International Journal of Applied Mathematics
& Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 6, Issue 4, Jun - Jul 2017; 1 - 18

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

FAMILY OF POISSON DISTRIBUTION AND ITS APPLICATION

AYESHA FAZAL[1] & SHAKILA BASHIR[2]

[1]Roll of Honor M-Phil Statistics, Kinnaird College for Women, Lahore, Pakistan

[2]Assistant Professor, Department of Statistics, Forman Christian College, A Chartered University, Lahore, Pakistan

ABSTRACT

The purpose of this paper is to introduce three discrete distributions named Poisson exponential distribution, Poisson size biased exponential distribution and size biased Poisson exponential distribution. These distributions apply to biological data sets, traffic datasets and thunderstorm datasets. These distributions are introduced with some of its basic properties including moments, coefficient of skewness and kurtosis are discussed. The method of moments and maximum likelihood estimation of the parameters of discrete PED, PSBED and SBPED are investigated. It is found that the reciprocal of MOM and MLE estimator is unbiased for the proposed distributions. Applications of the three models to different discrete data sets are compared with Poisson distribution, size biased Poisson distribution, size biased generalized Poisson distribution size biased geometric distribution and size biased Poisson lindley distribution to test their goodness of fit and the fit shows that the proposed distributions can be an important tool for modelling biological, traffic and other discrete data sets.

KEYWORDS: Goodness of Fit, Estimation of Parameters, Exponential Distribution, PED, Poisson Distribution, PSBED, SBPD Moments, Size-Biased Exponential Distribution, Size-Biased Poisson Distribution

INTRODUCTION

Poisson distribution was derived by French mathematician Simeon-Denis Poisson in 1830with probability density function (pdf)

$$P(X) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0,1,2 \ldots .... \qquad \mathbf{(1.1)}$$

Where $\lambda > 0$ is the shape parameter

The exponential distribution is the distribution of continuous random variables with probability density function (pdf)

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0 \qquad \mathbf{(1.2)}$$

Where $\lambda > 0$ is the scale parameter

Weighted distribution was first introduced by Fisher (1934) and was later formalized by Rao (1965). When the probability of selection of an event is not equal weights are assigned to the distribution to remove bias.

The pdf of weighted distribution is

$$f_w(x:\theta) = \frac{w(x)f_o(x;\theta)}{E[w(x)]} \qquad \mathbf{(1.3)}$$

When $w(x) = x^r$ The weighted distribution are called moment distributions. For r=1 and 2 moment distribution is named as size-biased and area-biased distributions simultaneously.

When w (x) =x, the weighted distribution is called size biased/length biased distribution

$$f(x;\theta) = \frac{x f_0(x;\theta)}{\mu'_1} \quad \textbf{(1.4)}$$

$$E(x) = \mu'_1$$

Weighted distribution has a number of applications in environmental sciences, including ecology and forestry and fields of real life including medicine, ecology and reliability.

The pdf of the size biased Poisson distribution is

$$g(x/\lambda) = \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!} \quad \textbf{(1.5)}$$

X=1, 2, 3… $\lambda > 0$

The pdf of the size biased exponential distribution is

$$f(x;\theta) = x\theta^2 e^{-x\theta} \quad \textbf{(1.6)}$$

X=1, 2, 3… $\theta > 0$

The mixture of two or more probability distributions is known as mixture distribution. The probability distribution is called a mixture distribution if its pdf is written in the form

$$P(X = x) = \int f(x/\lambda) g_\lambda(\lambda) d\lambda \quad \textbf{(1.7)}$$

Shakila Bashir and Mujahid Rasul (2016) introduced a discrete distribution named Poisson are a biased Lindley distribution and its application to biological data. Rama Shankar, Hagos Fasshaye, Abrehe Yemane (2015) applied the SBPLD to model thunderstorm data. They proved that SBPLD gives much closer fit than SBPD and gives a better alternative to the SBPD and can be recommended for modelling thunderstorm events.

Shankar and Mishra (2015) discussed a size-biased two parameter Poisson Lindley distribution(SBTPPLD) by size biasing the two parameter Poisson Lindley of Shankar and Mishra (2014)which has been obtained by mixing Poisson distribution with two parameter Lindley distribution introduced by Shankar and Mishra (2013). Bhatti (2015) derived the new generalized Poisson Lindley distribution.

Porinita Dutta and Muhindra Borah (2014) further investigated some properties and applications of SBPLD, such as moments, cummulants, co-efficient of variation, harmonic mean, survival function, estimation of parameters by the method of moments etc. Khurshid Ahmed Mir and Munir Ahmed (2009) introduced size-biased discrete distribution, generalized size biased discrete distribution and their generalization. Ghitani et al (2008) obtained the size-biased Poisson Lindley distribution and its properties. Shaban (1981) derived the Poisson-inverse Gaussian distribution, by mixing the Poisson with the inverse Gaussian distribution. Bulmer (1974) discussed the Poisson lognormal distribution by mixing the Poisson distribution with lognormal distribution. Sankaran (1970) derived poison-Lindley distribution by mixing Poisson with Lindley distribution. He discussed estimation of parameters and applied data on proposed distribution. He proved that Poisson gives a better fit as compared to some other distributions.

In this paper, we consider the Poisson-exponential distribution PED, PSBED and SBPED which is obtained by integration and discretization of the mixture of Poisson and exponential distribution, Poisson and Size-biased exponential distribution and Size-biased Poisson and Size-biased exponential distribution.

