BAYESIAN ACCELERATED LIFE TESTS: EXPONENTIAL AND WEIBULL MODELS

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematical Statistics

Department of Statistics

Rhodes University

by

Sharkay Ruwade Izally

(Student number: 610I3478)

March 2016

Supervisor: Dr. L. Raubenheimer

Declaration

I, the undersigned, declare that the work contained in this thesis is my own work, except for references specifically indicated in the text, and that I have not previously submitted it elsewhere for degree purposes.

S. R. Izally

Date

Abstract

Reliability life testing is used for life data analysis in which samples are tested under normal conditions to obtain failure time data for reliability assessment. It can be costly and time consuming to obtain failure time data under normal operating conditions if the mean time to failure of a product is long. An alternative is to use failure time data from an accelerated life test (ALT) to extrapolate the reliability under normal conditions. In accelerated life testing, the units are placed under a higher than normal stress condition such as voltage, current, pressure, temperature, to make the items fail in a shorter period of time. The failure information is then transformed through an accelerated model commonly known as the time transformation function, to predict the reliability under normal operating conditions. The power law will be used as the time transformation function in this thesis. We will first consider a Bayesian inference model under the assumption that the underlying life distribution in the accelerated life test is exponentially distributed. The maximal data information (MDI) prior, the Ghosh Mergel and Liu (GML) prior and the Jeffreys prior will be derived for the exponential distribution. The propriety of the posterior distributions will be investigated. Results will be compared when using these non-informative priors in a simulation study by looking at the posterior variances. The Weibull distribution as the underlying life distribution in the accelerated life test will also be investigated. The maximal data information prior will be derived for the Weibull distribution using the power law. The uniform prior and a mixture of Gamma and uniform priors will be considered. The propriety of these posteriors will also be investigated. The predictive reliability at the use-stress will be computed for these models. The deviance information criterion will be used to compare these priors. As a result of using a time transformation function, Bayesian inference becomes analytically intractable and Markov Chain Monte Carlo (MCMC) methods will be used to alleviate this problem. The Metropolis-Hastings algorithm will be used to sample from the posteriors for the exponential model in the accelerated life test. The adaptive rejection sampling method will be used to sample from the posterior distributions when the Weibull model is considered.

Keywords: Accelerated life testing, Adaptive rejection sampling, Bayesian inference, Exponential distribution, Gibbs sampler, Gosh Mergel and Liu prior, Jeffreys prior, Log-concave, Maximal data information prior, Metropolis-Hastings sampler, Power law, Uniform prior, Weibull distribution.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to acknowledge and express my gratitude to my supervisor Dr. Raubenheimer. This thesis would not have been possible without her meaningful advice, patience and motivation. I can not thank her enough for all the efforts, support and guidance. Dr. Raubenheimer spent a lot of time, including weekends on this thesis and provided me with meaningful feedback in her comments regarding my work. I am grateful to Dr. Raubenheimer for sharing her excellent knowledge with me. I also want to thank my colleagues, Prof. Sarah Radloff, Mutsa and Jeremy from the department of Statistics who have been very supportive and encouraging throughout this journey. I thank my parents, Norman and Lydia for all the words of encouragement and prayers, my siblings, Heinrich and Sharday for always showing interest and my niece Shakira.

I would like to express my gratitude to the Rhodes University Levenstein bursary for funding received for the first year of my Masters and the SASA/NRF once-off research award received during the second year of my Masters.

Lastly, I would like to thank God for guiding me and pulling me through to complete my thesis.

List of Abbreviations

ALT	Accelerated Life Tests
ARS	Adaptive Rejection Sampling
CDF	Cumulative Distribution Function
DIC	Deviance Information Criterion
EST	Environmental Stress Testing
GML	Ghosh Mergel and Liu Prior
HALT	Highly Accelerated Life Tests
MCMC	Markov Chain Monte Carlo Methods
MDI	Maximal Data Information Prior
MH	Metropolis-Hastings
MLE	Maximum Likelihood Estimation
MTTF	Mean Time To Failure
PDF	Probability Density function
SSALT	Step-Stress Accelerated Life Testing
STRIFE	Stress Life Testing

Notation

x	observation from random variable X
$f(x)$	pdf of random variable X
$F(x)$	cdf of random variable X
$R(x)$	reliability function
$\lambda(x)$	failure rate
$E(X)$	mean of random variable X
$V(X)$	variance of random variable X
R_{eact}	reaction rate
T_{abs}	absolute temperature
L	a quantifiable measure such as the mean, median, mode, quantiles, etc.
A_F	acceleration factor
E_a	activation energy
log	natural logarithm
$\pi(\theta)$	prior of θ
$\pi(\theta \vert data)$	posterior distribution of θ
$I(\theta)$	Fisher information matrix
$P(\mathbf{X}_{t+1} \vert \mathbf{X}_t)$	transition kernel
$q(.\vert X_t)$	proposal distribution
D	domain of $f(x)$
$\mathbb{R}$	real space

Chapter 1

Introduction

1.1 Overview

Reliability theory is used when we have failure data to predict the life characteristics of a component under its normal operating conditions. It may be time consuming and costly to wait for a component to fail before we have failure data, and modern products are quite often designed to operate without failure for many years. An alternative method to use is accelerated life testing (ALT), whereby instead of waiting for a component to fail, a range of single stresses are applied to the component which allows failure to occur much quicker. The accelerated failure times will then have to be extrapolated to the components failure under normal operating conditions. This is achieved via the time transformation function. The time transformation function relates the failure of a product to the stresses involved. The choice of a time transformation function depends on the types of stresses involved. The two most commonly used time transformation functions are the power law and the Arrhenius law. The power law is mainly used when the stress involved is voltage and the Arrhenius law is mainly used when the stress is thermal in nature. The Bayesian approach will be used . We will specifically look at non-informative priors, namely the Jeffreys prior, the Gosh Mergel and Liu prior, the uniform prior and the maximal data information prior. These priors are discussed in Section 2.6. The objective Bayesian approach, that is, when a non-informative (vague) prior is used, often leads to a posterior distribution which is an unknown form, therefore Markov Chain Monte Carlo methods need to be used to sample from the posterior distribution. We will make use of the Metropolis-Hastings sampler, the Gibbs sampler and the adaptive rejection sampler to sample from the unknown posterior distributions. These methods are discussed in Section 2.9. When using non-informative priors one needs to investigate the propriety of the posterior distributions, else any inferences made will be invalid and give misleading results. The properness of the posteriors will be investigated in this thesis. Two lifetime models will be considered, the exponential distribution and the Weibull distribution. The exponential distribution is a commonly used model in reliability theory. Its applications are limited since it can only work with constant failure rates. A more flexible model is the Weibull model, we will consider this model as well since it does not require a constant failure rate, and it can model failure data more accurately.

1.2 Objectives

The main objectives of this thesis can be summarised as follows:

- to derive the Jeffreys prior using the power law for the exponential distribution;
- to derive the GML prior using the power law for the exponential distribution;
- to derive the MDI prior using the power law for the exponential distribution;
- to show the properness of the Jeffreys, GML and MDI posterior distributions using the power law for the exponential distribution;
- to obtain the marginal and conditional posteriors using the Jeffreys, GML and MDI priors under the power law for the exponential model;
- to obtain the pre-posterior variance in order to find the optimal design in an ALT for the exponential model and comparing the performance when the Jeffreys, GML and the MDI priors are used;
- to derive the MDI prior for the Weibull distribution using the power law time transformation function;
- to obtain the posterior distribution of the Weibull model under the power law when a uniform prior is used and showing that it is proper;
- to show that the conditional distributions of the Weibull model when the uniform prior is used are all log-concave;
- to show that the conditional distributions of the Weibull model when the MDI prior is used are all log-concave;
- Computing the predictive reliability under use-stress for the Weibull model and the deviance criterion using the various priors.

1.3 Contributions

Given the objectives, we can summarise the contribution of the thesis in the field of objective Bayesian statistics as the following:

- the derivation of the Jeffreys prior by Jeffreys (1961) for the exponential ALT model using the power law and showing that the posterior is proper.

- the derivation of the GML prior using the method by Ghosh et al. (2011) for the exponential ALT model using the power law and showing that the posterior is proper.
- the derivation of the MDI prior, which was developed by Zellner (1971), for the exponential ALT model using the power law and showing that the posterior is proper.
- the derivation of the MDI prior for the Weibull model under the power law.
- showing the posterior for the Weibull model under the power law when a uniform prior is used, is proper.

As far as we know, most of these non-informative priors have not been derived and used in the above mentioned settings. Erkanli & Soyer (2000) used independent Gamma priors for the exponential model. Zhang (1997) considered an approximate Bayesian approach for step-stress life testing under the exponential model. Tojeiro et al. (2004) considered independent normal priors for the exponential power law model. Kim et al. (2009) derived non-informative priors for the scale parameter of the exponential distribution in an accelerated life test setup, but they used a tampered random variables model. Soyer et al. (2008) considered a subjective Bayesian approach for the Weibull model. Xu & Tang (2011) derived the Jeffreys prior and reference prior for the Weibull model in an ALT setting. Mazzuchi et al. (1997) considered a linear Bayesian inference Weibull model.

1.4 Outline

Accelerated life testing is used when one is interested in failure data in a shorter period of time than usual. This is achieved by exposing the units of an item under severe stress levels. There are various models that can be used in ALT. We will consider the power law as the time transformation function in this thesis and consider the following non-informative priors: the Jeffreys prior, GML prior, the MDI prior and the uniform prior.

In **Chapter 3** the power law will be considered. Using the exponential distribution as a lifetime model, three non-informative priors will be derived. Using this time transformation function, the Jeffreys prior, GML prior and the MDI prior will be derived. The properness of the resulting posterior distributions will also be investigated. The work in this chapter is an extension of some of the work done in Erkanli & Soyer (2000). They considered Gamma priors for certain designs and investigated the posterior variance for each design. A similar approach will be followed in this chapter, but three non-informative priors will be used. Erkanli & Soyer (2000) used the Gibbs sampler to evaluate the pre-posterior variance. The three designs from Erkanli & Soyer (2000) will be considered and the performance of the non-informative priors will be compared by looking at their pre-posterior variances. The Metropolis-Hastings sampler will be used in this chapter, since the full conditional posteriors are not available.

In **Chapter 4** the power law will again be considered. Using the Weibull distribution as a lifetime model, two non-informative priors will be considered. Using this time transformation function, the MDI prior will be derived. The other non-informative prior that will be considered is the uniform prior. The properness of the resulting posterior distribution when using the uniform prior will also be investigated. The work in this chapter is an extension of the work done by Soyer et al. (2008) and the work done by Mazzuchi et al. (1997). Soyer et al. (2008) considered the parametric ALT model, using Gamma priors, a hierarchical exchangeable Bayesian model and a Markov dynamic model. They computed the predictive reliability at the use-stress environment for the various models. Mazzuchi et al. (1997) made use of linear Bayesian methods. A similar approach to that of Soyer et al. (2008) will be followed in this chapter, but two non-informative priors will be used, where they used Gamma priors for the parametric model. As in Soyer et al. (2008) the adaptive rejection sampling method will be used. Concluding remarks and avenues for future research will be discussed in **Chapter** 5.

The following **software** will be used in this thesis: for the computations in Chapter 3, the exponential distribution, MATLA*B*® (MATLAB, 2014) will be used; for computations in Chapter 4, the Weibull distribution, OpenBUGS (Lunn et al., 2000) and *R*® R Core Team (2014) will be used.

Appendix A contains some additional results for Chapter 3 and the derivation of the Fisher information matrix for the exponential model using the power law.

Appendix B contains derivations of log-concavity for the conditional distributions of the posterior distribution using the uniform prior and the MDI prior under the power law for the Weibull model in Chapter 4.

Appendix C contains MATLAB® code for Chapter 3

Appendix D contains the OpenBUGS code for Chapter 4.

Chapter 2

Literature Review

Reliability testing makes use of analysing times to failure data under normal operating conditions in order to quantify the life characteristics of a product or component. It can be difficult to obtain times to failure data if the lifetime or mean time to failure of the product is long. Most modern products are designed to operate without any failure for many years. Therefore under normal operating conditions, only a few units will fail in a test of practical length. An alternative approach is to use methods that will cause the products to fail quicker, these methods are referred to as accelerated life tests.

2.1 Background

2.1.1 Reliability

Let T be the time at which a failure event occurs, then the reliability function is given by

$$R(t) = P(T \geq t), \tag{2.1}$$

where $T \geq 0$, $R(t) \geq 0$, and $\lim_{t\to\infty} R(t) = 0$. For a given value of $t > 0$, $R(t)$ is the probability that the time to failure is greater than or equal to t. The probability that a failure occurs before time t is given by the cumulative distribution function (CDF) of the failure distribution, and is defined as

$$F(t) = 1 - R(t), \tag{2.2}$$

where $F(0) = 0$, and $\lim_{t\to\infty} F(t) = 1$. The probability density function (PDF) is commonly defined in reliability engineering as

$$f(t) = \frac{dF(t)}{dt} = -\frac{dR(t)}{dt}. \tag{2.3}$$

Another function commonly used, is the failure rate or hazard rate function, $\lambda(t)$, and is given by

$$\lambda(t) = \frac{f(t)}{1-F(t)} \tag{2.4}$$

$$= \frac{f(t)}{R(t)}. \tag{2.5}$$

The hazard rate can be interpreted as follows: suppose we have an item with lifetime T, knowing it has survived until time t and we are interested in knowing the probability that it does not survive for an additional time Δt. This is given by $P\{T \in (t, t+\Delta t | T > t)\}$. Now

$$\begin{aligned} P\{T \in (t,t+\Delta t|T>t)\} &= \frac{P\{T \in (t,t+\Delta t)\}}{p\{T>t\}} \\ &\approx \frac{f(t)\Delta t}{1-F(t)} = \lambda(t)\Delta t \end{aligned}$$

and this represents the conditional probability density that a $t-\text{year}$ old item will fail. The failure rate uniquely determines the distibution F. This will be shown as it was done in Ross (2007). From Equation 2.4 we can write $\lambda(t)$ as follows:

$$\lambda(t) = \frac{\frac{d}{dt}F(t)}{1-F(t)}. \tag{2.6}$$

Integrating both sides of Equation 2.6 with respect to t gives

$$\log(1-F(t)) = -\int_0^t \lambda(t)dt + c \tag{2.7}$$

where c is an unknown constant. Equation 2.7 can then be written as

$$1-F(t) = \exp\{c\}\exp\left\{-\int_0^t \lambda(t)dt\right\}. \tag{2.8}$$

Let $t=0$ then it implies that $c=0$ and thus

$$F(t) = 1-\exp\left\{-\int_0^t \lambda(t)dt\right\}. \tag{2.9}$$

Equation 2.9 shows how the cumulative distribution function can be determined using the failure rate. A common measurement of reliability is the mean time to failure (MTTF) which is given by,

$$\begin{aligned} MTTF = E(T) &= \int_0^\infty tf(t)dt \\ &= \int_0^\infty R(t)dt. \end{aligned} \tag{2.10}$$

2.1.2 Types of data

When dealing with time to failure data, often not all the information is available. Lifetime data can therefore be seperated into two types: complete data, whereby all information is available or censored data, where some of the information is missing. The commonly used type of censoring is right censoring. The term right censored implies that the event of interest is to the right of our data point. For example, if we tested seven units and only four had failed by the end of the test, we would have right censored data for the three units that did not fail and if these units were to keep on operating, the failure would occur at some time after our data point. We can further distinguish between different types of right censoring schemes.

- Type I censoring: the event is observed only if it occurs prior to some specified time and the remaining units are censored after that time.
- Type II censoring: the number of events observed is predetermined and the remaining units in the study are censored.
- Random censoring: this type of censoring takes place when we are interested in estimation of the marginal distribution of some event but some of the units under study may experience some competing event which causes them to be removed from the study which results in the event of interest being unobservable for these units and they get random right censored at that time as pointed out in Klein & Moeschberger (2003).

2.1.3 Accelerated life testing

Accelerated life testing (ALT) involves the acceleration of failures of a product to get information quicker on its life distribution. The units being tested are run under severe conditions to ensure quicker failure than they would have under normal operating conditions, Ebeling (1997) . ALT can further be divided into two types, qualitative ALT and quantitative ALT. In qualitative ALT, the objective is to identify failures or failure modes with no attempt to make any predictions about the lifetime of the product. In quantitative ALT, the sole purpose is to make predictions on the life of the product and obtain its life characteristics at normal use conditions.

2.1.3.1 Qualitative Accelerated life testing

A qualitative ALT involves testing units at higher than usual stress variables. If the unit survives then it passes the test. There are more specific names of these tests such as HALT (for highly accelerated life tests), STRIFE (stress life) and EST (environmental stress testing) as defined in Escobar & Meeker (2006). Qualitative accelerated tests is mainly used to identify probable failure modes and it is important to study and analyse the root cause and assess whether it will occur in actual use or not. When it has been established that this cause of failure will in fact occur in actual use then the product design needs to be changed or the manufacturing process needs to be looked at to eliminate such causes. Qualitative tests are not used to quantify the life characteristics of a product but rather they provide valuable information as to the types of stress levels one may wish to implement in a quantitative test.

2.1.3.2 Quantitative Accelerated life testing

A quantitative ALT involves testing units at higher than usual stress variables such as temperature, pressure, voltage, etc. Unlike qualitative ALT, the purpose of quantitative ALT is to quantify the life characteristics of a product and obtain reliability information such as the mean time to failure of a product under use conditions or the probability of failure under use conditions. In order to perform an accelerated life testing data analysis, one needs to choose an appropriate life distribution. The exponential, Weibull and lognormal distributions are the most popular distributions used in life data analysis. The exponential and Weibull distribution will be considered in this thesis.

2.2 Life distributions for Accelerated life testing

2.2.1 Exponential Distribution

A continuous random variable T is said to have an exponential distribution with parameter λ, $\lambda > 0$, if its probability density function is given by

$$f(t) = \begin{cases} \lambda \exp\{-\lambda t\}, & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{2.11}$$

The exponential distribution is often used in reliability since it is a one parameter model which makes it easy to work with mathematically and it has a property that makes it easy to analyse. This property is known as the memoryless property and it basically means that if the lifetime of an item is exponentially distributed then if an item has been in use for some time, it is the same as a new item in regards to the amount of time remaining until the item fails, as defined in Ross (2007). The mean of the exponential distribution is given by

$$E(T) = \frac{1}{\lambda}$$

and the variance is given by

$$V(T) = \frac{1}{\lambda^2}.$$

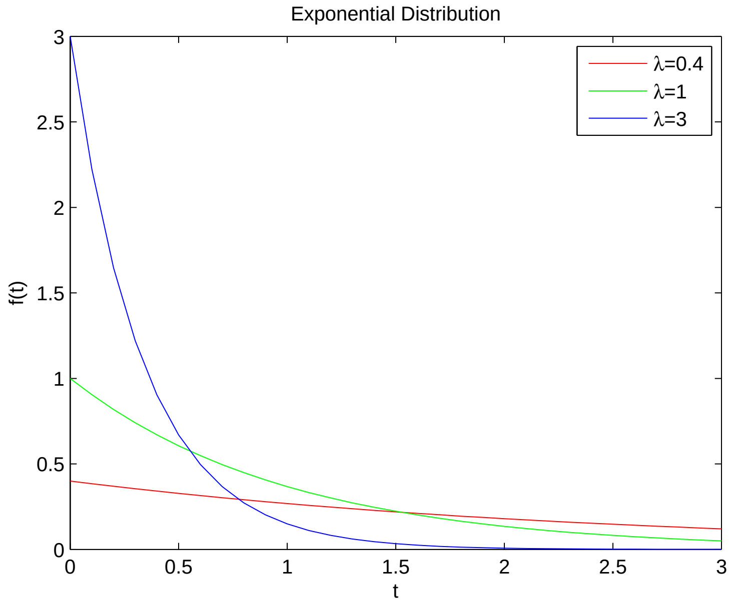


Figure 2.1: Exponential distribution for various values of parameter λ.

From Figure 2.1 we see that the exponential PDF has only one shape and it stretches to the right as the value of λ decreases. Also, as the values of t increases, $f(t)$ tends to zero. The reliability function of the exponential distribution is given by

$$\begin{aligned} R(t) &= 1 - F(t) \\ &= \exp\{-\lambda t\}. \end{aligned} \tag{2.12}$$

From Equation 2.5 we have that the failure rate $\lambda(t) = \lambda$. Thus the failure rate of an exponential distribution is constant and this limits the use of the exponential model in practice. For example, it would not be appropriate to use the exponential distribution to model the reliability of an external laptop hard drive since the constant failure rate of the exponential distribution requires that the hard drive is equally likely to crash when used the first time as it is when used after two years or so. This is clearly not a valid assumption, so in practice, care must be taken when using the exponential distribution as a model in reliability engineering.

2.2.2 Weibull Distribution

A continuous random variable T is said to have a Weibull distribution with shape parameter β and scale parameter λ, $\beta > 0$, $\lambda > 0$, if its probability density function is given by

$$f(t) = \begin{cases} \beta \lambda t^{\beta-1} \exp\left\{-\lambda t^{\beta}\right\}, & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{2.13}$$

The reliability function of the Weibull distribution is

$$\begin{aligned} R(t) &= 1 - F(t) \\ &= \exp\left\{-\lambda t^{\beta}\right\}. \end{aligned} \tag{2.14}$$

The mean of the Weibull distribution is given by

$$E(T) = \lambda^{-\frac{1}{\beta}} \Gamma\left(1 + \frac{1}{\beta}\right), \tag{2.15}$$

where $\Gamma(n) = \int_0^{\infty} t^{n-1} \exp(-t)\, dt = (n-1)!$ for any positive integer n, and the variance is given by

$$V(T) = \frac{1}{\lambda^{\frac{2}{\beta}}} \left\{\Gamma\left(1 + \frac{2}{\beta}\right) - \left[\Gamma\left(1 + \frac{1}{\beta}\right)\right]^2\right\}. \tag{2.16}$$

The failure rate is given by $\lambda(t) = \beta t^{\beta-1}\lambda$. The Weibull distribution is widely used as a lifetime distribution because it is flexible enough to reflect different failure rates. When the shape parameter β is greater than one, then the failure rate is an increasing function, when the shape parameter is less than one then it has a decreasing failure rate function. For $\beta = 1$, the Weibull distribution becomes the same as the exponential distribution with a constant failure rate.

2.3 Accelerated Models-Time transformation functions

After a model is fitted to the accelerated failure times using data collected from high stress conditons, an acceleration model is required. The acceleration model or time transformation function describes the relationship between the lifetimes and the stresses involved. Depending on the stress involved for example pressure, temperature, voltage, there are various kinds of accelerated models. Well-known models include the Arrhenius, Eyring and the inverse power law. These are the typical life stress models used in ALT, see Escobar & Meeker (2006). The inverse power law also referred to as the power model will be the life stress model used in this thesis.

2.3.1 Arrhenius Relationship

The Arrhenius relationship is used when the stress factor is thermal in nature. Particularly the model describes the effect of temperature on the rate of a simple chemical reaction. This relationship is defined as

$$R_{eact}(T) = \theta \exp\left\{-\frac{E_a}{k \times T_{abs}}\right\} \tag{2.17}$$

where R_{eact} is the reaction rate, T_{abs} is the absolute temperature in kelvin (K), k is the Boltzman's constant, $k = 8.6171 \times 10^{-5}$ in units of electronvolt per kelvin (eVK^{-1}). E_a is the activation energy and θ is an unknown non-thermal constant, these are both material characteristics. In a simple one step chemical reaction, the activation energy is described as the minimum amount of energy required for a reaction to take place as described in Escobar & Meeker (2006). The Arrhenius life stress model as given by Soyer et al. (2008) is

$$L(S_i) = \exp\{\theta_1 - \theta_2/S_i\}, \tag{2.18}$$

where L represents some quantifiable measure, such as the mean life, characteristic life, median life, etc. $\theta_1 > 0$ and θ_2 are the model parameters to be determined. S_i is the i^{th} stress level. In applications of ALT, the ratio of the acceleration characteristic between the use level and a higher stress level is defined as the acceleration factor and is given by

$$A_F = \frac{L_{use}}{L_{accelerated}}. \tag{2.19}$$

For the Arrhenius relationship, this factor is given by

$$\begin{aligned} A_F &= \frac{\exp\{\theta_1 - \theta_2/S_a\}}{\exp\{\theta_1 - \theta_2/S_u\}} \\ &= \exp\left\{\theta_2\left(\frac{1}{S_a} - \frac{1}{S_u}\right)\right\}, \end{aligned} \tag{2.20}$$

where S_u is the use stress level and S_a is the accelerated stress level.

2.3.2 Eyring Relationship

The Eyring model is also used when the stress factor is thermal in nature. It describes the effect that temperature has on a reaction rate. When written in terms of the reaction rate, the Eyring relationship as defined in Escobar & Meeker (2006) is given by

$$R(T) = \theta \times Z \times \exp\left\{-\frac{E_a}{k \times T}\right\} \tag{2.21}$$

where Z is a function of temperature depending on the specifics of the reaction dynamics. T is the absolute temperature in kelvin (K), k is the Boltzman's constant, $k = 8.6171 \times 10^{-5}$ in units of electronvolt per kelvin and E_a is the activation energy. Equation 2.21 is similar to the Arrhenius equation except that it accounts for the temperature that depends on the reaction dynamics.

2.3.3 Power law Relationship

The inverse power law model is used for non-thermal accelerated stresses such as voltage and pressure and is given by

$$L(S_i) = \theta_1 S_i^{\theta_2}, \tag{2.22}$$

where L represents some quantifiable measure, such as the mean life, characteristic life, median life, etc. S_i is the i^{th} stress level, $\theta_1 > 0$ and θ_2 are model parameters to be determined. The parameter θ_2 is a measure of the effect of the stress on life. As the absolute value of θ_2 increases, the greater the effect of stress. For the inverse power relationship, the acceleration factor is given by

$$\begin{aligned} A_F &= \frac{L_{use}}{L_{accelerated}} \\ &= \frac{\left(\theta_1 S_u^{\theta_2}\right)}{\left(\theta_1 S_a^{\theta_2}\right)} \\ &= \frac{(S_u)^{\theta_2}}{(S_a)^{\theta_2}}, \end{aligned} \tag{2.23}$$

where S_u is the use stress level and S_a is the accelerated stress level.

2.4 Designs of Accelerated Life Testing

Step-stress ALT (SSALT) is an alternative form of ALT, whereby instead of applying only a single higher than normal stress to the unit (constant-stress ALT), several stress levels are placed on the test units. In SSALT, the samples are tested under more than one level of stress and this allows for even quicker failures than constant-stress ALT. There are different types of ALT scenarios that are possible. There is the **fixed-stress** ALT, which is another name used for constant-stress ALT. When the stress levels are increased at each interval as time increases, then it is called **progressive step-stress** ALT. There is also **profile step-stress** ALT, this is when the stress levels can increase or decrease as time increases in the ALT. See Figure 2.2. The units will start at a lower or higher stress level and at a pre-determined time or failure number, the stress gets increased or decreased and the test continues. The test is stopped when all the units have failed, when a certain number of failures have been observed or after a certain time period has elapsed. A test with only two stress levels is called a simple step-stress accelerated life test, and if a test has more than two stress levels, it is known as a multiple step-stress accelerated life test as mentioned in Gouno & Balakrishnan (2001). If a test is ended at a pre-specified time, it is called Type-I censoring. A Type-II censoring test terminates the test after a certain number of failures were observed which causes limited applications in constant-stress ALT since the total testing time can not be controlled, (Lee & Pan, 2008).

