The International Journal's
Research Journal of Science & IT Management
ISSN:2251-1563

Multivariate Statistical Quality Control –Attribute Charts C Techniques in Designing

Mrs. G.Annalakshmi
Asst.prof of Mathematics and Research scholar , Dr.M.G.R Educational Research Institute University ,Madurayoal , Chennai-95

Dr. S.P Rajagopalan
Prof . of Emeritus , Dr.M.G.R Educational Research Institute University , Maduravoyal Chennai-95

Dr. A. Iyemperumal
Prof. of Mathematics Dr.M.G.R Educational Research Institute University ,Maduravoyal Chennai-95

Abstract:

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, we develop a new methodology to monitor multi-attribute processes in which the defects counts are important and different types of defects are dependent random variables. In order to do this, based on the symmetric square root transformation concept, first we transform multi-attribute data such that the correlation between variables either vanishes or becomes very small. Then, by simulation and bisection method, we find the symmetric control limits and form a symmetric rectangular region for control. In simulation studies, we present some numerical examples to illustrate the proposed method and to evaluate and compare its performance to the ones of the existing method.

Key words:
Multivariate C control charts, process monitoring, optimal control limits, symmetric control charts, symmetric square root

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, we 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 rates, all of which need to be monitored in a situation where some correlation may exist between them. In these cases, if we want 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 chart, for which many researchers have developed different methodologies. Early research on multivariate control charts goes back to Hotelling (1947), 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 (1995) have shown that a multivariate control scheme normally has better sensitivity than the one based on univariate control charts. Other multivariate control charts are multivariate form of the Shewhart charts presented by Golnabi and Houshmand (1999), the multivariate CUSUM charts proposed by Woodall and Ncube (1985), Healy (1987), Lucas and Crosier (1982), and Pignatiello and Runger (1990), and the multivariate exponentially weighted moving average (MEWMA) charts proposed by Lowry et al. (1992). Moreover, there are some other methods proposed by Runger (1996), Hawkins (1991), and Niaki and Abbasi (2005).

Despite the fact that the multi-attribute monitoring has many applications, almost all researchers have focused on the first category of control charting and only few methods have been proposed to monitor multi-attribute processes (see for example Xie and Goh (1992) and Montgomery (2005)). Furthermore, in many instances where we may not need exact measurements, it is easy to collect correlated discrete-type data. Patel (1973) 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 case, he assumes that if we choose 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. (1998) addressed the statistical design of multi-attribute control charts. They proposed a multivariate *np*-chart (MNP chart) to develop Shewhart charts based on an X statistic. They showed that this statistic reduced type two 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 (2000) 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 it is based on the assumption of independent attributes, a J approximation (Jolayemi (1994)), and Gibra's model (1978) for univariate *np* chart. When the proportions in each quality category are known or estimated using a base period, Marcucci (1985) 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. Larpkiattaworn (2003) proposed a back propagation neural network (BPNN) for two-attribute process control in bivariate Binomial and bivariate Poisson case. They 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. They also discussed different values of correlation between two variables and gave some suggestions on using three-attribute control charts (χ^2, MNP, and BPNN method). Gadre and Rattihalli (2005) with the assumption of multinomial distribution for multi attribute processes, used MP-test to determine if the parameters of the distribution have changed 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-attributes process mean in a multivariate C chart is proposed. This region is reached by first employing a transformation method that almost eliminates the correlation between the 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.

χ^2 and MNP charts as existing control charts for multi-attribute processes. Explains the concept of the transformation technique used in this research. The new method is developed. In order to understand the proposed method better,we present three numerical examples and evaluate its performance using the Average Run Length (*ARL*) criterion. The conclusion and recommendations for future research

Normal approximation method used in multi-attribute monitoring

In this section, we briefly explain the normal approximation of either the multi-variate Binomial or multi-variate Poisson distribution used in multi-attribute control charting (Patel (1973)).

Although Patel's method included both time independent and time dependent (auto-correlated) samples, we focus 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.

$$T^2 = (\mathbf{X} - \bar{\mathbf{X}})'\mathbf{S}^{-1}(\mathbf{X} - \bar{\mathbf{X}}) \qquad \text{……… ………..(1)}$$

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, $\mathbf{\Sigma}$, 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, $\chi_\alpha{}^2$, and α is a specified significance level. The lower control limit is equal to zero. In this method $\mathbf{X}$ may follow a multi-variate binomial or multi-variate Poisson distribution.

