International Journal of Statistics and Applications 2016, 6(6): 376-385
DOI: 10.5923/j.statistics.20160606.06

A Size-Biased Poisson-Amarendra Distribution and Its Applications

Rama Shanker[1,*], **Hagos Fesshaye**[2]

[1]Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea
[2]Department of Economics, College of Business and Economics, Halhale, Eritrea

Abstract In this paper, a size-biased Poisson-Amarendra distribution (SBPAD) has been proposed by size-biasing the Poisson-Amarendra distribution (PAD) of Shanker (2016 b), a Poisson mixture of Amarendra distribution introduced by Shanker (2016 a). The first four moments (about origin) and the central moments (about mean) have been obtained and expressions for coefficient of variation (C.V), skewness, kurtosis and index of dispersion have been given. The estimation of its parameter has been discussed using maximum likelihood estimation and method of moments. Three examples of real data-sets have been presented to test the goodness of fit of SBPAD over size-biased Poisson distribution (SBPD), size-biased Poisson-Lindley distribution (SBPLD) and size-biased Poisson-Sujatha distribution (SBPSD).

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

1. Introduction

Size-biased distributions are a particular class of weighted distributions which arise naturally in practice when observations from a sample are recorded with probability proportional to some measure of unit size. In field applications, size-biased distributions can arise either because individuals are sampled with unequal probability by design or because of unequal detection probability. Size-biased distributions come into play when organisms occur in groups, and group size influences the probability of detection. Fisher (1934) firstly introduced these distributions to model ascertainment biases which were later formalized by Rao (1965) in a unifying theory for problems where the observations fall in non-experimental, non-replicated and non-random categories. Size-biased distributions have applications in environmental science, econometrics, social science, biomedical science, human demography, ecology, geology, forestry etc. Van Duesen (1986) has detailed study about the applications of size-biased distributions for fitting distributions of diameter at breast height (DBH) data arising from horizontal point sampling (HPS). Later, Lappi and Bailey (1987) have applied size-biased distributions to analyze HPS diameter increment data. The 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 fields of knowledge by different 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 probability distribution $P_0(x;\theta)\,;x=0,1,2,...,\theta>0$. If sample units are weighted or selected with probability proportional to x^{α}, then the corresponding size-biased distribution of order α is given by its probability mass function

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

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

2. Size-Biased Poisson-Amarendra Distribution (SBPAD)

The Poisson-Amarendra distribution (PAD) having probability mass function (p.m.f.)

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

$$P_0(x;\theta)=\frac{\theta^4}{\theta^3+\theta^2+2\theta+6}\frac{x^3+(\theta+7)x^2+\left(\theta^2+5\theta+15\right)x+\left(\theta^3+4\theta^2+7\theta+10\right)}{(\theta+1)^{x+4}}\quad ;x=0,1,2,...,\ \theta>0 \tag{2.1}$$

has been introduced by Shanker (2016 b) for modeling count data in various fields of knowledge. Its various statistical properties, estimation of parameter and applications have been discussed in details by Shanker (2016 b) and shown that it is better than Poisson distribution (PD), Poisson-Lindley distribution (PLD) of Sankaran (1970) and Poisson-Sujatha distribution (PSD) introduced by Shanker (2016 d). It would be recall that PSD is a Poisson mixture of Sujatha distribution introduced by Shanker (2016 c). Shanker and Hagos (2016 a) have detailed study about the applications of PSD for modeling discrete data from biological sciences.

The PAD arises from the Poisson distribution when its parameter λ follows Amarendra distribution introduced by Shanker (2016 a) with probability density function (p.d.f.)

$$f_0(\lambda;\theta)=\frac{\theta^4}{\theta^3+\theta^2+2\theta+6}\left(1+\lambda+\lambda^2+\lambda^3\right)e^{-\theta\lambda}\quad ;\lambda>0,\ \theta>0 \tag{2.2}$$

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

$$P_1(x;\theta)=\frac{x\cdot P_0(x;\theta)}{\mu_1'}=\frac{\theta^5}{\theta^3+2\theta^2+6\theta+24}\frac{\left[x^4+(\theta+7)x^3+\left(\theta^2+5\theta+15\right)x^2+\left(\theta^3+4\theta^2+7\theta+10\right)x\right]}{(\theta+1)^{x+4}}\quad ; \tag{2.3}$$

$x=1,2,3,..,\theta>0$

where $\mu_1{}'=\dfrac{\theta^3+2\theta^2+6\theta+24}{\theta\left(\theta^3+\theta^2+2\theta+6\right)}$ is the mean of the PAD with p.m.f. (2.1).

