ISSN 0975-3303
MJS, Vol. 8, No. 2, July - December 2009, pp. 59-72
https://doi.org/10.12725/mjs.15.7

AN EMPIRICAL COMPARISON OF MAXIMUM LIKELIHOOD AND MOMENT ESTIMATORS OF PARAMETERS IN A ZERO-INFLATED POISSON MODEL

G. Nanjundan, A. Loganathan** and T. Raveendra Naika****

ABSTRACT

The zero-inflated Poisson model is discussed in the context of a real life situation. The moment estimators of the parameters in the model are obtained and they are compared with the respective maximum likelihood estimators through simulation.

Keywords: *Zero-inflated models, Poisson distribution, Maximum likelihood estimators, EM Algorithm, Moment estimators, Asymptotic normality.*

* Dept. of Statistics, Bangalore University, Bangalore 560 056, e-mail: nanzundan@gmail.com
** Dept. of Statistics, M.S. University, Tirunelveli 627 012
*** Maharani's Science College for Women, Bangalore 560 001.

1. Introduction

There is a surge of interest in zero-inflated models because they are readily applicable in many real life situations and also appropriate in count-regression models [see Lambert (1992), Hall (2000), Jiang and Paul (2009), Erdman et al (2008). Situations where data are from a mixture of two distributions such that one produces only zeros and the other produces non-negative integer valued observations are common. Mixtures of a singular or degenerate distribution at zero and a non-negative integer valued distribution like Poisson or negative binomial are suitable models in such situations. These mixture distributions are over dispersed and are also known as zero-inflated models.

A zero-inflated model which is a mixture of a distribution degenerate at zero and a Poisson distribution is considered here. The moment estimators of the parameters involved in the model are obtained and empirically compared with the maximum likelihood estimators.

Suppose that we are interested in the number of insects (X) per leaf in a tree. Usually insects live on leaves that are suitable for feeding. If a leaf is unsuitable for feeding, then no insect lives on it. Suppose that the proportion of unsuitable leaves in a tree is φ and the number of insects on a suitable leaf follows a Poisson distribution with mean θ. 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 or the chance variation allowed by the Poisson distribution. Then, the probability mass function (pmf) of X is

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

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

where $p_o(x) = \begin{cases} 1, \; if \; x = 0 \\ 0, \; if \; x \neq 0 \end{cases}$ and $p_1(x) = \dfrac{e^{-\theta}\theta^x}{x!}, \; x = 0, 1, 2, \ldots, \theta > 0.$

Thus the distribution of X is a mixture of a distribution degenerate at zero and a Poisson distribution. Also, this is known as a zero-inflated Poisson model.

2. Maximum Likelihood Estimation

If $\underline{X} = (X_1, X_2, X_3, \ldots, Xn)$ is a random sample on X with the pmf specified in (1.1), the likelihood function is given by

$$L(\theta, \varphi \mid \underline{x}) = \prod_{j=1}^{n} P(X_j = x_j)$$

$$= \prod_{j=1}^{n} \left\{\varphi + (1-\varphi)e^{-\theta}\right\}^{1-a_j} \left\{(1-\varphi)\frac{e^{-\theta}\theta^{x_j}}{x_j!}\right\}^{a_j}, \theta > 0,\ 0 \le \varphi \le 1,$$

where $a_j = \begin{cases} 0, \text{ if } x_j = 0 \\ 1, \text{ if } x_j \ge 1 \end{cases}$

Since the above likelihood function does not yield closed form expressions for the maximum likelihood estimators (MLEs) of θ and φ, numerical procedures like Newton-Raphson method can be employed to compute them. Yip (1988) has observed that due to the flat surface of $L(\theta, \varphi|\underline{x})$ and boundary problems it is difficult to find the MLEs by numerical procedures. Yip (1988), motivated by a criterion of Cox (1958), has obtained the conditional MLE of θ treating φ as a nuisance parameter. Yip (1991) has also extended the conditional maximum likelihood estimation procedure to the case of zero-inflated binomial and negative binomial models.

