QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2001; **17**: 119–124

STATISTICAL PROCESS MONITORING USING AN EMPIRICAL BAYES MULTIVARIATE PROCESS CONTROL CHART

CAROL J. FELTZ[1]* AND JYH-JEN HORNG SHIAU[2]
[1] *Division of Statistics, Northern Illinois University, DeKalb, IL USA*
[2] *Institute of Statistics, National Chiao Tung University, Hsinchu, Taiwan*

SUMMARY

In this paper, we describe the theory underlying an empirical Bayesian approach to monitoring two or more process characteristics simultaneously. If the data is continuous and multivariate in nature, often the multivariate normal distribution can be used to model the process. Then, using Bayesian theory, we develop techniques to implement empirical Bayes process monitoring of the multivariable process. Lastly, an example is given to illustrate the use of our techniques. Copyright © 2001 John Wiley & Sons, Ltd.

KEY WORDS: multivariate; quality control; on-line

1. INTRODUCTION

When monitoring a process, one often has data on several variables that simultaneously impact quality or the yield of the process. In addition, two or more variables can 'interact', so that even if each variable is within its specification limits, the product may not function. These situations make it imperative to monitor more than one variable at the same time.

Many authors have investigated methods of monitoring multivariate continuous data. In 1947, Hotelling [1] developed his multivariate 'T^2' statistic for quality control purposes. Multivariate generalizations of the CUSUM procedure have been studied by Woodall and Ncube [2] and Crosier [3]. Lowry *et al.* [4] developed and investigated multivariate exponentially weighted moving averages to identify quality problems. The use of multivariate exponentially weighted moving averages in monitoring multivariate data have been enhanced by Runger *et al.* [5]. Monitoring principal components of multivariate data has been studied by Mastrangelo *et al.* [6]. We propose developing empirical Bayesian techniques to monitor multivariate continuous data.

Using a Bayesian approach to monitor process data is not entirely new. In the univariate setting, Sturm *et al.* [7] developed empirical Bayesian techniques to monitor continuous data. For count data, Yousry *et al.* [8] used a binomial model with a beta prior to monitor yield and defect data. These techniques were found to be very useful in industrial settings. In the multivariate framework, Jain *et al.* [9] described a Bayesian approach to multivariate quality control. In their paper, they showed that their multivariate control chart procedure was better at identifying out-of-control processes than existing procedures.

In this paper, we describe the theory underlying an empirical Bayesian approach to monitoring multivariate continuous data, generalizing the approaches developed by Jain *et al.* [9] and Sturm *et al.* [7]. Using this theory, we develop methods to implement the empirical Bayes process monitoring for multivariate normal data. In addition, discussions on improving estimation of the process parameters and tips on how to implement the empirical Bayes technique in an industrial environment will be provided. Lastly, an example to illustrate the use of the empirical Bayes process monitoring is given.

2. EMPIRICAL BAYESIAN THEORY FOR MULTIVARIATE PROCESS CONTROL

Suppose that we want to monitor p process characteristics simultaneously. Define the value of the process characteristics at time t to be $\mathbf{X}_t$. This is a vector of observations, each an observation of one of the p process characteristics. Because our multivariate observation $\mathbf{X}_t$ has sampling error, assume that, at a given time t, it is normally distributed about a mean vector $\boldsymbol{\mu}_t$ with variance–covariance matrix $\boldsymbol{\Sigma}$. The

*Correspondence to: Carol J. Feltz, Division of Statistics, Northern Illinois University, DeKalb, IL 60115 USA.
Email: feltz@math.niu.edu

Received 12 July 2000
Revised 5 January 2001

probability density function (pdf) for $\mathbf{X}_t$ given $\boldsymbol{\mu}_t$ is

$$g_1(\mathbf{X}_t|\boldsymbol{\mu}_t) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|} \times \exp\left[-\frac{1}{2}(\mathbf{X}_t - \boldsymbol{\mu}_t)'\boldsymbol{\Sigma}^{-1}(\mathbf{X}_t - \boldsymbol{\mu}_t)\right] \quad (1)$$

where $\boldsymbol{\mu}_t$ is the average process response at time t and $\boldsymbol{\Sigma}$ is an unknown non-negative matrix. Although we allow the process average to change over time, we assume that the sampling variability $\boldsymbol{\Sigma}$ is constant.

