Chapter 21

Process Monitoring

Introduction

- Process monitoring also plays a key role in ensuring that the plant performance satisfies the operating objectives.

- The general objectives of process monitoring are:

  1. *Routine Monitoring*. Ensure that process variables are within specified limits.

  2. *Detection and Diagnosis*. Detect abnormal process operation and diagnose the root cause.

  3. *Preventive Monitoring*. Detect abnormal situations early enough so that corrective action can be taken before the process is seriously upset.
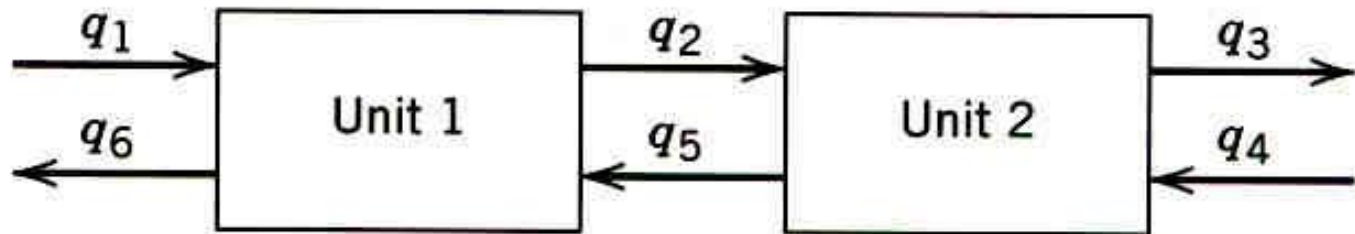
Figure 21.2 Countercurrent flow process.

Traditional Monitoring Techniques

Limit Checking

Process measurements should be checked to ensure that they are between specified limits, a procedure referred to as *limit checking*. The most common types of measurement limits are:

1. High and low limits

2. High limit for the absolute value of the rate of change

3. Low limit for the sample variance

The limits are specified based on safety and environmental considerations, operating objectives, and equipment limitations.

- In practice, there are physical limitations on how much a measurement can change between consecutive sampling instances.

- Both redundant measurements and conservation equations can be used to good advantage.

- A process consisting of two units in a countercurrent flow configuration is shown in Fig. 21.2.

- Three steady-state mass balances can be written, one for each unit plus an overall balance around both units.

- Although the three balances are not independent, they provide useful information for monitoring purposes.

- Industrial processes inevitably exhibit some variability in their manufactured produces regardless of how well the processes are designed and operated.

- In statistical process control, an important distinction is made between normal (random) variability and abnormal (nonrandom) variability.

- Random variability is caused by the cumulative effects of a number of largely unavoidable phenomena such as electrical measurement noise, turbulence, and random fluctuations in feedstock or catalyst preparation.

- The source of this abnormal variability is referred to as a *special cause* or an *assignable cause*.

Normal Distribution

- Because the *normal distribution* plays a central role in SPC, we briefly review its important characteristics.

- The normal distribution is also known as the *Gaussian distribution*.

Suppose that a random variable x has a normal distribution with a mean μ and a variance σ^2 denoted by $N\left(\mu, \sigma^2\right)$. The probability that x has a value between two arbitrary constants, a and b, is given by

$$P\left(a < x < b\right) = \int_a^b f\left(x\right) dx \qquad (21\text{-}4)$$

where $f(x)$ is the probability density function for the normal distribution:

$$f\left(x\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{\left(x-u\right)^2}{20^2}\right] \qquad (21\text{-}5)$$

The following probability statements are valid for the normal distribution (Montgomery and Runger, 2003),

$$P\left(\mu - \sigma < x < \mu + \sigma\right) \ \ = 0.6827$$

$$P\left(\mu - 2\sigma < x < \mu + 2\sigma\right) = 0.9545 \qquad (21\text{-}6)$$

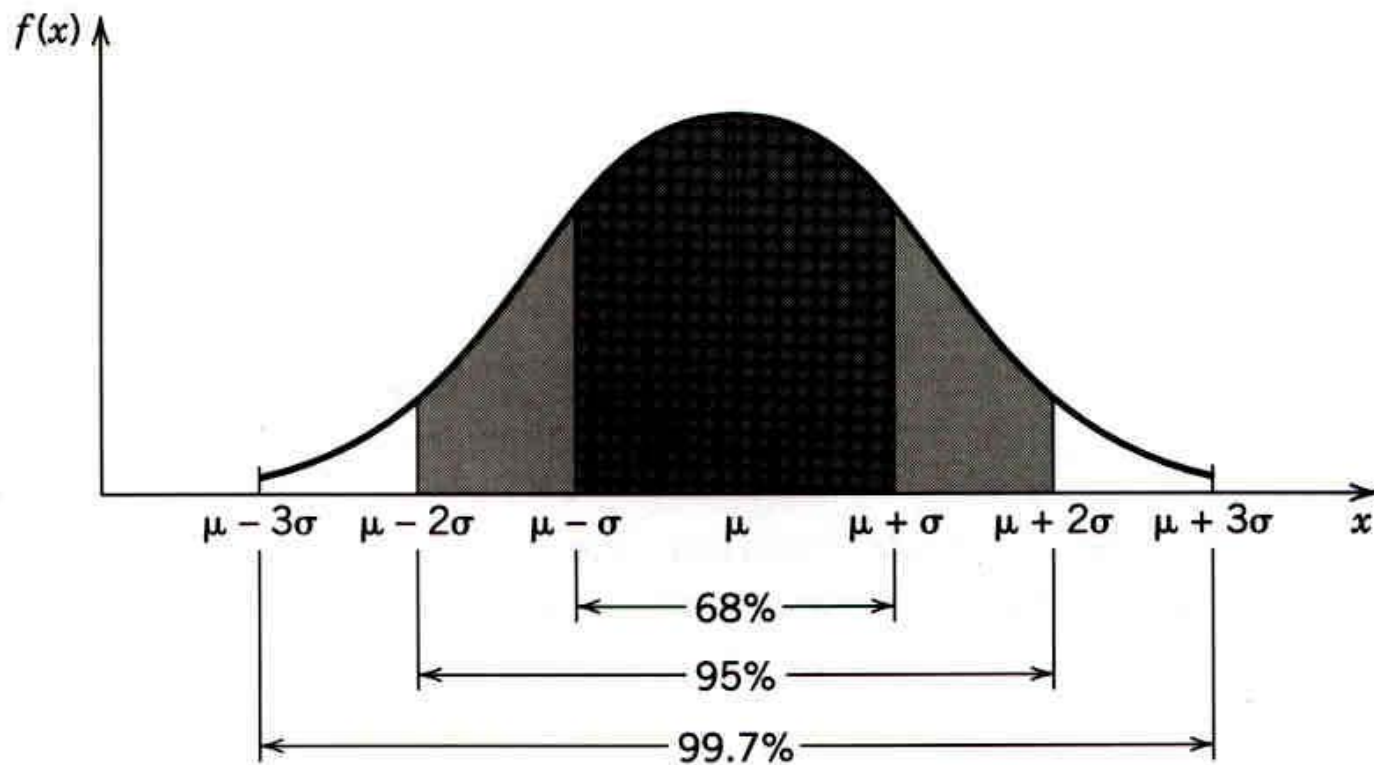$$P\left(\mu - 3\sigma < x < \mu + 3\sigma\right) = 0.9973$$

Figure 21.3 Probabilities associated with the normal distribution (From Montgomery and Runger (2003)).

- For the sake of generality, the tables are expressed in terms of the *standard normal distribution*, $N(0, 1)$, and the *standard normal variable*, $z \triangleq (x - \mu)/\sigma$.

- It is important to distinguish between the theoretical mean μ, and the sample mean $\bar{x}$.

- If measurements of a process variable are normally distributed, $N\left(\mu, \sigma^2\right)$ the sample mean is also normally distributed.

- However, for any particular sample, $\bar{x}$ is not necessarily equal to μ.

The $\bar{x}$ Control Chart

In statistical process control, *Control Charts* (or *Quality Control Charts*) are used to determine whether the process operation is normal or abnormal. The widely used $\bar{x}$ control chart is introduced in the following example.

9

This type of control chart is often referred to as a Shewhart Chart, in honor of the pioneering statistician, Walter Shewhart, who first developed it in the 1920s.