Symmetric square root method :

In order to eliminate the existing correlation between the quality characteristics in vector $\mathbf{X} = \left[X_1, X_2, ..., X_p\right]^T$, in a symmetric square root transformation method a new vector $\mathbf{Y} = \left[Y_1, Y_2, ..., Y_p\right]^T = \mathbf{CX}$ 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 $\mathbf{\Sigma}$, the correlation matrix of $\mathbf{X}$. However, before applying this method, first, we subtract $\boldsymbol{\mu} = \left[\mu_1, \mu_2, ..., \mu_p\right]^T$ from $\mathbf{X}$ and then do the transformation (Golnabi and Houshmand (1999)). In other words, the transformed vector becomes

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

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

Suppose vector $\mathbf{X} = \left[X_1, X_2\right]^T$ follows a bivariate Poisson distribution, in which the marginal probability mass distributions are Poisson with parameters $\lambda_1 = 4, \lambda_2 = 5$ with covariance equal to 1.75. In this case, using NORTA method (Cario and Nelson (1997) & Niaki and Abbasi (2006)), 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 $\boldsymbol{\mu}_{\mathbf{X}} = [\ 4.041, 5.011]^T$, $\text{Cov}(\mathbf{X}) = \begin{pmatrix} 4.111 & 1.752 \\ 1.752 & 4.883 \end{pmatrix}$.

Then, applying equation (2) to the generated observations yield $\boldsymbol{\mu}_{\mathbf{Y}} = [0.0103, 0.0003]$, $\text{Cov}(\mathbf{Y}) = \begin{pmatrix} 0.2524 & \text{-}0.0002 \\ \text{-}0.0002 & 0.1981 \end{pmatrix}$ on the transformed vector, in which we note that the covariance is close to zero.

Symmetric multi-attribute control charts for transformed data:

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

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, we need to determine an ARL_{0i} value for each transformed variable such that the overall average run length becomes 200. If we define α_i to be type-one error associated with the i^{th} control interval on the i^{th} transformed quality characteristic, then we will have:

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

And the corresponding ARL_{0_i} becomes:

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

In order to reach the ARL_{0_i} values for each control charts 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 $f(a_0)f(b_0) < 0$ we pick tolerance ε and then apply the following algorithm:

$k = 0$

While $\left|f(x_{k+1})\right| > \varepsilon$

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

If $\left(f(x_{k+1})f(a_k) < 0\right)$

Then

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

Else

$b_{k+1} = b_k$ 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 us 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, we select L_1 and L_2 such that if we use $\pm L_1$ as control limits for variable i, $ARL_0(L_1) < ARL_{0i}$, and if we employ $\pm L_2$ then $ARL_0(L_2) > ARL_{0i}$. We set $L = (L_1 + L_2)/2$. In the next step of the algorithm, if $ARL_{0i}(L\ )$ becomes greater than ARL_{0i} then L_2 is replaced with L, otherwise we replace L_1 with L. We continue until $ARL_0(L\ )$ approaches to ARL_{0i}. At the end, a L value for the symmetric limits on the i^{th} variable is selected.

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

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 means of $\lambda_1 = 5, \lambda_2 = 6$ with correlation of 0.18.

To monitor both attributes simultaneously, first 5000 random vector following a bi-variate 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

$\boldsymbol{\mu}_{\mathbf{Y}} = \left[0.0054, 0.0013\right]^T$ and $\mathrm{Cov}(\mathbf{Y}) = \begin{pmatrix} 0.200 & 0.000 \\ 0.000 & 0.166 \end{pmatrix}$.

Moreover, we used the in-control and out-of-control average run lengths criteria and performed an evaluation on the proposed method along with a comparison study with Patel's procedures. To do this, we note that the upper control limit of T^2 chart in Patel's method is $\chi^2_{0.995,2} = 10.59$. For the proposed method, first we selected L_1=2.5 and L_2=3.5 for each variables and then applied the bisection method 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. We see that when we use the original data in Patel's method, the ARL_0 value is very low, whereas when we transform the data, the ARL_0 will have an appropriate value as 205.75. Moreover, since the type one error of the Patel's method is much higher than the one from the proposed method, we are not able to compare its out-of-control average run length with the ones from the proposed method. Hence, we computed ARL_1 values of the proposed method for different shifts and summarized the results in Table (1). The results of Table (1) show that the proposed method has a good performance.

