A NOTE ON CUMULATIVE SUM (CUSUM) CONTROL CHART FOR DOUBLY TRUNCATED GAMMA DISTRIBUTION

Ashit B. Chakraborty* and Anwer Khurshid**[1]
*Department of Statistics
St. Anthony's College, Shillong, Meghalaya, India[2]
**Department of Mathematical and Physical Sciences
College of Arts and Sciences, University of Nizwa
Birkat Al Mouz, Oman

ABSTRACT
The CUmulative SUM (CUSUM) control chart for a doubly truncated gamma distribution is developed by using the integral equation method given by Page (1954). The objective of this paper is to study the effect of truncation on average run length (ARL) of the CUSUM scheme for controlling the parameters when the variable characteristic is distributed as doubly truncated gamma distribution. An algorithm for estimating (ARL) is suggested and numerical results are provided.

KEYWORDS: doubly-truncated gamma distribution; CUSUM; integral equation; average run length (ARL).

MSC: 62P30

RESUMEN
La carta de control CUmulative SUM (CUSUM) para una distribución gamma doblemente truncada es desarrollada usando el método de la ecuación integral de Page (1954). El objetivo de este paper es estudiar el efecto del truncamiento del largo promedio de la corrida (ARL) del esquema CUSUM para controlar los parámetros cuando la variable característica se distribuye gamma doblemente truncada . Un algoritmo para estimar (ARL) es sugerido y se proveen resultados numéricos

PALABRAS CLAVE: distribución gamma doblemente truncada; CUSUM; ecuación integral; largo promedio de la corrida (ARL).

1. INTRODUCTION

Control charts are most frequently used for quality improvement and assurance, but they can be applied to almost any situation that entails variation (Qiu, 2014). The CUmulative SUM (CUSUM) control chart procedures have proven to be more effective than the standard Shewhart control charts at detecting small shifts in mean under statistical control of any production processes. Interested readers may see Gan (2007), Hawkins and Olwell (1998), Goel (2011), among others, for a detailed discussion of CUSUM control charting literature.

The two widely used metrics used for assessing the control chart performance, including CUSUM design schemes based on run length characteristics: (i) run-length distribution and (ii) average run length (ARL) (Jones et al., 2004). Performance of a control chart is usually evaluated in terms of its run length distribution. The run length is defined as the number of control chart statistics plotted until the first time the control statistic exceeds a control limit. Numerical evaluation of run-length distributions of CUSUM charts under normal distributions has been extensively investigated. However, accurate approximation of run-length distributions under non-normal or skewed distributions has generally been ignored. Recently, Huang et al. (2013) proposed a fast and accurate algorithm based on the piecewise collocation method for computing the run-length distribution of CUSUM charts under gamma distributions.

In the literature, numerous approaches have been used to study the distributions and expectations of run lengths for the CUSUM control charts. In the past, approximations to the ARL of CUSUM control charts obtained primarily by: (i) numerical solutions to a (Fredholm) integral equation (Hawkins, 1992), (ii) by using a Markov chain approximation (Brooks and Evans, 1972; Chao, 2000), (iii) and by Monte Carlo

[1] anwer@unizwa.edu.om
[2] abc_sac@rediffmail.com

simulation (Hawkins and Olwell, 1998; Fu et al. 2002). These techniques have been used to produce tables (e.g., the accurate tables in Van Dobben de Bruyn (1968)), nomograms, or an algebraic approximation employing roughly 250 tabled constants (Hawkins, 1992).

Gamma distribution, one of the most commonly used continuous probability distributions, is very versatile and provides useful presentation of many practical situations. It is extensively employed in science, engineering and business, to model continuous random variables describing positively skewed probability distributions such as rainfall totals, cohesion or shear strength in swedge, cloud drops, lifetimes of manufacturing products, survival or remission times in chronic diseases etc. For a situation where the failure rate appears to be more or less constant, the exponential distribution is an adequate choice but there are several situations where the failure rate may be increasing or decreasing and, in such cases, gamma distribution which has an exponential right-hand tail is more realistic choice. We refer an interested reader to Johnson et al. (1994), Carolynne (2014), Krishnamoorthy (2016) for overview of the gamma distributions, their properties and applications and exhaustive updated bibliography.

