Skewness Reduction Approach in Multi-Attribute Process Monitoring

Communications in Statistics—Theory and Methods, 36: 2313–2325, 2007 **Quality Control**

SEYED TAGHI AKHAVAN NIAKI AND BABAK ABBASI

報 告 者：鍾昇倍

指導老師：童超塵 教授

Outline

1. Introduction
2. Existing Multi-Attribute Control Charts
3. The Proposed Normalizing Transformation
4. Multi-Attribute Control Chart Based on Transformed Data in T^2 chart
5. Simulation Experiments
6. Conclusion and Recommendations for Future Research
7. References

何謂偏態 (Skewness)＆峰態(Kurtosis)

一、**偏態量數**主要是衡量次數分配之偏斜情況，亦即衡量高峰偏向變量大或變量小之一方。

峰態量數是用來衡量次數分配圖形中高峰兩旁之次數是高峻或平坦之現象。

二、偏態係數

(一)動差法

$$\beta_1 = \frac{m_3}{s^3} = \frac{\Sigma(X_i-\bar{X})^3}{n}$$

若(1) m_3 為0，$\beta_1 = 0$ 為對稱分配。

(2) m_3 為正，$\beta_1 > 0$ 為右偏分配（正偏分配）

(3) m_3 為負，$\beta_1 < 0$ 為左偏分配（負偏分配）

三、峰態係數

$$\beta_2 = \frac{m_4}{s^4} = \frac{m_4}{(m_2)^2} = \frac{\Sigma(X_i-\bar{X})^4}{n}$$

若(1) $\beta_2 > 3$ 為高狹峰分配（Leptokurtic distribution）

(2) $\beta_2 = 3$ 為常態峰分配（Mesokurtic distribution）

(3) $\beta_2 < 3$ 為低闊峰分配（PlatyKurtic distribution）

1. Introduction

- In general, there are two broad categories in statistical control charts, namely **variable** and **attribute** control chart.

(1)Variable：指可量度出讀數之計量品質性質。

(2)**Attribute**：指品質性質有多少件數是符合規定之計數品質性質。

1. Introduction (Conti.)

學者	年代	相關文獻
Hotelling	1947	Multivariate control charts of the Shewhart type were first developed by him.
Lowry and Montgomery	1995	They have shown that a multivariate control scheme normally has better sensitivity than the one based on univariate control charts.
Woodall and Ncube	1985	Other multivariate control charts are the multivariate CUSUM charts proposed.
Healy	1987	
Lucas and Crosier	1982	
Pignatiello and Runger	1990	
Lucas and Saccucci	1990	the multivariate exponentially weighted moving average (MEWMA) charts proposed.
Lowry et al.	1992	

1. Introduction (Conti.)

學者	年代	相關文獻
Bourke,P. D.	1991	Detecting shift in fraction non conforming using run-length control chart with 100% inspection.
Montgomery	2003	Introduction to Statistical Quality Control.
Xie et al.	2000	Data transformation for geometrically distributed quality characteristics.
Patel, H. I.	1973	He proposed a Hotelling-type χ^2 chart to monitor observations from multivariate binomial or multivariate Poisson distribution.
Larpkiattaworn	2003	proposed a back propagation neural network (BPNN) for two-attribute process control in bivariate binomial and bivariate Poisson case.

multi-attribute processes

1. Introduction (Conti.)

- In this article, we propose a T^2 control chart based upon the Patel (1973)method to monitor multi-attribute processes. At first, we propose a data transformation technique and then we use a T^2 control chart.(提出一個資料轉換的技巧，之後使用T^2管制圖- Patel (1973)X^2管制圖主要用在多元變數製程的監控)

2. Existing Multi-Attribute Control Charts

2.1 *Normal Approximation of Multivariate Binomial Distribution*

- Although Patel's method included both time-independent and time-dependent samples, we focus on the time independent case.(本文的樣本主要是時間獨立的案例)

$$T^2 = (X - \overline{X})' S^{-1} (X - \overline{X}), \qquad (1)$$

X ：a multivariate binomial random vector

$\overline{\mathbf{X}}$ ：the sample mean vector,

S ：an estimator of the population covariance matrix

where T^2 has an approximate chi-square distribution with p degrees of freedom, The upper control limit of the control chart equals the upper quantile of a chi-squared distribution with p degrees of freedom. The lower control limit is equal to zero.(上管制界限由卡方分配在自由度(P)下求得，下管制界限等於0)

