P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : https://tci-thaijo.org/index.php/SciTechAsia
Science & Technology Asia
Vol. 27 No.4 October - December 2022
Page: [152-174]

Original research article

Bayesian Inference of Discrete Weibull Regression Model for Excess Zero Counts

Dusit Chaiprasithikul, Monthira Duangsaphon*

Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathum Thani 12120, Thailand

Received 2 November 2021; Received in revised form 24 December 2021
Accepted 9 March 2022; Available online 31 December 2022

ABSTRACT

This research aimed to study the use of Bayesian estimation for the zero-inflated and hurdle discrete Weibull regression models. Moreover, this study compared the performance of the Bayesian estimation with uniform noninformative priors and informative priors using the random walk Metropolis algorithm and the maximum likelihood estimation. A simulation study was conducted to compare the performance of three different estimation methods by using mean square error with three cases of a simple explanatory variable. A real dataset was analyzed to see how the model works in practice. The results from the simulation study showed that the Bayesian estimation with informative priors is more appropriate for the zero-inflated and hurdle discrete Weibull regression models than other methods. Moreover, the results from a real data application revealed that the Bayes estimators with informative priors for the zero-inflated and hurdle discrete Weibull regression models are the best fitting models.

Keywords: Bayesian estimation; Discrete count data; Hurdle model; Random walk Metropolis algorithm; Zero-inflated model

1. Introduction

In experimental and observational studies in many fields, including social sciences, industrial, economy, and public health, regression model is demonstrated to count response variables. For such counts, the number of times cardiac arrest happens over a fixed period of time, aside from the number of postoperative complications over a fixed period of time, the number of epileptic seizures experienced over a fixed period of time, the number of claims in an insurance company over a fixed period of time, and the number of recurrent circuit breaker failures over a fixed period of time. A Poisson regression model is commonly used to evaluate the relationship between the count response variable and explanatory variables, e.g., [1-4]. However, its use is limited of the equality of the mean and Variance assumption with real data. A negative binomial regression and the

***Corresponding author:** monthira@mathstat.sci.tu.ac.th

Poisson-inverse Gaussian model are used to account for over-dispersion, e.g., [5-7]. However, these models may be unable to fit with data that is excessive zero or high skewed data. On the other hand, Conway-Maxwell Poisson model is used to deal for under-dispersion, e.g., [8-10]. One of the count models examined to handle for under-dispersion and over-dispersion is a discrete Weibull model, e.g., [11-15].

A discrete Weibull regression model was proposed by Kalktawi [11]. The cumulative distribution function and the probability mass function of a discrete Weibull random variable Y are given by

$$F_Y(y;q,\beta)=\begin{cases}1-q^{(y+1)^{\beta}} & ;y=0,1,\dots\\ 0 & ;\text{otherwise}\end{cases} \tag{1.1}$$

and

$$p_Y(y;q,\beta)=\begin{cases}q^{y^{\beta}}-q^{(y+1)^{\beta}} & ;y=0,1,\dots\\ 0 & ;\text{otherwise}\end{cases} \tag{1.2}$$

respectively, where $0<q<1$ and $\beta>0$ are the shape parameters. Moreover, the parameter $q=1-p_Y(0;q,\beta)$, which is the probability of Y being more than zero. Kalktawi showed how a discrete Weibull regression model can be adapted to address over-dispersion and under-dispersion via the log-log link function under the parameter q.

The over-dispersion data may be caused by count data with excessive zeros that are common in many application areas. Several models have been proposed to handle count data with excessive zeros; the zero-inflated Poisson (ZIP) regression model and the hurdle Poisson (HP) regression model, e.g., [17-20], the zero-inflated negative binomial (ZINB) regression model and the hurdle negative binomial (HNB) regression model, e.g., [21-25], the zero-inflated discrete Weibull (ZIDW) regression model and the hurdle discrete Weibull (HDW) regression model, e.g., [11]. Zero counts can be classified into two types: structural zeros and sampling zeros. Structural zeros are zero responses that count response variables are always zero counts. In contrast, sampling zeros or random zeros occur to count response variables that can be greater than zero, but it appears to be zero counts, due to the sampling variability [26]. The excess zero counts models include the zero-inflated model and the hurdle model; the difference between these two models is zero counts in the zero-inflated model that can come from both types that is structural zeros and sampling zeros, whereas in the hurdle model can come from only structural zeros.

Methods to estimate the regression model parameters precisely and efficiently are very important. The maximum likelihood estimation is valid for an asymptotically large sample size of data. Additionally, the maximum likelihood estimation is used for only empirical knowledge from the likelihood function. Alternatively, the Bayesian estimation is an interesting method because it uses information from both prior knowledge about the parameters from the prior probability distribution and empirical knowledge from the likelihood function. Hence, the performance of the Bayesian estimation depends upon the prior distribution that defined. According to determining the prior distribution, it is very important in the Bayesian estimation, if the researchers have no prior knowledge of the parameters; then they can use the noninformative prior distribution. Contrastingly, the researchers use the informative prior distribution when knowing about prior knowledge of the parameters.

Furthermore, the Bayesian estimation is offered for small sample problems, e.g., [27]. However, the disadvantage of the Bayesian estimation is that it takes a long time to compute.

Kalktawi [11] performed the maximum likelihood for estimation of parameter based on the standard model, censoring model, and excessive zero models. Moreover, there are many paper works considering the Bayesian inference for estimation in discrete regression model with excess zeros, such as [13, 28]. Unfortunately, there are no conjugate priors in the context of discrete Weibull regression. Haselimashhadi et al. [13] had recently proposed the Bayesian implementation of the discrete Weibull regression model under a uniform noninformative prior. They showed the effectiveness of the Bayesian estimation procedure via a simulation study both in the case of data drawing a discrete Weibull regression model and in case of model misspecification; Poisson and negative binomial. In addition, the applicability of Bayesian discrete Weibull regression model to health data. It is often more natural to express prior information directly in term of the parameters, the regression coefficients that can be a real number which correspond to the possible values of a normal distribution. There are papers selecting the prior distribution of the regression coefficients as a normal distribution [29-31]. The parameter β from the discrete Weibull distribution is equivalent to the shape parameter β from the continuous Weibull distribution that $\beta > 0$ corresponds to the possible values of a Gamma distribution. There are papers selecting the prior distribution of β as a Gamma distribution [32, 33].

The main objective of this paper is to perform the Bayesian inference for the ZIDW and HDW regression models under uniform noninformative and informative prior distributions. This study constructed the estimators of the parameters under squared error loss function which is the expected value of the joint posterior density function. The main problem faced when dealing with the Bayesian estimation that comes from the integral of the posterior probability distribution without a closed form. Therefore, in this case, it chose a one of the Markov chain Monte Carlo (MCMC) methods which is the random walk Metropolis algorithm in order to estimate the parameters.

The remainder of this paper is organized as follows. In Section 2, it introduces the discrete Weibull regression, and present the Bayesian estimation via the random walk Metropolis algorithm for discrete Weibull regression with excess zero counts; the ZIDW and HDW regression models. In Section 3, it investigates the performance of the estimations through a simulation study and applies the computational methods to a real dataset. Finally, the findings are concluded in Section 4.

2. Materials and Methods

2.1 Discrete Weibull regression

Regression analysis for count data is a statistical process to measure the relationship between a count variable and one or more explanatory variables. The discrete Weibull regression can link the independent variables via the shape parameters q and β. In this paper, it linked the explanatory variables only via the shape parameter q.

This study determines Y_i as a count response variable which takes only the non-negative integer values with the k explanatory variables. Assume that the parameter q_i is related to k explanatory variables $\mathbf{x}_i$ via the log-log link function:

$$\log(-\log(q_i)) = \alpha_0 + \alpha_1 x_{i1} + \ldots + \alpha_k x_{ik} = \mathbf{x}_i \boldsymbol{\alpha} \qquad (2.1)$$

where $q_i = e^{-e^{(\mathbf{x}_i \boldsymbol{\alpha})}}$, $\mathbf{x}_i = (1 \quad x_{i1} \quad \cdots \quad x_{ik})$, and $\boldsymbol{\alpha} = (\alpha_0 \quad \alpha_1 \quad \cdots \quad \alpha_k)'$.

