Basic SPC Tools

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

Check Sheet

CHECK SHEET
DEFECT DATA FOR 2002–2003 YTD

Part No.: TAX-41
Location: Bellevue
Study Date: 6/5/03
Analyst: TCB

	2002												2003					
Defect	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	Total
Parts damaged		1		3	1	2		1		10	3		2	2	7	2		34
Machining problems			3	3				1	8		3		8	3				29
Supplied parts rusted			1	1		2	9											13
Masking insufficient		3	6	4	3	1												17
Misaligned weld	2																	2
Processing out of order	2															2		4
Wrong part issued		1						2										3
Unfinished fairing			3															3
Adhesive failure				1							1		2			1	1	6
Powdery alodine					1													1
Paint out of limits						1								1				2
Paint damaged by etching			1															1
Film on parts						3		1	1									5
Primer cans damaged								1										1
Voids in casting									1	1								2
Delaminated composite										2								2
Incorrect dimensions											13	7	13	1		1	1	36
Improper test procedure										1								1
Salt-spray failure													4			2		4
TOTAL	4	5	14	12	5	9	9	6	10	14	20	7	29	7	7	6	2	166

FIGURE 5.16 A check sheet to record defects on a tank used in an aerospace application.

How to Construct a Cause-and-Effect Diagram

1. Define the problem or effect to be analyzed.
2. Form the team to perform the analysis. Often the team will uncover potential causes through brainstorming.
3. Draw the effect box and the center line.
4. Specify the major potential cause categories and join them as boxes connected to the center line.
5. Identify the possible causes and classify them into the categories in step 4. Create new categories, if necessary.
6. Rank order the causes to identify those that seem most likely to impact the problem.
7. Take corrective action.

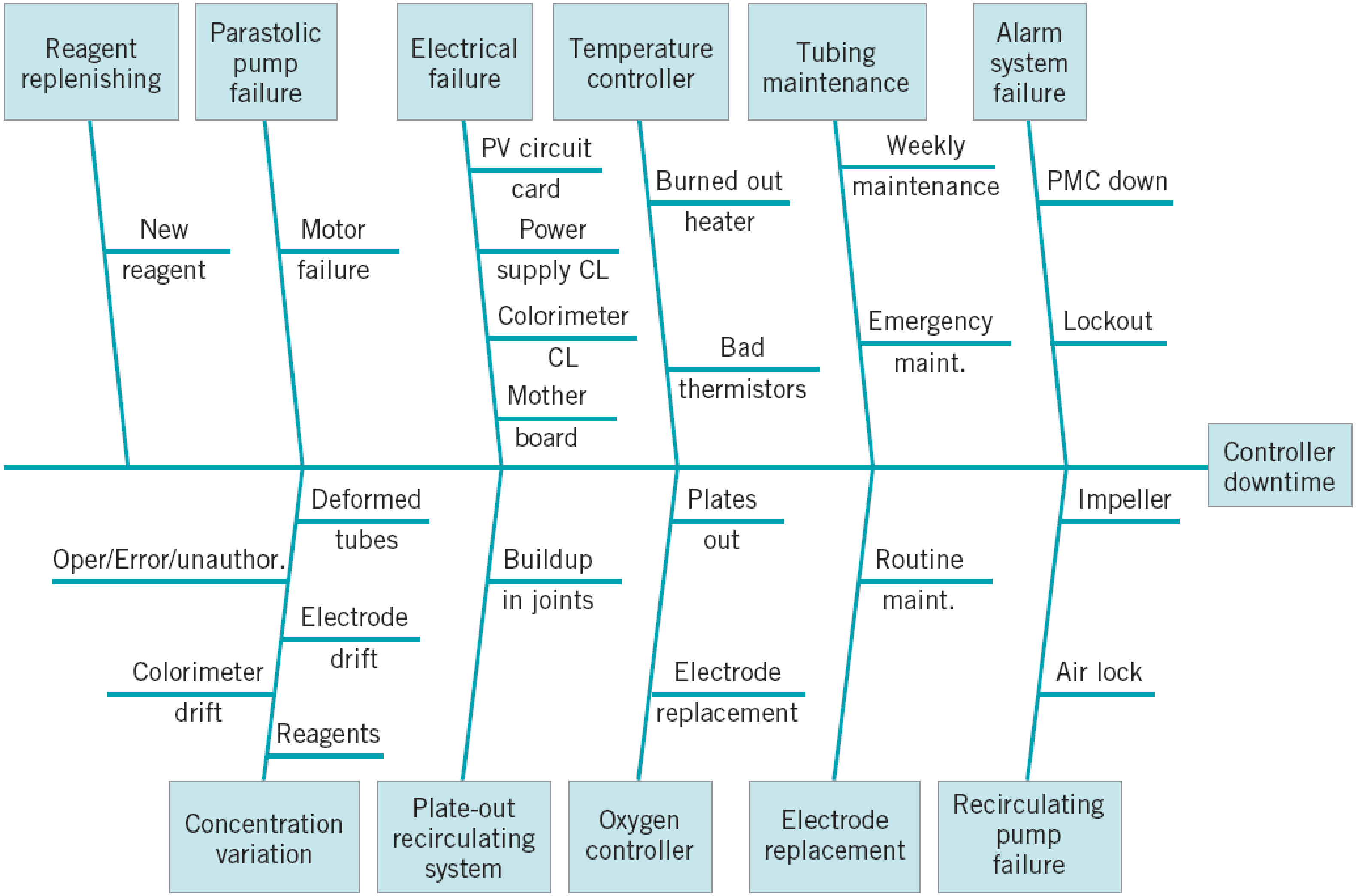


FIGURE 5.23 Cause-and-effect diagram for controller downtime.

