Communications in Mathematics and Statistics (2020) 8:255–277
https://doi.org/10.1007/s40304-018-00173-0

Estimation and Prediction for Flexible Weibull Distribution Based on Progressive Type II Censored Data

O. M. Bdair[1] · R. R. Abu Awwad[2] · G. K. Abufoudeh[2] · M. F. M. Naser[1]

Received: 30 August 2018 / Revised: 25 September 2018 / Accepted: 27 December 2018 /
Published online: 13 May 2019

Abstract

In this work, we consider the problem of estimating the parameters and predicting the unobserved or removed ordered data for the progressive type II censored flexible Weibull sample. Frequentist and Bayesian analyses are adopted for conducting the estimation and prediction problems. The likelihood method as well as the Bayesian sampling techniques is applied for the inference problems. The point predictors and credible intervals of unobserved data based on an informative set of data are computed. Markov Chain Monte Carlo samples are performed to compare the so-obtained methods, and one real data set is analyzed for illustrative purposes.

Keywords Flexible Weibull distribution · Progressive censoring data · Bayes estimation · Bayes prediction · Gibbs sampling · Simulation

Mathematics Subject Classification 62N02 · 62N01 · 62G30 · 62F15

1 Introduction

Over the last four decades, thousands of papers dealing with various extensions of the Weibull distribution and their applications have been proposed to enhance the Weibull distribution's capability to fit diverse lifetime data. A common factor among these generalized models is that the Weibull distribution is a special case of theirs. Either the distribution function $F(t)$ or the hazard rate function $h(t)$ of these modified models are related to the corresponding function of the Weibull distribution in someway. Although the hazard rate functions $h(t)$ of these distributions are more capable to model diverse problems than the Weibull model does, these distributions usually do not have simple

✉ M. F. M. Naser
mohammad.naser@bau.edu.jo

1 Faculty of Engineering Technology, Al-Balqa Applied University, Amman, Jordan

2 Department of Mathematics, Faculty of Arts and Sciences, University of Petra, Amman, Jordan

expressions. The hazard function of the Weibull distribution can only be increasing, decreasing or constant. Thus, it cannot be used to model lifetime data with a bathtub-shaped hazard function, such as human mortality and machine life cycles. A detail discussion of Weibull distribution has been provided by Johnson et al. [14].

The Weibull distribution has played an important role in analyzing skewed data, and it is an appropriate model in reliability and life testing problems such as: time to failure or life length of a component or a product. It is a quite useful in various fields ranging from engineering to medical scopes (Lawless [21]). Murthy et al. [26] discussed additional applications and gave a methodological review of the 'Weibull area'. They also suggested further study of various Weibull-type distributions, their properties, related plots, and model selection. Many life distributions for applications in engineering and lifetime analysis require increasing failure rate (IFR), increasing failure rate average (IFRA) or modified bathtub (MBT). The two-parameter flexible Weibull extension proposed by Bebbington et al. [5] has a hazard function that can be increasing, decreasing or bathtub shape. The new two-parameter aging distribution is a generalization of the Weibull distribution namely, the Flexible Weibull Distribution (FWD).

The probability density function (pdf) for FWD is given by:

$$f(x|\alpha,\lambda) = \left(\alpha + \frac{\lambda}{x^2}\right) e^{(\alpha x - \frac{\lambda}{x})}\, e^{-e^{(\alpha x - \frac{\lambda}{x})}}, \; x > 0, \; \alpha, \lambda > 0, \tag{1.1}$$

and the corresponding cumulative distribution function (cdf) is given by:

$$F(x|\alpha,\lambda) = 1 - e^{-e^{(\alpha x - \frac{\lambda}{x})}}. \tag{1.2}$$

The hazard rate of this distribution is given by:

$$h(x) = \left(\alpha + \frac{\lambda}{x^2}\right) e^{(\alpha x - \frac{\lambda}{x})}, \tag{1.3}$$

which is clearly much simpler than other generalizations of Weibull distribution. Bebbington et al. [5] showed that this distribution is able to model various aging classes of the lifetime distributions including IFR, IFRA and MBT. They also checked the goodness of fit of this distribution and found that this distribution gives the better fit among other various extensions of Weibull distribution. Therefore, this distribution can be considered as alternative lifetime distribution of the various well-known generalizations of the Weibull distribution. In fact, FWD has many important contributions like studying the survival analysis for secondary reactor pumps as given in Bebbington et al. [5]. Bebbington et al. [6] introduced another contribution in studying human mortality. Choqueta et al. [9] used the FWD in estimating stop over duration of animals in the presence of trap-effects.

The aim of this paper is mainly twofold. First, we compute the Bayes estimates of α and λ of the FWD under progressively type II censored data within the importance sampler under gamma priors with squared error loss function. We compare the performances of the Bayes estimators based on different censoring schemes with maximum likelihood estimators (MLEs) by extensive computer simulations. Our second aim of this paper is to consider the prediction of future unobservable data based on the current observed data. In this paper, we consider the estimation of the posterior predictive density of future censored data based on current informative data by implementing the importance and Metropolis–Hastings (M–H) algorithms and also construct predictive intervals (PIs) of the future censored data. The main difference of our work with the existing works is that we have considered a comprehensive statistical analysis on FWD under progressively type II censored data including frequentist and Bayesian analysis and moreover.

The progressively type II censored data is considered as a generalization of type II censored data, and it can be described as follows:

Suppose that n units are placed in a life testing experiment and only $m(< n)$ are observed until failure. The censoring occurs progressively in m stages. These m stages offer failure times of the m observed units. At the time of the first failure (the first stage) $X_{1:m:n}$, r_1 of the $n-1$ surviving units are randomly removed (censored) from the experiment. Similarly, at the time of the second failure (the second stage) $X_{2:m:n}$, r_2 of the $n-2-r_1$ surviving units are randomly removed (censored) from the experiment. Finally, at the time of the mth failure (the mth stage) $X_{m:m:n}$, all the remaining $r_m = n-m-(r_1+r_2+\cdots+r_{m-1})$ surviving units are removed from the experiment. Each censoring scheme of the different schemes we will consider in the simulation study is denoted by $(r_1, \ldots, r_m)$. Notice that this scheme includes the type II censoring scheme when $(r_1 = r_2 = \ldots = r_{m-1} = 0, r_m = n-m)$. The model of progressive type II censored data is of importance in the field of reliability and life testing. Progressive type II censored sampling is an important sampling method for obtaining data in lifetime studies where the removal of units prior to failure is preplanned in order to save time and money associated with testing. In other words, when some of the surviving units in the experiment are removed early, the importance of this model can be observed; see for example Cohen [10] who discussed the importance of progressive censoring in life testing reliability experiments.

The problem of prediction is a very important topic in statistics, different applications for the prediction problem were reported in the literature, for more details one may refer to Kaminsky and Rhodin [17], Al-Hussaini [3], and Madi and Raqab [24]. The prediction problem involves the prediction of the future order statistics depending on the observed (informative) sample. Singh et al. [29] obtained the maximum likelihood estimators (MLEs) and Bayes estimators of the parameters of FWD based on the type II censored data, they also used the one-sample and two-sample prediction to predict the missing data in the type II censored sample. Many authors have focused on the problems of estimation and prediction based on various types of censored data from different models (for example, Kim et al. [18], Kundu [19] and Kundu and Raqab [20]).

The scenario of this paper is as follows. In Sect. 2, we obtain the MLEs for the two parameters α and λ of FWD. The Bayes estimates for the two parameters α and λ are derived in Sect. 3. In Sect. 4, we introduce the Bayes prediction problem of the unknown observations from the censored sample. A data analysis and a Monte Carlo simulation that conduct numerical comparisons depending on both very large number of artificial samples and real-life data are performed in Sect. 5.

2 Frequentist Statistical Analysis

In this section, we consider the problem of estimating α, λ using the maximum likelihood (ML) method. Let $\mathbf{x} = (x_{1:m:n}\,,\ x_{2:m:n}\,,\ \ldots\ ,\ x_{m:m:n})$ with $x_{1:m:n} \leq x_{2:m:n} \leq \cdots \leq x_{m:m:n}$ denote the observed progressive type II censored sample of size m from a sample of size n drawn from a FWD with pdf and cdf given in (1.1) and (1.2). The likelihood function based on a progressive type II censored sample $\mathbf{x}$ is given by (see Balakrishnan and Aggrawala [4])

$$L(\alpha, \lambda|\mathbf{x}) = c\prod_{i=1}^{m} f(x_{i:m:n}|\alpha, \lambda)[1 - F(x_{i:m:n}|\alpha, \lambda)]^{r_i}, \tag{2.1}$$

where $c = n(n-1-r_1)(n-2-r_1-r_2)\ldots(n-m+1-r_1-\cdots-r_{m-1})$.

