Research Article

Inference for zero inflated truncated power series family of distributions

Abstract

Zero-inflated data indicates that the data set contains an excessive number of zeros. The word zero-inflation is used to emphasize that the probability mass at the point zero exceeds than the one allowed under a standard parametric family of discrete distributions. Gupta et al.,[1] Murat & Szynal,[2] Patil & Shirke[3] have contributed to estimation and testing of the parameters involved in Zero Inflated Power Series Distributions. If the data set under study does not contain observations after some known point in the support, we have to modify Zero Inflated Power Series Distribution (ZIPSD) accordingly in order to get better inferential properties. Zero Inflated Truncated Power Series Distribution (ZITPSD) is one of the better options. In the present work we address problem of estimation for ZITPSD with more emphasis on statistical tests. We provide three asymptotic tests for testing the parameter of ZITPSD, using an unconditional (standard) likelihood approach, a conditional likelihood approach and the sample mean, respectively. The performance of first two tests has been studied for Zero Inflated Truncated Poisson Distribution (ZITPD). Asymptotic Confidence Intervals for the parameter are also provided. The model has been applied to a real life data.

Keywords: zero inflation, zero inflated power series distribution, zero inflated truncated power series distribution, zero inflated truncated poisson distribution

MK Patil

Padmabhushan Vasantraodada Patil Mahavidyalaya, India

Correspondence: MK Patil, Padmabhushan Vasantraodada Patil Mahavidyalaya, Kavathe Mahankal, Dist. Sangli, India, Email mkpatil_stats@rediffmail.com

Introduction

In certain applications involving discrete data, we come across data having frequency of an observation 'zero' significantly higher than the one predicted by the assumed model. The problem of high proportion of zeros has been an interest in data analysis and modeling. There are many situations in the medical field, engineering applications, manufacturing, economics, public health, road safety epidemiology and in other areas leading to similar situations. In highly automated production process, occurrence of defects is assumed to be Poisson. However, we get no defectives in many samples. This leads to excess number of zeros. Models having more number of zeros significantly are known as zero-inflated models.

In the literature, numbers of researchers have worked on family of zero-inflated power series distributions. Gupta et al.,[1] have studied the structural properties and point estimation of parameters of Zero-Inflated Modified Power Series distributions and in particular for zero-inflated Poisson distribution. Murat & Szynal[2] have studied the class of inflated modified power series distributions where inflation occurs at any of the support points. Moments, factorial moments, central moments, the maximum likelihood estimators and variance-covariance matrix of the estimators are obtained. Murat & Szynal[2] extended the results of Gupta et al.,[1] to the discrete distributions inflated at any point S.

Zero Inflated Truncated Power Series Distribution contains two parameters. The first parameter indicates inflation (π) of zero and the other parameter (θ) is that of power series distribution. Literature survey reveals that many researchers devoted to the inflation parameter of the model. In the present study, we focus on the referential aspect of the basic parameter of the model. In this article, we provide maximum likelihood parameters, Fisher information and asymptotic tests for testing the parameter of the Zero Inflated Truncated Power Series Distribution. Additionally, asymptotic confidence interval for the parameter is provided.

In section 2.1 we report estimation of both the parameters of ZITPSD and corresponding asymptotic variances using full likelihood approach, conditional likelihood approach and method of moments. In section 2.2, we provide three asymptotic tests for testing the parameter of ZITPSD. Section 2.3 is devoted to asymptotic confidence intervals for the parameters of ZITPSD. In section 3.1 we report estimation of parameters involved in Zero Inflated Truncated Poisson Distribution (ZITPD) and inference related to the model. Section 3.2 is devoted to three asymptotic tests for testing the parameter of ZITPD and in section 3.3 we provide asymptotic confidence intervals for the parameters of ZITPD. Simulation study is carried out in section 4, to study the performance of the tests. Illustrative example is provided in section 5.

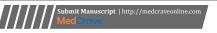## Zero-inflated truncated power series distribution (ZITPSD)

Before we define truncated ZIPSD, we first consider the Truncated Power Series Distribution (TPSD) truncated at the support point 't' onwards, where 't' is known. Then the probability mass function of TPSD is given by

$$P(X=x) = \frac{b_x \theta^x}{f(\theta)(1 - P(X > t))}, \quad for \ x = 0, 1, 2, \dots, t$$

$$= \frac{b_x \theta^x}{\left(\sum\limits_{y=0}^{t} b_y \theta^y \right)}, \qquad = \frac{b_x \theta^x}{G(\theta)}, \quad where \ G(\theta) = \sum\limits_{y=0}^{t} b_y \theta^y$$

It is clear that the truncated distribution is also Power series distribution. Based on the same, we define ZITPSD as follows:

Let the probability mass function of a random variable X is given by

$$P(X=x)=\begin{cases}1-\pi+\pi\dfrac{b_0}{G(\theta)} & for\ x=0\\[2mm]\dfrac{\pi b_x\theta^x}{G(\theta)} & for\ x=1,2,3,...,t\end{cases}\ ....(2.1)$$

where $G(\theta)=\sum\limits_{y=0}^{t} b_y\theta^y$

Estimation of π and θ

Estimation of the parameters using full likelihood function:
Suppose a random sample $X_1, X_2, ..., X_n$ of size n from ZITPSD is available. Then the likelihood function is given by

$$L(\theta,\pi\ ;\underline{x})=\prod_{i=1}^{n}\left(1-\pi+\frac{\pi b_0}{G(\theta)}\right)^{1-a_i}\left(\frac{\pi b x_i \theta^{x_i}}{G(\theta)}\right)^{a_i}\qquad \theta,\ \pi>0$$

where $a_i=0$ if $x_i=0$ and $a_i=1$ if $x_i=1,2,3,....t.$ ...(2.2)

then,

$$\log L(\theta,\pi\ ;\underline{x})=$$

$$=n_0\log\left(1-\pi+\frac{\pi b_0}{G(\theta)}\right)+\sum_{i=1}^{n}a_i\log\pi+\sum_{i=1}^{n}a_i\log bx_i+\sum_{i=1}^{n}a_i x_i\log(\theta)-\sum_{i=1}^{n}a_i\log G(\theta)$$

...(2.3)

Maximum likelihood estimators of θ and π are obtained by solving the following two equations

$$\hat{\pi}=\frac{(n-n_0)G(\hat{\theta})}{n(G(\hat{\theta})-b_0)}\qquad\qquad ...(2.4)$$

$$\frac{\sum\limits_{i=1}^{n}a_i x_i}{\hat{\theta}}=\frac{n_0\hat{\pi}b_0 G'(\hat{\theta})}{G(\hat{\theta})^2\left(1-\hat{\pi}+\hat{\pi}\frac{b_0}{G(\hat{\theta})}\right)}+\frac{\sum\limits_{i=1}^{n}a_i G'(\hat{\theta})}{G(\hat{\theta})},\qquad ...(2.5)$$

