The Discrete Poisson-Janardan Distribution with Applications

Rama Shanker, Shambhu Sharma, Uma Shanker, Ravi Shanker, Tekie Asehun Leonida

Abstract- In the present paper a discrete Poisson-Janardan distribution (PJD), of which the Sankaran's (1970) discrete Poisson-Lindley distribution (PLD) is a particular case, has been obtained by compounding Poisson distribution with the Janardan distribution of Shanker et al (2013). The first four moments of this distribution have been obtained and the estimation of its parameters using the method of maximum likelihood and the method of moments has been discussed. The distribution has been fitted to some data-sets to test its goodness of fit and its fitting of two data sets has been presented.

Keywords: Poisson-Lindley distribution, Janardan distribution, compounding, moments, estimation of parameters, goodness of fit.

I. INTRODUCTION

The discrete Poisson-Lindley distribution (PLD) given by its probability mass function

$$P(x) = \frac{\theta^2\left(x+\theta+2\right)}{\left(\theta+1\right)^{x+3}} \;, x = 0,\, 1,\, 2, \ldots, \quad \theta > 0. \qquad (1.1)$$

has been introduced by Sankaran (1970) to model count data. The distribution arises from a Poisson distribution when its parameter follows the Lindley (1958) distribution with its probability density function

$$f\left(x,\theta\right) = \frac{\theta^2}{\theta+1}\left(1+x\right) e^{-\theta x} \;, \qquad x > 0, \quad \theta > 0. \qquad (1.2)$$

The first four moments about origin of the discrete PLD (1.1) have been obtained as

$$\mu_1' = \frac{\theta+2}{\theta\left(\theta+1\right)} \qquad (1.3)$$

$$\mu_2' = \frac{\theta+2}{\theta\left(\theta+1\right)} + \frac{2\left(\theta+3\right)}{\theta^2\left(\theta+1\right)} \qquad (1.4)$$

$$\mu_3' = \frac{\theta+2}{\theta\left(\theta+1\right)} + \frac{6\left(\theta+3\right)}{\theta^2\left(\theta+1\right)} + \frac{6\left(\theta+4\right)}{\theta^3\left(\theta+1\right)} \qquad (1.5)$$

$$\mu_4' = \frac{\theta+2}{\theta\left(\theta+1\right)} + \frac{14\left(\theta+3\right)}{\theta^2\left(\theta+1\right)} + \frac{36\left(\theta+4\right)}{\theta^3\left(\theta+1\right)} + \frac{24\left(\theta+5\right)}{\theta^4\left(\theta+1\right)} \qquad (1.6)$$

Manuscript Received May 2014

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

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

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

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

Tekie Asehun Leonida, Department of Mathematics, Eritrea Institute of Technology, Asmara, Eritrea, India.

Ghitany et al (2009) discussed the estimation methods for the discrete Poisson- Lindley distribution (1.1) and its applications.

In this paper, a discrete Poisson-Janardan distribution (PJD) has been obtained by compounding a Poisson distribution with the Janardan distribution of Shanker et al (2013). The first four moments of this distribution have been obtained and the estimation of its parameters has been discussed. The discrete PJD has been fitted to some data sets used earlier by others for fitting the discrete PLD and it has been found to be more flexible than the discrete PLD for analyzing different types of count data.

II. THE DISCRETE POISSON- JANARDAN DISTRIBUTION (PJD)

Recently Shanker et al (2013) obtained a two-parameter Janardan distribution given by its probability density function

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

$$\theta > 0, \quad \alpha > 0 \qquad (2.1)$$

It can be seen that at $\alpha = 1$, it reduces to the one parameter Lindley distribution (1.2).This distribution has been found to be a better model than the one parameter Lindley distribution for analyzing waiting time, survival time and grouped mortality data.

Suppose that the parameter λ of the Poisson distribution follows the Janardan distribution (2.1). Thus, the two-parameter Janardan mixture of Poisson distribution can be obtained as

$$P\left(x\right) = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{\Gamma\left(x+1\right)} \frac{\theta^2}{\alpha\left(\theta+\alpha^2\right)}\left(1+\alpha\lambda\right) e^{-\frac{\theta}{\alpha}\lambda} d\lambda \qquad (2.2)$$

$$= \left(\frac{\theta}{\theta+\alpha}\right)^2 \left(\frac{\alpha}{\theta+\alpha}\right)^x \left[1+\frac{\alpha\left(1+\alpha x\right)}{\theta+\alpha^2}\right], x = 0,1,2,...;$$

$$\theta > 0, \quad \alpha > 0 \qquad (2.3)$$

We name this distribution as 'A discrete Poisson -Janardan distribution (PJD)'. It can be seen that at $\alpha = 1$, this distribution reduces to the discrete Poisson-Lindley distribution (PLD) (1.1).

