International Journal of Statistical Distributions and Applications
2016; 2(4): 42-48
http://www.sciencepublishinggroup.com/j/ijsda
doi: 10.11648/j.ijsd.20160204.11
ISSN: 2472-3487 (Print); ISSN: 2472-3509 (Online)

On the Zero-One Inflated Poisson Distribution

Rafid Saeed Abdulrazak Alshkaki

General Requirements Department, Ahmed Bin Mohammed Military College, Doha, Qatar

Email address:
rafid@abmmc.edu.qa

To cite this article:
Rafid Saeed Abdulrazak Alshkaki. On the Zero-One Inflated Poisson Distribution. *International Journal of Statistical Distributions and Applications.* Vol. 2, No. 4, 2016, pp. 42-48. doi: 10.11648/j.ijsd.20160204.11

Received: August 18, 2016; **Accepted:** September 10, 2016; **Published:** December 10, 2016

Abstract: In many sampling involving non negative integer data, the zeros are observed to be significantly higher than the expected assumed model. Such models are called zero-one inflated models. The zero inflated Poisson distribution was recently considered and studied due to its empirical needs and application. In this paper, an extension to the case of zero inflated case is considered, namely, the zero and one inflated Poisson distribution, along with some of its structural properties, and estimation of its parameters using the methods of moments and maximum likelihood estimators were obtained with three empirical examples as well.

Keywords: Poisson Distribution, Zero-One Inflated Model, Maximum Likelihood Estimator, Moments Estimator, Inflated Poisson Distribution

1. Introduction

The Poisson distribution is a well-known non negative integer valued discrete that has been studied by many researchers due to its member's empirical applications. Often, a zero-inflated Poisson (ZIP) model occurs when the zero counts in a sampling Poisson data, has a higher number that expected if the data were Poisson distributed.

Lambert [1] studied the zero-inflated Poisson (ZIP) regression model in analyzing defects in manufacturing equipment using simulations, and concluded that regression zero-inflated Poisson models are not only easy to interpret, but they can also lead to more refined data analyses. van den Broek [2] presented a score test for zero-inflation, comparing the ZIP model with a constant proportion of excess zeros to a standard Poisson regression model. Sharma [3] proposed an ZIP distribution to study the trends for out-migration, using the method of moments and the maximum likelihood method. For estimating the parameters, and testing the method using data from the Rural Development and Population Growth survey undertaken in Varanasi, India, in 1978. Ridout et al [4] consider the problem of modelling count data with excess zeros and review some possible models including the ZIP, and considered some aspects of model fitting and inference, and gave an example from horticultural research for illustration.

Hall and Berenhaut [5] proposed a general score test for the null hypothesis that variance components associated with these random effects are zero for a ZIP model with random intercept. Jansakul and Hinde [6] extended the score test of van den Broek [2] to the more general situation and evaluated its performance using a simulation study. Gupta et al [7] studied a zero adjusted generalized Poisson distribution and developed a score test, with and without covariates, to determine whether such an adjustment is necessary, and gave examples to illustrate the results. Rodrigues-Motta et al [8] used a ZIP model to account for correlated genetic effects, and to analyze the number of clinical mastitis cases in Norwegian Red cows. Naya et al [9] compared the performance of four Poisson and ZIP regression models under four simulation scenarios when studying the number of black spots in Corriedale sheep.

