ZERO-TRUNCATED NEW QUASI POISSON-LINDLEY DISTRIBUTION AND ITS APPLICATIONS

RAMA SHANKER
KAMLESH KUMAR SHUKLA
Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea

A zero-truncated new quasi Poisson-Lindley distribution (ZTNQPLD), which includes zero-truncated Poisson-Lindley distribution (ZTPLD) as a particular case, has been studied. Its probability mass function has also been obtained by compounding size-biased Poisson distribution (SBPD) with an assumed continuous distribution. The r^{th} factorial moment of ZTNQPLD have been derived and hence its raw moments and central moments have been presented. The expressions for coefficient of variation, skewness, kurtosis, and index of dispersion have been given and their nature and behavior have been studied graphically. The method of maximum likelihood estimation has been discussed for estimating the parameters of ZTNQPLD. Finally, the goodness of fit of ZTNQPLD has been discussed with some datasets and the fit has been found better as compared with zero – truncated Poisson distribution (ZTPD) and zero- truncated Poisson- Lindley distribution (ZTPLD).

Keywords: Zero-truncated distribution, New quasi Poisson-Lindley distribution, compounding, moments, Maximum Likelihood estimation, Goodness of fit.

1. INTRODUCTION

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

$$P_1(x;\theta)=\frac{P_0(x;\theta)}{1-P_0(0;\theta)} \quad ;x=1,2,3,... \tag{1.1}$$

In probability theory, zero-truncated distributions are certain discrete distributions having support the set of positive integers. Zero-truncated distributions are suitable models for modeling data when the data to be modeled originate from a mechanism which generates data excluding zero counts. Zero-truncated distributions are useful in situations such as the number of households having at least on migrant, the number of flower heads per fly egg, the number of claims per claimant, the number of occupants per car etc.

Shanker and Tekie (2014) has obtained the new quasi Poisson-Lindley distribution (NQPLD) defined by its pmf

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

The NQPLD is a Poisson mixture of new quasi Lindley distribution (NQLD) introduced by Shanker and Amanuel (2013) having pdf

$$f_1(x;\theta,\alpha)=\frac{\theta^2}{\theta^2+\alpha}\left(\theta+\alpha x\right)e^{-\theta x};x>0,\theta>0,\theta^2+\alpha>0 \tag{1.3}$$

Shanker and Tekie (2014) have discussed the mathematical and statistical properties, estimation of parameters and applications of NQPLD to model count data. It can be easily verified that at $\alpha=\theta$, NQPLD (1.2) reduces to Poisson-Lindley distribution (PLD) introduced by Sankaran (1970) having pmf

$$P_2(x;\theta)=\frac{\theta^2\left(x+\theta+2\right)}{\left(\theta+1\right)^{x+3}};x=0,1,2,...,\theta>0 \tag{1.4}$$

It should be noted that PLD is also a Poisson mixture of Lindley distribution, introduced by Lindley (1958). Shanker and Hagos (2015) have discussed the applications of PLD for modeling data from biological sciences and observed that in biological science it a suitable model.

The main motivation of considering ZTNQPLD is two-folds: (1) NQLD gives better fit than both exponential and Lindley distributions for modeling lifetime data and NQPLD gives better fit than both Poisson and Poisson-Lindley distributions for count data. It is of interest to examine the goodness of fit of ZTNQPLD and is expected to give better fit than both ZTPD and ZTPLD in count datasets excluding zero-counts and (2) since both ZTPD and ZTPLD are of one parameter; ZTNQPLD is expected to give better fit than both ZTPD and ZTPLD due to the presence of additional parameter.

