American Journal of Mathematics and Statistics 2017, 7(3): 99-107
DOI: 10.5923/j.ajms.20170703.02

On Size- Biased Two Parameter Poisson-Lindley Distribution and Its Applications

Rama Shanker[1,*], A. Mishra[2]

[1]Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea
[2]Department of Statistics, Patna University, Patna, India

Abstract A size - biased version of the two parameter Poisson- Lindley distribution introduced by Shanker and Mishra (2014) has been proposed of which the Ghitany and Al Mutairi's (2008) size - biased one parameter Poisson-Lindley distribution is a particular case. A general expression for its *r*th factorial moment about origin has been derived and hence its raw moments and central moments are obtained. The expressions for its coefficient of variation, skewness, kurtosis and index of dispersion have also been given. The method of maximum likelihood and the method of moments for the estimation of its parameters have been discussed. The applications and the goodness of fit of the proposed distribution have been discussed with three data sets excluding zero counts and the fit has been compared with that of size-biased Poisson and size-biased Poisson-Lindley distributions.

Keywords Size-biased distributions, Two-parameter Poisson-Lindley distribution, Poisson-Lindley distribution, Size-biased distributions, Moments, Estimation of Parameters, Goodness of fit

1. Introduction

The size - biased distributions arise when the observations generated from a random process do not have equal probability of being recorded and are recorded according to some weight function. When the sampling mechanism is such that the sample units are selected with probability proportional to some measure of the unit size, the resulting distribution is called 'size-biased distribution'. Fisher (1934) first introduced such distributions to model ascertainment bias and Rao (1965) formulated these in a unifying theory. Patil and Ord (1975) studied the size-biased sampling and the related form-invariant weighted distribution whereas Van Deusen (1986) arrived at size - biased distribution theory independently and applied it to fitting distributions of diameter at breast height (DBH) data arising from horizontal point sampling (HPS). Later, Lappi and Bailey (1987) analyzed HPS diameter increment data using size- biased distribution. Patil and Rao (1977, 1978) examined some general models leading to size - biased distributions. The results were applied to the analysis of data relating to human populations and wild life management. Gove (2003) reviewed some of the recent results on size- biased distributions pertaining to parameter estimation in forestry with special emphasis on Weibull distribution. Simoj and Maya (2006) introduced some fundamental relationships between weighted and unique variables in the context of maintainability function and inverted repair rate. Mir and Ahmad (2009), Das and Roy (2011) and Ducey and Gove (2015) have also studied the various aspects of size - biased distributions.

A simple size-biased version of a distribution $f(x;\theta)$ is given by its probability function $f^*(x;\theta) = x\,f(x;\theta)/\mu_1'$ where μ_1' is the mean of the distribution.

Ghitany and Al Mutairi (2008) obtained a size-biased Poisson-Lindley distribution (SBPLD) given by its probability mass function (p.m.f.)

$$P_1(x,\theta)=\frac{\theta^3}{\theta+2}\cdot\frac{x(x+\theta+2)}{(\theta+1)^{x+2}}\;;\;\theta>0,\quad x=1,2,3,... \tag{1.1}$$

by size biasing the Poisson -Lindley distribution of Sankaran (1970) having pmf

$$P_2(x;\theta)=\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}\;;\;x=0,1,2,\ldots,\quad \theta>0. \tag{1.2}$$

It is to be mentioned that Sankaran (1970) obtained the distribution (1.2) by mixing the Poisson distribution with the Lindley (1958) distribution having pdf

$$f_1(x,\theta)=\frac{\theta^2}{\theta+1}(1+x)\,e^{-\theta x}\;;\quad x>0,\quad \theta>0 \tag{1.3}$$

* Corresponding author:
shankerrama2009@gmail.com (Rama Shanker)
Published online at http://journal.sapub.org/ajms