Yang et al [10] proposed score tests for overdispersion based on the ZIP model for zero-inflated count data. Hall and Shen [11] proposed a robust estimation approach for ZIP regression. Saffari and Adnan [12] introduced a ZIP regression model on censored data and study the effects of right censoring in terms of parameters estimation and their standard errors via simulation and an example. Momeni [13] used the ZIP distribution model to fit a real data related to the number of accident insurance claims in Mazandaran Province. Kibria et al [14] proposed a ridge regression estimator for the ZIP model using a simulation study to compare the performance of the estimators. Mouatassim and Ezzahid [15] used Poisson and ZIP regression when studying private health insurance data. Sharma and Landge [16] used a ZIP regression model to estimate the accident frequencies for the heavy vehicle traffic accident data. Gupta et al [17] identify the optimal modeling strategy for highly skewed, zero- inflated data often observed in the clinical data of children with an EP disorder, using ZIP model among other models. Beckett et al [18] estimate the parameters of a ZIP distribution to model some natural calamities' data. Stewart [19] studied and compared the method of moment and maximum-likelihood estimator of various multi-point inflated Poisson distributions along with the standard Poisson distribution. Zhang et al [20] studied zero-and-one Inflated Poisson distribution, obtained maximum likelihood estimates of parameters using both the Fisher scoring and expectation-maximization algorithms, and provided Bootstrap confidence intervals for parameters of interest and testing hypotheses under large sample sizes. Hassanzadeh and Kazemi [21] extends regression modeling of positive count data to deal with excessive proportion of one counts and proposed one-inflated positive Poisson and negative binomial regression models and present some of their properties.

Alshkaki [22] introduced an extension to the zero-inflated power series distributions, in which the Poisson distribution is one of them, in which not only the number of frequencies with zeros is inflated, but the number of frequencies with ones are also inflated as well. He called such models zero-one inflated models, he studied its structure properties, as well as its relation to the standard and the zero inflated cases.

In this paper, we give in Section 2, the definition of the Poisson distribution, then, in Section 3, we introduce the class of zero-one inflated Poisson distribution, and some of its structural properties, namely, its mean, variance, and generating functions, were given in Section 4. Then in Section 5, we consider moment estimators method of its parameters, followed by the maximum likelihood estimators method for its parameters also in Section 6. Finally, empirical examples consist of estimation of the parameters of the zero-one inflated Poisson distribution as well as fitting its frequencies were presented in Section 7, using three different sets of data representing; accident insurance claims data, stillbirths of New Zealand white rabbits data, and heavy vehicle traffic accident data. Finally, some concluding remarks were given in Section 8.

2. Poisson Distribution

Let $\theta \epsilon \Omega = \{\theta; 0 < \theta < \omega\}$, where ω is the radius of convergence of e^{θ}, then the discrete random variable (rv) X having probability mass function (pmf);

$$P(X = x) = \frac{\theta^x}{x!} e^{-\theta}, \quad x = 0, 1, 2, \tag{2.1}$$

is said to have a Poison distribution (PD) with parameter space Ω, and will denoted that by writing $X \sim PD(\theta)$.

3. Zero-One Inflated Poisson Distribution

Let $X \sim PD(\theta)$ as given in (2.1), let $\alpha \in (0,1)$ be an extra proportion added to the proportion of zero of the rv X, and let $\beta \in (0,1)$ be an extra proportion added to the proportion of ones of the rv X, such that $0 < \alpha + \beta < 1$, then the rv Z defined by;

$$P(Z = z) = \begin{cases} \alpha + (1 - \alpha - \beta)\, e^{-\theta}, & z = 0 \\ \beta + (1 - \alpha - \beta)\theta e^{-\theta}, & z = 1 \\ (1 - \alpha - \beta)\frac{\theta^z}{z!} e^{-\theta}, & z = 2, 3, 4, \ldots \\ 0 \; otherwise, & \end{cases} \tag{3.1}$$

is said to have a zero-one inflated Poisson distribution, and will denoted that by writing $Z \sim ZOIPD(\theta; \alpha, \beta)$.

Note that, if $\beta \to 0$, the (3.1) reduces to the form of the zero-inflated Poisson distribution (ZIPD). Similarly, the case with $\alpha \to 0$ and $\beta \to 0$, reduces to the standard case of PD.

4. Some Structural Properties of the ZOIPD

Let the rv $Z \sim ZOIPD(\theta; \alpha, \beta)$, then it is easy to find that;

$$E(Z) = \beta + (1 - \alpha - \beta)\theta \tag{4.1}$$

$$= \beta + (1 - \alpha - \beta)\, E(X)$$

$$= \beta(1 - \theta) + E(Y)$$

where $X \sim PD(\theta)$ as given by (2.1), and $Y \sim ZOIPD(\theta; \alpha, 0)$, a zero-inflated PD. Similarly,

$$Var(Z) = \beta(1 - \beta) + (1 - 2\beta)(1 - \alpha - \beta)\theta + (\alpha + \beta)(1 - \alpha - \beta)\theta^2 \tag{4.2}$$

$$= \beta(1 - \beta) + (1 - 2\beta)(1 - \alpha - \beta)Var(X) + (\alpha + \beta)(1 - \alpha - \beta)[Var(X)]^2$$

$$= Var(Y) + \beta(1 - \beta) - (3 - 2\alpha - 2\beta)\beta\theta + \beta(1 + \beta)\theta^2$$

