6 Variables Control Charts

Chapter Outline

6.2 Control Charts for $\bar{x}$ and R

6.2.1 Statistical Basis of the Charts

Suppose that a quality characteristic is normally distributed with mean μ and standard deviation σ, where both μ and σ are known. If $x_1, x_2, \ldots, x_n$ is a sample of size n, then the average of this sample is

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

and we know that $\bar{x}$ is normally distributed with mean μ and standard deviation $\sigma_{\bar{x}} = \sigma/\sqrt{n}$. Furthermore, the probability is $1 - \alpha$ that any sample mean will fall between

$$\mu + Z_{\alpha/2}\sigma_{\bar{x}} = \mu + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \quad \text{and} \quad \mu - Z_{\alpha/2}\sigma_{\bar{x}} = \mu - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \tag{6.1}$$

Therefore, if μ and σ are known, equation 6.1 could be used as upper and lower control limits on a control chart for sample means. As noted in Chapter 5, it is customary to replace $Z_{\alpha/2}$ by 3, so that three-sigma limits are employed. If a sample mean falls outside of these limits, it is an indication that the process mean is no longer equal to μ.

Subgroup Data with Unknown μ and σ

In practice, we usually will not know μ and σ. Therefore, they must be estimated from preliminary samples or subgroups taken when the process is thought to be in control. These estimates should usually be based on at least 20 to 25 samples. Suppose that m samples are available, each containing n observations on the quality characteristic. Typically, n will be small, often either 4, 5, or 6. These small sample sizes usually result from the construction of rational subgroups and from the fact that the sampling and inspection costs associated with variables measurements are usually relatively large. Let $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m$ be the average of each sample. Then the best estimator of μ, the process average, is the grand average—say,

$$\bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \cdots + \bar{x}_m}{m} \qquad (6.2)$$

Thus, $\bar{\bar{x}}$ would be used as the center line on the $\bar{x}$ chart.

To construct the control limits, we need an estimate of the standard deviation σ. Recall from Chapter 4 (Section 4.2) that we may estimate σ from either the standard deviations or the ranges of the m samples. For the present, we will use the range method. If $x_1, x_2, \ldots, x_n$ is a sample of size n, then the range of the sample is the difference between the largest and smallest observations—that is,

$$R = x_{\max} - x_{\min}$$

Let $R_1, R_2, \ldots, R_m$ be the ranges of the m samples. The average range is

$$\overline{R} = \frac{R_1 + R_2 + \cdots + R_m}{m} \tag{6.3}$$

We may now give the formulas for constructing the control limits on the $\bar{x}$ chart. They are as follows:

Control Limits for the $\bar{x}$ Chart

$$\begin{aligned} \text{UCL} &= \bar{\bar{x}} + A_2\bar{R} \\ \text{Center line} &= \bar{\bar{x}} \\ \text{LCL} &= \bar{\bar{x}} - A_2\bar{R} \end{aligned} \qquad (6.4)$$

The constant A_2 is tabulated for various sample sizes in Appendix Table VI.

Process variability may be monitored by plotting values of the sample range R on a control chart. The center line and control limits of the R chart are as follows:

Control Limits for the R Chart

$$\begin{aligned} \text{UCL} &= D_4\bar{R} \\ \text{Center line} &= \bar{R} \\ \text{LCL} &= D_3\bar{R} \end{aligned} \qquad (6.5)$$

The constants D_3 and D_4 are tabulated for various values of n in Appendix Table VI.

$$\hat{\sigma} = \frac{\overline{R}}{d_2} \tag{6.6}$$

If we use $\overline{\overline{x}}$ as an estimator of μ and $\overline{R}/d_2$ as an estimator of σ, then the parameters of the $\overline{x}$ chart are

$$\begin{aligned} \text{UCL} &= \overline{\overline{x}} + \frac{3}{d_2\sqrt{n}}\overline{R} \\ \text{Center line} &= \overline{\overline{x}} \\ \text{LCL} &= \overline{\overline{x}} - \frac{3}{d_2\sqrt{n}}\overline{R} \end{aligned} \tag{6.7}$$

If we define

$$A_2 = \frac{3}{d_2\sqrt{n}} \tag{6.8}$$

then equation 6.7 reduces to equation 6.4.

Now consider the R chart. The center line will be $\overline{R}$. To determine the control limits, we need an estimate of σ_R. Assuming that the quality characteristic is normally distributed, $\hat{\sigma}_R$ can be found from the distribution of the relative range $W = R/\sigma$. The standard deviation of W, say d_3, is a known function of n. Thus, since

$$R = W\sigma$$

the standard deviation of R is

$$\sigma_R = d_3\sigma$$

Since σ is unknown, we may estimate σ_R by

$$\hat{\sigma}_R = d_3 \frac{\overline{R}}{d_2} \tag{6.9}$$

Consequently, the parameters of the R chart with the usual three-sigma control limits are

$$\begin{aligned} \text{UCL} &= \overline{R} + 3\hat{\sigma}_R = \overline{R} + 3d_3 \frac{\overline{R}}{d_2} \\ \text{Center line} &= \overline{R} \\ \text{LCL} &= \overline{R} - 3\hat{\sigma}_R = \overline{R} - 3d_3 \frac{\overline{R}}{d_2} \end{aligned} \tag{6.10}$$

If we let

$$D_3 = 1 - 3\frac{d_3}{d_2} \quad \text{and} \quad D_4 = 1 + 3\frac{d_3}{d_2}$$

equation (6.10) reduces to equation (6.5).

