Journal of Modern Applied Statistical Methods

Volume 22 | Issue 2 Article 4

A Generalization of Generalized Poisson-Lindley Distribution and its Applications

Rama Shanker
Department of Statistics, Assam University Silchar, India,
shakerrama2009@gmail.com

Kamlesh Kumar Shukla
Department of Mathematics, Jaypee Institute of Information Technology, India,
kkshukla22@gmail.com

Received: 28-09-2023; Revised: 06-11-2023; Accepted: 11-12-2023; Published: 15-12-2023.

Recommended Citation

Rama Shanker, Kamlesh Kumar Shukla (2023). A Generalization of Generalized Poisson-Lindley Distribution and its Applications. Journal of Modern Applied Statistical Methods, 22(2), https://doi.org/10.56801/Jmasm.V22.i2.4

Journal of Modern Applied Statistical Methods
2023, Vol. 22, No. 2,
Doi: 10.56801/Jmasm.V22.i2.4
Copyright © 2023 JMASM
ISSN 1538 – 9472

A Generalization of Generalized Poisson-Lindley Distribution and its Applications

Rama Shanker
Department of Statistics, Assam University Silchar, India

Kamlesh Kumar Shukla
Department of Mathematics, Jaypee Institute of Information Technology, India

In this paper we proposed a generalization of generalized Poisson-Lindley distribution which includes generalized Poisson-Lindley distribution, Poisson-Lindley distribution, Poisson-weighted Lindley distribution, negative binomial distribution and geometric distribution as special cases. Statistical properties based on moments, maximum likelihood estimation and applications of the distribution have been discussed.

Keywords: Poisson-Lindley distribution, Generalized Poisson-Lindley distribution, Negative binomial distribution, Descriptive measures, Estimation, Goodness of fit.

1. Introduction

Sankaran (1970) introduced a one parameter Poisson-Lindley distribution (PLD) to model count data having probability mass function (pmf)

$$P_1(x;\theta)=\frac{\theta^2(x+\theta+2)}{(\theta+1)^{x+3}}\ ;x=0,1,2,...,\theta>0\ . \tag{1}$$

This distribution arises from the Poisson distribution when its parameter λ follows Lindley distribution introduced by Lindley (1958) with probability density function (pdf)

$$f_1(\lambda;\theta)=\frac{\theta^2}{\theta+1}(1+\lambda)e^{-\theta\lambda}\quad ;\lambda>0,\theta>0 \tag{2}$$

Ghitany et al (2008) studied various properties including shapes, moments, generating functions, hazard rate and mean residual life functions, stochastic ordering, order statistics, Renyi entropy measure, mean deviations, Bonferroni and Lorenz curves, and stress-strength reliability along with estimation of parameter using both the method of moments and the method of maximum likelihood estimation and application of Lindley distribution to model waiting time data in a bank. Ghitany et al (2008, 2009) have also proposed the size-biased and zero-truncated versions of PLD to model count data when the data to be modeled originate from a generating mechanism that structurally excludes zero counts.

Mahmoudi and Zakerzadeh (2010) introduced a two-parameter generalized Poisson-Lindley distribution (GPLD) having pmf

$$P_2\left(x;\theta,\alpha\right)=\frac{\Gamma\left(x+\alpha\right)}{\Gamma\left(x+1\right)}\frac{\theta^{\alpha+1}}{\Gamma\left(\alpha+1\right)}\frac{x+\alpha\,\theta+2\alpha}{\left(\theta+1\right)^{x+\alpha+2}}\ ;x=0,1,2,...,\theta>0,\alpha>0 \tag{3}$$

It can be easily verified that (1) is a particular case of (3) at $\alpha=1$. The first four moments about origin and the variance of GPLD are given by

$$\mu_1'=\frac{\alpha\left(\theta+1\right)+1}{\theta\left(\theta+1\right)}$$

$$\mu_2'=\frac{\alpha\,\theta^2+\left(\alpha^2+2\alpha+1\right)\theta+\left(\alpha^2+3\alpha+2\right)}{\theta^2\left(\theta+1\right)}$$

$$\mu_3'=\frac{\alpha\,\theta^3+\left(3\alpha^2+4\alpha+1\right)\theta^2+\left(\alpha^3+6\alpha^2+11\alpha+6\right)\theta+\left(\alpha^3+6\alpha^2+11\alpha+6\right)}{\theta^3\left(\theta+1\right)}$$

$$\mu_4'=\frac{\begin{Bmatrix}\alpha\,\theta^4+\left(7\alpha^2+8\alpha+1\right)\theta^3+\left(6\alpha^3+25\alpha^2+33\alpha+14\right)\theta^2+\left(\alpha^4+12\alpha^3+47\alpha^2+72\alpha+36\right)\theta\\+\left(\alpha^4+10\alpha^3+35\alpha^2+50\alpha+24\right)\end{Bmatrix}}{\theta^4\left(\theta+1\right)}$$

$$\mu_2=\frac{\alpha\,\theta^3+\left(3\alpha+1\right)\theta^2+3\left(\alpha+1\right)\theta+\left(\alpha+1\right)}{\theta^2\left(\theta+1\right)^2}.$$

It can be easily verified that at $\alpha=1$, these moments reduce to the corresponding moments of PLD.

Zakerzadeh and Dolati (2009) have introduced a three parameter generalized Lindley distribution (GLD) having pdf

$$f_2\left(x;\theta,\alpha,\beta\right)=\frac{\theta^{\alpha+1}}{\theta+\beta}\frac{x^{\alpha-1}}{\Gamma\left(\alpha+1\right)}\left(\alpha+\beta\,x\right)e^{-\theta x}\ \ ;x>0,\theta>0,\alpha>0,\beta>0 \tag{4}$$

Lindley distribution, gamma distribution and weighted Lindley distribution of Gitany et al (2011) are particular cases of (1.4) at $\left(\alpha=\beta=1\right)$, $\left(\beta=0\right)$ and $\left(\beta=\alpha\right)$, respectively. Note that GPLD in (3) is a Poisson mixture of GLD in (4) taking $\beta=1$. Shanker and Shukla (2016) have detailed comparative study on modeling of real lifetime data from engineering and biomedical sciences using GLD and generalized gamma distribution (GGD) introduced by Stacy (1962) and concluded that there are several lifetime data where GGD gives much better fit than GLD.

