CUMULATIVE SUM CONTROL CHARTS FOR BINOMIAL PARAMETERS WHEN THE UNDERLYING DISTRIBUTION IS POISSON

Ashit B. Chakraborty and Anwer Khurshid[2]
Department of Mathematical and Physical Sciences, College of Arts and Science, University of Nizwa, P. O. Box 33, PC 616, Birkat Al Mouz, Oman

ABSTRACT
Cumulative sum control charts (CUSUM) are constructed for controlling the parameters of the binomial distribution when the underlying distribution is Poisson. It is observed that the parameters of the V-mask and the Average Run Length (ARL) change considerably for a slight shift in the parameters (or ratio) of the distribution understudy.

KEYWORDS: Sequential Probability Ratio Test; Cumulative Sum control chart; Average Run Length; binomial distribution; Poisson distribution

RESUMEN
Cartas de control se la suma acumulativa (CUSUM) son construidas para controlar los parámetros una distribución binomial cuando la subyacente es una Poisson. Se observa que los parámetros de la V-máscara y del Largo Promedio de la Rachas (ARL) cambian considerablemente para un ligero cambio en el ajuste en los parámetros (o razones) de la distribución bajo estudio.

MSC: 62F03

1. INTRODUCTION

A common problem in the application of statistical methods to the quality of material produce by a continuous process is that of ensuring that the proportion of defective product does not exceed a specified limit. The CUSUM technique proposed by Page (1954, 1961) is a valuable tool to study such problems, as this is much more sensitive than the standard Shewhart's schemes especially to small deviations from the target (see also Montgomery, 2005). Johnson (1961) showed that the CUSUM control charts can be interpreted as a modified form of a pair of Sequential Probability Ratio Test (SPRT) treated simultaneously and gave mathematical procedures for constructing the CUSUM charts. Johnson and Leone (1962) constructed CUSUM charts for controlling the means of the binomial and Poisson distributions.

Probability distributions often arise in quality control, life-testing, medical and demographic studies when the control charts for ratio of two Poisson means (Sahai and Khurshid, 1993) need to be constructed, as the situation may arise to control the ratio rather than the single parameter. In such situation binomial distribution being derived based on the ratio of two Poisson can be used to develop the CUSUM chart.

The purpose of this paper is to construct one-sided CUSUM control chart for controlling the different parameters of the distribution under study and to obtain ARL for detecting the shift of the process parameters (or the ratio). The method of Johnson (1961) is used while constructing the CUSUM chart.

2. CUSUM CHART

1

[2] Corresponding author, e-mail: anwer_khurshid@yahoo.com

Let X and Y be two independent Poisson variates with parameters λ and μ respectively. Then the conditional distribution of X given $X+Y$ follows a binomial distribution (Lehmann and Romano, 2005). Thus,

$$
\begin{aligned}
P\left[X = x \middle| (X+Y=n)\right] &= P(X = x \cap Y = n-x)\left[P(X+Y=n)\right]^{-1} \\
&= \binom{n}{x}\left(\frac{\lambda}{\lambda+\mu}\right)^{x}\left(\frac{\mu}{\lambda+\mu}\right)^{n-x}
\end{aligned}
\tag{2.1a}
$$

where $x = 0,1,2,\ldots,n$. The mean and variance of X are given by

$$E(X) = \frac{n\lambda}{\lambda+\mu},\ Var(X) = \frac{n\mu\lambda}{(\lambda+\mu)^2}.$$

Charts are constructed separately for controlling the parameters λ and μ of (2.1a). The ARLs also calculated for a number of combinations of the values of the parameters.

