Sankhyā : The Indian Journal of Statistics
2014, Volume 76-B, Part 1, pp. 120-139

Analysis of Interval-Censored Data with Weibull Lifetime Distribution

Biswabrata Pradhan
Indian Statistical Institute, Kolkata, India
Debasis Kundu
Indian Institute of Technology, Kanpur, India

Abstract

In this work the analysis of interval-censored data, with Weibull distribution as the underlying lifetime distribution has been considered. It is assumed that censoring mechanism is independent and non-informative. As expected, the maximum likelihood estimators cannot be obtained in closed form. In our simulation experiments it is observed that the Newton-Raphson method may not converge many times. An expectation maximization algorithm has been suggested to compute the maximum likelihood estimators, and it converges almost all the times. The Bayes estimates of the unknown parameters under gamma priors are considered. If the shape parameter is known, the Bayes estimate of the scale parameter can be obtained in explicit form. When both the parameters are unknown, the Bayes estimators cannot be obtained in explicit form. Lindley's approximation, importance sampling procedures and Metropolis Hastings algorithm are used to compute the Bayes estimates. Highest posterior density credible intervals of the unknown parameter are obtained using importance sampling technique. Small simulation experiments are conducted to investigate the finite sample performance of the proposed estimators, and the analysis of two data sets; one simulated and one real life, have been provided for illustrative purposes.

AMS (2000) *subject classification.* Primary 62F10, 62F15; Secondary 62N02.
Keywords and phrases. EM algorithm, Gibbs sampling, HPD credible interval, Lindley's approximation, importance sampling.

1 Introduction

Lifetime data analysis is used in various fields for analyzing data involving the duration between two events. It is also known as *event history analysis*, *survival data analysis*, *reliability analysis* or *time to event analysis* etc. In lifetime data analysis often the data are censored. The event time is *right censored*, when follow-up is curtailed without observing the event.

Left censoring arises when the event occurs at some unknown time prior to a known specified time point. The event time is considered to be *interval censored* when an event occurs within some interval of time, but the exact time of the event is unknown, see Kalbfleish and Prentice (2002) and the survey article by Gomez, Calle and Oller (2004) and the references therein. Jammalamadaka and Mangalam (2003) termed it as *middle censoring*, see also Jammalamadaka and Iyer (2004) and Iyer, Jammalamadaka and Kundu (2009) in this respect.

The interval censoring scheme can be described as follows. Suppose n identical items are put on a life test and let $T_1, \ldots, T_n$ be the lifetime of these items. For the i-th item, there is a random censoring interval (L_i, R_i), which follows some unknown bivariate distribution. Here L_i and R_i denote the left and right random end point, respectively, of the censoring interval. The life time of the i-th item, T_i, is observable only if $T_i \notin [L_i, R_i]$, otherwise it is not observable. Define $\delta_i = I(T_i \notin [L_i, R_i])$, then $\delta_i = 1$ implies the observation is not censored. In that case the actual value of T_i is observed. When $\delta_i = 0$, only, the censoring interval $[L_i, R_i]$ is observed. For all the n items, the observe data is of the form $(y_i, \delta_i), i = 1, \ldots, n$, where

$$(y_i, \delta_i) = \begin{cases} (T_i, 1) & \text{if } T_i \notin [L_i, R_i] \\ \\ ([L_i, R_i], 0) & \text{otherwise,} \end{cases} \tag{1.1}$$

see for example Sparling, Younes and Lachin (2006) or Jammalamadaka and Mangalam (2003). Note that the interval censoring is a generalization of the existing left censoring and right censoring. Examples of interval censored data arise in diverse fields, such as biology, demography, economics, engineering, epidemiology, medicine, and in public health. Several examples where interval-censored data can arise have been cited by Gomez et al. (2004), Jammalamadaka and Mangalam (2003) also provided a nice example in the sociological context.

Recently, Iyer et al. (2009) considered the analysis of interval/middle censored data when T_i's are exponentially distributed. Although, exponential distribution has been used quite extensively in analyzing lifetime data, but it can have only constant hazard function, and also the PDF is always a decreasing function. While analyzing one interval censored data (Example 5.2), it is observed that the exponential distribution may not provide a good fit to that data set.

In this work, it is assumed that $T_1, \ldots, T_n$ are independent identically distributed (*i.i.d.*) Weibull random variables with the probability density function

$$f(t;\alpha,\lambda) = \begin{cases} \alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}} & \text{if} \quad t > 0 \\ 0 & \text{if} \quad t \le 0, \end{cases} \tag{1.2}$$

here $\alpha > 0$, $\lambda > 0$ are the shape and scale parameters respectively. From now on, the Weibull distribution with the PDF defined as (1.2) will be denoted as WE(α,λ). Also it is assumed that the random censoring times L_i and R_i are independent of T_i, and it does not have any information regarding the population parameters α and λ. The assumption that the censoring mechanism is non-informative is not very uncommon in the survival or in the reliability analysis, and it happens quite naturally in many real life applications, see for example Gomez et al. (2004).

First we consider the maximum likelihood estimation (MLE) of α and λ. It is observed that the maximum likelihood estimators (MLEs) do not exist in closed form, and they have to be obtained by solving two non-linear equations. Moreover, the standard Newton-Raphson algorithm may not even converge sometimes. Due to this reason, we have proposed to use the expectation maximization (EM) algorithm. It is observed that at each 'E' step of the EM algorithm, the corresponding 'M' step can be performed by solving a one dimensional optimization process, and we have provided a simple fixed point type algorithm to perform that. It is observed in our simulation experiment that EM algorithm converges almost all the times.

Next we compute the Bayes estimate of the unknown parameters under the assumption of gamma priors for both the shape and scale parameters. Note that the assumptions of gamma priors for the Weibull parameters are not very uncommon, see for example Berger and Sun (1993) or Kundu (2008). Non-informative priors also can be obtained as a special case of the gamma priors. With respect to the gamma priors, it is not possible to compute the Bayes estimates in explicit forms. They have to be obtained in terms of integrations only. We suggest to use Lindley's approximation and importance sampling procedure to compute approximate Bayes estimates. We also provide highest posterior density (HPD) credible interval for α and λ using importance sampling procedure. We perform some simulation experiments to see the performance of the different estimators, and analyze two data sets for illustrative purposes.

Rest of the paper is organized as follows. We provide the MLEs in Section 2. The Bayes estimation of the unknown parameters is discussed in Section 3. Simulation results are presented in Section 4. Two data sets are analyzed in Section 5. Finally we conclude the paper in Section 6.

