Computers & Industrial Engineering 105 (2017) 329–337

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Combined Shewhart CUSUM charts using auxiliary variable

Ridwan A. Sanusi [a,c,*], Mu'azu Ramat Abujiya [b], Muhammad Riaz [a], Nasir Abbas [a]

[a] Department of Mathematics and Statistics, King Fahad University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
[b] Preparatory Year Mathematics Program, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia
[c] Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong

ARTICLE INFO

Article history:
Received 12 May 2016
Received in revised form 6 December 2016
Accepted 18 January 2017
Available online 21 January 2017

Keywords:
Auxiliary information
Average run length
Combined Shewhart-CUSUM
Control chart
Extra quadratic loss

ABSTRACT

A control chart is an important statistical tool for monitoring disturbances in a statistical process, and it is richly applied in the industrial sector, the health sector and the agricultural sector, among others. The Shewhart chart and the Cumulative Sum (CUSUM) chart are traditionally used for detecting large shifts and small shifts, respectively, while the Combined Shewhart-CUSUM (CSC) monitors both small and large shifts. Using auxiliary information, we propose new CSC (M_iCSC) charts with more efficient estimators (the Regression-type estimator, the Ratio estimator, the Singh and Tailor estimator, the power ratio-type estimator, and the Kadilar and Cingi estimators) for estimating the location parameter. We compare the charts using average run length, standard deviation of the run length and extra quadratic loss, with other existing charts of the same purpose and found out that some of the M_iCSC charts outperform their existing counterparts. At last, a real-life industrial example is provided.

1. Introduction

The most widely known quality control chart, the Shewhart chart, was proposed by Shewhart (1924). It detects shifts in a production process by signaling when a process goes beyond some particular threshold limits known as control limits. Shewhart chart makes use of the information when the process goes out of the control limits and ignores the information when the process is within the control limits, i.e. in-control. Due to this fact, the chart is sensitive for detecting large shifts (or disturbances) in a process. Roberts (1959) and Page (1954) proposed the Exponentially Weighted Moving Average (EWMA) chart and the Cumulative Sum (CUSUM) chart, respectively, which make use of the information when the process gets out-of-control and even when the process is in-control, hence, these charts are sensitive to small and moderate shifts in a process. Other modifications of these charts have been proposed to increase their efficiency in terms of time, cost, and simplicity of usage and expression.

The plotting statistic of CUSUM chart assumes normality. What if the plotting statistic is not normally distributed or its normality is altered? Nazir, Riaz, Does, and Abbas (2013) answered these questions by suggesting some charts which are not normally distributed or their normality has been altered. They aimed at finding charts that perform practically well under normal, contaminated normal, non-normal, and special cause contaminated parent cases. Based on mean, median, Hodge-Lehman, midrange and trimean statistics, they proposed different CUSUM charts for phase II monitoring of the location parameter and computed their performance measure using the average run length (ARL) approach. Abujiya, Lee, and Riaz (2015) suggested the use of well-structured sampling techniques, such as the double ranked set sampling, the median-double ranked set sampling, and the double-median ranked set sampling, to significantly improve the performance of the CUSUM chart, without inflating the false alarm rate. They compared their proposed charts with some existing charts and found out that their charts perform better.

Due to the advancement in technology and industrial processes, emphasis has been made on the implementation of the CUSUM chart to existing Levey-Jennings or Shewhart control charts (Westgard, Groth, Aronsson, & de Verdier, 1977). These can be done manually using control charts or in computerized quality control systems. Westgard et al. (1977) applied this concept to improve quality control in clinical chemistry. A combination of the Shewhart chart and CUSUM chart was observed by Lucas (1982), after which some scholars improved the chart by proposing more efficient charts. Combined Shewhart-CUSUM (hereafter called "CSC") for location parameter can be optimized over the entire mean shift range by adding an extra parameter (w), known as the exponential of the sample mean shift, to the structure of the CSC. This will improve its performance and it will not increase the

* Corresponding author at: Department of Systems Engineering and Engineering Management, City University of Hong Kong, Kowloon, Hong Kong.
E-mail address: amhigher2010@yahoo.com (R.A. Sanusi).

http://dx.doi.org/10.1016/j.cie.2017.01.018
0360-8352/

difficulty level of understanding and implementing the chart (Wu, Yang, Jiang, & Khoo, 2008). The CSC, which has a wide range of applications, attracts the attention of Environmentalists, and it is the only quality control chart directly recommended by the United States Environment Protection Agency for intra-well monitoring (Gibbons, 1999). Abujiya, Riaz, and Lee (2013) replaced the traditional simple random sampling in the plotting statistic of the CSC with ranked set sampling.

The control statistics of the classical Shewhart, CUSUM, and CSC charts for monitoring location parameter are based on the usual unbiased simple mean estimator ($\bar{x} = (1/n)\sum_{i=1}^{n} x_i$) for estimating the population mean. However, in the field of sample survey, different authors have suggested some estimators for estimating the population mean, which are more efficient than the simple mean estimator, in terms of their mean squared error (MSE). Some of these estimators require the use of auxiliary variable(s) which are cheap, easy and affordable to get, and also, with known population parameters (Cochran, 1977). According to Cochran (1977), the correlation between the study variable and the auxiliary variable will serve as an advantage to increase the precision of estimation. Sukhatme and Sukhatme (1970) proposed the regression estimator for estimating the mean, while the power ratio-type estimator and modified ratio-type estimator were suggested by Srivastava (1967) and Ahmad, Abbasi, Riaz, and Abbas (2014), respectively. Interested readers can see Singh and Tailor (2003), Kadilar and Cingi (2004, 2006a, 2006b), Gupta and Shabbir (2008) and Adebola, Adegoke, and Sanusi (2015) for different forms of a transformed ratio estimator.

