Available online at http://scik.org
J. Math. Comput. Sci. 11 (2021), No. 6, 6811-6828
https://doi.org/10.28919/jmcs/5704
ISSN: 1927-5307

A GENERALIZATION OF SIZE-BIASED POISSON-SUJATHA DISTRIBUTION AND ITS APPLICATIONS

RAMA SHANKER[1], KAMLESH KUMAR SHUKLA[2], MANOJ TIWARI[3,4,*], TEKIE ASEHUEN LEONIDA[5]

[1]Department of Statistics, Assam University Silchar, India

[2]Department of Community Medicine, Noida International Institute of Medical Sciences, Noida International University, G. B. Nagar, India

[3]Department of Statistics, Sultan Qaboos University, Muscat, Oman

[4]Department of Statistics, Panjab University, Chandigarh – 160 014, India

[5]Department of Applied Mathematics, University of Twente, The Netherlands

Abstract: In this paper, a generalization of size-biased Poisson-Sujatha distribution (AGSBPSD) which includes both the size-biased Poisson-Lindley distribution (SBPLD) and the size-biased Poisson-Sujatha distribution (SBPSD) as particular cases has been proposed and studied. Its moments based measures including coefficients of variation, skewness, kurtosis, and index of dispersion have been derived and their shapes have been discussed with varying values of the parameters. The estimation of its parameters has been discussed using maximum likelihood estimation. Some applications of the proposed distribution have been explained using two count datasets.

Keywords: size-biased distribution; size-biased Poisson-Lindley distribution; size-biased Poisson-Sujatha distribution; compounding; skewness; kurtosis; maximum likelihood estimation; applications.

2010 AMS Subject Classification: 62E05, 62E99

---

*Corresponding author
E-mail addresses: m.tiwari@squ.edu.om; mtiwari@pu.ac.in
Received March 16, 2021

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 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. It was Fisher [1] who firstly introduced these distributions to model ascertainment biases which were later reformulated by Rao [2]. Size-biased distributions have applications in environmental science, econometrics, social science, biomedical science, human demography, ecology, geology, forestry etc. A summary of weighted distributions and size-biased distributions and their applications are available in Patil and Rao ([3],[4]).

Suppose 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 defined by its probability mass function (pmf)

$$P(x;\theta)=\frac{x^{\alpha}\cdot P_0(x;\theta)}{\mu'_{\alpha}} \tag{1.1}$$

where $\mu'_{\alpha}=E\left(X^{\alpha}\right)=\sum_{x=0}^{\infty}x^{\alpha}P_0(x;\theta)$. If $\alpha=1$, (1.1) is known as simple size-biased distribution and is applicable for size-biased sampling whereas for $\alpha=2$, (1.1) is known as area-biased distribution and is applicable for area-biased sampling.

The modeling and statistical analysis of count data excluding zero counts using size-biasing of Poisson mixture of continuous distributions are crucial in almost every fields of knowledge including biological science, insurance, medical science, demography and finance, some amongst others. During a short period of time, a number of Poisson mixtures of continuous distributions and their size biased versions have been introduced in Statistics literature. For example, Greenwood and Yule[5] proposed negative binomial distribution (NBD) as a Poisson mixture of gamma distribution and size-biased version of NBD were studied by Mir and Ahmad [6] whereas Sankaran[7] proposed the discrete Poisson-Lindley distribution (PLD) as a Poisson mixture of Lindley[8] distribution and size-biased version of PLD were studied by Ghitany and Al-Mutairi[9]. During recent decades several one parameter and two-parameter lifetime distributions and their Poisson mixture have been introduced in statistics literature

The probability density function (pdf) of Lindley distribution, Sujatha distribution, generalized Lindley distribution (GLD) and a generalization of Sujatha distribution (AGSD) along with their introducers and years are presented in table 1. The Poisson mixture of these lifetime distributions namely, Poisson-Lindley distribution (PLD), Poisson-Sujatha distribution (PSD), generalized Poisson-Lindley distribution (GPLD) and a generalization of Poisson-Sujatha distribution (AGPSD) their mean, introducers and years are given in table 2. The size biased version of these Poisson mixture of lifetime distributions namely, size-biased Poisson-Lindley distribution (SBPLD), size-biased Poisson-Sujatha distribution (SBPSD), and generalized size-biased Poisson-Lindley distribution (GSBPLD)

Table 1: Probability density function of distributions

Name of distributions	Pdf	Introducers(years)
Lindley	$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}\quad ;x>0,\ \theta>0$	Lindley [8]
Sujatha	$f(x;\theta)=\frac{\theta^3}{\theta^2+\theta+2}\left(1+x+x^2\right)e^{-\theta x}\quad ;x>0\ ,\theta>0$	Shanker [10]
GLD	$f(x;\theta,\alpha)=\frac{\theta^{\alpha+1}}{\theta+1}\frac{x^{\alpha-1}}{\Gamma(\alpha+1)}(\alpha+x)e^{-\theta x}\quad ;x>0,\ \theta>0,\alpha>0$	Zakerzadeh and Dolati [11]
AGSD	$f(x;\theta,\alpha)=\frac{\theta^3}{\theta^2+\theta+2\alpha}\left(1+x+\alpha x^2\right)e^{-\theta x}\quad ;x>0,\ \theta>0,\alpha>0$	Shanker et al [12]

Table 2: Pmf of discrete distributions and its mean and introducers (years)

