Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 9, Number 4 (2013), pp. 325-334

http://www.ripublication.com/gjpam.htm

Size-Biased Discrete Two Parameters Poisson-Lindley Distribution and its Applications

Rama Shanker

Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea
Email-shankerrama2009@gmail.com

Abstract

In this paper, a size- biased discrete two parameters Poisson- Lindley distribution, of which size- biased discrete one parameter Poisson- Lindley distribution is a particular case, has been obtained by size biasing the discrete two parameters Poisson - Lindley distribution of Shanker et al (2012). The first four moments of the distribution have been obtained and the estimation of its parameters has been discussed using the method of maximum likelihood and the method of moments. The distribution has been fitted to some data sets to test its goodness of fit.

Keywords: Poisson-Lindley distribution, size-biased distribution, estimation of parameters, goodness of fit.

1. INTRODUCTION

The size biased discrete one parameter Poisson -Lindley distribution given by its probability mass function (p.m.f.)

$$f(x,\theta)=\frac{\theta^3}{\theta+2}\cdot\frac{x(x+\theta+2)}{(\theta+1)^{x+2}}\;;\theta>0,x=1,2,3,... \qquad (1.1)$$

has been introduced by Ghitany et al (2008) by size biasing the discrete Poisson-Lindley distribution of Sankaran (1970). Its various properties and applications has been discussed by Ghitany et al (2008) and showed that (1.1) can also be obtained from the size biased Poisson distribution with p.m.f.

$$P(x,\lambda)=\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\;\;;\lambda>0,x=1,2,3,... \qquad (1.2)$$

when its parameter λ follows the size biased Lindley distribution with probability density function (p.d.f.)

$$f(x,\theta)=\frac{\theta^3}{\theta+2}x(1+x)e^{-\theta x}\ ;\theta>0,x>0 \tag{1.3}$$

The first four moments about origin and the variance of the size-biased discrete one parameter Poisson-Lindley distribution have been obtained as

$$\mu_1'=1+\frac{2(\theta+3)}{\theta(\theta+2)} \tag{1.4}$$

$$\mu_2'=1+\frac{6(\theta+3)}{\theta(\theta+2)}+\frac{6(\theta+4)}{\theta^2(\theta+2)} \tag{1.5}$$

$$\mu_3'=1+\frac{14(\theta+3)}{\theta(\theta+2)}+\frac{36(\theta+4)}{\theta^2(\theta+2)}+\frac{24(\theta+5)}{\theta^3(\theta+2)} \tag{1.6}$$

$$\mu_4'=1+\frac{30(\theta+3)}{\theta(\theta+2)}+\frac{126(\theta+4)}{\theta^2(\theta+2)}+\frac{240(\theta+5)}{\theta^3(\theta+2)}+\frac{120(\theta+6)}{\theta^4(\theta+2)} \tag{1.7}$$

$$\mu_2=\frac{2\left(\theta^3+6\theta^2+12\theta+6\right)}{\theta^2(\theta+2)^2} \tag{1.8}$$

Ghitany et al (2008) showed that in many ways (1.1) is a better model for some applications than size-biased Poisson distribution.

In this paper, a size-biased discrete two parameters Poisson-Lindley distribution, of which the size biased discrete one-parameter Poisson-Lindley distribution (1.1) is a particular case, has been obtained by compounding the size - biased Poisson distribution with the size biased two parameters Lindley distribution of Shanker et at (2013). The first four moments about origin of this distribution have been obtained and the estimation of its parameters has been discussed. The distribution has been fitted to some data sets and it has been found that to almost all these data- sets it provides closer fits than the size -biased one parameter Poisson-Lindley distribution. This shows that the size-biased discrete two parameters Poisson-Lindley distribution is more flexible than the Ghitany et al (2008) size-biased one parameter Poisson-Lindley distribution for analyzing different types of count data.

2. SIZE-BIASED DISCRETE TWO PARAMETERS POISSON-LINDLEY DISTRIBUTION

The discrete two parameters Poisson-Lindley distribution (PLD) given by its probability mass function (p.m.f)

$$f_0(x;\theta,\alpha)=\frac{\theta^2}{(\theta+1)^{x+2}}\left(1+\frac{\alpha x+1}{\theta+\alpha}\right)\ ;x=0,1,2,...;\ \theta>0,\alpha>-\theta \tag{2.1}$$

has been introduced by Shanker et al (2012) to model count data. It can be easily seen that the discrete one parameter Poisson-Lindley distribution (1.1) of Sankaran (1970) is a particular case for $\alpha = 1$. Shanker et al (2013) has shown that (2.1) is a better model than the one parameter discrete Poisson-Lindley distribution of sankaran (1970) for analyzing different types of count data. The distribution arises from the Poisson distribution when its parameter λ follows the Shanker et al (2013) two parameters Lindley distribution for modeling waiting and survival times data with its probability density function (p.d.f)

$$g_0(x;\theta,\alpha) = \frac{\theta^2}{\theta+\alpha}(1+\alpha x)e^{-\theta x}; x > 0, \theta > 0, \alpha > -\theta \tag{2.2}$$

