Janardan Distribution and its Application to Waiting Times Data

KEYWORDS: Lindley distribution, moments, failure rate function, mean residual life function, stochastic ordering, estimation of parameters, goodness of fit

Rama Shanker
Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea

Shambhu Sharma
Department of Mathematics, Dayalbagh Educational Institute, Agra, India

Uma Shanker
Department of Mathematics, K.K.College of Engineering & Management, Bihar Sharif (Nalanda) Bihar, India

Ravi Shanker
Department of Mathematics, G.L.A.College, N.P University, Daltonganj, Jharkhand, India

ABSTRACT *A two parameter continuous distribution named "Janardan distribution (JD)", of which the Lindley distribution (LD) is a particular case, has been introduced. Its moments, failure rate function, mean residual life function and stochastic orderings have been discussed. The maximum likelihood method and the method of moments have been discussed for estimating its parameters. The superiority of the proposed distribution has been illustrated with an application to a real data set.*

1. INTRODUCTION

Lindley (1958) introduced a one-parameter distribution, known as Lindley distribution, given by its probability density function

$$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)\,e^{-\theta x};\qquad x>0,\quad \theta>0 \quad (1.1)$$

It can be seen that this distribution is a mixture of exponential (θ) and gamma $(2,\theta)$ distributions. Its cumulative distribution function has been obtained as

$$F(x)=1-\frac{\theta+1+\theta x}{\theta+1}e^{-\theta x};\ x>0,\ \theta>0 \quad (1.2)$$

Ghitany et al (2008) have discussed various properties of this distribution and showed that in many ways (1.1) provides a better model for some applications than the exponential distribution. The first four moments about origin of the Lindley distribution have been obtained as

$$\mu_1'=\frac{\theta+2}{\theta(\theta+1)},\quad \mu_2'=\frac{2(\theta+3)}{\theta^2(\theta+1)},\quad \mu_3'=\frac{6(\theta+4)}{\theta^3(\theta+1)},\quad \mu_4'=\frac{24(\theta+5)}{\theta^4(\theta+1)} \quad (1.3)$$

and its central moments have been obtained as

$$\mu_2=\frac{\theta^2+4\theta+2}{\theta^2(\theta+1)^2},\quad \mu_3=\frac{2(\theta^3+6\theta^2+6\theta+2)}{\theta^3(\theta+1)^3},\quad \mu_4=\frac{3(3\theta^4+24\theta^3+44\theta^2+32\theta+8)}{\theta^4(\theta+1)^4} \quad (1.4)$$

Ghitany et al (2008) studied various properties of this distribution. A discrete version of this distribution has been suggested by Deniz and Ojeda (2011) having its applications in count data related to insurance. Sankaran (1970) obtained the Lindley mixture of Poisson distribution.

In this paper, a two parameter continuous distribution named "Janardan distribution (JD)", of which the Lindley distribution (LD) (1.1) is a particular case, has been suggested. Its first four moments and some of the related measures have been obtained. Its failure rate, mean residual rate and stochastic ordering have also been studied. Estimation of its parameters has been discussed and the distribution has been fitted to some of those data sets where the Lindley distribution has earlier been fitted by others, to test its goodness of fit.

2. JANARDAN DISTRIBUTION

A two parameter continuous distribution with parameters α and θ is defined by its probability density function (p.d.f.)

$$f(x;\theta,\alpha)=\frac{\theta^2}{\alpha(\theta+\alpha^2)}(1+\alpha x)\,e^{-\frac{\theta}{\alpha}x};\ x>0,\ \theta>0,\ \alpha>0 \quad (2.1)$$

We would call this two parameter continuous distribution as "Janardan distribution (JD)". It can easily be seen that Lindley distribution (LD) is a particular case of (2.1) for $\alpha=1$. The p.d.f. (2.1) can be shown as a mixture of exponential $\left(\frac{\theta}{\alpha}\right)$ and gamma $\left(2,\frac{\theta}{\alpha}\right)$ distributions as follows:

$$f(x;\theta,\alpha)=pf_1(x)+(1-p)f_2(x) \quad (2.2)$$

where $p=\frac{\theta}{\theta+\alpha^2}$, $f_1(x)=\frac{\theta}{\alpha}e^{-\frac{\theta}{\alpha}x}$ and $f_2(x)=\frac{\theta^2}{\alpha^2}x\,e^{-\frac{\theta}{\alpha}x}$.

