Bayesian Analysis of Progressively Censored Competing Risks Data

Debasis Kundu [*]& Biswabrata Pradhan [†]

Abstract

In this paper we consider the Bayesian inference of the unknown parameters of the progressively censored competing risks data, when the lifetime distributions are Weibull. It is assumed that the latent cause of failures have independent Weibull distributions with the common shape parameter, but different scale parameters. In this article, it is assumed that the shape parameter has a log-concave prior density function, and for the given shape parameter, the scale parameters have Beta-Dirichlet priors. When the common shape parameter is known, the Bayes estimates of the scale parameters have closed form expressions, but when the common shape parameter is unknown, the Bayes estimates do not have explicit expressions. In this case we propose to use MCMC samples to compute the Bayes estimates and highest posterior density (HPD) credible intervals. Monte Carlo simulations are performed to investigate the performances of the estimators. Two data sets are analyzed for illustration. Finally we provide a methodology to compare two different censoring schemes and thus find the optimum Bayesian censoring scheme.

Key Words and Phrases: Latent failure time model; Type-II progressive censoring scheme; Markov Chain Monte Carlo; Credible interval; Optimum censoring scheme.

[*]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India, and a visiting Professor at King Saud University, Saudi Arabia, Corresponding author, e-mail: kundu@iitk.ac.in

[†]SQC & OR Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata, Pin 700108, India

1 Introduction

In many life-testing studies, often the failure of items or individuals may be associated to more than one cause. These 'risk factors' in some sense compete with each other for the failure of the experimental unit. Due to this reason, in the statistical literature it is well known as the competing risks model. Several examples can be found, see for example Crowder [10], where failure may occur due to more than one cause. In analyzing such data set, the investigator is naturally interested in the assessment of a specific risk in presence of other risk factors.

In analyzing the data for competing risks model, ideally the data consists of a failure time and the associated cause of failure. The causes of failure may be assumed to be independent or dependent. Although the assumption of dependence seems more reasonable, but there is some concern about the identifiability issue of the competing risks model. Several authors, see for example Kalbfleish and Prentice [13], Crowder [10], argued that without the information of covariates, it is not possible to test the assumption of the failure time distributions of the competing causes, just based on the observed data.

In this paper we use the latent failure time modeling of Cox [9] for analyzing competing risks data. In the latent failure time modeling, it is assumed that the competing causes of failures are independently distributed. Here it is further assumed that the lifetime distributions of the competing causes follow Weibull distributions with the same shape parameter, but different scale parameters. It may be mentioned that the assumption of the common shape parameter for the Weibull distribution in case of competing risks of model is not very unrealistic, see for example Rao *et al.* [26], Mukherjee and Basu [19], Kundu and Basu [16] and the references therein.

Therefore, if T_i denotes the lifetime of the i-th individual then

$$T_i = \min\{X_{i1}, \cdots, X_{iM}\},$$

where $X_{i1}, \cdots, X_{iM}$ are the latent failure times of the M different causes for the i-th individual. According to the latent failure time model assumption, $X_{i1}, \cdots, X_{iM}$ are independently distributed. Moreover, $X_{i1}, \cdots, X_{iM}$ are not observable, only T_i is observable and the indicator J such that $X_{iJ} = \min\{X_{i1}, \cdots, X_{iM}\}$ is observable. In this paper it is further assumed that X_{ij} for $j = 1, \cdots, M$, follows a Weibull distribution with the probability density function (PDF)

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

here $\alpha > 0$, $\lambda_j > 0$ are the shape and scale parameters of the Weibull distribution with the PDF (1) and it will be denoted as WE(α, λ_j). In this paper, from now on it is assumed that $M = 2$ for notational convenience, although all the results presented here are valid for general M.

Censoring is very common in most of the life-testing and reliability studies, because quite often the experimenter is unable to obtain complete information on lifetimes of all the items/individuals. Although, Type-I and Type-II censoring schemes are two most popular censoring schemes, but recently progressive censoring scheme has received considerable attention in the statistical literature due to its wide scale applicability, see for example Viveros and Balakrishnan, [27], Balasooriya *et al.* [6], Ng *et al.* [21], Kundu [15], Pradhan and Kundu [22], the review article by Balakrishnan [3] and the references cited therein. Although, extensive work has been done on progressive censoring scheme, not much work has been done in the competing risk set up.

In this paper we consider the Bayesian analysis of the competing risks data, when the lifetime distributions are Weibull with the same shape but different scale parameters. For the Bayesian inference of the unknown parameters, we need to assume some priors on the unknown parameters. If the common shape parameter α is known, the convenient but quite general conjugate priors on the scale parameters are the Beta-Gamma (for $M > 2$, it is Dirichlet-Gamma) priors, see Pena and Gupta [25] in this respect. In this case, the explicit expressions of the Bayes estimates can be obtained under the squared error loss function. But when the common shape parameter is unknown, it is known that in this case the continuous conjugate priors do not exist, see for example Kaminskiy and Krivtsov [14] in this connection. We use the same conjugate priors on the scale parameters, even when α was unknown. We have not assumed any specific prior on α, it is simply assumed that the support of the prior on α is on $(0, \infty)$, and that it has a log-concave density function. Note that the assumption of log-concave density function of the prior distribution is quite common in Bayesian analysis, see Berger and Sun [5], and many common distribution functions, for example normal, log-normal, gamma and Weibull may have log-concave density functions.

Based on the above prior distributions, we obtain the joint posterior distribution of the unknown parameters. As expected the Bayes estimates cannot be obtained in explicit forms. We propose to use MCMC samples to compute Bayes estimates and approximate highest posterior density (HPD) credible intervals of the unknown parameters. One can also apply Lindley's approximation to compute Bayes estimates. It is observed that our method can be easily extended even when some of the causes of failures are unknown. We compare the performances of Bayes estimates and HPD credible intervals with the classical maximum likelihood estimators (MLEs). It is observed that if we do not have any prior information, the performances of the MLEs and the Bayes estimators are quite comparable. But with informative priors, the Bayes estimates behave much better than the MLEs, as expected. Although we have derived the Bayes estimates of the unknown parameters based on progressively censored data, the proposed method is easily extendable for other censoring schemes. In survival analysis, data are mainly random censored. We have analyzed such a data set in Section 4.2 to illustrate our methodology.

The second aim of this manuscript is to provide the methodology to compare two different sampling schemes, and hence in turn to compute the optimal censoring scheme in presence of competing risks. Finding the optimal progressive censoring scheme is an important problem, and it has received considerable attention in the recent statistical literature due to its practical applicability. In the progressive censoring scheme, an optimal censoring scheme means, for fixed n and m, the choice of $\{R_1, \cdots, R_m\}$, which provides the maximum information of the unknown parameters. Unfortunately, not much work has been done in this direction in presence of competing risks.

Using the idea of Zhang and Meeker [28], we have proposed two information measure of the unknown parameters for a given progressive censoring scheme, when the competing causes of failure are present. We have provided the optimal censoring schemes based on different criteria and compared the results with the traditional Type-II censoring also. It is observed that the relative efficiencies of the Type-II censoring schemes are quite close to one.

The rest of the paper is organized as follows. In section 2 we provide the model formulation and prior assumptions. Posterior analysis and Bayesian inference of the unknown parameters are provided in section 3. Numerical simulation results and the analysis of two data sets are presented in section 4. In section 5 we provide the optimal censoring plan, and finally we conclude the paper in section 6.

2 Model Formulation and Prior Assumptions

In this section we introduce the model, and the available data. We also provide the necessary prior assumptions for further development.

2.1 Model Formulation and Available Data

Suppose n identical items are put on a test at the time point zero, with the lifetime of the n-items are denoted by $T_1, \cdots, T_n$. It is assumed that for each i, $T_i = \min\{X_{i1}, X_{i2}\}$, where $X_{i1} \sim \text{WE}(\alpha, \lambda_1)$, $X_{i2} \sim \text{WE}(\alpha, \lambda_2)$ and they are independently distributed. Therefore, $T_i \sim \text{WE}(\alpha, \lambda_1 + \lambda_2)$, and moreover it is assumed that $T_1, \cdots, T_n$ are independently distributed. At the time of each failure, the failure time and the corresponding cause of failure is observed.

