Estimation of Parameters in a Zero-inflated Power Series Model

G. Nanjundan[1], T. Raveendra Naika[2]

[1]Department of Statistics, Bangalore University, Bangalore, Karnataka, INDIA.

[2]Maharani's Science College for Women, Bangalore, Karanataka, INDIA.

Corresponding Addresses:

[1]nanzundan@gmail.com, [2]raviisec@yahoo.co.uk

Research Article

Abstract: This paper discusses the maximum likelihood and the method of moments estimation of parameters in a zero-inflated power series model. This model is shown to belong to two parameter exponential family. The maximum likelihood and the moment estimators are asymptotically compared. Further, the special cases of the zero-inflated power series model viz. zero-inflated Poisson, binomial, and negative binomial models are also discussed.

Keywords: Zero-inflated power series model, ZIP, ZIB and ZINB models, Two parameter exponential family, Maximum likelihood and moment estimators, EM algorithm, Asymptotic relative efficiency.

1. Introduction

Suppose that we are interested in the distribution of the number of insects on a leaf of a tree. Insects live on leaves that are suitable for feeding and not on those which are unsuitable for feeding. Assume that the proportion of unsuitable leaves in the tree is φ. The number of insects on a suitable leaf can be a non-negative integer valued random variable with the probability mass function (p.m.f.) $p_1(x,\theta)$, $\theta\in\Theta\subset R$. If any insect is found on a leaf, then it is suitable for feeding. If a leaf has no insect on it, then it may be due to the unsuitability of the leaf or the chance variation allowed by the distribution with the p.m.f. $p_1(x,\theta)$. The p.m.f. of the number of insects (X) on any observed leaf is

$$p(x,\theta,\varphi)=\begin{cases}\varphi+(1-\varphi)p_1(0,\theta), & x=0\\ (1-\varphi)p_1(x,\theta), & x=1,2,3,\ldots\end{cases} \tag{1.1}$$

$$=\varphi\, p_0(x)+(1-\varphi)p_1(x,\theta)$$

where $p_0(x)=\begin{cases}1, & if\ \ x=0\\ 0, & if\ \ x\neq 0\end{cases}$. (1.2)

Thus the distribution of X is a mixture. When

$$p_1(x,\theta)=\frac{a(x)\theta^x}{b(\theta)},\ x=0,1,2,3,\ldots,a(x)>0,\ a(0)\neq 0$$

$p(x,\theta,\varphi)$ is called a zero-inflated power series (ZIPS) model.

1.1 Maximum Likelihood Estimation

Let $\underline{X}=(X_1, X_2, \ldots, X_n)$ be a random sample on X. Then the likelihood function is given by

$$L(\theta,\varphi\,|\,\underline{x})=\prod_{j=1}^{n}\{\varphi+(1-\varphi)p_1(0,\theta)\}^{a_j}\{(1-\varphi)p_1(x_j,\theta)\}^{1-a_j} \tag{1.3}$$

where $a_j=\begin{cases}1, & if\ \ x_j=0\\ 0, & if\ \ x_j\geq 1\end{cases}$.

It is evident that the likelihood equations $\frac{\partial\log L}{\partial\theta}=\frac{\partial\log L}{\partial\varphi}=0$ do not yield closed form expressions for the maximum likelihood (ML) estimators of the parameters θ and φ. Hence the ML estimates have to be computed by a numerical procedure. But the Newton - Raphson method might fail due to boundary problem [see McLeish and Small [4] and Sprott [9]]. Yip [10] has obtained the conditional maximum likelihood estimator of the parameter θ treating φ as a nuisance parameter. Even the conditional likelihood estimator has no closed form expression and there is loss of information. Yip [10] has also discussed the loss of information in estimating θ by the conditional likelihood approach. Kale [2] has shown that $p(x,\theta,\varphi)$ belongs to Cramer family and $\log p(x,\theta,\varphi)$ is twice differentiable w.r.t. θ and φ. He has also obtained the optimal estimating equation for θ treating φ as a nuisance parameter. This optimal estimating equation does not yield closed form expression for the estimator of θ even when $p_1(x;\theta)$ is the p.m.f. of a Poisson distribution with mean θ. But the Newton-Raphson method works well in computing the estimate of θ from this optimal estimating equation [see Nanjundan [6]]. Nanjundan [5] has also computed the ML estimates of θ and φ using the EM algorithm when $p_1(x;\theta)$ is the p.m.f. of a Poisson distribution with mean θ. In this section, the adaptation of the EM algorithm is discussed for computing the ML estimates of θ and φ in $p(x,\theta,\varphi)$ specified in (1.1). The EM algorithm is an iterative procedure to compute the ML estimates of the parameters involved in a model when the likelihood equations do not admit closed form solutions. There are E- and M-steps at each of the iteration. To implement the EM algorithm the likelihood has to be rewritten so as to accommodate missing data. The details of the EM algorithm can be found in Dempster et al. [1] and McLachlan and Krishnan [3]. Let $Z_j = 0$ or 1 according as the j-th observed leaf is unsuitable or suitable. If $X_j > 0$, then $Z_j = 1$. On the other hand, when $X_j = 0$, then $Z_j = 0$ or 1. Therefore $\{Z_j : X_j = 0\}$ becomes the set of missing observations. When $(X_1, X_2, \ldots, X_n)$ is augmented with $(Z_1, Z_2, \ldots, Z_n)$, $((X_1,Z_1), (X_2,Z_2), \ldots, (X_n,Z_n))$ becomes the complete data set. The likelihood function of the complete data is given by