Kale (1998) has derived, in the spirit of Godambe (1976), the optimal estimating equation for θ ignoring φ when $p_1(x)$ in (1.1) is the pmf of a general power series distribution. The estimators obtained by Yip (1988, 1991) and Kale (1998) turn out to be the same.

Both Yip (1988, 1991) and Kale (1998) have estimated θ treating φ as a nuisance parameter. But if one is interested in the spread of a disease which is responsible for leaves becoming unsuitable for feeding, then φ is of significant interest and it can not be treated as a nuisance parameter.

When the likelihood functions have complicated structures and their maximization by numerical methods is difficult, a popular and remarkably simple alternative procedure is given by the Expectation Maximization (EM) algorithm. It is an iterative procedure and in each of the iterations there are two steps, the Expectation step

(E-step) and the Maximization step (M-step). This algorithm was developed by Dempster, Laird, and Rubin (1977) who synthesized an earlier formulation in many particular cases and gave a general method of finding the MLEs in a variety of problems. See McLachlan and Krishnan (1997) and Krishnan (2004) for a detailed discussion.

Nanjundan (2006) has obtained the E- and M-steps by rewriting the likelihood function so as to accommodate missing data. Let

$$Z_i = \begin{cases} 1, \text{ if the j-th sampled leaf is suitable} \\ 0, \text{ otherwise.} \end{cases}$$

$$\text{Then, } P(Z_i = 1) = 1-\varphi = 1 - P(Z_i = 0), \; i = 1, 2, 3, \ldots, n.$$

Suppose that $\underline{X} = (X_1, X_2, X_3, \ldots, X_n)$ is a sample observed on X. Then, $[(X_1, Z_1), (X_2, Z_2), \ldots, (X_n, Z_n)]$, becomes a complete sample when $(X_1, X_2, X_3, \ldots, X_n)$ is augmented with $(Z_1, Z_2, Z_3, \ldots, Z_n)$. Note that if $X_i > 0$, then $Z_i = 1$ and if $X_i = 0$, then $Z_i = 0$ or 1. In other words, for $X_i = 0$, we have no information on Z_i and hence $\{Z_i : X_i = 0\}$ can be treated as missing data.

The likelihood function of the complete data is given by

$$L_c(\theta, \varphi \mid \underline{x}, \underline{u}) = \prod_{i=1}^{n} \varphi^{1-u_i} \left\{ (1-\varphi) \frac{e^{-\theta} \theta^{x_i}}{x_i!} \right\}^{u_i},$$

where $u_i = 1$ if $X_i > 0$ and $u_i = Z_i$ if $X_i = 0$.

In the E-step, the expectation of the likelihood function of the complete data is taken and $E(Z_i)$ is replaced by the conditional Expectation $E(Z_i \mid \theta_o, \varphi_o, X_i = 0)$, where θ and φ are respectively the initial estimates of θ and φ. In the M-step, $E[L_c(\theta, \varphi \mid x, u)]$ is maximized with respect to θ and φ. If θ_1 and φ_1 are the values of θ and φ which maximize $E[L_c(\theta, \varphi \mid x, u)]$, then the E-step is repeated using θ_1 and φ_1. Nanjundan (2006) has discussed these steps much in detail.

The computational details of these steps can be summarized as follows:

a) Choose the initial estimates θ_o and φ_o.

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

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

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

where n_g and n_o are respectively the number of observations greater than zero and equal to zero.

d) Repeat the steps b) and c) by fixing $\theta_o = \theta_1$ and $\varphi_o = \varphi_1$

A reasonable initial estimate of φ is n_o/n and the mean of the observed sample can be taken as an initial estimate of θ. If $\{\theta_n\}_{n=1}^{\infty}$ and $\{\varphi_n\}_{n=1}^{\infty}$ are respectively the sequence of iterates of the estimates of θ and φ and they converge, then their limits are the MLEs of θ and φ. For proof see Dempster et al. (1977). Nanjundan (2006) has observed that the above sequences of iterates converge for every sample simulated for various combinations of θ and φ. Nanjundan (2007) has further compared the maximum likelihood and conditional likelihood estimates of θ and has argued that the estimate suggested by Yip (1988) for the nuisance parameter φ may turn out to be negative.

