Statistics and Its Interface Volume 11 (2018) 339–351

A zero-and-one inflated Poisson model and its application

Wenchen Liu, Yincai Tang[*,†], and Ancha Xu[*,†]

To model count data with excess zeros and excess ones, Melkersson and Olsson (1999) proposed a zero-and-one-inflated Poisson (ZOIP) distribution. Zhang, Tian and Ng (2016) studied the properties and likelihood-based inference methods on ZOIP model. However, they only propose some estimation methods for the ZOIP model. In this paper, the maximum likelihood estimation (MLE) and Bayesian estimation for this model are investigated and some properties are derived. The reference prior and the Jeffreys prior are derived for this model. It is further shown that they are second-order matching priors and the posterior distributions based on these priors are proper under a relatively mild condition. And the zero-and-one-inflated Poisson regression model has also been discussed. A simulation study based on proposed sampling algorithm is conducted to assess the performance of the proposed estimation for various sample sizes. Finally, two real data sets are analyzed to illustrate the practicability of the proposed method.

Keywords and phrases: Zero-and-one-inflated Poisson model, Objective Bayes, Reference prior, Metropolis-Hastings algorithm.

1. INTRODUCTION

Count data with excess zeros arise frequently in various fields when dealing with manufacturing defects (Lambert [18]), patent applications (Crepon & Duguet [7]), road safety (Miaou [21]), species abundance (Welsh et al. [27]; Faddy [9]), use of recreational facilities (Gurmu & Trivedi [13]; Shonkwiler & Shaw [26]) and Legionellosis infection (Xu et al. [29]), etc. Conventional models such as Poisson or negative binomial distribution may not fit these data well, and seriously underestimate the zero-count probability, which is an important indicator (of production quality in manufacturing for example). Various methods have been developed to address this issue, in which zero-inflated Possion (ZIP) model proposed by Lambert [18] plays an important part. For modeling complete female fertility Melkersson and Rooth [20] proposed a zero-and-two-inflated count data model, which accounts for a relative excess of both zero and two children. However, in many cases, count data may contain excess zeros and ones simultaneously. For example, it is most probable that in a shopping trip one does not buy or just buy one item at a clothing store; one may be infected by some virus for at most one time due to the generation of corresponding antibodies once after the infection. Melkersson and Olsson [19] extended the zero-inflated Poisson distribution to a zero-and-one-inflated Poisson (ZOIP) distribution to analyse the number of visits to a dentist in a year for a sample of adult Swedes. The major goal of Melkersson and Olsson [19] is to fit the dentist visiting data in Sweden. They only considered the covariates with the parameter of Poisson distribution. Zhang et al. [31] studied the properties and likelihood-based inference methods on ZOIP model. They constructed five equivalent stochastic representations without covariates for the ZOIP random variable and maximum likelihood estimates of parameters were obtained by both the Fisher scoring and expectation-maximization algorithms. At the end of their article, testing hypotheses under large sample sizes are provided.

A random variable Y in a zero-and-one-inflated Possion (ZOIP) model can be represented as $Y = V(1 - B_1) + B_1(1 - B_2)$, where B_1 is a Bernoulli random variable with success probability p_0, B_2 is a Bernoulli random variable with success probability p_1, V follows a Poisson distribution with rate parameter θ and B_1, B_2 and V are mutually independent. The relation between Y and (B_1, B_2, V) is

$$(1)\qquad \begin{cases} (Y=0) \Leftrightarrow (V=0, B_1=0) \cup (B_1=1, B_2=1) \\ (Y=1) \Leftrightarrow (V=1, B_1=0) \cup (B_1=1, B_2=0) \\ (Y=k) \Leftrightarrow (V=k, B_1=0), k=2,3,\dots \end{cases}$$

Then the probability mass function of the nonnegative integer-valued random variable Y is

$$(2)\qquad \Pr(Y=k) = \begin{cases} p_0 p_1 + (1-p_0)\,\mathrm{e}^{-\theta}, & \text{if } k=0, \\ p_0(1-p_1) + (1-p_0)\,\theta\mathrm{e}^{-\theta}, & \text{if } k=1, \\ (1-p_0)\,\frac{\theta^k}{k!}\mathrm{e}^{-\theta}, & \text{if } k\ge 2, \end{cases}$$

with $0 \le p_0 \le 1$, $0 \le p_1 \le 1$, and $\theta > 0$. We denote this zero-and-one-inflated Poisson model as ZOIP (p_0, p_1, θ). When $p_0 = 0$, this model is a Poisson model. When $p_1 = 1$, $p_0 \ge 0$, ZOIP becomes zero inflated Poisson model which is also called a "with-zeros Poisson" model by Mullahy [23]. Broek [5] proposed a score test model to test whether a

*Corresponding authors.

†Supported by NSF of China (11271136, 81530086, 11671303, 11201345). The 111 project (B14019). Supported by NSF of Zhejiang province (LY15G010006). Supported by CPSF (2015M572598).

count variable was from a ZIP model or from a Poisson model. Regression analysis based on this ZIP model has been reported by Lambert [18], Lam et al. [16], Bae et al. [1], Hasan and Sneddon [14], and Bassil et al. [2]. Regression analysis with Bayesian techniques has been considered in Ghosh et al. [11], Chen [6], Dagne [8], and Musio et al. [24].

Alternatively, we can develop the second form of the zero-and-one inflated Poisson model by the following transformation of the first form. Denoted by q_0 and q_1 the probability of Y being zero and one respectively, i.e.,

$$\begin{cases} q_0 = p_0 p_1 + (1-p_0)\mathrm{e}^{-\theta} \\ q_1 = p_0(1-p_1) + (1-p_0)\theta\mathrm{e}^{-\theta}. \end{cases} \tag{3}$$

Then the probability mass function (2) becomes

$$\Pr(Y=k) = \begin{cases} q_0, & \text{if } k=0, \\ q_1, & \text{if } k=1, \\ \frac{1-q_0-q_1}{1-\mathrm{e}^{-\theta}-\theta\mathrm{e}^{-\theta}} \frac{\theta^k \mathrm{e}^{-\theta}}{k!}, & \text{if } k\geq 2, \end{cases} \tag{4}$$

where $q_0 \geq 0, q_1 \geq 0, q_0+q_1 \leq 1$, and $\theta > 0$. When $q_1 = \frac{(1-q_0)}{\mathrm{e}^{\theta}-1}\theta$, this model was called a "hurdle Poisson" model by Mullahy [23] and King [15]. In this article, the Jeffreys and reference priors are derived for the second form of the ZOIP model which used the similar method used by Xu et al. [29] and Xu and Tang [28]. The maximum likelihood estimation and Bayesian estimation for zero-and-one-inflated Poisson regression model are also discussed in our paper.

We have listed two forms or representations of the ZOIP model, (2) and (4), in terms of (p_0, p_1, θ) and (q_0, q_1, θ) respectively. In this article, the parameter estimation of the second form of the ZOIP model is mainly studied as few analyses have been reported based on this form. The remaining part of this article is organized as follows. In Section 2, the maximum likelihood estimates (MLEs) of the parameters are obtained and shown to be unique under a mild condition. We focus our attention on the Bayesian estimation in Section 3. The reference prior and the Jeffreys prior are derived and are shown to be second-order matching priors when θ is the parameter of interest. We derive the closed forms of posterior distributions based on these priors and prove that they are proper under a relatively mild condition. The zero-and-one-inflated Poisson regression model is discussed in Section 4. A simulation study is conducted in Section 5 to compare the performance of MLE and Bayesian estimation. Finally, two real data sets are analyzed in Section 6 to illustrate the practicability of the proposed method. Our conclusions are presented in the final section. The proofs of lemmas and theorems are given in the appendix.

2. MAXIMUM LIKELIHOOD ESTIMATION

Given a random sample $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$ of size n from the ZOIP model (4), the likelihood function of (q_0, q_1, θ) is

$$\begin{aligned} L(q_0, q_1, \theta|\boldsymbol{Y}) \propto & q_0^{S_0} q_1^{S_1} (1-q_0-q_1)^{n-S_0-S_1} \\ & \frac{\theta^S}{(1-\mathrm{e}^{-\theta}-\theta\mathrm{e}^{-\theta})^{n-S_0-S_1}} \mathrm{e}^{-(n-S_0-S_1)\theta}, \end{aligned} \tag{5}$$

where $S_0 = S_0(\boldsymbol{Y}) = \sharp\{i : Y_i = 0\}, S_1 = S_1(\boldsymbol{Y}) = \sharp\{i : Y_i = 1\}, S = S(\boldsymbol{Y}) = \sum_{Y_i \geq 2} Y_i$. Here $\sharp X$ is defined to be the number of elements of the set X.

Note that under model (4), the mean and variance of Y are given by

$$\mathrm{E}(Y) = q_1 + \frac{(1-q_0-q_1)\theta}{1-\mathrm{e}^{-\theta}-\theta\mathrm{e}^{-\theta}}\left(1-\mathrm{e}^{-\theta}\right), \tag{6}$$

$$\begin{aligned} \mathrm{Var}(Y) = & q_1 + \frac{(1-q_0-q_1)\theta}{1-\mathrm{e}^{-\theta}-\theta\mathrm{e}^{-\theta}}\left(1+\theta-\mathrm{e}^{-\theta}\right) \\ & - \left(q_1 + \frac{(1-q_0-q_1)\theta}{1-\mathrm{e}^{-\theta}-\theta\mathrm{e}^{-\theta}}\left(1-\mathrm{e}^{-\theta}\right)\right)^2. \end{aligned} \tag{7}$$

According to the likelihood function (5), the maximum likelihood estimates of q_0 and q_1 are

$$\hat{q}_i = \frac{S_i}{n}, i = 0, 1 \tag{8}$$

and the MLE of θ, $\hat{\theta}$, is the solution of the following equation:

$$S\left(\mathrm{e}^{\theta}-\theta-1\right) - (n-S_0-S_1)\,\theta\left(\mathrm{e}^{\theta}-1\right) = 0, \tag{9}$$

which can be solved numerically according to the Newton-Raphson iterative algorithm. A sufficient condition for the existence and uniqueness of $\hat{\theta}$ is given in the following theorem.

Theorem 2.1. *If at least one observation is larger than one, i.e., $n - S_0 - S_1 > 0$, then there is a unique solution of θ for Equation (9).*

Parameters p_0 and p_1 in the ZOIP model (2) can be easily expressed by (q_0, q_1, θ). Based on the invariance property for the maximum likelihood estimation and the one-to-one transformation (3), we can obtain the MLEs of p_0 and p_1 as follows:

$$\hat{p}_0 = \frac{\hat{q}_0 + \hat{q}_1 - (1+\hat{\theta})\mathrm{e}^{-\hat{\theta}}}{1-(1+\hat{\theta})\mathrm{e}^{-\hat{\theta}}},$$

$$\hat{p}_1 = \frac{\hat{q}_0 - (1-\hat{p}_0)\mathrm{e}^{-\hat{\theta}}}{\hat{p}_0}.$$

3. BAYESIAN ESTIMATION

In this section, we give the Jeffereys prior and two reference priors of model (4) under the Bayesian framework.

