BIASED EXPONENTIATED GAMMA DISTRIBUTIONS AND ITS VARIATES: STATISTICAL PROPERTIES AND ESTIMATION

[*1]Ikegwu E. M., [1]Amusa S. O., [2]Kuranga Lola J., [1]Ayodeji A.A., and [1]Alakija T. O.

[1]Department of Statistics, Yaba College of Technology, Yaba Lagos
[2]Department of Statistics, Institute of Applied Sciences, Kwara State Polytechnic, Ilorin
*__Corresponding Author__: emmanuel.ikegwu@yabatech.edu.ng, **Phone number:** +23480669408154

Manuscript received on the 5/07/23 and accepted on the 28/07/23

ABSTRACT

This study proposed the Exponentiated Gamma Distribution (EGD) derived from the amalgamation of two existing distribution the exponential distribution and the gamma distribution and explored its variates like the length - biased exponentiated gamma distribution (LeBEGD) and area biased exponentiated gamma distribution (ABEGD). The mathematical expressions of these distributions were also provided. The probability density functions, cumulative distribution functions, hazard and survival functions of these distributions were generated and their statistical properties like the mean, variance and other moments, skewness and kurtosis equally estimated, tested and compared using the Markov Chain Monte Carlo (MCMC) simulated data. The result from simulated data shows that the skewness and kurtosis reduce and tends to normality with higher biases. The study concluded that weighting existing distributions makes them more flexible and amenable to real-life situations.

Keywords: Area-Biased, Exponentiated gamma distribution, Length-Biased, Markov Chain Monte Carlo, Skewness

1.0 INTRODUCTION

The exponentiated gamma distribution has been recognised as one of the important families of distributions. Shawky & Bakoban (2009) introduced a mixture of two – component exponentiated gamma distribution. Ghanizadeh, Pazira, & Lotfi (2011) estimated the parameters of exponentiated gamma distributions in the presence of k outliers. Babokan (2012) considered a finite mixture of exponentiated Raleigh and exponentiated exponential distributions and made inferences using theoritical and numerical methods based on order statistics. Hussain (2014) introduced a a new distribution called the Transmuted exponential distribution describing it as more flexible with some interesting properties. Al-Babtain, Merovci, & Elbatal (2015) introduced a new class of distributions called the Mc-Donald generalised distributions and proposed a five parameter McDonald exponentiated gamma distribution (McEG) deriving its statistical properties and applied same to braking strength of glass fibres. Signh, Singh, & Yadav (2015) proposed and estimated the parameters and reliability function of exponentiated gamma distribution under progressive type – II censored samples using Bayesian procedures under different approximations like the T-K, Lindley and MCMC techniques. Kumar, Singh, Singh, & Bhattacharyya (2015) proposed the classical and Bayesian approaches to estimate the parameter of exponentiated gamma distributions under the general entropy loss functions (GELF) and squared error loss functions (SELF) and compared for progressive type – II censored data with binomial removals.

ISSN 2814-0036

Biased distributions are a special class of weighted distributions introduced by Fisher (1934) and used generally for analysis of lifetime data which occur frequently in researches related to reliability, biomedicine, ecology, environmetrics disease control (Economou, *et al.*, 2019; Saghir, Hamedani, Tazeem, & Khadim, 2017; Das & Roy, 2011).

Al-kadim & Hussein (2014) defined weighted distribution as a pdf satisfying the following:

$$f_w(x) = \frac{w(x)f(x)}{\int_{-\infty}^{\infty} w(x)f(x)dx \quad -\infty \leq x \leq \infty,} \qquad a \leq x \leq b \tag{1}$$

When the weight function depends on the lengths of units of interest (i.e. w(x) = x), the resulting distribution is said to length – biased (Al-kadim & Hussein, 2014) and the statistical properties and interpretation of such distributions with application to wildlife populations and human families was estimated by (Patil & Rao, 1978). Oluyede & George (2002) studied the characteristics of many lengths – biased distributions comparing weighted and length – biased distributions. Arashi, Bekker, & van Niekerk (2019) introduced and discussed the properties of the weighted the distributions of the eigenvalues of the Wishart distribution. Ayeesha (2017) introduces a size biased Lindley distribution as a special case of weighted distributions and compared with the existing distribution. Nkemnole & Ikegwu (2020) proposed and estimated the parameters and properties of the polynomial weighted exponentiated gamma distribution while applying it to glycated haemoglobin of diabetic patients in General Hospital Lagos.

Area – biased distributions are special weighted distributions where the area of units of interest were used as weight functions. Ikegwu (2016) proposed and considered the statistical properties of the new Area Biased Rayleigh Distribution (ABRD) with its applications.

In this study, we proposed the length and area biases of the exponentiated gamma distribution called length – biased exponentiated distribution (LeBEGD) and the area – biased exponentiated gamma distribution (ABEGD) and considered and compared the properties of the biased – distributions with the original mixed distribution.

