J. Ind. Soc. Agril. Statist.
61(1), 2007 : 18-27

A New Sequential Estimator for ZIP Regression Model

A.R.S. Bhatta and K. Aruna Rao[1]
University of Agricultural Sciences, Dharwad
(Received : February, 2005)

SUMMARY

Non-convergence problem is common in zero inflated Poisson distribution. To overcome this problem sequential estimators are proposed for parameters which are free from non-convergence. Simulated result reveals that performance of sequential estimators are better as compared to one-step, two-steps and regular maximum likelihood estimators.

Key words : Zeros Inflated Poisson Regression (ZIP), Maximum likelihood estimator, Sequential estimator, Small sample comparison.

1. INTRODUCTION

Maximum likelihood estimator is widely used in statistical work. The reason for this is the fact that asymptotically, this estimator possesses several optimum properties (Rao 1973, Lehman 1983). Many authors have described the maximum likelihood estimation procedure in a general context, which turn out to be iterative in nature, as closed form solutions do not exist. Ghosh *et al.* (1999) point out some problems with classical statistical inference procedures such as maximum likelihood estimation and the large sample approximation theory for making inferences of parameters in the zero inflated models. They proposed a Bayesian procedure to model data with exceed zero. Lambart (1992) suggested to use EM algorithm to find the maximum likelihood estimate. Li *et al.* (1999) used the Powell's algorithm that is grid search method with the moment estimates as the initial estimates for finding the maximum likelihood estimates. There are other algorithms such as Newton-Raphson to maximize the log-likelihood. Usually Newton-Raphson algorithm is faster than EM algorithm. EM often converges very slowly. Thomas and Gan (1997) put prior distribution on the variance covariance matrix when they analysed educational data in order to avoid the extreme slowness of computation, moreover, as Lambart (1992) pointed out that if θ is a function of λ, EM algorithm fails. Due to this, we have used Newton-Raphson algorithm to maximize the log likelihood, even though it would lead to lots of tedious work involved in the computation of first order and second order derivatives (Searborough 1996). As the non-convergence problem is common in zero inflated models, it motivated us to propose sequential estimators for θ and the regression parameter β. Such estimators have also been proposed by Morel and Koehler (1995) for modeling data with extraneous variation.

2. ZERO INFLATED POISSON (ZIP) DISTRIBUTION

Let $Y_1, \ldots, Y_n$ be n independently distributed Poisson variates with parameters λ_i and θ (the proportion of non zeros in the distribution). Y has a zero-inflated Poisson (ZIP) distribution of the form given by

$$P(Y = y) = \begin{cases} (1-\theta) + \theta e^{-\lambda} & \text{if } y = 0 \\ \theta e^{-\lambda} \lambda^y / y! & \text{if } y > 0 \end{cases} \quad (2.1)$$

In (2.1), we have considered the proportion of zero in the distribution is $(1 - \theta)$ and not the θ as considered by the several researchers. It is possible for θ in (2.1) to assume negative, given a zero deflated distribution. Zero deflated data seldom arise in practice and we assume $0 < \theta < 1$ in the present study. Zero inflated

[1] *Mangalore University, Mangalore*

form of the other count distribution can be defined in the similar way, Gupta *et al.* (1996), for example, investigated the zero inflated form of the generalized Poisson distribution.

For zero inflated Poisson distribution of the form (2.1)

$$\mu_i = E(y_i) = \theta\lambda_I$$

and $$V(y_i) = \mu_I + W\mu_i^2$$

where $$W = \left(\frac{1-\theta}{\theta}\right)$$

$W > \mu_i$ as $0 < \theta < 1$

The results of (2.2) and (2.3) show that the variance of the zero inflated Poisson distribution is greater than the ordinary Poisson distribution with the same mean. Lambart (1992) considered the model in which $\log \lambda = X\beta$ as the link function, where β is vector of regression parameters and $\lambda = (\lambda_1, \ldots \lambda_n)'$ and X is matrix of covariates. Martin Ridout *et al.* (1998) pointed out that a great variety of alternative models can be generated by using different link functions for λ and/or θ. Details of zero inflated negative binomial regression models, which are analogous to ZIP models, are available in Greene (1994).