The pmf of Poisson-weighted Lindley distribution (P-WLD) introduced by El-Monsef and Sohsah (2014) is given by

$$P_3\left(x;\theta,\alpha\right)=\frac{\Gamma\left(x+\alpha\right)}{\Gamma\left(x+1\right)\Gamma\left(\alpha\right)}\frac{\theta^{\alpha+1}}{\left(\theta+\alpha\right)}\frac{x+\theta+\alpha+1}{\left(\theta+1\right)^{x+\alpha+1}};x=0,1,2,...,\theta>0,\alpha>0 \tag{5}$$

It can be easily verified that PLD is a particular case of P-WLD for $\alpha=1$. Shanker and Shukla (2017) have detailed study on P-WLD and obtained raw moments and central moments and studied the nature and behavior of coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion. Further, Shanker and Shukla (2017) discussed the maximum likelihood estimation for estimating its parameters and discussed the goodness of fit of P-WLD and concluded that P-WLD competes well with Poisson distribution (PD), PLD, NBD and GPLD. Note that P-WLD in (5) is a Poisson mixture of a two-parameter weighted Lindley distribution (WLD) proposed by Ghitany et al (2011) having pdf.

$$f_3\left(x;\theta,\alpha\right)=\frac{\theta^{\alpha+1}}{\left(\theta+\alpha\right)}\frac{x^{\alpha-1}}{\Gamma\left(\alpha\right)}\left(1+x\right)e^{-\theta x}\ ;x>0,\ \theta>0,\ \alpha>0 \tag{6}$$

Note that Lindley distribution is a particular case of WLD at $\alpha=1$. Shanker et al (2016) discussed various moments based properties including coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion of weighted Lindley distribution and its applications to model lifetime data from biomedical sciences and engineering. Shanker et al (2017) have proposed a three-parameter weighted Lindley distribution (TPWLD) which includes a two-parameter weighted Lindley distribution as particular cases and discussed its various structural properties, estimation of parameters and applications for modeling lifetime data from engineering and biomedical sciences.

The main motivation of this paper is to obtain a generalization of generalized Poisson-Lindley distribution (GGPLD) by taking a Poisson mixture of three-parameter GLD (4), which includes generalized Poisson-Lindley distribution (GPLD), Poisson-Lindley distribution (PLD), Poisson-weighted Lindley distribution (P-WLD), negative binomial distribution (NBD) and geometric distribution (GD) as particular cases. Its moments and moments associated measures have been derived and discussed. The estimation of its parameters has been discussed using maximum likelihood estimation. The applications of the distribution have been discussed with some count datasets and the goodness of fit of the distribution has also been compared with other discrete distributions.

2. A Generalization of Generalized Poisson-Lindley Distribution

Assuming that the parameter λ of the Poisson distribution follows a three-parameter GLD (4), the Poisson mixture of three-parameter GLD can be obtained as

$$P_3\left(x;\theta,\alpha,\beta\right)=\int_0^\infty \frac{e^{-\lambda}\lambda^x}{x!}\frac{\theta^{\alpha+1}}{\theta+\beta}\frac{\lambda^{\alpha-1}}{\Gamma\left(\alpha+1\right)}\left(\alpha+\beta\lambda\right)e^{-\theta\lambda}d\lambda \tag{7}$$

$$=\frac{\theta^{\alpha+1}}{\left(\theta+\beta\right)\Gamma\left(\alpha+1\right)x!}\left[\alpha\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x+\alpha-1}d\lambda+\beta\int_0^\infty e^{-(\theta+1)\lambda}\lambda^{x+\alpha+1-1}d\lambda\right]$$

$$=\frac{\theta^{\alpha+1}}{\left(\theta+\beta\right)\Gamma\left(\alpha+1\right)x!}\left[\frac{\alpha\Gamma\left(x+\alpha\right)}{\left(\theta+1\right)^{x+\alpha}}+\frac{\beta\Gamma\left(x+\alpha+1\right)}{\left(\theta+1\right)^{x+\alpha+1}}\right]$$

$$=\frac{\Gamma\left(x+\alpha\right)}{\Gamma\left(x+1\right)\Gamma\left(\alpha+1\right)}\frac{\theta^{\alpha+1}}{\left(\beta+\theta\right)}\frac{\beta x+\alpha\left(\theta+\beta+1\right)}{\left(\theta+1\right)^{x+\alpha+1}};x=0,1,2,...,\theta>0,\alpha>0,\beta>0 \tag{8}$$

We would call this distribution in (8) as the generalization of generalized Poisson-Lindley distribution (GGPLD). We denote it by GGPLD$\left(\theta,\alpha,\beta\right)$. It can be easily verified that GPLD in (4) and PLD in (1) are special cases of GGPLD for $\left(\beta=1\right)$, and $\left(\alpha=\beta=1\right)$ respectively. Further, the Poisson - Weighted Lindley distribution (P-WLD), negative binomial distribution (NBD) with parameters $r=\alpha$ and $p=\frac{\theta}{\theta+1}$ and the geometric distribution with parameter $p=\frac{\theta}{\theta+1}$ are also particular cases of (2.2) for $\left(\beta=\alpha\right),\left(\beta=0\right)$ and $\left(\alpha=1,\beta=0\right)$, respectively.

The nature and behavior of GGPLD for varying values of the parameters $\theta,\alpha,$ and β have been explained graphically in figure 1.

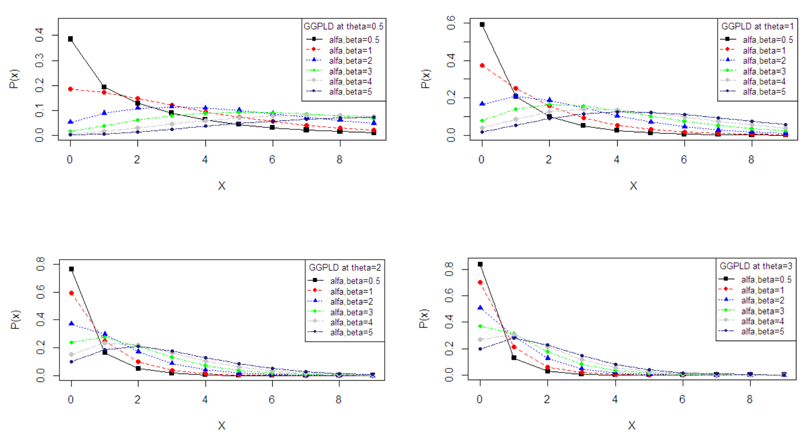

Figure 1. Probability mass function plot of GGPLD for varying values of parameters $\theta,\alpha,$ and β