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

$$\mu_1' = \frac{\theta^2 + 4\theta + 6}{\theta(\theta+2)} \tag{1.4}$$

$$\mu_2' = \frac{\theta^3 + 8\theta^2 + 24\theta + 24}{\theta^2(\theta+2)} \tag{1.5}$$

$$\mu_3' = \frac{\theta^4 + 16\theta^3 + 78\theta^2 + 168\theta + 120}{\theta^3(\theta+2)} \tag{1.6}$$

$$\mu_4' = \frac{\theta^5 + 32\theta^4 + 240\theta^3 + 840\theta^2 + 1320\theta + 720}{\theta^4(\theta+2)} \tag{1.7}$$

and so its variance as

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

Shanker and Mishra (2014) obtained a two-parameter Poisson-Lindley distribution (TPPLD) given by its pmf

$$P_3(x;\theta,\alpha) = \frac{\theta^2}{(\theta+1)^{x+2}}\left(1 + \frac{\alpha + x}{\theta\alpha + 1}\right) \quad ; x = 0,1,2,...; \ \theta > 0, \ \theta\alpha > -1 \tag{1.9}$$

having its mean,

$$\mu_1' = \frac{\theta\alpha + 2}{\theta(\theta\alpha + 1)} \tag{1.10}$$

It can be seen that the PLD (1.2) is a particular case of it at $\alpha = 1$. Shanker and Mishra (2014) have shown that (1.9) is a better model than the PLD of Sankaran (1970) for analyzing different types of count data. This distribution arises from the Poisson distribution when its parameter λ follows the Shanker and Mishra (2013) two parameter Lindley distribution having probability density function (p.d.f)

$$f_2(x;\theta,\alpha) = \frac{\theta^2}{\theta\alpha + 1}(\alpha + x)e^{-\theta x}; \quad x > 0, \ \theta > 0, \theta\alpha > -1 \tag{1.11}$$

In this paper, a size -biased two parameter Poisson-Lindley distribution (SBTPPLD), of which the SBPLD (1.1) is a particular case, has been obtained. A general expression for its *r*th factorial moment about origin has been obtained and hence its first four moments about origin and central moments are obtained. The expressions for coefficients of variation, skewness, kurtosis and index of dispersion have also been given. The method of maximum likelihood and the method of moments for the estimation of its parameters have been discussed. The distribution has been fitted to some data sets to show that it provides closer fit than the size-biased Poisson distribution (SBPD) and SBPLD. This makes one believe that SBTPPLD is more flexible than the SBPD and SBPLD for analyzing different count data.

2. A Size- Biased Two Parameter Poisson-Lindley Distribution

A size -biased version of the two parameter Poisson-Lindley distribution (SBTPPLD) with parameters α and θ can be obtained as

$$P_4(x;\theta,\alpha) = \frac{xP_3(x;\theta,\alpha)}{\mu_1'} \tag{2.1}$$

Taking $P_3(x;\theta,\alpha)$ from (1.9) and μ_1' from (1.10), we get

$$P_4(x;\theta,\alpha)=\frac{\theta^3}{\theta\alpha+2}\frac{x(x+\theta\alpha+\alpha+1)}{(\theta+1)^{x+2}}\ ;\ x=1,2,3,...,\ \theta>0,\ \theta\alpha>-2 \tag{2.2}$$

It can be easily seen that at $\alpha=1$, SBTPPLD (2.2) reduces to SBPLD (1.1). The SBTPPLD (2.2) can also be obtained from the size- biased Poisson distribution when its parameter λ follows a size-biased two parameter Lindley distribution of Shanker and Mishra (2013) with p.d.f.

$$f_3(\lambda;\theta,\alpha)=\frac{\theta^3}{\theta\alpha+2}\lambda(\alpha+\lambda)e^{-\theta\lambda}\quad ;\lambda>0,\ \theta>0,\ \theta\alpha>-2 \tag{2.3}$$

We have

$$P(X=x)=\int_0^\infty\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\cdot\frac{\theta^3}{\theta\alpha+2}\lambda(\alpha+\lambda)e^{-\theta\lambda}d\lambda \tag{2.4}$$

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

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

$$=\frac{\theta^3}{\theta\alpha+2}\frac{x(x+\theta\alpha+\alpha+1)}{(\theta+1)^{x+2}}\quad ;x=1,2,3,... \tag{2.5}$$

which is the SBTPPLD, as obtained in (2.2).

