A comparison of life distributions in Bayesian reliability theory

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

Master of Science

in

Mathematical Statistics

Department of Statistics

Rhodes University

by

Sinoyolo Ntanjana

Student number: 19N8018

ORCID iD:0000-0001-5559-1141

December 2025

Supervisor: Dr. S. R. Izally ORCID iD:0000-0002-768656-138X

Co-Supervisor: Prof. L. Raubenheimer ORCID iD:0000-0003-0756-7459

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained herein is my own, original work, that I am the sole author hereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Rhodes University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

S. Ntanjana | Date

Acknowledgements

I would like to begin by giving thanks to God for His guidance, strength, and wisdom throughout this journey. I would like to express my sincere gratitude to Dr S. R. Izally and Prof. L. Raubenheimer for their continuous support, guidance, and meaningful feedback and meetings. I would also like to express my sincere gratitude to my family and the Department of Statistics. I would like to express my sincere gratitude to the NRF Innovation Master's Scholarship for providing financial support throughout my two-year Master's programme. Their funding made this work possible.

Abstract

Reliability analysis plays a fundamental role in assessing the lifetime behaviour of components, systems, and materials. In the Bayesian approach, the uncertainty about model parameters can be measured using the posterior distribution. This study presents a comparative analysis of two life distributions, the Weibull and Birnbaum-Saunders distributions under Bayesian reliability theory. The study focuses on the derivation of posterior distributions using a range of objective priors, including the Jeffreys prior, divergence prior, reference prior, and the probability matching prior, for both complete and type I right censoring cases. These priors are derived from the Fisher information matrix for both models, and the properness of the resulting posterior distributions is examined both graphically and analytically. Markov Chain Monte Carlo techniques, including the Metropolis-Hastings sampler, Gibbs sampler, and the Metropolis-within-Gibbs algorithm, are employed to simulate from the posterior distributions of the model parameters. Convergence of the posterior samples is assessed using standard diagnostics such as the trace plots, the Gelman-Rubin convergence diagnostic, and the Geweke diagnostic. Simulation studies are conducted to assess model performance across different sample sizes and priors, with evaluation based on coverage rates and mean interval lengths. Predictive reliability analyses are performed to analyse the ability of both distributions to predict future lifetimes. Applications include fitting and evaluating two fatigue lifetime datasets using both the Weibull and the Birnbaum-Saunders distributions. Bayesian estimation is carried out, and posterior summaries are analysed to assess parameter behaviour, credible intervals, and overall model fit. Model comparison using the deviance information criterion (DIC) is performed to determine which distribution provides a better fit and more stable parameter estimates.

Keywords: Bayesian inference, Bayesian reliability theory, Birnbaum-Saunders distribution, Censoring, Coverage rate, Credibility Interval, Divergence prior, Gibbs sampler, Jeffreys prior, Likelihood, MCMC, Mean interval lengths, Metropolis-Hastings sampler, Posterior convergence, Posterior divergence, Predictive reliability, Reference prior, Weibull distribution.

Contents

List of Figures

List of Tables

List of Abbreviations

CDF	Cumulative Density Function
CI	Credibility Interval
DIC	Deviance Information Criterion
MCMC	Markov Chain Monte Carlo
MH	Metropolis-Hastings
PDF	Probability Density Function
PMC	Population Monte Carlo
MvN	Multivariate Normal Distribution

List of Distributions

- **Weibull distribution**

  Parameters α, β where $\alpha > 0$ and $\beta > 0$.

  Probability density function:

  $f(t|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\},\ t > 0\ .$

  Moments

  $E(T) = \alpha\Gamma\left(1+\frac{1}{\beta}\right).$

  $Var(T) = \alpha^2\left\{\Gamma\left(1+\frac{2}{\beta}\right) - \left[\Gamma\left(1+\frac{1}{\beta}\right)\right]^2\right\}.$

- **Birnbaum-Saunders distribution**

  Parameters α, β where $\alpha > 0$ and $\beta > 0$.

  Probability density function:

  $f(t|\alpha,\beta) = \frac{\sqrt{\frac{t}{\beta}}+\sqrt{\frac{\beta}{t}}}{2\sqrt{2\pi}\alpha t}\exp\left\{-\frac{1}{2\alpha}\left(\sqrt{\frac{t}{\beta}}+\sqrt{\frac{\beta}{t}}\right)^2\right\}, t > 0.$

  Moments

  $E(T) = \beta\left(1+\frac{\alpha^2}{2}\right).$

  $Var(T) = \left(\alpha^2\beta^2\right)\left(1+\frac{5\alpha^2}{4}\right).$

Notation

t	observation from random variable T
$\underline{t}$	denotes observed data
$f(t)$	PDF of random variable T
$F(t)$	CDF of random variable T
$R(t)$	reliability function
$L(\theta \| data)$	likelihood function
$H(\theta)$	Fisher information matrix
$\pi(\theta)$	prior of θ
$\pi(\theta \| data)$	posterior distribution of θ
erf	error function
$\Phi(\cdot)$	CDF for the standard normal distribution
$\phi(\cdot)$	PDF for the standard normal distribution
$\lvert \cdot \rvert$	determinant
$\lvert\lvert \cdot \rvert\rvert$	length/ absolute value
n	sample size
τ	time
r	number of failures
ℓ	log likelihood
MvN	multivariate normal distribution
T	denotes transpose

Chapter 1

Introduction

1.1 Overview

This thesis looks at how Bayesian methods can be used to study the reliability of systems or components. Reliability analysis plays a crucial role in many industrial and engineering applications, where accurate modelling of system lifetimes is essential for maintenance planning and risk assessment. In many real-life cases, data may be censored, which makes traditional statistical methods less effective. Such challenges commonly arise in practical settings such as manufacturing systems and reliability testing, motivating the need for flexible Bayesian approaches that can adequately account for uncertainty and limited data. In industries like aerospace and manufacturing, where lifetime predictions of critical components are essential for safety and cost efficiency, censored data often complicates traditional reliability analyses. By using Bayesian methods, which can handle uncertainty from incomplete data, this study aims to provide more accurate models for real-world applications such as turbine failure prediction and quality control in production lines. The Bayesian approach helps to deal with these problems by combining background knowledge with observed data to form a posterior distribution. The focus of this thesis is on comparing two important life distributions, the Weibull distribution and the Birnbaum-Saunders distribution. Different types of priors are studied and derived, including the Jeffreys prior, the divergence prior, the reference prior, and the probability matching prior. The corresponding posterior distributions are derived, and their properness is evaluated. Predictive reliability is also calculated from the posterior distributions to see how the reliability of a system changes over time. Furthermore, a simulation study was conducted to evaluate the frequentist properties of the Bayesian estimators. In particular, the study focused on the coverage rate and mean interval length of the 95% credible intervals for the model parameters under different prior specifications and sample sizes. The simulation results provide insight into the performance of each prior, especially for small sample sizes. In the application section, two lifetime examples were conducted to illustrate the practical implementation of the Bayesian methods under complete data. Posterior summaries, credibility intervals, and posterior curves were presented for each case.

1.2 Objectives

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

- This study aims to compare different life distribution models within the Bayesian framework.
- To investigate the Fisher information matrix for both the Weibull and Birnbaum-Saunders distribution, which serves as the foundation for constructing objective priors.
- To derive the reference prior, Jeffreys prior, probability matching prior and the divergence prior for each distribution.
- To derive and analyse the relevant posterior distributions under complete data and type I right censoring.
- To conduct simulation studies, focusing on coverage rates and mean interval lengths.
- To derive the predictive reliability.
- To demonstrate the practical application using two real lifetime data examples.
- Conduct a sensitivity analysis for the choice of prior.

1.3 Contributions

Given the objectives, the principal contributions of the study are summarised as follows

- The derivation of the divergence prior for the Weibull distribution from the Fisher information and proving that the corresponding posterior is proper under no censoring and type I right censoring.
- The derivation of the reference prior for the Weibull distribution using the algorithm by Berger and Bernardo (1992) and proving that the corresponding posterior is proper under no censoring and type I right censoring.
- The derivation of the probability matching prior for the Weibull distribution, which was shown to be equivalent to the reference prior.
- The derivation of the full Fisher information matrix for the Birnbaum–Saunders distribution for complete data. The derivation presents every algebraic step omitted in previous work by Lemoine (2019) and provides closed-form expressions for the expected second derivatives of the log-likelihood. These results are reported in full in Appendix A.2.

- The evaluation of the objective priors for the Birnbaum-Saunder distribution using the Fisher information matrix, and the demonstration that the priors lead to improper posterior distribution under complete data.
- Assessment of posterior convergence using Markov Chain Monte Carlo (MCMC) diagnostics, including trace plots, the Gelman-Rubin diagnostic, and the Geweke diagnostic test.
- Evaluation of the coverage rate and mean interval length of the posterior credible intervals to assess the accuracy and efficiency of Bayesian inference under different prior specifications.

1.4 Outline

This study focuses on the comparison of two important life distributions: the Weibull and Birnbaum–Saunders distributions within a Bayesian reliability framework. Both models are widely used in reliability analysis and fatigue life modeling. The objective of this research is to investigate the posterior behaviour, parameter estimation, and reliability performance of these models using simulation and real data applications. The comparison is performed under complete data conditions, using non-informative and weakly informative priors, and inference is conducted through MCMC methods implemented in OpenBUGS (version 3.2.3) and R (version 4.5.1). All simulations and analyses were conducted using R (version 4.5.1), R. Core Team (2025) and OpenBUGS (version 3.2.3), Lunn et al. (2009).

In **Chapter** 2, a detailed literature review is presented, outlining the theoretical developments and previous studies related to Bayesian reliability analysis, life distributions, and objective prior construction, as well as the implementation of MCMC methods and assessments of MCMC convergence.

In **Chapter** 3, the Weibull distribution is introduced and discussed in detail. The chapter presents its probability density and reliability functions. Posterior distributions were derived using non-informative priors. The properness of the posterior is evaluated both numerically and graphically both under complete data and type I censoring to ensure valid inference. Simulation studies are conducted to assess parameter coverage probabilities, and mean interval lengths.

In **Chapter** 4, the Birnbaum-Saunders distribution is introduced and discussed in detail. The chapter presents its probability density and reliability functions. Posterior distributions were derived using both non-informative and weak informative priors. The properness of the posterior is evaluated both numerically and graphically both under complete data and type I censoring to ensure valid inference. Simulation studies are conducted to assess parameter coverage probabilities, and interval lengths.

In **Chapter** 5, the application of the Weibull and Birnbaum-Saunders models to complete lifetime datasets. The objective is to estimate α and β for both distributions and to evaluate distributions using the deviance information criterion (DIC). Two examples are analyzed: the first example uses a small dataset to illustrate the Bayesian estimation process, and the second example applies the same procedure to a larger dataset of 101 fatigue lifetimes. Posterior summaries, credible intervals, density plots, and Q-Q plots are presented for both distributions.

In **Chapter 6,** the main findings of the study are summarised, including the performance of different objective priors for the Weibull and Birnbaum-Saunders distributions across varying sample sizes. The chapter also discusses MCMC convergence issues observed in small sample sizes. Finally, **Chapter 6** presents recommendations for future research.

Appendix A contains the derivations of the Fisher information matrix for both the Weibull and Birnbaum–Saunders distributions. Detailed calculations are provided for the complete data case, including the second-order derivatives of the log-likelihood functions with respect to the distribution parameters.

Appendix B contains the evaluation of properness of the posterior distribution when using the reference prior for the Birnbaum–Saunders distribution when there is no censoring.

Appendix C contains the R and OpenBUGS code for Bayesian simulation and application of the Weibull and Birnbaum-Saunders distributions.

Chapter 2

Literature Review

2.1 Background and Related Studies

The modelling of lifetime data through parametric distribution is fundamental in reliability theory, survival analysis, and fatigue life modelling. Among various lifetime distributions, the Weibull and Birnbaum-Saunders distributions are widely studied due to their ability to capture failure mechanisms.

2.1.1 Bayesian Analysis of the Weibull Distribution

The Weibull distribution, defined by a shape and scale parameter, is one of the most commonly used lifetime models due to its flexibility in representing increasing, decreasing, or constant rates. Bayesian approaches for the Weibull parameter estimation often rely on assigning prior distributions for the parameter, such as Gamma, Jeffreys, or reference priors, and then deriving posterior inferences using computational methods, including Gibbs sampling or other MCMC techniques, Kundu and Mitra (2016).

- Kundu and Mitra (2016) provide a detailed Bayesian inference framework for the Weibull model with left truncated and right censored data, using Gibbs sampling to compute Bayesian estimates of both shape and scale parameters, along with credible intervals and predictive densities.
- Ramos et al. (2020) investigate objective priors, including the Jeffreys prior and the reference prior, and derive sufficient conditions under which improper priors produce a proper posterior distribution for the Weibull parameters when data is censored.

Objective Bayesian methods for Weibull parameters and functions have been discussed in the context of reliability analysis, highlighting derivations of the Jeffreys prior, the reference prior, and other objective priors to improve inference under limited or censored data, Consonni et al. (2018).

Although Bayesian methods for estimating Weibull parameters have been widely studied, including posterior properties and simulation-based approaches, there are still few studies that compare the posterior behaviour of the Weibull distribution with other lifetime distributions, under different priors and censoring schemes. In particular, comparisons focusing on predictive reliability and coverage performance are limited.

2.1.2 Bayesian Analysis of the Birnbaum-Saunders Distribution

The Birnbaum-Saunders distribution, defined by a shape and scale parameter, is a lifetime model originally developed to describe fatigue failure processes, Birnbaum and Saunders (1969a). It is particularly suitable for modelling positively skewed lifetime data arising from cumulative damage mechanisms. The Birnbaum-Saunders distribution was proposed to model fatigue life, based on crack growth under cyclic loading. It has since been widely used in reliability and lifetime studies and has been the focus of numerous Bayesian estimation approaches.

- In the literature, Bayesian methods have been used for inference in the Birnbaum-Saunders fatigue life model. These studies focus on parameter estimation, posterior analysis, and reliability measures, and demonstrate that the Birnbaum-Saunders distribution is well-suited for modelling fatigue life data, Achcar (1993). In addition, Bayesian analysis of the Birnbaum-Saunders distribution has been considered in situations where only partial information about the data is available, Xu and Tang (2011). These studies develop Bayesian inference procedures for parameter estimation and reliability analysis when complete failure times are not fully observed, such as under censoring or limited data conditions. The results show that the Birnbaum–Saunders distribution remains flexible and effective for modelling fatigue life data even when information is incomplete.

- Early work on Bayesian Birnbaum-Saunders estimation under reference priors and computational methods demonstrated that Bayesian estimators can outperform classical methods, including maximum likelihood estimation for certain parameter settings, Xu and Tang (2010).

- Recent studies also investigate improved Bayesian inferences for right-censored Birnbaum-Saunders data, leveraging Gibbs sampling under the Jeffreys and reference priors and comparing performance across varying sample sizes and censoring levels, Jayalath (2024). These studies conclude that Bayesian inference for the Birnbaum-Saunders distribution under right-censored data, implemented via Gibbs sampling with the Jeffreys or reference priors, provides stable and accurate estimates, with performance improving as sample size increases and censoring decreases.

- Recent scientometric reviews show that the Birnbaum-Saunders distribution has been widely adopted for lifetime analysis, extended to regression and multivariate models, and studied using Bayesian and other estimation techniques, Victor (2025).

- Smit (2021) developed Bayesian dual-stress accelerated life testing models using the generalised Eyring model with one thermal and one non-thermal stressor. The Weibull and Birnbaum-Saunders distributions were used as lifetime models, with likelihoods accommodating uncensored, type I, and type II censored data. Posterior distributions were mathematically intractable, requiring MCMC methods for inference and convergence checks via trace plots and the Brooks-Gelman-Rubin diagnostic. Model fit was assessed using the DIC, and predictive reliability and Bayes factors were explored. The study highlights sensitivity to prior choices, recommending flat priors when prior knowledge is limited.

In industries like aerospace and manufacturing, where lifetime predictions of critical components are essential for safety and cost efficiency, censored data often complicates traditional reliability analyses. By using Bayesian methods, which can handle uncertainty from incomplete data, this study aims to provide more accurate models for real-world applications such as turbine failure prediction and quality control in production lines. Although both the Weibull and Birnbaum-Saunders distributions are widely used in reliability analysis, direct Bayesian comparisons of these models under different censoring schemes and objective priors remain limited. In particular, the behaviour of posterior coverage and predictive reliability has not been thoroughly examined, especially for small sample sizes or heavily censored data. Furthermore, the application of Bayesian model selection criteria, such as the DIC, across real-world and simulated reliability scenarios is still sparse. By addressing these gaps, the present study contributes to the reliability literature by systematically comparing the posterior inference and predictive performance of the Weibull and Birnbaum-Saunders distributions within a Bayesian framework using objective priors.

2.1.3 Reliability Function

The function $R(t)$ is the probability that the time to failure is higher than or equal to t for a given value of $t \geq 0$. Let T be the time at which a failure event occurs. If T is a non-negative random variable which has a (continuous) cumulative distribution function (CDF) denoted by $F(t)$, and a corresponding probability density function (PDF) denoted by $f(t)$, then the reliability function is given by

$$R(t) = P(T \geq t) = 1 - F(t),$$

where $T \geq 0$, $R(t) \geq 0$, $F(0) = 0$, and $\lim_{t \to \infty} F(t) = 1$. Another significant function related to the life distribution is the failure rate, or hazard function, $h(t)$, and is given by

$$\begin{aligned} h(t) &= \frac{f(t)}{1 - F(t)} \\ &= \frac{f(t)}{R(t)}. \end{aligned}$$

Reliability plays a crucial role in the design of products and systems, and to estimate how dependable a proposed design will be, engineers employ various modeling approaches, including probabilistic models that assess system reliability, Ruggeri et al. (2008). As stated by Ross (2014), the reliability theory is mainly concerned with the determination of the probability that a system, consisting possibly of several components, will operate adequately for a given period of time in its intended application.

2.1.3.1 Predictive Reliability

The predictive reliability predicts the probability that a future unit will survive after a specified time, given the observed data and prior information. The predictive reliability is given by

$$R(t_u|data) = \int R(t_u|\theta)\,\pi(\theta|data)\,d\theta, \tag{2.1}$$

where $R(t_u)$ is a reliability function at some time t_u.
The evaluation of $R(t_u|data)$ can be done as follow

1. Sample θ from the posterior a sufficiently large number, say N times.
2. The integral in equation 2.1 can be calculated by the Monte Carlo average

$$R(t_u|data) \approx \frac{1}{N}\sum_{n=1}^{N} R\left(t_u|\theta^{(n)}\right), \tag{2.2}$$

which is the expected reliability at time t_u, using the posterior sample $\left\{\theta^{(n)}\right\}, n = 1,...,N.$

2.1.4 Censoring

According to Leung et al. (1997), censoring occurs when incomplete information is available about the survival time of some systems. Censoring is a common phenomenon in survival analysis and reliability studies, where the exact time of an event of interest is not observed or unavailable, Klein and Moeschberger (2003). These methods allow researchers to account for censored observations and estimate the survival function, hazard rate, or other quantities of interest. There are several types of censoring, but the most common ones are: right censoring (type I and type II), left censoring and interval censoring.

- Tian et al. (2024) defines right censoring arises when one or more units have not failed when the data are analyzed and occur for different reasons. Right censoring is a widely utilized approach in survival analysis and reliability studies, which involves several distinct types of censoring schemes.

- Type I right censoring is defined by Gijbels (2010) as censoring that occurs if an experiment is started at a given time for a set of components or items, and the experiment is stopped at a predetermined time.

- Type II right censoring occurs when an experiment is continued until a predetermined number of the subjects under study have failed, Gijbels (2010).

- According to Klein and Moeschberger (2003), left censoring happens when the event of interest occurred between two time points, but the exact time is unknown. For example, if a subject reports that the event happened sometime last month.

- Smit (2021) defines interval censoring occurs when the exact failure time of an item is unknown, but it is known to lie within a specific time interval.

The focus of this thesis will be on no censoring and type I right censoring. Under type I right censoring, each observation i consists of the pair

$$data_i = (t_i, \delta_i),$$

where t_i is either an exact failure time or a censored time, and

$$\delta_i = \begin{cases} 1, & \text{if } t_i \text{ is an exact failure time,} \\ 0, & \text{if } t_i \text{ is a right censored observation.} \end{cases}$$

Let $f(t_i|\theta)$ denote the probability density function of the specified lifetime distribution with parameter vector θ, and let $R(t_i \mid \theta)$ denote the corresponding reliability function. The likelihood of each observation is then

$$\{f(t_i|\theta)\}^{\delta_i} \{R(t_i \mid \theta)\}^{\delta_i},$$

so the likelihood for the complete sample is

$$L(\alpha, \beta|\underline{t}) = \prod_{i=1}^{n} \{f(t_i|\alpha, \beta)\}^{\delta_i} \{R(t_i|\alpha, \beta)\}^{1-\delta_i}. \tag{2.3}$$

Let r denote number of failures in the sample, the data may be reordered such that the first r observations correspond to failures

$$\delta_1 = \cdots \delta_r = 1, \qquad \delta_{r+1} = \cdots \delta_n = 0.$$

Substituting these values into equation 2.3, the likelihood with respect to type I right censoring can be rewritten as

$$L(\alpha, \beta|\underline{t}) = \prod_{i=1}^{r} \{f(t_i|\alpha, \beta)\} \prod_{i=r+1}^{n} \{R(t_i|\alpha, \beta)\},$$

where r is the number of failures and n is the sample size.

2.2 Bayesian Statistics

The foundational concepts of Bayesian statistics originated with the work of Bayes (1763). In his essay, he described Bayes Theorem using inverse probability and explained how subjective views or a researcher's understanding may be used to make statistical inferences. This explanation gave rise to what are now referred to as posterior and prior beliefs. Prior beliefs are initially stated regarding an event, and these are subsequently adjusted or revised to form posterior beliefs as additional information becomes available. Spiegelhalter and Rice (2009) mentioned that, the Bayesian statistical methods are commonly used when there is no other choice but to incorporate quantitative prior judgements because there is a lack of data on a particular feature of a model or because it is necessary to make assumptions about the biases involved in order to accept the shortcomings of some evidence. Within the context of Bayesian statistics, existing knowledge is formulated as a prior distribution and combined with observed data as a likelihood function.

2.3 Bayesian Inference

Bayesian inference considers the data to be fixed, and unknowns parameters say θ, are treated as random variables and follow some distribution. The prior distribution, denoted as $\pi(\theta)$, expresses the belief or knowledge about the possible values of θ before any data is observed. Bayesian inference utilizes prior knowledge along with the observed data (likelihood), Ellison (2004). According to Gelman (2002), the prior distribution is a key part of Bayesian inference and represents the information about an unknown parameter that is combined with the probability distribution of new data to yield the posterior distribution which in turn is used for future inferences and decisions involving θ. The likelihood function is the joint probability density of observed data viewed as a function of the parameters of a statistical model. It is denoted as $L(\theta|data)$, where θ represents the model parameters and $data$ is the observed data. In Bayesian statistics, once a prior is selected, it is combined with the likelihood function to calculate the posterior distribution.

The posterior distribution for continuous random variables will be

$$\pi(\theta|data) = \frac{L(\theta|data)\pi(\theta)}{\int L(\theta|data)\pi(\theta)d\theta}. \tag{2.4}$$

The posterior distribution for discrete random variables will be

$$\pi(\theta|data) = \frac{L(\theta|data)\pi(\theta)}{\sum L(\theta|data)\pi(\theta)}, \tag{2.5}$$

where θ is the parameter vector, $\pi(\theta)$ denotes the prior distribution, $L(\theta|data)$ is the likelihood function and the denominator is the normalizing constant.

By using Bayes rule the posterior distribution can be written as

$$\pi(\theta|data) \propto \pi(\theta) \times L(\theta|data)$$
$$posterior \propto prior \times likelihood.$$

Bayesian inferences are derived from the posterior distribution. The posterior distribution state what can be said about unknown parameters of interest given available data and prior knowledge. The posterior can be used for making predictions about future events, van de Schoot et al. (2021). According to Spiegelhalter and Rice (2009), a complete posterior distribution report is the accurate and definitive outcome of a statistical analysis. Prior selection is one of the most difficult and challenging problems in Bayesian analysis.

2.4 Non-informative Prior Distributions

In this thesis the focus will be on non-informative priors. Non-informative priors are frequently improper, meaning they do not integrate to one, and are also referred to as objective, vague, or flat priors. Lemoine (2019) state that non-informative priors are distributions that are flat over the entire real number line and thus contain little or no information. Bayesians use non-informative priors when there is desire to analyze data without using any information that might be available to specify a prior distribution also known as objective Bayesian analysis, Tian et al. (2024). There is little effect of non-informative priors on the posterior distribution, Sadok et al. (2023). Sadok et al. (2023) state that using a non-informative prior is recommended since it is flexible and can cover a variety of distributions. The following objective priors will be considered in this thesis: the Jeffreys prior, divergence prior, reference prior, and the probability matching prior.

2.4.1 The Jeffreys Prior

In Bayesian statistics, the Jeffreys prior is regarded as a non-informative prior distribution for a parameter space. It was proposed by Jeffreys (1939) and is proportional to the square root of the determinant of the Fisher information matrix. The Jeffreys prior is expressed as

$$\pi_J(\theta) \propto \sqrt{|H(\theta)|}, \tag{2.6}$$

where $H(\theta)$ denotes the Fisher information matrix.

The Jeffreys prior, even if there is more than one parameter, is well known to be invariant to reparameterization. In Bayesian statistics, by saying “invariance” it means that if $\pi(\theta)$, the PDF of θ, is a specific prior distribution for the parameter of interest $t(\theta)$, and ψ is a one-to-one transformation of θ, then the transformed prior, $\pi(\psi)$ will also be that specific prior, Datta and Ghosh (1996). This is a well-known weak prior that is frequently applied in situations when the unknown parameters are poorly understood.

2.4.2 The Divergence Prior

The divergence prior was developed by Ghosh et al. (2011), it maximizes the distance between the prior and posterior by using a chi-square divergence distance metric. This prior has been successfully applied to a vector field of orientated unit tangents for vessel pathlines as a regularisation constraint, Zhang et al. (2019). The divergence prior is proportional to the fourth root of the determinant of the Fisher information matrix. The divergence prior is given by

$$\pi_D(\theta) \propto | H(\theta) |^{\frac{1}{4}} . \tag{2.7}$$

2.4.3 The Reference Prior

Bernardo (1979) proposed the reference prior, which was then further developed by Berger and Bernardo (1989). Finding appropriate, non-informative priors in multiparameter issues is challenging since typical non-informative priors, like Jeffreys prior, might have characteristics that unexpectedly drastically affect the posterior distribution, Harvey and Van der Merwe (2018). In recognition of this problem Bernardo (1979) proposed the reference prior, and further developed by Berger and Bernardo (1989). The reference prior method is also derived from the Fisher information matrix. The reference prior is the prior that, according to the experimental data, maximizes the predicted Kullback-Leibler distance between the posterior distribution and the prior distribution. It is invariant under a one-to-one translation in the parameters, Ramos et al. (2017). The algorithm of Berger and Bernardo (1992) may then be described as follows

The parametric statistical problem in which the random observation T which has a density $f(t|\theta)$, where $\theta \in \Theta \subset R^K$ is the unknown parameter is considered. It is assumed that the Fisher information matrix

$$H(\theta) = -E_{t|\theta}\left[\left\{\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log f(t|\theta)\right\}\right]$$

exists and has rank k, so that

$$S(\theta) = H^{-1}(\theta),$$

also exists. Often, we will just write H and S.

It is assumed that the θ_i are separated into m groups of sizes $n_1,...n_m$

$$\theta_{(1)} = (\theta_1, ...\theta_{n_1}), \theta_{(2)} = (\theta_{n_1+1}, ...\theta_{n_1+n_2}), ...,$$
$$\theta_{(1)} = \left(\theta_{N_{i-1}+1}, ...\theta_{n_1}\right), ..., \theta_{(m)} = \left(\theta_{N_{m-1}+1}, ...\theta_k\right),$$

where $N_j = n_1 + ... + n_j$ for $j = 1, ...m$.

Then S is denoted as

$$S = \begin{bmatrix} A_{11} & A_{21}^{\intercal} & . & . & . & A_{m1}^{\intercal} \\ A_{21} & A_{22} & . & . & . & A_{m2} \\ . & . & & & & . \\ . & . & & & & . \\ . & . & & & & . \\ A_{m1} & A_{m2} & . & . & . & A_{mm} \end{bmatrix}$$

Note that $h_1 \equiv H_1 \equiv A_{11}^{-1}$ and $B_j = \left(A_{j1}....A_{j(j-1)}\right)$, for $j = 2, ..., m$, iterative expressions for computing these quantities, in general, are

$$h_j = \left(A - B_j H_{j-1} B_j^{\intercal}\right)^{-1}$$
$$H_j = \begin{pmatrix} H_{j-1} + H_{j-1} B_j^{\intercal} h_j B_j H_{j-1} & -H_{j-1} B_j^{\intercal} h_j \\ -h_j B_j H_{j-1} & h_j \end{pmatrix}.$$

In the improper case we proceed by specifying a nested sequence $\Theta^1 \subset \Theta^2 \subset ...$ of compact subsets of Θ such that

$$\bigcup_{l=1}^{\infty} \Theta^l = \Theta,$$

where Θ represents the parameter space.

Define

$$\pi_m^l \left(\theta_{[\sim(m-1)]} | \theta_{[m-1]}\right) = \pi_m^l \left(\theta_{(m)} | \theta_{[m-1]}\right)$$
$$= \frac{| h_m^{\frac{1}{2}}(\theta) | \, 1_{\Theta^l\left(\theta_{[m-1]}\right)}^l \left(\theta_{(m)}\right)}{\int_{\Theta^l\left(\theta_{[m-1]}\right)} | h_m(\theta) |^{\frac{1}{2}} \, d\theta_{(m)}}.$$

For $j = m-1, m-2, ..., 1$, define

$$\pi_j^l\left(\theta_{[\sim(j-1)]}|\theta_{[j-1]}\right)=\frac{\pi_{j+1}^l\left(\theta_{[\sim j]}|\theta_{[j]}\right)\exp\left\{\frac{1}{2}E_j^l\left[\log\left(|\,|\,h_j(\theta)\,|\right)|\theta_{[j]}\right]\right\}1_{\Theta^l\left(\theta_{[j-1]}\right)}\left(\theta_{(j)}\right)}{\int_{\Theta^l\left(\theta_{[j-1]}\right)}\exp\left\{\frac{1}{2}E_j^l\left[\log\left(|\,h_j(\theta)\,|\right)|\theta_{[j]}\right]\right\}d\theta_{(j)}},$$

where $E_j^l\left[g(\theta)\,|\theta_{[j]}\right]=\int g(\theta)\,\pi_{j+1}^l\left(\theta_{[\sim j]}|\theta_{[j]}\right)d\theta_{[j]}$, where the integral is over the range $\left\{\theta\left[\sim j\right]:\theta_{[j]},\theta_{(\sim j)}\in\Theta^l\right\}$.

Define the m-group reference prior by

$$\pi_R(\theta)=\lim_{l\to\infty}\frac{\pi^l(\theta)}{\pi^l(\theta^*)},$$

where θ^* is some point in Θ^1.

2.4.4 The Probability Matching Prior

The probability matching prior was firstly proposed by Welch and Peers (1963). A probability matching prior is a prior distribution where the posterior probabilities of certain regions perfectly or substantially overlap with their coverage probabilities, Datta and Sweeting (2005). By using such a prior, Bayesian credible areas will be guaranteed to have exact or approximate frequentist coverage. Datta and Ghosh (1995) derived a differential equation which a prior must satisfy if the posterior probability of a one sided credibility interval for a parametric function and its frequentist probability agree up to $O\left(n^{-1}\right)$ where n is the sample size. They proved that the agreement between the posterior probability and the frequentist probability hold if and only if $\sum_{i=1}^{k}\frac{\partial}{\partial p_i}\left\{\eta_i\left(\underline{\theta}\right)\pi\left(\underline{\theta}\right)\right\}=0$. Reasons for using the probability matching prior is that it provides a methods of constructing accurate frequestist intervals and it could also be useful for comparative purposes in Bayesian analysis.

It is well known that the following method can be used to derive the probability matching prior:

1. Determine the inverse of the Fisher information matrix.
2. Suppose the goal is to find a probability matching prior for a function $t(\theta)$, then determine

$$\nabla_t^{\mathsf{T}}(\theta)=\left[\frac{\partial t(\theta)}{\partial\theta_1},\frac{\partial t(\theta)}{\partial\theta_2},\ldots\frac{\partial t(\theta)}{\partial\theta_k}\right] \tag{2.8}$$

and

$$\nabla_t(\theta)=\left[\frac{\partial t(\theta)}{\partial\theta_1},\frac{\partial t(\theta)}{\partial\theta_2},\ldots\frac{\partial t(\theta)}{\partial\theta_k}\right]^{\mathsf{T}}.$$

3. Define

$$\eta^{\intercal}(\theta) = \frac{\nabla_t^{\intercal}(\theta) H^{-1}(\theta)}{\sqrt{\nabla_t^{\intercal}(\theta) H^{-1}(\theta) \nabla_t(\theta)}}.$$

4. The prior $\pi(\theta)$ is a probability matching prior if and only if the differential equation $\sum_{i=1}^{k} \frac{\partial}{\partial\theta_i} \{\eta_i(\theta)\pi(\theta)\} = 0$ is satisfied.

2.5 Bayesian Application in Reliability Theory

Bayesian inference has been successfully applied in reliability analysis, where it facilitates the modeling and prediction of the lifetime or performance of systems and components, Ruggeri et al. (2008). Hamada et al. (2008) mentioned that Bayesian reliability presents modern methods and techniques for analysing reliability data from a Bayesian perspective. The authors mentioned that the adoption and application of Bayesian methods in virtually all branches of science and engineering have significantly increased over the past few years. This increase is largely due to advances in simulation-based computational tools for implementing Bayesian methods. Bayesian reliability methods permit the formal incorporation of pertinent supplementary information about the parameters of interest in a statistical analysis beyond that contained in the sample data, Martz (2014). According to Al-Bossly (2020), the development of Monte Carlo Markov Chain methods, lead to huge improvements in computational capabilities which in turn lead to increased use of Bayesian methods in reliability applications. Bayesian inference in reliability theory has several advantages, including the ability to incorporate prior information, account for uncertainty, and make probabilistic statements about unknown parameters.

2.6 MCMC Methods

The posterior distribution of many complex Bayesian models is intractable, meaning it cannot be expressed in closed form Smit (2021). As a result, MCMC techniques, which are computational algorithms that can produce samples from complex posterior distributions have been developed, Hamada et al. (2008). A Markov chain is a sequence of random variables $T^{(1)}, T^{(2)}, \ldots,$ where for some τ, the next state $T^{(\tau+1)}$only depends on the current $T^{(\tau)}$, and not on all the previous states $T^{(1)}, T^{(2)}, \ldots, T^{(\tau-1)}$. The Markov property can be expressed as

$$P\left(T^{(\tau+1)} = t \mid T^{(1)} = t^{(1)}, T^{(2)} = t^{(2)}, \ldots, T^{(\tau)} = t^{(\tau)}\right) = P\left(T^{(\tau+1)} = t \mid T^{(\tau)} = t^{(\tau)}\right).$$

For some arbitrary starting value $T^{(0)}$, a Markov chain $T^{(\tau)}$can be generated using a transition distribution, that has stationary distribution f, to guarantee that $\left\{T^{(\tau)}\right\}$ converges in distribution to a random variable from f, Robert et al. (1999). The MCMC methods discussed in this thesis include the MH sampler, and Gibbs-within-MH sampler.

2.6.1 Metropolis-Hastings (MH) Sampler

The MH algorithm is one of the most fundamental and widely used MCMC methods in Bayesian computation. It was originally proposed by Metropolis et al. (1953) to simulate the states of a physical system in statistical mechanics using a symmetric proposal distribution. Hastings (1970) expanded the method's use in Bayesian statistics and beyond by generalising it to accommodate asymmetric proposal distributions. The MH algorithm is used to generate samples from a complex target distribution $\pi(\theta|t)$, typically a posterior distribution, when direct sampling is infeasible.

Here is how the Metropolis-Hastings algorithm is used

1. Choose a starting value for which $\pi\left(\theta^{(0)}|\underline{t}\right) > 0$, represented by $\theta^{(0)}$.
2. In the sequence $\theta^{(\tau-1)}$, create a candidate point (or proposal) θ^* from the preceding state using a jumping distribution (or proposal distribution), which is a conditional density represented by $q\left(\theta^*|\theta^{(\tau-1)}\right)$ at time τ. The probability of jumping from θ^* back to $\theta^{(\tau-1)}$ is denoted by $q\left(\theta^{(\tau-1)}|\theta^*\right)$. The proposal distribution the proposed distribution must meet specific requirements, such as the resulting Markov chain being irreducible and a periodic, and it should be simple to simulate from.
3. Calculate the probability that the candidate point will be accepted as the next state in the sequence, known as the acceptance probability, denoted by ρ, and defined as

$$\rho = min\left(1, \frac{\pi(\theta^*|\underline{t})\, q\left(\theta^{(\tau-1)}|\theta^*\right)}{\pi\left(\theta^{(\tau-1)}|\underline{t}\right) q\left(\theta^*|\theta^{(\tau-1)}\right)}\right).$$

4. From the distribution $U(0,1)$, simulate a value u. As the subsequent state in the sequence, the candidate point is approved if $u \leq \rho$. If $u > \rho$, the candidate point is rejected and the sequence's state stays the same.

$$\theta^{(\tau)} = \begin{cases} \theta^* & \text{with probability } \rho \\ \theta^{(\tau-1)} & \text{with probability } 1-\rho \end{cases}$$

5. Set $\tau = \tau + 1$ and go back to Step 2 to proceed to the following state in the sequence. Repeat a significant number of times, say N, to produce the Markov chain $\theta^{(1)}, \theta^{(2)}, \ldots, \theta^{(N)}$.

2.6.2 Gibbs Sampler

The Gibbs sampler is one of the most widely used MCMC algorithms, and it was initially presented by Geman and Geman (1984) in the context of image processing. The Gibbs sampler creates a Markov chain by separating the parameter vector into subvectors, and then sampling each subvector conditional only on the current values of all the other subvectors. Let $\pi(\theta \mid \underline{t})$ be the posterior distribution with $d-$dimensional parameter vector $\theta = \{\theta_1, \theta_2, .., \theta_d\}$, and denote the full conditional posterior distributions by $\pi(\theta_\tau \mid \underline{t}, \theta_1, \theta_2, .., \theta_{i-1}, \theta_{t+1}, ..., \theta_d)$. Both the posterior and full conditional posterior distributions need only be defined at least up to proportionality. The Gibbs sampler is performed as follows

1. Select arbitrary starting values $\theta^{(0)} = \left\{\theta_1^{(0)}, \theta_2^{(0)}, ..., \theta_d^{(0)}\right\}$ for each parameter.

2. Sample new values for each parameter from the corresponding full conditional posterior distribution. The full conditional distribution are updated for each new parameter value sampled. This is done in succession as

$$
\begin{array}{rcl}
\text{Sample } \theta_1^{(1)} & \text{from} & \pi\left(\theta_1 \mid \underline{t}, \theta_2^{(0)}, \theta_3^{(0)}, ..., \theta_d^{(0)}\right) \\
\text{Sample } \theta_2^{(1)} & \text{from} & \pi\left(\theta_2 \mid \underline{t}, \theta_1^{(0)}, \theta_3^{(0)}, ..., \theta_d^{(0)}\right) \\
\cdots & \cdots & \cdots \\
\text{Sample } \theta_i^{(1)} & \text{from} & \pi\left(\theta_i \mid \underline{t}, \theta_1, \theta_2, .., \theta_{i-1}, \theta_{i+1}, ..., \theta_d\right) \\
\cdots & \cdots & \cdots \\
\text{Sample } \theta_d^{(1)} & \text{from} & \pi\left(\theta_d \mid \underline{t}, \theta_1^{(0)}, \theta_2^{(0)}, ..., \theta_{d-1}^{(0)}\right).
\end{array}
$$

   After a new value is sampled for each parameter to find $\theta_1^{(1)} = \left\{\theta_1^{(1)}, \theta_2^{(1)}, ..., \theta_d^{(1)}\right\}$, one iteration of the Gibbs sampler has been completed.

3. Repeat Step 2 a large number of times, say N times, conditioning only on the most recent values of all other parameters, to obtain a Markov chain of simulated parameter values $\theta^{(1)}, \theta^{(2)}, ..., \theta^{(N)}$.

2.7 MCMC Convergence

MCMC diagnostic techniques are required to determine whether Markov chains will converge to a stationary distribution, Roy (2020). Smit (2021) state that Markov chain was started to produce a sufficient number of samples, and then the initial subset of the samples is eliminated referred to as the burn-in. Therefore, determining a Markov chain's convergence is crucial for determining the chain's suitable burn-in. This section outlines a few MCMC diagnostics that can be used to determine whether Markov chain convergence or MCMC sampling is stopped. Convergence of the Markov Chain Monte

Carlo algorithm was assessed primarily through visual inspection of trace plots. Trace plots display the sampled values of each parameter across iterations and are a standard diagnostic tool to evaluate whether the chain has reached a stationary distribution. For each simulation run, trace plots of the posterior samples of the scale and shape parameter were examined. A well mixed trace plot is characterized by rapid oscillation and absence of apparent trends or drifts, is taken as evidence of convergence. The chains that exhibits poor mixing were discarded and re-run with adjusted tuning parameters or starting values. Additional diagnostics such as the Gelman-Rubin convergence diagnostic and the Geweke diagnostic were used to supplement visual checks where necessary. Autocorrelation plots were used to assess the degree of dependence in the samples. Rapid decay of autocorrelations to zero was considered a positive indication of good mixing. The Gelman-Rubin diagnostic Gelman and Rubin (1992) assesses convergence by analyzing two or more parallel MCMC chains. When a shrinkage factor, denoted by $\hat{R}$ is close to 1, it suggests that the chains have converged. Geweke (1992) suggested a diagnostic test for checking the convergence of the mean of each parameter separately from the sampled values of a single chain. This diagnostic applies a simple Z test to check whether the means estimated from two different subsamples of the total MCMC output are equal. The parameter with $||Z|| \leq 2$ suggest no significant difference between the means of the initial and final segments of the chain, indicating convergence.

2.7.1 Trace plot

Trace plots provide a simple visual method for evaluating a Markov chain's convergence. Trace plots are produced for each parameter separately and evaluate the chain, but it is also useful to monitor the Markov chains jointly, either the total parameter vector, Lesaffre and Lawson (2012). To check convergence of a chain, Gelfand and Smith (1990) proposed the thick pen test, if the trace plot can be covered by a thick pen then the chain has converged. Figure 2.1 shows a trace plot that passes the thick pen test, indicating good mixing and convergence. Figure 2.2 displays a poorly mixing chain that fails the thick pen test. This lack of convergence implies that the posterior was not fully explored.

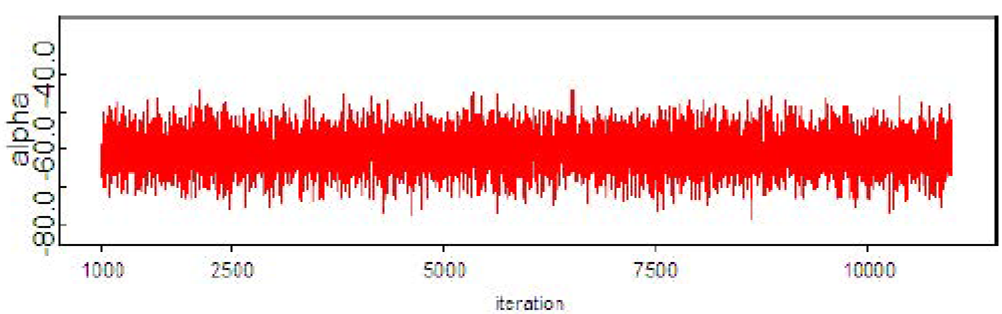

Figure 2.1: Trace plot that converges example (Image from Spiegelhalter et al. (2014).

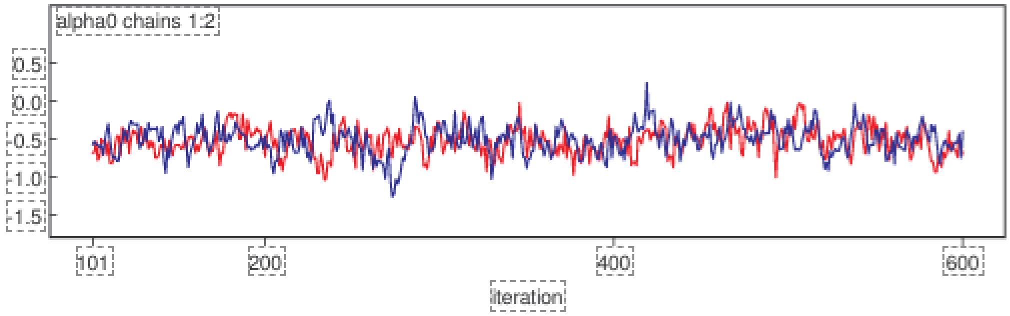

Figure 2.2: Trace plot that does not converges example (Image from Spiegelhalter et al. (2014)).

2.7.2 Autocorrelation

The autocorrelation function describes how values in a sequence are related at a specific lag m, and its estimates, denoted by $\hat{\rho}_m$, are graphically displayed using an autocorrelation plot, Lesaffre and Lawson (2012). When the autocorrelation decreases only slowly with increasing lag, the mixing rate is low. When autocorrection is close to zero, then MCMC sampling is done in an independent manner.

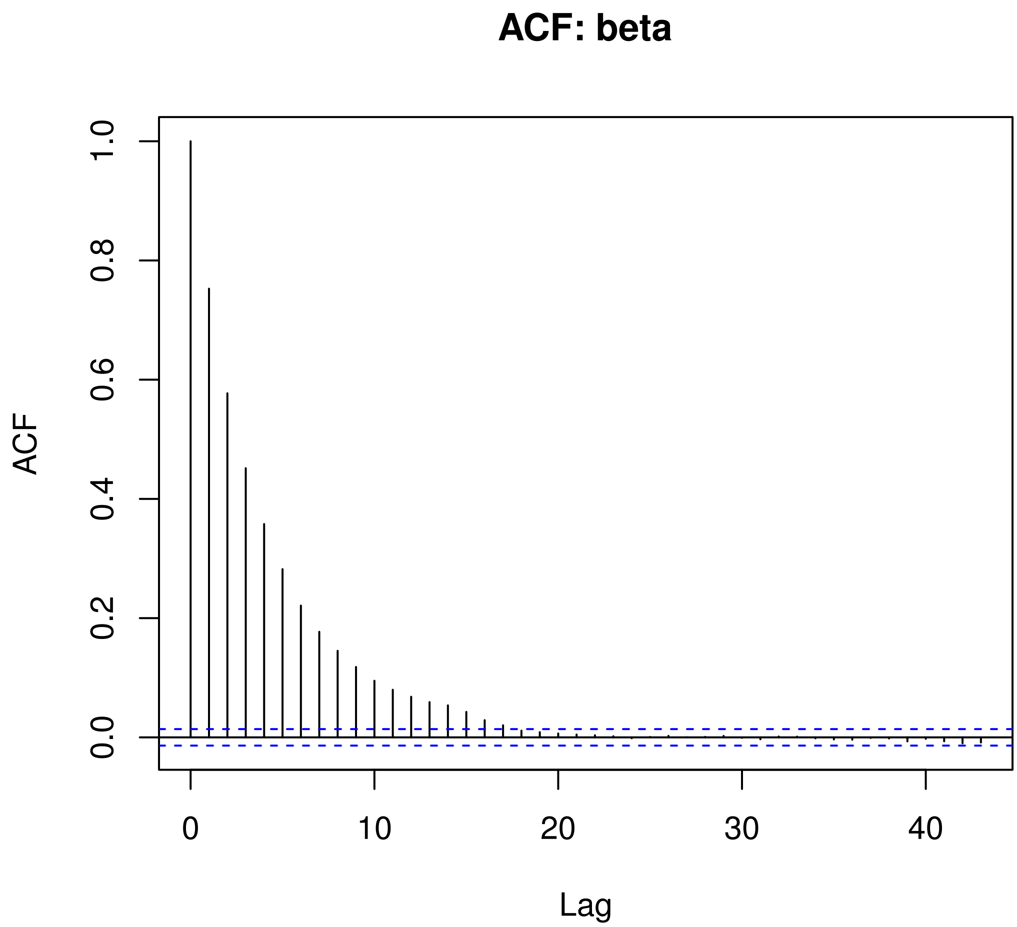

Figure 2.3: Example of autocorrelation when independency is high.

2.7.3 Gelman-Rubin Diagnostic

The Gelman-Rubin diagnostic Gelman and Rubin (1992) uses two or more samples created in parallel to verify if the chain has converged. This test is an ANOVA-type diagnostic, calculating a shrinking factor $\hat{R}$. A value around one signifies convergence. Figure 2.4 presents a Gelman-Rubin plot where clear convergence is achieved around 6500 iterations. The potential shrinking factor approaches 1, indicating alignment between chains.

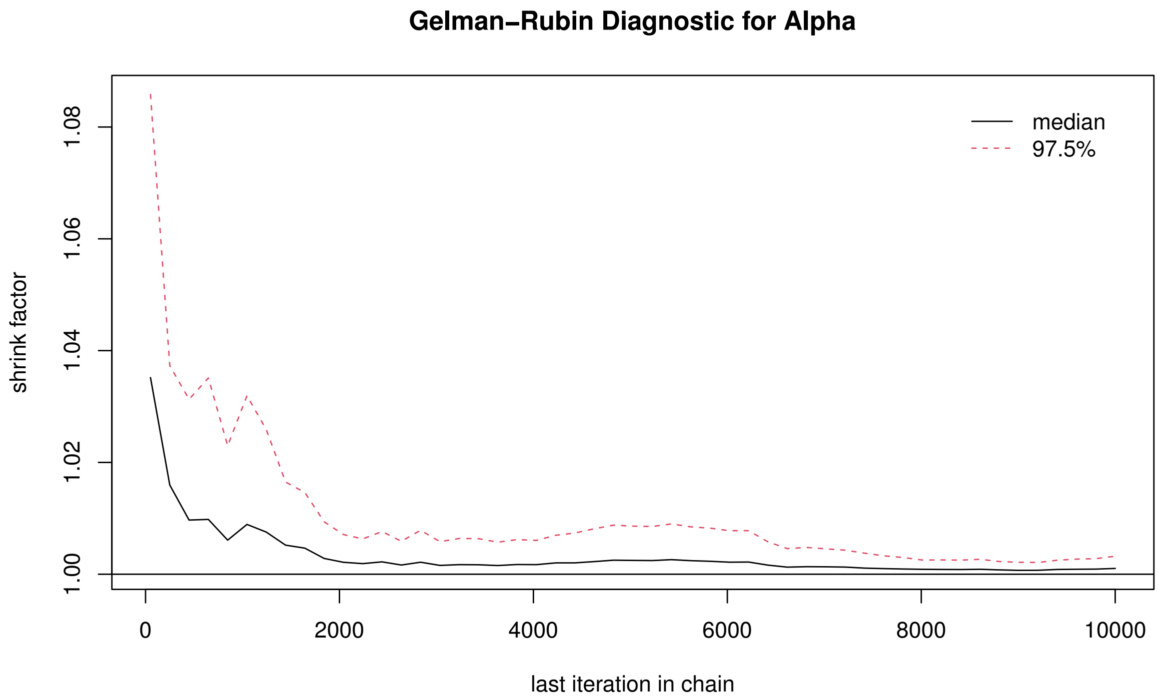

Figure 2.4: Example of Gelman-Rubin plot, where convergence occurs around 6500 iterations.

2.7.4 Geweke Diagnostic

A diagnostic test was proposed by Geweke (1992) to verify that the mean of each parameter, independently from the sampled values of a single chain converges. He suggested using MCMC output to see the set of simulated data as a time series in order to construct this test. To determine whether the means estimated from two distinct subsamples of the overall MCMC output are equal, this diagnostic uses a straightforward two-test. The observations from the start and finish of the created chain are referenced in these subsamples. In Figure 2.5, most Z-scores lie within the standard normal bound which is $+2$ and -2. This implies that the chain has reached its stationary distribution.

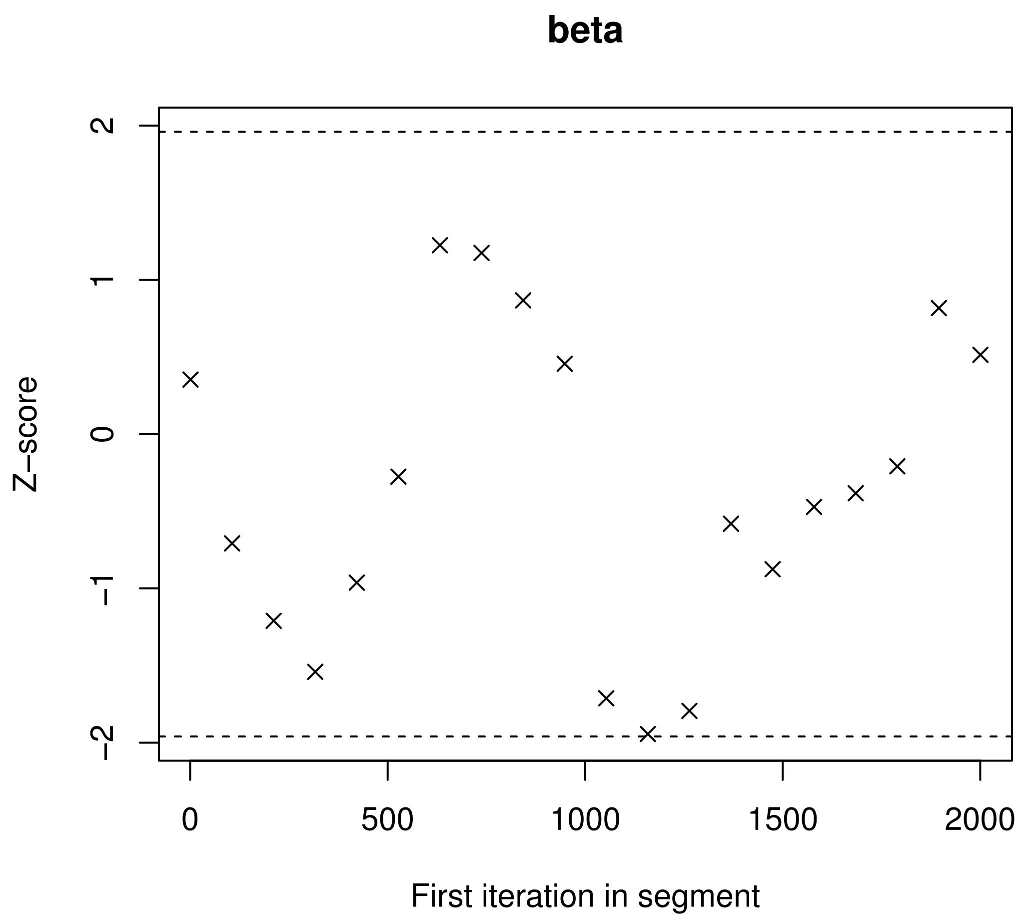

Figure 2.5: Example of Geweke plot that illustrate convergence.

2.8 Life Distributions in Reliability Analysis

A life distribution, according to Harris and Singpurwalla (1968), is an effort to represent the lifespan of a component. Life distributions constitute a group of probability distributions utilized in reliability engineering and survival analysis, Smit (2021). In this thesis, some of the most commonly used life distributions, namely the Weibull distribution and Birnbaum-Saunders distributions, will be considered and discussed. The Weibull distribution will be discussed in Chapter 3. The Birnbaum-Saunders distribution will be discussed in Chapter 4.

Chapter 3

The Weibull Distribution

Weibull (1951) provided the first formal definition of the Weibull distribution. The Weibull distribution is one of the most widely used and studied life distributions in reliability theory, Smit (2021). It is widely employed in the analysis of lifetime or survival data, as well as in characterizing phenomena that exhibit monotonic failure rates, Carrasco et al. (2008). This distribution is known to have special cases, such as the exponential and the Rayleigh distribution. For $\beta = 1$, the Weibull distribution is well-known to be the same as the exponential distribution with a constant failure rate. For $\beta = 2$, the Weibull distribution is well-known to be the same as the Rayleigh distribution with a monotonically increasing failure rate. If the shape parameter β is larger than one, the failure rate function is increasing; if the shape parameter is smaller than one, the failure rate function is decreasing, Li et al. (2011). Akram and Hayat (2014) state that this distribution has been proven to be a successful model for many product failure mechanisms because it is a flexible distribution.

A random variable that is continuous T is said to have a Weibull distribution if its PDF is defined by

$$f(t|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\},\ t > 0, \tag{3.1}$$

where β shape parameter and α scale parameter , $\beta > 0$, $\alpha > 0$. In the context of this model, the parameter vector $\theta = (\alpha, \beta)$.

The Weibull CDF is given as

$$F(t|\alpha,\beta) = 1 - \exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\},\ t > 0, \tag{3.2}$$

and the corresponding reliability function is given by

$$R(t|\alpha,\beta) = \exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\}. \tag{3.3}$$

The likelihood for the Weibull distribution with no censoring is given as

$$L(\alpha,\beta|\underline{t}) = \prod_{i=1}^{n} \frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}.$$

The likelihood for the Weibull distribution with respect to type I right censoring is given as

$$L(\alpha,\beta|\underline{t}) = \prod_{i=1}^{n} \left\{\frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\right\}^{\delta_i} \left\{\exp\left\{-\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\right\}^{1-\delta_i},$$

where, for observation i, $data_i = (t_i, \delta_i)$, t_i is either a failure time or a right censored time, $\delta_i = 1$ for an exact failure time, and $\delta_i = 0$ for a right censored observation.

The Fisher information matrix for the Weibull distribution is given by

$$H(\alpha,\beta) = \begin{bmatrix} \frac{n\beta^2}{\alpha^2} & -\frac{n(1+\gamma_1)}{\alpha} \\ -\frac{n(1+\gamma_1)}{\alpha} & \frac{n(1+2\gamma_1+\gamma_2)}{\beta^2} \end{bmatrix}, \tag{3.4}$$

where $\gamma_1 = \int_0^\infty log(y)\exp(-y)\,dy$ and $\gamma_2 = \int_0^\infty log^2(y)\exp(-y)\,dy.$

The Fisher information matrix for the Weibull distribution was also derived by Sun (1997) and the detailed derivations are provided in Appendix A.1.

Figure 3.1 shows possible shapes for the PDF of the Weibull distribution for various values of the parameters α and β, and the corresponding reliability functions are shown in Figure 3.2.

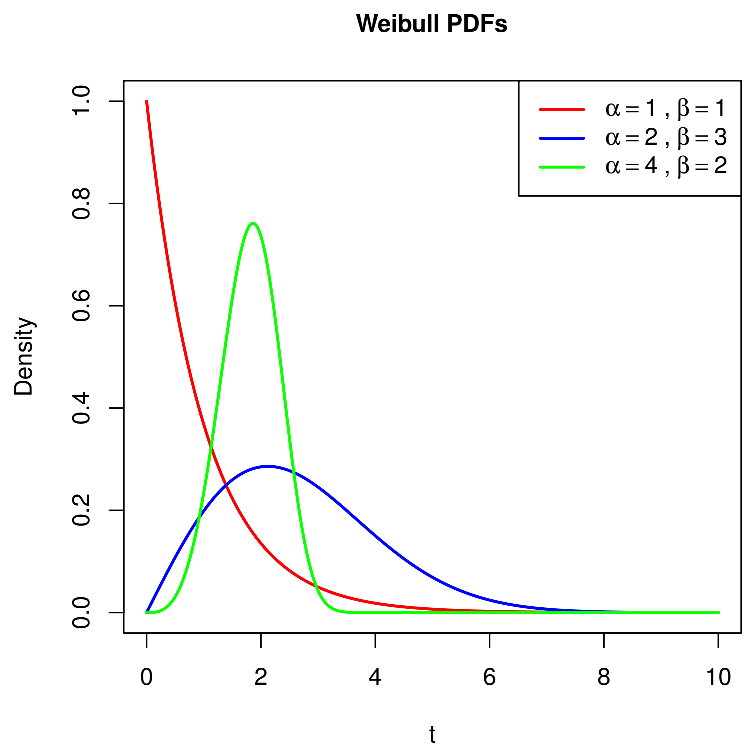

Figure 3.1: Probability density function of the Weibull distribution for various values of α and β.

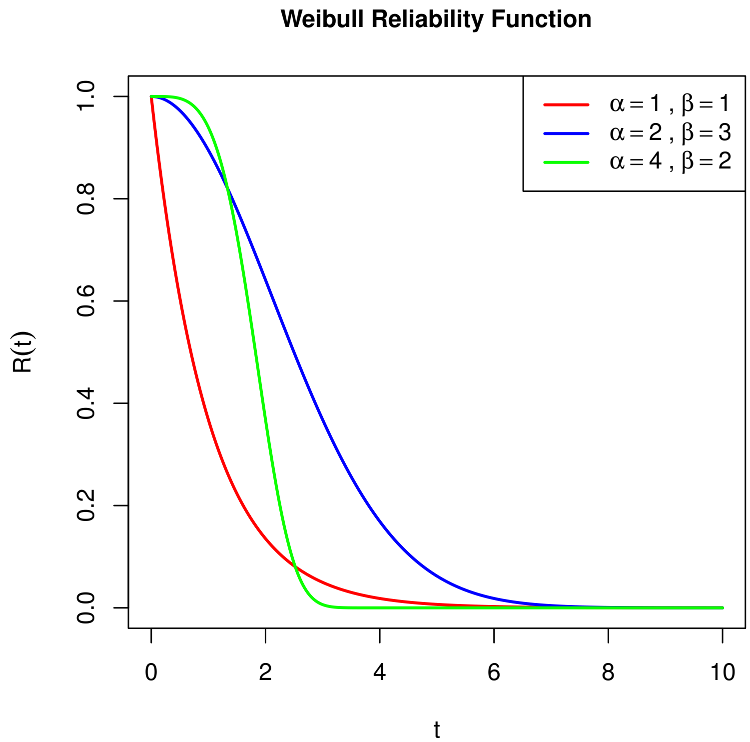

Figure 3.2: Reliability function of the Weibull distribution for various values of α and β.

3.1 The Posterior for the Weibull Distribution using the Jeffreys Prior

In this section, the posterior distribution for the Weibull distribution using the Jeffreys prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined.

Using the definition of Jeffreys prior in equation 2.6, the Jeffreys prior for Weibull distribution is given by

$$\pi_J(\alpha,\beta) \propto \frac{1}{\alpha}. \tag{3.5}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the Jeffreys prior when there is no censoring is given by

$$\pi_J(\alpha,\beta|\underline{t}) \propto \beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\}. \tag{3.6}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the Jeffreys prior when there is no censoring is proper, Sun (1997).

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_J(\alpha|\beta,\underline{t}) \propto \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}, \tag{3.7}$$

letting $u=\alpha^{\beta}$ in the previous expression, then $u=\alpha^{\beta} \sim Inv-G\left(n, \sum_{i=1}^{n} t_i^{\beta}\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_J(\beta|\alpha,\underline{t}) \propto \beta^{n}\alpha^{-(n\beta)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\}. \tag{3.8}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the Jeffreys prior under type I right censoring is given by

$$\pi_J(\alpha,\beta|\underline{t}) \propto \beta^{r}\alpha^{-r\beta-1}\left\{\prod_{i=1}^{r} t_i^{\beta-1}\right\} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}. \tag{3.9}$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_J(\alpha|\beta,\underline{t}) \propto \alpha^{-r\beta-1} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}, \tag{3.10}$$

letting $u=\alpha^{\beta}$ in the previous expression, then $u=\alpha^{\beta} \sim Inv-G\left(r, \sum_{i=1}^{n} t_i^{\beta}\right)$. The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_J(\beta|\alpha,\underline{t}) \propto \beta^{r}\alpha^{-r\beta}\left\{\prod_{i=1}^{r} t_i^{\beta-1}\right\} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}. \tag{3.11}$$

The posterior of α and β for the Weibull distribution given the observed data using the Jeffreys prior under type I right censoring is proper, it was shown by Southey (2015).

The conditional posterior of α given β and $\underline{t}$ has a known form, and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. Instead, a Gibbs sampler with a MH step (Metropolis-within-Gibbs) can be employed. This occurs both in the complete data case and under type I right censoring.

3.2 The Posterior for the Weibull Distribution using the Divergence Prior

In this section, the posterior distribution for the Weibull distribution using divergence prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined.

Using the definition of the divergence prior in equation 2.7, the divergence prior for Weibull distribution is given by

$$\pi_D(\alpha,\beta) \propto \frac{1}{\alpha^{\frac{1}{2}}}. \tag{3.12}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the divergence prior when there is no censoring is given by

$$\pi_D(\alpha,\beta|\underline{t}) \propto \beta^n \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\}. \tag{3.13}$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_J(\alpha|\beta,\underline{t}) \propto \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}, \tag{3.14}$$

letting $u=\alpha^{\beta}$ in the previous expression, then $u=\alpha^{\beta} \sim Inv-G\left(n-\frac{1}{2\beta}, \sum_{i-1}^{n} t_i^{\beta}\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_D(\beta|\alpha,\underline{t}) \propto \beta^n \alpha^{-(n\beta)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\}. \tag{3.15}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the divergence prior under type I right censoring is given by

$$\pi_D(\alpha,\beta|\underline{t}) \propto \beta^r \alpha^{-r\beta-\frac{1}{2}} \left\{\prod_{i=1}^{r} t_i^{\beta-1}\right\} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}. \tag{3.16}$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_D(\alpha|\beta,\underline{t}) \propto \alpha^{-(r\beta+1)} \exp\left\{-\frac{\sum_{i-1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\},$$

letting $u = \alpha^{\beta}$ in the previous expression, then $u = \alpha^{\beta} \sim Inv - G\left(r - \frac{1}{2\beta}, \sum_{i-1}^{n} t_i^{\beta}\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_D(\beta|\alpha,\underline{t}) \propto \beta^r \alpha^{-r\beta} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}.$$

The conditional posterior of α given β and $\underline{t}$ has a known form, and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. Instead, a Gibbs sampler with a MH step (Metropolis-within-Gibbs) can be employed. This occurs both in the complete data case and under type I right censoring.

Theorem 3.1. *The joint posterior of α and β for the Weibull distribution given the observed data using the divergence prior when there is no censoring is a proper distribution.*

Proof. For the posterior to be proper, this must hold

$$\int_0^{\infty}\int_0^{\infty} c\beta^n \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\} d\alpha d\beta = 1,$$

where c is the normalising constant.

For the above to be true, it must be shown that

$$\int_0^{\infty}\int_0^{\infty} \beta^n \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\} d\alpha d\beta < \infty.$$

Now

$$\int_0^{\infty}\int_0^{\infty} \beta^n \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\} d\alpha d\beta = \int_0^{\infty} \beta^n \left\{\prod_{i=1}^{n} t_i^{\beta}\right\} \int_0^{\infty} \alpha^{-\left(n\beta+\frac{1}{2}\right)} \times \exp\left\{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\} d\alpha d\beta.$$

Consider

$$\int_0^{\infty} \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\} d\alpha = \int_0^{\infty} \left(\alpha^{\beta}\right)^{-n} \alpha^{-\frac{1}{2}} \exp\left\{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\} d\alpha,$$

and let $u = \alpha^{\beta}$, then $du = \beta\alpha^{\beta-1} d\alpha$ and $d\alpha = \frac{\alpha du}{\beta u}$.

Now

$$\int_0^\infty \left(\alpha^\beta\right)^{-n} \alpha^{-\frac{1}{2}} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha = \int_0^\infty (u)^{-n} \alpha^{-\frac{1}{2}} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} \frac{\alpha du}{\beta u}$$

$$= \frac{1}{\beta}\int_0^\infty (u)^{-n-1} \alpha^{\frac{1}{2}} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du$$

$$= \frac{1}{\beta}\int_0^\infty u^{-n-1} u^{\frac{1}{2\beta}} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du$$

$$= \frac{1}{\beta}\int_0^\infty u^{-\left(n-\frac{1}{2\beta}+1\right)} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du.$$

The expression above follows Inverse Gamma with parameters $n-\frac{1}{2\beta}$ and $\sum_{i=1}^n t_i^\beta$, then this follows

$$\frac{1}{\beta}\int_0^\infty u^{-\left(n-\frac{1}{2\beta}+1\right)} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du = \frac{1}{\beta}\Gamma\left(n-\frac{1}{2\beta}\right)\left(\sum_{i=1}^n t_i^\beta\right)^{-\left(n-\frac{1}{2\beta}\right)}$$

Then

$$\int_0^\infty \int_0^\infty \beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^n \left\{\prod_{i=1}^n t_i^\beta\right\} \beta^{-1}\Gamma\left(n-\frac{1}{2\beta}\right)\left(\sum_{i=1}^n t_i^\beta\right)^{-\left(n-\frac{1}{2\beta}\right)} d\beta$$

$$= \int_0^\infty \beta^{n-1}\Gamma\left(n-\frac{1}{2\beta}\right) \frac{\left\{\prod_{i=1}^n t_i^\beta\right\}}{\left(\sum_{i=1}^n t_i^\beta\right)^{n-\frac{1}{2\beta}}} d\beta.$$

Note that for large β, $\Gamma\left(n-\frac{1}{2\beta}\right) \to \Gamma(n) < \infty$, then it must be shown that

$$\int_0^\infty \beta^{n-1} \frac{\left\{\prod_{i=1}^n t_i^\beta\right\}}{\left(\sum_{i=1}^n t_i^\beta\right)^{n-\frac{1}{2\beta}}} d\beta < \infty.$$

Let $t_k < \max(T_1, T_2, ..., T_n) = t_{max}$

$$\int_0^\infty \beta^{n-1} \frac{\left(\prod_{i=1}^n t_i^\beta\right)}{\left(\sum_{i=1}^n t_i^\beta\right)^{n-\frac{1}{2\beta}}} d\beta \leq \int_0^\infty \beta^{n-1} \frac{\left(\prod_{i=1}^n t_i^{n\beta}\right)}{\left(\sum_{i=1}^n t_i^\beta\right)^{n-\frac{1}{2\beta}}} d\beta$$
$$\leq \int_0^\infty \beta^{n-1} \frac{\left(t_k^{n\beta}\right)}{\left(t_{max}^\beta\right)^{n-\frac{1}{2\beta}}} d\beta$$
$$= \int_0^\infty \beta^{n-1} \frac{\left(t_k^{n\beta}\right)}{(t_{max})^{n\beta}\, t_{max}^{-\frac{1}{2}}} d\beta$$
$$\propto \int_0^\infty \beta^{n-1} \frac{\left(t_k^{n\beta}\right)}{(t_{max})^{n\beta}} d\beta$$
$$= \int_0^\infty \beta^{n-1} \frac{(t_k)^{n\beta}}{(t_{max})^{n\beta}} d\beta$$
$$= \int_0^\infty \beta^{n-1} \exp\left\{\beta n \log\left(\frac{t_k}{t_{max}}\right)\right\} d\beta.$$

Since $t_k < t_{max} \iff \frac{t_k}{t_{max}} < 1$ then $\log\left(\frac{t_k}{t_{max}}\right) < 0$, then the expression above follows Gamma with parameters n and $-n\log\left(\frac{t_k}{t_{max}}\right)$, it implies that

$$\int_0^\infty \beta^{n-1} \exp\left\{\beta n \log\left(\frac{t_k}{t_{max}}\right)\right\} d\beta < \infty,$$

as a result

$$\int_0^\infty \int_0^\infty \beta^n \alpha^{-\left(n\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta < \infty.$$

This completes the proof, and we can conclude that the posterior is proper. □

Theorem 3.2. *The joint posterior of α and β for the Weibull distribution given the observed data using the divergence prior under type I right censoring is a proper distribution.*

Proof. To show properness of $\pi_D(\alpha, \beta | \underline{t})$ the following should be true

$$\int_0^\infty \int_0^\infty c\beta^r \alpha^{-\left(r\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta = 1,$$

where c is the normalising constant.

For the above to be true, the following must be shown

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-\left(r\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta < \infty.$$

Now

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-\left(r\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^r \left\{\prod_{i=1}^r t_i^\beta\right\} \int_0^\infty \alpha^{-\left(r\beta+\frac{1}{2}\right)} \times \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha d\beta.$$

By letting $u = \alpha^\beta$ then $du = \beta\alpha^{\beta-1}d\alpha$, this follows

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-\left(r\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^{r-1} \left\{\prod_{i=1}^r t_i^\beta\right\} \int_0^\infty u^{-\left(r-\frac{1}{2\beta}+1\right)} \times \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du d\beta.$$

Consider the expression

$$\int_0^\infty u^{-\left(r-\frac{1}{2\beta}+1\right)} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du,$$

the expression above follows Inverse Gamma distribution with parameters $r-\frac{1}{2\beta}$ and $\sum_{i=1}^n t_i^\beta$. Then this follows

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-\left(r\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^{r-1} \left\{\prod_{i=1}^r t_i^\beta\right\} \Gamma\left(r-\frac{1}{2\beta}\right) \times \left\{\sum_{i=1}^n t_i^\beta\right\}^{-\left(r-\frac{1}{2\beta}\right)} d\beta.$$

For large β, $\Gamma\left(r-\frac{1}{2\beta}\right) \to \Gamma(r) < \infty$, then it remain to show that

$$\int_0^\infty \beta^{r-1}\left\{\prod_{i=1}^r t_i^\beta\right\}\left\{\sum_{i=1}^n t_i^\beta\right\}^{-\left(r-\frac{1}{2\beta}\right)} d\beta < \infty.$$

It is well known that

$$\int_0^\infty \beta^{r-1}\frac{\left\{\prod_{i=1}^r t_i^\beta\right\}}{\left\{\sum_{i=1}^n t_i^\beta\right\}^{\left(r-\frac{1}{2\beta}\right)}} d\beta \le \int_0^\infty \beta^{r-1}\frac{\left\{\prod_{i=1}^r t_i^{r\beta}\right\}}{\left\{\sum_{i=1}^n t_i^\beta\right\}^{\left(r-\frac{1}{2\beta}\right)}} d\beta.$$

If there is t_k such that $t_k < \max(T_1, T_2, ..., T_n) = t_{max}$, this follows

$$\int_0^\infty \beta^{r-1}\frac{\left\{\prod_{i=1}^r t_i^{r\beta}\right\}}{\left\{\sum_{i=1}^n t_i^\beta\right\}^{\left(r-\frac{1}{2\beta}\right)}} d\beta \le \int_0^\infty \beta^{r-1}\frac{\left\{t_k^{r\beta}\right\}}{\left\{t_{max}^\beta\right\}^{\left(r-\frac{1}{2\beta}\right)}} d\beta$$

$$= \int_0^\infty \beta^{r-1}\frac{t_k^{r\beta}}{\left(t_{max}^\beta\right)^{r-\frac{1}{2\beta}}} d\beta = \int_0^\infty \beta^{r-1}\frac{t_k^{r\beta}}{t_{max}^{r\beta}}\frac{1}{t_{max}^{-\frac{1}{2}}} d\beta,$$

since $\frac{1}{t_{max}^{-\frac{1}{2}}}$ is a constant, this is true

$$\int_0^\infty \beta^{r-1}\frac{t_k^{r\beta}}{t_{max}^{r\beta}}\frac{1}{t_{max}^{-\frac{1}{2}}} d\beta \propto \int_0^\infty \beta^{r-1}\left(\frac{t_k}{t_{max}}\right)^{r\beta} d\beta$$

$$= \int_0^\infty \beta^{r-1}\exp\left\{\beta r \log\left(\frac{t_k}{t_{max}}\right)\right\} d\beta.$$

The above expression follows Gamma distribution with parameters r and $-\log\left(\frac{t_k}{t_{max}}\right)$, since $t_k < t_{max} \Longleftrightarrow \log\left(\frac{t_k}{t_{max}}\right) < 0$, then,

$$\int_0^\infty \beta^{r-1}\left\{\prod_{i=1}^r t_i^\beta\right\}\left\{\sum_{i=1}^n t_i^\beta\right\}^{-\left(r-\frac{1}{2\beta}\right)} d\beta < \infty.$$

Which results in

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-\left(r\beta+\frac{1}{2}\right)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^{r} t_i^\beta\right\} d\alpha d\beta < \infty.$$

This completes the proof, and we can conclude that the posterior is proper. □

3.3 The Posterior for the Weibull Distribution using the Reference Prior

In this section, the posterior distribution for the Weibull distribution using the reference prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined. Sun (1997) derived the reference prior for the Weibull distribution using a different approach. In this work, I applied the one-at-a-time algorithm to derive the reference prior and obtained the same result. The reference prior is derived using the algorithm of Berger and Bernardo (1992).

