Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v8n1p57

Variable control charts based on Exponential-Gamma distribution

By B. Srinivasa Rao, Ch. Srinivasa Kumar

Published: 26 April 2015

Electronic Journal of Applied Statistical Analysis
Vol. 08, Issue 01, April 2015, 57-66
DOI: 10.1285/i20705948v8n1p57

Variable control charts based on Exponential-Gamma distribution

B.Srinivasa Rao*[a] and Ch.Srinivasa Kumar[b]

[a]*Associate Professor of Statistics,Department of Mathematics & Humanities, R.V.R & J.C College of Engineering,Guntur-522 019, Andhra Pradesh, India.*
[b]*Assistant Professor, Department of Mathematics,K.L.University, Vaddeswaram, Guntur-522 502, Andhra Pradesh, India.*

Published: 26 April 2015

The well known variable control charts for mean and range of subgroups for a skewed distribution, exponential-gamma are constructed by two different approaches. One from the first principles of using the percentiles of sampling distribution of sample mean and sample range and the other is based on the popular Shewart control limits. The coverage probabilities in both the approaches are computed through simulation and compared.

keywords: most probable, percentiles, equi-tailed,EGD,pdf,cdf.

1 Introduction

It is well known that in the theory of distributions, normal distribution and exponential distribution are the basic models exemplified in a number of theoretical results. Specifically exponential distribution is an invariable example for a number of theoretical concepts in reliability studies. It is characterised as constant failute rate (CFR) model also. In case of necessity for an IFR model, ordinarily the choice falls on Weibull model with shape parameter more than 1 (>1), in particular taken as 2. Similar in shape with common characteristics of Weibull , we have gamma distribution as another model. Though it is not as popular as Weibull for reliability studies, gamma distribution has its own prominence as a life testing model. Here, we propose to combine an exponential which is a CFR model and a gamma with shape parameter 2 which is an IFR model because such a combination and related works are not published in the available literature, we make an attempt to consider such a model for our study. The probability density function (pdf) of an exponential-gamma distribution is given by

*Corresponding author: boyapatisrinu@yahoo.com

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

$$f(x) = e^{-x(\lambda+\nu)}[\lambda(1+\nu x)+\nu^2 x], x \geq 0; \lambda > 0, \nu > 0 \tag{1}$$

Its cumulative distribution function (cdf) is

$$F(x) = 1 - e^{-x(\lambda+\nu)}(1+\nu x), x \geq 0; \lambda > 0, \nu > 0 \tag{2}$$

exponential-gamma distribution (EGD) is a skewed , unimodal distribution on the positive real line. Its mean and variance are respectively

$$E(X) = \frac{\lambda + 2\nu}{(\lambda+\nu)^2} \tag{3}$$

$$V(X) = \frac{\lambda^2 + 4\lambda\nu + 2\nu^2}{(\lambda+\nu)^4} \tag{4}$$

Variable control charts for mean and range are developed for non normal distributions by various authors. For example,Edgeman (1989)-Inverse Gaussian distribution, Gonzalez and Viles (2000)-Gamma distribution,Kantam et al. (2006)-Log logistic distribution, Betul Kan and Berna Yaziki (2006)-Burr distribution and references there in. Rao and Kantam (2012)-Half logistic distribution through the approach of percentiles of distributions for $\bar{X}$ and R charts.Rao and Sarath Babu (2012)-Control chart constants based on linear failure rate distribution. Chanand Heng (2003) have developed control chart constants for $\bar{X}$ and R charts in a unified way for a skewed distributions where the constants are dependent on the coefficient of skewness of the distribution. Since exponential-gamma distribution is a skewed distribution , this paper makes an attempt to study in a comparative manner the constants of exponential-gamma distribution and those of Shewart . The remaining part of the paper is planned as follows. Control chart constants through percentiles are given in Section 2.Comparison of two procedures through simulation is given in Section 3.Summary and Conclusions are given in Section 4.

2 Control chart constants through percentiles

2.1 $\bar{X}$ – chart

Let $X_1, X_2, ...X_n$ be a random sample of size n supposed to have been drawn from a exponential-gamma distribution (λ=0.7,ν=0.3, such that , $\lambda + \nu = 1$) . If this is considered as a subgroup of an industrial process data with a targeted population average, under repeated sampling the statistic $\bar{x}$ gives whether the process average is around the targeted mean or not. Statistically speaking, we have to find the 'most probable' limits within which $\bar{x}$ falls. Here the phrase 'most probable' is a relative concept which is to be considered in the population sense. As the existing procedures are for normal distribution only, the concept of 3σ limits is taken as the 'most probable' limits. It is well known that 3σ limits of normal distribution include 99.73% of probability. Hence, we have to search two limits of the sampling distribution of sample mean in exponential-gamma distribution such that the probability content of those limits is 0.9973. Symbolically we have to find L, U such that

