國 立 交 通 大 學

統 計 學 研 究 所

碩 士 論 文

聯合多維管制圖
Combined Multivariate Control Chart

研 究 生：吳采玲 (Cai-Ling Wu)
指導教授：陳鄰安 博士 (Dr. Lin-An Chen)

中華民國九十七年六月

聯合多維管制圖
Combined Multivariate Control Chart

研 究 生：吳采玲　　Student：Cai-Ling Wu
指導教授：陳鄰安 博士　　Advisor：Dr. Lin-An Chen

國 立 交 通 大 學
統計學研究所
碩 士 論 文

A Thesis
Submitted to Institute of Statistics
College of Science
National Chiao Tung University
In Partial Fulfillment of the Requirements
For the Degree of
Master
In
Statistics
June 2008
Hsinchu, Taiwan, Republic of China

中華民國九十七年六月

Combined Multivariate Control Chart

student：Cai-Ling Wu Advisor：Dr. Lin-An Chen

Institute of Statistics
National Chiao Tung University

ABSTRACT

A new combined control chart for multivariate distribution is proposed. This control chart may be applied on any distribution that its joint probability density function in terms of a random sample is known its distribution while the existed multivariate control charts are generally designed only for multivariate normal distribution. A comparison of this chart with a moving average control chart by Chen, Cheng and Xie (2005) for bivariate normal distribution shows that it is very competitive. When the joint probability density function is not known in its distribution, an approximate combined chart is proposed. Studies of ARLs for these two charts are performed and displayed.

聯合多維管制圖

學生：吳采玲　　　　　　　　　指導教授：陳鄰安博士

國立交通大學統計學研究所 碩士班

摘　　　要

本文提出一種新的聯合多維管制圖，只要隨機變數的聯合機率密度函數已知即可應用此管制圖。這解決了當今現存的多維管制圖普遍都只應用在多維常態分配上的不足。此管制圖與2005年 Chen, Cheng and Xie 提出的二維常態分配平均移動管制圖具有相當的競爭性。當隨機變數的聯合機率密度函數為未知時，可利用模擬方法建造近似聯合多維管制圖。本文也提出有關於這兩個聯合多維管制圖的平均連串長度之研究。

誌 謝

誠摯感謝指導教授陳鄰安老師，老師細心指導與諄諄教誨且不時的與我討論並指點我正確的方向，使我獲益匪淺。

謝謝口試委員許文郁老師、黃榮臣老師及洪慧念老師給予本論文的指正與建議，使整體論文更加充實。

最後，感謝一直在背後支持我，不時地給予我精神上鼓勵的家人與朋友。

謹將我的成果與喜悅獻給所有關心我的家人、老師與朋友。

采玲 於交大統計研究所

中華民國九十七年六月

Contents

Combined Multivariate Control Chart

Abstract

A new combined control chart for multivariate distribution is proposed. This control chart may be applied on any distribution that its joint probability density function in terms of a random sample is known its distribution where the existed multivariate control charts are generally designed only for multivariate normal distribution. A comparison of this chart with a moving average control chart by Chen, Cheng and Xie (2005) for bivariate normal distribution shows that it is very competitive. When the joint probability density function is not known in its distribution, an approximate combined chart is proposed. Studies of ARL's for these two charts are performed and displayed.

1. Introduction

We say that a process is in statistical control if the process distribution of quality characteristic of the product is constant over time and if there is change over time, the process is said to be out of control. A control chart provides the most popular technique for monitoring the process.

Sometimes we encounter process that involves numerous quality characteristics of interest. While control charts, one for each characteristic, may be constructed, using the multivariate control chart (Hotelling T^2) to avoid the incorrect (probability) limits has been general application and extensive research on it (see Mason, et al. (1997) and Wierda (1994) for a review).

However, in this process involving numerous characteristics, there can be another situation that causes the same problem of incorrect control limits. In some manufacturing processes, we are frequently dealing with these characteristics whose distribution not only involving some location parameters but also involving scale parameters. A common method to deal with this problem is to design a statistical process control scheme that has a probability α of type one error and to apply a multivariate control chart separately to location parameters and to scale parameters. However, incorrect control limits may also occur for this scheme. For example, when we consider a

Typeset by $\mathcal{AMS}$-TEX

multivariate normal data, the multivariate control chart for location classically uses the Hotelling T^2 that is composed with the sample mean $\bar{X}$ and sample covariance matrix and the control chart for scale uses S. These two charts for location and scale are not independent so that two test statistics are not independent, Hence, incorrect control limits occurs for this scheme. We note that univariate normal distribution by using $\bar{X}$ chart and R (or S) chart is free of this problem.

This problem has been solved in univariate control chart. Under univariate normal processes, the $\bar{X}$ chart and S-chart are generally used to monitor the process mean and process variance, respectively. Many authors have proposed to combine charts designed for different parameters into a single chart which can simultaneously monitor various parameters. In the case of univariate normal processes, for example, Repco (1986), Van Nuland (1992), Chao and Cheng (1996), Spiring and Cheng (1998), and Yeh and Lin (2002), these combined charts are designed to simultaneously monitor the process mean and variance. A good review of the existing works can be found in Cheng and Thaga (2005). Our interest in this paper is to generalize the combined chart for univariate problem to multivariate model.

2. Combined Control Chart

Let X be a p-variate random vector with sample space $\Re^p$ represents the process characteristics of the product. We consider the process with the following elements for our concern:

(a_1) The process distribution has a probability density function $f(x, \Theta_1, ..., \Theta_k)$, $x \in \Re^p$ where $\Theta_1, ..., \Theta_k$ are unknown matrices, vectors or constants.

