

On zero-truncation of poisson and poisson-lindley distributions and their applications

Abstract

In this paper, a general expression for the r th factorial moment of zero-truncated Poisson-Lindley distribution (ZTPLD) has been obtained and hence the first four moments about origin has been given. A very simple and alternative method for finding moments of ZTPLD has also been suggested. The expression for the moment generating function of ZTPLD has been obtained. Both ZTPD (Zero-truncated Poisson distribution) and ZTPLD have been fitted using maximum likelihood estimate to a number of data- sets from demography, biological sciences and social sciences and it has been observed that in most cases ZTPLD gives much closer fits than ZTPD while in some cases ZTPD gives much closer fits than ZTPLD.

Keywords: poisson-lindley distribution, zero-truncated distribution, moments, estimation of parameter, goodness of fit

Volume 2 Issue 6 - 2015

Rama Shanker,¹ Hagos Fesshaye,² Sujatha Selvaraj,³ Abrehe Yemane⁴

¹Department of Statistics, Eritrea Institute of Technology, Eritrea

²Department of Economics, College of Business and Economics, Eritrea

³Department of Banking and Finance, Jimma University, Ethiopia

⁴Department of Statistics, Eritrea Institute of Technology, Eritrea

Correspondence: Rama Shanker, Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea, Email shankerrama2009@gmail.com

Received: June 23, 2015 | **Published:** July 22, 2015

Abbreviations: ZTPLD, zero-truncated poisson-lindley distribution; ZTPD, zero-truncated poisson distribution; PLD, poisson-lindley distribution; PMF, probability mass function; PDF, probability density function; SBPD, size-biased poisson distribution; MLE, maximum likelihood estimate; MOME, method of moment estimate; MVUE, minimum variance unbiased estimation

Introduction

Zero-truncated distributions, in probability theory, are certain discrete distributions having support the set of positive integers. These distributions are applicable for the situations when the data to be modeled originate from a mechanism that generates data excluding zero-counts.

Let $P_0(x; \theta)$ is the original distribution with support non negative positive integers. Then the zero-truncated version of $P_0(x; \theta)$ with the support the set of positive integers is given by

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

The Poisson-Lindley distribution (PLD) with parameter θ and having probability mass function (p.m.f.)

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

has been introduced by Sankaran¹ to model count data. Recently, Shanker and Hagos² have done an extensive study on its applications to Biological Sciences and found that PLD provides a better fit than Poisson distribution to almost all biological science data. The PLD arises from the Poisson distribution when its parameter follows Lindley distribution³ with probability density function (p.d.f.).

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

Detailed study of Lindley distribution (1.3) has been done by Ghitany et al.,⁴ and shown that (1.3) is a better model than exponential

distribution. Recently, Shanker et al.,⁵ showed that (1.3) is not always a better model than the exponential distribution for modeling lifetimes data. In fact, Shanker et al.,⁵ has done a very extensive and comparative study on modeling of lifetimes data using exponential and Lindley distributions and discussed various lifetimes examples to show the superiority of Lindley over exponential and that of exponential over Lindley distribution. The PLD has been extensively studied by Sankaran¹ and Ghitany & Mutairi⁶ and they have discussed its various properties. The Lindley distribution and the PLD have been generalized by many researchers. Shanker and Mishra⁷ obtained a two parameter Poisson-Lindley distribution by compounding Poisson distribution with a two parameter Lindley distribution introduced by Shanker and Mishra.⁸ A quasi Poisson-Lindley distribution has been introduced by Shanker and Mishra.⁹ by compounding Poisson distribution with a quasi Lindley distribution introduced by Shanker and Mishra.¹⁰ Shanker et al.,¹¹ obtained a discrete two parameter Poisson-Lindley distribution by mixing Poisson distribution with a two parameter Lindley distribution for modeling waiting and survival times data introduced by Shanker et al.,¹² Further, Shanker and Tekie.¹³ obtained a new quasi Poisson-Lindley distribution by compounding Poisson distribution with a new quasi Lindley distribution introduced by Shanker and Amanuel.¹⁴

In this paper, the nature of zero-truncated Poisson distribution (ZTPD) and zero-truncated Poisson-Lindley distribution (ZTPLD) has been compared and studied using graphs for different values of their parameter. A general expression for the r th factorial moment of ZTPLD has been obtained and the first four moment about origin has been given. A very simple and easy method for finding moments of ZTPLD has been suggested. Both ZTPD and ZTPLD have been fitted to a number of data sets from different fields to study their goodness of fits and superiority of one over the other.

Zero-truncated poisson and poisson-lindley distribution

Zero-truncated poisson distribution (ZTPD)

Using (1.1) and the p.m.f. of Poisson distribution, the p.m.f. of zero-truncated Poisson distribution (ZTPD) given by

$$P_1(x; \theta) = \frac{\theta^x}{(e^\theta - 1)x!}; x = 1, 2, 3, \dots, \theta > 0 \quad (2.1.1)$$

was obtained independently by Plackett¹⁵ and David and Johnson.¹⁶ to model count data that structurally excludes zero counts. An extension of a truncated Poisson distribution and estimation in a truncated Poisson distribution when zeros and some ones are missing has been discussed by Cohen.^{17,18} Tate and Goen.¹⁹ has discussed minimum variance unbiased estimation (MVUE) for the truncated Poisson distribution.

