Biostatistics and Biometrics
Open Access Journal
ISSN: 2573-2633

Juniper Publishers
key to the Researchers

Research Article
Volume 3 Issue 1 - September 2017
DOI: 10.19080/BBOAJ.2017.03.555605

Biostat Biometrics Open Acc J

Zero-Truncated Poisson-Garima Distribution and its Applications

Rama Shanker* and Kamlesh Kumar Shukla

Department of Statistics, Eritrea Institute of Technology, Eritrea

Submission: August 23, 2017; **Published:** September 26, 2017

***Corresponding author:** Rama Shanker, Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea, Email: shankerrama2009@gmail.com; kkshukla22@gmail.com

Abstract

'In this paper, a zero-truncated Poisson-Garima distribution (ZTPGD) has been introduced by taking the zero-truncated version of Poisson-Garima distribution (PGD) of shanker [1]. The moments about origin and moments about mean, coefficients of variation, skewness and kurtosis, and index of dispersion of ZTPGD have been obtained. The method of moment and the method of maximum likelihood estimation have been discussed for estimating the parameter of ZTPGD. Goodness of fit of ZTPGD has been discussed with two real data sets and the fit has been compared with that of zero-truncated Poisson distribution (ZTPD) and zero-truncated Poisson-Lindley distribution (ZTPLD).

Abbreviations: ZTPGD: zero-truncated Poisson-Garima distribution; PGD: Poisson-Garima distribution; ZTPD: zero-truncated Poisson distribution; ZTPLD: zero-truncated Poisson-Lindley distribution

Introduction

Suppose $P_0(x;\theta)$ is the original discrete distribution. Then, the zero-truncated version of $P_0(x;\theta)$ is defined as

$$P(x;\theta)=\frac{P_0(x;\theta)}{1-P_0(0;\theta)} \quad ;x=1,2,3,... \qquad (1.1)$$

In probability theory, zero-truncated distributions are certain discrete distributions whose support is the set of positive integers. When the data to be modeled originate from a mechanism which generates data that structurally excludes zero counts, zero-truncated distribution is the appropriate choice.

The probability mass function (p.m.f.) of the Poisson-Garima distribution (PGD) given by

$$P_0(x;\theta)=\frac{\theta}{\theta+2}\frac{\theta x+(\theta^2+3\theta+1)}{(\theta+1)^{x+2}}\ ;\ x=0,1,2,3,...,\ \theta>0 \qquad (1.2)$$

was introduced by shanker [1] to model count data. This distribution arises from the Poisson distribution when its parameter follows Garima distribution introduced by shanker [2] with the probability density function

$$f(x;\theta)=\frac{\theta}{\theta+2}(1+\theta+\theta x)e^{-\theta x};\quad x>0,\quad \theta>0 \qquad (1.3)$$

This distribution has been extensively studied by Shanker [2] and it has been shown that (1.3) provides a better model for behavioral science data than many one parameter exponential, Lindley distribution introduced by Lindley[3], Shanker, Akash, Aradhana and Sujatha distributions introduced by [3-7].

In the present paper an attempt has been made to obtain a zero-truncated Poisson-Garima distribution (ZTPGD) by taking the zero-truncated version of Poisson-Garima distribution (PGD) of shanker [1]. Its moments and moments based properties including coefficients of variation, skewness and kurtosis, and index of dispersion of ZTPGD have been obtained and discussed graphically. For estimating the parameter the method of moment and the method of maximum likelihood estimation have been discussed. Goodness of fit of ZTPGD has been discussed with two real data sets and the fit has been compared with that of zero-truncated Poisson distribution (ZTPD) and zero-truncated Poisson-Lindley distribution (ZTPLD) [8].

F (ZTPGD)

Using (1.1) and (1.2), the p.m.f of zero-truncated Poisson-Garima distribution (ZTPGD) can be obtained as

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

Graphs of ZTPGD for varying values of parameter have been presented in Figure 1. It is clear from the graphs of ZTPGD that as the value of parameter increases, initially graph shift upward and as the value of increases graph decrease fast and becomes asymptote to the x-axis.

