Thailand Statistician
October 2025; 23(4): 916-936
http://statassoc.or.th
Contributed paper

Bayesian Estimation and Prediction for Zero-Inflated Discrete Weibull Distribution

Monthira Duangsaphon [a], Kamon Budsaba [a], Sudarat Nidsunkid [b] and Dusit Chaiprasithikul*[c]

[a] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathumthani, Thailand.

[b] Department of Statistics, Faculty of Science, Kasetsart University, Bangkok, Thailand.

[c] Department of Statistics, Faculty of Science, Silpakorn University, Nakhon Pathom, Thailand.

*Corresponding author; e-mail: chaiprasithikul_d@su.ac.th

Received: 25 October 2024
Revised: 14 June 2025
Accepted: 24 June 2025

Abstract

This paper proposes the Bayesian estimation of the zero-inflated discrete Weibull distribution assuming three prior distributions, namely Beta-Uniform-Uniform prior, Beta-Jeffreys' rule prior, and Beta-Beta-Gamma prior. It is commonly known that there is no compact form for the Bayes estimators. The Bayesian estimate of the model parameters has been performed through the random walk Metropolis algorithm. The Bayes estimates of the unknown parameters and the credible interval construction are established based on the generated samples. Moreover, the maximum likelihood estimation is considered, as well as the confidence interval estimation for the model parameters has been performed through normal approximation. The performance of the Bayes estimators has also been compared with the classical estimators through the Monte Carlo simulation study. Further, the Bayesian prediction of a future observation is proposed under the three prior distributions. The posterior predictive distribution of a future observation cannot be evaluated analytically. To obtain the estimate of future sample, the Metropolis-Hastings algorithm is used. Two datasets have been analyzed to show how the proposed model and the method work in practice.

Keywords: maximum likelihood estimation, predictive distribution, count data, Metropolis-Hastings algorithm, cross-validation.

1. Introduction

Many application fields, including medical, actuarial science, biostatistics, demography, economics, engineering, political science, and sociology, utilize count data. In medicine, examples; the number of heart attacks and the number of hospitalization days in medical studies. Additional instances include the number of absent students in education studies, the number of domestic abuse by parents against their child in social science studies, the number of victim from the accident, and the number of fatalities. The term of count variable refers to individual count data that is handled like a random variable, including the Poisson, negative binomial, and Conway-Maxwell Poisson distributions. A Poisson distribution's well-known property is that its mean equals its variance. Since the variance of the count distribution is typically greater than the mean (over-dispersion), many circumstances are frequently irrational (Gardner et al. 1995). To model count variables when the data handles over-dispersion, a negative binomial distribution is utilized (Gardner et al. 1995, Allison and Waterman 2002, Liu et al. 2005). Nonetheless, there are situations where the variance of the count distribution is smaller than the mean (under-dispersion). The Conway-Maxwell Poisson distribution is used to deal with under-dispersion and over-dispersion (Sellers and Shmueli 2010; Taveekal et al. 2023). Additionally, a discrete Weibull distribution is frequently employed to depict count data (Kalktawi et al. 2018; Haselimashhadi et al. 2018; Chaiprasithikul and Duangsaphon 2022; Duangsaphon et al. 2023; Duangsaphon et al. 2024). The handling of discrete data with under- and over-dispersion is addressed (Collins et al. 2020; Chaiprasithikul and Duangsaphon 2022).

The discrete Weibull distribution was proposed by Nakagawa and Osaki (1975). They looked at failure studies, where the number of cycles to failure is frequently used to assess the duration to failure, turning it into a discrete random variable. Failure data from failure studies are typically measured in discrete time, such as blows, rotations, cycles, or shocks, in failure analysis. The discrete Weibull distribution is very helpful to theorists and reliability engineers. Let Y be a discrete random variable which follows the discrete Weibull distribution with the parameters q and β, denoted by $Y \sim DW(q,\beta)$. The probability mass function of Y is given by

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

where $0 < q < 1$ and $\beta > 0$ are the shape parameters. In addition, the parameter $q = 1 - p_Y(0;q,\beta)$ which is the probability of Y being more than zero. In particular, the numerical analyses have approximately shown that; $0 < \beta \le 1$ is a case of over-dispersion, regardless of the value of q, $\beta > 2$ is a case of under-dispersion, regardless of the value of q, $1 < \beta < 2$ leads to both cases of over and under-dispersion depending on the value of q (Kalktawi 2017).

Count data in numerous experiments may contain a large number of zeros. Some studies may be more interested in predicting the frequency of zeros, the counts in some experiments are generated by two different mechanisms for zeros and non-zeros in the data. A modified method should be applied when the aim is to distinguish between zero and non-zero data generating processes which can be explained using mixture models which is zero-inflated model, and hence additional care is required to choose the applied model. Many applied studies with count data employ zero-inflated models. For example, Feng (2021) provided a better understanding of the differences between zero-inflated and hurdle models. The performances based on simulation study of the two models depend on the percentage of the zero-deflated data points in the data and the discrepancy in the data generating processes between the structural zeros and sampling zeros. Bekalo and Kebede (2021) showed that the zero-inflated Poisson, zero-inflated negative binomial and hurdle models were better fitted the number of antenatal care service visits data than Poisson and negative binomial. Altun et al. (2023) proposed a new count regression model for zero-inflated and over-dispersed count data sets based on the re-parametrization of the Poisson generalized-Lindley distribution, namely, zero-inflated Poisson generalized-Lindley linear model. In some situations, the zero-inflated model based on the discrete Weibull distribution may be a good fit due to various dispersions of the count data in the presence of a large number of zeros.

The zero-inflated discrete Weibull (ZIDW) model can be derived as a two-component mixture models, that is, mixing a point mass at zero and a count distribution, $p_Y\left(y;q,\beta\right)$. Thus, the probability mass function of ZIDW is given by

$$p\left(y;\pi,q,\beta\right)=\begin{cases}\pi+(1-\pi)(1-q) & ;y=0\\ (1-\pi)\left(q^{y^{\beta}}-q^{(y+1)^{\beta}}\right) & ;y=1,2,3\ldots\end{cases} \tag{2}$$

where y is the count variable, $0<\pi<1$ is a zero-inflation parameter (the probability or proportion of a structural zero). It is to be mentioned here that this model has not been considered under the classical and Bayesian estimations in the earlier literature.

The maximum likelihood estimation of parameters is valid for an asymptotically large sample size of data (Wang et al. 2023). One of the most common problems occurring in many parameters in model is the maximum likelihood estimates that become unstable with larger standard errors of the estimates which 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. Additionally, the Bayesian technique is employed to solve the problem of predicting future observations, whereas in classical statistics, one simply works with the model that is fitted and uses its mean, median, or mode to predict the future observation without any prior knowledge of the model parameters.

Recently, Singh et al. (2013) developed the Bayesian estimation and prediction procedure for flexible Weibull distribution under Type-II censoring scheme assuming different priors for the model parameters. Unhapipat et al. (2018) suggested the zero-inflated Poisson distribution to model of many examples with real-life datasets for predicting a future observation using Bayesian predictive inference. Duangsaphon et al. (2023) proposed the maximum likelihood and Bayesian approaches for two parameters estimation for the discrete Weibull distribution. The Bayesian procedure to the prediction problems of future observations which use the concept of Bayesian predictive posterior distribution. They discovered that the data dispersion and tail behavior of the relevant prediction distributions affect how well the Bayesian technique performs. The latter has probability concentrated at zero and is heavily skewed. Additionally, Irfan and Sharma (2024) proposed two approximation techniques for the Bayes estimators, namely, Lindley's approximation and Metropolis-Hastings within Gibbs sampler algorithm and also to construct the associate highest posterior density credible intervals. Furthermore, Bayesian prediction, predictive density, and predictive intervals are derived for future observation and decision.