2 Maximum Likelihood Estimator

In this section we provide the MLEs of α and λ. It is assumed that the observed data is as follows.

$$(T_1, 1), \ldots (T_{n_1}, 1), ([L_{n_1+1}, R_{n_1+1}], 0), \ldots, ([L_{n_1+n_2}, R_{n_1+n_2}], 0). \qquad (2.1)$$

Here n_1 and n_2, denote the number of uncensored and censored observations, respectively, and $n_1 + n_2 = n$. Based on the assumptions as described in the previous section, the likelihood function can then be written as

$$L(\alpha, \lambda | data) = c\alpha^{n_1} \lambda^{n_1} \prod_{i=1}^{n_1} t_i^{\alpha-1} e^{-\lambda \sum_{i=1}^{n_1} t_i^{\alpha}} \prod_{i=n_1+1}^{n_1+n_2} (e^{-\lambda l_i^{\alpha}} - e^{-\lambda r_i^{\alpha}}). \qquad (2.2)$$

Here c is the normalizing constant independent of α and λ. The log-likelihood function becomes

$$\ln L(\alpha, \lambda | data) = l(\alpha, \lambda) = \ln c + n_1 \ln \alpha + n_1 \ln \lambda + (\alpha - 1) \sum_{i=1}^{n_1} \ln t_i - \lambda \sum_{i=1}^{n_1} t_i^{\alpha} + \sum_{i=n_1+1}^{n_1+n_2} \ln(e^{-\lambda l_i^{\alpha}} - e^{-\lambda r_i^{\alpha}}).$$

The corresponding normal equations can be written as;

$$\frac{\partial l(\alpha, \lambda)}{\partial \alpha} = \frac{n_1}{\alpha} + \sum_{i=1}^{n_1} \ln t_i - \lambda \sum_{i=1}^{n_1} t_i^{\alpha} \ln t_i + \lambda \sum_{i=n_1+1}^{n_1+n_2} \frac{r_i^{\alpha} e^{-\lambda r_i^{\alpha}} \ln r_i - l_i^{\alpha} e^{-\lambda l_i^{\alpha}} \ln l_i}{e^{-\lambda l_i^{\alpha}} - e^{-\lambda r_i^{\alpha}}} = 0 \qquad (2.3)$$

$$\frac{\partial l(\alpha, \lambda)}{\partial \lambda} = \frac{n_1}{\lambda} - \sum_{i=1}^{n_1} t_i^{\alpha} + \sum_{i=n_1+1}^{n_1+n_2} \frac{r_i^{\alpha} e^{-\lambda r_i^{\alpha}} - l_i^{\alpha} e^{-\lambda l_i^{\alpha}}}{e^{-\lambda l_i^{\alpha}} - e^{-\lambda r_i^{\alpha}}} = 0. \qquad (2.4)$$

The MLEs of α and λ are obtained by solving (2.3) and (2.4) simultaneously. It is immediate that explicit solutions cannot be obtained from the above equations. Moreover, they cannot be reduced further, like Type-I, Type-II or progressive censoring cases. We need to apply a suitable numerical technique to solve the two non-linear equations. One may use Newton-Raphson or Gauss-Newton methods or their variants to solve these. The observed information matrix is given by

$$I(\alpha, \lambda) = \begin{pmatrix} U & V \\ V & W \end{pmatrix},$$

where the explicit expressions of U, V and W are provided in the Appendix. The observed information matrix may be used to construct the asymptotic confidence intervals of the unknown parameters. We have performed some simulation experiments to compute the MLEs using standard Newton-Raphson algorithm with different model parameters (the details are provided in the simulation section), and it is observed that the iteration converges between 82% to 85% of the times. Due to this reason we have proposed to use the following EM algorithm, and it is observed in our simulation experiments that it has a convergence rate almost 100% of the times.

2.1. EM Algorithm. We propose to use the EM algorithm to compute the MLEs of α and λ. In implementing the EM algorithm, we need to treat this problem as a missing value problem. The EM algorithm has two steps. The first step is the 'E'-step, where the 'pseudo-likelihood' function is formed from the likelihood function, by replacing the missing observations with their corresponding expected values. The second step of the EM algorithm is the 'M'-step, where the 'pseudo-likelihood' function is maximized to compute the parameters for the next iteration.

E-step: Suppose the censored observations are denoted by $\{Z_i; i = n_1 + 1, \ldots, n_1 + n_2\}$, then the pseudo likelihood function takes the form

$$L_c(\alpha, \lambda) = \alpha^n \lambda^n \prod_{i=1}^{n_1} t_i^{\alpha-1} \prod_{i=n_1+1}^{n_1+n_2} z_i^{\alpha-1} \times e^{-\lambda\left(\sum_{i=1}^{n_1} t_i^\alpha + \sum_{i=n_1+1}^{n_1+n_2} z_i^\alpha\right)},$$

where for $i = n_1 + 1, \cdots, n_1 + n_2$,

$$z_i = E(T|L_i < T < R_i) = \frac{\int_{L_i}^{R_i} \alpha\lambda x^\alpha e^{-\lambda x^\alpha} dx}{e^{-\lambda L_i^\alpha} - e^{-\lambda R_i^\alpha}}. \tag{2.5}$$

The pseudo log-likelihood function becomes

$$l_c(\alpha, \lambda) = n \ln \alpha + n \ln \lambda + (\alpha - 1)\left(\sum_{i=1}^{n_1} \ln t_i + \sum_{i=n_1+1}^{n_1+n_2} \ln z_i\right) - \lambda\left(\sum_{i=1}^{n_1} t_i^\alpha + \sum_{i=n_1+1}^{n_1+n_2} z_i^\alpha\right). \tag{2.6}$$

M-step: It involves maximization of pseudo log-likelihood function (2.6) with respect to α and λ to compute the next iterates. Let $(\alpha^{(k)}, \lambda^{(k)})$ be the estimate of (α, λ) at the k-th stage of the EM algorithm, then $(\alpha^{(k+1)}, \lambda^{(k+1)})$ can be obtained by maximizing

$$l_c^*(\alpha, \lambda) = n \ln \alpha + n \ln \lambda + (\alpha - 1)\left(\sum_{i=1}^{n_1} \ln t_i + \sum_{i=n_1+1}^{n_1+n_2} \ln z_i(\alpha^{(k)}, \lambda^{(k)})\right)$$

$$- \lambda \left(\sum_{i=1}^{n_1} t_i^{\alpha} + \sum_{i=n_1+1}^{n_1+n_2} z_i^{\alpha}(\alpha^{(k)}, \lambda^{(k)}) \right), \qquad (2.7)$$

with respect to α and λ, where $z_i(\alpha^{(k)}, \lambda^{(k)})$ can be obtained from (2.5) by replacing (α, λ) with $(\alpha^{(k)}, \lambda^{(k)})$. Note that for fixed α, the maximum of $l_c^*(\alpha, \lambda)$ with respect to λ occurs at $\lambda^{(k+1)}(\alpha)$, where

$$\lambda^{(k+1)}(\alpha) = \frac{n}{\sum_{i=1}^{n_1} t_i^{\alpha} + \sum_{i=n_1+1}^{n_1+n_2} z_i^{\alpha}(\alpha, \lambda^{(k)})}.$$

Clearly, for a given α, $\lambda^{(k+1)}(\alpha)$ is unique and it maximizes (2.7). Then $\alpha^{(k+1)}$ can be obtained by maximizing $l_c^*(\alpha, \lambda^{(k+1)}(\alpha))$, the 'pseudo-profile log-likelihood function', with respect to α. Using similar argument as in Theorem 2 of Kundu (2008), it can be shown that $l_c^*(\alpha, \lambda^{(k+1)}(\alpha))$ is an unimodal function of α, with an unique mode. Therefore, if $\alpha^{(k+1)}$ maximizes $l_c^*(\alpha, \lambda^{(k+1)}(\alpha))$, then $\alpha^{(k+1)}$ is unique. If $\alpha^{(k+1)}$ maximizes $l_c^*(\alpha, \lambda^{(k+1)}(\alpha))$, then it is immediate that $(\alpha^{(k+1)}, \lambda^{(k+1)}(\alpha^{(k+1)}))$ maximizes $l_c^*(\alpha, \lambda)$, as

$$l_c^*(\alpha, \lambda) \leq l_c^*(\alpha, \lambda^{(k+1)}(\alpha)) < l_c^*(\alpha^{(k+1)}, \lambda^{(k+1)}(\alpha^{(k+1)}),$$

and $(\alpha^{(k+1)}, \lambda^{(k+1)}(\alpha^{(k+1)})$ is the unique maximum of (2.7). The maximization of $l_c^*(\alpha, \lambda^{(k+1)}(\alpha))$ with respect to α can be performed by solving a fixed point type equation

$$g^{(k)}(\alpha) = \alpha, \qquad (2.8)$$

where

$$g^{(k)}(\alpha) = n \left[\frac{\sum_{i=1}^{n_1} t_i^{\alpha} \ln t_i + \sum_{i=n_1+1}^{n_1+n_2} z_i^{\alpha}(\alpha^{(k)}, \lambda^{(k)}) \ln z_i(\alpha^{(k)}, \lambda^{(k)})}{\sum_{i=1}^{n_1} t_i^{\alpha} + \sum_{i=n_1+1}^{n_1+n_2} z_i^{\alpha}(\alpha^{(k)}, \lambda^{(k)})} - \left(\sum_{i=1}^{n_1} \ln t_i + \sum_{i=n_1+1}^{n_1+n_2} \ln z_i(\alpha^{(k)}, \lambda^{(k)}) \right) \right]^{-1}.$$