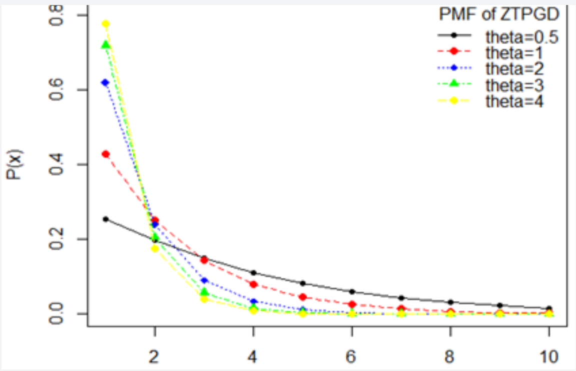

Figure1: Graphs of ZTPGD for varying values of parameter.

The ZTPGD can also be obtained from the size-biased Poisson distribution (SBPD) with p.m.f.

$$g(x\mid\lambda)=\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!};\quad x=1,2,3,...,\quad \lambda>0 \quad (2.2)$$

when its parameter follows a continuous distribution having p.d.f.

$$h(\lambda;\theta)=\frac{\theta}{\theta^2+4\theta+2}\left[\theta(\theta+1)\lambda+\left(\theta^2+3\theta+1\right)\right]e^{-\theta\lambda};\quad \lambda>0,\quad \theta>0 \quad (2.3)$$

The p.m.f. of ZTPGD can thus be obtained as

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

$$=\frac{\theta}{\theta^2+4\theta+2}\int_0^\infty\frac{e^{-\lambda}\lambda^{x-1}}{(x-1)!}\left[\theta(\theta+1)\lambda+\left(\theta^2+3\theta+1\right)\right]e^{-\theta\lambda}d\lambda;\quad \lambda>0,\quad \theta>0$$

(2.4)

$$=\frac{\theta}{\left(\theta^2+4\theta+2\right)(x-1)!}\int_0^\infty e^{-(\theta+1)\lambda}\cdot\left[\theta(\theta+1)\lambda^{x+1-1}+\left(\theta^2+3\theta+1\right)\lambda^{x-1}\right]d\lambda$$

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

$$=\frac{\theta}{\left(\theta^2+4\theta+2\right)}\cdot\left[\frac{\theta x+(\theta^2+3\theta+1)}{(\theta+1)^x}\right];\quad x=1,2,3,....;\quad \theta>0$$

which is the p.m.f. of ZTPGD.

The pmf of zero-truncated Poisson-Lindley distribution (ZTPLD) obtained by Ghitany et al. [9] is given by

$$P_3(x;\theta)=\frac{\theta^2}{\theta^2+3\theta+1}\frac{x+\theta+2}{(\theta+1)^x};x=1,2,3,...,\theta>0 \quad (2.5)$$

Recall that ZTPLD has been obtained by zero-truncating the discrete Poisson-Lindley distribution suggested by Sankaran [10] and the discrete Poisson-Lindley distribution is the Poisson mixture of Lindley distribution introduced by Lindley [3]. Shanker & Hagos [11] have detailed study on applications of Poisson-Lindley distribution for biological sciences. Shanker et al. [12] have detailed study on modeling of lifetime data using both exponential and Lindley distributions and concluded that both exponential and Lindley distributions compete each other.

The pmf of zero-truncated Poisson distribution (ZTPD) is given by

$$P_4(x;\theta)=\frac{e^{-\theta}\theta^x}{\left(1-e^{-\theta}\right)x!}\quad;x=1,2,3,...,\theta>0 \quad (2.6)$$

Moments and Related Measures

Using (2.4), the th factorial moment about origin of ZTPGD (2.1) can be obtained as

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

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

Taking $y=x-r$, we get

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

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

$$=\frac{r!(\theta+1)\left\{\theta^2+(r+3)\theta+(r+2)\right\}}{\theta^r\left(\theta^2+4\theta+2\right)};r=1,2,3,.... \quad (3.1)$$