Zero-truncated poisson-lindley distribution (ZTPLD)

Using (1.1) and (1.2), the p.m.f. of zero-truncated Poisson- Lindley distribution (ZTPLD) given by

$$P_2(x; \theta) = \frac{\theta^2}{\theta^2 + 3\theta + 1} \frac{x + \theta + 2}{(\theta + 1)^x}; x = 1, 2, 3, \dots, \theta > 0 \quad (2.2.1)$$

was obtained by David and Johnson.¹⁶ to model count data for the missing zeros. It has been shown by Ghitany et al.,²⁰ that ZTPLD can also arise from the size-biased Poisson distribution (SBPD) with p.m.f.

$$f(x | \lambda) = \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!}; x = 1, 2, 3, \dots, \lambda > 0 \quad (2.2.2)$$

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

$$h(\lambda; \theta) = \frac{\theta^2}{\theta^2 + 3\theta + 1} \left[(\theta + 1)\lambda + (\theta + 2) \right] e^{-\theta\lambda}; \lambda > 0, \theta > 0 \quad (2.2.3)$$

Thus the p.m.f. of ZTPLD is obtained as

$$\begin{aligned} P(X = x) &= \int_0^\infty f(x | \lambda) \cdot h(\lambda; \theta) d\lambda \\ &= \int_0^\infty \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} \cdot \frac{\theta^2}{\theta^2 + 3\theta + 1} \left[(\theta + 1)\lambda + (\theta + 2) \right] e^{-\theta\lambda} d\lambda \quad (2.2.4) \\ &= \frac{\theta^2}{(\theta^2 + 3\theta + 1)(x-1)!} \int_0^\infty e^{-(\theta+1)\lambda} \cdot \left[(\theta + 1)\lambda^x + (\theta + 2)\lambda^{x-1} \right] d\lambda \\ &= \frac{\theta^2}{\theta^2 + 3\theta + 1} \left[\frac{x}{(\theta + 1)^x} + \frac{(\theta + 2)}{(\theta + 1)^x} \right] \\ &= \frac{\theta^2}{\theta^2 + 3\theta + 1} \frac{x + \theta + 2}{(\theta + 1)^x}; x = 1, 2, 3, \dots, \theta > 0 \end{aligned}$$

which is the p.m.f. of ZTPLD with parameter θ .

To study the nature and behaviors of ZTPD and ZTPLD for different values of its parameter, a number of graphs of their probability functions have been drawn and presented in Figure 1.

Moments and related measures

Moments of ZTPD

The r th factorial moment of the ZTPD (2.1.1) can be obtained as

$$\mu_{(r)}' = E[X^{(r)}] = \frac{1}{e^\theta - 1} \sum_{x=1}^\infty x^{(r)} \frac{\theta^x}{x!}, \text{ where}$$

$$\begin{aligned} X^{(r)} &= X(X-1)(X-2)\dots(X-r+1) \\ &= \frac{\theta^r}{e^\theta - 1} \sum_{x=r}^\infty \frac{\theta^{x-r}}{(x-r)!} = \frac{\theta^r e^\theta}{e^\theta - 1} \end{aligned} \quad (3.1.1)$$

Substituting $r = 1, 2, 3$, and 4 in (3.1.1), the first four factorial moments can be obtained and, therefore, using the relationship between factorial moments and moments about origin, the first four moments about origin of ZTPD (2.1.1) are obtained as

$$\begin{aligned} \mu_1' &= \frac{\theta e^\theta}{e^\theta - 1} \\ \mu_2' &= \frac{\theta e^\theta}{e^\theta - 1} (\theta + 1) \\ \mu_3' &= \frac{\theta e^\theta}{e^\theta - 1} (\theta^2 + 3\theta + 1) \\ \mu_4' &= \frac{\theta e^\theta}{e^\theta - 1} (\theta^3 + 6\theta^2 + 7\theta + 1) \end{aligned}$$

Generating Function: The probability generating function of the ZTPD (2.1.1) is obtained as

$$P_X(t) = E(t^X) = \frac{1}{e^\theta - 1} \sum_{x=1}^\infty \frac{(\theta t)^x}{x!} = \frac{1}{e^\theta - 1} \left[\sum_{x=0}^\infty \frac{(\theta t)^x}{x!} - 1 \right] = \frac{e^{\theta t} - 1}{e^\theta - 1}$$

The moment generating function of the ZTPD (2.1.1) is thus given by

$$M_X(t) = E(e^{tX}) = \frac{e^{\theta e^t} - 1}{e^\theta - 1}$$

Moments of ZTPLD

The r th factorial moment of the ZTPLD (2.2.1) can be obtained as

$$\mu_{(r)}' = E[X^{(r)} | \lambda], \text{ where } X^{(r)} = X(X-1)(X-2)\dots(X-r+1)$$