To model the process average's changes over time, assume that $\boldsymbol{\mu}_t$ is distributed as a multivariate normal with mean vector $\boldsymbol{\mu}$ and variance–covariance matrix $\mathbf{G}$. The probability density function for $\boldsymbol{\mu}_t$ is

$$g_2(\boldsymbol{\mu}_t|\boldsymbol{\mu}) = \frac{1}{(2\pi)^{p/2}|\mathbf{G}|} \times \exp\left[-\frac{1}{2}(\boldsymbol{\mu}_t - \boldsymbol{\mu})'\mathbf{G}^{-1}(\boldsymbol{\mu}_t - \boldsymbol{\mu})\right] \quad (2)$$

Notice that the above model allows two sources of variability: (i) the sampling variability, indicating the amount of spread present between samples if the process is not changing; and (ii) process variability, indicating the amount of variability due to process changes over time. This generalizes the approach of Jain *et al.* [9]. In their 1993 paper, Jain and his coauthors assumed that the underlying process variability is the same as the sampling variability, i.e. $\mathbf{G} = \boldsymbol{\Sigma}$. In our work in the electronics industry, we found that by allowing the process to have its own variability, we gained information about the process behavior as well as information about how the sample behaves around the process mean. Allowing the process variability to be different from the sampling variability will make our estimation of process parameters more difficult, but also more rewarding.

Our problem is then to estimate the location of the process at time t, using the prior information of where the process was at time $t-1$ and the current observation $\mathbf{X}_t$. Using Bayes Theorem, the pdf of the conditional distribution of $\boldsymbol{\mu}_t$ given $\mathbf{X}_t$ is

$$f(\boldsymbol{\mu}_t|\mathbf{X}_t) = \frac{g_1(\mathbf{X}_t|\boldsymbol{\mu}_t)g_2(\boldsymbol{\mu}_t)}{\int g_1(\mathbf{X}_t|\boldsymbol{\mu}_t)g_2(\boldsymbol{\mu}_t)d\boldsymbol{\mu}_t} \quad (3)$$

This distribution is known as a posterior distribution in Bayesian terminology, with equation (2) called the prior distribution of $\boldsymbol{\mu}_t$. It is easy to derive that this posterior distribution, $f(\boldsymbol{\mu}_t|\mathbf{X}_t)$, is a multivariate normal distribution with mean

$$E(\boldsymbol{\mu}_t|\mathbf{X}_t) = \mathbf{X}_t - \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{G})^{-1}(\mathbf{X}_t - \boldsymbol{\mu}) \quad (4)$$

and variance–covariance matrix

$$COV(\boldsymbol{\mu}_t|\mathbf{X}_t) = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{G})^{-1}\boldsymbol{\Sigma} \quad (5)$$

Letting $\mathbf{W} = \boldsymbol{\Sigma}(\boldsymbol{\Sigma} + \mathbf{G})^{-1}$ be a weighting matrix, the posterior mean (4) can be rewritten as

$$\mathbf{W}\boldsymbol{\mu} + (\mathbf{I} - \mathbf{W})\mathbf{X}_t \quad (6)$$

Intuitively, it is difficult to see what effect $\mathbf{W}$ has on the posterior mean. However, if $\boldsymbol{\Sigma}$ and $\mathbf{G}$ are diagonal matrices (i.e. there is no correlation between process characteristics) then $\mathbf{W}$ is also a diagonal matrix, and it is easier to see the effect of $\mathbf{W}$. Those components of $\mathbf{W}$ which are large indicate that the corresponding process characteristics have a large sampling variation, which will pull the estimate of the process mean for those process characteristics at time t toward their prior means. Similarly, components of $\mathbf{W}$ which are small indicate a large process variation compared to the sampling variability will pull the estimate of the process mean at time t toward the current observation, $\mathbf{X}_t$. Usually in practice, there is correlation between process characteristics and the effect of $\mathbf{W}$ is not so clear.

With the components of the model in place, we can now focus on using the process data to estimate the different components.