The integer $m < n$ is pre-fixed, and $R_1, \cdots, R_m$ are pre-fixed integers such that

$$R_1 + \cdots + R_m = n - m. \tag{2}$$

At the time of the first failure, say t_1, R_1 of the remaining units are randomly chosen and removed. Similarly, at the time of the second failure, say t_2, R_2 of the remaining units are chosen and removed, and so on. Finally at the time of the m-th failure time, t_m, the rest of the units, $R_m = n - m - R_1 - \cdots - R_{m-1}$ are removed and the experiment stops. Therefore, a progressively censored competing risk data will be as follows:

$$\{(t_1, \delta_1, R_1), \cdots, (t_m, \delta_m, R_m)\}. \tag{3}$$

Here $\delta_1, \cdots, \delta_m$ denote the m causes of failures at the time points $t_1, \cdots, t_m$ respectively, and for each i, δ_i takes a value either 1 and 2. Therefore, for a given $R_1, \cdots, R_m$, we have the following m observations;

$$\{(t_i, 1); i \in I_1\}, \quad \text{and} \quad \{(t_i, 2); i \in I_2\}, \tag{4}$$

here

$$I_1 = \{i; \delta_i = 1\}, \quad \text{and} \quad I_2 = \{i; \delta_i = 2\}.$$

2.2 Prior Assumptions

When the common shape parameter α is known, the scale parameters have conjugate priors. Using the idea of Pena and Gupta [25], it is assumed that $\lambda = \lambda_1 + \lambda_2$ has a Gamma(a_0, b_0) prior, say $\pi_0(\cdot|\ a_0, b_0\cdot)$. Here the PDF of Gamma$(a_0, b_0)$ for $\lambda > 0$ is;

$$\pi_0(\lambda|\ a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \lambda^{a_0-1} e^{-b_0\lambda}, \tag{5}$$

and 0 otherwise. Given λ, λ_1/λ has Beta(a_1, a_2), say $\pi_1(\cdot|a_1, a_2)$ prior, *i.e.*

$$\pi(\lambda_1/\lambda|\ a_1, a_2) = \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\Gamma(a_2)} \left(\frac{\lambda_1}{\lambda}\right)^{a_1-1} \left(1 - \frac{\lambda_1}{\lambda}\right)^{a_2-1} \tag{6}$$

for $\lambda_1/\lambda > 0$, and 0 otherwise. Here all the hyper-parameters $a_0 > 0, b_0 > 0, a_1 > 0, a_2 > 0$. It will be shown that when the common shape parameter is known, the above priors are the conjugate priors. After simple transformation, the joint prior of λ_1 and λ_2 becomes;

$$\pi(\lambda_1, \lambda_2|a_0, b_0, a_1, a_2) = \frac{\Gamma(a_1 + a_2)}{\Gamma(a_0)} \times (b_0\lambda)^{a_0-a_1-a_2} \times \frac{b_0^{a_1}}{\Gamma(a_1)} \lambda_1^{a_1-1} e^{-b_0\lambda_1} \times \frac{b_0^{a_2}}{\Gamma(a_2)} \lambda_2^{a_2-1} e^{-b_0\lambda_2}. \tag{7}$$

This is the Beta-Dirichlet distribution, and it will be denoted by BD(b_0, a_0, a_1, a_2). Clearly, in general λ_1 and λ_2 will be dependent, but when $a_0 = a_1 + a_2$, they are independent. Therefore, independent priors can be obtained as a special case of (7). It may be easily observed that the covariance of λ_1 and λ_2 can be positive or negative depending on $a_0 > a_1 + a_2$ or $a_0 < a_1 + a_2$. The following result will be useful for further development, whose proof can be easily obtained from Theorem 2 of Pena and Gupta [25].

Result: If $(\lambda_1, \lambda_2) \sim$ BD(b_0, a_0, a_1, a_2), then for $i = 1, 2$,

$$E(\lambda_i) = \frac{a_0 a_i}{b_0(a_1 + a_2)} \quad \text{and} \quad V(\lambda_i) = \frac{a_0 a_i}{b_0^2(a_1 + a_2)} \times \left\{ \frac{(a_i + 1)(a_0 + 1)}{a_1 + a_2 + 1} - \frac{a_0 a_i}{a_1 + a_2} \right\}. \tag{8}$$

When the common shape parameter is known, the above priors are the conjugate priors. But when the shape parameter is not known, the conjugate priors do not exist. In this case it is assumed that λ_1 and λ_2 have the same Beta-Dirichlet prior as defined (7) and the prior on α is independent of (λ_1, λ_2). No specific form of prior on α has been assumed here. It is only assumed that the absolute continuous prior $\pi(\alpha)$ on α has a positive support on $(0, \infty)$ and the PDF of $\pi(\alpha)$ is log-concave and it is independent of (λ_1, λ_2). Although, in general the choice of the hyper-parameters are very important in practice, it is not pursued here.

3 Posterior Analysis and Bayes Inference

In this section, we provide the Bayes estimates of the unknown parameters and the corresponding credible intervals, when the common shape parameter is known and when it is unknown based on the priors assumed in the previous section. We mainly assume the squared error loss (SEL) function, although any other loss function can be easily considered, without much of a difficulty.

3.1 Common Shape Parameter Known

Based on the observed sample $\{(t_1, \delta_1, R_1), \cdots, (t_m, \delta_m, R_m)\}$, the likelihood function is;

$$l(data|\ \alpha, \lambda_1, \lambda_2) \propto \alpha^m \lambda_1^{m_1} \lambda_2^{m_2} e^{-(\lambda_1+\lambda_2)\sum_{i=1}^{m}(R_i+1)t_i^{\alpha}} \times \prod_{i=1}^{m} t_i^{\alpha-1}. \tag{9}$$

Here m_1 and m_2 denote the number of elements in the set I_1 and I_2 respectively. For known α, when λ_1 and λ_2 have the joint priors as given in section 2, it can be easily observed that the joint posterior distribution of λ_1 and λ_2, *i.e.*

$$l(\lambda_1, \lambda_2|data, \alpha) \propto \text{BD}\left(b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha}, a_0 + m_1 + m_2, a_1 + m_1, a_2 + m_2\right). \tag{10}$$

Therefore, with respect to the squared error loss function, the Bayes estimates of λ_1 and λ_2 are

$$\widehat{\lambda}_{1B} = \frac{(a_0 + m_1 + m_2)(a_1 + m_1)}{(b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha})(a_1 + a_2 + m_1 + m_2)} \tag{11}$$

and

$$\widehat{\lambda}_{2B} = \frac{(a_0 + m_1 + m_2)(a_2 + m_2)}{(b_0 + \sum_{i=1}^{m}(R_i + 1)t_i^{\alpha})(a_1 + a_2 + m_1 + m_2)}. \tag{12}$$

The corresponding posterior variances are

$$V(\widehat{\lambda}_{1B}) = A_1 \times B_1, \quad \text{and} \quad V(\widehat{\lambda}_{2B}) = A_2 \times B_2, \tag{13}$$

respectively, where for $j = 1, 2$,

$$A_j = \frac{(a_0 + m_1 + m_2)(a_j + m_j)}{(b_0 + \sum_{i=1}^{m}(R_i + 1)t_i^{\alpha})^2(a_1 + a_2 + m_1 + m_2)} \tag{14}$$

$$B_j = \left\{ \frac{(a_j + m_j + 1)(a_0 + m_1 + m_2 + 1)}{(a_1 + a_2 + m_1 + m_2 + 1)} - \frac{(a_j + m_j)(a_0 + m_1 + m_2)}{(a_1 + a_2 + m_1 + m_2)} \right\}. \tag{15}$$

Under the assumptions of non-informative priors, *i.e.* $a_0 = b_0 = a_1 = a_2 = 0$, the Bayes estimates of λ_1 and λ_2 become

$$\widehat{\lambda}_{1B} = \frac{m_1}{\sum_{i=1}^{m}(R_i + 1)t_i^{\alpha}}, \quad \text{and} \quad \widehat{\lambda}_{2B} = \frac{m_2}{\sum_{i=1}^{m}(R_i + 1)t_i^{\alpha}}, \tag{16}$$

and they can be easily seen to be the MLEs of λ_1 and λ_2 respectively.

