BIOMETRICS 57, 219–223
March 2001

A Score Test for Testing a Zero-Inflated Poisson Regression Model Against Zero-Inflated Negative Binomial Alternatives

Martin Ridout,[1] John Hinde,[2] and Clarice G. B. Demétrio[3]

[1]Horticulture Research International, East Malling, West Malling, Kent ME19 6BJ, U.K.
Current address: Institute of Mathematics and Statistics, University of Kent at Canterbury, Canterbury, Kent CT2 7NF, U.K.
email: M.S.ridout@ukc.ac.uk

[2]School of Mathematical Sciences, Laver Building, University of Exeter, North Park Road, Exeter EX4 4QE, U.K.
email: j.p.hinde@exeter.ac.uk

[3]DME/ESALQ, University of São Paulo, 13418-900 Piracicaba, SP, Brazil
email: clarice@carpa.ciagri.usp.br

SUMMARY. Count data often show a higher incidence of zero counts than would be expected if the data were Poisson distributed. Zero-inflated Poisson regression models are a useful class of models for such data, but parameter estimates may be seriously biased if the nonzero counts are overdispersed in relation to the Poisson distribution. We therefore provide a score test for testing zero-inflated Poisson regression models against zero-inflated negative binomial alternatives.

KEY WORDS: Count data; Negative binomial; Poisson regression model; Score test; Zero inflation.

1. Introduction

Poisson regression models provide a standard framework for the analysis of count data. In practice, however, count data often have a higher incidence of zero counts than is expected for the Poisson distribution. One approach to analyzing count data with many zeros is to use a zero-inflated Poisson distribution. This assumes that the population consists of two types of individual. The first type gives a Poisson-distributed count, which may be zero, whereas the second type always gives a zero count. The distribution has two parameters, the mean of the Poisson distribution (λ) and the proportion of individuals that are of the second type (ω).

In a regression context, either or both of these parameters may depend on covariates. Zero-inflated Poisson regression models originated in the econometrics literature (Mullahy 1986), but their use has become more widespread recently, particularly since the publication of Lambert (1992). Van den Broek (1995) has given a score test for testing a standard Poisson regression model against a zero-inflated Poisson alternative. A similar test was derived by Mullahy (1986), though the details were not included in the published paper.

Although the zero-inflated Poisson is more general than the standard Poisson distribution, it remains rather inflexible. The nonzero counts necessarily follow a zero-truncated Poisson distribution. In practice, count data are often overdispersed and alternative distributions, such as the zero-inflated negative binomial distribution, may be more appropriate than the zero-inflated Poisson. In the context of standard Poisson regression, overdispersion has little effect on parameter estimates but leads to underestimation of standard errors (Cox 1983). This underestimation may be corrected by the use of quasi-likelihood methods (e.g., Breslow 1984). However, zero-truncated and zero-inflated models depend explicitly on the probability of a zero count in the underlying distribution. Thus, if the underlying distribution is assumed to be Poisson but is in fact negative binomial, the wrong functional form will be used for the zero probability and this in turn will lead to inconsistent parameter estimates. Grogger and Carson (1991) discuss this point in detail in the context of zero-truncated regression models. These authors fitted zero-truncated Poisson regression models to data simulated from zero-truncated negative binomial distributions and found biases of up to 30% in estimated parameters. Similar arguments extend to zero-inflated models, and therefore the issue of overdispersion is much more important in zero-inflated and zero-truncated models than in standard models.

In this paper, we provide a score test for testing a zero-inflated Poisson regression model against a zero-inflated negative binomial alternative. Section 2 defines the null and alternative models. We use a general parameterization of the negative binomial distribution that incorporates two common variance–mean relationships. In Section 3, we give the score test. Our development parallels Gurmu (1991), who derived a score test for testing a zero-truncated Poisson regression model against a zero-truncated negative binomial alternative. Section 4 discusses an example with no covariates, where our test statistic coincides with Gurmu's statistic. Section 5 gives an example with covariates, where the test statistic is more complex than Gurmu's.

2. Models

We consider a zero-inflated Poisson regression model in which the response variable Y_i $(i = 1, \ldots, n)$ has the distribution

$$\Pr(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)\exp(-\lambda_i), & y_i = 0, \\ (1-\omega_i)\exp(-\lambda_i)\lambda_i^{y_i}/y_i!, & y_i > 0, \end{cases}$$

where the parameters λ_i and ω_i depend on (vectors of) covariates x_i and z_i, respectively. In this paper, we assume the specific models

$$\log(\lambda_i) = x_i^t\beta$$

and

$$\log\left(\frac{\omega_i}{1-\omega_i}\right) = z_i^t\gamma,$$

but similar results could be derived for other link functions. The mean and the variance of the distribution are $\mathrm{E}(Y_i) = (1-\omega_i)\lambda_i$ and $\mathrm{var}(Y_i) = (1-\omega_i)\lambda_i(1+\omega_i\lambda_i)$.