Name of the distribution	Pmf and its mean	Mixture of distributions	Introducers(years)
PLD	$P(x;\theta)=\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}\ ;x=0,1,2,3,...,\ \theta>0$ $\mu_1'=\frac{\theta+2}{\theta(\theta+1)}$	Mixture of Poisson and Lindley	Sankaran[7]
PSD	$P_0(x;\theta)=\frac{\theta^3}{\theta^2+\theta+2}\frac{x^2+(\theta+4)x+(\theta^2+3\theta+4)}{(\theta+1)^{x+3}}$ $;x=0,1,2,3,...,\ \theta>0$ $\mu_1'=\frac{\theta^2+2\theta+6}{\theta(\theta^2+\theta+2)}$	Mixture of Poisson and Sujatha	Shanker[13]
GPLD	$P(x;\theta,\alpha)=\frac{\theta^3\{\alpha x^2+(\theta+3\alpha+1)x+(\theta^2+3\theta+2\alpha+2)\}}{(\theta^2+\theta+2\alpha)(\theta+1)^{x+3}}\ ;x=0,1,2,3,...,\ \theta>0,\alpha>0$ $\mu_1'=\frac{\alpha(\theta+1)+1}{\theta(\theta+1)}$	Mixture of Poisson and GLD	Mahmoudi and Zakerzadeh[14]
AGPSD	$P(x;\theta,\alpha)=\frac{\theta^3\{\alpha x^2+(\theta+3\alpha+1)x+(\theta^2+3\theta+2\alpha+2)\}}{(\theta^2+\theta+2\alpha)(\theta+1)^{x+3}}\ ;x=1,2,3,...,\ \theta>0,\alpha>0$ $\mu_1'=\frac{\theta^2+2\theta+6\alpha}{\theta(\theta^2+\theta+2\alpha)}$	Mixture of Poisson and AGSD	Shanker and Shukla [15]

Table 3: Pmf of size-biased distributions along with introducers (years)

Name of distributions	Pmf	Introducers(years)
SBPLD	$P(x;\theta)=\frac{\theta^3}{\theta+2}\frac{x(x+\theta+2)}{(\theta+1)^{x+2}}\ ;x=1,2,3,...,\ \theta>0$	Ghitany and Al-Mutairi [9]
SBPSD	$P(x;\theta)=\frac{\theta^4\left\{x^3+(\theta+4)x^2+(\theta^2+3\theta+4)x\right\}}{(\theta^2+2\theta+6)(\theta+1)^{x+3}};$ $x=1,2,3,..,\theta>0$	Shanker and Hagos [16]
GSBPLD	$P(x;\theta,\alpha)=\frac{\Gamma(x+\alpha)}{\Gamma(x)\Gamma(\alpha+1)}\frac{\theta^{\alpha+2}}{(\alpha\theta+\alpha+1)}\frac{x+\alpha\theta+2\alpha}{(\theta+1)^{x+\alpha+1}}$ $;\theta>0,\alpha>0,x=1,2,3,...$	Shanker and Shukla[17]

The main motivation of this paper is to introduce a generalization of size-biased Poisson-Poisson Sujatha distribution (AGSBPSD) which includes size-biased Poisson-Sujatha distribution (SBPSD) and size-biased Poisson-Lindley distribution (SBPLD) as particular cases and to see its applications to model count data excluding zero counts. Moments and moments based measures have been derived. The coefficients of variation, skewness, kurtosis, index of dispersion have been explained graphically. Maximum likelihood estimation has been discussed for estimating the parameters of the distribution. Goodness of fit of the proposed distribution have been explained with several count datasets and compared with other size-biased distributions.

.

2. A GENERALIZATION OF SIZE-BIASED POISSON-SUJATHA DISTRIBUTION

Using (1.1) and the mean and the pmf of the AGPSD, the pmf of a generalization of size-biased Poisson-Sujatha distribution (AGSBPSD) can be defined as

$$P(x;\theta,\alpha)=\frac{\theta^4\left\{\alpha x^3+(\theta+3\alpha+1)x^2+(\theta^2+3\theta+2\alpha+2)x\right\}}{(\theta^2+2\theta+6\alpha)(\theta+1)^{x+3}}\quad ;x=1,2,3,...,\ \theta>0,\alpha>0 \qquad (2.1)$$

It can be easily verified that AGSBPSD reduces to SBPSD and SBPLD at $\alpha = 1$ and $\alpha = 0$, respectively. Since some of the characteristics of AGSBPSD including coefficients of variation, skewness, kurtosis and index of dispersion depend on central moments, it has been observed that it difficult and very complicated to obtain the moments of AGSBPSD from its pmf directly. Therefore, to obtain the moments of AGSBPSD easily, it has been observed that AGSBPSD can also be obtained as mixture of size-biased Poisson distribution (SBPD) with size-biased version of the generalization of Sujatha distribution (SBGSD).

Suppose X follows SBPD with parameter λ having pmf

$$P(X \mid \lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad x = 1, 2, 3, \ldots, \lambda > 0,$$

and the parameter λ follows SBGSD with parameters θ and α having pdf

$$h(\lambda \mid \theta, \alpha) = \frac{\theta^4}{\theta^2 + 2\theta + 6\alpha}\left(\lambda + \lambda^2 + \alpha\lambda^3\right)e^{-\theta\lambda} \quad ; \lambda > 0,\ \theta > 0, \alpha > 0.$$

