Journal of Quality Technology

A Quarterly Journal of Methods, Applications and Related Topics

ISSN: 0022-4065 (Print) 2575-6230 (Online) Journal homepage: https://www.tandfonline.com/loi/ujqt20

Identification and Quantification in Multivariate Quality Control Problems

Anthony J. Hayter & Kwok-Leung Tsui

To cite this article: Anthony J. Hayter & Kwok-Leung Tsui (1994) Identification and Quantification in Multivariate Quality Control Problems, Journal of Quality Technology, 26:3, 197-208, DOI: 10.1080/00224065.1994.11979526

To link to this article: https://doi.org/10.1080/00224065.1994.11979526

Published online: 21 Feb 2018.

Identification and Quantification in Multivariate Quality Control Problems

ANTHONY J. HAYTER and KWOK-LEUNG TSUI

Georgia Institute of Technology, Atlanta, GA 30332-0205

Many quality control problems are multivariate in character since the quality of a given product or object consists simultaneously of more than one variable. A good multivariate quality control procedure should possess three important properties, namely, the control of the overall error rate, the easy identification of errant variables, and the easy quantification of any changes in the variable means. In this paper a procedure is suggested based on the construction of exact simultaneous confidence intervals for each of the variable means that meets each of these three goals. Both parametric and nonparametric procedures are considered, and critical point evaluation through tables, numerical integration, and simulation is discussed. Various examples of the implementation of the procedure are given.

Introduction

In many quality control settings the product under examination may have two or more related quality characteristics, and the objective of the supervision is to investigate whether all of these characteristics are simultaneously behaving appropriately. In particular, a standard multivariate quality control problem is to consider whether an observed vector of measurements $\mathbf{x} = (x_1, \ldots, x_k)'$ from a particular sample exhibits any evidence of a location shift from a set of "satisfactory" or "standard" mean values $\boldsymbol{\mu}^0 = (\mu_1^0, \ldots, \mu_k^0)'$. The individual measurements will usually be correlated due to the nature of the problem so that their covariance matrix $\boldsymbol{\Sigma}$ will not be diagonal. In practice, the mean vector $\boldsymbol{\mu}^0$ and covariance matrix $\boldsymbol{\Sigma}$ may be estimated from an initial large pool of observations, and the problem is then to monitor further observations $\mathbf{x}$ in order to identify location shifts in any of the mean values.

If the assumption is made that the data are normally distributed, then the $\mathbf{x}$'s come from a distribution that is $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, and the problem is to assess the evidence that $\boldsymbol{\mu} \neq \boldsymbol{\mu}^0$. In the univariate setting ($k = 1$) this problem can be handled with a Shewhart control chart with control limits set to guarantee a specified error rate α. One might consider handling the multivariate problem by constructing individual α-level control charts for each of the k variables under consideration. However, it has long been realized that such an approach is unsatisfactory since it ignores the correlation between the variables and allows the overall error rate to be much larger than α. On the other hand, if individual error rates of α/k are used, then the Bonferroni inequality ensures that the overall error rate is less than the nominal level α. However, this procedure is not sensitive enough since the actual overall error rate tends to be much smaller than α because of the correlation between the variables.

It is clear that a basic property of a good procedure for this multivariate problem is that an overall error rate of a specified level α should be approximately maintained, so that the probability of *incorrectly* deciding that the process is out of control when it is actually in control is approximately α. Hotelling (1947) provided the first solution to this problem by suggesting the use of the statistic

$$T^2 = (\mathbf{x} - \boldsymbol{\mu}^0)'\widehat{\boldsymbol{\Sigma}}^{-1}(\mathbf{x} - \boldsymbol{\mu}^0)$$

where $\widehat{\boldsymbol{\Sigma}}$ is an estimate of the population covariance matrix $\boldsymbol{\Sigma}$. When the population covariance matrix is known, Hotelling's statistic is equivalent to the χ^2 statistic

$$\chi^2 = (\mathbf{x} - \boldsymbol{\mu}^0)'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}^0). \qquad (1)$$

Dr. Hayter is an Associate Professor in the School of Industrial and Systems Engineering.

Dr. Tsui is an Associate Professor in the School of Industrial and Systems Engineering.

When $\mu = \mu^0$, there is probability α that this statistic exceeds a critical point of $\chi^2_{k,\alpha}$, so that the overall error rate can be maintained exactly at the level α by triggering a warning only when

$$\chi^2 > \chi^2_{k,\alpha}.$$

It is common practice to suggest the use of a Shewhart type χ^2 control chart with an upper control limit of $\chi^2_{k,\alpha}$ (see, e.g., Alt (1985) and Montgomery (1991, Sec. 8-4.1)).

However, there is still the problem of deciding what conclusions can be drawn once $\chi^2 > \chi^2_{k,\alpha}$ and the alarm has been sounded. More specifically, once the experimenter has evidence via the χ^2 statistic that the process is no longer in control, how is it determined which location parameters have moved away from their control values μ_i^0? No completely satisfactory solution to this identification problem has been suggested, nor to the related problem of estimating the magnitudes of any differences in the location parameters from their standard values μ_i^0.

The earliest work on this identification problem suggested the use of principal components analysis to decompose the χ^2 statistic into various independent components so that the most "influential" linear combinations of the variables could be identified (see, e.g., Jackson and Mudholkar (1979) and Jackson (1985)). Although the identified linear combination of variables can sometimes have physical interpretations (see examples in Jackson (1991)), they are not always meaningful to the experimenter. For example, to be told that there is no evidence that either the weight or the length of a certain product has changed, but that the quantity "(1/4) weight + (3/4) length" is clearly not what it was before, is enough to cool the ardor of the most devoted quality control practitioner and is not likely to foster much enthusiasm for the further adoption and implementation of quality control techniques.

An alternative identification method proposed by Murphy (1987) is to try to identify which subset of the k variables contributes a "substantial" amount to the χ^2 statistic. While this is a sensible idea, it is necessarily complicated to implement, and its sensitivity and power properties are not clear. Alt (1985) and Doganaksoy et al. (1991) describe what is probably the simplest and best identification method proposed to date, which is to calculate individual p-values for each of the k variables under consideration and to compare them with the Bonferroni error rates α/k. Essentially, this method is equivalent to using k individual univariate Shewhart control charts, each with an error rate of α/k, but with an overall trigger based on the χ^2 statistic. The conservative nature of the Bonferroni error rates, which ignores the known correlations between the variables, implies that this approach is not very powerful or sensitive. Often $\chi^2 > \chi^2_{k,\alpha}$ so there is an indication that the process is out of control, but no variable is able to be identified as the culprit since no p-value is smaller than α/k. In an attempt to avoid this problem Doganaksoy et al. (1991) suggest an ad hoc choice of a larger error rate α_1 in the identification stage. It is also worth pointing out that because of the assumption of normality, the Dunn-Sidak inequality (see, e.g., Dunn (1958), Sidak (1967), or Hochberg and Tamhane (1987)) can be used to justify the error rate $1-(1-\alpha)^{1/k}$ in place of the Bonferroni error rate α/k. Since $1-(1-\alpha)^{1/k} > \alpha/k$, this results in a more powerful procedure, although in practice the difference is often negligible.