2. Existing Multi-Attribute Control Charts

2.2 *Multivariate NP-Chart (MNP chart)*

$$UCL\ \&\ LCL = n\sum_{i=1}^{m} d_i\sqrt{p_i} \pm 3\sqrt{n\left\{\sum_{i=1}^{m} d_i^2(1-p_i) + 2\sum_{i<j}\left(d_i d_j \sigma_{ij}\sqrt{(1-p_i)(1-p_j)}\right)\right\}} \quad (2)$$

$$CL = n\sum_{i=1}^{m} d_i\sqrt{p_i}.$$

C_i：the number of non conforming items of type i(不合格項目的數量)

m：the number of process attributes

P_i：the proportion non conforming of the ith quality characteristic(品質特性不合格比例)

d_i：the demerits of the severity of the non conformance in the ith quality characteristic(品質特性不合格的缺點數)

p_i ＆σ_{ij}：are unknown and being estimated from historical data (未知且由歷史資料估預而得)

3. The Proposed Normalizing Transformation

- usually data have binomial, Poisson, or geometric distribution and the assumption of approximate normality causes two problems.

(要將資料型態binomial, Poisson, geometric 等分配假設為常態，存在二個問題)

(1)The first and most important problem is the fact that these distributions have skewness.

(最重要的問題是這些分配有偏態的問題)

(2) The second problem arises from the discrete nature of these distributions.

(這些分配均是離散型的分配，而常態分配是連續型分配)

3. The Proposed Normalizing Transformation

- There are two approaches to diminish skewness in univariate attribute control charts:

(有兩個減小偏態的方法，在多元計數管制圖的應用)

(1) adding some correction values to the control limits based on the value of skewness

(基於偏態的數值對管制界限增加一些修正值)

(2) applying normalizing transformation.

(應用常態轉換的方法)

※ Most researchers prefer to use normalizing transformation.

3. The Proposed Normalizing Transformation

學者	年代	轉換方法
Box and Cox	1964	square root
Johnson and Kotz	1969	inverse 反置法
Ryan	1989	arcsin 反正弦法
Ryan and Schwertman	1997	parabolic inverse 拋物線反置法
Quesenberry	1995	Q-transformation
Xie et al.	2000	**double square root** transformation for Geometric distribution
Niaki and Abbasi	2007	rth root transformation

3. The Proposed Normalizing Transformation

- In the skewness-reduction method, if we define f(r) to be the amount of skewness on the r_{th} root transformed attribute x, (i.e., x^r), we want to find r such that f(r) becomes zero. Therefore, applying the bisection method, we try to find a root for f(r) = 0 in the interval (0,1).(應用二分法 ,去找出第r個根,在f(r) 為0，介於0~1的區間裡)
- In order to explain the proposed rth root transformation method, we present two numerical examples by NORTA algorithm (Cario and Nelson, 1997) and compare the results with the ones of the other transformation.(本文提出二個由演算法的數值分析例子，為了解釋第r個根的轉換方法)

3. The Proposed Normalizing Transformation

- For a true normal distribution, the sample skewness should be near zero and the sample kurtosis should be near three.

Table 1
A comparison of different transformation methods

Index \ Method		The proposed transformation	$\sqrt{X}$	arcsin(X)	Q-Transformation**
Example 1					
X_1	Skewness	0.0019	−0.7429	1.1099	0.0384
	Kurtosis	2.9596	4.6772	2.5539	2.8992
	P-value (JB test)	0.8337	0	0	0.1831
X_2	Skewness	0.0093	−0.6257	1.3510	0.0134
	Kurtosis	3.0032	4.3563	4.0303	2.9456
	P-value (JB test)	0.9642	0	0	0.6724
Example 2					
X_1	Skewness	0.0015	−0.2863	0.7987	0.0486
	Kurtosis	3.0213	3.2549	4.5368	3.0815
	P-value (JB test)	0.9579	0	0	0.1915
X_2	Skewness	−0.0042	−0.5293	1.4977	0.0325
	Kurtosis	2.9882	4.2086	5.5721	2.9419
	P-value (JB test)	0.9753	0	0	0.4465

**In *Q-Transformation* method each transformed variable (X_{new}) is $\Phi^{-1}(F(x))$, where $F(x)$ is cumulative probability distribution function (cdf) of X and Φ is the standard normal cdf.

