



Research article

Bayesian estimation for median discrete Weibull regression model

Monthira Duangsaphon, Sukit Sokampang and Kannat Na Bangchang*

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

* **Correspondence:** Email: kannat@mathstat.sci.tu.ac.th.

Abstract: The discrete Weibull model can be adapted to capture different levels of dispersion in the count data. This paper takes into account the direct relationship between explanatory variables and the median of discrete Weibull response variable. Additionally, it provides the Bayesian estimate of the discrete Weibull regression model using the random walk Metropolis algorithm. The prior distributions of the coefficient predictors were carried out based on the uniform non-informative, normal and Laplace distributions. The performance of the Bayes estimators was also compared with the maximum likelihood estimator in terms of the mean square error and the coverage probability through the Monte Carlo simulation study. Meanwhile, a real data set was analyzed to show how the proposed model and the methods work in practice.

Keywords: Bayesian estimation; deviance information criterion; discrete count data; generalized linear models; random walk Metropolis algorithm

Mathematics Subject Classification: 62F15, 62J99

1. Introduction

Generalized linear models are extensions of the linear regression model to avoid the selection of the normality response and linearity imposed by the linear regression model, which is impossible for binary or count responses. Regression for count data is widely performed by models, such as Poisson [1], negative binomial [2] and zero-inflated regressions [3–5]. The well-known property of a Poisson distribution shows its mean that is equal to the variance. This situation is often unrealistic as the distribution of counts tends to have a variance not equal to its mean. When data handles over-dispersion, a negative binomial distribution is utilized to model count variables. Zero-inflated models are used to model count data that have many zero counts. Moreover, the Conway-Maxwell Poisson distribution is used to deal with under-dispersion and over-dispersion [2]. Also, the discrete Weibull distribution is examined to handle under-dispersion and over-dispersion discrete data. This

model was first introduced by Nakagawa and Osaki [6]. The motivation for considering the discrete Weibull distribution stems from the vital role that the continuous Weibull distribution plays in the survival analysis and failure time study. Similarly, the continuous Weibull distribution is widely used in probabilistic modeling and a fatigue life prediction model [7,8]. However, there are many challenging things about this distribution that has yet to be proposed.

The inference for parameters of the discrete Weibull regression model has been investigated in a few studies based on a parameter affecting zero observation related to the explanatory variables through the log-log and logit link functions. Kalktawi [9] and Englehardt and Li [10] showed how a discrete Weibull regression model can be adapted to address over-dispersion and under-dispersion via the log-log link function. Moreover, Klakattawi et al. [11] proposed an ability to adapt in a simple way to different types of dispersions: Over-dispersion, under-dispersion and covariate-specific dispersion. Peluso and Vinciotti [12] conducted a simulation study linking two parameters to inspect a discrete Weibull regression model's level of flexibility.

The maximum likelihood estimation of parameters is valid for an asymptotically large sample size of data [13]. One of the most common problems occurring in the count regression model is the maximum likelihood estimates that become unstable with larger standard errors of the estimates that affect statistical inference when insufficiently large sample sizes manifest. To overcome the problem, various alternatives to the maximum likelihood estimation have been proposed, and the Bayesian estimation is one of them. However, the Bayes estimators depend on the prior distributions of the parameters in the model. Many researchers at different periods of time worked in this area of research proposed different prior distributions in the count regression model (see, for example, Haselimashhadi et al. [14], Gelman et al. [15], Fu [16] and Chanialidis et al. [17]). Recently, Chaiprasithikul and Duangsaphon [18,19] proposed the Bayesian estimation for censored data and the zero-inflated and hurdle discrete Weibull regression models via the log-log link function for a parameter that affects zero observation. Uniform non-informative and normal prior distributions were used to account for the regression coefficients. It was demonstrated that this suggestion performs well when applied to real datasets and in simulation studies. Alternatively, a regression structure for the discrete Weibull model through the median link function was proposed by Kalktawi [9]. In addition, there are many works that have developed the Bayesian estimation procedure for the reliability and life testing experiments (see, for example, Ahmadini et al. [20] and Okasha et al. [21]).

In the present article, the Bayesian estimation is examined, based on the random walk Metropolis algorithm for the median discrete Weibull regression model under the three different prior distributions: Uniform non-informative, normal and Laplace prior distributions. A simulation study is conducted to compare the performance of three different prior distributions and the maximum likelihood estimation in both under-dispersion and over-dispersion cases. Moreover, a real dataset is analyzed to see how the model works in practice.

The rest of paper is organized as follows. An overview of the median discrete Weibull regression model is presented in section two, along with the maximum likelihood estimation, Bayesian estimation, simulation study and real data analysis. In section three, results and discussion are explained. Finally, this paper is concluded in section four.

2. Materials and methods

2.1. The median discrete Weibull regression model

A discrete Weibull distribution (type one) and some properties were introduced by Nakagawa and Osaki [6]. The cumulative distribution function and the probability mass function of a random variable are given by

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

and

$$p_Y(y; q, \beta) = \begin{cases} q^{y^\beta} - q^{(y+1)^\beta}; & y = 0, 1, \dots \\ 0; & \text{otherwise,} \end{cases} \quad (2)$$

respectively, where $0 < q < 1$ and $\beta > 0$ are the shape parameters. When $y = 0$, the parameter $q = 1 - p_Y(0; q, \beta)$, which is the probability of Y more than zero. In other words, when q is small, an excessive zero case occurs. Parameter β indicates the skewness and controls the range of values of Y . Moreover, Kalktawi [9] proposed the parameter β reflects the dispersion of data through the numerical analyses; if $0 < \beta \leq 1$, the data is over-dispersion, if $\beta \geq 2$, the data is under-dispersion and if $1 < \beta < 2$, the data can be either over-dispersed or under-dispersed depending on the value of q . Thus, the discrete Weibull distribution is suitable for both over-dispersion and under-dispersion. Meanwhile, there are the special cases of a discrete Weibull distribution: Geometric distribution, the discrete Exponential distribution (see Sato et al. [22]), the discrete Rayleigh distribution (see Roy [23]) and the Bernoulli distribution.