A good solution to this multivariate quality control problem is one that both controls the overall error rate at a specified level α by triggering the out-of-control alarm only with probability α when the process is still in control, and also provides a simple and easily implementable mechanism for deciding which variables are responsible when the process is determined to be out of control. In addition, it is desirable to be able to quantify how much the variables may have changed in mean value. In this paper a simple procedure is suggested that meets all three of these criteria. The procedure operates by calculating a set of simultaneous confidence intervals for the variable means μ_i with an exact simultaneous coverage probability of $1-\alpha$. The process is deemed to be out of control whenever any of these confidence intervals does not contain its respective control value μ_i^0, and the identification of the errant variable or variables is immediate. Furthermore, this procedure continually provides confidence intervals for the "current" mean values μ_i, regardless of whether the process is in control or not or whether a particular variable is in control or not.

It is also important to note that this simple procedure possesses the three important properties of a multivariate technique as stated in Jackson (1985). These are (1) that the procedure gives a single answer to the question "Is the process in control?", (2) the specified type I error probability is exactly maintained, and (3) the procedure takes into account the relationships among the variables (given by Σ) in the determination of the exact critical value.

A graphical control chart display is available for this new procedure. The experimenter would use a multivariate chart to monitor the multivariate observations and maintain k individual control charts for each of the k variables. In practice, the experimenter only needs to watch the multivariate chart for an out-of-control signal. Once there is such a signal, then the experimenter can immediately examine the individual control charts to discover which variables are responsible for the out-of-control signal. More detail of this graphical procedure is given in the next section.

The outline of this paper is as follows. The new procedure is discussed in the next section. In the following section, the calculation of the critical points required for implementation is discussed. Both nonparametric and parametric approaches to the problem of critical point evaluation are considered, and options such as tables, numerical integration, and simulation are discussed. Various examples of the implementation of the procedure are also given. The paper concludes with a summary.

Description of the Procedure

Let $\mathbf{X} \sim N_k(\mathbf{0}, \mathbf{R})$, where $\mathbf{R}$ is a general correlation matrix with diagonal elements equal to one and off-diagonal elements given by ρ_{ij}, and define the critical point $C_{\mathbf{R},\alpha}$ by

$$\Pr\left(|X_i| \leq C_{\mathbf{R},\alpha} \text{ for } 1 \leq i \leq k\right) = 1 - \alpha. \quad (2)$$

In the more general case when $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for any general covariance matrix $\boldsymbol{\Sigma}$, let the diagonal elements of $\boldsymbol{\Sigma}$ be given by σ_i^2 and the off-diagonal elements by σ_{ij}. Then if $\mathbf{R}$ is the correlation matrix generated from $\boldsymbol{\Sigma}$, $\rho_{ij} = \sigma_{ij}/\sigma_i\sigma_j$, and it follows from equation (2) that

$$\Pr\left(|X_i - \mu_i|/\sigma_i \leq C_{\mathbf{R},\alpha} \text{ for } 1 \leq i \leq k\right) = 1-\alpha. \quad (3)$$

However, equation (3) can be rewritten as

$$\Pr\left(\mu_i \in [X_i - \sigma_i C_{\mathbf{R},\alpha}, X_i + \sigma_i C_{\mathbf{R},\alpha}] \text{ for } 1 \leq i \leq k\right) = 1 - \alpha \quad (4)$$

which means that the probability that the random interval

$$[X_i - \sigma_i C_{\mathbf{R},\alpha}, X_i + \sigma_i C_{\mathbf{R},\alpha}]$$

contains μ_i for $1 \leq i \leq k$ is equal to $1-\alpha$. Notice that the choice of $C_{\mathbf{R},\alpha}$ depends on the correlation matrix $\mathbf{R}$ since the X_i's are correlated. Therefore, the correlation structure among the random variables $\mathbf{X}$ affects the simultaneous confidence intervals in equation (4) through the critical point $C_{\mathbf{R},\alpha}$.

The multivariate quality control procedure operates as follows. For a known covariance structure $\boldsymbol{\Sigma}$ and a chosen error rate α, the experimenter first evaluates the critical point $C_{\mathbf{R},\alpha}$. Then, following any observation $\mathbf{x} = (x_1, \ldots, x_k)'$, the experimenter constructs confidence intervals

$$[x_i - \sigma_i C_{\mathbf{R},\alpha}, x_i + \sigma_i C_{\mathbf{R},\alpha}] \quad (5)$$

for each of the k variables. The process is considered to be in control as long as *each* of these confidence intervals contains the respective standard value μ_i^0. However, when an observation $\mathbf{x}$ is obtained for which one or more of the confidence intervals do not contain their respective standard values μ_i^0, then the process is stated to be out of control, and the variable or variables whose confidence intervals do not contain μ_i^0 are identified as those responsible for the aberrant behavior.

This simple procedure clearly meets the three goals set in the Introduction for a good solution to the multivariate quality control problem. Equation (4) ensures that an overall error rate of α is achieved, since when $\boldsymbol{\mu} = \boldsymbol{\mu}^0$, there is a probability of $1 - \alpha$ that each of the confidence intervals contains the respective value μ_i^0. Also, the identification of the errant variables is immediate and simple, and furthermore, the confidence intervals allow the experimenter to assess the new mean values of the out-of-control variables. This is particularly useful when the experimenter can judge the process to be still "good enough" and hence allow it to continue.

It is interesting to compare this new procedure with the standard procedure, which triggers an alarm when

$$\chi^2 = (\mathbf{x} - \boldsymbol{\mu}^0)'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}^0) > \chi^2_{k,\alpha}. \quad (6)$$

The new procedure can be thought of as triggering an alarm when

$$M = \max_{1 \leq i \leq k} |x_i - \mu_i^0|/\sigma_i > C_{\mathbf{R},\alpha}. \quad (7)$$

Both of these statistics control the overall error rate at exactly α, and usually they will reach the same conclusion, but it is possible for one to trigger and the other not to trigger. The question of which statistic is "better" is essentially one concerning the relative powers and sensitivities of the two statistics. However, in this multivariate setting neither one is uniformly more powerful than the other. Although one may be better than the other when the variable means $\boldsymbol{\mu}$ differ from $\boldsymbol{\mu}^0$ in a particular manner, and vice versa elsewhere (see Woodall (1986) for a similar discussion of this point).