In this paper, a zero-truncated new quasi Poisson-Lindley distribution (ZTNQPLD), of which zero-truncated Poisson-Lindley distribution (ZTPLD) is a particular case, has been obtained by compounding size-biased Poisson distribution (SBPD) with an assumed continuous distribution. Its raw moments and central moments have been given. The expressions for coefficient of variation, skewness, kurtosis, and index of dispersion have been obtained and their nature and behavior have been discussed graphically. Maximum likelihood estimation has been discussed for estimating the parameters of ZTNQPLD. The goodness of fit of ZTNQPLD has been discussed with some datasets and the fit has been compared with zero -truncated Poisson distribution (ZTPD) and zero- truncated Poisson- Lindley distribution (ZTPLD). The main objective of this paper is to derive the pmf of ZTNQPLD as a size-biased Poisson mixture of a continuous distribution so that the raw moments can be easily obtained.

2. ZERO-TRUNCATED NEW QUASI POISSON-LINDLEY DISTRIBUTION

Using (1.1) and (1.2), the pmf of zero-truncated new quasi Poisson-Lindley distribution (ZTNQPLD) can be obtained as

$$P_3(x;\theta,\alpha)=\frac{\theta^2}{\theta^3+\theta^2+2\theta\alpha+\alpha}\frac{\alpha x+\left(\theta^2+\theta+\alpha\right)}{\left(\theta+1\right)^x}\quad ; \tag{2.1}$$

$$x=1,2,3,...,\ \theta>0,\theta^3+\theta^2+2\theta\alpha+\alpha>0$$

The nature and behavior of ZTNQPLD for varying values of parameters θ and α are explained graphically in Fig.1. It is obvious from the graphs of the pmf of ZTNQPLD that a small change in the parameter θ is playing a major role in the shapes whereas even a bigger change in the parameter α is not playing a major role in the shapes. This means that the parameter θ is the most dominating parameter in ZTNQPLD.

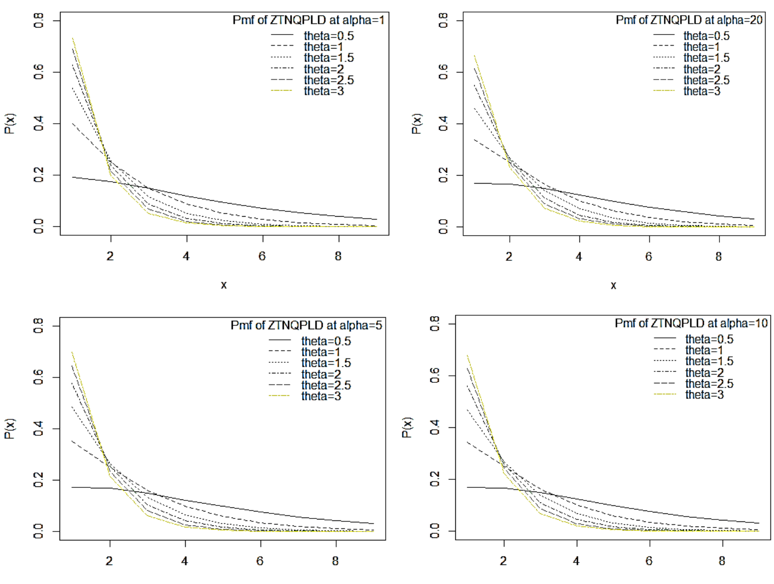

Fig.1. Graph of the Probability Mass Function of ZTNQPLD for Varying Values of Parameters

It has been found that it is tedious and very complicated to find the moments about origin of ZTNQPLD (2.1) directly. Therefore, to find moments of ZTNQPLD easily, firstly the pmf of ZTNQPLD has been obtained by considering a size-biased Poisson mixture of an assumed continuous distribution.

The ZTNQPLD (2.1) can also be obtained from size-biased Poisson distribution (SBPD) having pmf

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

when the parameter λ (> 0) of SBPD follows a continuous distribution having pdf

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

Thus, the pmf of ZTNQPLD can be obtained as

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

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

$$=\frac{\theta^2}{\left[\theta^3+\theta^2+2\theta\alpha+\alpha\right](x-1)!}\left[\frac{\alpha(\theta+1)\Gamma(x+1)}{(\theta+1)^{x+1}}+\frac{\left(\theta^2+\theta+\alpha\right)\Gamma(x)}{(\theta+1)^x}\right] \tag{2.4}$$

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

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

which is the pmf of ZTNQPLD with parameter θ and α as given in equation (2.1).

