ISSN 2229-5518

# The Distribution of the Average run Length (ARL) of the Cusum Control Charts for Binomial Parameters when Observation are Poisson Distributed

Edokpa I.W., Osabuohien-Irabor Osa., Ogbeide E. Michael

**Abstract** — In this study, the average run lengths (ARL) of the Cumulative Sum (CUSUM) charts for the binomial distribution when the underlying distribution is Poisson were obtained. It is observed that the parameters of the ARL changes considerably for a slight changes in the parameters o the underlying distribution.

**Index Terms** — Cumulative Sum chart, average run length, binomial distribution, Poisson distribution, distribution of ARL

## 1.0 INTRODUCTION

Cumulative sum (CUSUM) procedures are often used to monitor the quality of manufacturing processes. The major objective is to identify persistent causes of variation in the process average. This ability is attributed to the fact that they have memory as they are based on successive sums of the observations minus a constant. Generally, we can say that CUSUM charts are able to detect small to moderate shifts whereas Shewhart charts are able to detect large shifts. A major problem in the applications of the Statistical Process Control (SPC) to the quality of materials produced by its process is that of making sure that the proportion of defective produced does not exceed the specified limit. A valuable tool of achieving this goal was proposed by Page [1]. The principal feature of the CUSUM control chart is that it cumulates the difference between the observed value $X_i$ and the pre-determined target or reference value $k$, the CUSUM $S_i$ of the deviation of $X_i$ from $k$ is given in equation 1.1

$$S_i = \sum_{i=1}^{n}(X_i - k) \qquad 1.10$$


- ***Edokpa I.W.*** *is a senior Lecturer with a doctorate degree in statistics, Specializing in Quality Control from Ambrose Alli University, Ekpoma, Nigeria. PH: +2348035787519. E-mail: waziripl@gmail.com*
- ***Osabuohien-Irabor Osa.*** *holds a Masters degree in Applied Statistics from Nnamdi Azikiwe University, Nigeria, and currently pursuing a doctorate degree in Statistics with specialty in Time Series Analysis. He is a Lecturer at Ambrose Alli University, Ekpoma, Nigeria. E-mail: osabuohien247@gmail.com, PH: +2348034861146*
- ***Ogbeide E.M.*** *is a Lecturer at Ambrose Alli University, Ekpoma, Nigeria, and currently pursuing a doctorate degree program in Industrial Mathematics with an option in Statistics from the University of Benin, Nigeria. Email: ogbeideoutreach@yahoo.com, PH: +2348035910189*


If $S_i$ is plotted against the sample on a chart or recorded in tabulation, a change in the direction of the CUSUM path is usually taken to indicate a change in process level. For a one-sided CUSUM procedures in which the continuous or discrete random variables $X_1, X_2, X_3, \ldots$ given in equation (1.1) are taken successively and the cumulative sums

$$S_i = \max(0, S_{i-1} + X_i), i = 1,2,3, \ldots \qquad 1.2$$

are formed, where $S_0 = w\ (w \geq 0)$. The process is considered to being in - control until the first stage, N, such that $S_N \geq h\ (0 \leq w \leq h)$

The random variable N, referred to as the run length of the procedure and it is defined as the stage at which sampling terminates and necessary action is taken. We have to state here that in the case of standard normal data with $k = 3$ and $h = 3$ we end up with the classical Shewhart chart.

The applications of CUSUM charts have received considerable attentions. The recognition of a CUSUM control scheme as a sequence of Wald Sequential Probability Ratio Test (SPRT) allows the optimality properties of CUSUM procedures to be developed.

Johnson and Leone [2] gave a discussion on the CUSUM procedures using the relationship between SPRT'S and CUSUM. They also constructed the CUSUM charts for controlling the means of a Binomial and Poisson distribution. Lorden [3] gave the asymptotic optimality of CUSUM procedures for detecting a change in distribution.

Kennett and Pollack [4] showed then superiority of a CUSUM scheme for detecting a rare event over a non – CUSUM scheme proposed by Chen [5]. Kenett and Pollack [4] scheme can be improved upon by including the First Initial Response (FIR) feature, see Lucas and crosier[6].