For example, consider the two-dimensional case with population means equal to zero, variances equal to one, and correlation equal to 0.6. Figure 1 illustrates the critical regions of the two statistics at $\alpha = 0.05$. The region outside the rotated ellipsoid is the critical region for the χ^2 statistic, and the region outside the rectangle is the critical region for the M statistic. Clearly, for any observations inside the rectangle but outside the ellipsoid (regions A in Figure 1), the χ^2 statistic would trigger the alarm but the M statistic would not. Thus, the χ^2 statistic would be more powerful when the mean shifts to somewhere around this region. Similarly, for any observations outside the rectangle but inside the ellipsoid (shaded regions), the χ^2 statistic would not trigger the alarm but the M statistic would. Thus,

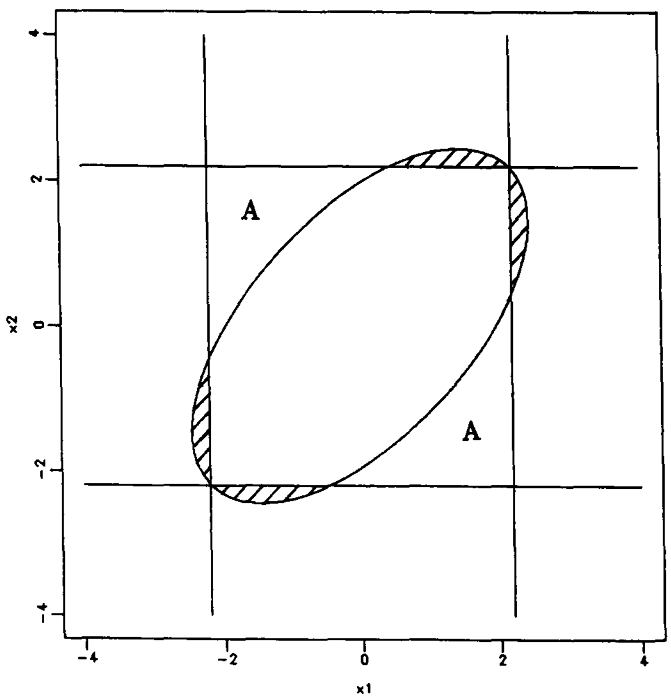

FIGURE 1. Critical Regions for the χ^2 and M Statistics at $\rho = 0.06$.

the χ^2 statistic would be less powerful when the mean shifts to somewhere around this region. In general, Doganaksoy et al. (1991, p. 2778) commented that "a counter correlational mean shift of given size is more readily detectable (by χ^2 statistic) than a shift of equal magnitude in a direction closer to the principal axis of the correlation structure". Alt (1985) provided additional examples comparing the two statistics, and Alt et al. (1980) investigated the power of the χ^2 statistic.

Since there is no real rationale per se for deciding between one statistic and the other, the motivation behind this paper is that the second stage identification problem should be paramount in influencing the choice. In other words, the motivation behind the use of the statistic (7) is that it is associated with the simultaneous confidence intervals given in equation (4). These intervals immediately and simply solve the identification and quantification problem, and thereby actually remove the need for an independent trigger mechanism. Historically, a reason for preferring statistic (6) over statistic (7) is that the required critical point $\chi^2_{k,\alpha}$ is independent of the correlation structure and is readily available. However, modern computing power allows the calculation of the critical point $C_{\mathbf{R},\alpha}$ one way or another, as discussed in the next section.

Finally, it is interesting to notice that the statistic (6) is associated with the simultaneous confidence intervals (see Hochberg and Tamhane (1987, p. 35)) obtained from

$$\Pr\left(\mathbf{l}'\boldsymbol{\mu} \in \mathbf{l}'\mathbf{X} \pm \sqrt{\chi^2_{k,\alpha}}\sqrt{\mathbf{l}'\boldsymbol{\Sigma}\mathbf{l}} \ \forall \ \mathbf{l} \in R^k\right) = 1 - \alpha. \tag{8}$$

In this equation, $\mathbf{l} = (l_1, ..., l_k)'$ is any k-dimensional vector of real numbers, so that simultaneous intervals are given for any linear combination

$$\mathbf{l}'\boldsymbol{\mu} = l_1\mu_1 + \cdots + l_k\mu_k$$

of the means. Notice also that $\mathbf{l}'\boldsymbol{\Sigma}\mathbf{l}$ is just the variance of $\mathbf{l}'\mathbf{X}$.

When $\chi^2 \leq \chi^2_{k,\alpha}$, then each of the confidence intervals obtained from equation (8) will contain the respective value $\mathbf{l}'\boldsymbol{\mu}^0$, whereas when $\chi^2 > \chi^2_{k,\alpha}$ there will be at least one vector $\mathbf{l} \in R^k$ for which the confidence interval for $\mathbf{l}'\boldsymbol{\mu}$ excludes $\mathbf{l}'\boldsymbol{\mu}^0$. It is not generally realized that the standard procedure based on the statistic (6) has these simultaneous confidence intervals associated with it. However, as was pointed out in the Introduction, the identification of a linear combination of the variables as being aberrant is not always useful to the experimenter.

Note that in reality the covariance matrix $\mathbf{\Sigma}$ may be unknown. A common practice is to estimate $\mathbf{\Sigma}$ from a large pool of past observations, which are often available when monitoring manufacturing processes. When the sample covariance matrix is used instead of $\mathbf{\Sigma}$, the χ^2 statistic becomes a T^2 statistic as mentioned in the Introduction, and the critical values of this statistic can be obtained from tables of the F distribution (see Alt (1985) or Doganaksoy et al. (1991) for details).

For a graphical control chart display of the new procedure, the experimenter could plot the M statistic for each multivariate observation and use the critical value $C_{\mathbf{R},\alpha}$ as the upper control limit. In addition, individual control charts for each of the k variables could be maintained with $\mu_i^0 \pm \sigma_i C_{\mathbf{R},\alpha}$ as the two-sided control limits. In practice, the experimenter only needs to watch the M statistic control chart for out-of-control signals. Once the M statistic goes above its control limit, then the experimenter can immediately examine the individual control charts to discover which variables are responsible for the out-of-control signal.

In summary, focusing on the identification and quantification aspects of the multivariate quality control problem leads to the simultaneous confidence intervals given in equation (4) as being the best solution to the problem. A discussion of the evaluation of the critical points $C_{\mathbf{R},\alpha}$ and some examples of the implementation of the procedure are given in the next section.

Critical Point Evaluation and Examples

In this section various methods of critical point evaluation are discussed together with some illustrative examples of the implementation of the procedure. When $k = 2$, the required critical points can be found from existing tables. For $k \geq 3$, critical point evaluation by numerical integration or by approximation from existing tables may be possible. Otherwise, a simple and quick simulation method is suggested. Finally, in those circumstances where a large pool of in-control observations is available to the experimenter, a nonparametric implementation of the procedure is suggested, which provides a solution to the problem without requiring the assumption of a normal distribution.

