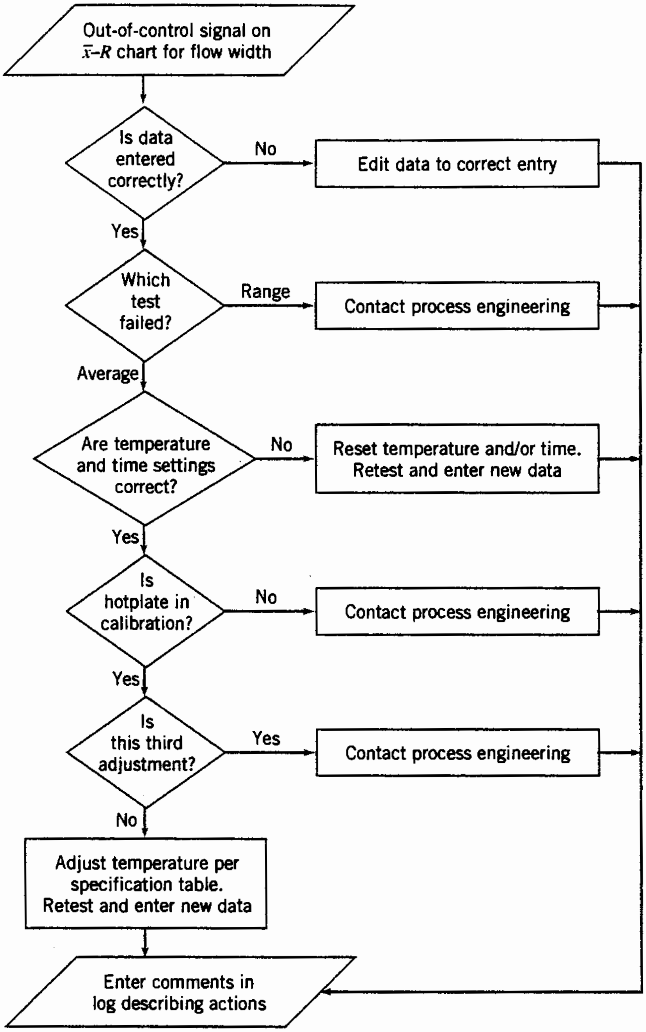

Figure 4-6 The out-of-control-action plan (OCAP) for the hard-bake process.

The OCAP for the hard-bake process is shown in Fig. 4-6. This process has two controllable variables, temperature and time. In this process, the mean flow width is monitored with an $\bar{x}$ control chart, and the process variability is monitored with a control chart for the range, or an R chart. Notice that if the R chart exhibits an out-of-control signal, operating personnel are directed to contact process engineering immediately. If the $\bar{x}$ control chart exhibits an out-of-control signal, operators are directed to check process settings and calibration and then make adjustments to temperature in an effort to bring the process back into a state of control. If these adjustments are unsuccessful, process engineering personnel are contacted.

We may also use the control chart as an **estimating device.** That is, from a control chart that exhibits statistical control, we may estimate certain process parameters, such as the mean, standard deviation, fraction nonconforming or fallout, and so forth. These estimates may then be used to determine the **capability** of the process to produce acceptable products. Such **process-capability studies** have considerable impact on many management decision problems that occur over the product cycle, including make or buy decisions, plant and process improvements that reduce process variability, and contractual agreements with customers or vendors regarding product quality.

Control charts may be classified into two general types. If the quality characteristic can be measured and expressed as a number on some continuous scale of measurement, it is usually called a **variable.** In such cases, it is convenient to describe the quality characteristic with a measure of central tendency and a measure of variability. Control charts for central tendency and variability are collectively called **variables control charts.** The $\bar{x}$ chart is the most widely used chart for controlling central tendency, whereas charts based on either the sample range or the sample standard deviation are used to control process variability. Control charts for variables are discussed in Chapter 5. Many quality characteristics are not measured on a continuous scale or even a quantitative scale. In these cases, we may judge each unit of product as either conforming or nonconforming on the basis of whether or not it possesses certain attributes, or we may count the number of nonconformities (defects) appearing on a unit of product. Control charts for such quality characteristics are called **attributes control charts** and are discussed in Chapter 6.

An important factor in control chart usage is the **design of the control chart.** By this we mean the selection of the sample size, control limits, and frequency of sampling. For example, in the $\bar{x}$ chart of Fig. 4-3, we specified a sample size of five measurements, three-sigma control limits, and the sampling frequency to be every hour. In most quality-control problems, it is customary to design the control chart using primarily statistical considerations. For example, we know that increasing the sample size will decrease the probability of type II error, thus enhancing the chart's ability to detect an out-of-control state, and so forth. The use of statistical criteria such as these along with industrial experience has led to general guidelines and procedures for designing control charts. These procedures usually consider cost factors only in an implicit manner. Recently, however, we have begun to examine control chart design from an **economic** point of view, considering explicitly the cost of sampling, losses from allowing defective product to be produced, and the costs of investigating out-of-control signals that are really "false alarms."

Another important consideration in control chart usage is the **type of variability** exhibited by the process. Fig. 4-7 presents data from three different processes. Figures 4-7*a* and 4-7*b* illustrate **stationary behavior.** By this we mean that the process data vary around a fixed mean in a stable or predictable manner. This is the type of behavior that Shewhart implied was produced by an **in-control process.**

Even a cursory examination of Figs. 4-7*a* and 4-7*b* reveals some important differences. The data in Fig. 4-7*a* are **uncorrelated;** that is, the observations give the appearance of having been drawn at random from a stable population, perhaps a normal

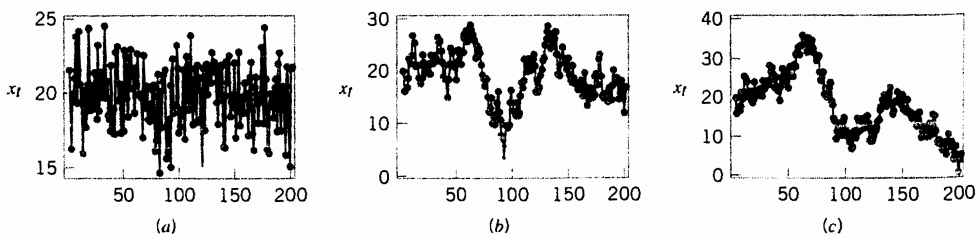

Figure 4-7 Data from three different processes. (*a*) Stationary and uncorrelated (white noise). (*b*) Stationary and autocorrelated. (*c*) Nonstationary.

distribution. This type of data is referred to by time series analysts as **white noise.** (Time-series analysis is a field of statistics devoted exclusively to studying and modeling time-oriented data.) In this type of process, the order in which the data occur does not tell us much that is useful to analyze the process. In other words, the past values of the data are of no help in predicting any of the future values.

Figure 4-7*b* illustrates stationary but **autocorrelated** process data. Notice that successive observations in these data are **dependent;** that is, a value above the mean tends to be followed by another value above the mean, whereas a value below the mean is usually followed by another such value. This produces a data series that has a tendency to move in moderately long "runs" on either side of the mean.

Figure 4-7*c* illustrates **nonstationary** variation. This type of process data occurs frequently in the chemical and process industries. Note that the process is very unstable in that it drifts or "wanders" about without any sense of a stable or fixed mean. In many industrial settings, we stabilize this type of behavior by using **engineering process control** (such as **feedback control**). This approach to process control is required when there are factors that affect the process that cannot be stabilized, such as environmental variables or properties of raw materials. When the control scheme is effective, the process output will *not* look like Fig. 4-7*c* but will hopefully resemble either Fig. 4-7*a* or 4-7*b*.

Shewhart control charts are most effective when the in-control process data look like Fig. 4-7*a*. By this we mean that the charts can be designed so that their performance is predictable and reasonable to the user, and that they are effective in reliably detecting out-of-control conditions. Most of our discussion of control charts in this chapter and in Chapters 5 and 6 will assume that the in-control process data are stationary and uncorrelated.

With some modifications, Shewhart control charts and other types of control charts can be applied to autocorrelated data. We discuss this in more detail in Part III of the book. We also discuss feedback control and the use of SPC in systems where feedback control is employed in Part III.

Control charts have had a long history of use in U.S. industries and in many offshore industries as well. There are at least five reasons for their popularity.

1. **Control charts are a proven technique for improving productivity.** A successful control chart program will reduce scrap and rework, which are the primary productivity killers in *any* operation. If you reduce scrap and rework, then productivity increases, cost decreases, and production capacity (measured in the number of *good* parts per hour) increases.
2. **Control charts are effective in defect prevention.** The control chart helps keep the process in control, which is consistent with the "do it right the first time" philosophy. It is never cheaper to sort out "good" units from "bad" units later on than it is to build it right initially. If you do not have effective process control, you are paying someone to make a nonconforming product.
3. **Control charts prevent unnecessary process adjustment.** A control chart can distinguish between background noise and abnormal variation; no other device including a human operator is as effective in making this distinction. If process operators adjust the process based on periodic tests unrelated to a control chart program, they will often overreact to the background noise and make unneeded adjustments. These unnecessary adjustments can actually result in a deterioration of process performance. In other words, the control chart is consistent with the "if it isn't broken, don't fix it" philosophy.

4. **Control charts provide diagnostic information.** Frequently, the pattern of points on the control chart will contain information of diagnostic value to an experienced operator or engineer. This information allows the implementation of a change in the process that improves its performance.
5. **Control charts provide information about process capability.** The control chart provides information about the value of important process parameters and their stability over time. This allows an estimate of process capability to be made. This information is of tremendous use to product and process designers.

Control charts are among the most important management control tools; they are as important as cost controls and material controls. Modern computer technology has made it easy to implement control charts in *any* type of process, as data collection and analysis can be performed on a microcomputer or a local area network terminal in real-time, on-line at the work center. Some additional guidelines for implementing a control chart program are given at the end of Chapter 6.

4-3.2 Choice of Control Limits

Specifying the control limits is one of the critical decisions that must be made in designing a control chart. By moving the control limits farther from the center line, we decrease the risk of a type I error—that is, the risk of a point falling beyond the control limits, indicating an out-of-control condition when no assignable cause is present. However, widening the control limits will also increase the risk of a type II error—that is, the risk of a point falling between the control limits when the process is really out of control. If we move the control limits closer to the center line, the opposite effect is obtained: The risk of type I error is increased, while the risk of type II error is decreased.

For the $\bar{x}$ chart shown in Fig. 4-3, where three-sigma control limits were used, if we assume that the flow width is normally distributed, we find from the standard normal table that the probability of type I error is 0.0027. That is, an incorrect out-of-control signal or false alarm will be generated in only 27 out of 10,000 points. Furthermore, the probability that a point taken when the process is in control will exceed the three-sigma limits in one direction only is 0.00135. Instead of specifying the control limit as a multiple of the standard deviation of $\bar{x}$, we could have directly chosen the type I error probability and calculated the corresponding control limit. For example, if we specified a 0.001 type I error probability in one direction, then the appropriate multiple of the standard deviation would be 3.09. The control limits for the $\bar{x}$ chart would then be

$$\text{UCL} = 1.5 + 3.09(0.0671) = 1.7073$$

$$\text{LCL} = 1.5 - 3.09(0.0671) = 1.2927$$

These control limits are usually called 0.001 **probability limits,** although they should logically be called 0.002 probability limits, because the total risk of making a type I error is 0.002. There is only a slight difference between the two limits.

