BAYESIAN ANALYSIS OF LEFT TRUNCATED RIGHT CENSORED DATA IN THE PRESENCE OF DEPENDENT COMPETING RISKS

A Project Report Submitted

for the Course

MA499 Project II

by

Kaushal Chhalani

(Roll No. 180123023)

Manav Chirania

(Roll No. 180123026)

to the

DEPARTMENT OF MATHEMATICS

INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI

GUWAHATI - 781039, INDIA

April 2022

CERTIFICATE

This is to certify that the work contained in this project report entitled "Bayesian Analysis of left truncated right censored data in the presence of dependent competing risks" submitted by Kaushal Chhalani (Roll No.: 180123023) & by Manav Chirania (Roll No.: 180123026) to the Department of Mathematics, Indian Institute of Technology Guwahati towards partial requirement of Bachelor of Technology in Mathematics and Computing has been carried out by them under my supervision.

It is also certified that, along with literature survey, a few new results are established by the student under the project.

Turnitin Similarity: 15 %

Guwahati - 781 039 (Dr. Ayon Ganguly)

April 2022 Project Supervisor

ABSTRACT

Significant work has been done in the area of prediction of lifetime data analysis in the presence of competing risks. A lot of this work has been carried out considering independent competing risks, and not much has been done considering dependent competing risks. In this paper, we will consider left-truncated right-censored data in the presence of dependent competing risks from a Bayesian perspective. The Block & Basu Bivariate Weibull (BBBW) distribution is used to model the dependence structure between competing risks. We will consider Bayesian estimators and the highest posterior density credible intervals based on conjugate priors, E-Bayesian and Hierarchical Bayesian methods for the point and interval estimation of the model parameters.

Contents

Chapter 1

Introduction

Life testing involves assessing the expected lifetime or durability of an item over time. The primary goal of any life testing experiment is to assess one or more of the experimental units' reliability characteristics. In a life testing experiment, usually, a number of identical items are tested under normal conditions and the *time to failure* is reported. The *time to failure* definition depends on the considered items. For example, in the testing of an electric item, the *time to failure* can be the number of hours before it stops working.

For this project, we have considered a real life data and also simulated data. The description of the data and the aim of the project is discussed in the following subsections.

1.1 Description of Data

1.1.1 Left Truncated Right Censored Data

In this project, we will analyze transformer lifetime data from an energy company in the US. The energy company began documenting records from 1975. The data contains information on transformers which were installed after 1975 and also on units which failed after 1^{st} January, 1975 but were installed before 1^{st} January, 1975. The transformers installed before 1975 represent the transformers sampled from truncated distributions. Thus, the data is left truncated. However, the data is only up to 2021. Therefore, the units still working after 2021 are right censored. Hence, the data is left truncated and right censored.

1.1.2 Competing risks

In lifetime analysis, a unit can fail for many reasons. Therefore, a unit can fail due to one reason and this may prevent us from observing the failure time due to other reasons. Hence, we need to investigate the effect of a specific risk in presence of other risks. This model is known as competing risks model.

1.2 Literature review and aim of the project

After [1], the LTRC data in the presence of competing risks has been addressed by [2]. [3] considered the Bayesian inference of Type-I progressively censored step-stress accelerated life test with dependent competing risks.

The main aim of the project is to combine the work of [2] and [3] to analyse the LTRC data in presence of dependent competing risks from a Bayesian perspective. In this project, we have assumed that there are two risks and the joint distribution of the latent failure times follow a Block and Basu bivariate Weibull (BBBW) distribution.

1.3 Distribution Used

We will be using the notation for the remaining paper. If X has a univariate Weibull distribution with the shape as $\alpha > 0$ and scale parameters as $\lambda > 0$, then for $x > 0$, the probability density function (PDF) is defined as follows;

$$f_{\mathrm{WE}}(x;\alpha,\lambda) = \alpha\lambda x^{\alpha-1}\mathrm{e}^{-\lambda x^{\alpha}}$$

The survival function (SF) is denoted by $S_{\mathrm{WE}}\left(x;\alpha,\lambda\right)$. Also, we denote a Weibull distribution with the scale parameter λ and the shape parameter α by $\mathrm{WE}(\alpha,\lambda)$.

Suppose U_0 follows ($\sim$) $\mathrm{WE}\left(\alpha,\lambda_0\right), U_1 \sim \mathrm{WE}\left(\alpha,\lambda_1\right)$ and $U_2 \sim \mathrm{WE}\left(\alpha,\lambda_2\right)$ and they are mutually independent. If $X_1 = \min\left\{U_0, U_1\right\}$ and $X_2 = \min\left\{U_0, U_2\right\}$, then the joint distribution function of $\left(X_1, X_2\right)$ is said to have the Marshall-Olkin bivariate Weibull distribution. The joint survival function of $\left(X_1, X_2\right)$ can be written as;