Two Variables

Consider first the basic multivariate quality control problem with $k = 2$ so that there are just two variables under consideration. In this case the required critical point $C_{\mathbf{R},\alpha}$ depends only on the error size α and the one correlation term $\rho_{12} = \rho$, say. In tables B.1-B.4 of Bechhofer and Dunnett (1988) values of the critical points are given for $\alpha = 0.20$, 0.10, 0.05, and 0.01, and for $\rho = 0(0.1)0.9$ (the required values for $C_{\mathbf{R},\alpha}$ correspond to the entries for $p = 2$ and $\nu = \infty$). More complete tables are given by Odeh (1982) who tabulated the required critical points for additional values of α and ρ (the values $C_{\mathbf{R},\alpha}$ for $k = 2$ correspond to the entries for $N = 2$). Interpolation within these tables can be used to provide critical values for other cases not given. An alternative method would be to use subroutines for evaluating the bivariate normal cumulative distribution functions, such as IMSL (1987).

As an example of the implementation of the procedure with $k = 2$, consider the problem outlined in Alt (1985) where a lumber manufacturing plant obtains measurements on both the *stiffness* and the *bending strength* of a particular grade of lumber. Samples of size 10 are averaged to produce an observation $\mathbf{x} = (x_1, x_2)'$, and standard values for these averaged observations are taken to be $\boldsymbol{\mu}^0 = (265, 470)'$ with a covariance matrix of

$$\mathbf{\Sigma} = \begin{bmatrix} 10 & 6.6 \\ 6.6 & 12.1 \end{bmatrix}.$$

In this case the correlation is $\rho = 0.6$, so that with an error rate of $\alpha = 0.05$, the tables referenced above give the critical point as $C_{\mathbf{R},\alpha} = 2.199$.

Following an observation $\mathbf{x} = (x_1, x_2)'$, the simultaneous confidence intervals for the current mean values $\boldsymbol{\mu} = (\mu_1, \mu_2)'$ are given by

$$\begin{aligned} \mu_1 &\in \left[x_1 - 2.199\sqrt{10}, x_1 + 2.199\sqrt{10}\right] \\ &= [x_1 - 6.95, x_1 + 6.95] \\ \mu_2 &\in \left[x_2 - 2.199\sqrt{12.1}, x_2 + 2.199\sqrt{12.1}\right] \\ &= [x_2 - 7.65, x_2 + 7.65]. \end{aligned} \tag{9}$$

These confidence intervals have a *joint* confidence level of 0.95. The process is considered to be in control as long as both of these confidence intervals contain their respective control values $\boldsymbol{\mu}^0 = (265, 470)'$, that is, as long as $258.05 \leq x_1 \leq 271.95$ and $462.35 \leq x_2 \leq 477.65$. However, following an observation $\mathbf{x} = (255, 465)'$ the process would be declared to be out of control, and the first variable stiffness would be identified as the culprit. Furthermore, the confidence interval for the mean stiffness level would be $\mu_1 \in (248.05, 261.95)$ so that the experimenter has an immediate quantification of the amount of change in the mean stiffness level.

It is instructive to compare the exact procedure above with other approximate procedures based on the Bonferroni and Dunn-Sidak inequalities. These approximate procedures provide the simultaneous confidence intervals

$$\mu_1 \in \left[x_1 - z_{\alpha^*}\sqrt{10}, x_1 + z_{\alpha^*}\sqrt{10}\right]$$
$$\mu_2 \in \left[x_2 - z_{\alpha^*}\sqrt{12.1}, x_2 + z_{\alpha^*}\sqrt{12.1}\right]$$

where z_{α^*} is the upper α^* quantile of the standard normal distribution, and $\alpha^* = \alpha/4$ for the Bonferroni procedure and $\alpha^* = (1 - \sqrt{1-\alpha})/2$ for the Dunn-Sidak procedure. For $\alpha = 0.05$, the values of z_{α^*} are 2.241 and 2.236, respectively. It is always the case that the critical point $C_{\mathbf{R},\alpha}$ (which is exact) will be smaller than the Dunn-Sidak critical point, which in turn will be smaller than the Bonferroni critical point. However, as this example shows, the differences may be quite small. The difference between the exact procedure and the approximate Dunn-Sidak procedure becomes more pronounced as either the error rate α or the correlation ρ increases. For example, when $\alpha = 0.10$ and $\rho = 0.6$, the exact critical point is $C_{\mathbf{R},\alpha} = 1.900$, whereas the Dunn-Sidak critical point is $z_{\alpha^*} = 1.949$. On the other hand, when $\alpha = 0.05$ and $\rho = 0.9$, the critical points are $C_{\mathbf{R},\alpha} = 2.108$ and $z_{\alpha^*} = 2.236$.

In summary, two points should be noted. First, when $k = 2$ the exact critical points $C_{\mathbf{R},\alpha}$ may be very close to the approximate Bonferroni and Dunn-Sidak critical points. Second, since the exact critical points $C_{\mathbf{R},\alpha}$ are well tabulated they should always be used, and they provide an exact solution to the problem through the construction of simultaneous confidence intervals for the variable mean levels, such as those given in equation (9).

Finally, a comparison can be made with the procedure that uses the χ^2 statistic. For the example discussed above the χ^2 statistic (1) is given by

$$\chi^2 = \frac{605(x_1 - 265)^2}{3872} + \frac{125(x_2 - 470)^2}{968} - \frac{165(x_1 - 265)(x_2 - 470)}{968}$$

and the $\alpha = 0.05$ level critical point is $\chi^2_{2,0.05} = 5.992$. As shown below, the identification problem may become unclear if the χ^2 statistic is used.

Consider an observation $\mathbf{x} = (269, 466)'$ for which $\chi^2 = 7.293$ so that the process is determined to be out of control. Now for this observation the confidence intervals (9) are

$$\mu_1 \in [262.05, 275.95]$$
$$\mu_2 \in [458.35, 473.65]$$

so that, in reality, the control values $\boldsymbol{\mu}^0 = (265, 470)'$ are still plausible. If the χ^2 statistic is used to trigger the alarm, then the identification problem is actually a difficult one since it is incorrect to conclude that either stiffness or bending strength has changed from its control value. The reason why the χ^2 statistic has triggered the alarm is that there is a linear combination of stiffness and bending strength which seems to have changed. Note that (x_1, x_2) is an observation from a multivariate normal distribution with correlation 0.6. If there is no shift in the mean, the probability of having x_1 far above μ_1^0 and x_2 far below μ_2^0 is quite small. In such a situation, the quality control supervisor should be alert when measuring the next batch of observations.

