Thailand Statistician
July 2023; 21(3): 529-551
http://statassoc.or.th
Contributed paper

Comparison of Estimates and Predictors using Joint Type-II Progressive Censored Samples from two Generalized Rayleigh Distribution

Mohammad Z. Raqab*[a], Omar M. Bdair [b, c] and Debasis Kundu [d]
[a] Department of Mathematics, The University of Jordan, Amman, Jordan
[b] Department of Mathematics and Statistics, McMaster University, Hamilton, ON, Canada
[c] Faculty of Engineering Technology, Al-Balqa Applied University, Amman, Jordan
[d] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, Pin India
* Corresponding author; Email: mraqab@ju.edu.jo; mraqab4@gmail.com

Received: 8 June 2021
Revised: 6 October 2021
Accepted: 7 October 2021

Abstract

Recently, the progressive Type-II censoring has been extended to conduct comparative life-testing experiment of different competing products, which tackles the lifetimes of two samples simultaneously. Here we consider the problem of the joint progressive censoring data coming from the two generalized Rayleigh distributions. The estimation of the unknown parameters and prediction of the life times of the censored units of the joint progressively censored sample are discussed. Frequentist and Bayesian analyses are adopted for conducting the estimation and prediction problems. The likelihood method, bootstrap methods as well as the Bayesian sampling techniques are applied for the inference problems. The point predictors and credible intervals of the times of future failure based on an informative observed censoring units can be developed. Monte Carlo simulations are performed to compare the so developed methods and one real data set is analyzed for illustrative purposes.

Keywords: Bayesian estimation and prediction; Gibbs and Metropolis sampling; highest posterior density credible interval; importance sampling; maximum likelihood estimation; prediction interval.

1. Introduction

The censoring data is of natural interest in survival, reliability and medical studies due to cost or time considerations. Experiment accidental breakage of units or leakage of individuals arise commonly in these studies [See, for example, Lawless (2002)]. The type-II progressively censoring scheme is one of the popular mechanisms of collecting data in lifetime analysis. This type of censoring schemes appears when the experimenter cannot avoid the loss of the test units at points other than the termination point. For this, it has wide industrial applications in reliability and quality. On a detailed discussion on progressive censoring subject, see for example, the monograph by Balakrishnan and Cramer (2014). For inferences on progressive censoring data, one may refer to Mohie El-din and Shafay (2013) and Kotb and Raqab (2019).

The progressive type-II censoring sample can be described as follows. Suppose that n units are placed on a life-testing experiment and only $k(< n)$ units are completely observed until failure.

The censoring occurs progressively in k stages. At the time of the first failure (the first stage), R_1 of the $n-1$ surviving units are randomly withdrawn from the experiment, R_2 of the $n-2-R_1$ surviving units are withdrawn at the time of the second failure (the second stage), and so on. Finally, at the time of the k-th failure (the k-th stage), the test stops with the removal of the remaining $R_k = n - k - \sum_{i=1}^{k-1} R_i$ surviving units. We will refer to this as progressive type-II right censoring scheme $(R_1, R_2, \ldots, R_k)$. The Type-II right censoring and complete sampling schemes are included as special cases by considering ($R_1 = R_2 = \ldots = R_{k-1} = 0$) and ($R_k = n - k$) and ($n = k$ and $R_1 = R_2 = \ldots = R_k = 0$), respectively. Rasouli and Balakrishnan (2010) introduced the joint progressive censoring (JPC) scheme when the data can be produced from two populations.

Under the JPC scheme, the two samples from Population-A (Pop-A) and Population-B (Pop-B) of sizes m and n, respectively, are combined and put on a life testing experiment. Let k be the total number of observed failures in the experiment with $R_1, \ldots, R_k$ being the number of removal units satisfying $\sum_{i=1}^{k}(R_i + 1) = m + n$ where $R_i = S_i + T_i$ with S_i and T_i being the number of removals at the i-th stage from Pop-A and Pop-B, respectively. Based on the combined sample, at the time of the first failure W_1, $R_1 = S_1 + T_1$ units are randomly withdrawn from the remaining $(m + n - 1)$ surviving units where S_1 and T_1 are the number of removed units from Pop-A and Pop-B, respectively. Similarly, at the second stage, $R_2 = S_2 + T_2$ units are withdrawn randomly from the remaining $m + n - 2 - R_1$ surviving units, and so on. Finally, at the time of the k-th failure, all the remaining $R_k = n + m - k - \sum_{i=1}^{k-1} R_i$ surviving units are withdrawn. In this context, the observed data consist of $(\boldsymbol{W}, \boldsymbol{Z}, \boldsymbol{S})$, where $\boldsymbol{W} = (W_1, \ldots, W_k), 1 \leq k < m + n$ being a pre-fixed integer, $\boldsymbol{Z} = (Z_1, \ldots, Z_k)$ with $Z_i = 1$ or 0 accordingly as W_i is taken from an X- or Y-sample and $\boldsymbol{S} = (S_1, \ldots, S_k)$. Let us denote the number of failed units from Pop-A and Pop-B by $k_1 = \sum_{i=1}^{k} Z_i$, and $k_2 = k - k_1 = \sum_{i=1}^{k}(1 - Z_i)$, respectively.

Under the JPC scheme, Parsi and Bairamov (2009) determined the expected number of failures in life testing experiment. Parsi et al. (2011) considered the conditional maximum likelihood and interval estimation of the parameters of two Weibull distributions. Mondal and Kundu (2019) addressed the problem of point and interval estimation of the unknown parameters of two Weibull distributions based on the Bayesian approach. For an elaborate treatment on JPC data and their inferences, one may refer to Balakrishnan et al. (2015), Mondal and Kundu (2020) and Mondal et al. (2020).

In addition to the estimation problem of the unknown parameters, the problem of predicting future occurrences based on the informative sample is another key topic in the statistical inference and it is used extensively in survival and industrial applications, especially, for two-sample prediction problem. Consider a situation where a manufacturer of a product is planning to set-up a warranty for the product to be sent to the market. It is quite useful to use the information based on one sample called informative sample to predict the future failure times in a future sample. An excellent review on the prediction development can be found in Ahsanullah (1980), Nagaraja (1986), Kaminsky and Nelson (1998), Raqab (2001), Barakat et al. (2014, 2018), and Valiollahi et al. (2018).

Although, extensive treatments for estimating parameters from different lifetime distributions based on joint progressive Type-II censored data are available, but no attempt has been made for the estimation and prediction based on joint progressive censored data from the generalized Rayleigh (GR). From Mudholkar et al. (1995), it follows that if $\alpha \leq 1/2$, the hazard function of GR(α, λ) is bathtub type and for $\alpha > 1/2$, it has an increasing hazard function. It is also well-known that for $\alpha \leq 1/2$, the density function is strictly decreasing and for $\alpha > 1/2$, it is unimodal. Shapes of the different probability densities of the GR distribution can be found in Raqab and Kundu (2006). Distributions with decreasing density appear naturally as forward or backward recurrence time distributions in renewal processes. These distributional properties allow a flexibility to the experimenter to fit practical situations where the model contains distributions with unimodal and bathtub failure rates and it is computationally convenient for censored data. The GR distribution with parameters α and $\lambda > 0$, has the cumulative distribution function (CDF) and probability density function (PDF), respectively,

$$F(x;\alpha,\lambda)=(1-e^{-(\lambda x)^2})^{\alpha},\ \alpha,\ \lambda\ >0, \tag{1}$$

and

$$f(x;\alpha,\lambda)=2\ \alpha\ \lambda^2 x\ e^{-(\lambda x)^2}(1-e^{-(\lambda x)^2})^{\alpha-1},\ \alpha,\ \lambda\ >0. \tag{2}$$

Here α and $\lambda>0$ are the shape and scale parameters respectively. From now on the GR distribution with parameters α and λ will be denoted by $GR\ (\alpha,\ \lambda)$. Several aspects of the GR distribution have been studied by Kundu and Raqab (2005) and Raqab and Madi (2011). For some general references on Burr type X distribution, the readers are referred to Sartawi and Abu-Salih (1991), Ahmad, et al.(1997), and the references cited there.

The aims of this paper are mainly described as follows. First we compute the Bayes estimates of all parameters of the GR joint progressive censored data using the importance sampling under gamma priors. We compare the performances of the Bayes estimators with maximum likelihood estimators (MLEs) by extensive computer simulations. We further compute the symmetric credible intervals (CRIs) and compare them with the confidence intervals (CIs) based on the asymptotic and bootstrap (Boot-t) arguments. Our second aim of this paper is to consider the prediction of the life lengths of removed units. In this paper, we also consider the estimation of the posterior predictive density of the removed failed units based on current informative data by implementing the importance and Metropolis-Hastings (M-H) algorithms and also construct prediction intervals (PIs) of the removed units.

The rest of the paper is organized as follows. In Section 2, we describe the Expectation-Maximization (EM) algorithm for determining the MLEs of the scale and shape parameters. Asymptotic properties of the MLEs are also discussed. The Bayes estimates for the shape and scale parameters, respectively are derived in Section 3 using importance sampling. In Section 4, we implement Gibbs and Metropolis sampling to develop sample-based estimates for the predictive density functions of the parameters as well as the times to failure of the $R_j(j=1,\ 2,\ ...,k)$ units still surviving at the time of observation $W_i\ (i=1,2,\ldots,k)$. Section 5 presents a data analysis and a Monte Carlo simulation that perform numerical comparisons. Finally, the findings of the paper are presented in Section 6.

2. Frequentist Statistical Methods

Let $X_1,\ldots,X_m$ be independent and identically distributed (iid) lifetimes of Pop-A with CDF $F(x;\boldsymbol{\theta}_1)$ and PDF $f(x;\boldsymbol{\theta}_1)$. Suppose $Y_1,\ldots,Y_n$ are iid lifetimes of Pop-B with CDF $G(x;\boldsymbol{\theta}_2)$ and PDF $g(x;\boldsymbol{\theta}_2)$. For given a censoring scheme $\boldsymbol{R}=(R_1,\ldots,R_k)$, let $(\boldsymbol{W},\boldsymbol{Z},\boldsymbol{S})=\{(w_1,z_1,s_1),\ldots,(w_k,z_k,s_k)\}$ denotes the JPC data from Pop-A and Pop-B. The likelihood function based on a progressive type II censored sample is given by

$$L(\boldsymbol{\theta}_1,\boldsymbol{\theta}_2;\boldsymbol{w},\boldsymbol{z},\boldsymbol{s})\ \propto\ \prod_{i=1}^{k}[f(w_i;\boldsymbol{\theta}_1)]^{z_i}\ [g(w_i;\boldsymbol{\theta}_2)]^{1-z_i}[1-F(w_i;\boldsymbol{\theta}_1)]^{S_i}[1-G(w_i;\boldsymbol{\theta}_2)]^{T_i},$$
$$w_1<w_2<\ldots<w_k. \tag{3}$$

In our set-up, let X-sample and Y-sample are taken from GR(α_1,λ) and GR(α_2,λ). From (1), (2), and (3), the likelihood function of α_2,α_2 and λ based on JPC sample can be written as

$$\begin{aligned} L(\alpha_1,\alpha_2,\lambda|\text{data})\ &\propto\ \alpha_1^{k_1}\alpha_2^{k_2}\lambda^{2k}\ exp\Bigg\{-\Bigg[\lambda^2\sum_{i=1}^{k}w_i^2+(\alpha_1-1)\sum_{i=1}^{k}z_iD_\lambda(w_i) \\ &+\ (\alpha_2-1)\sum_{i=1}^{k}(1-z_i)D_\lambda(w_i)+\sum_{i=1}^{k}S_i\ Q_{\alpha_1,\lambda}(w_i)+\sum_{i=1}^{k}\ T_iQ_{\alpha_2,\lambda}(w_i)\Bigg]\Bigg\}, \end{aligned}$$

where

$$D_\lambda(w_i) = -log(1-e^{-(\lambda w_i)^2}), \;\; Q_{\alpha_1,\lambda}(w_i) = -log\left[1-\left(1-e^{-(\lambda w_i)^2}\right)^{\alpha_1}\right],$$

and

$$Q_{\alpha_2,\lambda}(w_i) = -log\left[1-\left(1-e^{-(\lambda w_i)^2}\right)^{\alpha_2}\right].$$

Therefore the log-likelihood function can be written as

$$\begin{aligned} l(\alpha_1,\alpha_2,\lambda|\text{data}) &\propto k_1\, log\, \alpha_1 + k_2\, log\, \alpha_2 + 2k\, log\, \lambda - \lambda^2\sum_{i=1}^{k} w_i^2 - (\alpha_1-1)\sum_{i=1}^{k} z_i D_\lambda(w_i) \\ &- (\alpha_2-1)\sum_{i=1}^{k}(1-z_i)D_\lambda(w_i) - \sum_{i=1}^{k} S_i\, Q_{\alpha_1,\lambda}(w_i) - \sum_{i=1}^{k} T_i\, Q_{\alpha_2,\lambda}(w_i). \end{aligned}$$

