Generalized Verison of Size Biased Poisson-Ailamujia Distribution with Applications

Zahoor Ahmad[1], Adil Rahsid[2]and T. R. Jan[3]

P.G Department of Statistics, University of Kashmir, Srinagar

Abstract

The present paper provides a generalized version of size biased Poisson-Ailamujia distribution, which is obtained through commpouding technique by mixing size biased Poisson distribution with Ailamujia distribution. The explicit expressions for the factorial moments of the proposed model are derived that inturn generates the descriptive measures of the proposed distribution. Further the technique of maximum likelihood estimation has been used to estimate the moments of proposed distribution. The reliability of the proposed distibution has been tested by applying chi-square test of goodness of fit for modeling real world data sets from genetics.

Keywords: Compound distribution, size biased Poisson distributon, Ailamujia distribution, factorial moments, count data.

1 Introduction

In probability distribution theory there are vast number of discrete distributions that have specified applications but at times the observable data have distinct features that are not exhibited by these classical discrete distributions. So to overcome these limitations, researchers often develop generalized probability distributions so that these new distributions can be employed in those situations where classical distributions are not providing adequate fit. These new generalized distributions contain some additional parameters which makes it richer in terms of its flexibility for data modeling. But at times this addition of parameter(s) makes the parametric space of resulting distribution very large that may hinder the goodness of fit to a data set because the degrees of freedom becomes less which sometimes results in poor p-value for the proposed model. In this paper we consider size biased version of Poisson distribution which will be generalized by using compounding mechanism and the resulting distribution will have only one unknown parameter. Sankaran (1970) obtained a compound of Poisson distribution with that of Lindley distribution.Zamani and Ismail (2010) constructed a new compound distribution by compounding negative binomial with one parameter Lindley distribution.Aryuen and Bodhisuwan (2013) proposed a generalization of negative binomial distribution by compounding negative binomial distribution with exponential distribution.Adil and Jan (2014) explored a mixture of generalized negative binomial distribution with that of generalized exponential distribution**.** Recently Dey and Kumar (2017) provided an extension of generalized exponential distribution with applications in Ozone data**.** Some new compound distributions were obtained in the recent past by the numerous researchers for instance,Adil et.al (2018) obtained a new lifetime distribution to address issues related to lifetime of series system.

The present paper is organized as follows: In section (2), we present material and methods needed for the construction of proposed distribution. Factorial moments are mentioned in section (2). The parameter estimation by means of MLE has been discussed in section (3). Finally, real application and conclusion about new findings are respectively given in section (4) and (5)

2 Materials and Methods

Size biased Poisson distribution is defined by the probability mass function (p.m.f) given by

$$p_1(x) = \frac{e^{-\lambda}\,\lambda^{x-1}}{(x-1)!}\ , \qquad where \;\; x = 1, 2, \ldots, \lambda > 0 \qquad (1)$$

The parameter λ in (1) is itself a random variable having Ailamujia distribution with probability density function (pdf) given by

$$f_1(X = x, \alpha) = 4x\alpha^2 e^{-2\alpha x} \qquad , \qquad x \geq 0, \alpha > 0. \qquad (2)$$

Ailamujia distribution was developed by Lv.et.al (2002) for several engineering application. Usually the parameter λ in size biased Poisson distribution is fixed constant but here we have considered a problem in which the parameter λ is itself a random variable following Ailamujia distribution with p.d.f (2)

2.1 Definition of proposed model

If $X|\lambda$ follows size biased Poisson distribution symbolized as SP($X|\lambda$) where λ is itself a random variable following Ailamujia distribution then determining the distribution that results from marginalizing over λ will be known as a compound of size biased Poisson distribution with that of Ailamujia distribution denoted S $PAD(\lambda,\alpha)$. It may be noted here that proposed model will be a discrete since the parent distribution is discrete.