Using (2.2.4), we get

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

Taking $x + r$ in place of x , we get

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

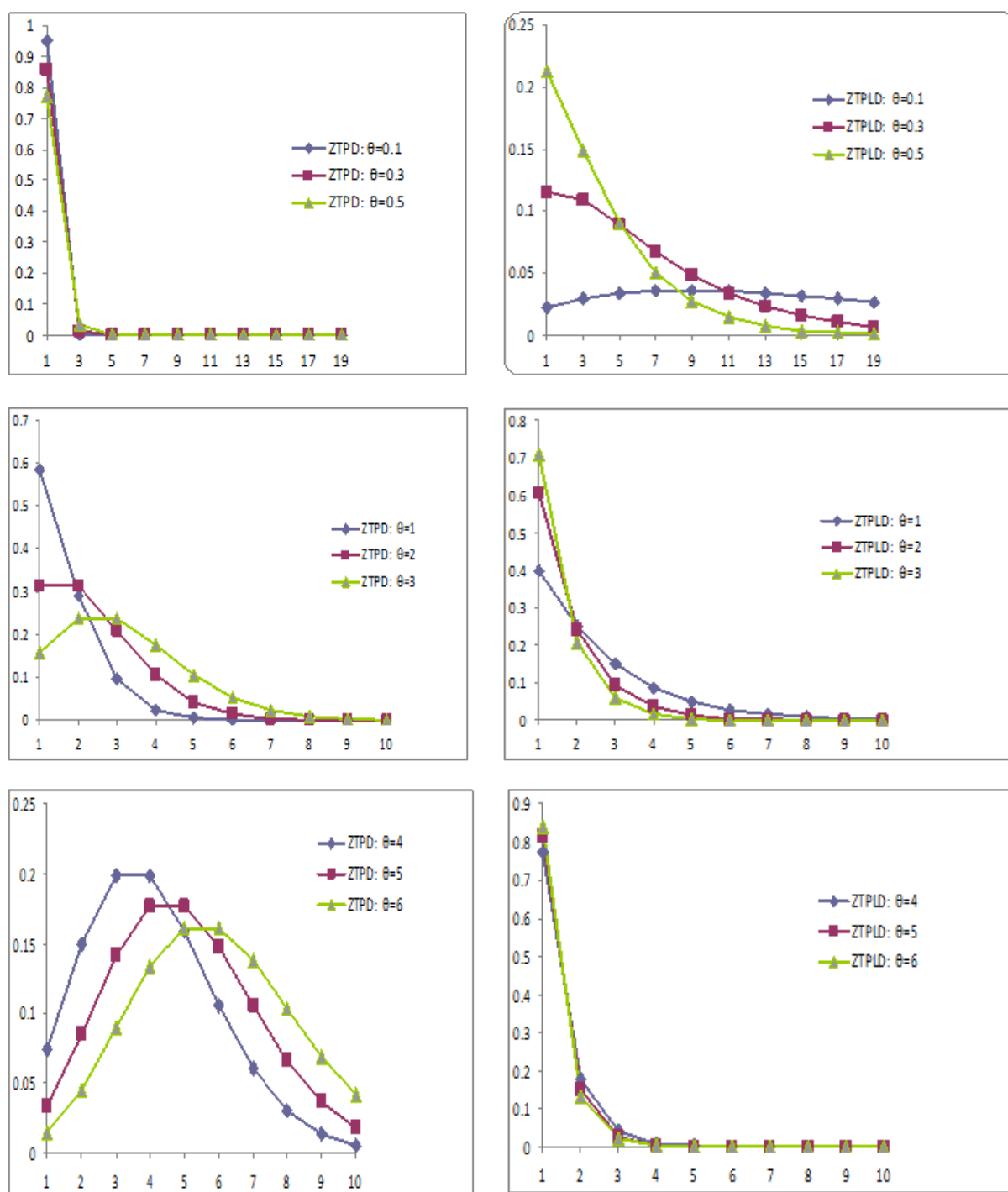


Figure 1 Graph of probability functions of ZTPD and ZTPLD for different values of their parameter. The left hand side graphs are for ZTPD and right hand side graphs are for ZTPLD.

It is obvious that the expression within the bracket is $\lambda + r$ and hence, we have

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

Using gamma integral and little algebraic simplification, we get finally, a general expression for the r th factorial moment of the ZTPLD (2.2.1) as

$$\mu_{(r)}' = \frac{r!(\theta + 1)^2 (r + \theta + 1)}{\theta^r (\theta^2 + 3\theta + 1)}; r = 1, 2, 3, \dots \quad (3.2.1)$$