It can be easily seen that the MLEs of α_1, α_2 and λ do not exist when $k_1 = 0$ or $k_2 = 0$. Under the assumption that $k_1, k_2 > 0$, the MLEs of these parameters can be obtained. Differentiating the log-likelihood function with respect to α_1, α_2 and λ, we obtain the likelihood equations:

$$\frac{k_1}{\alpha_1} - \sum_{i=1}^{k} z_i D_\lambda(w_i) + \sum_{i=1}^{k} S_i \frac{D_\lambda(w_i)}{e^{\alpha_1 D_\lambda(w_i)}-1} = 0, \tag{4}$$

$$\frac{k_2}{\alpha_2} - \sum_{i=1}^{k}(1-z_i) D_\lambda(w_i) + \sum_{i=1}^{k} T_i \frac{D_\lambda(w_i)}{e^{\alpha_2 D_\lambda(w_i)}-1} = 0, \tag{5}$$

and

$$\begin{aligned} \frac{k}{\lambda^2} &- \sum_{i=1}^{k} w_i^2 + (\alpha_1-1)\sum_{i=1}^{k} z_i w_i^2 (e^{D_\lambda(w_i)}-1) + (\alpha_2-1)\sum_{i=1}^{k}(1-z_i) w_i^2 (e^{D_\lambda(w_i)}-1) \\ &- \alpha_1 \sum_{i=1}^{k} S_i w_i^2 \frac{e^{D_\lambda(w_i)}-1}{e^{\alpha_1 D_\lambda(w_i)}-1} - \alpha_2 \sum_{i=1}^{k} T_i w_i^2 \frac{e^{D_\lambda(w_i)}-1}{e^{\alpha_2 D_\lambda(w_i)}-1} = 0. \end{aligned} \tag{6}$$

For complete sample ($R_i = 0$, for all $i = 1, 2, \ldots, k$), the last term in (4) and (5) have to be cancelled and the MLEs of α_1 and α_2 are obtained as functions of λ and (6) can be solved numerically by setting $S_i = T_i = 0$, for all i. In our JPC data, the MLEs of α_1, α_2 and λ have to be obtained by solving the three-dimensional equations in (4), (5) and (6). For this, it is more appropriate to propose the EM algorithm, suggested by Dempster et al. (1977) and used by Ng et al. (2002), to compute the MLEs of α_1, α_2 and λ. Let $\boldsymbol{U} = (\boldsymbol{u}_1, \ ..., \boldsymbol{u}_k)$ and $\boldsymbol{V} = (\boldsymbol{v}_1, \ ..., \boldsymbol{v}_k)$ with $\boldsymbol{u}_j = (u_{j1}, \ ..., \ u_{jS_j})$ and $\boldsymbol{v}_j = (v_{j1}, \ ..., \ v_{jT_j})$, $j = 1, 2, ..., k$, be the censored data from X- and Y-samples, respectively. We consider the censored data as missing data. The combination $(\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v})$, forms the complete data set. The log-likelihood function based on $(\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v})$ is

$$\begin{aligned} l_c(\alpha_1,\alpha_2,\lambda) &\propto m log\, \alpha_1 + n log\, \alpha_2 + 2(m+n) log\, \lambda - \lambda^2[\sum_{i=1}^{k} w_i^2 + \sum_{i=1}^{k}\sum_{j=1}^{S_i} u_{ij}^2 + \sum_{i=1}^{k}\sum_{j=1}^{T_i} v_{ij}^2] \\ &- (\alpha_1-1)[\sum_{i=1}^{k} z_i D_\lambda(w_i) + \sum_{i=1}^{k}\sum_{j=1}^{S_i} D_\lambda(u_{ij})] \\ &- (\alpha_2-1)[\sum_{i=1}^{k}(1-z_i) D_\lambda(w_i) + \sum_{i=1}^{k}\sum_{j=1}^{T_i} D_\lambda(v_{ij})]. \end{aligned} \tag{7}$$

In the E-step, firstly, we compute the pseudo-likelihood function by replacing any function of u_{ij}(say, $g_1(u_{ij})$) by $E(g_1(U_{ij})|U_{ij} > w_i)$ and any function of v_{ij}(say, $g_2(v_{ij})$) by $E(g_2(V_{ij})|V_{ij} > w_i)$. Therefore the pseudo log-likelihood function is

$$\begin{aligned} l_s(\alpha_1, \alpha_2, \lambda) &\propto mlog\ \alpha_1 + nlog\ \alpha_2 + 2(m+n)log\ \lambda \\ &- \lambda^2[\sum_{i=1}^{k} w_i^2 + \sum_{i=1}^{k} S_i E_{1i}(\alpha_1, \lambda) + \sum_{i=1}^{k} T_i E_{2i}(\alpha_2, \lambda)] \\ &- (\alpha_1 - 1)[\sum_{i=1}^{k} z_i D_\lambda(w_i) + \sum_{i=1}^{k} S_i\ E_{3i}(\alpha_1, \lambda)] \\ &- (\alpha_2 - 1)[\sum_{i=1}^{k} (1 - z_i) D_\lambda(w_i) + \sum_{i=1}^{k} T_i\ E_{4i}(\alpha_2, \lambda)], \end{aligned} \tag{8}$$

where $E_{1i}(\alpha_1, \lambda), E_{2i}(\alpha_2, \lambda), E_{3i}(\alpha_1, \lambda)$ and $E_{4i}(\alpha_2, \lambda)$ are defined as

$$\begin{aligned} E_{1i}(\alpha_1, \lambda) &= \int_{w_i}^{\infty} u^2 \frac{f(u, \alpha_1, \lambda)}{1 - F(w_i, \alpha_1, \lambda)} du,\ E_{2i}(\alpha_1, \lambda) = \int_{w_i}^{\infty} v^2 \frac{g(v, \alpha_2, \lambda)}{1 - G(w_i, \alpha_2, \lambda)} dv, \\ E_{3i}(\alpha_1, \lambda) &= \int_{w_i}^{\infty} D_\lambda(u) \frac{f(u, \alpha_1, \lambda)}{1 - F(w_i, \alpha_1, \lambda)} du,\ E_{4i}(\alpha_2, \lambda) = \int_{w_i}^{\infty} D_\lambda(v) \frac{g(v, \alpha_2, \lambda)}{1 - G(w_i, \alpha_2, \lambda)} dv, \end{aligned}$$

where f, F and g, G are the PDF and CDF of the GR distribution from Pop-A and Pop-B, respectively. Secondly, the M-step involves the maximization of the pseudo-likelihood function (8). Let us assume that at the l-th stage, the estimate of $(\alpha_1, \alpha_2, \lambda)$ is $(\alpha_1^{(l)}, \alpha_2^{(l)}, \lambda^{(l)})$. Then $(\alpha_1^{(l+1)}, \alpha_2^{(l+1)}, \lambda^{(l+1)})$ can be obtained by maximizing

$$\begin{aligned} l_s^*(\alpha_1, \alpha_2, \lambda) &\propto mlog\ \alpha_1 + nlog\ \alpha_2 + 2(m+n)log\ \lambda \\ &- \lambda^2[\sum_{i=1}^{k} w_i^2 + \sum_{i=1}^{k} S_i E_{1i}(\alpha_1^{(l)}, \lambda^{(l)}) + \sum_{i=1}^{k} T_i E_{2i}(\alpha_2^{(l)}, \lambda^{(l)})] \\ &- (\alpha_1 - 1)[\sum_{i=1}^{k} z_i D_\lambda(w_i) + \sum_{i=1}^{k} S_i E_{3i}(\alpha_1^{(l)}, \lambda^{(l)})] \\ &- (\alpha_2 - 1)[\sum_{i=1}^{k} (1 - z_i) D_\lambda(w_i) + \sum_{i=1}^{k} T_i E_{4i}(\alpha_2^{(l)}, \lambda^{(l)})], \end{aligned} \tag{9}$$

with respect to α_1, α_2 and λ. It follows from (9), the estimate of α_1 can be obtained as a function of $\alpha_1^{(l)}$ and $\lambda^{(l)}$ and the estimate of α_2 as a function of $\alpha_2^{(l)}$ and $\lambda^{(l)}$ as follows:

$$\begin{aligned} \hat{\alpha_1}_M(\lambda) &= \frac{m}{\sum_{i=1}^{k} z_i D_\lambda(w_i) + \sum_{i=1}^{k} S_i E_{3i}(\alpha_1^{(l)}, \lambda^{(l)})]}, \\ \hat{\alpha_2}_M(\lambda) &= \frac{n}{\sum_{i=1}^{k} (1 - z_i) D_\lambda(w_i) + \sum_{i=1}^{k} T_i E_{4i}(\alpha_2^{(l)}, \lambda^{(l)})]}. \end{aligned} \tag{10}$$

By plugging (10) into (9), we immediately have

$$\begin{aligned} l_s^*(\alpha_1, \alpha_2, \lambda) - l_s^*(\hat{\alpha_1}_M(\lambda), \hat{\alpha_2}_M(\lambda), \lambda) &= m\left\{ log\left(\frac{\alpha_1}{\hat{\alpha_1}_M(\lambda)}\right) - \left(\frac{\alpha_1}{\hat{\alpha_1}_M(\lambda)} - 1\right)\right\} \\ &+ n\left\{ log\left(\frac{\alpha_2}{\hat{\alpha_2}_M(\lambda)}\right) - \left(\frac{\alpha_2}{\hat{\alpha_2}_M(\lambda)} - 1\right)\right\}. \end{aligned}$$

By using the inequality $log\ x \leq (x-1)$ with equality being hold iff $x = 1$, we conclude that

$$l_s^*(\alpha_1, \alpha_2, \lambda) - l_s^*(\hat{\alpha_1}_M(\lambda), \hat{\alpha_2}_M(\lambda), \lambda) \leq 0,$$

and then the maximization of $l_s^*(\hat{\alpha_1}_M, \hat{\alpha_2}_M, \lambda)$ can be obtained by solving the following equation:

$$h(\lambda) = \frac{\lambda^2}{m+n}, \tag{11}$$

where

$$\begin{aligned} h(\lambda) &= \left[\sum_{i=1}^k w_i^2 + \sum_{i=1}^k S_i\ E_{1i}(\alpha_1^{(l)}, \lambda^{(l)}) + \sum_{i=1}^k T_i\ E_{2i}(\alpha_2^{(l)}, \lambda^{(l)})\right. \\ &- \left.(\hat{\alpha_1}_M - 1)\sum_{i=1}^k z_i w_i^2 (e^{D_\lambda(w_i)} - 1) - (\hat{\alpha_2}_M - 1)\sum_{i=1}^k (1-z_i) w_i^2 (e^{D_\lambda(w_i)} - 1)\right]^{-1}. \end{aligned}$$

Once $\lambda^{(l+1)}$ is obtained using (11), $\alpha_1^{(l+1)}$ and $\alpha_2^{(l+1)}$ are obtained as $\alpha_1^{(l+1)} = \hat{\alpha_1}(\lambda^{(l+1)})$ and $\alpha_2^{(l+1)} = \hat{\alpha_2}(\lambda^{(l+1)})$ based on (10). It is clearly noted that $\hat{\alpha_1}_M$, $\hat{\alpha_2}_M$ and $\hat{\lambda}_M$ are not in explicit forms. For this, the asymptotic variances based on the asymptotic normality of $\hat{\alpha_1}_M$, $\hat{\alpha_2}_M$ and $\hat{\lambda}_M$ can be proposed. From the log-likelihood function in (4), we have

$$\begin{aligned} \frac{\partial^2 l}{\partial \alpha_1^2} &= -\frac{k_1}{\alpha_1^2} - \sum_{i=1}^k S_i \frac{D_\lambda^2(w_i)\ e^{\alpha_1 D_\lambda(w_i)}}{(e^{\alpha_1 D_\lambda(w_i)} - 1)^2}, \quad \frac{\partial^2 l}{\partial \alpha_1 \partial \alpha_2} = 0, \\ \frac{\partial^2 l}{\partial \alpha_1 \partial \lambda} &= 2\lambda \sum_{i=1}^k z_i w_i^2 (e^{D_\lambda(w_i)} - 1) + 2\lambda \sum_{i=1}^k S_i w_i^2 \frac{(e^{D_\lambda(w_i)} - 1)[1 - e^{\alpha_1 D_\lambda(w_i)}(1 - \alpha_1 D_\lambda(w_i))]}{(e^{\alpha_1 D_\lambda(w_i)} - 1)^2}, \\ \frac{\partial^2 l}{\partial \alpha_2^2} &= -\frac{k_2}{\alpha_2^2} - \sum_{i=1}^k T_i \frac{D_\lambda^2(w_i)\ e^{\alpha_2 D_\lambda(w_i)}}{(e^{\alpha_2 D_\lambda(w_i)} - 1)^2}, \\ \frac{\partial^2 l}{\partial \alpha_2 \partial \lambda} &= 2\lambda \sum_{i=1}^k (1-z_i) w_i^2 (e^{D_\lambda(w_i)} - 1) + 2\lambda \sum_{i=1}^k T_i w_i^2 \frac{(e^{D_\lambda(w_i)} - 1)[1 - e^{\alpha_2 D_\lambda(w_i)}(1 - \alpha_2 D_\lambda(w_i))]}{(e^{\alpha_2 D_\lambda(w_i)} - 1)^2}, \\ \frac{\partial^2 l}{\partial \lambda^2} &= -\frac{2k}{\lambda^2} - 2\sum_{i=1}^k w_i^2 + 2(\alpha_1 - 1)\sum_{i=1}^k z_i w_i^2 (e^{D_\lambda(w_i)} - 1)(1 - 2\lambda^2 w_i^2 e^{D_\lambda(w_i)}) \\ &+ 2(\alpha_2 - 1)\sum_{i=1}^k (1-z_i) w_i^2 (e^{D_\lambda(w_i)} - 1)(1 - 2\lambda^2 w_i^2 e^{D_\lambda(w_i)}) \\ &- 2\alpha_1 \sum_{i=1}^k S_i w_i^2 \frac{e^{D_\lambda(w_i)} - 1}{(e^{\alpha_1 D_\lambda(w_i)} - 1)^2} \left\{\left[(e^{\alpha_1 D_\lambda(w_i)} - 1)(1 - 2\lambda^2 w_i^2 e^{D_\lambda(w_i)})\right]\right. \\ &+ \left. 2\alpha_1 \lambda^2 w_i^2 e^{\alpha_1 D_\lambda(w_i)} (e^{D_\lambda(w_i)} - 1)\right\} \\ &- 2\alpha_2 \sum_{i=1}^k T_i w_i^2 \frac{e^{D_\lambda(w_i)} - 1}{(e^{\alpha_2 D_\lambda(w_i)} - 1)^2} \left\{\left[(e^{\alpha_2 D_\lambda(w_i)} - 1)(1 - 2\lambda^2 w_i^2 e^{D_\lambda(w_i)})\right]\right. \\ &+ \left. 2\alpha_2 \lambda^2 w_i^2 e^{\alpha_2 D_\lambda(w_i)} (e^{D_\lambda(w_i)} - 1)\right\}. \end{aligned} \tag{12}$$