Phase I Application of $\overline{x}$ and R Charts

- Eqns 6.4 and 6.5 are **trial control limits**
  - Determined from *m* initial samples
    - Typically 20-25 subgroups of size *n* between 3 and 5
  - Any out-of-control points should be examined for assignable causes
    - If assignable causes are found, discard points from calculations and revise the trial control limits
    - Continue examination until all points plot in control
    - Adopt resulting trial control limits for use
    - If no assignable cause is found, there are two options
      1. Eliminate point as if an assignable cause were found and revise limits
      2. Retain point and consider limits appropriate for control
  - If there are many out-of-control points they should be examined for **patterns** that may identify underlying process problems

Example 6.1 The Hard Bake Process

TABLE 6.1

Flow Width Measurements (microns) for the Hard-Bake Process

Sample Number	Wafers 1	2	3	4	5	$\bar{x}_i$	R_i
1	1.3235	1.4128	1.6744	1.4573	1.6914	1.5119	0.3679
2	1.4314	1.3592	1.6075	1.4666	1.6109	1.4951	0.2517
3	1.4284	1.4871	1.4932	1.4324	1.5674	1.4817	0.1390
4	1.5028	1.6352	1.3841	1.2831	1.5507	1.4712	0.3521
5	1.5604	1.2735	1.5265	1.4363	1.6441	1.4882	0.3706
6	1.5955	1.5451	1.3574	1.3281	1.4198	1.4492	0.2674
7	1.6274	1.5064	1.8366	1.4177	1.5144	1.5805	0.4189
8	1.4190	1.4303	1.6637	1.6067	1.5519	1.5343	0.2447
9	1.3884	1.7277	1.5355	1.5176	1.3688	1.5076	0.3589
10	1.4039	1.6697	1.5089	1.4627	1.5220	1.5134	0.2658
11	1.4158	1.7667	1.4278	1.5928	1.4181	1.5242	0.3509
12	1.5821	1.3355	1.5777	1.3908	1.7559	1.5284	0.4204
13	1.2856	1.4106	1.4447	1.6398	1.1928	1.3947	0.4470
14	1.4951	1.4036	1.5893	1.6458	1.4969	1.5261	0.2422
15	1.3589	1.2863	1.5996	1.2497	1.5471	1.4083	0.3499
16	1.5747	1.5301	1.5171	1.1839	1.8662	1.5344	0.6823
17	1.3680	1.7269	1.3957	1.5014	1.4449	1.4874	0.3589
18	1.4163	1.3864	1.3057	1.6210	1.5573	1.4573	0.3153
19	1.5796	1.4185	1.6541	1.5116	1.7247	1.5777	0.3062
20	1.7106	1.4412	1.2361	1.3820	1.7601	1.5060	0.5240
21	1.4371	1.5051	1.3485	1.5670	1.4880	1.4691	0.2185
22	1.4738	1.5936	1.6583	1.4973	1.4720	1.5390	0.1863
23	1.5917	1.4333	1.5551	1.5295	1.6866	1.5592	0.2533
24	1.6399	1.5243	1.5705	1.5563	1.5530	1.5688	0.1156
25	1.5797	1.3663	1.6240	1.3732	1.6887	1.5264	0.3224
						$\Sigma\bar{x}_i = 37.6400$	$\Sigma R_i = 8.1302$
						$\bar{\bar{x}} = 1.5056$	$\bar{R} = 0.32521$

Introduction to Statistical Quality Control, 7th Edition by Douglas C. Montgomery.

$$\overline{R} = \frac{\sum_{i=1}^{25} R_i}{25} = \frac{8.1302}{25} = 0.32521$$

$$\text{LCL} = \overline{R} D_3 = 0.32521(0) = 0$$
$$\text{UCL} = \overline{R} D_4 = 0.32521(2.114) = 0.68749$$

$$\overline{\overline{x}} = \frac{\sum_{i=1}^{25} \overline{x}_i}{25} = \frac{37.6400}{25} = 1.5056$$

$$\text{UCL} = \overline{\overline{x}} + A_2 \overline{R} = 1.5056 + (0.577)(0.32521) = 1.69325$$

and

$$\text{LCL} = \overline{\overline{x}} - A_2 \overline{R} = 1.5056 - (0.577)(0.32521) = 1.31795$$

