A STUDY ON THE PERFORMANCES OF MEWMA AND MCUSUM CHARTS FOR SKEWED DISTRIBUTIONS

Michael B. C. Khoo[1], Abdu. M. A. Atta[2] and H. N. Phua[3]
School of Mathematical Sciences, Universiti Sains Malaysia, Penang, Malaysia
Email: mkbc@usm.my[1]

ABSTRACT

A multivariate chart, instead of separate univariate charts is used for a joint monitoring of several correlated variables. Two time weighted multivariate charts that are commonly used for a quick detection of small shifts in the mean vector are the multivariate exponentially weighted moving average (MEWMA) and multivariate cumulative sum (MCUSUM) charts. The MEWMA and MCUSUM charts use information from past data, which make them sensitive to small shifts. These charts require the assumption that the underlying process follows a multivariate normal distribution. This paper studies the robustness of the MEWMA and MCUSUM charts toward nonnormality by considering the multivariate Weibull and multivariate gamma distributions based on different sample sizes and correlation coefficients.

1. INTRODUCTION

In most process monitoring situations, the quality of a process is determined by two or more quality characteristics (Woodall and Montgomery, 1999). Process monitoring problems involving several related variables of interest are called multivariate statistical process control. The most useful tool used in the monitoring of a multivariate process is a multivariate control chart. The first step in constructing a multivariate chart involves the analysis of a preliminary set of data that is assumed to be in statistical control. This analysis is known as a Phase-I analysis and it is conducted to estimate process parameters that will be used for the monitoring of a future process, a.k.a., a Phase-II process.

Numerous multivariate charts and their extensions are presently available. These charts can be grouped into 3 broad categories, namely, the Hotelling's T^2, multivariate EWMA (MEWMA) and multivariate CUSUM (MCUSUM) charts. The Hotelling's T^2 chart was proposed by Hotelling (1947) for the detection of a large sustained shift. The MCUSUM chart was first suggested by Woodall and Ncube (1985) while the MEWMA chart was introduced by Lowry et al. (1992). However, the MCUSUM charts suggested by Crosier (1988) will be discussed in this paper as they are more widely used.

This paper is organized as follows: Section 2 reviews the MEWMA chart while Section 3 reviews the MCUSUM chart. In Section 4, a simulation study is conducted to compare the performances of MEWMA and MCUSUM charts for skewed distributions. Finally, conclusions are drawn in Section 5.

2. MEWMA CONTROL CHART

The MEWMA chart proposed by Lowry et al. (1992) is based on the following statistic:

$$\boldsymbol{Z}_t = \lambda \boldsymbol{X}_t + (1-\lambda)\boldsymbol{Z}_{t-1}, \text{ for } t = 1, 2, \ldots, \quad (1)$$

where $\boldsymbol{Z}_0 = \boldsymbol{\mu}_0$ and $0 < \lambda \leq 1$. $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots,$ are assumed to be independent multivariate normal random vectors, each with p quality characteristics. The control charting statistic of a MEWMA chart is (Lowry et al., 1992)

$$T_t^2 = \boldsymbol{Z}_t' \Sigma_{\boldsymbol{Z}_t}^{-1} \boldsymbol{Z}_t. \quad (2)$$

The chart signals a shift in the mean vector when $T_t^2 > h_1$, where h_1 is the limit chosen to achieve a desired in-control ARL (ARL_0) and

$$\Sigma_{\boldsymbol{Z}_t} = \frac{\lambda}{2-\lambda}\left[1-(1-\lambda)^{2t}\right]\Sigma_{\boldsymbol{X}} \quad (3)$$

is the variance-covariance matrix for $\boldsymbol{Z}_t$. Lowry et al. (1992) showed that the run length performance of the MEWMA chart depends on the off-target mean vector $\boldsymbol{\mu}_1$ and the covariance matrix of $\boldsymbol{X}_t$, i.e., $\Sigma_{\boldsymbol{X}_t}$ only through the value of the non-centrality parameter,

$$\delta = \left\{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \Sigma_{\boldsymbol{X}}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)\right\}^{1/2}, \quad (4)$$

where $\boldsymbol{\mu}_0$ denotes the in-control mean vector.

Lee and Khoo (2006a) provide a method based on the Markov chain approach for the selection of the optimal parameters, λ and h_1, which produce the minimum out-of-control ARL (ARL_1) for a desired size of a shift of interest based on a fixed ARL_0.