Cause-and-Effect Diagram

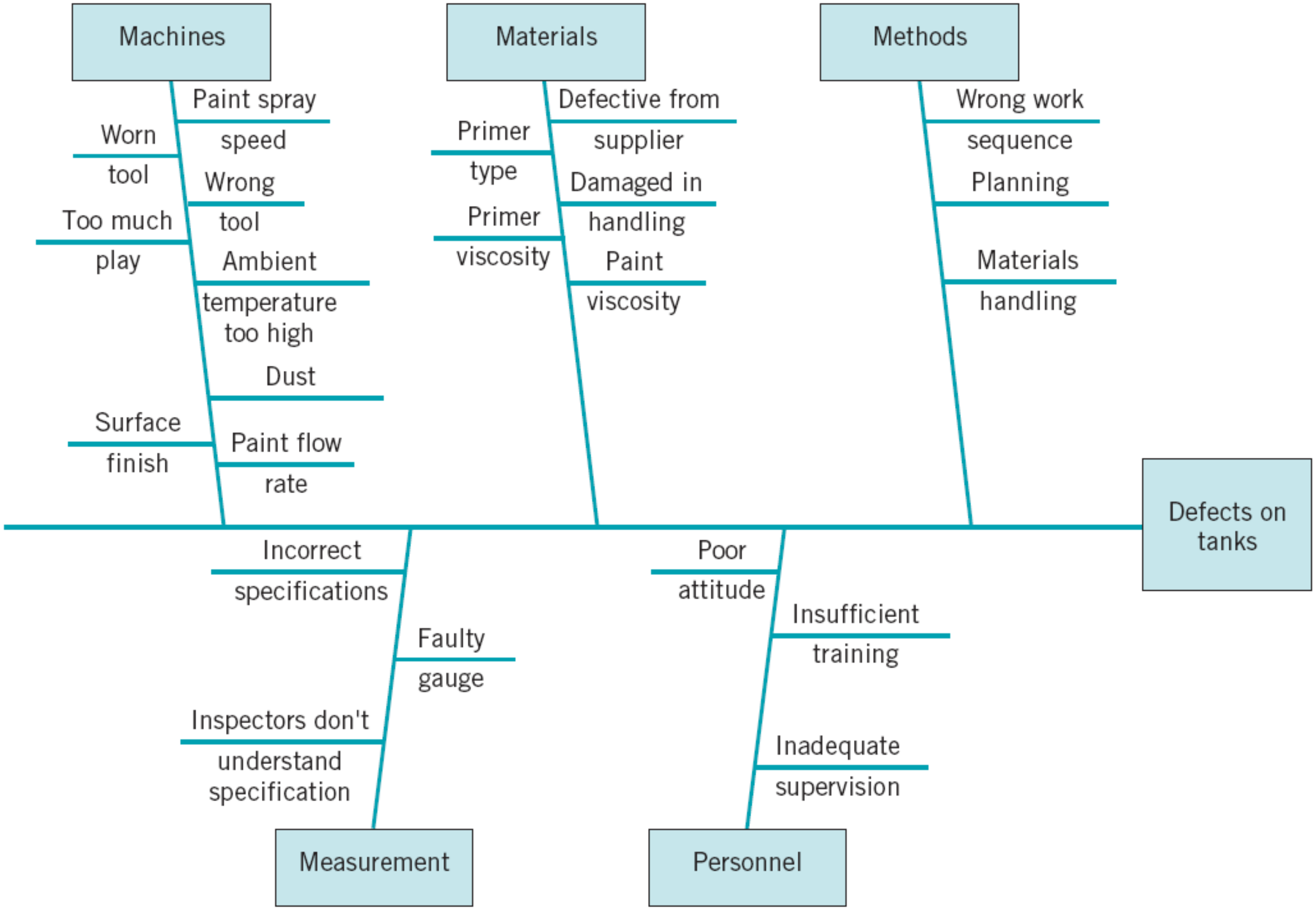


FIGURE 5.19 Cause-and-effect diagram for the tank defect problem.

Pareto Chart

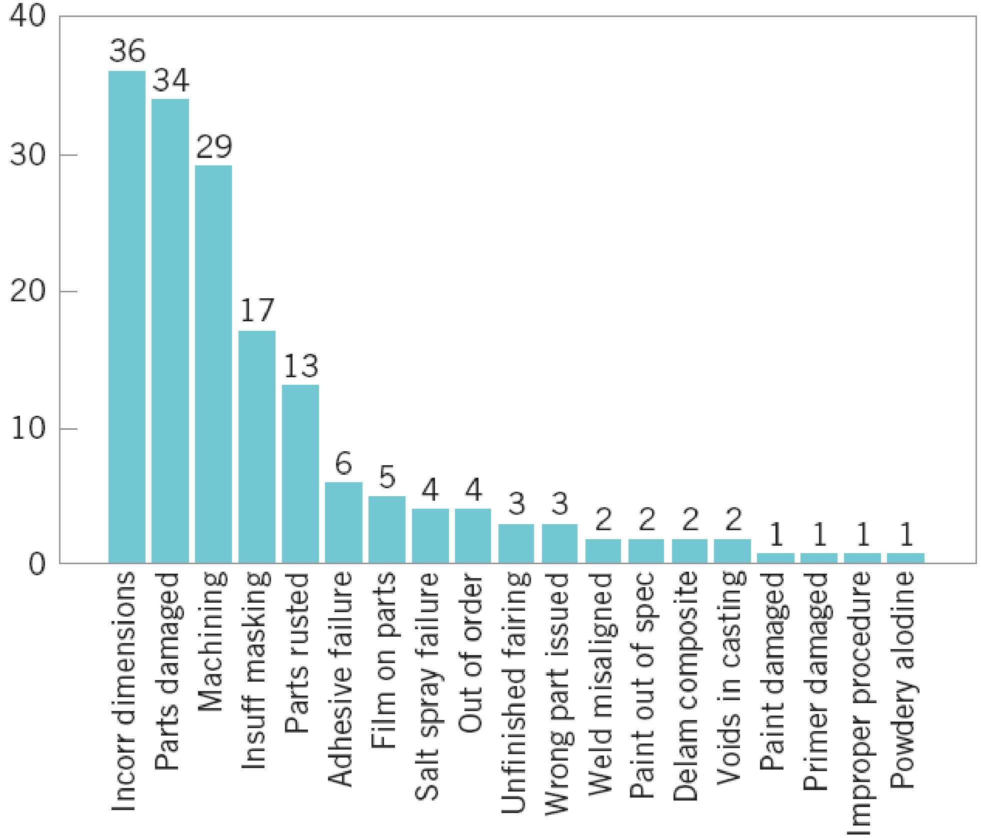


FIGURE 5.17 Pareto chart of the tank defect data.

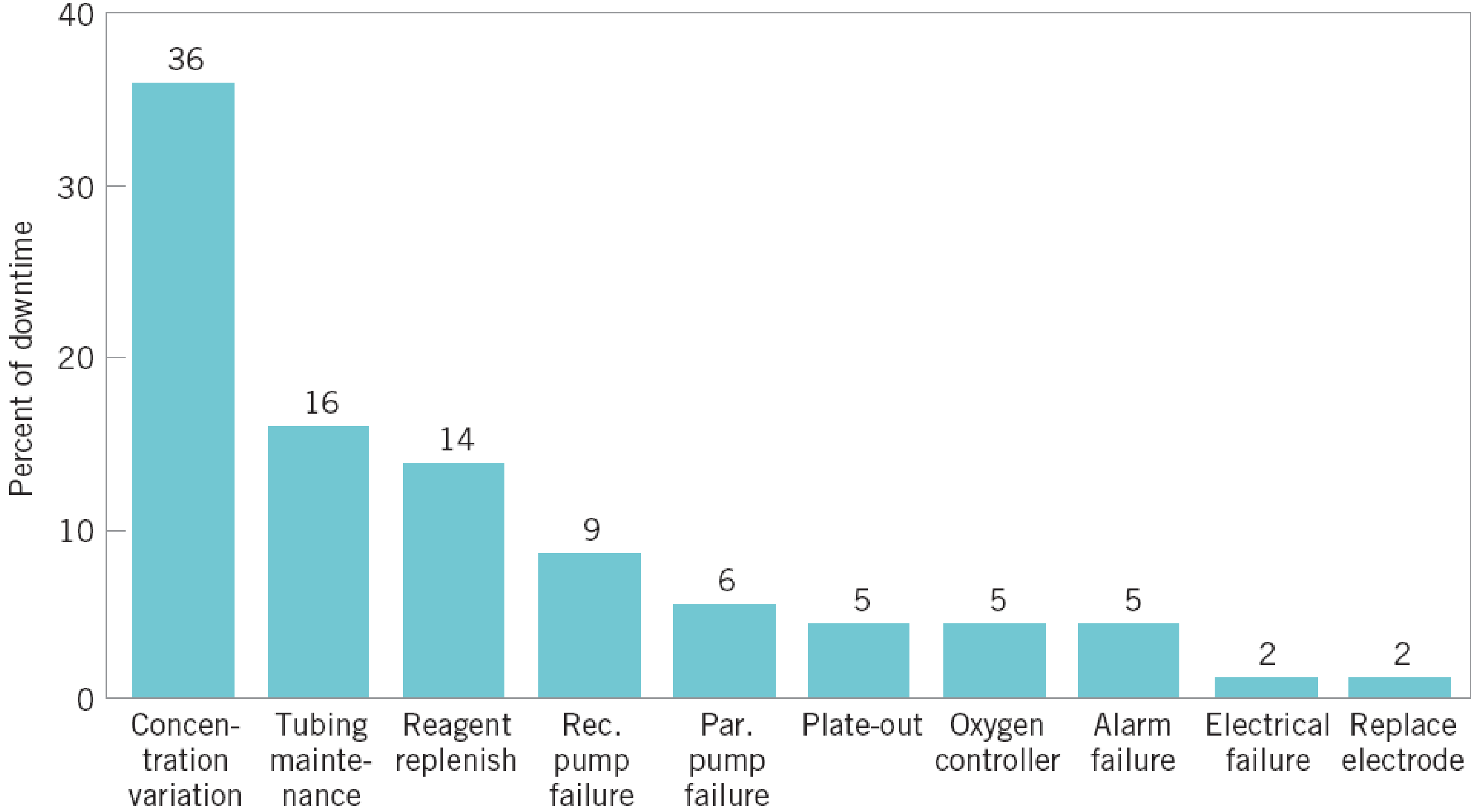


FIGURE 5.25 Pareto analysis of controller failures.