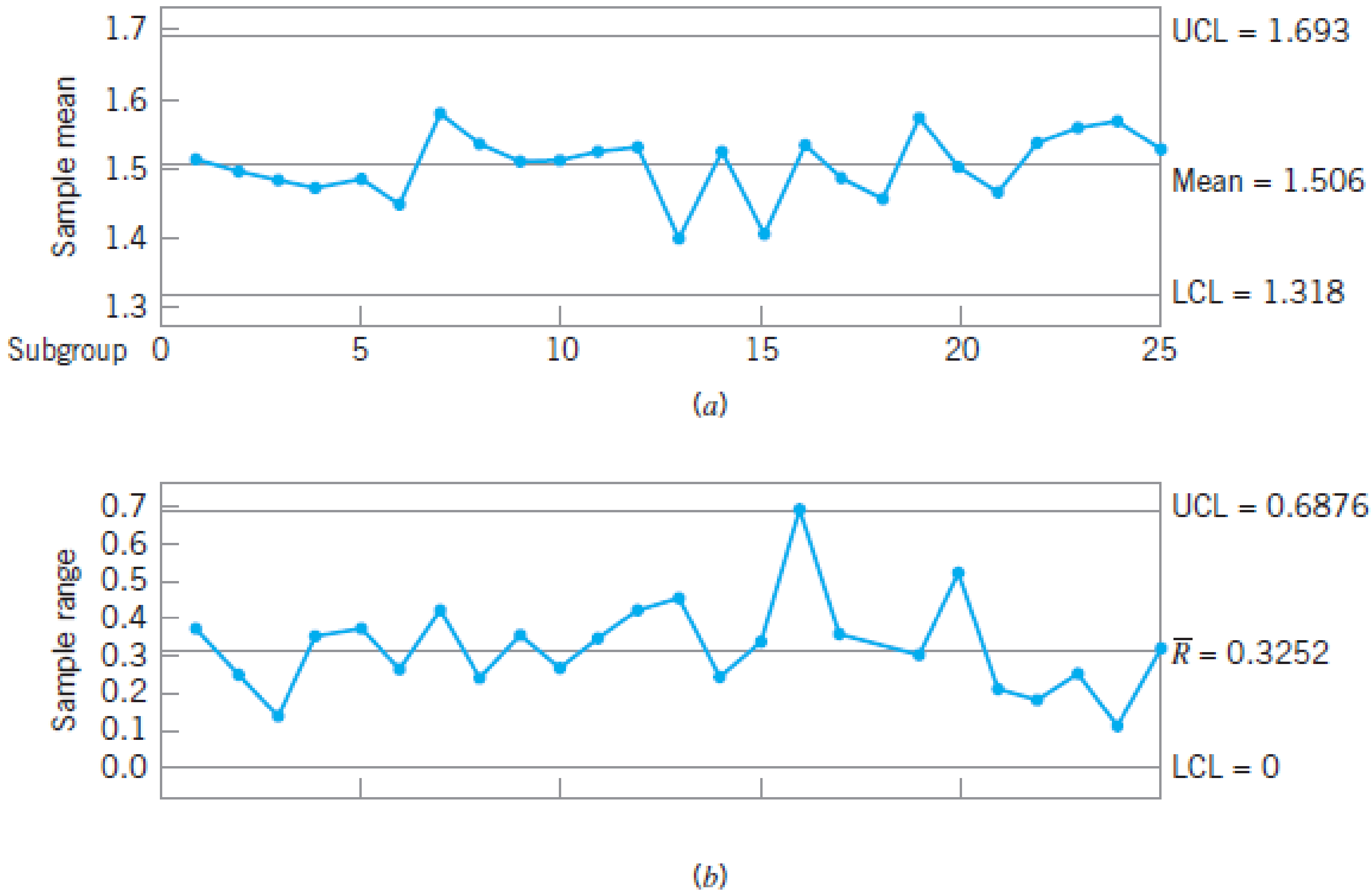


FIGURE 6.2 $\overline{x}$ and R charts (from Minitab) for flow width in the hard-bake process.

Estimating Process Capability

The $\bar{x}$ and R charts provide information about the performance or **capability** of the process. From the $\bar{x}$ chart, we may estimate the mean flow width of the resist in the hard-bake process as $\bar{\bar{x}} = 1.5056$ microns. The process standard deviation may be estimated using equation 5-6; that is,

$$\hat{\sigma} = \frac{\bar{R}}{d_2} = \frac{0.32521}{2.326} = 0.1398 \text{ microns}$$

where the value of d_2 for samples of size five is found in Appendix Table VI. The specification limits on flow width are 1.50 ± 0.50 microns. The control chart data may be used to describe the capability of the process to produce wafers relative to these specifications. Assuming that flow width is a normally distributed random variable, with mean 1.5056 and standard deviation 0.1398, we may estimate the fraction of nonconforming wafers produced as

$$\begin{aligned} p &= P\{x < 1.00\} + P\{x > 2.00\} \\ &= \Phi\left(\frac{1.00 - 1.5056}{0.1398}\right) + 1 - \Phi\left(\frac{2.00 - 1.5056}{0.1398}\right) \\ &= \Phi(-3.61660) + 1 - \Phi(3.53648) \\ &\simeq 0.00015 + 1 - 0.99980 \\ &\simeq 0.00035 \end{aligned}$$

That is, about 0.035 percent [350 parts per million (ppm)] of the wafers produced will be outside of the specifications.

Another way to express process capability is in terms of the **process capability ratio (PCR) C_p,** which for a quality characteristic with both upper and lower specification limits (USL and LSL, respectively) is

$$C_p = \frac{\text{USL} - \text{LSL}}{6\sigma} \tag{6.11}$$

Note that the 6σ spread of the process is the basic definition of process capability. Since σ is usually unknown, we must replace it with an estimate. We frequently use $\hat{\sigma} = \overline{R}/d_2$ as an estimate of σ, resulting in an estimate $\hat{C}_p$ of C_p. For the hard-bake process, since $\overline{R}/d_2 = \hat{\sigma} = 0.1398$, we find that

$$\hat{C}_p = \frac{2.00 - 1.00}{6(0.1398)} = \frac{1.00}{0.8388} = 1.192$$

This implies that the "natural" tolerance limits in the process (three-sigma above and below the mean) are inside the lower and upper specification limits. Consequently, a moderately small number of nonconforming wafers will be produced. The PCR C_p may be interpreted another way. The quantity

$$P = \left(\frac{1}{C_p}\right)100\%$$

is simply the percentage of the specification band that the process uses up. For the hard-bake process an estimate of P is

$$\hat{P} = \left(\frac{1}{\hat{C}_p}\right)100\% = \left(\frac{1}{1.192}\right)100\% = 83.89$$

That is, the process uses up about 84% of the specification band.

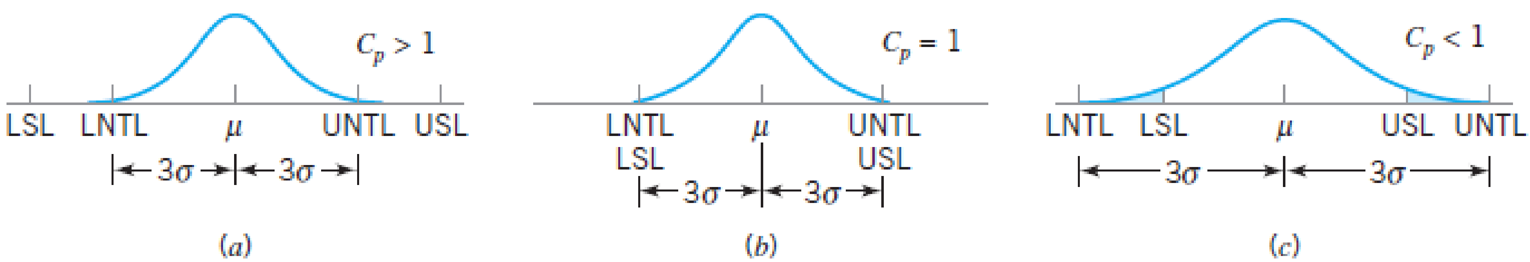


FIGURE 6.3 Process fallout and the process capability ratio C_p.

