Afrika Statistika

Afrika Statistika
Vol. 16 (2), 2021, pages 2767 - 2788.
DOI: http://dx.doi.org/10.16929/as/2021.2767.183

ISSN 2316-090X

Estimation in the zero-inflated bivariate Poisson model with an application to health-care utilization data

Konan Jean Geoffroy Kouakou [1,*], **Ouagnina Hili** [1] **and Jean-François Dupuy** [2]

[1] UMRI-Mathématiques et Nouvelles Technologies de l'Information
Institut National Polytechnique Félix Houphouët-Boigny (INP-HB) de Yamoussoukro
[2] Institut de recherche mathématique de Rennes (IRMAR), INSA de Rennes

Received on August 20, 2021; Accepted on September 09, 2021

Abstract. Data on the demand for medical care is usually measured by a number of different counts. These count data are most often correlated and subject to high proportions of zeros. However, excess zeros and the dependence between these data can jointly affect several utilization measures. In this paper, the zero-inflated bivariate Poisson regression model (ZIBP) was used to analyze health-care utilization data. First, the asymptotic properties of the maximum likelihood estimator (MLE) of this model were investigated theoretically. Then, a simulation study is conducted to evaluate the behaviour of the estimator in finite samples. Finally, an application of the ZIBP model to health care demand data is provided as an illustration.

Key words: Count data; excess zeros; demand for medical care; correlated data; simulations.
AMS 2010 Mathematics Subject Classification Objects : 62P10; 62F12; 62J12.

[*] Corresponding author: Konan Jean Geoffroy Kouakou (jeanosko@gmail.com)
Ouagnina Hili : o_hili@yahoo.fr
Jean-François Dupuy : Jean-Francois.Dupuy@insa-rennes.fr

Résumé. (Abstract in French) Les données sur la demande de soins médicaux sont généralement mesurées au moyen d'un certain nombre de comptes différents. Ces données de comptage sont le plus souvent corrélées et sujettes à de fortes proportions de zéros. Cependant, l'excès de zéros et la dépendance entre ces données peuvent affecter conjointement plusieurs de ces mesures d'utilisation. Dans cet article, le modèle de régression de Poisson bivarié à inflation de zéros (ZIBP) est utilisé pour analyser les données d'utilisation des soins de santé. Tout d'abord, les propriétés asymptotiques de l'estimateur du maximum de vraisemblance (EMV) de ce modèle ont été étudiées sur le plan théorique. Ensuite, une étude de simulation est réalisée pour évaluer le comportement de l'estimateur dans des échantillons finis. Enfin, une application du modèle ZIBP à des données de demandes de soins de santé est fournie à titre d'illustration.

Konan Jean Geoffroy Kouakou, M.Sc,. is preparing a Ph.D. thesis under the supervision of the second author and the collaboration of the third author, UMRI-Mathématiques et Nouvelles Technologies de l'Information, Institut National Polytechnique Félix Houphouët-Boigny (INP-HB) de Yamoussoukro, Côte d'Ivoire.

Ouagnina Hili, Ph.D., is a Full Professor of Statistics, UMRI-Mathématiques et Nouvelles Technologies de l'Information, Institut National Polytechnique Félix Houphouët-Boigny (INP-HB) de Yamoussoukro, Côte d'Ivoire.

Jean-François Dupuy, Ph.D., is a Full Professor of Statistics, Institut de recherche mathématique de Rennes (IRMAR), INSA de Rennes.

1. Introduction

Bivariate count models are used in situations where two dependent count variables are correlated and need to be jointly modeled. Bivariate count data are observed in many areas including marketing (number of purchases of different products), medical research (the number of seizures before and after treatment), epidemiology (incidents of different diseases in a series of districts), accident analysis (number of accidents in a site before and after infrastructure changes), econometrics (number of voluntary and involuntary job changes), sports (the number of goals scored by each one of the two opponent teams in soccer), just to name a few. In most cases, bivariate count data are modeled by bivariate Poisson models.

However, in many applications, the count data contain an excess of zeros, that is, a number of zeros that cannot be explained by standard models. A large number of statistical tools have been developed to solve this problem, such as zero-inflation regression models. These models account for excess zeros in count data by mixing a degenerate distribution with point mass of one at zero with a standard count regression model (Poisson, binomial or negative binomial when the response variable is univariate and bivariate Poisson, bivariate negative binomial when the response variables are bivariate, etc.). Several works have been carried out on zero-inflated univariate regression models, such as Lambert (1992), Dietz and Böhning (2000), Li (2011), Lim *et al.* (2014) and Monod (2014) for the ZIP (Zero Inflated Poisson) regression model, Ridout *et al.* (2001), Moghimbeigi *et al.* (2008), Mwalili *et al.* (2008), Garay *et al.* (2011) for ZINB (zero-inflated negative binomial) model, and Hall (2000), Hall and Berenhaut (2002), Diallo *et al.* (2017), Diallo *et al.* (2019) for the ZIB (zero-inflated binomial) regression model. But the ZIP, ZIB and ZINB models are not adapted to bivariate responses. Thus, several models have been proposed for bivariate count data with zero-inflation. For example, for zero-inflated bivariate negative binomial models, see Wang *et al.* (2003) and Faroughi and Ismail (2016), and for bivariate Poisson models, Li *et al.* (1999), Karlis and Ntzoufras (2003), Al Muhayfith *et al.* (2016), Yang *et al.* (2016), among others. In this paper, we consider the bivariate Poisson regression model with zero-inflation, which allows to model the correlation between the response variables and to handle a large number of observations $(0, 0)$ in the data set. Since its introduction by Li *et al.* (1999), the ZIBP (zero-inflated bivariate Poisson) model has been applied in a variety of contexts including marketing, epidemiology, accident analysis, medical research, sports, econometrics, etc. Hence, the consider estimation in ZIBP model. This work is also motivated by data from health economics. In health econometrics, health service utilization data are most often examined. Deb and Trivedi (2005) analyzed health service utilization data for individuals over age 65. These data from the National Medical Expenditure Survey (NMES) conducted in 1987 and 1988 and known as the NMES1988. These data contain measures of health care utilization such as the number of visits to a non-doctor health care professional (such as optician, physiotherapist, ...) in a office setting, the number of visits to a non-doctor in an outpatient setting. The proportions of zeros are high in these measures of health service use. This means that during the study period, these corresponding health services were not used by a large number of people. In addition, according to Gurmu and Elder (2000) and Wang *et al.* (2003), the health-care utilization measures are dependent. Therefore, a univariate analysis of these data would not be appropriate. To handle this problem, Diallo et al. (2018) proposed the ZIM (zero-inflated multinomial) regression model and applied it to the NMES1988 data. However, the ZIM regression model is restrictive, it is only suitable for bounded components.Thus, in this work, we are interested in the ZIBP model, which takes into account all the individuals in the population and takes into account the correlation between the health care demand data studied. First, we are interested in asymptotic properties in the ZIBP model. Second, an application of the ZIBP model allowed an assessment of the demand for medical care and the study of health care renunciation.

The outline of this paper is as follows. In Section 2, we present the ZIBP model and describe the maximum likelihood estimator. In Section 3, we first give some useful notations, then we indicate some regularity conditions and finally we establish the consistency and asymptotic normality of the MLE in the ZIBP regression. Section 4 presents the results of a simulation study. Section 5 describes an application of

ZIBP model to the analysis of health-care utilization by elderlies in United States. To conclude, a discussion and some perspectives are provided in Section 6.

2. The zero-inflated bivariate Poisson regression model

Consider random variables Z_1, Z_2 and U which follow independent Poisson distributions with parameters λ_1, λ_2 and μ respectively. Then the random variables $Y_1 = Z_1 + U$ and $Y_2 = Z_2 + U$ follow jointly a bivariate Poisson distribution BP$(\lambda_1, \lambda_2, \mu)$. Let $y_1 \wedge y_2 := \min(y_1, y_2)$. The joint distribution of the bivariate Poisson vector (Y_1, Y_2) is given by

$$\begin{aligned}
\mathbb{P}(Y_1 = y_1, Y_2 = y_2) &= \mathbb{P}(Z_1 + U = y_1, Z_2 + U = y_2) \\
&= \sum_{s=0}^{y_1 \wedge y_2} \mathbb{P}(U = s, Z_1 = y_1 - s, Z_2 = y_2 - s) \\
&= \sum_{s=0}^{y_1 \wedge y_2} \mathbb{P}(U = s)\mathbb{P}(Z_1 = y_1 - s)P(Z_2 = y_2 - s) \\
&= e^{-\mu-\lambda_1-\lambda_2}\varphi(y_1, y_2)
\end{aligned}$$

where

$$\varphi(y_1, y_2) = \sum_{s=0}^{y_1 \wedge y_2} \frac{\mu^s}{s!} \frac{\lambda_1^{y_1-s}}{(y_1-s)!} \frac{\lambda_2^{y_2-s}}{(y_2-s)!}.$$

Using the independence of Z_1, Z_2 and U, one has $\text{cov}(Y_1, Y_2) = \text{cov}(Z_1 + U, Z_2 + U) = \text{var}(U) = \mu$. Hence μ is a measure of dependence between Y_1 and Y_2. When μ =0, the bivariate Poisson distribution reduces to the product of two independent Poisson distributions (referred to as the double-Poisson distribution).