Substituting $\hat{\pi}=\frac{(n-n_0)G(\hat{\theta})}{n(G(\hat{\theta})-b_0)}$ in eq. (2.5) we get

$$\overline{x}=\frac{\hat{\theta}\ G'(\hat{\theta})}{(G(\hat{\theta})-b_0)},\qquad\qquad ...(2.6)$$

which is non-linear equation in θ, Using Newton-Raphson method first we find $\hat{\theta}$, substituting this value of $\hat{\theta}$ in Eq. (2.4) we find $\hat{\pi}$. The Fisher information matrix of $\underline{\delta}=(\pi,\theta)'$ is given by

$$I(\underline{\delta})=\begin{pmatrix}-E\left(\dfrac{\partial^2\log L}{\partial\pi^2}\right) & -E\left(\dfrac{\partial^2\log L}{\partial\pi\partial\theta}\right)\\[4mm]-E\left(\dfrac{\partial^2\log L}{\partial\pi\partial\theta}\right) & -E\left(\dfrac{\partial^2\log L}{\partial\theta^2}\right)\end{pmatrix}=\begin{pmatrix}I_{11} & I_{12}\\ I_{21} & I_{22}\end{pmatrix}$$

Where

$$I_{11}=\left(\frac{n(G(\theta)-b_0)}{\pi(G(\theta)-\pi\ G(\theta)+\pi\ b_0)}\right),\qquad ...(2.7)$$

$$I_{12}=\left(\frac{n\ b_0 G'(\theta)}{G(\theta)\big(G(\theta)-\pi G(\theta)+\pi b_0\big)}\right)\qquad ...(2.8)$$

and

$$I_{22}=\left(n\pi\ b_0\left(\left(\frac{G(\theta)^2 G''(\theta)-2G'(\theta)^2 G(\theta)}{G(\theta)^4}\right)+\frac{\pi\ b_0 G'(\theta)^2}{G(\theta)^4\left(1-\pi+\pi\frac{b_0}{G(\theta)}\right)}\right)\right.$$

$$\left.+n\pi\left(\frac{G'(\theta)}{\theta G(\theta)}\right)+\frac{n\pi\left(1-\frac{b_0}{G(\theta)}\right)\big(G(\theta)G''(\theta)-G'(\theta)^2\big)}{G(\theta)^2}\right)\ ...(2.9)$$

Assuming that conditions required for asymptotic normality for maximum likelihood estimators are satisfied, we have following theorem:

Theorem 2.1: Let $X_1, X_2,...X_n$ be a random sample from ZITPSD with parameters π and θ. Then the maximum likelihood estimator obtained by solving eq. (2.4) and eq. (2.6), have asymptotic bivariate normal distribution with mean vector $(\pi,\theta)'$ and dispersion matrix $I^{-1}(\underline{\delta})$ for n sufficiently large.

That is as $n\to\infty$, $\left(\sqrt{n}(\hat{\pi}-\pi),\sqrt{n}(\hat{\theta}-\theta)\right)\to N_2(0,I^{-1}(\underline{\delta}))$.

In the following we present conditional likelihood approach and obtain MLEs for θ.

Conditional likelihood function approach: We observe that the conditional density of X_i given $A_i=a_i$ is independent of inflation parameter π, since

$$P(X_i=x_i|A_i=a_i)=\left(\frac{b_{x_i}(\theta)^{x_i}}{G(\theta)-b_0}\right)^{a_i}\ ...(2.10)$$

Now the conditional log likelihood function is given by

$$\log L^*(\theta;\underline{x})=\sum_{i=1}^{n-n_0}\log b_{x_i}+\sum_{i=1}^{n-n_0}x_i\log(\theta)-\sum_{i=1}^{n-n_0}\log\big(G(\theta)-b_0\big)\ ...(2.11)$$

The mle $\tilde{\theta}$ of θ is the solution to an equation

$$\overline{x}=\frac{\tilde{\theta}G'(\tilde{\theta})}{(G(\tilde{\theta})-b_0)},\qquad\qquad ...(2.12)$$

where $\overline{x}=\dfrac{\sum\limits_{i=1}^{n-n_o}x_i}{n-n_0}$ is the mean of the positive observations only. We note that mle of θ based on full likelihood (eq. 2.6) and based on conditional likelihood (Eq. 2.12) are the same and

$$AV_{\tilde{\theta}}(\theta)=$$

$$\left((n-n_0)\left\{\left(\frac{G'(\theta)}{\theta(G(\theta)-b_0)}+\frac{G''(\theta)(G(\theta)-b_0)-G'(\theta)^2}{(G(\theta)-b_0)^2}\right)\right\}\right)^{-1}\ ...(2.13)$$

Assuming that Cramer-Huzurbazar conditions required for asymptotic normality for MLEs are satisfied, we have following theorem:

Theorem 2.2: Let $X_1, X_2, \dots X_n$ be a random sample from ZITPSD with parameters π and θ. Then the mle of θ is solution to the eq. (2.12) and has asymptotic normal distribution with mean θ and variance $AV_{\tilde{\theta}}(\theta)$ for n sufficiently large. That is as $n \to \infty$, $\left(\sqrt{n}(\tilde{\theta} - \theta)\right) \to N_1(0, AV_{\tilde{\theta}}(\theta))$.

In the following we present moment estimator of ZITPSD.

Moment estimator of ZITPSD: We have,

$$E(X) = \frac{\pi \theta G'(\theta)}{G(\theta)} = \pi \mu(\theta)$$

$$E(X^2) = \frac{\theta \pi}{G(\theta)}\left(\theta G''(\theta) + G'(\theta)\right)$$

and

$$Var(X) = \frac{\theta \pi}{G(\theta)}\left(\theta G''(\theta) + G'(\theta) - \frac{\pi \theta (G'(\theta))^2}{G(\theta)}\right),$$

$$= \sigma^2(\pi, \theta) \text{ say.}$$

Let,

$$\overline{X} = \pi \mu(\theta) \qquad \dots(2.14)$$

$$\frac{\sum_{i=1}^{n} x_i^2}{n} = \frac{\pi \theta}{G(\theta)}\left(\theta G''(\theta) + G'(\theta)\right), \qquad \dots(2.15)$$

Solving eq. (2.14) and eq. (2.15) we get moment estimators of π and θ.

Theorem 2.3: Let $X_1, X_2, \dots X_n$ be a random sample from ZITPSD with parameters π and θ. Then the moment estimator of π and θ are obtained by solving in the eq. (2.14) and eq. (2.15). The moment estimator of θ has asymptotic normal distribution with mean $\pi \mu(\theta)$ and variance $\dfrac{\sigma^2(\pi, \theta)}{n}$, for n sufficiently large. That is as $n \to \infty$, $\left(\sqrt{n}(\overline{\theta} - \theta)\right) \to N_1\left(0, \dfrac{\sigma^2(\pi, \theta)}{n}\right)$.