Phase II Operation of Charts

- Use of control chart for monitoring future production, once a set of reliable limits are established, is called *phase II* of control chart usage (Figure 6.4)
- A run chart showing individuals observations in each sample, called a **tolerance chart** or **tier diagram** (Figure 6.5), may reveal patterns or unusual observations in the data

TABLE 6.2
Additional Samples for Example 6.1

Sample Number	Wafers						
	1	2	3	4	5	$\bar{x}_i$	R_i
26	1.4483	1.5458	1.4538	1.4303	1.6206	1.4998	0.1903
27	1.5435	1.6899	1.5830	1.3358	1.4187	1.5142	0.3541
28	1.5175	1.3446	1.4723	1.6657	1.6661	1.5332	0.3215
29	1.5454	1.0931	1.4072	1.5039	1.5264	1.4152	0.4523
30	1.4418	1.5059	1.5124	1.4620	1.6263	1.5097	0.1845
31	1.4301	1.2725	1.5945	1.5397	1.5252	1.4724	0.3220
32	1.4981	1.4506	1.6174	1.5837	1.4962	1.5292	0.1668
33	1.3009	1.5060	1.6231	1.5831	1.6454	1.5317	0.3445
34	1.4132	1.4603	1.5808	1.7111	1.7313	1.5793	0.3181
35	1.3817	1.3135	1.4953	1.4894	1.4596	1.4279	0.1818
36	1.5765	1.7014	1.4026	1.2773	1.4541	1.4824	0.4241
37	1.4936	1.4373	1.5139	1.4808	1.5293	1.4910	0.0920
38	1.5729	1.6738	1.5048	1.5651	1.7473	1.6128	0.2425
39	1.8089	1.5513	1.8250	1.4389	1.6558	1.6560	0.3861
40	1.6236	1.5393	1.6738	1.8698	1.5036	1.6420	0.3662
41	1.4120	1.7931	1.7345	1.6391	1.7791	1.6716	0.3811
42	1.7372	1.5663	1.4910	1.7809	1.5504	1.6252	0.2899
43	1.5971	1.7394	1.6832	1.6677	1.7974	1.6970	0.2003
44	1.4295	1.6536	1.9134	1.7272	1.4370	1.6321	0.4839
45	1.6217	1.8220	1.7915	1.6744	1.9404	1.7700	0.3187