The alternative that we consider is that Y_i has a zero-inflated negative binomial distribution. Specifically, we consider the distribution

$$\Pr(Y_i = y_i) = \begin{cases} \omega_i + (1-\omega_i)(1+\alpha\lambda_i^c)^{-\lambda_i^{1-c}/\alpha}, & y_i = 0, \\ (1-\omega_i)\dfrac{\Gamma(y_i + \lambda_i^{1-c}/\alpha)}{y_i!\Gamma(\lambda_i^{1-c}/\alpha)} & \\ \quad\times(1+\alpha\lambda_i^c)^{-\lambda_i^{1-c}/\alpha} & \\ \quad\times(1+\lambda_i^{-c}/\alpha)^{-y_i}, & y_i > 0, \end{cases}$$

where α (≥ 0) is a dispersion parameter that is assumed not to depend on covariates. The mean and the variance of the distribution are $\mathrm{E}(Y_i) = (1-\omega_i)\lambda_i$ and $\mathrm{var}(Y_i) = (1-\omega_i)\lambda_i(1+\omega_i\lambda_i + \alpha\lambda_i^c)$. This distribution reduces to the zero-inflated Poisson distribution in the limit $\alpha \to 0$. The mean of the underlying negative binomial distribution is λ_i. The index c identifies the particular form of negative binomial distribution (Saha and Dong 1997); for $c = 0$, the variance of the underlying negative binomial distribution is $(1+\alpha)\lambda_i$ and, for $c = 1$, the variance is $\lambda_i + \alpha\lambda_i^2$. The parameterization therefore includes the two most common forms of the negative binomial (e.g., McCullagh and Nelder 1989, p. 199). This parameterization has been used by Gurmu (1991) in developing score tests for zero-truncated Poisson regression models against zero-truncated negative binomial alternatives.

3. The Score Test

We now derive the score test for testing the null hypothesis H_0: $\alpha = 0$ against the alternative H_1: $\alpha > 0$. Let L denote the log-likelihood function. The efficient score, S, obtained by evaluating the derivative $\partial L/\partial\alpha$ under the null hypothesis is

$$S = \frac{1}{2}\sum_i \hat{\lambda}_i^{c-1}\left\{\left[\left(y_i - \hat{\lambda}_i\right)^2 - y_i\right] - I_{(y_i=0)}\hat{\lambda}_i^2\hat{\omega}_i/\hat{p}_{0,i}\right\},$$

where symbols with hats denote estimates under the null hypothesis, $p_{0,i} = \Pr(Y_i = 0)$ from the zero-inflated model, and $I_{(y=0)}$ is an indicator variable taking the value one when $y = 0$ and zero otherwise. S can also be written as follows:

$$\frac{1}{2}\sum_i \hat{\lambda}_i^{c-1}\left\{\left[(y_i - \hat{\mu}_i)^2 - \hat{\mu}_i\left(1 + \frac{\hat{\omega}_i\hat{\mu}_i}{1-\hat{\omega}_i}\right)\right] - \left(1 + \frac{2\hat{\mu}_i}{1-\hat{\omega}_i}\right)(y_i - \hat{\mu}_i) - \frac{\hat{\omega}_i\hat{\mu}_i^2}{(1-\hat{\omega}_i)^2\hat{p}_{0,i}}\left(I_{(y_i=0)} - \hat{p}_{0,i}\right)\right\},$$

where $\hat{\mu}_i = (1-\hat{\omega}_i)\hat{\lambda}_i$ denotes the fitted mean under the null hypothesis. This makes it clear that the score test compares the first two sample moments and the zero frequency with the corresponding theoretical values under the zero-inflated model.