$$L_c(\theta,\varphi \mid \underline{x},\underline{u}) = \prod_{j=1}^{n} \varphi^{1-u_j}\left\{(1-\varphi)p_1(x_j,\theta,\varphi)\right\}^{u_j} \tag{1.4}$$

where $u_j = \begin{cases} 1, \text{ if } x_j > 0 \\ Z_j, \text{ if } x_j = 0 \end{cases}$.

Then the log of the complete data likelihood function becomes

$$\log L_c(\theta,\varphi \mid \underline{x},\underline{u}) = \sum_{j=1}^{n}(1-u_j)\log\varphi + \sum_{j=1}^{n}u_j \log\left\{(1-\varphi)p_1(x_j,\theta,\varphi)\right\}$$

Note that

$$\log L_c(\theta,\varphi \mid \underline{x},\underline{u}) = \sum_{j:x_j>0}\left[\log(1-\varphi)+\log p_1(x_j,\theta,\varphi)\right] + \sum_{j:x_j=0}(1-Z_j)\log\varphi$$

Hence

$$E[\log L_c(\theta,\varphi \mid \underline{x},\underline{u})] = \sum_{j:x_j>0}\left[\log(1-\varphi)+\log p_1(x_j,\theta,\varphi)\right] + \sum_{j:x_j=0}(1-E(Z_j))\log\varphi + \sum_{j:x_j=0}E(Z_j)\left[\log(1-\varphi)+\log p_1(x_j,\theta,\varphi)\right]. \tag{1.5}$$

E-Step: In this step of the algorithm, $E(Z_j)$ is replaced by the conditional expectation $E(Z_j \mid \theta_0,\varphi_0,X_j=0)$, where θ_0 and φ_0 are the initial estimates of θ and φ. We get

$$E(Z_j \mid \theta_0,\varphi_0,X_j=0) = 0.P(Z_j=0 \mid \theta_0,\varphi_0,X_j=0) + 1.P(Z_j=1 \mid \theta_0,\varphi_0,X_j=0) = P(Z_j=1 \mid \theta_0,\varphi_0,X_j=0).$$

Using the Bayes theorem,

$$P(Z_j=1 \mid \theta_0,\varphi_0,X_j=0) = \frac{P(X_j=0 \mid \theta_0,\varphi_0,Z_j=1)P(Z_j=1 \mid \theta_0,\varphi_0)}{\sum_{z_j=0,1}P(X_j=0 \mid \theta_0,\varphi_0,Z_j=z_j)P(Z_j=z_j \mid \theta_0,\varphi_0)}$$

That is

$$P(Z_j=1 \mid \theta_0,\varphi_0,X_j=0) = \frac{(1-\varphi)p_1(0,\theta_0,\varphi_0)}{\varphi+(1-\varphi)p_1(0,\theta_0,\varphi_0)} = w \text{ say.} \tag{1.6}$$

Note that this conditional probability is independent of j. Therefore

$$E[\log L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})] = \sum_{j:x_j>0}\left[\log(1-\varphi)+\log p_1(x_j,\theta,\varphi)\right] + \sum_{j:x_j=0}(1-w)\log\varphi + \sum_{j:x_j=0}w\left[\log(1-\varphi)+\log p_1(x_j,\theta,\varphi)\right]. \tag{1.7}$$