Example 21.1

A manufacturing plant produces 10,000 plastic bottles per day. Because the product is inexpensive and the plant operation is normally satisfactory, it is not economically feasible to inspect every bottle. Instead, a sample of n bottles is randomly selected and inspected each day. These n items are called a *subgroup*, and n is referred to as the *subgroup size*. The inspection includes measuring the toughness of x of each bottle in the subgroup and calculating the sample mean $\overline{x}$.
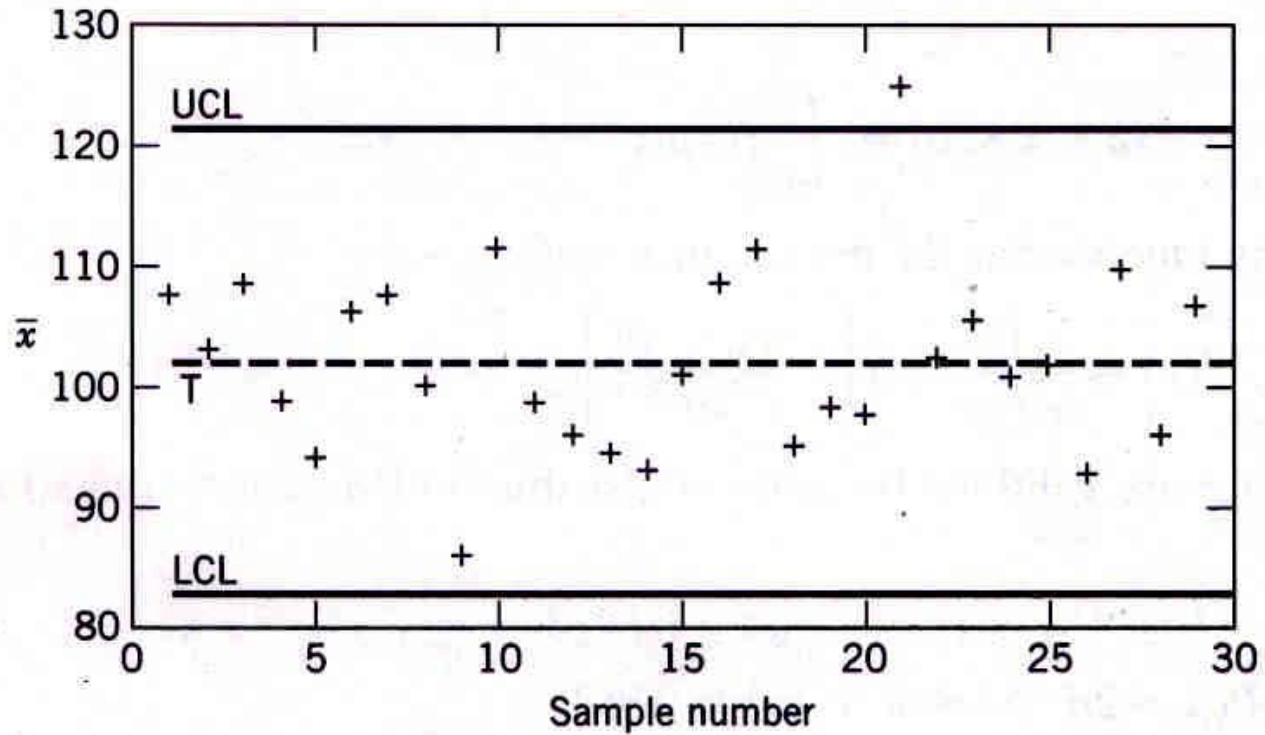
Figure 21.4 The $\bar{x}$ control chart for Example 21.1.

The $\bar{x}$ control chart in Fig. 21.4 displays data for a 30-day period. The control chart has a *target* (*T*), an *upper control limit* (*UCL*), and a *lower control limit* (*LCL*). The target (or *centerline*) is the desired (or *expected*) value for $\bar{x}$, while the region between UCL and LCL defines the range of normal variability, as discussed below. If all of the $\bar{x}$ data are within the control limits, the process operation is considered to be normal or "in a state of control". Data points outside the control limits are considered to be abnormal, indicating that the process operation is out of control. This situation occurs for the twenty-first sample. A single measurement located slightly beyond a control limit is not necessarily a cause for concern. But frequent or large chart violations should be investigated to determine a special cause.

Control Chart Development

- The first step in devising a control chart is to select a set of representative data for a period of time when the process operation is believed to be normal, that is, when the process is in a state of control.

- Suppose that these test data consist of N subgroups that have been collected on a regular basis (for example, hourly or daily) and that each subgroup consists of n randomly selected items.

- Let x_{ij} denote the *jth* measurement in the *ith* subgroup. Then, the subgroup sample means can be calculated:

$$\bar{x}_i \triangleq \frac{1}{n} \sum_{j=1}^{n} x_{ij} \quad (i = 1, 2, \ldots, N) \qquad (21\text{-}7)$$

The *grand mean* $\overline{\overline{x}}$ is defined to be the average of the subgroup means:

$$\overline{\overline{x}} \triangleq \frac{1}{N} \sum_{i=1}^{N} \overline{x}_i \qquad (21\text{-}8)$$

The general expressions for the control limits are

$$UCL \triangleq T + c\hat{\sigma}_{\overline{x}} \qquad (21\text{-}9)$$

$$LCL \triangleq T - c\hat{\sigma}_{\overline{x}} \qquad (21\text{-}10)$$

where $\hat{\sigma}_{\overline{x}}$ is an estimate of the standard deviation for $\overline{x}$ and c is a positive integer; typically, $c = 3$.

- The choice of $c = 3$ and Eq. 21-6 imply that the measurements will lie within the control chart limits 99.73% of the time, for normal process operation.

- The target T is usually specified to be either $\overline{\overline{x}}$ or the desired value of $\overline{x}$.

- The estimated standard deviation $\hat{\sigma}_{\bar{x}}$ can be calculated from the subgroups in the test data by two methods: (1) the standard deviation approach, and (2) the range approach (Montgomery and Runger, 2003).

- By definition, the *range R* is the difference between the maximum and minimum values.

- Consequently, we will only consider the standard deviation approach.

The average sample standard deviation $\bar{s}$ for the N subgroups is:

$$\bar{s} \triangleq \frac{1}{N} \sum_{i=1}^{N} s_i \qquad (21\text{-}11)$$

where the standard deviation for the *ith* subgroup is

$$s_i \triangleq \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}\left(x_{ij}-\bar{x}_i\right)^2} \qquad (21\text{-}12)$$

If the *x* data are normally distributed, then $\hat{\sigma}_{\bar{x}}$ is related to $\bar{s}$ by

$$\hat{\sigma}_{\bar{x}} = \frac{1}{c_4\sqrt{n}}\bar{s} \qquad (21\text{-}13)$$

where c_4 is a constant that depends on *n* and is tabulated in Table 21.1.

The *s* Control Chart

- In addition to monitoring average process performance, it is also advantageous to monitor process variability.

- The variability within a subgroup can be characterized by its range, standard deviation, or sample variance.

- Control charts can be developed for all three statistics but our discussion will be limited to the control chart for the standard deviation, the *s control chart*.

- The centerline for the *s* chart is $\overline{s}$, which is the average standard deviation for the test set of data. The control limits are

$$UCL = B_4\overline{s} \qquad\qquad (21\text{-}14)$$

$$LCL = B_3\overline{s} \qquad\qquad (21\text{-}15)$$

Constants B_3 and B_4 depend on the subgroup size n, as shown in Table 21.1.

Table 21.1 Control Chart Constants

	Estimation of σ		s Chart	
n	c_4		B_3	B_4
2	0.7979		0	3.267
3	0.8862		0	2.568
4	0.9213		0	2.266
5	0.9400		0	2.089
6	0.9515		0.030	1.970
7	0.9594		0.118	1.882
8	0.9650		0.185	1.815
9	0.9693		0.239	1.761
10	0.9727		0.284	1.716
15	0.9823		0.428	1.572
20	0.9869		0.510	1.490
25	0.9896		0.565	1.435

Example 21.2

In semiconductor processing, the photolithography process is used to transfer the circuit design to silicon wafers. In the first step of the process, a specified amount of a polymer solution, *photoresist,* is applied to a wafer as it spins at high speed on a turntable. The resulting photoresist thickness x is a key process variable. Thickness data for 25 subgroups are shown in Table 21.2. Each subgroup consists of three randomly selected wafers. Construct $\bar{x}$ and s control charts for these test data and critcially evaluate the results.