Tests for the parameter θ of ZITPS distribution

Test based on $\hat{\theta}$: Suppose we wish to test $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$. Let us assume that π is known. Therefore, under H_0, from Theorem (2.1) we have

$$\left(\hat{\theta} - \theta_0\right) \sim AN\left(0, AV_{\hat{\theta}}(\pi, \theta_0)\right). \qquad \dots(2.16)$$

Define a test statistic to be $Z_1 = \dfrac{\hat{\theta} - \theta_0}{\sqrt{AV_{\hat{\theta}}(\pi, \theta_0)}}$. Based on Z_1 we define the test ψ_1 which rejects H_0 at α level of significance, if $|Z_1| > z_{1-\alpha/2}$, where $z_{1-\alpha/2}$ is the upper $100(\alpha/2)^{\text{th}}$ percentile of SNV.

Let $\Phi(.)$ be the cumulative distribution function of SNV. Then the power of the test ψ_1 is given by

$$\beta_{\psi_1}(\pi, \theta) = 1 - \Phi(B) + \Phi(A),$$

where

$$A = \frac{\theta_0 - \theta - z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\pi, \theta_0)}}{\sqrt{AV_{\hat{\theta}}(\pi, \theta)}} \text{ and}$$

$$B = \frac{\theta_0 - \theta + z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\pi, \theta_0)}}{\sqrt{AV_{\hat{\theta}}(\pi, \theta)}}.$$

However, in practice π is unknown. Hence we modify the test statistic by replacing π by its maximum likelihood estimator ($\hat{\pi}_0$), when H_0 is true. By doing so, we define test $Z_1' = \dfrac{\hat{\theta} - \theta_0}{\sqrt{AV_{\hat{\theta}}(\hat{\pi}_0, \theta_0)}}$, where $\hat{\pi}_0 = \dfrac{(n - n_0)f(\theta_0)}{n\left(f(\theta_0) - b_0\right)}$

Based on Z_1', we propose a test ψ_1' rejects H_0 at α level of significance, if $|Z_1'| > Z_{1-\alpha/2}$.

The power of this test is given by

$$\beta_{\psi_1'}(\pi, \theta) = \sum_{k=0}^{n}\left(1 - \Phi(\hat{B}_k) + \Phi(\hat{A}_k)\right)P(n_0 = k), \qquad \dots(2.17)$$

where

$$\hat{B}_k = \frac{\theta_0 - \theta + z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\hat{\pi}_0, \theta_0)}}{\sqrt{AV_{\hat{\theta}}(\pi, \theta)}}, \quad \hat{A}_k = \frac{\theta_0 - \theta - z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\hat{\pi}_0, \theta_0)}}{\sqrt{AV_{\hat{\theta}}(\pi, \theta)}},$$

$$P(n_0 = k) = \binom{n}{k}P_0^k(1 - P_0)^{n-k}, \text{ with } P_0 = \left(1 - \pi + \pi\frac{b_0}{f(\theta)}\right). \qquad \dots(2.18)$$

Below we develop test based on $\tilde{\theta}$, estimator based on conditional likelihood approach.

Test based on $\tilde{\theta}$: Theorem (2.5) gives

$$\left(\tilde{\theta} - \theta_0\right) \sim AN\left(0, AV_{\tilde{\theta}}(\theta_0)\right). \qquad \dots(2.19)$$

Hence, we define test statistic $Z_2 = \dfrac{\tilde{\theta} - \theta_0}{\sqrt{AV_{\tilde{\theta}}(\theta_0)}}$. A test based on Z_2 which rejects H_0 α level of significance, if $|Z_2| > z_{1-\alpha/2}$.

The power of the test ψ_2 is given by

$$\beta_{\psi_2}(\pi, \theta) = 1 - \Phi(B) + \Phi(A), \qquad \dots(2.20)$$

where, $A = \dfrac{\theta_0 - \theta - z_{1-\alpha/2}\sqrt{AV_{\tilde{\theta}}(\theta_0)}}{\sqrt{AV_{\hat{\theta}}(\theta)}}$, $B = \dfrac{\theta_0 - \theta + z_{1-\alpha/2}\sqrt{AV_{\tilde{\theta}}(\theta_0)}}{\sqrt{AV_{\tilde{\theta}}(\theta)}}$.

Test based on the moment estimator $\overline{\theta}$ of θ: It is clear that the problem of testing $H_0: \theta = \theta_0$ vs $H_1: \theta \neq \theta_0$ is equivalent to testing $H_0: \mu(\theta) = \mu(\theta_0)$ vs $H_1: \mu(\theta) \neq \mu(\theta_0)$, where $\mu(\theta) = \dfrac{\theta G'(\theta)}{G(\theta)}$. We have from Theorem (2.3), sample mean is consistent and asymptotically normal for the population mean.

That is $\overline{X} \sim AN\left(\pi \mu(\theta), \dfrac{\sigma^2(\pi, \theta)}{n}\right)$.

Therefore, under H_0, we have

$$\sqrt{n}\left(\frac{\overline{X}}{\pi} - \mu(\theta_0)\right) \sim AN\left(0, \frac{\sigma^2(\pi, \theta_0)}{\pi^2}\right).$$

Define test statistic

$$Z_3 = \frac{\sqrt{n}\left(\dfrac{\overline{X}}{\pi} - \mu(\theta_0)\right)}{\sqrt{\dfrac{\sigma^2(\pi,\theta_0)}{\pi^2}}} \sim N(0,1) \text{, when } \pi \text{ is known.}$$

The test ψ_3 rejects H_0 at α level of significance if $|Z_3| > z_{1-\alpha/2}$.

That is, reject H_0 if $\left(\dfrac{\sqrt{n}\left|\dfrac{\overline{X}}{\pi} - \mu(\theta_0)\right|}{\sqrt{\dfrac{\sigma^2(\pi,\theta_0)}{\pi^2}}}\right) > z_{1-\alpha/2}$.