It can be seen that at $\alpha = 1$, (2.2) reduces to the one parameter Lindley distribution (1958). Shanker et al (2013) has shown that (2.2) is a better model than the one parameter Lindley distribution for analyzing waiting time, survival time and grouped mortality data. A size-biased discrete two parameters Poisson-Lindley distribution with parameters α and θ is defined by its probability mass function (p.m.f)

$$P(X=x) = \frac{x f_0(x;\theta,\alpha)}{\mu_1'} = \frac{\theta^3}{\theta+2\alpha}\cdot\frac{x(\alpha x+\theta+\alpha+1)}{(\theta+1)^{x+2}}\ ; x = 1,2,3,\ldots, \theta > 0, \alpha > -\frac{\theta}{2} \tag{2.3}$$

where $\mu_1' = \frac{\theta+2\alpha}{\theta(\theta+\alpha)}$ is the mean of the discrete two parameters Poisson-Lindley distribution given by (2.1). It can be easily seen that (2.3) reduces to (1.1) at $\alpha = 1$ .The size-biased discrete two parameters Poisson-Lindley distribution can also be obtained from the size -biased Poisson distribution with p.m.f. (1.2) when its parameter λ follows a size-biased two parameters Lindley distribution of Shanker et al (2013) with p.d.f.

$$f(\lambda;\theta,\alpha) = \frac{\theta^3}{\theta+2\alpha}\lambda\,(1+\lambda\alpha)e^{-\theta\lambda} \quad ; \lambda > 0,\ \theta > 0, \alpha > -\frac{\theta}{2} \tag{2.4}$$

We have

$$P(X=x) = \int_0^\infty g(x|\lambda) f(\lambda;\theta,\alpha)\,d\lambda = \int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\cdot\frac{\theta^3}{\theta+2\alpha}\lambda(1+\lambda\alpha)e^{-\theta\lambda}d\lambda \tag{2.5}$$

$$= \frac{\theta^3}{\theta+2\alpha}\frac{1}{(x-1)!}\int_0^\infty e^{-(\theta+1)\lambda}\left(\lambda^x+\alpha\lambda^{x+1}\right)d\lambda$$

$$= \frac{\theta^3}{\theta+2\alpha}\left[\frac{x}{(\theta+1)^{x+1}} + \frac{\alpha x(x+1)}{(\theta+1)^{x+2}}\right]$$

$$= \frac{\theta^3}{\theta+2\alpha}\frac{x(\alpha x+\theta+\alpha+1)}{(\theta+1)^{x+2}} \quad ; x = 1,2,3,\ldots \tag{2.6}$$

which is the size-biased discrete two parameters Poisson-Lindley distribution.

Since

$$\frac{P(x+1;\theta,\alpha)}{P(x;\theta,\alpha)}=\frac{1}{\theta+1}\left(1+\frac{1}{x}\right)\left(1+\frac{\alpha}{\alpha x+\theta+\alpha+1}\right) \tag{2.7}$$

is a decreasing function in x and thus $P(x;\theta,\alpha)$ is log-concave. Therefore, the size-biased discrete two parameters Poisson-Lindley distribution is unimodal, has an increasing failure rate (IFR), and hence increasing failure rate average (IFRA), new better than used (NBU), new better than used in expectation (NBUE) and decreasing mean residual life (DMRL), see Barlow and Proschan (1981) for more details about the definition of these ageing concepts.