The inverse of the Fisher information matrix is given by

$$\begin{aligned} H^{-1}(\alpha,\beta) &= \begin{bmatrix} \frac{\alpha^2(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)} & \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)} \\ \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)} & \frac{\beta^2}{n\left(\gamma_2-\gamma_1^2\right)} \end{bmatrix} \\ &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}. \end{aligned}$$

Thus,

$$\begin{aligned} h_1 =& A_{11}^{-1} \\ =& \frac{n\beta^2\left(\gamma_2-\gamma_1^2\right)}{\alpha^2\left(1+2\gamma_1+\gamma_2\right)}, \end{aligned}$$

and

$$\begin{aligned} h_2 &= \left(A_{22} - B_2 H_1 B_2^\intercal\right)^{-1} \\ &= \left[\frac{\beta^2}{n\left(\gamma_2-\gamma_1^2\right)} - \frac{\alpha\left(1+\gamma_1\right)}{n\left(\gamma_2-\gamma_1^2\right)} \frac{n\beta^2\left(\gamma_2-\gamma_1^2\right)}{\alpha^2\left(1+2\gamma_1+\gamma_2\right)} \frac{n\left(\gamma_2-\gamma_1^2\right)}{\alpha\left(1+\gamma_1\right)}\right]^{-1} \\ &= \frac{n\left(1+2\gamma_1+\gamma_2\right)}{\beta^2}. \end{aligned}$$

Let $\alpha \in (a,b)$ and $\beta \in (e,f)$, where a and b are the bounds of α parameter and e and f are the bounds of β parameter. The conditional reference prior for β given α is

$$\begin{aligned}\pi_2^l(\beta \mid \alpha) &= \frac{\mid h_2 \mid^{\frac{1}{2}}}{\int_e^f \mid h_2 \mid^{\frac{1}{2}} d\beta} \\ &= \frac{\mid \frac{n(1+2\gamma_1+\gamma_2)}{\beta^2} \mid^{\frac{1}{2}}}{\int_e^f \mid \frac{n(1+2\gamma_1+\gamma_2)}{\beta^2} \mid^{\frac{1}{2}} d\beta} \\ &= \frac{\beta^{-1}(n(1+2\gamma_1+\gamma_2))^{\frac{1}{2}}}{(n(1+2\gamma_1+\gamma_2))^{\frac{1}{2}} \int_e^f \frac{1}{\beta} d\beta} \\ &= \frac{\beta^{-1}}{\int_e^f \frac{1}{\beta} d\beta} = \frac{\beta^{-1}}{\log\left(\frac{f}{e}\right)}.\end{aligned}$$

The joint working prior is defined as

$$\pi^l(\alpha,\beta) = \pi_2^l(\beta \mid \alpha)\exp\left\{E\{\log\{h_1(\theta)\}\}^{\frac{1}{2}}\right\}.$$

Now the expectation can be simplified as

$$\begin{aligned}E\{\log\{h_1(\theta)\}\}^{\frac{1}{2}} &= \int_e^f \{\log\{h_1(\theta)\}\}\, \pi_2^l(\beta \mid \alpha)\, d\beta \\ &= \int_e^f \left\{\log\left\{\frac{n\beta^2(\gamma_2-\gamma_1^2)}{\alpha^2(1+2\gamma_1+\gamma_2)}\right\}^{\frac{1}{2}}\right\} \frac{\beta^{-1}}{\log\left(\frac{f}{e}\right)} d\beta \\ &= \int_e^f \log\left\{\beta\sqrt{n(\gamma_2-\gamma_1^2)}\right\} \frac{\beta^{-1}}{\log\left(\frac{f}{e}\right)} d\beta - \int_e^f \log\left\{\alpha\sqrt{\gamma+2\gamma_1+1}\right\} \frac{\beta^{-1}}{\log\left(\frac{f}{e}\right)} d\beta\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\{\log\{h_1(\theta)\}\}^{\frac{1}{2}} =& \int_e^f \log\{\beta\}\frac{\beta^{-1}}{\log\left(\frac{f}{e}\right)}d\beta + \int_e^f \log\left\{\sqrt{n\left(\gamma_2-\gamma_1^2\right)}\right\}\frac{\beta^{-1}}{\log\left(\frac{f}{e}\right)}d\beta \\
&- \frac{\log\left\{\alpha\sqrt{\gamma+2\gamma_1+1}\right\}}{\log\left(\frac{f}{e}\right)}\int_e^f \frac{1}{\beta}d\beta \\
=& \log\left(\frac{f}{e}\right)\int_e^f \frac{\log\{\beta\}}{\beta}d\beta + \frac{\log\left\{\sqrt{n\left(\gamma_2-\gamma_1^2\right)}\right\}}{\log\left(\frac{f}{e}\right)}\int_e^f \frac{1}{\beta}d\beta \\
&- \frac{\log\left\{\alpha\sqrt{\gamma+2\gamma_1+1}\right\}}{\log\left(\frac{f}{e}\right)}\int_e^f \frac{1}{\beta}d\beta \\
=& \log\left(\frac{f}{e}\right)\int_e^f \frac{\log\{\beta\}}{\beta}d\beta + \frac{\log\left\{\sqrt{n\left(\gamma_2-\gamma_1^2\right)}\right\}}{\log\left(\frac{f}{e}\right)}\log\left(\frac{f}{e}\right) \\
&- \frac{\log\left\{\alpha\sqrt{\gamma+2\gamma_1+1}\right\}}{\log\left(\frac{f}{e}\right)}\log\left(\frac{f}{e}\right) \\
=& \log\left(\frac{f}{e}\right)\int_e^f \frac{\log\{\beta\}}{\beta}d\beta + \log\left\{\sqrt{n\left(\gamma_2-\gamma_1^2\right)}\right\} - \log\left\{\alpha\sqrt{\gamma+2\gamma_1+1}\right\} \\
=& \log\left(\frac{f}{e}\right)\int_e^f \frac{\log\{\beta\}}{\beta}d\beta + \log\left\{\sqrt{n\left(\gamma_2-\gamma_1^2\right)}\right\} - \log(\alpha) - \log\left\{\sqrt{\gamma+2\gamma_1+1}\right\}.
\end{aligned}$$

Now evaluating $\int_e^f \frac{\log\{\beta\}}{\beta}d\beta$, let $u=\log\{\beta\}$, then $\frac{du}{d\beta}=\frac{1}{\beta}$

$$\int_e^f \frac{\log\{\beta\}}{\beta}d\beta = \int_{\log(e)}^{\log(f)} u\,du = \frac{u^2}{2}\Big|_{\log(e)}^{\log(f)} = C.$$

Now, let $C_1 = \log\left(\frac{f}{e}\right) C + \log\left\{\sqrt{n\left(\gamma_2 - \gamma_1^2\right)}\right\} - \log\left\{\sqrt{\gamma + 2\gamma_1 + 1}\right\}$, then this follows

$$\begin{aligned}
\pi^l(\alpha,\beta) &= \frac{\frac{\beta^{-1}}{\log\left(fe^{-1}\right)} \exp\{C_1 - \log(\alpha)\}}{\int_a^b \exp\{C_1 - \log(\alpha)\}\, d\alpha} \\
&= \frac{\frac{\beta^{-1}\alpha^{-1}}{\log\left(fe^{-1}\right)} \exp\{C_1\}}{\exp\{C_1\} \int_a^b \frac{1}{\alpha} d\alpha} \\
&= \frac{\frac{\beta^{-1}\alpha^{-1}}{\log\left(fe^{-1}\right)}}{\int_a^b \frac{1}{\alpha} d\alpha} \\
&= \frac{\frac{\beta^{-1}\alpha^{-1}}{\log\left(fe^{-1}\right)}}{\log\left(ba^{-1}\right)} \\
\pi^l(\alpha,\beta) &\propto \beta^{-1}\alpha^{-1}.
\end{aligned}$$

The reference prior is given by

$$\begin{aligned}
\pi_R(\alpha,\beta) &= \frac{\pi^l(\alpha,\beta)}{\pi^l(\alpha^*,\beta^*)} \\
&\propto \frac{\beta^{-1}\alpha^{-1}}{\log\left(ba^{-1}\right)} \\
&\propto \frac{1}{\alpha\beta}.
\end{aligned}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the reference prior when there is no censoring is given by

$$\pi_R(\alpha,\beta|\underline{t}) \propto \beta^{n-1}\alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\}. \tag{3.17}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the reference prior when there is no censoring is proper, Sun (1997).

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi(\alpha|\beta,\underline{t}) \propto \alpha^{-(n\beta+1)} \exp\left\{-\frac{\sum_{i-1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\},$$

letting $u = \alpha^{\beta}$ in the previous expression, then it is obtained that, $u = \alpha^{\beta} \sim Inv - G\left(n, \sum_{i-1}^{n} t_i^{\beta}\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi(\beta|\alpha,\underline{t}) \propto \beta^{n-1}\alpha^{-n\beta} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\}.$$

The joint posterior of α and β for the Weibull distribution given the observed data using the reference prior under type I right censoring is given by

$$\pi_R(\alpha,\beta|\underline{t}) \propto \beta^{r-1}\alpha^{-r\beta-1}\left\{\prod_{i=1}^{r} t_i^{\beta-1}\right\}\exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}. \tag{3.18}$$

The joint posterior of α and β for the Weibull distribution given the observed data using the reference prior under type I right censoring is proper.

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi(\alpha|\beta,\underline{t}) \propto \alpha^{-(r\beta+1)} \exp\left\{-\frac{\sum_{i-1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\},$$

letting $u = \alpha^{\beta}$ in the previous expression, then it is obtained that, $u = \alpha^{\beta} \sim Inv - G\left(r, \sum_{i-1}^{n} t_i^{\beta}\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi(\beta|\alpha,\underline{t}) \propto \beta^{r-1}\alpha^{-r\beta} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}.$$

The conditional posterior of α given β and $\underline{t}$ has a known form, and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. Instead, a Gibbs sampler with a MH step (Metropolis-within-Gibbs) can be employed. This occurs both in the complete data case and under type I right censoring.

3.4 The Posterior for the Weibull Distribution using the Probability Matching Prior

In this section, the posterior distribution for the Weibull distribution using the probability matching prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined. Sun (1997) established that the reference prior of the Weibull distribution is the third-order probability matching prior. I now derive the reference prior for the Weibull model using the Datta and Ghosh (1995) algorithm, the result coincides with Sun (1997) reference prior, which was showed to be third-order probability matching prior.

The inverse of the Fisher information is given by

$$H^{-1}(\alpha,\beta)=\begin{bmatrix}\frac{\alpha^2(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)} & \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)}\\ \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)} & \frac{\beta^2}{n\left(\gamma_2-\gamma_1^2\right)}\end{bmatrix}.$$

Suppose the parameter of interest is $t(\theta)=\alpha$, then

$$\nabla_t^{\mathsf{T}}(\theta)=\left[\frac{\partial t(\theta)}{\partial\alpha}\ \frac{\partial t(\theta)}{\partial\beta}\right]=[1\ 0].$$

Then,

$$\eta_t^{\mathsf{T}}(\theta)=\frac{\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)}{\sqrt{\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)\nabla_t(\theta)}},$$

the numerator is given by

$$\begin{aligned}\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)&=[1\ 0]\begin{bmatrix}\frac{\alpha^2(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)} & \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)}\\ \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)} & \frac{\beta^2}{n\left(\gamma_2-\gamma_1^2\right)}\end{bmatrix}\\ &=\left[\frac{\alpha^2(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)}\ \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)}\right],\end{aligned}$$

and

$$\begin{aligned}\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)\nabla_t(\theta)&=\left[\frac{\alpha^2(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)}\ \frac{\alpha(1+\gamma_1)}{n\left(\gamma_2-\gamma_1^2\right)}\right]\begin{bmatrix}1\\0\end{bmatrix}\\ &=\frac{\alpha^2(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)}\end{aligned}$$

as a results

$$\eta_t^{\mathsf{T}}(\theta)=\left[\frac{\alpha(1+2\gamma_1+\gamma_2)}{n\beta\left(\gamma_2-\gamma_1^2\right)\sqrt{\frac{1+2\gamma_1+\gamma_2}{n\left(\gamma_2-\gamma_1^2\right)}}}\ \frac{\beta(1+\gamma_1)}{n\sqrt{\frac{1+2\gamma_1+\gamma_2}{n\left(\gamma_2-\gamma_1^2\right)}}\left(\gamma_2-\gamma_1^2\right)}\right].$$

The probability matching prior is given byT

$$\pi_{PMP}(\alpha,\beta)\propto\frac{1}{\alpha\beta}.$$

Since

$$\frac{\partial}{\partial\alpha}\{\eta_1(\theta)\pi(\theta)\}+\frac{\partial}{\partial\beta}\{\eta_2(\theta)\pi(\theta)\}=\frac{\partial}{\partial\alpha}\left\{\frac{\alpha(1+2\gamma_1+\gamma_2)}{n\beta\left(\gamma_2-\gamma_1^2\right)\sqrt{\frac{1+2\gamma_1+\gamma_2}{n\left(\gamma_2-\gamma_1^2\right)}}}\frac{1}{\alpha\beta}\right\}$$
$$+\frac{\partial}{\partial\beta}\left\{\frac{\beta(1+\gamma_1)}{n\sqrt{\frac{1+2\gamma_1+\gamma_2}{n\left(\gamma_2-\gamma_1^2\right)}}\left(\gamma_2-\gamma_1^2\right)}\frac{1}{\alpha\beta}\right\},$$

simplified to

$$\frac{\partial}{\partial\alpha}\{\eta_1(\theta)\pi(\theta)\}+\frac{\partial}{\partial\beta}\{\eta_2(\theta)\pi(\theta)\}=\frac{\partial}{\partial\alpha}\left\{\frac{(1+2\gamma_1+\gamma_2)}{n\beta^2\left(\gamma_2-\gamma_1^2\right)\sqrt{\frac{1+2\gamma_1+\gamma_2}{n\left(\gamma_2-\gamma_1^2\right)}}}\right\}$$
$$+\frac{\partial}{\partial\beta}\left\{\frac{(1+\gamma_1)}{n\alpha\sqrt{\frac{1+2\gamma_1+\gamma_2}{n\left(\gamma_2-\gamma_1^2\right)}}\left(\gamma_2-\gamma_1^2\right)}\right\}$$
$$\frac{\partial}{\partial\alpha}\{\eta_1(\theta)\pi(\theta)\}+\frac{\partial}{\partial\beta}\{\eta_2(\theta)\pi(\theta)\}=0.$$

The joint posterior of α and β for the Weibull distribution given the observed data using the probability matching prior when there is no censoring is given by

$$\pi_{PMP}(\alpha,\beta|t)\propto\beta^{n-1}\alpha^{-(n\beta+1)}\exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n}t_i^{\beta}\right\}. \tag{3.19}$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi(\alpha|\beta,\underline{t})\propto\alpha^{-(n\beta+1)}\exp\left\{-\frac{\sum_{i-1}^{n}t_i^{\beta}}{\alpha^{\beta}}\right\},$$

letting $u=\alpha^{\beta}$ in the previous expression, then it is obtained that, $u=\alpha^{\beta}\sim Inv-G\left(n,\sum_{i-1}^{n}t_i^{\beta}\right)$.
The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi(\beta|\alpha,\underline{t})\propto\beta^{n-1}\alpha^{-n\beta}\exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n}t_i^{\beta}\right\}.$$

The joint posterior of α and β for the Weibull distribution given the observed data using the probability matching prior under type I censoring is given by

$$\pi_{PMP}(\alpha,\beta|\underline{t}) \propto \beta^{r-1}\alpha^{-r\beta-1}\left\{\prod_{i=1}^{r} t_i^{\beta-1}\right\}\exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}. \tag{3.20}$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi(\alpha|\beta,\underline{t}) \propto \alpha^{-(r\beta+1)}\exp\left\{-\frac{\sum\limits_{i-1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\},$$

letting $u=\alpha^{\beta}$ in the previous expression, then it is obtained that, $u=\alpha^{\beta}\sim Inv-G\left(r,\sum\limits_{i-1}^{n} t_i^{\beta}\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi(\beta|\alpha,\underline{t}) \propto \beta^{r-1}\alpha^{-r\beta}\exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}.$$

The conditional posterior of α given β and $\underline{t}$ has a known form, and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. Instead, a Gibbs sampler with a MH step (Metropolis-within-Gibbs) can be employed. This occurs both in the complete data case and under type I right censoring. These posterior distributions coincide with those obtained using the reference prior, and in both cases, they are proper.

3.5 Assessment of Posterior Convergence through MCMC for the Weibull Distribution

This section outlines the simulation procedure used to evaluate the performance of Bayesian inference for the Weibull distribution under various objective priors. The study focuses on estimating the shape and scale parameters of the Weibull distribution using MCMC techniques under different sample sizes. Simulations were conducted under no censoring or complete data and type I right censoring. The simulated datasets were generated using various combinations of the true Weibull parameters: scale parameters, $\alpha \in \{0.5,2,5\}$, and shape parameter, $\beta \in \{0.8,5,8\}$. A total of $T=20000$ iterations were performed for each simulation, with a burn-in period of $T/2$. The convergence of the chains was monitored using a trace plot, the Geweke diagnostic Z-score, and the Gelman–Rubin diagnostic. Autocorrelation plots were used to examine the dependence of the sampled β on the chosen initial values of the Markov chain. See Appendix C.1 for the R code.

Algorithm. 3.1 *Simulation procedure for the Weibull distribution*
Case 1: Complete data

- Step 0

  Simulate t from Weibull distribution with true parameters and fix the sample size n.

- Step 1

  Initialize

  Choose $\alpha^{(0)}$, $\beta^{(0)} > 0$, tuning parameter σ, iterations T.

- Step 2

  Sample $\alpha|\beta, \underline{t}$ (Gibbs step)

  Compute $S = \sum_{i=1}^{n} t_i^{\beta}$

  Draw $u \sim Gamma\,(n,\ rate = S)$

  Set $\alpha = u^{-\frac{1}{\beta}}$

- Step 3

  Update $\beta \mid \alpha, t$

  Propose

$$\log(\beta^*) \sim N\left(\log(\beta), \sigma^2\right)$$

  Compute the log-posterior kernel for β

$$\log(\pi(\beta|\alpha,\underline{t})) = n\ln(\beta) + (\beta-1)\sum_{i=1}^{n}\ln(t_i) - n\beta\ln(\alpha) - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}.$$

  Accept β^* with probability

$$min\{1,\ \exp[\log(\pi(\beta^*|\alpha,\underline{t})) - \log(\pi(\beta|\alpha,\underline{t}))]\}.$$

  repeat $\tau = 1, \ldots, T$. Discard burn-in and use remainder of samples for inference.

- Step 4

Repeat

Iterate Steps 2–3 for $\tau = 1, ..., T$.

After discarding burn-in iterations, use the remaining samples to estimate the posterior distributions of α and β.

Algorithm. 3.2 *Simulation Procedure for the Weibull distribution*
Case 2: Type I right censoring

- Step 0

Simulate t from Weibull distribution and choose sample size n, also note that r is the number of failures. A threshold of $C = 2.5$ was used.

- Step 1

Initialize

Choose $\alpha^{(0)}$, $\beta^{(0)} > 0$, tuning parameter σ, iterations T.

- Step 2

Sample $\alpha|\beta, \underline{t}$

Compute $S = \sum_{i=1}^{n} t_i^{\beta}$.

Draw $u \sim Gamma\left(r,\ rate = S\right)$.

Set $\alpha = u^{-\frac{1}{\beta}}$

- Step 3

Proposal density

Propose $\log\left(\beta^*\right) \sim N\left(\log\left(\beta\right), \sigma^2\right)$.

Compute

$\log\left(\pi\left(\beta|\alpha,\underline{t}\right)\right) = r\ln\left(\beta\right) + (\beta - 1)\sum_{i=1}^{n}\ln\left(t_i\right) - r\beta\ln\left(\alpha\right) - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}$.

Accept β^* with probability: $\min\{1,\ \exp\left[\log\left(\pi\left(\beta^*|\alpha,\underline{t}\right)\right) - \log\left(\pi\left(\beta|\alpha,\underline{t}\right)\right)\right]\}$.

- Step 4

Repeat

Iterate Steps 2–3 for $\tau = 1, ..., T$.

After discarding burn-in iterations, use the remaining samples to estimate the posterior distributions of α and β.

3.5.1 Complete Data

In this subsection, the convergence and posterior behaviour of the Weibull parameters α and β are evaluated under complete data. The convergence will be evaluated using the trace plots, the Gelman-Rubin diagnostic plot, and the Geweke diagnostic. These diagnostics are applied to both shape and scale parameters.

Figures 3.3, 3.4, and 3.5 represent the autocorrelation plots for β under various priors for $n = 10$, 40, and 200, respectively. For the small sample size $n = 10$, autocorrelation remains high across several lags, indicating slower mixing and stronger dependence between successive samples due to the influence of initial values. As the sample size increases, autocorrelation declines more rapidly toward zero, reflecting improved mixing and greater independence among posterior samples.

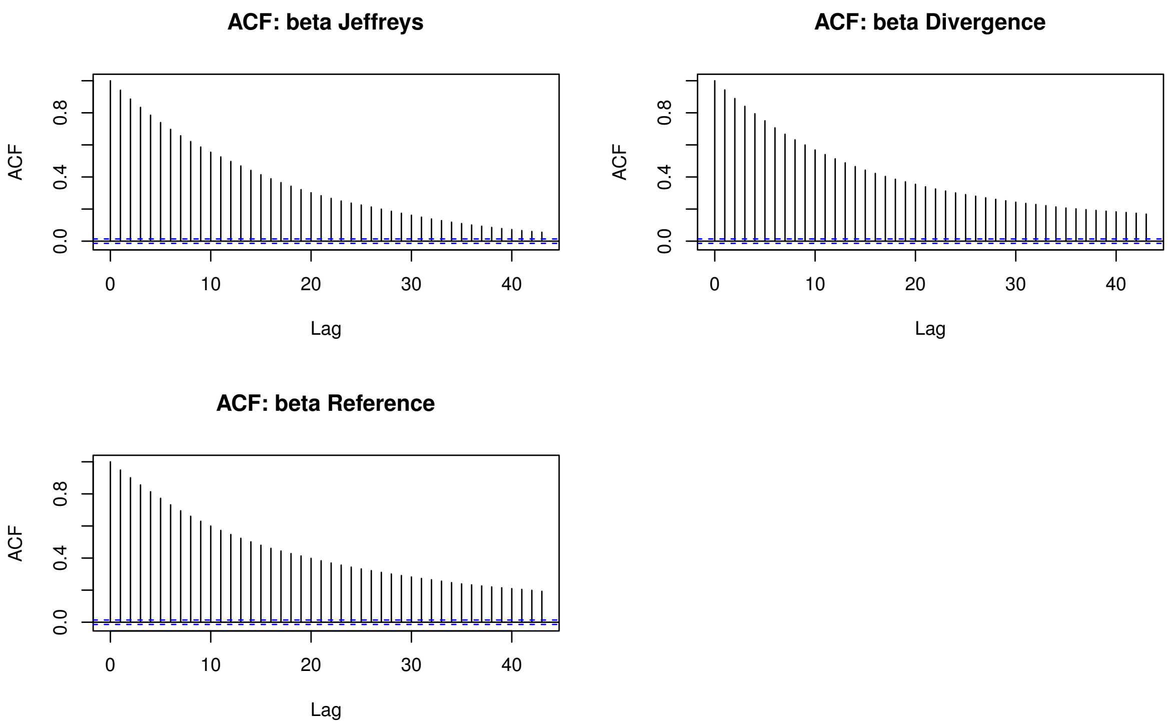

Figure 3.3: Autocorrelation for various priors when $n = 10$.

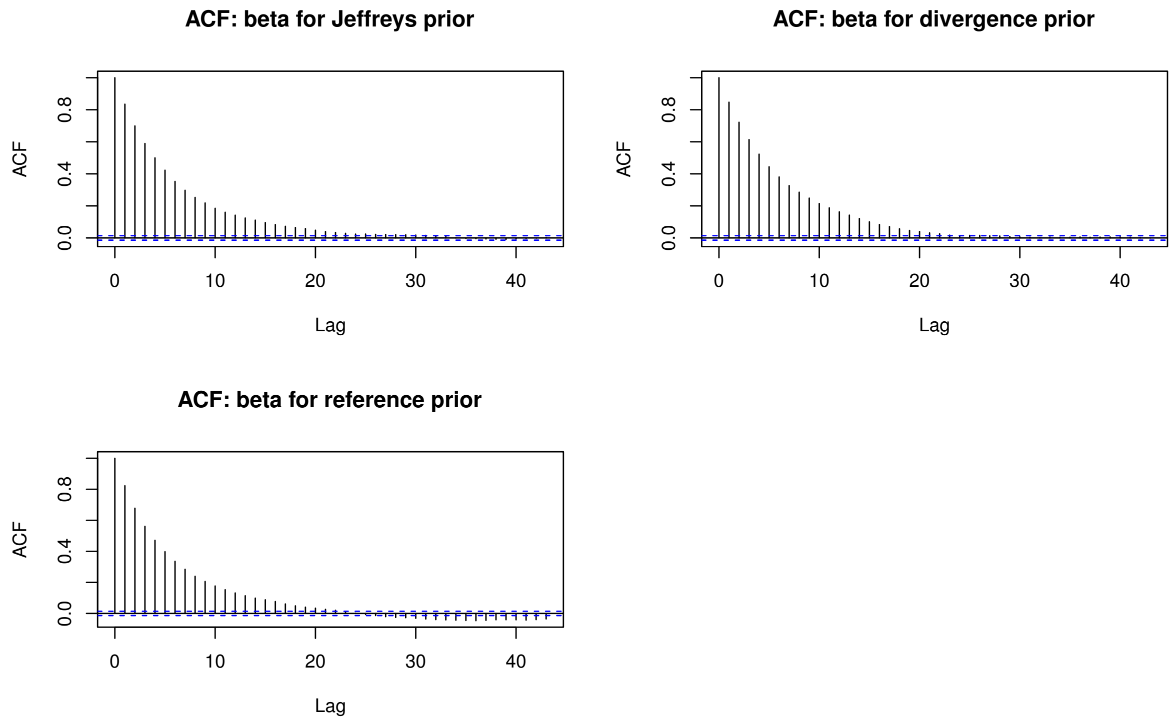

Figure 3.4: Autocorrelation for various priors when $n = 40$.

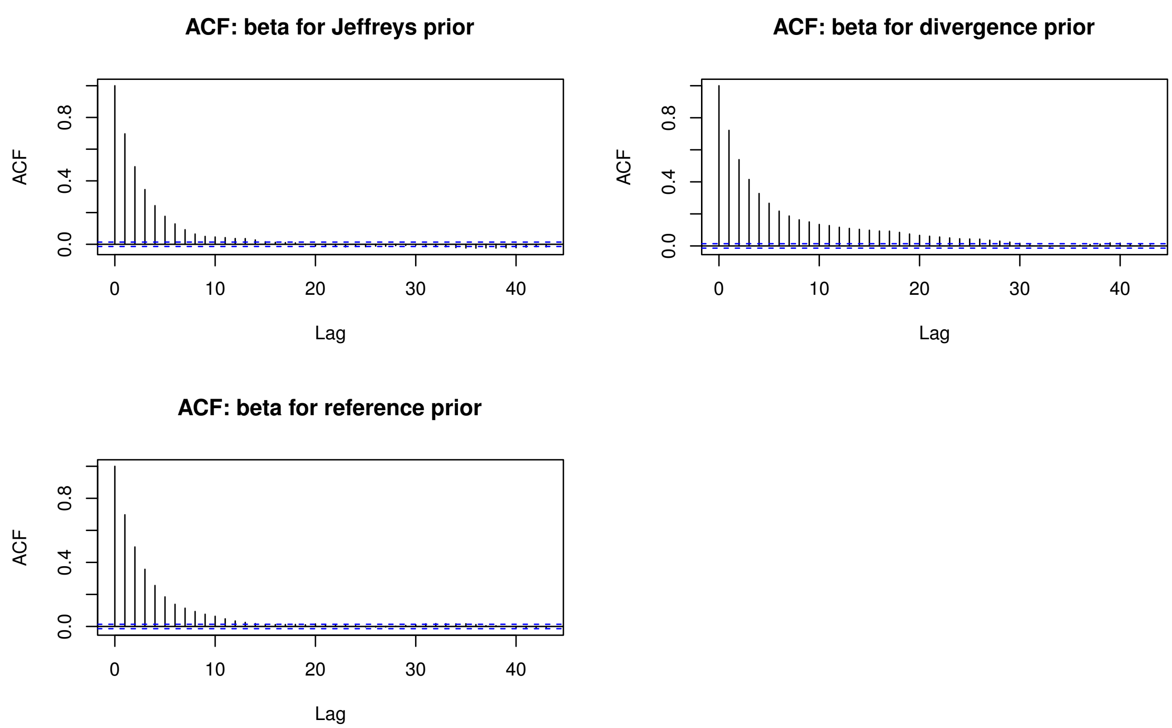

Figure 3.5: Autocorrelation for various priors when $n = 200$.

Figures 3.6, 3.7, and 3.8 represent the trace plots for α and β parameters under a sample size

$n = 10$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. For $n = 10$, the thick pen test confirms that the chains for both α and β have successfully converged when these three priors are used. It can be observed that the chains for α converge more quickly compared to those for β. Across all posterior distributions, the α chains exhibit rapid mixing and have reached a stable state around their posterior means. Figures 3.9, 3.10, and 3.11 represent the trace plots for the α and β parameters under a sample size of $n = 40$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. For $n = 40$, the thick pen test further confirms satisfactory convergence for both α and β, across all priors. The trace plots indicate that the chains mix well and stabilize more quickly than in the $n = 10$ case, reflecting the improved efficiency of posterior sampling with increased sample size. Overall, both parameters display consistent convergence behaviour and stable posterior means.

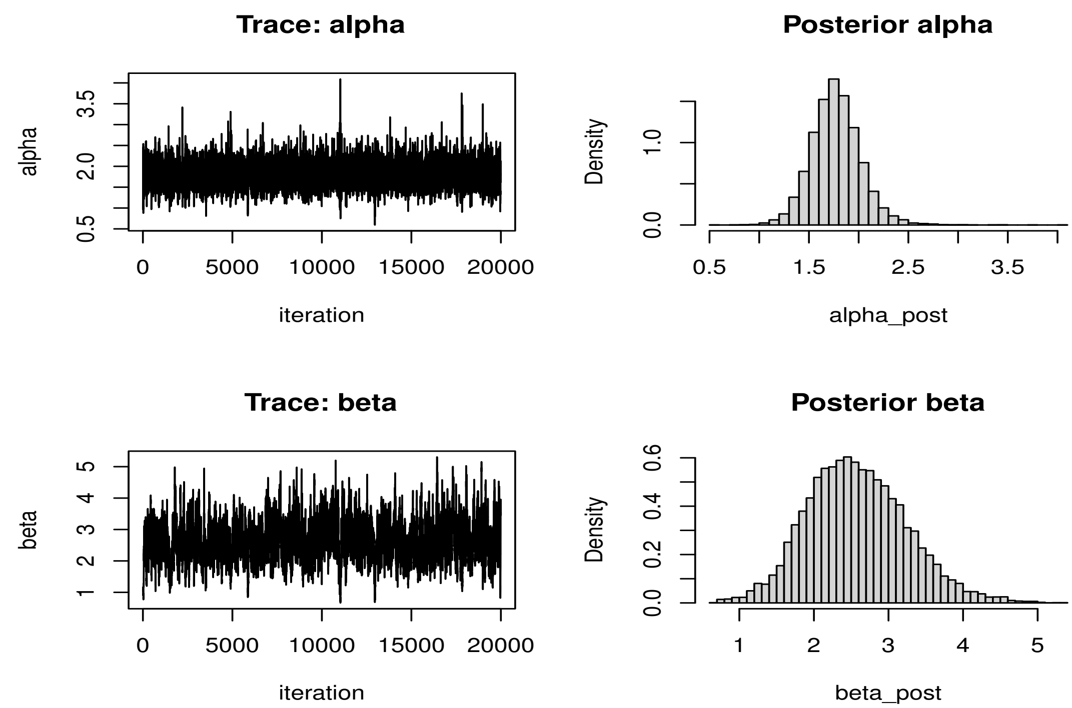

Figure 3.6: Trace plot and probability density function of α and β when n = 10 using the Jeffreys prior.

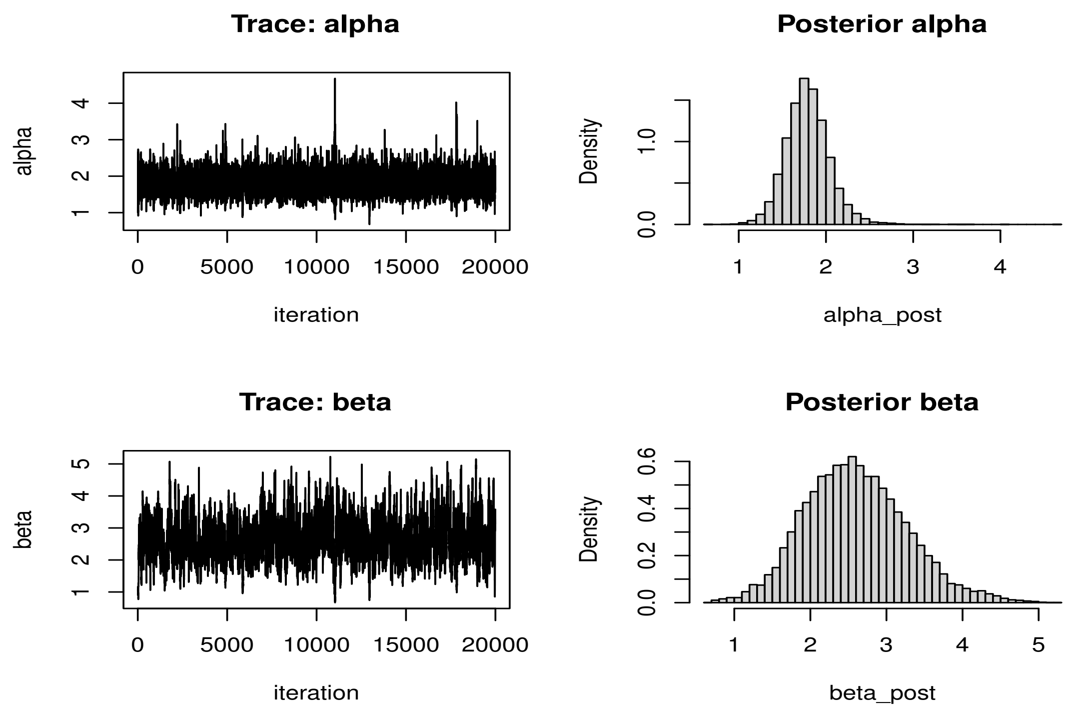

Figure 3.7: Trace plot and probability density function of α and β when $n = 10$ using divergence prior.

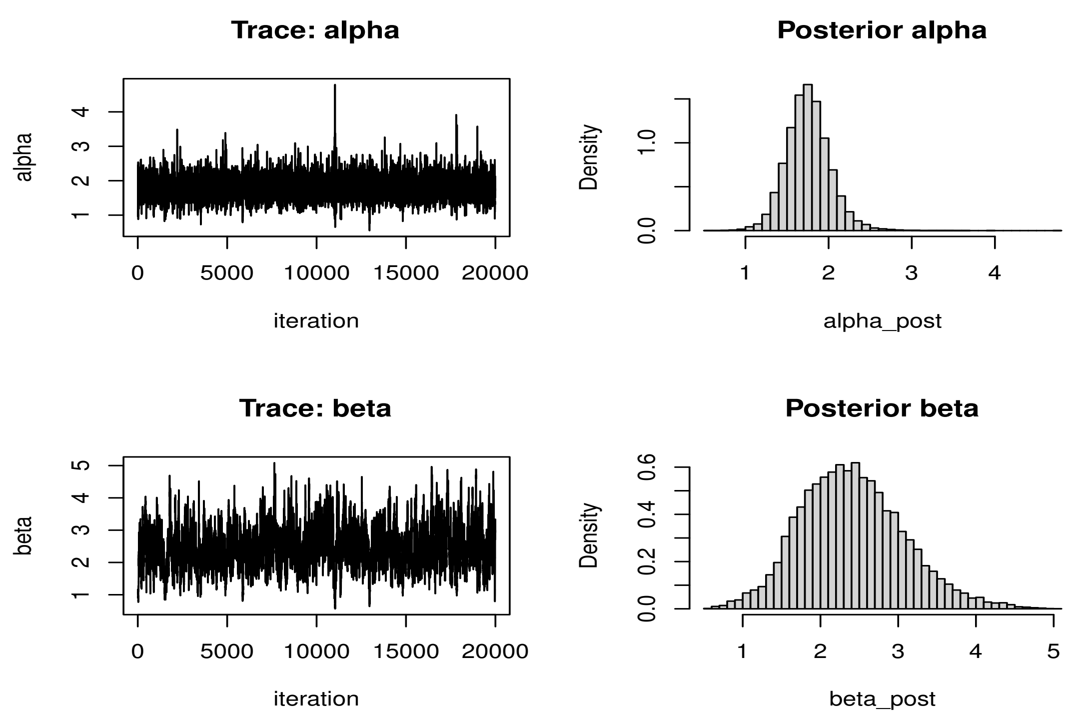

Figure 3.8: Trace plot and probability density function of α and β when n=10 using the reference prior.

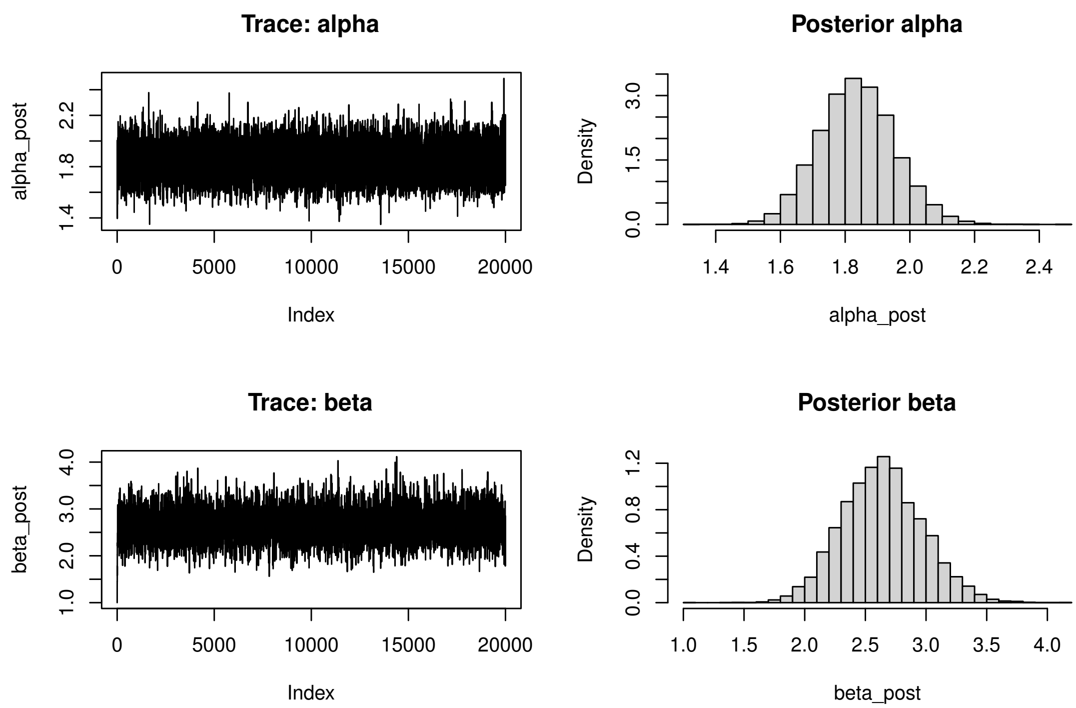

Figure 3.9: Trace plot and probability density function of α and β when $n = 40$ using the Jeffreys prior.

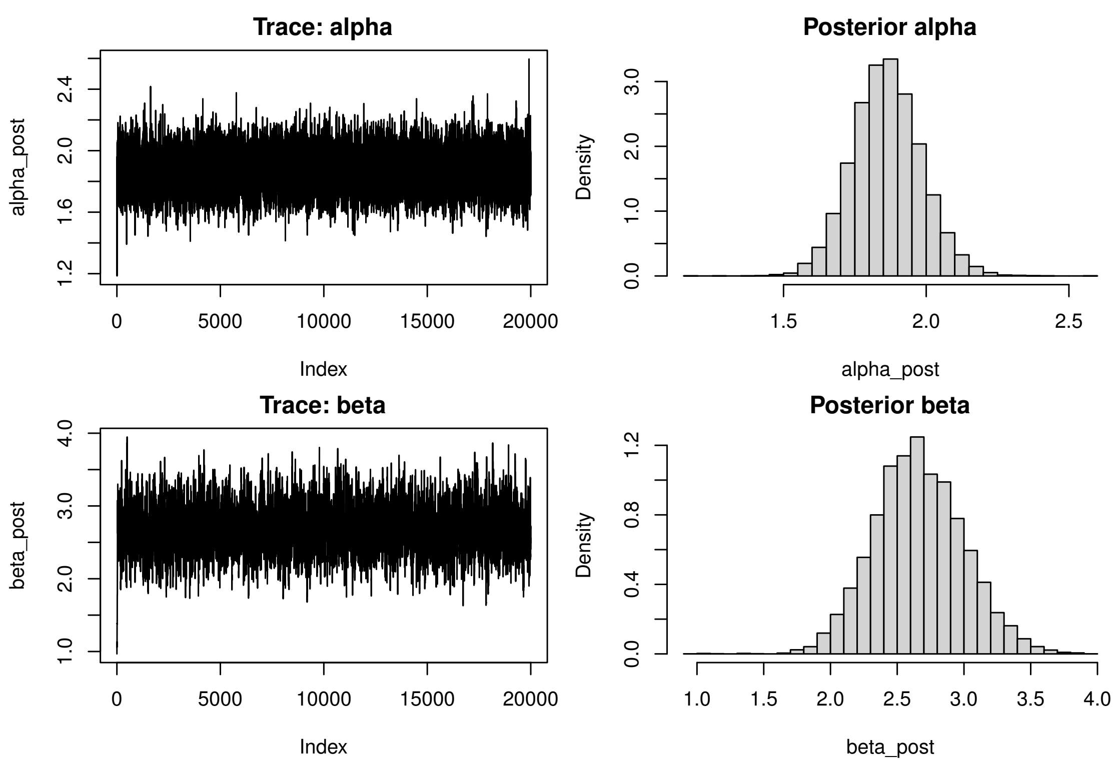

Figure 3.10: Trace plot and probability density function of α and β when n= 40 using the divergence prior.

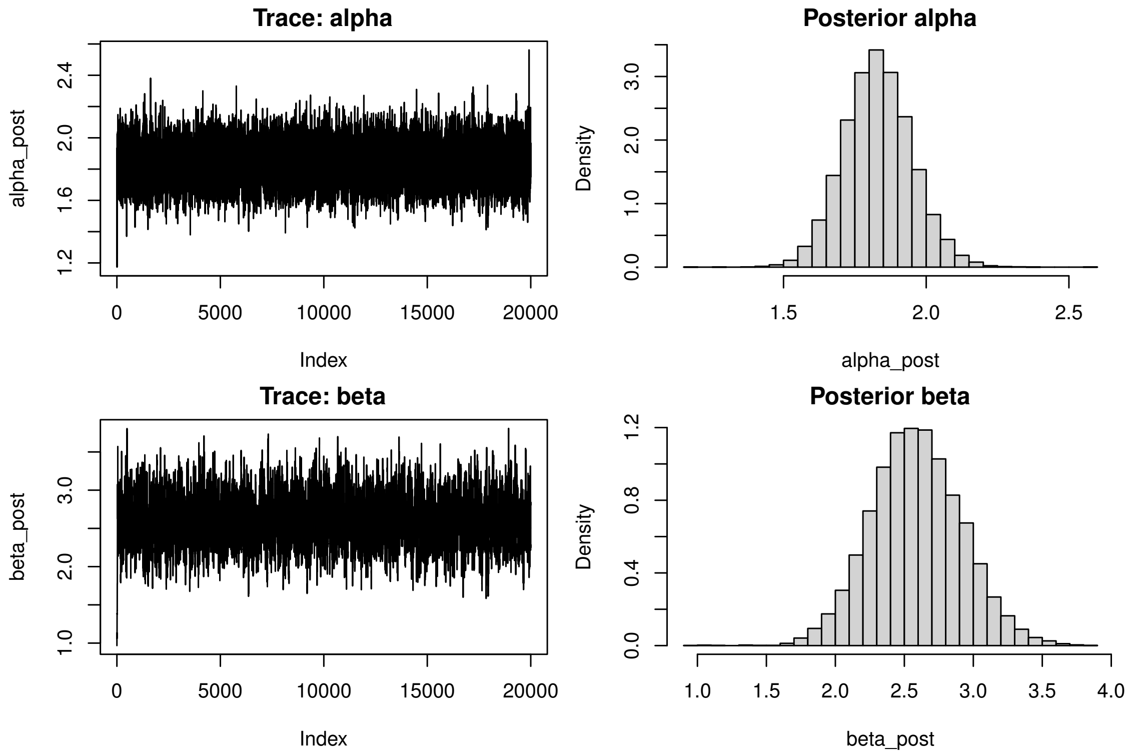

Figure 3.11: Trace plot and probability density function of α and β when n =40 using the reference prior.

Figures 3.12, 3.13, and 3.14 present the trace plots for the α and β parameters under a sample size of $n = 200$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. When $n = 200$, the thick pen test demonstrates an excellent convergence for both α and β. The MCMC chains from all three posterior distributions, producing a very solid and uniform band for each parameter around their posterior mean values. This indicates that the initial values have no influence , and the chains are efficiently sampling from the stationary posterior distribution. The large sample provides high precision in posterior estimates. There are no noticeable differences that are observed in the trace plot patterns when the Jeffreys prior, the divergence prior, or the reference prior are used; each yields similar chain behaviour for both parameters.

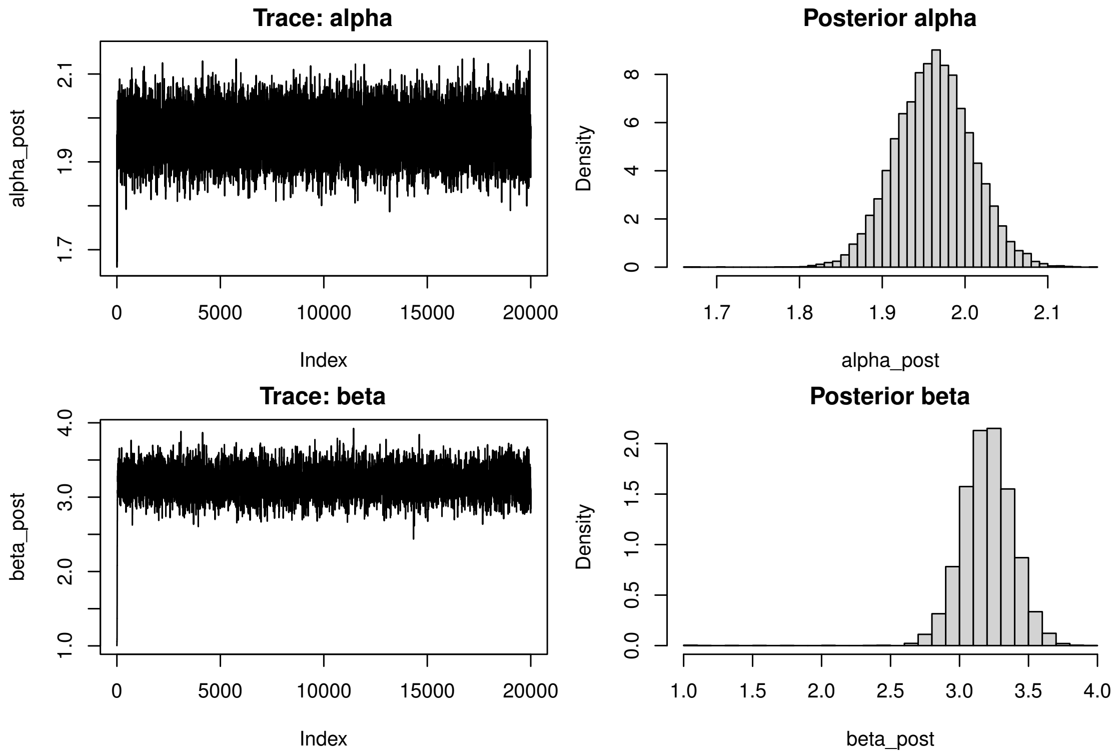

Figure 3.12: Trace plot and probability density function of α and β when $n = 200$ using the Jeffreys prior.

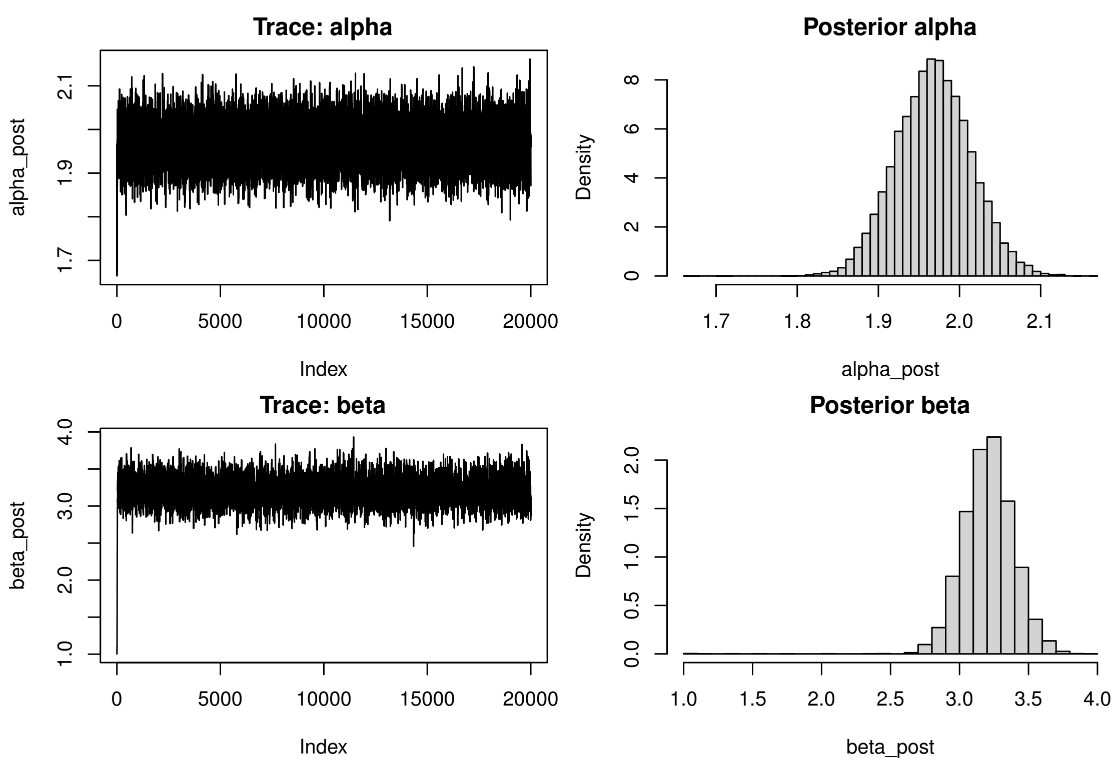

Figure 3.13: Trace plot and probability density function of α and β when n =200 using the divergence prior

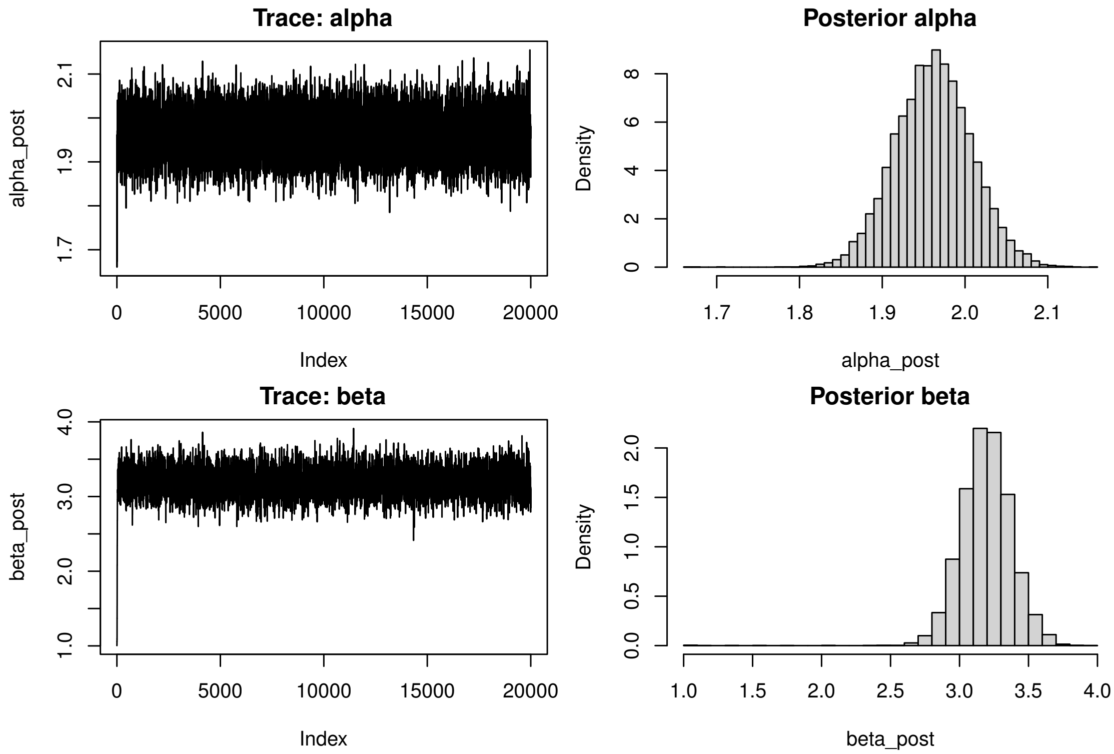

Figure 3.14: Trace plot and probability density function of α and β when n= 200 using the reference prior.

Figures 3.15, 3.16, and 3.17 present the Gelman-Rubin diagnostics for α and β when $n = 10$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. Figures 3.18, 3.19, and 3.20 present the Gelman-Rubin diagnostics for α and β for $n = 40$, when the Jeffreys prior, the divergence prior, and the reference prior are used, respectively. Figures 3.21, 3.22, and 3.23 present the Gelman-Rubin diagnostics for α and β when $n = 200$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively.

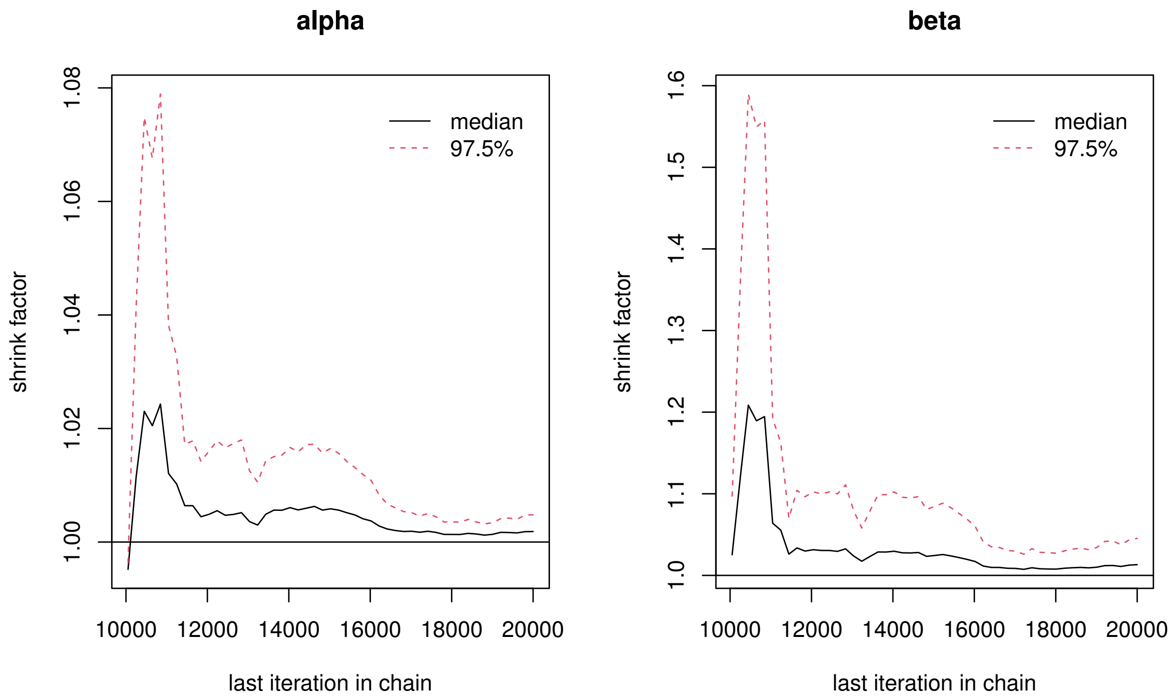

Figure 3.15: The Gelman-Rubin plots for α and β when $n = 10$ using the Jeffreys prior.

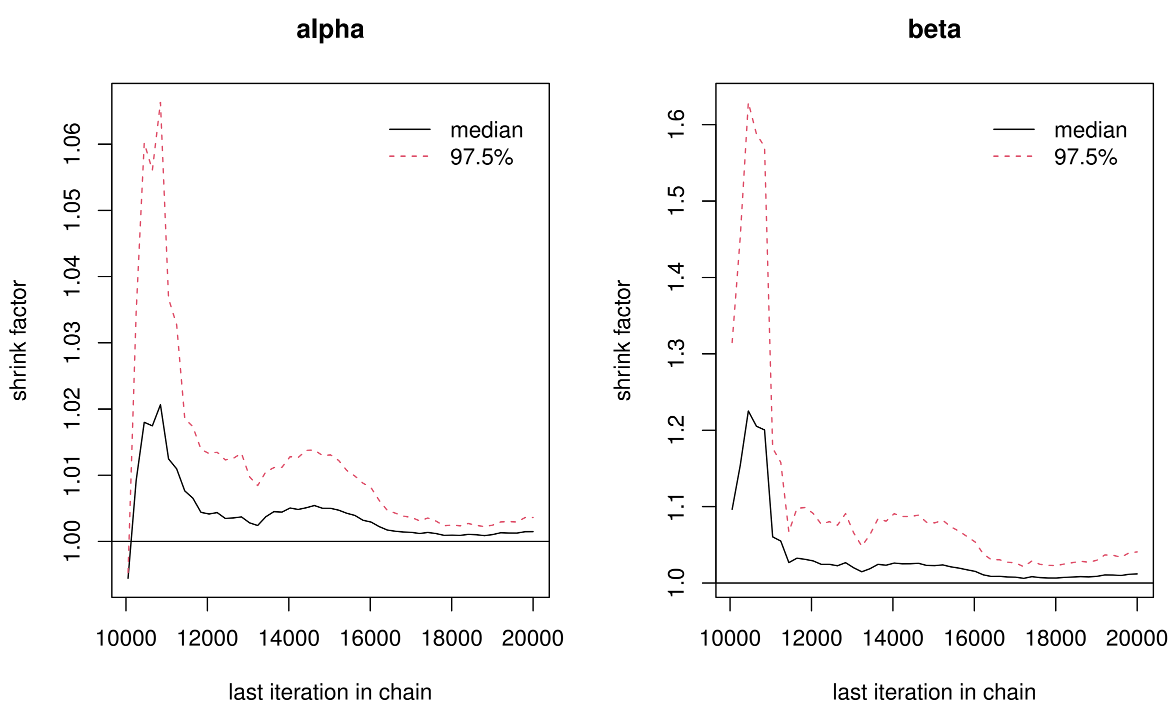

Figure 3.16: The Gelman-Rubin plots for α and β when n =10 using the divergence prior.

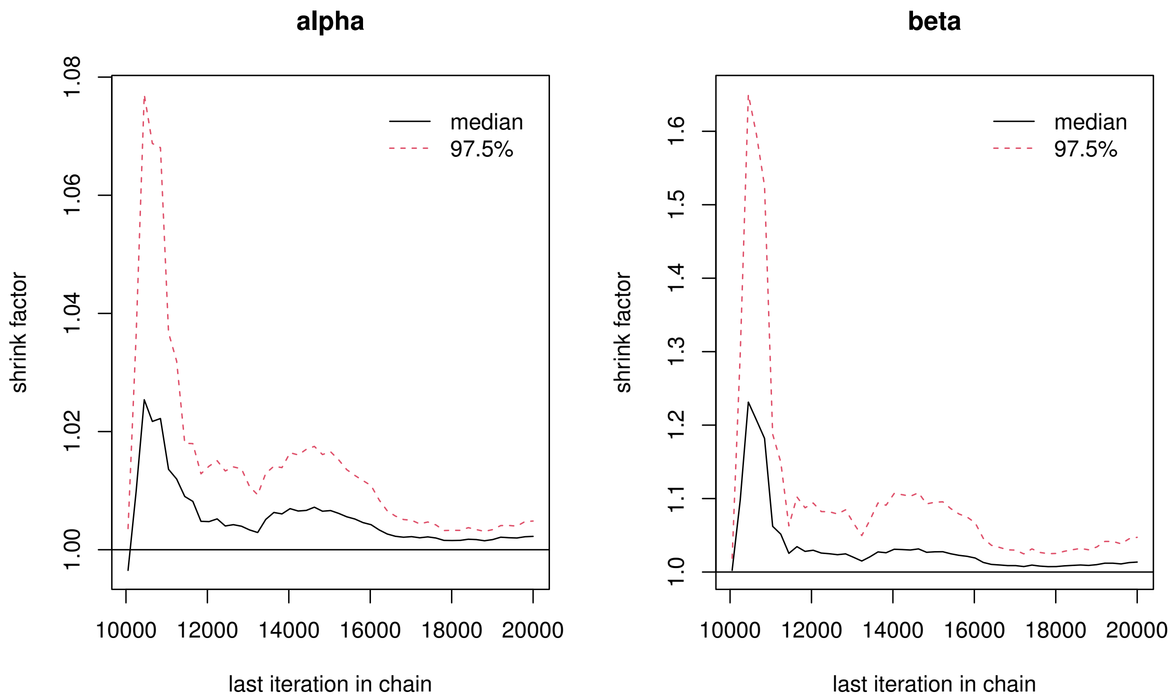

Figure 3.17: The Gelman-Rubin plots for α and β when $n = 10$ using the reference prior.

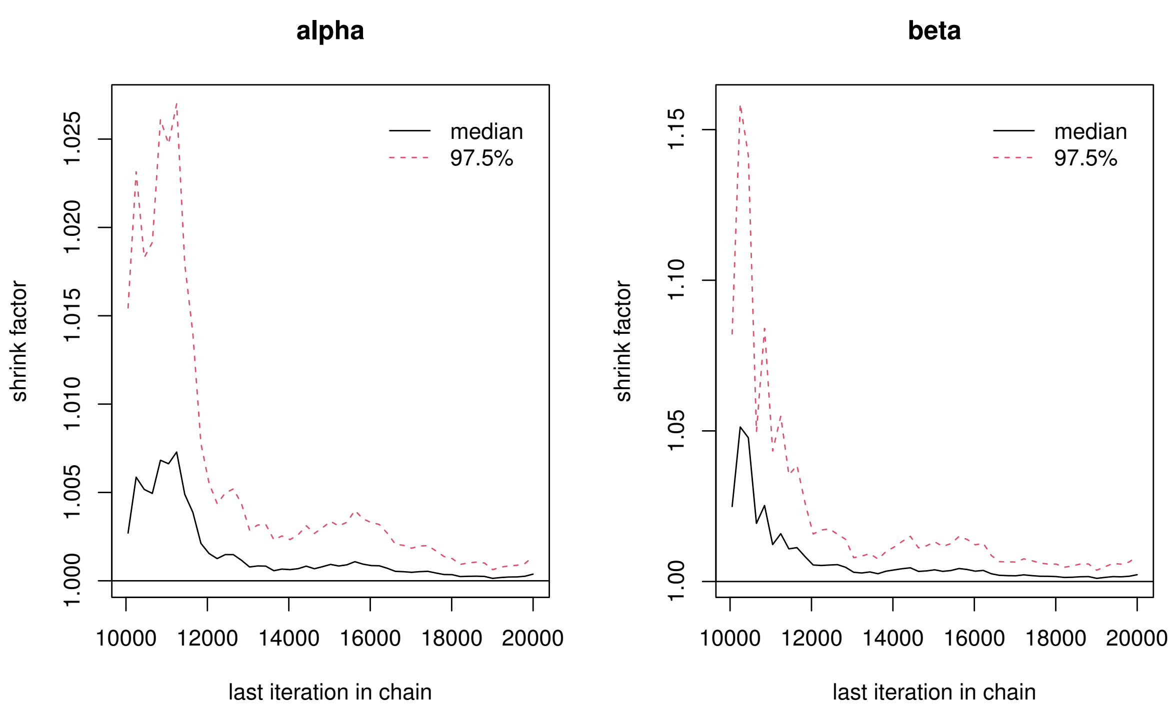

Figure 3.18: The Gelman-Rubin plots for α and β when $n = 40$ using the Jeffreys prior.

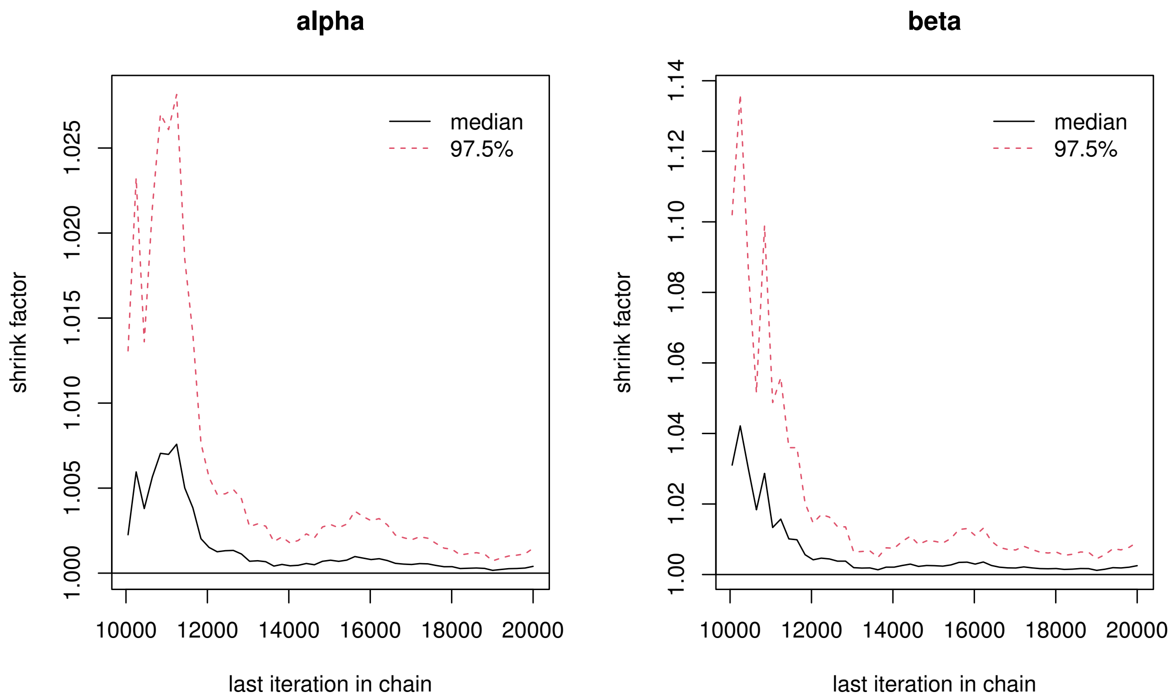

Figure 3.19: The Gelman-Rubin plots for α and β when $n = 40$ using the divergence prior.

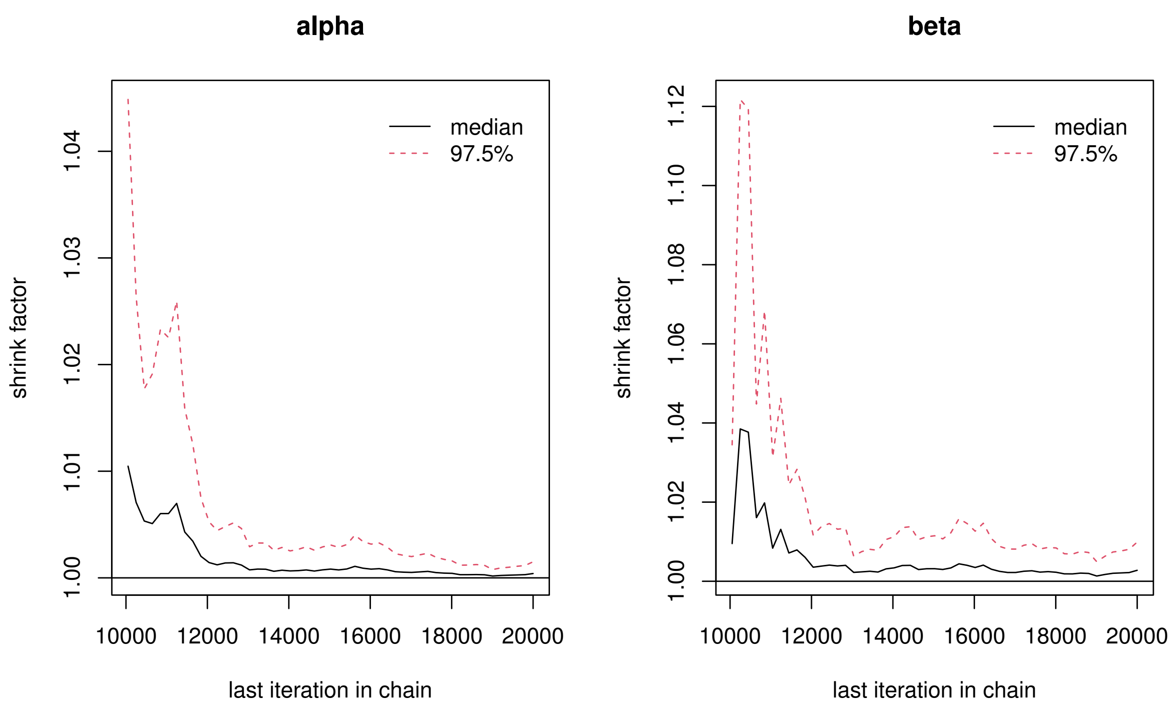

Figure 3.20: The Gelman-Rubin plots for α and β when $n = 40$ using the reference prior.

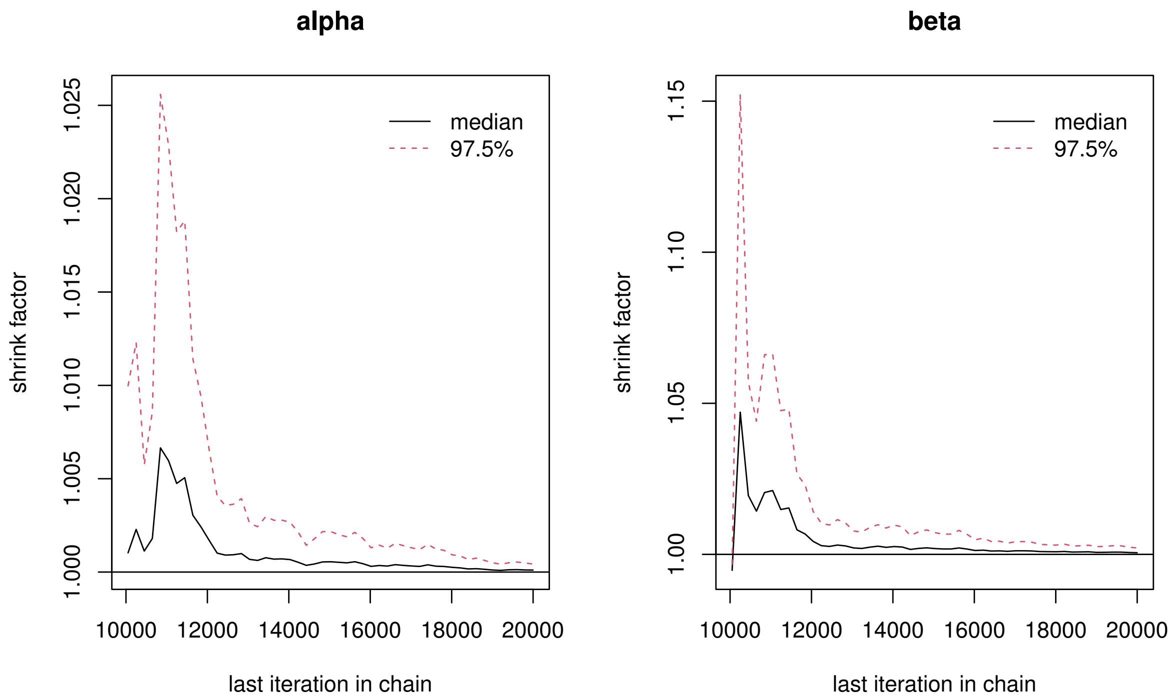

Figure 3.21: The Gelman-Rubin plots for α and β when n =200 using the Jeffreys prior.

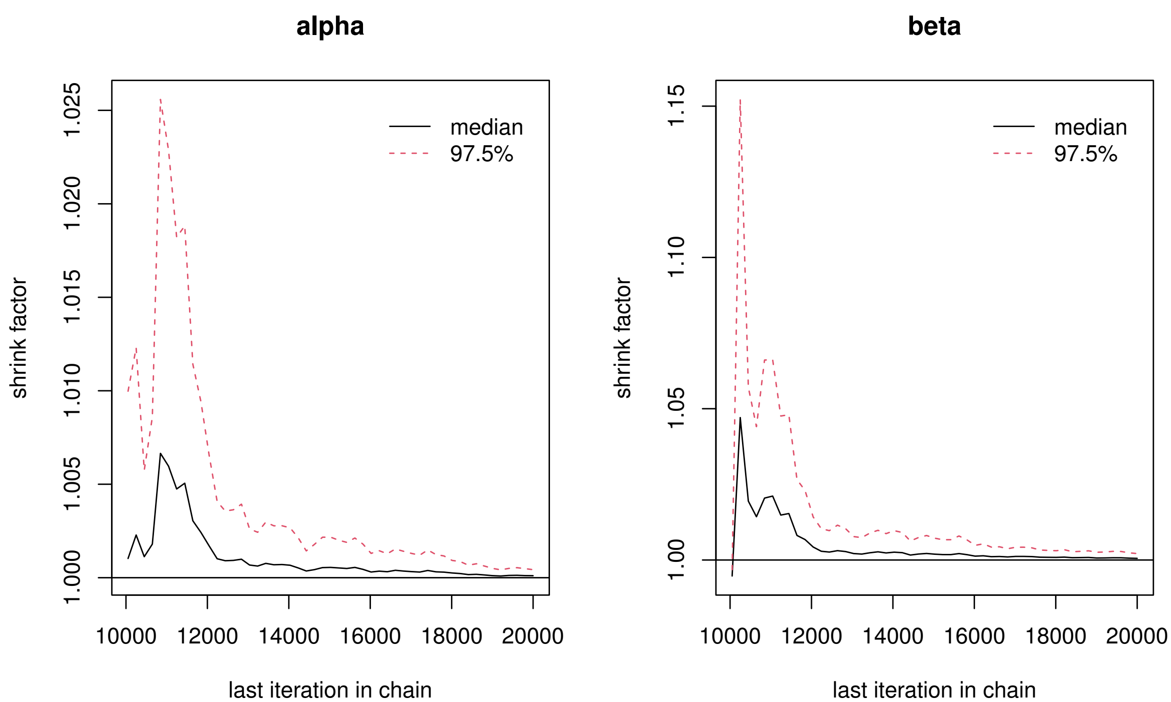

Figure 3.22: The Gelman-Rubin plots for α and β when $n = 200$ using the divergence prior.

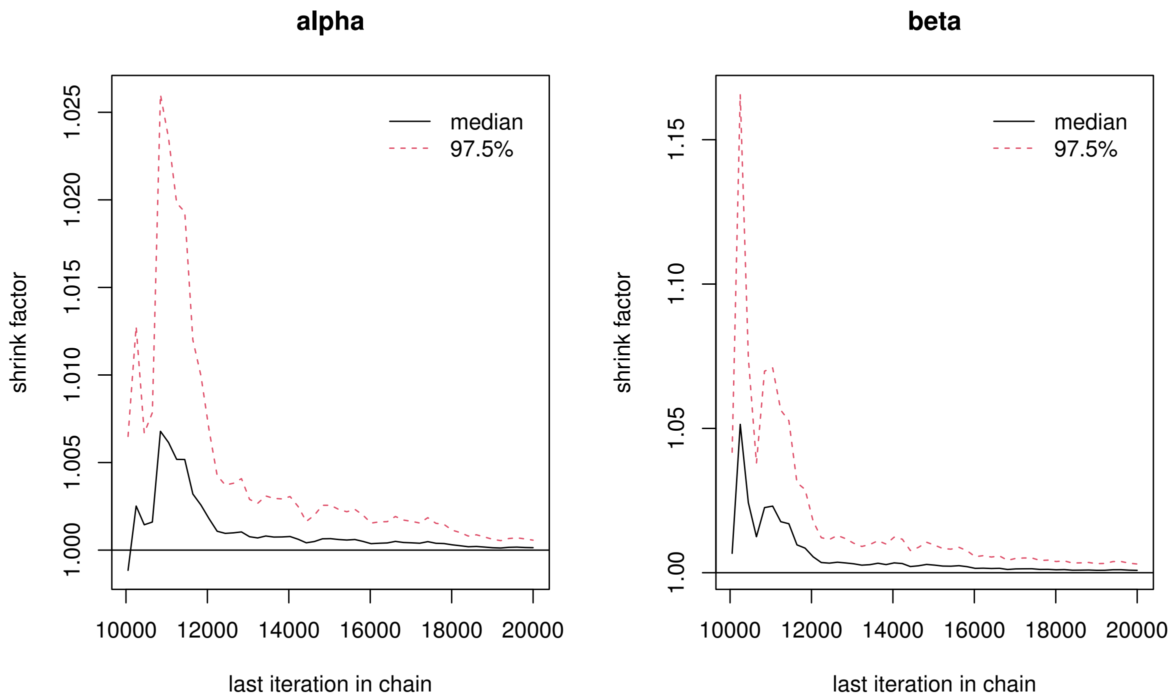

Figure 3.23: The Gelman-Rubin plots for α and β when $n = 200$ using the reference prior.

The Gelman-Rubin diagnostic plots were constructed above to further confirm the convergence of both the shape, β, and scale, α parameters under a sample size of 10, 40, and 200. The Gelman-Rubin diagnostic plots for both parameters α and β indicate that the shrink factor approaches 1 as the number of iterations increases, confirming convergence. For a small sample size $n = 10$, the shrink factor requires more iterations to be at a stationary point or to reach 1. As the sample size increases, the shrink factor for both parameters converges to 1 more quickly, indicating that a large sample size leads to clear convergence across chains.