3. ESTIMATING THE PROCESS PARAMETERS

In estimating the process parameters, consider first the overall process mean vector, $\boldsymbol{\mu}$.

3.1. *Estimating $\boldsymbol{\mu}$ (the overall process mean)*

To estimate $\boldsymbol{\mu}$, note that expected value of $\mathbf{X}_t$ is

$$E(\mathbf{X}_t) = E_{\boldsymbol{\mu}_t}E_{\mathbf{X}_t}(\mathbf{X}_t|\boldsymbol{\mu}_t) = \boldsymbol{\mu} \quad (7)$$

for any time t. Because each $\mathbf{X}_t$ is an unbiased estimate of $\boldsymbol{\mu}$, a reasonable estimator of $\boldsymbol{\mu}$, $\hat{\boldsymbol{\mu}}$, is the overall average of the $\mathbf{X}_t$'s. Assume that we have observed n data vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n$. Then

$$\hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{t=1}^{n}\mathbf{X}_t \quad (8)$$

Next we must estimate the sampling and process variability. If one had multiple independent observations at time t, one could estimate the sampling variance, $\boldsymbol{\Sigma}$, using standard formulas. However, in our experience, replications are usually not available.

To circumvent this obstacle, we need to assume that the $\boldsymbol{\mu}_t$ are fairly stable, i.e. the process mean is not shifting uncontrollably all of the time. Assuming that the process average at time t remains relatively constant over short time intervals, the correlation between $\boldsymbol{\mu}_t$ and $\boldsymbol{\mu}_{t+1}$ is close to 1 when the time intervals of taking data are small. If there is very little time lag between the $\mathbf{X}_t$'s, consecutive $\mathbf{X}_t$'s can be thought of as independent random variables from the same distribution. Under this assumption, one estimate of $\boldsymbol{\Sigma}$ is

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{2n}\sum_{t=1}^{n}(\mathbf{X}_t - \mathbf{X}_{t+1})(\mathbf{X}_t - \mathbf{X}_{t+1})' \quad (9)$$

3.2. *Estimating* G *(the process variance–covariance matrix)*

In estimating $\mathbf{G}$, we first note that the overall variance of $\mathbf{X}_t$ can be expressed by the variance–covariance matrix

$$\mathbf{V} = E(\mathbf{X}_t - \boldsymbol{\mu}_t)(\mathbf{X}_t - \boldsymbol{\mu}_t)' = \boldsymbol{\Sigma} + \mathbf{G} \quad (10)$$

Using $\hat{\boldsymbol{\mu}}$ in equation (8) as an estimator for $\boldsymbol{\mu}$, then the overall variance of $\mathbf{X}_t$ can be estimated in the usual way by

$$\hat{\mathbf{V}} = \frac{1}{n}\sum_{t=1}^{n}(\mathbf{X}_t - \hat{\boldsymbol{\mu}})(\mathbf{X}_t - \hat{\boldsymbol{\mu}})' \quad (11)$$

With our estimator of $\boldsymbol{\Sigma}$ and our estimator of the overall variation of $\mathbf{X}_t$, $\mathbf{V}$, we can estimate $\mathbf{G}$ by subtraction. That is, $\hat{\mathbf{G}} = \hat{\mathbf{V}} - \hat{\boldsymbol{\Sigma}}$.

The reader may note that we are using the maximum likelihood estimate of the variance and covariance, where the denominator is the number of data points instead of the unbiased estimator in which one is subtracted from the number of data points in the denominator. This will aid us in the weighted case to be presented in Section 4.

Although we have all the key estimators in place, implementing Bayesian process monitoring in a manufacturing environment requires some special features.

4. IMPLEMENTING EMPIRICAL BAYESIAN PROCESS CONTROL

When simulating the behavior of their Bayesian multivariate process monitor, Jain *et al.* [9] found that their estimate of the prior mean contained all previous data points from 1 to $n-1$. For many processes this could be a disaster, as the process mean would soon be weighted down with 'old' data. Sturm *et al.* [7] found a way around this problem in the univariate case by giving the current data more weight than older data. By using a weighting factor that is less than one, an exponentially weighted moving average is created. By incorporating the weighting into the empirical Bayes approach, we maintain our distributional structure, so that we can partition the variability into the sampling and process variability as well as using the distribution to identify shifts in the process mean. An added benefit is that we never have to delete data from the system. Old data is automatically weighted out.