1. POISSON EXPONENTIAL DISTRIBUTION

The Poisson exponential distribution (PED) arises from the Poisson distribution with pdf (1.1)

$$f(x/\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0,1,2, \dots$$

When its parameter λ follows the exponential distribution with pdf (1.2)

$$g(\lambda) = \theta e^{-\theta\lambda} \qquad \lambda = 0,1,2, \dots$$

So

$$P(X = x) = \int f(x/\lambda) g_\lambda(\lambda) d\lambda$$

$$\text{P(X=x)} = \int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\, \theta e^{-\theta\lambda}\, d\lambda$$

After simplification the pdf of the PED is

$$\text{P (X=x)} = \frac{\theta}{(1+\theta)^{x+1}}, \qquad \theta > 0, x = 0,1,2, \dots \dots \qquad \textbf{(2.1)}$$

PROPERTIES OF POISSON EXPONENTIAL DISTRIBUTION

The rth factorial moment of the PED is

$$\mu'_{(r)} = \frac{r\Gamma r}{\theta^r} \qquad \textbf{(2.2)}$$

For r=1, 2, 3& 4 the first four factorial moments of PED are

$$\mu'_{(1)} = \frac{1}{\theta}, \mu'_{(2)} = \frac{2}{\theta^2}, \mu'_{(3)} = \frac{6}{\theta^3}, \mu'_{(4)} = \frac{24}{\theta^4}$$

Therefore the first four raw moments of PED are

$$\mu'_1 = \frac{1}{\theta} = \bar{X}, \mu'_2 = \frac{2+\theta}{\theta^2}, \mu'_3 = \frac{\theta^2+6\theta+6}{\theta^3}, \mu'_4 = \frac{\theta^3+14\theta^2+36\theta+24}{\theta^4}$$

The mean moments of PED are

$$\mu_2 = \frac{1+\theta}{\theta^2}$$

$$\mu_3 = \frac{(\theta+2)(\theta+1)}{\theta^3}$$

$$\mu_4 = \frac{\theta^3+10\theta^2+18\theta+9}{\theta^4}$$

The co-efficient of skewness and kurtosis of PED are

$$\beta_1 = \frac{(\theta+2)^2}{\theta+1} \qquad \textbf{(2.3)}$$

$$\beta_2 = \frac{\theta^3+10\theta^2+18\theta+9}{\theta^2+2\theta+1} \quad \textbf{(2.4)}$$

The moment ratios for PED show that for different values of θ>0, β_1 increases and

For θ=0, β2=9 and as θ⟶∞ β2 is increasing.

The moment generating function of the PED is

$$\mathrm{M_x(t)} = \frac{\theta}{1+\theta-e^t} \quad \textbf{(2.5)}$$

The Cummulant generating function and the four cummulants of PED are

$$K_1 = \frac{1}{\theta}, K_2 = \frac{1+\theta}{\theta^2}, K_3 = \frac{(\theta+2)(\theta+1)}{\theta^3}, K_4 = \frac{\theta^3+7\theta^2+12\theta+6}{\theta^4}$$

The probability generating function of the PED is

$$g(t) = \frac{\theta}{1+\theta-t} \quad \textbf{(2.6)}$$

The Fisher information matrix of the PED is

$$F = \frac{1}{\theta^2(1+\theta)} \quad \textbf{(2.7)}$$

Some more properties of PED are

$$\mu = \sigma^2 - \frac{1}{\theta^2}$$

It follows that for all values of $\theta > 0$ the poisson-exponential distribution is over-dispersed unless for large values of θ.

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \frac{\{\theta|(1+\theta)^{x+2}\}}{\{\theta|(1+\theta)^{x+1}\}}$$

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \frac{1}{1+\theta}$$

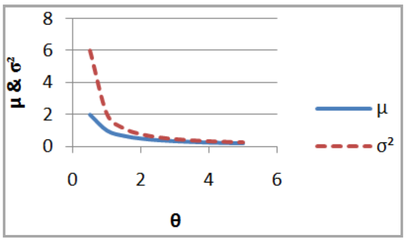

Figure 1.1: Graph of μ & σ² of PED

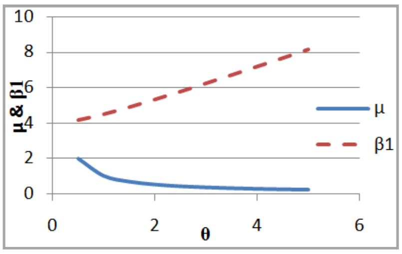

Figure 1.2: Graph of μ & β_1 of PED

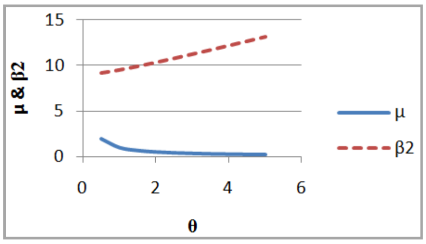

Figure 1.3: Graph of μ & β_2 of PED

- Figure (1.1) shows the trend of mean and variance of PED. As θ increases the mean and variance of PED decreases, so μ and σ^2 have the same trend.
- Figure (1.2) shows the trend of mean and coefficient of skewness of PED. As θ increases the mean decreases and coefficient of skewness increases.
- Figure (1.3) shows the trend of mean and coefficient of kurtosis of PED. As θ increases mean of PED decreases and coefficient of kurtosis for PED increases.
- For the PED model $(\beta_1, \beta_2) \rightarrow (4, 9)$ as $\theta \rightarrow 0$
- For the PED model the moment ratios increase as θ increases.

ESTIMATION OF PARAMETERS

In this section the parameter of the PED is estimated by MOM, MLE and properties of the estimates are derived.