2.1. Control of the parameter λ when μ is known

Let $x_1, x_2, \ldots, x_m$ be i.i.d. random variables each distributed with probability mass function (1.1). To test the null hypothesis $H_0 : \lambda = \lambda_0$ against the alternative hypothesis $H_1 : \lambda = \lambda_1 (> \lambda_0)$ assuming μ known, we use the likelihood ratio

$$\frac{f(x_1, x_2, \ldots, x_m | \lambda_1, \mu)}{f(x_1, x_2, \ldots, x_m | \lambda_0, \mu)} = \left(\frac{\lambda_1}{\lambda_0}\right)^{\sum_{i=1}^{m} x_i}\left[\frac{\lambda_0+\mu}{\lambda_1+\mu}\right]^{mn}. \tag{2.1b}$$

The continuation region of the SPRT discriminating between the two hypotheses $H_0 : \lambda = \lambda_0$ and $H_1 : \lambda = \lambda_1 (> \lambda_0)$ has the continuation region

$$\log\left(\frac{\beta}{1-\alpha}\right) < \sum_{i=1}^{m} x_i \log\left(\frac{\lambda_1}{\lambda_0}\right) + mn\log\left(\frac{\lambda_0+\mu}{\lambda_1+\mu}\right) < \log\left(\frac{1-\beta}{\alpha}\right) \tag{2.2}$$

where α is the probability of accepting H_1 when H_0 is true and β is the probability of accepting H_0 when H_1 is true.

Considering the right hand side inequality in (2.2), we get

$$\sum_{i=1}^{m} x_i \log\left(\frac{\lambda_1}{\lambda_0}\right) + mn\log\left(\frac{\lambda_0+\mu}{\lambda_1+\mu}\right) < \log\left(\frac{1-\beta}{\alpha}\right).$$

For a very small value of β, we have

$$\sum_{i=1}^{m} x_i < \frac{-\log\alpha + mn\log\left(\frac{\lambda_1+\mu}{\lambda_0+\mu}\right)}{\log\left(\frac{\lambda_1}{\lambda_0}\right)}.$$

For constructing the CUSUM chart (as shown in Figure 2.1), we plot the sum $S_m = \sum_{i=1}^{m} x_i$ against the number of observations m. Suppose here O is the last plotted point, P is the vertex of the mask and the point Q is obtained by drawing a perpendicular to the line OP. The change in the value of λ from λ_0 to λ_1 is detected if any plotted point falls below the line PQ.

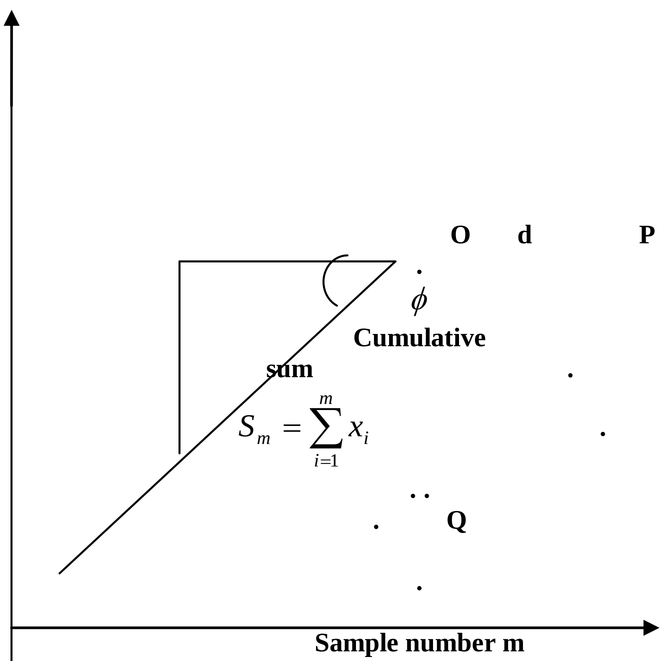

Figure 2.1 Cumulative sum control chart

In this case the parameters of the mask namely the lead distance d=OP and the angle of the mask $\phi = \angle OPQ$ are given by

$$d = -\log\alpha\left[n\log\left(\frac{\lambda_1+\mu}{\lambda_0+\mu}\right)\right]^{-1} \tag{2.3}$$

and

$$\phi = \tan^{-1}\left[\frac{n\log\left(\frac{\lambda_1+\mu}{\lambda_0+\mu}\right)}{\log\left(\frac{\lambda_1}{\lambda_0}\right)}\right]. \tag{2.4}$$