It is well known that univariate Poisson distributions are not appropriate for modeling counts with an excess of zeros. In such cases, zero-inflated regression models are most often suggested, see Lambert (1992). The same applies to the bivariate case when the data contain a high proportion of bivariate couples $(0, 0)$, see Li *et al.* (1999), Wang *et al.* (2003). The zero-inflated bivariate Poisson model was introduced by Li *et al.* (1999). Since then, this model was used by Wang *et al.* (2003) to analyze two types of occupational injuries, by Bermúdez (2009) in the field of automobile insurance, and by Yang *et al.* (2016) to model bivariate data in health economics, among others. According to Li *et al.* (1999), a ZIBP model is a mixture of a bivariate Poisson distribution and a point mass in $(0, 0)$. Thus, the ZIBP model is specified by the probability function:

$$f_{ZIBP}(y_1, y_2; \pi, \lambda_1, \lambda_2, \mu) = \begin{cases} \pi + (1-\pi)\exp(-\mu - \lambda_1 - \lambda_2), & (y_1, y_2) = (0, 0) \\ (1-\pi)\exp(-\mu - \lambda_1 - \lambda_2)\varphi(y_1, y_2), & (y_1, y_2) \neq (0, 0), \end{cases} \tag{1}$$

where $0 < \pi < 1$. When covariates are present, the model (1) can be extended to a regression model. For this purpose, for each $i = 1, ..., n$, we consider respectively $\mathbf{W}_i = (W_{i1}, ..., W_{iq})^\top$ and $\mathbf{X}_i = (X_{i1}, ..., X_{ip})^\top$ random vectors of covariates where $W_{i1} = X_{i1} = 1$. The mixing probability π_i is usually modeled by a logistic regression, that is:

$$\text{logit}(\pi_i) = \gamma^\top \mathbf{W}_i \tag{2}$$

and the Poisson parameters λ_{1i}, λ_{2i} and μ_i are modeled as:

$$\log(\lambda_{1i}) = \beta_1^\top \mathbf{X}_i, \;\; \log(\lambda_{2i}) = \beta_2^\top \mathbf{X}_i, \;\; \text{and} \;\; \log(\mu) = \eta \tag{3}$$

where $\gamma \in \mathbb{R}^q$, $\beta_1, \beta_2 \in \mathbb{R}^p$ and $\eta \in \mathbb{R}$ are unknown regression parameters and the symbol $\top$ denotes the transpose operator. One could also model μ as a function of the covariates, for example by $\log(\mu) = \beta_3^\top \mathbf{X}_i$. This generalisation is of no theoretical interest (it only makes the calculations more "painful"), we also think that in terms of interpretation, it is more relevant to have a "fixed" covariance.

Let $\theta = (\gamma^\top, \beta_1^\top, \beta_2^\top, \eta)^\top$ denote the vector k-dimensional ($k = 2p + q + 1$) parameters vector of the ZIBP model (1)-(2)-(3). The likelihood of θ, based on a sample of n independent observations $(Y_{1i}, Y_{2i}, \mathbf{X}_i, \mathbf{W}_i)$, $i = 1, \ldots, n$, is :

$$\begin{aligned}
L_n(\theta) &= \prod_{i=1}^{n} \left\{ \Big(\pi_i + (1-\pi_i) f_{BP}(0, 0, \lambda_{1i}, \lambda_{2i}, \mu)\Big)^{a_i} \times \Big((1-\pi_i) f_{BP}(Y_{1i}, Y_{2i}, \lambda_{1i}, \lambda_{2i}, \mu)\Big)^{1-a_i} \right\}, \\
&= \prod_{i=1}^{n} \left\{ \left(\frac{e^{\gamma^\top \mathbf{W}_i} + e^{-(e^\eta + e^{\beta_1^\top \mathbf{X}_i} + e^{\beta_2^\top \mathbf{X}_i})}}{1 + e^{\gamma^\top \mathbf{W}_i}} \right)^{a_i} \left(\frac{1}{1 + e^{\gamma^\top \mathbf{W}_i}} \right. \right. \\
&\qquad \left. \left. e^{-(e^\eta + e^{\beta_1^\top \mathbf{X}_i} + e^{\beta_2^\top \mathbf{X}_i})} \times \sum_{s=0}^{Y_{1i} \wedge Y_{2i}} \frac{(e^\eta)^s}{s!} \frac{(e^{\beta_1^\top \mathbf{X}_i})^{Y_{1i}-s}}{(Y_{1i}-s)!} \frac{(e^{\beta_2^\top \mathbf{X}_i})^{Y_{2i}-s}}{(Y_{2i}-s)!} \right)^{1-a_i} \right\},
\end{aligned}$$

where $a_i := 1_{(Y_{1i}=0, Y_{2i}=0)}$. Hence, the log-likelihood $\ell\ell_n(\theta) = \log\big(L_n(\theta)\big)$, is given by

$$\begin{aligned}
\ell\ell_n(\theta) &= \sum_{i=1}^{n} \left\{ a_i \log\big(e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)\big) - (1-a_i)\big(e^\eta + e^{\beta_1^\top \mathbf{X}_i} + e^{\beta_2^\top \mathbf{X}_i}\big) \right. \\
&\qquad \left. + (1-a_i) \log\left(\sum_{s=0}^{y_{1i} \wedge y_{2i}} \frac{(e^\eta)^s}{s!} \frac{(e^{\beta_1^\top \mathbf{X}_i})^{y_{1i}-s}}{(y_{1i}-s)!} \frac{(e^{\beta_2^\top \mathbf{X}_i})^{y_{2i}-s}}{(y_{2i}-s)!} \right) - \log(1 + e^{\gamma^\top \mathbf{W}_i}) \right\}, \\
\ell\ell_n(\theta) &:= \sum_{i=1}^{n} \ell_i(\theta),
\end{aligned}$$

where $h_i(\theta) = e^{-(e^\eta + e^{\beta_1^\top \mathbf{X}_i} + e^{\beta_2^\top \mathbf{X}_i})}$. The maximum likelihood estimator $\widehat{\theta}_n = \big(\widehat{\gamma}^\top, \widehat{\beta}_1^\top, \widehat{\beta}_2^\top, \widehat{\eta}\big)$ of θ is the solution of the k-dimensional score equation

$$U_n(\theta) = \frac{1}{\sqrt{n}} \frac{\partial \ell\ell_n(\theta)}{\partial \theta} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{\partial \ell_i(\theta)}{\partial \theta} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \dot{\ell}_i(\theta) = 0. \tag{4}$$

The components of the gradient vector are of the following form

$$\frac{\partial \ell_i(\theta)}{\partial \gamma_j} = \left(a_i \frac{e^{\gamma^\top \mathbf{W}_i}}{e^{\gamma_j^\top \mathbf{W}_i} + h_i(\theta)} - \frac{e^{\gamma_j^\top \mathbf{W}_i}}{1 + e^{\gamma_j^\top \mathbf{W}_i}} \right) W_{ij},$$

$$\frac{\partial \ell_i(\theta)}{\partial \beta_{1,\ell}} = \left(-a_i \frac{e^{\beta_{1,\ell}^\top \mathbf{X}_i} h_i(\theta)}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - (1-a_i) e^{\beta_{1,\ell}^\top \mathbf{X}_i} + (1-a_i) \frac{\sum_{s=0}^{Y_{1i} \wedge Y_{2i}} (Y_{1i} - s) g_i(s,\theta)}{\sum_{s=0}^{Y_{1i} \wedge Y_{2i}} g_i(s,\theta)} \right) X_{i\ell},$$

$$\frac{\partial \ell_i(\theta)}{\partial \beta_{2,\ell}} = \left(-a_i \frac{e^{\beta_{2,\ell}^\top \mathbf{X}_i} h_i(\theta)}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - (1-a_i) e^{\beta_{2,\ell}^\top \mathbf{X}_i} + (1-a_i) \frac{\sum_{s=0}^{Y_{1i} \wedge Y_{2i}} (Y_{2i} - s) g_i(s,\theta)}{\sum_{s=0}^{Y_{1i} \wedge Y_{2i}} g_i(s,\theta)} \right) X_{i\ell},$$

and

$$\frac{\partial \ell_i(\theta)}{\partial \eta} = -a_i \frac{e^{\eta} h_i(\theta)}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - (1-a_i) e^{\eta} + (1-a_i) \frac{\sum_{s=0}^{Y_{1i} \wedge Y_{2i}} s g_i(s,\theta)}{\sum_{s=0}^{Y_{1i} \wedge Y_{2i}} g_i(s,\theta)},$$

with

$$g_i(s,\theta) = \frac{(e^{\eta})^s}{s!} \times \frac{(e^{\beta_1^\top \mathbf{X}_i})^{Y_{1i}-s}}{(Y_{1i}-s)!} \times \frac{(e^{\beta_2^\top \mathbf{X}_i})^{Y_{2i}-s}}{(Y_{2i}-s)!},$$

for every $i = 1, \ldots, n$, $j = 1, \ldots, q$ and $\ell = 1, \ldots, p$. Furthermore, the estimation equation (4) can be solved by a Newton-Raphson algorithm.
In the next section, we establish the consistency and asymptotic normality of $\widehat{\theta}_n$.