Method of Moments (MOM)

The method of moment estimate $\tilde{\theta}$of the parameter θ is given by

$$\tilde{\theta} = \frac{1}{\bar{x}}$$

Maximum Likelihood Estimates (MLE)

The maximum likelihood estimate $\tilde{\theta}$ of the parameter θ is given by

$$\tilde{\theta} = \frac{1}{\bar{x}}$$

Theorem 1

If X follows PED with pdf $P(x) = \frac{\theta}{(\theta+1)^{x+1}}$ the MOM and MLE estimator $\frac{1}{\tilde{\theta}}$ is unbiased for $\frac{1}{\theta}$ and as $n \rightarrow \infty$ $var\left(\frac{1}{\tilde{\theta}}\right) \rightarrow 0$

Proof

The pdf of the Poisson - exponential distribution is

$$P(x) = \frac{\theta}{(\theta+1)^{x+1}}$$

$$\tilde{\theta} = \frac{1}{\bar{X}} \; gives \; \frac{1}{\tilde{\theta}} = \bar{X}(A)$$

Applying expectation on (A)

$E\left(\frac{1}{\bar{\theta}}\right) = E(\bar{X})$

After simplifications

$E\left(\frac{1}{\bar{\theta}}\right) = E(X)$

Since the mean of the PED is $\frac{1}{\theta}$

$$E\left(\frac{1}{\bar{\theta}}\right) = \frac{1}{\theta} \tag{2.8}$$

Which proves that $\frac{1}{\bar{\theta}}$ is unbiased for $\frac{1}{\theta}$ i.e the reciprocal of the estimator is unbiased.

$var\left(\frac{1}{\bar{\theta}}\right) = var(\bar{X})$

After simplification

$var\left(\frac{1}{\bar{\theta}}\right) = \frac{1}{n} var(X) \qquad (B)$

Variance of the PED is var (X) =$\frac{1+\theta}{\theta^2}$, putting this expression in equation (B) on RHS

$$var\left(\frac{1}{\bar{\theta}}\right) = \frac{1+\theta}{n\theta^2} \tag{2.9}$$

Hence $\lim_{n\to\infty} var\left(\frac{1}{\bar{\theta}}\right) = 0$

Which proves that the variance of the PED is minimum as $n \to \infty$

2. POISSON SIZE – BIASED EXPONENTIAL DISTRIBUTION

The Poisson size-biased exponential distribution arises from the Poisson distribution with pdf (1.1)

$f(x/\lambda) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad x = 0,1,2,...$

When the parameter λ follows the size-biased exponential distribution (SBED) with pdf (1.6)

$f(x;\theta) = x\theta^2 e^{-x\theta}$

So

$P(X = x) = \int f(x/\lambda) g_\lambda(\lambda) d\lambda$

$P(X = x) = \int_0^\infty \left(\frac{e^{-\lambda}\lambda^x}{x!}\right) \lambda\theta^2 \, e^{-\lambda\theta} d\lambda$

After simplifications the pdf of PSBED is

$$P(X = x) = \frac{\theta^2}{(1+\theta)^{x+2}}(x + 1) \tag{3.1}$$

θ>0, x=0, 1, 2, 3……..

PROPOERTIES OF POISSON SIZE - BIASED EXPONENTIAL DISTRIBUTION

The rth factorial moment of PSBED is

$$\mu'_{(r)} = \frac{\Gamma r+2}{\theta^r} \tag{3.2}$$

For r=1, 2, 3 & 4 the first four factorial moments of PSBED are

$$\mu'_{(1)} = \frac{2}{\theta},\ \mu'_{(2)} = \frac{6}{\theta^2},\ \mu'_{(3)} = \frac{24}{\theta^3},\ \mu'_{(4)} = \frac{120}{\theta^4}$$

Therefore the first four raw moments of PSBED are

$$\mu'_1 = \frac{2}{\theta},\ \mu'_2 = \frac{6+2\theta}{\theta^2},\ \mu'_3 = \frac{24+18\theta+2\theta^2}{\theta^3}\ ,\ \mu'_4 = \frac{2\theta^3+42\theta^2+144\theta+120}{\theta^4}$$

The mean moments of PSBED are

$$\mu_2 = \frac{2\theta+2}{\theta^2}$$

$$\mu_3 = \frac{2\theta^2+6\theta+28}{\theta^3}$$

$$\mu_4 = \frac{2\theta^3+26\theta^2+48\theta-72}{\theta^4}$$

The co-efficient of skewness and kurtosis of PSBED are

$$\beta_1 = \frac{4\theta^4+24\theta^3+148\theta^2+336\theta+784}{8\theta^3+24\theta^2+24\theta+8} \tag{3.3}$$

$$\beta_2 = \frac{\theta^3+13\theta^2+24\theta-36}{2\theta^2+4\theta+2} \tag{3.4}$$

The moment ratios for PSBED show that for different values of $\theta>0$, β_1 increases and for $\theta=0$, $\beta_2=98$ and as $\theta\to\infty$ β_2 is increasing

The moment generating function of PSBED is

$$M_x(t) = \frac{\theta^2}{\theta^2+2\theta-(2\theta+2)e^t+e^{2t}+1} \tag{3.5}$$

The probability generating function of PSBED is

$$g(t) = \frac{\theta^2}{\theta^2+2\theta+t^2-2t(\theta+1)+1} \tag{3.6}$$

The Fisher information matrix of PSBED is

$$F = \frac{2}{\theta^2(1+\theta)} \tag{3.7}$$

Some more properties of PSBED are

$$\mu = \sigma^2 - \frac{2}{\theta^2}$$

It follows that for all values of $\theta > 0$ the poisson size biased exponential distribution is over-dispersed unless for large values of θ.

