This article was downloaded by: [City University of Hong Kong Library] On: 26 August 2013, At: 19:40 Publisher: Taylor & Francis Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



### Journal of Applied Statistics

Publication details, including instructions for authors and subscription information: http://www.tandfonline.com/loi/cjas20

## Multivariate control chart based on the highest possibility region

Ting-Ting Gang<sup>a</sup>, Jun Yang<sup>a</sup> & Yu Zhao<sup>a</sup> <sup>a</sup> School of Reliability and Systems Engineering, Beihang University, Beijing, 100191, People's Republic of China Published online: 04 Jun 2013.

To cite this article: Ting-Ting Gang , Jun Yang & Yu Zhao (2013) Multivariate control chart based on the highest possibility region, Journal of Applied Statistics, 40:8, 1673-1681, DOI: <u>10.1080/02664763.2013.790007</u>

To link to this article: <u>http://dx.doi.org/10.1080/02664763.2013.790007</u>

#### PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <a href="http://www.tandfonline.com/page/terms-and-conditions">http://www.tandfonline.com/page/terms-and-conditions</a>

# Multivariate control chart based on the highest possibility region

Ting-Ting Gang, Jun Yang\* and Yu Zhao

School of Reliability and Systems Engineering, Beihang University, Beijing 100191, People's Republic of China

(Received 11 November 2012; accepted 24 March 2013)

The  $T^2$  control chart is widely adopted in multivariate statistical process control. However, when dealing with asymmetrical or multimodal distributions using the traditional  $T^2$  control chart, some points with relatively high occurrence possibility might be excluded, while some points with relatively low occurrence possibility might be accepted. Motived by the thought of the highest posterior density credible region, we develop a control chart based on the highest possibility region to solve this problem. It is shown that the proposed multivariate control chart will not only meet the false alarm requirement, but also ensure that all the in-control points are with relatively high occurrence possibility. The advantages and effectiveness of the proposed control chart are demonstrated by some numerical examples in the end.

**Keywords:** multivariate statistical process control; false alarm; highest possibility region; asymmetrical distributions; multimodal distributions;  $T^2$  control chart

#### 1. Introduction

In industrial applications, engineers often encounter the cases where several correlated quality characteristics need to be controlled and monitored simultaneously [3,12,13]. The multivariate control chart is widely utilized to solve this problem [15,19]. Due to the great efforts devoted by researchers, a large number of papers on control charts have been published in statistics and quality relevant journals [17,24,25,28,31]. For the recent advances in control chart design, such as economic design and optimal design, the readers can refer to Cheng and Mao [4] and Epprecht *et al.* [8] and the references therein for details.

In multivariate control fields, the traditional  $T^2$  control chart has been widely adopted in practice. When dealing with processes with symmetrical and unimodal distributions, the  $T^2$  control chart facilitates two preferred features: (1) the false alarm requirement is satisfied and (2) all the incontrol points have higher occurrence possibility than the out-of-control points. In order to monitor small changes in the processes, multivariate cumulative SUM control chart [7,10,21,27,28] and

<sup>\*</sup>Corresponding author. Email: tomyj2001@buaa.edu.cn

multivariate exponentially weighted moving average control chart [1,14,16,18,20,29,32,33] are often used.

However, those control charts are all under a very restrictive assumption, that is, the considered processes are all with multivariate normal distributions, which show great limitations in real quality control problems. Sun and Tsung [26] developed a new multivariate control chart named K-Chart using support vector machine (SVM), in which the restriction on the distributions of the processes is removed. Chiu *et al.* [5] studied how to obtain the proper parameters of SVM and apply SVM to construct a control chart to monitor the means of the processes with multivariate non-normal distributions. In Gani *et al.* [9], the K-Chart was applied to a real industrial problem successfully. However, the false alarm requirement is not considered in the K-Chart. An alternative way to obtain the control chart is by studying the distribution of the  $T^2$  statistics of multivariate non-normal distributions. In Chou *et al.* [6], a kind of kernel density estimation (KDE)-based  $T^2$  control charts are introduced. Phaladiganon *et al.* [22] proposed bootstrap-based  $T^2$  multivariate control charts, which performed comparably with the KDE-based  $T^2$  control charts. Nevertheless, none of the above methods consider the occurrence possibility of the in-control points.