Using (1.1), (1.2) and (2.1), we immediately obtain

$$L(\alpha, \lambda|\mathbf{x}) = c\prod_{i=1}^{m}(\alpha + \frac{\lambda}{x_i^2})\ e^{\sum_{i=1}^{m}\alpha x_i - \frac{\lambda}{x_i}}\ e^{-\sum_{i=1}^{m} e^{\alpha x_i - \frac{\lambda}{x_i}}}\ e^{-\sum_{i=1}^{m} r_i\, e^{\alpha x_i - \frac{\lambda}{x_i}}},$$

where x_i stands for $x_{i:m:n}$. Thus, the profile log-likelihood function can be written as:

$$ln\, L(\alpha, \lambda|\mathbf{x}) = lnc + \sum_{i=1}^{m} ln\left(\alpha + \frac{\lambda}{x_i^2}\right) + \sum_{i=1}^{m}\left(\alpha x_i - \frac{\lambda}{x_i}\right) - \sum_{i=1}^{m}(1 + r_i)\, e^{\alpha x_i - \frac{\lambda}{x_i}}. \tag{2.2}$$

By differentiating (2.2) with respect to α and λ and equating the results by zero, we obtain the normal equations

$$\sum_{i=1}^{m}\frac{1}{\alpha + \frac{\lambda}{x_i^2}} + \sum_{i=1}^{m} x_i - \sum_{i=1}^{m} x_i(1 + r_i)\, e^{\alpha x_i - \frac{\lambda}{x_i}} = 0$$

and

$$\sum_{i=1}^{m}\frac{\frac{1}{x_i^2}}{\alpha + \frac{\lambda}{x_i^2}} - \sum_{i=1}^{m}\frac{1}{x_i} + \sum_{i=1}^{m}\frac{1 + r_i}{x_i}\, e^{\alpha x_i - \frac{\lambda}{x_i}} = 0.$$

The simultaneous solution method of the above two nonlinear equations are the MLEs of α and λ ($\widehat{\alpha}$ and $\widehat{\lambda}$, respectively). Here, the solution cannot be obtained analytically, the well-known Newton-Raphson method is used to evaluate the solutions for both equations. The Newton-Raphson method depends on the Jacobian matrix (J) which contains partial derivatives of the normal equations with respect to both parameters α and λ. The ith iteration of this method can be concluded as follows:

$$\begin{bmatrix} \alpha_{i+1} \\ \lambda_{i+1} \end{bmatrix} = \begin{bmatrix} \alpha_i \\ \lambda_i \end{bmatrix} - J^{-1} * \begin{bmatrix} g_1(\alpha_i, \lambda_i) \\ g_2(\alpha_i, \lambda_i) \end{bmatrix},$$

where

$$J = \begin{bmatrix} \frac{\partial g_1(\alpha,\lambda)}{\partial \alpha} & \frac{\partial g_1(\alpha,\lambda)}{\partial \lambda} \\ \frac{\partial g_2(\alpha,\lambda)}{\partial \alpha} & \frac{\partial g_2(\alpha,\lambda)}{\partial \lambda} \end{bmatrix}_{(\alpha_i,\lambda_i)}$$

and $g_1(\alpha, \lambda)$, $g_2(\alpha, \lambda)$ are the two normal equations, respectively. It should be pointed out here that the numerical solution of the two normal equations exist. This can be verified simply since $\partial^2 lnL(\alpha, \lambda|\mathbf{x})/\partial\alpha^2 < 0$, $\partial^2 lnL(\alpha, \lambda|\mathbf{x})/\partial\lambda^2 < 0$ and $(\partial^2 lnL(\alpha, \lambda|\mathbf{x})/\partial\alpha^2)(\partial^2 lnL(\alpha, \lambda|\mathbf{x})/\partial\lambda^2) - (\partial^2 lnL(\alpha, \lambda|\mathbf{x})/\partial\alpha\partial\lambda)^2 > 0$. For more details, one may refer for example to Casella and Berger [7] page 322.

Since the MLEs of α and β are not in explicit forms, we propose to use the asymptotic normality of α and β to obtain the variances of these estimators. From the log-likelihood function in (2.2), we have

$$\frac{\partial^2 ln\, L}{\partial\alpha^2} = -\sum_{i=1}^{m} \frac{1}{\left(\alpha + \frac{\lambda}{x_i^2}\right)^2} - \sum_{i=1}^{m} x_i^2 (1 + r_i)\, e^{\alpha x_i - \frac{\lambda}{x_i}}$$

and

$$\frac{\partial^2 ln\, L}{\partial\lambda^2} = -\sum_{i=1}^{m} \frac{\frac{1}{x_i^4}}{\left(\alpha + \frac{\lambda}{x_i^2}\right)^2} - \sum_{i=1}^{m} \frac{1 + r_i}{x_i^2}\, e^{\alpha x_i - \frac{\lambda}{x_i}}. \tag{2.3}$$

When the number of censored sample is sufficiently large, then under very general conditions (Lehmann and Casella [22]), the asymptotic normality of the MLEs of $\varphi = (\alpha, \lambda)$ (say, $\widehat{\varphi} = (\widehat{\alpha}, \widehat{\beta})$) can be stated as $\widehat{\varphi} \xrightarrow{D} N_2(\varphi, \mathbf{I}^{-1}(\varphi))$, where $\xrightarrow{D}$ denotes convergence in distribution and $\mathbf{I}(\text{'})$ is the Fisher information matrix obtained by taking the negative of the expectations of the equations in (2.3). In practical applications, one may use the approximation $\widehat{\varphi} \sim N_2(\varphi, \mathbf{J}^{-1}(\widehat{\varphi}))$, where $\mathbf{J}(\widehat{\varphi})$ is the observed information matrix [the negative of the expressions in (2.3)]. Therefore, the $100(1-\gamma)\%$ approximate CIs for α and β are $\left(\widehat{\alpha} - z_{\gamma/2}\sqrt{V_{11}}, \widehat{\alpha} + z_{1-\gamma/2}\sqrt{V_{11}}\right)$ and $\left(\widehat{\beta} - z_{\gamma/2}\sqrt{V_{22}}, \widehat{\beta} + z_{1-\gamma/2}\sqrt{V_{22}}\right)$, respectively, where V_{11} and V_{22} are the elements of the main diagonal of $\mathbf{J}^{-1}(\widehat{\varphi})$ and z_γ is $(\gamma)100$-th lower percentile of the standard normal distribution.

3 Bayesian Statistical Analysis

In this section, we present the posterior densities of the parameters α and λ based on progressive type II censored data and then obtain the corresponding Bayes estimators (BEs) of these parameters. A natural choice for the priors of α and λ would be to assume that the two quantities are independent gamma $G(a, b)$ and $G(c, d)$ distributions, respectively, where $\alpha, \lambda > 0$ and a, b, c, d are chosen to reflect prior knowledge about α and λ. We have used flexible priors on α and λ which allow desirable choice from technical and computational points of view and it will give us adaptable ability to be calibrate between the posterior distribution and prior information. In the Bayesian analysis, the performance of the estimator or predictor depends on the prior distribution and also on the loss function used. For the FWD, if λ is known, then natural choice of the prior on α is the conjugate gamma prior. If the parameter λ is unknown, the continuous conjugate priors do not exist, in despite of the existence of a continuous-discrete joint prior distribution as suggested by Selim [28]. Because of the difficulties in evaluating the prior information in the real-life problems, this type of prior information has been widely criticized in the literature (see, for example, see Kaminsky and Krivtsov [15]). The densities of the priors have the forms:

$$\pi_1(\alpha|a,b) = \frac{b^a}{\Gamma(a)}\alpha^{a-1}e^{-b\,\alpha}, \quad \text{and} \quad \pi_2(\lambda|c,d) = \frac{c^d}{\Gamma(d)}\lambda^{c-1}e^{-d\,\lambda}. \tag{3.1}$$

The joint prior of α and λ is given by:

$$\pi(\alpha, \lambda) = \pi_1(\alpha|a, b) \times \pi_2(\lambda|c, d).$$

Therefore, the joint posterior density of α and λ is defined by:

$$\pi_{jp}(\alpha, \lambda|\mathbf{x}) = \frac{L(\alpha, \lambda|\mathbf{x})\ \pi(\alpha, \lambda)}{\int_0^\infty\int_0^\infty L(\alpha, \lambda|\mathbf{x})\ \pi(\alpha, \lambda)\, \mathrm{d}\alpha\, \mathrm{d}\lambda}. \tag{3.2}$$

The Bayes estimator (BE) of any function of α and λ, say $\theta = g(\alpha, \lambda)$ under the squared error loss (SEL) function $L_1(\delta, \theta) = (\delta - \theta)^2$ is defined by:

$$\theta_{BE} = E_{\pi_{jp}}(\theta|\mathbf{x}) = \frac{\int_0^\infty\int_0^\infty g(\alpha, \lambda)\ L(\alpha, \lambda|\mathbf{x})\ \pi(\alpha, \lambda)\, \mathrm{d}\alpha\, \mathrm{d}\lambda}{\int_0^\infty\int_0^\infty L(\alpha, \lambda|\mathbf{x})\ \pi(\alpha, \lambda)\, \mathrm{d}\alpha\, \mathrm{d}\lambda}. \tag{3.3}$$