The conditional probability mass function of Y_i given $\mathbf{x}_i$ can be written as

$$p_Y(y_i|\mathbf{x}_i) = \begin{cases} q_i^{y_i^\beta} - q_i^{(y_i+1)^\beta} & ; y_i = 0,1,\ldots \\ 0 & ;\text{otherwise} \end{cases} \qquad (2.2)$$

Given n observations y_i and $(x_{i1}, x_{i2}, \ldots, x_{ik})$, $i = 1,2,\ldots,n$, from Eq. (2.2) for the count response variable Y_i and k explanatory variables, respectively, the likelihood function and the log-likelihood function of the discrete Weibull regression model are given by

$$L_{DW} = \prod_{i=1}^{n} \left[q_i^{y_i^\beta} - q_i^{(y_i+1)^\beta} \right] \qquad (2.3)$$

and

$$l_{DW} = \sum_{i=1}^{n} \log\left[q_i^{y_i^\beta} - q_i^{(y_i+1)^\beta} \right] \qquad (2.4)$$

respectively.

2.2 Bayesian estimation for discrete Weibull regression with excess zero counts

In this section, it presents the Bayesian inference for the ZIDW and HDW regression models and defines the random walk Metropolis algorithm.

2.2.1 Zero-inflated discrete Weibull regression

The zero-inflated distribution can be expressed as two-component mixture distributions where there are one component degenerate distribution at zero and a regular discrete Weibull distribution. The mixing parameter π of these two distributions is thought to be completely unknown. Thus, the probability mass function of the ZIDW is

$$P(Y = y) = \begin{cases} \pi + (1-\pi) p_Y(0) & ; y = 0 \\ (1-\pi) p_Y(y) & ; y = 1,2,\ldots \end{cases} \qquad (2.5)$$

where $0 < \pi < 1$. The parameter π in Eq. (2.5), also known as the probability or proportion of a structural zero [34].

In the zero-inflated regression model, the proportion parameter π is related to k explanatory variables via any link function. This paper assumes the logit link function for the parameter π (Lambert, 1992). Let $\mathbf{z}_i$ be k explanatory variables and $\boldsymbol{\gamma} = (\gamma_0 \quad \gamma_1 \quad \cdots \quad \gamma_k)'$ represents the associated regression parameters vector. Hence, the parameter π_i can be related to k explanatory variables as follows:

$$\text{logit}(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i} \right) = \gamma_0 + \gamma_1 z_{i1} + \ldots + \gamma_k z_{ik} = \mathbf{z}_i \boldsymbol{\gamma} \qquad (2.6)$$

where $\pi_i = \left(e^{-\mathbf{z}_i\boldsymbol{\gamma}} + 1\right)^{-1}, \mathbf{z}_i = \left(1 \;\; z_{i1} \;\; \cdots \;\; z_{ik}\right)$.

The count response variable, Y_i, is determined to have a discrete Weibull distribution, where the parameter q_i is related to k explanatory variables via the log-log link function in Eq. (2.1). Thus, the conditional probability mass function of Y_i given $\mathbf{x}_i$ and $\mathbf{z}_i$ can be written as

$$p_Y\left(y_i \middle| \mathbf{x}_i, \mathbf{z}_i\right) = \begin{cases} \pi_i + \left(1-\pi_i\right) p_Y\left(0 \middle| \mathbf{x}_i\right) \\ ; y_i = 0 \\ \left(1-\pi_i\right) p_Y\left(y_i \middle| \mathbf{x}_i\right) \\ ; y_i = 1, 2, \ldots \end{cases} \quad (2.7)$$

Given n independent observations y_i, x_{ij}, and z_{ij}, $i = 1, 2, \ldots, n, j = 1, 2, \ldots, k$, from Eq. (2.7) where δ_i be the zero indicator that can be specified as

$$\delta_i = I\left(y_i = 0\right) = \begin{cases} 1 & ; y_i = 0 \\ 0 & ; y_i > 0 \end{cases} \quad (2.8)$$

The likelihood function and the log-likelihood function of the ZIDW regression model are given by

$$L_{ZIDW} = \prod_{i=1}^{n} w_{1i}\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right)^{\delta_i} \prod_{i=1}^{n} w_{2i}\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right)^{1-\delta_i} \quad (2.9)$$

and

$$l_{ZIDW} = \sum_{i=1}^{n} \delta_i \log\left[w_{1i}\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right)\right] + \sum_{i=1}^{n} \left(1-\delta_i\right) \log\left[w_{2i}\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right)\right] \quad (2.10)$$

respectively, where

$$w_{1i}\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right) = \pi_i + \left(1-\pi_i\right)\left(1-q_i\right),$$

and $w_{2i}\left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right) = \left(1-\pi_i\right)\left[q_i^{y_i^{\beta}} - q_i^{\left(y_i+1\right)^{\beta}}\right]$.

2.2.2 Hurdle discrete Weibull regression

The hurdle model was first used by [35] and the hurdle model as a modified count data model was proposed by [36]. This model was developed to deal with count data that has excessive zeros, as another option of the zero-inflated model. The hurdle model is two-part models that state a process for the zero counts and the positive counts. The basic idea of the hurdle model is the zero counts that are generated by some binary process, π, and the non-zero positive counts are observed with a probability based on a truncated parent count model $p_{Ptr}\left(y; \boldsymbol{\theta}\right) = \frac{p_P\left(y; \boldsymbol{\theta}\right)}{1 - p_P\left(0; \boldsymbol{\theta}\right)}$ that needs to be multiplied by $1-\pi$ to ensure probabilities sum to one.

The count response variable, Y_i, is given to have a discrete Weibull distribution, where the parameter q_i is related to k explanatory variables via the log-log link function in Eq. (2.1), and the parameter π_i to model the binary outcome $Y_i = 0$ versus $Y_i > 0$ is related to k explanatory variables via the logit link function in Eq. (2.6).

Thus, the conditional probability mass function of Y_i given $\mathbf{x}_i$ and $\mathbf{z}_i$ can be written as

$$p_Y\left(y_i \middle| \mathbf{x}_i, \mathbf{z}_i\right) = \begin{cases} \pi_i \;\; ; y_i = 0 \\ \left(1-\pi_i\right) \dfrac{p_Y\left(y_i \middle| \mathbf{x}_i\right)}{1 - p_Y\left(0 \middle| \mathbf{x}_i\right)} \\ ; y_i = 1, 2, \ldots \end{cases} \quad (2.11)$$

Given n independent observations y_i, x_{ij}, and z_{ij}; $i=1,2,\ldots,n, j=1,2,\ldots,k$, from Eq. (2.11) where δ_i be the zero indicator that can be specified as

$$\delta_i = I\left(y_i = 0\right) = \begin{cases} 1 & ; y_i = 0 \\ 0 & ; y_i > 0 \end{cases} \quad (2.12)$$

The likelihood function and the log-likelihood function of the HDW regression model are given by

$$L_{HDW} = \prod_{i=1}^{n} [\pi_i]^{\delta_i} \prod_{i=1}^{n} \left[(1-\pi_i) \frac{w_i(\boldsymbol{\alpha},\beta)}{\left(e^{-e^{(\mathbf{x}_i\boldsymbol{\alpha})}} \right)} \right]^{1-\delta_i} \quad (2.13)$$

and

$$l_{HDW} = \sum_{i=1}^{n} \delta_i \log(\pi_i) + \sum_{i=1}^{n} (1-\delta_i) \log(1-\pi_i) + \sum_{i=1}^{n} (1-\delta_i) \left[\log\left[w_i(\boldsymbol{\alpha},\beta) \right] + e^{(\mathbf{x}_i\boldsymbol{\alpha})} \right] \quad (2.14)$$

respectively where

$$w_i(\boldsymbol{\alpha},\beta) = q_i^{y_i^{\beta}} - q_i^{(y_i+1)^{\beta}}.$$

Moreover, the explanatory variables that affect the parameter q_i may or may not be the same as the explanatory variables that affect the parameter π_i.