Substituting in (3.1), the first four factorial moments can be obtained and using the relationship between moments about origin and factorial moments about origin, the first four moments about origin of ZTPGD can be obtained as

$$\mu_1'=\frac{(\theta+1)\left(\theta^2+4\theta+3\right)}{\theta\left(\theta^2+4\theta+2\right)}$$

$$\mu_2'=\frac{(\theta+1)\left(\theta^3+6\theta^2+13\theta+8\right)}{\theta^2\left(\theta^2+4\theta+2\right)}$$

$$\mu_3'=\frac{(\theta+1)\left(\theta^4+10\theta^3+39\theta^2+60\theta+30\right)}{\theta^3\left(\theta^2+4\theta+2\right)}$$

$$\mu_4'=\frac{(\theta+1)\left(\theta^5+18\theta^4+109\theta^3+296\theta^2+348\theta+144\right)}{\theta^4\left(\theta^2+4\theta+2\right)}$$

Using the relationship between moments about mean and moments about origin, the first four moments about mean of ZTPGD are thus obtained as

$$\mu_2=\frac{(\theta+1)\left(\theta^4+9\theta^3+26\theta^2+25\theta+7\right)}{\theta^2\left(\theta^2+4\theta+2\right)^2}$$

$$\mu_3=\frac{(\theta+1)\left(\theta^7+15\theta^6+92\theta^5+291\theta^4+495\theta^3+438\theta^2+186\theta+30\right)}{\theta^3\left(\theta^2+4\theta+2\right)^3}$$

$$\mu_4=\frac{(\theta+1)\begin{pmatrix}\theta^{10}+26\theta^9+289\theta^8+1776\theta^7+6578\theta^6+15138\theta^5+21722\theta^4+19175\theta^3+\\10027\theta^2+2835\theta+333\end{pmatrix}}{\theta^4\left(\theta^2+4\theta+2\right)^4}$$

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 of (γ) ZTPGD are thus obtained as

$$C.V=\frac{\sigma}{\mu_1'}=\frac{\sqrt{\theta^4+9\theta^3+26\theta^2+25\theta+7}}{\left(\sqrt{\theta+1}\right)\left(\theta^2+4\theta+3\right)}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{(\mu_2)^{3/2}}=\frac{\left(\theta^7+15\theta^6+92\theta^5+291\theta^4+495\theta^3+438\theta^2+186\theta+30\right)}{\left(\sqrt{\theta+1}\right)\left(\theta^4+9\theta^3+26\theta^2+25\theta+7\right)^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{\begin{pmatrix}\theta^{10}+26\theta^9+289\theta^8+1776\theta^7+6578\theta^6+15138\theta^5+21722\theta^4+19175\theta^3+\\10027\theta^2+2835\theta+333\end{pmatrix}}{(\theta+1)\left(\theta^4+9\theta^3+26\theta^2+25\theta+7\right)^2}$$

$$\gamma=\frac{\sigma^2}{\mu}=\frac{\theta^4+9\theta^3+26\theta^2+25\theta+7}{\theta\left(\theta^2+4\theta+2\right)\left(\theta^2+4\theta+3\right)}$$

Graphs of C.V, coefficient of skewness, coefficient of kurtosis and index of dispersion of ZTPGD are shown in Figure 2. From the graphs it is obvious that C.V, coefficient of skewness and

How to cite this article: Rama S, Kamlesh K S . Zero-Truncated Poisson-Garima Distribution and its Applications. Biostat Biometrics Open Acc J. 2017; 3(1): 555605. DOI: 10.19080/BBOAJ.2017.03.555605

index of dispersion are increasing while graph of coefficient of kurtosis are increasing for increasing values of the parameter θ .

The over-dispersion, under-dispersion and under-dispersion of ZTPGD and ZTPLD for values of parameter are presented in Table 1.