It is noted that both the KDE-based  $T^2$  control chart and the bootstrap-based  $T^2$  chart meet the false alarm requirement, i.e. the average run length which is reciprocal to the false alarm level, for asymmetrical or multimodal distributions. Nevertheless, there might be some out-of-control points with relatively high occurrence possibility and some in-control points with relatively low occurrence possibility [30]. This phenomenon does not coincide with quality control in practice, since producers only want to exclude the bad states of the processes and maintain the good ones. As mentioned in the co-author's recent work [30], the null hypothesis of control chart should focus on the occurrence possibility of each data point besides the false alarm requirement, where the possibility means the density function for continuous distributions and the probability mass function for discrete distributions, respectively.

In this paper, a multivariate control chart based on the highest possibility region (HPR) is proposed to ensure that all the in-control points have higher occurrence possibility than the outof-control points, where the HPR denotes the confidence region with the minimum area. In Section 2, the  $T^2$  control chart is introduced and some of its merits are analysed. Section 3 is devoted to descriptions of the in-control region and control limit based on the highest density region. The advantages of the proposed control chart over the existed control charts and its relation to the  $T^2$  control chart are discussed in Section 4. In Section 5, some concluding remarks are given.

#### 2. $T^2$ control limits for multivariate distribution

Consider a *p*-dimensional random vector  $X = (X_1, ..., X_p)^T$  which follows a multivariate normal distribution with known mean vector  $\mu$  and covariance matrix  $\Sigma$ . Then the  $T^2$  statistic

$$T^{2} = (X - \mu)^{\mathrm{T}} \Sigma^{-1} (X - \mu)$$
(1)

follows  $\chi^2$  distribution, and the upper control limit (UCL) for the multivariate normal distribution is given by

$$UCL = \chi^2_{1-\alpha}(p), \tag{2}$$

where  $\alpha$  is the false alarm level and  $\chi^2_{1-\alpha}(p)$  is the  $1-\alpha$  quantile of a  $\chi^2$  distribution with p freedom degree. Denote the  $T^2$  control region as  $\Theta_{T^2}$  for the sake of simplicity. The properties of  $T^2$  control chart are represented in the following theorem.

THEOREM 2.1 If the process vector  $X = (X_1, ..., X_p)^T$  follows a multivariate normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$ , and the false alarm level is  $\alpha$ , then the  $T^2$  control limit

satisfies

$$P(X \in \Theta_{T^2}) = \int \cdots \int_{X \in \Theta_{T^2}} f(x) dx = 1 - \alpha$$
(3)

and

$$f(x) > f(y), \quad \forall x \in \Theta_{T^2}, \quad \forall y \notin \Theta_{T^2},$$
(4)

where f(x) is the density function of multivariate normal distribution, and

$$\Theta_{T^2} = \{x | T^2(x) \le \chi^2_{1-\alpha}(p)\}, \text{ where } T^2(x) = (x-\mu)^T \Sigma^{-1}(x-\mu).$$

*Proof* For a given false alarm  $\alpha$ , the  $T^2$  control region is given by  $\Theta_{T^2} = \{x | T^2(x) \le \chi^2_{1-\alpha}(p)\}$ since  $T^2 \sim \chi^2(p)$ , thus we have  $P(T^2 \le T_0^2 = \chi^2_{1-\alpha}(p)) = 1 - \alpha$  as in Equation (3).