3. Asymptotic properties of the MLE

In this section, we first give some additional notations that we use in the rest of the work. Then we state some regularity conditions. Finally, we present the asymptotic properties of the estimator $\widehat{\theta}_n$ of θ.

3.1. Notations and regularity assumptions

For every $i = 1, \ldots, n$, let

$$A_i(\theta) = a_i \frac{e^{\gamma^\top \mathbf{W}_i}}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - \frac{e^{\gamma^\top \mathbf{W}_i}}{1 + e^{\gamma^\top \mathbf{W}_i}},$$

$$B_{1,i}(\theta) = -a_i \frac{e^{\beta_1^\top \mathbf{X}_i} h_i(\theta)}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - (1 - a_i) e^{\beta_1^\top \mathbf{X}_i} + (1 - a_i) \frac{\sum\limits_{s=0}^{Y_{1i} \wedge Y_{2i}} (Y_{1i} - s) g_i(s, \theta)}{\sum\limits_{s=0}^{Y_{1i} \wedge Y_{2i}} g_i(s, \theta)},$$

$$B_{2,i}(\theta) = -a_i \frac{e^{\beta_2^\top \mathbf{X}_i} h_i(\theta)}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - (1 - a_i) e^{\beta_2^\top \mathbf{X}_i} + (1 - a_i) \frac{\sum\limits_{s=0}^{Y_{1i} \wedge Y_{2i}} (Y_{2i} - s) g_i(s, \theta)}{\sum\limits_{s=0}^{Y_{1i} \wedge Y_{2i}} g_i(s, \theta)},$$

and

$$C_i(\theta) = -a_i \frac{e^{\eta} h_i(\theta)}{e^{\gamma^\top \mathbf{W}_i} + h_i(\theta)} - (1 - a_i) e^{\eta} + (1 - a_i) \frac{\sum\limits_{s=0}^{Y_{1i} \wedge Y_{2i}} s g_i(s, \theta)}{\sum\limits_{s=0}^{Y_{1i} \wedge Y_{2i}} g_i(s, \theta)}.$$

Now, we state the regularity assumptions under which we will establish the asymptotic properties of the maximum likelihood estimator $\widehat{\theta}_n$.

$(A1)$ The true parameter value $\theta_0 := (\gamma_0^\top, \beta_{1,0}^\top, \beta_{2,0}^\top, \eta_0)^\top$ lies in the interior of some known compact set of $\Theta \subset \mathbb{R}^k$.

$(A2)$ $\mathbb{E}\left[\left(\dot{\ell}_i(\theta)\right)\left(\dot{\ell}_i(\theta)\right)^\top\right]$ is positive definite in a neighborhood of θ_0.

$(A3)$ In a neighborhood of θ_0, the first and second derivatives of $U_n(\theta)$ with respect to θ are uniformly bounded above by a function of $(Y_1, Y_2, \mathbf{X}, \mathbf{W})$, whose expectations exist.

$(A4)$ For every $i = 1, \ldots, n$, $\mathbb{E}\left[\frac{\partial^2 \ell_i(\theta)}{\partial\theta\partial\theta^\top}\right]$ is finite and is negative definite in a neighborhood of θ_0. In addition, $-\frac{1}{\sqrt{n}} \frac{\partial U_n(\theta)}{\partial \theta^\top}$ converges to a positive definite matrix $\Sigma(\theta)$ as n tends to infinity.

Assumptions $(A1)$ - $(A4)$ are classical in zero-inflated regression models (see Diallo *et al.* (2017), Diallo et al. (2018); Lukusa *et al.* (2016), Lee *et al.* (2020)).
In the following, the $\mathbb{R}^k$ space of k-dimensional column vectors will be provided with the Euclidean norm $\|\cdot\|_2$ and the space of $(k \times k)$ real matrices will be provided with the norm $|||A|||_2 := max_{\|x\|_2=1}\|Ax\|_2$ (for notations simplicity, we will use $\|\cdot\|$ for both norms).

We are now in position to state our results:

3.2. Asymptotic results for the MLE

Theorem 1 (Existence and consistency). *Under assumptions* $(A1)$ - $(A4)$, $\widehat{\theta}_n$ *converges in probability to* θ_0 *when* n *tends to infinity.*

The consistency of $\widehat{\theta}_n$, can be proved by checking that the conditions of Foutz's inverse function theorem (see Foutz (1977)) are satisfied.

First, we have $\ell\ell_n(\theta)$ is twice differentiable with respect to θ and its second derivatives are continuous. Thus $\frac{\partial^2 \ell\ell_n(\theta)}{\partial\theta\partial\theta^\top}$ exists and is continuous in an open neighborhood of θ_0.
Condition 1 is therefore checked. ■

Secondly, let us show that $\frac{1}{n}\dot{\ell\ell}_n(\theta_0) = \frac{1}{n}\frac{\partial \ell\ell_n(\theta_0)}{\partial\theta}$ converges in probability to 0 when n tends to infinity. To justify this, we note

$$\begin{aligned}\frac{1}{n}\dot{\ell\ell}_n(\theta_0) \;=\; & \Big(\frac{1}{n}\sum_{i=1}^{n} W_{i1}A_i(\theta_0), \ldots, \frac{1}{n}\sum_{i=1}^{n} W_{iq}A_i(\theta_0), \frac{1}{n}\sum_{i=1}^{n} X_{i1}B_{1,i}(\theta_0), \ldots, \frac{1}{n}\sum_{i=1}^{n} X_{ip}B_{1,i}(\theta_0), \\ & \frac{1}{n}\sum_{i=1}^{n} X_{i1}B_{2,i}(\theta_0), \ldots, \frac{1}{n}\sum_{i=1}^{n} X_{ip}B_{2,i}(\theta_0), \frac{1}{n}\sum_{i=1}^{n} C_i(\theta_0)\Big)^\top\end{aligned}$$

Since the score vector is centered, it follows that $\mathrm{var}\Big[W_{il}A_i(\theta_0)\Big] = \mathbb{E}\Big[W_{il}^2 A_i^2(\theta_0)\Big]$.
In addition, for every $i = 1, \ldots, n$, $\mathbb{E}\big[-\frac{\partial^2 \ell_i(\theta_0)}{\partial\theta\partial\theta^\top}\big] = \mathbb{E}\big[\big(\dot{\ell}_i(\theta_0)(\dot{\ell}_i(\theta_0)\big)^\top\big]$.
Thus, by $(A2)$, it follows that $\mathrm{var}\big(W_{i\ell}A_i(\theta_0)\big) < \infty$.
Therefore, by the weak law of large numbers, we have $\frac{1}{n}\sum\limits_{i=1}^{n} W_{i\ell}A_i(\theta_0)$ converges in probability to 0 as $n \to \infty$, for every $\ell = 1, \ldots, q$.

By similary arguments, we show that for every $j = 1, ..., p$, $t \in \{1, 2\}$, $\frac{1}{n}\sum\limits_{i=1}^{n} X_{ij}B_{t,i}(\theta_0)$ and $\frac{1}{n}\sum\limits_{i=1}^{n} C_i(\theta_0)$ converge in probability to 0, when n tends to infinity.
Finally, we can conclude that $\frac{1}{n}\dot{\ell\ell}_n(\theta_0)$ converges in probability to $0_{(k,1)}$, when n tends to infinity.
Condition 2 is therefore checked. ■

Thirdly, let us show that $-\frac{1}{n}\frac{\partial^2 \ell\ell_n(\theta)}{\partial\theta\partial\theta^\top}$ converges uniformly in probability to the function $\Sigma(\theta)$ in an open neighbourhood of θ_0.

