Type I multivariate zero-inflated generalized Poisson distribution with applications

Xi-Fen Huang, Guo-Liang Tian*, Chi Zhang, and Xuejun Jiang

Excessive zeros in multivariate count data are often encountered in practice. Since the Poisson distribution only possesses the property of equi-dispersion, the existing Type I multivariate zero-inflated Poisson distribution (Liu and Tian, 2015, CSDA) [15] cannot be used to model multivariate zero-inflated count data with over-dispersion or under-dispersion. In this paper, we extend the univariate zero-inflated generalized Poisson (ZIGP) distribution to Type I multivariate ZIGP distribution via stochastic representation aiming to model positively correlated multivariate zero-inflated count data with over-dispersion or under-dispersion. Its distributional theories and associated properties are derived. Due to the complexity of the ZIGP model, we provide four useful algorithms (a very fast Fisher-scoring algorithm, an expectation/conditional-maximization algorithm, a simple EM algorithm and an explicit majorization–minimization algorithm) for finding maximum likelihood estimates of parameters of interest and develop efficient statistical inference methods for the proposed model. Simulation studies for investigating the accuracy of point estimates and confidence interval estimates and comparing the likelihood ratio test with the score test are conducted. Under both AIC and BIC, our analyses of the two data sets show that Type I multivariate ZIGP model is superior over Type I multivariate zero-inflated Poisson model.

Keywords and phrases: AIC, BIC, EM algorithm, Fisher scoring algorithm, MM algorithm, Multivariate zero-inflated generalized Poisson distribution, Zero-inflated count data.

1. INTRODUCTION

Count data with excessive zeros are frequently encountered in a number of research fields such as medicine, public health, agriculture, ecology, econometrics, manufacturing and so on. Several distributions of mixture including the zero-inflated Poisson (ZIP), zero-inflated binomial (ZIB), zero-inflated negative binomial (ZINB) have been proposed to handle such count data. For example, Lambert (1992) [12] introduced a ZIP regression model with an application to defects in manufacturing; Hall (2000) [7] described a ZIB regression model and incorporated random effects into ZIP and ZIB models; Lee et al. (2001) [13] generalized the ZIP model by incorporating the extent of individual exposure; and Mimani et al. (2007) [16] proposed the ZINB model and applied it to model the shark by catch data. Other existing models in the literature include the hurdle model (Mullahy, 1986) [17], the two-part model (Heibron, 1994) [8], and the semi-parametric model (Li, 2012) [14].

The equality of mean and variance characterizes the Poisson distribution. It has also been observed that in a population the probability of the occurrence of an event does not remain constant and changes with time and/or previous occurrences, resulting in unequal mean and variance in the data. As a useful generalization of the standard Poisson distribution, the generalized Poisson (GP) distribution was introduced firstly by Consul and Jain (1973) [1] as a limiting form of the generalized negative binomial distribution, implying that there is some changing tendency in the parameter with successive occurrences by adding an additional parameter. It is an important competitor to the negative binomial model when the count data are over-dispersed since the variance of the GP distribution could be greater than, equal to or smaller than its mean depending on if the additional parameter is positive, zero or negative. A non-negative integer valued random variable $X$ is said to have a GP distribution with parameter $\lambda \in \mathbb{R}_+$ and dispersion parameter $\theta$, if its probability mass function (pmf) is given by (Consul and Jain, 1973 [1]; Consul and Shoukri, 1985 [2])

\[
f(x; \lambda, \theta) = \begin{cases} 
\frac{\lambda(\lambda + \theta x)^{\theta - 1} e^{-\lambda - \theta x}}{x!}, & x = 0, 1, \ldots, \\
0, & \mbox{for } x > q \mbox{ when } \theta < 0,
\end{cases}
\]

where $\max(-1,-\lambda/q) < \theta \leq 1$ and $q \geq 4$ is the largest positive integer for which $\lambda + \theta q > 0$ when $\theta < 0$. We denote it by $X \sim GP(\lambda, \theta)$. When $\theta = 0$, the GP($\lambda, \theta$) distribution reduces to the Poisson($\lambda$) distribution with the property of equi-dispersion. When $\theta > 0$ (or $\theta < 0$), the GP($\lambda, \theta$) distribution can be used to model count data with over-dispersion (or under-dispersion). When $\lambda = 0$, the GP($\lambda, \theta$) reduces to the degenerate distribution Degenerate(0).

Based on (1.1), some researchers developed so-called zero-inflated generalized Poisson (ZIGP) and zero-adjusted gen-

*Corresponding author.
eralized Poisson (ZAGP) models as alternatives to ZIP, ZIB and ZINB models for the analysis of count data with extra zeros. For example, Gupta et al. (1996) [5] studied general zero-adjusted count data models, proposed a ZAGP distribution and investigated the relative error incurred by ignoring the adjustment. They also provided real data sets where the ZAGP distribution fits very well. Gupta et al. (2004) [6] studied the ZIGP regression model and developed a score test to determine whether an adjustment for zero inflation is necessary. Famoye and Singh (2006) [4] developed a ZIGP regression model to model domestic violence data. Xie and Wei (2010) [21] extended the ZIP mixed regression model to the ZIGP mixed regression model and Xie et al. (2014) [22] provided a Markov chain Monte Carlo method for dealing with the complexity of the ZIGP model.

Excessive zeros in multivariate count data are often encountered in practice, e.g., when events involve different types of defects in a manufacturing process near its perfect state, the univariate zero-inflated count distributions are no longer appropriate. To model dependent structure in multivariate count data, some authors have extended the univariate ZIP distribution to multivariate ZIP distribution, for example, Liu and Tian (2015) [15] introduced the Type I multivariate ZIGP distribution via stochastic representation aiming to model positively correlated multivariate zero-inflated count data with extra zeros. Since the Poisson distribution only possesses the property of equi-dispersion, the existing Type I multivariate ZIP distribution cannot be used to model multivariate zero-inflated count data with over-dispersion or under-dispersion. In this paper, we extend the univariate ZIGP distribution to Type I multivariate ZIGP distribution via stochastic representation aiming to model positively correlated multivariate zero-inflated count data with over-dispersion or under-dispersion. Its distributional theories and associated properties are derived. Due to the complexity of the ZIGP model, we will provide four useful algorithms (a very fast Fisher-scoring algorithm, an ECM algorithm, a simple EM algorithm and an explicit MM algorithm) for finding maximum likelihood estimates (MLEs) of parameters of interest and will develop efficient statistical inference methods for the proposed model.

The rest of the paper is organized as follows. In Section 2, we introduce the Type I multivariate ZIGP distribution, and study the distributional theories and corresponding properties. In Section 3, the likelihood-based statistical inferences about parameters of interest are provided. Simulation studies for investigating the accuracy of point estimates and confidence interval estimates and comparing the likelihood ratio test with the score test are conducted in Section 4. In Section 5, two real examples are used to illustrate the proposed methods and to compare with existing methods. A discussion is given in Section 6. Some detailed technical proofs are put in the Appendices.

2. TYPE I MULTIVARIATE ZERO-INFLATED GENERALIZED POISSON DISTRIBUTION

Let Z \sim Bernoulli(1 - \phi), X \sim GP(\lambda, \theta) and Z \perp X. The random variable Y \sim ZIGP(\phi, \lambda, \theta) has the following stochastic representation (SR):

\begin{equation}
Y \overset{d}{=} ZX = \begin{cases} 
0, & \text{with probability } \phi, \\
X, & \text{with probability } 1 - \phi,
\end{cases}
\end{equation}

where the symbol “\overset{d}{=}” means that the random variables on both sides of the equality have the same distribution. When \theta = 0, the ZIGP(\phi, \lambda, \theta) reduces to the zero-inflated Poisson distribution ZIP(\phi, \lambda). Alternative to (2.1), we obtain the following mixture representation:

Z \sim Bernoulli(1 - \phi) \quad \text{and} \quad Y|(Z = z) \sim GP(\lambda z, \theta).

From (2.1), we immediately obtain

\begin{equation}
\begin{aligned}
E(Y) &= (1 - \phi)\lambda, \\
E(Y^2) &= (1 - \phi)\lambda + (1 - \phi)^2\lambda^2, \\
\text{Var}(Y) &= (1 - \phi)\lambda + (1 - \phi)^2\lambda^2 - (1 - \phi)^2\lambda^2,
\end{aligned}
\end{equation}

where \theta < 1.

Motivated by the SR (2.1) of the univariate ZIGP distribution, we can extend it to the multivariate version by means of an SR in a vector form with a common Z to characterize the correlation structure among the components, as shown in the following definition.

**Definition 1.** Let Z \sim Bernoulli(1 - \phi), x = (X_1, \ldots, X_m)^\top, X_i \sim GP(\lambda_i, \theta_i) for i = 1, \ldots, m, and (Z, X_1, \ldots, X_m) are mutually independent. An m-dimensional discrete random vector y = (Y_1, \ldots, Y_m)^\top is said to have a Type I multivariate ZIGP distribution if

\begin{equation}
y \overset{d}{=} Z x = \begin{cases} 
0, & \text{with probability } \phi, \\
x, & \text{with probability } 1 - \phi,
\end{cases}
\end{equation}

where \phi \in [0, 1), \lambda = (\lambda_1, \ldots, \lambda_m)^\top \in \mathbb{R}_+^m, \theta = (\theta_1, \ldots, \theta_m)^\top, \text{max}(-1, -\lambda_i/q_i) < \theta_i < 1 and q_i \geq 4 is the largest positive integer for each \lambda_i + \theta_i > 0 when \theta_i < 0. We write y \sim ZIGP^{(1)}_m(\phi, \lambda, \theta) or y \sim ZIGP^{(1)}(\phi; \lambda_1, \ldots, \lambda_m, \theta_1, \ldots, \theta_m) and call x the base vector of the y.

