# Chapter 2

# Statistical Inference in Inverse Pareto Lifetime Model using Randomly Censored Data*

## 2.1 Introduction

The main objective of this chapter is to build classical and Bayesian inferences about the model parameters of the IP lifetime model using randomly censored data.

In the survival analysis, the entire lifetime of a person or an animal is not always observable. Some lifetimes may be censored, in that case, only a part of the lifetime is recorded. Therefore, censoring is a necessary part of life testing experiments. The units in these experiments are lost or removed, resulting in incomplete information. In the literature, there are different types of censoring schemes. Type-I and Type-II censoring schemes are the most extensively used censoring techniques in reliability and life testing experiments. The censoring time or the number of censored items are prefixed in these censoring techniques. Many scholars, such as Mann et al. (1974) and Sinha (1986), have examined these censoring techniques with various lifetime models extensively.

Random censoring is a common occurrence in real-world life testing experiments. For example, patients with leukemia enter into the study simultaneously after their treatments. We aim to track them throughout their lives, but censoring can take many forms, including loss to

*Part of this chapter has been published in the form of a research paper with the following details: Kumar, K., and Kumar, I. (2020). Parameter Estimation for Inverse Pareto Distribution with Randomly Censored Life Time Data. International Journal of Agricultural Statistical Sciences. 16 (1): 419-430.

follow-up (e.g., the patient may elect to relocate), drop out (e.g., inadequate side effects or an unfinished course of treatment), death from other conditions, or study layoff. That is, these random features are uncontrollable by the treatments, resulting in an independent random variable called a censoring time variable. This censoring scheme was introduced by Gilbert (1962) in literature. After that some early study on random censoring can be found in Breslow and Crowley (1974), Koziol and Green (1976), etc. Recently, several authors investigated the usefulness of random censoring in literature for different lifetime models like, Ghitany and Al-Awadhi (2002) discussed ML estimates of parameters for Burr Type XII distribution, the generalized inverted Rayleigh distribution is studied by Kumar and Garg (2014), Krishna et al. (2015) studied Maxwell distribution, the generalized inverted exponential distribution is studied by Garg et al. (2016), Krishna and Goel (2017) discussed geometric distribution, the log-logistic distribution is discussed by Kumar (2018), the Birnbaum-Saunders distribution is discussed by El-Sharkawy and Ismail (2020), EL-Sagheer et al. (2020) studied three parameters Burr Type XII distribution etc.

Mathematically, random censoring can be described as follows: suppose the failure times $X_1, X_2, \ldots, X_n$ are independent and identically distributed (iid) random variables with pdf $f_X, x > 0$ and survival function $S_X, x > 0$. Associated with these failure times, $T_1, T_2, \ldots, T_n$ are iid censoring times with pdf $f_T, t > 0$ and survival function $S_T, t > 0$. Now, suppose $X_i's$ and $T_i's$ mutually independent $\forall\ i = 1, 2, \ldots, n$. We observe failure or censored time $Y_i = \min(X_i's, T_i's)$ ; $i = 1, 2, \ldots, n$, and the corresponding censor indicators

$$D_i = \begin{cases} 1; \text{ if failure occurs} \\ 0; \text{ if censoring occurs.} \end{cases}$$

Some spacial cases of this censoring scheme are as follows: (i) It become complete sample case when $T_i = \infty\ \forall\ i = 1, 2, \ldots, n$. (i) It reduces to Type I censoring when $T_i = t_0\ \forall i = 1, 2, \ldots, n$, where, $t_0$ is the pre-fixed study period. Thus, the joint pdf of $Y$ and $D$ is given by

$$f_{Y,D}(y,d) = \{f_X(y)S_T(y)\}^d \{f_T(y)S_X(y)\}^{1-d}\ ;\ y > 0, d = 0, 1. \tag{2.1}$$

The marginal distribution of $Y$ and $D$ can be obtained as

$$f_Y(y) = f_X(y)S_T(y) + f_T(y)S_X(y),\ \ y > 0, \text{ and}$$

$$P[D = d] = p^d(1-p)^{1-d}\ \ ; d = 0, 1,$$

respectively, where, $p$ is the probability of observing a failure and it is given by

$$p = P[X \leq T] = \int_0^\infty S_T(y) f_X(y) dy.$$

There are numerous real-life situations in survival analysis where data requires a probability distribution with both decreasing and upside-down bathtub-shaped failure rate functions. For example, a disease's mortality may reach a high after a while and then gradually drop, as shown in Kundu and Howlader (2010). During the first few days after a heart transplant, while the body adjusts to the new organ, patients face an increasing failure rate of mortality. As the patient recovers, the failure rate reduces, as seen in Collett (2015). The failure function shaped like an upside-down bathtub would be acceptable in such cases.

The one parameter IP lifetime model has both the decreasing and upside-down bathtub-shaped failure rate functions depending on the true value of the parameter. Also, it has nice closed-form expressions of the cumulative distribution function (cdf) and failure rate function, both of which are useful in reliability theory or survival analysis. However, the IP lifetime model has a very nice closed form failure rate function, it has not gained much attention in the literature. Guo and Gui (2018) studied IP lifetime model based on stress-strength reliability in the case of both classical and Bayesian approaches. The application of IP lifetime model in extreme events is studied by Dankunprasert et al. (2021), Kumar et al. (2021) developed some estimation methods for associated parameter and reliability characteristics of IP lifetime model.

The main aim of this chapter is to develop the classical and Bayesian estimation procedures for the parameters of the IP lifetime model using randomly censored data. The rest of the chapter is laid out as follows: the IP lifetime model is discussed in Section 2.2. Also, a mathematical formulation is given for random censoring with failure and censoring time distributions. Section 2.3 deals with the ML estimation and ACIs of the parameters. Section 2.4 describes the formulation of Bayes estimation procedure using MCMC methods under LINEX loss function using gamma informative priors. The HPD credible intervals for the parameters are derived using MCMC techniques. Section 2.5 deals with an MC simulation study to explore the properties of various estimates developed in this chapter. Two real datasets are analyzed for illustration purposes in Section 2.6. Finally, concluding remarks are given in Section 2.7.