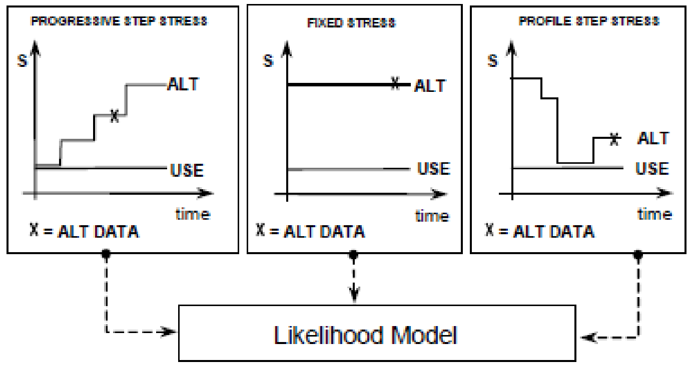


Figure 2.2: Different ALT scenarios, Van Dorp et al. (2006).

For a step stress pattern, there is a CDF for time to failure under the test from the observed data in the test. But one is usually interested in life under constant stress, not the life under step-stress. A model is needed to relate the distribution under step-stress to the distribution under constant stress. Nelson (1980) describes such a model by assuming that the remaining life of the test units depends only on the cumulative exposure that the units have seen and the units do not remember how such exposure was accumulated. This is known as the **cumulative exposure model**.

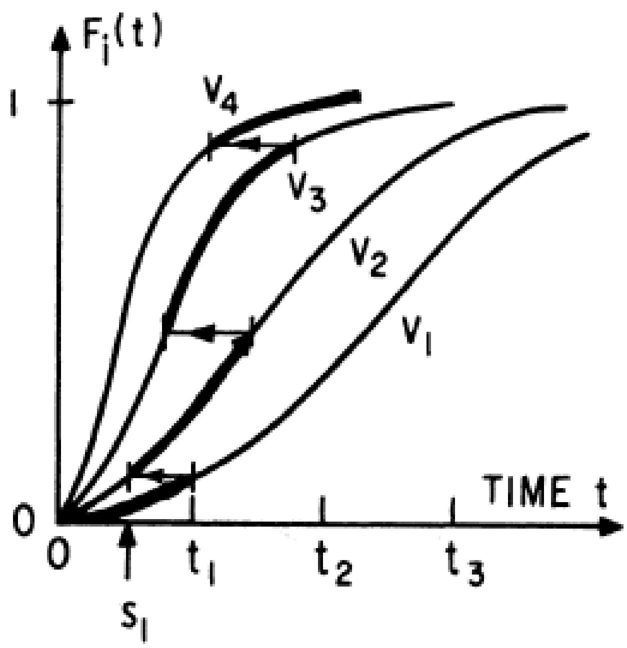


Figure 2.3: CDF under constant stress levels V_i, Nelson (1980).

Figure 2.3 shows such a model. It depicts four CDF's for the constant stresses V_i, $i = 1, ..., 4$. The arrows show that the items first follow the CDF for the stress level V_1 until time t_1. When the stress levels increase from V_1 to V_2, then the units that are still alive continue along the CDF for stress level V_2, starting at the accumulated fraction failed. The same is true when the stress level increases from V_2

to V_3, V_3 to V_4 and so on.

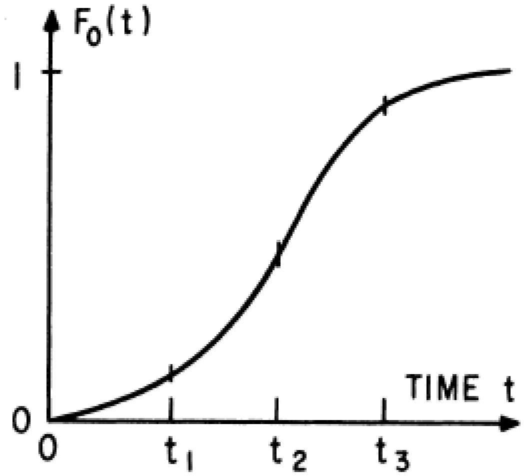


Figure 2.4: CDF for life under step-stress, Nelson (1980).

Figure 2.4 shows the CDF for a life under different step-stress patterns. If we compare Figure 2.4 to Figure 2.3, then we can see that $F_0(t)$ consists of segments of the CDF's for the constant stresses and this lead to the idea of the cumulative exposure model, see Nelson (1980) for further details. For a SSALT with m stress levels, the cumulative exposure model assumes that the failure times follow the following CDF ,

$$F_0(t) = \begin{cases} F_1(t), & 0 \leq t \leq \tau_1 \\ F_2(t-\tau_1+w_1), & \tau_1 \leq t \leq \tau_2 \\ F_3(t-\tau_2+w_2), & \tau_2 \leq t \leq \tau_3 \\ \cdots & \\ F_m(t), & \tau_m \leq t \leq \infty \end{cases} \tag{2.24}$$

where $F_0(t)$ is the CDF of time to failure under a particular step-stress pattern and τ_i is the time of changing the stress from the i^{th} stress level to the $(i+1)^{th}$ stress level. $F_i(t)$ is the CDF under the i^{th} stress level, and w_i is the solution of $F_i(w_{i-1}) = F_{i-1}(t_{i-1} - t_{i-2} + w_{i-2})$.

2.5 Maximum Likelihood Estimation

The method of maximum likelihood is a well known tool for parameter estimation given the sample data. Suppose that random variables $T_1, ..., T_n$ have a joint density $f(t_1, t_2, ..., t_n | \boldsymbol{\theta})$, where the vector $\boldsymbol{\theta} = (\theta_1, \theta_2, ... \theta_p)$, and θ_i for $i = 1, ...p$ are the unknown parameters. Assume T_i are independent and identically distributed random variables. Given the observed values $T_i = t_i$, for $i = 1, ..., n$, the likelihood function of $\boldsymbol{\theta}$ is given by

$$L(\boldsymbol{\theta}|t_1,t_2,\ ...,t_n) = \prod_{i=1}^{n} f(T_i|\boldsymbol{\theta}) \tag{2.25}$$

The maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$ is the value of $\boldsymbol{\theta}$ that maximises the likelihood or make the observed data most probable. It is usually easier to maximise the natural logarithm of the likelihood function and this is given by

$$\ell = \log L(\boldsymbol{\theta}|t_1,t_2,\ ...,t_n) = \sum_{i=1}^{n} \log\left(f(T_i|\boldsymbol{\theta}\right) \tag{2.26}$$

where "log" means the natural logarithm. In order to find MLE's the partial derivatives of the log likelihood function should be taken, equate them to zero and solve for the unknown parameters. This method works really well for large samples of data.

2.6 Bayesian inference

In the method of maximum likelihood, the unknown parameters are assumed to be fixed. The Bayesian approach assumes that these unknown parameters are random and follow some distribution and the uncertainties in the parameters are described by a prior distribution denoted by $\pi(\theta)$. The prior is formulated before the data are collected. The prior information can be based on historical data or expert opinion. Once the prior distribution is specified and with a given likelihood function, from Bayes rule we can express the posterior distribution as

$$\pi\left(\theta|data\right) = \frac{\pi(\theta)L(\theta|data)}{\int \pi(\theta)L(\theta|data)d\theta}. \tag{2.27}$$

Since the denominator of Equation 2.27 is not a function of θ, $\pi\left(\theta|data\right)$ can be written as

$$\pi\left(\theta|data\right) \propto \pi(\theta)L\left(\theta|data\right). \tag{2.28}$$

When Bayesian methods are used to estimate the parameters of the distribution of component lifetimes, it can be subjective in the sense that they rely highly on the priors for the parameter of the model with these priors also having their own parameters commonly known as hyperparameters. With little knowledge about the hyperparameters and limited time, the obtained priors can perform badly as mentioned by Xu & Tang (2011) and therefore objective Bayes will be used to derive priors for the unknown parameters. These priors are referred to as non-informative priors. They are generally used when little or no prior information is available. Non-informative priors are often called objective, vague or flat priors and they are often improper, meaning that they do not integrate to one. This does not pose a major problem as long as the posterior results in a proper density. There are a number of non-informitive priors that have been developed, Yang & Berger (1998) developed a catalogue that provides a list of non-informative priors that have been derived with their properties. They mainly put an emphasis on the Jeffreys prior and reference prior, although other approaches are also considered. We will consider four non-informative priors in this thesis: Zellner's maximal data information (MDI) prior, Jeffreys prior, a prior suggested by Ghosh et al. (2011), and a uniform prior.

2.6.1 Maximal Data Information (MDI) Prior

It is of interest that the data gives more information about the parameter than the information on the prior density, otherwise there would not be justification for the realisation of the experiment. The MDI prior was developed by Zellner (1971) and is chosen in such a way that the average information in the data density is maximised relative to that in the prior as pointed out in Zellner (1996). Thus the use of the MDI prior leads to an emphasis on the information in the data density or likelihood function. The MDI prior will be denoted by π_{MDI} and is given by

$$\pi_{MDI}(\theta) \propto \exp\left\{E\left(logL(\theta|t)\right\}\right. . \tag{2.29}$$

2.6.2 Jeffreys Prior

This was proposed by Jeffreys (1961), as a solution to the problem that the uniform prior does not yield an analysis invariant to choice of parameterization. It is invariant, meaning that if the Jeffreys prior in one parameterization is transformed to a different parameterization, then the transformed prior will be the Jeffreys prior in the new parameterization. This is a well known weak prior often used when there is little information about the unknown parameters. This prior will be denoted by $\pi_J(\theta)$ and is derived directly from the Fishers information matrix which will be denoted by $I(\theta)$. The Jeffreys prior is given by

$$\pi_J(\theta) \propto \sqrt{|I(\theta)|}. \tag{2.30}$$

2.6.3 GML Prior

Ghosh et al. (2011) set to find an objective prior which approximatley maximises the distance between the prior and the posterior. This was done under a general divergence criterion which they then based on a particular asymptotic expansion of this distance. The resulting prior was found to be the positive fourth root of the determinant of the Fishers information matrix. This prior will be denoted by π_{GML} and is given by

$$\pi_{GML}(\theta) \propto \sqrt[4]{|I(\theta)|}. \tag{2.31}$$

2.6.4 Uniform Prior

Bayes himself used a uniform prior for the binomial distribution. When using a uniform prior, one assigns a prior distribution to the unknown parameter on the interval (a,b) using the uniform distribution. The uniform prior will be denoted by π_U and is given by

$$\pi_U(\theta) \propto \text{constant}. \tag{2.32}$$

2.7 Accelerated Life Testing using the MLE Method

Nelson (1980) illustrated an application of estimating a model of life as a function of stress from step-stress data using the MLE method. The life distribution was assumed to be a Weibull distribution and the accelerated model was the inverse power law. The author suggested that the life stress model must take into account the cumulative effect of the applied stresses when dealing with SSALT data and proposed the cumulative exposure model to account for this effect and since then this model is widely used in SSALT. Balakrishnan et al. (2007) considered a simple step-stress model under Type-II censoring. The MLE's of the parameters were derived assuming the cumulative exposure model with lifetimes that are exponentially distributed. The exact distributions of the MLE's of the parameters were obtained using conditional moment generating functions. They used several methods to construct confidence intervals such as the parametric bootstrap, asymptotic distributions and the exact distributions which were all compared using a Monte Carlo simulation study. Kateri & Balakrishnan (2008) considered a simple SSALT model under Type-II censoring with Weibull distributed lifetimes. There was no accelerated model considered in this paper. However, they did assume a cumulative exposure model and derived the MLE's of the Weibull model and the observed Fisher information matrix. The authors also provided asymptotic and bootstrap confidence intervals for the parameters of the Weibull simple step-stress model.

2.8 Bayesian Accelerated Life Testing

The Bayesian approach to SSALT is becoming popular in literature. Van Dorp & Mazzuchi (2004) developed a general Bayes inference model for ALT assuming exponentially distributed failure times at each stress level. They considered regular life testing, constant stress testing, step-stress testing and profile stress testing. Type-I censoring and interval censoring were developed into the model. Their inference procedure relied on engineering judgment to specify the prior distrubution for inference and Markov Chain Monte Carlo (MCMC) methods were used to derive and estimate the posterior distribution. Lee & Pan (2008) presented a Bayes inference model for a SSALT using Type-II censoring. They assumed that the failure times at each stress level were exponentially distributed with a mean that is a log linear function of the natural stress level. An inverse beta distribution was used as a conjugate prior and they integrated engineering knowledge into the prior distribution of the parameters. It was found that by applying the Bayes approach to SSALT, the statistical precision of parameter inference was improved and the required number of samples was reduced. Van Dorp et al. (2006) focussed on comparing the different ALT designs such as constant stress ALT, profile ALT, progressive step-stress ALT and regressive ALT within a single Bayesian inference framework. They analysed the pre posterior variance of the use stress reliability based on a single failure over the course of the ALT for the different ALT designs. Van Dorp & Mazzuchi (2005) developed a Bayes inference model for ALT. The failures at constant stress was assumed to follow a Weibull distribution. Instead of specifying an accelerated model on the scale parameter of the Weibull distribution, they define a multivariate prior distribution on the scale parameter at various stress levels and a common shape parameter. Interval censoring and Type-I censoring was implemented under different ALT scenarios to do inference on the Bayes point estimates. In Erkanli & Soyer (2000), they assumed the life model was exponential and used the power law as the accelerated model and assigned independent priors on the unknown parameters in this model, one of them being a Gamma prior. The posterior variance was then computed to do an evaluation of the optimal design for the ALT.

Bayesian accelerated life testing has been investigated in a number of papers as can be seen above, however, in all these approaches they rely on some type of expert judgment to select a prior. Very little work has been done on objective Bayesian accelerated life testing, that is where a non-informative prior is used. Xu & Tang (2011) presents an objective Bayesian analysis of accelerated competing failure models under Type-I censoring. The time to failure due to a specific cause is assumed to follow a Weibull distribution. Two non-informative priors, the Jeffreys prior and the reference prior are used to estimate the parameters of the distribution of component lifetimes and Gibbs sampling procedures are used to estimate the posterior estimates of the parameters. The likelihood function was modified to ensure that the posterior densities are proper. Kim et al. (2009) presented an objective Bayesian accelerated life test where they derived non-informative priors for the scale parameter of the exponential distribution for data collected from a multiple step-stress ALT. The priors they derived were the reference prior and the probability matching prior. Simulations and examples were given to verify that their proposed Bayesian analysis performed well. In objective Bayesian accelerated life testing various non-informative priors were used, but we could not find anything in the literature where Zellner's MDI prior and the GML prior is used for inferences in ALT, and so we will mainly focus on developing Bayesian models using the MDI prior and GML prior in this thesis.

2.9 Sampling from the Posterior

When doing a Bayesian analysis using a conjugate prior the posterior will result in being a known density, but very often when non-informative priors are used then the posterior distribution results in an unknown and complicated density that is difficult to integrate or to obtain numerically. This can make it difficult to find the desired characteristics from the posterior in order to make Bayesian inferences. Fortunately, there are MCMC methods that are able to alleviate this problem and we will consider three of the most commonly used methods in MCMC, namely the Metropolis-Hastings sampler, the Gibbs sampler and the adaptive rejection sampling method.

2.9.1 Metropolis-Hastings Sampler

In order to understand the Metropolis-Hastings (MH) sampler, one needs to become familiar with the terminology of Markov chains. A Markov chain can be described as a sequence of random variables such that the next value or state depends only on the previous one. To introduce some notation, let $\mathbf{X}$ represent a vector of d random variables and it follows a distribution $\pi(\mathbf{x})$. For a Markov chain, we are generating a sequence of random variables, $\mathbf{X_0}, \mathbf{X_1}, \ldots$ such that the next state $\mathbf{X}_{t+1}$ with $t \geq 0$ is distributed according to $P(\mathbf{X}_{t+1}|\mathbf{X}_t)$, which is called a **transition kernel**. It is assumed that the transition kernel does not depend on t and this implies that the chain is time-homogenous, see Gamerman & Lopes (2006) for more details. Under certain conditions, the chain will forget its initial state and converge to the stationary distribution, denoted by φ, see Ross (2007) for these conditions and proofs of convergence to the stationary distribution. If the chain runs for b iterations and we assume that the sample points $\mathbf{X}_t$, $t = b+1, \ldots, n$ are distributed according to the stationary distribution, φ, then the first b iterations are discarded and the remaining $n-b$ samples can be used to obtain estimates via Monte Carlo integration. The number of discarded samples, b, is known as the **burn-in**. Metropolis-Hastings is a MCMC method used to simulate samples from a distribution that is unknown and analytically intractable. The method was explained and generalised in the context of statistics in a paper by Hastings (1970). The MH sampler obtains the state of the chain at $t+1$ by sampling an initial point $\mathbf{Z}$ from a proposal distribution, denoted by $q(.|X_t)$.When using the MH sampler, it is important to choose a proposal distribution that is easy to sample from. There are some conditions that need to be satisfied when choosing the proposal distribution, $q(.|X_t)$, namely irreducibility and aperiodicity. Irreducibility is when a Markov chain can reach any non-empty set with positive probability, no matter what starting point is used. Aperiodicity is when the chain does not oscillate between different states. These conditions are satisfied when the proposal distribution has a positive density in the domain of the target distribution. There is no specific rule of thumb on choosing the proposal distribution, but one has to keep these conditions in mind when using the MH sampler. Once the initial point has been sampled from the proposal distribution, it gets accepted as the next state of the chain with a probability given by

$$p(\mathbf{X_t}, \mathbf{Z}) = min\left\{1, \frac{\pi(\mathbf{Z})\, q(\mathbf{X_t}|\mathbf{Z})}{\pi(\mathbf{X_t})\, q(\mathbf{Z}|\mathbf{X_t})}\right\}. \tag{2.33}$$

If the point $\mathbf{Z}$ is not accepted, then the chain does not move and $\mathbf{X_{t+1}} = \mathbf{X_t}$. If we look at Equation 2.33 we see the target distribution $\pi(\mathbf{x})$ appears as a ratio so any constant from a distribution used in the algorithm will cancel out, this is one of the benefits of using MH sampler because one only needs to know a distribution upto a constant of proportionality to do the simulation study. The steps to the algorithm of the MH sampler are:

1. Initialise the chain to $\mathbf{X_0}$ and set $t = 0$.

2. Generate a point $\mathbf{Z}$ from the proposal distribution, $q(.|X_t)$.

3. Generate a U from a uniform distribution in the interval $(0,1)$.

4. If $U \leq p(\mathbf{X_t}, \mathbf{Z})$, then set $\mathbf{X_{t+1}} = \mathbf{Z}$,otherwise set $\mathbf{X}_{t+1} = \mathbf{X}_t$.

5. Set $t = t+1$ and repeat steps $2-5$.

2.9.2 Gibbs Sampler

The Gibbs sampler is a technique for generating random variables from a distribution or marginal distribution without having to calculate the density (Casella & George, 1992). When using the Gibbs sampler, it is important to note that one must know the full conditional distributions and they must be easy to sample from. As described in Casella & George (1992), lets consider a joint distribution $f(x_1, x_2)$ and using the same notation as in the previous section, consider $\mathbf{X}_t$ which will be a two element vector, $\mathbf{X}_t = (X_{t,1}, X_{t,2})$. First, we start with an initial point $\mathbf{X}_0 = (X_{0,1}, X_{0,2})$, then we generate a sample from $f(x_1|x_2)$ and $f(x_2|x_1)$. At each iteration, the elements of the random vector are obtained one at a time by alternately generating values from the conditional distribution. The algorithm for the bivariate Gibbs sampler is as follows:

1. Generate a starting point $\mathbf{X}_0 = (X_{0,1}, X_{0,2})$. Set $t = 0$.

2. Generate a point $X_{t,1}$ from
$$f(X_{t,1}|X_{t,2} = x_{t,2}).$$

3. Generate a point $X_{t,2}$ from
$$f(X_{t,2}|X_{t+1,1} = x_{t+1,1}).$$

4. Set $t = t+1$ and repeat steps $2-4$.

The fact that the Gibbs sampler requires that the full conditional distributions to be known limits its applications. There are hybrid algorithms that fixes this limitation such as the Metropolis-Hastings within Gibbs sampling, see Gamerman & Lopes (2006). There is also adaptive rejection sampling within Gibbs sampling, see Gilks & Wild (1992) for further details on adaptive rejection sampling within Gibbs sampling.

2.10 Rejection Sampling

2.10.1 Non-adaptive Rejection Sampling

Suppose we want to sample from an unknown complicated density $f(x)$. If there is a simpler method to sample from some density, $g(x)$, instead of the density we are seeking, then we can use $g(x)$ to generate from the desired density $f(x)$. A random number X is generated from $g(x)$ and the value is accepted with a probability proportional to $\frac{f(X)}{g(X)}$. If the constant c satisfies

$$\frac{f(x)}{g(x)} \leq c \tag{2.34}$$

for all x then we can generate the desired variates. The constant c is there to ensure that $g(x)$ is above $f(x)$. Points are generated from $cg(x)$ and those that are inside the curve $f(x)$ are accepted as coming from the desired density. This method is known as the rejection sampling method. The general algorithm as given by Ross (2007) is as follows:

1. Choose a density $g(x)$ that is easy to sample from.
2. Find a constant c such that Equation 2.34 is satisfied. This will usually require optimisation techniques from differential calculus.
3. Generate a random number Y from $g(x)$.
4. Generate a uniform random number U.
5. If
$$U \leq \frac{f(X)}{cg(X)},$$
then accept $X = Y$, else go to step 3.

The limitations in the rejection algorithm lies in finding the constant c and if we are dealing with a model that involves non-conjugacy in Bayesian inference then using the rejection algorithm at each step of the Gibbs sampler will be inefficient as mentioned in Gilks & Wild (1992).

2.10.2 Adaptive Rejection sampling

We will now give the adaptive rejection sampling (ARS) method as described by Gilks & Wild (1992). Let $g(x) = cf(x)$ where c is an unknown positive constant and differentiable everywhere in D, the domain of $f(x)$ and $h(x) = \log(g(x))$ is concave everywhere in D. See Figure 2.5 for an example of a log-concave function. Also, suppose $h(x)$ and $h'(x)$ were evaluated at k abscissae in $D: x_1 \leq x_2 \leq \ldots \leq x_k$. Let $T_k = \{x_i : i = 1, 2, ..., k\}$. Gilks & Wild (1992) defined the rejection envelope on T_k as

$\exp(u_k(x))$ where $u_k(x)$ is a piecewise linear upper hull formed from the tangents to $h(x)$ at the $x_i's$ in T_k, similar to Figure 2.5. For $j = 1,...,k-1$, the tangents at x_j and x_{j+1} intersect at

$$e_j = \frac{h(x_{j+1}) - h(x_j) - x_{j+1}h'(x_{j+1}) + x_j h'(x_j)}{h'(x_j) - h'(x_{j+1})} \tag{2.35}$$

So, for $x \in (e_{j-1}, e_j)$ and $j = 1,...,k$, we define

$$u_k(x) = h(x_j) + (x - x_j)\, h'(x_j) \tag{2.36}$$

where e_0 is the lower bound of D or $-\infty$ if D is not bounded from below and e_k is the upper bound of D or ∞, if not bounded above. They also defined

$$s_k = \frac{\exp(u_k(x))}{\int_D \exp(u_k(x'))\, dx'}. \tag{2.37}$$

The squeezing function is defined on T_k as $\exp(l_k(x))$ where $l_k(x)$ is a piecewise linear lower hull formed from the chords between adjacent abscissae in T_k , as depicted in the broken curve in Figure 2.5 that was taken from Gilks & Wild (1992). So for $x \in [x_j, x_{j+1}]$

$$l_k(x) = \frac{(x_{j+1} - x)\, h(x_j) + (x - x_j)\, h(x_{j+1})}{x_{j+1} - x_j} \tag{2.38}$$

for $j = 1,...,k-1$. For $x < x_1$ or $x > x_k$ they define $l_k(x) = -\infty$. The rejection envelope and the squeezing function are piecewise exponential functions. The concavity of $h(x)$ ensures that $l_k(x) \leq h(x) \leq u_k(x)$ for all x in D. The ARS method is more appealing than the non-adaptive rejection sampling since it does not require any maximisation techniques to find the supremum c and works well as long as the desired density function $f(x)$ is log-concave.

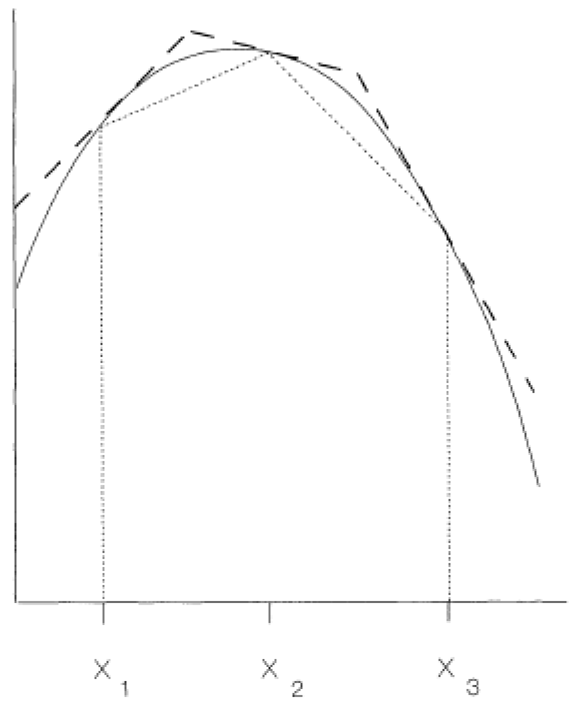


Figure 2.5: A concave log-density, bounded to the left, showing upper, and lower hulls based on three abscissae (x_1, x_2, x_3) : —,$h(x)$; - - -, upper hull; $\cdots$, lower hull.

Chapter 3

The Exponential Model

In this chapter the power law will be considered. The power law is a commonly used time transformation function in accelerated life testing. Using the exponential distribution as a lifetime model, three non-informative priors will be derived. Using this time transformation function, the Jeffreys prior, GML prior and the MDI prior will be derived. The properness of the resulting posterior distributions will also be investigated. The work in this chapter is an extension of some of the work done in Erkanli & Soyer (2000). They considered Gamma priors for certain designs and investigated the posterior variance for each design. A similar approach will be followed in this chapter, but three non-informative priors will be used. Erkanli & Soyer (2000) used the Gibbs sampler to evaluate the posterior variance. The Metropolis-Hastings sampler will be used in this chapter, since the full conditional posteriors are not available.

