CUMULATIVE SUM CONTROL CHART FOR LOG –LOGISTIC DISTRIBUTION

RRL Kantam[1] and G.Srinivasa Rao[2]
1. Department of Statistics, Acharya Nagarjuna University, Guntur-522510, INDIA.
2.Department of Science and Humanities, DVR & Dr. HS MIC College of Technology, Kanchikacherla, Vijayawada, India E-mail: gaddesrao@yahoo.com

ABSTRACT

The sequential probability ratio procedures of Statistical Inference are made use of in construction of a Cumulative Sum Control Chart for a variable process characteristic. The distribution of process variate is log-logistic distribution. The construction of mask and values of average run length are also presented. Such a chart is useful in early detection of shifts in the process average.

1. Introduction

In statistical quality control cumulative sum control charts (CUSUM Charts) have found importance as a parallel process control technique to the well known Shewhart control charts. Infact the two popular quality control methodologies namely (i) control charts (ii) sampling plans are direct applications of the two classical inferential procedures- confidence intervals for control charts, testing of hypothesis for sampling plans. An alternative method of testing statistical hypothesis parallel to Neyman’s theory is the well known sequential probability ratio test (SPRT) due to A.Wald. Johnson (1961) developed a simple theoretical approach to cumulative sum control charts. Johnson and Leone (1962) has made use of simultaneous applications of SPRT to test a simple H_o against two separate simple alternative hypotheses on either side of the null hypothesis. The concerned decision lines for both cases of alternative hypothesis to come to the respective rejections of the null hypothesis are taken to construct a cumulative sum control chart for the normal process variate. Johnson (1966) extended the same procedure to CUSUM chart for Weibull process variate. Edgeman (1989) studied Inverse Gaussian control charts. Nabar and Bilgi (1994) extended CUSUM chart procedure to the case of Inverse Gaussian distribution. Nabar (1999) studied these aspects to control the process dispersion for Inverse Gaussian distribution. We propose to study this for log-logistic distribution.

The probability density function and cumulative distribution function of a log – logistic distribution function as suggested by Balakrishnan *et. al.* (1987) are given by

$$f(x) = \frac{\frac{\beta}{\sigma}\left(\frac{x}{\sigma}\right)^{\beta-1}}{\left[1+\left(\frac{x}{\sigma}\right)^{\beta}\right]^{2}} \; ; \; x \geq 0, \; \beta > 1, \; \sigma > 0 \qquad (1.1)$$

$$F(x) = \frac{\left(\frac{x}{\sigma}\right)^{\beta}}{\left[1+\left(\frac{x}{\sigma}\right)^{\beta}\right]} \; ; \; x \geq 0, \; \beta > 1, \; \sigma > 0 \qquad (1.2)$$

This distribution is a survival model useful in reliability studies. Gupta *et al* (1999) made a study of log- logistic model in survival analysis. Ragab and Green (1984) developed order statistics from the log- logistic distribution and their properties. Kantam *et al* (2001) studied acceptance sampling based on life tests: log- logistic model. Kantam *et al* (2006) developed an economic reliability test plan: log- logistic model. Kantam *et al* (2006) studied variable control charts: log- logistic process variate.

In this paper we consider simultaneous application of SPRT to test $H_0 : \sigma = \sigma_0$ against the two alternatives $H_1 : \sigma = \sigma_1 \, (\sigma_1 > \sigma_0)$ as well as $H_{-1} : \sigma = \sigma_2 \, (\sigma_2 < \sigma_0)$. The construction of the associated CUSUM chart procedure is given in the following section.

2. The CUSUM Chart

CUSUM Chart procedure for detecting shifts in the process mean when the process variate is assumed to follow log-logistic variate is explained below.

We know that the expected value of x in the density (1.1) is given by

$$\mu = E(X) = \sigma \; \Gamma(1+\frac{1}{\beta}) \, \Gamma(1-\frac{1}{\beta})$$

