Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 11(1), 2016, pages 923–932.
DOI: http://dx.doi.org/10.16929/as/2016.923.83

A pseudo Lindley distribution and its application

Halim Zeghdoudi* **and Sihem Nedjar**

LaPS Laboratory, Badji-Mokhtar University, Box 12, Annaba, 23000, Algeria

Received February 6, 2016; Accepted June 21, 2016

Abstract. In this paper, we introduce a new distribution named as the Pseudo Lindley Distribution ($PsLD$) as a generalization of the Lindley distribution (LD). A full and detailed description are provided in terms of moments, cumulates, characteristic function, failure, rate function, stochastic ordering, distributions of sums, and parameters estimation. Simulations studies and data driven applications are also reported.

Résumé. Dans cet article, nous introduisons une nouvelle distribution de probability dénommée Pseudo-Lindely-Distribution ($PsLD$) comme une géné- ralisation de celle de Lindley (LD). Une description complète s'en suit, par rapport aux moments, à la fonction caractéristique, à la function de répartition et de survie, aux ordres stochastiques au sein de la famille. De plus l'estimation des paramètres par la méthode des moments et du maximum de vraisemblance est abordée. Une partie réservée aux simulations et à des applications sur des données réelles montrent la souplesse de cette loi pour expliquer certaines données de survie par rapport à la distribution de Lindley et certaines de ces généralisations.

Key words: Lindley distribution; Exponential distribution; Gamma distribution; Stochastic ordering; Maximum-likelihood estimation.
AMS 2010 Mathematics Subject Classification : 60E05; 62H20.

1. Introduction

We are concerned with a two-parameters generalization of the one-parameter Lindley distribution (Lindley , 1958) defined below by its probability density function, depending on

* Corresponding author Halim Zeghdoudi : halim.zeghdoudi@univ-annaba.dz
Sihem Nedjar: nedjar.sihem@yahoo.com

the parameter $\theta > 0$,

$$f(x;\theta) = \begin{cases} \dfrac{\theta^2(1+x)e^{-\theta x}}{1+\theta} & x > 0 \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

This distribution has attracted the interest of many researchers and has been generalized several times by various authors. First, Sankaran (1970) used (1) when the parameter follows a Poison Law to derive their discrete Poisson-Lindley Distribution (PLD) with density function

$$f_{PLD}(x;\theta) = \frac{\theta^2(x+\theta+2)e^{-\theta x}}{(1+\theta)^x}, x = 0, 1, \ldots$$

where $\theta > 0$ is a positive parameter. Recently, Asgharzadeh *et al.* (2013), Ghitany *et al.* (2008a) and Ghitany *et al.* (2008b) introduced new distributions bounded to 1, called Zero-truncated poisson-Lindley and Pareto Poisson-Lindley distributions. Also, Nedjar and Zeghdoudi (2016) and Zeghdoudi and Nedjar (2016) introduced a new distribution, named *gamma-Lindley distribution*, based on mixtures of gamma $(2, \theta)$ and one-parameter Lindley distributions.

Since the Lindsley distribution and its derived form we just described depend on one parameter, they may lack of flexibility in statistical modelling of different types of lifetime data. This motivated us to find out generalizations of more parameters that are easily handled. Yet Zakerzadah and Dolati (2010) introduced a generalization of with three parameters. But it happens that their distribution is difficult to use and is not as flexible as one can whish.

The idea of this paper is to use a mixture of ordinary exponential random variables (rv's) and $Gamma(2, \theta)$ ones. The found distribution is characterized through a number of properties concerning its characteristics and parameters : probability density function (pdf), cumulative distribution (cdf), survival and hazard rate functions, moment generating function (mgf), mean, variance and stochastic orderings. Also relevant plots are given as illustrations.

Moment estimates are also discussed and statistical applications treating goodness of fit are provided.

The paper is organized as follows. In Section 2, we introduce a generalization of the Lindley distribution, that we name as the Pseudo-Lindley distribution ($PsLD$) and give immediate properties as the mode. Section 3 is devoted to the study of survival and the hazard functions. In Section 4 we focus on the moments including skewness and kirtosis and stochastic ordering. In Section 5, we are interested in parameters estimation using both the maximum likelyhood and the moment method. In this last section, simulation studies are reported and as well, are provided datadriven applications allowing comparisons between our new law with the original Lindley law and with others of it generalizations. We finish the paper with a conclusive section.