Figures 3.24, 3.25, and 3.26 present the Geweke diagnostics for α and β for $n = 10$, when the Jeffreys prior, divergence prior, and reference prior are used, respectively.

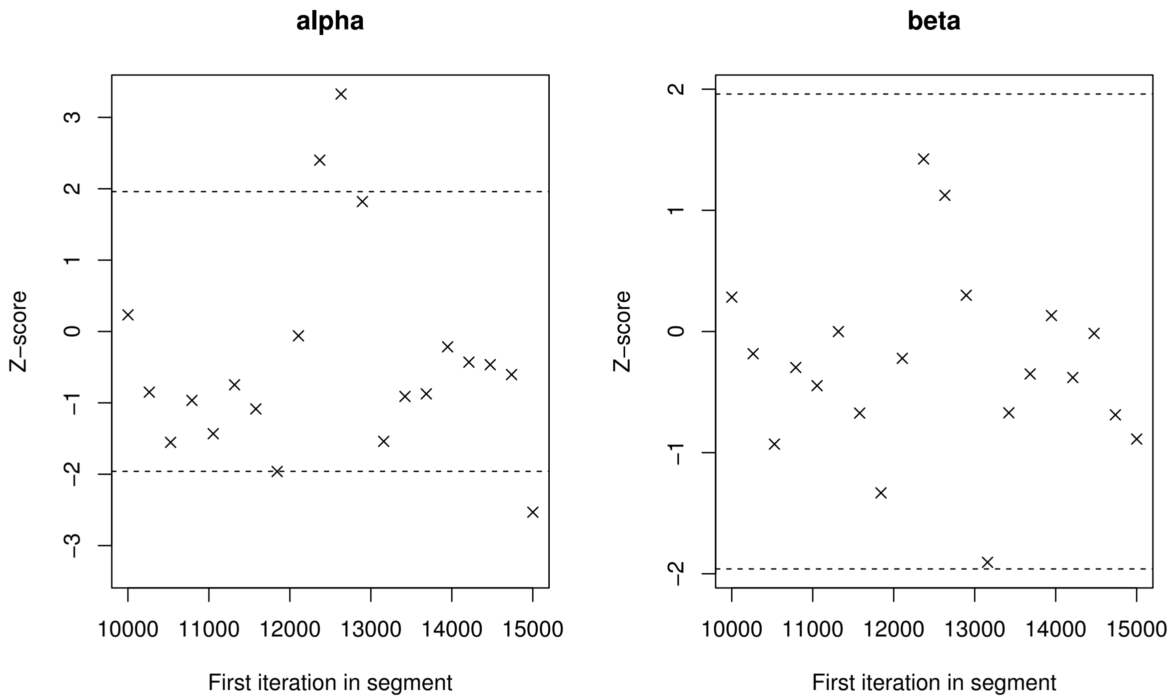

Figure 3.24: The Geweke plots for α and β when $n = 10$ using the Jeffreys prior.

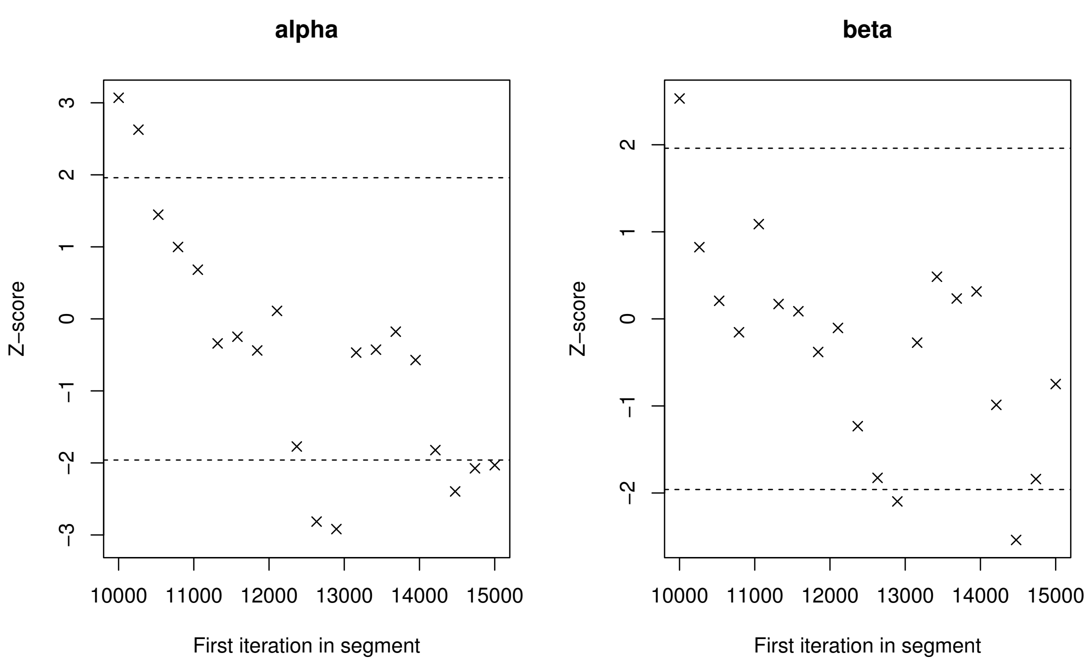

Figure 3.25: The Geweke plots for α and β when $n = 10$ using the divergence prior.

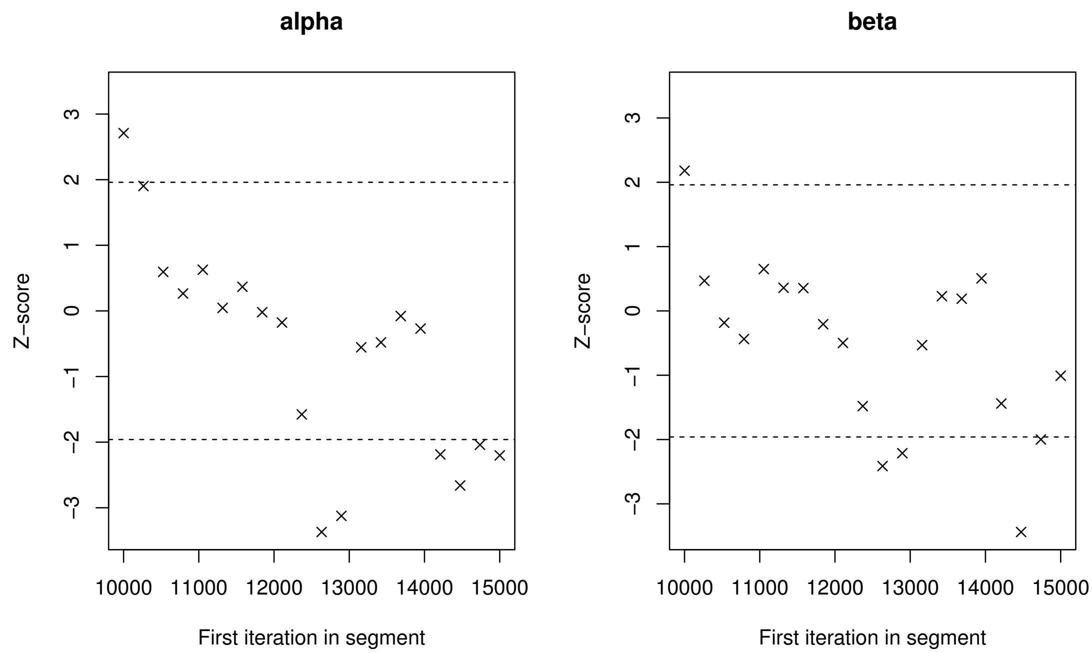

Figure 3.26: The Geweke plots for α and β when $n = 10$ using the reference prior.

Figures 3.27, 3.28, and 3.29 present the Geweke diagnostics for α and β for $n = 40$, when the Jeffreys prior, divergence prior, and reference prior are used, respectively.

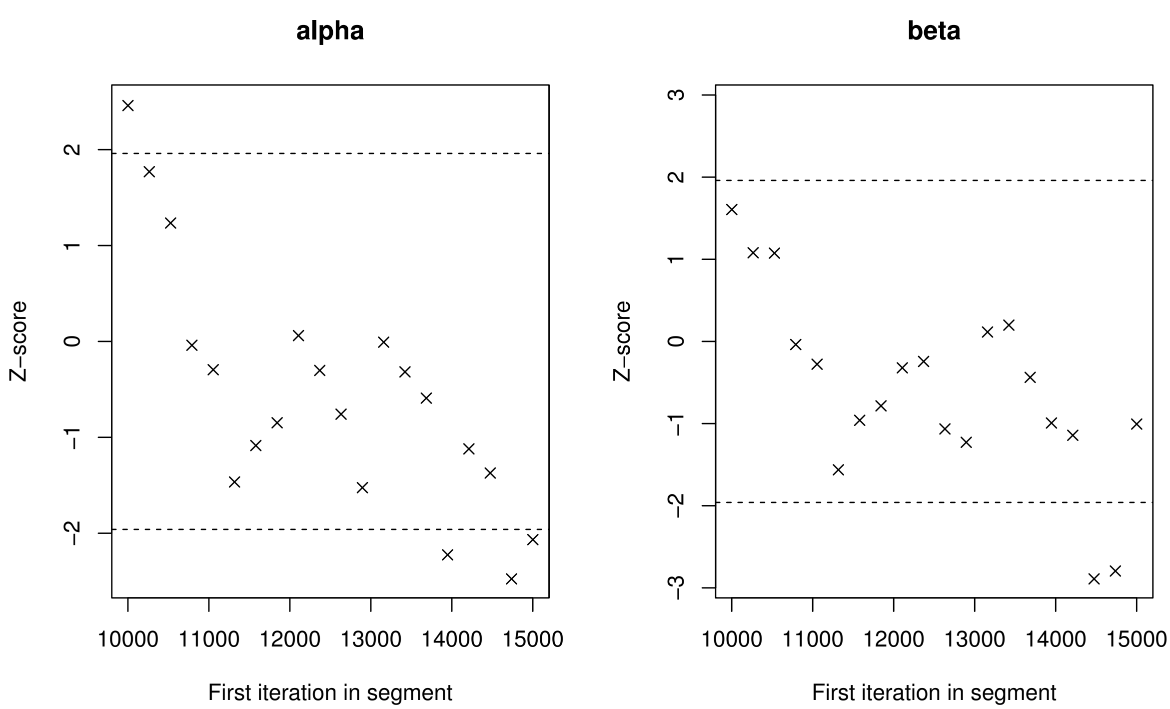

Figure 3.27: The Geweke plots for α and β when $n = 40$ using the Jeffreys prior.

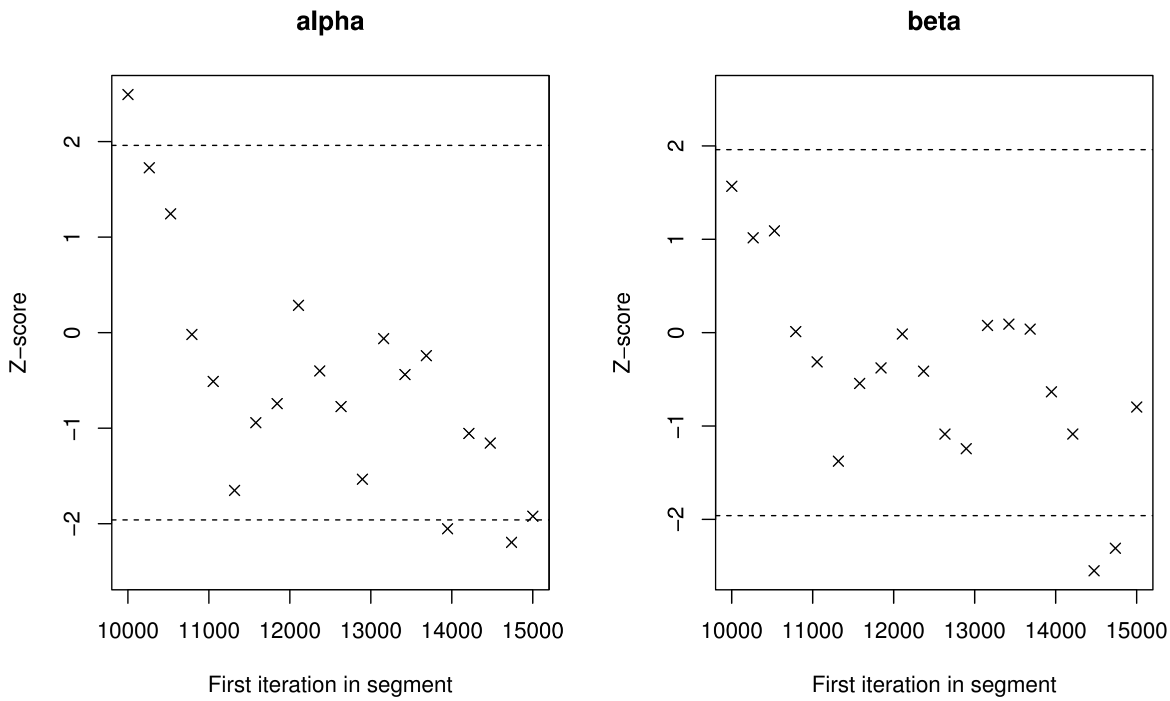

Figure 3.28: The Geweke plots for α and β when $n = 40$ using the divergence prior.

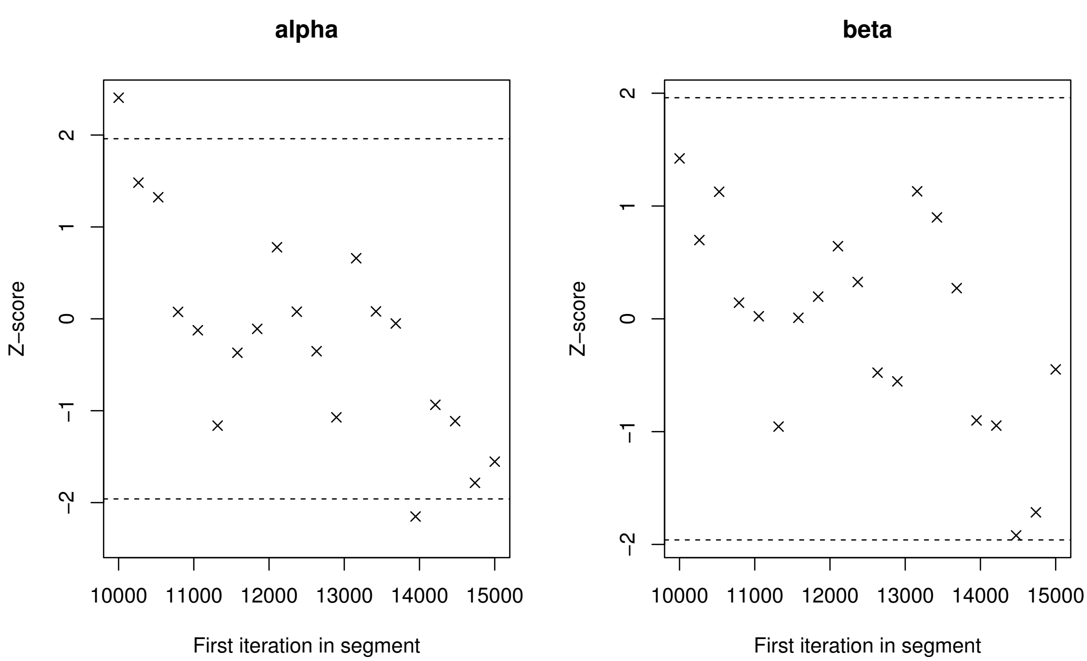

Figure 3.29: The Geweke plots for α and β when $n = 40$ using the reference prior.

Figures 3.30, 3.31, and 3.32 present the Geweke diagnostics for α and β for $n = 200$, when the Jeffreys prior, divergence prior, and reference prior are used, respectively.

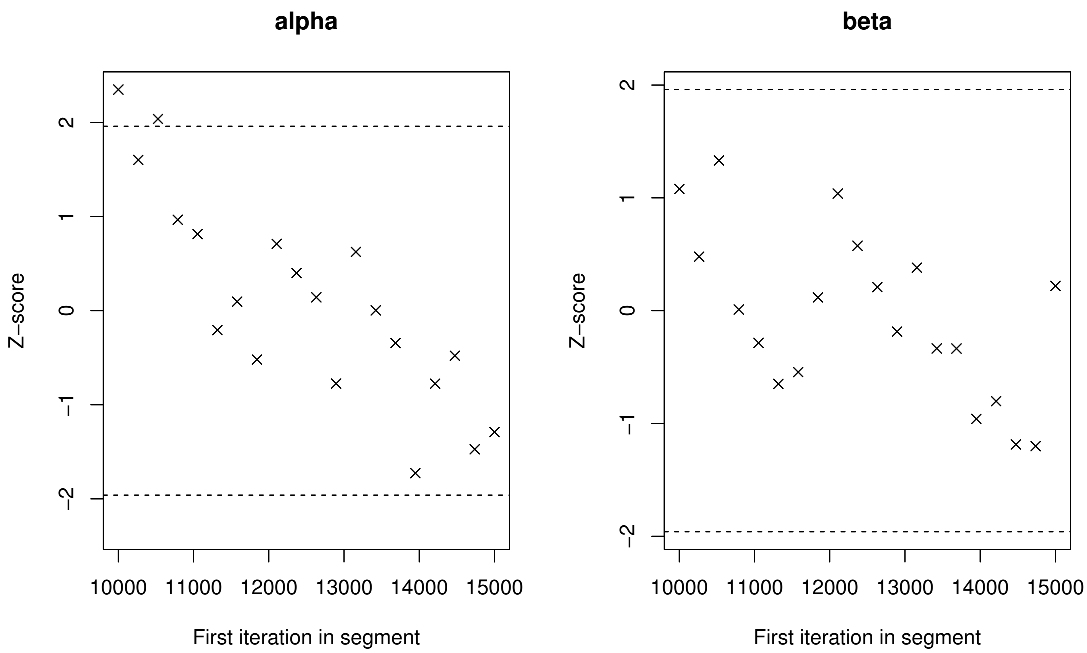

Figure 3.30: The Geweke plots for α and β when $n = 200$ using the Jeffreys prior.

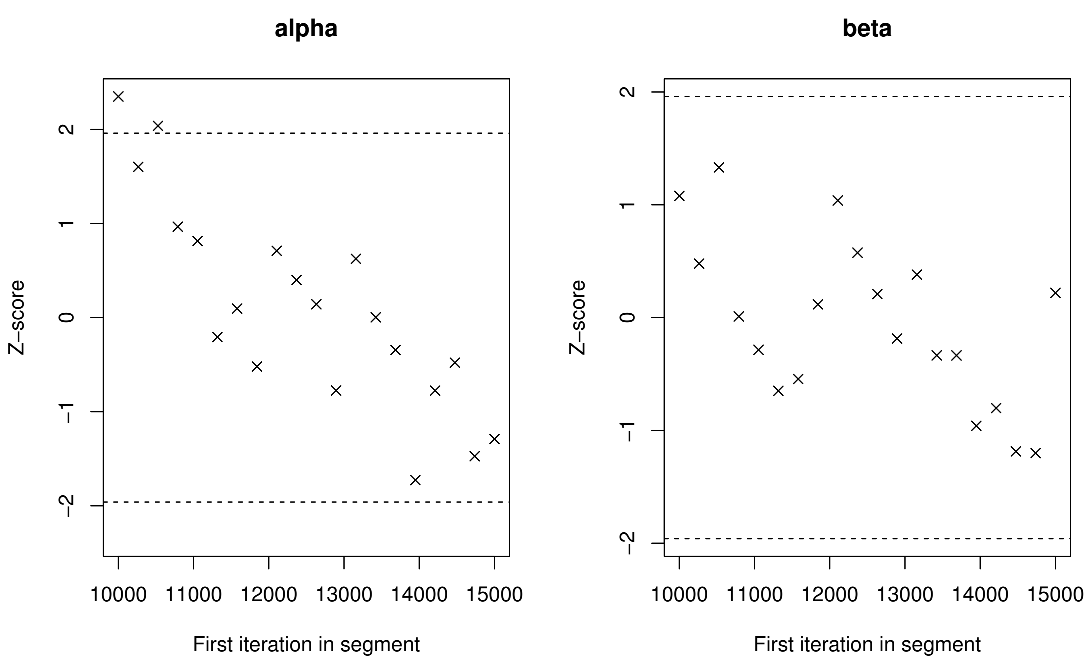

Figure 3.31: The Geweke plots for α and β when $n = 200$ using the divergence prior.

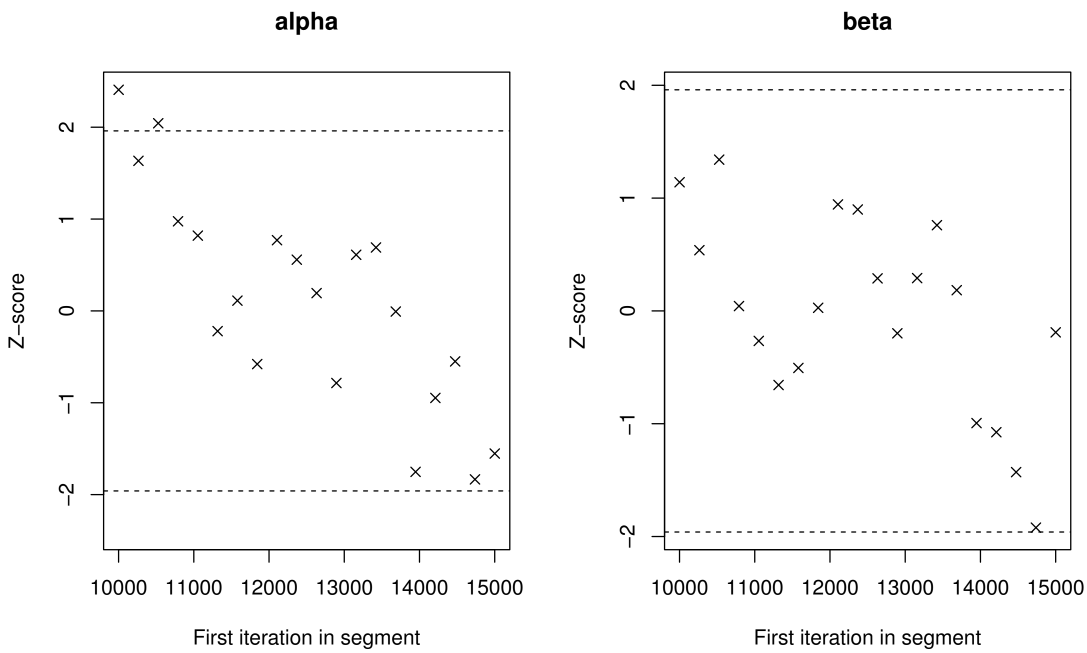

Figure 3.32: The Geweke plots for α and β when $n = 200$ using the reference prior.

For the Geweke diagnostic plots of α and β parameters, the results showed that as the sample size increased, the proportion of Z-scores falling within the $(-2,2)$ range also increased for all three priors used, indicating improved convergence. For smaller sample sizes, there were more instances of Z-score outside this range, indicating slower mixing. For a large sample size, the Z-scores were more in the acceptable bounds, reflecting strong evidence of convergence for both parameters.

3.5.2 Type I Right Censoring

In this subsection, the convergence of the MCMC chains for the Weibull model under type I right censored data will be assessed. The convergence will be evaluated using the trace plots, the Gelman-Rubin diagnostic plot, and the Geweke diagnostic. These diagnostics are applied to both shape and scale parameters.

Figures 3.33, 3.34, and 3.35 present the autocorrelation plots for various priors when $n = 10$, $r = 9$, $n = 40$, $r = 36$, and $n = 200$, $r = 175$, respectively. Under type I right censoring, a similar pattern is observed, when $n = 10$ and $r = 9$, the autocorrelation for β remains high across multiple lags, indicating slow mixing and strong dependence between samples. As the sample size increases or as the censoring threshold C allows more failures to be observed, the autocorrelation of β declines more rapidly toward zero, suggesting improved mixing and greater sample independence.

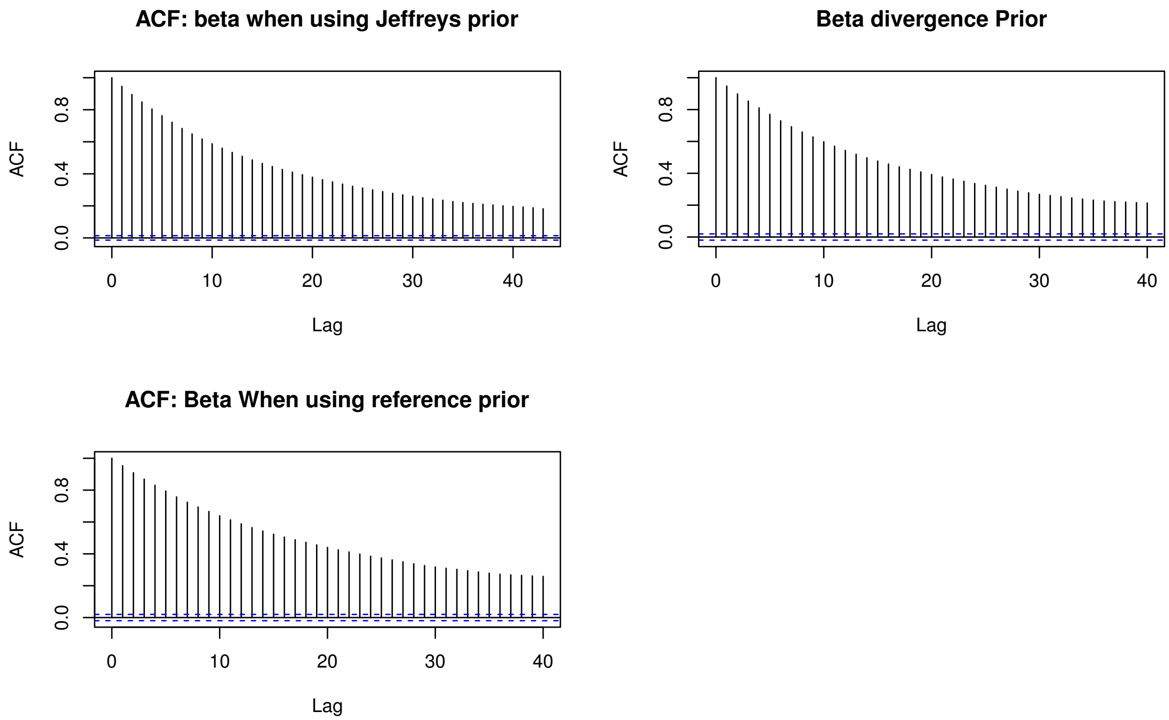

Figure 3.33: Autocorrelation for various priors when $n = 10$ and $r = 9$.

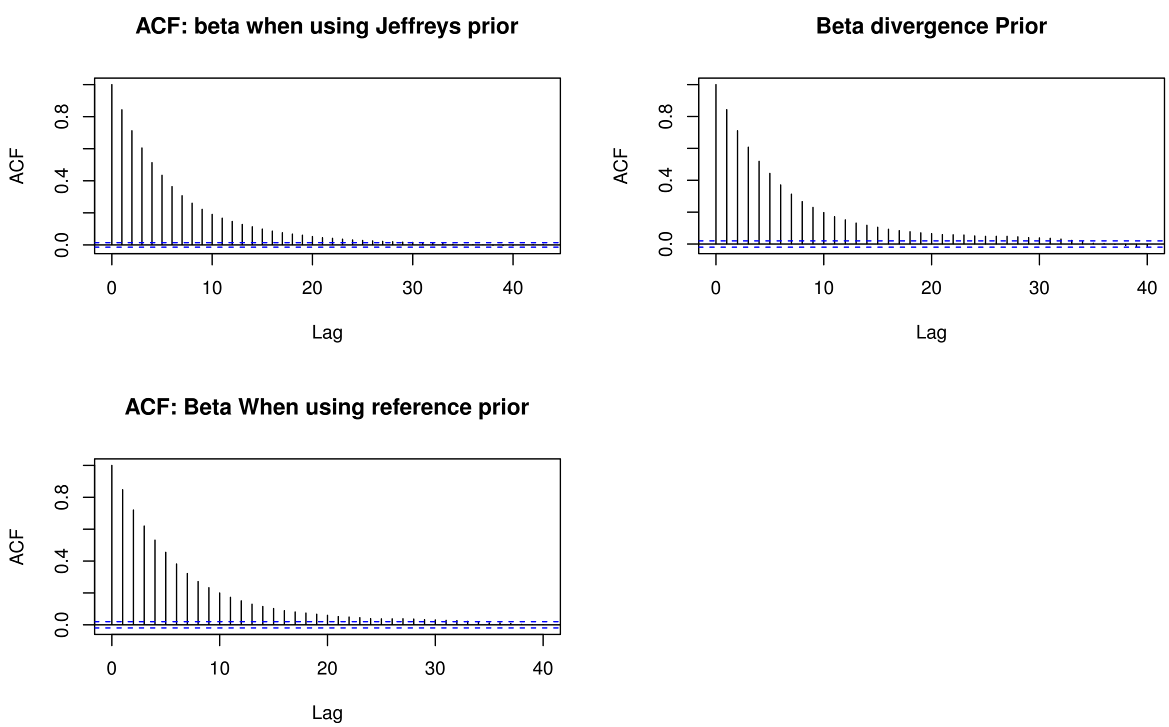

Figure 3.34: Autocorrelation for various priors when $n = 40$ and $r = 36$.

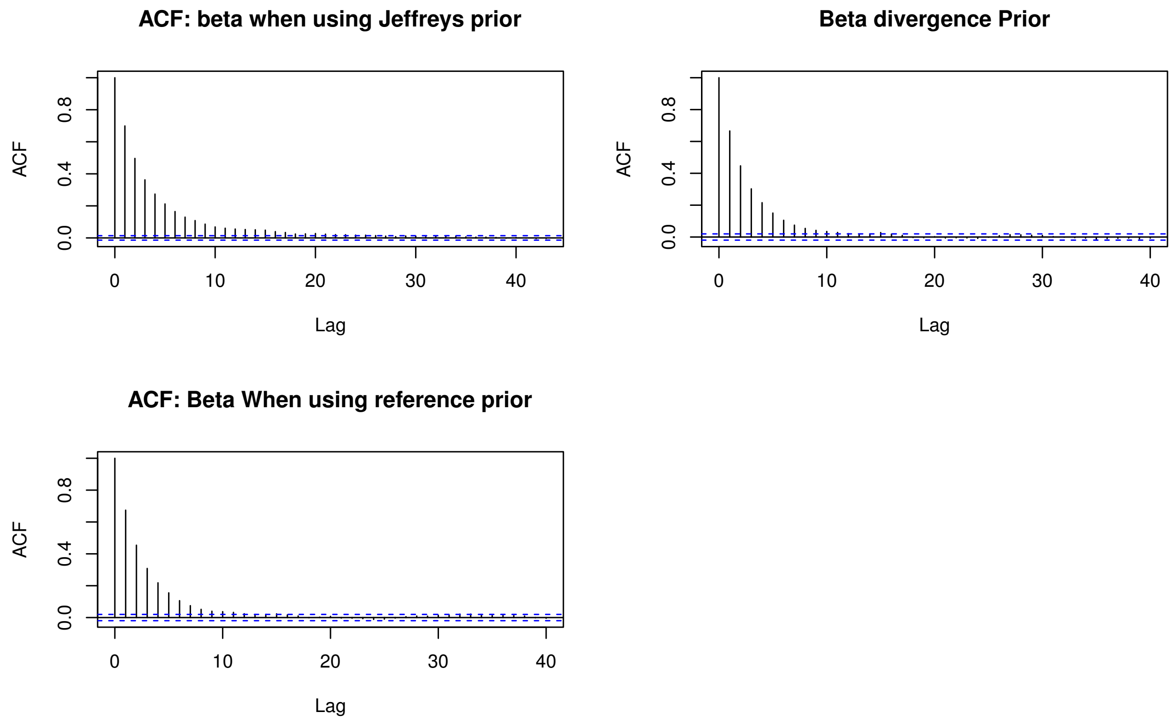

Figure 3.35: Autocorrelation for various priors when n = 200 and $r = 175$.

When there is type I right censoring, the same pattern holds. When $n = 10$ and $r = 9$, the autocorrelations for β, take more time to reach zero, indicating slow mixing and strong sample dependence. As the sample size increases or the censoring threshold C is high enough to observe more failures, the autocorrelation of beta decays more rapidly towards zero, indicating independence.

Figures 3.36, 3.37, and 3.38 present the trace plots for α and β parameters under a sample size $n = 10$ and failures $r = 9$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. For the case of type I right censoring, the trace plots for α and β under the three prior choices indicate that the chains for α converge quickly compared to those for β. Figures 3.39, 3.40, and 3.41 present the trace plots for α and β parameters under a sample size $n = 40$ and failures $r = 36$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. When $n = 40$ and the number of observed failures is 36, there are no noticeable differences that are observed in the trace plot patterns when the Jeffreys prior, divergence prior, or the reference prior is used, each yields similar chain behaviour for both parameters. The trace plots for both α and β become much tighter and show rapid mixing, the chains settle around their posterior means. The thick pen test is achieved for both chains of each posterior, meaning the posterior distribution has converged.

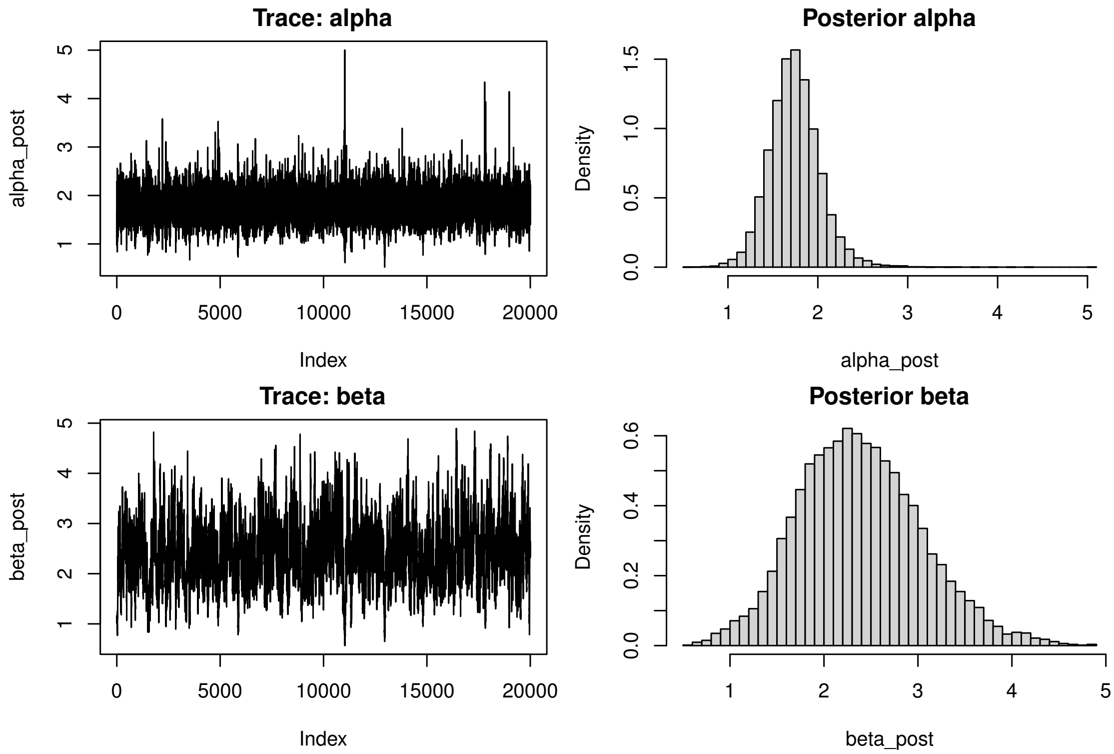

Figure 3.36: Trace plot and probability density function of α and β when n = 10 and $r = 9$ using Jeffreys prior.

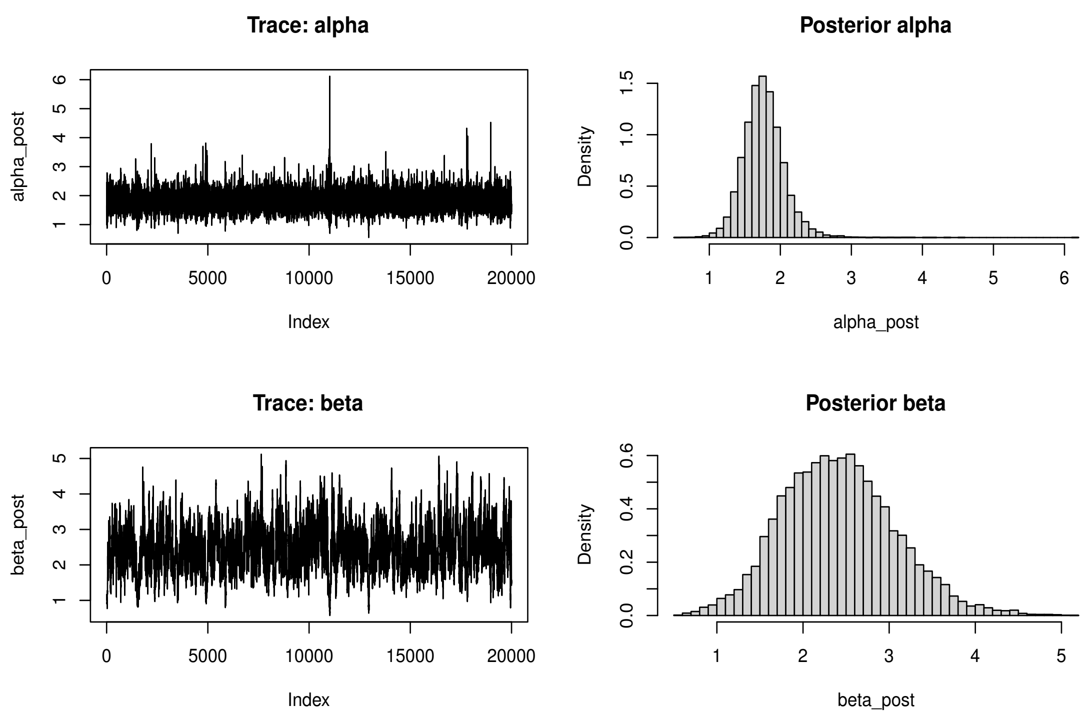

Figure 3.37: Trace plot and probability density function of α and β when n = 10 and $r = 9$ using divergence prior.

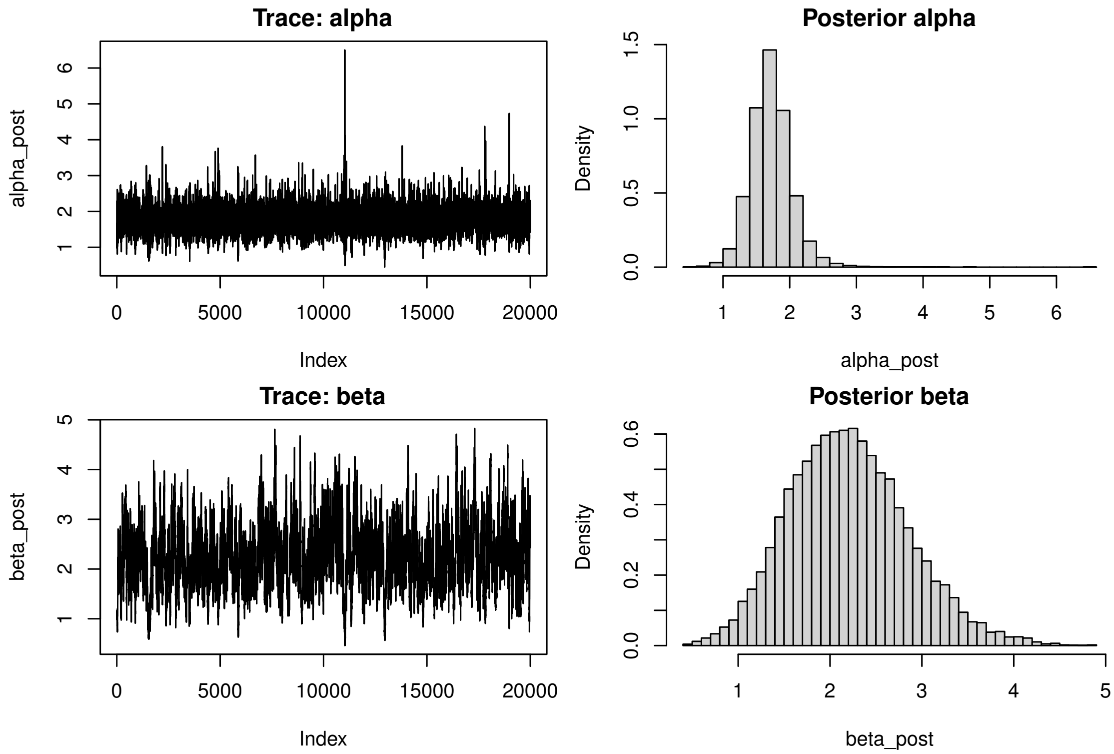

Figure 3.38: Trace plot and probability density function of α and β when $n = 10$ and $r = 9$ using reference prior.

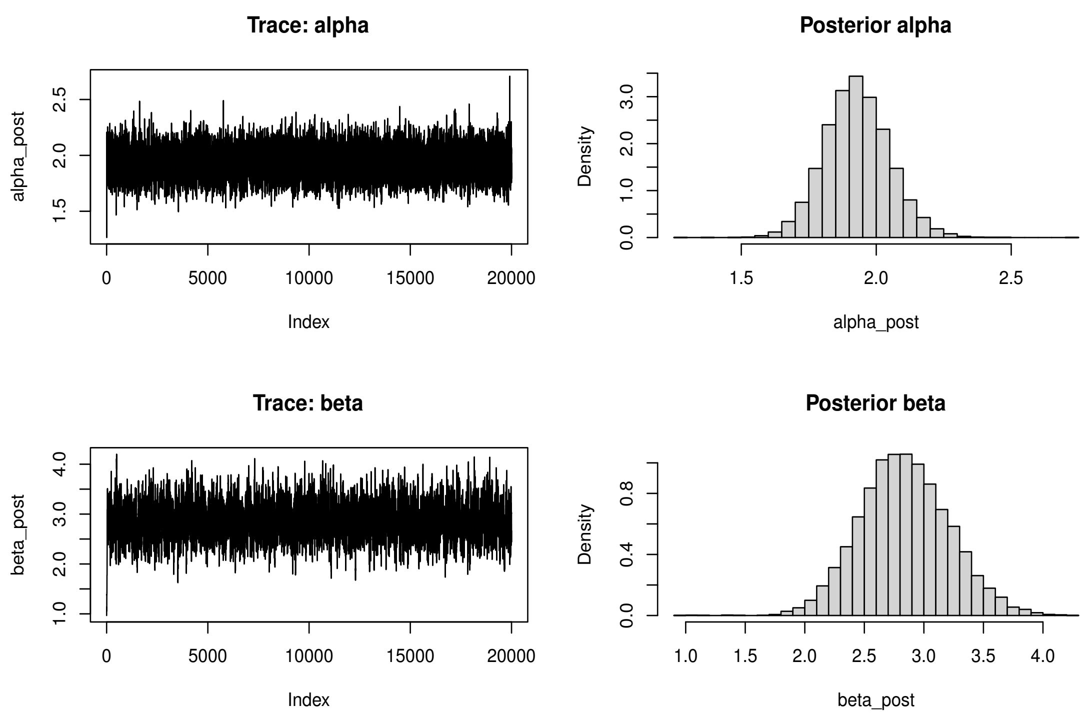

Figure 3.39: Trace plot and probability density function of α and β when $n = 40$ and $r = 36$ using the Jeffreys prior.

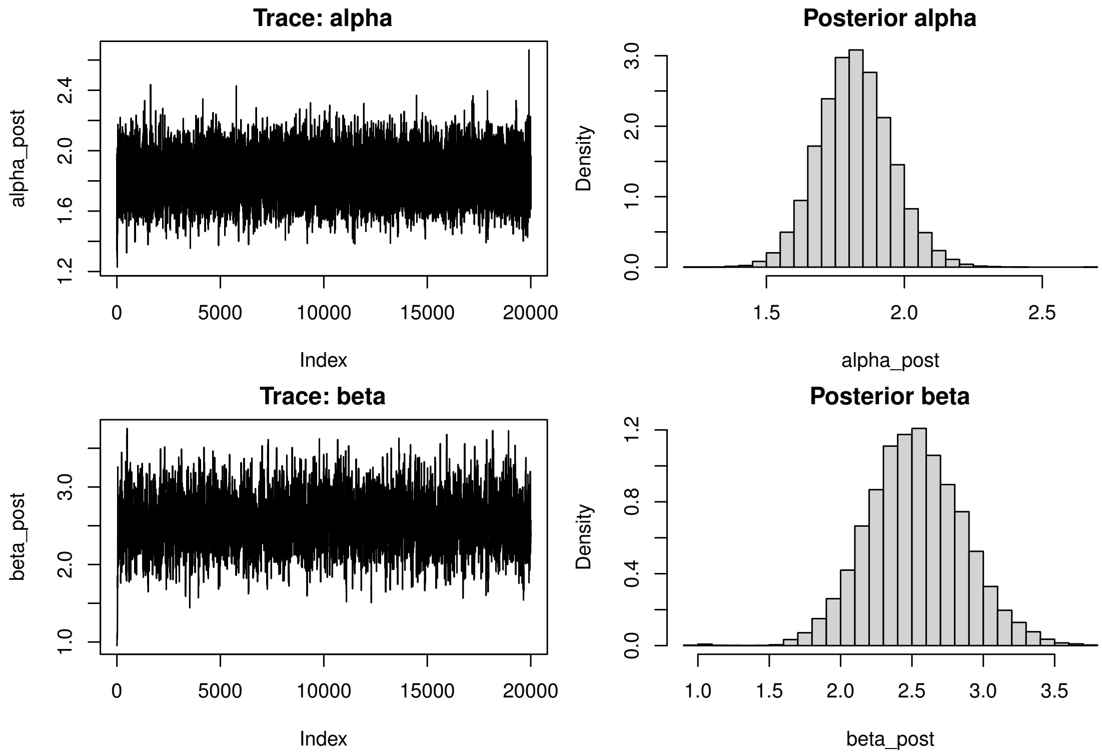

Figure 3.40: Trace plot and probability density function of α and β when $n = 40$ and $r = 36$ using the divergence prior.

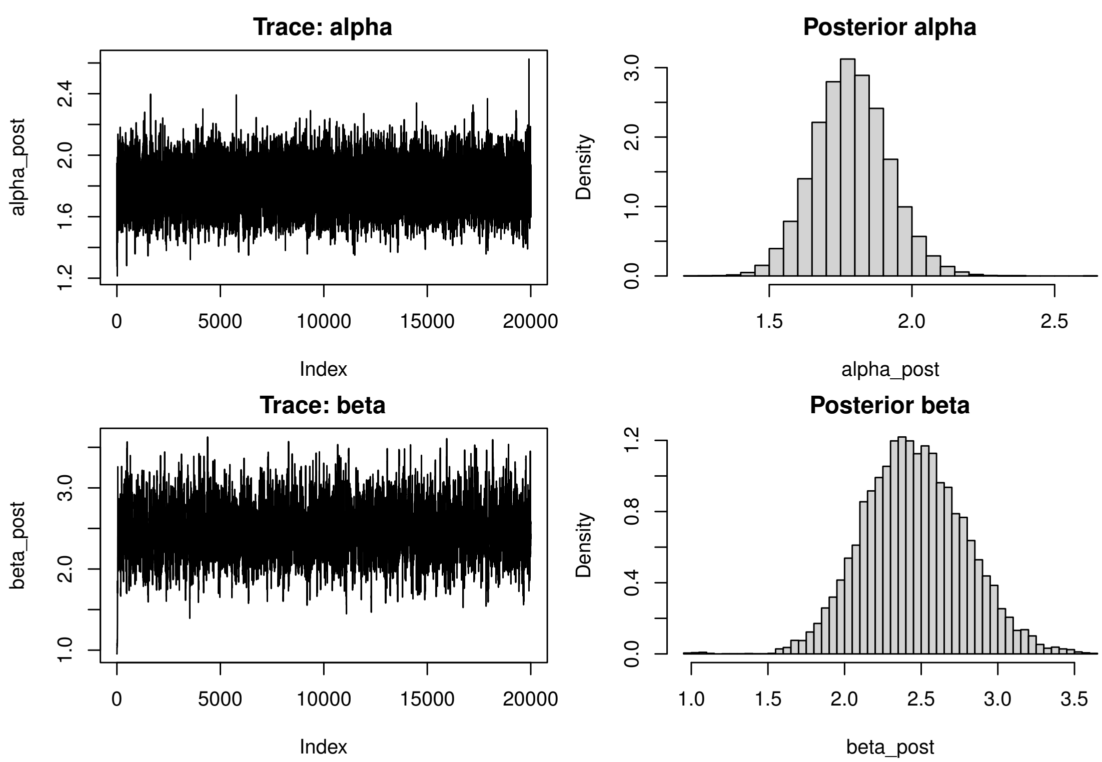

Figure 3.41: Trace plot and probability density function of α and β when $n = 40$ and $r = 36$ using the reference prior.

Figures 3.42, 3.43, and 3.44 present the trace plots for α and β parameters under a sample size $n = 200$ and failures $r = 175$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively.

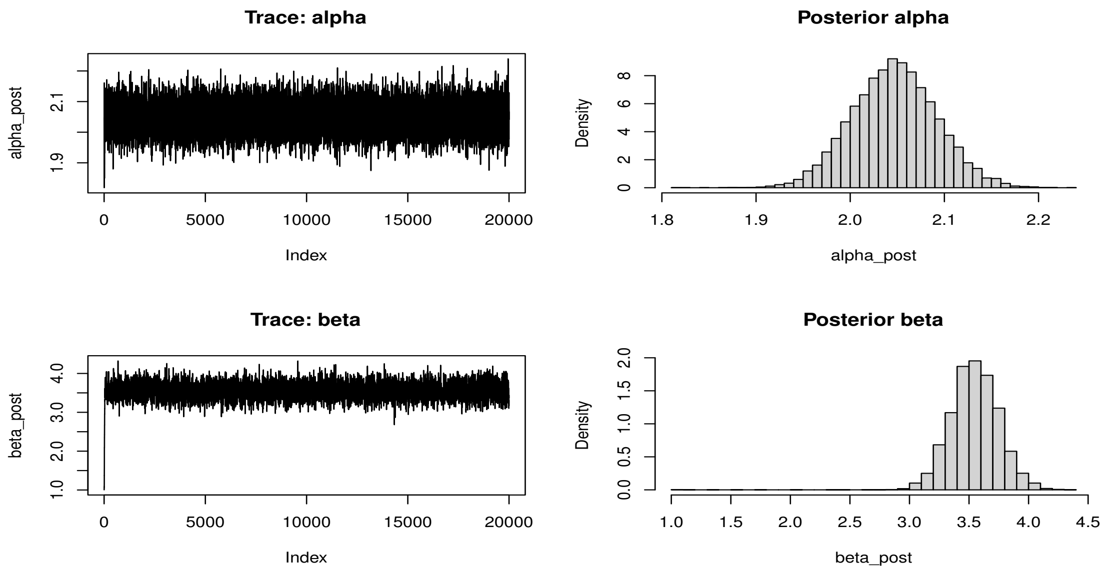

Figure 3.42: Trace plot and probability density function of α and β when $n = 200$ and $r = 175$ using the Jeffreys prior.

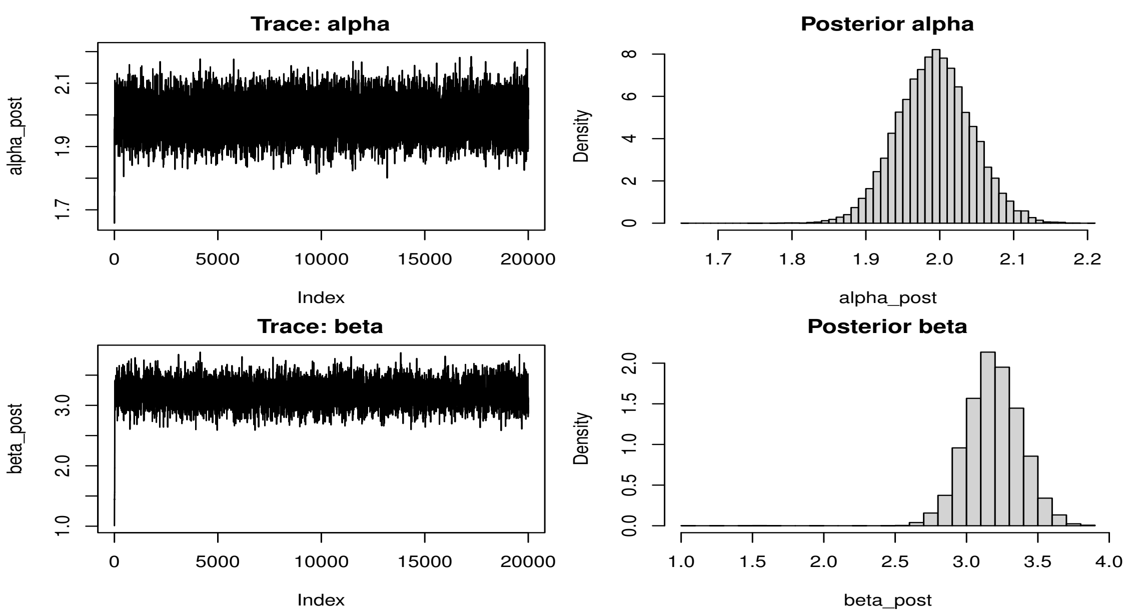

Figure 3.43: Trace plot and probability density function of α and β when $n = 200$ and $r = 175$ using the divergence prior.

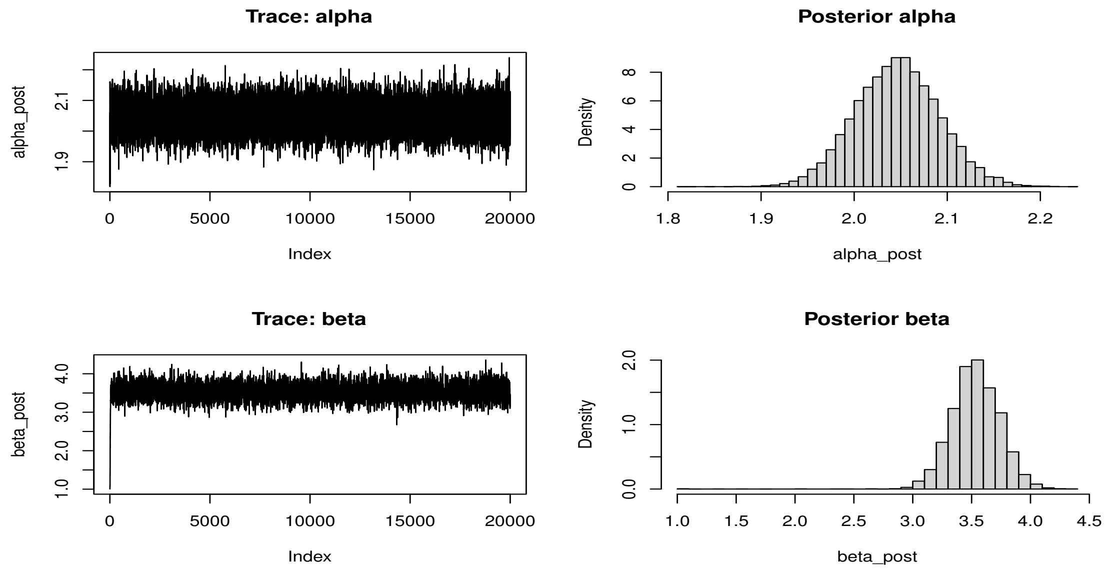

Figure 3.44: Trace plot and probability density function of α and β when n = 200 and $r = 175$ using the reference prior.

When $n = 200$ and $r = 175$, under type I right censoring, the trace plots for both α and β show excellent mixing, with chains fluctuating strongly around their posterior means. The thick pen test is achieved for each parameter, confirming that the distributions being sampled have converged.

Figures 3.45, 3.46, and 3.47 present the Gelman-Rubin plots for α and β when n = 10 and $r = 9$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. For the sample size $n = 10$ and $r = 9$, the shrink factor values approach 1 slowly, and more iterations are required before convergence is achieved. The shrink factor values for β approach 1 faster than α across all priors and sample sizes used in this thesis.

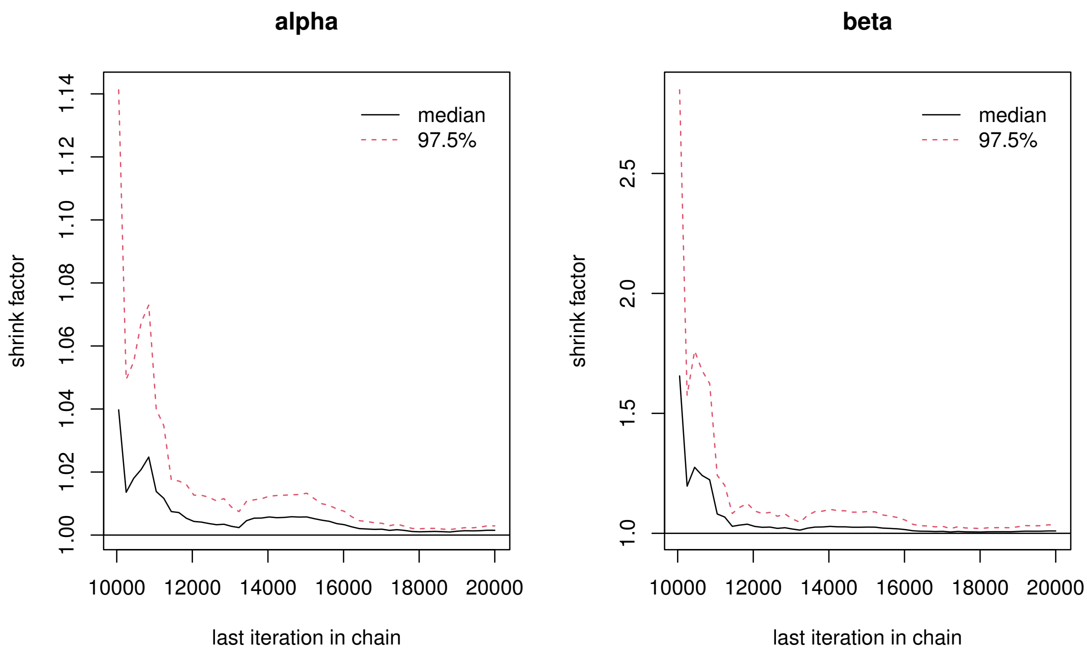

Figure 3.45: The Gelman-Rubin plots for α and β when $n = 10$ and $r = 9$ using the Jeffreys prior.

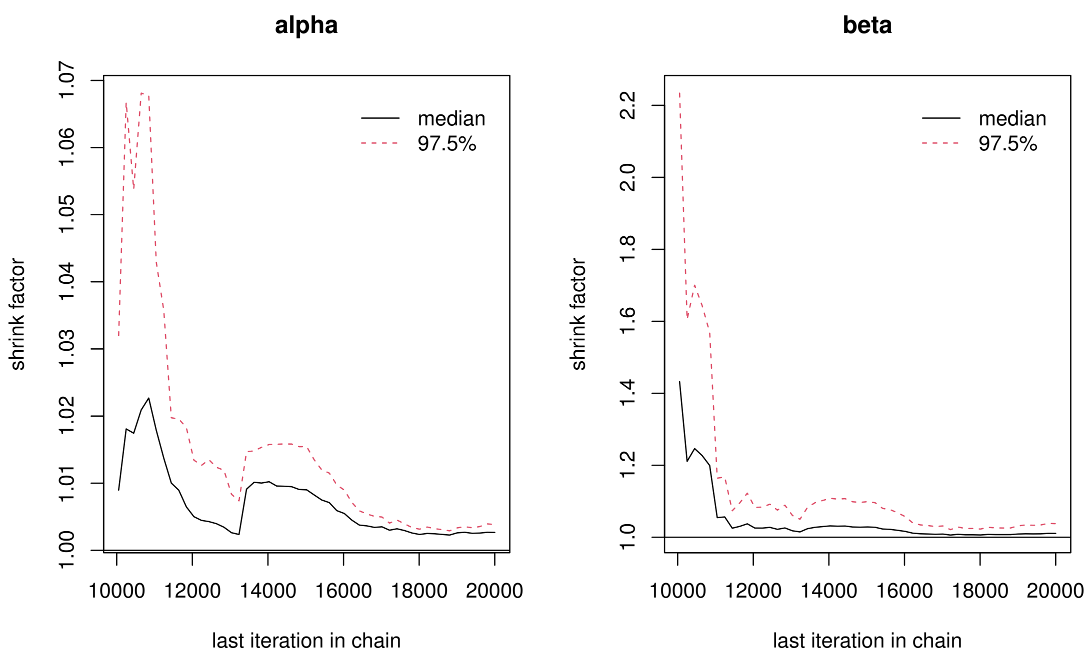

Figure 3.46: The Gelman-Rubin plots for α and β when $n = 10$ and $r = 9$ using the divergence prior.

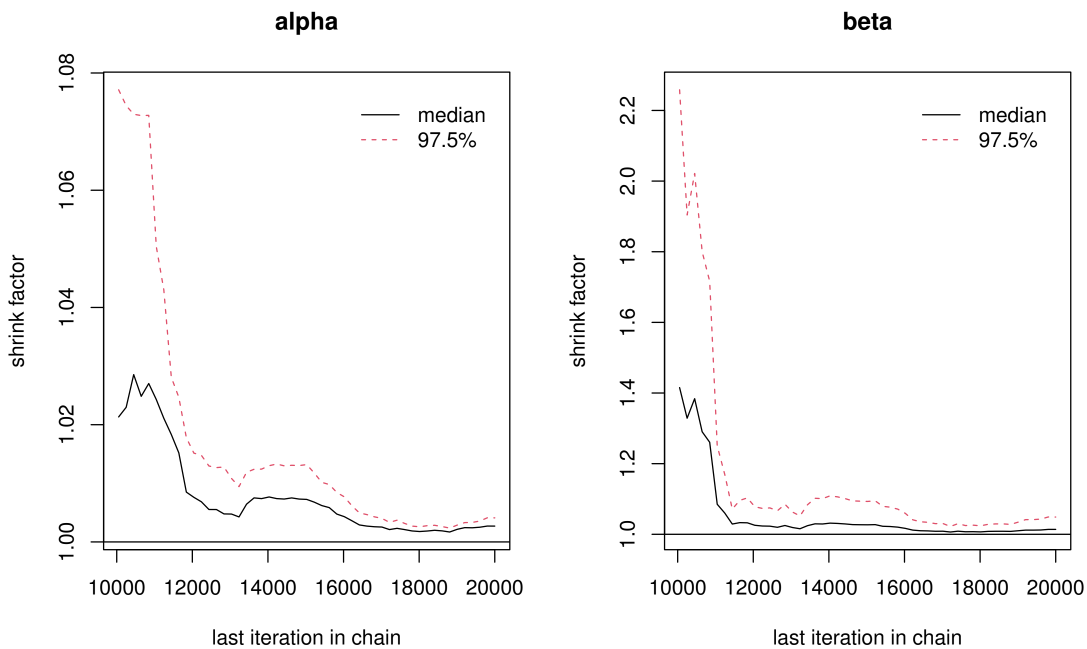

Figure 3.47: The Gelman-Rubin plots for α and β when $n = 10$ and $r = 9$ using the reference prior.

Figures 3.48, 3.49, and 3.50 present the Gelman-Rubin plots for α and β when $n = 40$ and $r = 36$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. For the sample size 40 and $r = 36$ failures, the shrink factor values approach 1 faster, indicating convergence.

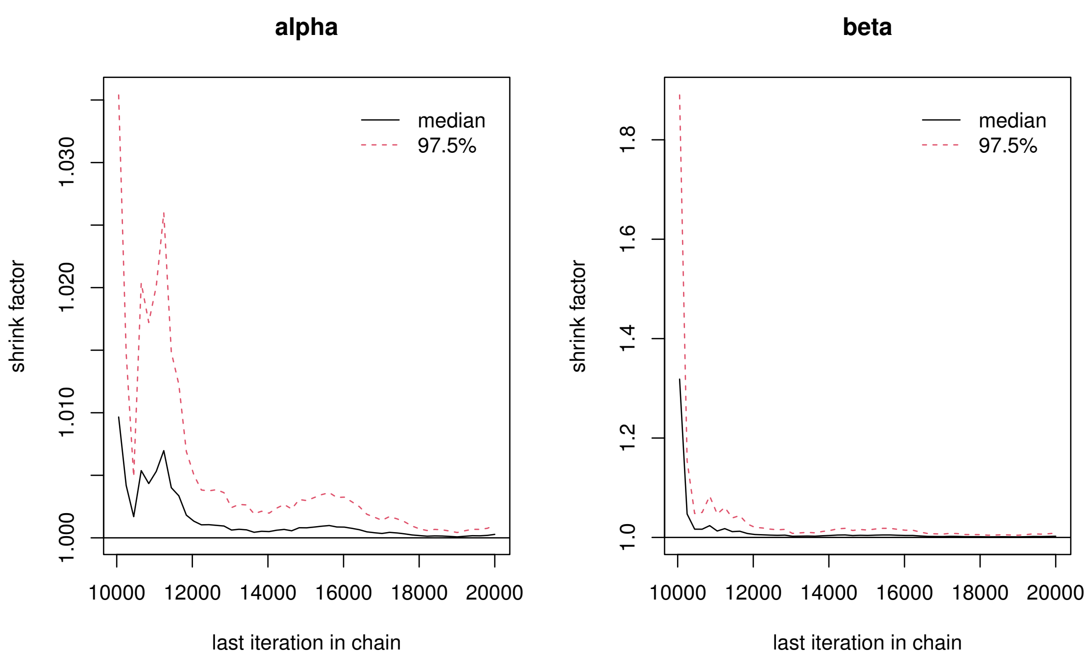

Figure 3.48: The Gelman-Rubin plots for α and β when $n = 40$ and $r = 36$ using the Jeffreys prior.

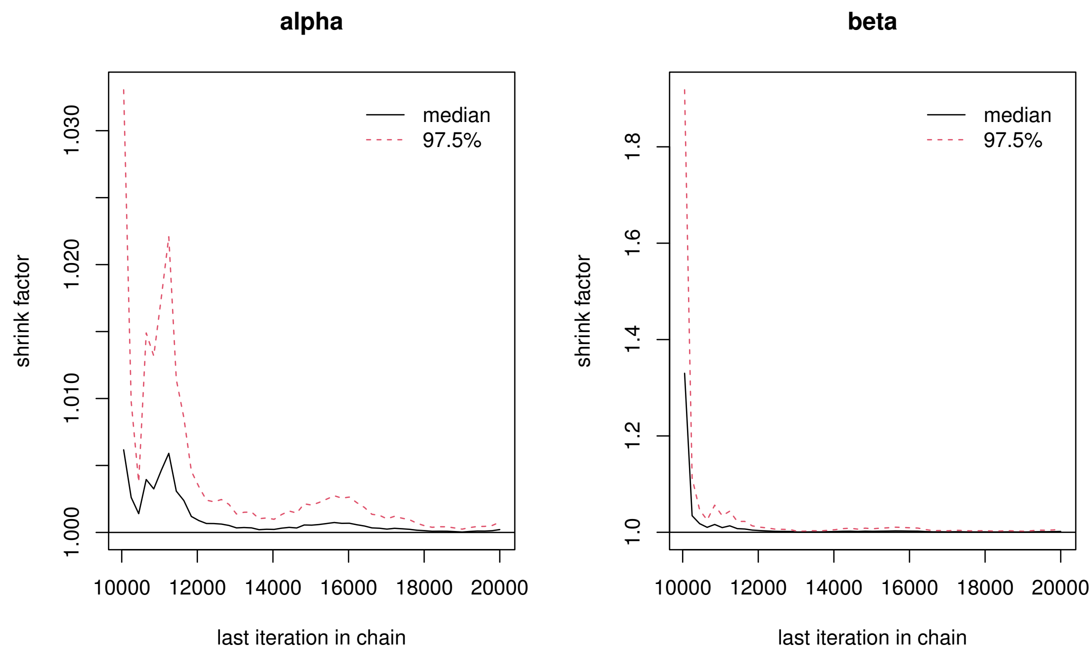

Figure 3.49: The Gelman-Rubin plots for α and β when $n = 40$ and $r = 36$ using the divergence prior.

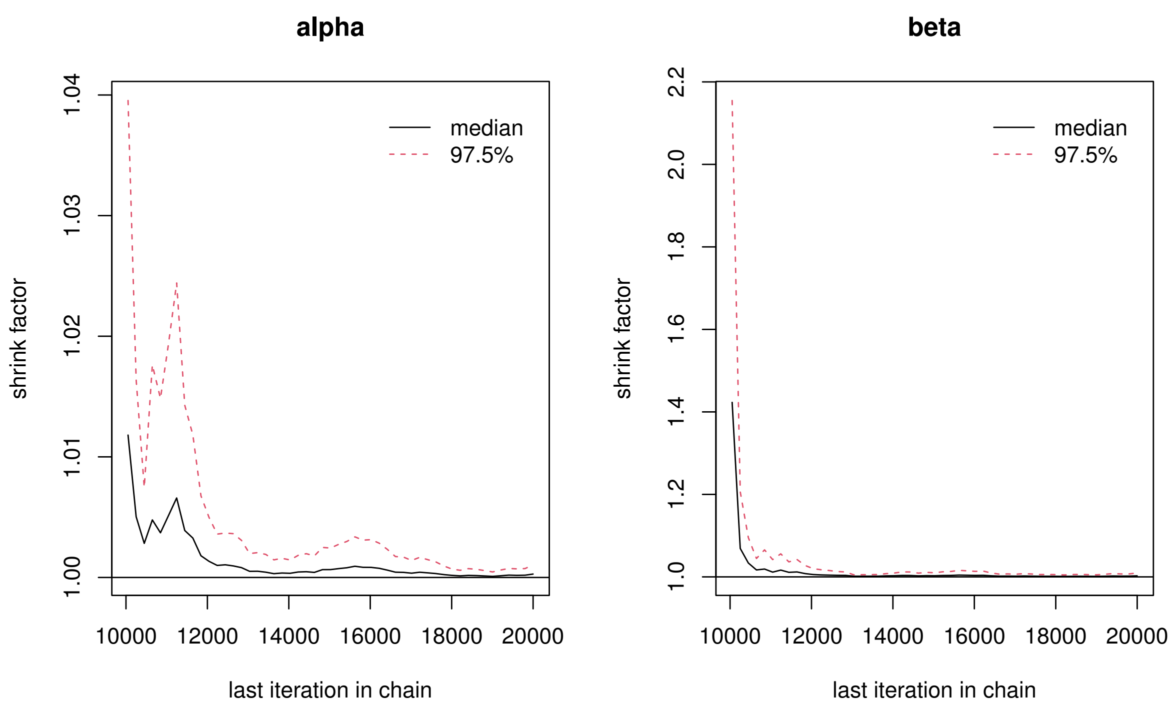

Figure 3.50: The Gelman-Rubin plots for α and β when $n = 40$ and $r = 36$ using the reference prior.

Figures 3.51, 3.52, and 3.53 present the Gelman-Rubin plots for α and β when $n = 200$ and $r = 175$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively.

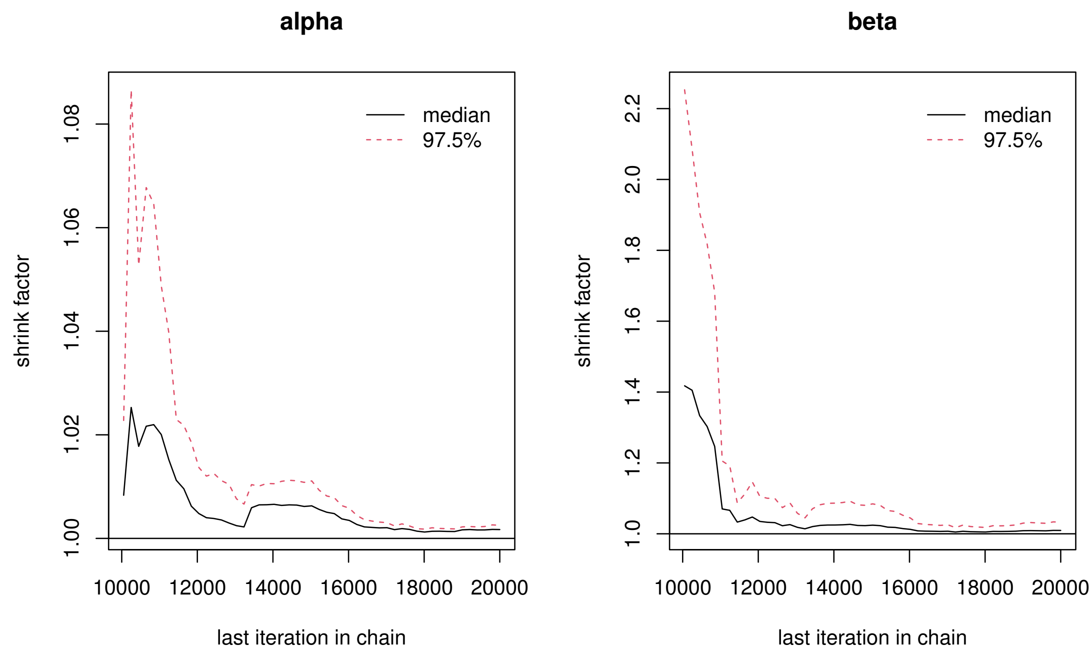

Figure 3.51: The Gelman-Rubin plots for α and β when $n = 200$ and $r = 175$ using the Jeffreys prior.

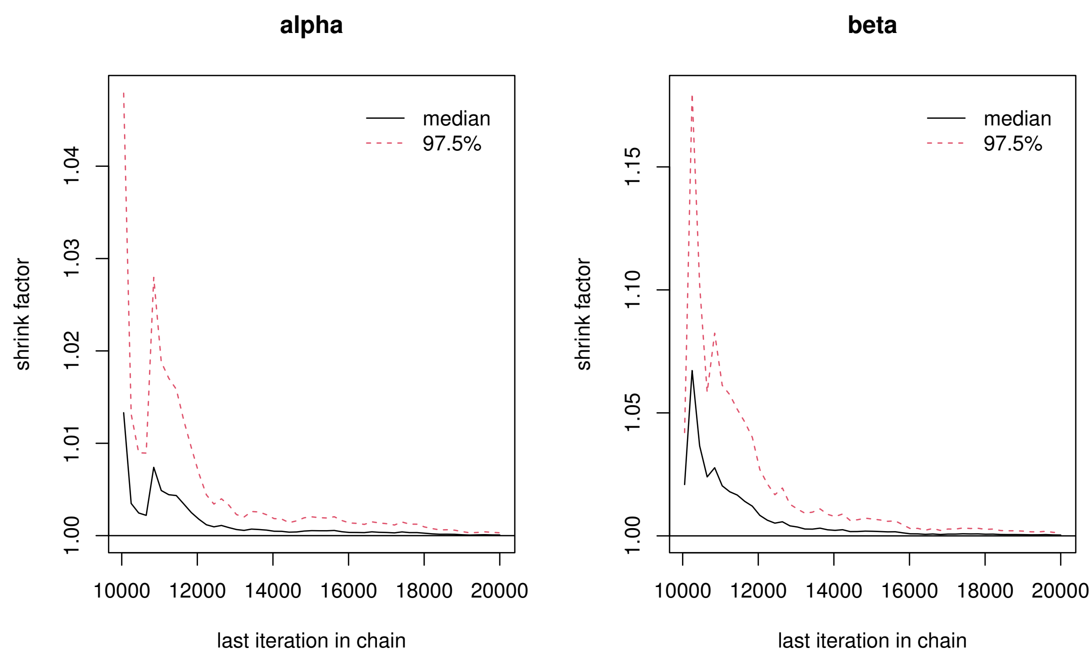

Figure 3.52: The Gelman-Rubin plots for α and β when $n = 200$ and $r = 175$ using the divergence prior.

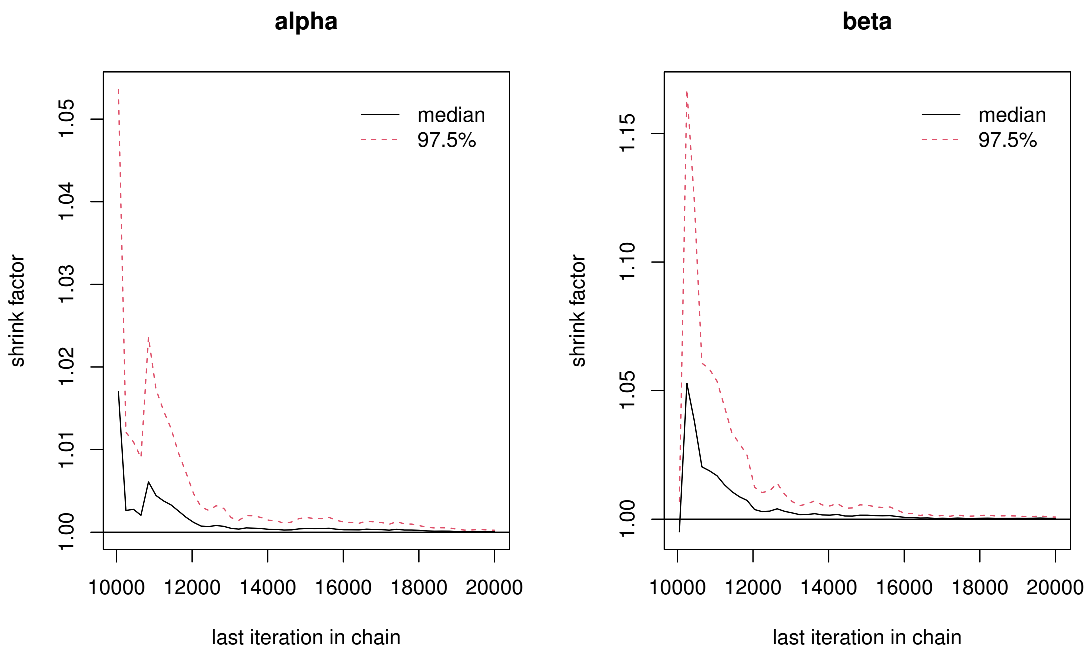

Figure 3.53: The Gelman-Rubin plots for α and β when $n = 200$ and $r = 175$ using the reference prior.