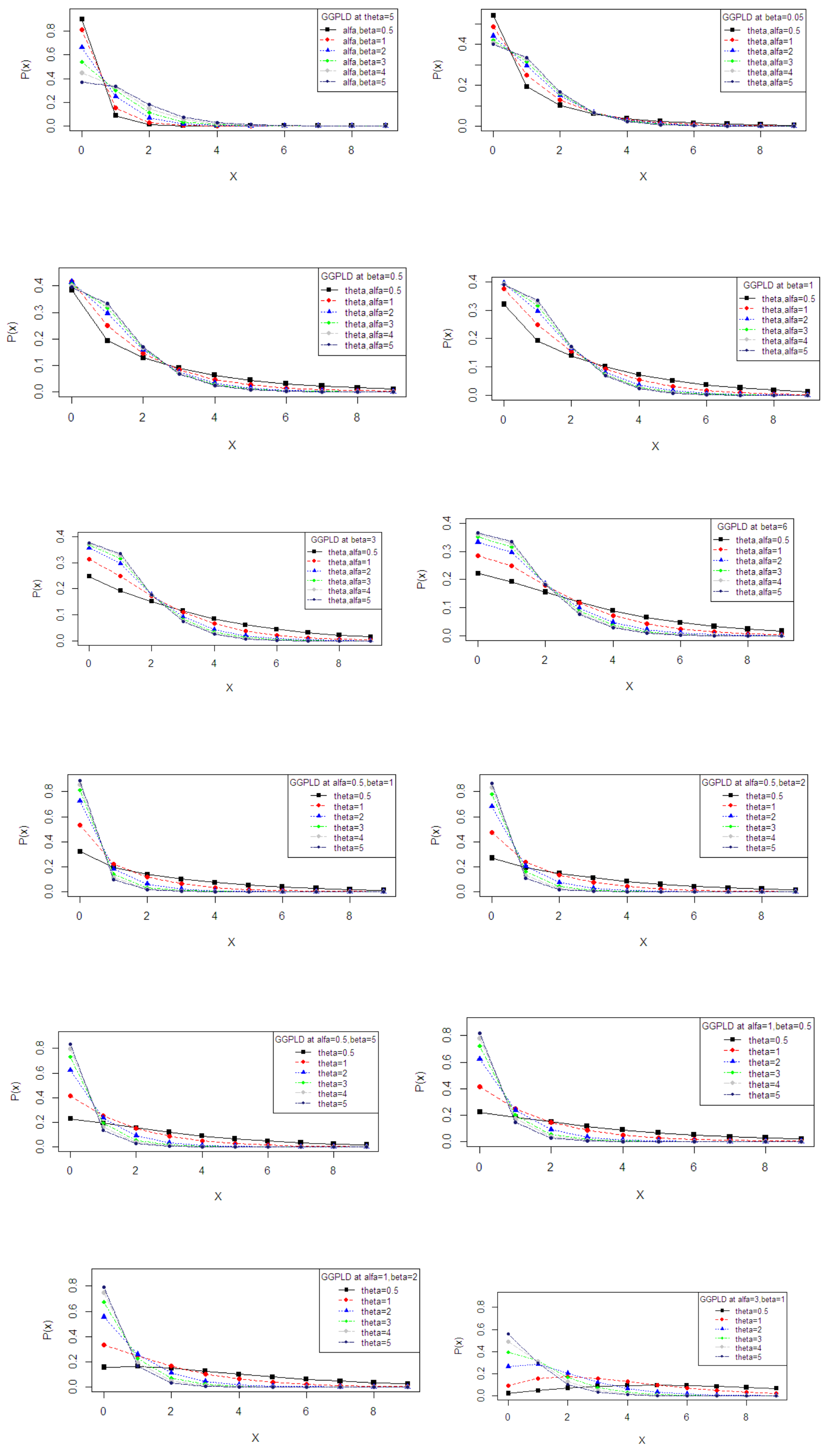

Figure 1. (Continues) Probability mass function plot of GGPLD for varying values of parameters $\theta, \alpha,$ and β

3. Moments and Associated Measures

Using (7), the r th factorial moment about origin of the GGPLD (8) can be obtained as

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

$$= \int_0^\infty \left[\sum_{x=1}^{\infty} x^{(r)} \frac{e^{-\lambda}\lambda^x}{x!}\right] \frac{\theta^{\alpha+1}}{(\theta+\beta)} \frac{\lambda^{\alpha-1}}{\Gamma(\alpha+1)} (\alpha+\beta\lambda) e^{-\theta\lambda} d\lambda$$

$$= \frac{\theta^{\alpha+1}}{(\theta+\beta)\Gamma(\alpha+1)} \int_0^\infty \left[\lambda^r \left\{\sum_{x=r}^{\infty} x \frac{e^{-\lambda}\lambda^{x-r}}{(x-r)!}\right\}\right] \lambda^{\alpha-1} (\alpha+\beta\lambda) e^{-\theta\lambda} d\lambda$$

Taking $x-r=y$, we get

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

$$= \frac{\theta^{\alpha+1}}{(\theta+\beta)\Gamma(\alpha+1)} \int_0^\infty \lambda^{\alpha+r-1} (\alpha+\beta\lambda) e^{-\theta\lambda} d\lambda$$

$$= \frac{\Gamma(\alpha+r)}{\Gamma(\alpha+1)} \frac{\left\{\alpha(\theta+\beta)+r\beta\right\}}{\theta^r(\theta+\beta)} \quad ; r=1,2,3,.... \tag{9}$$

Taking $r=1,2,3,\text{and}\,4$ in (9), the first four factorial moments about origin of GGPLD (8) can be obtained

$$\mu_{(1)}' = \frac{\alpha(\theta+\beta)+\beta}{\theta(\theta+\beta)}$$

$$\mu_{(2)}' = \frac{(\alpha+1)\left\{\alpha(\theta+\beta)+2\beta\right\}}{\theta^2(\theta+\beta)}$$

