Statistica Sinica **22** (2012), 847-868
doi:http://dx.doi.org/10.5705/ss.2009.239

METHODS FOR IDENTIFYING INFLUENTIAL VARIABLES IN AN OUT-OF-CONTROL MULTIVARIATE NORMAL PROCESS

Chia-Ling Yen and Jen Tang

National Tsing-Hua Unversity and Purdue University

Abstract: Hotelling's T^2 is a well-known statistic for testing the mean vector of a multivariate normal distribution. Control charts based on T^2 have been widely used in statistical process control for monitoring a multivariate process. Although it is a powerful tool, the T^2 statistic has a practical problem, namely, that a significant T^2-value that normally signals an overall out-of-control condition in the process mean vector does not provide direct information about which variable or group of variables may have caused this out-of-control condition. We propose a diagnostic method to identify the influential variable(s) for cases with and without a specified out-of-control mean vector. Our approach, based on the likelihood principle, computes the conditional likelihood of a variable or sub-group of variables causing or not causing the overall out-of-control condition. Unlike many existing methods, our method assumes that an out-of-control condition already exists; hence, all conditional likelihoods in this paper are based on non-central distributions of the monitoring/testing statistics. By comparing these conditional likelihoods, we identify the influential variable(s). We use an example from the literature to illustrate our method and to demonstrate its effectiveness.

Key words and phrases: Hotelling's T^2 statistic, hypothesis testing, influential variables, likelihood, mean vector, multivariate process control, out-of-control.

1. Introduction

To test a hypothesis about the mean vector of a multivariate normal distribution, Hotelling (1947) proposed a T^2 statistic that has been widely used in statistical process control (SPC) to monitor a multivariate normal process. Thus a process/population with p quality variables (characteristics) in $X = (X_1, \ldots, X_p)$ is assumed to follow a multivariate normal distribution with an unknown mean vector μ and unknown but constant (in-control) covariance matrix Σ. The process is said to be in-control in its mean (or simply in-control) at a given time if the hypothesis $H_0 : \mu = \mu_0$ cannot be rejected based on a random sample taken from the process at that time, where μ_0 represents the in-control process mean. On the other hand, μ can shift from μ_0 at an unknown time, and the main purpose of SPC is to detect this shift as soon as possible.

Since the main concern in SPC practice is the stability of the mean vector, μ_0 is normally not specified and is estimated using some reference/training data so a T^2 control chart can be set up in Phase I for future process monitoring in Phase II. Assume we have reference observations, $X_{R,i}$, $i = 1, \ldots, N$, from the in-control process in Phase I. We compute $\bar{X}_R = \sum_{i=1}^{N} X_{R,i}/N$ and $S_R = \sum_{i=1}^{N}(X_{R,i} - \bar{X}_R)(X_{R,i} - \bar{X}_R)'/(N-1)$ to estimate μ_0 and Σ, respectively. To see the decision rule for monitoring the process mean during Phase II, assume a random observation X is taken from the process $N_p(\mu, \Sigma)$. A monitoring statistic commonly used in practice (cf., Anderson (2003) and Hotelling (1947)) is

$$T^2 = \frac{N}{N+1}(X - \bar{X}_R)'S_R^{-1}(X - \bar{X}_R). \tag{1.1}$$

It is well known that $[(N-p)/((N-1)p)]T^2$ follows a non-central $F_{p,N-p,\lambda}$ distribution with non-centrality $\lambda = (N/(N+1))(\mu - \mu_0)'\Sigma^{-1}(\mu - \mu_0)$ (Anderson (2003)). Under $H_0 : \mu = \mu_0$, $\lambda = 0$ and the distribution is a central F distribution. Hence, our $100(1-\alpha)\%$ decision rule is: $H_0 : \mu = \mu_0$ is rejected and an out-control signal is triggered if $T^2 > t_0^2(\alpha) = [(N-1)p/(N-p)]F_{p,N-p,0}(\alpha)$, where $F_{p,N-p,0}(\alpha)$ is the $(1-\alpha)$ percentile of $F_{p,N-p,0}$. This $t_0^2(\alpha)$ is the $100(1-\alpha)\%$ control limit on a T^2 control chart.

When T^2 signals a change in the mean vector, corrective action is required. A T^2 value, however, does not provide direct information about which variable is responsible for the overall out-of-control condition. This information is of practical importance because engineers/analysts need to know which individual variable requires adjustments after the process is declared out-of-control. Identifying the influential variable(s) in a Hotelling's T^2 statistic has been studied by several authors. Alt (1985) proposed a set of Bonferroni control limits for each individual variable; Hayter and Tsui (1994) proposed a procedure to obtain control limits so that the overall alarm rate, α, is at or close to the desired value. A different approach, based on Rencher's (1993) decomposition of the T^2 statistic, can be found in the following references: the MTY decomposition of Mason, Tracy, and Young (1995, 1997) (more details can be found in Mason, Tracy, and Young (2002)); the union-intersection or step-down method of Roy (1958); the marginal contribution method of Murphy (1987); the numerical method of Doganaksoy, Faltin, and Tucker (1991); the regression-adjusted variables method of Hawkins (1991, 1993); and the finite intersection test of Timm's (1996).

Jackson (1991) and Fuchs and Benjamini (1994) recommended principal component analysis (*PCA*) for improving the interpretation of T^2. Kourti and MacGregor (1996) provided another approach based on *PCA* and partial least squares. Contribution plots proposed by Wasterhuis, Gurden, and Smilde (2000) can be constructed, for the normalized principal component scores with high values, to find the variables responsible for the out-of-control signal. Maravelakis et al. (2002) proposed a new method based on PCA to identify the variable or variables responsible for an out-of-control signal in the χ^2 control chart.

The cause-selecting chart (CSC) proposed by Zhang (1980, 1984, 1992) is a different approach to solving the problem of interpreting an out-of-control signal in the T^2 chart. Wade and Woodall (1993) suggested a modification of the CSC chart for diagnostic purposes and also investigated the relationship between cause-selection control and the multivariate T^2 chart. Sepulveda (1996) developed a Minimax control chart that can give evidence about which variable is causing the out-of-control signal. The Minimax control chart was again discussed in Sepulveda and Nachlas (1997) and is similar to the charts proposed by Hayter and Tsui (1994) and Timm (1996). Kalagonda and Kulkarni (2003) proposed a diagnostic procedure called 'D-technique,' using dummy variables in a multiple-regression equation.

An important adjunct to the statistical procedures is a suitable graphical scheme that can display the basic features of the data (see Iglewicz and Hoaglin (1987), Fuchs and Benjamini (1994), Sparks, Adolphson, and Phatak (1997), and Atienza, Tang, and Ang (1998)). More recent approaches based on an artificial neural network and decision tree can be found in Aparisi, Avendano, and Sanz (2006), Chen and Wang (2004), Guh and Shiue (2008), and references therein.

In this paper we propose a method based on the likelihood principle for identifying a variable or group of variables most likely to be responsible for the rejection of $H_0 : \mu = \mu_0$. We consider the cases with and without a specified out-of-control mean vector. When no out-of-control mean is specified, our method computes the conditional likelihood that an individual mean or a group of means is in-control, given that $H_0 : \mu = \mu_0$ is rejected. When an alternative hypothesis with a specified out-of-control mean is given, our method computes the conditional likelihood that an individual mean or a group of means has shifted in the direction specified by the overall alternative, given that $H_0 : \mu = \mu_0$ is rejected.

By comparing these conditional likelihoods, we identify the influential variable(s). Our method assumes that the process is already out-of-control and is therefore a diagnostic tool. In contrast, many existing methods assume that $H_0 : \mu = \mu_0$ is true when deriving the distributions for their monitoring statistics; e.g., the central F distributions for the decomposed statistics in Mason, Tracy, and Young (1995) and z-distributions for the regression-adjusted variables in Hawkins (1991). Note that, when $H_0 : \mu = \mu_0$ (in-control) is not rejected, we may not be interested in the identification problem.

This paper is organized as follows. Section 2 describes our proposed method and provides formulas for computing the conditional likelihoods used in the method. In Section 3, we illustrate our method using an example taken from the literature.

2. The Proposed Method Based on the Likelihood Principle

In Section 2.1, we describe our proposed method for identifying the influential variable(s) when $H_0 : \mu = \mu_0$ is rejected. Section 2.2 provides formulas for computing the likelihood of causing an out-of-control condition for each variable or group of variables.

2.1. Description of proposed method

We consider two cases for the alternative hypothesis: one specifies and the other does not specify the out-of-control mean. For the first case, let the specified out-of-control mean vector be $\mu_a^* \equiv (\mu_a^{*(1)'}, \ldots, \mu_a^{*(k)'})'$ $(\neq \mu_0)$, where $\mu_a^{*(j)}$ is a $p_j \times 1$ sub-vector and $\sum_{j=1}^{k} p_j = p$. This μ_a^* and its partition need to be determined before samples are taken from the process in Phase II, and they represent the user's belief or conjecture about the out-of-control mean vector. Hence, the two cases are

(A) $H_a : \mu = \mu_a^*$,

(B) H_a : not H_0.

Let X be a random sample taken from $N_p(\mu, \Sigma)$ during Phase II and x be the observed value. Assume that the observed $t^2 \equiv (N/(N+1))(x-\bar{X}_R)' S_R^{-1}(x-\bar{X}_R)$ of T^2 satisfies $t^2 > t_0^2(\alpha)$, so $H_0 : \mu = \mu_0$ is rejected at significance level α. Note that, when H_0 is rejected, X and $\bar{X}_R$ do not have the same mean.

