Bouchahed, Lahsen; Zeghdoudi, Halim

Article

A NEW AND UNIFIED APPROACH IN GENERALIZING THE LINDLEY'S DISTRIBUTION WITH APPLICATIONS

Statistics in Transition New Series

Provided in Cooperation with:
Polish Statistical Association

Suggested Citation: Bouchahed, Lahsen; Zeghdoudi, Halim (2018) : A NEW AND UNIFIED APPROACH IN GENERALIZING THE LINDLEY'S DISTRIBUTION WITH APPLICATIONS, Statistics in Transition New Series, ISSN 2450-0291, Exeley, New York, NY, Vol. 19, Iss. 1, pp. 61-74, https://doi.org/10.21307/stattrans-2018-004

STATISTICS IN TRANSITION new series, March 2018
Vol. 19, No. 1, pp. 61–74, DOI 10.21307/stattrans-2018-004

A NEW AND UNIFIED APPROACH IN GENERALIZING THE LINDLEY'S DISTRIBUTION WITH APPLICATIONS

Lahsen Bouchahed[1], **Halim Zeghdoudi**[2]

ABSTRACT

This paper proposes a new family of continuous distributions with one extra shape parameter called the generalized Zeghdoudi distributions (GZD). We investigate the shapes of the density and hazard rate function. We derive explicit expressions for some of its mathematical quantities. Various statistical properties like stochastic ordering, moment method, maximum likelihood estimation, entropies and limiting distribution of extreme order statistics are established. We prove the flexibility of the new family by means of applications to several real data sets.

Key words: Lindley distribution, exponential distribution, Gamma distribution, stochastic ordering, maximum-likelihood estimation.

1. Introduction

Statistical models are very useful in describing and predicting real-world phenomena. Numerous extended distributions have been extensively used over the last decades for modelling data in several areas. Recent developments focus on defining new families that extend well-known distributions and at the same time provide great flexibility in modelling data in practice. Thus, many lifetime distributions for modeling lifetime data such as Lindley, exponential, Gamma, Weibull, Lognormal, Akash, Shanker, Sujatha, Amarendra distributions have been proposed in the statistical literature.

The probability density function of Lindley distribution is given by

$$f_1(x,\theta)=\frac{\theta^2}{\theta+1}(1+x)\exp(-\theta x);\; x>0,\; \theta>0$$

It has been generalized many times by researchers including Zakerzadeh and Dolati (2009), Bakouch et al. (2012), Shanker et al. (2013), Elbatal et al. (2013), Ghitany et al. (2013), Suigh et al. (2014), Abouamoh et al. (2015).

Shanker used a procedure based on certain mixtures of the gamma and exponential distributions to obtain three new distributions:
the Akash distribution, see (Shanker $2015a$), whose probability density function is given by

$$f_2(x,\theta)=\frac{\theta^3\left(1+x^2\right)}{\theta^2+1}\exp(-\theta x);\; x>0,\; \theta>0$$

[1]LaPS laboratory, Badji-Mokhtar University, Box 12, Annaba, 23000,ALGERIA. E-mail: lbouchahed@hotmail.com

[2]LaPS laboratory, Badji-Mokhtar University, Box 12, Annaba, 23000,ALGERIA. E-mail: halim.zeghdoudi@univ-annaba.dz

the Aradhana distribution, see (Shanker $2016a$), whose probability density function is given by

$$f_2(x,\theta) = \frac{\theta^3 (1+x)^2}{\theta^2+2\theta+2} \exp(-\theta x);\ x>0,\ \theta>0$$

the Sujatha distribution, see (Shanker $2016b$), whose probability density function is given by

$$f_3(x,\theta) = \frac{\theta^3 \left(1+x+x^2\right)}{\theta^2+\theta+2} \exp(-\theta x);\ x>0,\ \theta>0$$

the Amarendra distribution, see (Shanker $2016c$), whose probability density function is given by

$$f_4(x,\theta) = \frac{\theta^4 \left(1+x+x^2+x^3\right)}{\theta^3+\theta^2+2\theta+6} \exp(-\theta x);\ x>0,\ \theta>0.$$

the Devya distribution, see (Shanker $2016d$), whose probability density function is given by

$$f_5(x,\theta) = \frac{\theta^5 \left(1+x+x^2+x^3+x^4\right)}{\theta^4+\theta^3+2\theta^2+6\theta+24} \exp(-\theta x);\ x>0,\ \theta>0.$$

the Shambhu distribution, see (Shanker $2016e$), whose probability density function is given by

$$f_6(x,\theta) = \frac{\theta^6 \left(1+x+x^2+x^3+x^4+x^5\right)}{\theta^5+\theta^4+2\theta^3+6\theta^2+24\theta+120} \exp(-\theta x);\ x>0,\ \theta>0.$$

In this paper we introduce a new one-parameter family of continuous distributions by considering a polynomial exponential family, which contains Akash, Sujatha, Amarendra, Devya and Shambhu distributions as particular cases.