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

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

Now using the relationship between factorial moments about origin and the moments about origin, the first four moments about origin of GGPLD (8) can be obtained as

$$\mu_1' = \frac{\alpha(\theta+\beta)+\beta}{\theta(\theta+\beta)}$$

$$\mu_2' = \frac{\alpha\theta^2 + \left(\alpha^2 + \alpha\beta + \alpha + \beta\right)\theta + \left(\alpha^2\beta + 3\alpha\beta + 2\beta\right)}{\theta^2\left(\theta+\beta\right)}$$

$$\mu_3' = \frac{\left\{\begin{matrix}\alpha\theta^3 + \left(3\alpha^2 + \alpha\beta + 3\alpha + \beta\right)\theta^2 + \left(\alpha^3 + 3\alpha^2\beta + 3\alpha^2 + 9\alpha\beta + 2\alpha + 6\beta\right)\theta \\ + \left(\alpha^3\beta + 6\alpha^2\beta + 11\alpha\beta + 6\beta\right)\end{matrix}\right\}}{\theta^3\left(\theta+\beta\right)}$$

$$\mu_4' = \frac{\left\{\begin{matrix}\alpha\theta^4 + \left(7\alpha^2 + \alpha\beta + 7\alpha + \beta\right)\theta^3 + \left(6\alpha^3 + 7\alpha^2\beta + 18\alpha^2 + 21\alpha\beta + 12\alpha + 14\beta\right)\theta^2 \\ + \left(\alpha^4 + 6\alpha^3 + 6\alpha^3\beta + 11\alpha^2 + 36\alpha^2\beta + 66\alpha\beta + 6\alpha + 36\beta\right)\theta \\ + \left(\alpha^4\beta + 10\alpha^3\beta + 35\alpha^2\beta + 50\alpha\beta + 24\beta\right)\end{matrix}\right\}}{\theta^4\left(\theta+\beta\right)}$$

Now, using the relationship between moments about mean and the moments about origin, the moments about mean of the GGPLD (8) can be obtained as

$$\mu_2 = \frac{\alpha\theta^3 + \left(\alpha + 2\alpha\beta + \beta\right)\theta^2 + \left(\alpha\beta + \beta + 2\alpha + 2\right)\beta\theta + \left(\alpha+1\right)\beta^2}{\theta^2\left(\theta+\beta\right)^2}$$

$$\mu_3 = \frac{\left\{\begin{matrix}\alpha\theta^5 + 2\left(3\alpha + 3\alpha\beta + \beta\right)\theta^4 + \left(3\alpha\beta^2 + 2\beta^2 + 9\alpha\beta + 2\alpha + 6\beta\right)\theta^3 \\ + \left(\alpha\beta^2 + 9\alpha\beta + \beta^2 + 9\beta + 6\alpha + 6\right)\beta\theta^2 + \left(3\alpha\beta + 3\beta + 6\alpha + 6\right)\beta^2\theta \\ +2\left(\alpha+1\right)\beta^3\end{matrix}\right\}}{\theta^3\left(\theta+\beta\right)^3}$$

$$\mu_4 = \frac{\left\{\begin{matrix}\alpha\theta^7 + \left(3\alpha^2 + 4\alpha\beta + 7\alpha + \beta\right)\theta^6 + \left(12\alpha^2\beta + 6\alpha\beta^2 + 34\alpha\beta + 6\alpha^2 + 3\beta^2 + 12\alpha + 14\beta\right)\theta^5 \\ + \left(4\alpha\beta^3 + 24\alpha^2\beta + 18\alpha^2\beta^2 + 60\alpha\beta^2 + 3\beta^3 + 3\alpha^2 + 38\beta^2 + 66\alpha\beta + 6\alpha + 37\beta\right)\theta^4 \\ + \left(\alpha\beta^3 + 12\alpha^2\beta^2 + 36\alpha^2\beta + 46\alpha\beta^2 + \beta^3 + 120\alpha\beta + 12\alpha^2 + 34\beta^2 + 36\alpha + 84\beta + 24\right)\beta\theta^3 \\ + \left(3\alpha^2\beta^2 + 13\alpha\beta^2 + 24\alpha^2\beta + 90\alpha\beta + 18\alpha^2 + 10\beta^2 + 66\alpha + 66\beta + 48\right)\beta^2\theta^2 \\ + \left(6\alpha^2\beta + 12\alpha^2 + 24\alpha\beta + 48\alpha + 18\beta + 36\right)\beta^3\theta \\ + \left(3\alpha^2 + 12\alpha + 9\right)\beta^4\end{matrix}\right\}}{\theta^4\left(\theta+\beta\right)^4}$$

The coefficient of variation $(C.V)$, coefficient of Skewness $\left(\sqrt{\beta_1}\right)$, coefficient of Kurtosis $\left(\beta_2\right)$ and index of dispersion $\left(\gamma\right)$ of the GGPLD (2.2)) are thus obtained as

$$C.V=\frac{\sigma}{\mu_1'}=\frac{\sqrt{\alpha\,\theta^3+\left(\alpha+2\alpha\,\beta+\beta\right)\theta^2+\left(\alpha\,\beta+\beta+2\alpha+2\right)\beta\theta+\left(\alpha+1\right)\beta^2}}{\alpha\left(\theta+\beta\right)+\beta}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{\mu_2^{3/2}}=\frac{\left\{\begin{array}{l}\alpha\,\theta^5+2\left(3\alpha+3\alpha\,\beta+\beta\right)\theta^4+\left(3\alpha\beta^2+2\beta^2+9\alpha\beta+2\alpha+6\beta\right)\theta^3\\+\left(\alpha\,\beta^2+9\alpha\,\beta+\beta^2+9\beta+6\alpha+6\right)\beta\,\theta^2+\left(3\alpha\beta+3\beta+6\alpha+6\right)\beta^2\theta\\+2\left(\alpha+1\right)\beta^3\end{array}\right\}}{\left\{\alpha\,\theta^3+\left(\alpha+2\alpha\,\beta+\beta\right)\theta^2+\left(\alpha\,\beta+\beta+2\alpha+2\right)\beta\theta+\left(\alpha+1\right)\beta^2\right\}^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^{\,2}}=\frac{\left\{\begin{array}{l}\alpha\,\theta^7+\left(3\alpha^2+4\alpha\,\beta+7\alpha+\beta\right)\theta^6+\left(12\alpha^2\beta+6\alpha\beta^2+34\alpha\,\beta+6\alpha^2+3\beta^2+12\alpha+14\beta\right)\theta^5\\+\left(4\alpha\beta^3+24\alpha^2\beta+18\alpha^2\beta^2+60\alpha\,\beta^2+3\beta^3+3\alpha^2+38\beta^2+66\alpha\,\beta+6\alpha+37\beta\right)\theta^4\\+\left(\alpha\beta^3+12\alpha^2\beta^2+36\alpha^2\beta+46\alpha\beta^2+\beta^3+120\alpha\,\beta+12\alpha^2+34\beta^2+36\alpha+84\beta+24\right)\beta\theta^3\\+\left(3\alpha^2\beta^2+13\alpha\,\beta^2+24\alpha^2\beta+90\alpha\,\beta+18\alpha^2+10\beta^2+66\alpha+66\beta+48\right)\beta^2\theta^2\\+\left(6\alpha^2\beta+12\alpha^2+24\alpha\,\beta+48\alpha+18\beta+36\right)\beta^3\theta\\+\left(3\alpha^2+12\alpha+9\right)\beta^4\end{array}\right\}}{\left\{\alpha\,\theta^3+\left(\alpha+2\alpha\,\beta+\beta\right)\theta^2+\left(\alpha\,\beta+\beta+2\alpha+2\right)\beta\theta+\left(\alpha+1\right)\beta^2\right\}^2}$$

$$\gamma=\frac{\sigma^2}{\mu_1'}=\frac{\alpha\,\theta^3+\left(\alpha+2\alpha\,\beta+\beta\right)\theta^2+\left(\alpha\,\beta+\beta+2\alpha+2\right)\beta\theta+\left(\alpha+1\right)\beta^2}{\theta\left(\theta+\beta\right)\left\{\alpha\left(\theta+\beta\right)+\beta\right\}}.$$