For Case (A), to detect the out-of-control individual mean(s), we similarly partition $\mu = (\mu^{(1)'}, \ldots, \mu^{(k)'})'$, $\mu_0 = (\mu_0^{(1)'}, \ldots, \mu_0^{(k)'})'$, $X \equiv (X^{(1)'}, \ldots, X^{(k)'})'$, and $x = (x^{(1)'}, \ldots, x^{(k)'})'$. Define $H_{0j} : \mu^{(j)} = \mu_0^{(j)}$ and $H_{aj} : \mu^{(j)} = \mu_a^{*(j)}$, for $j = 1, \ldots, k$. Then, the question is: Which of the hypotheses H_{aj} is most likely to be true according to the data, given that $H_a : \mu = \mu_a^*$ is accepted? We compute the conditional maximum likelihood, $\ell_j(H_{aj} \mid H_a)$, for each H_{aj}. By comparing these likelihoods, we identify which mean vector is most likely responsible for the overall out-of-control condition. Our method is different from other critical-value types of approach (e.g., Mason, Tracy, and Young (1995) and Hawkins (1991)).

The conditional maximum likelihood of H_{aj} is calculated as

$$\ell_j(H_{aj} \mid H_a) = \max_{\mu^{(j)} \in H_{aj}} f_{\mu^{(j)}}(X^{(j)} = x^{(j)} \mid T^2 = t^2), \tag{2.1}$$

where $f_{\mu^{(j)}}(X^{(j)} = x^{(j)} \mid T^2 = t^2)$ denotes the likelihood of $\mu^{(j)}$, also the conditional pdf of $X^{(j)}$ at $x^{(j)}$, given that $H_0 : \mu = \mu_0$ is rejected with $T^2 = t^2$. Since H_{aj} in Case (A) is a simple alternative, (2.1) reduces to

$$\ell_j(H_{aj} \mid H_a) = f_{\mu_a^{*(j)}}(X^{(j)} = x^{(j)} \mid T^2 = t^2), \qquad j = 1, \ldots, k. \tag{2.2}$$

The largest ℓ-value corresponds to the sub-vector most likely to have caused the overall out-of-control condition.

For Case (B), since no p_j's are pre-specified, we need to consider all possible partitions of the mean vectors, μ and μ_0. For each partition, our question is: which hypothesis H_{0j} is least likely to be true according to the data, given that H_0 is rejected with $T^2 = t^2$ $(> t_0^2(\alpha))$. The conditional likelihood of H_{0j} is similar to (2.1) and can be calculated as (since H_{0j} is simple):

$$\ell_j'(H_{0j} \mid H_a) = f_{\mu_0^{(j)}}(X^{(j)} = x^{(j)} \mid T^2 = t^2). \tag{2.3}$$

For each partition, we first select the sub-set with the smallest ℓ'-value. Then, we identify the variable or variables that have appeared in most or all of the selected sub-sets as the out-of-control variable(s).

2.2. Calculations of conditional likelihoods

To calculate the conditional likelihoods in (2.2) and (2.3), it is sufficient to consider only $\ell_1'(H_{01} \mid H_a)$ and $\ell_1(H_{a1} \mid H_a)$ for $X^{(1)}$ with p_1 $(1 \le p_1 \le p)$ univariate means, because conditional likelihoods for other sub-vectors can be computed similarly by rearranging and renaming the variables in X. To simplify the notation, we group all sub-vectors other than $X^{(1)}$ to obtain

$$X = (X^{(1)\prime}, X^{(2)\prime})', \quad \bar{X}_R = (\bar{X}_R^{(1)\prime}, \bar{X}_R^{(2)\prime})', \quad \text{and} \quad S_R = \begin{bmatrix} S_{R,11} & S_{R,12} \\ S_{R,21} & S_{R,22} \end{bmatrix},$$

where $X^{(1)}$ and $\bar{X}_R^{(1)}$ are $p_1 \times 1$ and $X^{(2)}$ and $\bar{X}_R^{(2)}$ are $q_1 \times 1$ with $q_1 = p - p_1$. Similarly, μ, μ_0, and Σ are partitioned as

$$\mu = (\mu^{(1)\prime}, \mu^{(2)\prime})', \quad \mu_0 = (\mu_0^{(1)\prime}, \mu_0^{(2)\prime})', \quad \text{and} \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.$$

Define the partial (conditional) covariance matrices

$$\Sigma_{22\cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}, \quad \text{and} \quad S_{R,22\cdot 1} = S_{R,22} - S_{R,21}S_{R,11}^{-1}S_{R,12}.$$

Then, T^2 can be decomposed as (see, for example, Mason, Tracy, and Young (1995)):

$$\begin{aligned}T^2 &= \frac{N}{N+1}(X-\bar{X}_R)'S_R^{-1}(X-\bar{X}_R)\\ &= \frac{N}{N+1}(X^{(1)}-\bar{X}_R^{(1)})'S_{R,11}^{-1}(X^{(1)}-\bar{X}_R^{(1)})\\ &\quad+\frac{N}{N+1}[(X^{(2)}-\bar{X}_R^{(2)})-S_{R,21}S_{R,11}^{-1}(X^{(1)}-\bar{X}_R^{(1)})]'\\ &\quad\times S_{R,22\cdot 1}^{-1}[(X^{(2)}-\bar{X}_R^{(2)})-S_{R,21}S_{R,11}^{-1}(X^{(1)}-\bar{X}_R^{(1)})]\\ &\equiv T_1^2+T_{2\cdot 1}^2, \qquad \text{say.}\end{aligned}$$

Theorem 1. (Proof in Appendix A)

(i) T_1^2 *follows a non-central* $[(N-1)p_1/(N-p_1)]F_{p_1,N-p_1,\lambda_1}$ *distribution with non-centrality parameter* $\lambda_1 = [N/(N+1)](\mu^{(1)}-\mu_0^{(1)})'\Sigma_{11}^{-1}(\mu^{(1)}-\mu_0^{(1)})$*; and*

(ii) *The conditional distribution of* $T_{2\cdot 1}^{2(*)} \equiv T_{2\cdot 1}^2/(1+T_1^2/(N-1))$*, given* $T_1^2 = t_1^2 = [N/(N+1)](x^{(1)}-\bar{x}_R^{(1)})'s_{R,11}^{-1}(x^{(1)}-\bar{x}_R^{(1)})$*, is a non-central* $[(N-1)q_1/(N-p)]F_{q_1,N-p,\lambda_2}$ *distribution with*

$$\begin{aligned}\lambda_2 &= \frac{N}{N+1}\frac{1}{1+t_1^2/(N-1)}[(\mu^{(2)}-\mu_0^{(2)})-\Sigma_{21}\Sigma_{11}^{-1}(\mu^{(1)}-\mu_0^{(1)})]'\\ &\quad\times\Sigma_{22\cdot 1}^{-1}[(\mu^{(2)}-\mu_0^{(2)})-\Sigma_{21}\Sigma_{11}^{-1}(\mu^{(1)}-\mu_0^{(1)})]\\ &= \frac{\lambda-\lambda_1}{(1+t_1^2/(N-1))}.\end{aligned}$$

Note that the unconditional distribution of $T_{2\cdot 1}^2$ can be obtained from (i) and (ii) above, and is given by $\int_{t_1^2>0} f_{T_{2\cdot 1}^2|T_1^2}(t_{2\cdot 1}^2 \mid t_1^2)f_{T_1^2}(t_1^2)dt_1^2$. This unconditional distribution can be used to replace the central F-distributions for the decomposed monitoring statistics in the MTY decomposition and can thus be used as a diagnostic tool (assuming that H_0 is rejected).

To compute the conditional likelihood $f_{\mu^{(1)}}(X^{(1)} = x^{(1)} \mid T^2 = t^2)$ when H_0 is rejected, we first note from (i) and (ii) of Theorem 1 that the conditional distribution of T^2, given $(X^{(1)}, \bar{X}_R^{(1)}, S_{R,11}) = (x^{(1)}, \bar{x}_R^{(1)}, s_{R,11})$, is $(1+t_1^2/(N-1))[(N-1)q_1/(N-p)]F_{q_1,N-p,\lambda_2} + t_1^2$, which depends on $(x^{(1)}, \bar{x}_R^{(1)}, s_{R,11})$ but through $t_1^2 = [N/(N+1)](x^{(1)}-\bar{x}_R^{(1)})'s_{R,11}^{-1}(x^{(1)}-\bar{x}_R^{(1)})$. From the fact that $X^{(1)}$, $\bar{X}_R^{(1)}$, and $S_{R,11}$ are independently distributed as $N_{p_1}(\mu^{(1)}, \Sigma_{11})$, $N_{p_1}(\mu_0^{(1)}, (1/N)\Sigma_{11})$, and $W_{p_1}(N-1,[1/(N-1)]\Sigma_{11})$, respectively, we have

$$\begin{aligned}
&f_{\mu^{(1)}}(X^{(1)}=x^{(1)} \mid t^2)=\frac{f_{(X^{(1)},T^2)}(x^{(1)},t^2)}{f_{T^2}(t^2)}\\
&=\frac{1}{f_{T^2}(t^2)}\int_{s_{R,11}>0}\int_{\bar{x}_R^{(1)}} f_{(X^{(1)},T^2)|(\bar{X}_R^{(1)},S_{R,11})}(x^{(1)},t^2 \mid \bar{x}_R^{(1)},s_{R,11})f_{\bar{X}_R^{(1)}}(\bar{x}_R^{(1)})\\
&\quad\times f_{S_{R,11}}(s_{R,11})d\bar{x}_R^{(1)}ds_{R,11}\\
&=\frac{f_{X^{(1)}}(x^{(1)})}{f_{T^2}(t^2)}\int_{s_{R,11}>0}\int_{\bar{x}_R^{(1)}} f_{T^2|(X^{(1)},\bar{X}_R^{(1)},s_{R,11})}(t^2 \mid x^{(1)},\bar{x}_R^{(1)},s_{R,11})f_{\bar{X}_R^{(1)}}(\bar{x}_R^{(1)})\\
&\quad\times f_{S_{R,11}}(s_{R,11})d\bar{x}_R^{(1)}ds_{R,11}.
\end{aligned}$$