To incorporate weighting into our estimators, assume the $\mathbf{X}_t$'s are ordered in time $(\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_T)$ so that $\mathbf{X}_T$ is the most recent observation. Then let

$$\hat{\boldsymbol{\mu}} = \frac{\sum_{t=1}^{T}\lambda^{T-t}\mathbf{X}_t}{\sum_{t=1}^{T}\lambda^{T-t}} \quad (12)$$

$$\hat{\boldsymbol{\Sigma}} = \frac{\sum_{t=1}^{T}\lambda^{T-t}(\mathbf{X}_t - \mathbf{X}_{t-1})(\mathbf{X}_t - \mathbf{X}_{t-1})'}{2\sum_{t=1}^{T}\lambda^{T-t}} \quad (13)$$

$$\hat{\mathbf{V}} = \frac{\sum_{t=1}^{T}\lambda^{T-t}(\mathbf{X}_t - \hat{\boldsymbol{\mu}})(\mathbf{X}_t - \hat{\boldsymbol{\mu}})'}{\sum_{t=1}^{T}\lambda^{T-t}} \quad (14)$$

Here λ defines the weight given to each time period, where λ is an arbitrary number (usually $0.80 < \lambda < 1.0$). The choice of λ depends on the process. If data taken 50, 100 or 200 observations ago are no longer relevant to where the process is currently, then the λ should be chosen appropriately. To weight away data that is 50 observations ago, λ is chosen to be 0.832. Similarly, to weight away data that is 100 or 200 observations ago, $\lambda = 0.912$ and $\lambda = 0.955$, respectively.

To make the analysis more computationally efficient, the above weighted estimators can be written in the form of recursive equations. Then one need only take the previous estimate of the process mean, sampling variance–covariance and process variance–covariance matrix and update with the current observation. The recursive equations are given below.

Given the observation vectors $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_t$, for the ith variable, denote the sample mean by $\bar{x}_{i,t}$ and the sample variance by $\hat{v}_{i,t}^2$, and the covariance of the ith and jth variables by $\hat{v}_{ij,t}$. $\hat{v}_{ij,t}$ is the (i, j)th element of the matrix $\hat{\mathbf{V}}$ at the time t and $\hat{v}_{i,t}^2 = \hat{v}_{ii,t}$. Without the weighting factor, recursive estimation formulas of the mean and overall variance at time t for the ith variable are given by:

$$\bar{x}_{i,t} = \bar{x}_{i,t-1} + \frac{x_{i,t} - \bar{x}_{i,t-1}}{t} \quad (15)$$

and

$$\hat{v}_{i,t}^2 = \frac{(t-1)\hat{v}_{i,t-1}^2}{t} + \frac{(t-1)(x_{i,t} - \bar{x}_{i,t-1})^2}{t^2} \quad (16)$$

The covariance of the ith and jth variables at time t are:

$$\hat{v}_{ij,t} = \frac{(t-1)\hat{v}_{ij,t-1}}{t} + \frac{(t-1)(x_{i,t}-\bar{x}_{i,t-1})(x_{j,t}-\bar{x}_{j,t-1})}{t^2} \quad (17)$$

Also, for the (i, j)th element of the matrix $\hat{\boldsymbol{\Sigma}}$ at time t, denoted by $\hat{\sigma}_{ij,t}$, the recursive formula is given by

$$\hat{\sigma}_{ij,t} = \frac{2(t-1)\hat{\sigma}_{ij,t-1} + (x_{i,t}-x_{i,t-1})(x_{j,t}-x_{j,t-1})}{2t} \quad (18)$$

