

## Structural properties of size biased Gamma distribution

A.Ahmed<sup>1</sup>, J.A.Reshi<sup>1</sup>, K.A.Mir<sup>2</sup>

<sup>1</sup> Department of Statistics, University of Kashmir, Srinagar, India.

<sup>2</sup> Department of Statistics, Govt. Degree college Bemina Srinagar, India.

---

**Abstract:** In this paper, a new class of size-biased Gamma distribution is derived. A size-biased gamma distribution, a particular case of weighted gamma distribution, taking the weights as the variate values has been defined. A new distribution which contains as a special case is introduced. The characterizing properties of the model are derived. A Baye's estimator of Size-biased gamma distribution (SBGMD) has also been obtained by using different priors. Also, other important properties including Shannon entropy and Fisher's information matrix which are the measure of the uncertainty in this class of the distribution are derived and studied.

**Keys Words:** Gamma distribution, Size-biased gamma distribution, Uniform prior, Non-Informative priors. Shannon's entropy and Fisher's Information matrix.

---

### I. Introduction

The gamma distribution is used as a lifetime model Gupta and Groll [1], though not, nearly as much as the Weibull distribution. It does fit a widely variety of lifetime adequately, besides failure process models that leads to it. It also arises in some situations involving the exponential distribution. Inference for gamma model has been considered by Engelhard and Bain[2], Chao and Glaser[3], Jamali et al[4] , Lawless [5] and Kalbfleisch and Prentic [6] have made significant contributions.

The gamma distribution has a probability density function of the form:

$$f(x; \alpha, \beta) = \frac{\alpha^\beta}{\Gamma\beta} e^{-\alpha x} x^{\beta-1}, \alpha, \beta > 0 \quad (1)$$

Where  $\alpha > 0$  and  $\beta > 0$  are parameters;  $\alpha$  is a scale parameter and  $\beta$  is sometimes called the index or shape parameter,

$$\text{Also, } E(x) = \frac{\beta}{\alpha} \text{ and } V(x) = \frac{\beta}{\alpha^2}$$

The distribution like the exponential distribution as a special case ( $\beta = 1$ ). The distribution with  $\alpha = 1$  is called One parameter Gamma distribution and has probability density function:

$$f(x; \beta) = \frac{1}{\Gamma\beta} e^{-x} x^{\beta-1}, \beta > 0 \quad (2)$$

Bayesian analysis is an important approach to statistics, which formally seeks use of prior information and Bayes Theorem provides the formal basis for using this information. In this approach, parameters are treated as random variables and data is treated fixed. Ghafoor et al[7] and Rahul et al[8] have discussed the application of Bayesian methods.

The concept of Shannon entropy [9] is the central role of information theory, some-times referred as measure of uncertainty. The entropy of a random variable is defined in terms of its probability distribution and can be shown to be a good measure of randomness or uncertainty. Henceforth we assume that log is to the base 2 and entropy is expressed in bits.

The Fisher information is that a random variable 'X' contains about the parameter  $\theta$  is given by

$$I(\theta) = E \left[ \frac{\partial}{\partial \theta} \log(f(x; \theta)) \right]^2$$

Now, if  $\log f(x; \theta)$  is twice differentiable with respect to  $\theta$  under certain regularity conditions, Fisher's information is given by:

$$I(\theta) = E_\theta \left[ \frac{\partial^2}{\partial \theta^2} \log(f(x; \theta)) \right]$$

## II. Materials And Methods

Suppose  $X$  is a non-negative random variable with its natural probability density function  $f(x; \theta)$ , where the natural parameter is  $\theta$ . Suppose a realization  $x$  of  $X$  under  $f(x; \theta)$  enters the investigator's record with probability proportional to  $w(x; \beta)$ , so that the weight function  $w(x; \beta)$  is a non-negative function with the parameter  $\beta$  representing the recording mechanism. Clearly, the recorded  $x$  is not an observation on  $X$ , but on the random variable  $X_w$ , having a pdf

$$f_w(x; \theta, \beta) = \frac{w(x; \beta) f(x; \theta)}{\int_0^{\infty} w(x) f(x) dx} \quad x > 0 \quad (3)$$

Assuming that  $E(X) = \int_0^{\infty} w(x) f(x) dx$  i.e the first moment of  $w(x)$  exists.