Since

$$\frac{P_4(x+1;\theta,\alpha)}{P_4(x;\theta,\alpha)}=\frac{1}{\theta+1}\left(1+\frac{1}{x}\right)\left(1+\frac{1}{x+\theta\alpha+\alpha+1}\right) \tag{2.6}$$

is a decreasing function of x, $P_4(x;\theta,\alpha)$ is log-concave. This implies that the SBTPPLD is unimodal, has an increasing failure rate (IFR) and so increasing failure rate average (IFRA). It is new better than used (NBU), new better than used in expectation (NBUE) and has decreasing mean residual life (DMRL). Details about the definitions and relationship of these aging concepts can be seen in Barlow and Proschan (1981).

3. Moments and Related Measures

The *r*th factorial moment about origin of the SBTPPLD (2.2) can be obtained as

$\mu_{(r)}'=E\left[E\left(X^{(r)}\mid\lambda\right)\right]$, where $X^{(r)}=X(X-1)(X-2)...(X-r+1)$.

From (2.4), we get

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

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

Taking $(x+r)$ in place of x, we get

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

Clearly the expression within the bracket is $(\lambda+r)$ and hence we have

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

Using gamma integral and a little algebraic simplification, a general expression for the *r*th factorial moment about origin of SBTPPLD is obtained as

$$\mu_{(r)}' = \frac{r!\left[r\theta(\theta\alpha+r+1)+(r+1)(\theta\alpha+r+2)\right]}{\theta^r(\theta\alpha+2)}\;;r=1,2,3,... \tag{3.1}$$

Substituting $r=1,2,3,\text{and }4$ in (3.1), the first four factorial moments about origin can be obtained and then using the relationship between factorial moments and moments about origin, the first four moments about origin of SBTPPLD are obtained as

$$\mu_1' = \frac{\alpha(\theta^2+2\theta)+(2\theta+6)}{\theta(\theta\alpha+2)} \tag{3.2}$$

$$\mu_2' = \frac{\alpha(\theta^3+6\theta^2+6\theta)+(2\theta^2+18\theta+24)}{\theta^2(\theta\alpha+2)} \tag{3.3}$$

$$\mu_3' = \frac{\alpha(\theta^4+14\theta^3+36\theta^2+24\theta)+(2\theta^3+42\theta^2+144\theta+120)}{\theta^3(\theta\alpha+2)} \tag{3.4}$$

$$\mu_4' = \frac{\alpha(\theta^5+30\theta^4+150\theta^3+240\theta^2+120\theta)+(2\theta^4+90\theta^3+600\theta^2+1200\theta+720)}{\theta^4(\theta\alpha+2)} \tag{3.5}$$

Using the relationship between moments about mean and the moments about origin, the moments about mean of SBTPPLD can be obtained as

$$\mu_2 = \frac{2\left[\alpha^2(\theta^3+\theta^2)+\alpha(5\theta^2+6\theta)+6(\theta+1)\right]}{\theta^2(\theta\alpha+2)^2} \tag{3.6}$$

$$\mu_3 = \frac{2\begin{bmatrix}\alpha^3(\theta^5+3\theta^4+2\theta^3)+\alpha^2(7\theta^4+24\theta^3+18\theta^2)+\alpha(16\theta^3+54\theta^2+36\theta)\\+12(\theta^2+3\theta+2)\end{bmatrix}}{\theta^3(\theta\alpha+2)^3} \tag{3.7}$$

$$\mu_4 = \frac{2\begin{bmatrix}\alpha^4(\theta^7+13\theta^6+24\theta^5+12\theta^4)+\alpha^3(9\theta^6+130\theta^5+264\theta^4+144\theta^3)+\\\alpha^2(30\theta^5+460\theta^4+936\theta^3+504\theta^2)+\alpha(44\theta^4+696\theta^3+1368\theta^2+720\theta)\\+24(\theta^3+16\theta^2+30\theta+15)\end{bmatrix}}{\theta^4(\theta\alpha+2)^4} \tag{3.8}$$

