

EWMA Control Charts to Monitor Correlation Coefficients

Duk Joon Chang¹⁾, Gyo-Young Cho²⁾ and Jae Man Lee³⁾

Abstract

Multivariate EWMA control charts to simultaneously monitor correlation coefficients of correlated quality characteristics under multivariate normal process are proposed. Performances of the proposed charts are measured in terms of average run length(ARL). Numerical results show that small values for smoothing constant with accumulate-combine approach are preferred for detecting small shifts.

1. Introduction

In statistical process for quality control, we often need to monitor several correlated quality characteristics simultaneously. In this case, multivariate statistical process control charts are efficient to detect assignable causes that are poorly detected by univariate control charts on individual quality variables. And shifts in correlation coefficients of quality variables are important when the strength of linear relationship between two or more characteristics largely affect the quality of product. Especially in chemical industry, changing in correlation coefficients of quality variables in production process is usually important.

Shewhart chart has been used to detect large shift in the process. However, small or moderate changes of the process are expected in many practical situations. In this case, EWMA and CUSUM charts have been used. The original work on multivariate chart which is based on a measure of distance of $\underline{X}' = (X_1, X_2, \dots, X_p)$, random vector of p quality variables at each sampling point, to the target mean vector $\underline{\mu}_0$ was introduced by Hotelling(1947) and the multivariate approach became popular in recent years. Alt(1982) and Jackson(1985) reviewed much of the articles on multivariate chart.

A multivariate EWMA(MEWMA) chart for mean vector $\underline{\mu}$ with accumulate-combine technique was proposed by Lowry et. al.(1992). By simulation, they showed that the performances of MEWMA procedure performs better than the multivariate CUSUM procedures of Crosier(1988) and Pignatiello and Runger(1990). Little research to monitor correlation

1) Associate Professor, Dept. of Statistics, Changwon National University, Changwon, 641-773, Korea

2) Associate Professor, Dept. of Statistics, Kyungpook National University, Taegu, 702-701 Korea

3) Associate Professor, Dept. of Statistics, Andong National University, Andong, 760-749, Korea

coefficients of correlated quality variables has appeared in quality control literatures. In this paper, we propose multivariate control statistics and EWMA charts to simultaneously monitor correlation coefficients of several correlated quality characteristics under multivariate normal process.

2. Constructing Multivariate Control Statistics

The underlying probability distribution of p quality characteristics is assumed to be multivariate normal with mean vector $\underline{\mu}$ and covariance matrix Σ . We obtain a sequence of random vectors $\underline{X}_1, \underline{X}_2, \dots$ to judge the state of the process where $\underline{X}_t = (\underline{X}'_{t1}, \dots, \underline{X}'_{tp})'$ is an observation vector of each sampling point t and $\underline{X}_{tj} = (X_{tj1}, \dots, X_{tjp})'$.

Let $\underline{\theta}_0 = (\underline{\mu}_0, \Sigma_0)$ be the known target process parameters for $\underline{\theta} = (\underline{\mu}, \Sigma)$, where $\underline{\theta}_0$ is represented as

$$\underline{\mu}_0 = \begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{p0} \end{bmatrix} \quad \text{and} \quad \Sigma_0 = \begin{bmatrix} \sigma_{10}^2 & \rho_{120}\sigma_{10}\sigma_{20} & \cdots & \rho_{1p0}\sigma_{10}\sigma_{p0} \\ & \sigma_{20}^2 & \cdots & \rho_{2p0}\sigma_{20}\sigma_{p0} \\ & & \ddots & \vdots \\ & & & \sigma_{p0}^2 \end{bmatrix}.$$

A univariate EWMA chart for ρ_{12} which is based on $r_{12} = \sum_{j=1}^n (x_j - \mu_1)(y_j - \mu_2) / n\sigma_1\sigma_2$, an estimator of ρ_{12} , can be constructed as

$$Y_t = (1 - \lambda)Y_{t-1} + \lambda \frac{\sum_{j=1}^n (X_{tj1} - \mu_{10})(X_{tj2} - \mu_{20})}{n\sigma_{10}\sigma_{20}}, \quad (2.1)$$