Substituting $r = 1, 2, 3$, and 4 in (3.2.1), the first four factorial moment can be obtained and then using the relationship between factorial moments and moments about origin, the first four moments

about origin of the ZTPLD (2.2.1) are given by

$$\begin{aligned}\mu_1' &= \frac{(\theta+1)^2(\theta+2)}{\theta(\theta^2+3\theta+1)} \\ \mu_2' &= \frac{(\theta+1)^2(\theta^2+4\theta+6)}{\theta^2(\theta^2+3\theta+1)} \\ \mu_3' &= \frac{(\theta+1)^2(\theta^3+8\theta^2+24\theta+24)}{\theta^3(\theta^2+3\theta+1)} \\ \mu_4' &= \frac{(\theta+1)^2(\theta^4+16\theta^3+78\theta^2+168\theta+120)}{\theta^4(\theta^2+3\theta+1)}\end{aligned}$$

Generating function: The probability generating function of the ZTPLD (2.2.1) is obtained as

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

The moment generating function of the ZTPLD (2.2.1) is thus given by

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

A Simple method of finding moments of ZTPLD

Using (2.2.4), the r th moment about origin of ZTPLD (2.2.1) can be obtained as

$$\begin{aligned}\mu_r' &= E\left[E(X^r | \lambda)\right] \\ &= \frac{\theta^2}{\theta^2+3\theta+1} \int_0^{\infty} \sum_{x=1}^{\infty} x^r \frac{e^{-\lambda} \lambda^{x-1}}{(x-1)!} \cdot [(\theta+1)\lambda + (\theta+2)] e^{-\theta\lambda} d\lambda\end{aligned}\quad (4.1)$$

It is obvious that the expression under the bracket in (4.1) is the r th moment about origin of the SBPD. Taking $r=1$ in (4.1) and using the first moment about origin of the SBPD, the first moment about origin of the ZTPLD (2.2.1) is obtained as

$$\begin{aligned}\mu_1' &= \frac{\theta^2}{\theta^2+3\theta+1} \int_0^{\infty} (\lambda+1) [(\theta+1)\lambda + (\theta+2)] e^{-\theta\lambda} d\lambda \\ &= \frac{(\theta+1)^2(\theta+2)}{\theta(\theta^2+3\theta+1)}\end{aligned}\quad (4.2)$$

Again taking $r=2$ in (4.1) and using the second moment about origin of the SBPD, the second moment about origin of the ZTPLD (2.2.1) is obtained as

$$\begin{aligned}\mu_2' &= \frac{\theta^2}{\theta^2+3\theta+1} \int_0^{\infty} (\lambda^2+3\lambda+1) [(\theta+1)\lambda + (\theta+2)] e^{-\theta\lambda} d\lambda \\ &= \frac{(\theta+1)^2(\theta^2+4\theta+6)}{\theta^2(\theta^2+3\theta+1)}\end{aligned}\quad (4.3)$$

Similarly, taking $r=3$ and 4 in (4.1) and using the respective moments of SBPD, we get finally, after a little simplification, the third and the fourth moments about origin of the ZTPLD (2.2.1) as

$$\mu_3' = \frac{(\theta+1)^2(\theta^3+8\theta^2+24\theta+24)}{\theta^3(\theta^2+3\theta+1)}\quad (4.4)$$

$$\mu_4' = \frac{(\theta+1)^2(\theta^4+16\theta^3+78\theta^2+168\theta+120)}{\theta^4(\theta^2+3\theta+1)}\quad (4.5)$$

Estimation of parameter

Estimation of parameter of ZTPD

Maximum likelihood estimate (MLE): Let x_1, x_2, \dots, x_n be a random sample of size n from the ZTPD (2.1.1). The MLE $\hat{\theta}$ of θ of ZTPD (2.1.1) is given by the solution of the following non linear equation.

$$e^{\theta}(\bar{x} - \theta) - \bar{x} = 0, \text{ where } \bar{x} \text{ is the sample mean}$$

Method of moment estimate (MOME): Let x_1, x_2, \dots, x_n be a random sample of size n from the ZTPD (2.1.1). Equating the first population moment about origin to the corresponding sample moment, the MOME $\hat{\theta}$ of θ of ZTPD (2.1.1) is the solution of the following non linear equation.

$$e^{\theta}(\bar{x} - \theta) - \bar{x} = 0, \text{ where } \bar{x} \text{ is the sample mean}$$

Thus both MLE and MOME give the same estimate of the parameter θ of ZTPD (2.1.1).

Estimation of parameter of ZTPLD

Maximum likelihood estimate (MLE): Let x_1, x_2, \dots, x_n be a random sample of size n from the ZTPLD (2.2.1) and let f_x be the observed frequency in the sample corresponding to $X = x (x = 1, 2, 3, \dots, k)$ such that $\sum_{x=1}^k f_x = n$, where k is the largest observed value having non-zero frequency. The likelihood function L of the ZTPLD (2.2.1) is given by

$$L = \left(\frac{\theta^2}{\theta^2+3\theta+1} \right)^n \frac{1}{(\theta+1)^{\sum_{x=1}^k x f_x}} \prod_{x=1}^k (x+\theta+2)^{f_x} \quad (5.1.1)$$

The log likelihood function is given by

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

and the log likelihood equation is thus obtained as

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

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

$$\frac{2n}{\theta} - \frac{n(2\theta + 3)}{\theta^2 + 3\theta + 1} - \frac{n\bar{x}}{\theta + 1} + \sum_{x=1}^k \frac{f_x}{x + \theta + 2} = 0 \quad (5.1.2)$$

where \bar{x} 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. Ghitany et al.²⁰ showed that the MLE $\hat{\theta}$ of θ is consistent and asymptotically normal.

