A Study of Monitoring Non-normal Multivariate Process Using Support Vector Machine

Jing-Er Chiu, Wan-Chi Huang, Yu-Ying Chen and Chih-Hsin Tsai

Abstract—**This paper focuses on the monitoring techniques in multivariate processes when the underlying distribution of the quality characteristics differs from normality. Hotelling T^2 control chart is the most common used control chart for multivariate process, however, it is based on the assumption of normality. Normality assumption is not always reasonable.**

Researchers gradually applied support vector machine (SVM) to monitor non-normal multivariate process. By using SVM, the selection of parameters in SVM will affect the classification accuracy of SVM. It is an important issue of choosing the SVM parameters.

The purpose of this research is to apply SVM in statistical quality control. By simulating bivariate *t* and bivariate gamma distributions, we study the relationship between the distributions and parameters of SVM to obtain the best classification rate. After obtaining the proper parameters of SVM, we will applied SVM to construct a control chart to monitor non-normal process mean and study the performance of the new chart.

Keywords—**Support Vector Machine, Parameter design, Multivariate process control, Non-normal distribution**

I. INTRODUCTION

Multivariate processes are very important in modern industries for multiple quality characteristics. When monitoring the process means shifts, traditional control charts for multivariate processes are not applicable if the underlying distributions are not normal. Several attempts have been made in the literatures to extend traditional statistical process control (SPC) techniques to deal with non-normality.

Researchers gradually applied support vector machine (SVM) to monitor non-normal multivariate process. By using SVM, the selection of parameters in SVM will affect the classification accuracy of SVM. It is an important issue of choosing the SVM parameters. Sun and Tsung (2003) developed a new multivariate control chart using the approach of support vector machine (SVM), named K-Chart,

J.-E. Chiu is with the Department of Industrial Engineering and Management, National Yunlin University of Science & Technology, Yunlin 64002, Taiwan; R.O.C. (e-mail: chiuje@yuntech.edu.tw).

W.-C. Huang is with the Department of Industrial Engineering and Management, National Yunlin University of Science & Technology, Yunlin 64002, Taiwan; R.O.C.

Y.-Y. Chen is with the Department of Industrial Engineering and Management, National Yunlin University of Science & Technology, Yunlin 64002, Taiwan; R.O.C.

C.-H. Tsai is with the Department of Industrial Engineering and Management, National Yunlin University of Science & Technology, Yunlin 64002, Taiwan; R.O.C. (e-mail: g9521803@yuntech.edu.tw).

without restriction of distribution assumptions. However, they did not investigate the performance of K-Chart and no comparisons were made between K-Chart and other multivariate control charts.

The primary objective of this paper is to find out the best setting of parameters on SVM before establishing K-Chart. We simulate bivariate *t* and bivariate gamma distributions.

Then we study the relationship between the distributions and parameters of SVM to obtain the best classification rate. After obtaining the proper parameters of SVM, we will apply SVM to construct K Chart to monitor non-normal process mean. A secondary objective is to investigate the properties and performances of K-Chart by different type of shifts in mean vector and different correlations between quality characteristics.

The remainder of this paper is organized as follows. First, the fundamental of support vector machines is reviewed in section 2. The construction of K-Chart using support vector machine is introduced in section 3. The review of non-normal multivariate process is in section 4. Section 5 presents simulation study. Section 6 includes the results. Finally, the conclusions with some discussions are made in final section.

II. SUPPORT VECTOR MACHINE

A. Review of the support vector machine

The support vector machine where initially proposed for classifications between two classes (Vapnik 1995). The purpose of SVM is to seek the optimal linear hyperplane to separate two classes of data. The basic assumption that data is linearly separable, and also deal with linear and non-linear information. In this paper, we would discuss about non-linear process data. A description of SVM algorithm is follows. Let $\{(x_1, y_1),...,(x_i, y_i)\}, x_i \in R^n, y_i \in \{-1,1\}, i = 1,2,...l$ be the training set with input vectors and labels. Here, l is the number of sample observations and n is the dimension of each observation. The algorithm is to seek the hyperpalne $w \cdot x_i + b = 0$ to separate the data from two classes with a maximal margin width $2/\|w\|^2$, and the all points under the boundary is named support vector (For example, in figure 1). In order to optimal the hyperplane that SVM was to solve the optimization problem was following.