2.1 Joint pmf and joint cumulative distribution function

The joint pmf of y \sim ZIGP^{(1)}_m(\phi, \lambda, \theta) is denoted by f(y|\phi, \lambda, \theta) = \Pr(y = y) = \Pr(ZX = y), 1 \leq i \leq m). If y = 0_m, we have
\[ f(y|\phi, \lambda, \theta) = \text{Pr}(ZX_i = 0, 1 \leq i \leq m) \]
\[ = \text{Pr}(Z = 0) + \text{Pr}(Z = 1, X_i = 0, 1 \leq i \leq m) \]
\[ (2.3) = \phi + (1 - \phi)e^{-\lambda_i}, \]
where \( \lambda_i = \sum_{i=1}^{m} \lambda_i. \) If \( y \neq 0, \) we have
\[ f(y|\phi, \lambda, \theta) = \text{Pr}(Z = 1, X_i = y_i, 1 \leq i \leq m) \]
\[ = (1 - \phi)e^{-\lambda_i - \sum_{i=1}^{m} \theta_i y_i} \prod_{i=1}^{m} \frac{\lambda_i(\lambda_i + \theta_i y_i)^{y_i-1}}{y_i!} \]
\[ (2.4) = (1 - \phi)a, \]

By combining (2.3) with (2.4), we obtain
\[ f(y|\phi, \lambda, \theta) = (\phi + (1 - \phi)e^{-\lambda_i})[I(y = 0) + (1 - \phi)I(y \neq 0)] = \phi \text{Pr}(\xi = y) + (1 - \phi) \text{Pr}(x \leq y), \]
where \( \xi = (\xi_1, \ldots, \xi_m)^T \) and \( \{\xi_i\}_{i=1}^{m} \text{iid} \sim \text{Degenerate}(0). \)

Let \( y \sim \text{ZIGP}^{(m)}(\phi, \lambda, \theta). \) For any non-negative real vector \( y = (y_1, \ldots, y_m)^T, \) the joint cumulative distribution function of \( y \) is given by
\[ \text{Pr}(y \leq y) = \phi \text{Pr}(\xi = 0) + (1 - \phi) \text{Pr}(x \leq y) \]
\[ = \phi + (1 - \phi) \prod_{i=1}^{m} \text{Pr}(X_i \leq y_i) \]
\[ = \phi + (1 - \phi) \prod_{i=1}^{m} \sum_{k_i=0}^{y_i} \frac{\lambda_i(\lambda_i + \theta_i k_i)^{k_i-1} e^{-\lambda_i - \theta_i k_i}}{k_i!} \]
for \( y_1, \ldots, y_m \geq 0. \)

2.2 Mixed moments and moment generating function

From (2.2), it is not difficult to obtain
\[ \begin{align*}
E(y) &= (1 - \phi)\alpha, \\
E(yy^T) &= (1 - \phi)[\text{diag}(\beta) + \alpha\alpha^T], \\
\text{Var}(y) &= (1 - \phi)[\text{diag}(\beta) + \phi\alpha\alpha^T],
\end{align*} \]
where \( \alpha = (\alpha_1, \ldots, \alpha_m)^T, \quad \alpha_i = \lambda_i/(1 - \theta_i), \quad \beta = (\beta_1, \ldots, \beta_m)^T, \quad \beta_i = \lambda_i/(1 - \theta_i)^3, \quad \theta_i < 1 \text{ for } i = 1, \ldots, m. \)

Thus we have
\[ \text{Corr}(Y_i, Y_j) = \frac{\lambda_i \lambda_j (1 - \theta_i)(1 - \theta_j)}{\sqrt{[(1 - \theta_i)\lambda_i + 1/\phi][(1 - \theta_j)\lambda_j + 1/\phi]}} \]
for \( i \neq j. \) In particular, when \( \lambda_i = \lambda_j = \lambda \) and \( \theta_i = \theta_j = \theta, \) we obtain
\[ \text{Corr}(Y_i, Y_j) = \frac{\phi\lambda(1 - \theta)}{1 + \phi\lambda(1 - \theta)}, \quad i \neq j. \]

By using the formula of \( E(\xi) = E[E(\xi|y)] \), we can obtain the moment generating function of \( y \sim \text{ZIGP}^{(1)}(\phi; \lambda_1, \ldots, \lambda_m, \theta_1, \ldots, \theta_m), \) given by
\[ M_y(t) = E[\exp(t^T y)] = E[\exp(Z \cdot t^T x)] \]
\[ = E\left\{ E[\exp(Z^T y)|Z]\right\} = E\left[ \prod_{i=1}^{m} M_{X_i}(t, Z) \right] \]
\[ = \phi \prod_{i=0}^{m} M_{X_i}(0) + (1 - \phi) \prod_{i=1}^{m} M_{X_i}(t_i), \]
where
\[ M_{X_i}(0) = \exp\left\{ \frac{\lambda_i}{\theta_i} \left[ W(-\theta_i e^{-\theta_i}) + \theta_i \right] \right\} \text{ and } \]
\[ M_{X_i}(t_i) = \exp\left\{ \frac{-\lambda_i}{\theta_i} \left[ W(-\theta_i e^{-\theta_i + t_i}) + \theta_i \right] \right\} \]
for \( i = 1, \ldots, m; \) the Lambert \( W(\cdot) \) function is defined by \( W(x) \exp[W(x)] = x, \) for more details about this function see Corless et al. (1996) [3].

2.3 Marginal distributions

Let \( y \sim \text{ZIGP}^{(1)}(\phi; \lambda_1, \ldots, \lambda_m, \theta_1, \ldots, \theta_m). \) Partition \( y \) into two parts
\[ y = \left( \begin{array}{c} y^{(1)} \\ y^{(2)} \end{array} \right), \quad \text{where } y^{(1)} = \left( \begin{array}{c} Y_1 \\ \vdots \\ Y_{r-1} \end{array} \right), \quad y^{(2)} = \left( \begin{array}{c} Y_r \\ \vdots \\ Y_m \end{array} \right), \]
We can partition \( x \) in the same fashion. According to Definition 1, we obtain
\[ \begin{cases} y^{(1)} \overset{\text{d}}{=} Zx^{(1)} \sim \text{ZIGP}^{(1)}(\phi; \lambda_1, \ldots, \lambda_r, \theta_1, \ldots, \theta_r) \quad \text{and} \\ y^{(2)} \overset{\text{d}}{=} Zx^{(2)} \sim \text{ZIGP}^{(1)}(\phi; \lambda_{r+1}, \ldots, \lambda_m, \theta_{r+1}, \ldots, \theta_m). \end{cases} \]
In fact, for any positive integers \( i_1, \ldots, i_r \) satisfying \( 1 \leq i_1 < \ldots < i_r \leq m, \) we have
\[ \left( \begin{array}{c} Y_{i_1} \\ \vdots \\ Y_{i_r} \end{array} \right) \overset{\text{d}}{=} Z \left( \begin{array}{c} X_{i_1} \\ \vdots \\ X_{i_r} \end{array} \right) \overset{\text{d}}{=} \text{ZIGP}^{(1)}(\phi; \lambda_{i_1}, \ldots, \lambda_{i_r}, \theta_{i_1}, \ldots, \theta_{i_r}). \]

2.4 Conditional distributions

2.4.1 Conditional distribution of \( y^{(1)}|y^{(2)} \)

From (2.5) and (2.6), the conditional distribution of \( y^{(1)}|y^{(2)} \) is given by
\[ \Pr(y^{(1)} = y^{(1)}|y^{(2)} = y^{(2)}) = \frac{f(y|\phi, \lambda, \theta)}{\Pr(y^{(2)} = y^{(2)})} \]

Type I multivariate ZIGP distribution 293
\[ y^{(1)} = \frac{\phi + (1 - \phi)e^{-\lambda_1}}{\phi + (1 - \phi)e^{-\lambda_2}} \] for \( y^{(2)} = \emptyset \), where \( \phi \) is defined by (2.4), and \( \lambda_2^{(2)} = \sum_{i=r+1}^{m} \lambda_i = \lambda_r - \lambda^{(1)} \).

We consider two cases. Case I: \( y^{(2)} \neq \emptyset \). Under Case I, it is obvious that \( y \neq \emptyset \). From (2.8) and (2.9), it is easy to obtain

\[ \Pr(y^{(1)} = y^{(2)}) = \prod_{i=1}^{r} \lambda_i(\lambda_i + \theta_i y_i) y_i! = \frac{a}{a_2} \hat{a}_1. \]

This implies \( y^{(1)} | (y^{(2)} = y^{(2)}) \) is \( \chi \) taking on \( Z \). In other words, given \( y^{(2)} \neq \emptyset \), \( (Y_1, \ldots, Y_r) \) are mutually independent and \( Y_i | (y^{(2)} = y^{(2)}) \) and \( Y_i \sim GP(\lambda_i, \theta_i) \), being free from \( \phi \).

Case II: \( y^{(2)} = \emptyset \). Under Case II, it is possible that \( y^{(1)} = \emptyset \) or \( y^{(1)} \neq \emptyset \). When \( y^{(1)} = \emptyset \), from (2.8), we have

\[ \Pr(y^{(1)} = \emptyset | y^{(2)} = \emptyset) = \frac{\phi (1 - \phi)^{r-1} e^{-\lambda_2}}{\phi (1 - \phi)^{r-1} e^{-\lambda_2}} \]

where \( \phi = \phi e^{\lambda_2} / (\phi e^{\lambda_2}) 

By combining (2.11) with (2.12), we obtain

\[ \Pr(y^{(1)} = y^{(1)} | y^{(2)} = \emptyset) = \frac{(1 - \phi)e^{-\lambda_2} \sum_{i=1}^{r} \theta_i y_i \prod_{i=1}^{r} \lambda_i(\lambda_i + \theta_i y_i) y_i!}{(1 - \phi)e^{-\lambda_2} \sum_{i=1}^{r} \theta_i y_i \prod_{i=1}^{r} \lambda_i(\lambda_i + \theta_i y_i) y_i!} \]

\[ = \frac{1 - \hat{\phi}_1 a_1}{(1 - \hat{\phi}_1) a_1}. \]

Therefore,

\[ Z | (y = y) \sim \begin{cases} \text{Bernoulli}(\psi), & \text{if } y = \emptyset, \\ \text{Degenerate}(1), & \text{if } y \neq \emptyset, \end{cases} \]

where

\[ \psi = \frac{(1 - \phi)e^{-\lambda_2}}{\phi + (1 - \phi)e^{-\lambda_2}}. \]

2.4.2 Conditional distribution of \( Z \mid y \)

Since \( Z \sim \text{Bernoulli}(1 - \phi) \), \( Z \) only takes the value 0 or 1. Note that

\[ \Pr(Z = 1 | y = y) = \frac{\Pr(Z = 1, x = y)}{\Pr(y = y)} = \frac{\Pr(Z = 1, x = y)}{\int f(y | \phi, \lambda, \theta) f(x | \psi, \lambda, \theta) dy} \]

\[ = \begin{cases} \frac{e^{-\lambda_1} I(x = 0)}{\phi (1 - \phi)e^{-\lambda_2}}, & \text{if } y = \emptyset, \\ \frac{e^{-\lambda_2} I(x = 0)}{\phi (1 - \phi)e^{-\lambda_2}}, & \text{if } y \neq \emptyset. \end{cases} \]

2.4.3 Conditional distribution of \( x \mid y \)

If \( y = \emptyset \), we have

\[ \Pr(x = x | y = \emptyset) = \frac{\Pr(x = x, y = \emptyset)}{\Pr(y = \emptyset)} = \frac{\Pr(x = x, y = \emptyset)}{\int f(y | \phi, \lambda, \theta) f(x | \psi, \lambda, \theta) dy} \]

\[ = \begin{cases} \frac{e^{-\lambda_1} I(x = 0)}{\phi (1 - \phi)e^{-\lambda_2}}, & \text{if } y = \emptyset, \\ \frac{e^{-\lambda_2} I(x = 0)}{\phi (1 - \phi)e^{-\lambda_2}}, & \text{if } y \neq \emptyset. \end{cases} \]

i.e.,

\[ x | (y = \emptyset) \sim \text{ZIGP}^{(1)}(\psi; \lambda_1, \ldots, \lambda_m, \theta_1, \ldots, \theta_m). \]

If \( y \neq \emptyset \), we have

\[ \Pr(x = x | y = y) = \frac{\Pr(x = x, y = y)}{\Pr(y = y)} = \frac{\Pr(x = x, y = y)}{\int f(y | \phi, \lambda, \theta) f(x | \psi, \lambda, \theta) dy} = 1. \]
Thus, given $y = y \neq 0$, $(X_1, \ldots, X_m)$ are independent and 

$$X_i | (y = y \neq 0) \sim \text{Degenerate}(y_i), \quad i = 1, \ldots, m. \tag{2.16}$$