Note that although the Bayes estimates can be obtained in explicit forms, the corresponding highest posterior density (HPD) credible intervals cannot be obtained explicitly. But it is possible to generate MCMC samples by direct sampling from the joint posterior density function, and they can be used to construct HPD credible intervals of λ_1 and λ_2. The details will be explained later.

3.2 Common Shape Parameter Unknown

In this subsection we consider the important case when the common shape parameter is unknown, which is most likely to happen in practice. In this case based on the priors on λ_1, λ_2 and α, as it has been assumed in the previous section, the joint density function of on λ_1,

λ_2, α and *data* is

$$l(data|\ \alpha, \lambda_1, \lambda_2)\pi_1(\lambda_1, \lambda_2|\ a_0, b_0, a_1, b_1)\pi_2(\alpha). \tag{17}$$

Based on (7), the joint posterior density of α, λ_1 and λ_2 is

$$l(\alpha, \lambda_1, \lambda_2|\ data) = \frac{l(data|\ \alpha, \lambda_1, \lambda_2)\pi_1(\lambda_1, \lambda_2|\ a_0, b_0, a_1, b_1)\pi_2(\alpha)}{\int_0^\infty \int_0^\infty \int_0^\infty l(data|\ \alpha, \lambda_1, \lambda_2)\pi_1(\lambda_1, \lambda_2|\ a_0, b_0, a_1, b_1)\pi_2(\alpha) d\alpha d\lambda_1 d\lambda_2}. \tag{18}$$

Therefore, the Bayes estimate of any function of α, λ_1 and λ_2, say $g(\alpha, \lambda_1, \lambda_2)$ under the squared error loss function can be obtained as

$$\begin{aligned} \widehat{g}_B(\alpha, \lambda_1, \lambda_2) &= E_{\alpha,\lambda_1,\lambda_2|data}(g(\alpha, \lambda_1, \lambda_2)) \\ &= \frac{\int_0^\infty \int_0^\infty \int_0^\infty g(\alpha, \lambda_1, \lambda_2) l(data|\ \alpha, \lambda_1, \lambda_2)\pi_1(\lambda_1, \lambda_2|\ a_0, b_0, a_1, b_1)\pi_2(\alpha) d\alpha d\lambda_1 d\lambda_2}{\int_0^\infty \int_0^\infty \int_0^\infty l(data|\ \alpha, \lambda_1, \lambda_2)\pi_1(\lambda_1, \lambda_2|\ a_0, b_0, a_1, b_1)\pi_2(\alpha) d\alpha d\lambda_1 d\lambda_2}. \end{aligned} \tag{19}$$

Note that in most of the situation, it is not possible to compute (19) explicitly. There are several methods available to approximate (19), but they do not provide the HPD credible intervals. We propose to use MCMC samples generated by direct sampling method to approximate (19) and also to compute the HPD credible intervals of the unknown parameters. The following results will be used for generating samples.

Theorem 1: The joint conditional distribution of λ_1 and λ_2 given α is

$$\text{BD}\left(b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha}, a_0 + m_1 + m_2, a_1 + m_1, a_2 + m_2\right). \tag{20}$$

Proof: It is trivial and it has already been mentioned in the previous subsection.

Theorem 2: The conditional density of α given the *data* is log-concave.

Proof: Observe that

$$l(\alpha|\ data) \propto \pi_2(\alpha)\alpha^m \prod_{i=1}^{m} t_i^{\alpha-1} \times \frac{1}{(b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha})^{a_0+m_1+m_2}}. \tag{21}$$

Now the result follows using Theorem 2 of Kundu [15].

Devroye [11] has suggested to generate samples from a general log-concave density function. Therefore, using Theorem 1 and Theorem 2, it is possible to generate MCMC samples from the joint posterior density function (18). We use the following algorithm:

Algorithm:

Step 1: Generate α from (21) using the method suggested by Devroye [11].

Step 2: For given α, generate (λ_1, λ_2) from (20), as suggested in Appendix A.

Step 3: Repeat Step 1 and Step 2, M times, and obtain $(\alpha_1, \lambda_{11}, \lambda_{21}), \cdots, (\alpha_M, \lambda_{1M}, \lambda_{2M})$.

Step 4: The Bayes estimates α, λ_1 and λ_2 with respect to SEL function respectively as

$$\widehat{\alpha}_B = \frac{1}{M}\sum_{k=1}^{M}\alpha_k, \quad \widehat{\lambda}_{1B} = \frac{1}{M}\sum_{k=1}^{M}\lambda_{1k} \quad \text{and} \quad \widehat{\lambda}_{2B} = \frac{1}{M}\sum_{k=1}^{M}\lambda_{2k}.$$

Step 5: The corresponding posterior variances can be obtained respectively as

$$\frac{1}{M}\sum_{k=1}^{M}\left(\alpha_k - \widehat{\alpha}_B\right)^2, \quad \frac{1}{M}\sum_{k=1}^{M}\left(\lambda_{1k} - \widehat{\lambda}_{1B}\right)^2 \quad \text{and} \quad \frac{1}{M}\sum_{k=1}^{M}\left(\lambda_{2k} - \widehat{\lambda}_{2B}\right)^2.$$

Step 6: To obtain credible interval of α, we order $\{\alpha_i\}$, as $\alpha_{(1)} < \alpha_{(2)} < \cdots < \alpha_{(M)}$. Then 100(1-2$\beta$)% credible intervals of α become

$$(\alpha_{(j)}, \alpha_{(j+M-2M\beta)}), \quad \text{for} \;\; j = 1, \cdots, 2M\beta.$$

Therefore, 100(1-2β)% HPD credible interval of α becomes $(\alpha_{(j^*)}, \alpha_{(j^*+M-2M\beta)})$, where j^* is such that

$$\alpha_{(j^*+M-2M\beta)} - \alpha_{(j^*)} \leq \alpha_{(j+M-2M\beta)} - \alpha_{(j)}$$

for all $j = 1, \cdots, 2M\beta$. Similarly, we obtain the HPD credible intervals for λ_1 and λ_2.

4 Numerical Results and Data Analysis

In this section we conduct a simulation study to investigate the performance of the proposed Bayes estimators and analyze a data set for illustration. The simulation study is carried out when both the shape and scale parameters are unknown.

4.1 Simulation study

We use different sample sizes n, different effective sample sizes m and the following sets of parameters $\{\alpha = 1, \lambda_1 = 0.6, \lambda_2 = 0.4\}$ and $\{\alpha = 2, \lambda_1 = 0.6, \lambda_2 = 0.4\}$. For the simulation study, it is assumed that α has gamma prior with the shape parameter a and scale parameter b. We consider non-informative and informative priors both for the shape and scale parameters. In case of non-informative prior we take $a = b = a_0 = b_0 = a_1 = a_2 = 0$. We call it as prior 0. The informative priors for the two sets of parameter values are given in Table 1. It may be noted that prior 1 and prior 3 are selected in such way that prior means are same as the original means. Hence prior 1 and prior 3 may be considered as the calibrated priors. Whereas prior 2 and prior 4 are not calibrated and they have been selected arbitrarily. Performance of the estimators are studied in terms of their biases and the mean squared errors (MSEs). We also compute the average lengths and the coverage percentages of the 95% HPD credible intervals based on MCMC samples generated by direct sampling.

