BAYES ANALYSIS OF A LIFETIME MODEL

by

Esmail Furjani

Submitted in partial fulfillment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
March 2017

Table of Contents

iii

List of Tables

List of Figures

Abstract

Lifetimes models have been developed in the last five centuries. These developments are purposed to fit lifetimes data. Models that accommodate different shapes of the hazard rate function are useful for analyzing the lifetimes data. Among these lifetimes models is the TN distribution with two parameters. The TN model has variety of shapes for hazard rate function. It accommodates increasing, decreasing, bathtub, unimodal, increasing-decreasing-increasing hazard shapes which allows it to fit variety of real lifetime data sets. The main aim of this thesis is to estimate the two parameters of the TN distribution using the classical and Bayesian methods. In the classical approach, we use the maximum likelihood estimation. While in the Bayesian approach, the rejection sampling algorithm and the Markov Chain Monte Carlo methods are applied to obtain the Bayes estimates of the TN parameters and the tow-sided Bayesian probability intervals of the parameters. Simulation study is performed to investigate the properties of the methods applied and compare the maximum likelihood and the Bayesian methods. In order to demonstrate the use of the methods used in this thesis, three data sets are analyzed using the maximum likelihood and Bayes methods.

List of Abbreviations and Symbols Used

α	Parameter of the TN distribution
β	Parameter of the TN distribution
θ	Vector's Parameters
$\mathcal{L}$	The log likelihood
$\hat{\alpha}$	Estimated α
$\hat{\beta}$	Estimated β
$\hat{\theta}$	Estimated θ
n	Sample Size
Pdf	Probability density function
RS	Rejection Sampling
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation
MSE	Mean Square Error

Acknowledgements

I would like to thank my great supervisor, Dr. Ammar Sarhan for his guidance and constant support over the past two years. It is an honour to me to be one of his students. I learn a lot from him, and I appreciate his time and ideas. The excellent instructions and valuable advices that encourage me to do this thesis. Also, I would like to thanks, Dr. Bruce Smith for his kind advices and support during my graduate study. I would like to give my thanks to, Dr. Michael Dowd for reading my thesis and providing kind suggestions for my thesis. My special thanks to every member of the Department of Statistics and Mathematical for their help and support. I would like to thanks my classmates and my professors in my Department. Finally, My thanks to my family who have supported me through my study.

Chapter 1

Introduction

The lifetime of humans, components, etc, are unanticipated and non-negative random variables, and amendable to statistical rules. In the mid half of the 20 century, there are many developments of models that deal with such kind of random variables [5]. These models are known as lifetime models since the main interest of these models is the lifetime. For instances, it can be the number of hours for a light bulb before it breaks down, it could be the period of time of the patients after they recovered from an operation, and it could be the number of cycles which are recorded for new product before the failure time. In lifetime data analysis, models with different shapes of hazard rate function are useful. As an example of common models that often being used to analyze lifetime data, are the exponential and Weibull distributions. However, the Weibull distribution has monotone hazard rate function whereas the hazard rate function of the exponential distribution is a constant. These two distributions are unable to accommodate the lifetime data sets that have non-monotone hazard rate shapes. For this reason, many lifetime models were introduced to address this lack. The generalized linear failure rate distribution was introduced by Sarhan and Kundu [10]. Gupta and Kundu suggested the generalized exponential distribution [7]. Pham and Lai [9] presented some of the lifetime distributions. The exponentiated modified Weibull extension was presented by Sarhan and Apaloo [11]. Sarhan et al. [12], proposed a new distribution and they referred it as the TN distribution, and it has two parameters which makes it as one of a few distributions that have only two parameters and shows different shapes of the hazard rate function. They discussed some statistical properties of the TN model such as skewness and kurtosis and proposed some algorithms that generate random samples from the TN distribution. Also, they used the maximum likelihood method to estimate its two unknown parameters. In this thesis, the Bayesian techniques will be applied to estimate the TN distribution parameters and compare it to the their maximum likelihood estimation.

1.1   The TN Distribution

Sarhan et al. [12] proposed a one parameter distribution and named it the new distribution and referred as $N(\beta)$ with the probability density function (pdf) given by

$$f(t) = \frac{\beta}{1+\beta}[\beta + (1+2\beta t)e^{-\beta t}]e^{-\beta t}, \ t \geq 0, \beta > 0. \tag{1.1}$$

The hazard rate function of the $N(\beta)$ is

$$h(t) = \frac{\beta[\beta + (1+2\beta x)e^{-\beta t}]}{\beta + (1+\beta t)e^{-\beta t}}, \ t \geq 0, \beta > 0. \tag{1.2}$$

The hazard rate function of the $N(\beta)$ is unimodal which makes the $N(\beta)$, in some cases, unable to fit some real data sets. Therefore, Sarhan et. al [12] used the power transformation $X = T^{\frac{1}{\alpha}}$, where $T \sim N(\beta)$ and $\alpha > 0$, to propose a new model, named transformed new distribution and they referred it as the TN (α, β) distribution. The probability density function of the $TN(\alpha, \beta)$ is

$$f(x) = \frac{\alpha\beta x^{\alpha-1}}{1+\beta}[\beta + (1+2\beta x^{\alpha})e^{-\beta x^{\alpha}}]e^{-\beta x^{\alpha}}, x \geq 0, \alpha, \beta > 0. \tag{1.3}$$

Figure 1.1 demonstrates the probability density function for different choices of α when $\beta = 1$. The probability density function can be skewed or symmetric. The hazard rate function of the $TN(\alpha, \beta)$ is

$$h(x) = \frac{\alpha\beta x^{\alpha-1}[\beta + (1+2\beta x^{\alpha})e^{-\beta x^{\alpha}}]}{\beta + (1+\beta x^{\alpha})e^{-\beta x^{\alpha}}}, x \geq 0, \alpha, \beta > 0. \tag{1.4}$$

Sarhan et al. [12] showed that the hazard rate function of the TN model may take different shapes, such as increasing, decreasing, unimodal, decreasing-increasing-decreasing or increasing-decreasing-increasing depending on the values of its parameters α and β. Figure 1.2 displays the hazard rate function of the $TN(\alpha, \beta)$ for different parameters values. The variety of the shapes of the hazard rate function of the $TN(\alpha, \beta)$ makes it a suitable model to fit many different real data sets.

1.2   Aim of Study

The main goal of the thesis is to discuss the Bayes inference of the TN distribution's parameters by using two different approaches. The first approach is the rejection
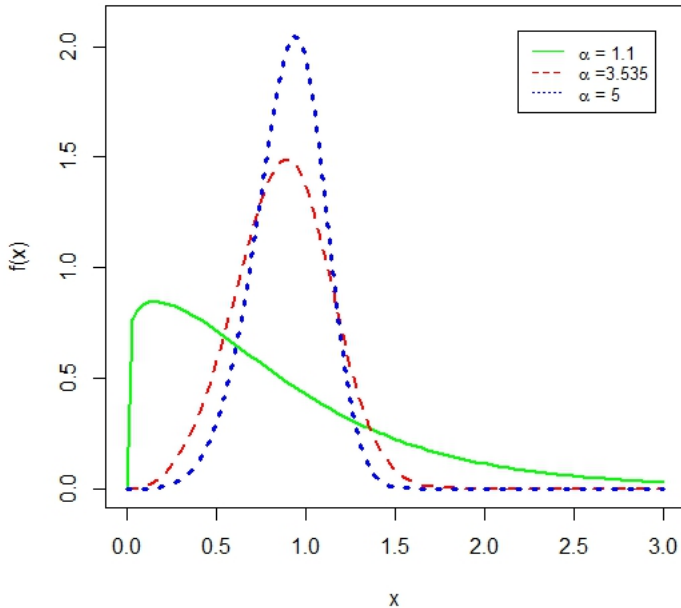
Figure 1.1: The density probability function of the TN (α, β) distribution with some values of α when $\beta = 1$.

sampling algorithm (RS) and the second method is the Markov Chain-Monte Carlo (MCMC). The maximum likelihood method is considered for comparison reason. We will use the model to analyze three different real data sets. Simulation study is performed to compare the Bayesian results with the maximum likelihood results.

1.3   The Structure of the Thesis

The rest of this thesis consists of four chapters. The next Chapter describes two estimation methods for the unknown parameters which are the classical and Bayesian methods. In the classical approach, maximum likelihood estimation is considered while in the Bayesian method, we study the rejection sampling algorithm and the Markov Chain Monte Carlo method. Simulation study is carried out using the three estimation methods and the results are presented in Chapter 3. In Chapter 4, three real data sets are analyzed using the $TN(\alpha, \beta)$ distribution. Chapter 5 concludes the thesis and proposes a possible future work.
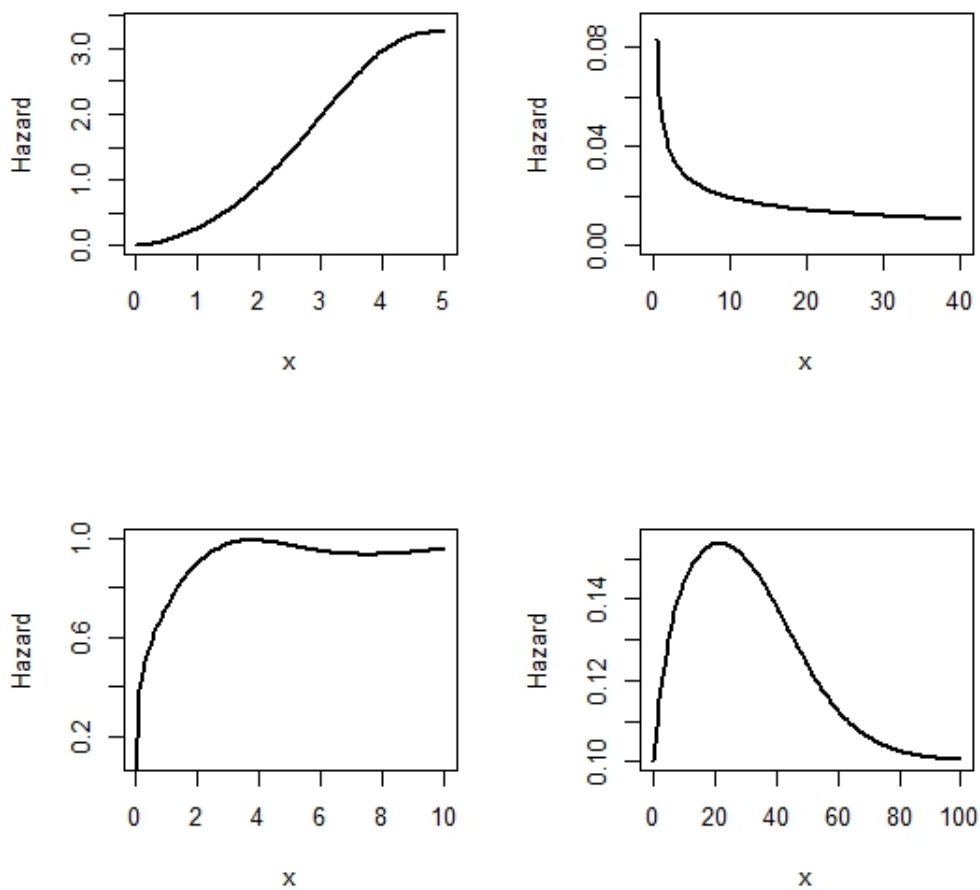
Figure 1.2: The hazard function of TN (α, β) with some different parameters values, $\alpha = 2.5$, $\beta = 0.1$ (top left), $\alpha = 0.5$, $\beta = 0.1$ (top right), $\alpha = 1.2$, $\beta = 0.5$ (bottom left), $\alpha = 1$, $\beta = 0.1$ (bottom right).

