New Trends in Mathematical Sciences
http://dx.doi.org/10.20852/ntmsci.2017.126

On Pseudo Lindley distribution: properties and applications

Sihem Nedjar and Halim Zeghdoudi

LaPS laboratory, Badji-Mokhtar University, Box 12, Annaba, Algeria

Received: 30 December 2015, Accepted: 8 December 2016
Published online: 23 January 2017.

Abstract: In this paper, we give a treatment of the mathematical properties for a new distribution named a pseudo Lindley distribution (PsLD)[14]. The properties studied include: moments, cumulates, characteristic function, failure rate function, mean residual life function, Lorenz curve, stochastic ordering, asymptotic distribution of the extreme order statistics, maximum likelihood estimation and simulation schemes. An application to waiting time data at a bank is described.

Keywords: Lindley distribution, gamma Lindley distribution, maximum-likelihood estimation.

1 Introduction

Let X be a random variable following the one-parameter distribution with the density function

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

introduced by Lindley(1958). Sankaran (1970) used (L) as mixing distribution of Poisson parameter which it named Poisson- Lindley distribution. Recently, Asgharzadeh et al. (2013) ,Ghitany et al. (2008a) and (2008b) rediscovered and studied the new distribution bounded to (L),
their derived is known Zero-truncated Poisson- Lindley and Pareto Poisson-Lindley distributions. This work offers a two parameters family of distributions which is PsLD because for Lindley distribution there is only one parameter and not flexible for analyzing an modeling different types of lifetime data. Moreover, Zakerzadah and Dolati (2010) introduced an other distribution with three parameters which Lindley distribution is a special case, but this distribution is difficult handled and not flexible. Recently, Zeghdoudi and Nedjar (2016a,2016b) introduced a new distribution, named gamma Lindley distribution, based on mixtures of gamma $(2,\theta)$ and one-parameter Lindley (θ) distributions.The idea of this paper is based on mixtures of the ordinary exponential (θ) and gamma $(2,\theta)$ distributions.

2 Pseudo Lindley distribution(PsLD) and some 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 $\beta\geq 0$, we consider the random variable $X=Y_1$ and $X=Y_2$ with same probability

* Corresponding author e-mail: halim.zeghdoudi@univ-annaba.dz
© 2017 BISKA Bilisim Technology

$\frac{\beta-1}{\beta}$ and $\frac{1}{\beta}$ respectively. Now, the density function of X is given by

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

Remark. If $\beta=\theta+1$, *this distribution is Lindley distribution.*

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 find easily the cumulative distribution function(c.d.f) of the PsLD :

$$F_{PsLD}(x)=1-\frac{(\beta+\theta x)\,e^{-\theta x}}{\beta};x,\theta>0,\beta\geq 1 \tag{2}$$

2.1 Survival and hazard rate function

Let

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

and

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

be the survival and hazard rate function, respectively.

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

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

3 The quantile function of the Pseudo lindley distribution

3.1 The Lambert W function

Theorem 1. *For any $\theta>0,\beta\geq 1$,the quantile function of the pseudo Lindley distribution X is*

$$Q_X\,(u)=Q_X\,(u)=-\frac{\beta}{\theta}-\frac{1}{\theta}W_{-1}\left(\beta e^{-\beta}\,(u-1)\right),\quad 0\prec u\prec 1, \tag{3}$$

where W_{-1} denotes negative branche of lambert W function.

Proof. The proof is omitted because it is very similar to the proof of Theorem 1 in Zeghdoudi and Nedjar (2016c).

Further the first three quantiles we obtained by substituting $u=\frac{1}{4},\frac{1}{2},\frac{3}{4}$ in equation (3)

$$Q_1=F^{-1}\left(\frac{1}{4},\theta,\beta\right)=-\frac{\beta}{\theta}-\frac{1}{\theta}-\frac{1}{\theta}LambertW\left(-1,\beta e^{-\beta}\left(\frac{1}{4}-1\right)\right)$$

© 2017 BISKA Bilisim Technology

$$Median = Q_2 = F^{-1}\left(\frac{1}{2},\theta,\beta\right) = -\frac{\beta}{\theta} - \frac{1}{\theta} - \frac{1}{\theta}LambertW\left(-1,\beta e^{-\beta}\left(\frac{1}{2}-1\right)\right)$$

$$Q_3 = F^{-1}\left(\frac{3}{4},\theta,\beta\right) = -\frac{\beta}{\theta} - \frac{1}{\theta} - \frac{1}{\theta}LambertW\left(-1,\beta e^{-\beta}\left(\frac{3}{4}-1\right)\right).$$

Table 1 displays the mode, mean and median for PsLD distribution for different choices of parameter θ and β. Also for any choice of θ and β it is observed that $Mean > Median > Mode$. Table 1. Mode, mean and median for PsLD for different values of θ and β.