When the number of observed items k is sufficiently large, then under very general conditions (Lehmann and Casella, 1998), the asymptotic normality of the MLE of $\varphi = (\alpha_1, \alpha_2, \lambda)$ (say,

$\hat{\varphi}_M = (\hat{\alpha_1}_M, \hat{\alpha_2}_M, \hat{\lambda}_M))$ can be stated as $\hat{\varphi}_M \xrightarrow{D} N_3(\varphi, \mathbf{I}^{-1}(\varphi))$, where $\xrightarrow{D}$ denotes convergence in distribution and $\mathbf{I}(\varphi)$ is the Fisher information matrix obtained by taking the negative of the expectations of the Eq.'s in (12). In practical point of view, it is appropriate to use the approximation $\hat{\varphi}_M \sim N_3(\varphi, \mathbf{J}^{-1}(\hat{\varphi}_M))$, where $\mathbf{J}(\hat{\varphi}_M)$ is the observed information matrix (the negative of the expressions in (12)). Therefore, the $100(1-\gamma)\%$ approximate CIs for α_1, α_2 and λ are $\left(\hat{\alpha_1}_M - z_{\gamma/2}\sqrt{V_{11}}, \hat{\alpha_1}_M + z_{1-\gamma/2}\sqrt{V_{11}}\right)$, $\left(\hat{\alpha_2}_M - z_{\gamma/2}\sqrt{V_{22}}, \hat{\alpha_2}_M + z_{1-\gamma/2}\sqrt{V_{22}}\right)$ and $\left(\hat{\lambda}_M - z_{\gamma/2}\sqrt{V_{33}}, \hat{\lambda}_M + z_{1-\gamma/2}\sqrt{V_{33}}\right)$, respectively, where V_{11} and V_{22} and V_{33} are the elements of the main diagonal of $\mathbf{J}^{-1}(\hat{\varphi}_M)$ and z_γ is 100γ-th lower percentile of the standard normal distribution. Usually, the CI based on the asymptotic results do not perform very well for small sample size. For this, we propose CI based on the bootstrap-t method (Boot-t method), see, for example, Ahmed (2014). The Boot-t algorithm can be described as follows:

Step 1: Estimate α_1, α_2 and λ using the maximum likelihood estimation based on the observed informative sample (say $\hat{\alpha_1}_M, \hat{\alpha_2}_M$ and $\hat{\lambda}_M$).

Step 2: Using $\hat{\alpha_1}_M, \hat{\alpha_2}_M$ and $\hat{\lambda}_M$ obtained in Step 1, generate a bootstrap sample and then obtain the first k observed censored units, $B_1, B_2, \ldots, B_k$ under the GR model. Then compute the corresponding MLEs $\hat{\alpha_1}^*_M, \hat{\alpha_2}^*_M$ and $\hat{\lambda}^*_M$ of α_1, α_2 and λ and the elements $(V^*_{11}, V^*_{22}, V^*_{33})$ of the main diagonal of ${J^*}^{-1}(\hat{\alpha_1}^*_M, \hat{\alpha_2}^*_M, \hat{\lambda}^*_M)$.

Step 3: Based on the bootstrap sample in Step 2, define an estimated bootstrap version

$$Q^*_1 = \frac{\hat{\alpha_1}^*_M - \hat{\alpha_2}_M}{\sqrt{V^*_{11}}},\ Q^*_2 = \frac{\hat{\alpha_2}^*_M - \hat{\alpha_2}_M}{\sqrt{V^*_{22}}} \text{ and } Q^*_3 = \frac{\hat{\lambda}^*_M - \hat{\lambda}_{ML}}{\sqrt{V^*_{33}}}.$$

Step 4: Generate M=1000 bootstrap samples and versions of Q^*_1, Q^*_2 and Q^*_3 and then obtain the 100γ-th $100(1-\gamma)$-th sample quantiles of Q^*_1, Q^*_2 and Q^*_3 (say $\hat{q}_{1,\gamma}, \hat{q}_{2,\gamma}$ and $\hat{q}_{3,\gamma}$).

Step 5: Compute the approximate $100(1-\gamma)\%$ CIs for α_1, α_2 and λ as $(\hat{\alpha_1}_M - \hat{q}_{1,1-\gamma/2}\sqrt{V_{11}}, \hat{\alpha_1}_M - \hat{q}_{1,\gamma/2}\sqrt{V_{11}})$, $(\hat{\alpha_2}_M - \hat{q}_{2,1-\gamma/2}\sqrt{V_{22}}, \widehat{\alpha_2}_M - \hat{q}_{2,\gamma/2}\sqrt{V_{22}})$ and $(\hat{\lambda}_M - \hat{q}_{3,1-\gamma/2}\sqrt{V_{33}}, \hat{\lambda}_M - \hat{q}_{3,\gamma/2}\sqrt{V_{33}})$.

3. Bayesian Estimation Methods

In this section, we obtain the posterior densities of the parameters α_1, α_2 and λ based on joint progressive censored sample from two-parameter GR distribution and then obtain the corresponding Bayes estimators of these parameters. To develop these estimates, we consider independent priors of α_1, α_2 and λ (Raqab and Madi (2009)). More specifically, we take $\pi_i(\alpha_i)$, $i = 1, 2$ and $\pi_3(\lambda)$ as gamma (denoted as G(a_i, b_i)) with hyper-parameters a_i and b_i, $i = 1, 2$ and generalized exponential power $GEP(a_3, b_3)$ distributions, respectively, where $\alpha_1, \alpha_2, \lambda > 0$ and a_i's and b_i's are chosen to reflect prior knowledge about α_1, α_2 and λ. These prior densities of $\alpha_i (i = 1, 2)$ and λ take the following forms:

$$p_{a_i,b_i}(\alpha_i) \propto \alpha_i^{a_i-1} e^{-b_i\,\alpha_i}, i = 1, 2,$$

and

$$q_{a_3,b_3}(\lambda) = \lambda^{2a_3-1}\, e^{-b_3\lambda^2}.$$

Note that λ^2 is gamma distribution with parameters c and d with the mean of the λ being $\Gamma(a_3 + 1/2)/\left(\Gamma(a_3)\sqrt{b_3}\right)$. It is important to point out the choices of the priors for α_i and λ are based on the mathematical structure of likelihood of the observed data which turns to be an attractable posterior form where the prior and posterior densities have similar mathematical forms.

By setting $\delta_\lambda(\boldsymbol{w}) = \sum_{i=1}^k z_i D_\lambda(w_i)$, $\bar{\delta}_\lambda(\boldsymbol{w}) = \sum_{i=1}^k (1-z_i) D_\lambda(w_i)$ and combining (4) with the prior densities, we obtain the joint posterior density of α_1, α_2 and λ as

$$\begin{aligned} \pi(\alpha_1, \alpha_2, \lambda|\text{data}) &\propto p_{k_1+a_1, b_1+\delta_\lambda(\boldsymbol{w})}(\alpha_1)\, p_{k_2+a_2, b_2+\bar{\delta}_\lambda(\boldsymbol{w})}(\alpha_2)\, q_{k+a_3, b_3+\sum_{i=1}^k w_i^2}(\lambda) \\ &\times \;(b_1+\delta_\lambda(\boldsymbol{w}))^{-(k_1+a_1)} \left(b_2+\bar{\delta}_\lambda(\boldsymbol{w})\right)^{-(k_2+a_2)} \\ &\times \; e^{\sum_{i=1}^k \left[D_\lambda(w_i) - S_i Q_{\alpha_1,\lambda}(w_i) - T_i Q_{\alpha_2,\lambda}(w_i)\right]}, \end{aligned} \tag{13}$$

where $p_{a,b}(\alpha)$ and $q_{c,d}(\lambda)$ denote the gamma density of α with hyper-parameters a and b and GEP density with hyper-parameters c and d, respectively. Based on the observed joint progressive sample, the marginal density of λ is obtained to be

$$\pi(\lambda|\text{data}) \propto q_{k+a_3, b_3+\sum_{i=1}^k w_i^2}(\lambda)\, \varphi_\lambda(\boldsymbol{w}), \tag{14}$$

where

$$\begin{aligned} \varphi_\lambda(\boldsymbol{w}) &= (b_1+\delta_\lambda(\boldsymbol{w}))^{-(k_1+a_1)} (b_2+\bar{\delta}_\lambda(\boldsymbol{w}))^{-(k_2+a_2)}\, e^{\sum_{i=1}^k D_\lambda(w_i)} \\ & \quad E_{\alpha_1}\left[e^{-\sum_{i=1}^k S_i\, Q_{\alpha_1,\lambda}(w_i)}\right] E_{\alpha_2}\left[e^{-\sum_{i=1}^k T_i\, Q_{\alpha_2,\lambda}(w_i)}\right], \end{aligned}$$

and E_{α_1} and E_{α_2} denote the expectation with respect to $G(k_1+a_1, b_1+\delta_\lambda(\boldsymbol{w}))$ and $G(k_2+a_2, b_2+\bar{\delta}_\lambda(\boldsymbol{w}))$, respectively. Using (14), the Bayes estimate of λ under the squared error loss (SEL) function is

$$\hat{\lambda}_B = \frac{E_\lambda[\lambda\varphi_\lambda(\boldsymbol{w})]}{E_\lambda[\varphi_\lambda(\boldsymbol{w}]}, \tag{15}$$

where E_λ denotes the expectation with respect to $GEP(k+a_3, b_3+\sum_{i=1}^k w_i^2)$. Given λ and data, the marginal density of α_1 and α_2 are, respectively,

$$\pi(\alpha_1|\lambda, \text{data}) \propto p_{k_1+a_1, b_1+\delta_\lambda(\boldsymbol{w})}(\alpha_1)\; e^{-\sum_{i=1}^k S_i\, Q_{\alpha_1,\lambda}(w_i)},$$

and

$$\pi(\alpha_2|\lambda, \text{data}) \propto p_{k_2+a_2, b_2+\bar{\delta}_\lambda(\boldsymbol{w})}(\alpha_2)\; e^{-\sum_{i=1}^k T_i\, Q_{\alpha_2,\lambda}(w_i)}.$$

This in turns out that

$$E(\alpha_1|\lambda, \text{data}) = \frac{E_{\alpha_1}\left[\alpha_1\, e^{-\sum_{i=1}^k S_i Q_{\alpha_1,\lambda}(w_i)}\right]}{E_{\alpha_1}\left[e^{-\sum_{i=1}^k S_i Q_{\alpha_1,\lambda}(w_i)}\right]}, \text{ and } E(\alpha_2|\lambda, \text{data}) = \frac{E_{\alpha_2}\left[\alpha_2\, e^{-\sum_{i=1}^k T_i Q_{\alpha_2,\lambda}(w_i)}\right]}{E_{\alpha_2}\left[e^{-\sum_{i=1}^k S_i Q_{\alpha_2,\lambda}(w_i)}\right]}.$$

As a result of that, the Bayes estimates (BEs) of α_1 and α_2 can be described as

$$\hat{\alpha}_{1B} = E(\alpha_1|\text{data}) = E_\lambda E_{\alpha_1|\lambda}(\alpha_1|\lambda, \text{data}) = \frac{E_\lambda\left[\varphi_\lambda(\boldsymbol{w})E(\alpha_1|\lambda, \text{data})\right]}{E_\lambda\left[\varphi_\lambda(\boldsymbol{w})\right]}, \tag{16}$$

and

$$\hat{\alpha}_{2B} = E(\alpha_2|\text{data}) = E_\lambda E_{\alpha_2|\lambda}(\alpha_2|\lambda, \text{data}) = \frac{E_\lambda\left[\varphi_\lambda(\boldsymbol{w})E(\alpha_2|\lambda, \text{data})\right]}{E_\lambda\left[\varphi_\lambda(\boldsymbol{w})\right]}. \tag{17}$$