Alt (1985) simulated some out-of-control observations $\mathbf{x}$ when μ_1 was *increased* by 5.0 and μ_2 was *decreased* by 5.5. These values are presented in the first five rows of Table 1 together with the corresponding χ^2 and M statistic values where

$$M = \max\left\{\frac{|\, x_1 - 265 \,|}{\sqrt{10}}, \frac{|\, x_2 - 470 \,|}{\sqrt{12.1}}\right\}.$$

With an error rate of $\alpha = 0.005$, the critical value for the χ^2 statistic is $\chi^2_{2,0.005} = 10.60$, and the critical value for the M statistic is $C_{\mathbf{R},0.005} = 3.01$, so it can be seen that the χ^2 statistic triggers an alarm for three out of the five observations, but the M statistic triggers an alarm for only one of the observations. Alt observed the same phenomenon when using Bonferroni critical points.

However, these simulation results should not be taken as evidence that the χ^2 statistic is generally

TABLE 1. Simulated Observations
(μ_1^0 =265, μ_2^0 =470)

(μ_1, μ_2)	(x_1, x_2)	χ^2	M
(270.0, 464.5)	(270.0, 465.2)	10.91*	1.58
(270.0, 464.5)	(268.2, 468.5)	2.76	1.02
(270.0, 464.5)	(272.9, 467.6)	13.54*	2.48
(270.0, 464.5)	(269.9, 466.2)	8.93	1.56
(270.0, 464.5)	(278.8, 474.2)	22.17*	4.36*
(270.0, 475.5)	(274.8, 474.9)	9.98	3.11*
(270.0, 475.5)	(275.5, 472.0)	14.16*	3.32*
(270.0, 475.5)	(264.6, 470.6)	0.11	0.17
(270.0, 475.5)	(274.3, 481.8)	12.77*	3.39*
(270.0, 475.5)	(269.8, 474.0)	2.44	1.53

* denotes alarm triggered at $\alpha = 0.005$

more sensitive or more powerful than the M statistic. For example, the results can be reversed by considering a different shift in the mean values. The last five rows in Table 1 represent simulated observations when μ_1 was *increased* by 5.0 and μ_2 was *increased* by 5.5. In this case three out of the five observations trigger an alarm using the M statistic, while only two trigger an alarm using the χ^2 statistic. These simple simulation results are reflections of the theoretical fact that neither statistic is uniformly more powerful than the other, and they should be considered equally sensitive to detecting an out-of-control process unless the experimenter wishes to detect a specific departure from the control status. Additional comparisons between the χ^2 statistic and an M-type statistic can be found in Bozzelo (1989).

Three or More Variables

For $k \geq 3$ the critical points $C_{\mathbf{R},\alpha}$ are not available from tables except in the special case when the correlations ρ_{ij} are all equal (see Bechhofer and Dunnett (1988) and Odeh (1982)). For a given size α and correlation matrix $\mathbf{R}$, numerical integration techniques can be used to find $C_{\mathbf{R},\alpha}$ for $k = 3$ and 4. For $k \geq 5$, these techniques will in general not be feasible due to the dimensionality of the integration region.

In general, a practical solution to the problem of critical point evaluation is the use of simulation. (Law and Kelton (1982, p. 505) provide information on the simulation of multivariate normal random variables.) In particular, the critical point $C_{\mathbf{R},\alpha}$ can be estimated as follows.

1. Generate a large number (N) of vector realizations from a multivariate normal distribution with zero means and covariance matrix $\mathbf{R}$, say $\mathbf{X}^1, \ldots, \mathbf{X}^N$.
2. Compute the M statistics for each of these vectors, $\mathbf{X}^i = (X_1^i, \ldots, X_k^i)$. That is, compute
$$M^i = \max_{1 \leq j \leq k} \mid X_j^i \mid$$
for $i = 1, \ldots, N$.
3. Find the $(1 - \alpha)^{\text{th}}$ percentile of the sample $\{M^1, \ldots, M^N\}$ and use it as an estimate of the critical point $C_{\mathbf{R},\alpha}$.

A simple computer program suffices to perform these simulations, and a copy may be requested from the authors. In practice, it is sensible to plot a graph of the empirical cumulative distribution function $F(t)$ = (number of statistics $M \leq t)/N$ since this will provide critical points $C_{\mathbf{R},\alpha}$ at all size levels α, and can be used to calculate the p-values of the process being in control for particular observations.

Our experience suggests that $N = 100,000$ simulations is a sufficient number to obtain accurate estimates of the critical points, and only a couple of minutes is needed between the input of a correlation matrix $\mathbf{R}$ and the output of the empirical cumulative distribution function. For example, for the two-dimensional example discussed above where the exact critical point $C_{\mathbf{R},0.05} = 2.199$ can be obtained from tables, 100,000 simulations resulted in an estimate of 2.198. Additional checks between tabulated critical points for $k = 2$ and simulated critical points based on 100,000 simulations indicated negligible differences.

As an example of the implementation of the procedure with $k = 4$ and the use of this simulation method, consider the example discussed in Doganaksoy et al. (1991, sec. 6), which is taken from Jackson (1980), concerning the testing of ballistic missiles. The data are subtracted from their means so that $\boldsymbol{\mu}^0 = (0, 0, 0, 0)$, and the covariance matrix (which is estimated from a pool of 40 observations) is

$$\boldsymbol{\Sigma} = \begin{bmatrix} 102.74 & 88.67 & 67.04 & 54.06 \\ 88.67 & 142.74 & 86.56 & 80.03 \\ 67.04 & 86.56 & 84.57 & 69.42 \\ 54.06 & 80.03 & 69.42 & 99.06 \end{bmatrix}.$$

For the related correlation matrix $\mathbf{R}$, the empirical cumulative distribution function of 100,000 sim-

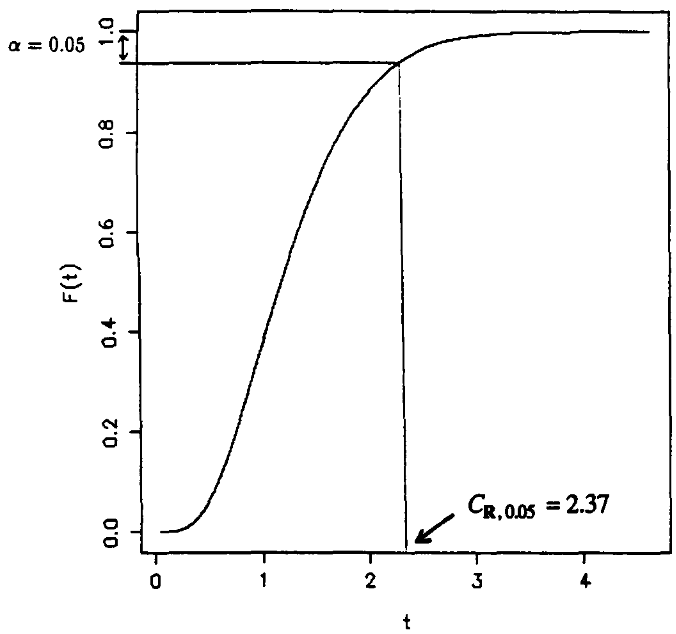

FIGURE 2. Simulated Empirical Cumulative Distribution Function for Ballistic Missiles Example.