By taking weight  $w(x) = x$  we obtain length biased distribution. Where  $w$  is the normalizing factor obtained to make the total probability equal to unity by choosing  $w = E[w(x, \beta)]$ . The variable  $X_w$  is called weighted version of  $X$ , and its distribution is related to that of  $X$  and is called the weighted distribution with weight function  $w$ . For example, when  $w(x; \beta) = x$ , in which case  $w = \mu$  is called the size-biased version of  $X$ . The distribution of  $X^*$  is called the size-biased distribution with pdf

$$f^*(x; \theta) = \frac{x f(x; \theta)}{\mu} \quad (4)$$

## III. Size -Biased Gamma Distribution

A size- biased Gamma distribution (SBGMD) is obtained by applying the weights  $x^c$ , where  $c=1$  to the Gamma distribution.

We have from relation (1) and (4)

$$\mu'_1 = \int_0^{\infty} x f(x; \theta) dx = \frac{\beta}{\alpha}$$

$$\int_0^{\infty} \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} e^{-\alpha x} x^{\beta} dx = 1$$

Where  $f(x; \alpha, \beta + 1)$  represents a probability density function. This gives the size -biased gamma distribution (SBGMD) as

$$f(x; \alpha, \beta + 1) = \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} e^{-\alpha x} x^{\beta} \quad (5)$$

$$= 0, \text{ otherwise; } \alpha > 0, \beta \geq 0 \quad 0 < x < \infty$$

### 2.1 Special Cases:

Case 1: When  $\beta = 0$ , then Size-biased gamma distribution (SBGMD) (5) reduces to exponential distribution (EPD) as

$$f(x; \alpha) = \alpha e^{-\alpha x}; \quad 0 < x < \infty \quad (6)$$

Case 2: When  $\beta = 1$ , then Size-biased gamma distribution (SBGMD) (5) reduces to size biased exponential distribution (SBEPD) as

$$f(x; \alpha, 1) = \alpha^2 x e^{-\alpha x}; \quad 0 < x < \infty \quad (7)$$

Case 3: When  $\alpha = 1$ , then Size-biased gamma distribution (SBGMD) (5) reduces to Poisson distribution (PD) as

$$f(x; \beta + 1) = \frac{e^{-x} x^\beta}{\beta!}, \beta \geq 0 \tag{8}$$

#### IV. Properties of the Size-biased Gamma Distribution:

In this section, we derive some properties of Size-biased gamma distribution.

##### 3.1 Moments of Size -biased gamma distribution.

The rth moment of Size- biased gamma distribution (5) about origin is obtained as

$$\begin{aligned} \mu'_r &= \int_0^\infty x^r f(x; \alpha, \beta + 1) dx \\ &= \frac{\alpha^{\beta+1}}{\Gamma(\beta + 1)} \int_0^\infty e^{-\alpha x} x^{\beta+r} dx \\ \mu'_r &= \frac{\alpha^{\beta+1}}{\Gamma(\beta + 1)} \frac{\Gamma(\beta + r + 1)}{\alpha^{\beta+r+1}} \end{aligned} \tag{9}$$

Using above relation (9), the mean and variance of the SBGMD is given as

$$\mu'_1 = \frac{\beta + 1}{\alpha} \tag{10}$$

$$\mu_2 = \frac{(\beta + 1)}{\alpha^2} \tag{11}$$

##### 1.2 Shannon’s entropy of size-biased Gamma Distribution

For deriving be entropy of the size-biased Gamma distribution, we need the following two definitions that more details of them can be found in [10].

Definition 3.2.1: The entropy of the discrete alphabet random variable f defined on the probability space is defined by:

$$H_p(f) = -\sum_{i=1}^n p(f = a) \log(p(f = a))$$

It is obvious that

$$H_p(f) \geq 0$$

Definition 3.2.1: The oblivious generalizations of the definition of entropy for a probability density function f defined on the real line as:

$$H(f) = -\int_0^\infty f(x) \log f(x) = E(-\log(x))$$

Provided the integral exists.

