

# A Size-Biased Poisson-Aradhana Distribution with Applications

Rama Shanker<sup>1,\*</sup>, Kamlesh Kumar Shukla<sup>1</sup>, Ravi Shanker<sup>2</sup>, Tekie Asehun Leonida<sup>3</sup>

<sup>1</sup>Department of Statistics, College of Science, Eritrea Institute of Technology, Asmara, Eritrea

<sup>2</sup>Department of Mathematics, G.L.A. College, N.P University, Daltonganj, Jharkhand, India

<sup>3</sup>Department of Applied Mathematics, University of Twente, The Netherlands

**Abstract** In this present paper, a size-biased Poisson-Aradhana distribution (SBPAD) has been proposed and its nature has been studied graphically. Its moments and moments based measures including coefficients of variation, skewness, kurtosis and index of dispersion have been obtained and their natures have been discussed graphically. The unimodality and increasing hazard rate function of the distribution has been discussed. Method of moments and the method of maximum likelihood have been discussed for estimating the parameter. Applications of SBPAD have been explained through three examples and it gives much better fit over size-biased Poisson distribution (SBPD) and size-biased Poisson-Lindley distribution (SBPLD).

**Keywords** Size-biasing, Aradhana distribution, Poisson-Aradhana distribution, Moments, Skewness, Kurtosis, Estimation, Applications

## 1. Introduction

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

$$P_1(x; \theta) = \frac{x^\alpha \cdot P_0(x; \theta)}{\mu'_\alpha}, \quad (1.1)$$

where  $\mu'_\alpha = E(X^\alpha) = \sum_{x=0}^{\infty} x^\alpha P_0(x; \theta)$ . The simple size-biased distribution and area-biased distribution are the particular cases of (1.1) for  $\alpha = 1$  and  $\alpha = 2$  respectively. Note that simple size-biased distribution has applications in size-biased sampling and area-biased distribution has been used in area-biased sampling.

When organism occurs in groups and the size of the group influences the probability of detection, size-biased distributions are the appropriate choice. Size-biased distributions are a special class of weighted distributions which arise naturally in many real life situations when observations from a sample are recorded with probability proportional to some measure of unit size, known as probability proportional to size (PPS). In the field applications, size-biased distributions can arise due to sampling of individuals with unequal probability by design and unequal detection probability. The concept of weighted distributions can be traced to the study of the effect of methods of ascertainment upon frequencies by Fisher (1934). In extending the basic ideas of Fisher, Rao (1965) have seen the need for a unifying concept and identified various sampling situations that can be modeled by weighted distributions. Size-biased distributions have applications in almost every branch of knowledge namely, social science, econometrics, environmental science, biomedical science, human demography,

\* Corresponding author:

shankerrama2009@gmail.com (Rama Shanker)

Published online at <http://journal.sapub.org/ajms>

Copyright © 2018 The Author(s). Published by Scientific & Academic Publishing

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>

ecology, geology, forestry, are some among others. The fitting of distributions of diameter at breast height (DBH) data arising from horizontal point sampling (HPS) using size-biased distributions has been discussed by Van Duesen (1986). Similarly, the analysis of HPS diameter increment data using size-biased distributions has been discussed by Lappi and Bailey (1987). Patil and Rao (1977, 1978) have detailed discussions and applications of size-biased distributions to the analysis and modeling of observed data relating to human population and ecology. Patil (1991, 1996, and 1997) has pursued weighted distributions for the purpose of encountered data analysis, equilibrium population analysis subject to harvesting and predation, meta-analysis incorporating publication bias and heterogeneity, modeling cluster and extraneous variation, etc., and has detailed discussion on applications of size-biased distributions in biostatistics, ecology, environment and risk assessment. A number of papers have appeared during a short period of time implicitly using the concept of weighted and size-biased distributions and their applications in various fields of knowledge by researchers in statistics, namely, Scheaffer (1972), Patil and Ord (1976), Singh and Maddala (1976), Patil (1981), McDonald (1984), Drummer and McDonald (1987), Gove (2000, 2003), Correa and Wolfson (2007), Ducey (2009), Alavi and Chinipardaz (2009), Ducey and Gove (2015), are some among others.