Regardless of the distribution of the quality characteristic, it is standard practice in the United States to determine the control limits as a multiple of the standard deviation of the statistic plotted on the chart. The multiple usually chosen is three; hence, three-sigma limits are customarily employed on control charts, regardless of the type of chart employed.

In the United Kingdom and parts of Western Europe, probability limits are used, with the standard probability level in each direction being 0.001.

We typically justify the use of three-sigma control limits on the basis that they give good results in practice. Moreover, in many cases, the true distribution of the quality characteristic is not known well enough to compute exact probability limits. If the distribution of the quality characteristic is reasonably approximated by the normal distribution, then there will be little difference between three-sigma and 0.001 probability limits.

Warning Limits on Control Charts

Some analysts suggest using two sets of limits on control charts, such as those shown in Fig. 4-8. The outer limits—say, at three-sigma—are the usual **action limits;** that is, when a point plots outside of this limit, a search for an assignable cause is made and corrective action is taken if necessary. The inner limits, usually at two-sigma, are called **warning limits.** In Fig. 4-8, we have shown the three-sigma upper and lower control limits for the $\bar{x}$ chart for flow width. The upper and lower warning limits are located at

$$\text{UWL} = 1.5 + 2(0.0671) = 1.6342$$

$$\text{LWL} = 1.5 - 2(0.0671) = 1.3658$$

When probability limits are used, the action limits are generally 0.001 limits and the warning limits are 0.025 limits.

If one or more points fall between the warning limits and the control limits, or very close to the warning limit, we should be suspicious that the process may not be operating properly. One possible action to take when this occurs is to increase the sampling frequency and/or the sample size so that more information about the process can be obtained quickly. Process control schemes that change the sample size and/or the sampling frequency depending on the position of the current sample value are called **adaptive** or **variable sampling interval** (or **variable sample size,** etc.) schemes. These techniques have been used in practice for many years and have recently been studied extensively by researchers in the field. We will discuss this technique again in Part III of this book.

The use of warning limits can increase the **sensitivity** of the control chart; that is, it can allow the control chart to signal a shift in the process more quickly. One of their disadvantages is that they may be confusing to operating personnel. This is not usually a serious objection, however, and many practitioners use warning limits routinely on control charts. A more serious objection is that although the use of warning limits can improve the

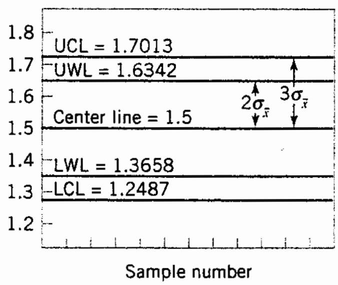

Figure 4-8 An $\bar{x}$ chart with two-sigma warning limits.

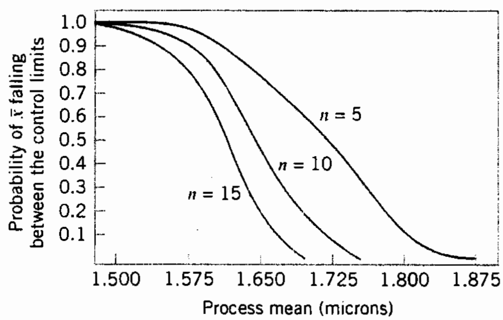

Figure 4-9 Operating-characteristic curves for an $\bar{x}$ chart.

sensitivity of the chart, they also result in an **increased risk of false alarms.** We will discuss the use of sensitizing rules (such as warning limits) more thoroughly in Section 4-3.6.

4-3.3 Sample Size and Sampling Frequency

In designing a control chart, we must specify both the **sample size** to use and the **frequency of sampling.** In general, larger samples will make it easier to detect small shifts in the process. This is demonstrated in Fig. 4-9, where we have plotted the operating characteristic curve for the $\bar{x}$ chart in Fig. 4-3 for various sample sizes. Note that the probability of detecting a shift from 1.500 microns to 1.650 microns (for example) increases as the sample size n increases. When choosing the sample size, we must keep in mind the size of the shift that we are trying to detect. If the process shift is relatively large, then we use smaller sample sizes than those that would be employed if the shift of interest were relatively small.

We must also determine the frequency of sampling. The most desirable situation from the point of view of detecting shifts would be to take large samples very frequently; however, this is usually not economically feasible. The general problem is one of **allocating sampling effort.** That is, either we take small samples at short intervals or larger samples at longer intervals. Current industry practice tends to favor smaller, more frequent samples, particularly in high-volume manufacturing processes, or where a great many types of assignable causes can occur. Furthermore, as automatic sensing and measurement technology develops, it is becoming possible to greatly increase sampling frequencies. Ultimately, every unit can be tested as it is manufactured. Automatic measurement systems and microcomputers with SPC software applied at the work center for real-time on-line process control is an effective way to apply statistical process control.

Another way to evaluate the decisions regarding sample size and sampling frequency is through the **average run length (ARL)** of the control chart. Essentially, the ARL is the average number of points that must be plotted before a point indicates an out-of-control condition. If the process observations are uncorrelated, then for any Shewhart control chart, the ARL can be calculated easily from

$$\text{ARL} = \frac{1}{p} \tag{4-2}$$

where p is the probability that any point exceeds the control limits. This equation can be used to evaluate the performance of the control chart.

To illustrate, for the $\bar{x}$ chart with three-sigma limits, $p = 0.0027$ is the probability that a single point falls outside the limits when the process is in control. Therefore, the average run length of the $\bar{x}$ chart when the process is in control (called ARL_0) is

$$\text{ARL}_0 = \frac{1}{p} = \frac{1}{0.0027} = 370$$

That is, even if the process remains in control, an out-of-control signal will be generated every 370 samples, on the average.

The use of average run lengths to describe the performance of control charts has been subjected to criticism in recent years. The reasons for this arise because the distribution of run length for a Shewhart control chart is a geometric distribution (refer to Section 2-2.4). Consequently, there are two concerns with ARL: (1) the standard deviation of the run length is very large, and (2) the geometric distribution is very skewed, so the mean of the distribution (the ARL) is not necessarily a very "typical" value of the run length.

For example, consider the Shewhart $\bar{x}$ control chart with three-sigma limits. When the process is in control, we have noted that $p = 0.0027$ and the in-control ARL_0 is $ARL_0 = 1/p = 1/0.0027 = 370$. This is the mean of the geometric distribution. Now the standard deviation of the geometric distribution is

$$\sqrt{(1-p)}/p = \sqrt{(1-0.0027)}/0.0027 \cong 370$$

That is, the standard deviation of the geometric distribution in this case is approximately equal to its mean. As a result, the actual ARL_0 observed in practice for the Shewhart $\bar{x}$ control chart will likely vary considerably. Furthermore, for the geometric distribution with $p = 0.0027$, the 10th and 50th percentiles of the distribution are 38 and 256, respectively. This means that approximately 10% of the time the in-control run length will be less than or equal to 38 samples and 50% of the time it will be less than or equal to 256 samples. This occurs because the geometric distribution with $p = 0.0027$ is quite skewed to the right.

It is also occasionally convenient to express the performance of the control chart in terms of its **average time to signal (ATS).** If samples are taken at fixed intervals of time that are h hours apart, then

$$\text{ATS} = \text{ARL}h \tag{4-3}$$

Consider the hard-bake process discussed earlier, and suppose we are sampling every hour. Equation 4-3 indicates that we will have a **false alarm** about every 370 hours on the average.

Now consider how the control chart performs in detecting shifts in the mean. Suppose we are using a sample size of $n = 5$ and that when the process goes out of control the mean shifts to 1.725 microns. From the operating characteristic curve in Fig. 4-9 we find that if the process mean is 1.725 microns, the probability of $\bar{x}$ falling between the control limits is approximately 0.50. Therefore, p in equation 4-2 is 0.50, and the out-of-control ARL (called ARL_1) is

$$\text{ARL}_1 = \frac{1}{p} = \frac{1}{0.5} = 2$$

That is, the control chart will require two samples to detect the process shift, on the average, and since the time interval between samples is $h = 1$ hour, the average time required to detect this shift is

$$\text{ATS} = \text{ARL}_1 h = 2(1) = 2 \text{ hours}$$

Suppose that this is unacceptable, because production of wafers with mean flow width of 1.725 microns results in excessive scrap costs and can result in further upstream manufacturing problems. How can we reduce the time needed to detect the out-of-control condition? One method is to sample more frequently. For example, if we sample every half hour, then the average time to signal for this scheme is ATS = $\text{ARL}_1\, h = 2(\frac{1}{2}) = 1$; that is, only one hour will elapse (on the average) between the shift and its detection. The second possibility is to increase the sample size. For example, if we use $n = 10$, then Fig. 4-9 shows that the probability of $\bar{x}$ falling between the control limits when the process mean is 1.725 microns is approximately 0.1, so that $p = 0.9$, and from equation 4-2 the out-of-control ARL or ARL_1 is

$$\text{ARL}_1 = \frac{1}{p} = \frac{1}{0.9} = 1.11$$

and, if we sample every hour, the average time to signal is

$$\text{ATS} = \text{ARL}_1 h = 1.11(1) = 1.11 \text{ hours}$$

Thus, the larger sample size would allow the shift to be detected about twice as quickly as the old one. If it became important to detect the shift in the (approximately) first hour after it occurred, two control chart designs would work:

Design 1	**Design 2**
Sample Size: $n = 5$	Sample Size: $n = 10$
Sampling Frequency: every half hour	Sampling Frequency: every hour

To answer the question of sampling frequency more precisely, we must take several factors into account, including the cost of sampling, the losses associated with allowing the process to operate out of control, the rate of production, and the probabilities with which various types of process shifts occur. We discuss various methods for selecting an appropriate sample size and sampling frequency for a control chart in the next four chapters.

4-3.4 Rational Subgroups

A fundamental idea in the use of control charts is the collection of sample data according to what Shewhart called the **rational subgroup** concept. To illustrate this concept, suppose that we are using an $\bar{x}$ control chart to detect changes in the process mean. Then the rational subgroup concept means that subgroups or samples should be selected so that if assignable causes are present, the chance for differences *between* subgroups will be maximized, while the chance for differences due to these assignable causes *within* a subgroup will be minimized.

When control charts are applied to production processes, the time order of production is a logical basis for rational subgrouping. Even though time order is preserved, it is still possible to form subgroups erroneously. If some of the observations in the sample are taken at the end of one shift and the remaining observations are taken at the start of the next shift, then any differences between shifts might not be detected. Time order is frequently a good basis for forming subgroups because it allows us to detect assignable causes that occur over time.

