⊠연구논문

# Control Chart for Correlation Coefficients of Correlated Quality Variables

Jae-Joo Kim
Dept. of Statistics, Seoul National University
Duk-Joon Chang
Dept. of Statistics, Changwon National University

## Abstract

Exponentially weighted moving average(EWMA) control chart to simultaneously monitor correlation coefficients of several correlated quality variables under multivariate normal process are proposed. Performances of the proposed control charts are measured in terms of average run length(ARL) by simulation. Numerical results show that smaller values of smoothing constant are more efficient in terms of ARL.

## 1. Introduction

Control charts are used for continuously monitoring the production process to quickly detect and eliminate assignable causes that may produce any deterioration in the quality of the product. The ability of a control chart to detect process changes is determined by the length of time required for the chart to signal when the process is out-of-control, and frequency of false alarm. Therefore a good control chart should quickly detect changes in the production process while producing few false alarms.

In many industries, there exist multiple variables to define the quality of output. And increasing in correlation coefficients of correlated quality variables is often important when the strength of linear relationship between two or more variables effect the quality. Especially in food production process such as snacks, change of correlation coefficient between shade and flavor variables largely affects the quality of the output.

The EWMA chart, first introduced by Roberts(1959), has approximately equivalent performances to the CUSUM chart and has a good ability when we are interested in detecting small or moderate shifts. And the original work on multivariate chart was introduced by Hotelling(1947). Alt(1982) and Jackson(1985) reviewed much of the articles on multivariate chart.

Lowry et al.(1992) presented a multivariate EWMA(MEWMA) chart for mean vector with accumulate-combine technique. 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), and stated that MEWMA procedure is easy to design and implement. Cho(1991) presented a multivariate EWMA chart for monitoring the variance components of dispersion matrix. Lucas and Saccucci(1990) discussed the choice of design parameters for the univariate EWMA chart in detail. Prabhu and Runger(1997) stated that good choices for smoothing constant $\lambda$ depend on the number of variables in the control scheme and the size of the shift in multivariate EWMA charts.

Up to the present, we cannot found the works on control charts for monitoring correlation coefficients of correlated quality variables in the quality control literatures. In this paper, we present an EWMA control chart to simultaneously monitor correlation coefficients of several correlated quality characteristics under multivariate normal process.

## 2. Univariate Control Chart

Assume that the quality vector of interest is $\underline{X}=(X_1, X_2, \cdots, X_p)'$ and we take a sequence of random vectors $\underline{X}_1, \underline{X}_2, \cdots$, where $\underline{X}_t=(\underline{X}'_{t1}, \cdots, \underline{X}'_{tn})'$ is a sample of observations at the sampling period $t$ and $\underline{X}_{tj}=(X_{tj1}, \cdots, X_{tjp})'$. It will be also assumed that the successive observation vectors are distributed independent multivariate normal distribution with $N_p(\underline{\mu}, \Sigma)$.

Let $\underline{\theta}_o=(\underline{\mu}_0, \Sigma_0)$ be the known target process parameters for $\underline{\theta}=(\underline{\mu}, \Sigma)$ of $p$ related quality variables. For simplicity, it is assumed that $\underline{\mu}_0=\underline{0}'$, all diagonal and off-diagonal elements of $\Sigma_0$ are 1 and 0.2, respectively. In this section, we consider a univariate EWMA chart for correlation coefficient $\rho_{12}$ of two quality variables $X_1$ and $X_2$.

Stuart(1955) proposed an estimator of the correlation parameter $\rho_{12}$ for the bivariate normal population $N_2(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho_{12})$ as

$$r_{12} = \sum_{j=1}^{n} \frac{(x_j - \mu_1)(y_j - \mu_2)}{n\sigma_1\sigma_2}, \tag{2.1}$$

where the other four parameters being known.

An EWMA chart based on the control statistic $r_{12}$ in (2.1) for $\rho_{12}$ can be constructed by using the statistic

$$Y_t = (1-\lambda)Y_{t-1} + \lambda \frac{\sum_{j=1}^{n}(X_{tj1} - \mu_{10})(X_{tj2} - \mu_{20})}{n\sigma_{10}\sigma_{20}}, \tag{2.2}$$

$t = 1, 2, 3, \cdots$, where $0 < \lambda \le 1$. By repeated substititution in (2.2), it can be shown that

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

If the process is in-control with $Y_0 = \omega_0$ and $Z_1, Z_2, \cdots$ are independent, we can obtain the mean and variance of $Y_t$ as

$$E(Y_t) = (1-\lambda)^t \omega_0 + [1-(1-\lambda)^t]\rho_{120}$$

and

$$V(Y_t) = \frac{\lambda[1-(1-\lambda)^{2t}]}{2-\lambda} \cdot \frac{(1+\rho_{120})^2}{n}.$$

This chart signals whenever $Y_t \le h_1$ or $Y_t \ge h_2$ $(h_1 < h_2)$. The parameters $h_1$ and $h_2$ can be obtained to satisfy a specified ARL by simulation.