Recall that the p.d.f. (2.3) of SBPAD can also be obtained from the size-biased Poisson distribution (SPBD) with p.m.f.

$$g(x\,|\,\lambda)=\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\quad ;x=1,2,3,...,;\lambda>0 \tag{2.4}$$

when its parameter λ follows the size-biased Amarendra distribution (SBAD) with p.d.f.

$$h(\lambda;\theta)=\frac{\theta^5}{\theta^3+2\theta^2+6\theta+24}\lambda\left(1+\lambda+\lambda^2+\lambda^3\right)e^{-\theta\lambda}\quad ;\lambda>0,\ \theta>0 \tag{2.5}$$

We have

$$\begin{aligned}
P(X=x)&=\int_0^\infty g(x\,|\,\lambda)\cdot h(\lambda;\theta)\,d\lambda\\
&=\int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\frac{\theta^5}{\theta^3+2\theta^2+6\theta+24}\lambda\left(1+\lambda+\lambda^2+\lambda^3\right)e^{-\theta\lambda}d\lambda \qquad (2.6)\\
&=\frac{\theta^5}{\left(\theta^3+2\theta^2+6\theta+24\right)(x-1)!}\int_0^\infty e^{-(\theta+1)\lambda}\left(\lambda^x+\lambda^{x+1}+\lambda^{x+2}+\lambda^{x+3}\right)d\lambda\\
&=\frac{\theta^5}{\left(\theta^3+2\theta^2+6\theta+24\right)(x-1)!}\left[\frac{\Gamma(x+1)}{(\theta+1)^{x+1}}+\frac{\Gamma(x+2)}{(\theta+1)^{x+2}}+\frac{\Gamma(x+3)}{(\theta+1)^{x+3}}+\frac{\Gamma(x+4)}{(\theta+1)^{x+4}}\right]\\
&=\frac{\theta^5}{\theta^3+2\theta^2+6\theta+24}\frac{\begin{bmatrix}x^4+(\theta+7)x^3+\left(\theta^2+5\theta+15\right)x^2\\+\left(\theta^3+4\theta^2+7\theta+10\right)x\end{bmatrix}}{(\theta+1)^{x+4}}\quad ;x=1,2,3,..,\theta>0
\end{aligned}$$

which is the p.m.f of SBPAD.

The p.m.f. of size-biased Poisson-Sujatha distribution (SBPSD) given by

$$P_3(x;\theta)=\frac{\theta^4}{\theta^2+2\theta+6}\frac{x^3+(\theta+4)x^2+(\theta^2+3\theta+4)x}{(\theta+1)^{x+3}}\quad ;x=1,2,3,..,\theta>0 \tag{2.7}$$

has been introduced by Shanker and Hagos (2016 b), which is a size biased version of Poisson-Sujatha distribution (PSD) proposed by Shanker (2016 d).