Two general approaches to constructing rational subgroups are used. In the first approach, each sample consists of units that were produced at the same time (or as closely together as possible). Ideally, we would like to take **consecutive** units of production. This approach is used when the primary purpose of the control chart is to detect process shifts. It minimizes the chance of variability due to assignable causes *within* a sample, and it maximizes the chance of variability *between* samples if assignable causes are present. It also provides a better estimate of the standard deviation of the process in the case of variables control charts. This approach to rational subgrouping essentially gives a "snapshot" of the process at each point in time where a sample is collected.

Figure 4-10 illustrates this type of sampling strategy. In Fig. 4-10*a* we show a process for which the mean experiences a series of sustained shifts, and the corresponding observations obtained from this process at the points in time along the horizontal axis, assuming that five consecutive units are selected. Figure 4-10*b* shows the $\bar{x}$ **control chart** and an ***R* chart** (or **range chart**) for these data. The center line and control limits on the R chart are constructed using the range of each sample in the upper part of the figure (details will be given in Chapter 5). Note that although the process mean is shifting, the process variability is stable. Furthermore, the within-sample measure of variability is used to construct the control limits on the $\bar{x}$ chart. Note that the $\bar{x}$ chart in Fig. 4-10*b* has points out of control corresponding to the shifts in the process mean.

In the second approach, each sample consists of units of product that are representative of *all* units that have been produced since the last sample was taken. Essentially, each subgroup is a **random sample of all process output over the sampling interval.** This method of rational subgrouping is often used when the control chart is employed to make decisions about the acceptance of all units of product that have been produced since the last sample. In fact, if the process shifts to an out-of-control state and then back in control again *between* samples, it is sometimes argued that the first method of rational subgrouping defined above will be ineffective against these types of shifts, and so the second method must be used.

When the rational subgroup is a random sample of all units produced over the sampling interval, considerable care must be taken in interpreting the control charts. If the process mean drifts between several levels during the interval between samples, this may cause the range of the observations within the sample to be relatively large, resulting in wider limits on the $\bar{x}$ chart. This scenario is illustrated in Fig. 4-11. In fact, **we can often make any process appear to be in statistical control just by stretching out the interval between observations in the sample.** It is also possible for shifts in the process average to cause points on a control chart for the range or standard deviation to plot out of control, even though there has been no shift in process variability.

There are other bases for forming rational subgroups. For example, suppose a process consists of several machines that pool their output into a common stream. If we sample from this common stream of output, it will be very difficult to detect whether or not some of the machines are out of control. A logical approach to rational subgrouping here is to apply control chart techniques to the output for each individual machine. Sometimes this concept needs to be applied to different heads on the same machine, different work stations, different operators, and so forth. In many situations the rational subgroup will consist of a single observation. This situation occurs frequently in the chemical and process industries where the quality characteristic of the product changes relatively slowly and samples taken very close together in time are virtually identical, apart from measurement or analytical error.

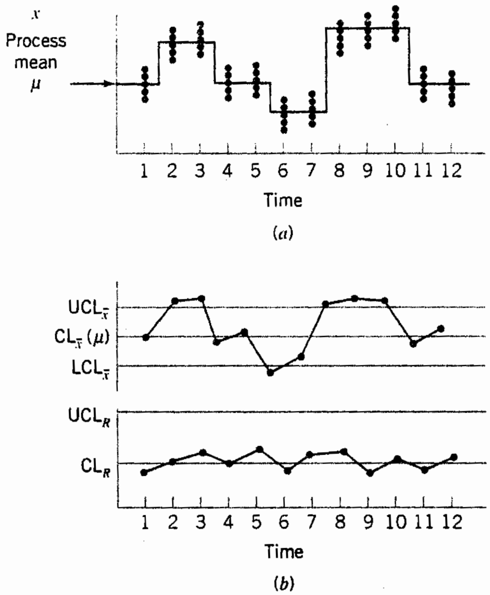

Figure 4-10 The "snapshot" approach to rational subgroups. (*a*) Behavior of the process mean. (*b*) Corresponding $\bar{x}$ and R control charts.

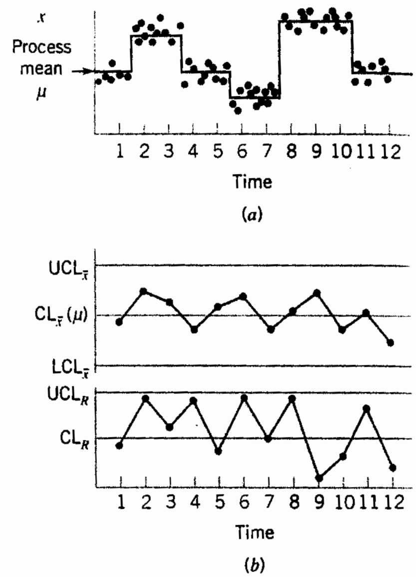

Figure 4-11 The random sample approach to rational subgroups. (*a*) Behavior of the process mean. (*b*) Corresponding $\bar{x}$ and R control charts.

The rational subgroup concept is very important. The proper selection of samples requires careful consideration of the process, with the objective of obtaining as much useful information as possible from the control chart analysis.

4-3.5 Analysis of Patterns on Control Charts

A control chart may indicate an out-of-control condition either when one or more points fall beyond the control limits or when the plotted points exhibit some nonrandom pattern of behavior. For example, consider the $\bar{x}$ chart shown in Fig. 4-12. Although all 25 points fall within the control limits, the points do not indicate statistical control because their pattern is very nonrandom in appearance. Specifically, we note that 19 of 25 points plot below the center line, while only six of them plot above. If the points are truly random, we should expect a more even distribution of them above and below the center line. We also observe that following the fourth point, five points in a row increase in magnitude. This arrangement of points is called a **run.** Since the observations are increasing, we could call this a *run up*. Similarly, a sequence of decreasing points is called a *run down*. This control chart has an unusually long run up (beginning with the fourth point) and an unusually long run down (beginning with the eighteenth point).

In general, we define a run as a sequence of observations of the same type. In addition to runs up and runs down, we could define the types of observations as those above and below the center line, respectively, so that two points in a row above the center line would be a run of length 2.

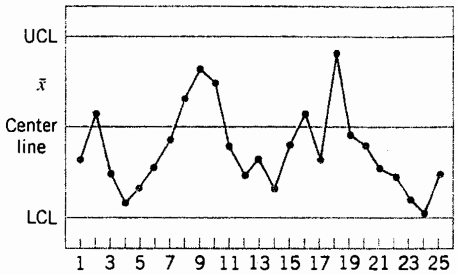

Figure 4-12 An $\bar{x}$ control chart.

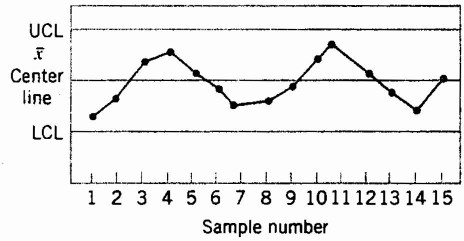

Figure 4-13 An $\bar{x}$ chart with a cyclic pattern.

A run of length 8 or more points has a very low probability of occurrence in a random sample of points. Consequently, any type of run of length 8 or more is often taken as a signal of an out-of-control condition. For example, eight consecutive points on one side of the center line may indicate that the process is out of control.

Although runs are an important measure of nonrandom behavior on a control chart, other types of patterns may also indicate an out-of-control condition. For example, consider the $\bar{x}$ chart in Fig. 4-13. Note that the plotted sample averages exhibit a cyclic behavior, yet they all fall within the control limits. Such a pattern may indicate a problem with the process such as operator fatigue, raw material deliveries, heat or stress buildup, and so forth. Although the process is not really out of control, the yield may be improved by elimination or reduction of the sources of variability causing this cyclic behavior (see Fig. 4-14).

The problem is one of **pattern recognition**—that is, recognizing systematic or nonrandom patterns on the control chart and identifying the reason for this behavior. The ability to interpret a particular pattern in terms of assignable causes requires experience and knowledge of the process. That is, we must not only know the statistical principles of control charts, but we must also have a good understanding of the process. We discuss the interpretation of patterns on control charts in more detail in Chapter 5.

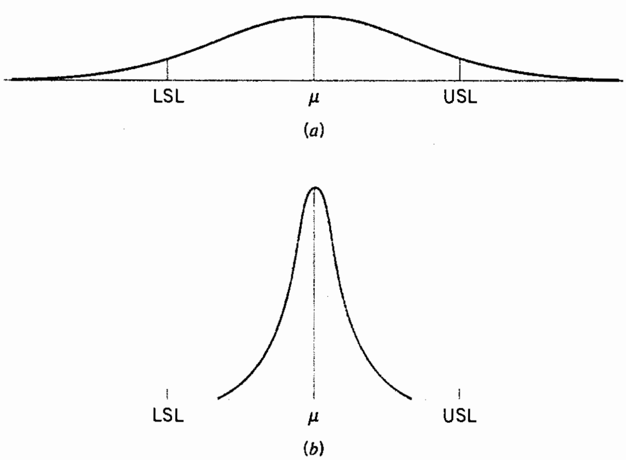

Figure 4-14 (*a*) Variability with the cyclic pattern. (*b*) Variability with the cyclic pattern eliminated.

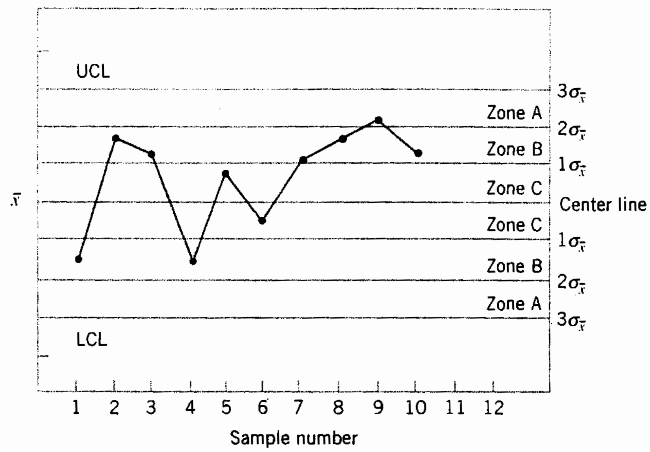

Figure 4-15 The Western Electric or zone rules, with the last four points showing a violation of the rule 3.

The Western Electric Handbook (1956) suggests a set of decision rules for detecting nonrandom patterns on control charts. Specifically, it suggests concluding that the process is out of control if either

1. One point plots outside the three-sigma control limits;
2. Two out of three consecutive points plot beyond the two-sigma warning limits;
3. Four out of five consecutive points plot at a distance of one-sigma or beyond from the center line;

or

4. Eight consecutive points plot on one side of the center line.

Those rules apply to one side of the center line at a time. Therefore, a point above the *upper* warning limit followed immediately by a point below the *lower* warning limit would not signal an out-of-control alarm. These are often used in practice for enhancing the sensitivity of control charts. That is, the use of these rules can allow smaller process shifts to be detected more quickly than would be the case if our only criterion was the usual three-sigma control limit violation.