Theorem 1: The probability mass function of a compound of size biased Poisson Ailamujia distribution is

$$f_{PAD}(X=\alpha)=\frac{4\alpha^2 x}{(1+2\alpha)^{x+1}}, \qquad x=1,2,....,\alpha\geq 0.$$

Proof: Using the definition (1.2), the p.m.f of a compound of $P(\lambda)$ with $AD(\alpha)$

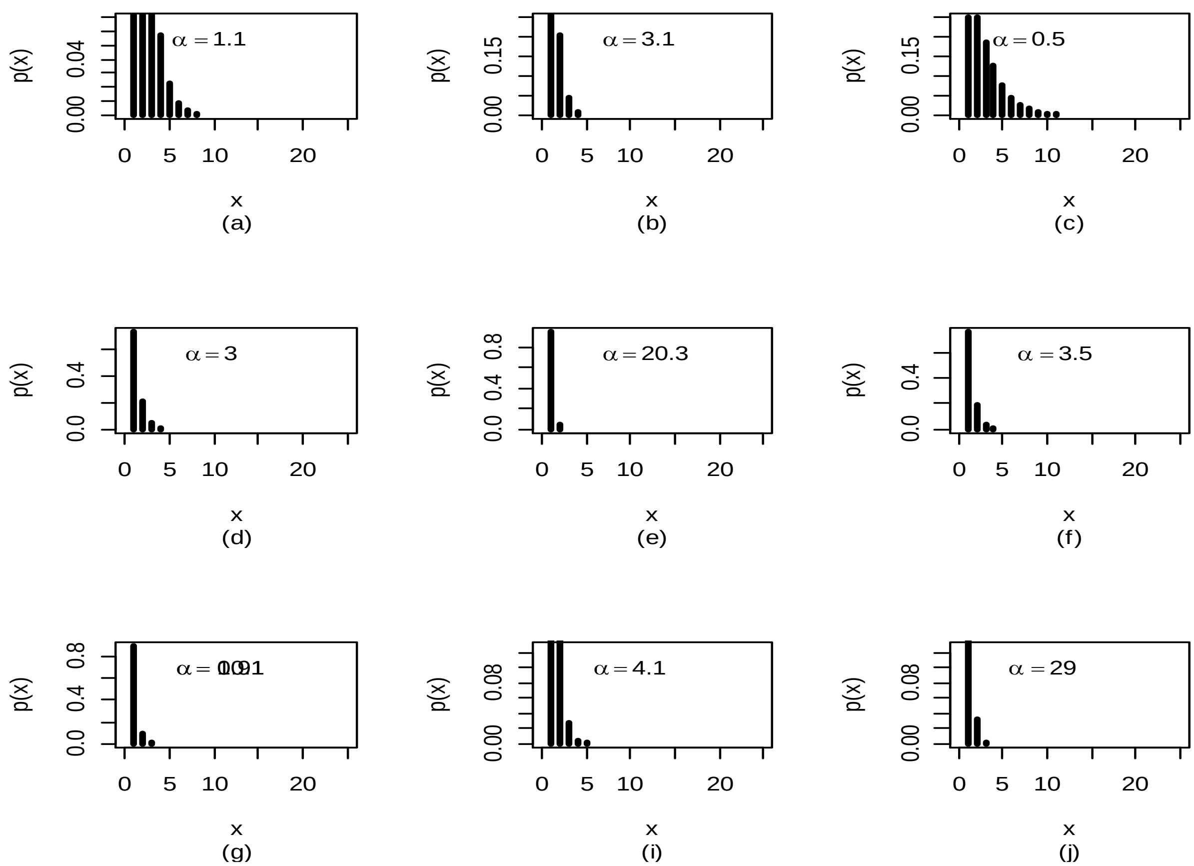
can be obtained as

$$f_{SPAD}(X;\alpha)=\int_0^\infty p_1(x|\lambda)\, f_1(x,\lambda)\, d\lambda$$

$$f_{SPAD}(X;\alpha)=\int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad 4\lambda\alpha^2 e^{-2\alpha\lambda}\, d\lambda$$

$$f_{SPAD}(X;\alpha)=\frac{4\alpha^2}{(x-1)!}\int_0^\infty e^{-(1+2\alpha)\lambda}\,\lambda^x \quad d\lambda$$

$$f_{SPAD}(X;\alpha)=\frac{4\alpha^2 x}{(1+2\alpha)^{x+1}}, \qquad x=1,2,....,\alpha\geq 0. \qquad (3)$$



3 Factorial moments of proposed model

Here we shall obtain factorial moments of the proposed model that will lead us to obtain some descriptive measures like mean, variance, standard deviation etc.