The mean and variance of a discrete Weibull distribution are no closed-form expressions, but the numerical approximations can be obtained (see Barbiero [24]). Another property of a discrete Weibull distribution is the quantile function $Q(\tau)$. Let's say the τ -th ($0 < \tau < 1$) quantile that is the smallest value of y for which $F(y) \geq \tau$. The quantile $Q(\tau)$ has a closed-form expression as given by

$$Q(\tau) = \left\lceil \left[\left(\frac{\ln(1-\tau)}{\ln q} \right)^{1/\beta} - 1 \right] \right\rceil. \quad (3)$$

The quantile formula provided in Eq (3) can be applied. The median for discrete distributions can be defined as any value of y that $F(y) \geq 0.5$, then the median can be easily obtained from the closed form as given by

$$M = \left\lceil \left(-\frac{\ln 2}{\ln q} \right)^{1/\beta} - 1 \right\rceil. \quad (4)$$

The presence of regression analysis for count data is a statistical process to measure the relationship between a count variable and one or more explanatory variables. Klakattawi et al. [11] showed how a discrete Weibull regression model can be adapted to address over-dispersion and under-dispersion via the log-log link function to the parameter q . Moreover, Haselimashhadi et al. [14] applied the logit link function to the parameter q , which is commonly used in classification problems for probabilities that are

bounded between zero and one. Furthermore, they performed β that depends on explanatory variables through the log link function. The proposed discrete Weibull regression model, unlike generalized linear models in which the conditional mean is central to the interpretation, has the advantage that the conditional quantiles can be easily extracted from the fitted model. Moreover, the regression coefficients can be easily interpreted in terms of changes in the conditional median. According to the median, it has a closed-form expression and is more appropriate than the mean because of the skewness and outliers commonly found for counting data. Kalktawi [9] performed the median link function to a discrete Weibull regression model and used the maximum likelihood method for parameters estimation.

This study determines $Y_i, i = 1, 2, \dots, n$ as a count response variable, which takes only the non-negative integer values with the k explanatory variables, $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ik})$ and a vector composed of regression coefficients as $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k)'$. It is assumed that the parameter $q_i = q(\mathbf{x}_i)$ is related to k explanatory variables \mathbf{x}_i via the median link function as follows

$$g(M_i) = \mathbf{x}_i \alpha = \alpha_0 + \alpha_1 x_{i1} + \dots + \alpha_k x_{ik}. \quad (5)$$

In the context, it is useful to assume that

$$g(M_i) = \ln(M_i + 1). \quad (6)$$

Thus,

$$M_i + 1 = e^{\mathbf{x}_i \alpha}. \quad (7)$$

Equation (7) is substituted to Eq (4); hence, the parameter q_i can be obtained as

$$q(\mathbf{x}_i) = e^{\left(\frac{-\ln 2}{e^{\beta \mathbf{x}_i \alpha}}\right)}. \quad (8)$$

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

$$p_Y(y_i | \mathbf{x}_i) = \begin{cases} e^{-(\ln 2)\left(\frac{y_i}{e^{\mathbf{x}_i \alpha}}\right)^\beta} - e^{-(\ln 2)\left(\frac{y_i+1}{e^{\mathbf{x}_i \alpha}}\right)^\beta}; & y = 0, 1, \dots \\ 0; & \text{otherwise.} \end{cases} \quad (9)$$

2.2. Maximum likelihood estimation

In this section, the maximum likelihood estimation for the discrete Weibull regression model is performed by linking only parameter q . The likelihood function of the median discrete Weibull regression is given by

$$L(\alpha, \beta | \mathbf{y}, \mathbf{x}) = \prod_{i=1}^n \left(e^{-(\ln 2)\left(\frac{y_i}{e^{\mathbf{x}_i \alpha}}\right)^\beta} - e^{-(\ln 2)\left(\frac{y_i+1}{e^{\mathbf{x}_i \alpha}}\right)^\beta} \right). \quad (10)$$

The log-likelihood function of the discrete Weibull regression model via median is given by

$$l(\alpha, \beta | \mathbf{y}, \mathbf{x}) = \sum_{i=1}^n \ln \left(e^{-(\ln 2)\left(\frac{y_i}{e^{\mathbf{x}_i \alpha}}\right)^\beta} - e^{-(\ln 2)\left(\frac{y_i+1}{e^{\mathbf{x}_i \alpha}}\right)^\beta} \right). \quad (11)$$

The maximum likelihood estimation of the parameters is obtained by setting the first partial derivatives of the log-likelihood function with respect to each unknown parameter equal to zero;

$$\frac{\partial l}{\partial \alpha_j} = \sum_{i=1}^n \frac{(\ln 2)\beta x_{ij} e^{-\mathbf{x}_i \boldsymbol{\alpha}}}{w_i(\boldsymbol{\alpha}, \beta)} \left(y_i \left(\frac{y_i}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^{\beta-1} e^{-(\ln 2) \left(\frac{y_i}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta} - (y_i + 1) \left(\frac{y_i + 1}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^{\beta-1} e^{-(\ln 2) \left(\frac{y_i + 1}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta} \right),$$

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^n \frac{-\ln 2}{w_i(\boldsymbol{\alpha}, \beta)} \left(\left(\frac{y_i}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta e^{-(\ln 2) \left(\frac{y_i}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta} [\ln y_i - \mathbf{x}_i \boldsymbol{\alpha}] - \left(\frac{y_i + 1}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta e^{-(\ln 2) \left(\frac{y_i + 1}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta} [\ln(y_i + 1) - \mathbf{x}_i \boldsymbol{\alpha}] \right),$$

where $w_i(\boldsymbol{\alpha}, \beta) = e^{-(\ln 2) \left(\frac{y_i}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta} - e^{-(\ln 2) \left(\frac{y_i + 1}{e^{\mathbf{x}_i \boldsymbol{\alpha}}} \right)^\beta}$.

The maximum likelihood estimators do not have a closed form solution because of the complex form of the likelihood equations. It is very difficult to prove that the solution to the normal equations gives a global maximum. Therefore, the maximum likelihood estimators are estimated by using the numerical method applied in the function *optim()* from package *stats* in R language, which minimizes the negative log-likelihood function of the median discrete Weibull regression model.

Let $I(\boldsymbol{\alpha}, \beta)$ be the observed Fisher’s information matrix for the $(k + 2) \times (k + 2)$ unknown parameters that contain negative of the second derivative of the log-likelihood function; hence, the variance-covariance matrix is the inverse of the observed Fisher’s information matrix,

$$\Sigma = I^{-1}(\boldsymbol{\alpha}, \beta). \tag{14}$$

The maximum likelihood estimators are substituted, thus resulting in an estimator of Σ denoted by $\hat{\Sigma}$,

$$\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{\alpha_0 \alpha_0} & \dots & \hat{\sigma}_{\alpha_0 \alpha_k} & \hat{\sigma}_{\alpha_0 \beta} \\ \vdots & \ddots & \vdots & \vdots \\ \hat{\sigma}_{\alpha_0 \alpha_k} & \dots & \hat{\sigma}_{\alpha_k \alpha_k} & \hat{\sigma}_{\alpha_k \beta} \\ \hat{\sigma}_{\alpha_0 \beta} & \dots & \hat{\sigma}_{\alpha_k \beta} & \hat{\sigma}_{\beta \beta} \end{pmatrix}. \tag{15}$$

This matrix can be obtained by inverting the Hessian matrix from the function *hessian()* in R language. The Hessian matrix contains the second derivative of the negative log-likelihood, i.e., moreover, the Hessian matrix is the observed Fisher’s information matrix.