2. The Pseudo-Lindley distributions and immediate properties

In this section, we give the pseudo lindley distribution and study its properties. Let $Y_1 \sim Exp(\theta)$ and $Y_2 \sim \Gamma(2, \theta)$ be two independent random variables. For $\beta \geq 1$, we consider the mixture random variable variable $X = Y_1$ and $X = Y_2$ with probability $(\beta - 1)/\beta$ and $1/\beta$ respectively. The density function of X is given by:

$$f_{PsLD}(x;\theta,\beta) = \begin{cases} \frac{\theta(\beta-1+\theta x)e^{-\theta x}}{\beta}, & x,\theta,\beta \geq 1 \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Remark 1. Let us make these two remarks.

(1) If $\beta = \theta + 1$, this distribution is lindley distribution.
(2) If $\beta = 1$,, this distribution is a $\Gamma(2,\theta)$ distribution.

To find the mode, we may see that the first and second derivatives of $f_{PsLD}(x)$ are given by

$$\frac{d}{dx} f_{PsLD}(x) = \frac{\theta^2 (2 - \beta - \theta x) e^{-\theta x}}{\beta} = 0 \text{ gives } x = \frac{2-\beta}{\theta} \text{ and } \frac{d^2}{dx^2} f_{PsLD}(\hat{x}) < 0.$$

For $1 \leq \beta < 2$, $\hat{x} = (2-\beta)/\theta$ is the unique critical point which $f_{PsLD}(x;\theta,\beta)$ is maximum.

For $\beta \geq 2$,

$$\frac{d}{dx} f_{PsLD}(x;\theta,\beta) \leq 0,$$

and then then the density function $f_{PsLD}(x;\theta,\beta)$ is decreasing in x. Therefore, the mode of PsLD is given by

$$\text{mode}(X) = \begin{cases} \frac{2-\beta}{\theta} & \text{for } 1 \leq \beta < 2 \\ 0 & \text{otherwise.} \end{cases}$$

We can easily find the cumulative distribution function(cdf) of the PsLD :

$$F_{PsLD}(x) = 1 - \frac{(\beta + \theta x) e^{-\theta x}}{\beta}; x, \theta > 0 > 0, \beta \geq 1. \tag{3}$$

3. Pseudo lindley distribution(PsLD) and some properties

3.1. Survival and hazard rate function

These two functions are

$$S_{PsLD}(x) = 1 - F_{PsLD}(x) = \frac{(\beta + \theta x) e^{-\theta x}}{\beta}$$

and

$$h_{PsLD}(x) = \frac{f_{PsLD}(x)}{1 - F_{PsLD}(x)} = \frac{\theta (\beta + \theta x - 1)}{\beta + \theta x}$$

be the survival and hazard rate function, respectively.

Proposition 1. *Let $h_{PsLD}(x)$ be the hazard rate function of X. Then $h_{PsL}(x)$ is increasing.*

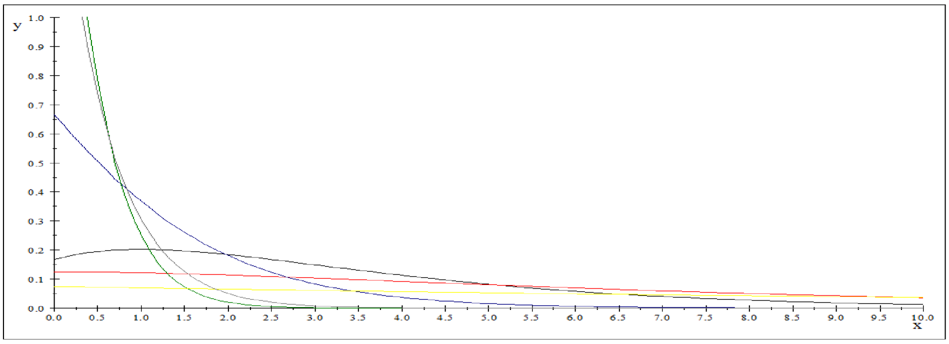

Fig. 1. Plots of the density function for some parameter values, black(0.5,1.5); red(0.25,2); bleu(1,3); green(3,3); yellow(0.1,4); gray(2,8)