Shannon’s entropy is defined as:

$$H[f(x; \alpha, \beta + 1)] = E[-\log\{f(x; \alpha, \beta + 1)\}]$$

$$H[f(x; \alpha, \beta + 1)] = E\left[-\log\left\{\frac{\alpha^{\beta+1}}{\Gamma(\beta + 1)} e^{-\alpha x} x^\beta\right\}\right]$$

$$H[f(x; \alpha, \beta + 1)] = E\left[-\log\left(\frac{\alpha^{\beta+1}}{\Gamma(\beta + 1)}\right)\right] + \alpha E(X) - \beta E(\log X)$$

$$H[f(x; \alpha, \beta + 1)] = \log \frac{\Gamma(\beta + 1)}{\alpha^{\beta+1}} + \alpha \left(\frac{\beta + 1}{\alpha}\right) - \beta E(\log X) \tag{12}$$

Now,

$$E(\log(x)) = \int_0^{\infty} \log x f(x; \alpha, \beta + 1) dx$$

$$E(\log(x)) = \int_0^{\infty} \log x \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} e^{-\alpha x} x^{\beta} dx$$

$$E(\log(x)) = \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} \int_0^{\infty} \log x e^{-\alpha x} x^{\beta} dx$$

Put

$$\alpha x = t, x = \frac{t}{\alpha} \quad dx = \frac{dt}{\alpha}, x \rightarrow 0, t \rightarrow 0. x \rightarrow \infty, t \rightarrow \infty$$

$$E(\log(x)) = \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} \int_0^{\infty} \log \frac{t}{\alpha} e^{-t} \left(\frac{t}{\alpha}\right)^{\beta} \frac{dt}{\alpha}$$

$$E(\log(x)) = \frac{1}{\Gamma(\beta+1)} \int_0^{\infty} (\log t - \log \alpha) e^{-t} t^{\beta} dt$$

$$E(\log(x)) = \frac{1}{\Gamma(\beta+1)} \int_0^{\infty} \log t e^{-t} t^{\beta} dt - \frac{\log \alpha}{\Gamma(\beta+1)} \int_0^{\infty} e^{-t} t^{\beta} dt$$

$$E(\log x) = \frac{\Gamma'(\beta+1)}{\Gamma(\beta+1)} - \log \alpha$$

$$E(\log x) = \Psi(\beta+1) - \log \alpha \tag{13}$$

Substitute the value of equation (16) in equation (15), we get

$$H[f(x; \alpha, \beta + 1)] = \log \frac{\Gamma(\beta+1)}{\alpha^{\beta+1}} + \alpha \left( \frac{\beta+1}{\alpha} \right) - \beta (\Psi(\beta+1) - \log \alpha)$$

$$H[f(x; \alpha, \beta + 1)] = \log \frac{\Gamma(\beta+1)}{\alpha^{\beta+1}} - \beta (\Psi(\beta+1) - \log \alpha) + \beta + 1 \tag{14}$$

The above relation (14) indicates the Shannon's entropy of Size-Biased Gamma Distribution.

### 1.3 Fisher's information matrix of size-biased Gamma Distribution

The Fisher information is that a random variable 'X' contains about the parameter  $\theta$  is given by

$$I(\theta) = E \left[ \frac{\partial}{\partial \theta} \log(f(x; \theta)) \right]^2$$

Now, if  $\log f(x; \theta)$  is twice differentiable with respect to  $\theta$  under certain regularity conditions, Fisher's information is given by:

$$I(\theta) = E_{\theta} \left[ \frac{\partial^2}{\partial \theta^2} \log(f(x; \theta)) \right] \tag{15}$$

The gamma distribution has a probability density function of the form:

$$f(x; \alpha, \beta + 1) = \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} e^{-\alpha x} x^{\beta} \tag{16}$$

Where  $\alpha > 0$  and  $\beta \geq 0$  are parameters;  $\alpha$  is a scale parameter and  $\beta$  is sometimes called the index or shape parameter,

$$\text{Also, } E(x) = \frac{\beta + 1}{\alpha} \text{ and } V(x) = \frac{\beta + 1}{\alpha^2}$$

Applying log on both sides in equation (19), we have

$$\log f(x; \alpha, \beta + 1) = \log \left( \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)} \right) - \alpha x + \beta \log x \tag{17}$$

Differentiating equation (20) partially with respect to  $\alpha, \beta$  we get

$$\frac{\partial}{\partial \alpha} \log f(x; \alpha, \beta + 1) = \frac{\beta + 1}{\alpha} - x \tag{18}$$

$$\frac{\partial^2}{\partial \alpha^2} \log f(x; \alpha, \beta + 1) = -\left( \frac{\beta + 1}{\alpha^2} \right) \tag{19}$$

$$\frac{\partial}{\partial \beta \partial \alpha} \log f(x; \alpha, \beta + 1) = \frac{1}{\alpha} \tag{20}$$

$$\frac{\partial}{\partial \beta} \log f(x; \alpha, \beta + 1) = \log \alpha - \Psi(\beta + 1) + \log x \tag{21}$$