Thus

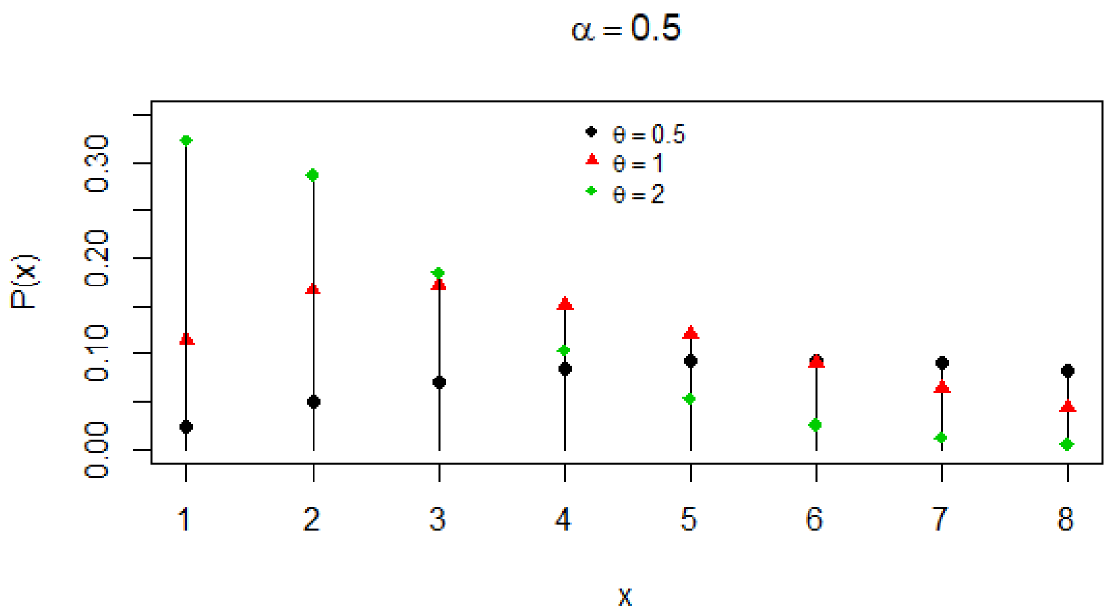
$$P(X = x) = \int_0^\infty P(X \mid \lambda)\, h(\lambda \mid \theta, \alpha)\, d\lambda$$

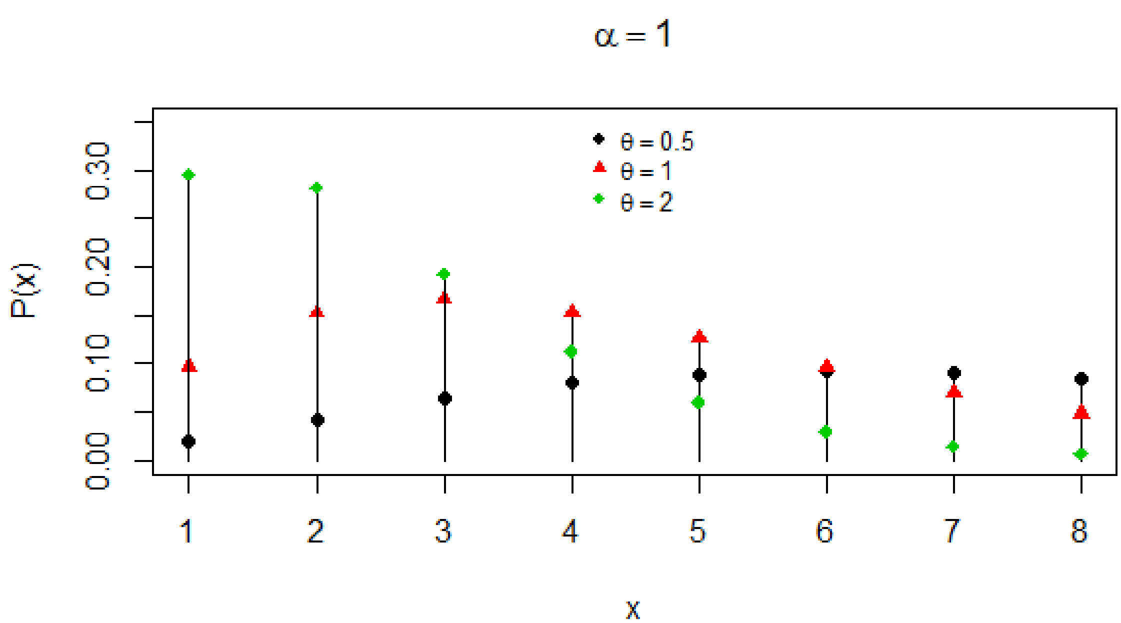
$$= \int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\, \frac{\theta^4}{\theta^2 + 2\theta + 6\alpha}\left(\lambda + \lambda^2 + \alpha\lambda^3\right)e^{-\theta\lambda}\, d\lambda \tag{2.2}$$

$$= \frac{\theta^4}{\left(\theta^2 + 2\theta + 6\alpha\right)(x-1)!}\int_0^\infty e^{-(\theta+1)\lambda}\left(\lambda^x + \lambda^{x+1} + \alpha\lambda^{x+2}\right)d\lambda$$

$$= \frac{\theta^4}{\left(\theta^2 + 2\theta + 6\alpha\right)(x-1)!}\left[\frac{\Gamma(x+1)}{(\theta+1)^{x+1}} + \frac{\Gamma(x+2)}{(\theta+1)^{x+2}} + \frac{\alpha\,\Gamma(x+3)}{(\theta+1)^{x+3}}\right]$$

$$= \frac{\theta^4}{\left(\theta^2 + 2\theta + 6\alpha\right)}\,\frac{\alpha x^3 + (\theta + 3\alpha + 1)x^2 + \left(\theta^2 + 3\theta + 2\alpha + 2\right)x}{(\theta+1)^{x+3}} \quad ; x = 1, 2, 3, \ldots$$