2.2.3 Bayesian inference

This study investigates the performance of the estimation through both noninformative and informative prior distributions. Firstly, if no prior information is available, we can resort to a default flat prior then it's easy to focus on the uniform noninformative prior distribution, i.e. $\pi(\boldsymbol{\theta}) \propto 1$, proposed by [13]. On the other hand, if prior information is available, we can perform to the informative prior distribution. Typically, the prior distribution should include all possible values of parameter. The possible values of α_j and γ_j are real number which correspond to the possible values of a normal distribution. This study selected the prior distribution of α_j and γ_j as a normal distribution with the hyperparameters are $\left(\mu_{\alpha_j}, \sigma^2_{\alpha_j}\right)$ and $\left(\mu_{\gamma_j}, \sigma^2_{\gamma_j}\right)$, $j=0,1,\ldots,k$, respectively and the prior distribution of β, that $\beta > 0$ corresponds to the possible values of a Gamma distribution, as a Gamma distribution with the hyperparameters are (a,b). The joint prior distribution of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β under the independence assumption is

$$\pi(\boldsymbol{\theta}) = \pi(\alpha_0)\cdots\pi(\alpha_k)\pi(\gamma_0)\cdots\pi(\gamma_k)\pi(\beta) \quad (2.15)$$

The choice of the hyperparameters' values will generally be modified by available information of dataset to improve the Bayes estimators. At this moment, they are left unspecified.

The joint posterior density function of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β can be written as:

$$p\left(\boldsymbol{\theta} \middle| \mathbf{y},\mathbf{x},\mathbf{z},\boldsymbol{\delta}\right) = \frac{L\left(\boldsymbol{\theta} \middle| \mathbf{y},\mathbf{x},\mathbf{z},\boldsymbol{\delta}\right)\pi(\boldsymbol{\theta})}{\iint\cdots\int L\left(\boldsymbol{\theta} \middle| \mathbf{y},\mathbf{x},\mathbf{z},\boldsymbol{\delta}\right)\pi(\boldsymbol{\theta})d\alpha_0\cdots d\gamma_k d\beta} \propto L\left(\boldsymbol{\theta} \middle| \mathbf{y},\mathbf{x},\mathbf{z},\boldsymbol{\delta}\right)\pi(\boldsymbol{\theta}) \quad (2.16)$$

The Bayes estimator of function $h(\boldsymbol{\theta})$ of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β under squared error loss function is the expected value of function $h(\boldsymbol{\theta})$ under the joint posterior density function. Therefore, the Bayes estimator of function $h(\boldsymbol{\theta})$ is given by

$$\hat{h}(\boldsymbol{\theta}) = \iint \cdots \int h(\boldsymbol{\theta}) p\left(\boldsymbol{\theta} \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) d\alpha_0 \cdots d\gamma_k d\beta \tag{2.17}$$

Since the integral in Eq. (2.17) does not have a closed form, this study chose the Metropolis-Hastings algorithm to estimate the Bayes estimators.

The Metropolis-Hastings algorithm is a MCMC method for simulating a sample from a probability distribution that is the target distribution from which direct sampling is difficult. This algorithm is similar to acceptance-rejection method; the proposal (candidate) value can be generated from the proposal distribution. Then, the proposal value is accepted with an acceptance probability. Moreover, the Metropolis-Hastings algorithm is converging to the target distribution itself. In this paper, it chose a random walk Metropolis algorithm, which is a special case of a Metropolis-Hastings algorithm.

This study determines the joint posterior density function of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β, $p\left(\boldsymbol{\theta} \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right)$, in Eq. (2.16) as the target distribution, while $\boldsymbol{\theta}$ is the current state value, and $\boldsymbol{\theta}^*$ is the proposal value generated from the proposal distribution $q\left(\boldsymbol{\theta}^* \middle| \boldsymbol{\theta}\right)$. Then, the proposal value $\boldsymbol{\theta}^*$ is accepted with the probability $p = \min\left(1, R_{\boldsymbol{\theta}}\right)$, where

$$R_{\boldsymbol{\theta}} = \frac{L\left(\boldsymbol{\theta}^* \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) \pi(\boldsymbol{\theta}^*)}{L\left(\boldsymbol{\theta} \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) \pi(\boldsymbol{\theta})} \times \frac{q\left(\boldsymbol{\theta} \middle| \boldsymbol{\theta}^*\right)}{q\left(\boldsymbol{\theta}^* \middle| \boldsymbol{\theta}\right)} \tag{2.18}$$

In the random walk Metropolis algorithm, the proposal distribution is symmetrical, depending only on the distance between the current state value and the proposal value. Then, the proposal value $\boldsymbol{\theta}^*$ is accepted with probability $p = \min\left(1, R_{\boldsymbol{\theta}}\right)$, where

$$R_{\boldsymbol{\theta}} = \frac{L\left(\boldsymbol{\theta}^* \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) \pi\left(\boldsymbol{\theta}^*\right)}{L\left(\boldsymbol{\theta} \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) \pi\left(\boldsymbol{\theta}\right)} \tag{2.19}$$

The iterative steps of the random walk Metropolis algorithm can be described as follows:

Step 1: Initialize the parameters $\boldsymbol{\theta}^{(0)} = \left(\boldsymbol{\alpha}^{(0)}, \boldsymbol{\gamma}^{(0)}, \beta^{(0)}\right)$ for the algorithm using the maximum likelihood estimation (MLE) of the parameters $\boldsymbol{\theta} = \left(\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta\right)$.

Step 2: For $l = 1, 2, \ldots, L$ repeat the following steps;

a. Generate random error vector $\boldsymbol{\varepsilon}$ from a multivariate normal distribution with a zero-mean vector and variance-covariance matrix as a diagonal matrix in which the diagonal elements are the diagonal of the inverse of the observed Fisher's information matrix; $\boldsymbol{\varepsilon} \sim \mathcal{N}\left(\boldsymbol{\mu} = \mathbf{0}, \boldsymbol{\Sigma} = diag\left(I^{-1}\left(\boldsymbol{\theta}\right)\right)\right)$.

Then, set $\boldsymbol{\theta}^* = \boldsymbol{\theta}^{(l-1)} + \boldsymbol{\varepsilon}$.

b. Calculate $p = \min\left(1, R_{\boldsymbol{\theta}}\right)$ where $R_{\boldsymbol{\theta}} = \dfrac{L\left(\boldsymbol{\theta}^* \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) \pi\left(\boldsymbol{\theta}^*\right)}{L\left(\boldsymbol{\theta} \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right) \pi\left(\boldsymbol{\theta}\right)}$.

c. Generate u from a uniform distribution; $u \sim U(0,1)$.

If $u \le p$, accept $\boldsymbol{\theta}^*$ and set $\boldsymbol{\theta}^{(l)} = \boldsymbol{\theta}^*$ with probability p.

If $u > p$, reject $\boldsymbol{\theta}^*$ and set $\boldsymbol{\theta}^{(l)} = \boldsymbol{\theta}^{(l-1)}$ with probability $1-p$.

Step 3: Remove B of the chain for *burn-in*.

Step 4: Calculate the estimated values of the Bayes estimators of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β from the average of the generated values given by

$$\hat{\theta}_{Bayes} = \frac{1}{L-B} \sum_{l=B+1}^{L} \theta^{(l)} \qquad (2.20)$$

where θ is a parameter in vector $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\gamma}, \beta)$.

3. Results and Discussion

3.1 Simulation study

In this section, a Monte Carlo simulation is conducted to assess and compare the performance of the maximum likelihood estimation and the Bayesian estimation for the ZIDW and HDW regression models with various selected sample sizes (n) are 60, 90, 120, 150, and 180. The three cases of a simple explanatory variable are considered: a Bernoulli distribution with probability of success 0.4 $(x \sim Ber(0.4))$, a uniform distribution that lies between 0 and 3 $(x \sim U(0,3))$, and a normal distribution with mean 2 and variance 1 $(x \sim N(2,1))$. Moreover, the explanatory variable that affects the parameter q_i is the same as the explanatory variable that affects the parameter π_i; $z = x$. In particular, this study selected $(\alpha_0, \alpha_1, \gamma_0, \gamma_1, \beta) = (-2, -1.7, 1.5, -1.7, 2.2)$ for $x \sim Ber(0.4)$, $(\alpha_0, \alpha_1, \gamma_0, \gamma_1, \beta) = (-2, -1.7, 1.5, -0.9, 2.2)$ for $x \sim U(0,3)$ and $x \sim N(2,1)$. It also fixed the hyperparameters' values of α_j and γ_j, $j = 0,1$, as the maximum likelihood estimators and the variance of the maximum likelihood estimators. Also, this study fixed the hyperparameters' values of β as 1 and the maximum likelihood estimator. It also computed q_i and π_i for each type of data from the log-log link function in Eq. (2.1) and the logit link function in Eq. (2.6), respectively. Then this study generated the count response variables $Y_1, Y_2, \ldots, Y_n$ for the ZIDW regression model from Eq. (2.7) and the HDW regression model from Eq. (2.11) as follows:

a. the ZIDW regression model

Generate u from a uniform distribution; $u \sim U(0,1)$.