## 2.2 The Model

If a random variable $X$ follows the IP lifetime model with parameter $\alpha$ denoted by IP($\alpha$), the pdf of IP lifetime model is given by

$$f_X(x;\alpha) = \frac{\alpha x^{\alpha-1}}{(1+x)^{\alpha+1}} \quad ; \ \alpha > 0, \ x > 0. \tag{2.2}$$

Figure 2.1 shows the pdf of IP lifetime model for distinct values of $\alpha$, say 0.25, 0.75, 1.5 and

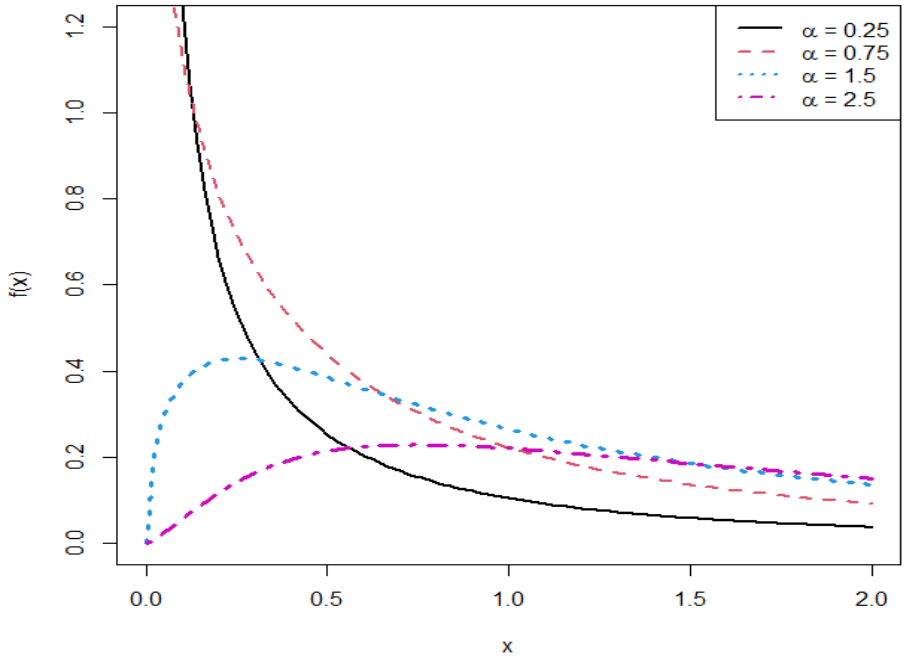


FIGURE 2.1: Plot of pdf of IPD.

2.5. Also, the corresponding cdf, survival and failure rate functions are, respectively, given by

$$F_X(x;\alpha) = \left(\frac{x}{1+x}\right)^{\alpha} \quad ; \ \alpha > 0, \ x > 0, \tag{2.3}$$

$$S(x;\alpha) = P(X > x) = 1 - \left(\frac{x}{1+x}\right)^{\alpha} \quad ; \ x > 0, \ \alpha > 0, \ \text{and} \tag{2.4}$$

$$h(x;\alpha) = \frac{\alpha x^{\alpha-1}}{(1+x)^{\alpha+1}\left[1-\left(\frac{x}{1+x}\right)^{\alpha}\right]} \quad ; \ \alpha > 0, x > 0. \tag{2.5}$$

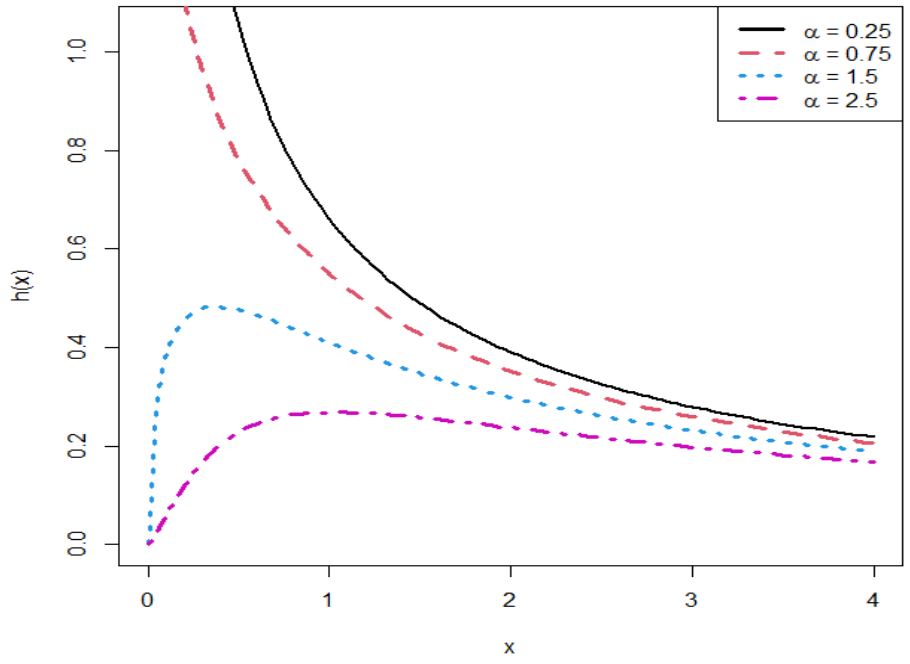


FIGURE 2.2: Plot of failure rate function of IPD.

Figure 2.2 shows the failure rate function of IP lifetime model for distinct values of $\alpha$, say 0.25, 0.75, 1.5 and 2.5. From the figure 2.2, it is clear that IP lifetime model holds both decreasing and upside-down bathtub shaped failure rate functions.

Next, suppose the failure time $X$ folow IP lifetime model with parameter $\alpha$, say IP($\alpha$), and censoring time $T$ follows IP lifetime model with parameter $\beta$, say IP($\beta$). Then using equation (2.1), the joint pdf of randomly censored IP lifetime model is given by

$$f_{Y,D}(y,d,\alpha,\beta) = \frac{\alpha^d \beta^{1-d} y^{d(\alpha-\beta)+\beta-1}}{(1+y)^{d(\alpha-\beta)+\beta+1}} \left[1-\left(\frac{y}{1+y}\right)^{\beta}\right]^d \left[1-\left(\frac{y}{1+y}\right)^{\alpha}\right]^{1-d}; \quad y>0, \alpha>0, \beta>0, d=0,1, \tag{2.6}$$

and the probability of observing a failure is given by

$$p = \int_0^{\infty} S_T(y) f_X(y) dy = \frac{\beta}{\alpha+\beta}.$$

## 2.3 Maximum Likelihood Estimation

In this section, we derive the ML estimates, $\hat{\alpha}$ and $\hat{\beta}$ of $\alpha$ and $\beta$, respectively. For the observed sample $(\mathbf{y},\ \mathbf{d}) = (y_1,\ d_1),(y_2,\ d_2),\ldots(y_n,\ d_n)$ of size $n$. Also, compute ACIs of the parameters based on observed Fisher information matrix. The likelihood function can be written as

$$L(\mathbf{y},\ \mathbf{d},\ \alpha,\ \beta) = \prod_{i=1}^{n} \frac{\alpha^{d_i}\beta^{1-d_i} y_i^{d_i(\alpha-\beta)+\beta-1}}{(1+y_i)^{d_i(\alpha-\beta)+\beta+1}} \left[1-\left(\frac{y_i}{1+y_i}\right)^{\beta}\right]^{d_i} \left[1-\left(\frac{y_i}{1+y_i}\right)^{\alpha}\right]^{1-d_i}. \quad (2.7)$$