Shanker (2017) has introduced Poisson-Aradhana distribution (PAD) defined by pmf

$$P_0(x; \theta) = \frac{\theta^3}{\theta^2 + 2\theta + 2} \frac{x^2 + (2\theta + 5)x + (\theta^2 + 4\theta + 5)}{(\theta + 1)^{x+3}}; x = 0, 1, 2, \dots, \theta > 0 \quad (1.2)$$

Its important statistical properties, estimation of parameter and applications have been discussed by Shanker (2017). It has been shown by Shanker (2017) that PAD gives much better fit than both Poisson distribution and Poisson-Lindley distribution (PLD). It should be noted that PAD is a Poisson mixture of Aradhana distribution when the parameter  $\lambda$  of the Poisson distribution follows Aradhana distribution introduced by Shanker (2016) having probability density function (pdf.)

$$f_0(\lambda; \theta) = \frac{\theta^3}{\theta^2 + 2\theta + 2} (1 + 2\lambda + \lambda^2) e^{-\theta\lambda}; \lambda > 0, \theta > 0 \quad (1.3)$$

The pmf of the size-biased Poisson-Aradhana distribution (SBPAD) with parameter  $\theta$  can be obtained as

$$P_2(x; \theta) = \frac{x \cdot P_0(x; \theta)}{\mu_1'} = \frac{\theta^4}{\theta^2 + 4\theta + 6} \frac{x^3 + (2\theta + 5)x^2 + (\theta^2 + 4\theta + 5)x}{(\theta + 1)^{x+3}}; x = 1, 2, 3, \dots, \theta > 0 \quad (1.4)$$

where  $\mu_1' = \frac{\theta^2 + 4\theta + 6}{\theta(\theta^2 + 2\theta + 2)}$  is the population mean of the PAD having pmf. (1.2). The pmf of SBPAD (1.4) can also be

obtained from the size-biased Poisson distribution (SPBD) having pmf

$$g(x | \lambda) = \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!}; x = 1, 2, 3, \dots; \lambda > 0 \quad (1.5)$$

when the parameter  $\lambda$  of SBPD follows the size-biased Aradhana distribution (SBAD) having pdf

$$h(\lambda; \theta) = \frac{\theta^4}{\theta^2 + 4\theta + 6} \lambda (1 + 2\lambda + \lambda^2) e^{-\theta\lambda}; \lambda > 0, \theta > 0 \quad (1.6)$$

Thus, we have

$$\begin{aligned} P(X = x) &= \int_0^\infty g(x | \lambda) \cdot h(\lambda; \theta) d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} \frac{\theta^4}{\theta^2 + 4\theta + 6} \lambda (1 + 2\lambda + \lambda^2) e^{-\theta\lambda} d\lambda \\ &= \frac{\theta^4}{(\theta^2 + 4\theta + 6)(x-1)!} \int_0^\infty e^{-(\theta+1)\lambda} (\lambda^x + 2\lambda^{x+1} + \lambda^{x+2}) d\lambda \end{aligned} \quad (1.7)$$

$$\begin{aligned}
&= \frac{\theta^4}{(\theta^2 + 4\theta + 6)(x-1)!} \left[ \frac{\Gamma(x+1)}{(\theta+1)^{x+1}} + \frac{2\Gamma(x+2)}{(\theta+1)^{x+2}} + \frac{\Gamma(x+3)}{(\theta+1)^{x+3}} \right] \\
&= \frac{\theta^4}{\theta^2 + 4\theta + 6} \frac{x^3 + (2\theta+5)x^2 + (\theta^2 + 4\theta + 5)x}{(\theta+1)^{x+3}} ; x=1,2,3,\dots,\theta>0
\end{aligned}$$

which is the pmf of SBPAD obtained earlier in (1.4).

Ghitany and Mutairi (2008), introduced the size-biased Poisson-Lindley distribution (SBPLD) having pmf

$$P_3(x; \theta) = \frac{\theta^3}{\theta+2} \frac{x(x+\theta+2)}{(\theta+1)^{x+2}} ; x=1,2,3,\dots; \theta>0 \quad (1.8)$$

Note that SBPLD is a simple size-biased version of Poisson–Lindley distribution (PLD), introduced by Sankaran (1970) and it is a Poisson mixture of Lindley (1958) distribution. Ghitany and Mutairi (2008) have discussed its statistical properties, estimation of the parameter, and goodness of fit. Shanker *et al* (2015) has detailed study on applications of SBPLD for modeling data on thunderstorms and shown that in majority of datasets, SBPLD gives much better fit than SBPD.