3. LIKELIHOOD FUNCTION AND NEW SEQUENTIAL ESTIMATOR

Let $Y_1, \ldots, Y_n$ be independently distributed ZIP random variables and $X_1, \ldots, X_n$ being the values taken by a co-variate X. i.e. $Y_i \sim \text{ZIP}(\theta, \lambda_i)$

Then likelihood function

$$l(Y, X, \theta, \lambda) = [1-\theta+\theta g(0,\lambda_i)]^{\Sigma\delta(y_i=0)}\theta^{\Sigma\delta(y_i>0)} \prod_{y_i>0} g(y_i, \lambda_i) \quad (3.1)$$

where

$$g(y_i, \lambda_i) = e^{-\lambda_i}\lambda_i^{y_i}/y_i!$$

$$g(y_i) = \begin{cases} 0 & \text{if } y_i = 0 \\ 1 & \text{if } y_i > 0 \end{cases}$$

and log likelihood function is

$$L(Y, X, \theta, \lambda) = \sum_{i=1}^{n}\delta(y_i)\log\left[(1-\theta)+\theta e^{-\lambda}\right] + \sum_{i=1}^{n}[1-\delta(y_i)]\log\left[\theta e^{-\lambda}\lambda_i^{y_i}/y_i!\right] \quad (3.2)$$

As we know that $E(y_i) = ql_i$, using the canonical link $\log l_i = x_i b$, it follows that

$$\log y_i = \log q + x_i b \quad (3.3)$$

However, when y_i's takes the value 0, above expression (3.3) is not defined, to overcome this we have defined

$$z_i = \log(y_i + 0.5) - \log q \quad (3.4)$$

Using this transformation and if θ is known, the regression parameter β can be estimated using the least square theory and is given by

$$\beta_s = (X'X)^{-1}X'Z \quad (3.5)$$

where

$Z = (\log(y_1+0.5) - \log\theta_s, \ldots, \log(y_n+0.5) - \log\theta_s)$

$X = n \times p$ design matrix and

$\beta = (\beta_1, \ldots \beta_p)'$, the vector of regression parameters

Second step of the sequential procedure involves the estimation of the parameter θ_s. For this, we have used the estimator based on the proportion of zeros.

$$\text{let } U_i = \begin{cases} 0 & \text{if } y_i = 0 \\ 1 & \text{if } y_i > 0 \end{cases} \qquad i = 1, \ldots, n$$

The U_i are independently distributed Bernoulli random variables with

$$E(U_i) = P(y_i = 0) = 1 - \theta + \theta e_i^{-\lambda}; \; i = 1, \ldots, n$$

Then the moment estimator of θ based on U_i is given by

$$\sum_{i=1}^{n} U_i = n(1-\theta) + \theta\sum_{i=1}^{n} e_i^{-\lambda}$$

leading to

$$\hat{\theta} = \frac{1-\overline{U}}{1-\frac{\sum_{i=1}^{n} e^{-\lambda_i}}{n}} \quad (3.6)$$

The sequential estimation starts with an initial value for θ in (3.5) and estimating β, which in turn is used to update the value of θ in (3.6). The iterative procedure is continued until convergence takes place. Our concern is in the regression paramater β in the presence of the nuisance parameter θ. Under suitable regularity conditions, it follows that

$$\sqrt{n(\hat{\beta}_s - \beta)}\ \text{a.d.}\ N(0, \Sigma)$$

in the case of maximum likelihood estimator for small samples, the estimator may not be unbiased, although the bias may be negligible.

where

$$\Sigma = \lim \frac{1}{n}(X'X)^{-1} X'V(U)X(X'X)^{-1}$$

$$V(U) = \text{diag.}\ [V(U_1), \ldots, V(U_n)]$$

and

$$V(U_i) = \frac{\theta\lambda_i + \left(\frac{1-\theta}{\theta}\right)(\theta\lambda_i)^2}{(\theta\lambda_i + 0.5)^2}, i = 1, \ldots, n$$

$$= \frac{\theta\lambda_i + \theta(1-\theta)(\lambda_i)^2}{(\theta\lambda_i + 0.5)^2}$$