Recall that the pmf of zero-truncated Poisson- Lindley distribution (ZTPLD) given by

$$P_4(x;\theta)=\frac{\theta^2}{\theta^2+3\theta+1}\frac{x+\theta+2}{(\theta+1)^x}\quad ;x=1,2,3,....,\theta>0 \tag{2.5}$$

has been introduced by Ghitany et al. (2008). It can be easily verified that ZTPLD is a particular case of ZTNQPLD for $\alpha=\theta$. Shanker et al. (2015) have done extensive study on the comparison between ZTPD and ZTPLD with respect to their applications in datasets excluding zero-counts and showed that in demography and biological sciences ZTPLD gives better fit than ZTPD while in social sciences ZTPD gives better fit than ZTPLD.

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

$$P_5(x;\theta)=\frac{e^{-\theta}\,\theta^x}{\left(1-e^{-\theta}\right)x!}\quad ;x=1,2,3,....,\theta>0 \tag{2.6}$$

It should be noted that ZTPD has been used by various researchers to fit data excluding zero counts namely, Finnery and Varley (1955), Keith and Meslow (1968), Singh and Yadava (1971), Mathews and Appleton (1973), are some among others.

3. MOMENTS OF ZTNQPLD

The r^{th} factorial moment about origin, denoted by $\mu_{(r)}{}'$, of ZTNQPLD (2.1) can be obtained as

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

Using (2.4), we have

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

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

Taking $y=x-r$, we get

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

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

Using gamma integral and a little algebraic simplification, we get the expression for the r^{th} factorial moment about origin of ZTNQPLD (2.1) as

$$\mu_{(r)}' = \frac{r!(\theta+1)^2\left\{\theta^2+(r+1)\alpha\right\}}{\theta^r\left(\theta^3+\theta^2+2\theta\alpha+\alpha\right)}\ ; r=1,2,3,... \tag{3.1}$$

Substituting $r = 1,2,3$ and 4 in equation (3.1), the first four factorial moments about origin can be obtained and using the relationship between moments about origin and factorial moments about origin , the first four moments about origin $\left(\mu_1', \mu_2', \mu_3', \mu_4'\right)$ of ZTNQPLD (2.1) are obtained as

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

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

$$\mu_3' = \frac{(\theta+1)^2\left(\theta^4+6\theta^3+6\theta^2+2\theta^2\alpha+18\theta\alpha+24\alpha\right)}{\theta^3\left(\theta^3+\theta^2+2\theta\alpha+\alpha\right)}$$

$$\mu_4' = \frac{(\theta+1)^2\left(\theta^5+14\theta^4+36\theta^3+2\theta^3\alpha+24\theta^2+42\theta^2\alpha+144\theta\alpha+120\alpha\right)}{\theta^4\left(\theta^3+\theta^2+2\theta\alpha+\alpha\right)}$$

It can be easily verified that at $\alpha=\theta$, these expressions reduce to the corresponding expressions of ZTPLD

Since the expressions for the raw moments (moments about origin) μ_1', μ_2', μ_3' and μ_4' are big, it is tedious and complicated to obtain expressions for central moments (moments about mean) μ_2, μ_3 and μ_4. In the present paper, a mathematical software, 'MAPLE 16' has been used to obtain the expressions for central moments.