In this paper, our aim is to develop a simple procedure of obtaining the ARL of a CUSUM control chart for Binomial parameters when the observations are Poisson distributed using a method similar to [2] and compare the changes in the distribution of the ARL with the changes in the values of the CUSUM parameters as the sample sizes varies graphically.

## 2.0 The Average Run Length (ARL) of the CUSUM Control Chart

Cusum control schemes are usually evaluated by calculating their average run length (ARL), the ARL is defined as the average number of samples taken before an out- of- control signal is obtained. The ARL should be large when the process is at the desired level and small when the process shifts to an undesirable level. The in-control ARL's are often approximately closely by the geometric distribution [7]. For a standard CUSUM, the ARL distribution is nearly geometric except that there is a lower probability of extremely short run lengths. When the FIR feature is used, the distribution is nearly geometric except that there is an increased probability of short run lengths due to the head start [7]. There are two majors methods for the computation of the ARL, there are the integral and Markov chain approach. Page[1] used integral equations for the computation of the ARL. Let $F(x)$ be the distribution function of a single score $x \sim N(m,1)$,where N(m,1) denotes a normal distribution with mean m and variance 1.0 and $L(z)$ be the ARL of the one sided case, if $L(0)$ is the ARL with an initial value of zero. Then, for $0 \leq z < h$

$$L(z) = 1 + L(0)F(-z) + \int_0^h f(x-z)L(x)dx \qquad 2.10$$

Van Dobben and de Bruyn [8] gave a discussion on the derivation of this equation. Additionally, Wetherill [9] gave an almost identical relationship but from a somewhat different way of thinking. Others that have dealt with the same problem are Ewan and Kemp [10] and recently Champ,Rigdon and Scharngi [11] gave a general method for obtaining integral equations used in the evaluation of many control charts.

Brook and Evans [7] were the first to propose the new method for computing the ARL based on a Markov chain. This method applies to both discrete and continuous variables.

The Markov Chain approach begins by approximating the problem of obtaining the average run length (ARL) and then obtains an exact solution to the approximate problem. Champ [12] compared integral and Markov chain approaches. They propose the integral equation approaches are more accurate than the Markov-chain approach, but it less versatile. The Markov-chain can calculate both the ARL and the distribution of the run length. It has also been used to calculate the properties of the modified CUSUM schemes such as the robust CUSUM schemes [6].

Assume that $X_1, X_2, X_3, \ldots$ are independent and identically distributed random variables that are observed sequentially. Let $X_1, X_2, X_3, \ldots, X_n$ have (in-control) distributionfunction $F_0$ and $X_{n+1}, X_{n+2}, X_{n+3}, \ldots, X_k$ have (out-of-control) distribution function $F_1$ where $F_0 \neq F_1$ The two distributions are known but the time of change is assumed unknown.

Many schemes can detect such a change (e.g. Shewhart charts). These schemes are categorised by the expected time until the process signals while it remains in-control (false alarm rate). Among all procedures with the same false alarm rates, the optimal procedure is the one that detects changes quicker. Or we could say that among all procedures with the same in-control expected number of samples until signal, the optimal procedure has the smallest expected time until it signals a change when the process shifts to the out-of-control state.

## 2.1 Probability of the CUSUM Chart

Let $X_1$ and $X_2$ be two independent Poisson variables with parameters $\lambda_1$ and $\lambda_2$ respectively. Then the conditional distribution of $X_1$ given $X_1 + X_2$ is given as

$$P(X_1 = r \mathrm{I}(X_1 + X_2 = n)) = \frac{P(X_1 = r \cap X_2 = n-r)}{P(X_1 + X_2 = n)} = \binom{n}{r}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^r \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^{n-r}. \qquad 2.11$$

Where $r = 0,1,2,\ldots,n$, then the mean of and variance of $X_1$ are given by

$E(X_1) = np$ and variance $Var(X_1) = npq$ where $p = \frac{\lambda_1}{\lambda_1+\lambda_2}$ and $q = \frac{\lambda_2}{\lambda_1+\lambda_2}$