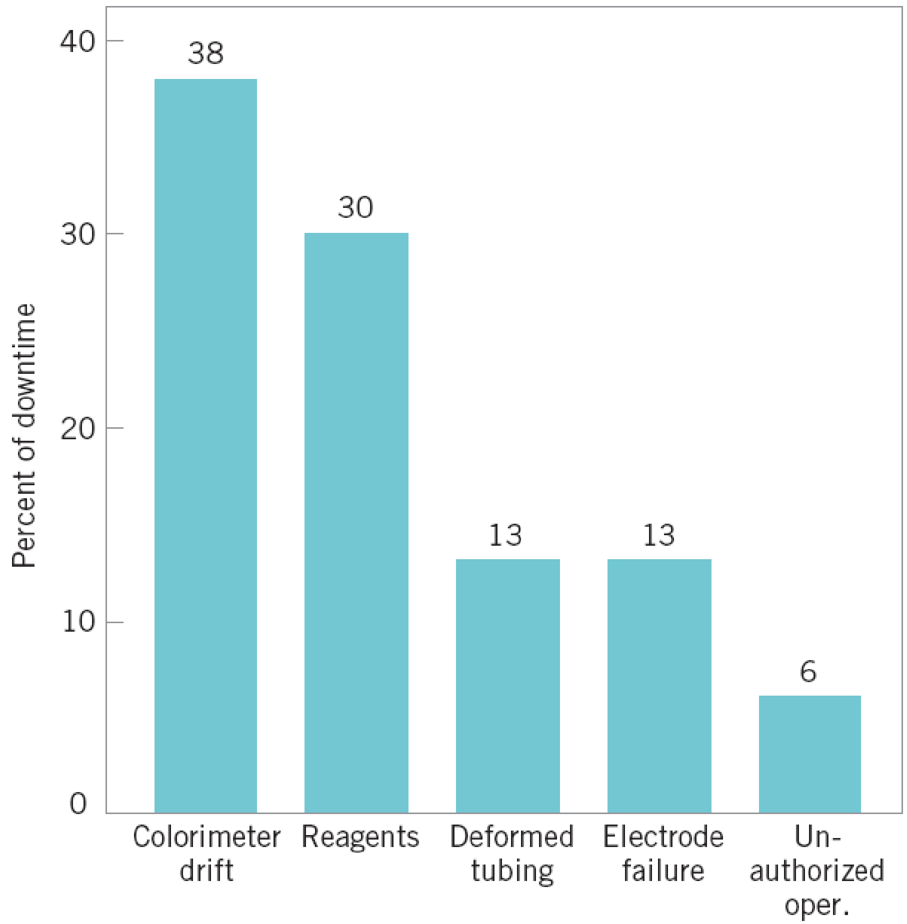


FIGURE 5.26 Pareto analysis of concentration variation.

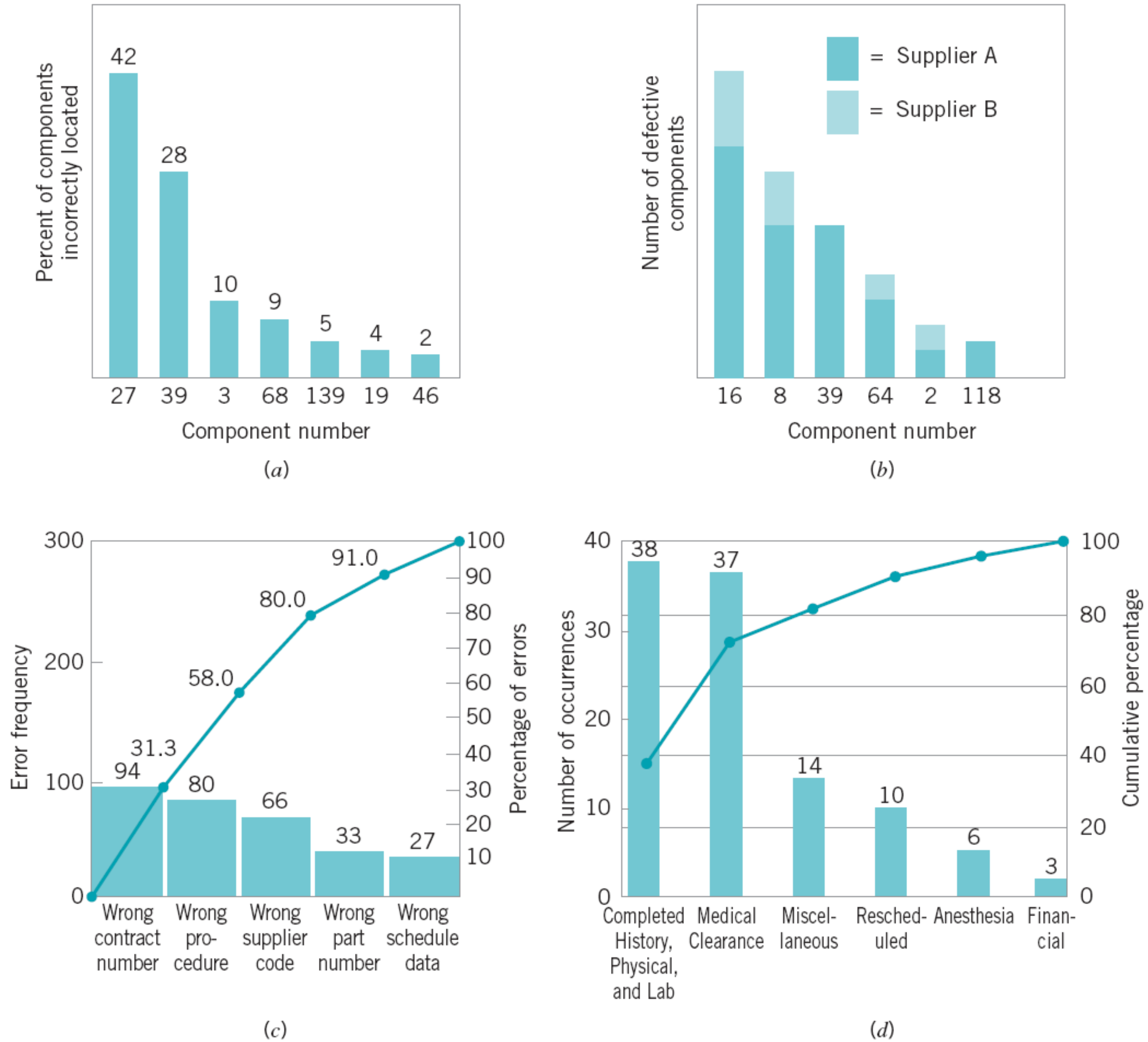


FIGURE 5.18 Examples of Pareto charts.

Scatter Diagram

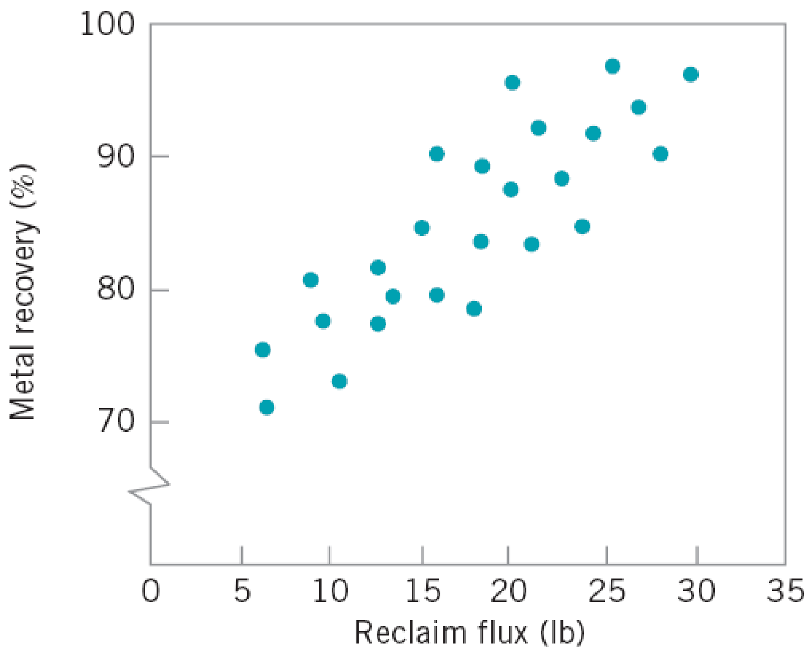


FIGURE 5.22 A scatter diagram.