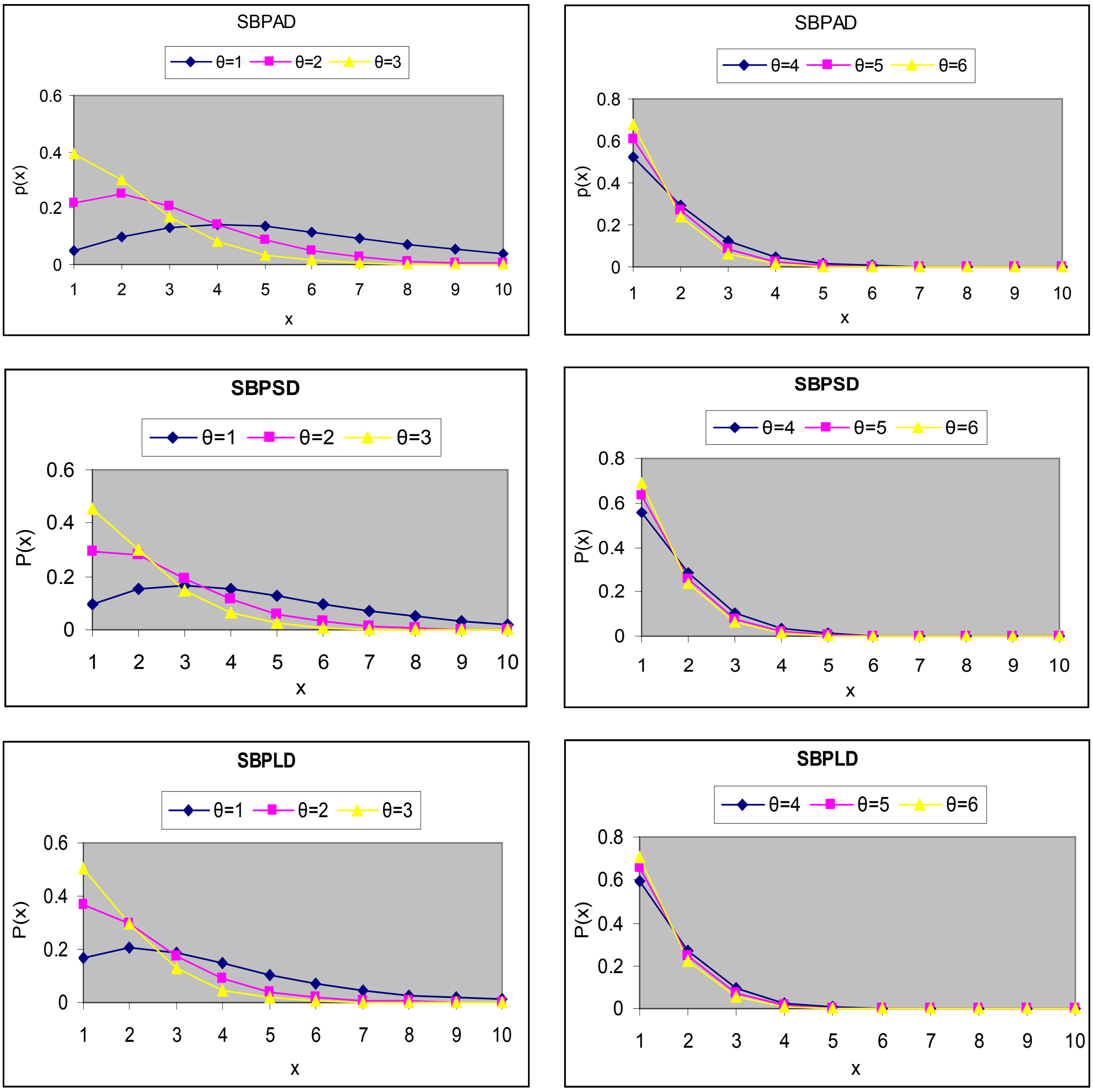

Figure 1. Graphs of pmf of SBPAD, SBPSD and SBPLD for selected values of the parameter θ

It would be recall that the p.m.f of size-biased Poisson-Lindley distribution (SBPLD) given by

$$P_2(x;\theta)=\frac{\theta^3}{\theta+2}\frac{x(x+\theta+2)}{(\theta+1)^{x+2}}\quad ;x=1,2,3,...,;\theta>0 \tag{2.8}$$

has been introduced by Ghitany and Mutairi (2008), which is a size-biased version of Poisson-Lindley distribution introduced by Sankaran (1970). The Poisson-Lindley distribution (PSD) is a Poisson mixture of Lindley (1958) distribution. Ghitany and Mutairi (2008) have discussed its various mathematical and statistical properties, estimation of the parameter using maximum likelihood estimation and the method of moments, and goodness of fit. Shanker *et al* (2015) have critical study on 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 graphs of the probability mass functions of SBPAD, SBPSD and SBPLD for selected values of their parameter θ are shown in the figure 1.

3. Moments and Associated Measures

Using (2.6), the r th factorial moment about origin of the SBPAD (2.3) can be obtained as

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

$$= \int_0^\infty \left[\sum_{x=1}^\infty x^{(r)} \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\right] \cdot \frac{\theta^5}{\theta^3+2\theta^2+6\theta+24} \lambda\left(1+\lambda+\lambda^2+\lambda^3\right)e^{-\theta\lambda} d\lambda$$

$$= \int_0^\infty \left[\lambda^{r-1}\left\{\sum_{x=r}^\infty x \frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right\}\right] \cdot \frac{\theta^5}{\theta^3+2\theta^2+6\theta+24} \lambda\left(1+\lambda+\lambda^2+\lambda^3\right)e^{-\theta\lambda} d\lambda$$