## 2.2 Controlling the parameter $\lambda_1$ where $\lambda_2$ is unknown.

Let $X_1, X_2, \ldots, X_k$ be identical, independent distributed Binomial random variable using the likelihood ratio test to test the hypothesis $H_0: \lambda_1 = \lambda_0$ against the alternative hypothesis$H_a: \lambda_1 = \lambda_a (> \boldsymbol{\lambda_0})$, where $\boldsymbol{\lambda_2}$ is assumed known.

$$\frac{L(\lambda_a)}{L(\lambda_0)} = \frac{f(x_1,\ldots,x_k I \lambda_a,\lambda_2)}{f(x_1,\ldots,x_k I \lambda_0,\lambda_2)} = \left[\frac{\lambda_a}{\lambda_0}\right]^{S_k}\left[\frac{\lambda_0+\lambda_2}{\lambda_a+\lambda_2}\right]^{kn}, \qquad 2.21$$

where $k = \sum_{i=1}^{k} X_i$

The region of the SPRT discriminates between $H_0: \lambda_1 = \lambda_0$ aganst $H_a: \lambda_1 = \lambda_a (> \lambda_0)$ has the continuation region of

$$\ln\left(\frac{\beta}{1-\alpha}\right) < S_k \ln\left(\frac{\lambda_a}{\lambda_0}\right) + kn\ln\left(\frac{\lambda_0+\lambda_2}{\lambda_a+\lambda_2}\right) < \ln\left(\frac{1-\beta}{\alpha}\right) \qquad 2.22$$

Where $\alpha$ =probability of accepting $H_a$ when $H_0$ is true) and $\beta$ =probability (accepting $H_0$ when $H_a$ is true). From 2.22 inequalities, we have

$$S_k \ln\left(\frac{\lambda_a}{\lambda_0}\right) + kn\ln\left(\frac{\boldsymbol{\lambda_0}+\lambda_2}{\lambda_a+\lambda_2}\right) < \ln\left(\frac{1-\beta}{\alpha}\right) \qquad 2.23$$

If $\beta$ tends to zero, then equation 2.23 can be written as

$$S_K < \frac{\ln\left(\frac{1}{\alpha}\right) + kn\ln\left(\frac{\lambda_a+\lambda_2}{\lambda_0+\lambda_2}\right)}{\ln\left(\frac{\lambda_a}{\lambda_1}\right)} \qquad 2.24$$

The inequality 2.24 can be re-written as

$$S_K < ln\left[\left(\frac{1}{\alpha}\right)\left(\frac{\lambda_a+\lambda_2}{\lambda_0+\lambda_2}\right)^{kn}\right]\left[\ln\left(\frac{\lambda_a}{\lambda_1}\right)\right]^{-1} \qquad 2.25$$

On plotting the sum $S_k$ against the sample number $k$, let A be the last plotted point on the CUSUM chart, and B be the vertex of the mask if C is obtained be drawing a perpendicular line to line AB as shown in Fig 1,

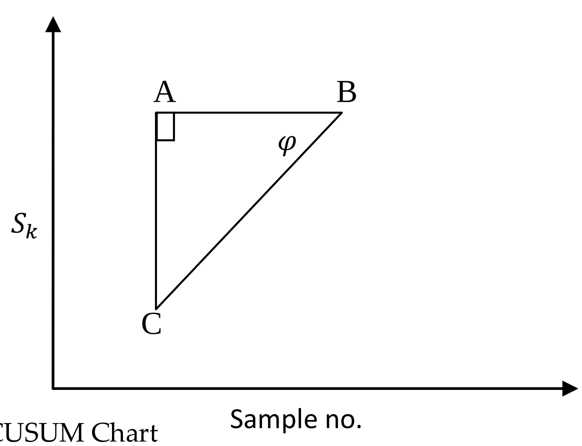


**Fig 1:** CUSUM Chart

The change in the direction of the mean from $\lambda_0$ *to* $\lambda$ is observed, if a point is plotted below line BC. The distance between points AB is denoted by

$$\sigma = \frac{\ln\left(\frac{1}{\alpha}\right)}{n\ln\left[\frac{\lambda_a+\lambda_2}{\lambda_0+\lambda_2}\right]} \qquad 2.26$$