Defect Concentration Diagram

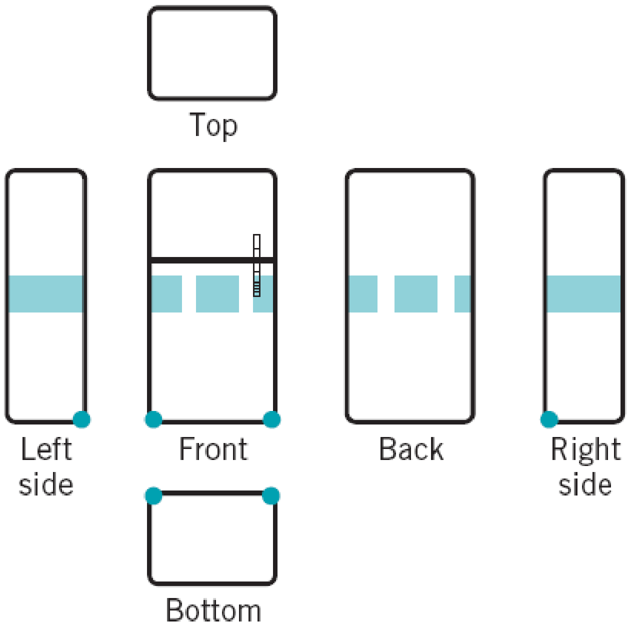


FIGURE 5.20 Surface-finish defects on a refrigerator.

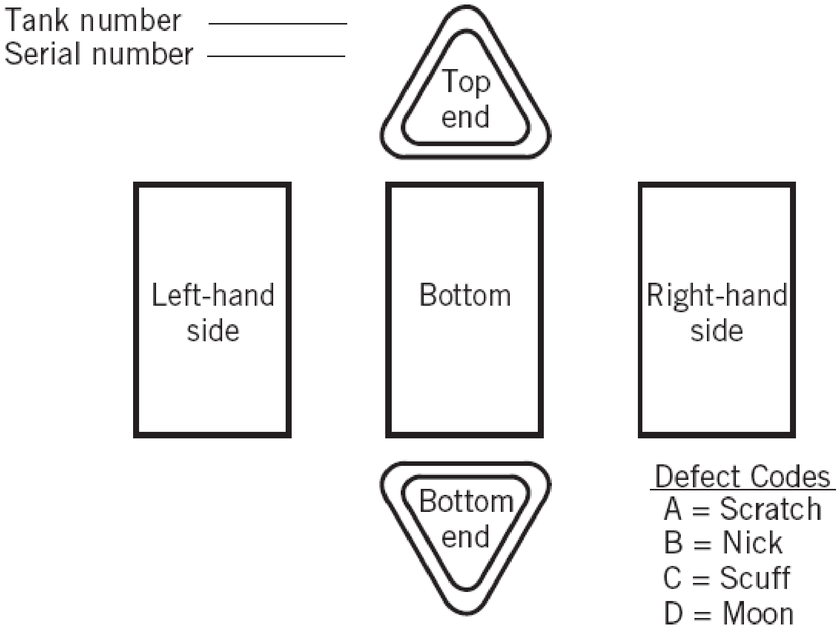


FIGURE 5.21 Defect concentration diagram for the tank.

5.2 Chance and Assignable Causes of Variation

- A process is operating with only **chance causes of variation** present is said to be **in statistical control**.
- A process that is operating in the presence of **assignable causes** is said to be **out of control**.

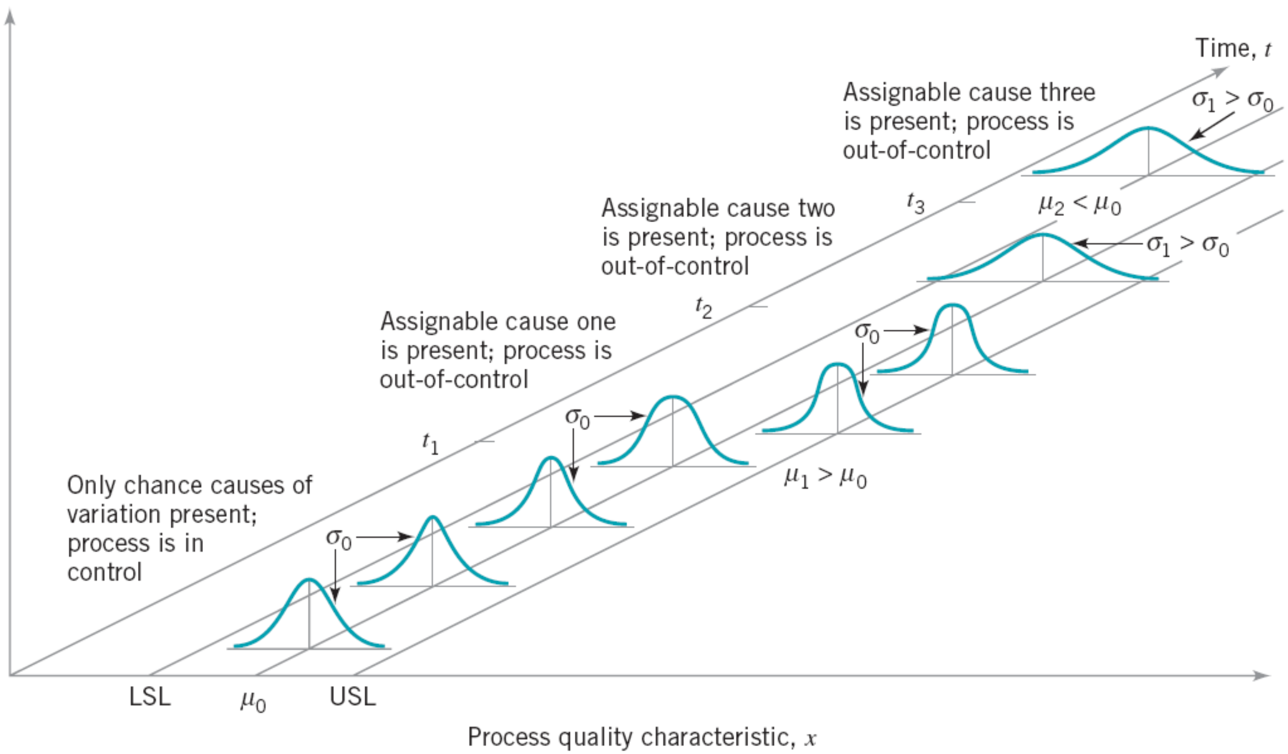
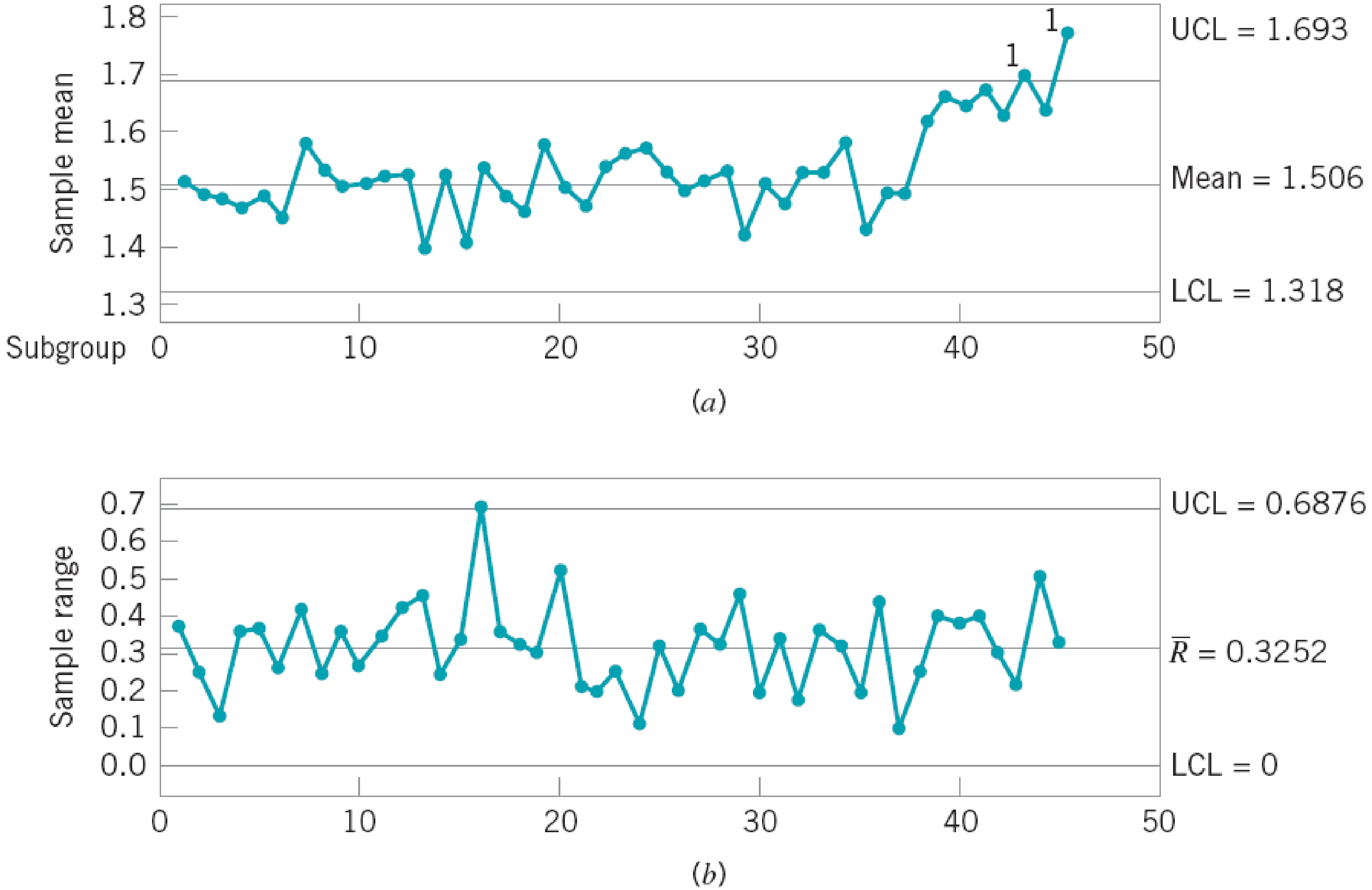