The rest of the paper is outlined as follows. Sections (2-10) are devoted to present various statistical properties like stochastic ordering, moment method, maximum likelihood estimation, entropies, and limiting distribution of extreme order statistics are established. In section 11 we present the new distribution called Zeghdoudi distribution, and in section 12 we give numerical examples.

We derive explicit expressions for some of its mathematical quantities. Various statistical properties like stochastic ordering, moment method, maximum likelihood estimation, entropies and limiting distribution of extreme order statistics are established. We prove the flexibility of the new family by means of applications to several real data sets.

Basic Theory

Suppose that X is random variable taking values in $]0,\infty)$, and that the distribution of X depends on an unspecified parameter θ taking values in $]0,\infty)$. So, the distribution of X might be absolutely continuous or discrete. In both cases, let f_θ be the probability distribution function with respect to the Lebesgue measure or to the counting measure on a countable set including discontinuity jumps of f_θ.

The distribution of X is a one-parameter polynomial exponential family and the probability density function can be written as

$$f_{GZD}(x;\theta) = h(\theta)\, p(x) \exp(-\theta x), \qquad \theta, x > 0$$

where $h(\theta)$ is real-valued functions on $]0,\infty)$, and where $p(x) = \sum_{k=0}^{n} a_k x^k$ and $a_n \neq 0$, is polynomial functions on $]0,\infty)$. Moreover, k is a positive integer and a_k is positive real number with $a_n \neq 0$. We can check immediately:

1) It is non-negative for $x > 0$;

2) $P[a < x < b] = \int_a^b f_\theta(x)\, dx$;

3) $\int_0^\infty f_\theta(x)\, dx = 1$ for $h(\theta) = \dfrac{1}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}}$.

Now, the probability density function of Generalized Zeghdoudi Distribution X is:

$$f_{GZD}(x;\theta) = \frac{\sum_{k=0}^{n} a_k x^k \exp(-\theta x)}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}}, \qquad \theta, x > 0 \tag{1}$$

Examples and Special Cases

Many of the special distributions studied in this work are general exponential families, at least with respect to some of their parameters. On the other hand, most commonly, a parametric family fails to be a general exponential family because the support set depends on the parameter.

2. General one-parameter distribution and some properties

In this section, we give the general one-parameter distribution and study its properties.

The first and second derivatives of $f_{GZD,\theta}(x)$

$$\frac{d}{dx} f_{GZD}(x;\theta) = \frac{\left[(a_1 - \theta a_0) + \ldots + (n a_n - \theta a_{n-1}) x^{n-1} + a_n x^n\right] \exp(-\theta x)}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}} = 0$$

gives $x_1, x_{2,\ldots,} x_n$ solutions.

We can find easily the cumulative distribution function (c.d.f) of the general one-parameter distribution:

$$F_{GZD}(x) = 1 - \frac{\sum_{k=0}^{n} \dfrac{a_k \Gamma(k+1, x\theta)}{\theta^{k+1}}}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}}; x, \theta > 0 \tag{2}$$

2.1. Survival and hazard rate function

Let

$$S_{GZD}(x) = 1 - F_{GZD}(x) = \frac{\sum_{k=0}^{n} \frac{a_k \Gamma(k+1, x\theta)}{\theta^{k+1}}}{\sum_{k=0}^{n} a_k \frac{k!}{\theta^{k+1}}}; x, \theta > 0$$

and

$$h_{GZD}(x) = \frac{f_{GZD}(x)}{1 - F_{GZD}(x)} = \frac{\sum_{k=0}^{n} a_k x^k \exp(-\theta x)}{\sum_{k=0}^{n} \frac{a_k \Gamma(k+1, x\theta)}{\theta^{k+1}}}$$

be the survival and hazard rate function, respectively.