The power of the test ψ_3 is given by

$$\beta_{\psi_3}(\pi,\theta) = 1 - \Phi(B') + \Phi(A') \text{ , } \quad \dots(2.21)$$

where $A' = \dfrac{\pi\left(\mu(\theta_0) - z_{1-\alpha/2}\sqrt{\dfrac{\sigma^2(\pi,\theta_0)}{n\pi^2}}\right) - \pi\mu(\theta)}{\sqrt{\dfrac{\sigma^2(\pi,\theta)}{n}}}$ and

$$B' = \frac{\pi\left(\mu(\theta_0) + z_{1-\alpha/2}\sqrt{\dfrac{\sigma^2(\pi,\theta_0)}{n\pi^2}}\right) - \pi\mu(\theta)}{\sqrt{\dfrac{\sigma^2(\pi,\theta)}{n}}},$$

If π is unknown, we modify the test statistic by replacing π by its estimate $(\hat{\pi}_0)$ under H_0. By doing so, we define test statistic

$$Z_3' = \frac{\sqrt{n}\left(\dfrac{\overline{X}}{\hat{\pi}_0} - \mu(\theta_0)\right)}{\sqrt{\dfrac{\sigma^2(\hat{\pi}_0,\theta_0)}{\hat{\pi}_0^2}}}, \quad \dots(2.22)$$

where $\hat{\pi}_0$ is given by $\hat{\pi}_0 = \dfrac{\overline{X}}{\mu(\theta_0)}$.

Based on Z_3' we propose a test ψ_3' which rejects H_0 at α level of significance if $|Z_3'| > z_{1-\alpha/2}$.

The power of the test is given by

$$\beta_{\psi_3'}(\pi,\theta) = \sum_{k=0}^{n}\left(1 - \Phi(B_k'') + \Phi(A_k'')\right)P(n_0=k) \text{ , } \dots(2.23)$$

where

$$A_k'' = \frac{\hat{\pi}_0\left(\mu(\theta_0) - z_{1-\alpha/2}\sqrt{\dfrac{\sigma^2(\hat{\pi}_0,\theta_0)}{n\pi^2}}\right) - \pi\mu(\theta)}{\sqrt{\dfrac{\sigma^2(\pi,\theta)}{n}}}$$

$$B_k'' = \frac{\hat{\pi}_0\left(\mu(\theta_0) + z_{1-\alpha/2}\sqrt{\dfrac{\sigma^2(\hat{\pi}_0,\theta_0)}{n\pi^2}}\right) - \pi\mu(\theta)}{\sqrt{\dfrac{\sigma^2(\pi,\theta)}{n}}}$$

and $P(n_0=k) = \binom{n}{k}P_0^k(1-P_0)^{n-k}$, with $P_0 = \left(1 - \pi + \pi\dfrac{a_0}{f(\theta)}\right)$.

Using the tests developed above, we can define two sided asymptotic confidence intervals for θ, by inverting acceptance regions of the tests appropriately. Below we report the same.

Asymptotic confidence interval for the parameter θ

Asymptotic confidence interval for θ based on the test ψ_1 is given by

$$\left(\hat{\theta} - z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\hat{\pi},\hat{\theta})}, \quad \hat{\theta} + z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\hat{\pi},\hat{\theta})}\right) \quad \dots(2.24)$$

where, $AV_{\hat{\theta}}(\hat{\pi},\hat{\theta})$ is an estimate of asymptotic variance of $\hat{\theta}$ and asymptotic confidence interval for θ based on the test ψ_2 is given by

$$\left(\tilde{\theta} - z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\tilde{\theta})}, \quad \tilde{\theta} + z_{1-\alpha/2}\sqrt{AV_{\hat{\theta}}(\tilde{\theta})}\right) \quad \dots(2.25)$$

where $AV_{\hat{\theta}}(\tilde{\theta})$ is an estimate of the asymptotic variance of $\tilde{\theta}$ as given in the eq. (2.13).

Asymptotic confidence interval for θ based on the test ψ_3 is given by

$$\left(\frac{\overline{X}}{\hat{\pi}} - z_{1-\alpha/2}\sqrt{AV_{\overline{\theta}}(\hat{\pi},\overline{\theta})}, \quad \frac{\overline{X}}{\hat{\pi}} + z_{1-\alpha/2}\sqrt{AV_{\overline{\theta}}(\hat{\pi},\overline{\theta})}\right), \quad \dots(2.26)$$

where $AV_{\overline{\theta}}(\hat{\pi},\overline{\theta}) = \dfrac{\sqrt{n}\left(\dfrac{\overline{X}}{\pi} - \mu(\theta)\right)}{\sqrt{\dfrac{\sigma^2(\pi,\theta)}{\pi^2}}}$.

In the following we study inference for zero-inflated Truncated Poisson distribution using results reported in the earlier.

Zero-inflated truncated poisson distribution

Truncated samples from discrete distributions arise in numerous situations where counts of zero are not observed. As an example, consider the distribution of the number of children per family in developing nations, where records are maintained only if there is at least a child in the family. The number of childless families remains unknown. The resulting sample is thus truncated with zero class missing. In continuous distribution, a sample of this type would be described as singly left truncated. In other situations, sample from discrete distributions might be censored on the right.

In this section, we consider zero-inflated truncated Poisson distribution truncated at right at the support point 't' onwards, where 't' is known. Moments, maximum likelihood estimators, Fisher information matrix for full and conditional likelihood are provided. We provide three tests for testing the parameter of the ZITPD.

Consider the probability mass function of truncated Poisson distribution (TPD) truncated at the support point 't' onwards. The probability mass function of TPD is given by

$$P(X=x) = \frac{e^{-\theta}\theta^x}{x!(1-P(X>t))}, \quad for\ x=0,1,2,\dots,t$$

$$= \frac{e^{-\theta}\theta^x}{x!\left(\sum_{y=0}^{t}\dfrac{e^{-\theta}\theta^y}{y!}\right)},$$

Citation: Patil MK. Inference for zero inflated truncated power series family of distributions. *Biom Biostat Int J.* 2016;4(7):119–122.
DOI: 10.15406/bbij.2016.04.00115

$$=\frac{\theta^x}{x!\,A(\theta)}, \quad \text{where} \quad A(\theta)=\left(\sum_{y=0}^{t}\frac{\theta^y}{y!}\right)$$

Using this truncated distribution, we define the zero-inflated truncated Poisson distribution truncated at $'t'$ onwards.