According to the parameter inferences performed using the maximum likelihood method, under some regularity conditions [25], these estimators enjoy standard asymptotic properties. Thus, by the asymptotic normality of maximum likelihood estimators, the $100(1 - \alpha)\%$ confidence intervals for parameters $\alpha_j, j = 0, 1, 2, \dots, k$ and β , respectively are

$$\hat{\alpha}_j \pm z_{\alpha/2} \sqrt{\hat{\sigma}_{\alpha_j \alpha_j}} \text{ and } \hat{\beta} \pm z_{\alpha/2} \sqrt{\hat{\sigma}_{\beta \beta}}, \tag{16}$$

where $z_{\alpha/2}$ is the upper $\alpha/2 - th$ quantile of the standard normal distribution.

2.3. Bayesian estimation

2.3.1. Random walk Metropolis algorithm

The Metropolis-Hastings (MH) algorithm is the most popular example of a Markov chain Monte Carlo (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 the 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 MH algorithm is converging to the target distribution itself. For more details on the MH algorithm, see Hastings [26] and Gilks et al. [27].

Given $\mathbf{y} = (y_1, y_2, \dots, y_n)$ is the vector of the observed values of a random sample Y_1, Y_2, \dots, Y_n , let $p(\boldsymbol{\theta}|\mathbf{y})$ be the target distribution, while $\boldsymbol{\theta}$ is the vector of current state values (parameters) and $\boldsymbol{\theta}^*$ is the proposal value generated from the proposal distribution $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$. Then, the proposal value $\boldsymbol{\theta}^*$ is accepted with the probability $p = \min(1, R_{\boldsymbol{\theta}})$, where

$$R_{\boldsymbol{\theta}} = \frac{p(\boldsymbol{\theta}^*|\mathbf{y})}{p(\boldsymbol{\theta}|\mathbf{y})} \times \frac{q(\boldsymbol{\theta}|\boldsymbol{\theta}^*)}{q(\boldsymbol{\theta}^*|\boldsymbol{\theta})}.$$

The iterative steps of the MH algorithm can be described as follows

Step 1: Initialize the parameter $\boldsymbol{\theta}^{(0)}$ for the algorithm.

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

a. Generate $\boldsymbol{\theta}^* \sim q(\boldsymbol{\theta}^*|\boldsymbol{\theta}^{(l-1)})$.

b. Calculate $p = \min(1, R_{\boldsymbol{\theta}})$.

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

If $u \leq p$, 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 $1-p$.

A random walk Metropolis algorithm is a special case of the MH algorithm. 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 $p = \min(1, R_{\boldsymbol{\theta}})$, where

$$R_{\boldsymbol{\theta}} = \frac{p(\boldsymbol{\theta}^*|\mathbf{y})}{p(\boldsymbol{\theta}|\mathbf{y})}.$$

The algorithm of random walk Metropolis can be summarized followed by the above steps with adjusting Step 2. Generate random error ϵ from a multivariate normal distribution with a zero-mean vector and variance-covariance Σ .

2.3.2. Bayesian estimation for median discrete Weibull regression model

This section performs the Bayes estimators for the median discrete Weibull regression model based on three schemes of prior distributions as follows:

i) Uniform non-informative prior distribution: If no prior information is available, a default flat prior can be resorted to, then it is easy to focus on the uniform non-informative prior distribution. The following the prior distributions are

$$\pi(\alpha_j) \propto 1, j = 0, 1, \dots, k, \text{ and } \pi(\beta) \propto 1.$$

ii) Normal prior distribution: As stated earlier, the possible values of α_j are real numbers, which corresponds to the possible values of a normal distribution; this study selects the prior distribution

of α_j ; that is, a normal distribution with the hyperparameters as $(\mu_{\alpha_j}, \sigma_{\alpha_j}^2)$, $j = 0, 1, \dots, k$. For parameter β , this study selects the prior distribution that is Gamma distribution with the hyperparameters as (a, b) . The following prior distributions are

$$\pi(\alpha_j) = \frac{1}{\sqrt{2\pi\sigma_{\alpha_j}^2}} e^{-\frac{1}{2\sigma_{\alpha_j}^2}(\alpha_j - \mu_{\alpha_j})^2}, \mu_{\alpha_j} \in R, \sigma_{\alpha_j}^2 > 0, j = 0, 1, 2, \dots, k$$

and

$$\pi(\beta) = \frac{1}{b^a \Gamma(a)} \beta^{a-1} e^{-\beta/b}, a, b > 0.$$

iii) Laplace prior distribution: If prior information is available, this study can perform the informative prior distribution that should include all possible values of parameter. The possible values of α_j are real numbers which corresponds to the possible values of a Laplace distribution; it selects the prior distribution of α_j ; that is, a Laplace distribution with the hyperparameters as $(0, 1/\lambda)$. Similarly, it selects the prior distribution, which is Gamma distribution with the hyperparameters as (a, b) . The following prior distributions are

$$\pi(\alpha_j) = \frac{\lambda}{2} e^{-\lambda|\alpha_j|}, \lambda > 0, j = 0, 1, 2, \dots, k,$$

and

$$\pi(\beta) = \frac{1}{b^a \Gamma(a)} \beta^{a-1} e^{-\beta/b}, a, b > 0.$$

The joint prior distributions of the parameters α and β under the independence assumption is

$$\pi(\theta) = \pi(\alpha_0) \dots \pi(\alpha_k) \pi(\beta), \quad (17)$$

where $\theta = (\alpha_0, \alpha_1, \dots, \alpha_k, \beta)$.

The choice of the hyperparameters' values is generally modified by available information of dataset to improve the Bayes estimators. For example, it fixes the hyperparameters' values of α_j , $j = 0, 1, 2, \dots, k$ of normal prior distribution with mean zero and high variance. For Laplace prior distribution, it fixes the hyperparameters' values of α_j , $j = 0, 1, 2, \dots, k$ with some $\lambda > 0$. In addition, the hyperparameters' values of β are considered by the maximum likelihood estimator of β with the mean of Gamma distribution. The joint posterior density function of the parameters α and β can be written as:

$$p(\theta|\mathbf{y}, \mathbf{x}) = \frac{L(\theta|\mathbf{y}, \mathbf{x})\pi(\theta)}{\int \int \dots \int L(\theta|\mathbf{y}, \mathbf{x})\pi(\theta) d\alpha_0 \dots d\alpha_k d\beta} \propto L(\theta|\mathbf{y}, \mathbf{x})\pi(\theta), \quad (18)$$

where $L(\theta|\mathbf{y}, \mathbf{x})$ is the likelihood function of the median discrete Weibull regression model in Eq (10).

The Bayes estimator of each parameter under the squared error loss function is the expected value of each parameter under the joint posterior density function. Therefore, the Bayes estimators are given by

$$\hat{\alpha}_j = \int \int \dots \int \alpha_j p(\theta|\mathbf{y}, \mathbf{x}) d\alpha_0 \dots d\alpha_k d\beta \quad (19)$$

and

$$\hat{\beta} = \int \int \dots \int \beta p(\boldsymbol{\theta}|\mathbf{y}, \mathbf{x}) d\alpha_0 \dots d\alpha_k d\beta, \quad (20)$$

where $j = 0, 1, 2, \dots, k$.