Proposition1. *Let* $h_\theta(x)$ *be the hazard rate function of* X. *Then* $h_\theta(x)$ *is increasing for* $\sum_0^m (k+1)(m-2k) a_{m-k} a_{k+1} \geq 0,\ m = 0, ..., 2n-1$

Proof. According to Glaser (1980) and from the density function (2) we have

$$\rho(x) = -\frac{f'_{GZD}(x;\theta)}{f_{GZD}(x;\theta)} = -\frac{\sum_{k=1}^{n} k a_k x^{k-1}}{\sum_{k=0}^{n} a_k x^k} + \theta$$

After simple computations we obtain

$$\rho'(x) = \frac{\sum_{m=0}^{2n-1} \sum_{k=0}^{m} (k+1)(m-2k) a_{m-k} a_{k+1} x^{m-1}}{\left(\sum_{k=0}^{n} a_k x^k\right)^2}.$$

Which implies that $h_\theta(x)$ is increasing *for* $\sum_{k=0}^{m} (k+1)(m-2k) a_{m-k} a_{k+1} \geq 0,\ m = 0, ..., 2n-1$

3. Moments and related measures

The kth moment about the origin of the GZD is :

$$\mathbb{E}(X^i) = \frac{\sum_{k=0}^{n} \frac{a_k}{\theta^{k+i+1}} (k+i)!}{\sum_{k=0}^{n} a_k \frac{k!}{\theta^{k+1}}}, i = 1, 2, ...$$

Remark 2 The kth moment about the origin of the Lindley distribution is

$$\mathbb{E}(X^i) = \frac{i!\,(\theta + i + 1)}{\theta^i (\theta + 1)}$$

Corollary 1. *Let* $X \sim GZD(\theta)$, *the mean of* X *is:*

$$\mathbb{E}(X) = \frac{\sum_{k=0}^{n} \frac{a_k}{\theta^{k+2}} (k+1)!}{\sum_{k=0}^{n} a_k \frac{k!}{\theta^{k+1}}}.$$

Theorem 1. *Let* $X \sim GZD(\theta)$, me *=median*(X) *and* $\mu = E(X)$. *Then* $me < \mu$

Proof. According to the increasingness of $F(x)$ for all x and θ,

$$F_{GZD}(me) = \frac{1}{2}$$

and

$$F_{GZD}(\mu) = 1 - h(\theta)\sum_{k=0}^{n} \frac{a_k \Gamma\left(k+1, \theta h(\theta)\sum_{k=0}^{n} \frac{a_k}{\theta^{k+2}}(k+1)!\right)}{\theta^{k+1}}$$

Note that $\frac{1}{2} < F(\mu) < 1$. It is easy to check that $F(me) < F(\mu)$. To this end, we have $me < \mu$.■.

The coefficients of variation γ , skewness and kurtosis of the GZD have been obtained as

$$\begin{aligned} \gamma &= \frac{\sqrt{Var(X)}}{\mathbb{E}(X)} \\ skewness &= \frac{\mathbb{E}(X^3)}{(Var(X))^{\frac{3}{2}}} \\ kurtosis &= \frac{\mathbb{E}(X^4)}{(Var(X))^2}. \end{aligned}$$

4. Estimation of parameter

Let $X_1, ..., X_n$ be a random sample of GZD. The ln-likelihood function, $\ln l(x_i;\theta)$ is given by :

$$\ln l(x_i;\theta) = n\ln h(\theta) + \sum_{i=1}^{n} \ln\left(\sum_{k=0}^{m} a_k x_i^k\right) - \theta\sum_{i=1}^{n} x_i.$$

The derivative of $\ln l(x_i;\theta)$ with respect to θ is :

$$\frac{d\ln l(x_i;\theta)}{d\theta} = \frac{n\dot{h}(\theta)}{h(\theta)} - \sum_{i=1}^{n} x_i.$$

From the Zeghdoudi distribution (2), the method of moments (MoM) and the maximum likelihood (ML) estimators of the parameter θ are the same and it can be obtained by solving the following non-linear equation

$$\frac{\dot{h}(\theta)}{h(\theta)} - \bar{x} = 0, \text{ where } \dot{h}(\theta) = \frac{dh(\theta)}{d\theta} \tag{4}$$

The equation to be solved is

$$\sum_{k=0}^{m} \frac{a_k k!}{\theta^k}\left((k+1)-\bar{x}\theta\right)=0 \tag{5}$$

Note that we can solve the equation (5) exactly for $m \leq 4$ and for $m \geq 5$ the equation (5) is to be solved numerically.