The unknown parameter β may be interpreted in the following way. The scaled observations of a process variate can be consider to have come from a log- logistic population with mean $E(X) = \Gamma(1+ 1/\beta)\Gamma(1- 1/\beta)$ provided the model is a good fit for the data. Accordingly β can be estimated by interesting the equation. Sample mean= $\Gamma(1+ 1/\beta)\Gamma(1- 1/\beta)$ which intern can be use in all other equations involving β and un-scaled observations. Instead of doing so, we have developed our work for a known β, because all the equations involved β. We confine our study when β=2 only (other β values are available with the authors). The process mean is a constant multiplier of population scale parameter σ. Hence a shift in process mean is equivalent to a shift in σ only. Let $\sigma = \sigma_0$ be the target value, let $\sigma = \sigma_1 \, (\sigma_1 > \sigma_0)$ be the changed value. The SPRT procedure will terminate by rejection or acceptance or continue by one more sample observation into sample according as L_1/L_0 is outside or in between two constants A and B. The process terminates by rejecting H_o if (L_1/L_0) >A, this whole procedure gives a rejection line $\sigma_1 > \sigma_0$. Similarly if we adopt SPRT with the same strength to the case of $\sigma_2 < \sigma_0$ we get another rejection line. These two rejection lines will give a geometrical nature of masking. The observations into the sample enter in a sequential way. This is called mask. Here we develop the mask of a CUSUM Chart. For the density given in (1.1) we know that $(L_1/L_0) \geq$ A implies

$$\sum_{i=1}^{m} X_i^2 \geq \frac{\log A - 2m \log(\sigma_0 / \sigma_1)}{2\left[\dfrac{\sigma_1^2 - \sigma_0^2}{\sigma_1^2\,\sigma_0^2}\right]}$$

$$\Rightarrow \sum_{i=1}^{m} X_i^2 \geq C + mD \qquad (2.1)$$

$$\text{where } C = \frac{\sigma_0^2\sigma_1^2 \log A}{2(\sigma_1^2 - \sigma_0^2)} = -\frac{\sigma_0^2\sigma_1^2 \log \alpha}{2(\sigma_1^2 - \sigma_0^2)} \qquad (2.2)$$

$$D = -\frac{\sigma_0^2\sigma_1^2 \log(\sigma_0 / \sigma_1)}{(\sigma_1^2 - \sigma_0^2)} = \frac{\sigma_0^2\sigma_1^2 \log(\sigma_1 / \sigma_0)}{(\sigma_1^2 - \sigma_0^2)} \qquad (2.3)$$

Similarly the rejection line of the procedure when $(\sigma_2 < \sigma_0)$ is given by

$$\sum_{i=1}^{m} X_i^2 \le \frac{\log A - 2m\log(\sigma_0/\sigma_2)}{2\left[\dfrac{\sigma_2^2-\sigma_0^2}{\sigma_2^2\,\sigma_0^2}\right]}$$

$$\Rightarrow \sum_{i=1}^{m} X_i^2 \le C^* + mD^* \qquad (2.4)$$

where $C^* = \dfrac{\sigma_0^2\sigma_2^2 \log A}{2(\sigma_2^2-\sigma_0^2)} = -\dfrac{\sigma_0^2\sigma_2^2 \log\alpha}{2(\sigma_2^2-\sigma_0^2)}$ (2.5)

$$D^* = -\frac{\sigma_0^2\sigma_2^2 \log(\sigma_0/\sigma_2)}{(\sigma_2^2-\sigma_0^2)} = \frac{\sigma_0^2\sigma_2^2 \log(\sigma_2/\sigma_0)}{(\sigma_2^2-\sigma_0^2)} \qquad (2.6)$$

Equations (2.1) and (2.4) are the regions above and below in the plane $\left(m, \sum_{i=1}^{m} x_i^2\right)$. Allowing 'm' sequentially, at some stage we get Σx_i^2 satisfying either equation (2.1) or equation (2.4). Till then the procedure continues. From the slopes, intercepts of these two lines we get the parameters of a CUSUM chart called the angle and the lead distance. These are given by

tan θ_1 = slop of the line P_1Q_1 = D and lead distance OP_1 is d_1

where $d_1 = \dfrac{\log(\alpha)}{2\log(\sigma_1/\sigma 0)}$ when $\sigma_1 > \sigma_0$

tan θ_{-1} = slop of the line $P_{-1}Q_{-1}$ = D^* and lead distance OP_1 is d_{-1}