2.0 METHODOLOGY

This section shows the developed models and the characteristics

2.1 The pdf of the Exponential Distribution

The exponential distribution (ED) is given as:

$$f(x) = \frac{1}{b}e^{-(x-\theta)/b}, \qquad x > \theta,\ b > 0 \tag{2}$$

Where b is the scale parameter and θ is the threshold parameter and the model is used for time before failure.

2.2 The pdf of the Gamma Distribution

The gamma distribution is also given as:

$$f(x) = \frac{1}{\Gamma a b^a}(x-\theta)^{a-1}e^{-(x-\theta)/b}, \qquad x > \theta;\ \theta, a, b > 0 \tag{3}$$

Where a is the shape parameter, b is the scale parameter θ is the threshold parameter, Γ is the Gamma function and is used to model positively scaled data.

2.3 The pdf of the Exponentiated Gamma Distribution (EGD)

ISSN 2814-0036

The Exponentiated Gamma Distribution (EGD), a mix or an amalgam of the existing exponential distribution (2) and the gamma distribution of equation 3 above is given as:

$$f(x,\ a,b,\theta) = \frac{2^a}{\Gamma a b^a}(x-\theta)^{a-1}e^{-2(x-\theta)/b}, \qquad x > \theta;\ \theta, a, b > 0 \qquad (4)$$

To show that the EGD is a true pdf, we integrate the function and determine if it is equal to 1.

$$\int_0^\infty f(x, \quad a,b,\theta) = \int_0^\infty \frac{2^a}{\Gamma a b^a}(x-\theta)^{a-1}e^{-2(x-\theta)/b}\,dx$$

$$= \frac{2^a}{\Gamma a b^a}\int_0^\infty (x-\theta)^{a-1}e^{-2(x-\theta)/b}\,dx = \frac{2^a}{\Gamma a b^a}\int_0^\infty (\frac{bt}{2})^{a-1}e^{-t}\frac{b}{2}\,dt$$

$$= \frac{b^a 2^a}{2^a \Gamma a b^a}\int_0^\infty t^{a-1}e^{-t}\,dt = \frac{b^a 2^a}{2^a \Gamma a b^a}\Gamma a = 1$$

Hence, the EGD is true pdf. With a cumulative density function (cdf) given as:

$$F(X=t) = \frac{\Gamma(\frac{2(t-\theta)}{b}, a)}{\Gamma(a)} \qquad (5)$$

2.4 The pdf of the Length – Biased Exponentiated Gamma Distribution (LeBEGD)

The Length - Biased Exponentiated Gamma Distribution (LeBEGD) is given as:

$$f(x,a,b,\theta) = \frac{2^{a+1}}{\Gamma(a).ab^{a+1}}(x-\theta)^a e^{-2(x-\theta)/b}, \qquad x > \theta;\ \ a,b > 0 \qquad (6)$$

To show that the LeBEGD is a true pdf

$$\int_0^\infty f(x,a,b,\theta) = \int_0^\infty \frac{2^{a+1}}{\Gamma(a).ab^{a+1}}(x-\theta)^a e^{-2(x-\theta)/b}dx$$

$$= \frac{2^{a+1}}{\Gamma(a).ab^{a+1}}\int_0^\infty (\frac{bt}{2})^a e^{-2(x-\theta)/b}\frac{b}{2}dt = \frac{2^{a+1}b^{a+1}}{2^{a+1}\Gamma(a).ab^{a+1}}\int_0^\infty t^a e^{-t}\,dt$$

$$= \frac{2^{a+1}b^{a+1}}{2^{a+1}\Gamma(a).ab^{a+1}}\Gamma(a+1) = \frac{2^{a+1}b^{a+1}}{2^{a+1}\Gamma(a).ab^{a+1}}\Gamma(a).a = 1$$

Hence, the LeBEGD is a true pdf.

With a cumulative density function (cdf) given as:

$$F(X=t) = \frac{\Gamma(\frac{2(t-\theta)}{b}, (a+1))}{\mathrm{a} * \Gamma(a)} \qquad (7)$$

2.5 The pdf of the Area – Biased Exponentiated gamma Distribution (ABEGD)

The pdf Area Biased Exponentiated Gamma Distribution (ABEGD) is given as:

$$f(x,a,b,\theta) = \frac{2^{a+2}}{\Gamma(a).a(a+1)b^{a+2}}(x-\theta)^{a+1}e^{-2(x-\theta)/b}, \quad x > \theta; \; \theta, a, b > 0 \tag{8}$$

To show that the ABEGD is a true pdf

$$\int_0^\infty f(x,a,b,\theta) = \int_0^\infty \frac{2^{a+2}}{\Gamma(a).a(a+1)b^{a+2}}(x-\theta)^{a+1}e^{-2(x-\theta)/b}dx$$

$$= \frac{2^{a+2}}{\Gamma(a).a(a+1)b^{a+2}}\int_0^\infty (\frac{bt}{2})^{a+1}e^{-t}\frac{b}{2}dt = \frac{2^{a+2}b^{a+2}}{2^{a+2}\Gamma(a).a(a+1)b^{a+2}}\int_0^\infty t^{a+1}e^{-t}\,dt$$

$$= \frac{2^{a+2}b^{a+2}}{2^{a+2}\Gamma(a).a(a+1)b^{a+2}}\Gamma(a+1)a(a+1) = \frac{2^{a+2}b^{a+2}}{2^{a+2}\Gamma(a).a(a+1)b^{a+2}}\Gamma(a).a(a+1) = 1$$

Hence, the ABEGD is a true pdf.