The probability generating function $G_Z(s)$ and the moment generating function $M_Z(t)$, are respectively, given by:

$$G_Z(s) = E(s^Z)$$

$$= \alpha + \beta s + (1-\alpha-\beta) e^{\theta(s-1)} \tag{4.3}$$

$$= \alpha + \beta s + (1-\alpha-\beta)\, G_X(s)$$

and

$$M_Z(t) = E(e^{tZ})$$

$$= \alpha + \beta e^t + (1-\alpha-\beta)\, e^{\theta(e^t-1)}$$

$$= \alpha + \beta e^t + (1-\alpha-\beta)\, M_X(t)$$

5. Moment Estimators of the Parameters

Using the moment generating function, or obtaining them directly, the second and three distribution moments about the origin for the ZOIPD can be found to be,

$$\mu_2' = \beta + (1-\alpha-\beta)\,\theta(1+\theta)$$

and,

$$\mu_3' = \beta + (1-\alpha-\beta)\,\theta(1+3\theta+\theta^2)$$

with the first moment about the origin as given by (4.1),

$$\mu_1' = \beta + (1-\alpha-\beta)\theta$$

Let z_1, z_2, , z_n be a random sample from $ZOIPD$ as given by (3.1), and let,

$$m_k' = \frac{\sum_{i=1}^n z_i^k}{n}, k = 1,2,3.$$

be their sample moments about the origin, then solving the following simultaneous:

$$m_1' = \beta + (1-\alpha-\beta) \tag{5.1}$$

$$m_2' = \beta + (1-\alpha-\beta)\,\theta(1+\theta)$$

$$m_3' = \beta + (1-\alpha-\beta)\theta(1+3\theta+\theta^2)$$

for θ, α, and β give us the following moments estimators (ME) for these parameters:

$$\hat{\alpha} = 1 - m_1' + (m_3' - 4m_2' + 3m_1')\left(\frac{m_2'-m_1'}{m_3'-3m_2'+2m_1'}\right)^2 \tag{5.2}$$

$$\hat{\beta} = m_1' - \frac{(m_2'-m_1')^2}{m_3'-3m_2'+2m_1'} \tag{5.3}$$

$$\hat{\theta} = \frac{m_3'-3m_2'+2m_1'}{m_2'-m_1'} \tag{5.4}$$

A derivation of (5.2) to (5.4) is given in the following: from (4.3), it is easily find that, the k^th factorial moment of Z defined by;

$$\mu_{[r]} = E[Z(Z-1)(Z-2)\ldots(Z-k+1)]$$

for k= 2, 3, , is given by,

$$\mu_{[k]} = (1-\alpha-\beta)\theta^k,$$

in particular,

$$\mu_{[2]} = (1-\alpha-\beta)\theta^2 \tag{5.5}$$

$$\mu_{[3]} = (1-\alpha-\beta)\theta^3 \tag{5.6}$$

and let for $k = 2, 3, \ldots$

$$m_{[k]}' = \frac{\sum_{i=1}^n z_i(z_i-1)(z_i-2)\ldots(z_i-k+1)}{n}, \tag{5.7}$$

be their sample factorial moments, then equating the distributional factorial moments $\mu_{[2]}$ and $\mu_{[3]}$ given by (5.5) and (5.6), respectively, with their sample factorial moments given by (5.7), we have,