III. MOMENTS

The r th moment about origin of the discrete Poisson - Janardan distribution (2.3) can be obtained as

$$\mu_r' = E\left[E\left(X^r \mid \lambda\right)\right] \qquad (3.1)$$

From (2.2) we thus get

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

$$\mu_r' = \int_0^\infty \left[\sum_{x=0}^{\infty} x^r \frac{e^{-\lambda}\lambda^x}{\Gamma(x+1)} \right] \frac{\theta^2}{\alpha(\theta+\alpha^2)} (1+\alpha\lambda) e^{-\frac{\theta}{\alpha}\lambda} d\lambda \qquad (3.2)$$

Obviously the expression under bracket is the r th moment about origin of the Poisson distribution. Taking $r=1$ in (3.2) and using the mean of the Poisson distribution, the mean of the discrete PJD is obtained as

$$\mu_1' = \frac{\theta^2}{\alpha(\theta+\alpha^2)} \int_0^\infty \lambda(1+\alpha\lambda) e^{-\frac{\theta}{\alpha}\lambda} d\lambda$$

$$= \frac{\theta^2}{\alpha(\theta+\alpha^2)} \left[\frac{\theta\alpha^2 + 2\alpha^4}{\theta^3} \right] = \frac{\alpha(\theta+2\alpha^2)}{\theta(\theta+\alpha^2)} \qquad (3.3)$$

Taking $r=2$ in (3.2) and using the second moment about origin of the Poisson distribution, the second moment about origin of the discrete PJD is obtained as

$$\mu_2' = \frac{\theta^2}{\alpha(\theta+\alpha^2)} \int_0^\infty (\lambda^2+\lambda)(1+\alpha\lambda) e^{-\frac{\theta}{\alpha}\lambda} d\lambda =$$

$$\frac{\theta^2}{\alpha(\theta+\alpha^2)} \left[\frac{\alpha^2(\theta+2\alpha^2)}{\theta^3} + \frac{2\alpha^3(\theta+3\alpha^2)}{\theta^4} \right]$$

$$= \frac{\alpha(\theta+2\alpha^2)}{\theta(\theta+\alpha^2)} + \frac{2\alpha^2(\theta+3\alpha^2)}{\theta^2(\theta+\alpha^2)} \qquad (3.4)$$

Similarly, taking $r=3$ & 4 in (3.2) and using the respective moments of the Poisson distribution, we get finally, after a little simplification, the third and the fourth moments about origin of the discrete PJD as

$$\mu_3' = \frac{\alpha(\theta+2\alpha^2)}{\theta(\theta+\alpha^2)} + \frac{6\alpha^2(\theta+3\alpha^2)}{\theta^2(\theta+\alpha^2)} + \frac{6\alpha^3(\theta+4\alpha^2)}{\theta^3(\theta+\alpha^2)} \qquad (3.5)$$

$$\mu_4' = \frac{\alpha(\theta+2\alpha^2)}{\theta(\theta+\alpha^2)} + \frac{14\alpha^2(\theta+3\alpha^2)}{\theta^2(\theta+\alpha^2)} + \frac{36\alpha^3(\theta+4\alpha^2)}{\theta^3(\theta+\alpha^2)} + \frac{24\alpha^4(\theta+5\alpha^2)}{\theta^4(\theta+\alpha^2)}$$

(3.6)

It can be seen that at $\alpha=1$ these moments reduce to the respective moments of the discrete PLD.

Probability Generating Function: The probability generating function of the discrete PJD is obtained as

$$P_X(t) = E(t^X) = \left(\frac{\theta}{\theta+\alpha}\right)^2 \sum_{x=0}^{\infty} \left(\frac{\alpha t}{\theta+\alpha}\right)^x \left[1 + \frac{\alpha(1+\alpha x)}{\theta+\alpha^2}\right]$$

$$= \left(\frac{\theta}{\theta+\alpha}\right)^2 \left[\sum_{x=0}^{\infty} \left(\frac{\alpha t}{\theta+\alpha}\right)^x + \frac{\alpha}{\theta+\alpha^2} \sum_{x=0}^{\infty} (1+\alpha x) \left(\frac{\alpha t}{\theta+\alpha}\right)^x \right]$$

$$= \frac{\theta^2}{(\theta+\alpha)(\theta+\alpha^2)} \left[\frac{\theta+\alpha+\alpha^2}{\theta+\alpha-\alpha t} + \frac{\alpha^3 t}{(\theta+\alpha-\alpha t)^2} \right]$$

$$= \frac{\theta^2\left[\theta+\alpha(\alpha+1-t)\right]}{(\theta+\alpha^2)(\theta+\alpha-\alpha t)^2} \qquad (3.7)$$

