International Journal of Probability and Statistics 2017, 6(3): 33-44
DOI: 10.5923/j.ijps.20170603.01

A Size-Biased Poisson-Shanker Distribution and Its Applications

Rama Shanker

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

Abstract In this paper, a size-biased Poisson-Shanker distribution (SBPSD) has been obtained by size-biasing the Poisson-Shanker distribution (PSD) introduced by Shanker (2016). Its raw moments and central moments have been obtained and hence expressions for coefficient of variation (C.V.), skewness, kurtosis and index of dispersion have also been given. The method of maximum likelihood and the method of moments have been discussed for estimating its parameter. The goodness of fit of SBPSD has been discussed with two real data sets and the fit shows quite satisfactory fit over size-biased Poisson distribution (SBPD) and size-biased Poisson-Lindley distribution (SBPLD).

Keywords Size-biasing, Poisson-Shanker distribution, Moments, Estimation of parameter, Goodness of fit

1. Introduction

Shanker (2016) introduced Poisson-Shanker distribution (PSD) having probability mass function (p.m.f.)

$$P_0(x;\theta)=\frac{\theta^2}{\theta^2+1}\frac{x+\theta^2+\theta+1}{(\theta+1)^{x+2}}\ ;x=0,1,2,...,\ \theta>0 \tag{1.1}$$

for modeling count data from different fields of knowledge. The PSD arises from the Poisson distribution when its parameter λ follows a lifetime distribution named Shanker distribution introduced by Shanker (2015) having probability density function (p.d.f.)

$$f_0(x;\theta)=\frac{\theta^2}{\theta^2+1}(\theta+x)e^{-\theta x}\ ;x>0,\ \theta>0 \tag{1.2}$$

Size-biased distributions are special class of weighted distributions and arise in practice when observations from a sample are recorded with probability proportional to some measure of unit size. Fisher (1934) firstly introduced these distributions to model ascertainment biases which were later formalized by Rao (1965) in a unifying theory. Size-biased observations occur in many research areas and its fields of applications includes econometrics, environmental science, medical science, sociology, psychology, ecology, geological sciences etc. The applications of size-biased distribution theory in fitting distributions of diameter at breast height (DBH) data arising from horizontal point sampling (HPS) has been discussed by Van Deusen (1986). Further, Lappi and Bailey (1987) have studied and applied size-biased distributions to analyze HPS diameter increment data. The detailed statistical applications of size-biased distributions to the analysis of data relating to human population and ecology can be found in Patil and Rao (1977, 1978). A number of research have been done relating to size-biased distributions and their applications in different areas of knowledge by many researchers including Scheaffer (1972), Patil and ord (1976), Singh and Maddala (1976), Patil, (1981), McDonald (1984), Gove (2000, 2003), Correa and Wolfson (2007), Drummer and McDonald (1987), Ducey (2009), Alavi and Chinipardaz (2009), Ducey and Gove (2015), are some among others.

Let a random variable X has original probability distribution $P_0(x;\theta);x=0,1,2,...,\theta>0$. Suppose the sample units are weighted or selected with probability proportional to some measure x^α, where α is a positive integer. Then the corresponding size-biased probability distribution of order α can be defined by its probability mass function

$$P_1(x;\theta)=\frac{x^\alpha\cdot P_0(x;\theta)}{\mu'_\alpha}$$

where $\mu'_\alpha=E(X^\alpha)=\sum_{x=0}^{\infty}x^\alpha P_0(x;\theta)$. If $\alpha=1$, then the distribution is known as simple size-biased and is applicable for size-biased sampling in sampling theory. If $\alpha=2$, then the distribution is known as area-biased distribution and is applicable for area-biased sampling in forestry.

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

The p.m.f. of the size-biased Poisson-Shanker distribution (SBPSD) with parameter θ can thus be be obtained as

$$P_2(x;\theta)=\frac{x\cdot P_0(x;\theta)}{\mu_1'}=\frac{\theta^3}{\theta^2+2}\frac{x\left(x+\theta^2+\theta+1\right)}{(\theta+1)^{x+2}};x=1,2,3,..,\theta>0 \tag{1.3}$$

where $\mu_1'=\dfrac{\theta^2+2}{\theta\left(\theta^2+1\right)}$ is the mean of the Poisson-Shanker distribution (PSD) with p.m.f. (1.1).

