Communication in Physical Sciences, 2022, 8(3):406-412

Statistical Properties and Application of Bagui-Liu-Zhang Distribution

Emmanuel Wilfred Okereke*, Sunday Ngozi Gideon, Kingsley Uchendu, Johnson Ohakwe and Onyinyechi Monica Sunday

Received: 27 March 2022/Accepted26 August 2022/Published online:28 August 2022

Abstract: *This paper extended the work of Bagiu et al. (2020) who defined the probability density function of a new one-parameter continuous distribution through the moment generating function approach. The new distribution called Bagiu-Liu-Zhang distribution is the distribution of the exponential mixture of the shifted exponential random variable. Properties of the distribution such as its cumulative distribution function (cdf), moments, coefficients of skewness and kurtosis, reliability function and hazard rate function were derived. The maximum likelihood estimator of the model parameter was also determined. We illustrated the usefulness of the distribution by comparing its fit to a real data set to the fit of the exponential distribution to the same data. The numerical results obtained indicate that the distribution can be a more suitable model for some continuous data than the exponential distribution and several one-parameter distributions.*

Keywords: *Exponential distribution, maximum likelihood method, mixture model, moment generating function, shifted exponential distribution*

Emmanuel Wilfred Okereke
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
Email: okereke.emmanuel@mouau.edu.ng
Orcid id: 0000-0002-6578-8324

Sunday Ngozi Gideon,
Department of Statistics, Abia Polytechnic, Aba, Abia State, Nigeria
Email: dearngos2012@gmail.com

Kingsley Uchendu
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria
Email: kingsleyuchendu2@yahoo.com
Orcid id: 0000-0002-5986-5557

Johnson Ohakwe
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

Onyinyechi Monica Sunday
Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

1.0 Introduction

Statistical distributions are widely used to model several random phenomena. Consequently, the exponential distribution and its extension, that is the shifted exponential distribution have been accepted as tools for finding solutions to the modelling of lifetime data (Lawless, 1977), quality control (Santiago and Smith, 2013; Albazi *et al.*, 2020, Chong *et al.*, 2021) and queuing theory (Noor *et al.*, 2020) among others.

Familiar and commonly applied probability distributions are discrete distributions and continuous distributions. Discrete distributions are used to model discrete data while continuous distributions are suitable for modelling continuous data. The exponential and shifted exponential distributions are continuous distributions. To find an appropriate distribution for a data set, we first determine whether it is discrete or continuous. If the data are continuous, an appropriate distribution can be fitted to the data by studying some characteristics such as skewness and kurtosis of the data. Distributions like gamma and Weibull distributions are positively skewed and can be fitted to the right-skewed data. Negatively skewed distributions are also available for modelling the left-skewed data. Also, the normal distribution is among the symmetric distributions in the literature. If we classify distributions based on their coefficients of kurtosis, then we haave platykurtic, mesokurtic and leptokurtic distributions.

In previous studies, authors such as Karlis and Xekalaki (2005) and Villa and Escobar (2006) emphasized the need to fit a mixture model to a data set in a statistical problem with a mixture structure. They posited that mixture models possess the flexibility level that is needed to provide good fits to data that occur in multiple stages. There are several mixture models in statistical science literature. For instance, the negative binomial distribution is the distribution of the Poisson mixture of gamma random variable. This discrete distribution play a key role in the analysis of overdispersed count data.

The derivation of mixture distributions from hierarchical models has attracted the attention of some authors (Casella and Berger, 2002; Karlis and Xekalaki, 2005). In a more recent publication, Bagui *et al*. (2020) derived the probability density function (PDF) of the exponential mixture of the shifted exponential distribution using the moment generating function approach. This new distribution will thereafter be called the Bagiu -Liu-Zhan distribution. To the best of our knowledge, only the PDF of the new distribution is known. Important properties, such as the cumulative distribution function (CDF), reliability function, hazard rate function and moments of the distribution have not been discussed. The present study seeks to extend the work of Bagiu *et al*. (2020) through the consideration of several properties of the distribution. Also considered are the point estimation of the parameter for the distribution as well as the comparison of the fit of the distribution to some data with that of the exponential distribution.