The main purpose of this paper is to propose a size-biased Poisson-Aradhana distribution (SBPAD) as a size-biased Poisson mixture of size-biased Aradhana distribution and investigate some of its properties. Its statistical properties based on moments, unimodality, and increasing hazard rate function have been discussed. The estimation of parameter has been discussed using both the method of moments and the maximum likelihood. Applications of the distribution have been explained through three real observed datasets from biological sciences and the fit has been compared with SBPD and SBPLD.

The nature of SBPAD for some values of parameter  $\theta$  has been shown in figure 1.

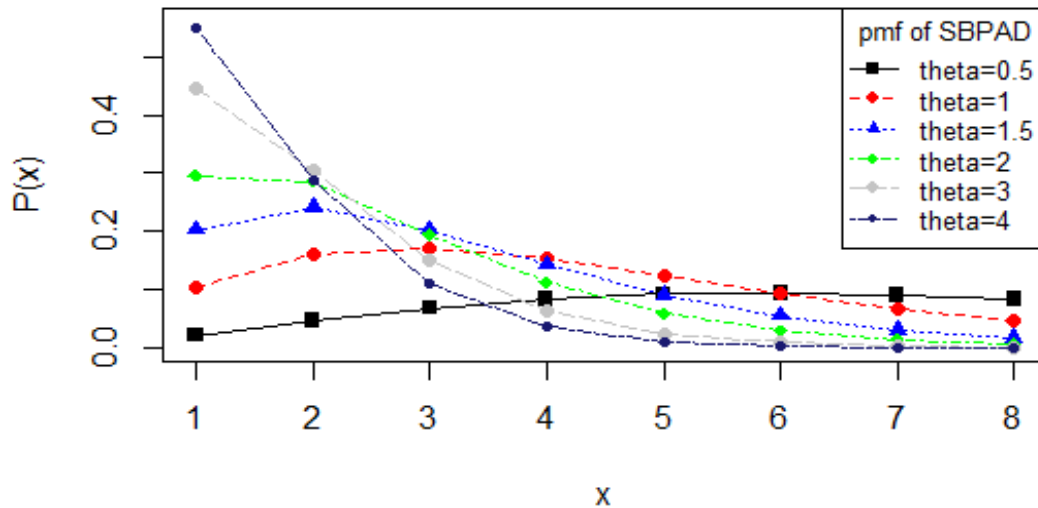


Figure 1. Nature of SBPAD for some values of the parameter  $\theta$

## 2. Moments, Skewness, Kurtosis and Index of Dispersion

The  $r$ th factorial moment about origin, using (1.7), of the SBPAD (1.4) can be obtained as

$$\begin{aligned}
\mu_{(r)}' &= E \left[ E \left( X^{(r)} \mid \lambda \right) \right], \text{ where } X^{(r)} = X(X-1)(X-2)\dots(X-r+1) \\
&= \int_0^\infty \left[ \sum_{x=1}^\infty x^{(r)} \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} \right] \cdot \frac{\theta^4}{\theta^2 + 4\theta + 6} \lambda (1 + 2\lambda + \lambda^2) e^{-\theta\lambda} d\lambda \\
&= \int_0^\infty \left[ \lambda^{r-1} \left\{ \sum_{x=r}^\infty x \frac{e^{-\lambda} \lambda^{x-r}}{(x-r)!} \right\} \right] \cdot \frac{\theta^4}{\theta^2 + 4\theta + 6} \lambda (1 + 2\lambda + \lambda^2) e^{-\theta\lambda} d\lambda
\end{aligned}$$

Taking  $y = x - r$ , we get

$$\begin{aligned}\mu_{(r)}' &= \int_0^\infty \left[ \lambda^{r-1} \left\{ \sum_{y=0}^\infty (y+r) \frac{e^{-\lambda} \lambda^y}{y!} \right\} \right] \cdot \frac{\theta^4}{\theta^2 + 4\theta + 6} \lambda (1 + 2\lambda + \lambda^2) e^{-\theta \lambda} d\lambda \\ &= \frac{\theta^4}{\theta^2 + 4\theta + 6} \int_0^\infty \lambda^{r-1} (\lambda + r) \lambda (1 + 2\lambda + \lambda^2) e^{-\theta \lambda} d\lambda \\ &= \frac{\theta^4}{\theta^2 + 4\theta + 6} \int_0^\infty (\lambda^{r+1} + r\lambda^r) (1 + 2\lambda + \lambda^2) e^{-\theta \lambda} d\lambda\end{aligned}$$