For the small sample comparison, we have restricted our attention to a model consisting of single covariate, say X. Then the expression for regression estimators from (3.5) is

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(z_i - \bar{z})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \qquad (3.7)$$

and $\alpha = \bar{z} - \beta\bar{x}$

where $\bar{z} = \sum_{i=1}^{n} \frac{Z_i}{n}$, $\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n}$, and $\hat{\theta}$ as given in (3.6) and z_i's are defined in (3.4).

4. MAXIMUM LIKELIHOOD ESTIMATORS

For the single covariate x, the maximum likelihood estimators of regression parameter (α, β) and inflation parameter θ (the proportion of zeros in the models) are obtained by solving the likelihood equation (3.1). Closed form solution do not exists and maximum likelihood estimators of parameters are estimated by the iterative solution of

$$\begin{pmatrix} \hat{\theta}_{i+1} \\ \hat{\alpha}_{i+1} \\ \hat{\beta}_{i+1} \end{pmatrix} = \begin{pmatrix} \theta_i \\ \alpha_i \\ \beta_i \end{pmatrix} - \begin{pmatrix} \frac{\partial^2 L}{\partial \theta^2} & \frac{\partial^2 L}{\partial \theta \partial \alpha} & \frac{\partial^2 L}{\partial \alpha \partial \beta} \\ \frac{\partial^2 L}{\partial \theta \partial \alpha} & \frac{\partial^2 L}{\partial \alpha^2} & \frac{\partial^2 L}{\partial \beta^2} \\ \frac{\partial^2 L}{\partial \theta \partial \beta} & \frac{\partial^2 L}{\partial \alpha \partial \beta} & \frac{\partial^2 L}{\partial \beta^2} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial L}{\partial \theta} \\ \frac{\partial L}{\partial \alpha} \\ \frac{\partial L}{\partial \beta} \end{pmatrix} \qquad (4.1)$$

In the expression (4.1) $\frac{\partial L}{\partial \theta} = 0, \frac{\partial L}{\partial \alpha} = 0$ and $\frac{\partial L}{\partial \beta} = 0$ are the maximum likelihood equations, for the parameters θ, α and β respectively and L is the log likelihood equation defined in (3.2). One-step maximum likelihood equation is obtained by putting i = 0 and two-step maximum likelihood equation by putting i = 1 in (4.1). The initial estimators $\hat{\theta}_0, \hat{\alpha}_0, \hat{\beta}_0$ are sequential estimators used for construction of one-step and two-step estimators.

5. SMALL SAMPLE COMPARISON OF THE ESTIMATORS

Sequential estimators of our proposal were compared with one-step, two-step maximum likelihood estimator and regular maximum likelihood estimator. Our concern is on the regression co-efficient β, which determines the strength of the relationship between the dependent and independent variables. Three types of covariates viz. normal, uniform and binary were used to know the effect of nature of independent variable on dependent variable. Three small sample sizes of n = 10, 20 and 30 were used in the study. Five values of β=–1.00, –0.50, 0, +0.50, 1.00, were studied. The intercept of α is –0.5 and +0.5. The inflated parameter θ which determines the effect of zero frequency on the Poisson model is taken as 0.30, 0.50 and 0.70. Relative bias and relative mean square errors are computed using 1000 simulations.

Simulation result reveals that in majority of cases the mean relative mean square error is minimum for proposed new sequential estimator and in few cases it is minimum for one-step or two-steps maximum likelihood estimator. Similar results are noticed for all the three covariates. With this result, we conclude that the sequential estimator which we have proposed in Section 3 performs very well as compared to other estimators.