A difficulty to the implementation of Bayesian procedure is that of obtaining the posterior distribution. The process often requires the integration which is very difficult to calculate especially when dealing with complex and high-dimensional models. In such a situation, MH algorithms are highly helpful in this case to model deviations from the posterior density and generate accurate approximations [26,27].

Since the integral in Eqs (19) and (20) does not have a closed form, this study chose the random walk MH algorithm to estimate the Bayes estimators. It also determines the joint posterior density function of the parameters $\boldsymbol{\alpha}$ and β in Eq (18) as the target distribution, while $\boldsymbol{\theta}$ is the current state value and $\boldsymbol{\theta}^*$ is the proposal value generated from the proposal distribution $q(\boldsymbol{\theta}^*|\boldsymbol{\theta})$. Then, the proposal value $\boldsymbol{\theta}^*$ is accepted with the probability $p = \min(1, R_{\boldsymbol{\theta}})$, where

$$R_{\boldsymbol{\theta}} = \frac{L(\boldsymbol{\theta}^*|\mathbf{y}, \mathbf{x})\pi(\boldsymbol{\theta}^*)}{L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{x})\pi(\boldsymbol{\theta})} \times \frac{q(\boldsymbol{\theta}|\boldsymbol{\theta}^*)}{q(\boldsymbol{\theta}^*|\boldsymbol{\theta})}. \quad (21)$$

For 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 $p = \min(1, R_{\boldsymbol{\theta}})$, where

$$R_{\boldsymbol{\theta}} = \frac{L(\boldsymbol{\theta}^*|\mathbf{y}, \mathbf{x})\pi(\boldsymbol{\theta}^*)}{L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{x})\pi(\boldsymbol{\theta})}. \quad (22)$$

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

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

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

a. Generate random error vector $\boldsymbol{\epsilon}$ 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; $\boldsymbol{\epsilon} \sim N(\boldsymbol{\mu} = 0, \boldsymbol{\Sigma} = \text{diag}(I^{-1}(\boldsymbol{\theta})))$.

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

b. Calculate $p = \min(1, R_{\boldsymbol{\theta}})$ where $R_{\boldsymbol{\theta}} = \frac{L(\boldsymbol{\theta}^*|\mathbf{y}, \mathbf{x})\pi(\boldsymbol{\theta}^*)}{L(\boldsymbol{\theta}|\mathbf{y}, \mathbf{x})\pi(\boldsymbol{\theta})}$.

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

If $u \leq p$, 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 $\boldsymbol{\alpha}$, and β from the average of the generated values as given by

$$\hat{\theta}_{Bayes} = \frac{1}{L - B} \sum_{l=B+1}^L \theta^{(l)}, \quad (23)$$

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

The construction of the highest posterior density (HPD) credible intervals of the parameters α_j , $j =$

$0, 1, 2, \dots, k$, and β follows the Monte Carlo procedure. Given an MCMC sample $\theta^{(l)}, l = B + 1, B + 2, \dots, L$, the HPD interval for θ can be shown as follows:

Step 1: Sort $\theta^{(l)}, l = B + 1, B + 2, \dots, L$ to obtain the ordered value

$$\theta_{(1)} \leq \theta_{(2)} \leq \dots \leq \theta_{(L-B)}.$$

Step 2: Compute the $100(1 - \alpha)\%$ HPD credible intervals

$$R_i(L - B) = (\theta_{(i)}, \theta_{(i + [(1 - \alpha)(L - B)])}), i = 1, 2, \dots, (L - B) - [(1 - \alpha)(L - B)], \quad (24)$$

where $[(1 - \alpha)(L - B)]$ is the integer part of $(1 - \alpha)(L - B)$ and θ is a parameter in vector $\boldsymbol{\theta} = (\alpha, \beta)$.

2.4. Simulation study

In this section, the Monte Carlo simulation is conducted to assess and compare the performance of the Bayesian estimation via the random walk Metropolis algorithm for the median discrete Weibull regression model under difference three prior distributions for the regression parameters: Uniform non-informative prior (Bayes(U)), normal prior (Bayes(N)) and Laplace prior (Bayes(L)). Moreover, the MLE is considered. The various selected sample sizes (n) are 50, 100 and 200. The three explanatory variables are considered: A standard normal distribution ($x_1 \sim N(0, 1)$), a uniform distribution that lies between -0.3 and 0.3 ($x_2 \sim U(-0.3, 0.3)$) and a Bernoulli distribution with probability of success 0.4 ($x_3 \sim Ber(0.4)$). In particular, this study selects the regression parameters to take values $(\alpha_0, \alpha_1, \alpha_2, \alpha_3) = (1.5, 0.4, -0.2, 0.8)$ and $\beta = 0.9$ for over-dispersion, $\beta = 2.5$ for under-dispersion and $\beta = 1.6$ for either over-dispersed or under-dispersed, depending on the value of q .

The parameters are estimated by using the numerical method. In this paper, the Nelder-Mead method in the function *optim()* from package *stats* in R is applied to estimate parameters of the median discrete Weibull regression model. For Bayesian estimation, it fixes the hyperparameters' values of α_j and $j = 0, 1, 2, 3$ of normal prior distribution with mean zero and variance 100^2 and the hyperparameters' values of β as one and the maximum likelihood estimator. In addition, it fixes the hyperparameters' values of α_j and $j = 0, 1, 2, 3$ of Laplace prior distribution as 0.5 and the hyperparameters' values of β as one and the maximum likelihood estimator. Additionally, this study considers 10,000 iterations of the sampler and uses the first 10% of the data as burn-in. This simulation study is repeated 1,000 times. The measures of accuracy for the estimators are

(i) the estimates of the parameters (Est.)

$$\text{Est.} = \sum_{l=1}^{1,000} \hat{\alpha}_j^{(l)} / 1,000, \quad (25)$$

(ii) the mean square error (MSE)

$$\text{MSE} = \sum_{l=1}^{1,000} (\hat{\alpha}_j^{(l)} - \alpha_j)^2 / 1,000, \quad (26)$$

(iii) the coverage probability (CP)

$$\text{CP} = \#(LCL_{\alpha_j} < \alpha_j < UCL_{\alpha_j}) / 1,000, \quad (27)$$

(iv) the average length (AL)

$$AL = \sum_{l=1}^{1,000} (UCL_{\alpha_j}^{(l)} - LCL_{\alpha_j}^{(l)}) / 1,000, \quad (28)$$

where $\hat{\alpha}_j$ is the j -th estimator $LCL_{\alpha_j}^{(l)}$, $UCL_{\alpha_j}^{(l)}$ are the j -th lower bound and upper bound for the 95% confidence interval of the l -th time and $LCL_{\alpha_j} < \alpha_j < UCL_{\alpha_j}$ is the total of the number of times that α_j is inside the confidence interval. The same measure of accuracy has been applied for the estimators of parameter β . Tables 1–3 report the estimates of the parameters (Est.) together with the MSE. In addition, Tables 4–6 report the 95% CP and the AL.