Therefore, simple iterative process can be used to compute $(\alpha^{(k+1)}, \lambda^{(k+1)})$ from $(\alpha^{(k)}, \lambda^{(k)})$. At the k-th step, first solve (2.8), by using an iteration of the type $\alpha^{[i+1]} = g^{(k)}(\alpha^{[i]})$, the iteration continues until it converges. Once $\alpha^{(k+1)}$ is obtained, $\lambda^{(k+1)}$ is obtained as $\lambda^{(k+1)}(\alpha^{(k+1)})$.

3 Bayes Estimation

In this section we consider the Bayesian inference of α and λ, when we have the interval censored data as in (2.1). First we consider the Bayes estimation of the scale parameter λ when the shape parameter α is known and then we consider the case when both are unknown. We also consider the highest posterior density credible intervals of both the parameters. Before progressing further, first we make the prior selection.

3.1. Prior Information. When α is known, it is well known that the scale parameter has a conjugate gamma prior. Therefore, for known α, gamma prior on λ is the most natural one. If the shape parameter is also unknown, it is well known that the Weibull distribution does not have a continuous conjugate priors, although there exists continuous-discrete joint prior distributions, see Soland (1969). Here the continuous part corresponds to the scale parameter, and the discrete part corresponds to the shape parameter. This method has been highly criticized due to its difficulty in interpretations and applications, see Kaminskiy and Krivtsov (2005). Following the approaches of Berger and Sun (1993) and Kundu (2008), when both the parameters are unknown, it is assumed that both α and λ have gamma priors and they are independently distributed.

It is assumed that α and λ have the following independent gamma prior distributions.

$$\pi_1(\alpha|a,b) = f_{GA}(\alpha;a,b), \quad \alpha > 0, \tag{3.1}$$

$$\pi_2(\lambda|c,d) = f_{GA}(\lambda;c,d), \quad \lambda > 0. \tag{3.2}$$

Here for $\beta > 0$ and $\theta > 0$,

$$f_{GA}(x;\beta,\theta) = \frac{\theta^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\theta x}, \quad x > 0,$$

and it will be denoted by gamma(β,θ). Here all the hyper parameters a, b, c and d are assumed to be known.

3.2. Shape Parameter Known. Here we obtain the Bayes estimate of λ when α is known. Based on the observed sample, the likelihood function is given by

$$L(\alpha,\lambda|data) = c\alpha^{n_1}\lambda^{n_1}\prod_{i=1}^{n_1} t_i^{\alpha-1} e^{-\lambda\sum_{i=1}^{n_1} t_i^\alpha} \prod_{i=n_1+1}^{n_1+n_2} (1-e^{-\lambda v_i}) e^{-\lambda\sum_{i=n_1+1}^{n_1+n_2} l_i^\alpha}, \tag{3.3}$$

where $v_i = r_i^\alpha - l_i^\alpha$.

As in Iyer et al. (2009), by a slight abuse of notation, writing $v_i = v_{n_1+i}$ and $l_i = l_{n_1+i}$ we can write (3.3) as

$$L(\alpha,\lambda|data) = c\alpha^{n_1}\lambda^{n_1}\prod_{i=1}^{n_1} t_i^{\alpha-1} e^{-\lambda\sum_{i=1}^{n_1} t_i^\alpha} \prod_{i=1}^{n_2} (1-e^{-\lambda v_i}) e^{-\lambda\sum_{i=1}^{n_2} l_i^\alpha},$$

Based on the prior (3.2), the posterior distribution of λ is given by

$$\begin{aligned}\pi(\lambda|\alpha, data) &= \frac{L(\alpha,\lambda|data)\pi_2(\lambda|c,d)}{\int_0^\infty L(\alpha,\lambda|data)\pi_2(\lambda|c,d)d\lambda} \\ &= \frac{\lambda^{n_1+c-1}e^{-\lambda(d+\sum_{i=1}^{n_1} t_i^\alpha+\sum_{i=1}^{n_2} l_i^\alpha)}\prod_{i=1}^{n_2}(1-e^{-\lambda v_i})}{\int_0^\infty \lambda^{n_1+c-1}e^{-\lambda(d+\sum_{i=1}^{n_1} t_i^\alpha+\sum_{i=1}^{n_2} l_i^\alpha)}\prod_{i=1}^{n_2}(1-e^{-\lambda v_i})d\lambda}.\end{aligned} \tag{3.4}$$

To simplify (3.4), we use the expansion

$$\prod_{i=1}^{n_2}(1-e^{-\lambda v_i}) = \sum_{P_j}(-1)^{|P_j|}e^{-\lambda(v.P_j)}, \tag{3.5}$$

where P_j is a vector of length n_2 and each entry of P_j is either a 0 or a 1. Here $|P_j|$ denotes the sum of elements of P_j and $v = (v_1, \ldots, v_{n_2})$. The summation on the right-hand side of (3.5) is over 2^{n_2} elements and $(v.P_j)$ denote the dot product between the two vectors of equal lengths. Using (3.5), (3.4) can be written as (See Iyer et al., 2009)

$$\pi(\lambda|\alpha, data) = \frac{\sum_{P_j}(-1)^{|P_j|}\lambda^{n_1+c-1}e^{-\lambda(d+\sum_{i=1}^{n_1} t_i^\alpha+\sum_{i=1}^{n_2} l_i^\alpha+(v.P_j))}}{\sum_{P_j}(-1)^{|P_j|}\Gamma(c+n_1)/(d+\sum_{i=1}^{n_1} t_i^\alpha+\sum_{i=1}^{n_2} l_i^\alpha+(v.P_j))^{c+n_1}}. \tag{3.6}$$

Therefore, the Bayes estimate of λ under squared error loss function is

$$E(\lambda|data) = \frac{\sum_{P_j}(-1)^{|P_j|}/(d+\sum_{i=1}^{n_1} t_i^\alpha+\sum_{i=1}^{n_2} l_i^\alpha+(v.P_j))^{c+n_1+1}}{\sum_{P_j}(-1)^{|P_j|}/(d+\sum_{i=1}^{n_1} t_i^\alpha+\sum_{i=1}^{n_2} l_i^\alpha+(v.P_j))^{c+n_1}}. \tag{3.7}$$

Although (3.7) can be evaluated easily for small n_2, it is difficult to compute numerically for large n_2. In general we propose Gibbs sampling technique to compute $E(\lambda|data)$. It is easily seen that when $n_2 = 0$, then

$$\pi(\lambda|\alpha, data) \sim \text{Gamma}(c+n_1, d+\sum_{i=1}^{n_1} t_i^\alpha). \tag{3.8}$$

So we can choose the initial value for the parameter by generating observation from the above distribution. Also, the conditional density of T, given $T \in (L,R)$, is

$$f_{T|T\in(L,R)}(x|\lambda) = \frac{\alpha\lambda t^{\alpha-1}e^{-\lambda t^\alpha}}{e^{-\lambda L^\alpha}-e^{-\lambda R^\alpha}} \quad \text{if } L < x < R. \tag{3.9}$$

This conditional distribution can be used to update the value of λ in Gibbs sampling scheme. Hence using (3.8) and (3.9), the algorithm for Gibbs sampling procedure to generate λ from the posterior distribution is described below.