$$\begin{aligned} S_{MO}\left(x_1,x_2\right) &= P\left(X_1 > x_1, X_2 > x_2\right) = P\left(U_0 > z, U_1 > x_1, U_2 > x_2\right) \\ &= S_{\mathrm{WE}}\left(x_1;\alpha,\lambda_1\right) S_{\mathrm{WE}}\left(x_2;\alpha,\lambda_2\right) S_{\mathrm{WE}}\left(z;\alpha,\lambda_0\right) \quad \text{where } z = \max\left\{x_1,x_2\right\} \\ &= \begin{cases} S_{\mathrm{WE}}\left(x_1;\alpha,\lambda_1\right) S_{\mathrm{WE}}\left(x_2;\alpha,\lambda_0+\lambda_2\right) & \text{if } 0 < x_1 < x_2 < \infty \\ S_{\mathrm{WE}}\left(x_1;\alpha,\lambda_0+\lambda_1\right) S_{\mathrm{WE}}\left(x_2;\alpha,\lambda_2\right) & \text{if } 0 < x_2 < x_1 < \infty \\ S_{\mathrm{WE}}\left(x;\alpha,\lambda_0+\lambda_1+\lambda_2\right) & \text{if } 0 < x_1 = x_2 = x < \infty. \end{cases} \end{aligned}$$

The joint survival function of $\left(X_1, X_2\right)$ can be written as a mixture of an absolutely continuous part and a singular part as follows;

$$S_{MO}\left(x_1,x_2\right) = \frac{\lambda_1+\lambda_2}{\lambda_0+\lambda_1+\lambda_2} S_a\left(x_1,x_2\right) + \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} S_s\left(x_1,x_2\right),$$

where $S_a(\cdot,\cdot)$ is the absolutely continuous part and $S_S(\cdot,\cdot)$ is the singular part. For $z = \max\{x_1, x_2\}$,

$$S_S(x_1, x_2) = S_{\text{WE}}(z; \alpha, \lambda_0 + \lambda_1 + \lambda_1)$$

and therefore, $S_a(x_1, x_2)$ can be obtained as

$$S_a(x_1, x_2) = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} \mathrm{e}^{-\lambda_1 x_1^\alpha} \mathrm{e}^{-\lambda_2 x_2^\alpha} \mathrm{e}^{-\lambda_0 z^\alpha} - \frac{\lambda_0}{\lambda_1 + \lambda_2} \mathrm{e}^{-(\lambda_0+\lambda_1+\lambda_2) z^\alpha}$$

When $x_1 = x_2 = x$,

$$\begin{aligned}
S_a(x, x) &= \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} \mathrm{e}^{-\lambda_1 x^\alpha} \mathrm{e}^{-\lambda_2 x^\alpha} \mathrm{e}^{-\lambda_0 x^\alpha} - \frac{\lambda_0}{\lambda_1 + \lambda_2} \mathrm{e}^{-(\lambda_0+\lambda_1+\lambda_2) x^\alpha} \\
&= \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} \mathrm{e}^{-(\lambda_0+\lambda_1+\lambda_2) x^\alpha} - \frac{\lambda_0}{\lambda_1 + \lambda_2} \mathrm{e}^{-(\lambda_0+\lambda_1+\lambda_2) x^\alpha} \\
&= \mathrm{e}^{-(\lambda_0+\lambda_1+\lambda_2) x^\alpha}
\end{aligned}$$

Now, we know that BBBW distribution can be obtained from MOBW distribution by removing the singular part and keeping the continuous part. The joint PDF of BBBW can be written as

$$f_{\text{BB}}(y_1, y_2) = \begin{cases} c f_1(y_1, y_2) = c f_{\text{WE}}(y_1; \alpha, \lambda_1) f_{\text{WE}}(y_2; \alpha, \lambda_0 + \lambda_2) & \text{if } 0 < y_1 < y_2 \\ c f_2(y_1, y_2) = c f_{\text{WE}}(y_1; \alpha, \lambda_0 + \lambda_1) f_{\text{WE}}(y_2; \alpha, \lambda_2) & \text{if } 0 < y_2 < y_1, \end{cases}$$

here c is the normalizing constant and $c = \frac{\lambda_0+\lambda_1+\lambda_2}{\lambda_1+\lambda_2}$. Therefore, the joint PDF of (Y_1, Y_2) will be denoted by BBBW $(\alpha, \lambda_0, \lambda_1, \lambda_2)$ and the joint survival function of Y_1 and Y_2 is $S_a(\cdot,\cdot)$.

Chapter 2

Bayesian Inference

Statistical inference for the characteristics of an item based on failure data is a crucial issue when the failure data is obtained via a life test. In this chapter, we aim to use the Bayesian method for estimating the model parameters.

2.1 Bayesian estimation methods

We will be using Bayesian analysis for estimating the model parameters. In Bayesian Analysis, the unknown parameters are considered as RVs which follow some particular prior distributions. Then, based on the prior distributions $\pi(\theta)$ and the observed data $\boldsymbol{x}$, we can obtain the *posterior* distributions of the unknown parameters $\pi(\theta|\boldsymbol{x})$. By Bayes' theorem,

$$\pi(\theta|\boldsymbol{x}) = \frac{f_{\boldsymbol{x}}(\boldsymbol{x}|\theta)\pi(\theta)}{f_{\boldsymbol{x}}(\boldsymbol{x})}$$

$$\therefore \pi(\theta|\boldsymbol{x}) \propto f_{\boldsymbol{x}}(\boldsymbol{x}|\theta) \times \pi(\theta)$$

$$\textit{posterior} \propto \textit{likelihood} \times \textit{prior}$$

But, the above method depends on the choice of prior distribution(or its hyper-parameters). Therefore, some other methods such as the E-Bayesian and hierarchical Bayesian methods are also considered to overcome this issue.