M-step: In this step, $E[\log L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]$ is maximized for θ and φ. Since $E[\log L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]$ is differentiable w.r.t. θ and φ, the stationary values are given by the equations

$$\frac{\partial E[\log(L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\theta} = 0 \text{ and}$$

$$\frac{\partial E[\log(L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\varphi} = 0$$

We get

$$\frac{\partial E[\log(L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\theta} = \sum_{j:x_j>0}\left[\frac{\partial \log p_1(x_j,\theta_0,\varphi_0)}{\partial\theta}\right] + \sum_{j:x_j=0}w\left[\frac{\partial \log p_1(x_j,\theta_0,\varphi_0)}{\partial\theta}\right] \tag{1.8}$$

and

$$\begin{aligned}\frac{\partial E[\log(L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\varphi} &= \sum_{j:x_j>0}\left[\frac{\partial \log p_1(x_j,\theta_0,\varphi_0)}{\partial\varphi} - \frac{1}{1-\varphi_0}\right] + \sum_{j:x_j=0}\left(\frac{1-w}{\varphi_0}\right) + \sum_{j:x_j=0}w\left(\frac{\partial \log p_1(x_j,\theta_0,\varphi_0)}{\partial\varphi} - \frac{1}{1-\varphi_0}\right) \\ &= \sum_{j:x_j>0}\left[\frac{\partial \log p_1(x_j,\theta_0,\varphi_0)}{\partial\varphi} - \frac{1}{1-\varphi_0}\right] + n_0 w\left[\frac{\partial \log p_1(0,\theta_0,\varphi_0)}{\partial\varphi}\right] + \frac{n_0(1-\varphi_0-w)}{\varphi_0(1-\varphi_0)}.\end{aligned} \tag{1.9}$$

If θ_1 and φ_1 are the values of θ and φ that maximize $E[\log L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]$, the E-step is repeated by taking $\theta_0 = \theta_1$ and $\varphi_0 = \varphi_1$. After each iteration, the value of the likelihood $L(\theta,\varphi \mid \underline{x})$ $[or \log L(\theta,\varphi \mid \underline{x})]$ specified in (1.2) can be evaluated and observed whether it is increasing. The iterative procedure can be terminated as soon as $L(\theta,\varphi \mid \underline{x})$ $[or \log L(\theta,\varphi \mid \underline{x})]$ converges to a value correct to required number of decimal places and the corresponding θ_1 and φ_1 are the MLEs.

In case the equations

$$\frac{\partial E[\log(L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\theta}=0$$

and $\dfrac{\partial E[\log(L_c(\theta,\varphi \mid \theta_0,\varphi_0,\underline{x},\underline{u})]}{\partial\varphi}=0$

do not yield closed form expressions for the stationary values, a numerical procedure like Newton-Raphson method can be used since the second derivatives of $E[\log L_c(\theta,\varphi|\theta_0,\varphi_0,\underline{x},\underline{u})]$ w.r.t. θ and φ exist.

1.2 Two Parameter Exponential Family

In this section, along the lines of Kale [2], we show that the zero-inflated model (1.1) belongs to two parameter exponential family.

The probability mass function specified in (1.1) can be written as

$$p(x;\theta,\varphi)=\{\varphi+(1-\varphi)p_1(0,\theta)\}^{t(x)}\{(1-\varphi)p_1(x,\theta)\}^{(1-t(x))} \quad (1.10)$$

where $t(x)=\begin{cases}1, & if\ \ x=0\\ 0, & if\ \ x\geq 1\end{cases}$.

Taking log on both sides in the above equation, we get

$$\log p(x;\theta,\varphi)=t(x)\{\log(\varphi+(1-\varphi)p_1(0,\theta))\} + (1-t(x)\{\log(1-\varphi)+\log p_1(x,\theta)\}. \quad (1.11)$$

After simple rearrangement of terms, the above expression can be written as

$$p(x;\theta,\varphi)=w(x)v(\theta,\varphi)\exp[t(x)u_1(\theta,\varphi)+(1-t(x))u_2(\theta,\varphi)],$$

or

$$p(x;\theta,\varphi)=\exp[t(x)u(\theta,\varphi)+w(x)+v(\theta,\varphi)],$$

(1.12)

which is the general form of two parameter exponential family. Hence the zero-inflated power series model belongs to two parameter exponential family.