To see this, let $\mathcal{V}_{\theta_0}$ be an open neighbourhood of θ_0 and let $\theta \in \mathcal{V}_{\theta_0}$. Let $H_{i,(j,\ell)}(Y_1, Y_2, \mathbf{X}_i, \mathbf{W}_i, \theta) = -\frac{\partial^2 \ell_i(\theta)}{\partial\theta_j \partial\theta_\ell}$ for $j, \ell \in \{1, \ldots 2p + q + 1\}$.
By $(A3)$, it exists a function $N_i(.)$ such that for $\theta, \widetilde{\theta} \in \mathcal{V}_{\theta_0}$, $i \in \{1, \ldots, n\}$, we have

$$\left|H_{i,(j,\ell)}(Y_{1i}, Y_{2i}, \mathbf{X}_i, \mathbf{W}_i, \theta) - H_{i,(j,\ell)}(Y_{1i}, Y_{2i}, \mathbf{X}_i, \mathbf{W}_i, \widetilde{\theta})\right| \leq N_i(Y_{1i}, Y_{2i}, \mathbf{X}_i, \mathbf{W}_i)\|\theta - \widetilde{\theta}\|,$$

also we have $\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}[N_i(Y_{1i}, Y_{2i}, \mathbf{X}_i, \mathbf{W}_i)] = O(1)$. Furthermore, by assumption (A4), $-\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ell_i(\theta)}{\partial\theta_j \partial\theta_\ell}$ converges in probability to $\Sigma_{(j,\ell)}(\theta)$ as n tends to infinity, where $\Sigma_{(j,\ell)}(\theta)$ denotes the (j, ℓ)-th element of $\Sigma(\theta)$, for $j, \ell \in \{1, \ldots, 2p + q + 1\}$. Hence, using corrolary 3.1 of Newey (1991) under the assumptions $(A1)$ - $(A4)$, it follows that $-\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ell_i(\theta)}{\partial\theta \partial\theta^\top}$ converges uniformly in probability to a positive definite matrix $\Sigma(\theta)$ on $\mathcal{V}_{\theta_0}$.

Condition 3 is therefore checked. ■

The three conditions of Foutz (1977) inverse function theorem are verified. Thus, we conclude that $\widehat{\theta}_n$ converges in probability to θ_0.

Theorem 2 (Asymptotic normality). *Under Assumptions* $(A1)$ - $(A4)$, $\Sigma(\widehat{\theta}_n)^{1/2}\sqrt{n}(\widehat{\theta}_n - \theta_0)$ *converges in distribution to the Gaussian vector* $\mathcal{N}(0, I_k)$, *as* $n \longrightarrow \infty$, *where* I_k *is the* k*-dimensional identity matrix.*

Proof of Theorem 2. The proof of Theorem 2 is classical, we omit it.

4. Simulation study

In this section, we assess finite-sample properties of the maximum likelihood estimator $\widehat{\theta}_n$.

4.1. Study design

We generate data from the following ZIBP regression model :

$$\begin{cases} \text{logit}(\pi_i) = \gamma^\top \mathbf{W}_i \\ \log(\lambda_{1i}) = \beta_1^\top \mathbf{X}_i, \quad \log(\lambda_{2i}) = \beta_2^\top \mathbf{X}_i, \quad \text{and} \quad \log(\mu) = \eta, \end{cases}$$

with $\mathbf{X}_i = (1, X_{i2}, X_{i3}, X_{i4}, X_{i5}, X_{i6})^\top$ and $\mathbf{W}_i = (1, W_{i2}, W_{i3}, W_{i4}, W_{i5})$ where

- $X_{i1} = 1$ and $X_{i2}, \ldots, X_{i6}$ are independently drawn from normal $\mathcal{N}(0, 0.1)$, uniform $\mathcal{U}(-1, 1)$, exponential $\mathcal{E}(1)$, binomial $\mathcal{B}(1, 0.8)$ and binomial $\mathcal{B}(1, 0.4)$ distributions, respectively,

- $W_{i1} = 1$ and W_{i3}, W_{i4}, W_{i5} are independently drawn from $\mathcal{B}(1, 0.3)$, normal $\mathcal{N}(-1, 1)$ and binomial $\mathcal{N}(1, 0.5)$ distributions, respectively,
- by letting $W_{i2} = X_{i2}$.
- The regression parameters β_1, β_2 and η are chosen as follows: $\beta_1 = (-0.3, 0.85, 0.1, 0.25, -0.1, -0.05)^\top$, $\beta_2 = (0.8, -0.74, -0.1, -0.1, 0.15, -0.1)^\top$, $\eta = 0.4$
- We consider two cases for the regression parameter γ :
  Case 1 $\gamma = (-0.55, -0.75, -1, 0.45, 0)^\top$ for 25% zero-inflation
  Case 2 $\gamma = (-0.25, -0.4, 0.8, 0.45, 0)^\top$ for 50% zero-inflation

Using these values, in the case 1 (respectively, case 2) the average proportions of zero-inflation in the simulated data sets is 25% (respectively, 50%).

For each combination of the simulation design parameters (`sample size` and `proportion of zero-inflation`), we simulate $N = 1000$ samples and we calculate the maximum likelihood estimate $\widehat{\theta}_n$ of $\theta = (\gamma^\top, \beta_1^\top, \beta_2^\top, \eta)$. Several authors have developed EM-type estimation algorithms in zero-inflation models (for example, see Wang *et al.* (2003)). Other authors perform direct maximization using Newton-Raphson algorithms see Diallo *et al.* (2017), Diallo *et al.* (2019). In our simulation study, we use a Newton-Raphson algorithm implemented in the `R package maxLik` developed by Henningsen and Toomet (2011).

4.2. Results

For each combination `sample size` × `zero-inflation proportion` of the simulation parameters, we calculate the MLE $\widehat{\theta}_n$ and the average bias and average relative bias (expressed as a percentage) of the estimates $\widehat{\gamma}_{i,n}$, $\widehat{\beta}_{1,j,n}$, $\widehat{\beta}_{2,k,n}$ and $\widehat{\eta}_n$ over the N simulated samples. For example, the relative bias of $\widehat{\gamma}_{i,n}$ is obtained as $\frac{1}{N}\sum_{t=1}^{N} \frac{\widehat{\gamma}_{i,n}^{(t)} - \gamma_i}{\gamma_i} \times 100$ where $\hat{\gamma}_{i,n}^{(t)}$ denotes the MLE of γ_i in the t-th simulated sample. We also calculate the average standard error (SE), empirical standard deviation (SD) and root mean square error (RMSE) for each $\widehat{\gamma}_{i,n}$ $(i = 1, \ldots, 5)$, $\widehat{\beta}_{1,j,n}$, $\widehat{\beta}_{2,k,n}$ $(j, k = 1, \ldots, 6)$ and $\hat{\eta}_n$. SE is calculated as the average of the standard errors across the N simulated samples. For example, $\widehat{\gamma}_{i,n}$, SE is obtain as $\frac{1}{N}\sum_{t=1}^{N} s.e.\big(\widehat{\gamma}_{i,n}^{(t)}\big)$, while SD (respectively RMSE) is the square root of the empirical variance (respectively RMSE) of $\big(\widehat{\gamma}_{i,n}^{(1)}, \ldots, \widehat{\gamma}_{i,n}^{(N)}\big)$. Moreover, we provide the empirical coverage probability (CP) and average length of 95%-level confidence intervals for the γ_i, $\beta_{1,j}$, $\beta_{2,k}$ and η. Results are given in Table 1 (for the case 1) and Table 2 (for the case 2). In Table 1, we provide results for $n = 500$, $n = 2000$ and 25% of zero-inflation. Table 2 provides results for $n = 500$, $n = 2000$ and 50% of zero-inflation.

From the results obtained, we observe that the bias and relative bias are fairly small. Second, the bias, relative bias, SE, SD, and ℓ(CI) of all estimators decrease as the sample size increases. In addition, for γ_i, $\beta_{1,j}$, $\beta_{2,k}$ and η empirical coverage probabilities are close to the nominal confidence level in every case. On the other hand, we observe that the MLE of the $\beta_{1,j}$s, $\beta_{2,k}$s and η_ns (respectively, γ_is) performs better when the zero inflation proportion decreases (respectively, increases).