3. MCUSUM CONTROL CHART

Crosier (1988) suggested two multivariate CUSUM charts. The one with the better ARL performance is based on the following statistics:

$$C_t = \left\{(\boldsymbol{S}_{t-1} + \boldsymbol{X}_t - \boldsymbol{a})' \Sigma_{\boldsymbol{X}}^{-1} (\boldsymbol{S}_{t-1} + \boldsymbol{X}_t - \boldsymbol{a})\right\}^{1/2}, \text{ for } t = 1, 2, \ldots, \quad (5)$$

where

$$\boldsymbol{S}_t = \begin{cases} \boldsymbol{0}, & \text{if } C_t \leq k \\ (\boldsymbol{S}_{t-1} + \boldsymbol{X}_t - \boldsymbol{a})\left(1 - \dfrac{k}{C_t}\right), & \text{if } C_t > k \end{cases}. \quad (6)$$

Note that $\boldsymbol{S}_0 = \boldsymbol{0}$, $k > 0$ is the reference value and $\boldsymbol{a}$ is the aim point or target value for the mean vector. The control charting statistic for the MCUSUM chart is (Crosier, 1988)

$$Y_t = \left(\boldsymbol{S}_t' \Sigma_X^{-1} \boldsymbol{S}_t\right)^{1/2} \tag{7}$$

A shift in the mean vector is signalled when $Y_t > h_2$, where h_2 represents the limit of the chart. The MCUSUM procedure assumes that the multivariate observations $\boldsymbol{X}_t$, for t = 1, 2, ..., follow an independently and identically distributed (i.i.d.) multivariate normal distribution.

Lee and Khoo (2006b) give an approach based on the Markov chain method in determining the optimal parameters, k and h_2 that give the minimum out-of-control ARL (ARL_1) for a size of shift of interest based on a fixed ARL_0.

4. A SIMULATION STUDY

The assumption of the underlying process having i.i.d. multivariate normal random variates is required for both the MEWMA and MCUSUM charts. Since many multivariate processes, such as chemical processes come from populations that are skewed, it is difficult to satisfy the multivariate normality assumption. In this section, the performances of the MEWMA and MCUSUM charts will be studied when the multivariate normality assumption is violated.

The performances of the MEWMA and MCUSUM charts are compared based on the false alarm rates when the process is in-control for multivariate skewed distributions, such as the Lee's multivariate Weibull (Lee, 1979) and Cheriyan and Ramabhadran's multivariate gamma distributions (Cheriyan, 1941 and Ramabhadran, 1951). For the sake of comparison, the multivariate normal distribution is also considered. For convenience, the bivariate case, i.e., the number of quality characteristics, p = 2 is considered. Note that the bivariate Weibull distribution can represent various skewnesses and correlations but the bivariate gamma can only represent some positive correlations (Kotz et al., 2000).

SAS programs are used to compute the false alarm rates for the three multivariate distributions considered. Each false alarm rate is computed based on 5000 simulation trials. The nominal false alarm rate is assumed to be α = 0.0027 when the underlying distribution is bivariate normal. The MEWMA and MCUSUM charts are designed for a quick detection of a shift in the mean vector of size δ = 1. The optimal smoothing constant, λ = 0.13 and limit h_1 = 10.55 are found for the MEWMA chart using the approach described in Lee and Khoo (2006a). Similarly, using the procedure given in Lee and Khoo (2006b), the optimal parameters are found to be k = 0.5 and h_2 = 6.227 for the MCUSUM chart.

The correlation coefficients, ρ = 0.3, 0.5 and 0.8 are considered for the bivariate distributions. For ease of computation, the sacle parameters of (1,1) for (X_1, X_2) are selected for the Weibull and gamma distributions. The shape parameters for (X_1, X_2) are chosen so that the desired skewnesses (γ_1, γ_2) = {(1,1), (1,2), (1,3), (2,2), (2, 3), (3,3)} for these parameters are attained. The sample sizes, n = 3, 5 and 7 are considered.

The false alarm rates for the MEWMA and MCUSUM charts are given in Tables 1 and 2, respectively. Note that the false alarm rates, marked as "*" in Tables 1 and 2 for the Cheriyan and Ramabhadran's bivariate gamma distribution cannot be computed because the corresponding shape parameters of one of the gamma distributed components, used in the transformation to compute variate X_2 have negative values.