$$P(L \leq \bar{x} \leq U) = 0.9973$$

Taking the equi-tailed concept L,U are respectively 0.00135 and 0.99865 percentiles of the sampling distribution of $\bar{x}$ .But sampling distribution of $\bar{x}$ is not mathematically tractable in the case of exponential-gamma distribution. We therefore resorted to the empirical sampling distribution of $\bar{x}$ through simulation thereby computing its percentiles. These are given in the Table 2.1.1

Table 2.1.1: Percentiles of Mean in EGD

n	0.99865	0.99	0.975	0.95	0.05	0.025	0.01	0.00135
2	1.7202	1.4303	1.1988	1.0614	0.1889	0.1759	0.1639	0.1109
3	1.4612	1.2072	1.0887	0.9777	0.2065	0.1935	0.1794	0.1508
4	1.2887	1.0836	0.9731	0.8946	0.2195	0.2060	0.1913	0.1701
5	1.1784	1.0143	0.9191	0.8377	0.2298	0.2156	0.2022	0.1792
6	1.1224	0.9617	0.8834	0.8076	0.2404	0.2239	0.2078	0.1888
7	1.0746	0.9347	0.8521	0.7888	0.2519	0.2343	0.2177	0.1925
8	1.0314	0.8951	0.8303	0.7669	0.2617	0.2435	0.2250	0.2001
9	0.9779	0.8673	0.7979	0.7442	0.2709	0.2494	0.2333	0.2091
10	0.9689	0.8483	0.7857	0.7283	0.2789	0.2557	0.2372	0.2183

The required percentiles from Table 2.1.1 are used in the following manner to get the control limits for sample mean.

$$P(Z_{0.00135} \leq \bar{z} \leq Z_{0.99865}) = 0.9973 \tag{5}$$

where $\bar{z}$ is the mean of sample of size n from a standard exponential-gamma distribution and Z_p are given in Table 2.1.1 for selected values of p. If $\bar{x}$ is the mean of a data following an exponential-gamma distribution with scale parameter σ,then $\bar{x} = \sigma\bar{z}$. Using this in Equation(5) we get

$$P(Z_{0.00135} \leq \frac{\bar{x}}{\sigma} \leq Z_{0.99865}) = 0.9973 \tag{6}$$

$$P(\sigma Z_{0.00135} \leq \bar{x} \leq \sigma Z_{0.99865}) = 0.9973 \tag{7}$$

Unbiased estimate of σ is $\frac{\bar{x}}{\frac{\lambda+2\nu}{(\lambda+\nu)^2}}$. From Equation(7) over repeated sampling, for the i^{th} subgroup mean we can have

$$P(Z_{0.00135}\frac{\overline{\overline{x}}}{\frac{\lambda+2\nu}{(\lambda+\nu)^2}} \leq \bar{x}_i \leq Z_{0.99865}\frac{\overline{\overline{x}}}{\frac{\lambda+2\nu}{(\lambda+\nu)^2}}) = 0.9973 \quad (8)$$

this can be written as

$$P(A^*_{2p} \times \overline{\overline{x}} \leq \bar{x}_i \leq A^{**}_{2p} \times \overline{\overline{x}}) = 0.9973 \quad (9)$$

where $\overline{\overline{x}}$ is grand mean, $\bar{x}_i$ is i^{th} subgroup mean. Thus A^*_{2p}, A^{**}_{2p} are the constants of $\bar{x}$ chart for a exponential-gamma distribution process data. These are given in Table 2.1.2 and are named as percentile constants of $\bar{x}$-chart.

Table 2.1.2
Percentile constants
of $\bar{x}$-chart

n	A^*_{2p}	A^{**}_{2p}
2	0.1731	2.6836
3	0.2353	2.2795
4	0.2653	2.0104
5	0.2796	1.8384
6	0.2946	1.7509
7	0.3004	1.6764
8	0.3121	1.6091
9	0.3263	1.5255
10	0.3405	1.5127

2.2 R-chart

It is well known that 3σ limits of normal distribution include 99.73% of probability. On par with this property, we have to search two limits of the sampling distribution of sample range in exponential-gamma distribution such that the probability content of these limits is 0.9973. Symbolically, we have to find L, U such that

$$P(L \leq R \leq U) = 0.9973$$

where R is the range of sample of size n. Taking the equi-tailed concept L,U are respectively 0.00135 and 0.99865 percentiles of the sampling distribution of range. But sampling distribution of range is not mathematically tractable in the case of exponential-gamma distribution. We therefore resorted to the empirical sampling distribution of range through simulation thereby computing its percentiles. These are given in the Table 2.2.1