In the present article, the Bayesian estimation is examined, based on Monte Carlo Markov chain (MCMC) methods, namely, the random walk Metropolis algorithm for the zero-inflated discrete Weibull distribution under the three different prior distributions; Beta-Uniform-Uniform prior, Beta-Jeffreys' rule prior, and Beta-Beta-Gamma prior. Moreover, the posterior predictive probability mass function of the future observations is estimated via the Metropolis-Hastings algorithm and a numerical method using the *hcubatur()* function in *cubature* package of the R language. A simulation study is conducted to compare the performance of the Bayes estimators and also with the maximum likelihood estimators. Two real datasets are illustrated the proposed estimations.

The remainder of this paper is organized as follows. The maximum likelihood estimation is presented in Section 2. In Section 3, the Bayesian estimation procedure and prediction are presented. A simulation study is reported in Section 4 to investigate the ability of different estimations. This model is applied to two real datasets in Section 5. Results and concluding remarks are given in Section 6 to 7, respectively.

2. Maximum Likelihood Estimation

Given a random sample $Y_1, Y_2, ..., Y_n$ from the ZIDW distribution in (2), then the likelihood function of the observed sample $\underset{\sim}{y} = (y_1, y_2, ..., y_n)$ can be written as

$$L(\pi, q, \beta \mid \underset{\sim}{y}, \underset{\sim}{\delta}) = \prod_{i=1}^{n} [\pi + (1-\pi)(1-q)]^{\delta_i} \left[(1-\pi)\left(q^{y_i^{\beta}} - q^{(y_i+1)^{\beta}} \right) \right]^{1-\delta_i} \tag{3}$$

where $\delta_i = I(y_i = 0) = \begin{cases} 1 & y_i = 0 \\ 0 & y_i > 0. \end{cases}$

The log-likelihood function of from the ZIDW model is given by

$$l(\pi, q, \beta \mid \underset{\sim}{y}, \underset{\sim}{\delta}) = \sum_{i=1}^{n} \ln\left([\pi + (1-\pi)(1-q)]^{\delta_i} \left[(1-\pi)\left(q^{y_i^{\beta}} - q^{(y_i+1)^{\beta}} \right) \right]^{1-\delta_i} \right)$$

$$= \ln[1-(1-\pi)q] \sum_{i=1}^{n} \delta_i + \sum_{i=1}^{n} (1-\delta_i) \ln\left((1-\pi)\left(q^{y_i^{\beta}} - q^{(y_i+1)^{\beta}} \right) \right). \tag{4}$$

The maximum likelihood estimators $\hat{\pi}_{MLE}$, $\hat{q}_{MLE}$, and $\hat{\beta}_{MLE}$ can be obtained by setting the first partial derivatives of the log-likelihood function with respect to each unknown parameter equal to zero;

$$\frac{\partial l(\pi, q, \beta \mid \underset{\sim}{y}, \underset{\sim}{\delta})}{\partial \pi} = \frac{q}{1-(1-\pi)q} \sum_{i=1}^{n} \delta_i - \sum_{i=1}^{n} \frac{(1-\delta_i)}{(1-\pi)} = 0, \tag{5}$$

$$\frac{\partial l(\pi, q, \beta \mid \underset{\sim}{y}, \underset{\sim}{\delta})}{\partial q} = \frac{(\pi-1)}{1-(1-\pi)q} \sum_{i=1}^{n} \delta_i - \sum_{i=1}^{n} \frac{(1-\delta_i)\left(y_i^{\beta} q^{y_i^{\beta}-1} - (y_i+1)^{\beta} q^{(y_i+1)^{\beta}-1} \right)}{q^{y_i^{\beta}} - q^{(y_i+1)^{\beta}}} = 0, \text{ and} \tag{6}$$

$$\frac{\partial l(\pi, q, \beta \mid \underset{\sim}{y}, \underset{\sim}{\delta})}{\partial \beta} = \sum_{i=1}^{n} \frac{(1-\delta_i) \ln q \left(q^{y_i^{\beta}} y_i^{\beta} (\ln y_i) - q^{(y_i+1)^{\beta}} (y_i+1)^{\beta} (\ln(y_i+1)) \right)}{q^{y_i^{\beta}} - q^{(y_i+1)^{\beta}}} = 0. \tag{7}$$

It can be seen that the above equations cannot be solved explicitly and one needs iterative method to solve them. Equations can be solved by some numerical method, like Newton-Raphson or Gauss-Newton method or their variants. Here, we calculate the maximum likelihood estimators by minimizing the negative log-likelihood function of ZIDW model using function *optim()* in R, it presents the BFGS method (Dai 2013, R Core Team and Contributors Worldwide 2022, CRAN Team 2023).

Let $I(\pi, q, \beta)$ be the observed Fisher's information matrix for the 3×3 unknown parameters with contain negative members of the second derivative of the log-likelihood; hence, the variance-covariance matrix is the inverse of the observed Fisher's information matrix,

$$\Sigma = I^{-1}(\pi, q, \beta). \tag{8}$$

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

$$\hat{\Sigma} = \begin{pmatrix} \hat{\sigma}_{\pi\pi} & \hat{\sigma}_{\pi q} & \hat{\sigma}_{\pi\beta} \\ \hat{\sigma}_{q\pi} & \hat{\sigma}_{qq} & \hat{\sigma}_{q\beta} \\ \hat{\sigma}_{\beta\pi} & \hat{\sigma}_{\beta q} & \hat{\sigma}_{\beta\beta} \end{pmatrix}. \tag{9}$$

This matrix can be obtained by inverting the Hessian matrix from the function *hessian* () in R language (R Core Team and Contributors Worldwide 2022). 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 are performed using the maximum likelihood method, then under some regularity conditions (Serfling 1980), these estimators enjoy standard asymptotic properties. Thus, by the asymptotic normality of maximum likelihood estimators, the $100(1-\alpha)\%$ confidence intervals for parameters $\pi, q,$ and β respectively as

$$\hat{\pi} \pm z_{\frac{\alpha}{2}}\sqrt{\hat{\sigma}_{\pi\pi}}, \quad \hat{q} \pm z_{\frac{\alpha}{2}}\sqrt{\hat{\sigma}_{qq}} \text{ and } \hat{\beta} \pm z_{\frac{\alpha}{2}}\sqrt{\hat{\sigma}_{\beta\beta}}, \tag{10}$$

where $z_{\frac{\alpha}{2}}$ is the upper $\frac{\alpha}{2}^{\text{th}}$ quantile the of standard normal distribution.

3. Bayesian Estimation and Prediction

This section is focused on estimating parameters and predicting a future observation of ZIDW distribution.

3.1. Prior selection

In Bayesian procedure, it is essential to select the prior distributions of the unknown parameters to take into account uncertainty of the parameters. It is assumed that the parameters π, q and β are independent. Therefore, the joint prior of π, q and β is given by

$$p(\pi, q, \beta) = p_1(\pi)p_2(q)p_3(\beta).$$

This paper applies three forms of prior distributions of π, q and β for our Bayesian approach. The first scenario focuses on Beta prior for π, uniform non-informative priors for q and β. The second scenario focuses on Beta prior for π and considers Jeffrey's prior (Ashour and Muiftah 2019) for q and β. The third scenario focuses on the Beta priors for π and q, and gamma prior for β.