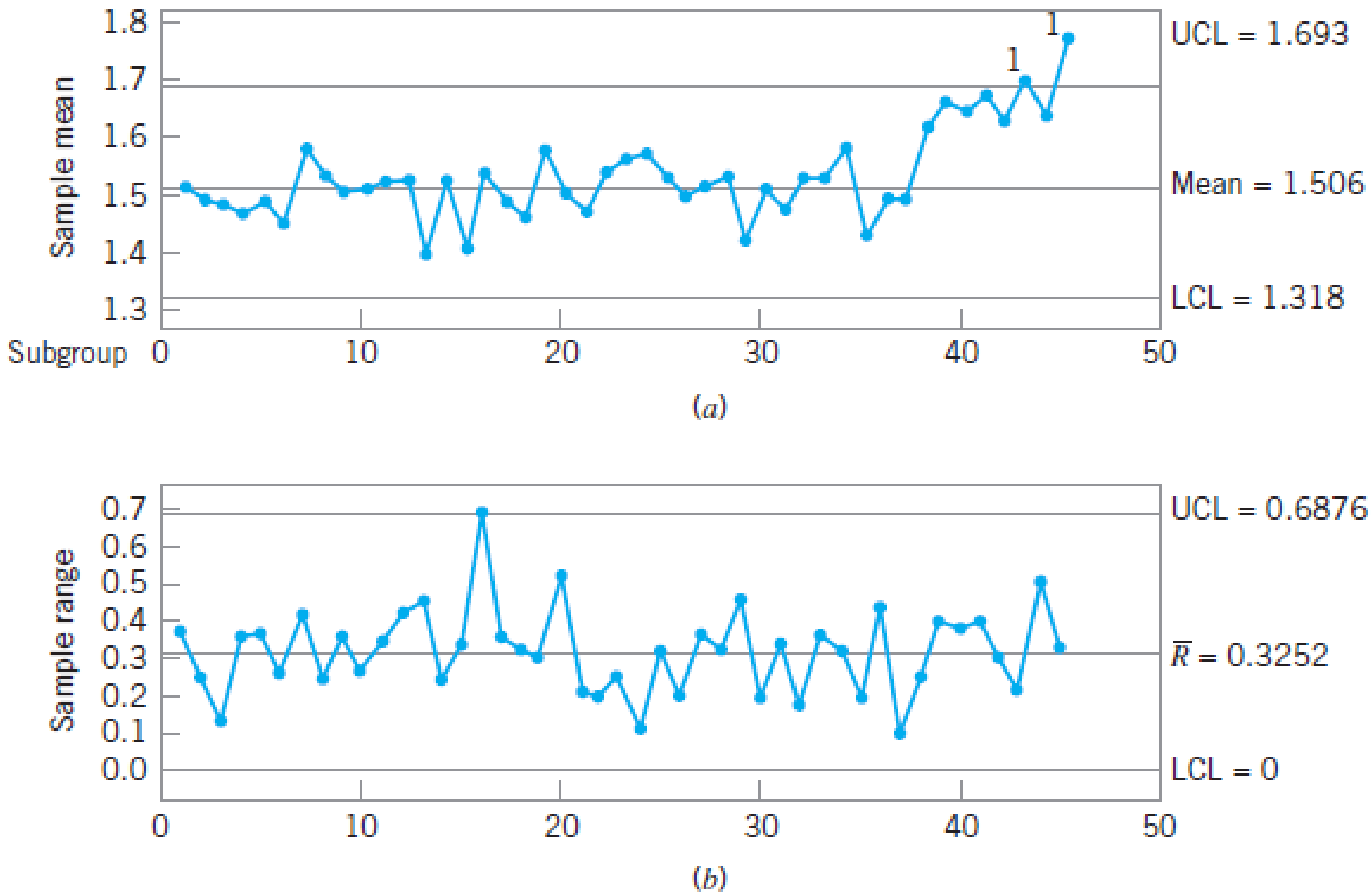
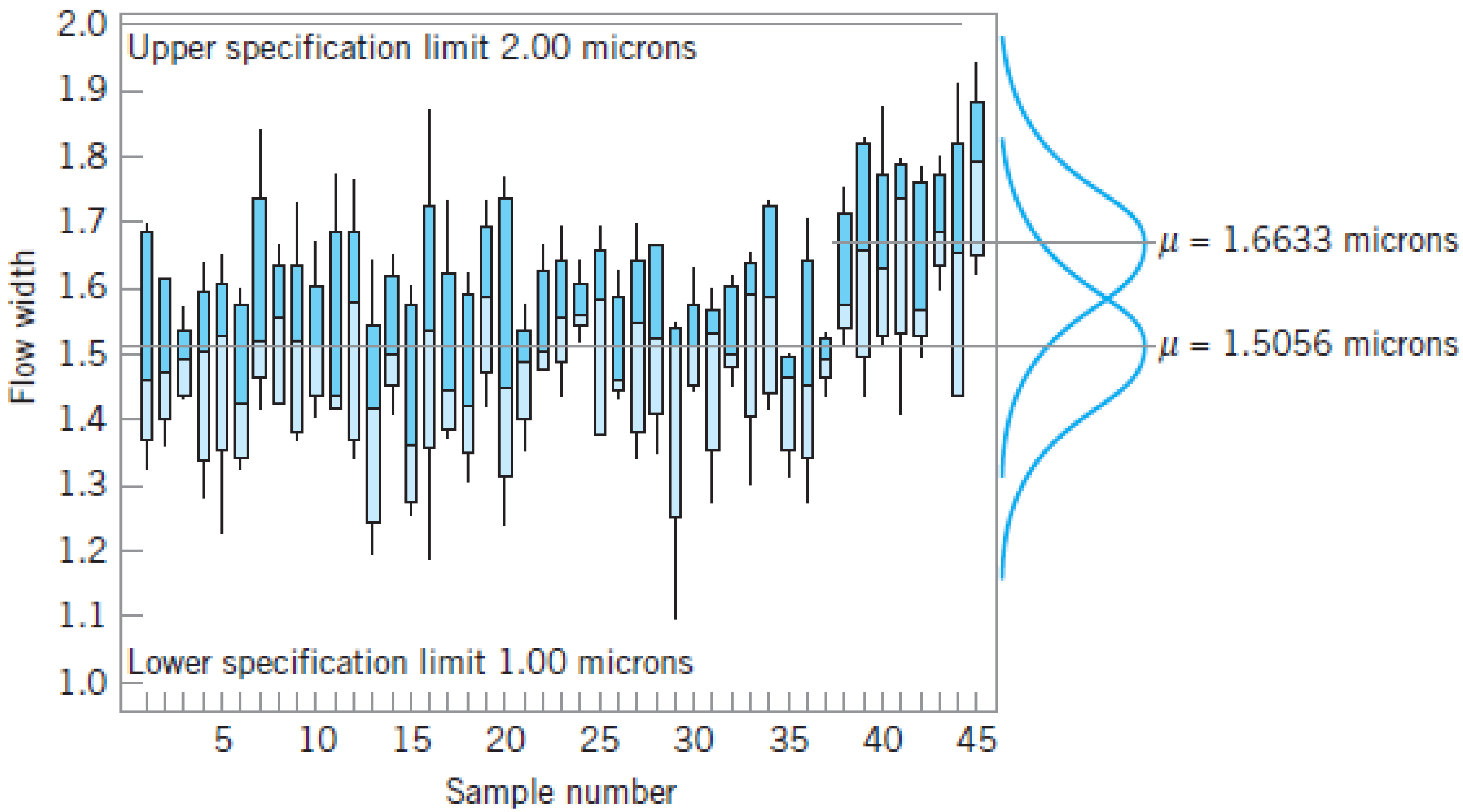


FIGURE 6.4 Continuation of the $\overline{x}$ and R charts in Example 6.1.



■ **FIGURE 6.5** Tier chart constructed using the Minitab box plot procedure for the flow width data.

Control vs. Specification Limits

- **Control** limits are derived from natural process variability, or the **natural tolerance** limits of a process

- **Specification** limits are determined externally, for example by customers or designers

- There is no mathematical or statistical relationship between the control limits and the specification limits

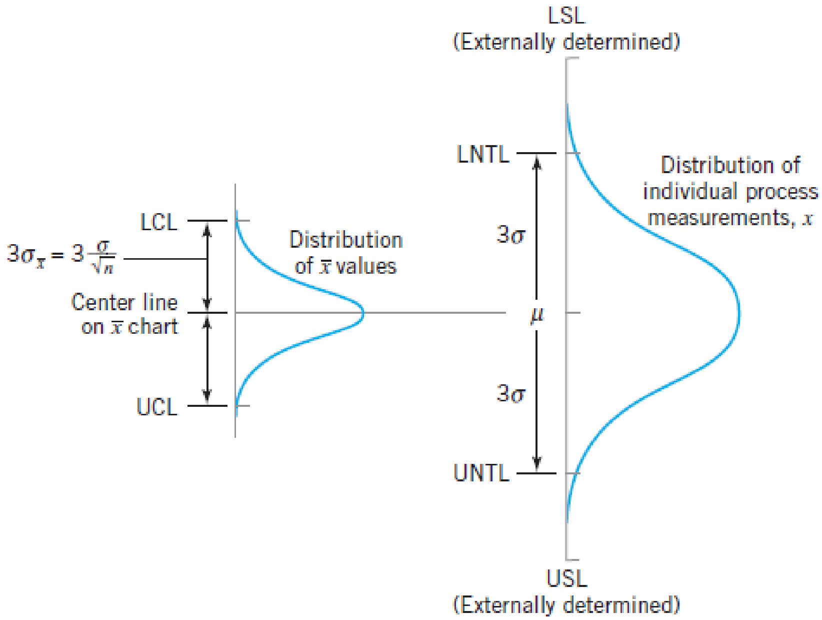


FIGURE 6.6 Relationship of natural tolerance limits, control limits, and specification limits.