Chapter 2

Methods of Estimation

Let us assume that $x_1, \ldots, x_n$ be an independent and identically distributed (iid) random sample from the $TN(\alpha, \beta)$ distribution with a probability density function (1.3). Our main goal is to do statistical inference on the two unknown parameters included in the TN model. We will apply two approaches, classical and Bayesian methods, to obtain the point and interval estimates of the two parameters. First we briefly discuss the maximum likelihood approach, then the two Bayesian approaches "the Rejection Sampling (RS)" and "the Markov Chain Monte Carlo (MCMC)" will be discussed.

2.1 Maximum Likelihood Method

Maximum likelihood method is a popular and long-standing approach for parameter estimation. Suppose that n independent and identical items are put on the life test. The test will be terminated when all items have failed. The life of these items, say $x_1, \ldots, x_n$, are observed and they follow the TN (α, β).

The likelihood function, L, for α and β, give the data $x = (x_1, x_2, \cdots, x_n)$, can be obtained as follows

$$L(\alpha, \beta | x) = \prod_{i=1}^{n} f(x_i | \alpha, \beta)$$

Substituting (1.3) into the above equation, we get

$$L(\alpha, \beta | x) = \frac{\alpha^n \beta^n}{(1 + \beta)^n} [P(x)]^{\alpha - 1} e^{-\beta \tau(\alpha, x)} \times \prod_{i=1}^{n} [\beta + (1 + 2\beta x_i^\alpha) e^{-\beta x_i^\alpha}], \qquad (2.1)$$

where

$$P(x) = \prod_{i=1}^{n} x_i \text{ and } \tau(\alpha, x) = \sum_{i=1}^{n} x_i^\alpha.$$

Definition 2.1

Give a set of random sample $x_1, x_2, \cdots, x_n$ from a distribution indexed by a vector of unknown parameters $\theta = (\theta_1, \theta_2, \cdots, \theta_k)$, where $k \geq 1$, the maximum likelihood estimator for θ, say $\hat{\theta}(x)$, is the value of θ that maximizes the likelihood function of θ (Casella, 1990).

Under a regular set of conditions, among them the likelihood function must be differentiable in θ, the common approach of maximizing the likelihood function can be done by solving

$$\frac{\partial}{\partial \theta_i} L(\theta|x) = 0, \quad i = 1, \cdots, k.$$

To simplify differentiation, the log of the likelihood function is used. Therefore, we will need to solve the following system of equations in θ_i, $i = 1, 2, \cdots, k$

$$\frac{\partial \mathcal{L}}{\partial \theta_i} = 0, \quad i = 1, \cdots, k,$$

where $\mathcal{L} = \ln L(\theta|x)$. The above k equations are called the likelihood equations that should be solved to obtain the maximum likelihood estimators of the parameters.

The log-likelihood function, $\mathcal{L}$, for the model discussed here is

$$
\begin{aligned}
\mathcal{L} =\ & n \ln \alpha + n \ln \beta - n \ln(1 + \beta) + (\alpha - 1) \sum_{i=1}^{n} \ln x_i \\
& - \beta \sum_{i=1}^{n} x_i^{\alpha} + \sum_{i=1}^{n} \ln\left[\beta + (1 + 2\beta x_i^{\alpha})e^{-\beta x_i^{\alpha}}\right].
\end{aligned}
\tag{2.2}
$$

The first partial derivatives of $\mathcal{L}$, with respect to α and β, are

$$
\frac{\partial \mathcal{L}}{\partial \alpha} = \frac{n}{\alpha} + \ln[P(x)] - \beta\tau(\alpha, x) \ln x_i + \sum_{i=1}^{n} \frac{C_{i,\alpha}(\alpha, \beta)}{C_i(\alpha, \beta)},
$$

$$
\frac{\partial \mathcal{L}}{\partial \beta} = \frac{n}{\beta(1 + \beta)} - \tau(\alpha, x) + \frac{C_{i,\beta}(\alpha, \beta)}{C_i(\alpha, \beta)},
\tag{2.3}
$$

where

$$
C_{i,\alpha}(\alpha, \beta) = \frac{\partial C_i(\alpha, \beta)}{\partial \alpha} = \beta x_i^{\alpha}(1 - 2\beta x_i^{\alpha}) \ln x_i e^{-\beta x_i^{\alpha}},
$$

$$C_{i,\beta}(\alpha, \beta) = \frac{\partial C_i(\alpha, \beta)}{\partial \beta} = 1 + (1 - 2\beta x_i^\alpha)x_i^\alpha e^{-\beta x_i^\alpha}.$$

The second partial derivatives of $\mathcal{L}$ are

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha^2} = -\frac{n}{\alpha^2} - \beta \sum_{i=1}^n x_i^\alpha (\ln x_i)^2 + \sum_{i=1}^n \frac{C_i(\alpha, \beta)C_{i,\alpha^2}(\alpha, \beta) - [C_{i,\alpha}(\alpha, \beta)]^2}{[C_i(\alpha, \beta)]^2},$$

$$\frac{\partial^2 \mathcal{L}}{\partial \alpha \beta} = -\sum_{i=1}^n x_i^\alpha \ln x_i + \sum_{i=1}^n \frac{C_i(\alpha, \beta)C_{i,\alpha\beta}(\alpha, \beta) - C_{i,\alpha}(\alpha, \beta)C_{i,\beta}(\alpha, \beta)}{[C_i(\alpha, \beta)]^2},$$

$$\frac{\partial^2 \mathcal{L}}{\partial \beta^2} = -\frac{n}{\beta^2} + \frac{n}{(1+\beta)^2} + \sum_i \frac{C_i(\alpha, \beta)C_{i,\beta^2}(\alpha, \beta) - [C_{i,\beta}(\alpha, \beta)]^2}{[C_i(\alpha, \beta)]^2}, \qquad (2.4)$$

where

$$C_{i,\alpha^2}(\alpha, \beta) = \beta x_i^\alpha (\ln x_i)^2 e^{-\beta x_i^\alpha}(1 - 5\beta x_i^\alpha + 2\beta^2 x_i^{2\alpha}),$$

$$C_{i,\alpha\beta}(\alpha, \beta) = x_i^\alpha (\ln x_i)e^{-\beta x_i^\alpha}(1 - 5\beta x_i^\alpha + 2\beta^2 x_i^{2\alpha}),$$

$$C_{i,\beta^2}(\alpha, \beta) = -x_i^{2\alpha}e^{-\beta x_i^\alpha}(3 - 2\beta x_i^\alpha).$$

The Observed Information Matrix

The second partial derivatives of the log-likelihood function provide the elements of the observed information matrix. That is,

$$I(\hat{\alpha}_{MLE}, \hat{\beta}_{MLE}) = - \begin{bmatrix} \frac{\partial^2 \mathcal{L}}{\partial \alpha^2} & \frac{\partial^2 \mathcal{L}}{\partial \alpha \beta} \\ \frac{\partial^2 \mathcal{L}}{\partial \alpha \beta} & \frac{\partial^2 \mathcal{L}}{\partial \beta^2} \end{bmatrix}_{(\alpha, \beta) = (\hat{\alpha}_{MLE}, \hat{\beta}_{MLE})} \qquad (2.5)$$

where $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$ are the MLE of the parameters α and β which are the solution of the likelihood equations (2.4) such that $I(\hat{\alpha}_{MLE}, \hat{\beta}_{MLE})$ is positive definite.

There is no analytic solution for the system of the likelihood equations, hence numerical approaches, such as the Newton-Raphson, have to be used.

Asymptotic Confidence Intervals

Given that the MLEs of the parameters α and β do not exist here in closed form, some difficulty arises in obtaining the exact distribution for $\hat{\alpha}_{MLE}$ and $\hat{\beta}_{MLE}$. To obtain the asymptotic confidence interval, we use the result that states that $(\hat{\alpha}_{mle}$

, $\hat{\beta}_{mle}$) are asymptotically normally distributed with mean equal to the true values of the two parameters α and β and variance-covariance equals to the inverse of the observed information matrix. That is,

$$\begin{pmatrix} \hat{\alpha}_{MLE} \\ \hat{\beta}_{MLE} \end{pmatrix} \sim N \left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}, [I(\hat{\alpha}_{MLE}, \hat{\beta}_{MLE})]^{-1} \right), \qquad (2.6)$$

where $[I(\hat{\alpha}_{MLE}, \hat{\beta}_{MLE})]^{-1}$ is the inverse of $I(\hat{\alpha}_{MLE}, \hat{\beta}_{MLE})$.

The result (2.6) can be used to approximate the $(1-\vartheta)100\ \%$ confidence intervals for the parameters α and β as

$$\hat{\alpha}_{MLE} \pm Z_{\frac{\vartheta}{2}} \sqrt{Var(\hat{\alpha})}, \hat{\beta}_{MLE} \pm Z_{\frac{\vartheta}{2}} \sqrt{Var(\hat{\beta})}, \qquad (2.7)$$

where $Z_{\frac{\vartheta}{2}}$ is the $(1-\frac{\vartheta}{2})100^{th}$ percentile of the standard normal distribution.

2.2 Bayesian Estimation

In this section, we use Bayesian estimation methods. Bayesian estimation is a well-known and respected approach. The significant elements of the Bayesian analysis are the likelihood function, that represents the information about the parameters included in the data, and the prior density, which reflects the knowledge about the parameters before the data are observed. Combining the prior density and the likelihood function produces the posterior distribution. The posterior distribution reflects knowledge about the parameters from the data and the prior information. The posterior distribution can be summarized by the mean, the mode and the credible intervals.

2.2.1 Bayes' Theorem

Suppose that the data $x = (x_1, \ldots, x_n)$ are independent. In order to get the posterior distribution $g(\theta|x)$, we apply the basic property of the Bayes' rule:

$$g(\theta|x) = \frac{L(x|\theta)q(\theta)}{p(x)} \propto L(x|\theta)q(\theta),$$

where

θ : is a vector parameters of interest.

$L(x|\theta)$: is a the likelihood function.

$q(\theta)$: is a prior distribution.

$p(x)$: is a marginal distribution of x.

It is easy to obtain the posterior distribution which is the likelihood function multiplied by the prior density and determine the constant part that does not depend on θ which impose the constant term to integrate to 1. A good technique for calculating the posterior density is to exclude any constant from likelihood function and the prior density that does not depend on θ, and then compute the normalizing constant.

Prior Distribution

After the data model is selected, a Bayesian estimation needs the assertion of a prior density for unknown distribution parameters. The prior density represents the current state of uncertainty or knowledge about the parameters prior to data being observed. There are essentially two main methods for choosing the prior density. The first method is an informative prior density [6]. In this approach, the researcher applies his or her knowledge about the essential issue based on a previous data, associated with expert suggestion if possible, to construct the prior density which reflects his thinking about the unknown parameters. The second approach is a noninformative prior density that does not require prior experience on the parameters. Also, this kind of density is called diffuse or vague prior density. Selecting a noninformative prior density is trying to work as if there is no knowledge about the parameters before recognizing the data. This can be performed by specifying the same probability for every value of the parameter, or at least it is close to the same probability on extent of the parameters. However, the noninformative prior densities have problems. First, some common approaches that construct the noninformative prior density supposing a uniform distribution, produce inconsistency. Second, there are many cases where formulating the noninformative prior density produce a probability function which integrates to infinity . Generally, the prior density has an impact in small sample; however, this is not disadvantage since the posterior distribution represents the low available information about the parameters when the sample size is a small. A technique that can be used for a distribution with multiple parameters is to assume for each parameter individually and construct the joint prior density. In this thesis, we assume the prior density for α and β are independent and following the gamma prior

density,

$$q_1(\alpha) \propto \alpha^{a_1-1} e^{-b_1\alpha}, \alpha > 0.$$

$$q_2(\beta) \propto \beta^{a_2-1} e^{-b_2\beta}, \beta > 0.$$

By combining the two prior densities for α and β, we can write the joint prior distribution as follows :

$$q(\alpha, \beta) \propto \alpha^{a_1-1} \beta^{a_2-1} e^{-b_1\alpha-b_2\beta}, \alpha, \beta > 0. \tag{2.8}$$

The hyperparameters a_1, a_2, b_1, b_2 are supposed to be known and non-negative.