Again it can be easily verified that at $\beta=1$ and $\alpha=\beta=1$, these moments based measures reduce to the corresponding measures of GPLD and PLD.

Nature and behavior of coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion of GGPLD for varying values of parameters $\theta,\alpha,$ and β have been shown graphically in figure 2.

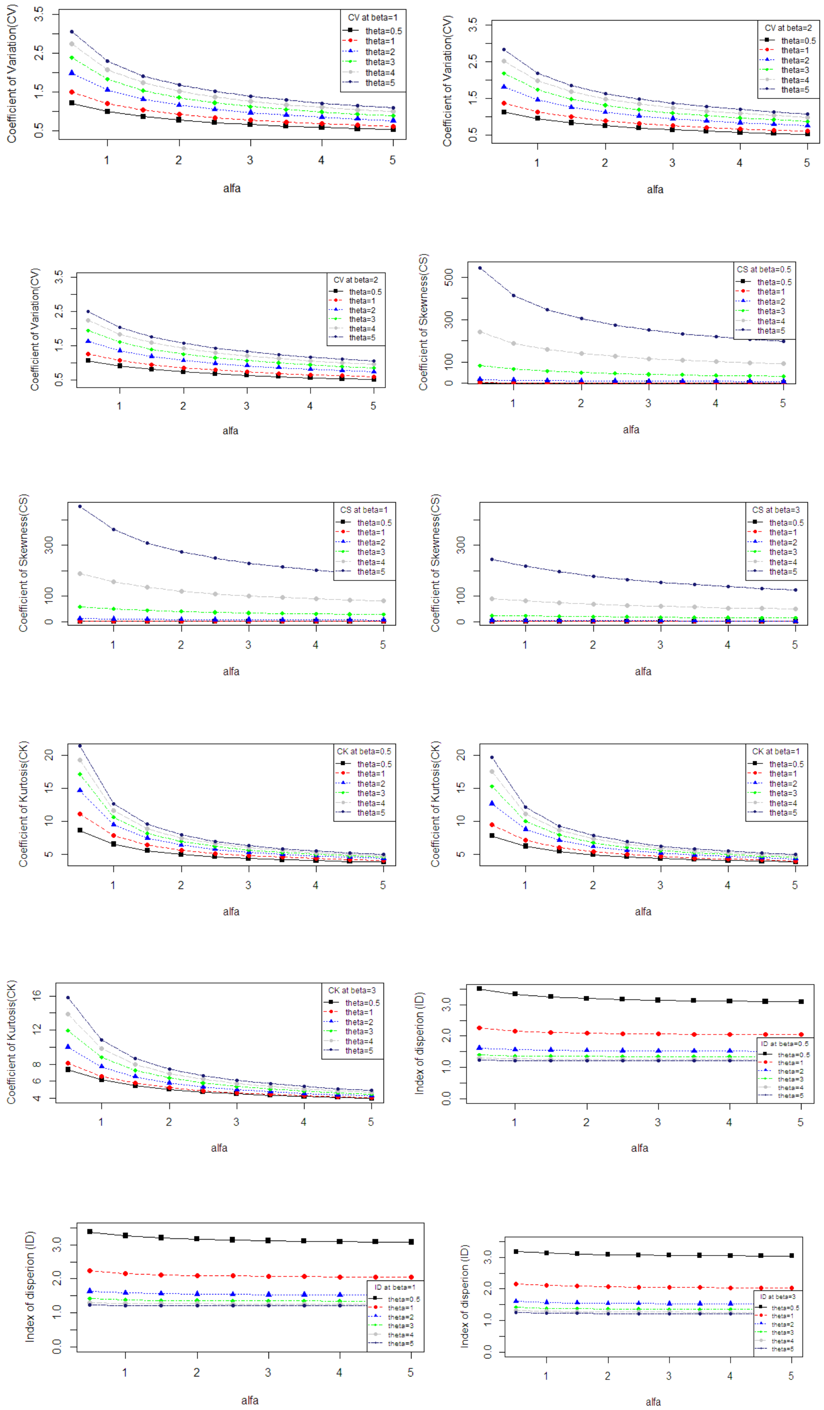

Figure 2. Nature and behavior of coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion of GGPLD for varying values of parameters θ, α, and β.

4. Maximum Likelihood Estimation of Parameters

Let $(x_1, x_2, ..., x_n)$ be a random sample of size n from the GGPLD (2.2) 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 log likelihood function of GGPLD (2.2) can be given by

$$\log L = n\left[(\alpha+1)\log\theta - (\alpha+1)\log(\theta+1) - \log\Gamma(\alpha+1) - \log(\theta+\beta)\right] - \sum_{x=1}^{k} x\, f_x \log(\theta+1)$$
$$+\sum_{x=1}^{k} f_x \left[\log\Gamma(x+\alpha) - \log\Gamma(x+1)\right] + \sum_{x=1}^{k} f_x \log\left[\beta x + \alpha(\theta+\beta+1)\right]$$

The maximum likelihood estimates $(\hat{\theta}, \hat{\alpha}, \hat{\beta})$ of (θ, α, β) of GGPLD (2.2) is the solutions of the following log likelihood equations

$$\frac{\partial \log L}{\partial \theta} = \frac{n(\alpha+1)}{\theta(\theta+1)} - \frac{n}{\theta+\beta} - \frac{n\bar{x}}{\theta+1} + \sum_{x=1}^{k} \frac{\alpha f_x}{\left[\beta x + \alpha(\theta+\beta+1)\right]} = 0$$

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

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

where $\bar{x}$ is the sample mean and $\psi(\alpha+1) = \frac{d}{d\alpha}\log\Gamma(\alpha+1)$ and $\psi(x+\alpha) = \frac{d}{d\alpha}\log\Gamma(x+\alpha)$ are digamma functions.