The first derivative of (2.1) is obtained as

$$f'(x)=\frac{\theta^2}{\alpha(\theta+\alpha^2)}\left[\frac{(\alpha^2-\theta)-\theta\alpha x}{\alpha}\right]e^{-\frac{\theta}{\alpha}x}$$

and so $f'(x)=0$ gives $x=\frac{\alpha^2-\theta}{\theta\alpha}$. From this it follows that

(i) for $\alpha>\theta$, $x_0=\frac{\alpha^2-\theta}{\theta\alpha}$ is the unique critical point at which $f(x)$ is maximum.

(ii) for $\alpha<\theta$, $f'(x)\le 0$ i.e. $f(x)$ is decreasing in x.

Therefore, the mode of the distribution (2.1) is given by

$$\text{Mode}=\begin{cases}\frac{\alpha^2-\theta}{\theta\alpha}, & \alpha>\theta\\ 0, & \text{otherwise}\end{cases} \quad (2.3)$$

The cumulative distribution function of the distribution (2.1) is given by

$$F(x)=1-\frac{\alpha(\theta+\alpha^2)+\theta\alpha^2x}{\alpha(\theta+\alpha^2)}e^{-\frac{\theta}{\alpha}x}\quad ;\ x>0,\theta>0,\alpha>0 \quad (2.4)$$

3. MOMENTS AND RELATED MEASURES

The rth moment about origin of the Janardan distribution has been obtained as

$$\mu_r'=r!\left(\frac{\alpha}{\theta}\right)^r\left(1+\frac{r\alpha^2}{\theta+\alpha^2}\right);\ r=1,2,3,\ldots \quad (3.1)$$

Taking $r=1,2,3$ and 4 in (3.1), the first four moments about origin of Janardan distribution are thus obtained as

$$\mu_1'=\frac{\alpha}{\theta}\left(\frac{\theta+2\alpha^2}{\theta+\alpha^2}\right),\mu_2'=2\left(\frac{\alpha}{\theta}\right)^2\left(\frac{\theta+3\alpha^2}{\theta+\alpha^2}\right),\mu_3'=6\left(\frac{\alpha}{\theta}\right)^3\left(\frac{\theta+4\alpha^2}{\theta+\alpha^2}\right),\mu_4'=24\left(\frac{\alpha}{\theta}\right)^4\left(\frac{\theta+5\alpha^2}{\theta+\alpha^2}\right) \quad (3.2)$$

It can be easily verified that for $\alpha=1$, the moments about origin of the Janardan distribution reduce to the respective moments of the Lindley distribution. Further, the mean of the distribution is always greater than the mode, the distribution is positively skewed. The central moments of the Janardan distribution have thus been obtained as

$$\mu_2=\frac{\alpha^2\left(\theta^2+4\theta\alpha^2+2\alpha^4\right)}{\theta^2\left(\theta+\alpha^2\right)^2} \quad (3.3)$$

$$\mu_3=\frac{2\alpha^3\left(\theta^3+6\theta^2\alpha^2+6\theta\alpha^4+2\alpha^6\right)}{\theta^3\left(\theta+\alpha^2\right)^3} \quad (3.4)$$

$$\mu_4=\frac{3\alpha^4\left(3\theta^4+24\theta^3\alpha^2+44\theta^2\alpha^4+32\theta\alpha^6+8\alpha^8\right)}{\theta^4\left(\theta+\alpha^2\right)^4} \quad (3.5)$$

It can be easily verified that for $\alpha=1$, the central moments of the Janardan distribution reduce to the respective moments of the Lindley distribution.