$t = 1, 2, \dots$ and $0 < \lambda \leq 1$. By repeated substitution, it can be shown that

$$Y_t = (1 - \lambda)^t Y_0 + \sum_{k=1}^t \lambda(1 - \lambda)^{t-k} \frac{\sum_{j=1}^n (X_{kj1} - \mu_{10})(X_{kj2} - \mu_{20})}{n\sigma_{10}\sigma_{20}}. \quad (2.2)$$

To simultaneously monitor the correlation coefficients of p correlated quality variables, if we let the control statistic for ρ_{lm} be r_{lm} by suitable modification of the simple expression in (2.2), then the vectors of EWMA's for $\underline{\rho} = (\rho_{12}, \rho_{13}, \dots, \rho_{1p}, \rho_{23}, \dots, \rho_{2p}, \dots, \rho_{p-1,p})'$ can be defined as

$$\begin{aligned} \underline{Y}'_t &= (r_{12}, r_{13}, \dots, r_{1p}, r_{23}, \dots, r_{2p}, \dots, r_{p-2,p-1}, r_{p-1,p}) \\ &= (Y_{t1}, Y_{t2}, \dots, Y_{t,p-1}, Y_{tp}, \dots, Y_{t,2p-3}, \dots, Y_{t,s-1}, Y_{ts}). \end{aligned}$$

Then the vector \underline{Y}_t can be rewritten as

$$\underline{Y}_t = \begin{bmatrix} (1-\lambda_1)^t Y_{t0} + \sum_{k=1}^t \lambda_1 (1-\lambda_1)^{t-k} Z_{k12} \\ \vdots \\ (1-\lambda_{p-1})^t Y_{t,p-1,0} + \sum_{k=1}^t \lambda_{p-1} (1-\lambda_{p-1})^{t-k} Z_{k1p} \\ (1-\lambda_p)^t Y_{t,p,0} + \sum_{k=1}^t \lambda_p (1-\lambda_p)^{t-k} Z_{k23} \\ \vdots \\ (1-\lambda_{2p-3})^t Y_{t,2p-3,0} + \sum_{k=1}^t \lambda_{2p-3} (1-\lambda_{2p-3})^{t-k} Z_{k2p} \\ \vdots \\ (1-\lambda_s)^t Y_{t,s,0} + \sum_{k=1}^t \lambda_s (1-\lambda_s)^{t-k} Z_{k,p-1,p} \end{bmatrix}, \quad (2.3)$$

where $s = p(p-1)/2$, $0 < \lambda_a \leq 1$ ($a = 1, 2, \dots, s$) and

$$Z_{kmu} = \frac{\sum_{j=1}^n (X_{kjm} - \mu_{m0})(X_{kju} - \mu_{u0})}{n\sigma_{m0}\sigma_{u0}} - \rho_{mu}, \quad (m \neq u)$$

Another control statistic can be obtained by using the likelihood ratio test (LRT) statistic for testing $H_0: \Sigma = \Sigma_0$ vs $H_1: \Sigma \neq \Sigma_0$ where μ_0 is known, because a control chart can be viewed as repeated tests of significance. By simple calculation, we obtain the multivariate control statistic as

$$W_t = \text{tr}(A_t \Sigma_0^{-1}) - n \ln |A_t| + n \ln |\Sigma_0| + np \ln n - np, \quad (2.4)$$

where $A_t = \sum_{j=1}^n (\underline{X}_{tj} - \mu_0)(\underline{X}_{tj} - \mu_0)'$.

Therefore, we take the statistics \underline{Y}_t and W_t as control statistics to monitor s correlation coefficients of p quality characteristics.