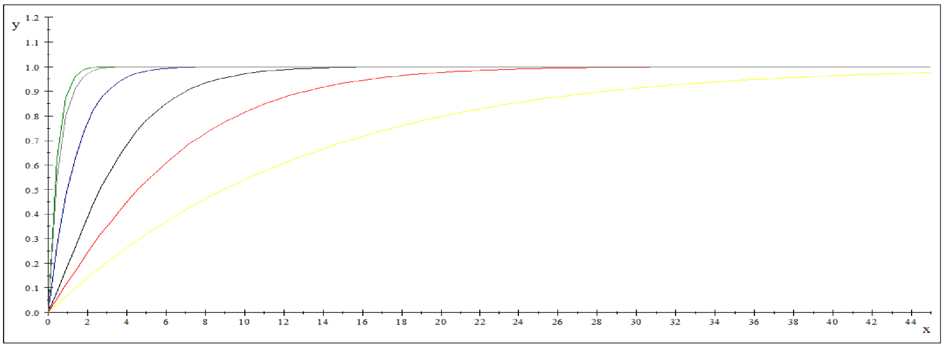

Fig. 2. Plots of the distribution function for some parameter values, black(0.5,1.5); red(0.25,2); bleu(1,3); green(3,3); yellow(0.1,4); gray(2,8)

Proof. It easy to check that

$$\frac{d}{dx}h_{PsLD}(x) = \frac{\theta^2}{(\beta+\theta x)^2} > 0.$$

□

4. Moments and related measures

The kth moment about the origin of the PsLD is :

$$\mathbb{E}(X^k) = \frac{k!\,(\beta+k)}{\theta^k \beta}, k = 1, 2, ...$$

Proposition 2. *Let $X_1, X_2, ..., X_n$ be n independent random variables from $PsLD(\beta, \theta)$ distribution. Then the moment generating function(mgf) of $S = \sum_{i=1}^{n} X_i$, is given by*

$$M_S(t) = \frac{\theta^n\left((1-\beta)\,t+\theta\beta\right)^n}{\beta^n\,(t-\theta)^{2n}} \tag{4}$$

and

$$M_X(t) = \mathbb{E}(e^{tX}) = \frac{\theta\left((1-\beta)\,t + \theta\beta\right)}{\beta\,(t-\theta)^2} \quad (5)$$

Remark 2. The kth moment about the origin of the exponential distribution is

$$\mathbb{E}(X^k) = \frac{k!}{\theta^k}$$

Furthermore, the moment generating function of X and S exists, that is $\mathbb{E}e^{tX} < +\infty$, if $t < \theta$.

Corollary 1. *Let $X \sim PsLD(\beta, \theta)$, the mean and variance of X are :*

$$\mathbb{E}(X) = \frac{\beta+1}{\theta\beta}, Var(X) = \frac{\beta^2 + 2\beta - 1}{\beta^2\theta^2}$$

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

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

$$F(M) = \begin{cases} 1 - \frac{2e^{-(2-\beta)}}{\beta} & \text{for } 1 \le \beta < 2 \\ 0 & \text{otherwise} \end{cases}, \; F(me) = \frac{1}{2}$$

and

$$F(\mu) = 1 - \frac{\left(\beta^2 + \beta + 1\right) e^{-\left(\frac{\beta+1}{\beta}\right)}}{\beta^2}$$

Note that $F(M)$ is a decreasing function for all $\beta \ge 1$. It easy to check that $F(M) < F(me) < F(\mu)$.To this end, we have $M < me < \mu$. □

The coefficients of variation γ, skewness and kurtosis of the PsLD are obtained as follows

$$\begin{aligned} \gamma &= \frac{\sqrt{Var(X)}}{\mathbb{E}(X)} = \frac{\sqrt{\beta^2 + 2\beta - 1}}{\beta + 1} \\ skewness &= \frac{\mathbb{E}(X^3)}{(Var(X))^{\frac{3}{2}}} = \frac{6\beta^2\,(\beta+3)}{(\beta^2 + 2\beta - 1)^{\frac{3}{2}}} \\ kurtosis &= \frac{\mathbb{E}(X^4)}{(Var(X))^2} = \frac{24\beta^3\,(\beta+4)}{(\beta^2 + 2\beta - 1)^2} \end{aligned}$$

Remark 3. All these expressions are independent of the parameter θ and depend upon the parameter β only.