Table 1: Different informative priors for the two sets of parameters

Parameters	Prior
$\alpha = 1, \lambda_1 = 0.6, \lambda_2 = 0.4$	Prior 1: a=b=5, a_0=b_0=2, $a_1 = 0.6$ and $a_2 = 0.4$
	Prior 2: a= 2.5, b=1, a_0= 1, b_0=1.5, $a_1 = 0.8$ and $a_2 = 0.9$
$\alpha = 2, \lambda_1 = 0.6, \lambda_2 = 0.4$	Prior 3: a= 5, b=2.5, a_0=b_0=2, $a_1 = 0.6$ and $a_2 = 0.4$
	Prior 4: a= 5, b=1.5, a_0= 2.5 b_0= 1, $a_1 = 1.5$ and $a_2 = 0.6$

For both the parameter sets, we considered the following sampling schemes:

- Scheme 1 (CS-1): $n = 15$, $m = 12$, $R_1 = \cdots = R_{11} = 0, R_{12} = 3$;
- Scheme 2 (CS-2): $n = 30$, $m = 15$, $R_1 = \cdots = R_{14} = 0, R_{15} = 15$;
- Scheme 3 (CS-3): $n = 40$, $m = 15$, $R_1 = \cdots = R_{14} = 0, R_{15} = 25$;
- Scheme 4 (CS-4): $n = 40$, $m = 30$, $R_1 = \cdots = R_{29} = 0, R_{30} = 10$;
- Scheme 5 (CS-5): $n = 40$, $m = 30$, $R_2 = \cdots = R_{30} = 0, R_1 = 10$;
- Scheme 6 (CS-6): $n = 40$, $m = 30$, $R_2 = \cdots = R_{29} = 0, R_1 = R_{30} = 5$;
- Scheme 7 (CS-7): $n = 40$, $m = 30$, $R_1 = \cdots = R_{14} = R_{17} = \cdots = R_{30} = 0, R_{15} = R_{16} = 5$.

Note that the sampling schemes CS-1, CS-2, CS-3 and CS-4 are the usual Type-II censoring schemes, that is $n - m$ items are removed at the time of the m-th failure. The sampling scheme CS-5 is just the opposite of the Type-II censoring scheme and in this case $n - m$ items are removed at the time of the first failure itself. This particular progressive censoring scheme is known as first step-censoring and this is a particular case of one-step censoring introduced by Balakrishnan *et al.* [4]. It is well known that for fixed n and m, the expected experimental time of the Type-II censoring scheme is less than the corresponding first step-censoring scheme. In fact for fixed n and m, the expected experimental time of any other censoring schemes are always between these two extremes. For example, the expected experimental time of the sampling scheme CS-6 and CS-7 will be between the schemes CS-4 and CS-5.

For generating the progressively censored Weibull samples, we use the algorithm suggested-in Balakrishnan and Sandhu [1]. For each data point we assigned the cause of failure as 1 or 2 with probability $\lambda_1/(\lambda_1 + \lambda_2)$ and $\lambda_2/(\lambda_1 + \lambda_2)$, respectively. In each case we calculate Bayes estimates using 10000 MCMC samples and also the 95% HPD credible intervals. We also compute maximum likelihood estimate (MLE) and 95% confidence interval for the purpose of comparison with the Bayes estimates. We replicate the process for 1000 times and compute average estimates, mean squared errors, average length of HPD credible intervals/confidence interval and coverage percentage. The results are reported in Tables 2-5.

Now we compare the performance of the Bayes estimators under different scenarios and MLEs in terms of biases, MSEs and length of credible intervals. From the simulation results, it can be seen that the performance of the Bayes estimates with respect to non-informative priors and MLEs are quite close, as expected. So, if we have no prior information on the unknown parameters, it is better to use MLEs than Bayes estimators, because the Bayes estimators are computationally more expensive. The performance of the Bayes estimators under informative calibrated priors (prior 1 and prior 3) is better than the MLEs or the Bayes estimators under non-informative prior and informative uncalibrated priors (prior 2 and prior 4), as expected.

4.2 Data Analysis

In this section, we analyze two data sets for illustrative purpose.

Example 1: We consider the competing risks data set of Nelson [20], which consist of failure or censoring times for 139 appliances (36 in Group I, 51 in Group II, and 52 in Group III) subjected to a manual lifetime test. This example has been used by several authors for illustration, see for example, Crowder [10] and Park and Kulasekera [24]. Although there are three groups, we consider group II for illustration. There are 51 appliances in Group II. Fifteen failures are due to mode 11, 6 failures are due to other modes and remaining are censored observations. For illustrative purpose, we consider failure mode 11 as cause 1 and other failure failure modes are considered as cause 2. So there are 15 failures due to cause 1 and 6 failures due to cause 2. We first analyze the data assuming that the latent cause of failures have independent Weibull distributions with the different shape (α_1 and α_2) and scale parameters (λ_1 and λ_2). The maximum likelihood estimate of the shape parameters are $\hat{\alpha}_1 = 1.433398$ and $\hat{\alpha}_2 = 0.983995$ with standard errors 0.29338 and 0.35120, respectively. When we perform the following testing of hypothesis problem: $H_0 : \alpha_1 = \alpha_2$, we cannot reject the null hypothesis, this justifies that shape parameters can be taken as equal.

Next we analyze the data by generating progressively type-II censored data from the original data set. We generate progressively type-II censored sample with $m = 12$ and censoring schemes $R_1 = 5$, $R_2 = 2$, $R_3 = 2$, $R_4 = 2$, $R_5 = 14$, $R_6 = 0$, $R_7 = 0$, $R_8 = 0$, $R_9 = 3$, $R_{10} = 0$, $R_{11} = 6$ and $R_{12} = 5$. The progressively type-II sample is (45, 2), (47, 1), (73, 1), (145, 1), (281, 2), (311, 1), (471, 2), (490, 1), (569, 2), (575, 1), (630, 1), (838, 1). Here we have $m_1 = 8$ failure due to cause 1 and $m_2 = 4$ failures due to cause 2. The maximum likelihood estimates of α, λ_1 and λ_2 with standard error in parentheses are 1.34094 (0.31988), 0.000051 (0.00010) and 0.000025 (0.000052), respectively. The asymptotic 95% confidence intervals of α, λ_1 and λ_2 based on observed information matrix are (0.71397, 1.96790), (-0.00015, 0.00028), (-0.00008, 0.000128).

Next, we compute the Bayes estimate of α, λ_1 and λ_2. We compute Bayes estimate under non-informative priors, since we have no prior information about the unknown parameters. Based on data, the posterior density function of α can be approximated by the gamma density function by equating the first two moments. The scale and shape parameters of the approximate gamma distribution are 17.3962 and 13.0559, respectively. Based on 10,000 MCMC samples, the Bayes estimates of α, λ_1 and λ_2 are 1.33406 (0.32239), 0.00025 (0.00061) and 0.00012 (0.00030). The 95% HPD credible intervals are (0.74162, 1.97743), (7.587188e-010, 0.00105) and (3.793594e-010, 0.00052), respectively.

Example 2: Here we consider another set of data to illustrate our methodology. We analyze the data of a lung cancer clinical trial being conducted by the Eastern Cooperative Oncology Group, Lagakos [18]. There are 194 cases with 83 deaths from cause 1 (local spread of disease) and 44 from cause 2 (metastatic spread), the remaining 67 times being right-censored. We analyze this data set how the proposed method can be applied for random censored data. The present data set is an example of competing risks data with exactly two causes of failure. The latent failure times independently Weibull distributed with same shape parameters but different scale parameters (See Crowder [10]). The data set was analyzed with different shape parameters to check whether the shape parameters are equal or not. Let α_1 and α_2 be the shape parameters of two latent time distributions. The maximum likelihood estimate of α_1 and α_2 are $\hat{\alpha}_1$= 1.30374 and $\hat{\alpha}_2$= 1.36784 with standard error 0.05731 and 0.08895, respectively. It is clear that the there is no significant difference between the shape parameter values. In this case also the null hypothesis $H_0 : \alpha_1 = \alpha_2$ cannot be rejected. Hence the shape parameter for both the Weibull distribution can be taken as equal. The justification for Weibull distribution is given in Crowder [10].

The maximum likelihood estimates of α, λ_1 and λ_2 with standard errors in parentheses are 1.32547 (0.08547), 0.00767 (0.00240) and 0.00407 (0.00134), respectively. The asymptotic 95% confidence intervals of α, λ_1 and λ_2 based on observed information matrix are (1.15795, 1.49299), (0.00297, 0.01237) and (0.00144, 0.0067). Next we compute Bayes estimates of α, λ_1 and λ_2 under non-informative priors. Based on data, the posterior density function of α can be approximated by the gamma density function by equating the first two moments. The scale and shape parameters of the approximate gamma distribution are 240.0418 and 181.3844, respectively. Based on 10,000 MCMC samples, the Bayes estimates of α, λ_1 and λ_2 with standard errors in parentheses are 1.32493 (0.08584), 0.00800 (0.00236) and 0.00424 (0.00125). The 95% HPD credible intervals are (1.16285, 1.49920), (0.00397, 0.01277) and (0.00210, 0.00677).

