BIAS-REDUCED MAXIMUM LIKELIHOOD ESTIMATION OF THE ZERO-INFLATED POISSON DISTRIBUTION

JACOB SCHWARTZ[1] AND DAVID E. GILES[2]

Revised, March 2013: *Forthcoming in Communications in Statistics - Theory & Methods*

Abstract. We investigate the small-sample quality of the maximum likelihood estimators (MLE) of the parameters of a zero-inflated Poisson distribution (ZIP). The finite-sample bias of the MLE is determined to $O(n^{-1})$ using an analytic bias-reduction methodology based on the work of Cox and Snell (1968) and Cordeiro and Klein (1994). Monte Carlo simulations show that the MLEs have very small percentage biases for this distribution, but the analytic bias reduction methods essentially eliminate the bias without adversely affecting the mean squared errors of the estimators. The analytic adjustment compares favourably with the parametric bootstrap bias-corrected estimator, in terms of bias reduction itself, as well as with respect to mean squared error and Pitman's nearness measure.

Keywords: Maximum Likelihood Estimation, Zero Inflated Poisson Distribution, Bias Reduction, Finite Sample Properties.

Mathematics Subject Classification: 62F10; 62F40; 62N02; 62N05.

[1]Department of Economics, University of British Columbia, Vancouver, B.C., Canada, V6T 1Z1. js@interchange.ubc.ca.

[2]Contact Author. Department of Economics, University of Victoria, Victoria, B.C., Canada, V8W 2Y2. dgiles@uvic.ca.

1. Introduction

In many statistical applications involving the use of "count" (i.e., non-negative integer) data, a characteristic of the sample is that it contains far too many zero-valued observations to be plausibly modelled with the Poisson or negative binomial distributions. This poses a challenge for the researcher, and various approaches have been suggested to deal with this phenomenon. These include, among others, the use of 'hurdle' models (Cragg, 1971; Mullahy, 1986) and the application of quantile regression methods to 'smoothed' count data as outlined by Machado and Santos Silva (2005). The zero-inflated Poisson (ZIP) model, proposed by Lambert (1992) following Cohen (1963) and Johnson, Kemp and Kotz (2005), is often an attractive choice in such situations. This model posits two regimes, one in which the data are Poisson-distributed, and one in which only zero values are generated, and nature chooses between them randomly. The model easily allows for the introduction of covariates, both through the Poisson mean, and through the regime selection mechanism. An attractive by-product of the model is that can account for "over-dispersed" data – that is, data for which the variance exceeds the mean, which is also a very commonly encountered phenomenon.

The ZIP model is typically estimated by the method of maximum likelihood, so sharp inferences are assured if the sample size is sufficiently large. However, the properties of the maximum likelihood estimator (MLE) for the ZIP model have not previously been investigated for the case where the sample size is small, and so small-sample bias is a potential concern. This paper addresses this issue using both analytic and simulation-based methods.

Specifically, we find that the MLE for the parameters of the ZIP model exhibits very little bias, even in relatively small samples. This finding is very positive for practitioners. In addition, the bias that is present can be almost eliminated by using the general first-order analytic bias correction proposed by Cox and Snell (1968), and others. Alternatively, a parametric bootstrap bias correction can be used with approximately equal success. Both of these approaches deal with the bias without compromising the mean squared error of the MLE.

In the next section we provide a more formal description of the ZIP model. Section 3 discusses the analytic bias reduction procedure of Cox and Snell (1968), and we apply this to the ZIP model in section 4. The performance of this bias-adjustment technique is compared, in section 5, with that of a bootstrap-based bias adjustment; and a simple illustrative application is provided in section 6. Our conclusions appear in section 7, and some supplementary mathematical results are given in the appendix.

2. Zero-Inflated Poisson Process

Lambert (1992) introduced the zero-inflated Poisson regression model with covariates to model the number of manufacturing defects from a soldering experiment conducted by AT&T Bell Laboratories. Although the standard Poisson model allows for the presence of some zeros, the zero-inflated Poisson model allows excess zeros to arise from a special data generating processes encompassing two regimes. To use Lambert's example, in the first regime, manufacturing defects do not occur because the equipment is set up properly. In the second regime, defects, while not inevitable, are possible and their occurence follows a Poisson process. Another example of a process that may be characterised by excess zeros is the demand for recreation (Cameron and Trivedi, 1998, p. 123). For instance, someone may report 'zero' fishing trips in a particular month because they do not fish, or they do fish, but didn't happen to go fishing during the time frame of interest.

Formally, suppose that in the first regime, R_I, every observation is a zero count, while in the second regime, R_{II}, observations are generated according to a Poisson process. In addition, suppose that:

$$P(y_i) = \begin{cases} \omega & y_i \in R_I \\ (1-\omega) & y_i \in R_{II} \end{cases} \quad i = 1, 2, \ldots n.$$

The zero-inflated Poisson distribution is given by:

$$P(y_i) = \begin{cases} \omega + (1-\omega)e^{-\lambda} & y_i = 0 \\ (1-\omega)e^{-\lambda}\lambda^y/y! & y_i \neq 0 \end{cases} \quad i = 1, 2, \ldots n$$

As with the Poisson distribution, covariates can be introduced through the conditional mean of the Poisson process. That is, we can assign $\lambda_i = exp(x_i^{'}\beta)$, where x_i is the i^{th} observation on the vector of covariates, and β is a $(k \times 1)$ vector of parameters. The mean and variance of the ZIP model are $E(Y_i|x_i) = (1-\omega)\lambda_i$, $Var(Y_i|x_i) = (\lambda_i + \omega\lambda_i^2)$. The probability that an observation comes from R_I can also be determined by covariates, by assigning $\omega = \omega_i = exp(z_i'\gamma)/(1 + exp(z_i'\gamma))$, where z_i is the i^{th} observation on a vector of (possibly different) covariates and γ is a $(p \times 1)$ vector of parameters. Some recent studies using a ZIP specification include Crépon and Duguet (1997), Böhning *et al.* (1998), and Carrivick *et al.* (2003).

3. Bias-Reduced Maximum Likelihood Estimation

Bartlett (1953a) showed that, for a single parameter log-likelihood function satisfying the usual regularity conditions, it is possible to analytically approximate the bias of the maximum likelihood estimator, $\hat{\theta}$, to $O(n^{-1})$ - even when $\hat{\theta}$ does not admit a closed-form expression. Haldane and Smith (1956), and Shenton and Bowman (1963) also derive expressions for this bias of the MLE for the one-parameter case. Bartlett (1953b) and Haldane (1953) obtain analytic approximations for two-parameter log-likelihood functions. The methods undertaken by these researchers typically involve Taylor-series approximations that can become cumbersome for the multi-parameter case - see Shenton and Bowman (1977). The latter authors call the methods employed by Cox and Snell (1968) *"an adjusted order of magnitude process"*. Giles (2012) argues that this method is easy to apply to the multi-parameter case, especially with respect to first-order bias approximations.

