Pak. J. Statist.
2021 Vol. 37(1), 19-38

ON SIZE BIASED POISSON AILAMUJIA DISTRIBUTION AND ITS APPLICATIONS

Anwar Hassan[1], Showkat Ahmad Dar[1§], Peer Bilal Ahmad[2] and **Bilal Ahmad Para[2]**

[1] Department of Statistics, University of Kashmir, Srinagar, (J&K), India.
Email: anwar.hassan2007@gmail.com

[2] Department of Mathematical Sciences, Islamic University of Science & Technology, Awantipora, Pulwama (J&K), India.
Email: peerbilal@yahoo.co.in; parabilal@gmail.com

[§] Corresponding author Email: darshowkat2429@gmail.com

ABSTRACT

In this paper, we obtained a new model for count data by compounding of size biased Poisson distribution(SBPD) with Size biased Ailamujia distribution(SBAD). Important mathematical and statistical properties of the distribution have been derived and discussed. The expression for coefficient of variation, skewness and kurtosis has been obtained. Then, parameter estimation is discussed using momemts method and maximum likelihood method of estimation. Finally, real data set is analyzed to investigate the suitability of the proposed distribution in modeling count data.

KEYWORDS

Size Biased Poisson distribution, Size Biased Ailamujia Distribution, Compound Distribution, Count Data.

1. INTRODUCTION

Compounding a discrete distribution with a continuous one is a technique to create new distribution. Discrete compound probability distributions are of great significance in several applications for theoretical research and applied fields such as biological, engineering, insurance, medical and life testing. Compound probability distribution provides great flexibility in modelling data in practice. By compounding technique we can obtain both continuous and discrete probability distributions. Green wood and Yule started work in this field in 1920 by establishing relation between Poisson and negative binomial distribution through compounding technique by setting rate parameter in Poisson distribution as gamma variable. Skellam (1948) derived a probability distribution from the binomial distribution by regarding the probability of success as a beta variable between sets of trials. Sankaran (1970) establishing a relation between Poisson and Lindley distribution for modeling count data. Gerstenkorn (1993, 1996) obtain the relationship between gamma distribution and exponential distribution through compounding technique by setting the parameters in gamma distribution as exponential variable and also constructed polya with beta distribution through compounding technique. Shanker and Hagos (2016) constructed a new Size biased Poisson Sujatha distribution by setting the Probability parameters in Size biased Poisson distribution as a random variable for Size biased Sujatha distribution. Hassan, Dar and Ahmad (2019) introduced a new compounding probability model for count data, by compounding Poisson distribution with Ishita distribution that finds its applications in epileptic seizure. Gupta and Ong (2004) constructed a new generalized negative binomial distribution , this distribution arises from Poisson distribution if the rate parameter follows generalized gamma distribution; the resulting distribution so obtained was applied to various data sets and can be used as better alternative to negative binomial distribution. Zamani and Ismail (2010) introduced a new count data model by compounding the negative binomial with one parameter Lindley distribution. Panger and Willmot (1981) obtained a compounding model by mixing Negative Binomial distribution with Exponential distribution and showed the flexibility of this particular model in fitting count data. Lord and Geedipall (2011) showed that Poisson distribution tends to under estimate the number of zeros given the mean of the data while the Negative Binomial distribution over estimates zero, but under estimate observations with a count. Altun (2019) introduced Poisson cousil lindly regression model for analyzing over dispersed count data.

In this paper we propose a new compounding distribution by compounding SBPD with SBAD, as there is a need to find more flexible model for analyzing statistical data.

2. SIZE BIASED POISSON AILAMUJIA DISTRIBUTION

If $Z|\nu \sim \text{SBP}(\nu)$, ν is itself a random variable having size biased Ailamujia distribution, then the resulting distribution obtained by marginalizing over ν will be recognized as a compound of size biased Poisson distribution (SBPD) with that of size biased Ailamujia distribution (SBAD) and is denoted by $SBPA(Z;\eta)$. The given model will be a discrete as parent distribution is discrete

Theorem 2.1:

The probability mass function (pmf) of a Size Biased Poisson Ailamujia Distribution i.e., $SBPA(Z;\eta)$ is given by

$$P(Z=z)=\frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}} \quad ;z=1,2,3,...,;\eta>0$$

Proof:

The pmf of a SBPD is given by

$$j(z\,|\,\nu)=\frac{e^{-\nu}\nu^{z-1}}{(z-1)!}\ ;\ z=1,2,3,...,;\nu>0$$

When its parameter ν follows SBAD with pdf

$$k(\nu;\eta)=4\nu^2\eta^3 e^{-2\eta\nu}\ ;\nu>0,\eta>0$$

The pmf of a compound SBP distribution and SBA distribution can be obtained as

$$P(Z=z)=\int_0^\infty j(z\mid\nu)k(\nu;\eta)d\nu$$

$$P(Z=z)=\int_0^\infty \frac{e^{-\nu}\nu^{z-1}}{(z-1)!}4\nu^2\eta^3 e^{-2\eta\nu}d\nu$$