Algorithm:

- Step 1: Generate $\lambda_{1,1}$ from Gamma$(c+n_1, d+\sum_{i=1}^{n_1} t_i^{\alpha})$.
- Step 2: Generate t^{n_1+i} for $i = 1, \ldots, n_2$ from $f_{T|T\in(l_{n_1+i}, r_{n_1+i})}(\cdot|\lambda_{1,1})$.
- Step 3: Generate $\lambda_{2,1}$ from Gamma$(c+n_1, d+\sum_{i=1}^{n_1} t_i^{\alpha}+\sum_{i=n_1+1}^{n_1+n_2}(t^{(i)})^{\alpha})$.
- Step 4: Go back to Step 2, and replace $\lambda_{1,1}$ by $\lambda_{2,1}$ and repeat Steps 2 and 3 for N times.

The Bayes estimate of λ under squared error loss function is then given by

$$\frac{1}{N-M}\sum_{j=M+1}^{N}\lambda_{2,j},$$

where M is the burn-in sample. Hence HPD credible interval can also be obtained using the method of Chen and Shao (1999).

3.3. Both Parameters Unknown. We compute the Bayes estimate of the unknown parameters under squared error loss function using the priors in (3.1) and (3.2). Using the likelihood function (2.2), the joint distribution of data, α, and λ can be written as

$$L(\alpha, \lambda|data)\pi_1(\alpha|a, b)\pi_2(\lambda|c, d).$$

Then, the joint posterior density of α and λ given the data is

$$\pi(\alpha, \lambda|data) = \frac{L(\alpha, \lambda|data)\pi_1(\alpha|a, b)\pi_2(\lambda|c, d)}{\int_0^{\infty}\int_0^{\infty} L(\alpha, \lambda|data)\pi_1(\alpha|a, b)\pi_2(\lambda|c, d)d\alpha d\lambda}.$$

Let $g(\alpha, \lambda)$ be any function of α and λ. Then, the Bayes estimate of $g(\alpha, \lambda)$ under squared error loss function is given by

$$\begin{aligned}\widehat{g}_B(\alpha, \lambda) &= E_{\alpha,\lambda|data}(g(\alpha, \lambda)) \\ &= \frac{\int_0^{\infty}\int_0^{\infty} g(\alpha, \lambda)L(\alpha, \lambda|data)\pi_1(\alpha|a, b)\pi_2(\lambda|c, d)d\alpha d\lambda}{\int_0^{\infty}\int_0^{\infty} L(\alpha, \lambda|data)\pi_1(\alpha|a, b)\pi_2(\lambda|c, d)d\alpha d\lambda}.\end{aligned} \tag{3.10}$$

It is clear from the expression (3.10) that there is no closed form of the estimators. We suggest Lindley's approximation and Importance Sampling procedure to compute the Bayes estimates.

3.4. Lindley's Approximation. Lindley's approximation has been used to approximate the ratio of two integrals in many occasion. Using Lindley's approximation, the Bayes estimates of α and λ are

$$\begin{aligned}\widehat{\alpha}_B &= \widehat{\alpha} + \frac{1}{2}\left(l_{30}\tau_{11}^2 + l_{03}\tau_{21}\tau_{22} + 3l_{21}\tau_{11}\tau_{12} + l_{12}(\tau_{22}\tau_{11} + 2\tau_{21}^2)\right) \\ &\quad + \left(\frac{a-1}{\widehat{\alpha}} - b\right)\tau_{11} + \left(\frac{c-1}{\widehat{\lambda}} - d\right)\tau_{12} \qquad (3.11)\end{aligned}$$

and

$$\begin{aligned}\widehat{\lambda}_B &= \widehat{\lambda} + \frac{1}{2}\left(l_{30}\tau_{12}\tau_{11} + l_{03}\tau_{22}^2 + l_{21}(\tau_{11}\tau_{22} + 2\tau_{12}^2) + 3l_{12}\tau_{22}\tau_{21}\right) \\ &\quad + \left(\frac{a-1}{\widehat{\alpha}} - b\right)\tau_{21} + \left(\frac{c-1}{\widehat{\lambda}} - d\right)\tau_{22}. \qquad (3.12)\end{aligned}$$

Here $\widehat{\alpha}$ and $\widehat{\lambda}$ are the MLEs of α and λ respectively, and a, b, c, d are the known hyper-parameters. The explicit expressions of τ_{11}, τ_{12}, τ_{21}, τ_{22}, τ_{30}, τ_{03}, l_{21}, l_{12} are provided in Appendix.

3.5. Importance Sampling. In this section, we apply importance sampling procedure to compute Bayes estimates and HPD credible intervals. The joint posterior density function of α and λ using the priors (3.1) and (3.2) can be written as

$$\begin{aligned}\pi(\alpha,\lambda|data) &\propto \alpha^{a+n_1-1}e^{-\alpha b}\lambda^{c+n_1-1}e^{-\lambda(d+\sum_{i=1}^{n_1}t_i^\alpha)}\prod_{i=1}^{n_1}t_i^{\alpha-1} \\ &\quad\times\prod_{i=n_1+1}^{n_1+n_2}(e^{-\lambda l_i^\alpha} - e^{-\lambda r_i^\alpha}) \\ &\propto f_{GA}(\alpha : a+n_1, b-\sum_{i=1}^{n_1}\ln t_i)f_{GA}(\lambda; c+n_1, d+\sum_{i=1}^{n_1}t_i^\alpha + \sum_{i=n_1+1}^{n_1+n_2}l_i^\alpha) \\ &\quad\times h(\alpha,\lambda), \qquad (3.13)\end{aligned}$$

where $h(\alpha,\lambda) = \dfrac{\prod_{i=n_1+1}^{n_1+n_2}(1-e^{-\lambda(r_i^\alpha - l_i^\alpha)})}{(d+\sum_{i=1}^{n_1}t_i^\alpha + \sum_{i=n_1+1}^{n_1+n_2}l_i^\alpha)^{c+n_1}}$.

Let us denote the right-hand side of (3.13) as $\pi_N(\alpha,\lambda|data)$. Note that $\pi(\alpha,\lambda|data)$ and $\pi_N(\alpha,\lambda|data)$ differ only by the proportionality constant. The Bayes estimate of $g(\alpha,\lambda)$ under squared error loss function is given by

$$\widehat{g}_B(\alpha,\lambda) = \frac{\int_0^\infty\int_0^\infty g(\alpha,\lambda)\pi_N(\alpha,\lambda|data)d\alpha d\lambda}{\int_0^\infty\int_0^\infty \pi_N(\alpha,\lambda|data)d\alpha d\lambda}. \qquad (3.14)$$

It is clear from (3.14) that to approximate $\widehat{g}_B(\alpha,\lambda)$, using the importance sampling procedure one need not compute the normalizing constant. We use the following procedure:

- Step 1: Generate

$$\alpha_1 \sim \text{gamma}\left(a+n_1, b-\sum_{i=1}^{n_1}\ln t_i\right) \quad \text{and}$$
$$\lambda_1|\alpha_1 \sim \text{gamma}\left(c+n_1, d+\sum t_i^{\alpha}+\sum_{i=n_1+1}^{n_1+n_2} l_i^{\alpha}\right).$$

- Step 2: Repeat this procedure to obtain $(\alpha_1,\lambda_1),\ldots,(\alpha_N,\lambda_N)$.