Scenario 1: Beta-Uniform-Uniform prior

$$p_1(\pi) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\pi^{a-1}(1-\pi)^{b-1}, a>0, b>0, \quad p_2(q) \propto 1, \text{ and } p_3(\beta) \propto 1.$$

Scenario 2: Beta-Jeffreys' rule prior

$$p_1(\pi) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\pi^{a-1}(1-\pi)^{b-1}, a>0, b>0, \quad p_2(q) \propto \frac{1}{\sqrt{q(q-1)}}, \text{ and } p_3(\beta) \propto \frac{1}{\sqrt{\beta}}.$$

Scenario 3: Beta-Beta-Gamma prior

$$p_1(\pi) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\pi^{a-1}(1-\pi)^{b-1}, a>0, b>0,$$

$$p_2(q)=\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)}q^{c-1}(1-q)^{d-1}\text{ , }c>0,d>0,\text{ and }p_3(\beta)=\frac{b^a}{\Gamma(a)}\beta^{a-1}e^{-b\beta}\text{ , }a>0,b>0.$$

Consequently, the joint posterior distribution of the ZIDW parameters (π,α,β) can be readily defined as

$$p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})=\frac{L(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})p(\pi,q,\beta)}{\int_0^\infty\int_0^1\int_0^1 L(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})p(\pi,q,\beta)\,d\pi dq d\beta}\propto L(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})p(\pi,\alpha,\beta). \quad (11)$$

Then, the posterior distribution of the ZIDW parameters (π,α,β) corresponding to the three scenarios are

Posterior for Scenario 1: Beta-Uniform-Uniform prior

$$p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})\propto\prod_{i=1}^{n}\left[\pi+(1-\pi)(1-q)\right]^{\delta_i}\left[(1-\pi)\left(q^{y_i^\beta}-q^{(y_i+1)^\beta}\right)\right]^{1-\delta_i}\pi^{a-1}(1-\pi)^{b-1}. \quad (12)$$

Posterior for Scenario 2: Beta-Jeffreys' rule prior

$$p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})\propto\prod_{i=1}^{n}\left[\pi+(1-\pi)(1-q)\right]^{\delta_i}\left[(1-\pi)\left(q^{y_i^\beta}-q^{(y_i+1)^\beta}\right)\right]^{1-\delta_i}$$
$$\times\pi^{a-1}(1-\pi)^{b-1}\times\frac{1}{\sqrt{q(q-1)}}\times\frac{1}{\sqrt{\beta}}. \quad (13)$$

Posterior for Scenario 3: Beta-Beta-Gamma prior

$$p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})\propto\prod_{i=1}^{n}\left[\pi+(1-\pi)(1-q)\right]^{\delta_i}\left[(1-\pi)\left(q^{y_i^\beta}-q^{(y_i+1)^\beta}\right)\right]^{1-\delta_i}$$
$$\times\pi^{a-1}(1-\pi)^{b-1}\times q^{c-1}(1-q)^{d-1}\times\beta^{a-1}e^{-b\beta}. \quad (14)$$

3.2. Parameter estimation

The Bayes estimator for function $h(\pi,q,\beta)$ of the parameters $\pi,q,$ and β under squared error loss function is the expected value of function $h(\pi,q,\beta)$ under the joint posterior density function. Then, the Bayes estimator for function $h(\pi,q,\beta)$ is given by

$$\hat{h}(\pi,q,\beta)=\int_0^\infty\int_0^1\int_0^1 h(\pi,q,\beta)p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})d\pi dq d\beta. \quad (15)$$

Therefore, the Bayes estimator of $\pi,q,$ and β are

$$\hat{\pi}=\int_0^\infty\int_0^1\int_0^1 \pi\, p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})d\pi dq d\beta,$$
$$\hat{q}=\int_0^\infty\int_0^1\int_0^1 q p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})d\pi dq d\beta,\text{ and}$$
$$\hat{\beta}=\int_0^\infty\int_0^1\int_0^1 \beta p(\pi,q,\beta\big|\underset{\sim}{y},\underset{\sim}{\delta})d\pi dq d\beta. \quad (16)$$

Since the three intractable integrals in (16) above cannot be obtained in nice closed form, this study chose the random walk Metropolis-Hastings algorithm to estimate the Bayes estimators.

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 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 MH algorithm see Hastings (1970) and Gilks et al. (1996).

Furthermore, this study determines the joint posterior density function of the parameters $\boldsymbol{\theta} = (\pi, q, \beta)$, $p(\boldsymbol{\theta} \mid \underset{\sim}{y}, \underset{\sim}{\delta})$, in (11) as the target distribution, while $\boldsymbol{\theta}$ is the current state value, and $\boldsymbol{\theta}^*$ is the proposal value generated from the proposal distribution $q(\boldsymbol{\theta}^* \mid \boldsymbol{\theta})$. Then, the proposal value $\boldsymbol{\theta}^*$ is accepted with the probability $p = \min\left(1, R_{\boldsymbol{\theta}}\right)$, where

$$R_{\boldsymbol{\theta}} = \frac{L\left(\boldsymbol{\theta}^* \middle| \underset{\sim}{y}, \underset{\sim}{\delta}\right) p(\boldsymbol{\theta}^*)}{L\left(\boldsymbol{\theta} \middle| \underset{\sim}{y}, \underset{\sim}{\delta}\right) p(\boldsymbol{\theta})} \times \frac{q\left(\boldsymbol{\theta} \middle| \boldsymbol{\theta}^*\right)}{q\left(\boldsymbol{\theta}^* \middle| \boldsymbol{\theta}\right)}. \tag{17}$$

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

$$R_{\boldsymbol{\theta}} = \frac{L\left(\boldsymbol{\theta}^* \middle| \underset{\sim}{y}, \underset{\sim}{\delta}\right) p(\boldsymbol{\theta}^*)}{L\left(\boldsymbol{\theta} \middle| \underset{\sim}{y}, \underset{\sim}{\delta}\right) p(\boldsymbol{\theta})}. \tag{18}$$

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

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

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 $R_{\boldsymbol{\theta}} = \dfrac{L\left(\boldsymbol{\theta}^* \middle| \underset{\sim}{y}, \underset{\sim}{\delta}\right) p(\boldsymbol{\theta}^*)}{L\left(\boldsymbol{\theta} \middle| \underset{\sim}{y}, \underset{\sim}{\delta}\right) p(\boldsymbol{\theta})}$.

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

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 π, q, and β from the average of the generated values given by

$$\hat{\theta}_{Bayes} = \frac{1}{L - B} \sum_{l=B+1}^{L} \theta^{(l)}, \tag{19}$$

where $\theta^{(l)}$ is Bayes estimator for a parameter in vector $\boldsymbol{\theta}=(\pi,q,\beta)$ of the l^{th} time.

Given an MCMC sample $\theta^{(l)}, l=B+1,B+2,...,L,$ the credible intervals interval (Ghosh 2006) for θ can be shown as follows:

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

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

Step 2. The $100(1-\alpha)\%$ credible intervals of θ are constructed as

$$\left(\theta_{((L-B)\frac{\alpha}{2})}, \theta_{((L-B)(1-\frac{\alpha}{2}))}\right), \tag{20}$$

where $\theta_{((L-B)\frac{\alpha}{2})}$ is the $\frac{\alpha}{2}^{\text{th}}$ quantile the of $\theta_{(1)},\theta_{(2)},...,\theta_{(L-B)}$ and $\theta_{((L-B)(1-\frac{\alpha}{2}))}$ is the $\left(1-\frac{\alpha}{2}\right)^{\text{th}}$ quantile the of $\theta_{(1)},\theta_{(2)},...,\theta_{(L-B)}$, and θ is a parameter in vector $\boldsymbol{\theta}=(\pi,q,\beta)$.

Moreover, the three intractable integrals above expressions in (16) are evaluated via a numerical analysis which is the *hcubatur()* function in *cubature* package of the R language (R Core Team and Contributors Worldwide 2023).

3.3. Predictive distribution of a future observation

The posterior predictive distribution of a future observation, y_{n+1}, $P\left(Y_{n+1}=y \middle| y_1,y_2,...,y_n\right)$, can be calculated directly by integration,

$$P\left(Y_{n+1}=y \middle| y_1,y_2,...,y_n\right)=\int_0^\infty\int_0^1\int_0^1 p\left(y;\pi,q,\beta\right)p\left(\pi,q,\beta \middle| \underset{\sim}{y},\underset{\sim}{\delta}\right)d\pi\, dq\, d\beta \tag{21}$$

where $p(y;\pi,q,\beta)$ is the probability mass function of ZIDW and $y=0,1,2,\ldots$.