2.4.4 Conditional distribution of $X_i | (Y_i = y_i) = 0), i = 1, \ldots, m$. 

From (2.7), we have $Y_i \sim \text{ZIGP}(\phi, \lambda_i, \theta_i)$. Thus, 

$$\begin{align*}
\Pr(X_i = x_i | Y_i = 0) &= \frac{\Pr(X_i = x_i, Y_i = 0)}{\Pr(Y_i = 0)} \\
&= \frac{\Pr(X_i = x_i, Y_i = 0)}{\int \phi(x_i | \lambda_i, \theta_i) d\lambda_i} \\
&= \frac{\Pr(Y_i = 0) \int \phi(x_i | \lambda_i, \theta_i) d\lambda_i}{\Pr(Y_i = 0)} \\
&= \frac{\phi(x_i | \lambda_i, \theta_i)}{\phi}.
\end{align*}$$

so that the log-likelihood function is 

$$\ell = \ell(\phi, \lambda, \theta | Y_{\text{obs}}) = n_0 \log[\phi + (1 - \phi)e^{-\lambda}]$$

$$+ \sum_{j=1}^{n_0} \log(\lambda + \theta y_j - \theta y_j - \theta y_j).$$

3. MLEs via the Fisher scoring algorithm

In this subsection, the Fisher scoring algorithm is employed to calculate the MLEs of $\phi$, $\lambda$, and $\theta$. The score vector $\nabla \ell$ and the Hessian matrix $\nabla^2 \ell$ are given by 

$$\nabla \ell = \left( \frac{\partial \ell}{\partial \phi}, \frac{\partial \ell}{\partial \lambda}, \frac{\partial \ell}{\partial \theta} \right)^T$$

$$\nabla^2 \ell = \left( \begin{array}{ccc}
\frac{\partial^2 \ell}{\partial \phi^2} & \frac{\partial^2 \ell}{\partial \phi \partial \lambda} & \frac{\partial^2 \ell}{\partial \phi \partial \theta} \\
\frac{\partial^2 \ell}{\partial \lambda \partial \phi} & \frac{\partial^2 \ell}{\partial \lambda^2} & \frac{\partial^2 \ell}{\partial \lambda \partial \theta} \\
\frac{\partial^2 \ell}{\partial \theta \partial \phi} & \frac{\partial^2 \ell}{\partial \theta \partial \lambda} & \frac{\partial^2 \ell}{\partial \theta^2} \\
\end{array} \right).$$

respectively, where 

$$\frac{\partial \ell}{\partial \phi} = \frac{n_0(1 - e^{-\lambda})}{\phi + (1 - \phi)e^{-\lambda}} - \frac{n - n_0}{1 - \phi},$$

$$\frac{\partial \ell}{\partial \lambda} = -\frac{n_0(1 - \phi)e^{-\lambda}}{\phi + (1 - \phi)e^{-\lambda}} - \frac{(n - n_0)}{\lambda_i} + \sum_{j=1}^{n} \left( \frac{1}{\lambda_i + \theta y_j} - y_j \right),$$

$$\frac{\partial^2 \ell}{\partial \phi^2} = \frac{n_0(1 - \phi)e^{-\lambda}}{(\phi + (1 - \phi)e^{-\lambda})^2} - \frac{n - n_0}{(1 - \phi)^2},$$

$$\frac{\partial^2 \ell}{\partial \phi \partial \lambda} = \frac{n_0(1 - \phi)e^{-\lambda}}{(\phi + (1 - \phi)e^{-\lambda})^2} - \frac{\sum_{j=1}^{n} \left( \frac{1}{\lambda_i + \theta y_j} - y_j \right)^2}{(1 - \phi)^2},$$

$$\frac{\partial^2 \ell}{\partial \lambda^2} = \sum_{j=1}^{n} \left( \frac{1}{\lambda_i + \theta y_j} - y_j \right)^2,$$

$$\frac{\partial^2 \ell}{\partial \lambda \partial \theta} = \frac{n_0(1 - \phi)e^{-\lambda}}{(\phi + (1 - \phi)e^{-\lambda})^2} - \frac{\sum_{j=1}^{n} \left( \frac{1}{\lambda_i + \theta y_j} - y_j \right)^2}{(1 - \phi)^2},$$

$$\frac{\partial^2 \ell}{\partial \theta^2} = \frac{n_0e^{-\lambda}}{(\phi + (1 - \phi)e^{-\lambda})^2}.$$

3.1 Type I multivariate ZIGP distribution 295
\[
\begin{align*}
\frac{\partial^2 \ell}{\partial \lambda_i \partial \lambda_k} &= \frac{n_0 \phi (1 - \phi) e^{-\lambda_i}}{[\phi + (1 - \phi) e^{-\lambda_i}]}^2, \\
\frac{\partial^2 \ell}{\partial \lambda_i \partial \theta_{ij}} &= - \sum_{j=1}^{n} \frac{y_{ij} (y_{ij} - 1)}{\lambda_i + \theta_{ij}}, \\
\frac{\partial^2 \ell}{\partial \lambda_i \partial \theta_{ik}} &= \frac{\partial^2 \ell}{\partial \theta_{ij} \partial \theta_{ik}} = 0,
\end{align*}
\]
for \(i, k = 1, \ldots, m\) and \(i \neq k\). By replacing \(n_0\),
\[
\begin{align*}
\left\{ \frac{y_{ij} - 1}{\lambda_i + \theta_{ij}} \right\}_{i=1}^{m}, \quad \left\{ \frac{y_{ij} - y_{ij}^2}{\lambda_i + \theta_{ij}} \right\}_{i=1}^{m}, \\
\left\{ \frac{y_{ij}^2 - y_{ij}}{\lambda_i + \theta_{ij}} \right\}_{i=1}^{m}
\end{align*}
\]
in the above second partial derivatives with their expectations (see Appendix A)
\[(3.2)
\]
\[
\begin{align*}
E \left[ \sum_{j=1}^{n} I(y_j) \right] &= n \left[ \phi (1 - \phi) e^{-\lambda_i} \right], \\
E \left[ \frac{Y_{ij} - 1}{\lambda_i + \theta_{ij}} \right] &= \frac{1 - \phi}{\lambda_i} - \frac{1 - \theta_{ij} (1 - \phi)}{\lambda_i + 2 \theta_{ij}}, \\
E \left[ \frac{Y_{ij}^2 - Y_{ij}}{\lambda_i + \theta_{ij}} \right] &= \frac{\lambda_i (1 - \phi) + 2 \lambda_i (1 - \phi)}{1 - \theta_{ij}} + \frac{\lambda_i + 2 \theta_{ij}}{1 - \theta_{ij}}, \\
E \left[ \frac{Y_{ij}^3 - Y_{ij}}{\lambda_i + \theta_{ij}} \right] &= \frac{\lambda_i (1 - \phi)}{\lambda_i + 2 \theta_{ij}},
\end{align*}
\]
we can calculate the Fisher information matrix
\[
J(\phi, \lambda, \theta) = E[-\nabla^2 \ell(\phi, \lambda, \theta|Y_{\text{obs}})].
\]
Let \((\phi^{(0)}, \lambda^{(0)}, \theta^{(0)})\) be the initial values of the MLEs \((\hat{\phi}, \hat{\lambda}, \hat{\theta})\). If \((\phi^{(t)}, \lambda^{(t)}, \theta^{(t)})\) denote the \(t\)-th approximations of \((\hat{\phi}, \hat{\lambda}, \hat{\theta})\), then their \((t+1)\)-th approximations can be obtained by the following Fisher scoring algorithm:
\[(3.3)\]
\[
\begin{align*}
\left( \frac{\phi^{(t+1)}}{\lambda^{(t+1)}} \right) &= \left( \frac{\phi^{(t)}}{\lambda^{(t)}} \right) \left( \theta^{(t)} \right) \\
&+ J^{-1}(\phi^{(t)}, \lambda^{(t)}, \theta^{(t)}) \nabla \ell(\phi^{(t)}, \lambda^{(t)}, \theta^{(t)}|Y_{\text{obs}}).
\end{align*}
\]
The standard errors of the MLEs \((\hat{\phi}, \hat{\lambda}, \hat{\theta})\) are the square roots of the diagonal elements \(J_{kk}\) of the inverse Fisher information matrix \(J^{-1}(\phi, \lambda, \theta)\). Thus the \((1 - \alpha)\)100\% asymptotic Wald confidence intervals (CIs) of \(\phi\), \(\{\lambda_i\}_{i=1}^{m}\) and \(\{\theta_{ij}\}_{i=1}^{m}\) are given by
\[(3.4)\]
\[
\begin{align*}
\hat{\phi} &\pm z_{\alpha/2} \sqrt{\lambda_{\phi}}, \\
\hat{\lambda}_i &\pm z_{\alpha/2} \sqrt{\lambda_{\lambda_i}}, \\
\hat{\theta}_{ij} &\pm z_{\alpha/2} \sqrt{\lambda_{\theta_{ij}}},
\end{align*}
\]
for \(i = 1, \ldots, m\), respectively, where \(z_{\alpha}\) denotes the \(\alpha\)-th upper quantile of the standard normal distribution.

### 3.2 MLEs via two EM-type algorithms

Although we derive the Fisher scoring algorithm to estimates the parameters in the Type I multivariate ZIGP model, it is sensitive to the choice of initial values. In other words, the Fisher scoring algorithm may be divergent if a poor initial value is chosen. Thus in this subsection we develop two EM-type algorithms: the first one is an ECM algorithm and the second one is an EM algorithm.

#### 3.2.1 An ECM algorithm based on SR

For each \(y_j = (y_{1j}, \ldots, y_{mj})\) with \(j \in \{1, \ldots, n\}\), based on the SR (2.2) we introduce independent latent variables \(Z_j \sim \text{Bernoulli}(1 - \phi)\), \(X_{ij} \sim \text{GP}(\lambda_i, \theta_i)\) for \(i = 1, \ldots, m\). We denote the latent/missing data by \(Y_{\min} = \{z_j, x_{ij}\}_{i=1}^{m}\) such that \(y_j = z_j x_j\) and the complete data are \(Y_{\text{com}} = \{Y_{\text{obs}}, Y_{\min}\} = Y_{\text{mis}},\) where \(x_j = (x_{1j}, \ldots, x_{mj})\), \(z_j\) and \(x_{ij}\) denote the realizations of \(Z_j\) and \(X_{ij}\), respectively. The complete-data log-likelihood function is
\[(3.5)\]
\[
L_1(\phi, \lambda, \theta|Y_{\text{com}}) = \prod_{j=1}^{n} \left[ (1 - \phi)^{z_j} \phi^{(1 - z_j)} \prod_{i=1}^{m} \frac{\lambda_i (\lambda_i + \theta_i x_{ij}) x_{ij} - 1}{x_{ij}} \right],
\]
so that the complete-data log-likelihood function is
\[
\ell_1(\phi, \lambda, \theta|Y_{\text{com}}) = c_1 + \sum_{j=1}^{n} \left[ z_j \log(1 - \phi) + (1 - z_j) \log(\phi) \right]
\]
\[
+ \sum_{j=1}^{n} \sum_{i=1}^{m} \left[ \log \lambda_i + (x_{ij} - 1) \log(\lambda_i + \theta_i x_{ij}) - \lambda_i - \theta_i x_{ij} \right].
\]
Then, the complete-data MLEs of \(\phi\) and \(\{\lambda_i\}_{i=1}^{m}\) are given by
\[
\begin{align*}
\phi &= \frac{n - \sum_{j=1}^{n} z_j}{n}, \quad \lambda_i = \frac{\sum_{j=1}^{n} x_{ij}}{n} (1 - \theta_i),
\end{align*}
\]
for \(i = 1, \ldots, m\), while the complete-data MLE of \(\theta_i\) is the root of the equation
\[
\begin{align*}
H_i(\theta_i|\lambda_i) &= \sum_{j=1}^{n} \frac{x_{ij}^2 - x_{ij}}{\lambda_i + \theta_i x_{ij}} - \sum_{j=1}^{n} x_{ij} = 0,
\end{align*}
\]
for \(i = 1, \ldots, m\). The E-step is to replace \(\{z_j\}_{j=1}^{n}, \{x_{ij}\}_{j=1}^{n}\) and \(\left\{ x_{ij}^2 - x_{ij} \right\}_{j=1}^{n}\) in (3.5)–(3.6) by their conditional expectations:
\[(3.7)\]
\[
E(Z_j|Y_{\text{obs}}, \phi, \lambda, \theta) = \psi I(y_j = \mathbf{0}) + I(y_j \neq \mathbf{0}),
\]
\[(2.13)\]
A simple EM algorithm by introducing only one latent variable

In the previous subsection, we proposed an ECM algorithm by introducing \( n(1 + m) \) latent variables. It is well known that an ECM algorithm generally converges much slower than the corresponding EM algorithm. Thus in this subsection, we will provide a simple EM algorithm by introducing only one latent variable.