FIGURE 5.1 Chance and assignable causes of variation.



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

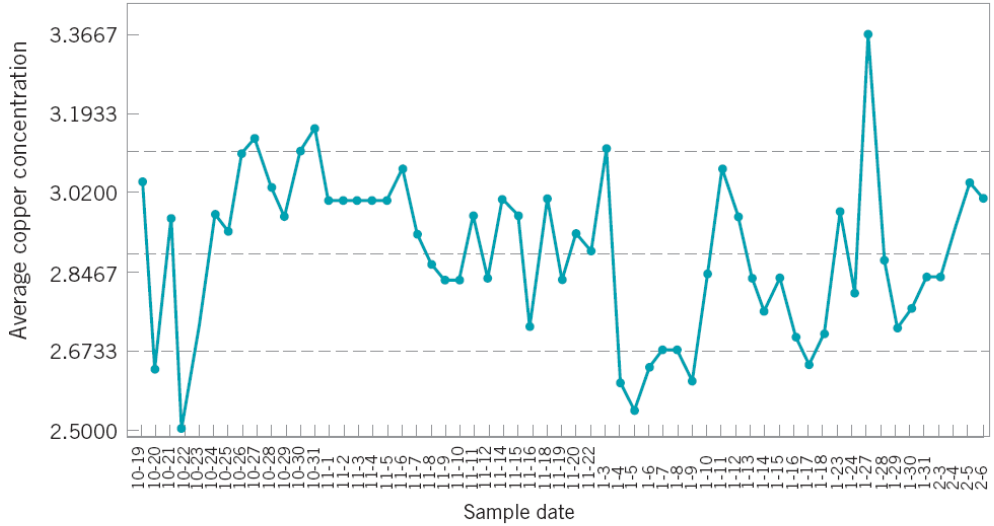


FIGURE 5.27 $\bar{x}$ chart for the average daily copper concentration.

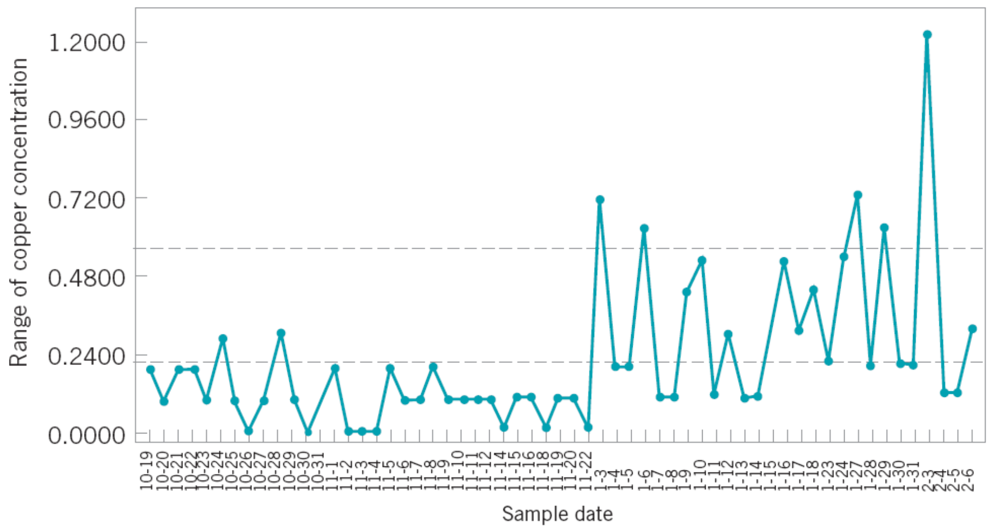


FIGURE 5.28 *R* chart for daily copper concentration.

5.3.4 Rational Subgroups

- The **rational subgroup** concept means that subgroups or samples should be selected so that if assignable causes are present, chance for differences *between* subgroups will be maximized, while chance for difference due to assignable causes *within* a subgroup will be minimized.

- Two general approaches for constructing rational subgroups:
  1. Sample consists of units produced at the same time – **consecutive** units
     - Primary purpose is to detect process shifts
  2. Sample consists of units that are representative of all units produced since last sample – **random sample of all process output over sampling interval**
     - Often used to make decisions about acceptance of product
     - Effective at detecting shifts to out-of-control state and back into in-control state *between* samples
     - Care must be taken because **we can often make any process appear to be in statistical control just by stretching out the interval between observations in the sample**.