α = 0.5
θ = 0.5
θ = 1
θ = 2
P(x)
0.00
0.10
0.20
0.30
1
2
3
4
5
6
7
8
x


α = 1
θ = 0.5
θ = 1
θ = 2
P(x)
0.00
0.10
0.20
0.30
1
2
3
4
5
6
7
8
x

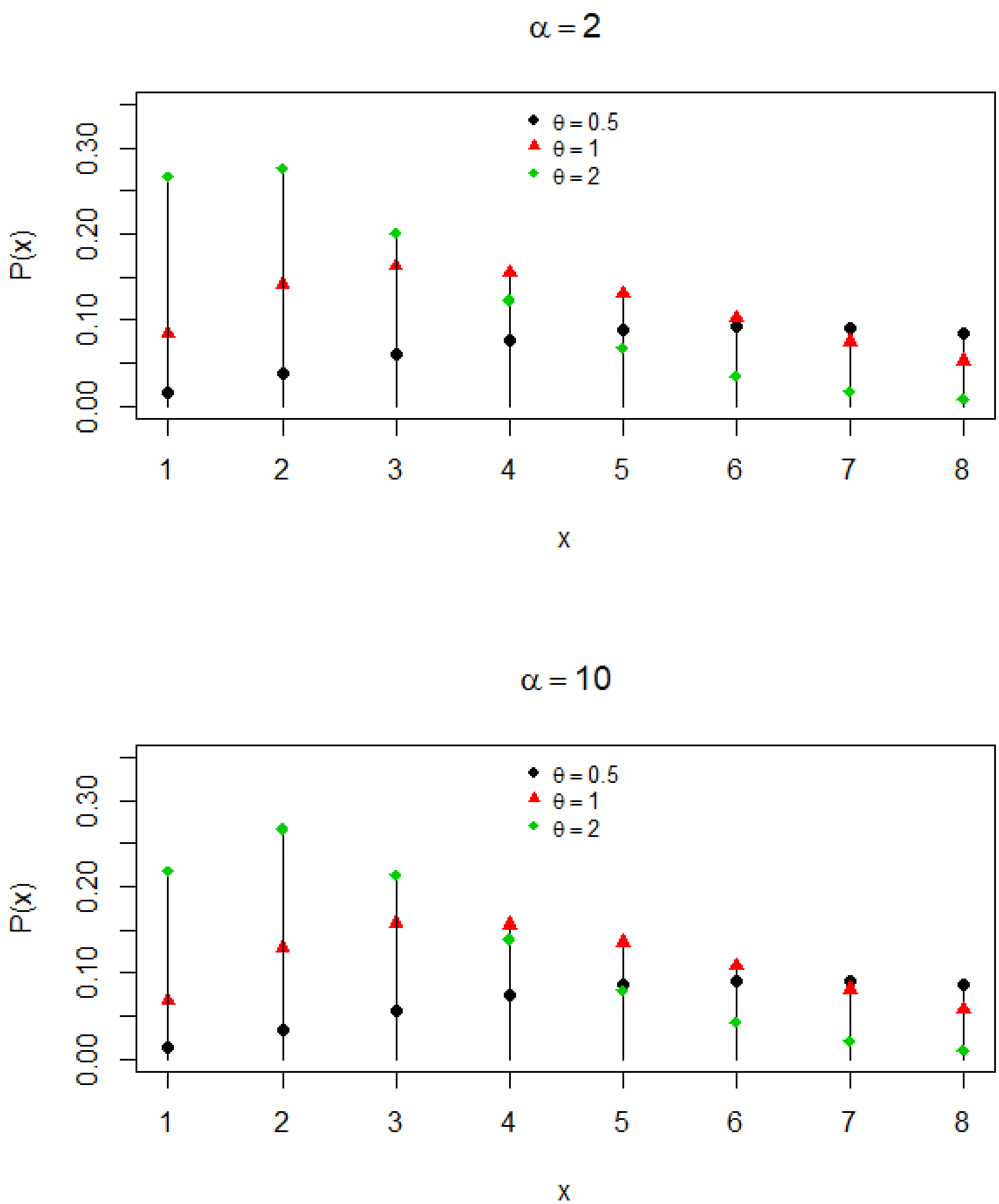


Fig. 1: Probability mass function of AGSBPSD for varying values of θ and α

3. MOMENTS

Using (2.2), the r th factorial moment about origin $\mu_{(r)}'$ of AGSBPSD can be obtained as

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

We have

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

$$= \int_0^\infty \left[\lambda^{r-1}\left\{\sum_{x=1}^\infty x\frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right\}\right] \frac{\theta^4}{\theta^2+2\theta+6\alpha}\left(\lambda+\lambda^2+\alpha\lambda^3\right)e^{-\theta\lambda}\,d\lambda .$$

Taking $y = x - r$, we have

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

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

Using gamma integral and some tedious algebraic simplification, the r th factorial moment about origin $\mu_{(r)}'$ of AGSBPSD can be expressed as

$$\mu_{(r)}' = \frac{r!\left\{r\theta^3+(r+1)^2\theta^2+(r+1)(r+2)(r\alpha+1)\theta+(r+1)(r+2)(r+3)\alpha\right\}}{\theta^r\left(\theta^2+2\theta+6\alpha\right)} \;;\; r = 1,2,3,... \qquad (3.1)$$

Taking $r = 1,2,3$ and 4 in (3.1), the first four factorial moments about origin of AGSBPSD can be obtained as

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

$$\mu_{(2)}' = \frac{2\left\{2\theta^3+9\theta^2+12(2\alpha+1)\theta+60\alpha\right\}}{\theta^2\left(\theta^2+2\theta+6\alpha\right)}$$

$$\mu_{(3)}' = \frac{6\left\{3\theta^3+16\theta^2+20(3\alpha+1)\theta+120\alpha\right\}}{\theta^3\left(\theta^2+2\theta+6\alpha\right)}$$

$$\mu_{(4)}' = \frac{24\left\{4\theta^3+25\theta^2+30(4\alpha+1)\theta+210\alpha\right\}}{\theta^4\left(\theta^2+2\theta+6\alpha\right)} .$$

Using relationship between moments about origin and factorial moments about origin, the moments about origin of AGSBPSD are obtained as

$$\mu_1' = \frac{\theta^3 + 4\theta^2 + 6(\alpha+1)\theta + 24\alpha}{\theta(\theta^2 + 2\theta + 6\alpha)}$$

$$\mu_2' = \frac{\theta^4 + 8\theta^3 + (6\alpha+24)\theta^2 + (72\alpha+24)\theta + 120\alpha}{\theta^2(\theta^2 + 2\theta + 6\alpha)}$$

$$\mu_3' = \frac{\theta^5 + 16\theta^4 + (6\alpha+78)\theta^3 + (168\alpha+168)\theta^2 + (720\alpha+120)\theta + 720\alpha}{\theta^3(\theta^2 + 2\theta + 6\alpha)}$$

$$\mu_4' = \frac{\theta^6 + 32\theta^5 + (6\alpha+240)\theta^4 + (360\alpha+840)\theta^3 + (3000\alpha+1320)\theta^2 + (7200\alpha+720)\theta + 5040\alpha}{\theta^4(\theta^2 + 2\theta + 6\alpha)}.$$