It can be easily checked that (15), (16) and (17) cannot be determined analytically and then importance sampler technique can be employed to approximate these expressions and produce consistent sample-based estimates for α_1, α_2 and λ and then construct the corresponding credible intervals. The Bayes point estimates of α_1, α_2 and λ can be computed using the importance sampler algorithm as follows:

Importance sampler algorithm:

1. Generate M λ_i values from $GEP(k+a_3, b_3+\sum_{i=1}^{k} w_i^2)$, $i=1,2,\ldots,M$;
2. For each λ_i generated in Step 1, generate M values of α_1 and α_2 from $G(k_1+a_1, b_1+\delta_\lambda(\boldsymbol{w}))$ and $G(k_2+a_2, b_2+\bar{\delta}_\lambda(\boldsymbol{w}))$, respectively;
3. Compute $E_{\alpha_1}\left[e^{-\sum_{i=1}^{k} S_i Q_{\alpha_1,\lambda_i}(w_i)}\right]$, and $E_{\alpha_1}\left[\alpha_1\ e^{-\sum_{i=1}^{k} S_i Q_{\alpha_1,\lambda_i}(w_i)}\right]$ with respect to the simulated α_1 values in Step 2;
4. Compute $E_{\alpha_2}\left[e^{-\sum_{i=1}^{k} T_i Q_{\alpha_2,\lambda_i}(w_i)}\right]$, and $E_{\alpha_2}\left[\alpha_2\ e^{-\sum_{i=1}^{k} T_i Q_{\alpha_2,\lambda_i}(w_i)}\right]$, with respect to the simulated α_2 values in Step 2;
5. Compute $\varphi_{\lambda_i}(\boldsymbol{w})$, $E(\alpha_1|\lambda_i, \text{data})$ and $E(\alpha_2|\lambda_i, \text{data})$;
6. Average the numerators and the denominators of (15), (16) and (17) with respect to the λ values simulated in Step 1.

Now, we aim at developing two-sided Bayes CRIs as well as highest posterior density (HPD) CRIs for the α_1, α_2 and λ. Based on the simulated values, $(1-\gamma)100\%$ the Bayes CRIs for these $\alpha_1, \alpha_2, \lambda$ may be computed. That is, a $100(1-\gamma)\%$ Bayes CRI for $\theta(\theta = \alpha_1, \alpha_2,$ or $\lambda)$ is $\left(\theta^{(\gamma/2)}, \theta^{(1-\gamma/2)}\right)$, where $\theta^{(\gamma)}$ is 100γ-th percentile of the θ values simulated via the previous algorithm. Since the previous CRIs do not specify whether the values of θ within these intervals have highest probability than that of the values outside the intervals, we propose the HPD CRI for θ. It is defined as the one of the shortest width such that the posterior density of any point outside the interval is less than that for any point inside the intervals. To compute the HPD CRI for any function of the parameters involved (say, $\rho(\alpha_1, \alpha_2, \lambda)$), we use a Monte Carlo method developed by Chen and Shao (1999) for using importance sampling to compute HPD CRIs for the parameters involved. By ordering the simulated values of $\rho(\alpha_1, \alpha_2, \lambda)$, we obtain

$$\rho_{(1)} \leq \rho_{(2)} \leq \ldots \leq \rho_{(M)},$$

then compute the ratios

$$\zeta_i = \frac{\varphi(\lambda_{(i)})}{\sum_{i=1}^{M} \varphi(\lambda_{(i)})}, i = 1, 2, \ldots, M.$$

For M sufficiently large, the $100(1-\gamma)\%$ shortest-width Chen and Shao CRI (C-S CRI) for ρ is the shortest interval among the intervals I_j for $j = 1, 2, \ldots, M - [(1-\gamma)M]$, with

$$I_j = \left(\rho^{(\frac{j}{M})}, \rho^{(\frac{j+[(1-\gamma)M]}{M})}\right),$$

where $[x]$ is the integer part of x and $\rho^{(\gamma)}$ is 100γ-th percentile of ρ, which can be obtained as follows:

$$\rho^{(\gamma)} = \begin{cases} \rho_{(1)} & \text{if } \gamma = 0, \\ \rho_{(i)} & \text{if } \sum_{j=1}^{i-1} \zeta_j \leq \gamma \leq \sum_{j=1}^{i} \zeta_j, \end{cases}$$

Therefore, the C-S CRIs of α_1, α_2 and λ can be obtained accordingly.

Next, we consider M-H algorithm as explained in Tierney (1994). For this, we opt for stochastic simulation procedures, namely, the Gibbs and Metropolis samplers (Gilks, Richardson, and Spiegelhalter, 1995) to generate samples from the posterior distributions. The following algorithm describes the steps to generate sample-based estimates of α_1, α_2 and λ. The MLEs of α_1, α_2 and λ can be considered as initial values. The M-H algorithm proceeds as follows:

M-H algorithm for estimation problem:

1. Start with initial values $(\alpha_1^{(0)}, \alpha_2^{(0)}, \lambda^{(0)})$;
2. Set J = 1;
3. Generate $\alpha_1^{(J)}$ from $G(k_1 + a_1, b_1 + \delta_\lambda(\boldsymbol{w}))$, and Generate $\alpha_2^{(J)}$ from $G(k_2 + a_2, b_2 + \bar{\delta}_\lambda(\boldsymbol{w}))$;
4. Given $\lambda^{(J-1)}$, generate λ from $\pi(\lambda|\alpha_1^{(J)}, \alpha_2^{(J)}, \boldsymbol{w})$ described in (14) with the $N(\lambda^{(J-1)}, S_\lambda^2)$ proposal distribution, where S_λ^2 is the variance of λ which can be chosen to be the inverse of Fisher information. The values of λ can be updated as follows:
   a. Generate ξ_J from $N(\lambda^{(J-1)}, S_\lambda^2)$ and u from $U(0, 1)$
   b. If $u < \min(1, \kappa)$ then let $\lambda^{(J)} = \xi_J$, else go to (a), where $\kappa = \frac{\pi(\xi_J|\boldsymbol{w})}{\pi(\lambda^{(J-1)}|\boldsymbol{w})}$.
5. Set $J = J + 1$.
6. Repeat steps 3-5, M times.

4. Bayesian Prediction Methods

In this section, we address the problem of predicting the censored units in the JPC from the GR distribution. Precisely, at stage j, we are interested in the posterior density of the s-th order statistic from a sample of size S_j removed items and the t-th order statistic from a sample of size T_j removed items, where $R_j = S_j + T_j$. In other words, we wish to predict $U_{s:S_j}(s = 1, 2, \ldots, S_j)$ and $V_{t:T_j}(t = 1, 2, \ldots, T_j)$, based on observing the joint progressively type-II right censored sample, $\boldsymbol{W} = (w_1, w_2, ..., w_k)$. The posterior predictive density of $Y_1 = U_{s:S_j}$ and $Y_2 = V_{t:T_j}$ given the observed censored data is

$$p(y_1, y_2|\text{data}) = \int_0^\infty \int_0^\infty \int_0^\infty f_{Y_1|\text{data}}(y_1|\alpha_1, \lambda)\, g_{Y_2|\text{data}}(y_2|\alpha_2, \lambda) \times \pi(\alpha_1, \alpha_2, \lambda|\text{data})\, d\alpha_1\, d\alpha_2\, d\lambda, \quad y_1, y_2 > w_j. \tag{18}$$

Here $f_{Y_1|\text{data}}(y_1|\alpha_1, \lambda)$ and $g_{Y_2|\text{data}}(y_2|\alpha_2, \lambda)$ are the conditional densities of Y_1 and Y_2 given $\boldsymbol{W} = \boldsymbol{w}$, respectively. By Markovian property of progressively type-II right censored order statistics, this conditional density is the conditional distribution of Y_1 and Y_2 given $W_j = w_j$. That is, the pdf of the s-th and t-th order statistics out of S_j and T_j with $R_j = S_j + T_j$, from F and G right truncated at w_j, respectively (see, for example, Arnold et al. (1998)). Precisely, it can be rewritten as follows:

$$h_{Y_1,Y_2}(y_1, y_2) \propto \sum_{i=0}^{s-1}\sum_{\nu=0}^{t-1}\sum_{k=0}^{S_j-s}\sum_{l=0}^{T_j-t}(-1)^{i+k+l+\nu}\binom{s-1}{i}\binom{t-1}{\nu}\binom{S_j-s}{k}\binom{T_j-t}{l}\lambda^4\alpha_1 \alpha_2 y_1 y_2\, e^{-\lambda^2(y_1^2+y_2^2)} e^{-\alpha_1[(s-i+k)D_\lambda(y_1)+iD_\lambda(w_j)]} e^{-\alpha_2[(t-\nu+l)D_\lambda(y_2)+\nu D_\lambda(w_j)]} e^{D_\lambda(y_1)+D_\lambda(y_2)+S_jQ_{\alpha_1,\lambda}(w_j)+T_jQ_{\alpha_2,\lambda}(w_j)}, \quad y_1, y_2 > w_j. \tag{19}$$

From (13), (18) and (19), we obtain the posterior predictive density of $(\boldsymbol{U}, \boldsymbol{V})$ given the observed JPC data. It is evident that the Bayes predictive estimates $E(Y_1|\boldsymbol{W} = \boldsymbol{w})$ or $E(Y_2|\boldsymbol{W} = \boldsymbol{w})$ cannot be computed directly from (18). Therefore, we propose the Monte Carlo (MC) simulation procedure to generate samples from the predictive distributions. Under the SEL function, the Bayes predictors (BPs) of Y_1 and Y_2 can be obtained as

$$\hat{Y}_1 = E_{posterior}(Y_1|\boldsymbol{w}) = \int_{w_j}^\infty y_1\, p(y_1|\boldsymbol{w})dy_1, \quad \hat{Y}_2 = E_{posterior}(Y_2|\boldsymbol{w}) = \int_{w_j}^\infty y_2\, p(y_2|\boldsymbol{w})dy_2.$$

Based on MC samples $(\alpha_1, \alpha_2, \lambda),\ \ j = 1, 2, ..., M$, the simulation consistent estimator of $p(y|\boldsymbol{w})$ can be obtained as

$$\hat{p}(y_1|\boldsymbol{w}) = \frac{1}{M}\sum_{j=1}^{M} f(y_1|w_j;\alpha_{1j},\lambda_j),\ \hat{p}(y_2|\boldsymbol{w}) = \frac{1}{M}\sum_{j=1}^{M} g(y_2|w_j;\alpha_{2j},\lambda_j). \tag{20}$$

Hence the sample-based predictors of Y_1 and Y_2 can be simplified in the following forms:

$$\hat{Y}_1 = \frac{1}{M}\sum_{j=1}^{M}\frac{1}{\lambda_j}\int_0^{\infty}\left\{-log\left[1-\left(u+(1-u)(1-e^{-(\lambda_j w_j)^2})^{\alpha_{1j}}\right)^{1/\alpha_{1j}}\right]\right\}\phi(u)du,$$

and

$$\hat{Y}_2 = \frac{1}{M}\sum_{j=1}^{M}\frac{1}{\lambda_j}\int_0^{\infty}\left\{-log\left[1-\left(u+(1-u)(1-e^{-(\lambda_j w_j)^2})^{\alpha_{2j}}\right)^{1/\alpha_{2j}}\right]\right\}\phi(u)du,$$

where $\phi(.)$ is the PDF of gamma distribution with parameters s and $S_j - s$. Further, the approximate estimator in (20) can be also used to obtain a two-sided prediction intervals (PIs) for $Y_1 = U_{s:S_j}$ $(s = 1, 2, \ldots, S_j)$ and $Y_2 = V_{t:T_j}, (t = 1, 2, \ldots, T_j)$. The $(1-\gamma)100\%$ PIs of Y_1 and Y_2 are (L_1, U_1) and (L_2, U_2), respectively, where L_1, U_1, L_2, and U_2 are computed numerically from the following equations:

$$P(Y_1 > L_1|\boldsymbol{w}) = \int_{L_1}^{\infty}\hat{p}(y_1|\boldsymbol{w})dy_1 = 1-\frac{\gamma}{2},\ \text{and}\ P(Y_1 > U_1|\boldsymbol{w}) = \int_{U_1}^{\infty}\hat{p}(y_1|\boldsymbol{w})dy_1 = \frac{\gamma}{2}, \tag{21}$$

and

$$P(Y_1 > L_2|\boldsymbol{w}) = \int_{L_2}^{\infty}\hat{p}(y_2|\boldsymbol{w})dy_2 = 1-\frac{\gamma}{2},\ \text{and}\ P(Y_1 > U_2|\boldsymbol{w}) = \int_{U_2}^{\infty}\hat{p}(y_2|\boldsymbol{w})dy_2 = \frac{\gamma}{2}. \tag{22}$$