2.1.1 Bayesian estimation

Lets assume λ_k and α are independent variables, and the prior distributions of λ_k and α are given by $\pi(\lambda_k)$ and $\pi(\alpha)$ with hyper-parameters a_k, b_k and a, b for $k = 0, 1, 2$.

Now, according to the Bayes theorem, the joint posterior density of α and $\lambda_k, k = 0, 1, 2$ given the observed data $\boldsymbol{x}$ is

$$\pi(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{\pi(\lambda_0)\pi(\lambda_1)\pi(\lambda_2)\pi(\alpha)L(\boldsymbol{\theta}|\boldsymbol{x})}{\int_{\boldsymbol{\Theta}} \pi(\lambda_0)\pi(\lambda_1)\pi(\lambda_2)\pi(\alpha)L(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta}}$$

where Θ is the product spaces of ranges of the unknown parameters $\boldsymbol{\theta} = (\lambda_0, \lambda_1, \lambda_2, \alpha)$. Subsequently, the posterior densities of the unknown parameters can be obtained as

$$\pi(\lambda_k|\boldsymbol{x}) = \int_{\boldsymbol{\Theta}\setminus Ran(\lambda_k)} \pi(\boldsymbol{\theta}|\boldsymbol{x})d\{\boldsymbol{\theta} \setminus \lambda_k\}, k = 0, 1, 2$$

$$\pi(\alpha|\boldsymbol{x}) = \int_{\boldsymbol{\Theta}\setminus Ran(\alpha)} \pi(\boldsymbol{\theta}|\boldsymbol{x})d\{\boldsymbol{\theta} \setminus \alpha\}$$

Taking the squared loss function as $L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$, we can obtain the Bayesian estimators of the unknown parameters as-

$$\tilde{\lambda}_{kB} = \int_{Ran(\lambda_k)} \lambda_k \pi(\lambda_k|\boldsymbol{x})d\lambda_k, k = 0, 1, 2$$

$$\tilde{\alpha}_B = \int_{Ran(\alpha)} \alpha\pi(\alpha|\boldsymbol{x})d\alpha$$

2.1.2 E-Bayesian estimation

The prior distributions of λ_k and α are $\pi(\lambda_k)$ and $\pi(\alpha)$ with the hyper-parameters a_k, b_k, and a, b for $k = 0, 1, 2$. According to Han, the prior hyper-parameters a_k, b_k should be selected such that $\pi(\lambda_k)$ is a decreasing function of λ_k. Also, the hyper-parameters a_k, b_k are assumed to be independent with the following density function -

$$\pi(a_k, b_k) = \pi(a_k)\pi(b_k)$$

Then, the E-Bayesian estimator of λ_k and α

$$\tilde{\lambda}_{kEB} = E(\lambda_k|\boldsymbol{x}) = \int\int_{D_k} \tilde{\lambda}_{kB}(a_k, b_k)\pi(a_k, b_k)da_k db_k$$

$$\tilde{\alpha}_{EB} = \int\int_{D_\alpha} \tilde{\alpha}_B(a, b)\pi(a, b)dadb$$

where $\tilde{\lambda}_{kB}$ and $\tilde{\alpha}_B$ are the Bayesian estimates of λ_k and α respectively as given in Subsection 2.1.1, D_k is the domain of a_k and b_k and D_α is the domain of a and b.

2.1.3 Hierarchical Bayesian estimation

In hierarchical Bayesian estimation, if the prior density function of λ_k is $\pi(\lambda_k) = \pi(\lambda_k|a_k, b_k)$, the corresponding hierarchical PDF of λ_k, $\pi_h(\lambda_k)$, can be obtained. Now, by Bayes' theorem, the hierarchical posterior PDF of λ_k is given by

$$\pi_h(\lambda_k|\boldsymbol{x}) = \frac{\pi_h(\lambda_k)L(\boldsymbol{\theta}|\boldsymbol{x})}{\int_0^\infty \pi_h(\lambda_k)L(\boldsymbol{\theta}|\boldsymbol{x})d\lambda_k}$$

The hierarchical Bayesian estimator of λ_k -

$$\tilde{\lambda}_{kHB} = \int_0^\infty \lambda_k \pi_h(\lambda_k|\boldsymbol{x})d\lambda_k$$

Similarly, for α -

$$\tilde{\alpha}_{HB} = \int_0^\infty \alpha(a,b)\pi_h(\alpha_k|\boldsymbol{x})d\alpha$$

2.1.4 HPD credible intervals

Given the observed data $\boldsymbol{x}$, let $\pi(\lambda|\boldsymbol{x})$ be the posterior density function of λ. Then, given a significance level γ, if there exists an interval $C_{\lambda|\boldsymbol{x}}$ satisfying the following conditions -

1. $P_{\lambda|\boldsymbol{x}}(C_{\lambda|\boldsymbol{x}}) > 1 - \gamma$,

2. For any $\lambda_1 \in C_{\lambda|\boldsymbol{x}}$ and $\lambda_2 \notin C_{\lambda|\boldsymbol{x}}$, $\pi(\lambda_1|\boldsymbol{x}) \geq \pi(\lambda_2|\boldsymbol{x})$ is true.