1.3 Moment Estimators

The first and the second theoretical moments of X having the p.m.f. (1.1) are

$$E(X)=\sum_{x_j=0}^{n}x_j p(x_j;\theta,\varphi)$$

$$=(1-\varphi)\sum_{x_j=1}^{n}x_j p_1(x_j;\theta)=M_{1n}$$

$$E(X^2)=\sum_{x_j=0}^{n}x_j^2 p(x_j;\theta,\varphi)$$

$$=(1-\varphi)\left[\sum_{x_j=1}^{n}x_j(x_j-1)p_1(x_j;\theta)+\sum_{x_j=1}^{n}x_j p_1(x_j;\theta)\right]=M_{2n}$$

where $M_{1n}=\dfrac{1}{n}\sum_{j=1}^{n}X_j$ and $M_{2n}=\dfrac{1}{n}\sum_{j=1}^{n}X_j^{\,2}$.

The method of moments estimators (MMEs) of θ *and* φ are obtained by using the above simultaneous equations in the following ways:

$$\frac{M_{2n}}{M_{1n}}=\frac{(1-\varphi)\left[\sum_{x_j=1}^{n}x_j(x_j-1)p_1(x_j;\theta)+\sum_{x_j=1}^{n}x_j p_1(x_j;\theta)\right]}{(1-\varphi)\left[\sum_{x_j=1}^{n}x_j p_1(x_j;\theta)\right]}$$

$$\frac{M_{2n}}{M_{1n}}=1+\frac{\left[\sum_{x_j=1}^{n}x_j(x_j-1)p_1(x_j;\theta)\right]}{\left[\sum_{x_j=1}^{n}x_j p_1(x_j;\theta)\right]}. \quad (1.13)$$

Since the zero-inflated Power series (ZIPS) model belongs to two parameter exponential family,

$$\left(\sum_{i=1}^{n}t(X_i),\ \sum_{i=1}^{n}X_i(1-t(X_i))\right)' \quad (1.14)$$

is minimal sufficient and complete for (θ, φ).

Let $\hat{\theta}_m$ and $\hat{\varphi}_m$ be the MMEs of θ and φ . Since the ZIPS model belongs to two parameter exponential family and the MMEs are based on these minimal sufficient statistics,

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_m\\ \hat{\varphi}_m\end{pmatrix}\xrightarrow{L}Z'\sim N\left(\begin{pmatrix}\theta\\ \varphi\end{pmatrix},\ \Sigma^{-1}\right),\ \ as\ n\to\infty, \quad (1.15)$$

where Σ is the Fisher information matrix and it is obtained in the next section.

1.4 Fisher Information Matrix

The logarithm of the p.m.f. specified in (1.1) becomes

$$\log p(x;\theta,\varphi)=\begin{cases}\log(\varphi+(1-\varphi)p_1(0;\theta)), & x=0\\ \log(1-\varphi)+\log p_1(x;\theta), & x=1,2,3,\ldots\end{cases} \quad (1.16)$$

$\log p(x;\theta,\varphi)$ is twice differentiable w. r. t. both θ and φ.

We get the following partial derivatives:

$$\frac{\partial\log p(x;\theta,\varphi)}{\partial\theta}=\begin{cases}\dfrac{(1-\varphi)\dfrac{\partial p_1(0;\theta)}{\partial\theta}}{p(0;\theta,\varphi)}, & x=0\\ \dfrac{\partial\log p_1(x;\theta)}{\partial\theta}, & x=1,2,3,....\end{cases} \quad \text{and}$$

$$\frac{\partial\log p(x;\theta,\varphi)}{\partial\varphi}=\begin{cases}\dfrac{1-p_1(0;\theta)}{p(0;\theta,\varphi)}, & x=0\\ \dfrac{-1}{(1-\varphi)}, & x=1,2,3,....\end{cases}$$

Using above expressions, we can verify that

$$E\left[\left(\frac{\partial\log p(x;\theta,\varphi)}{\partial\theta}\right)\right]=0\ \ and\ \ E\left[\left(\frac{\partial\log p(x;\theta,\varphi)}{\partial\varphi}\right)\right]=0$$