Solution

The following sample statistics can be calculated from the data in Table 21.2: $\bar{\bar{x}} = 199.8$ Å, $\bar{s} = 10.4$ Å. For $n = 3$ the required constants from Table 21.1 are $c_4 = 0.8862$, $B_3 = 0$, and $B_4 = 2.568$. Then the $\bar{x}$ and s control limits can be calculated from Eqs. 21-9 to 21-15.

The traditional value of $c = 3$ is selected for Eqs. (21-9) and (21-10). The resulting control limits are labeled as the "original limits" in Fig. 21.5.

Figure 21.5 indicates that sample #5 lies beyond both the UCL for both the $\bar{x}$ and s control charts, while sample #15 is very close to a control limit on each chart. Thus, the question arises whether these two samples are "outliers" that should be omitted from the analysis. Table 21.2 indicates that sample #5 includes a very large value (260.0), while sample #15 includes a very small value (150.0). However, unusually large or small numerical values by themselves do not justify discarding samples; further investigation is required.

Suppose that a more detailed evaluation has discovered a specific reason as to why measurements #5 and #15 should be discarded (e.g., faulty sensor, data misreported, etc.). In this situation, these two samples should be removed and control limits should be recalculated based on the remaining 23 samples.

These modified control limits are tabulated below as well as in Fig. 21.5.

	Original Limits	Modified Limits (omit samples #5 and #15)
$\bar{x}$ **Chart Control Limits**		
UCL	220.1	216.7
LCL	179.6	182.2
s **Chart Control Limits**		
UCL	26.6	22.7
LCL	0	0

Table 21.2 Thickness Data (in Å) for Example 21.2

| No.	x Data			$\bar{x}$	s	No.	x Data			$\bar{x}$	s
1	209.6	20.76	211.1	209.4	1.8	14	202.9	210.1	208.1	207.1	3.7
2	183.5	193.1	202.4	193.0	9.5	15	198.6	195.2	150.0	181.3	27.1
3	190.1	206.8	201.6	199.5	8.6	16	188.7	200.7	207.6	199.0	9.6
4	206.9	189.3	204.1	200.1	9.4	17	197.1	204.0	182.9	194.6	10.8
5	260.0	209.0	212.2	227.1	28.6	18	194.2	211.2	215.4	206.9	11.0
6	193.9	178.8	214.5	195.7	17.9	19	191.0	206.2	183.9	193.7	11.4
7	206.9	202.8	189.7	199.8	9.0	20	202.5	197.1	211.1	203.6	7.0
8	200.2	192.7	202.1	198.3	5.0	21	185.1	186.3	188.9	186.8	1.9
9	210.6	192.3	205.9	202.9	9.5	22	203.1	193.1	203.9	200.0	6.0
10	186.6	201.5	197.4	195.2	7.7	23	179.7	203.3	209.7	197.6	15.8
11	204.8	196.6	225.0	208.8	14.6	24	205.3	190.0	208.2	201.2	9.8
12	183.7	209.7	208.6	200.6	14.7	25	203.4	202.9	200.4	202.2	1.6
13	185.6	198.9	191.5	192.0	6.7						
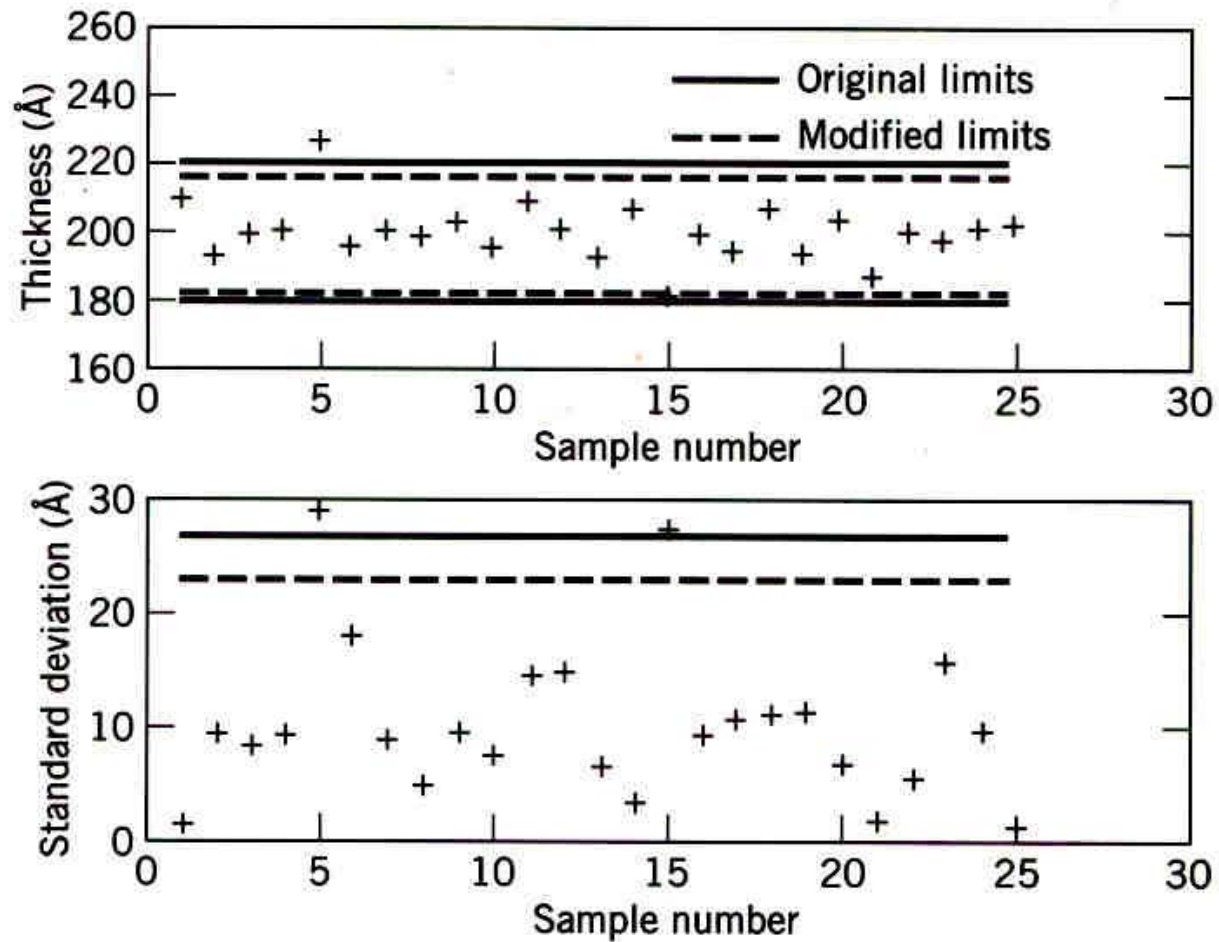
Figure 21.5 The $\bar{x}$ and s control charts for Example 21.2.

Theoretical Basis for Quality Control Charts

The traditional SPC methodology is based on the assumption that the natural variability for "in control" conditions can be characterized by random variations around a constant average value,

$$x(k) = x^* + e(k) \qquad\qquad (21\text{-}16)$$

where $x(k)$ is the measurement at time k, x^* is the true (but unknown) value, and $e(k)$ is an additive error. Traditional control charts are based on the following assumptions:

1. Each additive error, $\{e(k), k = 1, 2, \dots\}$, is a zero mean, random variable that has the same normal distribution, $N\left(0, \sigma^2\right)$.

2. The additive errors are statistically independent and thus uncorrelated. Consequently, $e(k)$ does not depend on $e(j)$ for $j \neq k$.

3. The true value of x^* is constant.

4. The subgroup size n is the same for all of the subgroups.

The second assumption is referred to as the *independent, identically, distributed* (*IID*) assumption. Consider an individuals control chart for x with x^* as its target and "3σ control limits":

$$UCL \triangleq x^* + 3\sigma \qquad\qquad (21\text{-}17)$$

$$LCL \triangleq x^* - 3\sigma \qquad\qquad (21\text{-}18)$$

- These control limits are a special case of Eqs. 21-9 and 21.10 for the idealized situation where σ is known, $c = 3$, and the subgroup size is $n = 1$.

- The typical choice of $c = 3$ can be justified as follows.