To assess the quality of the Gaussian approximation stated in Theorem 2, we provide normal Q-Q plots of the estimates and histograms of the normalized estimates $(\widehat{\gamma}_{i,n} - \gamma_i)/$s.e.$(\widehat{\gamma}_{i,n})$, $j = 1, \ldots, 5$, $(\widehat{\beta}_{1,j,n} - \beta_{1,j})/$s.e.$(\widehat{\beta}_{1,j,n})$, $j = 1, \ldots, 6$, $(\widehat{\beta}_{2,k,n} - \beta_{2,k})/$s.e.$(\widehat{\beta}_{2,k,n})$, $k = 1, \ldots, 6$ and $(\hat{\eta}_n - \eta_l)/$s.e.$(\hat{\eta}_n)$. We provide these graphs for $n = 2000$ and an average sample proportion of zero-inflation equal to 25% (**Fig.** 1 to 4 provide Q-Q plots for $\big(\widehat{\gamma}_{1,n}, \ldots, \widehat{\gamma}_{5,n}\big)$, $\big(\widehat{\beta}_{1,1,n}, \ldots, \widehat{\beta}_{1,6,n}\big)$, $\big(\hat{\beta}_{2,1,n}, \ldots, \widehat{\beta}_{2,6,n}\big)$, $\widehat{\eta}_n$, respectively; **Fig.** 5 to 8 provide histograms of the normalized $\big(\widehat{\gamma}_{1,n}, \ldots, \widehat{\gamma}_{5,n}\big)$, $\big(\widehat{\beta}_{1,1,n}, \ldots, \widehat{\beta}_{1,6,n}\big)$, $\big(\widehat{\beta}_{2,1,n}, \ldots, \widehat{\beta}_{2,6,n}\big)$, $\widehat{\eta}_n$, respectively). The plots of the other simulated scenarios are similar and are therefore not given. From these figures, it appears that the Gaussian approximation of the distribution of the MLE in the ZIBP is reasonably satisfied, even when the sample size is moderate and the proportion of zero-inflation is as high as 50%.

Table 1. Simulation results for number of replications $N = 1000$, sample size $n = 500$, and $n = 2000$ with 25% zero-inflation zero.

		$\widehat{\gamma}$					$\widehat{\beta_1}$						$\widehat{\beta_2}$						$\widehat{\eta}$
n		$\widehat{\gamma}_1$	$\widehat{\gamma}_2$	$\widehat{\gamma}_3$	$\widehat{\gamma}_4$	$\widehat{\gamma}_5$	$\widehat{\beta}_{1,1}$	$\widehat{\beta}_{1,2}$	$\widehat{\beta}_{1,3}$	$\widehat{\beta}_{1,4}$	$\widehat{\beta}_{1,5}$	$\widehat{\beta}_{1,6}$	$\widehat{\beta}_{2,1}$	$\widehat{\beta}_{2,2}$	$\widehat{\beta}_{2,3}$	$\widehat{\beta}_{2,4}$	$\widehat{\beta}_{2,5}$	$\widehat{\beta}_{2,6}$	$\widehat{\eta}$
500	bias	-0.0112	0.0232	-0.0362	0.0086	0.0098	-0.0602	-0.0149	0.0010	-0.0040	0.0266	0.0094	-0.0170	-0.0340	-0.0023	-0.0067	0.0113	0.0002	0.0052
	rel. bias	2.0278	3.0954	3.6227	1.9176	-	20.0556	1.7503	1.0411	1.5938	26.6454	18.7259	2.1222	4.5889	2.2930	6.7037	7.5364	0.2397	1.3125
	SD	0.1972	1.1554	0.3052	0.1184	0.2318	0.2426	0.8371	0.1498	0.0625	0.2016	0.1721	0.1275	0.4329	0.0731	0.0493	0.1179	0.0917	0.0786
	SE	0.1951	1.2121	0.2987	0.1190	0.2278	0.2380	0.8632	0.1445	0.0618	0.2042	0.1707	0.1243	0.4410	0.0748	0.0498	0.1161	0.0892	0.0765
	RMSE	0.2775	1.6743	0.4285	0.1681	0.3251	0.3451	1.2023	0.2080	0.0879	0.2881	0.2426	0.1788	0.6188	0.1046	0.0704	0.1658	0.1279	0.1098
	CP	0.9560	0.9530	0.9440	0.9560	0.9450	0.9460	0.9610	0.9490	0.9520	0.9580	0.9520	0.9430	0.9480	0.9500	0.9590	0.95300	0.9350	0.9500
	ℓ(CI)	0.7639	4.6888	1.1650	0.4658	0.8925	0.9237	3.3416	0.5632	0.2378	0.7926	0.6655	0.4854	1.7225	0.2928	0.1944	0.4535	0.3493	0.2977
2000	bias	-0.0043	0.0182	-0.0064	0.0026	0.0032	-0.0095	-0.0185	-0.0001	-0.0007	0.0043	-0.0049	0.0013	-0.0105	-0.0010	-0.0025	-0.0018	-0.0019	0.0020
	rel. bias	0.7788	-2.4311	0.6356	0.5684	-	3.1738	-2.1723	-0.0677	-0.2777	-4.2934	9.8046	0.1637	1.4257	0.9654	2.4769	-1.2300	1.8964	0.5013
	SD	0.0942	0.5721	0.1456	0.0571	0.1094	0.1152	0.3886	0.0698	0.0279	0.0977	0.0849	0.0605	0.2169	0.0361	0.0241	0.0560	0.0437	0.0377
	SE	0.0962	0.5929	0.1448	0.0585	0.1124	0.1121	0.4126	0.0696	0.0282	0.0962	0.0822	0.0607	0.2163	0.0369	0.0242	0.0566	0.0440	0.0379
	RMSE	0.1347	0.8239	0.2054	0.0818	0.1568	0.1610	0.5669	0.0985	0.0396	0.1371	0.1183	0.0856	0.3064	0.0516	0.0342	0.0796	0.0621	0.0535
	CP	0.9470	0.9530	0.9510	0.9640	0.9560	0.9320	0.9560	0.9530	0.9520	0.9440	0.9400	0.9440	0.9500	0.9550	0.9480	0.9530	0.9530	0.9530
	ℓ(CI)	0.3770	2.3005	0.5668	0.2293	0.4406	0.4383	1.6088	0.2725	0.1098	0.3762	0.3220	0.2376	0.8466	0.1447	0.0947	0.2218	0.1725	0.1485

SD: empirical standard deviation. SE: average standard error. RMSE: empirical root mean square error. CP: empirical coverage probability of 95% level confidence intervals. ℓ(CI) : average length of the confidence intervals.

Table 2. Simulation results for number of replications $N = 1000$, sample size $n = 500$, and $n = 2000$ with 50% zero-inflation zero.

		$\widehat{\gamma}$					$\widehat{\beta_1}$						$\widehat{\beta_2}$						$\widehat{\eta}$
n		$\widehat{\gamma}_1$	$\widehat{\gamma}_2$	$\widehat{\gamma}_3$	$\widehat{\gamma}_4$	$\widehat{\gamma}_5$	$\hat{\beta}_{1,1}$	$\widehat{\beta}_{1,2}$	$\widehat{\beta}_{1,3}$	$\widehat{\beta}_{1,4}$	$\widehat{\beta}_{1,5}$	$\widehat{\beta}_{1,6}$	$\widehat{\beta}_{2,1}$	$\widehat{\beta}_{2,2}$	$\widehat{\beta}_{2,3}$	$\widehat{\beta}_{2,4}$	$\widehat{\beta}_{2,5}$	$\widehat{\beta}_{2,6}$	$\widehat{\eta}$
	bias	0.0041	-0.0002	0.0132	0.0078	0.0019	-0.0759	0.0825	0.0071	-0.0052	0.0193	-0.0043	-0.0210	0.0077	-0.0070	-0.0114	0.0134	-0.0005	0.0076
	rel. bias	1.6350	0.0431	1.6454	1.7245	-	25.3001	9.7067	7.0622	-2.0873	-19.2686	8.5766	-2.6278	-1.0451	6.9572	11.4178	8.9470	0.5032	1.8895
	SD	0.1827	1.0000	0.2103	0.1026	0.1852	0.3116	1.0775	0.1819	0.0800	0.2651	0.2292	0.1642	0.5596	0.0942	0.0645	0.1461	0.1114	0.0995
500	SE	0.1757	0.9939	0.2115	0.1000	0.1903	0.3075	1.1625	0.1864	0.0823	0.2633	0.2210	0.1577	0.5655	0.0945	0.0636	0.1473	0.1127	0.0962
	RMSE	0.2535	1.4096	0.2985	0.1435	0.2655	0.4441	1.5868	0.2605	0.1148	0.3740	0.3183	0.2286	0.7954	0.1336	0.0913	0.2079	0.1585	0.1386
	CP	0.9420	0.9480	0.9560	0.9380	0.9620	0.9460	0.9470	0.9670	0.9590	0.9500	0.9440	0.9430	0.9410	0.9470	0.9570	0.9570	0.9570	0.9330
	ℓ(CI)	0.6885	3.8698	0.8286	0.3918	0.7458	1.1868	4.3976	0.7240	0.3137	1.0169	0.8583	0.6146	2.1952	0.3696	0.2473	0.5742	0.4411	0.3723
	bias	-0.0024	0.0174	0.0028	0.0003	-0.0012	-0.0192	0.0017	-0.0038	0.0003	0.0040	-0.0024	-0.0039	-0.0054	-0.0002	-0.0026	0.0021	0.0001	0.0031
	rel. bias	-0.9420	-4.3542	0.3477	0.0727	-	6.3964	0.1977	-3.7725	0.1000	-4.0352	4.7782	-0.4815	0.7335	0.1866	2.6356	1.3974	-0.0507	0.7705
	SD	0.0838	0.4657	0.0992	0.0505	0.0937	0.1471	0.4906	0.0921	0.0362	0.1242	0.1041	0.0757	0.2692	0.0472	0.0315	0.0714	0.0550	0.0472
2000	SE	0.0870	0.4904	0.1046	0.0495	0.0943	0.1417	0.5365	0.0878	0.0360	0.1218	0.1038	0.0761	0.2751	0.0463	0.0304	0.0712	0.0552	0.0474
	RMSE	0.1208	0.6764	0.1441	0.0707	0.1329	0.2051	0.7269	0.1273	0.0510	0.1739	0.1470	0.1074	0.3849	0.0661	0.0438	0.1008	0.0779	0.0670
	CP	0.9630	0.9600	0.9490	0.9470	0.9500	0.9390	0.9610	0.9320	0.9440	0.9480	0.9510	0.9570	0.9560	0.9440	0.9390	0.9440	0.9530	0.9530
	ℓ(CI)	0.3410	1.9134	0.4098	0.1939	0.3696	0.5534	2.0763	0.3436	0.1398	0.4758	0.4060	0.2980	1.0741	0.1814	0.1191	0.2788	0.2162	0.1855