2.0 Statistical Properties of Bagiu-Liu -Zhan Distribution

Here, we derive some properties of the distribution. Specifically, attention is given to the cdf, raw moments, reliability function and hazard rate function. It is possible to determine these properties if the pdf of the distribution is known. A continuous random variable X follows the Bagiu-Liu-Zhang distribution if its pdf is of the form (Bagiu *et al*., 2020):

$$f(x) = \lambda(1+\lambda)e^{-\lambda x}(1-e^{-x}), x \geq 0.$$

The two essential conditions under which the function defined in (2.1) is said to be a pdf of a continuous random variable X are $f(x) \geq 0$ and the integral of $f(x)$ over the range of values of x is 1. These conditions can easily be verified. Figure 1 shows the plots of the pdf of Bagui-Liu-Zhang distribution. It can be observed that the distribution is unimodal.

2.1 The CDF of Bagiu-Liu-Zhang distribution

Knowledge of the CDF of a continuous random variable enables us to determine other properties of the distribution like the reliability function and hazard rate function. Let $F(x)$ be the cdf of the Bagui-Liu-Zhang random variable X. Then

$$F(x) = \int_0^x f(t)dt,$$

where

$f(t) = \lambda(1+\lambda)e^{-\lambda t}(1-e^{-t})\,dt$. Thus

$$F(x) = \int_0^x \lambda(1+\lambda)e^{-\lambda t}(1-e^{-t})\,dt$$

$$= \lambda(1+\lambda)\int_0^x e^{-\lambda t}(1-e^{-t})\,dt$$

$$= \lambda(1+\lambda)\int_0^x e^{-\lambda t} - e^{-(1+\lambda)t}\,dt$$

Evaluating the integral on the right hand side of the equation above leads to

$$F(x) = 1 - e^{-\lambda x} - \lambda e^{-\lambda x} + \lambda e^{-(1+\lambda)x}.$$

The CDF of the Bagiu-Liu-Zhang distribution is graphed in Fig. 2.