Note that $f_{T^2|(X^{(1)},\bar{X}_R^{(1)},S_{R,11})}(t^2 \mid x^{(1)},\bar{x}_R^{(1)},s_{R,11})=0$ if $t_1^2>t^2$. Furthermore, this last conditional distribution is the conditional distribution of $T^2\equiv T_1^2+T_{2\cdot 1}^2$ or $t_1^2+T_{2\cdot 1}^2$ and, from the proof of Theorem 1, it depends on $x^{(1)},\bar{x}_R^{(1)},s_{R,11}$ but through t_1^2. A simple transformation gives

$$\begin{aligned}
&f_{T^2|(X^{(1)},\bar{X}_R^{(1)},S_{R,11})}(t^2 \mid x^{(1)},\bar{x}_R^{(1)},s_{R,11})\\
&\quad=f_{T_{2\cdot 1}^2|(X^{(1)},\bar{X}_R^{(1)},S_{R,11})}(t^2-t_1^2 \mid x^{(1)},\bar{x}_R^{(1)},s_{R,11})\\
&\quad=f_{T_{2\cdot 1}^{2(*)}|(X^{(1)},\bar{X}_R^{(1)},S_{R,11})}\Big(\frac{t^2-t_1^2}{(1+t_1^2/(N-1))} \mid x^{(1)},\bar{x}_R^{(1)},s_{R,11}\Big)\frac{1}{(1+t_1^2/(N-1))},
\end{aligned}$$

where the conditional distribution of $T_{2\cdot 1}^{2(*)}$, given $T_1^2=t_1^2$, can be found in Theorem 1 (ii). Thus, we have

$$\begin{aligned}
&f_{\mu^{(1)}}(X^{(1)}=x^{(1)} \mid t^2)\\
&=\frac{f_{X^{(1)}}(x^{(1)})}{f_{T^2}(t^2)}\int_{s_{R,11}>0}\int_{\bar{x}_R^{(1)}} f_{T_{2\cdot 1}^{2(*)}|(X^{(1)},\bar{X}_R^{(1)},S_{R,11})}\Big(\frac{t^2-t_1^2}{(1+t_1^2/(N-1))} \mid x^{(1)},\bar{x}_R^{(1)},s_{R,11}\Big)\\
&\quad\times\frac{1}{(1+t_1^2/(N-1))}f_{\bar{X}_R^{(1)}}(\bar{x}_R^{(1)})f_{S_{R,11}}(s_{R,11})d\bar{x}_R^{(1)}ds_{R,11}\\
&=\frac{f_{X^{(1)}}(x^{(1)})}{f_{T^2}(t^2)}E^*_{(\bar{X}_R^{(1)},S_{R,11})}\\
&\quad\Bigg(f_{T_{2\cdot 1}^{2(*)}|(X^{(1)},\bar{X}_R^{(1)},S_{R,11})}\Big(\frac{t^2-t_1^2}{(1+t_1^2/(N-1))}\Big|x^{(1)},\bar{X}_R^{(1)},S_{R,11}\Big)\frac{1}{(1+t_1^2/(N-1))}\Bigg), \qquad (2.4)
\end{aligned}$$

where the conditional expectation, E^*, is with respect to $\bar{X}_R^{(1)}$ and $S_{R,11}$, subjected to $t_1^2=[N/(N+1)](x^{(1)}-\bar{X}_R^{(1)})'S_{R,11}^{-1}(x^{(1)}-\bar{X}_R^{(1)})$ and $t_1^2\le t^2$.

Since it is generally difficult to compute the conditional expectation E^* in (2.4), we propose the following asymptotic results for the conditional expectation (proof in Appendix B):

$$\iint_{s_{R,11}>0,\bar{x}_R^{(1)},t_1^2\le t^2} \Big(\sum_{h=0}^{m} f_{asymp}^{(h)} N^{-h}+O(N^{-(m+1)})\Big) f_{\bar{X}_R^{(1)}}(\bar{x}_R^{(1)}) f_{S_{R,11}}(s_{R,11}) d\bar{x}_R^{(1)} ds_{R,11}, \tag{2.5}$$

where

$$f_{asymp}^{(h)} = \sum\nolimits^*_{\{j_1+\cdots+j_7=h\}} e^{-(\lambda-\lambda_1)/2+(t_1^2-t^2)/2} \beta_{j_1}^{(1)}(t_1^2)\beta_{j_2}^{(2)}(1-t^2,-\frac{1}{2})\beta_{j_3}^{(3)}(1-t_1^2,\frac{1}{2})$$
$$\times\Bigg[\sum_{\beta=0}^{\infty}\frac{(t^2-t_1^2)^{q_1/2-1+\beta}}{\Gamma(q_1/2+\beta)\beta!}\Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}\beta_{j_4}^{(4)}(\beta)2^{-(q_1/2+\beta)}\beta_{j_5}^{(5)}(\beta)\beta_{j_6}^{(6)}(t^2,\beta)\beta_{j_7}^{(7)}(t_1^2,\beta)\Bigg], \tag{2.6}$$

and the summation $\sum^*$ is over j_i's such that all $\beta_{j_i}^{(d)}\neq 0$. For example, when $m=2$, (2.5) becomes

$$\iint_{s_{R,11}>0,\bar{x}^{(1)},t_1^2\le t^2} \Big(f_{asymp}^{(0)}+f_{asymp}^{(1)}N^{-1}+f_{asymp}^{(2)}N^{-2}+O(N^{-3})\Big) f_{\bar{X}_R^{(1)}}(\bar{x}_R^{(1)})$$
$$f_{S_{R,11}}(s_{R,11}) d\bar{x}_R^{(1)} ds_{R,11},$$

with

$$f_{asymp}^{(0)} = e^{-(\lambda-\lambda_1)/2+(t_1^2-t^2)/2}\Bigg(\sum_{\beta=0}^{\infty}\frac{(t^2-t_1^2)^{q_1/2-1+\beta}}{\Gamma(q_1/2+\beta)\beta!}\Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}2^{-(q_1/2+\beta)}\Bigg), \tag{2.7}$$

$$\begin{aligned} f_{asymp}^{(1)} &= e^{-(\lambda-\lambda_1)/2+(t_1^2-t^2)/2}\sum_{\beta=0}^{\infty}\frac{(t^2-t_1^2)^{q_1/2-1+\beta}}{\Gamma(q_1/2+\beta)\beta!}\Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}2^{-(q_1/2+\beta)} \\ &\quad\times\Big(\frac{(\lambda-\lambda_1)t_1^2}{2}-\frac{(1-t_1^2)(1-t_1^2-2p)}{4}-\frac{(1-t^2)(1-t^2-2p_1)}{4} \\ &\quad+\frac{q_1}{2}(\frac{q_1}{2}-p-1)-(t^2+t_1^2+p_1-\beta)\beta\Big), \end{aligned} \tag{2.8}$$

and $f_{asymp}^{(2)}$ given in Appendix B (B.16).

To compute both conditional null and alternative likelihoods, we note that (2.5) depends on certain unknown model parameters that need to be estimated from the reference or current data. First, the in-control μ_0 and Σ are respectively estimated by $\bar{X}_R$ and S_R from the reference data. For $\ell_1(H_{a1}\mid H_a)$ in Case (A),

$\mu = (\mu^{(1)'}, \mu^{(2)'})'$ is replaced by $\mu_a^* = (\mu^{*(1)'}, \mu^{*(2)'})'$; for $\ell_1'(H_{01} \mid H_a)$ in Case (B), $\mu = (\mu^{(1)'}, \mu^{(2)'})'$ is estimated by $(\mu_0^{(1)'}, X^{(2)'})'$. All other parameters are known, and variables such as t^2 and t_1^2 are given. Furthermore, since the training sample size is normally $N = 20$ or 25 in a univariate case and much larger in a multivariate case, the two-term approximation in (2.5) suffices. We demonstrate this in the next section.

3. Illustration of Method Using Example

We use an example from the literature to illustrate and evaluate our proposed method. Out-of-control variables are identified for Case (B) and for Case (A).