$$\text{Min } \Phi(x) = 1/2\|w\|^2$$

$$\text{s.t. } y_i(w^T x_i + b) \geq 1, i = 1,2,...,l$$

It is difficult to solve regard equation, must to transform the optimization problem to be dual problem by Lagrange

Method. The value of α in the Lagrange method must be non-negative real coefficients. The solutions of the following programming problem,
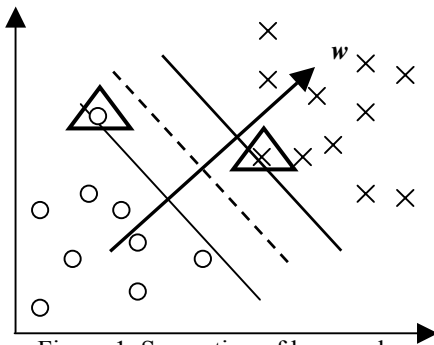

Figure 1. Separation of hyperpalne

$$Max \ \Phi(w,b,\xi,\alpha,\beta) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1,j=1}^{l} \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$s.t. \qquad \sum_{j=1}^{l} \alpha_j y_j = 0$$

$$0 \le \alpha_i \le C, i = 1,2,...,l$$

In order to separate two class exactly, add a slack variable (ξ) in the Lagrange equation to make equation $y_i\left(w^T x_i + b\right) \ge 1 - \xi_i \ \& \ \xi_i > 0$. An objective of slack variable is to increase the flexible buffer of boundary.

B. Kernel function

In general, it couldn't find the linear separate hyperplane in all application data. In the non-linear data, it must transform the original data to higher dimension of linear separate is the best solution (in figure 2). The higher dimension is called "Feature Space", it would improve that the date separated by linear classification.
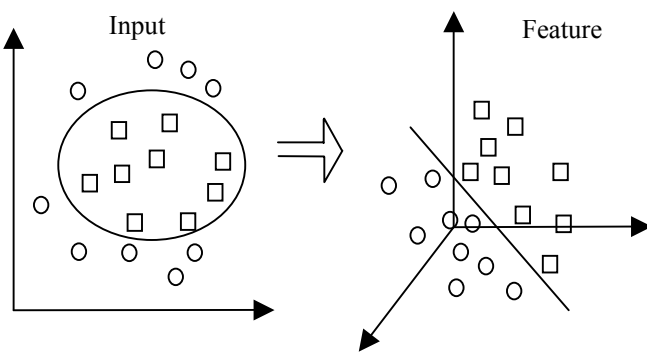

Figure 2. Non-linear mapping

The common kernel functions are linear, polynomial, radial basis function; RBF and sigmoid. The descriptions of kernel function are following.
(1) Linear Kernel Function:
$$K(x_i, x_j) = x_i^T x_j$$

(2) Polynomial Kernel Function:
$$K(x_i, x_j) = \left(\gamma x_i^T x_j + r\right)^d$$

(3) Gaussian Radial Basis Kernel Function; RBF:

$$K(x_i, x_j) = \exp\left(\frac{\left\|x_i - x_j\right\|^2}{2\sigma^2}\right) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right)$$

(4) Sigmoid Kernel Function:
$$K(x_i, x_j) = \tanh\left(\gamma x_i^T x_j + r\right)$$

There has each parameters in different kernel function, it obtains different correct ratio with different parameters of kernel function. While the dual problem appends the kernel function that the dual problem would change is:

$$Max \ \Phi(w,b,\xi,\alpha,\beta) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i=1,j=1}^{l} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$s.t. \qquad \sum_{j=1}^{l} \alpha_j y_j = 0$$

$$0 \le \alpha_i \le C, i = 1,2,...,l$$

In this paper, we will discuss the RBF kernel function, the reason that is the RBF could separate data of non-linear and higher dimensions. To using RBF kernel function would only adjust two parameters are C and γ . This is the first choice that user append the kernel function

III. K-CHART

K-chart was base on the support vector machine that developed to monitor the multivariate process. The control chart would search the support vector by support vector machine and obtain the kernel distance to establish control chart (Sun & Tsung 2003). The control limits of k-chart were decided by the boundary of support vector. Therefore, there is no assumption of distributions in k-chart. The illustration of two dimensions is following called Hypersphere (in figure 3).
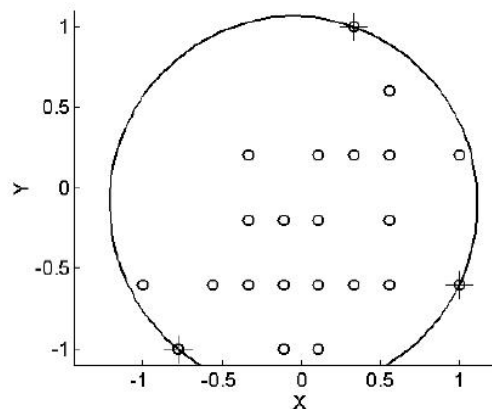
Figure 3. Hypersphere of two dimensions
Sun & Tsung (2003)