$$m_{[2]}' = (1-\alpha-\beta)\theta^2 \tag{5.8}$$

$$m_{[3]}' = (1-\alpha-\beta)\theta^3,$$

and hence,

$$\hat{\theta} = \frac{m_{[3]}'}{m_{[2]}'} \tag{5.9}$$

From (5.8), we have that,

$$\frac{m_{[2]}'}{\theta^2} = 1-\alpha-\beta, \tag{5.10}$$

therefore, it follows from (5.1) with the using of (5.10) that,

$$\beta = m_1' - \frac{m_{[2]}'}{\theta}, \tag{5.11}$$

and hence with the using of (5.9), we have that,

$$\hat{\beta} = m_1' - \frac{{m_{[2]}'}^2}{m_{[3]}'} \tag{5.12}$$

Similarly, (5.10) can be written with the using of (5.12) as,

$$\hat{\alpha} = 1 - m_1' + \frac{{m_{[2]}'}^2}{m_{[3]}'} - \frac{{m_{[2]}'}^3}{{m_{[3]}'}^2} \tag{5.13}$$

Now, subsisting $m_{[2]}'$ and $m_{[3]}'$ in terms of m_1', m_2', and m_3' into (5.13), (5.12) and (5.9) we get the forms given by (5.2) to (5.4), respectively.

6. Maximum Likelihood Estimators of the Parameters

Let $z_1, z_2, \ldots, z_n$ be a random sample from $ZOIPD$ as given by (3.1), and let for i=1, 2, … n,

$$\alpha_i = \begin{cases} 1 \; if \; z_i = 0, \\ 0 \; otherwise \end{cases}$$

and

$$\beta_i = \begin{cases} 1 \; if \; z_i = 1, \\ 0 \; otherwise \end{cases}$$

then, for i=1, 2, … , n, (3.1) can be written in the following form;

$$P(Z_i = z_i) = \left\{\alpha + (1-\alpha-\beta)\, e^{-\theta}\right\}^{\alpha_i}\left\{\beta + (1-\alpha-\beta)\theta e^{-\theta}\right\}^{\beta_i}\left\{(1-\alpha-\beta)\frac{\theta^{z_i}}{z_i!}e^{-\theta}\right\}^{1-\alpha_i-\beta_i},$$

hence, the likelihood function $L = L(\theta, \alpha, \beta; z_1, z_2, \ldots, z_n)$ will be,

$$L = \prod_{i=1}^{n}\left\{\alpha + (1-\alpha-\beta)\, e^{-\theta}\right\}^{\alpha_i}\left\{\beta + (1-\alpha-\beta)\theta e^{-\theta}\right\}^{\beta_i}\left\{(1-\alpha-\beta)\frac{\theta^{z_i}}{z_i!}e^{-\theta}\right\}^{1-\alpha_i-\beta_i}$$

$$= \left\{\alpha + (1-\alpha-\beta)\, e^{-\theta}\right\}^{n_0}\left\{\beta + (1-\alpha-\beta)\theta e^{-\theta}\right\}^{n_1}\prod_{i=1}^{n}\left\{(1-\alpha-\beta)\frac{\theta^{z_i}}{z_i!}e^{-\theta}\right\}^{c_i}$$

where $c_i = 1 - \alpha_i - \beta_i$, $n_0 = \sum_{i=1}^{n}\alpha_i$ and $n_1 = \sum_{i=1}^{n}\beta_i$. Note that n_0 and n_1 represents, respectively, the number of zeros and the number of ones in the sample. Therefore,

$$log\, L = n_0 log\left\{\alpha + (1-\alpha-\beta)\, e^{-\theta}\right\} + n_1 log\left\{\beta + (1-\alpha-\beta)\theta e^{-\theta}\right\} + (n-n_0-n_1)log(1-\alpha-\beta) + \sum_{i=1}^{n} c_i\, z_i \log(\theta)$$
$$- \sum_{i=1}^{n} c_i \log(z_i!) - (\mathrm{n} - \mathrm{n}_0 - \mathrm{n}_1)\theta$$

It follows that,

$$\frac{\partial}{\partial\alpha} logL = \frac{n_0[1-e^{-\theta}]}{\alpha+(1-\alpha-\beta)e^{-\theta}} - \frac{n_1\theta e^{-\theta}}{\beta+(1-\alpha-\beta)\theta e^{-\theta}} - \frac{(n-n_0-n_1)}{1-\alpha-\beta} \tag{6.1}$$