3.1. Identifying out-of-control variables for Case (B)

We use the data in Flury and Riedwyl (1988, p.151), where five dimensions of switch drums are measured: X_1 is the inside diameter of a drum; X_2, X_3, X_4, and X_5 are the distances from the head to the edges of four sectors cut in the drum, respectively. Hawkins (1991) treated the mean and covariance matrix computed from these data as the in-control population mean and covariance matrix:

$$\mu_0 = (17.960, 10.3, 13.76, 11.08, 11.08, 8.26)', \tag{3.1}$$

$$\Sigma = \sigma' R \sigma, \tag{3.2}$$

where

$$R = \begin{bmatrix} 1 & & & & \\ 0.1388 & 1 & & & \\ 0.3496 & 0.7324 & 1 & & \\ 0.0829 & 0.9130 & 0.6824 & 1 & \\ 0.2652 & 0.6932 & 0.8214 & 0.7640 & 1 \end{bmatrix}$$

is the (symmetric) correlation matrix with standard deviations $\sigma = (1.8622, 1.7053, 1.7090, 1.8718, 2.2114)'$. Hawkins then simulated $N = 35$ training observations from $N_5(\mu_0, \Sigma)$ and 15 observations after adding an upward shift of $0.5\sigma_1$ and $0.25\sigma_5$ to the in-control mean of X_1 and X_5, respectively, while keeping all other process parameters unchanged. The reference/training sample mean and covariance matrix based on the first $N = 35$ observations are respectively (in our notation)

$$\bar{X}_R = (17.6289, 10.3365, 13.6189, 11.1776, 8.2437)', \tag{3.3}$$

where

$$S_R = \begin{bmatrix} 2.7355 & & & & \\ 0.5193 & 2.4673 & & & \\ 1.3496 & 1.6465 & 2.2259 & & \\ 0.8029 & 2.5275 & 1.9026 & 3.4201 & \\ 1.4865 & 2.0266 & 2.4228 & 2.9601 & 4.5689 \end{bmatrix} \quad (3.4)$$

A T^2 statistic is computed for each of the observations 36−50. There is an out-of-control signal at $\alpha = 5\%$ for observation 48; namely, $X = (13.065, 11.625, 14.923, 12.589, 12.446)'$, with the observed $T^2 = t^2 = 22.2447 > t_0^2(0.05) = (5(35-1)/(35-5))F_{5,35-5}(0.95) = 14.3568$. Note that the individual distances of variables between this X and the estimated in-control mean $\bar{X}_R$ are: $-2.7594\sigma_1$, $0.8203\sigma_2$, $0.8741\sigma_3$, $0.7632\sigma_4$, and $1.966\sigma_5$. The complete data is given in Hawkins (1991).

For Case (B) where no alternative is specified, Table 1 gives the (approximated) conditional null likelihood, $\ell'_j(H_{0j} \mid H_a)$, for each $X^{(j)}$. Only two terms in (2.5) are used to compute each likelihood. First, to see the accuracy of our approximation, we used simulation to estimate the exact likelihood because the exact likelihood in (2.4) is difficult to compute (each simulated likelihood was obtained based on 10,000 iterations). The simulation procedure is given in Appendix C. The data in Table 1 indicate that our approximations are quite accurate.

We interpret the results in Table 1. If we believe, for example, that there is only one out-of-control variable (so $p_j = p_1 = 1$), then X_1 is the most likely one and its likelihood of being in-control is about 1/7 of the in-control likelihood for the next variable, X_5. On the other hand, if we think there are two out-of-control variables ($p_1 = 2$), then X_1 and X_5 are the most likely pair. The same interpretation applies to the cases for $p_1 = 3$ and 4.

From Table 1, we see that the sub-groups containing the two out-of-control variables X_1 and X_5 always rank high under each p_1 (from 1 to 4), which indicates that our method is quite effective. Furthermore, we can also compare results from different p_1 values. For example, while the individual conditional in-control likelihood for X_1 and X_5 is 3.189E-0.3 and 2.145E-02, respectively, the joint conditional likelihood that X_1 and X_5 are simultaneously in-control is significantly lower (at 4.355E-07), which is reasonable because these two variables are indeed simultaneously out-of-control in this example. Furthermore, when an in-control variable X_4 is added (to obtain the first case under $p_1 = 3$), the joint in-control conditional likelihood becomes smaller (5.299E-08), but not significantly smaller (note that X_4 is in-control but the observed value is $0.7632\sigma_4$ from its in-control mean). If one is to identify three variables, they are: X_1, X_4

Table 1. Conditional Null Likelihood, $\ell_j'(H_{0j} \mid H_a)$, for The Example in Section 3.1.

p_j	Variable				Simulation	(std error)*	Approximation
1				1	3.289E-03	(7.876E-06)	3.189E-03
				5	2.324E-02	(1.999E-05)	2.145E-02
				4	1.649E-01	(1.872E-05)	1.553E-01
				2	1.849E-01	(2.438E-05)	1.740E-01
				3	1.854E-01	(2.722E-05)	1.744E-01
2			1	5	3.946E-07	(3.401E-09)	4.355E-07
			1	3	6.850E-06	(4.716E-08)	7.654E-06
			1	4	1.899E-04	(6.734E-07)	1.944E-04
			1	2	2.840E-04	(9.285E-07)	2.890E-04
			4	5	4.242E-03	(4.917E-06)	3.951E-03
			3	5	6.125E-03	(6.625E-06)	5.707E-03
			2	5	7.043E-03	(5.910E-06)	6.534E-03
			3	4	5.607E-02	(6.789E-06)	5.314E-02
			2	3	6.596E-02	(8.465E-06)	6.250E-02
			2	4	8.381E-02	(8.866E-06)	7.960E-02
3		1	4	5	4.898E-08	(4.075E-10)	5.299E-08
		1	2	5	1.161E-07	(9.284E-10)	1.260E-07
		1	3	5	1.810E-07	(1.433E-09)	1.987E-07
		1	2	3	1.783E-06	(1.214E-08)	1.954E-06
		1	3	4	2.139E-06	(1.390E-08)	2.366E-06
		1	2	4	1.093E-04	(3.662E-07)	1.123E-04
		3	4	5	1.505E-03	(1.805E-06)	1.409E-03
		2	4	5	1.894E-03	(2.180E-06)	1.776E-03
		2	3	5	2.373E-03	(2.259E-06)	2.218E-03
		2	3	4	3.060E-02	(2.782E-06)	2.929E-02
4	1	3	4	5	1.461E-08	(1.037E-10)	1.438E-08
	1	2	4	5	2.924E-08	(1.963E-10)	2.938E-08
	1	2	3	5	3.392E-08	(2.345E-10)	3.427E-08
	1	2	3	4	9.502E-07	(5.611E-09)	1.032E-06
	2	3	4	5	5.155E-04	(7.543E-07)	4.931E-04

*Simulated averages (with std errors) are based on 10,000 iterations. See Appendix C for the simulation procedure.

and X_5. The reason that X_4 was included before X_2 or X_3 is less clear, and one has to consider the trade-off between some of the distances in (-2.7594σ_1, 0.8203σ_2, 0.8741σ_3, 0.7632σ_4, 1.966σ_5), the correlations, and the regression coefficients in the non-centrality of the conditional likelihood. Nevertheless, the null likelihoods of the first three subgroups under $p_1 = 3$ are fairly close.

According to our procedure in Section 2.1, $\{X_1\}$, $\{X_1, X_5\}$, $\{X_1, X_4, X_5\}$, and $\{X_1, X_3, X_4, X_5\}$ are first selected (for each p_j in Table 1). Since X_1 and

X_5 appear in almost all of the sub-sets, they are (correctly) identified as the out-of-control variables.

We would like to point out that the ranking of the variables based on our conditional likelihood approach may be different from that obtained using the marginal T^2 statistic (Mason, Tracy, and Young (1995)) and z-statistic (Hawkins (1991)) for each individual variable. This is mainly because their methods do not assume that an overall out-of-control condition existed; hence, all their non-centralities were assumed to be zero. We are interested in the identification problem when $H_0 : \mu = \mu_0$ is rejected.

3.2. Identifying out-of-control variables for Case (A)

Since our method for Case (A) can easily be illustrated in terms of shifts in means, we define the parameter $\Delta \equiv \mu - \mu_0 = (\Delta^{(1)'}, \ldots, \Delta^{(k)'})'$ and the hypothesized shifts $\Delta_a^* \equiv \mu_a^* - \mu_0 = (\Delta_a^{*(1)'}, \ldots, \Delta_a^{*(k)'})'$. Then, H_{aj} can be rewritten as $H_{aj} : \Delta^{(j)} = \Delta_a^{*(j)}$. Because Δ_a^* is only a hypothesized vector pre-specified by the user before the process monitoring in Phase II begins, it is possible that the process shifts in a direction different from the Δ_a^* during Phase II. To study the ability of the proposed method to detect out-of-control mean(s), let $\mu_a = (\mu_a^{(1)'}, \ldots, \mu_a^{(k)'})'$ be the actual out-of-control mean and $\Delta_a \equiv \mu_a - \mu_0 = (\Delta_a^{(1)'}, \ldots, \Delta_a^{(k)'})'$ be the actual shift.

We use the example from Section 3.1 to illustrate and evaluate our method. With $\Delta_a = (2.5\sigma_1, 0, 0, 0, 0)'$ and $p_j = 1$, we simulated a random sample from the out-of-control process $N_5(\mu_a, \Sigma)$, with $\mu_a = \mu_0 + \Delta_a$, to obtain $X = (23.19104, 10.53652, 13.89620, 11.01731, 9.57183)'$ with $t^2 = 17.99087 > t_0^2(0.05) = 14.3568$. Hence, H_0 was rejected. Here μ_0 and Σ are given in (3.1) and (3.2), respectively. First, we assume the pre-specified shift is the same as the true shift, $\Delta_a^* = \Delta_a$. For $p_j = 1$, the individual conditional likelihoods, $\ell_j(H_{aj} \mid H_a)$, for X_1 to X_5 are: 0.31519, 0.25097, 0.26626, 0.21544, and 0.15556, respectively.