The expected Fisher information matrix may be partitioned as follows:

$$J = \begin{bmatrix} J_{\alpha\alpha} & J_{\alpha\beta} & J_{\alpha\gamma} \\ J_{\alpha\beta}^t & J_{\beta\beta} & J_{\beta\gamma} \\ J_{\alpha\gamma}^t & J_{\beta\gamma}^t & J_{\gamma\gamma} \end{bmatrix},$$

where $J_{\alpha\alpha}$ is a scalar and the other elements are, in general, matrices with dimensions determined by the dimensions of the parameter vectors β and γ. It may be shown that, in the limit as $\alpha \to 0$, typical elements of the information matrix are

$$\begin{aligned} J_{\alpha\alpha} &= \frac{1}{4}\sum_i \lambda_i^{2c}\{2(1-\omega_i) - \lambda_i\kappa_i\} \\ J_{\alpha\beta_j} &= \frac{1}{2}\sum_i \lambda_i^{c+1}\kappa_i x_{i,j} \\ J_{\alpha\gamma_j} &= \frac{1}{2}\sum_i \lambda_i^{c}\kappa_i z_{i,j} \\ J_{\beta_j\beta_k} &= \sum_i \lambda_i\{(1-\omega_i) - \kappa_i\}x_{i,j}x_{i,k} \\ J_{\beta_j\gamma_k} &= -\sum_i \kappa_i x_{i,j}z_{i,k} \\ J_{\gamma_j\gamma_k} &= \sum_i \frac{\omega_i^2(1-p_{0,i})}{p_{0,i}}z_{i,j}z_{i,k}, \end{aligned}$$

where

$$\kappa_i = \lambda_i\omega_i\left(1 - \frac{\omega_i}{p_{0,i}}\right).$$

The score statistic for testing H_0 is then

$$T = S\sqrt{\hat{J}^{\alpha\alpha}},$$

where $\hat{J}^{\alpha\alpha}$ is the upper left-hand element of the inverse information matrix evaluated at the maximum likelihood estimates under H_0. Asymptotically, under H_0, T has a standard normal distribution. A one-sided test is appropriate because the alternative hypothesis is H_1: $\alpha > 0$, with large positive values of T providing evidence against the null hypothesis. The element $\hat{J}^{\alpha\alpha}$ may be computed from standard formulas for the inverse of a partitioned matrix (e.g., Healy 1986, p. 24).

Table 1
Number of movements made by a fetal lamb in 240 consecutive 5-second intervals, from Leroux and Puterman (1992); fitted frequencies are from a zero-inflated Poisson distribution

	Number of movements							
	0	1	2	3	4	5	6	7
Number of intervals	182	41	12	2	2	0	0	1
Fitted frequencies	182.0	36.9	15.6	4.4	0.9	0.2	0.0	0.0

It is also possible to write the numerator S as

$$S = \frac{1}{2}\sum_i I_{(y_i>0)}\hat{\lambda}_i^{c-1} \times \left\{\left(y_i - \hat{\mu}_i^{PP}\right)^2 - y_i + \left[\left(y_i - \hat{\mu}_i^{PP}\right) + y_i\right]\hat{\delta}_i\right\} - \frac{1}{2}\sum_i \hat{\lambda}_i^{c-1}\left\{\hat{\lambda}_i\hat{\delta}_i\left(1 - I_{(y_i=0)}/\hat{p}_{0,i}\right)\right\},$$

with $\hat{\mu}_i^{PP}$ denoting the estimated mean of the nonzero counts and $\hat{\delta}_i = \hat{\mu}_i^{PP} - \hat{\lambda}_i$. The first term corresponds to the numerator of Gurmu's test statistic (Gurmu 1991, equation (2.7)) based on zero-truncated data. In certain very special cases, the second term in the above expression is zero (as a consequence of the form of the estimating equations for γ) and the relevant entries in the inverse information matrices also become equal, giving identical test statistics. In particular, this happens when the models for λ and ω are full-interaction factorial models, with the ω model at least as complex as the λ model, and there is zero inflation for each combination of levels within the ω model. The example in the next section, with no covariates, is the simplest example of this equivalence. In general, e.g., when there are continuous covariates, the tests will differ, although in many practical applications they will give similar results. The test given here is for use specifically within the context of zero-inflated models, while the Gurmu test says nothing about the zero process and, indeed, the full data may be zero-deflated at some or all levels of the covariates, which would be evidence against overdispersion.

4. An Example with No Covariates

We consider data from Leroux and Puterman (1992) that give the number of movements made by a fetal lamb in each of 240 consecutive 5-second intervals (Table 1). The data were discussed further by Douglas (1994), who fitted many different distributions to the data, including the zero-inflated Poisson distribution. Subsequently, Gupta, Gupta, and Tripathi (1996) showed that a zero-inflated form of the generalized Poisson distribution provides a very good fit to the data. Here, for illustration, we ignore the temporal structure of the data and proceed as if the 240 observations were independent; Leroux and Puterman (1992) consider some alternative models that incorporate the temporal structure.