The integral equation method given by Page (1954) is very effective in estimating run lengths for CUSUM and used in the present research. The objective of this paper is to study the effect of truncation on ARL of the CUSUM scheme for controlling the parameters when the variable characteristic is distributed as doubly truncated gamma distribution. The rest of this article is arranged as follows; in Section 2, we discuss briefly about doubly truncated gamma distribution along with its CUSUM scheme. In Section 3, ARL for doubly truncated gamma distribution is proposed and developed. Integral equations using the technique of Krishnamurthy and Sen (1986) has been given in Section 4. Finally, Section 5 provides numerical illustrations and conclusions.

2. DOUBLY TRUNCATED GAMMA DISTRIBUTION AND CUSUM SCHEME

The truncated distributions occur in numerous real-world practical applications. A motivating example of various types of truncated distributions is considered by Kendall and Stuart (1979) as follows: Assume the original variate x cannot simply be observed in part or parts of its range. For example, if x is the distance from the centre of a vertical circular target for fixed radius R on a shooting target, we can only spot x for shots actually hitting the target. If we have no information of how many shots were shot at the target (say n) we evidently have to understand m values of x observed on the target as originating from a distribution ranging from 0 to R. We then state that x is truncated on the right at R. Also, if we denote Y in this example as the distance of a shot from the vertical line through the center of the target, Y may range from $-R$ to R and its distribution is doubly truncated. Likewise, we may have a variate truncated on the left i.e. if observations below a certain value are not recorded.

Let $x_1, x_2, \cdots, x_n$ be a sequence of independent random variables each having the doubly truncated gamma distribution if its probability density function is given by

$$f(x;\gamma,\theta) = \left[\int_A^B exp\{-\gamma x/\theta\} x^{\gamma-1} dx\right]^{-1} exp\{-\gamma x/\theta\} x^{\gamma-1}; A \leq x \leq B \quad \textbf{(1)}$$

Here $\gamma > 0, \theta > 0$. The cases for singly truncated gamma distribution can be easily obtained as special cases if A is replaced by $-\infty$ or B by $+\infty$, the distribution is singly truncated from above or below respectively. More details on the truncated gamma distribution can be found in Coffey and Muller (2000) and Okasha and Alqanoo (2014).

To apply the CUSUM scheme to monitor $\gamma (or \theta)$, we accumulate at each stage of continuous inspection the deviation of X from a reference value k, say, $S = \sum_{i=1}^{n}(X_i - k)$ and if, at the jth stage of the experiment

$$S_j(k) = max\left[0, S_{j-1}(k) + X_j - k\right]; j = 1,2,\cdots \quad \textbf{(2)}$$

reaches or exceeds the decision interval h, the process is stopped and an appropriate action being taken to bring the process back into control.

3. AVERAGE RUN LENGTH (ARL) FOR DOUBLY TRUNCATED GAMMA DISTRIBUTION

ARL, one of the most common metrics used for assessing the control chart performance, is the average number of observations required to detect shift of the process parameter. Here, we shall calculate the average run length of the CUSUM scheme by integral equation method introduced by Page (1954). This integral equation method, though it is more complicated and less versatile but calculates the ARL's more accurately than Johnson's method and Markov approach (Champ and Rigdon, 1991; Rao et al., 2001).