Let $l(\theta)$ be the log-likelihood function where the p-dimensional vector of parameters, θ, is to be estimated using a sample of n observations. Assume that the log-likelihood function is well behaved and satisfies the usual regularity conditions (Duguét, 1937; Cramér, 1946).

The joint cumulants of the derivatives of $l(\theta)$ are:

$$\begin{aligned} k_{ij} =& \; E\left(\frac{\partial^2 l}{\partial\theta_i\theta_j}\right); \quad i,\, j = 1,\, 2, \ldots,\, p \;. \\ k_{ijl} =& \; E\left(\frac{\partial^3 l}{\partial\theta_i\partial\theta_j\partial\theta_l}\right); \quad i,\, j,\, l = 1, 2, \ldots,\, p \;. \\ k_{ij,l} =& \; E\left[\left(\frac{\partial^2 l}{\partial\theta_i\partial\theta_j}\right)\left(\frac{\partial l}{\partial\theta_l}\right)\right]; \quad i,\, j,\, l = 1,\, 2, \ldots, p. \end{aligned}$$

The derivatives of the cumulants are denoted:

$$k_{il}^{(l)} = \partial k_{ij}/\partial\theta_l; \;\; i,\, j,\, l = 1,\, 2, \ldots,\, p.$$

Fisher's information matrix is $K = \{-k_{ij}\}$, each element of which is $O(n)$. Cox and Snell (1968) showed that with a sample of independent data that is not necessarily identically distributed, the bias of the s^{th} element of $\hat{\theta}_s$ can be written as:

$$Bias(\hat{\theta_s}) = \sum_{i=1}^{p}\sum_{j=1}^{p}\sum_{l=1}^{p} k^{si}k^{jl}[\frac{1}{2}k_{ijl} + k_{ij,l}] + O(n^{-2}); \;\; s = 1,\, 2, \ldots,\, p$$

where k^{ij} is the $(i,j)^{th}$ element of the inverse of the information matrix, K. Furthermore, Cordeiro and Klein (1994) showed that the previous equation can be written in the following convenient form, and can be applied even when the sample data are non-independent :

$$Bias(\hat{\theta_s}) = \sum_{s=1}^{p} k^{si}\sum_{j=1}^{p}\sum_{l=1}^{p}[k_{ij}^{(l)} - \frac{1}{2}k_{ijl}] + O(n^{-2}); \;\; s = 1,\, 2, \ldots,\, p.$$

Define $a_{ij}^{(l)} = k_{ij}^{(l)} - \frac{1}{2}k_{ijl}$ $\forall i,j,l = 1,\, 2, \ldots,\, p$ and define the matrices $A^{(l)} = \{a_{ij}^{(l)}\}$ $\forall i,\, j,\, l = 1,\, 2, \ldots,\, p$. After concatenating the matrices, $A = [A^{(1)}|A^{(2)}|\ldots|A^{(p)}]$, we are able to write the bias of $\hat{\theta}$ in the following way (Cordeiro and Klein, 1994):

$$Bias(\hat{\theta}) = K^{-1}A\, vec(K^{-1}) + O(n^{-2}).$$

Here, $vec(.)$ denotes the "vectorization" operator, which stacks the columns of the matrix in question one above the other, forming one extended column vector. Finally, define the bias adjusted-MLE as:

$$\tilde{\theta} = \hat{\theta} - \hat{K}^{-1}\hat{A}\, vec(\hat{K}^{-1}),$$

where $\hat{K} = K|_{\hat{\theta}}$ and $\hat{A} = A|_{\hat{\theta}}$. One of the advantages of this method is that these expressions can be evaluated when the likelihood equations for the concerned problem do not admit a closed-form, analytic solution. In such situations we can obtain bias-corrected MLE easily by means of conventional numerical methods, and $\tilde{\theta}$ is unbiased $O(n^{-2})$.

The analytic bias-reduction techniques of Cox and Snell (1968) and Cordeiro and Klein (1994) have been applied successfully to a range of statistical problems. For instance, Cordeiro and McCullagh (1991) apply it to bias-correct the parameters of generalized linear models. Cordeiro and Klein (1994) use the analytic methods to investigate the bias of the MLE in ARMA models. Cordeiro *et al.* (1996) derive closed-form bias-corrected MLE estimators of the parameters of the beta distribution. Similarly, Cribari-Neto and Vasconcellos (2002) derive bias-corrected expressions for the MLE of the parameters of a restricted beta distribution. Recently, this analytic bias-reduction strategy has been used for the half-logistic distribution (Giles, 2012). In the above mentioned cases, the researchers find by means of Monte Carlo simulations that the analytic methods compare favourably to the bootstrap bias-reduced MLEs in terms of both bias reduction and mean-squared error (MSE).

4. Bias-reduced MLE For the Zero-Inflated Poisson Distribution

This paper considers bias-reduction for the MLE for the parameters of the zero-inflated Poisson distribution. The benchmark model for this paper is inspired by Lambert (1992), though the author cites the influence of work by Cohen (1963) and other authors. Assuming independent sampling, the generic log-likelihood function for this problem is given as:

$$l(\lambda,\gamma) = \sum_{i=1}^{n} I_0 log\left[exp(\gamma) + exp\left(-\lambda\right)\right]$$

$$+\sum_{i=1}^{n}(1-I_0)\left[y_i log(\lambda) - \lambda - log(y_i!)\right] - \sum_{i=1}^{n} log(1+exp(\gamma)),$$

where the indicator function, I_0, takes the value unity when $y_i = 0$, and zero otherwise. The probability of an observation being in R_I, ω, is modelled with a logit specification, $\omega = exp(\gamma)/(1+exp(\gamma))$, and we focus on a zero inflated Poisson distribution without covariates in order to focus on the core problem.

