

Bayesian analysis of multiple-inflation Poisson models and its application to infection data

Duchwan Ryu^a, Devrim Bilgili^b, Önder Ergönül^c and Nader Ebrahimi^a

^a*Northern Illinois University*

^b*University of North Florida*

^c*Koç University*

Abstract. In this article we propose a multiple-inflation Poisson regression to model count response data containing excessive frequencies at more than one non-negative integer values. To handle multiple excessive count responses, we generalize the zero-inflated Poisson regression by replacing its binary regression with the multinomial regression, while Su et al. [*Statist. Sinica* **23** (2013) 1071–1090] proposed a multiple-inflation Poisson model for consecutive count responses with excessive frequencies. We give several properties of our proposed model, and do statistical inference under the fully Bayesian framework. We perform simulation studies and also analyze the data related to the number of infections collected in five major hospitals in Turkey, using our methodology.

1 Introduction

The Poisson distribution is used to analyze count data under the generalized linear model. However, the count data often contain an excessive zeros and are over dispersed relative to the Poisson distribution (Cox, 1983). To properly handle the problem of excessive zeros several flexible methods have been proposed, including the following three methods. The first method is to use zero-inflated Poisson (ZIP) models which use the mixture of binomial and Poisson distributions, see Lambert (1992), Li (2012). The second method is to use the negative binomial distribution instead of Poisson distribution which is referred to as the zero-inflated negative binomial (ZINB) models (Greene, 1994). The third method is to separate the zeros from the data and to apply an incomplete Poisson distribution to the non-zero data in hurdle models (Mullahy, 1986). See Hu, Pavlicova and Nunes (2011) for comparison of ZIP, ZINB and hurdle models.

All models assume the count data containing a single value, specifically zero, with high frequency. However, there are situations that count data contain excessive frequencies at multiple values. See, for example, Huang and Chin (2010). In this paper, we consider a multiple-inflation Poisson (MIP) model to account for the

Key words and phrases. Bayesian generalized linear model, EM algorithm, excessive count response, likelihood function, zero-inflated poisson model.

Received March 2016; accepted October 2016.

excessive frequencies in several responses. Specifically, if the excessive frequencies occur consecutively, for example, the consecutive numbers from zero to some integer, then we treat them as ordinal outcomes. MIP with ordinal responses also considered by Su et al. (2013). If the excessive frequencies occur randomly on any non-negative integer, then we treat them as nominal outcomes. Using a mixture distribution of multinomial and Poisson distributions, we propose an MIP model for both ordinal and nominal responses.

This research was motivated by a need to assess the relationship between the number of infections per week, which includes excessive zeros and ones, and several covariates using the infection data collected by five hospitals in Turkey. We develop MIP regression models under Bayesian framework and apply them to the infection data analysis. It should be noted that Su et al. (2013) based their inference on Frequentists approaches.

The paper is organized as follows. Section 2 describes the multiple-inflation Poisson distribution, gives several properties and discusses Bayesian analysis of our model. In Section 3, we perform simulation studies to compare the MIP regression models with the ZIP regression model and apply our method to infection data. Finally, Section 4 gives conclusions.

2 Multiple-inflation Poisson distribution, its properties and statistical inference

Consider a non-negative discrete random variable Y that takes excessive probabilities at multiple values, say c_1, \dots, c_m , comparing to the Poisson distribution. Then we say Y has a multiple-inflation Poisson (MIP) distribution, at c_1, \dots, c_m , denoted as $\text{MIP}[c_1, \dots, c_m; \pi_1, \dots, \pi_m; \lambda]$, if its probability mass function is given by

$$P(Y = y) = \sum_{j=1}^m I\{y = c_j\} \pi_j + \left(1 - \sum_{j=1}^m \pi_j\right) \frac{\lambda^y}{y!} e^{-\lambda}, \quad y = 0, 1, 2, \dots, \quad (2.1)$$

for $0 \leq \pi_j \leq 1$, $j = 1, \dots, m$, $\sum_{j=1}^m \pi_j \leq 1$ and $\lambda > 0$, as an extension of zero-inflated Poisson (ZIP) distribution (Lambert, 1992).

Now, based on the feature of c_1, \dots, c_m we propose two types of MIP: (Type 1) when c_1, \dots, c_m are consecutive integers, and (Type 2) when they are any non-negative integers. They are referred to as an *ordinal* MIP, see Su et al. (2013), and a *nominal* MIP, respectively.

Several comments are in order with regard to (2.1):

(A) In the equation (2.1), if $m = 1$ and $c_1 = 0$, then we get ZIP distribution.

- (B) The random variable Y follows a mixture of the multinomial distribution and the Poisson distribution. While ZIP model only considers the excessive probability at zero as a mixture of Binomial and Poisson distributions (Lambert, 1992), MIP model allows excessive probabilities at several non-negative integer values and can be considered as a mixture of the multinomial distribution with one trial and the Poisson distribution.
- (C) The first and the second moments of Y are

$$E(Y) = \sum_{y=0}^{\infty} yP(Y = y) = \sum_{j=1}^m c_j \pi_j + \left(1 - \sum_{j=1}^m \pi_j\right) \lambda,$$

$$E(Y^2) = \sum_{y=0}^{\infty} y^2 P(Y = y) = \sum_{j=1}^m c_j^2 \pi_j + \left(1 - \sum_{j=1}^m \pi_j\right) \lambda(\lambda + 1).$$

Consequently,

$$\begin{aligned} \text{var}(Y) &= E(Y^2) - \{E(Y)\}^2 \\ &= \lambda^2 \left(1 - \sum_{j=1}^m \pi_j\right) \sum_{j=1}^m \pi_j + \lambda \left(1 - \sum_{j=1}^m \pi_j\right) \left(1 - 2 \sum_{j=1}^m c_j \pi_j\right) \\ &\quad + \sum_{j=1}^m c_j^2 \pi_j - \left(\sum_{j=1}^m c_j \pi_j\right)^2. \end{aligned}$$

As a special case when $c_1 = 0$, $c_2 = 1$ and $m = 2$ we have

$$E(Y) = \pi_2 + (1 - \pi_1 - \pi_2) \lambda \quad \text{and}$$

$$\text{var}(Y) = \lambda^2 (1 - \pi_1 - \pi_2)(\pi_1 + \pi_2) + \lambda(1 - \pi_1 - \pi_2)(1 - 2\pi_2) + \pi_2 - \pi_2^2.$$

To calculate general moments of Y , we use the Stirling number of the second kind defined by

$$S(r, t) = \frac{1}{t!} \sum_{s=0}^t (-1)^s \binom{t}{s} (t-s)^r,$$

which is the number of ways of partitioning a set of r elements into t non-empty sets. See Dobiński (1877) for more detail. We may have the triangle of $S(r, t)$ shown in Table 1. The $S(r, t)$ has the following property (generating function)

$$y^r = \sum_{t=0}^r S(r, t) y(y-1) \cdots (y-t+1).$$

Now, using $S(r, t)$, the following lemma gives the r th moment of $\text{MIP}[c_1, \dots, c_m; \pi_1, \dots, \pi_m; \lambda]$.

Table 1 Values for $S(r, t)$

r	t					
	1	2	3	4	5	6
1	1					
2	1	1				
3	1	3	1			
4	1	7	6	1		
5	1	15	25	10	1	
6	1	31	90	65	15	1

Lemma 1. For $MIP[c_1, \dots, c_m; \pi_1, \dots, \pi_m; \lambda]$ the r th moment is

$$\begin{aligned}
 E(Y^r) &= \sum_{j=1}^m c_j^r \pi_j + \left(1 - \sum_{j=1}^m \pi_j\right) \sum_{y=0}^{\infty} y^r \frac{\lambda^y}{y!} e^{-\lambda} \\
 &= \sum_{j=1}^m c_j^r \pi_j + \left(1 - \sum_{j=1}^m \pi_j\right) \sum_{t=1}^r \lambda^t S(r, t).
 \end{aligned}$$

An interesting application of Lemma 1 is that one can use it to come up with estimates of unknown parameters. For example, suppose we start with $MIP[0, 1; \pi_1, \pi_2; \lambda]$, where (π_1, π_2, λ) are the parameters needed to be estimated from a sample $\{y_1, \dots, y_n\}$. Using the method of moments, we can estimate the parameters from their first three sample moments by solving the followings

$$\begin{aligned}
 \frac{1}{n} \sum_{i=1}^n y_i &= \pi_2 + (1 - \pi_1 - \pi_2)\lambda, \\
 \frac{1}{n} \sum_{i=1}^n y_i^2 &= \pi_2 + (1 - \pi_1 - \pi_2)\lambda(\lambda + 1) \quad \text{and} \\
 \frac{1}{n} \sum_{i=1}^n y_i^3 &= \pi_2 + (1 - \pi_1 - \pi_2)\lambda(\lambda^2 + 3\lambda + 1).
 \end{aligned}$$

2.1 Models with MIP

We first consider the ordinal MIP for count data with excessive frequencies at consecutive m non-negative integers. For notational convenience, we consider the integers from 0 to $m - 1$, and assume $Y \sim MIP[0, 1, \dots, m - 1; \pi_1, \dots, \pi_m; \lambda]$. Let

$$P_j = \log \left\{ \frac{P(Y \leq j)}{P(Y > j)} \right\} = \log \left(\frac{\sum_{k=1}^j \pi_k}{1 - \sum_{k=1}^j \pi_k} \right), \quad j = 1, \dots, m, \quad (2.2)$$

then, in terms of p covariates we assume the following models for the cumulative logit, P_j , $j = 1, \dots, m$, and a logistic model for λ :

$$P_j(\mathbf{x}_i) = \mathbf{x}_i \boldsymbol{\beta}_j, \quad j = 1, \dots, m, i = 1, \dots, n \tag{2.3}$$

and

$$\log\{\lambda(\mathbf{x}_i)\} = \mathbf{x}_i \boldsymbol{\gamma}, \quad i = 1, \dots, n, \tag{2.4}$$

where n is the number of observations, $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip})$ and x_{il} is the i th observation on the l th covariate, $l = 1, \dots, p$, $\boldsymbol{\beta}_j$ is $(p + 1)$ regression coefficients corresponding to P_j , $j = 1, \dots, m$, and $\boldsymbol{\gamma}$ is $(p + 1)$ regression coefficients corresponding to λ . In this paper, as in [Hedger and Gibbons \(2006\)](#), we model the cumulative logit with common slopes, $\beta_1^*, \dots, \beta_p^*$, for p covariates over different categories of integers with excessive frequencies under the ordinal MIP such that

$$\begin{aligned} \boldsymbol{\beta}_j &= (\beta_{0j}, \beta_{1j}, \dots, \beta_{pj})^T \\ &= (\beta_{0j}, \beta_1^*, \dots, \beta_p^*)^T, \quad j = 1, \dots, m, \end{aligned} \tag{2.5}$$

where β_{0j} is the threshold for P_j . It should be noted that, under the cumulative logit link function, the thresholds β_{0j} are in ascending order, that is, $\beta_{01} \leq \beta_{02} \leq \dots \leq \beta_{0m}$.