Method of moment estimate (MOME): Let x_1, x_2, \dots, x_n be a random sample of size n from the ZTPLD (2.2.1). Equating the first population moment about origin to the corresponding sample moment, the MOME $\hat{\theta}$ of θ of ZTPLD (2.2.1) is the solution of the following cubic equation.

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

where \bar{x} is the sample mean. Ghitany et al. [20] showed that the MOME $\hat{\theta}$ of θ is consistent and asymptotically normal.

Applications

In this section, both ZTPD and ZTPLD have been fitted to a number of data-sets using maximum likelihood estimates relating to demography, biological sciences, and social sciences to test their goodness of fit and it has been observed that in most of the cases ZTPLD gives much closer fit than ZTPD and in some cases ZTPD gives much closer fit than ZTPLD.

Mortality

Mortality does not depend only on biological factors; it depends upon the prevailing health conditions, medical facilities, the socio-economic and cultural factors. In developing and under-developed countries, the mortality among infants and children is found much higher than that among youngsters. The high infant mortality has thrown a serious challenge to the medical personnel and is considered as one of the sensitive position of existing medical and health facilities in the population.

In this section, an attempt has been made to test the suitability of ZTPD and ZTPLD in describing the neonatal deaths as well as of infant and child deaths experienced by mothers. The data-sets considered here are the data of Sri Lanka and India. The data-sets of Meegama²¹ have been used as the data of Sri Lanka whereas the data from the survey conducted by Lal²² and the survey of Kadam Kuan, Patna, conducted in 1975 and referred to by Mishra²³ have been used as the data of India. It is obvious from the fittings of ZTPD and the ZTPLD that ZTPLD gives much closer fits in almost all cases except in Table 2. Hence, in case of demographic data, ZTPLD is a better alternative than ZTPD to model count data.

Table 1 The number of mothers of the rural area having at least one live birth and one neonatal death

Number of neonatal deaths	Observed number of mothers	Expected Frequency	
		ZTPD	ZTPLD
1			
2	409	399.7	408.1
3	88	102.3	89.4
4	19	17.5	19.3
5	5	2.2	4.1
	1	0.3	1.1
Total	522	522.0	522.0
ML Estimate		$\hat{\theta} = 0.512047$	$\hat{\theta} = 4.199697$
χ^2		3.464	0.145
d.f.		1	2
P-value		0.0627	0.9301

Table 2 The number of mothers of the estate area having at least one live birth and one neonatal death

Number of Neonatal Deaths	Observed Number of Mothers	Expected Frequency	
		ZTPD	ZTPLD
1	71	66.5	
2	32	35.1	72.3
3	7		28.4
4	5	12.3	10.9
5	3	3.3	4.1
		0.8	2.3
Total	118	118.0	118.0
ML Estimate		$\hat{\theta} = 1.055102$	$\hat{\theta} = 2.049609$
χ^2		0.696	2.274
d.f.		1	2
P-value		0.4041	0.3208

Table 3 The number of mothers of the urban area with at least two live births by the number of infant and child deaths

Number of Infant and Child Deaths	Observed Number of Mothers	Expected Frequency	
		ZTPD	ZTPLD
1	176	164.3	171.6
2	44	61.2	51.3
3	16	15.2	15.0
4	6	2.8	4.3
5	2	0.5	1.7
Total	244	244.0	244.0
ML Estimate		$\hat{\theta} = 0.744522$	$\hat{\theta} = 2.209411$
χ^2		7.301	1.882
d.f.		1	2
P-value		0.0069	0.3902

Table 4 The number of mothers of the rural area with at least two live births by the number of infant and child deaths

Number of Neonatal Deaths	Observed Number of Mothers	Expected Frequency	
		ZTPD	ZTPLD
1	745	708.9	738.1
2	212	255.1	214.8
3	50	61.2	61.3
4	21	11.0	17.2
5	7	1.6	4.8
6	3	0.2	1.8
Total	1038	1038.0	1038.0
ML Estimate		$\hat{\theta} = 0.719783$	$\hat{\theta} = 3.007722$
χ^2		37.046	4.773
d.f.		2	3
P-value		0.0	0.1892

Table 5 The number of literate mothers with at least one live birth by the number of infant deaths

Number of Neonatal Deaths	Observed Number of Mothers	Expected Frequency	
		ZTPD	ZTPLD
1	683	659.0	674.4
2	145	177.4	154.1
3	29	31.8	34.6
4	11	4.3	7.7
5	5	0.5	2.2
Total	873	873.0	873.0
ML Estimate		$\hat{\theta} = 0.538402$	$\hat{\theta} = 4.00231$
		8.718	5.310
d.f.		1	2
P-value		0.0031	0.0703