Zhang (1992) suggested the cause-selecting control chart, while Riaz (2008b) popularised the use of auxiliary information at the estimation stage, for monitoring dispersion parameter. He concluded that the chart is better than the R chart, the S chart and the S^2 chart. Furthermore, Riaz (2008a) suggested similar chart for location parameter estimation, which was also superior to the Shewhart chart, the regression chart, and the cause-selecting control chart. Assuming stability of parameters, Ahmad, Riaz, Abbasi, and Lin (2014) proposed new Shewhart charts based on auxiliary information for non-cascading processes. The charts monitor a dispersion parameter in an efficient way. The superiority of the charts over competing charts was shown using the ARL, relative average run length (RARL) and extra quadratic loss (EQL) under t and normal distributed process environment. Similar work was also done for location parameter monitoring, and it was found out that there is an improvement in the detection ability of Shewhart chart based on the level of correlation between the concerned variables (Riaz, 2015).

Since most of the estimators are more efficient than the simple mean estimator based on a simple random sample, their introduction to the plotting statistic(s) of the Shewhart chart, the CUSUM chart, and the CSC chart would result in efficient control charts. Hence, this study aims at optimizing the CSC chart by introducing some efficient estimators to its plotting statistics. These estimators use auxiliary information in the sampling stage. This is helpful whenever there is no information about the population of the variable of interest, but there is information about a closely related variable(s) which is cheap and affordable to get. Cheap and affordable in the sense that little or no resources (money or time) are needed to get the extra information. In statistical process control, one of the relevancies of introducing an auxiliary variable can be found in the platinum refinery, where the quality of the process generally depends on the quantity of platinum metal that is correlated with the magnitude of other metals (Ahmad, Abbasi, et al., 2014; Hawkins, 1991).

The rest of this article is organized as follows: Location estimators and their properties are explained in the next section; The general structure of the proposed charts is explained in Section 3; Section 4 explains the performance measures for evaluating the proposed charts and compares the proposed charts with their existing counterparts; Section 5 gives an illustrative example; and finally, conclusions and recommendations are given in Section 6.

2. Location estimators and their properties

We assume that a process has a quality characteristic of interest X and an auxiliary quality characteristic A. Let the population parameters of X and A, respectively, be represented as $\bar{X}$ and $\bar{A}$ for the means; σ_X^2 and σ_A^2 for the variances; $C_X = \sigma_X/\bar{X}$ and $C_A = \sigma_A/\bar{A}$ for the coefficient of variations; $\beta_{2(X)}$ and $\beta_{2(A)}$ for the coefficient of kurtoses; σ_{XA} for the covariance between X and A; and ρ_{XA} for the correlation coefficient. Let the sample statistics of X and A, respectively, be represented as $\bar{x}$ and $\bar{a}$ for the means; s_x^2 and s_a^2 for the variances; c_x and c_a for the coefficient of variations; s_{xa} for the covariance; and r_{xa} for the correlation coefficient. Let x_i and (x_i, a_i) be univariate and bivariate sample respectively, where $i = 1, 2, \ldots, n$ and $n =$ sample size. From the sample statistics, we have $\bar{x} = \sum_{i=1}^{n} x_i/n$, $\bar{a} = \sum_{i=1}^{n} a_i/n$, $s_x^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2/(n-1)$, $s_a^2 = \sum_{i=1}^{n}(a_i - \bar{a})^2/(n-1)$, $c_x = s_x/\bar{x}$, $c_a = s_a/\bar{a}$ and $r_{xa} = s_{xa}/s_x s_a$. Based on this introduction, some efficient estimators with one auxiliary variable for estimating the mean of a quality process characteristic, assuming sampling with replacement, are presented in Eqs. (1)–(10) with their respective bias (B) and MSE. The estimators differ from one another in terms of their efficiency and simplicity of their structures. They perform better than the simple mean estimator for different cases. The regression estimator performs better than the simple mean estimator when there is a non-zero correlation between the study variable and the auxiliary variable. In addition, the ratio estimator is more efficient than the simple mean estimator when the correlation between the study variable, and the auxiliary variable is greater than 0.5. Further information about the estimators can be found in Srivastava (1967), Cochran (1977), Singh and Tailor (2003) and Kadilar and Cingi (2004).

(i) *The Simple Random Sampling Estimator* (Cochran, 1977)

$$M_1 = \sum_{i=1}^{n} x_i/n \tag{1}$$

with $B(M_1) = 0$ and $MSE(M_1) = \sigma_X^2/n$.

(ii) *The Regression-Type Estimator (Difference Estimator)* (Cochran, 1977)

$$M_2 = \bar{x} + b_{XA}(\bar{A} - \bar{a}) \tag{2}$$

where $b_{XA} = -\rho_{XA}\sigma_X/\sigma_A$, with $B(M_2) = 0$ and $MSE\,(M_2) = \sigma_X^2(1 - \rho_{XA}^2)/n$.

The bias and the MSE of the next estimators are given up to the first order approximation.

(iii) *The Ratio Estimator* (Cochran, 1977)

$$M_3 = \bar{x}\frac{\bar{A}}{\bar{a}} \tag{3}$$

with $B(M_3) = \bar{X}\left(C_A^2 - \rho_{XA}C_X C_A\right)$ and $MSE(M_3) = \bar{X}^2\left(C_X^2 + C_A^2 - 2\rho_{XA}C_X C_A\right)$.

(iv) *The Singh and Tailor Estimator* (Singh & Tailor, 2003)

$$M_4 = \bar{x}\left(\frac{\bar{A} + \rho_{XA}}{\bar{a} + \rho_{XA}}\right) \tag{4}$$