It is obvious that (3.3) cannot be evaluated analytically and then the explicit forms of the Bayes estimators for α and λ cannot be obtained. For this, we develop an algorithm using the Gibbs sampler method and Metropolis Hasting algorithm to compute the Bayes estimator above and also to construct the credible intervals. To apply the Gibbs sampler method, we need the full conditional posterior distributions of α and λ. The full conditional posterior distribution of α given λ and $\mathbf{x}$ is given by:

$$\pi_1(\alpha|\lambda, \mathbf{x}) \propto \prod_{i=1}^{m} \left(\alpha + \frac{\lambda}{x_i^2}\right) e^{-\alpha(b-\sum_{i=1}^{m} x_i)} e^{-\sum_{i=1}^{m} e^{\alpha x_i - \frac{\lambda}{x_i}}} e^{-\sum_{i=1}^{m} r_i e^{\alpha x_i - \frac{\lambda}{x_i}}} \alpha^{a-1}, \tag{3.4}$$

while the full conditional posterior distribution of λ given α and $\mathbf{x}$ is given by:

$$\pi_2(\lambda|\alpha, \mathbf{x}) \propto \prod_{i=1}^{m} \left(\alpha + \frac{\lambda}{x_i^2}\right) e^{-\lambda(d+\sum_{i=1}^{m} \frac{1}{x_i})} e^{-\sum_{i=1}^{m} e^{\alpha x_i - \frac{\lambda}{x_i}}} e^{-\sum_{i=1}^{m} r_i e^{\alpha x_i - \frac{\lambda}{x_i}}} \lambda^{c-1}. \tag{3.5}$$

It is clear that the above full conditional distributions cannot be reduced to well-known distributions and therefore we can not generate α and λ from these distributions directly by standard methods. So, to generate from these distributions, we use the M-H algorithm (see Metropolis et al. [25] and Hastings [13]) with normal proposal distribution. The idea here is to decrease the rate of rejections as much as possible. The algorithm below depends on using M-H algorithm based on choosing the normal distribution as a proposal distribution is used to find the BEs and also to construct the credible intervals for α and λ. The Gibbs sampler is a Markov Chain Monte Carlo (MCMC) technique which simulates a Markov chain based on the above full conditional distributions. To draw MCMC samples $\{(\alpha_l, \lambda_l), l = 1, 2, \ldots, N\}$ from the joint posterior distribution π_{jp}, and in turn obtain an approximation of θ_{BE} and the corresponding credible intervals, we use the Gibbs sampler according to the following algorithm:

- Step 1: Start with initial values $\alpha = \alpha^{(0)}$ and $\lambda = \lambda^{(0)}$.
- Step 2: Set $j = 1$.
- Step 3: Generate $\alpha^{*(j)}$ and $\lambda^{*(j)}$ from the proposal normal distribution based on $\alpha^{(j-1)}$ and $\lambda^{(j-1)}$, respectively.
- Step 4: Compute the acceptance ratio at j for α:

$$R_\alpha = \frac{\pi_1(\alpha^{*(j)}|\lambda^{(j-1)}, \mathbf{x})}{\pi_1(\alpha^{(j-1)}|\lambda^{(j-1)}, \mathbf{x})}.$$

- Step 5: Generate uniform random number u_α.
- Step 6: If $u_\alpha < R_\alpha$, then $\alpha^{(j)} = \alpha^{*(j)}$ else $\alpha^{(j)} = \alpha^{(j-1)}$.
- Step 7: Compute the acceptance ratio at j for λ:

$$R_\lambda = \frac{\pi_2(\lambda^{*(j)}|\alpha^{(j)}, \mathbf{x})}{\pi_2(\lambda^{(j-1)}|\alpha^{(j)}, \mathbf{x})}.$$

- Step 8: Generate uniform random number u_λ.
- Step 9: If $u_\lambda < R_\lambda$, then $\lambda^{(j)} = \lambda^{*(j)}$ else $\lambda^{(j)} = \lambda^{(j-1)}$.

- Step 10: Set $j = j + 1$.
- Step 11: Repeat steps 3 to 10, N times and obtain MCMC samples $\{(\alpha_i, \lambda_i);\ i = 1, 2, \ldots, N\}$.

Now, depending on the previous algorithm:
We obtain the Bayes estimate of $\theta = g(\alpha, \lambda)$ based on SEL function L_1 as

$$\hat{\theta}_{BE} = \frac{1}{N}\sum_{i=1}^{N} g(\alpha_i, \lambda_i),$$

and obtain the posterior variance of $\theta = g(\alpha, \lambda)$ as

$$\hat{V}ar(\theta|\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N}(\theta_i - \hat{\theta}_{BE})^2.$$

In particular, if $\theta = g(\alpha, \lambda) = \alpha$ (or λ), then

$$\hat{\alpha}_{BE} = \frac{1}{N}\sum_{i=1}^{N}\alpha_i, \quad \text{and} \quad \hat{V}ar(\alpha|\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N}(\alpha_i - \hat{\alpha}_{BE})^2$$

or

$$\hat{\lambda}_{BE} = \frac{1}{N}\sum_{i=1}^{N}\lambda_i, \quad \text{and} \quad \hat{V}ar(\lambda|\mathbf{x}) = \frac{1}{N}\sum_{i=1}^{N}(\lambda_i - \hat{\lambda}_{BE})^2.$$

To compute the credible interval of $\theta = g(\alpha, \lambda) = \alpha$, we order $\alpha_1, \alpha_2, \ldots, \alpha_N$ as $\alpha_{(1)} < \alpha_{(2)} < \cdots < \alpha_{(N)}$ and then the $(1-\beta)100\%$ symmetric credible interval of α is given by

$$\left(\alpha_{([\frac{N\beta}{2}])}, \alpha_{([N(1-\frac{\beta}{2})])}\right),$$

and in the same way, the $(1-\beta)100\%$ symmetric credible interval of λ is given by

$$\left(\lambda_{([\frac{N\beta}{2}])}, \lambda_{([N(1-\frac{\beta}{2})])}\right),$$

where $[x]$ denotes the largest integer less than or equal to x.

4 Bayesian Prediction

Predicting the unobserved or censored observations, from the early observed observations in the same sample, has incredible attention in many fields of life testing and reliability situations. See, for example, Kaminsky and Nelson [16]. The Bayes prediction of the unknown observation from the future sample based on current available (observed) sample, known as informative sample, is an important feature in Bayes analysis. We mainly consider the estimation of posterior predictive density of the kth order $Y_{k:r_j}$ $(k = 1, 2, \ldots, r_j;\ j = 1, 2, \ldots, m)$ based on observed progressive type II censored sample, $\mathbf{x} = (\mathbf{x_{1:m:n}}, \mathbf{x_{2:m:n}}, \ldots, \mathbf{x_{m:m:n}})$ with $x_{1:m:n} \le x_{2:m:n} \le \ldots \le x_{m:m:n}$. Our objective is to provide the prediction of the unknown observation of an experiment based on the results obtained from an informative experiment. The posterior predictive density of $Y_{k:r_j}$ given the observed censored data $\mathbf{x}$ is given by

$$p\left(y_{k:r_j} \mid \text{data}\right) = \int_0^\infty \int_0^\infty f_{Y_{k:r_j}|X}\left(y_{k:r_j} \mid \alpha, \lambda\right) \pi\left(\alpha, \lambda \mid X\right) \mathrm{d}\alpha\, \mathrm{d}\lambda\,,\ y_{k:r_j} > x_{j:m:n},$$

where $f_{Y_{k:r_j}|\mathbf{x}}\left(y_{k:r_j} \mid \alpha, \lambda\right)$ is the conditional density of $Y_{k:r_j}$ given α, λ and data $\mathbf{x}$, see for example Chen et al. [8].