Then , $C_{\lambda|\boldsymbol{x}}$ is called the HPD credible interval of λ with the credible level 1 - γ. In many cases, the closed forms of the posterior distributions cannot be obtained. We then usually consider numerical methods to overcome this problem.

2.2 Model description

The assumption in our model is that the item can fail due to two risks. We have assumed two risks for the sake of simplicity, though, the same can be extended for more than two risks. Let T_i denote the *time to failure* of an item under the risk $i = 1, 2$. We can see that the *time to failure* of an item under a specific risk, *i.e,* T_i, is not observable. Only the lifetime of an item, which is $T = min(T_1, T_2)$ is observable. The cause of failure is recorded whenever an item fails. We will also assume that the item is put on the life test at time $t = 0$. For each item, there is a left truncation time τ_L and a right censoring time τ_R, which may vary with experimental units. If $T > \tau_L$, the information about the item is available. The lifetime of the item is exactly known if $\tau_l < T < \tau_R$. For the data which was discussed in the above chapter, τ_L will be the difference between 1975 and the installation time of the item, and τ_R will be the difference 2021 and the installation time. We use the following notations for the rest of the report.

n : Number of items put on the test

τ_{iL} : Left truncation time for i^{th} unit

τ_{iR} : Right censoring time of i^{th} unit

T_i : RV denoting the lifetime of i^{th} unit

y_i : Observed failure time of the i^{th} unit.

T_{ji} : RV denoting the failure time of i^{th} unit due to cause j, where $j = \{1, 2\}$

ν_i : Truncation indicator of i^{th} unit. It takes the value 0 if the i^{th} unit is truncated and 1 otherwise.

δ_i : Indicator variable as described below:

$$\delta_i = \begin{cases} 0 & \text{if } i^{th} \text{ item is censored} \\ 1 & \text{if the } i^{th} \text{ unit failed due to cause 1} \\ 2 & \text{if the } i^{th}\text{unit failed due to cause 2} \end{cases}$$

I_j : Set of indices of units which have failed due to cause $j \in \{1, 2\}$

I_0 : Set of indices of units which are censored

m_j : $|I_j|$, the cardinality of I_j

As discussed earlier, we have assumed that (T_{1i}, T_{2i}) follows a $BBBW\ (\alpha, \lambda_0, \lambda_1, \lambda_2)$. We also assume that $(T_{11}, T_{21}), (T_{12}, T_{22}), \ldots, (T_{1n}, T_{2n})$ are independent. As

$T_i = \min\{T_{1i}, T_{2i}\}$ for $i = 1, 2, \ldots, n; T_1, T_2, \ldots, T_n$ are also independent.

2.3 Likelihood Function

All the units are assumed to be put on the test at the time 0. The likelihood contributions of a observation (y_i, δ_i, ν_i) are listed below for different cases.

Case 1: For $\delta_i = 0$ and $\nu_i = 1, S_{X_1,X_2}(y_i, y_i)$.

Case 2: For $\delta_i = 1$ and $\nu_i = 1, cf_{\mathrm{WE}}(y_i; \alpha, \lambda_1) S_{\mathrm{WE}}(y_i; \alpha, \lambda_0 + \lambda_2)$.

Case 3: For $\delta_i = 2$ and $\nu_i = 1, cf_{\mathrm{WE}}(y_i; \alpha, \lambda_2) S_{\mathrm{WE}}(y_i; \alpha, \lambda_0 + \lambda_1)$.

Case 4: For $\delta_i = 0$ and $\nu_i = 0, \frac{S_{X_1,X_2}(y_i,y_i)}{S_{X_1,X_2}(\tau_{iL},\tau_{iL})}$

Case 5: For $\delta_i = 1$ and $\nu_i = 0, \frac{1}{S_{X_1,X_2(\tau_{iL}\tau_{iL})}} cf_{\mathrm{WE}}(y_i; \alpha, \lambda_1) S_{\mathrm{WE}}(y_i; \alpha, \lambda_0 + \lambda_2)$.

Case 6: For $\delta_i = 2$ and $\nu_i = 0, \frac{1}{S_{X_1,X_2(\tau_{iL},\tau_{iL})}} cf_{\mathrm{WE}}(y_i; \alpha, \lambda_2) S_{\mathrm{WE}}(y_i; \alpha, \lambda_0 + \lambda_1)$.

The cases 1 and 4 are trivial. Cases 2, 3 are similar and cases 5,6 are similar. Here, we have given the detail of cases 2 and 5 below.