(a_2) A training sample that represents an in-control data of m samples of sample size n so that a set of estimates $\hat{\Theta}_1, ..., \hat{\Theta}_k$ is available.

(a_3) A random sample $X_1, ..., X_n$ that is drawn from the distribution $f(x, \Theta_1, ..., \Theta_k)$ for constructing a test statistic.

Woodall (2000) pointed out that a statistical quality control is to test the hypothesis for that the process distribution is assumed to be known along with values of the in-control parameters. With estimates $\hat{\Theta}_1, ..., \hat{\Theta}_k$ computed from the in-control observations, it is appropriate to interprete a process is in-control when the assumption of the process distribution in the following as

$$H_0 : f(x, \Theta_1, ..., \Theta_k) = f(x, \hat{\Theta}_1, ..., \hat{\Theta}_k) \tag{2.1}$$

is true.

The joint probability density function (pdf) of a ramdom sample $X_1, ..., X_n$ when H_0 is true is

$$L(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k) = \pi_{i=1}^{n} f(x_i, \hat{\Theta}_1, ..., \hat{\Theta}_k).$$

Consider that $(x_1^a, ..., x_n^a)$ and $(x_1^b, ..., x_n^b)$ are two observations and we suppose that $(x_1^a, ..., x_n^a)$ is already classified as an in-control sample point. Then there is no reason not to classify $(x_1^b, ..., x_n^b)$ also as an in-control sample point if $L(x_1^a, ..., x_n^a, \hat{\Theta}_1, ..., \hat{\Theta}_k) \leq L(x_1^b, ..., x_n^b, \hat{\Theta}_1, ..., \hat{\Theta}_k)$ is true. Suppose that there is a constant $\ell(\hat{\Theta}_1, ..., \hat{\Theta}_k)$ satisfying

$$1 - \alpha = P_{\hat{\Theta}_1,...,\hat{\Theta}_k}(L(X_1, ..., X_n, \hat{\Theta}_1, ..., \hat{\Theta}_k) \geq \ell(\hat{\Theta}_1, ..., \hat{\Theta}_k)) \tag{2.2}$$

where $L(X_1, ..., X_n, \hat{\Theta}_1, ..., \hat{\Theta}_k)$ serves as a test statistic for testing hypothesis H_0. We then may set a control chart as

$$LCL = \ell(\hat{\Theta}_1, ..., \hat{\Theta}_k)$$
$$\text{Test statistic function: } L(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k).$$

If sample points $x_1, ..., x_n$ from the multivariate distribution has a value $L(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k)$ lying above the lower limit $\ell(\hat{\Theta}_1, ..., \hat{\Theta}_k)$ and does not exhibit any systematic pattern, we may say that the process is in statistical control at the level $1 - \alpha$. Classically we choose $1 - \alpha = 0.9, 0.95$ or 0.9973.

There is a situation that the test statistic is in a simpler form. Suppose that there is a statistic $T = t(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k)$ and a constant $t^*(\hat{\Theta}_1, ..., \hat{\Theta}_k)$ such that

$$\begin{array}{c} L(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k) \leq \ell(\hat{\Theta}_1, ..., \hat{\Theta}_k) \text{ if and only if} \\ t(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k) \leq t^*(\hat{\Theta}_1, ..., \hat{\Theta}_k), \text{ or } \geq t^*(\hat{\Theta}_1, ..., \hat{\Theta}_k) \end{array} \tag{2.3}$$

If inequality $\leq$ holds in (2.3), an alternative combined chart is

$$LCL = t^*(\hat{\Theta}_1, ..., \hat{\Theta}_k)$$
$$\text{Test statistic function} : T = t(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k).$$

On the other hand, if $\geq$ holds in (2.3), an alternative combined chart is

$$UCL = t^*(\hat{\Theta}_1, ..., \hat{\Theta}_k)$$
$$\text{Test statistic function} : T = t(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k).$$

In the chart with lower control limit, if sample points $x_1, ..., x_n$ from the multivariate distribution has a value $t(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k)$ lying below LCL, we may claim that there is potential assignable cause occurring. On the other hand, in the chart with upper control limit, if sample points $x_1, ..., x_n$ has a value $t(x_1, ..., x_n, \hat{\Theta}_1, ..., \hat{\Theta}_k)$ lying above UCL, we may claim that there is potential assignable cause occurring.

3. Combined Multivariate Normal Control Chart

Suppose that the vector X of quality characteristics is multivariate normal $N_p(\mu, \Sigma)$ where p-vector μ and $p \times p$ covariance matrix Σ are unknown. With a random sample $X_1, ..., X_n$, the joint pdf is

$$L(x_1, ..., x_n, \mu, \Sigma) = \frac{1}{(2\pi)^{pn/2}|\Sigma|^{n/2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)}.$$