Using the Markovian property of the conditional order statistics, see David and Nagaraja [12], and the fact that the conditional density of $Y_{k:r_j}$ given data $\mathbf{x}$ is just the conditional distribution of the kth order statistic obtained from a sample of size r_j from $G(y) = [F(y) - F(x_{j:m:n})]/[1 - F(x_{j:m:n})],\ y > x_{j:m:n}$, we have

$$f_{Y_{k:r_j}|\mathbf{x}}\left(y_{k:r_j} \mid \alpha, \lambda\right) = c\left(\alpha + \frac{\lambda}{y_{k:r_j}^2}\right)\left[e^{-e^{\left(\alpha x_{j:m:n} - \frac{\lambda}{x_{j:m:n}}\right)}} - e^{-e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)}}\right]^{k-1}$$
$$\times e^{-(r_j - k)e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)} + \left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)} e^{-\left[e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)} - r_j e^{\left(\alpha x_{j:m:n} - \frac{\lambda}{x_{j:m:n}}\right)}\right]}. \tag{4.1}$$

Using the well-known binomial expansion and after some mathematical simplifications, Eq. (4.1) becomes

$$f_{Y_{k:r_j}|\mathbf{x}}\left(y_{k:r_j} \mid \alpha, \lambda\right) = c\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{k-i-1} e^{(r_j - i)e^{\left(\alpha x_{j:m:n} - \frac{\lambda}{x_{j:m:n}}\right)}} e^{(i - r_j)e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)}}$$
$$\times\left(\alpha + \frac{\lambda}{y_{k:r_j}^2}\right) e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)}. \tag{4.2}$$

From (3.3) and (4.2), the posterior predictive density of $Y_{k:r_j}$ given the observed censored data is

$$p\left(y_{k:r_j} \mid \text{data}\right) = c\int_0^\infty \int_0^\infty \sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{k-i-1} e^{(r_j - i)e^{\left(\alpha x_{j:m:n} - \frac{\lambda}{x_{j:m:n}}\right)}} e^{(i - r_j)e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)}}$$
$$\times\left(\alpha + \frac{\lambda}{y_{k:r_j}^2}\right) e^{\left(\alpha y_{k:r_j} - \frac{\lambda}{y_{k:r_j}}\right)} \pi\left(\alpha, \lambda \mid \text{data}\right) \mathrm{d}\alpha\ \mathrm{d}\lambda. \tag{4.3}$$

Under the SEL function L1, the Bayes predictor (BP) of $Y = Y_{k:r_j}$ can be obtained as

$$\begin{aligned} Y_{k:r_j}^{BP1} &= E_p(Y \mid \text{data}) \\ &= c\int_{x_{j:m:n}}^{\infty} y \int_0^{\infty}\int_0^{\infty} \\ &\quad \sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{k-i-1} e^{(r_j-i)e^{\left(\alpha x_{j:m:n}-\frac{\lambda}{x_{j:m:n}}\right)}} e^{(i-r_j)e^{\left(\alpha y-\frac{\lambda}{y}\right)}} \\ &\quad \times\left(\alpha+\frac{\lambda}{y^2}\right) e^{\left(\alpha y-\frac{\lambda}{y}\right)}\pi\,(\alpha,\lambda \mid \text{data})\ \mathrm{d}\alpha\ \mathrm{d}\lambda\ \mathrm{d}y. \end{aligned}$$

The form of the posterior predictive density in (4.3) is not attractable and the computation of the predictive Bayes estimates $E_p(Y \mid \text{data})$ is not an easy task. Consequently, we use the MCMC samples described in Sect. 3 to generate samples from the predictive distributions. Based on MCMC samples $\{(\alpha_l, \lambda_l) : l = 1, 2, \ldots, M\}$ obtained using the Gibbs sampling and M-H methods, the sample-based predictor $\hat{Y}_{k:r_j}^{BP1}$ of $Y = Y_{k:r_j}$ is given by

$$\begin{aligned} \hat{Y}_{k:r_j}^{BP1} &= \frac{c}{M}\sum_{l=1}^{M}\sum_{i=0}^{k-1}\binom{k-1}{i}(-1)^{k-i-1} e^{(r_j-i)e^{\left(\alpha_l x_{j:m:n}-\frac{\lambda_l}{x_{j:m:n}}\right)}} \\ &\quad \times\int_{x_{j:m:n}}^{\infty} y\left(\alpha_l+\frac{\lambda_l}{y^2}\right) e^{\left(\alpha_l y-\frac{\lambda_l}{y}\right)} e^{(i-r_j)e^{\left(\alpha_l y-\frac{\lambda_l}{y}\right)}}\ \mathrm{d}y. \end{aligned} \tag{4.4}$$

Now, let

$$w(x_{j:m:n}, \alpha_l, \lambda_l) = \int_{x_{j:m:n}}^{\infty} y\left(\alpha_l+\frac{\lambda_l}{y^2}\right) e^{\left(\alpha_l y-\frac{\lambda_l}{y}\right)} e^{(i-r_j)e^{\left(\alpha_l y-\frac{\lambda_l}{y}\right)}}\ \mathrm{d}y. \tag{4.5}$$

Clearly, the integral $w(x_{j:m:n}, \alpha_l, \lambda_l)$ is difficult to be written in a closed form. For this, we use Romberg's method to evaluate the integral numerically. To apply Romberg's method, we use the transformation $t = \frac{y-x_{j:m:n}}{1+y-x_{j:m:n}}$ and for simplicity we use x_j instead of $x_{j:m:n}$, this will lead the integral $w(x_{j:m:n}, \alpha_l, \lambda_l)$ to

$$\begin{aligned} &w(x_j, \alpha_l, \lambda_l) \\ &= \int_0^1\left(\alpha_l(x_j+\frac{t}{1-t})+\frac{\lambda_l}{x_j+\frac{t}{1-t}}\right) e^{\left(\alpha_l(x_j+\frac{t}{1-t})-\frac{\lambda_l}{x_j+\frac{t}{1-t}}\right)} \\ &\quad e^{(i-r_j)e^{\left(\alpha_l(x_j+\frac{t}{1-t})-\frac{\lambda_l}{x_j+\frac{t}{1-t}}\right)}}\ \frac{\mathrm{d}t}{(1-t)^2}. \end{aligned} \tag{4.6}$$

Now, the integral $w(x_j, \alpha_l, \lambda_l)$ is simply evaluated using a modification of Romberg's method. The Romberg's method was initially suggested by Romberg [27]; then, the method has many modifications see for example Dahlquist and Björck [11] and the references there in.

Another important aspect of prediction problem is to construct a two-sided predictive interval for $Y = Y_{k:r_j}(k = 1, 2, \ldots, r_j;\ j = 1, 2, \ldots, m)$. For this, we need to obtain the predictive survival function of $Y = Y_{k:r_j}$ at any point $y > x_{j:m:n}$, which is defined as

$$S_{Y|\text{data}}(y \mid \alpha, \lambda) = \int_y^\infty f_{Y|\text{data}}(z \mid \alpha, \lambda)\,\mathrm{d}z$$
$$= c\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{k-i-1}}{(r_j-i)} e^{(r_j-i)e^{\left(\alpha x_{j:m:n}-\frac{\lambda}{x_{j:m:n}}\right)}} e^{(i-r_j)e^{\left(\alpha y-\frac{\lambda}{y}\right)}}.$$

Under the SEL function L1, the predictive survival function of $Y = Y_{k:r_j}$ is presented as

$$S^P_{Y|\text{data}}(y \mid \alpha, \lambda) = \int_0^\infty\int_0^\infty c\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{k-i-1}}{(r_j-i)} e^{(r_j-i)e^{\left(\alpha x_{j:m:n}-\frac{\lambda}{x_{j:m:n}}\right)}} e^{(i-r_j)e^{\left(\alpha y-\frac{\lambda}{y}\right)}}$$
$$\pi(\alpha, \lambda \mid \text{data})\,\mathrm{d}\alpha\mathrm{d}\lambda.$$

It is clear that the above expression of the predictive survival function cannot be expressed in a closed form, and hence, it cannot be evaluated analytically. Using the MCMC samples $\{(\alpha_i, \lambda_i)\,;\, i = 1, 2, \ldots sM\}$, the simulation estimator for the predictive survival function is given by

$$\hat{S}{}^P_{Y|\text{data}}(y \mid \alpha, \lambda) = \frac{c}{M}\sum_{l=1}^{M}\left[\sum_{i=0}^{k-1}\binom{k-1}{i}\frac{(-1)^{k-i-1}}{(r_j-i)} e^{(r_j-i)e^{\left(\alpha_l x_{j:m:n}-\frac{\lambda_l}{x_{j:m:n}}\right)}} e^{(i-r_j)e^{\left(\alpha_l y-\frac{\lambda_l}{y}\right)}}\right].$$

Now, the $(1-\beta)\,100\%$ predictive interval of $Y = Y_{k:r_j}$ can be found by solving the following nonlinear equations given in (4.7) for the lower bound (L) and upper bound (U) using a suitable numerical technique

$$\hat{S}{}^P_{Y_s|\text{data}}(L) = 1 - \frac{\beta}{2} \text{ and } \hat{S}{}^P_{Y_s|\text{data}}(U) = \frac{\beta}{2}. \tag{4.7}$$

5 Simulations and Data Analysis

To illustrate the above procedures, we present the analysis of two data sets. The first data set is artificial based on very large number of generated samples and the second one is a real-life one. The computations are performed using Mathematica and R packages. It is worth mentioning here that the procedures can be easily applied for any data set.