From these likelihoods, $H_{a1} : \Delta^{(1)} = 2.5\sigma_1$ for X_1 is first and correctly identified by our sample as the most likely sub-group alternative. Furthermore, since $\Delta_a^{(j)} = 0$ for $j = 2, \ldots, 5$, the sub-group alternative $H_{aj} : \Delta^{(j)} = 0$ (which is $\Delta_a^{*(j)}$), $j = 2, \ldots, 5$, should be true. Indeed, they also receive high ℓ_j values. So, by confirming that X_2 to X_5 are in-control, we are able to single out X_1 as the out-of-control variable. We look to see if we still can detect X_1 when the pre-specified shift is $\Delta_a^* = (0, 2.5\sigma_2, 0, 0, 0)' \neq \Delta_a$. Because the true mean shift is $\Delta_a = (2.5\sigma_1, 0, 0, 0, 0)'$, $H_{a1} : \Delta^{(1)} = 0$, and $H_{a2} : \Delta^{(2)} = 2.5\sigma_2$ are expected to be the two least likely alternative hypotheses. Indeed, the $\ell_j(H_{aj} \mid H_a)$-values for X_1 through X_5 are 0.00011, 0.03088, 0.31536, 0.25647, and 0.17074, respectively. These values indicate that X_3-X_5 are most likely to be in-control and X_2 is unlikely to be out-of-control. Again, we are able to detect X_1 as the out-of-control variable.

This example illustrates our method with only one sample. To extend it, we also considered $\Delta_a = (2.5\sigma_1, 0, 0, 0, 2.5\sigma_5)'$ with shifted means in both X_1 and X_5. We simulated 500 random samples, each of size of 1, from $N_5(\mu_0 + \Delta_a, \Sigma)$. Note that for a pre-specified Δ_a^*, each of the 500 simulated samples gives $C_2^5 = 10$ ℓ_j-values for comparison if $p_j = 2$. Assume $\Delta_a^* = (2.5\sigma_1, 2.5\sigma_2, 2.5\sigma_3, 2.5\sigma_4, 2.5\sigma_5)'$. Based on the ℓ_j-values, $H_{a(1,5)} : (\Delta^{(1)}, \Delta^{(5)}) = (2.5\sigma_1, 2.5\sigma_5)$ was first and correctly picked as the most likely alternative in 90.2% of the 500 simulated samples. If $\Delta_a^* = (0, 0, 0, 2.5\sigma_4, 0)'$, our method first identified $H_{a(2,3)} : (\Delta^{(2)}, \Delta^{(3)}) = (0, 0)$ as the most likely sub-group alternative in 86.6% of the 500 simulated cases. This result is reasonable because the true and hypothesized means are equal for (X_2, X_3); thus, our methods eliminated X_2 and X_3 as out-of-control variables. Other alternative hypotheses for pairs of two variables were considered unlikely, because their hypothesized shift vectors were not equal to the respective true shift vectors. For example, the ℓ_j values of $H_{a(1,4)} : (\Delta^{(1)}, \Delta^{(4)}) = (0, 2.5\sigma_4)$, $H_{a(2,4)} : (\Delta^{(2)}, \Delta^{(4)}) = (0, 2.5\sigma_4)$, and $H_{a(4,5)} : (\Delta^{(4)}, \Delta^{(5)}) = (2.5\sigma_4, 0)$ were ranked first in only 0.2%, 0.0%, and 0.0% in the 500 simulated cases, respectively. From this, we see that X_4 is unlikely to be out-of-control, and we again identify the out-of-control variables X_1 and X_5. Yen (2008) has conducted a more extensive simulation study and found that our method is effective in identifying out-of-control variable(s) in all scenarios considered.

Appendix A: Proof of Theorem 1

Part (i) follows immediately from Anderson (2003, p.143). Next, we find the distribution of $T_2^2 \mid \{(X^{(1)}, \bar{X}_R^{(1)}, S_{R,11}) = (x^{(1)}, \bar{x}_R^{(1)}, s_{R,11})\}$. Since (Theorem 3.3.9 of Gupta and Nagar (2000))

$$S_{R,21} \mid \{S_{R,11} = s_{R,11}\} \sim N_{q_1,p_1}(\Sigma_{21}\Sigma_{11}^{-1}s_{R,11}, \frac{1}{N-1}\Sigma_{22\cdot1} \otimes s_{R,11}),$$

we have

$$\begin{aligned} &S_{R,21}S_{R,11}^{-1}(X^{(1)} - \bar{X}_R^{(1)}) \mid \Big\{(X^{(1)}, \bar{X}_R^{(1)}, S_{R,11}) = (x^{(1)}, \bar{x}_R^{(1)}, s_{R,11})\Big\} \\ &\quad \sim N_{q_1}\Big(\Sigma_{21}\Sigma_{11}^{-1}s_{R,11}s_{R,11}^{-1}(x^{(1)} - \bar{x}_R^{(1)}), \frac{1}{N}\frac{N+1}{N-1}t_1^2\Sigma_{22\cdot1}\Big), \end{aligned} \tag{A.1}$$

with $t_1^2 = (N/(N+1))(x^{(1)} - \bar{x}_R^{(1)})' s_{R,11}^{-1}(x^{(1)} - \bar{x}_R^{(1)})$. Furthermore, since X, $\bar{X}_R$, and S_R are independent, from (A.1) and the fact that

$$(X^{(2)}-\bar{X}_R^{(2)}) \mid \{(X^{(1)},\bar{X}_R^{(1)})=(x^{(1)},\bar{x}_R^{(1)})\} \\ \sim N_{q_1}\Big(\mu^{(2)}-\mu_0^{(2)}-\Sigma_{21}\Sigma_{11}^{-1}(x^{(1)}-\bar{x}_R^{(1)}-(\mu^{(1)}-\mu_0^{(1)})),\frac{N+1}{N}\Sigma_{22\cdot 1}\Big)$$

we have

$$\Big((X^{(2)}-\bar{X}_R^{(2)})-S_{R,21}S_{R,11}^{-1}(X^{(1)}-\bar{X}_R^{(1)})\Big)\Big|\Big\{(X^{(1)},\bar{X}_R^{(1)},S_{R,11})=(x^{(1)},\bar{x}_R^{(1)},s_{R,11})\Big\} \\ \sim N_{q_1}(\mu^{(2)}-\mu_0^{(2)})-\Sigma_{21}\Sigma_{11}^{-1}(\mu^{(1)}-\mu_0^{(1)}),\frac{N+1}{N}(1+\frac{1}{N-1}t_1^2)\Sigma_{22\cdot 1}\Big). \tag{A.2}$$

From (A.2), Anderson (2003, p.143), and the fact that $S_{R,22\cdot 1}\sim W_{q_1}(N-p_1-1,(1/(N-1))\Sigma_{22\cdot 1})$ is independent of $(S_{R,21},S_{R,11})$, the conditional distribution of $T_{2\cdot 1}^{2(*)}\equiv T_{2\cdot 1}^2/(1+t_1^2/(N-1))$, given $(X^{(1)},\bar{X}_R^{(1)},S_{R,11})=(x^{(1)},\bar{x}_R^{(1)},s_{R,11})$, is a non-central $[(N-1)q_1/(N-p)]F_{q_1,N-p,\lambda_2}$ distribution with non-centrality λ_2 as given in the theorem. Since this conditional distribution depends on $(x^{(1)},\bar{x}_R^{(1)},s_{R,11})$ but through t_1^2, it is also the conditional distribution of $T_{2\cdot 1}^{2(*)}$, given t_1^2, so we have proved Theorem 1.

Appendix B: Approximation to Expectation in (2.4)

From Theorem 1(ii), the integrand in (2.4) can be written as

$$\frac{e^{-\lambda_2/2}}{N-1}\sum_{\beta=0}^{\infty}\frac{(\lambda_2/2)^{\beta}\Big[(t^2-t_1^2)/[(N-1)(1+t_1^2/(N-1))]\Big]^{q_1/2-1+\beta}}{B(\frac{N-p}{2},\frac{q_1}{2}+\beta)\beta!\Big[1+\frac{t^2-t_1^2}{(N-1)(1+t_1^2/(N-1))}\Big]^{(N-p_1)/2+\beta}}\cdot\frac{1}{1+t_1^2/(N-1)} \\ =e^{-(N-1)(\lambda-\lambda_1)/[2(N-1+t_1^2)]}\frac{(N-1+t_1^2)^{N/2}}{(N-1+t^2)^{N/2}} \\ \times\sum_{\beta=0}^{\infty}\frac{((N-1)(\lambda-\lambda_1)/2)^{\beta}(t^2-t_1^2)^{q_1/2-1+\beta}}{B((N-p)/2,q_1/2+\beta)\beta!}\frac{(N-1+t_1^2)^{-p/2-\beta}}{(N-1+t^2)^{-p_1/2+\beta}}, \tag{B.1}$$

where

$$\Big[B\Big(\frac{N-p}{2},\frac{q_1}{2}+\beta\Big)\Big]^{-1}=\frac{\Gamma\Big((N-p)/2+q_1/2+\beta\Big)}{\Gamma\Big((N-p)/2\Big)\Gamma\Big(q_1/2+\beta\Big)}. \tag{B.2}$$

We find the asymptotic expansion in N^{-1} for (B.1) by finding the asymptotic expansion for each factor in (B.1).