When there are no covariates for λ and ω, the score test simplifies to

$$T = \frac{\sum_i \left[\left(y_i - \hat{\lambda}\right)^2 - y_i\right] - n\hat{\lambda}^2\hat{\omega}}{\hat{\lambda}\sqrt{n(1-\hat{\omega})\left(2 - \dfrac{\hat{\lambda}^2}{e^{\hat{\lambda}} - 1 - \hat{\lambda}}\right)}}.$$

This statistic does not depend on the parameter c. This is because, when there are no covariates and the data are thus from a single distribution, different values of c just give different parameterizations of this distribution. Consequently, T does not depend on the value of c. Also, the fitted frequency of zeros from the zero-inflated Poisson distribution is equal to the observed frequency, and the second term in our final reexpression of S is zero, giving Gurmu's statistic. Thus, in the absence of covariates, zero counts are irrelevant for assessing goodness-of-fit. Therefore, testing the zero-inflated Poisson distribution against a zero-inflated negative binomial alternative is equivalent to testing the fit of the zero-truncated Poisson distribution to the nonzero counts against a zero-truncated negative binomial alternative.

For the fetal lamb movements, the parameter estimates for the zero-inflated Poisson distribution are $\hat{\lambda} = 0.847$ and $\hat{\omega} = 0.577$, implying that $\hat{p}_0 = 0.758$. The score statistic is $T = 4.72$, providing considerable evidence against the zero-inflated Poisson distribution. However, the test statistic is influenced strongly by the single time period in which seven movements were observed. When this observation is excluded, the test statistic is reduced to $T = 1.44$.

5. An Example with Covariates

Ridout, Demétrio, and Hinde (1998, Table 1; reproduced here as Table 2) give the number of roots produced by 270 micropropagated shoots of the columnar apple cultivar Trajan. The shoots had been produced under an 8- or 16-hour photoperiod in culture systems that utilized one of four different concentrations of the cytokinin BAP in the culture medium. There were 30 or 40 shoots of each of these eight treatment combinations. Of the 140 shoots produced under the 8-hour photoperiod, only 2 failed to produce roots, but 62 of the 130 shoots produced under the 16-hour photoperiod failed to root. Ridout et al. (1998) fitted various models to these data, based on the Poisson and negative binomial distributions (with $c = 1$) and their zero-inflated counterparts.

Here we consider specifically a zero-inflated Poisson model in which there is a distinct value of λ for each of the eight treatment combinations, but ω differs only for the two photoperiods. With $c = 1$, the score statistic is $T = 3.58$. The

Table 2

Frequency distributions of the number of roots produced by 270 shoots of the apple cultivar Trajan, classified by the experimental conditions (BAP concentration and photoperiod) under which the shoots were reared; shown are the numbers of shoots that produced 0, 1, ..., 12 roots; counts that exceeded 12 are shown individually

	Photoperiod							
	8				16			
BAP (μM):	2.2	4.4	8.8	17.6	2.2	4.4	8.8	17.6
Number of roots								
0	0	0	0	2	15	16	12	19
1	3	0	0	0	0	2	3	2
2	2	3	1	0	2	1	2	2
3	3	0	2	2	2	1	1	4
4	6	1	4	2	1	2	2	3
5	3	0	4	5	2	1	2	1
6	2	3	4	5	1	2	3	4
7	2	7	4	4	0	0	1	3
8	3	3	7	8	1	1	0	0
9	1	5	5	3	3	0	2	2
10	2	3	4	4	1	3	0	0
11	1	4	1	4	1	0	1	0
12	0	0	2	0	1	1	1	0
>12	13, 17	13	14, 14	14				
Number of shoots	30	30	40	40	30	30	30	40
Mean	5.8	7.8	7.5	7.2	3.3	2.7	3.1	2.5
Variance	14.1	7.6	8.5	8.8	16.6	14.8	13.5	8.5
Overdispersion index[a]	1.42	−0.03	0.13	0.22	4.06	4.40	3.31	2.47

[a] Overdispersion index = (variance − mean)/mean.

square of this statistic (12.8) agrees well with the likelihood ratio statistic for comparing the zero-inflated Poisson model with the zero-inflated negative binomial with $c = 1$ (12.0, from Table 2 of Ridout et al. 1998). For $c = 0$, the score test statistic is $T = 4.31$. Both score tests indicate clearly that the zero-inflated Poisson model is unsuitable for these data.