Recall that the probability mass function of size-biased Poisson-Lindley distribution (SBPLD) with parameter θ, obtained by Ghitany and Mutairi (2008), is given by

$$P_3(x;\theta)=\frac{x\cdot P_0(x;\theta)}{\mu_1'}=\frac{\theta^3}{\theta+2}\frac{x(x+\theta+2)}{(\theta+1)^{x+2}};x=1,2,3,..,\theta>0 \tag{1.4}$$

The SBPLD is a size-biased version of Poisson-Lindley distribution introduced by Sankaran (1970) and the Poisson-Lindley distribution is a Poisson mixture of Lindley distribution, introduced by Lindley (1958). The statistical and mathematical properties, estimation of parameter using both maximum likelihood estimation and method of moments, and goodness of fit of SBPLD have been discussed by Ghitany and Mutairi (2008). Shanker *et al* (2015) has discussed the applications of SBPLD for modeling data on thunderstorms and found that SBPLD is a better model for thunderstorms than size-biased Poisson distribution (SBPD).

The comparative nature and behavior of SBPSD and SBPLD are shown in the figure 1.

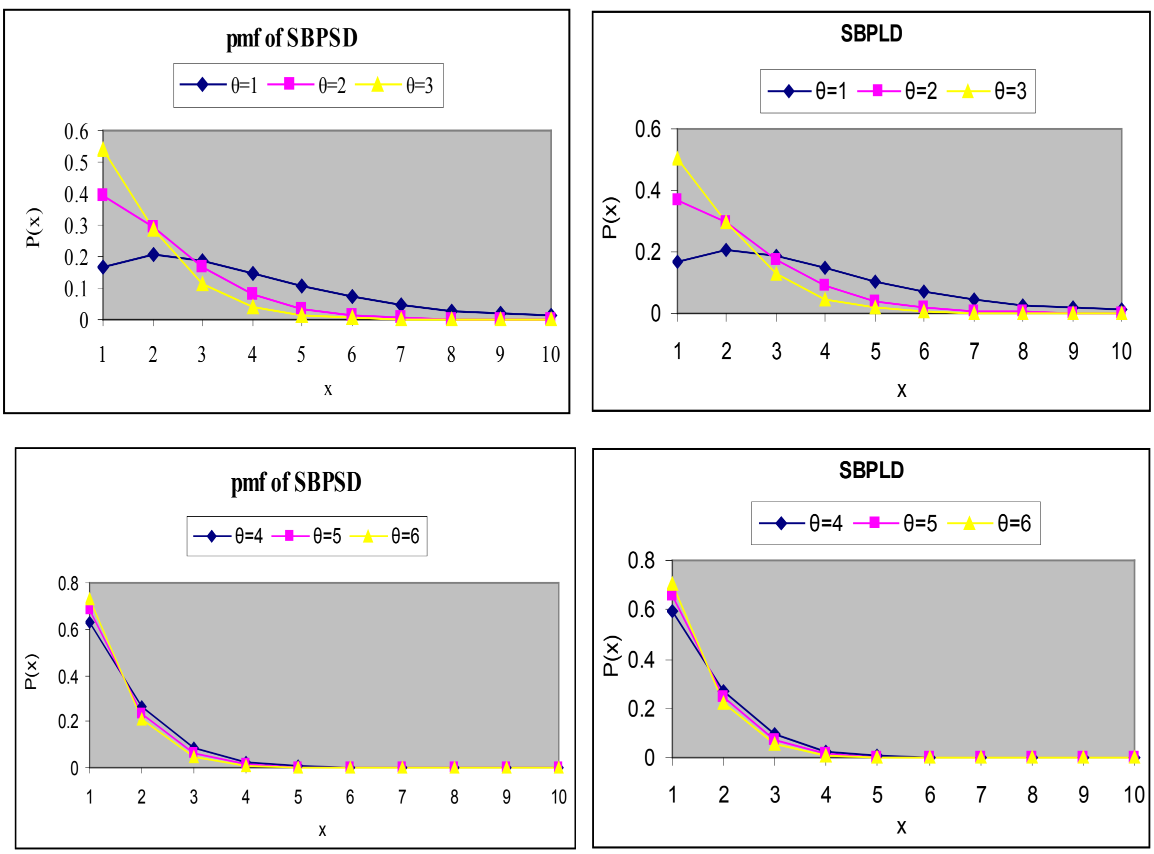

Figure 1. Graphs of SBPSD and SBPLD for varying values of parameter θ

The independent graphs of SBPSD for varying values of the parameter θ have been shown in figure 2.