We need the following result to derive several asymptotic expansions.

Lemma B.1. *If* $g(x) = \sum_{k=1}^{m} \alpha_k x^{-k}$ $(1 \le m \le \infty)$*, then* $\exp(g(x)) = \sum_{j=0}^{\infty} \beta_j(m)x^{-j}$*, where* β_j*'s satisfy the following recursive relation:*

$$\beta_0(m) = 1, \quad \beta_j(m) = \frac{1}{j} \sum_{k=0}^{\min(j,m)} k\alpha_k \beta_{j-k}(m) x^{-j}, \qquad j = 1, 2, \ldots. \tag{B.3}$$

Note that the β_j's depend on m only through the number of terms in the two sums in Lemma B.1. This m will be omitted and replaced by, if any, the variables (but not the fixed model parameters) affecting the values of β_j's. When $m \ge 3$, we have: $\beta_0 = 1$, $\beta_1 = \alpha_1$, $\beta_2 = 2\alpha_2 + \alpha_1^2$, and $\beta_3 = 3\alpha_3 + 4\alpha_1\alpha_2 + \alpha_1^3$. The β_j's can be computed rather easily because of the recursive relation in (B.3).

Begin with the first term in (B.1). We note

$$\begin{aligned} -\frac{1}{2}\frac{(N-1)(\lambda-\lambda_1)}{N-1+t_1^2} &= -\frac{\lambda-\lambda_1}{2}(N-1)(N-1+t_1^2)^{-1} \\ &= -\frac{\lambda-\lambda_1}{2}(1-\frac{1}{N})\Big[\sum_{j=0}^{\infty}\Big(\frac{1-t_1^2}{N}\Big)^j\Big] \\ &= -\frac{\lambda-\lambda_1}{2}\sum_{j=1}^{\infty}\frac{(\lambda-\lambda_1)t_1^2(1-t_1^2)^{j-1}}{2}\Big(\frac{1}{N}\Big)^j, \end{aligned}$$

and hence (from Lemma B.1),

$$\begin{aligned} e^{-(N-1)(\lambda-\lambda_1)/[2(N-1+t_1^2)]} &= e^{-(\lambda-\lambda_1)/2} e^{\sum_{k=1}^{\infty}(1/2)(\lambda-\lambda_1)t_1^2(1-t_1^2)^{k-1}(1/N)^k} \\ &= e^{-(\lambda-\lambda_1)/2}\Big(\sum_{j_1=0}^{m}\beta_{j_1}^{(1)}(t_1^2)N^{-j_1} + O(N^{-(m+1)})\Big), \end{aligned} \tag{B.4}$$

with $\alpha_k \equiv (\lambda-\lambda_1)t_1^2(1-t_1^2)^{k-1}/2$; $\beta_0^{(1)}(t_1^2) = 1$, $\beta_1^{(1)}(t_1^2) = (\lambda-\lambda_1)t_1^2/2$, and $\beta_2^{(1)}(t_1^2) = (\lambda-\lambda_1)t_1^2(1-t_1^2)/2 + ((\lambda-\lambda_1)t_1^2/2)^2$.

Next, using the Taylor series expansion for the log function, we obtain

$$\log\Big(1-\frac{a}{N}\Big)^{bN} = -ba + \sum_{k=1}^{\infty}\Big(\frac{-ba^{k+1}}{k+1}\Big)N^{-k}$$

and, from Lemma B.1, we obtain

$$\Big(1-\frac{a}{N}\Big)^{bN} = e^{-ba}\Big(\sum_{j_2=0}^{m}\beta_{j_2}^{(2)}(a,b)N^{-j_2} + O(N^{-(m+1)})\Big), \tag{B.5}$$

where $\beta_{j_2}^{(2)}(a,b)$'s satisfy (B.3) with $\alpha_k = -ba^{k+1}/(k+1)$. For example, $\beta_0^{(2)}(a,b) = 1$, $\beta_1^{(2)}(a,b) = -ba^2/2$, and $\beta_2^{(2)}(a,b) = -2ba^3/3 + b^2a^4/4$. For the second term in (B.1), we now apply (B.5) to obtain:

$$\frac{1}{(N-1+t^2)^{N/2}} = (N-1+t^2)^{-N/2} = N^{-N/2}\Big(1-\frac{1-t^2}{N}\Big)^{-N/2}$$
$$= N^{-N/2}e^{(1-t^2)/2}\Big(\sum_{j_2=0}^{m}\beta_{j_2}^{(2)}(1-t^2,-1/2)N^{-j_2}+O(N^{-(m+1)})\Big), \qquad \text{(B.6)}$$

where $\beta_0^{(2)}(1-t^2,-1/2) = 1$, $\beta_1^{(2)}(1-t^2,-1/2) = (1-t^2)^2/4$, and $\beta_2^{(2)}(1-t^2,-1/2) = (1-t^2)^3/3+(1-t^2)^4/16$. Similarly,

$$(N-1+t_1^2)^{N/2} = N^{N/2}\Big(1-\frac{1-t_1^2}{N}\Big)^{N/2}$$
$$= N^{N/2}e^{-(1-t_1^2)/2}\Big(\sum_{j_3=0}^{m}\beta_{j_3}^{(3)}(1-t_1^2,\frac{1}{2})N^{-j_3}+O(N^{-(m+1)})\Big), \qquad \text{(B.7)}$$

with $\beta_{j_3}^{(3)}$'s satisfying the same recursive relation as $\beta_{j_2}^{(2)}$ in (B.5); for example, $\beta_0^{(3)}(1-t_1^2,1/2) = 1$, $\beta_1^{(3)}(1-t_1^2,1/2) = -(1-t_1^2)^2/4$, and $\beta_2^{(3)}(1-t_1^2,1/2) = -(1-t_1^2)^3/3+(1-t_1^2)^4/16$.

Next, for the first term inside the summation in (B.1), we note that

$$\Big(\frac{(N-1)(\lambda-\lambda_1)}{2}\Big)^{\beta} = \Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}N^{\beta}(1-\frac{1}{N})^{\beta}$$
$$= \Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}N^{\beta}\sum_{j_4=0}^{\beta}\Big((-1)^{j_4}C_{j_4}^{\beta}\Big)N^{-j_4}, \qquad \text{(B.8)}$$

where $C_{j_4}^{\beta}$ is the Binomial coefficient. If we define $\beta_{j_4}^{(4)}(\beta) = (-1)^{j_4}C_{j_4}^{\beta}I_{j_4\le\beta}$, where indicator function $I_{j_4\le\beta} = 1$ for $j_4 \le \beta$, and $= 0$ otherwise, (B.9) can be rewritten as (for all m)

$$\Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}N^{\beta}\Big\{\sum_{j_4=0}^{m}\beta_{j_4}^{(4)}(\beta)N^{-j_4}+N^{-(m+1)}\Big\}. \qquad \text{(B.9)}$$

For (B.2), we use the asymptotic expansion of the log gamma function (Anderson (2003, p.318) and Lemma B.1 to obtain, for each $\beta \ge 0$,

$$\frac{\Gamma(N/2-p/2+q_1/2+\beta)}{\Gamma(N/2-p/2)} = \Big(\frac{N}{2}\Big)^{q_1/2+\beta}\exp\Big(\sum_{k=1}^{m}\alpha_k N^{-k}+O(N^{-(m+1)})\Big),$$
$$= \Big(\frac{N}{2}\Big)^{q_1/2+\beta}\Big(\sum_{j_5=0}^{m}\beta_{j_5}^{(5)}(\beta)N^{-j_5}+O(N^{-(m+1)})\Big), \text{(B.10)}$$

where $\alpha_r \equiv \alpha_r(\beta) = [(-2)^r/(r(r+1))](B_{r+1}(-p/2) - B_{r+1}((q_1-p)/2+\beta))$ and the $\beta^{(5)}_{j_5}(\beta)$ satisfy the recursive relation in (B.3). Here, $B_{r+1}(\cdot)$ is the Bernoulli polynomial of degree $r+1$ and order of unity defined by $\tau e^{h\tau}/(e^\tau - 1) = \sum_{r=0}^\infty (\tau^r/r!)B_r(h)$ (see Anderson (2003, p.318). For example,

$$\begin{aligned}
\beta^{(5)}_0(\beta) &= 1,\\
\beta^{(5)}_1(\beta) &= \alpha_1 = \frac{(-2)}{(1+1)}\Big(B_2(\frac{-p}{2}) - B_2(\frac{q_1-p}{2}+\beta)\Big)\\
&= -\Big[\Big((\frac{-p}{2})^2 - (\frac{-p}{2}) + \frac{1}{6}\Big) - \Big((\frac{q_1-p}{2}+\beta)^2 - (\frac{q_1-p}{2}+\beta) + \frac{1}{6}\Big)\Big]\\
&= (\frac{q_1}{2}+\beta)(\frac{q_1}{2} - p + \beta - 1),\\
\beta^{(5)}_2(\beta) &= 2\alpha_2 + \alpha_1^2,
\end{aligned}$$

where

$$\begin{aligned}
\alpha_2 &= \frac{(-2)^2}{2(2+1)}\Big(B_3(\frac{-p}{2}) - B_3(\frac{q_1-p}{2}+\beta)\Big)\\
&= (\frac{q_1}{2}+\beta)\Big(\frac{1}{3}(\frac{q_1}{2} - p + \beta)(3 - q_1 + p - 2\beta) - \frac{p^2}{6} - \frac{1}{3}\Big).
\end{aligned}$$