When the process have a new sample point that to calculate the kernel distance. First is to calculate the distance between new point and central point, the equation is:

$$d = \sqrt{(z-O)^T (z-O)}$$

In above equation that z is a new sample point, O is the central point. If d is over than the radius(R) that shows the sample point is out-of-the-control. The central point will renew by the following equation.

$$d = \sqrt{\left(z^T z\right) - 2\sum_{i=1}^{l} \alpha_i (z, x_i) + \sum_{i,j=1}^{l} \alpha_i \alpha_j \left(x_i, x_j\right)}$$

Then, the kernel distance (denote: kd) will be obtain:

$$kd = \sqrt{\left(z^T z\right) - 2\sum_{i=1}^{l} \alpha_i K(z, x_i) + \sum_{i,j=1}^{l} \alpha_i \alpha_j K\left(x_i, x_j\right)}$$

There is a few point α_i exceed zero, above the equation will be:

$$kd = \sqrt{\left(z^T z\right) - 2\sum_{i \in S}^{l} \alpha_i K(z, \boldsymbol{x}_i) + \sum_{i,j \in S}^{l} \alpha_i \alpha_j K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)}$$

S is the support vector in above equation. If the z is the support vector, the kernel distance (kd) would be control limit.

Let the all sample points of obtained kernel distance to draw the chart is the k-chart (in figure 4).
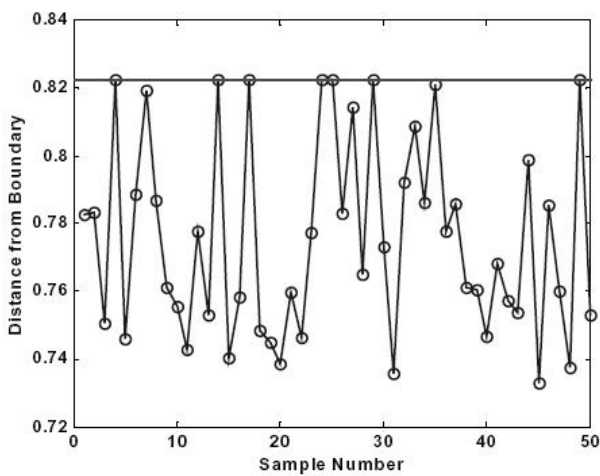

Figure 4. K-chart
Sun & Tsung (2003)

Support vectors were selected of all sample point would use for solving the quadratic programming problem is following:

$$max \sum_{i=1}^{l} \alpha_i \boldsymbol{K}\left(\boldsymbol{x}_i, \boldsymbol{x}_i\right) - \sum_{i,j=}^{l} \alpha_i \alpha_j \boldsymbol{K}\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right)$$

$$s.t. \; \alpha_i \geq 0 \; and \; \sum_{i=1}^{l} \alpha_i = 1$$

If the unusual point appeared, the support vector would involve the points, it had influence of established in k-chart.

In order to eliminate the problem, suggesting enhanced the restriction were:

$$s.t. \; 0 \leq \alpha_i \leq B \; and \; \sum_{i=1}^{l} \alpha_i = 1$$

B would control parameters of α_i is the positive. When B became smaller, the type I error will be bigger. Otherwise, B became bigger; the type II error will be bigger. This is also the important issues in choice of B.

IV. NON-NORMAL MULTIVARIATE PROCESS

In this study, we will focus on non-normal multivariate process that simulation under the multivariate t distributions and multivariate gamma distributions. This section describes two types of distributions.