ulations of the statistic $\max_{1\leq i\leq 4} \mid X_i \mid$ for $\mathbf{X} \sim N_4(\mathbf{0}, \mathbf{R})$ is given in Figure 2, and the size $\alpha = 0.05$ level critical point is estimated as $C_{\mathbf{R},0.05} = 2.37$. To obtain an indication of the accuracy of this critical point, nine additional independent simulations (each of size 100,000) were performed and the critical point $C_{\mathbf{R},0.05}$ was estimated in each case. For each simulation the critical point estimate was within 0.01 of the initial estimate, which suggests that a value of $C_{\mathbf{R},0.05} = 2.37$ can be considered an accurate value. Figure 2 can also be used to provide critical points for additional size α levels.

For a given observation $\mathbf{x} = (x_1, x_2, x_3, x_4)'$, the 0.95 confidence level simultaneous confidence intervals for the four variable means are given in Table 2. The process will be declared to be out of control whenever any of these confidence intervals does not contain zero. For example, with an observation $\mathbf{x} = (30, -12, -25, 10)'$, the process will be declared to be out of control due to changes in variables 1 and 3, and more specifically, the experimenter can conclude with 95% confidence that $\mu_1 \in [6.0, 54.0]$ and $\mu_3 \in [-46.8, -3.2]$.

Doganaksoy et al. (1991) considered an observation $\mathbf{x} = (15, 10, 20, -5)'$ for which the χ^2 statistic declares the process to be out of control at the $\alpha = 0.05$ error rate. However, it is clear from the confidence intervals in Table 2 that there is a problem with identification since in reality none of the four variables can be identified as having changed from their control value at the 0.95 confidence level. The reason why the χ^2 statistic has triggered the alarm is that it has found a linear combination of the variables which seems to have changed. Note that for this observation all of the variables have shifted above the mean with the exception of x_4. Since these variables are highly correlated, if there is no shift in the mean, the inconsistency of x_4 from other variables may be due to a shift of the covariance matrix. However, if there is no shift in the covariance matrix, the inconsistency of x_4 from other variables may be due to a shift of the means of the other three variables but not of x_4, which could be detected by the M statistic with a larger significance level. Of course it is possible that x_4 may simply be an outlier.

In fact, Figure 2 can be used to calculate the p-value of the process being in control given this observation. At $\mathbf{x} = (15, 10, 20, -5)'$,

$$\begin{aligned} M &= \max_{1\leq i\leq 4} \mid x_i \mid/\sigma_i \\ &= \max\left\{\frac{15}{\sqrt{102.74}}, \frac{10}{\sqrt{142.74}}, \frac{20}{\sqrt{84.57}}, \frac{5}{\sqrt{99.06}}\right\} \\ &= \frac{20}{\sqrt{84.57}} = 2.175 \end{aligned}$$

so that the p-value is $1 - F(2.175) \doteq 0.08$. In other words, the process will be considered to be in control for all error rates $\alpha \leq 0.08$. However, if $\alpha = 0.10$ say, then Figure 2 gives a critical point of $C_{\mathbf{R},0.10} = 2.08$,

TABLE 2. 95% Simultaneous Confidence Intervals

$$\begin{aligned} \mu_1 &\in \left[x_1 - 2.37\sqrt{102.74}, x_1 + 2.37\sqrt{102.74}\right] = [x_1 - 24.0, x_1 + 24.0] \\ \mu_2 &\in \left[x_2 - 2.37\sqrt{142.74}, x_2 + 2.37\sqrt{142.74}\right] = [x_2 - 28.3, x_2 + 28.3] \\ \mu_3 &\in \left[x_3 - 2.37\sqrt{84.57}, x_3 + 2.37\sqrt{84.57}\right] = [x_3 - 21.8, x_3 + 21.8] \\ \mu_4 &\in \left[x_4 - 2.37\sqrt{99.06}, x_4 + 2.37\sqrt{99.06}\right] = [x_4 - 23.6, x_4 + 23.6] \end{aligned}$$

and the process is declared to be out of control due to variable 3, which has confidence interval

$$\mu_3 \in \left[20 - 2.08\sqrt{84.57}, 20 + 2.08\sqrt{84.57}\right] = [0.87, 39.13].$$

Finally, it is again instructive to compare the exact critical points $C_{\mathbf{R},\alpha}$ with the conservative Dunn-Sidak and Bonferroni critical points. In general, if $\mathbf{R}$ is an identity matrix $\mathbf{I}$, the Dunn-Sidak critical point is exact, that is, $C_{\mathbf{I},\alpha} = z_{\alpha^*}$, where $\alpha^* = (1-(1-\alpha)^{1/k})/2$. For $k = 4$ and $\alpha = 0.05$, $z_{\alpha^*} = 2.491$, which is not too much larger than the exact critical point $C_{\mathbf{R},0.05} = 2.37$ (for $k = 4$ and $\alpha = 0.05$ the Bonferroni critical point is $z_{0.05/8} = 2.498$). It is clear that in this case it would not be overly wasteful to use these conservative critical points in place of the exact simulated critical points. However, for $\alpha = 0.10$ the exact and approximate Dunn-Sidak critical points are $C_{\mathbf{R},0.10} = 2.08$ and $z_{\alpha^*} = 2.226$, so that for the example discussed above where $M = 2.175$, the process is detected as being out of control with the exact critical point but not with the approximate critical point.

Nonparametric Method

Up to this point we have assumed that the mean vector $\boldsymbol{\mu}^0$ and the covariance matrix $\boldsymbol{\Sigma}$ were known and that the observations $\mathbf{x}$ had a multivariate normal distribution. The values of $\boldsymbol{\mu}^0$ and $\boldsymbol{\Sigma}$ are generally calculated from a pool of prior observations. If this is a large pool, say 500 observations or more, then a nonparametric approach to the problem can be adopted, which does not require the assumption of normality.