With a cumulative density function (cdf) is given as:

$$F(X=t) = \frac{\Gamma(\frac{2(t-\theta)}{b},(a+2)}{\mathrm{a}*(\mathrm{a}+1)*\Gamma(a)} \tag{9}$$

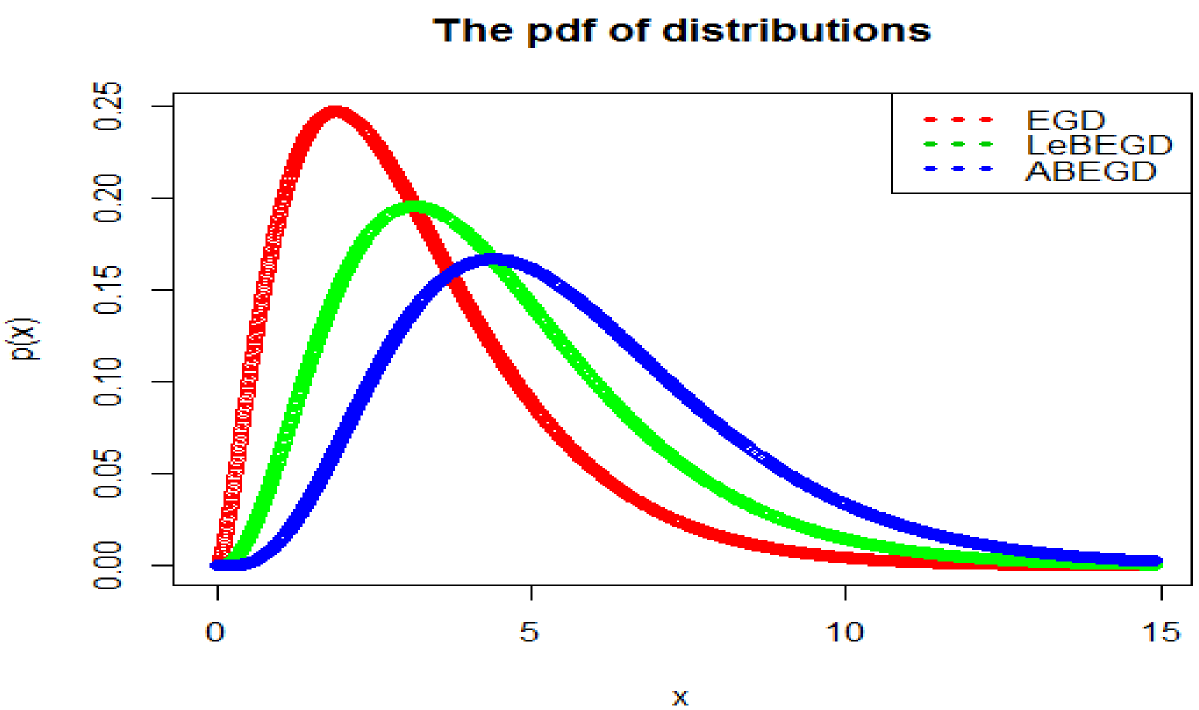

Fig. 1: The pdf of the distributions at n = 181, 0 ≤ x ≤ 15, step 0.01, a = 2.5, b = 2.5