The coefficients of variation(γ), skewness$(\sqrt{\beta_1})$and the kurtosis$(\beta_2)$of the Janardan distribution are given by

$$\gamma=\frac{\sigma}{\mu_1'}=\frac{\sqrt{\theta^2+4\theta\alpha^2+2\alpha^4}}{\theta+2\alpha^2} \quad (3.6)$$

$$\sqrt{\beta_1}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{2\left(\theta^3+6\theta^2\alpha^2+6\theta\alpha^4+2\alpha^6\right)}{\left(\theta^2+4\theta\alpha^2+2\alpha^4\right)^{3/2}} \quad (3.7)$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{3\left(3\theta^4+24\theta^3\alpha^2+44\theta^2\alpha^4+32\theta\alpha^6+8\alpha^8\right)}{\left(\theta^2+4\theta\alpha^2+2\alpha^4\right)^2} \quad (3.8)$$

It can be easily seen for $\alpha=1$ the coefficients of variation (γ), skewness$(\sqrt{\beta_1})$and the kurtosis$(\beta_2)$of the Janardan distribution reduce to the corresponding measures of the Lindley distribution.

4. FAILURE RATE AND MEAN RESIDUAL LIFE

For a continuous distribution with p.d.f. $f(x)$and c.d.f. $F(x)$, the failure rate function (also known as the hazard rate function) and the mean residual life function are respectively defined as

$$h(x)=\lim_{\Delta x\to 0}\frac{P\left(X<x+\Delta x\,|\,X>x\right)}{\Delta x}=\frac{f(x)}{1-F(x)} \quad (4.1)$$

and $m(x)=E\left[X-x\,|\,X>x\right]=\frac{1}{1-F(x)}\int_x^\infty\left[1-F(t)\right]dt$ (4.2)

The corresponding failure rate function, $h(x)$and the mean residual life function, $m(x)$of the Janardan distribution are thus given by

$$h(x)=\frac{\theta^2(1+\alpha x)}{\alpha\left(\theta+\alpha^2\right)+\theta\alpha^2 x} \quad (4.3)$$

And $m(x)=\frac{1}{\left[\alpha\left(\theta+\alpha^2\right)+\theta\alpha^2 x\right]e^{-\frac{\theta}{\alpha}x}}\int_x^\infty\left[\alpha\left(\theta+\alpha^2\right)+\theta\alpha^2 t\right]e^{-\frac{\theta}{\alpha}t}dt$

$$=\frac{\alpha}{\theta}\left[\frac{\left(\theta+\alpha^2\right)+\left(\theta\alpha x+\alpha^2\right)}{\left(\theta+\alpha^2\right)+\theta\alpha x}\right] \quad (4.4)$$

It can be easily verified that $h(0)=\frac{\theta^2}{\alpha(\theta+1)}=f(0)$and $m(0)=\frac{\alpha(\theta+2)}{\theta(\theta+1)}=\mu_1'$ .It is also obvious that $h(x)$is an increasing function of x, α and θ, whereas $m(x)$is a decreasing function of x, $\alpha$and θ. For $\alpha=1$, (4.3) and (4.4) reduce to the corresponding measures of the Lindley distribution. The failure rate function and the mean residual life function of the Janardan distribution show its flexibility over Lindley distribution and exponential distribution.

5. STOCHASTIC ORDERINGS

Stochastic ordering of positive continuous random variables is an important tool for judging the comparative behaviour. A random variable X is said to be smaller than a random variable Y in the

(i) stochastic order $(X\le_{st}Y)$if $F_X(x)\ge F_Y(x)$for all x
(ii) hazard rate order $(X\le_{hr}Y)$if $h_X(x)\ge h_Y(x)$ for all x

(iii) mean residual life order $(X\le_{mrl}Y)$if $m_X(x)\le m_Y(x)$for all x
(iv) likelihood ratio order $(X\le_{lr}Y)$if $\frac{f_X(x)}{f_Y(x)}$ decreases in x.