For non-consecutive and non-negative integers, c_1, \dots, c_m , with excessive frequencies, we assume the nominal MIP denoted by $Y \sim \text{MIP}[c_1, \dots, c_m; \pi_1, \dots, \pi_m; \lambda]$ and let

$$F_j = \log\left(\frac{\pi_j}{\pi_0}\right), \quad j = 1, 2, \dots, m, \tag{2.6}$$

where $\pi_0 = 1 - \sum_{j=1}^m \pi_j$. That is, in the mixture of Poisson and multinomial distributions, we denote the portion of the Poisson distribution as the reference probability π_0 . As before, we assume the following models for F_j , $j = 1, \dots, m$, and λ :

$$F_j(\mathbf{x}_i) = \mathbf{x}_i \boldsymbol{\beta}_j, \quad j = 1, \dots, m, i = 1, \dots, n \tag{2.7}$$

and

$$\log\{\lambda(\mathbf{x}_i)\} = \mathbf{x}_i \boldsymbol{\gamma}, \quad i = 1, \dots, n, \tag{2.8}$$

where n , \mathbf{x}_i and $\boldsymbol{\gamma}$ are the same as in equations (2.3) and (2.4), $\boldsymbol{\beta}_j$ is $(p + 1)$ regression coefficients corresponding to F_j .

It is worth noting that the nominal MIP model is more general than the ordinal MIP model, in the sense that (i) the nominal MIP allows different regression coefficients, $\boldsymbol{\beta}_j$, $j = 1, \dots, m$, in all their components, while the ordinal MIP allows different coefficients only in their first component and (ii) the nominal MIP model can handle the consecutive non-negative integers with excessive frequencies.

2.2 Likelihoods of MIP models

Let Y_1, \dots, Y_n be independent random samples from $MIP[c_1, \dots, c_m; \pi_1, \dots, \pi_m; \lambda]$ and let h_j denote the frequency of $c_j, j = 1, \dots, m$. We first consider the ordinal MIP, i.e. $(c_1, \dots, c_m) = (0, 1, \dots, m - 1)$. Without loss of generality, for $n_0 = 0$ and $n_j = h_1 + \dots + h_j$, we rearrange samples such that

$$Y_i = \begin{cases} 0, & i = 1, \dots, n_1, \\ 1, & i = n_1 + 1, \dots, n_2, \\ \vdots & \\ m - 1, & i = n_{m-1} + 1, \dots, n_m, \text{ and} \\ y_i, & i = n_m + 1, \dots, n, \end{cases} \tag{2.9}$$

where y_i 's are integers greater than or equal to m . When the covariates \mathbf{x}_i are available ($\dim(\mathbf{x}_i) = p + 1 < \min\{h_1, \dots, h_m, n - n_m\}$), under the models (2.3)–(2.5) the log-likelihood is given by

$$l(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m, \boldsymbol{\gamma}) = \sum_{j=1}^m \sum_{i=n_{j-1}+1}^{n_j} \log P(Y_i = j - 1 | \mathbf{x}_i) + \sum_{i=n_m+1}^n \log P(Y_i = y_i | \mathbf{x}_i),$$

where

$$P(Y_i = j - 1 | \mathbf{x}_i) = \pi_j(\mathbf{x}_i) + \left\{ 1 - \sum_{k=1}^m \pi_k(\mathbf{x}_i) \right\} \frac{\lambda(\mathbf{x}_i)^{j-1}}{(j - 1)!} e^{-\lambda(\mathbf{x}_i)}, \quad j = 1, \dots, m,$$

$$P(Y_i = y_i | \mathbf{x}_i) = \left\{ 1 - \sum_{k=1}^m \pi_k(\mathbf{x}_i) \right\} \frac{\lambda(\mathbf{x}_i)^{y_i}}{y_i!} e^{-\lambda(\mathbf{x}_i)}, \quad y_i \geq m,$$

and $\pi_1(\mathbf{x}_i) = \frac{1}{1 + e^{-P_1(\mathbf{x}_i)}}$, $\pi_j(\mathbf{x}_i) = \frac{1}{1 + e^{-P_j(\mathbf{x}_i)}} - \frac{1}{1 + e^{-P_{j-1}(\mathbf{x}_i)}}$, $j = 2, \dots, m$, and $\lambda(\mathbf{x}_i)$ can be obtained from the equations (2.3) and (2.4), respectively.

Since the probabilities of integer values with excessive frequencies are given by

$$P(Y_i = 0 | \mathbf{x}_i) = \frac{1}{1 + e^{P_m(\mathbf{x}_i)}} \left(\frac{1 + e^{P_m(\mathbf{x}_i)}}{1 + e^{-P_1(\mathbf{x}_i)}} + e^{-\lambda(\mathbf{x}_i)} \right)$$

and

$$P(Y_i = j - 1 | \mathbf{x}_i) = \frac{1}{1 + e^{P_m(\mathbf{x}_i)}} \left(\frac{1 + e^{P_m(\mathbf{x}_i)}}{1 + e^{-P_j(\mathbf{x}_i)}} - \frac{1 + e^{P_m(\mathbf{x}_i)}}{1 + e^{-P_{j-1}(\mathbf{x}_i)}} + \frac{\lambda(\mathbf{x}_i)^{j-1}}{(j - 1)!} e^{-\lambda(\mathbf{x}_i)} \right),$$

$j = 2, \dots, m$, the log-likelihood becomes

$$\begin{aligned}
 & l(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m, \boldsymbol{\gamma}) \\
 &= - \sum_{i=1}^n \log(1 + e^{\mathbf{x}_i \boldsymbol{\beta}_m}) + \sum_{i=1}^{n_1} \log \left\{ \frac{1 + e^{\mathbf{x}_i \boldsymbol{\beta}_m}}{1 + e^{-\mathbf{x}_i \boldsymbol{\beta}_1}} + \exp(-e^{\mathbf{x}_i \boldsymbol{\gamma}}) \right\} \\
 &+ \sum_{j=2}^m \sum_{i=n_{j-1}+1}^{n_j} \log \left(\frac{1 + e^{\mathbf{x}_i \boldsymbol{\beta}_m}}{1 + e^{-\mathbf{x}_i \boldsymbol{\beta}_j}} - \frac{1 + e^{\mathbf{x}_i \boldsymbol{\beta}_m}}{1 + e^{-\mathbf{x}_i \boldsymbol{\beta}_{j-1}}} \right. \\
 &\quad \left. + \frac{\exp\{(j-1)(\mathbf{x}_i \boldsymbol{\gamma})\}}{\exp[e^{\mathbf{x}_i \boldsymbol{\gamma}} + \log\{(j-1)!]\}} \right) \\
 &+ \sum_{i=n_m+1}^n \{y_i \mathbf{x}_i \boldsymbol{\gamma} - e^{\mathbf{x}_i \boldsymbol{\gamma}} - \log(y_i!)\}.
 \end{aligned} \tag{2.10}$$

Now, for the nominal MIP and for any non-negative integers c_1, \dots, c_m , we rearrange the sample like ordinal MIP and define

$$Y_i = \begin{cases} c_1, & i = 1, \dots, n_1, \\ c_2, & i = n_1 + 1, \dots, n_2, \\ \vdots & \\ c_m, & i = n_{m-1} + 1, \dots, n_m, \text{ and} \\ y_i, & i = n_m + 1, \dots, n. \end{cases} \tag{2.11}$$

Here y_i 's are integers different from c_1, \dots, c_m .

As before, with respect to the models (2.7) and (2.8), the log-likelihood is given by

$$l(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m, \boldsymbol{\gamma}) = \sum_{j=1}^m \sum_{i=n_{j-1}+1}^{n_j} \log P(Y_i = c_j | \mathbf{x}_i) + \sum_{i=n_m+1}^n \log P(Y_i = y_i | \mathbf{x}_i),$$

where

$$P(Y_i = c_j | \mathbf{x}_i) = \pi_j(\mathbf{x}_i) + \pi_0(\mathbf{x}_i) \frac{\lambda(\mathbf{x}_i)^{c_j}}{c_j!} e^{-\lambda(\mathbf{x}_i)}, \quad j = 1, \dots, m,$$

$$P(Y_i = y_i | \mathbf{x}_i) = \pi_0(\mathbf{x}_i) \frac{\lambda(\mathbf{x}_i)^{y_i}}{y_i!} e^{-\lambda(\mathbf{x}_i)}, \quad y_i \neq c_1, \dots, c_m,$$

and $\pi_j(\mathbf{x}_i)$, $j = 1, \dots, m$, and $\lambda(\mathbf{x}_i)$ can be obtained from the equations (2.7) and (2.8), respectively.