The coefficient of variation $(C.V)$, coefficient of skewness $(\sqrt{\beta_1})$, coefficient of kurtosis (β_2) and index of dispersion (γ) of SBTPPLD are thus given by

$$C.V=\frac{\sigma}{\mu_1'}=\frac{\sqrt{2\left[\alpha^2\left(\theta^3+\theta^2\right)+\alpha\left(5\theta^2+6\theta\right)+6\left(\theta+1\right)\right]}}{\alpha\left(\theta^2+2\theta\right)+\left(2\theta+6\right)} \tag{3.9}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{\left[\begin{array}{l}\alpha^3\left(\theta^5+3\theta^4+2\theta^3\right)+\alpha^2\left(7\theta^4+24\theta^3+18\theta^2\right)+\alpha\left(16\theta^3+54\theta^2+36\theta\right)\\+12\left(\theta^2+3\theta+2\right)\end{array}\right]}{\sqrt{2}\left[\alpha^2\left(\theta^3+\theta^2\right)+\alpha\left(5\theta^2+6\theta\right)+6\left(\theta+1\right)\right]^{3/2}} \tag{3.10}$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{\left[\begin{array}{l}\alpha^4\left(\theta^7+13\theta^6+24\theta^5+12\theta^4\right)+\alpha^3\left(9\theta^6+130\theta^5+264\theta^4+144\theta^3\right)+\\\alpha^2\left(30\theta^5+460\theta^4+936\theta^3+504\theta^2\right)+\alpha\left(44\theta^4+696\theta^3+1368\theta^2+720\theta\right)\\+24\left(\theta^3+16\theta^2+30\theta+15\right)\end{array}\right]}{2\left[\alpha^2\left(\theta^3+\theta^2\right)+\alpha\left(5\theta^2+6\theta\right)+6\left(\theta+1\right)\right]^2} \tag{3.11}$$

$$\gamma=\frac{\sigma^2}{\mu_1'}=\frac{2\left[\alpha^2\left(\theta^3+\theta^2\right)+\alpha\left(5\theta^2+6\theta\right)+6\left(\theta+1\right)\right]}{\theta\left(\theta\alpha+2\right)\left[\alpha\left(\theta^2+2\theta\right)+\left(2\theta+6\right)\right]} \tag{3.12}$$

It can be seen that at $\alpha=1$ these expressions reduce to the respective expressions of the SBPLD (1.1).

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 SBTPPLD (2.2) 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 SBTPPLD (2.2) is given by

$$L=\left(\frac{\theta^3}{\theta\alpha+2}\right)^n\frac{1}{\left(\theta+1\right)^{\sum_{x=1}^{k}(x+2)f_x}}\prod_{x=1}^{k}\left[x^2+x\left(\theta\alpha+\alpha+1\right)\right]^{f_x} \tag{4.1}$$

and so the log likelihood function is obtained as

$$\log L=n\log\left(\frac{\theta^3}{\theta\alpha+2}\right)-\sum_{x=1}^{k} f_x\left(x+2\right)\log\left(\theta+1\right)+\sum_{x=1}^{k} f_x\log\left[x^2+x\left(\theta\alpha+\alpha+1\right)\right] \tag{4.2}$$

The two log likelihood equations are thus obtained as

$$\frac{\partial\log L}{\partial\theta}=\frac{3n}{\theta}-\frac{n\alpha}{\theta\alpha+2}-\frac{n\left(\bar{x}+2\right)}{\theta+1}+\sum_{x=1}^{k}\frac{\alpha f_x}{\left[x+\left(\theta\alpha+\alpha+1\right)\right]}=0 \tag{4.3}$$

$$\frac{\partial\log L}{\partial\alpha}=\frac{-n\theta}{\theta\alpha+2}+\sum_{x=1}^{k}\frac{\left(\theta+1\right)f_x}{\left[x+\left(\theta\alpha+\alpha+1\right)\right]}=0 \tag{4.4}$$

