Computers & Industrial Engineering 98 (2016) 350–359

Contents lists available at ScienceDirect

Computers & Industrial Engineering

journal homepage: www.elsevier.com/locate/caie

A control scheme for autocorrelated bivariate binomial data

Zhen He [a], Zhiqiong Wang [a], Fugee Tsung [b], Yanfen Shang [a,*]

[a] College of Management and Economics, Tianjin University, Tianjin, China

[b] Department of Industrial Engineering and Logistics Management, Hong Kong University of Science and Technology, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 18 August 2015
Received in revised form 2 April 2016
Accepted 1 June 2016
Available online 1 June 2016

Keywords:
Attribute control chart
Autocorrelation
ANSS
Bivariate binomial process
EWMA
Statistical process control

ABSTRACT

The applications of attribute data cover a wide variety of manufacturing and service processes. Moreover, the need for attribute control charts suitable for multivariate and autocorrelated processes has been well recognized. However, there is a scarcity of methods for addressing the autocorrelation among the multivariate attribute data. This paper focuses on autocorrelated bivariate binomial data and proposes a control chart based on the bivariate binomial autoregressive model by integrating a log-likelihood-ratio statistic into Exponentially Weighted Moving Average (EWMA). The performance of the proposed chart in detecting sustained shifts in parameters is evaluated by the Average Number of Samples to Signal (ANSS). The simulation results show that the proposed chart is generally more robust and sensitive to small and moderate shifts compared with the existing charts. In addition, a real example from laser gyroscope manufacturing is used to illustrate the implementation of the proposed approach.

1. Introduction

Statistical Process Control (SPC) has been demonstrated to be an important tool for monitoring process or product quality. In modern manufacturing and service processes, it is more common to monitor two or more correlated quality characteristics simultaneously rather than a single one. In such situations, multivariate control charts must be considered. In general, there are two types of control charts in multivariate SPC – those used for measurement data and those used for attributes. In the past decades, some multivariate measurement control charts have been proposed, such as T^2 control charts (Hotelling, 1947; Khoo, Wu, Castagliola, & Lee, 2013; Sparks, 2015), Multivariate Cumulative Sum (MCUSUM) control charts (Crosier, 1988; Dai, Luo, Li, & Wang, 2011; Pignatiello & Runger, 1990; Woodall & Ncube, 1985), Multivariate Exponentially Weighted Moving Average (MEWMA) control charts (Lowry, Woodall, Champ, & Rigdon, 1992; Shen, Tsung, & Zou, 2014; Zou & Tsung, 2008) and multivariate control charts based on neural networks (Hwarng, 2008; Cheng & Cheng, 2011). However, fewer multivariate control charts dealing with attribute data have been developed, despite the fact that the multivariate attributes are common in many applications.

The quality characteristics of the multivariate attribute processes could be modeled by either multivariate Poisson distribution or multivariate binomial/multinomial distribution (Chiu & Kuo, 2008). This study focuses on the binomial approach and the Poisson could be subject of future research. Lu (1998) proposed a Shewhart-type multivariate np control chart (MNP control chart) to monitor multivariate binomial observations. They defined a new statistic, which is the weighted sum of nonconforming counts of each quality characteristic, while considering the correlations between the attributes. Chiu and Kuo (2010) presented a new control chart to monitor a bivariate binomial process. They simultaneously took the false alarms (type I errors) and the correlation between attributes into account. Based on log-linear models, Li, Tsung, and Zou (2014) proposed a Phase II control chart, which can be applied to the unified framework of multivariate binomial and multivariate multinomial processes. A recent review on the multinomial and multi-attribute control charts has been given by Topalidou and Psarakis (2009).