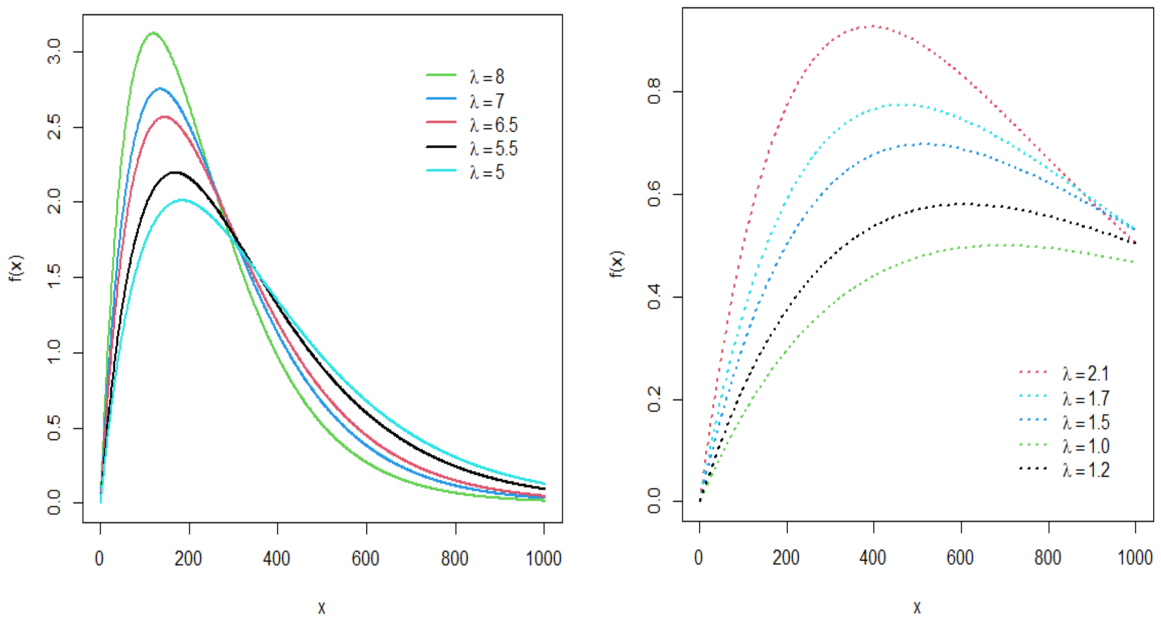

Fig.1: Graphical Representation of the pdf of Bagui-Liu-Zhang distribution for different values of λ.

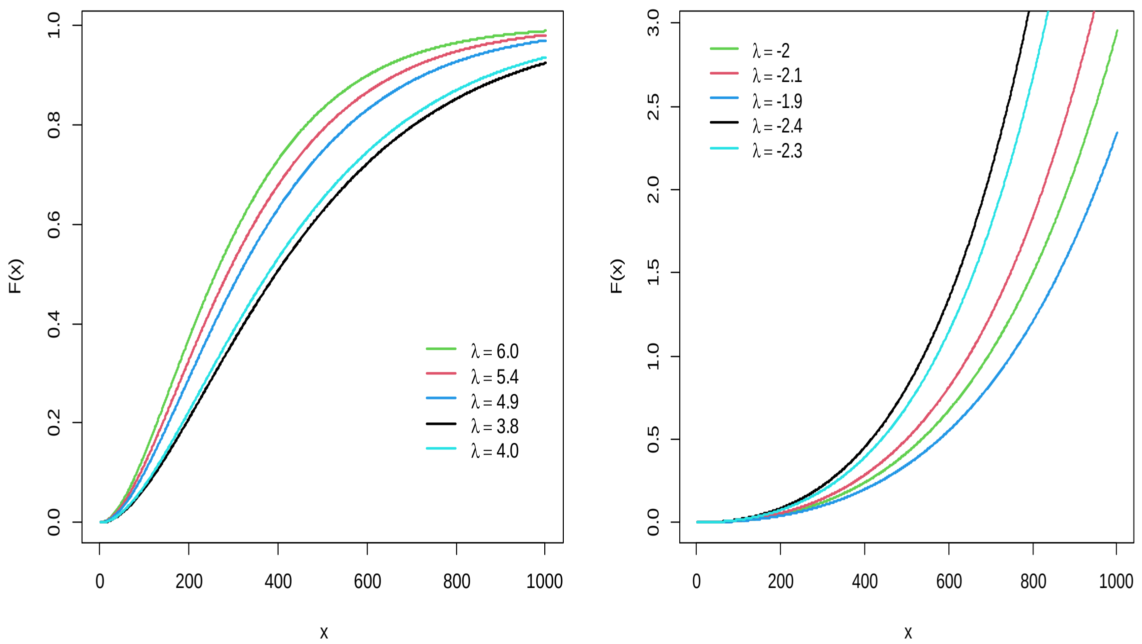

Fig. 2: Graphical representation of the CDF of Bagui-Liu-Zhang distribution for different values of λ

2.2 *Raw moments*

The rth raw moment of the Bagui-Liu-Zhang distribution is given by

$$E(X^r) = \int_0^\infty x^r f(x)dx$$

$$= \int_0^\infty x^r \lambda(1+\lambda)e^{-\lambda x}(1 - e^{-x})\,dx$$

$$= \lambda(1+\lambda)\int_0^\infty x^r e^{-\lambda x}\,dx - \int_0^\infty x^r e^{-(1+\lambda)x}\,dx$$

$$= \frac{[(1+\lambda)^{r+1} - \lambda^{r+1}]\Gamma(r+1)}{\lambda^r\,(1+\lambda)^r}. \qquad (2)$$