3. MOMENTS

The rth moment about origin of the size- biased discrete two parameters Poisson-Lindley distribution (2.3) can be obtained as

$$\mu_r' = E\left[\,E\left(X^r/\lambda\right)\right] \tag{3.1}$$

From (2.5), we thus get

$$\mu_r' = \int_0^\infty \left[\sum_{x=0}^\infty x^r \,\frac{e^{-\lambda}\,\lambda^{x-1}}{(x-1)!}\right]\frac{\theta^3}{\theta+2\alpha}\lambda(1+\lambda\alpha)e^{-\theta\lambda}\,d\lambda \tag{3.2}$$

Obviously the expression under bracket is the rth moment about origin of the size-biased Poisson distribution. Taking r = 1 in (3.2) and using the mean of the size - biased Poisson distribution, the mean of the size-biased discrete two parameters Poisson- Lindley distribution is obtained as

$$\mu_1' = \frac{\theta^3}{\theta+2\alpha}\int_0^\infty (\lambda+1)\lambda(1+\lambda\alpha)\,e^{-\theta\lambda}\,d\lambda$$

$$=\frac{\theta^3}{(\theta+2\alpha)}\left[\frac{6\alpha}{\theta^4}+\frac{2(\alpha+1)}{\theta^3}+\frac{1}{\theta^2}\right]=1+\frac{2(\theta+3\alpha)}{\theta(\theta+2\alpha)} \tag{3.3}$$

Again, taking r = 2 in (3.2) and using the second moment about origin of the size-biased Poisson distribution, the second moment about origin of the size - biased discrete two parameters Poisson -Lindley distribution is obtained as

$$\mu_2' = \frac{\theta^3}{\theta+2\alpha}\int_0^\infty \left(\lambda^2+3\lambda+1\right)\lambda(1+\lambda\alpha)\,e^{-\theta\lambda}\,d\lambda$$

$$=\frac{\theta^3}{(\theta+2\alpha)}\left[\frac{24\alpha}{\theta^5}+\frac{6(3\alpha+1)}{\theta^4}+\frac{2(\alpha+3)}{\theta^3}+\frac{1}{\theta^2}\right]$$

$$=1+\frac{6\left(\theta+3\alpha\right)}{\theta\left(\theta+2\alpha\right)}+\frac{6\left(\theta+4\alpha\right)}{\theta^2\left(\theta+2\alpha\right)} \tag{3.4}$$

Similarly, taking r = 3 & 4 and using the respective moments of the size- biased Poisson distribution, we get finally after a little simplification, the third and the fourth moments about origin of the size- biased discrete two parameters Poisson- Lindley distribution as

$$\mu_3'=1+\frac{14\left(\theta+3\alpha\right)}{\theta\left(\theta+2\alpha\right)}+\frac{36\left(\theta+4\alpha\right)}{\theta^2\left(\theta+2\alpha\right)}+\frac{24\left(\theta+5\alpha\right)}{\theta^3\left(\theta+2\alpha\right)} \tag{3.5}$$

$$\mu_4'=1+\frac{30\left(\theta+3\alpha\right)}{\theta\left(\theta+2\alpha\right)}+\frac{126\left(\theta+4\alpha\right)}{\theta^2\left(\theta+2\alpha\right)}+\frac{240\left(\theta+5\alpha\right)}{\theta^3\left(\theta+2\alpha\right)}+\frac{120\left(\theta+6\alpha\right)}{\theta^4\left(\theta+2\alpha\right)} \tag{3.6}$$

It can be seen that at $\alpha=1$, these moments reduce to the respective moments of the size- biased discrete one parameter Poisson-Lindley distribution.

4. ESTIMATION OF PARAMETERS

4.1. Maximum Likelihood Estimates: Let $(x_1, x_2, \ldots, x_n)$ be a random sample of size n from the size- biased discrete two parameters Poisson - Lindley distribution (2.3) and let f_x be the observed frequency in the sample corresponding to $X=x$ $\left(x=1,2,...,k\right)$ 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 size- biased discrete two parameters Poisson-Lindley distribution (2.3) is given by

$$L=\left(\frac{\theta^3}{\theta+2\alpha}\right)^n \frac{1}{\left(\theta+1\right)^{\sum_{x=1}^{k}(x+2)f_x}} \prod_{x=1}^{k}\left[\alpha x^2+x\left(\theta+\alpha+1\right)\right]^{f_x} \tag{4.1.1}$$

and so the log likelihood function is obtained as

$$\log L = n\log\left(\frac{\theta^3}{\theta+2\alpha}\right)-\sum_{x=1}^{k} f_x\left(x+2\right)\log\left(\theta+1\right)+\sum_{x=1}^{k} f_x \log\left[\alpha x^2+x\left(\theta+\alpha+1\right)\right] \tag{4.1.2}$$