Rational Subgroups

- $\overline{x}$ charts monitor **between-sample variability**
- R charts measure **within-sample variability**
- Standard deviation estimate of σ used to construct control limits is calculated from **within-sample variability**
- It is not correct to estimate σ using

$$s = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - \overline{\overline{x}}\right)^2}{mn-1}}$$

6.2.3 Charts Based on Standard Values

$$\begin{aligned}\text{UCL} &= \mu + 3\frac{\sigma}{\sqrt{n}}\\ \text{Center line} &= \mu\\ \text{LCL} &= \mu - 3\frac{\sigma}{\sqrt{n}}\end{aligned} \tag{6.14}$$

$$\begin{aligned}\text{UCL} &= \mu + A\sigma\\ \text{Center line} &= \mu\\ \text{LCL} &= \mu - A\sigma\end{aligned} \tag{6.15}$$

$$\begin{aligned}\text{UCL} &= d_2\sigma + 3d_3\sigma\\ \text{Center line} &= d_2\sigma\\ \text{LCL} &= d_2\sigma - 3d_3\sigma\end{aligned} \tag{6.16}$$

6.2.4 Interpretation of Control Charts

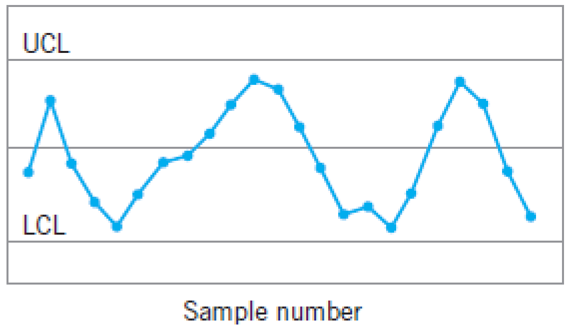


FIGURE 6.8 Cycles on a control chart.

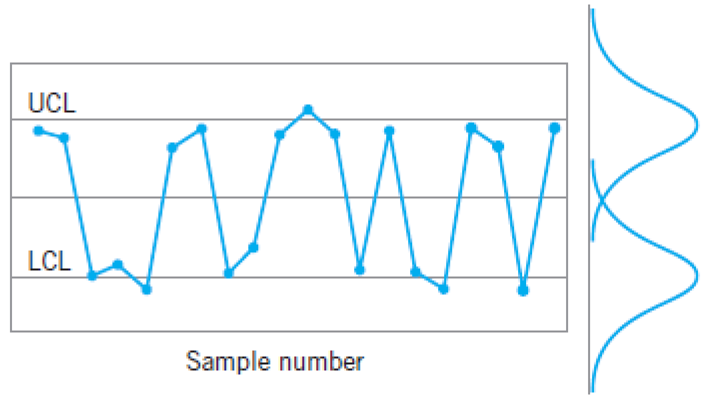


FIGURE 6.9 A mixture pattern.

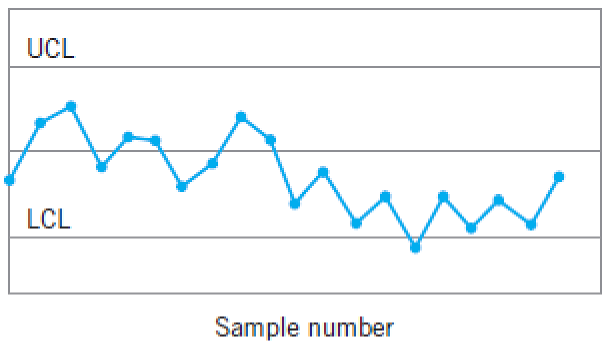


FIGURE 6.10 A shift in process level.

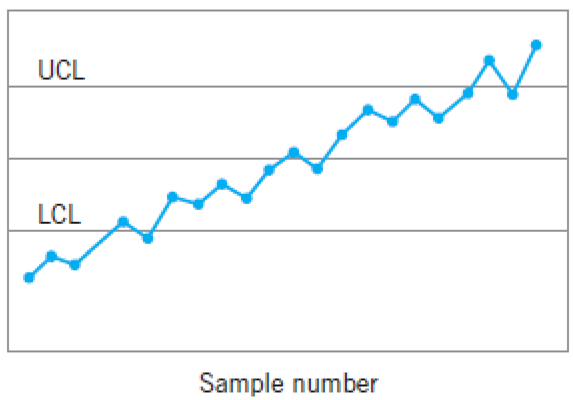
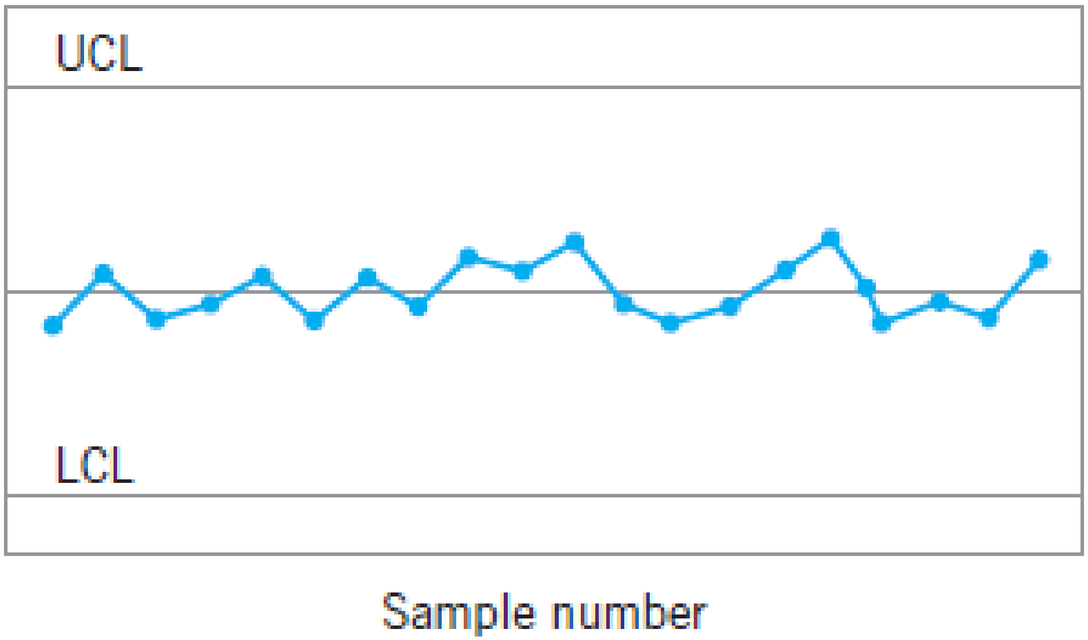