## 3. Multivariate Control Chart

Multivariate control chart for monitoring correlation coefficients of quality variables can be constructed by forming multivariate extension of the univariate

EWMA chart statistic in the previous section. In multivariate case, if we let the control statistic for $\rho_{lm}$ be $r_{lm}$, then the vectors of EWMA's can be defined as

$$\begin{aligned} \underline{Y}_t' &= (r_{12}, r_{13}, \cdots, r_{1p}, r_{23}, \cdots, r_{2p}, \cdots, r_{p-1,p-2}, r_{p-1,p})' \\ &= (Y_{t1}, Y_{t2}, \cdots, Y_{t,p-1}, Y_{tp}, \cdots, Y_{t,2p-3}, \cdots, Y_{t,s-1}, Y_{t,s})'. \end{aligned} \tag{3.1}$$

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

$$\underline{Y}_t = \begin{bmatrix} (1-\lambda_1)^t Y_{t10} + \sum_{k=1}^{t} \lambda_1 (1-\lambda_1)^{t-k} Z_{k12} \\ \vdots \\ (1-\lambda_{p-1})^t Y_{t,p-1,0} + \sum_{k=1}^{t} \lambda_{p-1} (1-\lambda_{p-1})^{t-k} Z_{k1p} \\ (1-\lambda_p)^t Y_{t,p,0} + \sum_{k=1}^{t} \lambda_p (1-\lambda_p)^{t-k} Z_{k23} \\ \vdots \\ (1-\lambda_{2p-3})^t Y_{t,2p-3,0} + \sum_{k=1}^{t} \lambda_{2p-3} (1-\lambda_{2p-3})^{t-k} Z_{k2p} \\ \vdots \\ (1-\lambda_s)^t Y_{t,s,0} + \sum_{k=1}^{t} \lambda_s (1-\lambda_s)^{t-k} Z_{k,p-1,p} \end{bmatrix},$$

$t = 1, 2, 3 \cdots$, where $s = p(p-1)/2$ and $0 < \lambda_a \leq 1$ $(a = 1, 2, \cdots, s)$.

Then the multivariate EWMA vector can be expressed as

$$\underline{Y}_t = \sum_{k=1}^{t} \Lambda (I-\Lambda)^{t-k} \underline{Z}_k + (I-\Lambda)^t \underline{Y}_0 , \tag{3.2}$$

where

$$\begin{aligned} \underline{Z}_k &= (Z_{k1}, Z_{k2}, \cdots, Z_{ks})' \\ &= (Z_{k12}, Z_{k13}, \cdots, Z_{k1p}, Z_{k23}, \cdots, Z_{k2p}, \cdots, Z_{k,p-1,p})', \end{aligned}$$

and

$$Z_{kmu} = \frac{\sum_{j=1}^{n} (X_{kjm} - \mu_{m0})(X_{kju} - \mu_{u0})}{n \sigma_{m0} \sigma_{u0}} - \rho_{mu0} ,$$

$\Lambda = diag(\lambda_1, \cdots, \lambda_s)$, $0 < \lambda_j \leq 1$ $(j = 1, 2, \cdots, s)$, $m \neq u$ .

If there is no distinct reason to take different smoothing constants for $s$ correlation coefficients being monitored, all diagonal elements of the matrix $\Lambda$ can be set equal values. Under the assumption that $\lambda_1=\lambda_2=\cdots=\lambda_s=\lambda$, the multivariate EWMA vectors in (3.2) can be written as

$$\begin{aligned} \underline{Y}_t &= (1-\lambda)\,\underline{Y}_{t-1}+\lambda\,\underline{Z}_t \\ &= \sum_{k=1}^{t}\lambda(1-\lambda)^{t-k}\,\underline{Z}_k+(1-\lambda)^t\,\underline{Y}_0 . \end{aligned} \tag{3.3}$$

A multivariate EWMA chart for $\underline{\rho}=(\rho_{12},\rho_{13},\cdots,\rho_{1p},\rho_{23},\cdots,\rho_{2p},\cdots,\rho_{p-1,p})'$ signals as soon as

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

where $\Sigma_{\underline{Y}_t}$ is the variance-covariance matrix of $\underline{Y}_t$ with dimension $s\times s$ and is given in (3.5) of Theorem 3.2. below. The parameter $h$ can be obtained to satisfy a specified ARL by simulation.