The following results due to Shaked and Shanthikumar (1994) are well known for establishing stochastic ordering of distributions

$$X\le_{lr}Y\Rightarrow X\le_{hr}Y\Rightarrow X\le_{mrl}Y \quad (5.1)$$
$$\Downarrow$$
$$X\le_{st}Y$$

The Janardan distribution is ordered with respect to the strongest 'likelihood ratio' ordering as shown in the following theorem:

Theorem: Let $X\sim$ Janardan distribution(θ_1,α_1) and $Y\sim$ Janardan distribution(θ_2,α_2). If $\alpha_1=\alpha_2$ and $\theta_1\ge\theta_2$(or if $\theta_1=\theta_2$ and $\alpha_1\le\alpha_2$), then $X\le_{lr}Y$and hence $X\le_{hr}Y$, $X\le_{mrl}Y$ and $X\le_{st}Y$.

Proof: We have

$$\frac{f_X(x)}{f_Y(x)}=\left(\frac{\theta_1}{\theta_2}\right)^2\frac{\alpha_2\left(\theta_2+\alpha_2^2\right)}{\alpha_1\left(\theta_1+\alpha_1^2\right)}\left(\frac{1+\alpha_1 x}{1+\alpha_2 x}\right)e^{-\frac{(\theta_1\alpha_2-\theta_2\alpha_1)}{\alpha_1\alpha_2}x}\ ;\ x>0$$

Now

$$\log\frac{f_X(x)}{f_Y(x)}=2\log\left(\frac{\theta_1}{\theta_2}\right)+\log\left[\frac{\alpha_2\left(\theta_2+\alpha_2^2\right)}{\alpha_1\left(\theta_1+\alpha_1^2\right)}\right]+\log\left(1+\alpha_1 x\right)-\log\left(1+\alpha_2 x\right)-\frac{\left(\theta_1\alpha_2-\theta_2\alpha_1\right)}{\alpha_1\alpha_2}x$$

Thus

$$\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)}=\frac{\alpha_1}{1+\alpha_1 x}-\frac{\alpha_2}{1+\alpha_2 x}-\frac{\left(\theta_1\alpha_2-\theta_2\alpha_1\right)}{\alpha_1\alpha_2}$$
$$=\frac{\alpha_1-\alpha_2}{(1+\alpha_1 x)(1+\alpha_2 x)}-\frac{\left(\theta_1\alpha_2-\theta_2\alpha_1\right)}{\alpha_1\alpha_2} \quad (5.2)$$

Case (i) If $\alpha_1=\alpha_2$ and $\theta_1\ge\theta_2$, then $\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)}<0$.

This means that $X\le_{lr}Y$and hence $X\le_{hr}Y$, $X\le_{mrl}Y$and $X\le_{st}Y$.

Case (ii) If $\theta_1=\theta_2$ and $\alpha_1\le\alpha_2$, then $\frac{d}{dx}\log\frac{f_X(x)}{f_Y(x)}<0$.

This means that $X\le_{lr}Y$and hence $X\le_{hr}Y$, $X\le_{mrl}Y$and $X\le_{st}Y$.

This theorem shows the flexibility of the Janardan distribution over Lindley and exponential distributions.