3. Moment Estimators

When X has the zero-inflated Poisson distribution specified in (1.1), it is easy to observe that

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

When $\underset{\sim}{X} = (X_1, X_2, X_3, \ldots, X_n)$ is a random sample on X, the moment estimators of θ and φ are given by the simultaneous equations

$$(1-\varphi)\,\theta = M_{1n} \text{ and } (1-\varphi)\,\theta\,(1+\theta) = M_{2n}$$

where $M_{1n} = \frac{1}{n}\sum_{j=1}^{n} X_j$ and $M_{2n} = \frac{1}{n}\sum_{j=1}^{n} X_j^2$ are the first two sample moments.

Note that $\frac{M_{2n}}{M_{1n}} = 1+\theta$. Hence, the moment estimator of θ is

$$\hat{\theta}_m = \left\{\frac{M_{2n}}{M_{1n}} - 1\right.$$

Similarly, the moment estimator of φ is

$$\hat{\varphi}_m = 1 - \frac{M_{1n}^2}{M_{2n} - M_{1n}}.$$

In both these estimators, we have the problem of inadmissibility when either $M_{1n} = 0$ or $M_{1n}^2 > M_{2n} - M_{1n}$. Hence it is reasonable to redefine these estimators as follows

$$\hat{\theta}_m = \begin{cases} \frac{M_{2n}}{M_{1n}} - 1, & \text{if } M_{1n} \neq M_{2n} \\ 0, \text{ if } M_{1n} = 0 \text{ or } M_{1n} = M_{2n} \end{cases}$$

and

$$\hat{\varphi}_m = \begin{cases} 1 - \frac{M_{1n}^2}{M_{2n} - M_{1n}}, \text{if } M_{1n} \neq M_{2n} \\ 0, \quad \text{if } M_{1n} = M_{2n} \text{ or } M_{1n}^2 > M_{2n} - M_{1n} \end{cases}$$

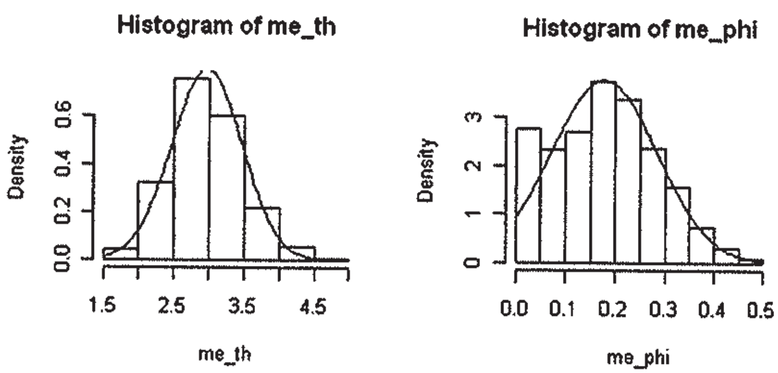
Note that for a given sample the above estimates can easily be computed even with simple devices such as pocket calculators.

4. Comparison of ML and Moment Estimators

Generally MLEs are preferred because of their properties and the most important among them is the asymptotic normality which gives way for large sample inference. In the case of regular distributions like binomial, Poisson, exponential etc., the moment estimators coincide with the MLEs. Moment estimators are always easy to compute whereas it is quite common that MLEs do not have closed form expressions and sophisticated computing facilities are required to compute them.