After a simple algebraic simplification using gamma integral, the  $r$ th factorial moment about origin of SBPAD (1.4) can be expressed as

$$\mu_{(r)}' = \frac{r! \{ r\theta^3 + (r+1)(2r+1)\theta^2 + (r+1)(r+2)^2\theta + (r+1)(r+2)(r+3) \}}{\theta^r (\theta^2 + 4\theta + 6)}; \quad r = 1, 2, 3, \dots \quad (2.1)$$

Taking  $r = 1, 2, 3$ , and  $4$  in (2.1) the first four factorial moments can be obtained. Then using the relationship between moments about origin (raw moments) and factorial moments, the first four moments about origin of the SBPAD (1.4) are thus obtained as

$$\begin{aligned}\mu_1' &= \frac{\theta^3 + 6\theta^2 + 18\theta + 24}{\theta(\theta^2 + 4\theta + 6)} \\ \mu_2' &= \frac{\theta^4 + 10\theta^3 + 48\theta^2 + 120\theta + 120}{\theta^2(\theta^2 + 4\theta + 6)} \\ \mu_3' &= \frac{\theta^5 + 18\theta^4 + 126\theta^3 + 480\theta^2 + 960\theta + 720}{\theta^3(\theta^2 + 4\theta + 6)} \\ \mu_4' &= \frac{\theta^6 + 34\theta^5 + 336\theta^4 + 1800\theta^3 + 5520\theta^2 + 8640\theta + 5040}{\theta^4(\theta^2 + 4\theta + 6)}\end{aligned}$$

Now, using the relationship  $\mu_r = E(Y - \mu_1')^r = \sum_{k=0}^r \binom{r}{k} \mu_k' (-\mu_1')^{r-k}$  between central moments and raw moments, the central moments of the SBPAD (1.4) are given by

$$\begin{aligned}\mu_2 &= \frac{2(\theta^5 + 11\theta^4 + 54\theta^3 + 138\theta^2 + 168\theta + 72)}{\theta^2(\theta^2 + 4\theta + 6)^2} \\ \mu_3 &= \frac{2(\theta^8 + 17\theta^7 + 138\theta^6 + 652\theta^5 + 1904\theta^4 + 3412\theta^3 + 3648\theta^2 + 2304\theta + 864)}{\theta^3(\theta^2 + 4\theta + 6)^3} \\ \mu_4 &= \frac{2 \left( \theta^{11} + 31\theta^{10} + 434\theta^9 + 3718\theta^8 + 21316\theta^7 + 85052\theta^6 + 240120\theta^5 \right. \\ &\quad \left. + 478968\theta^4 + 658848\theta^3 + 585504\theta^2 + 286848\theta + 46656 \right)}{\theta^4(\theta^2 + 4\theta + 6)^4}\end{aligned}$$

The expressions for coefficient of variation ( $C.V$ ), coefficient of Skewness ( $\sqrt{\beta_1}$ ), coefficient of Kurtosis ( $\beta_2$ ) and index of dispersion ( $\gamma$ ) of the SBPAD (1.4) are thus obtained as

$$C.V = \frac{\sigma}{\mu_1'} = \frac{\sqrt{2(\theta^5 + 11\theta^4 + 54\theta^3 + 138\theta^2 + 168\theta + 72)}}{(\theta^3 + 6\theta^2 + 18\theta + 24)}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\theta^8 + 17\theta^7 + 138\theta^6 + 652\theta^5 + 1904\theta^4 + 3412\theta^3 + 3648\theta^2 + 2304\theta + 8644}{\sqrt{2(\theta^5 + 11\theta^4 + 54\theta^3 + 138\theta^2 + 168\theta + 72)}^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\left( \theta^{11} + 31\theta^{10} + 434\theta^9 + 3718\theta^8 + 21316\theta^7 + 85052\theta^6 + 240120\theta^5 + 478968\theta^4 + 658848\theta^3 + 585504\theta^2 + 286848\theta + 46656 \right)}{2(\theta^5 + 11\theta^4 + 54\theta^3 + 138\theta^2 + 168\theta + 72)^2}$$