It is immediate that $P(Y_{n+1}=y|y_1,y_2,...,y_n)$ cannot be expressed in closed form and hence it cannot be evaluated analytically. Therefore, to obtain the estimate of $P(Y_{n+1}=y|y_1,y_2,...,y_n)$, we used the Metropolis-Hastings algorithm, to draw the sample from (21) (Singh et al. 2013). We can estimate the future observations under squared error loss function as the mean of simulated sample drawn from (21). The consistent estimate of $P(Y_{n+1}=y|y_1,y_2,...,y_n)$ is obtained as

$$P\left(Y_{n+1}=y \middle| y_1,y_2,...,y_n\right)=\frac{1}{L-B}\sum_{l=B+1}^{L} p\left(y;\pi^{(l)},q^{(l)},\beta^{(l)}\right) \tag{22}$$

where $\pi^{(l)},q^{(l)},\beta^{(l)}$ are the Bayes estimators of the l^{th} time in (19).

Moreover, the *hcubatur()* function in *cubature* package of the R language can be used to obtain an intractable integral in (21).

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 zero-inflated discrete Weibull model under the difference three prior distributions; Beta-Uniform-Uniform prior (Bayes(BUU)), Beta-Jeffreys' rule prior (Bayes(BJ)), and Beta-Beta-Gamma prior (Bayes(BBG)). Moreover, the maximum likelihood estimation (MLE) is considered. The various selected sample sizes (n) are 100, 200, and 300. In particular, this study generates $Y_i \sim ZIDW(\pi, q, \beta)$; $\pi = 0.3, 0.5,$ $q = 0.5,$ and $\beta = 0.5, 1.5$ as follows:

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

If $u \le \pi,$ set $y_i = 0$ with probability π.

If $u > \pi,$ generate y_i using function *rwd()* from package *DWreg* in R language (R Core Team and Contributors Worldwide 2016).

Then, we receive the response variables $y_1, y_2, ..., y_n$ as observed data for the ZIDW model from (2).

The determination of the hyperparameters for prior distribution q of Scenario 3 is considered by the fact that shows the mean of beta distribution, $E(q) = c/(c+d)$ and the variance, $Var(q) = cd \Big/ \left((c+d+1)(c+d)^2\right)$. For example, it is believed that the mean of q is 0.5 and the variance of q is 0.1; then, c and d from solving of the mean and variance equations are respectively obtained. Similarly, the hyperparameters is selected for prior distribution π. Moreover, the determination of the hyperparameters for prior distribution β of Scenarios 3 is considered by the fact that shows the mean, $E(\beta) = a/b$ and the variance, $Var(\beta) = a/b^2$. For example, it is believed that the mean of β is 1 and the variance of β is 1; then, a and b from solving of the mean and variance equations are obtained, respectively. Additionally, this study considers 10,000 iterations of the sampler and uses the first 10% of the data as burn-in for MCMC algorithm.

The simulation study is repeated 1,000 times. The measures of accuracy for estimators are

(i) the estimates of the parameters (Est.) $= \sum_{l=1}^{1,000} \hat{\pi}_l \Big/ 1,000,$ (23)

(ii) the mean square error (MSE) $= \sum_{l=1}^{1,000} \left(\hat{\pi}_l - \pi\right)^2 \Big/ 1,000,$ (24)

(iii) the coverage probability (CP) $= \#\{LCL_\pi < \pi < UCL_\pi\} / 1,000,$ and (25)

(iv) the average length (AL) $= \sum_{l=1}^{1,000} \left(UCL_{\pi l} - LCL_{\pi l}\right) \Big/ 1,000,$ (26)

where $\hat{\pi}_l$ is an estimator, $LCL_{\pi l}$ and $UCL_{\pi l}$ are the l^{th} lower bound and upper bound for the 95% confidence interval of the l^{th} time, and $\#\{LCL_\pi < \pi < UCL_\pi\}$ is the total of the number of times that π inside the confidence interval. The same measure of accuracy has been applied for the estimators of parameters q and β across each approach, encompassing maximum likelihood estimation (MLE) and Bayesian methods, while considering various prior distributions: Beta-Uniform-Uniform prior (Bayes(BUU)), Beta-Jeffreys' rule prior (Bayes(BJ)), and Beta-Beta-Gamma prior (Bayes(BBG)). The accuracy measures for the estimators include: the parameter estimates (Est.) being in proximity to the true parameter values. A larger sample size leads to lower estimated MSE values. Additionally, the minimum MSE value. The CP typically aligns with the nominal confidence level (95%). With larger sample sizes, the AL of the 95% confidence intervals diminishes. Furthermore, the shortest AL value is also noted.

5. Real Data Application

In this section, we are going to apply the two Bayesian techniques to obtain the predictive distributions for two datasets. Effectiveness of these three prior distributions will be evaluated in detail in the following section. The first dataset is the number of Elephantiasis patients from the Phetchaburi province with 22 observations. The second dataset is the number of visits to a doctor by pregnant women in the first three months of their pregnancies with 189 observations in Baystate Medical Center, Springfield, Massachusetts in 1986.

Moreover, a sample prediction is evaluated by using cross validation check to compare the three prior distributions of Bayesian approch. Given observations $Y_1, Y_2, ..., Y_n,$ the i^{th} observation Y_i is moved from the dataset, $i = 1, 2, ..., n.$ Thus, the remaining data appear as $Y_1, Y_2, ..., Y_{i-1}, Y_{i+1}, ..., Y_n.$ Next, this paper computes the predictive distribution of Y_i for each $i = 1, 2, ..., n$ based on the three scenarios of prior distributions and then calculates the mean of random variable. To compare the three scenarios of prior distributions, it is important to define the following accuracy characteristics (mean of absolute values between observation and estimates (MAVOE)) as

$$MAVOE_{mean} = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - mean_i\right|. \tag{27}$$

Dataset 1 Elephantiasis data from Phuket, Thailand

This data is available under Unhapipat et al. (2018), including the yearly number of new Elephantiasis patients reported in Phuket province from 1992 to 2013, is presented in with 22 observations in the Royal Thai Ministry of Public Health, 2017, as presented in Table 1.

Table 1 Frequency distribution of new Elephantiasis patients in Phuket (1992 to 2013)

Number of new Elephantiasis patients	0	1	2	3	Total
Frequency	14	5	1	2	22

First, we apply the standard chi-square goodness of fit (GOF) test to see if the zero-inflated discrete Weibull distribution can be used to model the above dataset. The hypotheses are given as:

H_0 : The number of new Elephantiasis patients/year follows the zero-inflated discrete Weibull distribution.

H_1 : The number of new Elephantiasis patients/year does not follow the zero-inflated discrete Weibull distribution.

The sample size of this dataset is small, we provide the traceplot, autocorrelation for sampled values and posterior densities for generated $\pi, q,$ and β values by the random walk Metropolis algorithm.

Dataset 2 The low birth weight (lbw) data

This data is available under the *"COUNT"* package, Usage *data(lbw)* in R, including the number of visits to a doctor by pregnant women in the first three months of their pregnancies with 189 observations in Baystate Medical Center, Springfield, Massachusetts in 1986 (Hosmer and Lemeshow 2004) as presented in Table 2.

Table 2 Frequency distribution of the number of visits to a doctor by pregnant women in the first three months

Number of visits	0	1	2	3	4	5	6	Total
Frequency	100	47	30	7	4	0	1	189

First, we apply the standard chi-square goodness of fit (GOF) test to see if the zero-inflated discrete Weibull distribution can be used to model the above dataset. The hypotheses are given as:

H_0 : The number of visits to a doctor by pregnant women in the first three months of their pregnancies follows the zero-inflated discrete Weibull distribution.

H_1 : The number of visits to a doctor by pregnant women in the first three months of their pregnancies does not follow the zero-inflated discrete Weibull distribution.

6. Results

6.1. Results of simulation study

Tables 3-6 provide the estimates of the parameters (Est.) and the mean squared error (MSE), while Tables 7-10 display the 95% coverage probability (CP) and the average length (AL).