The asymptotic properties of the MLEs and the moment estimators of θ and φ in the zero-inflated Poisson model have been investigated by a sufficiently large scale simulation carried through the *R* software. One thousand samples of size $n = 25$, *50, 100* and *250* each were generated from the zero-inflated Poisson distribution for two different combinations $(\theta = 3.0,\ \varphi = 0.2)$ and $(\theta = 3.0,\ \varphi = 0.7)$ of the model parameters. The MLEs and the moment estimates of θ (mle_th and me_th) and φ (mle_phi and me_phi) have been computed for each sample. The histograms of the estimates are presented in Figures 1 through 8.

A close perusal of the histograms indicates that the sampling distributions of the MLEs and the moment estimates of the model parameters converge to normal distributions as the sample size increases. The mean of the normal distribution in each case is the true value of the parameter used for simulating the samples. The sampling distributions of MLEs of θ and φ exhibit normality even when the sample size is so moderate as 25.



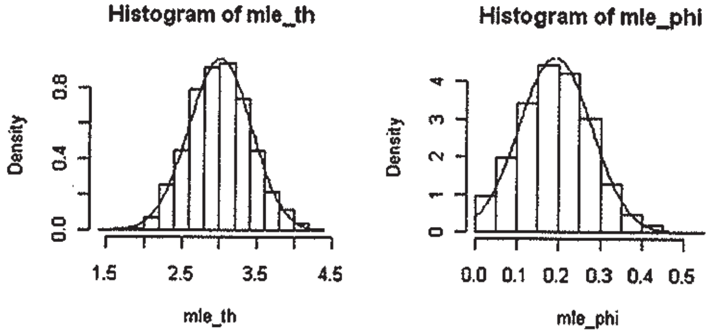


Figure - 1

Histograms of the moment and maximum likelihood estimates of θ and φ based on 1000 samples of size 25 each drawn from the distribution 0.2p0(x) + 0.8p1(x,3)

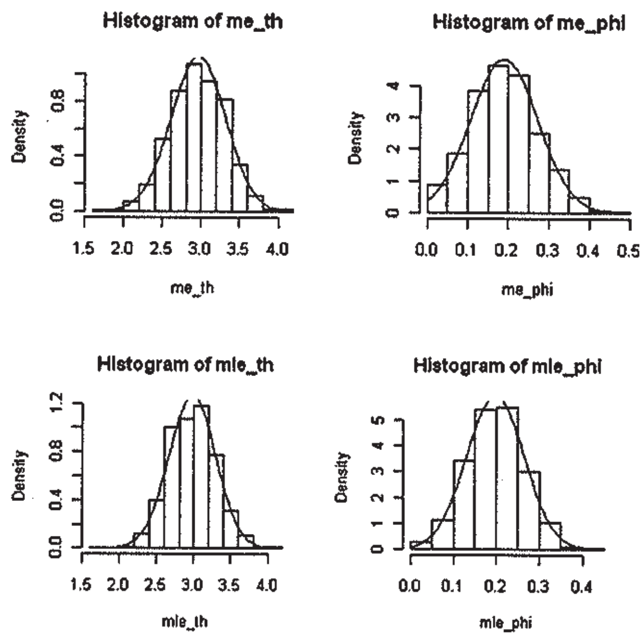


Figure - 2

Histograms of the moment and maximum likelihood estimates of θ and φ based on *1000* samples of size *50* each drawn from the distribution $0.2p_0(x) + 0.8p_1(x,3)$