Where

$$\Psi(\beta + 1) = \frac{\Gamma'(\beta + 1)}{\Gamma(\beta + 1)}$$

$$\frac{\partial}{\partial \alpha \partial \beta} \log f(x; \alpha, \beta + 1) = \frac{1}{\alpha} \tag{22}$$

$$\frac{\partial^2}{\partial \beta^2} \log f(x; \alpha, \beta + 1) = -\Psi'(\beta + 1) + (\Psi(\beta + 1))^2 \tag{23}$$

Taking expectations on both sides of the equations (19),(20),(22) and (23), we get

$$-E \left[ \frac{\partial^2}{\partial \alpha^2} \log f(x; \alpha, \beta + 1) \right] = \frac{\beta + 1}{\alpha^2} \tag{I(1,1)}$$

$$-E \left[ \frac{\partial^2}{\partial \beta \partial \alpha} \log f(x; \alpha, \beta + 1) \right] = \frac{-1}{\alpha} \tag{I(1,2)}$$

$$-E \left[ \frac{\partial^2}{\partial \alpha \partial \beta} \log f(x; \alpha, \beta + 1) \right] = \frac{-1}{\alpha} \tag{I(2,1)}$$

$$-E \left[ \frac{\partial^2}{\partial \beta^2} \log f(x; \alpha, \beta + 1) \right] = \Psi'(\beta + 1) - [\Psi(\beta + 1)]^2 \tag{I(2,2)}$$

Now, the Fisher's information matrix of size-biased Gamma Distribution is given as:

$$I(\alpha, \beta + 1) = \begin{pmatrix} -E \left( \frac{\partial^2}{\partial \alpha^2} \log f(x; \alpha, \beta + 1) \right) & -E \left( \frac{\partial^2}{\partial \alpha \partial \beta} \log f(x; \alpha, \beta + 1) \right) \\ -E \left( \frac{\partial^2}{\partial \beta \partial \alpha} \log f(x; \alpha, \beta + 1) \right) & -E \left( \frac{\partial^2}{\partial \beta^2} \log f(x; \alpha, \beta + 1) \right) \end{pmatrix}$$

$$I(\alpha, \beta + 1) = \begin{pmatrix} \frac{\beta + 1}{\alpha^2} & -\frac{1}{\alpha} \\ -\frac{1}{\alpha} & \Psi'(\beta + 1) - [\Psi(\beta + 1)]^2 \end{pmatrix} \quad (24)$$

The above relation (24) is the Fisher's information matrix of Size-biased Gamma Distribution.

**1.4 BAYESIAN ANALYSIS OF SIZE-BIASED GAMMA DISTRIBUTION BY USING UNIFORM PRIOR:**

Let  $X = (x_1, x_2, x_3, \dots, x_n)$  be an identical independently distributed sample from size-biased gamma distribution (5) and then likelihood is defined as:

$$L(x; \alpha : \beta + 1) = \frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha \sum_{i=1}^n x_i} \prod_{i=1}^n x_i^\beta \quad (25)$$

$$L(x; \alpha : \beta + 1) = \frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta \quad (26)$$

Where  $\sum_{i=1}^n x_i = y$

Since  $0 < x < \infty$ , therefore we assume that uniform prior about  $\alpha$ , when  $\beta$  is known

$$g(\alpha) = 1$$

The posterior distribution is given by:

$$\begin{aligned} \Pi(\alpha/x) &= \frac{Lg(\alpha)}{\int_0^\infty Lg(\alpha)d\alpha} \\ \Pi(\alpha/x) &= \frac{\frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta}{\int_0^\infty \frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta d\alpha} \end{aligned} \quad (27)$$

Now, Baye's estimator of  $\alpha$  is:

$$\begin{aligned} \hat{\alpha} &= \int_0^\infty \alpha \Pi(\alpha/x) d\alpha \\ \hat{\alpha} &= \int_0^\infty \alpha \frac{\frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta}{\int_0^\infty \frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta d\alpha} d\alpha \\ \hat{\alpha} &= \frac{n\beta + n + 1}{\sum_{i=1}^n x_i} \end{aligned} \quad (28)$$