4. Multi-Attribute Control Chart Based on Transformed Data in T^2 chart

- T^2 control charts, due to its excellent performance in multivariable quality control environments, may perform well in multi-attribute processes in which we transform their attributes to have multi-normal distribution.

(將計數值轉換成常態分配的， T^2 管制圖應也會有較佳的表現)

- To avoid the problems in Patel's method (1973),

(1) we eliminate the skewness of marginal distributions.

(排除邊際分配的偏態)

(2) estimate the covariance matrix of the transformed variables.

(估計轉換變數的共變異矩陣)

(3) the control limits in multivariate control process.

(在多元變數管制製程建置管制界限)

5. Simulation Experiments

5.1.1 *Simulation Experiment 1*

- simulation experiment with three attributes. based on available historical data, sample size of 30, probability vector of $\hat{\mathbf{p}}$ = 〔$p_1 = 0.1$ $p_2 = 0.15$ $p_3 = 0.18$〕

$$\hat{\Sigma} = \begin{pmatrix} 2.6 & 0.6 & 0.48 \\ 0.6 & 3.8 & 1 \\ 0.48 & 1 & 4.5 \end{pmatrix}.$$

To monitor all attributes simultaneously, first we generate 5,000 observations on MBiniomial〔〔$n_1 = 30$ $n_2 = 30$ $n_3 = 30$〕,$\hat{\mathbf{p}}$, $\hat{\Sigma}$〕 random vector. Then we find appropriate transformations such that the marginal distributions are approximately normal. Based on the proposed transformation method we have r1 = 0.76, r2 = 0.76,r3 = 0.75,

$$\hat{\boldsymbol{\mu}}_{\text{new}} = [2.26, 3.37, 3.48] \quad \text{and} \quad \hat{\Sigma}_{\text{new}} = \begin{pmatrix} 0.94 & 0.19 & 0.15 \\ 0.19 & 1.11 & 0.26 \\ 0.15 & 0.26 & 1.16 \end{pmatrix}$$

and the upper control limit of T^2 chart is $\chi^2_{0.995,3} = 1283$.

5. Simulation Experiments

5.1.2 *Simulation Experiment 1*

Table 2

ARL_1 values for different shifts in simulation experiment 1

Shift→	$(\sigma_1, 0, 0)$	$(0, \sigma_2, 0)$	$(0, 0, \sigma_3)$	$(\sigma_1, \sigma_2, 0)$	$(\sigma_1, 0, \sigma_3)$	$(0, \sigma_2, \sigma_3)$	$(\sigma_1, \sigma_2, \sigma_3)$
Proposed method	24.990	33.265	19.110	15.909	10.747	13.310	8.784
MNP method	69.926	80.059	84.742	22.5130	24.822	27.104	9.873
Shift→	$(2\sigma_1, 0, 0)$	$(0, 2\sigma_2, 0)$	$(0, 0, 2\sigma_3)$	$(2\sigma_1, 2\sigma_2, 0)$	$(2\sigma_1, 0, 2\sigma_3)$	$(0, 2\sigma_2, 2\sigma_3)$	$(2\sigma_1, 2\sigma_2, 2\sigma_3)$
Proposed method	5.476	5.903	3.745	2.781	2.112	2.320	1.929
MNP method	22.132	24.874	29.194	4.948	5.049	5.780	2.184
Shift→	$(3\sigma_1, 0, 0)$	$(0, 3\sigma_2, 0)$	$(0, 0, 3\sigma_3)$	$(3\sigma_1, 3\sigma_2, 0)$	$(3\sigma_1, 0, 3\sigma_3)$	$(0, 3\sigma_2, 3\sigma_3)$	$(3\sigma_1, 3\sigma_2, 3\sigma_3)$
Proposed method	2.283	2.142	1.673	1.296	1.166	1.241	1.106
MNP method	9.369	10.891	13.026	2.007	1.994	2.185	1.204
Shift→	$(-\sigma_1, \sigma_2, 0)$	$(-\sigma_1, 0, \sigma_3)$	$(0, -\sigma_2, \sigma_3)$	$(2\sigma_1, -2\sigma_2, 0)$	$(2\sigma_1, 0, -2\sigma_3)$	$(0, 2\sigma_2, -2\sigma_3)$	$(\sigma_1, \sigma_2, -2\sigma_3)$
Proposed method	13.247	9.817	7.612	1.002	1.379	1.215	1.505
MNP method	463.193	547.985	395.338	449.893	306.384	430.089	320.916