Figure 4-15 shows an $\bar{x}$ control chart with the one-sigma, two-sigma, and three-sigma limits used in the Western Electric procedure. Note that these limits partition the control chart into three zones A, B, and C on each side of the center line. Consequently, the Western Electric rules are sometimes called the **zone rules** for control charts. Note that the last four points fall in zone B or beyond. Thus, since four of five consecutive points exceed the one-sigma limit, the Western Electric procedure will conclude that the pattern is nonrandom and the process is out of control.

4-3.6 Discussion of Sensitizing Rules for Control Charts

As may be gathered from earlier sections, several criteria may be applied simultaneously to a control chart to determine whether the process is out of control. The basic criterion is one or more points outside of the control limits. The supplementary criteria are sometimes used to increase the **sensitivity** of the control charts to a small process shift so that we may

Table 4-1 Some Sensitizing Rules for Shewhart Control Charts

Standard Action Signal:	1. One or more points outside of the control limits.	Western Electric Rules
	2. Two of three consecutive points outside the two-sigma warning limits but still inside the control limits.	Western Electric Rules
	3. Four of five consecutive points beyond the one-sigma limits.	Western Electric Rules
	4. A run of eight consecutive points on one side of the center line.	Western Electric Rules
	5. Six points in a row steadily increasing or decreasing.	
	6. Fifteen points in a row in zone C (both above and below the center line).	
	7. Fourteen points in a row alternating up and down.	
	8. Eight points in a row on both sides of the center line with none in zone C.	
	9. An unusual or nonrandom pattern in the data.	
	10. One or more points near a warning or control limit.	

respond more quickly to the assignable cause. Some of the **sensitizing rules** that are widely used in practice are shown in Table 4-1. For a good discussion of some of these rules, see Nelson (1984). Frequently, we will inspect the control chart and conclude that the process is out of control if any one or more of the criteria in Table 4-1 are met.

When several of these sensitizing rules are applied simultaneously, we often use a **graduated response** to out-of-control signals. For example, if a point exceeded a control limit, we would immediately begin to search for the assignable cause, but if one or two consecutive points exceeded only the two-sigma warning limit, we might increase the frequency of sampling from every hour—say, to every 10 minutes. This **adaptive sampling** response might not be as severe as a complete search for an assignable cause, but if the process were really out of control, it would give us a high probability of detecting this situation more quickly than we would by maintaining the longer sampling interval.

In general, care should be exercised when using several decision rules simultaneously. Suppose that the analyst uses k decision rules and that criterion i has type I error probability α_i. Then the overall type I error or false-alarm probability for the decision based on all k tests is

$$\alpha = 1 - \prod_{i=1}^{k}\left(1 - \alpha_i\right) \qquad (4\text{-}4)$$

provided that all k decision rules are independent. However, the independence assumption is not valid with the usual sensitizing rules. Furthermore, the value of α_i is not always clearly defined for the sensitizing rules, because these rules involve several observations.

Champ and Woodall (1987) investigated the average run length performance for the Shewhart control chart with various sensitizing rules. They found that the use of these rules does improve the ability of the control chart to detect smaller shifts, but the in-control average run length can be substantially degraded. For example, assuming independent process data and using a Shewhart control chart with the Western Electric rules results in an in-control ARL of 91.25, in contrast to 370 for the Shewhart control chart alone.

Some of the individual Western Electric rules are particularly troublesome. An illustration is the rule of several (usually seven or eight) consecutive points which either increase or decrease. This rule is very ineffective in detecting a trend, the situation for which it was designed. It does, however, greatly increase the false-alarm rate. See Davis and Woodall (1988) for more details.

4-3.7 Phase I and Phase II of Control Chart Application

Standard control chart usage involves two distinct phases, with two different objectives. In phase I, a set of process data is gathered and analyzed all at once in a **retrospective** analysis, constructing **trial control limits** to determine if the process has been in control over the period of time where the data were collected, and to see if reliable control limits can be established to monitor future production. This is typically the very first thing that is done when control charts are applied to any process. Control charts are used primarily in phase I to assist operating personnel in bringing the process into a state of statistical control. Phase II beings after we have a "clean" set of process data gathered under stable conditions and representative of in-control process performance. In phase II, we use the control chart to **monitor** the process by comparing the sample statistic for each successive sample as it is drawn from the process to the control limits.

Thus in phase I, we are comparing a collection of, say, m points to a set of control limits computed from those points. Typically $m = 20$ or 25 subgroups are used in phase I. It is fairly typical in phase I to assume that the process is initially out of control, so the objective of the analyst is to bring the process into a state of statistical control. Control limits are calculated based on the m subgroups and the data plotted on the control charts. Points that are outside the control limits are investigated, looking for potential assignable causes. Any assignable causes that are identified are worked on by engineering and operating personnel in an effort to eliminate them. Points outside the control limits are then excluded and a new set of revised control limits calculated. Then new data are collected and compared to these revised limits. Sometimes this type of analysis will require several cycles in which the control chart is employed, assignable causes are detected and corrected, revised control limits are calculated, and the out-of-control action plan is updated and expanded. Eventually the process is stabilized, and a clean set of data that represents in-control process performance is obtained.

Generally, Shewhart control charts are very effective in phase I because they are easy to construct and interpret, and because they are effective in detecting both large, sustained shifts in the process parameters and outliers (single excursions that may have resulted from assignable causes of short duration), measurement errors, data recording and/or transmission errors, and the like. Furthermore, patterns on Shewhart control charts are often easy to interpret and have direct physical meaning. The sensitizing rules discussed in the previous sections are also easy to apply to Shewhart charts. (This is an optional feature in most SPC software.) The types of assignable causes that usually occur in Phase I result in fairly large process shifts—exactly the scenario in which the Shewhart control chart is most effective. Average run length is not usually a reasonable performance measure for phase I; we are typically more interested in the probability that an assignable cause will be detected than in the occurrence of false alarms. For good discussions of phase I control chart usage and related matters, see the papers by Woodall (2000), Borror and Champ (2001), Boyles (2000), and Champ and Chou (2003), and the standard ANSI/ASQC B1–133–1996 Quality Control Chart Methodologies (this can be downloaded at http://e-standards.asq.org).

In phase II we usually assume that the process is reasonably stable. Often, the assignable causes that occur in phase II result in smaller process shifts, because (hopefully) most of the really ugly sources of variability have been systematically removed during phase I. Our emphasis is now on **process monitoring,** not on bringing an unruly process into control. Average run length is a valid basis for evaluating the performance of a control chart in phase II. Shewhart control charts are much less likely to be effective in phase II because they are not very sensitive to small to moderate size process shifts; that is, their ARL performance is relatively poor. Attempts to solve this problem by employing sensitizing rules such as those discussed in the previous section are likely to be unsatisfactory, because the use of these supplemental sensitizing rules increases the false-alarm rate of the Shewhart control chart. (Recall the discussion of the Champ and Woodall (1987) paper in the previous section.) The routine use of sensitizing rules to detect small shifts or to react more quickly to assignable causes in phase II should be discouraged. The cumulative sum and EWMA control charts discussed in Chapter 8 are much more likely to be effective in phase II.

4-4 THE REST OF THE "MAGNIFICENT SEVEN"

Although the control chart is a very powerful problem-solving and process improvement tool, it is most effective when its use is fully integrated into a comprehensive SPC program. The seven major SPC problem-solving tools should be widely taught throughout the organization and used routinely to identify improvement opportunities and to assist in reducing variability and eliminating waste. These "magnificent seven," introduced in Section 4-1, are listed again here for convenience:

1. Histogram or stem-and-leaf plot
2. Check sheet
3. Pareto chart
4. Cause-and-effect diagram
5. Defect concentration diagram
6. Scatter diagram
7. Control chart

We have already introduced the histogram and the stem-and-leaf plot (Chapter 2), and control chart. In this section we will briefly illustrate the rest of the tools.

Check Sheet

In the early stages of an SPC implementation, it will often become necessary to collect either historical or current operating data about the process under investigation. A **check sheet** can be very useful in this data collection activity. The check sheet shown in Fig. 4-16 was developed by an engineer at an aerospace firm who was investigating the various types of defects that occurred on a tank used in one of their products with a view toward improving the process. The engineer designed this check sheet to facilitate summarizing

10 Multivariate Process Monitoring and Control

CHAPTER OUTLINE

The supplemental material is on the textbook website www.wiley.com/college/montgomery.

CHAPTER OVERVIEW AND LEARNING OBJECTIVES

In previous chapters we have addressed process monitoring and control primarily from the univariate perspective; that is, we have assumed that there is only one process output variable or quality characteristic of interest. In practice, however, many if not most process monitoring and control scenarios involve several related variables. Although applying univariate control charts to each individual variable is a possible solution, we will see that this is inefficient and can lead to erroneous conclusions. Multivariate methods that consider the variables jointly are required.

This chapter presents control charts that can be regarded as the multivariate extensions of some of the univariate charts of previous chapters. The Hotelling T^2 chart is the analog of the Shewhart $\bar{x}$ chart. We will also discuss a multivariate version of the EWMA control chart, and some methods for monitoring variability in the multivariate case. These multivariate control charts work well when the number of process variables is not too large—say, 10 or fewer. As the number of variables grows, however, traditional multivariate control charts lose efficiency with regard to shift detection. A popular approach in these situations is to reduce the dimensionality of the problem. We show how this can be done with principal components.

After careful study of this chapter you should be able to do the following:

1. Understand why applying several univariate control charts simultaneously to a set of related quality characteristics may be an unsatisfactory monitoring procedure
2. How the multivariate normal distribution is used as a model for multivariate process data
3. Know how to estimate the mean vector and covariance matrix from a sample of multivariate observations
4. Know how to set up and use a chi-square control chart
5. Know how to set up and use the Hotelling T^2 control chart
6. Know how to set up and use the multivariate exponentially weighted moving average (MEWMA) control chart
7. Know how to use multivariate control charts for individual observations
8. Know how to find the phase I and phase II limits for multivariate control charts
9. Use control charts for monitoring multivariate variability
10. Understand the basis of the regression adjustment procedure and know how to apply regression adjustment in process monitoring
11. Understand the basis of principal components and know how to apply principal components in process monitoring
12. Understand the basis of profile monitoring

10-1 THE MULTIVARIATE QUALITY-CONTROL PROBLEM

There are many situations in which the **simultaneous monitoring** or control of two or more related quality characteristics is necessary. For example, suppose that a bearing has both an inner diameter (x_1) and an outer diameter (x_2) that together determine the usefulness of the part. Suppose that x_1 and x_2 have independent normal distributions. Because both quality characteristics are measurements, they could be monitored by applying the usual $\bar{x}$ chart to each characteristic, as illustrated in Fig. 10-1. The process is considered to be in control only if the sample means $\bar{x}_1$ and $\bar{x}_2$ fall within their respective control limits. This is equivalent to the pair of means ($\bar{x}_1$, $\bar{x}_2$) plotting within the shaded region in Fig. 10-2.