Again using relationship between moments about the mean and the moments about origin, the moments about the mean of AGSBPSD are obtained as

$$\mu_2 = \frac{2\left\{\theta^5 + 6\theta^4 + (18\alpha+12)\theta^3 + (72\alpha+6)\theta^2 + \left(72\alpha^2 + 48\alpha\right)\theta + 72\alpha^2\right\}}{\theta^2(\theta^2 + 2\theta + 6\alpha)^2}$$

$$\mu_3 = \frac{2\left\{\begin{array}{l}\theta^8 + 10\theta^7 + (24\alpha+42)\theta^6 + (216\alpha+84)\theta^5 + \left(180\alpha^2 + 660\alpha + 72\right)\theta^4 \\ + \left(1152\alpha^2 + 720\alpha + 24\right)\theta^3 + \left(432\alpha^3 + 1728\alpha^2 + 288\alpha\right)\theta^2 + \left(1296\alpha^3 + 864\alpha^2\right)\theta \\ + 864\alpha^3\end{array}\right\}}{\theta^3(\theta^2 + 2\theta + 6\alpha)^3}$$

$$\mu_4 = \frac{2\left\{\begin{array}{l}\theta^{11} + 22\theta^{10} + (30\alpha+184)\theta^9 + (708\alpha+780)\theta^8 + \left(324\alpha^2 + 5160\alpha + 1800\right)\theta^7 \\ + \left(7488\alpha^2 + 16752\alpha + 2256\right)\theta^6 + \left(1512\alpha^3 + 42192\alpha^2 + 26544\alpha + 1440\right)\theta^5 \\ + \left(32400\alpha^3 + 93024\alpha^2 + 20160\alpha + 360\right)\theta^4 + \left(2592\alpha^4 + 122688\alpha^3 + 88128\alpha^2 + 5760\alpha\right)\theta^3 \\ + \left(49248\alpha^4 + 155520\alpha^3 + 29376\alpha^2\right)\theta^2 + \left(93312\alpha^4 + 62208\alpha^3\right)\theta + 46656\alpha^4\end{array}\right\}}{\theta^4(\theta^2 + 2\theta + 6\alpha)^4}$$

The statistical constants based on central moments including coefficient of variation (C.V), coefficient of Skewness $\left(\sqrt{\beta_1}\right)$, coefficient of Kurtosis (β_2) and Index of dispersion (γ) of AGSBPSD are obtained as

$$C.V. = \frac{\sigma}{\mu_1'} = \frac{\sqrt{2\left\{\theta^5 + 6\theta^4 + (18\alpha+12)\theta^3 + (72\alpha+6)\theta^2 + \left(72\alpha^2 + 48\alpha\right)\theta + 72\alpha^2\right\}}}{\theta^3 + 4\theta^2 + 6(\alpha+1)\theta + 24\alpha}$$

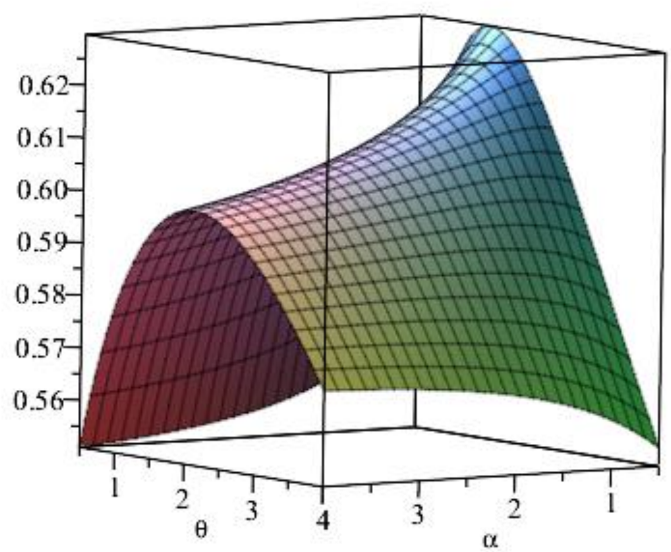
$$\sqrt{\beta_1}=\frac{\mu_3}{(\mu_2)^{3/2}}=\frac{\left\{\begin{array}{l}\theta^8+10\theta^7+(24\alpha+42)\theta^6+(216\alpha+84)\theta^5+\left(180\alpha^2+660\alpha+72\right)\theta^4\\+\left(1152\alpha^2+720\alpha+24\right)\theta^3+\left(432\alpha^3+1728\alpha^2+288\alpha\right)\theta^2+\left(1296\alpha^3+864\alpha^2\right)\theta\\+864\alpha^3\end{array}\right\}}{\sqrt{2}\left\{\theta^5+6\theta^4+(18\alpha+12)\theta^3+(72\alpha+6)\theta^2+\left(72\alpha^2+48\alpha\right)\theta+72\alpha^2\right\}^{3/2}}$$