Note that the observed zero vectors from a Type I multivariate ZIGP distribution can be classified into two categories: One is called the extra zero vectors resulted from degenerate distribution at point zero because of population variability; while the other is called the structural zero vectors came from the independent ordinary GP distributions. Thus, we can partition

\[
J_0 = \{j|y_j = 0, j = 1, \ldots, n\}
\]
as the union of \( J_{\text{extra}} \) and \( J_{\text{structural}} \). The major obstacle for obtaining explicit solutions of MLEs of parameters from \((3.1)\) is the first term of \((3.1)\). To overcome this difficulty, we augment \( Y_{\text{obs}} \) with a latent variable \( W \) that denotes the number of \( J_{\text{extra}} \) to split \( n_0 \) into \( W \) and \( n_0 - W \). The resultant conditional predictive distribution of \( W \) given \( Y_{\text{obs}} \) and \((\phi, \lambda, \theta)\) is

\[
W|Y_{\text{obs}}, \phi, \lambda, \theta \sim \begin{cases} \text{Binomial} \left( n_0, \frac{\phi}{\phi + (1 - \phi)e^{-\lambda_+}} \right) & \text{Binomial}(n_0, 1 - \psi), \\
\end{cases}
\]

where \( \psi \) is defined by \((2.14)\). The complete-data likelihood

\[
L_2(\phi, \lambda, \theta|Y_{\text{com}}) = \phi^w[(1 - \phi)e^{-\lambda_+}]^{n_0 - w}(1 - \phi)^{n - n_0}e^{-(n - n_0)\lambda_+} \times \prod_{j=1}^{n} \prod_{i=1}^{m} \lambda_i, (\lambda_i^t + \theta_i y_{ij})^{y_{ij} - 1}e^{-\theta_i y_{ij}},
\]

so that the complete-data log-likelihood function is

\[
\ell_2(\phi, \lambda, \theta|Y_{\text{com}}) = w \log \phi + (n - w) \log(1 - \phi) - (n - w)\lambda_+ + \sum_{j=1}^{n} \sum_{i=1}^{m} \left[ \log \lambda_i + (y_{ij} - 1) \log(\lambda_i + \theta_i y_{ij}) - \theta_i y_{ij} \right].
\]

Hence, the complete-data MLEs of \( \phi \) and \( \{\lambda_i\}_{i=1}^{m} \) are given by

\[
\hat{\phi} = \frac{w}{n}, \quad \hat{\lambda}_i = \frac{n \bar{y}_i (1 - \hat{\theta}_i)}{n - w}, \quad i = 1, \ldots, m,
\]

where \( \bar{y}_i \) is defined by \((3.13)\), and complete-data MLE of \( \theta_i \) is the root of the equation

\[
H_i(u) = \sum_{j=1}^{n} u y_{ij}, (\bar{y}_i - u) + w y_{ij}, (n - w) - n \bar{y}_i = 0,
\]

for \( i = 1, \ldots, m \). Thus, the E-step is to replace \( w \) in the above expressions by its conditional expectation

\[
E(W|Y_{\text{obs}}, \phi, \lambda, \theta) = \frac{n_0 \phi}{\phi + (1 - \phi)e^{-\lambda_+}} = n_0(1 - \psi).
\]

**Type I multivariate ZIGP distribution**
3.3 MLEs via the MM algorithm

Although the proposed two EM-type algorithms provided relatively simple iterations to find the MLEs of parameters in the Type I multivariate ZIGP distribution, we have to solve the root of \( m \) one-dimensional nonlinear equations specified by (3.11) or (3.15) at each step by employing the Newton’s method, whose convergence depends on the choice of initial values. The situation becomes much complicated when such EM-type algorithms are utilized to calculate the confidence intervals of parameters via bootstrap methods as shown in the next subsection. In other words, we do not know how to specify so many initial values in these Newton’s methods such that they can converge. In this subsection, we will develop a novel MM algorithm with explicit expressions at each iteration through constructing a \( Q \) function to separate the parameters \( \phi, \lambda \) and \( \theta \).

For convenience, we first define

\[
\mathcal{J}_0 = \{ j | y_j = \mathbf{0}, \ j = 1, \ldots, n \},
\]

\[
n_0 = \sum_{j=1}^{n} \mathcal{I}(y_j = \mathbf{0}) = \# \{\mathcal{J}_0\},
\]

\[
\mathcal{J} = \{ j | y_j \neq \mathbf{0}, \ j = 1, \ldots, n \},
\]

\[
\mathcal{J}_i = \{ j | y_{ij} \neq 0, \ j = 1, \ldots, n \}, \quad i = 1, \ldots, m,
\]

\[
\mathcal{J}_{i0} = \{ j | y_{ij} = 0, \ y_j \neq \mathbf{0}, \ j = 1, \ldots, n \}, \quad n_{i0} = \# \{\mathcal{J}_{i0}\}.
\]

Then, we have \( \# (\mathcal{J}) = n - n_0 \) and

\[
\# \{\mathcal{J}_i\} = \# (\mathcal{J}) - \# (\mathcal{J}_{i0}) = n - n_0 - n_{i0}.
\]

The observed-data likelihood function can be rewritten as

\[
L(\phi, \lambda, \theta | Y_{\text{obs}}) = [\phi + (1 - \phi)e^{-\lambda_+}]^{n_0} (1 - \phi)^{n-n_0} \times \prod_{j \in \mathcal{J}} \prod_{i=1}^{m} \lambda_i(y_{ij} - 1)e^{-\lambda_+ - \theta_iy_{ij}} \prod_{j \in \mathcal{J}_i} \lambda_i(y_{ij} - 1)e^{-\lambda_i - \theta_iy_{ij}} \prod_{j \in \mathcal{J}_{i0}} \lambda_i(y_{ij} - 1)e^{-\lambda_i - \theta_iy_{ij}}.
\]

so that the log-likelihood function is

\[
\ell(\phi, \lambda, \theta | Y_{\text{obs}}) = n_0 \log(\phi + (1 - \phi)e^{-\lambda_+}) + (n - n_0) \log(1 - \phi) - \sum_{i=1}^{m} n_{i0} \lambda_i + \sum_{i=1}^{m} \sum_{j \in \mathcal{J}_i} \left[ \log(\lambda_i) + (y_{ij} - 1) \log(\lambda_i + \theta_iy_{ij}) - \lambda_i - \theta_iy_{ij} \right].
\]

298 X.-F. Huang et al.
where the parameters $\phi$, $\lambda$, $\theta$ are separated, $C$ is a constant not involving $(\phi, \lambda, \theta)$, and

\[
\begin{align*}
Q_1(Y) &= \frac{n_0 \phi^{(t)}}{\beta^{(t)}} \log(\phi) + \left( n - n_0 \phi^{(t)} \right) \log(1 - \phi), \\
Q_2(Y) &= -\frac{n_0 (\beta^{(t)} - \phi^{(t)})}{\beta^{(t)}} \lambda + \sum_{i=1}^{m} \left( \frac{y_{ij} - 1}{\lambda_i} \right) \log(\lambda_i) \\
&\quad + \sum_{i=1}^{m} \left[ (n - n_0 - n_{i0}) \log(\lambda_i) - (n - n_0) \lambda_i \right], \\
Q_3(Y) &= \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left( \frac{\theta^{(t)}_i y_{ij}(y_{ij} - 1)}{\lambda_i + \theta^{(t)}_i y_{ij}} \right) \log(\theta_i) - \theta_i y_{ij}.
\end{align*}
\]

Therefore, the explicit MM iterations are given by

\[
\begin{align*}
\phi^{(t+1)} &= \frac{n_0 \phi^{(t)}}{n^{(t+1)}}, \\
\lambda^{(t+1)} &= \frac{n - n_0 - n_{i0} + \sum_{j=1}^{n_i} (y_{ij} - 1) \lambda^{(t)}_i}{n - n_0 \phi^{(t+1)}}, \\
\theta^{(t+1)}_i &= \frac{\sum_{j=1}^{n_i} \left( \frac{\theta^{(t)}_i y_{ij}(y_{ij} - 1)}{\lambda_i^{(t)} + \theta^{(t)}_i y_{ij}} \right)}{\sum_{j=1}^{n_i} y_{ij}},
\end{align*}
\]

for $i = 1, \ldots, m$.

### 3.4 Bootstrap confidence intervals for small sample sizes

The Wald confidence interval (CI) of $\phi$ specified by (3.4) may fall outside the unit interval $[0, 1]$. The Wald CIs of $\{\lambda_i\}_{i=1}^{m}$ and $\{\theta_i\}_{i=1}^{m}$ given by (3.4) are reliable only for large sample sizes. For small sample sizes, the bootstrap method is a useful tool to find CI for an arbitrary function of $\phi$, $\{\lambda_i\}_{i=1}^{m}$ and $\{\theta_i\}_{i=1}^{m}$, say, $\hat{\vartheta} = h(\hat{\phi}, \hat{\lambda}_1, \ldots, \hat{\lambda}_m, \hat{\theta}_1, \ldots, \hat{\theta}_m)$. Let $\vartheta = h(\hat{\phi}, \lambda_1, \ldots, \lambda_m, \theta_1, \ldots, \theta_m)$ denote the MLE of $\vartheta$, where $\hat{\phi}, \{\hat{\lambda}_i\}_{i=1}^{m}$ and $\{\hat{\theta}_i\}_{i=1}^{m}$ represent the respective MLEs of $\phi, \{\lambda_i\}_{i=1}^{m}$ and $\{\theta_i\}_{i=1}^{m}$ calculated by means of the second EM algorithm (3.14)–(3.16) or the MM algorithm (3.19). Based on the obtained MLEs $\hat{\phi}, \{\hat{\lambda}_i\}_{i=1}^{m}$ and $\{\hat{\theta}_i\}_{i=1}^{m}$, we can generate

\[
y_1^*, \ldots, y_m^* \overset{iid}{\sim} \text{ZIGP}^{(t)}(\hat{\phi}, \hat{\lambda}_1, \ldots, \hat{\lambda}_m, \hat{\theta}_1, \ldots, \hat{\theta}_m).
\]

Having obtained $Y_{\text{obs}}^* = \{y_1^*, \ldots, y_m^*\}$, we can calculate the bootstrap replication $\hat{\vartheta}^* = \{\hat{\lambda}_1^*, \ldots, \hat{\lambda}_m^*, \hat{\theta}_1^*, \ldots, \hat{\theta}_m^*\}$. Independently repeating this process $G$ times, we obtain $G$ bootstrap replications $\{\hat{\vartheta}^*_g\}_{g=1}^{G}$. Consequently, the standard error, $\hat{se}(\hat{\vartheta})$, of $\hat{\vartheta}$ can be estimated by the sample standard deviation of the $G$ replications, i.e.,

\[
\hat{se}(\hat{\vartheta}) = \left\{ \frac{1}{G-1} \sum_{g=1}^{G} \left( \hat{\vartheta}^*_g - \hat{\vartheta} \right)^2 \right\}^{1/2}.
\]

If $\{\hat{\vartheta}^*_g\}_{g=1}^{G}$ is approximately normally distributed, the first $(1-\alpha)100\%$ bootstrap CI for $\vartheta$ is

\[
\left[ \hat{\vartheta} - z_{\alpha/2} \cdot \hat{se}(\hat{\vartheta}), \hat{\vartheta} + z_{\alpha/2} \cdot \hat{se}(\hat{\vartheta}) \right].
\]

Alternatively, if $\{\hat{\vartheta}^*_g\}_{g=1}^{G}$ is non-normally distributed, the second $(1-\alpha)100\%$ bootstrap CI of $\vartheta$ can be obtained as

\[
\left[ \hat{\vartheta}_L, \hat{\vartheta}_U \right],
\]

where $\hat{\vartheta}_L$ and $\hat{\vartheta}_U$ are the $100(\alpha/2)$ and $100(1-\alpha/2)$ percentiles of $\{\hat{\vartheta}^*_g\}_{g=1}^{G}$, respectively.