Therefore, we obtain the following asymptotic expansion:

$$\begin{aligned}
&\Big[B\Big(\frac{N-p}{2}, \frac{q_1}{2}+\beta\Big)\Big]^{-1}\\
&= \Big[\Gamma\Big(\frac{q_1}{2}+\beta\Big)\Big]^{-1}\Big(\frac{N}{2}\Big)^{q_1/2+\beta}\Big(\sum_{j_5=0}^{m}\beta^{(5)}_{j_5}(\beta)N^{-j_5} + O(N^{-(m+1)})\Big). \quad \text{(B.11)}
\end{aligned}$$

For the last ratio in (B.1), we note that

$$\begin{aligned}
(N-1+t^2)^{p_1/2-\beta} &= N^{p_1/2-\beta}\Big(1 - \frac{1-t^2}{N}\Big)^{p_1/2-\beta}\\
&= N^{p_1/2-\beta}\Big(\sum_{j_6=0}^{m}\beta^{(6)}_{j_6}(t^2,\beta)N^{-j_6} + O(N^{-(m+1)})\Big), \quad \text{(B.12)}
\end{aligned}$$

where $\beta^{(6)}_{j_6}$'s satisfy (B.3) with $\alpha_k \equiv -(1-t^2)^k(p_1/2-\beta)/k$. For example, $\beta^{(6)}_0(t^2,\beta) = 1$, $\beta^{(6)}_1(t^2,\beta) = -(1-t^2)(p_1/2-\beta)$, and $\beta^{(6)}_2(t^2,\beta) = (1-t^2)^2(p_1/2-\beta)(p_1/2-\beta-1)$. Similarly,

$$\begin{aligned}
(N-1+t_1^2)^{-p/2-\beta} &= N^{-p/2-\beta}\Big(1 - \frac{1-t_1^2}{N}\Big)^{-p/2-\beta}\\
&= N^{-p/2-\beta}\Big(\sum_{j_7=0}^{m}\beta^{(7)}_{j_7}(t_1^2,\beta)N^{-j_7} + O(N^{-(m+1)})\Big), \quad \text{(B.13)}
\end{aligned}$$

with $\alpha_k \equiv (1-t_1^2)^k(p/2+\beta)/k$. For example $\beta_0^{(7)}(t_1^2,\beta) = 1$, $\beta_1^{(7)}(t_1^2,\beta) = (1-t_1^2)(p/2+\beta)$, and $\beta_2^{(7)}(t_1^2,\beta) = (1-t_1^2)^2(p/2+\beta)(p/2+\beta+1)$.

Finally, using (B.4)−(B.13), (B.1) can be written

$$\begin{aligned}\sum_{\beta=0}^{\infty}\Big\{&\frac{(t^2-t_1^2)^{q_1/2-1+\beta}}{\Gamma(q_1/2+\beta)\beta!}e^{(t_1^2-t^2)/2}e^{-(\lambda-\lambda_1)/2}\Big(\sum_{j_1=0}^{m}\beta_{j_1}^{(1)}(t_1^2)N^{-j_1}+O(N^{-(m+1)})\Big)\\ &\times\Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}N^{\beta}\Big(\sum_{j_4=0}^{m}\beta_{j_4}^{(4)}(\beta)N^{-j_4}+N^{-(m+1)}\Big)\\ &\times N^{-N/2}\Big(\sum_{j_2=0}^{m}\beta_{j_2}^{(2)}(1-t^2,-\frac{1}{2})N^{-j_2}+O(N^{-(m+1)})\Big)\\ &\times N^{N/2}\Big(\sum_{j_3=0}^{m}\beta_{j_3}^{(3)}(1-t_1^2,\frac{1}{2})N^{-j_3}+O(N^{-(m+1)})\Big)\\ &\times\Big(\frac{N}{2}\Big)^{q_1/2+\beta}\Big(\sum_{j_5=0}^{m}\beta_{j_5}^{(5)}(\beta)N^{-j_5}+O(N^{-(m+1)})\Big)\\ &\times N^{p_1/2-\beta}\Big(\sum_{j_6=0}^{m}\beta_{j_6}^{(6)}(t^2,\beta)N^{-j_6}+O(N^{-(m+1)})\Big)\\ &\times N^{-p/2-\beta}\Big(\sum_{j_7=0}^{m}\beta_{j_7}^{(7)}(t_1^2,\beta)N^{-j_7}+O(N^{-(m+1)})\Big)\Big\}.\end{aligned} \tag{B.14}$$

If we group the powers of N, (B.1) and (B.14) become

$$\begin{aligned}&\sum_{h=0}^{m}\Big\{\sum_{\{j_1+\cdots+j_7\}=h}e^{-(\lambda-\lambda_1)/2+(t_1^2-t^2)/2}\beta_{j_1}^{(1)}(t_1^2)\beta_{j_2}^{(2)}(1-t^2,-\frac{1}{2})\beta_{j_3}^{(3)}(1-t_1^2,\frac{1}{2})\\ &\cdot\Big[\sum_{\beta=0}^{\infty}\frac{(t^2-t_1^2)^{q_1/2-1+\beta}}{\Gamma(q_1/2+\beta)\beta!}\Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta}\beta_{j_4}^{(4)}(\beta)2^{-(q_1/2+\beta)}\beta_{j_5}^{(5)}(\beta)\beta_{j_6}^{(6)}(t^2,\beta)\beta_{j_7}^{(7)}(t_1^2,\beta)\Big]\Big\}\\ &\quad\times N^{-h}+O(N^{-(m+1)})\\ &=\sum_{h=0}^{m}f_{asymp}^{(h)}N^{-h}+O(N^{-(m+1)}),\end{aligned} \tag{B.15}$$

where $f_{asymp}^{(h)}$ is given in (2.6). We need to note that the summation in (2.6) is over all j_i's such that their beta coefficients are different from zero. The beta's are easy to compute because they satisfy the recursive relation in (B.3). The $f_{asymp}^{(h)}$'s are also easy to calculate. As examples, $f_{asymp}^{(0)}$ and $f_{asymp}^{(1)}$ are given in

(2.7) and (2.8), respectively, and

$$
\begin{aligned}
f_{asymp}^{(2)} = e^{-(\lambda-\lambda_1)/2+(t_1^2-t^2)/2} \sum_{\beta=0}^{\infty} \frac{(t^2-t_1^2)^{q_1/2-1+\beta}}{\Gamma(q_1/2+\beta)\beta!} \Big(\frac{\lambda-\lambda_1}{2}\Big)^{\beta} 2^{-(q_1/2+\beta)} \\
\times\Big\{ \Big(\beta_2^{(1)}(t_1^2) + \beta_2^{(2)}(1-t^2,-\frac{1}{2}) + \beta_2^{(3)}(1-t_1^2,\frac{1}{2}) \\
+\beta_2^{(4)}(\beta) + \beta_2^{(5)}(\beta) + \beta_2^{(6)}(t^2,\beta) + \beta_2^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(1)}(t_1^2)\Big(+ \beta_1^{(2)}(1-t^2,-\frac{1}{2}) + \beta_1^{(3)}(1-t_1^2,\frac{1}{2}) \\
+\beta_1^{(4)}(\beta) + \beta_1^{(5)}(\beta) + \beta_1^{(6)}(t^2,\beta) + \beta_1^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(2)}(1-t^2,-\frac{1}{2})\Big(\beta_1^{(1)}(t_1^2) + \beta_1^{(3)}(1-t_1^2,\frac{1}{2}) \\
+\beta_1^{(4)}(\beta) + \beta_1^{(5)}(\beta) + \beta_1^{(6)}(t^2,\beta) + \beta_1^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(3)}(1-t_1^2,\frac{1}{2})\Big(\beta_1^{(1)}(t_1^2) + \beta_1^{(2)}(1-t^2,-\frac{1}{2}) \\
+\beta_1^{(4)}(\beta) + \beta_1^{(5)}(\beta) + \beta_1^{(6)}(t^2,\beta) + \beta_1^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(4)}(\beta)\Big(\beta_1^{(1)}(t_1^2) + \beta_1^{(2)}(1-t^2,-\frac{1}{2}) \\
+\beta_1^{(3)}(1-t_1^2,\frac{1}{2}) + \beta_1^{(5)}(\beta) + \beta_1^{(6)}(t^2,\beta) + \beta_1^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(5)}(\beta)\Big(\beta_1^{(1)}(t_1^2) + \beta_1^{(2)}(1-t^2,-\frac{1}{2}) \\
+\beta_1^{(3)}(1-t_1^2,\frac{1}{2}) + \beta_1^{(4)}(\beta) + \beta_1^{(6)}(t^2,\beta) + \beta_1^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(6)}(t^2,\beta)\Big(\beta_1^{(1)}(t_1^2) + \beta_1^{(2)}(1-t^2,-\frac{1}{2}) \\
+\beta_1^{(3)}(1-t_1^2,\frac{1}{2}) + \beta_1^{(4)}(\beta) + \beta_1^{(5)}(\beta) + \beta_1^{(7)}(t_1^2,\beta)\Big) \\
+\beta_1^{(7)}(t_1^2,\beta)\Big(\beta_1^{(1)}(t_1^2) + \beta_1^{(2)}(1-t^2,-\frac{1}{2}) \\
+\beta_1^{(3)}(1-t_1^2,\frac{1}{2}) + \beta_1^{(4)}(\beta) + \beta_1^{(5)}(\beta) + \beta_1^{(6)}(t_1^2,\beta)\Big)\Big\}. \qquad \text{(B.16)}
\end{aligned}
$$

Finally, we integrate (B.15) or (B.16) with respect to $\bar{X}_R^{(1)}$ and $S_{R,11}$ to obtain (2.5).