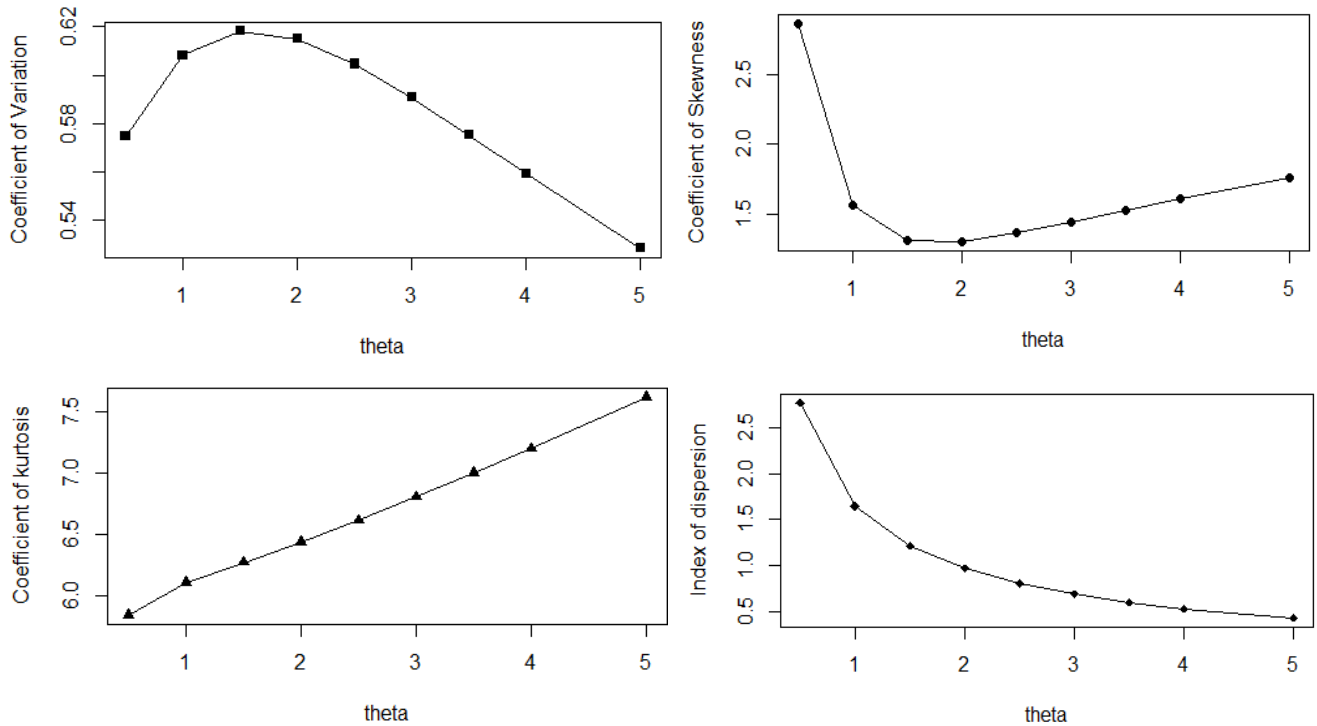
$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{2(\theta^5 + 11\theta^4 + 54\theta^3 + 138\theta^2 + 168\theta + 72)}{\theta(\theta^2 + 4\theta + 6)(\theta^3 + 6\theta^2 + 18\theta + 24)}$$

The dispersion (over-dispersion, equi-dispersion and under-dispersion) of SBPAD and SBPLD for parameter  $\theta$  are presented in table 1.

**Table 1.** Over-dispersion, equi-dispersion and under-dispersion of SBPAD and SBPLD for parameter  $\theta$

Distributions	Over-dispersion ( $\mu < \sigma^2$ )	Equi-dispersion ( $\mu = \sigma^2$ )	Under-dispersion ( $\mu > \sigma^2$ )
SBPAD	$\theta < 1.916770$	$\theta = 1.916770$	$\theta > 1.916770$
SBPLD	$\theta < 1.671162$	$\theta = 1.671162$	$\theta > 1.671162$

The nature of coefficient of variation (C.V), coefficient of Skewness ( $\sqrt{\beta_1}$ ), coefficient of Kurtosis ( $\beta_2$ ) and index of dispersion ( $\gamma$ ) of SBPAD for varying values of parameter  $\theta$  are shown graphically in figure 2



**Figure 2.** Nature of coefficient of variation, coefficient of Skewness, coefficient of Kurtosis, and index of dispersion of SBPAD for varying values of parameter  $\theta$

### 3. Statistical Properties

#### 3.1. Unimodality and Increasing Failure Rate

Since

$$\frac{P_2(x+1;\theta)}{P_2(x;\theta)} = \left(\frac{1}{\theta+1}\right) \left(1 + \frac{1}{x}\right) \left(1 + \frac{2(x+\theta+3)}{x^2 + (2\theta+5)x + (\theta^2 + 4\theta + 5)}\right)$$

is a decreasing function of  $x$ ,  $P_2(x;\theta)$  is log-concave and this means that SBPAD is unimodal, has an increasing failure rate (IFR), and hence increasing failure rate average (IFRA). Also, it is new better than used in expectation (NBUE) and has decreasing mean residual life (DMRL). A discussion about definitions, concepts and interrelationship between these reliability concepts are available in Barlow and Proschan (1981).

