PRES LUNAM
Ecole Doctorale STIM
Sciences et Technologies de l'Information et Mathématiques

Spécialité : Au, Ro, TS[1]
Laboratoire : IRCCyN
Equipe : SLP

RUN RULES c CHART WHEN PROCESS PARAMETER ARE UNKNOWN

Shu WU
Mél : shu.wu@univ-nantes.fr

Résumé : Run rules are used to indicate out-of-statistical control situations. It can increase the power of the standard Shewhart chart for detecting small shift. The waiting time distribution for a special run given a sequence of independent success, the parameters of the distribution are usually evaluated under the assumption of known parameters. However, the process parameters are rarely known and need to be estimated from an in-control Phase I data set. This paper will discuss the run rules c chart when the process parameter is estimated and compare them in the case where the process paramaters are assumed known.

Mots clés : *attribute control charts ; run rules ; average run length ; Markov chain.*
Collaborations : Institut Universitaire de Technologie de Nantes Département Qualité Logistique Industrielle et Organisation.

1 Introduction

As one of the Statistical process control (SPC) tools, control charts have been proved to be usefull for indentifying diversification in the processes. An indispensable assumption for the development of control charts is that the process parameters are assumed known. In practice, the distribution of the data and the process parameters are rarely known. The process parameters are usually estimated from an in-control historical data set (Phase I). When the parameters are estimated, the performance of the control charts differs from the known parameters case due to the variability of the estimators used during the Phase I. Concerning attribute control charts with estimated process parameters, as far as we know, very little has been done and the only significant contributions were proposed by [1],[2], [3], [4], [5], [6]. Most of the research are about $\bar{X}$ type control charts. An indispensable assumption for the development to control charts is that the process parameters are assumed known. No previous work already investigated Run Rules c charts when the in-control process parameters are estimated. Therefore, the goal of this paper is to cover this unstudied topic.

2 Run rules c chart with known parameter c_0

Let $\{Y_{i,1}, \dots, Y_{i,n}\}$, $i = 1, 2, \dots$, be a Phase II sample of independent random variables $Y_{i,j}$ such that $Y_i = \sum_{j=1}^{n} Y_{i,j} \sim P(c_1)$, i.e. a Poisson distribution of parameter c_1, where c_1 is an out-of-control number of nonconformities. The 2-out-of-3 Run Rules c chart an out-of-control signal is obtained if two out of three successive values Y_i are plotted above an upper control limit UCL_c or two out of three successive points are plotted below a lower control limit LCL_c with

$$LCL_c = \lceil c_0 - K_{c,2,3}\sqrt{c_0} \rceil \quad (1)$$
$$UCL_c = \lfloor c_0 + K_{c,2,3}\sqrt{c_0} \rfloor \quad (2)$$

where $\lceil \dots \rceil$ and $\lfloor \dots \rfloor$ denote the rounded up and rounded down integer, respectively, c_0 is the *known* in-control number of nonconformities and $K_{c,2,3} > 0$ is a constant.(The $\mathrm{RR}_{2,3} - c$ chart may easily be extended to "longer" Run Rules like, for instance, the 3-out-of-4 and the 4-out-of-5 Run Rules c chart)

A new point Y_i will fall into one of the above seven states or into the 8th state where 2 out of the 3 successive points are plotted out of the control limits, i.e. an out-of-control status is signalled. The run length properties (p.d.f., c.d.f, ARL, $SDRL$) of the $\mathrm{RR}_{2,3} - c$ with known parameter could be obtained

1. Automatique, Robotique, Traitement du Signal

using the following Markov Chain matrix $\mathbf{P}_{(8\times 8)}$ corresponding to the eight previously defined states (where the 8th state corresponds to the absorbing state) :

$$\mathbf{P} = \left(\begin{array}{cc} \mathbf{Q} & \mathbf{r} \\ \mathbf{0}^T & 1 \end{array} \right) = \left(\begin{array}{ccccccc|c} 0 & 0 & 0 & p_C & p_L & 0 & 0 & p_U \\ 0 & 0 & 0 & 0 & 0 & 0 & p_C & p_L + p_U \\ p_C & p_L & 0 & 0 & 0 & 0 & 0 & p_U \\ 0 & 0 & p_U & p_C & p_L & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_U & p_C & p_L \\ p_C & 0 & 0 & 0 & 0 & 0 & 0 & p_L + p_U \\ 0 & 0 & p_U & p_C & 0 & 0 & 0 & p_L \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