A. Multivariate t distributions

The bivariate t distribution was extended from unvariate t distribution (Kotz & Nadarajah 2004). Let $X=(x_i, x_j)$ from bivariate t distribution, denote $X \sim t_2(\upsilon)$ was degree of ν. The probability density function is following equation:

$$f(X) = \frac{1}{2\pi |\Sigma|^{1/2}} \left(1 + \sum_{i,j=1}^{2} \frac{X_i X_j}{\upsilon |\Sigma|^{1/2}}\right)^{-\upsilon+2/2}$$

$$; \Sigma = \begin{bmatrix} 1 & \rho_{12} \\ \rho_{21} & 1 \end{bmatrix}$$

B. Multivariate gamma distributions

There are many methods of simulation on gamma distribution. Base on the method proposed, we simulate bivariate gamma distributions by Sim (1993). Let $Z=(Z_1, Z_2)$ from bivariate gamma distribution denote $Z \sim g(\alpha_1, \alpha_2, \beta_1, \beta_2)$. The parameters of shape were α_1 and α_2, the parameters of scale were β_1 and β_2.

If $Z_1 = X_1$ were from unvariate gamma distribution then, the variable of $Z_2 = \frac{\beta_1}{\beta_2} B_1 X_1 + X_2$. The B_1 must be Beta $(\beta_1 \beta_2 S_{12}, \alpha_1 - \beta_1 \beta_2 S_{12})$, otherwise, X_2 is from $g(\alpha_2 - \beta_1 \beta_2 S_{12}, \beta_2)$. About S_{12} was convariance of two dimensions and correlation coefficient (ρ) and convariance (S_{12}) were transformed from $\rho = \frac{Cov(Z_1, Z_2)}{\sqrt{V(Z_1)}\sqrt{V(Z_2)}}$.

V. SIMULATION

A. Simulation process

In this study, we discuss setting parameters of combination of C and γ in SVM. We simulate data of non-normal multivariate process. It uses SVM to find the combination of parameters. Second, it analyzes the performance of K-Chart in monitoring non-normal multivariate process. The procedure for simulation process in this study is shown in Figure 5.

Therefore, we use the LIBSVM program from the guide of SVM (Hsu 2003). It utilizes the LIBSVM be the program interface of this study. We transform the data according to the type of LIVSVM and to find the optimal combination of parameters under cross-validation.
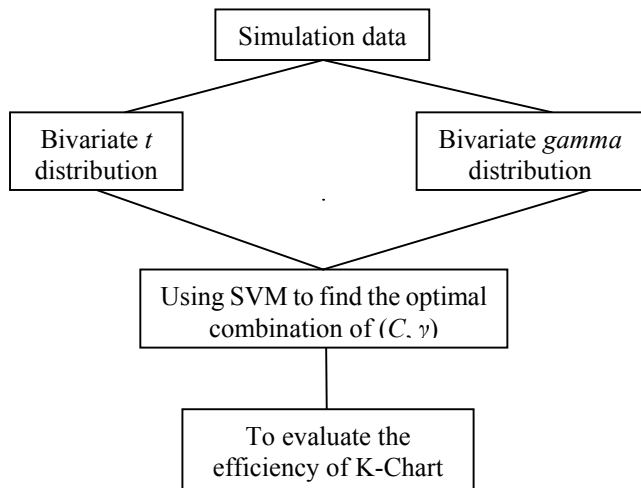
Figure 5. The procedure for simulation

Table 1. Accuracy of optimal parameter by SVM (n=100)

Type	ρ	(C, γ)	Accuracy (%)
$t_2(6)$	0	(1.4~1.5, 1.1~2)	82.553
	0.2	(0.1~1, 0.1)	91.176
	0.5	(1.1~2, 1.4~1.6)	85.294
$t_2(1000)$	0	(0.4~1, 0.1~0.2)	97.059
	0.2	(1.9~2, 0.5)	88.235
	0.5	(0.3~1, 0.1)	91.176
$g(4,4,1,1)$	0	(0.1~0.9, 7~17)	70.588
	0.2	(0.1~0.9, 7~25)	85.294
	0.5	(0.1~0.9, 10~30)	82.353
$g(16,16,1,1)$	0	(0.1~0.6, 2.3~2.5)	79.412
	0.2	(0.1~0.9, 1.3~2)	79.412
	0.5	(0.1~1, 1.8~2)	76.471
$g(1024,1024,1,1)$	0	(0.1~1, 0.1)	85.294
	0.2	(0.1~1, 0.1)	76.471
	0.5	(0.1~0.9, 0.1)	82.353

B. Simulation

In this study, we set the number of quality characteristic is two and four observations. From the idea of Stomumbos & Sullivan (2002), setting different degree is 6 and 1000 in *t* distribution. The training data are 100, 500, 1500, 3000 and the testing data including 34, 168, 500, 1000 respectively. In the simulation data, we would simulate different shift in traing and testing data.