Thus, the log-likelihood function becomes

$$\begin{aligned} l(\alpha,\beta|\text{data}) &= m \ln \alpha + (n-m)\ln\beta + (\alpha-\beta)\sum_{i=1}^{n} d_i \ln y_i + (\beta-1)\sum_{i=1}^{n} \ln y_i \\ &\quad - (\alpha-\beta)\sum_{i=1}^{n} d_i \ln(1+y_i) - (\beta+1)\sum_{i=1}^{n} \ln(1+y_i) \\ &\quad + \sum_{i=1}^{n} d_i \ln\left[1-\left(\frac{y_i}{1+y_i}\right)^{\beta}\right] + \sum_{i=1}^{n}(1-d_i)\ln\left[1-\left(\frac{y_i}{1+y_i}\right)^{\alpha}\right], \end{aligned} \quad (2.8)$$

where, $m = \sum_{i=1}^{n} d_i$. The corresponding normal equations of the log-likelihood function obtain as follows:

$$\frac{\partial\, l(\alpha,\beta|\text{data})}{\partial\alpha} = \frac{m}{\alpha} + \sum_{i=1}^{n} d_i \ln y_i - \sum_{i=1}^{n} d_i \ln(1+y_i) - \sum_{i=1}^{n}(1-d_i)\frac{\left(\frac{y_i}{1+y_i}\right)^{\alpha}\ln\left(\frac{y_i}{1+y_i}\right)}{\left[1-\left(\frac{y_i}{1+y_i}\right)^{\alpha}\right]} = 0 \quad (2.9)$$

$$\begin{aligned} \frac{\partial\, l(\alpha,\beta|\text{data})}{\partial\beta} &= \frac{n-m}{\beta} + \sum_{i=1}^{n} \ln y_i - \sum_{i=1}^{n} d_i \ln y_i + \sum_{i=1}^{n} d_i \ln(1+y_i) - \sum_{i=1}^{n} \ln(1+y_i) - \\ &\quad \sum_{i=1}^{n} d_i \frac{\left(\frac{y_i}{1+y_i}\right)^{\beta}\ln\left(\frac{y_i}{1+y_i}\right)}{\left[1-\left(\frac{y_i}{1+y_i}\right)^{\beta}\right]} = 0 \end{aligned} \quad (2.10)$$

The ML estimates $\hat{\alpha}$ and $\hat{\beta}$ of the parameters $\alpha$ and $\beta$, respectively, are the solutions of the non-linear equations (2.9) and (2.10). Here, equations (2.9) and (2.10) do not have closed form solutions, any iterative method can be used to solve these equations for $\alpha$ and $\beta$, respectively. Here, for the computation purpose, *nlm* or *optim* or *maxLik* functions of statistical software *R* can be used.

### 2.3.1 Asymptotic Confidence Intervals

As the ML estimates of the unknown model parameters are not in closed form, driving the exact distributions of the ML estimates is difficult. As a result, we build the ACIs of the parameters

based on the observed Fisher information matrix using the asymptotic distribution of ML estimates. Let $\hat{\theta} = (\hat{\alpha},\ \hat{\beta})$, be the MLE of $\theta = (\alpha,\ \beta)$, the observed Fisher information matrix is given by

$$I(\hat{\theta}) = \left[\begin{array}{cc} -\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\alpha^2} & -\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\alpha\partial\beta} \\ -\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\beta\partial\alpha} & -\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\beta^2} \end{array}\right]_{\theta=\hat{\theta}}$$

where,

$$\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\alpha^2} = -\frac{m}{\alpha^2} - \sum_{i=1}^{n}(1-d_i)\frac{\left(\frac{y_i}{1+y_i}\right)^{\alpha}\left(\ln\left(\frac{y_i}{1+y_i}\right)\right)^2}{\left[1-\left(\frac{y_i}{1+y_i}\right)^{\alpha}\right]^2},$$

$$\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\beta^2} = -\frac{n-m}{\beta^2} - \sum_{i=1}^{n} d_i\frac{\left(\frac{y_i}{1+y_i}\right)^{\beta}\left(\ln\left(\frac{y_i}{1+y_i}\right)\right)^2}{\left[1-\left(\frac{y_i}{1+y_i}\right)^{\beta}\right]^2},$$

$$\frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\alpha\partial\beta} = \frac{\partial^2 l(\alpha,\beta|\text{data})}{\partial\beta\partial\alpha} = 0.$$

The asymptotic distribution of ML estimates $\hat{\theta}$ follows a bivariate normal distribution i.e. $\hat{\theta} \sim N(\theta, I^{-1}(\hat{\theta}))$, see, Lawless (2003). Consequently, two sided equal tailed $100(1-\xi)\%$ ACIs of parameters $\alpha$ and $\beta$ are given by

$$\left(\hat{\alpha} \pm z_{\xi/2}\sqrt{\hat{Var}(\hat{\alpha})}\right) \text{ and } \left(\hat{\beta} \pm z_{\xi/2}\sqrt{\hat{Var}(\hat{\beta})}\right),$$

respectively. Here, $\hat{Var}(\hat{\alpha})$ and $\hat{Var}(\hat{\beta})$ are diagonal elements of the observed Fisher information matrix $I^{-1}(\hat{\theta})$ and $z_{\xi/2}$ is the upper $(\xi/2)^{th}$ percentile of the standard normal distribution N(0,1). Also, the coverage probability (CPs) for the parameters are given by

$$CP_{\alpha} = \left[\left|\frac{\hat{\alpha}-\alpha}{\sqrt{\hat{Var}(\hat{\alpha})}}\right| \le z_{\xi/2}\right] \text{ and } CP_{\beta} = \left[\left|\frac{\hat{\beta}-\beta}{\sqrt{\hat{Var}(\hat{\beta})}}\right| \le z_{\xi/2}\right].$$