5.1 Simulations

In this section, we conduct a simulation study to examine the performances of the Bayes estimators comparing with the classical estimators obtained via the maximum likelihood approach based on progressive type II flexible Weibull censored data. Also, we compute the Bayes prediction of the future or missing data based on observed progressive type II censored data. In Bayes estimation, we assume $\{(\alpha, \lambda) = (1, 1), (1, 2), (2, 1), (2, 2), (3, 3)\}$ to generate progressive type II censored data. While in prediction, we assume fixed values of $\alpha = 2, \lambda = 1$, to generate progressive type II censored data. The progressive type II censored data are generated using the algorithm proposed by Balakrishnan and Aggarwala [4]. We compute the Bayes estimates for the flexible Weibull parameters α, λ, under the SEL function (L1). In computing the Bayes estimates, we assume $\pi_1(\alpha)$ and $\pi_2(\lambda)$, the priors of α, λ, have gamma densities with the shape and scale parameters a, c and b, d, respectively. In fact, we consider two types of priors for both α and λ: first prior is the non-informative prior, i.e., $a = b = c = d = 0$, we call this prior as Prior 0, second prior is the informative prior, namely $a = b = 1; c = 2; d = 1$, we call this prior as Prior 1. We compute the mean square errors (MSEs) for the different Bayes estimators based on 10,000 replications. The credible intervals and the coverage percentage for the flexible Weibull parameters α and λ are also computed. For the computations of predictors, we consider only Prior 1 under the SEL function. Based on progressive type II censored data, we obtain the point predictors and the 95% prediction intervals for the missing order statistics $Y_{k:r_j}$, $k = 1, \ldots, r_j$, $j = 1, \ldots, m$ in the prediction problem. The results of the Bayes estimators for the flexible Weibull parameters α and λ and their corresponding CI lengths, when Prior 0 and 1 are used, as well as the results of prediction problem are reported in Tables 1–11 and under the consideration of the following schemes:

- Scheme 1: $n = 25, m = 10, (25, 10, 15, 9 * 0)$
  We remove 15 of the survived items after recording the first failure lifetime, then we record the last 9 failure lifetimes.
- Scheme 2: $n = 25, m = 15, (25, 15, 6 * 0, 5, 5, 7 * 0)$
  We remove 5 of the items after recording the 7th failure lifetime and then we remove additional 5 items after recording the 8th failure lifetime, finally we record the remaining 7 failure lifetimes.
- Scheme 3: $n = 25, m = 20, (25, 20, 19 * 0, 5)$
  We remove the last survived 5 items after recording the 20th failure lifetime.
- Scheme 4: $n = 30, m = 20, (30, 20, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0)$
  We remove 1 of the survived items after recording every new failure lifetime.
- Scheme 5: $n = 40, m = 30, (40, 30, 5, 28 * 0, 5)$
  We remove 5 of the survived items after recording the 1st failure lifetime and still observing the remaining item and then remove additional 5 items after recording the 30th lifetime.

In Tables 1, 2, 3, 4, 5, we present the MLEs for α and λ as well as the Bayes estimates of α and λ, under the loss functions L1, when Prior 0 and Prior 1 are used. Numerical

Table 1 The average estimate and MSEs (in the brackets) for the MLEs and the Bayes of α and λ for scheme (1): $n = 25, m = 10, (15, 9 * 0)$

Parameters	MLE (MSE)	Prior 0	Prior 1
		Bayes (MSE)	Bayes (MSE)
$\alpha = 1$	1.1841 (0.1698)	1.0282 (0.1235)	1.0208 (0.1090)
$\lambda = 1$	1.1026 (0.0926)	1.0427 (0.0779)	1.0299 (0.0692)
$\alpha = 1$	1.1641 (0.1433)	1.0473 (0.1376)	0.9841 (0.0847)
$\lambda = 2$	2.1735 (0.3400)	2.0388 (0.2789)	2.0133 (0.2523)
$\alpha = 2$	2.2708 (0.4680)	2.1078 (0.4296)	1.9606 (0.2975)
$\lambda = 1$	1.0879 (0.0810)	1.0387 (0.0751)	1.0267 (0.0641)
$\alpha = 2$	2.2866 (0.4700)	2.0788 (0.3040)	1.9651 (0.1685)
$\lambda = 2$	2.1814 (0.3453)	1.9506 (0.1502)	2.0271 (0.1413)
$\alpha = 3$	3.3622 (0.8194)	2.9695 (0.2109)	2.9732 (0.2085)
$\lambda = 3$	3.2951 (0.7563)	3.0373 (0.2051)	2.9787 (0.1827)

results of Bayes estimates for α and λ, and their corresponding MSEs, are computed using Algorithm presented in Sect. 3. In Table 6, we present numerical comparisons between the average lengths of the credible intervals along with the coverage percentage of the parameters α and λ when Prior 0 and 1 are used for all schemes considered and for fixed choice of $\alpha = 2$ and $\lambda = 1$. In Tables 7, 8, 9, 10, 11, we present the Bayes point predicted values and the prediction intervals for the missing kth order statistics $Y_{k:r_j}, k = 1, \ldots, r_j, j = 1, \ldots, m$, based on the observed sample of size m with censoring scheme $(r_1, r_2, \ldots, r_m)$, for all schemes described above under the loss function L1 for fixed choice of $\alpha = 2$ and $\lambda = 1$. Based on MCMC samples $\{(\alpha_i, \lambda_i), i = 1, 2, \ldots, M\}$, $M = 10{,}000$, the Bayes point prediction for the missing order statistics $Y_{k:r_j}$ in censoring stage j, $k = 1, 2, \ldots, r_j$, are computed under the loss function L1. The 95% lower bound L and upper bound U of prediction interval for the missing kth order statistics $Y_{k:r_j}$ are also computed.

From Tables 1, 2, 3, it is clear that as m increases, the performances of MLEs of α and λ become better in terms of the MSEs values. The Bayes estimates of α and λ obtained using Prior 0 based on SEL function L1 perform much better than the MLEs of α and λ in terms of MSEs for all schemes considered and under the different choices of α and λ. The Bayes estimates of α and λ obtained using Prior 1 (informative prior) under SEL function L1 are much better than those calculated using Prior 0 (non-informative prior) for α and λ in terms of MSEs for all schemes considered. It could be pointed out here that the Bayesian method in estimating the distribution's parameters is better than the classical method (MLE) of estimation. It is already expected that when additional information are reported through informative priors (like Prior 1), further refinements of the results have to be observed and this is what we have noticed through the studying of Tables 1, 2, 3.

It is also observed from Tables 1, 2, 3, 4, 5 that Bayes estimators are generally overestimate of α and λ except some restricted cases. When fixing the sample size n, the MSEs of each Bayes estimator decrease as the progressive sample size m increases.

Table 2 The average estimate and MSEs (in the brackets) for the MLEs and the Bayes of α and λ for scheme (2): $n = 25, m = 15, (6 * 0, 5, 5, 7 * 0)$

Parameters	MLE (MSE)	Prior 0	Prior 1
		Bayes (MSE)	Bayes (MSE)
$\alpha = 1$	1.1489 (0.1202)	1.0246 (0.1118)	1.0118 (0.1021)
$\lambda = 1$	1.0797 (0.0670)	1.0154 (0.0606)	1.0224 (0.0578)
$\alpha = 1$	1.1440 (0.1126)	1.0120 (0.0730)	1.0175 (0.0618)
$\lambda = 2$	2.1660 (0.2867)	2.0267 (0.1827)	2.0103 (0.1480)
$\alpha = 2$	2.2502 (0.3889)	2.0926 (0.2275)	1.9735 (0.2031)
$\lambda = 1$	1.0778 (0.0619)	1.0305 (0.0581)	1.0208 (0.0379)
$\alpha = 2$	2.2418 (0.3503)	2.0582 (0.1833)	1.9801 (0.1532)
$\lambda = 2$	2.1687 (0.2633)	2.0214 (0.1453)	2.0219 (0.1243)
$\alpha = 3$	3.3371 (0.6815)	2.9735 (0.1538)	2.9809 (0.1389)
$\lambda = 3$	3.2780 (0.5817)	3.0276 (0.1477)	2.9908 (0.1440)

Table 3 The average estimate and MSEs (in the brackets) for the MLEs and the Bayes of α and λ for scheme (3): $n = 25, m = 20, (19 * 0, 5)$

Parameters	MLE (MSE)	Prior 0	Prior 1
		Bayes (MSE)	Bayes (MSE)
$\alpha = 1$	1.1084 (0.0848)	1.0192 (0.1054)	1.0101 (0.0978)
$\lambda = 1$	1.0707 (0.0631)	1.0122 (0.0538)	1.0204 (0.0514)
$\alpha = 1$	1.1102 (0.0746)	0.9903 (0.0578)	1.0159 (0.0580)
$\lambda = 2$	2.1424 (0.2552)	2.0080 (0.1676)	2.0057 (0.1270)
$\alpha = 2$	2.2206 (0.2945)	2.0179 (0.1973)	1.9809 (0.1666)
$\lambda = 1$	1.0507 (0.0563)	1.0259 (0.0420)	1.0129 (0.0353)
$\alpha = 2$	2.1948 (0.2605)	2.0193 (0.1125)	1.9983 (0.1019)
$\lambda = 2$	2.1520 (0.2419)	2.0155 (0.1183)	1.9939 (0.0877)
$\alpha = 3$	3.3290 (0.5628)	3.0102 (0.1316)	2.9935 (0.0939)
$\lambda = 3$	3.2441 (0.5106)	3.0245 (0.1143)	2.9976 (0.1041)