and $\varphi$ which is equal or less than angle $ABC$ is given by

$$\varphi = \tan^{-1}\left[\frac{n\ln\left(\frac{\lambda_a+\lambda_2}{\lambda_0+\lambda_2}\right)}{\ln\left(\frac{\lambda_a}{\lambda_0}\right)}\right] \qquad 2.27$$

$$\sigma = \frac{2}{g^2}\ln\left(\frac{1}{\alpha}\right) \quad \text{where } g^2 = 2n\ln\left(\frac{\lambda_a+\lambda_2}{\lambda_0+\lambda_2}\right)$$

Varying the parameter $\lambda_a$ when $\lambda_2$ is a constant. From Johnson[2], the approximate formula for ARL to detect a shift of $\lambda$ from $\lambda_0$ to $\lambda_a$ is given to be,

$$ARL = \ln\frac{\left(\frac{1}{\alpha}\right)}{A}, \quad \text{where } A = E\left[\ln\frac{f(xI\lambda_a)}{f(xI\lambda_0)} IA\right]$$

The numerical values when $\lambda_2$ is constant (at 0.6), see table 1 for, n=5,10,15,…,60 when $\alpha$ varies between 0.01 and 0.1.The numerical values for varying $\lambda_2$ were obtained by the authors and can be given on request.