5 Optimal Censoring Scheme

Till this point we have discussed the Bayesian inference of the unknown parameters for progressively censored competing risks data. It is assumed so far that the progressive censoring is pre-fixed, *i.e.* n, m and $R_1, \cdots, R_m$ are known apriori. But in practice the natural question is, whether we should choose the progressive censoring scheme based on convenience or based on some scientific criterion.

In the last few years finding the "optimum" censoring scheme has received considerable attention in the recent statistical literature, see for example Balakrishnan and Aggarwala [2] (Chapter 10), Ng *et al.* [21], Balasooriya *et al.* [6], Burkschat [7], Burkschat *et al.* [8] and the references cited therein. Optimum censoring scheme means, among all the possible progressive censoring schemes, find the scheme which is "best", with respect to a certain criterion. Naturally we would like to choose that censoring scheme which provides maximum "information" of the unknown parameters. Intuitively, it is quite clear that if we do not have any restriction on n and m, we should choose $n = m$ and n should be as large as possible. In practice often we do not have any choice of n and m, and hence from now onwards if nothing is mentioned it is assumed that n and m are fixed.

Now all possible progressive censoring schemes, for fixed n and m, means all possible choices of $R_1, \cdots, R_m$, such $R_1 + \cdots + R_m = n - m$. Therefore, in this section, if nothing is mentioned it is assumed that n and m are fixed. There are two issues involved in dealing with the optimum censoring scheme, namely (i) find a proper criterion, (ii) with respect to that criterion find the best censoring scheme. Interestingly both are quite important and none of them is an trivial issue in this case.

5.1 Precision Criterion

To find a proper criterion, one needs to define an "information measure" of a given sampling scheme $R = \{R_1, \cdots, R_m\}$. Once the information measure of a given sampling scheme is properly defined, it is possible to compare two different sampling schemes, and hence the "optimum" censoring scheme can be obtained. It is clear that "optimum" censoring scheme will depend on the "information" measure defined for a given scheme R.

First let us discuss how the "information" measure can be defined for a given R, and how it can be used to compare two different censoring schemes. In this respect, the Fisher information of the given censoring scheme seems to be a natural choice. If only one parameter is unknown, then comparison between the two Fisher information of two censoring schemes, simply boils down to compare two real numbers. But if more than one parameter is unknown, then comparison of two censoring schemes means to compare two Fisher information matrices, which may not be a trivial issue. Different methods have been proposed in the literature to compare two Fisher information matrices, for example the trace or determinant of the Fisher information matrix. But unfortunately it is known that they are not scale invariant.

Alternatively, the variances of the $100p$-th quantile estimator can be used as an information measure for a given censoring scheme, and hence can be used quite effectively to compare two different censoring schemes. Zhang and Meeker [28] and Ng *et al.* [21] used this criterion to find optimum censoring schemes in two different problems. It may be mentioned that the above criterion is scale invariant, but it may depend on p, and it may be chosen based on some practical consideration. Zhang and Meeker [28] and Ng *et al.* [21] used $p = 0.95$ and 0.99, but other values of p also can be chosen.

Recently, Pareek *et al.* [23] proposed the following two information measures for a given progressive censoring scheme R, in the frequentist context and when the competing risks data are observed. Suppose the p-quantile points of the two latent lifetime distributions are

$$T_{p,1} = \left[-\frac{1}{\lambda_1}\ln(1-p)\right]^{\frac{1}{\alpha}}, \quad \text{and} \quad T_{p,2} = \left[-\frac{1}{\lambda_2}\ln(1-p)\right]^{\frac{1}{\alpha}} \tag{22}$$

for cause 1 and cause 2 respectively. Then for fixed $0 \le w \le 1$, the Criterion 1 is defined as

$$C_1(R) = wVar\left(\ln \widehat{T}_{p,1}\right) + (1-w)Var\left(\ln \widehat{T}_{p,2}\right) \tag{23}$$

here $Var\left(\ln \widehat{T}_{p,1}\right)$ and $Var\left(\ln \widehat{T}_{p,2}\right)$ are the asymptotic variance of the MLEs of $\ln \widehat{T}_{p,1}$ and $\ln \widehat{T}_{p,2}$ respectively, and although they might depend on the censoring scheme R, but we are not making it explicit for brevity. The weight w should be chosen depending on the importance the practitioner wants to put on cause 1 or cause 2.

Similarly, Criterion 2 has been proposed as

$$C_2(R) = w\int_0^1 Var\left(\ln \widehat{T}_{p,1}\right) dW_1(p) + (1-w)\int_0^1 Var\left(\ln \widehat{T}_{p,2}\right) dW_2(p), \tag{24}$$

here w, $Var\left(\ln \widehat{T}_{p,1}\right)$ and $Var\left(\ln \widehat{T}_{p,2}\right)$ are same as defined before. Moreover, the weight functions $W_1(p) \ge 0$, $W_2(p) \ge 0$ and they satisfy

$$\int_0^1 dW_1(p) = \int_0^1 dW_2(p) = 1. \tag{25}$$

The two weight functions $W_1(\cdot)$ and $W_2(\cdot)$ have to be decided before hand depending on the problem, see for example Pareek *et al.* [23] for discussions.

Following the procedure of Kundu [15], the natural modifications of $C_1(R)$ and $C_2(R)$ will be

Criterion 1: Let for any censoring scheme R

$$C_1^B(R) = wE_{data}\left[V_{posterior}\left(\ln T_{p,1}\right)\right] + (1-w)E_{data}\left[V_{posterior}\left(\ln T_{p,2}\right)\right]. \tag{26}$$

Here $V_{posterior}\left(\ln \widehat{T}_{p,1}\right)$ and $V_{posterior}\left(\ln \widehat{T}_{p,2}\right)$ denote the posterior variance of $\ln T_{p,1}$ and $\ln T_{p,2}$ respectively. Moreover, $E_{data}(\cdot)$ means unconditional expectation with respect to the *data*.

If we have two censoring schemes say $R^1 = \{R_1^1, \cdots, R_m^1\}$ and $R^2 = \{R_1^2, \cdots, R_m^2\}$, we say R^1 (R^2) is better than R^2 (R^1) if $C_1^B(R_1) < (>)$ $C_1^B(R_2)$.

Criterion 2: If for any censoring scheme R

$$C_2^B(R) = wE_{data}\left[\int_0^1 V_{posterior}\left(\ln T_{p,1}\right) dW_1(p)\right] + (1-w)E_{data}\left[\int_0^1 V_{posterior}\left(\ln T_{p,2}\right) dW_2(p)\right], \tag{27}$$

then we say R^1 (R^2) is better than R^2 (R^1) if $C_2^B(R_1) < (>)$ $C_2^B(R_2)$.

It is clear that the computation of (26) and (27) are not very easy. We have used Monte Carlo simulation technique to approximate (26) and (27), as proposed in Kundu [15]. For illustrative purposes, we present the optimal sampling scheme for different objective functions (OF) for selected combination of m and n. Here it is assumed $a_1 = a_2 = b_1 = b_2 = 1.0$, and $a_0 = b_0 = 2$. We have used the minimum trace criterion (OF-1), minimum determinant criterion (OF-2), and the minimum sum of the variances of the p-th percentile estimators for different choices of p, namely $p = 0.1$ (OF-3), $p = 0.99$ (OF-4). Finally we have also computed the optimum scheme based on the minimum sum of the weighted variances (OF-5) under the assumption that $W_1(p) = W_2(p) = 1$ for all $0 \le p \le 1$. Moreover in calculating OF-3, OF-4 and OF-5, it is assumed that $w = 1/2$.

Note that in all cases the optimization has to be performed numerically. They are discrete optimization problem. For a given n and m, the optimum censoring scheme with respect to to a given criterion can be found by exhaustive search for all possible values of R_i's satisfying (2). The results are reported in Tables 6 and 7. Since Type-II censoring scheme is a particular choice of the general progressive censoring scheme, and it is one of the most used censoring scheme, we report the relative efficiency and also the relative expected experimental time of the Type-II censoring scheme with respect to the optimal censoring scheme. To compute relative efficiency, we need expected Fisher information matrix. The expression of expected Fisher information matrix is given in Appendix B. It is clear that the relative expected experimental time of the Type-II censoring scheme is significantly smaller than the optimal censoring scheme, but the relative efficiencies are quite high. That justifies the overwhelming popularity of the Type-II censoring scheme.