### 3.5 Testing hypotheses for large sample sizes

#### 3.5.1 Likelihood ratio test for zero inflation

Suppose we want to test the null hypothesis

\[
H_0: \phi = 0 \quad \text{against} \quad H_1: \phi > 0.
\]

Under $H_0$, the likelihood ratio test (LRT) statistic (Jansakul and Hinde, 2002, p. 78 [9]; Joe and Zhu, 2005, p. 225 [10])

\[
T_1 = -2 \left\{ \ell(0, \hat{\lambda}_0, \hat{\theta}_0|Y_{\text{obs}}) - \ell(\hat{\phi}, \hat{\lambda}, \hat{\theta}|Y_{\text{obs}}) \right\} \sim 0.5 \chi^2(0) + 0.5 \chi^2(1),
\]

where $\hat{\lambda}_0$ and $\hat{\theta}_0$ are the MLEs of $\lambda$ and $\theta$ under $H_0$, $(\hat{\phi}, \hat{\lambda}, \hat{\theta})$ are the unconstrained MLEs of $(\phi, \lambda, \theta)$, and $\chi^2(0)$ denotes the degenerate distribution with all mass at zero. The corresponding $p$-value is

\[
p_{01} = \Pr(T_1 > t_1|H_0) = \frac{1}{2} \Pr(\chi^2(1) > t_1),
\]

where $t_1$ is the realization of the LRT statistic $T_1$.

#### 3.5.2 Score test for zero inflation

In this subsection, we will develop a score test for testing zero inflation in the Type I multivariate ZIGP model by reparametrization. Let

\[
\gamma = \frac{\phi}{1 - \phi},
\]

then, testing $H_0$ specified by (3.23) is equivalent to testing $H^*_0: \gamma = 0$. The observed-data log-likelihood function now becomes

\[
\ell^* = \ell(\gamma, \lambda, \theta|Y_{\text{obs}}).
\]

*Type I multivariate ZIGP distribution* 299
\[ U(\gamma, \lambda, \theta) = \left( \frac{\partial \ell^*}{\partial \gamma}, \frac{\partial \ell^*}{\partial \lambda}, \frac{\partial \ell^*}{\partial \theta} \right)^T \]
and the Fisher information matrix is
\[ J(\gamma, \lambda, \theta) = (J_{jk}) = E[\mathbf{I}(\gamma, \lambda, \theta|Y_{\text{obs}})], \]
see Appendix C.

Under \( H_0^* \), the score test statistic
\[ (3.27) \quad T_2 = U^T(\hat{\gamma}_0, \hat{\lambda}_0, \hat{\theta}_0)J^{-1}(\hat{\gamma}_0, \hat{\lambda}_0, \hat{\theta}_0)U(\gamma, \lambda, \theta) \sim \chi^2(1), \]
where \( \gamma_0 = 0, \lambda_0 \) and \( \theta_0 \) denote the MLEs of \( \lambda \) and \( \theta \) under \( H_0^* \). The corresponding \( p \)-value is given by
\[ (3.28) \quad p_{CV} = \Pr(T_2 > t_2|H_0) = \Pr(\chi^2(1) > t_2), \]
where \( t_2 \) is the realization of the score test statistic \( T_2 \).

### 3.5.3 Likelihood ratio test for testing equality of all \( \lambda \)'s
Suppose we want to test the null hypothesis
\[ (3.29) \quad H_0: \lambda_1 = \cdots = \lambda_m = \lambda \quad \text{vs} \quad H_1: H_0 \text{ is not true.} \]
Under \( H_0 \), the LRT statistic
\[ (3.30) \quad T_3 = -2\{\ell(\hat{\phi}_0, \hat{\lambda}_0, \hat{\theta}_0|Y_{\text{obs}}) - \ell(\hat{\phi}, \hat{\lambda}, \hat{\theta}|Y_{\text{obs}})\} \sim \chi^2(m-1), \]
where \( (\hat{\phi}_0, \hat{\lambda}_0, \hat{\theta}_0) \) are the MLEs of \( (\phi, \lambda, \theta) \) under \( H_0 \), and \( (\hat{\phi}, \hat{\lambda}, \hat{\theta}) \) are the unconstrained MLEs of \( (\phi, \lambda, \theta) \). The corresponding \( p \)-value is given by
\[ (3.31) \quad p_{CV} = \begin{cases} 2 \min\{\Pr(T_3 > t_3|H_0), \Pr(T_3 \leq t_3|H_0)\}, & \text{if } m > 3, \\ \Pr(T_3 > t_3|H_0), & \text{if } m = 2, 3, \end{cases} \]
where \( t_3 \) is the realization of the LRT statistic \( T_3 \). When \( p_{CV} > \alpha \), we cannot reject the null hypothesis \( H_0 \) at the \( \alpha \) level of significance. However, if \( H_0 \) specified by (3.29) is rejected, we could consider to test \( H_0^* \), where \( i, j = 1, \ldots, m; i \neq j \), and the corresponding test statistic follows \( \chi^2(1) \).

### 3.5.4 Score test for testing equality of all \( \lambda \)'s
Let \( \gamma \) be defined by (3.26), then, we apply score test to test \( H_0 \) specified by (3.29). Under \( H_0 \), the score test statistic
\[ (3.32) \quad T_4 = U^T(\gamma_0, \hat{\lambda}_0, \hat{\theta}_0)J^{-1}(\gamma_0, \hat{\lambda}_0, \hat{\theta}_0)U(\gamma_0, \hat{\lambda}_0, \hat{\theta}_0) \sim \chi^2(m-1), \]
where \( \hat{\lambda}_0 = \hat{\lambda}^1 \), and \( (\gamma_0, \hat{\lambda}_0, \hat{\theta}_0) \) are the MLEs of \( (\gamma, \lambda, \theta) \) under \( H_0 \). Hence, the \( p \)-value is
\[ (3.33) \quad p_{CV} = \begin{cases} 2 \min\{\Pr(T_4 > t_4|H_0), \Pr(T_4 \leq t_4|H_0)\}, & \text{if } m > 3, \\ \Pr(T_4 > t_4|H_0), & \text{if } m = 2 \text{ or } 3, \end{cases} \]
where \( t_4 \) is the realization of the score test statistic \( T_4 \).
Table 1. MLEs and bootstrap CIs of parameters for $m = 2$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Value</th>
<th>MLE Width</th>
<th>CP</th>
<th>True Value</th>
<th>MLE Width</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.1</td>
<td>0.1004</td>
<td>0.1280</td>
<td>0.937</td>
<td>0.2</td>
<td>0.1998</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>2</td>
<td>2.0285</td>
<td>0.8970</td>
<td>0.960</td>
<td>3</td>
<td>3.0551</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>2</td>
<td>2.0437</td>
<td>0.9021</td>
<td>0.941</td>
<td>3</td>
<td>3.0396</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.3</td>
<td>0.2881</td>
<td>0.2604</td>
<td>0.933</td>
<td>0.4</td>
<td>0.3875</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.3</td>
<td>0.2834</td>
<td>0.2692</td>
<td>0.925</td>
<td>0.4</td>
<td>0.3891</td>
</tr>
</tbody>
</table>

Note: MLE is the mean of the 1000 point estimates via the EM algorithm (3.14)–(3.16); width and CP are the average width and coverage proportion of 1000 bootstrap CIs.

Table 2. MLEs and bootstrap CIs of parameters for $m = 3$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>True Value</th>
<th>MLE Width</th>
<th>CP</th>
<th>True Value</th>
<th>MLE Width</th>
<th>CP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.1</td>
<td>0.1004</td>
<td>0.1150</td>
<td>0.937</td>
<td>0.2</td>
<td>0.1994</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>2</td>
<td>2.0218</td>
<td>0.8579</td>
<td>0.941</td>
<td>3</td>
<td>3.0389</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>4</td>
<td>4.0471</td>
<td>1.4900</td>
<td>0.941</td>
<td>5</td>
<td>5.0745</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>6</td>
<td>6.1214</td>
<td>2.1171</td>
<td>0.930</td>
<td>5</td>
<td>5.7129</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>0.3</td>
<td>0.2898</td>
<td>0.2555</td>
<td>0.920</td>
<td>0.2</td>
<td>0.1867</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>0.5</td>
<td>0.4926</td>
<td>0.1865</td>
<td>0.934</td>
<td>0.4</td>
<td>0.3919</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>0.7</td>
<td>0.6911</td>
<td>0.1197</td>
<td>0.928</td>
<td>0.6</td>
<td>0.5912</td>
</tr>
</tbody>
</table>

Note: MLE is the mean of the 1000 point estimates via the EM algorithm (3.14)–(3.16); width and CP are the average width and coverage proportion of 1000 bootstrap CIs.

4. SIMULATION STUDIES

To evaluate the performance of the proposed statistical methods in Section 3 for the Type I multivariate ZIGP distribution, we first investigate the accuracy of point estimates and confidence interval estimates for different parameter settings via simulation studies. Second, we assess the performance of the LRT with the score test by comparing their type I error rates and powers.

4.1 Accuracy of point estimates and interval estimates

In this subsection, we compare the accuracy of point estimates and confidence intervals by considering both cases of two-dimensional (i.e., $m = 2$) and three-dimensional (i.e., $m = 3$). When $m = 2$, the parameters $(\phi, \lambda_1, \lambda_2, \theta_1, \theta_2)$ are set to be $(0.1, 2, 2, 0.3, 0.3)$ and $(0.2, 3, 3, 0.4, 0.4)$. When $m = 3$, the parameters $(\phi, \lambda_1, \lambda_2, \lambda_3, \theta_1, \theta_2, \theta_3)$ are set to be $(0.1, 2, 4, 6, 0.3, 0.5, 0.7)$ and $(0.2, 3, 5, 7, 0.2, 0.4, 0.6)$. For each parameter configuration, we generate

\[
\{y_j\}_{j=1}^n \overset{\text{iid}}{\sim} \text{ZIGP}_m(\phi, \lambda, \theta)
\]

with $n = 100$, and calculate the MLEs via the second EM algorithm (3.14)–(3.16) and the 95% bootstrap CIs with $G = 1,000$. Here, we independently repeat this process 1,000 times and report the corresponding mean of the MLEs, the average width and the coverage probability (CP) of the bootstrap CIs in Tables 1 and 2, respectively.