ISSN 2229-5518

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | $\alpha$ | | | | |
| | $\lambda_0$ | $\lambda_a$ | **0.01** | **0.025** | **0.05** | **0.075** | **0.1** |
| **n = 05** | 0.5 | 0.52 | 190.9418 | 152.9501 | 124.2105 | 107.3989 | 95.47089 |
| | 0.5 | 0.54 | 78.99678 | 63.27879 | 51.38859 | 44.43327 | 39.49839 |
| | 0.5 | 0.56 | 44.90989 | 35.97417 | 29.21456 | 25.26044 | 22.45495 |
| | 0.5 | 0.58 | 29.35623 | 23.51522 | 19.09667 | 16.51198 | 14.67812 |
| | 0.5 | 0.60 | 20.80956 | 16.66908 | 13.53693 | 11.70474 | 10.40478 |
| **n = 10** | 0.5 | 0.52 | 95.47089 | 76.47505 | 62.10525 | 53.69945 | 47.73545 |
| | 0.5 | 0.54 | 39.49839 | 31.63939 | 25.69429 | 22.21663 | 19.74919 |
| | 0.5 | 0.56 | 22.45495 | 17.98709 | 14.60728 | 12.63022 | 11.22747 |
| | 0.5 | 0.58 | 14.67812 | 11.75761 | 9.548335 | 8.255991 | 7.339059 |
| | 0.5 | 0.60 | 10.40478 | 8.33454 | 6.768465 | 5.85237 | 5.20239 |
| **n = 15** | 0.5 | 0.52 | 63.64726 | 50.98337 | 41.4035 | 35.79964 | 31.82363 |
| | 0.5 | 0.54 | 26.33226 | 21.09293 | 17.12953 | 14.81109 | 13.16613 |
| | 0.5 | 0.56 | 14.96996 | 11.99139 | 9.738186 | 8.420146 | 7.484982 |
| | 0.5 | 0.58 | 9.785412 | 7.838408 | 6.365557 | 5.503994 | 4.892706 |
| | 0.5 | 0.60 | 6.936519 | 5.55636 | 4.51231 | 3.90158 | 3.46826 |
| **n = 20** | 0.5 | 0.52 | 47.73545 | 38.23753 | 31.05262 | 26.84973 | 23.86772 |
| | 0.5 | 0.54 | 19.74919 | 15.8197 | 12.84715 | 11.10832 | 9.874597 |
| | 0.5 | 0.56 | 11.22747 | 8.993543 | 7.30364 | 6.31511 | 5.613737 |
| | 0.5 | 0.58 | 7.339059 | 5.878806 | 4.774168 | 4.127996 | 3.669529 |
| | 0.5 | 0.60 | 5.20239 | 4.16727 | 3.384232 | 2.926185 | 2.601195 |
| **n = 25** | 0.5 | 0.52 | 38.18836 | 30.59002 | 24.8421 | 21.47978 | 19.09418 |
| | 0.5 | 0.54 | 15.79936 | 12.65576 | 10.27772 | 8.886653 | 7.899678 |
| | 0.5 | 0.56 | 8.981979 | 7.194834 | 5.842912 | 5.052088 | 4.490989 |
| | 0.5 | 0.58 | 5.871247 | 4.703045 | 3.819334 | 3.302397 | 2.935623 |
| | 0.5 | 0.60 | 4.161912 | 3.333816 | 2.707386 | 2.340948 | 2.080956 |
| **n = 30** | 0.5 | 0.52 | 31.82363 | 25.49168 | 20.70175 | 17.89982 | 15.91182 |
| | 0.5 | 0.54 | 13.16613 | 10.54646 | 8.564765 | 7.405545 | 6.583065 |
| | 0.5 | 0.56 | 7.484982 | 5.995695 | 4.869093 | 4.210073 | 3.742491 |
| | 0.5 | 0.58 | 4.892706 | 3.919204 | 3.182778 | 2.751997 | 2.446353 |