$$P(Z=z)=\frac{4\eta^3}{(z-1)!}\int_0^\infty e^{-(1+2\eta)\nu}\nu^{z+1}\,d\nu$$

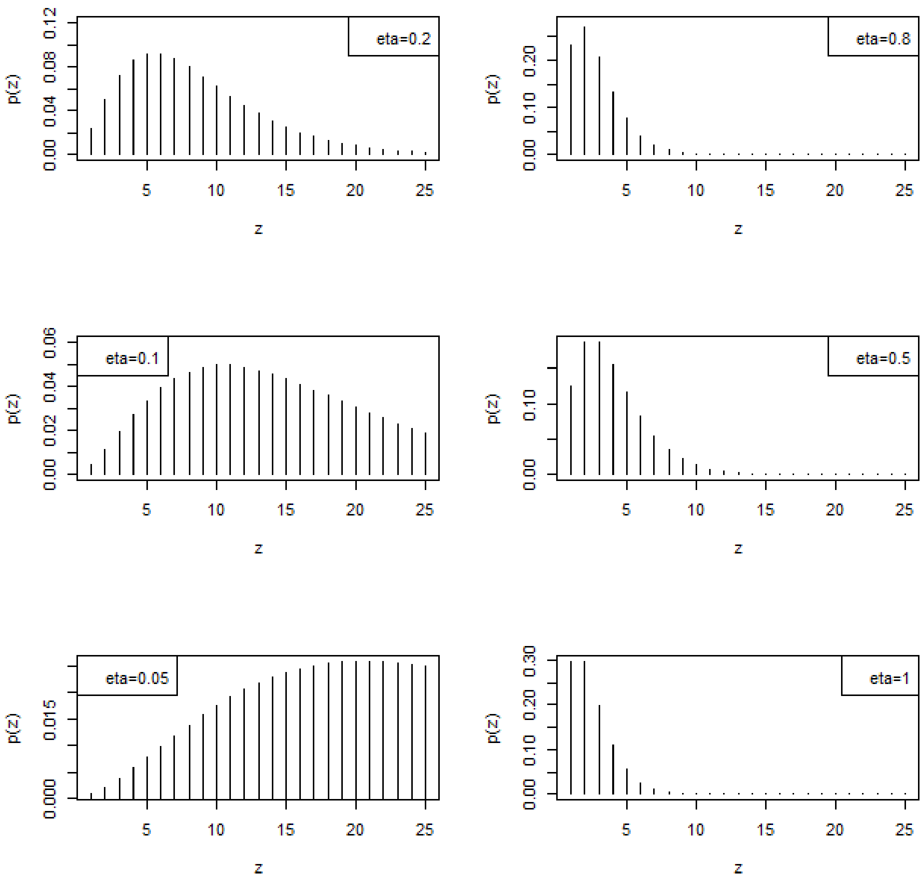
$$P(Z=z)=\frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}};z=1,2,3,...,;\eta>0$$

Which is the p.m.f. of SBPAD.

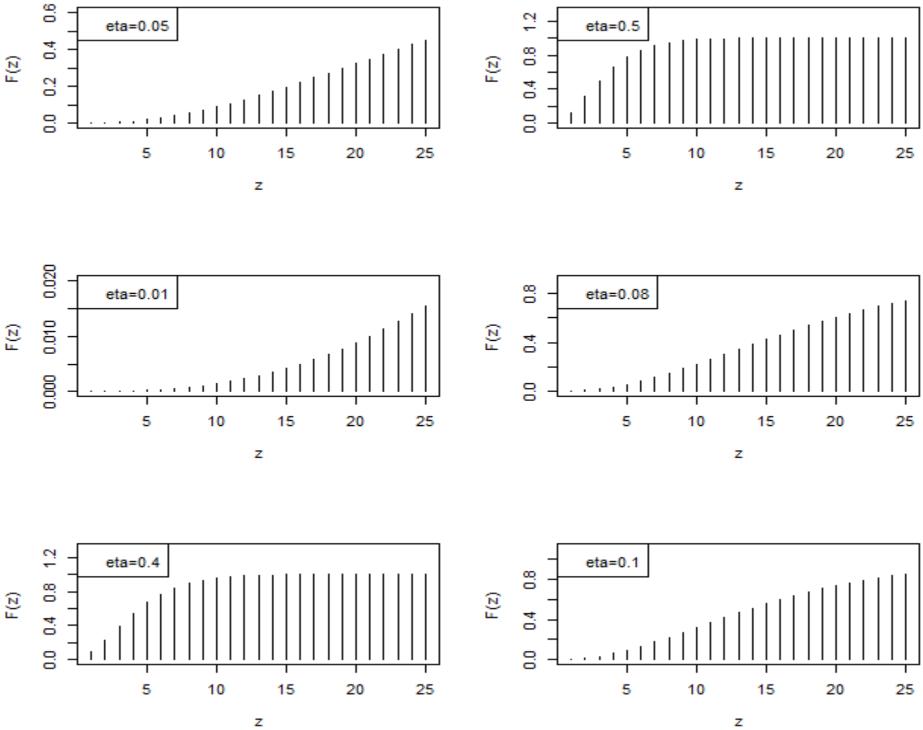
The corresponding c.d.f of SBPAD is obtained as:

$$F_Z(z)=\sum_{n=1}^{z}\frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}}$$

$$1-\frac{2z^2\eta^2+6z\eta^2+4\eta^2+2z\eta+4\eta+1}{(1+2\eta)^{z+2}};\ z=1,2,3\ldots,\ \eta>0\ .$$



Figures 1: pmf Plot of SBPAD for different Values of η



Figures 2: cdf Plot of SBPAD for different Values of η

3. STATISTICAL PROPERTIES

This section explains the various structural characteristics of the SBPA model. This section comprises of moments, moment generating function and probability generating function.

3.1 Moments

The s^{th} factorial moment about origin of the SBPAD can be obtained as

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

$$\mu_{(s)}' = \int_0^\infty \left[\sum_{z=1}^\infty z^{(s)} \frac{e^{-\nu}\nu^{z-1}}{(z-1)!}\right].4\eta^3\nu^2 e^{-2\eta\nu} d\nu$$

$$\mu_{(s)}' = \int_0^\infty \left[\nu^{s-1}\left(\sum_{z=s}^\infty z \frac{e^{-\nu}\nu^{z-s}}{(z-s)!}\right)\right] 4\eta^3\nu^2 e^{-2\nu\eta} d\nu .$$

Taking $z = a + s$, we get

$$\mu_{(s)}' = \int_0^\infty \left[\nu^{s-1} \left(\sum_{a=0}^{\infty} (a+s) \frac{e^{-\nu}\nu^a}{a!} \right) \right] 4\eta^3 \nu^2 e^{-2\eta\nu} d\nu$$

$$\mu_{(s)}' = 4\eta^3 \int_0^\infty \nu^{s+2-1} (\nu + s) e^{-2\eta\nu} d\nu$$

$$\mu_{(s)}' = 4\eta^3 (s+1)! \left[\frac{(s+2) + s(2\eta)}{(2\eta)^{s+3}} \right]. \tag{3.1}$$