Otherwise, simulation method of gamma distribution is the same as the *t* distributions. In order to study the correlation of simulation data, we simulate 3 type of correlation ($\rho_{12} = 0$; $\rho_{12} = 0.2$; $\rho_{12} = 0.5$).

C. Evaluation

In this study, we monitor the non-normal multivariate process under the SVM tool. SVM is to classify two different categories, so the measurement is ratio of classification accuracy.

Table 2. Accuracy of optimal parameter by SVM (n=500)

Type	ρ	(C, γ)	Accuracy (%)
$t_2(6)$	0	(1.1~2,0.1)	92.857
	0.2	(0.8,0.1)	91.667
	0.5	(0.3,0.2)	89.286
$t_2(1000)$	0	(1,0.1)	95.238
	0.2	(0.4~0.5,0.1)	91.667
	0.5	(1.7~2.5,0.7)	86.31
$g(4,4,1,1)$	0	(1.6~1.7,0.1)	93.452
	0.2	(1.1,0.1)	92.262
	0.5	(1.1,0.1)	87.5
$g(16,16,1,1)$	0	(1.1,0.1)	91.071
	0.2	(1.5~2,0.1)	88.095
	0.5	(1.9~2,0.1)	77.976
$g(1024,1024,1,1)$	0	(0.1~0.9,0.1)	76.19
	0.2	(0.1~0.9,0.1)	81.548
	0.5	(0.1~0.9,0.1)	75

In table 3 and 4, it shows that the ratio enhance gradually by data size. By the way, the accuracy ratio decreases by the correlation. In the simulation of parameter, it offers that combination of parameters (*C*, γ) on different type of distributions.

VI. RESULTS

A. Parameter selection

The simulation result of parameter selection is shown from table 1 to table 4. When the sample size is 100 that detect the ratio of gamma distribution is smaller than the others. Otherwise, there is no rule on different correlation of parameter combinations. When the data size is 500 that display the ratio is higher than others in most situations on table 2.

Table 3. Accuracy of optimal parameter by SVM (n=1500)

Type	ρ	(C, γ)	Accuracy (%)
$t_2(6)$	0	(1.3~1.5,0.3)	92.2
	0.2	(1~1.2,0.1)	91
	0.5	(0.6~0.7,0.1)	89.4
$t_2(1000)$	0	(0.1~1,0.6)	89.8
	0.2	(0.5~0.9,0.1)	90.4
	0.5	(0.1,0.1)	86
$g(4,4,1,1)$	0	(1.2,0.2)	94.2
	0.2	(1.4~1.6,0.1)	91.6
	0.5	(1.6~1.8,0.1)	89.2
$g(16,16,1,1)$	0	(1.5~1.6,0.1)	88
	0.2	(2,0.1)	88.6
	0.5	(2,0.1)	87
$g(1024,1024,1,1)$	0	(0.1~0.9,0.1)	86.4
	0.2	(0.1~0.9,0.1)	80.4
	0.5	(0.6~0.9,0.2)	75.6

Table 4. Accuracy of optimal parameter by SVM (n=3000)

Type	ρ	(C, γ)	Accuracy (%)
	0	(0.1,1~1.1)	94
$t_2(6)$	0.2	(0.1,0.7)	91.8
	0.5	(0.1,1)	90.1
	0	(0.6,0.4)	92
$t_2(1000)$	0.2	(0.2,0.1~0.2)	88.9
	0.5	(0.1,0.7~1)	87.7
	0	(0.1,0.5~0.6)	95.5
$g(4,4,1,1)$	0.2	(0.1,0.9)	94
	0.5	(0.1,0.4~0.7)	90.5
	0	(0.1,0.3)	90
$g(16,16,1,1)$	0.2	(0.1,0.6~0.9)	87.1
	0.5	(0.1,0.1)	84.7
	0	(0.2,0.1~0.9)	85.5
$g(1024,1024,1,1)$	0.2	(0.1,0.1~0.3)	80.7
	0.5	(0.2,0.1~0.7)	78.2

B. Performance of K-Chart

In this section, we simulated the efficiency of K-Chart to monitor non-normal multivariate process. The simulation result displayed from table 5 to 11. In table 5 to 8 is about that monitor the bivariate t distribution and the other is about bivariate gamma distribution.