FIGURE 6.11 A trend in process level.



■ FIGURE 6.12 A stratification pattern.

6.2.5 The Effect of Non-normality

- An assumption in performance properties is that the underlying distribution of quality characteristic is **normal**
  - If underlying distribution is not normal, sampling distributions can be derived and exact probability limits obtained
- Burr (1967) notes the usual normal theory control limits are very robust to normality assumption
- Schilling and Nelson (1976) indicate that in most cases, samples of size 4 or 5 are sufficient to ensure reasonable robustness to normality assumption for $\bar{x}$ chart
- Sampling distribution of *R* is **not** symmetric, thus symmetric 3-sigma limits are an approximation and α-risk is not 0.0027. *R* chart is more sensitive to departures from normality than $\bar{x}$ chart.
- Assumptions of normality and independence are not a primary concern in phase I

6.2.6 The Operating Characteristic Function

Shift in the process mean to: $\mu_1 = \mu_0 + k\sigma$,

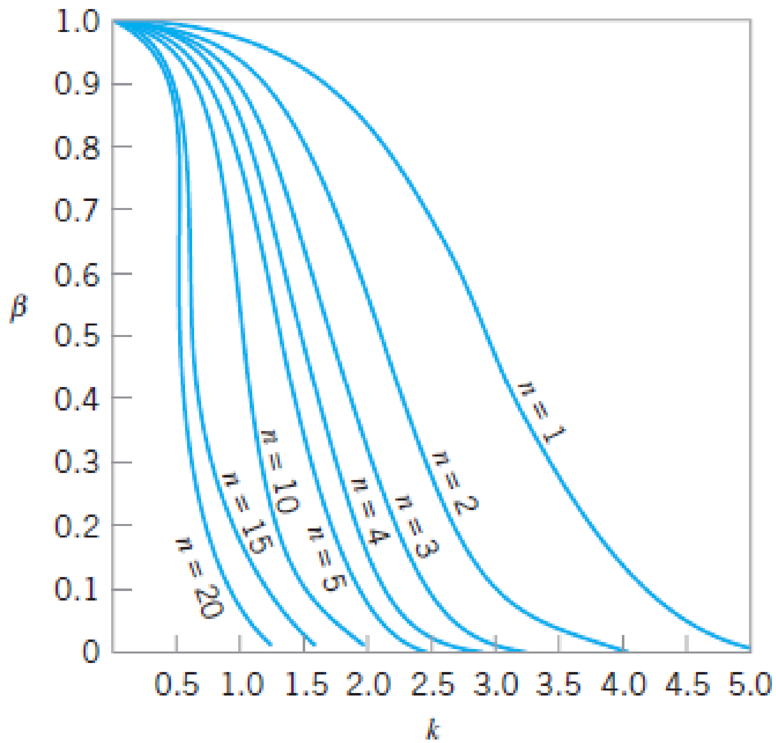
$$\beta = P\{\text{LCL} \le \bar{x} \le \text{UCL} | \mu = \mu_1 = \mu_0 + k\sigma\}$$

$$\beta = \Phi\left(L - k\sqrt{n}\right) - \Phi\left(-L - k\sqrt{n}\right) \qquad (6.19)$$

To illustrate the use of equation 6.19, suppose that we are using an $\bar{x}$ chart with $L = 3$ (the usual three-sigma limits), the sample size $n = 5$, and we wish to determine the probability of detecting a shift to $\mu_1 = \mu_0 + 2\sigma$ on the first sample following the shift. Then, since $L = 3$, $k = 2$, and $n = 5$, we have

$$\begin{aligned}\beta &= \Phi\left[3 - 2\sqrt{5}\right] - \Phi\left[-3 - 2\sqrt{5}\right] \\ &= \Phi(-1.47) - \Phi(-7.37) \\ &\cong 0.0708\end{aligned}$$

This is the β-risk, or the probability of not detecting such a shift. The probability that such a shift *will* be detected on the first subsequent sample is $1 - \beta = 1 - 0.0708 = 0.9292$.



If the shift is 1.0σ and the sample size is n = 5, then β = 0.75.

FIGURE 6.13 Operating-characteristic curves for the $\overline{x}$ chart with three-sigma limits. $\beta = P$ (not detecting a shift of $k\sigma$ in the mean on the first sample following the shift).

In general, the expected number of samples taken before the shift is detected is simply the **average run length,** or

$$\text{ARL} = \sum_{r=1}^{\infty} r\beta^{r-1}(1-\beta) = \frac{1}{1-\beta}$$

Therefore, in our example, we have

$$\text{ARL} = \frac{1}{1-\beta} = \frac{1}{0.25} = 4$$

In other words, the expected number of samples taken to detect a shift of 1.0σ with $n = 5$ is 4.