If $u \le \pi_i$, set $y_i = 0$ with probability π_i.

If $u > \pi_i$, generate y_i using function `rdw( )` from package `DWreg` in R.

b. the HDW regression model

Generate a random sample y_{Ber} from a Bernoulli distribution; $y_{Ber} \sim Ber(1-\pi_i)$.

If $y_{Ber} = 0$, set $y_i = 0$ with probability π_i.

If $y_{Ber} = 1$, generate a sample y_{tr} from the zero-truncated discrete Weibull using function `rdw( )` from package `DWreg` in R. Then, if a zero is generated, drop it and re-sample again until a non-zero sample is generated.

Then, this study received the response variables $y_1, y_2, \ldots, y_n$ as observed data for the ZIDW regression model from Eq. (2.7) and the HDW regression model from Eq. (2.11) and the indicator $\delta_1, \delta_2, \ldots, \delta_n$ is the zero indicator for the ZIDW regression model from Eq. (2.8) and the HDW regression model from Eq. (2.12).

This study calculated the maximum likelihood estimators of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β by minimizing the negative log-likelihood function of the ZIDW regression model in Eq. (2.10) and the HDW regression model in Eq. (2.14). Then, it got $\hat{\boldsymbol{\theta}}_{ML}^{(m)}$ using function `optim( )` from package `stats` in R. Next, this study calculated the Bayes estimators of the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β with uniform noninformative priors and informative priors under the squared error loss function by using the random walk Metropolis algorithm with $L = 10{,}000$ replicates and 10% of the chain for *burn-in*; $B = 1{,}000$. Finally, the Bayes estimators is obtained, resulting in $\hat{\boldsymbol{\theta}}_{Bayes(U)}^{(m)}$ and $\hat{\boldsymbol{\theta}}_{Bayes}^{(m)}$ from Eq. (2.20). Moreover, it was found that $R_{\boldsymbol{\theta}}$ in Step 2(b.) for uniform noninformative priors becomes $R_{\boldsymbol{\theta}} = \dfrac{L\left(\boldsymbol{\theta}^* \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right)}{L\left(\boldsymbol{\theta} \middle| \mathbf{y}, \mathbf{x}, \mathbf{z}, \boldsymbol{\delta}\right)}$.

This study performed the parameter estimates (Est.) and the mean squared error (MSE) of estimators based on $M = 1{,}000$ from the MLE and the Bayesian estimation with uniform noninformative priors (Bayes(Uniform)) and informative priors (Bayes(Informative)) which are reported in Table 1 to Table 3 for the ZIDW regression model when $x \sim Ber(0.4)$, $x \sim U(0,3)$, and $x \sim N(2,1)$, respectively. Moreover, Table 4 to Table 6 for the HDW regression model show $x \sim Ber(0.4)$, $x \sim U(0,3)$, and $x \sim N(2,1)$ respectively.

Table 1. Est. and MSE for ZIDW; $x \sim Ber(0.4)$ and $\boldsymbol{\theta} = (-2, -1.7, 1.5, -1.7, 2.2)$.

n (% zeros)	parameter	MLE Est.	MLE MSE	Bayes(Uniform) Est.	Bayes(Uniform) MSE	Bayes(Informative) Est.	Bayes(Informative) MSE
60 (68.72%)	α_0	-2.1387	1.1820	-1.1790	207.8896	-2.1262	**1.0881**
	α_1	-2.0605	1.0013	-2.4038	206.3418	-1.9794	**0.8409**
	γ_0	1.4501	**0.4504**	-1.1389	64.8343	1.4111	0.7812
	γ_1	-1.6981	**0.7052**	-3.0575	45.7127	-1.6970	0.8483
	β	2.5058	0.5812	1.9865	3.2538	2.4334	**0.5508**
90 (68.90%)	α_0	-2.2065	0.6695	-2.1108	1.1356	-2.1888	**0.6585**
	α_1	-1.9178	0.4531	-1.8432	0.6561	-1.8898	**0.4242**
	γ_0	1.5286	**0.2075**	1.0682	5.3707	1.5232	0.2233
	γ_1	-1.7453	**0.3230**	-2.0282	5.3475	-1.7570	0.3429
	β	2.4454	0.3289	2.3120	0.8115	2.4061	**0.3135**
120 (68.82%)	α_0	-2.1203	0.4276	-2.2024	0.5056	-2.1070	**0.4069**
	α_1	-1.8334	0.2950	-1.8401	0.3213	-1.8078	**0.2648**
	γ_0	1.5155	0.1515	1.4890	0.4629	1.5138	**0.1337**
	γ_1	-1.7332	0.2439	-1.8266	0.8740	-1.7421	**0.2285**
	β	2.3509	0.2174	2.3804	0.2585	2.3186	**0.2032**
150 (69.07%)	α_0	-2.0780	0.3292	-2.0964	0.3369	-2.0696	**0.3096**
	α_1	-1.8242	0.2097	-1.7970	0.2211	-1.8039	**0.2027**
	γ_0	1.4981	0.2945	1.4913	0.1771	1.5075	**0.1015**
	γ_1	-1.7030	0.1823	-1.7673	0.3190	-1.7099	**0.1703**
	β	2.3200	0.1580	2.3072	0.1655	2.2961	**0.1492**
180 (69.09%)	α_0	-2.0430	0.2310	-2.1138	0.2780	-2.0341	**0.2180**
	α_1	-1.7843	0.1701	-1.7884	**0.1612**	-1.7694	0.1655
	γ_0	1.5200	0.1076	1.5249	0.1258	1.5219	**0.0831**
	γ_1	-1.7286	0.1470	-1.7522	0.1722	-1.7361	**0.1368**
	β	2.2735	0.1093	2.3117	0.1257	2.2532	**0.1035**

Note: the boldface identifies the smallest MSE for each cases.

Table 2. Est. and MSE for ZIDW; $x \sim U(0,3)$ and $\boldsymbol{\theta} = (-2,-1.7,1.5,-0.9,2.2)$.

n (% zeros)	parameter	MLE Est.	MLE MSE	Bayes(Uniform) Est.	Bayes(Uniform) MSE	Bayes(Informative) Est.	Bayes(Informative) MSE
60 (53.61%)	α_0	-2.1335	0.6213	-2.2379	0.7035	-2.1278	**0.6074**
	α_1	-1.8567	0.2310	-1.8799	0.2545	-1.8422	**0.2265**
	γ_0	1.5617	**0.4791**	1.6120	1.0144	1.5684	0.4819
	γ_1	-0.9456	**0.1542**	-0.9946	0.1947	-0.9562	0.1576
	β	2.3842	0.2299	2.4257	0.2639	2.3592	**0.2194**
90 (53.84%)	α_0	-2.0680	0.3580	-2.1432	0.3743	-2.0668	**0.3483**
	α_1	-1.8111	0.1305	-1.8280	0.1405	-1.8017	**0.1268**
	γ_0	1.5422	0.3052	1.5940	0.3366	1.5474	**0.3011**
	γ_1	-0.9332	**0.1000**	-0.9678	0.1132	-0.9404	**0.1000**
	β	2.3232	0.1286	2.3542	0.1433	2.3075	**0.1231**
120 (54.35%)	α_0	-2.0779	0.2803	-2.1367	0.2841	-2.0780	**0.2678**
	α_1	-1.7796	0.0850	-1.7876	0.0863	-1.7704	**0.0816**
	γ_0	1.5440	0.2130	1.5787	0.2163	1.5478	**0.2042**
	γ_1	-0.9187	0.0655	-0.9428	0.0698	-0.9241	**0.0644**
	β	2.2996	0.0908	2.3200	0.0952	2.2866	**0.0866**
150 (54.01%)	α_0	-2.0422	0.2183	-2.0954	0.2044	-2.0474	**0.2005**
	α_1	-1.7574	0.0623	-1.7641	0.0643	-1.7503	**0.0608**
	γ_0	1.5068	0.1660	1.5361	0.1630	1.5100	**0.1556**
	γ_1	-0.9070	0.0507	-0.9265	0.0514	-0.9112	**0.0487**
	β	2.2657	0.0686	2.2852	0.0704	2.2582	**0.0654**
180 (53.96%)	α_0	-2.0389	0.1776	-2.0754	0.1642	-2.0396	**0.1613**
	α_1	-1.7464	0.0495	-1.7537	0.0513	-1.7376	**0.0473**
	γ_0	1.5062	0.1484	1.5554	0.1556	1.5084	**0.1354**
	γ_1	-0.9065	0.0473	-0.9368	0.0494	-0.9079	**0.0451**
	β	2.2559	0.0509	2.2681	0.0551	2.2447	**0.0467**

Note: the boldface identifies the smallest MSE for each cases.