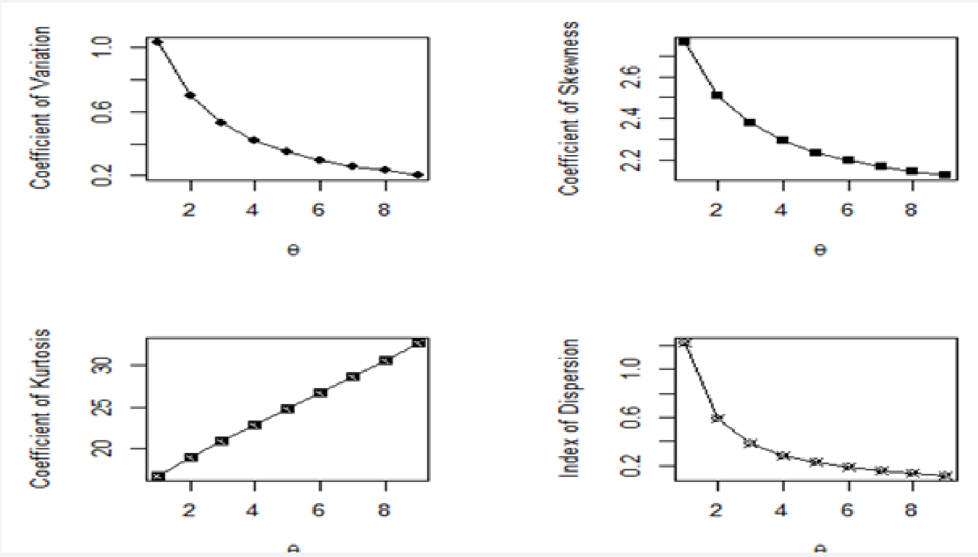

Figure 2: Graphs of C.V, coefficient of skewness, coefficient of kurtosis and index of dispersion of ZTPGD for varying values of parameter .

Table 1: Over-dispersion, under-dispersion and under-dispersion of ZTPGD and ZTPLD.

Distributions	Over-Dispersion	Equi-Dispersion	Under-Dispersion
	$\left(\mu<\sigma^2\right)$	$\left(\mu=\sigma^2\right)$	$\left(\mu>\sigma^2\right)$
ZTPGD	$\theta<1.20895$	$\theta=1.20895$	$\theta>1.20895$
ZTPLD	$\theta<1.25863$	$\theta=1.25863$	$\theta=1.25863$

Unimodality and Increasing Hazard Rate: Since

$$\frac{P(x+1;\theta)}{P(x;\theta)}=\left(\frac{1}{\theta+1}\right)\left[1+\frac{\theta}{\theta x+\left(\theta^2+3\theta+1\right)}\right]$$

is a decreasing function of , is log-concave. Therefore, ZTPGD is unimodal, has increasing failure rate (IFR), and hence increasing failure rate average (IFRA). It is new better than used (NBU), new better than used in expectation (NBUE), and has decreasing mean residual life (DMRL). Detailed discussions about the definitions of these aging concepts are available in Barlow & Proschan [13].

Generating Function

Probability Generating Function: The probability generating function of the ZTPGD (2.1) is obtained as

$$P_X(t)=E\left(t^X\right)=\frac{\theta}{\theta^2+4\theta+2}\left[\theta\sum_{x=1}^{\infty}x\left(\frac{t}{\theta+1}\right)^x+\left(\theta^2+3\theta+1\right)\sum_{x=1}^{\infty}\left(\frac{t}{\theta+1}\right)^x\right]$$

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

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

Moment Generating Function: The moment generating function of the ZTPGD (2.1) is given by

$$M_X(t)=E\left(e^{tX}\right)=\frac{\theta e^t}{\theta^2+4\theta+2}\left[\frac{\theta(\theta+1)}{\left(\theta+1-e^t\right)^2}+\frac{\theta^2+3\theta+1}{\left(\theta+1-e^t\right)}\right]$$

Parameter Estimation

Method of Moment Estimate (MOME)

Let be a random sample of size from the ZTPGD (2.1). Equating the population mean to the corresponding sample mean, MOME of is the solution of the following non-linear equation

How to cite this article: Rama S, Kamlesh K S . Zero-Truncated Poisson-Garima Distribution and its Applications. Biostat Biometrics Open Acc J. 2017; 3(1): 555605. DOI: 10.19080/BBOAJ.2017.03.555605

$$(\bar{x}-1)\theta^3+(4\bar{x}-5)\theta^2+(2\bar{x}-7)\theta-3=0$$

where $\bar{x}$ is the sample mean.