4.1. Stochastic orders

Definition 1. Consider two random variables X and Y. We define these four stochastic orders.

(a) X is said to be smaller than Y in stochastic order, denoted

$$X \prec_s Y,$$

if and only if

$$F_X(t) \geq F_Y(t), \text{ for all } t.$$

b) X is said to be smaller than Y in convex order, denoted

$$X \leq_{cx} Y,$$

if for any convex function ϕ and provided expectations exist,

$$E[\phi(X)] \leq E[\phi(Y)]$$

c) X is said to be smaller than Y in hazard rate order, denoted

$$X \prec_{hr} Y$$

if and only if

$$h_X(t) \geq h_Y(t), \text{ for all} t$$

(d) X is said smaller than Y in likelihood ratio order, denoted

$$X \prec_{lr} Y,$$

if and only if $f_X(t)/f_Y(t)$ is decreasing in t.

Remark 4. Likelihood ratio order $\Rightarrow$ Hazard rate order $\Rightarrow$ Stochastic order.
If $E[X] = E[Y]$, then Convex order $\Leftrightarrow$ Stochastic order.

Theorem 2. *Let $X_i \sim PsLD(\beta_i, \theta_i), i = 1, 2$ be two random variables. If $\theta_1 = \theta_2$ and $\beta_1 \geq \beta_2$, 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

$$\frac{f_{X_1}(t)}{f_{X_2}(t)} = \frac{\theta_1\beta_2(\beta_1 - 1 + \theta_1 t)}{\theta_2\beta_1(\beta_2 - 1 + \theta_2 t)} e^{-(\theta_1 - \theta_2)t}.$$

For simplicity's sake, we use $\ln\left(\frac{f_{X_1}(t)}{f_{X_2}(t)}\right)$. Now, we can find

$$\frac{d}{dt}\ln\left(\frac{f_{X_1}(t)}{f_{X_2}(t)}\right) = -(\theta_1 - \theta_2) + \frac{\theta_1(\beta_2 - 1) - \theta_2(\beta_1 - 1)}{(\beta_1 - 1 + \theta_1 t)(\beta_2 - 1 + \theta_2 t)}$$

To this end, if $\theta_1 = \theta_2$ and $\beta_1 \geq \beta_2$, 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 4, the theorem is proved. □

5. Estimation of parameters

5.1. Maximum Likelihood Estimates (MLE)

Let $X_i \sim PsLD(\beta,\theta), i = \overline{1,n}$ be n random variables. The ln-likelihood function, $\ln l(x_i;\beta,\theta)$ is:

$$\ln l(x_i;\beta,\theta) = n\ln\theta - n\ln\beta + \sum_{i=1}^{n}\ln(\beta-1+\theta x_i) - \theta\sum_{i=1}^{n}x_i.$$

The derivatives of $Lnl(x_i;\beta,\theta)$ with respect to θ and β are:

$$\frac{\partial\ln l(x_i;\beta,\theta)}{\partial\theta} = \frac{n}{\theta} - \sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\left(\frac{x_i}{\beta-1+\theta x_i}\right) \tag{6}$$

$$\frac{\partial\ln l(x_i;\beta,\theta)}{\partial\beta}0 = \frac{-n}{\beta} + \sum_{i=1}^{n}\left(\frac{1}{\beta-1+\theta x_i}\right) \tag{7}$$

The two equations (6) and (7) cannot be solved directly, we must used the Fisher scoring method. We have

$$\begin{bmatrix} \frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\theta^2} & \frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\theta\partial\beta} \\ \frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\beta\partial\theta} & \frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\beta^2} \end{bmatrix}_{\substack{\hat\theta=\theta_0\\ \hat\beta=\beta_0}} \begin{bmatrix} \hat\theta-\theta_0 \\ \hat\beta-\beta_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial Lnl(x_i;\beta,\theta)}{\partial\theta} \\ \frac{\partial Lnl(x_i;\beta,\theta)}{\partial\beta} \end{bmatrix}_{\substack{\hat\theta=\theta_0\\ \hat\beta=\beta_0}} \tag{8}$$

where, $\frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\theta^2} = -\frac{n}{\theta^2} - \sum_{i=1}^{n}\left(\frac{x_i^2}{(\beta-1+\theta x_i)^2}\right)$; $\frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\beta^2} = \frac{n}{\beta^2} - \sum_{i=1}^{n}\left(\frac{1}{(\beta-1+\theta x_i)^2}\right)$ and $\frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\beta\partial\theta} = \frac{\partial^2 Lnl(x_i;\beta,\theta)}{\partial\theta\partial\beta} = -\sum_{i=1}^{n}\left(\frac{x_i}{(\beta-1+\theta x_i)^2}\right)$.