3. Multivariate EWMA Chart

Multivariate EWMA chart based on the vector (2.3) can be expressed as

$$\underline{Y}_t = \sum_{k=1}^t \Lambda (I - \Lambda)^{t-k} \underline{Z}_k + (I - \Lambda)^t \underline{Y}_0, \quad (3.1)$$

where $\underline{Z}_k' = (Z_{k12}, Z_{k13}, \dots, Z_{k1p}, Z_{k23}, \dots, Z_{k2p}, \dots, Z_{k,p-1,p})$, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_s)$ and $0 < \lambda_j \leq 1$ ($j = 1, 2, \dots, s$).

Unless there is any reason to differently weight the elements of smoothing matrix Λ , all diagonal elements of Λ can be set to an equal value. Prabhu and Runger (1997) stated that good choices for λ depend on the number of variables in the control scheme and the size of

the shift in MEWMA chart for $\underline{\mu}$ and they stated that values for λ from 0.1 to 0.5 are good choices. Under the assumption that $\lambda_1 = \lambda_2 = \dots = \lambda_s = \lambda$, the multivariate EWMA vector in (3.1) can be written as

$$\begin{aligned}\underline{Y}_t &= (1-\lambda) \underline{Y}_{t-1} + \lambda \underline{Z}_t \\ &= \sum_{k=1}^t \lambda(1-\lambda)^{t-k} \underline{Z}_k + (1-\lambda)^t \underline{Y}_0.\end{aligned}\quad (3.2)$$

A multivariate EWMA chart for correlation coefficients signals whenever

$$T^2 = \underline{Y}_t' \Sigma_{Y_t}^{-1} \underline{Y}_t > h.$$

The parameter h can be obtained to satisfy a specified in-control ARL and the dispersion matrix Σ_{Y_t} is given in Kim and Chang(1998) as

$$\Sigma_{Y_t} = \left\{ \frac{\lambda[1-(1-\lambda)^{2t}]}{2-\lambda} \right\} \cdot \Sigma_Z \quad (3.3)$$

and

$$\Sigma_Z = \begin{pmatrix} \text{Var}(Z_{n2}) & \text{Cov}(Z_{n2}, Z_{n3}) & \cdots & \text{Cov}(Z_{n2}, Z_{t,p-1,p}) \\ & \text{Var}(Z_{n3}) & \cdots & \text{Cov}(Z_{n3}, Z_{t,p-1,p}) \\ & & \ddots & \vdots \\ & & & \text{Var}(Z_{t,p-1,p}) \end{pmatrix}, \quad (3.4)$$

where

$$\text{Var}(Z_{tpq}) = \frac{1 + \rho_{pq0}^2}{n}$$

$$\text{Cov}(Z_{tpq}, Z_{tpr}) = \frac{\rho_{qr0} + \rho_{pq0}\rho_{pr0}}{n}$$

$$\text{Cov}(Z_{tpq}, Z_{trw}) = \frac{\rho_{pr0}\rho_{qw0} + \rho_{pw0}\rho_{qr0}}{n}$$

and the subscripts p, q, r and w are different each other.

Theorem 3.1 Let the p -component vectors $\underline{X}_1, \underline{X}_2, \dots$ be independent and identically distributed according to $N_p(\underline{\mu}_0, \Sigma_0)$ and $\underline{Y}_0 = \underline{0}$. Then $\{ \Sigma_{Y_t}^{-1/2} \cdot \underline{Y}_t, t \geq 1 \}$ converges in distribution to a multivariate normal distribution with mean vector $\underline{0}$ and variance-covariance matrix I_s as $t \rightarrow \infty$, $\lambda \rightarrow 0$ and $t\lambda \rightarrow 1$.