Table 3. Est. and MSE for ZIDW; $x \sim N(2,1)$ and $\boldsymbol{\theta} = (-2, -1.7, 1.5, -0.9, 2.2)$.

n (% zeros)	parameter	MLE Est.	MLE MSE	Bayes(Uniform) Est.	Bayes(Uniform) MSE	Bayes(Informative) Est.	Bayes(Informative) MSE
60 (44.28%)	α_0	-2.0926	0.4969	-2.1905	0.5296	-2.1009	**0.4813**
	α_1	-1.8359	0.1514	-1.8664	0.1682	-1.8289	**0.1470**
	γ_0	1.6613	0.6351	1.7533	0.7675	1.6607	**0.6196**
	γ_1	-0.9860	0.1564	-1.0444	0.1942	-0.9952	**0.1555**
	β	2.3566	0.1677	2.4024	0.1921	2.3445	**0.1610**
90 (44.30%)	α_0	-2.0447	0.2984	-2.1143	0.2967	-2.0520	**0.2816**
	α_1	-1.7739	0.0779	-1.7926	0.0834	-1.7681	**0.0756**
	γ_0	1.6252	0.5011	1.6706	0.5275	1.6218	**0.4744**
	γ_1	-0.9634	0.1121	-0.9946	0.1225	-0.9683	**0.1087**
	β	2.2854	0.0853	2.3154	0.0929	2.2769	**0.0820**
120 (44.07%)	α_0	-1.9942	0.2392	-2.0657	0.2265	-2.0091	**0.2200**
	α_1	-1.7498	0.0561	-1.7677	0.0589	-1.7485	**0.0543**
	γ_0	1.6026	0.3738	1.6071	0.3545	1.5869	**0.3385**
	γ_1	-0.9516	0.0840	-0.9622	0.0822	-0.9507	**0.0786**
	β	2.2471	0.0618	2.2794	0.0654	2.2465	**0.0586**
150 (44.10%)	α_0	-2.0005	0.1649	-2.0594	0.1544	-2.0134	**0.1486**
	α_1	-1.7368	0.0396	-1.7512	0.0411	-1.7349	**0.0383**
	γ_0	1.5753	0.2646	1.5778	0.2392	1.5650	**0.2348**
	γ_1	-0.9390	0.0593	-0.9463	0.0553	-0.9383	**0.0544**
	β	2.2349	0.0426	2.2616	0.0439	2.2345	**0.0399**
180 (44.09%)	α_0	-1.9942	0.1470	-2.0529	0.1314	-2.0078	**0.1308**
	α_1	-1.7385	0.0371	-1.7499	0.0370	-1.7370	**0.0354**
	γ_0	1.5882	0.2514	1.5696	**0.2027**	1.5708	0.2113
	γ_1	-0.9441	0.0541	-0.9421	**0.0464**	-0.9411	0.0480
	β	2.2334	0.0400	2.2580	0.0393	2.2337	**0.0368**

Note: the boldface identifies the smallest MSE for each cases.

Table 4. Est. and MSE for HDW; $x \sim Ber(0.4)$ and $\boldsymbol{\theta} = (-2, -1.7, 1.5, -1.7, 2.2)$.

n (% zeros)	parameter	MLE Est.	MLE MSE	Bayes(Uniform) Est.	Bayes(Uniform) MSE	Bayes(Informative) Est.	Bayes(Informative) MSE
60 (67.05%)	α_0	-2.2587	1.0396	-2.0905	1.0826	-2.2464	**0.8719**
	α_1	-2.0509	0.9532	-1.7306	0.7430	-1.9340	**0.7049**
	γ_0	1.5674	0.2292	1.3376	0.2398	1.5366	**0.2176**
	γ_1	-1.7854	0.4210	-1.5302	0.4118	-1.7519	**0.4052**
	β	2.5596	0.5525	2.2297	0.6516	2.4652	**0.5141**
90 (66.85%)	α_0	-2.1468	0.4844	-2.2407	0.5488	-2.1302	**0.4522**
	α_1	-1.9053	0.3983	-1.8761	0.3943	-1.8799	**0.3583**
	γ_0	1.5587	0.1534	1.5410	**0.1530**	1.5584	0.1543
	γ_1	-1.7839	0.2792	-1.7631	**0.2730**	-1.7849	0.2860
	β	2.4032	0.2530	2.4216	0.3208	2.3660	**0.2401**
120 (67.14%)	α_0	-2.1121	0.3829	-2.1772	**0.3563**	-2.0978	0.3676
	α_1	-1.8281	0.2464	-1.8305	0.2612	-1.8091	**0.2223**
	γ_0	1.5441	0.1107	1.5644	0.1260	1.5467	**0.1104**
	γ_1	-1.7507	**0.1970**	-1.7690	0.2272	-1.7558	0.1992
	β	2.3428	0.1821	2.3650	0.1927	2.3141	**0.1750**
150 (67.05%)	α_0	-2.0979	0.2745	-2.1454	0.2788	-2.0847	**0.2741**
	α_1	-1.7914	0.1796	-1.8166	0.2021	-1.7807	**0.1775**
	γ_0	1.5215	**0.0785**	1.5415	0.0852	1.5254	0.0792
	γ_1	-1.7302	0.1451	-1.7397	0.1547	-1.7343	**0.1450**
	β	2.3160	0.1450	2.3404	0.1505	2.2950	**0.1412**
180 (67.02%)	α_0	-2.0882	0.2161	-2.1320	0.2341	-2.0763	**0.2139**
	α_1	-1.7903	0.1470	-1.7986	0.1543	-1.7821	**0.1449**
	γ_0	1.5122	**0.0686**	1.5320	0.0726	1.5154	0.0688
	γ_1	-1.7123	**0.1298**	-1.7350	0.1342	-1.7159	0.1310
	β	2.3061	0.1124	2.3256	0.1217	2.2880	**0.1088**

Note: the boldface identifies the smallest MSE for each cases.

Table 5. Est. and MSE for HDW; $x \sim U(0,3)$ and $\boldsymbol{\theta} = (-2, -1.7, 1.5, -0.9, 2.2)$.

n (% zeros)	parameter	MLE Est.	MLE MSE	Bayes(Uniform) Est.	Bayes(Uniform) MSE	Bayes(Informative) Est.	Bayes(Informative) MSE
60 (53.42%)	α_0	-2.1534	0.6747	-2.2884	0.7586	-2.1564	**0.6553**
	α_1	-1.8188	0.1783	-1.8519	0.1956	-1.8059	**0.1706**
	γ_0	1.6103	0.4467	1.6887	0.5288	1.6180	**0.4433**
	γ_1	-0.9604	0.1508	-1.0091	0.1793	-0.9682	**0.1507**
	β	2.3663	0.2098	2.4265	0.2405	2.3452	**0.1963**
90 (53.43%)	α_0	-2.0686	0.3550	-2.1541	0.3741	-2.0696	**0.3410**
	α_1	-1.8069	0.1207	-1.8263	0.1302	-1.7978	**0.1170**
	γ_0	1.5864	0.2832	1.6548	0.3075	1.5958	**0.2733**
	γ_1	-0.9508	0.0909	-0.9920	0.0999	-0.9581	**0.0895**
	β	2.3158	0.1257	2.3524	0.1404	2.3016	**0.1202**
120 (53.27%)	α_0	-2.0397	0.2501	-2.1192	0.2549	-2.0474	**0.2376**
	α_1	-1.7726	0.0870	-1.7877	0.0899	-1.7659	**0.0835**
	γ_0	1.5244	0.2315	1.5625	0.2339	1.5281	**0.2235**
	γ_1	-0.9099	0.0744	-0.9342	0.0758	-0.9143	**0.0724**
	β	2.2788	0.0840	2.3127	0.0919	2.2713	**0.0805**
150 (53.34%)	α_0	-2.0515	0.2254	-2.1048	0.2188	-2.0540	**0.2101**
	α_1	-1.7692	0.0658	-1.7746	0.0651	-1.7617	**0.0627**
	γ_0	1.5129	0.1544	1.5516	0.1547	1.5185	**0.1468**
	γ_1	-0.9062	0.0521	-0.9307	0.0531	-0.9116	**0.0503**
	β	2.2784	0.0684	2.2967	0.0689	2.2692	**0.0638**
180 (53.39%)	α_0	-2.0234	0.1520	-2.0732	0.1531	-2.0275	**0.1433**
	α_1	-1.7525	0.0529	-1.7576	0.0527	-1.7473	**0.0507**
	γ_0	1.5246	0.1343	1.5494	0.1351	1.5267	**0.1276**
	γ_1	-0.9108	0.0425	-0.9272	0.0436	-0.9145	**0.0415**
	β	2.2531	0.0513	2.2703	0.0531	2.2467	**0.0486**

Note: the boldface identifies the smallest MSE for each cases.