3.1 Introduction

We assume that life length X_i is exponential with failure rate λ_i, denoted by $X_i|\lambda_i \sim Exp(\lambda_i)$ with probability density function given by

$$f(x_i|\lambda_i) = \lambda_i \exp\{-\lambda_i x_i\} \tag{3.1}$$

where $x_i \geq 0$ and λ_i is the failure rate. It is common in accelerated life testing to assume an acceleration model that describes the relationship between the failure rate and the stress level. Depending on the stress involved, for example, temperature, pressure, voltage, etc., there are various kinds of acceleration models that can be used, see Escobar & Meeker (2006). In this chapter we will make use of the power law. Under the power law, the relationship between the failure rate and the stress level in the i^{th} testing environment is given by

$$\lambda_i = \alpha S_i^{\beta} \tag{3.2}$$

where $\alpha > 0$ and $\beta \in \mathbb{R}$ are the model parameters to be determined. This implies $X_i|\alpha,\beta,S_i \sim Exp(\alpha S_i^{\beta})$ and the density is given by

$$f(x_i|\alpha,\beta,S_i) = \alpha S_i^{\beta} \exp\left\{-\alpha S_i^{\beta} x_i\right\}. \tag{3.3}$$

The parameter β is a measure of the effect of the stress on life. As the absolute value of β increases, the greater the effect of stress. Let m denote the number of distinct stress levels used for the ALT and n_i denote the number of items tested at the stress level S_i and $n = \sum_{i=1}^{m} n_i$ is the pre determined number of items to be used in the ALT. We are interested in making inference about $\lambda_u = \alpha S_u^{\beta}$, the failure rate at the use stress environment. We assume that there is no censoring in the ALT. For fixed value of n, the design problem is to select $m \leq n$, the number of distinct stress levels, accelerated stress levels S_i, $i = 1,\ldots m$ and the number of items tested at each stress level, n_i, $i = 1,\ldots m$, in such a way that the expected loss is minimised. From Equation 3.3, the likelihood function is given by

$$L(\alpha,\beta|data) = \alpha^n \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right) \exp\left\{-\alpha \sum_{i=1}^{m} S_i^{\beta} T_i\right\} \tag{3.4}$$

where $T_i = \sum_{j=1}^{n_i} X_j$ and represents the total time on test at stress environment S_i. If $X_j \sim Exp(\lambda_i)$, then $T_i = \sum_{j=1}^{n_i} X_j$ will follow a Gamma distribution, $T_i \sim Gamma(n_i,\lambda_i)$. See proof in Appendix A.1. It can be shown that the Fisher information matrix, denoted by $I(\alpha,\beta)$ is given by

$$I(\alpha,\beta) = \begin{bmatrix} \frac{n}{\alpha^2} & \frac{1}{\alpha}\sum_{i=1}^{m} n_i logS_i \\ \frac{1}{\alpha}\sum_{i=1}^{m} n_i logS_i & \sum_{i=1}^{m} n_i (logS_i)^2 \end{bmatrix}.$$

See Appendix A.2 for the derivation of the Fisher information matrix.

3.2 Prior and Posterior Distributions

In Bayesian inference, the unknown parameters are random and follow some distribution and the uncertainties in the parameters are described by a prior distribution that is formulated before the failure data are collected. The prior information can be obtained from historical data or expert opinion. When there is little or no prior information available, objective Bayes can be used to derive priors for the unknown parameters.

3.2.1 Jeffreys Prior and Posterior Distribution

A well known prior to represent a situation with little information about the parameters was proposed by Jeffreys (1961). This prior, known as the Jeffreys prior denoted by $\pi_J(\alpha,\beta)$ is derived from the Fisher information matrix and is given by

$$\pi_J(\alpha,\beta) \propto \sqrt{|I(\alpha,\beta)|}. \tag{3.5}$$

Thus, from Equation 3.5, the Jeffreys prior for (α,β) under the power law is given by

$$\pi_J(\alpha,\beta) \propto \frac{1}{\alpha}. \tag{3.6}$$

The joint posterior distribution of (α,β) using the Jeffreys prior under the power law is given by

$$\pi_J(\alpha,\beta|data) \propto \alpha^{n-1}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum S_i^{\beta}T_i\right\}. \tag{3.7}$$

There is no general theory providing simple conditions under which an improper prior yields a proper posterior for a particular model, so this must be investigated case-by-case, as mentioned in Northrop & Attalides (2014). It is also mentioned in Northrop & Attalides (2014) that in order to ensure posterior propriety, we can make use of a diffuse prior or by truncating an improper prior by placing a lower bound on a unknown parameter. We will now show that the joint posterior distribution of (α,β) using the Jeffreys prior under the power law is proper.

Theorem 3.1. *Based on the Jeffreys prior $\pi_J(\alpha,\beta)$ and the observed data, the joint posterior distribution of (α,β) is proper.*

Proof. To show that $\pi_J(\alpha,\beta|data)$ is proper, the following should hold

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} c\alpha^{n-1}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum S_i^{\beta}T_i\right\}d\alpha d\beta = 1, \text{where } c \text{ is the normalising constant.}$$

For the above to be true, we need to show that

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} \alpha^{n-1}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum S_i^{\beta}T_i\right\}d\alpha d\beta < \infty. \tag{3.8}$$

Now

$$\int_{-\infty}^{\infty}\int_{0}^{\infty}\alpha^{n-1}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum S_i^{\beta}T_i\right\}d\alpha d\beta$$
$$=\Gamma(n)\int_{-\infty}^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-n}d\beta.$$

Since $0<\Gamma(n)<\infty$, we need to show that

$$\int_{-\infty}^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-n}d\beta<\infty. \tag{3.9}$$

But, to ensure the propriety of the posterior, we will place a lower bound on β *a priori*, constraining it to $(0,\infty)$. Hence, we need to show that

$$\int_{0}^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-n}d\beta<\infty. \tag{3.10}$$

To do this, we first need to show that

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-n}\leq\left(\prod_{i=1}^{m}S_i^{-\beta n_i}\right). \tag{3.11}$$

Now if Equation 3.11 holds, then this implies that

$$\left(\prod_{i=1}^{m}S_i^{2\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-n}\leq 1. \tag{3.12}$$

The LHS of Equation 3.12 can be written as

$$\frac{1}{\exp\left\{n\log\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)-\beta\sum_{i=1}^{m}\log\left(S_i^{2n_i}\right)\right\}}$$

which is less than or equal to 1 under the conditions that $\log\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)>\frac{\sum_{i=1}^{m}\log\left(S_i^{2\beta n_i}\right)}{n}$. Therefore,

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-n}\leq\left(\prod_{i=1}^{m}S_i^{-\beta n_i}\right)$$

for β, S_i, T_i, and$n_i>0$, $m>1$. The RHS of the inequality in Equation 3.11 can be written as

$$\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^{-1} = \left(\prod_{i=1}^{m} S_i^{n_i}\right)^{-\beta} = \beta^{1-1}\exp\left\{-\beta ln\left(\prod_{i=1}^{m} S_i^{n_i}\right)\right\}$$

$$\beta \sim Gamma\left(1, ln\left(\prod_{i=1}^{m} S_i^{n_i}\right)\right).$$

Since Equation 3.11 holds, the following will be true

$$\int_0^{\infty}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-n} d\beta \leq \int_0^{\infty}\left(\prod_{i=1}^{m} S_i^{-\beta n_i}\right) d\beta,$$

hence Equation 3.10 holds and this completes the proof. □

3.2.2 GML Prior and Posterior Distribution

The GML prior is derived from the Fishers information matrix so using Equation 2.31 we have the GML prior under the power law given by

$$\pi_{GML}(\alpha,\beta) \propto \frac{1}{\sqrt{\alpha}}. \tag{3.13}$$

The joint posterior distribution of (α,β) using the GML prior under the power law is given by

$$\pi_{GML}(\alpha,\beta|data) \propto \alpha^{n-\frac{1}{2}}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum_{i=1}^{m} S_i^{\beta} T_i\right\}. \tag{3.14}$$

The properness of the joint posterior of (α,β) using the GML prior under the power law is shown in the theorem that follows.

Theorem 3.2. *Based on the GML prior $\pi_{GML}(\alpha,\beta)$ and the observed data, the joint posterior distribution of (α,β) is proper.*

Proof. To show that $\pi_{GML}(\alpha,\beta|data)$ is proper, the following should hold

$$\int_{-\infty}^{\infty}\int_0^{\infty} c\alpha^{n-\frac{1}{2}}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum_{i=1}^{m} S_i^{\beta} T_i\right\} d\alpha d\beta = 1, \text{where } c \text{ is the normalising constant.}$$

For the above to be true, we need to show that

$$\int_{-\infty}^{\infty}\int_0^{\infty} \alpha^{n-\frac{1}{2}}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum_{i=1}^{m} S_i^{\beta} T_i\right\} d\alpha d\beta < \infty \tag{3.15}$$

Now

$$\int_{-\infty}^{\infty}\int_{0}^{\infty}\alpha^{n-\frac{1}{2}}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\exp\left\{-\alpha\sum_{i=1}^{m}S_i^{\beta}T_i\right\}d\alpha d\beta$$
$$=\Gamma(n+\frac{1}{2})\int_{-\infty}^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-\left(n+\frac{1}{2}\right)}d\beta$$

Since $0<\Gamma(n+\frac{1}{2})<\infty$, we need to show that

$$\int_{-\infty}^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-\left(n+\frac{1}{2}\right)}d\beta<\infty \tag{3.16}$$

But, to ensure the propriety of the posterior, we will place a lower bound on β *a priori*, constraining it to $(0,\infty)$. Hence, we need to show that

$$\int_{0}^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-\left(n+\frac{1}{2}\right)}d\beta<\infty. \tag{3.17}$$

To do this, we first need to show that

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-\left(n+\frac{1}{2}\right)}\leq\left(\prod_{i=1}^{m}S_i^{-\beta n_i}\right) \tag{3.18}$$

Now if Equation 3.18 holds, then this implies that

$$\left(\prod_{i=1}^{m}S_i^{2\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-\left(n+\frac{1}{2}\right)}\leq 1. \tag{3.19}$$

The LHS of Equation 3.19 can be written as

$$\frac{1}{\exp\left\{\left(n+\frac{1}{2}\right)\log\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)-\beta\sum_{i=1}^{m}\log\left(S_i^{2n_i}\right)\right\}}$$

which is less than or equal to 1 under the conditions that $\log\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)>\frac{\sum_{i=1}^{m}\log\left(S_i^{2\beta n_i}\right)}{\left(n+\frac{1}{2}\right)}$. Therefore,

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-\left(n+\frac{1}{2}\right)}\leq\left(\prod_{i=1}^{m}S_i^{-\beta n_i}\right)$$

for β, S_i, T_i, and$n_i>0$, $m>1$. The RHS of the inequality in Equation 3.18 can be written as

$$\begin{aligned}\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^{-1} &= \left(\prod_{i=1}^{m} S_i^{n_i}\right)^{-\beta} = \beta^{1-1}\exp\left\{-\beta ln\left(\prod_{i=1}^{m} S_i^{n_i}\right)\right\} \\ \beta &\sim Gamma\left(1, ln\left(\prod_{i=1}^{m} S_i^{n_i}\right)\right)\end{aligned}$$

Since Equation 3.18 holds, the following will be true

$$\int_0^\infty \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-\left(n+\frac{1}{2}\right)} d\beta \leq \int_0^\infty \left(\prod_{i=1}^{m} S_i^{-\beta n_i}\right) d\beta$$

hence Equation 3.17 holds and this completes the proof. □

3.2.3 Maximal Data Information (MDI) Prior and Posterior Distribution

It is of interest that the data gives more information about the parameter than the information on the prior density, otherwise there would not be justification for the realisation of the experiment. The MDI prior was developed by Zellner (1971) and is chosen in such a way that the average information in the data is maximised relative to that in the prior. Thus the use of the MDI prior leads to an emphasis on the information given in the data density or likelihood function. The MDI Prior will be derived using Equation 2.29 in the following theorem.

Theorem 3.3. *The MDI prior, $\pi_{MDI}(\alpha,\beta)$ for the exponential distribution under the power law is given by:*

$$\pi_{MDI}(\alpha,\beta) \propto \alpha^n \prod_{i=1}^{m} S_i^{\beta n_i}. \tag{3.20}$$

Proof. Firstly the log of the likelihood is given by

$$logL(\alpha,\beta|data) = log\left(\alpha^n \prod_{i=1}^{m} S_i^{\beta n_i}\right) - \alpha\sum_{i=1}^{m} S_i^{\beta} T_i \tag{3.21}$$

and the expected value of the log of the likelihood is

$$E\left[logL(\alpha,\beta|data)\right] = log\left(\alpha^n \prod_{i=1}^{m} S_i^{\beta n_i}\right) - n.$$

Now, using Equation 2.29 we have that the MDI prior for the parameters α and β is given by

$$\pi_{MDI}(\alpha,\beta) \propto \alpha^n \prod_{i=1}^{m} S_i^{\beta n_i}.$$

□

The joint posterior distribution of (α,β) using the MDI prior under the power law is

$$\pi_{MDI}(\alpha,\beta|data) \propto \alpha^{2n} \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2 \exp\left\{-\alpha \sum_{i=1}^{m} S_i^{\beta} T_i\right\}. \tag{3.22}$$

The MDI prior often results in being improper and an improper prior can lead to an improper posterior, which is not a desirable property when we want to do Bayesian inferences. We will investigate the propriety of the joint posterior distribution of (α,β) using the MDI prior in the following theorem.

Theorem 3.4. *Based on the MDI prior $\pi_{MDI}(\alpha,\beta)$ and the observed data, the joint posterior distribution of (α,β) is proper.*

Proof. To show that $\pi_{MDI}(\alpha,\beta|data)$ is proper, the following should hold

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} c\alpha^{2n} \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2 \exp\left\{-\alpha \sum_{i=1}^{m} S_i^{\beta} T_i\right\} d\alpha d\beta = 1 \text{, where } c \text{ is the normalising constant.}$$

For the above to be true, we need to show that

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} \alpha^{2n} \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2 \exp\left\{-\alpha \sum_{i=1}^{m} S_i^{\beta} T_i\right\} d\alpha d\beta < \infty. \tag{3.23}$$

Now

$$\begin{aligned} &\int_{-\infty}^{\infty}\int_{0}^{\infty} \alpha^{2n} \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2 \exp\left\{-\alpha \sum_{i=1}^{m} S_i^{\beta} T_i\right\} d\alpha d\beta \\ &= \Gamma(2n+1) \int_{-\infty}^{\infty} \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2 \left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-(2n+1)} d\beta. \end{aligned}$$

Since $0 < \Gamma(2n+1) < \infty$, we need to show that

$$\int_{-\infty}^{\infty} \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2 \left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-(2n+1)} d\beta < \infty. \tag{3.24}$$

But, to ensure the propriety of the posterior, we will place a lower bound on β *a priori*, constraining it to $(0,\infty)$. Hence, we need to show that

$$\int_0^{\infty}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^2\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-(2n+1)}d\beta<\infty. \tag{3.25}$$

It is easy to see that

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^2\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-(2n+1)}\leq\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^2. \tag{3.26}$$

Now the RHS of Equation 3.26 can be expressed as

$$\begin{aligned}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^2 &= \left(\prod_{i=1}^{m}S_i^{n_i}\right)^{2\beta}=\beta^{1-1}\exp\left\{-\beta ln\left(\prod_{i=1}^{m}S_i^{n_i}\right)^{-2}\right\}\\ \beta &\sim Gamma\left(1,ln\left(\prod_{i=1}^{m}S_i^{n_i}\right)^{-2}\right).\end{aligned}$$

But, the Gamma distribution can not have a negative parameter and therefore we first need to show that

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^2\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-(2n+1)}\leq\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^{-2} \tag{3.27}$$

so that we can ensure that all the parameters in the Gamma distribution is positive. Now if Equation 3.27 holds, then this implies that

$$\left(\prod_{i=1}^{m}S_i^{4\beta n_i}\right)\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-(2n+1)}\leq 1. \tag{3.28}$$

The LHS of Equation 3.28 can be written as

$$\frac{1}{\exp\left\{(2n+1)\log\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)-\beta\sum_{i=1}^{m}\log\left(S_i^{4n_i}\right)\right\}}$$

which is less than or equal to 1 under the conditions that $\log\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)>\frac{\sum_{i=1}^{m}\log\left(S_i^{4\beta n_i}\right)}{(2n+1)}$. Therefore,

$$\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^2\left(\sum_{i=1}^{m}S_i^{\beta}T_i\right)^{-(2n+1)}\leq\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^{-2}$$

for β, S_i, T_i, and$n_i>0$, $m>1$. The RHS of the inequality in Equation 3.27 can be written as

$$\left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^{-2} = \left(\prod_{i=1}^{m} S_i^{n_i}\right)^{-2\beta} = \beta^{1-1}\exp\left\{-\beta ln\left(\prod_{i=1}^{m} S_i^{n_i}\right)^{2}\right\}$$
$$\sim Gamma\left(1, ln\left(\prod_{i=1}^{m} S_i^{n_i}\right)^{2}\right).$$

Since Equation 3.27 holds, the following will be true

$$\int_0^\infty \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^{2}\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-(2n+1)} d\beta \leq \int_0^\infty \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^{-2} d\beta,$$

hence Equation 3.25 holds and this completes the proof. □

3.3 Marginal and Conditional Posterior Distributions

The posterior distributions in Equations 3.7, 3.14, and 3.22 cannot be obtained in closed form and therefore a Markov Chain Monte Carlo method can be used to sample from these posteriors by using their marginal and conditional posterior distributions. Therefore, we will obtain the marginal and conditional distributions for each of the posterior distributions. The marginal posterior of β using the Jeffreys prior is given by

$$\begin{aligned}\pi_J(\beta|data) &= \int \pi_J(\alpha,\beta|data)d\alpha \\ &= \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\int_0^\infty \alpha^{n-1}\exp\left\{-\alpha\sum_{i=1}^{m} S_i^{\beta} T_i\right\}d\alpha \\ &= \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-n}\Gamma(n), \end{aligned} \tag{3.29}$$

and the conditional posterior of α given β is

$$\pi_J(\alpha|\beta,data) = \frac{\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{n}}{\Gamma(n)}\alpha^{n-1}\exp\left\{-\alpha\sum_{i=1}^{m} S_i^{\beta} T_i\right\}. \tag{3.30}$$

The marginal posterior of β using the GML prior is given by

$$\begin{aligned}\pi_{GML}(\beta|data) &= \int \pi_{GML}(\alpha,\beta|data)d\alpha \\ &= \int_0^\infty \alpha^{n-\frac{1}{2}} \left(\prod_{i=1}^m S_i^{\beta n_i}\right) \exp\left\{-\alpha \sum_{i=1}^m S_i^\beta T_i\right\} d\alpha \\ &= \left(\prod_{i=1}^m S_i^{\beta n_i}\right) \left(\sum_{i=1}^m S_i^\beta T_i\right)^{-\left(n+\frac{1}{2}\right)} \Gamma\left(n+\frac{1}{2}\right), \end{aligned} \tag{3.31}$$

and the conditional posterior of α given β is

$$\pi_{GML}(\alpha|\beta,data) = \frac{\left(\sum_{i=1}^m S_i^\beta T_i\right)^{n+\frac{1}{2}}}{\Gamma\left(n+\frac{1}{2}\right)} \alpha^{n-\frac{1}{2}} \exp\left\{-\alpha \sum_{i=1}^m S_i^\beta T_i\right\}. \tag{3.32}$$

The marginal posterior of β using the MDI prior is given by

$$\begin{aligned}\pi_{MDI}(\beta|data) &= \int \pi_{MDI}(\alpha,\beta|data)d\alpha \\ &= \left(\prod_{i=1}^m S_i^{\beta n_i}\right)^2 \int_0^\infty \alpha^{2n} \exp\left\{-\alpha \sum_{i=1}^m S_i^\beta T_i\right\} d\alpha \\ &= \left(\prod_{i=1}^m S_i^{\beta n_i}\right)^2 \left(\sum_{i=1}^m S_i^\beta T_i\right)^{-(2n+1)} \Gamma\left(2n+1\right), \end{aligned} \tag{3.33}$$

and the conditional posterior of α given β is

$$\pi_{MDI}(\alpha|\beta,data) = \frac{\left(\sum_{i=1}^m S_i^\beta T_i\right)^{(2n+1)}}{\Gamma\left(2n+1\right)} \alpha^{2n} \exp\left\{-\alpha \sum_{i=1}^m S_i^\beta T_i\right\}. \tag{3.34}$$

The conditional posteriors given in Equations 3.30, 3.32 and 3.34 all follow a Gamma distribution.

3.4 Posteriors Under Censoring

It is common in reliability theory that we have censored or incomplete data. The likelihood function will differ depending on the type of censoring involved. Under the assumptions that the censoring times and lifetimes are independent, we give an expression for a general likelihood function for various types of censoring schemes by incorporating the following components :

- exact lifetimes - $f(x)$

- right-censored observations - $R(C_+)$
- left-censored observations - $(1-R(C_-))$
- interval-censored observations - $[R(-)-R(+)]$

The likelihood function as given in Klein & Moeschberger (2003) is

$$L \propto \prod_{i\in fail} f(x_i) \prod_{i\in +} R(C_+) \prod_{i\in -} (1-R(C_-)) \prod_{i\in I} [(R(-)-R(+))] \tag{3.35}$$

where $'+'$ indicates right censoring, $'-'$ indicates left censoring, $fail$ is the set of failure times, I is the set of interval censored observations and R is the reliability function. If we consider a pair of random variables (T,δ), where δ indicates that an event X occurred $(\delta=1)$ or wether an event has not occurred $(\delta=0)$ and T is equal to X if the event occurred and equal to C_+ if it is right censored, so $T=\min(X,C_+)$. Using this notation, if we have pairs $(T_i,\delta_i), i=1,...,n$ then the likelihood function under **Type-I censoring** can be written as

$$L=\prod_{i=1}^{n}[f(t_i)]^{\delta_i}[R(t_i)]^{1-\delta_i}. \tag{3.36}$$

We will now make use of Equation 3.36 to write down the likelihood function under **Type-I censoring** for the exponential distribution using the power law. From Equations 3.3 and 3.36, the likelihood of α and β given the observed data is given by

$$\begin{aligned} L(\alpha,\beta|data) &= \prod_{i=1}^{m}\prod_{j=1}^{n_i}\left[\alpha S_i^{\beta}\exp\left\{-\alpha S_i^{\beta}x_{ij}\right\}\right]^{\delta_i}\left[\exp\left\{-\alpha S_i^{\beta}x_{ij}\right\}\right]^{1-\delta_i} \\ &= \prod_{i=1}^{m}\left(\alpha S_i^{\beta}\right)^{\delta_i n_i}\exp\left\{-\alpha\sum_{i=1}^{m}S_i^{\beta}\sum_{j=1}^{n_i}x_{ij}\right\} \\ &= \alpha^{\sum_{i=1}^{m}\delta_i n_i}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^{\delta_i}\exp\left\{-\alpha\sum_{i=1}^{m}S_i^{\beta}T_i\right\}. \end{aligned} \tag{3.37}$$

Using Bayes rule and Equation 3.37 we find that the posterior distribution of α and β given the observed data using the Jeffreys prior under **Type-I censoring** is given by

$$\pi_J(\alpha,\beta|data) \propto \alpha^{\sum_{i=1}^{m}\delta_i n_i-1}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^{\delta_i}\exp\left\{-\alpha\sum_{i=1}^{m}S_i^{\beta}T_i\right\}. \tag{3.38}$$

The posterior of α and β given the observed data using the GML prior under **Type-I censoring** is given by

$$\pi_{GML}\left(\alpha,\beta|data\right)\propto\alpha^{\sum_{i=1}^{m}\delta_i n_i-\frac{1}{2}}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^{\delta_i}\exp\left\{-\alpha\sum_{i=1}^{m}S_i^{\beta}T_i\right\}. \tag{3.39}$$

The posterior of α and β given the observed data using the MDI prior under **Type-I censoring** is given by

$$\pi_{MDI}\left(\alpha,\beta|data\right)\propto\alpha^{\sum_{i=1}^{m}\delta_i n_i+n}\left(\prod_{i=1}^{m}S_i^{\beta n_i}\right)^{\delta_i+1}\exp\left\{-\alpha\sum_{i=1}^{m}S_i^{\beta}T_i\right\}. \tag{3.40}$$

3.5 Monte Carlo Method for a Given Design

Erkanli & Soyer (2000) obtained the pre-posterior variance to find the optimal design. The posterior variance, $V\left(\lambda_u|data\right)$, can be approximated by Monte Carlo sample averages, where $\lambda_u=\alpha S_u^{\beta}$ and S_u is the use stress level. The pre-posterior variance for a specific design can be calculated as follows

$$E_{data}\left[V\left(\lambda_u|data\right)\right]=\int\int V\left(\lambda_u|data\right)f\left(\mathbf{x}|\lambda_u\right)\pi\left(\lambda_u\right)d\lambda_u d\mathbf{x} \tag{3.41}$$

where $V\left(\lambda_u|data\right)$ will be evaluated via the Metropolis-Hastings sampler for a given design, $d=\{m,S_i,n_i,i=1,...,m\}$. Erkanli & Soyer (2000) used the Gibbs sampler to evaluate $V\left(\lambda_u|data\right)$ where they used the adaptive rejection sampling method of Gilks & Wild (1992), here the posterior should be log-concave. We will make use of the Metropolis-Hastings step instead, since for the Metropolis-Hastings, the requirement of log-concavity is not needed and one does not need the full conditional posterior.

Steps for the Monte Carlo evaluation of $E_{data}\left[V\left(\lambda_u|data\right)\right]$:

1. Specify the design, i.e. specify the number of stress levels, m, the size of each stress level, n_i, and the stress levels S_i, where $i=1,...,m$.

2. Generate (α,β) from the prior

$$\begin{aligned}\pi_J(\alpha,\beta) &\propto \frac{1}{\alpha}\\ \pi_{GML}(\alpha,\beta) &\propto \frac{1}{\sqrt{\alpha}}\\ \pi_{MDI}(\alpha,\beta) &\propto \alpha^n\prod_{i=1}^{n}S_i^{\beta n_i}.\end{aligned}$$

3. For each stress level, $i=1,...,m$, generate T_i from $f(T_i|\lambda_i)$. It is shown in Appendix A.1 that $T_i\sim Gamma\left(n_i,\lambda_i\right)$ where n_i is the number of items tested at stress level S_i and when using the power law $\lambda_i=\alpha S_i^{\beta}$.

4. Evaluate $V(\lambda_u|data)$ by using the Metropolis-Hastings sampler. First sample β from the marginal posterior $\pi(\beta|data)$, using 5000 Metropolis-Hastings samples. The MATLAB® function mhsample is used here. For the Jeffreys prior, the marginal posterior is

$$\pi_J(\beta|data) \propto \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-n}$$

and $\beta|data \sim Gamma\left(1, \log\left(\prod_{i=1}^{m} S_i^{n_i}\right)\right)$ is used as the proposal distribution. For the GML prior, the marginal posterior is

$$\pi_{GML}(\beta|data) \propto \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-\left(n+\frac{1}{2}\right)}$$

and $\beta|data \sim Gamma\left(1, \log\left(\prod_{i=1}^{m} S_i^{n_i}\right)\right)$ is used as the proposal distribution. For the MDI prior, the marginal posterior is

$$\pi_{MDI}(\beta|data) \propto \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)^2\left(\sum_{i=1}^{m} S_i^{\beta} T_i\right)^{-(2n+1)}$$

and $\beta|data \sim Gamma\left(1, \log\left(\prod_{i=1}^{m} S_i^{2n_i}\right)\right)$ is used as the proposal distribution. After obtaining β, α can be sampled from the conditional posterior $\pi(\alpha|\beta, data)$:

$$\begin{aligned}
\alpha|\beta, data_J &\sim Gamma\left(n, \sum_{i=1}^{m} S_i^{\beta} T_i\right)\\
\alpha|\beta, data_{GML} &\sim Gamma\left(n+\frac{1}{2}, \sum_{i=1}^{m} S_i^{\beta} T_i\right)\\
\alpha|\beta, data_{MDI} &\sim Gamma\left(2n+1, \sum_{i=1}^{m} S_i^{\beta} T_i\right).
\end{aligned}$$

Now $V(\lambda_u|data)$ can be determined, where $\lambda_u = \alpha S_u^{\beta}$.

5. Repeat steps $2-4$, $R = 2000$ times for the given design, and use the Monte Carlo average to compute the pre-posterior variance

$$\frac{1}{R}\sum_{r=1}^{R} V(\lambda_u|data).$$

Steps $1-5$ can be repeated for a different design. The design and prior yielding the smallest pre-posterior variance is the desired one.