Also we notice that the Bayes estimators obtained using Prior 1 perform much better than the MLEs of α and λ in all schemes considered. From Table 6, we observe that the average length of the credible intervals for α and λ, when Prior 1 is used, becomes smaller as expected, and decreases as m increases. For both Prior 0 and 1, the simulated probabilities are quite close to 0.95 and it is better when Prior 1 is used over Prior 0. It is observed from Tables 7, 8, 9, 10, 11 that the predicted values for the missing kth order statistics $Y_{k:r_j}$ based on the loss functions L1 are quite close to each other and it is increasing. It is also observed that the predictive values for the missing kth order statistics $Y_{k:r_j}$ fall in their corresponding 95% predictive intervals for all schemes considered. One may also observe that both the lower bounds L and the upper bounds

Table 4 The average estimate and MSEs (in the brackets) for the MLEs and the Bayes of α and λ for scheme (4): $n = 30$, $m = 20$, (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0)

Parameters	MLE (MSE)	Prior 0	Prior 1
		Bayes (MSE)	Bayes (MSE)
$\alpha = 1$	1.1049 (0.0808)	0.9766 (0.0780)	0.9863 (0.0689)
$\lambda = 1$	1.0589 (0.0620)	1.0141 (0.0493)	1.0061 (0.0315)
$\alpha = 1$	1.0973 (0.0633)	0.9780 (0.0496)	0.9839 (0.0424)
$\lambda = 2$	2.1324 (0.1851)	2.0370 (0.1715)	2.0229 (0.1520)
$\alpha = 2$	2.2167 (0.2743)	1.9708 (0.1875)	1.9815 (0.1589)
$\lambda = 1$	1.0833 (0.0540)	0.9782 (0.0422)	1.0213 (0.0354)
$\alpha = 2$	2.1803 (0.2310)	1.9723 (0.1239)	1.9902 (0.1041)
$\lambda = 2$	2.1489 (0.2003)	2.0357 (0.0953)	2.0142 (0.0821)
$\alpha = 3$	3.2345 (0.4274)	2.9623 (0.1205)	2.9841 (0.1053)
$\lambda = 3$	3.2044 (0.4065)	3.0112 (0.1062)	3.0104 (0.0983)

Table 5 The average estimate and MSEs (in the brackets) for the MLEs and the Bayes of α and λ for scheme (5): $n = 40$, $m = 30$, (5, 28*0, 5)

Parameters	MLE (MSE)	Prior 0	Prior 1
		Bayes (MSE)	Bayes (MSE)
$\alpha = 1$	1.0838 (0.0679)	1.0423 (0.0539)	1.0148 (0.0474)
$\lambda = 1$	1.0451 (0.0633)	1.0382 (0.0373)	1.0203 (0.0272)
$\alpha = 1$	1.0601 (0.0338)	0.9820 (0.0259)	0.9874 (0.0214)
$\lambda = 2$	2.0983 (0.1429)	2.0290 (0.1232)	2.0152 (0.0983)
$\alpha = 2$	2.1320 (0.1535)	1.9731 (0.1448)	1.9887 (0.1353)
$\lambda = 1$	1.0442 (0.0341)	1.0149 (0.0315)	1.0071 (0.0262)
$\alpha = 2$	2.0982 (0.1250)	2.0201 (0.0721)	2.0063 (0.0617)
$\lambda = 2$	2.0853 (0.1256)	2.0121 (0.0756)	2.0074 (0.0639)
$\alpha = 3$	3.1709 (0.2872)	2.9780 (0.0842)	2.9886 (0.0741)
$\lambda = 3$	3.1554 (0.2959)	3.0101 (0.0601)	3.0037 (0.0583)

U of prediction interval for the missing kth order statistics $Y_{k:r_j}$ increase as $0 \leq k \leq r_j$ increase within each r_j.

5.2 Data Analysis

Here, we discuss the analysis of progressively censored data extracted from practical experiments and conduct a comprehensive simulation study to examine the performances of the sample-based estimates and predictors of the censored data (Table 12–15).

Example 1 In this example, we analyze the data set of time between failures of secondary reactor pumps. Suprawhardana et al. [30] have originally discussed this data set. The chance of the failure of the secondary reactor pump is of the increasing nature

Table 6 Average credible intervals lengths based on progressive type II censored data and the coverage percentages, when Prior 0 and Prior 1 are used

Scheme	Bayes (Prior 0)		Bayes (Prior 1)	
	α	λ	α	λ
Scheme 1: $n = 25, m = 10$	2.1886	1.0386	2.0685	1.0290
	(0.93)	(0.92)	(0.94)	(0.95)
Scheme 2: $n = 25, m = 15$	1.8544	0.9729	1.8248	0.9471
	(0.94)	(0.96)	(0.94)	(0.94)
Scheme 3: $n = 25, m = 20$	1.6932	0.9608	1.6300	0.9270
	(0.98)	(0.98)	(0.98)	(0.97)
Scheme 4: $n = 30, m = 20$	1.7858	0.8971	1.6505	0.8915
	(0.96)	(0.91)	(0.95)	(0.96)
Scheme 5: $n = 40, m = 30$	1.5255	0.8475	1.5006	0.8445
	(0.91)	(0.94)	(0.92)	(0.95)

The first entry represents the average length of the credible intervals, while the corresponding coverage percentages is given between the parentheses

Table 7 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$, for scheme (1): $n = 25$, $m = 10$, $(15, 9 * 0)$

Scheme 1: $n = 25, m = 10$	Predicted values	95% PI
$Y_{1:1:25}$	0.3523	(0.2861, 0.3847)
$Y_{2:1:25}$	0.3953	(0.2966, 0.4003)
$Y_{3:1:25}$	0.4009	(0.3259, 0.4446)
$Y_{4:1:25}$	0.4235	(0.3520, 0.4854)
$Y_{5:1:25}$	0.4384	(0.3784, 0.5239)
$Y_{6:1:25}$	0.4455	(0.4017, 0.5633)
$Y_{7:1:25}$	0.4585	(0.4337, 0.5995)
$Y_{8:1:25}$	0.4717	(0.4601, 0.6337)
$Y_{9:1:25}$	0.5209	(0.4853, 0.6656)
$Y_{10:1:25}$	0.5406	(0.5193, 0.7080)
$Y_{11:1:25}$	0.5984	(0.5425, 0.7617)
$Y_{12:1:25}$	0.6249	(0.5895, 0.8026)
$Y_{13:1:25}$	0.6611	(0.6264, 0.8517)
$Y_{14:1:25}$	0.7044	(0.6637, 0.9200)
$Y_{15:1:25}$	0.7959	(0.7253, 1.0065)

in early stage of the experiment and after that it decreases. The data are studied by several authors, for example, Bebbington et al. [5] showed that flexible Weibull distribution works quite well for these failure time data. The progressively censored sample we suggest is as follows:

i	1	2	3	4	5	6	7
r_i	1	2	2	1	0	2	2
y_i	0.062	0.101	0.273	0.402	0.605	0.614	1.06

The maximum likelihood estimates based on the above data are determined to be $\hat{\alpha} = 0.4682$, $\hat{\lambda} = 0.2902$. The Bayes estimates using prior 1 parameter values turned out to be $\tilde{\alpha}_{\text{Bayes}} = 0.4507$, $\tilde{\lambda}_{\text{Bayes}} = 0.3273$. Further, 95% prediction intervals for α and λ were determined to be, respectively, (0.0526, 1.0556) and (0.1600, 0.5359). The Bayes point predictors and 95% prediction intervals for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$, when prior 1 is used, are presented in Table 13. It is observed that all predicted values with respect to the SEL function are all ordered and fall in their corresponding prediction intervals.