Considering the following reformulation

$$P(Y_i = c_j | \mathbf{x}_i) = \pi_0(\mathbf{x}_i) \left[\frac{\pi_j(\mathbf{x}_i)}{\pi_0(\mathbf{x}_i)} + \frac{\lambda(\mathbf{x}_i)^{c_j}}{c_j!} e^{-\lambda(\mathbf{x}_i)} \right], \quad j = 1, \dots, m,$$

and using $1 = \frac{1}{\pi_0(\mathbf{x}_i)} - \sum_{j=1}^m \frac{\pi_j(\mathbf{x}_i)}{\pi_0(\mathbf{x}_i)}$, $\frac{\pi_j(\mathbf{x}_i)}{\pi_0(\mathbf{x}_i)} = e^{\mathbf{x}_i \boldsymbol{\beta}_j}$ from equations (2.6) and (2.7), and $\lambda(\mathbf{x}_i) = e^{\mathbf{x}_i \boldsymbol{\gamma}}$ from the equation (2.8), the log-likelihood reduces to

$$\begin{aligned}
 l(\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m, \boldsymbol{\gamma}) &= - \sum_{i=1}^n \log \left(1 + \sum_{j=1}^m e^{\mathbf{x}_i \boldsymbol{\beta}_j} \right) \\
 &\quad + \sum_{j=1}^m \sum_{i=n_{j-1}+1}^{n_j} \log \left\{ e^{\mathbf{x}_i \boldsymbol{\beta}_j} + \frac{1}{c_j!} \exp(c_j \mathbf{x}_i \boldsymbol{\gamma} - e^{\mathbf{x}_i \boldsymbol{\gamma}}) \right\} \quad (2.12) \\
 &\quad + \sum_{i=n_m+1}^n \{ y_i \mathbf{x}_i \boldsymbol{\gamma} - e^{\mathbf{x}_i \boldsymbol{\gamma}} - \log(y_i!) \}.
 \end{aligned}$$

When $m = 1$ and $c_1 = 0$ we observe that both ordinal MIP and nominal MIP reduce to ZIP model

$$\begin{aligned}
 l(\boldsymbol{\beta}_1, \boldsymbol{\gamma}) &= \sum_{i=1}^n \log \{ 1 - \pi_1(\mathbf{x}_i) \} + \sum_{i=1}^{n_1} \log \left\{ \frac{\pi_1(\mathbf{x}_i)}{1 - \pi_1(\mathbf{x}_i)} + e^{-\lambda(\mathbf{x}_i)} \right\} \\
 &\quad + \sum_{i=n_1+1}^n \{ y_i \log \lambda(\mathbf{x}_i) - \lambda(\mathbf{x}_i) - \log y_i! \},
 \end{aligned}$$

where $\pi_1(\mathbf{x}_i) = P(Y_i = 0 | \mathbf{x}_i)$ and $\lambda(\mathbf{x}_i)$ is the mean of Poisson distribution.

It should be noted that the log-likelihood of the nominal MIP (2.12) allows excessive frequencies at non-consecutive as well as consecutive non-negative integers and models the ratio of Poisson distribution and nominal multinomial distribution, while the log-likelihood of the ordinal MIP (2.10) deals with the Poisson distribution and excessive frequencies at consecutive non-negative integer values.

2.3 Bayesian inference

By utilizing available prior information about unknown parameters we seek a fully Bayesian approach. To avoid computational difficulties, we conveniently use latent variables for the link functions of the models and use multivariate Gaussian priors for regression coefficients.

Specifically, under the ordinal MIP model we consider latent variables U and V for (2.3) and (2.4) with (h_1, \dots, h_m) numbers of observations corresponding to the responses $(0, \dots, m - 1)$, respectively, such that

$$\begin{aligned}
 U_{ij} &= \mathbf{x}_i \boldsymbol{\beta}_j + \varepsilon_{ij}, & i = n_{j-1} + 1, \dots, n_{j+1}, j = 1, \dots, m - 1, \\
 U_{im} &= \mathbf{x}_i \boldsymbol{\beta}_m + \varepsilon_{ij}, & i = 1, \dots, n, \\
 V_i &= \mathbf{x}_i \boldsymbol{\gamma} + \delta_i, & i = 1, \dots, n,
 \end{aligned} \quad (2.13)$$

where $n_j = \sum_{k=1}^j h_k$ are the numbers of excessive integers less than j and $n_0 = h_0 = 0$ as in (2.9), ε_{ij} and δ_i are mutually independent jittering variables and

follow normal distributions with zero mean and small variances, $\varepsilon_{ij} \overset{\text{ind}}{\sim} N(0, \sigma_{u_j}^2)$ and $\delta_i \overset{\text{ind}}{\sim} N(0, \sigma_v^2)$. Here, by the cumulative logit link function, we consider a constraint $U_{ij} \leq U_{ij'}$, for $j < j'$.

Under the nominal MIP, we use latent variables Z for (2.7) as well as W for (2.8)

$$\begin{aligned} Z_{ij} &= \mathbf{x}_i \boldsymbol{\beta}_j + \xi_{ij}, & i = 1, \dots, n, j = 1, \dots, m, \\ W_i &= \mathbf{x}_i \boldsymbol{\gamma} + \eta_i, & i = 1, \dots, n, \end{aligned} \tag{2.14}$$

where ξ_{ij} and η_i are mutually independent jittering variables and follow normal distributions with zero mean and small variances, $\xi_{ij} \overset{\text{ind}}{\sim} N(0, \sigma_{z_j}^2)$ and $\eta_i \overset{\text{ind}}{\sim} N(0, \sigma_w^2)$.

For regression coefficients $\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m$ and $\boldsymbol{\gamma}$ in (2.3) and (2.4) for ordinal MIP and in (2.7) and (2.8) for nominal MIP, we use multivariate Gaussian priors expressed as

$$\boldsymbol{\beta}_j \overset{\text{ind}}{\sim} \text{MVN}[\mathbf{0}, B_b^2 \mathbf{I}], \quad j = 1, \dots, m \quad \text{and} \quad \boldsymbol{\gamma} \sim \text{MVN}[\mathbf{0}, B_c^2 \mathbf{I}],$$

where $\mathbf{0}$ is the zero vector of size $(p + 1)$ and \mathbf{I} is $(p + 1) \times (p + 1)$ identity matrix. We consider large values for B_b^2 and B_c^2 to get flexible priors. Further, hierarchically, we assign conjugate inverse gamma hyperpriors on the variances of latent variables: $\sigma_{u_j}^2 \overset{\text{ind}}{\sim} \text{IG}[A_{u_j}, B_{u_j}]$, $\sigma_v^2 \sim \text{IG}[A_v, B_v]$, $\sigma_{z_j}^2 \overset{\text{ind}}{\sim} \text{IG}[A_{z_j}, B_{z_j}]$ and $\sigma_w^2 \sim \text{IG}[A_w, B_w]$.

Now, using the log-likelihoods in (2.10) and (2.12) and specified priors we can find the full conditional distributions of parameters as well as latent variables. Under the ordinal MIP, in the [Appendix](#) we have the full conditional distributions of latent variables U and V are presented in (A.1), the full conditional distributions of regression coefficients are presented in (A.2) and the full conditional distributions of nuisance parameters are presented in (A.3).

Under the nominal MIP presented the full conditional distributions of latent variables Z and W are in (A.4), the full conditional distributions of regression coefficients are presented in (A.5) and the full conditional distributions of nuisance parameters are presented in (A.6).

Now, using the full conditional distributions and the Metropolis-within-Gibbs algorithm one can collect Markov Chain Monte Carlo (MCMC) samples of parameters and latent variables from their joint posterior distribution. More details are given the next section.

3 Simulation and application to infection data

3.1 Simulation studies

We carried out simulations to compare the performance of multiple-inflation Poisson (MIP) regression model with zero-inflated Poisson (ZIP) regression model.

With a covariate ($p = 1$) which is randomly selected number between zero and one, that is, $x_i \in [0, 1]$, $i = 1, \dots, n$. Using the cumulative logistic distribution function $H(v) = 1/(1 + e^{-v})$, we also generated the corresponding response Y_i according to the following three true models.

M0. Excessive frequency of response only at 0:

$$Y_i|x_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, \pi_1(x_i), \lambda(x_i)],$$

$$H^{-1}\{\pi_1(x_i)\} = -(\log 4) - (\log 4)x_i,$$

$$\log\{\lambda(x_i)\} = -(\log 3) + (\log 6)x_i.$$

M1. Excessive frequencies at 0 and 1 with parallel logistic functions:

$$Y_i|x_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1; \pi_1(x_i), \pi_2(x_i); \lambda(x_i)],$$

$$H^{-1}\{\pi_1(x_i)\} = -(\log 4) - (\log 4)x_i,$$

$$H^{-1}\{\pi_2(x_i)\} = -(\log 4)x_i,$$

$$\log\{\lambda(x_i)\} = -(\log 3) + (\log 6)x_i.$$

M2. Excessive frequencies at 0 and 1 with non-parallel logistic functions:

$$Y_i|x_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1; \pi_1(x_i), \pi_2(x_i); \lambda(x_i)],$$

$$H^{-1}\{\pi_1(x_i)\} = -(\log 4)x_i,$$

$$H^{-1}\{\pi_2(x_i)\} = -(\log 4) + (\log 2)x_i,$$

$$\log\{\lambda(x_i)\} = -(\log 2) + (\log 6)x_i.$$

Here 200 simulation data sets with 100 sample size ($n = 100$) under each model were generated. Figure 1 shows a set of simulated data under M1 with the true regression function $m(x) = E(Y|x)$.