Using (2.2), we obtain

$$E[X] = \frac{1+2\lambda}{\lambda(1+\lambda)} \qquad (3)$$

$$E[X^2] = \frac{[(1+\lambda)^{r+1} - \lambda^{r+1}]\Gamma(r+1)}{\lambda^r\,(1+\lambda)^r}. \qquad (4)$$

$$E[X^3] = \frac{[(1+\lambda)^{3+1} - \lambda^{3+1}]\Gamma(3+1)}{\lambda^3\,(1+\lambda)^3} \qquad (5)$$

and

$$E[X^4] = \frac{[(1+\lambda)^{4+1} - \lambda^{4+1}]\Gamma(4+1)}{\lambda^4\,(1+\lambda)^4}$$

$$= \frac{24 + 120\lambda + 240\lambda^2 + 240\lambda^3 + 120\lambda^4}{\lambda^4\,(1+\lambda)^4}.$$

Let the moment based coefficient of skewness be S. Then

$$S = \frac{E(X-\mu)^3}{\mu_2^{3/2}} = \frac{4\lambda^3+6\lambda^2+6\lambda+2}{\left(\sqrt{2\lambda^2+2\lambda+1}\right)^3} \qquad (6)$$

Consequently, the coefficient of kurtosis is given as

$$K = \frac{E(X-\mu)^4}{\mu_2^2} = \frac{24\lambda^4+48\lambda^3+60\lambda^2+60\lambda-9}{(2\lambda^2+2\lambda+1)^2} \qquad (7)$$

2.3 *Survival function and hazard rate function*

The survival function for the Bagiu-Liu-Zhang distribution is given by;

$s(x) = 1 - F(x)$

Therefore

$$s(x) = e^{-\lambda x} + \lambda e^{-\lambda x} - \lambda e^{-(1+\lambda)x}. \qquad (8)$$

For the hazard rate function of the Bagiu-Liu-Zhang distribution, we have

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

$$= \frac{\lambda(1+\lambda)e^{-\lambda x}(1-e^{-x})}{1-[1-e^{-\lambda x} - \lambda e^{-\lambda x} + \lambda e^{-(1+\lambda)x}]}$$

$$= \frac{\lambda(1+\lambda)e^{-\lambda x}(1-e^{-x})}{e^{-\lambda x}+\lambda e^{-\lambda x} - \lambda e^{-(1+\lambda)x}}. \qquad (9)$$

Fig.3 shows the graphical representation of the hazard rate function of Bagui-Liu-Zhang distribution. We can deduce that the hazard rate function of the distribution can be an increasing function.

3.0 Maximum Likelihood Estimation

The likelihood function that is associated with the Bagui-Liu-Zhang distribution can be written as

$$L(x_1, x_2, \dots, x_n; \lambda) = \prod_{i=1}^{n}[f(x_i;\lambda)]$$

$$= \prod_{i=1}^{n}\lambda(1+\lambda)e^{-\lambda x_i}(1-e^{-x_i}) = [\lambda(1+\lambda)]^{\boldsymbol{n}}e^{-\lambda D}\prod_{i=1}^{n}(1-e^{-x_i}),$$

where, $D = \sum_{i=1}^{n} x_i$.

The log-likelihood function is given by

$$\log L(\lambda) = n\log\lambda + n\log(1+\lambda) - \lambda D + \sum_{i=0}^{n}\log(1-e^{-x_i}).$$

The maximum likelihood estimates of λ, say $\hat{\lambda}$ is the solution of the equation;

$$\frac{\partial \log L(\lambda)}{\partial\lambda} = 0.$$

For $\frac{\partial \log L(\lambda)}{\partial\lambda} = 0$, we have

$$\frac{n}{\lambda} + \frac{n}{1+\lambda} - D = 0.$$

Multiplying both sides by $\lambda(1+\lambda)$ and simplifying the result lead to

$$-\lambda^2 D + (2n - D)\lambda + n = 0$$

Multiplying both sides of the above by -1 yields

$\lambda^2 D - (2n - D)\lambda - n = 0.$

Solving the quadratic equation by formula method and considering the fact that parameter of the distribution is a positive quantity, the estimate is found to be