Real Application in Agriculture

For real application in agriculture, the data used for the analysis is obtained from an experiment conducted at the Department of Agricultural Entomology, UAS, Dharwad. One of the objectives of the experiment is to assess the Ovipositional repellent property of acetone extracts of botanicals under laboratory condition against spiraling whitefly. The common problem encountered is the presence of many zeros for several treatments. Conventionally, the data is analysed using logarithmic transformation, namely log $(y_i + 0.5)$, where y_i is the observed count. Such an analysis did not account for a zero inflated model and when we try to analyse the data using zero inflated Poisson regression model, we face the following problems as

(i) Newton–Rephson procedure in these settings (i.e. having many treatments) is computationally complex.

(ii) Since the regression parameters were many, in many cases, the iterative procedure did not converge.

Hence, the data were analysed by using sequential estimator and wald test. The model used is

$$\log (y_{ij} + 0.5) = \log \theta + \mu + \alpha_i , \ i = 1, \ldots 13$$

where θ is the proportion of zeros in the distribution, $\beta = (\mu, \alpha_1, \ldots \alpha_{12})$ be the parameter to be estimated and y_{ij} be the j^{th} observation in the i^{th} observation. The null hypothesis of interest is

$$H_0: \alpha_1 = \alpha_2 \ldots = \alpha_{12} = 0$$

The data of the experiment represents Ovipositional repellent property of botanicals against Aleurodicus disprsus under laboratory conditions.

Treatment sprayed	Number of egg masses observed per plant after 15th day of treatment spraying		
	1st plant	2nd plant	3rd plant
Aeorus calamus	0	0	0
Adathoda vasica	0	0	0
Allium salivum	0	0	0
Annona squamosa	0	1	3
Azadiraehta indica	0	0	0
Catheranthus roseus	0	0	0
Clerodendron inerme	0	1	2
Hyptis sauveolens	0	0	0
Nelium oleander	0	0	1
Thevetia nerrifolia	0	0	1
Vitex negundo	0	0	2
Triazophos	0	0	0
Control (untreated)	7	8	12

Estimated value of variance – covariance matrix

0.7027	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360
0.360	0.7027	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360
0.360	0.360	0.7027	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360
0.360	0.360	0.360	0.2776	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360
0.360	0.360	0.360	0.360	0.7027	0.360	0.360	0.360	0.360	0.360	0.360	0.360
0.360	0.360	0.360	0.360	0.360	0.7027	0.360	0.360	0.360	0.360	0.360	0.360
0.360	0.360	0.360	0.360	0.360	0.360	0.3063	0.360	0.360	0.360	0.360	0.360
0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.7027	0.360	0.360	0.360	0.360
0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.4983	0.360	0.360	0.360
0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.4983	0.360	0.360
0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.4259	0.360
0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.360	0.7027

Estimated Value of the Parameter

$$\hat{\alpha} = \begin{bmatrix} -2.9200 \\ -2.9200 \\ -2.9200 \\ -1.9052 \\ -2.9200 \\ -2.9200 \\ -2.0474 \\ -2.9200 \\ -2.5538 \\ -2.5538 \\ -2.3836 \\ -2.9200 \end{bmatrix}$$

Table of Test Statistic

Test statistic	Computed value of Test statistic	χ^2 (5% level of significance)	F (5% level of significance)
Wald	12.2895	21.0261	—
ANOVA F	1.4249	—	2.215

Inference of the Real Data Analysis

Table of estimated value of the parameter shows the effect of 12 acetone extract of botanical (treatment) sprayed on controlling spiraling whitefly. Diagonal values of the table of estimated value of variance covariance matrix shows the variance and above or below the diagonal value is the co-variance of the 12 parameters to be tested. By using the values of this table, Wald test statistics was computed and it showed non-significant differences between the sprays. Similarly, usual analysis of variance (ANOVA F) also showed a non-significant result. Hence, it can be concluded that all the 12 acetone extracts of botanical have similar effect on controlling the spiraling whitefly.

REFERENCES

Gan, N. (1999). General zero inflated models and their applications. Unpublished Ph.D. thesis submitted to North Carolina State University.

Ghosh, S.K., Mukhopadhyay, P. and Lu, J.C. (1999). Bayesion analysis of zero-inflated regression models. Article submitted to Technometrics. (Source: Gan, N. (1999)).