3.6 Application to a Fixed Point Design

In this section, the fixed points design from Erkanli & Soyer (2000) will be considered. The following three designs will be considered:

Table 3.1: Fixed points design.

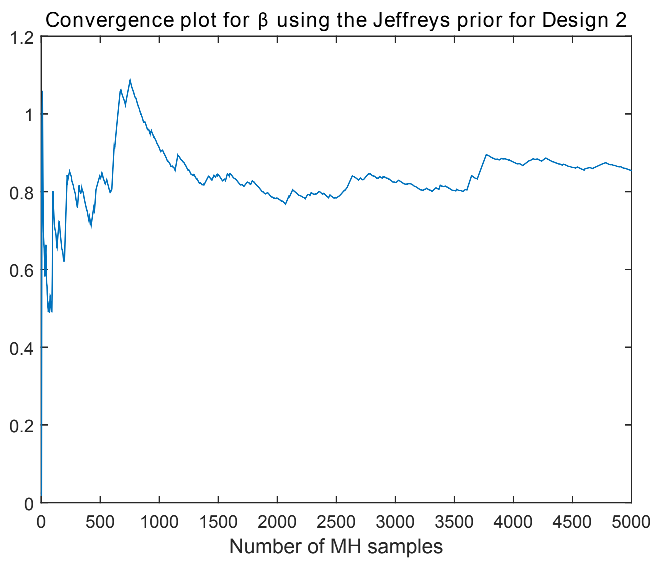
	Design 1	Design 2	Design 3
n_1	4	2	1
n_2	0	2	3
S_1	1.05	1.05	1.05
S_2	5.4	5.4	5.4

where $S_u = 1.05$ is the use-stress. Erkanli & Soyer (2000) assumed independent gamma priors for α and β, $\alpha \sim Gamma(20, 1000)$ and $\beta \sim Gamma(3, 1)$. Our focus in this chapter is to investigate the performance of the three non-informative priors, namely the Jeffreys prior, MDI prior and GML prior. See Appendix B.3 for the MATLA$B^{®}$ code. The results are given in Table 3.2.

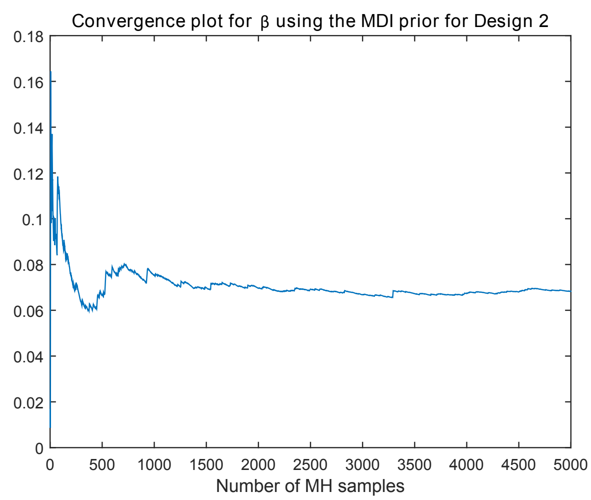
Table 3.2: Comparison of designs and priors, by looking at $E_{data}\left[V\left(\lambda_u|data\right)\right]$.

	Design 1	Design 2	Design 3
Jeffreys prior	2.1451	3.5437	7.7977
MDI prior	1.6904	1.7808	1.7652
GML prior	0.0106	0.0205	0.0483

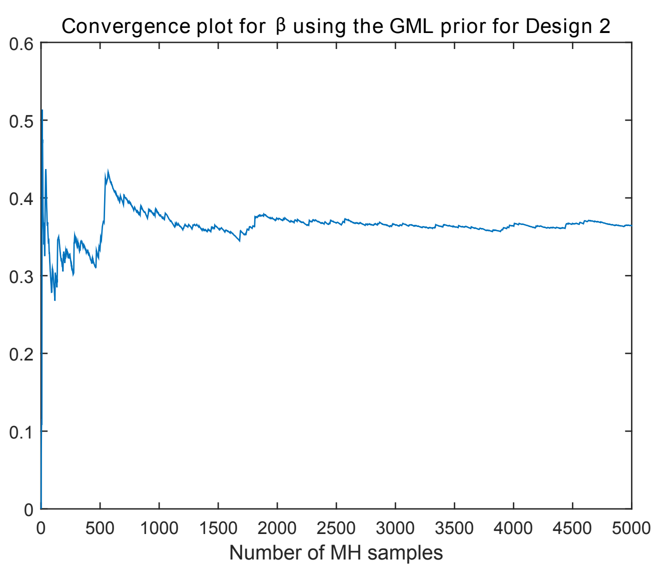
The desired design, is the one which yields the smallest value for $E_{data}\left[V\left(\lambda_u|data\right)\right]$. From Table 3.2 we see that design 1, which is a one-point design ($n_1 = 4$, $n_2 = 0$) performs better than the two-point designs, design 2 ($n_1 = 2$, $n_2 = 2$) and design 3 ($n_1 = 1$, $n_2 = 3$), when the Jeffreys prior, MDI prior and the GML prior is used. We also see that design 2 performs better than design 3 when the Jeffreys prior and GML prior is used, but when the MDI prior is used, design 2 does not perform better than design 3. It is clear that the GML prior yielded the smallest pre-posterior variance for all the three designs, and therefore performed the best. The MDI prior yielded smaller pre-posterior variances than the Jeffreys prior for all three designs. So, the Jeffreys prior did not perform as well as the GML prior and the MDI prior for this study. We will illustrate convergence plots for design 2 using each non-informative prior in Figure 3.1.



(a)



(b)



(c)

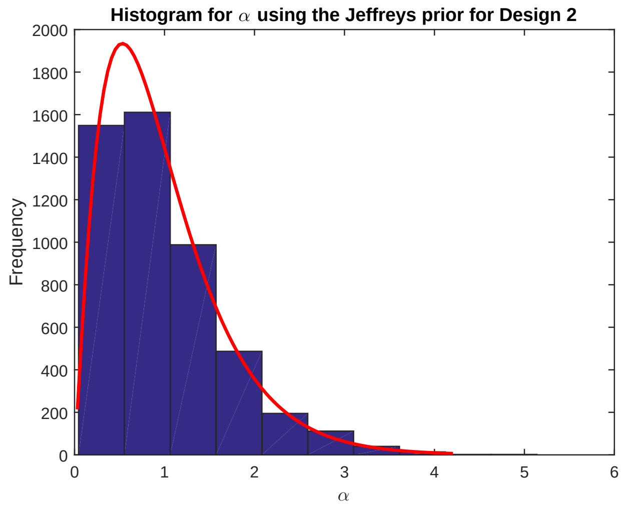
Figure 3.1: Convergence plots for design 2 using, (a) Jeffreys prior, (b) MDI prior and (c) GML prior.

In Figure 3.1 we present convergence plots for design 2 , when the Jeffreys prior (a), MDI prior (b) and the GML prior (c) is used. The number of simulated designs, $R = 2000$ is considered and the 5000 MH samples are shown for each prior used. We see from Figure 3.1 that all three priors that were used converged. The Jeffreys prior clearly converges slower than the MDI prior and the GML prior. The MDI prior and the GML prior both started converging around 1000 MH samples.

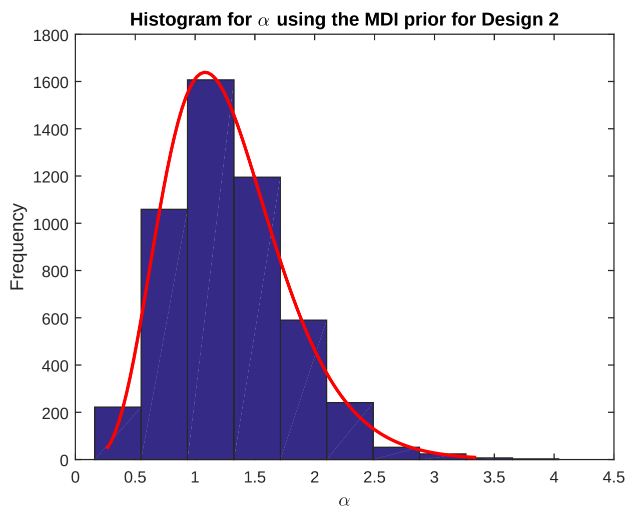
From Figure 3.2 (a) and (b), we see that the distribution for α for both the Jeffreys prior and MDI prior is around 1, implying that α for these priors does not contribute to the failure rate. From Figure 3.2 (c), the distribution of α for the GML prior is concentrated between 0.01 and 0.02, implying a decrease in failure rate at any stress level.

From Figure 3.3 (a), (b) and (c), we see that the distribution of β is skewed to the right for the Jeffreys prior, MDI prior and the GML prior. The distribution for β for the Jeffreys prior has values between 0 and 2.5. So for values of β less than 1, this would imply that there is a decrease in failure rate as the stress levels increase which is not what one would would expect to happen to the failure rate when stress increases. For values of β larger than 1, there will be an increase in the failure rate as the stress levels increase, which is expected. The distribution for β using the GML prior displays the same trend as the distribution for β for the Jeffreys prior. The distribution for β using the MDI prior has values between 0 and 1, which implies a decrease in failure rates for values less than 1 and no increase for the value 1, for any value of stress.

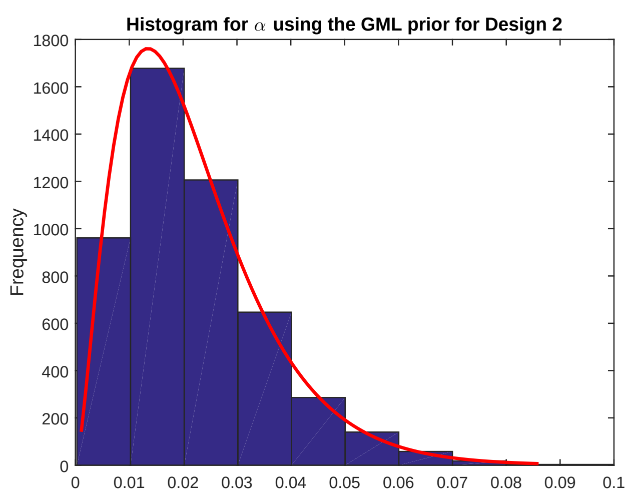
Figure 3.2 illustrates the conditional posterior of α given β for each non-informative prior used in design 2.



(a)



(b)



(c)

Figure 3.2: The histogram of the samples for α using, (a) Jeffreys prior, (b) MDI prior and (c) GML prior for design 2.

Figure 3.3 illustrates the marginal posterior of β for each non-informative prior used in design 2.

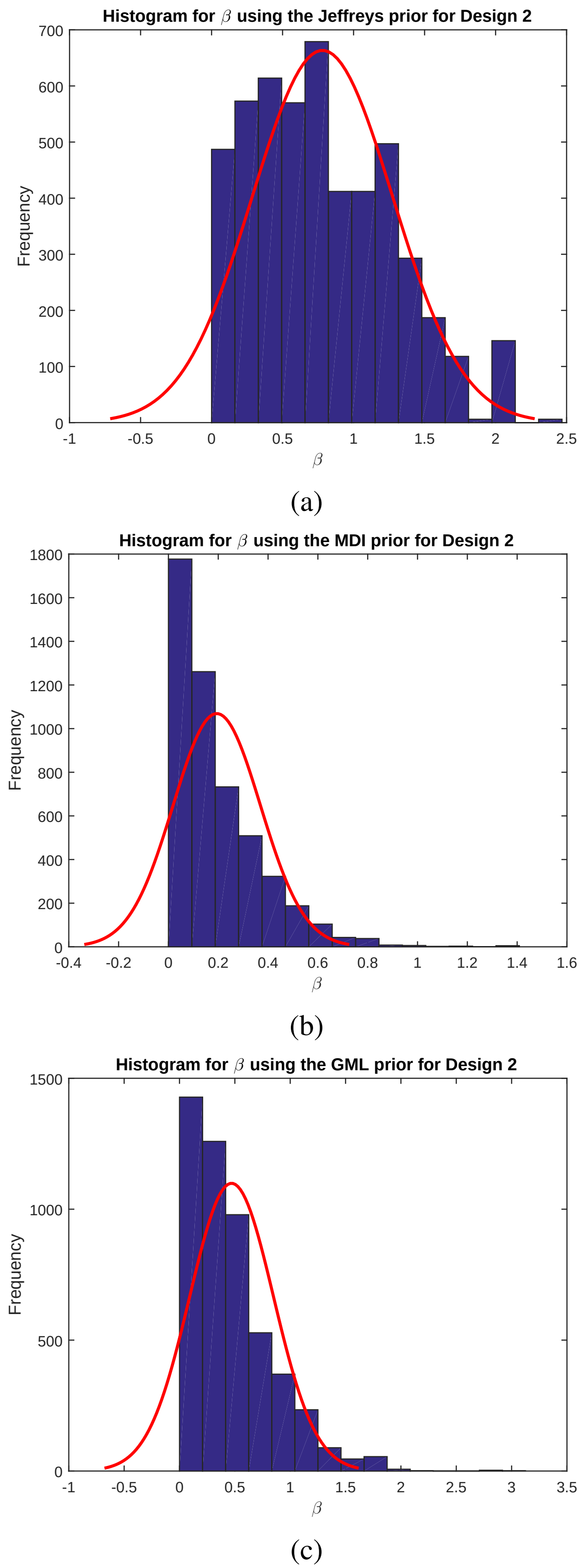
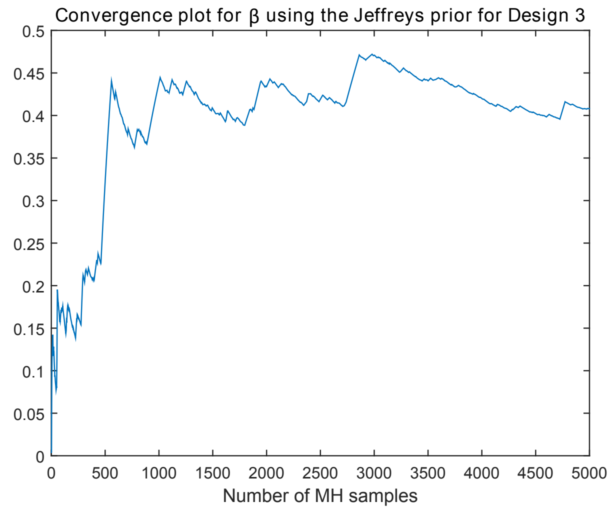
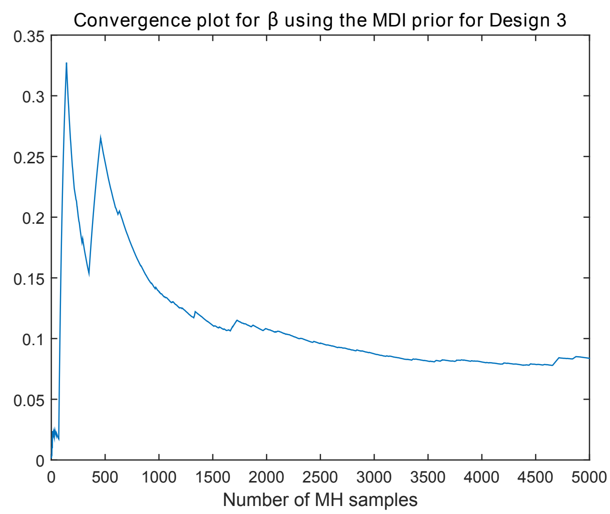


Figure 3.3: The histogram of the samples for β using, (a) Jeffreys prior, (b) MDI prior and (c) GML prior for design 2.

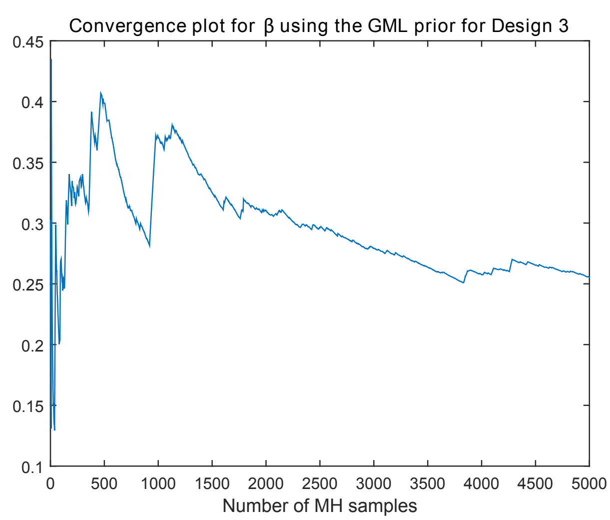
We will illustrate convergence plots from MH sampler for design 3 using all three non-informative priors used.



(a)



(b)



(c)

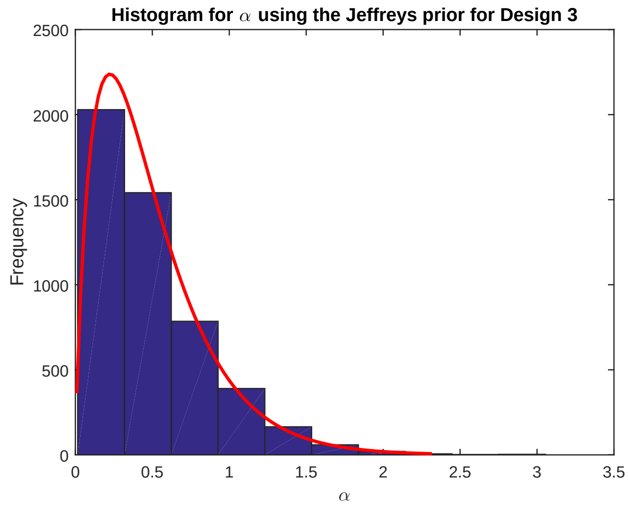
Figure 3.4: Convergence plots for design 3 using, (a) Jeffreys prior, (b) MDI prior and (c) GML prior.

In Figure 3.4 we present convergence plots for design 3 , when the Jeffreys prior (a), MDI prior (b) and the GML prior (c) is used. The number of simulated designs, $R = 2000$ is considered and the 5000 MH samples are shown for each prior used. We see from Figure 3.4 that all three priors that were used converged. We see The Jeffreys prior converges slower than the MDI prior and the GML prior. The MDI prior converged quicker than the GML prior.

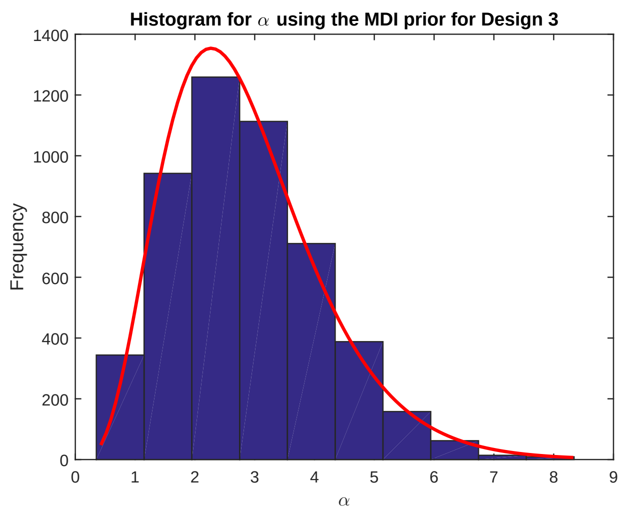
From Figure 3.5 (a) and (c), we see the distribution for α for both the Jeffreys prior and GML prior is around 0.2, implying a decreasing failure rate at any stress level. From Figure 3.5 (b), the distribution of α for the MDI prior is concentrated between 2 and 3, implying an increase in failure rate at any stress level. All three plots show that the conditional distribution of α given β is skewed to the right.

From Figure 3.6 (a), (b) and (c), we see that the distribution of β is skewed to the right for the Jeffreys prior, MDI prior and the GML prior. The distribution for β for the Jeffreys prior has values between 0 and 1.5. So for values of β less than 1, this would imply that there is a decrease in failure rate as the stress levels increase which is not what one would would expect to happen to the failure rate when stress increases. For values of β larger than 1, there will be an increase in the failure rate as the stress levels increase, which is expected. The distribution for β using the GML prior displays the same trend as the distribution for β for the Jeffreys prior. The distribution for β using the MDI prior has values between 0 and 0.8, which implies a decrease in failure rates for any value of stress.

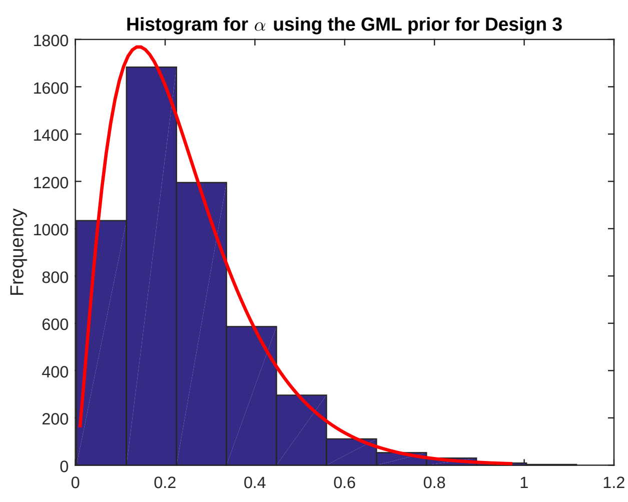
Figure 3.5 illustrates the conditional posterior of α given β for each non-informative prior used in design 3.



(a)



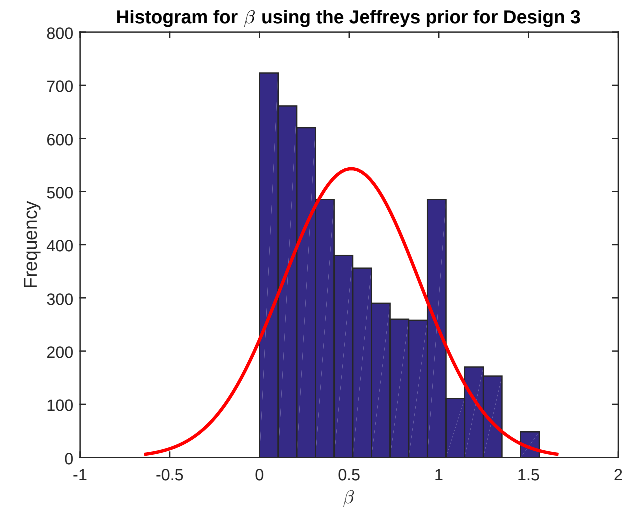
(b)



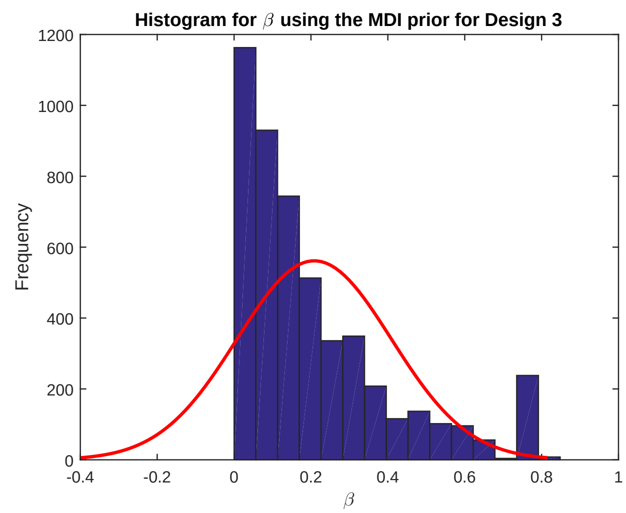
(c)

Figure 3.5: The histogram of the samples for α using, (a) Jeffreys prior, (b) MDI prior and (c) GML prior for design 3.

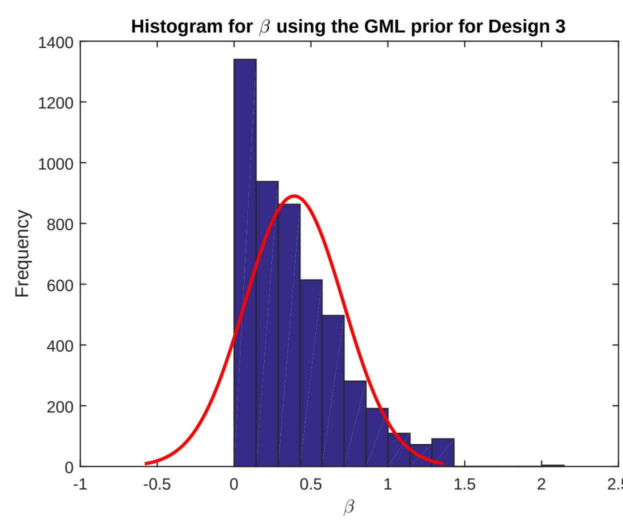
Figure 3.6 illustrates the marginal posterior of β for each non-informative prior used in design 2.



(a)



(b)



(c)

Figure 3.6: The histogram of the samples for β using, (a) Jeffreys prior, (b) MDI prior and (c) GML prior for design 3.

3.7 Conclusion

In this chapter we considered the exponential distribution as a lifetime model and used the power law as the time transformation function. The Jeffreys prior, GML prior and the MDI prior was derived and the posterior distribution using each prior considered was shown to be proper. The marginal and conditional posteriors were derived using each of the priors considered. We also derived the posterior distributions using the Jeffreys prior, MDI prior and GML prior under Type-I censoring. We then used a simulation study to find the optimal design using the different priors by looking at the pre-posterior variance for each design. The fixed points design from Erkanli & Soyer (2000) was used. The prior with the smallest pre-posterior variance is the one that performed the best. The GML prior had the smallest pre-posterior variance for each design that we considered and so performed the best. The MDI prior performed better than the Jeffreys prior for all three designs. Design 1 had the smallest pre-posterior variance regardless of which prior was used and therefore the optimal design was at the use-stress level, $S_u = 1.05$. As for the two-point designs, design 2 performed better than design 3 when the Jeffreys prior and GML prior is used. Design 3 only performed better than design 2 when the MDI prior was used.

Chapter 4

The Weibull Model

In this chapter the power law will again be considered. Using the Weibull distribution as a lifetime model, two non-informative priors will be considered. Using this time transformation function, the MDI prior will be derived. The other non-informative prior that will be considered is the uniform prior. It will be shown that the resulting posterior distributions are log-concave. The work in this chapter is an extension of the work done by Soyer et al. (2008) and the work done by Mazzuchi et al. (1997). Soyer et al. (2008) considered the parametric ALT model, using a mixture of uniform and Gamma priors, a hierarchical exchangeable Bayesian model and a Markov dynamic model. They computed the predictive reliability at the use-stress environment for the various models. Mazzuchi et al. (1997) made use of linear Bayesian methods . A similar approach to that of Soyer et al. (2008) will be followed in this chapter, but two non-informative priors will be used. As in Soyer et al. (2008) the adaptive rejection sampling method will be used.