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \frac{\theta^2(2x+1)/(\theta+1)^{x+3}}{\theta^2(x+1)/(\theta+1)^{x+2}}$$

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \frac{(2+{}^1/_x)}{(1+{}^1/_x)(\theta+1)}$$

This shows that $\frac{f(x+1;\theta)}{f(x;\theta)}$ is a decreasing function in x, that is f(x;θ) is logarithmically concave function. Therefore PSBED is unimodal.

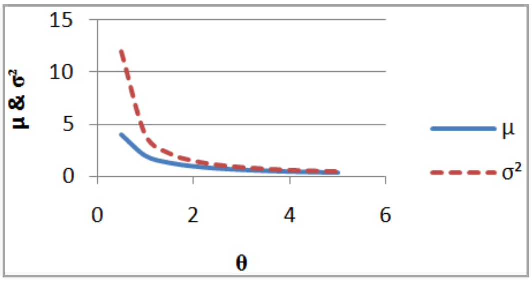

Figure 2.1: Graph of μ & σ^2 of PSBED

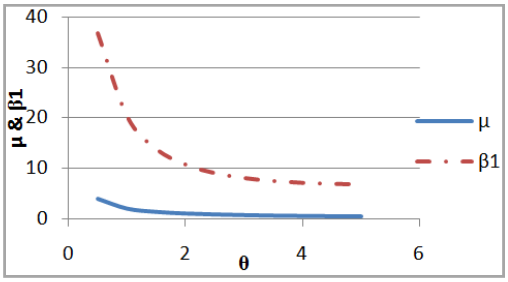

Figure 2.2: Graph of μ & β_1 of PSBED

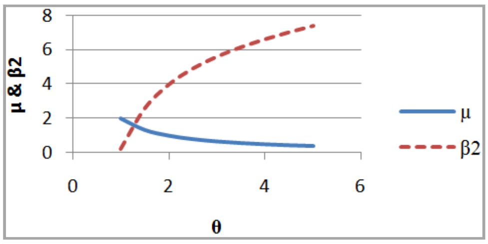

Figure 2.3: Graph of μ & β_2 of PSBED

- Figure (2.1) shows the trend of mean and variance of PSBED. As θ increases the mean and variance of PSBED decreases, so μ and σ^2 have the same trend.
- Figure (2.2) shows the trend of mean and coefficient of skewness of PSBED. As θ increases the mean decreases and coefficient of skewness also decreases, therefore, mean and coefficient of skewness has the same trend.
- Figure (2.3) shows the trend of mean and coefficient of kurtosis of PSBED. As θ increases mean of PSBED decreases and coefficient of kurtosis for PSBED increases.
- For the PSBED model $(\beta_1, \beta_2) \rightarrow (98, -18)$ as $\theta \rightarrow 0$

ESTIMATION OF PARAMETERS

In this section the parameter of PSBED is estimated by MOM, MLE and properties of the estimates are derived.

Method of Moments (MOM)

The method of moment estimate $\tilde{\theta}$ of the parameter θ is given by

$$\tilde{\theta} = \frac{2}{\bar{x}}$$

Maximum Likelihood Estimates (MLE)

The maximum likelihood estimate $\tilde{\theta}$ of the parameter θ is given by

$$\tilde{\theta} = \frac{2}{\bar{x}}$$

Theorem 2

If X follows PSBED with pdf $P(x) = \frac{\theta^2(x+1)}{(\theta+1)^{x+2}}$ the MOM and MLE estimator $\frac{1}{\tilde{\theta}}$ is unbiased for $\frac{1}{\theta}$ and as $n \to \infty$ $var\left(\frac{1}{\tilde{\theta}}\right) \to 0$

Proof

The pdf of Poisson size biased exponential distribution is

$$P(x) = \frac{\theta^2(x+1)}{(\theta+1)^{x+2}}$$

$$\tilde{\theta} = \frac{2}{\bar{X}} \; gives \; \frac{1}{\tilde{\theta}} = \frac{\bar{X}}{2} (A)$$

Applying expectation on (A)

$$E\left(\frac{1}{\tilde{\theta}}\right) = E\left(\frac{\bar{X}}{2}\right)$$

After simplifications

$$E\left(\frac{1}{\tilde{\theta}}\right) = \frac{1}{2} E(X)$$

Since mean of PSBED is $\frac{2}{\theta}$ Therefore, putting this value in equation above

$$E\left(\frac{1}{\tilde{\theta}}\right) = \frac{1}{\theta} \quad \textbf{(3.8)}$$

Which proves that $\frac{1}{\tilde{\theta}}$ is unbiased for $\frac{1}{\theta}$ i.e the reciprocal of the estimator is unbiased.

$$var\left(\frac{1}{\tilde{\theta}}\right) = var\left(\frac{\bar{X}}{2}\right)$$

After simplification

$$var\left(\frac{1}{\tilde{\theta}}\right) = \frac{1}{4n} var(X) \qquad (B)$$

Variance of PSBED is var (X)=$\frac{2\theta+2}{\theta^2}$, putting this expression in equation (B) on RHS

$$var\left(\frac{1}{\bar{\theta}}\right)=\frac{(1+\theta)}{2n\theta^2} \quad \textbf{(3.9)}$$