For the sample size 200 and $r = 175$ failures, convergence is achieved faster, with shrink factor values equal to 1 after some few iterations. Across all sample size, there is no noticeble difference in the Gelman-Rubin diagnostic plots for both parameters when comparing the Jeffreys prior, divergence prior, and the reference prior effect.

Figures 3.54, 3.55, and 3.56 present the Geweke plots for α and β when $n = 10$ and $r = 9$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. Figures 3.57, 3.58, and 3.59 present the Geweke plots for α and β when $n = 40$ and $r = 36$, using the Jeffreys prior, divergence prior, and the reference prior, respectively. Figures 3.60, 3.61, and 3.62 present the Geweke plots for α and β when $n = 200$ and $r = 175$, using the Jeffreys prior, divergence prior, and the reference prior, respectively.

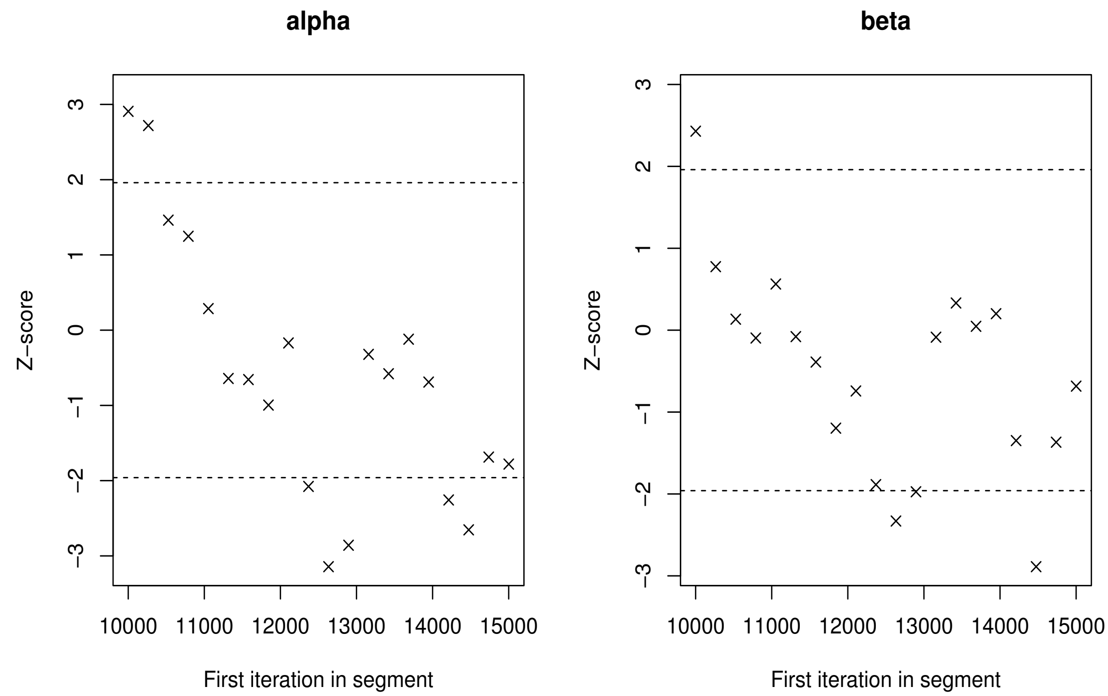

Figure 3.54: The Geweke plots for α and β when $n = 10$ and $r = 9$ using the Jeffreys prior.

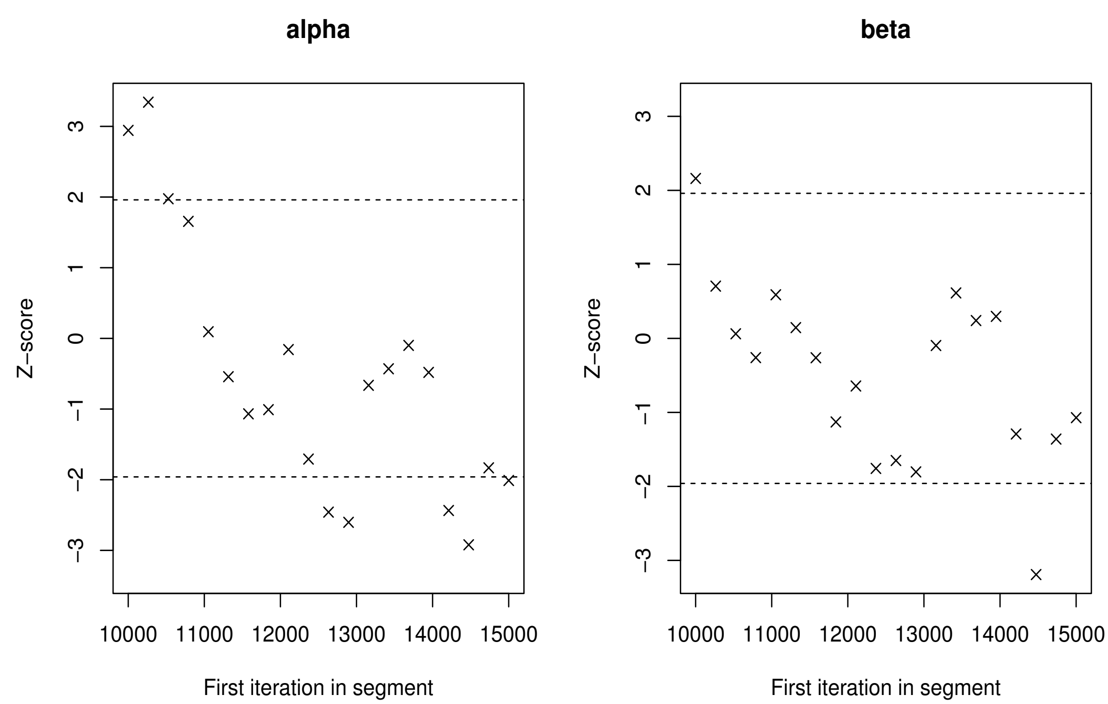

Figure 3.55: The Geweke plots for α and β when $n = 10$ and $r = 9$ using the divergence prior.

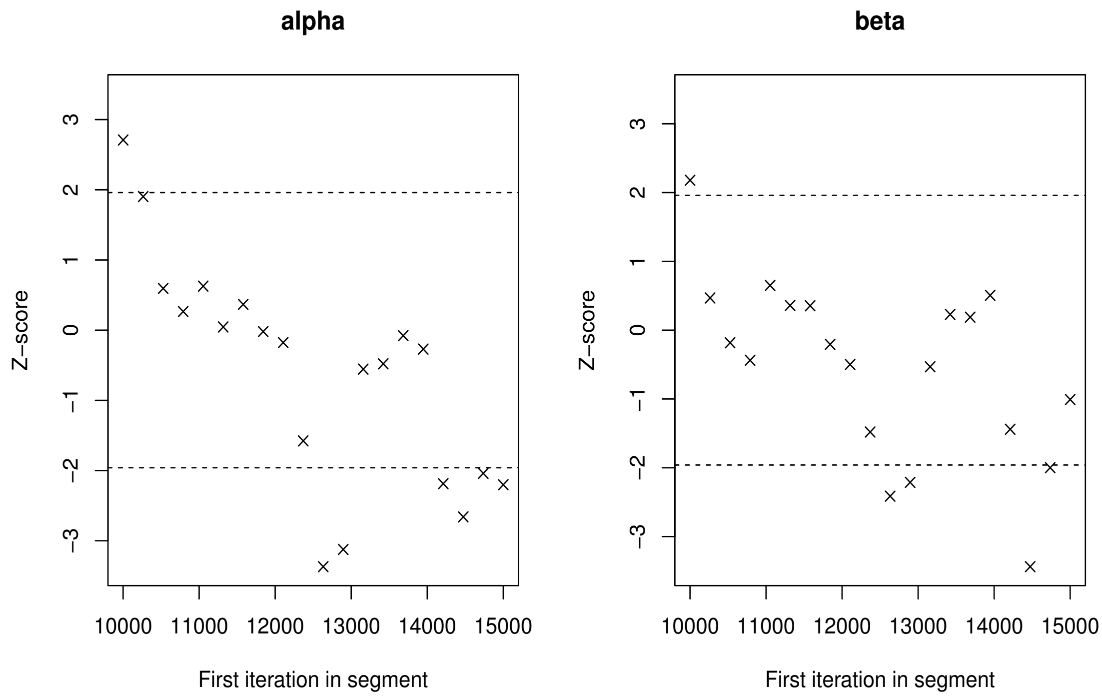

Figure 3.56: The Geweke plots for α and β when $n = 10$ and $r = 9$ using the reference prior.

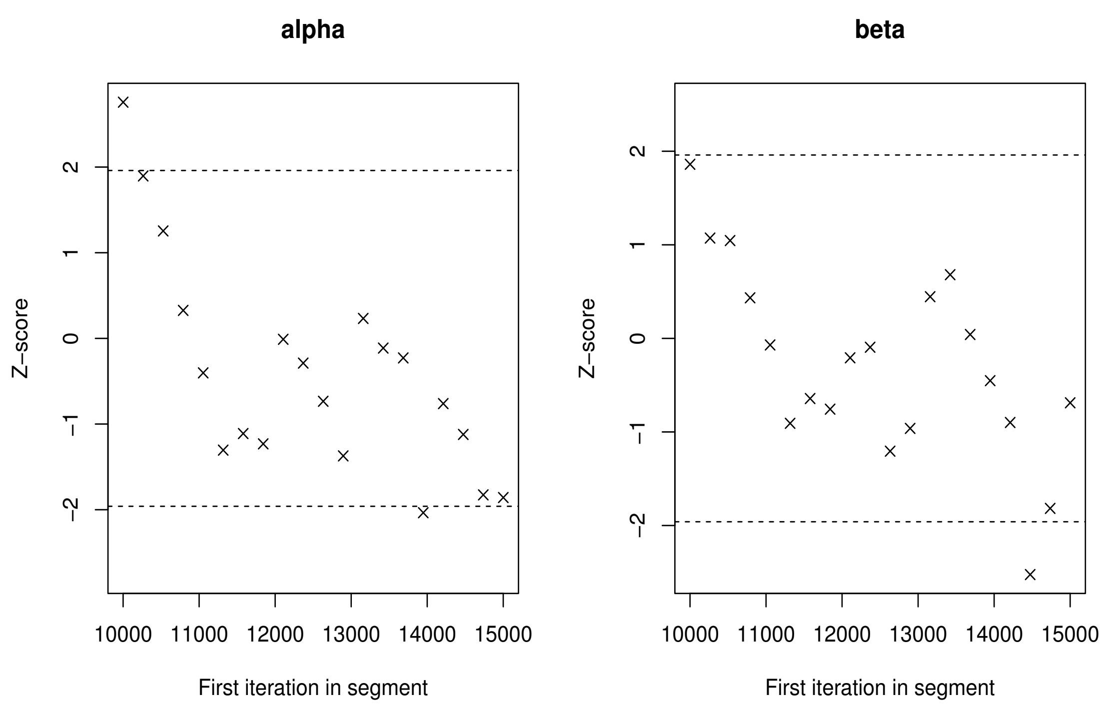

Figure 3.57: The Geweke plots for α and β when $n = 40$ and $r = 36$ using the Jeffreys prior.

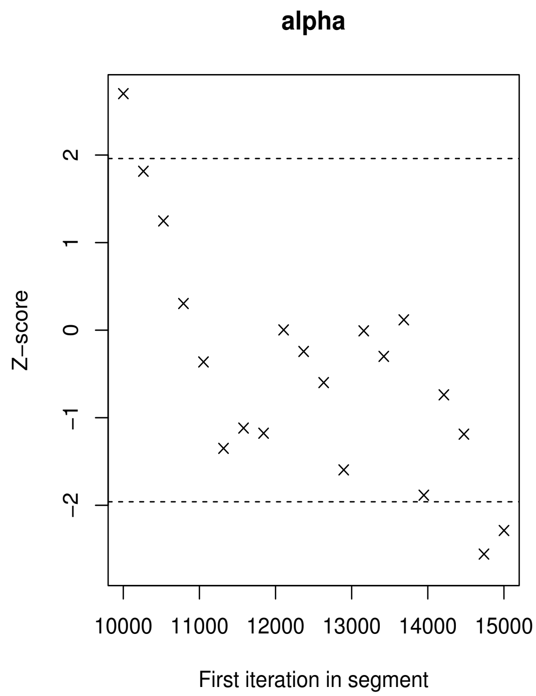

Figure 3.58: The Geweke plots for α and β when n=40 using the divergence prior.

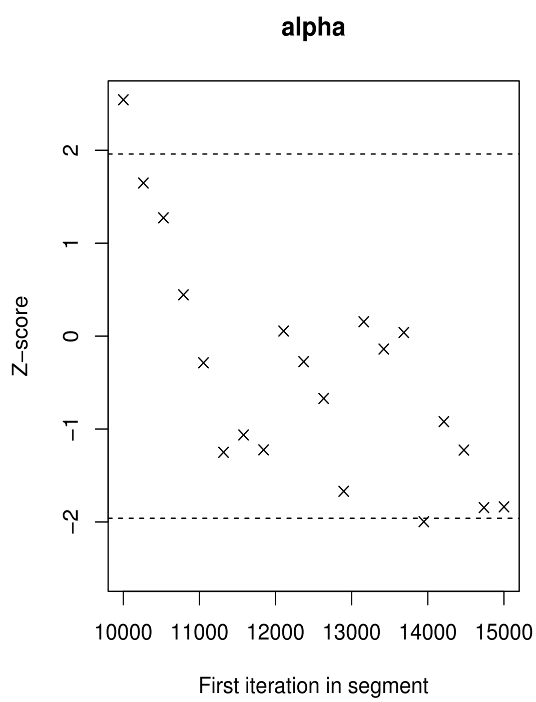

Figure 3.59: The Geweke plots for α and β when n=40 using the reference prior.

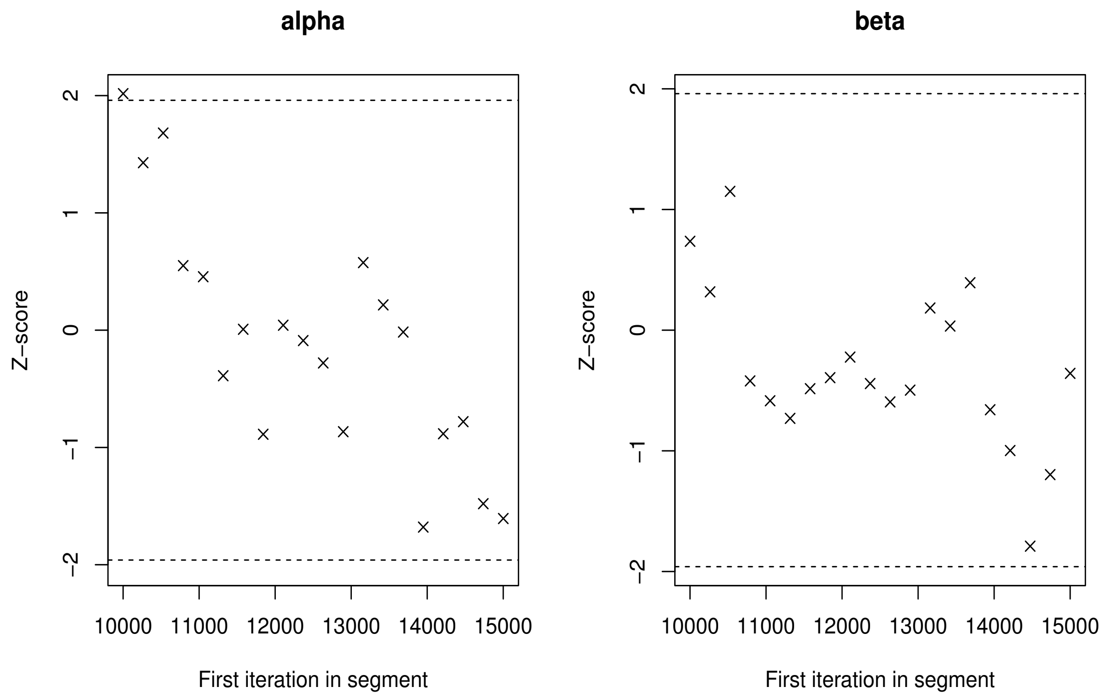

Figure 3.60: The Geweke plots for α and β when $n = 200$ and $r = 175$ using the Jeffreys prior.

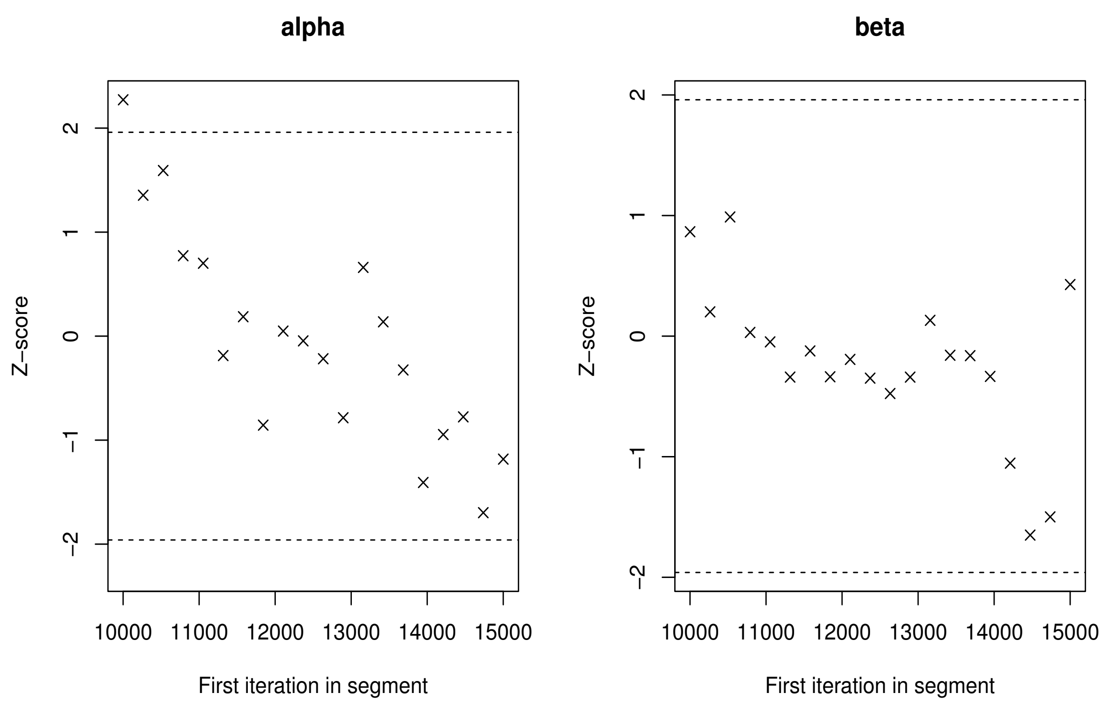

Figure 3.61: The Geweke plots for α and β when $n = 200$ and $r = 175$ using the divergence prior.

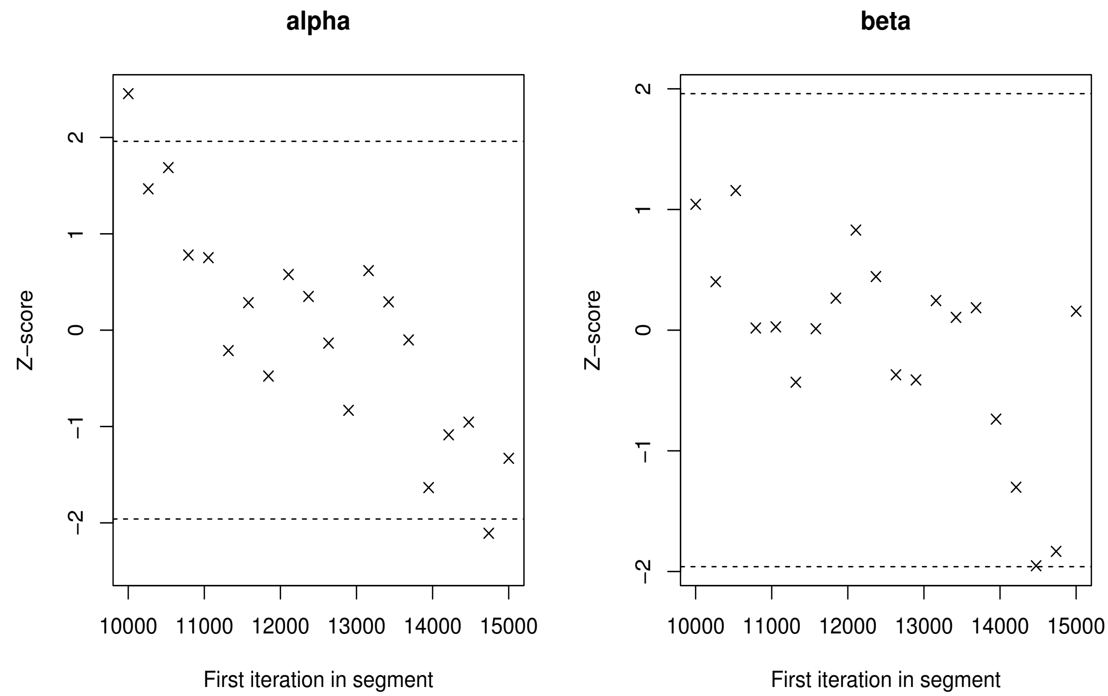

Figure 3.62: The Geweke plots for α and β when n = 200 and $r = 175$ using the reference prior.

When there is type I right censoring, the Geweke diagnostic results follow a similar trend, but the presence of censoring slightly delays convergence due to reduced information from censored observations. For $n = 10$ and $r = 9$, the high proportion of Z-scores for both α and β falling outside the acceptable range $(-2,2)$, indicating slower mixing across all three priors specified. As the sample size or observed failures increase, the proportion of Z-score falls outside the range $(-2,2)$ decreases noticeably. This pattern reflects clear convergence, as larger sample sizes yields most Geweke diagnostic Z-scores for both parameters within the acceptable bounds, suggesting satisfactory mixing and convergence.

3.6 Simulation Study for the Weibull Distribution

This section presents the simulation study conducted to evaluate the performance of Bayesian inference for the Weibull distribution under both complete data and type I right censoring scenarios. For each simulation, data were generated by sampling t from the Weibull distribution with specified true parameters, with scale parameter $\alpha \in \{0.5, 2, 5\}$, and shape parameter, $\beta \in \{0.8, 5, 8\}$. The focus of the simulation study is on evaluating the performance of different priors in terms of coverage rates and mean interval lengths for the Weibull parameters, which serve as the primary criteria for assessing efficiency and accuracy under different sample sizes and censoring conditions. The selected sample sizes are commonly used in simulation studies to represent small, moderate, and large samples, allowing comparison with existing reliability and Bayesian inference studies in the literature.

3.6.1 Complete Data

In this subsection, the performance of the Jeffreys prior, divergence prior, and the reference prior will be compared for different sample sizes, for the chosen sample sizes of $n = 10, 40$, and 200. Their performance will be evaluated using the coverage rate and the mean interval length. A model with the shortest mean interval length and coverage very close to 95% also known as the nominal level, is considered to be the best. For each simulated true parameter values (α, β), the coverage probabilities were recorded as C_α for α and C_β for β. The mean interval lengths were recorded as L_α and L_β, respectively.

Table 3.1: Coverage rate and interval length for α and β when n = 10 using the Jeffreys prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9369	0.9425	0.9275	0.8702
2.0	0.8	0.9387	0.9386	3.6911	0.8723
5.0	0.8	0.9351	0.9392	9.2380	0.8726
0.5	5.0	0.9340	0.9388	0.1453	5.4063
2.0	5.0	0.9329	0.9400	0.5768	5.4603
5.0	5.0	0.9359	0.9457	1.4416	5.4488
0.5	8.0	0.9335	0.9402	0.0910	8.7513
2.0	8.0	0.9379	0.9409	0.3664	8.6834
5.0	8.0	0.9356	0.9419	0.9147	8.7010
Average		0.9395	0.9409	1.9325	5.0074

Table 3.2: Coverage rate and interval length for α and β when n = 10 using the divergence prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9294	0.9363	1.2547	0.9284
2.0	0.8	0.9294	0.9381	5.0466	0.92505
5.0	0.8	0.9309	0.9354	89.0336	0.9327
0.5	5.0	0.9391	0.9353	0.1382	5.7795
2.0	5.0	0.9349	0.9348	0.5567	5.7334
5.0	5.0	0.9360	0.9330	1.3924	5.7401
0.5	8.0	0.9341	0.9390	0.0863	9.1980
2.0	8.0	0.9305	0.9391	0.3467	9.1556
5.0	8.0	0.9301	0.9369	0.8712	9.1299
Average		0.9327	0.9364	10.9696	5.2803.

Table 3.3: Coverage rate and interval length for α and β when $n = 10$ using the reference prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9494	0.9300	0.9969	0.8489
2.0	0.8	0.9495	0.9294	3.9686	0.8513
5.0	0.8	0.9464	0.9309	9.9302	0.8518
0.5	5.0	0.9483	0.9269	0.1578	5.2720
2.0	5.0	0.9469	0.9278	0.6261	5.3272
5.0	5.0	0.9492	0.9363	1.5641	5.3130
0.5	8.0	0.9465	0.9307	0.0989	8.5342
2.0	8.0	0.9515	0.9276	0.3986	8.4709
5.0	8.0	0.9496	0.9307	0.9949	8.4886
Average		0.9486	0.9300	2.0820	4.8840

Table 3.4: Coverage rate and interval length for α and β when $n = 40$ using Jeffreys prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9470	0.9477	0.4160	0.3967
2.0	0.8	0.9482	0.9496	1.6662	0.3971
5.0	0.8	0.9447	0.9476	4.1599	0.3983
0.5	5.0	0.9446	0.9456	0.0665	2.4845
2.0	5.0	0.9425	0.9486	0.2663	2.4818
5.0	5.0	0.9427	0.9489	0.6663	2.4814
0.5	8.0	0.9449	0.9486	0.0418	3.9692
2.0	8.0	0.9482	0.9475	0.1669	3.9739
5.0	8.0	0.9477	0.9483	0.4177	3.9704
Average		0.9457	0.9481	0.8884	2.2837

Table 3.5: Coverage rate and interval length for α and β when $n = 40$ using the divergence prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9414	0.9473	0.4374	0.4031
2.0	0.8	0.9410	0.9439	1.7571	0.4029
5.0	0.8	0.9436	0.9454	4.3873	0.4036
0.5	5.0	0.9447	0.9496	0.0661	2.5105
2.0	5.0	0.9443	0.9438	0.2641	2.5174
5.0	5.0	0.9450	0.9472	0.6610	2.5111
0.5	8.0	0.9428	0.9483	0.0413	4.0106
2.0	8.0	0.9474	0.9439	0.1651	4.0191
5.0	8.0	0.9471	0.9457	0.4135	4.0115
Average		0.9441	0.9461	0.9103	2.3100

Table 3.6: Coverage rate and interval length for α and β when n = 40 using the reference prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9493	0.9464	0.4194	0.3950
2.0	0.8	0.9505	0.9465	1.6795	0.3954
5.0	0.8	0.9463	0.9484	4.1929	0.3965
0.5	5.0	0.9460	0.9446	0.0675	2.4740
2.0	5.0	0.9432	0.9461	0.2701	2.4721
5.0	5.0	0.9451	0.9444	0.6760	2.4709
0.5	8.0	0.9461	0.9454	0.0424	3.9527
2.0	8.0	0.9501	0.9438	0.1694	3.9566
5.0	8.0	0.9494	0.9473	0.4240	3.9543
Average		0.9473	0.9459	0.8824	2.2742

Table 3.7: Coverage rate and interval length for α and β when n = 200 using Jeffreys prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9456	0.9497	0.1815	0.1721
2.0	0.8	0.9460	0.9466	0.7266	0.1722
5.0	0.8	0.9455	0.9489	1.8147	0.1720
0.5	5.0	0.9462	0.9454	0.0291	1.0775
2.0	5.0	0.9479	0.9436	0.1166	1.0764
5.0	5.0	0.9478	0.9438	0.2916	1.0769
0.5	8.0	0.9469	0.9443	0.0182	1.7221
2.0	8.0	0.9470	0.9478	0.0729	1.7228
5.0	8.0	0.9495	0.9448	0.1824	1.7209
Average		0.9469	0.9465	0.3815	0.9903

Table 3.8: Coverage rate and interval length for α and β when n = 200 using the divergence prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9525	0.9489	0.1835	0.1739
2.0	0.8	0.9470	0.9518	0.7343	0.1738
5.0	0.8	0.9535	0.9481	1.8368	0.1737
0.5	5.0	0.9481	0.9512	0.0293	1.0849
2.0	5.0	0.9506	0.9457	0.1171	1.0849
5.0	5.0	0.9516	0.9498	0.2928	1.0860
0.5	8.0	0.9513	0.9525	0.0183	1.7356
2.0	8.0	0.9488	0.9502	0.0732	1.7368
5.0	8.0	0.9489	0.9467	0.1829	1.7380
Average		0.9503	0.9497	0.3854	1.1086

Table 3.9: Coverage rate and interval length for α and β when n = 200 using the reference prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9497	0.9506	0.1831	0.1735
2.0	0.8	0.9504	0.9478	0.7330	0.1736
5.0	0.8	0.9490	0.9498	1.8304	0.1733
0.5	5.0	0.9488	0.9472	0.0293	1.0860
2.0	5.0	0.9493	0.9450	0.1174	1.0851
5.0	5.0	0.9500	0.9453	0.2936	1.0854
0.5	8.0	0.9495	0.9464	0.0184	1.7361
2.0	8.0	0.9474	0.9508	0.0734	1.7367
5.0	8.0	0.9505	0.9476	0.1837	1.7337
Average		0.9494	0.9480	0.3842	1.1087

When the sample size is $n = 10$, the performance of the three objective priors differs in terms of the coverage and mean interval length. From Table 3.1, it can be observed that the coverage rates are close to the nominal level for both α and β, and the shortest mean interval lengths for α are obtained when the Jeffreys prior is used. The shortest interval length is obtained for α when $\alpha = 0.5$ and $\beta = 8.0$, with $L_\alpha = 0.0910$. When the divergence prior is used, undercoverage and mean interval lengths are wide, see Table 3.2. When the reference prior is used, the coverage rates are close to the nominal level, and mean interval lengths are shorter for α, and undercoverage for β, but better than those in Table 3.2, see Table 3.3. This means the Jeffreys prior is recommended when we have small sample sizes and the interest is in both α and β.

When the sample size is increased to $n = 40$, the coverage rates become closer to 95% when the Jeffreys prior, divergence prior, and the reference prior. When the Jeffreys prior and the reference prior are used, they provide the shortest mean interval lengths for both α and β, meaning the Jeffreys prior and the reference prior are performing well. See Tables 3.4, 3.5, and 3.6. When the sample size is large, $n = 200$, the coverage rates for all priors are very close to the nominal level of 95% and the mean interval lengths become shorter for both α and β. See Tables 3.7, 3.8, and 3.9.

The priors were also compared using the average of coverage probabilities of α and β, respectively, along with the averages of their corresponding mean interval lengths. When $n = 10$ and the Jeffreys prior is used, the coverage is close to the nominal 95% for both parameters, with a 0.9395 coverage rate for α and 0.9409 for β, with a mean interval length of 1.9325 for α and 5.0074 for β. The mean interval lengths increase with a combination of larger parameter values, indicating more uncertainty as scale or shape parameters increase. When the divergence prior is used, the lowest coverage rate is produced, 0.9364 for α and 0.9364 for β, it generates wider intervals for α, 10.9696, and a longer mean interval length for β compared to others. When the reference prior is used, the best coverage for α is obtained at 0.9486, with a mean interval length of 2.0820, though coverage for β is lower at 0.9300, with a mean interval length of 4.8840. For small sample sizes, when the Jeffreys prior is used, it provides a good performance for both parameters; the divergence prior is unreliable for small sample sizes.

When the sample size is increased to $n = 40$, the results for all three priors used became more consistent, with a coverage rate close to the nominal rate of 95% and the mean interval lengths decrease compared to the sample size of $n = 10$. When the Jeffreys prior is used, the coverage rate is 0.9457 for α and 0.9481 for β, with relatively short intervals (0.8884 for α and 2.2837 for β), reflecting accurate inferences. When the divergence prior is used, the coverage is very similar, 0.9441 for α and 0.9461 for β, with slightly wider intervals of 0.9103 for α and 2.3100 for β. When the reference prior is used, the coverage reaches 0.9473 for α and 0.9459 for β, with the shortest intervals (0.8824 for α and 2.2742 for β). Thus, when $n = 40$, all three priors yield nearly identical results, with the reference prior offering a slight advantage in efficiency. When the sample size is large, $n = 200$, the coverage rates for all priors are very close to the nominal rate of 95% and the mean interval lengths become shorter. When Jeffreys prior is used, the coverage for α is 0.9469 and for α is 0.9465, with a very short mean interval length of (0.3815 for α and 0.9903 for β). When the divergence prior is used, the coverage rates are again nearly similar to those when the Jeffreys prior is used, with minor differences in the mean interval length. When the reference prior is used, the results are also closely aligned. Overall, when the sample size is large, $n = 200$, the choice of prior is negligible.

3.6.2 Type I Right Censoring

In this subsection, the performance of the Jeffreys prior, divergence prior, and the reference prior will be compared for different sample sizes when there is type I right censoring. Their performance will be evaluated using the coverage rate and the mean interval length. A model with the shortest interval length and coverage very close to 95% also known as the nominal level, is considered to be the best. For each simulated true parameter values (α, β), the coverage probabilities were recorded as C_α for α and C_β for β. The mean interval lengths were recorded as L_α and L_β, respectively.

Table 3.10: Coverage rate and interval length for α and β when $n = 10$ and $r = 9$ using the Jeffreys prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9368	0.9431	1.0069	0.9332
2.0	0.8	0.9316	0.9378	3.9817	0.9349
5.0	0.8	0.9350	0.9406	9.9815	0.9360
0.5	5.0	0.9351	0.9397	0.1544	5.8465
2.0	5.0	0.9379	0.9428	0.6209	5.7841
5.0	5.0	0.9350	0.9411	1.5435	5.8484
0.5	8.0	0.9358	0.9457	0.0979	9.2700
2.0	8.0	0.9340	0.9417	0.3909	9.3027
5.0	8.0	0.9330	0.9399	0.9754	9.3351
Average		0.9349	0.9414	2.0837	5.3545

Table 3.11: Coverage rate and interval length for α and β when n = 10 and $r = 9$ using the divergence prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9299	0.9349	2.4641	1.0019
2.0	0.8	0.9243	0.9401	5.8529	0.9952
5.0	0.8	0.9251	0.9342	2.2666	1.0024
0.5	5.0	0.9305	0.9352	1.4773	6.2037
2.0	5.0	0.9253	0.9314	5.9112	6.1886
5.0	5.0	0.9286	0.9362	1.4813	6.1789
0.5	8.0	0.9307	0.9376	9.2075	9.9856
2.0	8.0	0.9338	0.9359	3.6839	9.8772
5.0	8.0	0.9307	0.9320	9.1971	9.9047
Average		0.9288	0.9353	4.6158	5.7048

Table 3.12: Coverage rate and interval length for α and β when n = 10 and $r = 9$ using the reference prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9306	0.9298	1.10254	0.9065
2.0	0.8	0.9311	0.9269	4.36217	0.9084
5.0	0.8	0.9482	0.9303	10.93085	0.9091
0.5	5.0	0.9504	0.9273	0.1699	5.6781
2.0	5.0	0.9527	0.9263	0.6837	5.6226
5.0	5.0	0.9492	0.9255	1.6995	5.6813
0.5	8.0	0.9523	0.9339	0.1081	9.0097
2.0	8.0	0.9481	0.9272	0.4315	9.0389
5.0	8.0	0.9451	0.9301	1.0768	9.0698
Average		0.9453	0.9286	2.2850	5.2027

Table 3.13: Coverage rate and interval length for α and β when n = 40 and $r = 36$ using the Jeffreys prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9434	0.9486	0.4374	0.4210
2.0	0.8	0.9472	0.9460	1.7623	0.4201
5.0	0.8	0.9441	0.9483	4.4118	0.4199
0.5	5.0	0.9430	0.9422	0.0704	2.6232
2.0	5.0	0.9472	0.9462	0.2821	2.6181
5.0	5.0	0.9466	0.9487	0.7026	2.6342
0.5	8.0	0.9462	0.9448	0.0441	4.2012
2.0	8.0	0.9455	0.9492	0.1763	4.2046
5.0	8.0	0.9457	0.9479	0.4420	4.1930
Average		0.9454	0.9469	0.9254	2.4150

Table 3.14: Coverage rate and interval length for α and β when $n = 40$ and $r = 36$ using the divergence prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9455	0.9500	0.4461	0.4262
2.0	0.8	0.9438	0.9465	1.8654	0.4271
5.0	0.8	0.9470	0.9471	4.6572	0.4261
0.5	5.0	0.9466	0.9442	0.0697	2.6619
2.0	5.0	0.9473	0.9440	0.2789	2.6618
5.0	5.0	0.9451	0.9433	0.6964	2.6638
0.5	8.0	0.9449	0.9461	0.0435	4.2557
2.0	8.0	0.9429	0.9432	0.1743	4.2543
5.0	8.0	0.9389	0.9500	0.4354	4.2602
Average		0.9448	0.9450	0.9630	2.4486

Table 3.15: Coverage rate and interval length for α and β when $n = 40$ and $r = 36$ using the reference prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9453	0.9471	0.4414	0.4188
2.0	0.8	0.9492	0.9450	1.7785	0.4180
5.0	0.8	0.9477	0.9464	4.4530	0.4179
0.5	5.0	0.9463	0.9403	0.0716	2.6111
2.0	5.0	0.9481	0.9405	0.2867	2.6052
5.0	5.0	0.9488	0.9454	0.7141	2.6203
0.5	8.0	0.9496	0.9410	0.0449	4.1823
2.0	8.0	0.9476	0.9473	0.1793	4.1834
5.0	8.0	0.9484	0.9447	0.4494	4.1731
Average		0.9480	0.9442	0.9354	2.4033

Table 3.16: Coverage rate and interval length for α and β when $n = 200$ and $r = 175$ using the Jeffreys prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9471	0.9510	0.1953	0.1858
2.0	0.8	0.9505	0.9476	0.7830	0.1857
5.0	0.8	0.9463	0.9495	1.9558	0.1858
0.5	5.0	0.9506	0.9480	0.0313	1.1628
2.0	5.0	0.9460	0.9511	0.1252	1.1611
5.0	5.0	0.9488	0.9472	0.3131	1.1612
0.5	8.0	0.9480	0.9507	0.0196	1.8613
2.0	8.0	0.9534	0.9471	0.0783	1.8600
5.0	8.0	0.9464	0.9481	0.1958	1.8594
Average		0.9486	0.9489	0.4108	1.0692

Table 3.17: Coverage rate and interval length for α and β when $n = 200$ and $r = 175$ using the divergence prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9514	0.9475	0.1968	0.1860
2.0	0.8	0.9453	0.9453	0.7864	0.1860
5.0	0.8	0.9466	0.9487	1.9651	0.1860
0.5	5.0	0.9471	0.9511	0.0313	1.1628
2.0	5.0	0.9489	0.9545	0.1252	1.1609
5.0	5.0	0.9502	0.9487	0.3129	1.1618
0.5	8.0	0.9496	0.9525	0.0196	1.8597
2.0	8.0	0.9495	0.9465	0.0783	1.8618
5.0	8.0	0.9490	0.9490	0.1957	1.8607
Average		0.9486	0.9493	0.4124	1.0695

Figure 3.63: Coverage rate and interval length for α and β when $n = 200$ and $r = 175$ using the reference prior.

α	β	C_α	C_β	L_α	L_β
0.5	0.8	0.9476	0.9498	0.1956	0.1856
2.0	0.8	0.9507	0.9486	0.7842	0.1857
5.0	0.8	0.9472	0.9491	1.9587	0.1855
0.5	5.0	0.9512	0.9477	0.0314	1.1619
2.0	5.0	0.9458	0.9490	0.1256	1.1602
5.0	5.0	0.9494	0.9471	0.3141	1.1600
0.5	8.0	0.9486	0.9496	0.0196	1.8601
2.0	8.0	0.9535	0.9454	0.0786	1.8582
5.0	8.0	0.9469	0.9485	0.1964	1.8578
Average		0.9490	0.9483	0.4116	1.0683

When the sample size is $n = 10$ and $r = 9$, the performance of the three priors differs in terms of the coverage and mean interval length. From Table 3.10 and Table 3.12, it can be observed that, for parameter α, the reference prior provides the coverage rates averaging at 0.9453 which is very close to the nominal level with a mean interval lengths averaging at 2.2850. Jeffreys prior has slightly lower coverage rate (0.9345) with the shortest mean interval length (2.0837), when the divergence prior is used, it provides undercoverage and wider mean interval lengths, see Table 3.11. For β,the Jeffreys prior performs best in terms of coverage rate and produces reasonably short interval lengths. The reference prior has undercoverage (0.93) with shorter mean interval length of 5.2027, whereas the divergence prior shows undercoverage averaging at 0.9353 and the widest interval lengths averaging at 5.7045. These results indicate that no single prior is uniformly best for both parameters. The Jeffreys prior tends to produce shorter intervals, particularly for small samples, while the reference prior may give better coverage for α. The divergence prior generally produces wider intervals and slightly lower coverage for small samples. Therefore, the choice of prior should be considered separately for each parameter rather than generalized.

When type I right censoring with r failures is considered, the performance of the three objective priors is evaluated based on the sample size and number of observed failures. When the sample size is $n = 10$ and $r = 9$, when the Jeffreys prior is used, it produces coverage rates that are close to the nominal level for both parameters, with 0.9349 for α and 0.9414 for β, and the shortest mean interval lengths of 2.0837 for α. When the divergence prior is used, it produces the coverage of 0.9288 for α and 0.9353 for β, and produces wider intervals of 4.6158 for α and 5.7048 for β compared to others. When the reference prior is used, it provides the best coverage of 0.9453 for α and low for β at 0.93, and the shortest mean interval lengths of 5.2027 for β. For a small sample size or few observed failures, the Jeffreys prior performs well for both parameters; the divergence prior is not recommended for a small sample size or when there are few failures, and the reference prior is recommended when the focus is on α. When $n = 40$ and $r = 36$, the results for all priors become close to each other. The coverage rates are close to the nominal level of 95%, and interval lengths become shorter compared to $n = 10$ and $r = 9$. When Jeffreys prior is used, coverage is 0.9454 for α and 0.9469 for β, with the shortest mean interval lengths of 0.9254 for α and 2.4150 for β. When the divergence prior is used, the coverage is very similar, 0.9448 for α and 0.9450 for β, with slightly wider mean interval lengths of 0.9630 for α and 2.4486 for β. When the reference prior is used, it provides an average coverage rate of 0.9454 for α and 0.9469 for β, with the shorter mean interval lengths of 0.9254 for α and 2.4150 for β. This shows that as the number of observed failures increases relative to the sample size, the inference becomes much better and the differences between objective priors vanish. When the sample size is large, $n = 200$ with $r = 175$, coverage rates for all priors are very close to the nominal level, and interval lengths become shorter. In conclusion, the Jeffreys prior and the reference prior are performing well for all sample sizes.

3.7 Predictive Reliability for the Weibull Distribution

This section presents the predictive reliability analysis based on the Weibull distribution under both complete and type I right censoring. The objective is to evaluate how well the Weibull distribution predicts future lifetimes using the posterior estimates of the model parameters. Predictive reliability values are computed using the Jeffreys prior, the divergence prior, and the reference prior for different sample sizes. The results are summarised and compared in tables and plots to assess the impact of prior choice and sample size on predictive performance. The predictive reliability for the Weibull distribution is given by

$$R(t_u|\underline{t}) = \int_0^\infty \int_0^\infty \pi(\alpha, \beta|\underline{t}) R(t_u \mid \alpha, \beta) \, d\alpha d\beta, \tag{3.21}$$

where $R(t_u \mid \alpha, \beta)$ is the Weibull distribution reliability function at time t_u, given in equation 3.21.

This expression is analytically challenging. Given the posterior draws $\left(\alpha^{(n)}, \beta^{(n)}\right)$, $n = 1, 2, ..., N$, from the MCMC simulation, the integral in equation 3.21 can be evaluated by the Monte Carlo average

$$R\left(t_u|\underline{t}\right) \approx \frac{1}{N}\sum_{n=1}^{N} \exp\left\{-\left(\frac{t_u}{\alpha^{(n)}}\right)^{\beta^{(n)}}\right\}.$$

3.7.1 Complete Data

In this section, the predictive reliability for the Weibull distribution will be evaluated within a Bayesian framework using the prior distribution specified above for the shape β, and scale α parameter.

Tables 3.18 and 3.19 summarise the predictive reliability values for sample size $n = 10$ and $n = 40$, respectively, when the Jeffreys prior, divergence prior, and the reference prior are used.

Table 3.18: Predictive reliability table for various priors when $n = 10$.

Time	Jeffreys prior	Divergence prior	Reference prior
1	1	1	1
2	1	1	1
3	0.9991	0.9997	1
.	.	.	.
.	.	.	.
.	.	.	.
145	0.0582	0.0762	0.1925
146	0.0561	0.0732	0.1839
147.	0.0541	0.0703	0.1755
.	.	.	.
.	.	.	.
299	0.0387	0.0403	0.0454
300	0.0376	0.0392	0.0443

Table 3.19: Predictive reliability table for various priors when $n = 40$.

Time	Jeffreys prior	Divergence prior	Reference prior
1	1	1	1
2	1	1	1
3	1	0.9999	1
.	.	.	.
.	.	.	.
.	.	.	.
145	0.0938	0.0582	0.0823
146	0.0892	0.0554	0.0782
147	0.0847	0.0527	0.0743
.	.	.	.
.	.	.	.
299	0.0220	0.0218	0.0224
300	0.0200	0.0204	0.0212

Figures 3.64 and 3.65 illustrate predictive reliability curves for the Weibull distribution using three non-informative priors specified when sample size are 10 and 40 respectively. The predictive reliability curves decreases from one to zero as time increases.

When the sample size is 10, the predictive reliability curves show difference among three priors, the predictive reliability when using the divergence prior yields slightly higher estimates than the predictive reliability when using the Jeffreys prior across time, whereas predictive reliability curve when the reference prior is used lies just below the predictive reliability curve when the Jeffreys prior is used. This spread reflects the great uncertainty when sample size is small. As the sample size increases, $n = 40$, the predictive reliability curves for all three are overlapping almost perfectly. This indicates that larger samples size lead to more precise reliability predictions.

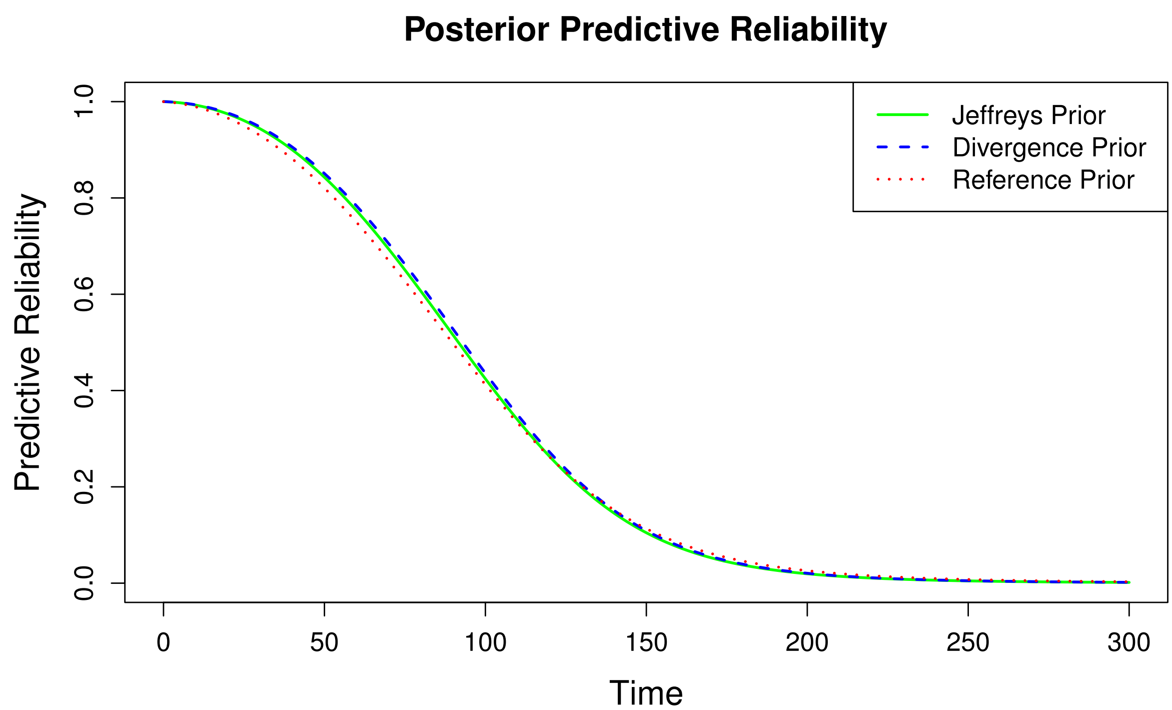

Figure 3.64: Predictive reliability when $n = 10$.

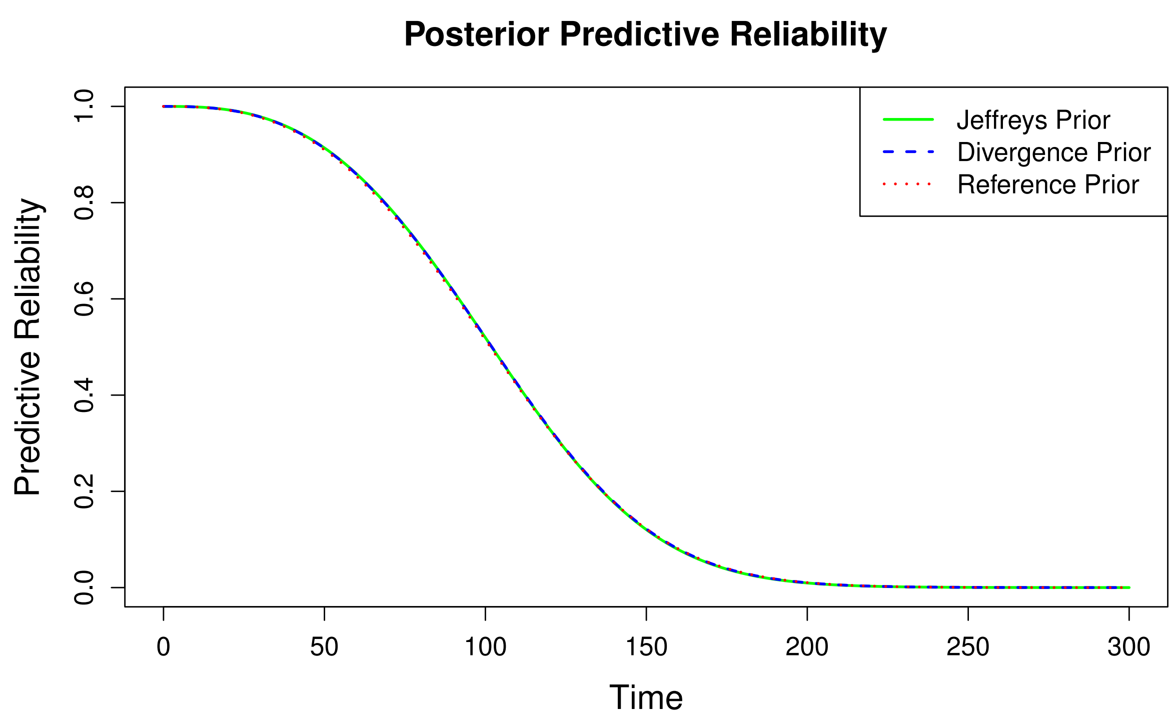

Figure 3.65: Predictive reliability when $n = 40$.

3.7.2 Type I Right Censoring

In this subsection, the predictive reliability for the Weibull distribution under type I censoring will be evaluated, using the Jeffreys prior, divergence prior, and the reference prior for the shape β, and scale α parameters. The posterior predictive reliability curves and tables were generated for sample sizes $n = 10$ and $n = 40$.

Tables 3.20 and 3.21 summarise the predictive reliability values for the sample sizes $n = 10$ and $n = 40$, respectively, when the Jeffreys prior, divergence prior, and the reference prior are used.

Table 3.20: Predictive Reliability for various priors when $n = 10$ and $r = 9$.

Time	Jeffreys prior	Divergence prior	Reference prior
1	1	1	1
2	0.9998	0.9998	0.9996
3	0.9995	0.9995	0.9990
.	.	.	.
.	.	.	.
.	.	.	.
145	0.1932	0.1983	0.1955
146	0.1880	0.1929	0.1906
147.	0.1829	0.1876	0.1858
.	.	.	.
.	.	.	.
299	0.0387	0.0403	0.0454
300	0.0376	0.0392	0.0443

Table 3.21: Predictive Reliability for various priors when $n = 40$ and $r = 36$.

Time	Jeffreys prior	Divergence prior	Reference prior
1	1	1	1
2	1	1	1
3	1	1	1
.	.	.	.
.	.	.	.
.	.	.	.
145	0.2351	0.2363	0.2338
146	0.2285	0.2297	0.2273
147.	0.2221	0.2232	0.2209
.	.	.	.
.	.	.	.
299	0.0320	0.0318	0.0324
300	0.0300	0.0304	0.0312

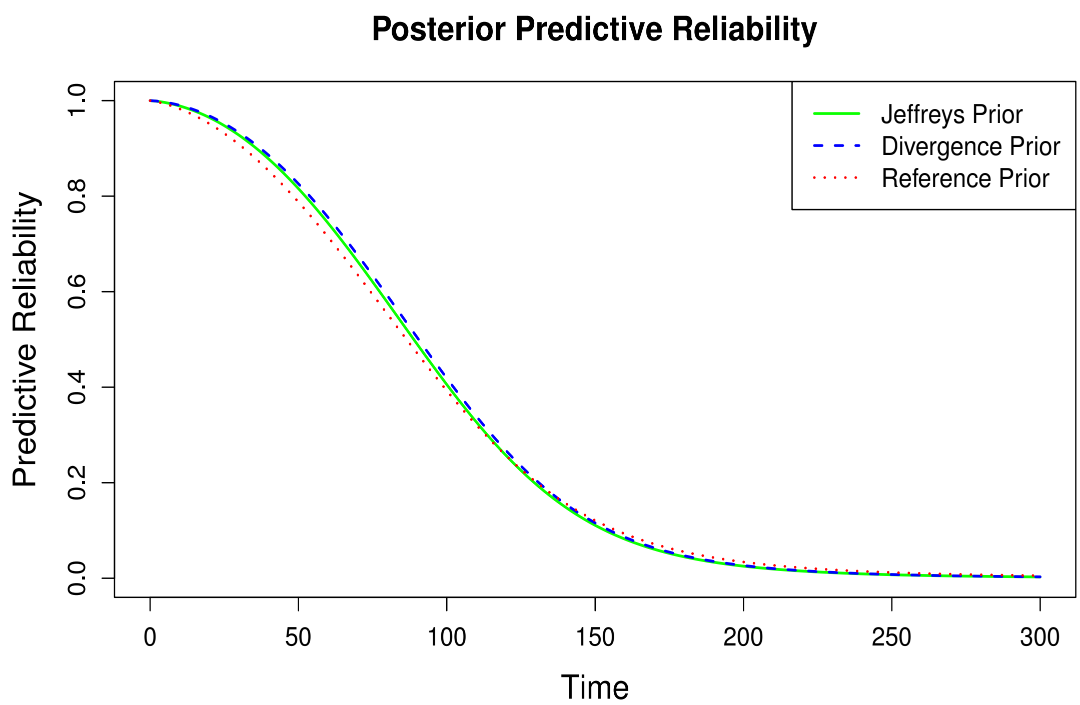

Figure 3.66: Predictive Reliability when $n = 10$ and $r = 9$.

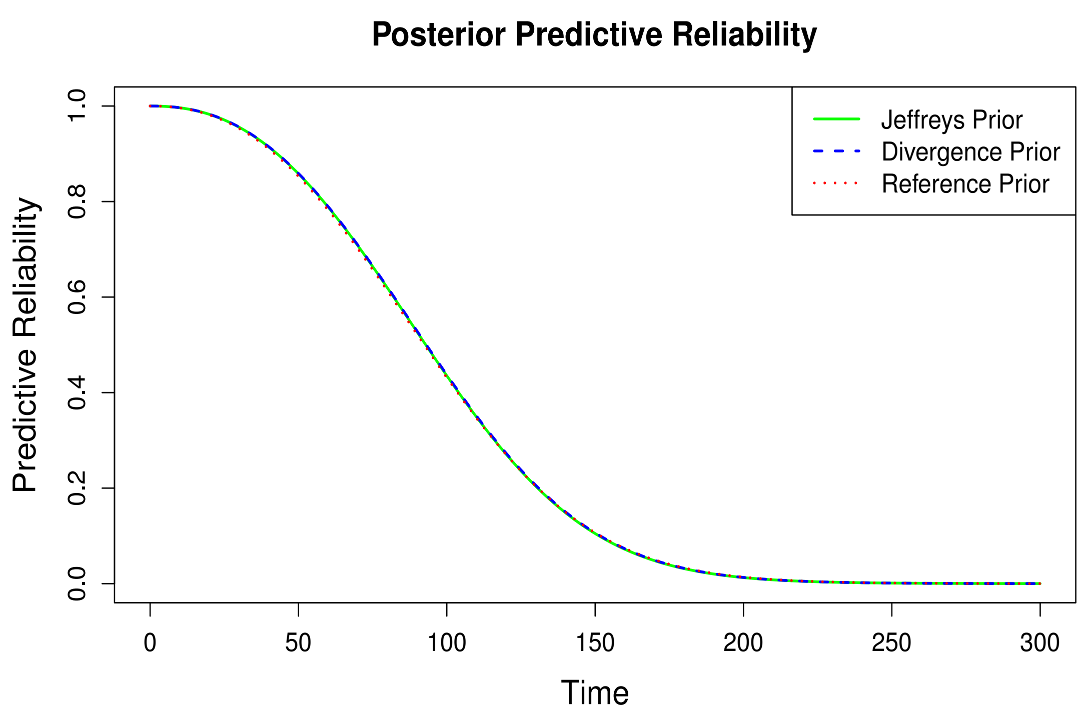

Figure 3.67: Predictive Reliability when $n = 40$ and $r = 36$.

Figures 3.66 and 3.67 illustrate predictive reliability curves for the Weibull distribution using three non-informative priors specified when sample size are 10 and 40, respectively. Across all non-informative priors specified in this thesis and sample sizes, the predictive reliability curves decrease from 1 (at time zero) towards 0 with increasing time, as expected for survival functions.

For $n = 10$ and $r = 9$, the predictive reliability when using the divergence prior yields slightly higher estimates than the predictive reliability when using the Jeffreys prior across time, whereas the predictive reliability curve when the reference prior is used lies just below the predictive reliability curve when the Jeffreys prior is used. When $n = 40$ and $r = 36$, all three curves overlap closely. As the sample size increases, and the number of observed failures increases, the difference vanishes, and the choice among the Jeffreys, the reference, and the divergence priors has a negligible impact on the predictive reliability curve.

Chapter 4

The Birnbaum-Saunders Distribution

The Birnbaum–Saunders distribution, also known as the fatigue life distribution, is a probability distribution used extensively in reliability applications to model failure times. It is named after Birnbaum and Saunders (1969b). This distribution was developed to model failures due to cracks. In reliability applications, the Birnbaum-Saunders distribution is often used to simulate failure times, Achcar (1993). According to Xu and Tang (2011), the Birnbaum-Saunders distribution was created with the exhaustion process taken into account, something that many other popular life distributions occasionally fail to achieve.

The PDF of the Birnbaum–Saunders distribution is given by

$$f(t|\alpha,\beta) = \frac{\sqrt{\frac{t}{\beta}} + \sqrt{\frac{\beta}{t}}}{2\sqrt{2\pi}\alpha t} \exp\left\{-\frac{1}{2\alpha^2}\left(\sqrt{\frac{t}{\beta}} - \sqrt{\frac{\beta}{t}}\right)^2\right\},\ t > 0. \tag{4.1}$$

In the context of this model, the parameter vector $\theta = (\alpha,\beta)$. The CDF of a two-parameter Birnbaum–Saunders random variable T can be written as

$$F(t|\alpha,\beta) = \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right],\ 0 < t < \infty,\ \alpha,\beta > 0, \tag{4.2}$$

where the parameters α and β are the shape and scale parameters, respectively.

The corresponding reliability function is given by

$$R(t|\alpha,\beta) = 1 - \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t}\right)^{\frac{1}{2}}\right\}\right],$$

where $\Phi(\cdot)$ denotes the CDF of the standard normal distribution. The likelihood for the Birnbaum-

Saunders with no censoring is given as

$$L(\alpha,\beta|\underline{t}) = \prod_{i=1}^{n} \frac{\sqrt{\frac{t_i}{\beta}}+\sqrt{\frac{\beta}{t_i}}}{2\sqrt{2\pi}\alpha t_i} \exp\left\{-\frac{1}{2\alpha^2}\left(\sqrt{\frac{t_i}{\beta}}-\sqrt{\frac{\beta}{t_i}}\right)^2\right\}.$$

The likelihood for the Birnbaum-Saunders with no censoring can be rewritten as

$$L(\alpha,\beta|\underline{t}) \propto \alpha^{-n}\beta^{-n}\prod_{i=1}^{n}\left\{\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right\}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

For the detailed derivations see Appendix B.1. The likelihood for the Birnbaum-Saunders distribution with respect to type I right censoring is given as

$$L(\alpha,\beta|\underline{t}) = \prod_{i=1}^{n}\left[\frac{\sqrt{\frac{t_i}{\beta}}+\sqrt{\frac{\beta}{t_i}}}{2\sqrt{2\pi}\alpha t_i}\exp\left\{-\frac{1}{2\alpha^2}\left(\sqrt{\frac{t_i}{\beta}}-\sqrt{\frac{\beta}{t_i}}\right)^2\right\}\right]^{\delta_i}\left[1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right]^{1-\delta_i},$$

where, for observation i, $data_i = (t_i,\delta_i)$, t_i is either a failure time or a right censored time, $\delta_i = 1$ for an exact failure time, and $\delta_i = 0$ for a right censored observation. The Fisher information matrix for Birnbaum-Saunders was derived by Lemoine (2019) and is given by

$$H(\alpha,\beta) = \begin{bmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & \frac{n\left(1+\left(\alpha\sqrt{\frac{\pi}{2}}-\pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1-\Phi\left(\frac{2}{\alpha}\right)\right]\right)(2\pi)^{-\frac{1}{2}}\right)}{\alpha^2\beta^2} \end{bmatrix}. \tag{4.3}$$

The detailed derivations are provided in Appendix A.2.

Figure 4.1 shows possible shapes for the PDF of the Birnbaum-Saunders distribution for various values of the parameters α and β, and the corresponding reliability functions are shown in Figure 4.2.

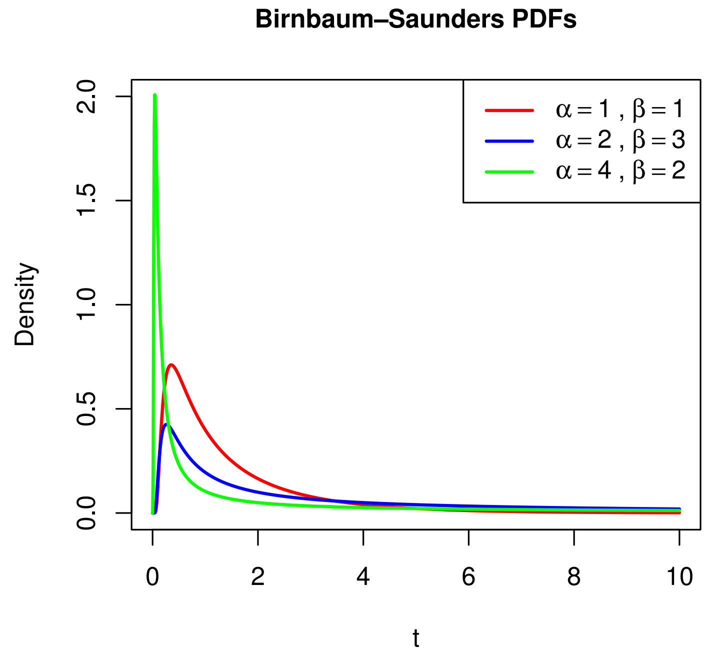

Figure 4.1: Probability density function of Birnbaum Saunders distribution for various values of α and β.

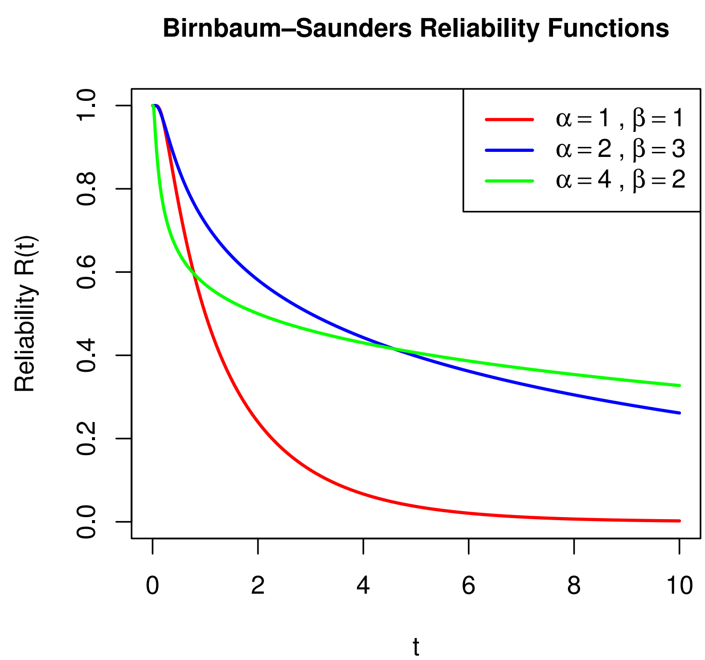

Figure 4.2: Reliability function of the Birnbaum-Saunders distribution for various values of α and β.

4.1 The Posterior for the Birnbaum-Saunders Distribution using the Jeffreys Prior

In this section, the posterior distribution for the Birnbaum-Saunders distribution using the Jeffreys prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined.

Using the definition of Jeffreys prior in equation 2.6, the Jeffreys prior for Birnbaum-Saunders distribution is given by

$$\pi_J(\alpha,\beta) \propto \frac{1}{\alpha\beta}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{2}},$$

see Lemoine (2019).

The joint posterior distribution of α and β given the observed data using the Jeffreys prior when there is no censoring is given by

$$\pi_J(\alpha,\beta|\underline{t}) \propto \alpha^{-n-1}\beta^{-n-1}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{2}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}. \tag{4.4}$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_J(\alpha|\beta,\underline{t}) \propto \alpha^{-n-1}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{2}}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}. \tag{4.5}$$

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_J(\beta|\alpha,\underline{t}) \propto \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)\right\}. \tag{4.6}$$

The joint posterior distribution of α and β given the observed data using the Jeffreys prior under type I censoring is given by

$$\pi_J(\alpha,\beta|\underline{t}) \propto \alpha^{-r-1}\beta^{-r-1}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{2}}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}$$
$$\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.$$

The conditional posterior of α given β and $\underline{t}$ has an unknown form and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. This situation motivates the use of alternative MCMC techniques, such as the MH algorithm, to sample from the joint posterior distribution.

Theorem 4.1. *The joint posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the Jeffreys prior when there is no censoring is an improper distribution.*

Proof. If the following result holds

$$\int_0^\infty \int_0^\infty \pi_J(\alpha, \beta | t)\, d\alpha d\beta = \infty,$$

then the posterior distribution is improper.

Using the well-known binomial approximation, the term

$$\left(\frac{1}{\alpha^2} + \frac{1}{4}\right)^{\frac{1}{2}},$$

can be approximated as

$$\frac{1}{\alpha}\left(1 + \frac{\alpha^2}{8}\right) = \left(\frac{1}{\alpha} + \frac{\alpha}{8}\right) \geq \frac{1}{\alpha}.$$

Thus,

$$\begin{aligned}
\int_0^\infty \int_0^\infty \pi_J(\alpha, \beta | t)\, d\alpha d\beta \propto & \int_0^\infty \int_0^\infty \alpha^{-n-1} \beta^{-n-1} \left(\frac{1}{\alpha^2} + \frac{1}{4}\right)^{\frac{1}{2}} \prod_{i=1}^n \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \\
& \times \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\} d\alpha d\beta \\
\geq & \int_0^\infty \int_0^\infty \alpha^{-n-1} \beta^{-n-1} \frac{1}{\alpha} \prod_{i=1}^n \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \\
& \times \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\} d\alpha d\beta.
\end{aligned}$$

Rewrite the integral as

$$\int_0^\infty \int_0^\infty \alpha^{-n-2} \beta^{-n-1} \prod_{i=1}^n \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^n \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\} d\alpha d\beta = \int_0^\infty \int_0^\infty H_J(\alpha, \beta)\, d\alpha d\beta.$$

Let $u=\alpha^2$ then $\alpha=u^{\frac{1}{2}}$, $\mathrm{d}\alpha=\frac{u^{-\frac{1}{2}}du}{2}$, then this follows

$$\int_0^\infty\int_0^\infty H_J(\alpha,\beta)\,d\alpha d\beta = \frac{1}{2}\int_0^\infty\int_0^\infty \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]u^{-\left(\frac{n+1}{2}+1\right)}\exp\left\{-\frac{1}{2u}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}dud\beta$$

$$==\frac{1}{2}\int_0^\infty\int_0^\infty \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left[\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n+1}{2}}}d\beta.$$

Let $C=\frac{\Gamma\left(\frac{n+1}{2}\right)}{2\left(\frac{1}{2}\right)^{\frac{n+1}{2}}}$, then

$$\frac{1}{2}\int_0^\infty\int_0^\infty \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\frac{\Gamma\left(\frac{n+1}{2}\right)}{\left[\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n+1}{2}}}d\beta=\int_0^\infty\frac{C\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n+1}{2}}\beta^{n+1}}d\beta.$$

Define

$$h(\beta)=\frac{C\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n+1}{2}}\beta^{n+1}}=\frac{C\prod_{i=1}^{n}\left[\beta^{\frac{3}{2}}\left\{\frac{1}{\beta t_i^{\frac{1}{2}}}+\frac{1}{t_i^{\frac{3}{2}}}\right\}\right]}{\left[\sum_{i=1}^{n}\beta\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{2}{\beta}\right)\right]^{\frac{n+1}{2}}\beta^{n+1}}$$

$$=\frac{C\beta^{\frac{3n}{2}}\prod_{i=1}^{n}\left\{\frac{1}{\beta t_i^{\frac{1}{2}}}+\frac{1}{t_i^{\frac{3}{2}}}\right\}}{\beta^{\frac{n+1}{2}}\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{2}{\beta}\right)\right]^{\frac{n+1}{2}}\beta^{n+1}}$$

$$=\frac{C\prod_{i=1}^{n}\left\{\frac{1}{\beta t_i^{\frac{1}{2}}}+\frac{1}{t_i^{\frac{3}{2}}}\right\}}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{2}{\beta}\right)\right]^{\frac{n+1}{2}}\beta^{\frac{3}{2}}}.$$

The simplification of $h(\beta)$ is obtained by factoring out powers of β from both the numerator and the denominator.

Taking limits,

$$\lim_{\beta\to\infty}\frac{h(\beta)}{\beta^{-\frac{3}{2}}}=\frac{C\prod_{i=1}^{n}\left[\frac{1}{t_i^{\frac{3}{2}}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)\right]^{\frac{n+1}{2}}}.$$

Thus, as $\beta \to \infty$,

$$h(\beta) = O\left(\frac{1}{\beta}\right).$$

For large K

$$\int_0^\infty h(\beta)\,d\beta \geq \int_K^\infty h(\beta)\,d\beta \to \infty.$$

This completes the proof, and we can conclude that the posterior is improper. □

4.2 The Posterior for the Birnbaum-Saunders Distribution using the Divergence Prior

In this section, the posterior distribution for the Birnbaum-Saunders distribution using the divergence prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined.

Using the definition of divergence prior in equation 2.7, the divergence prior for the Birnbaum-Saunders distribution is given by

$$\pi_D(\alpha,\beta) \propto \left(\frac{2n^2}{\alpha^2\beta^2}\left[\frac{1}{\alpha^2}+\frac{1}{4}\right]\right)^{\frac{1}{4}},$$

where $\frac{2n^2}{\alpha^2\beta^2}\left[\frac{1}{\alpha^2}+\frac{1}{4}\right]$ is the determinant of the Fisher information matrix for the Birnbaum-Saunders distribution.

The divergence prior for the Birnbaum-Saunders distribution can be rewritten as

$$\pi_D(\alpha,\beta) \propto \frac{1}{\alpha^{\frac{1}{2}}\beta^{\frac{1}{2}}}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{4}}. \tag{4.7}$$

The joint posterior distribution of α and β given the observed data using the divergence prior when there is no censoring is given by

$$\pi_D(\alpha,\beta|\underline{t}) \propto \alpha^{-n-\frac{1}{2}}\beta^{-n-\frac{1}{2}}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{4}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_D(\alpha|\beta,\underline{t}) \propto \alpha^{-n-\frac{1}{2}}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{4}}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}. \tag{4.8}$$

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_D(\beta|\alpha,\underline{t}) \propto \beta^{-n-\frac{1}{2}} \prod_{i=1}^{n} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)\right\}. \tag{4.9}$$

The joint posterior distribution of α and β given the observed data using the divergence prior under type I right censoring is given by

$$\pi_D(\alpha,\beta|\underline{t}) \propto \alpha^{-r-\frac{1}{2}}\beta^{-r-\frac{1}{2}}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{4}} \prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}$$
$$\times \prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.$$

The conditional posterior of α given β and $\underline{t}$ has an unknown form and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. This situation motivates the use of alternative MCMC techniques, such as the MH algorithm, to sample from the joint posterior distribution.

Theorem 4.2. *The joint posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the divergence prior when there is no censoring is an improper distribution.*

Proof. If the following result holds

$$\int_0^\infty\int_0^\infty \pi_D(\alpha,\beta|t)\,d\alpha d\beta = \infty,$$

then the posterior distribution is improper.

Using the well-known binomial approximation, the term

$$\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{4}},$$

can be approximated as

$$\frac{1}{\alpha^{\frac{1}{2}}}\left(1+\frac{\alpha^2}{16}\right) = \left(\frac{1}{\alpha^{\frac{1}{2}}}+\frac{\alpha^{\frac{3}{2}}}{16}\right) \geq \frac{1}{\alpha^{\frac{1}{2}}}.$$

Thus,

$$\begin{aligned}\int_0^\infty\int_0^\infty \pi_D(\alpha,\beta|\underline{t})\,d\alpha d\beta &\propto \int_0^\infty\int_0^\infty \alpha^{-n-\frac{1}{2}}\beta^{-n-\frac{1}{2}}\left(\frac{1}{\alpha^2}+\frac{1}{4}\right)^{\frac{1}{4}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta\\ &\geq \int_0^\infty\int_0^\infty \alpha^{-n-\frac{1}{2}}\beta^{-n-\frac{1}{2}}\frac{1}{\alpha^{\frac{1}{2}}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta\\ &= \int_0^\infty\int_0^\infty \alpha^{-n-1}\beta^{-n-\frac{1}{2}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta.\end{aligned}$$

Rewrite the integral as

$$\int_0^\infty\int_0^\infty \alpha^{-n-1}\beta^{-n-\frac{1}{2}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta = \int_0^\infty\int_0^\infty H_D(\alpha,\beta)\,d\alpha d\beta.$$

Let $u=\alpha^2$ then $\alpha=u^{\frac{1}{2}}$, $\mathrm{d}\alpha=\frac{u^{-\frac{1}{2}}du}{2}$, then this follows

$$\begin{aligned}\int_0^\infty\int_0^\infty H_D(\alpha,\beta)\,d\alpha d\beta &= \frac{1}{2}\int_0^\infty\int_0^\infty \beta^{-n-\frac{1}{2}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]u^{-\left(\frac{n}{2}+1\right)}\exp\left\{-\frac{1}{u}\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}du d\beta\\ &= \frac{1}{2}\int_0^\infty\int_0^\infty \beta^{-n-\frac{1}{2}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\frac{\Gamma\left(\frac{n}{2}\right)}{\left[\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}d\beta.\end{aligned}$$

Let $C=\frac{\Gamma\left(\frac{n}{2}\right)}{2\left(\frac{1}{2}\right)^{\frac{n}{2}}}$, then

$$\frac{1}{2}\int_0^\infty\int_0^\infty \beta^{-n-\frac{1}{2}}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\frac{\Gamma\left(\frac{n}{2}\right)}{\left[\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}d\beta = \int_0^\infty \frac{C\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}\beta^{n+\frac{1}{2}}}d\beta.$$

Define

$$h(\beta) = \frac{C\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right]^{\frac{n}{2}}\beta^{n+\frac{1}{2}}} = \frac{C\prod_{i=1}^{n}\left[\beta^{\frac{3}{2}}\left\{\frac{1}{\beta t_i^{\frac{1}{2}}} + \frac{1}{t_i^{\frac{3}{2}}}\right\}\right]}{\left[\sum_{i=1}^{n}\beta\left(\frac{t_i}{\beta^2} + \frac{1}{t_i} - \frac{2}{\beta}\right)\right]^{\frac{n}{2}}\beta^{n+\frac{1}{2}}}$$

$$= \frac{C\beta^{\frac{3n}{2}}\prod_{i=1}^{n}\left\{\frac{1}{\beta t_i^{\frac{1}{2}}} + \frac{1}{t_i^{\frac{3}{2}}}\right\}}{\beta^{\frac{n}{2}}\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2} + \frac{1}{t_i} - \frac{2}{\beta}\right)\right]^{\frac{n}{2}}\beta^{n+\frac{1}{2}}}$$

$$= \frac{C\prod_{i=1}^{n}\left\{\frac{1}{\beta t_i^{\frac{1}{2}}} + \frac{1}{t_i^{\frac{3}{2}}}\right\}}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2} + \frac{1}{t_i} - \frac{2}{\beta}\right)\right]^{\frac{n}{2}}\beta^{\frac{1}{2}}}.$$

Taking limits,

$$\lim_{\beta\to\infty}\frac{h(\beta)}{\beta^{-\frac{3}{2}}} = \frac{C\prod_{i=1}^{n}\left[\frac{1}{t_i^{\frac{3}{2}}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)\right]^{\frac{n}{2}}}.$$

Thus, as $\beta \to \infty$,

$$h(\beta) = O\left(\frac{1}{\beta}\right).$$

For large K

$$\int_{0}^{\infty} h(\beta)\,d\beta \geq \int_{K}^{\infty} h(\beta)\,d\beta \to \infty.$$

This completes the proof, and we can conclude that the posterior is improper.