#### 3.2. Generating Function

The probability generating function of the SBPAD (1.4) can be obtained as

$$\begin{aligned} P_X(t) &= \frac{\theta^4}{(\theta^2 + 4\theta + 6)(\theta + 1)^3} \left[ \sum_{x=1}^{\infty} x^3 \left(\frac{t}{\theta+1}\right)^x + (2\theta+5) \sum_{x=1}^{\infty} x^2 \left(\frac{t}{\theta+1}\right)^x + (\theta^2 + 4\theta + 5) \sum_{x=1}^{\infty} x \left(\frac{t}{\theta+1}\right)^x \right] \\ &= \frac{\theta^4 t}{(\theta^2 + 4\theta + 6)(\theta + 1)^2} \left[ \frac{t^2 + 4(\theta+1)t + (\theta+1)^2}{(\theta+1-t)^4} + \frac{(2\theta+5)(\theta+1+t)}{(\theta+1-t)^3} + \frac{(\theta^2 + 4\theta + 5)}{(\theta+1-t)^2} \right] \end{aligned}$$

Thus, the moment generating function of the SBPAD (1.4) is given by

$$M_X(t) = \frac{\theta^4 e^t}{(\theta^2 + 4\theta + 6)(\theta + 1)^2} \left[ \frac{(e^t)^2 + 4(\theta+1)e^t + (\theta+1)^2}{(\theta+1-e^t)^4} + \frac{(2\theta+5)(\theta+1+e^t)}{(\theta+1-e^t)^3} + \frac{(\theta^2 + 4\theta + 5)}{(\theta+1-e^t)^2} \right]$$

### 4. Estimation

#### 4.1. Estimation by Method of Moments

Since SBPAD has only one parameter to be estimated, equating the population mean to the corresponding sample mean, method of moments estimate (MOME)  $\tilde{\theta}$  of the parameter  $\theta$  of SBPAD is the solution of the following cubic equation in  $\theta$

$$(1 - \bar{x})\theta^3 + 2(3 - 2\bar{x})\theta^2 + 6(3 - \bar{x})\theta + 24 = 0,$$

where  $\bar{x}$  is the sample mean. This equation can be easily solved using Newton-Raphson method for MOME estimate  $\tilde{\theta}$  of  $\theta$  of SBPAD.

#### 4.2. Estimation by Maximum Likelihood Method

Let  $(x_1, x_2, \dots, x_n)$  be a random sample of size  $n$  from the SBPAD (1.4). Suppose  $f_x$  be the observed frequency in the sample corresponding to  $X = x$  ( $x = 1, 2, 3, \dots, k$ ) such that  $\sum_{x=1}^k f_x = n$ ,  $k$  being the largest observed value having non-zero frequency. Then, the log likelihood function,  $\log L$ , of the SBPAD (1.4) can be expressed as

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

The first derivative of the log likelihood function can be given by

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

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

The maximum likelihood estimate (MLE)  $\hat{\theta}$  of the parameter  $\theta$  of SBPAD (1.4) is thus the solution of the following log likelihood equation

$$\frac{4n}{\theta} - \frac{2n(\theta+2)}{\theta^2+4\theta+6} - \frac{n(\bar{x}+3)}{\theta+1} + \sum_{x=1}^k \frac{2(x+\theta+2)f_x}{x^2+(2\theta+5)x+(\theta^2+4\theta+5)} = 0.$$

This non-linear log likelihood equation can be solved by any iterative methods. Here we used Newton-Raphson method to solve above equation. Note that the initial value for Newton-Raphson method is the value given by MOME of the parameter  $\theta$ .