In most practical situations, it is required to predict the removed (missing) observations jointly or any function of these removed observations. For example, let us consider that we wish to predict the average of the removed units at the stage j. For this, our aim is to predict the removed units $(\boldsymbol{U}, \boldsymbol{V})$, jointly. In this case, we need to find the full Bayesian model of $(\boldsymbol{U}, \boldsymbol{V})$ given $\boldsymbol{W} = \boldsymbol{w}$ that allows us to be able to implement the Gibbs sampling algorithm. By the likelihood of complete data and the joint prior of α_1, α_2 and λ, the full Bayesian model can be described as

$$\pi(\alpha_1, \alpha_2, \lambda, \boldsymbol{u}, \boldsymbol{v}|\boldsymbol{w}) \propto \phi_1(\alpha_1|\lambda, \boldsymbol{w})\ \phi_2(\alpha_2|\lambda, \boldsymbol{w})\ \phi_3(\lambda|\boldsymbol{w}), \tag{23}$$

where $\phi_1(\alpha_1|\lambda, \boldsymbol{w})$ and $\phi_2(\alpha_2|\lambda, \boldsymbol{w})$ are the PDF of gamma distributions $G(m + a_1, b_1 + \delta_\lambda(\boldsymbol{w}, \boldsymbol{u}))$, $G(n + a_2, b_2 + \bar{\delta}_\lambda(\boldsymbol{w}, \boldsymbol{v}))$ and $\phi_3(\lambda|\boldsymbol{w})$ is given by

$$\phi_3(\lambda|\boldsymbol{w}) = \lambda^{2(m+n)+a_3-1}\ e^{-[b_3+\eta(\boldsymbol{w},\boldsymbol{u},\boldsymbol{v})]\lambda^2}\ e^{\delta_\lambda(\boldsymbol{w},\boldsymbol{u})+\bar{\delta}_\lambda(\boldsymbol{w},\boldsymbol{v})}, \tag{24}$$

where

$$\eta(\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v}) = \sum_{i=1}^{k} w_i^2 + \sum_{i=1}^{k}\sum_{j=1}^{S_i} u_{ij}^2 + \sum_{i=1}^{k}\sum_{j=1}^{T_i} v_{ij}^2,$$

and

$$\delta_\lambda(\boldsymbol{w}, \boldsymbol{u}) = \delta_\lambda(\boldsymbol{w}) + \sum_{i=1}^{k}\sum_{j=1}^{S_i} D_\lambda(u_{ij}),\ \bar{\delta}_\lambda(\boldsymbol{w}, \boldsymbol{v}) = \bar{\delta}_\lambda(\boldsymbol{w}) + \sum_{i=1}^{k}\sum_{j=1}^{T_i} D_\lambda(v_{ij}).$$

It is obvious that the full conditional distribution of α_1 given $\lambda, \boldsymbol{w}$ and the full conditional distribution of α_2 given $\lambda, \boldsymbol{w}$ are $G(m + a_1, b_1 + \delta_\lambda(\boldsymbol{w}, \boldsymbol{u}))$, $G(n + a_2, b_2 + \bar{\delta}_\lambda(\boldsymbol{w}, \boldsymbol{v}))$, but the the form of full conditional distribution of λ given $\boldsymbol{w}$ in (24) is not well-known distribution. By setting $\boldsymbol{U}_{j(s)} = (U_{1:S_j},\ldots,U_{s-1:S_j},U_{s+1:S_j},U_{S_j:S_j})$, $\boldsymbol{V}_{j(t)} = (V_{1:T_j},\ldots,V_{t-1:T_j},V_{t+1:T_j},V_{T_j:T_j})$ and using (24), we immediately obtain the conditional probability density function of Y_1 and Y_2 as

$$\pi(y_1,y_2|\boldsymbol{w},\boldsymbol{u}_{j(s)},\boldsymbol{v}_{j(t)},\boldsymbol{\theta}) = \begin{cases} \frac{f(y_1)g(y_2)\ I_{\{u_{s-1:S_j}<y_1<u_{s+1:S_j},v_{t-1:T_j}<y_2<v_{t+1:T_j}\}}}{\left[F(u_{s+1:S_j})-F(u_{s-1:S_j})\right]\left[G(v_{t+1:T_j})-G(v_{t-1:T_j})\right]}, & s\neq S_j, t\neq T_j,\\ \frac{f(y_1)g(y_2)\ I_{\{y_1>u_{s-1:S_j},y_2>v_{t-1:T_j}\}}}{\left[1-F(u_{s-1:S_j})\right]\left[1-G(v_{t-1:T_j})\right]}, & s=S_j, t=T_j, \end{cases}$$

where $f(.),F(.)$ and $g(.),G(.)$ were defined as the PDF and CDF for X-data and Y-data, respectively. Hence the values of $\boldsymbol{u}$ and $\boldsymbol{v}$ can be updated based on the transformations:

$$\begin{aligned} Y_1 &= \frac{1}{\lambda}\left\{-\ln\left[1-\left((1-U)(1-e^{-(\lambda u_{s-1:S_j})^2})^{\alpha_1}+U(1-e^{-(\lambda u_{s+1:S_j})^2})^{\alpha_1}\right)^{1/\alpha_1}\right]\right\}^{1/2},\ s\neq S_j,\\ Y_1 &= \frac{1}{\lambda}\left\{-\ln\left[1-\left((1-U)(1-e^{-(\lambda u_{s-1:S_j})^2})^{\alpha_1}+U\right)^{1/\alpha_1}\right]\right\}^{1/2},\ s=S_j, \end{aligned} \tag{25}$$

and

$$\begin{aligned} Y_2 &= \frac{1}{\lambda}\left\{-\ln\left[1-\left((1-U)(1-e^{-(\lambda v_{t-1:T_j})^2})^{\alpha_2}+U(1-e^{-(\lambda v_{t+1:T_j})^2})^{\alpha_2}\right)^{1/\alpha_2}\right]\right\}^{1/2},\ t\neq T_j,\\ Y_2 &= \frac{1}{\lambda}\left\{-\ln\left[1-\left((1-U)(1-e^{-(\lambda v_{t-1:T_j})^2})^{\alpha_2}+U\right)^{1/\alpha_2}\right]\right\}^{1/2},\ t=T_j, \end{aligned} \tag{26}$$

where U is standard uniform $U(0,1)$ variate. The Gibbs sampler is used to estimate the posterior distribution by sampling for α_1,α_2 and (Y_1,Y_2) from the full conditional distributions. For λ, it can be updated via M-H algorithm using normal distribution as a proposal distribution. Algorithm similar to the M-H algorithm for estimation problem can be implemented for the prediction problem using the full conditional distributions provided in (23) and (24). The steps can be described as follows:

M-H algorithm for prediction problem:

1. Start with initial values $(\alpha_1^{(0)},\alpha_2^{(0)},\lambda^{(0)},\boldsymbol{u}^{(0)},\boldsymbol{v}^{(0)})$;
2. Set $J = 1$;
3. Generate $\alpha_1^{(J)}$ from $G(m+a_1,b_1+\delta_\lambda(\boldsymbol{w},\boldsymbol{u})$, and Generate $\alpha_2^{(J)}$ from $G(n+a_2,b_2+\bar{\delta}_\lambda(\boldsymbol{w},\boldsymbol{v})$;
4. Given $\lambda^{(J-1)}$, generate λ from $\phi_3(\lambda|\alpha_1^{(J)},\alpha_2^{(J)},\boldsymbol{w})$ described in (24) with the $N(\lambda^{(J-1)},S_\lambda^2)$ proposal distribution, where S_λ^2 is the variance of λ which can be chosen to be the inverse of Fisher information. The values of λ can be updated as follows:
   a. Generate ξ_J from $\Delta(.|\lambda^{(J-1)},S_\lambda^2)\equiv N(\lambda^{(J-1)},S_\lambda^2)$ and u from $U(0,1)$
   b. If $u<\min(1,\kappa^*)$ then let $\lambda^{(J)}=\xi_J$, else go to (a), where
   $$\kappa^* = \frac{\phi_3(\Delta_J|\boldsymbol{w})}{\phi_3(\xi^{(J-1)}|\boldsymbol{w})}\frac{\Delta(\lambda^{(J-1)}|\xi_J,S_\lambda^2)}{\Delta(\xi_J|\lambda^{(J-1)},S_\lambda^2)}.$$
5. Set $J=J+1$.
6. Repeat steps 3-5, M times.

5. Simulation Results and Data Analysis

Here in this section, we perform a comprehensive simulation study to assess the performances of the sample-based estimates and predictors developed in the previous sections and discuss the analysis of the JPC data extracted from a practical data with GR fitting distribution. All computations are performed using R software.

5.1. Data analysis

In this section we present the analysis of real data sets to illustrate the performance of the obtained methods. The data represent the breaking strength (in MPa) of jute fibre at 5 mm and 15 mm gauge lengths. The data were originally reported by Xia et al. (2009). The sets of data are:

Data set 1 (5 mm gauge length):

```
129.08 167.87 168.2  178.25 185.42 187.68 218.86 226.53 254.29 260.97
268.20 270.79 304.84 306.99 315.33 360.8  367.70 370.02 441.87 495.51
496.28 516.48 537.45 546.11 554.61 566.31 583.97 618.57 809.23 823.03
```

Data set 2 (15 mm gauge length):

```
42.66  70.09  72.24  76.38  80.4   106.73 127.81 135.09 156.67 168.37
193.42 200.76 202.75 225.65 355.56 339.22 457.71 468.47 489.66 497.94
550.42 562.39 569.07 574.86 594.4  640.48 678.06 716.3  748.75 813.87
```

Table 1: MLEs, K-S and CvM goodness-of-fit tests.

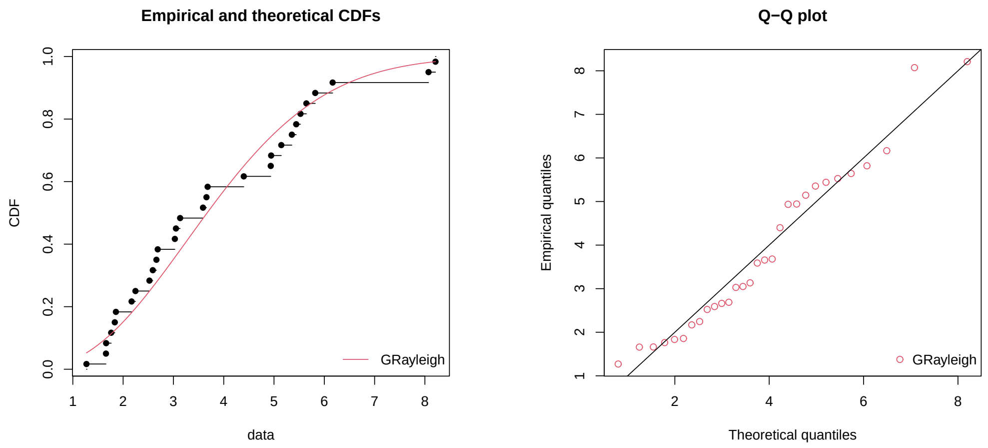
Data Set	Scale Parameter	Shape Parameter	K-S(p-value)	CvM(p-value)
1	0.2252157	1.290241	0.13333(0.9578)	0.14865(0.9238)
2	0.2252157	0.6278746	0.14825(0.9435)	0.14241(0.9360)

Table 1 presents the MLEs of the unknown parameters, the goodness-of-fit tests based on Kolmogrov-Smirnov (K-S) and Cramer-von Mises (CvM) statistics. It is easily seen that the GR model fits both data sets very well. This conclusion is also supported by diagnostic plots of the empirical and fitted distribution functions in Figures 1 and 2. In addition, it is of interest to study the null hypothesis $H_0 : \lambda_1 = \lambda_2 = \lambda$ (i.e., the scale parameters are equal) versus the alternative hypothesis $H_1 : \lambda_1 \neq \lambda_2$ using the likelihood ratio test. For the given data, the test statistic is computed as $\Delta = L1/L2 = 0.4123$, $-2\log(\Delta) = 1.7722$ and the p-value of the test is 0.5209. Hence, the assumption of equality of the scale parameters cannot be rejected.