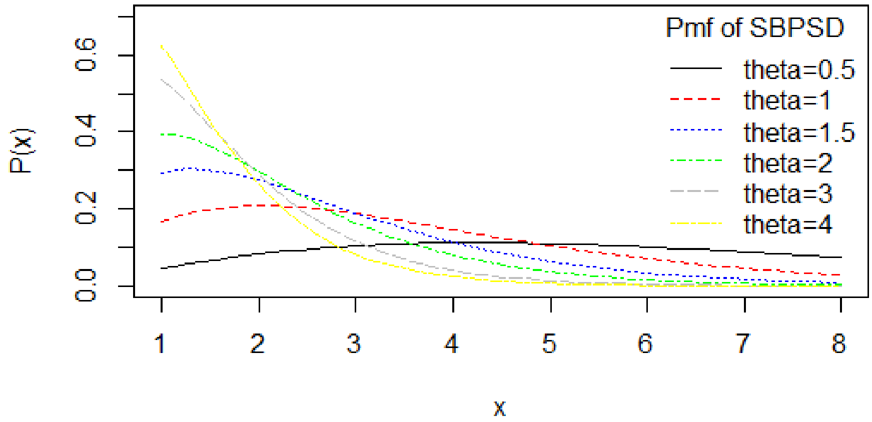

Figure 2. Graphs of SPBSD for varying values of parameter θ

The probability mass function of SBPSD can also be obtained from the size-biased Poisson distribution (SPBD) with p.m.f.

$$g(x\mid\lambda)=\frac{e^{-\lambda}\lambda^{x-1}}{\Gamma(x)}\quad ;x=1,2,3,...,\lambda>0 \tag{1.5}$$

when its parameter λ follows size-biased Shanker distribution (SBSD) with p.d.f.

$$h(\lambda;\theta)=\frac{\theta^3}{\theta^2+2}\lambda(\theta+\lambda)e^{-\theta\lambda}\quad ;x>0,\ \theta>0 \tag{1.6}$$

Thus, the pmf of SBPSD can be obtained as

$$P(X=x)=\int_0^\infty g(x\mid\lambda)\cdot h(\lambda;\theta)\,d\lambda$$

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

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

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

$$=\frac{\theta^3}{\theta^2+2}\frac{x\left(x+\theta^2+\theta+1\right)}{(\theta+1)^{x+2}};x=1,2,3,..,\theta>0 \tag{1.7}$$

which is the p.m.f of SBPSD as given in (1.3).

2. Moments and Related Measures of SBPSD

Using (1.7), th rth factorial moment about origin of the SBPSD (1.3) can be obtained as

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

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

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

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

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

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

$$= \frac{r!\left\{r\theta^3+(r+1)\theta(\theta+r)+(r+1)(r+2)\right\}}{\theta^r\left(\theta^2+2\right)} \;;r=1,2,3,... \tag{2.1}$$

Substituting $r=1,2,3$ and 4 in (2.1), first four factorial moments about origin can be obtained and then using the relationship between factorial moments about origin and moments about origin, the first four moments about origin of SBPSD (1.3) can be obtained as

$$\mu_1' = \frac{\theta^3+2\theta^2+2\theta+6}{\theta\left(\theta^2+2\right)}$$

$$\mu_2' = \frac{\theta^4+6\theta^3+8\theta^2+18\theta+24}{\theta^2\left(\theta^2+2\right)}$$

$$\mu_3' = \frac{\theta^5+14\theta^4+38\theta^3+66\theta^2+144\theta+120}{\theta^3\left(\theta^2+2\right)}$$

$$\mu_4' = \frac{\theta^6+30\theta^5+152\theta^4+330\theta^3+720\theta^2+1200\theta+720}{\theta^4\left(\theta^2+2\right)}$$

Again using the relationship between moments about mean and the moments about origin, the moments about mean of the SBPSD (1.3) are obtained as

$$\mu_2 = \sigma^2 = \frac{2\left(\theta^5+\theta^4+5\theta^3+6\theta^2+6\theta+6\right)}{\theta^2\left(\theta^2+2\right)^2}$$

$$\mu_3 = \frac{2\left(\theta^8+3\theta^7+9\theta^6+24\theta^5+34\theta^4+54\theta^3+48\theta^2+36\theta+24\right)}{\theta^3\left(\theta^2+2\right)^3}$$

$$\mu_4 = \frac{2\begin{pmatrix}\theta^{11}+13\theta^{10}+33\theta^9+142\theta^8+294\theta^7+604\theta^6+980\theta^5+1200\theta^4+1392\theta^3 \\ +1104\theta^2+720\theta+360\end{pmatrix}}{\theta^4\left(\theta^2+2\right)^4}$$

The coefficient of variation $(C.V)$, coefficient of Skewness $\left(\sqrt{\beta_1}\right)$, coefficient of Kurtosis (β_2) and index of dispersion (γ) of the SBPSD (1.3) are thus obtained as

$$C.V = \frac{\sigma}{\mu_1'} = \frac{\sqrt{2\left(\theta^5+\theta^4+5\theta^3+6\theta^2+6\theta+6\right)}}{\theta^3+2\theta^2+2\theta+6}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\theta^8+3\theta^7+9\theta^6+24\theta^5+34\theta^4+54\theta^3+48\theta^2+36\theta+24}{\sqrt{2}\left(\theta^5+\theta^4+5\theta^3+6\theta^2+6\theta+6\right)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\left(\begin{array}{l}\theta^{11}+13\theta^{10}+33\theta^9+142\theta^8+294\theta^7+604\theta^6+980\theta^5+1200\theta^4+1392\theta^3\\+1104\theta^2+720\theta+360\end{array}\right)}{2\left(\theta^5+\theta^4+5\theta^3+6\theta^2+6\theta+6\right)^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{2\left(\theta^5+\theta^4+5\theta^3+6\theta^2+6\theta+6\right)}{\theta\left(\theta^2+2\right)\left(\theta^3+2\theta^2+2\theta+6\right)}$$