3.1 Fisher information matrix

In this section, we provide a detailed derivation of the Fisher information matrix for the parameters (q_0, q_1, θ). Assuming that only one sample is observed, the corresponding likelihood function is

$$\begin{aligned} L_1(q_0, q_1, \theta|Y) \propto & q_0^{I\{Y=0\}} q_1^{I\{Y=1\}} \\ & (1-q_0-q_1)^{1-I\{Y=0\}-I\{Y=1\}} \\ & \frac{\theta^{Y-I\{Y=1\}}}{(\mathrm{e}^\theta-\theta-1)^{1-\mathrm{I}\{\mathrm{Y}=0\}-\mathrm{I}\{\mathrm{Y}=1\}}}, \end{aligned}$$

which indicate that the information matrix of (q_0, q_1) and θ is block diagonal. By calculation, we have

$$\begin{aligned} -E\left(\frac{\partial^2 \ln L}{\partial q_0^2}\right) &= \frac{1-q_1}{q_0(1-q_0-q_1)}, \\ -E\left(\frac{\partial^2 \ln L}{\partial q_0 \partial q_1}\right) &= \frac{1}{(1-q_0-q_1)}, \\ -E\left(\frac{\partial^2 \ln L}{\partial q_1^2}\right) &= \frac{1-q_0}{q_1(1-q_0-q_1)}, \\ -E\left(\frac{\partial^2 \ln L}{\partial \theta^2}\right) &= (1-q_0-q_1)k(\theta), \end{aligned}$$

where $k(\theta) = \frac{\mathrm{e}^{2\theta}-\theta^2\mathrm{e}^\theta-2\mathrm{e}^\theta+1}{\theta(\mathrm{e}^\theta-\theta-1)^2}$. Thus the Fisher information matrix of (q_0, q_1, θ) for one observation is

$$\boldsymbol{H_1}(q_0, q_1, \theta) = \begin{pmatrix} \boldsymbol{h_1} & \boldsymbol{0} \\ \boldsymbol{0} & h_2 \end{pmatrix}, \tag{10}$$

where

$$\boldsymbol{h_1} = \begin{pmatrix} \frac{1-q_1}{q_0(1-q_0-q_1)} & \frac{1}{(1-q_0-q_1)} \\ \frac{1}{(1-q_0-q_1)} & \frac{1-q_0}{q_1(1-q_0-q_1)} \end{pmatrix}$$

and

$$h_2 = (1-q_0-q_1)k(\theta).$$

3.2 Jeffreys prior

Lemma 3.1. *$k(\theta)$ is positive for $\theta > 0$.*

Jeffreys prior (Jeffreys, 1961) is proportional to the square root of the determinant of the Fisher information matrix. Accordingly, we can get the Jeffreys prior for (q_0, q_1) and θ

$$\pi_J(q_0, q_1, \theta) \propto q_0^{-1/2} q_1^{-1/2} k(\theta)^{1/2}, \tag{11}$$

with $\theta > 0$, $0 \le q_0 \le 1$ and $0 \le q_1 \le 1 - q_0$.

3.3 Reference prior

Jeffreys prior has been successfully applied to one-dimensional problems but can experience difficulties when multi-dimensional ones are considered. As mentioned by Berger, Bernado and Sun [4], "in multi-parameter models, reference priors typically depend on the parameter or quantity of interest, and it is well known that this is necessary to produce objective posterior distributions with optimal properties". The reference prior can be obtained according to the following algorithm, which was proposed by Berger and Bernado [3].

Here we take two-group parameters for an example to illustrate the algorithm of reference prior. Berger and Bernado [3] indicated that the parameters in the model are ordered in terms of importance in the inference. We assume that X is a random variable with density function $p(x|\boldsymbol{\eta})$, where $\boldsymbol{\eta} = (\boldsymbol{\eta_1}, \boldsymbol{\eta_2})$ denotes an unknown two-group parameter vector and $\boldsymbol{\eta_1}$ is the group parameter of interest. Denote the dimension of $\boldsymbol{\eta_1}$ and $\boldsymbol{\eta_2}$ by n_1 and n_2 respectively. Let

$$\boldsymbol{H} = \boldsymbol{H}(\boldsymbol{\eta}) = -E_{X|\boldsymbol{\eta}}\left[\frac{\partial^2 \log\, p(X|\boldsymbol{\eta})}{\partial \boldsymbol{\eta'} \partial \boldsymbol{\eta}}\right]$$

be the Fisher information matrix for $p(X|\boldsymbol{\eta})$. Suppose $\boldsymbol{H}$ is invertible and define $\boldsymbol{S} = \boldsymbol{S}(\boldsymbol{\eta}) = \boldsymbol{H^{-1}}(\boldsymbol{\eta})$. Let $N_1 = n_1$, $N_2 = n_1 + n_2$. Denote by $\boldsymbol{S_j}(\boldsymbol{\eta})$ the upper left $N_j \times N_j$ matrix of $\boldsymbol{S}(\boldsymbol{\eta})$, with $\boldsymbol{S_2}(\boldsymbol{\eta}) \equiv \boldsymbol{S}(\boldsymbol{\eta})$, and $\boldsymbol{H_j}(\boldsymbol{\eta}) \equiv \boldsymbol{S_j^{-1}}(\boldsymbol{\eta})$; the matrix $\boldsymbol{h_j}(\boldsymbol{\eta})$ is defined as the low right $n_j \times n_j$ corner of $\boldsymbol{H_j}(\boldsymbol{\eta})$. Then for two groups of parameters (Berger & Bernardo [3]) the algorithm can be described as follows:

1. Choose a nested sequence of compact subsets of Θ^l, such that $\cup_{l=1}^{\infty}\Theta^l = \Theta$. $\Theta^l(\boldsymbol{\eta_1}) = \{\boldsymbol{\eta_1} : (\boldsymbol{\eta_1}, \boldsymbol{\eta_2}) \in \Theta^l\}$, $\Theta^l(\boldsymbol{\eta_2}) = \{\boldsymbol{\eta_2} : (\boldsymbol{\eta_1}, \boldsymbol{\eta_2}) \in \Theta^l\}$.
2. For each l, let
$$\pi_2^l(\boldsymbol{\eta_2}|\boldsymbol{\eta_1}) = \frac{|\boldsymbol{h_2}(\boldsymbol{\eta})|^{1/2} I_{\Theta^l(\boldsymbol{\eta_2})}(\boldsymbol{\eta_2})}{\int_{\Theta^l(\boldsymbol{\eta_2})} |\boldsymbol{h_2}(\boldsymbol{\eta})|^{1/2} \mathrm{d}\boldsymbol{\eta_2}}.$$
3. Find
$$\pi_1^l(\boldsymbol{\eta}) = \frac{\pi_2^l(\boldsymbol{\eta_2}|\boldsymbol{\eta_1}) \exp\{\frac{1}{2}E^l[\log|\boldsymbol{h_1}(\boldsymbol{\eta})||\boldsymbol{\eta_1}]\} I_{\Theta^l(\boldsymbol{\eta_1})}(\boldsymbol{\eta_1})}{\int_{\Theta^l(\boldsymbol{\eta_1})} \exp\{\frac{1}{2}E^l[\log|\boldsymbol{h_1}(\boldsymbol{\eta})||\boldsymbol{\eta_1}]\} \mathrm{d}\boldsymbol{\eta_1}}, \tag{12}$$
where
$$E^l[g(\boldsymbol{\eta})|\boldsymbol{\eta_1}] = \int_{\Theta^l(\boldsymbol{\eta_2})} g(\boldsymbol{\eta}) \pi_2^l(\boldsymbol{\eta_2}|\boldsymbol{\eta_1}) \mathrm{d}\boldsymbol{\eta_2},$$
and
$$I_{\Omega(x)} = \begin{cases} 1, & x \in \Omega; \\ 0, & \text{otherwise}. \end{cases} \tag{13}$$
It is easy to check that $\{\pi_j^l, j = 1, 2\}$ defines a probability distribution. Define the 2-group reference prior, assuming it yields a proper posterior, by
$$\pi(\boldsymbol{\eta}) = \lim_{l\to\infty} \frac{\pi_1^l(\boldsymbol{\eta})}{\pi_1^l(\boldsymbol{\eta^*})}$$
where $\boldsymbol{\eta^*}$ is some point in Θ^1, which is the first nested compact set.

The calculation of the 2-group reference prior is greatly simplified under the condition

$$|\boldsymbol{h_1}(\boldsymbol{\eta})| \text{ depends only on } \boldsymbol{\eta_1}. \tag{14}$$

Lemma 3.2. *If (14) holds, then*

$$\pi^l(\boldsymbol{\eta}) = \left(\prod_{i=1}^{2} \frac{|\boldsymbol{h_i}(\boldsymbol{\eta})|^{\frac{1}{2}}}{\int |\boldsymbol{h_i}(\boldsymbol{\eta})|^{\frac{1}{2}} \mathrm{d}\boldsymbol{\eta_i}}\right) \boldsymbol{I}_{\Theta^l}(\boldsymbol{\eta}), \tag{15}$$

where the integral is over the range $\Theta^l(\boldsymbol{\eta_i})$.

Here we use this lemma to obtain the reference prior when (q_0, q_1) is the parameter of interest.

Theorem 3.1. *The reference prior when* (q_0, q_1) *or* θ *is the parameter of interest is given by*

$$\pi_R(q_0, q_1, \theta) \propto q_0^{-1/2} q_1^{-1/2} (1 - q_0 - q_1)^{-1/2} k(\theta)^{1/2}, \tag{16}$$

with $\theta > 0$, $0 \le q_0 \le 1$ *and* $0 \le q_1 \le 1 - q_0$.

A probability matching prior is a prior distribution which describes the posterior confidence or credible regions with exact or approximate frequent validity. Let $\hat{\vartheta}_n(\alpha)$ denote the posterior lower α-quantile of a parameter ϑ based on n observations $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$. A prior π is called an ith order matching prior for ϑ, if it satisfies

$$P_\vartheta(\vartheta \le \hat{\vartheta}_n(\alpha)) = \alpha + O(n^{-i/2}).$$

Here the left-hand side is the frequentist probability of $\boldsymbol{Y}$ satisfying $\vartheta \le \hat{\vartheta}_n(\alpha)$ given ϑ.

Lemma 3.3. *Suppose* k_{ij} *be the* (i, j) *element of* $[\mathbf{H}(\boldsymbol{\eta})]^{-1}$, *where* $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)$ *and* $[\mathbf{H}(\boldsymbol{\eta})]$ *is the Fisher information matrix of* $\boldsymbol{\eta}$. *If* η_1 *and* $(\eta_2, \ldots, \eta_k)$ *are orthogonal (Cox & Reid, 1987), then* $\pi(\boldsymbol{\eta})$ *is a second-order matching prior for* η_1 *iff* $\pi(\boldsymbol{\eta})$ *satisfy*

$$\pi(\boldsymbol{\eta}) \propto k_{11}(\boldsymbol{\eta})^{-1/2} K(\eta_2, \ldots, \eta_k).$$

This Lemma comes from Peers [25].