Table 3 presents Est. and MSE for $ZIDW(0.3, 0.5, 0.5)$. From the results reveal that the parameter estimates (Est.) obtained from all methods closely align with the true parameter values across various sample sizes. The MSE for almost of estimators decreases as sample increases, except the MSE of parameter estimate π for the three Bayesian approaches; Bayes(BUU), Bayes(BJ), Bayes(BBG). Table 4 displays Est. and MSE for $ZIDW(0.5, 0.5, 0.5)$, indicating that the parameter estimates obtained from all methods are in close agreement with the true parameter values across different sample sizes. The MSE for almost of estimators decreases as sample increases, except the MSE of parameter estimate π for the three Bayesian approaches; Bayes(BUU), Bayes(BJ), Bayes(BBG) as well as the MSE of parameter estimate q for the Bayes(BBG). Additionally, Tables 5-6 present Est. and MSE for $ZIDW(0.3, 0.5, 1.5)$ and $ZIDW(0.5, 0.5, 1.5)$, respectively. The results indicate that the parameter estimates obtained from all methods are in close alignment with the true parameter values across different sample sizes. The MSE analysis indicates that all estimators exhibit a monotonic trend, where an increase in sample sizes corresponds to a decrease in estimated MSE values.

Overall, the Bayes estimators exhibit a lower MSE compared to MLE. The MSE of the Bayes(BBG) consistently surpasses other methods across all scenarios. Furthermore, it has been noted that the performance of Bayes estimates derived from Bayes(BUU) and Bayes(BJ) exhibits considerable similarity. Additionally, the MSE of estimators of q and β in case of $\pi = 0.5$ more than in case of $\pi = 0.3$.

Table 3 Est. and MSE for $ZIDW(0.3, 0.5, 0.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
100	π	0.3936	0.0343	0.3404	0.0050	0.3431	0.0047	0.3271	**0.0025**
	q	0.6035	0.0282	0.5541	0.0076	0.5534	0.0075	0.5310	**0.0037**
	β	0.6262	0.0415	0.5749	0.0144	0.5724	0.0141	0.5430	**0.0067**
200	π	0.3361	0.0234	0.3203	0.0041	0.3226	0.0038	0.3151	**0.0027**
	q	0.5449	0.0141	0.5293	0.0042	0.5293	0.0041	0.5193	**0.0026**
	β	0.5529	0.0144	0.5381	0.0058	0.5371	0.0057	0.5252	**0.0036**
300	π	0.3179	0.0185	0.3093	0.0042	0.3110	0.0039	0.3059	**0.0030**
	q	0.5322	0.0106	0.5239	0.0038	0.5238	0.0036	0.5171	**0.0026**
	β	0.5370	0.0099	0.5294	0.0046	0.5285	0.0044	0.5209	**0.0031**

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

Table 4 Est. and MSE for $ZIDW(0.5, 0.5, 0.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
100	π	0.5470	0.0314	0.5168	0.0068	0.5060	0.0059	0.4834	**0.0041**
	q	0.6059	0.0397	0.5635	0.0121	0.5504	0.0107	0.5117	**0.0036**
	β	0.6450	0.0654	0.5961	0.0258	0.5820	0.0228	0.5288	**0.0070**
200	π	0.5089	0.0237	0.4931	0.0062	0.4856	0.0061	0.4757	**0.0051**
	q	0.5507	0.0215	0.5330	0.0079	0.5252	0.0077	0.5077	**0.0044**
	β	0.5661	0.0231	0.5509	0.0115	0.5433	0.0109	0.5214	**0.0057**
300	π	0.5000	0.0204	0.4855	0.0066	0.4792	0.0069	0.4734	**0.0057**
	q	0.5335	0.0171	0.5202	0.0072	0.5140	0.0072	0.5027	**0.0049**
	β	0.5458	0.0160	0.5356	0.0086	0.5299	0.0084	0.5162	**0.0053**

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

Table 5 Est. and MSE for $ZIDW(0.3, 0.5, 1.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
100	π	0.4514	0.0519	0.3733	0.0142	0.3767	0.0142	0.3376	**0.0060**
	q	0.6834	0.0670	0.5991	0.0245	0.5997	0.0246	0.5452	**0.0086**
	β	2.5191	3.1408	1.9991	0.8573	2.0214	0.9707	1.6922	**0.1584**
200	π	0.3872	0.0338	0.3474	0.0079	0.3497	0.0076	0.3297	**0.0039**
	q	0.6048	0.0325	0.5646	0.0115	0.5647	0.0114	0.5388	**0.0050**
	β	1.9338	0.8300	1.7701	0.3085	1.7731	0.3385	1.6456	**0.0869**
300	π	0.3603	0.0259	0.3336	0.0056	0.3360	0.0053	0.3233	**0.0033**
	q	0.5719	0.0206	0.5467	0.0069	0.5474	0.0067	0.5320	**0.0038**
	β	1.7481	0.2915	1.6710	0.1232	1.6704	0.1184	1.6107	**0.0541**

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

Table 6 Est. and MSE for $ZIDW(0.5, 0.5, 1.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		Est.	MSE	Est.	MSE	Est.	MSE	Est.	MSE
100	π	0.6044	0.0433	0.5538	0.0134	0.5392	0.0123	0.4827	**0.0079**
	q	0.7227	0.0989	0.6311	0.0376	0.6131	0.0349	0.5204	**0.0110**
	β	3.0150	5.5439	2.2542	1.5115	2.2394	1.7285	1.6550	**0.1716**
200	π	0.5447	0.0302	0.5207	0.0078	0.5095	0.0072	0.4828	**0.0048**
	q	0.6101	0.0458	0.5743	0.0179	0.5619	0.0162	0.5179	**0.0057**
	β	2.0660	1.4744	1.8750	0.5645	1.8455	0.5875	1.6157	**0.1020**
300	π	0.5192	0.0249	0.5026	0.0058	0.4939	0.0057	0.4789	**0.0042**
	q	0.5660	0.0286	0.5452	0.0104	0.5365	0.0099	0.5111	**0.0045**
	β	1.7861	0.5249	1.7115	0.2189	1.6924	0.2327	1.5802	**0.0650**

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

Table 7 CP and AL for $ZIDW(0.3, 0.5, 0.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		CP	AL	CP	AL	CP	AL	CP	AL
100	π	0.7800	0.9649	0.9940	0.4879	0.9950	0.4869	0.9990	**0.4686**
	q	0.8490	0.7189	0.9900	0.4275	0.9930	0.4337	0.9980	**0.3938**
	β	0.9480	0.7389	0.9810	0.4750	0.9820	0.4832	0.9890	**0.4137**
200	π	0.8650	0.8388	0.9960	0.4424	0.9970	0.4407	0.9990	**0.4273**
	q	0.9170	0.5615	0.9930	0.3508	0.9960	0.3539	0.9980	**0.3325**
	β	0.9620	0.4964	0.9870	0.3416	0.9920	0.3452	0.9970	**0.3177**
300	π	0.8820	0.7065	0.9920	0.4086	0.9930	0.4074	0.9960	**0.3974**
	q	0.9130	0.4651	0.9870	0.3090	0.9930	0.3112	0.9960	**0.2974**
	β	0.9610	0.3976	0.9850	0.2881	0.9870	0.2902	0.9920	**0.2733**

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

Table 8 CP and AL for $ZIDW(0.5, 0.5, 0.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		CP	AL	CP	AL	CP	AL	CP	AL
100	π	0.7670	0.9844	0.9820	**0.5410**	0.9900	0.5557	0.9980	0.5416
	q	0.8170	0.8325	0.9780	0.5242	0.9850	0.5363	0.9970	**0.4843**
	β	0.9470	0.8849	0.9750	0.6038	0.9850	0.6138	0.9990	**0.4977**
200	π	0.8440	0.7978	0.9860	0.4934	0.9890	0.5028	0.9970	**0.4875**
	q	0.8890	0.6482	0.9820	0.4448	0.9880	0.4512	0.9950	**0.4209**
	β	0.9590	0.5906	0.9800	0.4395	0.9840	0.4435	0.9930	**0.3969**
300	π	0.8670	0.6571	0.9760	0.4630	0.9810	0.4694	0.9910	**0.4586**
	q	0.9100	0.5458	0.9730	0.4023	0.9780	0.4063	0.9920	**0.3864**
	β	0.9480	0.4773	0.9750	0.3749	0.9780	0.3774	0.9880	**0.3496**