- Because x is $N\left(0, \sigma^2\right)$, the probability p that a measurement lies outside the 3σ control limits can be calculated from Eq. 21-6: $p = 1 - 0.9973 = 0.0027$.

- Thus on average, approximately 3 out of every 1000 measurements will be outside of the 3σ limits.

- The average number of samples before a chart violation occurs is referred to as the *average run length* (ARL).

- For the normal ("in control") process operation,

$$\text{ARL} \triangleq \frac{1}{p} = \frac{1}{0.0027} 370 \qquad\qquad (21\text{-}19)$$

- Thus, a Shewhart chart with 3σ control limits will have an average of one control chart violation every 370 samples, even when the process is *in a state of control*.

- Industrial plant measurements are not normally distributed.

- However, for large subgroup sizes ($n > 25$), $\overline{x}$ is approximately normally distributed even if x is not, according to the famous *Central Limit Theorem* of statistics (Montgomery and Runger, 2003).

- Fortunately, modest deviations from "normality" can be tolerated.

- In industrial applications, the control chart data are often *serially correlated* because the current measurement is related to previous measurements.

- Standard control charts such as the $\bar{x}$ and s charts can provide misleading results if the data are serially correlated.

- But if the degree of correlation is known, the control limits can be adjusted accordingly (Montgomery, 2001).

Pattern Tests and the Western Electric Rules

- We have considered how abnormal process behavior can be detected by comparing individual measurements with the $\bar{x}$ and s control chart limits.

- However, the pattern of measurements can also provide useful information.

27

- A wide variety of pattern tests (also called *zone rules*) can be developed based on the IID and normal distribution assumptions and the properties of the normal distribution.

- For example, the following excerpts from the *Western Electric Rules* indicate that the process is *out of control* if one or more of the following conditions occur:

1. One data point is outside the 3σ control limits.

2. Two out of three consecutive data points are beyond a 2σ limit.

3. Four out of five consecutive data points are beyond a 1σ limit and on one side of the center line.

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

- Pattern tests can be used to augment Shewhart charts.

- Although Shewhart charts with 3σ limits can quickly detect large process changes, they are ineffective for small, sustained process changes (for example, changes smaller than 1.5σ)

- Two alternative control charts have been developed to detect small changes: the CUSUM and EWMA control charts.

- They also can detect large process changes (for example, 3σ shifts), but detection is usually somewhat slower than for Shewhart charts.

CUSUM Control Chart

- The *cumulative sum* (*CUSUM*) is defined to be a running summation of the deviations of the plotted variable from its target.

- If the sample mean is plotted, the cumulative sum, *C(k)*, is

$$C(k) = \sum_{j=1}^{k} \left(\overline{x}(j) - T \right) \qquad (21\text{-}20)$$

where T is the target for $\bar{x}$.

- During normal process operation, $C(k)$ fluctuates around zero.

- But if a process change causes a small shift in $\bar{x}$, $C(k)$ will drift either upward or downward.

- The CUSUM control chart was originally developed using a graphical approach based on *V-masks*.

- However, for computer calculations, it is more convenient to use an equivalent algebraic version that consists of two recursive equations,

$$C^+(k) = \max\left[0, \bar{x}(k) - (T + K) + C^+(k-1)\right] \qquad (21\text{-}21)$$

$$C^-(k) = \max\left[0, (T - K) - \bar{x}(k) + C^-(k-1)\right] \qquad (21\text{-}22)$$

where C^+ and C^- denote the sums for the high and low directions and K is a constant, the *slack parameter*.

- The CUSUM calculations are initialized by setting $C^+(0) = C^-(0) = 0$.

- A deviation from the target that is larger than K increases either C^+ or C^-.

- A control limit violation occurs when either C^+ or C^- exceeds a specified control limit (or *threshold*), H.

- After a limit violation occurs, that sum is reset to zero or to a specified value.

- The selection of the threshold H can be based on considerations of average run length.

- Suppose that we want to detect whether the sample mean $\bar{x}$ has shifted from the target by a small amount, δ .

- The slack parameter K is usually specified as $K = 0.5\delta$.

- For the ideal situation where the normally distributed and IID assumptions are valid, ARL values have been tabulated for specified values of δ, K, and H (Ryan, 2000; Montgomery, 2001).

Table 21.3 Average Run Lengths for CUSUM Control Charts

Shift from Target (in multiples of $\sigma_{\bar{x}}$)	ARL for $H = 4\sigma_{\bar{x}}$	ARL for $H = 5\sigma_{\bar{x}}$
0	168.	465.
0.25	74.2	139.
0.50	26.6	38.0
0.75	13.3	17.0
1.00	8.38	10.4
2.00	3.34	4.01
3.00	2.19	2.57

EWMA Control Chart

- Information about past measurements can also be included in the control chart calculations by exponentially weighting the data.

- This strategy provides the basis for the *exponentially-weighted moving-average* (*EWMA*) *control chart.*

- Let $\bar{x}$ denote the sample mean of the measured variable and z denote the EWMA of $\bar{x}$. A recursive equation is used to calculate $z(k)$,

$$z(k) = \lambda \bar{x}(k) + (1 - \lambda) z(k-1) \qquad (21\text{-}23)$$

  where λ is a constant, $0 \le \lambda \le 1$.

- Note that Eq. 21-27 has the same form as the first-order (or exponential) filter that was introduced in Chapter 17.

- The EWMA control chart consists of a plot of $z(k)$ vs. k, as well as a target and upper and lower control limits.

- Note that the EWMA control chart reduces to a Shewhart chart for $\lambda = 1$.

- The EWMA calculations are initialized by setting $z(0) = T$.

- If the $\bar{x}$ measurements satisfy the IID condition, the EWMA control limits can be derived.

- The theoretical 3σ limits are given by

$$T \pm 3\sigma_x \sqrt{\frac{\lambda}{2 - \lambda}} \qquad (21\text{-}24)$$

where σ_x is determined from a set of test data taken when the process is *in a state of control*.

- The target T is selected to be either the desired value of $\bar{x}$ or the grand mean for the test data, $\bar{\bar{x}}$.

- Time-varying control limits can also be derived that provide narrower limits for the first few samples, for applications where early detection is important (Montgomery, 2001; Ryan, 2000).

- Tables of ARL values have been developed for the EWMA method, similar to Table 21.3 for the CUSUM method (Ryan, 2000).

- The EWMA performance can be adjusted by specifying λ.

- For example, $\lambda = 0.25$ is a reasonable choice because it results in an ARL of 493 for no mean shift ($\delta = 0$) and an ARL of 11 for a mean shift of $\sigma_x \left(\delta = 1 \right)$.

- EWMA control charts can also be constructed for measures of variability such as the range and standard deviation.

Example 21.3

In order to compare Shewhart, CUSUM, and EWMA control charts, consider simulated data for the tensile strength of a phenolic resin. It is assumed that the tensile strength x is normally distributed with a mean of $\mu = 70$ MPa and a standard deviation of $\sigma = 3$ MPa. A single measurement is available at each sampling instant. A constant $(\delta = 0.5\sigma = 1.5)$ was added to $x(k)$ for $k \geq 10$ in order to evaluate each chart's ability to detect a small process shift. The CUSUM chart was designed using $K = 0.5\sigma$ and $H = 5\sigma$, while the EWMA parameter was specified as $\lambda = 0.25$.