Table 1

	$\theta = 0.5, \beta = 1.5$	$\theta = 1.5, \beta = 1.5$	$\theta = 3, \beta = 1.5$
$Median = Q_2$	2.6537	0.88456	0.44228
$Mean$	3.33	1.11	0.55
$mode$	1	0.33	0.16
	$\theta = 1.2, \beta = 1$	$\theta = 1.2, \beta = 1.2$	$\theta = 1.2, \beta = 1.9$
$Median = Q_2$	1.3986	1.2554	0.97867
$Mean$	1.66	1.52	1.271
$mode$	0.83	0.66	0.083

4 Lorenz curve

The Lorenz curve for a positive random variable X is defined as the graph of the ratio

$$L(F(x)) = \frac{E(X|X \le x)F(x)}{E(X)} \tag{4}$$

against $F(x)$ with the properties $L(p) \le p, L(0) = 0$ and $L(1) = 1$. If X represents annual income,$L(p)$ is the proportion of total income that accrues to individuals having the $100p\%$ 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, see Gail and Gastwirth [2] for extensive discussion of Lorenz curves. 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 PsLD distribution in (2),

$$E(X|X \le x)F(x) = \frac{\beta+1}{\theta\beta} - \frac{e^{-\theta x}}{\beta\theta}\left[\beta(x\theta+1) + x^2\theta^2 + x\theta + 1\right].$$

Thus, from (4) we obtain the Lorenz curve for the pseudo Lindley distribution as

$$L(p) = 1 - \frac{(1-p)\beta\left(\beta(x\theta+1) + x^2\theta^2 + x\theta + 1\right)}{(\beta+1)(\beta+\theta x)}$$

where $x = F^{-1}(p)$ with $F(\cdot)$ given by (2).

© 2017 BISKA Bilisim Technology

5 Extreme order statistics

If $X_1,...,X_n$ is a random sample from (1) and if $\overline{X} = (X_1+\cdots+X_n)/n$ denotes the sample mean then by the usual central limit theorem $\frac{\sqrt{n}(\overline{X}-E(X))}{\sqrt{Var(X)}}$ approaches the standard normal distribution as $n \to \infty$. Sometimes one would be interested in the asymptotics of the extreme values $M_n = max(X_1,...,X_n)$ and$m_n = min(X_1,...,X_n)$. For the c.d.f. in (2), it can be seen that

$$\lim_{t\to\infty} \frac{1-F(t+x)}{1-F(t)} = \exp(-\theta x)$$

and

$$\lim_{t\to 0} \frac{F(tx)}{F(t)} = x.$$

Thus, it follows from Theorem 1.6.2 in Leadbetter et al. [6] that there must be norming constants $a_n > 0, b_n, c_n > 0$ and d_n such that

$$Pr\{a_n(M_n - b_n) \leq x\} \to -\exp\{-\theta x)\} \tag{5}$$

and

$$Pr\{c_n(m_n - d_n) \leq x\} \to 1-\exp(-x) \tag{6}$$

as $n \to \infty$. The form of the norming constants can also be determined. For instance, using Corollary 1.6.3 in Leadbetter et al. [6], one can see that $a_n = 1$ and $b_n = F^{-1}(1-1/n)$ with $F(\cdot)$ given by (2).

5.1 Maximum Likelihood Estimates (MLE)

In this section we shall discuss the point and interval estimation on the parameters that index the $PsL(\theta,\beta)$. Let the log-likelihood function of single observation(say x_i) for the vector of parameter (θ,β) can be written as

$$\ln l(x;\beta,\theta) = \ln\theta - \ln\beta + \ln(\beta - 1 + \theta x_i) - \theta x.$$

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

$$\frac{\partial \ln l(x;\beta,\theta)}{\partial\theta} = \frac{1}{\theta} - x + \left(\frac{x}{\beta - 1 + \theta x}\right) \tag{2}$$

$$\frac{\partial \ln l(x;\beta,\theta)}{\partial\beta} = \frac{-1}{\beta} + \left(\frac{1}{\beta - 1 + \theta x}\right) \tag{3}$$