Although we are concerned mainly with bias reduction for the estimation of the "core" parameters of the model, γ and ω, we may also consider bias reduction for the estimation of ω, the probability of an observation in R_I. The log-likelihood for this problem is given by:

$$l(\lambda,\omega) = \sum_{i=1}^{n} I_0 log\left[\omega + (1-\omega)exp(-\lambda)\right]$$
$$+\sum_{i=1}^{n}(1-I_0)\left[y_i log(\lambda) - \lambda + log(1-\omega) - log(y_i!)\right].$$

It should be noted that although $\hat{\omega}$ can be derived directly for $\hat{\gamma}$, by invariance, the non-linear relationship between ω and γ precludes a similar simple manipulation of the bias. To conserve space, the analytic expressions for bias-reduced MLE involving λ and ω are given in the Appendix.

To proceed, we require the derivatives of the log-likelihood function up to the third order. Define $e^{\gamma} + e^{-\lambda} \equiv \alpha$ and $1 + e^{\gamma} \equiv \beta$. Then:

$$\frac{\partial l}{\partial \lambda} = \sum_{i=1}^{n}\left\{-\frac{I_o e^{-\lambda}}{\alpha} + (1-I_o)\frac{y_i - \lambda}{\lambda}\right\}$$
$$\frac{\partial l}{\partial \gamma} = \sum_{i=1}^{n}\left\{\frac{I_o e^{\gamma}}{\alpha} - \frac{e^{\gamma}}{\beta}\right\}$$
$$\frac{\partial^2 l}{\partial \lambda^2} = \sum_{i=1}^{n}\left\{-\frac{I_o e^{-2\lambda}}{\alpha^2} + \frac{I_0 e^{-\lambda}}{\alpha} - (1-I_o)\frac{y_i}{\lambda^2}\right\}$$
$$\frac{\partial^2 l}{\partial \lambda \partial \gamma} = \sum_{i=1}^{n}\left\{\frac{I_0 e^{\gamma-\lambda}}{\alpha^2}\right\}$$
$$\frac{\partial^2 l}{\partial \gamma^2} = \sum_{i=1}^{n}\left\{\frac{I_0 e^{\gamma}}{\alpha} - \frac{e^{\gamma}}{\beta} - \frac{I_0 e^{2\gamma}}{\alpha^2} + \frac{e^{2\gamma}}{\beta^2}\right\}$$
$$\frac{\partial^3 l}{\partial \lambda^3} = \sum_{i=1}^{n}\left\{-\frac{2I_0 e^{-3\lambda}}{\alpha^3} + \frac{3I_0 e^{-2\lambda}}{\alpha^2} - \frac{I_0 e^{-\lambda}}{\alpha} + 2(1-I_o)\frac{y_i}{\lambda^3}\right\}$$
$$\frac{\partial^3 l}{\partial \lambda^2 \partial \gamma} = \sum_{i=1}^{n}\left\{\frac{2I_0 e^{\gamma-2\lambda}}{\alpha^3} - \frac{I_0 e^{\gamma-\lambda}}{\alpha^2}\right\}$$
$$\frac{\partial^3 l}{\partial \gamma^2 \partial \lambda} = \sum_{i=1}^{n}\left\{\frac{I_0 e^{\gamma-\lambda}}{\alpha^2} - \frac{2I_0 e^{2\gamma-\lambda}}{\alpha^3}\right\}$$
$$\frac{\partial^3 l}{\partial \gamma^3} = \sum_{i=1}^{n}\left\{\frac{2I_0 e^{3\gamma}}{\alpha^3} - \frac{3I_0 e^{2\gamma}}{\alpha^2} + \frac{I_0 e^{\gamma}}{\alpha} - \frac{2e^{3\gamma}}{\beta^3} + \frac{3e^{2\gamma}}{\beta^2} - \frac{e^{\gamma}}{\beta}\right\}.$$

We now determine the joint cumulants of the log likelihood function, first noting that $E(y_i) = \frac{\lambda}{1+e^{\gamma}}$ since $E(y_i) = (1-\omega_i)\lambda_i$ and $\omega_i = \frac{e^{\gamma}}{1+e^{\gamma}} \equiv \frac{e^{\gamma}}{\beta}$. Note further that, $E(I_0) = \omega + (1-\omega)e^{\lambda} = \frac{e^{\gamma}+e^{-\lambda}}{1+e^{\gamma}} \equiv \frac{\alpha}{\beta}$.

$$k_{11} = n\left(-\frac{e^{-2\lambda}}{\beta\alpha} + \frac{e^{-\lambda}}{\beta} - \frac{1}{\beta\lambda}\right)$$
$$k_{12} = \frac{ne^{\gamma-\lambda}}{\beta\alpha} = k_{21}$$
$$k_{22} = n\left(-\frac{e^{2\gamma}}{\beta\alpha} + \frac{e^{2\gamma}}{\beta^2}\right)$$
$$k_{111} = n\left(-\frac{2e^{-3\lambda}}{\beta\alpha^2} + \frac{3e^{-2\lambda}}{\beta\alpha} - \frac{e^{-\lambda}}{\beta} + \frac{2}{\lambda^2\beta}\right)$$
$$k_{112} = n\left(\frac{2e^{\gamma-2\lambda}}{\beta\alpha^2} - \frac{e^{\gamma-\lambda}}{\beta\alpha}\right) = k_{121} = k_{211}$$
$$k_{122} = n\left(\frac{e^{\gamma-\lambda}}{\beta\alpha} - \frac{2e^{2\gamma-\lambda}}{\beta\alpha^2}\right) = k_{212} = k_{221}$$
$$k_{222} = n\left(-\frac{3e^{2\gamma}}{\beta\alpha} + \frac{3e^{3\gamma}}{\beta^2} + \frac{2e^{3\gamma}}{\beta\alpha^2} - \frac{2e^{3\gamma}}{\beta^3}\right)$$

$$k_{11}^{(1)} = n\left(-\frac{e^{-3\lambda}}{\beta\alpha^2} + \frac{2e^{-2\lambda}}{\beta\alpha} - \frac{e^{-\lambda}}{\beta} + \frac{1}{\beta\lambda^2}\right)$$
$$k_{11}^{(2)} = n\left(\frac{e^{\gamma-2\lambda}}{\beta^2\alpha} + \frac{e^{\gamma-2\lambda}}{\beta\alpha^2} - \frac{e^{\gamma-\lambda}}{\beta^2} + \frac{e^{\gamma}}{\beta^2\lambda}\right)$$
$$k_{12}^{(1)} = n\left(\frac{e^{\gamma-2\lambda}}{\beta\alpha^2} - \frac{e^{\gamma-\lambda}}{\beta\alpha}\right)$$
$$k_{12}^{(2)} = n\left(\frac{e^{\gamma-\lambda}}{\beta\alpha} - \frac{e^{2\gamma-\lambda}}{\beta^2\alpha} - \frac{e^{2\gamma-\lambda}}{\beta\alpha^2}\right)$$
$$k_{22}^{(1)} = -\frac{ne^{2\gamma-\lambda}}{\beta\alpha^2}$$
$$k_{22}^{(2)} = n\left(\frac{e^{3\gamma}}{\beta\alpha^2} + \frac{e^{3\gamma}}{\beta^2\alpha} - \frac{2e^{2\gamma}}{\beta\alpha} - \frac{2e^{3\gamma}}{\beta^3} + \frac{2e^{2\gamma}}{\beta^2}\right)$$