The equation (8) can be solved iteratively where θ_0, β_0 are the initial values of θ, β.

5.2. Moments estimates

Using the first moment m and the variance s^2 about PsLD, we have

$$\begin{cases} m = \frac{\beta+1}{\theta\beta} \\ s^2 = \frac{\beta^2+2\beta-1}{\theta^2\beta^2} \end{cases} \tag{9}$$

We solve this nonlinear system we find the couple $\left(\hat\theta,\hat\beta\right)$, where $\left(\hat\theta,\hat\beta\right) > 0$ for $s>0, m>0$.

$$\hat\theta = \frac{2m+\sqrt{2}\sqrt{m^2-s^2}}{m^2+s^2} \text{ et } \hat\beta = \frac{m^2+s^2}{m^2-s^2+\sqrt{2}m\sqrt{m^2-s^2}} \tag{10}$$

Theorem 3. *The estimator(MM) $\hat\theta$ of θ is positively biased i.e, $E(\hat\theta) > \theta$*

Proof. Let $\hat{\theta} = N(m)$ and $N(t) = \dfrac{\beta+1}{t\beta}$, for $t > 0$ we have for t and fixed β

$$\frac{d^2}{dt^2}N(t) = \frac{2(\beta+1)}{\beta t^3} > 0$$

$N(t)$ is strictly convex. Now, by Jensen's inequality, we have $\mathbb{E}(N(m)) > N(\mathbb{E}(m))$. Thus, $\mathbb{E}(m) = N(\mu) = N\left(\frac{\beta+1}{\theta\beta}\right) = \theta$, $E(\hat{\theta}) > \theta$. Then $\hat{\theta}$ is *positively* biased. □

Theorem 4. *The estimator(MM) $\hat{\beta}$ of β is positively biased.*

Proof. Let $\hat{\beta} = g(m)$ and $g(t) = \dfrac{1}{\theta t - 1}$, for $t > 0$ we have for t and fixed θ

$$\frac{d^2}{dt^2}g(t) = \frac{2\theta^2}{(\theta t - 1)^3} > 0$$

$g(t)$ is strictly convex. Now, by Jensen's inequality, we have $\mathbb{E}(g(m)) > g(\mathbb{E}(m))$.

Thus, $\mathbb{E}(m) = g(\mu) = g\left(\frac{\beta+1}{\theta\beta}\right) = \beta$, $E(\hat{\beta}) > \beta$. Then $\hat{\beta}$ is *positively* biased. □

5.3. Illustrative examples

Example 1. In this section, we give some simulation for four series of moments estimator parameters distribution which Lindley, PL, Exponential and Gamma distribution. For the exponential distribution; $\hat{\theta} = \frac{1}{m}$, Lindley distribution; $\hat{\theta} = \frac{-(m-1)+\sqrt{(m-1)^2+8m}}{2m}$ and gamma distribution; $\hat{\theta} = \frac{s^2}{m}$, $\hat{\alpha} = \frac{m^2}{s^2}$, see Tables 1 and 2.

$Distr(LD)$	m	s	$\hat{\theta}$
Serie1	100	71	0.0198
Serie2	50	36	0.0392
Serie3	15	12	0.1258
Serie4	3.32	2.34	0.5017

$Distr(PsLD)$	m	s	$\hat{\theta}$	$\hat{\beta}$
Serie1	100	71	0.01998	1.0083
Serie2	50	36	0.03927	1.0378
Serie3	15	12	0.11579	1.357
Serie4	3.32	2.34	0.60436	1.003

Table 1.

$Distr(Exp)$	m	s	$\hat{\theta}$
Serie1	100	71	0.0100
Serie2	50	36	0.0200
Serie3	15	12	0.0666
Serie4	3.32	2.34	0.3012

$Distr(Gamma)$	m	s	$\hat{\theta}$	$\hat{\alpha}$
Serie1	100	71	50. 41	1.9837
Serie2	50	36	25. 92	1.9290
Serie3	15	12	9. 6	1.5625
Serie4	3.32	2.34	1. 649 3	2. 013

Table 2.