Theorem 2: The r^{th} order factorial moments of proposed model is given by

$$\mu_{[r]}(x) = 4\alpha^2 \left[\frac{r\Gamma(r+1)}{(2\alpha)^{r+1}} + \frac{\Gamma(r+2)}{(2\alpha)^{r+2}} \right]$$

Proof: The factorial moment of the proposed model can be obtained from the following definition $\mu_{[r]}(x) = E_\lambda\,[m_r(X \mid \lambda)]$

where $m_r(X \mid \lambda)$ is the r^{th}orderfactorial moments of size biased Poisson distribution which can be obtained as

$$m_r(X \mid \lambda) = E\left[\mathrm{X}^{(r)}\right]$$

$$= \sum_{x=1}^{\infty} x^{(r)} e^{-\lambda} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= e^{-\lambda}\lambda^{r-1}(\lambda e^{\lambda} + re^{\lambda})$$

$$= \lambda^{r-1}(\lambda + r) \quad for \quad r = 1, 2, ...$$

Therefore one obtains the factorial moments of a size biased Poisson-Ailamujia distribution, if we let λ as Ailamujiavariate

$$\mu_{[r]}(x) = E_{\lambda}\left[\lambda^{r-1}(\lambda + r)\right]$$

$$\mu_{[r]}(x) = 4\alpha^2 \int_0^{\infty} \lambda^r (\lambda + r) e^{-2\alpha\lambda} d\lambda$$

$$\mu_{[r]}(x) = 4\alpha^2 \left[\frac{r\Gamma(r+1)}{(2\alpha)^{r+1}} + \frac{\Gamma(r+2)}{(2\alpha)^{r+2}}\right]$$

For r=1, we get the mean of proposed distribution

$$\mu_{[1]}(x) = 4\alpha^2 \left[\frac{\Gamma(2)}{(2\alpha)^2} + \frac{\Gamma(3)}{(2\alpha)^3}\right]$$

$$= 4\alpha^2 \left[\frac{2\alpha + 2}{(2\alpha)^3}\right]$$

$$= \frac{\alpha + 1}{\alpha} = \mu_1'$$

$$\mu_{[2]}(x) = 4\alpha^2 \left[\frac{8\alpha + 6}{(2\alpha)^4} \right]$$

$$\mu_{[2]}(x) = \frac{4\alpha + 3}{2\alpha^2}$$

$$\mu_{[2]}(x) = \mu_2' - \mu_1'$$

$$\mu_2' = \frac{4\alpha + 3}{2\alpha^2} + \frac{\alpha + 1}{\alpha}$$

$$\mu_2' = \frac{2\alpha^2 + 6\alpha + 3}{2\alpha^2}$$

$$\text{Variance} = \frac{2\alpha^2 + 6\alpha + 3}{2\alpha^2} - \left(\frac{\alpha + 1}{\alpha} \right)^2$$

$$= \frac{2\alpha^2 + 6\alpha + 3 - 2\alpha^2 - 4\alpha - 2}{2\alpha^2} = \frac{2\alpha + 1}{2\alpha^2}$$

$$SD = \frac{\sqrt{2\alpha + 1}}{\alpha\sqrt{2}} = \frac{1}{\alpha}\sqrt{(2\alpha + 1)/2}$$