In this example, the number of nonconforming items on each of the quality characteristics has means of $\lambda_1 = 7, \lambda_2 = 6$ with a higher correlation value of 0.59.

Table 1: ARL_1 values for different shifts

Mean Shift→	$(0,0)$	$(\sigma_1,0)$	$(0,\sigma_2)$	(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→	$(2\sigma_1,2\sigma_2)$	$(3\sigma_1,0)$	$(0,3\sigma_2)$	$(\sigma_1,3\sigma_2)$	$(3\sigma_1,\sigma_2)$	$(3\sigma_1,3\sigma_2)$
ARL_1	3.6315	2.4243	2.2665	2.4368	2.5068	1.6952

To monitor both attributes simultaneously, first 5000 random vector on bi-variate 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

$$\boldsymbol{\mu}_{\mathbf{Y}} = [-0.0101, -0.0089]^T \text{ and } \mathrm{Cov}(\mathbf{Y}) = \begin{pmatrix} 1 & -0.0112 \\ -0.0112 & 1 \end{pmatrix}.$$

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

Table 2: ARL_1 values for different shifts

Mean Shift→	$(0,0)$	$(\sigma_1,0)$	$(0,\sigma_2)$	(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→	$(2\sigma_1,2\sigma_2)$	$(3\sigma_1,0)$	$(0,3\sigma_2)$	$(\sigma_1,3\sigma_2)$	$(3\sigma_1,\sigma_2)$	$(3\sigma_1,3\sigma_2)$

***ARL*₁**	4.702	1.7513	1.8161	2.1645	2.1507	2.1268

This example contains three attributes following a multivariate Poisson distribution with parameters $\lambda_1 = 4, \lambda_2 = 6, \lambda_3 = 3$ with the correlation matrix of $\mathbf{\Sigma} = \begin{pmatrix} 1 & 0.18 & 0.38 \\ 0.18 & 1 & 0.49 \\ 0.38 & 0.49 & 1 \end{pmatrix}$.

To monitor all attributes simultaneously, first 5000 random vector 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

$$\boldsymbol{\mu}_{\mathbf{Y}} = [0.0253\ ,\text{-}0.0054, 0.0027]^T \text{ and } Cov(\mathbf{Y}) = \begin{pmatrix} 1 & 0.0189 & \text{-}0.0074 \\ 0.0189 & 1 & \text{-}0.0390 \\ \text{-}0.0074 & \text{-}0.0390 & 1 \end{pmatrix}.$$

The upper control limit of T^2 chart is $\chi^2_{0.995,2} = 10.59$. In in-control and out-of-control average run length study, we selected L_1=2 and L_2=4 and at the end of the bisection method we reached the *L*-values of 3.4968, 3.4688, and 3.4687 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 the Patel's method, respectively. The results indicate one more time that the transformation technique is very useful. Furthermore, we estimated ARL_1 values of the proposed method for different shifts and summarized the results 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.

Table 3: ARL_1 values for different shifts

Mean Shift→	$(\sigma_1,0,0)$	$(0,\sigma_2,0)$	$(0,0,\sigma_3)$	$(\sigma_1,\sigma_2,0)$	$(\sigma_1,0,\sigma_3)$	$(0,\sigma_2,\sigma_3)$	$(\sigma_1,\sigma_2,\sigma_3)$
***ARL*₁**	26.8564	23.6643	18.6054	14.2781	15.3644	17.4370	14.6709
Mean Shift→	$2(\sigma_1,0,0)$	$2(0,\sigma_2,0)$	$2(0,0,\sigma_3)$	$2(\sigma_1,\sigma_2,0)$	$2(\sigma_1,0,\sigma_3)$	$2(0,\sigma_2,\sigma_3)$	$2(\sigma_1,\sigma_2,\sigma_3)$
***ARL*₁**	6.0101	5.0849	4.4204	3.2923	4.0277	4.6011	3.9850
Mean Shift→	$3(\sigma_1,0,0)$	$3(0,\sigma_2,0)$	$3(0,0,\sigma_3)$	$3(\sigma_1,\sigma_2,0)$	$3(\sigma_1,0,\sigma_3)$	$3(0,\sigma_2,\sigma_3)$	$3(\sigma_1,\sigma_2,\sigma_3)$
***ARL*₁**	2.5124	2.1866	2.0628	1.5692	1.9118	2.1124	1.8873

Conclusions and recommendation for future research

In this research, based on the symmetric square root transformation concept, first we transformed data obtained from multi-attributes quality control systems such that the correlation between variables either vanished or became very small. Then, by simulation and bisection method, we found the symmetric rectangular control region to monitor all attributes simultaneously. In simulation studies, we presented some numerical examples 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 term of in-control *ARL* criterion in all cases.