4.2 Comparison of the LRT with the score test

4.2.1 Tests for zero inflation

In this subsection, we compare the corresponding type I error rates (with $H_0: \phi = 0$) and powers (with $H_1: \phi > 0$) between the LRT and the score test for various sample sizes via simulations, where the values of $\phi$ in $H_1$ are chosen to be 0.01, 0.03, 0.05, 0.07, 0.10, 0.15. For a given pair of $(n, \phi)$, we first draw

\[
Z_{1l}^{(i)}, \ldots, Z_{nl}^{(i)} \overset{\text{iid}}{\sim} \text{Bernoulli}(1 - \phi)
\]

for $l = 1, \ldots, L (L = 1,000)$, and then independently generate

\[
X_{11}^{(i)} \ldots X_{1n}^{(i)} \overset{\text{iid}}{\sim} \text{GP}(\lambda_1, \theta_1)
\]

and

\[
X_{21}^{(i)} \ldots X_{2n}^{(i)} \overset{\text{iid}}{\sim} \text{GP}(\lambda_2, \theta_2),
\]

where only $\lambda_1 = 5, \theta_1 = 0.4$ and $\lambda_2 = 3, \theta_2 = 0.6$ are considered. Finally, we set

\[
y_j^{(i)} = \begin{pmatrix} y_{j1}^{(i)} \\ y_{j2}^{(i)} \end{pmatrix} = Z_{j1}^{(i)} \begin{pmatrix} X_{11}^{(i)} \\ X_{21}^{(i)} \end{pmatrix}, \quad j = 1, \ldots , n.
\]

All hypothesis testings are conducted at the significant level $\alpha = 0.05$. Let $r_k$ denote the number of rejecting the null hypothesis $H_0: \phi = 0$ by the test statistics $T_k$ ($k = 1, 2$) given by (3.24) and (3.27), respectively. Hence, the actual significance level can be estimated by $r_k/L$ with $\phi = 0$ and

Type I multivariate ZIGP distribution 301
the power of the test statistic $T_k$ can be estimated by $r_k/L$ with $\phi > 0$.

Figure 1 shows that the comparison of type I error rates between the LRT and the score test. In general, we can see the LRT test have a explicitly better performance in controlling its type I error rates around the pre-chosen nominal level.

Figure 2 gives the comparison of powers between the LRT and the score test for different values of $\phi > 0$. It is not difficult to find that there is no significant difference between the powers of the two tests when $\phi$ is larger than 0.03. But when $\phi = 0.01$, the score test is slightly more powerful than the LRT.

4.2.2 Tests for equality of $\lambda_1$ and $\lambda_2$

In this subsection, we compare the respective type I error rates (with $H_0: \lambda_1 = \lambda_2$) and powers (with $H_1: \lambda_1 \neq \lambda_2$) between the LRT and the score test for various sample sizes and different combinations of $(\lambda_1, \lambda_2)$ via simulations, where the values of $(\lambda_1, \lambda_2)$ are set to be (4, 4) and (5, 8). For a given combination of $(n, \lambda_1, \lambda_2)$, we first generate

$$Z_{1}^{(l)}, \ldots, Z_{n}^{(l)} \sim \text{Bernoulli}(1 - \phi)$$

for $l = 1, \ldots, L (L = 1,000)$, and then independently generate

$$X_{11}^{(l)}, \ldots, X_{1n}^{(l)} \sim \text{GP}(\lambda_1, \theta_1)$$

and

$$X_{21}^{(l)}, \ldots, X_{2n}^{(l)} \sim \text{GP}(\lambda_2, \theta_2),$$

where only $\phi = 0.5, \theta_1 = 0.4, \theta_2 = 0.6$ are considered. Then, we have

$$Y_{j}^{(l)} = \left( Y_{1j}^{(l)}, Y_{2j}^{(l)} \right) = Z_{j}^{(l)} \left( X_{1j}^{(l)}, X_{2j}^{(l)} \right), \quad j = 1, \ldots, n.$$ 

All hypothesis testing are conducted at the significant level $\alpha = 0.05$. Let $r_k$ denote the number of rejecting the null hypothesis $H_0: \lambda_1 = \lambda_2$ by the statistics $T_k$ ($k = 3, 4$) given by (3.30) and (3.32), respectively. Hence, the actual significance level can be estimated by $r_k/L$ with $\lambda_1 = \lambda_2$ and the power of the test statistic $T_k$ can be estimated by $r_k/L$ with $\lambda_1 \neq \lambda_2$.

Figure 3 shows that some comparison of type I error rates between the LRT and the score test. In general, there is no significance difference between the two tests’ performances in controlling their type I error rates around the pre-chosen nominal level.

Figure 4 gives the comparison of powers between the LRT and the score test for one case with $\lambda_1 \neq \lambda_2$. It is not difficult to find that the LRT almost has the same power as the score test is, no matter the sample size is small or large.

4.2.3 Tests for equality of $\theta_1$ and $\theta_2$

In this subsection, we compare the respective type I error rates (with $H_0: \theta_1 = \theta_2$) and powers (with $H_1: \theta_1 \neq \theta_2$) between the LRT and the score test for various sample sizes and different combinations of $(\theta_1, \theta_2)$ via simulations, where the values of $(\theta_1, \theta_2)$ are set to be (0.6, 0.6) and (0.3, 0.7).
For a given combination of \((n, \theta_1, \theta_2)\), we first generate
\[ Z^{(l)}_1, \ldots, Z^{(l)}_n \overset{iid}{\sim} \text{Bernoulli}(1 - \phi) \]
for \(l = 1, \ldots, L \) \((L = 1,000)\), and then independently generate
\[ X^{(l)}_{11}, \ldots, X^{(l)}_{1n} \overset{iid}{\sim} \text{GP}(\lambda_1, \theta_1) \]
and
\[ X^{(l)}_{21}, \ldots, X^{(l)}_{2n} \overset{iid}{\sim} \text{GP}(\lambda_2, \theta_2), \]
where only \(\phi = 0.5, \lambda_1 = 5, \lambda_2 = 8\) are considered. Then, we have
\[ y^{(l)}_j = \begin{pmatrix} y^{(l)}_{1j} \\ y^{(l)}_{2j} \end{pmatrix} = Z^{(l)}_j \begin{pmatrix} X^{(l)}_{1j} \\ X^{(l)}_{2j} \end{pmatrix}, \quad j = 1, \ldots, n. \]

All hypothesis testings are conducted at the significant level \(\alpha = 0.05\). Let \(r_k\) denote the number of rejecting the null hypothesis \(H_0: \theta_1 = \theta_2\) by the statistics \(T_k\) \((k = 3, 4)\) given by (3.35) and (3.37), respectively. Hence, the actual significance level can be estimated by \(r_k/L\) with \(\theta_1 = \theta_2\) and the power of the test statistic \(T_k\) can be estimated by \(r_k/L\) with \(\theta_1 \neq \theta_2\).

Figure 3 shows that some comparison of type I error rates between the LRT and the score test. In general, there is no significance difference between the two tests’ performances in controlling their type I error rates around the pre-chosen nominal level.

Figure 6 gives the comparison of powers between the LRT and the score test for one case with \(\theta_1 \neq \theta_2\). It is not difficult to find that the LRT almost has the same power as the score test is, no matter the sample size is small or large.

5. TWO REAL EXAMPLES

In this section, two real data sets are used to illustrate the proposed methods, where the Newton–Raphson algorithm for finding the MLEs of parameters does not work for the two examples because the corresponding observed information matrices are nearly singular, while the Fisher-scoring algorithm is always sensitive to the initial values. Unfortunately, the first EM algorithm does not work in the second example. As expected, the second EM algorithm and the MM algorithm work well in the two examples.

5.1 The children’s absenteeism data in Indonesia

In a survey of Indonesian family life conducted by Strauss et al. (2004) [18], the participants included 7,000 households sampled from 321 communities randomly selected from 13 of the nation’s 26 Provinces, in which 83% of the Indonesian population lived. Among those households with one child per household, 437 household heads were asked questions about the health of their children. Let \(Y_1\) denote the number of days the children missed their primary activities due to illness in the last four weeks and \(Y_2\) denote the number of days the children spent in bed due to illness in the last four weeks. Table 3 shows the children’s absenteeism data from this survey.
Figure 5. Comparison of type I error rates between the LRT (solid line) and the score test (dotted line).

Table 3. The children’s absenteeism data in the Indonesian family life survey (Cheung and Lam, 2006)

<table>
<thead>
<tr>
<th></th>
<th>Y_1</th>
<th>Y_2</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>323</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>323</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>3</td>
<td>19</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>391</td>
<td>12</td>
<td>16</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>437</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.1.1 Likelihood-based inferences

Let y_{i1}, \ldots, y_{in} \sim \text{ZIGP}^{(I)}(\phi, \lambda_1, \lambda_2, \theta_1, \theta_2), where y_j = (Y_{1j}, Y_{2j})^T for j = 1, \ldots, n (n = 437). To find the MLEs of (\phi, \lambda_1, \lambda_2, \theta_1, \theta_2), we randomly choose (\phi^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)}, \theta_1^{(0)}, \theta_2^{(0)}) = (0.5, 0.1, 0.2, 0.2, 0.2) as their initial values of the two EM algorithms and MM algorithm, and carefully choose initial values (\phi^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)}, \theta_1^{(0)}, \theta_2^{(0)}) = (0.5, 1, 1, 0.2, 0.2) for the Fisher-scoring algorithm. The MLEs of (\phi, \lambda_1, \lambda_2, \theta_1, \theta_2) converged to (\hat{\phi}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta}_1, \hat{\theta}_2) as shown in the second column of Table 4 in 13 iterations for the Fisher-scoring algorithm (3.3), in 105 iterations for the first EM algorithm (3.10)–(3.12), in 17 iterations for the second EM algorithm (3.14)–(3.16) and in 89 iterations for the MM algorithm (3.19). The standard errors of the MLEs (\hat{\phi}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta}_1, \hat{\theta}_2) are given in the third column and 95% asymptotic Wald CIs (i.e., (3.4)) of the five parameters are listed in the fourth column of Table 4. With G = 10,000 bootstrap replications, the two 95% bootstrap CIs of (\phi, \lambda_1, \lambda_2, \theta_1, \theta_2) are shown in the sixth and seventh columns of Table 4.

Suppose that we want to test the null hypothesis H_0: \phi = 0 against the alternative hypothesis H_1: \phi > 0. According to (3.24) and (3.27), we calculate the values of the LRT statistic and score test statistic, which are given by t_1 = 182.6755 and t_2 = 178.9367, respectively. Then from (3.25) and (3.28), we have p_{t_1} = p_{t_2} \approx 0 < \alpha = 0.05. Thus, we should reject H_0.

If we want to test the null hypothesis H_0: \lambda_1 = \lambda_2 against the alternative hypothesis H_1: \lambda_1 \neq \lambda_2. According to (3.30) and (3.32), we calculate the values of the LRT statistic and score test statistic, which are given by t_3 = 96.44369 and t_4 = 85.29277, respectively. Then from (3.31) and (3.33), we have p_{t_3} = p_{t_4} \approx 0 < \alpha = 0.05. As a result, the H_0 should be rejected.