Table 6. Est. and MSE for HDW; $x \sim N(2,1)$ and $\boldsymbol{\theta} = (-2,-1.7,1.5,-0.9,2.2)$.

n (% zeros)	parameter	MLE Est.	MLE MSE	Bayes(Uniform) Est.	Bayes(Uniform) MSE	Bayes(Informative) Est.	Bayes(Informative) MSE
60 (43.81%)	α_0	-2.0745	0.5337	-2.1820	0.5512	-2.0829	**0.4981**
	α_1	-1.8169	0.1451	-1.8543	0.1586	-1.8106	**0.1404**
	γ_0	1.6538	0.5772	1.7737	0.7278	1.6640	**0.5714**
	γ_1	-0.9805	**0.1384**	-1.0466	0.1773	-0.9911	0.1389
	β	2.3378	0.1624	2.3931	0.1865	2.3270	**0.1562**
90 (43.88%)	α_0	-2.0397	0.3163	-2.1181	0.3023	-2.0486	**0.2889**
	α_1	-1.7666	0.0791	-1.7920	0.0839	-1.7649	**0.0772**
	γ_0	1.6136	0.4164	1.6599	0.4298	1.6060	**0.3763**
	γ_1	-0.9529	0.0945	-0.9821	0.0976	-0.9544	**0.0862**
	β	2.2770	0.0886	2.3167	0.0958	2.2737	**0.0842**
120 (43.74%)	α_0	-2.0437	0.2414	-2.1204	0.2322	-2.0591	**0.2211**
	α_1	-1.7443	0.0542	-1.7631	0.0556	-1.7410	**0.0519**
	γ_0	1.5769	0.3045	1.6121	0.2869	1.5739	**0.2743**
	γ_1	-0.9385	0.0716	-0.9605	0.0693	-0.9409	**0.0665**
	β	2.2560	0.0599	2.2912	0.0636	2.2544	**0.0567**
150 (43.81%)	α_0	-1.9851	0.1874	-2.0526	**0.1597**	-2.0006	0.1602
	α_1	-1.7394	0.0456	-1.7569	0.0460	-1.7389	**0.0430**
	γ_0	1.5882	0.2516	1.6053	0.2340	1.5799	**0.2238**
	γ_1	-0.9410	0.0561	-0.9543	0.0530	-0.9410	**0.0512**
	β	2.2347	0.0493	2.2664	0.0485	2.2360	**0.0445**
180 (43.74%)	α_0	-1.9872	0.1527	-2.0484	0.1326	-2.0028	**0.1317**
	α_1	-1.7234	0.0381	-1.7371	0.0376	-1.7231	**0.0362**
	γ_0	1.5648	0.2176	1.5714	**0.1744**	1.5558	0.1771
	γ_1	-0.9311	0.0489	-0.9384	**0.0398**	-0.9301	0.0412
	β	2.2217	0.0429	2.2490	0.0414	2.2237	**0.0392**

Note: the boldface identifies the smallest MSE for each cases.

The results of these simulation studies show that all of estimators have monotonic behaviors according to the MSE, namely, when n increases, the estimated MSE values decrease. Almost all cases of the Bayes estimators with informative priors show the best performance in terms of the MSE except: Table 1 to Table 3 for the ZIDW regression model when $x \sim Ber(0.4)$, the MLE shows the best performance for parameters γ_0, γ_1 at $n = 60, 90$, and the Bayes estimator with uniform noninformative priors shows the best performance for parameter α_1 at $n = 180$, when $x \sim U(0,3)$, the MLE shows the best performance for parameters γ_0, γ_1 at $n = 60$, and the MLE and the Bayes estimator with informative priors shows the best performance for parameter γ_1 at $n = 90$, and when $x \sim N(2,1)$, the Bayes estimators with uniform noninformative priors shows the best performance for parameter γ_0, γ_1 at $n = 180$. Additionally, Table 4 to Table 6 for the HDW regression model when appearing $x \sim Ber(0.4)$, the MLE shows the best performance for parameters γ_0 at $n = 150,\ 180$ and γ_1 at $n = 120, 180$, and the Bayes estimator with uniform noninformative priors shows the best performance for parameters γ_0, γ_1 at $n = 90$ and α_0 at $n = 120$, and when concerning $x \sim N(2,1)$, the MLE shows the best performance for parameter γ_1 at $n = 60$, and the Bayes estimator with uniform noninformative priors shows the best performance for parameters α_0 at $n = 150$ and γ_0, γ_1 at $n = 180$. Moreover, it can be observed that the performance of the coefficient estimators from a uniform noninformative priors provided better than the MLE when showing the large sample sizes in cases of $x \sim N(2,1)$. Note that the performance of estimators has high sensitivity to small sample size and type of explanatory variable.

3.2 Application to real dataset

In this section, a real dataset is applied to show the performance of the Bayes estimators with informative priors for the ZIDW and HDW regression models and is compared with the two popular models for zero-inflated and hurdle models, the ZIP and ZINB regression models, and the HP and HNB regression models respectively. For the Poisson and negative binomial, the zero-inflated and hurdle models are applied by using the same configuration as with the certain approach. The state wildlife biologists gathered how many fish are being caught by fishermen at a state park; the dataset is available at https://stats.idre.ucla.edu/stat/data/fish.dat. This dataset has 250 groups of visitors that went to the park, and each group was questioned before leaving the park. The response variable is the number of fish that they caught and the three explanatory variables are whether or not they brought a camper to the park (camper), the number of people in the group (persons), and the number of children in the group (child). Moreover, the explanatory variable that affects the parameter q_i is the same as the explanatory variable that affects the parameter π_i, which is $z = x$. The response variable has 56.80% of zeros, which is in the case of excessive zeros data.

To confirms simulation study, this study calculated the parameter estimates (Est.) and the standard error (SE) of estimators from the MLE and the Bayesian estimation with uniform noninformative priors (Bayes(Uniform)) and informative priors (Bayes(Informative)) for the ZIDW and HDW regression models for each of the three cases of a simple explanatory variable camper, persons, and child which are reported in Table 7 and Table 8.

To demonstrate how the proposed Bayesian method under the informative priors can be used in practice, this study constructed the model in the simple regression model. It calculated the parameter estimates and the 95% highest posterior density (HPD) interval of the parameters with informative priors under the squared error loss function using the random walk Metropolis algorithm with $L = 10{,}000$ replicates and 10% of the chain for *burn-in*; $B = 1{,}000$ for the ZIP, ZINB, ZIDW and the HP, HNB, HDW regression models for each of the three cases of a simple explanatory variable camper, persons, and child which are reported in Table 9 and Table 10 respectively. Moreover, the three information criteria; the

Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the deviance information criterion (DIC), are reported in these tables.

The results from Table 7 and Table 8 show that all of the three cases of a simple explanatory variable camper, persons, and child, the performance of the Bayes estimators with informative priors for the ZIDW and HDW regression models is better than other methods in terms of the SE of the estimators. Besides, the result of the application to the fish data from the state wildlife biologists is close to the simulation results.

In addition, the results from Table 9 and Table 10 show that all of the three cases of a simple explanatory variable camper, persons, and child, the Bayes estimators with informative priors for the ZIDW and HDW regression models provided better fitting than both ZIP and ZINB models and HP and HNB models respectively, according to the lowest AIC, BIC and DIC values.

Table 7. Est. and SE for the ZIDW regression model.