6 Concluding remarks

In this work, we have considered the Bayesian analysis of the competing risks data when they are Type-II progressively censored. Latent failure time modeling of Cox [9] has been assumed in analyzing the competing risks data. The latent lifetime distributions are assumed to be Weibull with the same shape but different scale parameters. Although, in this paper we have mainly dealt with two causes of failure only, the work can be extended to more than two causes of failure. Prior selection has been suggested using the idea of Pena and Gupta [25], which is a very flexible prior.

We have proposed different criteria for selecting optimal censoring schemes and presented some optimal censoring scheme for selected choice of n and m. We further compute the relative efficiency of the Type-II censoring schemes with the optimal censoring schemes. It is observed that the relative efficiency of the Type-II censoring schemes is quite high. It justifies the usefulness of the Type-II censoring scheme in practice.

Acknowledgements:

The auhtors would like to thank one anonymous referee for many constructive suggestions.

Appendix A: Generation from Beta-Dirichlet Distribution

In this appendix we will mention how to generate samples from (20). It has already been shown that

$$
\begin{aligned}
&(\lambda_1, \lambda_2) \sim \text{BD}\left(b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha}, a_0 + m_1 + m_2, a_1 + m_1, a_2 + m_2\right) \Leftrightarrow \\
&(\lambda_1 + \lambda_2) \sim \text{Gamma}\left(a_0 + m_1 + m_2, b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha}\right) \quad \text{and} \\
&\lambda_1 | (\lambda_1 + \lambda_2) \sim \text{Beta}(a_1 + m_1, a_2 + m_2).
\end{aligned}
$$

Therefore, first we generate from a gamma distribution with the shape and scale parameters as $a_0 + m_1 + m_2$ and , $b_0 + \sum_{i=1}^{m}(R_i+1)t_i^{\alpha}$ respectively, and then we generate from a Beta distribution with the parameters as $a_1 + m_1$ and $a_2 + m_2$ respectively. Note that the generation from a Beta distribution can be easily performed using the acceptance-rejection principle, where the dominating function can be taken as the uniform (0, 1).

Appendix B: Expected Fisher Information Matrix

For completeness purposes we have provided the expected Fisher information matrix. Let us use the following notation;

$$
r_j = m - j + 1 + \sum_{i=j}^{m} R_i, \quad c_{j-1} = \prod_{i=1}^{j} r_i, \quad a_{1,1} = 1, \quad a_{i,j} = \prod_{k=1, k\neq i}^{j} \frac{1}{r_k - r_i}.
$$

If the 3×3 expected Fisher information matrix is denoted by E, then

$$
E = \begin{bmatrix} E_{11} & E_{12} & E_{13} \\ E_{21} & E_{22} & E_{23} \\ E_{31} & E_{32} & E_{33} \end{bmatrix} = -E \begin{bmatrix} \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \alpha^2} & \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \alpha \partial \lambda_1} & \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \alpha \partial \lambda_2} \\ \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \lambda_1 \partial \alpha} & \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \lambda_1^2} & \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \lambda_1 \partial \lambda_2} \\ \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \lambda_2 \partial \alpha} & \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \lambda_2 \partial \lambda_1} & \frac{\partial^2 \ln l(\alpha,\lambda_1,\lambda_2)}{\partial \lambda_2^2} \end{bmatrix}, \qquad (28)
$$

where

$$
E_{11} = \frac{m}{\alpha^2} + \frac{(\lambda_1 + \lambda_2)}{\alpha^2} \sum_{j=1}^{m} (R_j + 1) h_{1j}
$$

$$
\begin{aligned}
E_{22} &= \frac{1}{\lambda_1(\lambda_1+\lambda_2)}, \qquad E_{33} = \frac{1}{\lambda_2(\lambda_1+\lambda_2)} \\
E_{23} &= E_{32} = 0 \\
E_{12} &= E_{21} = E_{13} = E_{31} = \frac{1}{\alpha}\sum_{j=1}^{m}(R_j+1)h_{2j},
\end{aligned}
$$

and

$$
\begin{aligned}
h_{1j} &= \frac{c_{j-1}}{(\lambda_1+\lambda_2)}\sum_{i=1}^{j}\frac{a_{i,j}}{r_i^2}\left[\frac{\pi^2}{6} - 2\gamma + \gamma^2 - 2(1-\gamma)\ln\{r_i(\lambda_1+\lambda_2)\} + (\ln\{r_i(\lambda_1+\lambda_2)\})^2\right] \\
h_{2j} &= \frac{c_{j-1}}{(\lambda_1+\lambda_2)}\sum_{i=1}^{j}\frac{a_{i,j}}{r_i^2}\left[1-\gamma-\ln(r_i(\lambda_1+\lambda_2))\right].
\end{aligned}
$$

References

[1] Balakrishnan, N. and Sandhu, R.A. (1995), "A simple algorithm for generating progressively Type-II generated samples", *American Statistician*, vol. 49, 229 - 230.

[2] Balakrishnan, N. and Aggarwala, R. (2000), *Progressive Censoring: Theory, Methods and Applications*, Birkhäuser, Boston.

[3] Balakrishnan, N. (2007), "Progressive censoring methodology: an appraisal", (with discussions), *Test*, vol. 16, 211 - 289.

[4] Balakrishnan, N., Burkschat, M., Cramer, E. and Hoffman, G. (2008), "Fisher informa- tion based progressive censoring plans", *Computational Statistics and Data Analysis*, vol. 53, 366 - 380.

[5] Berger, J. O. and Sun, D. (1993), "Bayesian analysis for the Poly-Weibull distribution", *Journal of the American Statistical Association*, vol. 88, 1412 - 1418.

[6] Balasooriya, U. Saw, S.L.C., and Gadag, V.G. (2000), "Progressively censored reliability sampling plan for Weibull distribution", *Technometrics*, vol. 42 160 - 168.

[7] Burkschat, M. (2008), "On optimality of extremal schemes in progressive type-II censoring", *Journal of Statistical Planning and Inference*, vol. 138, 1647 - 1659.

[8] Burkschat, M., Cramer, E. and Kemps, U. (2007), "On optimal schemes in progressive censoring", *Statistics and Probability Letters*, vol. 76, 1032 - 1036.

[9] Cox, D.R. (1959), "The analysis of exponentially distributed lifetimes with two types of failures", *Journal of the Royal Statistical Society*, Ser. B, vol. 21, 411 - 421.

[10] Crowder, M. (2001), *Classical Competing Risks Model*, Chapman & Hall/ CRC, New York.

[11] Devroye, L. (1984), "A simple algorithm for generating random variables using log-concave density", *Computing*, vol. 33, 247 - 257.

[12] Hoel, D. G. (1972), "A representation of mortality data by competing risks", *Biometrics*, vol. 28, 475-488.

[13] Kalbfleish, J.D. and Prentice, R.L. (1980), *The Statistical Analysis of Failure Data*, New York, Wiley.

[14] Kaminskiy, M. P. and Krivtsov, V. V. (2005), "A simple procedure for Bayesian estimation of the Weibull distribution", *IEEE Transactions on Reliability*, vol. 54, 612 - 616.

[15] Kundu, D. (2008), "Bayesian inference and life testing plan for the Weibull distribution in presence of progressive censoring", *Technometrics*, vol. 50, 144–154.

[16] Kundu, D. and Basu, S. (2000), "Analysis of incomplete data in presence of competing risks", *Journal of Statistical Planning and Inference*, vol. 87, 221 - 239.

[17] Kundu, D., Kannan, N. and Balakrishnan, N. (2004), "Analysis of Progressively Censored Competing Risks Data", *Handbook of Statistics on Survival Analysis* , Editors: C.R. Rao and N. Balakrishnan, Elsevier Publications, vol. 23, 331-348.

[18] Lagakos, S. W. (1978), "A covariate model for partially censored data subject to competing causes of failure", *Applied Statistics*, vol. 27, 235-241.

[19] Mukhopadhyay, C. and Basu, A. P. (1997), " Bayesian analysis of incomplete time and cause of failure data", *Journal of Statistical Planning and Inference*, vol. 59, 79-100.