The two equations (4.3) and (4.4) do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. We have

$$\frac{\partial^2 \log L}{\partial \theta^2} = \frac{-3n}{\theta^2} + \frac{n\alpha^2}{(\theta\alpha+2)^2} + \frac{n(\bar{x}+2)}{(\theta+1)^2} - \sum_{x=1}^{k} \frac{\alpha^2 f_x}{\left[x+(\theta\alpha+\alpha+1)\right]^2} \tag{4.5}$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = \frac{-2n}{(\theta\alpha+2)^2} - \sum_{x=1}^{k} \frac{(x+1) f_x}{\left[x+(\theta\alpha+\alpha+1)\right]^2} \tag{4.6}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = \frac{n\theta^2}{(\theta\alpha+2)^2} - \sum_{x=1}^{k} \frac{(\theta+1)^2 f_x}{\left[x+(\theta\alpha+\alpha+1)\right]^2} \tag{4.7}$$

For the maximum likelihood estimates $(\hat{\theta},\hat{\alpha})$ of (θ,α) of SBTPPLD (2.2), following equations can be solved

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial \theta^2} & \frac{\partial^2 \log L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \log L}{\partial \theta \partial \alpha} & \frac{\partial^2 \log L}{\partial \alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta}-\theta_0 \\ \hat{\alpha}-\alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial \theta} \\ \frac{\partial \log L}{\partial \alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \tag{4.8}$$

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

4.2. Estimates from Moments

The SBTPPLD has two parameters to be estimated and so the first two moments about origin are required to get the estimates of its parameters by the method of moments.

From (3.2) and (3.3) we have

$$\frac{(\mu_2'-1)-3(\mu_1'-1)}{(\mu_1'-1)^2} = \frac{3(\theta\alpha+4)(\theta\alpha+2)}{2(\theta\alpha+3)^2} = K \text{ (say)} \tag{4.9}$$

Taking $\theta\alpha = \beta$, we get

$$\frac{3(\beta+4)(\beta+2)}{2(\beta+3)^2} = K \tag{4.10}$$

which gives

$$\beta = \sqrt{\frac{1}{1-\frac{2}{3}K}} - 3. \tag{4.11}$$

Replacing the first two population moments by the respective sample moments in (4.9) an estimate k of K can be obtained and using it in (4.11), an estimate $\tilde{b}$ of β can be obtained.

Again, substituting $\theta\alpha = \beta$ in (3.2) and replacing the population mean by the sample mean $\bar{x}$ and β by b, moment estimate $\tilde{\theta}$ of θ is obtained as

$$\tilde{\theta} = \frac{2(b+3)}{(b+2)(\bar{x}-1)} \tag{4.12}$$

and so the moment estimate $\tilde{\alpha}$ of α is obtained as

$$\tilde{\alpha} = \frac{b}{\tilde{\theta}} = \frac{b(b+2)(\bar{x}-1)}{2(b+3)} \tag{4.13}$$

5. Goodness of Fit

The SBTPPLD has been fitted to a number of data sets related to a number of observations of the size distribution of 'freely –forming' small groups in various public places reported by James (1953), Coleman and James (1961) and Simonoff (2003), and it was found that to almost all these data sets, the SBTPPLD provides closer fit than SBPD and SBPLD. Here, the goodness of fit of the SBTPPLD to three such data sets has been presented along with the goodness of fit given by SBPD and SBPLD.