These three log likelihood equations do not seem to be solved directly. However, the Fisher's scoring method can be applied to solve these equations. We have

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{n(\alpha+1)}{\theta^2} + \frac{n(\alpha+1)}{(\theta+1)^2} + \frac{n\bar{x}}{(\theta+1)^2} + \sum_{x=1}^{k} \frac{\alpha^2 f_x}{\left[\beta x + \alpha(\theta+\beta+1)\right]^2}$$

$$\frac{\partial^2 \log L}{\partial \alpha^2} = -n\psi'(\alpha+1) + \sum_{x=1}^{k} f_x \psi'(x+\alpha) - \sum_{x=1}^{k} \frac{(\theta+\beta+1)^2 f_x}{\left[\beta x + \alpha(\theta+\beta+1)\right]^2}$$

$$\frac{\partial^2 \log L}{\partial \beta^2} = \frac{n}{(\theta+\beta)^2} - \sum_{x=1}^{k} \frac{(x+\alpha)^2 f_x}{\left[\beta x + \alpha(\theta+\beta+1)\right]^2}$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = \frac{n}{\theta(\theta+1)} + \sum_{x=1}^{k} \frac{\beta x f_x}{\left[\beta x + \alpha(\theta+\beta+1)\right]^2} = \frac{\partial^2 \log L}{\partial \alpha \partial \theta}$$

$$\frac{\partial^2 \log L}{\partial\theta\,\partial\beta} = \frac{n}{(\theta+\beta)^2} - \sum_{x=1}^{k} \frac{\alpha(x+\alpha)\,f_x}{\left[\beta x+\alpha(\theta+\beta+1)\right]^2} = \frac{\partial^2 \log L}{\partial\beta\,\partial\theta}$$

$$\frac{\partial^2 \log L}{\partial\alpha\,\partial\beta} = -\sum_{x=1}^{k} \frac{(\theta+1)\,x\,f_x}{\left[\beta x+\alpha(\theta+\beta+1)\right]^2} = \frac{\partial^2 \log L}{\partial\beta\,\partial\alpha},$$

where $\psi'(\alpha+1) = \frac{d}{d\alpha}\psi(\alpha+1)$ and $\psi'(x+\alpha) = \frac{d}{d\alpha}\psi(x+\alpha)$ are trigamma functions.

The maximum likelihood estimates $(\hat\theta,\hat\alpha,\hat\beta)$ of (θ,α,β) of GGPLD (2.2) is the solution of the following equations

$$\begin{bmatrix} \frac{\partial^2 \log L}{\partial\theta^2} & \frac{\partial^2 \log L}{\partial\theta\,\partial\alpha} & \frac{\partial^2 \log L}{\partial\theta\,\partial\beta} \\ \frac{\partial^2 \log L}{\partial\alpha\,\partial\theta} & \frac{\partial^2 \log L}{\partial\alpha^2} & \frac{\partial^2 \log L}{\partial\alpha\,\partial\beta} \\ \frac{\partial^2 \log L}{\partial\beta\,\partial\theta} & \frac{\partial^2 \log L}{\partial\beta\,\partial\alpha} & \frac{\partial^2 \log L}{\partial\beta^2} \end{bmatrix}_{\substack{\hat\theta=\theta_0\\ \hat\alpha=\alpha_0\\ \hat\beta=\beta_0}} \begin{bmatrix} \hat\theta-\theta_0 \\ \hat\alpha-\alpha_0 \\ \hat\beta-\beta_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \log L}{\partial\theta} \\ \frac{\partial \log L}{\partial\alpha} \\ \frac{\partial \log L}{\partial\beta} \end{bmatrix}_{\substack{\hat\theta=\theta_0\\ \hat\alpha=\alpha_0\\ \hat\beta=\beta_0}}$$

where $\theta_0,\alpha_0,$ and β_0 are the initial values of θ,α and β respectively. These equations are solved iteratively till sufficiently close values of $\hat\theta,\hat\alpha$ and $\hat\beta$ are obtained.

5. Applications

In this section the applications of the GGPLD has been discussed with some count datasets from biological sciences and thunderstorms events. The dataset in table 1 is the data regarding the number of red mites on apple leaves, available in Bliss (1953). The dataset in tables 2 and 3 are the Mammalian Cytogenetic dosimetry Lesions in Rabbit Lymphoblast induced by streponigrin (NSC-45383), available in Catcheside et al (1946). The dataset in table 4 is the number of micronuclei after exposure at dose 4 Gy of γ irradiation, counted using the cytochalasin B method, available in Piug and Valero (2006). The dataset in tables 5 and 6 are the frequencies of the observed number of days that experienced X thunderstorm events at Cape kennedy, Florida for the 11-year period of record in the month of June and July, January 1957 to December 1967 and are available in Falls et al (1971) and Carter (2001). The goodness of fit of GGPLD has been compared with the goodness of fit given by Poisson distribution (PD), PLD, NBD and GPLD. Note that the estimates of the parameters are based on maximum likelihood estimates for all the considered distributions. Based on the values of chi-square $\left(\chi^2\right)$, $-2\log L$ and AIC (Akaike Information criterion), it is obvious that GGPLD gives closer fit over the considered distributions. Note that AIC has been calculated using the formula $AIC = -2\log L + 2k$, where k is the number of parameters involved in the distribution.

Table 1. Observed and Expected number of European red mites on Apple leaves, available in Bliss (1953)

Number of Red mites per leaf	Observed frequency	Expected frequency				
		PD	PLD	NBD	GPLD	GGPLD
0	70	47.6	67.2	69.5	69.8	69.9
1	38	54.6	38.9	37.6	36.7	36.6
2	17	31.3	21.2	20.1	20.1	20.2
3	10	11.9	11.1	10.7	10.9	11.0
4	9	3.4	5.7	5.7	5.8	5.8
5	3	0.8	2.8	3.0	3.1	3.1
6	2	0.2	1.4	1.6	1.6	1.6
7	1	0.1	0.9	0.9	0.8	0.8
8	0	0.1	0.8	0.9	1.2	1.0
Total	150	150.0	150.0	150.0	150.0	150.0
ML estimates		$\hat{\theta}=1.14666$	$\hat{\theta}=1.26010$	$\hat{\theta}=1.11914$ $\hat{\alpha}=1.02459$	$\hat{\theta}=1.09620$ $\hat{\alpha}=0.78005$	$\hat{\theta}=1.09860$ $\hat{\alpha}=0.69603$ $\hat{\beta}=1.42003$
Standard Errors		0.08743	0.11390	0.40136 0.42097	0.25400 0.31550	0.25887 0.97563 5.22472
χ^2		26.50	2.49	2.91	2.43	2.45
d.f		2	4	3	3	2
p-value		0.0000	0.5595	0.4057	0.4880	0.2937
$-2\log L$		485.61	445.02	469.38	444.62	444.61
AIC		487.61	447.02	447.02	448.62	450.61