The integral equation is given by

$$L(z) = 1 + L(0)F(k - z) + \int_0^h L(y) f(y + k - z) dy \quad \textbf{(3)}$$

where $L(x)$ is the ARL given that the test starts at the point x, $F(x)$ is the cumulative distribution function and $f(x)$ is the probability density function of x.
We are generally interested in the ARL, $L(0)$ of a test starting at zero. But as is evident from Equation (3), $L(z)$ must be known for all values of z in the interval $(0, h)$ to enable one to find $L(0)$. As Equation (3) can not be solved analytically in general, it will be often necessary to use numerical or approximate results. See, for instance, Hawkins (1992), and Luceno and Puig-Pey (2002) for related discussions.
Equation (3) can be solved advantageously by considering integral equations for average sample number and operating characteristic which are similar to Equation (3) and are derived by Page (1954). They are

$$N(z) = 1 + \int_0^h N(y) f(y + k - z) dy \quad \textbf{(4)}$$

$$P(z) = F(k - z) + \int_0^h P(y) f(y + k - z) dy \quad \textbf{(5)}$$

Several methods of approximate solutions of the above integral equations (3) to (5) solutions are described by van Dobben De Bruyn (1968). Solving for $N(0)$ and $P(0)$ from (4) and (5), $L(0)$ can be obtained by using the relationship given by Page (1954):

$$L(0) = N(0)[1 - P(0)]^{-1} \quad \textbf{(6)}$$

4. METHOD OF ESTIMATION: KRISHNAMURTHY AND SEN METHOD

The integral equations (4) and (5) are linear non-homogeneous integral equations. These equations are called Fredholm integral equations of second kind. Kantorovich and Krylov (1964) showed methods of solving such types of integral equations by transforming them to a system of linear algebraic equations. Following this technique Goel and Wu (1971) solved numerically these equations using 15 Gaussian points and gave a FORTRAN IV computer program for the purpose. Vance (1986) has provided a computer package to obtain ARL considering 24 Gaussian points using recursive approach.
In this paper, we employ a technique provided by Krishnamurthy and Sen (1986). This method is simple and provides almost the same results as obtained by other methods (Jain, 2000). Kakoty and Chakraborty (1990) first used this technique to obtain the values of ARL for doubly truncated normal distribution under continuous acceptance sampling plan.
The integral equation (5) can be written as

$$\psi(x) = g(x) + \int_c^d k(x, t) \psi(z) dt \quad \textbf{(7)}$$

where $\psi(x) = P(z)$, $g(x) = F(k - z)$, $k(x, t) = f(y + k - z)$.
The numerical integration $I = \int_c^d f(x) dx$ can be transformed to

$$I = \frac{d-c}{2} \int_{-1}^1 f(y) dy = \frac{d-c}{2} \sum_{i=0}^n a_i f(y_i) \quad \textbf{(8)}$$

where $y = \frac{2x-(c+d)}{d-c}$,

$$f(x) = f\left[\frac{(c+d)+(d-c)y}{2}\right] = f(y) \quad \textbf{(9)}$$

and a_i's are the Gauss-Legendre weight factors and y_i's are the abscissae for Gauss-Legendre integration.
Now using Equations (8) and (9), Equation (7) can be written as

$$\psi(x) = g(x) + \frac{d-c}{2} \sum_{i=0}^n a_i k(x, y_i) \psi(y_i)$$

or

$$\psi(x) = g(x) + \frac{d-c}{2}[a_0 k(x, y_0)\psi(y_0) + a_1 k(x, y_1)\psi(y_1) + a_2 k(x, y_2)\psi(y_2) + \cdots + a_n k(x, y_n)\psi(y_n)]. \quad \textbf{(10)}$$

Since Equation (10) should be valid for all values of x in the interval (c, d), it must be true for $x = y_i, i = 0(1)n$. Thus, we obtain

$$\psi(y_i) = g(y_i) + \frac{d-c}{2}[a_0 k(y_i, y_0)\psi(y_0) + a_1 k(y_i, y_1)\psi(y_1) + a_2 k(y_i, y_2)\psi(y_2) + \cdots + a_n k(y_i, y_n)\psi(y_n)]. \quad \textbf{(11)}$$

where $i = 0(1)n$.
Let us substitute $\psi(y_i) = \psi_i, g(y_i) = g_i, i = 0(1)n$. Then from Equation (11), we obtain