Hence $\lim_{n\to\infty} var\left(\frac{1}{\bar{\theta}}\right)=0$

Which proves that variance of PSBED is minimum as $n\to\infty$

3. SIZE – BIASED POISSON EXPONENTIAL DISTRIBUTION

The size-biased Poisson exponential distribution (SBPED) arises by giving weights, W(x) =x to Poisson exponential distribution, Using the pdf of the weighted distribution from (1.4)

$$f(x;\theta)=\frac{xf_0(x;\theta)}{\mu'_1}\ with\ E(x)=\mu'_1$$

The pdf of size-biased Poisson exponential distribution is

$$f_w(x;\theta)=\frac{x\theta^2}{(1+\theta)^{x+1}}$$

$\theta>0$, X=1, 2, 3…∞

Above is the pdf of SBPED which can be derived by integrating mixture of size biased Poisson and size biased exponential distribution from (1.5) and (1.6) also

The pdf of size-biased Poisson distribution (SBPD) is

$$g(x/\lambda)=\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}$$

When its parameter λ follows a size-biased exponential (SBED) model with pdf

$$h(\lambda,\theta)=\theta^2\lambda e^{-\theta\lambda}$$

then

$$P(x)=\int_0^\infty g(x/\lambda)h(\lambda,\theta)d\lambda$$

$$=\int_0^\infty \frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\theta^2\lambda e^{-\theta\lambda}d\lambda$$

After simplifications the pdf of size-biased Poisson exponential distribution is

$$P(x)=\frac{x\theta^2}{(1+\theta)^{x+1}} \quad \textbf{(4.1)}$$

$\theta>0$,X=1, 2, 3,…….∞

Therefore, when Poisson exponential distribution was given weights its pdf is the same when we made a mixture of SBPD and SBED.

PROPERTIS OF SIZE – BIASED POISSON EXPONENTIAL DISTRIBUTION

The first four raw moments of size-biased Poisson exponential distribution are

$$\mu'_1 = \frac{\theta+2}{\theta},\ \mu'_2 = \frac{\theta^2+6\theta+6\theta}{\theta^2},\ \mu'_3 = \frac{\theta^3+14\theta^2+36\theta+24}{\theta^3},\ \mu'_4 = \frac{\theta^4+30\theta^3+150\theta^2+240\theta+120}{\theta^4}$$

The mean moments of PSBED are

$$\mu_2 = \frac{2\theta+2}{\theta^2}$$

$$\mu_3 = \frac{2\theta^2+36\theta+4}{\theta^3}$$

$$\mu_4 = \frac{24+48\theta+86\theta^2-58\theta^3}{\theta^4}$$

The co-efficient of skewness and kurtosis of SBPED are

$$\beta_1 = \frac{4\theta^4+144\theta^3+1312\theta^2+288\theta+16}{8\theta^3+24\theta^2+96\theta+8} \quad \textbf{(4.2)}$$

$$\beta_2 = \frac{12+24\theta+43\theta^2-29\theta^3}{2\theta^2+4\theta+2} \quad \textbf{(4.3)}$$

The moment ratios for SBPED show that for different values of $\theta>0$, β_1 increases and for $\theta=0$, $\beta_2=6$

The moment generating function of SBPED is

$$M_x(t) = \frac{\theta^2 e^t}{\theta^2+2\theta-(2\theta+2)e^t+e^{2t}+1} \quad \textbf{(4.4)}$$

The probability generating function of SBPED is

$$g(t) = \frac{\theta^2 t}{\theta^2+2\theta+t^2-2t(\theta+1)+1} \quad \textbf{(4.5)}$$

The Fisher information matrix of SBPED is

$$F = \frac{2}{\theta^2(1+\theta)} \quad \textbf{(4.6)}$$

Some more properties of SBPED are

$$\mu - \sigma^2 = 1 + \frac{2}{\theta} - \frac{4}{\theta^2} - \frac{8}{\theta^3} - \frac{4}{\theta^4}$$

It follows that for all values of $\theta > 0$ the size-biased Poisson exponential distribution is over-dispersed unless for large values of θ.

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \frac{\theta^2(x+1)}{(1+\theta)^{x+1+1}}$$

$$\frac{f(x+1;\theta)}{f(x;\theta)} = \left(1 + {}^{1}\!/_{x}\right)\left(\frac{1}{1+\theta}\right)$$

The result above shows that $\frac{f(x+1;\theta)}{f(x;\theta)}$ is a decreasing function in x, therefore PSBED is unimodal.

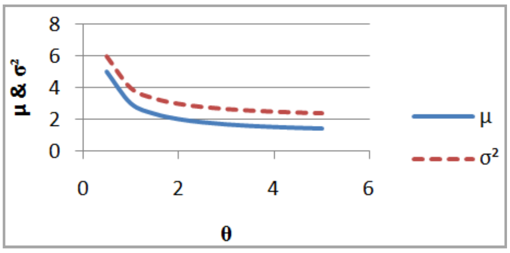

Figure 3.1: Graph of μ & σ^2 of SBPED

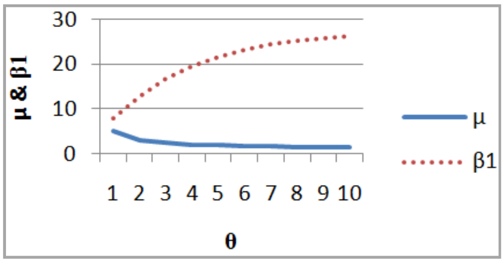

Figure 3.2: Graph of μ & β-1 of SBPED

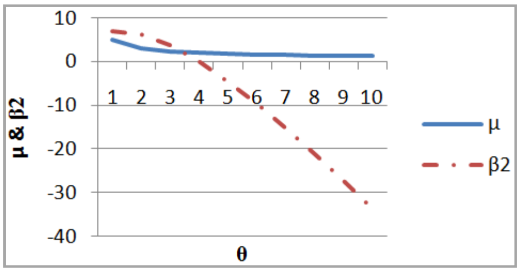

Figure 3.3: Graph of μ & β-2 of SBPED

- Figure (3.1) shows the trend of mean and variance of SBPED. As θ increases the mean and variance of SBPED decreases, so μ and σ^2 have the same trend.
- Figure (3.2) shows the trend of mean and coefficient of skewness of SBPED. As θ increases the mean decreases and coefficient of skewness increases.
- Figure (3.3) shows the trend of mean and coefficient of kurtosis of SBPED. As θ increases mean of SBPED decreases and coefficient of kurtosis for SBPED also decreases.
- For the SBPED model $(\beta_1, \beta_2) \rightarrow (2, 6)$ as $\theta \rightarrow 0$