Posterior Distribution of the TN Model

In order to learn about parameters of a distribution, we need the posterior density. In this study, the posterior density of (α, β) can be formulated as a product of the likelihood function and the joint gamma prior density (2.1), (2.8) respectively, up to normalized constant,

$$g(\alpha, \beta | x) \propto \frac{\alpha^n \beta^n}{(1+\beta)^n} [P(x)]^{\alpha-1} e^{-\beta\tau(\alpha,x)} \times \prod_{i=1}^{n} [\beta + (1 + 2\beta x_i^\alpha) e^{-\beta x_i^\alpha}]$$

$$\times \alpha^{a_1-1} \beta^{a_2-1} e^{-b_1\alpha-b_2\beta}, \alpha, \beta > 0.$$

The posterior density of (α, β) can be written as

$$g(\alpha, \beta | x) \propto \frac{\alpha^{a_1+n-1} \beta^{a_2+n-1}}{(1+\beta)^n} [P(x)]^{\alpha-1} e^{-\beta\tau(\alpha,x)-b_1\alpha-b_2\beta}$$

$$\times \prod_{i=1}^{n} [\beta + (1 + 2\beta x_i^\alpha) e^{-\beta x_i^\alpha}], \alpha, \beta > 0. \tag{2.9}$$

The normalizing constant is

$$p(x) = \int_0^\infty \int_0^\infty \frac{\alpha^{a_1+n-1} \beta^{a_2+n-1}}{(1+\beta)^n} [P(x)]^{\alpha-1} e^{-\beta\tau(\alpha,x)-b_1\alpha-b_2\beta}$$

$$\times \prod_{i=1}^{n} [\beta + (1 + 2\beta x_i^\alpha) e^{-\beta x_i^\alpha}] \, d\alpha \, d\beta. \tag{2.10}$$

There is no analytic solution for (2.10). Therefore, we should use numerical techniques to calculate the posterior distribution of $\theta = (\alpha, \beta)$. Among those methods, we will use sampling methods to draw samples from the posterior distribution without knowing the normalizing constant. We will use

1. Rejection sampling algorithm.

2. Markov Chain Monte Carlo.

2.3 Simulations Methods

Since we can not generate samples from the posterior distribution due to it is hard to compute the normalizing constant of the posterior distribution of $\theta = (\alpha, \beta)$. Therefore, we need to use simulation approaches to get draws form the posterior distribution. There are many simulation methods that can be applied to obtain draws from the posterior distribution, and we consider two methods which are the rejection sampling algorithm and the Markov Chain Monte Carlo.

2.3.1 Rejection Sampling

The rejection sampling (RS) is a straight-forward and practical approach to gain simulated samples from a distribution when the normalizing constant can not be obtained [2]. Suppose that we need to find simulated draws from the posterior density of $\theta = (\alpha, \beta)$ which the normalizing constant can not be known. In such an instance, implementing rejection sampling would be ideal. To apply the rejection sampling algorithm, the first step is to locate a proposal density $q(\theta)$ that is simple to derive samples from; and should also resemble the posterior density of $g(\theta|x)$ and constant K that satisfies $g(\theta|x) \leq Kq(\theta)$, for all θ ,where $K > 1$.

In stage two of the algorithm, nominee samples are derived from the proposal density $q(\theta)$. Hence, the probability of an acceptance is

$$r = g(\theta|x)/Kq(\theta).$$

We continue with this process of drawing samples from the proposal density $q(\theta)$ until we achieve the acceptance rate. It is worth noting here that the higher the acceptance rate is, the more accurate and efficient is the rejection sampling algorithm [2]. Keep in mind that the posterior density of $\theta = (\alpha, \beta)$ in (2.9) is our target density. To that end, we choose the proposal density as the multivariate t-distribution with mean and number of degrees of freedom. Our assumption here is that the proposal density will cover the posterior density of the TN distribution after we multiply it by

a suitable constant. The following algorithm describes how the draws are obtained using rejection sampling algorithm.

Algorithm 2.1 (Rejection Sampling)

1. Simulate θ from the proposal density $q(\theta)$.

2. Generate a random variable U uniformally distributed on the interval $[0, 1]$.

3. Compute the acceptance probability $r = g(\theta|x)/Kq(\theta)$.

4. If $U \leq r$ accept θ, otherwise reject θ.

5. Repeat steps 1 to 4 of the algorithm until the efficient number of acceptance rate is obtained.

2.3.2 Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is an alternative method that can be mainly used to generate random samples from a multivariate probability distribution function with a higher dimension. However, it can be used for a low dimension situation as in this thesis. Markov Chain Monte Carlo methods have been applied in many fields since 1990s. These methods are general and flexible along with the huge development of computing facilities. One approach to construct MCMC is the Metropolis-Hastings algorithm which was introduced by Metropolis et al. (1953). Hastings (1970) has made a generalization of Metropolis-Hastings algorithm. There are two different approaches of the Metropolis-Hastings algorithm: (1) the random walk chain and (2) the independent chain. In the following we will discuss the general algorithm for the Metropolis-Hastings. Then present the algorithm for the random walk chain approach that will be used throughout the thesis.

2.3.3 Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm is a common technique to simulate random sequence from the posterior probability density, $g(\theta|x)$. The Metropolis-Hastings algorithm starts with an initial point of the sequence, θ^0, and identifies a rule for sampling the mth element in the sequence, θ^m, given the $(m-1)^{th}$ element in the sequence,

θ^{m-1}. This rule requires a proposal density, from which we simulate a candidate random point, θ^*, and calculates an acceptance probability p, that depend on the posterior distribution and the proposal density, which gives the probability that the random candidate will be accepted as the following sampled point in the sequence. Below, is the general set-up of the Metropolis-Hastings algorithm:

Algorithm 2.2 (Metropolis-Hastings algorithm)

1. Specify the number of iterations M

2. Set a starting value $\theta^0 = (\alpha^0, \beta^0)$.

3. For $m = 1, 2, \cdots, M$, do the following

    (a) Simulate a candidate θ^* from a proposal density $q(\theta, \theta^{m-1})$

    (b) Calculate the ratio
    $$r = \frac{g(\theta^*|x)q(\theta^{m-1}|\theta^*)}{g(\theta^{m-1}|x)q(\theta^*|\theta^{m-1})}$$

    (c) Calculate the acceptance probability
    $$p = \min(r, 1)$$

    (d) Generate a random value U from a uniform distribution on $(0, 1)$ interval

    (e) Set
    $$\theta^m = \begin{cases} \theta^* & \text{if } U \leq p\,, \\ \theta^{m-1} & \text{otherwise}\,. \end{cases}$$

Under few conditions on the proposal density $q(\theta^*|\theta^{m-1})$, the sequence of random simulated draws $\theta^1, \theta^2, \cdots, \theta^M$ should converge to a random sample from the actual posterior distribution $g(\theta|x)$ Albert (2009). In the MCMC, the proposal density $q(\theta^*|\theta)$ plays a significant rule in the convergence process. A poor selection of the proposal density might lead to a delay in the convergence process.

Random Walk Chain

In the random walk chain, the proposal density is symmetric and should satisfies the following condition:
$$q(\theta^*|\theta) = q(\theta|\theta^*), \text{ for all } \theta^* \text{ and } \theta$$

Therefore, the ratio r, in the Metropolis-Hasings algorithm, for the random walk chain becomes

$$r = \frac{g(\theta^*|x)}{g(\theta^{m-1}|x)}.$$

One of the possible choices of the proposal can be the multivariate normal distribution with a mean of θ^{m-1} and variance-covariance matrix scale I, where I is the identity matrix and scale is a scale parameter that controls the spread of the candidates from the current mean, θ^{m-1}. This can be expressed as

$$\theta^* = \theta^{m-1} + scale \quad Z,$$

where Z is multivariate normal with a mean vector of zeros, and the variance-covariance matrix and a scale should be a positive scale parameter. As a special case from Algorithm 2.2, the random walk chain algorithm can be described as follows:

Algorithm 2.3 (Random Walk Chain -Metropolis-Hastings algorithm)

1. Specify the number of iterations M and the value of the scale

2. Set a starting value θ^0

3. For $m = 1, 2, \cdots, M$, do the following

   (a) Simulate $Z \sim N(0, I)$

   (b) Simulate a candidate $\theta^* = \theta^{m-1} + \text{scale } Z$

   (c) Calculate the ratio
   $$r = \frac{g(\theta^*|x)}{g(\theta^{m-1}|x)}$$

   (d) Calculate the acceptance probability
   $$p = \min(r, 1)$$

   (e) Generate a random value U from a uniform distribution on $(0, 1)$ interval

   (f) Set
   $$\theta^m = \begin{cases} \theta^* & \text{if } U \leq p, \\ \theta^{m-1} & \text{otherwise}. \end{cases}$$

Chapter 3

Simulation Study

In this chapter, we use the simulation method to test the performance of the methods applied in the thesis to estimate the parameters (α, β) of the TN distribution and compare them. As criteria of comparisons between the three methods used (the maximum likelihood, the rejections sampling and the Markov Chain Mote Carlo), we use

1. the mean square errors (MSE) that measures the spread of the point estimates from the actual value of the parameter.

2. the mean of the bias that estimates the deviation of the point estimate from the actual value of the parameter, and

3. the coverage probability of the interval estimation that measures the reliability of the interval estimation of the parameter. The coverage probability is the average of an indicator that the estimated interval captures the true value of the parameter.

This simulation study is conducted based on the following scheme:

1. Set the parameters' values of $\theta = (\alpha, \beta)$.

2. Set the sample size n.

3. Simulate a random sample with size n from the $TN(\alpha, \beta)$ model.

4. Calculate the point estimates of α and β using the MLE method.

5. Calculate the Bayes estimates of α and β using the rejection sampling and Markov Chain Monte Carlo algorithms.

6. Calculate deviations of the point estimate of α and β from their actual values, specified in step 2, using the three estimation methods.

7. Calculate a 95% confidence interval for α and β using the maximum likelihood method.

8. Calculate a 95% Bayesian probability interval for α and β using both rejection sampling and Markov Chain Monte Carlo methods.

9. Calculate an indicator C as

$$
C = \begin{cases} 1 & \text{if interval estimate capture the true parameter value} \\ 0 & \text{otherwise} \end{cases}
$$

for the three types of the interval estimates.

10. Repeat steps 3 to 9 N times.

11. Compute the MSE, the mean absolute bias of the point estimates and the coverage probability of the interval estimates using the three techniques:

    (a) The MSE of the $\hat{\theta}$ is

    $$
    MSE(\hat{\theta}) = \sum_{i=1}^{N} \frac{(\hat{\theta}^{(i)} - \theta)^2}{N},
    $$

    where $\hat{\theta}^{(i)} = (\hat{\alpha}^{(i)}, \hat{\beta}^{(i)})$ is the point estimate of θ using the generated sample in the ith iteration.

    (b) The mean of the bias is

    $$
    \frac{\sum_{i=1}^{N} \theta - \hat{\theta}^{(i)}}{N}
    $$

    (c) The coverage probability (CP) is

    $$
    CP = \frac{\sum_{i=1}^{N} C_i}{N}
    $$

    where $C_i = 1$ if the confidence interval in the ith iteration captures the true parameter and zero otherwise.

12. Repeat steps 2-11, using different sample sizes.

This simulation study was performed when the number of iterations $N = 10,000$, the sample size $n = 25, 50, 75$, and 100, and the true values of the parameters' values of (α, β) are (1,0.5), (1,1), (2,1), (1.3,0.5), (1,1.2), (2,1.2), (1.3,1.5), (0.9,1.2). The main reason behind setting these true values of (α, β) is that they give different shapes of the hazard function of the model such as decreasing, increasing, upside-down bathtub, and increasing-decreasing-increasing, that have a wide applications in reliability analysis and survival analysis.