4.1 Introduction

We assume that the life length X_i follows a Weibull distribution with scale parameter λ_i and shape parameter β. Assume that under the i^{th} accelerated test environment, the failure behaviour of the items can be described by a Weibull model with density given by

$$f(x_i|\lambda_i) = \beta \lambda_i x_i^{\beta-1} \exp\left\{-\lambda_i x_i^{\beta}\right\} \tag{4.1}$$

with scale parameter, $\lambda_i > 0$ and shape parameter $\beta > 0$. We will denote the above model by $X_i|\lambda_i, \beta \sim Wei(\lambda_i, \beta)$. The model above implies that the scale parameter λ_i, depends on the stress environment, but the shape parameter, β, does not. This is a common assumption in the literature, see Nelson (1980), Soyer et al. (2008) and Kateri & Balakrishnan (2008) . It is common to assume a functional relationship between the failure rate and the applied stress level. In this chapter we will consider the power law as the acceleration function. Under the power law, the relationship between the failure rate and the stress level in the i^{th} testing environment is given by

$$\lambda_i = \theta_1 S_i^{\theta_2} \tag{4.2}$$

where S_i denotes the i^{th} accelerated stress environment, $\theta_1 > 0$ and $\theta_2 \in \mathbb{R}$ are the unknown model parameters to be determined. The aging effect is captured by the shape parameter β, and the effect of the stress environment on the failure behaviour is captured by the parameters θ_1 and θ_2, as mentioned in Soyer et al. (2008). This implies that $X_i|\theta_1, \theta_2, \beta \sim Wei(\theta_1 S_i^{\theta_2}, \beta)$ and the density is given by

$$f(x_i|\theta_1, \theta_2, \beta) = \beta \theta_1 S_i^{\theta_2} x_i^{\beta-1} \exp\left\{-\theta_1 S_i^{\theta_2} x_i^{\beta}\right\}. \tag{4.3}$$

Let D_i denote the test data from the i^{th} accelerated stress environment, that is,

$$D_i = \{n_i, r_i, x_{i1}, ..., x_{ir_i}\}$$

where n_i is the number of items tested, r_i is the number of failures observed during the observation period, x_{ij} is the time to failure of the j^{th} item under the i^{th} environment, $j = 1, 2, \ldots, r_i \leq n_i$. Our main aim is to make inferences about the failure behaviour of the items at the use-stress environment, S_u. We assume that there is no censoring in the ALT, thus $r_i = n_i$. For mathematical convenience we will consider the likelihood function of λ_i and β from the i^{th} accelerated stress environment. Assuming conditional independence of the failure times x_{ij} given the stress levels S_i, and the parameters β, θ_1, and θ_2 and using Equation 4.3, the likelihood function is given by

$$L(\theta_1, \theta_2, \beta|data) = \left(\beta^{n_i} \theta_1^{n_i} S_i^{\theta_2 n_i}\right) \left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right) \exp\left\{-\theta_1 S_i^{\theta_2} \sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}. \tag{4.4}$$

The likelihood function when k stress levels are used is given by

$$L(\theta_1, \theta_2, \beta|data) = (\beta\theta_1)^{\sum_{i=1}^{k} n_i} \left(\prod_{i=1}^{k} S_i^{n_i\theta_2}\right) \left(\prod_{i=1}^{k}\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right) \exp\left[-\theta_1 \sum_{i=1}^{k} \left(S_i^{\theta_2} \sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]. \tag{4.5}$$

4.2 Priors and Posterior Distributions

4.2.1 Prior and Posterior from Soyer et al. (2008)

Soyer et al. (2008) assumed that $\theta_1 \sim Gamma(0.01, 0.01)$, $\theta_2 \sim Uni(0, 100)$ and $\beta \sim Uni(0, 10)$. It was further assumed that θ_1, θ_2, and β are independent. Under the power law, the joint prior distribution for $(\theta_1, \theta_2, \beta)$ is given by

$$\pi_{Soy}(\theta_1, \theta_2, \beta) \propto \theta_1^{0.01-1} \exp\{-0.01 \times \theta_1\} \times \frac{1}{100} \times \frac{1}{10}. \tag{4.6}$$

The joint posterior distribution is therefore given by

$$\begin{aligned}\pi_{Soy}(\theta_1,\theta_2,\beta|data) &\propto \beta^{\sum_{i=1}^{k} n_i}\theta_1^{\sum_{i=1}^{k} n_i+0.01-1}\left(\prod_{i=1}^{k} S_i^{n_i\theta_2}\right)\left(\prod_{i=1}^{k}\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\\ &\times \exp\left[-\theta_1\left(0.01+\sum_{i=1}^{k}\left(S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right)\right]. \quad (4.7)\end{aligned}$$

4.2.2 Uniform Prior and Posterior Distribution

The uniform prior denoted by $\pi_U(\theta_1,\theta_2,\beta)$ for the Weibull distribution under the power law at the i^{th} stress level is given by

$$\pi_U(\theta_1,\theta_2,\beta) \propto \text{constant}. \quad (4.8)$$

The joint posterior distribution of θ_1,θ_2 and β using the uniform prior for the Weibull model under the power law is given by

$$\pi_U(\theta_1,\theta_2,\beta|data) \propto \left(\beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}. \quad (4.9)$$

In Theorem 4.1 it will be shown that the posterior distribution when using the uniform prior is proper.

Theorem 4.1. *Based on the uniform prior $\pi_U(\theta_1,\theta_2,\beta)$, and the observed data, the joint posterior distribution of $(\theta_1,\theta_2,\beta)$ is proper.*

Proof. To show propriety of $\pi_U(\theta_1,\theta_2,\beta|data)$ the following should be true:

$$\int_0^\infty\int_{-\infty}^\infty\int_0^\infty c\left(\beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}d\theta_1 d\theta_2 d\beta = 1,$$

where c is the normalising constant. For the above to be true, we need to show that

$$\int_0^\infty\int_{-\infty}^\infty\int_0^\infty \left(\beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}d\theta_1 d\theta_2 d\beta < \infty. \quad (4.10)$$

Now

$$\begin{aligned}LHS &= \int_0^\infty\int_{-\infty}^\infty\int_0^\infty \left(\beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}d\theta_1 d\theta_2 d\beta\\ &= \int_0^\infty\int_{-\infty}^\infty\int_0^\infty \left(\beta^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\theta_1^{(n_i+1)-1}\exp\left\{-\theta_1\left[S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]\right\}d\theta_1 d\theta_2 d\beta\\ &= \int_0^\infty\int_{-\infty}^\infty \left(\beta^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\Gamma(n_i+1)\left[S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]^{-(n_i+1)} d\theta_2 d\beta\end{aligned}$$

But, to ensure the propriety of the posterior, we will place a lower bound on θ_2 *a priori*, constraining it to $(0,\infty)$. Hence, we need to show that

$$\int_0^\infty \int_0^\infty \left(\beta^{n_i} S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\Gamma\left(n_i+1\right)\left[S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]^{-(n_i+1)} d\theta_2 d\beta < \infty. \tag{4.11}$$

Since $0 < \Gamma\left(n_i+1\right)\left(\prod_{j=1}^{n_i} x_{ij}^{-1}\right) < \infty$, then it must be shown that

$$\int_0^\infty \int_0^\infty \left(\beta^{n_i} S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left[S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]^{-(n_i+1)} d\theta_2 d\beta < \infty. \tag{4.12}$$

Now

$$\begin{aligned} LHS &= \int_0^\infty \int_0^\infty \left(\beta^{n_i} S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left[S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]^{-(n_i+1)} d\theta_2 d\beta \\ &= \int_0^\infty \int_0^\infty \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} S_i^{-\theta_2} d\theta_2 d\beta \\ &= \int_0^\infty \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} \left(\log S_i\right)^{-1} d\beta. \end{aligned}$$

Since $0 < \left(\log S_i\right)^{-1} < \infty$, we need to show that

$$\int_0^\infty \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} d\beta < \infty \tag{4.13}$$

Now

$$\beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} \le \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{-\beta}\right) \le \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right),$$

for β, n_i, and $x_{ij} > 0$. Since

$$\begin{aligned} \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{-\beta}\right) &= \beta^{(n_i+1)-1}\exp\left\{\log\left(\prod_{j=1}^{n_i} x_{ij}^{-\beta}\right)\right\} \\ &= \beta^{(n_i+1)-1}\exp\left\{\log\left(\prod_{j=1}^{n_i} x_{ij}\right)^{-\beta}\right\} \\ &= \beta^{(n_i+1)-1}\exp\left\{-\beta\left(\log\left(\prod_{j=1}^{n_i} x_{ij}\right)\right)\right\} \end{aligned}$$

$$\beta \sim Gamma\left(n_i+1, \log\left(\prod_{j=1}^{n_i} x_{ij}\right)\right)$$

Now if

$$\beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} \leq \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{-\beta}\right),$$

then

$$\int_0^\infty \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} d\beta \leq \int_0^\infty \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{-\beta}\right) d\beta.$$

Therefore

$$\int_0^\infty \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta}\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)^{-(n_i+1)} d\beta < \infty$$

which results in

$$\int_0^\infty\int_0^\infty\int_0^\infty \left(\beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\right)\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right\} d\theta_1 d\theta_2 d\beta < \infty.$$

This completes the proof, and we can conclude that $\pi_U(\theta_1, \theta_2, \beta|data)$ is proper. □

Corollary 4.2. *The posterior distribution using the uniform prior under k stress levels is given by*

$$\pi_U(\theta_1, \theta_2, \beta|data) \propto (\beta\theta_1)^{\sum_{i=1}^{k} n_i}\left(\prod_{i=1}^{k} S_i^{n_i\theta_2}\right)\left(\prod_{i=1}^{k}\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left[-\theta_1\sum_{i=1}^{k}\left(S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right]. \tag{4.14}$$

4.2.3 Mixture of Uniform and Gamma Priors and Posteriors

We will consider a uniform prior on β, that is $\beta \sim Uni(0,5)$ and vague gamma priors on θ_1 and θ_2, where $\theta_1 \sim Gamma(0.001, 0.001)$ and $\theta_2 \sim Gamma(0.001, 0.001)$. The parameters θ_1, θ_2 and β are assumed to be independent, therefore the joint prior using the mixture of uniform and Gamma priors is given by

$$\pi_{MIX}(\theta_1, \theta_2, \beta) \propto \theta_1^{0.001-1}\exp\{-\theta_1(0.001)\}\,\theta_2^{0.001-1}\exp\{-\theta_2(0.001)\} \times \frac{1}{5} \tag{4.15}$$

and the posterior distribution is given by

$$\begin{aligned}\pi_{MIX}\left(\theta_1,\theta_2,\beta|data\right) \propto\ & \beta^{\sum\limits_{i=1}^{k} n_i}\theta_1^{\sum\limits_{i=1}^{k} n_i+0.001-1}\theta_2^{0.001-1}\left(\prod_{i=1}^{k} S_i^{n_i\theta_2}\right)\left(\prod_{i=1}^{k}\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\\ \times\ & \exp\left\{-\theta_2\left(0.001\right)\right\}\exp\left\{-\theta_1\left[\sum_{i=1}^{k}\left(S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)+0.001\right]\right\} \qquad (4.16)\end{aligned}$$

4.2.4 Maximal Data Information (MDI) Prior and Posterior Distribution

The MDI prior will be derived using the Weibull model under the power law in the following theorem.

Theorem 4.3. *The MDI prior, $\pi_{MDI}(\theta_1,\theta_2,\beta)$ for the Weibull distribution under the power law at the i^{th} stress level is given by :*

$$\pi_{MDI}(\theta_1,\theta_2,\beta) \propto \beta^{n_i}\theta_1^{\frac{n_i}{\beta}}S_i^{\frac{n_i\theta_2}{\beta}}\exp\left\{\frac{n_i\gamma}{\beta}\right\}. \qquad (4.17)$$

Proof. Firstly the log of the likelihood is given by

$$\ell = \log L\left(\theta_1,\theta_2,\beta|data\right) = n_i\log\beta + n_i\log\theta_1 + n_i\theta_2\log S_i + (\beta-1)\sum_{j=1}^{n_i}\log x_{ij} - \theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}$$

and the expected value of the log of the likelihood is

$$E\left(\ell\right) = \log\beta^{n_i} + \log\theta_1^{n_i} + \log S_i^{n_i\theta_2} + (\beta-1)\sum_{j=1}^{n_i} E\left(\log X_{ij}\right) - \theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} E\left(X_{ij}^{\beta}\right). \qquad (4.18)$$

If $X_{ij}\sim Wei\left(\lambda_i,\beta\right)$, then $X_{ij}^{\beta}\sim Exp\left(\lambda_i\right)$ so,

$$\begin{aligned}E\left(X_{ij}^{\beta}\right) &= \frac{1}{\lambda_i}\\ &= \frac{1}{\theta_1 S_i^{\theta_2}}.\end{aligned}$$

Now from Equation 4.18 we see that we need to find $E\left(\log X_{ij}\right)$ in order to evaluate $E\left(\ell\right)$. Therefore,

$$
\begin{aligned}
E\left(\log X_{ij}\right) &= \int_0^\infty \left(\log x_{ij}\right)\beta\theta_1 S_i^{\theta_2} x_{ij}^{\beta-1}\exp\left\{-\theta_1 S_i^{\theta_2} x_{ij}^{\beta}\right\}dx_{ij} \\
&= \beta\theta_1 S_i^{\theta_2}\int_0^\infty \left(\log x_{ij}\right) x_{ij}^{\beta-1}\exp\left\{-\theta_1 S_i^{\theta_2} x_{ij}^{\beta}\right\}dx_{ij} \\
&= \beta\theta_1 S_i^{\theta_2}\int_0^\infty \left(\log\left(x_{ij}^{\beta}\right)^{\frac{1}{\beta}}\right) x_{ij}^{\beta-1}\exp\left\{-\theta_1 S_i^{\theta_2} x_{ij}^{\beta}\right\}dx_{ij}.
\end{aligned}
$$

Let $x_{ij}^{\beta} = y_i$ then $dy_i = \beta x_{ij}^{\beta-1}dx_{ij}$, therefore

$$
\begin{aligned}
E\left(\log X_{ij}\right) &= \theta_1 S_i^{\theta_2}\int_0^\infty \frac{\log\left(y_i\right)}{\beta}\exp\left\{-\theta_1 S_i^{\theta_2} y_i\right\}dy_i \\
&= \frac{\theta_1 S_i^{\theta_2}}{\beta}\int_0^\infty \log\left(y_i\right)\exp\left\{-\theta_1 S_i^{\theta_2} y_i\right\}dy_i,
\end{aligned}
$$

let $u_i = \theta_1 S_i^{\theta_2} y_i$ then $du_i = \theta_1 S_i^{\theta_2} dy_i$, so

$$
\begin{aligned}
E\left(\log X_{ij}\right) &= \frac{1}{\beta}\int_0^\infty \log\left(\frac{u_i}{\theta_1 S_i^{\theta_2}}\right)\exp\left\{-u_i\right\}du_i \\
&= \frac{1}{\beta}\left[\int_0^\infty \log\left(u_i\right)\exp\left\{-u_i\right\}du_i - \log\left(\theta_1 S_i^{\theta_2}\right)\int_0^\infty \exp\left\{-u_i\right\}du_i\right] \\
&= \frac{1}{\beta}\left[-\gamma - \log\left(\theta_1 S_i^{\theta_2}\right)\right]
\end{aligned}
$$

where γ is Eulers constant. Substituting back into Equation 4.18 we get

$$
E\left(\ell\right) = \log\left(\beta\theta_1 S_i^{\theta_2}\right)^{n_i} + \frac{n_i\left(\beta-1\right)}{\beta}\left[-\gamma - \log\left(\theta_1 S_i^{\theta_2}\right)\right] - n_i.
$$

Using Equation 2.29 we get

$$
\begin{aligned}
\pi_{MDI}(\theta_1,\theta_2,\beta) &\propto \left(\beta\theta_1 S_i^{\theta_2}\right)^{n_i}\exp\left(\frac{n_i\left(\beta-1\right)}{\beta}\left[-\gamma-\log\left(\theta_1 S_i^{\theta_2}\right)\right]\right)\exp\left\{-n_i\right\} \\
&\propto \left(\beta\theta_1 S_i^{\theta_2}\right)^{n_i}\exp\left(\frac{n_i\left(\beta-1\right)}{\beta}\left[-\gamma-\log\left(\theta_1 S_i^{\theta_2}\right)\right]\right) \\
&= \beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\exp\left(\left[1-\frac{1}{\beta}\right]\left[-\gamma n_i - n_i\log\left(\theta_1 S_i^{\theta_2}\right)\right]\right) \\
&= \beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\exp\left(-\gamma n_i + \log\left(\theta_1\right)^{-n_i} + \log\left(S_i\right)^{-n_i\theta_2} + \frac{\gamma n_i}{\beta} + \log\left(\theta_1\right)^{\frac{n_i}{\beta}} + \log\left(S_i\right)^{\frac{n_i\theta_2}{\beta}}\right)
\end{aligned}
$$

$$\begin{aligned}\pi_{MDI}(\theta_1,\theta_2,\beta) &\propto \beta^{n_i}\theta_1^{n_i}S_i^{\theta_2 n_i}\theta_1^{-n_i}S_i^{-\theta_2 n_i}\exp\left(\frac{\gamma n_i}{\beta}\right)\theta_1^{\frac{n_i}{\beta}}S_i^{\frac{n_i\theta_2}{\beta}}\\ &= \beta^{n_i}\theta_1^{\frac{n_i}{\beta}}S_i^{\frac{n_i\theta_2}{\beta}}\exp\left(\frac{\gamma n_i}{\beta}\right).\end{aligned}$$

□

The joint posterior distribution of θ_1,θ_2 and β using the MDI prior for the Weibull model under the power law is given by

$$\pi_{MDI}(\theta_1,\theta_2,\beta|data) \propto \prod_{j=1}^{n_i} x_{ij}^{\beta-1}\beta^{2n_i}\theta_1^{\frac{n_i}{\beta}+n_i}S_i^{\frac{\theta_2 n_i}{\beta}+\theta_2 n_i}\exp\left\{\frac{n_i\gamma}{\beta}\right\}\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i}x_{ij}^{\beta}\right\}. \tag{4.19}$$

Corollary 4.4. *The joint prior using the MDI prior under k stress levels is given by*

$$\pi_{MDI}(\theta_1,\theta_2,\beta) \propto \beta^{\sum_{i=1}^{k}n_i}\theta_1^{\sum_{i=1}^{k}n_i}\prod_{i=1}^{k}S_i^{n_i\theta_2}\exp\left\{-\left(\frac{\beta-1}{\beta}\right)\sum_{i=1}^{k}n_i\left(\gamma+\log\theta_1 S_i^{\theta_2}\right)\right\} \tag{4.20}$$

and the joint posterior is given by

$$\begin{aligned}\pi_{MDI}(\theta_1,\theta_2,\beta|data) &\propto (\beta\theta_1)^{2\sum_{i=1}^{k}n_i}\left(\prod_{i=1}^{k}S_i^{n_i\theta_2}\right)^2\left(\prod_{i=1}^{k}\prod_{j=1}^{n_i}x_{ij}^{\beta-1}\right)\\ &\times \exp\left\{-\left[\left(\frac{\beta-1}{\beta}\right)\sum_{i=1}^{k}n_i\left(\gamma+\log\theta_1 S_i^{\theta_2}\right)+\theta_1\sum_{i=1}^{k}S_i^{\theta_2}\sum_{j=1}^{n_i}x_{ij}^{\beta}\right]\right\}. \end{aligned}\tag{4.21}$$

The properness of the posterior distribution using the MDI prior will be investigated in future research.

4.3 Adaptive Rejection Sampling

In order to use the ARS method, we need conditional distributions that are log-concave. In Appendix A.2 we show that the conditional posteriors using the uniform and MDI priors are log-concave. The general steps involved with the ARS algorithm, using the notation from Section 2.10 are:

1. Initialise the abscissae in T_k. If D is unbounded on the left then choose x_1 such that $h'(x) > 0$. If D is unbounded on the right then choose x_k such that $h'(x) < 0$. After k starting points have been defined, calculate $u_k(x)$,$s_k(x)$ and $l_k(x)$.

2. Sample a value x^* from $s_k(x)$ and sample a value w independently from the uniform $(0,1)$ distribution. Then, perform the following squeezing test if

$$w \leq \exp\{l_k(x^\star) - u_k(x^\star)\}$$

then accept $x^\star$. Otherwise evaluate $h(x^\star)$ and $h'(x^\star)$ and perform the following rejection test: if

$$w \leq \exp\{h(x^\star) - u_k(x^\star)\}$$

then accept $x^\star$, otherwise reject $x^\star$.

3. If $h(x^\star)$ and $h'(x^\star)$ were evaluated at step 2, include $x^\star$ in T_k to form T_{k+1} and relabel the elements of T_{k+1} in ascending order. Construct $u_{k+1}(x)$, $s_{k+1}(x)$ and $l_{k+1}(x)$ on the basis of T_{k+1} and increment k. Return to step 2 if n parts have not yet been accepted.

4.4 Reliability

Soyer et al. (2008) obtained the predictive reliability at the use-stress environment using the equation given by

$$R(x_u|D) = \int R(x_u|\theta_1, \theta_2, \beta) \times \pi(\theta_1, \theta_2, \beta|D)\, d\theta_1 d\theta_2 d\beta \tag{4.22}$$

where

$$R(x_u|\theta_1, \theta_2, \beta) = \exp\left\{-\lambda_u x_u^\beta\right\} \tag{4.23}$$

and

$$\lambda_u = \theta_1 S_u^{\theta_2}. \tag{4.24}$$

The steps involved for evaluating $R(x_u|D)$ is given by :

1. Sample θ_1, θ_2 and β from the posterior densities.

2. The integral in Equation 4.22 can then be computed using the Monte Carlo average of the posterior sample $\left\{\theta_1^{(j)}, \theta_2^{(j)}, \beta^{(j)}\right\}_{j=1}^{J}$ as

$$R(x_u|D) \approx \frac{1}{J}\sum_{j=1}^{J} R\left(x_u|\theta_1^{(j)}, \theta_2^{(j)}, \beta^{(j)}\right) \tag{4.25}$$

which is the expected reliability at mission time x_u.

4.5 Application

In this section we will compare the prior used in Soyer et al. (2008), the uniform prior and the mixture of uniform and Gamma priors presented in Section 4.2. The performance of the MDI prior will be investigated in future research. Comparing these models using Bayes factors is difficult since the marginal likelihoods for the competing models, which are needed to compute the Bayes factors can not be directly approximated from the Gibbs sampler as mentioned in Soyer et al. (2008). For this reason we will make use of the deviance information criterion (DIC). The DIC was introduced by Spiegelhalter et al. (2002) as a model assessment and comparison tool. It is defined in Spiegelhalter et al. (2002) as a Bayesian measure of fit, penalised by an additional complexity term. The effective number of parameters in the model, p_D is defined as

$$p_D = \overline{D(\theta)} - D(\overline{\theta}). \tag{4.26}$$

The Bayesian deviance is given by

$$D(\theta) = -2\log\left[L(\theta|data)\right] + 2\log\left[f(data)\right], \tag{4.27}$$

where $f(data)$ is some function of the data and $L(\theta|data)$ is the maximised likelihood value over the unknown parameters. Therefore, from Equation 4.27, the posterior mean deviance is given by

$$\overline{D(\theta)} = E_\theta\left(-2\log\left[L(\theta|data)\right] + 2\log\left[f(data)\right]\right) \tag{4.28}$$

and the deviance of the means, $D(\overline{\theta})$ is given by

$$D(\overline{\theta}) = -2\log\left[L\left(E\left(\theta|data\right)\right)\right] + 2\log\left[f(data)\right]. \tag{4.29}$$

The DIC as given in Spiegelhalter et al. (2002) is defined as

$$DIC = \overline{D(\theta)} + p_D \tag{4.30}$$

$$= D(\overline{\theta}) + 2p_D. \tag{4.31}$$

The model that has the smallest DIC is the best model. We will make use of an example used in Soyer et al. (2008) and use the ALT data published in Nelson (1972). The data is given in Table 4.1 and represents the breakdown of an insulating fluid subjected to various voltage levels. The accelerated stress levels are given by 26, 28, 30, 32, 34, 36, and 38 Kv, where Kv represents kilovolts. We are interested in making inference at the use-stress of 22 Kv. A power law model is assumed for the data, as in Nelson (1972). We consider three models to analyse this data, the prior used in Soyer et al. (2008), the uniform prior and the mixture of uniform and Gamma priors. The ARS method is implemented at each iteration of the Gibbs sampler. Trace plots are given and the posterior distributions of θ_1, θ_2 and β are also given, see Appendix D for the WinBUGS and R® code to this study.

Table 4.1: Times to breakdown of an insulating fluid (in minutes) under various values of the stress, Soyer et al. (2008).

38 Kv	36 Kv	34 Kv	32Kv	30Kv	28Kv	26Kv
0.09	0.35	0.19	0.27	7.74	68.85	5.79
0.39	0.59	0.78	0.40	17.05	108.29	1579.52
0.47	0.96	0.96	0.69	20.46	110.59	2323.70
0.73	0.99	1.31	0.79	21.02	426.07	
0.74	1.69	2.78	2.75	22.66	1067.6	
1.13	1.97	3.16	3.91	43.40		
1.40	2.07	4.15	9.88	47.30		
2.38	2.59	4.67	13.95	139.07		
	2.71	4.85	15.93	141.12		
	2.90	6.50	27.80	175.88		
	3.67	7.35	53.24	194.90		
	3.99	8.01	82.85			
	5.35	8.27	89.29			
	13.77	12.06	100.58			
	25.50	31.75	215.10			
		32.52				
		33.91				
		36.71				
		72.89				

The trace plots for each model considered is given in Figures 4.1 to 4.3 to see convergence of the unknown parameters of each model under consideration. The plots for the jump, history and running quantiles can be found in Appendix B.3 . Trace plots are used to visually check convergence. If no patterns or irregularities are observed, convergence can be assumed, as mentioned in Ntzoufras (2009).

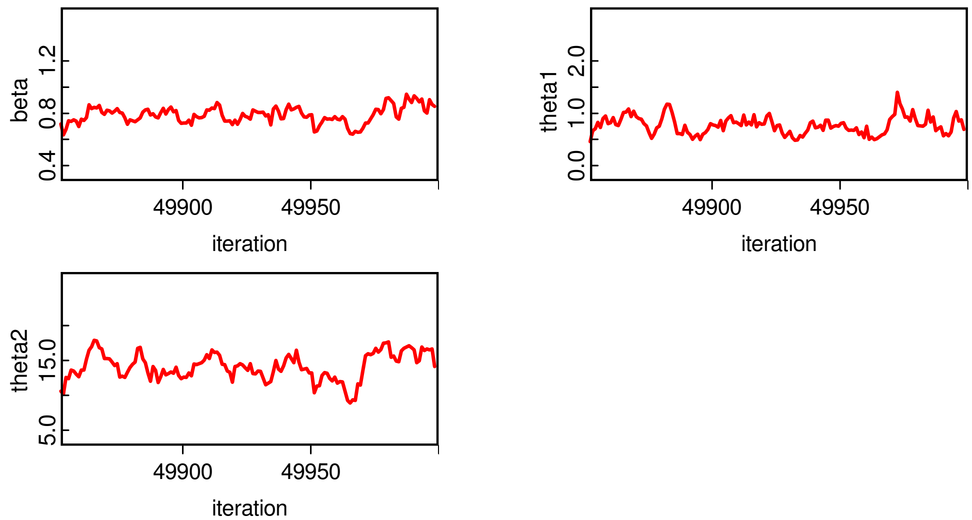


Figure 4.1: Trace plots of posteriors using priors from Soyer et al. (2008)

From Figure 4.1 the MCMC has been run for 50000 iterations and the chains are mixing well for β, θ_1 and θ_2, therefore all three parameters have converged for the priors used in Soyer et al. (2008).