Table 5. Accuracy of K-Chart ($t_2(6)$; ρ =0.2)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	97.04	2.96	9.00
(0.5,0.5)	97.20	4.80	8.58
(0.75,0.75)	97.28	7.32	8.50
(1,1)	97.04	16.12	8.67
(1.25,1.25)	97.40	23.36	8.00
(1.5,1.5)	97.16	41.00	8.08
(1.75,1.75)	97.16	64.52	8.17
(2,2)	97.20	82.00	8.08
(2.25,2.25)	97.40	92.52	8.17
(2.5,2.5)	96.72	98.00	7.50
(2.75,2.75)	96.96	99.04	8.33
(3,3)	97.48	99.96	8.42

Table 6. Accuracy of K-Chart ($t_2(6)$; ρ =0.8)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	97.60	2.64	7.83
(0.5,0.5)	97.76	2.68	7.67
(0.75,0.75)	97.24	5.92	7.08
(1,1)	97.24	9.16	7.25
(1.25,1.25)	97.44	19.12	7.67
(1.5,1.5)	97.40	33.68	8.00
(1.75,1.75)	98.12	49.20	7.25
(2,2)	97.44	65.00	8.08
(2.25,2.25)	97.24	84.76	8.00
(2.5,2.5)	97.36	93.68	7.67
(2.75,2.75)	97.84	97.40	7.33
(3,3)	97.28	98.96	7.08

In the simulation of $t_2(6)$ distribution exhibited that the accuracy enhanced by the process shift, especial in the large shift. When correlation of process was increase, the number of support vectors was also increase. If there is less quality characteristic, SVM would include the less supports to the process data.

Table 7. Accuracy of K-Chart ($t_2(20)$; ρ =0.2)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	97.12	3.04	8.08
(0.5,0.5)	96.92	6.28	8.08
(0.75,0.75)	97.32	14.00	8.25
(1,1)	96.88	28.44	8.17
(1.25,1.25)	97.36	45.72	8.50
(1.5,1.5)	97.60	67.48	8.50
(1.75,1.75)	96.92	86.72	8.17
(2,2)	97.64	94.96	8.17
(2.25,2.25)	97.28	98.24	8.83
(2.5,2.5)	97.80	99.76	7.83
(2.75,2.75)	97.12	100.00	8.58
(3,3)	97.40	100.00	8.00

Table 8. Accuracy of K-Chart ($t_2(20)$; ρ =0.8)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	97.68	3.16	7.25
(0.5,0.5)	97.72	5.68	7.58
(0.75,0.75)	97.28	9.92	7.83
(1,1)	98.12	21.56	7.50
(1.25,1.25)	97.36	33.00	7.58
(1.5,1.5)	97.56	54.00	8.00
(1.75,1.75)	97.80	70.32	7.83
(2,2)	98.12	83.96	7.50
(2.25,2.25)	97.64	90.76	7.67
(2.5,2.5)	96.96	98.00	7.42
(2.75,2.75)	97.68	99.28	7.92
(3,3)	97.80	99.92	7.58

In the analysis of $t_2(20)$, the accuracy enhanced by the process shift (60%~100%). The SVM would include a few support vectors when ρ was increase. Whatever the type of multivariate t process, the K-Chart would monitor process well.

Otherwise, the simulation of gamma distribution were including two type of degree and different correlation coefficient (ρ =0.2,0.8,) show in table 9 to table 12.

There was lower accuracy of classification in small process shift (less than 2), especially in $g_2(4,4,1,1)$ form. If the process shift was more than 2, the accuracy enhances of $g_2(16,16,1,1)$ type. Otherwise, the number of support vector is less than 23 in gamma distribution.

Table 9. Accuracy of K-Chart ($g_2(4,4,1,1,\rho)$; ρ =0.2)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	97.24	2.12	10.16
(0.5,0.5)	96.48	3.84	9.64
(0.75,0.75)	96.68	6.12	9.82
(1,1)	96.52	11.76	9.78
(1.25,1.25)	96.60	17.04	9.74
(1.5,1.5)	96.68	30.24	10.24
(1.75,1.75)	97.04	44.36	9.92
(2,2)	97.04	67.76	10.02
(2.25,2.25)	96.48	84.76	10.04
(2.5,2.5)	97.24	93.52	9.98
(2.75,2.75)	97.24	99.08	9.78
(3,3)	97.04	99.92	9.70