□

4.3 The Posterior for the Birnbaum–Saunders Distribution using the Reference prior

In this section, the posterior distribution for the Birnbaum-Saunders distribution using the reference prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined.

The reference prior was derived by Xu and Tang (2010), using Berger and Bernardo (1992) algorithm. The reference prior for the Birnbaum Saunders distribution is given as

$$\pi_R(\alpha,\beta) \propto \alpha^{-1}\beta^{-1},$$

by Xu and Tang (2010).

The joint posterior distribution of α and β given the observed data using reference prior when there is no censoring is given by

$$\pi_R(\alpha,\beta|\underline{t}) \propto \alpha^{-n-1}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

The joint posterior distribution when using the reference prior is improper, Xu and Tang (2011).

The conditional posterior of α given β and $\underline{t}$ is given by

$$\pi_R(\alpha|\beta,\underline{t}) \propto \alpha^{-n-1}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}, \tag{4.10}$$

letting $u=\alpha^2$in the previous expression, then it is obtained that, $u=\alpha^2 \sim Inv-G\left(\frac{n}{2},\ \frac{1}{2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right)$.

The conditional posterior of β given α and $\underline{t}$ is given by

$$\pi_R(\beta|\alpha,\underline{t}) \propto \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)\right\}. \tag{4.11}$$

The joint posterior distribution of α and β given the observed data using reference prior under type I censoring is given by

$$\begin{aligned}\pi_R(\alpha,\beta|\underline{t}) \propto & \alpha^{-r-1}\beta^{-r-1}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}\\ &\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.\end{aligned}$$

The conditional posterior of α given β and $\underline{t}$ has a known form and the conditional posterior of β given α and $\underline{t}$ has an unknown form therefore, a standard Gibbs sampler cannot be used. Instead, a Gibbs sampler with a MH step (Metropolis-within-Gibbs) can be employed.

4.4 The Posterior for the Birnbaum–Saunders Distribution using the Probability Matching Prior

In this section, the posterior distribution for the Birnbaum-Saunders distribution using probability matching prior will be derived, when there is no censoring and when there is type I right censoring. Additionally, the properness of the resulting posterior distribution will be examined. Kang and Kim (2025) established that the reference prior of the Birnbaum-Saunders distribution. I now derive the reference prior for the Birnbaum-Saunders model using the Datta and Ghosh (1995) algorithm, the result contradicts with Kang and Kim (2025).

The inverse of the Fisher information is given by

$$H^{-1}(\alpha,\beta)=\begin{bmatrix}\frac{\alpha^2}{2n} & 0\\ 0 & \frac{\alpha^2\beta^2}{n\left(1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right)}\end{bmatrix}.$$

Suppose the parameter of interest is $t(\theta)=\alpha$, then this is the case

$$\nabla_t^{\mathsf{T}}(\theta)=\left[\frac{\partial t(\theta)}{\partial\alpha}\ \frac{\partial t(\theta)}{\partial\beta}\right]=[1\ 0].$$

Now

$$\eta_t^{\mathsf{T}}(\theta)=\frac{\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)}{\sqrt{\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)\nabla_t(\theta)}}$$

The numerator part

$$\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)=[1\ 0]\begin{bmatrix}\frac{\alpha^2}{2n} & 0\\ 0 & \frac{\alpha^2\beta^2}{n\left(1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right)}\end{bmatrix}=\left[\frac{\alpha^2}{n}\ 0\right],$$

and the numerator part without a square root

$$\nabla_t^{\mathsf{T}}(\theta)H^{-1}(\theta)\nabla_t(\theta)=\left[\frac{\alpha^2}{n}\ 0\right]\begin{bmatrix}1\\ 0\end{bmatrix}=\frac{\alpha^2}{n},$$

as a results

$$\eta_t^{\mathsf{T}}(\theta)=\left[\frac{\alpha}{n\sqrt{\frac{1}{n}}}\ 0\right].$$

The probability matching prior is given by

$$\pi_{PMP}(\alpha,\beta) \propto \frac{1}{\alpha\beta}.$$

since

$$\begin{aligned}
\frac{\partial}{\partial\alpha}\{\eta_1(\theta)\pi(\theta)\}+\frac{\partial}{\partial\beta}\{\eta_2(\theta)\pi(\theta)\} =&\frac{\partial}{\partial\alpha}\left\{\frac{\alpha}{n\sqrt{\frac{1}{n}}}\frac{1}{\alpha\beta}\right\}+\frac{\partial}{\partial\beta}\left\{0\frac{1}{\alpha\beta}\right\}\\
=&\frac{\partial}{\partial\alpha}\left\{\frac{1}{n\beta\sqrt{\frac{1}{n}}}\right\}+\frac{\partial}{\partial\beta}\{0\}\\
\frac{\partial}{\partial\alpha}\{\eta_1(\theta)\pi(\theta)\}+\frac{\partial}{\partial\beta}\{\eta_2(\theta)\pi(\theta)\} =&0.
\end{aligned}$$

The joint posterior distribution of α and β given the observed data using probability matching prior when there is no censoring is given by

$$\pi_{PMP}(\alpha,\beta|\underline{t}) \propto \alpha^{-n-1}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

The joint posterior distribution of α and β given the observed data using probability matching prior under type I censoring is given by

$$\begin{aligned}
\pi_{PMP}(\alpha,\beta|\underline{t}) \propto&\alpha^{-r-1}\beta^{-r-1}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}\\
&\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.
\end{aligned}$$

The probability matching prior for the Birnbaum–Saunders distribution, derived using the algorithm of Datta and Ghosh (1995) was found to coincide with the reference prior for this model. Since the posterior distribution obtained under the reference prior is improper, the corresponding posterior distribution arising from the probability matching prior is also improper.

4.5 Assessment of Posterior Convergence through MCMC for the Birnbaum-Saunders Distribution

In this section, the posterior distributions corresponding to the objective priors, include the Jeffreys prior, divergence prior, and the reference prior are examined for different sample sizes using MCMC methods. The selected sample sizes are commonly used in simulation studies to represent small, moderate, and large samples, allowing comparison with existing reliability and Bayesian inference studies in the literature. It is well established in the Bayesian literature that the use of objective priors for the Birnbaum-Saunders distribution often leads to improper posterior distributions. These results have been showed mathematically, now the purpose of the present convergence assessment is therefore to numerically verify these results using MCMC simulation. In Bayesian inference, verifying the properness of the posterior distribution is a fundamental requirement. A posterior distribution must integrate to one in order to be considered valid for inferences, if it is improper, any resulting estimates, credible intervals, or predictive assessments become meaningless and should not be used. See Appendix C.3 for the R code.

4.5.1 Model Definition

This subsection defines the Birnbaum-Saunders model, its reparameterization into log-space, and how the stable log-likehood is computed to avoid numerical issues. The model assumes that $(t_1,...,t_n > 0)$ are i.i.d from the Birnbaum-Saunders variables with shape $(\alpha > 0)$ and scale $(\beta > 0)$.

4.5.1.1 Reparameterization

To simplify, parameters are trasnformed to an unconstrained space

$\eta = \log(\alpha)$, $\zeta = \log(\beta)$, with Jacobian $|J| = \alpha\beta = e^{\eta+\zeta}$.

The computations are done in log-space for the following reasons

- Numerical Stability: Logarithms convert products into sums, preventing underflow in posterior calculations.

- Simplified Acceptance Computation: Acceptance probabilities depend only on log differences, which are numerically safer.

- Removal of Constraints: Transforming to $(\eta,\zeta) = (\log(\alpha), \log(\beta))$ eliminates positivity constraints and allows Gaussian proposals.

- Interpretation in Log-Space: Log-space corresponds to multiplicative changes in (α,β), aligning with the natural scaling of the Birnbaum-Saunders model.

- Convergence is invariant under smooth one-to-one transformations. You run the chain on (η,ζ) $= (\log(\alpha), \log(\beta))$. If the chain in (η,ζ) converges (or drifts), then the transformed variables $(\alpha,\beta) = \left(e^{\eta}, e^{\zeta}\right)$ also converge (or drift).

4.5.2 Log Posterior (unnormalized) for each Prior

This section explains how the log-posterior is derived for each of the three priors after transforming to log-space.

Let $\ell(\eta,\zeta)=\sum_i \log f\left(t_i|\alpha=e^{\eta},\beta=e^{\zeta}\right)$.

4.5.2.1 Reference Prior

$\pi_R(\alpha,\beta)\propto 1/\alpha\beta$. After the Jacobian cancels

$$\log\widetilde{\pi_R}=\ell(\eta,\zeta).$$

4.5.2.2 Jeffreys Prior

$\pi_J(\alpha,\beta)\propto(\alpha\beta)^{-1}\sqrt{\alpha^{-2}+1/4}$. After Jacobian cancellation

$$\log\widetilde{\pi_J}=\ell(\eta,\zeta)+\frac{1}{2}\log\left(e^{-2\eta}+1/4\right).$$

Stable computation

If $(\eta>0)$: $\left(-\log 2+0.5\log\left(1+4e^{-2\eta}\right)\right)$.

If $(\eta<0)$: $\left(-\eta+0.5\log\left(1+0.25e^{2\eta}\right)\right)$.

4.5.2.3 Divergence Prior

$\pi_D(\alpha,\beta)\propto(\alpha\beta)^{-1/2}\left(\alpha^{-2}+1/4\right)^{1/4}$. After Jacobian cancellation

$$\log\widetilde{\pi_D}=\ell(\eta,\zeta)+\frac{1}{2}\eta+\frac{1}{2}\zeta+4\log\left(e^{-2\eta}+1/4\right).$$

The last term is half of the Jeffreys adjustment term.

4.5.3 Random-Walk MH

This section outlines how the random-walk MH sampler explores the posterior in log-space using Gaussian proposals.

4.5.3.1 Initialization

Set $\eta_0=\log(\alpha_0)$, $\zeta_0=\log(\beta_0)$, and compute initial $\log\widetilde{\pi}(\eta_0,\zeta_0)$.

4.5.3.2 Proposal and Acceptance Step

For each iteration $(t = 1, \ldots, T)$

- Propose $\eta^{'} = \eta + N\left(0, s_{\eta}^{2}\right)$, $\zeta^{'} = \zeta + N\left(0, s_{\zeta}^{2}\right)$.
- Compute $\log \widetilde{\pi}\left(\eta^{'}, \zeta^{'}\right)$.
- Accept with probability $\min\left(1, e^{\log \widetilde{\pi}\left(\eta^{'}, \zeta^{'}\right) - \log \widetilde{\pi}(\eta, \zeta)}\right)$.
- Store $\alpha_t = e^{\eta}$, $\beta_t = e^{\zeta}$ and return the full trace of (α_t, β_t).

4.5.4 Role of Tuning Parameters

This section explains how the sampler's tuning parameters influence exploration and stability.

4.5.4.1 Base Step Size

A base step size for $\eta = \log(\alpha)$, controls how aggressively proposals move in $\log(\alpha)$. Larger values increase tail exploration but can reduce acceptance.

4.5.4.2 S_eta_multipliers

Per-chain scaling factors applied to base step size . Used to vary proposal scales across chains for convergence testing.

4.5.4.3 Proposal Standard Deviation

Proposal standard deviation for $\zeta = \log(\beta)$; smaller for stability as β typically varies on a slower scale.

4.5.4.4 Over-Dispersed Initial Values

Chains start from widely separated (α, β) pairs to reveal divergence if the posterior is improper.

4.5.5 Multiple Chains and Diagnostics

This section describes how multiple chains and diagnostics are used to assess convergence and detect drift.

4.5.5.1 Running Multiple Chains

Independent chains with different initializations and proposal scales are run to test for convergence.

4.5.5.2 Diagnostic Visualization

- Trace plots for $\log(\alpha)$ and $\log(\beta)$.
- Sequential $\widehat{R}$ plots to detect late-stage divergence.

The sequential $\widehat{R}$ diagnostic is adopted as a tool for determining when the Markov chains have reached convergence and whether they stay converged.

Figure 4.3 presents the traces plots of $\log\alpha$ and $\log\beta$ for a sample size $n = 10$, corresponding to the Jeffreys prior, divergence prior, and the reference prior, respectively. For $\log\alpha$, the six chains all show a strong upward linear trend without mixing, whereas for $\log\beta$, the chains are double in length compared to those of $\log\alpha$, and pointing downwards without mixing. The posterior chains when the Jeffreys prior, divergence prior, and the reference prior are used for the Birnbaum-Saunders distribution cannot be covered by the thick pen, indicating that the posteriors fail to converge.

The trace plots of $\log\alpha$ and $\log\beta$ for $n = 10$ under the three priors show unstable and non-stationary behavior, indicating divergence or lack of convergence. Figure 4.4 presents the traces plots of $\log\alpha$ and $\log\beta$ for a sample size $n = 40$, corresponding to the Jeffreys prior, divergence prior, and the reference prior, respectively. The trace plots of $\log\alpha$ and $\log\beta$ for $n = 40$ under the three priors show unstable and non-stationary behavior, indicating divergence and lack of convergence. Figures 4.5 presents the trace plots of $\log\alpha$ and $\log\beta$ for a sample size is $n = 200$, corresponding to the Jeffreys prior, divergence prior, and the reference prior, respectively.

The trace plots of $\log\alpha$ and $\log\beta$ for $n = 200$ under the three priors show unstable and non-stationary behaviour, indicating divergence or lack of convergence. As the sample size increases, the chains tend to diverge further from one another. The chains cannot be covered by the thick pen test, illustrating that, even as the sample size increases, the chains remain poorly mixed and fail to converge. These results indicate that posterior inference for the Birnbaum-Saunders distribution using objective priors is unreliable, as the chains fail to converge and do not adequately explore the parameter space. This has important implications for the study, as it suggests that estimates of α and β may be biased or highly variable, and alternative sampling strategies or priors may be necessary for robust inference.

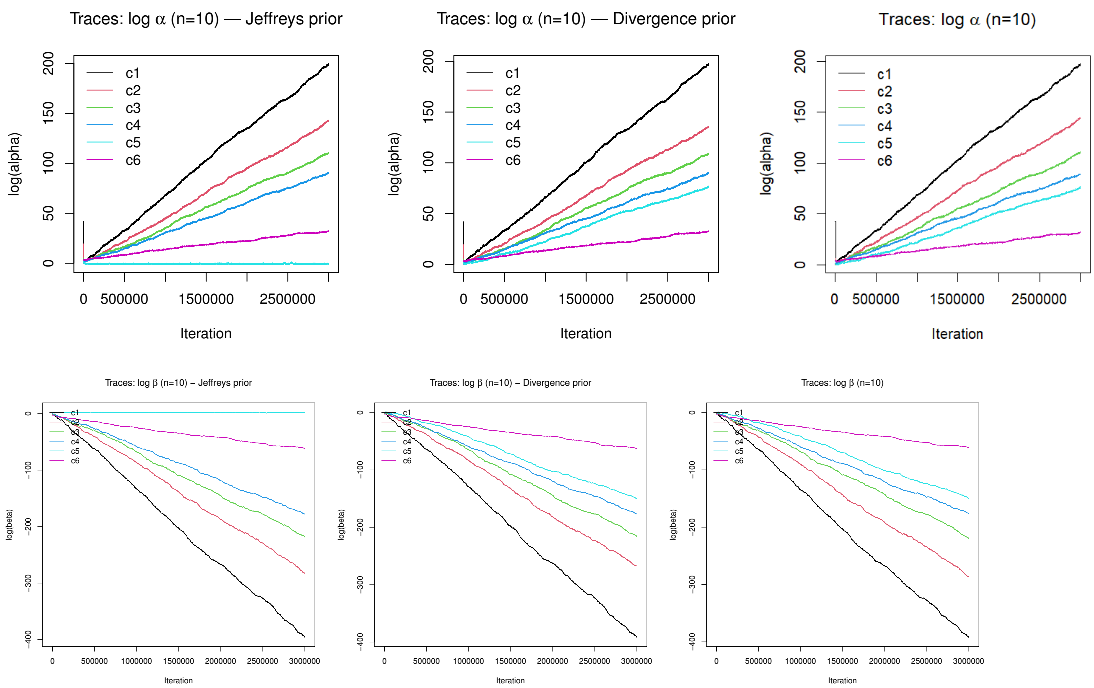

Figure 4.3: Trace plots of $\log\alpha$ and $\log\beta$ when n=10 using the Jeffreys prior, divergence prior, and the reference prior, respectively.

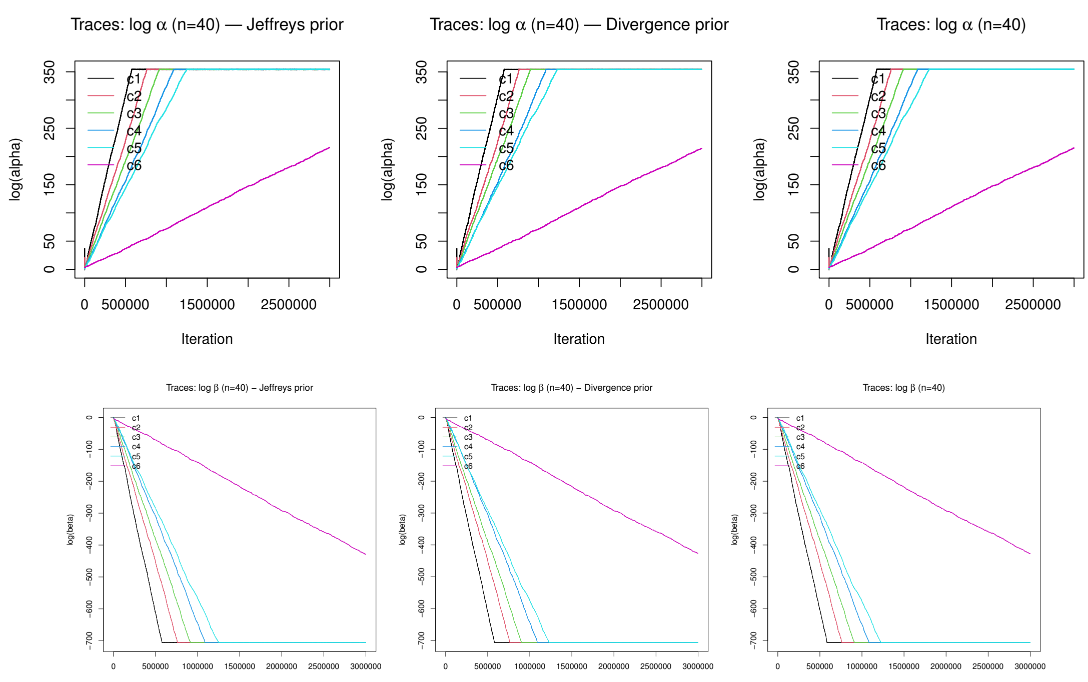

Figure 4.4: Trace plot of $\log\alpha$ and $\log\beta$ when n =40 using the Jeffreys prior, divergence prior, and the reference prior, respectively.

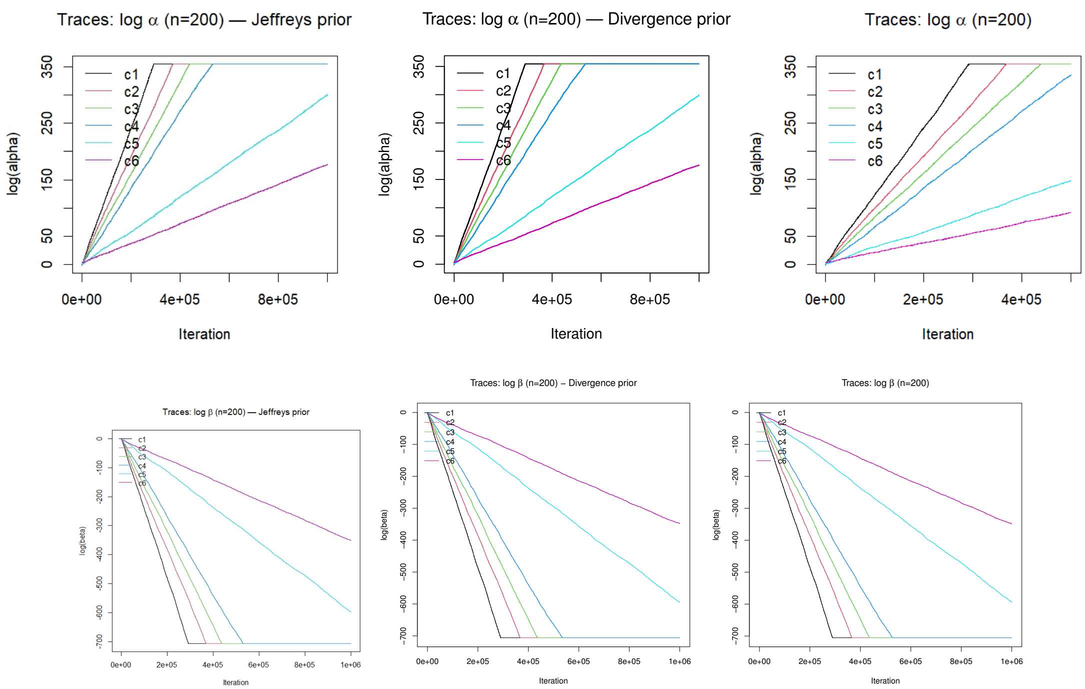

Figure 4.5: Trace plot of $\log\alpha$ and $\log\beta$ when $n = 200$ using the Jeffreys prior, divergence prior, and the reference prior, respectively.

Figures 4.6, 4.7, and 4.8 present the sequential $\hat{R}$ plots of $\log\alpha$ for a sample size $n = 10$, when the Jeffreys prior, divergence prior, and the reference prior are used, respectively. From Figure 4.6, it can be observed that $\widehat{R}$ values start high, then drop to 1, and from 2000000, they increase again, showing clear divergence. From Figures 4.7 and 4.8, it can be observed that $\widehat{R}$ values start high, then drop to 1, and from 2250000, they increase again, showing clear divergence.

Figures 4.9, 4.10, and 4.11 present the sequential $\hat{R}$ plots of $\log\alpha$ for a sample size $n = 40$, when the Jeffreys prior, the divergence prior, and the reference prior are used, respectively. From Figures 4.9, 4.10, and 4.11, it can be observed that $\widehat{R}$ values start high, then drop to 1, and from 1500000, they increase again, showing clear divergence.

Figures 4.12, 4.13, and 4.14 present the sequential $\hat{R}$ plots of $\log\alpha$ for a sample size $n = 200$, when the Jeffreys prior, the divergence prior, and the reference prior are used, respectively. It can be observed that as the sample size increases, the majority of $\hat{R}$ values fluctuate above 1 throughout the iteration, rather than staying close to 1. It is clear that the convergence is not met.

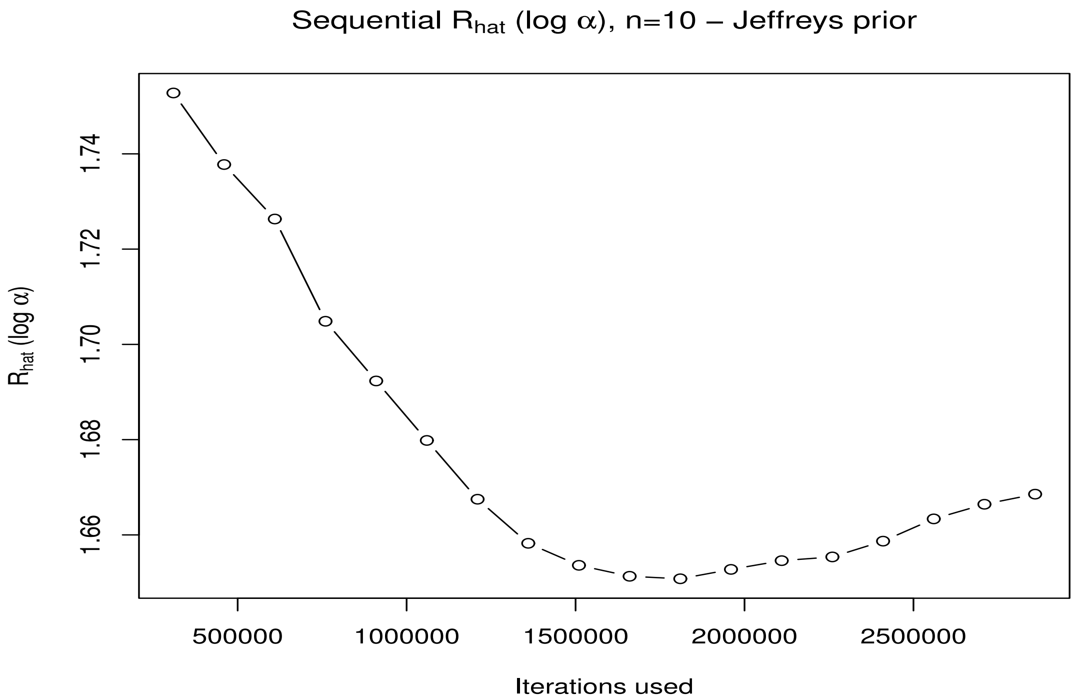

Figure 4.6: The sequential $\hat{R}$ plot of $\log \alpha$ when n=10 using the Jeffreys prior.

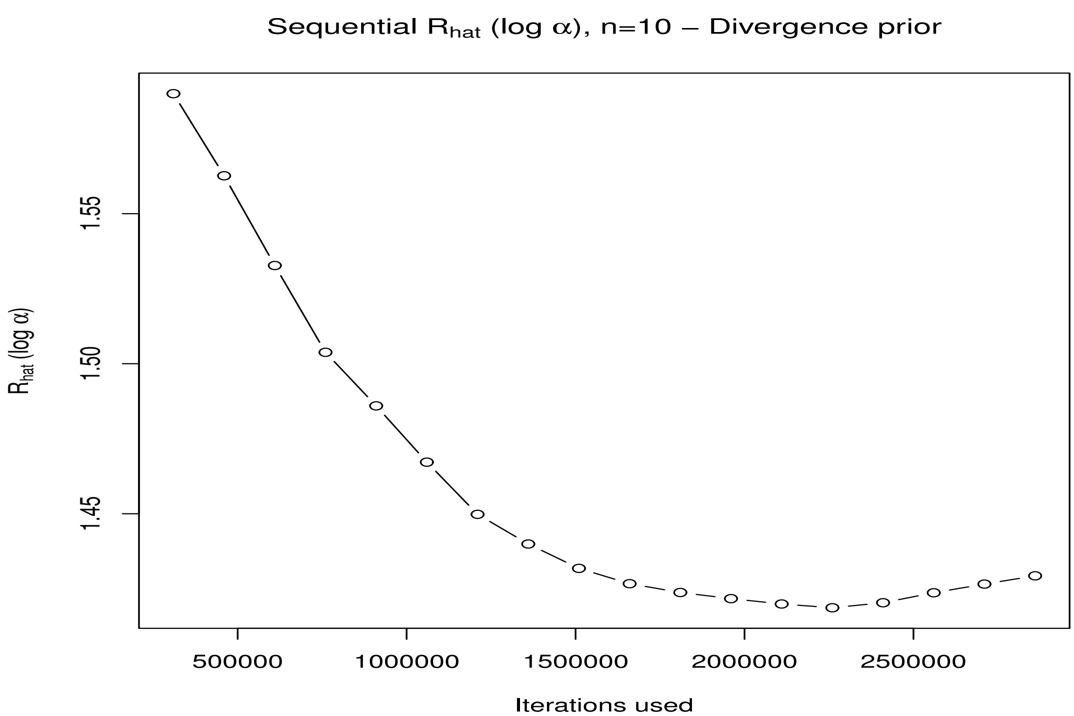

Figure 4.7: The sequential $\hat{R}$ plot of $\log \alpha$ when n=10 using the divergence prior.

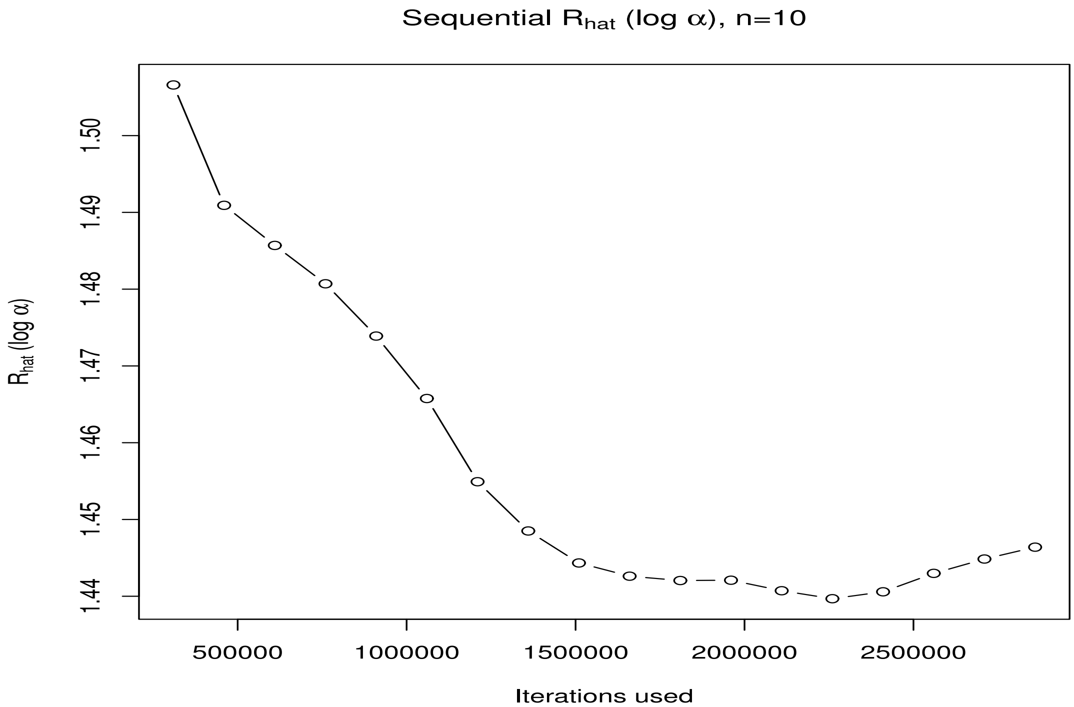

Figure 4.8: The sequential $\hat{R}$ plot of $\log\alpha$ when $n = 10$ using the reference prior.

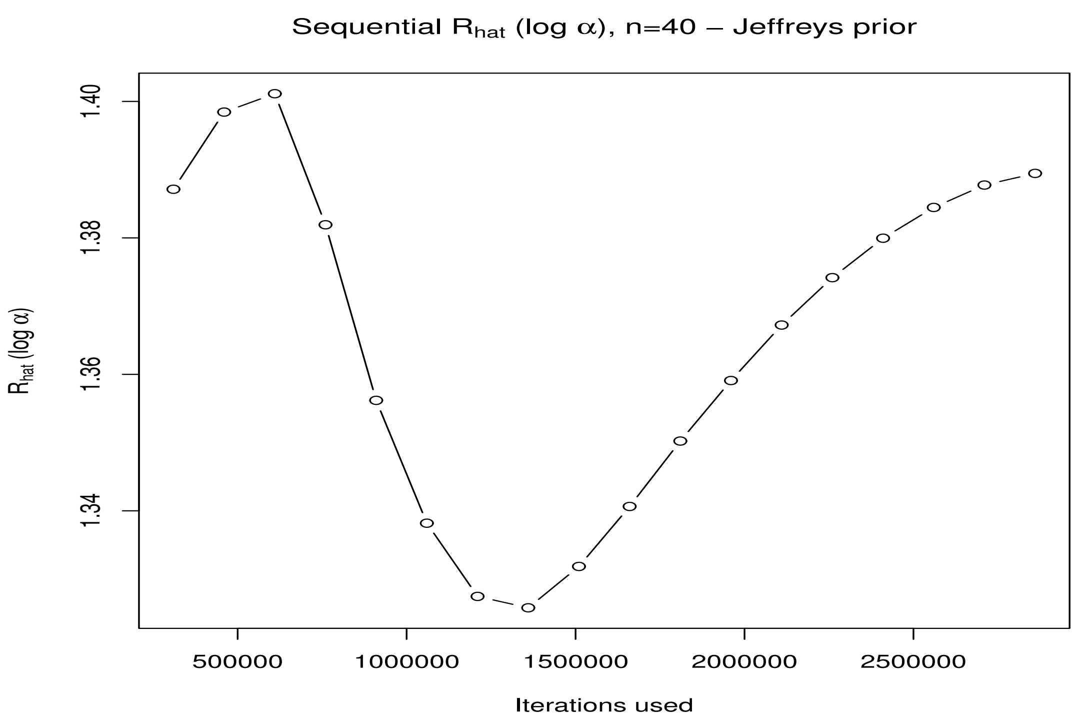

Figure 4.9: The sequential $\hat{R}$ plot of $\log\alpha$ when $n = 40$ using the Jeffreys prior.

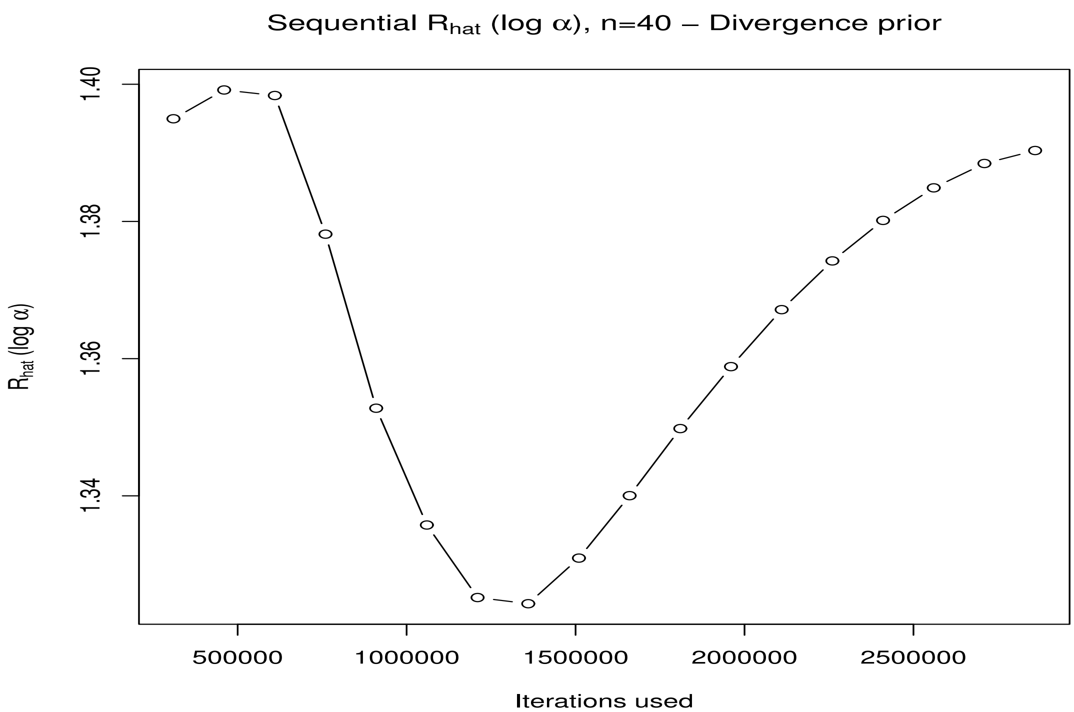

Figure 4.10: The sequential $\hat{R}$ plot of $\log \alpha$ when $n = 40$ using the divergence prior.

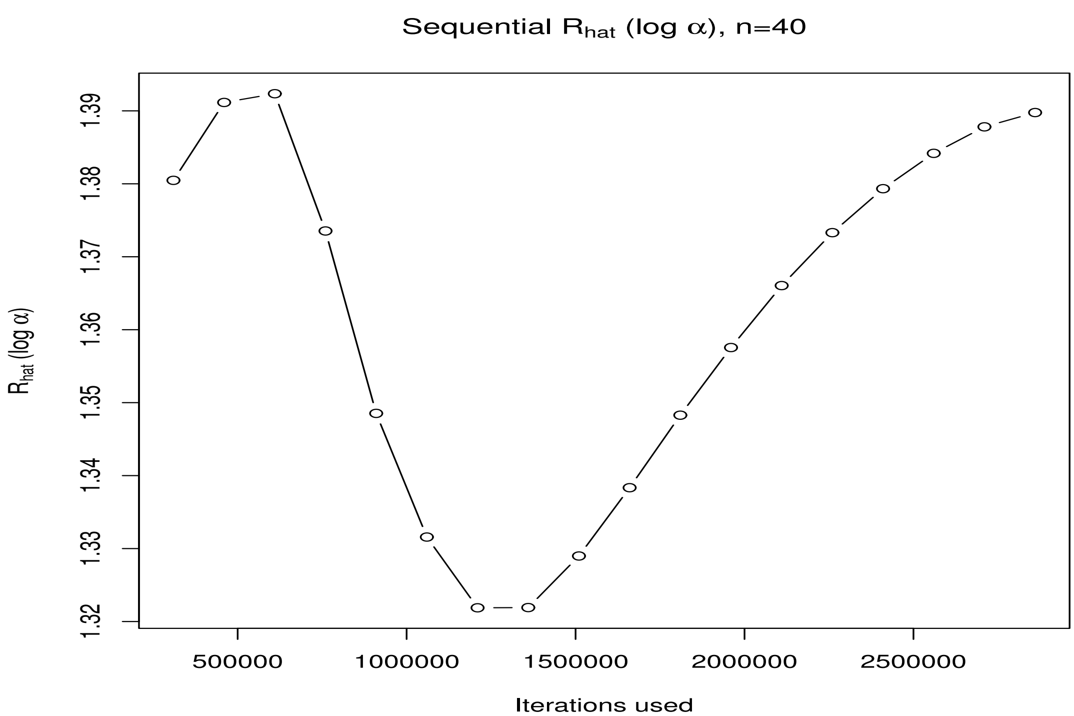

Figure 4.11: The sequential $\hat{R}$ plot of $\log \alpha$ when $n = 40$ using the reference prior.

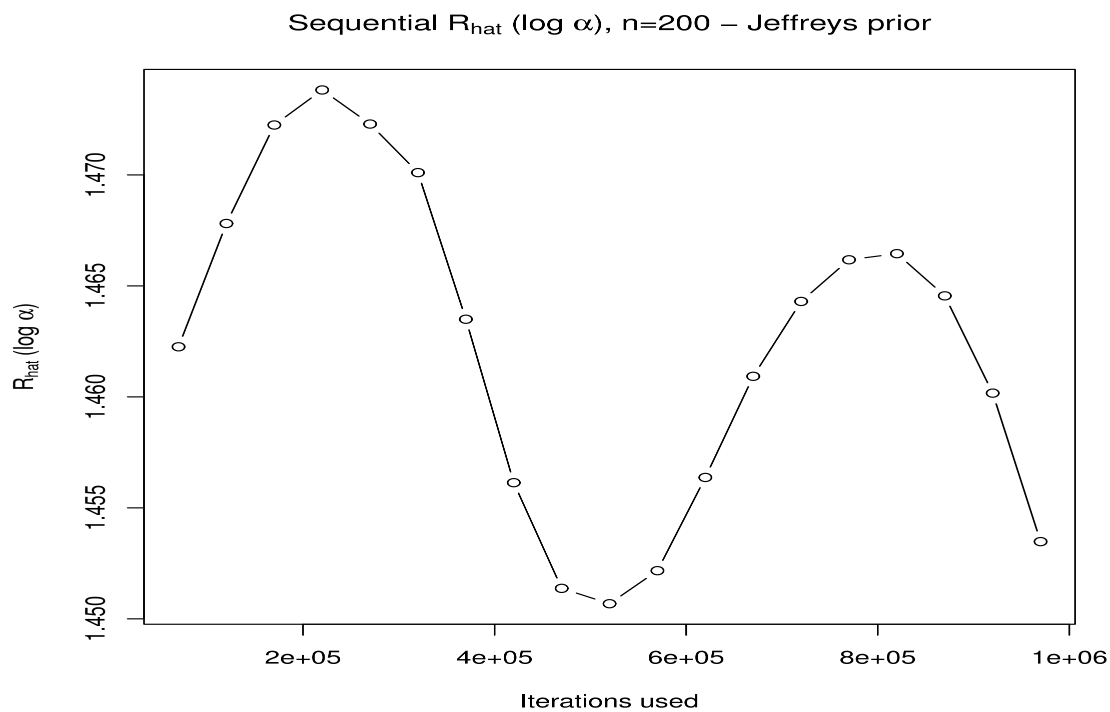

Figure 4.12: The sequential $\widehat{R}$ plot of $\log \alpha$ when $n = 200$ using the Jeffreys prior.

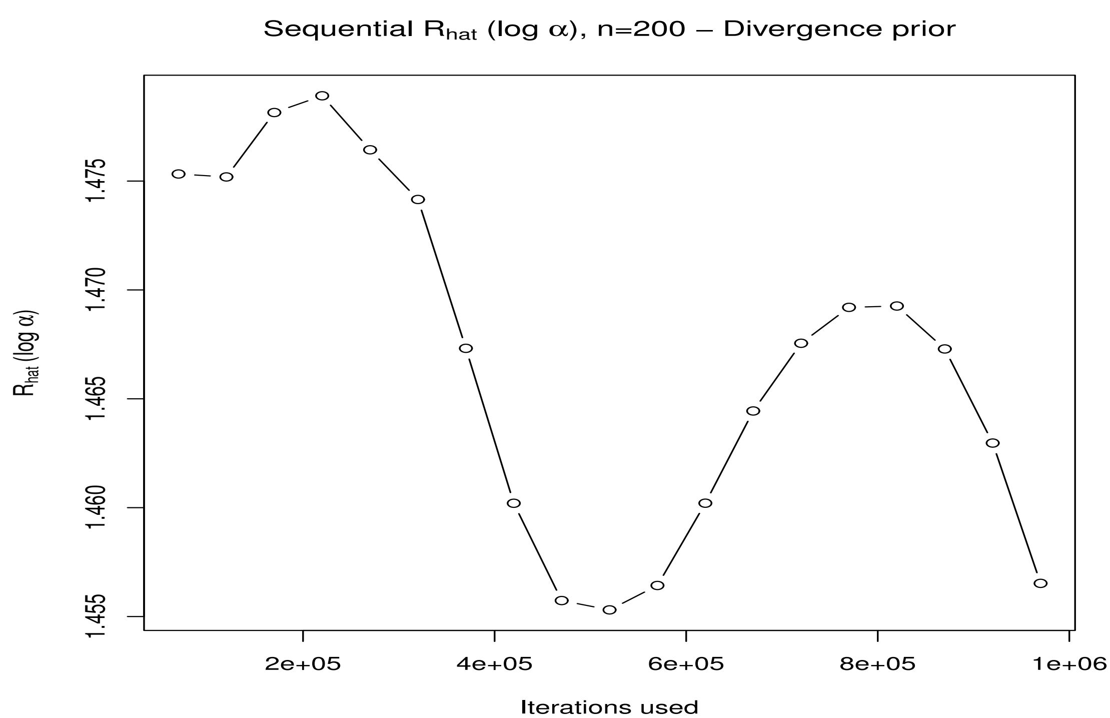

Figure 4.13: The sequential $\hat{R}$ plot of $\log \alpha$ when $n = 200$ using the divergence prior.

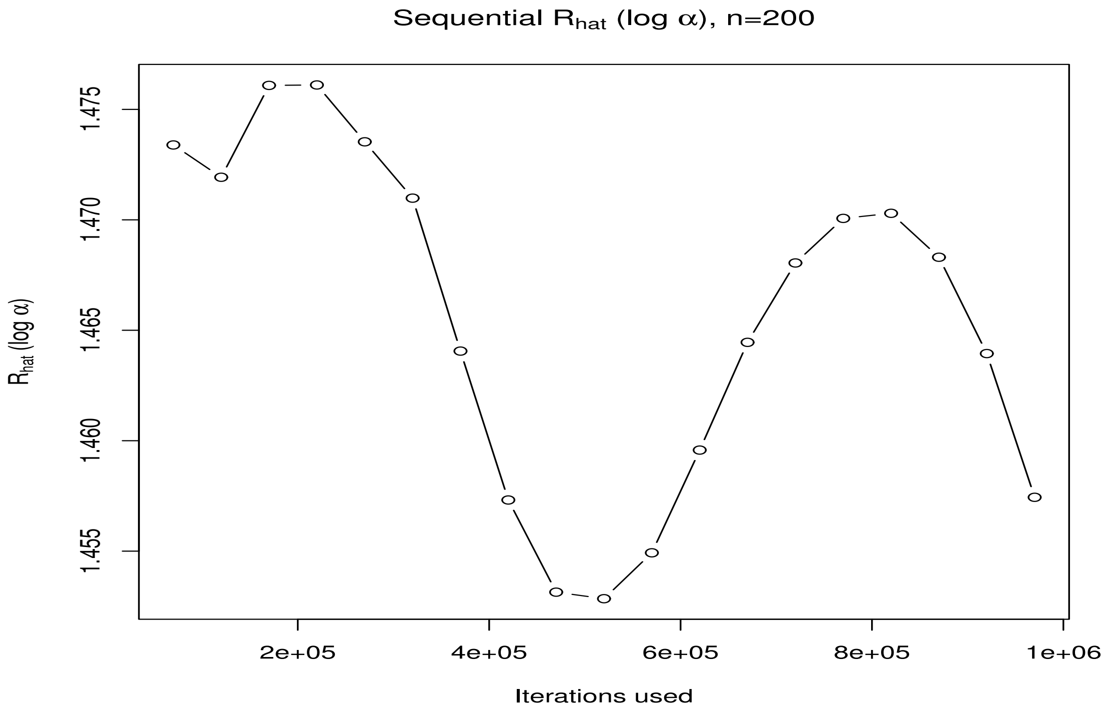

Figure 4.14: The sequential $\hat{R}$ plot of $\log\alpha$ when $n = 200$ using the reference prior.

4.6 Alternative Prior Distributions and the Corresponding Posterior Distributions for the Birnbaum–Saunders Distribution

In this section, the posterior behaviour of the Birnbaum–Saunders distribution under two different prior distribution choices is examined, since it has been shown in the Bayesian literature and in the previous section that the use of non-informative priors leads to an improper posterior distribution, making them unsuitable for Bayesian inference. To overcome this issue, the weak informative priors proposed by Xu and Tang (2011) are considered, and it is shown in this section that the resulting posterior distributions are proper and, therefore, valid for Bayesian analysis. Xu and Tang (2011) established posterior distribution properness under type II censoring. In this section, their approach is used to examine the properness of the corresponding posterior distributions, this is done on both the no censoring case and the type I right censoring case.

These priors are given by

$$\pi_1(\alpha,\beta) = \frac{\pi(\beta)}{\alpha} \tag{4.12}$$

and

$$\pi_2(\alpha,\beta) = \frac{\pi(\alpha)}{\beta}. \tag{4.13}$$

It is assumed that

$$\pi(\alpha) \propto \alpha^{-2a_1-2} \exp\left(-\frac{b_1}{\alpha^2}\right), \quad \pi(\beta) \propto \beta^{-a_2-1} \exp\left(-\frac{b_2}{\beta}\right),$$

where a_1,a_2, b_1, $b_2 > 0$ are hyperparameters controlling the prior information.
Then

$$\pi_1(\alpha,\beta) \propto \alpha^{-1}\beta^{-a_2-1} \exp\left(-\frac{b_2}{\beta}\right), \quad \pi_2(\alpha,\beta) \propto \beta^{-1}\alpha^{-2a_1-2} \exp\left(-\frac{b_1}{\alpha^2}\right).$$

The joint posterior distribution of α and β given the observed data using the prior in equation 4.12 when there is no censoring is given by

$$\pi_1(\alpha,\beta|\underline{t}) \propto \pi(\beta)\,\alpha^{-n-1}\beta^{-n} \prod_{i=1}^{n} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^{n} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\}.$$

The conditional posterior of α given β and the data for the above joint posterior distribution is given by

$$\pi_1(\alpha|\beta,\underline{t}) \propto \alpha^{-n-1} \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^{n} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\}.$$

The conditional posterior of β given α and the data for the above joint posterior distribution is given by

$$\pi_1(\beta|\alpha,\underline{t}) \propto \pi(\beta)\,\beta^{-n} \prod_{i=1}^{n} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^{n} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\}.$$

The joint posterior distribution of α and β given the observed data using the prior in equation 4.12 under type I censoring is given by

$$\pi_1(\alpha,\beta|\underline{t}) \propto \pi(\beta)\,\alpha^{-r-1}\beta^{-r} \prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^{r} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\}$$
$$\times \prod_{i=r+1}^{n} \left\{1 - \Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}} - \left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.$$

The joint posterior distribution of α and β given the observed data using the prior in equation 4.13 when there is no censoring is given by

$$\pi_2(\alpha,\beta|\underline{t}) \propto \pi(\alpha)\,\alpha^{-n}\beta^{-n-1} \prod_{i=1}^{n} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2} \sum_{i=1}^{n} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\}.$$

The conditional posterior of α given β and the data for the above joint posterior distribution is given by

$$\pi_2(\alpha|\beta,\underline{t}) \propto \pi(\alpha)\,\alpha^{-n}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

The conditional posterior of β given α and the data for the above joint posterior distribution is given by

$$\pi_2(\beta|\alpha,\underline{t}) \propto \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

The posterior distribution of α and β given the observed data using the prior in equation 4.13 under type I censoring is given by

$$\begin{aligned}\pi_2(\alpha,\beta|\underline{t}) \propto&\,\pi(\alpha)\,\alpha^{-r}\beta^{-r-1}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}\\ &\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.\end{aligned}$$

Theorem 4.3. *The joint posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.12 under no censoring is a proper distribution.*

Proof. The joint posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.12 under no censoring is given by

$$\pi_1(\alpha,\beta|\underline{t}) \propto \pi(\beta)\,\alpha^{-n-1}\beta^{-n}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

To show that the posterior is proper, this must be true

$$\int_0^\infty\int_0^\infty c\pi(\beta)\,\alpha^{-n-1}\beta^{-n}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta = 1,$$

where c is the normalising constant.

For the above to be true, the following must be shown

$$\int_0^\infty\int_0^\infty \pi(\beta)\,\alpha^{-n-1}\beta^{-n}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta < \infty.$$

Taking integration of $\pi_1(\alpha,\beta|\underline{t})$ with respect to α, it is obtained that

$$\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,d\alpha \propto \int_0^\infty \pi(\beta)\,\alpha^{-n-1}\beta^{-n}\prod_{i=1}^n\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^n\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha$$
$$=\pi(\beta)\,\beta^{-n}\prod_{i=1}^n\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\int_0^\infty \alpha^{-n-1}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^n\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha$$

By letting $\lambda=\alpha^2$, the integral is transformed to

$$\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,d\alpha \propto \pi(\beta)\,\beta^{-n}\prod_{i=1}^n\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\int_0^\infty \lambda^{-\frac{n}{2}-1}\exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^n\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\lambda$$
$$=\frac{\Gamma\left(\frac{n}{2}\right)\pi(\beta)\,\beta^{-n}\prod_{i=1}^n\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\frac{1}{2}\sum_{i=1}^n\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}$$
$$=\frac{c_1\pi(\beta)\,\beta^{-n}\prod_{i=1}^n\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^n\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}=H(\beta),$$

where $c_1=\frac{\Gamma\left(\frac{n}{2}\right)}{\left[\frac{1}{2}\right]^{\frac{n}{2}}}$. It remains to be shown that

$$\int_0^\infty H(\beta)\,d\beta<\infty.$$

By examining the limits

$$\lim_{\beta\to 0}\frac{\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^n\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}=\lim_{\beta\to 0}\frac{\beta^{\frac{n}{2}}\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}+\frac{\beta}{t_i^{\frac{3}{2}}}\right]}{\beta^n\left[\beta^{-1}\sum_{i=1}^{n}\left(t_i+\frac{\beta^2}{t_i}-2\beta\right)\right]^{\frac{n}{2}}}$$

$$=\lim_{\beta\to 0}\frac{\beta^{\frac{n}{2}}\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}+\frac{\beta}{t_i^{\frac{3}{2}}}\right]}{\beta^{n-\frac{n}{2}}\left[\sum_{i=1}^{n}\left(t_i+\frac{\beta^2}{t_i}-2\beta\right)\right]^{\frac{n}{2}}}$$

$$=\frac{\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}\right]}{\left[\sum_{i=1}^{n}(t_i)\right]^{\frac{n}{2}}}.$$

$$\lim_{\beta\to\infty}\frac{\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^n\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}=\lim_{\beta\to\infty}\frac{\beta^{\frac{3n}{2}}\prod_{i=1}^{n}\left[\frac{1}{\beta t_i^{\frac{1}{2}}}+\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^{n+\frac{n}{2}}\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{2}{\beta}\right)\right]^{\frac{n}{2}}}$$

$$=\frac{\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)\right]^{\frac{n}{2}}}.$$

Hence as $\beta\to 0$ or $\beta\to\infty$, $H(\beta)=\mathcal{O}(\pi(\beta))$.

Since $\pi(\beta)$ is integrable, then

$$\int_0^\infty H(\beta)\,d\beta<\infty,$$

proving that the posterior is proper. □

Theorem 4.4. *The joint posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.13 under no censoring is a proper distribution.*

Proof. The posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.13 under no censoring is given by

$$\pi_2(\alpha,\beta|\underline{t})\propto\pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

To show that the posterior is proper, this must be true

$$\int_0^\infty \int_0^\infty c\pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\beta\, d\alpha = 1,$$

where c is the normalising constant.

For this to hold, it must be shown that

$$\int_0^\infty \int_0^\infty \pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\beta\, d\alpha < \infty.$$

For any $0 < \varepsilon < 1$, the integral can be decomposed as

$$\begin{aligned}P = \int_0^\infty \int_0^\infty \pi_2(\alpha,\beta|\underline{t})\,d\beta\, d\alpha &= \int_0^\infty \int_0^\varepsilon \pi_2(\alpha,\beta|\underline{t})\,d\beta\, d\alpha + \int_0^\infty \int_\varepsilon^{\frac{1}{\varepsilon}} \pi_2(\alpha,\beta|\underline{t})\,d\beta\, d\alpha + \int_0^\infty \int_{\frac{1}{\varepsilon}}^\infty \pi_2(\alpha,\beta|\underline{t})\,d\beta\, d\alpha \\ &= P_1 + P_2 + P_3.\end{aligned}$$

Firstly it must be shown that $P_1 < \infty$. Thus, for any $\varepsilon < \min\left(1, \sum_{i=1}^{n}\frac{t_i}{4n}\right)$,

$$\begin{aligned}P_1 &= \int_0^\infty \int_0^\varepsilon \pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\beta\, d\alpha \\ &\le \int_0^\infty \int_0^\varepsilon \pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}+\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}-2\right)\right\}d\beta\, d\alpha \\ &\le \int_0^\infty \int_0^\varepsilon \pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}+\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{2\beta}\right)\right\}d\beta\, d\alpha.\end{aligned}$$

Letting $k = \prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}} + \left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]$, it follows that

$$\begin{aligned}
P_1 &\leq \int_0^\infty \int_0^\infty \pi(\alpha)\,\alpha^{-n}\beta^{-\frac{n}{2}-1} k \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{2\beta}\right)\right\} d\beta\, d\alpha \\
&= \int_0^\infty \pi(\alpha)\,\alpha^{-n} k \int_0^\infty \beta^{-\frac{n}{2}-1} \exp\left\{-\frac{1}{\beta\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{4}\right)\right\} d\beta\, d\alpha \\
&= k\int_0^\infty \pi(\alpha)\,\alpha^{-n} \frac{\Gamma\left(\frac{n}{2}\right)}{\left[\frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{4}\right)\right]^{\frac{n}{2}}} d\alpha \\
&\propto \int_0^\infty \pi(\alpha)\,\alpha^{-n} \frac{1}{\alpha^{-n}} d\alpha.
\end{aligned}$$

Hence,

$$P_1 \propto \int_0^\infty \pi(\alpha)\, d\alpha < \infty,$$

since $\pi(\alpha)$is integrable.

Next, it is required to show that $P_2 < \infty$.

Since $\pi(\alpha)$ is integrable, a constant $c_2 > 0$ exists such that $\pi(\alpha) < c_2$. Then,

$$\begin{aligned}
P_2 &< \int_\varepsilon^{\frac{1}{\varepsilon}} \int_0^\infty \beta^{-n-1} c_2 \prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\} d\alpha\, d\beta \\
&= \int_\varepsilon^{\frac{1}{\varepsilon}} \frac{c_2 c_3 \prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}} d\beta < \infty,
\end{aligned}$$

where $c_3 = \frac{\Gamma\left(\frac{n}{2}\right)}{2\left[\frac{1}{2}\right]^{\frac{n}{2}}}$. For $\varepsilon < min\left(1, \frac{1}{t_i}, \sum_{i=1}^{n}\frac{t_i}{4n}\right)$,the results follows.

Next, it is required to show that $P_3 < \infty$.

$$P_3 = \int_0^\infty \int_{\frac{1}{\varepsilon}}^\infty \pi(\alpha)\,\alpha^{-n}\beta^{-n-1} \prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\} d\beta\, d\alpha.$$

Note that if $0 < \varepsilon <$ then, $1 < \frac{1}{\varepsilon} < \infty$. For large beta,

$$\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \le \prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] = \prod_{i=1}^{n}\left[2\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right].$$

It is obvious that,

$$\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right) \ge \left(\frac{\beta}{t_i}-2\right).$$

Thus,

$$\begin{aligned}
P_3 &\le \int_0^\infty \int_{\frac{1}{\varepsilon}}^\infty \pi(\alpha)\,\alpha^{-n}\beta^{-n-1}\prod_{i=1}^{n}\left[2\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{t_i}-2\right)\right\}d\beta\, d\alpha \\
&\propto \int_0^\infty \int_{\frac{1}{\varepsilon}}^\infty \pi(\alpha)\,\alpha^{-n}\beta^{-n+\frac{3n}{2}-1}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{t_i}-2\right)\right\}d\beta\, d\alpha \\
&\le \int_0^\infty \int_{\frac{1}{\varepsilon}}^\infty \pi(\alpha)\,\alpha^{-n}\beta^{-n+\frac{3n}{2}-1}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{2t_i}\right)\right\}d\beta\, d\alpha,
\end{aligned}$$

since

$$\exp\left\{\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{t_i}-2\right)\right\} \ge \exp\left\{\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{2t_i}\right)\right\}$$

implies that,

$$\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{t_i}-2\right)\right\} \le \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta}{2t_i}\right)\right\}.$$

Now

$$\begin{aligned}
P_3 &< \int_0^\infty \int_0^\infty \pi(\alpha)\,\alpha^{-n}\beta^{\frac{n}{2}-1}\exp\left\{-\frac{\beta}{\alpha^2}\sum_{i=1}^{n}\left(\frac{1}{4t_i}\right)\right\}d\beta\, d\alpha \\
&= \int_0^\infty \pi(\alpha)\,\alpha^{-n}\frac{\Gamma\left(\frac{n}{2}\right)}{\left[\frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{1}{4t_i}\right)\right]^{\frac{n}{2}}}d\beta\, d\alpha,
\end{aligned}$$

using Gamma function.

$$P_3 \propto \int_0^\infty \pi(\alpha)\,\alpha^{-n}\alpha^n d\alpha$$
$$= \int_0^\infty \pi(\alpha)\,d\alpha < \infty.$$

since $\pi(\alpha)$ is integrable. This completes the proof. □

Theorem 4.5. *The posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.12 when there type I right censoring is a proper distribution.*

Proof. The posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.12 when there is type I right censoring is given by

$$\pi_1(\alpha,\beta|\underline{t}) \propto \pi(\beta)\,\alpha^{-r-1}\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}$$
$$\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.$$

To show that the posterior is proper, this must be true

$$\int_0^\infty\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,c\,d\alpha d\beta = 1,$$

where c is the normalising constant.

For the above to be true, the following must be shown

$$\int_0^\infty\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,d\alpha d\beta < \infty.$$

Now

$$\int_0^\infty\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,d\alpha d\beta \propto \int_0^\infty\int_0^\infty \pi(\beta)\,\alpha^{-r-1}\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}$$
$$\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}d\alpha d\beta.$$
$$\leq \int_0^\infty\int_0^\infty \pi(\beta)\,\alpha^{-r-1}\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha d\beta.$$

Taking integration of $\pi_1(\alpha,\beta|\underline{t})$ with respect to α, it is obtained that

$$\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,d\alpha \le \int_0^\infty \pi(\beta)\,\alpha^{-r-1}\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha$$

$$=\pi(\beta)\,\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\int_0^\infty \alpha^{-r-1}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha.$$

By letting $\lambda=\alpha^2$, the integral is transformed to

$$\int_0^\infty \pi_1(\alpha,\beta|\underline{t})\,d\alpha \propto \pi(\beta)\,\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\frac{1}{2}\int_0^\infty \left(\lambda^{\frac{1}{2}}\right)^{-r-1}\lambda^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\lambda$$

$$\propto \pi(\beta)\,\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\int_0^\infty \lambda^{-\frac{r}{2}-1}\exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\lambda$$

$$=\frac{\Gamma\left(\frac{r}{2}\right)\pi(\beta)\,\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\frac{1}{2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}}$$

$$=\frac{c_1\pi(\beta)\,\beta^{-r}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{r}{2}}}=H(\beta),$$

where $c_1=\frac{\Gamma\left(\frac{r}{2}\right)}{\left[\frac{1}{2}\right]^{\frac{r}{2}}}$. It remains to be shown that

$$\int_0^\infty H(\beta)\,d\beta<\infty.$$

By examining the limits

$$\lim_{\beta\to 0}\frac{\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^r\left[\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{r}{2}}}=\lim_{\beta\to 0}\frac{\beta^{\frac{r}{2}}\prod_{i=1}^{r}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}+\frac{\beta}{t_i^{\frac{3}{2}}}\right]}{\beta^r\left[\beta^{-1}\sum_{i=1}^{n}\left(t_i+\frac{\beta^2}{t_i}-2\beta\right)\right]^{\frac{r}{2}}}$$

$$=\lim_{\beta\to 0}\frac{\beta^{\frac{r}{2}}\prod_{i=1}^{r}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}+\frac{\beta}{t_i^{\frac{3}{2}}}\right]}{\beta^{n-\frac{r}{2}}\left[\sum_{i=1}^{r}\left(t_i+\frac{\beta^2}{t_i}-2\beta\right)\right]^{\frac{r}{2}}}=\frac{\prod_{i=1}^{r}\left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}}\right]}{\left[\sum_{i=1}^{r}(t_i)\right]^{\frac{r}{2}}}.$$

$$\lim_{\beta\to\infty}\frac{\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^r\left[\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{r}{2}}}=\lim_{\beta\to\infty}\frac{\beta^{\frac{3r}{2}}\prod_{i=1}^{r}\left[\frac{1}{\beta t_i^{\frac{1}{2}}}+\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^{r+\frac{r}{2}}\left[\sum_{i=1}^{r}\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{2}{\beta}\right)\right]^{\frac{n}{2}}}$$

$$=\frac{\prod_{i=1}^{r}\left[\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{r}\left(\frac{1}{t_i}\right)\right]^{\frac{r}{2}}}.$$

Hence as $\beta\to 0$ or $\beta\to\infty$, $H(\beta)=\mathcal{O}(\pi(\beta))$.

Since $\pi(\beta)$ is integrable, then

$$\int_0^\infty H(\beta)\,d\beta<\infty,$$

proving that the posterior is proper. □

Theorem 4.6. *The posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.13 when there is type I right censoring is a proper distribution.*

Proof. The posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the prior in equation 4.13 when there is type I right censoring is given by

$$\pi_2(\alpha,\beta|\underline{t})\propto\pi(\alpha)\,\alpha^{-r}\beta^{-r-1}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}$$
$$\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}.$$

To show that the posterior is proper, this must be true

$$\int_0^\infty \int_0^\infty c\pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha = 1,$$

where c is the normalising constant.

For this to hold, it must be shown that

$$\int_0^\infty \int_0^\infty \pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha < \infty.$$

Now

$$\begin{aligned}\int_0^\infty \int_0^\infty \pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha &\propto \int_0^\infty \int_0^\infty \pi(\alpha)\,\alpha^{-r}\beta^{-r-1}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}\\ &\quad\times\prod_{i=r+1}^{n}\left\{1-\Phi\left[\frac{1}{\alpha}\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}-\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\}\right]\right\}d\alpha\, d\beta\\ &\leq \int_0^\infty \int_0^\infty \pi(\alpha)\,\alpha^{-r}\beta^{-r-1}\prod_{i=1}^{r}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\\ &\quad\times\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha\, d\beta.\end{aligned}$$

For any $0<\varepsilon<1$, the integral can be decomposed as

$$\begin{aligned}P=\int_0^\infty \int_0^\infty \pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha &= \int_0^\infty \int_0^\varepsilon \pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha+\int_0^\infty \int_\varepsilon^{\frac{1}{\varepsilon}} \pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha+\int_0^\infty \int_{\frac{1}{\varepsilon}}^\infty \pi_2(\alpha,\beta|\underline{t})\, d\beta\, d\alpha\\ &= P_1+P_2+P_3.\end{aligned}$$

Firstly it must be shown that $P_1 < \infty$. Thus, for any $\varepsilon < \min\left(1, \sum_{i=1}^{r} \frac{t_i}{4r}\right)$,

$$\begin{aligned}
P_1 &= \int_0^\infty \int_0^\varepsilon \pi(\alpha)\alpha^{-r}\beta^{-r-1} \prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)\right\} d\beta d\alpha \\
&\le \int_0^\infty \int_0^\varepsilon \pi(\alpha)\alpha^{-r}\beta^{-r-1} \prod_{i=1}^{r} \left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}} + \left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta} - 2\right)\right\} d\beta d\alpha \\
&\le \int_0^\infty \int_0^\varepsilon \pi(\alpha)\alpha^{-r}\beta^{-r-1} \prod_{i=1}^{r} \left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}} + \left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right] \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{2\beta}\right)\right\} d\beta d\alpha.
\end{aligned}$$

Letting $k = \prod_{i=1}^{r} \left[\left(\frac{1}{t_i}\right)^{\frac{1}{2}} + \left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]$, it follows that

$$\begin{aligned}
P_1 &\le \int_0^\infty \int_0^\infty \pi(\alpha)\alpha^{-r}\beta^{-\frac{r}{2}-1} k \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{2\beta}\right)\right\} d\beta d\alpha \\
&= \int_0^\infty \pi(\alpha)\alpha^{-r} k \int_0^\infty \beta^{-\frac{r}{2}-1} \exp\left\{-\frac{1}{\beta\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{4}\right)\right\} d\beta d\alpha \\
&= k \int_0^\infty \pi(\alpha)\alpha^{-r} \frac{\Gamma\left(\frac{r}{2}\right)}{\left[\frac{1}{\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{4}\right)\right]^{\frac{r}{2}}} d\alpha \\
&\propto \int_0^\infty \pi(\alpha)\alpha^{-r}\frac{1}{\alpha^{-r}} d\alpha.
\end{aligned}$$

Hence,

$$P_1 \propto \int_0^\infty \pi(\alpha)\, d\alpha < \infty,$$

since $\pi(\alpha)$ is integrable. Next, it is required to show that $P_2 < \infty$.

Since $\pi(\alpha)$ is integrable, a constant $c_2 > 0$ exists such that $\pi(\alpha) < c_2$, then this holds

$$P_2 < \int_{\varepsilon}^{\frac{1}{\varepsilon}} \int_{0}^{\infty} \beta^{-r-1} c_2 \prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right] \exp\left\{ -\frac{1}{2\alpha^2} \sum_{i=1}^{r} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) \right\} d\alpha d\beta$$

$$= \int_{\varepsilon}^{\frac{1}{\varepsilon}} \frac{c_2 c_3 \prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right]}{\left[\sum_{i=1}^{r} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) \right]^{\frac{r}{2}}} d\beta < \infty,$$

where $c_3 = \frac{\Gamma\left(\frac{r}{2}\right)}{2\left[\frac{1}{2}\right]^{\frac{r}{2}}}$. For $\varepsilon < min\left(1, \frac{1}{t_i}, \sum_{i=1}^{r} \frac{t_i}{4r}\right)$, the results follows.

Next, it is required to show that $P_3 < \infty$.

$$P_3 = \int_{0}^{\infty} \int_{\frac{1}{\varepsilon}}^{\infty} \pi(\alpha) \alpha^{-r} \beta^{-r-1} \prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right] \exp\left\{ -\frac{1}{2\alpha^2} \sum_{i=1}^{r} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) \right\} d\beta d\alpha.$$

Note that if $0 < \varepsilon <$ then, $1 < \frac{1}{\varepsilon} < \infty$. For large beta,

$$\prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right] \leq \prod_{i=1}^{r} \left[\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right]$$

$$= \prod_{i=1}^{r} \left[2\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right].$$

It is obvious that,

$$\left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2 \right) \geq \left(\frac{\beta}{t_i} - 2 \right).$$

Thus,

$$P_3 \leq \int_{0}^{\infty} \int_{\frac{1}{\varepsilon}}^{\infty} \pi(\alpha) \alpha^{-r} \beta^{-r-1} \prod_{i=1}^{r} \left[2\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}} \right] \exp\left\{ -\frac{1}{2\alpha^2} \sum_{i=1}^{r} \left(\frac{\beta}{t_i} - 2 \right) \right\} d\beta d\alpha$$

$$\propto \int_{0}^{\infty} \int_{\frac{1}{\varepsilon}}^{\infty} \pi(\alpha) \alpha^{-r} \beta^{-r+\frac{3r}{2}-1} \exp\left\{ -\frac{1}{2\alpha^2} \sum_{i=1}^{r} \left(\frac{\beta}{t_i} - 2 \right) \right\} d\beta d\alpha$$

$$\leq \int_{0}^{\infty} \int_{\frac{1}{\varepsilon}}^{\infty} \pi(\alpha) \alpha^{-r} \beta^{-r+\frac{3r}{2}-1} \exp\left\{ -\frac{1}{2\alpha^2} \sum_{i=1}^{r} \left(\frac{\beta}{2t_i} \right) \right\} d\beta d\alpha,$$

since

$$\exp\left\{\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{\beta}{t_i}-2\right)\right\} \geq \exp\left\{\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{\beta}{2t_i}\right)\right\}$$

implies that,

$$\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{\beta}{t_i}-2\right)\right\} \leq \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{\beta}{2t_i}\right)\right\}.$$

Now

$$\begin{aligned} P_3 &< \int_0^\infty\int_0^\infty \pi(\alpha)\,\alpha^{-r}\beta^{\frac{r}{2}-1}\exp\left\{-\frac{\beta}{\alpha^2}\sum_{i=1}^{r}\left(\frac{1}{4t_i}\right)\right\}d\beta d\alpha \\ &= \int_0^\infty \pi(\alpha)\,\alpha^{-r}\frac{\Gamma\left(\frac{r}{2}\right)}{\left[\frac{1}{\alpha^2}\sum_{i=1}^{r}\left(\frac{1}{4t_i}\right)\right]^{\frac{r}{2}}}d\beta d\alpha, \end{aligned}$$

using Gamma function.

$$\begin{aligned} P_3 &\propto \int_0^\infty \pi(\alpha)\,\alpha^{-r}\alpha^{r}d\alpha \\ &= \int_0^\infty \pi(\alpha)\,d\alpha < \infty. \end{aligned}$$

since $\pi(\alpha)$ is integrable. This completes the proof. □

4.7 Assessment of Posterior Convergence through MCMC for the Birnbaum-Saunders Distribution

This section outlines the procedure used to evaluate the performance of Bayesian inference for the Birnbaum-Saunders distribution under two weak informative priors. The study focuses on estimating the shape and scale parameters of the Birnbaum-Saunders distribution using Markov Chain Monte Carlo techniques under different sample sizes, $n = 10, n = 40$, $n = 200$. Simulations were conducted under no censoring or complete data, and type I right censoring. The simulated datasets were generated using various combinations of the true Birnbaum-Saunders parameters: shape parameters, $\alpha \in \{0.1, 0.25, 0.5\}$ and scale parameter, $\beta \in \{0.2, 1, 0.75\}$ to evaluate the performance of Bayesian inference across different distribution shapes and scales. By considering combinations of α and β, the study examines estimator behaviour and posterior performance under diverse scenarios, ensuring that the results are robust and not specific to a single parameter setting. The hyperparameters for the prior distributions were chosen to be 10^{-3}, representing weakly informative settings. A total of $T = 20000$ iterations were performed for each simulation, with a burn-in period of $T/2$. The convergence of the chains was monitored using the trace plot, the Geweke diagnostic Z-score, and the Gelman–Rubin diagnostic. See Appendix C.2 for the R code.

Algorithm. 4.1 *Simulation procedure for the Birnbaum–Saunders model using the prior in equation 4.12.*

- Step 0

  Simulate t from Birnbaum-Saunders distribution with true parameters and fix the sample size n.

- Step 1

  Initialize

  Choose $\alpha^{(0)}$, $\beta^{(0)} > 0$, tuning parameter σ, iterations T.

- Step 2

  Sample $\alpha|\beta,\underline{t}$ (Gibbs step)

  Compute $S = \sum_{i=1}^{n} \left(\frac{t_i}{\beta} + \frac{\beta}{t_i} - 2\right)$

  Draw $u \sim Gamma\left(\frac{n}{2}, rate = S/2\right)$ and set $\alpha = 1/\sqrt{u}$

- Step 3

  Update $\beta \mid \alpha, t$

  Propose

$$\log(\beta^*) \sim N\left(\log(\beta), \sigma^2\right)$$

  Compute the log-posterior kernel for β

$$\log(\pi(\beta|\alpha,\underline{t})) = -(n+a_2+1)\log(\beta) - \frac{b_2}{\beta} + \sum_{i=1}^{n} \log\left\{\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right\} - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i}\right).$$

  Accept $\beta^*$with probability

$$min\{1, \exp[\log(\pi(\beta^*|\alpha,\underline{t})) - \log(\pi(\beta|\alpha,\underline{t}))]\}.$$

- Step 4

Repeat

Iterate Steps 2–3 for $\tau = 1,...,T$. After discarding burn-in iterations, use the remaining samples to estimate the posterior distributions of α and β.

When using the prior in equation 4.13, Substitute $u \sim Gamma\left(a_1 + \frac{n}{2}, rate = b_1 + S/2\right)$ and

$$\log\left(\pi\left(\beta|\alpha,\underline{t}\right)\right) = (n+1)\log\left(\beta\right) + \sum_{i=1}^{n}\log\left\{\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right\} - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta} + \frac{\beta}{t_i}\right).$$

Algorithm. 4.2*Simulation procedure for the Birnbaum-Saunders model using the prior in 4.12 under type I right censoring.*

- Step 0

Simulate a complete datase t_{full} from the Birnbaum-Saunders distribution with true parameters and fix the sample size n.

- Step 1

Determine the type I censoring threshold C as the 70th percentile of the simulated data

$$C = quantile\left(t_{full}, 0.7\right).$$

- Step 2

Split the data into

t_{fail}=$\{t_i : t_i \leq C\}$

t_{cens}=$\{t_i : t_i > C\}$.

$r = length\left(t_{fail}\right)$.

- Step 3

Initialize

Choose $\alpha^{(0)}$,$\beta^{(0)}$, iteration T.

- Step 4

Compute the log-posterior kernel for both α and β

$$\log\left(\pi\left(\alpha,\beta|\underline{t}\right)\right) = -\left(a+r+1\right)\log\left(\beta\right) - \left(r+1\right)\log\left(\alpha\right) - \frac{b}{\beta} + \sum_{i=1}^{r}\log\left\{\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}} + \left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right\}$$
$$-\frac{1}{2\alpha^2}\sum_{i=1}^{r}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right) + \sum\left(\log\left(1-\Phi\left(*\right)\right)\right),$$

where $\Phi\left(*\right) = pnorm\left(\left(\frac{1}{\alpha}\right)\times\left(\left(\frac{t_{cens}}{\beta}\right)^{\frac{1}{2}} - \left(\frac{t_{cens}}{\beta}\right)^{-\frac{1}{2}}\right)\right)$.

- Step 5

  Propose

$$\log\left(\alpha^*,\beta^*\right) \sim MvN\left(\begin{array}{c}\log\left(\alpha^*\right)\\ \log\left(\beta^*\right)\end{array},\Sigma\right),$$

  where Σ is a covariance matrix.

- Step 6

  Acceptance Region

  Accept α and β with probability

$$\min\left\{1,\ \exp\left(\log\frac{\pi\left(\alpha^*,\beta^*|\underline{t}\right)}{\pi\left(\alpha,\beta|\underline{t}\right)}\right)\right\}.$$

- Step 7

  Iteration

  Repeat Steps 4–6 for $\tau = 1,2,...,T$ to obtain posterior samples for α and β.