Appendix C: Simulation Procedure for Table 1

To obtain the simulated values of (2.4) in Table 1, we first note that $f_{X^{(1)}}(x^{(1)})$ and $f_{T^2}(t^2)$ in (2.4) are, respectively, the pdfs of $N_{p_1}(\mu^{(1)},\Sigma_{11})$ and $[((N -$

$1)p)/(N-p)]F_{p,N-p,\lambda}$ with $\lambda = (N/(N+1))(\mu-\mu_0)'\Sigma^{-1}(\mu-\mu_0)$. The function inside (2.4), $f_{T^{2(*)}_{2\cdot 1}|T^2_1}((t^2-t^2_1)/(1+t^2_1/(N-1))) \mid t^2_1)$, can be obtained from Theorem 1(ii) with $\lambda_2 = (\lambda-\lambda_1)/(1+t^2_1/(N-1))$, where $\lambda_1 = (N/(N+1))(\mu^{(1)} - \mu^{(1)}_0)'\Sigma^{-1}_{11}(\mu^{(1)}-\mu^{(1)}_0)$.

To simulate the conditional expectation in (2.4), we took a random sample from $\bar{X}^{(1)}_R \sim N_{p_1}(\mu^{(1)}_0, \Sigma_{11}/N)$ and $S_{R,11} \sim W_{p_1}(N-1, \Sigma_{11}/(N-1))$, respectively, and computed $t^2_1 = (N/(N+1))(x^{(1)}-\bar{X}^{(1)}_R)'S^{-1}_{R,11}(x^{(1)}-\bar{X}^{(1)}_R)$. If this t^2_1 was less than or equal to the given t^2 value, we computed the value of $f_{T^{2(*)}_{2\cdot 1}|T^2_1}((t^2 - t^2_1)/(1+t^2_1/(N-1))) \mid t^2_1)/[1+t^2_1/(N-1)]$ and, along with $f_{X^{(1)}}(x^{(1)})$ and $f_{T^2}(t^2)$, we obtained a value for (2.4). The non-centrality λ was estimated by the observed value of $(N/(N+1))(X-\bar{X}_R)'S^{-1}_R(X-\bar{X}_R)$ when t^2 was computed and an out-of-control signal occurred. Since some of the parameters in these distributions depend on unknown $\mu = (\mu^{(1)'}, \mu^{(2)'})'$, $\mu_0 = (\mu^{(1)'}_0, \mu^{(2)'}_0)'$, and $\Sigma = (\Sigma_{ij})$ in our example, they were estimated by the observed/given values of $X = (\bar{X}^{(1)'}_R, X^{(2)'})'$, $\bar{X}_R = (\bar{X}^{(1)'}_R, \bar{X}^{(2)'}_R)'$, and $S_R = (S_{R,ij})$, respectively. Note that, for Case (B) considered in Table 1, $\mu^{(1)} = \mu^{(1)}_0$, hence both were estimated by the given value of $\bar{X}^{(1)}_R$.

We repeated the procedure described above 10,000 times. The average and the standard error of the 10,000 simulated values of (2.4) for each case are given in Table 1.

References

Alt, F. B. (1985). Multivariate quality control. In Encyclopedia of Statistical Sciences **6** (Edited by S. Kotz, N. L. Johnson, and C. B. Read), 110-112. Wiley, New York.

Anderson, T. W. (2003). *An Introduction to Multivariate Statistical Analysis.* 3rd edition. John Wiley and Sons, New York.

Aparisi, F., Avendano, G. and Sanz, J. (2006). Techniques to interpret T^2 control chart signals. *IIE Trans.* **38**, 647-657.

Atienza, O. O., Tang, L-C. and Ang, B-W. (1998). Simultaneous monitoring of univariate and multivariate SPC information using boxplots. *Internat. J. Quality Sci.* **3**, 194-204.

Chen, L. H. and Wang, T. Y. (2004). Artificial neural networks to classify mean shifts from multivariate chart signals. *Comput. Indust. Eng.* **47**, 195-205.

Doganaksoy, N., Faltin, F. W. and Tucker, W. T. (1991). Identification of out-of-control quality characteristic in a multivariate manufacturing environment. *Comm. Statist. Theory Methods* **20**, 2775-2790.

Flury, B. and Riedwyl, H. (1988). *Multivariate Statistics: A Practical Approach.* Chapman and Hall, London.

Fuchs, C. and Benjamini, Y. (1994). Multivariate profile charts for statistical process control. *Technometrics* **36**, 182-195.

Guh, R. S. and Shiue, Y. R. (2008). An effective application of decision tree learning for on-line detection of mean shifts in multivariate control charts. *Comput. Indust. Eng.* **28**, 475-493.

Gupta, A. K. and Nagar, D. K. (2000). *Matrix Variate Distributions.* Chapman and Hall, Boca Raton.

Hawkins, D. M. (1991). Multivariate quality control based on regression-adjusted variables. *Technometrics* **33**, 61-75.

Hawkins, D. M. (1993). Regression adjustment for variables in multivariate quality Control. *J. Quality Tech.* **25**, 170-182.

Hayter, A. J. and Tsui, K. L. (1994). Identification and quantification in multivariate quality control problems. *J. Quality Tech.* **26**, 197-208.

Hotelling, H. (1947). Multivariate quality control - illustrated by the air testing of sample bombsights. In *Techniques of Statistical Analysis* (Edited by C. Eisenthart, M. W. Hastay, and W. Wallis). McGraw Hill, New York.

Iglewicz, B. and Hoaglin, D. C. (1987). Use of boxplots for process evaluation. *J. Quality Tech.* **19**, 180-190.

Jackson, J. E. (1991). *A User Guide to Principal Components.* John Wiley and Sons, New York.

Kalagonda, A. A. and Kulkarni, S. R. (2003). Diagnosis of multivariate control chart signal based on dummy variable regression technique. *Comm. Statist. Theory Methods* **32**, 1665-1684.

Kourti, T. and MacGregor, J. F. (1996). Multivariate SPC methods for process and product monitoring. *J. Quality Tech.* **28**, 409-428.

Maravelakis, P. E., Bersimis, S., Panaretos, J. and Psarakis, S. (2002). Identify the out of control variable in a multivariate control chart. *Comm. Statist. Theory Methods* **31**, 2391-2408.

Mason, R. L., Tracy, N. D. and Young, J. C. (1995). Decomposition of T^2 for multivariate control chart interpretation. *J. Quality Tech.* **27**, 99-108.

Mason, R. L., Tracy, N. D. and Young, J. C. (1997). A practical approach for interpreting multivariate T^2 control chart signals. *J. Quality Tech.* **29**, 396-406.

Mason, R. L., Tracy, N. D. and Young, J. C. (2002). *Multivariate Statistical Process Control with Industrial Applications.* ASA and SIAM.

Murphy, B. J. (1987). Selecting out of control variables with the T^2 multivariate quality control procedure. *The Statistician* **36**, 571-583.

Rencher, A. C. (1993). The contribution of individual variables to Hotelling's T^2, Wilks' Λ, and R^2. *Biometrics* **49**, 479-489.

Roy, J. (1958). Step-down procedure in multivariate analysis. *Ann. Math. Statist.* **29**, 1177-1187.

Sepulveda, A. (1996). The minimax control chart for multivariate quality control. Unpublished Ph.D. dissertation, Virginia Polytechnic Institute and State University, Department of Industrial and Systems Engineering.

Sepulveda, A. and Nachlas, J. A. (1997). A simulation approach to multivariate quality control. *Comput. Indust. Eng.* **33**, 113-116.

Sparks, R .S., Adolphson, A. and Phatak, A. (1997). Multivariate process monitoring using the dynamic biplot. *Internat. Statist. Rev.* **65**, 325-349.

Timm, N. H. (1996). Multivariate quality control using finite intersection tests. *J. Quality Tech.* **28**, 233-243.

Wade, M. R. and Woodall, W. H. (1993). A review and analysis of case-selecting control charts. *J. Quality Tech.* **25**, 161-169.

Wasterhuis, J. A., Gurden, S. P. and Smilde, A. K. (2000). Generalized contribution plots in multivariate statistical process monitoring. *Chemometrics and Intelligent Laboratory Systems* **51**, 95-114.

Yen, C. L. (2008). A study on multivariate process monitoring. Doctoral Thesis, Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan.

Zhang, G. X. (1980). A new type of quality control charts allowing the presence of assignable causes-the cause-selecting control chart. *Acta Electronica Sinica* **2**, 1-10.

Zhang, G. X. (1984). A new type of control charts and a theory of diagnosis with control chart. World Quality Congress Transactions, American Society for Quality Control, 175-185.

Zhang, G. X. (1992). Cause-Selecting Control Chart and Diagnosis, Theory and Practice. Aarhus School of Business, Department of Total Quality Management, Aarhus, Denmark.

Institute of Statistics, National Tsing-Hua Unversity, 30010 Hsin-Chu City, Taiwan.

E-mail: clyen@stat.nthu.edu.tw

Krannert Graduate School of Management, Purdue University, West Lafayette, IN 47907 U.S.A.

E-mail: jtang@purdue.edu

(Received September 2009; accepted March 2011)