

A Transformation Technique in Designing Multi-Attribute C Control Charts

S.T. Akhavan Niaki* and B. Abbasi¹

In a production process, when the quality of a product depends on more than one characteristic, multivariate quality control techniques are used. Although multivariate statistical process control is receiving increased attention in the literature, little work has been done to deal with multi-attribute processes. In this paper, a new methodology has been developed to monitor multi-attribute processes, in which the defect counts are important and different types of defect are dependent random variables. In order to do this, based on the symmetric square root transformation concept, first, multi-attribute data is transformed, such that the correlation between variables either vanishes or becomes very small. Then, by a simulation and bisection method, the symmetric control limits are found and a symmetric rectangular region is formed for control. In simulation studies, some numerical examples are presented to illustrate the proposed method and to evaluate and compare its performance to the ones of the existing method.

INTRODUCTION AND LITERATURE REVIEW

In many quality control environments, the process or product under consideration has two or more correlated quality characteristics. Today, with modern data-acquisition equipment, sensors and online computers, one can easily monitor these quality characteristics simultaneously. For example, the quality of a chemical process may be a function of process temperature, pressure and flow rate, all of which need to be monitored in a situation where some correlation may exist between them. In these cases, if one wants to monitor these quality characteristics separately, there will be some error associated with the out-of-control detection procedure.

In general, there are two broad categories in statistical control charts, namely; variable and attribute control charts, for which many researchers have developed different methodologies. Early research on multivariate control charts goes back to Hotelling [1], who introduced the problem of correlation between the quality characteristics of a process and came up with the well-known T^2 statistic, to identify whether

the whole process is out of control. Lowery and Montgomery [2] have shown that a multivariate control scheme normally has better sensitivity than the ones based on univariate control charts. Other multivariate control charts are multivariate forms of the Shewhart charts presented by Golnabi and Houshmand [3], the multivariate CUSUM charts, proposed by Woodall and Ncube [4], Healy [5], Lucas and Crosier [6] and Pignatiello and Runger [7] and the Multivariate Exponentially Weighted Moving Average (MEWMA) charts, proposed by Lowry et al. [8]. Moreover, there are some other methods, proposed by Runger [9], Hawkins [10] and Niaki and Abbasi [11].

Despite the fact that multi-attribute monitoring has many applications, almost all researchers have focused on the first category of control charting and only a few methods have been proposed to monitor multi-attribute processes (see e.g. [12]). Furthermore, in many instances where exact measurements are not needed, it is easy to collect correlated discrete-type data. Patel [13] proposed a Hotelling-type χ^2 chart to monitor observations from multivariate binomial or multivariate Poisson distribution (for time independent and time dependent samples). In fact, similar to Shewhart attribute control charts for univariate cases, Patel assumed that, if one chooses an appropriate sample size, the vectors will have multivariate normal distribution. Therefore, it is correct to use the concept of multivariate normal control charting. Lu et al. [14] addressed the statistical design of multi-attribute con-

*. Corresponding Author, Department of Industrial Engineering, Sharif University of Technology, P.O. Box 11155-9414, Tehran, I.R. Iran.

1. Department of Industrial Engineering, Sharif University of Technology, P.O. Box 11365-9414, Tehran, I.R. Iran.

trol charts. They proposed a multivariate np-chart (MNP chart) to develop Shewhart charts based on an X statistic and showed that this statistic reduced type-II errors better than individual np charts, since the correlation between the attributes was taken into account. However, in their research, there was no discussion on the Average Run Length (ARL) of the MNP chart and the distribution of the statistic used in this chart. Jolayemi [15] developed a model for an optimal design of multi-attribute control charts for processes with multiple assignable causes. This model addresses the economic design of control charting and is based on the assumption of independent attributes, a J approximation [16] and Gibra's model [17] for a univariate np chart. When the proportions in each quality category are known or estimated using a base period, Marucci [18] used a multinomial distribution to develop a control chart. Since not all multi-attribute processes follow a multinomial distribution, this method may not be applicable. Larpiattaworn [19] proposed a Back Propagation Neural Network (BPNN) for two-attribute process control in bivariate Binomial and bivariate Poisson cases. He detected the out of control condition by an artificial neural network, where the output was one, if the process was under control, and zero, otherwise. He also discussed different values of correlation between two variables and gave some suggestions on the use of three-attribute control charts (χ^2 , MNP and BPNN method). Gadre and Rattihalli [20], with the assumption of multinomial distribution for multi-attribute processes, used an MP-test to determine if the parameters of the distribution would change or not. In their method, the values of the parameters of interest must be known in advance.