Special cases

For $\mathbf{m}=\mathbf{0}$, we have $\hat{\theta}_{MV}=\frac{1}{\bar{x}}$

For $\mathbf{m}=\mathbf{1}$, we have $\hat{\theta}_{MV}=\frac{1}{2xa_0}\left(a_0-xa_1+\sqrt{x^2a_1^2+a_0^2+6xa_0a_1}\right)$

For $\mathbf{m}=\mathbf{2}$, $\hat{\theta}_{MV}$ is one of the two solutions :

$$\left\{\frac{\left(-a_1+\sqrt{x^2a_2^2+a_1^2-6a_0a_2+4xa_1a_2}+xa_2\right)}{a_0-xa_1},-\frac{\left(a_1+\sqrt{x^2a_2^2+a_1^2-6a_0a_2+4xa_1a_2}-xa_2\right)}{a_0-xa_1}\right\}$$

For $\mathbf{m}=\mathbf{3}$ and $\mathbf{m}=\mathbf{4}$, we can solve exactly equation (5) using methods such as Cardan and Ferrari method

For $\mathbf{m}\geq\mathbf{5}$, according to Galois theorem, there is no general method to solve exactly equation (5).

5. Stochastic orders

Definition 1. *Consider two random variables X and Y. Then X is said to be smaller than Y in the: a) Stochastic order ($X \prec_s Y$), if $F_X(t) \geq F_Y(t)$, $\forall t$.*
b) Convex order ($X \leq_{cx} Y$), if for all convex functions ϕ and provided expectation exist, $E[\phi(X)] \leq E[\phi(Y)]$.
c) Hazard rate order ($X \prec_{hr} Y$), if $h_X(t) \geq h_Y(t)$, $\forall t$.
d) Likelihood ratio order ($X \prec_{lr} Y$), if $\frac{f_X(t)}{f_Y(t)}$ is decreasing in t.

Remark 3 *Likelihood ratio order⇒Hazard rate order⇒Stochastic order.*
If $E[X]=E[Y]$, then Convex order⇔Stochastic order.

Theorem 4. *Let $X_i \sim GZD(\theta_i), i=1,2$ be two random variables. If $\theta_1 \geq \theta_2$, then $X_1 \prec_{lr} X_2, X_1 \prec_{hr} X_2, X_1 \prec_s X_2$ and $X_1 \leq_{cx} X_2$.*

Proof. We have

$$\frac{f_{X_1}(t)}{f_{X_2}(t)}=\frac{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta_2^{k+1}}}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta_1^{k+1}}}e^{-(\theta_1-\theta_2)t}.$$

For simplification, we use $\ln\left(\frac{f_{X_1}(t)}{f_{X_2}(t)}\right)$. Now, we can find

$$\frac{d}{dt}\ln\left(\frac{f_{X_1}(t)}{f_{X_2}(t)}\right)=-\left(\theta_1-\theta_2\right)$$

To this end, if $\theta_1 \geq \theta_2$, we have $\frac{d}{dt}\ln\left(\frac{f_{X_1}(t)}{f_{X_2}(t)}\right) \leq 0$. This means that $X_1 \prec_{lr} X_2$. Also, according to *Remark 3* the theorem is proved.

6. Mean Deviations

These are two mean deviation: about the mean and about the median, defined as $MD_1 = \int_0^\infty |x-\mu| f(x)\,dx$ and $MD_2 = \int_0^\infty |x-me| f(x)\,dx$ respectively, where $\mu = E(X)$

and me =Median(X). The measures MD_1 and MD_2 can be computed using the following simplified formulas

$$MD_1 = 2\mu F(\mu) - 2\int_0^\mu x f(x)\,dx$$

$$MD_2 = \mu - 2\int_0^{me} x f(x)\,dx$$

7. Extreme domain of attraction

As to the extreme value stability, the cdf F_{GZD} is in the Gumbel extreme value domain of attraction, that is, there exist two sequences $(a_n)_{n\geq 0}$ and $(b_n)_{n\geq 0}$ of real numbers such that for any $x \in \mathbb{R}$, we have

$$\lim_{n\to+\infty} \mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = \lim_{n\to+\infty} F_{GZD}(a_n x + b_n)^n = \exp(-\exp(-x)).$$

This follows from Formula 1.2.4 in theorem 1.2.1 (De Haan and Ferreira (2006)) since we have

$$\begin{aligned}
\lim_{t\to\infty} \frac{1 - F_{GZD}(t + xf(t))}{1 - F_{GZD}(t)} &= \lim_{t\to\infty} \frac{f_{GZD}(t + xf(t))}{f_{GZD}(t)} \\
&= \lim_{t\to\infty} \frac{\sum_{k=0}^{n} a_k (xf(t)+t)^k e^{(-\theta(xf(t)+t))}}{\sum_{k=0}^{n} a_k t^k e^{(-\theta t)}} = \exp(-x),
\end{aligned}$$