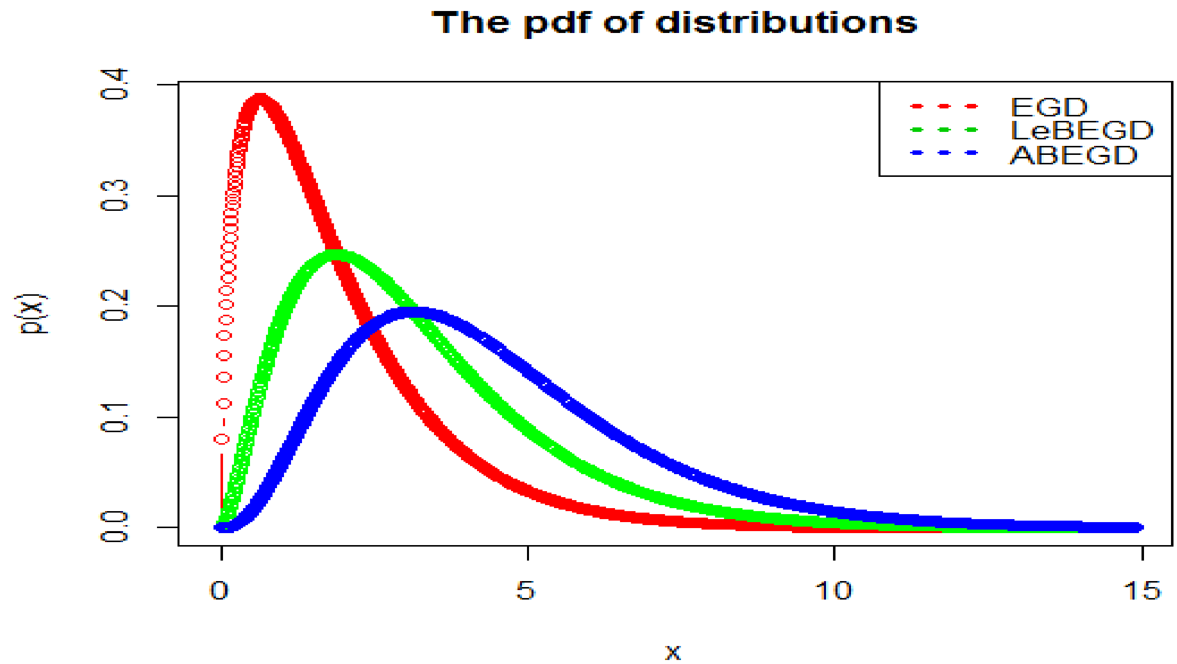

Fig. 2: The pdf of the distributions at n = 181, $0 \leq x \leq 15$, step 0.01, a = 1.5, b = 2.5

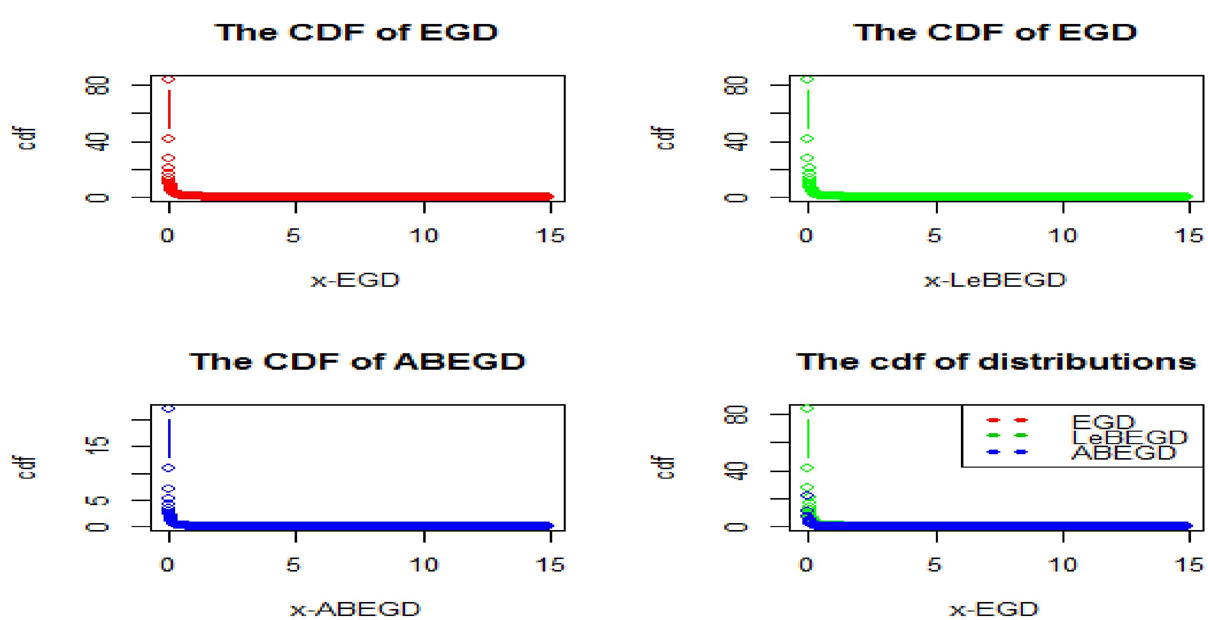

Fig. 3: The cdf of the distributions at n = 181, $0 \leq x \leq 15$, step 0.01, a = 1.5, b = 2.5