It can be easily verified that SBPSD is over-dispersed $\left(\mu<\sigma^2\right)$, equi-dispersed $\left(\mu=\sigma^2\right)$ and under-dispersed $\left(\mu>\sigma^2\right)$ for $\theta<(=)>\theta^*=1.634877$. The graphs of coefficient of variation, coefficient of skewness, coefficient of kurtosis, and index of dispersion of SBPSD for varying values of parameter θ are shown in figure 3.

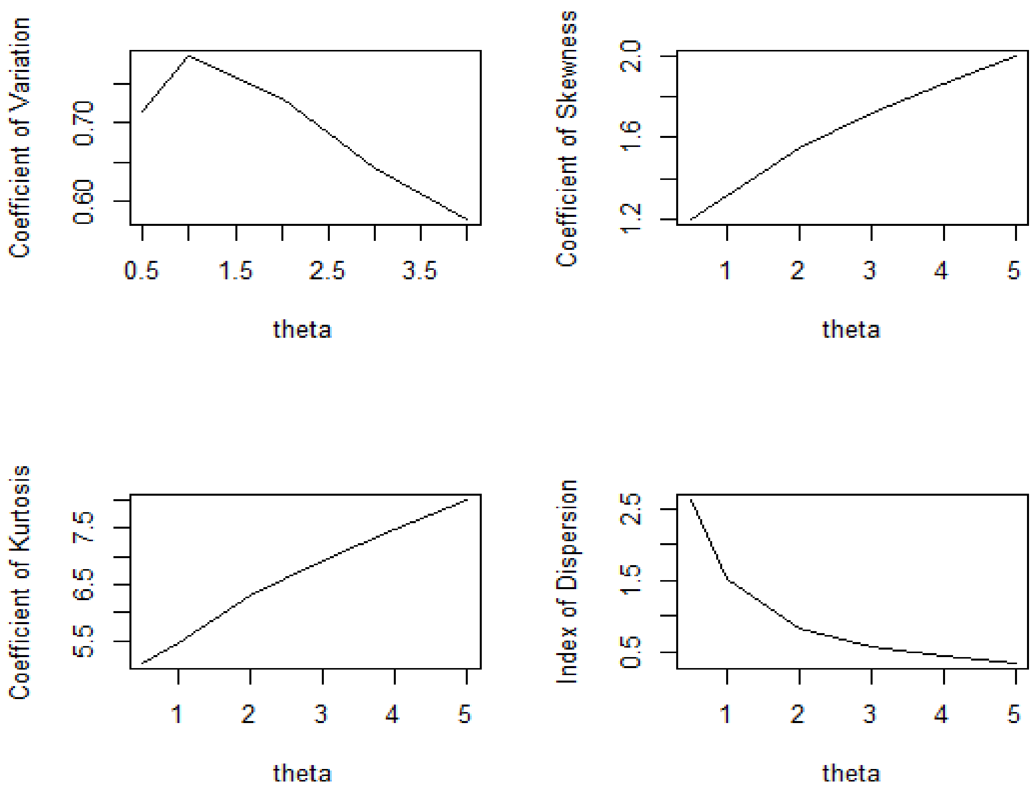

Figure 3. Graphs of C.V, $\sqrt{\beta_1}$, β_2 and γ of the SBPSD for varying values of parameter θ

To study the comparative nature and behavior of SBPSD and SBPLD, a table for μ_1', μ_2, C.V, $\sqrt{\beta_1}$, β_2, and γ has been prepared for varying values of the parameter θ and presented in table 1.