6. Discussion

The principle advantage of the score test in comparison with the likelihood ratio test is that the score test does not require the more complex model to be fitted. Lambert (1992) gave an EM algorithm for fitting zero-inflated Poisson regression models; this is straightforward to implement in any statistical package that includes facilities for fitting weighted generalized linear models for Poisson and binomial data. The additional computations required for the score statistic are also straightforward to implement, provided that the package has facilities for handling matrices. Copies of our program, using the package Genstat5, are available on request.

The principle disadvantage of the score test in comparison with the likelihood ratio test is that the asymptotic distribution of the score statistic is often approached more slowly than that of the likelihood ratio statistic. Thus, significance levels derived from the score statistic can be misleading, particularly in small samples. For testing Poisson regression models against negative binomial alternatives, the score test tends to give too few significant results (Dean and Lawless 1989). This is also true when the distributions are zero-truncated (Gurmu 1991). Both of these papers gave correction factors to improve the small-sample distribution of the score statistic, though Dean and Lawless (1989) note that, even with the correction, a sample size of perhaps 50 is required for reasonable accuracy. The adjusted statistics derived by Dean and Lawless (1989) and Gurmu (1991) correct for the negative bias of the test statistic but not for its skewness. Corrections that consider skewness are much more complicated (e.g., Cordeiro, Ferrari, and Cysneiros 1998). Derivation of a bias correction factor for the score test in this paper is complicated by the presence of separate submodels for λ and ω and will be pursued elsewhere. From a practical perspective, we suggest that, if the uncorrected score test gives even a weak indication that the zero-inflated Poisson model is inappropriate, say at the 10% significance level, a zero-inflated negative binomial model should be fitted to the data. The two models may then be compared by the likelihood ratio statistic, which is likely to be more accurate. On the other hand, if the score test gives no indication of lack of fit, inferences based on the zero-inflated Poisson model can be made with increased confidence. Correction factors are not greatly important in our examples because they involved large sample sizes and gave highly significant values of the test statistic.

ACKNOWLEDGEMENTS

We are grateful to the Royal Society, to CNPq, and to FAPESP for funding exchange visits that made this work possible and to the referees and associate editor for helpful comments.

Résumé

Les données de comptage montrent souvent une plus forte incidence de comptages nuls que l'on en attendrait pour des données correspondant à un modèle de Poisson. Les modèles de régression de Poisson à expansion de zéros représentent une classe utile de modèles pour de telles données, mais les estimateurs des paramètres peuvent être sérieusement biaisés si les comptages non nuls sont surdispersés par rapport à l'hypothèse poissonienne. Nous proposons alors un test du score pour comparer des modèles de régression de Poisson à expansion de zéros, à des alternatives binomiales négatives à expansion de zéros.

References

Breslow, N. E. (1984). Extra-Poisson variation in log-linear models. *Applied Statistics* **33,** 38–44.

Cordeiro, G. M., Ferrari, S. L. P., and Cysneiros, A. H. M. A. (1998). A formula to improve score test statistics. *Journal of Statistical Computation and Simulation* **62,** 123–136.

Cox, D. R. (1983). Some remarks on over-dispersion. *Biometrika* **70,** 269–274.

Dean, C. B. and Lawless, J. F. (1989). Tests for detecting overdispersion in Poisson regression models. *Journal of the American Statistical Association* **84,** 467–472.

Douglas, J. B. (1994). Empirical fitting of discrete distributions. *Biometrics* **50,** 576–579.

Grogger, J. and Carson, R. (1991). Models for truncated counts. *Journal of Applied Econometrics* **6,** 225–238.

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

Gurmu, S. (1991). Tests for detecting overdispersion in the positive Poisson regression model. *Journal of Business and Economic Statistics* **9,** 215–222.

Healy, M. J. R. (1986). *Matrices for Statistics.* Oxford: Clarendon Press.

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

Leroux, B. G. and Puterman, M. L. (1992). Maximum-penalized-likelihood estimation for independent and Markov-dependent mixture models. *Biometrics* **48,** 545–558.

McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models,* 2nd edition. London: Chapman and Hall.

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

Ridout, M. S., Demétrio, C. G. B., and Hinde, J. P. (1998). Models for count data with many zeros. *Proceedings of the XIXth International Biometrics Conference,* Cape Town, Invited Papers, pp. 179–192.

Saha, A. and Dong, D. (1997). Estimating nested count data models. *Oxford Bulletin of Economics and Statistics* **59,** 423–430.

van den Broek, J. (1995). A score test for zero inflation in a Poisson distribution. *Biometrics* **51,** 738–743.

Received August 1999. *Revised May* 2000.
Accepted May 2000.