Put s =1, 2, 3 and 4 in (3.1), we get the first four factorial moments about origin of SBPAD are as under

$$\mu_{(1)}' = \frac{3+2\eta}{2\eta}$$

$$\mu_{(2)}' = \frac{3(1+\eta)}{\eta^2}$$

$$\mu_{(3)}' = \frac{3(5+6\eta)}{2\eta^3}$$

$$\mu_{(4)}' = \frac{15(3+4\eta)}{2\eta^4}.$$

The first four moments about origin of SBPAD are given as

$$\mu_1' = \frac{3+2\eta}{2\eta}$$

$$\mu_2' = \frac{2\eta^2 + 9\eta + 6}{2\eta^2}$$

$$\mu_3' = \frac{2\eta^3 + 21\eta^2 + 36\eta + 15}{2\eta^3}$$

$$\mu_4' = \frac{2\eta^4 + 45\eta^3 + 150\eta^2 + 110\eta + 15}{2\eta^4}.$$

Using the relationship $\mu_s = E(Y - \mu_1')^s = \sum_{J=0}^{s} \binom{s}{J} \mu_k' (-\mu_1')^{s-J}$ we get, the moments about the mean of the SBPAD is given as

$$\mu_2 = \frac{3+6\eta}{(2\eta)^2}$$

$$\mu_3 = \frac{6\eta^2 - 57\eta + 3}{4\eta^3}$$

$$\mu_4 = \frac{24\eta^3 - 288\eta^2 - 240\eta + 525}{16\eta^4}.$$

The coefficient of variation (C.V), coefficient of skewness $(\sqrt{\beta_1})$, coefficient of kurtosis (β_2) and index of dispersion (γ) of the SBPAD are thus obtained as

$$C.V = \frac{\sigma}{\mu_1'} = \frac{\sqrt{3+6\eta}}{3+2\eta}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{12\eta^2 - 114\eta + 6}{(3+6\eta)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{8\eta^3 - 96\eta^2 - 80\eta + 175}{12\eta^2 + 12\eta + 3}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{6\eta+3}{6\eta+4\eta^2}.$$

3.2 Moment Generating Function (MGF) and Probability Generating Function (PGF) of Size Biased Poisson Ailamujia Distribution

We will derive PGF and MGF of SBPAD in this section.

Theorem 3.2.1:

Let Z follows SBPAD(η), then Z has Probability generating function $P_Z(t)$ is given by

$$P_Z(t) = \frac{8\eta^3}{1+2\eta}\left[\frac{t+2\eta t}{(1+2\eta-t)^3}\right].$$

Proof:

From the renowned definition of the PGF, we have

$$P_z(t) = \sum_{z=1}^{\infty} t^z \frac{4\eta^3(z^2+z)}{(1+2\eta)^{z+2}}$$

$$P_Z(t) = \frac{4\eta^3}{(1+2\eta)^2}\left[\sum_{z=1}^{\infty} z^2\left(\frac{t}{1+2\eta}\right)^z + \sum_{z=1}^{\infty} z\left(\frac{t}{1+2\eta}\right)^z\right]$$

$$P_Z(t) = \frac{4\eta^3}{(1+2\eta)^2}\left[\frac{t(1+2\eta)(1+2\eta+t)}{(1+2\eta-t)^3} + \frac{t(1+2\eta)}{(1+2\eta-t)^2}\right]$$

$$P_Z(t) = \frac{4\eta^3 t}{(1+2\eta)}\left[\frac{(1+2\eta+t)+(1+2\eta-t)}{(1+2\eta-t)^3}\right]$$

$$P_Z(t) = \frac{8\eta^3}{(1+2\eta)}\left[\frac{t+2\eta t}{(1+2\eta-t)^3}\right].$$

Theorem 3.2.2:

Let Z follows SBPAD (η), then Z has moment generating function $M_Z(t)$ given as

$$M_Z(t) = \frac{8\eta^3}{(1+2\eta)}\left[\frac{e^t + 2\eta e^t}{(1+2\eta-e^t)^3}\right].$$

Proof:

From the renowned definition of the MGF, we have

$$M_Z(t) = \sum_{z=1}^{\infty} e^{tz} \frac{4\eta^3(z^2+z)}{(1+2\eta)^{z+2}}$$

$$M_Z(t) = \frac{4\eta^3}{(1+2\eta)^2}\left[\sum_{z=1}^{\infty} z^2\left(\frac{e^t}{1+2\eta}\right)^z + \sum_{z=1}^{\infty} z\left(\frac{e^t}{1+2\eta}\right)^z\right]$$

$$M_Z(t) = \frac{4\eta^3}{(1+2\eta)^2}\left[\frac{e^t(1+2\eta)(1+2\eta+e^t)}{(1+2\eta-e^t)^3} + \frac{e^t(1+2\eta)}{(1+2\eta-e^t)^2}\right]$$

$$M_Z(t) = \frac{4\eta^3 e^t}{(1+2\eta)}\left[\frac{(1+2\eta+e^t)+(1+2\eta-e^t)}{(1+2\eta-t)^3}\right]$$

$$M_Z(t) = \frac{8\eta^3}{(1+2\eta)}\left[\frac{e^t + 2\eta e^t}{(1+2\eta-e^t)^3}\right].$$