**Theorem 3.1** If the process is in-control and $\underline{Y}_0=\underline{0}$, then $E(\underline{Y}_t)=\underline{0}$.

**Proof** It is easy to show that

$$\begin{aligned} &E(Z_{kmu}\,|\mu_l=\mu_{l0},\sigma_l=\sigma_{l0},\rho_{mu}=\rho_{mu0}) \\ &=E\Big[\frac{\sum_{j=1}^{n}(X_{kjm}-\mu_{m0})(X_{kju}-\mu_{u0})}{n\sigma_{m0}\sigma_{u0}}-\rho_{mu0}\,\Big|\,\mu_l=\mu_{l0},\sigma_l=\sigma_{l0},\rho_{mu}=\rho_{mu0}\Big] \\ &=0 \qquad l=m,u. \end{aligned}$$

Thus, given that $\underline{\mu}=\underline{\mu}_0$ and $\Sigma=\Sigma_0$, the expected value of the random vector $\underline{Y}_t$ is

$$\begin{aligned} E[\underline{Y}_t|\underline{\mu}=\underline{\mu}_0,\Sigma=\Sigma_0] &= \sum_{k=1}^{t}\lambda(1-\lambda)^{t-k}E[\underline{Z}_k|\underline{\mu}=\underline{\mu}_0,\Sigma=\Sigma_0] \\ &= \underline{0}. \end{aligned}$$

□

**Theorem 3.2** If the process is in-control and $\underline{Y}_0=\underline{0}$, then the dispersion matrix of $\underline{Y}_t$ is given by

$$\Sigma_{Y_t} = \left\{ \frac{\lambda[1-(1-\lambda)^{2t}]}{2-\lambda} \right\} \cdot \Sigma_Z \tag{3.4}$$

and

$$\Sigma_Z = \begin{pmatrix} Var(Z_{t12}) & Cov(Z_{t12}, Z_{t13}) & \cdots & Cov(Z_{t12}, Z_{t,p-1,p}) \\ & Var(Z_{t13}) & \cdots & Cov(Z_{t13}, Z_{t,p-1,p}) \\ & & \ddots & \vdots \\ & Sym & & Var(Z_{t,p-1,p}) \end{pmatrix}, \tag{3.5}$$

where

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

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

$$Cov(Z_{tpq}, Z_{trw}) = \frac{\rho_{pr0}\rho_{qu0} + \rho_{pu0}\rho_{qr0}}{n}$$

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

**Proof**

$$\Sigma_{Y_t} = \sum_{k=1}^{t} Cov[\lambda(1-\lambda)^{t-k} \underline{Z}_k]$$

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

To show the form of $\Sigma_Z$, it is necessary to derive that the following results by using the moment generating function(MGF) of quality vectors under multivariate normal distribution when the process is in-control such that

$$Var(Z_{tpq}) = \frac{1}{n^2\sigma^2_{p0}\sigma^2_{q0}} Var[\sum_{j=1}^{n}(X_{tjp}-\mu_{p0})(X_{tjq}-\mu_{q0})]$$

$$= \frac{1}{n(\sigma_{p0}\sigma_{q0})^2}(\sigma^2_{p0}\sigma^2_{q0} + \rho^2_{pq0}\sigma^2_{p0}\sigma^2_{q0})$$

$$= \frac{1+\rho^2_{pq0}}{n}$$

$$
\begin{aligned}
Cov(Z_{pq}, Z_{pr}) &= Cov\left[\frac{\sum_{j=1}^{n}(X_{tjp}-\mu_{p0})(X_{tjq}-\mu_{q0})}{n\sigma_{p0}\sigma_{q0}}, \frac{\sum_{j=1}^{n}(X_{tjp}-\mu_{p0})(X_{tjr}-\mu_{r0})}{n\sigma_{p0}\sigma_{r0}}\right] \\
&= \frac{1}{n\sigma_{p0}^{2}\sigma_{q0}\sigma_{r0}} Cov[\ (X_{tjp}-\mu_{p0})(X_{tjq}-\mu_{q0})\ ,\ (X_{tjp}-\mu_{p0})(X_{tjr}-\mu_{r0})\ ] \\
&= \frac{\rho_{qr0}+\rho_{pq0}\rho_{pr0}}{n}
\end{aligned}
$$

and

$$
\begin{aligned}
Cov(Z_{pq}, Z_{rw}) &= \frac{1}{n\sigma_{p0}\sigma_{q0}\sigma_{r0}\sigma_{u0}} Cov[\ (X_{tjp}-\mu_{p0})(X_{tjq}-\mu_{q0})\ ,\ (X_{tjr}-\mu_{r0})(X_{tjw}-\mu_{u0})\ ] \\
&= \frac{\rho_{pr0}\rho_{qu0}+\rho_{pu0}\rho_{qr0}}{n} .
\end{aligned}
$$