Monitoring these two quality characteristics independently can be very misleading. For example, note from Fig. 10-2 that one observation appears somewhat unusual with respect to the others. That point would be inside the control limits on both of the univariate $\bar{x}$ charts for x_1 and x_2, yet when we examine the two variables **simultaneously,** the unusual behavior of the point is fairly obvious. Furthermore, note that the probability that either $\bar{x}_1$ or $\bar{x}_2$ exceeds three-sigma control limits is 0.0027. However, the joint probability that both variables exceed their control limits simultaneously when they are both in control is (0.0027)(0.0027) = 0.00000729, which is considerably smaller than 0.0027. Furthermore, the probability that both $\bar{x}_1$ and $\bar{x}_2$ will simultaneously plot inside the control limits when the process is really in control is (0.9973)(0.9973) = 0.99460729. Therefore, the use of two independent $\bar{x}$ charts has distorted the simultaneous monitoring of $\bar{x}_1$ and $\bar{x}_2$, in that the type I error and the probability of a point correctly plotting in control are not equal to their advertised levels for the individual control charts. However, note that because the variables are independent the univariate control chart limits could be adjusted to account for this.

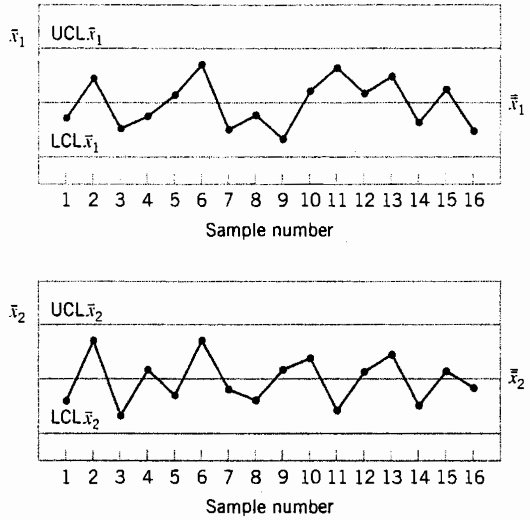

Figure 10-1 Control charts for inner ($\bar{x}_1$) and outer ($\bar{x}_2$) bearing diameters.

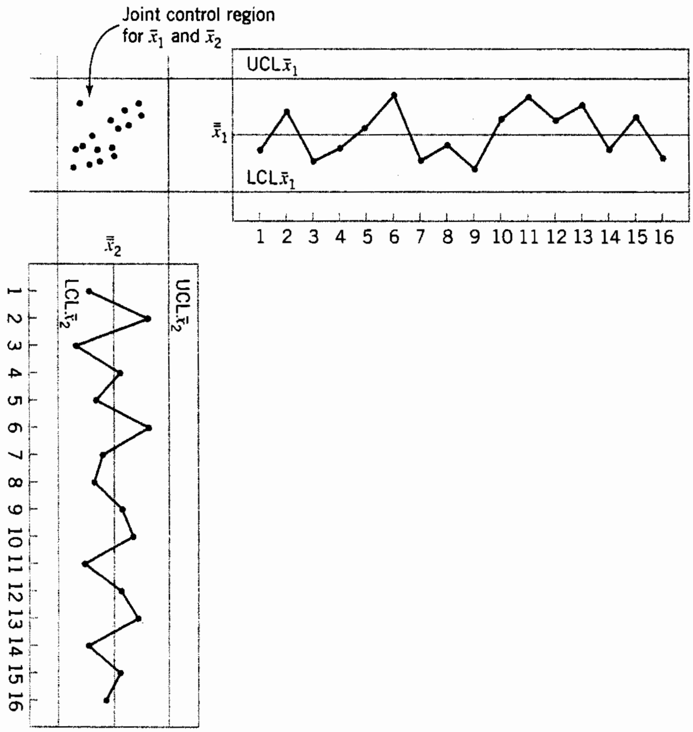

Figure 10-2 Control region using independent control limits for $\bar{x}_1$ and $\bar{x}_2$.

This distortion in the process-monitoring procedure increases as the number of quality characteristics increases. In general, if there are p statistically independent quality characteristics for a particular product and if an $\bar{x}$ chart with $P\{\text{type I error}\} = \alpha$ is maintained on each, then the true probability of type I error for the joint control procedure is

$$\alpha' = 1 - (1 - \alpha)^p \tag{10-1}$$

and the probability that all p means will simultaneously plot inside their control limits when the process is in control is

$$P\{\text{all } p \text{ means plot in control}\} = (1 - \alpha)^p \tag{10-2}$$

Clearly, the distortion in the joint control procedure can be severe, even for moderate values of p. Furthermore, if the p quality characteristics are not independent, which usually would be the case if they relate to the same product, then equations 10-1 and 10-2 do not hold, and we have no easy way even to measure the distortion in the joint control procedure.

Process-monitoring problems in which several related variables are of interest are sometimes called **multivariate quality-control** (or **process-monitoring**) problems. The original work in multivariate quality control was done by Hotelling (1947), who applied his procedures to bombsight data during World War II. Subsequent papers dealing with control procedures for several related variables include Hicks (1955), Jackson (1956, 1959, 1985), Crosier (1988), Hawkins (1991, 1993b), Lowry et al. (1992), Lowry and Montgomery (1995), Pignatiello and Runger (1990), Tracy, Young, and Mason (1992), Montgomery and Wadsworth (1972); and Alt (1985). This subject is particularly important today, as automatic inspection procedures make it relatively easy to measure many parameters on each unit of product manufactured. Many chemical and process plants and semiconductor manufacturers (for examples) routinely maintain manufacturing databases with process and quality data on hundreds of variables. Often the total size of these databases is measured in millions of individual records. Monitoring or analysis of these data with univariate SPC procedures is often ineffective. The use of multivariate methods has increased greatly in recent years for this reason.

10-2 DESCRIPTION OF MULTIVARIATE DATA

10-2.1 The Multivariate Normal Distribution

In univariate statistical quality control, we generally use the **normal distribution** to describe the behavior of a continuous quality characteristic. The univariate normal probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty \tag{10-3}$$

The mean of the normal distribution is μ and the variance is σ^2. Note that (apart from the minus sign) the term in the exponent of the normal distribution can be written as follows:

$$(x - \mu)\left(\sigma^2\right)^{-1}(x - \mu) \tag{10-4}$$

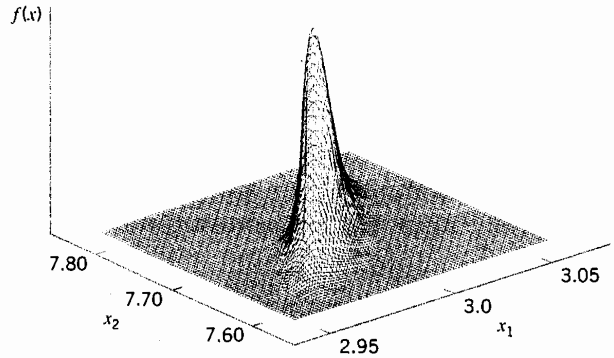

Figure 10-3 A multivariate normal distribution with $p = 2$ variables (bivariate normal).

This quantity measures the squared standardized distance from x to the mean μ, where by the term "standardized" we mean that the distance is expressed in standard deviation units.

This same approach can be used in the multivariate case. Suppose that we have p variables, given by $x_1, x_2, \ldots, x_p$. Arrange these variables in a p-component vector $\mathbf{x}' = [x_1, x_2, \ldots, x_p]$. Let $\boldsymbol{\mu}' = [\mu_1, \mu_2, \ldots, \mu_p]$ be the vector of the means of the x's, and let the variances and covariances of the random variables in $\mathbf{x}$ be contained in a $p \times p$ **covariance matrix** Σ. The main diagonal elements of Σ are the variances of the x's and the off-diagonal elements are the covariances. Now the squared standardized (generalized) distance from $\mathbf{x}$ to $\boldsymbol{\mu}$ is

$$(\mathbf{x}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu}) \tag{10-5}$$

The multivariate normal density function is obtained simply by replacing the standardized distance in equation 10-4 by the multivariate generalized distance in equation 10-5 and changing the constant term $1/\sqrt{2\pi\sigma^2}$ to a more general form that makes the area under the probability density function unity regardless of the value of p. Therefore, the **multivariate normal** probability density function is

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{m})'\mathbf{S}^{-1}(\mathbf{x}-\mathbf{m})} \tag{10-6}$$

where $-\infty < x_j < \infty$, $j = 1, 2, \ldots, p$.

A multivariate normal distribution for $p = 2$ variables (called a **bivariate normal**) is shown in Fig. 10-3. Note that the density function is a surface. The correlation coefficient between the two variables in this example is 0.8, and this causes the probability to concentrate closely along a line.

10-2.2 The Sample Mean Vector and Covariance Matrix

Suppose that we have a random sample from a multivariate normal distribution—say,

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$$

where the ith sample vector contains observations on each of the p variables $x_{i1}, x_{i2}, \ldots, x_{ip}$. Then the sample mean vector is

$$\bar{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i \tag{10-7}$$

and the sample covariance matrix is

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})' \tag{10-8}$$

That is, the sample variances on the main diagonal of the matrix **S** are computed as

$$s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)^2 \tag{10-9}$$

and the sample covariances are

$$s_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ij} - \bar{x}_j\right)\left(x_{ik} - \bar{x}_k\right) \tag{10-10}$$

We can show that the sample mean vector and sample covariance matrix are unbiased estimators of the corresponding population quantities; that is,

$$E(\bar{\mathbf{x}}) = \boldsymbol{\mu} \quad \text{and} \quad E(\mathbf{S}) = \boldsymbol{\Sigma}$$

10-3 THE HOTELLING T^2 CONTROL CHART

The most familiar multivariate process-monitoring and control procedure is the Hotelling T^2 control chart for monitoring the mean vector of the process. It is a direct analog of the univariate Shewhart $\bar{x}$ chart. We present two versions of the Hotelling T^2 chart: one for subgrouped data, and another for individual observations.

10-3.1 Subgrouped Data

Suppose that two quality characteristics x_1 and x_2 are jointly distributed according to the bivariate normal distribution (see Fig. 10-3). Let μ_1 and μ_2 be the mean values of the quality characteristics, and let σ_1 and σ_2 be the standard deviations of x_1 and x_2, respectively. The covariance between x_1 and x_2 is denoted by σ_{12}. We assume that σ_1, σ_2, and σ_{12} are known. If $\bar{x}_1$ and $\bar{x}_2$ are the sample averages of the two quality characteristics computed from a sample of size n, then the statistic

$$\chi_0^2 = \frac{n}{\sigma_1^2\sigma_2^2 - \sigma_{12}^2}\left[\sigma_2^2(\bar{x}_1 - \mu_1)^2 + \sigma_1^2(\bar{x}_2 - \mu_2)^2 - 2\sigma_{12}(\bar{x}_1 - \mu_1)(\bar{x}_2 - \mu_2)\right] \tag{10-11}$$

will have a chi-square distribution with 2 degrees of freedom. This equation can be used as the basis of a control chart for the process means μ_1 and μ_2. If the process means remain at the values μ_1 and μ_2, then values of χ_0^2 should be less than the upper control limit $\text{UCL} = \chi_{\alpha,2}^2$ where $\chi_{\alpha,2}^2$ is the upper α percentage point of the chi-square distribution with 2 degrees of freedom. If at least one of the means shifts to some new (out-of-control) value, then the probability that the statistic χ_0^2 exceeds the upper control limit increases.