In the weighted case, the recursive formulas are more complicated. For the mean computation, we simply multiply the previous average with the old sum of the weights, multiply by the weighting factor, add the new data point, and divide by the new weight. Operations are similar for computing the variance and covariance. Denote the sum of the weights at time t by $w_t = \sum_{k=1}^{t} \lambda^{t-k}$. Then the recursive formulas for the mean, variance and covariance are

$$\bar{x}_{i,t} = \frac{\lambda w_{t-1}\bar{x}_{i,t-1} + x_{i,t}}{w_t} \quad (19)$$

$$\hat{v}_{i,t}^2 = \frac{\lambda w_{t-1}(\hat{v}_{i,t-1}^2 + (\bar{x}_{i,t}-\bar{x}_{i,t-1})^2) + (x_{i,t}-\bar{x}_{i,t})^2}{w_t} \quad (20)$$

$$\hat{v}_{ij,t} = \frac{\lambda w_{t-1}(\hat{v}_{ij,t-1} + (\bar{x}_{i,t}-\bar{x}_{i,t-1})(\bar{x}_{j,t}-\bar{x}_{j,t-1}))}{w_t} + \frac{(x_{i,t}-\bar{x}_{i,t})(x_{j,t}-\bar{x}_{j,t})}{w_t} \quad (21)$$

The (i, j)th element of $\hat{\boldsymbol{\Sigma}}$ at time t can be recursively computed by

$$\hat{\sigma}_{ij,t} = \frac{2\lambda w_{t-1}\hat{\sigma}_{ij,t-1} + (x_{i,t}-x_{i,t-1})(x_{j,t}-x_{j,t-1})}{2w_t} \quad (22)$$

To identify process drifting, the decision rule developed by Jain *et al.* [9] will be used. Let $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ be the mean and covariance matrix obtained from a set of in-control data. Then, denoting our posterior mean as

$$\begin{aligned}\boldsymbol{\mu}_t^* &= \mathbf{X}_t - \boldsymbol{\Sigma}(\boldsymbol{\Sigma}+\mathbf{G})^{-1}(\mathbf{X}_t - \boldsymbol{\mu}) \\ &= \mathbf{X}_t - \boldsymbol{\Sigma}\mathbf{V}^{-1}(\mathbf{X}_t - \boldsymbol{\mu}) \end{aligned} \quad (23)$$

the test statistic is the quadratic form

$$B_t = (\boldsymbol{\mu}_t^* - \boldsymbol{\mu}^*)'(\boldsymbol{\Sigma}^*)^{-1}(\boldsymbol{\mu}_t^* - \boldsymbol{\mu}^*). \quad (24)$$

The test statistic is then compared to a critical value, $C_p = \chi^2{}_{p,0.9973}$, the 99.73th percentile of a chi-square with p degrees of freedom, which corresponds to the regular 3-sigma control chart limits and where p is the number of variables being monitored simultaneously. In the empirical Bayes setting, all the prior parameters in equation (23) and (24) are replaced by the corresponding estimates in equations (12), (13) and (14), which can be computed recursively by equations (15) to (17) for the unweighted case, and by equations (18) to (22) for the weighted case.

An example of the technique is given in the following section.

Table 1. Variables and their target means

Characteristic	Variable name	Target value
Outside diameter	X_1	90.0
Width	X_2	19.7
Seat height	X_3	25.2
Seat angle	X_4	0.48
Seat concentricity	X_5	4.52

Table 2. The sample covariance matrix $\hat{\mathbf{V}}$

Variable	X_1	X_2	X_3	X_4	X_5
X_1	8.990	0.137	0.223	0.067	−0.055
X_2	0.137	0.830	−0.122	−0.030	−0.050
X_3	0.223	−0.122	2.220	0.589	0.041
X_4	0.067	−0.030	0.589	0.310	0.004
X_5	−0.055	−0.050	0.041	0.004	0.830

5. AN ILLUSTRATIVE EXAMPLE

The example presented in Jain *et al.* [9] will be used to illustrate the multivariate empirical Bayesian technique. In this example, data from a machining operation for valve seat inserts is presented. The variables and their target means are given in Table 1.

Using all the data, Jain *et al.* [9] computed the sample covariance matrix shown in Table 2.