Greene, W. (1994). Accounting for excess zeros and sample selection in poisson and negative binomial regression models. Working paper Ec-94-10, Department of Economics, New York University.

Gupta, P., Gupta, R. and Tripathi, R. (1996). Analysis of zero adjusted count data. *Computational Statistics and Data Analysis*, **23**, 207-218.

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

Lehmann, E.L. (1983). *Theory of Point Estimation*. John Wiley, New York.

Li, C.S., Lu, J.C., Park, J., Kim, K., Brinkely, P.A. and Peterson, J.P. (1999). Multivariate zero-inflated poisson models and their applications. *Technometrics*, **41**, 29-38.

Morel, J.G. and Koehler, K.J. (1995). A one-step Gauss Newton estimator for modeling categorical data with extraneous variation. *Appl. Statist.*, **44**, 187-200.

Rao, C.R. (1973). *Linear Statistical Inferences and its Applications*. 2nd ed., Wiley, New York.

Ridout, M.S., Demetrio, C.G.B. and Hinde, J.P. (1998). Models for counts data with many zeros. *Proc. of the XIXth International Biometric Conference, Cape Town*, Invited papers, 179-192.

Table 1. Estimated mean relative m.s.e (%) of parameter q in ZIP regression model

Cova-riate	Parameter		n = 10				n = 20				n = 30			
	θ	α	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e
Normal	0.30	–0.5	16.21	16.71	16.82	16.74	52.04	56.71	56.39	117.46	54.15	49.02	45.22	110.87
		+0.5	44.84	47.57	48.89	50.06	31.21	35.34	38.58	43.43	28.61	27.77	29.55	30.64
	0.50	–0.5	21.57	20.22	20.80	20.97	10.93	14.61	16.04	16.77	4.15	4.67	4.83	4.12
		+0.5	51.72	50.17	49.88	48.97	8.48	9.73	10.20	10.46	34.63	34.98	34.98	34.46
	0.70	–0.5	27.52	26.69	26.17	26.03	28.18	24.30	32.05	31.42	34.20	20.38	25.64	27.40
		+0.5	10.11	11.21	11.26	12.01	12.73	19.26	18.48	18.21	11.76	15.26	14.87	14.77
Uniform	0.30	–0.5	41.03	40.22	40.71	40.03	53.51	43.87	54.00	138.03	53.22	33.91	43.89	121.87
		+0.5	30.26	33.21	36.72	39.22	31.60	35.38	39.42	46.99	27.08	24.77	28.94	32.02
	0.50	–0.5	29.39	30.51	30.09	31.72	27.92	29.56	30.16	30.47	22.95	23.40	23.53	23.58
		+0.5	32.46	39.02	42.50	45.57	17.83	19.32	19.90	20.21	26.57	26.35	26.19	25.97
	0.70	–0.5	31.21	31.12	30.67	30.51	34.96	24.87	30.80	33.06	35.85	21.50	26.43	30.36
		+0.5	10.07	11.51	11.62	12.66	11.97	18.18	17.99	17.81	8.31	11.89	14.06	13.96
Binary	0.30	–0.5	61.22	59.71	58.72	63.26	53.21	54.76	49.38	121.14	53.20	56.16	44.08	109.79
		+0.5	63.28	58.71	57.72	68.77	27.95	30.66	34.62	57.64	27.33	27.09	28.74	35.24
	0.50	–0.5	51.26	55.71	56.28	71.28	61.22	67.21	69.11	75.28	49.27	56.21	55.71	54.22
		+0.5	15.71	16.22	16.13	17.71	11.27	9.89	9.65	10.07	9.66	10.22	11.21	10.27
	0.70	–0.5	41.28	43.20	43.27	44.26	32.65	28.22	29.07	32.59	35.17	25.46	25.76	30.12
		+0.5	11.21	11.07	11.13	12.52	12.90	19.32	19.98	20.08	11.35	15.03	14.88	14.96

Table 2. Estimated mean relative m.s.e (%) of parameter a in ZIP regression model