$$a_{11}^{(1)} = k_{11}^{(1)} - \frac{1}{2}k_{111} = -\frac{ne^{\gamma-\lambda}}{2\beta\alpha} = a_{12}^{(1)} = a_{21}^{(1)} = a_{22}^{(1)}$$
$$a_{11}^{(2)} = k_{11}^{(2)} - \frac{1}{2}k_{112} = \frac{n(\lambda e^{\gamma-\lambda} - \lambda e^{2\gamma-\lambda} + 2e^{\gamma-\lambda} + 2e^{2\gamma})}{2\beta^2\alpha\lambda}$$
$$a_{12}^{(2)} = k_{21}^{(2)} - \frac{1}{2}k_{122} = \frac{n(e^{\gamma-\lambda} - e^{2\gamma-\lambda})}{2\beta^2\alpha} = a_{21}^{(2)}$$
$$a_{22}^{(2)} = k_{22}^{(2)} - \frac{1}{2}k_{222} = \frac{n(e^{3\gamma} - e^{2\gamma} + e^{2\gamma-\lambda} - e^{3\gamma-\lambda})}{2\beta^3\alpha}$$

$$A^{(1)} = -\frac{n}{2}\begin{bmatrix} \frac{e^{\gamma-\lambda}}{\beta\alpha} & \frac{e^{\gamma-\lambda}}{\beta\alpha} \\ \frac{e^{\gamma-\lambda}}{\beta\alpha} & \frac{e^{\gamma-\lambda}}{\beta\alpha} \end{bmatrix}$$
$$A^{(2)} = \frac{n}{2}\begin{bmatrix} \frac{(\lambda e^{\gamma-\lambda} - \lambda e^{2\gamma-\lambda} + 2e^{\gamma-\lambda} + 2e^{2\gamma})}{\beta^2\alpha\lambda} & \frac{(e^{\gamma-\lambda} - e^{2\gamma-\lambda})}{\beta^2\alpha} \\ \frac{(e^{\gamma-\lambda} - e^{2\gamma-\lambda})}{\beta^2\alpha} & \frac{(e^{3\gamma} - e^{2\gamma} + e^{2\gamma-\lambda} - e^{3\gamma-\lambda})}{\beta^3\alpha} \end{bmatrix}.$$

Finally, we are able to write the bias using the result of Cordeiro and Klein (1994):

$$Bias\begin{pmatrix}\hat{\lambda}\\ \hat{\gamma}\end{pmatrix} = K^{-1}AVec(K^{-1}) + O(n^{-2}).$$

The bias-adjusted estimators are given as :

$$\begin{pmatrix}\tilde{\lambda}\\ \tilde{\gamma}\end{pmatrix} = \begin{pmatrix}\hat{\lambda}\\ \hat{\gamma}\end{pmatrix} - \hat{K}^{-1}\hat{A}Vec(\hat{K}^{-1}),$$

where $\hat{K} = K|_{\hat{\theta}}$ and $\hat{A} = A|_{\hat{\theta}}$. Numerical evaluations of these estimators are reported in the following section. It should be noted that it is quite possible for the analytic bias adjustment to result in a negative value for $\tilde{\lambda}$. Clearly, such an estimate would be inappropriate. This was *not* an issue in the simulation experiment, as is discussed further below.

5. SIMULATION EXPERIMENT

The following Monte Carlo experiment compares the performance of the analytic bias-reduced MLEs to that of the parametric bootstrap bias-adjusted estimator. The actual bias and mean squared error (MSE) of the original MLEs themselves have also been simulated in the Monte Carlo experiment. For a generic parameter, θ , and its estimator, θ^*, we report the percent bias, $\frac{100}{|\theta|}\left[\left(\frac{1}{N}\sum_{i=1}^{N}\theta_i^*\right) - \theta\right]$, and the percent MSE, $\frac{100}{\theta^2}\left[\frac{1}{N}\sum_{i=1}^{N}(\theta_i^* - \theta)^2\right]$, based on N Monte Carlo replications.

Simulations were conducted using version 2.11.1 of the statistical software environment R (2008). The zero-inflated Poisson random variables were generated using the VGAM package (Yee, 2010) and the log likelihood function was maximized using the Nelder-Mead algorithm in the MaxLik package (Toomet, 2008).

Each experiment comprises 100,000 Monte Carlo replications. In addition to obtaining the biases and MSEs for the MLE and the analytic bias-adjusted MLE, we obtain another bias-adjusted MLE using the bootstrap resampling method (Efron, 1979). The parametric bootstrap-bias-adjusted estimator is given as $\breve{\theta} = 2\hat{\theta} - \frac{1}{N_B}\sum_{j=1}^{N_B}\hat{\theta}_j$, where $\hat{\theta}$ is the MLE of θ from the j^{th} of the $N_B = 999$ bootstrap samples. We did not encounter any negative values for $\widetilde{\lambda}$ for the experimental design that we considered, so the potential problem alluded to at the end of section 4 did not arise. Some additional investigation showed that this does not become an issue unless sample sizes as small as $n = 10$ are considered in conjunction with the parameter values assigned in Tables 1 to 3 below.

In addition to MSE, we also report a statistic suggested by Pitman (1939), which is generated within our Monte Carlo experiment. This measure serves as a further means by which the performance of the analytic bias-reduced estimator can be compared with the parametric bootstrap estimator. Specifically, Pitman's nearness

(PN) measure is the probability that the analytic bias-adjusted estimator is closer to the value of the true parameter than is the bootstrap bias-adjusted estimator:

$$PN = Pr(|\tilde{\theta} - \theta| < |\breve{\theta} - \theta|).$$

This probability is consistently estimated by comparing the magnitude of the absolute difference between the bias-adjusted estimators and the true parameter values in every repetition of the Monte Carlo simulation. The estimator $\tilde{\theta}$ is preferred to the estimator $\breve{\theta}$ if $PN \geqslant 0.5$, for all values of θ, and $PN > 0.5$ for at least one value of θ.