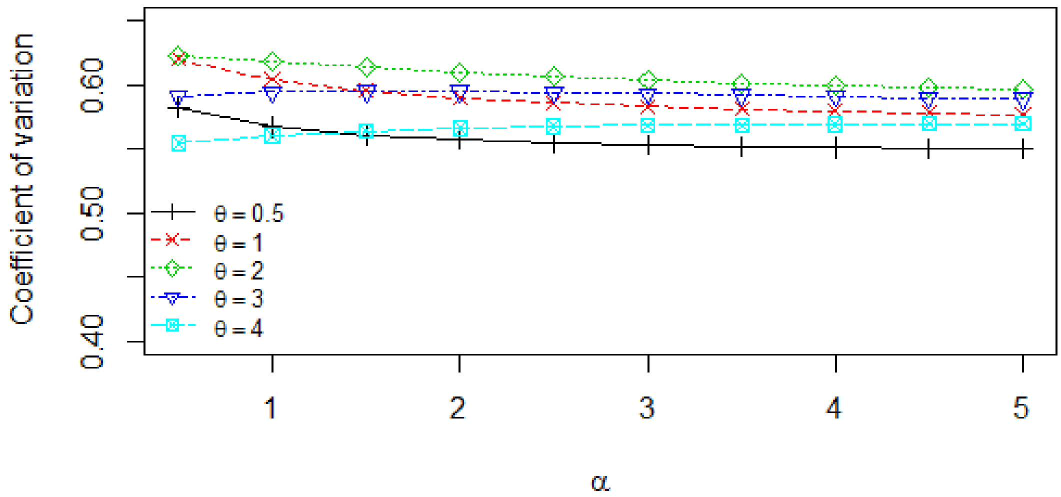
$$\beta_2=\frac{\mu_4}{\mu_2^{\,2}}=\frac{\left\{\begin{array}{l}\theta^{11}+22\theta^{10}+(30\alpha+184)\theta^9+(708\alpha+780)\theta^8+\left(324\alpha^2+5160\alpha+1800\right)\theta^7\\+\left(7488\alpha^2+16752\alpha+2256\right)\theta^6+\left(1512\alpha^3+42192\alpha^2+26544\alpha+1440\right)\theta^5\\+\left(32400\alpha^3+93024\alpha^2+20160\alpha+360\right)\theta^4+\left(2592\alpha^4+122688\alpha^3+88128\alpha^2+5760\alpha\right)\theta^3\\+\left(49248\alpha^4+155520\alpha^3+29376\alpha^2\right)\theta^2+\left(93312\alpha^4+62208\alpha^3\right)\theta+46656\alpha^4\end{array}\right\}}{2\left\{\theta^5+6\theta^4+(18\alpha+12)\theta^3+(72\alpha+6)\theta^2+\left(72\alpha^2+48\alpha\right)\theta+72\alpha^2\right\}^2}$$

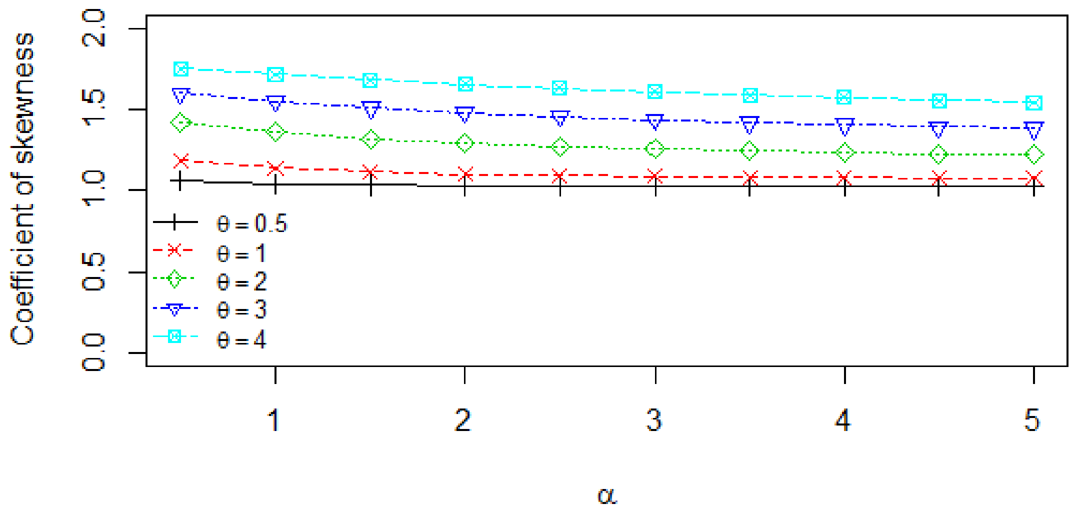
$$\gamma=\frac{\sigma^2}{\mu}=\frac{2\left\{\theta^5+6\theta^4+(18\alpha+12)\theta^3+(72\alpha+6)\theta^2+\left(72\alpha^2+48\alpha\right)\theta+72\alpha^2\right\}}{\theta\left(\theta^2+2\theta+6\alpha\right)\left\{\theta^3+4\theta^2+6(\alpha+1)\theta+24\alpha\right\}}$$

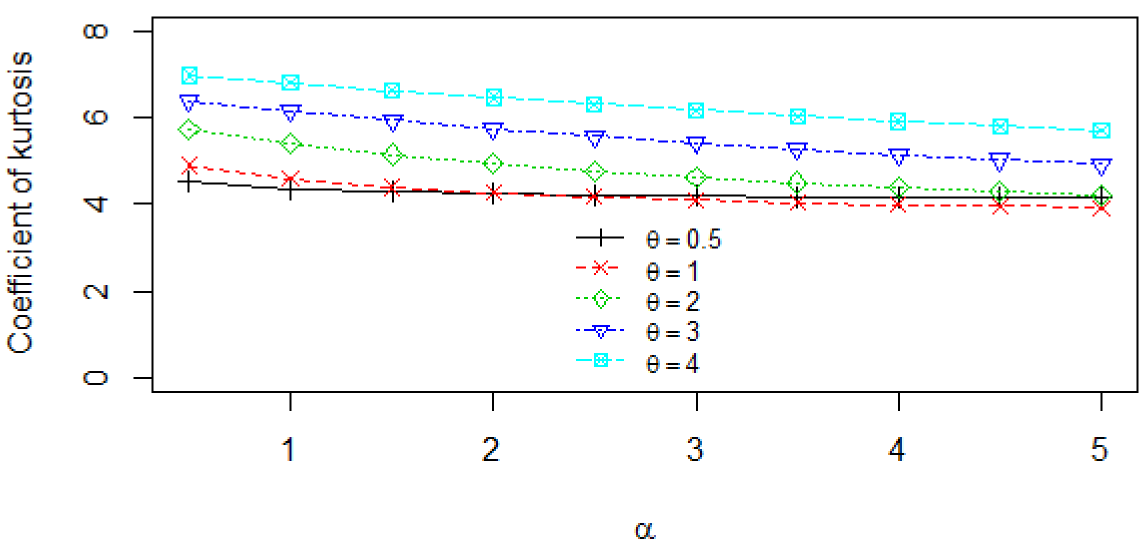
The shapes of coefficient of variation, skewness, kurtosis and index of dispersion for varying values of parameters are shown in figure 2.