(such formula is called Γ-variation). Then, F_{GZD} lies in the Gumbel extreme domain of attraction. In his case, $f(t) = \frac{1}{\theta}$. So, for (as in the invoked theorem) $a_n = f(F_{GZD}^{-1}(1-1/n)) = \frac{1}{\theta}$ and $b_n = F_{GZD}^{-1}(1-1/n)$, we have

$$\lim_{n\to+\infty} F_{GZD}(a_n x + b_n)^n = \exp(-\exp(-x))$$

8. Estimation of the Stress-Strength Parameter

The stress-strength parameter (R) plays an important role in the reliability analysis as it measures the system performance. Moreover, R provides the probability of a system failure, the system fails whenever the applied stress is greater than its strength, i.e. $R = P(X > Y)$. Here $X \sim GZD(\theta_1)$ denotes the strength of a system subject to stress Y, and $Y \sim GZD(\theta_2)$, X and Y are independent of each other. In our case, the stress-strength parameter R is given by

$$\begin{aligned} R &= P(X > Y) = \int_0^{\infty} S_X(y) f_Y(y)\, dy \\ &= \frac{\int_0^{\infty} \sum_{k=0}^{n} \dfrac{a_k \Gamma(k+1, y\theta_1)}{\theta_1^{k+1}} \sum_{k=0}^{n} a_k y^k \exp(-\theta_2 y)\, dy}{\left(\sum_{k=0}^{n} a_k \dfrac{k!}{\theta_1^{k+1}}\right)\left(\sum_{k=0}^{n} a_k \dfrac{k!}{\theta_2^{k+1}}\right)} \end{aligned}$$

9. Lorenz curve

The Lorenz curve is often used to characterize income and wealth distributions. The Lorenz curve for a positive random variable X is defined as the graph of the ratio

$$L(F(x)) = \frac{E(X|X \leq x)F(x)}{E(X)}$$

against $F(x)$ with the properties $L(p) \leq p, L(0) = 0$ and $L(1) = 1$. If X represents annual income, $L(p)$ is the proportion of total income that accrues to individuals having the $100\ p\%$ lowest incomes. If all individuals earn the same income then $L(p) = p$ for all p. The area between the line $L(p) = p$ and the Lorenz curve may be regarded as a measure of inequality of income, or more generally, of the variability of X. For the exponential distribution, it is well known that the Lorenz curve is given by

$$L(p) = p\{p + (1-p)log(1-p)\}.$$

For the GZ distribution in (3),

$$E(X|X \leq x)F_{GZD}(x) = \frac{\sum_{k=0}^{n} \frac{a_k}{\theta^{k+2}}(k+1)!}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}} \left(1 - \frac{\sum_{k=0}^{n} \dfrac{a_k \Gamma(k+1, x\theta)}{\theta^{k+1}}}{\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}}\right)$$

10. Entropies

It is well known that entropy and information can be considered as measures of uncertainty of probability distribution. However, there are many relationships established on the basis of the properties of entropy.

An entropy of a random variable X is a measure of variation of the uncertainty. Rényi entropy is defined by

$$J(\gamma) = \frac{1}{1-\gamma} log \left\{ \int f^{\gamma}(x)\, dx \right\}$$

where $\gamma > 0$ and $\gamma \neq 1$. For the GZ distribution in (1), note that, for γ integer we have

$$\begin{aligned}
\int f_{GZD}^{\gamma}(x)\, dx &= \frac{\int \left(\sum_{k=0}^{n} a_k x^k\right)^{\gamma} \exp(-\theta\gamma x)\, dx}{\left(\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}\right)^{\gamma}} \\
&= \frac{\sum_{k=0}^{n} b_k(\gamma) \int x^{k\gamma} \exp(-\theta\gamma x)\, dx}{\left(\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}\right)^{\gamma}}
\end{aligned}$$

where, $\int x^{k\gamma} \exp(-\theta\gamma x)\, dx = -\frac{1}{\theta\gamma(\theta\gamma)^{k\gamma}} \Gamma(k\gamma+1, x\theta\gamma)$ and $b_k(\gamma)$ in function the a_k and γ.

Now, the Rényi entropy as given by

$$J(\gamma) = \frac{1}{1-\gamma} \log \left[\frac{\frac{(k\gamma)!}{\theta\gamma(\theta\gamma)^{k\gamma}} \sum_{k=0}^{n} b_k(\gamma)}{\left(\sum_{k=0}^{n} a_k \dfrac{k!}{\theta^{k+1}}\right)^{\gamma}} \right]$$

if $\gamma \to \infty$, we find Shannon entropy.