Theorem 3.2. *The Jeffreys prior* (11) *and the reference prior* (16) *are second-order matching prior for* θ.

Theorem 3.3. *When* θ *is the parameter of interest, the reference prior for* (θ, p_0, p_1) *is given by*

(17)

$$\begin{aligned}\pi_R(\theta, p_0, p_1) \propto& \left[p_0 p_1 \mathrm{e}^\theta + (1 - \mathrm{p}_0)\right]^{-1/2} \\ & \left[p_0(1 - p_1)\mathrm{e}^\theta + (1 - \mathrm{p}_0)\theta\right]^{-1/2} (1 - p_0)^{-1/2} p_0 \\ & \left(\mathrm{e}^\theta - \theta - 1\right) k(\theta)^{1/2},\end{aligned}$$

with $0 \le p_0 \le 1$, $0 \le p_1 \le 1$, *and* $\theta > 0$.

Note that when p_0 and p_1 are the parameters of interest, the condition of Theorem 1 in Yang [30] is not satisfied. Then the reference prior for (p_0, p_1, θ) can not be calculated by this method.

3.4 Posterior properties

The posterior distributions of (q_0, q_1, θ) using the Jeffreys prior and the reference prior are respectively

$$\begin{aligned}\pi_J(q_0, q_1, \theta|\boldsymbol{Y}) \propto& q_0^{S_0 - \frac{1}{2}} q_1^{S_1 - \frac{1}{2}} (1 - q_0 - q_1)^{n - S_0 - S_1} \\ & \frac{\theta^S}{(\mathrm{e}^\theta - \theta - 1)^{n - S_0 - S_1}} k^{\frac{1}{2}}(\theta),\end{aligned}$$

and

$$\begin{aligned}\pi_R(q_0, q_1, \theta|\boldsymbol{Y}) \propto& q_0^{S_0 - \frac{1}{2}} q_1^{S_1 - \frac{1}{2}} (1 - q_0 - q_1)^{n - S_0 - S_1 - \frac{1}{2}} \\ & \frac{\theta^S}{(\mathrm{e}^\theta - \theta - 1)^{n - S_0 - S_1}} k^{\frac{1}{2}}(\theta).\end{aligned}$$

Now we show the posterior distributions are proper under a relatively mild condition.

Theorem 3.4. *The posterior distribution of* (q_0, q_1, θ) *using either the Jeffreys prior or the reference prior is proper when* $S \ge 2$, *and improper when* $S < 2$.

Theorem 3.5. *When* θ *is the parameter of interest, the posterior distribution of* (θ, p_0, p_1) *using the reference prior is proper when* $S \ge 2$, *and improper when* $S < 2$.

3.5 Posterior sampling

The marginal distributions of (q_0, q_1) using the Jeffreys prior and the reference prior are respectively

$$\pi_J(q_0, q_1|\boldsymbol{Y}) \propto q_0^{S_0 - \frac{1}{2}} q_1^{S_1 - \frac{1}{2}} (1 - q_0 - q_1)^{n - S_0 - S_1},$$

and

$$\pi_R(q_0, q_1|\boldsymbol{Y}) \propto q_0^{S_0 - \frac{1}{2}} q_1^{S_1 - \frac{1}{2}} (1 - q_0 - q_1)^{n - S_0 - S_1 - \frac{1}{2}}.$$

They are the Dirichlet distributions with shape parameters $(S_0 + \frac{1}{2}, S_1 + \frac{1}{2}, n - S_0 - S_1 + 1)$ and $(S_0 + \frac{1}{2}, S_1 + \frac{1}{2}, n - S_0 - S_1 + \frac{1}{2})$ respectively. They can be easily sampled using the $\mathbf{rdirichlet(N, alpha)}$ function in the R package gtools.

The marginal distribution of θ using the Jeffreys prior and the reference prior is identical. And the distribution is

$$\pi(\theta|\boldsymbol{Y}) \propto \frac{\theta^S}{(\mathrm{e}^\theta - \theta - 1)^{n - S_0 - S_1}} k^{\frac{1}{2}}(\theta).$$

The Bayesian inference of the ZOIP model (4) about θ can be performed using the Metropolis-Hastings sampling procedure below.

1. Set initial value for $\theta^{(0)} > 0$.
2. For $t = 1, 2, \ldots$,
   (a) Set $\theta = \theta^{(t-1)}$.
   (b) Propose a new value $\theta^{'}$ from $N(\theta, \sigma^2)$ and set $\theta^{'} = |\theta^{'}|$, where σ is a tuning parameter.
   (c) Calculate $\log\alpha = \min(0, A)$ with $A = \log\frac{\pi(\theta^{'}|Y)}{\pi(\theta|Y)}$.
   (d) Set $\theta^{(t)} = \theta^{'}$ with probability α and $\theta^{(t)} = \theta$ with the remaining probability.

From the posterior sample of (θ, q_0, q_1), the posterior sample of (θ, p_0, p_1) can be obtained though the transformation (23).

4. ZERO-AND-ONE-INFLATED POISSON REGRESSION MODEL

In this section we consider ZOIP model (4) with covariates. The covariates are usually linked to model parameters q_0, q_1 and θ. Denote $\boldsymbol{q_0} = (q_{01}, \dots, q_{0n})$, $\boldsymbol{q_1} = (q_{11}, \dots, q_{1n})$ and $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$. Assume that the independent responses Y_i are sampled from ZOIP$(q_{0i}, q_{1i}, \theta_i)$ and the parameters $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$, $\boldsymbol{q_0} = (q_{01}, \dots, q_{0n})$, and $\boldsymbol{q_1} = (q_{11}, \dots, q_{1n})$ are linked to the covariates $\boldsymbol{Z} = \boldsymbol{(z_1, \dots, z_n)^{\mathrm{T}}}$, $\boldsymbol{W_0} = \boldsymbol{(\omega_{01}, \dots, \omega_{0n})^{\mathrm{T}}}$ and $\boldsymbol{W_1} = \boldsymbol{(\omega_{11}, \dots, \omega_{1n})^{\mathrm{T}}}$ in forms like

$$
\begin{cases}
q_{0i} = \frac{\exp(\boldsymbol{\omega_{0i}^{\mathrm{T}}\gamma_0})}{1+\exp(\boldsymbol{\omega_{0i}^{\mathrm{T}}\gamma_0})+\exp(\boldsymbol{\omega_{1i}^{\mathrm{T}}\gamma_1})},\\
q_{1i} = \frac{\exp(\boldsymbol{\omega_{1i}^{\mathrm{T}}\gamma_1})}{1+\exp(\boldsymbol{\omega_{0i}^{\mathrm{T}}\gamma_0})+\exp(\boldsymbol{\omega_{1i}^{\mathrm{T}}\gamma_1})},\\
\theta_i = \exp(\boldsymbol{z_i^{\mathrm{T}}\beta}),
\end{cases}
\tag{18}
$$

where $\boldsymbol{\gamma_0}$, $\boldsymbol{\gamma_1}$ and $\boldsymbol{\beta}$ are vectors of regression parameters; $\boldsymbol{z_1, \dots, z_n}$ are the covariates of the same length as $\boldsymbol{\beta}$; $\boldsymbol{\omega_{j1}, \dots, \omega_{jn}}$ are the covariates of the same length as $\boldsymbol{\gamma_j}$, $j = 0, 1$, and the first element of $\boldsymbol{\omega_{0i}, \omega_{1i}}$ and $\boldsymbol{z_i}$ being 1 corresponding to the intercept. We denote this zero-and-one-inflated Poisson regression model as ZOIP regression model.

4.1 Maximum likelihood estimation

Let $\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_n)$ be a sample from the ZOIP regression model (18). The log-likelihood function under regression case is

$$
\begin{aligned}
\ell(\boldsymbol{\gamma_0, \gamma_1, \beta | Y}) = &\sum_{i=1}^{n} I\{Y_i = 0\} \ln q_{0i} + \sum_{i=1}^{n} I\{Y_i = 1\} \ln q_{1i} \\
&+ \sum_{i=1}^{n} I\{Y_i \geq 2\} \\
&\ln\left[\frac{1 - q_{0i} - q_{1i}}{1 - \mathrm{e}^{-\theta_i} - \theta_i \mathrm{e}^{-\theta_i}} \frac{\theta_i^{Y_i}\mathrm{e}^{-\theta_i}}{Y_i!}\right],
\end{aligned}
\tag{19}
$$

where q_{0i}, q_{1i} and θ_i are given by (18). We use the Newton-Raphson iteration algorithm to get the MLEs of $\boldsymbol{\gamma_0, \gamma_1}$ and $\boldsymbol{\beta}$. Firstly set the initial values of $\boldsymbol{\gamma_0^{(0)}, \gamma_1^{(0)}}$ and $\boldsymbol{\beta^{(0)}}$. Then given the the values of $\boldsymbol{\gamma_0^{(k)}, \gamma_1^{(k)}}$ and $\boldsymbol{\beta^{(k)}}$ the $(k+1)$-th iteration of the Newton-Raphson iteration algorithm is

$$
\begin{aligned}
\hat{\boldsymbol{\gamma_0}}^{\boldsymbol{(k+1)}} = &\hat{\boldsymbol{\gamma_0}}^{\boldsymbol{(k)}} + \left[\sum_{i=1}^{n} q_{0i}^{(k)}(1 - q_{0i}^{(k)})\boldsymbol{\omega_{0i}\omega_{0i}^{\mathrm{T}}}\right]^{-1} \\
&\sum_{i=1}^{n}\left[\left(I\{Y_i = 0\} - q_{0i}^{(k)}\right)\boldsymbol{\omega_{0i}}\right],
\end{aligned}
$$

$$
\begin{aligned}
\hat{\boldsymbol{\gamma_1}}^{\boldsymbol{(k+1)}} = &\hat{\boldsymbol{\gamma_1}}^{\boldsymbol{(k)}} + \left[\sum_{i=1}^{n} q_{1i}^{(k)}(1 - q_{1i}^{(k)})\boldsymbol{\omega_{1i}\omega_{1i}^{\mathrm{T}}}\right]^{-1} \\
&\sum_{i=1}^{n}\left[\left(I\{Y_i = 1\} - q_{1i}^{(k)}\right)\boldsymbol{\omega_{1i}}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
&\hat{\boldsymbol{\beta}}^{\boldsymbol{(k+1)}} \\
&= \hat{\boldsymbol{\beta}}^{\boldsymbol{(k)}} - \left\{\left[\frac{\partial^2 \ell(\boldsymbol{\gamma_0, \gamma_1, \beta | Y})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^{\mathrm{T}}}\right]^{-1} \frac{\partial \ell(\boldsymbol{\gamma_0, \gamma_1, \beta | Y})}{\partial\boldsymbol{\beta}}\right\}_{\hat{\boldsymbol{\beta}}^{\boldsymbol{(k)}}},
\end{aligned}
$$

where

$$
\frac{\partial \ell(\boldsymbol{\gamma_0, \gamma_1, \beta | Y})}{\partial\boldsymbol{\beta}} = \sum_{i=1}^{n} I\{Y_i \geq 2\}\left[Y_i - \frac{\theta_i^{(k)}(\mathrm{e}^{\theta_i^{(k)}} - 1)}{\mathrm{e}^{\theta_i^{(k)}} - \theta_i^{(k)} - 1}\right]\boldsymbol{z_i}
$$

and

$$
\begin{aligned}
&\frac{\partial^2 \ell(\boldsymbol{\gamma_0, \gamma_1, \beta | Y})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^{\mathrm{T}}} \\
&= -\sum_{i=1}^{n} I\{Y_i \geq 2\} \\
&\times \left[\frac{\mathrm{e}^{2\theta_i^{(k)}} - \theta_i^{(k)^2}\mathrm{e}^{\theta_i^{(k)}} - 2\mathrm{e}^{\theta_i^{(k)}} + 1}{(\mathrm{e}^{\theta_i^{(k)}} - \theta_i^{(k)} - 1)^2}\right]\theta_i^{(k)}\boldsymbol{z_i z_i^{\mathrm{T}}}.
\end{aligned}
$$

We continue this procedure until convergence.