Proof For $t \geq 1$, let

$$A_t = \frac{1}{t} \Sigma_{Y_t} = \frac{1}{t} \sum_{i=1}^t \text{Cov}\{\lambda(1-\lambda)^{t-i} \underline{Z}_i\} = \frac{1}{t} \left\{ \frac{\lambda[1-(1-\lambda)^{2t}]}{2-\lambda} \right\} \cdot \Sigma_Z$$

$$B_t = \text{the symmetric positive definite matrix satisfying } B_t^2 = A_t^{-1}$$

and

$$\gamma_t = \text{the smallest eigenvalue of } A_t.$$

Then $\Sigma^{-1/2} \mathbf{Y}_t \cdot \mathbf{Y}_t$ and γ_t can be expressed as

$$\Sigma^{-1/2} \mathbf{Y}_t \cdot \mathbf{Y}_t = t^{-1/2} B_t \sum_{i=1}^t \lambda(1-\lambda)^{t-i} \mathbf{Z}_i \quad (3.5)$$

$$\gamma_t = \left\{ \frac{1}{t} \frac{\lambda[1-(1-\lambda)^{2t}]}{2-\lambda} \right\} \gamma, \quad (3.6)$$

where γ is the smallest eigenvalue of Σ_Z .

To verify the asymptotic normality of $\Sigma^{-1/2} \mathbf{Y}_t \cdot \mathbf{Y}_t$, we need only to show that the following Liapounov's condition is satisfied. The condition is given by

$$\Theta_t(\lambda) = t^{-3/2} \sum_{i=1}^t E \|B_t \lambda(1-\lambda)^{t-i} \mathbf{Z}_i\|^3 \rightarrow 0 \text{ as } t \rightarrow \infty, \lambda \rightarrow 0, t\lambda \rightarrow 1. \quad (3.7)$$

(See Bhattacharya and Rao(1976,p185))

From the equation (17.63) given in Bhattacharya and Rao(1976,p177), we have

$$\|B_t \mathbf{Z}_i\|^3 \leq \gamma_t^{-3/2} \|\mathbf{Z}_i\|^3 \quad (3.8)$$

and thus

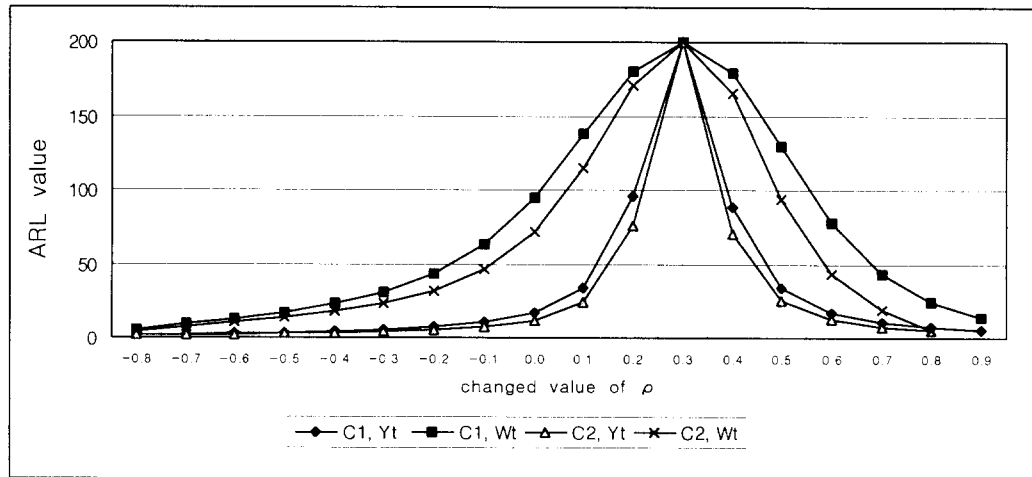
$$\begin{aligned} \Theta_t(\lambda) &= t^{-3/2} \lambda^3 \sum_{i=1}^t (1-\lambda)^{3(t-i)} E \|B_t \mathbf{Z}_i\|^3 \\ &\leq t^{-3/2} \lambda^3 \sum_{i=1}^t (1-\lambda)^{3(t-i)} \gamma_t^{-3/2} E \|\mathbf{Z}_i\|^3 \\ &\leq \left\{ \frac{\lambda[1-(1-\lambda)^{2t}]}{2-\lambda} \gamma \right\}^{-3/2} \lambda^3 \sum_{i=1}^t (1-\lambda)^{3(t-i)} E \|\mathbf{Z}_i\|^3 \\ &= \left\{ \frac{2-\lambda}{1-(1-\lambda)^{2t}} \right\}^{3/2} \left(\frac{\lambda}{\gamma} \right)^{3/2} (1-\lambda)^{3t} \sum_{i=1}^t (1-\lambda)^{-3i} E \left\{ \sum_{l=1}^s \mathbf{Z}_{il}^2 \right\}^{3/2} \\ &\leq \left\{ \frac{2-\lambda}{1-(1-\lambda)^{2t}} \right\}^{3/2} \left(\frac{\lambda}{\gamma} \right)^{3/2} (1-\lambda)^{3t} \sum_{i=1}^t (1-\lambda)^{-3i} s^{1/2} \sum_{l=1}^s E |\mathbf{Z}_{il}|^3 \\ &= \left\{ \frac{2-\lambda}{1-(1-\lambda)^{2t}} \right\}^{3/2} \left(\frac{\lambda}{\gamma} \right)^{3/2} (1-\lambda)^{3t} \left\{ \frac{(1-\lambda)^{-3t}-1}{\lambda(\lambda^2-3\lambda+3)} \right\} s^{3/2} E |\mathbf{Z}_{il}|^3 \\ &= \left\{ \frac{2-\lambda}{1-(1-\lambda)^{2t}} \right\}^{3/2} \left(\frac{s}{\gamma} \right)^{3/2} \lambda^{1/2} \left\{ \frac{1-(1-\lambda)^{3t}}{\lambda^2-3\lambda+3} \right\} E |\mathbf{Z}_{il}|^3. \end{aligned} \quad (3.9)$$