**1.5 BAYESAIN ANALYSIS OF SIZE-BIASED GAMMA DISTRIBUTION BASED ON NON-INFORMATIVE PRIOR ( $g(\alpha) = \frac{1}{\alpha}; \alpha > 0$ ):**

Let  $X = (x_1, x_2, x_3, \dots, x_n)$  be an identical independently distributed sample from size-biased gamma distribution (5) and then likelihood is defined as:

$$L(X; \alpha : \beta + 1) = \frac{\alpha^{n\beta + n}}{n\Gamma\beta + 1} e^{-\alpha \sum_{i=1}^n x_i} \prod_{i=1}^n x_i^\beta \quad (29)$$

$$L(y; \alpha : \beta + 1) = \frac{\alpha^{n\beta+n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta \tag{30}$$

Where  $\sum_{i=1}^n x_i = y$

Since  $0 < x < \infty$ , therefore we assume that prior information about  $\alpha$ , when  $\beta$  is known

We know that  $g(\alpha) = \frac{1}{\alpha}; \alpha > 0$

The posterior distribution is given by:

$$\begin{aligned} \Pi(\alpha/x) &= \frac{Lg(\alpha)}{\int_0^\infty Lg(\alpha)d\alpha} \\ \Pi(\alpha/x) &= \frac{\frac{1}{\alpha} \frac{\alpha^{n\beta+n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta}{\int_0^\infty \frac{1}{\alpha} \frac{\alpha^{n\beta+n}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta d\alpha} \end{aligned} \tag{31}$$

Now, Baye's estimator of  $\alpha$  is:

$$\begin{aligned} \hat{\alpha} &= \int_0^\infty \alpha \Pi(\alpha/x) d\alpha \\ \hat{\alpha} &= \int_0^\infty \alpha \frac{\frac{\alpha^{n\beta+n-1}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta}{\int_0^\infty \frac{\alpha^{n\beta+n-1}}{n\Gamma\beta + 1} e^{-\alpha y} \prod_{i=1}^n x_i^\beta d\alpha} d\alpha \\ \hat{\alpha} &= \frac{n\beta + n}{\sum_{i=1}^n x_i} \\ \hat{\alpha} &= \frac{\beta + 1}{\bar{x}} \end{aligned} \tag{32}$$

### V. Conclusion

Size- biased Gamma distribution has been studied. The structural properties including moments- mean variance, mode, median, coefficient of variation, skewness and kurtosis have been discussed and derived. A Baye's estimator of Size-biased gamma distribution (SBGMD) has also been obtained by using uniform prior, non-informative and gamma prior distributions. Also, other important properties including Shannon entropy and Fisher's information matrix which are the measure of the uncertainty in this class of the distribution are derived and studied.

### References

- [1] Gupta.S.S and P.A.Groll, 1961.Gamma distribution in acceptance sampling based on life tests. J. Am.Stat.Assoc. 56:942-970.
- [2] Engelaradt, M. and L.J.Bain, 1978.Prediction intervals for the Weibull process.Technometrics, 20:167-69.
- [3] Chao, M.and R.E.Glaser, 1978.The extract distribution of Bartlett's test statistic for homogeneity of variances with unequal sample sizes.J.Am.Stat.Assoc. 73: 422-426.
- [4] Jamali, A.S., L.J.Lin AND d.Yingzhuo, 2006.Effect of scale parameters in the performance of Shewhart control chart with interpretation rules.J.Applied Sci., 6: 2676-2678.
- [5] Lawless, J.F., 2003.Statistical Models and Methods for Lifetime data.Wiley-Interscience, UK.
- [6] Kalbfleisch, J.D and R.L. The statistical analysis of failure time data.2<sup>nd</sup> Edn., Wiley-Interscience, New York,ISBN; 047136357X Prentice, 2002..
- [7] Ghafoor, A., F.Muhammad and I.A.Arshad, 2005. Bayesian regression with prior non-sample information on mash yield. J. Applied Sci., 5: 187-191.
- [8] Rahul, GP. Singh and O.P. Singh, 2009. Population project OF Kerala using Bayesian methodology .Asian J. Applied Sci., 2: 402-413.
- [9] Jamal Shannon C.E (1948) "A Mathematical theory of communication," Bell system Technical Journal, 27, 623-659.
- [10] Cover T. M., Thomas J.A., Elements of Information Theory, New York, Wiley, 1991.