4.2 Bayesian inference

In this paper, we choose normal prior for the Bayesian inference. We assume that

$$
\boldsymbol{\beta} \sim N_q(\boldsymbol{\beta_0}, \sigma_\beta^2 \boldsymbol{I_q}) \quad \text{and} \quad \boldsymbol{\gamma_i} \sim N_{r_i}(\boldsymbol{\gamma_{0i}}, \sigma_{\gamma_i}^2 \boldsymbol{I_{r_i}}), (i = 0, 1)
$$

where $\boldsymbol{\beta_0}, \boldsymbol{\gamma_{0i}}, \sigma_\beta^2, \sigma_{\gamma_i}^2$ are known constants. It is further assumed that the parameters $\boldsymbol{\gamma_0}$, $\boldsymbol{\gamma_1}$ and $\boldsymbol{\beta}$ are mutually independent. And denote $\Pi(\boldsymbol{\gamma_0, \gamma_1, \beta})$ as the joint prior of the parameters $\boldsymbol{\gamma_0, \gamma_1}$ and $\boldsymbol{\beta}$. The posterior distribution being proper is easy to prove in that the function about $(\boldsymbol{\gamma_0, \gamma_1, \beta})$ in the likelihood function is bounded.

For the regression case, the posterior distribution for $\boldsymbol{\gamma_0, \gamma_1}$ and $\boldsymbol{\beta}$ has a nonstandard density with a complicated expression. MCMC method is used to sample from the posterior distribution. In particular, the Gibbs sampling method has been used to obtain a large number of random variates from the posterior distribution. Any distributional summary (such as mean, median or quantiles) of the posterior distribution can then be approximated by their corresponding sample analogue.

Let $\boldsymbol{Y} = (Y_1, Y_2, \dots, Y_n)$ be a sample from the ZOIP regression model (18). Then the posterior density of $(\boldsymbol{\gamma_0, \gamma_1, \beta})$ given $\boldsymbol{Y}$ is

$$\pi(\boldsymbol{\beta},\boldsymbol{\gamma_0},\boldsymbol{\gamma_1}|\boldsymbol{Y}) \propto \prod_{i=1}^{n} q_{0i}^{I\{Y_i=0\}} q_{1i}^{I\{Y_i=1\}} \left[\frac{1-q_{0i}-q_{1i}}{1-\mathrm{e}^{-\theta_{\mathrm{i}}}-\theta_{\mathrm{i}}\mathrm{e}^{-\theta_{\mathrm{i}}}} \frac{\theta_i^{Y_i}\mathrm{e}^{-\theta_{\mathrm{i}}}}{\mathrm{Y}_i!}\right]^{I\{Y_i\geq 2\}} \times \Pi(\boldsymbol{\beta},\boldsymbol{\gamma_0},\boldsymbol{\gamma_1}). \tag{20}$$

In order to implement the Gibbs sampling algorithm, the full conditional distributions of $\pi[\boldsymbol{\gamma_0}|\boldsymbol{\gamma_1},\boldsymbol{\beta},\boldsymbol{Y}]$, $\pi[\boldsymbol{\gamma_1}|\boldsymbol{\gamma_0},\boldsymbol{\beta},\boldsymbol{Y}]$ and $\pi[\boldsymbol{\beta}|\boldsymbol{\gamma_0},\boldsymbol{\gamma_1},\boldsymbol{Y}]$ are needed. The full conditional distributions of $\boldsymbol{\gamma_0}$, $\boldsymbol{\gamma_1}$ and $\boldsymbol{\beta}$ are not standard distributions. It is easy to see that these conditional densities are log-concave. So the adaptive rejection sampling(ARS) (see Gilks and Wild [12]) can be used to sample $\boldsymbol{\gamma_0}$, $\boldsymbol{\gamma_1}$ and $\boldsymbol{\beta}$ from their respective full conditional distributions.

Thus the Bayesian inference of the ZOIP regression model (18) can be performed using the Gibbs sampling procedure below.

1. Set initial values for $\boldsymbol{\gamma_0}^{(0)}, \boldsymbol{\gamma_1}^{(0)}$ and $\boldsymbol{\beta}^{(0)}$.
2. For $t = 1, 2, \ldots$, perform the following iterative update.
   (a) Sample $\boldsymbol{\gamma}_0^{(t)}$ using ARS, given the sampled values of $\boldsymbol{\gamma}_2^{(t-1)}$, $\boldsymbol{\beta}^{(t-1)}$ and $\boldsymbol{Y}$.
   (b) Sample $\boldsymbol{\gamma}_1^{(t)}$ using ARS, given the sampled values of $\boldsymbol{\gamma}_0^{(t)}$, $\boldsymbol{\beta}^{(t-1)}$ and $\boldsymbol{Y}$.
   (c) Sample $\boldsymbol{\beta}^{(t)}$ using ARS, given the sampled values of $\boldsymbol{\gamma}_0^{(t)}$, $\boldsymbol{\gamma}_1^{(t)}$ and $\boldsymbol{Y}$.

5. SIMULATION STUDY

In this section we will assess the performance of model (4). The sample size was set to n = 20, 30, 50, and 100, the value of q_0 was set to 0.15, 0.2, the value of q_1 was set to 0.3, 0.4, the value of θ was set to 3, 5, and 8, the confidence level α was set to 95% and all simulations are replicated for 10,000 times. When we do simulation, according to ZOIP model (4), the sample of size n is reserved only when $n - S_0 - S_1 > 0$, considering the existence of MLE of θ. Here the equal tail two-sided confidence intervals is used. The comparison results for the root mean squared error and the coverage probabilities are recorded in Tables 1 and 2. In the tables, the subscript M represents the maximum likelihood estimation. The subscripts J and R represent Bayesian estimation with the Jeffreys prior and the reference prior respectively.

For the point estimate of θ, the MLE performs slightly better than the two Bayesian estimates when θ is small but they are similar when θ is large. For the point estimates of q_0 and q_1 the MLE performs slightly worse than the Bayesian estimates when n is small and the MLE and Bayesian estimate performs similarly when n is large. For the point estimates and interval estimates of q_0 and q_1 the Jeffreys prior and the reference prior perform similarly. For the interval estimates of q_0 and q_1 the coverage probabilities based on the Bayesian estimates are generally more accurate than that based on the MLE. As the sample size increases, the accuracy of all the estimates increases, and the accuracy differences among the estimates becomes smaller.

To compare the performance between Bayesian estimation and MLE when covariates are considered, a simulation study was carried out. The sample size was set to $n = 50$, and 100. Throughout, there are three covariates matrix ($\boldsymbol{Z}$, $\boldsymbol{W_0}$, $\boldsymbol{W_1}$) and matrix generation method is the same. The covariates matrix is the sum of a fixed matrix generated by $(-1, 0, 1)$ and a random matrix which comes from $N(0, 2)$. And we standardized these matrix by column. Set the real $\boldsymbol{\beta} = (1.5, -2)$, $\boldsymbol{\gamma_0} = (1, -2)$ and $\boldsymbol{\gamma_1} = (1, -1)$. For prior distribution, we assume $\boldsymbol{\beta_0} = \boldsymbol{\gamma_{00}} = \boldsymbol{\gamma_{01}} = (0, 0)$ and $\sigma_\beta^2 = \sigma_{\gamma_0}^2 = \sigma_{\gamma_1}^2 = 1000$. Each simulation is replicated for 2,000 times. The simulation results are listed in Table 3. As the sample size increases, RMSE of all the estimates decreases, and the difference among them becomes smaller. And the MLE is slightly better than Bayesian estimate.

6. REAL DATA ANALYSIS

6.1 Singapore Legionnaires' disease data

In this subsection, one example about Legionnaires' disease in Singapore from the healthcare industry is presented to illustrate our method and this data set was analyzed by Xu et al. [29]. Legionellosis (Legionnaires' disease and Pontiac fever) is an acute respiratory infection caused by gram negative, rod-shaped bacteria of the genus Legionella (Lam et al. [17]). In Singapore, Legionnaires' disease has been recognized as a potential public health threat. In order to make relevant control polices, it is useful to know the distribution of the counts of Legionellosis cases. For illustration, here we apply our model in the study of the weekly Legionellosis count data in the year 2005. The data were reported by the Ministry of Health of Singapore. Xu et al. [29] derived the Jeffreys prior and reference prior for the ZIP model and presented the Bayesian fitted frequencies and compared with likelihood method for both the ZIP and pure Poisson models. See Table 4 for their detailed results. The estimation results of ZOIP model (4) is presented in Table 5. In these tables, PE represents the point estimation and CI represents the confidence interval.