The process monitoring procedure may be represented graphically. Consider the case in which the two random variables x_1 and x_2 are independent; that is, $\sigma_{12} = 0$. If $\sigma_{12} = 0$, then equation 10-11 defines an ellipse centered at (μ_1, μ_2) with principal axes parallel to the $\bar{x}_1$, $\bar{x}_2$ axes, as shown in Fig. 10-4. Taking χ_0^2 in equation 10-11 equal to $\chi_{\alpha,2}^2$ implies that a pair of sample averages $(\bar{x}_1, \bar{x}_2)$ yielding a value of χ_0^2 plotting inside the ellipse indicates that the process is in control, whereas if the corresponding value of χ_0^2 plots outside the ellipse the process is out of control. Figure 10-4 is often called a **control ellipse.**

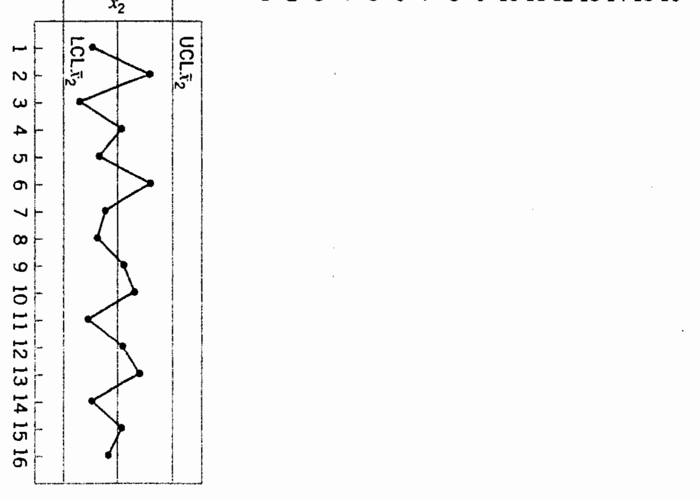

Figure 10-4 A control ellipse for two independent variables.

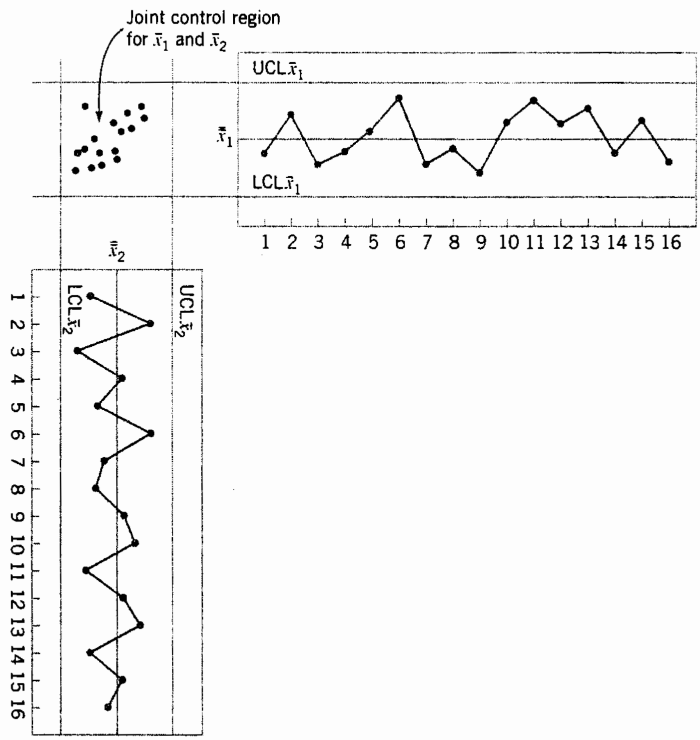

Figure 10-5 A control ellipse for two dependent variables.

In the case where the two quality characteristics are dependent, then $\sigma_{12} \neq 0$, and the corresponding control ellipse is shown in Fig. 10-5. When the two variables are dependent, the principal axes of the ellipse are no longer parallel to the $\bar{x}_1, \bar{x}_2$ axes. Also, note that sample point number 11 plots outside the control ellipse, indicating that an assignable cause is present, yet point 11 is inside the control limits on both of the individual control charts for $\bar{x}_1$ and $\bar{x}_2$. Thus there is nothing apparently unusual about point 11 when viewed individually, yet the customer who received that shipment of material would quite likely observe very different performance in the product. It is nearly impossible to detect an assignable cause resulting in a point such as this one by maintaining individual control charts.

Two disadvantages are associated with the control ellipse. The first is that the time sequence of the plotted points is lost. This could be overcome by numbering the plotted points or by using special plotting symbols to represent the most recent observations. The second and more serious disadvantage is that it is difficult to construct the ellipse for more than two quality characteristics. To avoid these difficulties, it is customary to plot the values of χ_0^2 computed from equation 10-11 for each sample on a control chart with only an upper control limit at $\chi^2_{\alpha, 2}$, as shown in Fig. 10-6. This control chart is usually called the **chi-square control chart.** Note that the time sequence of the data is preserved by this control chart, so that runs or other nonrandom patterns can be investigated. Furthermore, it has the additional advantage that the "state" of the process is characterized by a single number (the value of the statistic χ_0^2). This is particularly helpful when there are two or more quality characteristics of interest.

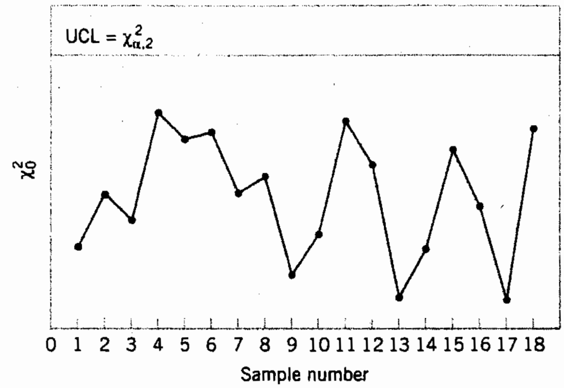

Figure 10-6 A chi-square control chart for $p = 2$ quality characteristics.

It is possible to extend these results to the case where p related quality characteristics are controlled jointly. It is assumed that the joint probability distribution of the p quality characteristics is the p-variate normal distribution. The procedure requires computing the sample mean for each of the p quality characteristics from a sample of size n. This set of quality characteristic means is represented by the $p \times 1$ vector

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}$$

The test statistic plotted on the chi-square control chart for each sample is

$$\chi_0^2 = n(\bar{\mathbf{x}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \qquad \text{(10-12)}$$

where $\boldsymbol{\mu}' = \mu_1, \mu_2, \ldots, \mu_p]$ is the vector of in-control means for each quality characteristic and $\boldsymbol{\Sigma}$ is the covariance matrix. The upper limit on the control chart is

$$\text{UCL} = \chi^2_{\alpha,p} \qquad \text{(10-13)}$$

Estimating μ and Σ

In practice, it is usually necessary to estimate μ and Σ from the analysis of preliminary samples of size n, taken when the process is assumed to be in control. Suppose that m such samples are available. The sample means and variances are calculated from each sample as usual; that is,

$$\bar{x}_{jk} = \frac{1}{n}\sum_{i=1}^{n} x_{ijk} \qquad \begin{cases} j = 1,2,\ldots,p \\ k = 1,2,\ldots,m \end{cases} \tag{10-14}$$

$$s_{jk}^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ijk} - \bar{x}_{jk}\right)^2 \qquad \begin{cases} j = 1,2,\ldots,p \\ k = 1,2,\ldots,m \end{cases} \tag{10-15}$$

where x_{ijk} is the ith observation on the jth quality characteristic in the kth sample. The covariance between quality characteristic j and quality characteristic h in the kth sample is

$$s_{jhk} = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ijk} - \bar{x}_{jk}\right)\left(x_{ihk} - \bar{x}_{hk}\right) \qquad \begin{cases} k = 1,2,\ldots,m \\ j \neq h \end{cases} \tag{10-16}$$

The statistics $\bar{x}_{jk}$, s_{jk}^2, and s_{jhk} are then averaged over all m samples to obtain

$$\bar{\bar{x}}_j = \frac{1}{m}\sum_{k=1}^{m}\bar{x}_{jk} \qquad j = 1,2,\ldots,p \tag{10-17a}$$

$$\bar{s}_j^2 = \frac{1}{m}\sum_{k=1}^{m} s_{jk}^2 \qquad j = 1,2,\ldots,p \tag{10-17b}$$

and

$$\bar{s}_{jh} = \frac{1}{m}\sum_{k=1}^{m} s_{jhk} \qquad j \neq h \tag{10-17c}$$

The $\{\bar{\bar{x}}_j\}$ are the elements of the vector $\bar{\bar{\mathbf{x}}}$, and the $p \times p$ average of sample covariance matrices $\mathbf{S}$ is formed as

$$\mathbf{S} = \begin{bmatrix} \bar{s}_1^2 & \bar{s}_{12} & \bar{s}_{13} & \cdots & \bar{s}_{1p} \\ & \bar{s}_2^2 & \bar{s}_{23} & \cdots & \bar{s}_{2p} \\ & & \bar{s}_3^2 & & \vdots \\ & & & \ddots & \bar{s}_p^2 \end{bmatrix} \tag{10-18}$$

The average of the sample covariance matrices $\mathbf{S}$ is an unbiased estimate of Σ when the process is in control.

The T^2 Control Chart

Now suppose that $\mathbf{S}$ from equation 10-18 is used to estimate $\boldsymbol{\Sigma}$ and that the vector $\bar{\bar{\mathbf{x}}}$ is taken as the in-control value of the mean vector of the process. If we replace $\boldsymbol{\mu}$ with $\bar{\bar{\mathbf{x}}}$ and $\boldsymbol{\Sigma}$ with $\mathbf{S}$ in equation 10-12, the test statistic now becomes

$$T^2 = n\left(\bar{\mathbf{x}} - \bar{\bar{\mathbf{x}}}\right)' \mathbf{S}^{-1}\left(\bar{\mathbf{x}} - \bar{\bar{\mathbf{x}}}\right) \tag{10-19}$$

In this form, the procedure is usually called the **Hotelling T^2 control chart.**

Alt (1985) has pointed out that in multivariate quality control applications one must be careful to select the control limits for Hotelling's T^2 statistic (equation 10-19) based on how the chart is being used. Recall that there are two distinct phases of control chart usage. **Phase I** is the use of the charts for establishing control; that is, testing whether the process was in control when the m preliminary subgroups were drawn and the sample statistics $\bar{\bar{\mathbf{x}}}$ and $\mathbf{S}$ computed. The objective in phase I is to obtain an in-control set of observations so that control limits can be established for phase II, which is the monitoring of future production. Phase I analysis is sometimes called a **retrospective analysis.**

The phase I control limits for the T^2 control chart are given by

$$\text{UCL} = \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$
$$\text{LCL} = 0 \tag{10-20}$$

In **phase II,** when the chart is used for monitoring future production, the control limits are as follows:

$$\text{UCL} = \frac{p(m+1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1}$$
$$\text{LCL} = 0 \tag{10-21}$$

Note that the UCL in equation 10-21 is just the UCL in equation 10-20 multiplied by $m + 1/m - 1$.