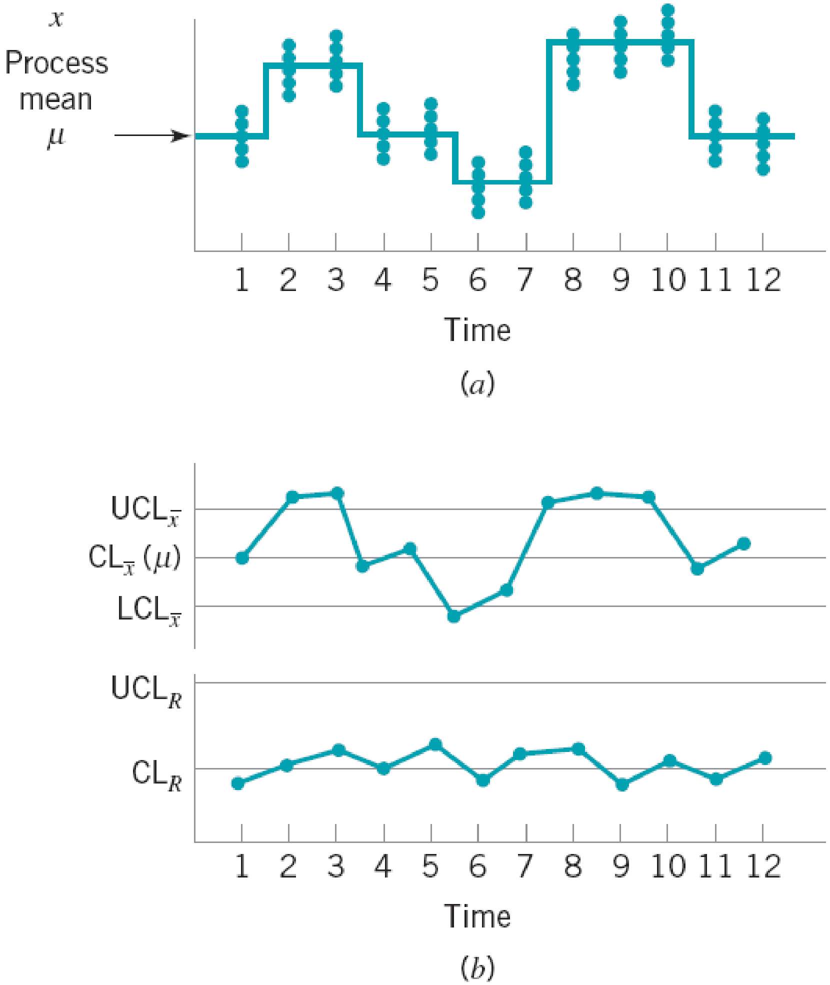


FIGURE 5.10 The snapshot approach to rational subgroups. (*a*) Behavior of the process mean. (*b*) Corresponding $\bar{x}$ and R control charts.

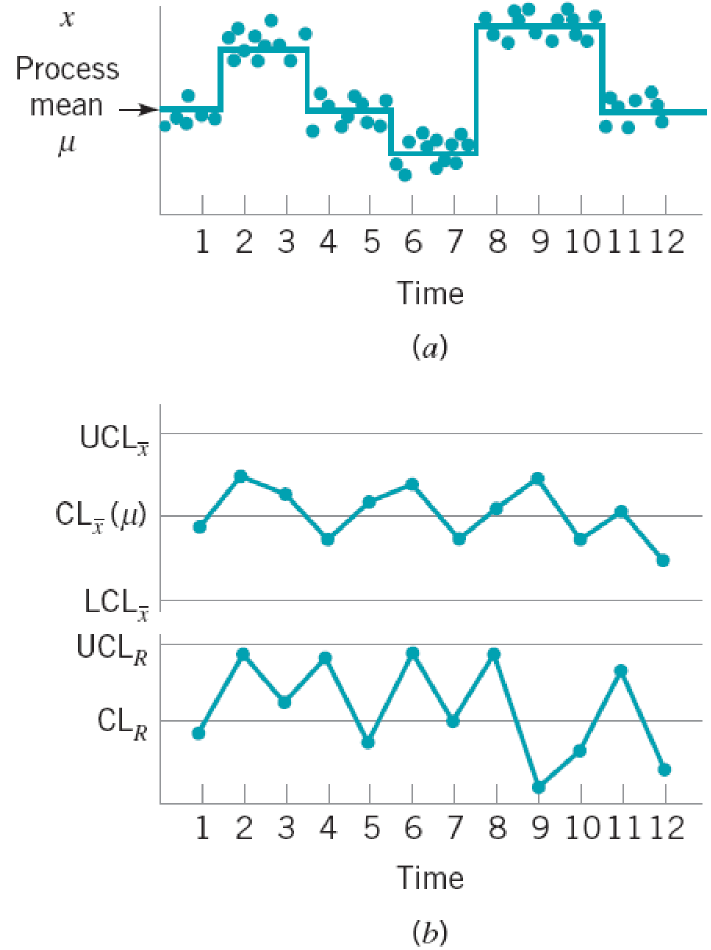


FIGURE 5.11 The random sample approach to rational subgroups. (*a*) Behavior of the process mean. (*b*) Corresponding $\bar{x}$ and R control charts.

5.3 Statistical Basis of the Control Chart

- A control chart contains
  - A **center line**
  - An **upper control limit**
  - A **lower control limit**
- A point that plots within the control limits indicates the process is in control
  - No action is necessary
- A point that plots outside the control limits is evidence that the process is out of control
  - Investigation and corrective action are required to find and eliminate assignable cause(s)
- There is a close connection between control charts and **hypothesis testing**

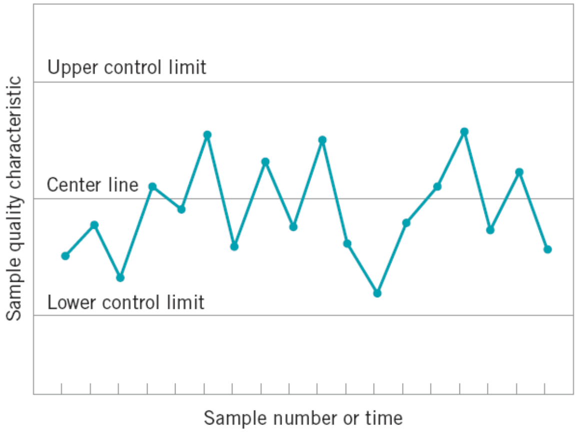


FIGURE 5.2 A typical control chart.

Shewhart Control Chart Model

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

$$\begin{aligned} \text{UCL} &= \mu_w + L\sigma_w \\ \text{Center line} &= \mu_w \\ \text{LCL} &= \mu_w - L\sigma_w \end{aligned} \tag{5.1}$$

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**.

Photolithography Example

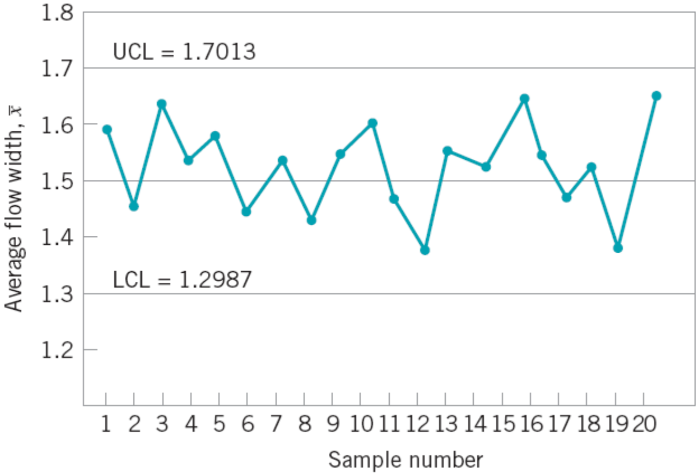


FIGURE 5.3 $\bar{x}$ control chart for flow width.

- Important quality characteristic in hard bake is resist flow width
- Process is monitored by average flow width
  - Sample of 5 wafers
  - Process mean is 1.5 microns
  - Process standard deviation is 0.15 microns
- Note that all plotted points fall inside the control limits
  - Process is considered to be in statistical control