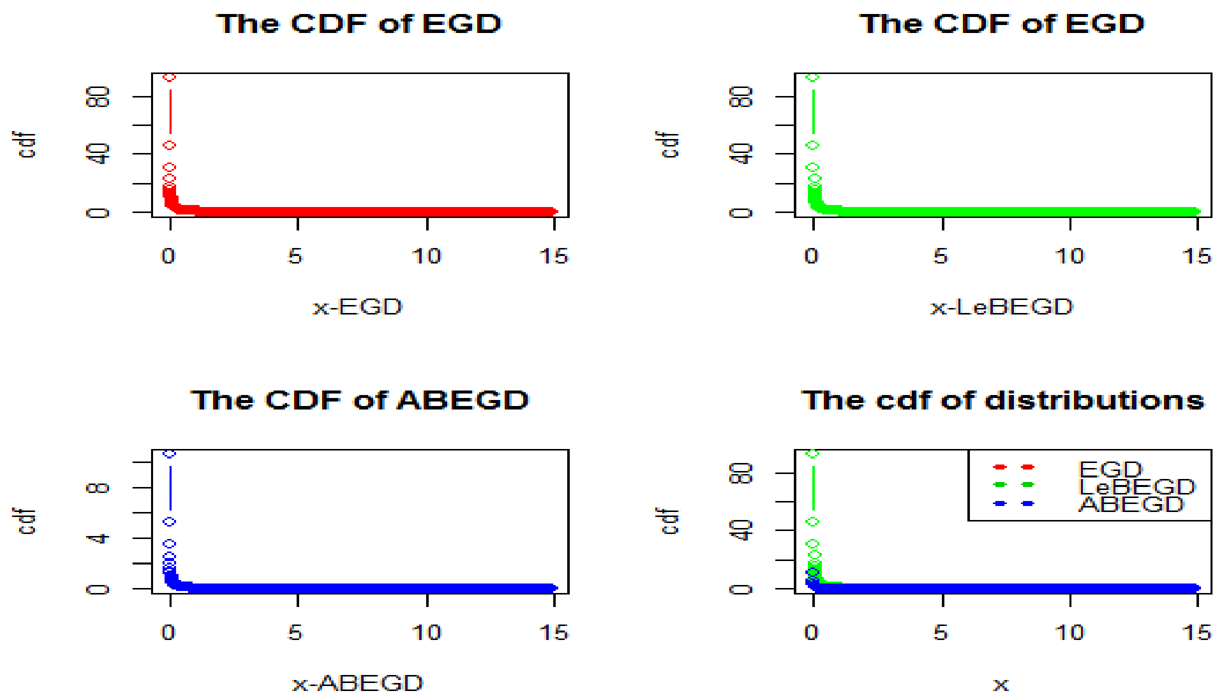

Fig. 4: The cdf of the distributions at n = 181, $0 \leq x \leq 15$, step 0.01, a = 2.5, b = 2.5

2.6 The Survival functions of the Exponentiated Gamma Distribution (EGD) and its Variates

The survival function, S(x) is given as:

$$S(x) = 1 - F(x) \qquad (10)$$

Where F(x) is the cumulative distribution function of the distribution.

The survival function of the Exponentiated Gamma Distribution (EGD) with the cdf, F(x) given in eqn. (5) is given as:

$$S(x) = 1 - \frac{\Gamma(\frac{2(t-\theta)}{b}, a)}{\Gamma(a)} \qquad (11)$$

The survival function of the Length Biased Exponentiated Gamma Distribution (LeBEGD) with the cdf, F(x) given in eqn. (7) is given as:

$$S(x) = 1 - \frac{\Gamma(\frac{2(t-\theta)}{b}, (a+1))}{a * \Gamma(a)} \qquad (12)$$

While the survival function of the Area Biased Exponentiated Gamma Distribution (ABEGD) with the cdf, F(x) given in eqn. (9) is given as:

$$S(x) = 1 - \frac{\Gamma(\frac{2(t-\theta)}{b}, (a+2)}{a * (a+1) * \Gamma(a)} \qquad (13)$$

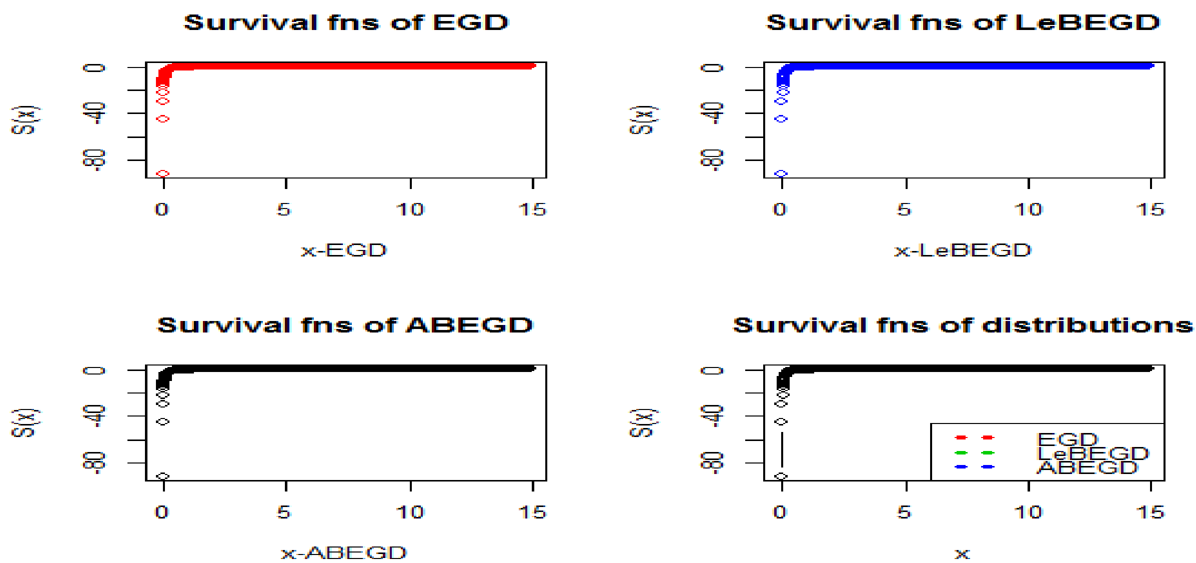

Fig. 5: The survival functions of distributions at n = 181, $0 \leq x \leq 15$, step 0.01, a = 2.5, b = 2.5