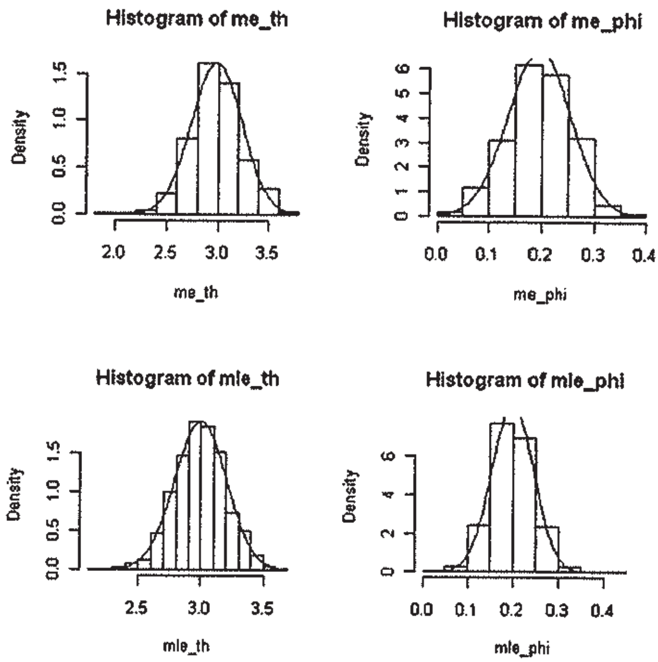
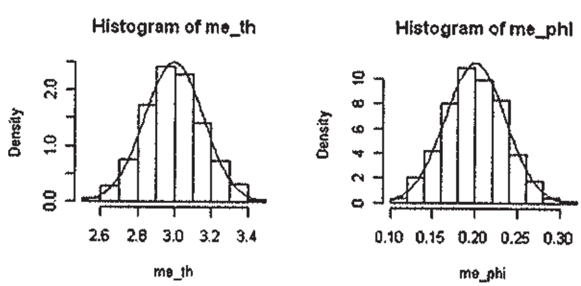


Figure - 3

Histograms of the moment and maximum likelihood estimates of θ and φ based on 1000 samples of size 100 each drawn from the distribution 0.2p0(x) + 0.8p1(x,3)



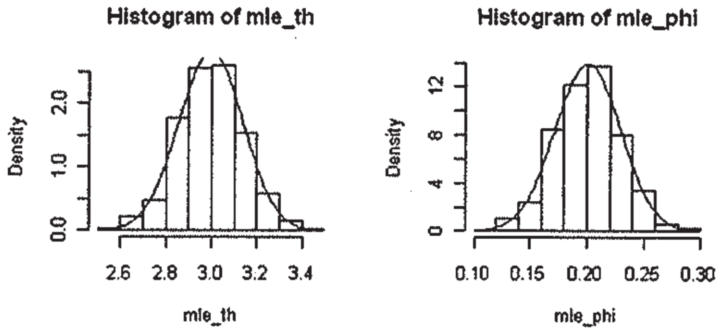


Figure - 4

Histograms of the moment and maximum likelihood estimates of θ and φ based on 1000 samples of size 250 each drawn from the distribution 0.2p0(x) + 0.8p1(x,3)

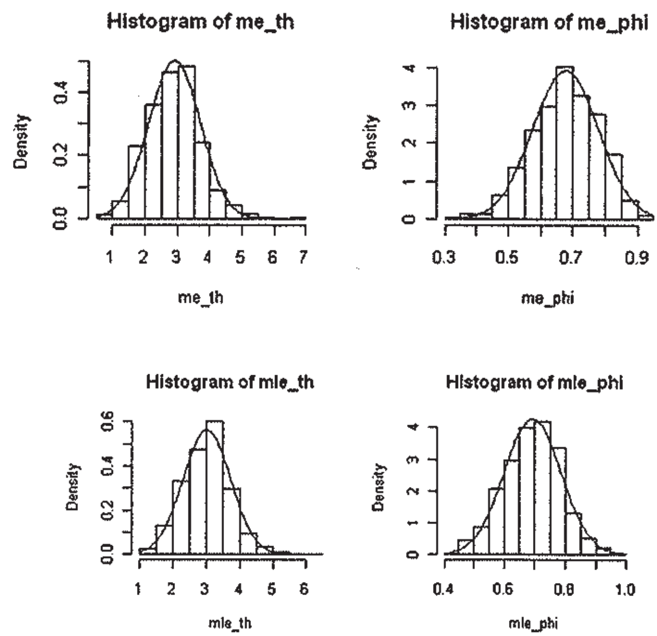


Figure - 5

Histograms of the moment and maximum likelihood estimates of θ and φ based on 1000 samples of size 25 each drawn from the distribution 0.7p0(x) + 0.3p1(x,3)