The two log likelihood equations are thus obtained as

$$\frac{\partial \log L}{\partial\theta}=\frac{3n}{\theta}-\frac{n}{\theta+2\alpha}-\sum_{x=1}^{k}\frac{\left(x+2\right)f_x}{\theta+1}+\sum_{x=1}^{k}\frac{x f_x}{\left[\alpha x^2+x\left(\theta+\alpha+1\right)\right]} \tag{4.1.3}$$

$$\frac{\partial \log L}{\partial\alpha}=\frac{-2n}{\theta+2\alpha}+\sum_{x=1}^{k}\frac{\left(x^2+x\right)f_x}{\left[\alpha x^2+x\left(\theta+\alpha+1\right)\right]} \tag{4.1.4}$$

The two equations (4.1.3) and (4.1.4) do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. For, we have

$$\frac{\partial^2 \log L}{\partial \theta^2} = \frac{-3n}{\theta^2} + \frac{n}{(\theta+2\alpha)^2} + \sum_{x=1}^{k} \frac{(x+2) f_x}{(\theta+1)^2} - \sum_{x=1}^{k} \frac{x^2 f_x}{\left[\alpha x^2 + x(\theta+\alpha+1)\right]^2} \tag{4.1.5}$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = \frac{2n}{(\theta+2\alpha)^2} - \sum_{x=1}^{k} \frac{x(x^2+x) f_x}{\left[\alpha x^2 + x(\theta+\alpha+1)\right]^2} \tag{4.1.6}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{4n}{(\theta+2\alpha)^2} - \sum_{x=1}^{k} \frac{(x^2+x)^2 f_x}{\left[\alpha x^2 + x(\theta+\alpha+1)\right]^2} \tag{4.1.7}$$

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

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \log L}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta}-\theta_0 \\ \hat{\alpha}-\alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \tag{4.1.8}$$

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.

4.2. Estimates from Moments: The size- biased discrete two parameters Poisson-Lindley distribution 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

$$\frac{(\mu_2'-1)-3(\mu_1'-1)}{(\mu_1'-1)^2} = K\,(\text{say}) = \frac{3(\theta+4\alpha)(\theta+2\alpha)}{2(\theta+3\alpha)^2} \tag{4.2.1}$$

Taking $\alpha = b\theta$ in (4.2.1), we get

$$\frac{3(1+4b)(1+2b)}{2(1+3b)^2} = K \tag{4.2.2}$$

which gives a quadratic equation in b as

$$g(b) = (24-18K)b \;\; + (18-12K)b + (3-2K) = 0\,. \tag{4.2.3}$$

Replacing the first two population moments by the respective sample moments in (4.2.1) an estimate k of K can be obtained and using it in (4.2.3), an estimate $\hat{b}$ of b can be obtained.

Again substituting $\alpha = b\theta$ in the expression for mean we get $\bar{X} - 1 = \frac{2(1+3b)}{\theta(1+2b)}$, and thus an estimate of θ is obtained as

$$\hat{\theta} = \frac{2(1+3b)}{(1+2b)(\bar{X}-1)} \tag{4.2.4}$$

Finally an estimate of α can be obtained as

$$\hat{\alpha} = b\hat{\theta} = \frac{2b(1+3b)}{(1+2b)(\bar{X}-1)} \tag{4.2.5}$$

5. GOODNESS OF FIT

The size-biased discrete two parameters Poisson-Lindley distribution has been fitted to a number of data- sets and it was found that to almost all these data -sets, the size-biased discrete two parameters Poisson- Lindley distribution provides closer fit than the size-biased discrete one parameter Poisson-Lindley distribution. Here the fittings of the size-biased discrete two parameters Poisson-Lindley distribution to three such data-sets have been presented in the following tables.

The first data sets in table 1 are due to Finney and Varley (1955) who gave counts of flower heads with 1, 2, 3, …, 9 fly eggs having corresponding counts 22, 18, 18, 11, 9, 6, 3, 0, and 1.

Table-1: Counts of flower heads with number of fly eggs

Number of fly eggs	Observed frequency	Expected frequency: Size-biased one parameter Poisson-Lindley distribution	Expected frequency: Size-biased two parameters Poisson-Lindley distribution
1	22	20.4	20.3
2	18	22.0	21.5
3	18	17.2	17.0
4	11	11.6	11.8
5	9	7.3	7.6
6	6	4.3	4.6
7	3	2.4	2.7
8	0	1.3	1.5
9	1	1.5	1.0
Total	88	88.0	88.0
Moments Estimates of Parameters		$\hat{\theta} = 1.2828$	$\hat{\theta} = 0.7238$, $\hat{\alpha} = -0.1250$
χ^2		1.342	1.087
d.f		4	3