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

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

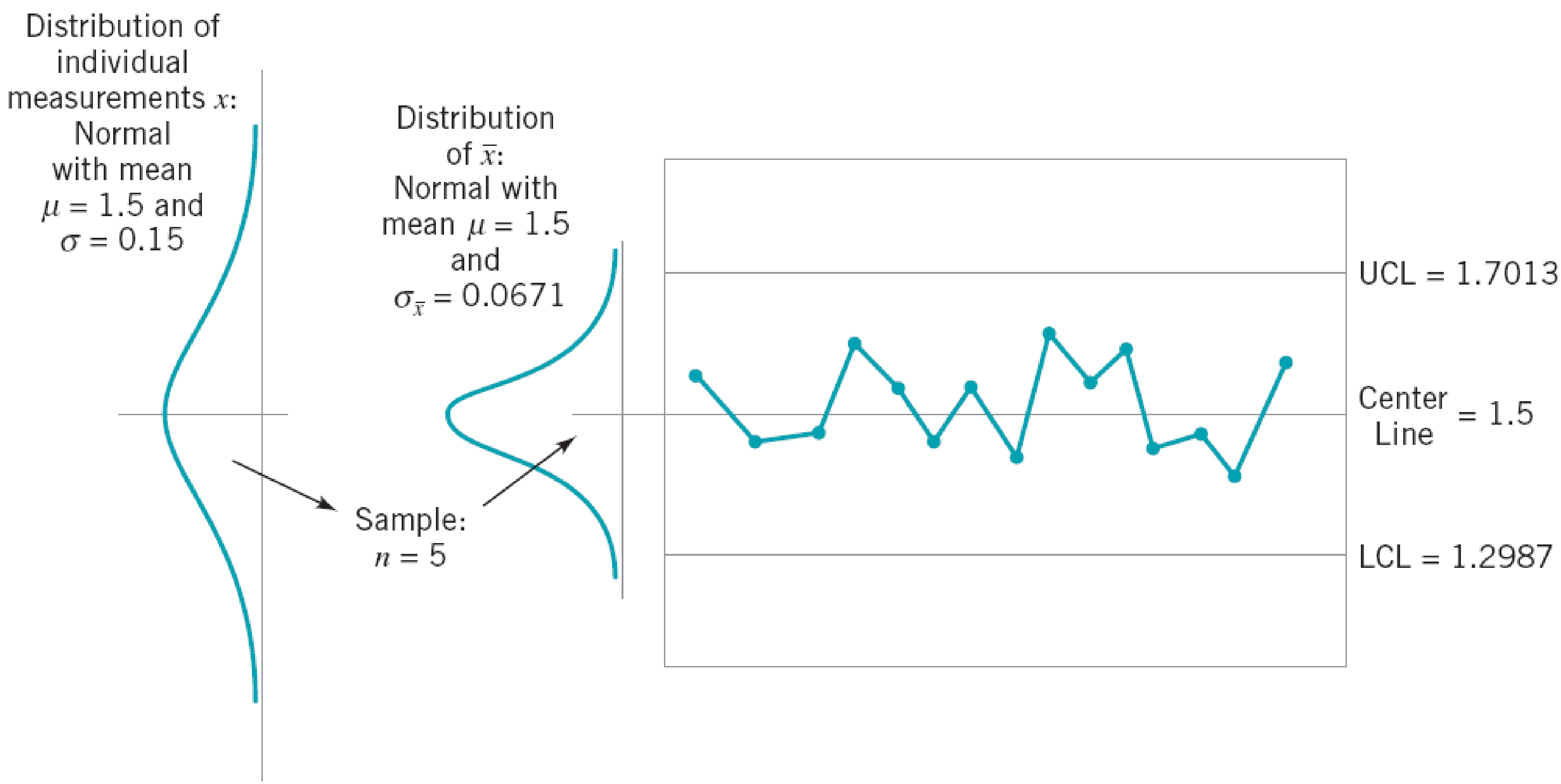
Therefore, if the process is in control with a mean flow width of 1.5 microns, then by using the central limit theorem to assume that $\bar{x}$ is approximately normally distributed, we would expect $100(1 - \alpha)\%$ of the sample means $\bar{x}$ to fall between $1.5 + Z_{\alpha/2}(0.0671)$ and $1.5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



■ **FIGURE 5.4** How the control chart works.

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.

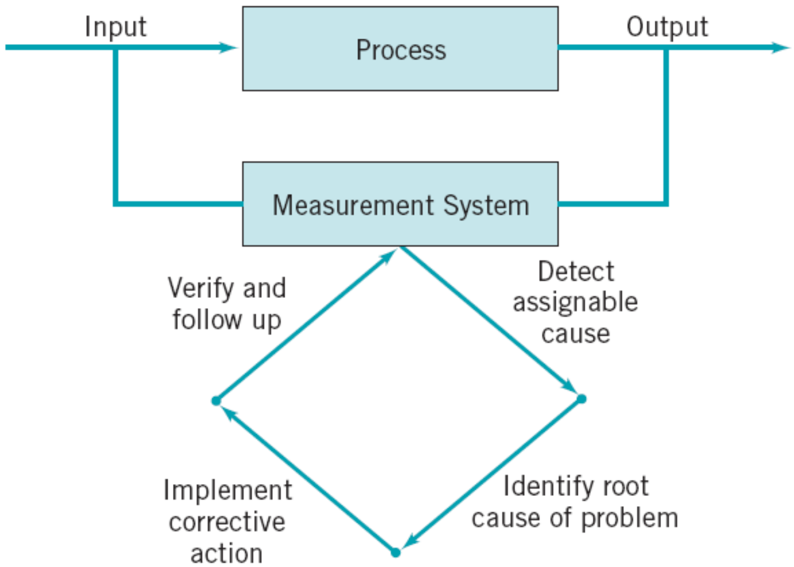


FIGURE 5.5 Process improvement using the control chart.

More Basic Principles

- Charts may be used to estimate process parameters, which are used to determine **capability**
- Two general types of control charts
  - Variables (Chapter 6)
    - Continuous scale of measurement
    - Quality characteristic described by central tendency and a measure of variability
  - Attributes (Chapter 7)
    - Conforming/nonconforming
    - Counts
- **Control chart design** encompasses selection of sample size, control limits, and sampling frequency

Types of Process Variability

- **Stationary** and **uncorrelated** – data vary around a fixed mean in a stable or predictable manner
- **Stationary** and **autocorrelated** – successive observations are dependent with tendency to move in long runs on either side of mean
- **Nonstationary** – process drifts without any sense of a stable or fixed mean

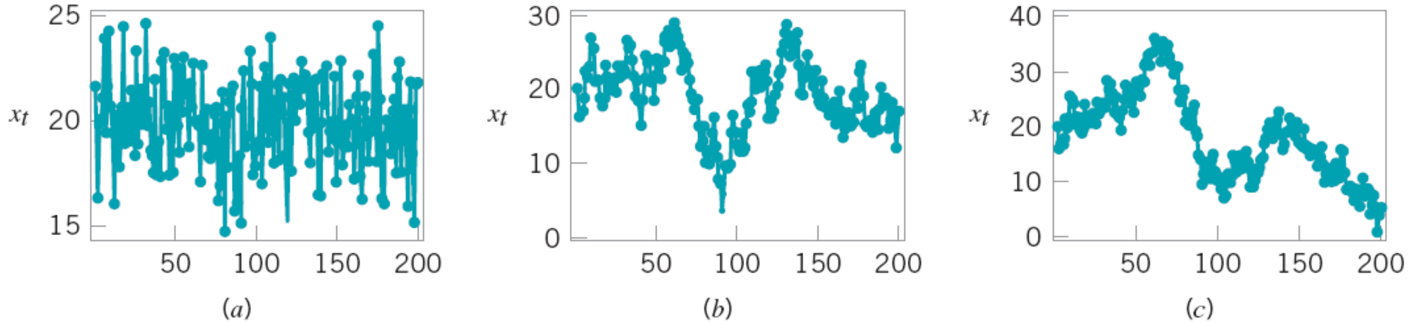


FIGURE 5.7 Data from three different processes. (*a*) Stationary and uncorrelated (white noise). (*b*) Stationary and autocorrelated. (*c*) Nonstationary.

Reasons for Popularity of Control Charts

1. Control charts are a proven technique for improving productivity.
2. Control charts are effective in defect prevention.
3. Control charts prevent unnecessary process adjustment.
4. Control charts provide diagnostic information.
5. Control charts provide information about process capability.

4-3.2 Choice of Control Limits

- 3-Sigma Control Limits
  - Probability of type I error is 0.0027
- Probability Limits
  - Type I error probability is chosen directly
  - For example, 0.001 gives 3.09-sigma control limits
- Warning Limits
  - Typically selected as 2-sigma limits

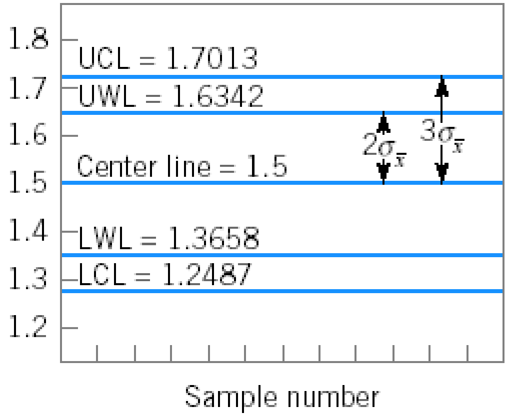


Figure 4-8 An $\bar{x}$ chart with two-sigma warning limits.