Further, we get

$$I_{\varphi\varphi} = E\left[\left(\frac{\partial \log p(x;\theta,\varphi)}{\partial \varphi}\right)\right]^2$$

$$I_{\varphi\varphi} = \frac{1-p_1(0;\theta)}{(1-\varphi)\,p(0;\theta,\varphi)}. \qquad (1.17)$$

$$I_{\varphi\theta} = I_{\theta\varphi} = E\left(\frac{\partial^2 \log p(x;\theta,\varphi)}{\partial \varphi\,\partial \theta}\right)$$

$$I_{\varphi\theta} = \frac{1-p_1(0;\theta)}{p(0;\theta,\varphi)}\;(1-\varphi)\frac{\partial p_1(0;\theta)}{\partial \theta} + \sum_{x=1}^{\infty}\frac{\partial \log p_1(x;\theta)}{\partial \theta}\;p_1(x;\theta). \qquad (1.18)$$

$$I_{\theta\theta} = E\left[\left(\frac{\partial \log p(x;\theta,\varphi)}{\partial \theta}\right)\right]^2.$$

After simplification, we get

$$I_{\theta\theta} = (1-\varphi)\left[\frac{\left(\frac{\partial p_1(0;\theta)}{\partial \theta}\right)^2\big((1-\varphi)p_1(0;\theta)-p(0;\theta,\varphi)\big)^2}{p(0;\theta,\varphi)\;\;p_1(0;\theta)}\right]$$

$$+(1-\varphi)\sum_{x_j=0}^{\infty}\left(\frac{\partial p_1(x_j;\theta)}{\partial \theta}\right)^2 p_1(x_j;\theta). \qquad (1.19)$$

Therefore the Fisher information matrix becomes

$$\Sigma = \begin{bmatrix} I_{\varphi\varphi} & I_{\theta\varphi} \\ I_{\varphi\theta} & I_{\theta\theta}\end{bmatrix}. \qquad (1.20)$$

The inverse of the Fisher information matrix is the variance co-variance matrix of MLEs .

Asymptotic relative efficiency of MLEs over MMEs of the parameters has been discussed in Section 5.

The specific cases of Poisson, Binomial and negative binomial distributions are discussed in the following sections

2. Zero-inflated Poisson Model

When $p_1(x;\theta)$ in (1.1) is the p.m.f. of a Poisson distribution, we get a zero-inflated Poisson model (ZIP). Then, (1.1) turns out to be

$$p(x;\theta,\varphi) = \begin{cases}\varphi + (1-\varphi)e^{-\theta}, & x = 0\\ (1-\varphi)\dfrac{e^{-\theta}\theta^x}{x!}, & x = 1,2,3,\ldots;\theta>0,\ 0<\varphi<1.\end{cases} \qquad (2.1)$$

The likelihood function of the complete data becomes

$$L_c(\theta,\varphi \mid \underline{x},\underline{u}) = \prod_{j=1}^{n}\varphi^{1-u_j}\left\{(1-\varphi)\frac{e^{-\theta}\theta^{x_j}}{x_j!}\right\}^{u_j}.$$

The computational steps of ML estimates of θ and φ using EM algorithm can be summarized as follows:

a) Choose the initial estimates of θ and φ to be

$$\theta_0 = \frac{1}{n}\sum_{j=1}^{n}x_j \quad \text{and} \quad \varphi_0 = \frac{n_0}{n} \qquad (2.2)$$

where $n = n_0 + n_g$ and $n_g =$ number of x_j's > 0 and $n_0 =$ number of x_j's equal to zero.

b) Compute $w = \dfrac{(1-\varphi_0)e^{-\theta_0}}{\varphi_0 + (1-\varphi_0)e^{-\theta_0}}$. (2.3)

c) Using the realization ($x_1, x_2, \ldots, x_n$) of the observed sample, compute the improved estimates of θ and φ by

$$\theta_1 = \frac{\sum_{j:x_j>0}^{n} x_j}{n_g + n_0 w} \quad and \quad \varphi_1 = \frac{n_0(1-w)}{n}. \qquad (2.4)$$

d) Repeat steps (b) and (c) until the difference between the successive $L(\theta,\varphi \mid \underline{x})$ [or $\log L(\theta,\varphi \mid \underline{x})$] values is less than a desired threshold level.

The corresponding values of θ_1 and φ_1 are the MLEs of θ and φ respectively.