Table 2.1: **Values of** d **for controlling the parameter** λ

(a) When $\mu = 0.5$ **and** $n = 24$						
		α				
λ_0	λ_1	0.05	0.025	0.01	0.005	0.001
0.4	0.43	3.808	4.688	5.853	6.734	8.780
0.4	0.46	1.934	2.383	2.975	3.423	4.463
0.4	0.49	1.310	1.612	2.012	2.315	3.019
0.4	0.52	0.997	1.228	1.533	1.764	2.300
0.4	0.55	0.887	0.996	1.243	1.430	1.865
(b) When $\mu = 0.6$ **and** $n = 24$						
		α				
λ_0	λ_1	0.05	0.025	0.01	0.005	0.001
0.4	0.43	4.223	5.200	6.491	7.468	9.738
0.4	0.46	2.142	2.368	3.293	3.788	4.939
0.4	0.49	1.447	1.782	2.224	2.559	3.337
0.4	0.52	1.102	1.358	1.695	1.950	2.542
0.4	0.55	0.777	0.955	1.193	1.372	1.789
(c) When $\mu = 0.6$ **and** $n = 20$						
		α				
λ_0	λ_1	0.05	0.025	0.01	0.005	0.001
0.4	0.43	5.068	6.234	7.782	8.954	11.67
0.4	0.46	3.491	4.316	5.388	6.199	8.08
0.4	0.49	1.738	2.140	2.671	3.073	4.01
0.4	0.52	1.322	1.627	2.031	2.336	3.05
0.4	0.55	1.072	1.321	1.649	1.897	2.47

2.2. Average Run Length to control the parameter λ when μ is known

Following Johnson (1961), the approximate formula for ARL detecting a shift for the parameter λ from λ_0 to λ_1 is given by $ARL = (-\log\alpha)E_{\lambda_1}^{-1}$ where $E_{\lambda_1} = E\left[\log\frac{f(x|\lambda_1)}{f(x|\lambda_0)}|\lambda_1\right]$.

Thus, we get

$$ARL = (-\log\alpha)\left[\frac{n\lambda_1}{\lambda_1+\mu}\log\left(\frac{\lambda_1}{\lambda_0}\right) + n\log\left(\frac{\lambda_0+\mu}{\lambda_1+\mu}\right)\right]^{-1}. \tag{2.5}$$

Table 2.2: Values of ϕ for controlling the parameter λ

(a) When $\mu = 0.5$ **and** $n = 24$					
	λ_1				
λ_0	0.43	0.46	0.49	0.52	0.55
0.3	84.31	84.42	84.52	84.61	84.69
0.4	84.75	84.84	84.93	85.01	85.08
(b) When $\mu = 0.6$ **and** $n = 24$					
	λ_1				
λ_0	0.43	0.46	0.49	0.52	0.55
0.3	83.66	83.79	83.91	84.02	84.12
0.4	84.18	84.29	84.39	84.49	84.58
(c) When $\mu = 0.6$ **and** $n = 20$					
	λ_1				
λ_0	0.43	0.46	0.49	0.52	0.55
0.3	82.40	82.56	82.70	82.83	82.95
0.4	83.03	83.16	83.28	83.40	83.50

To find the characteristic of ARL, we differentiate (2.5) with respect to λ_1 as

$$\frac{\partial}{\partial \lambda_1} ARL = (\log \alpha)\left[\frac{n\lambda_1}{\lambda_1+\mu}\left(\frac{\lambda_0}{\lambda_1}\right) + n\log\left(\frac{\lambda_1}{\lambda_0}\right)\frac{\mu}{(\lambda_1+\mu)^2} + \frac{n}{\lambda_1+\mu}\right] \div \left[\frac{n\lambda_1}{\lambda_1+\mu}\log\left(\frac{\lambda_1}{\lambda_0}\right) + n\log\left(\frac{\lambda_0+\mu}{\lambda_1+\mu}\right)\right]^2 \tag{2.6}$$

This will be negative provided the numerator is less than zero, implying that the ARL decreases as λ_1 increases from known (constant) value of μ. Tables (2.1), (2.2) and (2.3) depict some numerical values of d, ϕ and ARL for a number of combinations of the values of α , λ_0 , λ_1 and μ for specific n.