In this paper, a rectangular symmetric region to monitor multi-attribute processes mean in a multivariate C chart, is proposed. This region is reached by, first, employing a transformation method that almost eliminates the correlation between quality characteristics. Then, the control limits of optimal Shewhart-type control charts, for the transformed characteristics with a specific ARL_0 , are obtained using simulation and bisection methods. At the end, the control region, based on the values obtained for control limits, is constructed.

The structure of this paper is as follows. First, a brief background on χ^2 and MNP charts, as existing control charts for multi-attribute processes is presented. Then, the concept of the transformation technique used in this research is explained. After that, the new method is developed. In order to understand the proposed method better, three numerical examples are presented and its performance is evaluated using the Average Run Length (ARL) criterion. Finally, the conclusion and recommendations for future research are given.

NORMAL APPROXIMATION METHOD USED IN MULTI-ATTRIBUTE MONITORING

In this section, the normal approximation of either the multivariate binomial or multivariate Poisson distribution, used in multi-attribute control charting, is briefly explained [13].

Although Patel's method included both time independent and time dependent (auto-correlated) samples, this paper focuses on the time independent case. When sample size, n , is large, the statistic in Equation 1 forms the basics of the control charts used in multi-attribute quality control environments.

$$T^2 = (\mathbf{X} - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X} - \bar{\mathbf{X}}), \quad (1)$$

where T^2 has an approximate Chi-Square distribution with p degrees of freedom, $\mathbf{X}$ is a random vector from a population of interest, p is the number of process attributes and $\mathbf{S}$ is an estimator of the population covariance matrix, Σ , which is assumed to remain unchanged with time. The upper control limit of the control chart equals the critical point of a Chi-Squared distribution with p degrees of freedom, χ_{α}^2 , and α is a specified significance level. The lower control limit is equal to zero. In this method, $\mathbf{X}$ may follow a multivariate binomial or multivariate Poisson distribution.

SYMMETRIC SQUARE ROOT METHOD

In order to eliminate the existing correlation between the quality characteristics in vector $\mathbf{X} = [X_1, X_2, \dots, X_p]^T$, in a symmetric square root transformation method, a new vector, $\mathbf{Y} = [Y_1, Y_2, \dots, Y_p]^T = \mathbf{C}\mathbf{X}$, is obtained, such that Y_i s are almost uncorrelated random variables. Matrix $\mathbf{C}$ is a symmetric matrix which is the square root of Σ , the correlation matrix of $\mathbf{X}$. However, before applying this method, first, $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]^T$ is subtracted from $\mathbf{X}$, and then, the transformation is undertaken [3]. In other words, the transformed vector becomes:

$$\mathbf{Y} = (\Sigma)^{\frac{1}{2}} (\mathbf{X} - \boldsymbol{\mu}). \quad (2)$$

In production processes, in which the quality characteristics are counts on two different defect types and follow a bivariate Poisson distribution, the following numerical example is presented to illustrate the application of this transformation.

Example 1

Suppose vector $\mathbf{X} = [X_1, X_2]^T$ follows a bivariate Poisson distribution, in which the marginal probability mass distributions are Poisson with parameters $\lambda_1 = 4$

and $\lambda_2 = 5$, with covariance equal to 1.75. In this case, using the NORTA method [21,22], first, 5000 observations, on a random vector of size 2 for $\mathbf{X}$, from the above bivariate Poisson distribution, were generated. The estimated mean vector and the covariance matrix of the generated observations were:

$$\widehat{\mu}_{\mathbf{X}} = [4.041, 5.011]^T,$$

and:

$$\widehat{\text{Cov}}(\mathbf{X}) = \begin{pmatrix} 4.111 & 1.752 \\ 1.752 & 4.883 \end{pmatrix}.$$

Then, applying Equation 2 to the generated observations yields:

$$\widehat{\mu}_{\mathbf{Y}} = [0.0103, 0.0003],$$

and:

$$\widehat{\text{Cov}}(\mathbf{Y}) = \begin{pmatrix} 0.2524 & 0.0002 \\ 0.0002 & 0.1981 \end{pmatrix},$$

on the transformed vector, in which it is noted that the covariance is close to zero.