The maximum likelihood estimator $\hat{\theta}$ of θ and $\hat{\beta}$ of β is obtained by solving equation(7) and (8) numerically we give

$$\begin{cases} \hat{\theta} = \frac{1}{X} \\ \widehat{\beta} = \frac{1}{X-1} \end{cases} \tag{9}$$

and

$$\begin{cases} E\left(\hat{\theta}\right) = \frac{\beta+1}{\theta\beta} = m \\ E\left(\widehat{\beta}\right) = \frac{\beta+\theta+1}{\beta\theta} e^{-\theta} \end{cases} \tag{10}$$

© 2017 BISKA Bilisim Technology

6 Simulation

In this section, we investigate the behavior of the ML estimators for a nite sample size (n).Simulation study based on different $PsLD(\theta,\beta)$ is carried out. A simulation study consisting of following steps is being carried out for each triplet $(\beta,\theta\,;n)$, where $\theta = 0.5, 0.9, 1$, $\beta = 1.5, 5, 6$ and $n = 10, 30, 50$.

– Choose the initial values of θ_0, β_0 for the corresponding elements of the parameter vector $\Theta = (\theta,\beta)$ to specify GaL distribution;
– choose sample size n;
–generate N independent samples of size n from $PsLD(\theta,\beta)$;
–compute the ML estimate $\hat{\Theta}_n$ of Θ_0 for each of the N samples;
–compute the mean of the obtained estimators over all N samples,

$$average\ bias\,(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\Theta}_i - \Theta_0\right)$$

and the average square error (see tables 2 and 3)

$$MSE\,(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\Theta}_i - \Theta_0\right)^2.$$

Table 2: Average bias of the simulated estimates.

	$\theta = 1, \beta = 6$		$\theta = 0.5, \beta = 1.5$		$\theta = 0.9, \beta = 5$	
	bias(θ)	bias(β)	bias(θ)	bias(β)	bias(θ)	bias(β)
$n = 10$	1.6667×10^{-2}	0.5509	0.2833	9.2612×10^{-2}	4.3333×10^{-2}	0.4376
$n = 30$	5.5556×10^{-3}	0.1836	9.4444×10^{-2}	3.0871×10^{-2}	1.4444×10^{-2}	0.1458
$n = 50$	3.3333×10^{-3}	0.1101	5.666×10^{-2}	1.8522×10^{-2}	8.6667×10^{-3}	0.0875

Table 3: Average MSE of the simulated estimates.

	$\theta = 1, \beta = 6$		$\theta = 0.5, \beta = 1.5$		$\theta = 0.9, \beta = 5$	
	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)	MSE(θ)	MSE(β)
$n = 10$	2.7778×10^{-3}	3.0355	0.8027	0.08577	1.8778×10^{-2}	1.9155
$n = 30$	9.2593×10^{-4}	1.0118	0.2676	0.02859	6.2593×10^{-3}	0.6384
$n = 50$	5.5556×10^{-4}	0.6071	0.1605	0.0171	3.7556×10^{-3}	0.3831

7 Application to Real Data sets

In this section, we illustrate, the applicability of PsLD by considering two different data sets used by different researchers. We also fit generalized Lindley [11], quasi Lindley [9], two-parameter Lindley [8], gamma, Weibull and lognormal distributions.

In each of these distributions, the parameters are estimated by using the moment method, and for comparison we use negative log-likelihood values $(-LL)$, the Akaike information criterion (AIC) and Bayesian information criterion (BIC) which are defined by $-2LL + 2q$ and $-2LL + q\log(n)$, respectively, where q is the number of parameters estimated and n

© 2017 BISKA Bilisim Technology

is the sample size. Further $K-S$ (Kolmogorov-Smirnov) test statistic defined as $K-S = sup_x |F_n(x) - F(x)|$, where $F_n(x)$ is empirical distribution function and $F(x)$ is cumulative distribution function is calculated and shown for all the data sets.

Example 1. We consider from Lawless [5], 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.
According table 4, we can observe that pseudo Lindley distribution provide smallest k- S as compare to quasi Lindley,two-parameter Lindley, generalized Lindley, Weibull and lognormal distributions.

Table 4: Comparison between several distributions.