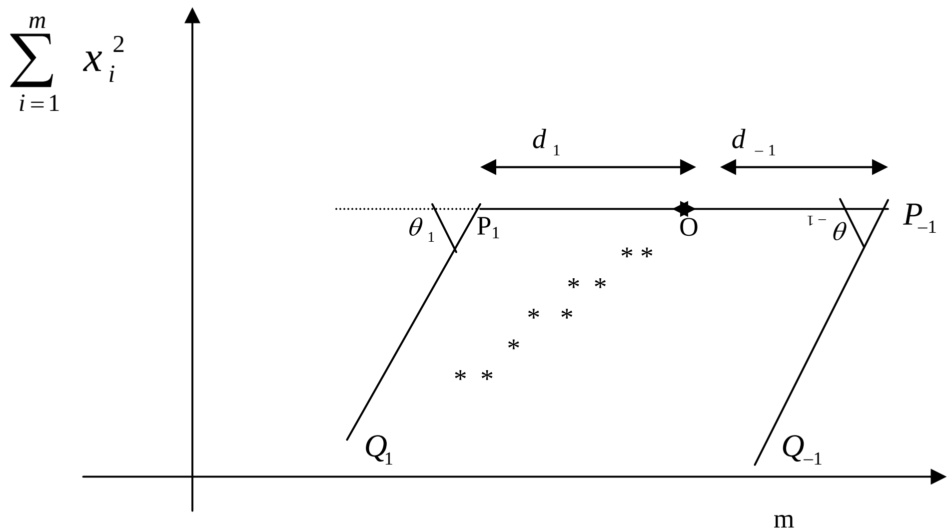
where $d_{-1} = \dfrac{\log(\alpha)}{2\log(\sigma_2/\sigma_0)}$ when $\sigma_2 < \sigma_0$

As a ready reference we have tabulated the values of the parameters angle θ and lead distance "d" for various values of α and choices of σ_1/σ_0 and are given in the Table 1.

Table 1. Parameters of CUSUM Chart

σ_1/σ_0	θ	d		
		α=0.05	α=0.025	α=0.010
0.25	5.28	1.08	1.33	1.66
0.50	13.00	2.16	2.66	3.32
0.60	16.03	2.93	3.61	4.50
0.75	20.29	5.20	6.41	8.00
0.90	24.18	14.21	17.50	21.85
1.10	28.77	15.71	19.35	24.15
1.25	31.79	6.71	8.26	10.31
1.50	36.12	3.69	4.54	5.67
1.60	37.64	3.18	3.92	4.11
1.75	39.72	2.67	3.29	4.11
2.00	42.74	2.16	2.66	3.32
2.25	45.30	1.84	2.27	2.83
2.50	47.48	1.63	2.01	2.51

In the given situation with sample data x_1, x_2,..,x_n from log-logistic distribution we plot the points $\left(m , \sum_{i=1}^{m} x_i^2 \right)$ with suitable scale. They ordinates of these points represent the cumulative sum of squared data. Equation (2.1) and (2.4) are the consequences of a shift in population parameter σ (and hence the shifts in the population average). Therefore the Figure indicates a considerable shift in the population average if Σx_i^2 falls outside the lines represented P_1Q_1 and $P_{-1}Q_{-1}$. The slopes and intercepts of such a lines are given in Table 1. Therefore a separate picture parallel to the concept of mask to be constructed using the table values. A typical structure of the chart is interpreted by place a mask over the chart as shown in Figure with the point O over the last plotted point on the chart with line P_1P_{-1} parallel to the axis of m. If any of the points lies below $P_{-1}Q_{-1}$ then it indicates a decrease in σ and if any of the points lie above P_1Q_1 it indicates an increase in σ.



The graph of the points $\left(m, \sum_{i=1}^{m} x_i^2\right)$ is super imposed with the above mask such that the point O of the mask and each point of CUSUM graph coincide with the horizontal line parallel to x-axis. If at any stage a point of the cumulative graph is covered in the lines of the mask it indicates that at that particular serial number of the sample the process went out of control, Hence assignable causes are to be searched for. In general the average number of trials required to detect a shift in the process average for the first time in called average run length (ARL).

3. Average Run Length

If α is the producer's risk then ARL is given by

$$ARL = \frac{-\log \alpha}{E(\log Z / \sigma = \sigma_1)} \qquad \text{where } Z = \frac{f(x;\sigma_1)}{f(x;\sigma_0)}$$

ARL when Shewhart's control chart is used is given by $ARL_S = 1 + (x_0/\sigma_1)^2$

Where x_0 is solution of the equation $1 - F(x_0) = \alpha$.

For various choices of α the ARL's of Shewarts control chart are shown in the Table 2.