- Step 3: The approximate value of (3.14) can be obtained as

$$\frac{\sum_{i=1}^N g(\alpha_i,\lambda_i)h(\alpha_i,\lambda_i)}{\sum_{i=1}^N h(\alpha_i,\lambda_i)}.$$

So the Bayes estimates of α and λ are

$$\hat{\alpha}_{IS} = \frac{\sum_{i=1}^N \alpha_i h(\alpha_i,\lambda_i)}{\sum_{i=1}^N h(\alpha_i,\lambda_i)} \text{ and } \hat{\lambda}_{IS} = \frac{\sum_{i=1}^N \lambda_i h(\alpha_i,\lambda_i)}{\sum_{i=1}^N h(\alpha_i,\lambda_i)}.$$

The corresponding posterior variances can be obtained as

$$\frac{\sum_{i=1}^N \alpha_i^2 h(\alpha_i,\lambda_i)}{\sum_{i=1}^N h(\alpha_i,\lambda_i)} - \hat{\alpha}_{IS}^2 \quad \text{and} \quad \frac{\sum_{i=1}^N \lambda_i^2 h(\alpha_i,\lambda_i)}{\sum_{i=1}^N h(\alpha_i,\lambda_i)} - \hat{\lambda}_{IS}^2.$$

Next we discuss how to obtain HPD credible intervals of α and λ. The advantage of importance sampling procedure is that the generated sample can be used to construct the HPD credible intervals. We illustrate the procedure for the parameter α, but it can be similarly obtained for any other functions of the parameters also. Suppose a_p is such that

$$P[\alpha \le a_p|data] = p.$$

Now consider the following function

$$g(\alpha,\lambda) = \begin{cases} 1 & \text{if} \quad \alpha \le a_p \\ 0 & \text{if} \quad \alpha > a_p. \end{cases}$$

Clearly,

$$E(g(\alpha, \lambda)|data) = p.$$

Therefore, an approximate Bayes estimate of a_p under the squared error loss function can be obtained from the generated sample $\{(\alpha_1, \lambda_1), \ldots, (\alpha_N, \lambda_N)\}$ as follows. Let

$$w_i = \frac{h(\alpha_i, \lambda_i)}{\sum_{i=1}^{N} h(\alpha_i, \lambda_i)}; \qquad i = 1, \ldots, N.$$

Rearrange $\{(\alpha_1, w_1), \ldots, (\alpha_N, w_N)\}$ as $\{(\alpha_{(1)}, w_{(1)}), \ldots, (\alpha_{(N)}, w_{(N)})\}$, where $\alpha_{(1)} < \ldots < \alpha_{(N)}$. Note that $w_{(i)}$'s are not ordered, they are just associated with $\alpha_{(i)}$. Then an Bayes estimate of a_p is

$$\widehat{a}_p = \alpha_{(N_p)},$$

where N_p is the integer satisfying

$$\sum_{i=1}^{N_p} w_{(i)} \leq p < \sum_{i=1}^{N_p+1} w_{(i)}.$$

Now using the above procedure a $100(1-\gamma)\%$ credible interval of α can be obtained as $(\widehat{a}_\delta, \widehat{a}_{\delta+1-\gamma})$, for $\delta = w_{(1)}, w_{(1)} + w_{(2)}, \ldots, \sum_{i=1}^{N_\gamma} w_{(i)}$. Therefore, an $100(1-\gamma)\%$ HPD credible interval of α becomes

$$(\widehat{a}_{\delta^*}, \quad \widehat{a}_{\delta^*+1-\gamma}),$$

where δ^* is such that

$$\widehat{a}_{\delta^*+1-\gamma} - \widehat{a}_{\delta^*} \leq \widehat{a}_{\delta+1-\gamma} - \widehat{a}_\delta \quad \text{for all} \quad \delta.$$

4 Simulation Study

We generate interval-censored observations from the Weibull distribution with $\alpha = 1.5$ and $\lambda = 1$ for different sample size. For the random censoring intervals, consider $Z_i = R_i - L_i$, L_i and Z_i are independent exponential random variables and they are independent of T_i. We generate observations from Z_i and L_i and hence get $R_i = Z_i + L_i$. We assume that L_i and Z_i have means $1/\theta_1$ and $1/\theta_2$, respectively. The simulation is carried out for sample sizes $n = 20$, 30, 50 and 100, and for different choices of (θ_1, θ_2). We choose $(\theta_1, \theta_2) = (0.50, 0.75)$ (scheme 1), (1.25, 1.0) (scheme 2) and (1.50, 0.25) (scheme 3) for the simulation study. These three schemes correspond to different proportion of censored observation. Note that The proportion of censored (PC) observation under an interval censoring scheme is given by

$$\begin{aligned} PC &= \int_0^\infty \int_0^\infty (e^{-\lambda l^\alpha} - e^{-\lambda(l+z)^\alpha})\theta_1\theta_2 e^{-(\theta_1 l+\theta_2 z)} dl dz \\ &= \theta_1 \int_0^\infty e^{-\lambda l^\alpha} e^{-\theta_1 l} dl - \theta_1\theta_2 \int_0^\infty \int_0^\infty e^{-\lambda(l+z)^\alpha} e^{-(\theta_1 l+\theta_2 z)} dl dz. \end{aligned}$$

The above integrals are computed using R software (Ri386 3.0.0). The proportion of censoring under different choices of θ_1 and θ_2 are given in Table 1.

For each set of the simulated data, we generate observations of the form (1.1) and calculate different estimates of α and λ. In each case, the MLEs are obtained by EM algorithm. We replicate the process 1000 times. We have computed the MLEs in all these cases by using standard Newton-Raphson algorithm and also by using the EM algorithm. Newton-Raphson algorithm converges between 82%–85% of the times, where as the EM algorithm converges all the times, except only two cases ((i) Scheme 1, $n = 50$ and (ii) Scheme 2, $n = 50$). We report the results based on the EM algorithm ignoring those two cases. We report the average bias (AB) and means squared error (MSE) in Table 1. From the simulation study, it is clear that as sample

Table 1: The average bias (AB) and mean squared error (MSE) in parentheses corresponding to MLEs of α and λ corresponding to different proportion of censoring (PC).

(θ_1, θ_2)	PC	n	α	λ
(0.50, 0.75)	0.22	20	0.2367 (0.2195)	0.0597 (0.1068)
(Scheme 1)		30	0.1634 (0.1013)	0.0307 (0.0575)
		50	0.1299 (0.0607)	0.0163 (0.0308)
		100	0.1029 (0.0305)	0.0035 (0.0148)
(1.25, 0.75)	0.37	20	0.2954 (0.2798)	0.0626 (0.1134)
(Scheme 2)		30	0.2157 (0.1329)	0.0363 (0.0657)
		50	0.1653 (0.0754)	0.0138 (0.0341)
		100	0.1417 (0.0436)	0.0014 (0.0160)
(1.50, 0.25)	0.54	20	0.7537 (1.1394)	0.4026 (1.3749)
(Scheme 3)		30	0.6047 (0.6253)	0.2274 (0.3383)
		50	0.5036 (0.3666)	0.1459 (0.1303)
		100	0.4564 (0.2610)	0.0991 (0.0489)

size increases the biases and MSEs decrease, as expected. The performances of the estimators are better when the proportion censored observation is less.

5 Applications

Here we illustrate our methodology with two examples. The first example deals with a simulated data set and the second one is based on real life data.