The MMEs are discussed in section 1.3 based on that a specific case of ZIP model specified in (2.1) the first and the second theoretical moments are arrived as follows

$$E(X) = (1-\varphi)\theta \quad \text{and} \quad E(X^2) = (1-\varphi)\theta(1+\theta). \qquad (2.5)$$

According to (1.13) the MMEs of θ and φ are respectively

$$\hat{\theta}_m = \frac{M_{2n}}{M_{1n}} - 1 \quad \text{and} \quad \hat{\varphi}_m = 1 - \frac{M_{1n}^2}{M_{2n} - M_{1n}}. \qquad (2.6)$$

It is easy to see that $P(M_{1n}=0) = \{\varphi + (1-\varphi)e^{-\theta}\}^n \to 0$, as $n \to \infty$. Similarly, $P(M_{1n} = M_{2n}) \to 0$, as $n \to \infty$. Hence the problem of division by zero in these MMEs doesn't arise when n is sufficiently large.

The Fisher information matrix is

$$\Sigma = \begin{bmatrix} \dfrac{1-e^{-\theta}}{(1-\varphi)(\varphi+(1-\varphi)e^{-\theta})} & \dfrac{e^{-\theta}}{\varphi+(1-\varphi)e^{-\theta}} \\ \dfrac{e^{-\theta}}{\varphi+(1-\varphi)e^{-\theta}} & \dfrac{(1-\varphi)(\varphi+(1-\varphi)e^{-\theta}+\theta\varphi)}{\theta(\varphi+(1-\varphi)e^{-\theta})}\end{bmatrix}$$

and

$$\Sigma^{-1} = \begin{bmatrix}\Sigma_{11}^{-1} & \Sigma_{12}^{-1} \\ \Sigma_{21}^{-1} & \Sigma_{22}^{-1}\end{bmatrix}$$

where

$$\Sigma_{11}^{-1} = \frac{(1-\varphi)(\varphi+(1-\varphi)e^{-\theta}+\theta\varphi)(\varphi+(1-\varphi)e^{-\theta})}{(1-e^{-\theta})(\varphi+(1-\varphi)e^{-\theta}+\theta\varphi)-\theta e^{-2\theta}}$$

$$\Sigma_{12}^{-1}=\Sigma_{21}^{-1}=\frac{-\theta e^{-\theta}(\varphi+(1-\varphi)e^{-\theta})}{(1-e^{-\theta})(\varphi+(1-\varphi)e^{-\theta}+\theta\varphi)-\theta e^{-2\theta}}$$

$$\Sigma_{22}^{-1}=\frac{(1-e^{-\theta})(\varphi+(1-\varphi)e^{-\theta})\theta}{(1-\varphi)[(1-e^{-\theta})(\varphi+(1-\varphi)e^{-\theta}+\theta\varphi)-\theta e^{-2\theta}]}.$$

3. Zero-inflated Binomial Model

The model (1.1) becomes a zero-inflated binomial model (ZIB), $p_1(x,\theta)$ is the p.m.f. of a binomial $b(N,\theta)$ – distribution. We get

$$p(x;\theta,\varphi)=\begin{cases}\varphi+(1-\varphi)(1-\theta)^N, & x=0\\ (1-\varphi)\binom{N}{x}\theta^x(1-\theta)^{N-x}, & x=1,2,3,\ldots\end{cases} \tag{3.1}$$

$0<\varphi<1,\ 0\le\theta\le 1.$

Here we assume that N is known.

The likelihood function of the complete data for ZIB model is given by

$$L_c(\theta,\varphi\,|\,\underline{x},\underline{u})=\prod_{j=1}^{n}\varphi^{1-u_j}\left\{(1-\varphi)\binom{N}{x_j}\theta^{x_j}(1-\theta)^{N-x_j}\right\}^{u_j}. \tag{3.2}$$

The computational procedure of the maximum likelihood estimates of the parameters θ and φ is as follows:

a) Choose the initial estimates

$$\theta_0=\frac{\sum_{j=1}^{n}x_j}{nN}=\frac{samplemean}{N}\text{ and }\varphi_0=\frac{n_0}{n}.$$

b) Compute $w=\dfrac{(1-\varphi_0)(1-\theta_0)^N}{\varphi_0+(1-\varphi_0)(1-\theta_0)^N}$. (3.3)

c) Using the realization $(x_1, x_2, \ldots, x_n)$ of the observed sample, compute

$$\theta_1=\frac{\sum_{j:x_j>0}x_j}{N(n_g+n_0w)}\quad\text{and}\quad\varphi_1=\frac{n_0(1-w)}{n}, \tag{3.4}$$