Jain *et al.* [9] then used ten observations to illustrate their multivariate Bayesian procedure. Because the covariance matrix above includes all the data, over a span of time, this covariance matrix would correspond to our $\hat{\mathbf{V}} = \hat{\boldsymbol{\Sigma}} + \hat{\mathbf{G}}$. Without additional information, let $\hat{\boldsymbol{\Sigma}} = \hat{\mathbf{G}} = \hat{\mathbf{V}}/2$. Similarly, let the target values be the initial prior estimate of the mean. Let the target values also be the $\boldsymbol{\mu}^*$ in the test statistic, B_1. Then letting our weighting factor, $\lambda = 0.9$, we can monitor the ten observations and use the results to get a more

Table 3. Results from ten observations

Observation	X_1	X_2	X_3	X_4	X_5	Posterior mean	B_i
Target						90.0, 19.7, 25.2, 0.48, 4.52	
1	93	20.0	24.0	0.0	5.0	91.6, 19.9, 24.6, 0.22, 4.77	1.3
2	90	18.0	25.0	0.0	5.0	90.6, 18.9, 24.9, 0.13, 4.87	3.1
3	90	19.0	26.0	1.0	6.0	90.1, 19.1, 25.5, 0.65, 5.42	3.0
4	94	18.0	26.0	1.0	3.0	92.1, 18.8, 25.8, 0.91, 4.35	4.5
5	91	20.0	27.0	1.0	6.0	91.9, 19.1, 26.1, 0.76, 4.46	2.4
6	88	20.0	25.0	0.0	6.0	89.9, 19.5, 25.1, 0.26, 4.91	0.9
7	95	21.0	25.0	0.0	5.0	91.9, 20.3, 25.1, 0.11, 5.12	4.0
8	91	20.0	28.0	2.0	5.0	91.9, 19.9, 26.3, 0.91, 4.64	2.2
9	93	19.0	25.0	1.0	4.0	91.4, 19.4, 25.6, 0.94, 4.74	2.6
10	92	21.0	25.0	0.0	3.0	90.8, 20.2, 25.4, 0.61, 4.11	1.2

Table 4. Estimate of $\mathbf{\Sigma}$

Variable	X_1	X_2	X_3	X_4	X_5
X_1	5.555	0.114	−0.743	−0.282	−0.861
X_2	0.114	0.723	0.025	−0.120	0.183
X_3	−0.743	0.025	1.39	0.581	0.243
X_4	−0.282	−0.120	0.581	0.372	0.115
X_5	−0.861	0.183	0.243	0.115	0.809

Table 5. Estimate of $\mathbf{G}$

Variable	X_1	X_2	X_3	X_4	X_5
X_1	0.519	0.094	0.711	0.327	0.196
X_2	0.094	0.166	−0.110	−0.020	−0.199
X_3	0.711	−0.110	0.181	0.037	−0.030
X_4	0.327	−0.020	0.037	0.026	−0.062
X_5	0.196	−0.199	−0.030	−0.062	0.211

up-to-date estimate of the process mean $\boldsymbol{\mu}_t$, $\hat{\mathbf{\Sigma}}$ and $\hat{\mathbf{G}}$ as shown in Table 3. (We chose the weighting factor to be $\lambda = 0.9$ to see how much the sampling covariance matrix $\hat{\mathbf{\Sigma}}$ and the process covariance matrix $\hat{\mathbf{G}}$ differ at the end of the eight runs. A larger λ moves the estimates more slowly from the prior distribution.)

The cut-off point for the test statistic is $\chi^2{}_{5,0.9973} = 18.2$, so none of the observations were close to being significant. After the ten observations, our estimates of $\mathbf{\Sigma}$ and $\mathbf{G}$ are shown in Table 4 and 5 respectively.

Note that the process variance for each of the variables is substantially smaller than the sampling variance. When implementing the empirical Bayes process monitoring in factories, we usually found the process variance was much smaller than the sampling variance.

6. CONCLUSION

This paper develops a method for monitoring continuous multivariate data using an empirical Bayes model. The empirical Bayes model gives us process information as well as sampling information about the process parameters being monitored. By putting a multivariate normal structure on the data, we can identify correlations between the variables that may have an effect on the quality of our product. By estimating both the process variation as well as the sampling variation, we get a more thorough understanding of our process.