Cova-	Parameter		n = 10				n = 20				n = 30			
riate	θ	α	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e
Normal	0.30	–0.5	281.24	113.94	64.22	56.53	237.84	274.36	228.77	262.90	236.80	179.18	127.65	282.36
		+0.5	151.50	87.10	63.26	70.02	119.54	126.35	112.98	217.00	121.27	82.61	81.91	92.28
	0.50	–0.5	170.06	71.74	77.07	74.45	195.65	83.07	69.72	67.61	164.34	79.56	82.28	81.34
		+0.5	24.17	76.47	57.81	66.20	29.11	29.48	31.06	31.53	32.08	28.91	25.04	25.95
	0.70	–0.5	76.27	105.19	89.90	87.74	78.14	108.46	98.11	167.38	77.99	74.24	75.56	95.29
		+0.5	87.37	137.69	85.83	94.14	30.39	53.86	58.74	58.18	27.01	40.40	43.96	44.03
Uniform	0.30	–0.5	285.00	163.82	127.68	194.53	235.15	249.95	227.24	266.61	235.61	179.87	142.10	227.64
		+0.5	18.00	68.08	42.90	18.70	76.81	105.72	106.33	183.64	76.87	61.27	77.54	101.36
	0.50	–0.5	173.58	109.51	74.32	49.79	151.43	84.40	74.90	100.27	155.60	58.33	63.20	62.90
		+0.5	37.26	29.46	28.33	29.07	39.82	108.44	108.56	108.52	22.64	14.60	16.30	24.60
	0.70	–0.5	88.20	34.88	114.77	102.32	113.54	126.99	95.38	174.46	112.76	81.44	83.17	101.77
		+0.5	80.62	141.44	66.43	83.86	30.03	54.84	60.12	59.87	28.43	40.41	43.42	43.54
Binary	0.30	–0.5	288.50	151.51	82.31	75.09	231.06	268.94	220.19	261.37	232.85	139.78	134.44	159.00
		+0.5	225.54	94.24	83.90	96.00	82.00	71.12	95.01	116.31	76.33	69.25	95.67	102.28
	0.50	–0.5	186.35	72.26	57.15	57.63	143.17	130.20	134.60	139.00	138.35	51.77	34.61	15.44
		+0.5	31.80	29.80	36.10	35.50	40.70	48.90	91.20	103.40	28.02	23.00	23.20	19.50
	0.70	–0.5	254.00	209.00	248.00	247.00	110.46	126.63	116.07	130.06	109.60	94.50	109.34	109.19
		+0.5	45.12	48.15	38.27	40.55	42.94	64.71	68.11	67.53	32.71	49.92	54.99	54.97

Table 3. Estimated mean relative m.s.e (%) of parameter b for normal covariate in ZIP regression model

Parameter		n = 10				n = 20				n = 30			
θ	β	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e
0.30	−1.00	33.73	43.62	61.51	63.07	78.98	138.34	101.33	111.50	79.14	97.15	62.76	67.95
	−0.50	75.32	130.20	135.63	141.32	80.21	171.16	127.59	130.00	78.50	130.31	90.76	131.98
	+0.50	90.35	110.23	125.28	115.26	85.42	146.57	131.18	146.34	84.19	113.36	92.15	99.40
	+1.00	110.29	95.63	92.69	85.23	84.45	155.23	148.51	210.36	80.27	62.45	44.58	30.90
0.50	−1.00	76.29	80.39	79.52	95.62	74.26	77.46	76.01	90.26	73.22	72.02	71.25	87.11
	−0.50	82.88	126.67	127.23	120.26	60.88	80.67	85.16	100.50	51.81	71.77	79.12	91.05
	+0.50	90.67	85.84	79.20	135.69	77.73	71.25	61.07	110.09	69.88	70.00	60.92	100.54
	+1.00	79.39	100.64	110.20	100.23	71.26	89.17	99.21	95.26	67.37	83.32	90.06	88.26
0.70	−1.00	68.67	97.14	71.03	68.25	55.25	50.75	41.37	39.91	50.94	32.94	23.07	20.42
	−0.50	63.21	35.26	40.47	59.21	55.32	91.49	62.36	76.26	54.67	60.60	44.46	48.27
	+0.50	69.46	44.23	43.26	44.30	66.41	67.22	70.35	115.96	65.04	52.15	50.40	52.19
	+1.00	65.41	47.10	49.23	51.29	60.42	39.43	36.00	34.66	57.61	25.09	21.99	21.44