3.3 Mean and Variance through MGF

$$\mu_1' = \frac{\delta}{\delta t} M_Z(t)\Big| t=0$$

$$\mu_1' = \frac{\delta}{\delta t} M_Z(t)\Big|t=0 \;\; = \frac{\delta}{\delta t}\left[\frac{8\eta^3}{(1+2\eta)}\left(\frac{e^t+2\eta e^t}{(1+2\eta-e^t)^3}\right)\right]\Big|t=0$$

$$\mu_1' = \frac{2\eta+3}{2\eta}$$

$$\mu_2' = \frac{\delta^2}{\delta t}\left[\frac{8\eta^3}{(1+2\eta)}\left(\frac{e^t+2\eta e^t}{(1+2\eta-e^t)}\right)\right]\Big|t=0$$

$$\mu_2' = \frac{2\eta^2+9\eta+6}{2\eta^2}$$

$$\mu_2 = \frac{3+6\eta}{(2\eta)^2}\,.$$

3.4 Recurrence Relation between Probabilities

If Z~SBPAD(η) then the pmf of Z is

$$P(Z=z) = \frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}}$$

$$P(Z=z+1) = \frac{4\eta^3(z+1)(z+2)}{(1+2\eta)^{z+3}}\,.$$

Dividing $P(Z=z+1)$ by $P(Z=z)$, we find the recurrence relation between probabilities

$$\frac{P(Z=z+1)}{P(Z=z)} = \frac{(z+2)}{z(1+2\eta)}$$

$$P(Z=z+1) = \frac{(z+2)}{z(1+2\eta)}P(Z=z)\,.$$

4. RECURRENCE RELATION BETWEEN RAW MOMENTS, CENTRAL MOMENTS AND FACTORIAL MOMENTS OF SBPAD

Theorem 4.1:

If Z has a SBPAD with parameter η and μ_s' is the s^{th} raw moment, prove that $\mu_{s+1}' = \frac{(1+2\eta)}{2}\left\|\frac{3}{\eta}\mu_s' - \frac{4}{(1+2\eta)}\mu_s' - \frac{\delta}{\delta\eta}\mu_s'\right\|.$

Proof:

Since Z follows SBPAD, The s^{th} raw moment is given by

$$\mu_s' = \frac{\sum_{z=1}^{\infty} z^s 4z(z+1)\eta^3}{(1+2\eta)^{z+2}}$$

$$\frac{\delta}{\delta\eta}\mu_s' = \sum_{z=1}^{\infty} 4z^s 4z(z+1)\left[\frac{(1+2\eta)^{z+2}\frac{\delta}{\delta\eta}\eta^3 - \eta^3\frac{\delta}{\delta\eta}(1+2\eta)^{z+2}}{\left\{(1+2\eta)^{z+2}\right\}^2}\right]$$

$$\frac{\delta}{\delta\eta}\mu_s' = \left[\frac{3}{\eta}\mu_s' - \frac{2\mu_{s+1}'}{(1+2\eta)} - \frac{4\mu_s'}{(1+2\eta)}\right]$$

$$\mu_{(s+1)}' = \frac{(1+2\eta)}{2}\left|\frac{3}{\eta}\mu_s' - \frac{4\mu_s'}{(1+2\eta)} - \frac{\delta}{\delta\eta}\mu_s'\right|.$$

Theorem 4.2:

If Z is a SBPAD with parameter η and μ_s' is the s^{th} central moment, prove that

$$\mu_{s+1} = \frac{(1+2\eta)}{2}\left[\frac{3}{\eta}\mu_s - s\mu_{s-1}\frac{\delta}{\delta\eta}\mu_1' - \frac{2(\mu_1'+2)\mu_s}{(1+2\eta)} - \frac{\delta}{\delta\eta}\mu_s\right].$$

Proof:

By definition

$$\mu_s = E(z-\mu_1')$$

$$\mu_s = \frac{\sum_{z=1}^{\infty}(z-\mu_1')^s 4z(z+1)\eta^3}{(1+2\eta)^{z+2}}$$

$$\frac{\delta}{\delta\eta}\mu_s = \sum_{z=1}^{\infty} 4z(z+1)\left[\frac{(1+2\eta)^{z+2}\frac{\delta}{\delta\eta}((z-\mu_1')\eta^3) - (z-\mu_1')^s\eta^3\frac{\delta}{\delta\eta}(1+2\eta)^{z+2}}{\{(1+2\eta)^{z+2}\}^2}\right]$$

$$\frac{\delta}{\delta\eta}\mu_s = \frac{3}{\eta}\mu_s - s\mu_{s-1}\frac{\delta}{\delta\eta}\mu_1' - \frac{2\mu_{s+1}}{(1+2\eta)} - \frac{2(\mu_1'+1)\mu_s}{(1+2\eta)}$$

$$\mu_{s+1} = \frac{(1+2\eta)}{2}\left[\frac{3}{\eta}\mu_s - s\mu_{s-1}\frac{\delta}{\delta\eta}\mu_1' - \frac{2(\mu_1'+2)\mu_s}{(1+2\eta)} - \frac{\delta}{\delta\eta}\mu_s\right].$$

Theorem 4.3:

If Z is a SBPAD with parameter η and $\mu^{[s]}$ is the s^{th} factorial moments, prove that

$$\mu^{[s+1]} = \frac{(1+2\eta)}{2}\left[\frac{3\mu^{[s]}}{\eta} - \frac{4\mu^{[s]}}{(1+2\eta)} - \frac{\delta}{\delta\eta}\mu^{[s]}\right].$$

Proof:

Since Z follows SBPAD, The sth factorial moment is given by

$$\mu^{[s]} = \sum_{z=1}^{\infty} z^{[s]} 4z(z+1)\frac{\eta^3}{(1+2)^{z+2}}$$

$$\frac{\delta}{\delta\eta}\mu^{[s]} = \sum_{z=1}^{\infty} z^{[s]} 4z(z+1)\left[\frac{3\eta^2(1+2\eta)^{z+2} - 2(z+2)(1+2\eta)^{z+1}\eta^3}{\left\{(1+2\eta)^{z+2}\right\}}\right]$$

$$\frac{\delta}{\delta\eta}\mu^{[s]} = \frac{3\mu^{[s]}}{\eta} - \frac{2\mu^{[s+1]}}{(1+2\eta)} - \frac{4\mu^{[s]}}{(1+2\eta)}$$

$$\mu^{[s+1]} = \frac{(1+2\eta)}{2}\left[\frac{3\mu^{[s]}}{\eta} - \frac{4\mu^{[s]}}{(1+2\eta)} - \frac{\delta}{\delta\eta}\mu^{[s]}\right].$$

5. RELIABILITY ANALYSIS

In this section we have introduced the reliability, hazard rate, reverse hazard rate and Mills ratio of the proposed SBPA model.