2.5. Real data analysis

In this section, the median discrete Weibull regression is applied to a real data set that shows the ability for over-dispersion data (see Kalktawi [9]). This data is available under the ‘‘COUNT’’ package in R from Hosmer and Lemeshow [28] and represents the number of visits to a doctor by pregnant women in the first three months of their pregnancies with 189 observations. The response variable is the number of physician visits in first trimester, and the three explanatory variables are history of mother smoking (1=history of mother smoking; 0=mother nonsmoker) (x_1), weight (lbs) at last menstrual period (x_2) and age of mother (x_3). For fitting the discrete Weibull distribution of the response variable, the Kolmogorov-Smirnov statistic is 0.0985 less than the critical value of 0.0989. Thus, this data can be modeled by the discrete Weibull distribution. Moreover, this data is modeled by the Poisson, negative binomial and discrete Weibull distributions. The results show that the Akaike information criterion (AIC) from the Poisson, negative binomial and discrete Weibull distributions are 476.59, 466.85 and 466.84, respectively. Additionally, the mean and variance of the data are 0.7937 and 1.1221, respectively, which indicates an over-dispersion case.

We estimate parameters and construct the 95% confidence intervals via the maximum likelihood estimation method. To demonstrate how the proposed Bayesian method under the three prior distributions can be used in practice, this study calculates parameter estimates and the 95% HPD interval of the parameters with $L = 10,000$ replicates and 10% of the chain for burn-in; $B = 1,000$. In addition, the three information criteria, namely, the AIC, the Bayesian information criterion (BIC) and the deviance information criterion (DIC) (see in Kalktawi [9] and Haselimashhadi et al. [14]) are applied to compare models with different estimates of parameters, which are models for the three explanatory variables and a subset of parameters of that significance. All results are reported in Table 7. Along with, the traceplot, autocorrelation for sampled values and posterior densities of significant independent variables are performed in Figure 1.

3. Results and discussion

3.1. Results and discussion of simulation study

An inspection of Tables 1–3 show that the estimates of the parameters (Est.) in the simulation study obtained by all methods seem to be close to the true parameter values. Moreover, all of the estimators have monotonic behaviors according to the MSE, namely, when n increases, the estimated MSE values decrease. The Bayes estimators have a smaller MSE than the estimators of MLE. The MSE of the Bayes(L) outperforms other methods in almost all situations. Additionally, note that the MSE for all estimators of the three Bayesian methods behave very similarly when $n = 200$. Conversely, the MLE presents the highest MSE but has a satisfactory performance when sample size increases. In addition,, note that the MSE of the estimators of α_2 are very high and may cause what we define as a strong effect on x_2 or the high variance of the estimators of α_2 . Furthermore, the explanatory variable affects the MSE of estimators. However, when n is large enough and β increases, the MSE of estimators will decrease and will no longer be very high anymore.

Table 1. Est. and MSE when $\theta=(1.5,0.4,-0.2,0.8,0.9)$.

n	parameter	MLE		Bayes(U)		Bayes(N)		Bayes(L)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
50	α_0	1.5070	0.0578	1.5006	0.0579	1.4884	0.0585	1.4803	0.0578
	α_1	0.4035	0.0316	0.4096	0.0308	0.4098	0.0309	0.3961	0.0296
	α_2	-0.3635	1.0364	-0.1902	0.9833	-0.1870	0.9819	-0.1390	0.5564
	α_3	0.7746	0.1228	0.8001	0.1195	0.8005	0.1196	0.7704	0.1110
	β	0.9622	0.0191	0.9274	0.0150	0.9139	0.0142	0.9109	0.0140
100	α_0	1.4906	0.0273	1.4898	0.0263	1.4836	0.0265	1.4813	0.0268
	α_1	0.3995	0.0155	0.4027	0.0149	0.4031	0.0150	0.3964	0.0148
	α_2	-0.3319	0.5356	-0.1948	0.4429	-0.1953	0.4420	-0.1637	0.3085
	α_3	0.8031	0.0583	0.8164	0.0566	0.8169	0.0569	0.7998	0.0559
	β	0.9286	0.0077	0.9125	0.0065	0.9060	0.0063	0.9042	0.0063
200	α_0	1.49830	0.0149	1.5013	0.0132	1.4983	0.0132	1.4970	0.0132
	α_1	0.3990	0.0070	0.4018	0.0065	0.4018	0.0064	0.3986	0.0064
	α_2	-0.3440	0.3627	-0.1686	0.2088	-0.1687	0.2088	-0.1483	0.1623
	α_3	0.7938	0.0284	0.8031	0.0253	0.8032	0.0252	0.7948	0.0251
	β	0.9095	0.0037	0.9058	0.0031	0.9026	0.0030	0.9015	0.0030

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

Table 2. Est. and MSE when $\theta=(1.5,0.4,-0.2,0.8,1.6)$.

n	parameter	MLE		Bayes(U)		Bayes(N)		Bayes(L)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
50	α_0	1.5009	0.0187	1.5026	0.0175	1.4962	0.0177	1.4951	0.0176
	α_1	0.4054	0.0102	0.4052	0.0097	0.4050	0.0097	0.4003	0.0096
	α_2	-0.2270	0.3470	-0.1931	0.3101	-0.1933	0.3108	-0.1647	0.2268
	α_3	0.7936	0.0398	0.8000	0.0378	0.8006	0.0377	0.7886	0.0373
	β	1.7116	0.0584	1.6525	0.0451	1.6303	0.0420	1.6281	0.0418
100	α_0	1.4952	0.0086	1.4955	0.0081	1.4922	0.0082	1.4916	0.0082
	α_1	0.4012	0.0051	0.4017	0.0048	0.4016	0.0048	0.3995	0.0048
	α_2	-0.2314	0.1710	-0.2003	0.1391	-0.1982	0.1372	-0.1798	0.1147
	α_3	0.8051	0.0193	0.8091	0.0180	0.8089	0.0180	0.8039	0.0179
	β	1.6521	0.0248	1.6248	0.0196	1.6140	0.0189	1.6126	0.0190
200	α_0	1.5006	0.0047	1.5007	0.0041	1.4991	0.0041	1.4987	0.0041
	α_1	0.4012	0.0022	0.4010	0.0021	0.4010	0.0021	0.4001	0.0021
	α_2	-0.2090	0.0871	-0.1821	0.0661	-0.1823	0.0661	-0.1693	0.0582
	α_3	0.7979	0.0093	0.8016	0.0081	0.8017	0.0081	0.7990	0.0080
	β	1.6195	0.0111	1.6090	0.0088	1.6040	0.0086	1.6027	0.0087

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

Table 3. Est. and MSE when $\theta=(1.5,0.4,-0.2,0.8,2.5)$.