Parameters	MLE		Bayes(Uniform)		Bayes(Informative)	
	Est.	SE	Est.	SE	Est.	SE
camper						
α_0	0.0748	0.1505	0.1062	0.1113	0.0658	**0.0934**
α_1	-0.4449	0.1621	-0.4469	0.1364	-0.4187	**0.0985**
γ_0	-2.2571	1.5754	-124.1471	49.1528	-2.9509	**1.2348**
γ_1	-2.9996	11.1311	-711.6784	171.9149	-9.8507	**7.1804**
β	0.4767	0.0380	0.4705	0.0334	0.4753	**0.0325**
persons						
α_0	0.6225	0.1906	0.5638	0.1866	0.6364	**0.1326**
α_1	-0.4538	0.1042	-0.3281	0.0915	-0.4622	**0.0778**
γ_0	-2.8735	1.6923	-12.7508	12.3856	-3.2492	**1.1151**
γ_1	0.5097	0.4092	-2.5387	3.3328	0.5600	**0.2663**
β	0.6240	0.0791	0.5348	**0.0552**	0.6237	0.0710
child						
α_0	-0.5514	0.1192	-0.5506	0.1084	-0.5550	**0.0825**
α_1	0.5379	0.1301	0.6193	**0.0965**	0.5061	0.0979
γ_0	-4.8773	1.6887	-113.3016	52.3060	-5.9873	**1.3262**
γ_1	3.7376	1.4286	5.9349	22.0883	2.9412	**0.8752**
β	0.5252	0.0382	0.5114	0.0353	0.5195	**0.0336**

Note: the boldface identifies the smallest SE for each cases.

Table 8. Est. and SE for the HDW regression model.

Parameters	MLE		Bayes(Uniform)		Bayes(Informative)	
	Est.	SE	Est.	SE	Est.	SE
camper						
α_0	0.5118	0.4043	0.5739	0.4084	0.5761	**0.2811**
α_1	-0.3333	0.2129	-0.3531	0.2357	-0.3794	**0.1369**
γ_0	0.7086	0.2096	0.6866	0.2384	0.7117	**0.1363**
γ_1	-0.7234	0.2667	-0.7130	0.2916	-0.7209	**0.1661**
β	0.3555	0.0917	0.3459	0.0980	0.3505	**0.0673**
persons						
α_0	1.1286	0.3354	1.0990	0.3527	1.1155	**0.2436**
α_1	-0.5145	0.0973	-0.5118	0.0972	-0.5167	**0.0639**
γ_0	0.7767	0.3239	0.7526	0.3322	0.7838	**0.1902**
γ_1	-0.1978	0.1161	-0.1868	0.1216	-0.2026	**0.0674**
β	0.5314	0.1019	0.5388	0.1021	0.5350	**0.0880**
child						
α_0	0.1480	0.4175	0.2712	0.5904	0.1722	**0.2912**
α_1	0.4341	0.1873	0.3148	0.2310	0.3831	**0.1332**
γ_0	-0.3842	0.1704	-0.3666	0.2265	-0.3843	**0.1139**
γ_1	1.1218	0.2060	1.0457	0.3555	1.1185	**0.1317**
β	0.3538	0.0924	0.3155	0.1890	0.3547	**0.0662**

Note: the boldface identifies the smallest SE for each cases.

Table 9. Parameter estimates and the 95% HPD intervals (in parentheses) for the ZIP, ZINB, ZIDW regression models.

Parameters	ZIP	ZINB	ZIDW
camper			
α_0	1.5139 (1.4188,1.6132)	0.4852 (0.1131,0.8760)	0.0658 (-0.1121,0.2451)
α_1	0.6858 (0.5803,0.7837)	1.0588 (0.6458,1.5008)	-0.4187* (-0.6203,-0.2234)
γ_0	0.7056 (0.4516,0.9825)	-3.4486 (-7.0760,-0.9232)	-2.9509* (-5.6694,-0.9952)
γ_1	-0.7139 (-1.0531,-0.3656)	-2.7336 (-7.0549,1.3302)	-9.8507 (-26.9134,0.8346)
r / β		0.2058 (0.1550,0.2673)	0.4753* (0.4160,0.5395)
AIC	2189.2686	927.2019	918.3932
BIC	2203.3544	944.8092	936.0005
DIC	2186.1540	922.7372	913.7076
persons			
α_0	-0.2607 (-0.4349,-0.0888)	-0.9956 (-1.3866,-0.5700)	0.6364* (0.3722,0.8837)
α_1	0.7410 (0.6893,0.7887)	0.8300 (0.6785,0.9835)	-0.4622* (-0.5992,-0.2927)
γ_0	0.4183 (-0.0482,0.8501)	-1.7332 (-2.8408,-0.6516)	-3.2492* (-5.8750,-1.6104)
γ_1	-0.0952 (-0.2589,0.0535)	0.1971 (-0.1017,0.4987)	0.5600* (0.0544,1.1445)
r / β		0.3964 (0.2371,0.6025)	0.6237* (0.4935,0.7726)
AIC	1858.6155	903.1250	901.7890
BIC	1872.7014	920.7323	919.3963
DIC	1856.1870	899.2652	898.7913
child			
α_0	2.2001 (2.1462,2.2555)	1.6436 (1.3943,1.8834)	-0.5550* (-0.7151,-0.3999)
α_1	-0.7116 (-0.8345,-0.5958)	-1.0413 (-1.4411,-0.6313)	0.5061* (0.2856,0.6884)
γ_0	-0.3764 (-0.5942,-0.1434)	-7.0574 (-12.7304,-3.5041)	-5.9873* (-8.9239,-3.6940)
γ_1	1.0508 (0.7554,1.3298)	3.5659 (1.5300,6.2092)	2.9412* (1.0978,4.5625)
r / β		0.2512 (0.1915,0.3229)	0.5195* (0.4528,0.5859)
AIC	2146.1583	893.7935	885.6576
BIC	2160.2442	911.4008	903.2649
DIC	2142.5286	889.5677	881.2140

Note: (*) denotes the 95% HPD interval for the ZIDW regression model does not contain zero (statistically significant).

Table 10. Parameter estimates and the 95% HPD intervals (in parentheses) for the HP, HNB, HDW regression models.

Parameters	HIP	HNB	HDW
camper			
α_0	1.5204	-2.9015	0.5761*
	(1.4195,1.6117)	(-4.1737,-1.9316)	(0.0105,1.0883)
α_1	0.6796	0.5078	-0.3794*
	(0.5716,0.7869)	(-0.3629,1.2973)	(-0.6689,-0.0890)
γ_0	0.7184	0.7459	0.7117*
	(0.4752,0.9793)	(0.4986,1.0152)	(0.4412,1.0073)
γ_1	-0.7334	-0.7672	-0.7209*
	(-1.0693,-0.3637)	(-1.1261,-0.4569)	(-1.0449,-0.4411)
r / β		0.0037	0.3505*
		(0.0007,0.0071)	(0.2345,0.5019)
AIC	2189.2618	921.7569	916.2778
BIC	2203.3476	939.3642	933.8851
DIC	2186.1235	917.2394	912.3456
persons			
α_0	-0.2933	-2.0637	1.1155*
	(-0.4822,-0.1000)	(-2.9861,-1.2358)	(0.6555,1.5650)
α_1	0.7494	0.9647	-0.5167*
	(0.6962,0.8020)	(0.7706,1.1492)	(-0.6423,-0.3817)
γ_0	0.8252	0.7819	0.7838*
	(0.5066,1.2221)	(0.4213,1.1880)	(0.3985,1.1525)
γ_1	-0.2170	-0.2026	-0.2026*
	(-0.3612,-0.0863)	(-0.3331,-0.0529)	(-0.3230,-0.0611)
r / β		0.1687	0.5350*
		(0.0339,0.4146)	(0.3780,0.6912)
AIC	1858.3974	899.5903	896.4451
BIC	1872.4832	917.1976	914.0524
DIC	1855.6746	895.5030	892.9249
child			
α_0	2.1986	-1.9529	0.1722
	(2.1495,2.2506)	(-3.4576,-0.6980)	(-0.3874,0.7834)
α_1	-0.7063	-0.9898	0.3831*
	(-0.8313,-0.5889)	(-1.5293,-0.4536)	(0.1309,0.6472)
γ_0	-0.3861	-0.3978	-0.3843*
	(-0.6141,-0.1512)	(-0.6237,-0.1646)	(-0.6231,-0.1694)
γ_1	1.1193	1.1408	1.1185*
	(0.8561,1.3904)	(0.8730,1.4269)	(0.8865,1.3792)
r / β		0.0058	0.3547*
		(0.0008,0.0177)	(0.2339,0.4880)
AIC	2146.8047	889.1705	884.7411
BIC	2160.8905	906.7778	902.3485
DIC	2143.0773	882.4158	880.6896

Note: (*) denotes the 95% HPD interval for the ZIDW regression model does not contain zero (statistically significant).