As is noted by Xu et al. [29], the difference between the estimation accuracy of the ZIP models and the Poisson model is not clear according to the fitted frequency distributions. According to the fitted frequency shown in Table 4, the frequency of one is underestimated overall in Xu et al. [29] (the estimated value is nearly half of the observed frequency). In our result, both the MLE and Bayes estimation for all frequencies are closer to the observed values. Besides, the

Table 1. RMSE of parameter estimation for model (4)

θ	q_0	q_1	n	θ_M	θ_J	θ_R	q_{0M}	q_{0J}	q_{0R}	q_{1M}	q_{1J}	q_{1R}
3	0.15	0.3	20	0.6484	0.7611	0.7591	0.0798	0.0731	0.0726	0.1029	0.0917	0.0914
			30	0.5013	0.5736	0.5649	0.0649	0.0680	0.0678	0.0836	0.0780	0.0806
			50	0.3841	0.5431	0.5586	0.0502	0.0514	0.0494	0.0620	0.0633	0.0609
			100	0.2861	0.4632	0.4522	0.0360	0.0349	0.0359	0.0452	0.0462	0.0457
		0.4	20	0.6793	1.0253	0.9865	0.0777	0.0733	0.0749	0.1081	0.1005	0.0995
			30	0.5754	0.7150	0.6821	0.0667	0.0623	0.0629	0.0879	0.0816	0.0867
			50	0.4381	0.4719	0.5023	0.0495	0.0490	0.0513	0.0693	0.0670	0.0658
			100	0.3047	0.4365	0.4535	0.0347	0.0350	0.0356	0.0497	0.0476	0.0489
	0.2	0.3	20	0.6836	0.7600	0.7951	0.0915	0.0801	0.0831	0.1059	0.0916	0.0925
			30	0.5665	0.6133	0.5894	0.0720	0.0668	0.0679	0.0802	0.0779	0.0776
			50	0.4096	0.4326	0.4399	0.0580	0.0537	0.0525	0.0654	0.0633	0.0611
			100	0.2873	0.4192	0.3972	0.0397	0.0395	0.0406	0.0457	0.0459	0.0467
		0.4	20	0.6888	0.8132	0.8418	0.1013	0.0834	0.0849	0.1289	0.1014	0.1010
			30	0.6179	0.7095	0.6939	0.0727	0.0919	0.0711	0.0900	0.0805	0.0853
			50	0.4785	0.5046	0.5259	0.0577	0.0411	0.0393	0.0672	0.0480	0.0492
			100	0.3283	0.4297	0.4401	0.0412	0.0393	0.0390	0.0486	0.0477	0.0482
5	0.15	0.3	20	0.7307	0.7811	0.7692	0.0787	0.0731	0.0764	0.0951	0.0888	0.0879
			30	0.6021	0.6428	0.6210	0.0670	0.0674	0.0717	0.0841	0.0789	08044
			50	0.4556	0.5308	0.5609	0.0508	0.0464	0.0478	0.0661	0.0624	0.0643
			100	0.3280	0.4727	0.4576	0.0356	0.0360	0.0357	0.0484	0.0449	0.0454
		0.4	20	0.8475	0.8321	0.8692	0.0795	0.0728	0.0760	0.1086	0.0968	0.1033
			30	0.6784	0.6259	0.6376	0.0678	0.0620	0.0628	0.0900	0.0845	0.0874
			50	0.4933	0.5414	0.5168	0.0501	0.0482	0.0512	0.0701	0.0693	0.0673
			100	0.3572	0.4165	0.3901	0.0366	0.0336	0.0360	0.0493	0.0484	0.0478
	0.2	0.3	20	0.7743	0.7931	0.8125	0.0868	0.0825	0.0852	0.1034	0.0930	0.0989
			30	0.6448	0.6374	0.6548	0.0703	0.0664	0.0689	0.0818	0.0776	0.0807
			50	0.4933	0.5250	0.4833	0.0571	0.0551	0.0556	0.0651	0.0635	0.0631
			100	0.3418	0.4149	0.4305	0.0396	0.0400	0.0411	0.0456	0.0446	0.0437
		0.4	20	0.8430	0.8655	0.8489	0.0838	0.0827	0.0873	0.1016	0.0998	0.1050
			30	0.7411	0.7185	0.6812	0.0719	0.0714	0.0715	0.0903	0.0776	0.0807
			50	0.5663	0.5887	0.5513	0.0571	0.0392	0.0393	0.0666	0.0478	0.0497
			100	0.3711	0.4050	0.3876	0.0383	0.0372	0.0383	0.0496	0.0477	0.0456
8	0.15	0.3	20	0.8810	0.9192	0.8827	0.0785	0.0764	0.0771	0.1033	0.0953	0.0988
			30	0.6966	0.6863	0.6671	0.0636	0.0678	0.0725	0.0809	0.0801	0.0768
			50	0.5739	0.6219	0.6352	0.0506	0.0481	0.0485	0.0650	0.0578	0.0607
			100	0.3867	0.4173	0.4409	0.0364	0.0351	0.0363	0.0460	0.0477	0.0447
		0.4	20	1.0159	0.9315	1.0989	0.0815	0.0716	0.0765	0.1117	0.0998	0.1006
			30	0.7900	0.7604	0.7740	0.0650	0.0615	0.0643	0.0909	0.0821	0.0864
			50	0.6125	0.6268	0.6218	0.0492	0.0497	0.0506	0.0686	0.0548	0.0671
			100	0.4206	0.4251	0.4375	0.0359	0.0346	0.0356	0.0489	0.0487	0.0494
	0.2	0.3	20	0.9317	0.9363	0.9225	0.0887	0.0796	0.0793	0.1435	0.1035	0.0961
			30	0.7697	0.7881	0.7606	0.0716	0.0693	0.0689	0.0825	0.0767	0.0811
			50	0.5771	0.6419	0.6321	0.0551	0.0561	0.0575	0.0642	0.0621	0.0642
			100	0.4114	0.4288	0.4566	0.0403	0.0398	0.0383	0.0447	0.0442	0.0445
		0.4	20	1.0770	0.9655	1.0424	0.0869	0.0819	0.0821	0.1368	0.0984	0.1042
			30	0.8436	0.8287	0.8322	0.0686	0.0663	0.0663	0.0878	0.0814	0.0810
			50	0.6408	0.6278	0.6439	0.0570	0.0553	0.0560	0.0705	0.0663	0.0707
			100	0.4529	0.4731	0.4793	0.0384	0.0383	0.0404	0.0480	0.0481	0.0494

estimation of parameter θ by ZOIP is nearly twice the estimation by ZIP. And the results of AIC (Akaike Information Criterion) is presented in Table 6. The AIC value of our results is smaller than the value of ZIP model. The results of DIC (Deviance Information Criterion) and WAIC (Watanabe-Akaike Information Criterion) are presented in Table 7. The criterion used by Gelman et al. [10] of DIC, WAIC1 and WAIC2 is used in our article. The results show that ZOIP model is more appropriate than ZIP model.

6.2 US Detroit accidental death data

In this section, one accidental data set from Detroit, Michigan is used to demonstrate the zero-and-one-inflated Poisson model introduced in the previous section. Accidental deaths have accounted for a large proportion of deaths in the event of death. In order to make relevant control polices, it is necesssary to know the distribution of the counts of accidental deaths. The NMMAPS data which is available in R (NMMAPSlite package) contains daily mortality, air

Table 2. Coverage probabilities of confidence intervals for model (4)

θ	q_0	q_1	n	θ_M	θ_J	θ_R	q_{0M}	q_{0J}	q_{0R}	q_{1M}	q_{1J}	q_{1R}
3	0.15	0.3	20	0.9321	0.9285	0.9197	0.8212	0.9451	0.9373	0.8913	0.9465	0.9417
			30	0.9392	0.9311	0.9291	0.9172	0.9243	0.9372	0.9483	0.9542	0.9444
			50	0.9472	0.9453	0.9529	0.9422	0.9276	0.9401	0.9445	0.9322	0.9312
			100	0.9479	0.9514	0.9538	0.9311	0.9513	0.9571	0.9486	0.9627	0.9584
		0.4	20	0.9361	0.9012	0.8993	0.8643	0.9414	0.9392	0.9177	0.9377	0.9517
			30	0.9270	0.9219	0.9287	0.9380	0.9202	0.9217	0.9439	0.9365	0.9313
			50	0.9516	0.9330	0.9409	0.9470	0.9401	0.9443	0.9404	0.9466	0.9472
			100	0.9488	0.9501	0.9528	0.9380	0.9475	0.9476	0.9455	0.9564	0.9486
	0.2	0.3	20	0.9247	0.8723	0.8909	0.8384	0.9473	0.9455	0.9117	0.9443	0.9544
			30	0.9215	0.9336	0.9358	0.9303	0.9292	0.9224	0.9323	0.9492	0.9523
			50	0.9493	0.9410	0.9497	0.9253	0.9661	0.9381	0.9273	0.9592	0.9501
			100	0.9495	0.9519	0.9483	0.9318	0.9462	0.9538	0.9524	0.9353	0.9464
		0.4	30	0.9243	0.8871	0.8902	0.8427	0.9514	0.9600	0.9366	0.9247	0.9535
			30	0.9372	0.9268	0.9205	0.9276	0.9206	0.9310	0.9215	0.9258	0.9483
			50	0.9294	0.9337	0.9411	0.9395	0.9537	0.9526	0.9504	0.9594	0.9525
			100	0.9495	0.9499	0.9520	0.9296	0.9461	0.9433	0.9503	0.9512	0.9546
5	0.15	0.3	20	0.9432	0.9480	0.9398	0.8164	0.9252	0.9280	0.9275	0.9437	0.9503
			30	0.9394	0.9385	0.9317	0.9244	0.9192	0.9161	0.9493	0.9611	0.9453
			50	0.9493	0.9411	0.9507	0.9338	0.9281	0.9424	0.9213	0.9356	0.9497
			100	0.9410	0.9490	0.9513	0.9362	0.9512	0.9535	0.9383	0.9600	0.9489
		0.4	20	0.9394	0.9319	0.9275	0.8235	0.9332	0.9370	0.9323	0.9185	0.9495
			30	0.9420	0.9451	0.9489	0.9318	0.9172	0.9313	0.9353	0.9384	0.9445
			50	0.9574	0.9406	0.9526	0.9387	0.9351	0.9491	0.9332	0.9573	0.9633
			100	0.9457	0.9528	0.9428	0.9190	0.9515	0.9412	0.9418	0.9472	0.9470
	0.2	0.3	20	0.9474	0.9377	0.9302	0.9363	0.9514	0.9555	0.9515	0.9515	0.9654
			30	0.9363	0.9488	0.9473	0.9502	0.9132	0.9155	0.9643	0.9562	0.9522
			50	0.9387	0.9513	0.9475	0.9410	0.9543	0.9532	0.9365	0.9231	0.9294
			100	0.9431	0.9551	0.9576	0.9437	0.9581	0.9600	0.9472	0.9505	0.9487
		0.4	20	0.9373	0.9208	0.9243	0.8528	0.9526	0.9552	0.9293	0.9296	0.9465
			30	0.9343	0.9367	0.9326	0.9454	0.9430	0.9411	0.9341	0.9163	0.9384
			50	0.9442	0.9399	0.9420	0.9452	0.9574	0.9442	0.9461	0.9471	0.9497
			100	0.9585	0.9543	0.9608	0.9401	0.9688	0.9546	0.9506	0.9596	0.9542
8	0.15	0.3	20	0.9526	0.9405	0.9488	0.8228	0.9372	0.9494	0.9517	0.9378	0.9479
			30	0.9591	0.9467	0.9509	0.9536	0.9265	0.9213	0.9536	0.9484	0.9673
			50	0.9364	0.9548	0.9562	0.9404	0.9301	0.9302	0.9394	0.9273	0.9552
			100	0.9547	0.9533	0.9496	0.9283	0.9516	0.9641	0.9560	0.9538	0.9371
		0.4	20	0.9383	0.9378	0.9416	0.8139	0.9575	0.9513	0.9209	0.9295	0.9501
			30	0.9522	0.9365	0.9451	0.9404	0.9142	0.9355	0.9275	0.9233	0.9310
			50	0.9480	0.9479	0.9402	0.9484	0.9311	0.9431	0.9498	0.9493	0.9565
			100	0.9561	0.9542	0.9504	0.9418	0.9586	0.9486	0.9495	0.9537	0.9592
	0.2	0.3	20	0.9433	0.9415	0.9355	0.9223	0.9463	0.9523	0.9355	0.9471	0.9387
			30	0.9452	0.9493	0.9447	0.9392	0.9395	0.9107	0.9581	0.9524	0.9363
			50	0.9334	0.9464	0.9512	0.9411	0.9513	0.9537	0.9458	0.9387	0.9497
			100	0.9462	0.9432	0.9498	0.9408	0.9662	0.9634	0.9603	0.9483	0.9622
		0.4	20	0.9446	0.9183	0.9215	0.8825	0.9660	0.9637	0.9143	0.9245	0.9483
			30	0.9545	0.9394	0.9418	0.9542	0.9221	0.9318	0.9391	0.9222	0.9454
			50	0.9472	0.9432	0.9468	0.9353	0.9395	0.9485	0.9421	0.9568	0.9661
			100	0.9495	0.9528	0.9513	0.9402	0.9467	0.9554	0.9478	0.9607	0.9542