Table 1. Values of $\mu_1', \mu_2, \text{C.V}, \sqrt{\beta_1}, \beta_2$ and γ of SBPSD and SBPLD for varying values of the parameter θ

	Values of θ for SBPSD					
	1	**2**	**3**	**4**	**5**	**6**
μ_1'	3.666667	2.166667	1.727273	1.527778	1.414815	1.342105
μ_2	5.555556	1.805556	0.986226	0.665895	0.500521	0.400508
CV	0.642824	0.620174	0.574946	0.534125	0.500048	0.47154
$\sqrt{\beta_1}$	1.318047	1.54948	1.721223	1.86595	1.997232	2.119503
β_2	5.4744	6.296095	6.926298	7.467793	7.982788	8.489255
γ	1.515152	0.833333	0.570973	0.435859	0.353772	0.298418

	Values of θ for SBPLD					
	1	**2**	**3**	**4**	**5**	**6**
μ_1'	3.666667	2.25	1.8	1.583333	1.457143	1.375
μ_2	5.555556	1.9375	1.093333	0.743056	0.556735	0.442708
CV	0.642824	0.61864	0.580903	0.544425	0.512061	0.483901
$\sqrt{\beta_1}$	1.318047	1.49478	1.649924	1.790721	1.921224	2.043701
β_2	5.4744	6.057232	6.599941	7.118613	7.625214	8.125813
γ	1.515152	0.861111	0.607407	0.469298	0.382073	0.32197

The comparative graphs of coefficient of variation, coefficient of skewness, coefficient of kurtosis, and index of dispersion of SBPSD and SBPLD are shown in figure 4.

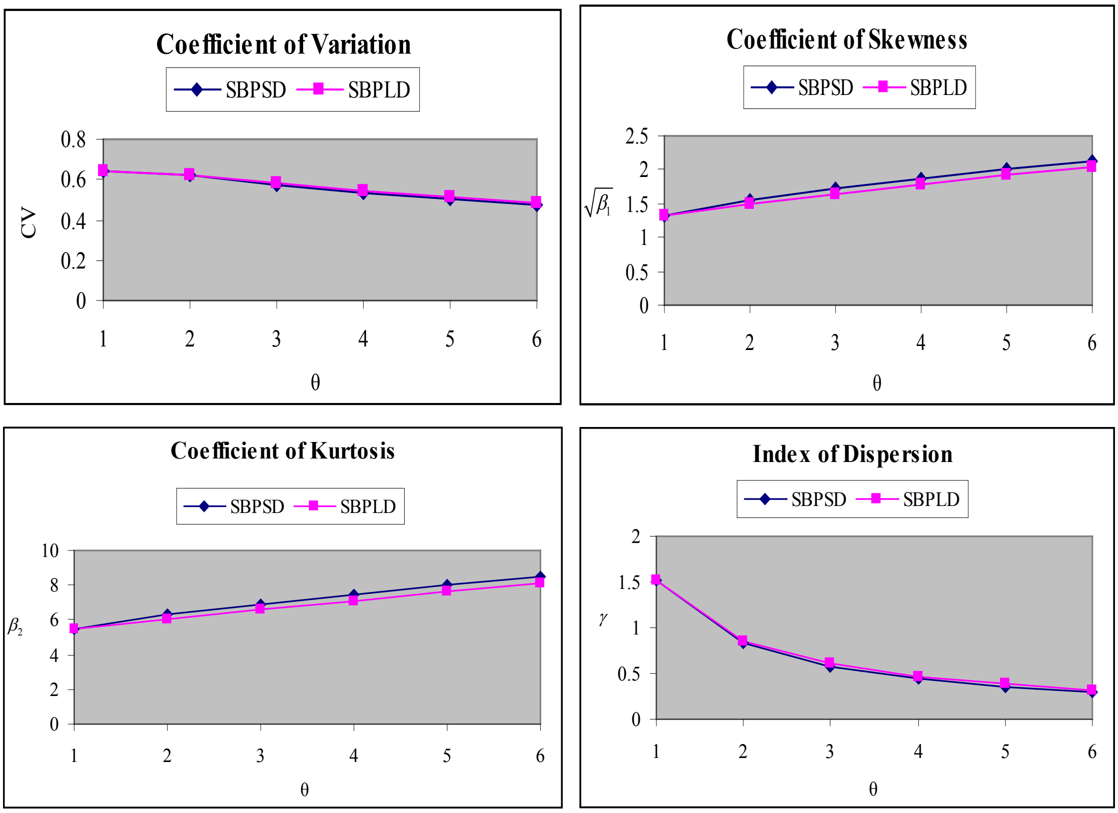

Figure 4. Graphs of C.V, $\sqrt{\beta_1}$, β_2 and γ of the SBPSD and SBPLD for varying values of parameter θ

3. Statistical Properties of SBPSD

3.1. Reliability Properties

Since

$$\frac{P_2(x+1;\theta)}{P_2(x;\theta)} = \left(\frac{1}{\theta+1}\right)\left(1+\frac{1}{x}\right)\left(1+\frac{1}{x+\theta^2+\theta+1}\right)$$

is a deceasing function of x, $P_2(x;\theta)$ is log-concave. Therefore, SBPSD is unimodal, has an increasing failure rate (IFR), and hence increasing failure rate average (IFRA). It is new better than used in expectation (NBUE) and has decreasing mean residual life (DMRL). Detailed discussion and definitions of these aging concepts can be seen in Barlow and Proschan (1981).