Coefficeint of variation




Coefficient of variation
0.40
0.50
0.60
1
2
3
4
5
α
θ = 0.5
θ = 1
θ = 2
θ = 3
θ = 4


Coefficient of skewness
0.0
0.5
1.0
1.5
2.0
1
2
3
4
5
α
θ = 0.5
θ = 1
θ = 2
θ = 3
θ = 4


Coefficient of kurtosis
0
2
4
6
8
1
2
3
4
5
α
θ = 0.5
θ = 1
θ = 2
θ = 3
θ = 4

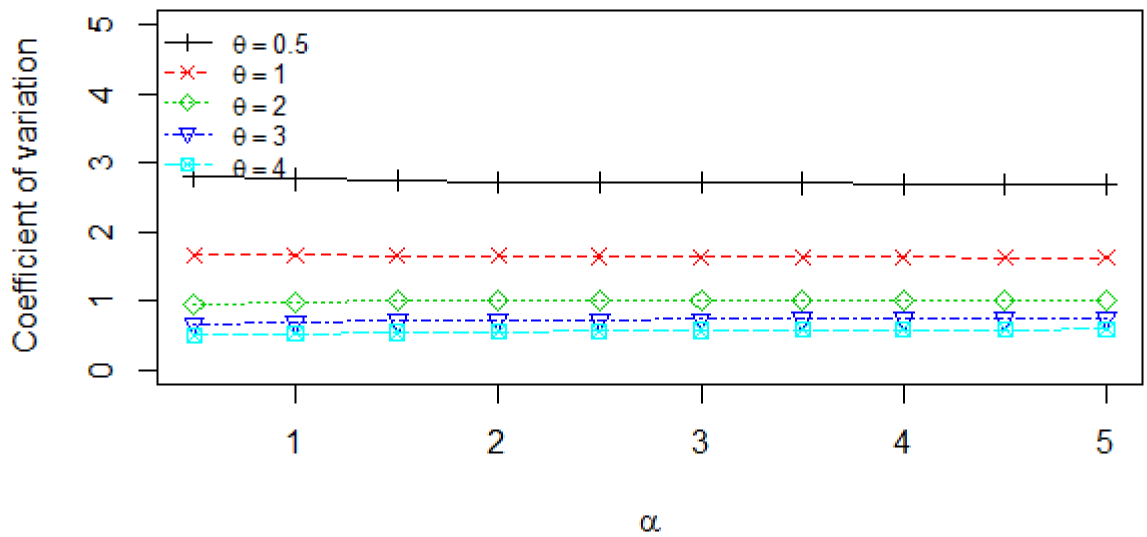


Fig. 2: Shapes of C.V, Skewness, Kurtosis and Index of dispersion of AGSBPSD for varying values of θ and α

4. Estimation of Parameters Using Maximum Likelihood

Let $(x_1, x_2, ..., x_n)$ be the observed values of a random sample $(X_1, X_2, ..., X_n)$ from AGSBPSD (θ, α) 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 AGSBPSD is given by

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

The log likelihood function is given by

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

The maximum likelihood estimates $\left(\hat{\theta},\hat{\alpha}\right)$ of parameters $\left(\theta,\alpha\right)$ of AGSBPSD (2.1) is the solutions of the following log likelihood equations

$$\frac{\partial \log L}{\partial \theta}=\frac{4n}{\theta}-\frac{2n\left(\theta+1\right)}{\theta^2+2\theta+6\alpha}-\frac{n\bar{x}}{\theta+1}+\sum_{x=1}^{k}\frac{\left(x^2+2\theta+3\right)f_x}{\alpha x^3+\left(\theta+3\alpha+1\right)x^2+\left(\theta^2+3\theta+2\alpha+2\right)x}=0$$

$$\frac{\partial \log L}{\partial \alpha}=\frac{-6n}{\theta^2+2\theta+6\alpha}+\sum_{x=1}^{k}\frac{\left(x^3+3x^2+2x\right)f_x}{\alpha x^3+\left(\theta+3\alpha+1\right)x^2+\left(\theta^2+3\theta+2\alpha+2\right)x}=0.$$

These two log-likelihood equations cannot be expressed in closed forms and seems difficult to solve directly. The MLE's $\left(\hat{\theta},\hat{\alpha}\right)$ of parameters $\left(\theta,\alpha\right)$ of AGSBPSD (2.1) can be computed directly by solving the log likelihood equations using Newton-Raphson iteration available in R-software till sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained. The initial values for parameters θ and α are generally taken as $\theta=0.5$ and $\alpha=1$.

5. GOODNESS OF FIT

In this section the goodness of fit of SBPD, SBPLD SBPSD, GSBPLD and AGSBPSD has been presented for three count data- sets. The goodness of fit of these distributions are based on maximum likelihood estimates of the parameter. The first dataset in table 4 is the number of counts of pairs of running shoes owned by 60 members of an athletic club, reported by Simonoff [18] and the second dataset in table 5 is the distribution of free-forming small group size available in James [19] and Coleman and James [20]. It is clear from these two tables that AGSBPSD gives best fit in table 4 while in table 5 GSBPLD gives best fit and these two distributions are competing well with the considered distributions and hence it can be considered as an important size-biased distribution.