pollution, and weather data originally assembled as part of the National Morbidity, Mortality, and Air Pollution Study (NMMAPS). The data have been updated and are available for 108 United States cities for the years 1987–2000. Here we apply our model in the study of the daily accidental deaths data of Detroit in the year 1994 available from the NMMAPS database. The original study examined 90 major cities for the years 1987–1994, including Detroit. The fitted frequency distributions based on the MLE of a Poisson model and zero-inflated Poisson model and Bayesian estimation of Xu et al. (2014) are presented in Table 8 and the MLE and Bayesian estimation results of ZOIP model (4) are presented in Table 9.

According to the fitted frequency shown in Table 9, the frequency of one is underestimated overall and the frequency of two is overestimated by ZIP model (the estimated value of two is nearly double of the true frequency). By ZOIP models, both the MLE and Bayesian estimation for all fre-

Table 3. The point estimation comparision between MLE and Bayes

sample size	estimator		β_1	β_2	γ_{11}	γ_{12}	γ_{21}	γ_{21}
100	MLE	Mean	1.4822	-1.9683	1.1284	-2.1071	1.2092	-1.2119
		Median	1.4903	-2.0125	1.0692	-2.0422	1.1156	-1.1940
		RMSE	0.5462	0.5492	0.6793	0.9789	0.9003	1.2722
	Bayes	Mean	1.4834	-1.8274	0.9611	-1.8294	1.3606	-0.9079
		Median	1.4847	-1.9253	0.9868	-1.8747	1.4349	-1.1560
		RMSE	0.6036	0.7030	0.7705	1.2853	0.9103	1.1949
200	MLE	Mean	1.4867	-1.9700	0.9533	-2.0693	1.1540	-0.8603
		Median	1.5152	-1.9921	1.0977	-2.1176	1.2136	-1.2195
		RMSE	0.5174	0.5239	0.5585	0.6580	0.5979	0.8738
	Bayes	Mean	1.4880	-1.9670	1.0288	-2.1297	1.1572	-1.2638
		Median	1.4905	-2.0360	0.9661	-1.9004	1.1431	-1.0781
		RMSE	0.6148	0.5069	0.6582	0.9266	0.9595	1.1045

Table 4. Fitted frequencies and estimation of θ and q_0 in Xu et al. (2014), Legionellosis data

	Frequency estimation Count of legionellosis cases					Estimation of θ		Estimation of q_0	
	0	1	2	3	4				
Observed frequency	36	23	3	0	1	PE	95% CI	PE	95%CI
MLE(Poisson)	34	14	3	0	0	0.423	(0.246,0.600)		
MLE(ZIP)	36	11	4	1	0	0.675	(0.160,1.190)	0.692	(0.567,0.818)
$Bayes_J$	35	11	4	1	0	0.725	(0.291,1.328)	0.682	(0.553,0.799)
$Bayes_R$	36	11	4	1	0	0.725	(0.291,1.328)	0.689	(0.559,0.805)

Table 5. Fitted frequencies and estimation of q_0, q_1 and θ using ZOIP model (4), Legionellosis data

	Frequency estimation Count of legionellosis cases					Estimation of θ		Estimation of q_0		Estimation of q_1	
	0	1	2	3	4						
Observed frequency	36	23	3	0	1	PE	95% CI	PE	95%CI	PE	95%CI
MLE(ZOIP)	36	23	3	1	0	1.229	(0.294,2.165)	0.571	(0.449,0.693)	0.365	(0.246,0.483)
$Bayes_J$(ZOIP)	35	23	3	1	0	1.242	(0.366,2.168)	0.561	(0.441,0.680)	0.361	(0.249,0.480)
$Bayes_R$(ZOIP)	36	23	3	1	0	1.242	(0.366,2.168)	0.566	(0.445,0.683)	0.365	(0.252,0.484)

Table 6. AIC comparison

AIC	MLE
ZIP	66.18
ZOIP	40.53

Table 7. DIC, WAIC1 and WAIC2 comparison

ZIP	$Bayes_J$	$Bayes_R$	ZOIP	$Bayes_J$	$Bayes_R$
DIC	-92.658	-92.544	DIC	-93.744	-93.528
WAIC1	1003.454	1002.881	WAIC1	987.142	984.356
WAIC2	1019.303	1014.649	WAIC2	992.581	994.208

quencies are closer to the true values. And the results of AIC is presented in Table 10. The AIC value of our results is also smaller than the value of ZIP model. The results of DIC and WAIC are presented in Table 11. The results show that ZOIP model is more appropriate than ZIP model.

In this paper, the covariates of day of week, maximum temperature, minimum temperature, mean relative humidity, and the difference between mean temperature and dew point temperature are considered. The covariate of day of week equals -1 when the day is Saturday or Sunday and equals 1 when the day is from Monday to Friday. The reason why the covariate of the difference between mean temperature and dew point temperature is considered is it is related to the formation of fog. Here we set $\boldsymbol{Z} = \boldsymbol{W_1} = \boldsymbol{W_2} = (\boldsymbol{z_1}, \boldsymbol{z_2}, \boldsymbol{z_3}, \boldsymbol{z_4}, \boldsymbol{z_5})$. The covariates $(\boldsymbol{z_1}, \boldsymbol{z_2}, \boldsymbol{z_3}, \boldsymbol{z_4}, \boldsymbol{z_5})$ represent day of week, the difference between mean temperature and dew point temperature, mean relative humidity, maximum temperature and minimum temperature respectively. We assume $\boldsymbol{\beta_0} = \boldsymbol{\gamma_{00}} = \boldsymbol{\gamma_{01}} = (0, 0, 0, 0, 0)$ and

Table 8. Fitted frequencies and estimation of θ and q_0 in Xu, Xie and Goh(2014), accidental death data

	Frequency estimation Count of accidental deaths								Estimation of θ		Estimation of q_0	
	0	1	2	3	4	5	6	7				
Observed frequency	181	122	28	25	5	2	1	1	PE	95% CI	PE	95%CI
MLE(Poisson)	132	53	14	3	0	0	0	0	0.8110	(0,2.5760)		
MLE(ZIP)	181	105	54	19	5	1	0	0	1.0402	(1.0236,1.0568)	0.4959	(0.4945,0.4972)
$Bayes_J$	181	104	54	19	5	1	0	0	1.0416	(1.0094,1.0718)	0.4962	(0.4943,0.4974)
$Bayes_R$	181	105	54	19	5	1	0	0	1.0397	(1.0098,1.0704)	0.4960	(0.4943,0.4976)

Table 9. Fitted frequencies and estimation of q_0, q_1 and θ using ZOIP model (4), accidental death data

	Frequency estimation Count of accidental deaths								Estimation of θ		Estimation of q_0		Estimation of q_1	
	0	1	2	3	4	5	6	7						
Observed frequency	181	122	28	25	5	2	1	1	PE	95% CI	PE	95%CI	PE	95%CI
MLE(ZOIP)	181	122	31	19	8	3	1	0	1.8168	(1.4225,2.2111)	0.4959	(0.4446,0.5472)	0.3342	(0.2859,0.3826)
$Bayes_J$(ZOIP)	181	122	31	19	9	3	1	0	1.8261	(1.3866,2.2957)	0.4948	(0.4441,0.5462)	0.3335	(0.2857,0.3829)
$Bayes_R$(ZOIP)	181	122	31	19	8	3	1	0	1.8128	(1.3728,2.2659)	0.4954	(0.4457,0.5474)	0.3340	(0.2865,0.3828)

Table 10. AIC comparison

AIC	MLE
ZIP	918.52
ZOIP	913.65

Table 11. DIC, WAIC1 and WAIC2 comparison

ZIP	$Bayes_J$	$Bayes_R$	ZOIP	$Bayes_J$	$Bayes_R$
DIC	-83.599	-83.054	DIC	-85.217	-86.342
WAIC1	951.773	953.098	WAIC1	905.620	907.121
WAIC2	952.512	950.060	WAIC2	906.355	909.890

$\sigma_\beta^2 = \sigma_{\gamma_0}^2 = \sigma_{\gamma_1}^2 = 1000$. The results of ZOIP regression model is presented in Table 12. In this table, the subscript M and B represent the maximum likelihood estimation and Bayesian estimation respectively. The estimations of β_1, β_3 and β_5 are negative which shows the day of Saturday and Sunday, the lower mean relative humidity and the lower minimum temperature lead to the higher the accidental deaths rate. And the estimation of β_4 is positive which shows the higher maximum temperature lead to the higher the accidental deaths rate. The smaller the difference between mean temperature and dew point temperature is easier to lead to a fog, however the sign of the difference is not changeless. Low humidity is more likely to trigger a fire. The signs of $\boldsymbol{\gamma_0}$ and $\boldsymbol{\gamma_1}$ are almost opposite of $\boldsymbol{\beta}$ except the variable of the difference between mean temperature and dew point temperature.

7. CONCLUSIONS

In this paper, we have listed two forms of zero-and-one-inflated Poisson models. The Jeffreys prior and reference priors of the second form are derived. Both of them are shown to be second order matching priors and the posterior distributions based on these priors are proper under a relatively mild condition. The Bayesian method is compared with MLE via Monte Carlo simulation. Simulation results show that the Bayesian estimates perform slightly better when the sample size is small or moderate. The zero-and-one-inflated Poisson regression model is also discussed. Two real data sets are analyzed using both MLE and Bayesian estimates. AIC, DIC and WAIC criterion show that ZOIP model performs better than ZIP model (adopted by Xu et al. [29]) in explaining the data.