The density mass function of multivariate normal distribution is given by

$$f(x) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-(1/2)(x-\mu)^{\mathrm{T}} \Sigma^{-1}(x-\mu)} = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-T^{2}(x)/2}.$$

One can observe that f(x) is strictly decreasing on  $T^2(x)$ . Then for  $\forall x \in \Theta_{T^2}$  and  $\forall y \notin \Theta_{T^2}$ , one has  $T^2(x) \le T_0^2 < T^2(y)$ , which implies

$$f(x) > f(y), \quad \forall x \in \Theta_{T^2}, \forall y \notin \Theta_{T^2}.$$

Thus the proof is completed.

*Example 2.1* Consider the multivariate normal distribution  $X \sim N(\mu, \Sigma)$ , i.e.  $f(x_1, x_2) = 1/2\pi |\Sigma|^{1/2} e^{-(1/2)(x-\mu)^{T}\Sigma^{-1}(x-\mu)}$ , where  $\mu = [0,0]^{T}$  and  $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$ . Suppose the false alarm level  $\alpha = 0.1$ . Then the  $T^2$  control region is given by  $\Theta_{T^2} = \{x | T^2(x) = x^{T}\Sigma^{-1}x \le 4.6052\}$ . From Theorem 2.1, we have

$$f(x) > f(y), \quad \forall x \in \Theta_{T^2}, \quad \forall y \notin \Theta_{T^2}.$$

If the process vector departures from a multivariate normal distribution, the UCL based on the  $\chi^2$  distribution might be inaccurate. In such a case, the kernel smoothing technique can be used to estimate the distribution of the  $T^2$  statistics, and the UCL of the  $T^2$  chart can thus be obtained. In the following, we will give the kernel estimator of the UCL when the density function f(x) is asymmetrical.

Generate a data set  $\{X^{(1)}, X^{(2)}, \dots, X^{(n)}\}$ , where  $X^{(i)} = (X_1^{(i)}, \dots, X_p^{(i)})^T$ , following a multivariate distribution f(x). The value of the  $T^2$  statistic for  $X^{(i)}$  is then given by

$$T_i^2 = (X^{(i)} - \mu)^{\mathrm{T}} \Sigma^{-1} (X^{(i)} - \mu).$$
(5)

The kernel estimator of the  $T^2$  distribution, which is denoted by FK, is given by

$$FK(y) = n^{-1} \sum_{j=1}^{n} \Phi\left\{\frac{y - T_j^2}{h}\right\},$$
(6)

where  $\Phi$  is a kernel function outlined in Polansky and Baker [23] and *h* is a smoothing parameter which varies with respect to different values of  $T^2$ . Since the UCL is the  $1 - \alpha$  quantile of the  $T^2$  distribution, the control limit UCL can be obtained as follows:

$$U\hat{C}L = FK^{-1}(1 - \alpha).$$
<sup>(7)</sup>

*Example 2.2* Chou *et al.* [6] considered a two-dimensional process X following a bivariate exponential distribution, where the mean vector is  $\mu = [1, 1]^T$  and covariance matrix is  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ . When  $\alpha = 0.0113$ , the  $T^2$  control limits are given by  $\Theta_{T^2} = \{x | (x_1 - 1)^2 + (x_2 - 1)^2 \le 19.44\}$ . One can observe that  $(3.14, 4) \in \Theta$  and  $(0.19, 5.49) \notin \Theta$ , however,

$$f(3.14,4) = 0.0008 < f(0.19, 5.94) = 0.0022.$$
(8)

The KDE-based  $T^2$  control limit of asymmetrical or multimodal distributions satisfies the false alarm level requirement, but do not have the good property as described in Equation (4). In other words, the KDE-based  $T^2$  control region may contain points with relatively low occurrence possibility, which is unreasonable in practice.

#### 3. Multivariate control chart based on the highest density region

As analysed in Section 2, although the KDE-based  $T^2$  control chart meets the false alarm requirement for asymmetrical or multimodal distributions, there might be some out-of-control points with relatively high occurrence possibility and some in-control points with relatively small occurrence possibility. From the decision procedure of a control chart [30], the decision is carried out by comparing each plotted data point with the control limits. In other words, the corresponding hypothesis test is actually conducted based on single data. Thus, the hypothesis test of control chart should also focus on the probability or density of individual data besides the false alarm probability.