Tables 1 to 3 illustrate the results of the Monte Carlo simulations for three different pairs of parameter values, and for various sample sizes. In each case, we report the percentage biases of the various estimators, together with the associated percentage MSEs (in square brackets). In addition, the Pitman nearness measure defined above is also reported in each table, to provide a different basis for comparing the estimators. These tables enable us to compare the relative merits of the analytic bias adjustment and the bootstrap bias adjustment in various ways. We see that the original, unadjusted, maximum likelihood estimators of λ and γ each exhibit rather small percentage bias, especially the MLE of λ. This suggests that the MLE for a ZIP model without covariates may be considered very reliable, in practice. Both the bootstrap and analytic approach are successful in reducing the relative bias and also lowering the relative MSE of the MLE for both parameters in the cases that we have considered. In keeping with othe results in the recent literature, we find that the analytic bias-reduced estimators perform favourably relative to the bootstrap estimators.

TABLE 1. RESULTS OF MONTE CARLO SIMULATIONS*. $\lambda = 3$, $\omega = 0.3$.

n	$\hat{\lambda}$	$\tilde{\lambda}$	$\breve{\lambda}$	PN	$\hat{\gamma}$	$\tilde{\gamma}$	$\breve{\gamma}$	PN
50	-0.1443	-0.0043	-0.0028	0.5082	4.3937	0.3840	2.1108	0.4298
	[1.1445]	[1.1410]	[1.1426]		[18.1438]	[16.8406]	[14.0278]	
60	-0.1980	-0.0815	-0.0799	0.5056	-3.5663	0.2455	1.0216	0.4525
	[0.9588]	[0.9562]	[0.9573]		[14.8074]	[13.3788]	[12.5959]	
75	-0.0512	0.0418	0.0426	0.5044	-2.7648	0.1754	0.4800	0.4767
	[0.7599]	[0.7584]	[0.7591]		[11.6132]	[10.7766]	[10.5612]	
100	-0.0742	-0.0046	-0.0044	0.5045	-2.1073	0.3568	0.1598	0.4897
	[0.5674]	[0.5666]	[0.5673]		[8.5405]	[8.8096]	[8.0518]	
125	-0.0615	-0.0059	-0.0054	0.5047	-1.6345	0.0497	0.1258	0.4959
	[0.4500]	[0.4494]	[0.4500]		[6.7238]	[6.4516]	[6.4382]	
150	0.0482	-0.0019	-0.0020	0.5046	-1.4100	-0.0202	0.0265	0.4996
	[0.3776]	[0.3772]	[0.3776]		[5.5465]	[5.3942]	[5.3900]	
200	-0.0337	0.0010	0.0012	0.5062	-1.0927	-0.0643	-0.0390	0.5013
	[0.2837]	[0.2835]	[0.2838]		[4.1380]	[4.0333]	[4.0333]	

**Percent bias of estimator, with percent MSE of estimator beneath in square brackets.*

The MLEs for γ are more biased than those for λ. For the former parameter,$\tilde{\gamma}$ exhibits a smaller bias than $\breve{\gamma}$, *albeit* at the cost of somewhat higher MSE. This variance-bias trade-off is especially evident in Table 1 and is also reflected in the superior PN for $\breve{\gamma}$. In Table 2, $\tilde{\gamma}$ performs especially well relative to $\breve{\gamma}$ in terms of bias reduction, MSE performance, and Pitman's nearness measure.

Table 2. Results of Monte Carlo Simulations*. $\lambda = 3$, $\omega = 0.7$.

n	$\hat{\lambda}$	$\tilde{\lambda}$	$\breve{\lambda}$	PN	$\hat{\gamma}$	$\tilde{\gamma}$	$\breve{\gamma}$	PN
50	-0.3889	-0.0476	-0.0152	0.5038	2.0631	-0.1523	-0.3083	0.5076
	[2.8205]	[2.7991]	[2.8022]		[15.5652]	[14.4796]	[14.8522]	
60	-0.3311	-0.0491	-0.0301	0.5018	1.0226	0.0494	0.1240	0.5008
	[2.3243]	[2.3096]	[2.3109]		[12.7806]	[12.0785]	[12.1637]	
75	-0.1767	0.0467	0.0578	0.5038	0.9620	0.1713	0.1856	0.50014
	[1.8337]	[1.8249]	[1.8273]		[10.0123]	[9.6346]	[9.6400]	
100	-0.1712	-0.0051	0.0003	0.5037	0.7462	0.1500	0.1520	0.5017
	[1.3558]	[1.3508]	[1.3520]		[7.4963]	[7.2859]	[7.2874]	
125	-0.0615	-0.0059	-0.0054	0.5047	-1.6345	0.0497	0.1258	0.4959
	[0.4500]	[0.4494]	[0.4500]		[6.7238]	[6.4516]	[6.4382]	
150	-0.0992	0.0105	0.0136	0.5067	0.4615	0.0625	0.0683	0.5044
	[0.8973]	[0.8952]	[0.8963]		[4.9448]	[4.8528]	[4.8576]	
200	-0.1122	-0.0303	-0.0290	0.5040	0.2769	-0.0222	-0.0228	0.50403
	[0.6695]	[0.6683]	[0.6690]		[3.7004]	[3.6494]	[3.6525]	

**Percent bias of estimator, with percent MSE of estimator beneath in square brackets.*

Table 3. Results of Monte Carlo Simulations*. $\lambda = 5$, $\omega = 0.7$.