EXAMPLE 5.1. In this example we consider a simulated data set for $n = 30$. Interval censored data were generated by taking $\alpha = 1.5$, $\lambda = 1$ and $(\theta_1,\ \theta_2) = (0.50,\ 0.75)$. In the generated data set, we have 24 complete observations and 6 censoring intervals. The complete observations are

0.8820	1.1739	0.4123	0.4565	1.9935	1.0662	1.3516	0.3130
1.3364	1.6493	0.3000	0.8187	0.0253	0.6841	0.2672	1.1791
0.3460	0.8371	0.9184	0.8331	0.5123	0.1045	0.2159	0.0992

and the censoring intervals are [0.7286, 2.7756], [0.4465, 1.7119], [0.0204, 2.7927], [0.6566, 1.9712], [1.5674, 2.4757], [0.1700, 2.3342].

The maximum likelihood estimate of the parameters with standard error in parentheses are $\hat{\alpha} = 1.4945$ (0.2300) and $\hat{\lambda} = 1.1864$ (0.2338). The asymptotic variance-covariance matrix based on observed information matrix is given by

$$I^{-1}(\hat{\alpha}, \hat{\lambda}) = \begin{bmatrix} 0.0530 & -0.0091 \\ -0.0091 & 0.0547 \end{bmatrix}.$$

The asymptotic 95% confidence intervals of α and λ are (1.0437, 1.9453) and (0.7281, 1.6447), respectively. Next we compute the Bayes estimates of α and λ. In the absence of any prior information, we compute Bayes estimates under non-informative prior. The Bayes estimate of α and λ by Lindley's approximation method are $\hat{\alpha}_B = 1.4944$ and $\hat{\lambda}_B = 1.1845$. We also compute Bayes estimate based on 1000 importance sample under non-informative prior. The estimates are $\hat{\alpha}_{IS} = 1.4374$ (0.2216) and $\hat{\lambda}_{IS} = 1.1745$ (0.2281). The 95% HPD credible intervals of α and λ are [1.0103, 1.8921] and [0.7705, 1.6334], respectively.

EXAMPLE 5.2. Here we consider a real life data set (Finkelstein, 1986; Lindsey and Ryan, 1998) from a retrospective study of patients with breast cancer. The study was designed to compare radiation therapy alone versus in combination with chemotherapy with respect to the time to cosmetic deterioration. This example has been analyzed by several authors to illustrate various methods for interval censored data. For illustration, we analyze the data set corresponding to combined radiotherapy and chemotherapy group.

Patients were seen initially every 4 to 6 months, with decreasing frequency over time. If deterioration was seen, it was known only to have occurred between two visits. Deterioration was not observed in all patients during the course of the trial, so some data were right-censored. The data set is, available in Lawless (2003, pp. 143), provided below for convenience.

(8, 12], (0, 22], (24, 31], (17, 27], (17, 23], (24, 30], (16, 24], (13, ∞), (11, 13], (16, 20], (18, 25], (17, 26], (32, ∞), (23, ∞), (44, 48], (10, 35], (0, 5], (5, 8], (12, 20], (11, ∞), (33, 40], (31, ∞), (13, 39], (19, 32], (34, ∞), (13, ∞), (16, 24], (35, ∞), (15, 22], (11, 17], (22, 32], (48, ∞), (30, 34], (13, ∞), (10, 17], (8, 21], (4, 9], (11, ∞), (14, 19], (4, 8], (34, ∞), (30, 36], (18, 24], (16, 60], (35, 39], (21, ∞), (11, 20].

The observations with $L = 0$ are left censored and with $R = \infty$ are right censored. Let n_1 be the number of left censored observation, n_2 the number of right censored observations and n_3 the number of interval censored observations. There are total 47 observations with $n_1 = 2$, $n_2 = 13$ and $n_3 = 32$. Note that there is no complete observation. The likelihood function can then be written as

$$L(\alpha, \lambda|data) = \prod_{i=1}^{n_1}(1 - e^{-\lambda r_i^\alpha}) \prod_{i=n_1+1}^{n_1+n_2} e^{-\lambda l_i^\alpha} \prod_{i=n_1+n_2+1}^{n_1+n_2+n_3} (e^{-\lambda l_i^\alpha} - e^{-\lambda r_i^\alpha}), \quad (5.1)$$

Note that the likelihood (5.1) is a special case of (2.2). The maximum likelihood estimate of α and λ are $\hat{\alpha} = 2.0234$ and $\hat{\lambda} = 0.0012$. The asymptotic variance-covariance matrix is given by

$$I^{-1}(\hat{\alpha}, \hat{\lambda}) = \begin{bmatrix} 0.08432 & -0.00033 \\ -0.00033 & 1.3175 \times 10^{-7} \end{bmatrix}.$$

The 95% asymptotic confidence intervals of α and λ are (1.8666, 2.1802) and (0.0005, 0.0019), respectively. Therefore, it is clear that the exponential distribution cannot be used to analyze this data set.

Next we obtain the Bayes estimate of the unknown parameters. The posterior distribution under gamma priors of α and λ is given by

$$\begin{aligned} \pi(\alpha, \lambda|data) &= \prod_{i=1}^{n_1}(1 - e^{-\lambda r_i^\alpha}) \prod_{i=n_1+1}^{n_1+n_2} e^{-\lambda l_i^\alpha} \prod_{i=n_1+n_2+1}^{n_1+n_2+n_3} (e^{-\lambda l_i^\alpha} - e^{-\lambda r_i^\alpha}) \\ &\quad \times \alpha^{a-1} e^{-b\alpha} \lambda^{c-1} e^{-d\lambda}. \end{aligned}$$

Since we do not have any prior information we consider only non-informative priors to obtain Bayes estimates. We obtain Bayes estimates both by Lindley's approximation and MCMC technique. The estimates by Lindley's approximation method are $\hat{\alpha}_B = 1.9642$ and $\hat{\lambda}_B = 0.0019$.

In this case the importance sampling technique can not be applied because it is not possible to generate importance sample. We obtain the Bayes estimates under squared error loss function using Metropolis-Hastings (M-H) algorithm. Let $\theta = (\alpha, \lambda)$. We consider symmetric proposal density of type $q(\theta'|\theta) = q(\theta|\theta')$. In particular, we take bivariate normal distribution as the proposal density. That is, we take

$$q(\theta'|\theta) \equiv N_2(\theta, S_\theta),$$

where S_θ is the variance-covariance matrix. It may be noted that if we generate observation from bivariate normal distribution, we may get negative observations which are not acceptable as the parameters under consideration are positive valued. Keeping this in mind, the M-H algorithm steps are given below.

1. Set initial values $\theta^{(0)}$.

2. For $t = 1, \ldots, T$ repeat the following steps.

   - Set $\theta = \theta^{(t-1)}$
   - Generate new candidate parameter values δ from $N_2(\log(\theta), S_\theta)$.
   - Set $\theta' = \exp(\delta)$
   - Calculate $\alpha = \min\left(1, \dfrac{\pi(\theta'|x)\theta'_1\theta'_2}{\pi(\theta|x)\theta_1\theta_2}\right)$
   - d. Update $\theta^{(t)} = \theta'$ with probability α; otherwise set $\theta^{(t)} = \theta$.