n	parameter	MLE		Bayes(U)		Bayes(N)		Bayes(L)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
50	α_0	1.4971	0.0089	1.5011	0.0074	1.4972	0.0074	1.4977	0.0075
	α_1	0.4025	0.0043	0.4034	0.0041	0.4014	0.0041	0.4035	0.0041
	α_2	-0.2189	0.1713	-0.1965	0.1302	-0.1782	0.1076	-0.1971	0.1300
	α_3	0.7958	0.0188	0.8004	0.0156	0.7952	0.0156	0.8002	0.0157
	β	2.6337	0.1496	2.5842	0.1114	2.5483	0.1030	2.5507	0.1033
100	α_0	1.4957	0.0045	1.4975	0.0034	1.4975	0.0034	1.4954	0.0034
	α_1	0.4016	0.0024	0.4018	0.0020	0.4017	0.0020	0.4008	0.0020
	α_2	-0.2216	0.1024	-0.2018	0.0599	-0.2015	0.0597	-0.1889	0.0540
	α_3	0.8022	0.0105	0.8056	0.0075	0.8057	0.0075	0.8034	0.0075
	β	2.5486	0.0729	2.5403	0.0471	2.5239	0.0454	2.5225	0.0454
200	α_0	1.4987	0.0024	1.5007	0.0018	1.4998	0.0018	1.4996	0.0018
	α_1	0.4002	0.0012	0.4006	0.0009	0.4005	0.0009	0.4001	0.0009
	α_2	-0.2118	0.0703	-0.1947	0.0410	-0.1949	0.0409	-0.1868	0.0393
	α_3	0.7987	0.0052	0.8016	0.0037	0.8014	0.0037	0.8004	0.0036
	β	2.5099	0.0401	2.5144	0.0239	2.5064	0.0236	2.5054	0.0236

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

Tables 4–6 show that when sample sizes were increased, the CP of all methods was generally close to the nominal confidence level. The CP obtained when using the three Bayesian methods are closer to the nominal level than using MLE method. Additionally, the CP for the three prior distributions of the Bayesian method behave remarkably similar. Regarding the AL, as sample sizes were increased, the AL of the 95% confidence intervals decreased for all methods. For cases $\beta = 0.9$ and $\beta = 1.6$, the AL based on the Bayes(L) were the shortest for almost all situations after the CP was considered, while the Bayes(U) were the shortest in the case of $\beta = 2.5$. Although the AL for the MLE method performs the shortest in some situations, but the CP is farthest from the nominal confidence level. Additionally, the results of the AL for the estimator α_2 are quite wide, which is the explanation given for the extremely high MSE values of the estimator α_2 .

Table 4. CP and AL when $\theta=(1.5,0.4,-0.2,0.8,0.9)$.

n	parameter	MLE		Bayes(U)		Bayes(N)		Bayes(L)	
		CP	AL	CP	AL	CP	AL	CP	AL
50	α_0	0.918	0.8825	0.939	0.9398	0.943	0.9523	0.948	0.9497
	α_1	0.925	0.6384	0.952	0.6802	0.956	0.6906	0.958	0.6776
	α_2	0.932	3.6636	0.943	3.8821	0.953	3.9383	0.975	3.3774
	α_3	0.923	1.2656	0.939	1.3614	0.941	1.3795	0.947	1.3446
	β	0.933	0.4552	0.932	0.4476	0.938	0.4428	0.938	0.4412
100	α_0	0.940	0.6229	0.952	0.6465	0.950	0.6499	0.952	0.6510
	α_1	0.918	0.4443	0.931	0.4573	0.928	0.4603	0.925	0.4584
	α_2	0.921	2.5582	0.944	2.6243	0.951	2.6481	0.964	2.4005
	α_3	0.935	0.8935	0.948	0.9242	0.945	0.9306	0.949	0.9235
	β	0.938	0.3066	0.948	0.3026	0.947	0.3013	0.945	0.3010
200	α_0	0.931	0.4491	0.949	0.4521	0.951	0.4538	0.946	0.4542
	α_1	0.947	0.3138	0.956	0.3166	0.951	0.3179	0.954	0.3170
	α_2	0.871	1.8096	0.948	1.8175	0.946	1.8231	0.959	1.7103
	α_3	0.939	0.6332	0.958	0.6405	0.958	0.6438	0.958	0.6430
	β	0.920	0.2115	0.937	0.2098	0.936	0.2094	0.936	0.2092

Table 5. CP and AL when $\theta=(1.5,0.4,-0.2,0.8,1.6)$.

n	parameter	MLE		Bayes(U)		Bayes(N)		Bayes(L)	
		CP	AL	CP	AL	CP	AL	CP	AL
50	α_0	0.915	0.4946	0.938	0.5214	0.944	0.5286	0.946	0.5278
	α_1	0.926	0.3605	0.944	0.3820	0.953	0.3863	0.950	0.3836
	α_2	0.917	2.0497	0.950	2.1733	0.951	2.2057	0.971	2.0250
	α_3	0.926	0.7132	0.939	0.7624	0.946	0.7728	0.945	0.7663
	β	0.926	0.7826	0.940	0.7683	0.940	0.7569	0.941	0.7562
100	α_0	0.935	0.3518	0.949	0.3601	0.946	0.3618	0.948	0.3624
	α_1	0.912	0.2509	0.923	0.2575	0.932	0.2588	0.931	0.2585
	α_2	0.921	1.4423	0.941	1.4748	0.946	1.4822	0.958	1.4122
	α_3	0.936	0.5041	0.945	0.5194	0.942	0.5211	0.948	0.5199
	β	0.930	0.5268	0.947	0.5210	0.953	0.5179	0.951	0.5171
200	α_0	0.936	0.2504	0.955	0.2532	0.950	0.2534	0.952	0.2536
	α_1	0.945	0.1769	0.955	0.1790	0.954	0.1794	0.953	0.1792
	α_2	0.929	1.0147	0.950	1.0240	0.945	1.0292	0.959	0.9931
	α_3	0.944	0.3567	0.958	0.3610	0.960	0.3619	0.962	0.3623
	β	0.917	0.3627	0.943	0.3617	0.941	0.3596	0.939	0.3594

Table 6. CP and AL when $\theta=(1.5,0.4,-0.2,0.8,2.5)$.