In ZIB model the first and the second moments are given by

$E(X)=(1-\varphi)N\theta$

and $E(X^2)=(1-\varphi)N\theta+(1-\varphi)N(N-1)\theta^2.$

The moment estimators of θ and φ are respectively given by

$$\hat{\theta}_m=\frac{M_{1n}-M_{2n}}{(N-1)M_{1n}}\quad\text{and}\quad\hat{\varphi}_m=1-\frac{M_{1n}^2(N-1)}{N(M_{1n}-M_{2n})}. \tag{3.5}$$

The Fisher information matrix is $\Sigma=\begin{bmatrix}I_{\varphi\varphi} & I_{\theta\varphi}\\ I_{\varphi\theta} & I_{\theta\theta}\end{bmatrix}$

where

$$I_{\varphi\varphi}=\frac{1-(1-\theta)^N}{(1-\varphi)(\varphi+(1-\varphi)(1-\theta)^N)}$$

$$I_{\theta\theta}=\frac{(1-\varphi)\left[(1-\theta)\left\{\varphi+(1-\varphi)(1-\theta)^N\right\}-\theta\varphi N(N-1)\right]}{\theta(1-\theta)^2(\varphi+(1-\varphi)(1-\theta)^N)}$$

$$I_{\theta\varphi}=I_{\varphi\theta}=\frac{N(1-\theta)^{N-1}\left\{(1-\varphi)+(1-\theta)^N\right\}}{\varphi+(1-\varphi)(1-\theta)^N}.$$

4. Zero-inflated Negative Binomial Model

While $p_1(x;\theta)$ in (1.1) is the p.m.f. of negative binomial distribution, we get a zero-inflated negative binomial model (ZINB). Then, (1.1) reduces to

$$p(x;\theta,\varphi)=\begin{cases}\varphi+(1-\varphi)\theta^r, & x=0\\ (1-\varphi)\binom{x+r-1}{x}\theta^r(1-\theta)^x, & x=1,2,3,\ldots\end{cases} \tag{4.1}$$

$$0<\varphi<1,\ 0<\theta<1.$$

The likelihood function of the complete data for ZINB model is given by

$$L_c(\theta,\varphi\,|\,\underline{x},\underline{u})=\prod_{j=1}^{n}\varphi^{1-u_j}\left\{(1-\varphi)\binom{x_j+r-1}{x_j}\theta^r(1-\theta)^{x_j}\right\}^{u_j}, \tag{4.2}$$

The computation of maximum likelihood estimates of φ and φ in this model are

a) Choose the initial estimates $\theta_0=\dfrac{sample\ mean}{sample\ variance}$ and

$$\varphi_0=\frac{n_0}{n}.$$

b) Compute $w=\dfrac{(1-\varphi_0)\theta_0^{-r}}{\varphi_0+(1-\varphi_0)\theta_0^{-r}}$. (4.3)

c) Using the realization $(x_1, x_2, \ldots, x_n)$ of the observed sample, compute

$$\theta_1=\frac{r(n_g+n_0w)}{\sum_{j:x_j>0}x_j+r(n_g+n_0w)}\quad\text{and}\quad\varphi_1=\frac{n_0(1-w)}{n}. \tag{4.4}$$

The MLEs of the parameters in a ZINB model are obtained based on the similar lines of previous sections.

In this model we get,

$$E(X)=(1-\varphi)r\left(\frac{1-\theta}{\theta}\right)$$

and $E(X^2)=(1-\varphi)r\left(\dfrac{1-\theta}{\theta}\right)\left(1+\dfrac{(1+r)\theta}{(1-\theta)}\right).$

The moment estimators of the parameters θ and φ are as follows

$$\hat{\theta}_m=\frac{M_{2n}-M_{1n}}{rM_{1n}+M_{2n}}\quad\textit{and}\quad\hat{\varphi}_m=1-\frac{M_{2n}-M_{1n}}{r(r+1)}. \tag{4.5}$$