Table 2. Average Run Length for Shewhart control chart ARL_S

σ_1/σ_0	α=0.05	α=0.025	α=0.010
0.25	305.00	625.00	1584.99
0.50	77.00	157.00	396.99
0.60	53.78	109.33	275.99
0.75	34.78	70.33	176.99
0.90	24.45	49.14	123.22
1.10	16.70	33.23	82.82
1.25	13.16	25.96	64.35
1.50	9.44	18.33	45.00
1.60	8.42	16.23	39.67
1.75	7.20	13.73	33.32
2.00	5.75	10.75	25.75
2.25	4.75	8.70	20.55
2.50	4.04	7.24	16.84

On the otherhand the ARL for the CUSUM chart requires numerical integration to find the expectation of Z given from the following equation.

$$E[\log Z/\sigma=\sigma_1]=2\log(\sigma_0/\sigma_1)+2E\left[\log\left\{\frac{1+(x/\sigma_0)^2}{1+(x/\sigma_1)^2}\right\}\right]$$

where $$E\left[\log\left\{\frac{1+(x/\sigma_0)^2}{1+(x/\sigma_1)^2}\right\}\right]=\int_0^\infty \log\left\{\frac{1+(x/\sigma_0)^2}{1+(x/\sigma_1)^2}\right\}\frac{2}{\sigma_1}\frac{(x/\sigma_1)}{[1+(x/\sigma_1)^2]^2}\,dx$$

We have used 10 point Gauss-Laguerre quadrature formula to find the numerical Integration and the values of $E(\log Z/\sigma=\sigma_1)$ are given in Table 3.

Table 3. $E(\log Z/\sigma=\sigma_1)$ **values**

σ_1/σ_0	E[log z]
0.25	-3335966.54
0.50	-6667609.97
0.60	-7073446.35
0.75	-6220547.69
0.90	-3278281.66
1.10	4425079.43
1.25	13365350.76
1.50	34906507.33
1.60	45999456.25
1.75	65433497.47
2.00	105597361.30
2.25	155926569.60
2.50	216850293.80

With these divergent values the ARL for CUSUM chart is almost zero. The practical indication of this is that the CUSUM chart for log-logistic process shifts under population average at the earliest and much faster than by a Shewhart chart. The procedure is developed for other values of β (β=3,4,5) also and it shows the same trend (due to the problem of constraint on the volume of the paper these are not given). Therefore we conclude that CUSUM Chart gives process shifts in the average much earlier than for a Shewhart chart in the case of log-logistic distribution.

References

1. Balakrishnan, N., Malik, H.J.and Puthenpura, S.(1987). Best linear unbiased estimation of location and scale parameters of the log-logistic distribution. Commun. Statist – Theor.Meth., 16(12), 3477 – 3495.
2. Edgeman, R.I. (1989). Inverse Gaussian control chart. Australian Journal of Statistics, 31,78-84.
3. Gupta, R.C., Akman,O. and Lvin, S. (1999). A study of log- logistic model in survival analysis. Biometrical Journal, 41(4), 431-443.
4. Johnson, N.L. (1961). A simple theoretical approach to Cumulative sum control charts. Journal of Amer. Statist. Assoc., 56, 835-840.
5. Johnson,N.L. (1966). Cumulative sum control charts and the Weibull distribution. Technometrics, 8 (3), 481-491.
6. Johnson,N.L and Leone,F.C. (1962). Cumulative sum control charts: Mathematical principles applied to their construction and use. Indust.Qual.Control.,I.vol.18(12),15-21 ; II.vol.19(1),29-36 ; III.vol.19(2),22-28.
7. Kantam, R.R.L., Rosaiah, K. and Srinivasa Rao, G.(2001). Acceptance sampling based on life tests: log-logistic model. Journal of Applied statistics, 28 (1), 121-128.
8. Kantam, R.R.L., Srinivasa Rao, G. and Sriram, B. (2006). An economic reliability test plan: log- logistic distribution. Journal of Applied Statistics, 33(3), 291-296.
9. Kantam, R.R.L., Vasudevarao, A. and Srinivasa Rao, G.(2006). Variable control chart: log- logistic process variate. To be appear in Economic Quality Control Vol. 21.

10. Montgomery, D.C. (2001) Introduction to Statistical Quality Control, Third edition, John Wiley & Sons, New York.
11. Nabar, S.P. (1994). A sequential sampling plan for the scale parameter of the Inverse Gaussian distribution, Journal of Indian Statistical Assoc. 32 (1),117-121.
12. Nabar,S.P. and Bilgi, S. (1994). Cumulative sum control chart for the Inverse Gaussian distribution, Journal of Indian statistical Assoc. 32, 9-14.
13. Nabar,S.P.(1999). Cumulative sum control chart for the scale parameter for inverse gaussian distribution, statistical inference and design of experiments, Narosa publishing House, 55-63.
14. Ragab, A. and Green, J. (1984). On order statistics from the log- logistic distribution and their properties, Commun. Statist – Theor.Methods, 3(21), 2713-2724.