Moment Generating Function: The moment generating function of the discrete PJD is thus obtained as

$$M_X(t) = E(e^{tX}) =$$

$$\frac{\theta^2\left[\theta+\alpha(\alpha+1-e^t)\right]}{(\theta+\alpha^2)(\theta+\alpha-\alpha e^t)^2} \qquad (3.8)$$

IV. ESTIMATION OF PARAMETERS

Moments Estimates of Parameters of PJD: The discrete PJD has two parameters to be estimated and so the first two moments are required to get the estimates of its parameters by the method of moments.

From (3.3) and (3.4) we have

$$\frac{\mu_2' - \mu_1'}{\mu_1'^2} = \frac{2(\theta+3\alpha^2)(\theta+\alpha^2)}{(\theta+2\alpha^2)^2} = K(\text{say}) \qquad (4.1)$$

Taking $\theta = b\alpha^2$ in (4.1), we get

$$\frac{2(b+3)(b+1)}{(b+2)^2} = K(\text{say})$$

which gives a quadratic equation in b as

$$(2-K)b^2 + (8-4K)b + (6-4K) = 0 \qquad (4.2)$$

Replacing the first two population moments by the respective sample moments in (4.1), an estimate k of K can be obtained and using it in (4.2), an estimate $\hat{b}$ of b can be obtained. Again substituting $\theta = b\alpha^2$ in the expression for mean we get $\frac{b+2}{\alpha b(b+1)} = \bar{X}$, and thus an estimate of α is obtained as

$$\hat{\alpha} = \frac{b+2}{b(b+1)\bar{X}} \qquad (4.3)$$

Finally, an estimate of θ can be obtained as

$$\theta = b\hat{\alpha}^2 = \frac{(b+2)^2}{b(b+1)^2(\bar{X})^2} \qquad (4.4)$$

Maximum Likelihood Estimates of Parameters of PJD: Let $x_1, x_2, \ldots, x_n$ be a random sample of size n from the discrete PJD (2.3) 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 discrete PJD (2.3) is given by

$$L=\left(\frac{\theta}{\theta+\alpha}\right)^{2n}\left(\frac{1}{\theta+\alpha^2}\right)^{n}\left(\frac{\alpha}{\theta+\alpha}\right)^{\sum_{x=1}^{k}x f_x}\prod_{x=1}^{k}\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]^{f_x}$$

and so the log likelihood function is obtained as

$$\log L=2n\log\left(\frac{\theta}{\theta+\alpha}\right)-n\log\left(\theta+\alpha^2\right)+\sum_{x=1}^{k}x f_x\log\left(\frac{\alpha}{\theta+\alpha}\right)+\sum_{x=1}^{k}f_x\log\left[\theta+\alpha^2+\alpha(1+\alpha x)\right] \quad (4.5)$$

The two log likelihood equations are thus obtained as

$$\frac{\partial\log L}{\partial\theta}=\frac{2n\alpha}{\theta(\theta+\alpha)}-\frac{n}{\theta+\alpha^2}-\sum_{x=1}^{k}\frac{x f_x}{\theta+\alpha}+\sum_{x=1}^{k}\frac{f_x}{\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]}=0$$

(4.6)

$$\frac{\partial\log L}{\partial\alpha}=-\frac{2n}{\theta+\alpha}-\frac{2n\alpha}{\theta+\alpha^2}+\sum_{x=1}^{k}\frac{\theta x f_x}{\alpha(\theta+\alpha)}+\sum_{x=1}^{k}\frac{f_x\left[1+2\alpha(1+x)\right]}{\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]}=0$$

(4.7)

The two equations (4.5) and (4.6) do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. For, we have

$$\frac{\partial^2\log L}{\partial\theta^2}=-\frac{2n\alpha(\alpha+2\theta)}{\theta^2(\theta+\alpha)^2}+\frac{n}{\left(\theta+\alpha^2\right)^2}+\sum_{x=1}^{k}\frac{x f_x}{(\theta+\alpha)^2}-\sum_{x=1}^{k}\frac{f_x}{\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]^2}$$