2.7 The Hazard Functions of the Exponentiated Gamma Distribution (EGD) and its Variates

The hazard function of probability distributions is given as:

$$h(x) = \frac{f(x)}{1 - F(x)} \tag{14}$$

Where f(x) is the p.d.f and F(x) the cdf of the distribution. The hazard function of the Exponentiated Gamma Distribution (EGD) with S(x) given in eqn. (11) is given as:

$$h(x) = \frac{2^{a}(x-\theta)^{a-1}e^{-2(x-\theta)/b}}{b^{a}\left(\Gamma(a) - \Gamma\left(\frac{2(t-\theta)}{b}, a\right)\right)} \tag{15}$$

The hazard function of the Length Biased Exponentiated Gamma Distribution (LeBEGD) with S(x) given in eqn. (12) is given as:

$$h(x) = \frac{2^{a+1}(x-\theta)^{a}e^{-2(x-\theta)/b}}{b^{a+1}\left(a*\Gamma(a) - \Gamma\left(\frac{2(t-\theta)}{b}, (a+1)\right)\right)} \tag{16}$$

While the hazard function of the Area Biased Exponentiated Gamma Distribution (ABEGD) with S(x) given in eqn. (13) is given as:

$$h(x) = \frac{2^{a+2}(x-\theta)^{a+1}e^{-2(x-\theta)/b}}{b^{a+@}\left(a*(a+1)*\Gamma(a) - \Gamma\left(\frac{2(t-\theta)}{b}, (a+1)\right)\right)} \tag{17}$$

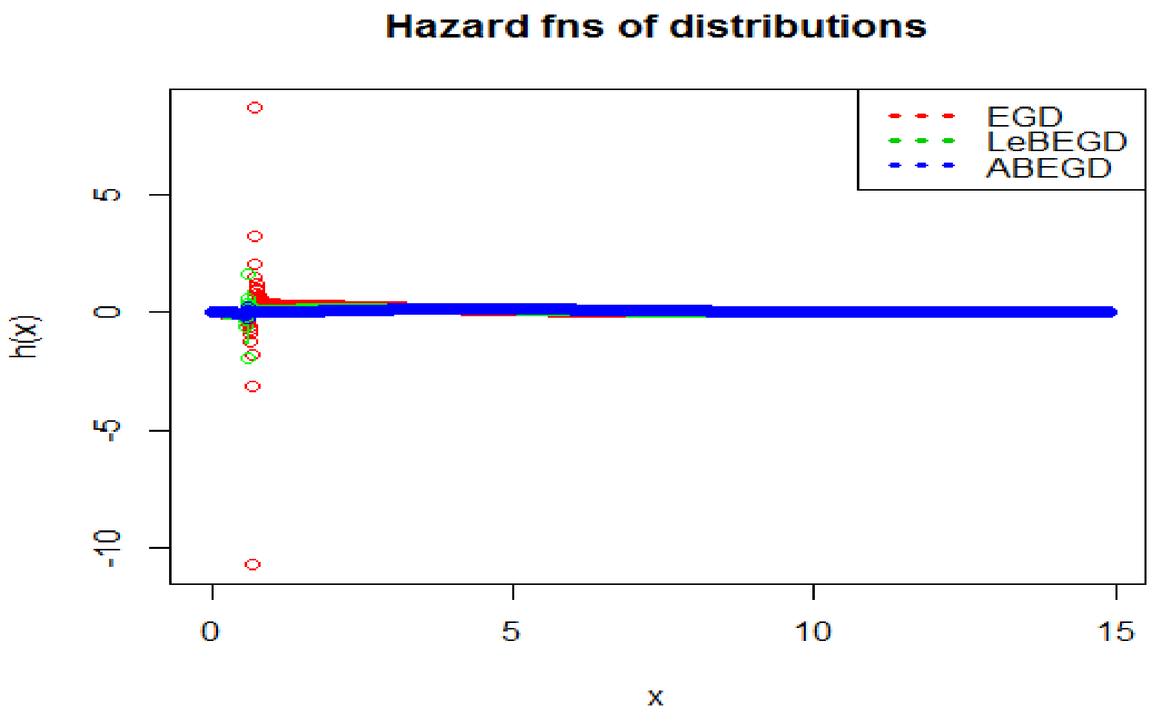

Fig. 6: The hazard functions of distributions at n = 181, $0 \leq x \leq 15$, step 0.01, a = 2.5, b = 2.5