Taking $x = x + r$, we get

$$\mu_{(r)}{}' = \frac{\theta^5}{\theta^3+2\theta^2+6\theta+24} \int_0^\infty \left[\lambda^{r-1}\left\{\sum_{x=0}^\infty (x+r) \frac{e^{-\lambda}\lambda^{x}}{x!}\right\}\right] \cdot \left(\lambda+\lambda^2+\lambda^3+\lambda^4\right)e^{-\theta\lambda} d\lambda$$

$$= \frac{\theta^5}{\theta^3+2\theta^2+6\theta+24} \int_0^\infty \lambda^{r-1}(\lambda+r)\left(\lambda+\lambda^2+\lambda^3+\lambda^4\right)e^{-\theta\lambda} d\lambda$$

$$= \frac{\theta^5}{\theta^3+2\theta^2+6\theta+24} \int_0^\infty \left(\lambda^r+r\lambda^{r-1}\right)\left(\lambda+\lambda^2+\lambda^3+\lambda^4\right)e^{-\theta\lambda} d\lambda$$

Using gamma integral and some small algebraic simplification, the r th factorial moment about origin of SBPAD (2.3) can be obtained as

$$\mu_{(r)}{}' = \frac{r!\left\{\begin{array}{l} r\theta^4+(r+1)^2\theta^3+(r+1)^2(r+2)\theta^2+(r+1)^2(r+2)(r+3)\theta \\ +(r+1)(r+2)(r+3)(r+4) \end{array}\right\}}{\theta^r\left(\theta^3+2\theta^2+6\theta+24\right)}; \quad r = 1,2,3,.... \tag{3.1}$$

Taking r = 1,2,3, and 4 in (3.1), the first four factorial moments about origin can be obtained and using the relationship between moments about origin and factorial moments about origin, the first four moments about origin of the SBPAD (2.3) are thus obtained as

$$\mu_1{}' = \frac{\theta^4+4\theta^3+12\theta^2+48\theta+120}{\theta\left(\theta^3+2\theta^2+6\theta+24\right)}$$

$$\mu_2{}' = \frac{\theta^5+8\theta^4+30\theta^3+120\theta^2+480\theta+720}{\theta^2\left(\theta^3+2\theta^2+6\theta+24\right)}$$

$$\mu_3{}' = \frac{\theta^6+16\theta^5+84\theta^4+360\theta^3+1680\theta^2+5040\theta+5040}{\theta^3\left(\theta^3+2\theta^2+6\theta+24\right)}$$

$$\mu_4{}' = \frac{\theta^7+32\theta^6+248\theta^5+1224\theta^4+6120\theta^3+25920\theta^2+55440\theta+40320}{\theta^4\left(\theta^3+2\theta^2+6\theta+24\right)}$$

Now, using the relationship between central moments and the moments about origin, the central moments of the SBPAD (2.3) are thus obtained as

$$\mu_2 = \frac{2\left(\theta^7 + 6\theta^6 + 30\theta^5 + 162\theta^4 + 504\theta^3 + 1008\theta^2 + 2160\theta + 1440\right)}{\theta^2\left(\theta^3 + 2\theta^2 + 6\theta + 24\right)^2}$$

$$\mu_3 = \frac{2\begin{pmatrix}\theta^{11} + 10\theta^{10} + 66\theta^9 + 408\theta^8 + 2040\theta^7 + 7320\theta^6 + 21744\theta^5 + 53136\theta^4 \\ +88992\theta^3 + 142560\theta^2 + 155520\theta + 69120\end{pmatrix}}{\theta^3\left(\theta^3 + 2\theta^2 + 6\theta + 24\right)^3}$$

$$\mu_4 = \frac{2\begin{pmatrix}\theta^{15} + 23\theta^{14} + 220\theta^{13} + 1650\theta^{12} + 10988\theta^{11} + 57852\theta^{10} + 246960\theta^9 + 918144\theta^8 \\ +2853504\theta^7 + 7482240\theta^6 + 17285184\theta^5 + 32716224\theta^4 + 50906880\theta^3 \\ +64696320\theta^2 + 52254720\theta + 17418240\end{pmatrix}}{\theta^4\left(\theta^3 + 2\theta^2 + 6\theta + 24\right)^4}$$