Let l_2 and l_5 denote the likelihood contributions for the cases 2 and 5, respectively. Also note that $T_i = \min\{T_{1i}, T_{2i}\}$.

$$
\begin{aligned}
l_2 &= P\left(x_1 = y_i, x_2 > y_i\right) \\
&= \int_{y_i}^{\infty} f_{X_1,X_2}\left(y_i, t\right) dt \\
&= \int_{y_i}^{\infty} c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_1\right) f_{\mathrm{WE}}\left(t; \alpha, \lambda_0 + \lambda_2\right) dt \\
&= c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_1\right) \int_{y_i}^{\infty} f_{\mathrm{WE}}\left(t; \alpha, \lambda_0 + \lambda_2\right) dt \\
&= c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_1\right) S_{\mathrm{WE}}\left(y_i; \alpha, \lambda_0 + \lambda_2\right)
\end{aligned}
$$

$$
\begin{aligned}
&l_5 = P\left(x_1 = y_i, x_2 > y_i \mid x_1 > \tau_{iL}, x_2 > \tau_{iL}\right) \\
&= \frac{P\left(x_1 = y_i, x_2 > y_i\right)}{P\left(x_1 > \tau_{iL}, x_2 > \tau_{iL}\right)} \\
&= \frac{1}{S_{X_1,X_2}\left(\tau_{iL}, \tau_{iL}\right)} c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_1\right) S_{\mathrm{WE}}\left(y_i; \alpha, \lambda_0 + \lambda_2\right) \qquad \text{from case 3}
\end{aligned}
$$

Based on the above contributions, the likelihood function of the parameters $\alpha, \lambda_0, \lambda_1$, and λ_2 can be written as-

$$
\begin{aligned}
&L\left(\alpha, \lambda_0, \lambda_1, \lambda_2\right) \\
&= \prod_{i \in I_0} \left[S_{X_1.X_2}\left(y_i, y_i\right)\right]^{\nu_i} \left[\frac{S_{X_1.X_2}\left(y_i, y_i\right)}{S_{X_1,X_2}\left(\tau_{iL}, \tau_{iL}\right)}\right]^{(1-\nu_i)} \\
&\quad \times \prod_{i \in I_1} \left[c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_1\right) S_{\mathrm{WE}}\left(y_i; \alpha, \lambda_0 + \lambda_2\right)\right]^{\nu_i} \left[\frac{c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_1\right) S_{\mathrm{WE}}\left(y_i; \alpha, \lambda_0 + \lambda_2\right)}{S_{X_1,X_2}\left(\tau_{iL}, \tau_{iL}\right)}\right]^{(1-\nu_i)} \\
&\quad \times \prod_{i \in I_2} \left[c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_2\right) S_{\mathrm{WE}}\left(y_i; \alpha, \lambda_0 + \lambda_1\right)\right]^{\nu_i} \left[\frac{c f_{\mathrm{WE}}\left(y_i; \alpha, \lambda_2\right) S_{\mathrm{WE}}\left(y_i; \alpha, \lambda_0 + \lambda_1\right)}{S_{X_1,X_2}\left(\tau_{iL}, \tau_{iL}\right)}\right]^{(1-\nu_i)} \\
&= c^m \alpha^m \lambda_1^{m_1} \lambda_2^{m_2} \left(\prod_{i \in I_1 \cup I_2} y_i^{\alpha-1}\right) e^{-(\lambda_0+\lambda_1+\lambda_2)\Lambda(\alpha)}
\end{aligned}
$$

$$
\begin{aligned}
\therefore L(\alpha,\lambda_0,\lambda_1,\lambda_2) &= \left(\frac{\lambda_0+\lambda_1+\lambda_2}{\lambda_1+\lambda_2}\right)^m \alpha^m \lambda_1^{m_1}\lambda_2^{m_2}\left(\prod_{i\in I_1\cup I_2} y_i^{\alpha-1}\right) e^{-(\lambda_0+\lambda_1+\lambda_2)\Lambda(\alpha)} \\
&= \left(\frac{1+\frac{\lambda_1}{\lambda_0}+\frac{\lambda_2}{\lambda_0}}{\lambda_1+\lambda_2}\right)^m \alpha^m \lambda_0^m \lambda_1^{m_1}\lambda_2^{m_2}\left(\prod_{i\in I_1\cup I_2} y_i^{\alpha-1}\right) e^{-(\lambda_0+\lambda_1+\lambda_2)\Lambda(\alpha)} \\
&= h^m \alpha^m \lambda_0^m \lambda_1^{m_1}\lambda_2^{m_2}\left(\prod_{i\in I_1\cup I_2} y_i^{\alpha-1}\right) e^{-(\lambda_0+\lambda_1+\lambda_2)\Lambda(\alpha)}
\end{aligned}
$$

where $h = \left(\frac{1+\frac{\lambda_1}{\lambda_0}+\frac{\lambda_2}{\lambda_0}}{\lambda_1+\lambda_2}\right)$, $m = m_1+m_2$ and $\Lambda(\alpha) = \sum_{i=1}^n {y_i}^\alpha - \sum_{i=1}^n (1-\nu_i)\, {\tau_{iL}}^\alpha$

2.4 Illustration

2.4.1 Bayesian Estimation

The likelihood function given the observed data $\boldsymbol{x}$ is

$$L(\boldsymbol{\theta}|\boldsymbol{x}) = h^m \alpha^m \lambda_0^m \lambda_1^{m_1}\lambda_2^{m_2}\left(\prod_{i\in I_1\cup I_2} y_i^{\alpha-1}\right) e^{-(\lambda_0+\lambda_1+\lambda_2)\Lambda(\alpha)}$$