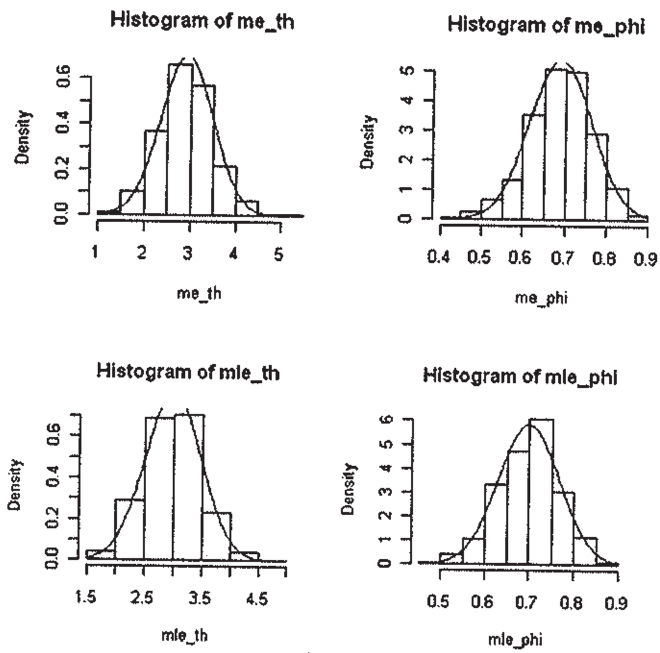
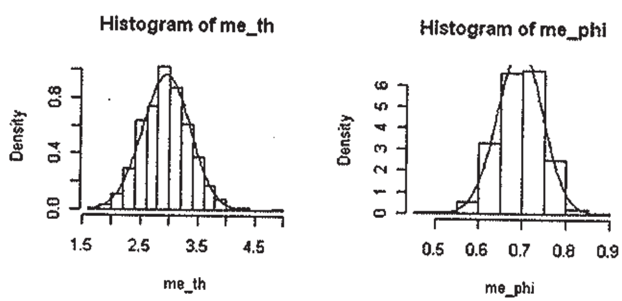


Figure - 6

Histograms of the moment and maximum likelihood estimates of θ and φ based on 1000 samples of size 50 each drawn from the distribution 0.7p0(x) + 0.3p1(x,3)



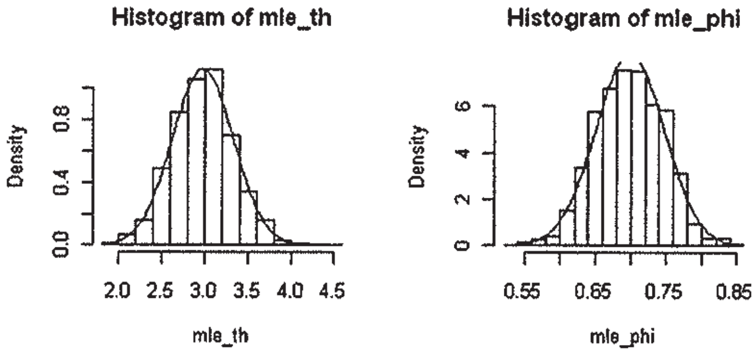


Figure - 7

Histograms of the moment and maximum likelihood estimates of and based on 1000 samples of size 100 each drawn from the distribution 0.7p0(x) + 0.3p1(x,3)