$$\hat{\lambda} = \frac{(2n-D)+\sqrt{(2n-D)^2+4nD}}{2D} \qquad (10)$$

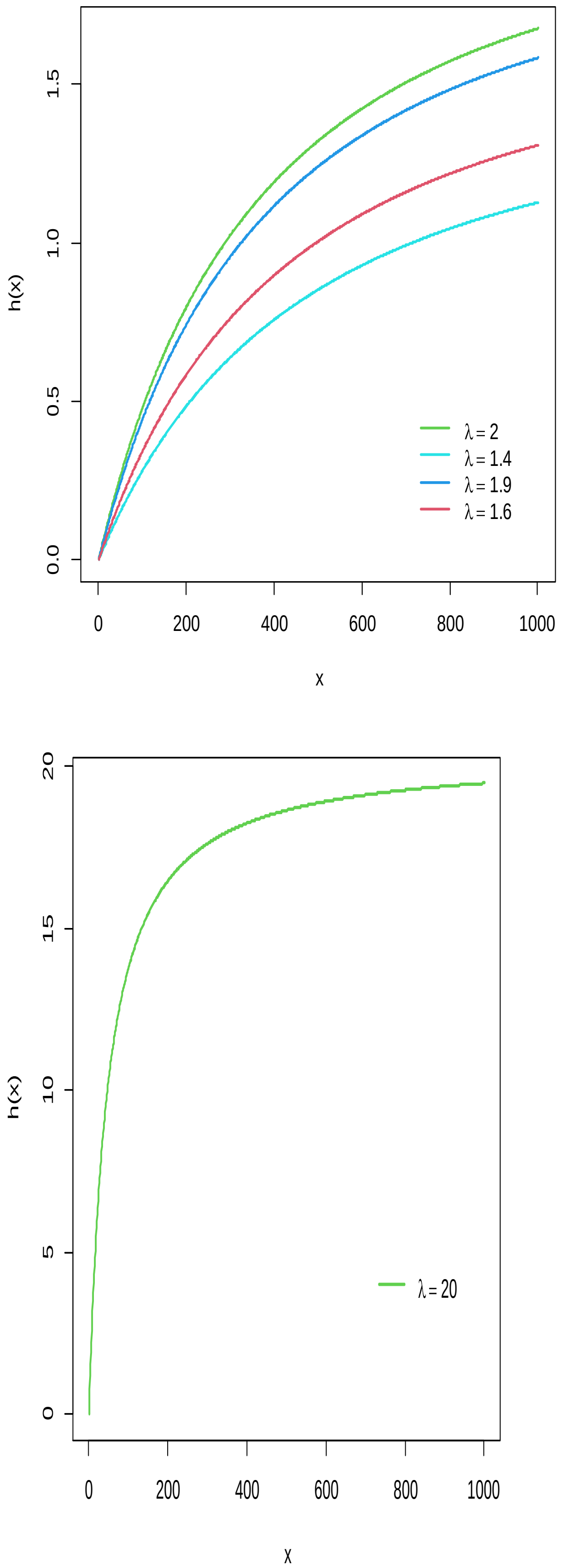

Fig.3: Plots of the hazard rate function of Bagui-Liu-Zhang distribution for different values of λ

4.0 Application of Bagui-Liu-Zhang Distribution

In this section, we illustrate the applicability of the Bagiu-Liu-Zhang distribution using a real data set. The data which constitute remission times of 128 bladder cancer patients were originally reported by Lee and Wang (2003) and are given below:

0.08, 2.09, 2.73, 3.48, 4.87, 6.94,8.66, 13.11, 23.63, 0.20, 2.22, 3.52, 4.98, 6.99, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06,7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82,0.51, 2.54,3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15,2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96,36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 15.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63,17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87,11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98,19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28,2.02, 3.36, 6.93, 8.65, 12.63, 22.69.