3. CONTROL OF THE PARAMETER μ WHEN λ IS KNOWN

Let $x_1, x_2, \ldots, x_m$ be i.i.d. random variables each distributed with probability mass function (1.1). To test the null hypothesis $H_0 : \mu = \mu_0$ against the alternative hypothesis $H_1 : \mu = \mu_1$ assuming λ known, we use the likelihood ratio

$$\frac{f(x_1, x_2, \ldots, x_m | \mu_1, \lambda)}{f(x_1, x_2, \ldots, x_m | \mu_0, \lambda)} = \left(\frac{\mu_1}{\mu_0}\right)^{mn - \sum_{i=1}^{m} x_i} \left[\frac{\lambda+\mu_0}{\lambda+\mu_1}\right]^{mn}. \tag{3.1}$$

The change in the value of μ will be detected if the inequality

$$\sum_{i=1}^{m} x_i > \frac{-\log\alpha + mn\left[\log\left(\frac{\mu_0}{\mu_1}\right) + \log\left(\frac{\lambda+\mu_1}{\lambda+\mu_0}\right)\right]}{\log\left(\frac{\mu_0}{\mu_1}\right)}$$

is hold .

Table 2.3: Values of ARL for controlling the parameter λ

(a) When $\mu = 0.5$ **and** $n = 24$						
		α				
λ_0	λ_1	0.05	0.025	0.01	0.005	0.001
0.4	0.43	192.42	236.92	295.75	340.26	443.66
0.4	0.46	51.36	63.24	78.94	90.82	118.41
0.4	0.49	24.31	29.93	37.36	42.99	56.05
0.4	0.52	14.53	17.89	22.33	25.70	33.50
0.4	0.55	9.86	12.14	15.16	17.44	22.74
(b) When $\mu = 0.6$ **and** $n = 24$						
		α				
λ_0	λ_1	0.05	0.025	0.01	0.005	0.001
0.4	0.43	197.11	242.70	302.97	348.56	434.48
0.4	0.46	52.40	64.52	80.54.94	92.65	120.81
0.4	0.49	24.71	30.42	37.98	43.69	56.97
0.4	0.52	14.72	18.12	22.62	26.02	33.93
0.4	0.55	9.95	12.26	15.30	17.60	22.95
(c) When $\mu = 0.6$ **and** $n = 20$						
		α				
λ_0	λ_1	0.05	0.025	0.01	0.005	0.001
0.4	0.43	236.53	291.24	363.56	418.27	545.38
0.4	0.46	62.88	77.42	96.64	111.19	144.97
0.4	0.49	29.65	36.51	45.57	52.43	68.36
0.4	0.52	17.66	21.74	27.14	31.23	40.72
0.4	0.55	11.94	14.71	18.36	21.12	27.54

To construct the CUSUM chart in this case also we plot the sum $S_m = \sum_{i=1}^{m} x_i$ against the number of observations m. The change of μ from μ_0 to μ_1 is detected if any plotted points fall below the line PQ as shown in Figure 2.1. The parameters of the mask (the lead distance d and the angle ϕ) are given by

$$d = -\log\alpha\left[n\log\left(\frac{\mu_0}{\mu_1}\right) + \log\left(\frac{\lambda+\mu_1}{\lambda+\mu_0}\right)\right]^{-1} \tag{3.2}$$

and

$$\phi = \tan^{-1}\left[\frac{n\log\left(\frac{\lambda+\mu_1}{\lambda+\mu_0}\right) - n\log\left(\frac{\mu_1}{\mu_0}\right)}{\log\left(\frac{\mu_0}{\mu_1}\right)}\right] \tag{3.3}$$