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 SBPAD (2.3) are finally obtained as

$$C.V = \frac{\sigma}{\mu_1'} = \frac{\sqrt{2\left(\theta^7 + 6\theta^6 + 30\theta^5 + 162\theta^4 + 504\theta^3 + 1008\theta^2 + 2160\theta + 1440\right)}}{\left(\theta^4 + 4\theta^3 + 12\theta^2 + 48\theta + 120\right)}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\begin{pmatrix}\theta^{11} + 10\theta^{10} + 66\theta^9 + 408\theta^8 + 2040\theta^7 + 7320\theta^6 + 21744\theta^5 + 53136\theta^4 \\ +88992\theta^3 + 142560\theta^2 + 155520\theta + 69120\end{pmatrix}}{\sqrt{2}\left(\theta^7 + 6\theta^6 + 30\theta^5 + 162\theta^4 + 504\theta^3 + 1008\theta^2 + 2160\theta + 1440\right)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\begin{pmatrix}\theta^{15} + 23\theta^{14} + 220\theta^{13} + 1650\theta^{12} + 10988\theta^{11} + 57852\theta^{10} + 246960\theta^9 + 918144\theta^8 \\ +2853504\theta^7 + 7482240\theta^6 + 17285184\theta^5 + 32716224\theta^4 + 50906880\theta^3 \\ +64696320\theta^2 + 52254720\theta + 17418240\end{pmatrix}}{2\left(\theta^7 + 6\theta^6 + 30\theta^5 + 162\theta^4 + 504\theta^3 + 1008\theta^2 + 2160\theta + 1440\right)^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{2\left(\theta^7 + 6\theta^6 + 30\theta^5 + 162\theta^4 + 504\theta^3 + 1008\theta^2 + 2160\theta + 1440\right)}{\theta\left(\theta^3 + 2\theta^2 + 6\theta + 24\right)\left(\theta^4 + 4\theta^3 + 12\theta^2 + 48\theta + 120\right)}$$

The conditions under which SBPAD, SBPSD and SBPLD are over-dispersed $\left(\mu < \sigma^2\right)$, equi-dispersed $\left(\mu = \sigma^2\right)$ and under-dispersed $\left(\mu > \sigma^2\right)$ are presented in table 3.1.

Table 3.1. Condition for over-dispersion, equi-dispersion, and under-dispersion for SBPAD, SBPSD and SBPLD

Distribution	Over-dispersion $\left(\mu < \sigma^2\right)$	Equi-dispersion $\left(\mu = \sigma^2\right)$	Under-dispersion $\left(\mu > \sigma^2\right)$
SBPAD	$\theta < 2.273739$	$\theta = 2.273739$	$\theta > 2.273739$
SBPSD	$\theta < 1.961384$	$\theta = 1.961384$	$\theta > 1.961384$
SBPLD	$\theta < 1.671162$	$\theta = 1.671162$	$\theta > 1.671162$

To study the characteristics of $\mu_1', \mu_2, C.V, \sqrt{\beta_1}, \beta_2$ and γ of SBPAD, SBPSD and SBPLD for varying values of their parameter θ, the numerical values of these characteristics have been presented in table 3.2.

Table 3.2. Characteristics of $\mu_1', \mu_2, C.V, \sqrt{\beta_1}, \beta_2$ and γ for SBPAD, SBPSD and SBPLD for selected values of the parameter θ

	Values of θ for SBPAD					
	1	**2**	**3**	**4**	**5**	**6**
μ_1'	5.606061	3.00000	2.149425	1.763889	1.558952	1.436782
μ_2	9.753903	3.307692	1.690316	1.034529	0.7129	0.53336
CV	0.557098	0.606235	0.604869	0.576634	0.541604	0.508299
$\sqrt{\beta_1}$	0.988214	1.208372	1.451176	1.657626	1.826565	1.970615
β_2	4.375715	4.939156	5.773739	6.609862	7.343342	7.986343
γ	1.739885	1.102564	0.786404	0.586505	0.457295	0.371218

	Values of θ for SBPSD					
	1	**2**	**3**	**4**	**5**	**6**
μ_1'	4.555556	2.571429	1.952381	1.666667	1.507317	1.407407
μ_2	7.580247	2.530612	1.346939	0.872222	0.630434	0.48834
CV	0.604366	0.61864	0.594442	0.560357	0.526763	0.496525
$\sqrt{\beta_1}$	1.138007	1.360076	1.554876	1.719786	1.864608	1.996131
β_2	4.820343	5.497561	6.191562	6.820256	7.394639	7.935729
γ	1.663957	0.984127	0.689895	0.523333	0.418249	0.346979

	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 graphs 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 $\left(\gamma\right)$ of SBPAD, SBPSD and SBPLD are shown in figure 2