4.7.1 Complete Data

This section presents the posterior distribution for the Birnbaum-Saunders with parameters α and β when prior choices $\pi_1\left(\alpha,\beta\right)$ and $\pi_2\left(\alpha,\beta\right)$ are used under no censoring. The analysis was conducted for sample sizes $n = 10$, $n = 40$, $n = 200$, for each case, the trace plots and the posterior densities, and the convergence diagnostics (Gelman-Rubin and Geweke) will be evaluated.

Figures 4.15, 4.16, and 4.17 present the trace plots and posterior density functions for α and β using the prior in equation 4.12 for $n = 10$, $n = 40$, and $n = 200$, respectively. The figures illustrate that the chains mix and pass the thick pen test, and the posterior densities become more concentrated around their means.

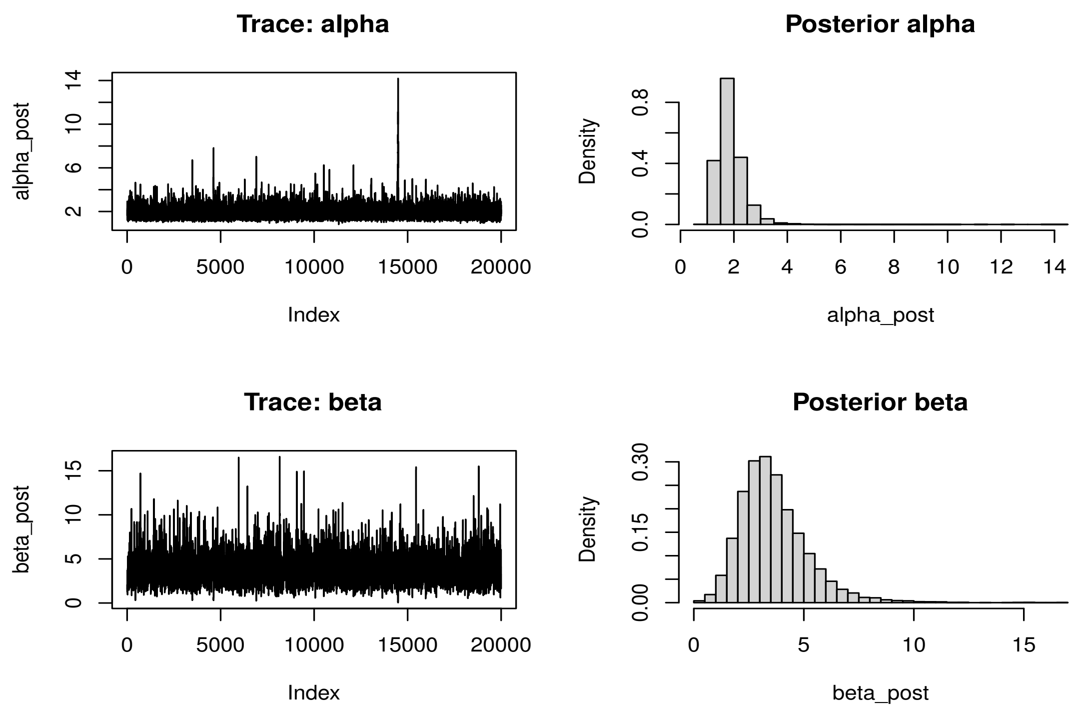

Figure 4.15: Trace plot and probability density function of α and β when $n = 10$.

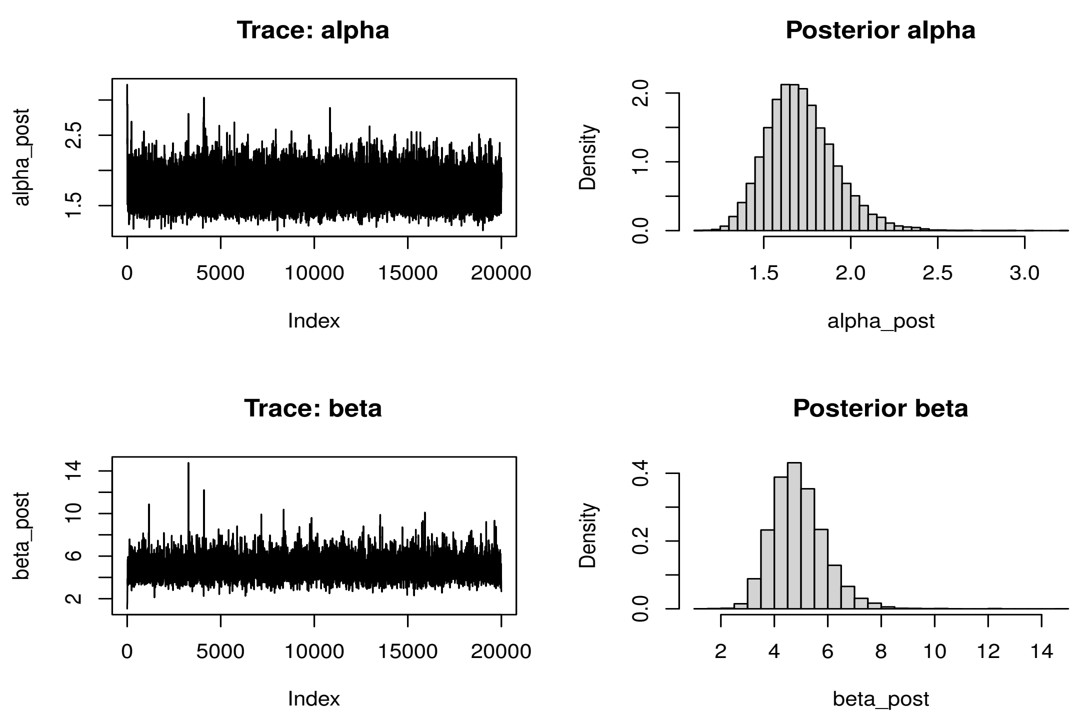

Figure 4.16: Trace plot and probability density function of α and β when $n = 40$.

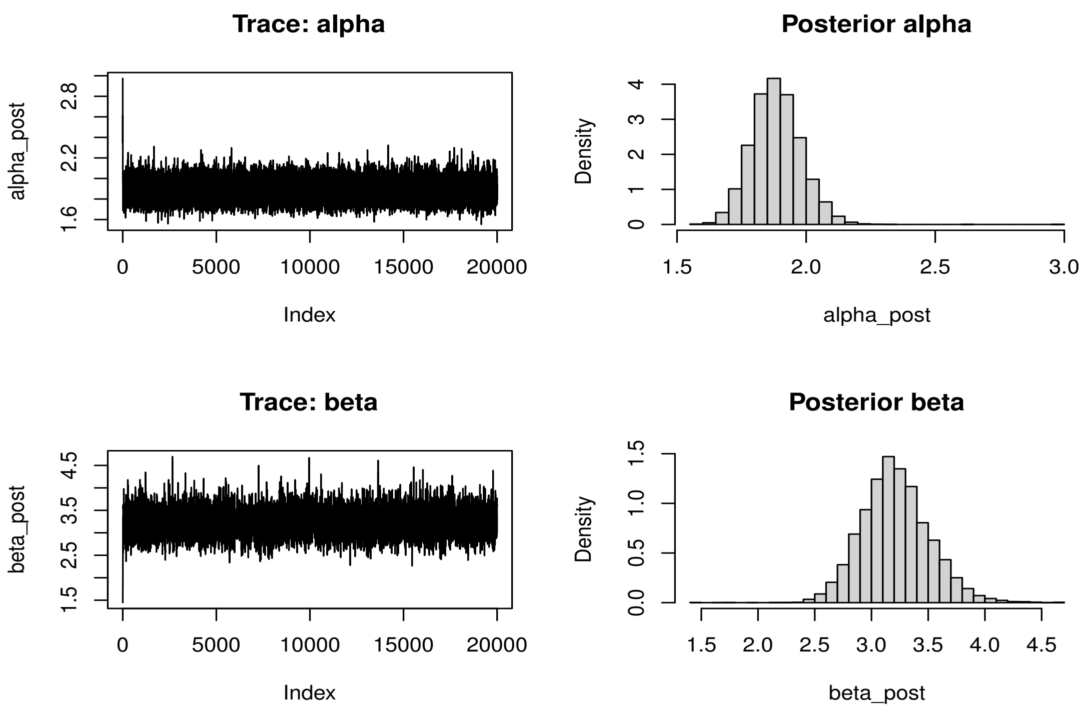

Figure 4.17: Trace plot and probability density function of α and β when $n = 200$.

Figures 4.18, 4.19, and 4.20 present the trace plots and posterior density functions for α and β when the prior in equation 4.13 is used for $n = 10$, $n = 40$, and $n = 200$, respectively. The trace plots for both α and β show that chains mix well and pass the thick pen test, which means that convergence has been reached. It is observed that as the sample size increases, there is no constant trend.

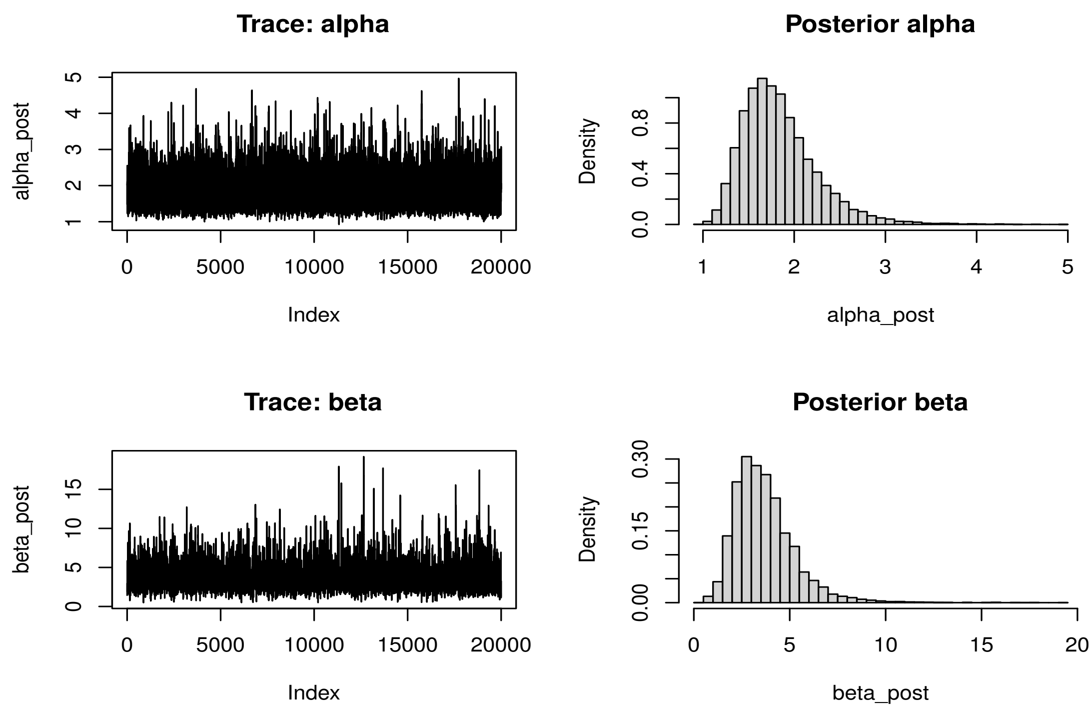

Figure 4.18: Trace plot and probability density function of α and β when $n = 10$.

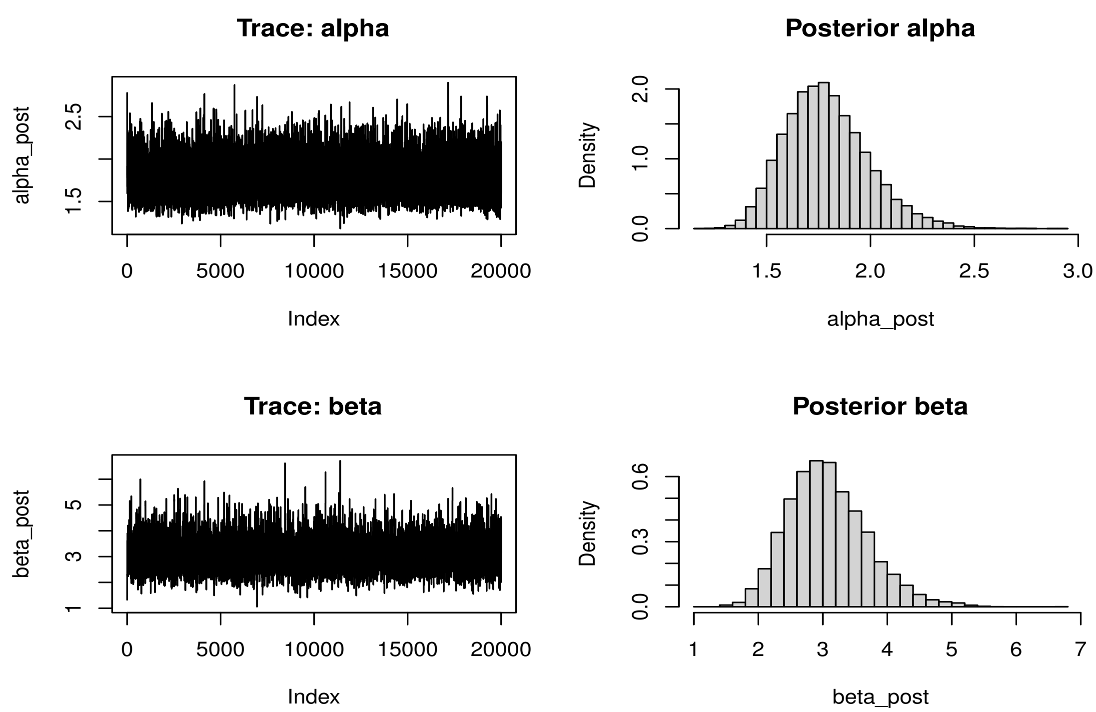

Figure 4.19: Trace plot and probability density function of α and β when $n = 40$.

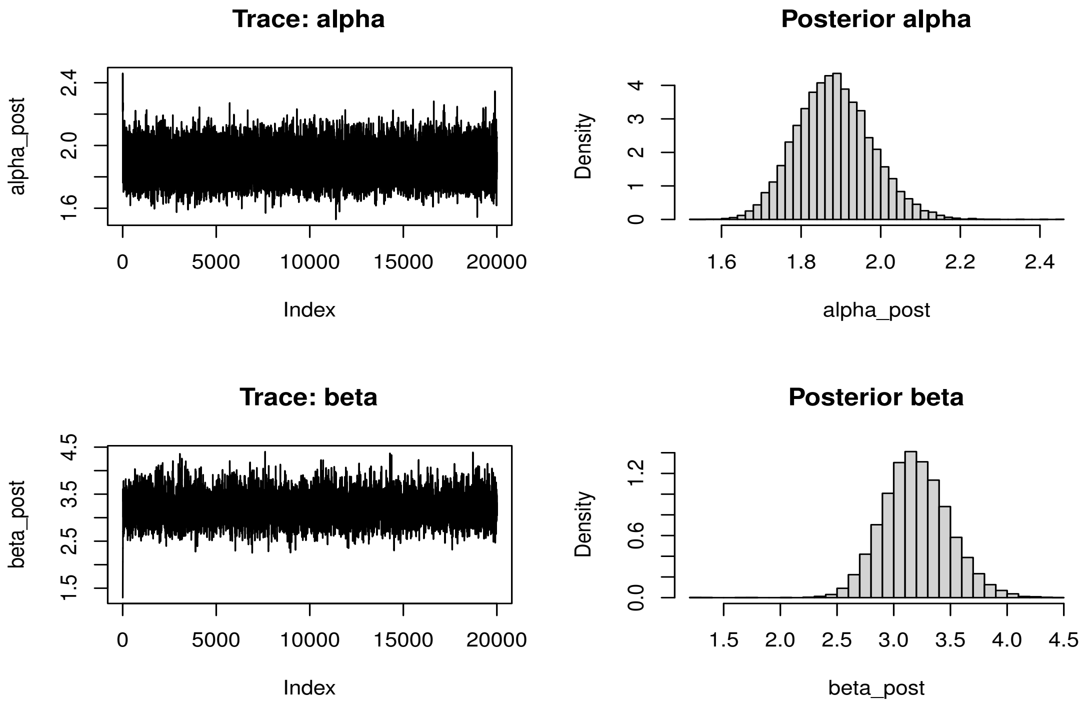

Figure 4.20: Trace plot and probability density function of α and β when $n = 200$.

Figures 4.21, 4.22, and 4.23 present the Gelman-Rubin plots for α and β using the prior in equation 4.12 for $n = 10$, $n = 40$, and $n = 200$, respectively. The plots indicate that the $\widehat{R}$ values approach 1 for both α and β, confirming satisfactory convergence of the chains across all sample sizes.

Figures 4.24, 4.25, and 4.26 present the Gelman-Rubin plots for α and β when the prior in equation 4.13 is used for $n = 10$, $n = 40$, and $n = 200$, respectively. The plots indicate that the $\widehat{R}$ values approach 1 for both α and β, confirming satisfactory convergence of the chains across all sample sizes.

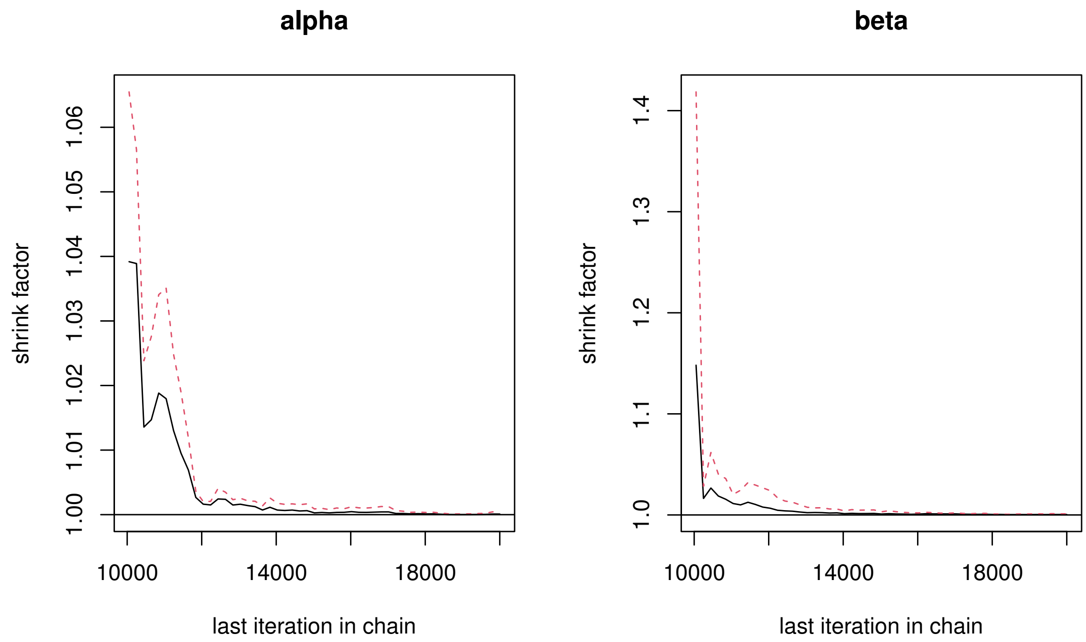

Figure 4.21: The Gelman-Rubin plots for α and β when $n = 10$.

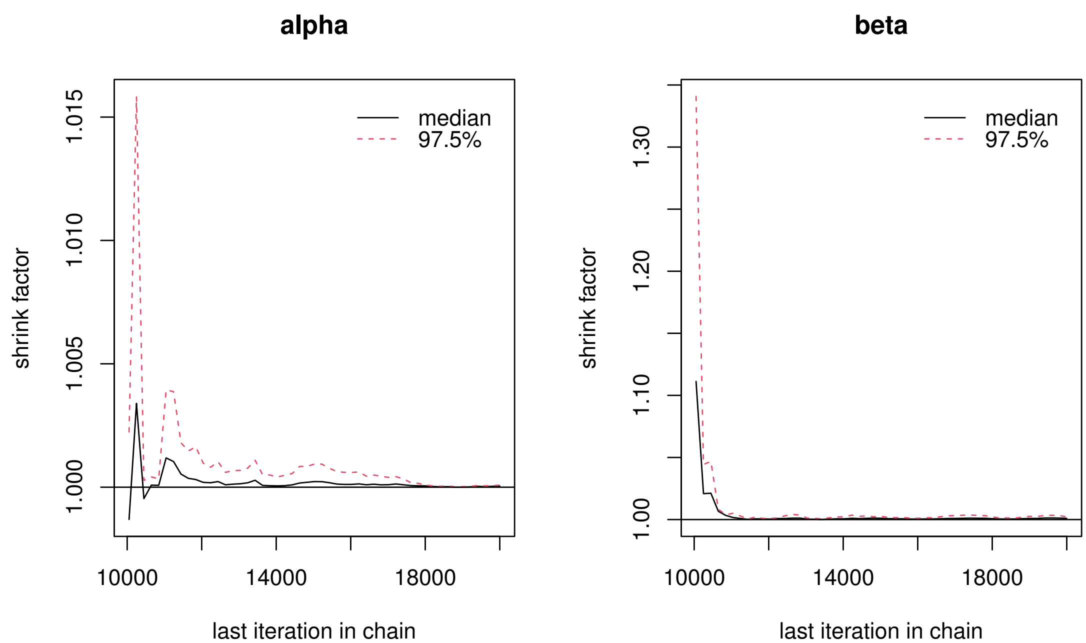

Figure 4.22: The Gelman-Rubin plots for α and β when $n = 40$.

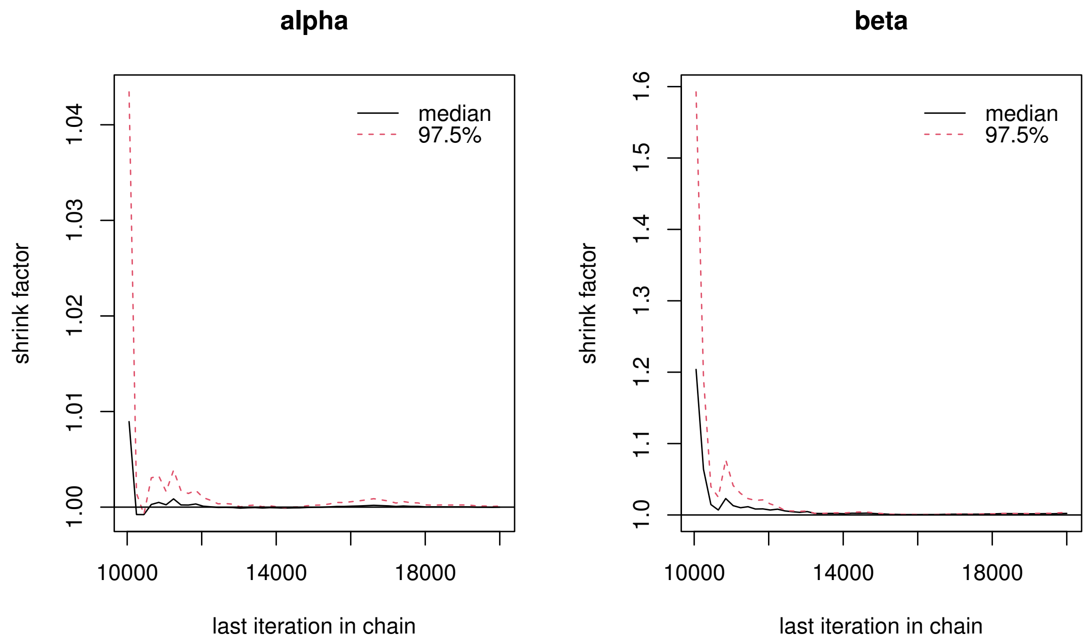

Figure 4.23: The Gelman-Rubin plots for α and β when $n = 200$.

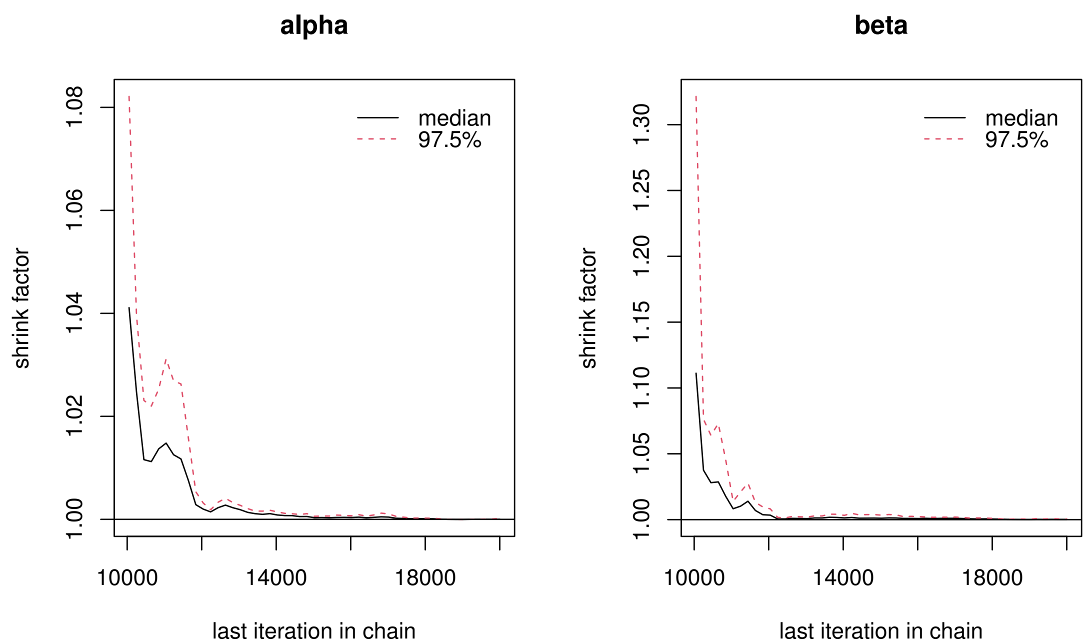

Figure 4.24: The Gelman-Rubin plots for α and β when $n = 10$.

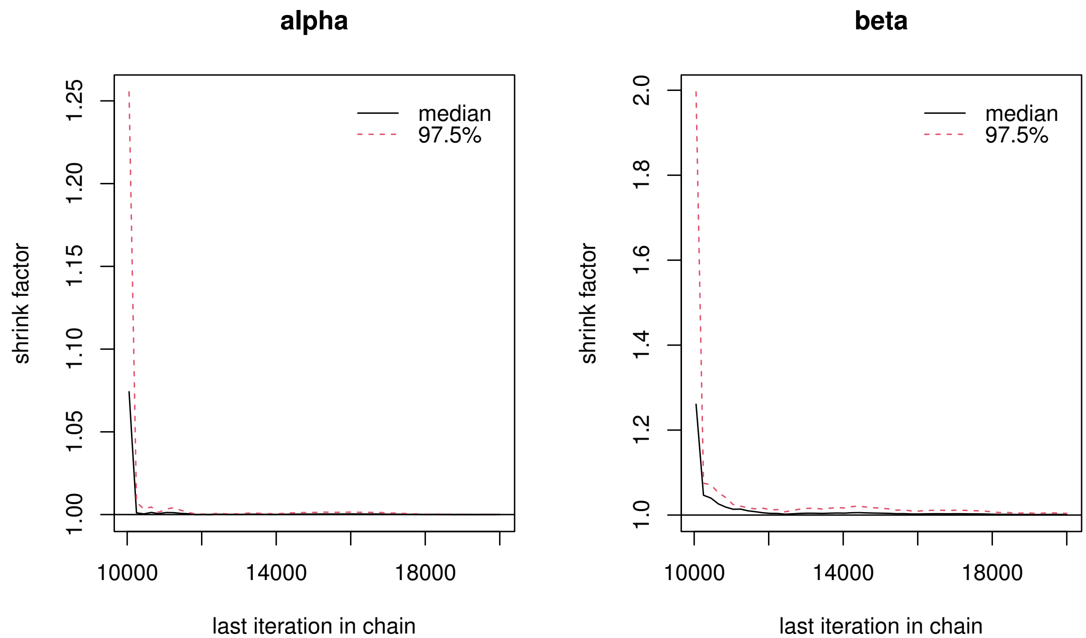

Figure 4.25: The Gelman-Rubin plots for α and β when $n = 40$.

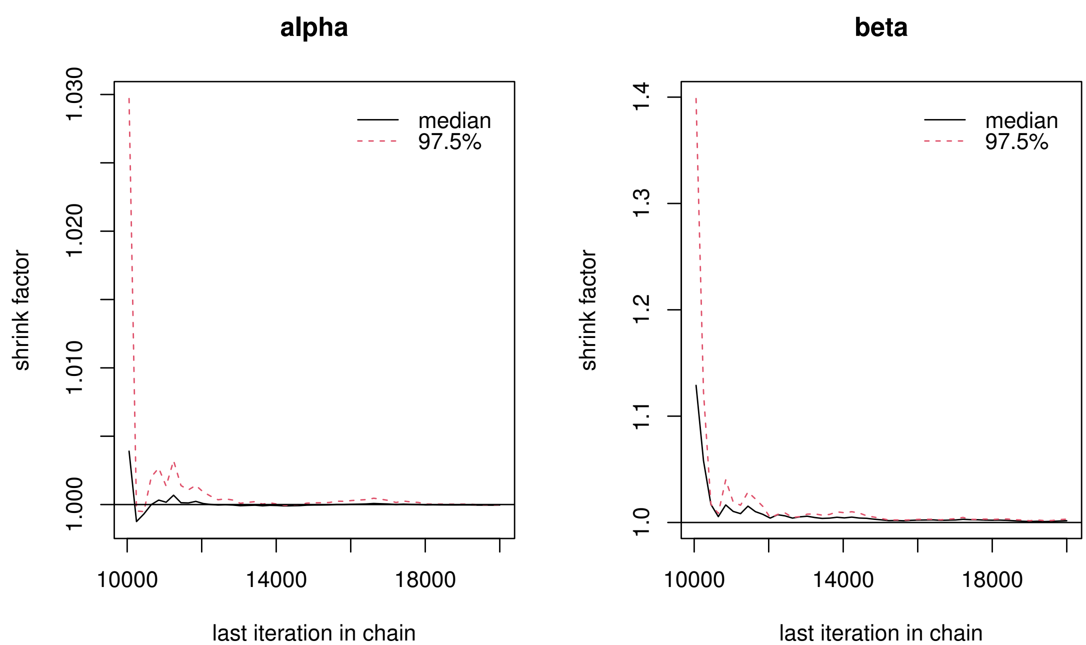

Figure 4.26: The Gelman-Rubin plots for α and β when $n = 200$.

Figures 4.27, 4.28, and 4.29 present the Geweke plots for α and β when the prior in equation 4.12 is used for $n = 10$, $n = 40$, and $n = 200$, respectively. Figures 4.30, 4.31, and 4.32 present the Geweke plots for α and β when the prior in equation 4.13 is used for $n = 10$, $n = 40$, and $n = 200$, respectively.

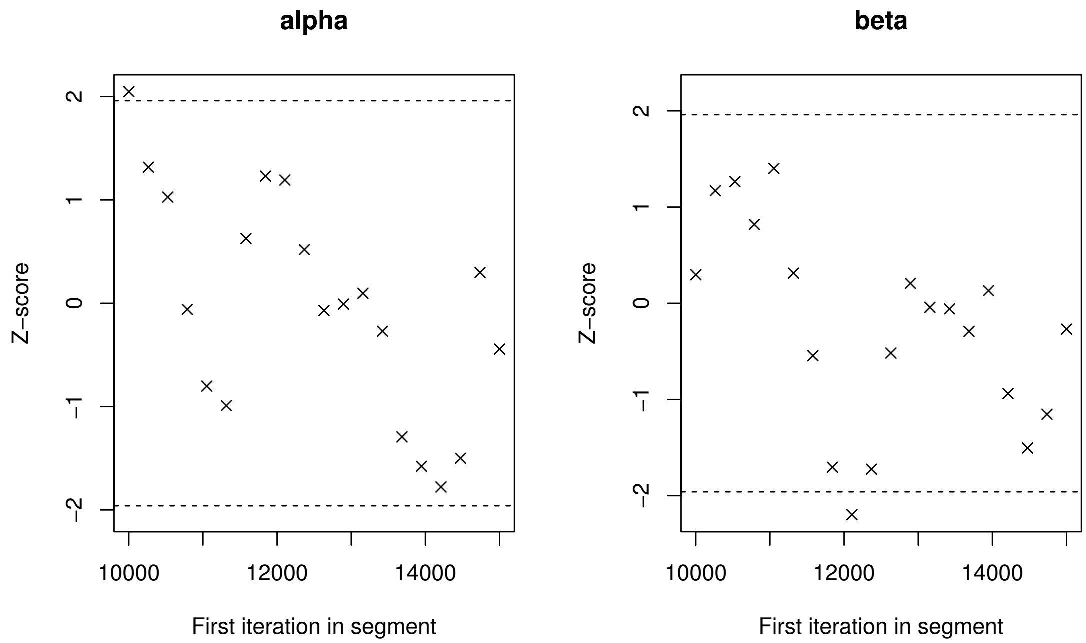

Figure 4.27: The Geweke plots for α and β when $n = 10$.

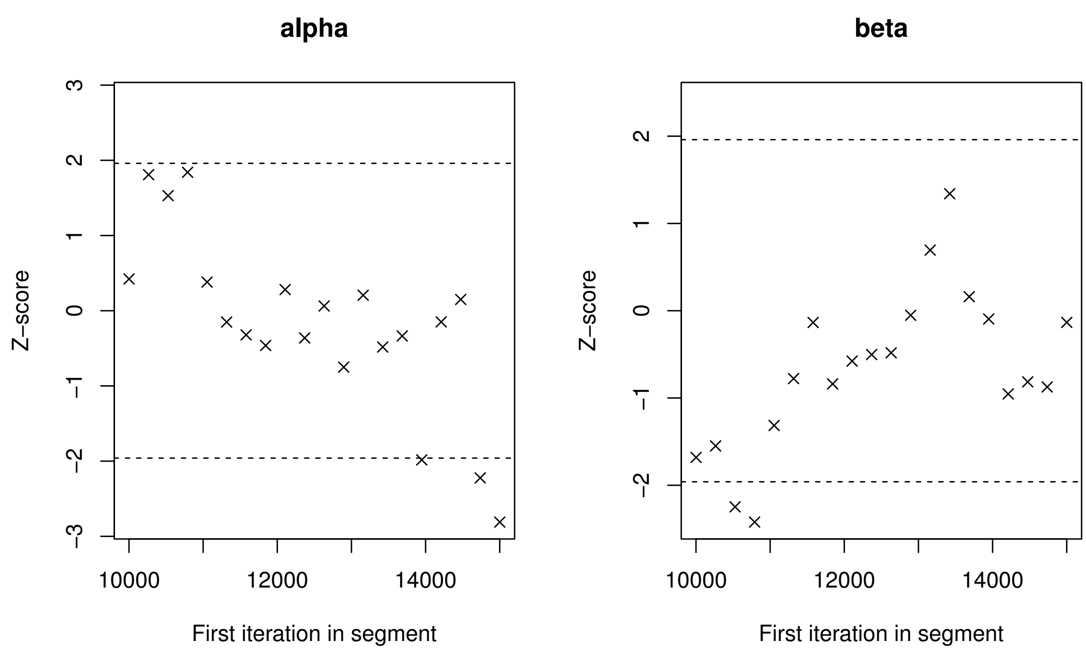

Figure 4.28: The Geweke plots for α and β when $n = 40$.

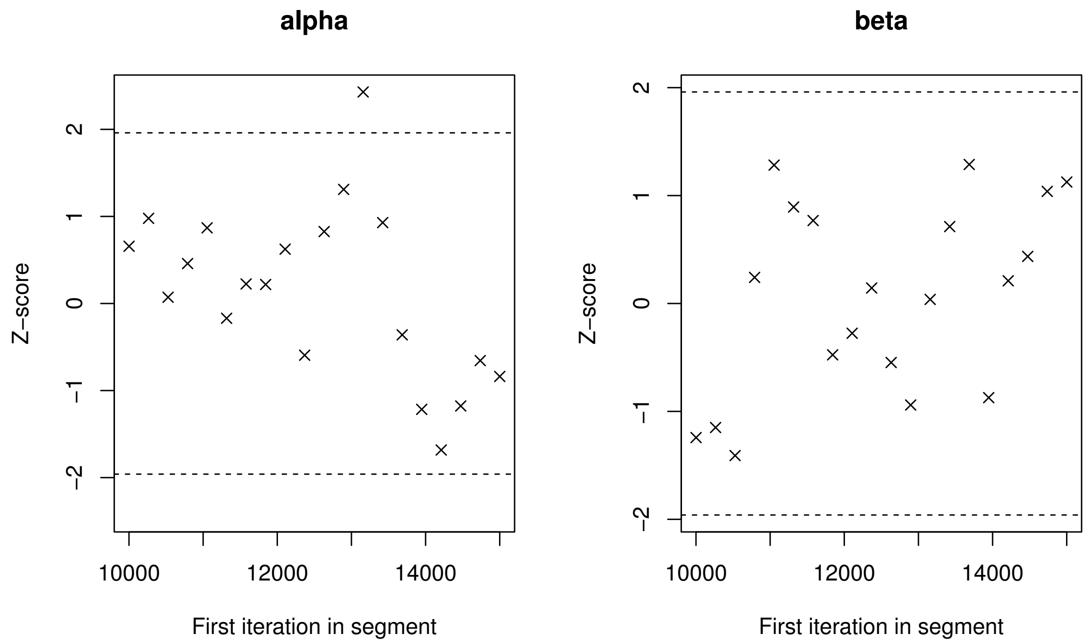

Figure 4.29: The Geweke plots for α and β when $n = 200$.

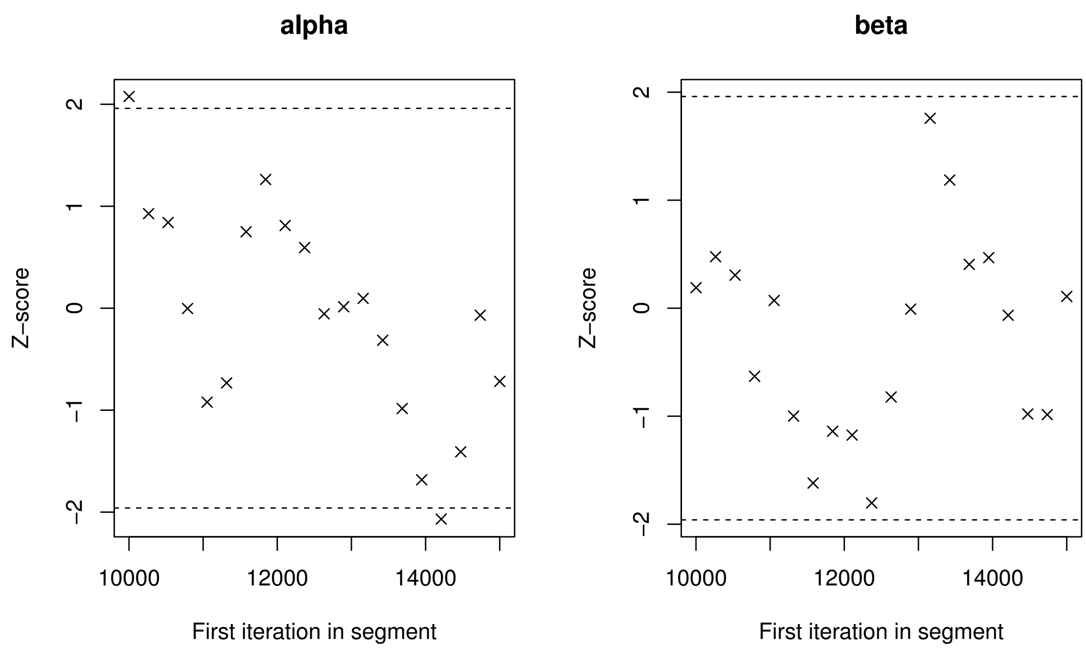

Figure 4.30: The Geweke plots for α and β when $n = 10$.

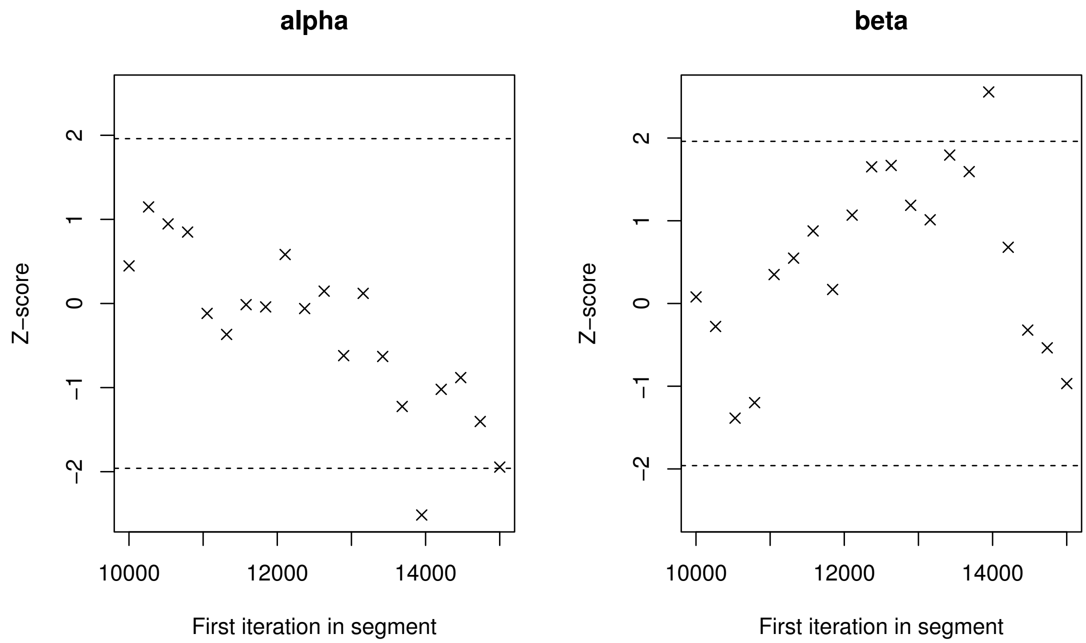

Figure 4.31: The Geweke plots for α and β when $n = 40$.

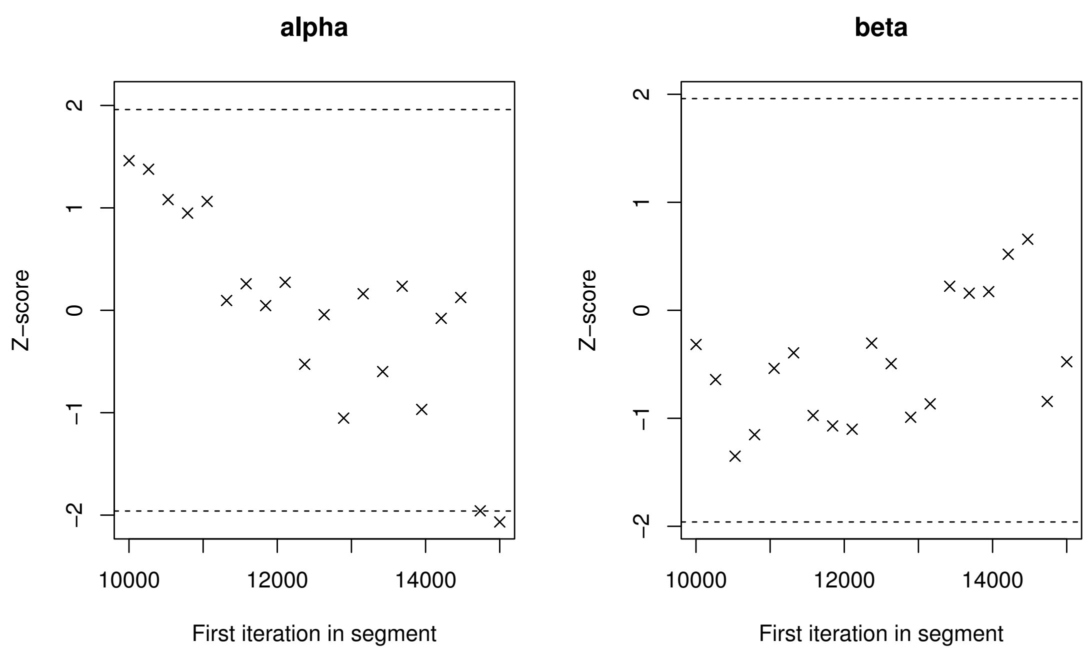

Figure 4.32: The Geweke plots for α and β when $n = 200$.

The Z-score for all sample sizes $(n = 10, 40, 200)$ mostly lie within the ± 2 range, indicating that the chains have reached convergence for when both priors from equation 4.12 and equation 4.13 are used.

4.7.2 Type I Right Censoring

This section presents the posterior distribution for the Birnbaum-Saunders with parameters α and β when prior choices $\pi_1(\alpha,\beta)$ and $\pi_2(\alpha,\beta)$ are used under type I right censoring. The analysis was conducted for sample sizes $n = 10$, $n = 40$, $n = 200$, for each case, the trace plots and the posterior densities, and the convergence diagnostics (Gelman-Rubin and Geweke) will be evaluated.

Figures 4.33, 4.34, and 4.35 present the trace plots and posterior density functions for α and β when the prior in equation 4.12 is used under type I right censoring for $n = 10$, $n = 40$, and $n = 200$, respectively. Figures 4.36, 4.37, and 4.38 present the trace plots and posterior density functions for α and β when the prior in equation 4.13 is used under type I right censoring for $n = 10$, $n = 40$, and $n = 200$, respectively. When both priors are used and under type I right censoring, the trace plots for α and β show that the chains mix well and pass the thick pen test, indicating that convergence has been reached. It is observed that as the sample size increases, there is no consistent trend in the posteriors, although the distributions become smoother and narrower, providing more reliable estimates.

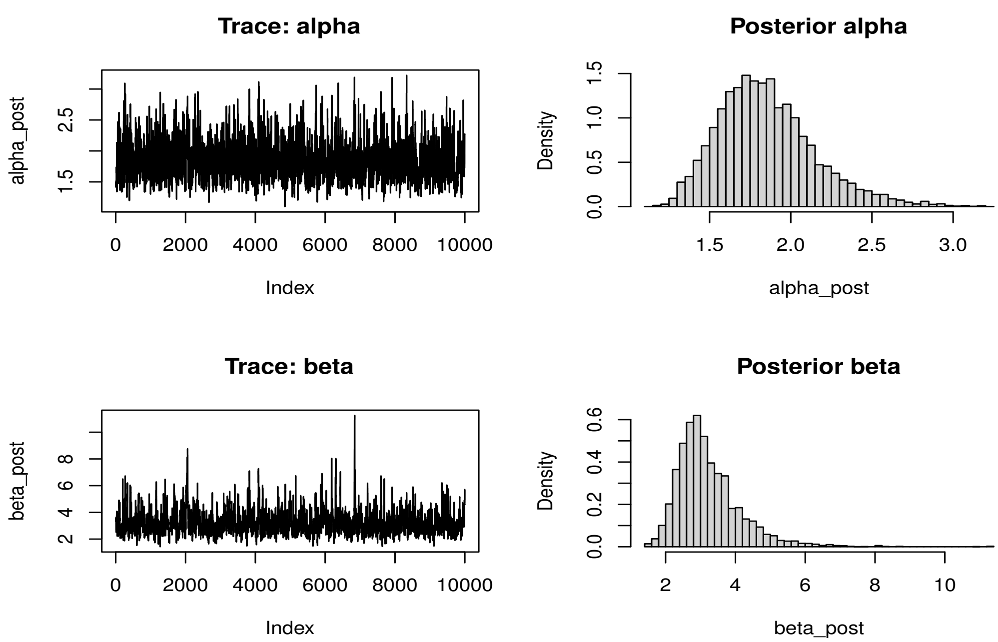

Figure 4.33: Trace plot and probability density function of α and β when $n = 10$.

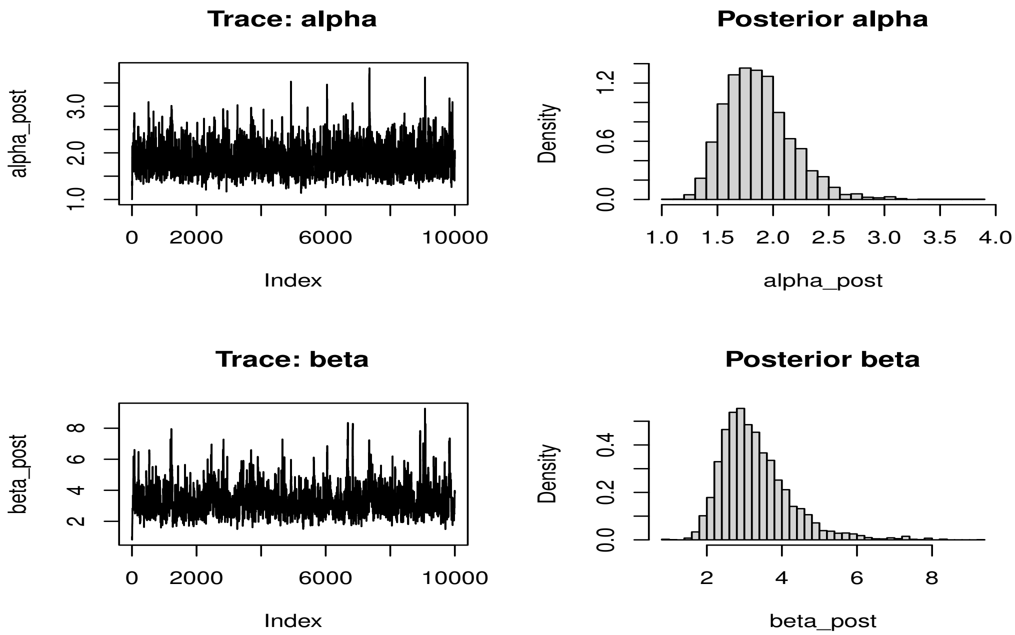

Figure 4.34: Trace plot and probability density function of α and β when $n = 40$.

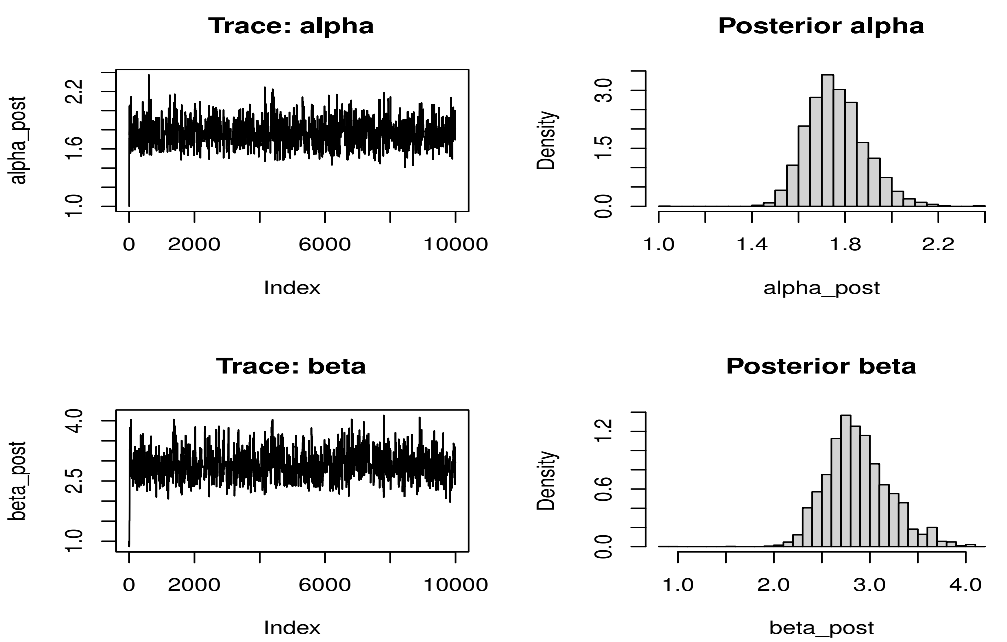

Figure 4.35: Trace plot and probability density function of α and β when $n = 200$.

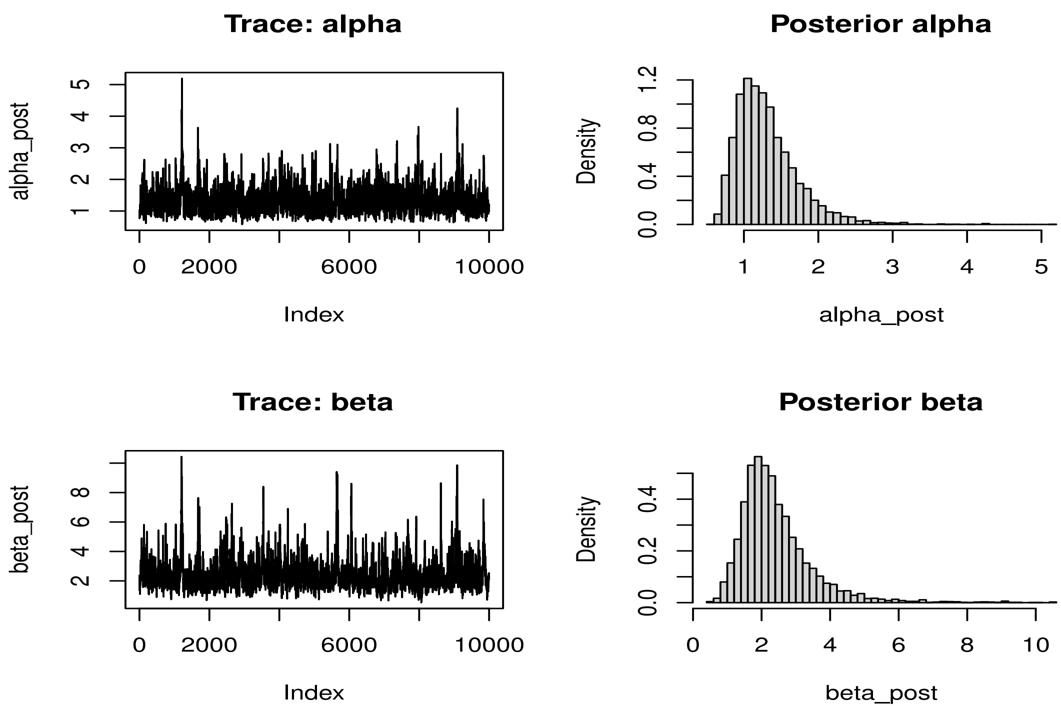

Figure 4.36: Trace plot and probability density function of α and β when $n = 10$.

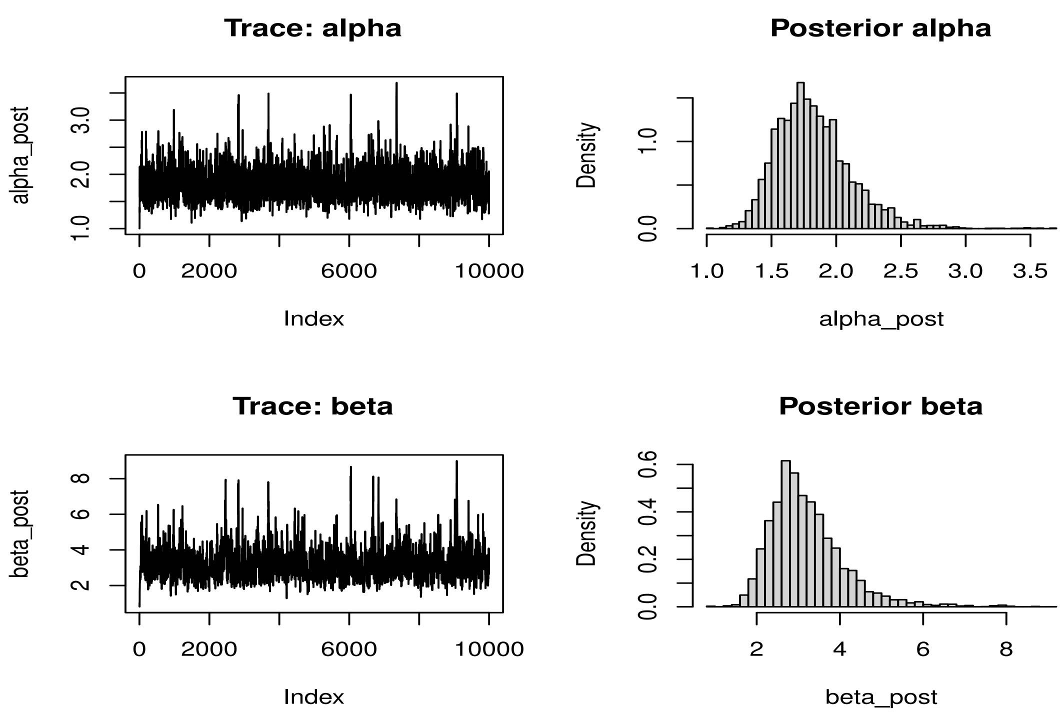

Figure 4.37: Trace plot and probability density function of α and β when $n = 40$.

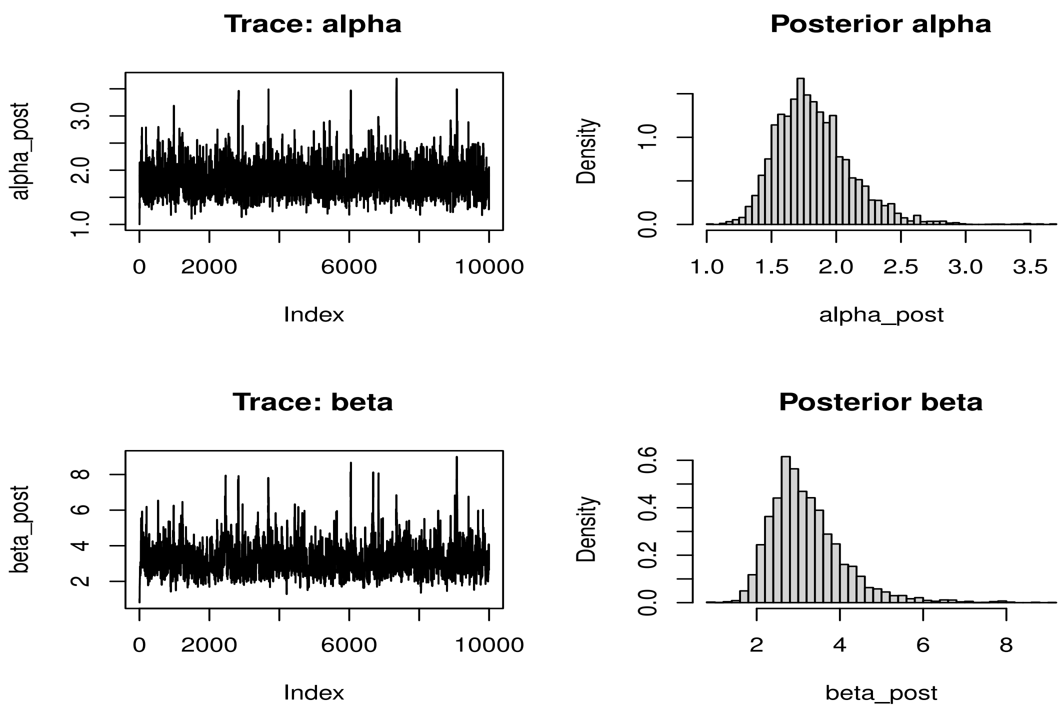

Figure 4.38: Trace plot and probability density function of α and β when $n = 200$.

Figures 4.39, 4.40, and 4.41 present the Gelman-Rubin plots for α and β when the prior in equation 4.12 is used for $n = 10$, $n = 40$, and $n = 200$, respectively.

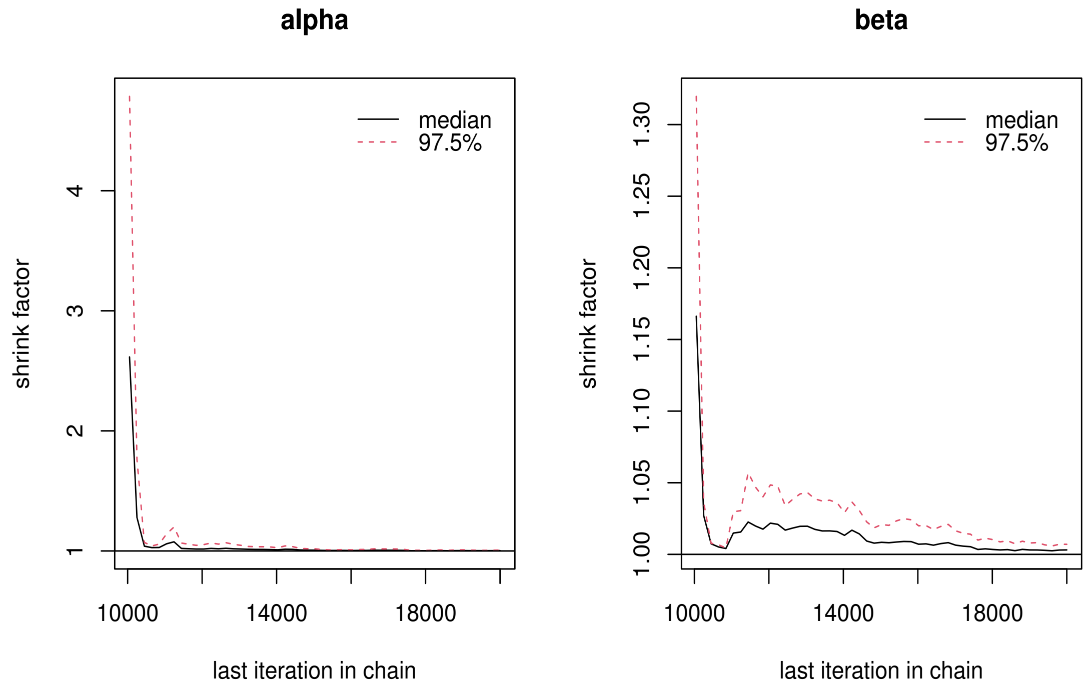

Figure 4.39: The Gelman-Rubin plots for α and β when $n = 10$.

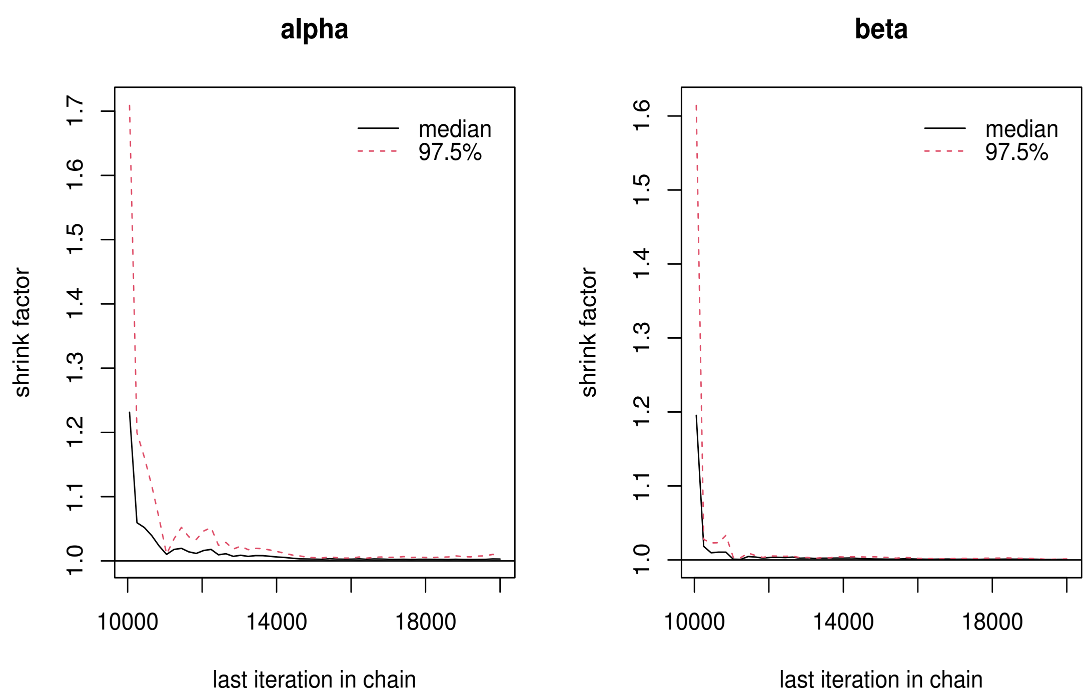

Figure 4.40: The Gelman-Rubin plots for α and β when $n = 40$.

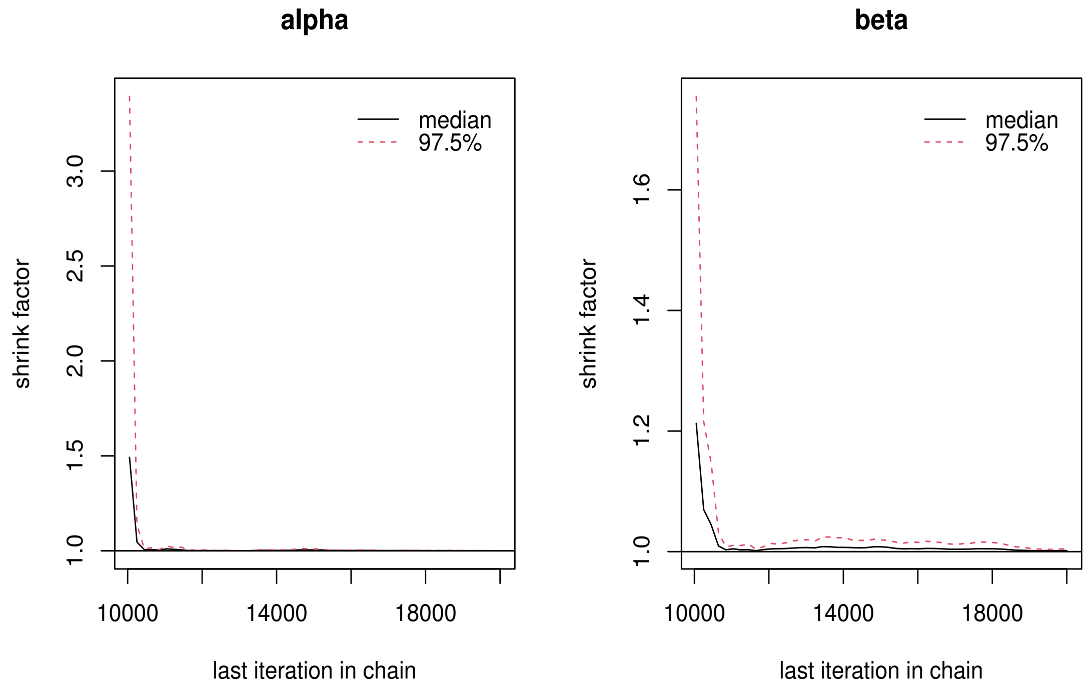

Figure 4.41: The Gelman-Rubin plots for α and β when $n = 200$.

Figures 4.42, 4.43, and 4.44 present the Gelman-Rubin plots for α and β when the prior in equation 4.13 is used for $n = 10$, $n = 40$, and $n = 200$, respectively.

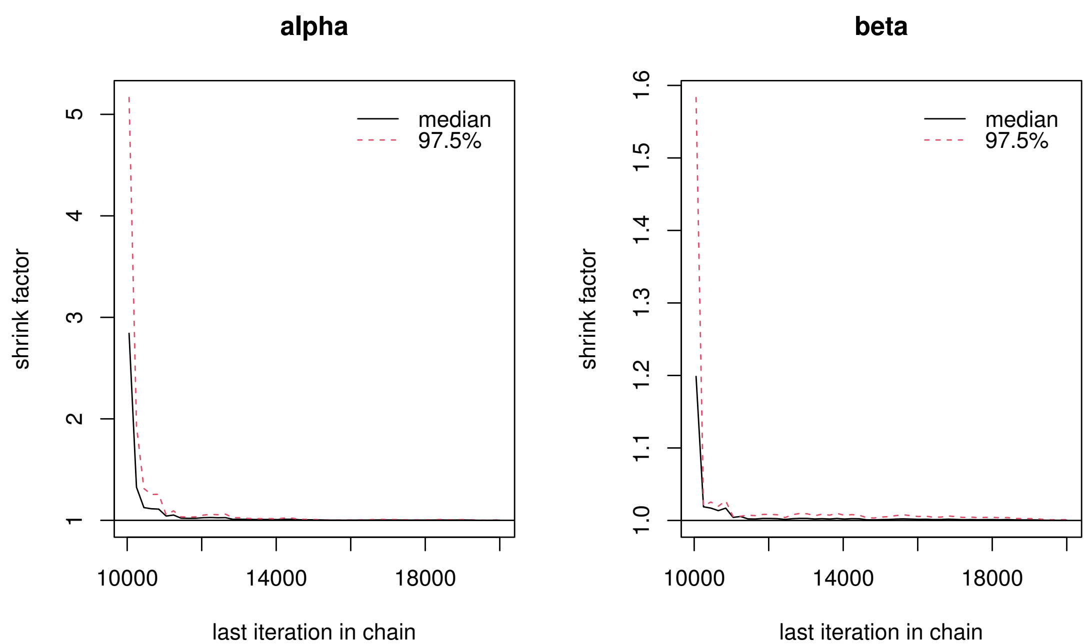

Figure 4.42: The Gelman-Rubin plots for α and β when $n = 10$.

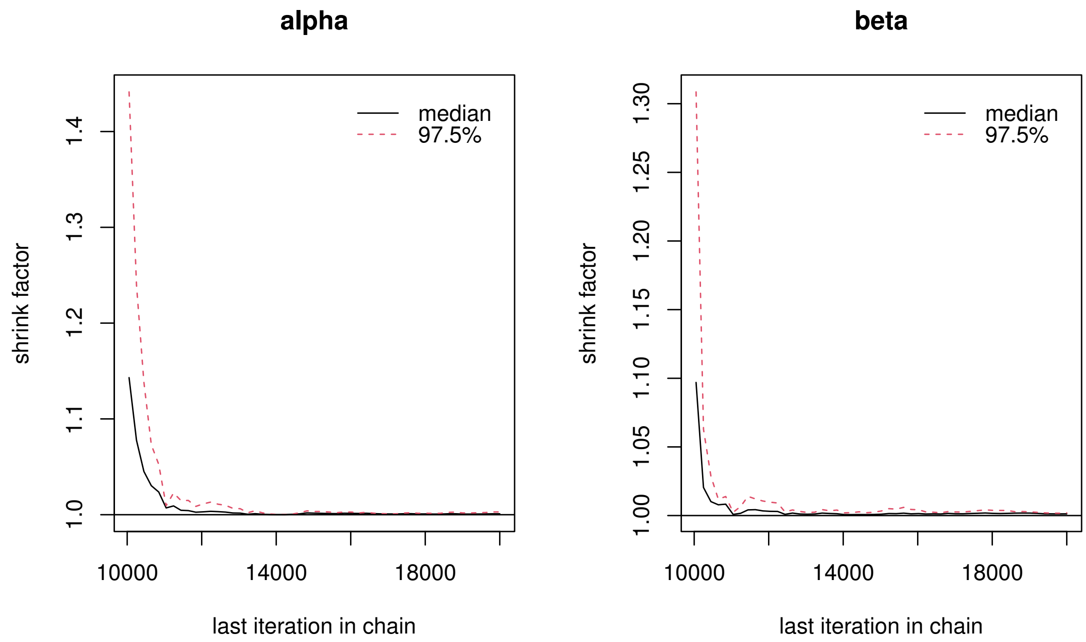

Figure 4.43: The Gelman-Rubin plots for α and β when $n = 40$.

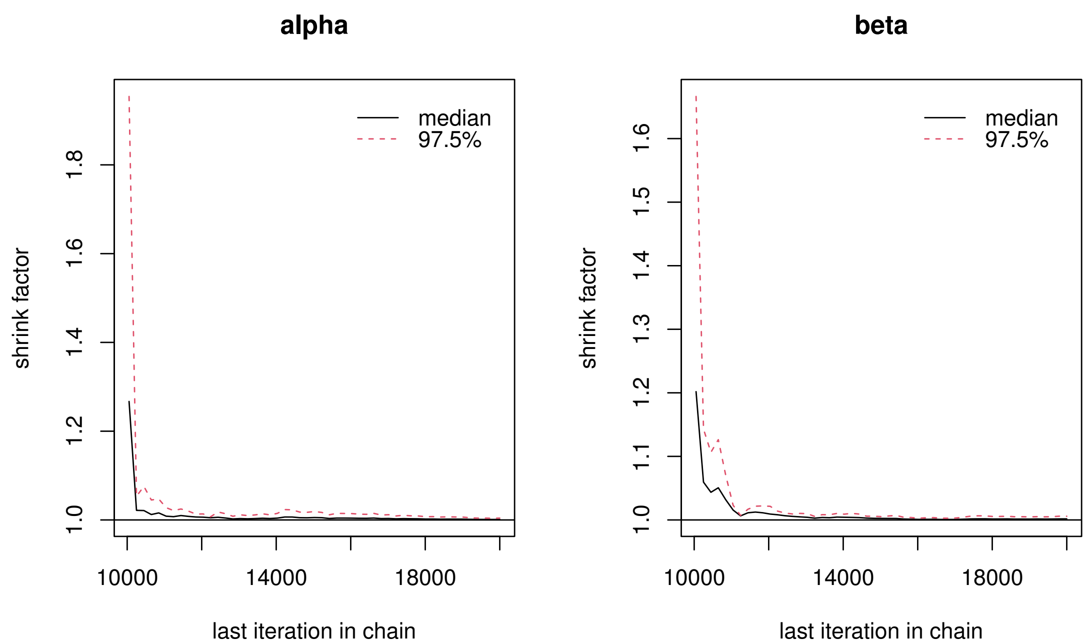

Figure 4.44: The Gelman-Rubin plots for α and β when $n = 200$.

For all sample sizes (n = 10, 40, 200) and when both priors 4.12 and 4.13 are used under type I right censoring, $\hat{R}$ values were very close to 1, indicating that the MCMC chains have converged.

Figures 4.45, 4.46, and 4.47 present the Geweke plots for α and β when the prior in equation 4.12 is used for $n = 10$, $n = 40$, and $n = 200$, respectively.

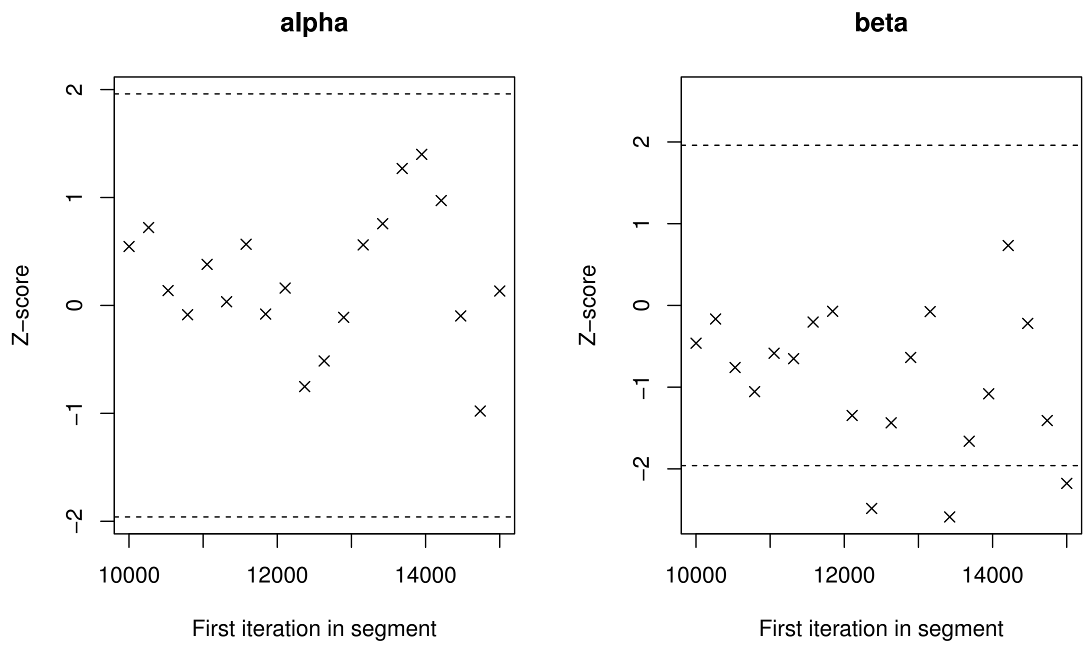

Figure 4.45: The Geweke plots for α and β when $n = 10$.

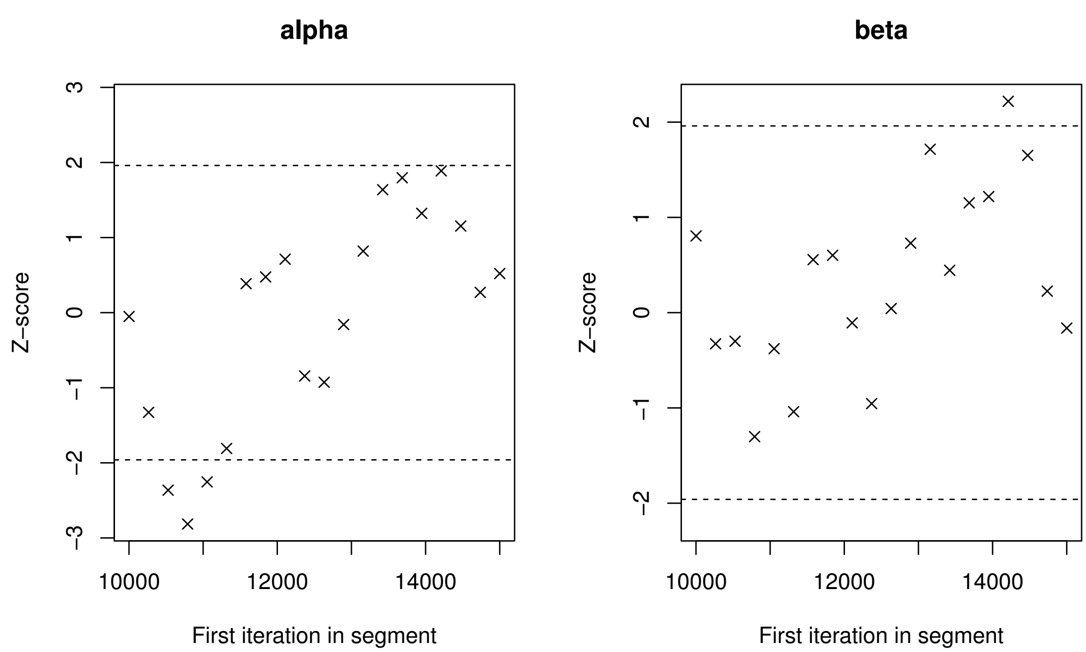

Figure 4.46: The Geweke plots for α and β when $n = 40$.