3.2. Generating Function

Probability Generating Function: The probability generating function of the SBPSD (1.3) can be obtained as

$$P_X(t) = E\left(t^X\right) = \frac{\theta^3}{\left(\theta^2+2\right)(\theta+1)^2}\left[\sum_{x=1}^{\infty} x^2\left(\frac{t}{\theta+1}\right)^x + \left(\theta^2+\theta+1\right)\sum_{x=1}^{\infty}\left(\frac{t}{\theta+1}\right)^x\right]$$

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

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

Moment Generating Function: The moment generating function of the SBPSD (1.3) is given by

$$M_X(t)=E\left(e^{tX}\right)=\frac{\theta^3 e^t}{\left(\theta^2+2\right)\left(\theta+1\right)}\left[\frac{\theta+1+e^t}{\left(\theta+1-e^t\right)^3}+\frac{\theta^2+\theta+1}{\left(\theta+1-e^t\right)^2}\right]$$

4. Estimation of Parameter

4.1. Maximum Likelihood Estimate (MLE): Let $x_1,x_2,...,x_n$ be a random sample of size n from the SBPSD (1.3) and let f_x be the observed frequency in the sample corresponding to $X=x\ (x=1,2,3,...,k)$ 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 SBPSD (1.3) is given by

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

The log likelihood function is obtained as

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

The first derivative of the log likelihood function is given by

$$\frac{d\log L}{d\theta}=\frac{n\left(\theta^2+6\right)}{\theta\left(\theta^2+2\right)}-\frac{n\left(\bar{x}+2\right)}{\theta+1}+\sum_{x=1}^{k}\frac{\left(2\theta+1\right)x f_x}{x^2+x\left(\theta^2+\theta+1\right)}$$

where $\bar{x}$ is the sample mean.

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ is the solution of the equation $\frac{d\log L}{d\theta}=0$ and is given by the solution of the non-linear equation

$$\frac{n\left(\theta^2+6\right)}{\theta\left(\theta^2+2\right)}-\frac{n\left(\bar{x}+2\right)}{\theta+1}+\sum_{x=1}^{k}\frac{\left(2\theta+1\right)x f_x}{x^2+x\left(\theta^2+\theta+1\right)}=0 \tag{4.1.1}$$

This non-linear equation can be solved by any numerical iteration methods such as Newton- Raphson method, Bisection method, Regula –Falsi method etc. The consistency and asymptotic normality of maximum likelihood estimator of SBPSD has been established in the following theorem.

Theorem: The ML estimator $\hat{\theta}$ of θ of the SBPSD is consistent and asymptotically normal. That is, $\sqrt{n}\left(\hat{\theta}-\theta\right)\xrightarrow{d}N\left[0,\ I^{-1}\left(\theta\right)\right]$, where

$$I\left(\theta\right)=\frac{4\theta^7-2\theta^6+7\theta^5-3\theta^4+6\theta^3+16\theta^2+12}{\theta^2\left(\theta+1\right)\left(\theta^2+2\right)^2}-\frac{\theta^3\left(2\theta+1\right)^2\left(\theta^2+\theta+1\right)}{\left(\theta+1\right)^2\left(\theta^2+2\right)}\int_0^1\frac{t^{\theta^2+\theta+1}}{\theta+1-t}dt$$ is the Fisher's information about θ.

Proof: The SBPSD satisfies the regularity conditions under which the ML estimator $\hat{\theta}$ of θ is consistent and asymptotically normal [see Hogg *et al* (2005), chapter 6]. Finally, we have

$$I(\theta) = E\left[-\frac{\partial^2}{\partial\theta^2}\ln f(X;\theta)\right]$$

$$= E\left[\frac{3}{\theta^2} - \frac{2(\theta^2-2)}{(\theta^2+2)^2} - \frac{X+2}{(\theta+1)^2} - \frac{2}{X+(\theta^2+\theta+1)} + \frac{(2\theta+1)^2}{\left\{X+(\theta^2+\theta+1)\right\}^2}\right]$$

$$= E\left[\frac{\theta^4+16\theta^2+12}{\theta^2(\theta^2+2)^2} - \frac{\mu+2}{(\theta+1)^2} - 2E\left[\frac{1}{X+(\theta^2+\theta+1)}\right] + (2\theta+1)^2 E\left[\frac{1}{\left\{X+(\theta^2+\theta+1)\right\}^2}\right]\right] \tag{4.1.2}$$