| | 0.5 | 0.60 | 3.46826 | 2.77818 | 2.256155 | 1.95079 | 1.73413 |
| **n = 35** | 0.5 | 0.52 | 27.2774 | 21.85001 | 17.74436 | 15.3427 | 13.6387 |
| | 0.5 | 0.54 | 11.28525 | 9.039827 | 7.341227 | 6.34761 | 5.642627 |
| | 0.5 | 0.56 | 6.415699 | 5.139167 | 4.173508 | 3.608634 | 3.20785 |
| | 0.5 | 0.58 | 4.193748 | 3.359318 | 2.728096 | 2.358855 | 2.096874 |
| | 0.5 | 0.60 | 2.972794 | 2.381297 | 1.933847 | 1.672106 | 1.486397 |
| **n = 40** | 0.5 | 0.52 | 23.86772 | 19.11876 | 15.52631 | 13.42486 | 11.93386 |
| | 0.5 | 0.54 | 9.874597 | 7.909849 | 6.423574 | 5.554158 | 4.937299 |
| | 0.5 | 0.56 | 5.613737 | 4.496771 | 3.65182 | 3.157555 | 2.806868 |
| | 0.5 | 0.58 | 3.669529 | 2.939403 | 2.387084 | 2.063998 | 1.834765 |
| | 0.5 | 0.60 | 2.601195 | 2.083635 | 1.692116 | 1.463092 | 1.300597 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **n = 45** | 0.5 | 0.52 | 21.21575 | 16.99446 | 13.80117 | 11.93321 | 10.60788 |
| | 0.5 | 0.54 | 8.77742 | 7.030976 | 5.709843 | 4.93703 | 4.38871 |
| | 0.5 | 0.56 | 4.989988 | 3.99713 | 3.246062 | 2.806715 | 2.494994 |
| | 0.5 | 0.58 | 3.261804 | 2.612803 | 2.121852 | 1.834665 | 1.630902 |
| | 0.5 | 0.60 | 2.312173 | 1.85212 | 1.504103 | 1.300527 | 1.156087 |
| **n = 50** | 0.5 | 0.52 | 19.09418 | 11.13542 | 9.043053 | 7.819098 | 6.950688 |
| | 0.5 | 0.54 | 6.047124 | 4.843928 | 3.933745 | 3.401322 | 3.023562 |
| | 0.5 | 0.56 | 3.567207 | 2.85744 | 2.320522 | 2.006445 | 1.783604 |
| | 0.5 | 0.58 | 2.39885 | 1.921551 | 1.560488 | 1.34928 | 1.199425 |
| | 0.5 | 0.60 | 1.739071 | 1.393048 | 1.131291 | 0.978174 | 0.869535 |
| **n = 55** | 0.5 | 0.52 | 17.35834 | 13.90455 | 11.29186 | 9.763537 | 8.679172 |
| | 0.5 | 0.54 | 7.181525 | 5.752617 | 4.67169 | 4.039388 | 3.590763 |
| | 0.5 | 0.56 | 4.082718 | 3.270379 | 2.655869 | 2.296404 | 2.041359 |
| | 0.5 | 0.58 | 2.668749 | 2.137748 | 1.736061 | 1.501089 | 1.334374 |
| | 0.5 | 0.60 | 1.891778 | 1.515371 | 1.23063 | 1.064067 | 0.945889 |
| **n = 60** | 0.5 | 0.52 | 15.91182 | 12.74584 | 10.35087 | 8.949909 | 7.955908 |
| | 0.5 | 0.54 | 6.583065 | 5.273232 | 4.282382 | 3.702772 | 3.291532 |
| | 0.5 | 0.56 | 3.742491 | 2.997848 | 2.434547 | 2.105037 | 1.871246 |
| | 0.5 | 0.58 | 2.446353 | 1.959602 | 1.591389 | 1.375999 | 1.223176 |
| | 0.5 | 0.60 | 1.73413 | 1.38909 | 1.128077 | 0.975395 | 0.867065 |