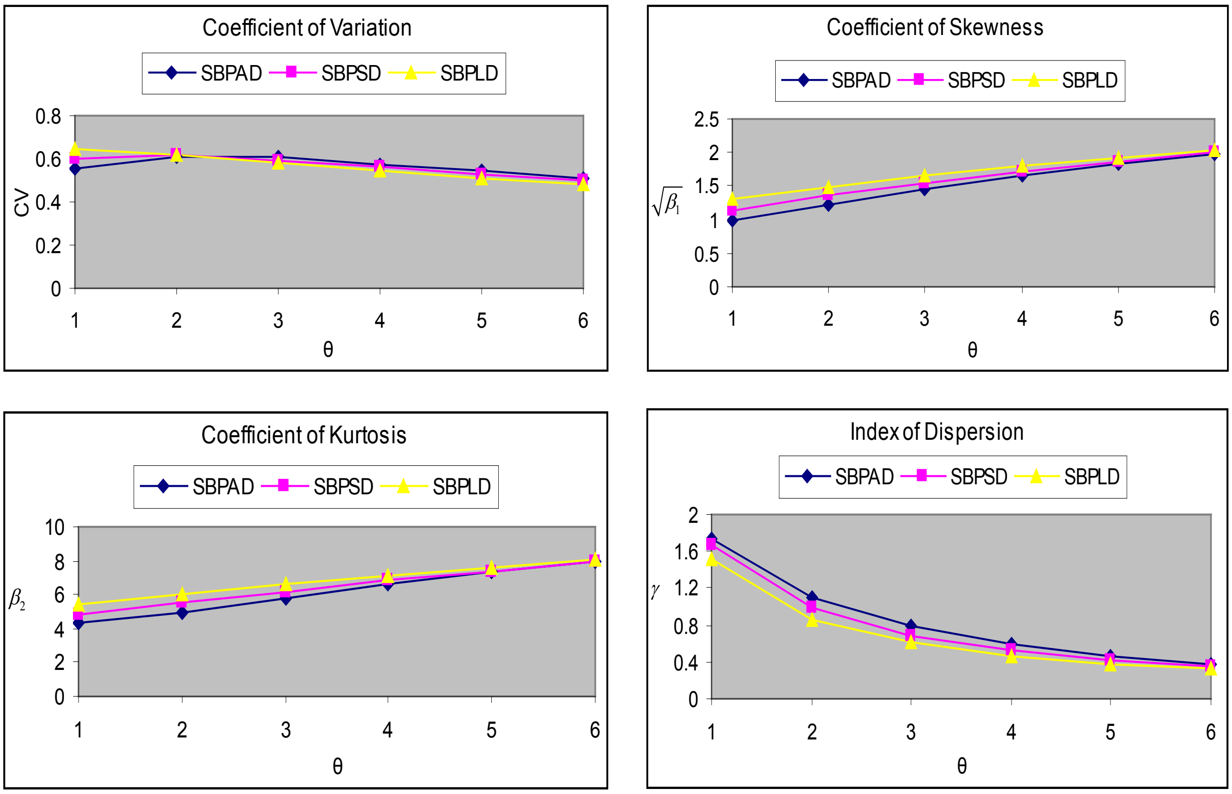

Figure 2. Graphs of coefficient of variation (C.V), coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis $\left(\beta_2\right)$, and index of dispersion $\left(\gamma\right)$ for SBPAD, SBPSD and SBPLD for selected values of their parameter θ

4. Unimodality and Increasing Failure Rate

Using p.m.f. of SBPAD from (2.3), we have

$$\frac{P_1(x+1;\theta)}{P_1(x;\theta)}=\left(\frac{1}{\theta+1}\right)\left(1+\frac{1}{x}\right)\left(1+\frac{3x^2+(2\theta+17)x+\left(\theta^2+6\theta+23\right)}{x^3+(\theta+7)x^2+\left(\theta^2+5\theta+15\right)x+\left(\theta^3+4\theta^2+7\theta+10\right)}\right)$$

which is a deceasing function of x, $P_1(x;\theta)$ is log-concave. Therefore, SBPAD 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). The definitions, concepts and interrelationship between these aging concepts have been discussed in Barlow and Proschan (1981).