## 5. Goodness of Fit

We know that when organism occurs in groups and the group size influences the probability of detection, size-biased distributions are the appropriate choice to model the datasets. In this section, three examples of real datasets, two from the size distribution of freely-forming small group at various public places, available in James (1953) and Coleman and James (1961) and one from the number of pairs of running shoes owned by 60 members of an athletic club, available in Simonoff (2003, p. 100), have been taken for testing the goodness of fit of SBPAD and compared with SBPD and SBPLD. The estimation of parameter of all distributions is based on MLE. The criterion for the selection of best distribution is based on the values of chi-square ( $\chi^2$ ),  $-2\log L$  and AIC (Akaike Information Criteria). The AIC is calculated using  $AIC = -2\log L + 2k$ , where  $k$  the number of parameters involved in the distribution. The best distribution is the distribution whose values of chi-square,  $-2\log L$  and AIC is the lowest. Clearly SBPAD gives a much better fit as compared with SBPD and SBPLD, and, therefore, it should be considered an important distribution over SBPD and SBPLD.

**Table 2.** Play Groups-Eugene, Spring, Public Playground A

Group Size	Observed frequency	Expected frequency		
		SBPD	SBPLD	SBPAD
1	306	292.2	309.4	308.6
2	132	155.2	131.2	132.2
3	47	41.2	41.1	41.3
4	10	7.3	11.3	11.2
5	2	1.1	4.0	3.7
Total	497	497.0	497.0	497.0
ML Estimate		$\hat{\theta} = 0.5312$	$\hat{\theta} = 4.3548$	$\hat{\theta} = 4.95006$
$\chi^2$		6.479	1.494	1.373
d.f.		2	2	2
p-value		0.039	0.4737	0.5033
$-2\log L$		2142.03	971.86	971.59
AIC		2144.03	973.86	973.59

**Table 3.** Play Groups-Eugene, Spring, Public Playground D

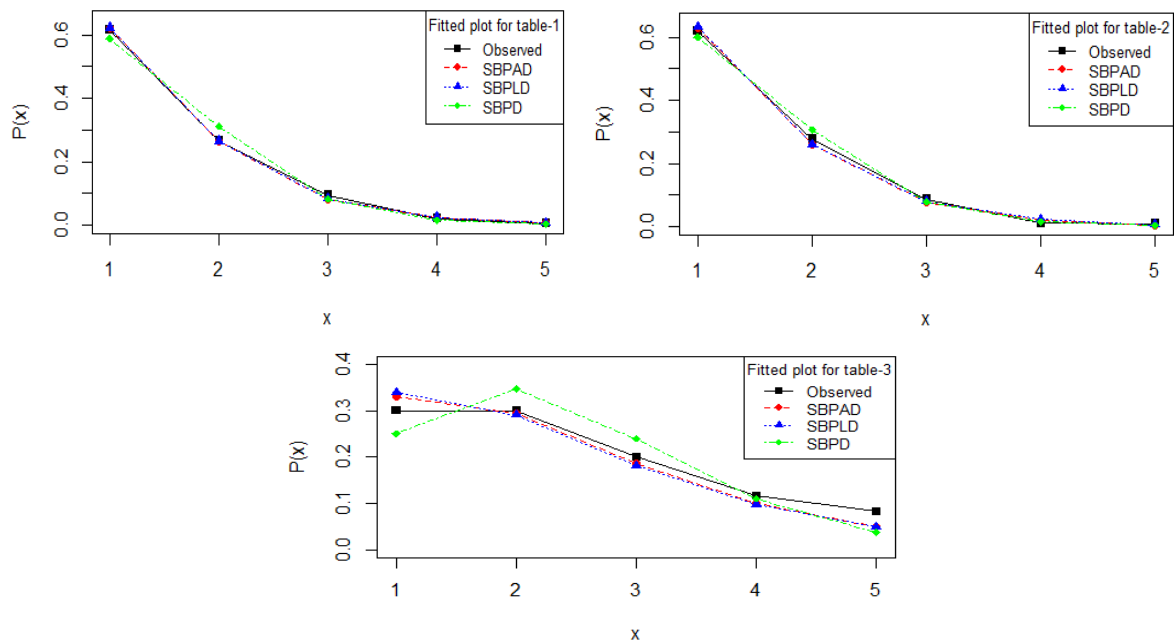
Group Size	Observed frequency	Expected frequency		
		SBPD	SBPLD	SBPAD
1	316	306.3	323.0	322.2
2	141	156.1	132.6	133.4
3	44	39.8	40.2	40.4
4	5	6.8	10.7	10.6
5	4	1.0	3.5	3.4
Total	510	510.0	510.0	510.0

Group Size	Observed frequency	Expected frequency		
		SBPD	SBPLD	SBPAD
ML Estimate		$\hat{\theta} = 0.50980$	$\hat{\theta} = 4.52061$	$\hat{\theta} = 5.12468$
$\chi^2$		2.395	2.947	2.658
d.f.		2	2	2
p-value		0.3019	0.2291	0.2647
$-2\log L$		2376.75	972.78	972.47
AIC		2378.75	974.78	974.47