All results obtained from the simulation study described above are presented in Tables 3.1 and 3.2. Table 3.1 shows all results for α while those for β are shown in Table 3.2. Based on these results, we can conclude the following:

1. As it was expected, the MSE of the point estimates of both two parameters using all the three methods applied here decreases.

2. The RS approach for estimating α gives slightly better estimate in term of having a smaller MSE, then the maximum likelihood then the MCMC.

3. The mean bias of the point estimate of α using either the maximum likelihood or the rejection sampling decreases as the sample size increases, while it fluctuates with the sample size for the point estimate obtained using the MCMC.

4. The MSE and the mean bias corresponding to the point estimate of β was a bit surprising. Although it decreases as sample size increases for all methods applied here. However, the MCMC approach has the least MSE and mean bias then the maximum likelihood then the Rejection sampling. Since the difference is very small, we can conclude that the MCMC provides a slightly better estimate for β. I cannot find a scientific justification of this result.

5. The coverage probability remains close to the nominal level of 95% for both two parameters α and β. This indicates that the estimation process using all the three approaches are consistent.

6. As a final conclusion, although considering non-informative prior distributions of the two parameters of interest, Bayes method provides a slightly better estimate. This also indicates that with a better prior information about the

parameters, the Bayes technique will provide a much better estimations of the parameters than the maximum likelihood method.

Table 3.1: The MSE, the mean bias, and the coverage probability for estimate parameter $\hat{\alpha}$

		MSE			The mean Bias			CP		
(α, β)	n	MLE	RS	MCMC	MLE	RS	MCMC	MLE	RS	MCMC
	25	0.0090247	0.0088399	0.0142309	-0.0307509	-0.0286229	-0.0781322	0.9530	0.9464	0.9505
(1,0.5)	50	0.0037691	0.0037249	0.0106144	-0.0143683	-0.0132126	-0.0828773	0.9537	0.9501	0.9566
	75	0.0024041	0.0023845	0.0113127	-0.0092368	-0.0084402	-0.0935309	0.9485	0.9462	0.9572
	100	0.0017887	0.0017768	0.0129918	-0.0073182	-0.0067159	-0.1045713	0.9495	0.9482	0.9600
	25	0.0364983	0.0357081	0.0430253	-0.0603781	-0.0559050	-0.1030578	0.9485	0.9428	0.9486
(1,1)	50	0.0149281	0.0147403	0.0223756	-0.0266472	-0.0242554	-0.0900569	0.9507	0.9496	0.9561
	75	0.0093926	0.0093062	0.0186259	-0.0185228	-0.0168493	-0.0972937	0.9543	0.9503	0.9581
	100	0.0068175	0.0067692	0.0178692	-0.0141456	-0.0129145	-0.1051270	0.9530	0.9520	0.9604
	25	0.1437911	0.1403823	0.1506751	-0.1193386	-0.1098520	-0.1522918	0.9526	0.9481	0.9501
(2,1)	50	0.0596186	0.0588360	0.0677441	-0.0535587	-0.0485227	-0.1077328	0.9517	0.9475	0.9527
	75	0.0380900	0.0377227	0.0477539	-0.0387416	-0.0353217	-0.1070604	0.9493	0.9480	0.9557
	100	0.0277322	0.0275222	0.0385884	-0.0285225	-0.0258701	-0.1087069	0.9508	0.9490	0.9559
	25	0.0571776	0.0559792	0.0639131	-0.0709390	-0.0656531	-0.1117291	0.9540	0.9500	0.9536
(1.3,0.5)	50	0.0244056	0.0241197	0.0324985	-0.0357619	-0.0329869	-0.0972728	0.9532	0.9506	0.9575
	75	0.0156964	0.0155596	0.0254889	-0.0249510	-0.0230708	-0.1018349	0.9488	0.9480	0.9569
	100	0.0114690	0.0113937	0.0227521	-0.0184662	-0.0170623	-0.1073845	0.9488	0.9468	0.9562
	25	0.0360937	0.0352500	0.0425862	-0.0595764	-0.0548064	-0.1019694	0.9498	0.9454	0.9486
(1,1.2)	50	0.0152623	0.0150576	0.0229047	-0.0287824	-0.0262472	-0.0920905	0.9495	0.9468	0.9541
	75	0.0098031	0.0097062	0.0189930	-0.0188302	-0.0170657	-0.0971266	0.9466	0.9443	0.9552
	100	0.0070438	0.0069857	0.0180967	-0.0141047	-0.0127394	-0.1053928	0.9499	0.9489	0.9573
	25	0.1448127	0.1413013	0.1510043	-0.1150852	-0.1049276	-0.1475068	0.9485	0.9421	0.9455
(2,1.2)	50	0.0605183	0.0596049	0.0687346	-0.0549594	-0.0494948	-0.1091730	0.9499	0.9468	0.9524
	75	0.0387183	0.0382950	0.0484883	-0.0397629	-0.0360067	-0.1087349	0.9496	0.9467	0.9537
	100	0.0282708	0.0280094	0.0394723	-0.0310087	-0.0281339	-0.1114966	0.9494	0.9474	0.9563
	25	0.0613497	0.0596079	0.0679775	-0.0790759	-0.0718990	-0.1178699	0.9525	0.9477	0.9510
(1.3,1.5)	50	0.0265185	0.0260557	0.0347633	-0.0394914	-0.0356047	-0.0995645	0.9473	0.9450	0.9506
	75	0.0162821	0.0160820	0.0258374	-0.0247889	-0.0221273	-0.1008386	0.9478	0.9456	0.9538
	100	0.0120187	0.0119109	0.0230399	-0.0180960	-0.0160913	-0.1060103	0.9495	0.9483	0.9588
	25	0.0300213	0.0293266	0.0362706	-0.0546920	-0.0504085	-0.0978933	0.9544	0.9480	0.9503
(0.9,1.2)	50	0.0125876	0.0123960	0.0203034	-0.0277164	-0.0253759	-0.0919415	0.9510	0.9481	0.9544
	75	0.0077636	0.0076857	0.0170981	-0.0176204	-0.0160360	-0.0974069	0.9510	0.9494	0.9567
	100	0.0056352	0.0055910	0.0166370	-0.0121738	-0.0109647	-0.1046352	0.9492	0.9477	0.9591

Table 3.2: The MSE, the mean bias, and the coverage probability for estimated parameter
$\hat{\beta}$

(α, β)	n	MSE			The mean Bias			CP		
		MLE	RS	MCMC	MLE	RS	MCMC	MLE	RS	MCMC
(1,0.5)	25	0.0625544	0.0610927	0.0588462	-0.0433460	-0.0442486	-0.0412347	0.9397	0.9414	0.9532
	50	0.0270002	0.0266452	0.0254013	-0.0204737	-0.0209810	-0.0162332	0.9438	0.9435	0.9586
	75	0.0163671	0.0162215	0.0153258	-0.0117637	-0.0121362	-0.0063836	0.9463	0.9489	0.9630
	100	0.0121887	0.0121135	0.0113671	-0.0107005	-0.0109817	-0.0040027	0.9483	0.9487	0.9659
(1,1)	25	0.062013	0.0605317	0.0589516	-0.0380127	-0.0389408	-0.0356700	0.9422	0.9420	0.9509
	50	0.0265728	0.0262218	0.0252829	-0.0215563	-0.0220700	-0.0175026	0.9458	0.9472	0.9567
	75	0.0164310	0.0162914	0.0155912	-0.0140454	-0.0143467	-0.0087601	0.9517	0.9509	0.9622
	100	0.0122497	0.0121709	0.0116231	-0.0091143	-0.0093780	-0.0028427	0.9459	0.9474	0.9591
(2,1)	25	0.0611058	0.0596419	0.0586096	-0.0397009	-0.0406865	-0.0376934	0.9427	0.9439	0.9485
	50	0.0262292	0.0258650	0.0251815	-0.0210613	-0.0215361	-0.0174173	0.9476	0.9456	0.9564
	75	0.0162195	0.0160765	0.0155805	-0.0127395	-0.0131036	-0.0082402	0.9472	0.9470	0.9569
	100	0.0120350	0.0119538	0.0115329	-0.0089177	-0.0092008	-0.0035711	0.9472	0.9464	0.9576
(1.3,0.5)	25	0.0170082	0.0167513	0.0152405	-0.0032900	-0.0110950	-0.0225355	0.9329	0.9456	0.9592
	50	0.0079490	0.0079029	0.0071729	-0.0025239	-0.0064917	-0.0180112	0.9426	0.9482	0.9650
	75	0.0052298	0.0052023	0.0047433	-0.0005788	-0.0032541	-0.0157270	0.9437	0.9484	0.9639
	100	0.0038011	0.0037807	0.0034592	0.0007868	-0.0012286	-0.0149074	0.9455	0.9526	0.9699
(1,1.2)	25	0.0961395	0.0937532	0.0912453	-0.0600467	-0.0581960	-0.0524105	0.9430	0.9414	0.9507
	50	0.0382755	0.0377222	0.0362583	-0.0319142	-0.0309557	-0.0232781	0.9474	0.9462	0.9569
	75	0.0235743	0.0233356	0.0223286	-0.0181629	-0.0175234	-0.0083835	0.9481	0.9488	0.9610
	100	0.0172101	0.0170841	0.0162885	-0.0126416	-0.0121964	-0.0016009	0.9488	0.9475	0.9624
(2,1.2)	25	0.0941778	0.0917435	0.0897167	-0.0659011	-0.0639281	-0.0590030	0.9477	0.9443	0.9504
	50	0.0390081	0.0384346	0.0373698	-0.0277905	-0.0268394	-0.0205051	0.9467	0.9461	0.9537
	75	0.0240347	0.0237904	0.0229680	-0.0185836	-0.0179766	-0.0104427	0.9463	0.9477	0.9570
	100	0.0171345	0.0169861	0.0163889	-0.0127128	-0.0122524	-0.0035855	0.9517	0.9507	0.9640
(1.3,1.5)	25	0.1626922	0.1588441	0.1542376	-0.1010802	-0.0953382	-0.0864913	0.9536	0.9435	0.9531
	50	0.0646044	0.0635965	0.0611348	-0.0459814	-0.0428938	-0.0314679	0.9479	0.9443	0.9556
	75	0.0400133	0.0395511	0.0376515	-0.0324990	-0.0303834	-0.0166876	0.9498	0.9480	0.9620
	100	0.0282662	0.0280135	0.0266310	-0.0211885	-0.0196463	-0.0040377	0.9506	0.9490	0.9630
(0.9,1.2)	25	0.0939258	0.0916782	0.0890387	-0.0663984	-0.0644528	-0.0585222	0.9498	0.9463	0.9550
	50	0.0392678	0.0386862	0.0371772	-0.0328858	-0.0319351	-0.0239525	0.9479	0.9458	0.9575
	75	0.0244968	0.0242495	0.0231568	0.0194617	-0.0188337	-0.0092447	0.9463	0.9473	0.9598
	100	0.0172366	0.0171081	0.0162729	-0.0134273	-0.0129684	-0.0020590	0.9488	0.9496	0.9643

Chapter 4

Applications

In this chapter, we present and analyze three data sets in order to demonstrate the use of methods applied in this thesis for three real data sets. The data sets are referred as Aarset Data, Recidivism Failure Times Data and Active Repair Data.

We used the rejection sampling and MCMC approaches to estimate the TN distribution parameters. The maximum likelihood method will be used to compare with the Bayesian methods. The point and the interval estimates for the two parameters are obtained. Note that, R-programming language was adapted here for the numerical calculations.