Table-2: Groups observed on a Spring Afternoon in Portland, Oregon

Size of group	Observed frequency	Expected frequency Size-biased one parameter Poisson-Lindley distribution	Expected frequency Size-biased two parameter Poisson-Lindley distribution
1	1486	1532.5	1504.4
2	694	630.6	658.0
3	195	191.9	203.1
4	37	51.3	49.9
5	10	12.8	8.9
6	1	3.9	1.3
Total	2423	2423.0	2423.0
Moments Estimates of Parameters		$\hat{\theta} = 4.5082$	$\hat{\theta} = 2.6784$, $\hat{\alpha} = -0.5172$
χ^2		13.766	5.916
d.f		3	2

Table-3: Counts of the pair of running shoes owned by 60 members of an athletics club

Number of pair of running shoes	Observed frequency	Expected frequency Size-biased one parameter Poisson-Lindley distribution	Expected frequency Size-biased two parameters Poisson-Lindley distribution
1	18	20.4	16.8
2	18	17.4	16.8
3	12	10.9	12.4
4	7	5.9	7.8
5	5	5.4	6.2
Total	60	60.0	60.0
Moments Estimates of Parameters		$\hat{\theta} = 1.8243$	$\hat{\theta} = 0.7277$, $\hat{\alpha} = -0.1813$
χ^2		0.649	0.499
d.f		3	2

The second data sets in table 2 are due to Coleman and James (1961) who gave the counts of groups of people in public places on a spring afternoon in Portland. For group sizes 1, 2, 3, 4, 5 and 6, the counts were 1486, 694, 195, 37, 10, and 1.

The third data sets in table 3 are due to Simonoff (2003) who gave counts of the pair of running shoes owned by 60 members of an athletics club.

The expected frequencies according to the size-biased discrete one parameter Poisson-Lindley distribution have also been given in these tables for ready comparison with those obtained by the size-biased discrete two parameters Poisson-Lindley distribution. The estimates of the parameters have been obtained by the method of moments.

It can be seen that the size-biased discrete two parameters Poisson-Lindley distribution gives much closer fits than the size-biased discrete one parameter Poisson-Lindley distribution and thus provides a better alternative to the size-biased discrete one parameter Poisson-Lindley distribution.

8. CONCLUSION

In this paper, we propose a size-biased discrete two parameters Poisson- Lindley distribution, of which the size-biased discrete one parameter Poisson-Lindley distribution is a particular case, to model various types of count data. Some of its properties such as modality, increasing failure rate etc. and the estimation of parameters by the method of maximum likelihood and the method of moments have been discussed. Finally, the proposed distribution has been fitted to a number of count data to test its goodness of fit to which earlier the size-biased discrete one parameter Poisson-Lindley distribution has been fitted and it is found that size-biased discrete two parameters Poisson-Lindley distribution provides better fits than those by the size-biased discrete one parameter Poisson-Lindley distribution.

REFERENCES

[1] Barlow, R.E. and Proschan, F. (1981): *Statistical Theory of Reliability and Life Testing*, Silver Spring, MD

[2] Coleman, J.S. and James, J. (1961): The equilibrium size distribution of freely forming groups, *Sociometry*, Vol. 24, No. 1, pp. 36 – 45

[3] D.V. Lindley (1958): Fiducial distributions and Bayes' theorem, *Journal of the Royal Statistical Society, Series B*, Vol. 20, pp. 102- 107.

[4] Finnay, D.J. and Varley, G.C. (1955): An example of the truncated Poisson distribution, *Biometrics*, Vol. 11, No. 3, pp. 387 – 394.

[5] Ghitany, M.E., Al-Mutairi, D.K. (2008): Size- biased Poisson-Lindley distribution and its applications, *Metron*, Vol. LXVI, No. 3, pp. 299 - 311

[6] Ghitany, M.E., Atieh, B, and Nadarajah, S. (2008): Lindley distribution and its Applications, *Mathematics and Computers in Simulation*, Vol.78 No.4, pp.493 – 506.

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

[8] Shanker, R., Sharma, S., and Shanker, R. (2013): A two parameter Lindley distribution for modeling waiting and survival times data, *Applied Mathematics*, Vol. 4 (2), pp.363 – 368, 2013, U.S.A

[9] Shanker, R., Sharma, S., and Shanker, R. (2012): A discrete two parameter Poisson Lindley distribution, *Journal of Ethiopian Statistical Association (JESA)*, Vol. 21, 2013, pp. 22 – 29.

[10] Simonoff, J.S. (2003): Analyzing categorical data, New York Springer