## 2.4 Bayesian Estimation

Here, we discussed the Bayes estimators of unknown parameters associated with the model in (2.6) under the LINEX loss function. In decision theory, a suitable loss function must be given in order to get the optimal decision. For this purpose, the squared error loss function (SELF) is commonly employed loss function in the literature. This loss function is appropriate, when overestimation and underestimation of equal magnitude have the same effects. When the true loss is not symmetric in terms of overestimation and underestimation, asymmetric

loss functions are employed to characterise the implications of various losses. Varian (1975) introduced an asymmetric loss function for the first time known as LINEX loss function and it is given as follows:

$$L(\phi,\hat{\phi}) = e^{k(\hat{\phi}-\phi)} - k(\hat{\phi}-\phi) - 1, \tag{2.11}$$

where, $\hat{\phi}$ is an estimate of parameter $\phi$, $k \neq 0$ is the known loss parameter. The sign and magnitude of the loss parameter $k$ reflects the direction and degree of asymmetry, respectively. When $k$ is positive, the over estimation is more serious than under estimation and the situation is reverse when $k$ is negative. The LINEX loss function reduces to SELF when magnitude of $k$ tends to zero, see, Zellner (1986). Under the LINEX loss function, the Bayes estimate of $\phi$ is given as follows

$$\hat{\phi}_{Bayes} = -\frac{1}{k}\ln E[e^{-k\phi}|\text{data}],$$

where, $E[e^{-k\phi}|\text{data}]$ is the posterior expectation which exist and finite. Further, we assume the prior belief of the unknown parameters $\alpha$ and $\beta$ follows gamma distributions with the following pdfs:

$$g_1(\alpha) = \frac{b_1^{a_1}}{\Gamma(a_1)}\alpha^{a_1-1}e^{-b_1\alpha}\ ;\alpha,\ a_1,\ b_1 > 0,$$
$$g_2(\beta) = \frac{b_2^{a_2}}{\Gamma(a_2)}\beta^{a_2-1}e^{-b_2\beta}\ ;\beta,\ a_2,\ b_2 > 0,\ \text{respectively.}$$

Thus, the joint prior distribution of $\alpha$ and $\beta$ can be written as

$$g(\alpha,\beta) \propto \alpha^{a_1-1}e^{-b_1\alpha}\beta^{a_2-1}e^{-b_2\beta},\ a_1,\ b_1, a_2,\ b_2 > 0. \tag{2.12}$$

The assumption of the piece-wise independent gamma priors is quite reasonable. It is noted that the non-informative priors are the special cases of independent gamma priors when hyper-parameters $a_1 = b_1 = a_2 = b_2 = 0$. Based on the observed randomly censored data, likelihood function in (2.7) and joint prior distribution of ($\alpha$, $\beta$) in (2.12), the joint posterior distribution of $\alpha$ and $\beta$ is given by

$$\pi(\alpha,\beta|\text{data}) = \frac{L(\text{data}|\alpha,\beta)g(\alpha,\beta)}{\int_0^\infty\int_0^\infty L(\text{data}|\alpha,\beta)g(\alpha,\beta)d\alpha d\beta}$$

$$\pi(\alpha,\beta|\text{data}) \propto \alpha^{m+a_1-1} e^{-\alpha\left[b_1-\sum_{i=1}^{n} d_i \ln\left(\frac{y_i}{1+y_i}\right)\right]} \prod_{i=1}^{n}\left[1-\left(\frac{y_i}{1+y_i}\right)^{\alpha}\right]^{1-d_i} \times \beta^{n-m+a_2-1} e^{-\beta\left[b_2-\sum_{i=1}^{n}(1-d_i)\ln\left(\frac{y_i}{1+y_i}\right)\right]} \prod_{i=1}^{n}\left[1-\left(\frac{y_i}{1+y_i}\right)^{\beta}\right]^{d_i} \tag{2.13}$$

From the joint posterior distribution of $\alpha$ and $\beta$ given in equation (2.13), we observe that the posterior distributions of $\alpha$ and $\beta$ are independent. Thus the marginal posterior distribution of $\alpha$ given data **(y, d)** is obtained as

$$\pi_1(\alpha \mid \text{data}) \propto \alpha^{m+a_1-1} e^{-\alpha\left[b_1-\sum_{i=1}^{n} d_i \ln\left(\frac{y_i}{1+y_i}\right)\right]} \prod_{i=1}^{n}\left[1-\left(\frac{y_i}{1+y_i}\right)^{\alpha}\right]^{(1-d_i)} ;\ \alpha>0 \tag{2.14}$$

Similarly, the marginal posterior distribution of $\beta$ given data $(\mathbf{y},\ \mathbf{d})$ is obtained as

$$\pi_2(\beta \mid \text{data}) \propto \beta^{n-m+a_2-1} e^{-\beta\left[b_2-\sum_{i=1}^{n}(1-d_i)\ln\left(\frac{y_i}{1+y_i}\right)\right]} \prod_{i=1}^{n}\left[1-\left(\frac{y_i}{1+y_i}\right)^{\beta}\right]^{d_i} ;\beta>0 \tag{2.15}$$

Thus, the expectations of any function of $\alpha$ say $\phi_1(\alpha)$ and $\beta$ say $\phi_2(\beta)$, respectively, are given by

$$E[\phi_1(\alpha) \mid \text{data}] = \int_0^{\infty} \phi_1(\alpha)\pi_1(\alpha \mid \text{data})d\alpha \tag{2.16}$$

$$\text{and} \quad E[\phi_2(\beta) \mid \text{data}] = \int_0^{\infty} \phi_2(\beta)\pi_1(\beta \mid \text{data})d\beta. \tag{2.17}$$

From the above equation (2.16) and (2.17), we observe that the closed form solutions are not available. The above integrals can be solved numerically. Here, we use Markov Chain Monte Carlo (MCMC) techniques like, the Metropolis-Hastings (M-H) algorithm to derive the Bayes estimates of the parameters $\alpha$ and $\beta$, respectively.

### 2.4.1 MCMC Technique

Here, we use MCMC techniques to generate sequences of samples from the marginal posterior distributions of the parameters. The M-H algorithm is used to obtain sample based Bayes estimates of the unknown parameters. For more details about MCMC and M-H algorithm techniques, one may refer, Gelman et al. (2013), Robert and Casella (2004), Metropolis et al.