where $\mathbf{0}_{(7\times 1)} = (0, 0, \ldots, 0)^T$, $\mathbf{Q}_{(7\times 7)}$ is the matrix of transient probabilities, vector $\mathbf{r}_{(7\times 1)}$ satisfies $\mathbf{r} = \mathbf{1} - \mathbf{Q}\mathbf{1}$ (i.e. row probabilities must sum to 1), with $\mathbf{1}_{(7\times 1)} = (1, 1, 1, 1, 1, 1, 1)^T$. By definition, the probabilities p_L, p_U and p_C are equal to

$$\begin{aligned} p_L &= P(Y_i < LCL) = P(Y_i \leq LCL - 1) \\ p_U &= P(Y_i > UCL) = 1 - P(Y_i \leq UCL) \\ P_C &= P(LCL \leq Y_i \leq UCL) = P(Y_i \leq UCL) - P(Y_i \leq LCL - 1) \end{aligned}$$

where $F_P(x|\lambda)$ is the c.d.f. of the Poisson distribution of parameter λ, Consequently, the probability density function $f_L(\ell)$ and the cumulative distribution function $F_L(\ell)$ of the run length L of the $\mathrm{RR}_{2,3} - c$ chart with known parameter are defined for $\ell = \{2, 3, \ldots\}$ and are equal to

$$f_L(\ell) = \mathbf{q}^T \mathbf{Q}^{\ell-1} \mathbf{r}, \qquad F_L(\ell) = 1 - \mathbf{q}^T \mathbf{Q}^{\ell} \mathbf{1},$$

where $\mathbf{q}_{(7\times 1)} = (0, 0, 0, 1, 0, 0, 0)^T$ is the vector of initial probabilities associated with the transient states (i.e. the initial state is the 4th one). The mean (ARL) and standard-deviation ($SDRL$) of the run length are equal to

$$ARL = \nu_1, \qquad SDRL = \sqrt{\mu_2},$$

with

$$\nu_1 = \mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{1}, \quad \nu_2 = 2\mathbf{q}^T (\mathbf{I} - \mathbf{Q})^{-2} \mathbf{Q}\mathbf{1}, \quad \mu_2 = \nu_2 - \nu_1^2 + \nu_1.$$

3 Run rules c chart with estimated parameters

We assume that we have a Phase I data set composed of $i = 1, \ldots, m$ samples $\{X_{i,1}, \ldots, X_{i,n}\}$ of size n. Let us assume that there is independence within and between samples, and $X_i = \sum_{j=1}^{n} X_{i,j} \sim P(c_0)$, i.e. a Poisson distribution of parameter c_0. An estimator $\hat{c}_0$ of c_0 is

$$\hat{c}_0 = \frac{1}{m} \sum_{i=1}^{m} X_i = \frac{X}{m}$$

where $X \sim P(mc_0)$, i.e. a Poisson distribution of parameter mc_0 defined for $x \in \{0, 1, \ldots\}$. When c_0 is estimated by $\hat{c}_0$, the control limits of the $\mathrm{RR}_{2,3} - c$ chart with estimated parameter becomes

$$\widehat{LCL}_c = \left\lceil \hat{c}_0 - K_{c,2,3}\sqrt{\hat{c}_0} \right\rceil, \qquad \widehat{UCL}_c = \left\lfloor \hat{c}_0 + K_{c,2,3}\sqrt{\hat{c}_0} \right\rfloor$$