For explanation purposes, we suggest the following joint progressive type-II censored sample with $m = n = 30$, $k = 15$, $R_i = 5$, $i = 1, \ldots, 6$, $R_i = 0$, $i = 7, \ldots, 10$ and $R_i = 3$, $i = 11, \ldots, 15$. The data set is as follows:

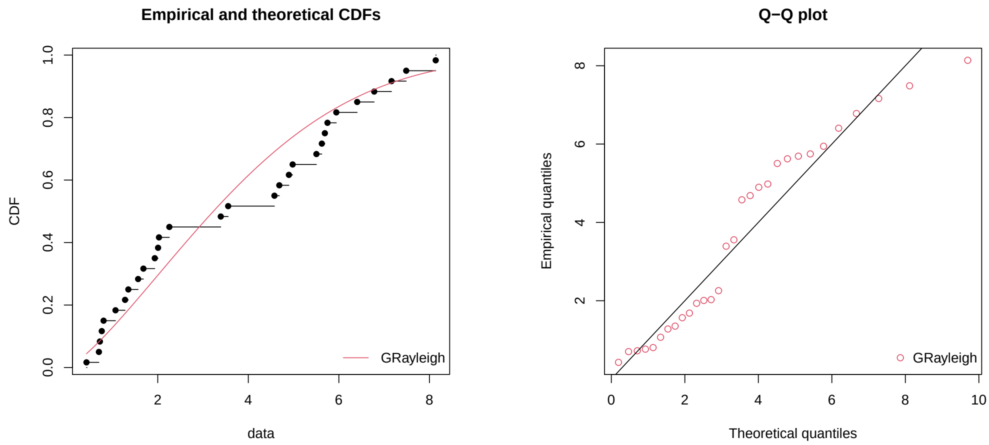
```
(42.66,0,2), (70.09,0,4), (72.24,0,1), (76.38,0,3), (80.4,0,2),
(106.73,0,4), (127.81,0,0), (129.08,1,0), (135.09,0,0), (156.67,0,0),
(167.87,1,3), (168.2,1,0), (168.37,0,2), (178.25,1,2), (185.42,1,2).
```

Based on the above observed joint progressive type II censored data, we obtain the MLEs and BEs of α_1, α_2 and λ. We have generated 5000 observations to compute the BEs of α_1, α_2 and λ based on the importance sampler. For computing the BEs and C-S CRIs, we assume that the priors of α_1, α_2 and λ are improper, i.e. $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$, since we do not have any prior information. The M-H algorithm is also used to compute the BEs of λ. These BEs are obtained after discarding the initial 500 burn-in samples. The histogram displayed in Figure 3 shows that the Gaussian distribution is an appropriate proposal distribution for the full conditional distributions of λ. Here, the best fitted model for the full conditional distribution can be concluded by managing the choice of the parameters for the proposal distribution. Therefore, to generate numbers from the target probability distribution, we use the M-H algorithm with Gaussian proposal distribution. We assumed the initial value of λ to be its MLE, $\widehat{\lambda}_M$ which is computed using EM-algorithm while the variance



(a) **(b)**

Figure 1 (a): Empirical and fitted distribution functions Plot for data set 1, (b): Q-Q Plot for data set 1



(a) **(b)**

Figure 2 (a): Empirical and fitted distribution functions Plot for data set 2, (b): Q-Q Plot for data set 2

of λ to be the reciprocal of Fisher information which is $S^2_\lambda = 0.00027$. Here, we generated $50,000$ random variates with $S^2_\lambda = 0.00027$ and we checked the acceptance rate for this choice of variance to be 70.77% which is quite satisfactory. We discarded the initial 5000 burn-in samples and computed the BEs based on the remaining observations.

Graphical diagnostics tools involving trace and ACF plots are used to check the convergence of M-H algorithm. Figure 3 shows the trace and ACF plots for λ. From the trace plot, we can easily observe a random scatter about some mean value represented by a solid line with a fine mixing of the chains for the simulated values of λ. The ACF plot shows that chains have very low autocorrelations. As a result, these plots indicate the rapid convergence of the M-H algorithm based on the proposed Gaussian distribution.

The results for MLEs and BEs using importance and M-H samplers along with the 95% Boot-t CI, Asymptotic CI and C-S CRIs for α_1, α_2 and λ are presented in Table 2.

Let us now consider the prediction of some censored values, based on the above 15 observed values. The point predicted values and PIs for censored values, are computed and displayed in Table 3. The PIs are established based on Bayesian methods (importance and M-H algorithms) under SEL function which are discussed in details in Section 4. It is clearly evident that the Bayesian M-H PI is the shortest interval for all cases.

Table 2: Point estimators and 95% CIs for α_1, α_2 and λ.

		SEL				
	MLE	Importance	M-H	Approximate CI	Boot-t CI	C-S CRI
α_1	6.9268	4.7495	4.8628	(5.5266, 7.7802)	(4.1257, 5.8196)	(4.3577, 5.9571)
α_2	1.8312	1.8959	1.8382	(0.7529, 2.9094)	(0.9751, 3.0608)	(0.7869, 2.4675)
λ	0.0075	0.0089	0.00827	(0.0053, 0.0096)	(0.0057, 0.0094)	(0.0059, 0.0092)

Table 3: Point predictors and 95% PIs of $U_{s:S_j}$ and $V_{t:T_j}$ based on Bayesian methods.

	Point predictors		PIs	
	Importance	M-H	Importance	M-H
n ↓				
$U_{1:S_1}$	43.56	44.62	(41.27, 45.48)	(43.17, 47.08)
$U_{2:S_1}$	45.39	46.82	(44.12, 48.77)	(44.84, 49.13)
$V_{1:T_1}$	44.12	45.94	(42.89, 47.22)	(44.21, 48.22)
$V_{3:T_1}$	46.95	47.12	(45.28, 50.09)	(46.41, 50.83)
$U_{1:S_2}$	72.36	73.62	(71.19, 75.53)	(72.03, 76.08)
$U_{4:S_2}$	77.68	79.12	(75.91, 80.62)	(77.82, 82.14)
$V_{1:T_2}$	73.11	73.95	(71.75, 76.14)	(71.97, 76.08)
$U_{1:S_3}$	74.75	75.52	(73.14, 77.62)	(73.88, 78.02)
$V_{1:T_3}$	75.48	76.33	(74.46, 78.91)	(75.26, 79.47)
$V_{3:T_3}$	78.69	79.09	(75.98, 80.67)	(77.69, 82.16)
$U_{1:S_4}$	78.17	78.89	(76.67, 81.23)	(77.13, 81.51)
$U_{3:S_4}$	81.66	82.64	(79.26, 84.04)	(80.37, 84.96)
$V_{1:T_4}$	80.09	81.34	(77.52, 82.13)	(78.62, 83.06)
$U_{1:S_5}$	82.99	84.45	(80.12, 84.94)	(82.45, 86.98)
$V_{1:T_5}$	84.32	85.97	(83.42, 88.36)	(84.28, 89.02)
$V_{3:T_5}$	86.43	87.82	(84.78, 89.86)	(85.63, 90.54)
$U_{1:S_6}$	108.61	109.72	(106.56, 111.61)	(107.72, 112.49)
$U_{3:S_6}$	113.28	115.11	(110.73, 115.00)	(113.47, 118.40)
$V_{1:T_6}$	110.71	111.65	(109.08, 114.25)	(109.67, 114.49)
$U_{1:S_{11}}$	171.62	170.23	(169.64, 175.48)	(168.83, 174.41)
$U_{3:S_{11}}$	177.52	176.82	(175.38, 181.44)	(173.86, 179.65)
$V_{1:T_{12}}$	173.66	172.19	(171.63, 177.72)	(170.76, 176.59)
$V_{3:T_{12}}$	179.42	178.61	(177.29, 183.50)	(175.64, 181.59)
$U_{1:S_{13}}$	175.01	173.19	(173.83, 180.15)	(171.39, 177.42)
$V_{1:T_{13}}$	177.49	177.39	(175.47, 181.95)	(176.12, 182.27)
$U_{1:S_{14}}$	182.39	183.07	(180.67, 187.24)	(181.48, 187.75)
$V_{1:T_{14}}$	183.47	184.62	(180.82, 187.55)	(181.79, 188.17)
$U_{1:S_{15}}$	189.70	190.68	(186.77, 193.56)	(188.52, 195.03)
$V_{1:T_{15}}$	191.58	192.49	(188.90, 195.82)	(189.62, 196.35)

5.2. Simulation results

Now, we compare the performances of the different methods of estimation and prediction based on Monte Carlo simulations. We compare the performances of the MLEs, and BEs in terms of biases

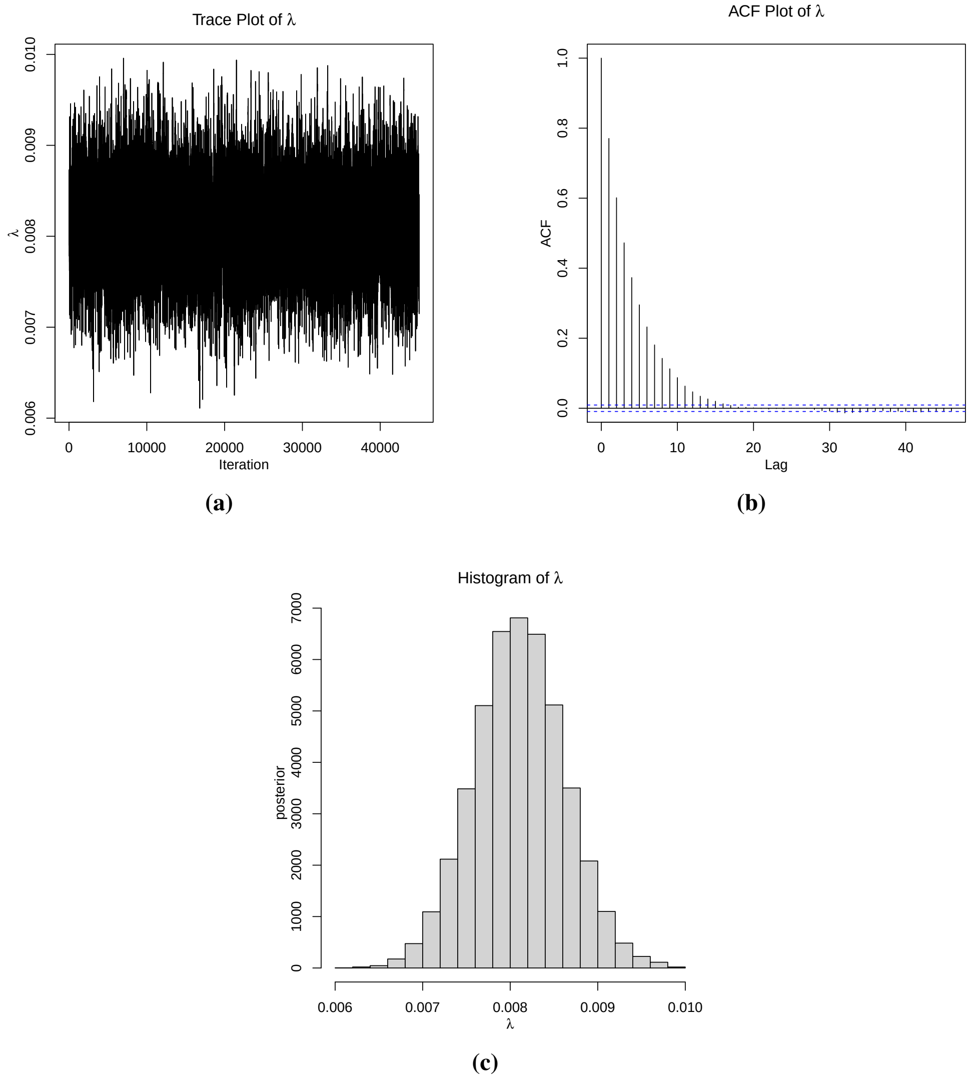


Figure 3 Plots of Metropolis-Hastings Markov chains for λ. (a): Trace Plot of λ, (b): ACF Plot of λ and (c): Histogram of λ

and mean squares errors (MSEs). In this simulation, the values of GR parameters are considered as $\alpha_1 = 2$, $\alpha_2 = 1.5$ and $\lambda = 2$. Here, we consider different effective sample sizes, $k = 20$, 25 and different censoring schemes. For conducting the Bayesian analysis, we assume three different priors. For the first prior (Prior 0), zero values were given to the prior parameters to reflect improper prior information. That is, we assumed that $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$. Then we assume two additional proper priors with same means but different variances to reflect the sensitivity of our inferences to variations in the specification of prior parameters. Prior 1 and Prior 2 correspond to $a_1 = a_2 = a_3 = 1$, $b_1 = b_2 = b_3 = 2$ and $a_1 = a_2 = 2$, $a_3 = 1$, $b_1 = b_2 = b_3 = 4$, respectively. For this reason, Prior 2 is more informative than Prior 1. This helps us to see how much does the informative prior effect contributes to the results obtained based on observed data. We use the following notation for a particular progressive censoring scheme. For example $k = 7$ and $R = (7, 0_{(5)}, 15)$ means $R_1 = 7$, $R_2 = R_3 = R_4 = R_5 = R_6 = 0, R_7 = 15$.

Table 4 presents the average biases and MSEs of the MLEs and BEs under SEL of α_1, α_2 and λ are computed over 5000 replications. The average lengths (ALs) and coverage probabilities (CPs) of 95% CIs for α_1, α_2 and λ based on Bootstrap-t, asymptotic maximum likelihood and Bayesian methods with improper prior (Prior 0) and informative priors (Prior 1 and Prior 2) under SEL function are displayed in Table 5. As seen in Table 4, the BEs perform well for $k = 20$ and $k = 25$ in the sense of bias and MSE. Clearly, the results of BEs are not very sensitive to the assumed values of the prior parameters, particularly the informative prior (Prior 2). From Table 5, it is observed that the C-S CRIs are shorter than the asymptotic and Boot-t CIs under all priors for $k = 20$ and $k = 25$. The performances of C-S CRIs tend to be high under the informative priors when compared to asymptotic and Boot-t CIs. While the Boot-t method performs well when compared to asymptotic method for estimating of all parameters. It can be also observed that all CIs are shorter for $k = 25$ when compared to $k = 20$. The simulated CPs are very close for all these intervals.