Maximum Likelihood Estimate (MLE)

Let $x_1, x_2, ..., x_n$ be a random sample of size from the ZTPGD (2.1) 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 is the largest observed value having non-zero frequency. The likelihood function of the ZTPGD (2.1) is given by

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

The log likelihood function is given by

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

and the log likelihood equation is thus obtained as

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

The maximum likelihood estimate $\hat{\theta}$ of θ is the solution of the equation $\frac{d\log L}{d\theta}=0$ and is given by the solution of the following non-linear equation

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

where is the sample mean. This non-linear equation can be solved by any numerical iteration methods such as Newton-Raphson method, Bisection method, Regula–Falsi method etc. The initial value of the parameter can be taken from the value of obtained by the method of moments.

Applications

The ZTPGD has been fitted to a number of data - sets to test its goodness of fit over ZTPD and ZTPLD. The maximum likelihood estimate (MLE) has been used to fit the ZTPGD. Two examples of observed data-sets, for which the ZTPD, ZTPLD and ZTPGD has been fitted, are presented. The first data-set in Table 2 is the number of European red mites on apple leaves, reported by Garman [14] and the second data-set in Table 3 is the animal abundance data of Keith & Meslow [15] regarding the distribution of snowshoe hares captured over 7 days (Figure 3).

Table 2: Number of European red mites on apple leaves, reported by Garman (1923).

Number of European red mites	Observed Frequency	Expected Frequency		
		ZTPD	ZTPLD	ZTPGD
1	38	28.7	36.1	36.7
2	17	25.7	20.5	20.2
3	10	15.3	11.2	10.9
4	9	6.9	5.9	5.9
5	3	6.9 2.5 0.7 0.2 0.1	3.1 1.6 0.8 0.8	3.0 1.7 1.1 0.5
6	2			
7	1			
8	0			
Total	80	80	80	80
ML Estimate		$\hat{\theta}=1.791615$	$\hat{\theta}=1.185582$	$\hat{\theta}=1.108926$
χ^2		9.827	2.467	2.241
d.f.		2	3	3
P-value		0.0073	0.4813	0.5239

Table 3 : Distribution of snowshoe hares captured over 7 days.

No. times Hares Caught	Observed Frequency	Expected Frequency		
		ZTPD	ZTPLD	ZTPGD
1	184	170.6	182.6	182.7
2	55	72.5	55.3	55.1
3	14	15.4 2.2 0.3	16.4	16.3
4	4		4.8 1.9	4.8 2.1
5	4			
Total	261	261	261	261

How to cite this article: Rama S, Kamlesh K S . Zero-Truncated Poisson-Garima Distribution and its Applications. Biostat Biometrics Open Acc J. 2017; 3(1): 555605. DOI: 10.19080/BBOAJ.2017.03.555605

ML estimate		$\hat{\theta} = 0.425$	$\hat{\theta} = 2.8639$	$\hat{\theta} = 2.778$
		6.216	0.61	0.49
d.f.		1	2	2
p-value		0.013	0.7371	0.7827

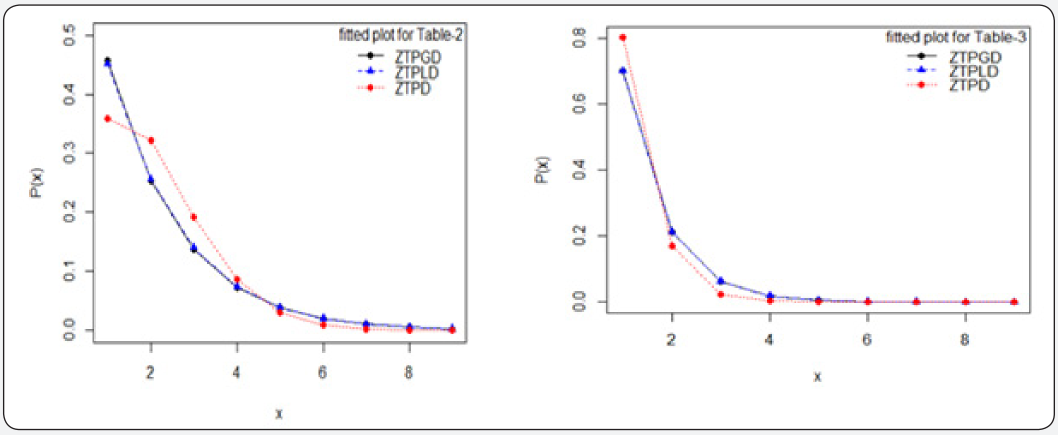

Figure 3: Fitted plots of the data in table 2 and 3 by ZTPGD, ZTPLD and ZTPD.