Using the relationship between moments about origin and moments about mean, the moments about mean of ZTNQPLD (2.1) are obtained as

$$\mu_2=\sigma^2=\frac{(\theta+1)^2\left(\theta^5+\theta^4+5\theta^3\alpha+4\theta^2\,\alpha+6\theta\,\alpha^2+2\alpha^3\right)}{\theta^2\left(\theta^3+\theta^2+2\theta\,\alpha+\alpha\right)^2}$$

$$\mu_3=\frac{(\theta+1)^2\begin{Bmatrix}\theta^9+4\theta^8+(7\alpha+5)\theta^7+(28\alpha+2)\theta^6+\left(16\alpha^2+33\alpha\right)\theta^5+\left(59\alpha^2+12\alpha\right)\theta^4\\+\left(12\alpha^3+54\alpha^2\right)\theta^3+\left(38\alpha^3+12\alpha^2\right)\theta^2+22\theta^3\alpha+4\alpha^3\end{Bmatrix}}{\theta^3\left(\theta^3+\theta^2+2\theta\,\alpha+\alpha\right)^3}$$

$$\mu_4=\frac{(\theta+1)^2\begin{Bmatrix}\theta^{13}+12\theta^{12}+(9\alpha+39)\theta^{11}+(114\alpha+55)\theta^{10}+\left(30\alpha^2+363\alpha+36\right)\theta^9\\+\left(389\alpha^2+492\alpha+9\right)\theta^8+\left(44\alpha^3+1147\alpha^2+306\alpha\right)\theta^7+\left(572\alpha^3+1376\alpha^2+72\alpha\right)\theta^6\\+\left(246\alpha^4+1497\alpha^3+720\alpha^2\right)\theta^5+\left(308\alpha^4+1508\alpha^3+132\alpha^2\right)\theta^4\\+\left(686\alpha^4+636\alpha^3\right)\theta^3+\left(554\alpha^4+96\alpha^3\right)\theta^2+192\theta^4\alpha+24\alpha^4\end{Bmatrix}}{\theta^4\left(\theta^3+\theta^2+2\theta\,\alpha+\alpha\right)^4}$$

It can be easily shown that at $\alpha=\theta$, these expressions reduce to the corresponding expressions of ZTPLD.

Finally, the coefficient of variation (C.V), coefficient of Skewness $\left(\sqrt{\beta_1}\right)$, coefficient of Kurtosis (β_2) and index of dispersion (γ) of ZTNQPLD (2.1) are obtained as

$$C.V.=\frac{\sigma}{\mu_1'}=\frac{\sqrt{\left(\theta^5+\theta^4+5\theta^3\alpha+4\theta^2\,\alpha+6\theta\,\alpha^2+2\alpha^3\right)}}{(\theta+1)\left(\theta^2+2\alpha\right)}$$

$$\sqrt{\beta_1}=\frac{\mu_3}{(\mu_2)^{3/2}}=\frac{\begin{Bmatrix}\theta^9+4\theta^8+(7\alpha+5)\theta^7+(28\alpha+2)\theta^6+\left(16\alpha^2+33\alpha\right)\theta^5+\left(59\alpha^2+12\alpha\right)\theta^4\\+\left(12\alpha^3+54\alpha^2\right)\theta^3+\left(38\alpha^3+12\alpha^2\right)\theta^2+22\theta^3\alpha+4\alpha^3\end{Bmatrix}}{(\theta+1)\left(\theta^5+\theta^4+5\theta^3\alpha+4\theta^2\,\alpha+6\theta\,\alpha^2+2\alpha^3\right)^{3/2}}$$

$$\beta_2=\frac{\mu_4}{\mu_2^{\,2}}=\frac{\begin{Bmatrix}\theta^{13}+12\theta^{12}+(9\alpha+39)\theta^{11}+(114\alpha+55)\theta^{10}+\left(30\alpha^2+363\alpha+36\right)\theta^9\\+\left(389\alpha^2+492\alpha+9\right)\theta^8+\left(44\alpha^3+1147\alpha^2+306\alpha\right)\theta^7+\left(572\alpha^3+1376\alpha^2+72\alpha\right)\theta^6\\+\left(246\alpha^4+1497\alpha^3+720\alpha^2\right)\theta^5+\left(308\alpha^4+1508\alpha^3+132\alpha^2\right)\theta^4\\+\left(686\alpha^4+636\alpha^3\right)\theta^3+\left(554\alpha^4+96\alpha^3\right)\theta^2+192\theta^4\alpha+24\alpha^4\end{Bmatrix}}{(\theta+1)^2\left(\theta^5+\theta^4+5\theta^3\alpha+4\theta^2\,\alpha+6\theta\,\alpha^2+2\alpha^3\right)^2}$$