For prediction problem, we have randomly generated six different joint progressive censored schemes from GR distribution and for $k = 20$ and $k = 25$, we then compute the average biases, mean square prediction errors (MSPEs) for the predictors and PIs for censored items. Under the non-informative and informative priors, the biases and MSPEs of BPs over 1000 replications are computed and provided in Tables 6 and 7 for various censoring schemes of $k = 20$ and $k = 25$. The ALs and CPs of 95% PIs based on importance, M-H and bootstrap methods are also reported in Table 8. It can be seen from Tables 6 and 7 that the BPs using M-H method perform well when compared to the importance sampler method. As expected, the MSPEs of the censored items tend to be larger when s gets larger and when k becomes smaller. From Table 8, there is a clear evidence that the Bayes PIs based on M-H sampling are the most preferred PIs in terms of AL criterion. It can also be observed that the Bayes PIs under the informative prior (Prior 2) perform better than PIs based on other priors. Moreover, the simulated CPs are remarkably high and tend to be the true prediction coefficient $1 - \gamma = 0.95$.

6. Conclusions

In this study, we have considered the estimation and prediction problems based on joint Type-II progressive censoring scheme when their lifetimes follow generalized Rayleigh distributions with different shape parameters. It is shown that the maximum likelihood estimators of the model parameters can be obtained by adopting the expectation-maximization algorithm. We have also proposed Bayesian procedures to estimating the parameters involved and predicting the life lengths of the removed units in multiple stages of the joint progressively censored sample. The corresponding estimation and prediction intervals involving importance, bootstrap and Metropolis-Hastings methods are used to develop prediction intervals for the future censored units. The performance of all methods presented in this paper are evaluated and compared via Monte Carlo simulations. It is observed that the Bayes estimates and predictors under Metropolis-Hastings method outperform the frequentist methods as well as the Bayes ones under the importance sampling in the sense of bias and mean square error. By considering the average length and coverage probability as optimality criteria, it is also evident that the highest posterior density credible intervals compete the approximate and Boot-t confidence intervals. In the context of prediction, the prediction intervals based on Metropolis-Hastings method are more efficient than the ones based on importance sampling.

Although, we have mainly restricted our attention to the joint Type-II progressive censoring scheme produced from the two populations, but the so developed procedures can be extended to more than two populations as well. More investigation is needed along this line.

Table 4: Biases and MSEs of the MLEs and BEs of α_1, α_2 and λ.

		MLE		Prior 0		Prior 1		Prior 2	
Censoring scheme ↓		Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
$k=20$, $R=(7,0_{(18)},15)$	α_1	0.0877	0.0694	-0.0762	0.0625	0.0781	0.0602	0.0708	0.0556
	α_2	0.1281	0.1316	0.0840	0.0905	0.0804	0.0848	0.0747	0.0759
	λ	0.0745	0.0628	0.0693	0.0581	0.0633	0.0549	0.0591	0.0527
$k=20$, $R=(0_{(9)},7,0_{(9)},15)$	α_1	0.0915	0.0773	-0.0899	0.0710	0.0821	0.0678	0.0799	0.0652
	α_2	-0.1320	0.1586	0.0948	0.1095	0.0857	0.0967	0.0817	0.0959
	λ	0.0870	0.0687	-0.0851	0.0640	0.0830	0.0637	0.0784	0.0617
$k=20$, $R=(0_{(18)},7,15)$	α_1	0.0958	0.0784	-0.0934	0.0714	0.0856	0.0689	0.0839	0.0672
	α_2	0.1542	0.1881	0.1103	0.1183	0.0871	0.0986	0.0847	0.0985
	λ	0.0918	0.0701	0.0903	0.6558	0.0835	0.0647	0.0815	0.0637
$k=25$, $R=(7,0_{(23)},10)$	α_1	0.0608	0.0507	-0.0584	0.0495	0.0561	0.0489	0.0542	0.0467
	α_2	0.0772	0.0761	0.0752	0.0738	0.0744	0.0725	-0.0724	0.0659
	λ	0.0527	0.0476	0.0511	0.0453	0.0503	0.0442	0.0488	0.0430
$k=25$, $R=(0_{(11)},7,0_{(12)},10)$	α_1	0.0622	0.0541	0.0601	0.0511	0.0588	0.0501	0.0566	0.0471
	α_2	0.0799	0.0810	0.0791	0.0742	0.0784	0.0727	-0.0775	0.0687
	λ	0.0549	0.0499	-0.0520	0.0467	0.0515	0.0453	0.0500	0.0467
$k=25$, $R=(0_{(23)},7,10)$	α_1	-0.0628	0.0542	-0.0602	0.0517	0.0594	0.0509	0.0584	0.0488
	α_2	0.0816	0.0832	0.0795	0.0753	0.0786	0.0739	-0.0724	0.0718
	λ	0.0547	0.0497	0.0525	0.0478	-0.0520	0.0492	0.0519	0.0475

Table 5: ALs and CPs of 95% CIs of α_1, α_2 and λ when $m=20$ and $n=22$.

			Boot-t CI	Approximate CI	C-S CRI		
Censoring scheme ↓					Prior 0	Prior 1	Prior 2
$k=20$, $R=(7,0_{(18)},15)$	α_1	AL	1.1514	1.2920	0.9714	0.8527	0.8156
		CP	0.9514	0.9502	0.9541	0.9546	0.9609
	α_2	AL	1.2299	1.3216	1.0148	0.9135	0.8853
		CP	0.9520	0.9511	0.9547	0.9558	0.9621
	λ	AL	1.1201	1.2278	0.9265	0.9107	0.8161
		CP	0.9501	0.9492	0.9530	0.9534	0.9577
$k=20$, $R=(0_{(9)},7,0_{(9)},15)$	α_1	AL	1.2927	1.5298	1.0147	0.9277	0.8826
		CP	0.9524	0.9512	0.9549	0.9552	0.9622
	α_2	AL	1.3647	1.6472	1.0954	0.9725	0.9043
		CP	0.9528	0.9518	0.9556	0.9569	0.9634
	λ	AL	1.1821	1.3080	0.9814	0.9146	0.8275
		CP	0.9519	0.9508	0.9536	0.9554	0.9585
$k=20$, $R=(0_{(18)},7,15)$	α_1	AL	1.4692	1.6842	1.0848	1.0276	0.8916
		CP	0.9528	0.9519	0.9554	0.9569	0.9625
	α_2	AL	1.5859	1.8228	1.1924	1.0752	0.9265
		CP	0.9536	0.9522	0.9562	0.9578	0.9644
	λ	AL	1.2142	1.3696	1.0122	0.9407	0.8521
		CP	0.9525	0.9514	0.9543	0.9568	0.9591
$k=25$, $R=(7,0_{(23)},10)$	α_1	AL	0.9836	1.0243	0.9246	0.8290	0.7837
		CP	0.9481	0.9474	0.9501	0.9512	0.9545
	α_2	AL	1.0098	1.0852	0.9425	0.8224	0.8009
		CP	0.9498	0.9489	0.9518	0.9532	0.9590
	λ	AL	0.9662	1.0109	0.8912	0.7710	0.7592
		CP	0.9472	0.9466	0.9488	0.9506	0.9520
$k=25$, $R=(0_{(11)},7,0_{(12)},10)$	α_1	AL	1.0554	1.1223	0.9452	0.8406	0.8065
		CP	0.9504	0.9496	0.9524	0.9535	0.9587
	α_2	AL	1.1562	1.1851	0.9712	0.8572	0.8232
		CP	0.9510	0.9501	0.9533	0.9542	0.9600
	λ	AL	0.9902	1.1087	0.9002	0.8084	0.7852
		CP	0.9493	0.9488	0.9518	0.9529	0.9571
$k=25$, $R=(0_{(23)},7,10)$	α_1	AL	1.1348	1.2656	0.9601	0.8256	0.8054
		CP	0.9519	0.9504	0.9537	0.9548	0.9594
	α_2	AL	1.2157	1.2842	1.0025	0.8995	0.8620
		CP	0.9519	0.9510	0.9541	0.9559	0.9607
	λ	AL	1.0821	1.2014	0.9084	0.8151	0.8006
		CP	0.9511	0.9501	0.9529	0.9534	0.9575

Table 6: Biases and MSPEs of BPs for censored items for $k = 20$.

Censoring Scheme ↓			Prior 0		Prior 1		Prior 2	
			Importance	M-H	Importance	M-H	Importance	M-H
$R = (7, 0_{(18)}, 15)$	$Y_{1:7}$	Bias	0.3816	-0.3625	0.3688	0.3527	0.3561	0.3441
		MSPE	0.8248	0.7845	0.7485	0.7098	0.5858	0.4929
	$Y_{3:7}$	Bias	0.4264	-0.4025	0.3928	0.3758	0.3784	0.3668
		MSPE	0.8559	0.8088	0.7679	0.7260	0.6009	0.5359
	$Y_{5:7}$	Bias	0.4760	0.4421	-0.4265	0.4052	0.3928	0.3735
		MSPE	0.8769	0.8368	0.8067	0.7565	0.7015	0.6086
	$Y_{1:15}$	Bias	-0.3554	0.3351	0.3364	0.3252	0.3169	-0.3055
		MSPE	0.8038	0.7652	0.7225	0.6811	0.5690	0.4661
	$Y_{5:15}$	Bias	0.3787	0.3491	0.3528	0.3350	0.3275	0.3117
		MSPE	0.8358	0.8015	0.7728	0.7301	0.6008	0.5182
	$Y_{10:15}$	Bias	0.4082	0.3942	0.3852	-0.3466	0.3426	0.3114
		MSPE	0.8729	0.8350	0.8010	0.7685	0.6328	0.5992
$R = (0_{(9)}, 7, 0_{(9)}, 15)$	$Y_{1:7}$	Bias	0.4006	0.3853	0.3871	0.3732	-0.3766	0.3661
		MSPE	0.8595	0.8187	0.7714	0.7328	0.6285	0.5220
	$Y_{3:7}$	Bias	0.4626	0.4243	-0.4119	0.4056	0.3974	0.3822
		MSPE	0.8845	0.8580	0.8074	0.7565	0.6449	0.5659
	$Y_{5:7}$	Bias	0.5092	0.4741	0.4836	-0.4654	0.4392	0.4135
		MSPE	0.9023	0.8645	0.8391	0.7955	0.7209	0.6460
	$Y_{1:15}$	Bias	0.3925	-0.3792	0.3655	0.3518	0.3498	-0.3213
		MSPE	0.8504	0.8250	0.7795	0.7428	0.6288	0.5463
	$Y_{5:15}$	Bias	0.4529	0.4182	0.4320	0.4016	0.4027	0.3682
		MSPE	0.8716	0.8655	0.8202	0.7667	0.6754	0.6018
	$Y_{10:15}$	Bias	0.4757	0.4448	0.4401	0.4212	0.4273	0.3989
		MSPE	0.9068	0.8611	0.8400	0.7912	0.7492	0.6754
$R = (0_{(18)}, 7, 15)$	$Y_{1:7}$	Bias	0.4692	0.4435	0.4528	0.4411	0.4284	0.4266
		MSPE	0.8911	0.8521	0.8173	0.7720	0.6632	0.5775
	$Y_{3:7}$	Bias	-0.5260	0.4874	0.4752	0.4505	0.4495	0.4468
		MSPE	0.9147	0.8790	0.8487	0.7948	0.6820	0.6065
	$Y_{5:7}$	Bias	-0.5557	0.5274	0.5375	0.5181	0.4919	0.4612
		MSPE	0.9345	0.8905	0.8714	0.8236	0.7562	0.6865
	$Y_{1:15}$	Bias	0.4501	0.4379	-0.4365	0.4181	0.4085	-0.3991
		MSPE	0.8754	0.8401	0.7825	0.7611	0.6488	0.5602
	$Y_{5:15}$	Bias	0.4752	0.4528	0.4563	0.4371	0.4216	0.4066
		MSPE	0.8954	0.8670	0.8296	0.7788	0.7059	0.6374
	$Y_{10:15}$	Bias	0.4929	0.4735	0.4751	0.4509	0.4490	0.4227
		MSPE	0.9111	0.8789	0.8574	0.8048	0.7839	0.7177

Table 7: Biases and MSPEs of BPs for censored items for $k = 25$.