where

$$\mu = \frac{\theta^3+2\theta^2+2\theta+6}{\theta(\theta^2+2)} \tag{4.1.3}$$

Now

$$E\left[\frac{1}{X+(\theta^2+\theta+1)}\right] = \sum_{x=1}^{\infty}\frac{1}{x+(\theta^2+\theta+1)}\cdot\frac{\theta^3}{\theta^2+2}\cdot\frac{x\left[x+(\theta^2+\theta+1)\right]}{(\theta+1)^{x+2}}$$

$$= \frac{\theta^3}{(\theta+1)^2(\theta^2+2)}\sum_{x=1}^{\infty}x\left(\frac{1}{\theta+1}\right)^x = \frac{\theta^2}{(\theta+1)(\theta^2+2)} \tag{4.1.4}$$

and

$$E\left[\frac{1}{\left\{X+(\theta^2+\theta+1)\right\}^2}\right] = \sum_{x=1}^{\infty}\frac{1}{\left\{x+(\theta^2+\theta+1)\right\}^2}\cdot\frac{\theta^3}{\theta^2+2}\cdot\frac{x\left[x+(\theta^2+\theta+1)\right]}{(\theta+1)^{x+2}}$$

$$= \frac{\theta^3}{(\theta+1)^2(\theta^2+2)}\sum_{x=1}^{\infty}\frac{x}{\left\{x+(\theta^2+\theta+1)\right\}(\theta+1)^x}$$

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

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

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

$$=\frac{\theta^2}{(\theta+1)^2(\theta^2+2)}-\frac{\theta^3(\theta^2+\theta+1)}{(\theta+1)^2(\theta^2+2)}\sum_{x=1}^{\infty}\frac{1}{\{x+(\theta^2+\theta+1)\}(\theta+1)^x}$$

$$=\frac{\theta^2}{(\theta+1)^2(\theta^2+2)}-\frac{\theta^3(\theta^2+\theta+1)}{(\theta+1)^2(\theta^2+2)}\int_0^1\frac{t^{\theta^2+\theta+1}}{\theta+1-t}dt \tag{4.1.5}$$

Using equations (4.1.3), (4.1.4), and (4.1.5) in (4.1.2), we get

$$I(\theta)=\frac{4\theta^7-2\theta^6+7\theta^5-3\theta^4+6\theta^3+16\theta^2+12}{\theta^2(\theta+1)(\theta^2+2)^2}-\frac{\theta^3(2\theta+1)^2(\theta^2+\theta+1)}{(\theta+1)^2(\theta^2+2)}\int_0^1\frac{t^{\theta^2+\theta+1}}{\theta+1-t}dt\,.$$

4.2. Method of Moment Estimate (MOME): Let $x_1,x_2,...,x_n$ be a random sample of size n from the SBPSD (1.3). Equating population mean to the corresponding sample mean, the MOME $\tilde{\theta}$ of θ of SBPSD is the solution of the following cubic equation in θ

$$(\overline{x}-1)\theta^3-2\theta^2+2(\overline{x}-1)\theta-6=0 \tag{4.2.1}$$

where $\overline{x}$ is the sample mean.

5. Goodness of Fit of SBPSD

The SBPSD has been fitted to a number of data - sets to test its goodness of fit over SBPD and SBPLD. The maximum likelihood estimate (MLE) has been used to fit the SBPSD. Two examples of observed data-sets, for which the SBPD, SBPLD and SBPSD has been fitted, are presented. The first data-set is immunogold assay data of Cullen *et al.* (1990) regarding the distribution of number of counts of sites with particles from immunogold assay data and the second data-set is animal abundance data of Keith and Meslow (1968) regarding the distribution of snowshoe hares captured over 7 days.

Table 2. Distribution of number of counts of sites with particles from Immunogold data

No. of sites with particles	Observed Frequency	Expected Frequency		
		SBPD	SBPLD	SBPSD
1	122	111.3	119.0	119.3
2	50	64.1	53.8	53.4
3	18	18.5	18.0	17.9
4	4	3.5	5.3	5.3
5	4	0.6	1.9	2.1
Total	198	198.0	198.0	198.0
ML estimate		$\hat{\theta}=0.576$	$\hat{\theta}=4.051$	$\hat{\theta}=3.696733$
χ^2		4.642	0.433	0.325
d.f.		1	2	2
p-value		0.031	0.805	0.8500

Table 3. Distribution of snowshoe hares captured over 7 days

No. times hares caught	Observed Frequency	Expected Frequency		
		SBPD	SBPLD	SBPSD
1	184	170.6	177.3	177.5
2	55	72.5	62.5	62.3
3	14	15.4	16.4	16.4
4	4	2.2	3.8	3.8
5	4	0.3	1.0	1.0
Total	261	261.0	261.0	261.0
ML estimate		$\hat{\theta} = 0.425$	$\hat{\theta} = 5.351$	$\hat{\theta} = 4.886676$
χ^2		6.216	1.183	1.123
d.f.		1	1	1
p-value		0.013	0.277	0.2892