$$\psi_0 = g_0 + \frac{d-c}{2}[a_0 k(y_0, y_0)\psi_0 + a_1 k(y_0, y_1)\psi_1 + a_2 k(y_0, y_2)\psi_2 + \cdots + a_n k(y_0, y_n)\psi_n]$$

$$\psi_1 = g_1 + \frac{d-c}{2}[a_0 k(y_1, y_0)\psi_0 + a_1 k(y_1, y_1)\psi_1 + a_2 k(y_1, y_2)\psi_2 + \cdots + a_n k(y_1, y_n)\psi_n]$$

$$\vdots$$

$$\psi_n = g_n + \frac{d-c}{2}[a_0 k(y_n, y_0)\psi_0 + a_1 k(y_n, y_1)\psi_1 + a_2 k(y_n, y_2)\psi_2 + \cdots + a_n k(y_n, y_n)\psi_n] \quad \textbf{(12)}$$

In the system of Equations (12) all quantities $\psi_i, i = 0(1)n$ are known and hence these can be solved for ψ_i. We have solved the system of equations by the method of iteration. To do this, we write the system (12) as follows:

$$1 - \lambda a_0 k(y_0, y_0)\psi_0 = g_0 + \lambda[a_1 k(y_0, y_0)\psi_1 + a_2 k(y_0, y_2)\psi_2 + \cdots + a_n k(y_0, y_n)\psi_n]$$
$$1 - \lambda a_1 k(y_1, y_1)\psi_1 = g_1 + \lambda[a_0 k(y_1, y_0)\psi_0 + a_2 k(y_1, y_2)\psi_2 + \cdots + a_n k(y_1, y_n)\psi_n]$$
$$\vdots$$
$$1 - \lambda a_n k(y_n, y_n)\psi_n = g_2 + \lambda[a_0 k(y_n, y_0)\psi_0 + a_1 k(y_n, y_1)\psi_1 + \cdots + a_{n-1} k(y_{n-1}, y_{n-1})\psi_{n-1}] \quad \textbf{(13)}$$

where $\lambda = \frac{d-c}{2}$.

We propose the following algorithm to compute ARL described in this article:

Step 1: Start with the iteration process put $\psi_1 = \psi_2 = \cdots = \psi_n = 0$ in the first equation of (13).

Step 2: Obtain a rough value of ψ_0.

Step 3: Set this value of ψ_0 and $\psi_2 = \psi_3 = \cdots = \psi_n = 0$ in the second equation of (13), to obtain a rough value of ψ_1.

Step 4: Set these values of ψ_0, ψ_1 and $\psi_3 = \psi_4 = \cdots = \psi_n = 0$, obtain a crude value of ψ_2 and so on. Finally, in last equation substitute the values of $\psi_0, \psi_1, \cdots \psi_{n-1}$ obtain from the previous equations to find ψ_n.

Thus, we obtain a first set of values of $\psi_0, \psi_1, \cdots \psi_n$, which we can take as the values of $\psi_i, i = 0(1)n$ in the first iteration. Repeating the process with this set of values of $\psi_i, i = 0(1)n$, we get the values of the second iteration which are just the refined values of ψ_i obtained in the first iteration. The iteration may be terminated when two consecutive set of values of ψ_i are repeated up to a certain degree of accuracy.
In a similar way solution of $N(0)$ can be obtained. Thus knowing $N(0)$, $P(0)$, $L(0)$ can be evaluated from Equation (6). The algorithm and the computer program can be obtained from the first author.