Let $m_3 = E |\mathbf{Z}_{il}|^3 < \infty$, $l=1, \dots, s$. Then the last quantity in (3.9) goes to zero as $t \rightarrow \infty$, $\lambda \rightarrow 0$ and $t\lambda \rightarrow 1$.

Corollary 3.1.1 $\{T^2, t \geq 1\}$ converges in distribution to a chi-squared distribution with s degrees of freedom as $t \rightarrow \infty$, $\lambda \rightarrow 0$ and $t\lambda \rightarrow 1$.

Proof Since the control statistic T^2 can be expressed as

$$T^2 = (\Sigma^{-1/2} \mathbf{Y}_t \cdot \mathbf{Y}_t)' (\Sigma^{-1/2} \mathbf{Y}_t \cdot \mathbf{Y}_t).$$

[Figure 1] ARL performances for $p=3$ ($\rho_0=0.3$)

From the results of Theorem 4.1 and Corollary of Serfling(1980,p25)

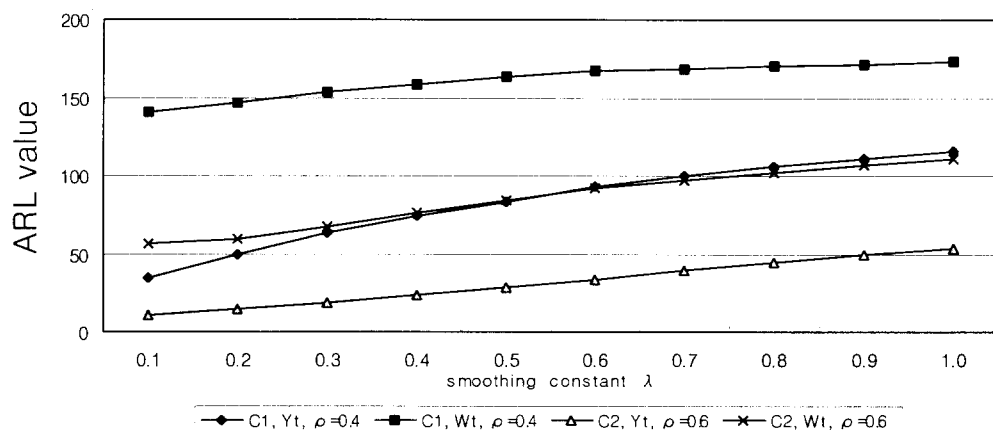
$$T^2 = \underline{Y}_t' \Sigma_{X_t}^{-1} \underline{Y}_t \xrightarrow{d} \chi^2(s) \text{ as } k \rightarrow \infty, \lambda \rightarrow 0, k\lambda \rightarrow 1.$$

This completes the proof.