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

Table 9 CP and AL for $ZIDW(0.3, 0.5, 1.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		CP	AL	CP	AL	CP	AL	CP	AL
100	π	0.6700	1.0876	0.8760	0.4756	0.8770	0.4768	0.9420	**0.4658**
	q	0.7590	0.9479	0.8600	0.4397	0.8620	0.4525	0.9240	**0.4005**
	β	0.9850	6.8851	0.8510	2.0242	0.8530	2.3244	0.9240	**1.3582**
200	π	0.7640	0.9481	0.9570	0.4623	0.9640	0.4626	0.9820	**0.4468**
	q	0.8370	0.7199	0.9570	0.4021	0.9590	0.4077	0.9770	**0.3723**
	β	0.9650	2.6401	0.9560	1.4062	0.9580	1.4625	0.9690	**1.1427**
300	π	0.8120	0.8455	0.9860	0.4458	0.9870	0.4451	0.9890	**0.4298**
	q	0.8780	0.6129	0.9830	0.3686	0.9790	0.3725	0.9870	**0.3460**
	β	0.9610	1.7532	0.9820	1.1319	0.9800	1.1471	0.9890	**1.0014**

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

Table 10 CP and AL for $ZIDW(0.5, 0.5, 1.5)$

n	Parameter	MLE		Bayes(BUU)		Bayes(BJ)		Bayes(BBG)	
		CP	AL	CP	AL	CP	AL	CP	AL
100	π	0.6230	1.1827	0.8160	**0.4943**	0.8360	0.5113	0.9370	0.5270
	q	0.7100	1.2461	0.7810	0.5071	0.8000	0.5284	0.8910	**0.4656**
	β	0.9780	12.9239	0.7750	2.8209	0.7950	3.2793	0.9150	**1.5880**
200	π	0.7570	0.9600	0.9280	**0.4999**	0.9330	0.5136	0.9630	0.5062
	q	0.8340	0.8543	0.9210	0.4920	0.9210	0.5028	0.9600	**0.4563**
	β	0.9810	3.9497	0.9180	1.8906	0.9250	1.9886	0.9530	**1.3738**
300	π	0.8120	0.8706	0.9700	0.4930	0.9760	0.5034	0.9860	**0.4908**
	q	0.8810	0.7310	0.9630	0.4644	0.9690	0.4724	0.9820	**0.4358**
	β	0.9760	2.3099	0.9640	1.4686	0.9740	1.5211	0.9830	**1.2337**

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

Tables 7-8 display the CP and AL for $ZIDW(0.3, 0.5, 0.5)$ and $ZIDW(0.5, 0.5, 0.5)$, respectively. The results demonstrate that the CP of parameters π and q of the MLE is significantly distant from the nominal confidence level. However, with an increase in sample sizes, the CP generally approaches the nominal confidence level, while the CP of parameter β stands close to the nominal confidence level. In all three Bayesian approaches, the CP for all parameters were generally higher than the nominal confidence level and typically approached the nominal confidence level. Furthermore, their behaviors exhibited notable similarities. Concerning the AL for all methods, the AL of the 95% confidence intervals for all parameters decreased with an increase in sample size. Additionally, it is significant to observe that the AL derived from the Bayes(BBG) was the shortest in practically all situations with the exception of case $ZIDW(0.5, 0.5, 0.5)$ at $n = 100$ for parameter π, where the Bayes(BUU) exhibited the shortest length.

Tables 9-10 present the CP and AL for $ZIDW(0.3, 0.5, 1.5)$ and $ZIDW(0.5, 0.5, 1.5)$, respectively. The results demonstrate that the CP of parameters π and q of the MLE is significantly distant from the nominal confidence level. However, with an increase in sample sizes, the CP generally approaches the nominal confidence level, while the CP of parameter β stands close to the nominal confidence level. In all three Bayesian approaches, the CP for almost all parameters were typically approached the nominal confidence level. With the exception of the CP for the Bayes (BUU) and Bayes (BJ) at $n=100$ for all three parameters, which were significantly below the nominal confidence level. Regarding the AL, the result indicates that as the sample size increases, the AL of the 95% confidence intervals for all parameters exhibits a decrease. Furthermore, it is important to note that the AL obtained from the Bayes(BBG) was the shortest in nearly all instances, except for case $ZIDW(0.5,0.5,1.5)$ at $n=100$ and 200 for parameter π, where the Bayes(BUU) showed the shortest length.

6.2. Results of real data application

In this subsection, the predictive distribution of a future observation is presented in Table 12 and Table 14 for dataset 1 and dataset 2 respectively.

Dataset 1 Elephantiasis data from Phuket, Thailand

The maximum likelihood estimates, Bayes estimates, and corresponding confidence intervals are presented in Table 11. The GOF statistic is 2.6984 less than critical values 3.841 by using the zero-inflated discrete Weibull distribution based on MLE (The GOF test value has a p-value of 0.1004, which is greater than 0.05 ($\hat{\pi}_{MLE}=0.4204$, $\hat{q}_{MLE}=0.6274$, $\hat{\beta}_{MLE}=1.49216$)). Thus, these data can be modeled by the zero-inflated discrete Weibull distribution. Additionally, there was a striking similarity in the outcomes of the two Bayesian computational approaches, which were based on the three prior distributions.

Table 11 Estimates and confidence intervals (in brackets) for dataset 1

Methods	π	q	β
MLE	0.4204 (-0.5140, 1.3550)	0.6274 (-0.3229, 1.5777)	1.4922 (-0.8732, 3.8575)
Bayes(BUU)-MH	0.4036 (0.0997, 0.6880)	0.6154 (0.3068, 0.9091)	1.5608 (0.7367, 2.8544)
Bayes(BJ)-MH	0.3989 (0.1029, 0.6876)	0.5909 (0.2787, 0.9112)	1.4844 (0.6704, 2.7886)
Bayes(BBG)-MH	0.3967 (0.1089, 0.6765)	0.5858 (0.3169, 0.8703)	1.4247 (0.6975, 2.3993)
Bayes(BUU)-hcubatur	0.3982	0.6217	1.5866
Bayes(BJ)-hcubatur	0.3961	0.6182	1.5753
Bayes(BBG)-hcubatur	0.3877	0.6040	1.4644

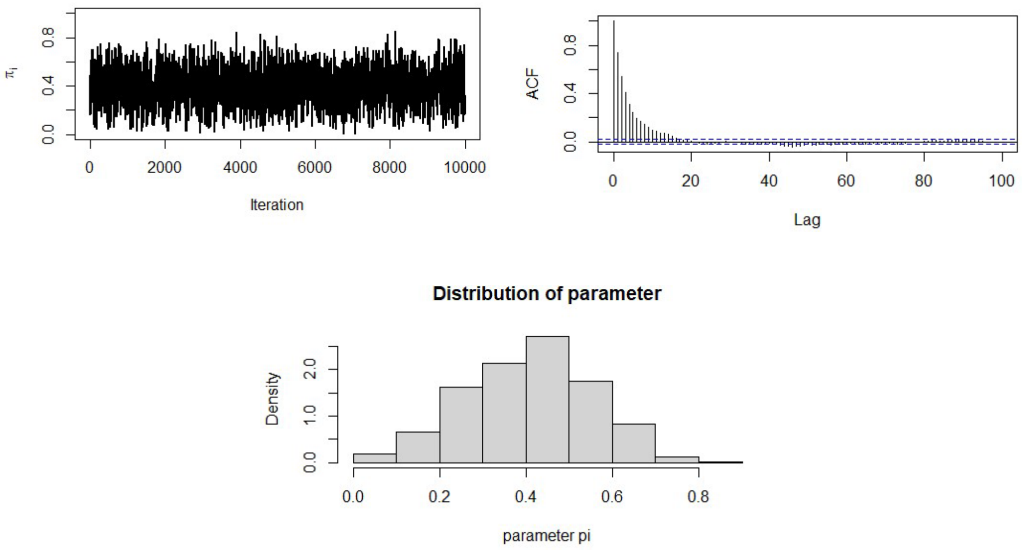

Figure 1 Traceplot, autocorrelation for sampled values and posterior densities for π