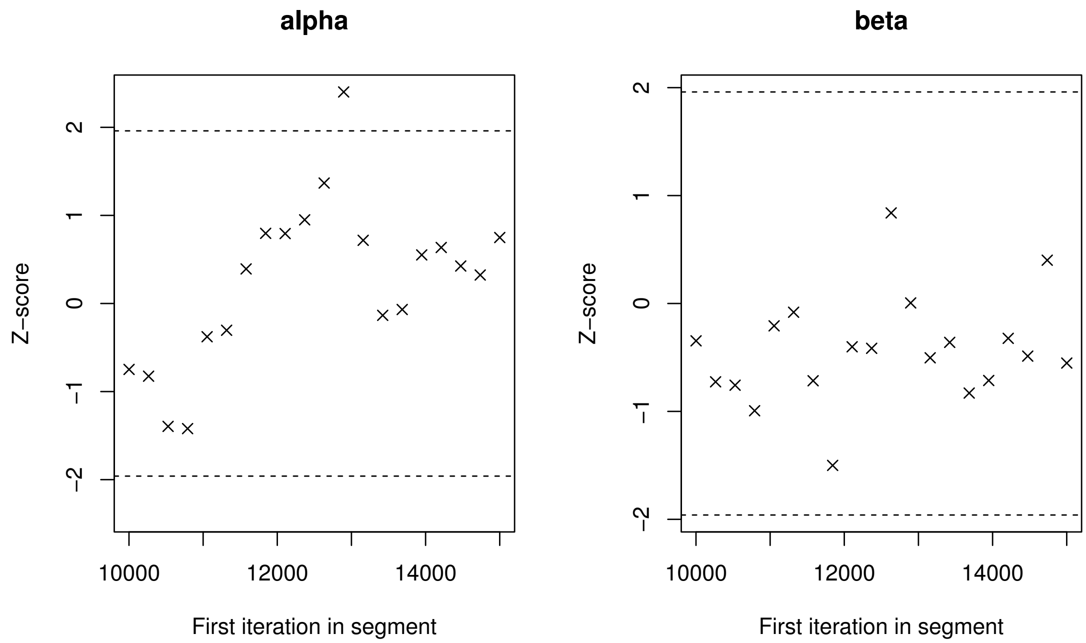

Figure 4.47: The Geweke plots for α and β when $n = 200$.

Figures 4.48, 4.49, and 4.50 present the Geweke plots for α and β when the prior in equation 4.13 is used for $n = 10$, $n = 40$, and $n = 200$, respectively.

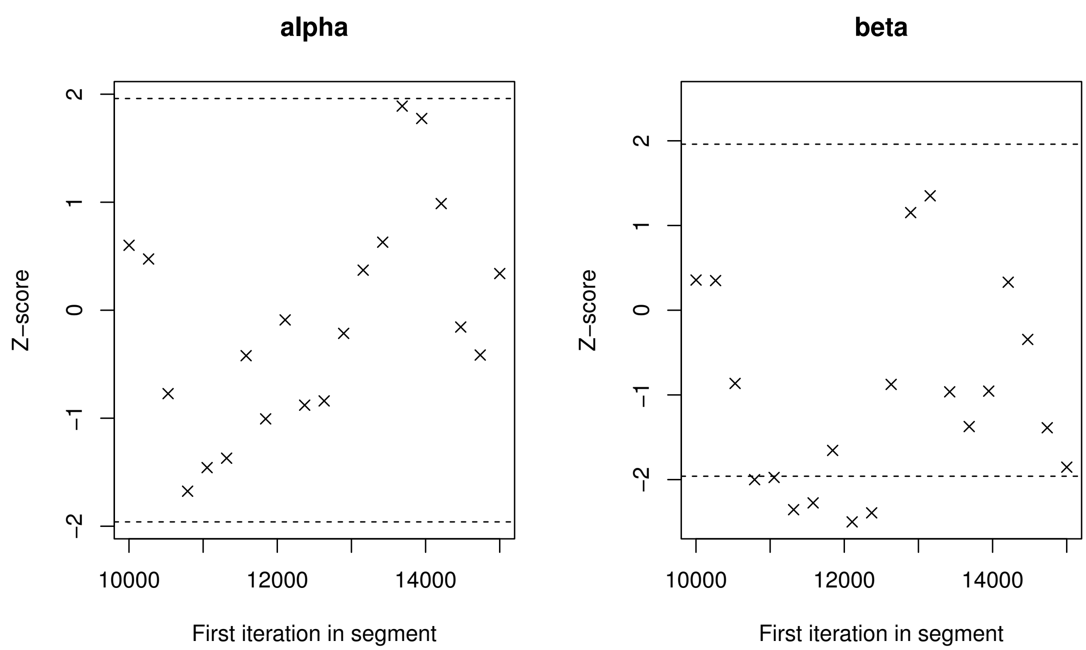

Figure 4.48: The Geweke plots for α and β when $n = 10$.

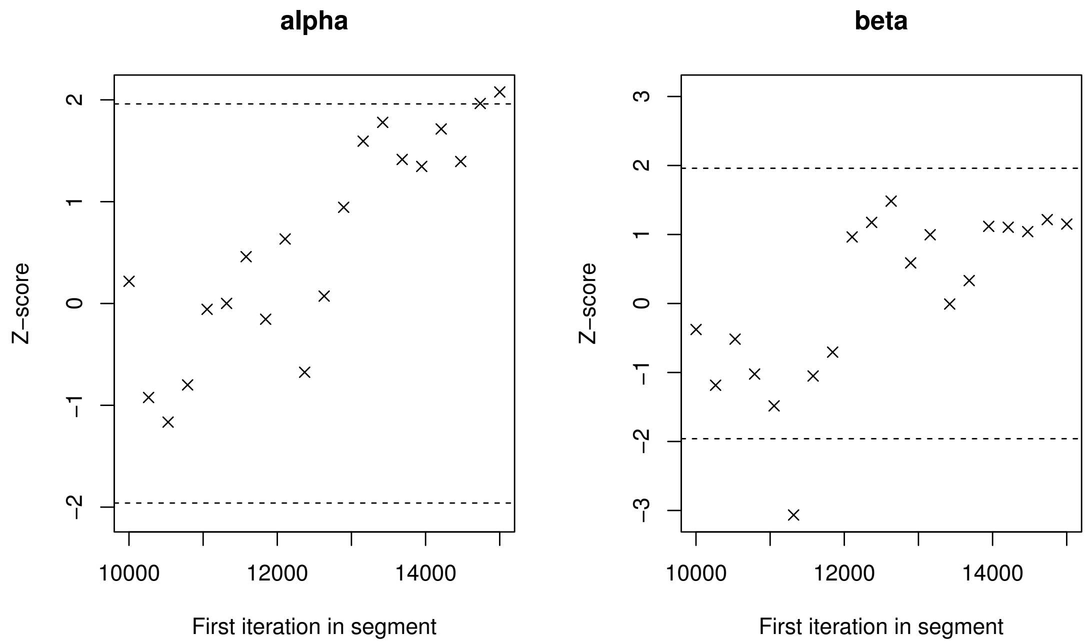

Figure 4.49: The Geweke plots for α and β when $n = 40$.

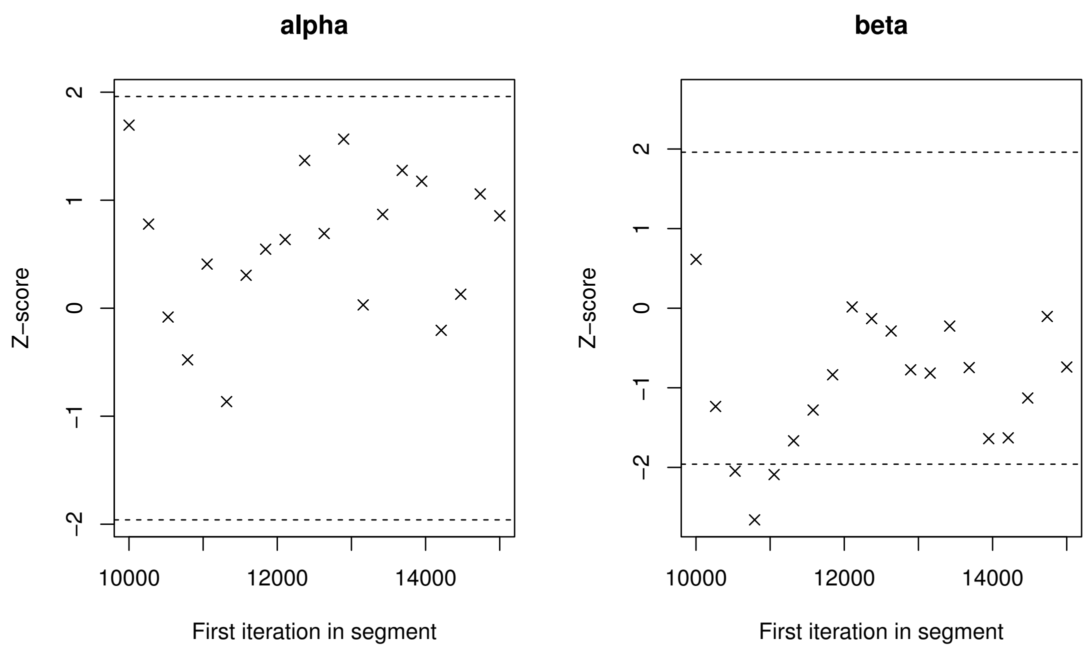

Figure 4.50: The Geweke plots for α and β when $n = 200$.

Geweke diagnostics were also computed for α and β to assess convergence within each MCMC chain. For all sample sizes ($n = 10, 40, 200$) under type I right censoring, the Z-scores mostly lie within the ±2 range, indicating that the chains have reached convergence for both when priors 4.12 and 4.13 are used.

4.8 Simulation Study for the Birnbaum-Saunders Distribution

This section presents the simulation study conducted to evaluate the performance of Bayesian inference for the Birnbaum-Saunders distribution under both complete data and type I right censoring scenarios. For each simulation, data were generated by sampling t from the Birnbaum-Saunders distribution with specified true parameters, shape parameters, $\alpha \in \{0.1, 0.25, 0.5\}$, and scale parameter, $\beta \in \{0.2, 1, 0.75\}$. The focus of the simulation study is on evaluating the performance of different priors in terms of coverage rates and mean interval lengths for the Birnbaum-Saunders parameters, under different sample sizes and censoring conditions.

4.8.1 Complete Data

In this subsection, the performance of the prior in equation 4.12 and the prior in equation 4.13 will be compared for different sample size, for the chosen sample sizes of $n = 10, 40$, and 200. Their performance will be evaluated using coverage rate and interval, a model with the shortest interval length and coverage very close to 95% also known as nominal level is considered to be the best. For each simulated true parameter values (α, β), the coverage probabilities were recorded as C_α for α and C_β for β. The mean interval lengths were recorded as L_α and L_β, respectively.

The following tables represent the coverage rate and mean lengths for α and β when the prior in equation 4.12 is used.

Table 4.1: Coverage rate and interval length for α and β when $n = 10$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9528	0.9407	0.1022	0.0269
0.25	0.20	0.9524	0.9412	0.2572	0.0680
0.50	0.20	0.9517	0.9401	1.1426	0.1351
0.10	1.00	0.9493	0.9393	0.1025	0.1349
0.25	1.00	0.9499	0.9409	0.2575	0.3407
0.50	1.00	0.9498	0.9365	2.8729	0.6760
0.10	0.75	0.9481	0.9342	0.1021	0.1008
0.25	0.75	0.9542	0.9441	0.2679	0.2547
0.50	0.75	0.9490	0.9336	2.2796	0.5034
Average		0.9508	0.9389	0.8205	0.2489

Table 4.2: Coverage rate and interval length for α and β when $n = 40$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9503	0.9378	0.0453	0.0124
0.25	0.20	0.9449	0.9448	0.1129	0.0310
0.50	0.20	0.9531	0.9488	0.2264	0.0609
0.10	1.00	0.9531	0.9375	0.0452	0.0617
0.25	1.00	0.9571	0.9430	0.1132	0.1554
0.50	1.00	0.9500	0.9457	0.2264	0.3049
0.10	0.75	0.9478	0.9343	0.0452	0.0462
0.25	0.75	0.9491	0.9463	0.1132	0.1166
0.50	0.75	0.9464	0.9433	0.2267	0.2288
Average		0.9502	0.9424	0.1283	0.1131

Table 4.3: Coverage rate and interval length for α and β when $n = 200$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9479	0.9499	0.0196	0.0051
0.25	0.20	0.9480	0.9433	0.0492	0.0132
0.50	0.20	0.9484	0.9488	0.0983	0.0262
0.10	1.00	0.9522	0.9442	0.0197	0.0256
0.25	1.00	0.9530	0.9385	0.0491	0.0658
0.50	1.00	0.9512	0.9438	0.0983	0.1308
0.10	0.75	0.9487	0.9493	0.0197	0.0191
0.25	0.75	0.9508	0.9454	0.0491	0.0494
0.50	0.75	0.9473	0.9489	0.0983	0.0981
Average		0.9497	0.9458	0.0557	0.0481

The following tables represent the coverage rate and mean lengths for α and β when the prior in equation 4.13 is used.

Table 4.4: Coverage rate and interval length for α and β when $n = 10$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9460	0.9348	0.1022	0.0269
0.25	0.20	0.9550	0.9428	2.5769	0.0681
0.50	0.20	0.9456	0.9334	1.6418	0.1346
0.10	1.00	0.9484	0.9399	1.0221	0.1346
0.25	1.00	0.9491	0.9405	2.5679	0.3400
0.50	1.00	0.9441	0.9339	2.3454	0.6687
0.10	0.75	0.9477	0.9321	1.0196	0.1007
0.25	0.75	0.9549	0.9425	3.0428	0.2548
0.50	0.75	0.9458	0.9397	3.2599	0.5020
Average		0.9485	0.9377	1.9532	0.2478

Table 4.5: Coverage rate and interval length for α and β when $n = 40$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9512	0.9301	0.0453	0.0121
0.25	0.20	0.9434	0.9392	0.1129	0.0306
0.50	0.20	0.9533	0.9461	0.2264	0.0605
0.10	1.00	0.9528	0.9488	0.0453	0.0605
0.25	1.00	0.9467	0.9407	0.1132	0.1535
0.50	1.00	0.9517	0.9425	0.2266	0.3028
0.10	0.75	0.9461	0.9443	0.0452	0.0453
0.25	0.75	0.9515	0.9440	0.1133	0.1153
0.50	0.75	0.9463	0.9406	0.2267	0.2269
Average		0.9492	0.9478	0.1283	0.1119

Table 4.6: Coverage rate and interval length for α and β when $n = 200$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9476	0.9490	0.0197	0.0132
0.25	0.20	0.9471	0.9463	0.0492	0.0132
0.50	0.20	0.9487	0.9472	0.0983	0.0256
0.10	1.00	0.9521	0.9375	0.0197	0.0256
0.25	1.00	0.9535	0.9390	0.0491	0.0658
0.50	1.00	0.9503	0.9477	0.0983	0.1308
0.10	0.75	0.9481	0.9452	0.0197	0.0191
0.25	0.75	0.9499	0.9494	0.0491	0.0494
0.50	0.75	0.9476	0.9490	0.0983	0.0981
Average		0.9495	0.9466	0.0557	0.0489

The coverage rates and mean interval lengths were found using the prior in equation 4.12 and the prior in equation 4.13, under different sample sizes, $n = 10$, $n = 40$, and $n = 200$. The coverage rates obtained from the simulation study are close to the nominal level of 95%, even for a small sample size, indicating that the Bayesian credible intervals are reliable. The mean interval lengths increase with a combination of larger parameter values, indicating more uncertainty as scale or shape parameters increase. The average interval lengths are getting smaller as the sample size increases. When $n = 10$ and the prior in equation 4.12 is used, the average coverage rate is 0.9508 for α and 0.9389 for β, closer to the nominal level for α compared to β. The average of the mean interval lengths is shorter for β at 0.2489 compared to α at 0.2489. On average, when $n = 40$ and prior in 4.12 is used, the coverage rate is slightly higher for $\alpha\,(0.9502)$ compared to $\beta\,(0.9424)$, indicating that the credible intervals for α tend to align more closely with the nominal level. The average mean interval lengths for $\alpha\,(0.1283)$ is wider than that for $\beta\,(0.1131)$. When $n = 200$ and the prior in 4.12 is used, the coverage probabilities remain close to the nominal level of 95% for both parameters, with averages of 0.9497 for α and 0.9458 for β. The mean interval lengths are shorter compared to smaller sample sizes, averaging 0.0557 for α and 0.0481 for β, which highlights the gain in precision as the sample size increases. Overall, the results confirm that the Bayesian credible intervals are both accurate and efficient in large samples. When $n = 10$ and the prior in equation 4.13 is used, the coverage rates are close to the nominal level, with averages of 0.9485 for α and 0.9377 for β, slightly lower than those when $n = 10$ and the prior in equation 4.12 is used. However, the mean interval lengths are slightly wider for α compared to those of β, averaging at 1.9532 for α and 0.2478 for β. The mean interval lengths when the prior in equation 4.13 is used are wider than the mean interval lengths when the prior in equation 4.12 is used. When $n = 40$, the average coverage rates are 0.9492 for α and 0.9478 for β, which are very close to the nominal 95% level, indicating reliable Bayesian credible intervals. The mean interval lengths are smaller than for $n = 10$, averaging 0.1283 for α and 0.1119 for β, reflecting increased precision as the sample size grows. Overall, these results demonstrate that both coverage accuracy and interval precision improve with larger samples, with β intervals slightly shorter than α. When $n = 200$, the average coverage rates are 0.9495 for α and 0.9466 for β, which are very close to the nominal level of 95%, indicating that the Bayesian credible intervals are highly reliable. The mean interval lengths are shorter than for $n = 10$ and $n = 40$, averaging 0.0557 for α and 0.0489 for β, reflecting increased precision with larger sample sizes. Overall, these results show that both coverage accuracy and interval precision improve as the sample size grows, with β intervals slightly shorter than those for α. The prior in equation 4.12 performs slightly better because it produces more consistent estimates with narrower intervals, even for small samples, while maintaining good coverage rates.

4.8.2 Type I Right Censoring

In this subsection, the performance of the prior in equation 4.12 and the prior in equation 4.13 will be compared for different sample size, for the chosen sample sizes of $n = 10, 40$, and 200 under type I right-censoring. Their performance will be evaluated using coverage rates and mean interval length. A model with the shortest interval length and coverage very close to 95% (also known as nominal level) is considered to be the best. For each simulated true parameter values (α, β), the coverage probabilities were recorded as C_α for α and C_β for β. The mean interval lengths were recorded as L_α and L_β, respectively. The following tables represent the coverage rates and mean lengths for α and β when the prior in equation 4.12 is used.

Table 4.7: Coverage rate and interval length for α and β when $n = 10$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9273	0.9341	0.1189	0.0294
0.25	0.20	0.9235	0.9347	0.2998	0.0735
0.50	0.20	0.9381	0.9322	0.6096	0.1416
0.10	1.00	0.9225	0.9328	0.1187	0.1467
0.25	1.00	0.9339	0.9338	0.2994	0.3673
0.50	1.00	0.9350	0.9317	0.6451	0.7384
0.10	0.75	0.9249	0.9332	0.1188	0.1104
0.25	0.75	0.9405	0.9406	0.3016	0.2781
0.50	0.75	0.9275	0.9276	0.6215	0.5507
Average		0.9304	0.9334	0.3482	0.2707

Table 4.8: Coverage rate and interval length for α and β when $n = 40$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9330	0.9449	0.0557	0.0134
0.25	0.20	0.9390	0.9394	0.1403	0.0335
0.50	0.20	0.9368	0.9377	0.2828	0.0664
0.10	1.00	0.9338	0.9421	0.0554	0.0668
0.25	1.00	0.9372	0.9415	0.1402	0.1672
0.50	1.00	0.9406	0.9443	0.2840	0.3334
0.10	0.75	0.9371	0.9418	0.0555	0.0502
0.25	0.75	0.9411	0.9423	0.1409	0.1260
0.50	0.75	0.9373	0.9409	0.2834	0.2493
Average		0.9373	0.9417	0.1598	0.1229

Table 4.9: Coverage rate and interval length for α and β when $n = 200$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9393	0.9421	0.0245	0.0058
0.25	0.20	0.9429	0.9479	0.0619	0.0146
0.50	0.20	0.9452	0.9446	0.1251	0.0290
0.10	1.00	0.9370	0.9448	0.0245	0.0293
0.25	1.00	0.9436	0.9461	0.0619	0.0732
0.50	1.00	0.9434	0.9460	0.1249	0.1450
0.10	0.75	0.9358	0.9405	0.0245	0.0219
0.25	0.75	0.9427	0.9421	0.0619	0.0548
0.50	0.75	0.9405	0.9463	0.1249	0.1087
Average		0.9412	0.9445	0.0705	0.0536

The following tables represent the coverage rate and mean lengths for α and β when the prior in equation 4.13 is used.

Table 4.10: Coverage rate and interval length for α and β when $n = 10$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9298	0.9424	1.2127	0.0301
0.25	0.20	0.9248	0.9366	3.0011	0.0735
0.50	0.20	0.9265	0.9314	4.7476	0.1462
0.10	1.00	0.9308	0.9402	1.2109	0.1499
0.25	1.00	0.9274	0.9364	3.0026	0.3687
0.50	1.00	0.9260	0.9301	3.1353	0.7375
0.10	0.75	0.9320	0.9418	1.2119	0.1127
0.25	0.75	0.9321	0.9404	3.0249	0.2791
0.50	0.75	0.9386	0.9281	2.1755	0.5503
Average		0.9287	0.9364	2.5274	0.2720

Table 4.11: Coverage rate and interval length for α and β when $n = 40$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9342	0.9466	0.0559	0.0135
0.25	0.20	0.9381	0.9402	0.1403	0.0335
0.50	0.20	0.9386	0.9386	0.2829	0.0665
0.10	1.00	0.9361	0.9432	0.0556	0.0671
0.25	1.00	0.9379	0.9407	0.1403	0.1674
0.50	1.00	0.9414	0.9453	0.2841	0.3336
0.10	0.75	0.9414	0.9453	0.0557	0.0504
0.25	0.75	0.9404	0.9438	0.1409	0.1260
0.50	0.75	0.9360	0.9408	0.2833	0.2493
Average		0.9382	0.9427	0.1599	0.1230

Table 4.12: Coverage rate and interval length for α and β when $n = 200$.

α	β	C_α	C_β	L_α	L_β
0.10	0.20	0.9394	0.9428	0.0245	0.0058
0.25	0.20	0.9439	0.9466	0.0619	0.0146
0.50	0.20	0.9451	0.9464	0.1252	0.0291
0.10	1.00	0.9388	0.9450	0.0245	0.0293
0.25	1.00	0.9419	0.9461	0.0619	0.0732
0.50	1.00	0.9438	0.9463	0.1249	0.1450
0.10	0.75	0.9370	0.9408	0.0245	0.0219
0.25	0.75	0.9437	0.9421	0.0618	0.0548
0.50	0.75	0.9413	0.9477	0.1249	0.1088
Average		0.9417	0.9459	0.0705	0.0536

The coverage rates and mean interval lengths under type I censoring were evaluated using the prior in equation 4.12and the prior in equation 4.13 for different sample sizes. The results indicate that the coverage rates are close to the nominal 95% level even for small samples. The mean interval lengths increase with larger parameter values, reflecting greater uncertainty as the scale or shape parameters increase. Additionally, average interval lengths decrease as the sample size increases.

When $n = 10$ and the prior in equation 4.12 is used, the average coverage rates were 0.9304 for α and 0.9334 for β, slightly below the nominal 95% level; in other words there is undercoverage. The corresponding average interval lengths were relatively wide, 0.3483 for α and 0.2707 for β, reflecting the higher uncertainty associated with small datasets. As the sample size increased to $n = 40$, coverage rates improved to 0.9373 for α and 0.9417 for β, with interval lengths decreasing to 0.1598 and 0.1229, respectively. For the largest sample size $n = 200$, the coverage rates approached the nominal level more closely, with 0.9412 for α and 0.9445 for β, while the intervals became substantially narrower (0.0705 for α and 0.0536 for β).

When $n = 10$ and the prior in equation 4.13 is used, the average coverage rates were 0.9287 for α and 0.9364 for β, there was undercoverage. The corresponding average interval lengths were 2.5274 for α and 0.2720 for β, the interval lengths for α are consistently larger than those for β. As the sample size increased to $n = 40$, the average coverage rates improved to 0.938 for α and 0.943 for β, while the average interval lengths decreased to 0.1599 for α and 0.1230 for β. For the largest sample size $n = 200$, the coverage rates approached the nominal level more closely, with 0.942 for α and 0.946 for β, and mean interval lengths got shorter to an average of 0.0705 for α and 0.0536 for β. This demonstrates high precision and strong posterior certainty with large sample sizes.

The prior in equation 4.12 is recommended, as it provides more stable coverage and narrower intervals compared to those when prior in equation 4.13 is used.

4.9 Predictive Reliability for the Birnbaum-Saunders Distribution

This section presents the predictive reliability analysis based on the Birnbaum-Saunders distribution under both complete and type I right censoring. The objective is to evaluate how well the Birnbaum-Saunders distribution predicts future lifetimes using the posterior estimates of the model parameters. Predictive reliability values are computed using the prior in equation 4.12 and the prior in equation 4.13 for different sample sizes. The results are summarised and compared in tables and plots to assess the impact of prior choice and sample size on predictive performance. The predictive reliability for the Birnbaum-Saunders distribution is given by

$$R(t_u|\underline{t}) = \int_0^\infty \int_0^\infty \pi(\alpha, \beta|\underline{t}) R(t_u \mid \alpha, \beta)\, d\alpha d\beta, \tag{4.14}$$

where $R(t_u \mid \alpha, \beta)$ is the Birnbaum-Saunders distribution reliability function at time t_u.

This expression is analytically challenging. Given the posterior draws $\left(\alpha^{(n)}, \beta^{(n)}\right)$, $n = 1, 2, ..., N$, from the MCMC simulation, the integral in equation 4.14 can be evaluated by Monte Carlo average

$$R\left(t_u|\underline{t}\right) \approx \frac{1}{N}\sum_{n=1}^{N}\left\{1 - \Phi\left[\frac{1}{\alpha^{(n)}}\left\{\left(\frac{t_u}{\beta^{(n)}}\right)^{\frac{1}{2}} - \left(\frac{\beta^{(n)}}{t_u}\right)^{\frac{1}{2}}\right\}\right]\right\}.$$

4.9.1 Complete Data

In this subsection, the predictive reliability for the Birnbaum-Saunders distribution will be evaluated within a Bayesian framework using the prior distribution specified above for the shape α, and scale β parameter.

The posterior predictive reliability curves and tables were generated for sample size $n = 10$ and $n = 40$.

Table 4.13: Predictive reliability table for various priors when $n = 10$.

Time	Prior1	Prior2
1	1	1
2	1	1
3	1	1
.		
.		
.		
401	0.0711	0.0616
402	0.0704	0.0611
403	0.0699	0.0605
.	.	.
.	.	.
599	0.0167	0.0125
600	0.0166	0.0124

Table 4.14: Predictive reliability table for various priors when $n = 40$.

Time	Prior1	Prior2
1	1	1
2	1	1
3	1	1
.		
.		
.		
401	0.0083	0.0074
402	0.0082	0.0073
403	0.0081	0.0072
.		
.		
599	0.0006	0.0005
600	0.0006	0.0005

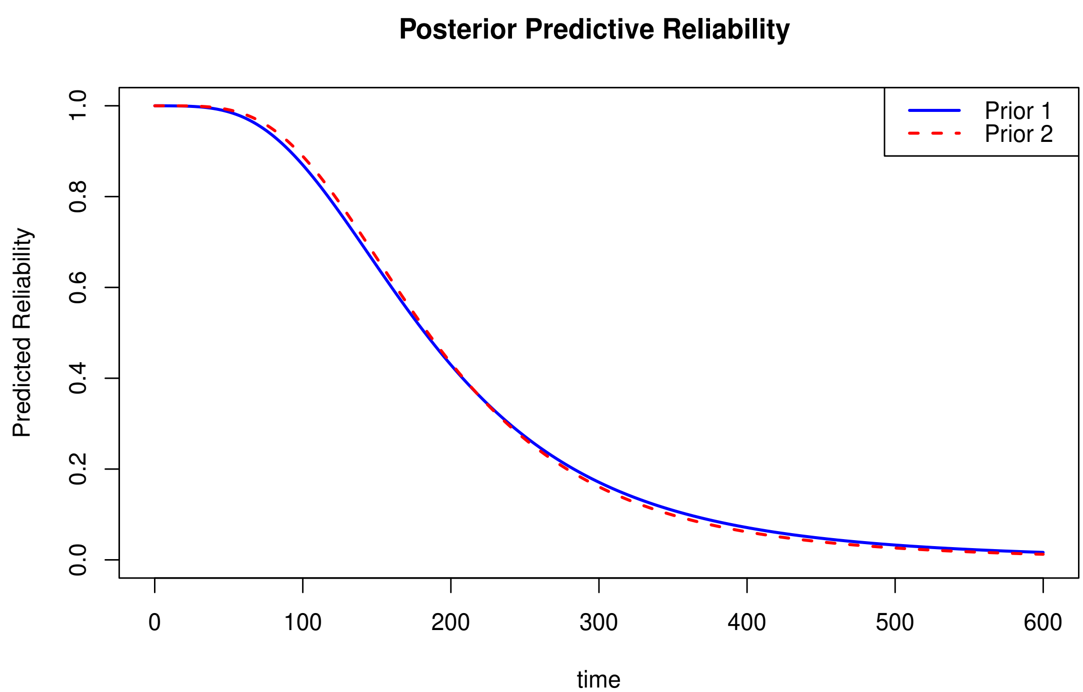

Figure 4.51: Predictive reliability when $n = 10$.

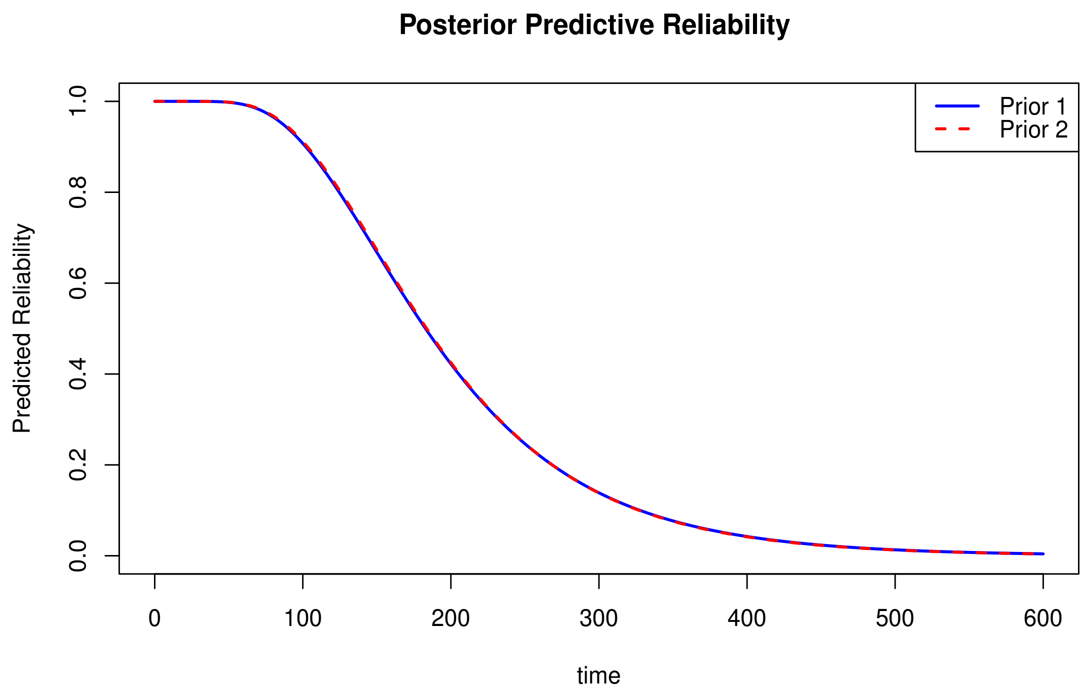

Figure 4.52: Predictive reliability when $n = 40$.

Figure 4.51 and Figure 4.52 illustrate predictive reliability curves for the Birnbaum-Saunders distribution using two priors specified where above when sample size are 10 and 40 respectively. The predictive reliability curve under the two priors from equation 4.12 and equation 4.13 exhibit very similar behavior across time, both starting at 1 and decreasing smoothly towards 0. For $n = 10$, in the early stage (approximately 50-200 seconds), the predictive reliability curve produced when the prior in equation 4.13 is used lies slightly above that of when the prior in equation 4.12 is used, suggesting that when the prior in equation 4.13 is used it yields higher short-term reliability. After about 201 seconds, the curves intersect, and predictive reliability produced when the prior in equation 4.12 is used remains higher as they both approach zero. For $n = 40$, the predictive reliability curve for both priors overlap throughout, showing that the influence of the prior diminishes as the sample size increases.

4.9.2 Type I Right Censoring

In this subsection, we will evaluate the predictive reliability for the Birnbaum-Saunders distribution under type I censoring, when the priors in equation 4.12 and equation 4.13 are used for the shape α, and scale β parameter. The posterior predictive reliability curves and tables were generated for sample size $n = 10$ and $n = 40$.

Table 4.15: Predictive Reliability for various priors when $n = 10$.

Time	Prior1	Prior2
1	1	1
2	1	1
3	1	1
.		
.		
.		
401	0.0326	0.0132
402	0.0323	0.0130
403	0.0321	0.0129
.	.	.
.	.	.
599	0.0086	0.0019
600	0.0086	0.0019

Table 4.16: Predictive Reliability for various priors when $n = 40$.

Time	Prior1	Prior2
1	1	1
2	1	1
3	1	1
.		
.		
.		
401	0.0507	0.0443
402	0.0502	0.0438
403	0.0496	0.0433
.	.	.
.	.	.
599	0.0072	0.0054
600	0.0071	0.0053

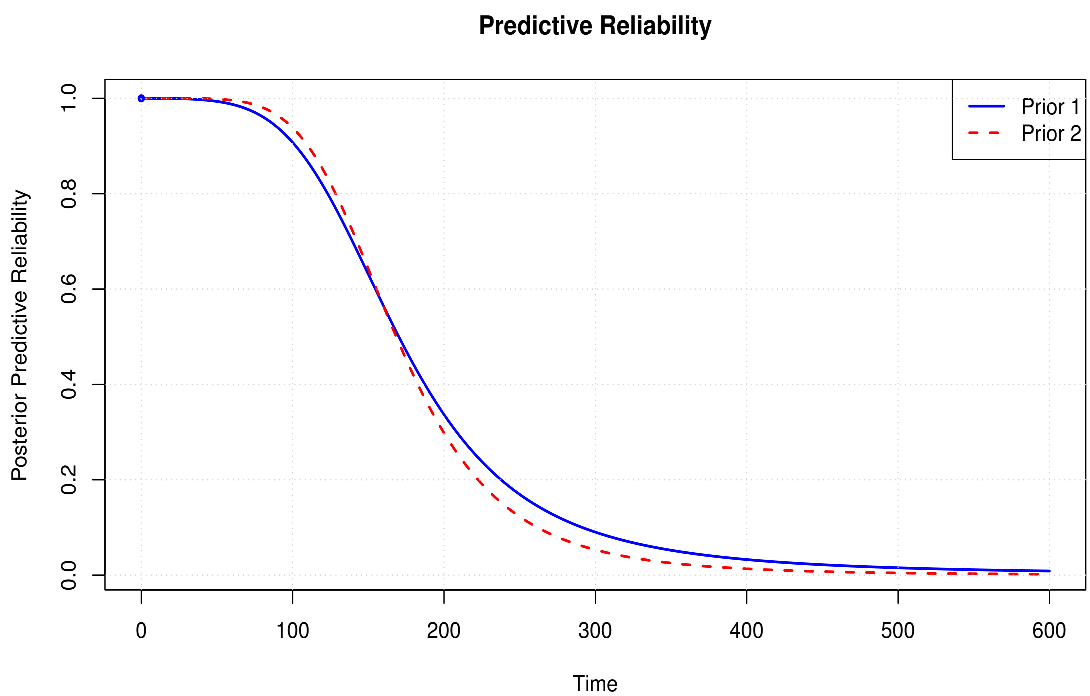

Figure 4.53: Predictive Reliability when $n = 10$.

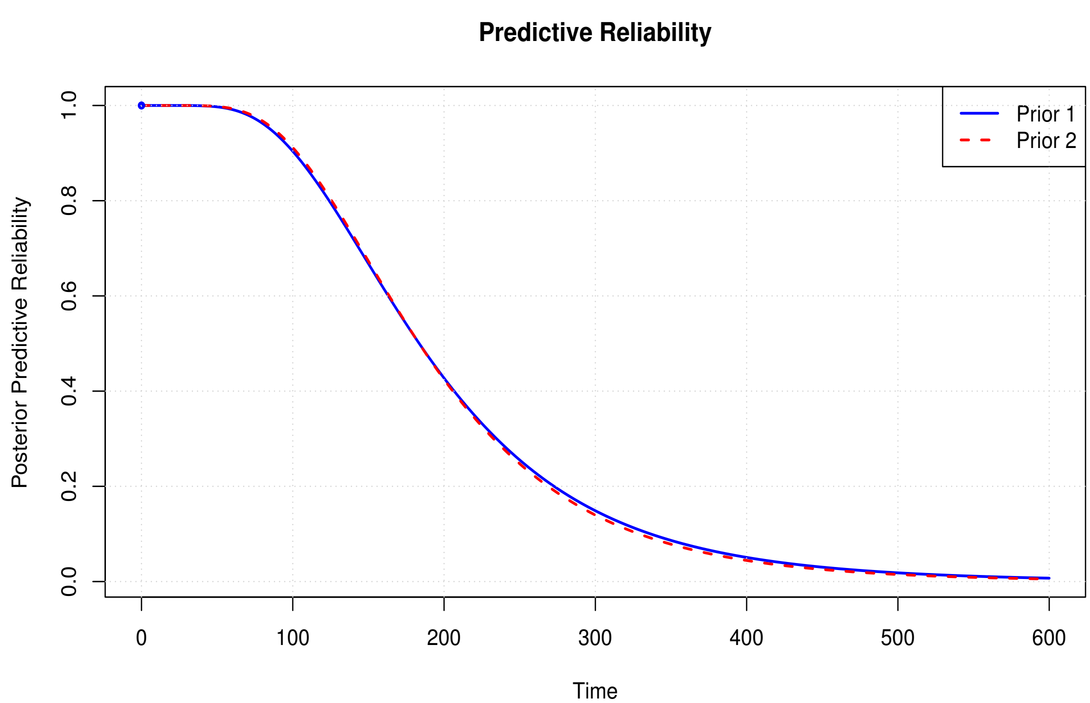

Figure 4.54: Predictive Reliability when $n = 40$.

The plots compare the predictive reliability under two priors for different sample sizes. For $n =$

10, the predictive reliability decreases over time, as expected. At the early stage (approximately 0-200 seconds), the reliability curve when the prior in equation 4.13 is used lies slightly above the reliability curve when the prior in equation 4.12 is used, indicating a higher short-term reliability, after intersecting the reliability curve when the prior in equation 4.12 is used remains higher as they both approach 0. When the sample size increases to $n = 40$, the reliability decreases smoothly as time goes. The two curves overlap, showing that as the sample size increases, the effect of the prior becomes negligible, and the posterior predictive reliability is nearly identical under both prior.

Chapter 5

Application

5.1 Application (Complete Data)

This section presents the application of both the Weibull and Birnbaum-Saunders distribution to complete lifetime datasets. The objective of this application is to estimate the shape β and scale α parameters for both distributions, together with their corresponding credibility intervals, and to evaluate the DIC for model assessment and comparison. In Bayesian model assessment, the DIC provides a measure of model fit that balances model adequacy with model complexity. DIC was proposed by Spiegelhalter et al. (2002) as a model assessment and comparison tool. The effective number of parameters in the model, p_D is defined as

$$p_D = \overline{D(\theta)} - D\left(\bar{\theta}\right),$$

where parameter vector $\theta = (\alpha, \beta)$.

The Bayesian deviance is given by

$$D(\theta) = -2\log\left[L(\theta \mid data)\right] + 2\log\left[f(data)\right], \tag{5.1}$$

where $f(data)$ is some function of the data and $L(\theta \mid data)$ is the maximsed likelihood value over the unknown parameters, Izally (2016). From equation 5.1, the posterior mean deviance is given by

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

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

$$D\left(\bar{\theta}\right) = -2\log\left[L(E(\theta \mid data))\right] + 2\log\left[f(data)\right].$$

As stated in Spiegelhalter et al. (2002), the definition of the DIC is

$$DIC = \overline{D(\theta)} + p_D. \tag{5.2}$$

The preferred model is the one with the lowest DIC value. Two examples are analyzed in this section: Example 5.1.1 uses a small dataset to illustrate the Bayesian estimation process, while Example 5.1.2 applies the same procedure to a larger dataset consisting of 101 observed lifetimes.

5.1.1 Example 1

This example uses a small fatigue lifetime dataset originally reported by Mccool (1975), consisting of observed failures times from fatigue experiment on mechanical components subjected to cyclic loading. The dataset is commonly used in reliability studies to demonstrate lifetime modelling under small sample sizes, where uncertainty is substantial and prior choice plays an important role. In this study, the data are analysed using both the Weibull and Birnbaum-Saunders models to illustrate the Bayesian estimation procedure, compare posterior behaviour, and assess the suitability of each model for fatigue related lifetime data. The example therefore serves as a practical illustration of the proposed Bayesian methodology rather than as a large scale empirical analysis.

Table 5.1: Fatigue lifetime data.

152.7	172.0	172.5	173.3	193.0	204.7	216.5	234.9	262.6	422.6

Figure 5.1 and 5.2 shows the quantile-quantile(Q-Q) plots for the Weibull and Birnbaum-Saunders distribution, respectively. The empirical quantiles of the observed data are plotted against the theoretical quantiles computed from the posterior mean and median estimates of the Weibull shape β and scale α parameters. The points lie approximately along the reference line, indicating that the Weibull distribution provides a reasonable description for this data. Figure 5.2 shows that the data is describe by the Birnbaum-Saunders distribution. This result is also support by the study of Wang et al. (2016).

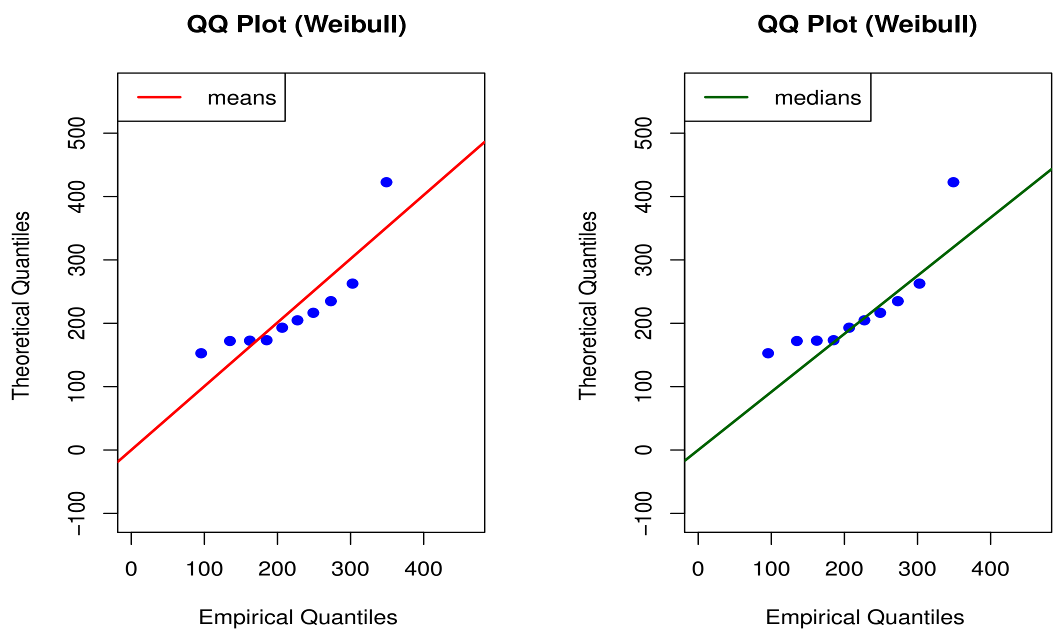

Figure 5.1: Q-Q plots for the fitted Weibull distribution with the means (left panel) and the median (right panel) for Example Mccool (1975).5.1.1.

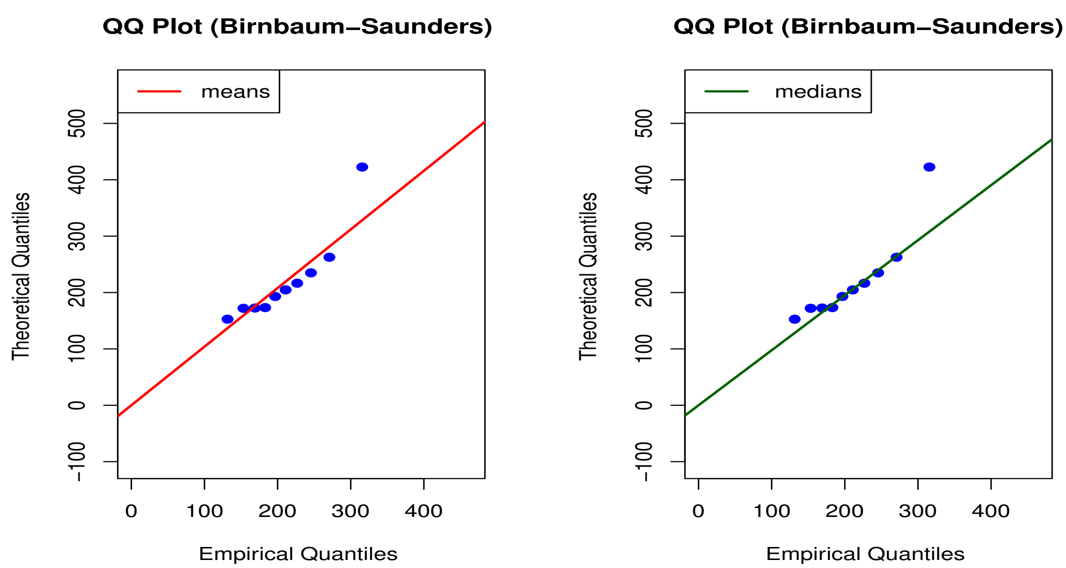

Figure 5.2: Q-Q plots for the fitted Birnbaum-Saunders distribution with the means (left panel) and the median (right panel) for Example 5.1.1.

The Bayesian Weibull and Birnbaum–Saunders models were implemented in OpenBUGS. For each model, three Markov chains were simulated, each with 300 000 iterations, of which the first 100 000 iterations were discarded as burn-in to remove the influence of initial values. Weakly-informative priors were assigned to the model parameters. Specifically, both α and β parameter were assumed to follow inverse gamma distributions with hyperparameters $a = b = 10^{-4}$.

Table 5.2 presents the posterior summaries for both distribution, including posterior means, medians, standard deviations, and credibility intervals. The Birnbaum–Saunders model shows slightly lower scale estimates and precise credible intervals for both parameters. Furthermore, the deviance information for the Weibull and Birnbaum-Saunders models is reported in Table 5.3. The DIC value for the Birnbaum-Saunders model is lower than that of the Weibull distribution, indicating better overall fit to the data.

Table 5.2: Posterior Summaries for the Weibull and Birnbaum-Saunders models.

Model	Parameter	Mean	Median	sd	MC_error	95% CI
Weibull	α	2.725	2.695	0.6332	0.0026	$(1.5641, 3.8859)$
	β	250.7	248.8	33.95	0.1175	$(188.4697, 312.9304)$
Birnbaum-Saunders	α	0.3268	0.3102	0.0882	0.0650	$(0.1651, 0.4885)$
	β	213.3	212.1	22.89	0.0498	$(171.3426, 255.2574)$

Table 5.3: Deviance information for Example 5.1.1.

Model	$\overline{D}$	$\hat{D}$	DIC	p_D
Weibull	116.8	115	118.7	1.841
Birnbaum-Saunders	98.29	96.47	100.1	1.825

Figure 5.3 and 5.4 present the posterior distributions of the parameters for the Weibull and Birnbaum-Saunders models fitted to the fatigue lifetime data. All posterior curves exhibit noticeable right skewness, reflecting the limited information content of the small dataset. Under such conditions, posterior uncertainty is expected, and relatively large lifetime values remain plausible.

For the Weibull model, the posterior means and medians of both parameters are close, with α having a mean of 2.725 and a median of 2.695, and β having a mean of 250.7 and a median of 248.8. The small differences between these summaries indicate mild right skewness. The posterior distribution of the scale parameter α is concentrated around 2.7, suggesting a characteristic lifetime of this magnitude, while the concentration of the shape parameter β around 250 implies a strongly increasing hazard function, consistent with wear-out or fatigue-driven failure behaviour. Similarly, the Birnbaum–Saunders posterior distributions are right-skewed, with α having a posterior mean of 0.3268 and median of 0.3102, and β having a mean of 213.3 and median of 212.1. The posterior concentration of α around 0.3 reflects moderate variability in the fatigue damage accumulation process, while the concentration of β near 213 indicates a characteristic fatigue life consistent with the observed data. The right skewness in both parameters again highlights the uncertainty induced by the small sample size.

See Appendix C.4 and C.5 for the OpenBUGS code.

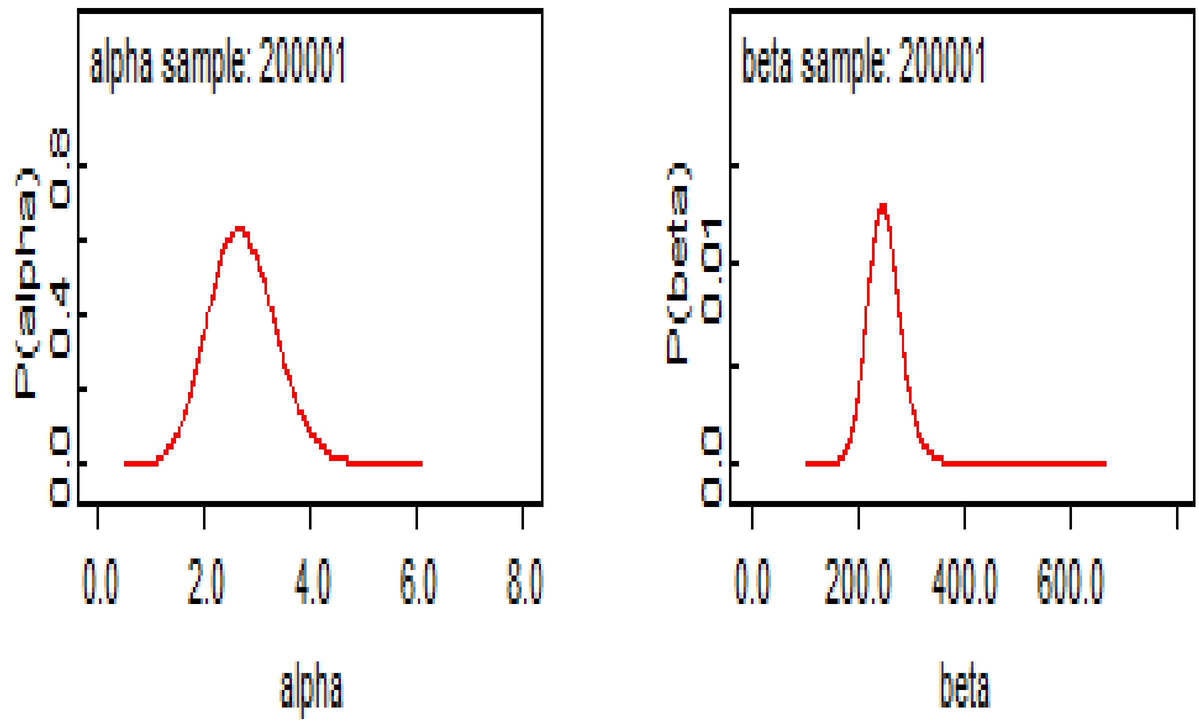

Figure 5.3: Posterior curve of α and β for the Weibull distribution.

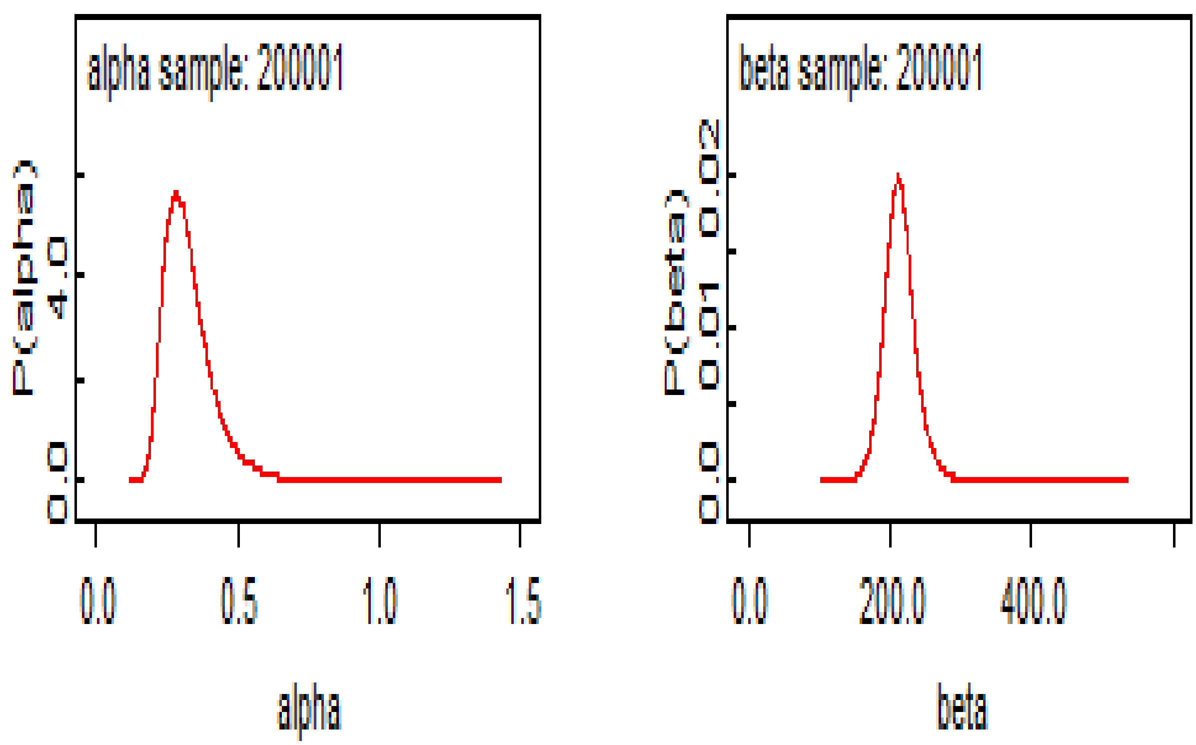

Figure 5.4: Posterior curve of α and β for the Birnbaum-Saunders distribution.

5.1.2 Example 2

In this example, a large dataset consisting of 101 observed fatigue lifetimes is analysed. The data, originally reported by Birnbaum and Saunders (1969a), comprise failure time measurements obtained under cyclic stress conditions and are listed in Table 5.4. The observed lifetimes range from approximately 70 to 212 units, with the majority of observations concentrated between 100 and 160, indicating moderate variability and a right-skewed distribution typical of fatigue failure data. The relatively large sample size allows for stable Bayesian inference and provides a suitable setting for comparing the Weibull and Birnbaum-Saunders models using the same criteria described previously.

Table 5.4: Fatigue lifetime data.

70	90	96	97	99	100	103	104	104	105	107	108	108
112	113	114	114	114	116	119	120	120	120	121	121	123
124	128	128	129	129	130	130	130	131	131	131	131	131
134	134	134	134	134	134	136	136	137	138	138	138	139
142	142	142	142	144	144	145	146	148	148	149	151	151
157	157	157	158	159	162	163	163	164	166	166	168	170
108	109	109	112	124	124	124	124	132	132	132	133	139
141	142	142	152	155	156	157	174	196	212			

The Bayesian Weibull and Birnbaum-Saunders models were implemented in OpenBUGS. For each model, three Markov chains were simulated, each with 500 000 iterations, of which the first 200 000 iterations were discarded as burn-in to remove the influence of initial values. Weakly-informative priors were assigned to the model parameters. Specifically, both α and β parameter were assumed to follow inverse gamma distributions with hyperparameters $a = b = 10^{-4}$.

Table 5.5 presents the posterior summaries for both distribution, including posterior means, medians, standard deviations, and credibility intervals. The Birnbaum–Saunders model shows slightly lower scale estimates and precise credible intervals for both parameters. The sample mean and median are close in value, indicating moderate right skewness, which is also evident from the upper tail of the distribution. This distributional shape is characteristic of fatigue life data and supports the use of asymmetric lifetime models such as the Weibull and Birnbaum–Saunders distributions. Furthermore, the deviance information for the Weibull and Birnbaum–Saunders models is reported in Table 5.6. The DIC value for the Birnbaum-Saunders model is lower than that of the Weibull distribution, indicating better overall fit to the data.

Table 5.5: Posterior Summaries for the Weibull and Birnbaum-Saunders models.

Model	Parameter	Mean	Median	sd	MC_error	95% CI
Weibull	α	6.0270	6.024	0.4229	0.1714	(5.3249, 6.7290)
	β	143.2000	143.2000	2.5220	0.0051	(139.0135, 147.3865)
Birnbaum-Saunders	α	0.1725	0.1718	0.0124	0.0158	(0.1463, 0.1987)
	β	131.8000	131.8000	2.2590	0.0043	(128.0501, 135.5499)

Table 5.6: Deviance information for Example 5.1.2

Model	$\overline{D}$	$\hat{D}$	DIC	p_D
Weibull	926.7	924.7	928.6	1.986
Birnbaum-Saunders	776.5	774.5	778.6	2.0090

Figures 5.5 and 5.6 illustrate the posterior curves for the Weibull and Birnbaum-Saunders models, respectively. The posterior curves for the Weibull and Birnbaum-Saunders parameters show a bell-shaped distribution, reflecting well-behaved, unimodal posteriors. For the Weibull model, the posterior of α is sharply peaked around 6.03, with the mean and median nearly identical, indicating a precise and symmetric estimate. The posterior for β is also concentrated around 143.2, reflecting low uncertainty and an increase in failure rate. In the Birnbaum–Saunders model, the posterior curve for α is sharply peaked near 0.172, showing a decrease in failure rate, while β is concentrated around 131.8 and this shows that there is an increase in failure rate.

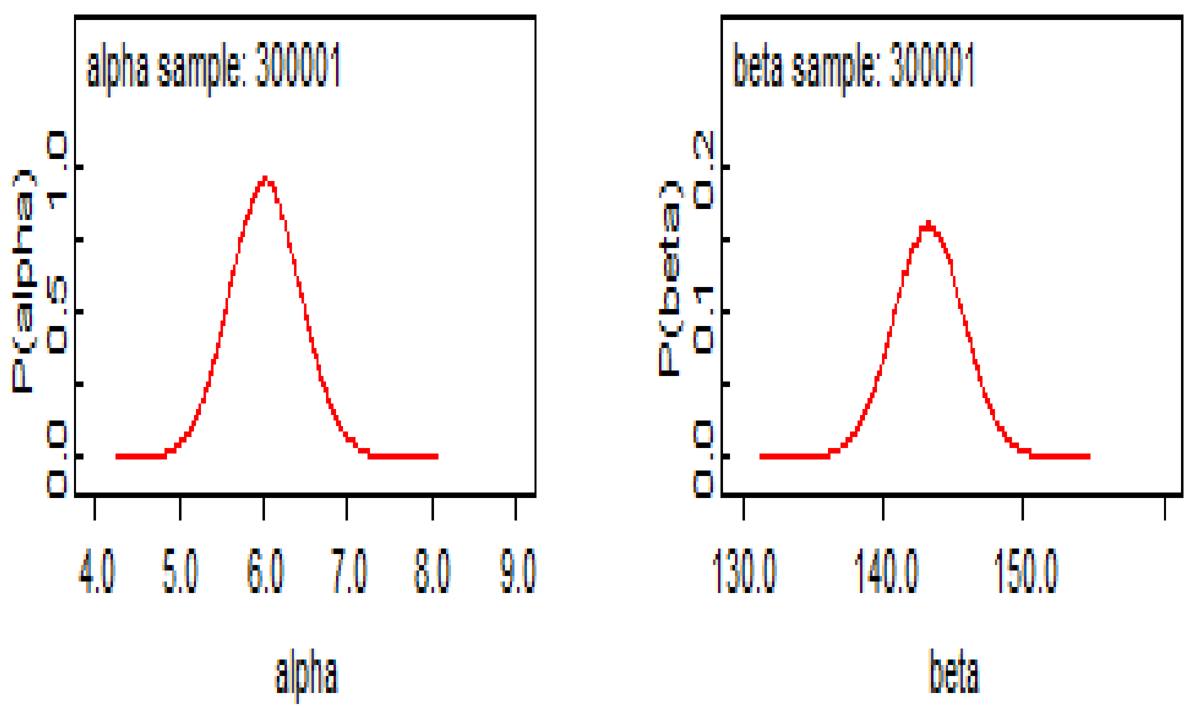

Figure 5.5: Posterior curve of α and β for the Weibull distribution .

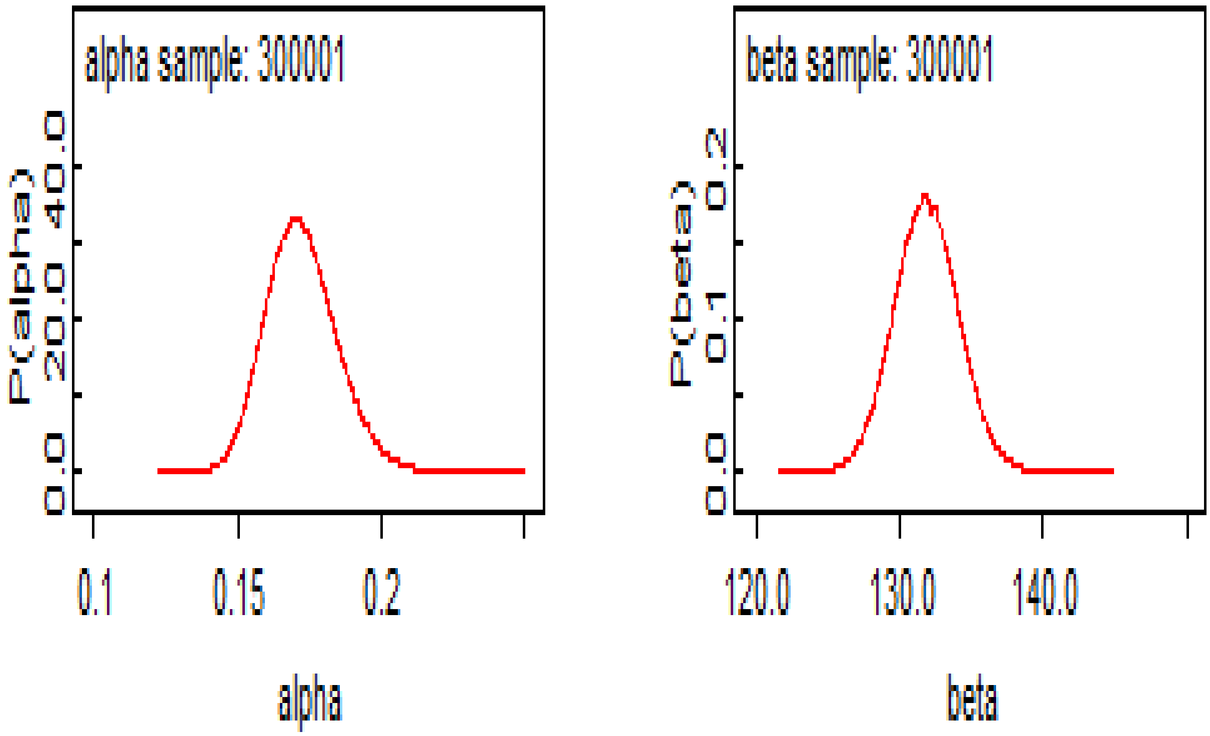

Figure 5.6: Posterior curve of α and β for the Birnbaum-Saunders distribution .

The application of Bayesian Weibull and Birnbaum-Saunders models to complete lifetime datasets demonstrated the effectiveness of Bayesian inference in estimating α and β parameters under both small and large sample sizes. In both examples, posterior distributions captured uncertainty and asymmetry, with small sample data exhibiting right-skewed posteriors and large sample data producing well-behaved, unimodal posteriors. Posterior summaries, including means, medians, and credible intervals, highlighted the differences in parameter estimates between the two distributions. The DIC consistently favoured the Birnbaum-Saunders model, indicating a better overall fit for fatigue lifetime data. Posterior curves for α and β reflected corresponding increases or decreases in failure rates, aligning with the characteristics of fatigue failure. Overall, the analysis illustrates that Bayesian methods provide a flexible and robust framework for reliability modelling and model comparison. The posterior behaviour observed for both models is consistent with fatigue lifetime data, where increasing failure risk over time is expected. The Birnbaum-Saunders model, being explicitly motivated by fatigue failure mechanisms, captures this behaviour naturally, while the Weibull model provides a flexible alternative.

Chapter 6

Concluding Remarks

This chapter presents a summary of conclusion of the key findings of this thesis and outlines potential directions for future research.

6.1 Summary of conclusion

This thesis examined Bayesian reliability analysis for lifetime distributions using objective priors, focusing on the Weibull and Birnbaum-Saunders distributions. The priors considered were, the Jeffreys prior, divergence prior, probability matching prior, and the reference prior. The corresponding posterior distributions were derived for the Weibull distribution, all priors produced proper posterior distributions for both no censoring and type I right censoring. For each distribution, the Fisher information matrix was derived. In particular, deriving the Fisher information matrix for the Birnbaum-Saunders distribution proved to be mathematically challenging due to the distribution's complex parameterization and dependence structure. However, the re-derivation provided important analytical insight and served as the foundation for constructing objective priors. In contrast, for the Birnbaum-Saunders distribution, the use of non-informative priors such as the Jeffreys, probability matching prior and reference priors resulted in improper posterior distributions under no censoring. Convergence of the posterior estimates was carefully examined through Gibbs within MH sampling and MH. Diagnostic measures such as trace plots, Geweke tests, and the Gelman-Rubin were used. The results show that the posterior chains for the Weibull distribution converge when using the non-informative priors considered in this thesis. For the Birnbaum-Saunders distribution, the posterior chains diverge when using the non-informative priors considered in this thesis. To address this issue, weakly informative priors were adopted, and proven that the corresponding posteriors are proper. Simulation studies were then performed to evaluate model performance in terms of coverage rates and mean interval lengths. The results indicated that the Jeffreys and reference priors performed consistently well for the Weibull model. The predictive reliability analyses were conducted, showing that both the Weibull and Birnbaum-Saunders distributions can predict future lifetimes effectively. Applications were done where two fatigue lifetime datasets were fitted and evaluated using both the Weibull and Birnbaum-Saunders distributions. The Bayesian estimation was performed using weakly informative gamma priors for α and β with hyperparameters 10^{-4}. The results showed that both distributions correctly described the observed lifetime data. The Birnbaum-Saunders model consistently produced lower DIC values, indicating a better overall fit. Posterior summaries for both datasets revealed well behaved, unimodal distributions for the parameters, with credible intervals demonstrating reasonable precision. Overall, the Birnbaum-Saunders model provided slightly more accurate and stable parameter estimates compared to the Weibull model, confirming its suitability for modeling fatigue data.

6.2 Future Research

Future research may consider extending this analysis to other life distributions, such as the log-normal or inverse Gaussian, to further assess robustness across various failure mechanisms. Moreover, exploring adaptive MCMC techniques or Hamiltonian Monte Carlo techniques could improve convergence efficiency for complex posterior structures. Evaluating the corresponding posterior behaviour obtained from non-informative priors under Type I censoring for the Birnbaum–Saunders distribution.

References

Abramowitz, M. and Stegun, I. A. (1965). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, volume 55. Courier Corporation.

Achcar, J. A. (1993). Inferences for the Birnbaum-Saunders fatigue life model using Bayesian methods. *Computational Statistics and Data Analysis*, 15(4):367–380.

Akram, M. and Hayat, A. (2014). Comparison of estimators of the Weibull distribution. *Journal of Statistical Theory and Practice*, 8(2):238–259.

Al-Bossly, A. (2020). Bayesian statistics application on reliability prediction and analysis. *Journal of Statistics and Probability*, 9(3):19–34.

Bayes, T. (1763). An Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society*, 53:370 – 418.

Berger, J. O. and Bernardo, J. M. (1989). Estimating a product of means: Bayesian analysis with reference priors. *Journal of the American Statistical Association*, 84(405):200–207.

Berger, J. O. and Bernardo, J. M. (1992). On the development of the reference prior method. *Bayesian Statistics*, 4(4):35–60.

Bernardo, J. M. (1979). Reference posterior distributions for Bayesian inference. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 41(2):113–128.

Birnbaum, Z. W. and Saunders, S. C. (1969a). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*, 6(2):328–347.

Birnbaum, Z. W. and Saunders, S. C. (1969b). A new family of life distributions. *Journal of Applied Probability*, 6(2):319–327.

Carrasco, J. M., Ortega, E. M., and Cordeiro, G. M. (2008). A generalized modified Weibull distribution for lifetime modeling. *Computational Statistics & Data Analysis*, 53(2):450–462.

Consonni, G., Fouskakis, D., Liseo, B., and Ntzoufras, I. (2018). Prior distributions for objective Bayesian analysis.

Datta, G. S. and Ghosh, M. (1995). Some remarks on noninformative priors. *Journal of the American Statistical Association*, 90(432):1357–1363.

Datta, G. S. and Ghosh, M. (1996). On the invariance of noninformative priors. *The Annals of Statistics*, 24(1):141–159.

Datta, G. S. and Sweeting, T. J. (2005). Probability matching priors. *Handbook of Statistics*, 25:91–114.

Ellison, A. M. (2004). Bayesian inference in ecology. *Ecology Letters*, 7(6):509–520.

Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.

Gelman, A. (2002). Prior distribution. *Encyclopedia of Environmetrics*, 3(4):1634–1637.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. *Bayesian statistics*, 4:196–193.

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

Gijbels, I. (2010). Censored Data. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(2):178–188.

Hamada, M. S., Wilson, A. G., and Reese, C. S. Martz, H. F. (2008). *Bayesian Reliability*. Springer, New York.

Harris, C. M. and Singpurwalla, N. D. (1968). Life distributions derived from stochastic hazard functions. *IEEE Transactions on Reliability*, 17(2):70–79.

Harvey, J. and Van der Merwe, A. J. (2018). A Bayesian approach to inference on the variance of lognormal data. *South African Statistical Journal*, 52(1):29–55.

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

Izally, S. R. (2016). Bayesian accelerated life tests: exponential and Weibull models. Masters Thesis, Rhodes University.

Jayalath, K. P. (2024). Improved Bayesian Inferences for Right-Censored Birnbaum–Saunders Data. *Mathematics*, 12(6):874.

Jeffreys, H. (1939). *Theory of Probability*. Oxford University Press, Oxford, 1st edition.

Kang, S. G. and Kim, Y. (2025). Objective Bayesian inference for Birnbaum–Saunders distributions. *Communications in Statistics-Simulation and Computation*, pages 1–20.

Klein, J. P. and Moeschberger, M. L. (2003). Censoring and truncation. *Survival Analysis: Techniques for Censored and Truncated Data, springer, New York.*, pages 63–90.

Kundu, D. and Mitra, D. (2016). Bayesian inference of Weibull distribution based on left truncated and right censored data. *Computational Statistics & Data Analysis*, 99:38–50.

Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, 128(7):912–928.

Lesaffre, E. and Lawson, A. B. (2012). *Bayesian Biostatistics*. John Wiley & Sons, United Kingdom.

Leung, K.-M., Elashoff, R. M., and Afifi, A. A. (1997). Censoring issues in survival analysis. *Annual Review of Public Health*, 18(1):83–104.

Li, H., Yuan, R., Peng, W., Liu, Y., and Huang, H.-Z. (2011). Bayesian inference of Weibull distribution based on probability encoding method. In *2011 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pages 365–369. IEEE.

Lunn, D., Spiegelhalter, D., Thomas, A., and Best, N. (2009). The BUGS project: Evolution, critique and future directions. *Statistics in Medicine*, 28(25):3049–3067.

Martz, H. F. (2014). Bayesian Reliability Analysis. *Wiley StatsRef: Statistics Reference Online*.

Mccool, J. I. (1975). Inferential Techniques for Weibull Populations. II. Technical report.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.

R. Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramos, E., Ramos, P. L., and Louzada, F. (2020). Posterior properties of the Weibull distribution for censored data. *Statistics & Probability Letters*, 166:108873.

Ramos, P. L., Achcar, J. A., Moala, F. A., Ramos, E., and Louzada, F. (2017). Bayesian analysis of the generalized gamma distribution using non-informative priors. *Statistics*, 51(4):824–843.

Robert, C. P., Casella, G., and Casella, G. (1999). *Monte Carlo statistical methods*, volume 2. Springer.

Ross, S. (2014). *Reliability Theory*, pages 559–606.

Roy, V. (2020). Convergence diagnostics for Markov Chain Monte Carlo. *Annual Review of Statistics and Its Application*, 7(1):387–412.

Ruggeri, F., Kenett, S. S., and Faltin, F. W., editors (2008). *Encyclopedia of Statistics in Quality and Reliability*, volume 2. John Wiley & Sons, Chichester.

Sadok, I., Zribi, M., and Masmoudi, A. (2023). Non-informative Bayesian estimation in dispersion models. *Hacettepe Journal of Mathematics and Statistics*, 53(1):251–268.

Smit, N. (2021). *Accelerated life testing using the Eyring model for the Weibull and Birnbaum-Saunders distributions* . PhD thesis, NorthWest University.

Southey, R. (2015). Objective Bayesian Weibull Survival Model. Honours project, unpublished.

Spiegelhalter, D. and Rice, K. (2009). Bayesian statistics. *Scholarpedia*, 4(8):5230.

Spiegelhalter, D., Thomas, A., Best, N., and Lunn, D. (2014). WinBUGS user manual.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and 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.

Sun, D. (1997). A note on noninformative priors for Weibull distributions. *Journal of Statistical Planning and Inference*, 61(2):319–338.

Tian, Q., Lewis-Beck, C., Niemi, J. B., and Meeker, W. Q. (2024). Specifying prior distributions in reliability applications. *Applied Stochastic Models in Business and Industry*, 40(1):5–62.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., and Willemsen, J. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, 1(1):1.

Victor, L. (2025). Mapping research on the Birnbaum–Saunders statistical distribution: Patterns, trends, and scientometric perspective. *Stats*, 8(4):116.

Wang, M., Sun, X., and Park, C. (2016). Bayesian analysis of Birnbaum–Saunders distribution via the generalized ratio-of-uniforms method. *Computational Statistics*, 31(1):207–225.

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics*, 18:287–293.

Welch, B. L. and Peers, H. W. (1963). On formulae for confidence points based on integrals of weighted likelihoods. *Journal of the Royal Statistical Society: Series B (Methodological)*, 25(2):318–329.

Xu, A. and Tang, Y. (2010). Reference analysis for Birnbaum–Saunders distribution. *Computational statistics & data analysis*, 54(1):185–192.

Xu, A. and Tang, Y. (2011). Bayesian analysis of Birnbaum-Saunders distribution with partial information. *Computational Statistics & Data Analysis*, 55(7):2324–2333.

Zhang, Z., Marin, D., Chesakov, E., Maza, M. M., Drangova, M., and Boykov, Y. (2019). Divergence prior and vessel-tree reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10216–10224.

Appendix A: Additional Theorems and Proofs

Proposition .1. *The Weibull distribution $f(t|\alpha,\beta)$ under trasformation $y = \left(\frac{t}{\alpha}\right)^{\beta}$ is an exponential distribution with mean 1.*

Proof.

Consider the Weibull distribution

$$f(t|\alpha,\beta) = f(t|\alpha,\beta) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\}.$$

let

$$y = \left(\frac{t}{\alpha}\right)^{\beta}$$

then

$$t = \alpha y^{\frac{1}{\beta}}.$$

$$\frac{dt}{dy} = \frac{\alpha}{\beta}y^{\frac{1}{\beta}-1}$$

$$\begin{aligned}
f(y|\alpha,\beta) =& f\left(t = \alpha y^{\frac{1}{\beta}}\right) \mid \frac{dt}{dy} \mid \\
=& \frac{\beta}{\alpha}\left(\frac{\alpha y^{\frac{1}{\beta}}}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{\alpha y^{\frac{1}{\beta}}}{\alpha}\right)^{\beta}\right\}\frac{\alpha}{\beta}y^{\frac{1}{\beta}-1} \\
=& \left(y^{\frac{1}{\beta}}\right)^{\beta-1}\exp\left\{-\left(y^{\frac{1}{\beta}}\right)^{\beta}\right\}y^{\frac{1}{\beta}-1} \\
=& y^{1-\frac{1}{\beta}}\exp\left(-y^{\frac{\beta}{\beta}}\right)y^{\frac{1}{\beta}-1} \\
=& \exp(-y).
\end{aligned}$$

The random variable $y = \left(\frac{t}{\alpha}\right)^{\beta}$ follows an exponential distribution with mean 1, where t is a random variable from a Weibull distribution with shape parameter β and scale parameter α.