The nonparametric approach operates in a manner similar to the simulation method of critical point evaluation discussed above, except that the empirical cumulative distribution function is calculated not from a set of values simulated under the assumption of normality, but rather from the pool of actual observations. Specifically, if $\mathbf{x}^1, \ldots, \mathbf{x}^N$ is a pool of observations on a process obtained under conditions that are to be designated as the control status, then the experimenter first calculates the sample mean vector $\bar{\mathbf{x}}$ and sample covariance matrix $\mathbf{S} = \{s_{ij}\}$. Next, the statistics

$$M^j = \max_{1\le i\le k} \mid x_i^j - \bar{x}_i \mid / \sqrt{s_{ii}}$$

are calculated for $1 \le j \le N$. The empirical cumulative distribution function $F(t)$ is formed using these values. This nonparametric empirical cumulative distribution function has the same interpretation as the simulated empirical cumulative distribution function discussed above, so that, for example, a size α level critical point c_α can be found as the solution to $F(c_\alpha) = 1 - \alpha$.

Given a new observation $\mathbf{x} = (x_1, \ldots, x_k)'$, a new statistic

$$M = \max_{1\le i\le k} \mid x_i - \bar{x}_i \mid / \sqrt{s_{ii}}$$

is formed, and for an error rate of α, the process is declared to be in control as long as $M \le c_\alpha$ and out of control if $M > c_\alpha$. More generally, a p-value for the process being in control can be calculated as $1 - F(M)$. Such a procedure controls the overall error rate at the desired level α since under the assumption that $\mathbf{x}^1, \ldots, \mathbf{x}^N$ are independent observations from some common "control" distribution, there is approximately probability α that a new independent observation from the same distribution will generate a statistic M that is within the top $N\alpha$ of the statistics $M^1, \ldots, M^N$. The procedure has only approximate level due to discreteness and to the correlations between the x_i^j and the sample statistics $\bar{x}_i$ and s_{ii}. These correlations will be negligible for reasonably large sample sizes N.

The advantage of this nonparametric procedure is that it can control the error rate without requiring the assumption that the variable measurements have a normal distribution. In the nonparametric procedure, a new observation is compared directly with the pool of previous control observations, whereas the normal assumption procedure compares a new observation with what would be expected for a normal random variate with the mean vector and covariance structure estimated from the pool of previous control observations. However, the nonparametric procedure is only applicable if a large enough initial pool of observations is available. In many quality control settings, however, there are large pools of observations that are available to the experimenter.

When this nonparametric procedure declares that the process is out of control, the identification of which variables are responsible can intuitively be done by seeing which of the variables give values of $\mid x_i - \bar{x}_i \mid / \sqrt{s_{ii}}$ that exceed the critical value c_α. A rough guide to the new values taken by the variable mean can be obtained from the intervals

$$\mu_i \in [x_i - c_\alpha\sqrt{s_{ii}}, x_i + c_\alpha\sqrt{s_{ii}}]$$

although these could be misleading if the actual distribution is very skewed.

As an example of the implementation of the nonparametric procedure, a sample of 500 bivariate observations $\mathbf{x}^j = (x_1^j, x_2^j)'$, $1 \leq j \leq 500$, was simulated by taking $x_1 = \max\{z_1, z_2\}$ and $x_2 = z_1^2 + z_2^2$ for independent standard normal random deviates z_1 and z_2. Clearly, this is a pool of 500 observations whose common distribution is not normal. The sample average was $\bar{\mathbf{x}} = (1.12, 1.94)'$ and the sample covariance matrix was

$$\mathbf{S} = \begin{bmatrix} 0.354 & 1.08 \\ 1.08 & 3.69 \end{bmatrix}.$$

The statistics $M^1, \ldots, M^{500}$ were formed by taking

$$M^j = \max\left\{\frac{|\, x_1^j - 1.12\,|}{\sqrt{0.354}}, \frac{|\, x_2^j - 1.94\,|}{\sqrt{3.69}}\right\}$$

and their empirical cumulative distribution function $F(t)$ is shown in Figure 3. The corresponding critical points are $c_{0.10} = 1.57$, $c_{0.05} = 2.08$, and $c_{0.01} = 3.80$. For a new observation $\mathbf{x} = (x_1, x_2)'$ and $\alpha = 0.05$, the process is considered to be in control as long as

$$\max\left\{\frac{|\, x_1 - 1.12\,|}{\sqrt{0.354}}, \frac{|\, x_2 - 1.94\,|}{\sqrt{3.69}}\right\} \leq 2.08$$

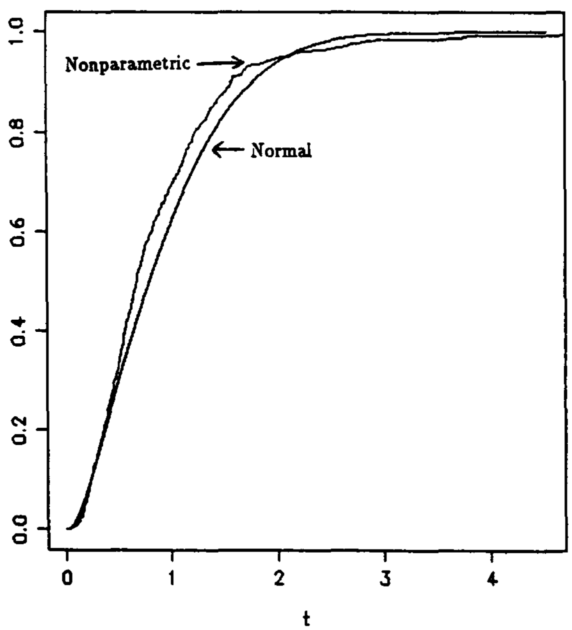

FIGURE 3. Nonparametric and Simulated Normal Empirical Cumulative Distribution Functions.

or, as long as $-0.12 \leq x_1 \leq 2.36$ and $0 \leq x_2 \leq 5.94$ (since necessarily $x_2 \geq 0$).

To illustrate the advantage of the nonparametric procedure, suppose that this problem had been approached with the (incorrect) assumption of normality. Then only the information given by the sample average $\bar{\mathbf{x}}$ and the covariance matrix $\mathbf{S}$ would have been extracted from the pool of 500 observations, and subsequent observations would be assumed to be normally distributed with this covariance matrix. Figure 3 also shows the empirical cumulative distribution function of the statistic

$$\max\{|\, Z_1\,|, |\, Z_2\,|\}$$

calculated from 100,000 simulations under the assumption that $(Z_1, Z_2)'$ come from a bivariate $N_2(\mathbf{0}, \mathbf{R})$ distribution, where $\mathbf{R}$ is the correlation matrix corresponding to the covariance matrix $\mathbf{S}$. The critical points $C_{\mathbf{R},\alpha}$ can be calculated from this second empirical cumulative distribution function. The differences between the two superimposed empirical cumulative distribution functions illustrate the fact that in this example the assumption of normality is invalid, and its adoption would lead to poor choices of the critical points. In fact, under the normality assumption, the critical points are $C_{\mathbf{R},0.10} = 1.77$, $C_{\mathbf{R},0.05} = 2.08$, and $C_{\mathbf{R},0.01} = 2.68$, which while coincidentally matching the nonparametric critical point at $\alpha = 0.05$ are substantially larger and smaller in the other cases.