6. Concluding Remarks

In this paper, a size-biased Poisson-Shanker distribution (SBPSD) has been proposed by size-biasing the Poisson-Shanker distribution (PSD) introduced by Shanker (2016). Its raw moments and central moments have been obtained and hence expressions for coefficient of variation (C.V.), skewness, kurtosis and index of dispersion have also been given. Its reliability properties have been explained and generating functions have been derived. The method of maximum likelihood and the method of moments have been discussed for estimating the parameter. Two real lifetime data sets have been presented to test the goodness of fit of SBPSD and the fit is quite satisfactory over size-biased Poisson distribution (SBPD) and size-biased Poisson-Lindley distribution (SBPLD).

ACKNOWLEDGEMENTS

The author expresses his thankfulness to the editor in-chief and the anonymous reviewer for their constructive comments which improved the presentation of the paper.

REFERENCES

[1] Alavi, S.M.R. and Chinipardaz, R. (2009): Form-invariance under weighted sampling, *Statistics*, 43, 81 – 90.

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

[3] Correa, J.A. and Wolfson, D.B. (2007): Length-bias: some Characterizations and applications, *Journal of Statistical Computation and Simulation*, 64, 209 – 219.

[4] Cullen, M.J., Walsh, J., Nicholson, L.V., and Harris, J.B. (1990): Ultrastructural localization of dystrophin in human muscle by using gold immunolabelling, *Proceedings of the Royal Society of London*, 20, 197-210.

[5] Ducey, M.J. (2009): Sampling trees with probability nearly proportional to biomass, *For. Ecol. Manage*., 258, 2110 – 2116.

[6] 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.

[7] Drummer, T.D. and MacDonald, L.L. (1987): Size biased in line transect sampling, *Biometrics*, 43, 13 – 21.

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

[9] Ghitany, M.E. and Al-Mutairi, D.K. (2008): Size-biased Poisson-Lindley distribution and Its Applications, *Metron - International Journal of Statistics*, LXVI (3), 299 – 311.

[10] Gove, J.H. (2000): Some observations on fitting assumed diameter distributions to horizontal point sampling data, *Can. J. For. Res.,* 30, 521 – 533.

[11] Gove, J.H. (2003): Estimation and applications of size-biased distributions in forestry. In Modeling Forest Systems. A Amaro, D. Reed and P. Soares (Eds), CABI Publishing, pp. 201 – 212.

[12] Hogg, R.V., Mckean, J.W. and Craig, A.T. (2005): Introduction to Mathematical Statistics, 6th edition, Pearson Prentice hall, New Jersey.

[13] Keith, L.B. and Meslow, E.C. (1968): Trap Response by snowshoe hares, *Journal of Wildlife Management*, 20, 795-801.

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

[15] Lindley, D.V. (1958): Fiducial distributions and Bayes theorem, *Journal of the Royal Statistical Society*, 20 (1), 102-107.

[16] McDonald, J.B. (1984): Some generalized functions for the size distribution of income, *Econometrics*, 52, 647 – 664.

[17] Patil, G.P. (1981): Studies in statistical ecology involving weighted distributions. In Applications and New Directions, J.K. Ghosh and J. Roy (eds). Proceeding of Indian Statistical Institute. Golden Jubliee, Statistical Publishing society, pp. 478 – 503.

[18] Patil, G.P. and Ord, J.K. (1976): On size-biased sampling and related form-invariant weighted distributions, *Sankhya Ser. B*, 38, 48 – 61.

[19] 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.

[20] 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

[21] Rao, C.R. (1965): On discrete distributions arising out of methods of ascertainment In: Patil, G.P. (eds) Classical and Contagious Discrete Distributions. Statistical Publishing Society, Calcutta, 320 – 332.

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

[23] Scheaffer, R.L. (1972): Size-biased sampling, *Technometrics*, 14, 635 – 644.

[24] Shanker, R. (2015): Shanker distribution and Its Applications, *International Journal of Statistics and Applications*, 5(6), 338 – 348.

[25] Shanker, R. (2016): The discrete Poisson-Shanker distribution, *Jacobs Journal of Biostatistics*, 1(1), 1 – 7

[26] Shanker, R., Hagos, F. and Abrehe, Y. (2015): On Size –Biased Poisson-Lindley Distribution and Its Applications to Model Thunderstorms, *American Journal of Mathematics and Statistics*, 5 (6), 354 – 360.

[27] Singh, S.K. and Maddala, G.S. (1976): A function for the size distribution of incomes, *Econometrica*, 44, 963 – 970.

[28] Van Deusen, P.C. (1986): Fitting assumed distributions to horizontal point sample diameters, *For. Sci*., 32, 146 -148.