4.1 Aarset Data

In this data, 50 devices were put on the life test and their lifetime "time to failure" (in days) were observed. This data were originally studied by Aarset [1]. The data (presented in days) are :

0.1, 0.2, 1, 1, 1, 1, 1, 2, 3, 6, 7, 11, 12, 18, 18, 18, 18, 18, 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85, 85, 86, 86.

Aarset [1] showed that the data have a bathtub shaped hazard rate. Sarhan et al. [12] used the TN distribution to re-analyze this data using the maximum likelihood method. Here, we assume that the Aarset data follow the TN distribution. We use the Bayesian methods to estimate the TN distribution parameters. In this study, it is assumed that α and β are independent with gamma prior distribution when all the hyperparameters are equal and equal to 0.001. To see the range of the plausible values of the two parameters α and β, we depict the contour of the likelihood function for α and β in Figure 4.1. We simulated 10,000 draws from the joint posterior distribution, using the RS and MCMC techniques. In the MCMC method, the first half of the simulated draws were ignored. Figures 4.2 and 4.3 show the posterior contours of α and β along with simulated draws using the RS and MCMC, respectively.

The simulated draws are used to :

1. Plot the marginal posterior densities of α and β as shown in figures 4.4 and 4.5.

2. Calculate the Bayesian point and credible interval estimates of α and β as provided in Table 4.1.

Table 4.1 provides the Bayesian point estimate, the 95 % credible interval, and the corresponding width for α and β using the RS, the MCMC, and the MLE. The results in Table 4.1 show that the two techniques provide quite similar results for the two parameters. However, the MCMC gives a slightly better estimation in the sense of the narrower credible intervals. The acceptance rate for the RS is 50.07 % and for the MCMC is 53.31 %. The MCMC has a high acceptance rate because the scale value is small which leads to a small spread of the proposed θ^* from the current $\theta^{(m-1)}$. To investigate the performance of the MCMC, we provide the trace plots and the autocorrelation plots for the simulated draws as shown in Figure 4.6 and 4.7, respectively. The trace plot of the simulated draws of $(\log \alpha, \log \beta)$ shows that the chain covers the same region of the log of the parameter space and finds the stationary distribution. The autocorrelation drops from lag 0 to lag 13, which indicates that the simulated draws in the chain are slightly correlated with the previous ones that indicates no sign for concern. The two plots tell us that the draws are approximately independent, random samples from the posterior distribution. From the marginal posterior density graphs, we can see that the marginal posterior density obtained form the two approaches are similar. Also, the posterior distribution of α is approximately symmetric while that for β is right skewed.

Method	Parameter	Point estimate	Interval	Width
RS	α	0.9783	(0.7928,1.1792)	0.3864
	β	0.0233	(0.0088,0.0451)	0.0363
MCMC	α	0.9748	(0.7978,1.1690)	0.3712
	β	0.0012	(0.0095,0.0441)	0.0346
MLE	α	0.9439	(0.7140,1.1738)	0.4598
	β	0.0207	(0.0001, 0.0413)	0.0412

Table 4.1: The point and interval estimates of the parameters along with the intervals' width for the Aarset data using the rejection sampling, the Markov Chain Monte Carlo, and the maximum likelihood estimation methods.
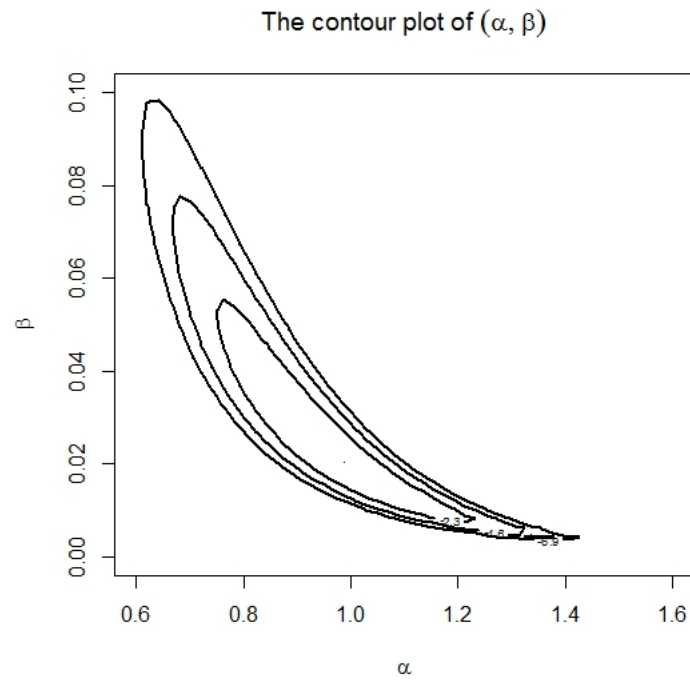
The contour plot of (α, β)

Figure 4.1: The contour plot for the likelihood function of (α, β) using Aarset data.
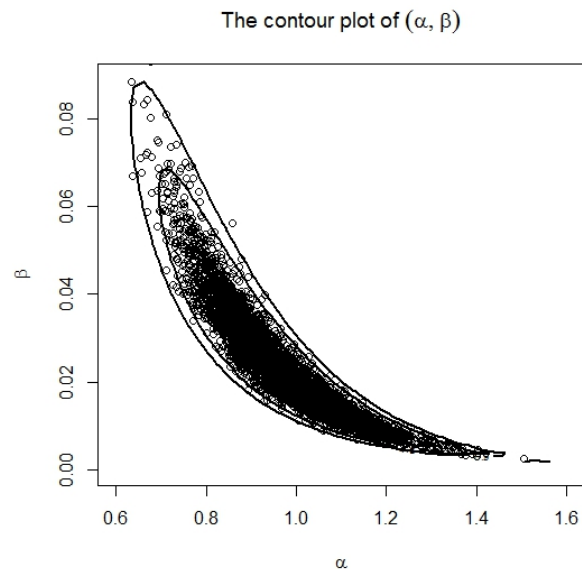
The contour plot of (α, β)

Figure 4.2: The contour plot of (α, β) along with the simulated draws using the rejection sampling algorithm.
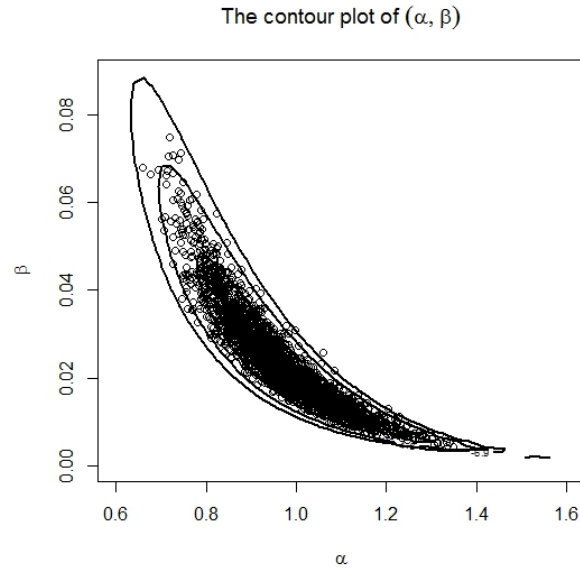
Figure 4.3: The contour plot of (α, β) along with the simulated draws using the Markov Chain Monte Carlo method after discarding the first half of the draws.
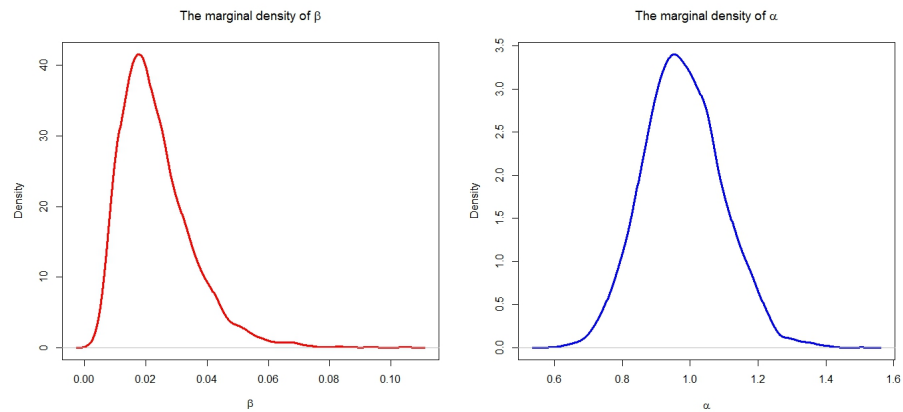

Figure 4.4: The marginal posterior density functions of α and β using simulated draws obtained from the rejection sampling algorithm.
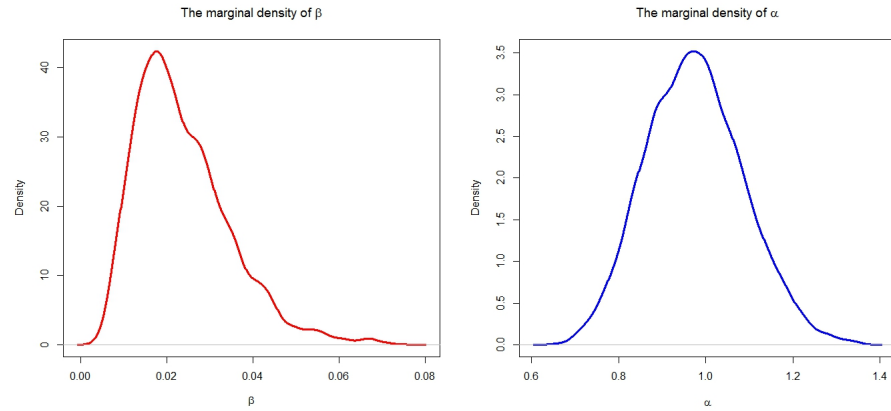
Figure 4.5: The marginal posterior density functions of α and β using the random samples obtained from the Markov Chain Monte Carlo approach after discarding the first 5000 draws.
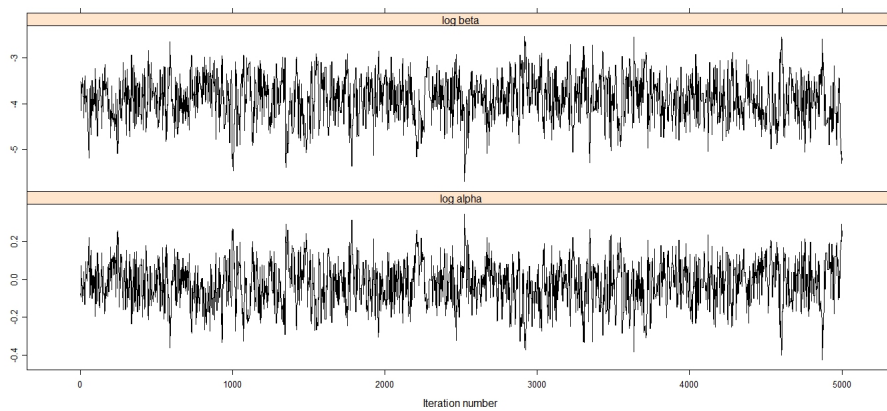
Figure 4.6: The trace plots of the simulated draws of $\log \alpha$ and $\log \beta$ using the MCMC algorithm after discarding the first half of draws.
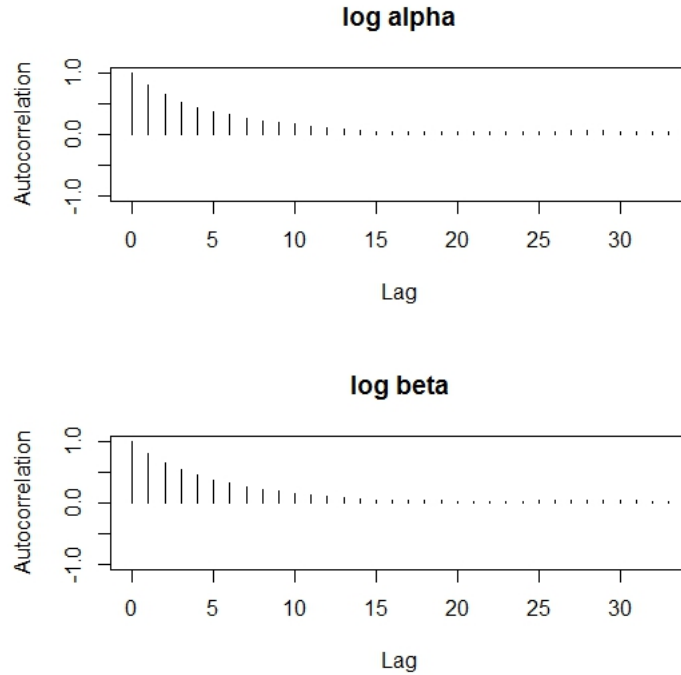
Figure 4.7: The autocorrelation plots of the simulated draws of $\log \alpha$ and $\log \beta$ using the MCMC method with good choices of the initial value and the scale factor.