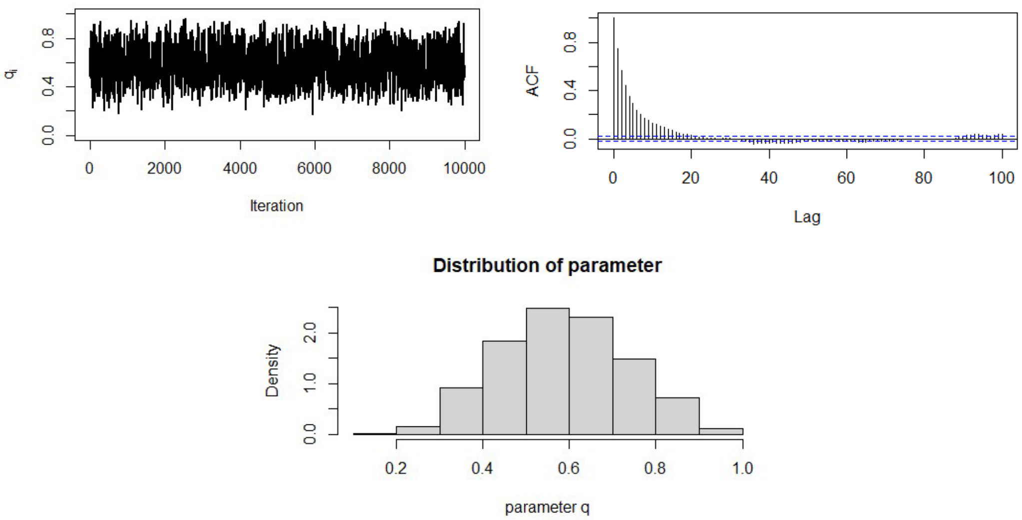

Figure 2 Traceplot, autocorrelation for sampled values and posterior densities for q

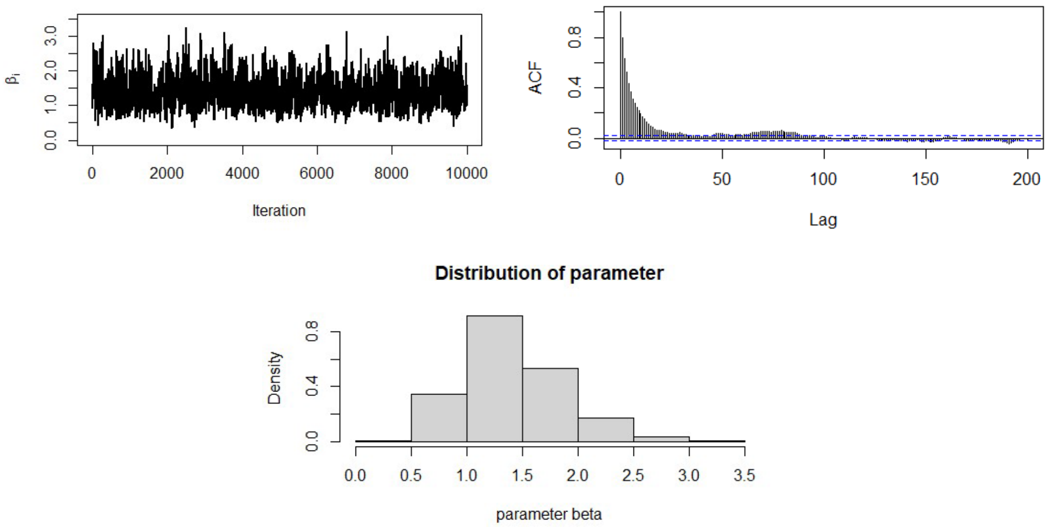

Figure 3 Traceplot, autocorrelation for sampled values and posterior densities for β

Figure 1-3 shows the traceplot, autocorrelation for sampled values and posterior densities for parameters π, q and β, respectively, based on the Beta-Beta-Gamma prior (BBG) of dataset 1. It can be seen that the trace plot showed adequate convergence. Moreover, it clear that sampled values are well mixed and exhibit adequate stability for autocorrelation.

Table 12 Three Baysian predictive distributions of y_{n+1} of dataset 1

y_{n+1}	Predictive distribution by method used the MH algorithm			Predictive distribution by method used the *hcubatur* function		
	Bayes(BUU)	Bayes(BJ)	Bayes(BBG)	Bayes(BUU)	Bayes(BJ)	Bayes(BBG)
0	0.6490	0.6613	0.6583	0.6427	0.6456	0.6443
1	0.1940	0.1863	0.1889	0.1949	0.1891	0.1914
2	0.0988	0.0941	0.0924	0.1019	0.1017	0.0980
3	0.0362	0.0356	0.0357	0.0377	0.0388	0.0387
4	0.0123	0.0125	0.0133	0.0128	0.0135	0.0146
5	0.0047	0.0049	0.0054	0.0049	0.0053	0.0060
6	0.0021	0.0022	0.0025	0.0022	0.0024	0.0028
7	0.0011	0.0011	0.0013	0.0011	0.0013	0.0015
8	0.0006	0.0006	0.0007	0.0006	0.0007	0.0008
9	0.0003	0.0004	0.0004	0.0004	0.0004	0.0005
10	0.0002	0.0002	0.0003	0.0002	0.0003	0.0003
≥ 11	0.0007	0.0008	0.0008	0.0006	0.0009	0.0011

From Table 12, the predictive distributions are heavily skewed with the probability concentrated at 0. Based on the three prior distributions, the behavior of the Bayesian predictive distribution is nearly the same. Besides, take note of how similar the prediction distributions are from the two Bayesian computational methods.

Dataset 2 The low birth weight (lbw) data

The maximum likelihood estimates, Bayes estimates, and corresponding confidence intervals are presented in Table 13. The GOF statistic is 1.96 less than critical values 5.991 by using the zero-inflated discrete Weibull distribution based on MLE (The GOF test value has a p-value of 0.3753, which is greater than 0.05 ($\hat{\pi}_{MLE} = 0.2659,\ \hat{q}_{MLE} = 0.6415,\ \hat{\beta}_{MLE} = 1.4759$)). Thus, these data can be modeled by the zero-inflated discrete Weibull distribution. Additionally, there was a striking similarity in the outcomes of the two Bayesian computational approaches, which were based on the three prior distributions.

Table 13 Estimates and confidence intervals (in brackets) for dataset 2

Methods	π	q	β
MLE	0.2659 (-0.0226, 0.5544)	0.6415 (0.4088, 0.8742)	1.4759 (0.9034, 2.0485)
Bayes(BUU)-MH	0.2476 (0.0615, 0.4205)	0.6309 (0.4831, 0.7837)	1.4677 (1.1227, 1.9117)
Bayes(BJ)-MH	0.2513 (0.0713, 0.4255)	0.6328 (0.4874, 0.7884)	1.4701 (1.1232, 1.9268)
Bayes(BBG)-MH	0.2534 (0.0791, 0.4217)	0.6350 (0.4956, 0.7789)	1.4762 (1.1397, 1.9246)
Bayes(BUU)-hcubatur	0.2286	0.6184	1.4371
Bayes(BJ)-hcubatur	0.2271	0.6170	1.4320
Bayes(BBG)-hcubatur	0.2272	0.6168	1.4289

Table 14 Three Baysian predictive distributions of y_{n+1} of dataset 2

y_{n+1}	Predictive distribution by method used the MH algorithm			Predictive distribution by method used the *hcubatur* function		
	Bayes(BUU)	Bayes(BJ)	Bayes(BBG)	Bayes(BUU)	Bayes(BJ)	Bayes(BBG)
0	0.5288	0.5288	0.5288	0.5288	0.5291	0.5290
1	0.2591	0.2591	0.2591	0.2591	0.2587	0.2586
2	0.1336	0.1336	0.1336	0.1336	0.1334	0.1333
3	0.0532	0.0532	0.0532	0.0532	0.0532	0.0533
4	0.0178	0.0178	0.0178	0.0178	0.0179	0.0180
5	0.0053	0.0053	0.0053	0.0053	0.0054	0.0055
6	0.0015	0.0015	0.0015	0.0015	0.0016	0.0016
7	0.0004	0.0004	0.0004	0.0004	0.0005	0.0005
≥ 8	0.0003	0.0003	0.0003	0.0003	0.0002	0.0002

From Table 14, the predictive distributions are heavily skewed with the probability concentrated at 0. Based on the three prior distributions, the behavior of the Bayesian predictive distribution is nearly the same. Besides, take note of how similar the prediction distributions are from the two Bayesian computational methods.