The conjugate prior distribution of λ_k, for $k = 0,1,2$ is considered to be Gamma distribution. We will denote the prior distribution by $\pi(\lambda_k) = Ga(\lambda_k; a_k, b_k)$. The PDF is given by

$$\pi(\lambda_k) = Ga(\lambda_k; a_k, b_k) = \frac{b_k^{a_k}}{\Gamma(a_k)}\lambda_k^{a_k-1} e^{-b_k c_k}$$

The prior distribution for α is $\pi(\alpha) \propto$ constant (Achcar *1995*). The joint posterior distribution of $\lambda_0, \lambda_1, \lambda_2, \alpha$ based on the priors, can be written as:

$$\pi(\lambda_0, \lambda_1, \lambda_2, \alpha|x) \propto h^m \times \frac{Ga(\lambda_0; a_0+m, b_0+\Lambda(\alpha))}{(b_0+\Lambda(\alpha))^{a_0+m}} \times \frac{Ga(\lambda_1; a_1+m_1, b_1+\Lambda(\alpha))}{(b_1+\Lambda(\alpha))^{a_1+m_1}}$$
$$\times \frac{Ga(\lambda_2; a_2+m_2, b_2+\Lambda(\alpha))}{(b_2+\Lambda(\alpha))^{a_2+m_2}} \times \alpha^m \times \left(\prod_{i\in I_1\cup I_2} y_i^{\alpha-1}\right)$$

To obtain the Bayesian estimators and HPD CrIs, the Metropolis Hastings algorithm can be used as follows:

Step I: Lets denote the joint posterior density as $l(\boldsymbol{\theta})$ which is the target distribution and we consider the following proposal distribution $q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)})$:

$$q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)}) = Tr(\lambda_0^{(k)}; \lambda_0^{(k-1)}, 1) \times Tr(\lambda_1^{(k)}; \lambda_1^{(k-1)}, 1) \times Tr(\lambda_2^{(k)}; \lambda_2^{(k-1)}, 1) \times Tr(\alpha^{(k)}; \alpha^{(k-1)}, 1)$$

where $\boldsymbol{\theta}^{(k)} = (\lambda_0^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)}, \alpha^{(k)})$, $Tr(x;\mu,\sigma)$ is the Truncated Normal distribution with mean μ, variance σ^2 with the support $x \in [0,\infty)$ and all the parameters are mutually independent.

Step II: Considering a starting point $\boldsymbol{\theta}^{(0)}$, we sample $\boldsymbol{\theta}^{(1)}$ from q, and accept it with the probability R, where:

$$R = min\left(1, \frac{l(\boldsymbol{\theta}^{(k)})q(\boldsymbol{\theta}^{(k-1)}|\boldsymbol{\theta}^{(k)})}{l(\boldsymbol{\theta}^{(k-1)})q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)})}\right)$$

If the sample is accepted, we similarly calculate $\boldsymbol{\theta}^{(2)}$. We repeat this step until we obtain $\boldsymbol{\theta}^{(N)}$.

Step III: From the above step, we have the ordered sequences $(\lambda_k^{(0)}, \lambda_k^{(1)}, \ldots, \lambda_k^{(N)})$ for $k = 0, 1, 2$ and $(\alpha^{(0)}, \alpha^{(1)}, \ldots, \alpha^{(N)})$.

Step IV: We reorder the sequences and denote the new sample by $(\lambda_k^{[0]}, \lambda_k^{[1]}, \ldots, \lambda_k^{[N]})$ for $k = 0, 1, 2$ and $(\alpha^{[0]}, \alpha^{[1]}, \ldots, \alpha^{[N]})$. The Bayesian estimators of the parameters under the squared loss function are given by:

$$\tilde{\lambda}_{kB} = \frac{1}{N}\sum_{i=1}^{N}\lambda_k^{[i]}$$

$$\tilde{\alpha} = \frac{1}{N}\sum_{i=1}^{N}\alpha^{[i]}$$

and the $100(1-\alpha)\%$ HPD CrIs of the unknown parameters and reliability are given by

$$\left[\tilde{\lambda}_k^{\left[i_k^*\right]}, \tilde{\lambda}_k^{\left[i_k^*+(1-\alpha)M\right]}\right], \left[\tilde{\alpha}^{[i^*]}, \tilde{\alpha}^{[i^*+(1-\alpha)M]}\right]$$

where i_k^*, i^* and j^* satisfy the following equations,

$$\tilde{\lambda}_k^{\left[i_k^*+(1-\alpha)M\right]} - \tilde{\lambda}_k^{\left[i_k^*\right]} = \min_i\left\{\tilde{\lambda}_k^{[i+(1-\alpha)M]} - \tilde{\lambda}_k^{[i]}\right\}$$
$$\tilde{\alpha}^{[i^*+(1-\alpha)M]} - \tilde{\alpha}^{[i^*]} = \min_i\left\{\tilde{\alpha}^{[i+(1-\alpha)M]} - \tilde{\alpha}^{[i]}\right\}$$

respectively, for $k = 0, 1, 2; 1 \leq i \leq M - (1-\alpha)M$.

In the following part, E-Bayesian and hierarchical Bayesian methods are adopted to analyze the hyper-parameters a_k and $b_k, k = 0, 1, 2$. As the prior distribution of α is proportional to a constant, here, we replace α by its Bayesian estimation $\tilde{\alpha}_B$. Then, the estimators of λ_k under squared loss function are deduced by using different Bayesian methods.