$$\gamma=\frac{\sigma^2}{\mu_1'}=\frac{\theta^5+\theta^4+5\theta^3\alpha+4\theta^2\alpha+6\theta\alpha^2+2\alpha^3}{\theta\left(\theta^3+\theta^2+2\theta\alpha+\alpha\right)\left(\theta^2+2\alpha\right)}$$

These expressions reduce to the corresponding expressions of ZTPLD at $\alpha=\theta$.

The nature of coefficient of variation, coefficient of skewness, coefficient of kurtosis, and index of dispersion of ZTNQPLD (2.1) are shown graphically in Fig. 2.

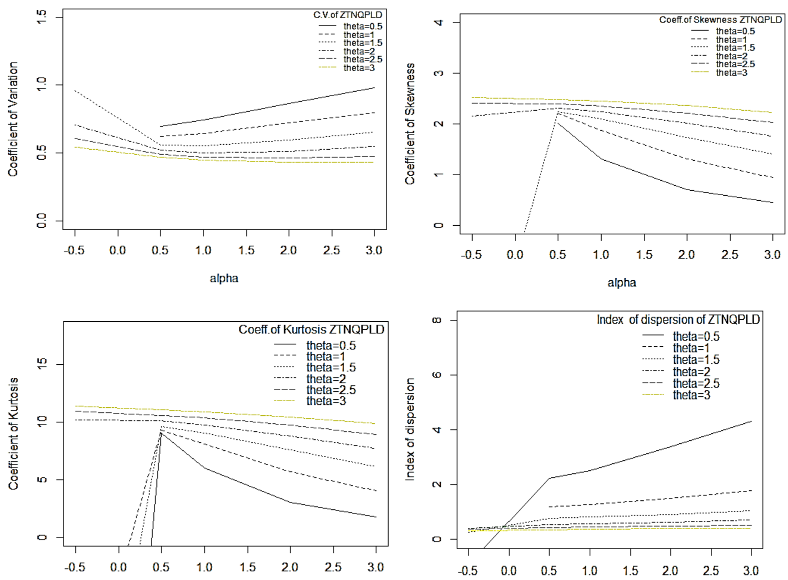

Fig.2. Coefficient of Variation (C.V.), Coefficient of Skewness, Coefficient of Kurtosis and Index of Dispersion Plot for Different Values of *α* and *θ*

It is also obvious from the graphs of Coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion that a small change in the parameter θ is playing a big role on the shape than a change in the parameter α.

4. MAXIMUM LIKELIHOOD ESTIMATION

Let $(x_1, x_2, \ldots, x_n)$ be a random sample of size n from the ZTNQPLD (2.1) and let f_x be the observed frequency in the sample corresponding to $X = x$ ($x = 1, 2, 3, \ldots, k$) 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 ZTNQPLD (2.1) is given by

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

.