Table 15 The $MAVOE$ for three Baysian methods of both datasets

Methods	Dataset 1	Dataset 2
Bayes(BUU)	0.7871	0.8446
Bayes(BJ)	0.7852	0.8445
Bayes(BBG)	0.7846	0.8443

Table 15 displays the accuracy characteristics according to the three scenarios of prior distributions. It is important to note that the Beta-Beta-Gamma prior (BBG) presents the best performance.

7. Conclusion and Discussion

This paper is focused on the Bayesian estimation and predictive distribution of a future observation for zero-inflated discrete Weibull distribution based on the three different scenarios of prior distributions; Beta-Uniform-Uniform, Beta-Jeffreys' rule, and Beta-Beta-Gamma. Augmented random walk Metropolis procedure is also proposed to compute the Bayes estimates of the unknown parameters, as well as a numerical method using the *hcubatur()* function in *cubature* package of the R language can be used. Moreover, the parameter estimation has been obtained by maximum likelihood estimation. The performance of among methods are compared by using the Monte Carlo simulation, based on the mean square error and coverage probabilities criteria. These criteria are calculated for different sample sizes, proportion of structural zero, and dispersion data. We have found that Bayesian procedure provides the precise estimates of the unknown parameters with smaller mean square error than classical approach. Estimated coverage probabilities of the three Bayesian approaches are considered as the criteria of a good confidence interval. Additionally, the average length of credible intervals is smaller than that of the asymptotic confidence interval. The Bayesian estimation using MCMC algorithms presents notable challenges when estimating parameters that fall within the interval (0,1). Considering insufficient sampler iterations can lead to approximation errors, resulting in poor convergence and numerical instability, particularly in cases of over-dispersion. The observed mean square error behavior might deviate from theoretical predictions due to their influence.

Along with, the application of the methods is illustrated by using two real datasets available on the literature to evaluate models with different estimates of parameters, as well as the predictive distribution is provided by using cross validation check to compare the three prior distributions of Bayesian approach. Besides, the results of real data analysis are coincided with those in the simulation study. Overall, the Beta-Beta-Gamma prior distribution outperforms other methods. As mentioned earlier, the performance of Bayesian method depends on proportion of structural zero, dispersion of data, and the tail behavior of the corresponding predictive distributions. It is observed that in estimating the parameters and predictive distribution of a future observation at any point, the random walk Metropolis procedure can be used quite effectively.

Acknowledgements

The authors gratefully acknowledge the financial support provided by Faculty of Science and Technology, Thammasat University, Contract No. SciGR 9/2567.

References

Altun E, Alqifari H, Eliwa MS. A novel approach for zero-inflated count regression model: Zero inflated Poisson generalized-Lindley linear model with applications. AIMS Math. 2023; 8(10): 23272-23290.

Allison PD, Waterman RP. Fixed-effects negative binomial regression models. Sociol Methodol. 2002; 32(1): 247-265.

Ashour SK, Muiftah MSA. Bayesian estimation of the parameters of discrete Weibull type (I) distribution. J Mod Appl Stat Methods. 2019; 18(2): 1-13.

Bekalo DB, Kebede DT. Zero-inflated models for count data: an application to number of antenatal care service visits. Ann Data Sci 2021; 8(4): 683-708.

Chaiprasithikul D, Duangsaphon M. Bayesian inference of discrete Weibull regression model for excess zero counts. Sci Technol Asia. 2022; 27(4): 152-174.

Chaiprasithikul D, Duangsaphon M. Bayesian inference for the discrete Weibull regression model with type-I right censored data. Thail Stat. 2022; 20(4): 791-811.

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): 153-162.

CRAN Team, The comprehensive R archive network, 2023. https://cran.r-project.org/.

Dai YH. A perfect example for the BFGS method. Math Program. 2013; 138: 501-530.

Duangsaphon M, Santimalai R, Volodin A. Bayesian estimation and prediction for discrete Weibull distribution. Lobachevskii J Math. 2023; 44(11): 4693-4703.

Duangsaphon M, Sokampang S, Na Bangchang K. Bayesian estimation for median discrete Weibull regression model. AIMS Math. 2024; 9(1): 270-288.

Feng CX. A comparison of zero-inflated and hurdle models for modeling zero-inflated count data. J Stat Distrib Appl. 2021: 8(8): 1-9.

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.

Ghosh JK, Delampady M, Samanta T. An introduction to Bayesian analysis: Theory and methods. New York: Springer; 2006.

Gilks WR, Richardson S, Spiegelhalter D. Markov chain Monte Carlo in practice. Interdisciplinary Statistics. London: Chapman and Hall; 1996.

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): 1085-1105.

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

Hosmer DW, Lemeshow JS. Applied logistic regression. John Wiley & Sons: New York; 2004.

Irfan M, A. K. Sharma MA. Bayesian estimation and prediction for inverse power Maxwell distribution with applications to tax revenue and health care data. J Mod Appl Stat Methods. 2024; 23(1): 1-28.

Kalktawi HS. Discrete Weibull regression model for count data. PhD [dissertation]. London: Brunel University London; 2017.

Kalktawi HS, Vinciotti V, Yu K. A simple and adaptive dispersion regression model for count data. Entropy. 2018; 20(2): 1-15.

Liu H, Davidson RA, Rosowsky DV, Stedinger JR. Negative binomial regression of electric power outages in hurricanes. J Infrastruct Syst. 2005; 11(4): 258-267.

Nakagawa T, Osaki S. The discrete Weibull distribution. IEEE Trans Reliab. 1975; R-24(5): 300-301.

R Core Team and Contributors Worldwide, Cubature, Adaptive multivariate integration over hypercubes, 2023. https://cran.r-project.org/web/packages/cubature/index.html.

R Core Team and Contributors Worldwide, General-Purpose Optimization, 2022.

R Core Team and Contributors Worldwide, Parametric regression for discrete response, 2016. https://CRAN.R-project.org/package=DWreg.

Sellers KF, Shmueli AG. A flexible regression model for count data. Ann Appl Stat. 2010; 4: 943-961.

Serfling RJ. Approximation theorems of mathematical statistics: New York: JohnWiley & Sons; 1980.

Singh KS, Singh U, Sharma VK. Bayesian estimation and prediction for flexible Weibull model under type-II censoring scheme. J Prob Stat. 2013; 2013(146140): 1-16.

Taveekal P, Rajchanuwong P, Wongwiangjan R, Lerdsuwansri R, Intrakul J, Simmachan T, Wongsai S. Modelling road accident injuries and fatalities in Suratthani province of Thailand using Conway-Maxwell-Poisson Regression. Thail Stat. 2023; 21(3): 569-579.

Unhapipat S, Tiensuwan M, Pal N. Bayesian predictive inference for zero-inflated Poisson (ZIP) distribution with applications. Am J Math Manag Sci. 2018; 37(1) 66-79.

Wang S, Chen W, Chen M, Zhou Y. Maximum likelihood estimation of the parameters of the inverse Gaussian distribution using maximum rank set sampling with unequal samples. Math Popul Stud. 2023; 30: 1-21.