4.2 Recidivism Failure Times

The next data set contains 61 observed recidivism failure times (in days) revealed by correctional institutions in Columbia USA [13]. The failure times data were

1, 6, 9, 29, 30, 34, 39, 41, 44, 45, 49, 56, 84, 89, 91, 100, 103, 104, 115, 119, 124, 138, 141, 146, 156, 162, 168, 183, 185, 198, 209, 217, 217, 228, 233, 238, 241, 252, 258, 271, 275, 276, 279, 282, 305, 313, 329, 331, 334, 336,336, 362, 384, 404, 408, 422, 438, 441, 465, 486, 556. Sarhan et. al [12] used the TN model to analyze this data set by using the maximum likelihood approach. In this study, we suppose that the recidivism failure times data follow the TN distribution. We compute the point estimates and the two sided credible intervals by using the Bayesian approach. The two parameters α and β are considered to be independent random variables with gamma prior distribution when all the hyperparameters are equal to 0.001. To demonstrate the spread of the possible values of the parameters, we present the contour plot of the likelihood for (α, β) in Figure 4.8. We use the RS, the MCMC, and the MLE methods to simulate 10,000 draws from the joint posterior distribution. However, the first half draws was discarded in the MCMC method. By applying the RS and MCMC

techniques, Figure 4.9 and 4.10 display the posterior contours of (α, β) along with the simulated draws. We use these draws to plot the marginal posterior densities for parameters as provided in Figures 4.11 and 4.12. Also, we compute the Bayesian point estimates, the credible intervals and the corresponding widths for α and β as shown in Table 4.2. For comparison, we provided the maximum likelihood estimates in Table 4.2.

From the results in Table 4.2, we can say that the two Bayesian approaches show almost similar results for α and β. However, based on the widths, we find that the MCMC provides slightly better estimation than the RS procedure. For the RS, the acceptance rate is 51.45 %, whereas for the MCMC, the acceptance rate is 52.33 %. The MCMC has a high acceptance rate because the scale value is small which leads to a small spread of the proposed θ^* from the current $\theta^{(m-1)}$. To evaluate the performance of the MCMC method, we illustrate the trace plots and autocorrelation plots for the simulated draws in figures 4.13 and 4.14 respectively. The trace plot of the simulated draws of $(\log \alpha, \log \beta)$ shows that the chain covers the same region of the the parameter space and has reached the stationary distribution. The autocorrelation drops from lag 0 to lag 13, which indicates that the simulated draws in the chain are slightly correlated with the previous ones that indicates no sign for concern. The two plots suggest that the draws are approximately independent, random samples from the posterior distribution. From the two marginal posterior densities plots, it is clear that the two Bayesian approaches are nearly the same. Furthermore, the posterior distribution of α is almost symmetric while the posterior distribution of β is right skewed.

Method	Parameter	Point estimate	Interval	Width
RS	α	1.2753	(1.0488,1.5117)	0.4629
	β	0.0012	(0.0002,0.0033)	0.0031
MCMC	α	1.2798	(1.0582,1.5166)	0.4584
	β	0.0012	(0.0002,0.0032)	0.0030
MLE	α	1.0623	(0.9975,1.1271)	0.1296
	β	0.0026	(0.0018,0.0033)	0.0015

Table 4.2: The point and interval estimates along with the width of the two parameters α and β using the Bayes and maximum likelihood techniques for recidivism failure times data.
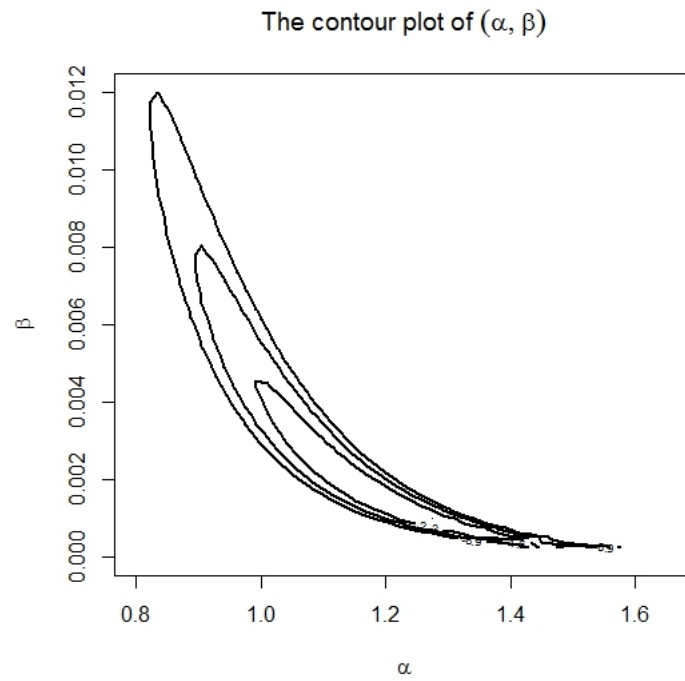
Figure 4.8: The contour plot of the likelihood function for parameters (α, β) in the TN distribution using recidivism failure times.
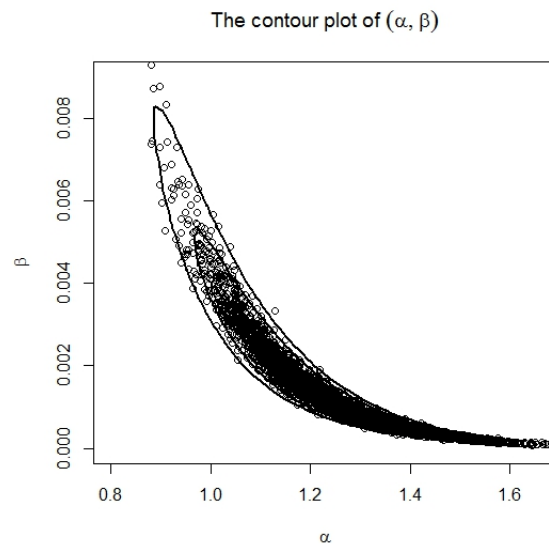

Figure 4.9: The contour plot of the posterior distribution along with the simulated draws of (α, β) in the TN model using the rejection sampling algorithm.
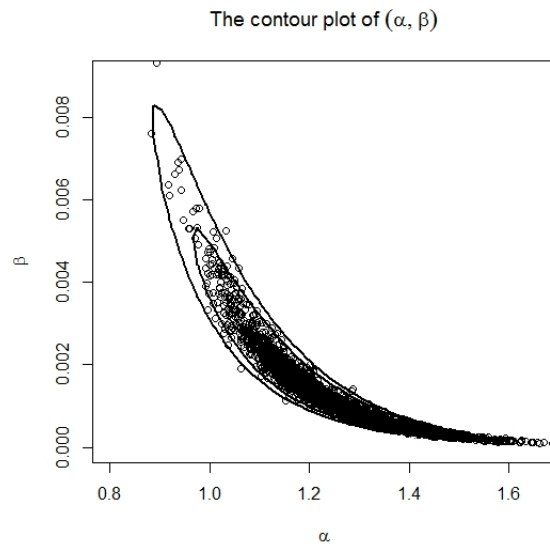
The contour plot of (α, β)

Figure 4.10: The contour plot of the posterior distribution along with the simulated draws of (α, β) using the Markov Chain Monte Carlo method. The second 5000 draws have been used.
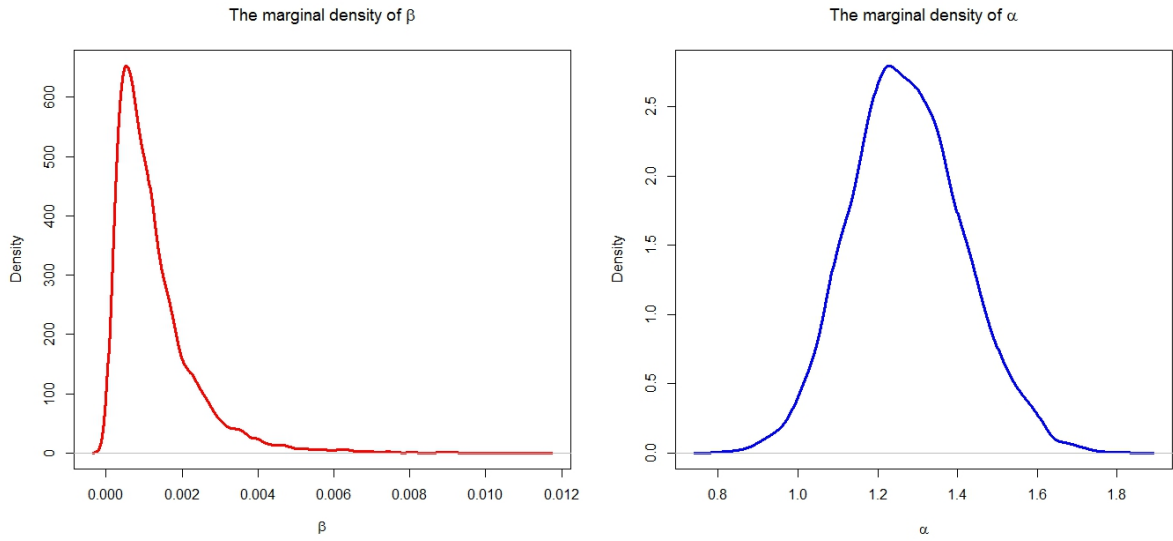
The contour plot of (α, β)
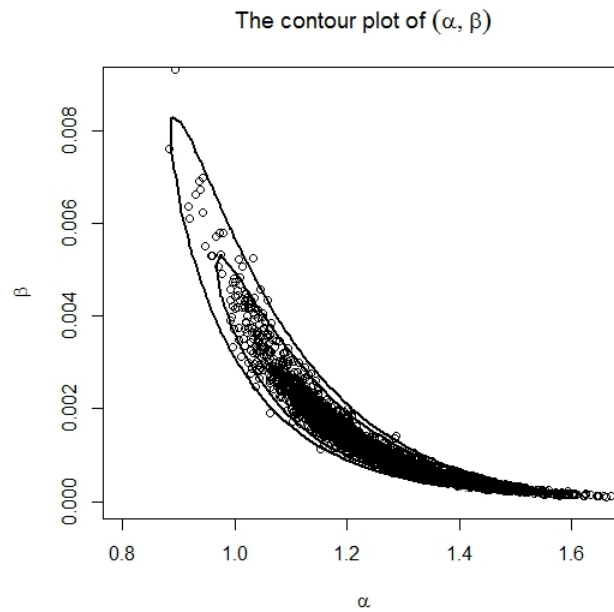
Figure 4.11: The contour plot of the posterior distribution along with the simulated draws of (α, β) using the Markov Chain Monte Carlo method.