When $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are estimated from a large number of preliminary samples, it is customary to use $\text{UCL} = \chi^2_{\alpha,p}$ as the upper control limit in both phase I and phase II.

Retrospective analysis of the preliminary samples to test for statistical control and establish control limits also occurs in the univariate control chart setting. For the $\bar{x}$ chart, it is well known that if we use $m \geq 20$ or 25 preliminary samples, the distinction between phase I and phase II limits is usually unnecessary, because the phase I and phase II limits will nearly coincide. However, with multivariate control charts, we must be careful.

Lowry and Montgomery (1995) show that in many situations a large number of preliminary samples would be required before the exact phase II control limits are well approximated by the chi-square limits. These authors present tables indicating the recommended minimum value of m for sample sizes of $n = 3$, 5, and 10 and for $p = 2$, 3, 4, 5, 10, and 20 quality characteristics. The recommended values of m are always greater than 20 preliminary samples, and often more than 50 samples.

EXAMPLE 10-1

The tensile strength and diameter of a textile fiber are two important quality characteristics that are to be jointly controlled. The quality engineer has decided to use $n = 10$ fiber specimens in each sample. He has taken 20 preliminary samples, and on the basis of these data he concludes that $\bar{\bar{x}}_1 = 115.59$ psi, $\bar{\bar{x}}_2 = 1.06\ (\times 10^{-2})$ inch, $\bar{s}_1^2 = 1.23$, $\bar{s}_2^2 = 0.83$, and $\bar{s}_{12} = 0.79$. Therefore, the statistic he will use for process-control purposes is

$$T^2 = \frac{10}{(1.23)(0.83)-(0.79)^2}\Big[0.83(\bar{x}_1 - 115.59)^2 + 1.23(\bar{x}_2 - 1.06)^2$$
$$-2(0.79)(\bar{x}_1 - 115.59)(\bar{x}_2 - 1.06)\Big]$$

The data used in this analysis and the summary statistics are in Table 10-1, panels (a) and (b).

Figure 10-7 presents the Hotelling T^2 control chart for this example. We will consider this to be phase I, establishing statistical control in the preliminary samples, and calculate the upper control limit from equation 10-20. If $\alpha = 0.001$, then the UCL is

$$\begin{aligned}
\text{UCL} &= \frac{p(m-1)(n-1)}{mn-m-p+1} F_{\alpha,p,mn-m-p+1} \\
&= \frac{2(19)(9)}{20(10)-20-2+1} F_{0.001,2,20(10)-20-2+1} \\
&= \frac{342}{179} F_{0.001,2,179} \\
&= (1.91)7.18 \\
&= 13.72
\end{aligned}$$

This control limit is shown on the chart in Fig. 10-7. Notice that no points exceed this limit, so we would conclude that the process is in control. Phase II control limits could be calculated from equation 10-21. If $\alpha = 0.001$, the upper control limit is UCL = 15.16. If we had used the approximate chi-square control limit, we would have obtained $\chi^2_{0.001,2} =$ 13.816, which is reasonably close to the correct limit for phase I but somewhat too small for phase II.

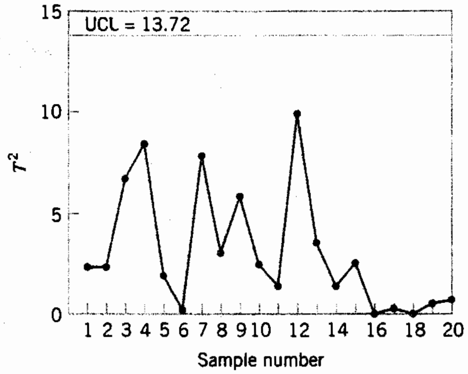

Figure 10-7 The Hotelling T^2 control chart for tensile strength and diameter, Example 10-1.

Table 10-1 Data for Example 10-1

Sample Number k	(a) Means		(b) Variances and Covariances			(c) Control Chart Statistics	
	Tensile Strength ($\bar{x}_{1k}$)	Diameter ($\bar{x}_{2k}$)	s_{1k}^2	s_{2k}^2	s_{12k}	T_k^2	$\lvert \mathbf{S}_k \rvert$
1	115.25	1.04	1.25	0.87	0.80	2.16	0.45
2	115.91	1.06	1.26	0.85	0.81	2.14	0.41
3	115.05	1.09	1.30	0.90	0.82	6.77	0.50
4	116.21	1.05	1.02	0.85	0.81	8.29	0.21
5	115.90	1.07	1.16	0.73	0.80	1.89	0.21
6	115.55	1.06	1.01	0.80	0.76	0.03	0.23
7	114.98	1.05	1.25	0.78	0.75	7.54	0.41
8	115.25	1.10	1.40	0.83	0.80	3.01	0.52
9	116.15	1.09	1.19	0.87	0.83	5.92	0.35
10	115.92	1.05	1.17	0.86	0.95	2.41	0.10
11	115.75	0.99	1.45	0.79	0.78	1.13	0.54
12	114.90	1.06	1.24	0.82	0.81	9.96	0.36
13	116.01	1.05	1.26	0.55	0.72	3.86	0.17
14	115.83	1.07	1.17	0.76	0.75	1.11	0.33
15	115.29	1.11	1.23	0.89	0.82	2.56	0.42
16	115.63	1.04	1.24	0.91	0.83	0.08	0.44
17	115.47	1.03	1.20	0.95	0.70	0.19	0.65
18	115.58	1.05	1.18	0.83	0.79	0.00	0.36
19	115.72	1.06	1.31	0.89	0.76	0.35	0.59
20	115.40	1.04	1.29	0.85	0.68	0.62	0.63
Averages	$\bar{\bar{x}}_1 = 115.59$	$\bar{\bar{x}}_2 = 1.06$	$\bar{s}_1^2 = 1.23$	$\bar{s}_2^2 = 0.83$	$\bar{s}_{12} = 0.79$		

The widespread interest in multivariate quality control has led to including the Hotelling T^2 control chart in some software packages. These programs should be used carefully, as they sometimes use an incorrect formula for calculating the control limit. Specifically, some packages use

$$\text{UCL} = \frac{p(m-1)}{m-p} F_{\alpha, p, m-p}$$

This control limit is obviously incorrect. This is the correct critical region to use in multivariate statistical hypothesis testing on the mean vector $\boldsymbol{\mu}$, where a sample of size m is taken at random from a p-dimensional normal distribution, but it is not directly applicable to the control chart for either phase I or phase II problems.

Interpretation of Out-of-Control Signals

One difficulty encountered with any multivariate control chart is practical **interpretation** of an out-of-control signal. Specifically, which of the p variables (or which *subset* of them) is responsible for the signal? This question is not always easy to answer. The standard practice is to plot **univariate** $\bar{x}$ charts on the individual variables $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_p$. However, this approach may not be successful, for reasons discussed previously. Alt (1985) suggests using $\bar{x}$ charts with Bonferroni-type control limits [i.e., replace $Z_{\alpha/2}$ in the $\bar{x}$ chart control limit calculation with $Z_{\alpha/(2p)}$]. This approach reduces the number of false alarms associated with using many simultaneous univariate control charts. Hayter and Tsui (1994) extend this idea by giving a procedure for exact simultaneous confidence intervals. Their procedure can also be used in situations where the normality assumption is not valid. Jackson (1980) recommends using control charts based on the p principal components (which are linear combinations of the original variables). Principal components are discussed in Section 10-7. The disadvantage of this approach is that the principal components do not always provide a clear interpretation of the situation with respect to the original variables. However, they are often effective in diagnosing an out-of-control signal, particularly in cases where the principal components do have an interpretation in terms of the original variables.

Another very useful approach to diagnosis of an out-of-control signal is to decompose the T^2 statistic into components that reflect the contribution of each individual variable. If T^2 is the current value of the statistic, and $T^2_{(i)}$ is the value of the statistic for all process variables except the ith one, then Runger, Alt, and Montgomery (1996b) show that

$$d_i = T^2 - T^2_{(i)} \tag{10-22}$$

is an indicator of the relative contribution of the ith variable to the overall statistic. When an out-of-control signal is generated, we recommend computing the values of d_i ($i = 1, 2, \ldots, p$) and focusing attention on the variables for which d_i are relatively large. This procedure has an additional advantage in that the calculations can be performed using standard software packages.

To illustrate this procedure, consider the following example from Runger, Alt, and Montgomery (1996*a*). There are $p = 3$ quality characteristics and the covariance matrix is known. Assume that all three variables have been scaled as follows:

$$y_{ij} = \frac{x_{ij} - \mu_j}{\sqrt{(m-1)\sigma_j^2}}$$

This scaling results in each process variable having mean zero and variance one. Therefore, the covariance matrix Σ is in **correlation form;** that is, the main diagonal elements are all one and the off-diagonal elements are the pairwise correlation between the process variables (the x's). In our example,

$$\Sigma = \begin{bmatrix} 1 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 \\ 0.9 & 0.9 & 1 \end{bmatrix}$$

The in-control value of the process mean is $\mu' = [0, 0, 0]$. Consider the following display:

Observation Vector y'	Control Chart Statistic $T_0^2 (= \chi_0^2)$	$d_i = T^2 - T_{(i)}^2$ d_1	d_2	d_3
(2, 0, 0)	27.14	27.14	6.09	6.09
(1, 1, −1)	26.79	6.79	6.79	25.73
(1, −1, 0)	20.00	14.74	14.74	0
(0.5, 0.5, 1)	15.00	3.69	3.68	14.74

Since Σ is known, we can calculate the upper control limit for the chart from a chi-square distribution. We will choose $\chi^2_{0.005,3} = 12.84$ as the upper control limit. Clearly all four observation vectors in the above display would generate an out-of-control signal. Runger, Alt, and Montgomery (1996b) suggest that an approximate cutoff for the magnitude of an individual d_i is $\chi^2_{\alpha,1}$. Selecting $\alpha = 0.01$, we would find $\chi^2_{0.01,1} = 6.63$, so any d_i exceeding this value would be considered a large contributor. The decomposition statistics d_i computed above give clear guidance regarding *which* variables in the observation vector have shifted.

Other diagnostics have been suggested in the literature. For example, Murphy (1987) and Chua and Montgomery (1992) have developed procedures based on discriminant analysis, a statistical procedure for classifying observations into groups. Tracy, Mason, and Young (1996) also use decompositions of T^2 for diagnostic purposes, but their procedure requires more extensive computations and uses more elaborate decompositions than equation 10-22.

10-3.2 Individual Observations

In some industrial settings the subgroup size is naturally $n = 1$. This situation occurs frequently in the chemical and process industries. Since these industries frequently have multiple quality characteristics that must be monitored, multivariate control charts with $n = 1$ would be of interest there.