From Tables 1 and 2, it is found that for the multivariate Weibull and gamma distributions, the false alarm rates of the MEWMA and MCUSUM charts increase as the level of skewness and correlation coefficient increase. This is because the covariance matrix of the multivariate observation, $\boldsymbol{X}$ is inflated as the skewness and correlation coefficient increase, hence making it easier for the MEWMA and MCUSUM charts to issue out-of-control signals. Also note that the false alarm rate decreases as the sample size increases. This is consistent with the multivariate central limit theorem, where the sample mean vector of a multivariate skewed distribution approaches multivariate normality as the sample size increases. A comparison of the false alarm rates of the two charts show that generally the MEWMA chart has lower false alarm rates than the MCUSUM chart for various levels of skewnesses. Thus, the MEWMA chart is more robust than the MCUSUM chart.

5. CONCLUSIONS

In this paper, we have studied the performance of the MEWMA and MCUSUM charts for multivariate normal and multivariate skewed distributions. We found that the false alarms of both charts are affected by the skewness of the underlying distribution. Also, the sample size and correlation of the quality characteristics have an impact on the false alarm rates of the charts. The simulation results show that the MEWMA chart has a lower false alarm rate than the MCUSUM chart when the underlying distribution is skewed. Since it is known that both the MEWMA and MCUSUM charts have equal performances in the detection of small shifts when the underlying process is multivariate normally distributed, the use of the MEWMA chart in process monitoring is recommended because the MEWMA chart is more robust towards skewed populations.

ACKNOWLEDGEMENTS

This research is supported by the Universiti Sains Malaysia, Research University (RU) grant, no. 1001/PMATHS/811024.

REFERENCES

1. Cheriyan, K.C. (1941). A bivariate correlated gamma-type distribution function. *Journal of the Indian Mathematical Society*, 5, 133-144.
2. Crosier, R.B. (1988). Multivariate generalizations of cumulative sum quality-control schemes. *Technometrics*, 30, 291–303.

3. Hotelling, H. (1947). "Multivariate quality control," *Techniques of Statistical Analysis*, Eisenhart, Hastay and Wallis (eds.), McGraw-Hill, New York.
4. Kotz, S., Balakrishnan, N. and Johnson, N.L. (2000). *Continuous multivariate distributions*, Vol. 1., 2^{nd} ed., John Wiley, New York.
5. Lee, L. (1979). Multivariate distributions having Weibull properties. *Journal of Multivariate Analysis*, 9, 267–277.
6. Lee, M.H. and Khoo, M.B.C. (2006a). Optimal statistical design of a multivariate EWMA chart based on ARL and MRL. *Communications in Statistics – Simulation and Computation*, 35, 831–847.
7. Lee, M.H. and Khoo, M.B.C. (2006b). Optimal statistical design of a multivariate CUSUM chart. *International Journal of Reliability, Quality and Safety Engineering*, 13, 479–497.
8. Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. (1992). A multivariate exponentially weighted moving average control chart. *Technometrics*, 34, 46–53.
9. Ramabhadran, V.R. (1951). A multivariate gamma-type distributions. *Sankhya*, 11, 45–46.
10. Woodall, W.H. and Montgomery, D.C. (1999). Research issues and ideas in statistical process control. *Journal of Quality Technology*, 31, 376-386.
11. Woodall, W.H. and Ncube, M.M. (1985). Multivariate CUSUM quality control procedures. *Technometrics*, 27, 285–292.

Table 1. False alarm rates for the MEWMA chart when $\lambda = 0.13$ and $h_1 = 10.55$