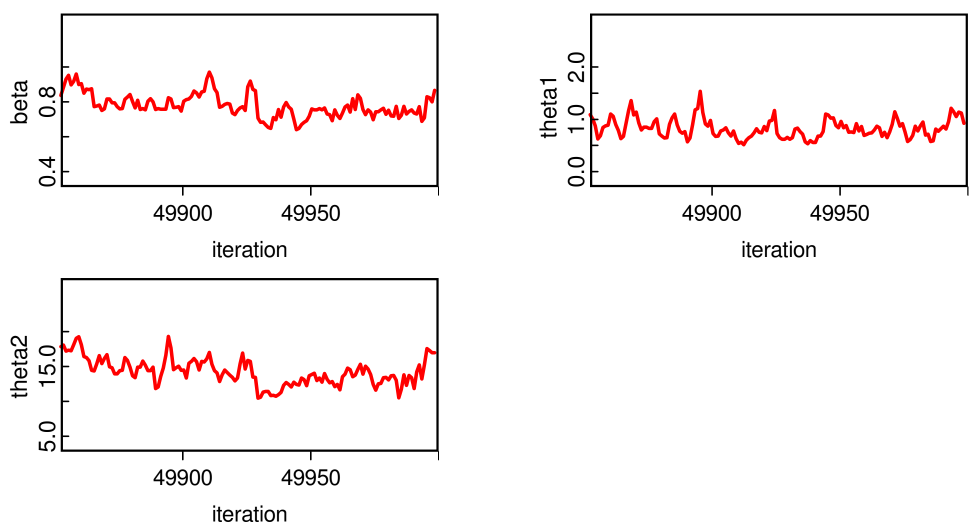


Figure 4.2: Trace plots of posteriors using uniform priors.

From Figure 4.2 we see that for θ_1 θ_2 and β the chain is moving well around the parameter space and establishes convergence after 50000 iterations.

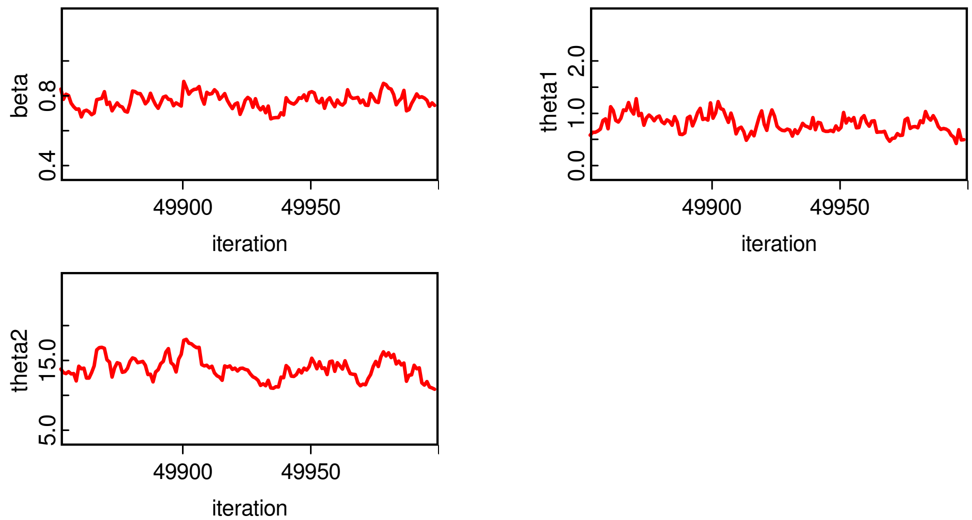


Figure 4.3: Trace plots of posteriors using mixture of uniform and Gamma priors.

It is clear from Figure 4.3 that θ_1, θ_2 and β have converged as the chain is mixing well after 50 000 iterations.

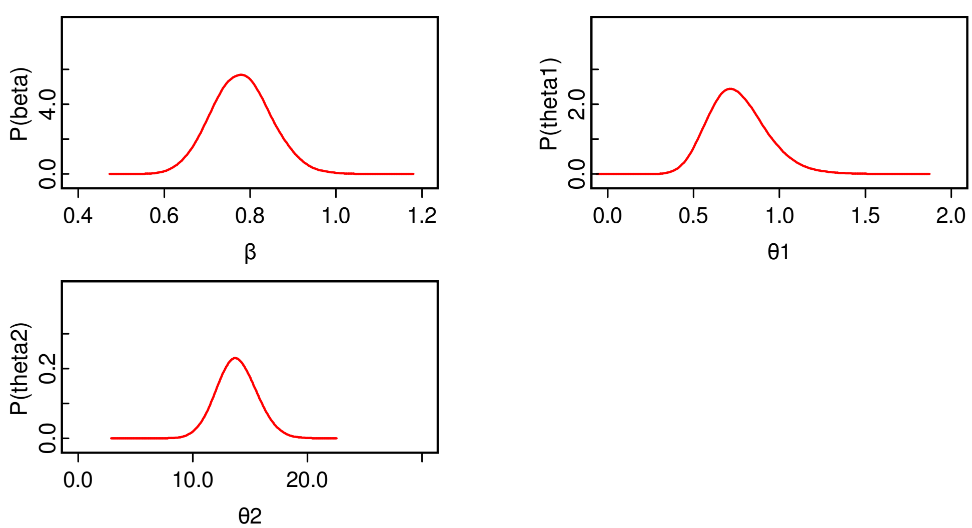


Figure 4.4: Posterior distributions of β, θ_1 and θ_2, using the priors from Soyer et al. (2008).

The posterior distribution of θ_1,θ_2 and β are given in Figure 4.4. The posterior distribution of β is concentrated around 0.8, which implies a decrease in failure rate at any stress level. The distribution of θ_2 is focussed around positive values, which implies that as the stress level increases the failure rate also increases, which is an expected result in ALT.

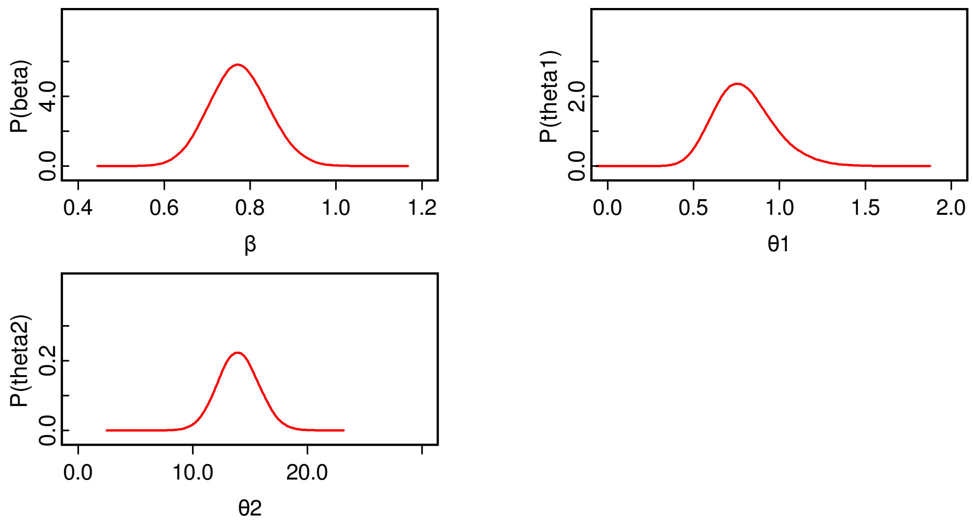


Figure 4.5: Posterior distributions of β, θ_1 and θ_2, using the uniform prior.

The posterior distributions of θ_1, θ_2 and β when the uniform prior is used are given in Figure 4.5. The posterior distribution of β is concentrated around 0.7 and this shows that there is a decrease in failure rate at any level of stress in the ALT. From Figure 4.5, we see the posterior distribution of θ_2 is around high positive integers and this implies that the higher the stress level then the failure rate will increase.

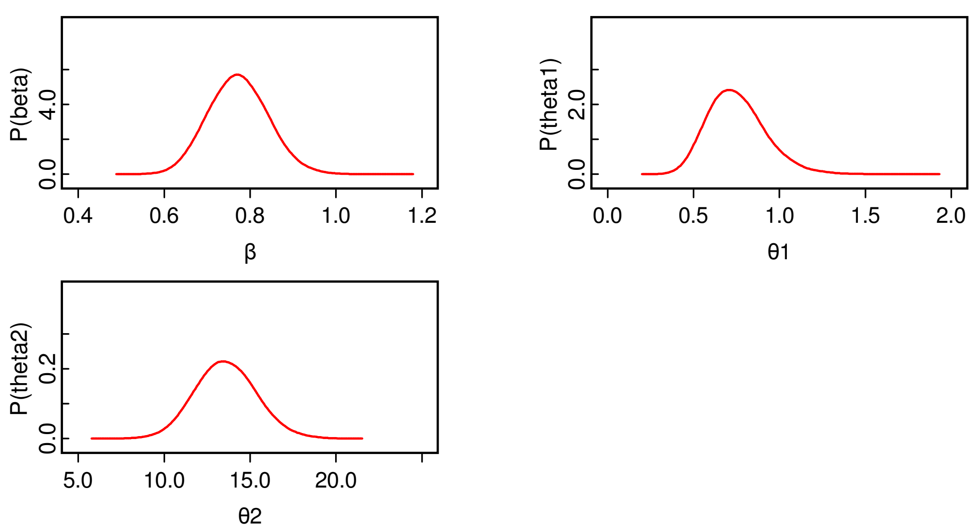


Figure 4.6: Posterior distributions of β, θ_1 and θ_2, using a mixture of uniform and Gamma priors.

Figure 4.6 shows the posterior distributions of θ_1, θ_2 and β when a mixture of a uniform prior and Gamma priors are used. From Figure 4.6 the values of β is concentrated around 0.75 and this shows that the failure rate increases at any stress level. For values of θ_1 greater than one, there is an increase in the failure rate as each stress level increases. Also, θ_2 sits around high positive values which implies that the higher the stress level the more the failure rate increases.

The results of the posterior values of β are summarised in Figure 4.7, where box plots are constructed. The central mark on the box is the median, the edges are the 25^{th} and the 75^{th} percentiles.

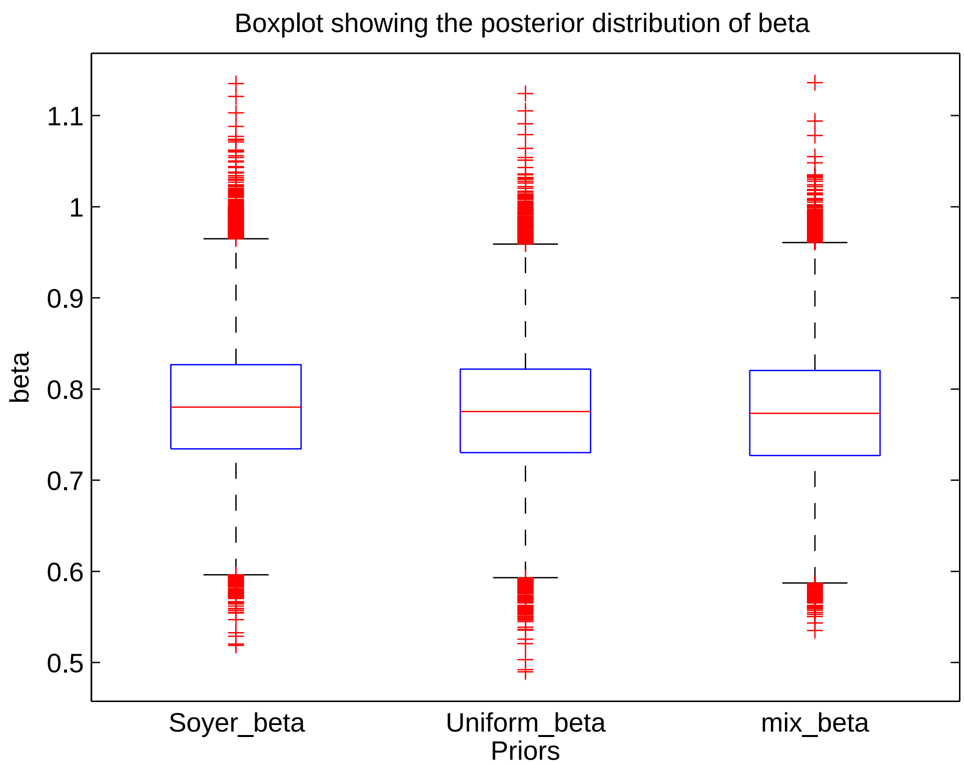


Figure 4.7: Boxplot showing the posterior of β for each model considered.

From Figure 4.7 we see the median values of β for all three models are around 0.8. From the above boxplot it seems the distribution for β where the prior Soyer et al. (2008) used, the uniform prior and the mixture of uniform and Gamma priors are very similar.

The results of the posterior values of θ_1 are summarised in Figure 4.8 where box plots are constructed.

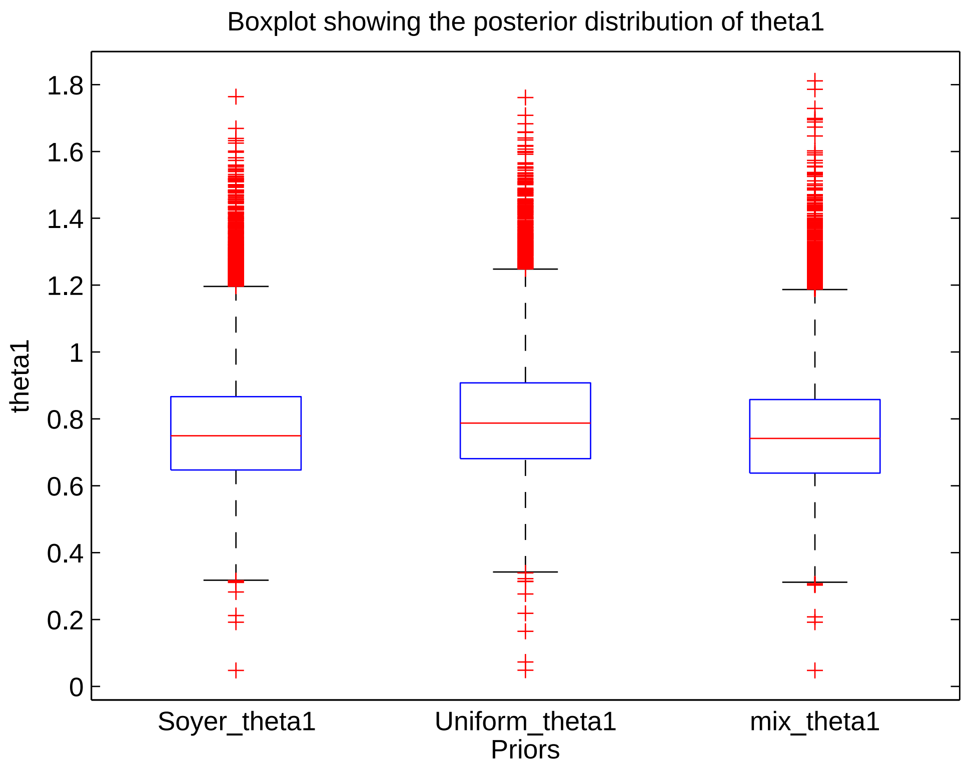


Figure 4.8: Boxplot showing the posterior of θ_1 for each model considered.

From Figure 4.8, we see the median values for θ_1 for the prior used by Soyer et al. (2008) and the mixture of uniform and Gamma priors are both just below 0.75. The median values for θ_1 when the uniform prior is used is below 0.8. The posterior distribution for θ_1 does not change much from one stress level to the next therfore the data suggests that the time transformation function does not vary with the stress level.

The results of the posterior values of θ_2 are summarised in Figure 4.9 where boxplots are constructed.

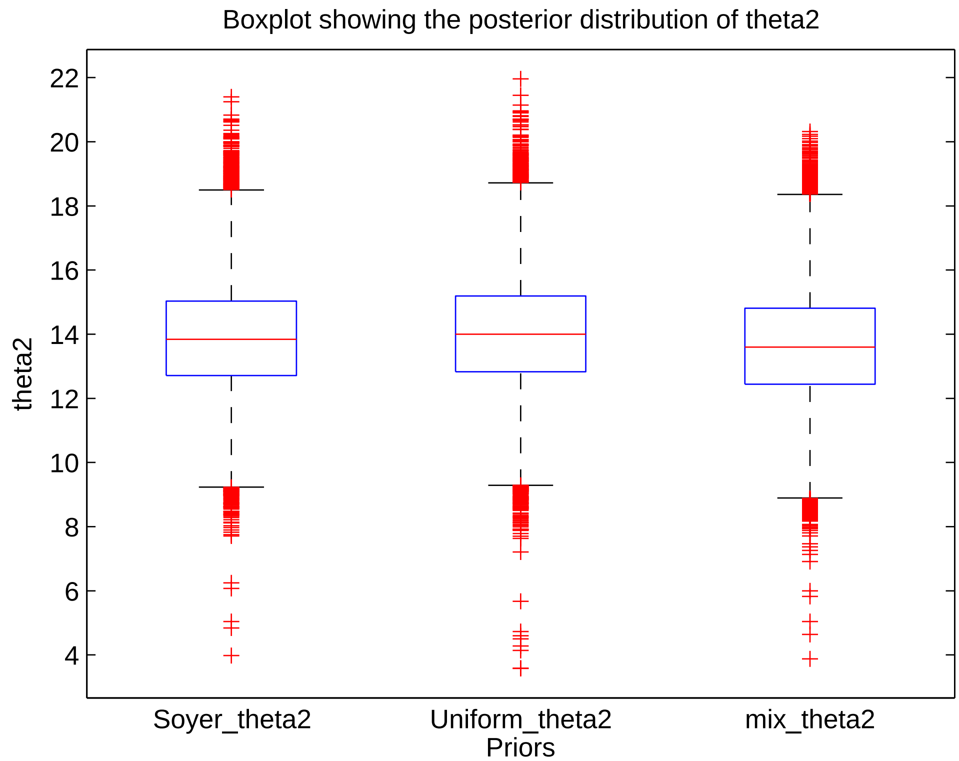


Figure 4.9: Boxplot showing the posterior of θ_2 for each model considered.

The median value of θ_2 for the prior used in Soyer et al. (2008) and the mixture of priors is less than 14. The median value of θ_2 for the uniform prior is exactly 14. Since θ_2 lies above zero, it suggests that the failure rate increases with the stress levels. The distribution for θ_2 is similar for all three models considered.

Table 4.2: Summary statistics for the Soyer et al. (2008) prior.

	mean	**standard deviation**	**Monte Carlo (MC) error**	**2.5% percentiles**	**median**	**97.5% percentiles**
β	0.7816	0.06861	9.04E-0.4	0.6518	0.7803	0.9202
θ_1	0.7651	0.168	0.001581	0.4823	0.7497	1.137
θ_2	13.89	1.726	0.02349	10.59	13.84	17.38

Table 4.3: Summary statistics for the uniform prior.

	mean	standard deviation	Monte Carlo (MC) error	2.5% percentiles	median	97.5% percentiles
β	0.7765	0.06787	9.42E-0.4	0.6466	0.7753	0.9124
θ_1	0.8035	0.1737	0.001683	0.5091	0.7878	1.191
θ_2	14.02	1.75	0.02438	10.67	14	17.48

Table 4.4: Summary statistics for mixture of uniform and Gamma priors.

	mean	standard deviation	Monte Carlo (MC) error	2.5% percentiles	median	97.5% percentiles
β	0.7747	0.06852	9.65E-0.4	0.6455	0.7735	0.9122
θ_1	0.7571	0.167	0.001779	0.4781	0.7418	1.124
θ_2	13.64	1.756	0.02571	10.31	13.6	17.2

The Monte Carlo error measures the variability of each estimate due to the simulation. The MC error must be low in order to calculate the parameter of interest with increased precision, Ntzoufras (2009). Small values of this error will indicate that the quantity of interest has been calculated with precision. From Tables 4.2 to 4.4 we can see that the MC errors are small, where the priors for Soyer et al. (2008) yielded the smallest MC error.

A comparison of the three models using the DIC is given in Table 4.5

Table 4.5: Model comparison using DIC.

Model	DIC	p_D
Prior from Soyer et al. (2008)	607.6	2.983
Uniform Prior	607.5	2.931
Mixture of priors	607.6	2.99

Table 4.5 shows that the DIC for the model from Soyer et al. (2008) and the model that used a mixture of a uniform and Gamma priors are the same. The model using the uniform prior gave the smallest DIC, based on the DIC is the preferred model. The effective number of parameters for all three models under consideration is 3. The three models gave very similar results and there is not much difference in their performance.

The predictive reliability will now be determined for this example, using Equations 4.23 and 4.25. The predictive reliability at the use stress S_u = 22 Kv when using the priors from Soyer *et al.* (2008) is given in Figure 4.10. The predictive reliability at the use stress S_u = 22 Kv when using the uniform prior is given in Figure 4.11. The predictive reliability at the use stress S_u = 22 Kv when using a mixture of uniform and Gamma priors is given in Figure 4.12. See Appendix B.3 for the R® code to this study. The results are very similar for the three priors, and as noted by Soyer *et al.* (2008) the items are quite reliable at the use stress level. Where reliability values greater than 0.5 are observed for mission time values less than 10 000.

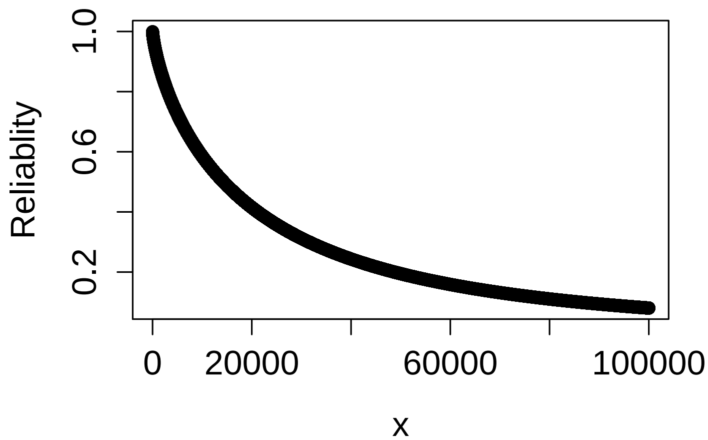


Figure 4.10: Predictive reliability function at use stress S_u = 22 Kv, when using the priors from Soyer *et al.* (2008).

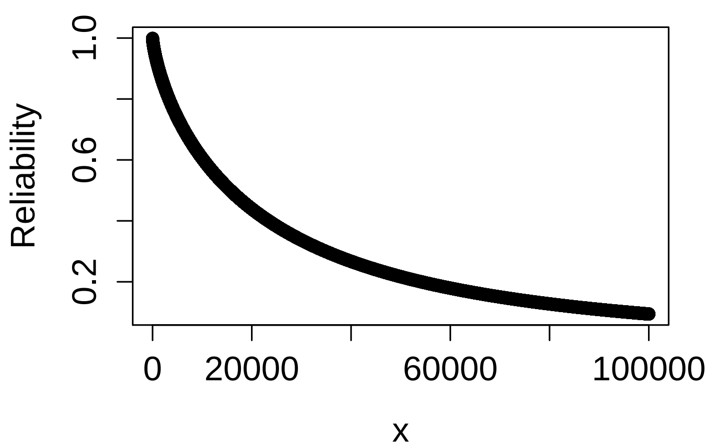


Figure 4.11: Predictive reliability function at use stress S_u = 22 Kv, when using the uniform prior.

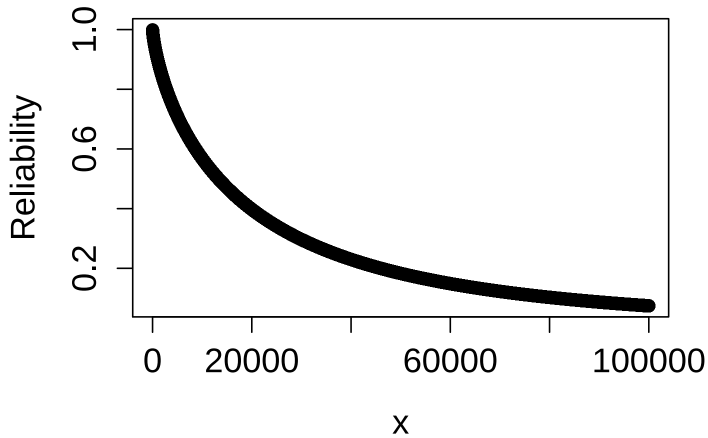


Figure 4.12: Predictive reliability function at use stress $S_u = 22$ Kv, when using a mixture of uniform and Gamma priors.

Table 4.6 shows further results on the predictive reliability for various values of mission times. Having a closer look at the actual values, it is clear that the predictive reliability is higher when the uniform prior is used.

Table 4.6: Predictive reliability function at use stress S_u = 22 Kv.

Mission time, x	**Soyer prior**	**Uniform prior**	**Mixture prior**
1	0.9994484	0.9994527	0.9993694
2	0.9990814	0.9990929	0.9989554
3	0.998761	0.9987798	0.9985956
4	0.9984673	0.9984935	0.9982667
5	0.9981921	0.9982256	0.9979591
6	0.9979306	0.9979714	0.9976673
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
51	0.9898242	0.9901594	0.9887279
52	0.9896757	0.990017	0.9885652
53	0.9895279	0.9898752	0.9884034
54	0.9893808	0.9897342	0.9882423
55	0.9892343	0.9895938	0.988082
56	0.9890886	0.9894541	0.9879225
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
101	0.9830564	0.983682	0.9813404
102	0.9829314	0.9835626	0.9812044
103	0.9828068	0.9834435	0.9810688
104	0.9826824	0.9833248	0.9809335
105	0.9825584	0.9832063	0.9807985
106	0.9824347	0.9830881	0.9806639
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
151	0.9771343	0.9780306	0.9749081

4.6 Conclusion

In this chapter the power law was considered using the Weibull distribution as a lifetime model. The MDI prior and posterior distribution were derived. The posterior when using the uniform prior was also derived and it was shown that the posterior distribution using the uniform prior under a single stress level is proper. It was also shown that the posterior distributions using the MDI prior and the uniform prior are both log-concave. An example was considered which looked at the breakdown of an insulating fluid subjected to various voltage stress levels. The models used to analyse the data was the prior considered in Soyer et al. (2008), the uniform prior and a mixture of the uniform and Gamma priors. The adaptive rejection sampling method was used to sample from the posterior distributions of each model. These models were compared using the DIC for model assessment. The uniform prior had the smallest DIC and therefore is the preferred model. The DIC for the model used in Soyer et al. (2008) and the mixture of uniform and Gamma priors model had the same DIC and thus had the same performance. The predictive reliability was computed at the use-stress level for the priors from Soyer et al. (2008), the uniform prior and the mixture of the uniform and Gamma priors. The results was very similar for the three priors and it was found that the items were quite reliable at the use-stress level, this was also noted in Soyer et al. (2008). The predictive reliability was found to be higher when the uniform prior was used, a further indication that the model where the uniform prior is used is the preferred one.

Chapter 5

Concluding Remarks

This chapter concludes with a summary of the conclusions of the chapters in this thesis, possibilities for future research and highlights some limitations of the study.