[20] Nelson, W. (1970), "Hazard plotting methods for analysis of life data with different failure modes", *Journal of Quality Technology*, vol. 2, 126-149.

[21] Ng, T., Chan, C.S. and Balakrishnan, N. (2004), "Optimal progressive censoring plans for the Weibull distribution", *Technometrics*, vol. 46, 470 - 481.

[22] Pradhan, B. and Kundu, D. (2009), "On progressively censored generalized exponential distribution", *Test*, vol. 18, 497 - 515.

[23] Pareek, B., Kundu, D. and Kumar, S. (2009), "On progressive censored competing risks data for Weibull distributions", *Computational Statistics and Data Analysis*, vol. 53, 4083 - 4094.

[24] Park, C. and Kulasekera, K. B. (2004), "Parametric inference of incomplete data with competing risks among several groups", *IEEE Transactions on Reliability*, vol. 53, 11-21.

[25] Pena, E. A. and Gupta, A. K. (1990), "Bayes estimation for the Marshall-Olkin exponential distribution", *Journal of the Royal Statistical Society*, Ser. B, vol. 52, 379 - 389.

[26] Rao, B.R., Talwalker, S. and Kundu, D. (1991), "Confidence intervals for the relative risk ratio parameters from survival data under a random epidemiology study", *Biometrical Journal*, vol. 33, 959 - 984.

[27] Viveros, R. and Balakrishnan, N. (1994), "Interval estimation of parameters of life from progressively censored data", *Technometrics*, vol. 36, 84 - 91.

[28] Zhang, Y. and Meeker, W.Q. (2005), "Bayesian life test planning for the Weibull distribution with the given shape parameter", *Metrika*, vol. 61, 237–249.

Table 2: The average values of the MLEs and Bayes estimates under different priors along with the MSE's in parentheses when $\alpha = 1$, $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$.

Estimator		CS-1	CS-2	CS-3	CS-4	CS-5	CS-6	CS-7
MLE	α	1.1398	1.1400	1.1455	1.0598	1.0476	1.0534	1.0450
		(0.1341)	(0.1079)	(0.1144)	(0.0347)	(0.0252)	(0.0304)	(0.0235)
	λ_1	0.6585	0.7289	0.7952	0.6227	0.6227	0.6172	0.6293
		(0.0965)	(0.1812)	(0.3743)	(0.0249)	(0.0256)	(0.0235)	(0.0281)
	λ_2	0.4491	0.4919	0.5372	0.4243	0.4160	0.4204	0.4211
		(0.0663)	(0.1030)	(0.2074)	(0.0197)	(0.0175)	(0.0185)	(0.0185)
Bayes	α	1.1621	1.0986	1.1308	1.0556	1.0412	1.0455	1.0417
Prior 0		(0.1431)	(0.0919)	(0.1053)	(0.0339)	(0.0255)	(0.0307)	(0.0273)
	λ_1	0.66735	0.6658	0.7893	0.6227	0.6133	0.6161	0.6102
		(0.1281)	(0.0980)	(0.2062)	(0.0228)	(0.0212)	(0.0236)	(0.0253)
	λ_2	0.4568	0.4604	0.5318	0.4132	0.4127	0.4135	0.4099
		(0.0625)	(0.0422)	(0.2442)	(0.0163)	(0.0164)	(0.0147)	(0.0154)
Bayes	α	1.0778	1.0630	1.0587	1.0469	1.0359	1.0423	1.0300
Prior 1		(0.0476)	(0.0444)	(0.0377)	(0.0263)	(0.0193)	(0.0244)	(0.0200)
	λ_1	0.6420	0.6541	0.6789	0.6232	0.6166	0.6139	0.6213
		(0.0557)	(0.0484)	(0.0571)	(0.0221)	(0.0207)	(0.0193)	(0.0222)
	λ_2	0.4329	0.4358	0.4475	0.4103	0.4120	0.4119	0.4177
		(0.0364)	(0.0335)	(0.0369)	(0.0141)	(0.0141)	(0.0135)	(0.0142)
Bayes	α	1.1960	1.1627	1.1415	1.0814	1.0665	1.0924	1.0676
Prior 2		(0.1231)	(0.0944)	(0.0768)	(0.0408)	(0.0333)	(0.0403)	(0.0330)
	λ_1	0.6155	0.6573	0.6703	0.6098	0.5937	0.5981	0.5995
		(0.0580)	(0.0640)	(0.0644)	(0.0237)	(0.0211)	(0.0207)	(0.0205)
	λ_2	0.4498	0.4775	0.4817	0.4120	0.4057	0.4079	0.4194
		(0.0452)	(0.0445)	(0.0463)	(0.0143)	(0.0136)	(0.0133)	(0.0150)

Table 3: The average confidence length/HPD credible length and coverage percentage in parentheses when $\alpha = 1$, $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$

Estimator		CS-1	CS-2	CS-3	CS-4	CS-5	CS-6	CS-7
MLE	α	1.1146	1.0691	1.0984	0.6631	0.5807	0.6268	0.5735
		(0.96)	(0.96)	(0.96)	(0.96)	(0.95)	(0.95)	(0.96)
	λ_1	0.9906	1.1384	1.4709	0.5832	0.5893	0.5769	0.5855
		(0.94)	(0.95)	(0.96)	(0.94)	(0.94)	(0.94)	(0.94)
	λ_2	0.8038	0.8866	1.1093	0.4783	0.4762	0.4730	0.4759
		(0.92)	(0.95)	(0.95)	(0.93)	(0.93)	(0.93)	(0.93)
Bayes Prior 0	α	1.1161	0.9982	1.0579	0.6549	0.5636	0.6148	0.5570
		(0.94)	(0.96)	(0.96)	(0.96)	(0.94)	(0.95)	(0.94)
	λ_1	0.9620	0.9576	1.3094	0.5668	0.5688	0.5623	0.5568
		(0.93)	(0.93)	(0.97)	(0.95)	(0.95)	(0.94)	(0.93)
	λ_2	0.7579	0.7401	0.9942	0.4546	0.4572	0.4554	0.4493
		(0.94)	(0.95)	(0.95)	(0.94)	(0.94)	(0.93)	(0.94)
Bayes Prior 1	α	0.9051	0.8468	0.8337	0.6118	0.5347	0.5836	0.5255
		(0.98)	(0.97)	(0.97)	(0.96)	(0.96)	(0.96)	(0.95)
	λ_1	0.8520	0.8365	0.6789	0.5517	0.5545	0.5449	0.5518
		(0.95)	(0.96)	(0.97)	(0.95)	(0.95)	(0.95)	(0.94)
	λ_2	0.6805	0.6550	0.4475	0.4421	0.4464	0.4411	0.4476
		(0.93)	(0.94)	(0.95)	(0.94)	(0.94)	(0.94)	(0.95)
Bayes Prior 2	α	1.0645	0.9778	0.9499	0.6598	0.5658	0.6273	0.5625
		(0.95)	(0.95)	(0.97)	(0.90)	(0.88)	(0.89)	(0.86)
	λ_1	0.8435	0.8839	0.9641	0.5464	0.5435	0.5364	0.5408
		(0.91)	(0.95)	(0.96)	(0.93)	(0.93)	(0.94)	(0.93)
	λ_2	0.6999	0.7201	0.7681	0.4419	0.4400	0.4365	0.4455
		(0.93)	(0.96)	(0.97)	(0.94)	(0.94)	(0.94)	(0.94)

Table 4: The average values of MLEs and the Bayes estimates under different priors along with the MSE's in parentheses when $\alpha = 2$, $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$.