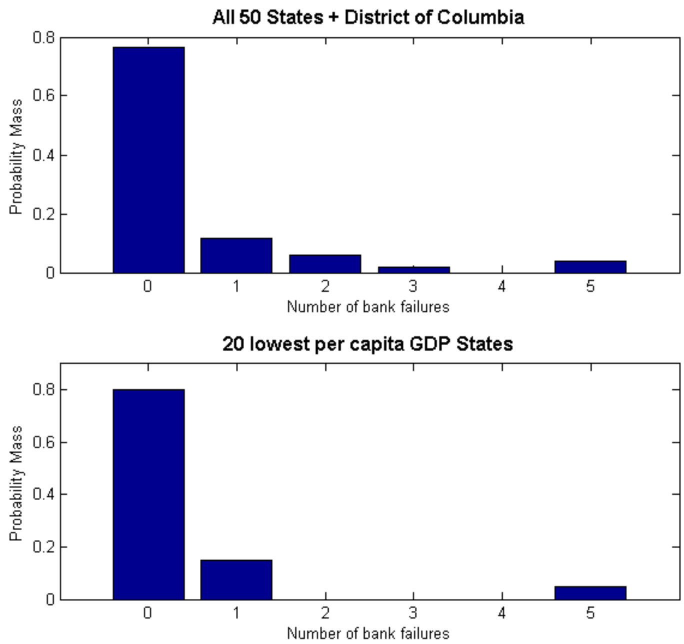
n	$\hat{\lambda}$	$\tilde{\lambda}$	$\breve{\lambda}$	PN	$\hat{\gamma}$	$\tilde{\gamma}$	$\breve{\gamma}$	PN
50	-0.1018	-0.0200	-0.0103	0.5001	2.0631	-0.1523	-0.3083	0.4799
	[1.4607]	[1.4540]	[1.4554]		[14.2616]	[13.3856]	[13.3195]	
60	-0.0862	-0.0191	-0.0144	0.4993	1.8800	0.0525	0.0508	0.4848
	[1.2076]	[1.2030]	[1.2039]		[11.7561]	[11.1552]	[11.1190]	
75	-0.0322	0.0205	0.0232	0.5042	1.5611	0.1157	0.0499	0.4946
	[0.9557]	[0.9528]	[0.9540]		[9.3067]	[8.9291]	[8.9126]	
100	-0.0230	0.0159	0.0174	0.5064	1.2292	0.1573	0.1212	0.4938
	[0.7043]	[0.7028]	[0.7033]		[6.9799]	[6.7686]	[6.7653]	
125	$-8.4783 \cdot 10^{-5}$	0.0307	0.03184	0.5042	0.9592	0.1084	0.0868	0.5126
	[0.5647]	[0.5637]	[0.5642]		[5.5387]	[5.4058]	[5.4078]	
150	-0.0478	-0.0223	-0.0216	0.5003	0.6655	-0.0386	-0.0544	0.4987
	[0.4728]	[0.4721]	[0.4724]		[4.5734]	[4.4842]	[4.4878]	
200	-0.0499	-0.0309	-0.0303	0.5077	0.4768	-0.0482	-0.0570	0.4989
	[0.3494]	[0.3491]	[0.3494]		[3.4201]	[3.3706]	[3.3730]	

**Percent bias of estimator, with percent MSE of estimator beneath in square brackets.*

6. Illustrative Application

The data for this application come from a list of bank failures from the United States Federal Deposit Insurance Corporation. The data comprise the number of banks that failed in each of the fifty states plus the District of Columbia in the year 2008 ($n = 51$), out of a total of 25 bank failures in that year. Figure 1 shows empirical mass functions for the full sample and for the 20 states with the lowest per capita GDP ($n = 20$). The sample means and variances are $\bar{x} = 0.5$ (0.4) and $var(x) = 1.32$ (1.31) for $n = 51$ ($n = 20$) respectively, typifying the overdispersion that is readily modelled with a ZIP distribution. We have fitted ZIP models to each sample, using MLE, and then applied the analytic bias adjustment to the point estimates. The results appear in Table 4. These are presented for two parameterizations of the model - once in terms of the paramters λ and γ, and once in terms of the parameters λ and ω. Ninety-five percent confidence intervals based on 999 bootstrap replications are reported in square brackets beneath the coefficient estimates in that table.

Figure 1. Bank Failure Data



Although the changes arising from bias correction are not dramatic in Table 4, the analytic corrections nevertheless improve the quality of our inferences for this example. It should also be noted that the bootstrap proved unstable in the $n = 20$ case. Bootstrap samples with too many zeros resulted in large outliers, requiring us to replace those problematic samples for the purposes of constructing sensible confidence intervals. This of course, also has adverse implications for the use of the bootstrap for bias correction itself, for this model with very small samples.

Table 4. Estimated ZIP Parameters for Bank Failure Data.

$n = 51$				
	$\hat{\lambda}$	$\tilde{\lambda}$	$\hat{\gamma}$	$\tilde{\gamma}$
(λ, γ):	1.7047	1.7202	0.9071	0.9236
	[0.8139, 2.6773]	[0.8040, 2.7105]	[0.0844, 1.7171]	[0.2019, 1.6896]
	$\hat{\lambda}$	$\tilde{\lambda}$	$\hat{\omega}$	$\tilde{\omega}$
(λ, ω):	1.7047	1.7243	0.7124	0.7241
	[0.8141, 2.6773]	[0.5575, 2.7093]	[0.5211, 0.8478]	[0.5036, 0.8557]
$n = 20$				
	$\hat{\lambda}$	$\tilde{\lambda}$	$\hat{\gamma}$	$\tilde{\gamma}$
(λ, γ):	1.5935	1.6402	1.0931	1.1423
	[0.9840, 3.3812]	[1.1560, 3.2957]	[-0.4966, 1.6852]	[0.15268, 1.5819]
	$\hat{\lambda}$	$\tilde{\lambda}$	$\hat{\omega}$	$\tilde{\omega}$
(λ, ω):	1.5937	1.6583	0.7490	0.7856
	[0.9842, 3.3807]	[0.7997, 3.6798]	[0.3784, 0.8436]	[0.3726, 0.8485]

7. Discussion and Conclusions

This paper shows that analytic-bias corrections to $O(n^{-1})$ compare favourably to parametric bootstrap corrections to the same order for the MLE for the zero-inflated Poisson model, though our results also show that the MLE is very reliable in small samples to begin with.

One benefit of the analytic approach is that the corrections can in principle be extended to higher orders approximations. This means that with some effort, the analytic approach may unambiguously dominate the standard parametric bootstrap, especially in situations where the bootstrap is reliable only to $O(n^{-1})$.

The analytic bias-reduction methods may also provide a useful alternative in areas in which the bootstrap fails. For example, Cordeiro and Klein (1994) show that it is still applicable if the data are non-*i.i.d.*, while MacKinnon (2002) argues that the bootstrap may perform poorly in this context. Horowitz (2000) also cautions against the use of bootstrap methods for dependent data and statistics that are not asymptotically pivotal. As noted in the last section, the bootstrap can also fail for other reasons when applied to the ZIP model with small samples.

We have compared a parametric bootstrap estimator to an analytic bias correction, but other resampling methods may also be worth exploring, such as the double bootstrap and its less computationally intensive cousin, the fast double bootstrap (FDB). The latter is described in Davidson and MacKinnon (2007), who argue that the FDB may be useful in situations where the ordinary bootstrap performs well but there exists room for further improvement.