We may apply the proposed method to multi-attribute binomial (MNP) control charts in future research.

References

1. Cario, M. C. and Nelson, B. L. (1997). “Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix”. Technical Report. Department of Industrial Engineering and Management Sciences, Northwestern University, U.S.A.
2. Gadre, M. P. and Rattihalli, R. N. (2005). “Some group inspection based multi-attribute control charts to identify process deterioration”. *Economic Quality Control*, **2**, 151-164.
3. Gibra, I. N. (1978). “Economically optimal determination of the parameters of np-control charts”. *Journal of Quality Technology,* **10**, 12 –19.
4. Golnabi, S. and Houshmand, A. A. (1999). “Multivariate shewhart x-bar chart.” *Inter Stat*, No. 4, —A web based journal: http://interstat.stat.vt.edu/interstat/index/Sep99.html
5. Hawkins, D. M. (1991). “Regression adjustment for variables in multivariate quality control”. *Journal of Quality Technology,* **25***,* 175-182.
6. Healy J. D. (1987). “A note on multivariate CUSUM procedures”. *Technometrics,* **29***,* 409–412.
7. Hotelling, H. (1947). “Multivariate quality control". *In Techniques of Statistical Analysis*, edited by Eisenhart, Hastay, and Wallis, McGraw-Hill, New York, NY.
8. Jolayemi, J. K. (1994). “Convolution of independent binomial variables: An approximation method and a comparative study”. *Computational Statistics & Data Analysis,* **18***,* 403–417.
9. Jolayemi, J. K. (2000). “An optimal design of multi-attribute control charts for processes subject to a multiplicity of assignable causes”. *Applied Mathematics and Computation,* **114**, 187-203.
10. Larpkiattaworn, S. (2003). “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, U.S.A.
11. Lowry, C. A. and Montgomery, D. C. (1995). “A review of multivariate control charts”. *IIE Transactions,* **27***,* 800-810.
12. Lowry, C. A., Woodall, W. H., Champ, C. W., and Erigdon, S. (1992). “A multivariate exponentially weighted moving average control chart”. *Technometrics,* **34***,* 46-53.
13. Lu, X. S., Xie, M., Goh, T. N., and Lai, C. D. (1998). “Control chart for multivariate attribute processes”. *International Journal of Production Research,* **36***,* 3477-3489.
14. Lucas, J. M. and Crosier, R. B. (1982). “Fast initial response for CUSUM quality control schemes: Give your CUSUM a head start”. *Technometrics,* **24***,* 199-2054.
15. Marcucci, M. (1985). “Monitoring multinomial processes”. *Journal of Quality Technology,* **17***,* 86-91.
16. Montgomery, D. C. (2005). “*Introduction to Statistical Quality Control*”. 5th ed., John Wiley & Sons, New York, NY.
17. Niaki, S. T. A. & Abbasi, B. (2005). “Fault diagnosis in multivariate control charts using artificial neural networks”. *Journal of Quality & Reliability Engineering International,* **21**, 825-840.
18. Niaki, S. T. A. & Abbasi, B. (2006). “NORTA and neural networks based method to generate random vectors with arbitrary marginal distributions and correlation matrix”. *In the proceedings of the17th IASTED International Conference on Modeling and Simulation*, Montreal, Canada, 234-239.

19. Patel, H. I. (1973). "Quality control methods for multivariate binomial and Poisson distributions". *Technometrics,* **15***,* 103-112.
20. Pignatiello J. J. and Runger, G. C. (1990). "Comparisons of multivariate CUSUM charts". *Journal of Quality Technology,* **22***,*173–186.
21. Runger, G. C. (1996). "Projections and the U2 multivariate control chart". *Journal of Quality Technology,* **2***,*313- 319.
22. Woodall W. H. and Ncube, M. M. (1985). "Multivariate CUSUM quality-control procedures". *Technometrics,* **27***,* 285–292.
23. Xie, M., Goh, T. N., and Tang, Y. (2000). "Data transformation for geometrically distributed quality characteristics". *Quality and Reliability Engineering International,* **16***,* 9-15.

***