Suppose that we want to test the null hypothesis H_0: \theta_1 = \theta_2 against the alternative hypothesis H_1: \theta_1 \neq \theta_2. According to (3.35) and (3.37), we calculate the values of the LRT statistic and score test statistic, which are given
Table 4. MLEs and CIs of parameters for the children’s absenteeism data in Indonesia

<table>
<thead>
<tr>
<th>Parameter</th>
<th>MLE</th>
<th>std(^{0})</th>
<th>95% Wald CI</th>
<th>std(^{0})</th>
<th>95% CI(^{1})</th>
<th>95% CI(^{0})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\phi)</td>
<td>0.7252</td>
<td>0.0225</td>
<td>[0.6811, 0.7693]</td>
<td>0.0147</td>
<td>[0.6964, 0.7538]</td>
<td>[0.6961, 0.7536]</td>
</tr>
<tr>
<td>(\lambda_1)</td>
<td>2.4618</td>
<td>0.2563</td>
<td>[1.9595, 2.9612]</td>
<td>0.1708</td>
<td>[2.1420, 2.8117]</td>
<td>[2.1553, 2.8269]</td>
</tr>
<tr>
<td>(\lambda_2)</td>
<td>0.5208</td>
<td>0.0758</td>
<td>[0.3721, 0.6694]</td>
<td>0.0501</td>
<td>[0.4247, 0.6210]</td>
<td>[0.4285, 0.6259]</td>
</tr>
<tr>
<td>(\theta_1)</td>
<td>0.2772</td>
<td>0.0581</td>
<td>[0.1633, 0.3911]</td>
<td>0.0389</td>
<td>[0.1965, 0.3489]</td>
<td>[0.1932, 0.3448]</td>
</tr>
<tr>
<td>(\theta_2)</td>
<td>0.5076</td>
<td>0.0708</td>
<td>[0.3688, 0.6464]</td>
<td>0.0477</td>
<td>[0.4087, 0.5957]</td>
<td>[0.4047, 0.5896]</td>
</tr>
</tbody>
</table>

std\(^{0}\): The square roots of the diagonal elements of the inverse Fisher information matrix \(J^{-1}(\phi, \lambda, \theta)\). std\(^{0}\): The sample standard deviation of the bootstrap samples, cf. (3.20). CI\(^{1}\): Normal-based bootstrap CI, cf. (3.21). CI\(^{0}\): Non-normal-based bootstrap CI, cf. (3.22).

Table 5. Comparisons for Type I multivariate ZIGP distribution and the Type I multivariate ZIP distribution

<table>
<thead>
<tr>
<th>Model</th>
<th>Criterion</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I multivariate ZIGP distribution</td>
<td>1324.855</td>
<td>1345.255</td>
<td></td>
</tr>
<tr>
<td>Type I multivariate ZIP distribution</td>
<td>1345.731</td>
<td>1447.971</td>
<td></td>
</tr>
</tbody>
</table>

by \(t_5 = 7.086542925\) and \(t_6 = 7.914848236\), respectively. Then from (3.36) and (3.38), we have \(p_{\lambda_1} = 0.007766492 < 0.05\), \(p_{\lambda_2} = 0.004903069 < 0.05\). Thus, we should reject \(H_0\).

5.1.2 Model comparison

Now we focus on the comparison between the Type I multivariate ZIGP model with the Type I multivariate ZIP model under AIC and BIC based on the full likelihood function. In Table 5, we can find that both AIC and BIC of the Type I multivariate ZIGP model are less than those of the Type I multivariate ZIP model, indicating that the proposed Type I multivariate ZIGP model is more appropriate to fit the data set.

5.2 Voluntary and involuntary job changes data


Table 6. Cross tabulation of voluntary and involuntary job changes (Jung and Winkelmann, 1993)

<table>
<thead>
<tr>
<th>(Y_1\backslash Y_2)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1227</td>
<td>319</td>
<td>109</td>
<td>27</td>
<td>20</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1713</td>
</tr>
<tr>
<td>1</td>
<td>150</td>
<td>83</td>
<td>23</td>
<td>10</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>274</td>
</tr>
<tr>
<td>2</td>
<td>34</td>
<td>16</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Total 1463 427 139 45 25 10 4 4 1 2 3 1 2124

According to (3.30) and (3.32), we have \(p_{\lambda_1} = 0.000106 < 0.05\), \(p_{\lambda_2} = 0.000019 < \alpha = 0.05\). Thus, we should reject \(H_0\).

If we want to test the null hypothesis \(H_0: \phi = 0\) against the alternative hypothesis \(H_1: \phi > 0\), According to (3.24) and (3.27), we calculate the values of the LRT statistic and score test statistic, which are given by \(t_1 = 15.02807\) and \(t_2 = 16.98990\), respectively. Then from (3.25) and (3.28), we have \(p_{\lambda_1} = 0.000106 < 0.05\), \(p_{\lambda_2} = 0.000019 < \alpha = 0.05\). Thus, we should reject \(H_0\).

Suppose that we want to test the null hypothesis \(H_0: \lambda_1 = \lambda_2\) against the alternative hypothesis \(H_1: \lambda_1 \neq \lambda_2\). According to (3.14) and (3.16), and in 199 iterations for the second EM algorithm (3.14)–(3.16) and in 219 iterations for the MM algorithm (3.19). The standard errors of the MLEs \((\hat{\phi}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{\theta}_1, \hat{\theta}_2)\) are given in the third column and 95% asymptotic Wald CIs (i.e., (3.4)) of the five parameters are listed in the fourth column of Table 7. With \(G = 10,000\) bootstrap replications, the two 95% bootstrap CIs of \((\phi, \lambda_1, \lambda_2, \theta_1, \theta_2)\) are shown in the sixth and seventh columns of Table 7.

Suppose that we want to test the null hypothesis \(H_0: \theta_1 = \theta_2\) against the alternative hypothesis \(H_1: \theta_1 \neq \theta_2\). According to (3.5) and (3.7), we calculate the values of the second EM algorithm (3.14)–(3.16) and in 219 iterations for the Fisher-scoring algorithm (3.3), in 180 iterations for the MM algorithm (3.19).

Type I multivariate ZIGP distribution 305
Criterion and (3.38), we have computational advantages.

The four proposed algorithms: from Fisher scoring to two vals and the bootstrap method can be easily calculated via inference, actually, the MLE procedure for the ZIGP model

the model parameters. Two data sets in literature have inference approaches via four different algorithms concern-

\[ \theta \]

Type I multivariate ZIP distribution

Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model</th>
<th>MLE</th>
<th>std</th>
<th>95% Wald CI (std)</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_1 )</td>
<td>Type I multivariate ZIGP distribution</td>
<td>0.2680</td>
<td>0.0208</td>
<td>[0.2272, 0.3089]</td>
<td>7882</td>
<td>8000</td>
</tr>
<tr>
<td>( \lambda_2 )</td>
<td>Type I multivariate ZIGP distribution</td>
<td>0.4738</td>
<td>0.0360</td>
<td>[0.4032, 0.5444]</td>
<td>94288</td>
<td>95470</td>
</tr>
<tr>
<td>( \theta_1 )</td>
<td>Type I multivariate ZIGP distribution</td>
<td>0.3928</td>
<td>0.0254</td>
<td>[0.3430, 0.4426]</td>
<td>0.0943</td>
<td>0.0396</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>Type I multivariate ZIGP distribution</td>
<td>0.2690</td>
<td>0.0219</td>
<td>[0.2260, 0.3120]</td>
<td>0.1493</td>
<td>[-0.0688, 0.5165]</td>
</tr>
</tbody>
</table>

\[ \text{std}^2: \text{The square roots of the diagonal elements of the inverse Fisher information matrix } J^{-1}(\phi, \lambda, \theta), \text{ std}^3: \text{The sample standard deviation of the bootstrap samples, cf. (3.20).} \]

\[ \text{CI}^1: \text{Normal-based bootstrap CI, cf. (3.21).} \]

\[ \text{CI}^2: \text{Non-normal-based bootstrap CI, cf. (3.22).} \]

**Table 7. MLEs and CIs of parameters for the voluntary and involuntary job changes data**

**Table 8. Comparisons for Type I multivariate ZIGP distribution and the Type I multivariate ZIP distribution**

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I multivariate ZIGP distribution</td>
<td>7182.795</td>
<td>7211.100</td>
</tr>
<tr>
<td>Type I multivariate ZIP distribution</td>
<td>7894.818</td>
<td>7911.801</td>
</tr>
</tbody>
</table>

the LRT statistic and score test statistic, which are given by \( t_5 = 15.7782 \) and \( t_6 = 16.94288 \), respectively. Then from (3.36) and (3.38), we have \( p_1 = 0.000071 \ll 0.05 \), \( p_2 = 0.000038 \ll 0.05 \). Thus, we should reject \( H_0 \) at 0.05 level of significance.

5.2.2 Model comparison

Now we focus on the comparison between the Type I multivariate ZIGP model with the Type I multivariate ZIP model under AIC and BIC based on the full likelihood function. In Table 8, we can see that both AIC and BIC of the Type I multivariate ZIGP model are less than that of the Type I multivariate ZIP model, indicating that the data set is fitted more appropriately by using the proposed Type I multivariate ZIGP model compared with the Type I multivariate ZIP model.

6. DISCUSSION

In this paper we have introduced a multivariate ZIGP distribution, called the Type I multivariate ZIGP model, and developed the distribution theory and its important properties. We have also investigated the efficient likelihood inference approaches via four different algorithms concerning the model parameters. Two data sets in literature have been used to illustrate the applications. For the likelihood inference, actually, the MLE procedure for the ZIGP model is difficult especially when the dimension is large. For the propose model, however, the MLEs, the confidence intervals and the bootstrap method can be easily calculated via the four proposed algorithms: from Fisher scoring to two EM algorithms, to MM algorithm, thus offering substantial computational advantages.

Xuejun Jiang’s research was partially supported by a grant from the Natural Science Foundation of China (No. 11101432). The authors contributed equally to this work.