Moreover

$$CV = \sqrt{\frac{2\alpha + 1}{2(\alpha + 1)^2}} \times 100$$

4 Maximum likelihood Estimation

The parameter estimation of proposed model via maximum likelihood estimation method requires the log likelihood function of the model. Let $X_1, X_{2,}..., X_n$ are i.i.d variables from S $PAD(\alpha)$ then the log likelihood function of the model is

$$L = \prod_{i=1}^{n} P(X = x, \alpha)$$

$$L=\prod_{i=1}^{n}\frac{4\alpha^2 x_i}{(1+2\alpha)^{x_i+1}}$$

$$L=\frac{(4\alpha^2)^n\prod_{i=1}^{n}x_i}{(1+2\alpha)^{\sum x_i+n}}$$

$$\log L=n\log 4+2n\log\alpha+\sum_{i=1}^{n}\log x_i-(\sum_{i=1}^{n}x_i+n)\log(1+2\alpha).$$

$$\frac{\partial\log L}{\partial\alpha}=\frac{2n}{\alpha}-\frac{2(\sum_{i=1}^{n}x_i+n)}{(1+2\alpha)}$$

For estimating the unknown parameter we set the first order partial derivative equal to zero

$$\frac{\partial\log L}{\partial\alpha}=\frac{2n}{\alpha}-\frac{2(\sum_{i=1}^{n}x_i+n)}{(1+2\alpha)}=0$$. Hence $\hat{\alpha}=\frac{1}{\overline{x}-1}$

5 Applications

Here we will discuss the potentially of the proposed model by applying it to some reported real life count data sets:

1) Data regarding playgroups by Coleman and James(1961)
2) Immunogold data by Mathews and Appleton (1993)
3) Data regarding number of observed migrants, reported by Singh and Yadav (1971)
4) Snowshoe hares captured over 7 days, available in Keith and Meslow (1968)

Abbreviation	**Meaning**
SBQPLD	Size biased quasi Poisson Lindley distribution
SBNQPLD	Size biased new quasi Poisson Lindley distribution
NTPSBPLD	New three parameter size biased Poisson Lindley distribution
SBPAD	Size biased Poisson Ailamujia distribution

Table 1: Play groups/eugene, spring, public playground-A, available in Coleman and James (1961)					
Group Size	**Observed Frequency**	**Fitted Distribution**			
		SBQPLD	**SBNQPLD**	**NTPSBPLD**	**SBPAD**
1	306	304.4	306.4	304.4	310.1
2	132	137.9	134.4	137.9	130.3
3	47	41.3	41.6	41.3	42.1
4	10	10.3	11.0	10.3	11.5
5	2	3.1	3.6	3.1	3.0
Total	497	497	497	497	497
ML Estimates		$\hat{\theta} = 5.71$ $\hat{\alpha} = -0.06$	$\hat{\theta} = 4.99$ $\hat{\alpha} = 25.69$	$\hat{\theta} = 5.71$ $\hat{\alpha} = -0.07$ $\hat{\beta} = 5.81$	$\hat{\alpha}$ =1.88
Chi-square		1.19	1.20	1.19	1.08
df		1	1	0	2
P-value		**0.275**	**0.273**	**-**	**0.580**

Table 2: Distribution of counts of sites with particles from immunogold data available in Mathews and Appleton (1993)

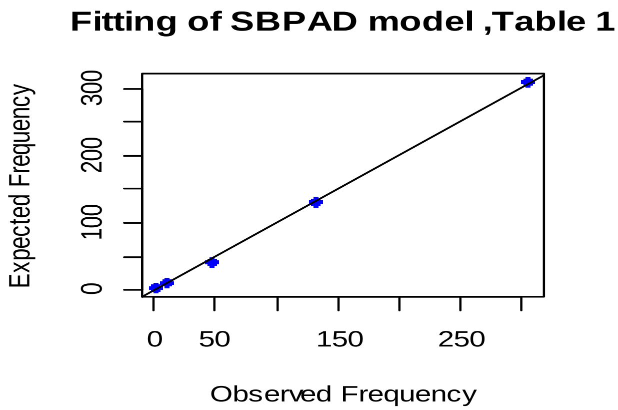
No. of sites with particles	Observed Frequency	Fitted Distribution			
		SBQPLD	SBNQPLD	NTPSBPLD	SBPAD
1	122	119.2	119.3	119.3	119.2
2	50	53.5	53.3	53.3	53.4
3	18	17.9	17.8	17.8	18.0
4	4	5.3	5.3	5.3	5.6
5	4	2.1	2.3	2.3	1.8
Total	198	198	198	198	198
ML Estimates		$\hat{\theta} = 3.75$ $\hat{\alpha} = -10.12$	$\hat{\theta} = 3.47$ $\hat{\alpha} = 0.0216$	$\hat{\theta} = 3.47$ $\hat{\alpha} = 1.396$ $\hat{\beta} = 0.0001$	$\hat{\alpha} = 1.73$
Chi-square		0.34	0.28	0.28	0.46
df		1	1	0	2
P-value		**0.559**	**0.596**	**-**	**0.794**