Table 1. Counts of groups of people in public places on a spring afternoon in Portland

Size of Groups	Observed Frequency	Expected Frequency		
		SBPD	SBPLD	SBTPPLD
1	1486	1452.4	1531.9	1485.5
2	694	743.3	630.8	697.2
3	195	190.2	192.1	189.7
4	37	32.4	51.4	41.1
5	10	4.1	12.7	7.9
6	1	0.6	4.1	1.6
Total	2423	2423.0	2423.0	2423.0
ML Estimates		$\hat{\theta} = 0.5118$	$\hat{\theta} = 4.5043$	$\hat{\theta} = 7.1400$ $\hat{\alpha} = -0.1107$
χ^2		7.369	13.786	0.807
d.f.		2	3	2

Table 2. Counts of Shopping Groups-Eugene, Spring, Department Store and Public Market

Size of Groups	Observed Frequency	Expected Frequency		
		SBPD	SBPLD	SBTPPLD
1	316	306.3	322.9	315.7
2	141	156.2	132.6	142.7
3	44	39.8	40.2	40.1
4	5	6.8	10.7	9.1
5	4	0.9	3.6	2.4
Total	510	510.0	510.0	510.0
ML Estimates		$\hat{\theta} = 0.5098$	$\hat{\theta} = 4.5206$	$\hat{\theta} = 6.5526$ $\hat{\alpha} = -0.0775$
χ^2		2.448	3.002	0.942
d.f.		2	2	1

Table 3. Counts of Play Groups-Eugene, Spring, Public Playground

Size of Groups	Observed Frequency	Expected Frequency		
		SBPD	SBPLD	SBTPPLD
1	305	296.5	314.2	304.3
2	144	159.0	134.4	148.3
3	50	42.7	42.5	42.3
4	5	7.6	11.8	9.7
5	2	1.0	3.0	1.9
6	1	0.2	1.1	0.5
Total	507	507.0	507.0	507.0
ML Estimates		$\hat{\theta} = 0.5365$	$\hat{\theta} = 4.3138$	$\hat{\theta} = 6.7078$ $\hat{\alpha} = -0.1116$
χ^2		2.983	6.214	2.914
d.f.		2	2	1

The expected frequencies according to the SBPD and SBPLD have also been given in these tables for ready comparison with those obtained by the SBTPPLD. The estimates of the parameters have been obtained by the method of maximum likelihood estimation. On the basis of the values of chi-square, it can be seen that the SBTPPLD gives much closer fit than those by the SBPD and SBPLD.

The values of coefficient of variation (C.V), coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis $\left(\beta_2\right)$ and index of dispersion $\left(\gamma\right)$ for estimated values of parameters for SBPD, SBPLD, and SBTPPLD and for original data for tables 1, 2 and 3 are presented in the following table 4.

Table 4. Values of C.V, $\sqrt{\beta_1}$, β_2 and γ of SBPD, SBPLD, and SBTPPLD for estimated values of parameters and for original data

Data set		Original Data	SBPD	SBPLD	SBTPPLD
Table 1	C.V	0.495104	0.473212	0.527558	0.495984
	$\sqrt{\beta_1}$	1.258915	1.397816	1.857652	1.615977
	β_2	2.797695	4.953888	7.375122	6.099772
	γ	0.370575	0.338537	0.420895	0.371874
Table 2	C.V	0.502414	0.472912	0.527031	0.503102
	$\sqrt{\beta_1}$	1.287003	1.400555	1.859776	1.672407
	β_2	3.156249	4.961554	7.383374	6.362748
	γ	0.381105	0.337661	0.419495	0.382140
Table 3	C.V	0.500704	0.476708	0.533799	0.502112
	$\sqrt{\beta_1}$	1.269157	1.365254	1.832651	1.598441
	β_2	3.408812	4.863933	7.278512	6.059832
	γ	0.385205	0.349170	0.437971	0.387356

It is also obvious from the analysis of the table 1, 2, 3, and 4 that the distribution which gives better fit in terms of chi-square values are the distribution whose index of dispersion for given values of the parameters is equal or nearer to the value of the index of dispersion of the original data. For example, it is clear that the index of dispersion of SBTPPLD is almost equal to the index of dispersion of the original data. Therefore, both the index of dispersion and the values of chi-square of SBTPPLD certify that the SBTPPLD is the best model than both SBPD and SBPLD for modeling data which structurally excludes zero counts.