To fit the simulated data, we used the ordinal and nominal MIP regressions under the Bayesian framework, and as a golden standard used the ZIP regression by considering the excessive frequency only at 0 such that

- ZIP model:

$$Y_i|x_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, \pi_1(x_i), \lambda(x_i)], \quad i = 1, \dots, n,$$

where $\log(\frac{\pi_1(x_i)}{1-\pi_1(x_i)}) = \beta_{0,1} + \beta_{1,1}x_i$ and $\log(\lambda(x_i)) = \gamma_0 + \gamma_1x_i$.

- Ordinal MIP model:

$$Y_i|x_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1, \pi_1(x_i), \pi_2(x_i), \lambda(x_i)], \quad i = 1, \dots, n,$$

where $\log(\frac{\pi_1(x_i)}{1-\pi_1(x_i)}) = \beta_{0,1} + \beta_{1,1}x_i$, $\log(\frac{\pi_1(x_i)+\pi_2(x_i)}{1-\pi_1(x_i)-\pi_2(x_i)}) = \beta_{0,2} + \beta_{1,2}x_i$, $\beta_{1,1} = \beta_{1,2}$ and $\log\{\lambda(x_i)\} = \gamma_0 + \gamma_1x_i$.

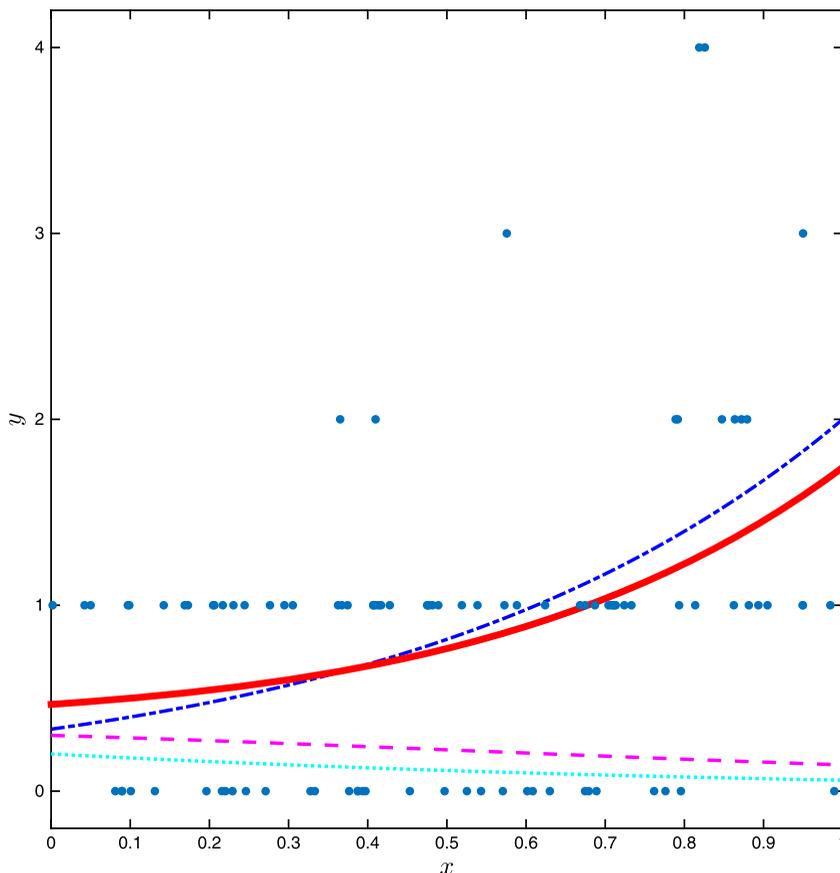


Figure 1 Simulation data. The red solid line indicates regression function $m(x)$; the pink blocked line and light blue dotted line show the excess frequency probabilities, $\pi_1(x)$ and $\pi_2(x)$, respectively; the blue block-dotted line describes the Poisson mean function $\lambda(x)$; and the blue dots indicate the count responses. The figure shows that $\lambda(x)$ is adjusted by $\pi_1(x)$ and $\pi_2(x)$ to produce $m(x)$.

- Nominal MIP model:

$$Y_i | x_i \stackrel{\text{ind}}{\sim} \text{MIP}[c_1, c_2, \pi_1(x_i), \pi_2(x_i), \lambda(x_i)], \quad i = 1, \dots, n,$$

where c_1 and c_2 are responses with excessive frequencies, $\log\left(\frac{\pi_1(x_i)}{\pi_0(x_i)}\right) = \beta_{0,1} + \beta_{1,1}x_i$, $\log\left(\frac{\pi_2(x_i)}{\pi_0(x_i)}\right) = \beta_{0,2} + \beta_{1,2}x_i$, $\pi_0(x_i) = 1 - \pi_1(x_i) - \pi_2(x_i)$ and $\log\{\lambda(x_i)\} = \gamma_0 + \gamma_1x_i$.

We considered the independent non-informative prior distributions for regression coefficients with large variances $\beta_{i,j} \stackrel{\text{i.i.d.}}{\sim} N(0, 100)$, $i = 0, 1$; $j = 1, 2$, and $\gamma_i \stackrel{\text{i.i.d.}}{\sim} N(0, 100)$, $i = 0, 1$, and assigned conjugate inverse gamma prior distributions for

the variance of the latent variables in (2.13) and (2.14) such that $\sigma_{u_j}^2 \sim \text{IG}[10, 10]$, $\sigma_v^2 \sim \text{IG}[10, 10]$, $\sigma_{z_j}^2 \sim \text{IG}[10, 10]$ and $\sigma_w^2 \sim \text{IG}[10, 10]$, for $j = 1, 2$.

From Lemma 1, it is clear that, for $\pi_0(x) = 1 - \pi_1(x) - \pi_2(x)$,

$$m(x) = E(Y|x) = \begin{cases} \left\{ \begin{array}{l} \{1 - \pi_1(x)\}\lambda(x) = \frac{4^{x+1}6^x}{3(1 + 4^{x+1})}, \\ \text{under M0,} \\ \pi_2(x) + \pi_0(x)\lambda(x) = \frac{3(1 + 4^{x+1}) + 6^x(4^{2x+1} - 1)}{3(1 + 4^{x+1})(1 + 4^x)}, \\ \text{under M1,} \\ \pi_2(x) + \pi_0(x)\lambda(x) = \frac{2(1 + 4^x) + 6^x(2^{x+2} - 1)}{2(1 + 4^x)(1 + 2^{-x+2})}, \\ \text{under M2.} \end{array} \right. \end{cases}$$

Using the Bayes estimates of regression coefficients, we can estimate $\pi_1(x)$, $\pi_2(x)$ and $\lambda(x)$ in the ZIP and MIP models. The estimates of the rate of Poisson distribution is directly calculated by $\hat{\lambda}(x) = e^{\hat{\gamma}_0 + \hat{\gamma}_1 x}$. The estimates of $\pi_1(x)$ and $\pi_2(x)$ vary in the model considered. Let $\hat{\rho}_j(x_i) = \hat{\beta}_{0,j} + \hat{\beta}_{1,j}x_i$, $j = 1, 2$, denote the estimated regression lines by the estimated coefficients. Under the ZIP model, the estimate of the excessive frequency is directly calculated by $\hat{\pi}_1(x) = 1/\{1 + e^{-\hat{\rho}_1(x)}\}$ with the assumption of $\hat{\pi}_2(x) \equiv 0$. Under the ordinal MIP model, the estimate $\hat{\pi}_1(x)$ is calculated as in ZIP model and the estimate of the second excessive frequency is calculated by $\hat{\pi}_2(x) = 1/\{1 + e^{-\hat{\rho}_2(x)}\} - \hat{\pi}_1(x)$. Under the nominal MIP model, the first excessive frequency is estimated by $\hat{\pi}_1(x) = e^{\hat{\rho}_1(x)}/\{1 + e^{\hat{\rho}_1(x)} + e^{\hat{\rho}_2(x)}\}$ and the second excessive frequency is estimated by $\hat{\pi}_2(x) = e^{\hat{\rho}_2(x)}/\{1 + e^{\hat{\rho}_1(x)} + e^{\hat{\rho}_2(x)}\}$, respectively. The estimated regression lines $\hat{m}(x)$ are evaluated by the estimated $\hat{\lambda}(x)$, $\hat{\pi}_1(x)$ and $\hat{\pi}_2(x)$, for each model. Figure 2 presents the convergence of MCMC samples for regression coefficients in a simulation.