Data	*Distribution*	β	θ	γ	log-likelihood	$K-S$	*AIC*	*BIC*
*Serie*1	Generalized Lindley	1.203	0.064	0.083	-64.080	0.095	134.16	136.28
n=15	*PsLD*	1.129	0.684		-62.075	0.082	128.15	129.57
m=27.546	*QLD*	4.016	-0.99		-1504	0.93	301.2	3013.4
s=20.059	*TwoPLD*	0.0704	1.110		-169.12	0.196	342.24	343.66
	Gamma	1.442	0.052		-64.197	0.102	132.39	133.81
	Weibull	1.306	0.034		-64.026	0.450	132.05	133.47
	Lognormal	1.061	2.931		-65.626	0.163	135.25	136.67
*Serie*2	Generalized Lindley	1.505	0.012	0.018	-152.369	0.137	310.74	314.39
n=25	*PsLD*	1.086	0.010		-150.232	0.128	304.464	306.9
m=178.32	*QLD*	0.0107	8.514		-104.59	0.93	213.18	215.62
$s = 131.097$	*TwoPLD*	0.0107	0,125		-283.41	0.232	570.82	573.26
	Gamma	1.794	0.010		-152.371	0.135	308.74	311.18
	Weibull	1.414	0.005		-152.440	0.697	308.88	310.7
	Lognormal	0.891	4.880		-154.092	0.155	312.18	314.62

8 Conclusion

In this work, we discussed more statistical properties of two parameter PsLD, including the quantile function, Lorenz curve and probability density of the order statistics. The maximum likelihood estimates of the two parameters index to the new distribution are discussed. The distribution includes the Lindley and the exponential distributions as special cases.Two real data sets are analyzed using the new distribution and it is compared with six immediate sub-models mentioned above in addition to another distributions (quasi Lindley ,Two-Parameter Lindley, Generalized Lindley, Weibull and Lognormal distributions). The results of the comparisons showed that the new distribution provides a better fit than those three mentioned distributions to the three data sets. We hope our new distribution might attract wider sets of applications in lifetime data reliability analysis and actuarial sciences. For future studies, we can explain the derivation of posterior distributions for the Pseudo Lindley distribution under Linex loss functions and squared error using non-informative and informative priors(the extension of Jeffreys and Inverted Gamma priors) respectively.

Acknowledgment

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

© 2017 BISKA Bilisim Technology

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have contributed to all parts of the article. All authors read and approved the final manuscript.

References

[1] A. Asgharzadeh , H.S. Bakouch, L. Esmaeili (2013). Pareto Poisson-Lindley distribution and its application. *Journal of Applied Statistics, pp. 1-18.*

[2] M. H. Gail and J. L.Gastwirth (1978). A scale-free goodness of fit test for the exponential distribution based on the Lorenz curve. *Journal of the American Statistical Association* **73** 787–793.

[3] M. E. Ghitany, B. Atieh, S. Nadarajah (2008a). Lindley distribution and its applications. *Math. Comput. Simulation , **78**, pp. 493-506* .

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

[5] J. F. Lawless (2003). Statistical models and methods for lifetime data. *Wiley, New York*.

[6] M.R. Leadbetter, G. Lindgren, H. Rootzén (1987). Extremes and Related Properties of Random Sequences and Processes, *Springer Verlag, New York*.

[7] D. V. Lindley (1958). Fiducial distributions and Bayes theorem. *Journal of the Royal Society, series B, **20**, pp. 102-107.*

[8] R.Shanker, S.Sharma, R.Shanker. (2013). A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics, vol. **4**, pp. 363-368.*

[9] R. Shanker and A. Mishra (2013). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research,Vol. 6(4), pp. 64-71.*

[10] M. Sankaran (1970). The discrete Poisson-Lindley distribution. *Biometrics, **26**, pp. 145-149.*

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

[12] H. Zeghdoudi, S. Nedjar (2016a). Gamma Lindley distribution and its application. *Journal of Applied Probability and Statistics. Vol. **11**, N° **1**, 129-138.*

[13] H. Zeghdoudi, S. Nedjar (2016b). On Gamma Lindley distribution :proprieties and simulation. *Journal of Computation and Applied Mathematics Probability and Statistics. 298, pp 167-174.*

[14] H. Zeghdoudi, S. Nedjar (2016c) A Pseudo Lindley distribution and its application.,*J. Afrika Statistika. Vol 11 (1) page 923-932*

[15] H. Zeghdoudi, S. Nedjar (2017). A Poissson pseudo Lindley distribution and its application . *Journal of Probability and Statistical Sciences. Vol. 15, N° 1.*

© 2017 BISKA Bilisim Technology