APPENDIX: PROOFS OF LEMMAS AND THEOREMS

Proof of Theorem 2.1. Let $v(\theta) = (n - S_0 - S_1)\theta e^\theta - Se^\theta - (n - S - S_0 - S_1)\theta + S$. Then,

$$v'(\theta) = (n - S_0 - S_1)\theta e^\theta + (n - S - S_0 - S_1)e^\theta - (n - S - S_0 - S_1),$$
$$v''(\theta) = (n - S_0 - S_1)\theta e^\theta + (2n - S - 2S_1 - 2S_0)e^\theta.$$

According to the definition of S and the condition that $n - S_0 - S_1 > 0$, we have $(2n - S - 2S_1 - 2S_0) < 0$. It immediately follows that $v''(\theta) > 0$ when $\theta > \frac{S+2S_1+2S_0-2n}{n-S_0-S_1}$ and $v''(\theta) < 0$ when $\theta < \frac{S+2S_1+2S_0-2n}{n-S_0-S_1}$. So $v'(\theta)$ is decreasing on $(0, \frac{S+2S_1+2S_0-2n}{n-S_0-S_1})$ and increasing on $(\frac{S+2S_1+2S_0-2n}{n-S_0-S_1}, +\infty)$. Additionally, it can be easily verified that $v'(0) = 0$ and $v'(+\infty) > 0$. Therefore there exists $a_1 > 0$ such that $v'(\theta) < 0$ when $\theta \in (0, a_1)$, and $v'(\theta) > 0$, when $\theta \in (a_1, +\infty)$. Now

Table 12. The parameter estimation of ZOIP regression model

	$ZOIP_M$	$ZOIP_B$		$ZOIP_M$	$ZOIP_B$		$ZOIP_M$	$ZOIP_B$
β_1	-3.6952	-4.1455	γ_{11}	5.9440	6.4126	γ_{21}	3.9052	4.2338
β_2	1.4016	1.1055	γ_{12}	2.6594	2.9803	γ_{22}	-1.7878	-1.2560
β_3	-5.7981	-5.0282	γ_{13}	7.2280	8.1645	γ_{23}	3.3588	3.5141
β_4	4.1642	4.3041	γ_{14}	-3.1861	-3.3418	γ_{24}	-4.9368	-5.3248
β_5	-0.1988	-0.2267	γ_{15}	2.8411	3.2901	γ_{25}	4.3301	3.6788

that we can obtain the trend of $v(\cdot)$ on $(0,+\infty)$, and it can be further shown that $v(0)=0$ and $v(+\infty)>0$. With the continuity of the function $v(\cdot)$, the result is derived that there is only one solution for Equation (9).

Proof of Lemma 3.1. It is easy to see that the denominator of $k(\theta)$ is positive. Denoting the numerator by $f(\theta)$, then we only need to prove $f(\theta)>0$.

Simple calculation yields

$$f'(\theta)=\mathrm{e}^{\theta}(2\mathrm{e}^{\theta}-\theta^2-2\theta-2).$$

Let $g(\theta)=2\mathrm{e}^{\theta}-\theta^2-2\theta-2$. Noting that $g''(\theta)>0$ for $\theta>0$ and $g'(0)=0$, we have $g'(\theta)>0$ for $\theta>0$, which combining with $g(0)=0$ yields $g(\theta)>0$ for $\theta>0$. Thus $f'(\theta)>0$ for $\theta>0$. Also it can be easily verified that $f(0)=0$. So $f(\theta)>0$ for $\theta>0$ holds.

Proof of Lemma 3.2. It is clear that

$$E^l[\log|\boldsymbol{h_1}(\boldsymbol{\eta})||\boldsymbol{\eta_1}]=\log|\boldsymbol{h_1}(\boldsymbol{\eta})|.$$

The result is immediate from (12). This lemma is a special case of Lemma 2.1 in Berger and Bernado [3].

Proof of Theorem 3.1. Suppose that (q_0,q_1) are the parameters of interest while θ is a nuisance parameter. Set $\boldsymbol{\eta}=(q_0,q_1,\theta)$, $\boldsymbol{\eta_1}=(q_0,q_1)$ as the first group, and $\boldsymbol{\eta_2}=\theta$ as the second group. Then the corresponding Fisher information matrix for $(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2)$ is

$$\boldsymbol{H_1}(q_0,q_1,\theta)=\begin{pmatrix}\boldsymbol{h_1} & \mathbf{0}\\ \mathbf{0} & h_2\end{pmatrix},$$

where $\boldsymbol{h_1}$ and h_2 are given in (10). Further calculation yields

$$\int_0^1\int_0^{1-q_1}\frac{1}{\sqrt{q_0q_1(1-q_0-q_1)}}\mathrm{d}q_0\mathrm{d}q_1=2\pi,$$
$$\int_0^1\int_0^{1-q_1}\frac{\log(1-q_0-q_1)}{\sqrt{q_0q_1(1-q_0-q_1)}}\mathrm{d}q_0\mathrm{d}q_1=-4\pi.$$

Let $\Theta^l(q_0,q_1,\theta)=\{0\le q_0\le 1, 0\le q_1\le 1-q_0, \theta\in[\frac{1}{l},2l]\}$, $l=1,2,\ldots$. Then $\{\Theta^l,l=1,2,\ldots\}$ is a nested sequence of compact subsets such that $\cup_{l=1}^{\infty}\Theta^l=\Theta$. It is easy to find that $|\boldsymbol{h_1}(\boldsymbol{\eta})|$ depends only on $\boldsymbol{\eta_1}$. It follows from Lemma 3.2 that,

$$\pi^l(\boldsymbol{\eta})=\frac{k(\theta)^{\frac{1}{2}}}{\int_{\frac{1}{l}}^{2l}k(\theta)\mathrm{d}\theta}\cdot\frac{1}{2\pi\sqrt{q_0q_1(1-q_0-q_1)}}\cdot I_{\Theta^l}(\boldsymbol{\eta}).$$

According to the algorithm above, the reference prior $\pi_R(\boldsymbol{\eta})\propto\lim_{l\to\infty}\pi^l(\boldsymbol{\eta})$. So we can obtain

$$\pi_R(q_0,q_1,\theta)\propto q_0^{-1/2}q_1^{-1/2}(1-q_0-q_1)^{-1/2}k(\theta)^{1/2},$$

with $\theta>0$, $0\le q_0\le 1$, and $0\le q_1\le 1-q_0$.

If θ is the parameter of interest, we set $\boldsymbol{\eta_1}=\theta$ and $\boldsymbol{\eta_2}=(q_0,q_1)$. Then the corresponding Fisher information matrix for $(\boldsymbol{\eta}_1,\boldsymbol{\eta}_2)$ is

$$\boldsymbol{H_2}(q_0,q_1,\theta)=\begin{pmatrix}h_2 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{h_1}\end{pmatrix},$$

where $\boldsymbol{h_1}$ and h_2 are the same as above, which are given in (10). Choose the same nested sequence of compact subsets of Θ as the above case. Then for each $l\in N^+$,

$$\pi_2^l(\boldsymbol{\eta_2}|\boldsymbol{\eta_1})=\frac{1}{4\pi\sqrt{q_0q_1(1-q_0-q_1)}}I_{\Theta^l}(\boldsymbol{\eta}),$$
$$E^l[\log(\boldsymbol{h_1}(\boldsymbol{\eta})|\boldsymbol{\eta_1})]=\log k(\theta)+2.$$

It immediately follows that

$$\pi^l(\boldsymbol{\eta})=\frac{1}{2\pi\sqrt{q_0q_1(1-q_0-q_1)}}\cdot\frac{k(\theta)^{\frac{1}{2}}}{\int_{\frac{1}{l}}^{2l}k(\theta)\mathrm{d}\theta}\cdot I_{\Theta^l}(\boldsymbol{\eta}).$$

Obviously, we can derive the same reference prior as in the first case.

Proof of Theorem 3.2. Let

$$k_{11}(\boldsymbol{\eta})^{-1/2}=h_{11}^{1/2}=(1-q_0-q_1)^{1/2}k(\theta)^{1/2}.$$

When Jeffreys prior is used, set $\boldsymbol{\eta}=(\theta,q_0,q_1)$ and let

$$K(q_0,q_1)=q_0^{-1/2}q_1^{-1/2}(1-q_0-q_1)^{-1/2}.$$

Then the Jeffreys prior (11) is a second-order matching prior for θ. When the reference prior with θ as the interesting parameter is used, let

$$K(q_0,q_1)=q_0^{-1/2}q_1^{-1/2}(1-q_0-q_1)^{-1}.$$

Then, according to Lemma (3.3), the reference prior (16) is a second-order matching prior for θ. When q_0 is the parameter of interest q_0 and (q_1,θ) are not orthogonal. According to Equation (32) given in Peers [25] or Equation (2.7) given in Mukerjee and Ghosh [22], it is easy to find that the Jeffreys prior and reference prior are not second-order matching prior when q_0 is the parameter of interest. When q_1 is the parameter of interest, the result is similar to q_0.

Proof of Theorem 3.3. According to Theorem 1 in Yang [30], when θ is the parameter of interest, the reference prior for (θ, p_0, p_1) can be calculated as follows,

$$\begin{aligned}\pi_R(\theta,p_0,p_1) =& \pi_R(\theta,q_0,q_1)\left|\frac{\partial(\theta,q_0,q_1)}{\partial(\theta,p_0,p_1)}\right| \\ =& \left[p_0p_1\mathrm{e}^{\theta}+(1-\mathrm{p}_0)\right]^{-1/2} \\ & \left[p_0(1-p_1)\mathrm{e}^{\theta}+(1-\mathrm{p}_0)\theta\right]^{-1/2} \\ & (1-p_0)^{-1/2}p_0\left(\mathrm{e}^{\theta}-\theta-1\right)k(\theta)^{1/2}.\end{aligned} \tag{21}$$

Proof of Theorem 3.4. It is easy to prove that

$$0<\int_0^1\int_0^{1-q_0}q_0^{S_0-\frac{1}{2}}q_1^{S_1-\frac{1}{2}}(1-q_0-q_1)^{n-S_0-S_1}\mathrm{d}q_0\mathrm{d}q_1 < +\infty$$

and

$$0<\int_0^1\int_0^{1-q_0}q_0^{S_0-\frac{1}{2}}q_1^{S_1-\frac{1}{2}}(1-q_0-q_1)^{n-S_0-S_1-\frac{1}{2}}\mathrm{d}q_0\mathrm{d}q_1 < +\infty.$$

As we all know, $\lim\limits_{\theta\to+\infty}\frac{\theta^m}{\mathrm{e}^\theta}=0$, for any given $m\in R$. And it is easy to see that

$$\frac{\theta^S}{\mathrm{e}^\theta-\theta-1}\left(\frac{\mathrm{e}^{2\theta}-\theta^2\mathrm{e}^\theta-2\mathrm{e}^\theta+1}{\theta(\mathrm{e}^\theta-\theta-1)^2}\right)^{1/2}=O\left(\frac{\theta^{S-\frac{1}{2}}}{\mathrm{e}^\theta}\right) \quad (\theta\to+\infty).$$

Then we have

$$\int_1^{+\infty}\frac{\theta^S}{\mathrm{e}^\theta-\theta-1}\left(\frac{\mathrm{e}^{2\theta}-\theta^2\mathrm{e}^\theta-2\mathrm{e}^\theta+1}{\theta(\mathrm{e}^\theta-\theta-1)^2}\right)^{1/2}\mathrm{d}\theta<+\infty.$$