SD: empirical standard deviation. SE: average standard error. RMSE: empirical root mean square error. CP: empirical coverage probability of 95% level confidence intervals. ℓ(CI) : average length of the confidence intervals.

5. Application

5.1. Data description and modeling

In this section we describe an application of ZIBP regression model to the analysis of health-care utilization by elderlies in the United States. We use data from the National Medical Expenditure Survey (NMES) conducted in 1987-1988 in the United States. These health survey data contain a set of 4406 observations of individuals aged 66 and over. This data set has been reviewed by Deb and Trivedi Deb and Trivedi (2005); see also Diallo *et al.* (2017) and is available in the `R` package `AER` under the name "NMES1988". In these data, we consider jointly health-care utilization measures: the number `ofnd` of consultations with a non-doctor in an office setting and the number `opnd` of consultations with a non-doctor in an outpatient setting. The frequency of individuals with zero occurring simultaneously in (`ofnd` and `opnd`) is 59.03%. The tests carried out with the `cor.test` function of `package stat` of `R` on the variables `ofnd` and `opnd` show that they are correlated and cov(`ofnd`,`opnd`)=0.8615743. Thus we propose to use ZIBP model to investigate the determinants of health-care utilization in this data set. Some covariates were recorded on each individual. They include : (i) socio-economic variables : gender (1 for female, 0 for male, denoted `gender`), age (in years, divided by 10, denoted by `age`), marital status (1 if married, 0 otherwise, denoted by `status`), educational level (number of years of education, denoted by `school`), family income (in ten-thousands of dollars, denoted by `income`); (ii) various measures of health status: number of chronic conditions (cancer, diabete, arthritis, . . . denoted by `chronic`) and a variable indicating self-perceived health level (poor, average, excellent), which we re-code as "`health1`" (1 if health is perceived as poor, 0 otherwise) and "`health`2" (1 health is perceived as excellent, 0 otherwise); (iii) `medicaid`, a binary variable that indicates whether the individual is covered by medicaid or not. We code it as 1 if the person is covered and 0 otherwise. We fit a ZIBP regression model incorporating all available covariates in (2)-(3) for each individual. Then, we used Wald tests to select significant covariates. The least significant covariate "at the level 5%" is removed and the model is fitted again, until all remaining covariates are significant; the BIC criterion decreases at each step of this procedure. Table3 presents the final ZIBP model.

5.2. Results

Results for the resulting ZIBP model are displayed in Table 3. Estimate, standard error (s.e.) and significance level (as : not significant, significant or very significant) of Wald test of nullity for each parameter are reported in Table 3. The results show that, number of chronic conditions, gender, educational level and medicaid status are identified by ZIBP as the most influencing factors of the decision of never resorting to non-physician health professional (office or outpatient). It appears that the probability of never resorting to non-doctor consultations decreases with the number of chronic conditions. This is justified by the fact that the more chronic the patient's condition, the more likely the patient is to favour visits to the doctor. Then, the probability of never having recourse to consultations with a

Table 3. Health-care data analysis: estimates (standard errors) and significance codes: $\star\star\star$ significant at the 0.1% level, $\star\star$ significant at the 1% level, $\star$ significant at the 5% level.

parameter	variable	estimate	s.e.	Pr$(> t)$	signif.
$\hat{\gamma}_1$	`intercept`	0.667400	0.202808	0.000999	$\star\star\star$
$\hat{\gamma}_2$	`chronic`	-0.165918	0.023503	$1.67e^{-12}$	$\star\star\star$
$\hat{\gamma}_3$	`gender`	0.314814	0.064692	$1.14e^{-06}$	$\star\star\star$
$\hat{\gamma}_4$	`education`	-0.088953	0.008967	$< 2e^{-16}$	$\star\star\star$
$\hat{\gamma}_5$	`medicaid`	0.397012	0.117432	0.000723	$\star\star\star$
$\hat{\beta}_{1,1}$	`intercept`	1.381952	0.209195	$3.95e^{-11}$	$\star\star\star$
$\hat{\beta}_{1,2}$	`health1`	0.083504	0.041901	0.046273	$\star$
$\hat{\beta}_{1,3}$	`health2`	0.133144	0.046970	0.004588	$\star\star$
$\hat{\beta}_{1,4}$	`chronic`	0.025524	0.009689	0.008428	$\star\star$
$\hat{\beta}_{1,5}$	`age`	-0.124267	0.021390	6.26^{e-09}	$\star\star\star$
$\hat{\beta}_{1,6}$	`gender`	-0.007129	0.027981	0.798901	
$\hat{\beta}_{1,7}$	`marital statuts`	0.004958	0.028733	0.862999	
$\hat{\beta}_{1,8}$	`education`	0.032632	0.003950	$< 2^{e-16}$	$\star\star\star$
$\hat{\beta}_{1,9}$	`income`	-0.018566	0.004775	0.000101	$\star\star\star$
$\hat{\beta}_{1,10}$	`medicaid`	0.205903	0.051070	$5.54e^{-05}$	$\star\star\star$
$\hat{\beta}_{2,1}$	`intercept`	7.573439	0.377543	$< 2e^{-16}$	$\star\star\star$
$\hat{\beta}_{2,2}$	`health1`	-0.168968	0.063093	0.007404	$\star\star$
$\hat{\beta}_{2,3}$	`health2`	-0.788500	0.144316	$4.66e^{-08}$	$\star\star\star$
$\hat{\beta}_{2,4}$	`chronic`	0.112977	0.015864	$1.07e^{-12}$	$\star\star\star$
$\hat{\beta}_{2,5}$	`age`	-0.491438	0.040842	$< 2e^{-16}$	$\star\star\star$
$\hat{\beta}_{2,6}$	`gender`	0.214599	0.048058	7.99^{e-06}	$\star\star\star$
$\hat{\beta}_{2,7}$	`marital statuts`	-0.116671	0.050193	0.020101	$\star$
$\hat{\beta}_{2,8}$	`education`	-0.103404	0.006346	$< 2e^{-16}$	$\star\star\star$
$\hat{\beta}_{2,9}$	`income`	-0.024264	0.010076	0.016033	$\star$
$\hat{\beta}_{2,10}$	`medicaid`	-1.758261	0.061534	$< 2e - 16$	$\star\star\star$

non-physician decreases with the number of years of education. Indeed, education may make individuals more informed consumers of health care services. This result further confirms those of Deb and Trivedi (2005). Medicaid beneficiaries tend to forego consultations with a non-physician. Because Medicaid is health insurance for poor people, recipients are limited in their choice of consultations. They are limited to visits to the doctor only. It is also found that women are more likely to be non-users of `ofnd` and `opnd`.