The probability mass function of ZITP distribution is given by

$$P(X=x)= \begin{cases} (1-\pi)+\dfrac{\pi}{A(\theta)} & \text{for } x=0, \\[2em] \dfrac{\pi\,\theta^x}{x!\,A(\theta)} & \text{for } x=1,2,\dots,t \text{ and } \theta>0, 0<\pi<1 \end{cases} \quad \dots(3.1)$$

Estimation of the parameters π and θ

Estimation of the parameters using full likelihood function

Let $X_1, X_2, \dots, X_n$ be a random sample observed from zero-inflated truncated Poisson distribution truncated at $'t'$ onwards, where $'t'$ is the point in the support defined in the above probability mass function. Then the likelihood function is given by

$$L(\theta,\pi;\underline{x})=\prod_{i=1}^{n}\left(1-\pi+\frac{\pi}{A(\theta)}\right)^{1-a_i}\left(\frac{\pi\,\theta^x}{x!\,A(\theta)}\right)^{a_i} \qquad \theta,\pi>0$$

The corresponding log likelihood function is given by

$$\log L(\theta,\pi;\underline{x}) = n_0\log\left(1-\pi+\frac{\pi}{A(\theta)}\right)+\sum_{i=1}^{n}a_i\log\pi$$

$$+\sum_{i=1}^{n}a_i x_i\log\theta-\sum_{i=1}^{n}a_i\log x_i!-\sum_{i=1}^{n}a_i\log A(\theta) \qquad \dots(3.2)$$

To find MLEs of θ and π, we differentiate the eq. (3.2) with respective π and θ, and then equating to zero we get

$$\hat{\pi}=\frac{(n-n_0)A(\hat{\theta})}{n\left(A(\hat{\theta})-1\right)} \qquad \dots(3.3)$$

and $$\frac{\sum_{i=1}^{n}a_i x_i}{\hat{\theta}}=\frac{\pi\,n_0\left(A'(\hat{\theta})\right)}{\left(1-\pi+\frac{\pi}{A(\hat{\theta})}\right)\left(A(\hat{\theta})\right)^2}+\frac{\sum_{i=1}^{n}a_i A'(\hat{\theta})}{A(\hat{\theta})}$$

$$=\frac{A'(\hat{\theta})}{A(\hat{\theta})}\left((n-n_0)+\frac{\hat{\pi}\,n_0}{\left(1-\pi+\frac{\hat{\pi}}{A(\hat{\theta})}\right)A(\hat{\theta})}\right)$$

$$\frac{\sum_{i=1}^{n}a_i x_i}{\theta}=\frac{A'(\theta)}{A(\theta)}\left(\frac{(n-n_0)(1-\pi)A(\theta)+n\pi}{(1-\pi)A(\theta)+\pi}\right) \qquad \dots(3.4)$$

Substituting $\hat{\pi}=\dfrac{(n-n_0)A(\hat{\theta})}{n\left(A(\hat{\theta})-1\right)}$ in the above equation we have

$$\frac{\sum_{i=1}^{n}a_i x_i}{\hat{\theta}}=\frac{(n-n_0)A'(\hat{\theta})}{\left(A(\hat{\theta})-1\right)},$$

$$\sum_{i=1}^{n}a_i x_i\left(A(\hat{\theta})-1\right)-(n-n_0)\hat{\theta}\,A'(\hat{\theta})=0, \qquad \dots(3.5)$$

which is non-linear equation in $\hat{\theta}$. Therefore, we use a numerical technique to solve it. Let

$$h(\hat{\theta})=\sum_{i=1}^{n}a_i x_i\left(A(\hat{\theta})-1\right)-(n-n_0)\hat{\theta}\,A'(\hat{\theta}) \quad \text{and}$$

$$h'(\hat{\theta})=\sum_{i=1}^{n}a_i x_i A'(\hat{\theta})-(n-n_0)\left(\hat{\theta}\,A''(\hat{\theta})+A'(\hat{\theta})\right).$$

Using Newton-Raphson iterative formula $\hat{\theta}_{i+1}=\hat{\theta}_i-\dfrac{h(\hat{\theta})}{h'(\hat{\theta})}, \quad i=0,1,2,\dots$ with suitable initial value of θ_0 we get $\hat{\theta}$. Substituting this value of $\hat{\theta}$ in eq. (3.3), we get the value of $\hat{\pi}$.

In the following we find the elements of Fisher information matrix

Here we have

$$\frac{\partial\log L}{\partial\pi}=\frac{n_0\left(-1+\frac{1}{A(\theta)}\right)}{\left(1-\pi+\frac{\pi}{A(\theta)}\right)}+\frac{\sum_{i=1}^{n}a_i}{\pi},$$

$$\frac{\partial^2\log L}{\partial\pi^2}=-\frac{n_0\left(-1+\frac{1}{A(\theta)}\right)^2}{\left(1-\pi+\frac{\pi}{A(\theta)}\right)^2}-\frac{\sum_{i=1}^{n}a_i}{\pi^2},$$

$$-E\left(\frac{\partial^2\log L}{\partial\pi^2}\right)=\frac{E(n_0)\left(-1+\frac{1}{A(\theta)}\right)^2}{\left(1-\pi+\frac{\pi}{A(\theta)}\right)^2}+\frac{E\left(\sum_{i=1}^{n}a_i\right)}{\pi^2},$$

$$=\frac{n\left(-1+\frac{1}{A(\theta)}\right)^2}{\left(1-\pi+\frac{\pi}{A(\theta)}\right)}+\frac{n\left(1-\frac{1}{A(\theta)}\right)}{\pi},$$

$$-E\left(\frac{\partial^2\log L}{\partial\pi^2}\right)=\frac{n\left(-1+\frac{1}{A(\theta)}\right)^2}{\left(1-\pi+\frac{\pi}{A(\theta)}\right)}-\frac{n\left(-1+\frac{1}{A(\theta)}\right)}{\pi},$$

$$I_{11}=\frac{n\left(A(\theta)-1\right)}{\pi\left(A(\theta)-\pi A(\theta)+\pi\right)}. \qquad \dots(3.6)$$

Now

$$\frac{\partial\log L}{\partial\pi}=\frac{n_0\left(-1+\frac{1}{A(\theta)}\right)}{\left(1-\pi+\frac{\pi}{A(\theta)}\right)}+\frac{\sum_{i=1}^{n}a_i}{\pi},$$

$$\frac{\partial^2 \log L}{\partial \pi \partial \theta} = \frac{-n_0 \dfrac{A'(\theta)}{(A(\theta))^2}\left(\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)-\pi\left(-1+\dfrac{1}{A(\theta)}\right)\right)}{\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)^2},$$

$$-E\left(\frac{\partial^2 \log L}{\partial \pi \partial \theta}\right) = \frac{E(n_0)\left(\dfrac{A'(\theta)}{A(\theta)^2}\right)\left(\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)-\pi\left(-1+\dfrac{1}{A(\theta)}\right)\right)}{\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)^2},$$

$$\frac{\sum_{i=1}^n a_i x_i}{\theta^2} - \frac{\sum_{i=1}^n a_i\left(A(\theta)A''(\theta)-A'(\theta)^2\right)}{A(\theta)^2}$$

$$-E\left(\frac{\partial^2 \log L}{\partial \pi \partial \theta}\right) = I_{12} = \frac{nA'(\theta)}{A(\theta)\left(A(\theta)-\pi A(\theta)+\pi\right)} \qquad \dots(3.7)$$

