Computers & Operations Research 40 (2013) 3021–3032

Contents lists available at ScienceDirect

Computers & Operations Research

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

Optimal linear combination of Poisson variables for multivariate statistical process control

Eugenio K. Epprecht [a], Francisco Aparisi [b,*], Sandra García-Bustos [c]

[a] Pontificia Universidade Católica de Río de Janeiro, Río de Janeiro, Brazil
[b] Departamento de Estadística e I. O. Aplicadas y Calidad, Universidad Politécnica de Valencia, 46022 Valencia, Spain
[c] Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador

ARTICLE INFO

Available online 17 July 2013

Keywords:
Control chart
Poisson
Genetic algorithms
Multivariate SPC

ABSTRACT

In this paper we analyze the monitoring of p Poisson quality characteristics simultaneously, developing a new multivariate control chart based on the linear combination of the Poisson variables, the LCP control chart. The optimization of the coefficients of this linear combination (and control limit) for minimizing the out-of-control ARL is constrained by the desired in-control ARL. In order to facilitate the use of this new control chart the optimization is carried out employing user-friendly Windows© software, which also makes a comparison of performance between this chart and other schemes based on monitoring a set of Poisson variables; namely a control chart on the sum of the variables (MP chart), a control chart on their maximum (MX chart) and an optimized set of univariate Poisson charts (Multiple scheme). The LCP control chart shows very good performance. First, the desired in-control ARL (ARL_0) is perfectly matched because the linear combination of Poisson variables is not constrained to integer values, which is an advantage over the rest of charts, which cannot in general match the required ARL_0 value. Second, in the vast majority of cases this scheme signals process shifts faster than the rest of the charts.

1. Introduction

The design of a process control chart for monitoring a single Poisson variable is an easy task: 3-standard-deviations control limits are given in closed form as a function of the in-control mean of the variable; alternatively, probability limits can be determined by inverting the cumulative distribution so as to achieve an acceptable false-alarm probability (although exactly matching a specified false-alarm probability is not in general possible, due to the discreteness of the Poisson variable, which makes its cumulative distribution discontinuous). For details, see, for instance, Montgomery [13]. When, to ensure the quality of the produced items, several quality characteristics need to be monitored simultaneously, the practitioner has two options: (i) a control scheme based on one chart for each variable (Multiple scheme) and (ii) a control scheme based on a single control chart (multivariate scheme). There is a large bibliography on multivariate and multiple statistical process control for continuous variables; see, for example, [3]. However, very little research has been done when the variables to be monitored are discrete, and in the specific case of this paper, when they follow the Poisson distribution.

Holgate [7] investigated the bivariate distribution of correlated Poisson variables. His model assumes that there is a common factor for all the variables, plus a unique factor for each of the observed variables. For example, the common factor in a cutting process can be the rotational speed of a saw when cutting wood panels. The speed affects the frequencies of two types of possible defects in the surface. The influence of this common factor in the number of defects of the observed variables X_i, $i=1, 2$, is represented by an unobservable variable Y_0, and the influence of each individual factor on the respective observed variable X_i is represented by an unobservable variable Y_i so that $X_i=Y_0+Y_i$. The common part Y_0 responds for the correlation between the observed variables X_i. This model can be easily extended to more than two variables, simply by making $i=1, 2, \ldots, p$, with $p>2$. This extended Holgate′s model is assumed throughout this paper. Let λ_i denote the mean of each Poisson variable Y_i. It is straightforward to obtain:

$$E(X_i)=\lambda_0+\lambda_i, \quad \mathrm{Cov}(X_i,\ X_j)=\lambda_0, \quad \rho(X_i,X_j)=\frac{\lambda_0}{\sqrt{(\lambda_0+\lambda_i)(\lambda_0+\lambda_j)}} \tag{1}$$

Another reference in multiple process control by attributes is Patel [14], who developed a multivariate control chart based on the multivariate normal approximation to the binomial distribution. More recently, Skinner et al. [16] proposed to employ the Deviance Residual of the general linear model as the statistic to monitor

* Corresponding author. Tel.: +34 96 3877494; fax: +34 96 3877499.
E-mail addresses: faparisi@eio.upv.es, profepaco@hotmail.com (F. Aparisi).

0305-0548/$ - see front matter
http://dx.doi.org/10.1016/j.cor.2013.07.007

several independent Poisson variables. Chiu and Kuo [4] proposed the multivariate monitoring of several Poisson variables by the sum of all of them, in what they named the MP control chart. They found the distribution of this sum and analyzed the performance under Holgate's model for correlation. Another multivariate proposal was presented by Ho and Costa [6], for the case of a bivariate Poisson. They proposed monitoring the variables by their difference (DX chart) and also by the maximum of them (MX chart). An exhaustive comparison of performance is presented in their paper.