And hence,

$$\frac{\partial^2}{\partial\alpha^2} logL = -\frac{n_0\left[1-e^{-\theta}\right]^2}{[\alpha+(1-\alpha-\beta)e^{-\theta}]^2} - \frac{n_1\left[\theta e^{-\theta}\right]^2}{[\beta+(1-\alpha-\beta)\theta e^{-\theta}]^2} - \frac{(n-n_0-n_1)}{(1-\alpha-\beta)^2}$$

Therefore, $\frac{\partial^2}{\partial\alpha^2} logL < 0$, which indicates that L has a local maximum at α. Similarly,

$$\frac{\partial}{\partial\beta} logL = -\frac{n_0[e^{-\theta}]}{\alpha+(1-\alpha-\beta)e^{-\theta}} + \frac{n_1[1-\theta e^{-\theta}]}{\beta+(1-\alpha-\beta)\theta e^{-\theta}} - \frac{(n-n_0-n_1)}{1-\alpha-\beta} \tag{6.2}$$

$$\frac{\partial^2}{\partial\beta^2} logL = -\frac{n_0\left[e^{-\theta}\right]^2}{[\alpha+(1-\alpha-\beta)e^{-\theta}]^2} - \frac{n_1\left[1-\theta e^{-\theta}\right]^2}{[\beta+(1-\alpha-\beta)\theta e^{-\theta}]^2} - \frac{(n-n_0-n_1)}{(1-\alpha-\beta)^2}$$

And hence, $\frac{\partial^2}{\partial\beta^2} logL < 0$, which indicates that L has a local maximum at β. And finally,

$$\frac{\partial}{\partial\theta}logL = -\frac{n_0(1-\alpha-\beta)e^{-\theta}}{\alpha+(1-\alpha-\beta)e^{-\theta}} + \frac{n_1(1-\alpha-\beta)\left(-\theta e^{-\theta}+e^{-\theta}\right)}{\beta+(1-\alpha-\beta)\theta e^{-\theta}} + \frac{\sum_{i=1}^{n} c_i\, z_i}{\theta}$$

$$-(n-n_0-n_1) \tag{6.3}$$

And,

$$\frac{\partial^2}{\partial\theta^2}logL = -n_0(1-\alpha-\beta)\frac{\partial}{\partial\theta}\left[\frac{e^{-\theta}}{\alpha+(1-\alpha-\beta)e^{-\theta}}\right] + n_1(1-\alpha-\beta)\frac{\partial}{\partial\theta}\left[\frac{-\theta e^{-\theta}+e^{-\theta}}{\beta+(1-\alpha-\beta)\theta e^{-\theta}}\right] - \frac{\sum_{i=1}^{n} c_i\, z_i}{\theta^2}$$

$$= -n_0(1-\alpha-\beta)\left[\frac{e^{-\theta}}{\alpha+(1-\alpha-\beta)e^{-\theta}}\right]\left[\frac{(1-\alpha-\beta)e^{-\theta}}{\alpha+(1-\alpha-\beta)e^{-\theta}}-1\right]$$
$$+ n_1(1-\alpha-\beta)\left\{\left[\frac{\theta e^{-\theta}-2e^{-\theta}}{\beta+(1-\alpha-\beta)\theta e^{-\theta}}\right] - (1-\alpha-\beta)\left[\frac{-\theta e^{-\theta}+e^{-\theta}}{\beta+(1-\alpha-\beta)\theta e^{-\theta}}\right]^2\right\} - \frac{\sum_{i=1}^{n} c_i\, z_i}{\theta^2}$$

Therefore, the local maximum of θ has to be explicitly examined.