APPENDIX A. SAS PROGRAM FOR MULTIVARIATE EMPIRICAL BAYES PROCESS MONITORING

```
     proc iml;
infile 'data.dat';
v={
8.990  0.137  0.223  0.067  -0.055,
0.137  0.830  -0.122  -0.030  -0.050,
0.223  -0.122  2.220  0.589  0.041,
0.067  -0.030  0.589  0.310  0.004,
-0.055  -0.050  0.041  0.004  0.830};
testsig=v/2.0;
sig=v/2.0;
lambda=.9;
w=1/(1.0-lambda);
gamma=sig;
print sig;
print gamma;
priormu={90.0, 19.7, 25.2, 0.48, 4.52};
x=priormu;
oldxbar=priormu;
oldx=priormu;
do data;
input x1 x2 x3 x4 x5;
x[1]=x1;
x[2]=x2;
```

```
x[3]=x3;
x[4]=x4;
x[5]=x5;
print x;
xbar=w*lambda*oldxbar+x;
xbar=xbar/w;
v=w*lambda*v+(xbar-oldxbar)*(xbar-oldxbar)
    '+(x-xbar)*(x-xbar)';
v=v/w;
sig=2.0*w*lambda*sig+(x-oldx)*(x-oldx)';
sig=sig/(2.0*w);
postmu=x-sig*(v**-1)*(x-xbar);
b1=(postmu-priormu)'*(testsig**-1)
    *(postmu-priormu);
oldx=x;
oldxbar=xbar;
print xbar;
print postmu;
print b1;
end;
print v;
print sig;
gamma=v-sig;
print gamma;
```

REFERENCES

1. Hotelling H. Multivariable quality control—illustrated by the air testing of sample bombsights. *Techniques of Statistical Analysis*, Eisenhart C, Hastay MW, Wallis WA (eds.). McGraw Hill: New York, 1947; 111–184.
2. Woodall WH, Ncube M. Multivariate CUSUM quality-control procedures. *Technometrics* 1985; **27**:285–292.
3. Crosier RB. Multivariate generalizations of cumulative sum quality-control schemes. *Technometrics* 1988; **30**:291–303.
4. Lowry CA, Woodall WH, Champ CW, Rigdon SE. A multivariate exponentially weighted moving average control chart. *Technometrics* 1992; **34**:46–53.
5. Runger GC, Keats JB, Montgomery DC, Scranton RD. Improving the performance of the multivariate exponentially weighted moving average control chart. *Quality and Reliability International* 1999; **15**:161–166.
6. Mastrangelo CM, Runger GC, Montgomery DC. Statistical process monitoring with principal components. *Quality and Reliability International* 1996; **12**:203–210.
7. Sturm G, Feltz C, Yousry M. An empirical Bayes strategy for analysing manufacturing data in real time. *Quality and Reliability International* 1991; **7**:159–168.
8. Yousry M, Sturm G, Feltz C, Noorossana R. Process monitoring in real time: empirical Bayes approach—discrete case. *Quality and Reliability International* 1991; **7**:123–132.
9. Jain K, Alt F, Grimshaw SD. Multivariate quality control—a Bayesian approach. *ASQC Quality Congress Transactions—Boston*, 1993; 645–651.

Authors' biographies:

Carol J. Feltz is an Associate Professor in the Division of Statistics at Northern Illinois University. After receiving her PhD in Statistics from the University of Missouri-Columbia, she worked at AT&T's Engineering Research Center in Princeton, New Jersey. While there, she developed methods of improving quality at AT&T's manufacturing facilities. In 1989, she moved to her current position. Her research interests include developing statistical methods for quality improvement, life-testing and goodness-of-fit tests.

Jyh-Jen Horng Shiau received her BS in mathematics from National Taiwan University, MS in applied mathematics from the University of Maryland Baltimore County, MS in computer science and PhD in statistics from the University of Wisconsin-Madison. She has taught at Southern Methodist University, the University of Missouri at Columbia and National Tsing Hua University. She worked for the Engineering Research Center of AT&T Bell Labs before she moved to Hsinchu, Taiwan where she is an Associate Professor in the Institute of Statistics at National Chiao Tung University. Her interests include industrial statistics and non-parametric function estimation. She is a member of the Institute of Mathematical Statistics.