Censoring Scheme ↓			Prior 0		Prior 1		Prior 2	
			Importance	M-H	Importance	M-H	Importance	M-H
$R = (7, 0_{(23)}, 10)$	$Y_{1:7}$	Bias	-0.3405	0.3292	0.3184	0.2932	-0.2945	0.2738
		MSPE	0.6721	0.6618	0.6228	0.6024	0.4542	0.4311
	$Y_{3:7}$	Bias	0.3660	-0.3509	-0.3429	0.3291	0.3240	0.3117
		MSPE	0.6950	0.6740	0.6531	0.6168	0.4734	0.4552
	$Y_{5:7}$	Bias	0.3867	-0.3731	0.3619	0.3547	0.3385	-0.3282
		MSPE	0.7061	0.6866	0.6825	0.6585	0.4932	0.4628
	$Y_{1:10}$	Bias	0.2879	0.2810	-0.2851	0.2762	0.2623	-0.2580
		MSPE	0.6502	0.6471	0.6020	0.5840	0.4413	0.4281
	$Y_{4:10}$	Bias	0.3097	0.2962	0.2922	0.2834	0.2758	0.2664
		MSPE	0.6661	0.6580	0.6376	0.6016	0.4659	0.4429
	$Y_{8:10}$	Bias	0.3124	0.3069	-0.3021	0.2951	0.2942	-0.2844
		MSPE	0.6852	0.6732	0.6654	0.6365	0.4827	0.4527
$R = (0_{(11)}, 7, 0_{(12)}, 10)$	$Y_{1:7}$	Bias	0.3457	0.3331	-0.3246	0.3102	-0.3047	0.2931
		MSPE	0.7065	0.6918	0.6379	0.6154	0.4722	0.4523
	$Y_{3:7}$	Bias	0.3714	0.3573	0.3510	-0.3384	0.3379	0.3228
		MSPE	0.7334	0.7227	0.6660	0.6375	0.4965	0.4735
	$Y_{5:7}$	Bias	-0.4011	0.3885	-0.3861	0.3621	0.3525	0.3408
		MSPE	0.7521	0.7405	0.6925	0.6731	0.5118	0.4730
	$Y_{1:10}$	Bias	0.3186	0.3045	-0.2916	0.2821	0.2744	-0.2676
		MSPE	0.6812	0.6750	0.6187	0.5922	0.4588	0.4356
	$Y_{4:10}$	Bias	0.3296	0.3143	-0.3092	0.2911	0.2864	0.2778
		MSPE	0.7146	0.6946	0.6563	0.6168	0.4852	0.4538
	$Y_{8:10}$	Bias	0.3351	-0.3246	0.3169	0.3013	0.3005	0.2952
		MSPE	0.7296	0.7061	0.6782	0.6551	0.4961	0.4665
$R = (0_{(23)}, 7, 10)$	$Y_{1:7}$	Bias	0.3661	0.3552	-0.3456	0.3271	0.3274	0.3182
		MSPE	0.7249	0.7007	0.6582	0.62448	0.4935	0.4691
	$Y_{3:7}$	Bias	-0.3852	0.3750	0.3698	0.3562	-0.3515	0.3374
		MSPE	0.7589	0.7307	0.6760	0.6411	0.5100	0.4805
	$Y_{5:7}$	Bias	-0.4259	0.4131	0.4054	0.3868	-0.3722	0.3681
		MSPE	0.7652	0.7188	0.7089	0.6812	0.5261	0.4962
	$Y_{1:10}$	Bias	0.3238	0.3111	-0.3034	0.2950	0.2951	-0.2846
		MSPE	0.7083	0.6895	0.6301	0.6035	0.4713	0.4463
	$Y_{4:10}$	Bias	0.3346	0.3214	0.3219	-0.3168	0.3126	0.2996
		MSPE	0.7409	0.7008	0.6642	0.6244	0.4932	0.4682
	$Y_{8:10}$	Bias	0.3502	0.3375	0.3354	0.3215	0.3188	0.3024
		MSPE	0.7562	0.7135	0.6924	0.6648	0.5172	0.4765

Table 8: ALs and CPs for PIs of censored items based on importance, M-H and Bootstrap samplers.

Censoring Scheme ↓							
$k = 20$		$Y_{1:7}$	$Y_{3:7}$	$Y_{5:7}$	$Y_{1:15}$	$Y_{5:15}$	$Y_{10:15}$
$R = (7, 0_{(18)}, 15)$	Bootstrap	1.8332(0.9512)	2.1486(0.9487)	2.3712(0.9464)	1.7064(0.9488)	2.0992(0.9461)	2.2551(0.9434)
	Importance	1.1836(0.9585)	1.3478(0.9554)	1.5775(0.9549)	1.1098(0.9574)	1.2254(0.9551)	1.4539(0.9546)
	M-H	0.9878(0.9612)	1.2138(0.9589)	1.4261(0.9561)	0.9581(0.9590)	1.1955(0.9564)	1.2956(0.9552)
$k = 20$		$Y_{1:7}$	$Y_{3:7}$	$Y_{5:7}$	$Y_{1:15}$	$Y_{5:15}$	$Y_{10:15}$
$R = (0_{(9)}7, 0_{(9)}, 15)$	Bootstrap	1.9034(0.9541)	2.2158(0.9523)	2.4935(0.9502)	1.8264(0.9523)	2.1972(0.9506)	2.3480(0.9484)
	Importance	1.2355(0.9548)	1.4862(0.9534)	1.6257(0.9520)	1.1980(0.9531)	1.3491(0.9517)	1.5952(0.9492)
	M-H	1.1685(0.9552)	1.3308(0.9540)	1.5534(0.9528)	1.0546(0.9539)	1.2654(0.9523)	1.3711(0.9499)
$k = 20$		$Y_{1:7}$	$Y_{3:7}$	$Y_{5:7}$	$Y_{1:15}$	$Y_{5:15}$	$Y_{10:15}$
$R = (0_{(18)}, 7, 15)$	Bootstrap	1.9802(0.9571)	2.3508(0.9550)	2.5128(0.9539)	1.9132(0.9563)	2.2836(0.9547)	2.4283(0.9539)
	Importance	1.2995(0.9579)	1.6105(0.9554)	1.7568(0.9549)	1.2287(0.9574)	1.4925(0.9551)	1.6862(0.9546)
	M-H	1.2044(0.9585)	1.4020(0.9573)	1.6218(0.9562)	1.1654(0.9579)	1.3400(0.9569)	1.4848(0.9555)
$k = 25$		$Y_{1:7}$	$Y_{3:7}$	$Y_{5:7}$	$Y_{1:10}$	$Y_{4:10}$	$Y_{8:10}$
$R = (7, 0_{(23)}, 10)$	Bootstrap	1.7352(0.9492)	2.0147(0.9473)	2.2252(0.9451)	1.6245(0.9465)	1.9512(0.9450)	2.1124(0.9412)
	Importance	1.0765(0.9548)	1.2675(0.9519)	1.4428(0.9504)	1.0098(0.9533)	1.1589(0.9511)	1.3357(0.9491)
	M-H	0.8809(0.9563)	1.1270(0.9538)	1.3151(0.9519)	0.8143(0.9552)	1.0569(0.9528)	1.1592(0.9503)
$k = 25$		$Y_{1:7}$	$Y_{3:7}$	$Y_{5:7}$	$Y_{1:10}$	$Y_{4:10}$	$Y_{8:10}$
$R = (0_{(11)}7, 0_{(12)}, 10)$	Bootstrap	1.7937(0.9498)	2.1908(0.9482)	2.3655(0.9474)	1.7039(0.9484)	2.0765(0.9462)	2.2826(0.9445)
	Importance	1.1866(0.9518)	1.3024(0.9503)	1.5139(0.9489)	1.0808(0.9498)	1.2190(0.9474)	1.4233(0.9458)
	M-H	0.9216(0.9526)	1.2002(0.9513)	1.4120(0.9504)	0.9692(0.9510)	1.1237(0.9487)	1.2829(0.9473)
$k = 25$		$Y_{1:7}$	$Y_{3:7}$	$Y_{5:7}$	$Y_{1:10}$	$Y_{4:10}$	$Y_{8:10}$
$R = (0_{(23)}, 7, 10)$	Bootstrap	1.8652(0.9512)	2.2654(0.9496)	2.4429(0.9485)	1.7492(0.9496)	2.1836(0.9478)	2.3360(0.9461)
	Importance	1.2252(0.9526)	1.4351(0.9512)	1.5623(0.9496)	1.1738(0.9516)	1.3285(0.9482)	1.4705(0.9477)
	M-H	1.0505(0.9532)	1.2824(0.9524)	1.4825(0.9520)	1.1127(0.9533)	1.2374(0.9493)	1.3621(0.9488)

Conflict of interest statement:

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgements

The authors are also grateful to the Editor and referees for many constructive suggestions which have helped to improve the paper.

References

Ahmed EA. Bayesian estimation based on progressive Type-II censoring from two-parameter bathtub-shaped lifetime model: an Markov chain Monte Carlo approach. J Appl Stat. 2014; 41: 752-768.

Ahmad KE, Fakhry ME, Jaheen ZF. Empirical Bayes estimation of $P(X < Y)$ and characterization of Burr type X model. J Stat Plan Inference. 1997; 64: 297 - 308.

Ahsanullah M. Linear prediction of record values for the two parameter exponential distribution, Ann Inst Stat Math. 1980; 32(3): 363-368.

Arnold BC, Balakrishnan N, Nagaraja HN. Records. New York: John Wiley and Sons; 1998.

Balakrishnan N, Cramer E. The art of progressive censoring: applications to reliability and quality. New York: Springer; 2014.

Balakrishnan N, Su F, Liu KY. Exact likelihood inference for k exponential population under joint progressive Type-II censoring. Commun Stat Simul Comput. 2015; 44(4): 902-923.

Barakat, HM, El-Adll ME, Aly AE. Prediction intervals of future observations for a sample of random size from any continuous distribution. Math Comput Simul. 2014; 97: 1-13.

Barakat HM, Nigm EM, El-Adll ME, Yusuf M. Prediction for future exponential lifetime based on random number of generalized order statistics under a general set-up. Stat Pap. 2018; 59(2): 605-631.

Chen MH, Shao QM. Monte Carlo estimation of Bayesian credible and HPD intervals. J Comput Graph Stat. 1999; 8: 69-92.

Dempster AP, Laird NM, Rubin DD. Maximum likelihood from incomplete data via EM algorithm. J R Stat Soc Series B. 1977; 39: 1-38.

Gilks WR, Richardson S, Spiegelhalter DJ. Markov chain Monte Carlo in practice. London, UK: Chapman & Hall; 1995.

Kaminsky KS, Nelson PI. Prediction intervals. In: Handbook of Statistics. Balakrishnan N, Rao CR (Eds.). Amesterdam, North Holland; 1998: 431-450.

Kotb MS, Raqab MZ. Statistical inference for modified Weibull distribution based on progressively type-II censored data. Math Comput Simul. 2019; 162: 233-248.

Kundu D, Raqab MZ. Generalized Rayleigh distribution; different methods of estimation. Comput Stat Data Anal. 2005; 49: 187 - 200.

Lawless JF. Statistical Models and Methods for Lifetime Data. 2^{nd} Ed. New York: Wiley; 2002.

Lehmann EL, Casella G. Theory of Point Estimation. 2^{nd} Ed. New York: Springer-Verlag; 1998.

Mohie El-din MM, Shafay AR. One- and two-sample Bayesian prediction intervals based on progressively Type-II censored data. Stat Pap. 2013; 54: 287-307.

Mondal S, Kundu D. Point and interval estimation of Weibull parameters based on joint progressively censored data. Sankhya Ser B. 2019; 81: 1-25.

Mondal S, Kundu D. On the joint Type-II progressive censoring scheme. Commun Stat Theo Meth. 2020; 49(4): 958-976.

Mondal S, Bhattacharya R, Pradhan B, Kundu D. Bayesian optimal life testing plan under the balanced two sample type-II progressive censoring scheme. Appl Stoch Models Bus Ind. 2020; 36(1): 628-640.

Mudholkar GS, Srivastava DK, Freimer M. The Exponentiated Weibull family: A re-analysis of the bus motor failure data. Technometrics. 1995; 37: 436-445.

Nagaraja HN. Comparison of estimators and predictors from two-parameter exponential distribution. Sankhy$\bar{a}$ Ser B. 1986; 48: 10-18.

Ng HKT, Chan CS, Balakrishnan N. Estimation of parameters from progressively ordered censored data using EM algorithm. Comput Stat Data Anal. 2002; 39: 371-386.

Parsi S, Bairamov I. Expected values of the number of failures for two populations under joint Type-II progressive censoring. Comput Stat Data Anal. 2009; 53(10): 3560-3570.

Parsi S, Ganjali M, Farsipour NS. Conditional maximum likelihood and interval estimation for two Weibull populations under joint Type-II progressive censoring. Commun Stat Theo Meth. 2011; 40(12): 2117-2135.

Raqab ZM. Optimal prediction-intervals for the exponential distribution based on generalized order statistics. IEEE Trans Reliab. 2001; 50(1): 112-115.

Raqab ZM, Kundu D. Burr type X distribution: revisited. J of Prob and Stat Sci. 2006; 4: 179-193.

Raqab ZM, Madi MT. Bayesian analysis for the exponentiated Rayleigh distribution. Metron. 2009; Vol. LXVII, 3: 269-288.

Raqab ZM, Madi MT. Inference for the generalized Rayleigh distribution based on progressively censored data. J Stat Plan Inference. 2011; 141: 3313-3322.

Sartawi HA, Abu-Salih MS. Bayes prediction bounds for the Burr type X model. Commun Stat Theo Meth. 1991; 20: 2307 - 2330.

Rasouli A, Balakrishnan N. Exact likelihood inference for two-exponential populations under joint progressive Type-II censoring. Commun Stat Theo Meth. 2010; 39(12): 2172-2191.

Valiollahi R, Raqab MZ, Asgharzadeh A, Alqallaf FA. Estimation and prediction for power Lindley distribution under progressively type II right censored samples. Math Comput Simul. 2018; 149: 32-47.

Tierney L. Markov chains for exploring posterior distributions. Ann Stat. 1994; 22: 1701-1728.

Xia ZP, Yu JY, Cheng LD, Liu LF, Wang WM. Study on the breaking strength of jute fibres using modified Weibull distribution. Compos Part A. 2009; 40: 5459.