Figure 4.12: The marginal posterior density of parameters α and β using the rejection sampling algorithm.
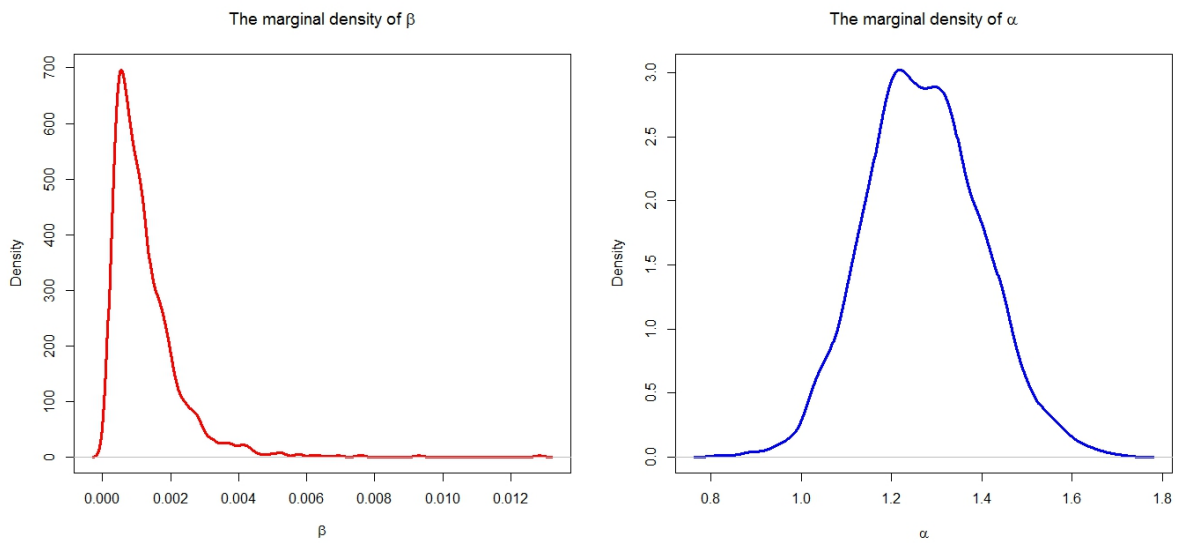

Figure 4.13: The marginal posterior density of parameters α and β using the Markov Chain Monte Carlo approach.
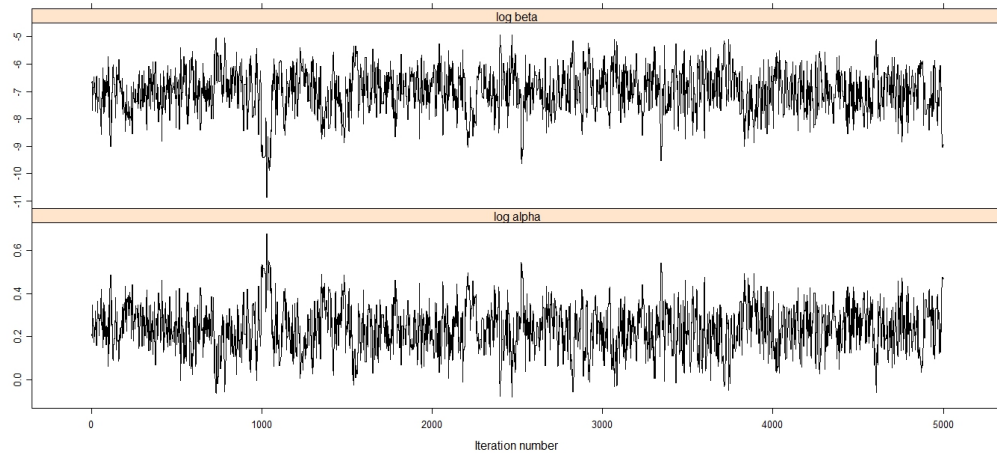
Figure 4.14: The trace plots of the simulated draws for log α and log β from the Markov Chain Monte Carlo method in the TN model problem with discarded the first 5000 draws.
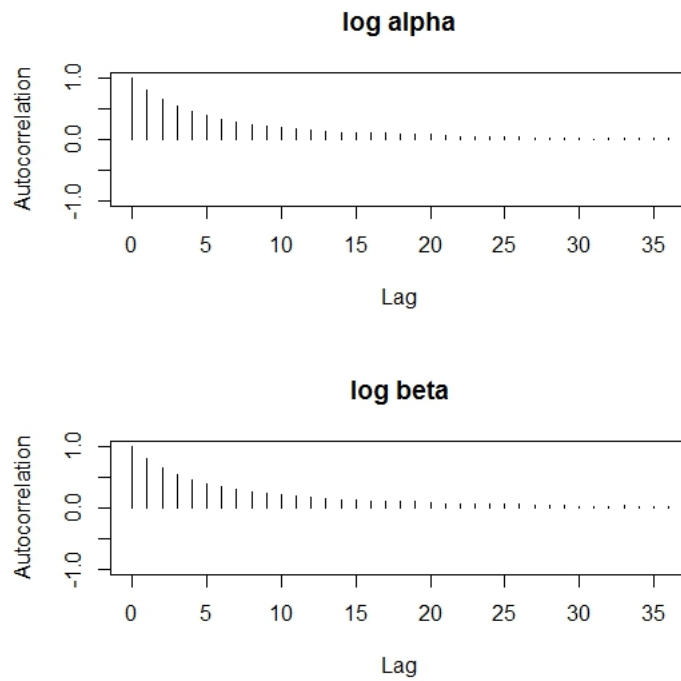

Figure 4.15: The autocorrelations plots of the simulated draws for log α and log β from Markov Chain Monte Carlo approach.

4.3 Active Repair Times Data

These data show the active repair times of 46 (in hours) airborne communication transceivers originally taken from Brillinger [3].

The data for the active repair times are presented below :

0.2, 0.3, 0.5, 0.5, 0.5, 0.5, 0.6, 0.6, 0.7, 0.7, 0.7, 0.8, 0.8, 1, 1, 1, 1, 1.1, 1.3, 1.5, 1.5, 1.5, 1.5, 2, 2, 2.2, 2.5, 2.7, 3, 3, 3.3, 3.3, 4, 4, 4.5 , 4.7, 5, 5.4, 5.4, 7, 7.5, 8.8, 9, 10.3, 22, 24.5.

It is assumed that the active repair data are independent and identically distributed, and we suppose that the active repairs data follow the TN distribution. Here, let us consider that the two parameters α and β are independent with gamma prior, and the values of all the hyperparameters are equal to 0.001. We plot the contour of the likelihood function for α and β in order to identify the range of the plausible values of the TN model parameters in Figure 4.15.

We use 10,000 iterations to simulate form the joint posterior distribution, by applying the two Bayesian approaches along with the maximum likelihood estimation. In the MCMC method, we use the second half of the simulated draws. Figure 4.16 shows the contour plot of the posterior density along with simulated draws of (α, β) using the RS method, and Figure 4.17 displays the posterior density along with the simulated draws of (α, β) from the MCMC method. The simulated draws are used to draw the marginal posterior densities of α and β parameters for the RS and the MCMC procedures as presented in Figures 4.18 and 4.19, respectively. Moreover, we calculate the point estimates, the Bayesian intervals and the associated widths as given in Table 4.3. Also, we calculate the point and interval estimates of the maximum likelihood method to compare it with the Bayesian methods. Table 4.3 reveals the results of the two Bayesian methods and maximum likelihood approach for α and β parameters, which shows close results for the both Bayesian approaches. However, the RS does not give better estimation results in the sense of the narrower intervals as the MCMC method. The acceptance rate of the RS is 50.07% and for the MCMC 50, 22%. The MCMC has a high acceptance rate because the scale value is small which leads to a small spread of the proposed θ^* from the current $\theta^{(m-1)}$. To see how the MCMC method is performed, we provide the trace plots and the autocorrelation plots for the simulated draws in figures 4.20 and 4.21, respectively. The trace plots of the simulated draws of $(\log \alpha, \log \beta)$ shows that the chain covers the same region of the parameter space and finds the stationary distribution. The autocorrelation drops from lag 0 to lag 13, which indicates that the simulated draws in the chain are slightly correlated with the previous ones that
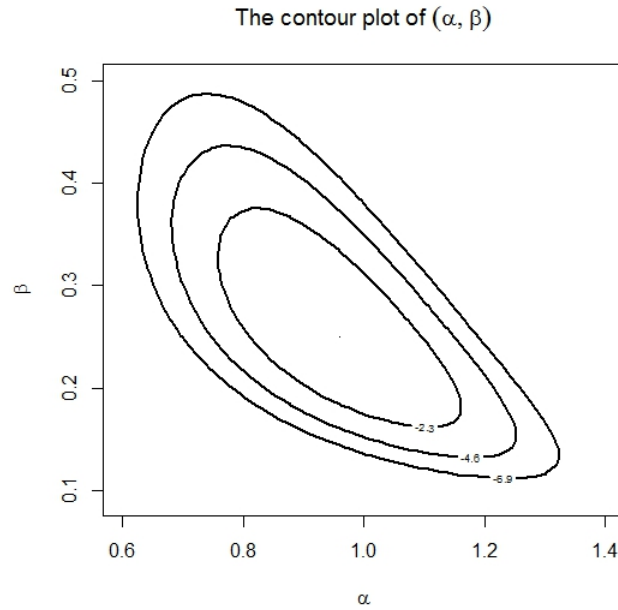
The contour plot of (α, β)

Figure 4.16: Graph of the contour plot for the likelihood function of parameters (α, β) in the TN model using active repair times data.

indicates no sign for concern. The two plots indicate that the draws are approximately independent, random samples from the posterior distribution which means that the MCMC is mixing appropriately. The two marginal posterior densities for the two methods depict that the two Bayesian techniques are close. Also, as can be easily seen, the marginal posterior distribution of α and β are symmetric using the RS and the MCMC methods.

Method	Parameter	Point estimate	Interval	Width
RS	α	0.9522	(0.7975,1.1079)	0.3104
	β	0.2558	(0.1816,0.3434)	0.1618
MCMC	α	0.9512	(0.8007,1.1023)	0.3016
	β	0.2557	(0.1856,0.3388)	0.1532
MLE	α	0.8389	(0.6598,1.0180)	0.3582
	β	0.3103	(0.1822,0.4384)	0.2562

Table 4.3: The point and interval estimates with width of the parameter α and β of the TN model using the RS, the MCMC, and the MLE methods for the active repair times data.
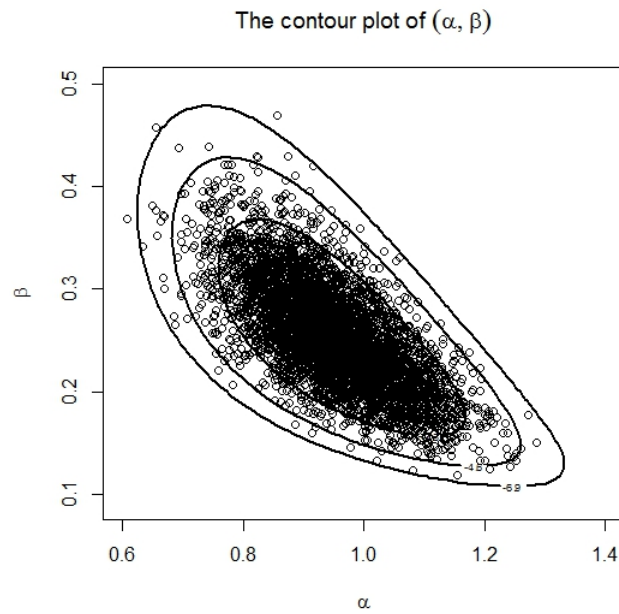
Figure 4.17: The contour plot of the posterior distribution along with the simulated draws of parameters (α, β) form the rejection sampling algorithm using the active repair times data.
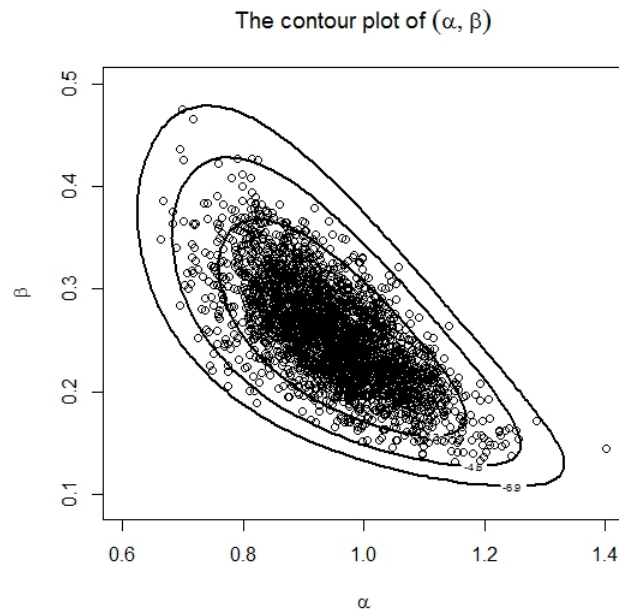