(4.8)

$$\frac{\partial^2\log L}{\partial\theta\partial\alpha}=\frac{2n}{(\theta+\alpha)^2}+\frac{2n\alpha}{\left(\theta+\alpha^2\right)^2}+\sum_{x=1}^{k}\frac{x f_x}{(\theta+\alpha)^2}-\sum_{x=1}^{k}\frac{f_x\left[1+2\alpha(1+x)\right]}{\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]^2}$$

(4.9)

$$\frac{\partial^2\log L}{\partial\alpha^2}=\frac{2n}{(\theta+\alpha)^2}-\frac{2n\left(\theta-\alpha^2\right)}{\left(\theta+\alpha^2\right)^2}+\sum_{x=1}^{k}\frac{\theta(\theta+2\alpha)x f_x}{\alpha^2(\theta+\alpha)^2}+\sum_{x=1}^{k}\frac{2(1+x)f_x}{\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]}-\sum_{x=1}^{k}\frac{\left[1+2\alpha(1+x)\right]f_x}{\left[\theta+\alpha^2+\alpha(1+\alpha x)\right]^2} \quad (4.10)$$

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

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

(4.11)

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

V. GOODNESS OF FIT

The discrete PJD has been fitted to a number of data- sets to which earlier the PLD has been fitted and it is found that to almost all these data-sets, the discrete PJD provides closer fits than those by the PLD. Here the fittings of the discrete PJD to two such data-sets have been presented in the following tables. The first data set is due to Kemp and Kemp (1965) which is regarding the distribution of mistakes in copying groups of random digits, the second is due to Beall (1940) regarding the distribution of Pyrausta nublilalis in 1937.

The expected frequencies according to the discrete PLD have also been given in these tables for ready comparison with those obtained by discrete PJD. The estimates of the parameters have been obtained by the method of moments. Cochran (1952, 1954) suggested that for goodness- of- fit tests of unimodal distributions the minimum expected frequency can be as low as 1. In calculating the degrees of freedom (d.f) for chi-square goodness-of-fit tests, the expected frequencies less than 1 are combined with adjacent categories to get the minimum expected frequency of 1.

Table 1: Distribution of mistakes in copying groups of random digits

No. of errors per group	Observed frequency	Expected frequency PLD	PJD
0	35	33.1	32.4
1	11	15.3	15.8
2	8	6.8	7.0
3	4	2.9	2.9
4	2	1.2	1.9
Total	60	59.3	60.0
$\bar{X}=0.7833$, $m_2'=1.8500$		$\hat{\theta}=1.7434$	$\hat{\theta}=5.223371$, $\hat{\alpha}=2.612042$
χ^2		2.20	2.10
d.f.		3	2

Table 2: Distribution of Pyrausta nublilalis in 1937

Number of Insects	Observed frequency	Expected frequency PLD	PJD
0	33	31.5	31.9
1	12	14.2	13.8
2	6	6.1	5.9
3	3	2.5	2.5
4	1	1.0	1.1
5	1	0.7	0.8
Total	56	56.0	56.0
$\bar{X}=0.7500$, $m_2'=1.8571$		$\hat{\theta}=1.8081$	$\hat{\theta}=0.392490$, $\hat{\alpha}=0.257251$
χ^2		0.53	0.33
d.f.		3	2

It can be seen that the discrete PJD gives much closer fits than the discrete PLD and thus provides a better alternative to the discrete PLD for modeling different types of count data.

REFERENCES

1. Beall, G. (1940): The fit and significance of contagious distributions when applied to observations on larval insects, Ecology, Vol. 21, 460-474
2. Cochran, W. G. (1952): The χ^2 test of goodness of fit, Annals of Mathematical Statistics, Vol. 23, pp. 315- 345.
3. Cochran, W.G. (1954): Some methods for strengthening the common χ^2 tests, Biometrics, Vol. 10, pp. 417 - 451.
4. Ghitany, M. E., and Al-Mutairi, D.K. (2009): Estimation Methods for the discrete Poisson-Lindley distribution, Journal of Statistical Computation and Simulation, Vol.79 (1), 1 – 9.
5. Kemp, C.D. and Kemp, A.W. (1965): Some properties of the Hermite distribution, Biometrika, Vol. 52, 381-394.
6. Lindley, D. V. (1958): Fiducial distributions and Bayes theorem, Journal of Royal Statistical Society, Ser. B, Vol.20, 102-107
7. Sankaran, M. (1970): The discrete Poisson-Lindley distribution, Biometrics, Vol. 26, 145-149.
8. Shanker, R., Sharma, S., Shanker, U., and Shanker, R. (2013): Janardan distribution and its Applications to Waiting times data, Indian Journal of Applied Research, Vol. 3, Issue 8, pp. 500 – 502.