※The results of Table 2 show that the proposed method on the transformed data performs better than MNP procedure in all scenarios. This fact is more obvious in situations where there are both positive and negative shifts around the mean.

5. Simulation Experiments

5.2 *Simulation Experiment 2*

- simulation experiment with three correlated attributes. sample size of 22 generated data sets, probability vector of $\hat{p}$ = （$p_1 = 0.11$ $p_2 = 0.12$ $p_3 = 0.16$ ）

Table 3

ARL$_1$ values for different shifts in simulation experiment 2

Shift→	$(\sigma_1, 0, 0)$	$(0, \sigma_2, 0)$	$(0, 0, \sigma_3)$	$(\sigma_1, \sigma_2, 0)$	$(\sigma_1, 0, \sigma_3)$	$(0, \sigma_2, \sigma_3)$	$(\sigma_1, \sigma_2, \sigma_3)$
Proposed method	34.764	56.448	32.220	18.258	35.438	15.899	16.574
MNP method	64.7150	77.3580	66.2040	21.3070	18.2190	21.1830	8.4030
Shift→	$(2\sigma_1, 0, 0)$	$(0, 2\sigma_2, 0)$	$(0, 0, 2\sigma_3)$	$(2\sigma_1, 2\sigma_2, 0)$	$(2\sigma_1, 0, 2\sigma_3)$	$(0, 2\sigma_2, 2\sigma_3)$	$(2\sigma_1, 2\sigma_2, 2\sigma_3)$
Proposed method	6.179	11.831	5.523	3.032	7.152	2.277	2.645
MNP method	19.5460	25.9220	21.4290	4.0440	3.8670	4.3720	1.7890
Shift→	$(3\sigma_1, 0, 0)$	$(0, 3\sigma_2, 0)$	$(0, 0, 3\sigma_3)$	$(3\sigma_1, 3\sigma_2, 0)$	$(3\sigma_1, 0, 3\sigma_3)$	$(0, 3\sigma_2, 3\sigma_3)$	$(3\sigma_1, 3\sigma_2, 3\sigma_3)$
Proposed method	2.3410	4.0100	1.8930	1.3330	2.5160	1.1880	1.2760
MNP method	7.9310	10.2320	10.0990	1.7670	1.8420	1.8510	1.1150
Shift→	$(-\sigma_1, \sigma_2, 0)$	$(-\sigma_1, 0, \sigma_3)$	$(0, -\sigma_2, \sigma_3)$	$1.5(\sigma_1, -\sigma_2, 0)^*$	$(2\sigma_1, 0, -2\sigma_3)$	$(0, 2\sigma_2, -2\sigma_3)$	$(\sigma_1, \sigma_2, -2\sigma_3)$
Proposed method	243.518	6.243	28.213	2.001	1.186	2.418	1.5510
MNP method	554.106	566.093	235.146	169.011	1015.6	1213.8	907.371

*In this experiment if p_2 shifts $-2\sigma_2$ then $p_2 = -0.0186$ (negative) therefore we shift p_2 to $1.5\sigma_2$ ($p_2 = 0.0161$).

※The results of Table 3 show that the proposed method performs better than MNP procedure in most of the mean-shift scenarios.