Table 12. Accuracy of K-Chart ($g_2(16,16,1,1,\rho)$; ρ =0.8)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	92.48	6.32	23.56
(0.5,0.5)	91.64	11.48	23.66
(0.75,0.75)	93.24	18.56	23.66
(1,1)	92.44	35.00	24.00
(1.25,1.25)	92.00	52.68	23.94
(1.5,1.5)	91.32	72.72	22.56
(1.75,1.75)	92.16	89.16	23.26
(2,2)	91.44	96.72	22.88
(2.25,2.25)	92.28	99.48	23.08
(2.5,2.5)	92.72	99.88	22.50
(2.75,2.75)	92.52	100.00	23.54
(3,3)	92.80	100.00	22.94

Table 10. Accuracy of K-Chart ($g_2(4,4,1,1,\rho)$; ρ =0.8)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	96.88	2.56	9.64
(0.5,0.5)	96.48	3.36	9.92
(0.75,0.75)	96.64	6.36	9.54
(1,1)	96.84	10.08	10.34
(1.25,1.25)	96.92	16.40	9.86
(1.5,1.5)	97.04	27.20	9.74
(1.75,1.75)	96.64	42.32	9.98
(2,2)	96.52	62.36	9.94
(2.25,2.25)	97.80	76.96	9.76
(2.5,2.5)	96.84	91.72	9.50
(2.75,2.75)	96.16	97.76	9.62
(3,3)	97.20	99.36	9.52

Table 11. Accuracy of K-Chart ($g_2(16,16,1,1,\rho)$; ρ =0.2)

Shift (μ_1, μ_2)	In-Control %	Out-of-Control %	Number of support vector
(0.25,0.25)	91.96	6.52	23.82
(0.5,0.5)	92.60	11.52	22.70
(0.75,0.75)	91.64	20.60	23.32
(1,1)	92.40	35.96	22.76
(1.25,1.25)	91.36	55.24	23.50
(1.5,1.5)	92.48	74.92	23.16
(1.75,1.75)	92.44	88.24	23.44
(2,2)	91.44	96.88	23.72
(2.25,2.25)	93.24	99.44	23.30
(2.5,2.5)	92.68	99.96	22.60
(2.75,2.75)	92.24	99.96	23.50
(3,3)	92.44	100.00	23.48

VII. CONCLUSION

In this paper, we presented a relationship between the non-normal multivariate distributions and parameters of SVM to obtain classification rate. After obtaining the proper parameters of SVM, we discuss the performance of control chart base on SVM. Based on an analysis of SVM parameters, the classification rate would increase by sample size. For various non-normal bivariate distributions, the combination of parameters (C, γ) will be stable by sample size. The interval of two parameters is between zero and one. The default parameter setting (C=0.5, γ =0.5) obtain the classification rate is fairly accurate under larger sample size. When the sample size is smaller (n=100), we obtain the best classification rate under more accurate parameter selection.

After obtaining the proper parameters of SVM, we discuss the performance of the control chart based on SVM to monitor non-normal process. The K-Chart obtains the better classification rate under monitoring process mean to large shift ($\geq 1.5\sigma$). In the various correlation coefficients, the K-Chart obtains the different classification rate. It will obtain the best rate when the process with larger ρ. The K-Chart will be able to monitor process well whatever the different non-normal multivariate distributions are.

In the future research the K-Chart will be applied in the time series process. We would discuss the performance of K-Chart in monitoring the time series model or extending to the multivariate time series.

REFERENCES

[1] C. H. Sim, "Generation of poisson and gamma random vectors with given marginals and covariance matrix," *Journal of the Statistical Computer Simulation.*, No. 47, 1993, pp. 1-10.
[2] V. N. Vapnik, *The Nature of Statistical Learning Theory.* New York: Springer, 1995.
[3] Z. G. Stoumbos, and J. H. Sullivan, "Robustness to non-normality of the multivariate EWMA control chart," *Journal of Quality Technology.*, vol. 34, 2002, pp. 260-276.
[4] C. W. Hsu, C. C. Chang, and C. J. Lin, *A Practical Guide to Support Vector Classification*, Information Engineering, National Taiwan University, 2003.
[5] R. Sun, and F. Tsung, "A kernel-distance-based multivariate control chart using support vector methods," *International Journal of Production Research.*, vol. 41, 2003, pp. 2975-2989.
[6] S. Kotz and S. Nadarajah, *Multivariate T distributions and their applications*, Cambridge University Press, 2004.