2.4.2 E-Bayesian Estimation

The E-Bayesian Estimation of λ_k can be esimated as:

$$\tilde{\lambda}_{kEB} = E(\lambda_k|\boldsymbol{x}) = \int\int_{D_k}\tilde{\lambda}_{kB}(a_k, b_k)\pi(a_k, b_k)da_k db_k$$

As the above integral cannot be computed explicitly, we will use Monte Carlo method for approximating the the E-Bayesian estimate. Assuming the hyper-parameters a_k, b_k are independent and considering the following prior distribution of a_k and b_k:

$$\pi(a_k, b_k) = \pi(a_k)\pi(b_k)$$

where $\pi(a_k), \pi(b_k) \sim Ga(c, d)$.

The E-Bayesian estimate can be calculated as follows:

Step I: For $i = 1, 2, \ldots, N$:

- Generate $a_k^i, b_k^i \sim Ga(c, d)$
- Calculate the Bayesian estimate $\tilde{\lambda}_{kB}^i$ considering $\pi(\lambda_k^i) = Ga(\lambda_k^i; a_k^i, b_k^i)$ as done in the above subsection with the following proposal distribution:

$$q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)}) = Tr(\lambda_0^{(k)}; \lambda_0^{(k-1)}, 1) \times Tr(\lambda_1^{(k)}; \lambda_1^{(k-1)}, 1) \times Tr(\lambda_2^{(k)}; \lambda_2^{(k-1)}, 1)$$

where $\boldsymbol{\theta}^{(k)} = (\lambda_0^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)})$, $Tr(x; \mu, \sigma)$ is the Truncated Normal distribution with mean μ, variance σ^2 with the support $x \in [0, \infty)$ and all the parameters are mutually independent.

Step II: The E-Bayesian estimates of the parameters are:

$$\tilde{\lambda}_{kEB} = \frac{1}{N} \sum_{i=i}^{N} \tilde{\lambda}_{kB}^i$$

2.4.3 Hierarchical Bayesian Estimation

According to Lindley and Smith (1972), if a_k and b_k are hyperparameters of λ_k, the prior density function of λ_k is $\pi(\lambda_k)$ as mentioned above and the joint prior density function of hyper-parameters a_k and b_k is

$$\pi(a_k, b_k) = \pi(a_k)\pi(b_k) = 1/p_k, \quad 0 < a_k < 1, 0 < b_k < p_k$$

then the corresponding hierarchical PDF of λ_k is given by

$$\pi_h\left(\lambda_k\right)=\int_0^1\int_0^{p_k}\pi\left(\lambda_k\right)\pi\left(a_k,b_k\right)db_k da_k=\int_0^1\frac{\gamma\left(a_k+1,p_k\right)}{\Gamma\left(a_k\right)\lambda_k^2 p_k}da_k$$

$$\therefore \pi_h\left(\lambda_k\right)\propto 1/\lambda_k^2$$

Therefore, the hierarchical posterior of λ_k is

$$\pi_h(\lambda_0,\lambda_1,\lambda_2|\boldsymbol{x},\tilde{\alpha_B})\propto h^m\times Ga(\lambda_0;m-1,\Lambda(\tilde{\alpha_B}))\times Ga(\lambda_1;m_1-1,\Lambda(\tilde{\alpha_B}))\times Ga(\lambda_2;m_2-1,\Lambda(\tilde{\alpha_B}))$$

To calculate the hierarchical Bayesian estimates, we will use the Metropolis Hastings algorithm as done in 2.4.1 with the following proposal distribution $q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)})$:

$$q(\boldsymbol{\theta}^{(k)}|\boldsymbol{\theta}^{(k-1)})=Tr(\lambda_0^{(k)};\lambda_0^{(k-1)},1)\times Tr(\lambda_1^{(k)};\lambda_1^{(k-1)},1)\times Tr(\lambda_2^{(k)};\lambda_2^{(k-1)},1)$$

where $\boldsymbol{\theta}^{(k)}=(\lambda_0^{(k)},\lambda_1^{(k)},\lambda_2^{(k)})$, $Tr(x;\mu,\sigma)$ is the Truncated Normal distribution with mean μ, variance σ^2 with the support $x\in[0,\infty)$ and all the parameters are mutually independent.

Chapter 3

Simulation Study and Data Analysis

3.1 Simulation Study

We performed extensive simulation to check the effectiveness of the method proposed in the previous chapter. The results are as follows.

For simulation study, we have fixed 1975 as the left truncation time and 2021 as right censoring time. First, we divide the installation year into two sets, one from 1960 to 1974 and another 1975 to 1999 . These sets correspond to truncated and non-truncated lifetimes respectively. Next, we fix the truncation percentage. For fixed truncation percentage, installation years of machines are accordingly sampled from these sets as per the following probability assignment. For each year between 1960 to 1974 are assigned probability uniformly . Similarly, the probability is assigned to every year between 1975 to 1999 uniformly. The lifetimes of the machines are then generated from BBBW distribution and added to installation year of the unit to obtain the date of failure of the unit.