In Bayesian statistics, for any given coverage rate, the highest posterior density credible region results in the minimum area credible region, which has the property that the possibility of all the points within the highest posterior density credible region is higher than the ones outside. Motivated by this idea, a multivariate control chart based on the HPR will be proposed in this section, which has the good properties described by Equations (3) and (4) for even asymmetrical and/or multimodal distributions.

#### 3.1 Control limits based on the highest density region

Consider a process vector  $X = (X_1, ..., X_p)^T$  which follows a continuous joint distribution with density function  $f(x) : \mathbb{R}^p \to \mathbb{R}$ . Assume that the false alarm level is  $\alpha$ . Then the corresponding null hypothesis is given by

$$H_0: P(X \in \Theta) = 1 - \alpha \quad \text{and} \quad f(x) > f(y), \, \forall x \in \Theta, \, \forall y \notin \Theta, \tag{9}$$

where  $\Theta$  is the acceptance region of the null hypothesis  $H_0$ .

Equivalently,  $\Theta$  is called the in-control region based on the highest density region, since any point x is in-control if and only if  $x \in \Theta$ . Furthermore, if the density function f(x) is continuous, the in-control region  $\Theta$  will be the solution of

$$P(X \in \Theta_{T^2}) = \int \cdots \int_{f(x_1, \dots, x_p) \ge c} f(x) \mathrm{d}x = 1 - \alpha.$$
(10)

Then the control limit  $E(\Theta) = \{x | f(x) = c\}$  is called the control limit based on the highest density region which is corresponding to an equal probability zone. Denote the HPR's control region as  $\Theta_{\text{HPR}}$  for sake of simplify.

For multivariate unimodal distribution,  $\Theta_{HPR}$  is a simple-connected region. However, if the continuous distribution is multimodal, the control region  $\Theta_{HPR}$  might be composed of a set of

distributed self-simple-connected regions, i.e.  $\Theta_{HPR} = \Theta_1 \cup \Theta_2 \cup \cdots \cup \Theta_k$ ,  $k \le m$ , where *m* is the number of peaks.

If the process variable X follows a multivariate discrete distribution with probability mass function p(x), and the false alarm level is  $\alpha$ , the null hypothesis corresponding to Equation (9) is then given by

$$H_0: P(X \in \Theta_{\text{HPR}}) \ge 1 - \alpha \quad \text{and} \quad p(x) > p(y), \, \forall x \in \Theta_{\text{HPR}}, \, \forall y \notin \Theta_{\text{HPR}}, \tag{11}$$

where  $\Theta_{\text{HPR}}$  is the in-control region based on the highest probability region for a multivariate discrete distribution. The control limit  $E(\Theta_{\text{HPR}}) = \{x | p(x) = c\}$  satisfies the following equation:

$$P(X \in \Theta_{\mathrm{HPR}}) = \sum_{p(x) \ge c} p(x) \ge 1 - \alpha.$$
(12)

According to the above constructing process of control limits based on the HPR, it is known that the proposed control limits not only satisfy the false alarm requirement for all distributions, but also ensure the in-control points have higher occurrence possibility than any out-of-control point, which are exactly the properties described by Equations (3) and (4). When p = 1, the proposed method reduced to that of Yang *et al.* [30]. In the next section, the calculation of HPR's control limit will be introduced.

#### 3.2 Calculation of HPR's control limit

For multivariate density function f(x), the HPR's control limit with false alarm level  $\alpha$  can be obtained from the definition described as in Equation (10). The procedure for solving the HPR's control limit can be summarized as follows:

Step 1. Choose a value d from the value range of f(x).

Step 2. Solve the equation  $f(x) \ge d$ , and the region  $\Theta = \{x | f(x) \ge d\}$  is obtained.

Step 3. Calculate the integral  $P = \int \cdots \int_{X \in \Theta} f(x) dx$ .