Now to see how good our method of estimation is we computed the mean integrated square error (MISE),

$$\text{MISE} = \int_0^1 \{\hat{m}(x) - m(x)\}^2 dx.$$

Table 2 shows the resulted MISEs and squared biases of the simulations, $\text{Bias}^2 = \int_0^1 \{\tilde{m}(x) - m(x)\}^2 dx$, where $\tilde{m}(x)$ is the mean curve of estimated regression lines $\hat{m}(x)$ over simulations. From Table 2, we observe that ZIP has the minimum MISE under M0, ordinal MIP has the minimum MISE under M1 and nominal MIP has the minimum MISE under M2, respectively. In addition, we find that ordinal and nominal MIP regressions have comparable MISE with ZIP regression, while they have smaller biases than ZIP regression. This is not surprising, because MIP regressions use more parameters to interpret the data. Also, the ordinal MIP regres-

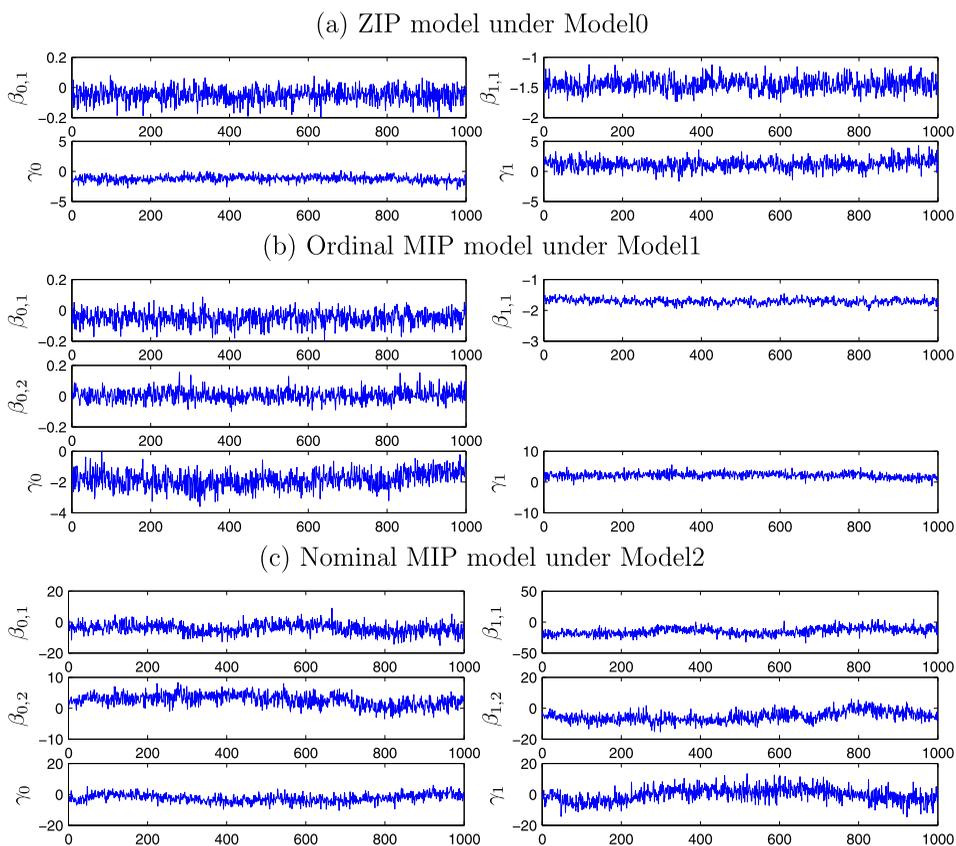


Figure 2 MCMC samples for regression coefficients in a simulation. After 10,000 burning iterations, 1000 MCMC samples show convergence. In (a), the ZIP model only has one excessive frequency and the plots in the first row display the MCMC samples of regression coefficients. In (b), the ordinal MIP model shares the slope with different intercepts and three plots in the first two rows present the MCMC samples of corresponding regression coefficient. In (c), the nominal MIP model allows different intercepts and different slope and the four plots in the first two row are the MCMC samples for regression coefficients. In (a), (b) and (c), two plots in the last row are regression coefficients for the rate of Poisson distribution.

sion showed slightly less bias than the nominal MIP regression. This is again understandable because the ordinal MIP model shares the slope of excess probabilities of zeros and ones and the ZIP model does not have excess probability of ones.

3.2 Application to infection data

Acinetobacter baumannii bacteria (*A. baumannii*) infection is an important problem in intensive care units in the hospitals as a leading agent of hospital infection. Among many studies related to *A. baumannii*, there was a large prospective cohort study conducted to explain its contamination routes of infections at five large

Table 2 Simulation results for MISEs and variances of regression lines

	Method	M0	M1	M2
MISE	ZIP	0.0418	0.0382	0.0957
	Ordinal MIP	0.0450	0.0360	0.0938
	Nominal MIP	0.0465	0.0372	0.0692
Bias ²	ZIP	0.0087	0.0196	0.0619
	Ordinal MIP	0.0018	0.0033	0.0398
	Nominal MIP	0.0087	0.0052	0.0312

Table 3 Weekly number of infections and covariates for 10 weeks

Infection(Y_i)	Age(x_{i1})	Apache(x_{i2})	Gender(x_{i3})	IMV(x_{i4})	ANT(x_{i5})
1	0.6296	0.4444	0.5185	0.4074	0.1111
0	0.6774	0.3871	0.4839	0.4194	0.0968
0	0.6571	0.3714	0.4571	0.4571	0.1143
0	0.6410	0.3590	0.4872	0.4103	0.1026
1	0.6279	0.3256	0.4651	0.4419	0.0930
2	0.6531	0.3673	0.4898	0.4082	0.0816
3	0.6316	0.3860	0.5088	0.4386	0.0702
1	0.6190	0.3968	0.4762	0.4286	0.0794
1	0.6119	0.4328	0.4925	0.4179	0.0746
0	0.6234	0.4545	0.5195	0.4416	0.0779

medical centers in Turkey namely Hacettepe University Hospital, Istanbul University Cerrahpaşa Medical School, GATA Haydarpaşa Hospital, Marmara University Hospital and Uludağ University Medical School.

The infection data consist of the response, the number of infected individuals per week, as well as several covariates, including Age (ratio of patients older than or equal to 65), Gender (ratio of female patients), Apache (ratio of patients with higher than or equal to 20 Apache score, where the higher the score higher the risk for an infection), IMV (ratio of patients with invasive mechanical ventilation) and ANT (ratio of patients with antibiotic use three months before admission). The number of infected patients and the covariate values are recorded for 49 weeks. In Table 3, we only give the data for the 10 weeks.

Let $Y_i, i = 1, \dots, n = 49$, denote the weekly number of infected patients. We assume that Y_i is affected by the 5 ratios during the week, that is, we have $(p + 1) \times 1$ vector of covariates x_i with $p = 5$. For example, in the first row in Table 3, we assume that one infection recorded during a week can be interpreted by the situation of intensive care unit where 62.96% of patients were older than or equal to 65, 44.44% had their Apache score more than or equal to 20, 51.85% of patients were female, 40.74% had IMV and 11.11% had antibiotic three months before admission. Comparing the histogram of infection data in Figure 3 and the probability

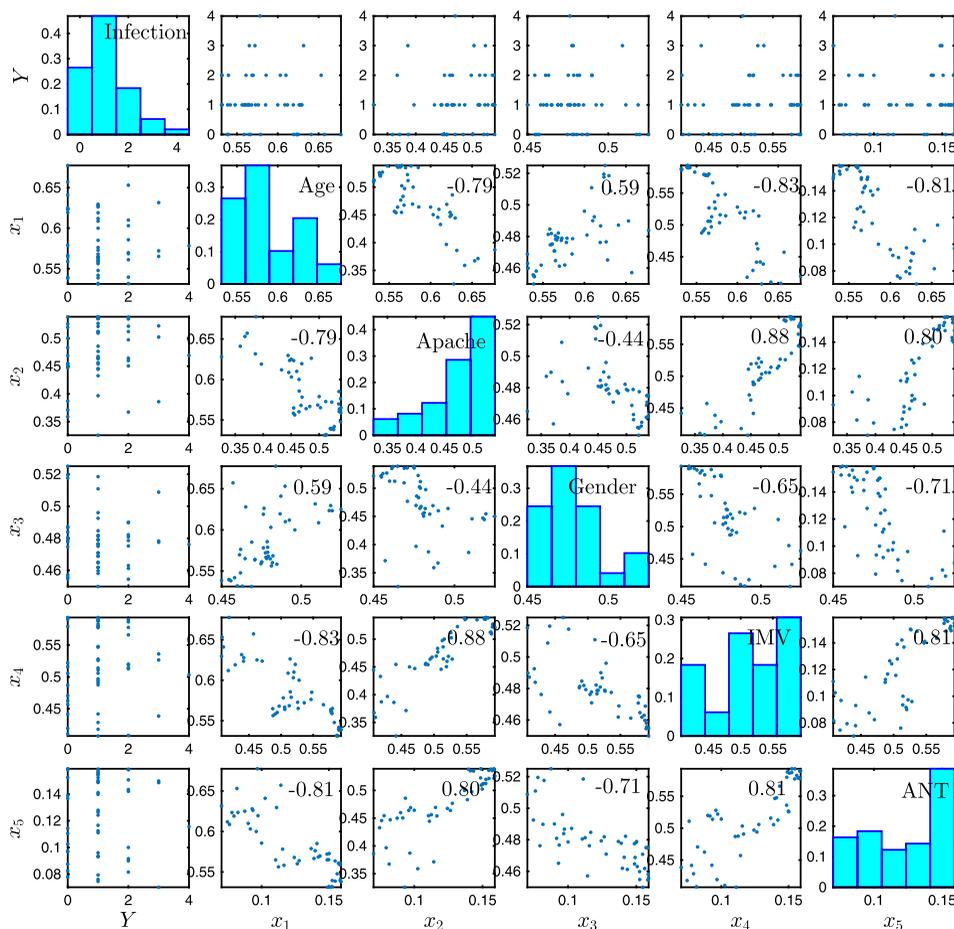


Figure 3 Matrix plot for infection data. Plots in diagonal are histograms of variables and plots in off-diagonal are scatterplots of corresponding pairs of variables with the sample Pearson correlation coefficients denoted inside of plots.

mass function of Poisson distribution which has the similar probabilities at values 2, 3 and 4 with the histogram, we observe excessive zeros and ones in infection data, and highly correlated covariates. In the sense of Higdon et al. (2008), we use principal components as predictors instead of highly correlated covariates. We first perform the principal component analysis, and take the first principal axis ϕ_i which explains 87.13% of total variation with the following loading equation:

$$\phi_i = -0.3758x_{i1} + 0.5893x_{i2} - 0.1266x_{i3} + 0.6396x_{i4} + 0.2939x_{i5},$$

for $i = 1, \dots, n$.