Concluding Remarks

A zero-truncated Poisson-Garima distribution (ZTPGD) has been introduced by taking the zero-truncated version of Poisson-Garima distribution (PGD) of Shanker [1]. To find the moments of the distribution ZTPGD has also been obtained by taking a size-biased mixture of an assumed continuous distribution. Its raw moments and central moments and moments based properties including coefficients of variation, skewness and kurtosis, and index of dispersion have been obtained and discussed graphically. Both the method of moment and the method of maximum likelihood estimation have been discussed for estimating the parameter. Finally two examples of real data sets have been presented to test the goodness of fit of ZTPGD and the fit shows quite satisfactory fit over ZTPD and ZTPL

References

1. Shanker R (2017) The discrete Poisson-Garima distribution. Biometrics & Biostatistics International Journal 5(2): 1-7.
2. Shanker R (2016) Garima distribution and its application to model behavioral science data. Biometrics & Biostatistics International Journal 4(7): 1-9.
3. Lindley DV (1958) A Two-Parameter Lindley Distribution for Modeling Waiting and Survival Times Data. Fiducial distributions and Bayes theorem, Journal of the Royal Statistical Society 20(1): 102-107.
4. Shanker R (2015a) Shanker Distribution and Its Applications. International Journal of Statistics and Applications 5(6): 338 -348.
5. Shanker R (2015b) Akash Distribution and Its Applications. International Journal of Probability and Statistics. International Journal of Probability and Statistics 4(3): 65-75.
6. Shanker R (2016a) Aradhana Distribution and Its Applications. International Journal of Statistics and Applications 6(1): 23-34.
7. Shanker R (2016b) Sujatha Distribution and Its Applications. Statistics in Transition-New series 17(3): 391-410.
8. Ghitany ME, Atieh B, Nadarajah S (2008) Lindley distribution and Its Applications. Mathematics Computation and Simulation 78(4): 493-506.
9. Ghitany ME, Al-Mutairi DK, Nadarajah S (2008) Zero-truncated Poisson-Lindley distribution and its Applications. Mathematics and Computers in Simulation 79(3): 279-287.
10. Sankaran M (1970) The discrete Poisson-Lindley distribution. Biometrics 26(1): 145-149.
11. Shanker R, Hagos F, Sujatha S (2015) On modeling of lifetimes data using exponential and Lindley distributions. Biometrics & Biostatistics International Journal 2(5): 1-9.
12. Shanker R, Hagos F (2015) On Poisson-Lindley distribution and Its Applications to Biological Sciences. Biometrics & Biostatistics International Journal 2(4): 1-5.
13. Barlow RE, Proschan F (1981) Statistical Theory of Reliability and Life Testing. Silver Spring MD.
14. Garman P (1923) The European red mites in Connecticut apple orchards, Connecticut Agri. Exper. Station Bull 252: 103-125.
15. Keith LB, Meslow EC (1968) Trap response by snowshoe hares. Journal of Wildlife Management 32(4): 795-801.

How to cite this article: Rama S, Kamlesh K S . Zero-Truncated Poisson-Garima Distribution and its Applications. Biostat Biometrics Open Acc J. 2017; 3(1): 555605. DOI: 10.19080/BBOAJ.2017.03.555605

DOI: 10.19080/BBOAJ.2017.03.555605

How to cite this article: Rama S, Kamlesh K S . Zero-Truncated Poisson-Garima Distribution and its Applications. Biostat Biometrics Open Acc J. 2017; 3(1): 555605. DOI: 10.19080/BBOAJ.2017.03.555605