(1953), Hastings (1970). The marginal posterior distributions of the parameters $\alpha$ and $\beta$ in equations (2.14) and (2.15), respectively, are not well known distributions. Therefore random numbers from these distributions can be generated by using M-H algorithm. The following steps are used to generate random numbers from the marginal posterior distribution in (2.14):

Step 1: Begin with an initial guess. $\alpha^{(0)}$.
Step 2: From the proposed density $\delta(\alpha^{(j)} \mid \alpha^{(j-1)})$, create a candidate point $\alpha_c^{(j)}$.
Step 3: Generate $u$ using the Uniform (0,1) distribution.
Step 4: Obtain $z(\alpha_c^{(j)} \mid \alpha^{(j-1)}) = \min\left\{\frac{\pi_1(\alpha_c^{(j)}|\text{data})\delta(\alpha^{(j)}|\alpha^{(j-1)})}{\pi_1(\alpha^{(j-1)}|\text{data})\delta(\alpha_c^{(j)}|\alpha^{(j-1)})}, 1\right\}$
Step 5: If $u \leq z$ set $\alpha^{(j)} = \alpha_c^{(j)}$ with acceptance probability $z$ otherwise $\alpha^{(j)} = \alpha^{(j-1)}$.
Step 6: To acquire the parameter sequence of $\alpha$ as $\left(\alpha^{(1)}, \alpha^{(2)}, \ldots, \alpha^{(M)}\right)$, repeat steps 2-5 for $j = 1, 2, \ldots, M,$.

Here, we consider proposal density as a normal distribution. The ML estimates and variance of ML estimates from posterior distribution of $\alpha$ are considered as mean and variance of the proposal normal distribution, see, (Ntzoufras, 2009, pp. 44-45). To get an independent sample from the stationary distribution of the Markov chain, which is generally the posterior distribution, we discard first $M_0$, $\alpha^{(j)}$'s ; $j = 1, 2, \ldots, M_0$, where, $M_0$ $(< M)$ is the burn-in-period. Now, the approximate posterior mean of $\phi_1(\alpha)$ using M-H algorithm is obtained as

$$\hat{\phi}_{1MH}(\alpha) = \frac{1}{M - M_0} \sum_{j=M_0+1}^{M} \phi_1(\alpha^{(j)}).$$

Similarly, the approximate posterior mean of $\phi_2(\beta)$ using M-H algorithm is obtained as

$$\hat{\phi}_{2MH}(\beta) = \frac{1}{M - M_0} \sum_{j=M_0+1}^{M} \phi_2(\beta^{(j)}).$$

Therefore, the Bayes estimates of the parameters $\alpha$ and $\beta$ under LINEX loss function using M-H algorithm are, respectively, given by

$$\hat{\alpha}_{MH} = -\frac{1}{k}\ln(\hat{\phi}_{1MH}(\alpha)) \text{ and } \hat{\beta}_{MH} = -\frac{1}{k}\ln(\hat{\phi}_{2MH}(\beta)).$$

### 2.4.2 HPD Credible Intervals

Here, we compute the HPD credible intervals of the parameters $\alpha$ and $\beta$ using the generated MCMC samples. Let $\alpha_{(1)} < \alpha_{(2)} < \cdots < \alpha_{(M-M_0)}$ denote the ordered values of $\alpha^{(M_0+1)}$, $\alpha^{(M_0+2)}, \ldots, \alpha^{(M)}$. Then, using the algorithm proposed by Chen and Shao (1999), the $100(1-\xi)\%$, where $0 < \xi < 1$, HPD credible interval for $\alpha$ is given by $(\alpha_{(j)}, \alpha_{(j+[(1-\xi)(M-M_0)])})$,

where $j$ is chosen such that

$$\alpha_{(j+[(1-\xi)(M-M_0)])} - \alpha_{(j)} = \min_{1\leq i\leq (M-M_0)} \left(\alpha_{(i+[(1-\xi)(M-M_0)])} - \alpha_{(i)}\right); j = 1,2,\ldots,M-M_0,$$

where, $[x]$ is the largest integer less than or equal to x. Similarly, we can construct the $100(1-\xi)\%$ HPD credible interval for $\beta$.

## 2.5 Numerical Computations

Here, we perform a MC simulation study to examine the different estimators created in the preceding sections. The simulation study considers six distinct sample sizes $n = 30, 40, 50, 60, 70, 80$ for different combinations of true parameters ($\alpha$, $\beta$) = (0.75, 1.5) and (1.5, 0.75), respectively. The unknown parameter $\alpha$ and $\beta$ are estimated using ML and Bayes estimation methods in each cases. The hyper-parameters $(a_1, b_1, a_2, b_2)$ = (3, 2, 3, 4) and (3, 4, 3, 2) are taken into account for gamma informative priors (Prior 1) in Bayesian calculations such that the prior means precisely the same as the true values of the parameters. In case of non-informative priors (Prior 0), the hyper-parameters are taken as $a_1 = b_1 = a_2 = b_2 = 0.0001$. Two distinct values of loss parameter $k$ = -1 and 1 are taken for LINEX loss function. For MCMC technique, $M = 10,000$ sequence of parameter samples are drawn from posterior distribution and $M_0 = 1,000$ taken as burn-in-period. The 95% ACIs based on OFI matrix and HPD credible intervals based on MCMC technique are computed. The entire procedure is replicated by 1000 times. The average estimates (AE) and their associated mean squared error (MSE) are estimated for various estimators. Also determined the average length (AL) and coverage probabilities (CP) of 95% ACI and HPD credible intervals. Tables 2.1, 2.2, and 2.3 describe the findings of the MC simulation study.

These observations can lead to the following conclusions: In almost every case, as sample size grows, AEs become closer to the real value of the parameters, while MSEs go lower. Similarly, when the sample size grows, the ALs of interval estimates shrink, demonstrating the estimators' asymptotic behaviour. CPs achieve the required levels of confidence in almost every case. Bayes estimators perform more effectively in the case of Prior 1 than Prior 0 or ML estimators in terms of biases. On average, HPD credible intervals are shorter AL than ACIs. When some prior information about parameters is provided or non-informative priors are used, we suggest Bayes estimators. ML estimators can also be utilised for rapid results in other situations.