With ϕ_i , we fitted ZIP model by assuming the excessive frequency only at zero, the ordinal MIP model by assuming the excessive frequencies at zero and

one and the nominal MIP model which is applicable to both consecutive and non-consecutive values with excessive frequencies. More specifically, denoting $H(v) = 1/(1 + e^{-v})$ as the cumulative distribution function of the logistic distribution, and using the principal axis instead of 5 covariates, three models for the infection data are reduced to the followings.

- ZIP model for data:

$$Y_i | \phi_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, \pi_1(\phi_i), \lambda(\phi_i)], \quad i = 1, \dots, n,$$

where $\log(\frac{\pi_1(\phi_i)}{1-\pi_1(\phi_i)}) = \beta_{0,1} + \beta_{1,1}\phi_i$ and $\log\{\lambda(\phi_i)\} = \gamma_0 + \gamma_1\phi_i$.

- Ordinal MIP model for data:

$$Y_i | \phi_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1; \pi_1(\phi_i), \pi_2(\phi_i); \lambda(\phi_i)], \quad i = 1, \dots, n,$$

where $\log(\frac{\pi_1(\phi_i)}{1-\pi_1(\phi_i)}) = \beta_{0,1} + \beta_{1,1}\phi_i$, $\log(\frac{\pi_1(\phi_i)+\pi_2(\phi_i)}{1-\pi_1(\phi_i)-\pi_2(\phi_i)}) = \beta_{0,2} + \beta_{1,1}\phi_i$ and $\log\{\lambda(\phi_i)\} = \gamma_0 + \gamma_1\phi_i$.

- Nominal MIP model for data:

$$Y_i | \phi_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1; \pi_1(\phi_i), \pi_2(\phi_i); \lambda(\phi_i)], \quad i = 1, \dots, n,$$

where c_1 and c_2 are responses with excessive frequencies, $\log(\frac{\pi_1(\phi_i)}{\pi_0(\phi_i)}) = \beta_{0,1} + \beta_{1,1}\phi_i$, $\log(\frac{\pi_2(\phi_i)}{\pi_0(\phi_i)}) = \beta_{0,2} + \beta_{1,2}\phi_i$, $\pi_0(\phi_i) = 1 - \pi_1(\phi_i) - \pi_2(\phi_i)$ and $\log\{\lambda(\phi_i)\} = \gamma_0 + \gamma_1\phi_i$.

We assign flat Gaussian priors for regression coefficients with large variances such that $\beta_{i,j} \stackrel{\text{i.i.d.}}{\sim} N(0, 100)$ and $\gamma_i \stackrel{\text{i.i.d.}}{\sim} N(0, 100)$, for $i = 0, 1$ and $j = 1, 2$, and conjugate inverse Gamma priors for the variance of the latent variables in (2.13) and (2.14) such that $\sigma_{u_j}^2 \sim \text{IG}[10, 10]$, $\sigma_v^2 \sim \text{IG}[0.1, 0.1]$, $\sigma_{z_j}^2 \sim \text{IG}[10, 10]$ and $\sigma_w^2 \sim \text{IG}[0.1, 0.1]$, $j = 1, 2$. Now, through the Metropolis-within-Gibbs algorithm we collect 10,000 Markov Chain Monte Carlo samples from the joint posterior distribution after 50,000 burn-in iterations. Denoting the standard errors as the small numbers in parentheses under the estimated coefficients, the estimated models are

- Estimated ZIP model for data:

$$Y_i | \phi_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, \hat{\pi}_1(\phi_i), \hat{\lambda}(\phi_i)],$$

$$\log\left(\frac{\hat{\pi}_1(\phi_i)}{1 - \hat{\pi}_1(\phi_i)}\right) = \begin{matrix} -0.0008 & - 2.5157\phi_i, \\ (0.0231) & (0.0970) \end{matrix}$$

$$\log\{\hat{\lambda}(\phi_i)\} = \begin{matrix} -0.3508 & + 0.6894\phi_i. \\ (1.1028) & (2.8411) \end{matrix}$$

- Estimated ordinal MIP model for data:

$$Y_i | \phi_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1; \hat{\pi}_1(\phi_i), \hat{\pi}_2(\phi_i); \hat{\lambda}(\phi_i)],$$

$$\log\left(\frac{\hat{\pi}_1(\phi_i)}{1 - \hat{\pi}_1(\phi_i)}\right) = \begin{matrix} -0.0078 & - & 2.8041\phi_i, \\ (0.0391) & & (0.1267) \end{matrix}$$

$$\log\left(\frac{\hat{\pi}_1(\phi_i) + \hat{\pi}_2(\phi_i)}{1 - \hat{\pi}_1(\phi_i) - \hat{\pi}_2(\phi_i)}\right) = \begin{matrix} 0.0238 & - & 2.8041\phi_i, \\ (0.0282) & & (0.1267) \end{matrix}$$

$$\log\{\hat{\lambda}(\phi_i)\} = \begin{matrix} -0.4549 & + & 0.8339\phi_i. \\ (1.1636) & & (3.0293) \end{matrix}$$

- Estimated nominal MIP model for data:

$$Y_i | \phi_i \stackrel{\text{ind}}{\sim} \text{MIP}[0, 1; \hat{\pi}_1(\phi_i), \hat{\pi}_2(\phi_i); \hat{\lambda}(\phi_i)],$$

$$\log\left(\frac{\hat{\pi}_1(\phi_i)}{\hat{\pi}_0(\phi_i)}\right) = \begin{matrix} 0.5001 & - & 4.5883\phi_i, \\ (0.1827) & & (0.3792) \end{matrix}$$

$$\log\left(\frac{\hat{\pi}_2(\phi_i)}{\hat{\pi}_0(\phi_i)}\right) = \begin{matrix} -0.9605 & + & 2.2203\phi_i, \\ (0.2337) & & (0.5532) \end{matrix}$$

$$\log\{\hat{\lambda}(\phi_i)\} = \begin{matrix} -0.5762 & + & 1.0656\phi_i, \\ (1.7326) & & (4.5291) \end{matrix}$$

where $\hat{\pi}_0(\phi_i) = 1 - \hat{\pi}_1(\phi_i) - \hat{\pi}_2(\phi_i)$.

As shown in Figure 4, the estimated regression coefficients and mean curves of three models indicate that there may be an increasing trend of number of infections on the principal axis ϕ , which has positive loadings on Apache, IMV and ANT and has negative loadings on Age and Gender, without statistical significance. Whereas, the excess probability of zeros is significantly decreasing on all tree model along with ϕ . Interestingly, the excess probability of ones is significantly increasing in the nominal MIP model while its excess probability of zeros is more rapidly decreased than other two models.

4 Conclusion and discussion

In this article, we have proposed and studied properties of multiple-inflation Poisson (MIP) distributions and MIP regression models that incorporate covariates. Our proposed models are the generalization of zero inflated Poisson by considering the mixture of multinomial and Poisson distributions.

We describe a Bayesian approach to make statistical inference on parameters of MIP regression models. In a Frequentist approach, using the expectation-maximization (EM) algorithm would confront computational difficulties to determine multiple sets of indicators which present the data from the zero state or from

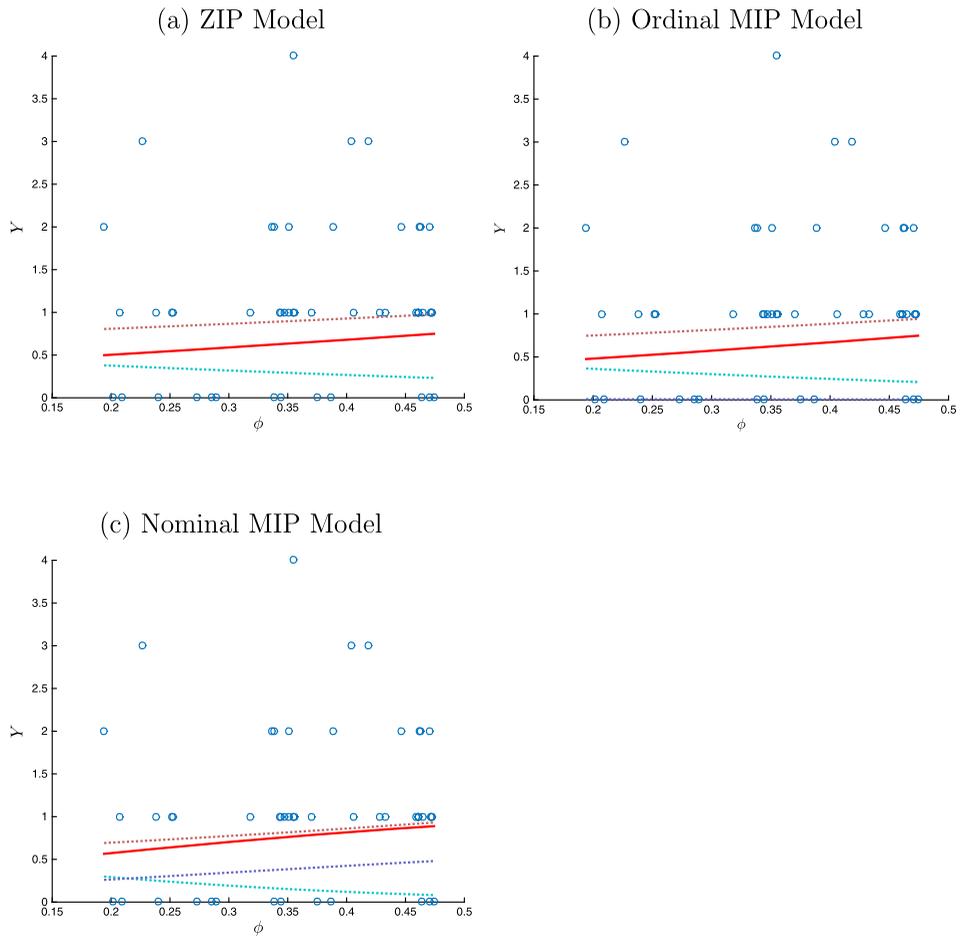


Figure 4 Fitted mean curves for infection data. Circles denote the number of infections on the principal component axis ϕ . Dotted lines are fitted components in the ZIP or MIP models; red dotted lines for Poisson components, light blue dotted lines for probability for excess zeros, and dark blue dotted lines for the excess probability for excess ones. The solid red lines are estimated mean curves by aggregating the components.

the Poisson state. However, by utilizing a data augmentation method in Bayesian models, we can successfully made statistical inference feasible.