3.1. Average Run Length to control the parameter μ when λ is known

The average run length for detecting a shift in μ from μ_0 to μ_1, following Johnson (1961) is approximately given by

$$\begin{aligned} ARL &= (-\log\alpha)E_\mu^{-1} \\ &= (-\log\alpha)\left[n\left(\frac{\mu_1}{\lambda+\mu_1}\right)\log\left(\frac{\mu_1}{\mu_0}\right) + \log\left(\frac{\lambda+\mu_0}{\lambda+\mu_1}\right)\right]^{-1} \end{aligned} \tag{3.4}$$

4. AN ILLUSTRATIVE EXAMPLE AND CONCLUSION

In this paper we have formulated the different criteria of CUSUM scheme like distance d and angle ϕ of the CUSUM chart along with the derivation of average run length. Computation of the distance d and angle ϕ of the V-mask of one-sided CUSUM chart for different combinations of the values of λ, μ, n and α for controlling the parameters λ (when μ is known) and μ (when λ is known) are shown in Tables 2.1, 2.2, 3.1 and 3.2.

An illustrative example of counts of diatoms (Sahai and Khurshid 1993) may be considered in such a situation where it is required to control the diatom concentration of two lake water and to detect the change in the ratio (λ / μ) of the concentration.

It has been observed from the Table 2.1 for all the cases (a, b, and c) for fixed α, the values of d decreases as the difference $(\lambda_1 - \lambda_0)$ increases, whereas for the same difference $(\lambda_1 - \lambda_0)$, the values of d increases as α decreases. It can also be interpreted as the ratio (λ/μ) increases, the lead distance d decreases. It is also observed from the Table 2.1 (b and c) that the distance d increases for fixed ratio (λ/μ) and α if n is decreased.

It has also been observed from Tables 2.2 and 3.1 that the angle of the mask increases as the ratio (λ_1/μ_0) increases and for fixed ratio (λ_1/μ_0), angle of the mask decreases as n is decreased. It can also be said that the angle of the mask increases as the ratio (λ/μ) increased (or decreases as the ratio (λ/μ) increases).

Table 2.3 shows the values of average run length for different combinations of α, μ, n and λ. Here it is interesting to note that the ARL values obtained for controlling the parameter λ when μ is known are same as that of controlling the parameter μ when λ is known.

It is evident from Table 2.3 (a, b and c) that for fixed α, the ARL decreases as the shift from λ_0 to λ_1 increases (or ratio λ/μ increases) and for fixed changed from λ_0 to λ_1, the ARL increases as the initial region α decreases. But for fixed change (or ratio λ/μ) and α, the ARL increases if the parameter n decreases.

ACKNOWLEDGEMENT: The authors wish to thanks the referees for their valuable comments and suggestions.

RECEIVED JANUARY 2010
REVISED JUNE 2010)

REFERENCES

[1] JOHNSON, N. L. (1961): A simple theoretical approach to cumulative sum chart. **Journal of the American Statistical Association**, 56, 835-840.

[2] JOHNSON, N. L. and LEONE, F. C. (1962): Cumulative sum control charts: Mathematical principles applied to their construction and use, Part II. **Industrial Quality Control**, xix, 22-28.

[3] LEHMANN, E. L. and ROMANO, J. P. (2005): **Testing Statistical Hypotheses**, Third Edition. Springer, New York.

[4] MONTGOMERY, D. C. (2005): **Introduction to Statistical Quality Control**, Fifth Edition, John Wiley, New York.

[5] PAGE, E. S. (1954): Continuous inspection schemes. **Biometrika**, 41, 100-115.

[6] PAGE, E. S. (1961): Cumulative sum charts. **Technometrics**, 3, 1-9.

[7] SAHAI, H. and KHURSHID, A. (1993): Confidence intervals for ratio of two Poisson means. **The Mathematical Scientist**, 18, 43-50.