The log likelihood function is thus obtained as

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

The maximum likelihood estimates $\left(\hat{\theta},\hat{\alpha}\right)$ of (θ, α) of ZTQPLD (2.1) are the solutions of the following log likelihood equations

$$\frac{\partial \log L}{\partial \theta}=\frac{2n}{\theta}-\frac{n\left(3\theta^2+2\theta+2\alpha\right)}{\theta^3+\theta^2+2\theta\alpha+\alpha}-\frac{n\bar{x}}{\theta+1}+\sum_{x=1}^{k}\frac{f_x}{\left[\alpha x+\left(\theta^2+\theta+\alpha\right)\right]}=0$$

$$\frac{\partial \log L}{\partial \alpha}=\frac{-n(2\theta+1)}{\theta^3+\theta^2+2\theta\alpha+\alpha}+\sum_{x=1}^{k}\frac{(x+1)f_x}{\left[\alpha x+\left(\theta^2+\theta+\alpha\right)\right]}=0$$

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

These two 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

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

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

$$\frac{\partial^2 \log L}{\partial \theta\,\partial \alpha}=\frac{n\left(4\theta^3+5\theta^2+2\theta\right)}{\left(\theta^3+\theta^2+2\theta\alpha+\alpha\right)^2}-\sum_{x=1}^{k}\frac{(x+1)f_x}{\left[\alpha x+\left(\theta^2+\theta+\alpha\right)\right]^2}=\frac{\partial^2 \log L}{\partial \alpha\,\partial \theta}$$

For the maximum likelihood estimates $\left(\hat{\theta},\hat{\alpha}\right)$ of (θ, α) of ZTNQPLD (2.1), following equations can be solved

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

where θ_0 and α_0 are the initial values of θ and α, respectively. These equations are solved iteratively till sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

5. GOODNESS OF FIT

The goodness of fit of ZTNQPLD has been discussed with several data sets and the fit has been compared with zero-truncated Poisson distribution (ZTPD) and zero-truncated Poisson-Lindley distribution (ZTPLD). In this section we present the goodness of fit of ZTPD, ZTPLD, and ZTNQPLD for four data sets. The first data set is due to Finney and Varley (1955) who gave counts of number of flower heads having number of fly eggs. The second data set is due to Singh and Yadav (1971) regarding the number of households having at least one migrant from households according to the number of migrants. The third data set is regarding the number of counts of sites with particles from Immunogold data, reported by Mathews and Appleton (1993). The fourth data set is regarding the number of snowshoe hares counts captured over 7 days, reported by Keith and Meslow (1968).

The main reasons for considering ZTNQPLD to fit these four datasets are as follows:

1. All four datasets considered here have been collected by respective authors which structurally excludes zero counts and the respective authors have proposed ZTPD as a suitable model.
2. Shanker et al. (2015) have conducted a comparative study on modeling of many datasets excluding zero counts from different fields of knowledge using ZTPD and ZTPLD and showed that ZTPLD is a suitable model than ZTPD in case of biological sciences and demography.
3. Since ZTPLD is a one parameter distribution and is a particular case of two-parameter ZTNQPLD for $\alpha = \theta$, it is of interest to compare the goodness of fit of ZTNQPLD with ZTPLD and ZTPD.

Table 5.1. Number of Flower Heads with Number of Fly Eggs, Reported by Finney and Varley (1955)

Number of fly eggs	Number of flower heads	Expected Frequency		
		ZTPD	ZTPLD	ZTNQPLD
1	22	15.3	26.8	24.9
2	18	21.8	19.8	20.3
3	18	20.8	14.0	14.8
4	11	14.9	9.5	10.1
5	9	8.5	6.3	6.6
6	6	4.0	4.2	4.2
7	3	1.7	2.7	2.7
8	0	0.6	1.7	1.6
9	1	0.4	3.0	2.8
Total	88	88.0	88.0	88.0
ML Estimate		$\hat{\theta} = 2.8604$	$\hat{\theta} = 0.7186$	$\hat{\theta} = 0.82437$ $\hat{\alpha} = 30.01272$
$\chi2$		6.648	3.780	2.38
d.f.		4	4	3
P-value		0.1557	0.4366	0.4974

Table 5.2. Number of households having at least one migrant according to the number of migrants, reported by Singh and Yadav (1971)