6. ESTIMATION OF PARAMETERS

6.1. Maximum Likelihood Estimates :Let $x_1,x_2,\ \ ,x$ be a random sample of size n from the Janardan (2.1) and let f_x be the observed frequency in the sample corresponding to $X=x$ $(x=1,2,...,k)$ such that $\sum_{x=1}^k f_x=n$, where k is the largest observed value having non-zero frequency. The likelihood function, L of the Janardan distribution (2.1) is given by

$$L=\left[\frac{\theta^2}{\alpha\left(\theta+\alpha^2\right)}\right]^n\prod_{x=1}^k(1+\alpha x)^{f_x}e^{-\frac{n\theta}{\alpha}\bar{X}} \quad (6.1.1)$$

and so the log likelihood function is obtained as

$$\log L=n\log\theta^2-n\log\left\{\alpha\left(\theta+\alpha^2\right)\right\}+\sum_{i=1}^k f_x\log\left(1+\alpha x\right)-\frac{n\theta}{\alpha}\bar{X} \quad (6.1.2)$$

The two log likelihood equations are thus obtained as

$$\frac{\partial\log L}{\partial\theta}=\frac{2n}{\theta}-\frac{n}{\theta+\alpha^2}-\frac{n\bar{X}}{\alpha}=0 \quad (6.1.3)$$

$$\frac{\partial\log L}{\partial\alpha}=-\frac{n}{\alpha}-\frac{2n\alpha}{\theta+\alpha^2}+\sum_{x=1}^k\frac{x f_x}{1+\alpha x}+\frac{n\theta\bar{X}}{\alpha^2}=0 \quad (6.1.4)$$

Volume : 3 | Issue : 8 | Aug 2013 | ISSN - 2249-555X

Equation (6.1.3) gives $\bar{X}=\frac{\alpha\left(\theta+2\alpha^2\right)}{\theta\left(\theta+\alpha^2\right)}$, which is the mean of the Janardan distribution. The two equations (6.1.3) and (6.1.4) do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. For, we have

$$\frac{\partial^2 \log L}{\partial \theta^2}=-\frac{2n}{\theta^2}+\frac{n}{\left(\theta+\alpha^2\right)^2} \quad (6.1.5)$$

$$\frac{\partial^2 \log L}{\partial \theta\, \partial \alpha}=\frac{2n\alpha}{\left(\theta+\alpha^2\right)^2}+\frac{n\bar{X}}{\alpha^2} \quad (6.1.6)$$

$$\frac{\partial^2 \log L}{\partial \alpha^2}=\frac{n}{\alpha^2}-\frac{2n\left(\theta-\alpha^2\right)}{\left(\theta+\alpha^2\right)^2}-\sum_{x=1}^{k}\frac{x^2 f_x}{(1+\alpha x)^2}-\frac{2n\theta\bar{X}}{\alpha^3} \quad (6.1.7)$$

The following equations for $\hat{\theta}$ and $\hat{\alpha}$ can be solved

$$\begin{bmatrix}\frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha}\\ \frac{\partial^2 \log L}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2}\end{bmatrix}_{\substack{\hat{\theta}=\theta_0\\ \hat{\alpha}=\alpha_0}}\begin{bmatrix}\hat{\theta}-\theta_0\\ \hat{\alpha}-\alpha_0\end{bmatrix}=\begin{bmatrix}\frac{\partial \log L}{\partial \theta}\\ \frac{\partial \log L}{\partial \alpha}\end{bmatrix}_{\substack{\hat{\theta}=\theta_0\\ \hat{\alpha}=\alpha_0}} \quad (6.1.8)$$

where θ_0 and α_0 are the initial values of θ and α respectively. These equations are solved iteratively till sufficiently close estimates of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

6.2. Estimates from Moments: The Janardan distribution has two parameters to be estimated. Using the first two moments about origin, we have

$$\frac{\mu_2'}{\mu_1'^2}=k\ \text{(say)}=\frac{2\left(\theta+3\alpha^2\right)\left(\theta+\alpha^2\right)}{\left(\theta+2\alpha^2\right)^2} \quad (6.2.1)$$