SYMMETRIC MULTI-ATTRIBUTE CONTROL CHARTS FOR TRANSFORMED DATA

One may obtain a symmetric rectangular region for the transformed data using some simple Shewhart-type control charts for each uncorrelated transformed variable. However, since the probability distribution of the transformed vector is unknown, symmetric control limits are sought, such that an overall in-control average run length, ARL_0 , becomes close to a pre-specified value. To do this, the bisection method was applied to the data generated by simulation.

As an example, for a specified value of $ARL_0 = 200$ (i.e. $\alpha = 0.005$), for a multi-attribute process monitoring, with the assumption of independent control intervals, one needs to determine an ARL_{0i} value for each transformed variable, such that the overall average run length becomes 200. If one defines α_i to be a type-I error associated with the i th control interval on the i th transformed quality characteristic, then, one will have:

$$\alpha_i = 1 - \sqrt[p]{1 - \alpha}; \quad i = 1, 2, \dots, p, \quad (3)$$

and the corresponding ARL_{0i} becomes:

$$ARL_{0,i} = 1 / (1 - \sqrt[p]{1 - \alpha}); \quad i = 1, 2, \dots, p. \quad (4)$$

In order to reach the ARL_{0i} values for each control chart, the bisection method was applied. The bisection method is based on the fact that a function will change sign when it passes through zero. By evaluating the

function at the middle of an interval and replacing whichever limit has the same sign, the bisection method can halve the size of the interval in each iteration and eventually find the root. For example, to find a root of $f(x) = 0$ in the interval of (a_0, b_0) with which $f(a_0)f(b_0) < 0$, one picks tolerance ε and, then, applies the following algorithm:

$$k = 0$$

$$\text{while } |f(x_{k+1})| > \varepsilon$$

$$x_{k+1} = \frac{a_k + b_k}{2}$$

$$\text{if } (f(x_{k+1})f(a_k) < 0)$$

then,

$$a_{k+1} = a_k \text{ and } b_{k+1} = x_{k+1}$$

else,

$$b_{k+1} = b_k \text{ and } a_{k+1} = x_k$$

end if

$$k = k + 1$$

end while

$$x^* = x_k.$$

In order to apply the bisection algorithm to find L as the symmetric limit for variable i , let one define $ARL_{0i}(L)$ to be the ARL_{0i} , when the control limits are obtained by $\pm L$ for variable i . In other words, $f(x = L)$ becomes $ARL_{0i}(L) - ARL_{0i}$. Then, L_1 and L_2 are selected, such that, if one uses $\pm L_1$ as control limits for variable i , $ARL_0(L_1) < ARL_{0,i}$ and if one employs $\pm L_2$, then, $ARL_0(L_2) > ARL_{0,i}$. One sets $L = (L_1 + L_2)/2$. In the next step of the algorithm, if $ARL_{0i}(L)$ becomes greater than ARL_{0i} , then, L_2 is replaced with L , otherwise, L_1 is replaced with L . One continues until $ARL_0(L)$ approaches ARL_{0i} . At the end, an L value for the symmetric limits on the i th variable is selected.

In summary, in the proposed method, first, transformation on the original quality characteristics is applied and, then, for each transformed variable, a Shewhart-type control chart satisfying the $\min_{L_i} |ARL_{0i}(L) - ARL_{0i}|$ relation is found.

PERFORMANCE EVALUATION

In this section, a simulation study containing three numerical examples is performed to evaluate the performance of the proposed method and to compare it with an existing method in different situations.

Example 2

Consider a manufacturing process, in which the product has two dependent quality characteristics measured as attributes. Based on historical data, the number of nonconforming items for the two quality characteristics has a means of $\lambda_1 = 5$ and $\lambda_2 = 6$ with a correlation of 0.18.