Now, we used Data of survival times (in months) of 94 guinea individus infected with Ebola virus, see Table 3.

Survival time m=3.17,s=2.095	*Obsf req*	*LD* $\hat{\theta}$=0.522	*PsLD* $\hat{\theta}$=0.772,$\hat{\beta}$=1.004
$[0, 2[$	32	38.217	32.651
$[2, 4[$	35	28.16	35.981
$[4, 6[$	17	15.089	16.37
$[6, 8[$	7	7.33	6.0458
$[8, 10]$	3	3.152	2.0285
Total	94	94	94
χ^2	–	2.9244	0.6796

Table 3.

Example 2. We consider from Lawless (2003), pp. 204 and 263 two series of real data. The first one, represents the failure times (mm) for a sample of fifteen electronic components in an acceleration life test : 1.4, 5.1, 6.3, 10.8, 12.1, 18.5, 19.7, 22.2, 23, 30.6, 37.3, 46.3, 53.9, 59.8, 66.2. The second set of data, are the number of cycles to failure for 25 100-cm specimens of yarn, tested at a particular strain level : 15, 20, 38, 42, 61, 76, 86, 98, 121, 146, 149, 157, 175, 176, 180, 180, 198, 220, 224, 251, 264, 282, 321, 325, 653, see Table 4.

Data	*Distribution*	β	θ	*log−likelihood*	*Kolmogrov−Smirnov*
*Serie*1	PsLD	1.129	0.684	−62.075	0.82
n=15	Gamma	1.442	0.052	−64.197	0.102
m=27.546	Weibull	1.306	0.034	−64.026	0.450
s=20.059	Lognormal	1.061	2.931	−65.626	0.163
*Serie*2	PsLD	1.086	0.010	−150.232	0.128
n=25	Gamma	1.794	0.010	−152.371	0.135
m=178.32	Weibull	1.414	0.005	−152.440	0.697
s=131.097	Lognormal	0.891	4.880	−154.092	0.155

Table 4.

6. Conclusion

In this work, we proposed a two parameter *PsLD*, of which the *LD* is a particular case. Several properties have been discussed : moments, cumulates, characteristic function, failure rate function, stochastic ordering, distributions of sums, the maximum likelihood method and the method of moments. The *LD* does not provide enough flexibility for analyzing and modeling different types of lifetime data and survival analysis. But the *PsLD* is flexible , simple and easy to handle. Many properties and applications are given which confirm the goodness of fit and it is better than Lindley, Exponential, Gamma, Weibull, Lognormal distributions.

Acknowledgements

The authors are grateful for the comments and suggestions by the referee and the Editor. Their comments and suggestions greatly improved the article.

References

Asgharzadeh, A., Bakouch, Hassan, S. and Esmaeili, L., 2013. Pareto Poisson-Lindley distribution with applications. J. Appl. Stat. **40(8)**, 1717-1734.

Ghitany, M. E., Al-Mutairi, D.K. and Nadarajah,S., 2008. Zero-truncated Poisson–Lindley distribution and its application, Math. Comput. Simulation **79**, 279-287.

Ghitany, M. E., Atieh, B. and Nadarajah, S., 2008. Lindley distribution and its applications. Mathematics and Computers in Simulation, **78**, 493-506.

Lawless, J., 2003. Statistical models and methods for lifetime data. Wiley, New York.

Lindley, D. V., 1958. Fiducial distributions and Bayes'theorem. Journal of the Royal Society, series B, **20**, 102-107.

Nedjar, S. and Zeghdoudi, H., 2016. Gamma Lindley distribution and its application. *Journal of Applied Probability and Statistics*, **11**(1), 129-138.

Sankaran, M., 1970. The discrete Poisson-Lindley distribution. Biometrics, **26**, 145-149.

Shanker, R. and Sharma, S., 2013. A two-parameter Lindley distribution for modeling waiting and survival times data. Applied Mathematics, **4**, 363-368.

Zakerzadah, H. and Dolati, A., 2010. Generalized Lindley distribution. J. Math. Ext, **3(2)**, 13-25.

Zeghdoudi, H. and Nedjar, S., 2016. On gamma Lindley distribution : Properties and Simulations. *Journal of Computational and Applied Mathematics*, **298**, 167-174.