2.8 The Moments of the Exponentiated Gamma Distribution (EGD) and its Variates

From eqn. (4), the moment of the EGD is

$$E[(x-\theta)^r] = \int_0^\infty (x-\theta)^r \frac{2^a}{\Gamma a b^a}(x-\theta)^{a-1} e^{-2(x-\theta)/b} \tag{17}$$

$$= \frac{b^r}{2^r\Gamma(a)}\Gamma(a+r) \tag{18}$$

The first moment (r = 1) of EGD is given as:

$$E(x-\theta) = \frac{a.b}{2}$$

The second moment (r = 2) of EGD is

$$E[(x-\theta)^2] = \frac{a(a+1)b^2}{2^2}$$

And the variance of the EGD is

$$Var\,(x-\theta) = ¼ab^2$$

The third moment (r = 3) of EGD is given as:

$$E[(x-\theta)^3] = \frac{a(a+1)(a+2)b^3}{2^3}$$

The fourth moment (r = 4) of EGD is given as:

$$E[(x-\theta)^4] = \frac{a(a+1)(a+2)(a+3)b^4}{2^4}$$

ISSN 2814-0036

2.9 The Moments of the LeBEGD

The moments of the LeBEGD is given as:

$$E[(x-\theta)^r] = \int_0^\infty (x-\theta)^r \frac{2^{a+1}}{\Gamma(a).ab^{a+1}} (x-\theta)^a e^{-2(x-\theta)/b} \tag{19}$$

$$= \frac{b^r}{2^r a\Gamma(a)} (a+r)\Gamma(a+r) \tag{20}$$

The first moment (r = 1) i.e. the mean of LeBEGD is given as:

$$E(x-\theta) = \frac{(a+1)b}{2}$$

The second moment (r = 2) of LeBEGD is

$$E[(x-\theta)^2] = \frac{(a+1)(a+2)b^2}{2^2}$$

And the variance of the LeBEGD is

$$Var\ (x-\theta) = \frac{(a+1)b^2}{4}$$

The third moment (r = 3) of LeBEGD is

$$E[(x-\theta)^3] = \frac{(a+1)(a+2)(a+3)b^3}{2^3}$$

The fourth moment (r = 4) of LeBEGD is

$$E[(x-\theta)^4] = \frac{(a+1)(a+2)(a+3)(a+4)b^4}{2^4}$$

2.10 The Moments of the ABEGD

The moments of ABEGD from the pdf in eqn. 8 is given as:

$$E[(x-\theta)^r] = \int_0^\infty (x-\theta)^r \frac{2^{a+2}}{\Gamma(a).a(a+1)b^{a+2}} (x-\theta)^{a+1} e^{-2(x-\theta)/b}, \tag{21}$$

$$= \frac{b^r}{2^r a(a+1)\Gamma(a)} (a+r+1)\Gamma(a+r+1) \tag{22}$$

The first moment (r = 1) i.e. the mean of ABEGD is given as:

$$E(x-\theta) = \frac{(a+2)b}{2}$$

The second moment (r = 2) of ABEGD is

$$E[(x-\theta)^2] = \frac{(a+2)(a+3)b^2}{2^2}$$

And the variance of the ABEGD is

$$Var\ (x-\theta) = \frac{(a+2)b^2}{4}$$

ISSN 2814-0036

The third moment (r = 3) of the ABEGD is

$$E[(x - \theta)^3] = \frac{(a+2)(a+3)(a+4)b^3}{2^3}$$

The fourth moment (r = 4) of the LeBEGD is

$$E[(x - \theta)^3] = \frac{(a+2)(a+3)(a+4)(a+5)b^4}{2^4}$$

3.0 RESULTS

The exponentiated gamma distribution and its variates were explored using a Markov Chain Monte Carlo (MCMC) simulated data at 181 sample sizes and at different shape parameters (a) and value of the centred theta. The analysis was done using r-Studio. The graphical representations and results are shown together for ease of comparison as was one of the objectives of the study. However, because it is still an evolving work, real-life application was not explored. However, the real-life application was explored and applied in Nkemnole & Ikegwu (2020 a & b).

Statistical Properties of the EGD

Table 1: Statistical properties of the three distributions for $0 \le x \le 9$ at 0.05 increment, n = 181, a = 2.5, b = 2.5 and for $0 \le x \le 9$ at 0.15 increment, n = 61, a = 1.5, b = 2.5.

Distributions	EGD	LeBEGD	ABEGD
Mean	1.875	3.125	4.375
Variance	2.344	3.906	5.469
Skewness	3.615	1.521	1.386
Kurtosis	4.200	2.829	2.270
3rd Moment	51.269	76.904	169.189
4th Moment	144.196	528.717	1374.664

Table 1 reveals that the mean value and the variance increase with every additional weight to the exponentiated gamma distribution. The values obtained from the simulated data were exactly the same when n = 61 with a = 1.5 and b = 2.5 and when n = 181, a = 2.5 and b = 2.5.

While the mean, variance, and moments increase with additional weight, the skewness and kurtosis decrease with additional weights to the exponentiated gamma distribution. This implies that weighting in the length-biased and area-biased exponentiated gamma distributions helps manage the complexities inherent in real-life data and moves the exponentiated gamma distribution closer to normality. The third and fourth moments are also given in the table.