With respect to the Multiple scheme, designing a set of univariate Poisson control charts is not an easy task. Aparisi et al. [2] developed a procedure to design this set of charts, considering that the set has to achieve a required in-control ARL. They provided user-friendly software to perform the ARL calculations and optimize the charts parameters. The difficulty is that, since the in-control means of the variables are fixed values, then the only parameter in each chart that one has freedom to change (as a decision variable) is the upper control limit. Since, however, the Poisson variable can take only integer values, its cumulative distribution is discontinuous, which prevents (except by a lucky chance) adjusting the limits of this set of charts for a specified false-alarm probability so as to match the required in-control ARL. Sometimes a close value can be obtained, but the majority of times there is a large distance between the desired value and the closest ARL obtained. For that reason, in many cases, we have to choose from a scheme where the number of false alarms is high, or from a scheme that is not powerful for detecting the process shifts, because the in-control ARL is too large (which is tantamount to saying that the control limit is too high for providing good power for the chart). A similar problem occurs with all previously cited charts based on the Poisson distribution.

Laungrungrong et al. [11] developed a multivariate EWMA control chart for Poisson variables (the MPEWMA control chart) assuming Holgate's model as well. The MPEWMA chart was compared with the traditional MEWMA control chart [12] applied to the Poisson variables. They show that the use of the MEWMA chart only produces reasonable results when the mean of the Poisson variables have a value of 5 of more. When this condition is not satisfied, the MEWMA chart tends to produce more false alarms.

One of the applications where often there is the need of monitoring several discrete (and often Poisson) variables is health surveillance. There is a vast bibliography about this field, where the new control chart developed in this paper may be applied. A good review of the use of control charts in health-care and public-health surveillance is Woodall [17]. Other interesting papers, where the interested reader can find more information and references are Joner et al. [9] and Jiang et al. [8].

In this paper we propose a new multivariate control chart, the Linear Combination of Poisson counts (LCP) chart that can be optimized to obtain the required in-control ARL solving one of the problems of multivariate control charts for Poisson variables. In order to promote the use of this chart, user-friendly software (available by request) has been developed. This software finds the best linear combination of the Poisson variables and the best control limits in order to minimize the out-of-control ARL. Moreover, this computer program also optimizes the multivariate MP and MX control charts and the Multiple scheme, and makes a complete comparison of performance among all the charts. Therefore, the end-user can determine which is the most efficient control scheme for his/her particular needs.

Using the program, we optimized the LCP chart and the other three control schemes just mentioned for a large number of cases, in order to analyze and compare their performances. The LCP chart has shown to be, the large majority of times, the most powerful one to detect process shifts.

The remainder of this paper is organized as follows: Section 2 presents the theoretical basis of the (optimized) Multiple scheme (several univariate control charts) and of the multivariate schemes focused here (sum, maximum and linear combination of the variables). Section 3 gives a formal definition of the optimization problem. Section 4 describes a computer program that has been developed with the aim of helping the final user with selecting the best control scheme for his/her process. This software allows obtaining the control limits for the charts mentioned in Section 2, when $p=2$, 3 or 4 Poisson variables are monitored, in order to obtain the desired in-control ARL, or the closest possible. In addition, the software carries out a complete performance comparison between the different control charts. A sensitivity analysis appears in Section 5. A general comparison of performance is shown in Section 6. Finally, Section 7 summarizes the conclusions of this paper.

2. Controlling several Poisson variables: multivariate and multiple approaches

In this section all the Poisson control charts that will be considered in the comparison of performance are introduced. The performance measure that will be used is the Average Run Length (ARL), which is the most used one to make this type of comparisons. It is the average number of points on the control chart until the chart signals, i.e., until a point is plot outside the control limits or just on the control limits. A signal will be a false alarm if the process is in control, that is, with no shifts in its parameters. Therefore a large in-control ARL is usually required. On the other hand, shifts in the process parameters (out-of-control state) must be detected quickly, therefore small out-of-control ARLs are desired. In the type of charts analyzed in this paper, where the statistics (points) to be plotted are independent, the number of samples until a signal follows a geometric distribution, and the ARL is the reciprocal of the probability of a signal, formally, $\text{ARL}=1/P(\text{signal})$.

2.1. The multiple scheme: multiple univariate Poisson control charts

The Multiple scheme consists of monitoring each Poisson variable by one control chart. Aparisi et al. [2] offer a computer program that optimizes this set of univariate control charts to minimize the ARL for a given shift. The authors also analyze the performance of this scheme for a large number of cases. In that paper, as in the present one, it is assumed that each of these Poisson's charts has only the upper control limit, as normally the practitioner is only interested in detecting shifts that increase the number of defects. Therefore, if p Poisson variables are to be monitored, p upper control limits should be determined. The search for their optimal values is conducted taking into account that a specified value, ARL_0, is required for the joint in-control ARL of the set of p control charts. However, it is generally not feasible to achieve the required value of ARL_0 given that the Poisson variables can take only integer values. Aparisi et al. [2] search the control limits following the objective of obtaining the closest possible value of in-control ARL to the required ARL_0, but always larger than ARL_0 if the exact value cannot be matched. This criterion will be employed in this article as well.

The statistic to be plotted in each chart is the observed value of the variable, X_i. One chart shows a signal when the observed value is greater or equal than the control limit, i.e., $X_i \geq \text{UCL}_i$. Hence, this set of p control charts shows an out-of-control signal when one or more control charts signal.

2.2. Multivariate schemes

As it was previously mentioned, the multivariate approach consists in employing a unique statistic for all the p variables to be monitored. For example, in the case of Poisson variables, Chiu