Further differentiating eq. (3.2) twice with respect to θ, we get

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{n_0 \pi\left\{\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)\left(\dfrac{A(\theta)^2 A''(\theta)-2A'(\theta)^2 A(\theta)}{f(\theta)^4}\right)-\dfrac{A'(\theta)^2}{A(\theta)^2}\left(-\dfrac{\pi}{A(\theta)^2}\right)\right\}}{\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)^2}$$

Therefore,

$$-E\left(\frac{\partial^2 \log L}{\partial \theta^2}\right) =$$

$$\left(n\pi\left(\left(\dfrac{A(\theta)^2 A''(\theta)-2A'(\theta)^2 A(\theta)}{A(\theta)^4}\right)+\dfrac{\pi A'(\theta)^2}{A(\theta)^4\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)}\right)\right.$$

$$\left.+\dfrac{n\pi A'(\theta)}{\theta A(\theta)}+\dfrac{n\pi\left(1-\dfrac{1}{A(\theta)}\right)\left(A(\theta)A''(\theta)-A'(\theta)^2\right)}{A(\theta)^2}\right).$$

Hence,

$$I_{22} = \left(n\pi\left(\left(\dfrac{A(\theta)^2 A''(\theta)-2A'(\theta)^2 A(\theta)}{A(\theta)^4}\right)+\dfrac{\pi A'(\theta)^2}{A(\theta)^4\left(1-\pi+\dfrac{\pi}{A(\theta)}\right)}\right)\right.$$

$$\left.+\dfrac{n\pi A'(\theta)}{\theta A(\theta)}+\dfrac{n\pi\left(1-\dfrac{1}{A(\theta)}\right)\left(A(\theta)A''(\theta)-A'(\theta)^2\right)}{A(\theta)^2}\right).$$

The asymptotic variance of $\hat{\pi}$ and $\hat{\theta}$ are

$$AV_{\hat{\pi}}(\pi,\theta)=I^{11}=\left(I_{11}-\frac{I_{12}}{I_{22}}\right)^{-1}.$$

$$AV_{\hat{\theta}}(\pi,\theta)=I^{22}=\left(I_{22}-\frac{I_{12}}{I_{11}}\right)^{-1}. \qquad \dots (3.8)$$

I.   Conditional likelihood function approach

The conditional likelihood function is given by

$$L^*(\theta;\underline{x})=\prod_{i=1}^{n}\left(\frac{\theta^{x_i}}{x_i!(A(\theta)-1)}\right)^{a_i}, \qquad \theta>0 \qquad \dots(3.9)$$

The corresponding log likelihood function is given by

$$\log L^*(\theta;\underline{x}) = \sum_{i=1}^{n-n_0} x_i \log(\theta) - (n-n_0)\log(A(\theta)-1) - \sum_{i=1}^{n-n_0}\log x_i!$$
$$\dots(3.10)$$

The corresponding mle $\tilde{\theta}$ is the solution to an equation

$$\bar{x}=\frac{\tilde{\theta}A(\tilde{\theta})}{A(\tilde{\theta})-1} \qquad \dots(3.11)$$

Now consider,

$$\frac{\partial^2 \log L^*}{\partial \theta^2}=-\sum_{i=1}^{n-n_0}\frac{x_i}{\theta^2}-\sum_{i=1}^{n-n_0}\frac{\left(A(\theta)-1\right)A''(\theta)-A'(\theta)^2}{\left(A(\theta)-1\right)^2}$$

$$-E\left(\frac{\partial^2 \log L^*}{\partial \theta^2}\right) = E\left(\sum_{i=1}^{n-n_0}\frac{x_i}{\theta^2}\right)+E\left(\sum_{i=1}^{n-n_0}\frac{\left(A(\theta)-1\right)A''(\theta)-A'(\theta)^2}{\left(A(\theta)-1\right)^2}\right)$$

$$=\frac{(n-n_0)\theta A'(\theta)}{\theta^2\left(A(\theta)-1\right)}+\frac{(n-n_0)\left\{\left(A(\theta)-1\right)A''(\theta)-A'(\theta)^2\right\}}{\left(A(\theta)-1\right)^2}$$

$$=\frac{(n-n_0)}{\left(A(\theta)-1\right)}\left(\frac{A'(\theta)}{\theta}+\frac{A''(\theta)(A(\theta)-1)-A'(\theta)^2}{\left(A(\theta)-1\right)}\right). \qquad \dots (3.12)$$

Therefore, asymptotic variance of $\hat{\theta}$ is different than the asymptotic variance of estimate of θ based on the standard likelihood approach. The same is given by

$$AV_{\tilde{\theta}}(\theta)=\left(\frac{(n-n_0)}{\left(A(\theta)-1\right)}\left(\frac{A'(\theta)}{\theta}+\frac{(A(\theta)-1))A''(\theta)-(A'(\theta)^2}{\left(A(\theta)-1\right)}\right)\right)^{-1} \qquad (3.13)$$

II.   Moment estimator of ZITP distribution

$$\text{Mean} = E(X) = \frac{\pi\theta A'(\theta)}{A(\theta)} \qquad \dots(3.14)$$

$$E(X^2) = \frac{\theta\pi}{A(\theta)}\left(\theta A''(\theta)+A'(\theta)\right)$$

$$Var(X) = \frac{\theta \pi}{A(\theta)} \left(\theta A''(\theta) + A'(\theta) - \frac{\pi \theta A'(\theta)^2}{A(\theta)} \right) = \sigma^2(\pi, \theta) \text{ say } \dots(3.15)$$

$$\overline{x} = \frac{\pi \theta A'(\theta)}{A(\theta)} \qquad \dots(3.16)$$

$$\frac{\sum_{i=1}^{n} x_i^2}{n} = \frac{\theta \pi}{A(\theta)} \left(\theta A''(\theta) + A'(\theta) \right) \qquad \dots(3.17)$$

Solving eq. (3.16) and eq. (3.17), we get moment estimators of π and θ.

Tests for the parameter θ of ZITP distribution

Suppose we want to test $H_0 : \theta = \theta_0$ vs $H_1 : \theta \neq \theta_0$, (assuming π is unknown) [4]

Test based on $\hat{\theta}$

$$Z_4 = \frac{\hat{\theta} - \theta_0}{\sqrt{AV_{\hat{\theta}}(\hat{\pi}_0, \theta_0)}} \qquad \dots(3.18)$$

where $AV_{\hat{\theta}}(\hat{\pi}_0, \theta_0)$ is defined in eq. (3.8). The test ψ_4 rejects H_0, if $|Z_4| > z_{1-\alpha/2}$.

Test based on $\tilde{\theta}$

The test statistic here is $Z_5 = \dfrac{\tilde{\theta} - \theta_0}{\sqrt{AV_{\tilde{\theta}}(\theta_0)}}$, $\dots(3.19)$

Where, $AV_{\tilde{\theta}}(\theta_0)$ is as defined in eq. (3.13). The test ψ_5 rejects H_0 if $|Z_5| > z_{1-\alpha/2}$.