4. Conclusion

In this article, it considers the classical and Bayesian inference for the ZIDW and HDW regression models where the parameters q and π are related to explanatory variables via the log-log and logit links respectively.

Moreover, this study chooses the random walk Metropolis algorithm to estimate the Bayes estimators with uniform noninformative priors and informative priors.

The results of the simulation showed that as n increases the MSE decreases for all methods, indicating that the estimators are consistent. The Bayes estimators with informative priors for the parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β are more appropriate for both the ZIDW and HDW regression models than other methods in terms of the MSE. Moreover, the results of an application to the fish data from the state wildlife biologists revealed that the Bayes estimators with informative priors for parameters $\boldsymbol{\alpha}$, $\boldsymbol{\gamma}$, and β for the ZIDW and HDW regression models show the best fitting model in terms of the AIC, BIC, and DIC. These results confirm that using the Bayesian method under informative prior distributions for the ZIDW and HDW regression models work alternatively better than the Poisson and negative binomial.

Hence, it was recommended that the Bayesian regression model be under this informative prior where the data is fitted ZIDW and HDW. However, there are some computational challenges to be faced while implementing the Bayesian approach which is the selection of hyperparameters' values that may affect the parameter estimates. Future research will explore other link functions on parameters and construct the censored response with too many zero counts.

Acknowledgements

The research of Dr. Duangsaphon is currently supported by the Thammasat University Research Unit in Theoretical and Computational Statistics. The authors gratefully acknowledge financial support from Thammasat University.

References

[1] Frome EL. The analysis of rates using Poisson regression models. Biometrics. 1983;39(3):665-74.

[2] Frome EL, Checkoway H. Epidemiologic programs for computers and calculators. Use of Poisson regression models in estimating incidence rates and ratios. Am J Epidemiol. 1985;121(2):309-23.

[3] Lovett A, Flowerdew R. Analysis of count data using Poisson regression. Prof Geogr. 1989;41(2):190-8.

[4] Hutchinson MK, Holtman MC. Analysis of count Data using Poisson regression. Res Nurs Health. 2005;28(5),408-18.

[5] Gardner W, Mulvey EP, Shaw EC. Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. Psychol Bull. 1995;118(3):392-404.

[6] Allison PD, Waterman RP. Fixed-effects negative binomial regression models. Sociol Methodol. 2002;32(1):247-65.

[7] Liu H, Davidson RA, Rosowsky DV, Stedinger JR. Negative binomial regression of electric power outages in hurricanes. J Infrastruct Syst. 2005;11(4): 258-67.

[8] Shmueli G, Minka TP, Kadane JB, Borle S, Boatwright P. A useful distribution for fitting discrete data: revival of the Conway-Maxwell-Poisson distribution. J R Stat Soc Ser C Appl Stat. 2005;54(1):127-42.

[9] Lord D, Guikema SD, Geedipally SR. Application of the Conway-Maxwell-Poisson generalized linear model for analyzing motor vehicle crashes. Accid Anal Prev. 2008;40(3):1123-34.

[10] Lord D, Geedipally SR, Guikema SD. Extension of the application of Conway-Maxwell-Poisson models: analyzing traffic crash data exhibiting underdispersion. Risk Anal. 2010;30(8): 1268-76.

[11] Kalktawi HS. Discrete Weibull regression model for count data. [Ph.D. thesis]. London: Brunel University London; 2017.

[12] Kalktawi HS, Vinciotti V, Yu K. A simple and adaptive dispersion regression Model for count data. Entropy. 2018;20(2):142.

[13] Haselimashhadi H, Vinciotti V, Yu K. A novel Bayesian regression model for counts with an application to health data. J Appl Stat. 2018;45(6):1085-105.

[14] Barbiero A. Discrete Weibull regression for modelling football outcomes. Int J Bus Intell Data Min. 2020;17(1):76-100.

[15] Collins K, Waititu A, Wanjoya A. Discrete Weibull and artificial neural network models in modelling over-dispersed count data. Int J Data Sci Anal. 2020;6(5):153-62.

[16] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. Biometrika. 1970;57(1):97-109.

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

[18] Zorn CJW. An analytic and empirical examination of zero-inflated and hurdle Poisson specifications. Sociol Methods Res. 1998;26(3):368-400.

[19] Hall DB. Zero-inflated Poisson and binomial regression with random effects: a case study. Biometrics. 2000;56(4): 1030-9.

[20] Saffari SE, Adnan R, Greene W. Parameter estimation on hurdle Poisson regression model with censored data. J Teknol. 2013;57(1):189-98.

[21] Yau KKW, Wang K, Lee AH. Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros. Biom J. 2003;45(4):437-52.

[22] Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. J Biopharm Stat. 2006;16(4):463-81.

[23] Minami M, Lennert-Cody CE, Gao W, Román-Verdesoto M. Modeling shark bycatch: The zero-inflated negative binomial regression model with smoothing. Fish Res. 2007;84(2):210-21.

[24] Garay AM, Hashimoto EM, Ortega EMM, Lachos VH. On estimation and influence diagnostics for zero-inflated negative binomial regression models. Comput Stat Data Anal. 2011;55(3):1304-18.

[25] Saffari SE, Adnan R, Greene W. Hurdle negative binomial regression model with right censored count data. Sort-Stat Oper Res T. 2012;36(2):181-94.

[26] Blasco-Moreno A, Pérez-Casany M, Puig P, Morante M, Castells E. What does a zero mean? Understanding false, random and structural zeros in ecology. Methods Ecol Evol. 2019;10(7):949-59.

[27] Obeidat M, Al-Nasser A, Al-Omari AI. Estimation of Generalized Gompertz Distribution Parameters under Ranked-Set Sampling. J Probab Stat. 2020;ID 7362657:14.

[28] Sunday AO, Samson OA, Modupe OD. Bayesian optimization for parameter of discrete Weibull regression. J Adv Math Comput Sci. 2019;34(6):1-13.

[29] Gelman A, Jakulin A, Pittau MG, Su Y. A weakly informative default prior distribution for logistic and other regression models. Ann Appl Stat. 2008;2(4):1360-83.

[30] Fu S. Hierarchical Bayesian LASSO for a negative binomial regression. J Stat Comput Simul. 2016; 86(11):2182-203.

[31] Chanialidis C, Evers L, Neocleous T, Nobile A. Efficient Bayesian inference for COM-Poisson regression models. Stat Comput. 2018;28(1):595-608.

[32] Aslam M, Kazmi SMA, Ahmad I, Shah SH. Bayesian estimation for parameters of the Weibull distribution. Sci Int (Lahore). 2014;26(5):1915-20.

[33] Chacko M, Mohan R. Bayesian analysis of Weibull distribution based on progressive type-II censored competing risks data with binomial removals. Comput Stat. 2019;34(4):233-52.

[34] Cameron AC, Trivedi PK. Regression Analysis of Count Data. 2nd ed. Cambridge University Press; 2013.

[35] Cragg JG. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica. 1971;39(5):829-44.

[36] Mullahy J. Specification and testing of some modified count data models. J Econom. 1986;33(3):341-65.

[37] Hilbe JM. COUNT: Functions, Data and Code for Count Data. R package version 1.3.4.; 2016 [Internet]. [cited 2020 Feb 15]. Available from: https://CRAN.R-project.org/package=COUNT.

[38] Lee ET, Wang JW. Statistical methods for survival data analysis. 3rd ed. Hoboken: John Wiley & Sons, Inc.; 2013. p. 259-63.

[39] Nakagawa T, Osaki S. The discrete Weibull distribution. IEEE Trans Reliab. 1975;R-24(5):300-01.

[40] Nash JC. Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation. 2nd ed. Bristol: Adam Hilger; 1990.

[41] R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria; 2020 [Internet]. [cited 2020 Feb 15]. Available from: https://www.R-project.org/.

[42] Vinciotti V. DWreg: Parametric Regression for Discrete Response. R package version 2.0.; 2016 [Internet]. [cited 2020 Feb 15]. Available from: https://CRAN.R-project.org/package=DWreg.