5. Parameter Estimation

5.1. Maximum Likelihood Estimate (MLE): Let $x_1,x_2,...,x_n$ be a random sample of size n from the SBPAD (2.3) 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 SBPAD (2.3) is given by

$$L=\left(\frac{\theta^5}{\theta^3+2\theta^2+6\theta+24}\right)^n \frac{1}{(\theta+1)^{\sum_{x=1}^{k} f_x(x+4)}}\prod_{x=1}^{k}\left[\begin{array}{l}x^4+(\theta+7)x^3+\left(\theta^2+5\theta+15\right)x^2\\+\left(\theta^3+4\theta^2+7\theta+10\right)x\end{array}\right]^{f_x}$$

The log likelihood function can be obtained as

$$\log L = n\log\left(\frac{\theta^5}{\theta^3+2\theta^2+6\theta+24}\right)-\sum_{x=1}^{k} f_x\left(x+4\right)\log\left(\theta+1\right)$$

$$+\sum_{x=1}^{k} f_x \log\left[x^4+\left(\theta+7\right)x^3+\left(\theta^2+5\theta+15\right)x^2+\left(\theta^3+4\theta^2+7\theta+10\right)x\right]$$

The first derivative of the log likelihood function is thus given by

$$\frac{d\log L}{d\theta}=\frac{5n}{\theta}-\frac{n\left(3\theta^2+4\theta+6\right)}{\theta^3+2\theta^2+6\theta+24}-\frac{n\left(\bar{x}+4\right)}{\theta+1}+\sum_{x=1}^{k}\frac{\left[x^2+\left(2\theta+5\right)x+\left(3\theta^2+8\theta+7\right)\right]f_x}{\left[x^3+\left(\theta+7\right)x^2+\left(\theta^2+5\theta+15\right)x+\left(\theta^3+4\theta^2+7\theta+10\right)\right]}$$

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

The maximum likelihood estimate (MLE), $\hat{\theta}$ of θ of SBPAD (2.3) is the solution of the equation $\frac{d\log L}{d\theta}=0$ and is given by the solution of the following non-linear equation

$$\frac{5n}{\theta}-\frac{n\left(3\theta^2+4\theta+6\right)}{\theta^3+2\theta^2+6\theta+24}-\frac{n\left(\bar{x}+4\right)}{\theta+1}+\sum_{x=1}^{k}\frac{\left[x^2+\left(2\theta+5\right)x+\left(3\theta^2+8\theta+7\right)\right]f_x}{\left[x^3+\left(\theta+7\right)x^2+\left(\theta^2+5\theta+15\right)x+\left(\theta^3+4\theta^2+7\theta+10\right)\right]}=0$$

This non-linear equation can be solved by any numerical iteration methods such as Newton- Raphson method, Bisection method, Regula –Falsi method etc. In the present paper, Newton-Raphson method has been used to solve the above non-linear equation to find MLE of the parameter.

5.2. Method of Moment Estimate (MOME): Equating the population mean to the corresponding sample mean, the method of moment estimate (MOME) $\tilde{\theta}$ of θ of SBPAD (2.3) is the solution of the following fourth degree polynomial equation in θ

$$\left(1-\bar{x}\right)\theta^4+2\left(2-\bar{x}\right)\theta^3+6\left(2-\bar{x}\right)\theta^2+24\left(2-\bar{x}\right)\theta+120=0$$

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

6. Goodness of Fit

In this section the goodness of fit of the SBPAD, SBPSD, SBPLD and SBPD have been presented for three count data- sets. The fitting of these distributions are based on maximum likelihood estimates of the parameter. 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, the second data-set is animal abundance data of Keith and Meslow (1968) regarding the distribution of snowshoe hares captured over 7 days, and the third data-set is number of counts of pairs of running shoes owned by 60 members of an athletics club, reported by Simonoff (2003).

Table 6.1. Distribution of number of counts of sites with particles from Immunogold data

No. of sites with particles	Observed Frequency	Expected Frequency			
		SBPD	SBPLD	SBPSD	SBPAD
1	122	111.3	119.0	119.3	119.4
2	50	64.1	53.8	53.4	53.3
3	18	18.5	18.0	17.9	17.9
4	4	3.5	5.3	5.3	5.3
5	4	0.6	1.9	2.1	2.1
Total	198	198.0	198.0	198.0	198.0
ML estimate		$\hat{\theta}=0.575758$	$\hat{\theta}=4.050987$	$\hat{\theta}=4.511904$	$\hat{\theta}=4.898958$
χ^2		4.64	0.43	0.32	0.29
d.f.		1	2	2	2
p-value		0.0312	0.8065	0.8521	0.8650