5.1 Reliability Function

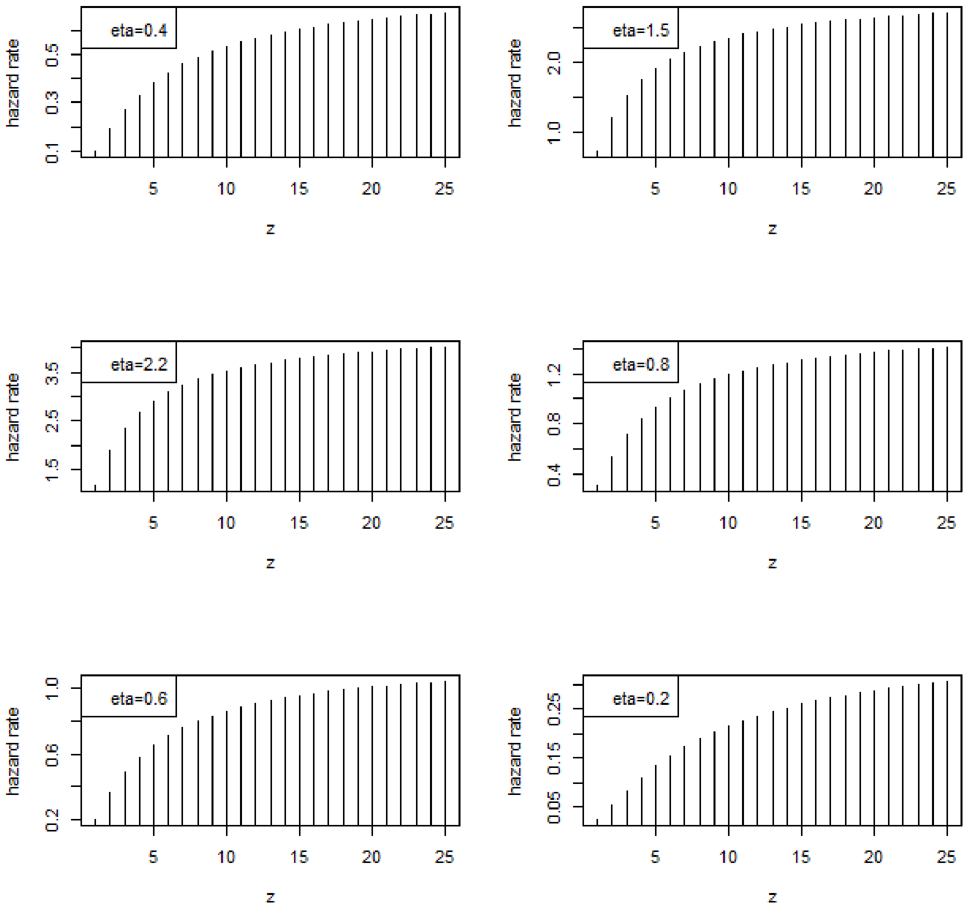
Reliability for any system may be defined as probability that beyond a certain time period system will function. The reliability function or the survival function of SBPAD is calculated as:

$$R(z,\eta) = \left(\frac{2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1}{(1+2\eta)^{z+2}}\right).$$

5.2 Hazard Function (H.F)

H.F is also known as hazard rate, instantaneous failure rate or force of mortality for a system of given age Z is defined as rate of death of the system and is calculated as below

$$\text{H.R=h }(z,\eta) = \frac{f(z,\eta)}{R(z,\eta)} = \left(\frac{4\eta^3 z(z+1)}{2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1}\right).$$



Figures 3: Hazard Rate Plot of SBPAD for different Values of η

5.3 Reverse Hazard Rate (R.H.R) and Mills Ratio

The reverse hazard rate and the mills ratio of Size biased Poisson Ailamujia distribution (SBPAD) are respectively given as:

$$\text{R.H.R} = h(z,\eta) = \frac{f(z,\eta)}{F(z,\eta)} = \frac{4\eta^3 z(z+1)}{(1+\eta)^{z+2}(2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1)}$$

$$\text{Mills ratio} = \frac{(1+\eta)^{z+2}(2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1)}{4\eta^3 z(z+1)}.$$

6. ORDER STATISTICS

Let $Z_{(1)}, Z_{(2)}, Z_{(3)}...., Z_{(n)}$ be the ordered statistics of the random sample $Z_1, Z_2, Z_3...Z_n$ drawn from the SBPAD with c.d.f $F_Z(Z)$ and p.m.f. $P_Z(Z)$, then the p.m.f. of sth order statistics $Z_{(s)}$ is given by:

$$f_{z(s)}(z,\eta) = \frac{n!}{(s-1)!(n-s)!} P(z)\left[F(z)\right]^{s-1}\left[1-F(z)\right]^{n-s}. \; s=1, 2, 3,\ldots,n \quad (6.1)$$

Put the value of c.d.f and p.m.f. in equation (6.1) we get the sth order statistics of SBPAD is given as

$$f_{(s)}(z,\eta) = \frac{n!}{(s-1)!(n-s)!} \frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}} \left[1 - \frac{2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1}{(1+2\eta)^{z+2}}\right]^{S-1} \left[\frac{2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1}{(1+2\eta)^{z+2}}\right]^{n-S}$$

Then, the p.m.f of first order $Z_{(1)}$ SBPA distribution is given by:

$$f_1(z,\eta) = n \frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}} \left[\frac{2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1}{(1+2\eta)^{z+2}}\right]^{n-1}.$$

And the pmf of nth order $Z_{(n)}$ SBPA model is given as:

$$f_{(n)}(z,\eta) = n \frac{4\eta^3 z(z+1)}{(1+2\eta)^{z+2}} \left[1 - \frac{2z^2\eta^2 + 6z\eta^2 + 4\eta^2 + 2z\eta + 4\eta + 1}{(1+2\eta)^{z+2}}\right]^{n-1}.$$

7. ESTIMATION OF PARAMETERS

In this section, the parameters of the SBPAD have been obtained by using different techniques.