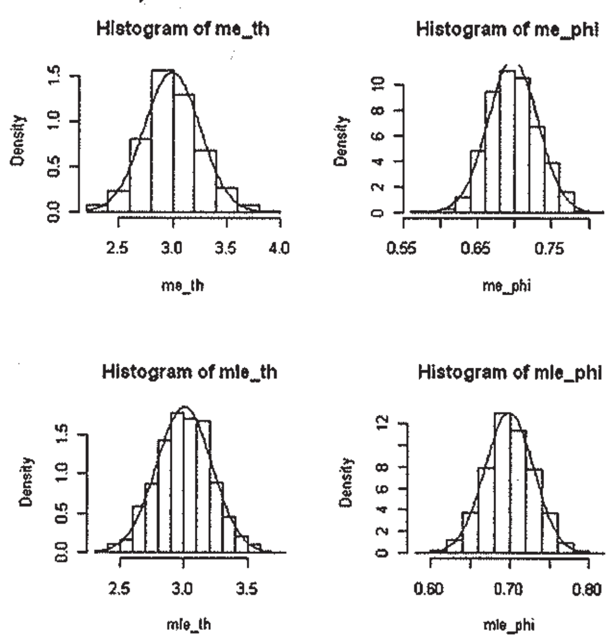


Figure - 8

Histograms of the moment and maximum likelihood estimates of θ and φ based on 1000 samples of size 250 each drawn from the distribution 0.7p0(x) + 0.3p1(x,3)

5. Conclusion

The moment estimators of the parameters of the zero-inflated Poisson model have closed form expressions and can be computed with simple computational devices like pocket calculators. Whereas the maximum likelihood estimators have no closed form expressions and the computations of the estimates for a given sample requires sophisticated computers and programming skill. Though both the moment and maximum likelihood estimators are asymptotically normal, the maximum likelihood estimators exhibit normality even for moderately small samples. Hence, the maximum likelihood estimators are preferable to the moment estimators when sample sizes are not large enough. But the moment estimators are easy to compute and perform equally well when the sample size is large.

References

1. Cox, D.R. (1958). Some problems connected with statistical inference. *Ann. Math. Statist.*, 29, 357–372.

2. Dempster, A.P. , Laird, N.M., and Rubin, D.B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. Ser. B*, 39, 1-38.

3. Erdman, D., Jackson, L., and Sinko, A. (2008). Zero-inflated Poisson and Zero-inflated negative binomial models using the countreg procedure, Paper 322- 2008, SAS Institute Inc, Cary, NC.

4. Hall, D.B. (2000). Zero-inflated Poisson and binomial regression with random effects: A case study. *Biometrics*, 56, 1030-1039.

5. Godambe, V.P. (1976). Conditional likelihood and unconditional optimum estimating equations. *Biometrika*, 63, 277–284.

6. Jiang, X. and Paul, S.R. (2009). A zero-inflated bivariate Poisson regression model and application to some dental epidemiological data, Calcutta Statistical Association Bulletin, 61 (Special Sixth Triennial Symposium Volume).

7. Kale, B.K. (1998). Optimal estimating equations for discrete data with higher frequencies at a point. *J. Ind. Stat. Assoc.*, 36, 125–136.

8. Krishnan, T. (2004). The EM Algorithm. In *Statistical Computing*, Eds. D. Kundu and A. Basu, Narosa, New Delhi.

9. Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, 34, No.1, 1-14.

10. McLachlan, G.J. and Krishnan, T. (1997). *The EM Algorithm and Extensions*. John Wiley and Sons, New York.

11. Nanjundan, G. (2006). An EM algorithmic approach to maximum likelihood estimation in a mixture model. *Vignana Bharathi*, Vol. 18, 7–13.

12. Nanjundan, G. (2007). On the computation of the maximum likelihood estimates of the parameters in a mixture model. *Mapana*, Vol. 6, No. 2, 57–66.

13. Yip, P. (1988). Inference about the mean of a Poisson distribution in the presence of a nuisance parameter. *Austral. J. Statist.*, 30 (3), 299-306.

14. Yip, P. (1991). Conditional inference on a mixture model for the analysis of count data. *Commn. in Statist. – Theory and Methods*, 20 (7), 2045–2057.