6. Conclusions

In this paper, a size-biased two parameter Poisson-Lindley distribution (SBTPPLD), of which the size-biased Poisson-Lindley distribution (SBPLD) is a particular case, has been introduced to model count data which structurally excludes zero counts. The first four moments about origin, moments about mean, and expressions for coefficient of variation, skewness, kurtosis and index of dispersion have been obtained. The estimation of its parameters has been discussed using the method of maximum likelihood and the method of moments. The goodness of fit of the distribution has been presented to three data sets and it has been found that to all these data sets it provides much closer fit than both SBPD and SBPLD. The SBTPPLD has been found more general in nature and wider in scope than SBPD and SBPLD. Since SBTPPLD provides much closer fit to the observed data sets than those provided by the SBPD and SBPLD, SBTPPLD should be preferred over SBPD and SBPLD for modeling count data sets which structurally excludes zero counts.

ACKNOWLEDGEMENTS

The authors are grateful to the Editor-In-Chief of the journal and the anonymous reviewer for constructive and helpful comments which lead to the improvement in the quality of the paper.

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*, 24, 36 – 45.

[3] Das, K.K. and Roy, T.D. (2011): On some length biased weighted Weibull distribution, *Advances in Applied Science Research,* 2(5), 465 – 475.

[4] Ducey, M.J. and Gove, J.H. (2015): Size-biased distributions in the generalized Beta distribution family with Applications to forestry, Forestry-An International Journal of Forest Research, 88. 143 – 151.

[5] Fisher, R.A. (1934): The effects of methods of ascertainment upon the estimation of frequency, *Ann. Eugenics,* 6, 13-25.

[6] Ghitany, M.E., Al-Mutairi, D.K. (2008): Size- biased Poisson-Lindley distribution and its applications, *Metron*, LXVI, (3), 299 – 311.

[7] Gove, H. J. (2003): Estimation and application of size-biased distributions in forestry, *In Modeling Forest Systems*, A. Amaro, D. Reed and P.Soares, CAB International, Wallingford, U.K., 201 – 212.

[8] James, J. (1953): The Distribution of Free-Forming Small Group Size, *American Sociological Review*, 18, 569 – 570.

[9] Lappi, J. and Bailey, R. L. (1987): Estimation of diameter increment function or other relations using angle-count samples, *Forest Science*, 33, 725 – 739.

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

[11] Mir, K.A. and Ahmad, M. (2009): Size-biased distributions and Their Applications, *Pakistan Journal of Statistics*, 25 (3), 283 – 294.

[12] Patil, G. P. and Rao, C. R. (1977): The weighted distributions: A survey and their applications, *In Applications of Statistics, (Ed. P.R. Krishnaiah*), 383 – 405, North Holland Publications co., Amsterdam.

[13] Patil, G. P. and Rao, C. R. (1978): Weighted distributions and size-biased sampling with applications to wild life populations and human families, *Biometrics*, 34, 179 - 189

[14] Rao, C. R. (1965): On discrete distributions arising out of methods on ascertainment, *Classical and Contagious Discrete Distribution,* Patil, G.P. (Ed), Statistical Publishing Society, Calcutta, 320-332.

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

[16] Shanker, R. and Mishra, A. (2013): A two parameter Lindley distribution, *Statistics in Transition-New Series,* 14(1), 45 – 56.

[17] Shanker, R. and Mishra, A. (2014): A two parameter Poisson-Lindley distribution, *International Journal of Statistics and Systems*, 9 (1), 79 – 85.s

[18] Simoj, S.M. and Maya, S.S. (2006): Some properties of weighted distributions in the context of reparable systems, *Communications in Statistics-Theory and Methods*, 35, 223 – 228.

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

[20] Van Deusen, P.C. (1986): Fitting assumed distributions to horizontal point sample diameters, *Forest Science*, 32, 146 – 148.