Appendix: Full conditional distributions

In this Appendix, we present full conditional distributions for the latent variables and parameters in MIP models. Using the Bayes rule, the likelihood of data, distribution of latent variables and the prior distributions for parameters in the model lead to a joint posterior distribution of parameters and latent variables. By virtue

of the Gibbs sampler, we generate Markov Chain Monte Carlo (MCMC) samples for parameters and latent variables from the joint posterior distribution by using the full conditional distributions of parameters and latent variables instead of using the joint posterior distribution. Considering the full conditional distribution of the latent variables we have $m + 1$ building blocks of log-likelihood according to the responses with excessive frequencies as the two blocks in the ZIP regression models (Lambert, 1992).

Under the ordinal MIP, the likelihood of data (2.10), distributions of latent variables (2.13) and the prior distributions for parameters in the MIP model lead to the joint posterior distribution of parameters and latent variables. Let $\theta^O = (\beta_1, \dots, \beta_m, \gamma, \sigma_{u_1}^2, \dots, \sigma_{u_m}^2, \sigma_v^2)$ denote all parameters in the model, then the joint posterior distribution is expressed by

$$\begin{aligned}
 &P(\theta^O, U_{11}, \dots, U_{nm}, V_1, \dots, V_n | Y_1, \dots, Y_n, \mathbf{x}_1, \dots, \mathbf{x}_n) \\
 &\propto P(\theta^O) \prod_{i=1}^n P(Y_i | \mathbf{x}_i, U_{i1}, \dots, U_{im}, V_i, \theta^O) P(V_i | \sigma_v^2, \gamma, \mathbf{x}_i) \\
 &\quad \times \prod_{j=1}^m P(U_{ij} | \sigma_{u_j}^2, \beta_j, \mathbf{x}_i),
 \end{aligned}$$

where $P(\theta^O)$ is the prior distribution of parameters. Denoting $U_{ij}^* = e^{U_{ij}}$ and $V_i^* = e^{V_i}$ in the models (2.13), we have the following logarithms of full conditional distributions of latent variables U_{ij} and $V_i, i = 1, \dots, n, j = 1, \dots, m$, from the joint posterior distribution such that

$$\begin{aligned}
 &\log P(U_{i1} | \cdot) \\
 &\propto \log N(\mathbf{x}_i \beta_1, \sigma_{u_1}^2), \quad i = 1, \dots, n_2, \\
 &\quad + \begin{cases} \log\left(\frac{1 + U_{im}^*}{1 + U_{i1}^{*-1}} + e^{-V_i^*}\right), & i = 1, \dots, n_1, \\ \log\left(\frac{1 + U_{im}^*}{1 + U_{i2}^{*-1}} - \frac{1 + U_{im}^*}{1 + U_{i1}^{*-1}} + \frac{V_i^*}{e^{V_i^*}}\right), & i = n_1 + 1, \dots, n_2, \end{cases} \\
 &\log P(U_{ij} | \cdot) \\
 &\propto \log N(\mathbf{x}_i \beta_j, \sigma_{u_j}^2), \quad i = n_{j-1} + 1, \dots, n_{j+1}, \\
 &\quad + \begin{cases} \log\left(\frac{1 + U_{im}^*}{1 + U_{ij}^{*-1}} - \frac{1 + U_{im}^*}{1 + U_{i,j-1}^{*-1}} + \frac{V_i^{*j-1}}{e^{V_i^*} (j-1)!}\right), & \\ i = n_{j-1} + 1, \dots, n_j, \\ \log\left(\frac{1 + U_{im}^*}{1 + U_{i,j+1}^{*-1}} - \frac{1 + U_{im}^*}{1 + U_{ij}^{*-1}} + \frac{V_i^{*j}}{e^{V_i^*} j!}\right), & \\ i = n_j + 1, \dots, n_{j+1}, j = 2, \dots, m - 1, \end{cases} \tag{A.1}
 \end{aligned}$$

$$\begin{aligned} & \log P(U_{im}|\cdot) \\ & \propto \log N(\mathbf{x}_i \boldsymbol{\beta}_m, \sigma_{u_m}^2) - \log(1 + U_{im}^*), \quad i = 1, \dots, n, \\ & + \begin{cases} \log\left(\frac{1 + U_{im}^*}{1 + U_{i1}^{*-1}} + e^{-V_i^*}\right), \\ \quad i = 1, \dots, n_1, \\ \log\left(\frac{1 + U_{im}^*}{1 + U_{ij}^{*-1}} - \frac{1 + U_{im}^*}{1 + U_{i,j-1}^{*-1}} + \frac{V_i^{*j-1}}{e^{V_i^*} (j-1)!}\right), \\ \quad i = n_{j-1} + 1, \dots, n_j, j = 2, \dots, m, \end{cases} \end{aligned}$$

$$\begin{aligned} & \log P(V_i|\cdot) \\ & \propto \log N(\mathbf{x}_i \boldsymbol{\gamma}, \sigma_v^2), \quad i = 1, \dots, n, \\ & + \begin{cases} \log\left(\frac{1 + U_{im}^*}{1 + U_{i1}^{*-1}} + e^{-V_i^*}\right), \\ \quad i = 1, \dots, n_1, \\ \log\left(\frac{1 + U_{im}^*}{1 + U_{ij}^{*-1}} - \frac{1 + U_{im}^*}{1 + U_{i,j-1}^{*-1}} + \frac{V_i^{*j-1}}{e^{V_i^*} (j-1)!}\right), \\ \quad i = n_{j-1} + 1, \dots, n_j, j = 2, \dots, m, \\ y_i V_i - V_i^*, \\ \quad i = n_m + 1, \dots, n, \end{cases} \end{aligned}$$

where $N(\mu, \sigma^2)$ indicate the density of Normal distribution with mean μ and variance σ^2 .

For the full conditional distributions of regression coefficients, we assume different thresholds β_{0j} and common slopes $\boldsymbol{\beta}^* = (\beta_1^*, \dots, \beta_p^*)^T$ for $j = 1, \dots, m$ in (2.5). Denoting $\mathbf{U}_j = (U_{n_{j-1}+1,j}, \dots, U_{n_{j+1},j})^T$, $\mathbf{x}_i^* = (x_{i1}, \dots, x_{ip})$, $\mathbf{X}_j^* = (\mathbf{x}_{n_{j-1}+1}^{*T}, \dots, \mathbf{x}_{n_{j+1}}^{*T})^T$, $j = 1, \dots, m - 1$, $\mathbf{U}_m = (U_{1m}, \dots, U_{nm})^T$, $\mathbf{X}_m^* = (\mathbf{x}_1^{*T}, \dots, \mathbf{x}_n^{*T})^T$, $\mathbf{X} = (\mathbf{x}_1^T, \dots, \mathbf{x}_n^T)^T$ and $\mathbf{V} = (V_1, \dots, V_n)^T$ we have the full conditional distributions

$$\begin{aligned} \beta_{0j}|\cdot & \sim N\left[\left(\frac{h_j^*}{\sigma_{u_j}^2} + \frac{1}{B_b^2}\right)^{-1} \mathbf{1}_{h_j^*}^T (\mathbf{U}_j - \mathbf{X}_j^* \boldsymbol{\beta}^*), \left(\frac{h_j^*}{\sigma_{u_j}^2} + \frac{1}{B_b^2}\right)^{-1}\right], \\ & j = 1, \dots, m, \\ \boldsymbol{\beta}^*|\cdot & \sim N\left[\left(\sum_{j=1}^m \frac{\mathbf{X}_j^{*T} \mathbf{X}_j^*}{\sigma_{u_j}^2} + \frac{\mathbf{I}_p}{B_b^2}\right)^{-1} \left(\sum_{j=1}^m \frac{\mathbf{X}_j^{*T} (\mathbf{U}_j - \mathbf{1}_{h_j^*} \beta_{0j})}{\sigma_{u_j}^2}\right), \right. \\ & \left. \left(\sum_{j=1}^m \frac{\mathbf{X}_j^{*T} \mathbf{X}_j^*}{\sigma_{u_j}^2} + \frac{\mathbf{I}_p}{B_b^2}\right)^{-1}\right], \end{aligned} \tag{A.2}$$

$$\boldsymbol{\gamma}|\cdot \sim N\left[\left(\frac{1}{\sigma_v^2}\mathbf{X}^T\mathbf{X} + \frac{1}{B_c^2}\mathbf{I}_{p+1}\right)^{-1}\left(\frac{1}{\sigma_v^2}\mathbf{X}^T\mathbf{V}\right),\right. \\ \left.\left(\frac{1}{\sigma_v^2}\mathbf{X}^T\mathbf{X} + \frac{1}{B_c^2}\mathbf{I}_{p+1}\right)^{-1}\right],$$