Suppose that we have a training sample $x_{ij}, i = 1, ..., n, j = 1, ..., m$ of m groups of size n from an in-control multivariate normal distribution. We can then calculate m sample $p \times 1$ means $\bar{x}_j$ and m sample $p \times p$ covariance matrices s_j, $j = 1, ..., m$; as well as their averages $\bar{\bar{x}} = \frac{1}{m}\sum_{j=1}^{m}\bar{x}_j$ and $\bar{s} = \frac{1}{m}\sum_{j=1}^{m}s_j$. The appropriate hypothesis for this in control process distribution is

$$H_0 : X \sim N_p(\bar{\bar{x}}, \bar{s}).$$

When we consider that $\bar{\bar{x}}$ and $\bar{s}$ as the true mean and covariance matrix, we have $\sum_{i=1}^{n}(X_i - \bar{\bar{x}})'\bar{s}^{-1}(X_i - \bar{\bar{x}}) \sim \chi^2(pn)$. The inequality

$$L(X_1, ..., X_n, \bar{\bar{x}}, \bar{s}) \geq \ell(\bar{\bar{x}}, \bar{s})$$

subjected to $1 - \alpha = P_{\bar{\bar{x}},\bar{s}}(L(X_1, ..., X_n, \bar{\bar{x}}, \bar{s}) \geq \ell(\bar{\bar{x}}, \bar{s})$ yields $\ell(\bar{\bar{x}}, \bar{s}) = \frac{1}{(2\pi)^{pn/2}|\bar{s}|^{n/2}} e^{-\frac{\chi^2_{1-\alpha}}{2}}$ where $\chi^2_{1-\alpha}$ satisfies $1 - \alpha = P(\chi^2(pn) \leq \chi^2_{1-\alpha})$. We then have a new control chart when the process distribution is multivariate normal as

$$LCL = \frac{1}{(2\pi)^{pn/2}|\bar{s}|^{n/2}} e^{-\frac{\chi^2_{1-\alpha}}{2}}$$
$$\text{Test statistic function: } L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s}) = \frac{1}{(2\pi)^{pn/2}|\bar{s}|^{n/2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i-\bar{\bar{x}})'\bar{s}^{-1}(x_i-\bar{\bar{x}})}$$

For a given observation $x_1, ..., x_n$, we compare its density value $L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s})$ with the lower control limit LCL. The control chart indicates being out of control if $L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s}) < LCL$.

We denote $\chi_0^2 = \sum_{i=1}^n (X_i - \bar{\bar{x}})' \bar{s}^{-1} (X_i - \bar{\bar{x}})$ which has distribution $\chi^2(pn)$ when H_0 is true. From the relation $L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s}) \leq LCL$ if and only if $\chi_0^2 \geq \chi_{1-\alpha}^2$, we may see that the new control chart is exactly a chi-square control chart as

$$UCL = \chi_{1-\alpha}^2$$
$$LCL = 0$$
$$\text{Test statistic function: } \chi_0^2 = \sum_{i=1}^n (x_i - \bar{\bar{x}})' \bar{s}^{-1} (x_i - \bar{\bar{x}})$$

If a sequence of data so that their χ_0^2 fall within the control limits LCL and UCL, and it does not exhibit any systematic pattern, we say that the process is in statistical control.

Example 1. To monitor manufacturing fabric, it is interesting in factors of single-strand break (a measure of the breaking strength) and the weight of textile fibers (hanks per pound). A data of two characteristics for 20 samples of size $n = 4$ and a study of control charting for this samples have been provided by Amitava (1998). The Hotelling's T^2 chart at level $\alpha = 0.0054$ has been contructed that indicates sample number 9 an out of control point. For this bivariate samples, they also conducted two separate level $\alpha = 0.0054$ $\bar{X}$-charts to monitor the means of single-strand break and weight of textile fibers, however, they claimed that no sample is indicated as out of control point in either chart and then they conclude that multivariate control chart may observe out of control points that couldn't observe by separate single charts. It is interesting to see what may happen to apply the density control chart that we monitor the mean vector but also the dispersion matrix simultaneously.

We first conduct Hotelling's T^2 chart at level $\alpha = 0.01$ resulted the upper control limit $UCL = 9.629$ which indicates that sample numbers $9, 11, 14$ are suspected out of control points since they have T^2 values $15.25, 10.08, 10.66$ lie outside the upper control limit. The density control limit at level 0.01 is $LCL = 4.60681E - 11$. There is no sample point below this lower control limit LCL. That is, sample points showing potential assignable causes due to the T^2 chart are no-longer out of control points when we apply the density control chart. To search a reason for the difference between the performance between the density chart and the Hotelling's T^2 chart, we compute the sample covariance matrices of these 20 samples. One way to evaluate the size of covariance matrices is to evaluate their corresponding determinants. The following listed sequentially the determinants of these sample covariance matrices for comparison:

13.2472,	9,	5.4189,	15.1875,	22.5875,	2.6097,	5.5861,
8.5222,	2.0075,	3.7336,	1.84,	0.51,	25.2075,	4.2075
1.7822,	8.01,	48.5811,	4.0736,	51.6536,	5.84.	

We see that the determinants of sample numbers $9, 11, 14$ are, respectively, 2.0075, 1.84, 4.2075. It is clear that these three samples although have relative large T^2 values, however, their sizes in dispersion are moderate so that the combined performance is not siginificant in terms of density to be claimed as out of control points.

How can we execute on-line quality control based on density control chart. The constructed density control chart is

$$LCL = 4.60681E - 11$$

$$\text{Test statistic function: } L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s}) = 1.06153E - 06$$

$$e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \begin{pmatrix} 82.46 \\ 20.17 \end{pmatrix})' \begin{pmatrix} 7.51 & -0.35 \\ -0.35 & 3.29 \end{pmatrix}^{-1} (x_i - \begin{pmatrix} 82.46 \\ 20.17 \end{pmatrix})}$$

For a given observation $x_1, ..., x_n$, we compare its density value $L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s})$ with the lower control limit LCL. The control chart indicates being out of control if $L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s}) < LCL$.