The relative performance of the Shewhart, CUSUM, and EWMA control charts is compared in Fig. 21.6. The Shewhart chart fails to detect the 0.5σ shift in x. However, both the CUSUM and EWMA charts quickly detect this change because limit violations occur about ten samples after the shift occurs (at $k = 20$ and $k = 21$, respectively). The mean shift can also be detected by applying the Western Electric Rules in the previous section.
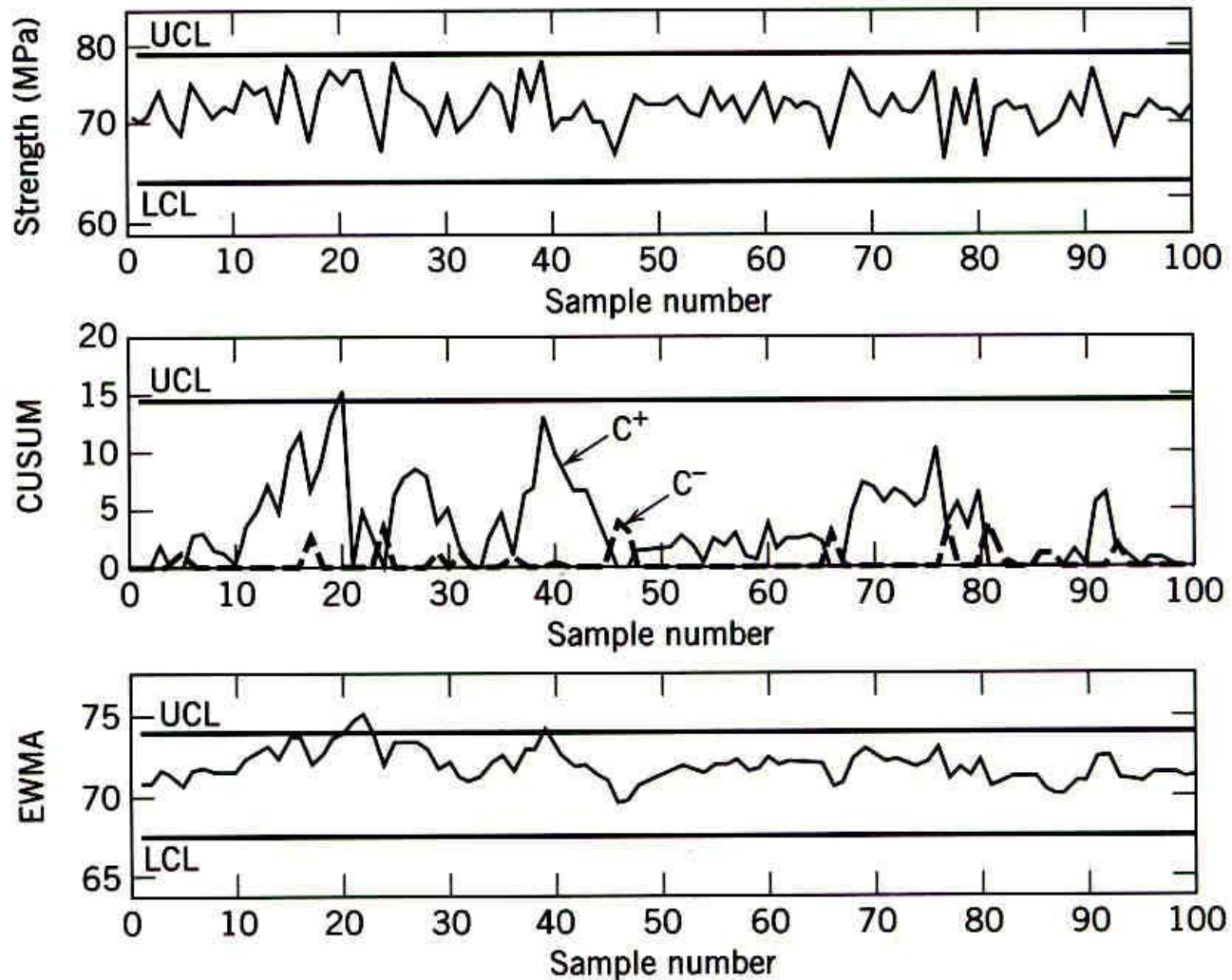
Figure 21.6 Comparison of Shewhart (top), CUSUM (middle), and EWMA (bottom) control charts for Example 21.3.

37

Process Capability Indices

- Process capability indices (or *process capability ratios*) provide a measure of whether an "in control" process is meeting its product specifications.

- Suppose that a quality variable *x* must have a volume between an *upper specification limit (USL)* and a *lower specification limit (LSL),* in order for product to satisfy customer requirements.

- The C_p *capability index* is defined as,

$$C_p \triangleq \frac{USL - LSL}{6\sigma} \qquad (21\text{-}25)$$

where σ is the standard deviation of *x*.

- Suppose that $C_p = 1$ and x is normally distributed.

- Based on Eq. 21-6, we would expect that 99.73% of the measurements satisfy the specification limits.

- If $C_p > 1$, the product specifications are satisfied; for $C_p < 1$, they are not.

- A second capability index C_{pk} is based on average process performance ($\bar{x}$), as well as process variability (σ). It is defined as:

$$C_{pk} \triangleq \frac{\min\left[\bar{x} - LSL, USL - \bar{x}\right]}{3\sigma} \qquad (21\text{-}26)$$

- Although both C_p and C_{pk} are used, we consider C_{pk} to be superior to C_p for the following reason.

- If $\bar{x} = T$, the process is said to be "centered" and $C_{pk} = C_p$.

- But for $\bar{x} \neq T$, C_p does not change, even though the process performance is worse, while C_{pk} decreases. For this reason, C_{pk} is preferred.

- In practical applications, a common objective is to have a capability index of 2.0 while a value greater than 1.5 is considered to be acceptable.

- Three important points should be noted concerning the C_p and C_{pk} capability indices:

1. The data used in the calculations do *not* have to be normally distributed.

2. The specification limits, USL and LSL, and the control limits, UCL and LCL, are *not* related. The specification limits denote the desired process performance, while the control limits represent actual performance during normal operation when the process is *in control*.

3. The numerical values of the C_p and C_{pk} capability indices in (21-25) and (21-26) are only meaningful when the process is *in a state of control*. However, other *process performance indices* are available to characterize process performance when the process is not *in a state of control*. They can be used to evaluate the incentives for improved process control (Shunta, 1995).

Example 21.4

Calculate the average values of the C_p and C_{pk} capability indices for the photolithography thickness data in Example 21.2. Omit the two outliers (samples #5 and #15) and assume that the upper and lower specification limits for the photoresist thickness are USL=235 Å and LSL = 185 Å.

Solution

After samples #5 and #15 are omitted, the grand mean is $\bar{\bar{x}} = 199\text{Å}$, and the standard deviation of $\bar{x}$ (estimated from Eq. (21-13) with $c_4 = 0.8862$) is

$$\hat{\sigma}_x = \frac{\bar{s}}{c_4 \sqrt{n}} = \frac{8.83}{0.8862\sqrt{3}} = 5.75\text{Å}$$

From Eqs. 21-25 and 21-26,

$$C_p = \frac{235 - 185}{6(5.75)} = 1.45$$

$$C_{pk} = \frac{\min[199.5 - 185, 235 - 199.5]}{3(5.75)} = 0.84$$

Note the C_{pk} is much smaller than the C_p because $\bar{\bar{x}}$ is closer to the LSL than the USL.

Six Sigma Approach

- Product quality specifications continue to become more stringent as a result of market demands and intense worldwide competition.

- Meeting quality requirements is especially difficult for products that consist of a very large number of components and for manufacturing processes that consist of hundreds of individual steps.

- For example, the production of a microelectronics device typically requires 100-300 batch processing steps.

- Suppose that there are 200 steps and that each one must meet a quality specification in order for the final product to function properly.

- If each step is independent of the others and has a 99% success rate, the overall yield of satisfactory product is $(0.99)^{200} = 0.134$ or only 13.4%.

Six Sigma Approach

•   This low yield is clearly unsatisfactory.

• Similarly, even when a processing step meets 3σ specifications (99.73% success rate), it will still result in an average of 2700 "defects" for every million produced.

• Furthermore, the overall yield for this 200-step process is still only 58.2%.

• Suppose that a product quality variable x is normally distributed, $N\left(\mu, \sigma^2\right)$.

• As indicated on the left portion of Fig. 21.7, if the product specifications are $\mu \pm 6\sigma$, the product will meet the specifications 99.999998% of the time.

• Thus, on average, there will only be two defective products for every billion produced.

- Now suppose that the process operation changes so that the mean value is shifted from $\overline{x} = \mu$ to either $\overline{x} = \mu + 1.5\sigma$ or $\overline{x} = \mu - 1.5\sigma$, as shown on the right side of Fig. 21.7.

- Then the product specifications will still be satisfied 99.99966% of the time, which corresponds to 3.4 defective products per million produced.

- In summary, if the variability of a manufacturing operation is so small that the product specification limits are equal to $\mu \pm 6\sigma$, then the limits can be satisfied even if the mean value of x shifts by as much as 1.5σ.