Using $\theta=b\alpha^2$ in (6.2.1), we get

$$\frac{2(b+3)(b+1)}{(b+2)^2}=k$$

This gives a quadratic equation in b as

$$(2-k)b^2+4(2-k)b+2(3-2k)=0 \quad (6.2.2)$$

Again substituting $\theta=b\alpha^2$ in the expression for mean

$\bar{X}=\frac{\alpha\left(\theta+2\alpha^2\right)}{\theta\left(\theta+\alpha^2\right)}$, we get

$\bar{X}=\frac{b+2}{\alpha b(b+1)}$ and thus $\hat{\alpha}=\frac{b+2}{b(b+1)\bar{X}}$ (6.2.3)

After getting the estimated value of α from equation (6.2.3), we finally get

$$\hat{\theta}=b\hat{\alpha}^2=\frac{1}{b}\left\{\frac{b+2}{(b+1)\bar{X}}\right\}^2 \quad (6.2.4)$$

7. APPLICATION

The Janardan distribution has been fitted to a number of data- sets to which earlier the Lindley distribution has been fitted by others and to almost all these data-sets the Janardan distribution provides closer fits than the Lindley distribution.

Here the fittings of the Janardan distribution to data-set relating to waiting times (in minutes) of 100 bank customers reported by Ghitany et al (2008) have been presented in the following table. The expected frequencies according to the Lindley distribution have also been given for ready comparison with those obtained by the Janardan distribution. The estimates of the parameters have been obtained by the method of moments.

It can be seen from table 1 that the Janardan distribution gives much closer fits than the Lindley distribution and thus provides a better alternative to the Lindley distribution for modeling waiting times data.

Table 1
Waiting times (in minutes) of 100 bank customers
Waiting Time Observed Expected frequency

(In minutes) frequency Lindley distribution Janardan distribution

Waiting Time (In minutes)	Observed frequency	Expected frequency: Lindley distribution	Expected frequency: Janardan distribution
0 – 5	30	30.39	30.16
5 – 10	32	30.69	30.92
10 – 15	19	19.21	19.32
15 – 20	10	10.28	10.28
20 – 25	5	5.08	5.05
25 – 30	1	2.40	2.37
30 – 35	2	1.09	1.07
35 – 40	1	0.86	0.83
Total	100	100.0	100.0
Estimates of parameters		$\hat{\theta}=0.1897$	$\hat{\theta}=0.213874$, $\hat{\alpha}=1.118919$
	χ^2	0.091	0.068
	d.f.	4	3

CONCLUSION

In this paper, a two-parameter continuous distribution, named "Janardan distribution (JD)", of which the Lindley distribution (LD) is a particular case, has been proposed. Several properties of the Janardana distribution such as moments, failure rate function, mean residual life function, stochastic orderings, estimation of parameters by the method of maximum likelihood and the method of moments have been discussed. Finally, an application of the proposed distribution has been given by fitting to data sets relating to waiting times to test its goodness of fit to which earlier the LD has been fitted and it is found that the Janardan distribution provides better fits than those by the LD. An application to a real data set indicates that the fit of the proposed distribution is superior to the fit of the Lindley distribution and we hope that the proposed distribution may be interesting for a wider range of statistics research.

REFERENCE

. Deniz, E.G., and Ojeda, E.C. (2011): The discrete Lindley distribution-Properties and Applications, Journal of Statistical Computation and Simulation, Vol. 81 (11), 1405 – 1416. | | 2. Ghitany, M. E., Atieh, B., Nadarajah, S. (2008): Lindley distribution and its Applications, Mathematics and Computers in Simulation, Vol.78 (4), 493 – 506. | | 3. Lindley, D.V. (1958): Fiducial distributions and Bayes' theorem, Journal of the Royal Statistical Society, Series B, 20, 102- 107. | | 4. Sankaran, M. (1970): The discrete Poisson-Lindley distribution, Biometrics, 26, 145 – 149. | | 5. Shaked,M. and Shanthikumar,J.G.(1994): Stochastic Orders and Their Applications, Academic Press, New York. | |