feature, the regression adjustment procedure has many possible applications in chemical and process plants where there are often cascade processes with several inputs but only a few outputs, and where many of the variables are highly autocorrelated.

10-6 CONTROL CHARTS FOR MONITORING VARIABILITY

Just as it is important to monitor the process mean vector $\boldsymbol{\mu}$ in the multivariate case, it is also important to monitor process variability. Process variability is summarized by the $p \times p$ *covariance matrix* $\boldsymbol{\Sigma}$. The main diagonal elements of this matrix are the variances of the individual process variables, and the off-diagonal elements are the covariances. Alt (1985) gives a nice introduction to the problem and presents two useful procedures.

The first procedure is a direct extension of the univariate s^2 control chart. The procedure is equivalent to repeated tests of significance of the hypothesis that the process covariance matrix is equal to a particular matrix of constants $\boldsymbol{\Sigma}$. If this approach is used, the statistic plotted on the control chart for the ith sample is

$$W_i = -pn + pn\ln(n) - n\ln\left(|\mathbf{A}_i|/|\boldsymbol{\Sigma}|\right) + \mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{A}_i\right) \qquad (10\text{-}34)$$

where $\mathbf{A}_i = (n-1)\mathbf{S}_i$, $\mathbf{S}_i$ is the sample covariance matrix for sample i, and tr is the trace operator. (The trace of a matrix is the sum of the main diagonal elements.) If the value of W_i plots above the upper control limit UCL $= \chi^2_{\alpha, p(p+1)/2}$, the process is out of control.

The second approach is based on the sample *generalized* variance, $|\mathbf{S}|$. This statistic, which is the determinant of the sample covariance matrix, is a widely used measure of multivariate dispersion. Montgomery and Wadsworth (1972) used an asymptotic normal approximation to develop a control chart for $|\mathbf{S}|$. Another method would be to use the mean and variance of $|\mathbf{S}|$—that is, $E(|\mathbf{S}|)$ and $V(|\mathbf{S}|)$—and the property that most of the probability distribution of $|\mathbf{S}|$ is contained in the interval $E|\mathbf{S}| \pm 3\sqrt{V(|\mathbf{S}|)}$. It can be shown that

$$E(|\mathbf{S}|) = b_1|\boldsymbol{\Sigma}| \qquad (10\text{-}35)$$

and

$$V(|\mathbf{S}|) = b_2|\boldsymbol{\Sigma}|^2$$

where

$$b_1 = \frac{1}{(n-1)^p}\prod_{i=1}^{p}(n-i)$$

and

$$b_2 = \frac{1}{(n-1)^{2p}} \prod_{i=1}^{p}(n-i)\left[\prod_{j=1}^{p}(n-j+2) - \prod_{j=1}^{p}(n-j)\right]$$

Therefore, the parameters of the control chart for $|\mathbf{S}|$ would be

$$\begin{aligned} \text{UCL} &= |\boldsymbol{\Sigma}|\left(b_1 + 3b_2^{1/2}\right) \\ \text{CL} &= b_1|\boldsymbol{\Sigma}| \\ \text{LCL} &= |\boldsymbol{\Sigma}|\left(b_1 - 3b_2^{1/2}\right) \end{aligned} \tag{10-36}$$

The lower control limit in equation 10-36 is replaced with zero if the calculated value is less than zero.

Usually, in practice $\boldsymbol{\Sigma}$ will be estimated by a sample covariance matrix **S**, based on the analysis of preliminary samples. If this is the case, we should replace $|\boldsymbol{\Sigma}|$ in equation 10-36 by $|\mathbf{S}|/b_1$, since equation 10-35 has shown that $|\mathbf{S}|/b_1$ is an unbiased estimator of $|\boldsymbol{\Sigma}|$.

EXAMPLE 10-2

To illustrate controlling process variability in the multivariate case, we will return to Example 10-1 and construct a control chart for the generalized variance. Based on the 20 preliminary samples in Table 10-1, the sample covariance matrix is

$$\mathbf{S} = \begin{bmatrix} 1.23 & 0.79 \\ 0.79 & 0.83 \end{bmatrix}$$

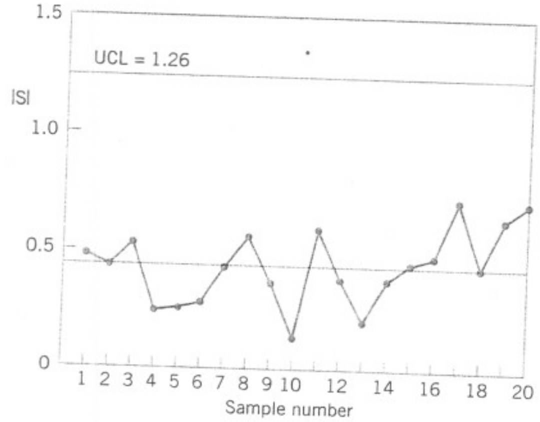

Figure 10-13 A control chart for the sample generalized variance, Example 10-2.

so

$$|\mathbf{S}| = 0.3968$$

The constants b_1 and b_2 are (recall that $n = 10$)

$$b_1 = \frac{1}{81}(9)(8) = 0.8889$$

$$b_2 = \frac{1}{6561}(9)(8)[(11)(10) - (9)(8)] = 0.4170$$

Therefore, replacing $|\boldsymbol{\Sigma}|$ in equation 10-36 by $|\mathbf{S}|/b_1 = 0.3968/0.8889 = 0.4464$, we find that the control chart parameters are

$$\text{UCL} = (|\mathbf{S}|/b_1)\left(b_1 + 3b_2^{1/2}\right) = 0.4464\left[0.8889 + 3(0.4170)^{1/2}\right] = 1.26$$

$$\text{CL} = |\mathbf{S}| = 0.3968$$

$$\text{LCL} = (|\mathbf{S}|/b_1)\left(b_1 - 3b_2^{1/2}\right) = 0.4464\left[0.8889 - 3(0.4170)^{1/2}\right] = -0.47 = 0$$

Figure 10-13 presents the control chart. The values of $|\mathbf{S}_i|$ for each sample are shown in the last column of panel (c) of Table 10-1.

Although the sample generalized variance is a widely used measure of multivariate dispersion, remember that it is a relatively simplistic scalar representation of a complex multivariable problem, and it is easy to be fooled if all we look at is $|\mathbf{S}|$. For example, consider the three covariance matrices:

$$\mathbf{S}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{S}_2 = \begin{bmatrix} 2.32 & 0.40 \\ 0.40 & 0.50 \end{bmatrix}, \text{ and } \mathbf{S}_3 = \begin{bmatrix} 1.68 & -0.40 \\ -0.40 & 0.50 \end{bmatrix}$$

Now $|\mathbf{S}_1| = |\mathbf{S}_2| = |\mathbf{S}_3| = 1$, yet the three matrices convey considerably different information about process variability and the correlation between the two variables. It is probably a good idea to use univariate control charts for variability in conjunction with the control chart for $|\mathbf{S}|$.

10-7 LATENT STRUCTURE METHODS

Conventional multivariate control-charting procedures are reasonably effective as long as p (the number of process variables to be monitored) is not very large. However, as p increases, the average run-length performance to detect a specified shift in the mean of these variables for multivariate control charts also increases, because the shift is "diluted" in the p-dimensional space of the process variables. To illustrate this, consider the ARLs of the MEWMA control chart in Table 10-3. Suppose we choose $\lambda = 0.1$ and the magnitude