ESTIMATION OF PARAMETERS

In this section the parameter of SBPED is estimated by MOM, MLE and properties of the estimates are derived.

Method of Moments (MOM)

The method of moment estimate $\tilde{\theta}$ of the parameter θ is given by

$$\tilde{\theta} = \frac{2}{\bar{x}-1}$$

Maximum Likelihood Estimates (MLE)

The maximum likelihood estimate $\tilde{\theta}$ of the parameter θ is given by

$$\tilde{\theta} = \frac{2}{\bar{x}-1}$$

Theorem 3

If X follows SBPED with pdf $P(x) = \frac{x\theta^2}{(\theta+1)^{x+1}}$ the MOM and MLE estimator $\frac{1}{\tilde{\theta}}$ is unbiased for $\frac{1}{\theta}$ and as $n \rightarrow \infty$ $var\left(\frac{1}{\tilde{\theta}}\right) \rightarrow 0$

Proof

The pdf of size biased Poisson exponential distribution is

$$P(x) = \frac{x\theta^2}{(\theta+1)^{x+1}}$$

$$\tilde{\theta} = \frac{2}{\bar{x}-1} \quad gives \; \frac{1}{\tilde{\theta}} = \frac{\bar{x}-1}{2} (A)$$

Applying expectation on (A)

$$E\left(\frac{1}{\tilde{\theta}}\right) = E\left(\frac{\bar{x}-1}{2}\right)$$

After simplifications

$$E\left(\frac{1}{\tilde{\theta}}\right) = \frac{1}{2}[E(x) - 1]$$

The mean of SBPED is $\mu = \frac{\theta+2}{\theta}$ Putting in the equation above

After simplifications

$$E\left(\frac{1}{\tilde{\theta}}\right) = \frac{1}{\theta} \quad \textbf{(4.6)}$$

Which proves that $\frac{1}{\tilde{\theta}}$ is unbiased for $\frac{1}{\theta}$ i.e the reciprocal of the estimator is unbiased.

$$var\left(\frac{1}{\tilde{\theta}}\right) = var\left(\frac{\bar{x}-1}{2}\right)$$

$$var\left(\frac{1}{\tilde{\theta}}\right) = \frac{1}{4n} var(x) \qquad (B)$$

Variance of SBPE distribution is var(X)=$\frac{2\theta+2}{\theta^2}$, putting this expression in equation (B) on RHS

$$var\left(\frac{1}{\tilde{\theta}}\right) = \frac{(1+\theta)}{2n\theta^2} \quad \textbf{(4.7)}$$

Hence $\lim_{n\rightarrow\infty} var\left(\frac{1}{\tilde{\theta}}\right) = 0$

Which proves that variance of SBPED is minimum as $n \rightarrow \infty$

Table 1

Measures	PD	PED	PSBED	SBPED
μ	λ	$^{1}/_{\theta}$	$^{2}/_{\theta}$	$\frac{\theta+2}{\theta}$
σ^2	λ	$\frac{1+\theta}{\theta^2}$	$\frac{2\theta+2}{\theta^2}$	$\frac{2\theta+2}{\theta^2}$
$\sqrt{\beta_1}$	$^{1}/_{\sqrt{\lambda}}$	$\sqrt{\frac{(\theta+2)^2}{\theta+1}}$	$\sqrt{\frac{\theta^4+6\theta^3+37\theta^2+84\theta+196}{2\theta^3+6\theta^2+6\theta+2}}$	$\frac{4\theta^4+144\theta^3+1312\theta^2+288\theta+16}{8\theta^3+24\theta^2+96\theta+8}$
β_2	$3+{}^{1}/_{\lambda}$	$\frac{\theta^3+10\theta^2+18\theta+9}{\theta^2+2\theta+1}$	$\frac{\theta^3+13\theta^2+24\theta-36}{2(\theta+1)^2}$	$\frac{24+48\theta+86\theta^2-58\theta^3}{4\theta^2+8\theta+4}$

The table above shows the mean (μ), variance (σ^2), coefficient of skewness ($\sqrt{\beta_1}$) and coefficient of kurtosis (β_2) for the Poisson distribution (PD), Poisson exponential distribution (PED), poisson size biased exponential distribution (PSBED) and size biased Poisson exponential distribution (SBPED).

APPLICATIONS

To illustrate the applications and to justify the suitability of the three distributions i.e: Poisson exponential distribution, Poisson size biased exponential distribution and size biased Poisson exponential distribution in practical applications, we are fitting these distributions to some datasets.

GOODNESS OF FIT FOR POISSON EXPONENTIAL DISTRIBUTION (PED) AND POISSON SIZE –BIASED EXPONENTIAL DISTRIBUTION (PSBED)

In this section we have applied Poisson exponential distribution (PED) and Poisson size biased exponential distribution (PSBED) to biological data and traffic datasets using the MOM and MLE estimates and compared with Poisson distribution (PD).