Summary

A good multivariate quality control procedure is one that, at a specified error rate α, triggers the out-of-control alarm only with probability α when the process is still in control, and triggers the alarm as quickly as possible when the process is out of control. In addition, it should provide a simple and easily implementable mechanism for deciding which of the variables are responsible when the process is determined to be out of control. Finally, it should allow the easy quantification of the amount by which the out-of-control variables have changed in mean value. The simple procedure proposed in this paper based on simultaneous confidence intervals for the variable means meets all of these criteria.

Previous simultaneous confidence interval approaches to the multivariate quality control problem have been based on the inefficient Bonferroni inequality, and consequently it has been observed that they are less powerful than the common Shewhart-type χ^2 control chart in triggering the out-of-control alarm. Under the assumption that in-control observations are $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, the procedure proposed in this paper produces exact simultaneous confidence intervals, from which Shewhart-type charts with overall size exactly α can be obtained. Neither the procedure based on the χ^2 statistic nor the one based on the M statistic is uniformly more powerful. As illustrated in Doganaksoy et al. (1991) and earlier in this paper, one procedure may be more powerful than the other when the variable mean shifts in a particular direction and the reverse may be true when the shift is in another direction. Therefore, it is not correct to argue that one procedure is better or more sensitive than the other in triggering out-of-control alarms.

However, it was also shown earlier in this paper that if the χ^2 statistic is used, then it is sometimes difficult to identify which of the variables has caused the alarm. The reason why the χ^2 statistic has triggered the alarm is that it has found a linear combination of the variables which seems to have changed. However, the identification of such a linear combination is not necessarily meaningful to the experimenter. On the other hand, the new procedure, which is associated with a set of simultaneous confidence intervals for the variable means, can immediately and simply identify the variables that are responsible for the out-of-control alarm.

An additional advantage of the new procedure is that it is easy to implement without the usual assumption of normality when a large pool of prior observations is available. As illustrated earlier in this paper, this nonparametric implementation can give more accurate results when the actual true distribution is quite different from the normal distribution. However, as with all nonparametric procedures, there is a slight loss of sensitivity and power if the distribution is close to normality. Woodall and Ncube (1985) proposed the use of modulus statistics similar to the M statistic for analyzing the cumulative sums of observations in a multivariate procedure. Actually, the M statistic can be considered as a special case of the modulus statistic. Similar extensions should also be possible in the EWMA setting.

Acknowledgments

The authors would like to express their appreciation to the Editor and two referees for their helpful comments and suggestions. Kwok-Leung Tsui's work was supported by National Science Foundation grants #DDM-9114554 and #DDM-9257918. Tony Hayter's work was supported by National Science Foundation grant #DMS-9224370.

References

ALT, F.B. (1985). "Multivariate Quality Control" in *Encyclopedia of Statistical Sciences* 6, edited by S. Kotz and N. L. Johnson. John Wiley & Sons, New York, NY.

ALT, F. B.; WALKER, J. W.; and GOODE, J. J. (1980). "A Power Paradox for Testing Bivariate Normal Means". *ASQC Annual Technical Conference Transactions,* pp. 754-759.

BECHHOFER, R. E. and DUNNETT, C. W. (1988). "Percentage Points of Multivariate Student t Distributions". *Selected Tables in Mathematical Statistics* 11, American Mathematical Society, Providence, RI.

BOZZELO, M. (1989). "A Comparative Run Length Analysis of Hotelling T^2 Control Chart and Multiple Shewhart Charts with Run Rules." Unpublished Master's Thesis, Southern Illinois University-Edwardsville, Department of Mathematics & Statistics, Edwardsville, IL.

DOGANAKSOY, N., FALTIN, F. W. and TUCKER, W. T. (1991). "Identification of Out of Control Quality Characteristics in a Multivariate Manufacturing Environment". *Communications in Statistics - Theory and Methods* 20, pp. 2775-2790.

DUNN, O. J. (1958), "Estimation of the Means of Dependent Variables", *Annals of Mathematical Statistics* 29, pp. 1095-1111.

HOCHBERG, Y. and TAMHANE, A. C. (1987). *Multiple Comparisons Procedures.* John Wiley & Sons, New York, NY.

HOTELLING, H. (1947). "Multivariate Quality Control" in *Techniques of Statistical Analysis,* edited by Eisenhart, Hastay, and Wallis. McGraw-Hill, New York, NY.

IMSL (1987). *International Mathematical and Statistical Library.* Houston, TX.

JACKSON, J. E. (1980). "Principal Components and Factor Analysis: Part I - Principal Components". *Journal of Quality Technology* 12, pp. 201-213.

JACKSON, J. E. (1985). "Multivariate Quality Control". *Communications in Statistics - Theory and Methods* 14, pp. 2657-2688.

JACKSON, J. E. (1991). *A User's Guide to Principle Component.* John Wiley & Sons, New York, NY.

JACKSON, J. E. and MUDHOLKAR, G. S. (1979). "Control Procedures for Residuals Associated with Principal Components Analysis". *Technometrics* 21, pp. 341-349.

LAW, A. M. and KELTON, W. D. (1982). *Simulation Modeling and Analysis.* McGraw-Hill, New York. NY.

MONTGOMERY, D. C. (1991). *Introduction to Statistical Quality Control,* 2nd. ed. John Wiley & Sons, New York, NY.

MURPHY, B. J. (1987). "Selecting Out of Control Variables with the T^2 Multivariate Quality Control Procedure". *The Statistician* 36, pp. 571-583.

ODEH, R. E. (1982). "Tables of Percentage Points of the Distribution of the Maximum Absolute Value of Equally Correlated Normal Random Variables". *Communications in Statistics - Simulation and Computation* 11, pp. 65-87.

SIDAK, Z. (1967), "Rectangular Confidence Regions for the Means of Multivariate Normal Distribution". *Journal of American Statistical Association* 62, pp. 626-633.

WOODALL, W. H. (1986). "The Use of Multivariate Control Charts". Paper presented at the International Research Conference on Reliability and Quality, University of Missouri-Columbia, Columbia, MO.

WOODALL, W. H. and NCUBE, M. M., (1985). "Multivariate CUSUM Quality Control Procedures". *Technometrics* 27, pp. 285-292.

Key Words: *Dunn-Sidak Inequality, Hotelling's T^2 Chart, Multivariate Quality Control, Simultaneous Confidence Intervals, χ^2 Control Chart.*