Table 2.2.1: Percentiles of Range in EGD

n	0.99865	0.99	0.975	0.95	0.05	0.025	0.01	0.00135
2	1.7456	1.6410	1.5447	1.4141	0.0111	0.0056	0.0023	0.0004
3	1.7616	1.6978	1.6374	1.5523	0.0482	0.0342	0.0204	0.0083
4	1.7744	1.7218	1.6744	1.6111	0.0861	0.0673	0.0506	0.0234
5	1.7768	1.7406	1.6995	1.6496	0.1154	0.0950	0.0732	0.0432
6	1.7783	1.7515	1.7171	1.6705	0.1454	0.1185	0.0940	0.0582
7	1.7981	1.7584	1.7264	1.6875	0.1672	0.1387	0.1121	0.0814
8	1.8115	1.7620	1.7346	1.7032	0.2269	0.1588	0.1329	0.0973
9	1.8309	1.7667	1.7447	1.7136	0.3188	0.1803	0.1470	0.1042
10	1.8498	1.7670	1.7491	1.7194	0.4027	0.2361	0.1623	0.1253

The required percentiles from Table 2.2.1 are used in the following manner to get the control limits for sample range. We know that for a exponential-gamma distribution with scale parameter σ, $\frac{\bar{R}}{\alpha_n - \alpha_1}$ is an unbiased estimator of σ where R is range of sample of size n and α_i is expected value of $\imath^{th}$ standard order statistic in a sample of size 'n' in exponential-gamma distribution. Under repeated sampling of size n, if $\bar{R}$ is the mean of ranges of all samples of the same size n, $\frac{\bar{R}}{\alpha_n - \alpha_1}$ is also an unbiased estimate of σ. From distribution of R, consider

$$P(Z_{0.00135} \leq R \leq Z_{0.99865}) = 0.9973 \quad (10)$$

where $Z_p, (0 < p < 1)$ are given in Table 2.2.1 for selected values of p. From Equation(10), for the $\imath^{th}$ subgroup range we can have

$$P(Z_{0.00135} \frac{\bar{R}}{\alpha_n - \alpha_1} \leq R_i \leq Z_{0.99865} \frac{\bar{R}}{\alpha_n - \alpha_1}) = 0.9973 \quad (11)$$

This can be written as

$$P(D_{3p}^{*}\bar{R} \leq R_i \leq D_{4p}^{**}\bar{R}) = 0.9973 \tag{12}$$

where $\bar{R}$ is mean of ranges, R_i is i^{th} subgroup range. Thus D_{3p}^{*}, D_{4p}^{**} are the constants of R chart for a exponential-gamma distribution process data. These are given in Table 2.2.2.and are named as percentile constants of Range chart.

Table 2.2.2
Percentile constants of Range-chart

n	D_{3p}^{*}	D_{4p}^{**}
2	0.0013	5.2368
3	0.0166	3.5232
4	0.0389	2.9574
5	0.0648	2.6652
6	0.0815	2.4897
7	0.1086	2.3975
8	0.1251	2.3291
9	0.1303	2.2886
10	0.1531	2.2242

3 Comparitive study

The control chart constants for the statistics mean and range, developed in section 2 are based on the population described by EGD. In order to use this for a data, the data is confirmed to follow EGD. Therefore the power of the control limits can be assessed through their application for a true EGD data in relation to the application of Shewart limits . With this back drop we have made this comparative study by Monte-Carlo simulation of 10,000 runs with random samples of size n=2,10 from EGD and calculated the control limits using the constants of section 2 for mean and range in succession. The number of statistic values that have fallen within the respective control limits is evaluated and is named as EGD coverage probability. Similar count for control limits using Shewart constants available in quality control manuals are also calculated. These are named as Shewart coverage probability. The coverage probabilities under the two schemes namely true EGD, Shewart limits are presented and the average run length (ARL) is also tabulated. It is defined as the reciprocal of probability of a point following the out of control limits. However, preferability of one set of limits is based on the ARL which indicates average number of subgroups required to detect the first out of control signal which can also be termed as the average waiting time to notice out of control. Therefore, a chart for which the ARL is less is preferable. Under this criterion the control limits based on percentiles of $\overline{X}$ and R would be preferable to other control limits like Shewart based and skewness based limits and they are tabulated in the tables 3.1 and 3.2 .