Table 6 The number of mothers of the completed fertility having experienced at least one child death

Number of Child Deaths	Observed Number of Mothers	Expected Frequency	
		ZTPD	ZTPLD
1	89	76.8	83.4
2	25	39.9	32.3
3	11	13.8	12.2
4	6	3.6	4.5
5	3	0.7	1.6
6	1	0.2	0.9
Total	135	135.0	135.0
ML Estimate		$\hat{\theta} = 1.038289$	$\hat{\theta} = 2.089084$
χ^2		7.90	3.428
d.f.		1	2
P-value		0.0049	0.1801

Table 7 The number of mothers having at least one neonatal death

Number Of Neonatal Deaths	Observed Number of Mothers	Expected Frequency	
		ZTPD	ZTPLD
1	567		561.4
2	135	545.8	139.7
3	28	162.5	34.2
4	11	32.3	8.2
5	5	4.8	2.6
		0.6	
Total	746	746.0	746.0
ML Estimate		$\hat{\theta} = 0.595415$	$\hat{\theta} = 3.625737$
χ^2		26.855	3.839
d.f.		2	2
P-value		0.0	0.1467

Biological sciences

In this section, an attempt has been made to test the goodness of fit of both ZTPD and ZTPLD on many data- sets relating to biological sciences. It has been observed that ZTPLD gives much closer fit than

ZTPD in almost all cases except in Table 11 regarding the distribution of the number of leaf spot grade of Ichinose variety of Mulberry. Thus in biological sciences ZTPLD is a better model than ZTPD to model zero-truncated count data.

Table 8 Number of European red mites on apple leaves, reported by Garman.²⁴

Number of European Red Mites	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	38	28.7	36.1
2	17	25.7	20.5
3	10	15.3	11.2
4	9	6.9	3.1
5	3	2.5	1.6
6	2	0.7	0.8
7	1	0.2	0.8
8	0	0.1	
Total	80	80.0	80.0
ML Estimate		$\hat{\theta} = 1.791615$	$\hat{\theta} = 1.185582$
χ^2		9.827	2.467
d.f.		2	3
P-value		0.0073	0.4813

Table 9 Number of yeast cell counts observed per mm square, reported by Student.²⁵

Number of Yeast Cells Counts Per Mm Square	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	128	121.3	127.6
2	37	49.2	40.9
3	18	13.3	12.8
4	3	2.7	4.0
5	1	0.4	1.2
6	0	0.1	0.5
Total	187	187.0	187.0
ML Estimate		$\hat{\theta} = 0.811276$	$\hat{\theta} = 2.667323$
χ^2		5.228	1.034
d.f.		1	1
P-value		0.0222	0.3092

Table 10 The number of counts of flower heads as per the number of fly eggs reported by Finney & Varley.²⁶

Number of Fly Eggs	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	22	15.3	26.8
2	18	21.9	19.8
3	18	20.8	13.9
4	11	14.9	9.5
5	9	8.5	6.4
6	6	4.1	4.2
7	3	1.7	2.7
8	0	0.6	1.7
9	1	0.3	3.0
Total	88	88.0	88.0
ML Estimate		$\hat{\theta} = 2.860402$	$\hat{\theta} = 0.718559$
χ^2		6.677	3.743
d.f.		4	4
P-value		0.1540	0.4419

Table 11 The number of leaf spot grade of Ichinose variety of Mulberry, reported by Khurshid.²⁷

Number Of Leaf Spot Grade	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	18	14.2	23.0
2	15	18.7	16.3
3	10	16.5	11.1
4	14	10.9	7.3
5	13	9.7	12.4
Total	70	70.0	70.0
ML Estimate		$\hat{\theta} = 2.639984$	$\hat{\theta} = 0.781902$
χ^2		6.311	7.476
d.f.		3	3
P-value		0.0974	0.0582

Table 12 The number of leaf spot grade of Kokuso-20 variety of Mulberry, reported by Khurshid²⁷

Number of Leaf Spot Grade	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	37	28.5	36.7
2	16	26.7	21.4
3	15	16.7	12.0
4	8	7.8	6.6
5	8	4.2	7.3
Total	84	84.0	84.0
ML Estimate		$\hat{\theta} = 1.874567$	$\hat{\theta} = 1.130211$

Table Continued

Number of Leaf Spot Grade	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
χ^2		8.329	2.477
d.f.		2	3
P-value		0.0155	0.4795

Table 13 The number of counts of sites with particles from Immuno gold data, reported by Mathews and Appleton²⁸

Number of Sites With Particles	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	122	115.9	124.8
2	50	57.4	46.8
3	18	18.9	17.1
4	4	4.7	6.1
5	4	1.1	3.2
Total	198	198.0	198.0
ML Estimate		$\hat{\theta} = 0.990586$	$\hat{\theta} = 2.18307$
χ^2		2.14	0.51
d.f.		2	2
P-value		0.3430	0.7749

Table 14 The number of snowshoe hares counts captured over 7 days, reported by Keith & Meslow²⁹

Number of Times Hares Caught	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	184	176.6	182.6
2	55	66.0	55.3
3	14	16.6	16.4
4	4	3.1	4.8
5	4	0.7	1.9
Total	261	261.0	261.0
ML Estimate		$\hat{\theta} = 0.756171$	$\hat{\theta} = 2.863957$
χ^2		2.45	0.61
d.f.		1	2
P-value		0.1175	0.7371

Social sciences

In this section, an attempt has been made to test the goodness of fit test of both ZTPD and ZTPLD on many data-sets relating to social

sciences, such as migration, number of accidents and free-forming small group size. It has been observed that the ZTPD gives much closer fits than ZTPLD in almost all cases except the distribution of the number of household having at least one migrant in Table 15.