7.1 Method of Moments

In order to obtain sample moments, we replace population moments with sample moments:

$$\mu_1' = \frac{\sum_{z=1}^{\infty} z_i}{n}$$

$$\bar{z} = \frac{2\eta+3}{2\eta}$$

$$\hat{\eta} = \frac{3}{2\bar{z}-2}.$$

Theorem 7.1.1:

The MOM estimator $\hat{\eta}$ of η is positively biased

Proof:

Let $\hat{\eta} = h(\bar{Z})$ where $h(t) = \dfrac{3}{2t-2,}$, $t > 1$

Since $h''(t) = \dfrac{24}{(2t-2)^3} > 0$

Then h(t) is strictly convex. Hence, by Jensen's inequality, we have

$$E\{h(\bar{z})\} > h\{E(\bar{z})\}$$

Finally, since $h\{E(\bar{z})\} = h(\mu) = h\left(\dfrac{2\eta+3}{2\eta}\right) = \eta,$ we obtain $E(\hat{\eta}) > \eta$

Theorem 7.1.2:

The MOM estimator $\hat{\eta}$ of η is consistent and asymptotically normal

$$\sqrt{n(\hat{\eta}-\eta)} \xrightarrow{d} N(0, \upsilon^2(\eta))$$

where $\upsilon^2(\eta) = \dfrac{(2\eta)^2(3+6\eta)}{36}$.

Proof:

Consistency: Since $\mu < \infty,$ then $\bar{Z} \xrightarrow{p} \mu$. Also since $h(t)$ is continuous function at $t = \mu,$ then $h(\bar{z}) \xrightarrow{p} h(\mu),$ i.e. $\hat{\eta} \xrightarrow{p} \eta$

Asymptotic normality: Since $\sigma^2 < \infty,$ then by the central limit theorem, we have

$$\sqrt{n}\ (\bar{Z} - \mu) \xrightarrow{d} N(0, \sigma^2)$$

Also, since $h(\mu)$ is differentiable and $h'(\mu) \neq 0,$ by the delta-method, we have

$$\sqrt{n}(h(\bar{z}) - h(\mu)) \xrightarrow{d} N(0, \upsilon^2(\eta))$$

where $\upsilon^2(\eta) = \dfrac{(2\eta)^2(3+6\eta)}{36}$

The theorem follows.

As a result of this, the asymptotic $100(1-\alpha)\%$ confidence interval for η is given by

$$\hat{\eta} \pm z_{\frac{\alpha}{2}} \frac{\upsilon(\hat{\eta})}{\sqrt{n}},$$

where $z_{\frac{\alpha}{2}}$ is the (1-$\dfrac{\alpha}{2}$) percentile of the standard normal distribution.

7.2 Method of Maximum Likelihood Estimation

Method of Maximum Likelihood Estimation is simple and most efficient method of estimation. Let $Z_1, Z_2 ... Z_n$ be the random size of sample n drawn from SBPAD, then the likelihood function of SBPAD is given as

$$L(z \mid \eta) = 4^n \eta^{3n} \left(\prod_{i=1}^{n} \frac{z(z+1)}{(1+2\eta)^{z+2}} \right)$$

$$\frac{\delta}{\delta\eta} \log L = \frac{3n}{\eta} - \frac{\sum_{i=1}^{n}(z_i + 2)}{(1+2\eta)} (2)$$

$$\hat{\eta} = \frac{3}{2\bar{z} - 2}.$$

8. APPLICATIONS OF SBPAD

In order to demonstrate the flexibility and applicability of the proposed distribution in modeling count datasets, we have analyzed the three data sets, so as to illustrate our claim that our proposed model fits well when compared to other competing models. The data set are given in Table 1, Table 4 and Table 7 respectively. We compute the expected frequencies for fitting Size biased Poisson distribution (SBPD), Size biased Poisson Lindley distribution (SBPLD) and Size biased Poisson Ailamujia distribution (SBPAD) with the help of R studio statistical software and Pearson's chi-square test is applied to check the goodness of fit of the models discussed. The calculated figures are given in the Table 3, 6 and 9. Based on the chi-square, we observe that Size biased Poisson Ailamujia distribution provides a satisfactorily better fit for the data sets. Thus we conclude that the proposed model fits the data better as compared to the other models. Also the parameters are estimated by using the ML method. We have analyzed the data using R software (3.5.2).

Furthermore, From Table 2, 5 and 8, it has been observed that the Size biased Poisson Ailamujia distribution have the lesser AIC and BIC values as compared to other competing models. Hence we can concluded that the SBPAD leads to a best fit as compared to other competing models for analyzing the data set given in Table 1, 4 and 7.

Table 1
Number of Households of Size Group 16 having at least One Migrant according to Number of Migrants Studied by Singh and Yadav (1971)

Number of Migrants (Z)	1	2	3	4	5	6	7	8
Frequency	18	21	16	9	2	2	1	1

Table 2
Model Comparison Criterion for Fitted Models to a Dataset Representing Number of Households

Criterion	SBPD	SBPLD	SBPAD
-logL	118.07	116.5	116.05
AIC	238.15	235	234.1
BIC	238.21	235.07	234.17

Table 3
Fitted Proposed Distribution and other Competing Models to a Dataset Representing Number of Households

Number of Migrants (Z)	Observed Frequency	SBPD	SBPLD	SBPAD
1	18	14.34	20.84	19.60
2	21	22.73	19.52	20.33
3	16	18.02	13.3	18.02
4	9	9.53	7.83	9.53
5	2	3.78	4.26	3.78
6	2	1.20	2.19	1.20
7	1	0.32	1.08	0.32
8	1	0.07	0.52	0.07
Degree of freedom		3	3	3
Parameter estimation		1.59	1.60	0.95
Standard error		0.15	0.18	0.11
Chi-Statistic-Value		1.40	1.75	0.85
P-Value		0.71	0.62	0.84