TABLE 2.1: ML and Bayes estimates of $\alpha$, for different values of $\alpha$.

| | | | | | $\hat{\alpha}_{\text{Bayes}}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $k=-1$ | | | | $k=1$ | | | |
| | | | $\hat{\alpha}_{\text{MLE}}$ | | Prior 0 | | Prior 1 | | Prior 0 | | Prior 1 | |
| $\alpha$ | $n$ | $m$ | AE | MSE | AE | MSE | AE | MSE | AE | MSE | AE | MSE |
| 0.75 | 30 | 21 | 0.7758 | 0.0233 | 0.7876 | 0.0256 | 0.7859 | 0.0195 | 0.7712 | 0.0217 | 0.7662 | 0.0170 |
| | 40 | 24 | 0.7780 | 0.0170 | 0.7863 | 0.0183 | 0.7692 | 0.0143 | 0.7686 | 0.0171 | 0.7590 | 0.0139 |
| | 50 | 35 | 0.7679 | 0.0134 | 0.7745 | 0.0142 | 0.7679 | 0.0113 | 0.7598 | 0.0122 | 0.7554 | 0.0105 |
| | 60 | 42 | 0.7681 | 0.0112 | 0.7736 | 0.0118 | 0.7686 | 0.0096 | 0.7608 | 0.0105 | 0.7586 | 0.0096 |
| | 70 | 43 | 0.7666 | 0.0095 | 0.7712 | 0.0099 | 0.7656 | 0.0093 | 0.7556 | 0.0088 | 0.7525 | 0.0073 |
| | 80 | 57 | 0.7593 | 0.0079 | 0.7632 | 0.0081 | 0.7611 | 0.0075 | 0.7590 | 0.0076 | 0.7578 | 0.0069 |
| 1.5 | 30 | 12 | 1.5911 | 0.1553 | 1.6783 | 0.2422 | 1.6187 | 0.1337 | 1.5483 | 0.1320 | 1.5240 | 0.0842 |
| | 40 | 16 | 1.5688 | 0.0971 | 1.6268 | 0.1288 | 1.5951 | 0.0947 | 1.5343 | 0.0869 | 1.5220 | 0.0687 |
| | 50 | 13 | 1.5510 | 0.0762 | 1.5949 | 0.0942 | 1.5977 | 0.0743 | 1.5146 | 0.0635 | 1.5217 | 0.0539 |
| | 60 | 18 | 1.5479 | 0.0655 | 1.5838 | 0.0783 | 1.5622 | 0.0586 | 1.5233 | 0.0555 | 1.5170 | 0.0468 |
| | 70 | 24 | 1.5432 | 0.0548 | 1.5730 | 0.0637 | 1.5785 | 0.0585 | 1.5168 | 0.0442 | 1.5092 | 0.0394 |
| | 80 | 23 | 1.5217 | 0.0409 | 1.5467 | 0.0460 | 1.5540 | 0.0444 | 1.5226 | 0.0408 | 1.5170 | 0.0421 |

TABLE 2.2: ML and Bayes estimates of $\beta$, for different values of $\beta$.

| | | | | | $\hat{\beta}_{\text{Bayes}}$ | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | $k=-1$ | | | | $k=1$ | | | |
| | | | $\hat{\beta}_{\text{MLE}}$ | | Prior 0 | | Prior 1 | | Prior 0 | | Prior 1 | |
| $\beta$ | $n$ | $m$ | AE | MSE | AE | MSE | AE | MSE | AE | MSE | AE | MSE |
| 1.5 | 30 | 21 | 1.5976 | 0.1560 | 1.6876 | 0.2703 | 1.6361 | 0.1497 | 1.5349 | 0.1200 | 1.5407 | 0.0883 |
| | 40 | 24 | 1.5686 | 0.0993 | 1.6266 | 0.1308 | 1.5989 | 0.0930 | 1.5389 | 0.0887 | 1.5172 | 0.0665 |
| | 50 | 35 | 1.5429 | 0.0709 | 1.5862 | 0.0870 | 1.5848 | 0.0721 | 1.5272 | 0.0726 | 1.5086 | 0.0510 |
| | 60 | 42 | 1.5369 | 0.0620 | 1.5721 | 0.0735 | 1.5747 | 0.0611 | 1.5246 | 0.0544 | 1.5076 | 0.0485 |
| | 70 | 43 | 1.5391 | 0.0490 | 1.5688 | 0.0573 | 1.5556 | 0.0496 | 1.5187 | 0.0512 | 1.5215 | 0.0422 |
| | 80 | 57 | 1.5294 | 0.0422 | 1.5547 | 0.0479 | 1.5532 | 0.0443 | 1.5196 | 0.0399 | 1.5238 | 0.0388 |
| 0.75 | 30 | 12 | 0.7762 | 0.0210 | 0.7881 | 0.0231 | 0.7853 | 0.0198 | 0.7666 | 0.0206 | 0.7664 | 0.0182 |
| | 40 | 16 | 0.7661 | 0.0157 | 0.7744 | 0.0167 | 0.7845 | 0.0170 | 0.7653 | 0.0155 | 0.7657 | 0.0146 |
| | 50 | 13 | 0.7687 | 0.0140 | 0.7755 | 0.0148 | 0.7650 | 0.0117 | 0.7627 | 0.0132 | 0.7558 | 0.0097 |
| | 60 | 18 | 0.7651 | 0.0105 | 0.7706 | 0.0110 | 0.7722 | 0.0103 | 0.7575 | 0.0103 | 0.7584 | 0.0101 |
| | 70 | 24 | 0.7607 | 0.0087 | 0.7651 | 0.0091 | 0.7627 | 0.0079 | 0.7594 | 0.0089 | 0.7556 | 0.0079 |
| | 80 | 23 | 0.7564 | 0.0077 | 0.7604 | 0.0079 | 0.7687 | 0.0076 | 0.7577 | 0.0080 | 0.7602 | 0.0070 |

TABLE 2.3: The AL and CPs of 95% ACIs and HPD credible intervals for the different values parameters ($\alpha$, $\beta$).