To monitor both attributes simultaneously, first, 5000 random vectors, following a bivariate Poisson distribution with the given parameters, were generated. Then, the vectors were transformed to the new vector using Equation 2. The mean and the covariance of the uncorrelated transformed vector were:

$$\widehat{\mu_Y} = [0.0054, 0.0013]^T,$$

and:

$$\widehat{\text{Cov}(\mathbf{Y})} = \begin{pmatrix} 0.200 & 0.000 \\ 0.000 & 0.166 \end{pmatrix}.$$

Moreover, the in-control and out-of-control average run length criteria were used to evaluate the proposed method, along with a comparison study using Patel's procedures. To do this, it is noted that the upper control limit of the T^2 chart in Patel's method is $\chi_{0.995,2}^2 = 10.59$. For the proposed method, first, $L_1 = 2.5$ and $L_2 = 3.5$ were selected for each variable and, then, the bisection method was applied to reach the L values of 3.2658 and 3.2620 for the first and the second variables, respectively. For an in-control ARL study, a replication of 10000 data sets resulted in an ARL_0 value of 205.756, ($\alpha = 0.0049$), for the proposed method. The corresponding value is 121.834, ($\alpha = 0.0082$), for Patel's method. It is seen that when the original data in Patel's method is used, the ARL_0 value is very low, whereas, when the data is transformed, the ARL_0 will have an appropriate value of 205.75. Moreover, since the type-I error of Patel's method is much higher than the ones from the proposed method, it is not possible to compare its out-of-control average run length with the ones from the proposed method. Hence, the ARL_1 values of the proposed method were computed for different shifts and the results were summarized in Table 1. The results of Table 1 show that the proposed method has a good performance.

Example 3

In this example, the mean numbers of nonconforming items on each of the quality characteristics are $\lambda_1 = 7$ and $\lambda_2 = 6$, with a higher correlation value of 0.59.

To monitor both attributes simultaneously, first, 5000 random vectors on the bivariate Poisson distribution, with the above parameters, were generated. Then, the original vector was transformed to the new vector using Equation 2. The estimated mean vector and the covariance matrix of the transformed vector were:

$$\widehat{\mu_Y} = [0.0101, 0.0089]^T,$$

and:

$$\widehat{\text{Cov}(\mathbf{Y})} = \begin{pmatrix} 1 & 0.0112 \\ 0.0112 & 1 \end{pmatrix}.$$

The upper control limit of the T^2 chart is $\chi_{0.995,3}^2 = 12.84$. In addition, both in-control and out-of-control average run lengths (ARL) criteria were used to evaluate the proposed method, along with a comparison study with Patel's procedures. First, $L_1 = 2$ and $L_2 = 3.5$ were selected for each variable and, at the end of the bisection method, L -values of 3.2275 and 3.2656 were reached for the first and the second variable, respectively. Then, a replication of 10000 observations was generated, which resulted in ARL_0 values of 198.59 and 142.101 for the proposed and Patel's method, respectively. It is again seen that, when one uses the original data in Patel's method, the ARL_0 value is very low, whereas, when one transforms the data, the ARL_0 will have a more appropriate value. Moreover, since type-I errors were different, only ARL_1 values of the proposed method were computed for different shifts and the results were summarized in Table 2. Once again, one can see the proper performance of the proposed method.

Example 4

This example contains three attributes, following a multivariate Poisson distribution with parameters $\lambda_1 = 4$, $\lambda_2 = 6$ and $\lambda_3 = 3$ with the correlation matrix of:

$$\Sigma = \begin{pmatrix} 1 & 0.18 & 0.38 \\ 0.18 & 1 & 0.49 \\ 0.38 & 0.49 & 1 \end{pmatrix}.$$

Table 1. ARL_1 values for different shifts in Example 2.

Mean Shift $\rightarrow$	(0,0)	($\sigma_1, 0$)	(0, σ_2)	(σ_1, σ_2)	($2\sigma_1, 0$)	(0, $2\sigma_2$)
ARL_1	205.75	26.1902	23.3513	15.4439	5.8754	5.3081
Mean Shift $\rightarrow$	($2\sigma_1, 2\sigma_2$)	($3\sigma_1, 0$)	(0, $3\sigma_2$)	($\sigma_1, 3\sigma_2$)	($3\sigma_1, \sigma_2$)	($3\sigma_1, 3\sigma_2$)
ARL_1	3.6315	2.4243	2.2665	2.4368	2.5068	1.6952

Table 2. ARL_1 values for different shifts in Example 3.