□

## 4. Computational Results and Conclusion

In this paper, we propose EWMA charts for monitoring correlation coefficients of correlated quality variables under multivariate normal process. Since it is difficult to obtain the distribution of (2.2) and (3.3), the parameter $h$ and ARL values of the charts can be evaluated by simulation with 10,000 iterations under the parameter of the process is on-target or changed.

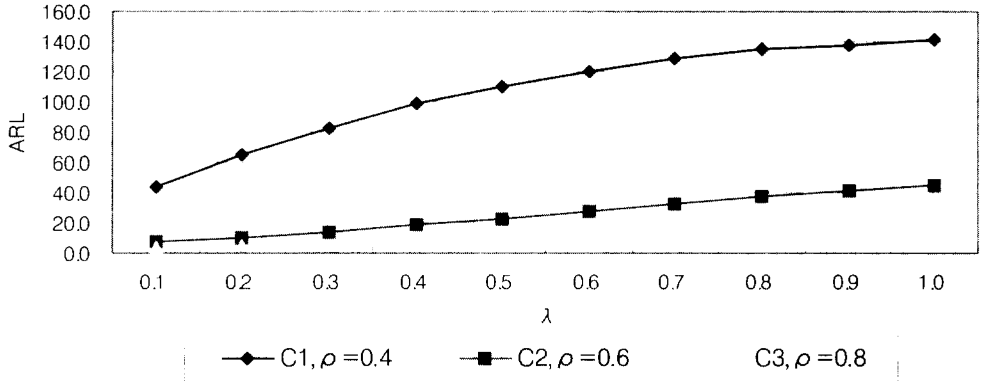
In our computation, the results were obtained when ARL of the in-control state was approximately fixed to be 200 and the sample size for each variable was 5 for p=2 and p=4. The types of shifts in the process parameters when the process is out-of-control is stated as follows :

(C1) $\rho_{12}$ and $\rho_{21}$ have increased

(C2) $\rho_{1i}$ and $\rho_{i1}(i=2,\cdots,p)$ have increased

(C3) all $\rho_{ij}{'}s(i,j=1,\cdots,p$ and $i\neq j)$ have increased

Numerical results shows that smaller values of smoothing constant are more effective in detecting any scale of changes for both univariate and multivariate EWMA chart for various $p$. In <Figure 1>, we can see the trends of ARL performances according to $\lambda$ when the process is out-of-control. The ARL values of the out-of-control state were stated in <Table 1> through <Table 3> when $p$ is 2 and 4.




< Figure 1 > Trends of ARL performances according to $\lambda$ ( $p=4$, $\rho_0=0.2$)

< Table 1 > ARL values for $p=2$ ($\rho_o=0.2$)

| | $\lambda=0.1$ | $\lambda=0.2$ | $\lambda=0.3$ | $\lambda=0.6$ | $\lambda=1.0$ |
|---|---|---|---|---|---|
| in-control | 199.8 | 200.0 | 200.0 | 200.0 | 200.0 |
| $\rho$ = 0.3 | 86.9 | 109.3 | 124.9 | 153.4 | 171.9 |
| $\rho$ = 0.4 | 37.4 | 44.1 | 53.7 | 80.8 | 109.5 |
| $\rho$ = 0.5 | 22.0 | 23.1 | 27.4 | 43.8 | 65.4 |
| $\rho$ = 0.6 | 15.3 | 14.6 | 16.6 | 26.6 | 41.1 |
| $\rho$ = 0.7 | 11.8 | 10.5 | 11.3 | 17.1 | 27.6 |
| $\rho$ = 0.8 | 9.5 | 8.1 | 8.3 | 11.9 | 19.5 |
| $\rho$ = 0.9 | 8.0 | 6.7 | 6.6 | 8.8 | 14.2 |