Step 4. If  $|P - (1 - \alpha)|$  is less than the prescribed small positive value  $\varepsilon$ , then  $\Theta$  will be the control region; if  $(1 - \alpha) - P$ , set d = d - e where e is a small positive constant and return to Step 2; and if  $P - (1 - \alpha)$ , set d = d + e and return to Step 2.

In many cases, it is difficult to obtain an analytic solution to the integral in *Step* 3. And the following approximating algorithm based on the Monte Carlo method [11] can be used instead.

Step 1'. Generate a set of independent random data  $x^{(i)} = \{x_1^{(i)}, \ldots, x_p^{(i)}\}, i \in \{1, \ldots, n\}$ , following the density function f(x).

Step 2'. Calculate the corresponding density value  $f(x^{(i)})$  and arrange the set  $f(x^{(1)}), \ldots, f(x^{(n)})$  in increasing order. The obtained set is denoted by  $f_1, \ldots, f_n$ .

Step 3'. In order to obtain the control limit  $f_{\alpha}$  which satisfies  $P(f(x) \ge f_{\alpha}) = 1 - \alpha$ , set  $j = \lfloor n * (1 - \alpha) \rfloor$ , where  $\lfloor x \rfloor$  denotes the maximum integer not exceeding x. If  $\lfloor j/n - (1 - \alpha) \rfloor < \varepsilon$ , the control limit is given by  $f_j$ , otherwise set n = n + d, where d is determined with respect to the false alarm requirement and return to *Step* 1'. Generally, d should be made large if the false alarm requirement is small.

For a multivariate discrete probability mass function P(x), the above procedure still works by replacing the integral in *Step* 3 by the sum of probability mass function P(x).

#### 4. Comparison between HPR's control chart and some existed control charts

In this section, the HPR's control limits will be compared with some existed control charts, for example, the multivariate normal  $T^2$  control chart and KDE-based  $T^2$  multivariate control

chart. And HPR's control limit of mixed multivariate normal distribution and mixture Poisson distribution will be calculated as the examples of multimodal and discrete distributions, respectively.

For multivariate normal distribution, the HPR's control chart is identified with traditional  $T^2$  control chart based on Theorem 2.1, which is summarized in the following corollary.

COROLLARY 3.1 Consider a process vector X which follows multivariate normal distribution with density function f(x). Then for any given false alarm level  $\alpha$  and  $T_0^2 = \chi_{1-\alpha}^2(p)$ , the HPR's control region  $\Theta_{\text{HPR}} = \{x | f(x) \ge f_{\alpha} = (1/(2\pi)^{p/2} |\Sigma|^{1/2}) e^{-T_0^2/2}\}$ , where  $\mu$  and  $\Sigma$  are the mean and covariance matrix of the distribution, respectively, is equivalent to the  $T^2$  control region  $\Theta_{T^2} = \{x | T^2(x) = (X - \mu)^T \Sigma^{-1}(X - \mu) \le T_0^2\}$ .

*Example 3.1* (Example 2.1 continued) Consider a process with two-dimensional normal distribution  $f(x_1, x_2)$ . One can obtain the HPR's control region  $\Theta_{\text{HPR}} = \{(x_1, x_2, f(x_1, x_2)) | f(x_1, x_2) \ge 0.0184\}$  by using the algorithms shown in Section 3.2. The obtained HPR's control region and the  $T^2$  control region  $\Theta_{T^2} = \{(x_1, x_2, 0) | T^2 = \frac{4}{3}(x_1^2 - x_1x_2 + x_2^2) \le 4.6052\}$  are both illustrated in Figure 1. It is found that  $\Theta_{T^2}$  is actually the projection of  $\Theta_{\text{HPR}}$  on the *X*-*Y* coordinate plane. In other words, the obtained two regions are equivalent on the *X*-*Y* coordinate plane in this case.