**Appendix A: The derivation of (3.2)**

Since \( y_1, \ldots, y_n \overset{iid}{\sim} \text{ZIGP}(1)(\phi, \lambda, \theta) \), we have

\[ E \left[ \sum_{j=1}^{n} I(y_j = 0) \right] = nE[I(Y_1 = 0)] = nPr(Y_1 = 0) \]

\[ = n[\phi + (1 - \phi)e^{-\lambda y}] \]

which implies the first formula of (3.2). In the follows, we assume that \( Y \sim \text{ZIGP}(\phi, \lambda, \theta) \) and only need to prove

\[ E \left[ \frac{Y - 1}{(\lambda + \theta Y)^2} \right] = \frac{1 - \phi}{\lambda} - \frac{1}{\lambda^2} - \frac{\theta(1 - \phi)}{\lambda + 2\theta}, \]

\[ E \left[ \frac{Y^3 - Y^2}{(\lambda + \theta Y)^2} \right] = \frac{\lambda(1 - \phi)}{1 - \theta} + \frac{2\lambda(1 - \phi)}{\lambda + 2\theta}, \]

\[ E \left[ \frac{Y^2 - Y}{(\lambda + \theta Y)^2} \right] = \frac{\lambda(1 - \phi)}{\lambda + 2\theta}. \]

Let

\[ s = \sum_{y=0}^{\infty} \frac{(\lambda + \theta y)^{y-1}e^{-(\lambda + \theta y)}}{y!}. \]

Because

\[ \lambda \times s = \sum_{y=0}^{\infty} \frac{\lambda(\lambda + \theta y)^{y-1}e^{-(\lambda + \theta y)}}{y!} = 1, \]

then we obtain \( s = 1/\lambda \). On the one hand, we have

\[ s = \sum_{y=0}^{\infty} \frac{(\lambda + \theta y)^{y-1}e^{-(\lambda + \theta y)}}{y!} \]

\[ = \sum_{y=0}^{\infty} \frac{\lambda(\lambda + \theta y)^{y-2}e^{-(\lambda + \theta y)}}{y!} + \sum_{y=0}^{\infty} \frac{\theta y(\lambda + \theta y)^{y-2}e^{-(\lambda + \theta y)}}{y!} \]

306 X.-F. Huang et al.
\[
\lambda \sum_{y=0}^{\infty} \frac{(\lambda + \theta y)^{y-2}e^{-(\lambda + \theta y)}}{y!} + \theta \sum_{y=1}^{\infty} \frac{[\lambda + \theta + \theta(y - 1)]^{y-1}e^{-[\lambda + \theta + \theta(y-1)]}}{(y - 1)!}
= \lambda \sum_{y=0}^{\infty} \frac{(\lambda + \theta y)^{y-2}e^{-(\lambda + \theta y)}}{y!} + \theta \sum_{y=0}^{\infty} \frac{[\lambda + \theta + \theta(y - 1)]^{y-1}e^{-[\lambda + \theta + \theta(y-1)]}}{(y - 1)!}
\]

so that

\begin{equation}
(A.2) \quad s_1 = \frac{1}{\lambda^2} - \frac{\theta}{\lambda(\lambda + \theta)},
\end{equation}

On the other hand,

\[
\begin{align*}
\lambda \sum_{y=0}^{\infty} & \frac{(\lambda + \theta y)^{y-2}e^{-(\lambda + \theta y)}}{y!} \\
+ \theta & \sum_{y=0}^{\infty} \frac{[\lambda + \theta + \theta(y - 1)]^{y-1}e^{-[\lambda + \theta + \theta(y-1)]}}{(y - 1)!} \\
= & \lambda \sum_{y=0}^{\infty} \frac{(\lambda + \theta y)^{y-3}e^{-(\lambda + \theta y)}}{y!} + \theta \sum_{y=0}^{\infty} \frac{[\lambda + \theta + \theta(y - 1)]^{y-2}e^{-[\lambda + \theta + \theta(y-1)]}}{(y - 1)!}
\end{align*}
\]

so that

\begin{equation}
(A.2) \quad s_1 = \frac{1}{\lambda^2} - \frac{\theta}{\lambda(\lambda + \theta)},
\end{equation}

so that

\begin{equation}
(A.3) \quad s_2 = \frac{1}{\lambda^2} - \frac{\theta}{\lambda^2(\lambda + \theta)} - \frac{\theta}{\lambda(\lambda + \theta)^2} + \frac{\theta^2}{\lambda(\lambda + \theta)(\lambda + 2\theta)}.
\end{equation}

Based on (A.3) and (A.2), we can obtain

\begin{align*}
E \left[ \frac{1}{(\lambda + \theta Y)^2} \right] \\
= & \frac{1}{\lambda^2} \sum_{y=0}^{\infty} \frac{\lambda(\lambda + \theta y)^{y-1}e^{-(\lambda + \theta y)}}{y!} \\
= & \left(1 - \phi \right) \sum_{y=0}^{\infty} \frac{\lambda(\lambda + \theta y)^{y-3}e^{-(\lambda + \theta y)}}{y!} \\
= & \left(1 - \phi \right) \sum_{y=1}^{\infty} \frac{y\lambda(\lambda + \theta y)^{y-3}e^{-(\lambda + \theta y)}}{y!} \\
= & \left(1 - \phi \right) \lambda \sum_{y=1}^{\infty} \frac{y[\lambda + \theta + \theta(y - 1)]^{y-2}e^{-[\lambda + \theta + \theta(y-1)]}}{(y - 1)!} \\
= & \left(1 - \phi \right) \lambda \sum_{y=0}^{\infty} \frac{[\lambda + \theta + \theta y)^{y-2}e^{-(\lambda + \theta y)}}{y!} \\
= & \left(1 - \phi \right) \lambda \left[ \frac{1}{(\lambda + \theta)^2} - \frac{\theta}{(\lambda + \theta)(\lambda + 2\theta)} \right]
\end{align*}

(\textit{Type I multivariate ZIGP distribution})
(A.6)

By combining (A.5) with (A.6), we immediately obtain the third formula of (A.1). To obtain $E[Y^3/(\lambda + \theta Y)^2]$, we need to calculate

$$c_1 = \sum_{y=0}^{\infty} \frac{y^2(\lambda + \theta + \theta y)e^{-(\lambda + \theta + \theta y)}}{y!}$$

and

$$c_2 = \sum_{y=0}^{\infty} \frac{y(\lambda + \theta + \theta y)e^{-(\lambda + \theta + \theta y)}}{y!}$$

In fact,

$$c_1 = \sum_{y=0}^{\infty} \frac{y^2(\lambda + 2\theta + (y-1))e^{-(\lambda + 2\theta + (y-1))}}{(y-1)!}$$

$$= \sum_{y=0}^{\infty} \frac{y(\lambda + 2\theta + \theta y)e^{-(\lambda + 2\theta + \theta y)}}{y!}$$

$$+ \sum_{y=0}^{\infty} \frac{y(\lambda + 2\theta + \theta y)e^{-(\lambda + 2\theta + \theta y)}}{y!} + \frac{1}{1 - \theta} + \frac{1}{\lambda + 2\theta}.$$

Thus, we have

$$E\left[\frac{Y^3}{(\lambda + \theta Y)^2}\right] = \frac{\sum_{y=0}^{\infty} \frac{y^3(\lambda + \theta y)e^{-(\lambda + \theta y)}}{y!}}{E[Y^3/(\lambda + \theta Y)^2]}$$

$$= (1 - \phi)\lambda \sum_{y=0}^{\infty} \frac{y^3(\lambda + \theta y)e^{-(\lambda + \theta y)}}{y!}$$

By combining (A.6) with the above formula, we immediately obtain the second formula of (A.1).

**Appendix B: The derivation of (3.9)**

From (2.16), since $X_i/y = y \neq 0$ for $i = 1, \ldots, m$, we obtain that when $y = y \neq 0$. 

308 X.-F. Huang et al.
From (2.15), we know that when \( y = y = 0, X_i | Y \sim ZIGP(\psi; \lambda, \theta). \) To obtain (3.9), we assume that \( X \sim ZIGP(\psi, \lambda, \theta) \) and only need to prove

\[ E \left( \frac{X^2 - X}{\lambda + \theta X} \right) = \frac{(1 - \psi) \lambda}{1 - \theta}. \]

Note that

\[ E \left( \frac{X}{\lambda + \theta X} \right) = \frac{(1 - \psi) \lambda}{1 - \theta}. \]

and

\[ E \left( \frac{X^2}{\lambda + \theta X} \right) = \frac{(1 - \psi) \lambda}{1 - \theta}. \]

By combining the two formulae, we obtain (B.1).

**Appendix C: The score vector and Fisher information matrix in Section 3.5.2**

The elements in the score vector \( U(\gamma, \lambda, \theta) \) and observed information matrix \( I(\gamma, \lambda, \theta | Y_{\text{obs}}) \) are

\[ \frac{\partial U^*}{\partial \gamma} = -n + \frac{n_0}{1 + \gamma + e^{-\lambda+\theta}}, \]

\[ \frac{\partial U^*}{\partial \lambda} = -n e^{-\lambda+\theta} + (n - n_0) \gamma \]

\[ \frac{\partial U^*}{\partial \theta} = \frac{n}{(1+\gamma)^2} - \frac{n_0}{(\gamma + e^{-\lambda+\theta})^2} \]

\[ \frac{\partial^2 U^*}{\partial \gamma^2} = -\frac{n}{(1+\gamma)^2} + \frac{n_0}{(\gamma + e^{-\lambda+\theta})^2}, \]

\[ \frac{\partial^2 U^*}{\partial \lambda^2} = -\frac{n}{(1+\gamma)^2} + \frac{n_0}{(\gamma + e^{-\lambda+\theta})^2}, \]

\[ \frac{\partial^2 U^*}{\partial \lambda \partial \gamma} = \frac{n_0 e^{-\lambda+\theta}}{(\gamma + e^{-\lambda+\theta})^2}, \]

\[ \frac{\partial^2 U^*}{\partial \lambda \partial \theta} = \frac{n_0 e^{-\lambda+\theta}}{(\gamma + e^{-\lambda+\theta})^2}, \]

\[ \frac{\partial^2 U^*}{\partial \gamma} \frac{\partial U^*}{\partial \theta} = \frac{n_0 e^{-\lambda+\theta}}{(\gamma + e^{-\lambda+\theta})^2}, \]

for \( i, k = 1, \ldots, m \) and \( i \neq k. \) By using (3.2), we can calculate the Fisher information matrix \( J(\gamma, \lambda, \theta), \) whose ele-

**Type I multivariate ZIGP distribution**
ments are given by

\[ J_{ii} = -E \left( \frac{\partial^2 \ell^*}{\partial \gamma_i^2} \right) = -n \left( \frac{1}{1 + \gamma} \right)^2 + \frac{n}{(1 + \gamma)(\gamma + e^{-\lambda_i})}, \]

\[ J_{ij + 1, i + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \lambda_i \partial \lambda_j} \right) = -\frac{n \gamma e^{-\lambda_i}}{(1 + \gamma)(\gamma + e^{-\lambda_i})} + \frac{n}{(1 + \gamma)(\lambda_i + \theta_i)}, \]

\[ J_{i + m + 1, i + m + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \theta_i^2} \right) = \frac{n \lambda_i}{(1 - \theta_i)(1 + \gamma)} + \frac{2n \lambda_i}{(\lambda_i + 2 \theta_i)(1 + \gamma)}, \]

\[ J_{1, i + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \gamma \partial \lambda_i} \right) = -\frac{n e^{-\lambda_i}}{(1 + \gamma)(\gamma + e^{-\lambda_i})}, \]

\[ J_{1, i + m + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \gamma \partial \theta_i} \right) = 0, \]

\[ J_{i + 1, k + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \lambda_i \partial \lambda_k} \right) = -\frac{n \gamma e^{-\lambda_i}}{(1 + \gamma)(\gamma + e^{-\lambda_i})}, \]

\[ J_{i + 1, k + m + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \theta_i \partial \theta_k} \right) = 0, \]

\[ J_{i + 1, i + m + 1} = -E \left( \frac{\partial^2 \ell^*}{\partial \theta_i \partial \theta_i} \right) = \frac{n \lambda_i}{(\lambda_i + 2 \theta_i)(1 + \gamma)}, \]

for \( i, k = 1, \ldots, m \) and \( i \neq k \).

Received 21 May 2015

REFERENCES

Chi Zhang  
Department of Statistics and Actuarial Science  
The University of Hong Kong, Pokfulam Road  
Hong Kong  
P. R. China  
E-mail address: zangchi@hku.hk

Xuejun Jiang  
Department of Mathematics  
South University of Science and Technology of China  
Shenzhen, Guangdong  
P. R. China  
E-mail address: jiangxj@sustc.edu.cn