Table 4. Estimated mean relative m.s.e (%) of parameter β for uniform covariate in ZIP regression model

Parameter		n = 10				n = 20				n = 30			
θ	β	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e
0.30	−1.00	85.35	191.36	177.37	250.53	79.22	187.63	170.63	273.53	79.47	103.22	85.42	260.31
	−0.50	87.55	195.16	188.60	227.47	83.88	186.74	187.24	226.01	81.64	217.95	167.38	291.68
	+0.50	95.45	240.57	241.06	251.26	93.64	247.67	237.23	254.45	90.37	170.69	159.13	179.56
	+1.00	90.07	170.21	176.82	181.87	86.49	168.83	179.87	178.87	84.31	179.52	222.19	92.04
0.50	−1.00	77.26	75.29	74.15	70.22	75.31	73.85	72.66	70.07	74.17	72.69	72.71	69.12
	−0.50	82.51	80.37	80.22	81.56	81.20	78.10	78.48	80.22	80.29	77.87	77.42	80.30
	+0.50	86.77	84.13	84.96	85.27	86.33	82.68	82.81	83.03	84.36	82.35	81.97	81.02
	+1.00	81.28	78.28	79.17	80.27	80.17	76.12	77.79	78.32	80.12	76.17	77.03	79.11
0.70	−1.00	47.62	51.33	47.15	49.28	45.49	47.50	44.02	43.30	51.20	51.79	40.95	42.87
	−0.50	68.71	85.59	80.36	81.26	61.71	84.19	81.63	82.39	57.38	93.40	72.90	83.90
	+0.50	83.54	84.18	103.82	105.11	79.39	113.57	111.28	113.84	71.15	81.58	76.99	82.78
	+1.00	64.53	53.51	52.35	87.28	61.65	52.14	56.53	44.97	63.00	47.03	45.01	44.93

Table 5. Estimated mean relative m.s.e (%) of parameter b for binary covariate in ZIP regression model

Parameter		n = 10				n = 20				n = 30			
θ	β	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e	SE	One step	Two step	m.l.e
0.30	−1.00	83.41	105.02	119.11	145.27	81.98	101.91	117.57	140.85	83.44	112.97	124.66	157.25
	−0.50	89.28	140.22	171.21	142.37	84.62	146.96	184.56	240.18	89.09	144.09	174.48	121.86
	+0.50	95.11	110.27	129.59	135.81	88.16	108.74	126.59	132.68	86.20	146.16	174.81	138.47
	+1.00	81.21	107.27	125.11	65.13	77.95	54.46	59.62	64.12	81.59	71.87	84.79	75.17
0.50	−1.00	77.09	81.21	85.77	89.29	81.27	69.11	67.26	69.03	89.27	85.16	87.12	91.57
	−0.50	73.21	85.71	86.31	88.91	85.17	71.13	73.37	75.17	88.15	89.11	90.51	93.22
	+0.50	72.13	87.11	87.91	88.07	110.23	99.17	98.53	100.71	91.23	90.27	88.29	78.26
	+1.00	79.25	80.57	81.21	83.22	113.22	111.17	110.07	105.16	93.85	81.57	81.31	80.12
0.70	−1.00	71.72	69.17	69.92	81.32	59.78	88.45	63.32	83.70	63.08	91.94	60.94	138.76
	−0.50	66.54	61.27	63.25	117.63	69.72	142.75	193.56	111.06	66.27	135.59	94.40	157.08
	+0.50	77.63	79.21	91.23	119.29	82.18	115.28	103.84	111.52	71.80	80.25	75.72	75.81
	+1.00	55.61	55.58	55.41	55.13	56.80	56.06	47.50	48.07	54.96	38.23	34.03	34.44