Then letting $\frac{\partial}{\partial\alpha}logL = 0$, we have from (6.1) that

$$1-\alpha-\beta = \frac{(n-n_0-n_1)}{\frac{n_0}{p_0}\left(1-e^{-\theta}\right)-\frac{n_1}{p_1}\left(\theta e^{-\theta}\right)} \tag{6.4}$$

where,

$$p_0 = \alpha+(1-\alpha-\beta)e^{-\theta}\ , \tag{6.5}$$

and

$$p_1 = \beta+(1-\alpha-\beta)\theta e^{-\theta} \tag{6.6}$$

Setting $\frac{\partial}{\partial\theta}logL = 0$, then (6.3) reduces, with the using of (6.5) and (6.6), to;

$$-\frac{n_0}{p_0}(1-\alpha-\beta)e^{-\theta} + \frac{n_1}{p_1}(1-\alpha-\beta)(1-\theta)e^{-\theta} + \frac{\sum_{i=1}^{n} c_i z_i}{\theta} - (n-n_0-n_1) = 0 \tag{6.7}$$

Now, if we replace, p_0 and p_1 by their sample relative frequencies, i.e. by their sample estimates, the proportion of zeros and the proportion of ones in the sample, that is; $\widehat{p_0} = n_0/n$ and $\widehat{p_1} = n_1/n$, respectively, then (6.4) and (6.7) reduce to;

$$1-\alpha-\beta = \frac{(n-n_0-n_1)}{n\left(1-e^{-\theta}-\theta e^{-\theta}\right)} \tag{6.8}$$

and

$$-n(1-\alpha-\beta)e^{-\theta} + n(1-\alpha-\beta)(1-\theta)e^{-\theta} + \frac{\sum_{i=1}^{n} c_i z_i}{\theta} - (n-n_0-n_1) = 0 \tag{6.9}$$

Hence, using (6.8), (6.9) reduces to;

$$A(\theta) = 0 \tag{6.10}$$

Where,

$$A(\theta) = \left[\sum_{i=1}^{n} c_i\, z_i - (n-n_0-n_1)\theta\right]\left(e^{\theta}-1-\theta\right) - \theta^2$$

Hence, (6.10) can be solved by any numerical procedure, say, Newton Rapson, to obtain $\hat{\theta}$ numerically, i.e. $A(\hat{\theta}) = 0$.

Similarly, using (6.4), (6.5) and (6.8), α and β can be estimated to be;

$$\hat{\alpha} = \frac{1}{n}\left[n_0 - \frac{(n-n_0-n_1)}{e^{\hat{\theta}}-1-\hat{\theta}}\right] \tag{6.11}$$

and,

$$\hat{\beta} = \frac{1}{n}\left[n_1 - \frac{(n-n_0-n_1)\hat{\theta}}{e^{\hat{\theta}}-1-\hat{\theta}}\right] \tag{6.12}$$

Therefore, the maximum likelihood estimates (MLE) of the parameters θ , α and β are given by solving (6.10) numerically to find $\hat{\theta}$, with $\hat{\alpha}$ and $\hat{\beta}$ given by (6.9) and (6.10), respectively.

7. Empirical Examples

In this Section, three different sets of data will be used to estimate empirically the parameters of the zero-one inflated geometric distribution to illustrate the estimation results that are arrived to at Sections 5 and 6.

7.1. Accident Insurance Claims Data

This data set consists of the number of accident insurance claims on the basis of 16760 policies (number of policyholders) in Mazandaran Province.

Momeni [13] used the zero inflated Poisson (ZIP) distribution model to fit the number of accident insurance claims of 16760 policies (number of policyholders) in Mazandaran Province. Table (1) shows fitting a Momeni's ZIPD using the ML method and our ZOIPD fitting using both ML and MLE methods. The p-values for fitting a ZOIPD, as shown in the Table, indicate that this model provide also a higher accuracy than the ZIPD.

Table 1. *The parameters estimates of the ZIP and the ZOIP for the number accident insurance claims in Mazandaran Province.*

No. of Accident	Observed Frequencies	Expected Frequencies			
		ME		MLE	
		ZIP	ZOIP	ZIP	ZOIP
0	13772	13871	13772	13772	13662
1	2631	2616	2633	2620	2631
2	318	255	314	336	316
3	34	17	37	30	37
4+	5	1	4	2	4
Total	16760	16760	16760	16760	16760
Model Parameters	θ	0.1959	0.360001	0.25738	0.581356
	α	0.031	0.532613	0.21436	0.718869
	β	---	0.053044	---	0.0971889
	χ^2	49.3573	0.5457178	6.0438	0.5059015
	df	2	1	2	1
	p-value	0.0001	0.4601	0.0487	0.4769

7.2. Stillbirths of New Zealand White Rabbits Data

This data set consists of frequencies of stillbirths in 402 litters of New Zealand white rabbits, originally used by Morgan et al [22] when discussing score test statistic.