The fit of Bagiu-Liu-Zhang distribution (BLZD) to the data is compared with that of the exponential distribution (ED) using criteria Akaike information criteria (AIC), Bayesian information criteria (BIC) and others as indicated in Table 1. Notably, S.E, L, A^*, W^* and K-S represent the standard error, estimated log-likelihood function, Anderson-Darling statistic, Cramér-von Mises statistic and Kolmogorov-Smirnov statistic respectively. When distributions are fitted to the data under consideration, the distribution with the smallest value of each statistic is considered to be the best for modelling the data. The parameter of each of the distributions is estimated by the maximum likelihood method. Table 1 contains the maximum likelihood estimates of the parameters of the two distributions and their corresponding standard errors. The estimated values of the goodness of fit statistics are equally presented in Table 1.

Table 1: Maximum Likelihood Estimates of the Distributions fitted to Remission Time Data and Values of Some Goodness of Fit Statistics

Model	BLZD	ED
Estimate	0.11881	0.10740
SE	0.0101	0.00964
-L	3.99.7991	400.6669
AIC	801.5981	803.3337
BIC	804.4184	806.1540
A^2	0.56948	0.77192
W^2	0.09070	0.12753
K-S	0.06922	0.08213

From Table 1, the Bagiu-Liu-Zhang distribution (BLZD) has the minimum value of each of AIC, BIC , HQIC , K-S , W* and A*. Therefore, we conclude that the Bagiu-Liu-Zhang distribution fits the data better than the exponential distribution.

5.0 Conclusion

We have studied properties of a one-parameter continuous distribution called Bagiu-Liu-Zhang distribution. In particular, we derived the cdf, raw moments, coefficients of skewness and kurtosis, reliability and hazard rate functions and maximum likelihood estimate of the parameter of the distribution. The distribution is unimodal and has a nondecreasing hazard rate function. It is also positively skewed and leptokurtic. The maximum likelihood estimator and method of moments estimator of the parameter are equal. The numerical results obtained show that the distribution is capable of providing better fits to certain data than the exponential distribution and several one-parameter continuous distributions in the literature.

5.0 References

Albazi, A. O, Aslam, M and Dobbah, S. A. (2020). A control chart for exponentially distributed characteristics using modified multiple dependent state sampling. *Mathematical Problems in Engineering*, https://doi.org/10.1155/2020/5682587

Bagiu, S., Liu, J & Zhang, S. (2020). Deriving mixture distributions through moment generating functions. *Journal of Statistical Theory and Applications*, 19, 3, pp. 383-390.

Casella, G & Berger, R. L. (2002). *Statistical Inference, Second Edition. Duxbury*, Pacific Grove, CA, USA.

Chong, Z. L, Mukherjee, A & Marozzi, M. (2021). Simultaneous monitoring of origin and scale of a shifted exponential process with known and estimated parameters. *Quality and Reliability Engineering International*, 37, pp. 242-261.

Karlis, D & Xekalaki, E. (2005). Mixed Poisson distributions. *International Statistical Review*, 73, pp. 35-58.

Lawless, J. F. (1977). Prediction interval for the two parameter exponential distribution. *Technometrics*, 19, 4, pp. 469-472.

Lee, E. T & Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis*, Third Edition. John Wiley & Sons, Inc., Hoboken, New Jersey, USA.

Noor, F, Aslam, M & Sultana, T. (2020). Bayesian estimation and application of shifted exponential mixture distribution. *Thailand Statistician*, 18, 3, pp. 354-372.

Santiago, E & Smith, J. (2013). Control charts based on the exponential distribution: adapting runs rules for t chart. *Quality Engineering*, 25, 2, pp. 85-96.

Villa, E. R & Esocobar, L. A. (2006). Using moment generating functions to derive mixture distributions. *The American Statistician*, 60, 1, pp. 75-80.

Consent for publication

Not Applicable.

Availability of data and materials

The publisher has the right to make the data public.

Competing interests

The authors declared no conflict of interest.

Funding

There is no source of external funding.

Authors. ' Contribution

This paper is a product of the concerted effort of all the authors. All the authors participated in the literature review and the derivation of the theoretical properties of the distribution. The first and last authors took care of the numerical illustration.