Table 2. Mammalian Cytogenetic dosimetry Lesions in Rabbit Lymphoblast induced by streponigrin (NSC-45383), Exposure- 60 $\mu g \mid kg$

Class/Exposure ($\mu g \mid kg$)	Observed frequency	Expected frequency				
		PD	PLD	NBD	GPLD	GGPLD
0 1 2 3 4 5 6	413 124 42 15 5 0 2	374.0 177.4 42.1 6.6 0.8 0.1 0.0	405.7 133.6 42.6 13.3 4.1 1.2 0.5	412.7 124.9 41.5 14.2 4.9 1.7 1.1	412.9 124.1 42.0 14.3 4.9 1.6 1.2	413.0 124.1 42.0 14.4 4.9 1.6 1.0
Total	601	601.0	601.0	601.0	601.0	601.0
ML estimates		$\hat{\theta} = 0.47421$	$\hat{\theta} = 2.68537$	$\hat{\theta} = 1.76494$ $\hat{\alpha} = 0.83700$	$\hat{\theta} = 2.16876$ $\hat{\alpha} = 0.71287$	$\hat{\theta} = 2.17980$ $\hat{\alpha} = 0.70100$ $\hat{\beta} = 1.08585$
Standard Errors		0.02809	0.16467	0.40075 0.17964	0.38481 0.20487	0.95510 1.04202 7.45628
χ^2		48.17	1.34	0.12	0.10	0.06
d.f		2	3	2	2	1
p-value		0.0000	0.7206	0.94129	0.9520	0.8096
$-2\log L$		1165.35	1113.76	1112.39	1112.36	1112.36
AIC		1167.35	1115.76	1116.39	1116.36	1118.36

Table 3. Mammalian Cytogenetic dosimetry Lesions in Rabbit Lymphoblast induced by streponigrin (NSC-45383), Exposure- 90 $\mu g \mid kg$

Class/Exposure ($\mu g \mid kg$)	Observed frequency	Expected frequency				
		PD	PLD	NBD	GPLD	GGPLD
0 1 2 3 4 5 6	155 83 33 14 11 3 1	127.8 109.0 46.5 13.2 2.8 0.5 0.2	158.3 77.2 35.9 16.1 7.1 3.1 2.3	155.1 80.6 36.7 15.9 6.7 2.8 2.2	155.3 80.1 36.9 16.0 6.7 2.8 2.2	155.3 80.1 36.9 16.1 6.8 2.8 2.0
Total	300	300.0	300.0	300.0	300.0	300.0
ML estimates		$\hat{\theta} = 0.85333$	$\hat{\theta} = 1.61761$	$\hat{\theta} = 1.56009$	$\hat{\theta} = 1.80860$	$\hat{\theta} = 1.83019$

				$\hat{\alpha}=1.33128$	$\hat{\alpha}=1.18743$	$\hat{\alpha}=1.12587$ $\hat{\beta}=1.41395$
Standard Errors		0.05333	0.11327	0.41479 0.33752	0.40045 0.370072	0.501159 1.18510 7.80636
χ^2		24.969	1.51	1.60	1.69	1.78
d.f		2	3	2	2	1
p-value		0.0000	0.6799	0.4488	0.42955	0.1821
$-2\log L$		800.92	766.10	765.86	765.79	765.79
AIC		802.92	768.10	769.86	769.79	771.79

Table 4. Number of micronuclei after exposure at dose 4 Gy of γ irradiation, counted using the cytochalasin B method and available in Piug and Valero (2006)

Number of micronuclei	Observed frequency	Expected frequency				
		PD	PLD	NBD	GPLD	GGPLD
0 1 2 3 4 5 6 7	1974 1674 869 342 102 26 13 2	1816.0 1839.9 932.1 314.8 79.7 16.1 2.7 1.6	2396.8 1300.3 668.8 332.1 160.9 76.5 35.8 30.8	1966.2 1695.5 331.5 857.5 108.4 31.6 8.4 2.9	1964.9 1696.6 857.9 331.5 108.3 31.5 8.41 2.9	1966.5 1695.2 857.3 331.6 108.5 31.6 5.5 2.8
Total	5002	5002.0	5002.0	5002.0	5002.0	5002.0
ML estimates		$\hat{\theta}=1.01319$	$\hat{\theta}=1.38736$	$\hat{\theta}=5.79197$ $\hat{\alpha}=5.79197$	$\hat{\theta}=5.88560$ $\hat{\alpha}=5.81844$	$\hat{\theta}=5.91675$ $\hat{\alpha}=5.69437$ $\hat{\beta}=2.53870$
Standard Errors		0.01423	0.02251	0.82644 0.83264	0.82310 0.84660	1.31079 2.26181 40.79249
χ^2		62.21	337.08	3.37	3.36	3.37
d.f		4	5	4	4	3
p-value		0.0000	0.0000	0.4976	0.4995	0.3381
$-2\log L$		13535.82	13836.70	13471.80	13471.81	13471.8
AIC		13537.82	13836.70	13475.80	13475.81	13477.8

Table 5. Frequencies of the observed number of days that experienced X thunderstorm events at Cape kennedy, Florida for the 11-year period of record in the month of June, January 1957 to December 1967.

X	Observed frequency	Expected frequency				
		PD	PLD	NBD	GPLD	GGPLD
0	187	155.6	185.3	184.6	185.3	185.6
1	77	116.9	83.4	84.5	83.5	82.9
2	40	43.9	35.9	35.8	35.9	36.2
3	17	11.0	15.0	14.8	15.0	15.1
4	6	2.0	6.1	6.0	6.1	6.1
5	2	0.3	2.5	2.4	2.5	2.5
6	1	0.3	1.8	1.9	1.7	1.6
Total	330	330.0	330.0	330.0	330.0	330.0
ML estimates		$\hat{\theta}=0.75148$	$\hat{\theta}=1.80427$	$\hat{\theta}=1.55916$ $\hat{\alpha}=1.17172$	$\hat{\theta}=1.80780$ $\hat{\alpha}=1.00340$	$\hat{\theta}=1.79790$ $\hat{\alpha}=0.66060$ $\hat{\beta}=4.11542$
Standard Errors		0.04772	0.12573	0.41501 0.29696	0.39558 0.32657	0.42815 0.73695 6.62533
χ^2		31.6	1.43	1.68	1.42	1.21
d.f		2	3	2	2	1
p-value		0.0000	0.6985	0.4317	0.4916	0.2733
$-2\log L$		824.50	788.88	789.18	788.88	788.65
AIC		826.50	790.88	793.18	792.88	794.65

Table 6. Frequencies of the observed number of days that experienced X thunderstorm events at Cape kennedy, Florida for the 11-year period of record in the month of July, January 1957 to December 1967.