5.1 Summary of Conclusion

This thesis focused on objective Bayesian statistics in accelerated life testing by considering a number of non-informative priors. These priors were the Jeffreys prior, the MDI prior, the GML prior and the uniform prior. These priors are used when there is little or no prior information available about the experiment. In Chapter 3 we derived the Jeffreys prior, GML prior and the MDI prior using the power law as the time transformation function and the exponential distribution as the lifetime model. The posterior for each of the priors considered was derived and it was proven that all these posteriors were proper. The marginal and conditional posteriors were derived using each of the priors considered. We also derived the posterior distributions using the Jeffreys prior, MDI prior and GML prior under Type-I censoring. We then used a simulation study to find the optimal design using the different priors by looking at the pre-posterior variance for each design. The posterior variance was considered since we assumed the Bayes estimator under squared error loss and the minimum under squared error loss is the variance of the posterior. We looked at the fixed points design from Erkanli & Soyer (2000). The prior with the smallest pre-posterior variance is the preferred one. The GML prior had the smallest pre-posterior variance for each design that we considered and thus performed the best. The MDI prior performed better than the Jeffreys prior for all three designs. Design 1 had the smallest pre-posterior variance regardless of which prior was used and therefore the optimal design was at the use-stress level, $S_u = 1.05$. As for the two-point designs, design 2 performed better than design 3 when the Jeffreys prior and GML prior is used. Design 3 only performed better than design 2 when the MDI prior was used. In Chapter 4, the MDI prior and posterior distribution were derived. The posterior when using the uniform prior was also derived and it was shown that the posterior distribution using the uniform prior under a single stress level is proper. It was also shown that the posterior distributions using the MDI prior and the uniform prior are both log-concave. An example was considered which looked at the breakdown of an insulating fluid subjected to various voltage stress levels. The models used to analyse the data was the prior considered in Soyer et al. (2008), the uniform prior and a mixture of the uniform and Gamma priors. The adaptive rejection sampling method was used, since the posteriors are all log-concave, to sample from the posterior distributions of each model. These models were compared using the DIC for model assessment. The uniform prior had the smallest DIC and therefore is the preferred model. The DIC for the model used in Soyer et al. (2008) and the mixture of uniform and Gamma priors model had the same DIC and thus had the same performance. The predictive reliability of the three models were also computed at the use-stress level. It was found that the items were quite reliable at the use-stress level. The predictive reliability was found to be the highest when the uniform prior is used and therefore it is more reliable and the preferred model to use.

5.2 Future Research

Possible future research would be to consider a different time transformation function. In this thesis we only considered the power law and we would like to use objective Bayes to do Bayesian inference using the Arrhenius law as the time transformation function. In Chapter 4 of this thesis, we considered only the uniform prior and the MDI prior for the Weibull model. The properness and the performance of the MDI prior will be investigated in future research. It would be interesting to explore other non-informative priors. In Soyer et al. (2008) they mention that the Arrhenius law can be considered but they do not go into actually deriving any priors or posteriors using this model. To our knowledge we have not seen any papers that have considered the Arrhenius law as a time transformation function using objective Bayes in ALT. Also, we would like to incorporate censoring into the likelihood function since we may not always have complete data in real life applications. We may want to consider using the log-normal distribution as a lifetime model in future research. It would also be interesting to use the cumulative exposure model that is described in Nelson (1980) as the model that relates the distribution under step stress to the distribution under constant stress.

References

Balakrishnan, N., Kundu, D., Ng, H. K. T., & Kannan, N. (2007). Point and interval estimation for a simple step-stress model with type-II censoring. *Journal of Quality Technology*, 39(1), 35–47.

Casella, G. & George, E. (1992). Explaining the Gibbs sampler. *The American Statistician*, 46(3), 167–174.

Ebeling, C. E. (1997). *An Introduction to Reliability and Maintainability Engineering*. New York: McGraw-Hill, 1st edition.

Erkanli, A. & Soyer, R. (2000). Simulation-based designs for accelerated life tests. *Journal of Statistical Planning and Inference*, 90(2), 335–348.

Escobar, L. A. & Meeker, W. Q. (2006). A review of accelerated test models. *Statistical Science*, 21(4), 552–577.

Gamerman, D. & Lopes, H. (2006). *Markov Chain Monte Carlo- Stochastic Simulation for Bayesian Inference*. New York: Chapman and Hall, 2nd edition.

Ghosh, M., Mergel, V., & Liu, R. (2011). A general divergence criterion for prior selection. *Annals of the Institute of Statistical Mathematics*, 63(1), 43–58.

Gilks, W. & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Applied Statistics*, 41(2), 337–348.

Gouno, E. & Balakrishnan, N. (2001). Step-stress accelerated life test. *Handbook of Statistics*, 20, 623–693.

Hastings, W. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97–109.

Jeffreys, H. (1961). *Theory of Probability and Inference*. London: Cambridge University Press, 3rd edition.

Kateri, M. & Balakrishnan, N. (2008). Inference for a simple step-stress model with type-II censoring, and Weibull distributed lifetimes. *IEEE Transactions on Reliability*, 57(4), 616–626.

Kim, D. H., Kang, S. G., & Lee, W. D. (2009). An objective Bayesian analysis for multiple step-stress accelerated life tests. *Journal of the Korean Data & Information Science Society*, 20(3), 601–614.

Klein, J. P. & Moeschberger, M. L. (2003). *Survival Analysis: Techniques for Censored and Truncated Data*. Statistics for Biology and Health. Springer New York, 2nd edition.

Lee, J. & Pan, R. (2008). Bayesian inference model for step-stress accelerated life testing with type-II censoring. *Reliability and Maintainability Symposium*, 91-96.

Lunn, D., Thomas, A., Best, N., & Spiegelhalter, D. (2000). A Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10(4), 325–337.

MATLAB (2014). *MATLAB Student, MATLAB and Simulink Student Suite- 2014a*. The Mathworks Inc.

Mazzuchi, T. A., Soyer, R., & Vopatek, A. (1997). Linear Bayesian inference for accelerated Weibull model. *Lifetime Data Analysis*, 3(1), 63–75.

Nelson, W. (1972). Graphical analysis of accelerated life test data with the inverse power law model. *IEEE Transactions on Reliability*, R-21(1), 2–11.

Nelson, W. (1980). Accelerated life testing- step-stress models and data analyses. *IEEE Transactions on Reliability*, R-29(2), 103–108.

Northrop, P. & Attalides, N. (2014). Posterior propriety in objective Bayesian extreme value analyses. *Departmental research report 323, University College London*.

Ntzoufras, I. (2009). *Bayesian Modelling using WinBUGS*. New Jersey: John Wiley & Sons.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ross, S. M. (2007). *Introduction to Probability Models*. New York: Academic Press, 9th edition.

Soyer, R., Erkanli, A., & Merrick, J. R. (2008). Accelerated life tests: Bayesian models. *Encyclopedia of Statistics in Quality and Reliability*, 1, 20–30.

Spiegelhalter, D., Best, N., Carlin, B., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583–639.

Tojeiro, C., Louzada-Neto, F., & Bolfarine, H. (2004). A Bayesian analysis for accelerated lifetime tests under an exponential power law model with threshold stress. *Journal of Applied Statistics*, 31(6), 685–691.

Van Dorp, J. R. & Mazzuchi, T. A. (2004). A general Bayes exponential inference model for accelerated life testing. *Journal of Statistical Planning and Inference*, 119(1), 55–74.

Van Dorp, J. R. & Mazzuchi, T. A. (2005). A general Bayes Weibull inference model for accelerated life testing. *Reliability Engineering and System Safety*, 90(2-3), 140–147.

Van Dorp, J. R., Mazzuchi, T. A., & Garcidueñas, J. E. (2006). A comparison of accelerated life testing designs within a single Bayesian inferential framework. *Reliability and Maintainability Symposium*, 208-214.

Xu, A. & Tang, Y. (2011). Objective Bayesian analysis of accelerated competing failure models under type-I censoring. *Computational Statistics and Data Analysis*, 55(10), 2830–2839.

Yang, R. & Berger, J. O. (1998). *A catalog of noninformative priors*. Institute of Statistics and Decision Sciences, Duke University.

Zellner, A. (1971). *An Introduction to Bayesian inference in Econometrics*. New York: Wiley, 1st edition.

Zellner, A. (1996). Models, prior information, and Bayesian analysis. *Journal of Econometrics*, 75(1), 51–68.

Zhang, Z. (1997). Bayesian analysis of step-stress accelerated life testing under exponential distributions. *Chinese Applied Mathematics*, 12(2), 175–181.

Appendix A: Additional Results for Chapter 3

A.1 Distribution of T_i

If $X_j \sim Exp(\lambda_i)$, then $T_i = \sum\limits_{j=1}^{n_i} X_j \sim Gamma\left(n_i, \lambda_i\right)$. The moment-generating function of X_j will then be

$$M_{X_j}(t) = \frac{\lambda_i}{\lambda_i - t}.$$

Let $T_i = \sum_{j=1}^{n_i} X_j = X_1 + X_2 + \ldots + X_{n_i}$, then the moment-generating function of T_i will then be

$$\begin{aligned}
M_{T_i}(t) &= M_{X_1}(t)M_{X_2}(t)\ldots M_{X_{n_i}}(t) \\
&= \left(\frac{\lambda_i}{\lambda_i - t}\right)\left(\frac{\lambda_i}{\lambda_i - t}\right)\cdots\left(\frac{\lambda_i}{\lambda_i - t}\right) \\
&= \lambda^{n_i}\left(\lambda_i - t\right)^{-n_i} \\
&= \left(\frac{\lambda_i - t}{\lambda_i}\right)^{-n_i} \\
&= \left(1 - \frac{t}{\lambda_i}\right)^{-n_i},
\end{aligned}$$

which is the moment generating function of a Gamma distribution. Therefore $T_i \sim Gamma(n_i, \lambda_i)$.

A.2 Fisher Information Matrix for the Exponential Distribution

We will derive the Fisher information matrix for the exponential distribution using the power law. The likelihood function given the observed data is given by

$$L(\alpha, \beta | data) = \prod_{i=1}^{m}\prod_{j=1}^{n_i} \alpha S_i^{\beta} \exp\left\{-\alpha S_i^{\beta} x_{ij}\right\}$$

$$
\begin{aligned}
&= \prod_{i=1}^{m}\left(\alpha S_i^{\beta n_i}\right)\exp\left\{-\alpha S_i \sum_{i=1}^{n_i} x_{ij}\right\} \\
&= \alpha^n \left(\prod_{i=1}^{m} S_i^{\beta n_i}\right)\exp\left\{-\alpha \sum_{i=1}^{m} S_i^{\beta} T_i\right\} \quad \text{where } T_i = \sum_{j=1}^{n_i} x_{j.}
\end{aligned}
$$

The log-likelihood is

$$
\begin{aligned}
\ell = \log L(\alpha,\beta|data) &= n\log\alpha + \sum_{i=1}^{m} \log S_i^{\beta n_i} - \alpha \sum_{i=1}^{m} S_i^{\beta} T_i \\
&= n\log\alpha + \sum_{i=1}^{m} \beta n_i log S_i - \alpha \sum_{i=1}^{m} S_i^{\beta} T_i.
\end{aligned}
$$

The first and second order partial derivatives of the log-likelihood with respect to the unknown parameters are:

$$
\begin{aligned}
\frac{\partial \ell}{\partial \alpha} &= \frac{n}{\alpha} - \sum_{i=1}^{m} S_i^{\beta} T_i \\
\frac{\partial \ell}{\partial \beta} &= \sum_{i=1}^{m} n_i \log S_i - \alpha \sum_{i=1}^{m} \left(S_i^{\beta}\right)(\log S_i)(T_i)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2 \ell}{\partial \alpha^2} &= -\frac{n}{\alpha^2} \\
\frac{\partial^2 \ell}{\partial \beta^2} &= -\alpha \sum_{i=1}^{m} \left(S_i^{\beta}\right)(\log S_i)^2 (T_i) \\
\frac{\partial^2 \ell}{\partial \alpha \partial \beta} &= -\sum_{i=1}^{m} \left(S_i^{\beta}\right)(\log S_i)(T_i) = \frac{\partial^2 \ell}{\partial \beta \partial \alpha}.
\end{aligned}
$$

The Fisher information matrix is given by

$$
I(\alpha,\beta) = \begin{bmatrix} \frac{n}{\alpha^2} & \sum_{i=1}^{m} \left(S_i^{\beta}\right)(\log S_i) E(T_i) \\ \sum_{i=1}^{m} \left(S_i^{\beta}\right)(\log S_i) E(T_i) & \alpha \sum_{i=1}^{m} \left(S_i^{\beta}\right)(\log S_i)^2 E(T_i) \end{bmatrix}
$$

Since $T_i \sim Gamma(n_i, \lambda_i)$, $E(T_i) = \frac{n_i}{\lambda_i}$, and since $\lambda_i = \alpha S_i^{\beta}$, therefore $E(T_i) = \frac{n_i}{\alpha S_i^{\beta}}$. The Fisher information matrix will therefore be

$$I(\alpha,\beta) = \begin{bmatrix} \frac{n}{\alpha^2} & \frac{1}{\alpha}\sum_{i=1}^{m} n_i \log S_i \\ \frac{1}{\alpha}\sum_{i=1}^{m} n_i \log S_i & \sum_{i=1}^{m} n_i \left(\log S_i\right)^2 \end{bmatrix}.$$

Appendix B: Additional Results for Chapter 4

We will show that the conditional distributions of the posterior distribution using the uniform prior and the MDI prior under the power law are log-concave. We will now define log-concavity using the definition given in Gamerman & Lopes (2006).

Definition B.1. A density $\pi(x)$ is log-concave if $\frac{d\log\pi}{dx}$ is a non-increasing function of x or $\frac{d^2\log\pi}{dx^2}$ is a non positive function of x.

B.1 Log-Concavity of the Conditional Posteriors Using the Uniform Prior

The conditional posterior distribution of θ_1 is given by

$$\pi_U(\theta_1|\theta_2,\beta,data) \propto \theta_1^{n_i}\exp\left\{-\theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i}x_{ij}^{\beta}\right\}.$$

The log of this expression is given by

$$\log\left[\pi_U(\theta_1|\theta_2,\beta,data)\right] \propto n_i\log\theta_1 - \theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i}x_{ij}^{\beta}.$$

Then differentiating with respect to θ_1 :

$$\frac{\partial\log\left[\pi_U(\theta_1|\theta_2,\beta,data)\right]}{\partial\theta_1} = \frac{n_i}{\theta_1} - S_i^{\theta_2}\sum_{j=1}^{n_i}x_{ij}^{\beta}.$$

Taking the second derivative with respect to θ_1:

$$\frac{\partial^2\log\left[\pi_U(\theta_1|\theta_2,\beta,data)\right]}{\partial\theta_1^2} = -\frac{n_i}{\theta_1^2}$$

which is negative and therefore $\pi_U(\theta_1|\theta_2,\beta,data)$ is log-concave.

The conditional posterior of θ_2 is given by

$$\pi_U(\theta_2|\theta_1,\beta,data) \propto S_i^{\theta_2 n_i} \exp\left\{-\theta_1 S_i^{\theta_2} \sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}.$$

The log of this expression is given by

$$\log\left[\pi_U(\theta_2|\theta_1,\beta,data)\right] \propto n_i\theta_2 \log S_i - \theta_1 S_i^{\theta_2} \sum_{j=1}^{n_i} x_{ij}^{\beta}.$$

Then differentiating with respect to θ_2 :

$$\frac{\partial \log\left[\pi_U(\theta_2|\theta_1,\beta,data)\right]}{\partial\theta_2} = n_i \log S_i - \theta_1 S_i^{\theta_2}\left(\log S_i\right) \sum_{j=1}^{n_i} x_{ij}^{\beta}.$$

Taking the second derivative with respect to θ_2:

$$\frac{\partial^2 \log\left[\pi_U(\theta_2|\theta_1,\beta,data)\right]}{\partial\theta_2^2} = -\theta_1 S_i^{\theta_2}\left(\log S_i\right)^2 \sum_{j=1}^{n_i} x_{ij}^{\beta}$$

which is clearly negative hence $\pi_U(\theta_2|\theta_1,\beta,data)$ is log-concave.

The conditional posterior distribution of β is given by

$$\pi_U(\beta|\theta_1,\theta_2|data) \propto \beta^{n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\exp\left\{-\theta_1 S_i^{\theta_2} \sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}.$$

The log of this expression is given by

$$\log\left[\pi_U(\beta|\theta_1,\theta_2|data)\right] \propto n_i \log\beta + \beta\left(\sum_{j=1}^{n_i}\log x_{ij}\right) - \left(\sum_{j=1}^{n_i}\log x_{ij}\right) - \theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}.$$

Then differentiating with respect to β :

$$\frac{\partial \log\left[\pi_U(\beta|\theta_1,\theta_2|data)\right]}{\partial\beta} = \frac{n_i}{\beta} + \left(\sum_{j=1}^{n_i}\log x_{ij}\right) - \theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\log x_{ij}\right).$$

Taking the second derivative with respect to β:

$$\frac{\partial^2 \log\left[\pi_U(\beta|\theta_1,\theta_2|data)\right]}{\partial\beta^2} = -\frac{n_i}{\beta^2} - \theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\left(\log x_{ij}\right)^2\right)$$

which is negative and therefore $\pi_U(\beta|\theta_1,\theta_2|data)$ is log-concave.

B.2 Log-Concavity of the Conditional Posteriors Using the MDI Prior

The conditional posterior distribution of θ_1 is given by

$$\pi_{MDI}(\theta_1|\theta_2,\beta,data) \propto \theta_1^{\frac{n_i}{\beta}+n_i} \exp\left\{-\theta_1 S_i^{\theta_2} \sum_{j=1}^{n_i} x_{ij}^{\beta}\right\}.$$

The log of this expression is given by

$$\log\left[\pi_{MDI}(\theta_1|\theta_2,\beta,data)\right] \propto \left(\frac{n_i}{\beta}+n_i\right)\log\theta_1 - \theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right).$$

Then differentiating with respect to θ_1 :

$$\frac{\partial \log\left[\pi_{MDI}(\theta_1|\theta_2,\beta,data)\right]}{\partial\theta_1} = \frac{\frac{n_i}{\beta}+n_i}{\theta_1} - S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right).$$

Taking the second derivative with respect to θ_1:

$$\frac{\partial^2 \log\left[\pi_{MDI}(\theta_1|\theta_2,\beta,data)\right]}{\partial\theta_1^2} = -\frac{\left(\frac{n_i}{\beta}+n_i\right)}{\theta_1^2}$$

which is negative and therefore $\pi_{MDI}(\theta_1|\theta_2,\beta,data)$ is log-concave.

The conditional posterior of θ_2 is given by

$$\pi_{MDI}(\theta_2|\theta_1,\beta,data) \propto S_i^{\frac{\theta_2 n_i}{\beta}+\theta_2 n_i} \exp\left\{-\theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right\}.$$

The log of this expression is given by

$$\log\left[\pi_{MDI}(\theta_2|\theta_1,\beta,data)\right] \propto \left(\frac{\theta_2 n_i}{\beta}+\theta_2 n_i\right)\log S_i - \theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right).$$

Then differentiating with respect to θ_2 :

$$\frac{\partial \log\left[\pi_{MDI}(\theta_2|\theta_1,\beta,data)\right]}{\partial\theta_2} = \left(\frac{n_i}{\beta}+n_i\right)\log S_i - \theta_1 S_i^{\theta_2}\left(\log S_i\right)\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right).$$

Taking the second derivative with respect to θ_2:

$$\frac{\partial^2 \log\left[\pi_{MDI}(\theta_2|\theta_1,\beta,data)\right]}{\partial\theta_2^2} = -\theta_1 S_i^{\theta_2}\left(\log S_i\right)^2\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)$$

which is clearly negative hence $\pi_{MDI}(\theta_2|\theta_1,\beta,data)$ is log-concave.

The conditional posterior distribution of β is given by

$$\pi_{MDI}(\beta|\theta_1,\theta_2|data) \propto \beta^{2n_i}\left(\prod_{j=1}^{n_i} x_{ij}^{\beta-1}\right)\theta_1^{\frac{n_i}{\beta}} S_i^{\frac{\theta_2 n_i}{\beta}} \exp\left\{\frac{n_i\gamma}{\beta}\right\}\exp\left\{-\theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\right)\right\}.$$

The log of this expression is given by

$$\log\left[\pi_{MDI}(\beta|\theta_1,\theta_2|data)\right] \propto 2n_i\log\beta + (\beta-1)\left(\sum_{j=1}^{n_i}\log x_{ij}\right) + \frac{n_i}{\beta}\left(\log\theta_1 S_i^{\theta_2}+\gamma\right) - \theta_1 S_i^{\theta_2}\sum_{j=1}^{n_i} x_{ij}^{\beta}.$$

Then differentiating with respect to β :

$$\frac{\partial\log\left[\pi_{MDI}(\beta|\theta_1,\theta_2|data)\right]}{\partial\beta} = \frac{2n_i}{\beta} + \left(\sum_{j=1}^{n_i}\log x_{ij}\right) - \frac{n_i}{\beta^2}\left(\log\theta_1 S_i^{\theta_2}+\gamma\right) - \theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\log x_{ij}\right).$$

Taking the second derivative with respect to β:

$$\frac{\partial^2\log\left[\pi_{MDI}(\beta|\theta_1,\theta_2|data)\right]}{\partial\beta^2} = -\frac{2n_i}{\beta^2} + \frac{2n_i}{\beta^3}\left(\log\theta_1 S_i^{\theta_2}+\gamma\right) - \theta_1 S_i^{\theta_2}\left(\sum_{j=1}^{n_i} x_{ij}^{\beta}\left(\log x_{ij}\right)^2\right)$$

which is negative for $\beta > \left(\log\theta_1 S_i^{\theta_2}+\gamma\right)$ and $n_i > 0$. Therefore $\pi_{MDI}(\beta|\theta_1,\theta_2|data)$ is log-concave.

B.3 Additional Simulation Results

Figures B.1 to B.3 show jump plots created in OpenBUGS. This plots the mean square jumping distance at the nodes with updates averaged over batches of 100 iterations.

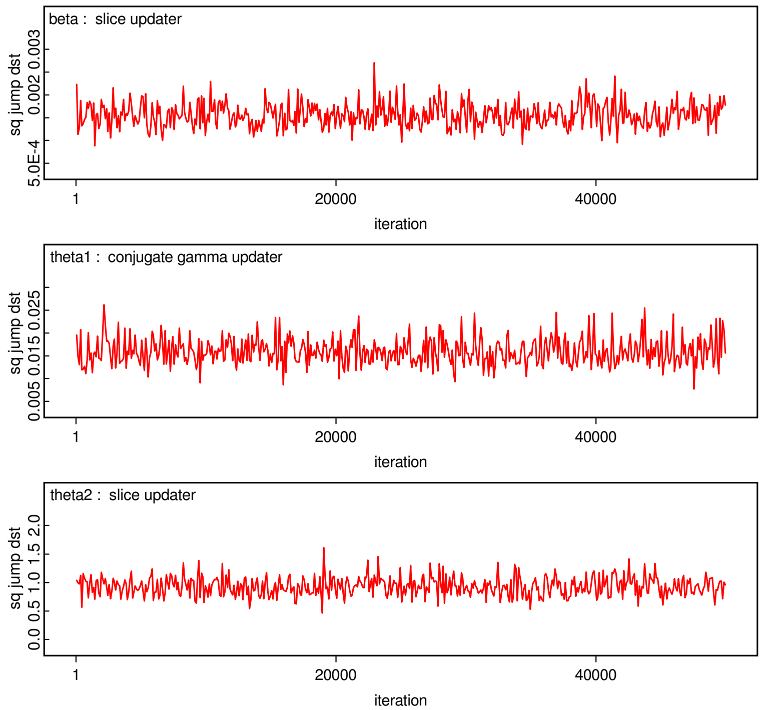


Figure B.1: Jump for prior used in Soyer et al. (2008).

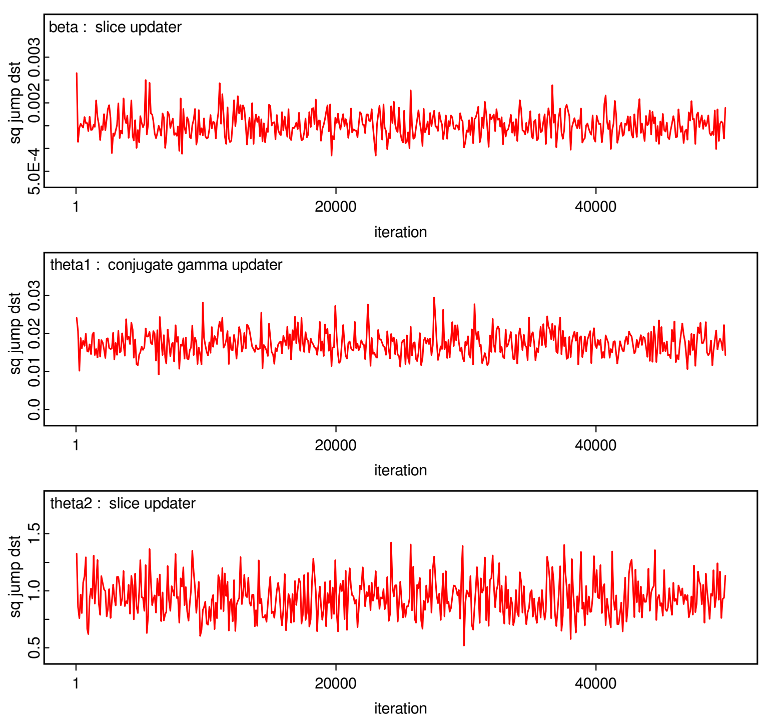


Figure B.2: Jump for uniform prior.

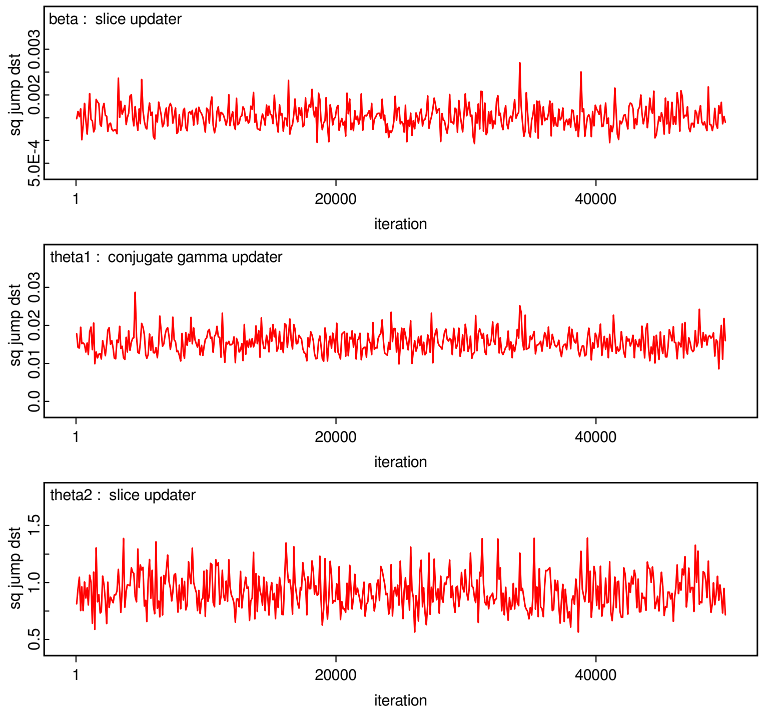


Figure B.3: Jump for mixture of priors.

Figures B.4 to B.6 show history plots. These plot out a complete trace for the variable. This is another method of visually checking the convergence of the chain.