IMPORTANT TERMS AND CONCEPTS

- Average run length
- Cascade process
- Chi-square control chart
- Control ellipse
- Covariance matrix
- Hotelling T^2 control chart
- Matrix of scatter plots
- Mean vector
- Monitoring multivariate variability
- Multivariate EWMA control chart
- Multivariate normal distribution
- Multivariate process control
- Partial least squares
- Phase I control limits
- Phase II control limits
- Principal component scores
- Principal components analysis (PCA)
- Profile monitoring
- Profiles
- Regression adjustment
- Residual control chart
- Sample covariance matrix
- Sample mean vector
- Subgroup data versus individual observations
- Trajectory plots

EXERCISES

The Student Resource Manual presents comprehensive annotated solutions to the odd-numbered exercises included in the Answers to Selected Exercises section in the back of this book.

10-1. The data shown here come from a production process with two observable quality characteristics, x_1 and x_2. The data are sample means of each quality characteristic, based on samples of size $n = 25$. Assume that mean values of the quality characteristics and the covariance matrix were computed from 50 preliminary samples:

$$\bar{\bar{\mathbf{x}}} = \begin{bmatrix} 55 \\ 30 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 200 & 130 \\ 130 & 120 \end{bmatrix}$$

Construct a T^2 control chart using these data. Use the phase II limits.

Sample Number	$\bar{x}_1$	$\bar{x}_2$
1	58	32
2	60	33
3	50	27
4	54	31
5	63	38
6	53	30
7	42	20
8	55	31
9	46	25
10	50	29
11	49	27
12	57	30
13	58	33
14	75	45
15	55	27

10-2. A product has three quality characteristics. The nominal values of these quality characteristics and their sample covariance matrix have been determined from the analysis of 30 preliminary samples of size $n = 10$ as follows:

$$\bar{\bar{\mathbf{x}}} = \begin{bmatrix} 3.0 \\ 3.5 \\ 2.8 \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} 1.40 & 1.02 & 1.05 \\ 1.02 & 1.35 & 0.98 \\ 1.05 & 0.98 & 1.20 \end{bmatrix}$$

The sample means for each quality characteristic for 15 additional samples of size $n = 10$ are shown next. Is the process in statistical control?

Sample Number	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$
1	3.1	3.7	3.0
2	3.3	3.9	3.1
3	2.6	3.0	2.4
4	2.8	3.0	2.5
5	3.0	3.3	2.8
6	4.0	4.6	3.5
7	3.8	4.2	3.0
8	3.0	3.3	2.7
9	2.4	3.0	2.2
10	2.0	2.6	1.8
11	3.2	3.9	3.0
12	3.7	4.0	3.0
13	4.1	4.7	3.2
14	3.8	4.0	2.9
15	3.2	3.6	2.8

10-3. Reconsider the situation in Exercise 10-1. Suppose that the sample mean vector and sample covariance matrix provided were the actual population parameters. What control limit would be appropriate for phase II for the control chart? Apply this limit to the data and discuss any differences in results that you find in comparison to the original choice of control limit.

10-4. Reconsider the situation in Exercise 10-2. Suppose that the sample mean vector and sample covariance matrix provided were the actual population parameters. What control limit would be appropriate for phase II of the control chart? Apply this limit to the data and discuss any differences in results that you find in comparison to the original choice of control limit.

10-5. Consider a T^2 control chart for monitoring $p = 6$ quality characteristics. Suppose that the subgroup size is $n = 3$ and there are 30 preliminary samples available to estimate the sample covariance matrix.

(a) Find the phase II control limits assuming that $\alpha = 0.005$.

(b) Compare the control limits from part (a) to the chi-square control limit. What is the magnitude of the difference in the two control limits?

(c) How many preliminary samples would have to be taken to ensure that the exact phase II control limit is within 1% of the chi-square control limit?

10-6. Rework Exercise 10-5, assuming that the subgroup size is $n = 5$.

10-7. Consider a T^2 control chart for monitoring $p = 10$ quality characteristics. Suppose that the subgroup size is $n = 3$ and there are 25 preliminary samples available to estimate the sample covariance matrix.

(a) Find the phase II control limits assuming that $\alpha = 0.005$.

(b) Compare the control limits from part (a) to the chi-square control limit. What is the magnitude of the difference in the two control limits?

(c) How many preliminary samples would have to be taken to ensure that the chi-square control limit is within 1% of the exact phase II control limit?

10-8. Rework Exercise 10-7, assuming that the subgroup size is $n = 5$.

10-9. Consider a T^2 control chart for monitoring $p = 10$ quality characteristics. Suppose that the subgroup size is $n = 3$ and there are 25 preliminary samples available to estimate the sample covariance matrix. Calculate both the phase I and the phase II control limits (use $\alpha = 0.01$).

10-10. Suppose that we have $p = 4$ quality characteristics, and in correlation form all four variables have variance unity and all pairwise correlation coefficients are 0.7. The in-control value of the process mean vector is $\mu' = [0, 0, 0, 0]$.

(a) Write out the covariance matrix Σ.

(b) What is the chi-square control limit for the chart, assuming that $\alpha = 0.01$?

(c) Suppose that a sample of observations results in the standardized observation vector $\mathbf{y}' = [3.5, 3.5, 3.5, 3.5]$. Calculate the value of the T^2 statistic. Is an out-of-control signal generated?

(d) Calculate the diagnostic quantities d_i, $i = 1, 2, 3, 4$ from equation 10-22. Does this information assist in identifying which process variables have shifted?

(e) Suppose that a sample of observations results in the standardized observation vector $\mathbf{y}' = [2.5, 2, -1, 0]$. Calculate the value of the T^2 statistic. Is an out-of-control signal generated?

(f) For the case in (e), calculate the diagnostic quantities d_i, $i = 1, 2, 3, 4$ from equation 10-22. Does this information assist in identifying which process variables have shifted?

10-11. Suppose that we have $p = 3$ quality characteristics, and in correlation form all three variables have variance unity and all pairwise correlation coefficients are 0.8. The in-control value of the process mean vector is $\mu' = [0, 0, 0]$.

(a) Write out the covariance matrix Σ.

(b) What is the chi-square control limit for the chart, assuming that $\alpha = 0.05$?

(c) Suppose that a sample of observations results in the standardized observation vector $\mathbf{y}' = [1, 2, 0]$. Calculate the value of the T^2 statistic. Is an out-of-control signal generated?

(d) Calculate the diagnostic quantities d_i, $i = 1, 2, 3$ from equation 10-22. Does this information assist in identifying which process variables have shifted?

(e) Suppose that a sample of observations results in the standardized observation vector $\mathbf{y}' = [2, 2, 1]$. Calculate the value of the T^2 statistic. Is an out-of-control signal generated?

(f) For the case in (e), calculate the diagnostic quantities d_i, $i = 1, 2, 3$ from equation 10-22. Does this information assist in identifying which process variables have shifted?

10-12. Consider the first two process variables in Table 10-5. Calculate an estimate of the sample covariance matrix using both estimators $\mathbf{S}_1$ and $\mathbf{S}_2$ discussed in Section 10-3.2.

10-13. Consider the first three process variables in Table 10-5. Calculate an estimate of the sample covariance matrix using both estimators $\mathbf{S}_1$ and $\mathbf{S}_2$ discussed in Section 10-3.2.

10-14. Consider all 30 observations on the first two process variables in Table 10-6. Calculate an estimate of the sample covariance matrix using both estimators $\mathbf{S}_1$ and $\mathbf{S}_2$ discussed in Section 10-3.2. Are the estimates very different? Discuss your findings.

10-15. Suppose that there are $p = 4$ quality characteristics, and in correlation form all four variables have variance unity and all pairwise correlation coefficients are 0.75. The in-control value of the process mean vector is $\boldsymbol{\mu}' = [0, 0, 0, 0]$, and we want to design an MEWMA control chart to provide good protection against a shift to a new mean vector of $\mathbf{y}' = [1, 1, 1, 1]$. If an in-control ARL_0 of 200 is satisfactory, what value of λ and what upper control limit should be used? Approximately, what is the ARL_1 for detecting the shift in the mean vector?

10-16. Suppose that there are $p = 4$ quality characteristics, and in correlation form all four variables have variance unity and that all pairwise correlation coefficients are 0.9. The in-control value of the process mean vector is $\boldsymbol{\mu}' = [0, 0, 0, 0]$, and we want to design an MEWMA control chart to provide good protection against a shift to a new mean vector of $\mathbf{y}' = [1, 1, 1, 1]$. Suppose that an in-control ARL_0 of 500 is desired. What value of λ and what upper control limit would you recommend? Approximately, what is the ARL_1 for detecting the shift in the mean vector?

10-17. Suppose that there are $p = 2$ quality characteristics, and in correlation form both variables have variance unity and the correlation coefficient is 0.8. The in-control value of the process mean vector is $\boldsymbol{\mu}' = [0, 0]$, and we want to design an MEWMA control chart to provide good protection against a shift to a new mean vector of $\mathbf{y}' = [1, 1]$. If an in-control ARL_0 of 200 is satisfactory, what value of λ and what upper control limit should be used? Approximately, what is the ARL_1 for detecting the shift in the mean vector?

10-18. Consider the cascade process data in Table 10-5.

(a) Set up an individuals control chart on y_2.

(b) Fit a regression model to y_2, and set up an individuals control chart on the residuals. Comment on the differences between this chart and the one in part (a).

(c) Calculate the sample autocorrelation functions on y_2 and on the residuals from the regression model in part (b). Discuss your findings.

10-19. Consider the cascade process data in Table 10-5. In fitting regression models to both y_1 and y_2 you will find that not all of the process variables are required to obtain a satisfactory regression model for the output variables. Remove the nonsignificant variables from these equations and obtain subset regression models for both y_1 and y_2. Then construct individuals control charts for both sets of residuals. Compare them to the residual control charts in the text (Fig. 10-11) and from Exercise 10-18. Are there any substantial differences between the charts from the two different approaches to fitting the regression models?

10-20. **Continuation of Exercise 10-19.** Using the residuals from the regression models in Exercise 10-19, set up EWMA control charts. Compare these EWMA control charts to the Shewhart charts for individuals constructed previously. What are the potential advantages of the EWMA control chart in this situation?

10-21. Consider the $p = 4$ process variables in Table 10-6. After applying the PCA procedure to the first 20 observations data (see Table 10-7), suppose that the first three principal components are retained.

(a) Obtain the principal component scores. (Hint: Remember that you must work in standardized variables.)

(b) Construct an appropriate set of pairwise plots of the principal component scores.

(c) Calculate the principal component scores for the last 10 observations. Plot the scores on the charts from part (b) and interpret the results.

10-22. Consider the $p = 9$ process variables in Table 10-5.

(a) Perform a PCA on the first 30 observations. Be sure to work with the standardized variables.

(b) How much variability is explained if only the first $r = 3$ principal components are retained?

(c) Construct an appropriate set of pairwise plots of the first $r = 3$ principal component scores.

(d) Now consider the last 10 observations. Obtain the principal component scores and plot them on the chart in part (c). Does the process seem to be in control?