**Table 4.** The number of pairs of running shoes owned by 60 members of an athletic club, available in Simonoff (2003, p. 100)

Number of pairs of running shoes	Observed frequency	Expected Frequency		
		SBPD	SBPLD	SBPAD
1	18	15.0	20.3	19.8
2	18	20.8	17.4	17.7
3	12	14.4	10.9	11.2
4	7	6.6	5.9	6.0
5	5	3.2	5.5	5.3
Total	60	60.0	60.0	60.0
ML Estimate		$\hat{\theta} = 1.383333$	$\hat{\theta} = 1.818978$	$\hat{\theta} = 2.20145$
$\chi^2$		1.87	0.64	0.39
d.f.		2	3	3
P-value		0.3926	0.8872	0.9423
$-2\log L$		556.13	187.08	186.54
AIC		558.13	189.08	188.54

The probability plots of SBPAD, SBPLD and SBPD for the considered datasets are shown in figure 3.



**Figure 3.** Fitted probability plots for distributions for considered datasets



## 6. Concluding Remarks

In this paper, a size-biased Poisson –Aradhana distribution (SBPAD) has been proposed by size-biasing the discrete Poisson- Aradhana distribution (PAD) suggested by Shanker (2017), a Poisson mixture of Aradhana distribution introduced by Shanker (2016). Its statistical constants including coefficients of variation, skewness, kurtosis, and index of dispersion have been studied. Both the method of moments and the maximum likelihood estimation has been discussed Applications of SBPAD have been explained with three examples of real datasets and the goodness of fit shows that SBPAD gives better fit over SBPD and SBPLD.

## ACKNOWLEDGEMENTS

Authors are grateful to the editor in chief of the journal and the anonymous reviewer for fruitful comments 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] Coleman, J.S. and James, J. (1961): The equilibrium size distribution of freely forming groups, *Sociometry*, 24, 13 – 25.
- [4] Correa, J.A. and Wolfson, D.B. (2007): Length-bias: some Characterizations and applications, *Journal of Statistical Computation and Simulation*, 64, 209 – 219.
- [5] Drummer, T.D. and MacDonald, L.L. (1987): Size biased in line transects sampling, *Biometrics*, 43, 13 – 21.
- [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, 201 – 212.
- [12] James, J. (1953): The distribution of free-forming small group size, *American Sociological Review*, 18, 569 – 570.
- [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] MacDonald, 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 Jubilee, Statistical Publishing society, 478 – 503.
- [17] Patil, G.P. (1991): Encountered data, Statistical ecology, environment statistics, and Weighted methods, *Environmetrics*, 2, 377 – 423.
- [18] Patil, G.P.(1996): Statistical ecology, environment statistics, and risk assessment, In *Advances in Biometry*, P. Armitage & H.A. David, eds, Wiley, New York, 213 – 240.
- [19] Patil, G.P. (1997): Weighted distributions, In *Encyclopedia of Biostatistics*, Vol. 6, P. Armitage & T. Colton, eds, Wiley, Chichester, 4735 – 4738.
- [20] Patil, G.P. and Ord, J.K. (1976): On size-biased sampling and related form-invariant weighted distributions, *Sankhya Ser. B*, 38, 48 – 61.

- [21] 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.
- [22] 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.
- [23] 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.
- [24] Sankaran, M. (1970): The discrete Poisson-Lindley distribution, *Biometrics*, 26, 145 – 149.
- [25] Scheaffer, R.L. (1972): Size-biased sampling, *Technometrics*, 14, 635 – 644.
- [26] Shanker, R. (2016): Aradhana Distribution and Its Applications, *International Journal of Statistics and Applications*, 6(1), 23 – 34.
- [27] Shanker, R. (2017): The discrete Poisson-Aradhana distribution, *Turkiye Klinikleri Journal of Biostatistics*, 9(1), 12 – 22.
- [28] 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.
- [29] Simonoff, J.S. (2003): Analyzing Categorical data, Springer, New York.
- [30] Singh, S.K. and Maddala, G.S. (1976): A function for the size distribution of incomes, *Econometrica*, 44, 963 – 970.
- [31] Van Deusen, P.C. (1986): Fitting assumed distributions to horizontal point sample diameters, *For. Sci.*, 32, 146 – 148.