**Table 1:** The distribution of the ARL for CUSUM chart of Binomial distribution when the underlying distribution is poisson



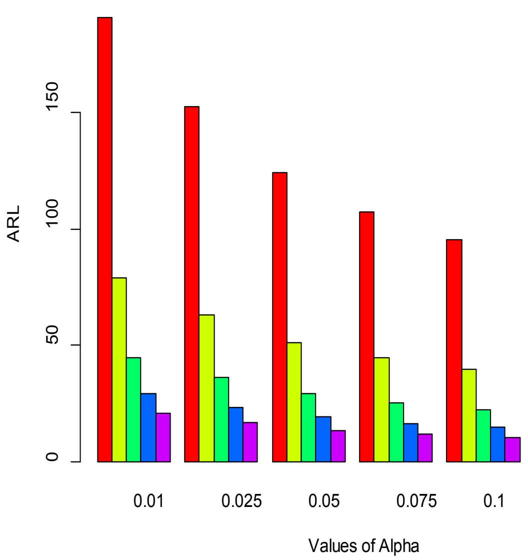


**Fig 2.1:** Dist. of ARL for varying values of θ for n = 05

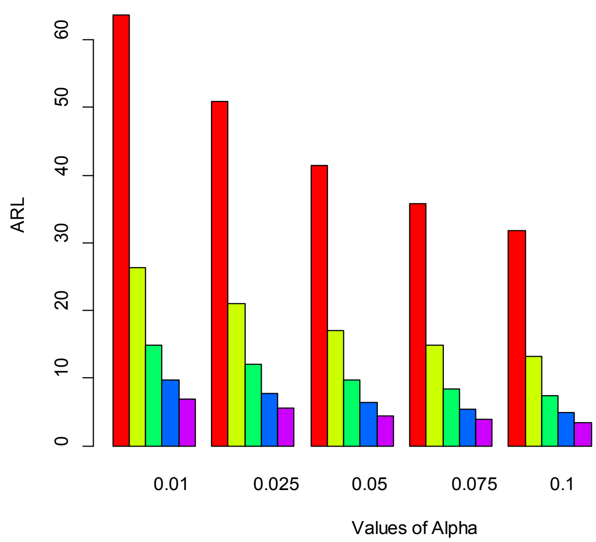


**Fig 2.2:** Dist. of ARL for varying values of θ for n = 15

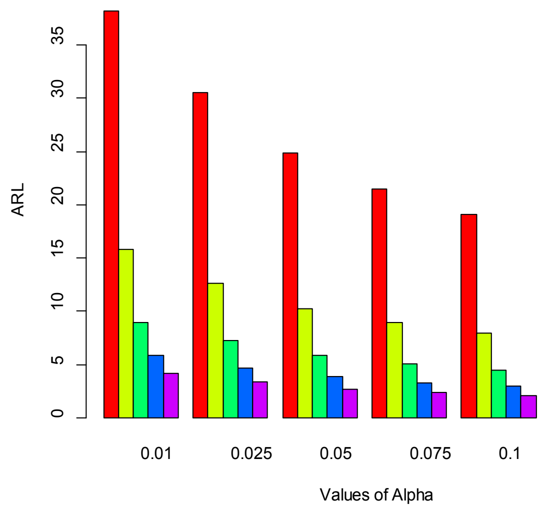


**Fig 2.3:** Dist. of ARL for varying values of θ for n = 25

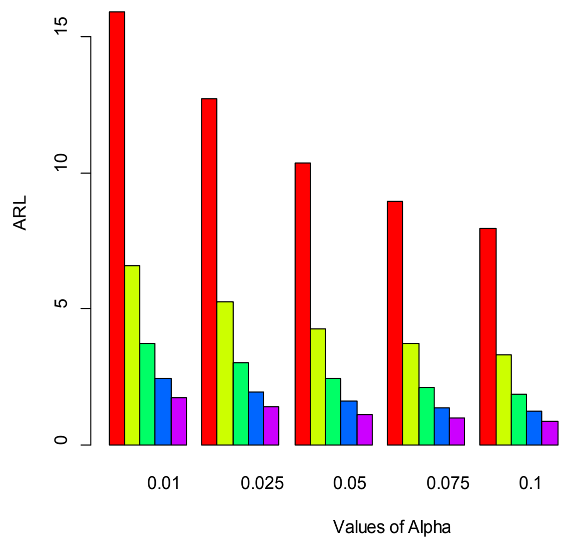


**Fig 2.4:** Dist. of ARL for varying values of θ for n =60

## CONCLUSION

It is observed that the parameters of the V-mask and ARL changes considerably for a slight shift in the parameters of the distribution and as n increases, the value of the ARL decreases. But for a fixed n, the value of the ARL decreases as $\lambda_a$ increases. This result compares favourably with the result of Ashit and Anwer [13] and the result of Johnson and Leone [2].

## REFERENCES

[1] Page, E.S. (1954). Contionous Inspection Schemes, Biometrika, 41, 100-114

[2] JOHNSON N.L and LEONE F.C.(1962), "Cumulative Sum Control Charts: Mathematical Principles Applied to Their Construction and use, part III", Industrial Quality Control, No 19(1), pp 22-28

[3] LORDEN G. (1971)."Procedures for Reacting to a Change in Distribution". Annals of Mathematical Statistics 42 pp 1887-1908.

[4] KENNETT, R. and POLLACK,M. (1983) "On a Sequential Detection of a Shift in the probability of a Rare Event". Journal of the American Statistical Association 78,pp389-395

[5] CHEN R(1978)." A Surveillance System for Congential Malformations". Journal of the American Statistical Association 73,pp 323-327.

[6] LUCAS, J. M AND CROSIER, R. B (1982), "Fast Initial Response for CUSUM Quality Control Schemes", Technometrics Vol. 24, pp 199 – 205.

[7] BROOK, D. AND EVANS, D.A. (1972). "An Approach to the Probability Distribution of Cusum Run Length", Biometrika, 59, 539-549.

[8] VAN DOBBEN DE BRUYN, C.S. (1968). "Cumulative Sum Tests: Theory and Practice". Griffin, London, United Kingdom.

[9] WETHERILL, B.G. (1977). "Sampling Inspection and Quality Control". Chapman and Hall, New York.

[10] EWAN, W.D. and KEMP, K.W. (1960). "Sampling Inspection of Continuous Processes With No Autocorrelation Between Successive Results", Biometrika, 47, 363-380.

[11] CHAMP, C.W., RIGDON, S.E. and SCHARNAGL, K.A. (2001). "A Method for Deriving Integral Equations Useful in Control Chart Performance Analysis, Nonlinear Analysis", 47, 2089- 2101.

[12] CHAMP C.W and RIGDON S.E.(1991). "The performance of exponentially weighted moving Average with estimated parameters. Technometrics 43,156-167.

[13] ASHIT B,CHAKRABORTY and ANWER KHURSHID: (2011) Cumulative Sum Control Charts for Binomial Parameters when the underly distribution is Poisson(2011) Revista Investigacion Operational Vol 32 no 1 pp 12-19