*Example 3.2* (Example 2.2 continued) By using the algorithms in Section 3.2, it is obtained that the HPR's control limit of the bivariate exponential distribution in Example 2.2 can be given by  $\Theta_{\text{HPR}} = \{(x_1, x_2) | f(x_1, x_2) \ge 0.0015\}$  when  $\alpha = 0.0113$ . Compared with the  $T^2$  control limits  $\Theta_{T^2} = \{(x_1, x_2) | (x_1 - 1)^2 + (x_2 - 1)^2 \le 19.44\}$  in Example 2.2, the minimum density of points in  $\Theta_{T^2}$  is 0.0002651, which is much less than the minimum density of points in  $\Theta_{\text{HPR}}$  (the value is 0.0015 in this example). In other words, there might exist some points within  $\Theta_{T^2}$  with smaller possibility than those outside  $\Theta_{T^2}$  and within  $\Theta_{\text{HPR}}$ , which is shown in Figure 2.

In Figure 2, the HPR's control region is given by  $\Theta_{\text{HPR}} = S_1 \cup S_2 \cup S_4$  and the  $T^2$  control area is given by  $\Theta_{T^2} = S_1 \cup S_3$ . Since the bivariate exponential distribution is strictly decreasing with



Figure 1. The HPR's control region and  $T^2$  control region of a multivariate normal distribution.



Figure 2. The difference between HPR's control region and  $T^2$  control region.

respect to  $x_1 + x_2$ , then one has that

$$f(x) > f(y), \quad \forall x \in S_2 \cup S_4, \forall y \in S_3.$$

It can be seen that for Example 2.2, the good property in Equation (4) cannot be guaranteed by using the KDE-based  $T^2$  control chart. However, Equation (4) is satisfied by using the HPR's control chart.

Furthermore, one can obtain that Area\_ $\Theta_{HPR} = 21.14$ , Area\_ $\Theta_{T^2} = 21.14$  and Area\_ $S_3 = 5.1413$ , where Area\_X represent the area of X region. Then, Area\_ $S_3$ /Area\_ $\Theta_{T^2} = 20.2\%$ . It is known that the density of points in  $S_3$  is smaller than the minimum density of points in  $\Theta_{HPR}$ . In the  $T^2$  control region, the points with lower density cover 20.2% area of  $\Theta_{T^2}$ . From the above comparison, one can observe that the HPR's method achieves better performance than the KDE-based  $T^2$  control chart for asymmetrical distributions.

In the following, it is shown that the HPR's control limit can also be applied to distributions with multimodal distributions or multivariate discrete distributions.

*Example 3.3* Consider a multivariate distribution, which is shown in Figure 3. Its density function is given by

$$f(x) = \frac{1}{4\pi} (|\Sigma_1|^{-1/2} e^{-(x-\mu_1)^T \Sigma_1^{-1} (x-\mu_1)/2} + |\Sigma_2|^{-1/2} e^{-(x-\mu_2)^T \Sigma_2^{-1} (x-\mu_2)/2}),$$

where  $\mu_1 = [0,0]^T$ ,  $\mu_2 = [3,4]^T$ ,  $\Sigma_1 = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$  and  $\Sigma_2 = \begin{pmatrix} 1 & -0.5 \\ -0.5 & 1 \end{pmatrix}$ . The false alarm level is chosen to be  $\alpha = 0.05$ . Then HPR's control limit can be given by  $\Theta_{\text{HPR}} = \{x | f(x) \ge 0.029\}$ , which is also shown in Figure 3. It can be seen that  $\Theta_{\text{HPR}}$  is not a simple-connected region and the property in Equation (4) is satisfied.

*Example 3.4* For the case of discrete distribution, we employ a mixture Poisson distribution  $X = (X_1, X_2)$ , which is given by

$$P(x_1, x_2) = 0.5 \frac{e^{-1}}{x_1!} + 0.5 \frac{e^{-15} \times 15^{x_2}}{x_2!}, \quad x_1 \ge 0, \, x_2 \ge 0.$$

Then the HPR's control limit can be obtained as  $\Theta = \{x | p(x) \ge 0.0404\}$  when  $\alpha = 0.01$ .