The M-H algorithm is also used to compute the BEs of α and λ. The plots of the full conditional distributions of α and λ in Figs. 1 and 2 show that the normal distribution is an appropriate proposal distribution. Here, we manage the choice of the parameters for the proposal distribution in a way that yields a best fitted model for the full conditional distribution. Therefore, the M-H algorithm with normal proposal distribution is used

Table 8 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$ for scheme (2): $n = 25$, $m = 15$, $(6 * 0, 5, 5, 7 * 0)$

Scheme 2: $n = 25$, $m = 15$	Predicted values	95% PI
$Y_{1:7:25}$	0.6335	(0.5610, 0.6260)
$Y_{2:7:25}$	0.7020	(0.5848, 0.7227)
$Y_{3:7:25}$	0.7386	(0.6282, 0.8258)
$Y_{4:7:25}$	0.9472	(0.7012, 0.9915)
$Y_{5:7:25}$	1.1182	(0.7683, 1.2960)
$Y_{1:8:25}$	0.6084	(0.5634, 0.6298)
$Y_{2:8:25}$	0.6888	(0.6081, 0.7215)
$Y_{3:8:25}$	0.7494	(0.6268, 0.8163)
$Y_{4:8:25}$	0.9501	(0.7192, 1.0276)
$Y_{5:8:25}$	1.1313	(0.7713, 1.3901)

Table 9 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$ for scheme (3): $n = 25$, $m = 20$, $(19 * 0, 5)$

Scheme 3: $n = 25$, $m = 20$	Predicted values	95% PI
$Y_{1:20:25}$	0.6717	(0.5981, 0.7351)
$Y_{2:20:25}$	0.7336	(0.6267, 0.7983)
$Y_{3:20:25}$	0.8050	(0.6560, 0.8328)
$Y_{4:20:25}$	0.9617	(0.7381, 0.9926)
$Y_{5:20:25}$	1.1473	(0.7951, 1.3001)

Table 10 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \le r \le m$ for scheme (4): $n = 30$, $m = 20$, (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0)

Scheme 4: $n = 30, m = 20$	Predicted values	95% PI
$Y_{1:1:30}$	0.434895	(0.3558, 0.8558)
$Y_{1:3:30}$	0.583846	(0.5004, 0.8917)
$Y_{1:5:30}$	0.635808	(0.5477, 0.9032)
$Y_{1:7:30}$	0.674156	(0.6000, 0.9271)
$Y_{1:9:30}$	0.774774	(0.7081, 0.9821)
$Y_{1:11:30}$	0.805874	(0.7286, 0.9938)
$Y_{1:13:30}$	0.826912	(0.7655, 1.0074)
$Y_{1:15:30}$	0.88302	(0.7816, 1.0232)
$Y_{1:17:30}$	0.926959	(0.8525, 1.0628)
$Y_{1:19:30}$	0.990984	(0.9388, 1.1275)

Table 11 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \le r \le m$ for scheme (5): $n = 40$, $m = 30$, (5, 28*0, 5)

Scheme 5: $n = 40, m = 30$	Predicted values	95% PI
$Y_{1:1:40}$	0.4049	(0.2567, 0.4283)
$Y_{2:1:40}$	0.5288	(0.3733, 0.5711)
$Y_{3:1:40}$	0.6455	(0.4919, 0.7028)
$Y_{4:1:40}$	0.7647	(0.5280, 0.8454)
$Y_{5:1:40}$	0.8852	(0.6475, 1.0380)
$Y_{1:30:40}$	0.9501	(0.9293, 0.9784)
$Y_{2:30:40}$	1.0184	(0.9194, 1.1297)
$Y_{3:30:40}$	1.0679	(0.9582, 1.1614)
$Y_{4:30:40}$	1.1068	(0.8851, 1.2074)
$Y_{5:30:40}$	1.1829	(1.0356, 1.2406)

Table 12 MLEs and Bayes estimates for the real data

	MLE		Bayes (Prior 1)	
	α	λ	α	λ
Scheme: $n = 17, m = 7$	0.4682 (0.0799)	0.2902 (0.0390)	0.4507 (0.0697)	0.3273 (0.0197)

to generate numbers from the target probability distribution. In this work, we assume the initial value of α and λ to be their MLEs, while the variances of α and λ to be the reciprocal of Fisher information. Generation of 10,000 random variates is performed. It is checked that the acceptance rates for these choices of variance are computed to be 73.34% and 78.56% which are quite satisfactory. We discarded the initial 500 burn-in samples and computed the BEs based on the remaining observations.

Table 13 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$ for the real data

$n = 17, m = 7$	Predicted values	95% PI
$Y_{1:1:17}$	0.2254	(0.0789, 0.6029)
$Y_{1:2:17}$	0.2567	(0.1082, 0.6369)
$Y_{2:2:17}$	0.3697	(0.1938, 0.6827)
$Y_{1:3:17}$	0.3523	(0.2819, 0.7097)
$Y_{2:3:17}$	0.4308	(0.3256, 0.8283)
$Y_{1:4:17}$	0.4688	(0.4259, 1.1776)
$Y_{1:6:17}$	0.5824	(0.5292, 1.2442)
$Y_{2:6:17}$	0.5882	(0.5439, 1.3605)
$Y_{1:7:17}$	1.1068	(1.0886, 2.6183)
$Y_{2:7:17}$	1.1829	(1.1356, 2.7406)

The graphical diagnostics tools like trace plots and ACF (autocorrelation function) plots are used to check the convergence of the M-H algorithm. Figures 1 and 2 show the trace and ACF plots for α and λ. From the trace plots, we can easily observe random scatter about solid lines with fine mixing of the chains for the simulated values of α and λ, while the ACF plots shows that chains have low autocorrelations (ACF is decaying to 0 quickly). Trace and ACF plots are good visual indicators of the mixing property, which indicates how fast the generated chain converges to the stationary distribution. If the trace plots are sticky and the ACF's decay to 0 fairly slowly, this means that the mixing property is not very satisfactory, see for example Li and Ghosh [23]. A high acceptance rate is another indication that the chain converges to the posterior distribution. In these plots, the ACF plots show low autocorrelation and the trace plots are not sticky with high acceptance rates of 73.34% and 78.56%, this means that the chains converge to the posterior distribution in a satisfactory speed based on our choice of normal distribution as a proposal distribution. As a result, these plots indicate the rapid convergence of the M-H algorithm.

Example 2 In this example, we analyze another set of real-life data which has been originally presented by Aarest [1]. The data set contains the times to failure of 50 devices put on life test at time 0. Ahmad and Hussain [2] recently studied this data set and they proved that the FWD fitting these failure time data very well. The progressively censored sample we suggest is as follows:

i	1	2	3	4	5	6	7	8	9	10	11
r_i	0	0	5	0	6	5	5	5	0	5	8
y_i	0.1	0.2	1	3	6	18	45	63	72	75	84

The MLEs based on the above data are determined to be $\hat{\alpha} = 0.4971$, $\hat{\lambda} = 0.0047$. The Bayes estimates using prior 1 parameter values turned out to be $\tilde{\alpha}_{\text{Bayes}} = 0.6031$, $\tilde{\lambda}_{\text{Bayes}} = 0.0054$. Further, 95% prediction intervals for α and λ were determined to

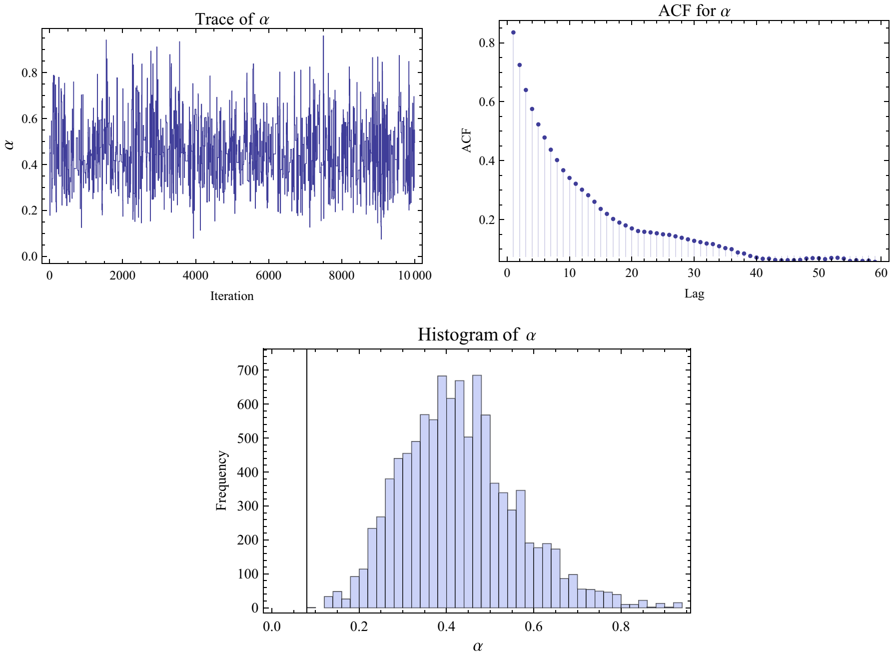

Fig. 1 Plots of Metropolis–Hastings Markov chains for α

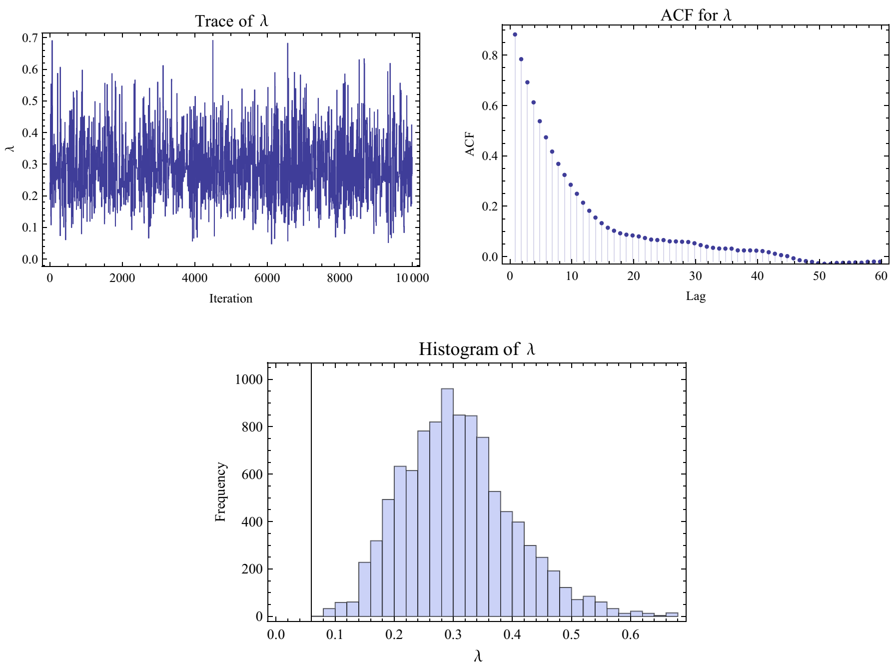

Fig. 2 Plots of Metropolis–Hastings Markov chains for λ

Table 14 MLEs and Bayes estimates for the real data

	MLE		Bayes (Prior 1)	
	α	λ	α	λ
Scheme: $n = 51, m = 11$	0.4971	0.0047	0.6031	0.0054
	(0.1610)	(0.0261)	(0.0602)	(0.0160)