Correlation coefficient	Multivariate distribution	Skewness coefficient (γ_1, γ_2)	Sample size, n		
			3	5	7
$\rho = 0.3$	Normal	(0,0)	0.0025830	0.0026450	0.0026600
	Weibull	(1,1)	0.0029040	0.0027250	0.0026130
		(1,2)	0.0035350	0.0031600	0.0030100
		(1,3)	0.0043070	0.0037940	0.0035010
		(2,2)	0.0041670	0.0035550	0.0033540
		(2,3)	0.0049280	0.0041320	0.0038270
		(3,3)	0.0056780	0.0047290	0.0043020
	Gamma	(1,1)	0.0029980	0.0028110	0.0027850
		(1,2)	0.0034710	0.0031680	0.0030860
		(1,3)	0.0039560	0.0035100	0.0033090
		(2,2)	0.0040630	0.0036180	0.0032550
		(2,3)	0.0047300	0.0040400	0.0036070
		(3,3)	0.0053240	0.0043790	0.0039400
$\rho = 0.5$	Normal	(0,0)	0.0025830	0.0026450	0.0026600
	Weibull	(1,1)	0.0028690	0.0026110	0.0024610
		(1,2)	0.0037170	0.0032160	0.0030640
		(1,3)	0.0047470	0.0040950	0.0038720
		(2,2)	0.0044490	0.0037830	0.0035030
		(2,3)	0.0053900	0.0045610	0.0042110
		(3,3)	0.0062970	0.0052450	0.0048160
	Gamma	(1,1)	0.0030340	0.0029600	0.0027750
		(1,2)	0.0036180	0.0031160	0.0030910
		(1,3)	*	*	*
		(2,2)	0.0041920	0.0035860	0.0033770
		(2,3)	0.0047770	0.0039130	0.0036190
		(3,3)	0.0055110	0.0046630	0.0042070
$\rho = 0.8$	Normal	(0,0)	0.0025830	0.0026450	0.0026600
	Weibull	(1,1)	0.0033460	0.0027720	0.0025900
		(1,2)	0.0049340	0.0041010	0.0036990
		(1,3)	0.0070200	0.0060950	0.0057130
		(2,2)	0.0057440	0.0046370	0.0041530
		(2,3)	0.0071680	0.0059480	0.0054110
		(3,3)	0.0081140	0.0068870	0.0062570
	Gamma	(1,1)	0.0032220	0.0030090	0.0029240
		(1,2)	*	*	*
		(1,3)	*	*	*
		(2,2)	0.0048070	0.0040280	0.0036410
		(2,3)	*	*	*
		(3,3)	0.0064080	0.0053320	0.0047210

Table 2. False alarm rates for the MCUSUM chart when $k = 0.5$ and $h_2 = 6.227$

Correlation coefficient	Multivariate distribution	Skewness coefficient (γ_1, γ_2)	Sample size, n		
			3	5	7
ρ = 0.3	Normal	(0,0)	0.0026930	0.0027220	0.0027220
	Weibull	(1,1)	0.0029440	0.0027420	0.0026800
		(1,2)	0.0035020	0.0031400	0.0030080
		(1,3)	0.0042500	0.0037290	0.0034920
		(2,2)	0.0036650	0.0033290	0.0031400
		(2,3)	0.0048150	0.0040180	0.0037430
		(3,3)	0.0055340	0.0045740	0.0041640
	Gamma	(1,1)	0.0026930	0.0027220	0.0027220
		(1,2)	0.0029440	0.0027420	0.0026800
		(1,3)	0.0035020	0.0031400	0.0030080
		(2,2)	0.0042500	0.0037290	0.0034920
		(2,3)	0.0036650	0.0033290	0.0031400
		(3,3)	0.0048150	0.0040180	0.0037430
ρ = 0.5	Normal	(0,0)	0.0026930	0.0027220	0.0027220
	Weibull	(1,1)	0.0028180	0.0025440	0.0024610
		(1,2)	0.0036490	0.0031830	0.0030670
		(1,3)	0.0046620	0.0040150	0.0038180
		(2,2)	0.0043290	0.0036840	0.0034060
		(2,3)	0.0052570	0.0044200	0.0040660
		(3,3)	0.0061670	0.0050810	0.0046200
	Gamma	(1,1)	0.0031260	0.0028970	0.0028610
		(1,2)	0.0034620	0.0032160	0.0029970
		(1,3)	*	*	*
		(2,2)	0.0039720	0.0035540	0.0033110
		(2,3)	0.0047190	0.0039000	0.0035950
		(3,3)	0.0055250	0.0045230	0.0040310
ρ = 0.8	Normal	(0,0)	0.0026930	0.0027220	0.0027220
	Weibull	(1,1)	0.0030700	0.0026280	0.0024780
		(1,2)	0.0047220	0.0039330	0.0036050
		(1,3)	0.0068540	0.0059950	0.0056410
		(2,2)	0.0054640	0.0044190	0.0039710
		(2,3)	0.0068990	0.0056910	0.0051910
		(3,3)	0.0078760	0.0065890	0.0059620
	Gamma	(1,1)	0.0032610	0.0029770	0.0028680
		(1,2)	*	*	*
		(1,3)	*	*	*
		(2,2)	0.0045670	0.0038680	0.0035700
		(2,3)	*	*	*
		(3,3)	0.0062740	0.0052470	0.0044940