4. 0 DISCUSSION

This proposed the exponentiated gamma distribution (EGD) and its variates Length-biased EGD and Area -Biased EGD by weighting the exponentiated gamma distribution. It determined that all three variates were proper probability density functions. The weighted distribution (length-biased and area-biased) provides a better fit than parent models while higher weights result in better fits agreeing with Mohiuddin, Dar, Khan, & Ahajeeth (2022). The study also agreed with Nkemnole & Ikegwu (2020b) that increasing the weight gives better fit to the model.

5.0 CONCLUSIONS

Having looked at the exponentiated gamma distribution and its length and area biased variates, it was very obvious that the biased variates where better than the generic distribution as their skewness and kurtosis where indeed lower than the generic distribution. The area biased distribution was the best with the least skewness and kurtosis though with higher variance. Its distribution can best estimate reducing values of financial commitments. Since the distribution is evolving, the model selection criteria such as: AIC, CAIC, BIC, HQIC; Goodness of fit Criteria like Shapiro Wilk test, Cramer Von test, Kolmogorov test and MSE, Bias and RMSE will be performed in the application stage of the work. All these are very important to make the paper stronger.

REFERENCES

Al-Babtain, A. A., Merovci, F., & Elbatal, I. (2015). The McDonald exponentiated gamma distribution and its statistical properties. *SpringerPlus [online], 4*(2). Retrieved January 12, 2016, from http://www.springerplus.com/content/4/1/2

Al-kadim, K. A., & Hussein, N. A. (2014). New Proposed Length-Biased Weighted Exponential and Rayleigh Distribution with Application. *Mathematical Theory and Modeling, 4*(7), 137 - 152.

Arashi, M., Bekker, A., & van Niekerk, J. (2019). Weighted distributions of eigenvalues. Linear Algebra and its Applications. *561*, 24 - 40. doi:10.1016/j.laa.2018.09.019

Ayeesha, A. (2017). Size Biased Lindley Distribution and Its Properties a Special Case of Weighted Distribution. *Applied Mathematics, 8*(6), 808 - 819. doi:10.4236/am.2017.86063

Babokan, R. A. (2012). A study on mixture of exponentiated Rayleigh and exponentiated exponential distributions based on order statistics. *Journal of Mathematics and System Science, 2*, 163 - 170.

Das, K. K., & Roy, T. D. (2011). Applicability of Length Biased Weighted Generalized Rayleigh Distribution. *Advances in Applied Science Research, 2*(4), 320 - 327.

Economou, P., Batsidis, A., Tzavelas, G., & Alexopoulos, P. (2019). Berkson's paradox and weighted distributions: An application to Alzheimer's disease. *Biometrical Journal*, 1 – 12. doi:10.1002/bimj.201900046

Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics, 6*, 13 - 25.

Ghanizadeh, A., Pazira, H., & Lotfi, R. (2011). Classical estimations of the exponentiated gamma distribution parameters with presence of k outliers. *Australian Journal of Basic Applied Sciences*, *5*(3), 571 - 579.

Hussain, M. A. (2014). Transmuted exponentiated gamma distribution: A generalisation of the Exponentiated Gamma Probability distribution. *Applied Mathematical Sciences [online]*, *8*. doi:10.12988/ams.2014.35258

Ikegwu, E. M. (2016). *Statistical properties of the Area Biased Rayleigh distribution with application.* Ibadan: University of Ibadan.

Kumar, D., Singh, U., Singh, S. K., & Bhattacharyya, G. (2015). Bayesian estimation of exponentiated gamma parameter for Progressive Type II Censored data with Binomial removals. *Journal of Statistical Applications and Probability*, *4*(2), 265 - 273.

Nkemnole, E. B., & Ikegwu, E. M. (2020a). Estimation of parameters of the polynomial weighted exponentiated gamma distribution with applications to glycated haemoglobin of diabetic patients. *Journal of Scientific Research and Development*, *19*(2), 126 - 144.

Nkemnole, E. B., & Ikegwu, E. M. (2020b). Poly-Weighted Exponentiated Gamma Distribution with Applications. *Journal of Statistical Theory and Applications*, *9*(3), 446 - 459. doi:10.2991/jsta.d.201016.002

Oluyede, B. O., & George, E. O. (2002). On Stochastic inequalities and comparisons of reliability measures for weighted distributions. *Mathematical Probability and Engineering*, *8*, 1 - 13.

Patil, G. P., & Rao, C. R. (1978). Weighted distributions and size-biased sampling with applications to wildlife populations and human families. *Biometrics*, *34*, 179 - 189.

Saghir, A., Hamedani, G. G., Tazeem, S., & Khadim, A. (2017). Weighted Distributions: A Brief Review, Perspective and Characterizations. *International Journal of Statistics and Probability*, *6*(3), 109. doi:10.5539/ijsp.v6n3p109

Shawky, A. I., & Bakoban, R. A. (2009). On finite mixture of two - component Exponentiated Gamma Distribution. *Journal of Applied Sciences Research*, *5*(10), 1351 - 1369.

Signh, S. K., Singh, U., & Yadav, A. S. (2015). Bayesian estimation of exponentiated gamma distribution under Progressive Type - II Censoring using different approximation techniques. *Journal of Data Science*, *13*, 551 - 568.