Estimator		CS-1	CS-2	CS-3	CS-4	CS-5	CS-6	CS-7
MLE	α	2.2799	2.2801	2.2910	2.1195	2.0923	2.1073	2.0876
		(0.5364)	(0.4314)	(0.4575)	(0.1388)	(0.0988)	(0.1214)	(0.0928)
	λ_1	0.6585	0.7289	0.7952	0.6227	0.6230	0.6172	0.6296
		(0.0965)	(0.1811)	(0.3743)	(0.0249)	(0.0257)	(0.0235)	(0.0278)
	λ_2	0.4491	0.4919	0.5372	0.4243	0.4177	0.4204	0.4207
		(0.0663)	(0.1030)	(0.2074)	(0.0972)	(0.0179)	(0.0185)	(0.0182)
Bayes Prior 0	α	2.2754	2.2645	2.2465	2.1112	2.0824	2.0910	2.0835
		(0.4719)	(0.4050)	(0.4186)	(0.1355)	(0.1021)	(0.1226)	(0.1090)
	λ_1	0.6763	0.7351	0.8126	0.6226	0.6133	0.6161	0.6102
		(0.1803)	(0.1604)	(0.3104)	(0.0228)	(0.0212)	(0.0236)	(0.0253)
	λ_2	0.4595	0.4991	0.5385	0.4132	0.4127	0.4135	0.4135
		(0.0620)	(0.1144)	(0.1623)	(0.0163)	(0.0164)	(0.0147)	(0.0154)
Bayes Prior 3	α	2.1554	2.1261	2.1173	2.0937	2.0718	2.0847	2.0601
		(0.1905)	(0.1774)	(0.1508)	(0.1051)	(0.0771)	(0.0976)	(0.0799)
	λ_1	0.6420	0.6541	0.6789	0.6232	0.6166	0.6139	0.6113
		(0.0557)	(0.0484)	(0.0571)	(0.0221)	(0.0207)	(0.0193)	(0.0222)
	λ_2	0.4329	0.4358	0.4475	0.4103	0.4120	0.4112	0.4177
		(0.0364)	(0.0335)	(0.0369)	(0.0141)	(0.0141)	(0.0135)	(0.0142)
Bayes Prior 4	α	2.4204	2.4818	2.5274	2.1774	2.1108	2.1713	2.1427
		(0.4894)	(0.5103)	(0.5547)	(0.1366)	(0.0963)	(0.1285)	(0.1051)
	λ_1	0.7464	0.8423	0.9306	0.6617	0.6554	0.6499	0.6532
		(0.1168)	(0.1760)	(0.2583)	(0.0309)	(0.0286)	(0.0240)	(0.6532)
	λ_2	0.4609	0.5220	0.5902	0.4331	0.4208	0.4262	0.4187
		(0.0518)	(0.0727)	(0.1174)	(0.0166)	(0.0155)	(0.0168)	(0.0158)

Table 5: The average confidence length/HPD credible length and coverage percentage in parentheses when $\alpha = 2$, $\lambda_1 = 0.6$ and $\lambda_2 = 0.4$

Estimator		CS-1	CS-2	CS-3	CS-4	CS-5	CS-6	CS-7
MLE	α	2.2293	2.1383	2.1969	1.3263	1.1607	1.2536	1.1465
		(0.96)	(0.96)	(0.96)	(0.95)	(0.96)	(0.96)	(0.96)
	λ_1	0.9906	1.1384	1.4709	0.5832	0.5900	0.5769	0.5856
		(0.94)	(0.95)	(0.96)	(0.94)	(0.94)	(0.94)	(0.94)
	λ_2	0.8038	0.8866	1.1094	0.4783	0.4772	0.4730	0.4756
		(0.92)	(0.95)	(0.95)	(0.93)	(0.93)	(0.93)	(0.93)
Bayes	α	2.1871	2.0830	2.1110	1.3098	1.1272	1.2359	1.1139
Prior 0		(0.95)	(0.95)	(0.95)	(0.96)	(0.94)	(0.95)	(0.94)
	λ_1	0.9763	1.0818	1.3769	0.5668	0.5688	0.5623	0.5568
		(0.93)	(0.95)	(0.97)	(0.95)	(0.95)	(0.94)	(0.93)
	λ_2	0.7649	0.8371	1.0087	0.4546	0.4572	0.4539	0.4493
		(0.93)	(0.93)	(0.94)	(0.94)	(0.94)	(0.94)	(0.94)
Bayes	α	1.8102	1.6936	1.6673	1.2236	1.0694	1.1673	1.0512
Prior 3		(0.98)	(0.96)	(0.97)	(0.95)	(0.96)	(0.96)	(0.95)
	λ_1	0.8520	0.8365	0.9267	0.5517	0.5545	0.5449	0.5518
		(0.95)	(0.96)	(0.97)	(0.95)	(0.95)	(0.95)	(0.94)
	λ_2	0.6805	0.6555	0.7075	0.4421	0.4464	0.4411	0.4476
		(0.93)	(0.94)	(0.95)	(0.94)	(0.94)	(0.95)	(0.94)
Bayes	α	1.9970	1.9603	1.9687	1.2617	1.0749	1.2032	1.0785
Prior 4		(0.93)	(0.90)	(0.89)	(0.94)	(0.94)	(0.95)	(0.94)
	λ_1	0.9589	1.0894	1.3283	0.5797	0.5807	0.5688	0.5695
		(0.94)	(0.95)	(0.96)	(0.94)	(0.94)	(0.96)	(0.95)
	λ_2	0.7244	0.8037	0.9694	0.4623	0.4562	0.4546	0.4493
		(0.92)	(0.95)	(0.97)	(0.94)	(0.95)	(0.94)	(0.94)

Table 6: The optimal censoring scheme for different objective functions when m = 5 and n = 10, 15, 20, 25 and 30. The relative efficiency (RE) and the relative expected time (RT) of Type-II censoring scheme with respect to the optimum censoring scheme are reported.

OF	n	R_1	R_2	R_3	R_4	R_5	RE	RT
1	10	1	4	0	0	0	96.1%	53.6%
	15	1	9	0	0	0	96.8%	46.0%
	20	0	15	0	0	0	98.7%	35.1%
	25	0	20	0	0	0	98.9%	33.5%
	30	0	0	25	0	0	99.5%	32.8%
2	10	4	1	0	0	0	92.9%	52.1%
	15	1	9	0	0	0	95.1%	46.0%
	20	0	15	0	0	0	96.8%	35.1%
	25	0	0	20	0	0	98.1%	37.8%
	30	0	0	25	0	0	98.8%	34.5%
3	10	0	0	0	0	5	100.0%	100.0%
	15	0	0	0	0	10	100.0%	100.0%
	20	0	0	0	0	15	100.0%	100.0%
	25	0	0	0	0	20	100.0%	100.0%
	30	0	0	0	1	24	99.9%	55.4%
4	10	5	0	0	0	0	89.4%	54.9%
	15	10	0	0	0	0	90.0%	43.0%
	20	15	0	0	0	0	91.3%	36.7%
	25	20	0	0	0	0	92.5%	32.5%
	30	25	0	0	0	0	93.4%	29.5%
5	10	5	0	0	0	0	98.0%	54.9%
	15	10	0	0	0	0	98.0%	43.0%
	20	15	0	0	0	0	98.4%	36.7%
	25	20	0	0	0	0	98.7%	32.5%
	30	25	0	0	0	0	98.9%	29.5%

Table 7: The optimal censoring scheme for different objective functions when $n = 15$, and $m = 6$, 8 and 10. The relative efficiency (RE) and the relative expected time (RT) of Type-II censoring scheme with respect to the optimum censoring scheme are reported.

$m = 6$

OF	R_1	R_2	R_3	R_4	R_5	R_6	RE	RT
1	1	8	0	0	0	0	96.8%	46.8%
2	1	8	0	0	0	0	95.7%	46.8%
3	0	0	0	0	0	9	100.0%	100.0%
4	9	0	0	0	0	0	90.0%	46.1%
5	9	0	0	0	0	0	98.2%	46.1%

$m = 8$

OF	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	RE	RT
1	1	6	0	0	0	0	0	0	98.6%	53.5%
2	1	6	0	0	0	0	0	0	93.9%	53.5%
3	0	0	0	0	0	0	0	7	100.0%	100.0%
4	7	0	0	0	0	0	0	0	92.1%	52.4%
5	0	7	0	0	0	0	0	0	97.8%	53.6%

$m = 10$

OF	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	RE	RT
1	0	1	4	0	0	0	0	0	0	0	97.9%	62.2%
2	0	5	0	0	0	0	0	0	0	0	94.9%	61.7%
3	0	0	0	0	0	0	0	0	0	5	100.0%	100.0%
4	5	0	0	0	0	0	0	0	0	0	95.1%	61.2%
5	0	2	3	0	0	0	0	0	0	0	98.4%	62.0%