X	Observed frequency	Expected frequency				
		PD	PLD	NBD	GPLD	GGPLD
0	177	142.3	177.7	171.8	172.7	173.2
1	80	124.3	87.9	94.0	92.8	92.1
2	47	54.3	41.5	43.3	43.2	43.3
3	26	15.8	18.9	18.7	18.8	20.0
4	9	3.5	8.4	7.8	8.0	7.9
5	2	0.8	6.6	5.4	5.4	4.5
Total	341	341.0	341.0	341.0	341.0	341.0

ML estimates		$\hat{\theta}=0.87390$	$\hat{\theta}=1.58353$	$\hat{\theta}=1.67672$ $\hat{\alpha}=1.46527$	$\hat{\theta}=1.86350$ $\hat{\alpha}=1.28028$	$\hat{\theta}=1.84560$ $\hat{\alpha}=0.92720$ $\hat{\beta}=4.02700$
Standard Errors		0.05062	0.10317	0.45068 0.37896	0.42561 0.40429	0.46436 0.70499 8.78609
χ^2		39.4	5.16	5.77	5.39	3.94
d.f		2	3	2	2	1
p-value		0.0000	0.1594	0.0558	0.0674	0.0471
$-2\log L$		911.00	880.50	880.35	879.93	879.55
AIC		913.00	882.50	884.35	883.93	885.55

From the goodness of fit given above it is obvious that GGPLD competes well with other distributions. In table 1, both GPLD and GGPLD gives equal fit. In table 2, GGPLD gives slightly better fit than GPLD. In table 3, PLD gives better fit. In table 4, NBD, GPLD and GGPLD gives almost identical fit. In table 5, GGPLD gives better fit as compared to other distributions. In table 6, GGPLD gives closer fits than other considered distributions.

6. Concluding Remarks

A generalization of generalized Poisson-Lindley distribution (GGPLD), which includes generalized Poisson-Lindley distribution (GPLD), Poisson-Lindley distribution (PLD), Poisson-weighted Lindley distribution (P-WLD), negative binomial distribution (NBD) and geometric distribution (GD) as particular cases, has been proposed and studied. Its moments and moments based measures have been derived and discussed. Estimation of parameters has been discussed using maximum likelihood estimation. The goodness of fit of the distribution has also been discussed with some count datasets and the fit has been compared with other discrete distributions.

6.1 Acknowledgements

Authors are grateful to the Editor-In –chief of the journal and the anonymous reviewer for their constructive comments which improved the quality and presentation of the paper.

References

Bliss, C.I. (1953): Fitting negative binomial distribution to biological data, Biometrics, 9, 177 – 200.

Carter M.C. A model for Thunderstorm Activity: Use of the Compound Negative Binomial-Positive Binomial Distribution, Journal of the Royal Statistical society, Series C (Applied Statistics), 2001; 2, 196 – 201.

Catcheside, D.G., Lea, D.E. & Thoday, J.M. (1946) : The production of chromosome structural changes in Tradescantia microspores in relation to dosage, intensity and temperature, Journal of Genetics, 47, 137-149.

El-Monsef, M.M.E. and Sohsah, N.M. (2014): Poisson-Weighted Lindley Distribution, Jokull Journal, 64(5), 192 – 202.

Falls, L.W., Wilford, W.O., and Carter M.C. Probability distribution for Thunderstorm Activity at Cape Kennedy, Florida, Journal of Applied, Meteorology, 1971; 10, 97 – 104.

Ghitany, M.E., Alqallaf ,F., Al-Mutairi, D.K., Husain, H.A. (2011): A two-parameter weighted Lindley distribution and its applications to survival data, Mathematics and Computers in simulation, 81, 1190-1201.

Ghitany, M.E., Atieh, B. and Nadarajah, S. (2008): Lindley distribution and its Application, Mathematics Computing and Simulation, 78, 493 – 506

Ghitany, M.E. and Al-Mutairi, D.K. (2008): Size-biased Poisson-Lindley distribution and Its Applications, Metron - International Journal of Statistics, LXVI (3), 299 – 311.

Ghitany, M.E., Al-Mutairi, D.K. and Nadarajah, S. (2009): Zero-truncated Poisson- Lindley distribution and its Applications, Mathematics and Computers in Simulation, 79 (3), 279 – 287.

Lindley, D. V., (1958). Fiducial distributions and Bayes' theorem, Journal of the Royal Statistical Society, Series B, 20, PP. 102- 107.

Mahmoudi, E. and Zakerzadeh, H. (2010): Generalized Poisson-Lindley distribution, Communications in Statistics-Theory & Methods, 39, 1785 – 1798.

Puig, P. and Valero, J. (2006): Count data distribution-Some characterizations with applications, Journal of the American Statistical Association, 101, 332 – 340.

Shanker, R., Shukla, K.K. and Hagos, F. (2016): On Weighted Lindley distribution and Its Applications to Model Lifetime Data, Jacobs Journal of Biostatistics,1(1), 1- 9.

Shanker, R. and Shukla, K.K. (2016): On modeling of lifetime data using three-parameter Generalized Lindley and Generalized Gamma distributions, Biometrics & Biostatistics International Journal, 4(7), 1 – 7.

Shanker, R., Shukla, K.K., and Mishra, A. (2017): A Three- parameter weighted Lindley distribution and its applications to model survival time data, Statistics in Transition-New Series, 18 (2), 291 –300.

Shanker, R. and Shukla, K.K. (2017): On Poisson-Weighted Lindley distribution, J. Sci. Res. 11 (1), 1-13.

Stacy, E.W. (1962): A generalization of the gamma distribution, Annals of Mathematical Statistics, 33, 1187 – 1192.

Zakerzadeh, H. and Dolati, A. (2009): Generalized Lindley distribution, Journal of Mathematical extension, 3 (2), 13 – 25.