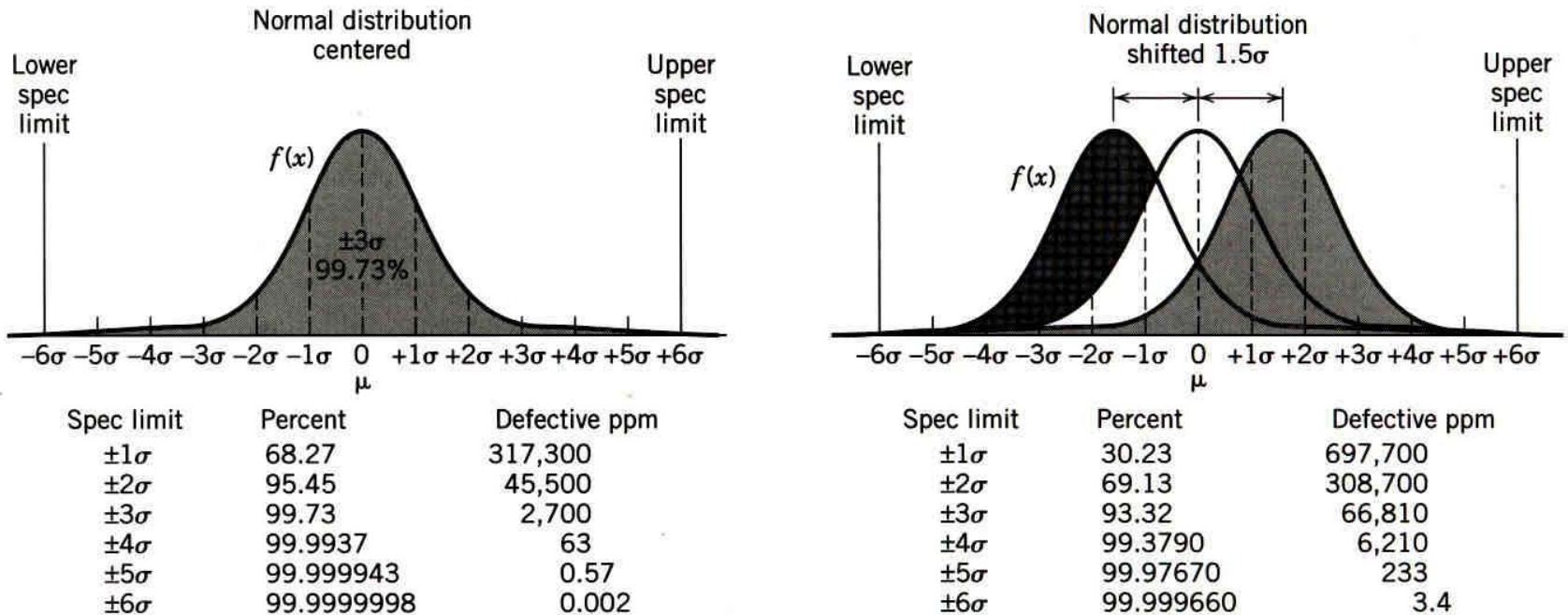- This very desirable situation of near perfect product quality is referred to as *six sigma quality*.

Figure 21.7 The Six Sigma Concept (Montgomery and Runger, 2003). Left: No shift in the mean. Right: 1.5σ shift.

Comparison of Statistical Process Control and Automatic Process Control

- Statistical process control and automatic process control (APC) are complementary techniques that were developed for different types of problems.

- APC is widely used in the process industries because no information is required about the source and type of process disturbances.

- APC is most effective when the measurement sampling period is relatively short compared to the process settling time and when the process disturbances tend to be deterministic (that is, when they have a sustained nature such as a step or ramp disturbance).

- In statistical process control, the objective is to decide whether the process is behaving normally and to identify a special cause when it is not.

- In contrast to APC, no corrective action is taken when the measurements are within the control chart limits.

- From an engineering perspective, SPC is viewed as a *monitoring* rather than a *control* strategy.

- It is very effective when the normal process operation can be characterized by random fluctuations around a mean value.

- SPC is an appropriate choice for monitoring problems where the sampling period is long compared to the process settling time, and the process disturbances tend to be random rather than deterministic.

- SPC has been widely used for quality control in both discrete-parts manufacturing and the process industries.

- In summary, SPC and APC should be regarded as complementary rather than competitive techniques.

- They were developed for different types of situations and have been successfully used in the process industries.

- Furthermore, a combination of the two methods can be very effective.

Multivariate Statistical Techniques

- For common SPC monitoring problems, two or more quality variables are important, and they can be highly correlated.

- For example, ten or more quality variables are typically measured for synthetic fibers.

- For these situations, multivariable SPC techniques can offer significant advantages over the single-variable methods discussed in Section 21.2.

- In the statistics literature, these techniques are referred to as *multivariate methods*, while the standard Shewhart and CUSUM control charts are examples of *univariate methods*.

Example 21.5

The effluent stream from a wastewater treatment process is monitored to make sure that two process variables, the biological oxidation demand (BOD) and the solids content, meet specifications. Representative data are shown in Table 21.4. Shewhart charts for the sample means are shown in parts (a) and (b) of Fig. 21.8. These univariate control charts indicate that the process appears to be in-control because no chart violations occur for either variable. However, the bivariate control chart in Fig. 21.8c indicates that the two variables are highly correlated because the solids content tends to be large when the BOD is large and vice versa. When the two variables are considered together, their joint confidence limit (for example, at the 99% confidence level) is an ellipse, as shown in Fig. 21.8c. Sample # 8 lies well beyond the 99% limit, indicating an out-of-control condition.

By contrast, this sample lies within the Shewhart control chart limits for both individual variables. This example has demonstrated that univariate SPC techniques such as Shewhart charts can fail to detect abnormal process behavior when the process variables are highly correlated. By contrast, the abnormal situation was readily apparent from the multivariate analysis.

Table 21.4 Wastewater Treatment Data

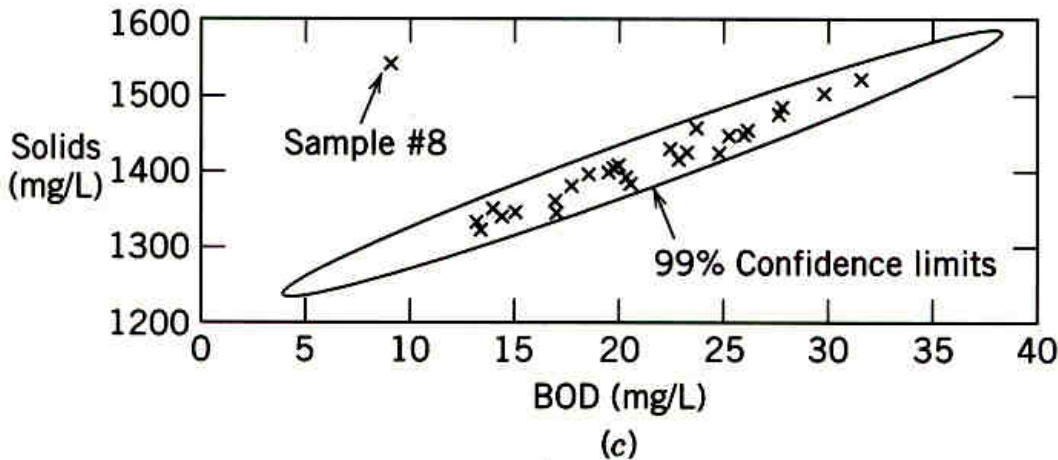| Sample Number	BOD (mg/L)	Solids (mg/L)	Sample Number	BOD (mg/L)	Solids (mg/L)
1	17.7	1380	16	16.8	1345
2	23.6	1458	17	13.8	1349
3	13.2	1322	18	19.4	1398
4	25.2	1448	19	24.7	1426
5	13.1	1334	20	16.8	1361
6	27.8	1485	21	14.9	1347
7	29.8	1503	22	27.6	1476
8	9.0	1540	23	26.1	1454
9	14.3	1341	24	20.0	1393
10	26.0	1448	25	22.9	1427
11	23.2	1426	26	22.4	1431
12	22.8	1417	27	19.6	1405
13	20.4	1384	28	31.5	1521
14	17.5	1380	29	19.9	1409
15	18.4	1396	30	20.3	1392
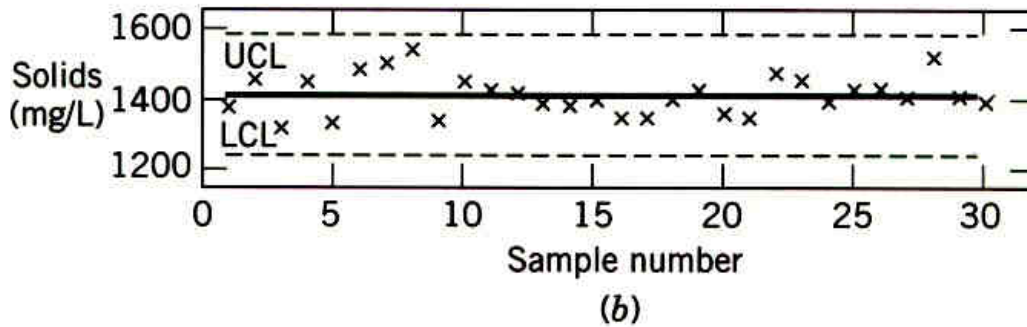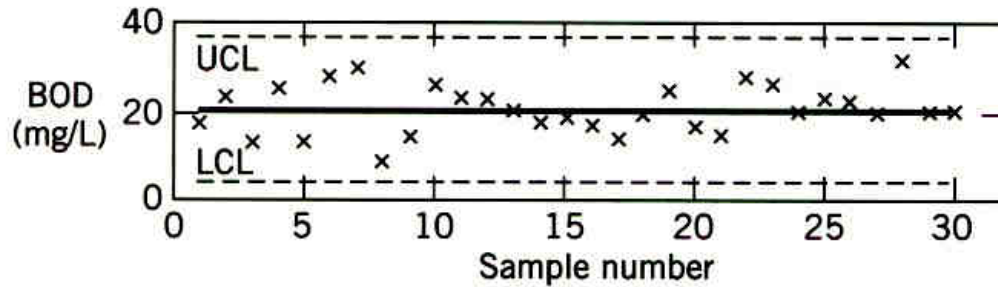
Figure 21.8 Confidence regions for Example 21.5 univariate (a) and (b), bivariate (c).