Table 6.2. Distribution of snowshoe hares captured over 7 days

No. times hares caught	Observed Frequency	Expected Frequency			
		SBPD	SBPLD	SBPSD	SBPAD
1	184	170.6	177.3	177.5	178.8
2	55	72.5	62.5	62.3	62.0
3	14	15.4	16.4	16.4	16.3
4	4	2.2	3.8	3.8	3.8
5	4	0.3	1.0	1.0	1.1
Total	261	261.0	261.0	261.0	261.0
ML estimate		$\hat{\theta} = 0.425287$	$\hat{\theta} = 5.351256$	$\hat{\theta} = 5.799735$	$\hat{\theta} = 6.131809$
χ^2		6.22	1.18	1.11	1.03
d.f.		1	1	1	1
p-value		0.0126	0.2773	0.2921	0.3101

Table 6.3. Number of counts of pairs of running shoes owned by 60 members of an athletics club, reported by Simonoff (2003, p. 100)

Number of pairs of running shoes	Observed frequency	Expected Frequency			
		SBPD	SBPLD	SBPSD	SBPAD
1	18	15.0	20.3	20.0	19.8
2	18	20.8	17.4	17.5	17.5
3	12	14.4	10.9	11.1	11.2
4	7	6.6	5.9	6.0	6.1
5	5	3.2	5.5	5.4	5.4
Total	60	60.0	60.0	60.0	60.0
ML Estimate		$\hat{\theta} = 1.383333$	$\hat{\theta} = 1.818978$	$\hat{\theta} = 2.208089$	$\hat{\theta} = 2.609819$
χ^2		1.87	0.64	0.47	0.39
d.f.		2	3	3	3
p-value		0.3926	0.8872	0.9254	0.9423

7. Concluding Remarks

A size-biased Poisson mixture of size-biased Amarendra distribution named, “size-biased Poisson-Amarendra distribution (SBPAD)” has been proposed by size-biasing the Poisson-Amarendra distribution (PAD) of Shanker (2016 b). Its moments and other structural properties including coefficient of variation, skewness, kurtosis, index of dispersion, hazard rate and unimodality have been studied. The estimation of its parameter has been discussed using method of maximum likelihood and that of moments. Three examples of real data-sets have been presented to test the goodness of fit of SBPAD over the size-biased Poisson distribution (SBPD), the size-biased Poisson-Lindley distribution (SBPLD) and the size-biased Poisson -Sujatha distribution (SBPSD). The fit of SBPAD over SBPD, SBPLD, and SBPSD shows that SBPAD can be considered as an important distribution for modeling data which structurally excludes zero counts.

ACKNOWLEDGEMENTS

The authors express their gratitude to the editor-in-chief and the anonymous reviewer for his constructive suggestions which improved the quality 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] Drummer, T.D. and MacDonald, L.L. (1987): Size biased in line transect sampling, *Biometrics*, 43, 13 – 21.

[5] 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.

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

[7] 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.

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

[13] 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.

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

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

[16] 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.

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

[18] Patil, G.P. and Rao, C.R. (1977): The Weighted distributions: A survey and their applications. In applications of Statistics (Ed P.R. Krishnaiah0, 383 – 405, North Holland Publications Co., Amsterdam.

[19] 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.

[20] 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.

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

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

[23] Shanker, R. (2016a): Amarendra distribution and Its Applications, *American Journal of Mathematics and Statistics*, 6(1), 44 – 56.

[24] Shanker, R. (2016b): The discrete Poisson-Amarendra distribution, to appear in *International Journal of Statistical distributions and Applications*, 2(2), 14 - 21

[25] Shanker, R. (2016c): Sujatha distribution and Its Applications, *Statistics in Transition-new Series*, 17 (3), 1 – 20.

[26] Shanker, R. (2016d): The discrete Poisson-Sujatha distribution, *International Journal of Probability and Statistics*, 5(1), 1 – 9

[27] 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

[28] Shanker, R. and Hagos, F. (2016a): On Poisson-Sujatha distribution and its Applications to Model count data from Biological Science, *Biometrics and Biostatistics International Journal,* 3(4), 1 – 7.

[29] Shanker, R. and Hagos, F. (2016b): Size-biased Poisson-Sujatha distribution with Applications, *American Journal of Mathematics and Statistics,* 6(4), 145 – 154.

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

[31] Simmonoff, J.S. (2003): Analyzing Categorical data, Springer, New york

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