Example 2. Ryan (2000) provides an illustrative example with data set to compare the Hotelling's T^2 chart and the individual $\bar{X}$ charts. The data set been studied is a bivariate case that contains 20 subgroups of sample size $n = 4$. In this analysis, the upper control limit for Hotelling's T^2 chart is $UCL = 11.04$ and it is observed that numbered 10 and 20 sample points are with T^2 values 63.76 and 13.4 respectively. Hence, from the T^2 chart, these two sample points are out of control sample points. This study also present two individual $\bar{X}$ charts that shows only sample point numbered 6 is out of control.

The author indicated a benefit in using the T^2 chart. The sample means for these two variables show that they are strongly and positively correlated. It is also seen that the sample means of two variables showing on the respective $\bar{X}$ charts are both above their respective control midlines or below their midlines, and the distances to the midline are about the same on each chart. However, there is one exception. The sample point numbered 10 exhibit far away this pattern, one is well below its midline whereas the other point is somewhat above its midline. The author declared this benefit of able detecting the out of trend sample point through the T^2 chart whereas the individual $\bar{X}$ charts do not share this benefit.

For this data set, we performed in contructing the normal density control chart that gives lower control limit $LCL = 3.26156E - 15$ which indicates that there is one sample point, numbered 10 with density value $2.46463E - 24$, lying below the lower control limit. Hence, based on the criterion of density, numbered 10 sample is an out of control point. We delet this sample point and re-construct the normal density control chart that yields $LCL = 1.03162E - 14$. Again, only sample point numbered 10 is an out of control point. The density control chart has exactly detected the sample point that does not follow the pattern of the major data set. This is interesting for use of density control chart when we compare it with the T^2 chart and individual $\bar{X}$ charts. First, the T^2 chart have to pay the price to classify a sample point numbered 20 that stay on the main trend of the data set as an out of control point although it has the ability to observe sample point numbered 10 as an out of control point. Second, the individual $\bar{X}$ charts do not observe sample point numbered 20 as an out of control point but also indicates a sample point numbered 6 as out of control point, unfortunately, which point also stays on the main trend of the data set.

With sample point numbered 10 deleted, the constructed density control chart is

$$LCL = 1.03162E-14$$

$$\text{Test statistic function: } L(x_1, ..., x_n, \bar{\bar{x}}, \bar{s}) = 5.45183E-10$$

$$e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \begin{pmatrix}61.38158\\18.36842\end{pmatrix})' \begin{pmatrix}226.7237 & 100.6974\\100.6974 & 49.50877\end{pmatrix}^{-1} (x_i - \begin{pmatrix}61.38158\\18.36842\end{pmatrix})}$$

for on-line process control.

In the case of normal processes, one can actually use the statistic

$$\sum_{i=1}^{n}(x_i - \bar{\bar{x}})'\bar{s}^{-1}(x_i - \bar{\bar{x}})$$

for diagnosing whether the mean or the covariance matrix or both are out of control. Rewrite the statistic as

$$\begin{aligned}&\sum_{i=1}^{n}(x_i - \bar{x} + \bar{x} - \bar{\bar{x}})'\bar{s}^{-1}(x_i - \bar{x} + \bar{x} - \bar{\bar{x}})\\&= SS_s + n(\bar{x} - \bar{\bar{x}})'\bar{s}^{-1}(\bar{x} - \bar{\bar{x}})\end{aligned}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $SS_s = \sum_{i=1}^{n}(x_i - \bar{x})'\bar{s}^{-1}(x_i - \bar{x})$. Given $\bar{\bar{x}}$ and $\bar{s}$ and α, the region

$$\{(\bar{x}, SS_s) : SS_s + n(\bar{x} - \bar{\bar{x}})'\bar{s}^{-1}(\bar{x} - \bar{\bar{x}}) \leq \chi^2_{1-\alpha}\} \tag{3.1}$$

Any sample which produces $(\bar{x}, SS_s)$ that lies outside of the semi-circle indicates that the sample is out of control. If s^2 is closer to 0 while $\overline{X}$ is away from $\widehat{\mu}_0$, this is an indication that there is a possible mean shift. On the other hand, if $\overline{X}$ is closer to $\widehat{\mu}_0$ while s^2 is far away from 0, it is an indication that the variance has changed.

4. ARL Study for Normal Density Control Chart

The average run length (ARL) tell us, for a given situation in the distribution, how long on the average successive control chart points will be plotted before we detect a point beyond the control chart. Suppose that the vector of quality characteristics when the process is under control is multivariate normal $N_p(\mu_0, \Sigma_0)$ where we may let $\mu_0 = \bar{\bar{x}}$ and $\Sigma_0 = \bar{s}$. We consiedr the ARL when X follows the same multivariate normal distribution but with mean $\mu = \mu_0 + a\ell$ and covariance matrix $\Sigma = b^2\Sigma_0$. When $a = 0$ and $b = 1$, this will be an in control process, otherwise, it is out of contol process. Now, suppose that we have a random sample $X_1, ..., X_n$ drawn from distribution $N_p(\mu_0 + a\ell, b^2\Sigma_0)$, the ARL under this setting is derived based on the followings:

$$p^* = P_{\mu_0+a\ell,b^2\Sigma_0}(L(X_1, ..., X_n, \mu_0, \Sigma_0) \leq (2\pi)^{-np/2}|\Sigma_0|^{-n/2}e^{-\chi^2_\alpha(np)/2})$$
$$= P_{\mu_0+a\ell,b^2\Sigma_0}(\sum_{i=1}^{n}(X_i - \mu_0)'\Sigma_0^{-1}(X_i - \mu_0) \geq \chi^2_\alpha(np))$$