In control area correctly indicated by both types of charts.

In control area incorrectly indicated as *out of control* by the univariate charts.

Out of control area incorrectly indicated as *in control* by the univariate charts.

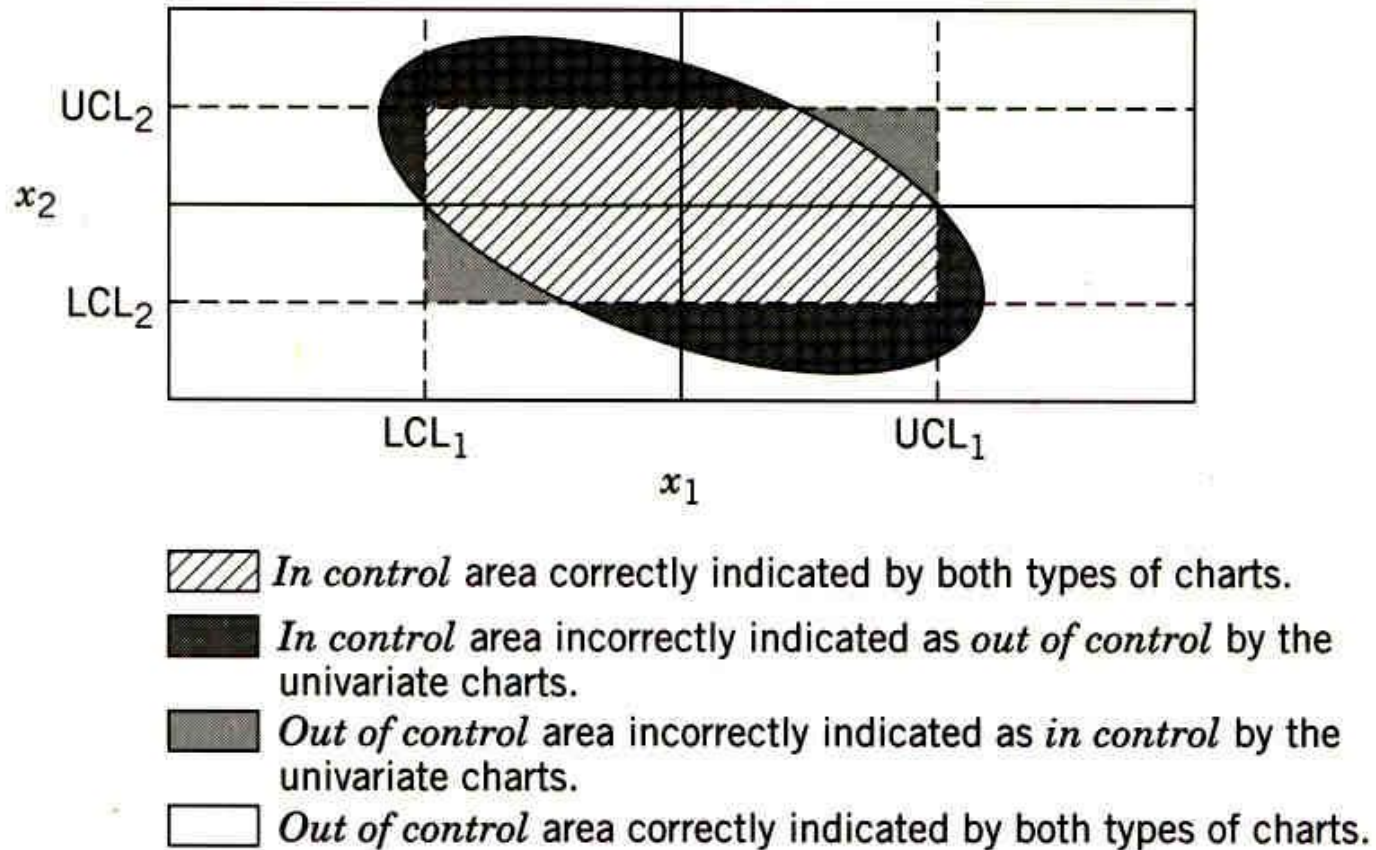*Out of control* area correctly indicated by both types of charts.

Figure 21.9 Univariate and bivariate confidence regions for two random variables, x_1 and x_2 (modified from Alt et al., 1998).

Hotelling's T^2 Statistic

- Suppose that it is desired to use SPC techniques to monitor p variables, which are correlated and normally distributed.

- Let $\boldsymbol{x}$ denote the column vector of these p variables,
$$\boldsymbol{x} = col\ [x_1,\ x_2,\ ...,\ x_p].$$

- At each sampling instant, a subgroup of n measurements is made for each variable.

- The subgroup sample means for the kth sampling instant can be expressed as a column vector:
$$\overline{x}(k) = col\left[\overline{x}_1(k), \overline{x}_2(k),\ldots,\overline{x}_p(k)\right]$$

- Multivariate control charts are traditionally based on *Hotelling's T^2 statistic,*

$$T^2(k) \triangleq n\left[\overline{x}(k) - \overline{\overline{x}}\right]^T S^{-1}\left[\overline{x}(k) - \overline{\overline{x}}\right] \qquad (21\text{-}27)$$

where $T^2(k)$ denotes the value of the T^2 statistic at the kth sampling instant.

- The vector of grand means $\bar{\bar{\mathbf{x}}}$ and the covariance matrix $\mathbf{S}$ are calculated for a test set of data for *in-control* conditions.

- By definition S_{ij}, the (i,j)-element of matrix $\mathbf{S}$, is the sample covariance of x_i and x_j:

$$S_{ij} \triangleq \frac{1}{N}\sum_{k=1}^{N}\left[\bar{x}_i(k)-\bar{\bar{x}}_i\right]^T\left[\bar{x}_j(k)-\bar{\bar{x}}_j\right] \qquad (21\text{-}28)$$

- In Eq. (21-28) N is the number of subgroups and $\bar{\bar{x}}_i$ denotes the grand mean for $\bar{x}_i$.

- Note that T^2 is a scalar, even though the other quantities in Eq. 21-27 are vectors and matrices.

- The inverse of the sample covariance matrix, $\mathbf{S}^{-1}$, scales the p variables and accounts for correlation among them.

- A multivariate process is considered to be out-of-control at the kth sampling instant if $T^2(k)$ exceeds an upper control limit, UCL.

- (There is no target or lower control limit.)