The p.d.f. and c.d.f. of the run length L are

$$f_L(\ell) = \sum_{x=0}^{\infty} f_P(x|mc_0)(\mathbf{q}^T \hat{\mathbf{Q}}^{\ell-1} \mathbf{r}), \qquad F_L(\ell) = 1 - \sum_{x=0}^{\infty} f_P(x|mc_0)\mathbf{q}^T \hat{\mathbf{Q}}^{\ell} \mathbf{1}$$

where $\hat{\mathbf{Q}}$ is the matrix $\mathbf{Q}$ where p_L, p_U and p_C have been replaced by $\hat{p}_L$, $\hat{p}_U$ and $\hat{p}_C$ and we have

$$
\begin{aligned}
\hat{p}_L &= F_P\left(\left\lceil \frac{x}{m} - K_{c,2,3}\sqrt{\frac{x}{m}}\right\rceil - 1 \middle| c_1\right) \\
\hat{p}_U &= 1 - F_P\left(\left\lfloor \frac{x}{m} + K_{c,2,3}\sqrt{\frac{x}{m}}\right\rfloor \middle| c_1\right) \\
\hat{p}_C &= F_P\left(\left\lfloor \frac{x}{m} + K_{c,2,3}\sqrt{\frac{x}{m}}\right\rfloor \middle| c_1\right) - F_P\left(\left\lceil \frac{x}{m} - K_{c,2,3}\sqrt{\frac{x}{m}}\right\rceil - 1 \middle| c_1\right)
\end{aligned}
$$

The ARL and $SDRL$ of the $\mathrm{RR}_{2,3} - c$ chart with *estimated* parameter are equal to

$$ARL = \sum_{x=0}^{\infty} f_P(x|mc_0)\nu_1, \qquad SDRL = \sqrt{E(L^2) - ARL^2}$$

with

$$E(L^2) = \sum_{x=0}^{\infty} f_P(x|mc_0)(\nu_1 + \nu_2), \quad \nu_1 = \mathbf{q}^T(\mathbf{I} - \hat{\mathbf{Q}})^{-1}\mathbf{1}, \quad \nu_2 = 2\mathbf{q}^T(\mathbf{I} - \hat{\mathbf{Q}})^{-2}\hat{\mathbf{Q}}\mathbf{1}$$

4 Conclusions

The run length performances of the run rules c chart is quite different when the process parameters are know and when the process parameters are estimated, unless the number of sample m is large enough. With the value of m increases, the difference in terms of in-control ARL's, within the known and the estimated parameter case tends to decrese but not always in a monotonic way, both 2-of-3 and 3-of -4 run rules charts. Depending on the values of c_0, the in-control $ARL's$ when $m < \infty$ are either smaller, either largerm than the in-control $ARL's$ corresponding to $m = \infty$, This result is also different from what can be observed in the case of the Run Rules $\bar{X}$ chart where the in-control ARL values corresponding to $m < \infty$ are always larger than the in-control ARL values corresponding to $m = \infty$. The use of alternative chart parameters K' especially dedicated to the estimated parameter case to be as close as possible to in- control ARL value corresponding to the known parameter case. With these specific K' that take the number m of Phase I samples into account also allow to reduce the in-control SDRL of the c chart with estimated parameter.

Références

[1] L. Tian, "Inferences on the common coefficient of variation", Statistics in Medicine. AP, vol. 24, pp. 2213–2220, 2005.

[2] S. Verrill and R.A. Johnson, "Confidence Bounds and Hypothesis Tests for Normal Distribution Coefficients of Variation", Communications in Statistics – Theory and Methods. AP, vol. 36, pp. 2187–2206, 2007.

[3] R. Mahmoudvand and H. Hassani, "Two New Confidence Intervals for the Coefficient of Variation in a Normal Distribution", Journal of Applied Statistics. AP, vol. 36, pp. 429–442, 2009.

[4] E. Del Castillo, "Evaluation of the Run Length Distribution of $\bar{X}$ Charts with Unknown Variance", Journal of Quality Technology. AP, vol. 28, pp. 116–122, 1996.

[5] G. Nedumaran and J.J. Pignatiello, Jr., "On Estimating $\bar{X}$ Control Chart Limits", Journal of Quality Technology. AP, vol. 33, pp. 206–212, 2001.

[6] C.W. Champ and L. Allison-Jones, "Designing Phase I $\bar{X}$ Charts with Small Sample Sizes", Quality and Reliability Engineering International. AP, vol. 20, pp. 497–510, 2004.