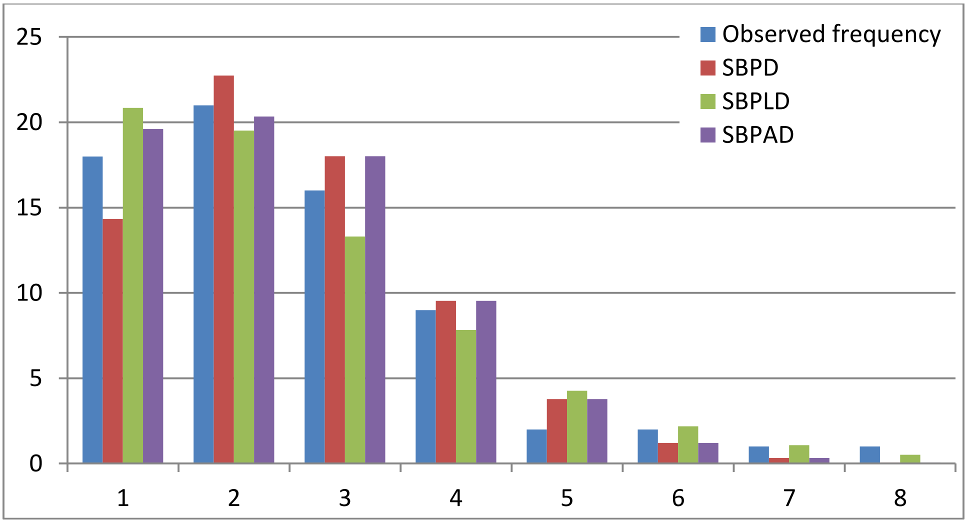


Figure 4: Graphical Overview of Fitted Data Models to Dataset Representing Number of Migrants

We have analyzed the date set regarding the number of groups of people in public places on spring studied by James (1953).

Table 4
Number of Groups of People in Public Places on Spring Studied by James (1953)

Size of groups (Z)	1	2	3	4	5	6
Frequencies	1486	694	195	37	10	1

Table 5
Model Comparison Criterion for Fitted Models to a Dataset Representing Number of Groups in Public Places

Criterion	**SBPD**	**SBPLD**	**SBPAD**
-logl	2308.69	2311.18	2306.01
AIC	4619.39	4624.36	4614.02
BIC	4619.17	4624.15	4613.80

Table 6
Fitted Proposed Distribution and other Competing Models to a Dataset Representing Number of Groups in Public Places

Z	**Observed Counts**	**SBPD**	**SBPLD**	**SBPAD**
1	1486	1452.40	1531.90	1510.33
2	694	743.30	630.80	660.5
3	195	190.20	192.10	192.6
4	37	32.40	51.40	46.8
5	10	4.10	12.70	10.22
6	1	0.60	4.10	2.08
Degree of freedom		2	3	3
Parameter Estimation		0.51	4.50	η =2.93
Standard error		0.014	0.13	0.09
Chi-Statistic Value		7.36	13.78	4.31
P-Value		0.025	0.0030	0.23

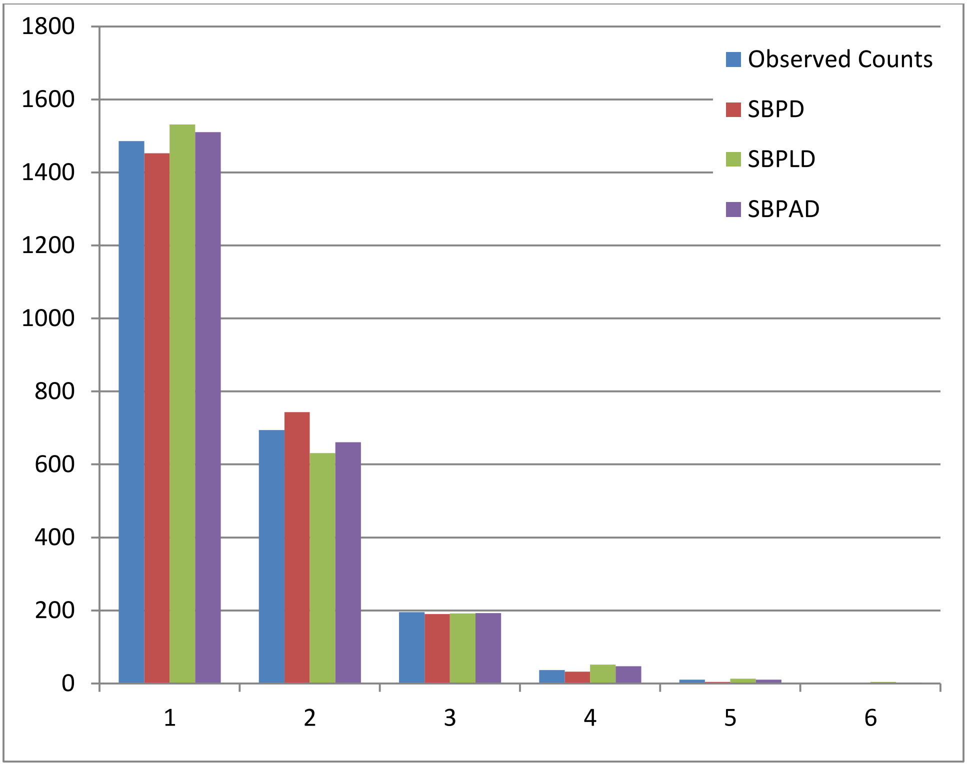


Figure 5: Graphical Overview of the Dataset regarding the Number of Groups of People in Public Places

We have analyzed the data set discussed by Coleman and James, The number of shopping groups-Eugene, spring, department store and public market (1961).