A: Shankar, et al (2015) gave data on Observed and expected number of Hemocytometer yeast cell counts per square observed by 'student' 1907.

Form table-2 it can be seen that PED and PSBED give much closer fit than the PD to the yeast cell data set, therefore PED and PSBED are suitable for modelling data in ecology.

B: Shankar, et al (2015) gave data on Observed and expected number of red mites on apple leaves.

From table-3 it can be seen that PED and PSBED gives a better fit than PD, therefore PED and PSBED are can be considered as an important tool for modelling datasets in ecology.

C: Gerlough, D.L (1955) studied the Analysis and comparison of accident data. He gave data on the observed and expected number of car accidents.

From table-4 it can be seen that PED and PSBED seems to be a good replacement of PD to model traffic data sets

Table 2: Goodness of Fit for PD, PED and PSBED on Number of Hemocytometer Yeast Cell Counts

Number of Cells Per Square	Observed Frequency	Expected Frequency		
		PD	PED	PSBED
0	128	118.1	128.0838	123.6145
1	37	54.3	40.35407	46.22106
2	183	12.5	12.71395	12.96199
3	3	1.9	4.005654	3.231105
4	1	0.2	1.262021	0.755095
5+	0	0.0	0.397612	0.169404
Total	**187**			
Estimation of parameters		$\tilde{\theta} = 0.459893$	$\tilde{\theta}$=2.174386	$\tilde{\theta}$=4.348838
χ^2		9.903	2.966108	3.38411
d.f		1	2	1
p-value		0.0016	0.226944	0.0658

Table 3: Goodness of Fit for PD, PED and PSBED on Number of Red Mites on Apple Leaves

Number of Mites Per Leaf	Observed Frequency	Expected Frequency		
		PD	PED	PSBED
0	38	25.3	37.2101	32.22676
1	17	29.1	19.90271	23.52404
2	10	16.7	10.64544	12.8786
3	9	6.4	5.693964	6.267186
4	3	1.8	3.045552	2.859222
5	2	0.4	1.628986	1.25226
6	1	0.2	0.871302	0.53322
7+	0	0.1	0.466036	0.222415
Total	**80**	**80**	**79.46**	**79.76**
Estimation of parameters		$\tilde{\theta}$=1.15	$\tilde{\theta}$=0.8696	1.739
χ^2		18.275	2.398823	4.82913
d.f		2	3	2
p-value		0.0001	0.4939	0.0894

Table 3: Goodness of Fit for PD, PED and PSBED on Number of Car Accidents

Number of Car Accidents X	Number of Roads Observed f_o	Expected Frequency		
		PD	PED	PSBED
0	18	14.5	21.09371	18.33412
1	14	16.4	11.20605	13.26296
2	7	9.3	5.953225	7.19585
3	4	3.5	3.162656	3.470334
4	1	1.0	1.680163	1.569032
5	0	0.2	0.892588	0.681026
6	0	0.04	0.474188	0.287383
7	1	0.007	0.251913	0.118796
8	0	0.001	0.133829	0.04834
Total	**45**	**45**	**44.8483**	**44.9678**
Estimation of parameters		$\tilde{\theta}$ =1.1333	$\tilde{\theta}$ =0.88235	$\tilde{\theta}$ =1.76471
χ^2	3.84	1.2922	1.3881	0.05733
d.f		1	2	2
P-value		0.25561	0.49955	0.9718

GOODNESS OF FIT FOR SIZE – BIASED POISSON EXPONENTIAL DISTRIBUTION

In this section we have tried to fit size-biased Poisson distribution (SBPD), size-biased geometric distribution (SBGD) and size-based Poisson exponential distribution (SBPED) to different datasets like thunderstorm data, number of workers N_i has I accidents and number of migrants aged 15 years and above (survey 1978 data) using the MOM and MLE estimates.

A: Shankar, et al (2015) gave data on Observed and expected frequency of thunderstorm events containing X thunderstorm at Cape Kennedy for August.

Table-4 shows that SBPED gives much closer fit than the SBPD and SBPLD to the thunderstorm dataset. Thus SBPED can be considered as an important tool for model thunderstorm data.

B: Mir, K A., Ahmed, M (2009) gave data on Observed and expected frequencies of the number of workers N_i having I accidents.

From table-5 it can be seen that SBPED gives much closer fit to model count datasets than SBPD and SBGPD, therefore SBPED provides a better alternative to SBPD and SBGPD to model count data sets.

C: Priti., Singh, B. P. (2015) gave data on Observed and expected number of households with at least one male migrant according to the number of male migrants aged 15-years and above (survey 1978 data) in the semi urban household.

Table-6 shows that SBPED gives a better fit than SBGD and can be considered as an important tool to model migration data sets.