Mean Shift $\rightarrow$	(0, 0)	(σ_1 , 0)	(0, σ_2)	(σ_1 , σ_2)	($2\sigma_1$, 0)	(0, $2\sigma_2$)
ARL_1	198.59	18.704	18.541	18.5426	3.7752	4.0156
Mean Shift $\rightarrow$	($2\sigma_1$, $2\sigma_2$)	($3\sigma_1$, 0)	(0, $3\sigma_2$)	(σ_1 , $3\sigma_2$)	($3\sigma_1$, σ_2)	($3\sigma_1$, $3\sigma_2$)
ARL_1	4.702	1.7513	1.8161	2.1645	2.1507	2.1268

Table 3. ARL_1 values for different shifts in Example 4.

Mean Shift $\rightarrow$	(σ_1 , 0, 0)	(0, σ_2 , 0)	(0, 0, σ_3)	(σ_1 , σ_2 , 0)	(σ_1 , 0, σ_3)	(0, σ_2 , σ_3)	(σ_1 , σ_2 , σ_3)
ARL_1	26.8564	23.6643	18.6054	14.2781	15.3644	17.4370	14.6709
Mean Shift $\rightarrow$	$2(\sigma_1, 0, 0)$	$2(0, \sigma_2, 0)$	$2(0, 0, \sigma_3)$	$2(\sigma_1, \sigma_2, 0)$	$2(\sigma_1, 0, \sigma_3)$	$2(0, \sigma_2, \sigma_3)$	$2(\sigma_1, \sigma_2, \sigma_3)$
ARL_1	6.0101	5.0849	4.4204	3.2923	4.0277	4.6011	3.9850
Mean Shift $\rightarrow$	$3(\sigma_1, 0, 0)$	$3(0, \sigma_2, 0)$	$3(0, 0, \sigma_3)$	$3(\sigma_1, \sigma_2, 0)$	$3(\sigma_1, 0, \sigma_3)$	$3(0, \sigma_2, \sigma_3)$	$3(\sigma_1, \sigma_2, \sigma_3)$
ARL_1	2.5124	2.1866	2.0628	1.5692	1.9118	2.1124	1.8873

To monitor all attributes simultaneously, first, 5000 random vectors, on the above multivariate Poisson distribution, were generated. Then, the vectors were transformed to the new vectors by Equation 4. The mean vector and the covariance matrix of the transformed variables became:

$$\widehat{\mu_Y} = [0.0253, 0.0054, 0.0027]^T,$$

and:

$$\widehat{\text{Cov}(\mathbf{Y})} = \begin{pmatrix} 1 & 0.0189 & 0.0074 \\ 0.0189 & 1 & 0.0390 \\ 0.0074 & 0.0390 & 1 \end{pmatrix}.$$

The upper control limit of the T^2 chart is $\chi_{0.995,3}^2 = 12.84$. In the in-control and out-of-control average run length study, $L_1 = 2$ and $L_2 = 4$ were selected and, at the end of the bisection method, the L -values of 3.4968, 3.4688 and 3.4687 were reached for the first, the second and the third variables, respectively. In the in-control ARL study, a replication of 10000 observations resulted in ARL_0 values of 192.0936 and 120.680, for the proposed and Patel's method, respectively. The results indicate, once again, that the transformation technique is very useful. Furthermore, the ARL_1 values of the proposed method were estimated for different shifts and the results were summarized in Table 3.

The results of Table 3 show that the proposed method performs well, even in situations where there are both positive and negative shifts around the mean.

CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE RESEARCH

In this research, based on the symmetric square root transformation concept, first, data obtained from

multi-attribute quality control systems, was transformed, such that the correlation between variables either vanished or became very small. Then, using simulation and bisection methods, the symmetric rectangular control region was found to monitor all attributes simultaneously. In simulation studies, some numerical examples were presented to illustrate the proposed method and to evaluate and compare its performances with the existing method in different scenarios. The results show that the proposed method performs better than the existing method, in terms of in-control ARL criterion, in all cases.