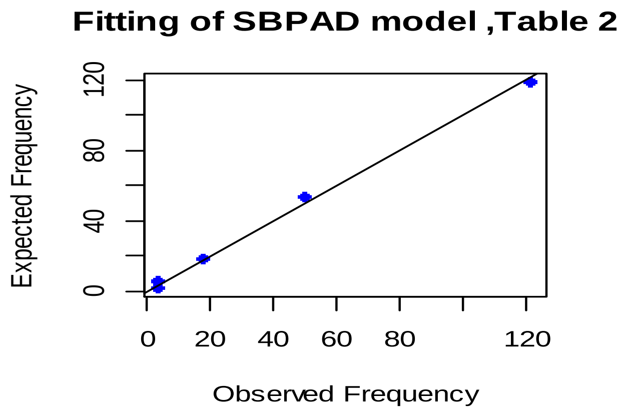
Table 3: No. of households having at least one migrant (X) according to the number of observed migrants, reported by Singh and Yadav (1971)

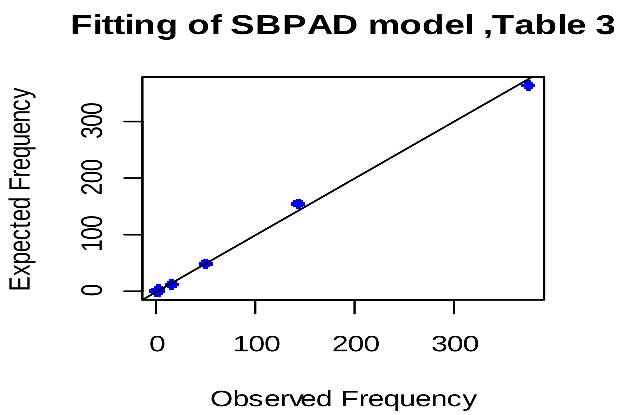
X	**Observed Frequency**	**Fitted Distribution**			
		SBQPLD	**SBNQPLD**	**NTPSBPLD**	**SBPAD**
1	375	363.3	363.6	363.6	363.1
2	143	156.5	156.3	156.3	156.5
3	49	50.4	50.4	50.4	50.6
4	17	14.4	14.4	14.4	14.4
5	2	3.9	3.8	3.8	3.9
6	2	1.0	1.0	1.0	1.0
7	1	0.2	0.2	0.2	0.3
8	1	0.4	0.3	0.3	0.2
Total	590	590	590	590	590
ML Estimates		$\hat{\theta} = 3.83$ $\hat{\alpha} = 17.29$	$\hat{\theta} = 3.65$ $\hat{\alpha} = 0.00067$	$\hat{\theta} = 3.65$ $\hat{\alpha} = 12.986$ $\hat{\beta} = -0.037$	$\hat{\alpha} = 1.82$
Chi-square		2.11	2.08	2.08	2.07
df		2	2	1	3
P-value		**0.348**	**0.353**	**0.149**	**0.559**