The MLE of θ is considered as initial value for θ. The choice of covariance matrix S_θ is an important issue, see Natzoufras (2009) for details. One choice for S_θ is the asymptotic variance-covariance matrix $I^{-1}(\hat{\alpha}, \hat{\lambda})$. We generate M-H samples with $S_\theta = I^{-1}(\hat{\alpha}, \hat{\lambda})$, but the acceptance rate for this choice of S_θ is about 10%. By acceptance rate, we mean the proportion of times a new sample is generated as the sampling proceeds. When acceptance rate is low, a good strategy is to run a small pilot run using diagonal S_θ to roughly estimate the correlation structure of the target posterior distribution (see Natzoufras, 2009) and then rerun the algorithm using the corresponding estimated variance-covariance matrix $\tilde{\Sigma}_\theta$. Gelmen et al. (1995, pp. 334–335) suggested that the proposal variance-covariance matrix can be taken as $S_\theta = c^2\tilde{\Sigma}_\theta$ with $c^2 \approx 5.8/d$, where d is the number of parameters. Here we first carry out a pilot run for $N = 2000$ with diagonal S_θ and obtain $\tilde{\Sigma}_\theta$ using these 2000 samples. Then we rerun the algorithm taking $S_\theta = 2.9\tilde{\Sigma}_\theta$ for

10000 times. The acceptance rate is 25.6%. We discard the initial 1000 burn-in sample and calculate the estimates based on the remaining observations. The estimates of the parameters are $\hat{\alpha}_{MH} = 2.1910$ and $\hat{\lambda}_{MH} = 0.0011$. The 95% HPD credible intervals for α and λ are [1.8092, 2.2460] and [0.0010, 0.0033], respectively.

6 Conclusions

In this work we have considered both classical and Bayesian analysis of the interval censored data, when the lifetime of the items follows Weibull distribution. The MLEs do not have explicit forms. EM algorithm has been used to compute the MLEs and it works quite well. The Bayes estimates under the squared error loss function also do not exist in explicit form. We have proposed to use importance sampling technique to compute the Bayes estimates when the shape and scale parameters have independent gamma priors. Although we have considered gamma prior on the shape parameter, but a more general prior, namely a prior which has the log-concave PDF may be used, and the method can be easily incorporated in that case. We have not considered any covariates in this paper. But in practice often the covariates may be present. It will be interesting to develop statistical inference of the unknown parameters in presence of covariates. More work is needed in that direction.

Acknowledgement. The authors are grateful to unknown reviewers for their valuable suggestions and comments. Part of the work of the first author is supported by the project "Optimization and Reliability Modeling" funded by Indian Statistical Institute. Part of the work of the second author is supported by a grant from the Department of Science and Technology, Government of India.

References

BERGER, J.O. and SUN, D. (1993). Bayesian analysis for the Poly-Weibull distribution. *J. Amer. Statist. Assoc.*, **88**, 1412–1418.

CHEN, M.-H. and SHAO, Q.-M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals. *J. Comput. Graph. Statist.*, **8**, 69–92.

FINKELSTEIN, D.M. (1986). A proportional hazards model for interval-censored failure time data. *Biometrics*, **42**, 845–865.

GELMEN, A., CARLIN, J., STERN, H. and RUBIN, D. (1995). *Bayesian data analysis. Text in Statistical Science*. Chapman & Hall, London.

GOMEZ, G., CALLE, M.L. and OLLER, R. (2004). Frequentest and Bayesian approaches for interval-censored data. *Statist. Papers*, **45**, 139–173.

IYER, S.K., JAMMALAMADAKA, S.R. and KUNDU, D. (2009). Analysis of middle-censored data with exponential lifetime distributions. *J. Statist. Plann. Inference*, **138**, 3550–3560.

JAMMALAMADAKA, S.R. and MANGALAM, V. (2003). Non-parametric estimation for middle censored data. *J. Nonparametr. Stat.*, **15**, 253–265.

JAMMALAMADAKA, S.R. and IYER, S.K. (2004). Approximate self consistency for middle censored data. *J. Statist. Plann. Inference*, **124**, 75–86.

KALBFLEISH, J.D. and PRENTICE, R.L. (2002). *The statistical analysis of failure time data*, 2nd edition. John Wiley and Sons, New York.

KAMINSKIY, M.P. and KRIVTSOV, V.V. (2005). A simple procedure for Bayesian estimation of the Weibull distribution. *IEEE Trans. on Rel.*, **54**, 612–616.

KUNDU, D. (2008). Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring. *Technometrics*, **50**, 144–154.

LAWLESS, J.F. (2003). *Statistical models and methods for lifetime data.* Wiley, New York.

LINDSEY, J.C. and RYAN, L.M. (1998). Tutorials in biostatistics: Methods for interval-censored data. *Stat. Med.*, **17**, 219–238.

NATZOUFRAS, I. (2009). *Bayesian modeling using WinBugs.* Wiley, New York.

SPARLING, Y.H., YOUNES, N. and LACHIN, J.M. (2006). Parametric survival models for interval-censored data with time-dependent covariates. *Biostatistics*, **7**, 599–614.

SOLAND, R. (1969). Bayesian analysis of the Weibull process with unknown scale and shape parameters. *IEEE Trans. on Rel.*, **18**, 181–184.

Appendix

Note that we have two parameters α and λ. Let $\pi_0(\alpha, \lambda)$ be the joint prior distribution of α and λ. Using the notation $(\lambda_1, \lambda_2) = (\alpha, \lambda)$, the Lindley's approximation can be written as

$$\widehat{g} = g(\widehat{\lambda}_1, \widehat{\lambda}_2) + \frac{1}{2}\left(A + l_{30}B_{12} + l_{03}B_{21} + l_{21}C_{12} + l_{12}C_{21}\right) + p_1A_{12} + p_2A_{21},$$

where

$$A = \sum_{i=1}^{2}\sum_{j=1}^{2} w_{ij}\tau_{ij}, \quad l_{ij} = \frac{\partial^{i+j} l(\lambda_1, \lambda_2)}{\partial\lambda_i^i \lambda_2^j}, i, j = 0, 1, 2, 3, \quad \text{and}$$

$$i + j = 3, \quad p_i = \frac{\partial p}{\partial\lambda_i}, \quad w_i = \frac{\partial g}{\partial\lambda_i},$$

$$w_{ij} = \frac{\partial^2 g}{\partial\lambda_i\partial\lambda_j}, \quad p = \ln\pi_0(\lambda_1, \lambda_2), \quad A_{ij} = w_i\tau_{ii} + w_j\tau_{ji}, \quad B_{ij} = (w_i\tau_{ii} + w_j\tau_{ij})\tau_{ii},$$

$$c_{ij} = 3w_i\tau_{ii}\tau_{ij} + w_j(\tau_{ii}\tau_{jj} + 2\tau_{ij}^2).$$

Now when $g(\alpha, \lambda) = \alpha$, we have $w_1 = 1,\ w_2 = 0,\ w_{ij} = 0,\ i, j = 1, 2$, then

$$A{=}0,\ B_{12} = \tau_{11}^2,\ B_{21} = \tau_{21}\tau_{22},\ C_{12} = 3\tau_{11}\tau_{12},\ C_{21} = (\tau_{22}\tau_{11}{+}2\tau_{21}^2)\ A_{12} = \tau_{11},\ A_{21} = \tau_{12}.$$

Now (3.11) follows by using

$$p_1 = \left(\frac{a-1}{\widehat{\alpha}} - b\right) \quad \text{and} \quad p_2 = \left(\frac{c-1}{\widehat{\lambda}} - d\right).$$

For (3.12), note that $g(\alpha,\lambda)=\lambda$; then

$$w_1=0,\ w_2=1,\ w_{ij}=0,\ i,j=1,2;$$

and

$$A=0,\ B_{12}=\tau_{12}\tau_{11},\ B_{21}=\tau_{22}^2,\ C_{12}=\tau_{11}\tau_{22}+2\tau_{12}^2,\ C_{21}=3\tau_{22}\tau_{21}\ A_{12}=\tau_{21},\ A_{21}=\tau_{22}.$$