Number of migrants	Observed frequency	Expected Frequency		
		ZTPD	ZTPLD	ZTNQPLD
1	375	354.0	379.0	374.6
2	143	167.7	137.2	142.3
3	49	52.9	48.4	49.3
4	17	12.5	16.7	16.2
5	2	2.4	5.7	5.1
6	2	0.4	1.9	1.6
7	1	0.1	0.6	0.9
8	1	0.0	0.5	0.0
Total	590	590.0	590.0	590.0
ML Estimate		$\hat{\theta} = 0.9475$	$\hat{\theta} = 2.22848$	$\hat{\theta} = 2.72764$ $\hat{\alpha} = 24.99516$
$\chi2$		8.922	1.138	0.37
d.f.		2	3	2
P-value		0.0115	0.7679	0.8311

Table 5.3. The number of counts of sites with particles from Immunogold data, reported by Mathews and Appleton (1993)

Number of Sites with Particles	Observed Frequency	Expected Frequency		
		ZTPD	ZTPLD	ZTNQPLD
1	122	115.8	124.7	123.1
2	50	57.4	46.7	48.5
3	18	18.9	17.0	17.4
4	4	4.7	6.1	5.9
5	4	5.9	3.5	3.1
Total	198	198.0	198.0	198.0
ML Estimate		$\hat{\theta} = 0.9906$	$\hat{\theta} = 2.1831$	$\hat{\theta} = 2.60526$ $\hat{\alpha} = 25.0081$
$\chi2$		2.140	0.617	0.17
d.f.		2	2	1
P-value		0.3430	0.7345	0.6801

Table 5.4. The Number of Snowshoe Hares Counts Captured over 7 Days, Reported by Keith and Meslow (1968)

Number of times hares caught	Observed frequency	Expected Frequency		
		ZTPD	ZTPLD	ZTNQPLD
1	184	174.6	182.6	183.1
2	55	66.0	55.3	54.6
3	14	16.6	16.4	16.3
4	4	3.2	4.8	**4.8**
5	4	0.6	1.9	**2.2**
Total	261	261.0	261.0	261.0
ML Estimate		$\hat{\theta} = 0.7563$	$\hat{\theta} = 2.8639$	$\hat{\theta} = 2.34598$ $\hat{\alpha} = 0.012549$
$\chi2$		2.464	0.615	0.46
d.f.		1	2	1
P-value		0.1165	0.7353	0.4976

It is obvious from the goodness of fit based on chi-square (χ^2) and p-value of ZTNQPLD, ZTPLD and ZTPD that in tables 5.1 and 5.2 ZTNQPLD gives much better fit than ZTPD and ZTPLD while in tables 5.3 and 5.4, ZTPLD gives better fit than ZTNQPLD and ZTPD. One of the reasons of better fit by ZTPLD in datasets 5.3 and 5.4 may be that the additional parameter α in ZTNQPLD is not playing a major role in adjusting values of probabilities. This is also obvious from the probability plots of ZTNQPLD, ZTPLD and ZTPD for datasets 5.3 and 5.4 in figure 3.

The nature of the pmf's of the fitted distributions, ZTNQPLD, ZTPLD, and ZTPD for four datasets has been shown graphically in the Fig. 3.