6. Conclusion and Recommendations for Future Research

- Monitoring multi attribute processes, where there are some correlations between attributes, is an important issue in statistical quality control. One of these methods is to approximatethe distribution of the correlated attributes with multi normal distribution(with i.i.d)

(計數之間存在相互關聯的,而我們在著假設其為常態(iid-相互獨立)

- In this article, we propose a new transformation technique to reduce the amount of skewness in the marginal first, and then we use a multivariate control charting (T^2 control Chart) on the transformed data.

6. Conclusion and Recommendations for Future Research (Conti.)

- Future research may consider processes with multivariate Poisson distribution and instead of T^2 control chart examine other multivariate control charts like Multivariate Cumulative Sum (MCUSUM) and Multivariate Exponentially Weighted Moving Average (MEWMA) control charts.

References

- Bourke, P. D. (1991). Detecting shift in fraction non conforming using run-length control
- chart with 100% inspection. *J. Qual. Technol.* 23:225–238.
- Box, G. E. P., Cox, D. R. (1964). An analysis of transformations. *J. Roy. Statist. Soc. B*
- 26:211–243.
- Cario, M. C., 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.
- Gibra, I. N. (1978). Economically optimal determination of the parameters of np-control
- charts. *J. Qual. Technol.* 10:12–19.
- Hawkins, D. M. (1991). Regression adjustment for variables in multivariate quality control.
- *J. Qual. Technol.* 25:175–182.
- Hayter, A. J., Tsui, K. L. (1994). Identification and qualification in multivariate quality
- control problems. *J. Qual. Technol.* 26:197–208.
- Healy, J. D. (1987). A note on multivariate CUSUM procedures. *Technometrics* 29:409–412.
- Hotelling, H. (1947). Multivariate quality control. In: Eisenhart, C., Hastay, M. W., Wallis,
- W. A., eds. *Techniques of Statistical Analysis*. New York, NY: McGraw-Hill.
- Jarque, C. M., Bera, A. K. (1987). A test for normality of observations and regression
- residuals. *Int. Statist. Rev.* 55:163–172.
- Johnson, N. L., Kotz, S. (1969). *Discrete Distributions*. New York, NY: John Wiley.
- Jolayemi, J. K. (1994). Convolution of independent binomial variables: an approximation
- method and a comparative study. *Computat. Statist. Data Anal.* 18:403–417.

References (Conti.)

- Jolayemi, J. K. (2000). An optimal design of multi-attribute control charts for processes
- subject to a multiplicity of assignable causes. *Appl. Math. Computat.* 114:187–203.
- Kourti, T., MacGregor, J. F. (1996). Multivariate SPC methods for process and product
- monitoring. *J. Qual. Technol.* 28:409–428.
- 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, Pittsburgh, PA.
- Lowry, C. A., Montgomery, D. C. (1995). A review of multivariate control charts. *IIE Trans.*
- 27:800–810.
- Lowry, C. A., Woodall, W. H., Champ, C. W., Erigdon, S. (1992). A multivariate
- exponentially weighted moving average control chart. *Technometrics* 34:46–53.
- Lu, X. S., Xie, M., Goh, T. N., Lai, C. D. (1998). Control chart for multivariate attribute
- processes. *Int. J. Product. Res.* 36:3477–3489.
- Lucas, J. M., Crosier, R. B. (1982). Fast initial response for CUSUM quality control
- schemes: give your CUSUM ahead start. *Technometrics* 24:199–205.
- Lucas, J. M., Saccucci, M. S. (1990). Exponentially weighted moving average control
- schemes: properties and enhancements. *Technometrics* 32:1–10.
- Marcucci, M. (1985). Monitoring multinomial processes. *J. Qual. Technol.* 17:86–91.
- Montgomery, D. C. (2003). *Introduction to Statistical Quality Control*. 5th ed., New York:
- John Wiley.
- Niaki, S. T. A., Abbasi, B. (2005). Fault diagnosis in multivariate control charts using artificial neural networks. *J. Qual. Reliab. Eng. Int.* 21:825–840.

The End