Table 7
Number of Shopping Groups-Eugene, Spring, Department Store and Public Market (1961)

Size of group (Z)	1	2	3	4	5
Frequency	316	141	44	5	4

Table 8
Model Comparison Criterion for Fitted Models to a Dataset Representing Number of Shopping Groups-Eugene, Spring

Criterion	**SBPD**	**SBPLD**	**SBPAD**
-logl	487.35	486.4	485.65
AIC	976.7	974.8	973.3
BIC	976.30	974.40	972.90

Table 9
Fitted Proposed Distribution and other Competing Models to a Dataset Representing Number of Shopping Groups-Eugene, Spring

Size of group (Z)	Frequency	SBPD	SBPLD	SBPAD
1	316	306.30	322.90	318.5
2	141	156.20	132.60	139
3	44	39.80	40.20	40.3
4	5	6.80	10.70	9.75
5	4	0.90	3.60	2.10
Degree of freedom		2	2	2
Parameter Estimation		0.50	4.52	2.94
Standard error		0.31	0.28	0.19
Chi-Statistic Value		2.44	3.002	1.07
P-Value		0.29	0.22	0.58

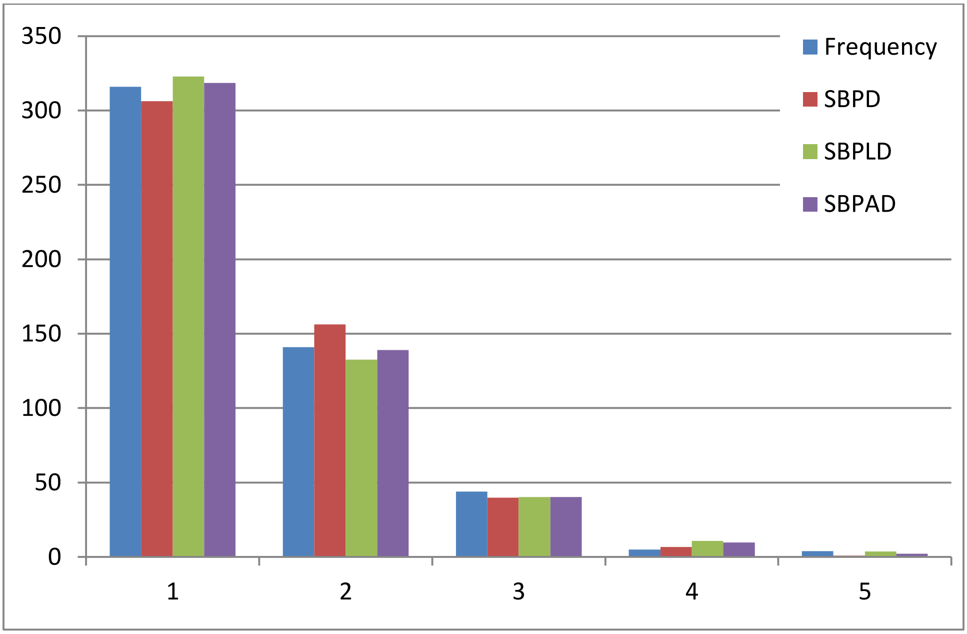


Figure 6: Graphical Overview of the Dataset regarding the Number of Shopping Groups

9. CONCLUSION

A new model has been built using compounding technique. Statistical and mathematical properties such as reliability, hazard rate, reverse hazard rate, Mills ratio, recurrence relation between probability and order statistics, have been discussed. Finally three real data sets are discussed to demonstrate the suitability and applicability of the proposed distribution in modeling count data sets.

REFERENCES

1. Altun, E. (2019). A new model for over-dispersed count data: Poisson quasi-Lindley regression model. Mathematical Sciences, 13(3), 241-247.
2. Coleman, J.S. and James, J. (1961). The equilibrium size distribution of freely forming groups, *Sociometry*, 45, 24-36.
3. Gerstenkorn, T. (1996). A compound of the Polya distribution with the beta one. *Random Oper. and Stoch. Equ.*, 4(2), 103-110.
4. Gerstenkorn, T. (1993). A compound of the generalized gamma distribution with the exponential one. *Recherches surles deformations*, 16(1), 5-10.
5. Greenwood, M. and Yule, G.U. (1920). An inquiry into the nature of frequency distribution representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *J. Roy. Stat. Soc.*, 83, 255-279.
6. Gupta, R.C. and Ong, S.H. (2004). A new generalization of the negative binomial distribution. *Journal of Computational Statistics and Data Analysis*, 45, 287-300.
7. Hassan, H., Dar, S.A. and Ahmad, P.B. (2019). Poisson Ishita distribution: A new compounding probability model. *IOSR Journal of Engineering (IOSRJEN),* 9(2), 38-46.
8. James, J. (1953). The distribution of free forming small group size. *American Sociological Review*, 18, 569-570.
9. Lord, D. and Geedipally, S.R. (2011). The negative binomial–Lindley distribution as a tool for analyzing crash data characterized by a large amount of zeros. *Accident Analysis & Prevention*, 43(5), 1738-1742.
10. Panjer, H.H. and Willmot, G.E. (1981). Finite sum evaluation of the negative binomial-exponential model. *ASTIN Bulletin: The Journal of the IAA*, 12(2), 133-137.
11. R Core Team (2019). R version 3.5.3: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
12. Sankaran, M. (1970). The discrete Poisson-Lindley distribution. *Biometrics*, 26, 145-149.
13. Shanker, R. and Hagos, F. (2016). Size-biased Poisson-Sujatha distribution with applications. *American Journal of Mathematics and Statistics*, 6(4), 145-154.
14. Sing, S.N. and Yadav, R.C. (1971). Trends in rural out-migration at household level, *Rural Demography*, 8, 53-61.
15. Skellam, J.G. (1948). A probability distribution derived from the binomial distribution by regarding the probability of success as variable between the sets of trials. *Journal of the Royal Statistical Society*, Series B, 10, 257-261.
16. Zamani, H. and Ismail, N. (2010). Negative Binomial–Lindley Distribution and Its Application. *Journal of Mathematics and Statistics*, 1, 4-9.