I.   Test based on sample mean

The test statistic

$$Z_6 = \frac{\sqrt{n} \left(\frac{\overline{X}}{\hat{\pi}_0} - \theta_0 \right)}{\sqrt{\hat{\pi}_0^{-2} AV_{\overline{X}}(\hat{\pi}_0, \theta_0)}} , \quad \dots(3.20)$$

where $\hat{\pi}_0 = \overline{X} \left(\dfrac{A(\theta_0)}{\theta_0 A'(\theta_0)} \right)$

Power of the test is given by

$$\beta_{\psi_6}(\hat{\pi}, \theta) = \sum_{k=0}^{n} \left(1 - \Phi(\hat{B}_k) + \Phi(\hat{A}_k) \right) P(n_0 = k)$$

where, $\hat{B}_k = \dfrac{\hat{\pi}_0 \left(\theta_0 + z_{1-\alpha/2} \sqrt{\hat{\pi}_0^{-2} AV_{\overline{X}}(\hat{\pi}_0, \theta_0)} \right) - \pi \theta}{\sqrt{AV_{\overline{X}}(\pi, \theta_0)}}$,

$$\hat{A}_k = \frac{\hat{\pi}_0 \left(\theta_0 - z_{1-\alpha/2} \sqrt{\hat{\pi}_0^{-2} AV_{\overline{X}}(\hat{\pi}, \theta_0)} \right) - \pi \theta}{\sqrt{AV_{\overline{X}}(\pi, \theta_0)}} \text{ and}$$

$$P(n_0 = k) = \binom{n}{k} P_0^k (1 - P_0)^{n-k} \quad \text{, with } P_0 = 1 - \pi + \frac{\pi}{A(\theta)}$$

Asymptotic confidence interval for the parameter θ

Asymptotic confidence interval for θ based on the test ψ_4 is given by

$$\left(\hat{\theta} - z_{1-\alpha/2} \sqrt{AV_{\hat{\theta}}(\hat{\pi}, \hat{\theta})}, \quad \hat{\theta} + z_{1-\alpha/2} \sqrt{AV_{\hat{\theta}}(\hat{\pi}, \hat{\theta})} \right) \dots(3.21)$$

where, $AV_{\hat{\theta}}(\hat{\pi}, \hat{\theta})$ is an estimate of asymptotic variance of $\hat{\theta}$ and asymptotic confidence interval for θ based on the test ψ_5 is given by

$$\left(\tilde{\theta} - z_{1-\alpha/2} \sqrt{AV_{\tilde{\theta}}(\tilde{\theta})}, \quad \tilde{\theta} + z_{1-\alpha/2} \sqrt{AV_{\tilde{\theta}}(\tilde{\theta})} \right) \dots(3.22)$$

where $AV_{\tilde{\theta}}(\tilde{\theta})$ is an estimate of the asymptotic variance of $\tilde{\theta}$ as given in the eq. (3.13).

Asymptotic confidence interval for θ based on the test ψ_6 is given by

$$\left(\frac{\overline{X}}{\hat{\pi}} - z_{1-\alpha/2} \sqrt{AV_{\overline{\theta}}(\hat{\pi}, \overline{\theta})}, \quad \frac{\overline{X}}{\hat{\pi}} + z_{1-\alpha/2} \sqrt{AV_{\overline{\theta}}(\hat{\pi}, \overline{\theta})} \right), \dots(3.23)$$

$$\text{where } AV_{\overline{\theta}}(\hat{\pi}, \overline{\theta}) = \frac{\sqrt{n} \left(\frac{\overline{X}}{\pi} - \mu(\theta) \right)}{\sqrt{\frac{\sigma^2(\pi, \theta)}{\pi^2}}} .$$

Simulation study

A simulation study is carried out to investigate the power of the two tests proposed in section 3.2. We generate 10000 samples of sizes 100 and 200 for different values of π, θ and truncation point t. Based on generated sample, the test statistics were calculated. Percentage of times the test statistics exceeds $Z_{1-\alpha/2}$ is computed, which is an estimate of power of the respective test. R programme is developed to find power of the test. The results for the case of $\theta_0 = 2$ and 4, $\pi = 0.3$, 0.4, 0.5, 0.6, 0.7, $\alpha = 0.05$ and truncation point t= 7 and 9 are presented in the Table 1 & Table 2.

From the simulation study reported in Table 1 & Table 2, we observe that

I. The test based on full likelihood approach is better than the one based on conditional likelihood approach when θ is small. For large θ, both the tests are equally good.

II. Probability of Type I error of the former test is more than that of later.

III. Since for large values of θ both the tests are equally good. We recommend the use of conditional likelihood approach, when θ is large, from the computational point of view.

IV. If θ is large, proportion of zeros corresponding the Poisson distribution are relatively low. Hence these zeros can be ignored while making inference about θ. However, for smaller values of θ, such ignorance will have effect on inference of θ.

Table 1 Power (in %) of the test ψ_4 and ψ_5 for θ_0 =2. t=7, n=100 and 200, α=0.05

π	θ	$n = 100$		$n = 200$	
		ψ_4	ψ_5	ψ_5	ψ_5
	2.0	6.57	4.28	6.57	4.63
	2.2	11.49	8.9	16.08	12.88
	2.4	26.85	22.24	45.08	39.93
	2.6	49.43	44.5	76.81	72.82
	2.8	71.27	66.99	93.49	91.72
	3	86.28	83.23	98.91	98.46
0.3	3.2	94.64	93.18	99.77	99.73
	3.4	98.08	97.58	99.99	99.99
	3.6	99.44	99.08	100	100
	3.8	99.8	99.76	100	100
	4	99.95	99.94	100	100
	4.2	99.98	99.97	100	100
	4.4	100	100	100	100
	2	6.44	4.24	6.29	4.46
	2.2	12.83	10.33	20.01	15.59
	2.4	33.16	28.94	56.94	50.24
	2.6	60.11	55.4	87.34	83.64
	2.8	81.87	78.72	97.82	97
0.4	3	93.8	92.31	99.83	99.79
	3.2	98.35	97.87	100	100
	3.4	99.62	99.54	100	100
	3.6	99.98	99.96	100	100
	3.8	99.99	99.97	100	100
	3.8	100	100	100	100