n	parameter	MLE		Bayes(U)		Bayes(N)		Bayes(L)	
		CP	AL	CP	AL	CP	AL	CP	AL
50	α_0	0.919	0.3255	0.948	0.3348	0.944	0.3375	0.947	0.3381
	α_1	0.928	0.2383	0.937	0.2466	0.947	0.2478	0.944	0.2496
	α_2	0.911	1.3709	0.944	1.4014	0.961	1.3439	0.944	1.4162
	α_3	0.920	0.4746	0.943	0.4896	0.936	0.4926	0.939	0.4948
	β	0.924	1.2153	0.941	1.1989	0.943	1.1721	0.941	1.1712
100	α_0	0.926	0.2311	0.946	0.2317	0.950	0.2327	0.946	0.2331
	α_1	0.915	0.1657	0.929	0.1665	0.929	0.1672	0.932	0.1673
	α_2	0.899	0.9562	0.943	0.9509	0.943	0.9554	0.946	0.9264
	α_3	0.924	0.3337	0.941	0.3343	0.944	0.3360	0.946	0.3356
	β	0.904	0.8210	0.946	0.8140	0.949	0.8042	0.947	0.8039
200	α_0	0.929	0.1630	0.955	0.1623	0.958	0.1623	0.957	0.1625
	α_1	0.921	0.1158	0.946	0.1153	0.948	0.1156	0.948	0.1157
	α_2	0.907	0.6652	0.946	0.6583	0.948	0.6591	0.949	0.6471
	α_3	0.934	0.2324	0.951	0.2318	0.952	0.2323	0.952	0.2324
	β	0.896	0.5652	0.942	0.5617	0.942	0.5584	0.940	0.5588

3.2. Results and discussion of real data analysis

Table 7 shows the results of significant explanatory variables that are selected from the three explanatory variables. We only report that an explanatory variable x_3 (age of mother) shows significant in all methods. Results from comparing models with different estimates of parameters suggest that a model for only x_3 provided better fitting than a model for the three explanatory variables, according to the AIC, BIC and DIC. Regarding the DIC of the three Bayesian methods for each of the two models, they exhibited remarkably similar behavior. Additionally, the AIC and BIC are included as well. This finding corresponds to the results of the simulation study in case $n = 200$. Figure 1 shows the traceplot, autocorrelation for sampled values and posterior densities for regression coefficient α_3 based on the Bayes(L). It can be seen that the trace plot showed adequate convergence. Moreover, it is clear that the sampled values are well mixed and exhibit adequate stability for autocorrelation.

Table 7. Parameter estimates, the 95% confidence intervals, and the three information criteria.

Models	Parameters	MLE	Bayes(U)	Bayes(N)	Bayes(L)	
x_1, x_2, x_3	α_0	-1.0983* (-1.8269,-0.3697)	-1.1566* (-2.0877,-0.4473)	-1.667* (-1.8724,-0.4639)	-1.0553* (-1.7395,-0.3875)	
	α_1	-0.6240 (-0.3173,0.1924)	-0.6690 (-0.3491,0.1849)	-0.6980 (-0.3169,0.1738)	-0.0478 (-0.3182,0.1795)	
	α_2	0.0029 (-0.0009,0.0069)	0.0031 (-0.0008,0.0072)	0.0030 (-0.0007,0.0070)	0.0029 (-0.0009,0.0074)	
	α_3	0.0295* (0.0058,0.0532)	0.0306* (0.0055,0.0585)	0.0314* (0.0083,0.0552)	0.0273* (0.0057,0.0490)	
	β	1.19312* (0.9986,1.3875)	1.1676* (0.9897,1.3639)	1.1564* (0.9684,1.3795)	1.1717* (0.9848,1.3579)	
	AIC	463.1398	463.2289	463.3156	463.2471	
	BIC	479.3486	479.4376	479.5244	479.4558	
	DIC	-	463.3949	463.0142	462.3886	
	x_3	α_0	-0.7974* (-1.3929,-0.2019)	-0.8387* (-1.5015,-0.2737)	-0.7974* (-1.3929,-0.2019)	-0.7791* (-1.4164,-0.1910)
		α_3	0.0321* (0.0083,0.0558)	0.0333* (0.0111,0.0574)	0.0321* (0.0083,0.0558)	0.0310* (0.0083,0.0555)
β		1.1733* (0.9850,1.3617)	1.1587* (0.9676,1.3525)	1.1733* (0.9850,1.3617)	1.1545* (0.9532,1.3618)	
AIC		459.5681**	461.6021**	461.7598**	461.6235**	
BIC		466.0515**	471.3273**	471.3521**	471.3487**	
DIC		-	461.5384**	461.7598**	461.7922**	

Note: (*) denotes the 95% confidence intervals does not contain zero (statistically significant) and (**) denotes the minimum value of each information criteria between models of and only.

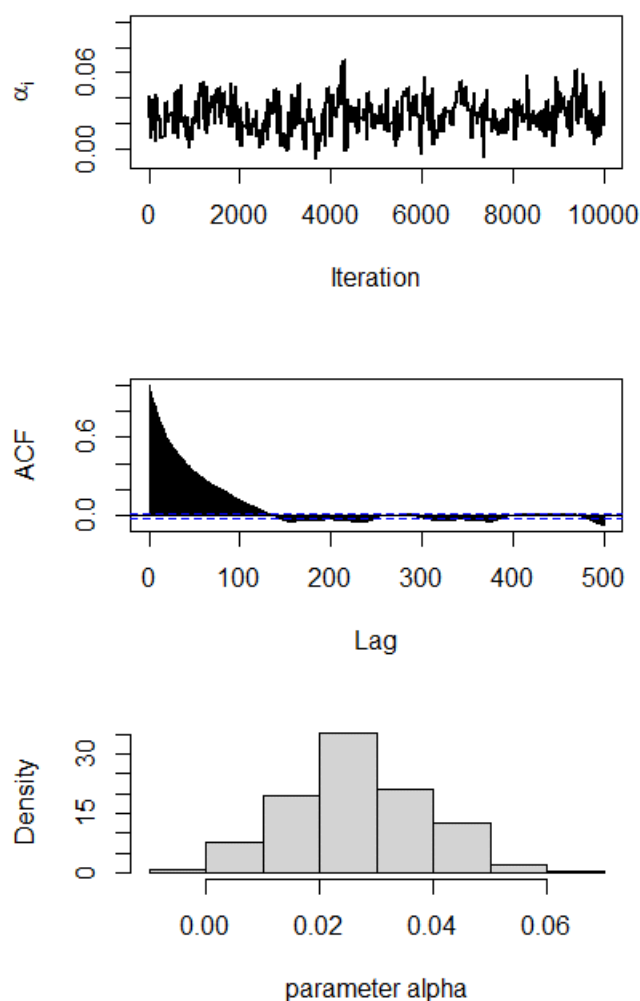


Figure 1. Traceplot, autocorrelation and posterior densities for regression coefficient of Bayes(L).