With the fact that $X_i - \mu_0 \sim N_p(\mu_0 + a\ell - \mu_0, b^2\Sigma_0)$, we see that $(b^2\Sigma_0)^{-1/2'}(X_i - \mu_0) \sim N_p(\frac{a}{b}\Sigma_0^{-1/2'}\ell, I_p)$. This further implies, from the independence assumption, that $\sum_{i=1}^{n}(X_i - \mu_0)'(b^2\Sigma_0)^{-1}(X_i - \mu_0)$ has a noncentral chi-square distribution $\chi^2_{np}(\frac{na^2}{b^2}\ell'\Sigma_0^{-1}\ell)$ where $\frac{na^2}{b^2}\ell'\Sigma_0^{-1}\ell$ is the noncentrality parameter. Hence, we further have

$$p^* = P_{\mu_0+a\ell,b^2\Sigma_0}(\sum_{i=1}^{n}(X_i - \mu_0)'(b^2\Sigma_0)^{-1}(X_i - \mu_0) \geq \frac{\chi^2_\alpha(np)}{b^2})$$
$$= P(\chi^2_{np}(\frac{na^2}{b^2}\ell'\Sigma_0^{-1}\ell) \geq \frac{\chi^2_\alpha(np)}{b^2}).$$

With the derivation of p^*, the ARL is formulated in the following

$$ARL = \frac{1}{p^*}. \tag{4.1}$$

For displaying its performance, we consider the case that $\ell = (1, ..., 1)'$ and $\Sigma_0 = I_p$ where this Σ_0 assumes that this multivariate normal distribution contains independent normal variables. In this situation, $p^* = P(\chi^2_{np}(\frac{na^2p}{b^2}) \geq \frac{\chi^2_\alpha(np)}{b^2})$. Restricting to bivariate distribution ($p = 2$) and for several cases of sample size n and combinations of (a, b), we display the ARL's in the following table.

Table 1. ARL for normal density control chart

	$n=2$	$n=3$	$n=5$	$n=20$
$a=0$				
$b=1$	370.37	370.37	370.37	370.37
$b=1.2$	42.48	32.88	22.46	5.69
$b=1.5$	8.02	5.60	3.47	1.17
$b=2$	2.51	1.84	1.33	1.00
$b=5$	1.04	1.00	1.00	1
$a=0.5$				
$b=1$	101.23	84.03	63.13	20.40
$b=1.2$	21.70	15.79	9.97	2.35
$b=1.5$	6.07	4.23	2.65	1.07
$b=2$	2.32	1.72	1.27	1.00
$b=5$	1.04	1.00	1.00	1
$a=1$				
$b=1$	15.14	9.98	5.61	1.34
$b=1.2$	6.90	4.66	2.80	1.07
$b=1.5$	3.45	2.43	1.62	1.00
$b=2$	1.91	1.46	1.15	1.00
$b=5$	1.04	1.00	1.00	1

In the next, we consider $\Sigma_0 = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$ which indicates that normal random variables forming the multivariate distribution are positivaly correlated. We denote $k = \ell'\Sigma_0^{-1}\ell$. With the same design for ℓ, we have $k = 1.25$.

We have several comments drawn from the above two tables:

(a) As usual, the ARL for in control process is 370.37. It is desired that all other situations that revealed out of control process' have ARL's smaller than 370.37.

(b) For a given $a, a = 0, 0.5, 1$, the ARL decreases when value b increases. On the other hand, when b is fixed the ARL is also nonincreasing when value a increases. This shows that the chart has better ability to detect a poorer process.

(c) Comparing the ARL's for a fixed combination of (n, a, b) in Tables 1 and 2, we see that ARL for correlated normal variables is relatively larger than it for uncorrelated one. This means that when correlation exists, the detection of out of control process is relatively more difficult.

Table 2. ARL for normal density control chart

	$n=2$	$n=3$	$n=5$	$n=20$
$a=0$				
$b=1$	370.37	370.37	370.37	370.37
$b=1.2$	54.69	44.10	31.98	9.52
$b=1.5$	11.73	8.47	13.96	1.55
$b=2$	3.66	2.64	2.66	1.01
$b=5$	1.16	1.05	1.01	1
$a=0.5$				
$b=1$	120.85	103.96	82.31	31.87
$b=1.2$	29.88	22.75	15.25	3.86
$b=1.5$	8.98	6.42	9.30	1.30
$b=2$	3.36	2.44	2.40	1.01
$b=5$	1.16	1.05	1.01	1
$a=1$				
$b=1$	23.18	16.12	9.55	2.00
$b=1.2$	10.51	7.33	4.47	1.30
$b=1.5$	5.16	3.65	4.21	1.05
$b=2$	2.71	2.00	1.88	1.00
$b=5$	1.15	1.05	1.01	1

A combined chart for multivariate normal distribution is not new. Chen, Cheng and Xie (2005) develoed a multivariate EWMA control chart, named the Max-MEWMA chart. They also compare this chart with the combination chart, the combination of χ^2 chart and the $|S|$ chart through the ARL's in a Monte Carlo study. They observed that this Max-MEWMA chart is more sensitive than the combination chart in detecting small to moderate changes in the mean vector and/or the variability. It is then interesting to compare this density control chart with Max-MEWMA chart and combination chart. Let's first define the indices for comparison:

$$\gamma_m = \frac{\text{ARL of Max-MEWMA}}{\text{ARL of density control chart}}, \quad \gamma_c = \frac{\text{ARL of Combination chart}}{\text{ARL of densitycontrol chart}}.$$

When the process is in-control, the γ values are expected to be 1 since these charts are designed to have the same first alarm rate. When the process is out of control, the smaller the γ values indicate the worser the density control chart than the other charts and the larger the γ values indicate the better (more sensitive) the density control chart than the other charts. The ARL's compuation of Chen, Cheng and Xie (2005) is based on distribution $N_2(\mu, \Sigma)$

with $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ a \end{pmatrix}$ and $\Sigma = b^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ with case $a = 0, b = 1$ the in-control distribution. We follow this design to compute the theoretical values of ARL's for the density control chart and compute their corresponding γ values. The results are listed in Table 3.