Table 15 Number of households having at least one migrant according to the number of migrants, reported by Singh and Yadav³⁰

Number of migrants	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	375	354.0	379.0
2	143	167.7	137.2
3	49	53.0	48.4
4	17	12.5	16.8
5	2	2.4	5.7
6	2	0.4	1.9
7	1	0.1	0.6
8	1	0.0	0.3
Total	590	590.0	590.0
ML Estimate		$\hat{\theta} = 0.947486$	$\hat{\theta} = 2.284782$
χ^2		8.933	1.031
d.f.		2	3
P-value		0.0115	0.7937

Table 16 Number of workers according to the number of accidents, reported by Mir & Ahmad³¹

Number of Accidents	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	2039	2034.2	2050.4
2	312	319.5	291.7
3	35	33.5	41.1
4	3	2.6	5.8
5	1	0.2	1.0
Total	2390	2390.0	2390.0
ML Estimate		$\hat{\theta} = 0.314125$	$\hat{\theta} = 6.749732$
χ^2		0.387	3.128
d.f.		1	1
P-value		0.5339	0.0769

Table 17 Number of counts of pairs of running shoes owned by 60 members of an athletics club, reported by Simonoff³²

Number of Pairs of Running Shoes	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	18	17.7	24.1
2	18	18.5	15.0
3	12	12.9	9.0
4	7	6.7	5.2
5	5	4.2	6.2

Table Continued

Number of Pairs of Running Shoes	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
Total	60	60.0	60.0
ML Estimate		$\hat{\theta} = 2.087937$	$\hat{\theta} = 1.004473$
χ^2		0.191	3.998
d.f.		2	3
P-value		0.9089	0.2617

Table 18 Number of free forming small group size according to the group size, reported by Coleman & James³³

Group Size	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	1486	1500.5	1592.8
2	694	669.6	551.8
3	195	199.2	186.5
4	37	44.4	61.9
5	10	7.9	20.3
6	1	1.3	9.6
Total	2423	2423.0	2423.0
ML Estimate		$\hat{\theta} = 0.892496$	$\hat{\theta} = 2.419103$
χ^2		2.702	66.155
d.f.		3	3
P-value		0.4399	0.0

Table 19 Number of free forming small group size according to the group size, reported by Coleman & James³³

Group Size	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	316	316.4	335.8
2	141	140.7	116.0
3	44	41.7	39.1
4	5	9.3	12.9
5	4	1.9	6.2
Total	510	510.0	510.0
ML Estimate		$\hat{\theta} = 0.889458$	$\hat{\theta} = 2.428125$
χ^2		0.558	12.481
d.f.		2	2
P-value		0.7565	0.0019

Table 20 Number of free forming small group size according to the group size, reported by Coleman & James³³

Group Size	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	306	302.5	322.5
2	132	139.5	114.6
3	47	42.9	39.7
4	10	9.9	13.5
5	2	2.1	6.8

Table Continued

Group Size	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
Total	497	497.0	497.0
ML Estimate		$\hat{\theta} = 0.922509$	$\hat{\theta} = 2.341269$
χ^2		0.834	8.220
d.f.		2	2
P-value		0.6590	0.0164

Table 21 Number of free forming small group size according to the group size, reported by Coleman & James³³

Group Size	Observed Frequency	Expected Frequency	
		ZTPD	ZTPLD
1	305	307.2	327.7
2	144	142.9	117.3
3	50	44.3	40.9
4	5	10.3	14.0
5	2	1.9	4.7
6	1	0.3	2.4
Total	507	507.0	507.0
ML Estimate		$\hat{\theta} = 0.930664$	$\hat{\theta} = 2.31943$
χ^2		2.376	17.806
d.f.		2	2
P-value		0.3048	0.0001

Conclusion

In this paper, the nature and behavior of ZTPD and ZTPLD have been studied by drawing different graphs of their probability functions for the different values of their parameter. A general expression for the r th factorial moment has been given and the first four moments about origin has been obtained. Also a very simple and easy method for finding moments of ZTPLD has been suggested. An attempt has been made to study the goodness of fit of both ZTPD and ZTPLD to count data relating to demography, biological sciences, and social sciences and it has been found that ZTPLD is a better model than the ZTPD in almost all data-sets relating to mortality and biological sciences whereas ZTPD is a better model than ZTPLD in almost all data-sets relating to social sciences. Thus, ZTPLD has an advantage over ZTPD for modeling zero-truncated count data in mortality and biological sciences whereas ZTPD has an advantage over ZTPLD for modeling zero-truncated count data in social sciences.