11. Zeghdoudi distribution(ZD) and immediate properties

In this section we will introduce a new distribution called Zeghdoudi distribution(ZD) (see Zeghdoudi and Messadia 2018), which is a member of the new family.

The density function of $X \rightsquigarrow$ Zeghdoudi distribution is given by:

$$f_{ZD}(x;\theta) = \begin{cases} \frac{\theta^3 x(1+x)e^{-\theta x}}{\theta+2} & x,\theta > 0 \\ 0, & \text{otherwise.} \end{cases}$$

We note that ZD distribution is a member of the new family where $n = 2, a_0 = 0, a_1 = a_2 = 1$ using formula (1). Therefore, the mode of ZD is given by

$$\text{mode}(X) = \frac{-\theta+2+\sqrt{\theta^2+4}}{2\theta} \quad \text{for } \theta > 0$$

We can find easily the cumulative distribution function(c.d.f) of the ZD :

$$F_{ZD}(x) = 1 - \left(\frac{\theta^2 x^2 + \theta(\theta+2)x + \theta + 2}{\theta+2}\right) e^{-\theta x}; x > 0, \theta > 0$$

From the Zeghdoudi distribution, the method of moments (MoM) and the maximum likelihood (ML) estimators of the parameter θ are the same and can be obtained by solving the following non-linear equation

$$\frac{3}{\theta} - \frac{1}{\theta+2} - \bar{x} = 0$$

$$\hat{\theta}_M = \hat{\theta}_{ML} = \left\{ \frac{1}{\bar{x}} \left(-\bar{x} + \sqrt{4\bar{x} + \bar{x}^2 + 1} + 1 \right) \right\}$$

We can show that the estimator of θ is positively biased.

12. Simulation and Goodness of Fit

12.1. Simulation study

We can see that the equation $F(x) = u$, where u is an observation from the uniform distribution on $(0,1)$, cannot be solved explicitly in x (cannot use **lambert W function in the case** $k \geq 2$), the inversion method for generating random data from the GZ distribution fails. However, we can use the fact that the GZ distribution is a mixture of gamma (k,θ) and gamma $(k+1,\theta)$ distributions:

$$f_{GZD}(x;\theta) = p(\theta)\ gamma(k,\theta) + (1-p(\theta))\, gamma(k+1,\theta), 0 < p(\theta) < 1$$

In this subsection, we investigate the behaviour of the ML estimators for a finite sample size (n). A simulation study consisting of the following steps is being carried out $N = 10000$ times for selected values of $(\theta;n)$, where $\theta = 0.01, 0.1, 0.5, 1, 3, 10$ and $n = 10, 30, 50$.

- Generate $U_i \sim Uniform(0,1)$, $i = 1, \ldots n$.
- Generate $Y_i \sim Gamma(k,\theta)$, $i = 1, \ldots n$.
- Generate $Z_i \sim Gamma(k+1,\theta)$, $i = 1, \ldots n$.
- If $U_i \leq p(\theta)$, then set X_i = Y_i, otherwise, set X_i = Y_i, $i = 1, \ldots n$,

$$average\ bais\,(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\theta}_i - \theta \right)$$

and the average square error.

$$MSE\,(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\theta}_i - \theta \right)^2, i = 1, \ldots N.$$

12.2. Applications

Example 1

Tables 1 and 2 represent the Data of survival times (in months) of (94,91) guinea individus infected with Ebola virus.

Table 1 Chi-Square Values for Lindley and Zeghdoudi distributions

Survival time $m = 3.17$	Obs freq	LD $\hat{\theta} = 0.522$	ZD $\hat{\theta} = 0.852$
$[0,2[$	32	38.217	33.252
$[2,4[$	35	28.16	32.366
$[4,6[$	17	15.089	17.799
$[6,8[$	7	7.33	7. 133
$[8,10]$	3	3.152	2.418
Total	94	94	94
χ^2	-	2.9244	**0.40020**

Table 2 Chi-Square Values for Lindley and Zeghdoudi distributions

Survival time $m = 3$	Obs freq	LD $\hat{\theta} = 0.548$	ZD $\hat{\theta} = 0.896$
$[0,2[$	35	39.100	34.56
$[2,4[$	32	27.390	31.497
$[4,6[$	16	13.92	16.206
$[6,8[$	6	6.247 5	6.069 7
$[8,10]$	2	2.618 9	1.921 8
Total	91	91	91
χ^2	-	1.616 5	**0.01995**