Table 3.1: Comparison between the count within the limits of Percentile constants and Shewart constants of $\bar{x}$-chart

		$\%^{le}$ limits				Shewart limits			
n	$\overline{\overline{x}}$	$A^{*}_{2p}\overline{\overline{x}}$	$A^{**}_{2p}\overline{\overline{x}}$	count	ARL	$\overline{\overline{x}} - A_2\overline{R}$	$\overline{\overline{x}} + A_2\overline{R}$	count	ARL
2	0.7545	0.1306	2.0249	**9981**	526.31	0.0360	1.5451	**9944**	178.57
3	0.7555	0.1778	1.7222	**9983**	588.23	0.1220	1.3890	**9977**	434.78
4	0.7560	0.2006	1.5200	**9822**	56.17	0.2005	1.3115	**9812**	53.19
5	0.7562	0.2115	1.3902	**9812**	53.19	0.2539	1.2584	**8920**	9.25
6	0.7563	0.2228	1.3243	**9763**	42.19	0.2928	1.2198	**8483**	6.59
7	0.7564	0.2272	1.2680	**9833**	59.88	0.3236	1.1891	**8274**	5.79
8	0.7573	0.2364	1.2187	**9822**	56.17	0.3490	1.1657	**8020**	5.05
9	0.7574	0.2471	1.1555	**9776**	44.64	0.3724	1.1425	**7665**	4.28
10	0.7600	0.2588	1.1725	**9725**	36.36	0.3944	1.1256	**7179**	3.54

Table 3.2: Comparison between the count within the limits of Percentile constants and Shewart constants of R-chart

		$\%^{le}$ limits				Shewart limits			
n	$\overline{R}$	$D^{*}_{3p}\overline{R}$	$D^{**}_{4p}\overline{R}$	count	ARL	$D_3\overline{R}$	$D_4\overline{R}$	count	ARL
2	0.4205	0.0005	2.2022	**9980**	500.00	0	1.3738	**9411**	16.97
3	0.6192	0.0102	2.1817	**9973**	370.37	0	1.5945	**9626**	26.73
4	0.7620	0.0297	2.2535	**9980**	500.00	0	1.7388	**9936**	156.25
5	0.8704	0.0322	2.3200	**9994**	1666.66	0	1.8410	**9991**	1111.11
6	0.9596	0.0459	2.3892	**9998**	5000.00	0	1.9231	**9992**	1250.00
7	1.0328	0.0784	2.4762	**9994**	1666.66	0.0785	1.9872	**9989**	909.09
8	1.0947	0.1369	2.5497	**9881**	84.03	0.1488	2.0405	**9816**	54.34
9	1.1425	0.1488	2.6149	**9892**	92.59	0.2102	2.0749	**9687**	31.94
10	1.1869	0.1818	2.6400	**9822**	56.18	0.2646	2.1092	**9718**	35.46

4 Summary & Conclusions

These tables show that for a true EGD if the Shewart limits are used in a mechanical way it would result in less confidence coefficient about the decision of process variation for mean and range charts . Hence if a data is confirmed to follow EGD , the usage of Shewart constants in all the above charts is not advisable and exclusive evaluation and application of EGD constants is preferable in statistical quality control.

Acknowledgments

The authors thank A.V.Rama Krishna, Assistant Professor of Mathematics, editor and the reviewers for their helpful suggestions,comments and encouragement, which helped in improving the final version of the paper.

References

Betul Kan and Berna Yaziki(2006). The Individual Control Charts for BURR distributed data *In proceedings of the ninth WSEAS International Conference on Applied Mathematics, Istanbul, Turkey*, pages 645–649, 2006

Chan.Lai.K., and Heng.J.Cui Skewness Correction $\bar{x}$ and R-charts for skewed distributions *Naval Research Logistics*, 50:1–19, 2003

Edgeman.R.L Inverse Gaussian control charts *Australian Journal of Statistics*, 31(1):435–446, 1989

Gonzalez.I., and Viles.I., Semi-Economic design of Mean control charts assuming Gamma Distribution *Economic Quality Control*, 15:109–118, 2000

Kantam.R.R.L., Vasudeva Rao.A., and Rao.G.S, Control Charts for Log-logistic Distribution *Economic Quality Control*, 21(1):77-86, 2006a

Rao.B.S, and Kantam.R.R.L., Mean and Range Charts for Skewed distributions- A comparison based on Half Logistic Distribution *Pakistan Journal of Statistics*, 28(4):437-444,2012

Rao.B.S, and Sarath Babu.G., Variable control Charts based on Linear Failure Rate Model *International Journal of Statistics and Systems*, 7(3):331-341,2012