< Table 2 > ARL values for $p=4$ ($\rho_o=0.2$)

| | $C_1$ | | | $C_2$ | | $C_3$ | | |
|---|---|---|---|---|---|---|---|---|
| | $\lambda=0.1$ | $\lambda=0.3$ | $\lambda=1.0$ | $\lambda=0.1$ | $\lambda=0.3$ | $\lambda=0.1$ | $\lambda=0.3$ | $\lambda=1.0$ |
| in-control | 200.0 | 200.1 | 200.0 | 200.0 | 200.1 | 200.0 | 200.1 | 200.0 |
| $\rho$ = 0.3 | 110.2 | 150.0 | 180.2 | 75.2 | 116.7 | 59.1 | 91.3 | 141.9 |
| $\rho$ = 0.4 | 43.6 | 82.9 | 141.2 | 25.1 | 51.6 | 20.6 | 34.5 | 78.0 |
| $\rho$ = 0.5 | 21.8 | 44.0 | 102.9 | 12.4 | 24.7 | 11.0 | 16.9 | 43.4 |
| $\rho$ = 0.6 | 13.4 | 25.2 | 72.8 | 7.7 | 13.7 | 7.3 | 10.1 | 26.2 |
| $\rho$ = 0.7 | 9.3 | 16.1 | 51.2 | - | - | 5.3 | 7.1 | 17.7 |
| $\rho$ = 0.8 | 7.0 | 11.1 | 36.9 | - | - | 4.2 | 5.3 | 12.6 |
| $\rho$ = 0.9 | 5.5 | 8.2 | 27.2 | - | - | 3.5 | 4.3 | 9.4 |

< Table 3 > ARL values according to $\rho_0$ ($p=4$, $C_1$)

| | $\rho_0=0.1$ | | | $\rho_0=0.3$ | | $\rho_0=0.5$ | | |
|---|---|---|---|---|---|---|---|---|
| | $\lambda=0.1$ | $\lambda=0.2$ | $\lambda=0.4$ | $\lambda=0.1$ | $\lambda=0.4$ | $\lambda=0.1$ | $\lambda=0.2$ | $\lambda=0.4$ |
| $\rho$ = 0.1 | 200.0 | 200.0 | 200.0 | - | - | - | - | - |
| $\rho$ = 0.2 | 112.5 | 138.3 | 162.5 | - | - | - | - | - |
| $\rho$ = 0.3 | 45.8 | 68.4 | 101.0 | 200.0 | 200.0 | - | - | - |
| $\rho$ = 0.4 | 23.1 | 34.4 | 58.9 | 104.2 | 158.9 | - | - | - |
| $\rho$ = 0.5 | 14.1 | 19.7 | 34.9 | 39.5 | 93.7 | 200.0 | 200.0 | 200.0 |
| $\rho$ = 0.6 | 9.7 | 12.9 | 22.1 | 19.7 | 51.9 | 82.8 | 116.3 | 149.4 |
| $\rho$ = 0.7 | 7.3 | 9.2 | 15.1 | 12.2 | 29.9 | 27.5 | 42.9 | 75.9 |
| $\rho$ = 0.8 | 5.7 | 7.0 | 10.9 | 8.4 | 18.8 | 13.7 | 19.7 | 36.6 |
| $\rho$ = 0.9 | 4.7 | 5.6 | 8.3 | 6.4 | 12.7 | 8.6 | 11.3 | 20.0 |

# 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] Cho, G.Y.(1991), "A Multivariate Control Charts for the Mean Vector and Variance-Covariance Matrix with Variable Sampling Intervals," Ph. D. dissertation, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

[3] Crosier, R.B.(1988), "Multivariate Generalization of Cumulative Sum Quality-Control Scheme," *Technometrics,* Vol. 30, pp. 291-303.

[4] Hotelling H.(1947), "Multivariate Quality Control, *Techniques of Statistical Analysis,*" McGraw-Hill, New York, pp. 111-184.

[5] Jackson, J.E.(1985), "Multivariate Quality Control," *Communications in Statistics-Theory and Method,* Vol. 14, pp. 2657-2688.

[6] 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, pp. 46-53.

[7] Lucas, J.M. and Saccucci, M.S.(1990), "Exponentially Weighted Moving Average Control Schemes:Properties and Enhencements," *Technometrics*, Vol. 32, pp. 1-12.

[8] Pignatiello, J.J., Jr. and Runger, G.C.(1990), "Comparisons of Multivariate CUSUM Charts," *Journal of Quality Technology,* Vol. 22, pp. 173-186.

[ 9 ] Prabhu, S.S. and Runger, G.C. (1997), "Designing a Multivariate EWMA Control Chart," *Journal of Quality Technology,* Vol. 29, pp. 8-15.

[10] Roberts, S.W.(1959), "Control Chart Tests based on Geometric Moving Averages," *Technometrics,* Vol. 1, pp. 239-250.

[11] Stuart, A.(1955), "A Paradox in Statistical Estimation," *Biometrika,* Vol. 42, pp. 527-529.