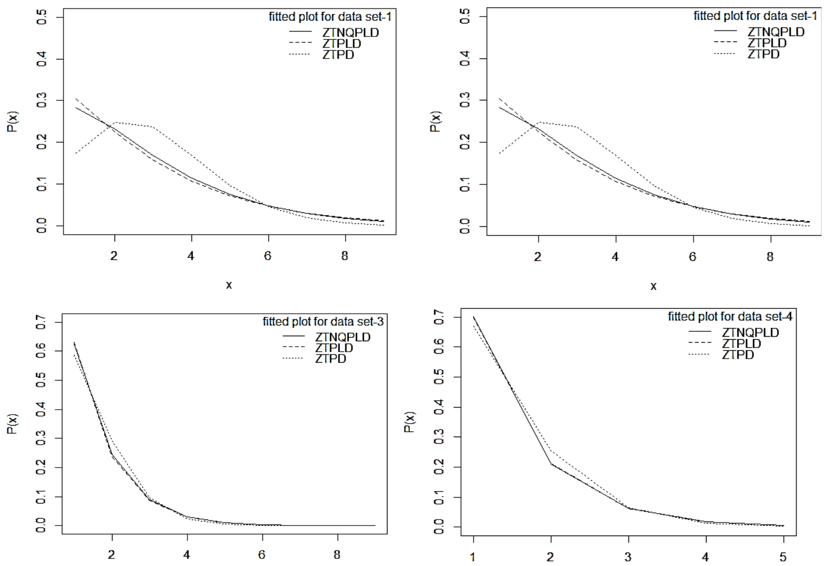

Fig. 3: Probability Plots of ZTNQPLD, ZTPLD and ZTPD for Fitted Data Sets in Tables 1, 2, 3, and 4.

6. CONCLUDING REMARKS

In this paper, a zero-truncated new quasi Poisson-Lindley distribution (ZTNQPLD), of which zero-truncated Poisson-Lindley distribution (ZTPLD) is a particular case, has been obtained by compounding size-biased Poisson distribution (SBPD) with a continuous distribution. Its moments, and moments based measures including coefficient of variation, skewness, kurtosis, and index of dispersion have been obtained and their nature and behavior have been studied graphically. Method of maximum likelihood estimation has been discussed for estimating the parameters of ZTNQPLD and the goodness of fit has been discussed with four datasets. In two datasets ZTNQPLD gives better fit than both ZTPD and ZTPLD while in two datasets ZTPLD gives better fit than ZTNQPLD and ZTPD and thus it can be concluded that ZTNQPLD is a competing with

ZTPLD. The nature of probability mass functions of fitted distribution for four data sets has also been discussed graphically.

ACKNOWLEDGEMENT

Authors are thankful to the anonymous reviewers for constructive comments which improved the quality of the paper.

REFERENCES

FINNEY, D.J. and VARLEY, G.C., 1955, An example of the truncated Poisson distribution, *Biometrics* 11, 387–394.

GHITANY, M.E., AL-MUTAIRI, D.K. and NADARAJAH, S., 2008, Zero-truncated Poisson-Lindley distribution and its Applications, *Mathematics and Computers in Simulation* 79 (3), 279–287.

KEITH, L.B. and MESLOW, E.C., 1968, Trap response by snowshoe hares, *Journal of Wildlife Management* 32, 795–801.

LINDLEY, D.V., 1958, Fiducial distributions and Bayes' Theorem, *Journal of the Royal Statistical Society* Series B, 20, 102–107.

MATHEWS, J.N.S. and Appleton, D.R., 1993, An application of the truncated Poisson distribution to Immunogold assay, *Biometrics* 49, 617–621.

SANKARAN, M., 1970, The discrete Poisson-Lindley distribution, *Biometrics* 26, 145–149.

SHANKER, R. and AMANUEL, H.G., 2013, A new quasi Lindley distribution, *International Journal of Statistics and Systems* 8(2), 143–156.

SHANKER, R. and TEKIE, A.L., 2014, A new quasi Poisson-Lindley distribution, *International Journal of Statistics and Systems* 9(1), 143–156.

SHANKER, R. and HAGOS, F., 2015, On Poisson-Lindley distribution and Its Applications to Biological Sciences, *Biometrics and Biostatistics International Journal*, 2(4), 1–5.

SHANKER, R., HAGOS, F., SUJATHA, S. and ABREHE, Y., 2015, On Zero-truncation of Poisson and Poisson-Lindley distributins and Their Applications, *Biometrics and Biostatistics International Journal* 2(6), 1–14.

SINGH, S.N. and YADAV, R.C., 1971, Trends in rural out-migration at household level, *Rural Demography* 8, 53–61.