Figure 3. The HPR's regions for a multimodal distribution.

From the above analysis and numerical examples, it is obvious that the HPR's control chart has the following properties.

- (1) For the given false alarm level  $\alpha$ , the HPR's control region occupies the smallest possible volume in the sample space, which means that all the in-control points have a higher occurrence probability than any point out-of-control [2].
- (2) The HPR's control limit actually represents an equal possibility zone, which can be described by a simple equation even for distributions with high dimensions.

#### 5. Conclusions

The traditional multivariate  $T^2$  control chart only focuses on the false alarm requirement and it is very inaccurate for non-normal distribution. In this paper, a control chart based on the highest probability region (HPR's control limit) is proposed. It is shown that the proposed HPR's control limit not only satisfies the false alarm requirement but also ensures that all the in-control points have higher occurrence possibility than out-of-control ones. It is also shown that the HPR's control limit is prominent for asymmetrical or multimodal distributions, where the traditional multivariate  $T^2$  control limits and KDE-based  $T^2$  control limits are both very inaccurate. Especially, when considering one-dimensional distribution, the HPR's control limit reduces to the control limit based on the narrowest confidence interval proposed by Yang *et al.* [30], which is the extension of the 3-Sigma control limits and probability limits. Therefore, it is believed that the proposed approach enjoys much wider applications. Future work might include applying the approach to monitoring mean, dispersion and multivariate attribute characteristics.

#### Acknowledgement

The work described in this paper was supported by the National Natural Science Foundation of China under grant 11001005.

#### References

 S. Alkahtani and J. Schaffer, A double multivariate exponentially weighted moving average (dMEWMA) control chart for a process location monitoring, Commun. Stat. Simul. Comput. 41(2) (2012), pp. 238–252.