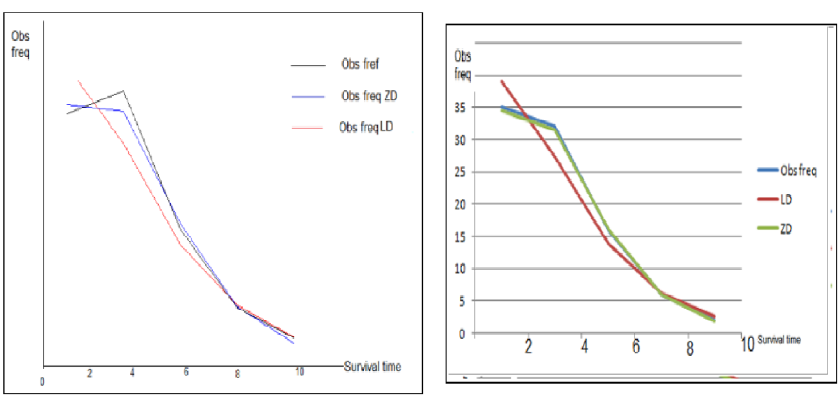

Figure 1: Plots of the density function of LD and ZD

Example 2

Now, we compare one-parameter (Aradhana, Akash, Shanker, Amarendra, Devya, Shambhu) distributions see Shanker (2015a, 2015b, 2016a, 2016b, 2016c, 2016d, 2016e) and two-parameter (Gamma, Weibull, Lognormal) distributions with Zeghdoudi distribution (see tables 3 and 4).

Table 3 Comparison of ZD with one-parameter distributions

Data	Distribution	θ	log-likelihood	Kolmogrov-Smirnov
data 1	ZD	1.365	**-52.4**	**0.292**
$n = 20$	Aradhana	1.123	-56.4	0.302
$m = 1.9$	Akash	1.157	-59.5	0.320
$s = 0.704$	Shanker	0.804	-59.7	0.315
	Amarendra	1.481	-55.64	0.286
	Devya	1.842	-54.50	0.268
	Shambhu	2.215	-53.9	0.254

Data	Distribution	θ	log-likelihood	Kolmogrov-Smirnov
data 2	ZD	0.095	**-239.7**	**0.251**
$n = 25$	Aradhana	0.094	-242.2	0.274
$m = 30.811$	Akash	0.097	-240.7	0.266
$s = 7.253$	Shanker	0.063	-252.3	0.326
	Amarendra	1.481	-233.41	0.225
	Devya	1.842	-227.68	0.193
	Shambhu	2.215	-223.40	0.167

Table 4 Comparison of ZD with two-parameter distributions

Data	Distribution	β	θ	log-likelihood	Kolmogrov-Smirnov
data 3	ZD	—	0.107	**-58.67**	**0.081**
$n = 15$	Gamma	1.442	0.052	-64.197	0.102
$m = 27.54$	Weibull	1.306	0.034	-64.026	0.450
$s = 20.06$	Lognormal	1.061	2.931	-65.626	0.163
data 4	ZD	—	0.0167	**-142.32**	**0.108**
$n = 25$	Gamma	1.794	0.010	-152.371	0.135
$m = 178.32$	Weibull	1.414	0.005	-152.440	0.697
$s = 131.09$	Lognormal	0.891	4.880	-154.092	0.155

According to tables 1, 2, 3, 4 and figures 1, 2, we can observe that Zeghdoudi distribution provide smallest -LL and K-S values as compared to one-parameter (Aradhana, Akash, Shanker, Amarendra, Devya, Shambhu) distributions, and two-parameter (Gamma, Weibull, Lognormal), and hence best fits the data among all the models considered.

13. Conclusions

In this work we propose a one-parameter family GZD. Several properties have been discussed: moments, cumulants, characteristic function, failure rate function, stochastic ordering, the maximum likelihood method and the method of moments. The LD does not provide enough flexibility for analyzing and modelling different types of lifetime data and survival analysis. But GZD is flexible, simple and easy to handle. Two real data sets are analyzed using the new distribution and are compared with five immediate sub-models mentioned above in addition to other distributions (Lindley, Exponential, Gamma, Weibull, Lognormal distributions). The results of the comparisons confirm the goodness of fit of GZ distribution. We hope our new distribution family might attract wider sets of applications in lifetime data reliability analysis and actuarial sciences. For future studies, we can go more generally by using $p(x;\theta)$. Also, we can explain the derivation of posterior distributions for the GZ distribution under Linex loss functions and squared error using non-informative and informative priors (the extension of Jeffreys and Inverted Gamma priors) respectively.