4.3.7 Phase I and Phase II of Control Chart Application

- Phase I is a **retrospective analysis** of process data to construct **trial control limits**
  - Charts are effective at detecting large, sustained shifts in process parameters, outliers, measurement errors, data entry errors, etc.
  - Facilitates identification and removal of assignable causes
- In phase II, the control chart is used to **monitor** the process
  - Process is assumed to be reasonably stable
  - Emphasis is on **process monitoring**, not on bringing an unruly process into control

5.3.3 Sample Size and Sampling Frequency

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

$$\mathrm{ARL} = \frac{1}{p} \tag{5.2}$$

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

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

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

$$\text{ATS} = \text{ARL}h \tag{5.3}$$

Consider the hard-bake process discussed earlier, and suppose we are sampling every hour. Equation (5.3) 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. 5.9 we find that if the process mean is 1.725 microns, the probability of $\bar{x}$ falling between the control limits is approximately 0.35. Therefore, p in equation (5.2) is 0.35, and the out-of-control ARL (called ARL_1) is

$$\mathrm{ARL}_1 = \frac{1}{p} = \frac{1}{0.35} = 2.86$$

That is, the control chart will require 2.86 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

$$\mathrm{ATS} = \mathrm{ARL}_1 h = 2.86\,(1) = 2.86 \text{ hours}$$

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 = $ARL_1\ h = 2.86(\frac{1}{2}) = 1.43$; that is, only 1.43 hours 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. 5.9 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 (5.2) the out-of-control ARL or ARL_1 is

$$ARL_1 = \frac{1}{p} = \frac{1}{0.9} = 1.11$$

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

$$ATS = ARL_1 h = 1.11(1) = 1.11 \text{ hours}$$

Thus, the larger sample size would allow the shift to be detected more quickly than with the smaller one.

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

5.3.5 Patterns on Control Charts

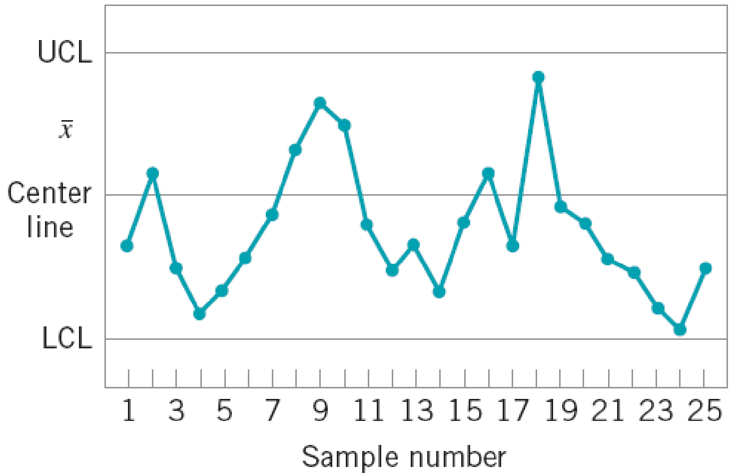


FIGURE 5.12 An $\bar{x}$ control chart.

- Pattern is very nonrandom in appearance
- 19 of 25 points plot below the center line, while only 6 plot above
- Following 4th point, 5 points in a row increase in magnitude, a *run up*
- There is also an unusually long *run down* beginning with 18th point

The Cyclic Pattern

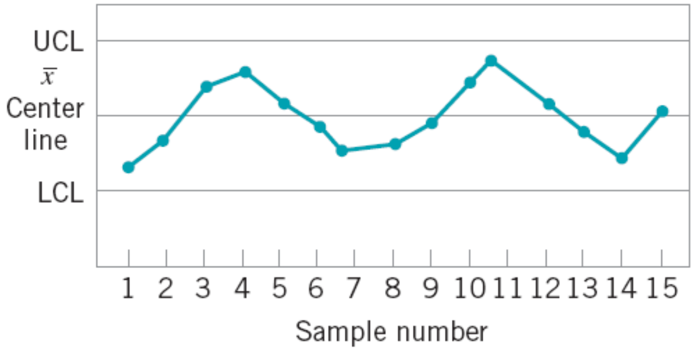


FIGURE 5.13 An $\bar{x}$ chart with a cyclic pattern.

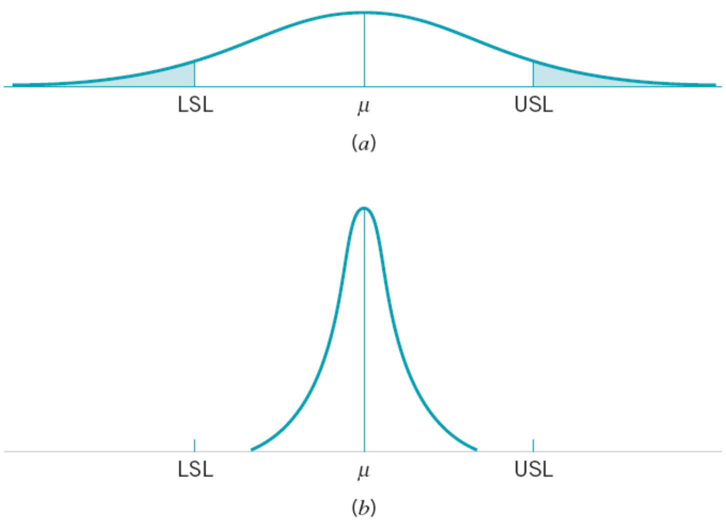


FIGURE 5.14 (*a*) Variability with the cyclic pattern. (*b*) Variability with the cyclic pattern eliminated.

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.

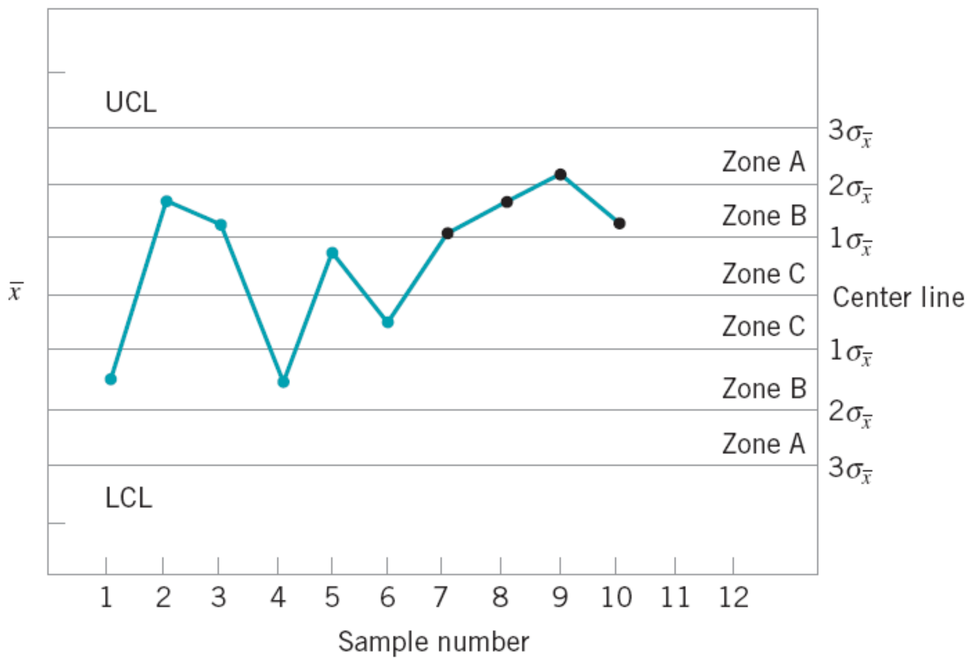


FIGURE 5.15 The Western Electric or zone rules, with the last four points showing a violation of rule 3.

5.3.6 Discussion of the Sensitizing Rules

■ **TABLE 5.1**

Some Sensitizing Rules for Shewhart Control Charts

Standard Action Signal:		
	1. One or more points outside of the control limits.	Western Electric Rules (rules 1–4)
	2. Two of three consecutive points outside the two-sigma warning limits but still inside the control limits.	
	3. Four of five consecutive points beyond the one-sigma limits.	
	4. A run of eight consecutive points on one side of the center line.	
	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.	

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

$$\alpha = 1 - \prod_{i=1}^{k} (1 - \alpha_i) \tag{5.4}$$

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.

See Champ and Woodall (1987)

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.