Although we have found the MLEs for the parameters of the zero-inflated Poisson model to be relatively unbiased, we have demonstrated that analytic methods can be used to bias-correct these MLEs very successfully. Moreover, this bias reduction does not come at the expense of increased MSE. Indeed, it can often result in a lower MSE for the bias-reduced MLE. In the cases examined, the analytic bias-corrected estimators are superior to the bootstrap-corrected estimators in terms of Pitman's nearness measure for the less-biased estimators, but they are not preferred to the bootstrap on this basis when the estimators are more biased, on account of a bias-variance tradeoff.

Our illustrative example demonstrates the ease and utility of applying these methods to empirical data. These procedures may yield more dramatic improvements if applied to other datasets that are also believed to follow a zero-inflated Poisson process, though it should be noted again that based on the findings of this paper, the MLE for the parameters of the ZIP model appears to be very reliable, even in the context of relatively small samples.

Acknowledgement. We are grateful to Judith Clarke, Ryan Godwin, an Associate Editor, and a referee for helpful comments on earlier versions of this work.

Appendix

Analytic bias reduction for the ZIP model with parameters λ and ω. The log-likelihood function for the model is given by:

$$l(\lambda,\omega)=\sum_{i=1}^{n} I_0 log\left[\omega+(1-\omega)exp(-\lambda)\right]$$
$$+\sum_{i=1}^{n}(1-I_0)\left[y_i log(\lambda)-\lambda+log(1-\omega)-log(y_i!)\right]$$

We require derivatives of the log-likelihood function up to the third order. Allowing $\omega+(1-\omega)exp(-\lambda)\equiv\alpha$,$(1-\omega)exp(-\lambda)\equiv\beta$,$(1-exp(-\lambda))\equiv\delta$,

$$\frac{\partial l}{\partial\lambda}=\sum_{i=1}^{n}\left\{-\frac{I_0\beta}{\alpha}+(1-I_0)\frac{y_i-\lambda}{\lambda}\right\}$$
$$\frac{\partial l}{\partial\omega}=\sum_{i=1}^{n}\left\{\frac{I_0\delta}{\alpha}-\frac{(1-I_0)}{(1-\omega)}\right\}$$
$$\frac{\partial^2 l}{\partial\lambda^2}=\sum_{i=1}^{n}\left\{\frac{I_0\beta}{\alpha}-\frac{I_0\beta^2}{\alpha^2}-(1-I_o)\frac{y_i}{\lambda^2}\right\}$$
$$\frac{\partial^2 l}{\partial\lambda\partial\omega}=\sum_{i=1}^{n}\left\{\frac{I_0\beta}{\alpha(1-\omega)}+\frac{I_0\beta\delta}{\alpha^2}\right\}$$
$$\frac{\partial^2 l}{\partial\omega^2}=\sum_{i=1}^{n}\left\{-\frac{I_0\delta^2}{\alpha^2}-\frac{(1-I_0)}{(1-\omega)^2}\right\}$$
$$\frac{\partial^3 l}{\partial\lambda^3}=\sum_{i=1}^{n}\left\{-\frac{2I_0\beta^3}{\alpha^3}+\frac{3I_0\beta^2}{\alpha^2}-\frac{I_0\beta}{\alpha}+2(1-I_o)\frac{y_i}{\lambda^3}\right\}$$
$$\frac{\partial^3 l}{\partial\lambda^2\partial\omega}=\sum_{i=1}^{n}\left\{-\frac{I_0\beta}{\alpha(1-\omega)}-\frac{I_0\delta\beta}{\alpha^2}+\frac{2I_0\beta^2}{\alpha^2(1-\omega)}+\frac{2I_0\delta\beta^2}{\alpha^3}\right\}$$
$$\frac{\partial^3 l}{\partial\lambda\partial\omega^2}=\sum_{i=1}^{n}\left\{-\frac{2I_0\delta\beta}{\alpha^2(1-\omega)}-\frac{2I_0\delta^2\beta}{\alpha^3}\right\}$$
$$\frac{\partial^3 l}{\partial\omega^3}=\sum_{i=1}^{n}\left\{\frac{2I_0\delta^3}{\alpha^3}-\frac{2(1-I_0)}{(1-\omega)^3}\right\}$$

Using the fact that $E(y) = (1-\omega)\lambda$, and $E(I_0) = \omega + (1-\omega)e^{-\lambda} \equiv \alpha$.

$$
\begin{aligned}
k_{11} &= n\left(\beta - \frac{\beta^2}{\alpha} - \frac{(1-\omega)}{\lambda}\right) \\
k_{12} &= n\left(\frac{\beta}{(1-\omega)} + \frac{\beta\delta}{\alpha}\right) = k_{21} \\
k_{22} &= n\left(-\frac{\delta^2}{\alpha} + \frac{(\alpha-1)}{(1-\omega)^2}\right) \\
k_{111} &= n\left(-\frac{2\beta^3}{\alpha^2} + \frac{3\beta^2}{\alpha} - \beta + \frac{2(1-\omega)}{\lambda^2}\right) \\
k_{112} &= n\left(-\frac{\beta}{(1-\omega)} - \frac{\delta\beta}{\alpha} + \frac{2\beta^2}{\alpha(1-\omega)} + \frac{2\delta\beta^2}{\alpha^2}\right) = k_{121} = k_{211} \\
k_{122} &= n\left(-\frac{2\delta^2\beta}{\alpha^2} - \frac{2\delta\beta}{\alpha(1-\omega)}\right) = k_{212} = k_{221} \\
k_{222} &= n\left(\frac{2\delta^3}{\alpha^2} + \frac{2(\alpha-1)}{(1-\omega)^3}\right) \\
k_{11}^{(1)} &= n\left(-\beta + \frac{2\beta^2}{\alpha} - \frac{\beta^3}{\alpha^2} + \frac{(1-\omega)}{\lambda^2}\right) \\
k_{11}^{(2)} &= n\left(\frac{2\beta^2}{\alpha(1-\omega)} - \frac{\beta}{(1-\omega)} + \frac{\beta^2\delta}{\alpha^2} + \frac{1}{\lambda}\right) \\
k_{12}^{(1)} &= n\left(\frac{\beta^2}{\alpha(1-\omega)} - \frac{\beta}{(1-\omega)} + \frac{\delta\beta^2}{\alpha^2} - \frac{\delta\beta}{\alpha}\right) \\
k_{12}^{(2)} &= n\left(-\frac{\beta\delta}{\alpha(1-\omega)} - \frac{\delta^2\beta}{\alpha^2}\right) \\
k_{22}^{(1)} &= n\left(-\frac{2\delta\beta}{\alpha(1-\omega)} - \frac{\delta^2\beta}{\alpha^2} - \frac{\beta}{(1-\omega)^2}\right) \\
k_{22}^{(2)} &= n\left(\frac{\delta^3}{\alpha^2} - \frac{2(1-\alpha)}{(1-\omega)^3} + \frac{\delta}{(1-\omega)^2}\right) \\
a_{11}^{(1)} &= \frac{n}{2}\left(\frac{\beta^2}{\alpha} - \beta\right) \\
a_{12}^{(1)} &= \frac{n}{2}\left(-\frac{\beta}{(1-\omega)} - \frac{\delta\beta}{\alpha}\right) = a_{21}^{(1)} \\
a_{22}^{(1)} &= n\left(-\frac{\delta\beta}{\alpha(1-\omega)} - \frac{\beta}{(1-\omega)^2}\right) \\
a_{11}^{(2)} &= n\left(\frac{\delta\beta}{2\alpha} - \frac{\beta}{2(1-\omega)} + \frac{\beta}{\alpha(1-\omega)} + \frac{1}{\lambda}\right) \\
a_{12}^{(2)} &= a_{21}^{(2)} = 0 \\
a_{22}^{(2)} &= -\frac{2n\beta}{(1-\omega)^3}
\end{aligned}
$$