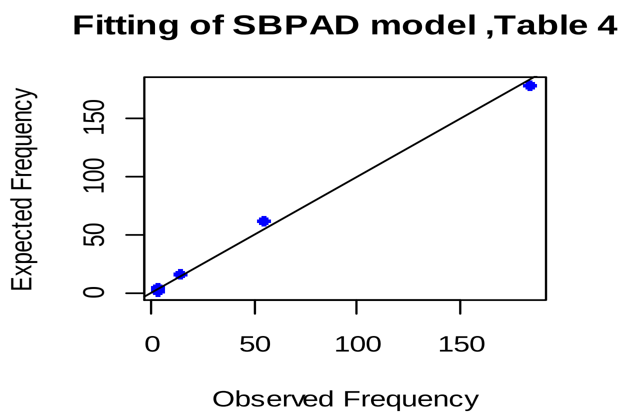
Table 4: Distribution of snowshoe hares captured over 7 days, available in Keith and Meslow (1968)

No. of times hares caught	Observed Frequency	Fitted Distribution			
		SBQPLD	**SBNQPLD**	**NTPSBPLD**	**SBPAD**
1	184	177.4	177.5	177.5	177.5
2	55	62.3	62.2	62.2	61.9
3	14	16.2	16.2	16.3	16.4
4	4	3.9	3.8	3.8	4.1
5	4	1.2	1.3	1.2	1.1
Total	261	261	261	261	261
ML Estimates		$\hat{\theta} = 4.98$ $\hat{\alpha} = 14.91$	$\hat{\theta} = 4.69$ $\hat{\alpha} = -0.030$	$\hat{\theta} = 4.69$ $\hat{\alpha} = 12.004$ $\hat{\beta} = -0.039$	$\hat{\alpha} = 2.3$
Chi-square		3.20	3.19	0.28	2.87
df		1	1	0	2
P-value		**0.073**	**0.074**	**-**	**0.239**









6 Conclusion

The present paper introduced a generalized version of size biased Poisson-Ailamujia distribution, which is stemmed through commpouding technique by composing size biased Poisson distribution with Ailamujia distribution. The exact expressions for the factorial moments of the proposed model are obtained, which inturn generates the descriptive measures of the proposed distribution. Apart from this, the technique of maximum likelihood estimation has been exploited to estimate the moments of proposed distribution. Finally, the application of the new model has been tested on several real life data sets and based on appropriate statistical tests we came to known that that the new model offers a superior fit as compared with other distributions. We hope that our new model will serve as an excellent competatorover various existing models available in statistics literature.

REFERENCES

1) Adil, R and Jan, T.R. A compound of Geeta distribution with the Generalized Beta Distribution. *Journal of Applied probability and Statistics Methods***, 9**(2), 2014b, 17-24.
2) Adil, R, Jan, T.R. A and Zahoor (2018). A new lifetime distribution for series system: Model, Properties and Application. *Journal of Modern Applied Statistical Methods,* 17(1), 1-25.
3) Coleman JS., James J. (1961). The equilibrium size distribution of freely forming groups, *Sociometry*, 24, 36–45.
4) Ghitany, M.E.,Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application, *Mathematics and Computers in Simulation* 78, 493–506.
5) Gove. J.H. (2003). Estimation and applications of size biased distribution in forestry. *Journal of Modeling Forest Systems*, 201-212.
6) Keith LB., Meslow EC. (1968). Trap response by snowshoe hares, *Journal of Wildlife Management*, 32(4), 795–801.
7) Mathews JNS., Appleton DR. (1993). An application of the truncated Poisson distribution to Immunogold assay*, Biometrics,* 49(2), 617–621.
8) Rama. S., and Kamlesh, K. (2020). A new three parameter size biased Poission Lindley distribution, *Biometrics and Biostatistics International Journal, 9(1) 01-14.*
9) Shanker R., (2017). A Size biased Poisson Shanker distribution and its applications, *International Journal of Probability and Statistics,* 6(3) 33–44.
10) Shanker, R., Hagos, F., and Abrehe, Y. (2015). On size biased Poission Lindley distribution and its application to model tunder storms, *Americian Jouranl of Mathematics and Statistics,* 5(6), 354–360.
11) Singh SN, Yadav RC (1971). Trends in rural out–migration at household level, *Rural Demogr,* 8(1), 53–61.