Table (2) shows the parameters estimates of the ZIP using the ML and the MLE methods, and that of the ZOIP using the MLE, for the data set of frequencies of stillbirths in 402 litters of New Zealand white rabbits, which indicates for the p-values of the goodness of fit the accuracy of the ZOIP.

Table 2. *The parameters estimates of the ZIP and the ZOIP for the number of Stillbirths of New Zealand White Rabbits Data.*

No. of Stillbirths	Observed Frequencies	Expected Frequencies			
		ME		MLE	
		ZIP	ZOIP	ZIP	ZOIP
0	314	328	312	314	314
1	48	22	56	36	48
2	20	23	11	28	16
3	7	16	9	15	11
4	5	8	6	6	6
5	2	4	3	2	3
6+	6	1	5	1	4
Total	402	402	402	402	402
Model Parameters	θ	2.06857	2.57747	1.57835	2.10417
	α	0.78955	0.773023	0.72419	0.768771
	β	---	0.119832	---	0.082012
	χ^2	63.90364	9.663758	35.71905	3.954545
	df	4	3	4	3
	p-value	0.0001	0.0217	0.0001	0.266419

7.3. Heavy Vehicle Traffic Accident Data

This data set consists of the accident frequencies for the heavy vehicle traffic accident data collected for the year 2010 by Sharma and Landge [16] from the National Highway No.6 commonly refer to as NH-6 or G.E. Road (Great Eastern Road) in India.

Sharma and Landge [16] used a ZIP regression model to estimate the accident frequencies for the heavy vehicle traffic accident data, as shown in table (1) below.

Using our ZOIP model gives accurate estimates for the given accident data frequencies using the ME and MLE methods for estimating its parameters, .as shown, in Table (3), from the goodness of fit p-values.

Table 3. *Estimates of the heavy vehicle traffic accident data frequencies using the ZINB Regression of Sharma and Landge [16] and the ZOIG.*

No. of Accidents	Observed Frequencies	Expected Frequencies			
		ME		MLE	
		ZIP	ZOIP	ZIP	ZOIP
0	55	65	56	55	55
1	26	7	24	16	26
2	4	8	4	13	4
3	3	7	4	8	4
4	3	5	4	3	2
5+	1	4	4	1	5
Total	96	96	96	96	96
Model Parameters	θ	2.51765	3.36449	1.69102	3.50686
	α	0.64332	0.580214	0.4764	0.567498
	β	---	0.222859	---	0.251832
	χ^2	58.4456	0.934524	31.60577	0.75
	df	3	2	3	2
	p-value	0.0001	0.6267	0.0001	0.6872

8. Conclusions

We consider estimation of the parameters of the zero-one inflated Poisson distribution by the method of moment estimators and maximum likelihood estimators. The method of maximum likelihood estimators is shown to have butter estimates on three real data sets representing; the An accident insurance claims data, a stillbirths of rabbits data, and a heavy vehicle traffic accident data. The zero-one inflated Poisson distribution is shown also to have a better fitting for that frequencies of the real data sets than the zero inflated Poisson distribution.

References

[1] Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. Technometrics, Vol. 34, No. 1, 1-14.

[2] Van den Broek, J. (1995). A score test for zero inflation in a Poisson distribution, Biometrics Vol. 51, No. 2, 738-743.

[3] Sharma, H. L. (1987). A Probability Distribution for Out-Migration. Janasamkhya; 5 (2): 95-101.

[4] Ridout, M., Dem´etrio, C. G. B., and Hinde, J. (1998). Models for count data with many zeros. International Biometric Conference, Cape Town, December 1998.

[5] Hall, D. B. and Berenhaut, K. S. (2000). Score tests for heterogeneity and overdispersion in zero-inflated Poisson and binomial regression models. The Canadian Journal of Statistics, Vol. 30, No. 3, 415-430.