| | | | $\hat{\alpha}_{\text{MLE}}$ | | $\hat{\alpha}_{\text{Bayes}}$ Prior 0 | | $\hat{\alpha}_{\text{Bayes}}$ Prior 1 | | $\hat{\beta}_{\text{MLE}}$ | | $\hat{\beta}_{Bayes}$ Prior 0 | | $\hat{\beta}_{Bayes}$ Prior 1 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $(\alpha,\beta)$ | $n$ | $m$ | AL | CP | AL | CP | AL | CP | AL | CP | AL | CP | AL | CP |
| (0.75,1.5) | 30 | 21 | 0.5638 | 0.954 | 0.5521 | 0.943 | 0.5256 | 0.955 | 1.3605 | 0.955 | 1.5308 | 0.976 | 1.3958 | 0.978 |
| | 40 | 24 | 0.4895 | 0.957 | 0.4790 | 0.950 | 0.4527 | 0.956 | 1.1435 | 0.958 | 1.2874 | 0.979 | 1.2102 | 0.983 |
| | 50 | 35 | 0.4319 | 0.953 | 0.4233 | 0.948 | 0.4078 | 0.956 | 1.0008 | 0.948 | 1.1279 | 0.970 | 1.0864 | 0.983 |
| | 60 | 42 | 0.3944 | 0.951 | 0.3862 | 0.942 | 0.3745 | 0.958 | 0.9093 | 0.955 | 1.0225 | 0.972 | 0.9904 | 0.968 |
| | 70 | 43 | 0.3643 | 0.949 | 0.3573 | 0.943 | 0.3468 | 0.937 | 0.8419 | 0.956 | 0.9482 | 0.972 | 0.9135 | 0.976 |
| | 80 | 57 | 0.3374 | 0.953 | 0.3308 | 0.942 | 0.3237 | 0.956 | 0.7817 | 0.952 | 0.8801 | 0.970 | 0.8588 | 0.971 |
| (1.5,0.75) | 30 | 12 | 1.3515 | 0.954 | 1.3268 | 0.948 | 1.2024 | 0.962 | 0.5643 | 0.967 | 0.6316 | 0.987 | 0.6000 | 0.982 |
| | 40 | 16 | 1.1462 | 0.961 | 1.1269 | 0.951 | 1.0514 | 0.947 | 0.4818 | 0.955 | 0.5397 | 0.971 | 0.5267 | 0.973 |
| | 50 | 13 | 1.0084 | 0.941 | 0.9914 | 0.933 | 0.9586 | 0.968 | 0.4324 | 0.940 | 0.4837 | 0.964 | 0.4629 | 0.973 |
| | 60 | 18 | 0.9176 | 0.956 | 0.9014 | 0.943 | 0.8603 | 0.964 | 0.3928 | 0.947 | 0.4397 | 0.969 | 0.4300 | 0.975 |
| | 70 | 24 | 0.8452 | 0.956 | 0.8323 | 0.951 | 0.8155 | 0.945 | 0.3614 | 0.949 | 0.4040 | 0.974 | 0.3942 | 0.973 |
| | 80 | 23 | 0.7773 | 0.95 | 0.7656 | 0.944 | 0.7489 | 0.962 | 0.3361 | 0.956 | 0.3762 | 0.974 | 0.3734 | 0.973 |

## 2.6 Real Data Analysis

With the help of two real datasets, we illustrate the estimation procedures discussed in the previous sections. Here, we consider two real datasets, namely leukemia patients' data (Data I) and Hodgkin's disease patients' data (Data II). These datasets are reported in (Lawless, 2003, pp. 139). Data I depicts the remission periods (in weeks) of a group of 30 leukemia patients who all got the same therapy. Data II considered the survival times (in months) of 15 patients with Hodgkin's disease who were treated with nitrogen mustards and received heavy prior therapy. Data I and Data II, respectively, are given below:

**Data I:** 1, 1, 2, 4, 4, 6, 6, 6, 7, 8, 9, 9, 10, 12, 13, 14, 18, 19, 24, 26, 29, 31+, 42, 45+, 50+, 57, 60, 71+, 85+, 91.
**Data II:** 1.05, 2.92, 3.61, 4.20, 4.49, 6.72, 7.31, 9.08, 9.11, 14.49+, 16.85, 18.82+, 26.59+, 30.26+, 41.34+.
The observations with + sign are censored times.

Before going further, we fit Data I and Data II to randomly censored IP lifetime and compare its fitting with some well-known lifetime models, namely, inverse exponential (IE) and generalized inverted exponential (GIE) lifetime models in case of random censoring. The pdfs of the competitive lifetime models are as follows:

$$\text{IE: } f(x,\theta) = \frac{\theta}{x^2} e^{-\theta/x} \quad x>0, \theta>0,$$
$$\text{GIE: } f(x,\alpha,\theta) = \frac{\alpha\theta}{x^2} e^{-\theta/x}\left(1-e^{-\theta/x}\right)^{\alpha-1} \quad x>0, \alpha, \theta>0.$$

We compute ML estimates of the unknown parameters along with some useful measure of goodness-of-fit tests and model comparison criteria for both datasets, namely, the negative log-likelihood -ln$L$, the AIC defined by AIC $= 2\times k - 2\times \ln L$, proposed by Akaike (1974) and Bayesian information criterion (BIC) defined by $BIC = k\times\ln(n) - 2\times\ln L$, proposed by Schwarz (1978), where $k$ is the number of associated parameters in the model, $n$ is the number of data points in the given datasets, $L$ is the maximised value of the likelihood function for the estimated model and the Kolmogorov-Smirnov (KS) statistics with its $p$-values. The best lifetime model corresponds to the lowest –ln$L$, AIC, BIC, and KS statistic and the highest $p$-value. The KS statistic with its $p$-values are obtained using *ks.test* function in statistical software R, see R Core Team (2021). The results of the ML estimates and measures of goodness-of-fit tests are reported in Table 2.4 and 2.5, respectively. From these results, we observed that the performance of the randomly censored IP lifetime model is the best choice for the considered datasets.

Moreover, ML and Bayes estimates with their corresponding 95% ACIs and HPD credible intervals of the unknown parameters associated with randomly censored IP lifetime model corresponding to the above real datasets (Data I and Data II) are computed and reported in Table 2.6. The Bayes estimates are computed using non-informative priors under SELF with the help of the MCMC technique. For the M-H algorithm, we generate Markov chain $M = 10,000$ from the posterior distribution. We also examine the convergence of their stationary distributions using graphical diagnostic tools such as trace and histogram plots with Gaussian kernel density plots, as shown in Figures 2.3 and 2.4. The trace plots indicate a random scatter and show the fine mixing of the chains. The histogram plots of the generated MCMC samples show that the marginal posterior distributions of the parameters are almost symmetrical i.e. we can take the mean as the best estimate for the parameters. These plots are hallmarks of rapid MCMC convergence. From these results, we see that ML and Bayes estimates of parameters based on MCMC techniques are quite closed.