4. Conclusions

This paper introduced the Bayesian estimation for the discrete Weibull regression model via the median link function. The Bayesian approach was considered on the three different prior distributions: Uniform non-informative prior, normal prior and Laplace prior. The augmented random walk Metropolis procedure was also proposed to compute the Bayes estimates of the unknown parameters. Moreover, the maximum likelihood estimation was compared. The performance of these methods was compared by using the Monte Carlo simulation based on the MSE and CP criteria. These criteria were calculated for different sample sizes based on both under-dispersion and over-dispersion data, along with the application of the methods illustrated by using a real dataset available on the literature

to compare models with different estimates of parameters via the Akaike information criterion, the Bayesian information criterion and the deviance information criterion. Based on MSE criterion, the Bayesian using Laplace prior distribution for estimating the parameters performs better than other approaches. Additionally, the three Bayesian methods behaved very similarly with the large sample size. Estimated coverage probabilities of the three Bayesian approaches were considered as the criteria of a good confidence interval. In addition, the results of real data analysis were coincided with those in the simulation study. Overall, the Bayesian estimation using Laplace prior distribution outperforms other methods for parameters estimation. However, the Bayesian estimation using all three prior distributions can be an effective alternative for this model.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors are grateful to the Editor in Chief and anonymous referee for the insightful and constructive comments that improve this paper.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. A. C. Cameron, P. K. Trivedi, *Regression analysis of count data*, Cambridge: Cambridge University Press, 2013. <https://doi.org/10.1017/CBO9781139013567>
2. K. F. Sellers, G. Shmueli, A flexible regression model for count data, *Ann. Appl. Stat.*, **4** (2010), 943–961. <https://doi.org/10.1214/09-AOAS306>
3. D. Lambert, Zero-inflated Poisson regression, with an application to defects in manufacturing, *Technometrics*, **34** (1992), 1–14. <https://doi.org/10.2307/1269547>
4. D. B. Hall, Zero-inflated Poisson and binomial regression with random effects: A case study, *Biometrics*, **56** (2000), 1030–1039. <https://doi.org/10.1111/j.0006-341X.2000.01030.x>
5. K. K. W. Yau, K. Wang, A. H. Lee, Zero-inflated negative binomial mixed regression modeling of over-dispersed count data with extra zeros, *Biomretical J.*, **45** (2003), 437–452. <https://doi.org/10.1002/bimj.200390024>
6. T. Nakagawa, S. Osaki, The discrete Weibull distribution, *IEEE T. Reliab.*, **24** (1975), 300–301. <https://doi.org/10.1109/TR.1975.5214915>
7. L. Dang, X. He, D. Tang, Y. Li, T. Wang, A fatigue life prediction approach for laser-directed energy deposition titanium alloys by using support vector regression based on pore-induced failures, *Int. J. Fatigue*, **159** (2022), 106748. <https://doi.org/10.1016/j.ijfatigue.2022.106748>

8. J. C. He, S. P. Zhu, C. Luo, X. Niu, Q. Wang, Size effect in fatigue modelling of defective materials: Application of the calibrated weakest-link theory, *Int. J. Fatigue*, **165** (2022), 10721. <https://doi.org/10.1016/j.ijfatigue.2022.107213>
9. H. S. Kalktawi, *Discrete Weibull regression model for count data*, PhD. thesis, London: Brunel University London, 2017.
10. J. D. Englehardt, R. Li, The discrete Weibull distribution: An alternative for correlated counts with confirmation for microbial counts in water, *Risk Anal.*, **31** (2011), 370–381. <https://doi.org/10.1111/j.1539-6924.2010.01520.x>
11. H. S. Klakattawi, V. Vinciotti, K. Yu, A simple and adaptive dispersion regression model for count data, *Entropy*, **20** (2018), 142. <https://doi.org/10.3390/e20020142>
12. A. Peluso, V. Vinciotti, Discrete weibull generalised additive model: An application to count fertility data, *J. Roy. Stat. Soc. C-Appl.*, 2018. <https://doi.org/10.1111/rssc.12311>
13. S. Wang, W. Chen, M. Chen, Y. Zhou, Maximum likelihood estimation of the parameters of the inverse Gaussian distribution using maximum rank set sampling with unequal samples, *Math. Popul. Stud.*, **30** (2023), 1–21. <https://doi.org/10.1080/08898480.2021.1996822>
14. H. Haselimashhadi, V. Vinciotti, K. Yu, A novel Bayesian regression model for counts with an application to health data, *J. Appl. Stat.*, **45** (2018), 1085–1105. <https://doi.org/10.1080/02664763.2017.1342782>
15. A. Gelman, A. Jakulin, M. G. Pittau, Y. Su, A weakly informative default prior distribution for logistic and other regression models, *Ann. Appl. Stat.*, **2** (2008), 1360–1383. <https://doi.org/10.1214/08-AOAS191>
16. S. Fu, Hierarchical Bayesian LASSO for a negative binomial regression, *J. Stat. Comput. Sim.*, **86** (2016), 2182–2203. <https://doi.org/10.1080/00949655.2015.1106541>
17. C. Chaniavidis, L. Evers, T. Neocleous, A. Nobile, Efficient Bayesian inference for COM-Poisson regression models, *Stat. Comput.*, **28** (2018), 595–608. <https://doi.org/10.1007/s11222-017-9750-x>
18. D. Chairprasithikul, M. Duangsaphon, Bayesian inference for the discrete Weibull regression model with Type-I right censored data, *Thail. Statist.*, **20** (2022), 791–811.
19. D. Chairprasithikul, M. Duangsaphon, Bayesian inference of discrete Weibull regression model for excess zero counts, *Sci. Tech. Asia*, **27** (2022), 152–174.
20. A. A. Ahmadini, A. S. Hassan, A. N. Zaky, S. S. Alshqaq, Bayesian inference of dynamic cumulative residual entropy from Pareto II distribution with application to COVID-19, *AIMS Math.*, **6** (2021), 2196–2216. <https://doi.org/10.3934/math.2021133>
21. H. Okasha, M. Nassar, S. A. Dobbah, E-Bayesian estimation of Burr Type XII model based on adaptive Type-II progressive hybrid censored data, *AIMS Math.*, **6** (2021), 4173–4196. <https://doi.org/10.3934/math.2021247>
22. H. Sato, M. Ikota, A. Sugimoto, H. Masuda, A new defect distribution metrology with a consistent discrete exponential formula and its applications, *IEEE T. Semiconductor M.*, **12** (1999), 409–418. <https://doi.org/10.1109/66.806118>
23. D. Roy, Discrete rayleigh distribution, *IEEE T. Reliab.*, **53** (2014), 255–260. <https://doi.org/10.1109/TR.2004.829161>

24. A. Barbiero, *DiscreteWeibull: Discrete Weibull distributions (Type 1 and 3)*, 2015. Available from: <http://CRAN.R-project.org/package=DiscreteWeibull.Rpackageversion1.0.1>.
25. R. J. Serfling, *Approximation theorems of mathematical statistics*, New York: John Wiley & Sons, 1980. <https://doi.org/10.1002/9780470316481>
26. W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57** (1970), 97–109. <https://doi.org/10.1093/biomet/57.1.97>
27. W. R. Gilks, S. Richardson, D. Spiegelhalter, *Markov chain Monte Carlo in practice*, London: Chapman and Hall, 1996. <https://doi.org/10.1201/b14835>
28. D. W. Hosmer, S. Lemeshow, *Applied logistic regression*, New York: John Wiley & Sons, 2004. <https://doi.org/10.1002/0470011815.b2a10029>



AIMS Press

© 2024 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)