$$A^{(1)} = n \left[\begin{array}{cc} \frac{1}{2}\left(\frac{\beta^2}{\alpha} - \beta\right) & \frac{1}{2}\left(-\frac{\beta}{(1-\omega)} - \frac{\delta\beta}{\alpha}\right) \\ \frac{1}{2}\left(-\frac{\beta}{(1-\omega)} - \frac{\delta\beta}{\alpha}\right) & \left(-\frac{\delta\beta}{\alpha(1-\omega)} - \frac{\beta}{(1-\omega)^2}\right) \end{array} \right]$$

$$A^{(2)} = n \left[\begin{array}{cc} \left(\frac{\delta\beta}{2\alpha} - \frac{\beta}{2(1-\omega)} + \frac{\beta}{\alpha(1-\omega)} + \frac{1}{\lambda}\right) & 0 \\ 0 & -\frac{2\beta}{(1-\omega)^3} \end{array} \right]$$

Now we can write:

$$Bias \left(\begin{array}{c} \hat{\lambda} \\ \hat{\omega} \end{array} \right) = K^{-1} AVec(K^{-1}) + O(n^{-2}).$$

The bias-adjusted estimators are given as :

$$\left(\begin{array}{c} \tilde{\lambda} \\ \tilde{\omega} \end{array} \right) = \left(\begin{array}{c} \hat{\lambda} \\ \hat{\omega} \end{array} \right) - \hat{K}^{-1} \hat{A} Vec(\hat{K}^{-1}).$$

References

Bartlett, M. S. (1953a). Approximate confidence intervals. Biometrika, 40, 12-19

Bartlett, M. S. (1953b). Approximate confidence intervals. II. More than one unknown parameter. Biometrika, 40, 306-317.

Böhning, D., Dietz, Schlattmann, P., Mendonca, L., and Kirchner, U. (1999). The zero-inflated Poisson model and the decayed, missing and filled teeth index in dental epidemiology. Journal of the Royal Statistical Society, 162, 195-205.

Cohen, A.C. (1963). Estimation in mixtures of discrete distributions. From Proceedings of the International Symposium on Discrete Distributions. Montreal, 373-378.

Cordeiro, G. M., and Klein, R. (1994). Bias correction in ARMA models. Statistics and Probability Letters, 19, 169-176.

Cordeiro, G. M., and McCullagh, P. (1991). Bias correction in generalized linear models. Journal of the Royal Statistical Society, Series B, 53, 629-643.

Cox, D. R., and Snell, E. J. (1968). A general definition of residuals. Journal of the Royal Statistical Society Series B., 30, 248-275.

Cragg, J.G. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica. 39, 829-844.
Cramér, H (1946). Mathematical Methods of Statistics. Princeton, NJ. Princeton University Press.

Cribari-Neto, F., and Vasconcellos, K. L. P. (2002). Nearly unbiased maximum likelihood estimation for the beta distribution. Journal of Statistical Computation and Simulation, 72, 107-118.

Davidson, R. MacKinnon, J. G. (2007). Improving the reliability of bootstrap tests with the fast double bootstrap. Computational Statistics & Data Analysis, 51, 3259-3281.

Dugué, D. (1937). Applications des propriétés de la limite au sens du calcul des probabilitiés à l'étude de diverses questions d'estimation. J. École Polytechnique, 3, 305-374.

Efron, B. (1979). Bootstrap methods: another look at the Jackknife. Annals of Statistics, 7, 1-26.

Federal Deposit Insurance Corporation: Failed Bank List.
http://www.fdic.gov/bank/individual/failed/banklist.html

Giles, D.E.A. (2012). Bias reduction for the maximum likelihood estimators of the parameters in the half-logistic distribution. Communications in Statistics - Theory and Methods, 41, 212-222.

Haldane, J. B. S. (1953). The estimation of two parameters from a sample, Sankhyā. 12, 313- 320.

Haldane, J. B. S. and Smith, S.M. (1956), The sampling distribution of a maximum likelihood estimate. Biometrika, 43, 96-103.

Johnson, N.L., Kemp, A. W. and Kotz, S. (2005). Univariate Discrete Distributions, 3rd. ed.. Wiley, Hoboken, N.J..

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics, 34, 1-14.

Machado, J.A.F., and Santos Silva, J.M.C. (2005). Quantiles for counts. Journal of the American Statistical Association. 100, 1226-1237.

MacKinnon, J.G. (2002). Bootstrap inference in econometrics, Canadian Journal of Economics. 35, 615-645.

Mullahy, J. (1986). Specification and testing of some modified count data models. Journal of Econometrics. 33, 341-365.

Pitman, E. J. G. (1939). The estimation of location and scale parameters of a continuous population of any given form. Biometrika, 30, 391-421.

R (2008). The R Project for Statistical Computing. http://www.r-project.org

Shenton, L. R., Bowman, K. (1963). Higher moments of a maximum-likelihood estimate, Journal of the Royal Statistical Society, B, 25, 305-317.

Shenton, L. R., Bowman, K. (1977). Maximum Likelihood Estimation in Small Samples. Griffin, London.

Toomet, O. and A. Henningsen (2008), maxLik: Maximum likelihood estimation, http://CRAN.R- project.org ; http://maxLik.org

Yee, T. (2010). The VGAM package for R. http://www.stat.auckland.ac.nz/˜yee/VGAM/