Figure 4.18: The contour plot of the posterior distribution along with the simulated draws of parameters (α, β) from Markov Chain Monte Carlo method after the first 5000 have been ignored.
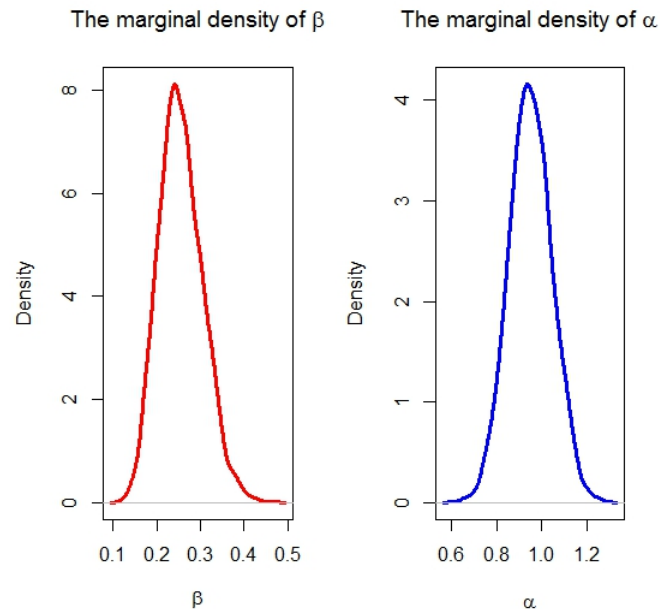
Figure 4.19: The marginal posterior density of parameters α and β in the TN distribution problem using the rejection sampling algorithm with $10,000$ iterations.
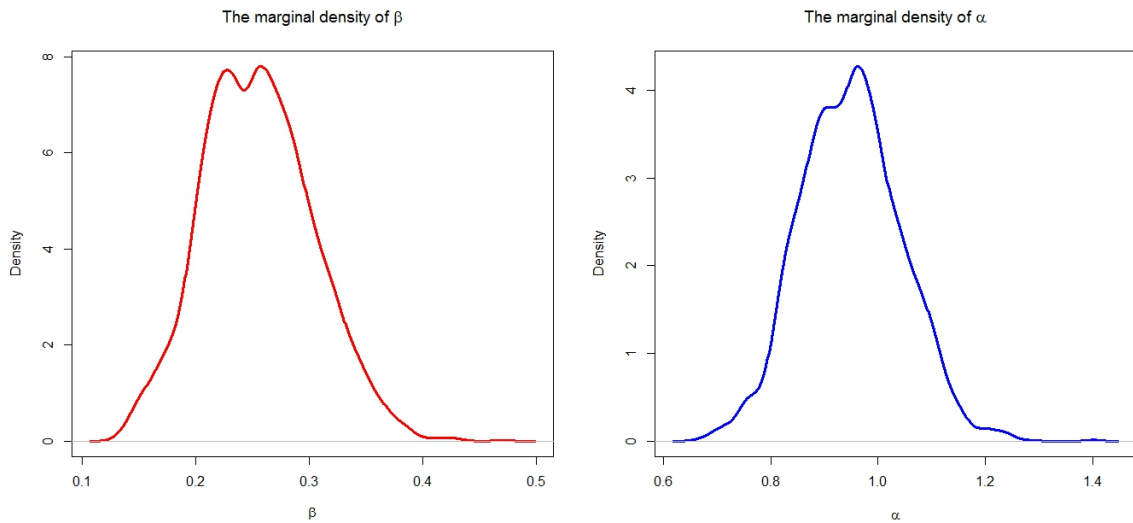

Figure 4.20: The marginal posterior density of parameters α and β using the Markov Chain Monte Carlo algorithm.
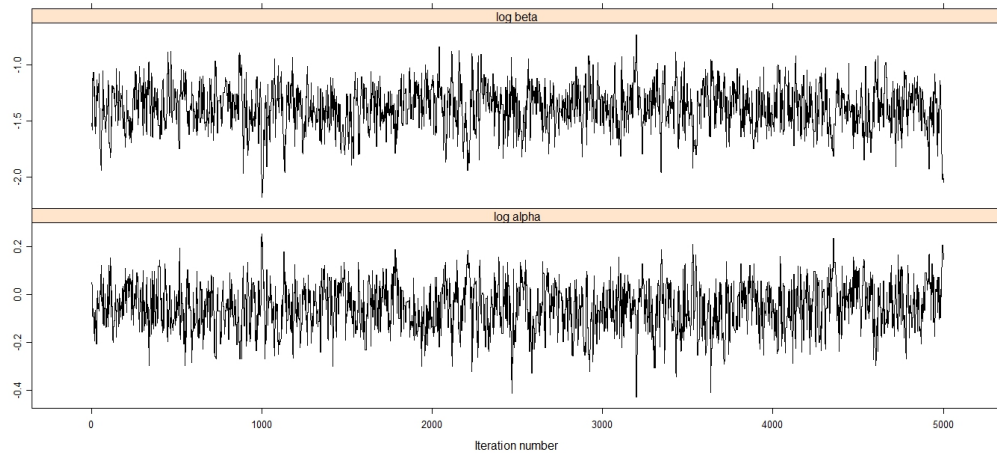
Figure 4.21: The trace plot of the simulated draws for log α and log β from the Markov Chain Monte Carlo method.
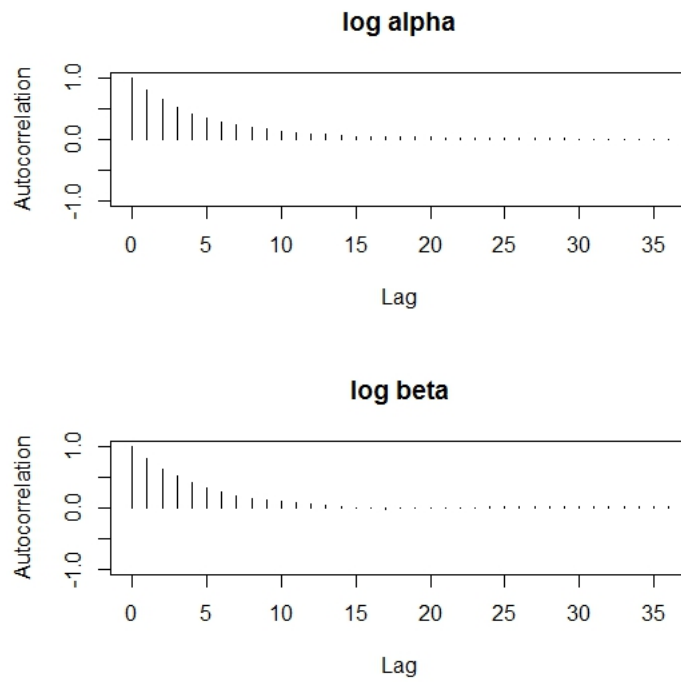

Figure 4.22: The autocorrelations plot of the simulated draws for log α and log β from the Markov Chain Monte Carlo algorithm.

Chapter 5

Conclusion

For recent years, many authors have been developed different approaches for generating lifetime models. These developments such as extension or modified forms of the lifetime models which can add more flexibility in data modelling. Generally, on the parameter estimation, the number of parameters can be crucial, and how to estimate the parameter easily can be another factor. Usually, a distribution with more than three parameters is unfavourable because it is mathematically hard to obtain the hazard rate function. This thesis is considered some Bayesian estimation methods that estimate the TN distribution which has defined by Sarhan et. al [12]. The significant of the TN distribution is the ability to fit monotone and non monotone data sets which are popular in the lifetime distributions. We used two different estimation methods to make inference on unknown parameters of the TN distribution which is described in chapter 2. On the classical approach, the maximum likelihood estimation is provided as well as the first, the second derivatives, and the fisher information matrix are presented which were addressed by Sarhan et. al [12]. On Bayesian estimation, we used two Bayesian approaches namely the rejection sampling algorithm (RS) and Markov Chain-Monte Carlo (MCMC). Also, we assumed the gamma prior distribution of unknown parameters. Algorithms of two approaches are presented. Chapter 3 describes the simulation study which is addressed to evaluate the performance of the estimation methods. We examined the mean square error (MSE), the mean bias, and the coverage probability. In testing the mean bias and the MSE of estimates, the bias tends to zero when the sample size is increased for the three methods, and the MSE decrease when the sample size is increased. Chapter 4 studies three real data sets which are applied to analyze the TN model. The TN model was estimated by maximum likelihood, the RS, the MCMC, and the results are presented. The thesis is concluded in Chapter 5.

5.1   Future Research

In the Bayesian analysis, we assumed that all hyperparameters are equal to 0.001. From posterior analysis, one can select the values of the hyperparametres in order to reflect the nature of the parameters α and β. Also, we may apply different types of data sets, such

as censored data, progressively censored data, and regression. This might help for further study on the Bayesian analysis of the two parameters.

Bibliography

[1] Aarset, M. V. (1987). How to identify bathtub hazard rate. IEEE Trans. Rel. 36(1):106-108.

[2] Albert, J. (2009). Bayesian Computation with R. Springer Science & Business Media.

[3] Jörgensen, B. (1982). Statistical Properties of the Generalized Inverse Gaussian Distribution, vol. 9, Springer, Heidelberg, Germany.

[4] Casella, G., and Berger, R. L. (2002). Statistical inference. Vol. 2. Pacific Grove, CA: Duxbury.

[5] Deshpande, J. V., and S. G. Purohit. "Life Time Data: Statistical Models and Methods. Series on Quality." Reliability and Engineering Statistics 11 (2005).

[6] Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2014). Bayesian Data Analysis. Vol. 2. Boca Raton, FL, USA: Chapman & Hall/CRC.

[7] Gupta, R. D., and Kundu, D. (2001). Generalized exponential distribution: different method of estimations. Journal of Statistical Computation and Simulation 69.4 (2001): 315-337.

[8] Hamada, M. S., Wilson, A., Reese, C. S. and Martz, H. (2008). Bayesian Reliability. Springer Science Business Media.

[9] Pham, H., and C. Lai, C. (2007). On recent generalizations of the Weibull distribution. IEEE Transactions on Reliability 56.3: 454-458.

[10] Sarhan, A. M., and Kundu, D. (2009). Generalized linear failure rate distribution. Communications in StatisticsTheory and Methods 38.5: 642-660.

[11] Sarhan, A. M., and Apaloo, J. (2013). Exponentiated modified Weibull extension distribution." Reliability Engineering   System Safety 112: 137-144.

[12] Sarhan, Ammar M., Lotfi Tadj, and David C. Hamilton. "A new lifetime distribution and its power transformation." Journal of Probability and Statistics 2014 (2014).

[13] Stollmack, S. and Harris, C. M. (1974). Failure-rate analysis applied to recidivism data. Operations Research, vol. 22, no. 6, pp. 1192-1205.