A multivariate EWMA chart based on LRT statistic in (2.4) is given by

$$Y_{W,t} = (1-\lambda)Y_{W,t-1} + \lambda W_t, \quad (3.10)$$

where $Y_{W,0} = \omega \cdot I_{(\omega \geq 0)}$ and $0 < \lambda \leq 1$. This chart signals if $Y_{W,t}$ exceeds an upper control limit(UCL) h_W , and the parameter h_W can be obtained to satisfy a specified in-control ARL.

[Figure 2] Trend of ARL performance according to λ ($p=3, \rho_0=0.2$)

4. Computational Results and Conclusion

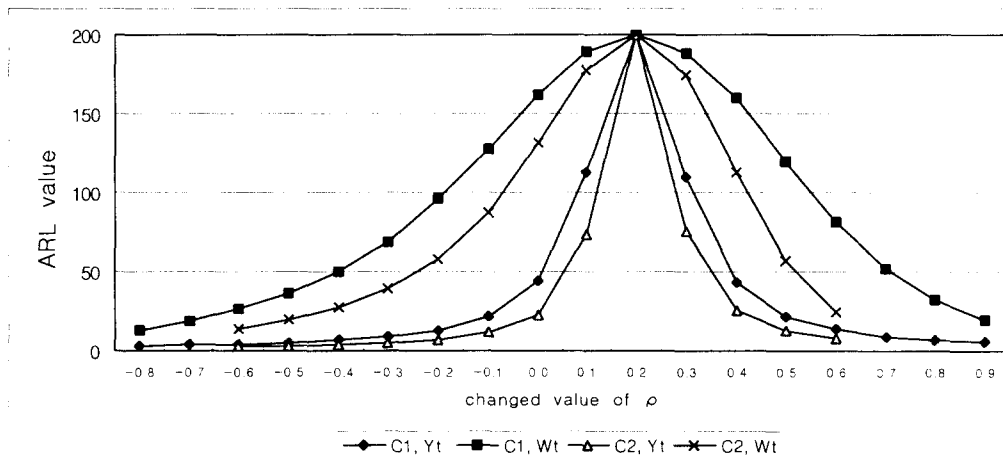
We propose multivariate EWMA charts to monitor correlation coefficients of correlated quality variables under multivariate normal process. Since it is difficult to obtain the distribution of \underline{Y}_t and $Y_{W,t}$, the process parameters h , h_w and ARL performances of the proposed charts were evaluated by simulation with 10,000 iterations.

For simplicity, it is assumed that in-control process mean vector is $\underline{\mu}_0 = \underline{0}'$, all diagonal elements of Σ_0 are 1 and off-diagonal elements of Σ_0 are 0.2 or 0.3. To compare the properties and performances of the charts, the charts should be set up so that both have the same ARL when the process is in-control. In our computation, the values in [Figure 1] through [Figure 4] were obtained when ARL of the in-control state was approximately fixed to be 200 and the sample size $n = 5$ on each sampling period for $p = 3$ and 4.