Table 15 Point predictors and PIs for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$ for the real data

$n = 50, m = 11$	Predictedvalues	95% PI
$Y_{1:3:17}$	1.1535	(0.0692, 3.917)
$Y_{3:3:17}$	1.3291	(0.0701, 4.124)
$Y_{5:3:17}$	1.7286	(0.0899, 5.616)
$Y_{1:5:17}$	11.987	(6.711, 18.884)
$Y_{3:5:17}$	16.324	(10.334, 23.627)
$Y_{5:5:17}$	18.803	(13.982, 28.029)
$Y_{1:6:17}$	18.113	(14.823, 27.376)
$Y_{3:6:17}$	31.925	(25.728, 35.757)
$Y_{5:6:17}$	33.919	(29.370, 39.509)
$Y_{1:7:17}$	49.155	(43.736, 57.126)
$Y_{3:7:17}$	55.247	(49.923, 61.728)
$Y_{5:7:17}$	57.949	(53.086, 68.117)
$Y_{1:8:17}$	66.098	(59.097, 73.392)
$Y_{3:8:17}$	71.479	(60.293, 81.709)
$Y_{5:8:17}$	73.022	(62.020, 84.991)
$Y_{1:10:17}$	76.745	(63.732, 87.723)
$Y_{3:10:17}$	80.726	(72.628, 93.167)
$Y_{5:10:17}$	86.196	(81.923, 101.427)
$Y_{1:11:17}$	85.752	(80.236, 100.562)
$Y_{3:11:17}$	87.488	(83.825, 103.452)
$Y_{5:11:17}$	88.559	(83.164, 107.639)
$Y_{7:11:17}$	90.936	(84.876, 111.139)

be, respectively, $(0.1947, 1.1836)$ and $(0.0017, 0.0391)$. The Bayes point predictors and 95% prediction intervals for the missing order statistics $Y_{r:m:n}$, $1 \leq r \leq m$, when prior 1 is used, are presented in Table 15. It is observed that all predicted values with respect to the SEL function are all ordered and fall in their corresponding prediction intervals.

6 Conclusion

We have studied the problem of estimating the parameters of the two-parameter Flexible Weibull distribution as well as the predicting of the censored values under progressive type II censored data. In the estimation problem, we have computed the classical estimators along with the Bayesian estimators for the model parameters. Based on the simulation part of the paper, we have proved that the Bayesian estimation method is better than the classical method of estimation. Furthermore, it has been proven that the Bayesian estimators under informative prior are better than the Bayesian estimators under non-informative and the classical estimators of the parameters. In the prediction problem, we have used the Gibbs sampling and M-H algorithm to predict the unknown observation of an experiment based on the results obtained from an informative experiment, $Y_{i:r_j},\ i = 1, \ldots, r_j$, which are censored after the last observed time. It has been also proven that the predicted values under more informative prior are better than the ones predicted under non-informative prior.

Acknowledgements The authors would like to thank the editor and referees for their comments and very constructive suggestions that contributed to the improvement of this version of the paper.

References

1. Aarset, M.V.: How to identify bathtub hazard rate. IEEE Trans. Reliab. **36**, 106–108 (1978)
2. Ahmad, Z., Hussain, Z.: Flexible Weibull distribution. J. Comput. Math. Sci. **8**(6), 251–260 (2017)
3. Al-Hussaini, E.K.: Predicting observable from a general class of distributions. J. Stat. Plan. Inference **79**, 79–81 (1999)
4. Balakrishnan, N., Aggrawala, R.: Progressive Censoring, Theory, Methods and Applications. Birkhauser, Boston (2000)
5. Bebbington, M., Lai, C.D., Zitikis, R.: A flexible Weibull extension. Reliab. Eng. Syst. Saf. **92**, 719–726 (2007a)
6. Bebbington, M., Lai, C.D., Zitikis, R.: Modeling human mortality using mixtures of bathtub shaped failure rate distributions. J. Theor. Biol. **245**, 528–538 (2007b)
7. Casella, G., Berger, R.: Statistical Inference, 2nd edn. Duxbury, New York (2002)
8. Chen, M.H., Shao, Q.M., Ibrahim, J.G.: Monte Carlo Methods in Bayesian Computation. Springer, New York (2000)
9. Choqueta, R., Guédonb, Y., Besnarda, A., Guillemainc, M., Pradel, R.: Estimating stop over duration in the presence of trap-effects R. Ecol. Modell. **250**, 111–118 (2013)
10. Cohen, A.C.: Progressively censored samples in life testing. Technometrics **5**, 327–329 (1963)
11. Dahlquist, G., Björck, Å.: Numerical methods in scientific computing, Volume I, SIAM (2007)
12. David, H.A., Nagaraja, H.N.: Order Statistics, 3rd edn. Wiley, New York (2003)
13. Hastings, W.K.: Monte Carlo sampling methods using Markov chains and their applications. Biometrika **57**(1), 97–109 (1970)
14. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distribution, 2nd edn. Wiley and Sons, New York (1995)
15. Kaminsky, M.P., Krivtsov, V.V.: A simple procedure for Bayesian estimation of the Weibull distribution. IEEE Trans. Reliab. **54**, 612–616 (2005)
16. Kaminsky, K.S., Nelson, P.I.: Prediction of order statistics. In: Balakrishnan, N., Rao, C.R. (eds.) Hand book of statistics, order statistics: applications, vol. 17, pp. 431–450. North-Holland, Amsterdam (1998)
17. Kaminsky, K.S., Rhodin, L.S.: Maximum likelihood prediction. Ann. Inst. Stat. Math. **37**, 707–717 (1985)
18. Kim, C., Jung, J., Chung, Y.: Bayesian estimation for the exponentiated Weibull model under Type II progressive censoring. Stat. Pap. **52**(1), 53–70 (2011)
19. Kundu, D.: Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. Technometrics **50**, 144–154 (2008)
20. Kundu, D., Raqab, M.Z.: Bayesian inference and prediction for a Type-II censored Weibull distribution. J. Stat. Plan. Inference **142**(1), 41–47 (2012)
21. Lawless, J.F.: Statistical Models and Methods for Life Time Data. Wiley, New York (1982)

22. Lehmann, E.L., Casella, G.: Theory of Point Estimation, 2nd edn. Springer, New York (1998)
23. Li, Y., Ghosh, S.K.: Efficient sampling methods for truncated multivariate normal and Student-t distributions subject to linear inequality constraints. J. Stat.Theory Pract. **9**(4), 712–732 (2015). https://doi.org/10.1080/15598608.2014.996690
24. Madi, M.T., Raqab, M.Z.: Bayesian prediction of temperature records using the Pareto model. Environmetrics **15**, 701–710 (2004)
25. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equations of state calculations by fast computing machines. J. Chem. Phys. **21**(6), 1087–1092 (1953)
26. Murthy, D.N.P., Bulmer, M., Eccleston, J.A.: Weibull model selection for reliability modelling. Reliab. Eng. Syst. Saf. **86**(3), 257–267 (2004)
27. Romberg, W.: Vereinfachte numerische integration. Det Kongelige Norske Videnskabers Selskab Forhandlinger (Trondheim) **28**(7), 30–36 (1955)
28. Selim, M.A.: Bayesian estimations from the two-parameter bathtub-shaped lifetime distribution based on record values. Pak. J. Stat. Oper. Res. **8**(2), 155–165 (2012)
29. Singh, S.K., Singh, U., Sharma, V.K.: Bayesian estimation and prediction for flexible Weibull model under TypeII censoring scheme. J. Probab. Stat. **2013**, 1–16 (2013)
30. Suprawhardana, M.S., Prayoto, S.: Total time on test plot analysis for mechanical components of the RSG-GAS reactor. Atom Indones **25**(2), 81–90 (1999)