Among patients who have not systematically given up consulting a non-physician health professional, the probability of having recourse to an `ofnp` or `opnp` consultation decreases with age and income. A patient's level of income will influence the nature and quality of the care he or she seeks, rather than the number of visits, which is consistent with Deb and Trivedi (2005). The probability of resorting to an `opnp` consultation decreases when patients feel that their health is no longer excellent, it has deteriorated. An elderly patient living away from cities may find

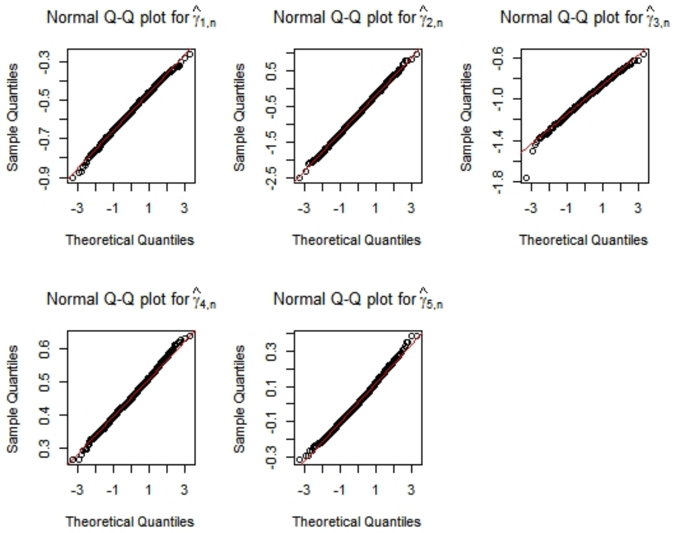

Fig. 1. Normal $Q-Q$ plots for $\widehat{\gamma}_{1,n},\ldots,\widehat{\gamma}_{5,n}$ with $n = 2000$ and 25% of zero-inflation.

it difficult to get to a consultation. Elderly patients most often have mobility problems. Married patients appear to renounce consulting a non-physician on an outpatient basis. Although better informed patients tend to diversify their use of health care, they seem to move away from the opnp health service to the ofnp service.

Therefore, by considering the variables ofnd and opnd simultaneously, while taking into account the correlation between them, allows the ZIBP model a better understanding of the elements that justify the use of different forms of medical care in order of use. We have found that patients with frail health, elderly, and covered by medicaid insurance prefer consultations with doctors than with non-doctors.

6. Conclusion

In this work, we theoretically and numerically evaluated the performance of the maximum likelihood estimator in the ZIBP model. An application of the ZIBP model to NMES data provided insight into the factors that promote the renunciation or use of some health care services. Now, several extensions of the ZIBP model should be developed to extend its scope. For example, we can first consider studying the properties of the MLE in the ZIBP models with randomly censored values on the right or left or by intervals. Second, consider inference in multivariate Poisson regression models with zero inflation. These topics could form the basis of our future work.

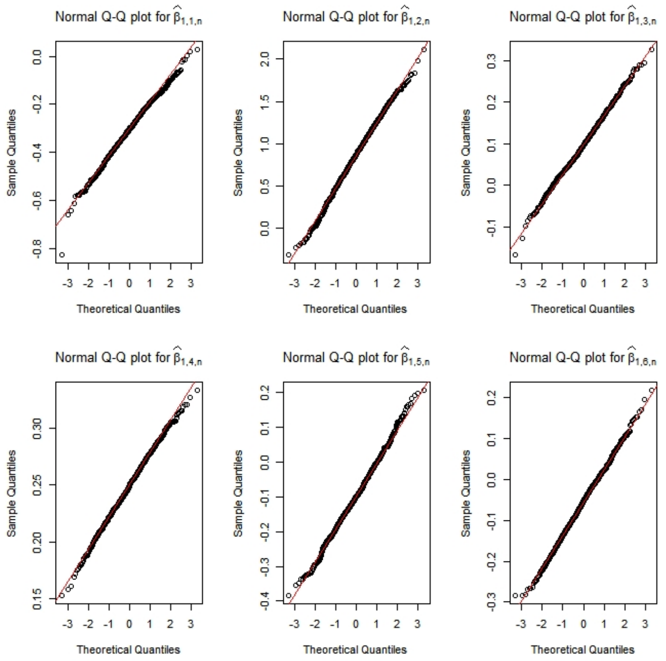

Fig. 2. Normal $Q-Q$ plots for $\widehat{\beta}_{1,1,n}, \ldots, \widehat{\beta}_{1,6,n}$ with $n = 2000$, 25% of zero-inflation.

Acknowledgments.
The authors acknowledge the Associate Editor and anonymous reviewers for their helpful comments and suggestions that led to an improved version of this paper.

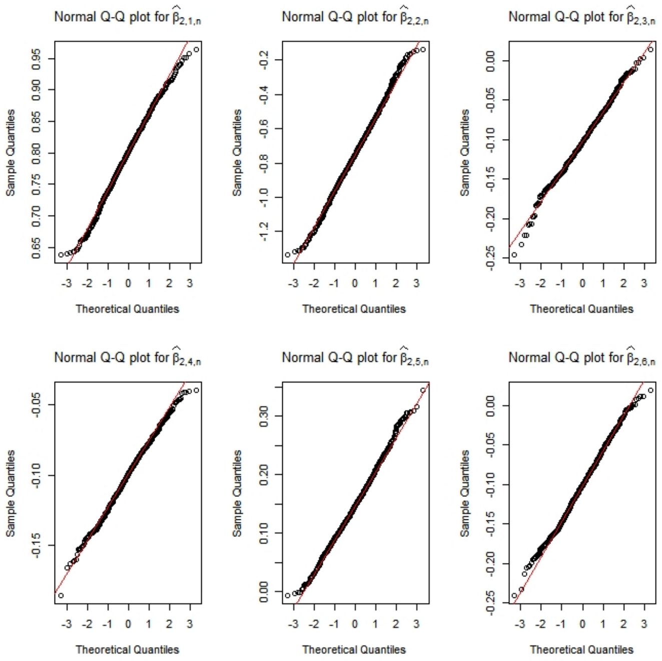

Fig. 3. Normal $Q-Q$ plots for $\widehat{\beta}_{2,1,n},\ldots,\widehat{\beta}_{2,6,n}$ with $n = 2000$, 25% of zero-inflation.

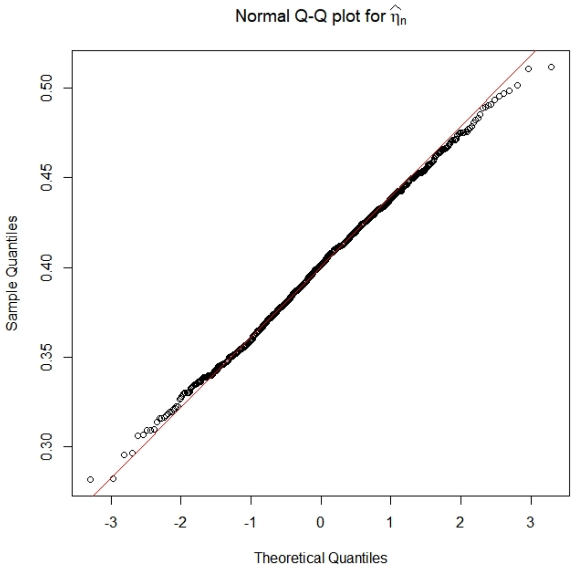

Fig. 4. Normal $Q-Q$ plots for $\widehat{\eta}_n$ with $n = 2000$, 25% of zero-inflation.

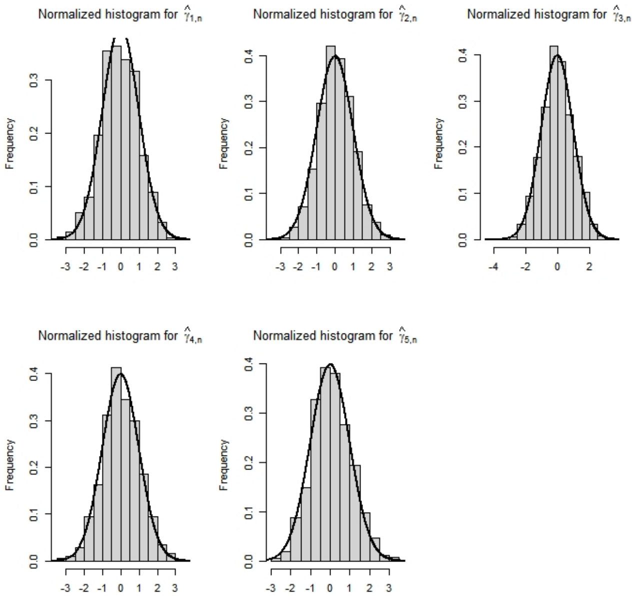

Fig. 5. Histograms of the normalized estimates $(\widehat{\gamma}_{j,n} - \gamma_j)/\text{s.e.}(\widehat{\gamma}_{j,n})$, $j = 1, \ldots, 5$ with $n = 2000$ and 25% of zero-inflation.