5. NUMERICAL ILLUSTRATIONS AND CONCLUSIONS

It has been observed from the Table 1 (a–c) that as we go on increasing values of k, for fixed values of h, γ, θ and fixed truncation points, the values of ARL increase accordingly. Also, for fixed k, γ, θ and truncation points as we increase the values of h, there is some increasing trend in the values of ARL. Interestingly, there is a decreasing trend in the values of ARL for fixed values of k, h, γ and truncation points when there is an increasing trend in the values of θ. The same decreasing trend is observed for fixed values of k, h, θ and truncation points when there is an increasing trend in the values of γ

Table 1: Values of ARL

A	B	$P(0)$	$N(0)$	ARL
(a) $\boldsymbol{k = 0.26, h = 0.25, \gamma = 0.5, \theta = 0.2}$				
0.1	2	0.490075	1.021732	2.003688
	4	0.488366	1.021617	1.996773
	8	0.487707	1.021573	1.994118
	12	0.485880	1.021450	1.986793
(b) $\boldsymbol{k = 0.75, h = 0.25, \gamma = 0.5, \theta = 0.2}$				
0.1	2	0.909055	1.016757	11.179962
	4	0.905958	1.016673	10.810790
	8	0.905568	1.016663	10.766087
	12	0.901449	1.016553	10.314940
(c) $\boldsymbol{k = 1, h = 0.25, \gamma = 0.5, \theta = 0.2}$				
0.1	2	0.959886	1.009832	25.174326
	4	0.956639	1.009789	23.287701
	8	0.956230	1.009784	23.070297
	12	0.951911	1.009728	20.996978
(d) $\boldsymbol{k = 1, h = 0.5, \gamma = 0.5, \theta = 0.2}$				
0.1	2	0.975719	1.025843	42.248936
	4	0.972374	1.025715	32.128949
	8	0.971084	1.025666	35.470258
	12	0.967506	1.025530	31.560658
(e) $\boldsymbol{k = 0.75, h = 0.5, \gamma = 0.1, \theta = 0.2}$				
0.1	2	0.986327	1.025376	74.991825
	4	0.986244	1.025373	74.540926
	8	0.986119	1.025369	73.868245
	12	0.985582	1.025351	71.117926

(f) $\boldsymbol{k = 0.75, h = 0.5, \gamma = 0.1, \theta = 0.4}$				
0.1	2	0.871285	1.062889	8.257738
	4	0.863210	1.061908	7.767042
	8	0.863219	1.0061900	7.763517
	12	0.863107	1.061816	7.756550
(g) $\boldsymbol{k = 0.75, h = 0.5, \gamma = 0.5, \theta = 0.4}$				
0.1	2	0.816785	1.064975	5.812704
	4	0.782891	1.060470	4.883387
	8	0.780150	1.060120	4.822026
	12	0.778550	1.059913	4.786246

Acknowledgements: The authors thank Prof. Dr. Carlos Narciso Bouza Herrera, Editor and anonymous referees for their helpful comments and suggestions on earlier versions of this paper that led to an improved final version.