We have several comments drawn from the above table:

(a) When the data is drawn from the in-control distribution ($a = 0$ and $b = 1$), the three charts are similarly sensitive with equal ARLs.

(b) The Max-MEWMA is designed for detecting smaller changes in the distribution. We see that when $b = 1$ and $a \leq 1$ this moving average method is more sensitive than the density control chart in all ρ's. In all cases, the density control chart are better than the moving average chart. When the samples are drawn from wider situations such as $b = 2$ the γ_m values may be as large as 2.

(c) The combination chart is more sensitive than the density chart in all cases when $b = 1$. Again, in all other situations, the density chart is better than the combination chart. This sensitivity is strong for moderate change in distribution ($b = 1.5, 2.0$) and less strong for wide change ($b = 2, 3$).

(d) From (b) and (c), we see that when moderate or large changes may occur in a process, the density control chart is a suitable choice.

Table 3. ARL comparison with Max-MEWMA chart and combination chart

	$a = 0$	$a = 0.5$	$a = 1.0$	$a = 2.0$
$b = 1.0$				
$\rho = 0.0$	$\begin{pmatrix} 200 \\ \gamma_m = 1 \\ \gamma_c = 1 \end{pmatrix}$	$\begin{pmatrix} 78.58 \\ \gamma_m = 0.14 \\ \gamma_c = 0.61 \end{pmatrix}$	$\begin{pmatrix} 14.45 \\ \gamma_m = 0.26 \\ \gamma_c = 0.45 \end{pmatrix}$	$\begin{pmatrix} 1.51 \\ \gamma_m = 1.32 \\ \gamma_c = 0.72 \end{pmatrix}$
$\rho = 0.3$	$\begin{pmatrix} 200 \\ \gamma_m = 1.00 \\ \gamma_c = 1.00 \end{pmatrix}$	$\begin{pmatrix} 72.76 \\ \gamma_m = 0.15 \\ \gamma_c = 0.59 \end{pmatrix}$	$\begin{pmatrix} 12.35 \\ \gamma_m = 0.29 \\ \gamma_c = 0.46 \end{pmatrix}$	$\begin{pmatrix} 1.37 \\ \gamma_m = 1.38 \\ \gamma_c = 0.80 \end{pmatrix}$
$\rho = 0.6$	$\begin{pmatrix} 200 \\ \gamma_m = 1.00 \\ \gamma_c = 1.00 \end{pmatrix}$	$\begin{pmatrix} 52.07 \\ \gamma_m = 0.14 \\ \gamma_c = 0.54 \end{pmatrix}$	$\begin{pmatrix} 6.69 \\ \gamma_m = 0.44 \\ \gamma_c = 0.47 \end{pmatrix}$	$\begin{pmatrix} 1.08 \\ \gamma_m = 1.48 \\ \gamma_c = 0.92 \end{pmatrix}$
$\rho = 0.9$	$\begin{pmatrix} 200 \\ \gamma_m = 1.00 \\ \gamma_c = 1.00 \end{pmatrix}$	$\begin{pmatrix} 9.06 \\ \gamma_m = 0.36 \\ \gamma_c = 0.46 \end{pmatrix}$	$\begin{pmatrix} 1.18 \\ \gamma_m = 1.52 \\ \gamma_c = 0.84 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m = 1.00 \\ \gamma_c = 1.00 \end{pmatrix}$
$b = 1.5$				
$\rho = 0.0$	$\begin{pmatrix} 2.91 \\ \gamma_m = 1.44 \\ \gamma_c = 1.75 \end{pmatrix}$	$\begin{pmatrix} 2.56 \\ \gamma_m = 1.48 \\ \gamma_c = 1.60 \end{pmatrix}$	$\begin{pmatrix} 1.92 \\ \gamma_m = 1.61 \\ \gamma_c = 1.35 \end{pmatrix}$	$\begin{pmatrix} 1.16 \\ \gamma_m = 1.63 \\ \gamma_c = 1.03 \end{pmatrix}$
$\rho = 0.3$	$\begin{pmatrix} 2.91 \\ \gamma_m = 1.44 \\ \gamma_c = 1.75 \end{pmatrix}$	$\begin{pmatrix} 2.53 \\ \gamma_m = 1.50 \\ \gamma_c = 1.58 \end{pmatrix}$	$\begin{pmatrix} 1.86 \\ \gamma_m = 1.61 \\ \gamma_c = 1.29 \end{pmatrix}$	$\begin{pmatrix} 1.13 \\ \gamma_m = 1.59 \\ \gamma_c = 1.06 \end{pmatrix}$
$\rho = 0.6$	$\begin{pmatrix} 2.91 \\ \gamma_m = 1.44 \\ \gamma_c = 1.75 \end{pmatrix}$	$\begin{pmatrix} 2.41 \\ \gamma_m = 1.53 \\ \gamma_c = 1.53 \end{pmatrix}$	$\begin{pmatrix} 1.64 \\ \gamma_m = 1.64 \\ \gamma_c = 1.21 \end{pmatrix}$	$\begin{pmatrix} 1.05 \\ \gamma_m = 1.52 \\ \gamma_c = 1.04 \end{pmatrix}$
$\rho = 0.9$	$\begin{pmatrix} 2.91 \\ \gamma_m = 1.44 \\ \gamma_c = 1.75 \end{pmatrix}$	$\begin{pmatrix} 1.75 \\ \gamma_m = 1.60 \\ \gamma_c = 1.25 \end{pmatrix}$	$\begin{pmatrix} 1.08 \\ \gamma_m = 1.57 \\ \gamma_c = 1.01 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m = 1.00 \\ \gamma_c = 1.00 \end{pmatrix}$
$b = 2.0$				
$\rho = 0.0$	$\begin{pmatrix} 1.26 \\ \gamma_m = 1.82 \\ \gamma_c = 1.34 \end{pmatrix}$	$\begin{pmatrix} 1.24 \\ \gamma_m = 1.85 \\ \gamma_c = 1.29 \end{pmatrix}$	$\begin{pmatrix} 1.17 \\ \gamma_m = 1.88 \\ \gamma_c = 1.19 \end{pmatrix}$	$\begin{pmatrix} 1.05 \\ \gamma_m = 1.71 \\ \gamma_c = 1.04 \end{pmatrix}$
$\rho = 0.3$	$\begin{pmatrix} 1.26 \\ \gamma_m = 1.82 \\ \gamma_c = 1.34 \end{pmatrix}$	$\begin{pmatrix} 1.23 \\ \gamma_m = 1.86 \\ \gamma_c = 1.30 \end{pmatrix}$	$\begin{pmatrix} 1.17 \\ \gamma_m = 1.79 \\ \gamma_c = 1.19 \end{pmatrix}$	$\begin{pmatrix} 1.04 \\ \gamma_m = 1.63 \\ \gamma_c = 1.05 \end{pmatrix}$
$\rho = 0.6$	$\begin{pmatrix} 1.26 \\ \gamma_m = 1.82 \\ \gamma_c = 1.34 \end{pmatrix}$	$\begin{pmatrix} 1.22 \\ \gamma_m = 1.88 \\ \gamma_c = 1.31 \end{pmatrix}$	$\begin{pmatrix} 1.14 \\ \gamma_m = 1.84 \\ \gamma_c = 1.14 \end{pmatrix}$	$\begin{pmatrix} 1.02 \\ \gamma_m = 1.47 \\ \gamma_c = 1.07 \end{pmatrix}$
$\rho = 0.9$	$\begin{pmatrix} 1.26 \\ \gamma_m = 1.82 \\ \gamma_c = 1.34 \end{pmatrix}$	$\begin{pmatrix} 1.15 \\ \gamma_m = 1.82 \\ \gamma_c = 1.21 \end{pmatrix}$	$\begin{pmatrix} 1.03 \\ \gamma_m = 1.55 \\ \gamma_c = 1.06 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m = 1.00 \\ \gamma_c = 1.00 \end{pmatrix}$