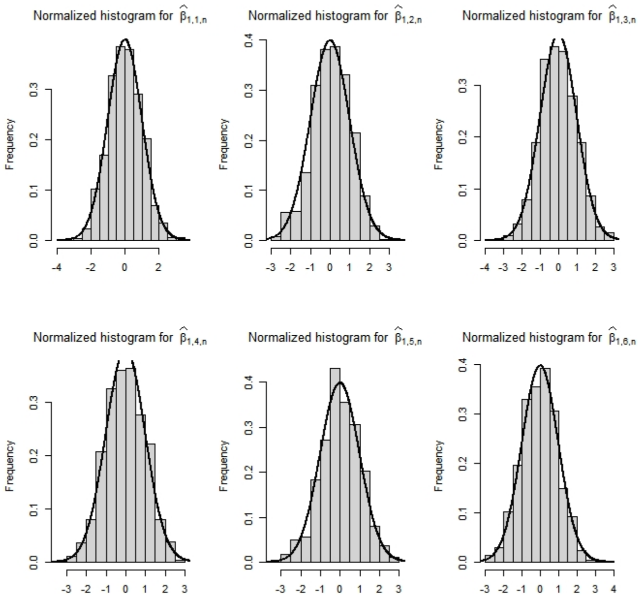

Fig. 6. Histograms of the normalized estimates $(\widehat{\beta}_{1,j,n} - \beta_{1,j})/\text{s.e.}(\widehat{\beta}_{1,j,n})$, $j = 1, \ldots, 6$ with $n = 2000$ and 25% of zero-inflation.

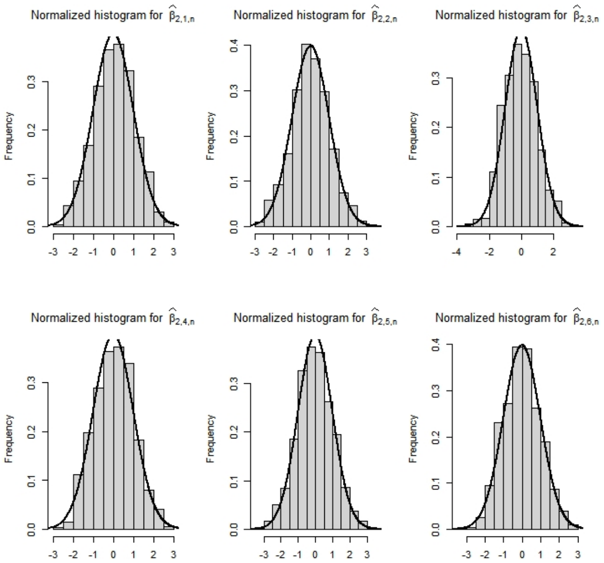

Fig. 7. Histograms of the normalized estimates $(\widehat{\beta}_{2,j,n} - \beta_{2,j})/\text{s.e.}(\widehat{\beta}_{2,j,n})$, $j = 1, \ldots, 6$ with $n = 2000$ and 25% of zero-inflation.

References

Al Muhayfith, F.E., Alzaid A.A., and Omair, M.A., 2016. *On bivariate Poisson regression models*. Journal of King Saud University - Science, 28(2):178-189.

Bermúdez, L., 2009. A priori ratemaking using bivariate Poisson regression models. *Insurance Math. Econom.* 44, 135-141.

Bermúdez, L., Karlis, D., 2012. A finite mixture of bivariate Poisson regression models with an application to insurance ratemaking. *Comput. Stat. Data Anal.* 56, 3988-3999.

Deb P., and Trivedi. P. K., 1997. Demand for medical care by the elderly : *a finite mixture approach. Journal of Applied Econometrics*, 12(3) :313-336.

Diallo, A., Diop, A., Dupuy, J.-F., 2017. Asymptotic properties of the maximum likelihood estimator in zero-inflated binomial regression. *Communications in Statistics - Theory and Methods*, 46(20), 9930-9948.

Diallo, A., Diop, A., Dupuy, J.-F., 2018. Analysis of multinomial counts with joint zero-inflation, with an application to health economics. *Journal of Statistical Planning and Inference*, 194, 85-105.

Diallo, A., Diop, A., Dupuy, J.-F., 2019. Estimation in zero-inflated binomial regression with missing covariates. *Statistics*, 53(4):839-865.

Dietz, E., Böhning, D., 2000. On estimation of the Poisson parameter in zero-modified Poisson models. *Computational Statistics & Data Analysis*, 34(4), 441-459.

Faroughi. P., and Ismail, N., 2016. Bivariate zero-inflated negative binomial regression model with applications. *Journal of Statistical Computation and Simulation*, 87(3), 457-477.

Foutz R.V., 1977. On the unique consistent solution to the likelihood equations. *Journal of the American Statistical Association*, 72(357):147-148.

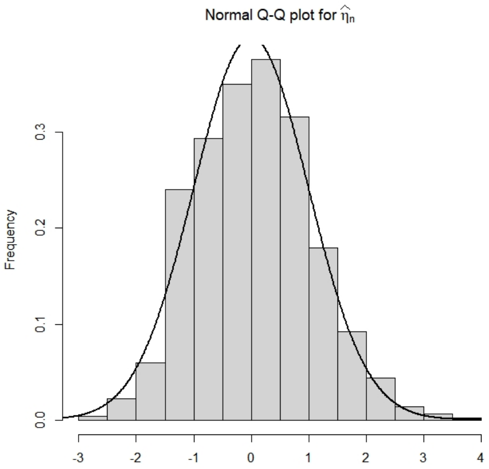

Fig. 8. Histograms of the normalized estimates $(\widehat{\eta}_n - \eta_j)/\text{s.e.}(\widehat{\eta}_n)$, with $n = 2000$ and 25% of zero-inflation.

Garay, A.M., Hashimoto E.M., Ortega, E.M., and Lachos, V.H., 2011. On estimation and influence diagnostics for zero-inflated negative binomial regression models. *Computational Statistics & Data Analysis*, 55(3):1304-1318.

Gurmu, S., Elder, J., 2000. Generalized bivariate count data regression models. *Economics Letters*. 68, 31-36

Hall D.B., 2000. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*, 56(4):1030-1039.

Hall, D.B., and Berenhaut, K.S., 2002. Score tests for heterogeneity and overdispersion in zero-inflated Poisson and binomial regression models. *Can. J. Statistics*, 30(3):415-430.

Hsieh SH, Lee SM, Shen PS. 2010. *Logistic regression analysis of randomized response data with missing covariates*. J Stat Plan Inference. 140(4):927-940.

Henningsen, A., and Toomet, O., 2011. maxLik: A package for maximum likelihood estimation in R. *Comput Stat*, 26(3):443-458.

Karlis D., Ntzoufras, I., 2003. Analysis of sports data by using bivariate Poisson models. *J Royal Statistical Soc D*, 52(3):381-393.

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

Lee S.-M., Lukusa M. T., Li C.-S., 2020. Estimation of a zero-inflated Poisson regression model with missing covariates via nonparametric multiple imputation methods. *Computational Statistics*, 35, 725-754

Li, C.-S., Lu, J.-C., Park, J., Kim, K., Brinkley, P.A., and Peterson J.P., 1999. Multivariate Zero-Inflated Poisson Models and Their Applications. *Technometrics*, 41(1):29-38.

Li, C.-S., 2011. A lack-of-fit test for parametric zero-inflated Poisson models. *Journal of Statistical Computation and Simulation*, 81(9):1081-1098.

Lim, H.K., Li, W.K., and Yu, P.L.H., 2014. Zero-inflated Poisson regression mixture model. *Computational Statistics & Data Analysis*, 71:151-158.

Lukusa, M.T., Lee, S.-M., and Li, C.-S., 2016. Semiparametric estimation of a zero-inflated Poisson regression model with missing covariates. *Metrika*, 79(4):457-483.

Moghimbeigi, A., Eshraghian, M.R., Mohammad K., and Mcardle, B., 2008. *Multilevel zero-inflated negative binomial regression modeling for over-dispersed count data with extra zeros. Journal of Applied Statistics*, 35(10):1193-1202.

Monod, A., 2014. Random Effects Modeling and the Zero-Inflated Poisson Distribution. *Communications in Statistics - Theory and Methods*, 43(4):664-680.

Mwalili, S., M., Lesaffre, E., and Declerck D., 2008. The zero-inflated negative binomial regression model with correction for misclassification: an example in caries research. *Stat. Methods Med. Res.* 17(2):123-139.

Newey, W., K., 1991. Uniform Convergence in Probability and Stochastic Equicontinuity. *Econometrica*. 59(4):1161.

Ridout, M., Hinde, J., and Demetrio, C.G.B., 2001. A Score Test for Testing a Zero-Inflated Poisson Regression Model Against Zero-Inflated Negative Binomial Alternatives. *Biometrics*. 57(1):219-223.

Wang, K., Lee A., H., Yau, K.K,W., and Carrivick, P.J.W., 2003. A bivariate zero-inflated Poisson regression model to analyze occupational injuries. *Accident Analysis & Prevention*, 35(4):625-629.

Yang, M., Das, K., and Majumdar A., 2016. Analysis of bivariate zero inflated count data with missing responses. *Journal of Multivariate Analysis*, 148:73-82.