- [2] G. Box and G. Tiao, Bayesian Inference in Statistical Analysis, Addison-Wesley Publishing Co., Reading, MA, 1973.
- [3] S. Chen and H.B. Nembhard, Multivariate Cuscore control charts for monitoring the mean vector in autocorrelated processes, IIE Trans. 43(4) (2011), pp. 291–307.
- [4] S.W. Cheng and H. Mao, *The economic design of multivariate MSE control chart*, Qual. Technol. Quant. Manage. 8(2) (2011), pp. 75–85.
- [5] J. Chiu, W. Huang, Y. Chen and C. Tsai, A Study of Monitoring Non-normal Multivariate Process using Support Vector Machine, Proceedings of the International Multiconference of Engineers and Computer Scientists, Vol. II, Hong Kong, March 18–20, 2009.
- [6] Y. Chou, R.L. Mason, and J.C. Young, The control chart for individual observations from a multivariate non-normal distribution, Commun. Stat. Theory Methods 30(8–9) (2001), pp. 1937–1949.
- [7] Y. Dai, Y. Luo, Z. Li, and Z. Wang, A new adaptive CUSUM control chart for detecting the multivariate process mean, Qual. Reliab. Eng. Int. 27(7) (2011), pp. 877–884.
- [8] E.K. Epprecht, M.A. de Luna, and F. Aparisi, Joint EWMA charts for multivariate process control: Markov chain and optimal design, Int. J. Prod. Res. 49(23) (2011), pp.7151–7169.
- [9] W. Gani, H. Taleb, and M. Liman, An assessment of the kernel-distance-based multivariate control chart through an industrial application, Qual. Reliab. Eng. Int. 27(4) (2011), pp. 391–401.
- [10] D.M. Hawkins, Regression adjustment for variables in multivariate quality control, J. Qual. Technol. 25 (1993), pp. 170–182.
- [11] R.J. Hyndman, Computing and graphing highest density regions, Am. Stat. 50(2) (1996), pp. 120–126.
- [12] M.E Johnson, Multivariate Statistical Simulation, John Wiley & Sons, New York, 1987.
- [13] A.A. Kalgonda and S.R. Kulkarni, Multivariate quality control chart for autocorrelated processes, J. Appl. Stat. 31(3) (2004), pp. 317–327.
- [14] K. Linderman and T.E. Love, Economic and economic statistical designs for MA control charts, J. Qual. Technol. 32 (2000), pp. 410–417.
- [15] C.A. Lowry and D.C. Montgomery, A review of multivariate control charts, IIE Trans. 27 (1995), pp. 800-810.
- [16] C.A. Lowry, W.H. Woodall, C.W. Champ, and S.E. Rigdon, A multivariate EWMA control chart, Technometrics 34 (1992), pp. 46–53.
- [17] R. Mason, Y.M. Chou, and J.C. Young, Detection and interpretation of a multivariate signal using combined charts, Commun. Stat. Theory Methods 40(5) (2011), pp. 942–957.
- [18] A.O. Memar and S.T.A. Nikia, Multivariate variability monitoring using EWMA control charts based on squared deviation of observations from target, Qual. Reliab. Eng. Int. 27(8) (2011), pp. 1069–1086.
- [19] D.C. Montgomery, Introduction to Statistical Quality Control (5e), John Wiley & Sons, Inc., New York, 2005.
- [20] X. Pan, An alternative approach to multivariate EWMA control chart, J. Appl. Stat. 32(7) (2005), pp. 695–705.
- [21] J.J. Pignatiello and G.C. Runger, Comparisons of multivariate CUSUM charts, J. Qual. Technol. 22 (1990), pp. 173–186.
- [22] P. Phaladiganon, S.B. Kim, V.C.P. Chen, J. Baek, and S. Park, *Bootstrap-based T<sup>2</sup> multivariate control charts*, Commun. Stat. Simul. Comput. 40(5) (2011), pp. 645–662.
- [23] A.M. Polansky and E.R. Baker, Multistage plug-in bandwidth selection for kernel distribution function estimates, J. Stat. Comput. Simul. 65 (2000), pp. 63–80.
- [24] A. Seif, A. Faraz, C. Heuchenne, E. Saniga, and M.B. Moghadam, A modified economic-statistical design of the T<sup>2</sup> control chart with variable sample sizes and control limits, J. Appl. Stat. 38(11) (2011), pp. 2459–2469.
- [25] Z.G. Stoumbos, M.R. Reynolds, T.P. Ryan, and W.H. Woodall, *The state of statistical process control as we proceed into the 21st century*, J. Am. Stat. Assoc. 95 (2000) pp. 992–998.
- [26] R. Sun and F. Tsung, A kernel-distance-based multivariate control chart using support vector methods, Int. J. Prod. Res. 41 (2003), pp. 2975–2989.
- [27] W.H. Woodall and M.M. Ncube, *Multivariate CUSUM quality control procedures*, Technometrics 27 (1985), pp. 285–292.
- [28] W.H. Woodall and D.C. Montgomery, Research issues and ideas in statistical process control, J. Qual. Technol. 31 (1999), pp. 376–386.
- [29] Y. Xie, M. Xie, and T.N. Goh, Two MEWMA charts for Gumbel's bivariate exponential distribution, J. Qual. Technol. 43(1) (2011), pp. 50–65.
- [30] J. Yang, M. Xie, and T.N. Goh, Control limits based on the narrowest confidence interval, Commun. Stat. Theory Methods 40(12) (2011), pp. 2172–2181.
- [31] C.L. Yen, J.H. Shiau, and A.B. Yeh, Effective control charts for monitoring multivariate process dispersion, Qual. Reliab. Eng. Int. 28(4) (2012), pp. 409–426.
- [32] C. Zou and F. Tsung, A multivariate sign EWMA control chart, Technometrics 53(1) (2011), pp. 84–97.
- [33] C. Zou, Z. Wang, and F. Tsung, A spatial rank-based multivariate EWMA control chart, Naval Res. Log. 59(2) (2012), pp. 91–110.