Here we have

$$l(\alpha,\lambda) = \ln c+n_1\ln\alpha+n_1\ln\lambda+(\alpha-1)\sum_{i=1}^{n_1}\ln t_i-\lambda\sum_{i=1}^{n_1}t_i^{\alpha} + \sum_{i=n_1+1}^{n_1+n_2}\ln(e^{-\lambda l_i^{\alpha}}-e^{-\lambda r_i^{\alpha}}),$$

$$\tau_{11}=\frac{W}{UW-V^2},\quad \tau_{12}=-\frac{V}{UW-V^2},\quad \text{and}\quad \tau_{22}=\frac{U}{UW-V^2}$$

where

$$U=-\frac{\partial^2 l(\alpha,\lambda)}{\partial\alpha^2}=\frac{n_1}{\alpha^2}+\lambda\sum_{i=1}^{n_1}t_i^{\alpha}(\ln t_i)^2-\sum_{n_1+1}^{n_1+n_2}\frac{\xi_i\xi_{\alpha i}'-\xi_{\alpha i}^2}{\xi_i^2},$$

$$V=-\frac{\partial^2 l(\alpha,\lambda)}{\partial\alpha\partial\lambda}=\sum_{i=1}^{n_1}t_i^{\alpha}\ln t_i-\sum_{n_1+1}^{n_1+n_2}\frac{\xi_i\xi_{\alpha\lambda i}-\xi_{\alpha i}\xi_{\lambda i}}{\xi_i^2},$$

$$W=-\frac{\partial^2 l(\alpha,\lambda)}{\partial\lambda^2}=\frac{n_1}{\lambda^2}-\sum_{n_1+1}^{n_1+n_2}\frac{\xi_i\xi_{\lambda i}'-\xi_{\lambda i}^2}{\xi_i^2},$$

$$l_{30}=\frac{\partial^3 l(\alpha,\lambda)}{\partial\alpha^3}=\frac{2n_1}{\alpha^3}-\lambda\sum_{i=1}^{n_1}t_i^{\alpha}(\ln t_i)^3+\sum_{n_1+1}^{n_1+n_2}\frac{2\xi_{\alpha i}^3-3\xi_i\xi_{\alpha i}\xi_{\alpha i}'+\xi_i^2\xi_{\alpha i}''}{\xi_i^3},$$

$$l_{03}=\frac{\partial^3 l(\alpha,\lambda)}{\partial\lambda^3}=\frac{2n_1}{\lambda^3}+\sum_{n_1+1}^{n_1+n_2}\frac{2\xi_{\lambda i}^3-3\xi_i\xi_{\lambda i}\xi_{\lambda i}'+\xi_i^2\xi_{\lambda i}''}{\xi_i^3},$$

$$l_{12}=\frac{\partial^3 l(\alpha,\lambda)}{\partial\alpha\partial\lambda^2}=\sum_{n_1+1}^{n_1+n_2}\frac{\xi_i(\xi_i\xi_{\alpha\lambda i}'-\xi_{\alpha i}\xi_{\lambda i}')-2\xi_{\lambda i}(\xi_i\xi_{\alpha\lambda i}-\xi_{\alpha i}\xi_{\lambda i})}{\xi_i^3},$$

$$l_{21}=-\sum_{i=1}^{n_1}t_i^{\alpha}(\ln t_i)^2+\sum_{n_1+1}^{n_1+n_2}\frac{\xi_i(\xi_{\alpha i}'\xi_{\lambda i}+\xi_i\xi_{\alpha\lambda i}''-2\xi_{\alpha i}\xi_{\alpha\lambda i})-2\xi_{\lambda i}(\xi_i\xi_{\alpha i}'-\xi_{\alpha i}^2)}{\xi_i^3},$$

$$\xi_i=e^{-\lambda l_i^{\alpha}}-e^{-\lambda r_i^{\alpha}},$$

$$\xi_{\alpha i}=\frac{\partial}{\partial\alpha}\xi_i=-\lambda l_i^{\alpha}e^{-\lambda l_i^{\alpha}}\ln l_i+\lambda r_i^{\alpha}e^{-\lambda r_i^{\alpha}}\ln r_i,$$

$$\xi_{\alpha i}'=\frac{\partial}{\partial\alpha}\xi_{\alpha i}=-\lambda(\ln l_i)^2(l_i^{\alpha}-\lambda l_i^{2\alpha})e^{-\lambda l_i^{\alpha}}+\lambda(\ln r_i)^2(r_i^{\alpha}-\lambda r_i^{2\alpha})e^{-\lambda r_i^{\alpha}},$$

$$\xi_{\alpha i}''=\frac{\partial}{\partial\alpha}\xi_{\alpha i}'=-\lambda(\ln l_i)^3(l_i^{\alpha}-3\lambda l_i^{2\alpha}+\lambda^2 l_i^{3\alpha})e^{-\lambda l_i^{\alpha}}+\lambda(\ln r_i)^3(r_i^{\alpha}-3\lambda r_i^{2\alpha}+\lambda^2 r_i^{3\alpha})e^{-\lambda r_i^{\alpha}},$$

$$\xi_{\alpha\lambda i} = \frac{\partial}{\partial\lambda}\xi_{\alpha i} = -l_i^{\alpha}\ln l_i e^{-\lambda l_i^{\alpha}}(1-\lambda l_i^{\alpha}) + r_i^{\alpha}\ln r_i e^{-\lambda r_i^{\alpha}}(1-\lambda r_i^{\alpha}),$$

$$\xi_{\alpha\lambda i}^{'} = \frac{\partial}{\partial\lambda}\xi_{\alpha\lambda i} = -l_i^{\alpha}\ln l_i e^{-\lambda l_i^{\alpha}}(\lambda l_i^{2\alpha} - 2l_i^{\alpha}) + r_i^{\alpha}\ln r_i e^{-\lambda r_i^{\alpha}}(\lambda r_i^{2\alpha} - 2r_i^{\alpha}),$$

$$\xi_{\alpha\lambda i}^{''} = \frac{\partial}{\partial\lambda}\xi_{\alpha i}' = -(\ln l_i)^2[l_i^{\alpha} - 3\lambda l_i^{2\alpha} + \lambda^2 l_i^{3\alpha}]e^{-\lambda l_i^{\alpha}} + (\ln r_i)^2[r_i^{\alpha} - 3\lambda r_i^{2\alpha} + \lambda^2 r_i^{3\alpha}]e^{-\lambda r_i^{\alpha}},$$

$$\xi_{\lambda i} = \frac{\partial}{\partial\lambda}\xi_i = -l_i^{\alpha}e^{-\lambda l_i^{\alpha}} + r_i^{\alpha}e^{-\lambda r_i^{\alpha}},$$

$$\xi_{\lambda i}^{'} = \frac{\partial}{\partial\lambda}\xi_{\lambda i} = l_i^{2\alpha}e^{-\lambda l_i^{\alpha}} - r_i^{2\alpha}e^{-\lambda r_i^{\alpha}},$$

$$\xi_{\lambda i}^{''} = \frac{\partial}{\partial\lambda}\xi_{\lambda i}^{'} = -l_i^{3\alpha}e^{-\lambda l_i^{\alpha}} + r_i^{3\alpha}e^{-\lambda r_i^{\alpha}}.$$

Biswabrata Pradhan
SQC & OR Unit
Indian Statistical Institute
203 B.T. Road
Kolkata 700108, India

Debasis Kundu
Department of Mathematics and Statistics
Indian Institute of Technology
Kanpur 208016, India
E-mail: kundu@iitk.ac.in

Paper received: 26 October 2012; revised: 23 October 2013.