Table 4: Goodness of Fit for SBPD, SBPLD and SBPEDs on Thunderstorm Event

X	Observed Frequency f_o	Expected Frequency f_e		
		SBPD	SBPLD	SBPED
1	201	194.2	199.5	199.663
2	60	68.3	59.9	59.68365
3	10	12.0	13.3	13.38056
4	3	1.4	2.6	2.666496
5	2	0.1	0.7	0.498171
Total	**276**	**276.0**	**276.0**	**275.8919**
Estimation of parameters		$\tilde{\theta} = 0.351449$	$\tilde{\theta} = 6.37477$	$\tilde{\theta} = 5.69071$
χ^2		1.414	0.165	0.15495
d.f		1	1	1
p-value		0.2344	0.6846	0.6939

Table 5: Goodness of Fit for SBPD, SBGPD and SBPED on Number of Workers Having Accidents

Number of Accidents	Number of Workers f_o	Expected Frequency f_e		
		SBPD	SBGPD	SBPED
1	2039	2034.27	2039.83	2039.078
2	312	319.48	309.76	311.2761
3	35	33.45	36.38	35.63846
4	3	2.63	3.66	3.626934
5	1	0.17	0.37	0.346044
Total	**2390**	**2390.0**	**2390.00**	**2389.965**
χ^2		0.772	0.069	0.0111245
d.f		1	1	1
p-value		0.379599	0.7928	0.916093

Table 6: Goodness of Fit for SBGD and SBPED on Number of Households with at Least One Male Migrant

Number of Migrants	Observed Frequency f_o	Expected Frequency f_e	
		SBGD	SBPED
1	95	86.54	86.52835
2	19	31.32	31.32305
3	10		8.50415
4	2	11.14	2.05232
5	3		0.464334
Total	**129**		**128.8722**
Estimation of parameters		$\tilde{\theta} = 0.819$	$\tilde{\theta} = 4.5249$
χ^2		7.01	3.1621
d.f		1	1
p-value		0.0082	0.075366

CONCLUSIONS

The PED, PSBED and SBPED are discrete distributions that are obtained by a mixture of Poisson and Exponential distribution (PED), a mixture of Poisson and Size biased exponential distribution (PSBED) and by applying weights on PED. Some important properties of the above distributions are derived. From Figure (1), Figure (2) and Figure (3) it can be seen that these distributions are positively skewed. The estimation of parameters is done by the method of moments (MOM) and maximum likelihood estimate (MLE) and it is proved that the reciprocal of MOM and MLE estimator is unbiased for PED, PSBED and SBPED. The discussion on estimation and applications of PED, PSBED and SBPED demonstrates that the above distributions give a better approach to the biological data sets, traffic datasets and is a good alternative to PD, SBPD, SBGD, SBGPD, SBPLD.

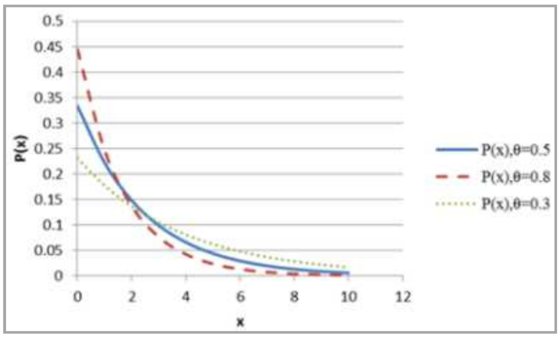

Figure 1: pdf Graph of PED

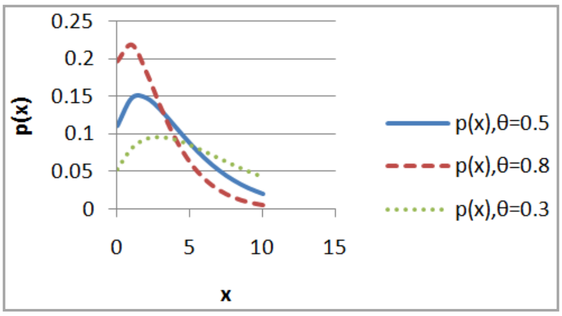

Figure 2: Graph of pdf PSBED

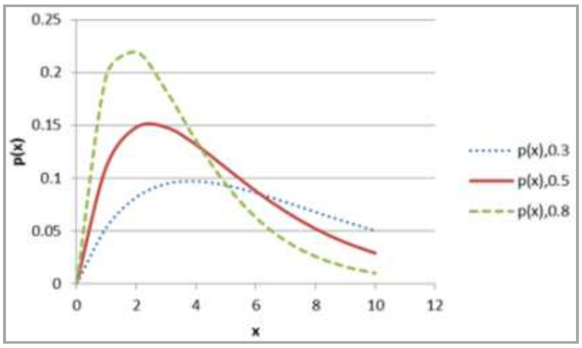

Figure 3: pdf Graph of SBPED

REFERENCES

1. Bashir. S., Rasul. M. (2016). Poisson Area-Biased Lindley Distribution and Its Applications on Biological Data. *Biometrics & Biostatistics International Journal, 3* (1), 1-10
2. Dutta, P., Borah, M. (2014). Some Properties and Application of Size Biased Poisson-Lindley Distribution. *International Journal of Mathematics Archive, 5* (1), 89-96.
3. Gerlough, D. L. (1955). *Use of Poisson Distribution in Highway Traffic. Analysis and Comparison of Accident Data:* Columbia University Press.
4. Ghitany, M. E., Al-Mutairi, D. K. (2008). Size-Biased Poisson-Lindley Distribution and Its Application. *International Journal of Statistics, LXVI* (3), 299-311.
5. Mir, K. A., Ahmed, M. (2009). Size Biased Distribution and there Applications. *Pakistan Journal of Statistics, 25*(3), 283-294.
6. Priti., Singh, B. P. (2015). A Size-Biased probability Distribution for the Number of Male Migrants. *Journal of Statistics Applications and Probability, An International Journal, 4*(3), 411-415.
7. Shanker, R., Fesshaye, H. (2015). On Poisson-Lindley Distribution and Its Applications to Biological Sciences. *Biometrics & Biostatistics International Journal, 2*(4), 1-5.
8. Shanker, R., Fesshaye, H., Yemane, A. (2015). On Size Biased Poisson-Lindley Distribution and Its Applications to Model Thunderstorms. *American Journal of Mathematics and Statistics, 5*(6), 354-360.