These classical control charts focus on the cases where multivariate binomial observations are time-independent. However, because of the fast sampling schemes relative to the process dynamics and the natural behavior of many real-life applications, the successive observations are often autocorrelated (Soleimani, Noorossana, & Amiri, 2009). One real example is the manufacturing of the laser gyroscope which can be found in Section 2. It is well known that autocorrelation might affect the performance of a control chart if the chart is designed for processes with independent observations (Hwarng & Wang, 2010). For this reason, effective statistical methodologies are required. Based on the two-state Markov chain model, some control charts were proposed for autocorrelated univariate binomial or Bernoulli data (e.g., Bhat & Lal, 1990; Lai, Xie, & Govindaraju, 2000; Mousavi & Reynolds, 2009; Shepherd, Champ, Rigdon, & Fuller, 2007; Wang & Reynolds, 2014; Weiß,

* Corresponding author.
E-mail address: syf8110@gmail.com (Y. Shang).

http://dx.doi.org/10.1016/j.cie.2016.06.001
0360-8352/

2009a). Alternatively, Weiß (2009b) investigated the binomial AR (1) model for correlated processes with binomial marginals and proposed some control schemes for such processes. In contrast to the numerous univariate approaches to autocorrelated binomial data, extensions of these approaches to multivariate control charts lags behind. Based upon the modified Elman neural network capabilities, Niaki and Nasaji (2011) developed a method for monitoring autocorrelated multivariate attribute data. Recently, Dokouhaki and Noorossana (2013) used the Markov approach and the copula approach for modeling the bivariate autocorrelated binary series, and developed a CUSUM chart. However, how to apply SPC methods in monitoring autocorrelated multivariate binomial processes remains a challenge and has not been investigated sufficiently in the literature.

The purpose of this article is to build an effective control chart for the autocorrelated bivariate binomial observations, which is based on the bivariate binomial autoregressive model and integrates the likelihood ratio test with the EWMA procedure, referred to as LR-EWMA chart. The rest of the paper is organized as follows. First, we introduce an example from manufacturing industry that motivates this research in Section 2. The bivariate binomial autoregressive model is introduced in Section 3. In Section 4, we discuss the design of the proposed control chart. Following that, the performance analysis and performance comparison are presented in Section 5. In Section 6, we demonstrate implementation of the proposed method using a real example. Finally, concluding remarks are offered to summarize the major contributions of the article and suggest issues for future research. Some technical details are given in Appendices A and B.

2. A motivating example

We use an example taken from the manufacture of a laser gyroscope to motivate this research. A laser gyroscope is a device which employs a ring laser to measure rotation speed. In the production of a laser gyroscope, there are some relevant tests involving in scale factor, maximum input angular rate, lock in threshold, input axis misalignment, bias and random walk coefficient. Among the tests, lock-in threshold test is the most important, and it is concentrated on two quality characteristics: In-Threshold and Out-Threshold. In-Threshold is based on the method of gradually decreasing laser rotation rate until the output signals disappear, while Out-Threshold is obtained by gradually increasing laser rotation rate until the output signals appear. Note that, as expected from engineering principles, the two quality characteristics are correlated to some extent. The lock-in threshold test are conducted by sampling inspection, and the inspection output is "defective" or "good" based on the specifications of the two quality characteristics. If the values of In-Threshold and Out-Threshold are both less than $0.15^\circ/s$, the laser gyroscope is a "good" one, otherwise it is a "defective" one. Thus, bivariate binomial data are collected after a certain time. During the process of inspection, In-Threshold and Out-Threshold are automatically measured by an electronic device at a very high speed. To this end, the collected bivariate binomial data is more likely to be autocorrelated.

To protect proprietary and confidential information, for illustrative purpose, we only sample 20 laser gyroscopes from each batch and obtain 100 samples. In-Threshold and Out-Threshold for each laser gyroscope are then inspected as conforming or nonconforming by some electronic devices. The number of nonconforming items per sample for the two attributes are denoted by $X_{t,1}$ $(t = 1, 2, \ldots, 100)$ and $X_{t,2}$ $(t = 1, 2, \ldots, 100)$, respectively. $X_{t,1}$ and $X_{t,2}$ are binomially-distributed random variables and $\boldsymbol{X}_t = [X_{t,1}\ X_{t,2}]'$ is a bivariate variable with bivariate binomial distribution. Empirical AutoCorrelation Functions (ACFs) and cross-correlation of the two characteristics are plotted in Figs. 1 and 2, respectively. Both ACFs show a first order autocorrelation around 0.4, indicating that a first-order model should be adequate to model the serial dependence in this case. Furthermore, the correlation between the two characteristics is around 0.7, clearly indicating strong dependence. Therefore, multivariate attribute processes with autocorrelated data, in particular, the autocorrelated bivariate binomial process do exist. In the remainder of this paper, we propose a control scheme for the autocorrelated bivariate binomial data and give a step-by-step demonstration of how to implement the proposed scheme in practice.