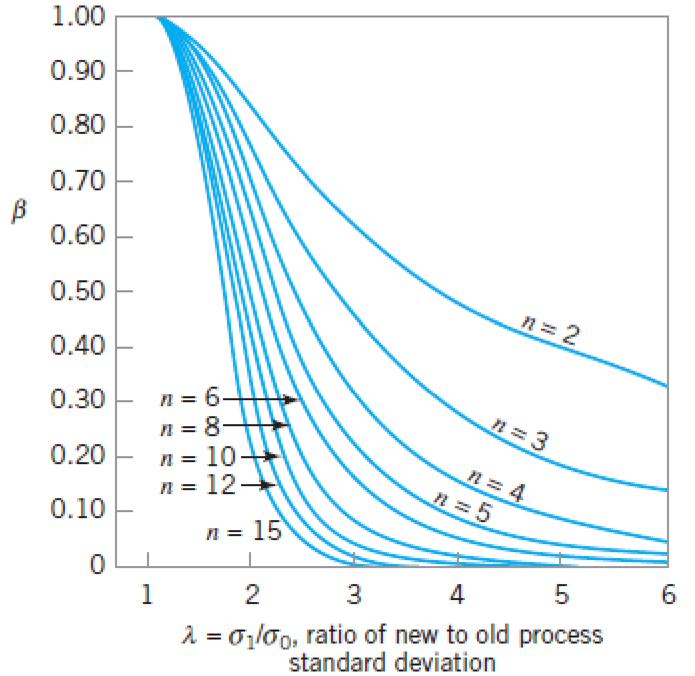


FIGURE 6.14 Operating-characteristic curves for the R chart with three-sigma limits. (Adapted from A. J. Duncan, "Operating Characteristics of R Charts," *Industrial Quality Control*, vol. 7, no. 5, pp. 40–41, 1951, with permission of the American Society for Quality Control.)

6.2.7 The Average Run Length for the $\bar{x}$ Chart

For any **Shewhart control chart,** we have noted previously that the ARL can be expressed as

$$\text{ARL} = \frac{1}{P(\text{one point plots out of control})}$$

or

$$\text{ARL}_0 = \frac{1}{\alpha} \tag{6.20}$$

for the in-control ARL and

$$\text{ARL}_1 = \frac{1}{1-\beta} \tag{6.21}$$

for the out-of-control ARL. These results are actually rather intuitive. If the observations plotted on the control chart are independent, then the number of points that must be plotted until the first point exceeds a control limit is a geometric random variable with parameter p (see Chapter 3). The mean of this geometric distribution is simply $1/p$, the average run length.

Two other performance measures based on ARL are sometimes of interest. The average time to signal is the number of time periods that occur until a signal is generated on the control chart. If samples are taken at equally spaced intervals of time h, then the **average time to signal** or the ATS is

$$\text{ATS} = \text{ARL}\, h \tag{6.22}$$

It may also be useful to express the ARL in terms of the expected number of individual *units* sampled—say, I—rather than the number of samples taken to detect a shift. If the sample size is n, the relationship between I and ARL is

$$I = n\, \text{ARL} \tag{6.23}$$

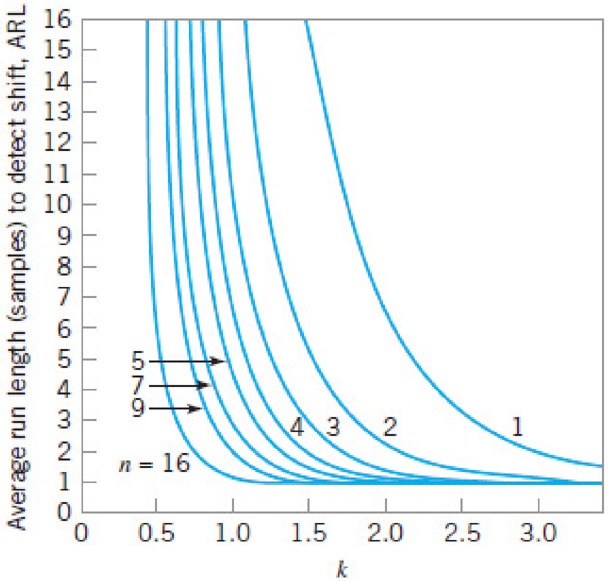


FIGURE 6.15 Average run length (samples) for the $\bar{x}$ chart with three-sigma limits, where the process mean shifts by $k\sigma$. (Adapted from *Modern Methods for Quality Control and Improvement*, by H. M. Wadsworth, K. S. Stephens, and A. B. Godfrey, 2nd edition, John Wiley & Sons, 2002.)

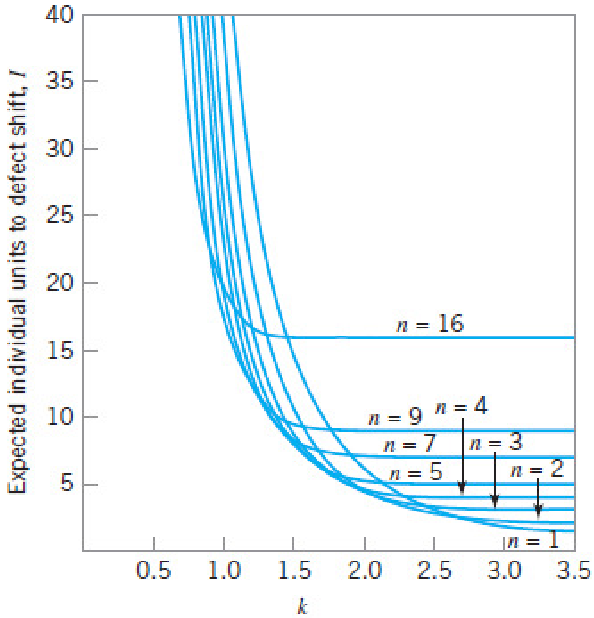


FIGURE 6.16 Average run length (individual units) for the $\bar{x}$ chart with three-sigma limits, where the process mean shifts by $k\sigma$. (Adapted from *Modern Methods for Quality Control and Improvement*, by H. M. Wadsworth, K. S. Stephens, and A. B. Godfrey, 2nd edition, John Wiley & Sons, 2002.)