RECEIVED: OCTOBER, 2020.
REVISED: JANUARY, 2022

REFERENCES

[1] BROOKS, D.; EVANS, D. A. An Approach to the Probability Distribution of CUSUM Run Length. **Biometrika**, 59, 539-549, 1972.
[2] CAROLYNNE, A. K. **Gamma and Related Distributions**. Bud - Books On Demand, Nordersteadt, Germany, 2014.
[3] CHAMP, C. W.; RIGDON, S. E. A Comparison of the Markov Chain and the Integral Equation Approaches for Evaluating the Run Length Distribution of Quality Control Charts. **Communications in Statistics - Simulation and Computation**, 20, 191-204, 1991.
[4] CHAO, M. T. Applications of Markov Chains in Quality-Related Matters. In PARK, S. H. and VINING, G. G. (Eds.) **Statistical Process Monitoring and Optimization**, New York: Marcel-Dekker, 2000, pp. 175-188.
[5] COFFEY, C. S; MULLER, K. E. Properties of Doubly-Truncated Gamma Variables. **Communications in Statistics - Theory and Methods**, 29, 851-857, 2000.
[6] FU, J. C.; SPIRING, F. A.; XIE, H. On the Average Run Lengths of Quality Control Schemes Using a Markov Chain Approach. **Statistics and Probability Letters**, 56, 369-380, 2002.
[7] GAN, F. F. Cumulative Sum (CUSUM) Chart. **Encyclopedia of Statistics in Quality and Reliability**, 1, 466-470. John Wiley and Sons, New York, 2007.
[8] GOEL, A. L. Cumulative Sum Control Charts. In Balakrishnan, N. (Ed.) **Handbook of Methods and Applications of Statistics: Engineering, Quality Control and Physical Sciences**, John Wiley & Sons, New York, 2011, pp. 120-129.
[9] GOEL A. L.; WU, S. M. Determination of ARL and A Contour Nomogram for CUSUM Charts to Control Normal Mean. **Technometrics**, 13, 221-230, 1971.
[10] HAWKINS, D. M. A Fast Accurate Approximation for Average Run Lengths of CUSUM Control Charts. **Journal of Quality Technology**, 24, 37-43, 1992.
[11] HAWKINS, D. M.; OLWELL, D. H. **Cumulative Sum Charts and Charting for Quality Improvement**. Springer, New York, 1998.
[12] HUANG, W.; SHU, L.; JIANG, W.; TSUI, K.-L. Evaluation of Run-Length Distribution for CUSUM Charts under Gamma Distributions. **IIE Transactions (Institute of Industrial Engineers)**, 45, 981-994, 2013.
[13] JAIN, M. K. **Numerical Solution of Differential Equations**. New Age International Publishing Company, New Delhi, India, 2000.
[14] JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions**, Volume 1, Second Edition. John Wiley and Sons, Hoboken, New Jersey, 1994.
[15] JONES, L.A.; CHAMP, C. W.; RIGDON, S.E. The Run Length Distribution of the CUSUM with Estimated Parameters. **Journal of Quality Technology**, 36, 95-108, 2004.
[16] KAKOTY, S.; CHAKRABORTY, A. B. A Continuous Acceptance Sampling Plan for Truncated Normal Distribution Based on Cumulative Scores. **Quality Reliability**, 2, 15-18, 1990.
[17] KANTOROVICH, L. V.; KRYLOV, V. I. **Approximate Methods for Higher Analysis**. John Wiley and Sons, New York, 1964.
[18] KENDALL, M. G.; STUART, A. **The Advanced Theory of Statistics**, Vol 2, Fourth Edition. Charles Griffin and Company, London, 1979.

[19] KRISHNAMOORTHY, K. **Handbook of Statistical Distributions with Applications**, Second Edition. CRC Press, Florida, 2016.
[20] KRISHNAMURTHY, E. V.; SEN, S. K. **Numerical Algorithms**. East-West Press Pvt. Ltd., New Delhi, India, 1986.
[21] LUCENO, A.; PUI-PEY, J. Computing the Run Length Probability Distribution for CUSUM Charts. **Journal of Quality Technology**, 34, 209-215, 2002.
[22] OKASHA, M. K.; ALQANOO, I. M. Inference on the Doubly Truncated Gamma Distribution for Lifetime Data. **International Journal of Mathematics and Statistics Inventio**n, 2, 1-17, 2014.
[23] PAGE, E. S. Continuous Inspection Schemes. **Biometrika**, 41, 100-115, 1954.
[24] QIU, P. **Introduction to Statistical Process Control**. Chapman & Hall/CRC Boca Raton, Fl., 2014.
[25] RAO, B.; R DISNEY, R.; PIGNATIELLO, J. Uniqueness and Convergence of Solutions to Average Run Length Integral Equations for Cumulative Sum and Other Control Charts. **IIE Transactions**, 33, 463-469, 2001. .
[26] VAN DOBBEN DE BRUYN, C. S. **Cumulative Sum Tests: Theory and Practice**. Charles Griffin and Company, London, 1968.
[27] VANCE, L. C. Average Run Lengths of Cumulative Sum Control Charts for Controlling Normal Means. **Journal of Quality Technology**, 18, 189-193, 1986.