Table 4: 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	GSBPLD	AGSBPSD
1	18	15.0	20.3	20.0	16.7	18.5
2	18	20.8	17.4	17.5	19.7	18.5
3	12	14.4	10.9	11.1	13.1	11.9
4	7	6.6	5.9	6.0	6.4	6.2
5	5	3.2	5.5	5.4	4.1	4.9
Total	60	60.0	60.0	60.0	60.0	60.0
ML Estimate		$\hat{\theta}=1.383333$	$\hat{\theta}=1.818978$	$\hat{\theta}=2.208089$	$\hat{\theta}=5.9389$ $\hat{\alpha}=7.0543$	$\hat{\theta}=2.7985$ $\hat{\alpha}=27.6732$
χ^2		1.87	0.64	0.47	0.55	0.10
d.f.		2	3	3	1	1
P-value		0.3926	0.8872	0.9254	0.4583	0.7518
$-2\log L$		147.10	187.08	186.69	185.02	185.56
AIC		149.10	189.08	188.69	189.02	189.56

Table 5: Pedestrians-Eugene, Spring, Morning

Group Size	Observed frequency	Expected frequency				
		SBPD	SBPLD	SBPSD	GSBPLD	AGSBPSD
1	1486	1452.4	1532.5	1532.4	1486.1	1502.0
2	694	743.3	630.6	629.6	694.5	670.6
3	195	190.2	191.9	192.3	192.4	193.6
4	37	32.4	51.3	51.7	41.1	45.3
5	10	4.1	12.8	12.9	7.5	9.3
6	1	0.6	3.9	4.1	1.5	2.2
Total	2423	2423.0	2423.0	2423.0	2423.0	2423.0
ML Estimate		$\hat{\theta}=0.5118$	$\hat{\theta}=4.5082$	$\hat{\theta}=4.9576$	$\hat{\theta}=10.7002$ $\hat{\alpha}=4.3729$	$\hat{\theta}=7.4281$ $\hat{\alpha}=96.9832$
χ^2		7.370	13.760	14.32	0.880	2.53
d.f.		2	3	3	2	2
p-value		0.0251	0.003	0.0025	0.644	0.2822
$-2\log L$		10445.34	4622.36	4622.86	4607.70	4609.81
AIC		10447.34	4624.36	4624.86	4611.70	4613.81

6. Concluding Remarks

A generalization of size-biased Poisson-Sujatha distribution (AGSBPSD) which includes both the size-biased Poisson-Lindley distribution (SBPLD) and the size-biased Poisson-Sujatha distribution (SBPSD) as particular cases has been proposed. Moments based measures including coefficients of variation, skewness, kurtosis, and index of dispersion have been derived and their behaviors have been discussed with varying values of the parameters. Maximum likelihood estimation has been discussed for estimating the parameters. Goodness of fit of the distribution has been presented for two count datasets along with the goodness of fit of SBPD, SBPLD, SBPSD and GSBPLD.

Acknowledgment

Authors are thankful to the Chief Editor and referee for given valuable comments to the improve the quality of papers

CONFLICT OF INTERESTS

The author(s) declare that there is no conflict of interests.

REFERENCES

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

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

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

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

[5] M. Greenwood and G.U. Yule, An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the multiple attacks of disease or of repeated accidents, J. R. Stat. Soci. 83(2) (1920), 115 – 121

[6] K. H. Mir and M. Ahmad, Size-biased distributions and Their Applications, Pakistan J. Stat. 25 (3) (2009), 283 – 294.

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

[8] D.V. Lindley, Fiducial distributions and Bayes theorem, J. R. Stat. Soc. 20 (1) (1958), 102 – 107.

[9] M.E. Ghitany and D. K. Al-Mutairi, Size-biased Poisson-Lindley distribution and Its Applications, Metron - Int. J. Stat. LXVI (3) (2008), 299 – 311.

[10] R. Shanker, Sujatha distribution and Its Applications, Stat. Transit. -New Ser. 17 (3) (2016), 391 – 410

[11] H. Zakerzadeh and A. Dolati, Generalized Lindley distribution, J. Math. Ext. 3 (2) (2009), 13 – 25.

[12] R. Shanker, K.K. Shukla and F. Hagos, A Generalization of Sujatha distribution and its Applications to Real lifetime data, J. Inst. Sci. Technol. 22(1) (2017), 77 – 94.

[13] R. Shanker, The discrete Poisson-Sujatha distribution, Int. J. Probab. Stat. 5(1) (2016), 1 – 9

[14] E. Mahmoudi and H. Zakerzadeh, Generalized Poisson-Lindley distribution, Commun. Stat.-Theory Meth. 39 (2010), 1785 – 1798.

[15] R. Shanker and K.K. Shukla, A generalization of Poisson-Sujatha Distribution and Its Applications to Ecology, Int. J. Biomath. 12 (2) (2019), 1- 11.

[16] R. Shanker, and F. Hagos, Size-biased Poisson-Sujatha distribution with Applications, Amer. J. Math. Stat. 6(4) (2016), 145 – 154.

[17] R. Shanker and K.K. Shukla, A generalized size-biased Poisson-Lindley distribution and its applications to model size distribution of freely-forming small group, J. Sci. Res. 10 (2) (2018), 145 – 157.

[18] J.S. Simmonoff, Analyzing Categorical data, Springer, New York, (2003).

[19] J. James, The distribution of free-forming small group size, Amer. Soc. Rev. 18 (1953), 569 – 570.

[20] J.S. Coleman and J. James, The equilibrium size distribution of freely forming groups, Sociometry, 24(1) (1961), 36 - 45.