Example 21.6

Construct a T^2 control chart for the wastewater treatment problem of Example 21.5. The 99% control chart limit is $T^2 = 11.63$. Is the number of T^2 control chart violations consistent with the results of Example 21.5?

Solution

The T^2 control chart is shown in Fig. 21.10. All of the T^2 values lie below the 99% confidence limit except for sample #8. This result is consistent with the bivariate control chart in Fig. 21.8c.
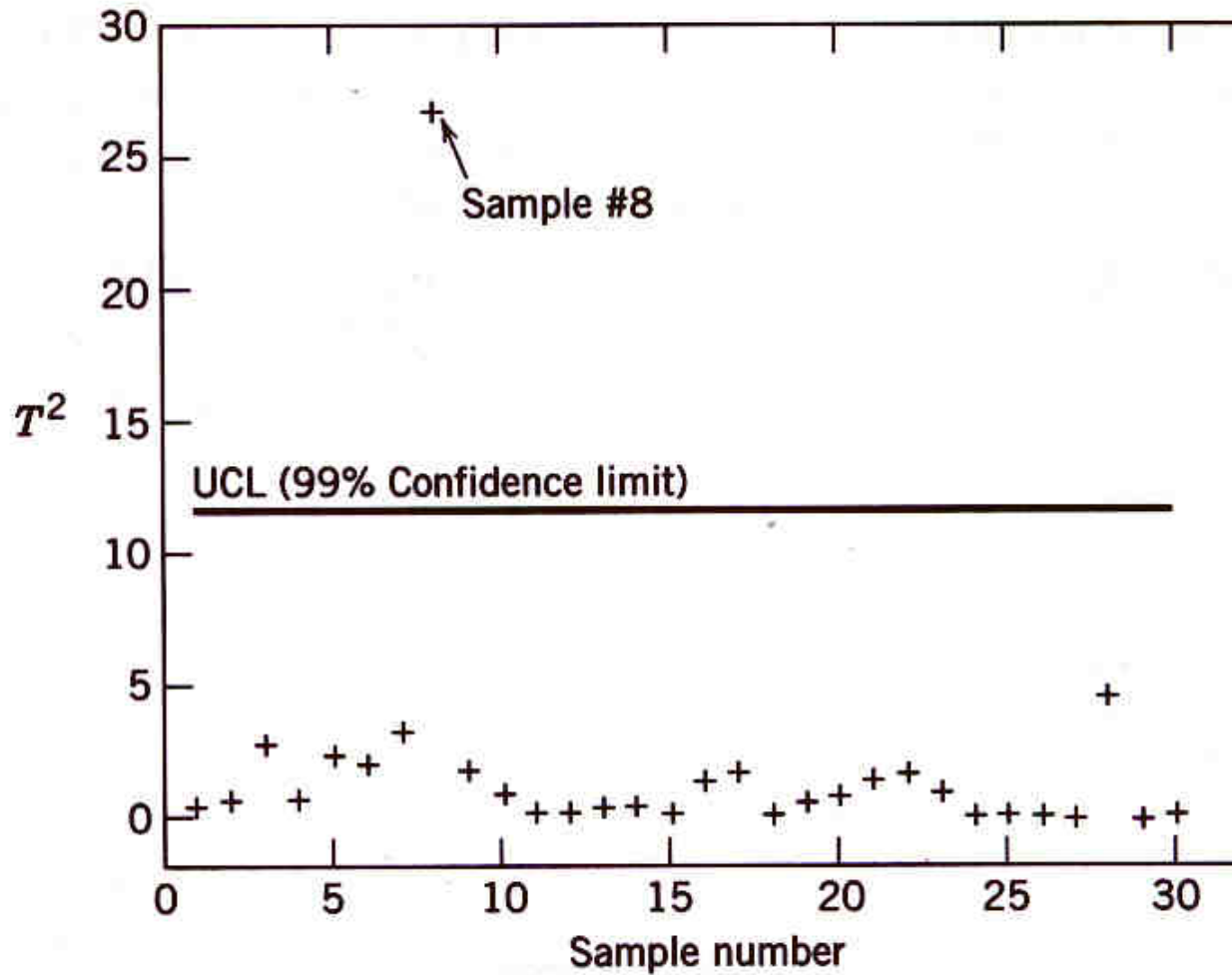
Figure 21.10 T^2 control chart for Example 21.5.

Principal Component Analysis and Partial Least Squares

- Multivariate monitoring based on Hotelling's T^2 statistic can be effective if the data are not highly correlated and the number of variables p is not large (for example, $p < 10$).

- For highly correlated data, the $\boldsymbol{S}$ matrix is poorly conditioned and the T^2 approach becomes problematic.

- Fortunately, alternative multivariate monitoring techniques have been developed that are very effective for monitoring problems with large numbers of variables and highly correlated data.

- The *Principal Component Analysis (PCA)* and *Partial Least Squares (PLS)* methods have received the most attention in the process control community.

Control Performance Monitoring

- In order to achieve the desired process operation, the control system must function properly.

- In large processing plants, each plant operator is typically responsible for 200 to 1000 loops.

- Thus, there are strong incentives for automated *control* (or *controller*) *performance monitoring* (*CPM*).

- The overall objectives of CPM are: (1) to determine whether the control system is performing in a satisfactory manner, and (2) to diagnose the cause of any unsatisfactory performance.

Basic Information for Control Performance Monitoring

- In order to monitor the performance of a single standard PI or PID control loop, the basic information in Table 21.5 should be available.

- *Service factors* should be calculated for key components of the control loop such as the sensor and final control element.

- Low service factors and/or frequent maintenance suggest chronic problems that require attention.

- The fraction of time that the controller is in the automatic mode is a key metric.

- A low value indicates that the loop is frequently in the manual mode and thus requires attention.

- Service factors for computer hardware and software should also be recorded.

- Simple statistical measures such as the sample mean and standard deviation can indicate whether the controlled variable is achieving its target and how much control effort is required.

- An unusually small standard deviation for a measurement could result from a faulty sensor with a constant output signal, as noted in Section 21.1.

- By contrast, an unusually large standard deviation could be caused by equipment degradation or even failure, for example, inadequate mixing due to a faulty vessel agitator.

- A high alarm rate can be indicative of poor control system performance.

- Operator logbooks and maintenance records are valuable sources of information, especially if this information has been captured in a computer database.

Table 21.5 Basic Data for Control Loop Monitoring

- Service factors (time in use/total time period)

- Mean and standard deviation for the control error (set point – measurement)

- Mean and standard deviation for the controller output

- Alarm summaries

- Operator logbooks and maintenance records

Control Performance Monitoring Techniques

- Chapters 6 and 12 introduced traditional control loop performance criteria such as rise time, settling time, overshoot, offset, degree of oscillation, and integral error criteria.

- CPM methods can be developed based on one or more of these criteria.

- If a process model is available, then process monitoring techniques based on monitoring the model residuals can be employed

- In recent years, a variety of statistically based CPM methods have been developed that do not require a process model.

- Control loops that are excessively oscillatory or very sluggish can be detected using correlation techniques.

- Other methods are based on calculating a standard deviation or the ratio of two standard deviations.

- Control system performance can be assessed by comparison with a *benchmark.*

- For example, historical data representing periods of satisfactory control could be used as a benchmark.

- Alternatively, the benchmark could be an ideal control system performance such as *minimum variance control.*

- As the name implies, a minimum variance controller minimizes the variance of the controlled variable when unmeasured, random disturbances occur.

- This ideal performance limit can be estimated from closed-loop operating data if the process time delay is known or can be estimated.

- The ratio of minimum variance to the actual variance is used as the measure of control system performance.

- This statistically based approach has been commercialized, and many successful industrial applications have been reported.

- For example, the Eastman Chemical Company has develop a large-scale system that assesses the performance of over 14,000 PID controllers in 40 of their plants (Paulonis and Cox, 2003).

- Although several CPM techniques are available and have been successfully applied, they also have several shortcomings.

- First, most of the existing techniques assess control system performance but do not diagnose the root cause of the poor performance.

- Thus busy plant personnel must do this "detective work".

- A second shortcoming is that most CPM methods are restricted to the analysis of individual control loops.

- The minimum variance approach has been extended to MIMO control problems, but the current formulations are complicated and are usually restricted to unconstrained control systems.

- Monitoring strategies for MPC systems are a subject of current research.