Algorithm to generate from BBBW$(\lambda_0, \lambda_1, \lambda_2, \alpha)$:

- Step 1: Generate $U_0 \sim \mathrm{WE}(\alpha, \lambda_0)$, $U_1 \sim \mathrm{WE}(\alpha, \lambda_1)$ and $U_2 \sim \mathrm{WE}(\alpha, \lambda_2)$ independently.

- Step 2: If $U_0 < U_1$ and $U_0 < U_2$ go back to Step 1 otherwise set $T_1 = \min\{U_0, U_1\}$ and $T_2 = \min\{U_0, U_2\}$.

Now, T_1, T_2 correspond to the lifetime of the unit under two different competing risks. Finally the generated lifetime of the unit is given by $T = \min\{T_1, T_2\}$. At this point, the cause of failure of the unit (which is minimum among T_1 and T_2) is also recorded.

Note that the year of left truncation is 1975. Hence if any unit fails before 1975, we would not even know about the unit. Such cases are discarded and a new lifetime and hence failure year are generated for the unit. Also note that 2021 is considered as the right censoring time. For any unit, if the failure time turn out to be more than 2021 , it is considered as right censored item.

For simulation, we consider the following two sets of parameters: $(\alpha, \lambda_0, \lambda_1, \lambda_2) = (1.5, 1.2, 0.5, 1.5)$ and $(2, 10, 8, 12)$. For Bayesian estimation, we take $a_k, b_k = 0.01$ for $k = 0, 1, 2$ and $c, d = 1, 0.5$ for E-Bayesian estimation.

The sample size is $n = 100$ and the truncation percentage is fixed at 20%. The Bayesian estimators (BEs), MSEs, the 95% CPs and the ALs of the HPD CrIs of the unknown parameters are shown in Table 3.1 and 3.2.

Table 3.1: n = 100, Truncation percentage = 20%

Parameter	True Value	BE	MSE	AL	CP
λ_0	1.2	1.0838	0.2179	3.334	99.8
λ_1	0.5	0.6831	0.1356	1.501	99.5
λ_2	1.5	2.013	0.8493	3.9322	99.7
α	2	2.05142	0.1027	1.2694	96.5

Table 3.2: n = 100, Truncation percentage = 20%

Parameter	True Value	BE	MSE	AL	CP
λ_0	10	8.89479	40.9799	16.3115	92
λ_1	8	8.42608	12.0942	12.2892	94.5
λ_2	12	12.81379	26.21231	15.97365	93.6
α	2	1.89226	0.03449	0.59061	94

3.2 Data Analysis

We have also analyzed a real life data set. The data set has been obtained from the R package '**etm**'. It represents the outcome of pregnancies which ended in spontaneous abortion or induced abortion, resulting in a competing risks situation . Also, the data is left truncated as women usually enter the study some weeks after conception.

Since, we do not have any information on the hyper-parameter values, we have considered the Bayesian estimation of the parameters, $\alpha, \lambda_0, \lambda_1, \lambda_2$ under non-informative priors. The results are shown below:

Table 3.3: Real life data set

Parameter	Bayesian	E-Bayesian	H-Bayesian	HPD
λ_0	5.26119	6.5497	6.3523	[0.1759, 17.6747]
λ_1	4.36491	3.4533	4.23759	[0.0105, 8.8655]
λ_2	15.05745	13.667	14.7238	[0.3284, 22.9647]
α	0.79659	0.79659	0.79659	[0.2, 1.3061]

Chapter 4

Conclusion

We have considered a left-truncated right-censored model in the presence of dependent competing risks in this project. The latent failure times are assumed to follow Block & Basu Weibull distribution and we have used Bayesian estimation of the unknown parameters. Since, Bayesian estimation also depends upon the choice of the prior distribution of the parameters, we have also considered methods such as E-Bayesian and Hierarchical Bayesian to overcome this. We observed that the Bayes estimate does not exist in closed form. So, we have obtained the Bayesian estimates under the square loss function and the HPD credible Intervals using the Metropolis-Hastings algorithm. We have also done a simulation study to judge the performance of the estimation methods proposed. It can be seen that the results are satisfactory with different values of parameters.The algorithms used are computationally heavy, so we recommend to use it for small sample sizes.

Bibliography

[1] Yili Hong, William Q. Meeker, and James D. McCalley. Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *The Annals of Applied Statistics*, 3(2):857 – 879, 2009.

[2] D. Kundu, D. Mitra, and A. Ganguly. Analysis of left truncated right censored competing risks data. *Computational Statistics and Data Analysis*, 108:12–26, 2017.

[3] Xuchao Bai, Yimin Shi, and Hon Keung Tony Ng. Statistical inference of type-i progressively censored step-stress accelerated life test with dependent competing risks. *Communications in Statistics - Theory and Methods*, 0(0):1–27, 2020.

[4] Debasis Kundu and Rameshwar D. Gupta. A class of absolutely continuous bivariate distributions. *Statistical Methodology*, 7:464–477, 2010.

[5] Jorge alberto Achcar. Inferences for accelerated life tests considering a bivariate exponential distribution. *Statistics*, 26:269–283, 1995.

[6] D V Lindley and A F M Smith. Bayes estimates for the linear model. *J. R. Stat. Soc.*, 34(1):1–18, 1972.