Table 3 (Continue). ARL comparison with Max-MEWMA chart and

combination chart

	$a=0$	$a=0.5$	$a=1.0$	$a=2.0$
$b=2.5$				
$\rho=0.0$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.04 \\ \gamma_m=1.82 \\ \gamma_c=1.05 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.68 \\ \gamma_c=1.08 \end{pmatrix}$
$\rho=0.3$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.04 \\ \gamma_m=1.82 \\ \gamma_c=1.05 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.58 \\ \gamma_c=1.08 \end{pmatrix}$
$\rho=0.6$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.03 \\ \gamma_m=1.84 \\ \gamma_c=1.06 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.50 \\ \gamma_c=1.00 \end{pmatrix}$
$\rho=0.9$	$\begin{pmatrix} 1.05 \\ \gamma_m=1.90 \\ \gamma_c=1.14 \end{pmatrix}$	$\begin{pmatrix} 1.03 \\ \gamma_m=1.84 \\ \gamma_c=1.06 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.48 \\ \gamma_c=1.00 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.00 \\ \gamma_c=1.00 \end{pmatrix}$
$b=3.0$				
$\rho=0.0$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.88 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.88 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.78 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.30 \\ \gamma_c=1.00 \end{pmatrix}$
$\rho=0.3$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.88 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.88 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.78 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.30 \\ \gamma_c=1.00 \end{pmatrix}$
$\rho=0.6$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.88 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.78 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.68 \\ \gamma_c=1.00 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.20 \\ \gamma_c=1.00 \end{pmatrix}$
$\rho=0.9$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.88 \\ \gamma_c=1.08 \end{pmatrix}$	$\begin{pmatrix} 1.01 \\ \gamma_m=1.78 \\ \gamma_c=1.00 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.10 \\ \gamma_c=1.00 \end{pmatrix}$	$\begin{pmatrix} 1.00 \\ \gamma_m=1.10 \\ \gamma_c=1.00 \end{pmatrix}$

5. Approximate Density Control Chart

In previous sections, we discussed situations in which the distribution of the plotting statistics of the density control chart can be readily derived. However, in other applications it may be too complicated to derive such a distribution. In general, given $\widehat{\Theta}_1$, $\widehat{\Theta}_2$, ..., $\widehat{\Theta}_p$, if the distribution of $L(X_1, X_2, ...X_n; \widehat{\Theta}_1, \widehat{\Theta}_2, ..., \widehat{\Theta}_p)$ is too complicated, one can repeatedly generate samples from $f(x; \widehat{\Theta}_1, \widehat{\Theta}_2, ..., \widehat{\Theta}_p)$ calculate the likelihood of each of the, say, k generated samples, which are denoted by l_1, l_2, ..., l_k. Then the LCL can be estimated as $\widehat{\text{LCL}} = l_{([k\alpha]+1)}$, where $l_{(1)} \leq l_{(2)} \leq \cdots \leq l_{(k)}$ are the ordered statistics and $[a]$ is the largest integer smaller than or equal to

a. Here we look at another distribution to illustrate the use of simulations to approximate the LCL.