Table Continued

π	θ	$n = 100$		$n = 200$	
		ψ_4	ψ_5	ψ_5	ψ_5
	2	6.17	4.46	6.25	4.2
	2.2	14.83	12.01	24.63	19.24
	2.4	40.31	34.76	66.99	60.39
	2.6	70.38	65.06	92.88	90.37
	2.8	90.13	87.37	99.52	99.14
0.5	3	97.23	96.45	99.99	99.97
	3.2	99.55	99.36	100	100
	3.4	99.96	99.94	100	100
	3.6	100	99.99	100	100
	3.8	100	100	100	100
	2	6.8	4.43	7.12	4.89
	2.2	18.04	13.52	28.35	21.91
	2.4	47.01	40.71	73.41	65.85
	2.6	77.65	72.17	96.33	94.68
0.6	2.8	94.05	91.78	99.86	99.74
	3	99.01	98.48	99.99	99.99
	3.2	99.85	99.74	100	99.99
	3.4	99.99	99.97	100	100
	3.6	100	100	100	100
	2	7.11	4.17	7.34	4.95
	2.2	19.69	14.15	32.46	24.21
	2.4	54.29	45.76	80.95	73.64
	2.6	84.17	78.59	98.35	97.26
0.7	2.8	96.91	95.1	99.95	99.9
	3	99.65	99.28	100	100
	3.2	99.99	99.97	100	100
	3.4	100	100	100	100

Table 2 Power (in %) of the test ψ_4 and ψ_5 for θ_0 =4, t=9, n=100 and 200, α=0.05

π	$n = 100$	$n = 100$		$n = 200$	
		ψ_4	ψ_5	ψ_4	ψ_5
	4	5.56	3.58	4.65	3.37
	4.2	9.33	4.71	12.38	5.58
	4.4	19.4	9.95	31.31	17.04
	4.6	33.8	20.81	56.28	38.36
	4.8	50.58	35.48	78.37	62.91
	5	68.14	53.07	92	82.84
0.3	5.2	80.88	67.97	97.5	93.47
	5.4	89.97	81.06	99.45	98.4
	5.6	95.31	89.59	99.83	99.53
	5.8	97.77	94.74	99.97	99.94
	6	99.05	97.48	100	99.99
	6.2	99.6	98.72	100	100
	6.4	99.85	99.5	100	100
	4	5.29	3.57	5.26	3.95
	4.2	10.24	4.86	13.8	5.8
	4.4	22.52	12.32	38.38	21.74
	4.6	41.49	26.14	68.09	49.91
	4.8	62.45	46.12	88.35	76.97
	5	78.69	65.52	97.45	92.41
0.4	5.2	90.17	81.34	99.52	98.26
	5.4	95.75	90.75	99.96	99.67
	5.6	98.55	96.16	99.99	99.97
	5.8	99.53	98.56	100	100
	6	99.88	99.52	100	100
	6.2	99.94	99.78	100	100
	6.4	99.96	99.94	100	100
	4	5.39	3.75	4.88	3.91
	4.2	11.78	5.51	15.75	6.94
	4.4	26.72	14.86	45.12	26.2
	4.6	49.72	33.44	76.69	59.34
	4.8	70.81	55	94.06	85.45
	5	86.58	75.44	98.94	96.71
0.5	5.2	95.84	89.47	99.95	99.49
	5.4	98.59	95.86	99.98	99.95
	5.6	99.62	98.82	100	100
	5.8	99.93	99.79	100	100
	6	99.98	99.86	100	100
	6.2	99.99	99.97	100	100
	6.4	100	100	100	100

Table Continued

π	$n = 100$	$n = 100$		$n = 200$	
		ψ_4	ψ_5	ψ_4	ψ_5
0.6	4	4.71	3.41	5.27	4.12
	4.2	13.45	5.88	20.35	8.06
	4.4	34.38	19.15	57.57	35.63
	4.6	62.82	45.15	89.27	74.97
	4.8	84.74	70.72	98.5	94.9
	5	95.58	88.95	99.96	99.56
	5.2	98.9	96.8	100	99.97
	5.4	99.77	99.19	100	100
	5.6	99.98	99.88	100	100
	5.8	100	100	100	100
0.7	4	4.71	3.41	5.27	4.12
	4.2	13.45	5.88	20.35	8.06
	4.4	34.38	19.15	57.57	35.63
	4.6	62.82	45.15	89.27	74.97
	4.8	84.74	70.72	98.5	94.9
	5	95.58	88.95	99.96	99.56
	5.2	98.9	96.8	100	99.97
	5.4	99.77	99.19	100	100
	5.6	99.98	99.88	100	100
	5.8	100	100	100	100

Illustrative example

Let us consider the data of Traffic Accident Research given by Kuan et al.,[5]

The data from the department of motor vehicles master driver license file

Traffic accidents	0	1	2	≥ 3
Number of drivers	4499	766	136	21

From the data we see that there is excess number of zero counts and the frequency of X is greater than or equal to 3 is 21. Generally such data is modeled by Poisson distribution. But Poisson distribution does not fit well for this data. We fit the above data for ZIPD. In ZIPD there are two parameters π and θ. In this problem $n_0 = 4499$, $n = 5422$ and estimated values of $\hat{\pi}=0.5583$ and $\hat{\theta}=0.3637$. Using these values we fit the ZIPD for the above data. The calculated Chi-square value is 0.4043 and table value of $\chi^2(1, 0.05)$ is 3.841459 and the P-value is 0.5249

Same data is fitted to ZITPD truncated at 4 and above. The parameters are $\hat{\pi}=0.5582$ and $\hat{\theta}=0.3646$ The calculated Chi-square value is 0.4018 and table value of $\chi^2(1, 0.05)$ is 3.8415 and the P-value is 0.5262. If the same data is fitted to ZITPD truncated at 5 and above. The parameters are $\hat{\pi}=0.5583$ and $\hat{\theta}=0.3638$. The calculated Chi-square value is 0.4012 and table value of $\chi^2(1, 0.05)$ is 3.8415 and the P-value is 0.5265. Here we prefer ZITPD to model the data.

Acknowledgement

None.

Conflict of interest

None.

References

1. Gupta PL, Gupta RL, Tripathi RC. Inflated Modified Power Series Distributions with Applications. *Comm Statist Theory Meth*. 1995;24(9):2355–2374.

2. Murat M, Szynal D. Non-Zero-Inflated Modified Power Series Distributions. *Commun Statist.Theory Meth*. 1998;27(12):3047–3064.

3. Patil MK, Shirke DT. Tests for equality of inflation parameters of two zero-inflated power series distributions. *Commun Statist Theory Meth*. 2011;40(14):2539–2553.

4. Patil MK, Shirke DT. Testing parameter of the power series distribution of a zero-inflated power series model. *Statistical Methodology*. 2007;4(4):393–406.

5. KuanJ, Peck RC, Janke MK. *Statistical Methods for Traffic Accidents Research*. In proceeding of the 1990 Taipei Symposium in statistics, June 28-30, 1990, (Eds), by Min-Te Chao and Philip E Cheng Taipei, Institute of Statistical Science; 1991.