The Fisher information matrix is given below

$$\Sigma=\begin{bmatrix}I_{\varphi\varphi} & I_{\theta\varphi}\\ I_{\varphi\theta} & I_{\theta\theta}\end{bmatrix}$$

where

$$I_{\varphi\varphi} = \frac{1-\theta^r}{(1-\varphi)(\varphi+(1-\varphi)\theta^r)}$$

$$I_{\theta\theta} = \frac{(1-\varphi)\Big((r(1-\theta)+\theta^2)(\varphi+(1-\varphi)\theta^r)-(1-\theta)\varphi r(r-1)\Big)}{\theta^2(1-\theta)(\varphi+(1-\varphi)\theta^r)}$$

$$I_{\theta\varphi} = I_{\varphi\theta} = \frac{\varphi\Big(\theta^r - r - \theta\Big) - \theta^{r+1}(1-\varphi)}{\theta(\varphi+(1-\varphi)\theta^r)}.$$

5. Asymptotic Relative Efficiency

An empirical comparison of MLEs and MMEs of the parameters in a ZIP model is made by Nanjundan, Loganathan and Naika [7]. The relative efficiency (ARE) of MMEs with respect to MLEs of the parameters are compared analytically in case of ZIP model [see Nanjundan and Naika [8]]. The MLEs of the parameters θ and φ based on ZIPS model in all the three cases do not yield closed form expressions but MMEs are obtained closed form expressions. According to (1.15) the estimators of the parameters are asymptotically normally distribution. Hence the asymptotic relative efficiencies of the estimators are compared analytically. Since the zero-inflated power series model in (1.1) belongs to two parameter exponential family, the MLEs of θ and φ are also asymptotically normal and

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_{mle}\\ \hat{\varphi}_{mle}\end{pmatrix} \xrightarrow{L} Z' \sim N\left(\begin{pmatrix}\theta\\ \varphi\end{pmatrix}, \Sigma^{-1}\right), \; as\ n \to \infty.$$

Hence the asymptotic relative efficiency of $\hat{\theta}_m$ with respect to $\hat{\theta}_{mle}$ is

$$ARE(\hat{\theta}_{mle}, \hat{\theta}_m) = \frac{AV(\theta_m)}{AV(\theta_{mle})}$$

$$= 1.$$

Therefore, the MMEs and the MLEs of θ are asymptotically equally efficient. Similarly results holds in the case of φ.

6. Conclusion

The MLEs of the parameters in the ZIP, ZIB and ZINB models have no closed form expressions and computing them even by the EM algorithm required very much computer aid. Whereas the computation of MMEs are simple and yields closed form expressions. The MMEs and the MLEs are asymptotically equally efficient.

References

1. Dempster, A.P., Laird, N.M., and Rubin, D.B., "Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion)", J. Roy. Statist. Soc. Ser. B, 39, 1-38, 1977.
2. Kale, B.K., "Optimal estimating equations for discrete data with higher frequencies at a point", J. Ind. Stat. Assoc., 36, 125 – 136, 1998.
3. McLachlan, G.J. and Krishnan, T., "The EM Algorithm and Extensions", John Wiley and Sons, New York, 1997.
4. McLeish, D.L. and Small, C.G., "Theory and Applications of Statistical Inference Functions", Springer Lecture Notes in Statistics 44, 1988.
5. Nanjundan, G., "An EM algorithmic approach to maximum likelihood estimation in a mixture model", Vignana Bharathi, Vol. 18, 7 – 13, 2006.
6. Nanjundan, G., "On the computation of the maximum likelihood estimates of the parameters in a mixture model", Mapana Journal of Sciences, Vol. 6, No. 2, 57 – 66, 2007.
7. Nanjundan, G, Loganathan and Raveendra Naika, T., "An empirical comparison of maximum likelihood and moment estimators of parameters in a zero-inflated Poisson model", Mapana Journal of Sciences, Vol.8, No.2, 59-72, 2009.
8. Nanjundan, G and Raveendra Naika, T., "Asymptotic Comparison of Method of Moments Estimators and Maximum Likelihood Estimators of Parameters in Zero-Inflated Poisson Model", Applied Mathematics, 3, 610-617, 2012.
9. Sprott, D. A., "Maximum likelihood and small samples: estimation in the presence of nuisance parameters", Biometrika 67, 515-523, 1980.
10. Yip, P., "Inference about the mean of Poisson distribution in the presence of a nuisance parameter", Austral. J. Statist., 30 (3), 299-306, 1988.