Suppose that a quality characteristic follows a multivariate t distribution having the following density function

$$f(x, \mu, \Sigma) = \frac{\Gamma(\frac{p+v}{2})}{\Gamma(\frac{v}{2})(\pi v)^{p/2}|\Sigma|^{1/2}}[1 + \frac{1}{v}(x-\mu)'\Sigma^{-1}(x-\mu)]^{-\frac{p+v}{2}}, x \in \Re^p$$

where μ and Σ are assumed to be unknown. However, it can be seen that μ and Σ are, respectively, the population mean and covariance matrix of random vector X. Suppose that μ_0 and Σ_0 are determined from a training sample. We carry out the process of generating a sample of size n from distribution with probability density function $f(x, \mu_0, \Sigma_0)$ $k = 1$ million times to form $\widehat{\text{LCL}}$. We conduct the process of generating lower limit $\widehat{\text{LCL}}$ 1 million times. By denoting ℓ^j the jth resulting limit, we let approximate LCL as $\hat{\ell} = \frac{1}{k}\sum_{j=1}^{1 \text{ million}} \ell^j$.

We then perform a 1 million replications to generate simulated ARL's. Let μ and Σ be the true parameters for the multivariate t distribution that sample $\mathbf{X}= (X_1, .., X_n)$ be drawn. For jth replication, we consecutively generate samples and compute their corresponding likelihood $L(x_1, ..., x_n, \mu_0, \Sigma_0)$. This generation stop when

$$L(x_1, ..., x_n, \mu_0, \Sigma_0) < \hat{\ell}. \tag{5.1}$$

We let ARL_j be the iteration number for that (5.1) occurs. The simulated ARL is then defined as

$$ARL = \frac{1}{1 \text{ million}} \sum_{j=1}^{1 \text{ million}} \text{ARL}_j.$$

We consider true parameters $\mu = \mu_0 + a$ and $\Sigma = b^2\Sigma_0$. The simulated ARL's are listed in the following table.

Table 4. Simulated ARL's for approximate density control chart

(a, b)	ARL	(a, b)	ARL	(a, b)	ARL
$(0, 1)$	390.52	$(0, 1)$	390.52	$(0, 1)$	390.52
$(0.3, 1)$	288.64	$(0, 1.1)$	86.220	$(0.2, 1.2)$	25.554
$(0.5, 1)$	158.07	$(0, 1.3)$	11.139	$(1, 2)$	1.2427
$(1, 1)$	24.186	$(0, 1.5)$	3.9138	$(1, 5)$	1.000075
$(1.5, 1)$	4.772	$(0, 2)$	1.3293	$(2, 2)$	1.0915
$(2.5, 1)$	1.0928	$(0, 3)$	1.0142	$(3, 2)$	1.0190
$(5, 1)$	1.0000	$(0, 5)$	1.000075	$(3, 5)$	1.0000

We have several comments for the simulation results in Table 4:

(a) The design $(a, b) = (0, 1)$ reveals that the process is in control and the theoretical value of ARL is approximately 370. The simulated ARL is 390.52 better than expected value.

(b) The cases in first column consider shift in location parameter μ but with scale parameter unchanged. The ARL decreases rapidly when value of μ moves away from μ_0. This fulfills our desirability. The ARL's perform in a similar manners when the location parameter is unchanged but the scale increases and both parameters move away from the control one.

(c) The performance of the ability of approximate simultaneous control chart showing in Table 4 seems to be satisfactory.

References

Amitava, M. (1998). *Fundamentals of quality control and improvement.* Prentice Hall: New York.

Chao, M. T. and Cheng, S. W. (1996). Semicircle control chart for variables data. *Quality Engineering.* **8**, 441-446.

Chen, G., Cheng, S. W. and Xie, H. (2005). A new multivariate control chart for monitoring both location and dispersion. *Communication in Statistics - Simulation and Computation.* **34**, 203-217. Cheng, S. W. and Thaga, K. (2005). Multivariate Max-CUSUM chart. *Quality Technology and Quantitative Management International.* **2**, 191-206.

Repco, J. (21986). Process capability plot. *The Proceedings of the 330th EQQC Conference.* 373-381.

Ryan, T. P. (2000). *Statistical Methods for Quality Improvement.* Wiley: New York.

Spiring, F. A. and Cheng, S. W. (1998). An alternative variables control chart: The univariate and multivariate case. *Statistica Sinica.* **8**, 273-287.

Van Nuland, Y. (1992). ISO 9002 and the circle technique. *Quality Engineering.* **5**, 269-291.

Woodall, W. H. (2000). Controversies and contradictions in statistical process control. *Journal of Quality Technology*, 32, 341-378.

Yeh, A. B. and Lin, D. K.-J. (2002). A new variables control chart for simultaneously monitoring multivariate process mean and variability. *International Journal of Reliability, Quality and Safety Engineering.* **9**, 41-59.