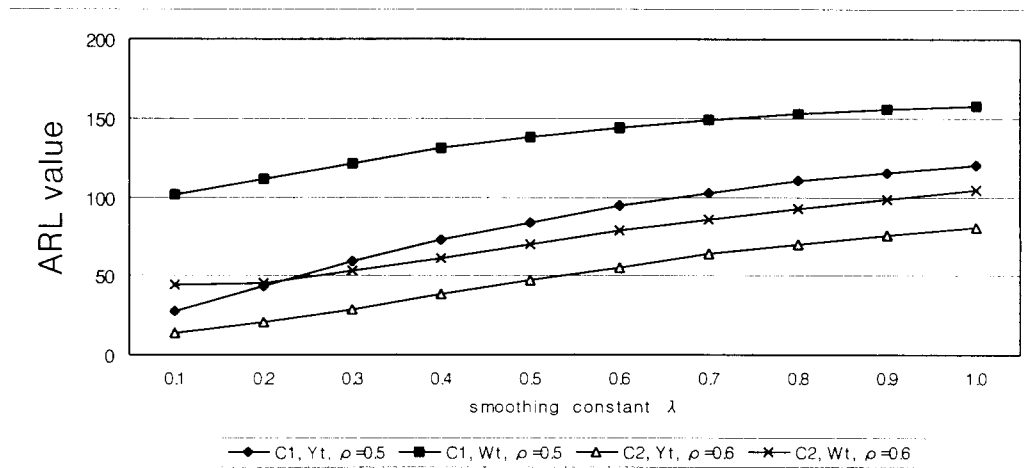
The types of shifts in the process parameters when the state is out-of-control is stated as follows :

- (C1) ρ_{12} and ρ_{21} have changed
- (C2) ρ_{1j} and ρ_{j1} ($j = 2, \dots, p$) have changed

Numerical results show that the chart based on \underline{Y}_t in (3.2) with accumulate-combine approach is more efficient than the chart based on W_t in (3.10) with combine-accumulate approach in terms of ARL performance. And we also found that smaller values for λ are preferred for detecting smaller shifts in multivariate EWMA chart and vice versa for various



[Figure 3] ARL performances for $p = 4$ ($\rho_0 = 0.2$)



[Figure 4] Trend of ARL performance according to λ ($p=4, \rho_0=0.3$)

p . The ARL performances for out-of-control state were stated in [Figure 1] and [Figure 3] when p is 3 and 4. In [Figure 2] and [Figure 4], we can see the trends of ARL performances according to smoothing constant λ when the process is out-of-control.

References

- [1] Alt, F. B. (1984). Multivariate Control Charts, in *The Encyclopedia of statistical Sciences*, edited by S. Kotz and N.L. Johnson, John Wiley, New York.
- [2] Bhattacharya, R. N. and Rao, R. R. (1976). *Normal Approximation and Asymptotic Expansions*, John Wiley & Sons, New York.
- [3] Crosier, R. B. (1988). "Multivariate Generalization of Cumulative Sum Quality-Control Scheme," *Technometrics*, Vol. 30, 291-303.
- [4] Hotelling H. (1947). "Multivariate Quality Control, *Techniques of Statistical Analysis*," McGraw-Hill, New York, 111-184.
- [5] Jackson, J. E. (1985). "Multivariate Quality Control," *Communications in Statistics-Theory and Method*, Vol. 14, 2657-2688.
- [6] Kim J. J. and Chang D. J. (1998). "Control Chart for Correlation Coefficients of Correlated Quality Variables," *Journal of the Korean Society for Quality Management*, Vol. 26, 51-66.
- [7] Lowry, C. A., Woodall, W. H., Champ, C. W. and Rigdon, S. E. (1992). "A Multivariate Exponentially Weighted Moving Average Control Charts," *Technometrics*, Vol. 34, 46-53.

- [8] Mason, R. L., Champ, C. W., Tracy, N. D., Wierda, S. J. and Young, J. C. (1997). "Assessment of Multivariate Process Control Techniques," *Journal of Quality Technology*, Vol. 29, 140-143.
- [9] Pignatiello, J. J., Jr. and Runger, G. C. (1990). "Comparisons of Multivariate CUSUM Charts," *Journal of Quality Technology*, Vol. 22, 173-186.
- [10] Prabhu, S. S. and Runger, G. C. (1997). "Designing a Multivariate EWMA Control Chart," *Journal of Quality Technology*, Vol. 29, 8-15.
- [11] Serfling, R. J. (1976). *Approximation Theorems of Mathematical Statistics*, John Wiley & Sons, New York.