Acknowledgement

None.

Conflict of interest

No conflict of interest

References

1. Sankaran M. The discrete Poisson–Lindley distribution. *Biometrics*. 1970;26:145–149.
2. Shanker R, Hagos F. On Poisson–Lindley distribution and Its Applications to Biological Sciences. *Biometrics and Biostatistics International Journal*. 2015;2(4):1–5.
3. Lindley DV. Fiducial distributions and Bayes' Theorem. *Journal of the Royal Statistical Society*. 1958;20:102–107.
4. Ghitany ME, Atieh B, Nadarajah S. Lindley distribution and Its Applications. *Mathematics and Computers in Simulation*. 2008a;78(4):493–506.
5. Ghitany ME, Al-Mutairi DK. Estimation Methods for the discrete Poisson–Lindley distribution. *Journal of Statistical Computation and Simulation*. 2009;79(1):1–9.
6. Shanker R, Mishra A. A two-parameter Poisson–Lindley distribution. *International Journal of Statistics and Systems*. 2014;9(1):79–85.
7. Shanker R, Mishra A. A two-parameter Lindley distribution. *Statistics in Transition new Series*. 2013 a;14(1):45–56.
8. Shanker R, Mishra A. A quasi Poisson–Lindley distribution (submitted). 2015.
9. Shanker R, Mishra A. A quasi Lindley distribution. *African journal of Mathematics and Computer Science Research*. 2013b;6(4):64–71.
10. Shanker R, Sharma S, Shanker R. A Discrete two-Parameter Poisson Lindley Distribution. *Journal of Ethiopian Statistical Association*. 2012;21:15–22.
11. Shanker R, Sharma S, Shanker R. A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics*. 2013;4:363–368.

12. Shanker R, Tekie AL. A new quasi Poisson–Lindley distribution. *International Journal of Statistics and Systems*. 2014;9(1):87–94.
13. Shanker R, Amanuel AG. A new quasi Lindley distribution. *International Journal of Statistics and Systems*. 2013;8(2):143–156.
14. Shanker R, Hagos F, Sujatha S. On modeling of Lifetimes data using exponential and Lindley distributions. *Biometrics and Biostatistics International Journal*. 2015;2(5):1–7.
15. Plackett RL. The truncated Poisson– distribution. *Biometrics*. 1953;9:185–188.
16. David FN, Johnson NL. The truncated Poisson. *Biometrics*. 1952;8:275–285.
17. Cohen AC. An extension of a truncated Poisson distribution. *Biometrics*. 1960a;16:446–450.
18. Cohen AC. Estimation in a truncated Poisson distribution when zeros and some ones are missing. *Journal of American Statistical Association*. 1960b;55:342–348.
19. Tate RF, Goen RL. MVUE for the truncated Poisson distribution. *Annals of Mathematical Statistics*. 1958;29:755–765.
20. Ghitany ME, Al-Mutairi DK, Nadarajah S. Zero–truncated Poisson–Lindley distribution and its Applications. *Mathematics and Computers in Simulation*. 2008b;79(3):279–287.
21. Meegama SA. *Socio–economic determinants of infant and child mortality in Sri Lanka, an analysis of postwar experience*. International Statistical Institute (World Fertility Survey), Netherland. 1980.
22. Lal DN. *Patna in 1955 A Demographic Sample Survey*. Demographic Research Center, Department of Statistics, Patna University, Patna, India. 1955.
23. Mishra A. *Generalizations of some discrete distributions*. Unpublished Ph.D thesis, Patna University, Patna, India. 1979.
24. Garman P. The European red mites in Connecticut apple orchards. *Connecticut Agri Exper Station Bull*. 1923;252:103–125.
25. Student. On the error of counting with a haemocytometer. *Biometrika*. 1907;5(3):351–360.
26. Finney DJ, Varley GC. An example of the truncated Poisson distribution. *Biometrics*. 1955;11:387–394.
27. Khursid AM. On size–biased Poisson distribution and its use in zero truncated case. *J KSLAM*. 2008;12(3):153–160.
28. Matthews JN, Appleton DR. An application of the truncated Poisson distribution to Immunogold assay. *Biometrics*. 1993;49(2):617–621.
29. Keith LB, Meslow EC. Trap response by snowshoe hares. *Journal of Wildlife Management*. 1968;32:795–801.
30. Singh SN, Yadava KN. Trends in rural out–migration at household level. *Rural Demogr*. 1981;8(1):53–61.
31. Mir KA, Ahmad M. Size–biased distributions and their applications. *Pak J Statist*. 2009;25(3):283–294.
32. Simmonoff JS. *Analyzing Categorical data*. Springer, New York. 2003.
33. Coleman JS, James J. The equilibrium size distribution of freely–forming groups. *Sociometry*. 1961;24(1):36–45.