where $h_j^* = n_{j+1} - n_{j-1} = h_j + h_{j+1}$, $j = 1, \dots, m - 1$, $h_m^* = n$, $\mathbf{1}_{h_j^*}$ is the column vector of h_j^* ones and \mathbf{I}_p is the $p \times p$ identity matrix. The full conditional distributions of nuisance parameters are also given by

$$\sigma_{u_j}^2|\cdot \sim \text{IG}\left[\frac{h_j^*}{2} + A_{u_j}, \left(\frac{\text{RSS}_{u_j}}{2} + \frac{1}{B_{u_j}}\right)^{-1}\right], \quad j = 1, \dots, m, \tag{A.3}$$

$$\sigma_v^2|\cdot \sim \text{IG}\left[\frac{n}{2} + A_v, \left(\frac{\text{RSS}_v}{2} + \frac{1}{B_v}\right)^{-1}\right],$$

where $\text{RSS}_{u_j} = (\mathbf{U}_j - \mathbf{1}_{h_j^*}\beta_{0j} - \mathbf{X}_j^*\boldsymbol{\beta}^*)^T(\mathbf{U}_j - \mathbf{1}_{h_j^*}\beta_{0j} - \mathbf{X}_j^*\boldsymbol{\beta}^*)$ and $\text{RSS}_v = (\mathbf{V} - \mathbf{X}\boldsymbol{\gamma})^T(\mathbf{V} - \mathbf{X}\boldsymbol{\gamma})$.

Similarly, under the nominal MIP, the likelihood of data (2.12), distributions of latent variables (2.14) and the prior distributions for parameters in the MIP model bring the joint posterior distribution of parameters and latent variables. Denoting $\boldsymbol{\theta}^N = (\boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_m, \boldsymbol{\gamma}, \sigma_{z_1}^2, \dots, \sigma_{z_m}^2, \sigma_w^2)$ as all parameters in the model, then the joint posterior distribution is given by

$$P(\boldsymbol{\theta}^N, Z_{11}, \dots, Z_{nm}, W_1, \dots, W_n | Y_1, \dots, Y_n, \mathbf{x}_1, \dots, \mathbf{x}_n) \\ \propto P(\boldsymbol{\theta}^N) \prod_{i=1}^n P(Y_i | \mathbf{x}_i, Z_{i1}, \dots, Z_{im}, W_i, \boldsymbol{\theta}^N) \\ \times P(W_i | \sigma_w^2, \boldsymbol{\gamma}, \mathbf{x}_i) \\ \times \prod_{j=1}^m P(Z_{ij} | \sigma_{z_j}^2, \boldsymbol{\beta}_j, \mathbf{x}_i),$$

where $P(\boldsymbol{\theta}^N)$ is the prior distribution of parameters. Denoting $Z_{ij}^* = e^{Z_{ij}}$ and $W_i^* = e^{W_i}$ in the models (2.14), the logarithms of full conditional distributions of latent variables Z_{ij} and W_i , $i = 1, \dots, n$, $j = 1, \dots, m$ are given by

$$\log P(Z_{ij}|\cdot) \propto \log N(\mathbf{x}_i\boldsymbol{\beta}_j, \sigma_{z_j}^2) - \log\left(1 + \sum_{j=1}^m Z_{ij}^*\right), \quad i = 1, \dots, n, \\ + \log\left\{Z_{ij}^* + \frac{1}{c_j!}W_i^{*c_j}e^{-W_i^*}\right\}, \\ i = n_{j-1} + 1, \dots, n_j, j = 1, \dots, m, \tag{A.4}$$

$$\log P(W_i|\cdot) \propto \log N(\mathbf{x}_i\boldsymbol{\gamma}, \sigma_w^2) + \begin{cases} \log\left(Z_{ij}^* + \frac{1}{c_j!} W_i^{*c_j} e^{-W_i^*}\right), \\ \quad i = n_{j-1} + 1, \dots, n_j, j = 1, \dots, m, \\ y_i W_i - W_i^*, \\ \quad i = n_m + 1, \dots, n, \end{cases}$$

where $N(\mu, \sigma^2)$ indicates the density of Normal distribution with mean μ and variance σ^2 . For the full conditional distribution of regression coefficients, denote $\mathbf{Z}_j = (Z_{1j}, \dots, Z_{hj})^T$, $\mathbf{X} = (\mathbf{x}_1^T, \dots, \mathbf{x}_n^T)^T$ and $\mathbf{W} = (W_1, \dots, W_n)^T$. Then we have the following full conditional distributions

$$\boldsymbol{\beta}_j|\cdot \sim N\left[\left(\frac{\mathbf{X}^T \mathbf{X}}{\sigma_{z_j}^2} + \frac{1}{B_b^2} \mathbf{I}_{p+1}\right)^{-1} \left(\frac{\mathbf{X}^T \mathbf{Z}_j}{\sigma_{z_j}^2}\right), \left(\frac{\mathbf{X}^T \mathbf{X}}{\sigma_{z_j}^2} + \frac{1}{B_b^2} \mathbf{I}_{p+1}\right)^{-1}\right],$$

$j = 1, \dots, m,$ (A.5)

$$\boldsymbol{\gamma}|\cdot \sim N\left[\left(\frac{1}{\sigma_w^2} \mathbf{X}^T \mathbf{X} + \frac{1}{B_c^2} \mathbf{I}_{p+1}\right)^{-1} \left(\frac{1}{\sigma_w^2} \mathbf{X}^T \mathbf{W}\right), \left(\frac{1}{\sigma_w^2} \mathbf{X}^T \mathbf{X} + \frac{1}{B_c^2} \mathbf{I}_{p+1}\right)^{-1}\right].$$

The full conditional distributions of nuisance parameters are given by

$$\sigma_{z_j}^2|\cdot \sim \text{IG}\left[\frac{n}{2} + A_{z_j}, \left(\frac{\text{RSS}_{z_j}}{2} + \frac{1}{B_{z_j}}\right)^{-1}\right], \quad j = 1, \dots, m,$$

$$\sigma_w^2|\cdot \sim \text{IG}\left[\frac{n}{2} + A_w, \left(\frac{\text{RSS}_w}{2} + \frac{1}{B_w}\right)^{-1}\right],$$

(A.6)

where $\text{RSS}_{z_j} = (\mathbf{Z}_j - \mathbf{X}\boldsymbol{\beta}_j)^T (\mathbf{Z}_j - \mathbf{X}\boldsymbol{\beta}_j)$ and $\text{RSS}_w = (\mathbf{W} - \mathbf{X}\boldsymbol{\gamma})^T (\mathbf{W} - \mathbf{X}\boldsymbol{\gamma})$.

Notice that the full conditional distributions of latent variables do not follow specific distributions, while the full conditional distributions of regression coefficients and variances follow Normal distributions and Inverse Gamma distributions, respectively. Using Metropolis-within-Gibbs algorithm we collect MCMC samples for parameters from the joint posterior distribution after some burn-in iterations and perform statistical inferences based on MCMC samples.

Acknowledgements

We thank the referees and Associate Editor for their helpful comments that led to improvement of our paper. Ebrahimi’s research was partially supported by National Science Foundation, DMS-12-08273.

References

- Cox, D. R. (1983). Some remarks on overdispersion. *Biometrika* **70**, 269–274. [MR0742997](#)
- Dobiński, G. (1877). Summierung der Reihe $\sum n^m/n!$ für $m = 1, 2, 3, 4, 5, \dots$. *Grunert Archiv (Arch. Math. Phys.)* **61**, 333–336.
- Greene, W. H. (1994). Accounting for excess zeros and sample selection in Poisson and negative binomial regression models. Working papers. Leonard N. Stern School of Business, Department of Economics, New York University.
- Hedger, D. and Gibbons, R. D. (2006). *Longitudinal Data Analysis*. Hoboken, NJ: John Wiley & Sons. [MR2284230](#)
- Higdon, D., Gattiker, J., Williams, J. B. and Rightly, M. (2008). Computer model calibration using high-dimensional output. *Journal of the American Statistical Association* **103**, 570–583. [MR2523994](#)
- Hu, M.-C., Pavlicova, M. and Nunes, E. V. (2011). Zero-inflated and hurdle models of count data with extra zeros: Examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse* **37**, 367–375.
- Huang, H. and Chin, H. C. (2010). Modeling road traffic crashes with zero-inflation and site-specific random effects. *Statistical Methods & Applications* **19**, 445–462. [MR2673353](#)
- Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34**, 1–14.
- Li, C.-S. (2012). Identifiability of zero-inflated Poisson models. *Brazilian Journal of Probability and Statistics* **26**, 306–312. [MR2911708](#)
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics* **33**, 341–365. [MR0867980](#)
- Su, X., Fan, J., Levine, R. A., Tan, X. and Tripathi, A. (2013). Multiple-inflation Poisson model with L_1 regularization. *Statistica Sinica* **23**, 1071–1090. [MR3114705](#)

D. Ryu
 N. Ebrahimi
 Division of Statistics
 Northern Illinois University
 DeKalb
 Illinois 60115
 USA
 E-mail: dryu@niu.edu
nebrahim@niu.edu

D. Bilgili
 Department of Mathematics and Statistics
 University of North Florida
 Jacksonville
 Florida 32224
 USA
 E-mail: devrim.bilgili@unf.edu

Ö Ergönül
 Infectious Diseases and Clinical Microbiology,
 and Public Health Departments
 School of Medicine
 Koç University
 Istanbul
 Turkey
 E-mail: oergoul@ku.edu.tr