Using Taylor expansion, we have $\mathrm{e}^\theta=1+\theta+\frac{\theta^2}{2}+\mathrm{O}(\theta^3)(\theta\to 0)$. There exists a sufficiently small neighborhood $U^+(0)$ and a positive number M such that any point θ within the neighborhood satisfies

$$\frac{\theta^{S-\frac{5}{2}}}{M}\le\frac{\theta^S}{\mathrm{e}^\theta-\theta-1}\left(\frac{\mathrm{e}^{2\theta}-\theta^2\mathrm{e}^\theta-2\mathrm{e}^\theta+1}{\theta(\mathrm{e}^\theta-\theta-1)^2}\right)^{1/2}\le M\cdot\theta^{S-\frac{5}{2}}.$$

When $S<2$,

$$\int_0^1\frac{\theta^S}{\mathrm{e}^\theta-\theta-1}\left(\frac{\mathrm{e}^{2\theta}-\theta^2\mathrm{e}^\theta-2\mathrm{e}^\theta+1}{\theta(\mathrm{e}^\theta-\theta-1)^2}\right)^{1/2}\mathrm{d}\theta\to+\infty.$$

When $S\ge 2$,

$$\int_0^1\frac{\theta^S}{\mathrm{e}^\theta-\theta-1}\left(\frac{\mathrm{e}^{2\theta}-\theta^2\mathrm{e}^\theta-2\mathrm{e}^\theta+1}{\theta(\mathrm{e}^\theta-\theta-1)^2}\right)^{1/2}\mathrm{d}\theta<+\infty.$$

Combining the above results, we can find that the posterior of (q_0, q_1, θ) with either the Jeffreys or reference prior is proper when $S\ge 2$ and vice versa. This condition is weak since it will be met if there exist at least one sample larger than one.

Proof of Theorem 3.5. The likelihood functions of (p_0, p_1, θ) is

$$\begin{aligned}L(p_0,p_1,\theta|\boldsymbol{Y})\propto&\left[p_0p_1\mathrm{e}^\theta+(1-p_0)\right]^{S_0} \\ &\left[p_0(1-p_1)\mathrm{e}^\theta+(1-p_0)\theta\right]^{S_1} \\ &\times(1-p_0)^{n-S_0-S_1}\theta^S\mathrm{e}^{-n\theta}.\end{aligned} \tag{22}$$

According to the likelihood function (22) and the reference prior (21), when θ is the parameter of interest, the posterior distribution of (θ, p_0, p_1) is

$$\begin{aligned}\pi_R(\theta,p_0,p_1|\boldsymbol{Y})=&\left[p_0p_1+(1-p_0)\mathrm{e}^{-\theta}\right]^{S_0-1/2} \\ &\left[p_0(1-p_1)+(1-p_0)\theta\mathrm{e}^{-\theta}\right]^{-1/2} \\ &(1-p_0)^{n-S_0-S_1-1/2}p_0(\mathrm{e}^\theta-\theta-1) \\ &k(\theta)^{1/2}\theta^S\mathrm{e}^{-(\mathrm{n}-\mathrm{S}_0-\mathrm{S}_1+1)\theta}.\end{aligned}$$

It suffices to prove that

$$\begin{aligned}\int_0^\infty\int_0^1\int_0^1&\left[p_0p_1+(1-p_0)\mathrm{e}^{-\theta}\right]^{S_0-1/2} \\ &\left[p_0(1-p_1)+(1-p_0)\theta\mathrm{e}^{-\theta}\right]^{-1/2} \\ &(1-p_0)^{n-S_0-S_1-1/2}p_0(\mathrm{e}^\theta-\theta-1) \\ &k(\theta)^{1/2}\theta^S\mathrm{e}^{-(n-S_0-S_1+1)\theta}\mathrm{d}p_0\mathrm{d}p_1\mathrm{d}\theta<\infty.\end{aligned}$$

With the following transformation

$$\begin{cases} p_0=\frac{q_0+q_1-(1+\theta)\mathrm{e}^{-\theta}}{1-(1+\theta)\mathrm{e}^{-\theta}}, \\ p_1=\frac{q_0-(1-p_0)\mathrm{e}^{-\theta}}{p_0}, \\ \theta=\theta, \end{cases} \tag{23}$$

the above integrand becomes

$$\int_0^\infty\int_0^1\int_0^{1-q_0}q_0^{S_0-\frac{1}{2}}q_1^{S_1-\frac{1}{2}}(1-q_0-q_1)^{n-S_0-S_1-\frac{1}{2}} \frac{\theta^S}{(\mathrm{e}^\theta-\theta-1)^{n-S_0-S_1}}k^{\frac{1}{2}}(\theta)\mathrm{d}q_0\mathrm{d}\mathrm{q}_1\mathrm{d}\theta<\infty.$$

The proof of this integrand is the same as Theorem 3.4.

ACKNOWLEDGEMENTS

The authors would like to thank the Editor, an Associate Editor and two referees for their helpful comments and useful suggestions on the previous drafts of the paper.

Received 10 April 2016

REFERENCES

[1] Bae, S.J., Hwang, J.Y. and Kuo, W. (2007). Yield prediction via spatial modeling of clustered defect counts across a wafer map. *IIE Transactions.* **39** 1073–1083.

[2] Bassil, K.L., Cole, D.C. Moineddin, R., Lou, W., Craig, A.M., Schwartz, B. and Rea, E. (2011). The relationship between temperature and ambulance response calls for heat-related illness in Toronto, Ontario, 2005. *Journal of Epidemiology & Community Health.* **65** 829–831.

[3] Berger, J.O. and Bernardo, J.M. (1992). Ordered group reference priors with applications to a multinomial problem. *Biometrika.* **79** 25–37. MR1158515

[4] Berger, J.O., Bernardo, J.M., and Sun, D. (2015). Overall objective priors. *Bayesian Analysis.* **10** 189–221. MR3420902

[5] Broek, J. (1995). A score test for zero inflation in a Poisson distribution. *Biometrics.* **85** 738–743. MR1349912

[6] Chen, X.D. (2009). Bayesian analysis of semiparametric mixed-effects models for zero inflated count data. *Communications in Statistics -Theory and Methods.* **38** 1815–1833. MR2552965

[7] Crepon, B. and Duguet, E. (1997). Research and development, competition and innovation pseudo-maximum likelihood and simulated maximum likelihood methods applied to count data models with heterogeneity. *Journal of Econometrics.* **79** 355–378.

[8] Dagne, G.A. (2010). Bayesian semiparametric zero–inflated Poisson model for longitudinal count data. *Mathematical Biosciences.* **224** 126–130. MR2655805

[9] Fletcher, D.F., Kavalieris, L. and Manly, B.F.J. (1998). *Statistics in Ecology and Environmental Monitoring 2: Decision Making and Risk Assessment in Biology.* University of Otago Press.

[10] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, B.D., Vehtari, A. and Rubin D.B. (2013). *Bayesian Data Analysis, Third Edition.* CRC Press. MR3235677

[11] Ghosh, S.K., Mukhopadhyay, P. and Lu, J.C. (2006). Bayesian analysis of zero-inflated regression models. *Journal of Statistical Planning and Inference.* **136** 1360–1375. MR2253768

[12] Gilks, W.R. and Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society: Series A* **41** 337–348.

[13] Gurmu, S. and Trivedi, P. (1996). Excess zeros in count models for recreational trips. *Journal of Business & Economic Statistics.* **14** 469–477.

[14] Hasan, M.T., Sneddon, G. and Ma, R. (2009). Pattern-Mixture Zero-Inflated Mixed Models for Longitudinal Unbalanced Count Data with Excessive Zeros. *Biometrical Journal.* **51** 946–960. MR2744449

[15] King, G. (2008). Event count models for international relations: generalizations and applications. *International Studies Quarterly.* **32** 123–147.

[16] Lam, K.F., Xue, H. and Cheung, Y.B. (2006). Semiparametric analysis of zero-inflated count data. *Biometrics.* **62** 996–1003. MR2297670

[17] Lam, M.C., Ang, L.W., Tan, A.L., James, L. and Goh, K.T. (2011). Epidemiology and control of Legionellosis, Singapore. *Emerging Infectious Diseases.* **17** 1209–1215.

[18] Lambert, D. (1992). Zero-infated Poisson regression with application to defects in manufacturing. *Technometrics.* **34** 1–14.

[19] Melkersson, M. and Olsson, C. (1999). Is visiting the dentist a good habit? Analyzing count data with excess zeros and excess ones. *Umeå Economic Studies.* **492** 1–18.

[20] Melkersson, M. and Rooth, D. (2000). Modeling Female Fertility Using Inflated Count Data Models. *Journal of Population Economics.* **13** 189–203.

[21] Miaou, S.P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention.* **26** 471–482. MR2693677

[22] Mukerjee, R. and Ghosh, M. (1997). Second-order probability matching priors. *Biometrika.* **84** 970–975. MR1625016

[23] Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics.* **33** 341–365. MR0867980

[24] Musio, M., Sauleau, E.A. and Buemi, A. (2010). Bayesian semiparametric ZIP models with space–time interactions: an application to cancer registry data. *Mathematical Medicine and Biology.* **27** 181–194. MR2657556

[25] Peers, H.W. (1965). On confidence sets and Bayesian probability points in the case of several parameters. *Journal of the Royal Statistical Society: Series B.* **27** 9–16. MR0191029

[26] Shonkwiler, J. and Shaw, W. (1996). Hurdle count-data models in recreation demand analysis. *Journal of Agricultural and Resource Economics.* **21** 210–219.

[27] Welsh, A., Cunningham, R. Donnelly, C. and Lindenmayer, D. (1996). Modeling the abundance of rare species: statistical models for counts with extra zeros. *Ecological Modelling.* **88** 297–308.

[28] Xu, A. and Tang, Y. (2011). Bayesian analysis of Birnbaum-Saunders distribution with partial information. *Computational Statistics & Data Analysis.* **55** 2324–2333. MR2786991

[29] Xu, H.Y., Xie, M. and Goh, T.N. (2014). Objective Bayes analysis of zero-inflated Poisson distribution with application to healthcare data. *IIE Transactions.* **46** 843–852.

[30] Yang, R. (1995). Invariance of the Reference Prior Under Reparametrization. *Test.* **4** 83–94. MR1365041

[31] Zhang, C., Tian, G.-L. and Ng, K.-W. (2016). Properties of the zero-and-one inflated Poisson distribution and likelihood-based inference methods. *Statistics and Its Interface.* **9** 11–32. MR3423307

Wenchen Liu
School of Statistics
East China Normal University
Shanghai 200241
China
E-mail address: liuwenchen321@126.com

Yincai Tang
School of Statistics
East China Normal University
Shanghai 200241
China
E-mail address: yctang@stat.ecnu.edu.cn

Ancha Xu
College of Mathematics and Information Science
Wenzhou University
Zhejiang, 325035
China
E-mail address: xuancha@wzu.edu.cn