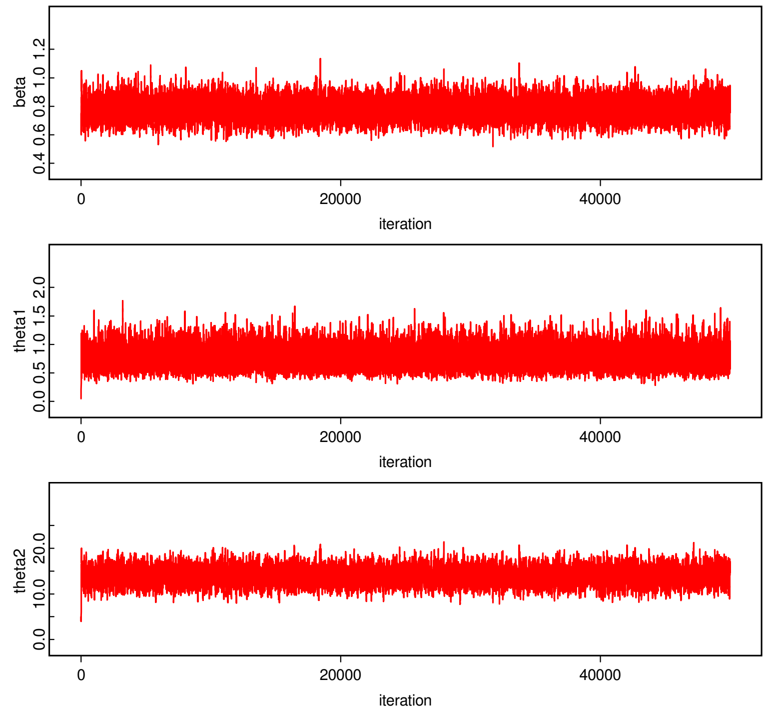


Figure B.4: History for prior used in Soyer et al. (2008).

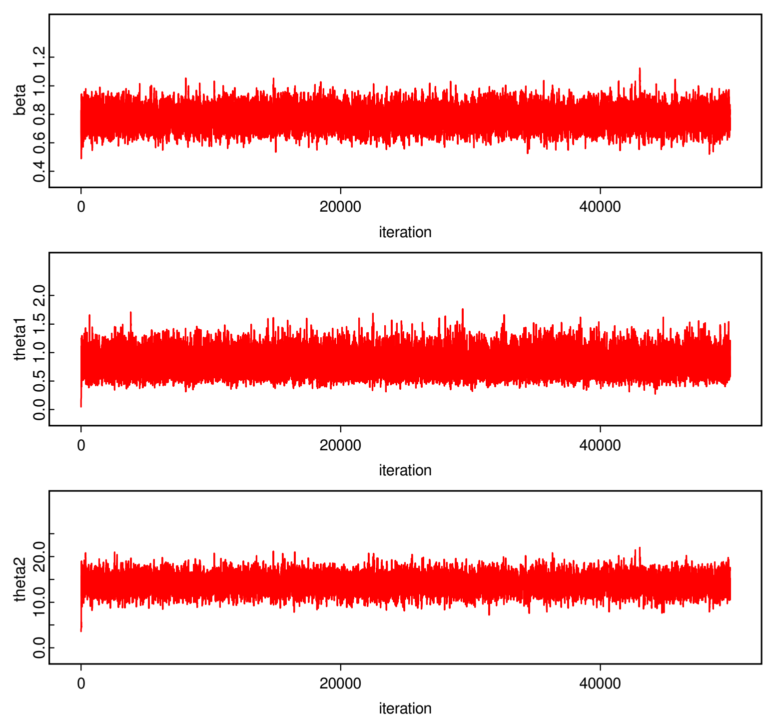


Figure B.5: History for uniform prior.

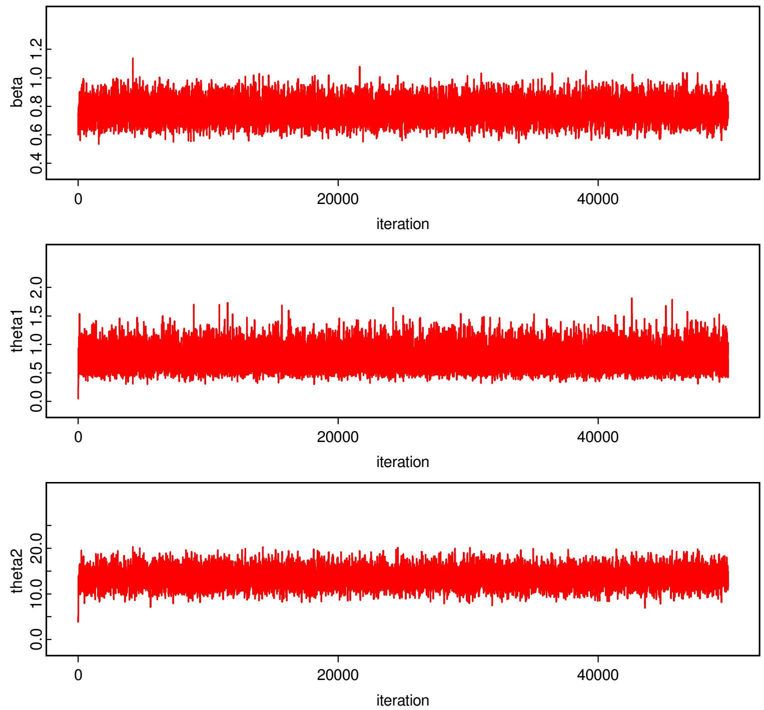


Figure B.6: History for mixture of priors.

Figures B.7 to B.9 show the quantile plots, which shows the mean with 95% confidence intervals against iteration number.

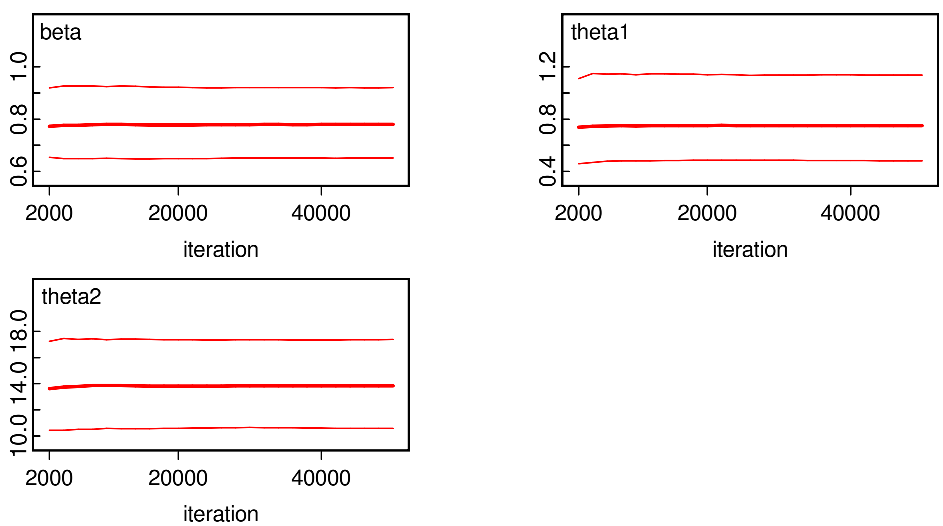


Figure B.7: Running quantiles for prior used in Soyer et al. (2008).

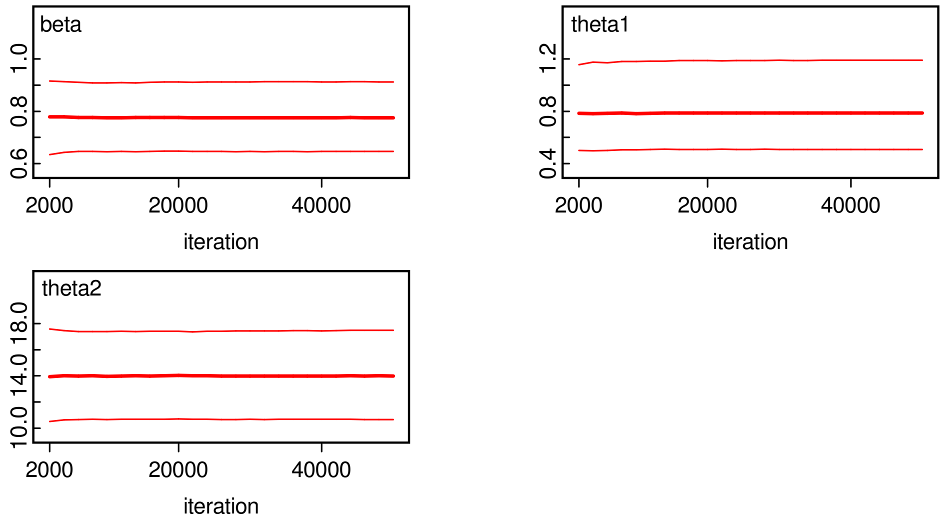


Figure B.8: Running quantiles for uniform prior.

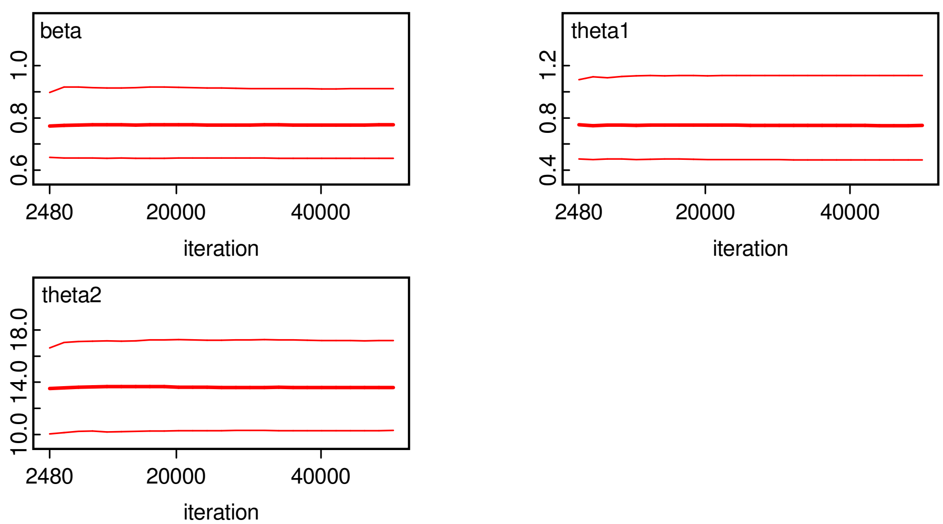


Figure B.9: Running quantiles for mixture of priors.

Appendix C: Code for Chapter 3

MATLAB® code for simulation-based design using non-informative priors

```
clc;
clear;
%Step 1 - Specifying the design
m = 2; %number of stress levels
S1 = 1.05; %stress level one
S2 = 5.4; %stress level two
S_use = 1.05;
S = [S1 S2];
n1 = 1; %number of items tested at stress level 1
n2 = 3; %number of items tested at stress level 2
ni = [n1 n2];
n = n1 + n2;
R = 2000; %number of generated values
varianceLAM = [];
varianceLAMM = [];
varianceLAMJ = [];
varianceLAMG = [];
for i = 1:R
  u = rand(1);
  %Step 2 - Specifying priors
  %Using prior from Erkanli and Soyer
  alpha = gamrnd(20, 1/1000); %a = 20 and b = 1000, generating 100 alpha values
  beta = gamrnd(3, 1/1); % a = 3 and b = 1, generating 100 beta values
  %Using the MDI prior
  alphaM = (u*(n+1))^(1/(n+1));
  betaM = (log(u*(log((S1^n1)*(S2^n2)))))/((S1^n1)*(S2^n2));
  %Using the Jeffreys prior
  alphaJ = exp(u);
```

```
  betaJ = u;

  %Using the GML prior
  alphaG = (0.5*u)^2;
  betaG = u;

  %Step 3 - Generate Ti using the power law lambdai = alpha*(Si^beta),
  where %Ti is a gamma density
  lambda1 = alpha.*S1.^beta;
  lambda2 = alpha.*S2.^beta;
  T1 = gamrnd(n1, 1/lambda1);
  T2 = gamrnd(n2, 1/lambda2);
  lambda1M = alphaM.*S1.^betaM;
  lambda2M = alphaM.*S2.^betaM;
  T1M = gamrnd(n1, 1/lambda1M);
  T2M = gamrnd(n2, 1/lambda2M);
  lambda1J = alphaJ.*S1.^betaJ;
  lambda2J = alphaJ.*S2.^betaJ;
  T1J = gamrnd(n1, 1/lambda1J);
  T2J = gamrnd(n2, 1/lambda2J);
  lambda1G = alphaG.*S1.^betaG;
  lambda2G = alphaG.*S2.^betaG;
  T1G = gamrnd(n1, 1/lambda1G);
  T2G = gamrnd(n2, 1/lambda2G);
%Step 4 - Generating the posteriors of alpha and beta
  nsamples = 5000;
  shape = 3;
  rate = 1 + log((S1^n1)*(S2^n2));
  shapeJ = 1;
  rateJ = log((S1^n1)*(S2^n2));
  shapeG = 1;
  rateG = log((S1^n1)*(S2^n2));
 shapeM = 1;
 rateM = log(((S1^n1)*(S2^n2))^2);
pdf_BETA=@(x)((1000+((S1^x)*T1+(S2^x)*T2))^((-1)*(20+n)))*((S1^(n1*x))
*(S2^(n2*x)))*(x^2)*(exp(-x));
pdf_BETAJ = @(x) ((((S1^x)*T1J+(S2^x)*T2J))^((-1)*n))*((S1^(n1*x))
```

```
*(S2^(n2*x)));
pdf_BETAG = @(x) ((((S1^x)*T1G+(S2^x)*T2G))^((-1)*(n+0.5)))*((S1^(n1*x))
*(S2^(n2*x)));
pdf_BETAM = @(x)((((S1^x)*T1M+(S2^x)*T2M))^((-1)*(2*n+1)))*(((S1^(n1*x))
*(S2^(n2*x)))^2);
proppdf_BETA= @(x,y) gampdf(x, shape,1/rate);
proprnd_BETA = @(x) gamrnd(shape,1/rate);
proppdf_BETAJ= @(x,y) gampdf(x, shapeJ,1/rateJ);
proprnd_BETAJ = @(x) gamrnd(shapeJ,1/rateJ);
proppdf_BETAG= @(x,y) gampdf(x, shapeG,1/rateG);
proprnd_BETAG = @(x) gamrnd(shapeG,1/rateG);
proppdf_BETAM= @(x,y) gampdf(x, shapeM,1/rateM);
proprnd_BETAM = @(x) gamrnd(shapeM,1/rateM);
[MH_BETA,accept_BETA]=mhsample(0.1,nsamples,'pdf',pdf_BETA,'proppdf',proppdf_BETA
,'proprnd',proprnd_BETA);
Pbeta = [MH_BETA];
BETA_post = mean(MH_BETA);
[MH_BETAJ,accept_BETAJ]=mhsample(0.1,nsamples,'pdf',pdf_BETAJ,'proppdf',
proppdf_BETAJ,'proprnd',proprnd_BETAJ);
PbetaJ = [MH_BETAJ];
BETA_postJ = mean(MH_BETAJ);
[MH_BETAG,accept_BETAG]=mhsample(0.1,nsamples,'pdf',pdf_BETAG,'proppdf',
proppdf_BETAG,'proprnd',proprnd_BETAG);
PbetaG = [MH_BETAG];
BETA_postG = mean(MH_BETAG);
[MH_BETAM,accept_BETAM]=mhsample(0.1,nsamples,'pdf',pdf_BETAM,'proppdf',
proppdf_BETAM,'proprnd',proprnd_BETAM);
PbetaM = [MH_BETAM];
BETA_postM = mean(MH_BETAM);
ALPHA = [];
for j = 1:length(MH_BETA)
shape_ALPHA = 20+n;
b = Pbeta(j);
rate_ALPHA = 1000 + T1*S1^b + T2*S2^b;
Palpha = gamrnd(shape_ALPHA,1/rate_ALPHA);
ALPHA = [ALPHA Palpha];
end
```

```
ALPHA_post = mean(ALPHA);
ALPHAJ = [];
for j = 1:length(MH_BETAJ)
shape_ALPHAJ = n;
bJ = PbetaJ(j);
rate_ALPHAJ = T1J*S1^bJ + T2J*S2^bJ;
PalphaJ = gamrnd(shape_ALPHAJ,1/rate_ALPHAJ);
ALPHAJ = [ALPHAJ PalphaJ];
end
ALPHA_postJ = mean(ALPHAJ);
ALPHAG = [];
for j = 1:length(MH_BETAG)
shape_ALPHAG = n+0.5;
bG = PbetaG(j);
rate_ALPHAG = T1G*S1^bG + T2G*S2^bG;
PalphaG = gamrnd(shape_ALPHAG,1/rate_ALPHAG);
ALPHAG = [ALPHAG PalphaG];
end
ALPHA_postG = mean(ALPHAG);
ALPHAM = [];
for j = 1:length(MH_BETAM)
shape_ALPHAM = 2*n+1;
bM = PbetaM(j);
rate_ALPHAM = T1M*S1^bM + T2M*S2^bM;
PalphaM = gamrnd(shape_ALPHAM,1/rate_ALPHAM);
ALPHAM = [ALPHAM PalphaM];
end
ALPHA_postM = mean(ALPHAM);
LAMBDA_post1 = ALPHA_post*(S1^BETA_post);
lamlam = ALPHA'.*(S_use.^MH_BETA);
VARIANCE = var(lamlam);
varianceLAM = [varianceLAM VARIANCE];
ACCEPTANCE = [accept_BETA];
LAMBDA_post1J = ALPHA_postJ*(S1^BETA_postJ);
lamlamJ = ALPHAJ'.*(S_use.^MH_BETAJ);
VARIANCEJ = var(lamlamJ);
varianceLAMJ = [varianceLAMJ VARIANCEJ];
```

```
ACCEPTANCEJ = [accept_BETAJ];
LAMBDA_post1G = ALPHA_postG*(S1^BETA_postG);
lamlamG = ALPHAG'.*(S_use.^MH_BETAG);
VARIANCEG = var(lamlamG);
varianceLAMG = [varianceLAMG VARIANCEG];
ACCEPTANCEG = [accept_BETAG];
LAMBDA_post1M = ALPHA_postM*(S1^BETA_postM);
lamlamM = ALPHAM'.*(S_use.^MH_BETAM);
VARIANCEM = var(lamlamM);
varianceLAMM = [varianceLAMM VARIANCEM];
ACCEPTANCEM = [accept_BETAM];
end
posteriorVAR = mean(varianceLAM);
posteriorVARJ = mean(varianceLAMJ);
posteriorVARG = mean(varianceLAMG);
posteriorVARM = mean(varianceLAMM);
FINAL = [posteriorVAR posteriorVARJ posteriorVARG posteriorVARM]
ACCEPTANCE = [ACCEPTANCEM  ACCEPTANCEJ  ACCEPTANCEG  ACCEPTANCEM ]
histfit(MH_BETA,15)
title('Histogram for \beta using a \Gamma prior for Design 3')
xlabel('\beta')
ylabel('Frequency')
figure
histfit(ALPHA,10, 'gamma')
title('Histogram for \alpha using a \Gamma prior for Design 3')
xlabel('\alpha')
ylabel('Frequency')
figure
hatBETA = cumsum(MH_BETA.^2)./(1:nsamples)';
plot(1:nsamples,hatBETA)
title('Convergence plot for \beta using a \Gamma prior for Design 3')
xlabel('Number of MH samples')
figure histfit(MH_BETAJ,15)
title('Histogram for \beta using the Jeffreys prior for Design 3')
xlabel('\beta')
ylabel('Frequency')
figure
```

```
histfit(ALPHAJ,10, 'gamma')
title('Histogram for \alpha using the Jeffreys prior for Design 3')
xlabel('\alpha')
ylabel('Frequency')
figure
hatBETAJ = cumsum(MH_BETAJ.^2)./(1:nsamples)';
plot(1:nsamples,hatBETAJ)
title('Convergence plot for \beta using the Jeffreys prior for Design 3')
xlabel('Number of MH samples')
figure
histfit(MH_BETAG,15)
title('Histogram for \beta using the GML prior for Design 3')
xlabel('\beta')
ylabel('Frequency')
figure
histfit(ALPHAG,10, 'gamma')
title('Histogram for \alpha using the GML prior for Design 3')
xlabel('\alpha')
ylabel('Frequency')
figure
hatBETAG = cumsum(MH_BETAG.^2)./(1:nsamples)';
plot(1:nsamples,hatBETAG)
title('Convergence plot for \beta using the GML prior for Design 3')
xlabel('Number of MH samples')
figure
histfit(MH_BETAM,15)
title('Histogram for \beta using the MDI prior for Design 3')
xlabel('\beta')
ylabel('Frequency')
figure
histfit(ALPHAM,10, 'gamma')
title('Histogram for \alpha using the MDI prior for Design 3')
xlabel('\alpha')
ylabel('Frequency')
figure
hatBETAM = cumsum(MH_BETAM.^2)./(1:nsamples)';
plot(1:nsamples,hatBETAM)
```

```
title('Convergence plot for \beta using the MDI prior for Design 3')
xlabel('Number of MH samples')
```

Appendix D: Code for Chapter 4

<u>**OpenBUGS code using the priors from Soyer**</u>

```
model
{ for ( i in 1:N){ lambda[i]<-theta1*pow(Si[i]/38,theta2)
TI[i]~dweib(beta, lambda[i]) }
beta~dunif(0,10) theta1~dgamma(0.01,0.01) theta2~dunif(0,100)
}
Data
list(TI=c(0.0900000, 0.3900000,0.4700000,0.7300000,0.7400000,1.1300000,1.4000000,
2.3800000,0.3500000,0.5900000,0.96, 0.9900000, 1.6900000,1.9700000,
2.0700000,  2.5800000,  2.7100000,  2.9000000,  3.6700000,
3.9900000,  5.3500000, 13.77000, 25.5000000,  0.1900000,
0.7800000,  0.9600000,  1.3100000,  2.7800000,  3.1600000,
4.1500000,  4.6700000,  4.850, 6.50,  7.35,  8.01,  8.27,
12.060000, 31.7500000, 32.520000, 33.910, 36.71, 72.89,
0.2700000,  0.40,  0.69,  0.79,  2.75,  3.91,  9.880000,
13.95, 15.93, 27.80, 53.24, 82.850, 89.290, 100.58, 215.10,
7.740, 17.05, 20.460, 21.02, 22.66, 43.40, 47.30, 139.07,
141.12, 175.880, 194.9,  68.85, 108.29,  110.29,  426.07,
1067.6,   5.79, 1579.52, 2323.7 ),
N=76,
Si=c(38, 38, 38, 38, 38, 38, 38,  38,  36, 36, 36, 36, 36,
36,  36, 36, 36, 36, 36,  36, 36, 36, 36, 34,  34, 34, 34,
34, 34, 34, 34, 34, 34, 34, 34,  34, 34,  34, 34, 34, 34,
34, 32,  32,  32,  32,  32, 32, 32, 32, 32,  32, 32, 32,
32,  32,  32,  30, 30, 30, 30, 30, 30, 30,  30, 30, 30,
30, 28,  28, 28, 28, 28, 26, 26, 26))
Initials
```

```
list(beta=1,theta1=0.001,theta2=5)
```

OpenBUGS code using the uniform prior

```
model
{ for ( i in 1:N){ lambda[i]<-theta1*pow(Si[i]/38,theta2)
TI[i]~dweib(beta, lambda[i]) }
beta~dunif(0,20) theta1~dunif(0,20) theta2~dunif(0,60)
}
Data
list(TI=c(0.0900000, 0.3900000,0.4700000,0.7300000,0.7400000,1.1300000,1.4000000,
2.3800000,0.3500000,0.5900000,0.96, 0.9900000, 1.6900000,1.9700000,
2.0700000,  2.5800000,  2.7100000,  2.9000000,  3.6700000,  3.9900000,
5.3500000, 13.77000, 25.5000000,  0.1900000,  0.7800000,  0.9600000,
1.3100000,  2.7800000,  3.1600000,  4.1500000,  4.6700000,  4.850,
6.50,  7.35,  8.01,  8.27, 12.060000, 31.7500000, 32.520000, 33.910,
36.71, 72.89, 0.2700000,  0.40,  0.69,  0.79,  2.75,  3.91,  9.880000,
13.95, 15.93, 27.80, 53.24, 82.850, 89.290, 100.58, 215.10,  7.740,
17.05, 20.460, 21.02, 22.66, 43.40, 47.30, 139.07, 141.12, 175.880,
194.9,  68.85, 108.29,  110.29,  426.07, 1067.6,   5.79, 1579.52,
2323.7 ),
N=76,
Si=c(38, 38, 38, 38, 38, 38, 38,  38,  36, 36, 36, 36, 36,  36,
36, 36, 36, 36, 36,  36, 36, 36, 36, 34,  34, 34, 34, 34, 34, 34,
34, 34, 34, 34, 34,  34, 34,  34, 34, 34, 34, 34, 32,  32,  32,
32,  32, 32, 32, 32, 32,  32, 32, 32, 32,  32,  32,  30, 30, 30,
30, 30, 30, 30,  30, 30, 30, 30, 28,  28, 28, 28, 28, 26, 26, 26))
Initials
list(beta=1,theta1=0.001,theta2=5)
```

OpenBUGS code using the mixture of uniform and Gamma priors

```
model
{ for ( i in 1:N){ lambda[i]<-theta1*pow(Si[i]/38,theta2)
TI[i]~dweib(beta, lambda[i]) }
beta~dunif(0,5) theta1~dgamma(0.001,0.001) theta2~dgamma(0.001,0.001)
}
Data
list(TI=c(0.0900000, 0.3900000,0.4700000,0.7300000,0.7400000,1.1300000,
```

```
1.4000000,2.3800000,0.3500000,0.5900000,0.96, 0.9900000, 1.6900000,
1.9700000, 2.0700000,  2.5800000,  2.7100000,  2.9000000,
3.6700000,  3.9900000,  5.3500000, 13.77000, 25.5000000,
0.1900000,  0.7800000,  0.9600000,  1.3100000,  2.7800000,
3.1600000,  4.1500000,  4.6700000,  4.850, 6.50,  7.35,  8.01,
8.27, 12.060000, 31.7500000, 32.520000, 33.910, 36.71, 72.89,
0.2700000,  0.40,  0.69,  0.79,  2.75,  3.91,  9.880000, 13.95,
15.93, 27.80, 53.24, 82.850, 89.290, 100.58, 215.10,  7.740, 17.05,
20.460, 21.02, 22.66, 43.40, 47.30, 139.07, 141.12, 175.880, 194.9,
68.85, 108.29,  110.29,  426.07, 1067.6,   5.79, 1579.52, 2323.7 ),
N=76,
Si=c(38, 38, 38, 38, 38, 38, 38,  38,  36, 36, 36, 36, 36,
36,  36, 36, 36, 36, 36,  36, 36, 36, 36, 34,  34, 34, 34,
34, 34, 34, 34, 34, 34, 34, 34,  34, 34,  34, 34, 34, 34,
34, 32,  32,  32,  32,  32, 32, 32, 32, 32,  32, 32, 32, 32,
32,  32,  30, 30, 30, 30, 30, 30, 30,  30, 30, 30, 30, 28,
28, 28, 28, 28, 26, 26, 26))
Initials
list(beta=1,theta1=0.001,theta2)
```

R® code for calculating the predictive reliability

```
data = read.table("C:/.../.....txt", sep="\t", header = T)
bet = data$beta
thet1 = data$theta1
thet2 = data$theta2
Su = 22
lam = thet1*((Su/38)^thet2)
x = seq(1,100000,1)
result = vector("list")
for (i in 1:100000)
{
rel = exp(-1*lam*(i^bet))
Pred_rel = mean(rel)
result[i]=(Pred_rel)
}
options(scipen=5)
plot(x,result, xlab = "x", ylab = "Reliability")
```