[6] Jansakul, N. and Hinde, J. P. (2002). Score tests for Zero-Inflated Poisson Models. Computational Statistics and data Analysis, 40, 75-96.

[7] Gupta, P. L., Gupta, R. C. and Tripathi, R. C. (2004). Score Test for Zero Inflated Generalized Poisson Regression Model. Communications In Statistics, Theory and Methods, Vol. 33, No. 1, 47–64.

[8] Rodrigues-Motta, M., Gianola, D., Heringstad, B., Rosa, G. J., and Chang, Y. M. (2007). A Zero-Inflated Poisson Model For Genetic Analysis Of The Number Of Mastitis Cases In Norwegian Red Cows. Journal of Dairy Science, 90 (11), 5306-5315.

[9] Naya, H., Urioste, J. I., Chang, Y., Rodrigues-Motta, Kremer, M. R. and Gianola, D. (2008). A comparison between Poisson and zero-inflated Poisson regression models with an application to number of black spots in Corriedale sheep. Genetics Selection Evolution 40, 379–394.

[10] Yang, Z., Hardin, J W., and Addy, C. L. (2009). Testing Overdispersion In The Zero-Inflated Poisson Model. Journal of Statistical Planning and Inference, Volume 139, Issue 9, 3340–3353.

[11] Hall, D. B. and Shen, J. (2010). Robust Estimation for Zero-Inflated Poisson Regression, Scandinavian Journal of Statistics, Volume 37, Issue 2, 237–252.

[12] Saffari, S. E. and Adnan, R. (2010). Zero-Inflated Poisson Regression Models with Right Censored Count Data. MATEMATIKA, Volume 27, Number 1, 21–29.

[13] Momeni, F. (2011). The Generalized Power Series Distributions and their Application. The Journal of Mathematics and Computer Science Vol .2, No. 4, 691-697.

[14] Kibria, B. M. G., Månsson, K. and Shukur, G. (2011). A Ridge Regression estimator for the zero-inflated Poisson model. CESIS Electronic Working Paper Series, The Royal Institute of technology, Paper No. 257.

[15] Mouatassim, Y. and Ezzahid, E. H. (2012). Poisson regression and Zero-inflated Poisson regression: application to private health insurance data. European Actuarial Journal, Volume 2, Issue 2, 187–204.

[16] Sharma, A. K. and Landge, V. S. (2013). Zero Inflated Negative Binomial For Modeling Heavy Vehicle Crash Rate On Indian Rural Highway. International Journal of Advances in Engineering & Technology, Vol. 5, Issue 2, 292-301.

[17] Gupta, R., Marino, B. S., Cnota, J. F. and Ittenbach, R. F. (2013). Finding The Right Distribution For Highly Skewed Zero-Inflated Clinical Data. Epidemiology Biostatistics and Public Health, Volume 10, Number 1, 1-15.

[18] Beckett, S., Jee, J., Ncube, T., Pompilus, S., Washington, Q., Singh, A. and Pal, N. (2014). Zero-Inflated Poisson (ZIP) Distribution: Parameter Estimation and Applications to Model Data from Natural Calamities. Involve, Vol. 7, No. 6, 751–767.

[19] Stewart, P. (2014). A generalized inflated Poisson distribution. (Doctoral dissertation, Marshall University).

[20] Zhang, C., Tian, G. L., Ng, K. W. (2016). Properties of the Zero-And-One Inflated Poisson Distribution And Likelihood-Based Inference Methods. Statistics And Its Interface, 9(1), 11-32.

[21] Hassanzadeh, F., & Kazemi, I. (2015). Regression Modeling Of One-Inflated Positive Count Data. Statistical Papers, 1-19.

[22] Alshkaki, R. S. A. (2016). An Extension to the Zero-Inflated Generalized Power Series Distributions. To be presented at the International Conference on Statistical Distributions and Applications 2016, October 14-16, Niagara Falls, Canada.

[23] Morgan, B. J. T., Palmer, K. J., and Ridout, M. S. (2007). Negative Score Test Statistics. The American Statistician, Vol. 61, 285–288.