Annexe

Data set 1: represents the lifetime data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by Gross and Clark (1975, P. 105). The data are as follows: 1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3.0, 1.7, 2.3, 1.6, 2.0

Data set 2: is the strength data of glass of the aircraft window reported by Fuller et al. (1994): 18.83, 20.80, 21.657, 23.03, 23.23, 24.05, 24.321, 25.50, 25.52, 25.80, 26.69, 26.77, 26.78, 27.05, 27.67, 29.90, 31.11, 33.20, 33.73, 33.76, 33.89, 34.76, 35.75, 35.91, 36.98, 37.08, 37.09, 39.58, 44.045, 45.29, 45.381

Acknowledgment

The authors acknowledge Editor-in-Chief, Pr. Patryk Barszcz, of this journal for the constant encouragement to finalize the paper. Further, the authors acknowledge profound thanks to anonymous referees for giving critical comments, which have immensely improved the presentation of the paper. Also, this work was financed by European Mathematical Society.

REFERENCES

ASGHARZADEH, A., BAKOUCH, H, S., ESMAEILI, L., (2013). Pareto Poisson-Lindley distribution and its application, Journal of Applied Statistics, pp. 1–18.

FULLER, E.J., FRIEMAN, S., QUINN, J., QUINN, G., CARTER, W. Fracture mechanics approach to the design of glass aircraft windows: A case study, SPIE Proc 2286, pp. 419–430 (1994).

GHITANY, M, E., AL-MUTAIRI, D, K., NADARAJAH, S., (2008a). Zero-truncated Poisson-Lindley distribution and its application, Math. Comput. Simulation, 79, pp. 279–287.

GHITANY, M, E., ATIEH, B., NADARAJAH, S., (2008b). Lindley distribution and its applications. Mathematics and Computers in Simulation, 78, pp. 493–506.

GROSS, A,J., CLARK, V, A.,(1975). Survival Distributions: Reliability Applications in the Biometrical Sciences, John Wiley, New York.

LAURENS DE HAAN, FERREIRA, A., (2006). Extreme value theory: An introduction, Springer.

LAWLESS, J, F., (2003). Statistical models and methods for lifetime data, Wiley, New York.

LEADBETTER,M,R., LINDGREN, G., ROOTZén, H., (1987). Extremes and Related Properties of Random Sequences and Processes, Springer Verlag, New York.

LINDLEY, D, V., (1958). Fiducial distributions and Bayes' theorem, Journal of the Royal Society, series B, 20, pp. 102–107.

SANKARAN, M., (1970). The discrete Poisson-Lindley distribution, Biometrics, 26, pp. 145–149.

SHANKER, R., SHARMA, S., SHANKER, R., (2013). A two-parameter Lindley distribution for modeling waiting and survival times data. Applied Mathematics, Vol. 4, pp. 363–368.

SHANKER, R., (2015 a). Akash distribution and Its Applications. International Journal of Probability and Statistics, 4 (3), pp. 65–75.

SHANKER, R., (2015 b). Shanker distribution and Its Applications, International Journal of Statistics and Applications, 5 (6), pp. 338–348.

SHANKER, R., (2016 a). Aradhana distribution and Its Applications, International Journal of Statistics and Applications, 6 (1), pp. 23–34.

SHANKER, R., (2016 b). Sujatha distribution and Its Applications, Statistics in Transition new series, 17 (3), pp. 391–410.

SHANKER, R., (2016 c). Amarendra distribution and Its Applications, American Journal of Mathematics and Statistics, 6 (1), pp. 44–56.

SHANKER, R., (2016 d). Devya distribution and Its Applications, International Journal of Statistics and Applications, 6 (4), pp. 189–202.

SHANKER, R., (2016 e). Shambhu Distribution and Its Applications, International Journal of Probability and Statistics, 5 (2), pp. 48–63.

ZAKERZADAH, H., DOLATI, A., (2010). Generalized Lindley distribution, J. Math. Ext, 3(2), pp. 13–25.

ZEGHDOUDI, H., NEDJAR, S., (2016a). Gamma Lindley distribution and its application, Journal of Applied Probability and Statistics, Vol. 11 (1), 129–138.

ZEGHDOUDI, H., NEDJAR, S., (2016b). On gamma Lindley distribution: Properties and Simulations, Journal of Computational and Applied Mathematics, Vol 298, pp 167–174.

ZEGHDOUDI, H., NEDJAR, S., (2016c). A pseudo Lindley distribution and its application, Journal Afrika Statistika, Vol. 11 (1), pp. 923–932.

ZEGHDOUDI, H., MESSADIA,H., (2018). Zeghdoudi Distribution and its Applications. International Journal of Computing Science and Mathematics, In Press.