The proposed method may be applied to multi-attribute binomial (MNP) control charts in future research.

REFERENCES

- Hotelling, H. "Multivariate quality control", In *Techniques of Statistical Analysis*, Eisenhart, C., Hastay, M.W. and Wallis, W.A., Eds., McGraw-Hill, New York, NY, USA (1947).
- Lowry, C.A. and Montgomery, D.C. "A review of multivariate control charts", *IIE Transactions*, **27**, pp 800-810 (1995).
- Golnabi, S. and Houshmand, A.A. "Multivariate Shewhart x-bar chart", *Inter. Stat.*, **4** (1999); A web based journal: <http://interstat.stat.vt.edu/interstat/index/Sep99.html>.
- Woodall, W.H. and Ncube, M.M. "Multivariate CUSUM quality-control procedures", *Technometrics*, **27**, pp 285-292 (1985).
- Healy, J.D. "A note on multivariate CUSUM procedures", *Technometrics*, **29**, pp 409-412 (1987).
- Lucas, J.M. and Crosier, R.B. "Fast initial response for CUSUM quality control schemes: Give your CUSUM a head start", *Technometrics*, **42**(1), pp 102-107 (2000).

7. Pignatiello, J.J. and Runger, G.C. "Comparisons of multivariate CUSUM charts", *Journal of Quality Technology*, **22**, pp 173-186 (1990).
8. Lowry, C.A., Woodall, W.H., Champ, C.W. and Erigdon, S. "A multivariate exponentially weighted moving average control chart", *Technometrics*, **34**, pp 46-53 (1992).
9. Runger, G.C. "Projections and the U2 multivariate control chart", *Journal of Quality Technology*, **2**, pp 313-319 (1996).
10. Hawkins, D.M. "Regression adjustment for variables in multivariate quality control", *Journal of Quality Technology*, **25**, pp 175-182 (1991).
11. Niaki, S.T.A. and Abbasi, B. "Fault diagnosis in multivariate control charts using artificial neural networks", *Journal of Quality & Reliability Engineering International*, **21**, pp 825-840 (2005).
12. Montgomery, D.C., *Introduction to Statistical Quality Control*, 5th Ed., John Wiley & Sons, New York, NY, USA (2005).
13. Patel, H.I. "Quality control methods for multivariate binomial and Poisson distributions", *Technometrics*, **15**, pp 103-112 (1973).
14. Lu, X.S., Xie, M., Goh, T.N. and Lai, C.D. "Control chart for multivariate attribute processes", *International Journal of Production Research*, **36**, pp 3477-3489 (1998).
15. Jolayemi, J.K. "An optimal design of multi-attribute control charts for processes subject to a multiplicity of assignable causes", *Applied Mathematics and Computation*, **114**, pp 187-203 (2000).
16. Jolayemi, J.K. "Convolution of independent binomial variables: An approximation method and a comparative study", *Computational Statistics & Data Analysis*, **18**, pp 403-417 (1994).
17. Gibra, I.N. "Economically optimal determination of the parameters of np-control charts", *Journal of Quality Technology*, **10**, pp 12-19 (1978).
18. Marcucci, M. "Monitoring multinomial processes", *Journal of Quality Technology*, **17**, pp 86-91 (1985).
19. Larpiattaworn, S. "A neural network approach for multi-attribute process control with comparison of two current techniques and guidelines for practical use", Ph.D. Thesis, University of Pittsburgh, USA (2003).
20. Gadre, M.P. and Rattihalli, R.N. "Some group inspection based multi-attribute control charts to identify process deterioration", *Economic Quality Control*, **2**, pp 151-164 (2005).
21. Cario, M.C. and Nelson, B.L. "Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix", Technical Report, Department of Industrial Engineering and Management Sciences, Northwestern University, USA (1997).
22. Niaki, S.T.A. and Abbasi, B. "NORTA and neural networks based method to generate random vectors with arbitrary marginal distributions and correlation matrix", In the *Proceedings of the 17th IASTED International Conference on Modeling and Simulation*, Montreal, Canada, pp 234-239 (2006).