TABLE 2.4: Summary fit of the leukemia patients data (Data I).

| Models | MLE | $-\ln L$ | AIC | BIC | KS-Test | |
|---|---|---|---|---|---|---|
| | | | | | KS-Statistic | $p$-value |
| $X \sim \text{IE}(\alpha)$<br>$T \sim \text{IE}(\beta)$ | $\hat{\alpha} = 6.4343$<br>$\hat{\beta} = 76.3664$ | 139.8547 | 283.7094 | 286.5118 | 0.1755 | 0.3137 |
| $X \sim \text{IP}(\alpha)$<br>$T \sim \text{IP}(\beta)$ | $\hat{\alpha} = 7.863$<br>$\hat{\beta} = 77.3696$ | 137.7025 | 279.4049 | 282.2073 | 0.1371 | 0.6256 |
| $X \sim \text{GIE}(\alpha, \beta)$<br>$T \sim \text{GIE}(\alpha, \lambda)$ | $\hat{\alpha} = 0.6619$<br>$\hat{\beta} = 4.7952$<br>$\hat{\lambda} = 63.1718$ | 138.2794 | 282.5587 | 286.7623 | 0.1599 | 0.4271 |

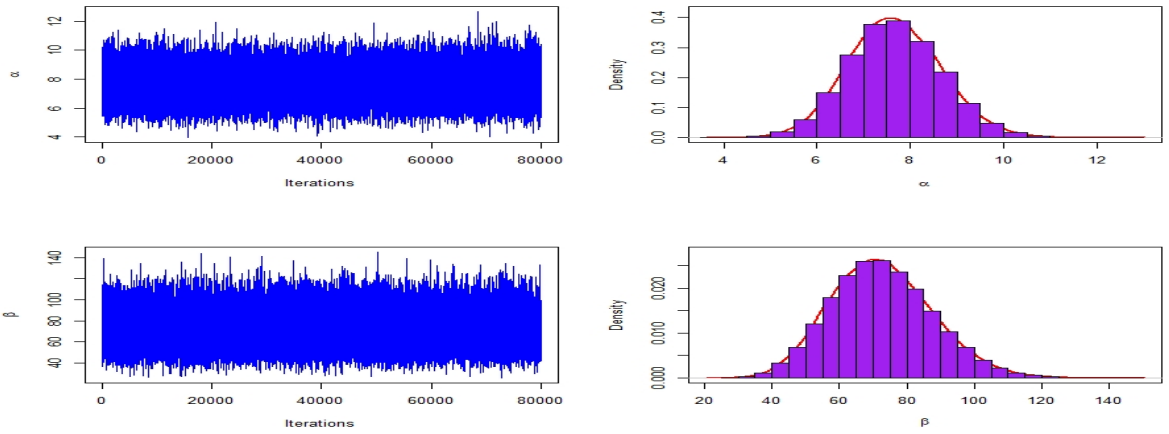


FIGURE 2.3: MCMC plot of leukemia patients data (Data I)

TABLE 2.5: Summary fit of the Hodgkin's disease patient data (Data II).

| Models | MLE | $-\ln L$ | AIC | BIC | KS-Test | |
|---|---|---|---|---|---|---|
| | | | | | KS-Statistic | $p$-value |
| $X \sim \text{IE}(\alpha)$<br>$T \sim \text{IE}(\beta)$ | $\hat{\alpha} = 5.5533$<br>$\hat{\beta} = 27.2313$ | 60.0676 | 124.1352 | 125.5513 | 0.1303 | 0.9323 |
| $X \sim \text{IP}(\alpha)$<br>$T \sim \text{IP}(\beta)$ | $\hat{\alpha} = 6.6481$<br>$\hat{\beta} = 28.1105$ | 59.7491 | 123.4982 | 124.9143 | 0.0966 | 0.9965 |
| $X \sim \text{GIE}(\alpha, \beta)$<br>$T \sim \text{GIE}(\alpha, \lambda)$ | $\hat{\alpha} = 0.9172$<br>$\hat{\beta} = 5.2465$<br>$\hat{\lambda} = 26.1409$ | 60.0400 | 126.0800 | 128.2041 | 0.1157 | 0.9740 |

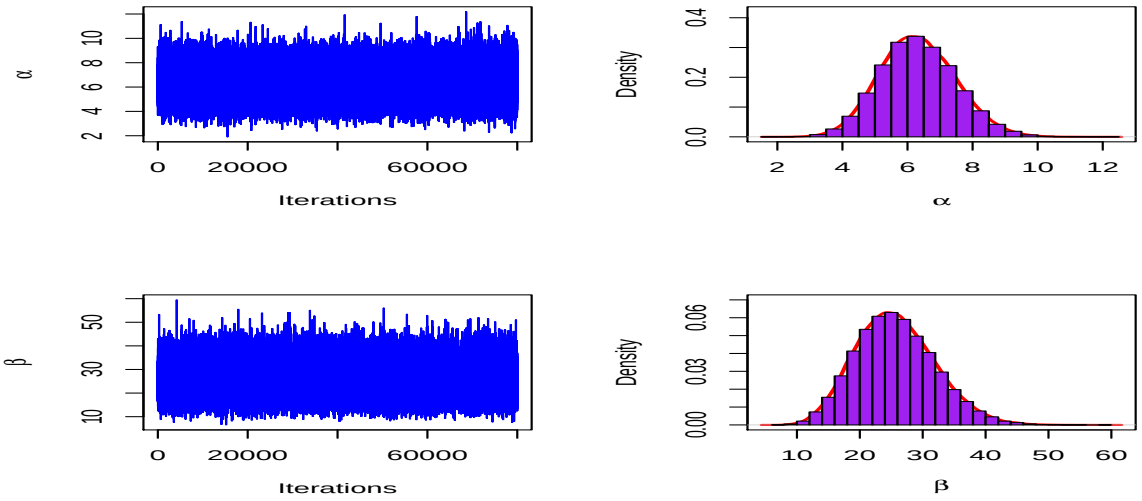


FIGURE 2.4: MCMC plot of Hodgkin's disease data (Data II)

TABLE 2.6: The ML, Bayes estimates and 95% asymptotic and HPD credible intervals of the unknown parameters corresponding to Data I and Data II, respectively.

| Datasets | Parameters | MLE | 95% CI | Bayes estimates | 95% HPD CI |
|---|---|---|---|---|---|
| Data I | $\alpha$ | 7.863 | (5.0484, 10.6775) | 7.6672 | (5.7135, 9.598) |
| | $\beta$ | 77.3696 | (31.6325, 123.1067) | 72.6036 | (40.328, 106.8078) |
| Data II | $\alpha$ | 6.6481 | (3.2792, 10.017) | 6.3191 | (4.0975, 8.6322) |
| | $\beta$ | 28.1105 | (8.3706, 47.8505) | 25.6675 | (12.6852, 40.3103) |

## 2.7 Concluding Remarks

The classical and Bayesian estimation techniques for the parameters of the IP lifetime model using randomly censored data were discussed in this chapter. The ML estimators and their corresponding ACIs based on the observed Fisher information matrix of the unknown parameters were derived. MCMC methods were used to approximate Bayes estimates of the parameters under the LINEX loss function. A comprehensive Monte Carlo simulation study was conducted to evaluate the performance of different estimators, and the results show that ML estimates may be employed easily with acceptable results. For more efficient estimators, the Bayesian estimation method with available prior information or convenient non-informative priors in the absence of prior information is appropriate and recommended.