3. The bivariate binomial autoregressive model

To model the bivariate time series of counts with a finite range in the motivating example, we use a bivariate binomial autoregressive model which could describe this kind of data properly. In this section, the bivariate binomial distribution is introduced first, and then the bivariate binomial autoregressive model is developed based on the bivariate binomial distribution and bivariate binomial thinning operation.

3.1. Bivariate binomial distribution

According to the notation of Marshall and Olkin (1985), a bivariate binomial distribution could be generated by the bivariate Bernoulli distribution. If $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_k$ are independent bivariate Bernoulli random variables and they all take the four possible values (1,1), (1,0), (0,1), (0,0) with probabilities $p_{11}, p_{10}, p_{01}, p_{00}$, respectively, then the sum $\boldsymbol{Y} = [Y_1\ Y_2]' = \boldsymbol{Z}_1 + \cdots + \boldsymbol{Z}_k$ follows a bivariate binomial distribution of Type I, abbreviated as $\mathrm{BVB_I}(k; p_1, p_2, \Phi)$ (Marshall & Olkin, 1985; Scotto, Weiß, Silva, & Pereira, 2014), where $p_1 = p_{11} + p_{10}$ and $p_2 = p_{11} + p_{01}$ are the marginal probabilities and $\Phi = (p_{11} - p_1 p_2)/\sqrt{p_1 p_2 (1-p_1)(1-p_2)}$ is the correlation coefficient between Y_1 and Y_2. Its marginals are univariately binomially distributed, $Y_1 \sim \mathrm{B}(k, p_1)$ and $Y_2 \sim \mathrm{B}(k, p_2)$. Scotto et al. (2014) proved that the range of Φ is restricted to

$$\max\left\{-\sqrt{\frac{p_1p_2}{(1-p_1)(1-p_2)}}, -\sqrt{\frac{(1-p_1)(1-p_2)}{p_1p_2}}\right\} < \Phi < \min\left\{\sqrt{\frac{p_1(1-p_2)}{p_2(1-p_1)}}, \sqrt{\frac{p_2(1-p_1)}{p_1(1-p_2)}}\right\} \tag{1}$$

As a generalization, we can let $\boldsymbol{Y}, U, V$ be independent random variables, where $\boldsymbol{Y} \sim \mathrm{BVB_I}(k; p_1, p_2, \Phi), U \sim \mathrm{B}(n_1 - k, p_1)$, $V \sim \mathrm{B}(n_2 - k, p_2)$ and $k = \min(n_1, n_2)$ (Biswas & Hwang, 2002; Marshall & Olkin, 1985). Then $\boldsymbol{X} = [X_1\ X_2]' = [Y_1 + U\ Y_2 + V]'$ is said to follow a bivariate binomial distribution of Type II, abbreviated as $\mathrm{BVB_{II}}(n_1, n_2, k; p_1, p_2, \Phi)$ (Scotto et al., 2014). The joint probability of the bivariate binomial distribution is given by Hamdan and Jensen (1976) and Marshall and Olkin (1985)

$$\begin{aligned} p_{(n_1,n_2;p_1,p_2,p_{11})} &= P(X_1 = x_1, X_2 = x_2) \\ &= \sum_{j_1=0}^{\min(x_1,n_1-k)} \sum_{j_2=0}^{\min(x_2,n_2-k)} \sum_{i=\max(0,x_1-j_1+x_2-j_2-k)}^{\min(x_1-j_1,x_2-j_2)} \binom{n_1-k}{j_1} p_1^{j_1}(1-p_1)^{n_1-k-j_1} \\ &\quad \times \binom{n_2-k}{j_2} p_2^{j_2}(1-p_2)^{n_2-k-j_2} \\ &\quad \times \binom{k}{i, x_1-j_1-i, x_2-j_2-i, k+i+j_1+j_2-x_1-x_2} \\ &\quad \times p_{11}^{i}(p_1-p_{11})^{x_1-j_1-i}(p_2-p_{11})^{x_2-j_2-i}(1+p_{11}-p_1-p_2)^{k+i+j_1+j_2-x_1-x_2}. \end{aligned} \tag{2}$$