Lemma .2. *For $i \geq 1$, define $\gamma_i = \int_0^{\infty}[\log(y)]^i \exp(-y)\,dy$, which is the i^{th}moment of log(y) where y is* *exponential distribution distribution with mean* 1, *Sun (1997).*

A.1 Fisher Information Matrix for the Weibull Distribution

The Fisher information matrix for the Weibull distribution will be derived. This is also derived by Sun (1997).

The Fisher information is given by

$$H(\alpha,\beta) = \begin{bmatrix} -E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha^2}\right) & -E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha\partial\beta}\right) \\ -E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\beta\partial\alpha}\right) & -E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\beta^2}\right) \end{bmatrix}.$$

The Fisher information matrix for the Weibull distribution in the absence of censoring is shown below.

$$\begin{aligned} L(\alpha,\beta|\underline{t}) &= \prod_{i=1}^{n} \frac{\beta}{\alpha}\left(\frac{t_i}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{t_i}{\alpha}\right)^{\beta}\right\} \\ &= \prod_{i=1}^{n} \frac{\beta}{\alpha}\left(\frac{1}{\alpha}\right)^{\beta-1} t_i^{\beta-1} \exp\left\{-\left(\frac{t_i}{\alpha}\right)^{\beta}\right\} \\ &= \beta^n \alpha^{-n\beta} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta-1}\right\} \end{aligned}$$

The next step is to apply *log* to the likelihood function.

$$logL(\alpha,\beta|\underline{t}) = nlog(\beta) - n\beta log(\alpha) - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} + (\beta-1)\sum_{i=1}^{n} t_i.$$

The first partial derivative with respect to α is

$$\frac{\partial logL(\alpha,\beta|\underline{t})}{\partial\alpha} = \frac{-n\beta}{\alpha} + \beta\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta-1}\frac{t_i}{\alpha^2}.$$

The second partial derivative with respect to α is

$$\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha^2} = \frac{n\beta}{\alpha^2} - \frac{\beta(\beta+1)}{\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}.$$

The first partial derivative with respect to β is

$$\frac{\partial logL(\alpha,\beta|\underline{t})}{\partial\beta} = \frac{n}{\beta} - nlog(\alpha) + \sum_{i=1}^{n} log(t_i) - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right).$$

The second partial derivative with respect to β is

$$\begin{aligned}\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\beta^2} &= \frac{-n}{\beta^2} - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right) log\left(\frac{t_i}{\alpha}\right)\\ &= \frac{-n}{\beta^2} - \sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right).\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha\partial\beta} &= \frac{-n}{\alpha} + \frac{1}{\alpha}\sum_{i=1}^{n}\left[\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right) + \left(\frac{t_i}{\alpha}\right)^{\beta}\right]\\ \frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha\partial\beta} &= \frac{-n}{\alpha} + \frac{1}{\alpha}\sum_{i=1}^{n}\left[\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right) + \left(\frac{t_i}{\alpha}\right)^{\beta}\right].\end{aligned}$$

Finding $-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha^2}\right)$.

$$\begin{aligned}-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha^2}\right) &= -E\left(\frac{n\beta}{\alpha^2} - \frac{\beta(\beta+1)}{\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\\ &= \frac{-n\beta}{\alpha^2} + \frac{\beta(\beta+1)}{\alpha^2}\sum_{i=1}^{n} E\left[\left(\frac{t_i}{\alpha}\right)^{\beta}\right]\\ &= \frac{-n\beta}{\alpha^2} + \frac{\beta(\beta+1)}{\alpha^2}\sum_{i=1}^{n} 1\\ &= \frac{-n\beta}{\alpha^2} + \frac{\beta(\beta+1)}{\alpha^2}n\\ &= \frac{-n\beta}{\alpha^2} + \frac{n\beta^2}{\alpha^2} + \frac{n\beta}{\alpha^2}\\ &= \frac{n\beta^2}{\alpha^2}.\end{aligned}$$

Finding $-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha\partial\beta}\right)$.

$$\begin{aligned}E\left[\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right)\right] &= \int_0^{\infty}\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right) f(t)\,dt\\ &= \int_0^{\infty}\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right)\frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}\exp\left\{-\left(\frac{t}{\alpha}\right)^{\beta}\right\} dt\end{aligned}$$

let $y = \left(\frac{t_i}{\alpha}\right)^{\beta}$ then $\frac{t_i}{\alpha} = y^{\frac{1}{\beta}}$ and $dt = \frac{\alpha}{\beta\left(\frac{t_i}{\alpha}\right)^{\beta-1}}dy$.

$$E\left[\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right)\right] = \int_0^\infty y\left[log\left(y^{\frac{1}{\beta}}\right)\right]^2 \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left(-y\right)\frac{\alpha}{\beta\left(\frac{t_i}{\alpha}\right)^{\beta-1}}dy$$
$$= \int_0^\infty y\left[\frac{1}{\beta}log\left(y\right)\right]^2 \exp\left(-y\right)dy$$
$$=\frac{1}{\beta^2}\int_0^\infty y\left[log\left(y\right)\right]^2 exp\left(-y\right)dy.$$

Applying integration by parts, let $u = y\left[log\left(y\right)\right]^2$ and let $dv = \exp\left(-y\right)dy$ then, $\frac{du}{dy}$=$\left[log\left(y\right)\right]^2 + 2log\left(y\right)$ and $v = -\exp\left(-y\right)$, then this follows

$$E\left[\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right)\right] = \frac{1}{\beta^2}\left[\frac{-ylog^2\left(y\right)}{\exp\left(y\right)}\,|_0^\infty + \int_0^\infty\left[log^2\left(y\right)+2log\left(y\right)\right]\exp\left(-y\right)dy\right]$$
$$=\frac{1}{\beta^2}\left[0+\int_0^\infty log^2\left(y\right)\exp\left(-y\right)dy+2\int_0^\infty log\left(y\right)\exp\left(-y\right)dy\right]$$
$$=\frac{1}{\beta^2}\left[\gamma_2+2\gamma_1\right].$$

By using Lemma .2.

$$E\left(\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right)\right) =\beta\int_0^\infty ylog\left(y^{\frac{1}{\beta}}\right)\exp\left(-y\right)dy$$
$$=\beta\int_0^\infty y\frac{1}{\beta}log\left(y\right)\exp\left(-y\right)dy$$
$$=\int_0^\infty ylog\left(y\right)\exp\left(-y\right)dy.$$

Let $u = ylog\left(y\right)$and $dv = \exp\left(-y\right)dy$, then $du = \left(log\left(y\right)+1\right)dy$ and $v = -\exp\left(-y\right)$, then this follows

$$E\left(\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right)\right) = \int_0^\infty ylog\left(y\right)\exp\left(-y\right)dy$$
$$=\frac{-ylog\left(y\right)}{\exp\left(y\right)}\,|_0^\infty + \int_0^\infty \exp\left(-y\right)\left(log\left(y\right)+1\right)dy$$
$$=0+\int_0^\infty \exp\left(-y\right)log\left(y\right)dy+\int_0^\infty \exp\left(-y\right)dy$$
$$=\gamma_1+1$$

$$
\begin{aligned}
-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\alpha\partial\beta}\right) =& -E\left(\frac{-n}{\alpha}+\frac{1}{\alpha}\sum_{i=1}^{n}\left[\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right)+\left(\frac{t_i}{\alpha}\right)^{\beta}\right]\right)\\
=&\frac{n}{\alpha}-\frac{1}{\alpha}\sum_{i=1}^{n}E\left[\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right)+\left(\frac{t_i}{\alpha}\right)^{\beta}\right]\\
=&\frac{n}{\alpha}-\frac{1}{\alpha}\sum_{i=1}^{n}\left[E\left(\beta\left(\frac{t_i}{\alpha}\right)^{\beta} log\left(\frac{t_i}{\alpha}\right)\right)+E\left(\left(\frac{t_i}{\alpha}\right)^{\beta}\right)\right]\\
=&\frac{n}{\alpha}-\frac{1}{\alpha}\sum_{i=1}^{n}\left(1+\gamma_1+1\right)\\
=&\frac{n}{\alpha}-\frac{n}{\alpha}\left(2+\gamma_1\right)\\
=&-\frac{n\left(1+\gamma_1\right)}{\alpha}=-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\beta,\partial\alpha}\right).
\end{aligned}
$$

Finding $-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\beta^2}\right)$.

$$
\begin{aligned}
-E\left(\frac{\partial^2 logL(\alpha,\beta|\underline{t})}{\partial\beta^2}\right) =& -E\left(\frac{-n}{\beta^2}-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right)\right)\\
=&\frac{n}{\beta^2}+\sum_{i=1}^{n}E\left[\left(\frac{t_i}{\alpha}\right)^{\beta} log^2\left(\frac{t_i}{\alpha}\right)\right]\\
=&\frac{n}{\beta^2}+\sum_{i=1}^{n}\frac{2\gamma_1+\gamma_2}{\beta^2}\\
=&\frac{n}{\beta^2}+n\frac{2\gamma_1+\gamma_2}{\beta^2}\\
=&\frac{n\left(1+2\gamma_1+\gamma_2\right)}{\beta^2}.
\end{aligned}
$$

$$
H\left(\alpha,\beta\right)=\left[\begin{array}{cc}\frac{n\beta^2}{\alpha^2} & -\frac{n(1+\gamma_1)}{\alpha}\\ -\frac{n(1+\gamma_1)}{\alpha} & \frac{n(1+2\gamma_1+\gamma_2)}{\beta^2}\end{array}\right]. \tag{1}
$$

The determinant of this Fisher information matrix is given by

$$
\begin{aligned}
\left|H\left(\alpha,\beta\right)\right| =& \frac{n\beta^2}{\alpha^2}\times\frac{n\left(1+2\gamma_1+\gamma_2\right)}{\beta^2}-\frac{n^2\left(1+\gamma_1\right)^2}{\alpha^2}\\
=&\frac{n^2+2n^2\gamma_1+n^2\gamma_2-n^2-2n^2\gamma_1-n^2\gamma_1^2}{\alpha^2}\\
=&\frac{n^2\left(\gamma_2-\gamma_1^2\right)}{\alpha^2}.
\end{aligned}
$$

A.2 Fisher Information Matrix for Birnbaum-Saunders Distribution

The Fisher information matrix for the Birnbaum–Saunders distribution will be derived. Lemoine (2019) also derive this Fishers information but some steps where missing, in this thesis every step is explained fully.

The firsts step is to find the log likelihood which is given as

$$\begin{aligned}\ln L(\alpha,\beta|\underline{t}) &= \ln\left\{\frac{\exp\left(n\alpha^{-2}\right)}{2\alpha\sqrt{2\pi\beta}}\right\} + \sum_{i=1}^{n}\ln\left\{t_i^{-\frac{1}{2}} + \beta t_i^{-\frac{3}{2}}\right\} - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)\\ &= n\alpha^{-2} - n\ln(\alpha) - n\ln\left(2\sqrt{2\pi}\right) - \frac{n}{2}\ln(\beta) + \sum_{i=1}^{n}\ln\left\{t_i^{-\frac{1}{2}} + \beta t_i^{-\frac{3}{2}}\right\} - \frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right).\end{aligned}$$

The first partial derivative with respect to α is

$$\frac{\partial \ln L(\alpha,\beta|\underline{t})}{\partial\alpha} = \frac{-2n}{\alpha^3} - \frac{n}{\alpha} + \frac{1}{\alpha^3}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right).$$

The second partial derivative with respect to α is

$$\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\alpha^2} = \frac{6n}{\alpha^4} + \frac{n}{\alpha^2} - \frac{3}{\alpha^4}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right).$$

Finding fisher information with respect to α

$$\begin{aligned}-E\left(\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\alpha^2}\right) &= -E\left(\frac{6n}{\alpha^4} + \frac{n}{\alpha^2} - \frac{3}{\alpha^4}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)\right)\\ &= -\frac{6n}{\alpha^4} - \frac{n}{\alpha^2} + \frac{3}{\alpha^4}E\left(\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}\right)\right)\\ &= -\frac{6n}{\alpha^4} - \frac{n}{\alpha^2} + \frac{3}{\alpha^4}\left(\sum_{i=1}^{n}\left(\frac{E\left(t_i\right)}{\beta}+\beta E\left(t_i^{-1}\right)\right)\right)\\ &= -\frac{6n}{\alpha^4} - \frac{n}{\alpha^2} + \frac{3}{\alpha^4}\left(\sum_{i=1}^{n}\left(\frac{\beta\left(1+\frac{\alpha^2}{2}\right)}{\beta}+\beta\beta^{-1}\left(1+\frac{\alpha^2}{2}\right)\right)\right)\\ &= -\frac{6n}{\alpha^4} - \frac{n}{\alpha^2} + \frac{3}{\alpha^4}\left(\sum_{i=1}^{n}\left(1+\frac{\alpha^2}{2}+1+\frac{\alpha^2}{2}\right)\right)\\ &= -\frac{6n}{\alpha^4} - \frac{n}{\alpha^2} + \frac{3}{\alpha^4}n\left(2+\alpha^2\right) = \frac{2n}{\alpha^2}.\end{aligned}$$

Cross partial derivative of the log-likelihood function is

$$\begin{aligned}\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\alpha\beta} =& \frac{-1}{\alpha^3}\sum_{i=1}^{n}\frac{t_i}{\beta^2}+\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{1}{t_i}\\ =&\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\alpha\beta}.\end{aligned}$$

$$\begin{aligned}-E\left(\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\alpha\beta}\right) =& -E\left(\frac{-1}{\alpha^3}\sum_{i=1}^{n}\frac{t_i}{\beta^2}+\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{1}{t_i}\right)\\ =&\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{E(t_i)}{\beta^2}-\frac{1}{\alpha^3}\sum_{i=1}^{n}E\left(t_i^{-1}\right)\\ =&\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{\beta\left(1+\frac{\alpha^2}{2}\right)}{\beta^2}-\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{1}{\beta}\left(1+\frac{\alpha^2}{2}\right)\\ =&\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{1}{\beta}\left(1+\frac{\alpha^2}{2}\right)-\frac{1}{\alpha^3}\sum_{i=1}^{n}\frac{1}{\beta}\left(1+\frac{\alpha^2}{2}\right)\\ =&0.\end{aligned}$$

The first partial derivative with respect to β is given by

$$\begin{aligned}\frac{\partial \ln L(\alpha,\beta|\underline{t})}{\partial\beta} =& \frac{-n}{2\beta}+\sum_{i=1}^{n}\frac{t_i^{\frac{-3}{2}}}{\left(t_i^{\frac{-1}{2}}+\beta t_i^{\frac{-3}{2}}\right)}-\sum_{i=1}^{n}\frac{1}{2\alpha^2}\left(\frac{-t_i}{\beta^2}+\frac{1}{t_i}\right)\\ =&\frac{-n}{2\beta}+\sum_{i=1}^{n}\frac{1}{(t_i+\beta)}-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{-t_i}{\beta^2}+\frac{1}{t_i}\right).\end{aligned}$$

The second partial derivative with respect to β is given by

$$\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\beta^2} = \frac{n}{2\beta^2}-\sum_{i=1}^{n}\frac{1}{(t_i+\beta)^2}-\frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta^3}\right).$$

Taking expectations yields

$$\begin{aligned}-E\left[\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\beta^2}\right] =& -E\left[\frac{n}{2\beta^2}-\sum_{i=1}^{n}\frac{1}{(t_i+\beta)^2}-\frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta^3}\right)\right]\\ =&-\frac{n}{2\beta^2}+\sum_{i=1}^{n}E\left\{\frac{1}{(t_i+\beta)^2}\right\}+\frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{E(t_i)}{\beta^3}\right)\end{aligned}$$

$$-E\left[\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\beta^2}\right] = -\frac{n}{2\beta^2}+\sum_{i=1}^{n} E\left\{\frac{1}{(t_i+\beta)^2}\right\}+\frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta\left(1+\frac{\alpha^2}{2}\right)}{\beta^3}\right).$$

Taking expectations yields

$$\begin{aligned}
E\left[\frac{1}{(t+\beta)^2}\right] &= \int_0^\infty (t+\beta)^{-2} f(t|\alpha,\beta)\,dt \\
&= \int_0^\infty (t+\beta)^{-2}\frac{\left(\frac{\beta}{t}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t}\right)^{\frac{3}{2}}}{2\sqrt{2\pi}\alpha\beta}\exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t}{\beta}+\frac{\beta}{t}-2\right)\right\}dt \\
&=\frac{1}{2\sqrt{2\pi}\alpha\beta}\int_0^\infty (t+\beta)^{-2}\left\{\left(\frac{\beta}{t}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t}\right)^{\frac{3}{2}}\right\}\exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t}{\beta}+\frac{\beta}{t}-2\right)\right\}dt
\end{aligned}$$

Let $u=\frac{t}{\beta}\Longleftrightarrow t=u\beta$ and $dt=\beta du$ then

$$\begin{aligned}
E\left[\frac{1}{(t+\beta)^2}\right] &=\frac{1}{2\sqrt{2\pi}\alpha\beta}\int_0^\infty (u\beta+\beta)^{-2}\left\{\frac{1}{u^{\frac{1}{2}}}+\frac{1}{u^{\frac{3}{2}}}\right\}\exp\left\{-\frac{1}{2\alpha^2}\left(u+u^{-1}-2\right)\right\}\beta du \\
&=\frac{1}{2\sqrt{2\pi}\alpha\beta^2}\int_0^\infty (u+1)^{-2}\left\{\frac{1}{u^{\frac{1}{2}}}+\frac{1}{u^{\frac{3}{2}}}\right\}\exp\left\{-\frac{1}{2\alpha^2}\left(u^{\frac{1}{2}}-u^{-\frac{1}{2}}\right)^2\right\}du \\
&=\frac{1}{2\sqrt{2\pi}\alpha\beta^2}\int_0^\infty \frac{u^{-\frac{3}{2}}(u+1)}{(u+1)^2}\exp\left\{-\frac{1}{2\alpha^2}\left(u^{\frac{1}{2}}-u^{-\frac{1}{2}}\right)^2\right\}du \\
&=\frac{1}{2\sqrt{2\pi}\alpha\beta^2}\int_0^\infty \frac{u^{-\frac{3}{2}}}{(u+1)}\exp\left\{-\frac{1}{2\alpha^2}\left(u^{\frac{1}{2}}-u^{-\frac{1}{2}}\right)^2\right\}du.
\end{aligned}$$

Now let $v=\sqrt{u}$ and $du=2vdv$ then

$$\begin{aligned}
E\left[\frac{1}{(t+\beta)^2}\right] &=\frac{1}{2\sqrt{2\pi}\alpha\beta^2}\int_0^\infty \frac{\left(v^2\right)^{-\frac{3}{2}}}{(v^2+1)}\exp\left\{-\frac{1}{2\alpha^2}\left(v-v^{-1}\right)^2\right\}2vdv \\
&=\frac{1}{\sqrt{2\pi}\alpha\beta^2}\int_0^\infty \frac{vv^{-3}}{(v^2+1)}\exp\left\{-\frac{1}{2\alpha^2}\left(v-v^{-1}\right)^2\right\}dv \\
&=\frac{1}{\sqrt{2\pi}\alpha\beta^2}\int_0^\infty \frac{v^{-2}}{(v^2+1)}\exp\left\{-\frac{1}{2\alpha^2}\left(v-v^{-1}\right)^2\right\}dv
\end{aligned}$$

$$E\left[\frac{1}{(t+\beta)^2}\right] = \frac{1}{\sqrt{2\pi}\alpha\beta^2}\int_0^\infty \frac{1}{v^2(v^2+1)}\exp\left\{-\frac{1}{2\alpha^2}\left(v-v^{-1}\right)^2\right\}dv$$
$$= \frac{h(\alpha)}{\sqrt{2\pi}\alpha\beta^2},$$

where $h(\alpha) = \int_0^\infty \frac{1}{v^2\left(v^2+1\right)}\exp\left\{-\frac{1}{2\alpha^2}\left(v-v^{-1}\right)^2\right\}dv$. Let $u = v - v^{-1}$ and $\frac{du}{dv} = \frac{1+v^2}{v^2}$ and $v = \frac{u+\sqrt{u^2+4}}{2}$, then this follows

$$h(\alpha) = \int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left\{\frac{u^2+4+u\sqrt{u^2+4}}{2}\right\}\left\{\frac{u^2+2+u\sqrt{u^2+4}}{2}\right\}} \frac{\frac{u^2+2+u\sqrt{u^2+4}}{2}}{\frac{u^2+4+u\sqrt{u^2+4}}{2}}du$$
$$= \int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left\{\frac{u^2+4+u\sqrt{u^2+4}}{2}\right\}\left\{\frac{u^2+4+u\sqrt{u^2+4}}{2}\right\}}du$$
$$= 4\int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left(u^2+4+u\sqrt{u^2+4}\right)\left(u^2+4+u\sqrt{u^2+4}\right)}du$$
$$= 4\int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left(2u^4+2u^3\sqrt{u^2+4}+16+12u^2+8u\sqrt{u^2+4}\right)}du$$
$$= 2\int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left(u^4+u^3\sqrt{u^2+4}+8+6u^2+4u\sqrt{u^2+4}\right)}du$$
$$= 2\int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left(u^4+8+6u^2+u^3\sqrt{u^2+4}+4u\sqrt{u^2+4}\right)}du$$
$$= 2\int_{-\infty}^\infty \frac{\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{\left((u^2+2)(u^2+4)+u(u^2+4)\sqrt{u^2+4}\right)}du$$
$$= 2\int_{-\infty}^\infty \frac{1}{(u^2+4)\left(u^2+2+u\sqrt{u^2+4}\right)} \frac{u^2+2-u\sqrt{u^2+4}}{u^2+2-u\sqrt{u^2+4}}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$
$$= 2\int_{-\infty}^\infty \frac{u^2+2-u\sqrt{u^2+4}}{(u^2+4)\left(u^2+2+u\sqrt{u^2+4}\right)\left(u^2+2-u\sqrt{u^2+4}\right)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$h(\alpha) = \int_{-\infty}^{\infty} \frac{u^2+2-u\sqrt{u^2+4}}{(u^2+4)\left(u^4-2u^2+u^3\sqrt{u^2+4}+2u^2+4+2u\sqrt{u^2+4}-u^3\sqrt{u^2+4}-2u\sqrt{u^2+4}-u^4\right)}$$

$$\times 2\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$=2\int_{-\infty}^{\infty} \frac{u^2+2-u\sqrt{u^2+4}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}}{(u^2+4)4}du$$

$$=\frac{1}{2}\int_{-\infty}^{\infty} \frac{u^2+2-u\sqrt{u^2+4}}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$=\frac{1}{2}\int_{-\infty}^{\infty} \frac{u^2+4-2-u\sqrt{u^2+4}}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$=\frac{1}{2}\int_{-\infty}^{\infty} \frac{u^2+4}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du-\frac{1}{2}\int_{-\infty}^{\infty} \frac{2}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$-\frac{1}{2}\int_{-\infty}^{\infty} \frac{u\sqrt{u^2+4}}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$=\frac{1}{2}\int_{-\infty}^{\infty} \exp\left\{\frac{-u^2}{2\alpha^2}\right\}du-\int_{-\infty}^{\infty} \frac{1}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du-\frac{1}{2}\int_{-\infty}^{\infty} \frac{u}{\sqrt{u^2+4}}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du$$

$$=\frac{\sqrt{2\pi\alpha^2}}{2}-\int_{-\infty}^{\infty} \frac{1}{(u^2+4)}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du-0,$$

$\int_{-\infty}^{\infty} \frac{u}{\sqrt{u^2+4}}\exp\left\{\frac{-u^2}{2\alpha^2}\right\}du = 0$, since the integral of an odd function over symmetric limits is 0.

The integral above was simplified by using well known techniques such as fraction cancellation, rationaling, factoring, expansion, and multiplication by the conjugate.

Let

$$g(x) = \int_{-\infty}^{\infty} \frac{1}{(u^2+4)}\exp\left\{\frac{-u^2x^2}{2}\right\}du, \tag{2}$$

where $x = \frac{1}{\alpha}$.

Then, differentiating $g(x)$ with respect to x gives

$$
\begin{aligned}
g^{'}(x) &= \int_{-\infty}^{\infty} \frac{-xu^2}{(u^2+4)} \exp\left\{\frac{-u^2x^2}{2}\right\} du \\
&= -x \int_{-\infty}^{\infty} \frac{u^2}{(u^2+4)} \exp\left\{\frac{-u^2x^2}{2}\right\} du \\
&= -x \int_{-\infty}^{\infty} \frac{u^2+4-4}{(u^2+4)} \exp\left\{\frac{-u^2x^2}{2}\right\} du \\
&= -x \int_{-\infty}^{\infty} \exp\left\{\frac{-u^2x^2}{2}\right\} du + 4x \int_{-\infty}^{\infty} \frac{1}{(u^2+4)} \exp\left\{\frac{-u^2x^2}{2}\right\} du \\
&= -x \int_{-\infty}^{\infty} \exp\left\{\frac{-u^2x^2}{2}\right\} du + 4xg(x) \\
&= -x\sqrt{2\pi x^{-2}} + 4xg(x) \\
&= -xx^{-1}\sqrt{2\pi} + 4xg(x) \\
&= -\sqrt{2\pi} + 4xg(x).
\end{aligned}
$$

Finding a solution to the differential equation

$$
\begin{aligned}
g^{'}(x) &= -\sqrt{2\pi} + 4xg(x) \\
g^{'}(x) - 4xg(x) &= -\sqrt{2\pi}.
\end{aligned}
$$

The integrating factor

$$
\begin{aligned}
\mu(x) &= \exp\left\{-\int 4x dx\right\} \\
&= \exp\left\{-2x^2\right\}.
\end{aligned}
$$

Multiplying by the integrating factor

$$
\begin{aligned}
\exp\left\{-2x^2\right\}\left(g^{'}(x) - 4xg(x)\right) &= -\exp\left\{-2x^2\right\}\left(\sqrt{2\pi}\right) \\
\left(\exp\left\{-2x^2\right\} g^{'}(x) - \exp\left\{-2x^2\right\} 4xg(x)\right) &= -\sqrt{2\pi}\exp\left\{-2x^2\right\} \\
\frac{d}{dx}\left\{\exp\left\{-2x^2\right\} g(x)\right\} &= -\sqrt{2\pi}\exp\left\{-2x^2\right\}.
\end{aligned}
$$

Integrate both sides

$$\int \frac{d}{dx}\left\{\exp\left\{-2x^2\right\}g(x)\right\}dx = -\int \sqrt{2\pi}\exp\left\{-2x^2\right\}dx$$
$$\exp\left\{-2x^2\right\}g(x) = -\sqrt{2\pi}\int \exp\left\{-2x^2\right\}dx.$$

Solving the integral on the right hand side, let $u = x\sqrt{2}$ and dx=$\frac{du}{\sqrt{2}}$ then

$$\begin{aligned}\exp\left\{-2x^2\right\}g(x) &= -\frac{\sqrt{2\pi}}{\sqrt{2}}\int \exp\left\{-u^2\right\}du \\ &= -\sqrt{\pi}\int \exp\left\{-u^2\right\}du \\ &= -\sqrt{\pi}\left[\frac{\sqrt{\pi}}{2}erf(u) + K\right],\end{aligned}$$

where K is a constant.
Hence,

$$\begin{aligned}g(x) &= \exp\left\{2x^2\right\}\left[\frac{-\pi}{2}erf(u) - \sqrt{\pi}K\right] \\ &= \exp\left\{2x^2\right\}\left[\frac{-\pi}{2}erf\left(x\sqrt{2}\right) - \sqrt{\pi}K\right].\end{aligned}$$

Note that

$$\begin{aligned}\Phi(2x) &= \frac{1}{2} + \frac{1}{2}erf\left(\frac{2x}{\sqrt{2}}\right) \\ &= \frac{1}{2} + \frac{1}{2}erf\left(x\sqrt{2}\right).\end{aligned}$$

Now

$$\begin{aligned}g(x) &= \exp\left\{2x^2\right\}\left[\frac{-\pi}{2}(2\Phi(2x) - 1) - \sqrt{\pi}K\right] \\ &= \exp\left\{2x^2\right\}\left[\frac{-\pi}{2}(2\Phi(2x)) + \frac{\pi}{2} - \sqrt{\pi}K\right] \\ &= \exp\left\{2x^2\right\}\left[-\pi\Phi(2x) + C\right],\end{aligned}$$

where $C = \frac{\pi}{2} - \sqrt{\pi}K$. To determine the value of C, the expression $g(x)$ is evaluated at $x = 0$

$$\begin{aligned}g(0) &= \exp\{0\}\left[-\pi\Phi(0) + C\right] \\ &= C - \frac{\pi}{2}.\end{aligned}$$

Using equation 2, the same function can be expressed as

$$g(0) = \int_{-\infty}^{\infty} \frac{1}{(u^2+4)} \exp\{0\}\, du$$
$$= \int_{-\infty}^{\infty} \frac{1}{(u^2+4)} du.$$

Let $u = 2x \Longleftrightarrow du = 2dx$ then

$$g(0) = \int_{-\infty}^{\infty} \frac{2}{\left((2x)^2+4\right)} dx$$
$$= \int_{-\infty}^{\infty} \frac{2}{4x^2+4} dx$$
$$= \frac{1}{2} \int_{-\infty}^{\infty} \frac{1}{x^2+1} dx.$$

Let $x = \tan(v) \Longleftrightarrow v = \arctan(x)$ and $\frac{dx}{dv} = \sec^2(v)$.
Remember that $\tan^2(v) + 1 = \sec^2(v)$ and $\arctan(\infty) = \frac{\pi}{2}$, $\arctan(-\infty) = \frac{-\pi}{2}$.
Then, evaluating $g(0)$ gives

$$g(0) = \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\sec^2(v)}{\tan^2(v)+1} dv$$
$$= \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{\sec^2(v)}{\sec^2(v)} dv$$
$$= \frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 1 dv$$
$$= \frac{\pi}{2}.$$

$\frac{\pi}{2} = C - \frac{\pi}{2} \Longleftrightarrow C = \pi.$

Define the auxiliary functions

$$g(x) = \exp\left\{2x^2\right\}\left[-\pi\Phi(2x) + \pi\right],$$

and by substituting $x = \frac{1}{\alpha}$, the following expression is obtained

$$g\left(\frac{1}{\alpha}\right) = \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right].$$

Next, define

$$\begin{aligned} h(\alpha) =& \frac{\sqrt{2\pi\alpha^2}}{2} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right] \\ =& \alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]. \end{aligned}$$

Hence, the expected value can be expressed as

$$E\left(\frac{1}{(t+\beta)^2}\right) = \frac{\alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]}{\alpha\beta^2\sqrt{2\pi}}$$

Then,

$$\begin{aligned} -E\left[\frac{\partial^2 \ln L(\alpha,\beta|\underline{t})}{\partial\beta^2}\right] =& -\frac{n}{2\beta^2} + \sum_{i=1}^{n}\left(\frac{\alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]}{\alpha\beta^2\sqrt{2\pi}}\right) + \frac{1}{\alpha^2}\sum_{i=1}^{n}\left(\frac{\beta\left(1 + \frac{\alpha^2}{2}\right)}{\beta^3}\right) \\ =& -\frac{n}{2\beta^2} + n\left(\frac{\alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]}{\alpha\beta^2\sqrt{2\pi}}\right) + n\left(\frac{\left(1 + \frac{\alpha^2}{2}\right)}{\alpha^2\beta^2}\right) \\ =& -\frac{n}{2\beta^2} + n\left(\frac{\alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]}{\alpha\beta^2\sqrt{2\pi}}\right) + \frac{n}{\alpha^2\beta^2} + \frac{n}{2\beta^2} \\ =& n\left(\frac{\alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]}{\alpha\beta^2\sqrt{2\pi}}\right) + \frac{n}{\alpha^2\beta^2} \\ =& \frac{n\left(1 + \left(\alpha\sqrt{\frac{\pi}{2}} - \pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1 - \Phi\left(\frac{2}{\alpha}\right)\right]\right)(2\pi)^{-\frac{1}{2}}\right)}{\alpha^2\beta^2}. \end{aligned}$$

The Fishers information matrix for the Birnbaum-Saunders distribution is given by

$$H(\alpha,\beta) = \begin{bmatrix} \frac{2n}{\alpha^2} & 0 \\ 0 & \frac{n\left(1+\left(\alpha\sqrt{\frac{\pi}{2}}-\pi\exp\left\{\frac{2}{\alpha^2}\right\}\left[1-\Phi\left(\frac{2}{\alpha}\right)\right]\right)(2\pi)^{-\frac{1}{2}}\right)}{\alpha^2\beta^2} \end{bmatrix}. \tag{3}$$

Note that, according to Abramowitz and Stegun (1965), the upper tail of the standard normal distribution can be approximated as

$$\begin{aligned} 1-\Phi\left(\frac{2}{\alpha}\right) &\approx \frac{1}{\sqrt{2\pi}}\frac{1}{\frac{2}{\alpha}}\exp\left\{-\frac{\left(\frac{2}{\alpha}\right)^2}{2}\right\} \\ &= \frac{\alpha}{2\sqrt{2\pi}}\exp\left\{-\frac{2}{\alpha^2}\right\}. \end{aligned}$$

The determinant of the Fisher information matrix is given by

$$\begin{aligned} |I(\alpha,\beta)| &\approx \frac{2n^2}{\alpha^4\beta^2}\left[1+\alpha^2(2\pi)^{-\frac{1}{2}}-\alpha(2\pi)^{-\frac{1}{2}}\pi\exp\left(\frac{2}{\alpha^2}\right)\frac{\alpha}{2\sqrt{2\pi}}\exp\left(-\frac{2}{\alpha^2}\right)\right] \\ &= \frac{2n^2}{\alpha^4\beta^2}\left[1+\frac{\alpha^2}{2}-\alpha^2(2\pi)^{-\frac{1}{2}}\frac{\pi}{2\sqrt{2\pi}}\exp\left(\frac{2}{\alpha^2}-\frac{2}{\alpha^2}\right)\right] \\ &= \frac{2n^2}{\alpha^4\beta^2}\left[1+\frac{\alpha^2}{2}-\frac{\alpha^2}{4}\right] \\ &= \frac{2n^2}{\alpha^4\beta^2}\left[1+\frac{\alpha^2}{4}\right] \\ &= \frac{2n^2}{\alpha^2\beta^2}\left[\frac{1}{\alpha^2}+\frac{1}{4}\right]. \end{aligned}$$

Appendix B Additional Proofs

Theorem .3. *The posterior of α and β for the Weibull distribution given the observed data using the Jeffreys prior when there is no censoring is a proper distribution .*

Proof. For the posterior to be proper, this must hold

$$\int_0^\infty \int_0^\infty c\beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta = 1,$$

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

$$\int_0^\infty \int_0^\infty \beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta < \infty.$$

Now

$$\int_0^\infty \int_0^\infty \beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^n \left\{\prod_{i=1}^n t_i^\beta\right\} \int_0^\infty \alpha^{-(n\beta+1)} \times \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha d\beta.$$

Consider

$$\int_0^\infty \alpha^{-(n\beta+1)} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha = \int_0^\infty \left(\alpha^\beta\right)^{-n} \alpha^{-1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha,$$

and let $u = \alpha^\beta$, then $du = \beta\alpha^{\beta-1} d\alpha$ and $d\alpha = \frac{\alpha du}{\beta u}$, then

$$\int_0^\infty \left(\alpha^\beta\right)^{-n} \alpha^{-1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha = \int_0^\infty (u)^{-n} \alpha^{-1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} \frac{\alpha du}{\beta u}$$

$$= \frac{1}{\beta}\int_0^\infty (u)^{-n-1} \alpha^{-1+1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du$$

$$= \frac{1}{\beta}\int_0^\infty u^{-n-1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du.$$

The expression above follows Inverse Gamma with parameters n and $\sum_{i=1}^n t_i^\beta$. Then the expression can be rewritten as

$$\frac{1}{\beta}\int_0^\infty u^{-n-1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\} du = \frac{1}{\beta}\Gamma(n)\left(\sum_{i=1}^n t_i^\beta\right)^{-n}.$$

Then,

$$\int_0^\infty\int_0^\infty \beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^n \left\{\prod_{i=1}^n t_i^\beta\right\}\beta^{-1}\Gamma(n)\left(\sum_{i=1}^n t_i^\beta\right)^{-n} d\beta$$

$$= \Gamma(n)\int_0^\infty \beta^{n-1}\left\{\prod_{i=1}^n t_i^\beta\right\}\left(\sum_{i=1}^n t_i^\beta\right)^{-n} d\beta.$$

Since $0 < \Gamma(n) < \infty$, then it must be shown that

$$\int_0^\infty \beta^{n-1}\left\{\prod_{i=1}^n t_i^\beta\right\}\left(\sum_{i=1}^n t_i^\beta\right)^{-n} d\beta < \infty.$$

Using a well known inequality between geometric and arithmetic means,

$$\frac{\left(\prod_{i=1}^n t_i^\beta\right)}{\left(\sum_{i=1}^n t_i^\beta\right)^n} \leq \left\{\prod_{i=1}^n t_i^{-\beta}\right\}$$

as a result

$$\beta^{n-1}\frac{\left(\prod_{i=1}^{n} t_i^{\beta}\right)}{\left(\sum_{i=1}^{n} t_i^{\beta}\right)^n} \leq \beta^{n-1}\left\{\prod_{i=1}^{n} t_i^{-\beta}\right\}.$$

Thus,

$$\begin{aligned}\int_0^{\infty} \beta^{n-1}\frac{\left(\prod_{i=1}^{n} t_i^{\beta}\right)}{\left(\sum_{i=1}^{n} t_i^{\beta}\right)^n} d\beta &\leq \int_0^{\infty} \beta^{n-1}\left\{\prod_{i=1}^{n} t_i^{-\beta}\right\} d\beta \\ &= \int_0^{\infty} \beta^{n-1} \exp\left\{\ln\left(\prod_{i=1}^{n} t_i^{-\beta}\right)\right\} d\beta \\ &= \int_0^{\infty} \beta^{n-1} \exp\left\{-\beta\sum_{i=1}^{n} \ln t_i\right\} d\beta.\end{aligned}$$

The expression above follows gamma distribution with parameters n and $\sum_{i=1}^{n} \ln t_i$, then it follows that

$$\int_0^{\infty} \beta^{n-1}\frac{\left(\prod_{i=1}^{n} t_i^{\beta}\right)}{\left(\sum_{i=1}^{n} t_i^{\beta}\right)^n} d\beta \leq \Gamma(n)\left(\sum_{i=1}^{n} \ln t_i\right)^{-n}$$

$$\Gamma(n)\left(\sum_{i=1}^{n} \ln t_i\right)^{-n} < \infty,$$

as a result

$$\int_0^{\infty}\int_0^{\infty} \beta^n \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{n} t_i^{\beta}\right\} d\alpha d\beta < \infty.$$

This completes the proof and we can conclude that the posterior distribution is proper. □

Theorem .4. *The posterior of α and β for the Weibull distribution given the observed data using the Jeffreys prior under type I right censoring is a proper distribution.*

Proof. To show properness of $\pi_J(\alpha,\beta|\underline{t})$ the following should be true

$$\int_0^{\infty}\int_0^{\infty} c\beta^r \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\} d\alpha d\beta = 1,$$

where c is the normalising constant. For the above to be true, the following must be shown

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\} d\alpha d\beta < \infty.$$

Now,

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\} d\alpha d\beta = \int_0^\infty \beta^r \left\{\prod_{i=1}^{r} t_i^{\beta}\right\} \int_0^\infty \alpha^{-(r\beta+1)} \times \exp\left\{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{\alpha^{\beta}}\right\} d\alpha d\beta.$$

Let $u = \alpha^{\beta}$ then $du = \beta\alpha^{\beta-1}d\alpha$, then it follows that

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\} d\alpha d\beta = \int_0^\infty \beta^{r-1} \left\{\prod_{i=1}^{r} t_i^{\beta}\right\} \int_0^\infty u^{-r-1} \times \exp\left\{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{u}\right\} du d\beta.$$

Consider the expression

$$\int_0^\infty u^{-r-1} \exp\left\{-\frac{\sum_{i=1}^{n} t_i^{\beta}}{u}\right\} du,$$

the above expression follows the Inverse Gamma distribution with parameters r and $\sum_{i=1}^{n} t_i^{\beta}$. Then this follows

$$\begin{aligned}\int_0^\infty \int_0^\infty \beta^r \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\} d\alpha d\beta &= \Gamma(r) \int_0^\infty \beta^{r-1} \left\{\prod_{i=1}^{r} t_i^{\beta}\right\}\left\{\sum_{i=1}^{n} t_i^{\beta}\right\}^{-r} d\beta \\ &= \Gamma(r) \int_0^\infty \beta^{r-1} \frac{\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}}{\left\{\sum_{i=1}^{n} t_i^{\beta}\right\}^{r}} d\beta.\end{aligned}$$

Since $0 < \Gamma(r) < \infty$, then it must be shown that

$$\int_0^\infty \beta^{r-1} \frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} d\beta < \infty.$$

Since

$$\frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} \leq \left\{\prod_{i=1}^{r} t_i^{-\beta}\right\},$$

then,

$$\beta^{r-1} \frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} \leq \beta^{r-1} \left\{\prod_{i=1}^{r} t_i^{-\beta}\right\},$$

for β, n, r and $t_i > 0$, $n > 1$ and $r < n$. Now,

$$\begin{aligned} \beta^{r-1} \left\{\prod_{i=1}^{r} t_i^{-\beta}\right\} &= \beta^{r-1} \exp\left\{\log\left\{\prod_{i=1}^{r} t_i^{-\beta}\right\}\right\} \\ &= \beta^{r-1} \exp\left\{\log\left\{\prod_{i=1}^{r} t_i\right\}^{-\beta}\right\} \\ &= \beta^{r-1} \exp\left\{-\beta \log\left\{\prod_{i=1}^{r} t_i\right\}\right\} \\ \beta^{r-1} \exp\left\{-\beta \log\left\{\prod_{i=1}^{r} t_i\right\}\right\} &\sim Gamma\left(r, \log\left\{\prod_{i=1}^{r} t_i\right\}\right), \end{aligned}$$

then this follows

$$\int_0^\infty \beta^{r-1} \frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} d\beta \leq \int_0^\infty \beta^{r-1} \left\{\prod_{i=1}^{r} t_i^{-\beta}\right\} d\beta = \Gamma(r) \left\{\log\left\{\prod_{i=1}^{r} t_i\right\}\right\}^{-r},$$

therefore

$$\int_0^\infty \beta^{r-1} \frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} d\beta < \infty.$$

Which implies that

$$\int_0^\infty \int_0^\infty \beta^r \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta < \infty.$$

The proof is finished and we can conclude that the posterior is proper. □

Theorem .5. *The joint posterior of* α *and* β *for the Weibull distribution given the observed data using the reference prior when there is no censoring is a proper distribution .*

Proof. For the posterior to be proper, this must hold

$$\int_0^\infty \int_0^\infty c\beta^{n-1} \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta = 1,$$

where c is the normalising constant.

For the above to be true, it must be shown that

$$\int_0^\infty \int_0^\infty \beta^{n-1} \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta < \infty.$$

Now

$$\int_0^\infty \int_0^\infty \beta^{n-1} \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^{n-1} \left\{\prod_{i=1}^n t_i^\beta\right\} \int_0^\infty \alpha^{-(n\beta+1)} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha d\beta.$$

Consider

$$\int_0^\infty \alpha^{-(n\beta+1)} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha = \int_0^\infty \left(\alpha^\beta\right)^{-n} \alpha^{-1} \exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\} d\alpha,$$

and let $u = \alpha^\beta$, then $du = \beta\alpha^{\beta-1} d\alpha$ and $d\alpha = \frac{\alpha du}{\beta u}$. The this follows

$$\int_0^\infty \left(\alpha^\beta\right)^{-n}\alpha^{-1}\exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{\alpha^\beta}\right\}d\alpha = \int_0^\infty (u)^{-n}\alpha^{-1}\exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\}\frac{\alpha du}{\beta u}$$

$$=\frac{1}{\beta}\int_0^\infty (u)^{-n-1}\alpha^{-1+1}\exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\}du$$

$$=\frac{1}{\beta}\int_0^\infty u^{-n-1}\exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\}du.$$

The expression above follows Inverse Gamma with parameters n and $\sum_{i=1}^n t_i^\beta$, it can be simplified to

$$\frac{1}{\beta}\int_0^\infty u^{-n-1}\exp\left\{-\frac{\sum_{i=1}^n t_i^\beta}{u}\right\}du = \frac{1}{\beta}\Gamma(n)\left(\sum_{i=1}^n t_i^\beta\right)^{-n}.$$

Then,

$$\int_0^\infty\int_0^\infty \beta^{n-1}\alpha^{-(n\beta+1)}\exp\left\{-\sum_{i=1}^n\left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^n t_i^\beta\right\}d\alpha d\beta = \int_0^\infty \beta^{n-1}\left\{\prod_{i=1}^n t_i^\beta\right\}\beta^{-1}\Gamma(n)\left(\sum_{i=1}^n t_i^\beta\right)^{-n}d\beta$$

$$=\Gamma(n)\int_0^\infty \beta^{n-2}\left\{\prod_{i=1}^n t_i^\beta\right\}\left(\sum_{i=1}^n t_i^\beta\right)^{-n}d\beta.$$

Since $0<\Gamma(n)<\infty$, then it must be shown that

$$\int_0^\infty \beta^{n-2}\left\{\prod_{i=1}^n t_i^\beta\right\}\left(\sum_{i=1}^n t_i^\beta\right)^{-n}d\beta<\infty.$$

Let $t_k < max(T_1, T_2, ...T_n) = t_{max}$, thus,

$$\int_0^\infty \int_0^\infty \beta^{n-1} \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta \le \int_0^\infty \beta^{n-2} \left\{\prod_{i=1}^n t_i^\beta\right\} \left(\sum_{i=1}^n t_i^\beta\right)^{-n} d\beta$$

$$\le \int_0^\infty \beta^{n-2} \left\{\prod_{i=1}^n t_i^{n\beta}\right\} \left(\sum_{i=1}^n t_i^\beta\right)^{-n} d\beta$$

$$\le \int_0^\infty \beta^{n-2} \frac{t_k^{n\beta}}{t_{max}^{n\beta}} d\beta$$

$$= \int_0^\infty \beta^{n-2} \left(\frac{t_k}{t_{max}}\right)^{n\beta} d\beta$$

$$= \int_0^\infty \beta^{n-2} \exp\left\{n\beta \log\left(\frac{t_k}{t_{max}}\right)\right\} d\beta.$$

Since $t_k < t_{max}$, then the expression above follows gamma distribution with parameters $n-1$ and $-n\log\left(\frac{t_k}{t_{max}}\right)$, then

$$\int_0^\infty \beta^{n-2} \left\{\prod_{i=1}^n t_i^\beta\right\} \left(\sum_{i=1}^n t_i^\beta\right)^{-n} d\beta \le \int_0^\infty \beta^{n-2} \exp\left\{n\beta \log\left(\frac{t_k}{t_{max}}\right)\right\} d\beta$$

$$\int_0^\infty \beta^{n-2} \exp\left\{n\beta \log\left(\frac{t_k}{t_{max}}\right)\right\} d\beta < \infty,$$

as a result

$$\int_0^\infty \int_0^\infty \beta^{n-1} \alpha^{-(n\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^n t_i^\beta\right\} d\alpha d\beta < \infty.$$

This completes the proof. □

Theorem .6. *The joint posterior of α and β for the Weibull distribution given the observed data using the reference prior under type I right censoring is a proper distribution.*

Proof. To show properness of $\pi_J(\alpha, \beta | \underline{t})$ the following should be true

$$\int_0^\infty \int_0^\infty c\beta^{r-1} \alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^n \left(\frac{t_i}{\alpha}\right)^\beta\right\} \left\{\prod_{i=1}^r t_i^\beta\right\} d\alpha d\beta = 1,$$

where c is the normalising constant.

For the above to be true, the following must be shown

$$\int_0^\infty \int_0^\infty \beta^{r-1}\alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^{r} t_i^\beta\right\} d\alpha d\beta < \infty.$$

Now,

$$\int_0^\infty \int_0^\infty \beta^{r-1}\alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^{r} t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^{r-1}\left\{\prod_{i=1}^{r} t_i^\beta\right\}\int_0^\infty \alpha^{-(r\beta+1)} \exp\left\{-\frac{\sum_{i=1}^{n} t_i^\beta}{\alpha^\beta}\right\} d\alpha d\beta.$$

Let $u=\alpha^\beta$ then $du=\beta\alpha^{\beta-1}d\alpha$, it follows that

$$\int_0^\infty \int_0^\infty \beta^{r-1}\alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^{r} t_i^\beta\right\} d\alpha d\beta = \int_0^\infty \beta^{r-2}\left\{\prod_{i=1}^{r} t_i^\beta\right\}\int_0^\infty u^{-r-1} \exp\left\{-\frac{\sum_{i=1}^{n} t_i^\beta}{u}\right\} du d\beta.$$

Consider the expression

$$\int_0^\infty u^{-r-1} \exp\left\{-\frac{\sum_{i=1}^{n} t_i^\beta}{u}\right\} du,$$

the above expression follows the Inverse Gamma distribution with parameters r and $\sum_{i=1}^{n} t_i^\beta$. Thus

$$\begin{aligned}\int_0^\infty \int_0^\infty \beta^{r-1}\alpha^{-(r\beta+1)} \exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^\beta\right\}\left\{\prod_{i=1}^{r} t_i^\beta\right\} d\alpha d\beta &= \Gamma(r)\int_0^\infty \beta^{r-2}\left\{\prod_{i=1}^{r} t_i^\beta\right\}\left\{\sum_{i=1}^{n} t_i^\beta\right\}^{-r} d\beta \\ &= \Gamma(r)\int_0^\infty \beta^{r-2}\frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} d\beta.\end{aligned}$$

Since $0<\Gamma(r)<\infty$, then it must be shown that

$$\int_0^\infty \beta^{r-2}\frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} d\beta < \infty.$$

Since

$$\frac{\left\{\prod_{i=1}^{r} t_i^\beta\right\}}{\left\{\sum_{i=1}^{n} t_i^\beta\right\}^r} \le \left\{\prod_{i=1}^{r} t_i^{-\beta}\right\}$$

then

$$\beta^{r-2}\frac{\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}}{\left\{\sum_{i=1}^{n} t_i^{\beta}\right\}^{r}} \leq \beta^{r-2}\left\{\prod_{i=1}^{r} t_i^{-\beta}\right\},$$

for β, n, r and $t_i > 0$, $n > 1$ and $r < n$.for β, n, r and $t_i > 0$, $n > 1$ and $r < n$. Now

$$\begin{aligned}
\beta^{r-2}\left\{\prod_{i=1}^{r} t_i^{-\beta}\right\} &= \beta^{r-2}\exp\left\{\log\left\{\prod_{i=1}^{r} t_i^{-\beta}\right\}\right\} \\
&= \beta^{r-2}\exp\left\{\log\left\{\prod_{i=1}^{r} t_i\right\}^{-\beta}\right\} \\
&= \beta^{r-2}\exp\left\{-\beta\log\left\{\prod_{i=1}^{r} t_i\right\}\right\}
\end{aligned}$$

$$\beta^{r-2}\exp\left\{-\beta\log\left\{\prod_{i=1}^{r} t_i\right\}\right\} \sim Gamma\left(r, \log\left\{\prod_{i=1}^{r} t_i\right\}\right).$$

Then,

$$\int_0^{\infty} \beta^{r-2}\frac{\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}}{\left\{\sum_{i=1}^{n} t_i^{\beta}\right\}^{r}} d\beta \leq \int_0^{\infty} \beta^{r-2}\left\{\prod_{i=1}^{r} t_i^{-\beta}\right\} d\beta = \Gamma(r)\left\{\log\left\{\prod_{i=1}^{r} t_i\right\}\right\}^{-r}.$$

Therefore

$$\int_0^{\infty} \beta^{r-2}\frac{\left\{\prod_{i=1}^{r} t_i^{\beta}\right\}}{\left\{\sum_{i=1}^{n} t_i^{\beta}\right\}^{r}} d\beta < \infty,$$

which implies that

$$\int_0^{\infty}\int_0^{\infty} \beta^{r-1}\alpha^{-(r\beta+1)}\exp\left\{-\sum_{i=1}^{n}\left(\frac{t_i}{\alpha}\right)^{\beta}\right\}\left\{\prod_{i=1}^{r} t_i^{\beta}\right\} d\alpha d\beta < \infty.$$

This completes the proof. □

B.1 Simplification of the Likelihood for the Birnbaum-Saunders Distribution

The likelihood for the Birnbaum-Saunders with no censoring is given as

$$L(\alpha,\beta|\underline{t}) = \prod_{i=1}^{n} \frac{\sqrt{\frac{t_i}{\beta}}+\sqrt{\frac{\beta}{t_i}}}{2\sqrt{2\pi}\alpha t_i} \exp\left\{-\frac{1}{2\alpha^2}\left(\sqrt{\frac{t_i}{\beta}}-\sqrt{\frac{\beta}{t_i}}\right)^2\right\}.$$

The previous expression can be rewritten as

$$\begin{aligned}
L(\alpha,\beta|\underline{t}) &= \prod_{i=1}^{n} \frac{\left\{\left(\frac{t_i}{\beta}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}\right\} t_i^{-1}}{2\sqrt{2\pi}\alpha} \exp\left\{-\frac{1}{2\alpha^2}\left(\sqrt{\frac{t_i}{\beta}}-\sqrt{\frac{\beta}{t_i}}\right)\left(\sqrt{\frac{t_i}{\beta}}-\sqrt{\frac{\beta}{t_i}}\right)\right\} \\
&= \prod_{i=1}^{n} \frac{\left\{\frac{t_i^{\frac{1}{2}}}{\beta^{\frac{1}{2}}}+\frac{\beta^{\frac{1}{2}}}{t_i^{\frac{1}{2}}}\right\} t_i^{-1}}{2\sqrt{2\pi}\alpha} \exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t_i}{\beta}-2\sqrt{\frac{t_i}{\beta}}\sqrt{\frac{\beta}{t_i}}+\frac{\beta}{t_i}\right)\right\} \\
&= \prod_{i=1}^{n} \frac{\left\{\frac{t_i^{-\frac{1}{2}}}{\beta^{\frac{1}{2}}}t_i^{-1}+\frac{\beta^{\frac{1}{2}}}{t_i^{\frac{1}{2}}}t_i^{-1}\right\}}{2\sqrt{2\pi}\alpha} \exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t_i}{\beta}-2+\frac{\beta}{t_i}\right)\right\}.
\end{aligned}$$

Taking β^{-1} as a common factor in the previous expression, lead to

$$\begin{aligned}
L(\alpha,\beta|\underline{t}) &= \prod_{i=1}^{n} \frac{\beta^{-1}\left\{\frac{t_i^{-\frac{1}{2}}}{\beta^{-\frac{1}{2}}}+\frac{\beta^{\frac{3}{2}}}{t_i^{\frac{3}{2}}}\right\}}{2\sqrt{2\pi}\alpha} \exp\left\{-\frac{1}{2\alpha^2}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\} \\
&= \frac{\beta^{-n}\prod_{i=1}^{n}\left\{\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right\}}{\left(2\sqrt{2\pi}\right)^n\alpha^n} \exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.
\end{aligned}$$

By taking out constant $\left(2\sqrt{2\pi}\right)^n$ in the previous expression, then

$$L(\alpha,\beta|\underline{t}) \propto \alpha^{-n}\beta^{-n}\prod_{i=1}^{n}\left\{\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right\}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}. \tag{4}$$

B.2 Deriving the Reference Prior for the Birnbaum-Saunders Distribution

The reference prior was derived by Xu and Tang (2010), using Berger and Bernardo (1992) algorithm. In this work, I applied the one-at-a-time algorithm to derive the reference prior and obtained the same result.

The procedure begins with the inverse of the Fisher information matrix

$$H^{-1}(\alpha,\beta) = \frac{\alpha^4\beta^2}{2n^2\left[1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right]}\begin{bmatrix} \frac{n\left(1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right)}{\alpha^2\beta^2} & 0 \\ 0 & \frac{2n}{\alpha^2}\end{bmatrix}$$
$$= \begin{bmatrix} \frac{\alpha^2}{2n} & 0 \\ 0 & \frac{\alpha^2\beta^2}{n\left(1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right)}\end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22}\end{bmatrix}.$$

Thus,

$$h_1 = A_{11}^{-1} = \frac{2n}{\alpha^2},$$

and

$$h_2 = \left(A_{22} - B_2H_1B_2^{\intercal}\right)^{-1} = \left(\frac{\alpha^2\beta^2}{n\left(1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right)} - 0\frac{2n}{\alpha^2}0^t\right)^{-1}$$
$$= \frac{n\left(1+\alpha(2\pi)^{-\frac{1}{2}}g(\alpha)\right)}{\alpha^2\beta^2}.$$

Now, let $\alpha \in (a,b)$ and $\beta \in (e,f)$, where a and b are the bounds of α parameter and e and f are the bounds of β parameter, this follows

$$\begin{aligned}
\pi_2^l(\beta \mid \alpha) =& \frac{\mid h_2 \mid^{\frac{1}{2}}}{\int_e^f \mid h_2 \mid^{\frac{1}{2}} d\beta} \\
=& \frac{\mid \frac{n\left(1+\alpha(2\pi)^{-\frac{1}{2}} g(\alpha)\right)}{\alpha^2\beta^2} \mid^{\frac{1}{2}}}{\int_e^f \mid \frac{n\left(1+\alpha(2\pi)^{-\frac{1}{2}} g(\alpha)\right)}{\alpha^2\beta^2} \mid^{\frac{1}{2}} d\beta} \\
=& \frac{\mid n\left(1+\alpha(2\pi)^{-\frac{1}{2}} g(\alpha)\right) \mid^{\frac{1}{2}} \alpha^{-1}\beta^{-1}}{\mid n\left(1+\alpha(2\pi)^{-\frac{1}{2}} g(\alpha)\right) \mid^{\frac{1}{2}} \int_e^f \frac{1}{\alpha\beta} d\beta} \\
=& \frac{\alpha^{-1}\beta^{-1}}{\alpha^{-1}\int_e^f \frac{1}{\beta} d\beta} = \frac{\beta^{-1}}{\log(fe^{-1})}.
\end{aligned}$$

Next, the conditional prior for α given β is

$$\pi_1^l(\alpha \mid \beta) = \frac{\pi_2^l(\beta \mid \alpha) \exp\left\{E\left[\log\{h_1(\theta)\}^{\frac{1}{2}}\right]\right\}}{\int_a^b \exp\left\{E\left[\log\{h_1(\theta)\}^{\frac{1}{2}}\right]\right\} d\alpha}.$$

Now,

$$\begin{aligned}
E\left[\log\{h_1(\theta)\}^{\frac{1}{2}}\right] =& E\left[\log\left\{\frac{2n}{\alpha^2}\right\}^{\frac{1}{2}}\right] \\
=& E\left[\log\left(\frac{\sqrt{2n}}{\alpha}\right)\right] \\
=& \int_e^f \log\left(\frac{\sqrt{2n}}{\alpha}\right) \frac{\beta^{-1}}{\log(fe^{-1})} d\beta \\
=& \frac{\log\left(\frac{\sqrt{2n}}{\alpha}\right)}{\log(fe^{-1})} \int_e^f \beta^{-1} d\beta
\end{aligned}$$

$$E\left[\log\{h_1(\theta)\}^{\frac{1}{2}}\right] = \log\left(\frac{\sqrt{2n}}{\alpha}\right).$$

Hence,

$$\pi^{l}(\alpha,\beta)=\frac{\frac{\beta^{-1}}{\log\left(fe^{-1}\right)}\exp\left(\log\left(\frac{\sqrt{2n}}{\alpha}\right)\right)}{\int_{a}^{b}\exp\left(\log\left(\frac{\sqrt{2n}}{\alpha}\right)\right)d\alpha}$$

$$=\frac{\frac{\beta^{-1}}{\log\left(fe^{-1}\right)}\sqrt{2n}\alpha^{-1}}{\sqrt{2n}\int_{a}^{b}\alpha^{-1}d\alpha}$$

$$=\frac{\frac{\beta^{-1}}{\log\left(fe^{-1}\right)}\alpha^{-1}}{\int_{a}^{b}\alpha^{-1}d\alpha}$$

$$=\frac{\frac{\alpha^{-1}\beta^{-1}}{\log\left(fe^{-1}\right)}}{\log\left(ba^{-1}\right)}.$$

Finally, the reference prior for the Birnbaum Saunders distribution is given as

$$\pi_R(\alpha,\beta)=\lim_{l\to\infty}\frac{\pi^{l}(\alpha,\beta)}{\pi^{l}(\alpha^{*},\beta^{*})}$$

$$\propto\lim_{l\to\infty}\frac{\alpha^{-1}\beta^{-1}}{\log\left(ba^{-1}\right)}$$

$$\pi_R(\alpha,\beta)\propto\alpha^{-1}\beta^{-1}.$$

Theorem .7. *The posterior of α and β for the Birnbaum-Saunders distribution given the observed data using the reference prior when there is no censoring is an improper distribution.*

Proof. If the following result holds

$$\int_{0}^{\infty}\int_{0}^{\infty}\pi_R(\alpha,\beta|t)\,d\alpha d\beta=\infty,$$

then the posterior distribution is improper.

The posterior distribution of α and β given the observed data using the reference prior when there is no censoring is given by

$$\pi_R(\alpha,\beta|\underline{t})\propto\alpha^{-n-1}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}.$$

Taking integral $\pi_R(\alpha,\beta|\underline{t})$ of with respect to α, then this follows

$$\int_0^\infty \pi_R(\alpha,\beta|\underline{t})\,d\alpha \propto \int_0^\infty \alpha^{-n-1}\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha$$

$$=\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\int_0^\infty \alpha^{-n-1}\exp\left\{-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\alpha.$$

If we let $\lambda=\alpha^2$, then we have

$$\int_0^\infty \pi_R(\alpha,\beta|\underline{t})\,d\alpha \propto \beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]\int_0^\infty \lambda^{-\frac{n}{2}-1}\exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}d\lambda.$$

From the above it can be observe that $\lambda^{-\frac{n}{2}-1}\exp\left\{-\frac{1}{2\lambda}\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right\}$ is a kernal of an inverse-gamma distribution,

then

$$\pi_R(\beta|\alpha,\underline{t}) \propto \frac{\beta^{-n-1}\prod_{i=1}^{n}\left[\left(\frac{\beta}{t_i}\right)^{\frac{1}{2}}+\left(\frac{\beta}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta}+\frac{\beta}{t_i}-2\right)\right]^{\frac{n}{2}}} = \frac{\beta^{-n-1+\frac{3n}{2}}\prod_{i=1}^{n}\left[\frac{1}{\beta(t_i)^{\frac{1}{2}}}+\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\beta^{\frac{n}{2}}\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{2}{\beta}\right)\right]^{\frac{n}{2}}}$$

$$=\frac{\beta^{-1}\prod_{i=1}^{n}\left[\frac{1}{\beta(t_i)^{\frac{1}{2}}}+\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{t_i}{\beta^2}+\frac{1}{t_i}-\frac{n}{\beta}\right)\right]^{\frac{n}{2}}}.$$

Note that,

$$\lim_{\beta\to\infty}\frac{\pi_R(\beta|\alpha,\underline{t})}{\beta^{-1}} = \frac{\prod_{i=1}^{n}\left[\left(\frac{1}{t_i}\right)^{\frac{3}{2}}\right]}{\left[\sum_{i=1}^{n}\left(\frac{1}{t_i}\right)\right]^{\frac{n}{2}}}.$$

Thus, as $\beta\to\infty$,

$$\pi_R(\beta|\alpha,\underline{t}) = O\left(\frac{1}{\beta}\right).$$

For large K,

$$\int_0^\infty \pi_R(\beta|\alpha,\underline{t})\,d\beta \geq \int_K^\infty \pi_R(\beta|\alpha,\underline{t})\,d\beta \to \infty.$$

This completes the proof.

This is also shown in Xu and Tang (2011).

Appendix C: Code

C.1 Weibull Distribution Code

```
#(Jeffreys Prior)
#Data and initialization
set.seed(123)
To select a specific sample size  n, simply remove the comment symbol (#)
#n <- 10
#n<-40
#n<-200
t <- rweibull(n, shape = 3, scale = 2)  # true α=2, β=3
T <- 20000
burnin <- 10000
sigma <- 0.1
alpha <- 1; beta <- 1
sumlogt <- sum(log(t))
alpha_samp <- numeric(T)
beta_samp <- numeric(T)
accept_beta <- logical(T)
#2. Log-posterior for β (given α and t)
logpost_beta <- function(b, alpha) {
if (b <= 0) return(-Inf)
(n-1)*log(b) +
(b - 1)*sumlogt - n*b*log(alpha) - sum((t/alpha)^b) }
#3. MCMC loop
for (k in seq_len(T))
{   # Gibbs for α
S <- sum(t^beta)
u <- rgamma(1, shape = n , rate = S)
alpha <- u^(-1 / beta)
```

```
# MH for β
logb_prop <- rnorm(1, mean = log(beta), sd = sigma)
 beta_prop <- exp(logb_prop)
log_r <- logpost_beta(beta_prop, alpha) - logpost_beta(beta, alpha)
 if (log(runif(1)) < log_r) {
 beta <- beta_prop
  accept_beta[k] <- TRUE
}
alpha_samp[k] <- alpha
beta_samp[k] <- beta
}
# 4. Diagnostics and plots
burn <- burnin + 1:T - burnin
alpha_post <- alpha_samp[burn]
beta_post <- beta_samp[burn]
# Trace & density
par(mfrow = c(2,2))
plot(alpha_post, type = 'l', main = 'Trace: alpha')
hist(alpha_post, breaks=40, main='Posterior alpha', prob=TRUE)
plot(beta_post, type = 'l', main = 'Trace: beta')
hist(beta_post, breaks=40, main='Posterior beta', prob=TRUE)
# Posterior estimates
cat("Alpha_hat:", mean(alpha_post), "\n")
cat("Beta_hat:", mean(beta_post), "\n")
# Acceptance rate
cat("Beta acceptance rate:", mean(accept_beta), "\n")
# Autocorrelation using base R
acf_vals <- acf(beta_post, plot=FALSE)
plot(acf_vals, main = "ACF: beta")
# Predictive Reliability
pre_reliability <- function(alpha_post, beta_post, x_vals) {
sapply(x_vals, function(x) { mean(exp(-(x / alpha_post)^beta_post)) }) }
x_vals <- seq(0, 5, by = 1/60)
rel_jeffreys <- Pred_reliability(alpha_post, beta_post, x_vals)
```

 # Convergence Diagnostics (Jeffreys prior)

```
library(coda)
```

```
# Single-chain Weibull MCMC sampler
weibull_chain <- function(t, T = 20000, sigma = 0.1,
alpha0 = 1, beta0 = 1) {
n <- length(t) sumlogt <- sum(log(t))
alpha <- alpha0; beta <- beta0 s
amples <- matrix(NA, nrow = T, ncol = 2)
colnames(samples) <- c("alpha", "beta")
logpost_beta <- function(b, alpha) {
if (b <= 0) return(-Inf)
-(n+1)*log(b) + (b - 1)*sumlogt - n*b*log(alpha) - sum((t/alpha)^b)
}
for (k in 1:T) {
S <- sum(t^beta)
u <- rgamma(1, shape = n + 1, rate = S)
alpha <- u^(-1 / beta)
logb_prop <- rnorm(1, mean = log(beta), sd = sigma)
beta_prop <- exp(logb_prop)
log_r <- logpost_beta(beta_prop, alpha) - logpost_beta(beta, alpha)
if (log(runif(1)) < log_r) beta <- beta_prop
samples[k, ] <- c(alpha, beta)
}
mcmc(samples)
}
# Simulate data
set.seed(123)
t <- rweibull(40, shape = 3, scale = 2)
# Run 3 chains with different starting values
chains <- mcmc.list(
weibull_chain(t, alpha0 = 0.5, beta0 = 0.5),
weibull_chain(t, alpha0 = 2, beta0 = 1),
weibull_chain(t, alpha0 = 5, beta0 = 5)
)
# Discard the first half as burn-in
early <- niter(chains) / 2
chains_trimmed <- window(chains, start = early + 1)
# Gelman-Rubin diagnostic
gr <- gelman.diag(chains_trimmed, autoburnin = FALSE)
```

```
print(gr)
gelman.plot(chains_trimmed, autoburnin = FALSE)
# Geweke diagnostic for first chain
ge <- geweke.diag(chains_trimmed[[1]])
print(ge)
geweke.plot(chains_trimmed[[1]])
# Trace & density plots plot(chains_trimmed)
# Effective sample sizes
cat("Effective sample sizes:\n")
print(effectiveSize(chains_trimmed))
# For coverage rate and mean interval length
library(coda)
# MCMC sampler function
weibull_chain <- function(t, T=5000, sigma=0.1, alpha0=1, beta0=1) {
n <- length(t); sumlogt <- sum(log(t))
alpha <- alpha0; beta <- beta0
samps <- matrix(NA, nrow=T, ncol=2)
colnames(samps) <- c("alpha","beta") l
logpost_beta <- function(b,a) {
if (b<=0) return(-Inf)
(n)*log(b)+(b-1)*sumlogt - n*b*log(a) - sum((t/a)^b)
}
for (i in 1:T) {
# Gibbs for alpha
S <- sum(t^beta)
u <- rgamma(1, shape=n+1, rate=S)
alpha <- u^(-1/beta)
# MH for beta
logb_prop <- rnorm(1, log(beta), sigma)
beta_prop <- exp(logb_prop) l
log_r <- logpost_beta(beta_prop, alpha) - logpost_beta(beta, alpha)
if (log(runif(1)) < log_r) beta <- beta_prop
samps[i,] <- c(alpha, beta)
}
mcmc(window(mcmc(samps), start=T/2)) # discard first half }
# Function to compute coverage & interval length per simulation
single_run <- function(alpha_true, beta_true, n=100) {
```

```
t <- rweibull(n, shape=beta_true, scale=alpha_true)
post <- weibull_chain(t, T=10000, sigma=0.1,
alpha0=alpha_true, beta0=beta_true)
ci_alpha <- HPDinterval(post[,"alpha"], prob=0.95)
ci_beta <- HPDinterval(post[,"beta"], prob=0.95)
cover_alpha<-(alpha_true>=ci_alpha[1,"lower"]&&
alpha_true<=ci_alpha[1,"upper"])
cover_beta<-(beta_true >= ci_beta[1,"lower"]&&
beta_true <= ci_beta[1,"upper"])
len_alpha <- ci_alpha[,"upper"] - ci_alpha[,"lower"]
len_beta <- ci_beta[,"upper"] - ci_beta[,"lower"]
return(c(cover_alpha, cover_beta, len_alpha, len_beta)) }
# Grid of true values
alphas_true <- c(0.5, 2, 5)
betas_true <- c(0.8, 5, 8)
set.seed(42)
reps <- 10000 # number of datasets per scenario
results <- expand.grid(alpha_t=alphas_true, beta_t=betas_true)
out <- data.frame(results, coverage_alpha=0, coverage_beta=0,
mean_len_alpha=0, mean_len_beta=0)
# Simulate for each scenario
for (i in seq_len(nrow(results))) {
a0 <- results$alpha_t[i]
b0 <- results$beta_t[i]
sim_res <- replicate(reps, single_run(a0, b0))
out$coverage_alpha[i] <- mean(sim_res[1,])
out$coverage_beta[i] <- mean(sim_res[2,])
out$mean_len_alpha[i] <- mean(sim_res[3,])
out$mean_len_beta[i] <- mean(sim_res[4,]) }
print(out)
Note that, when using  the divergence prior you change,u to be
u <- rgamma(1, shape=n -1/2*beta , rate=S)in the code above.
When using the reference prior you change, logpost_beta to be
logpost_beta <- function(b,a) {if (b<=0) return(-Inf)
(n -1)*log(b)+(b-1)*sumlogt - n*b*log(a) - sum((t/a)^b) }
```

C.2 Birnbaum-Saunders Distribution Code

```
#( First prior)
#  Data and initialization
set.seed(123)
#  To select a specific sample size  n, simply remove the comment symbol (#)
#n <-10
#n <-40
#n <-200
# Birnbaum-Saunders random generator
rBS <- function(n, alpha, beta) {
Z <- rnorm(n)
Y <- (beta / 4) * (alpha * Z + sqrt((alpha * Z)^2 + 4))^2
 return(Y) }
t <- rBS(n, alpha_true=0.5, beta_true=3)
T <- 20000
burnin <- 10000
sigma <- 0.35
alpha <- 1; beta <- 1
alpha_samp <- numeric(T)
beta_samp <- numeric(T)
accept_beta <- logical(T)
#2. Log-posterior for β (given α and t)
logpost_beta <- function(b, alpha) {
  if (b <= 0) return(-Inf)
val <- -(n + a_1 + 1) * log(b) - b_1 / b +
 sum(log((b / t)^0.5 + (b / t)^1.5)) -
(1 / (2 * alpha^2)) * sum(t / b + b / t)
return(val) }
#3. MCMC loop
for (k in seq_len(T))
{   # Gibbs for α
 S <- sum(t / beta_1 + beta_1 / t - 2)
u <- rgamma(1, shape =  (n)/2 , rate = S)
alpha <- 1/sqrt(u)
# MH for β
logb_prop <- rnorm(1, mean = log(beta), sd = sigma)
```

```
 beta_prop <- exp(logb_prop)
log_r <- logpost_beta(beta_prop, alpha) - logpost_beta(beta, alpha)
 if (log(runif(1)) < log_r) {
 beta <- beta_prop
  accept_beta[k] <- TRUE
}
alpha_samp[k] <- alpha
beta_samp[k] <- beta
}
# 4. Diagnostics and plots
burn <- burnin + 1:T - burnin
alpha_post <- alpha_samp[burn]
beta_post <- beta_samp[burn]
# Trace & density
par(mfrow = c(2,2))
plot(alpha_post, type = 'l', main = 'Trace: alpha')
hist(alpha_post, breaks=40, main='Posterior alpha', prob=TRUE)
plot(beta_post, type = 'l', main = 'Trace: beta')
hist(beta_post, breaks=40, main='Posterior beta', prob=TRUE)
# Posterior estimates
cat("Alpha_hat:", mean(alpha_post), "\n")
cat("Beta_hat:", mean(beta_post), "\n")
# Acceptance rate
cat("Beta acceptance rate:", mean(accept_beta), "\n")
# Autocorrelation using base R
acf_vals <- acf(beta_post, plot=FALSE)
plot(acf_vals, main = "ACF: beta")
# Predictive Reliability
predictive_reliability <- function(x, alpha_post, beta_post) {
mean(1 - pnorm((1 / alpha_post) * (sqrt(x / beta_post)
- sqrt(beta_post / x))))
}
x <- seq(0, 10, by=1/60)
pred_rel <- sapply(x,
predictive_reliability, alpha_post = alpha_post_1, beta_post = beta_post_1)
```

 # Convergence Diagnostics

```
library(coda)
# Single-chain Birnbaum-Saunders  MCMC sampler
BS_chain <- function(t, T = 20000, sigma = 0.4, alpha0 = 1, beta0 = 1) {
n <- length(t)
alpha <- alpha0; beta <- beta0
samples <- matrix(NA, nrow = T, ncol = 2)
colnames(samples) <- c("alpha", "beta")
logpost_beta <- function(b, alpha) {
if (b <= 0) return(-Inf)
val <- -(n + 1) * log(b) +
sum(log((b / t)^0.5 + (b / t)^1.5))
- (1 / (2 * alpha^2)) * sum(t / b + b / t)
return(val)
}
for (k in seq_len(T)) {
S <- sum(t / beta + beta / t - 2)
u <- rgamma(1, shape = a_1+(n+1) / 2, rate =b_1+ S / 2)
alpha <- 1 / sqrt(u)
logb_prop <- rnorm(1, mean = log(beta), sd = sigma)
beta_prop <- exp(logb_prop)
log_r <- logpost_beta(beta_prop, alpha) - logpost_beta(beta, alpha)
if (log(runif(1)) < log_r) beta <- beta_prop
samples[k, ] <- c(alpha, beta)
}
mcmc(samples)
}
# Simulate data
set.seed(123)
n=200
a_1<-b_1<-10^(-3)
lpha_true <- 0.5 beta_true <- 3
# Birnbaum-Saunders random generator
rBS <- function(n, alpha, beta) {
Z <- rnorm(n)
Y <- (beta / 4) * (alpha * Z + sqrt((alpha * Z)^2 + 4))^2
return(Y) }
t <- rBS(n, alpha_true, beta_true)
```

```
# Run 3 chains with different starting values
chains <- mcmc.list( B
S_chain(t, alpha0 = 0.5, beta0 = 0.5),
BS_chain(t, alpha0 = 2, beta0 = 1),
BS_chain(t, alpha0 = 5, beta0 = 5)
)
# Discard the first half as burn-in
early <- niter(chains) / 2
chains_trimmed <- window(chains, start = early + 1)
# Gelman-Rubin diagnostic
gr <- gelman.diag(chains_trimmed, autoburnin = FALSE)
print(gr)
gelman.plot(chains_trimmed, autoburnin = FALSE)
# Geweke diagnostic for first chain
ge <- geweke.diag(chains_trimmed[[1]])
print(ge)
geweke.plot(chains_trimmed[[1]])
# Trace & density plots
plot(chains_trimmed)
# Effective sample sizes
cat("Effective sample sizes:\n")
print(effectiveSize(chains_trimmed))
# For coverage rate and mean interval length
library(coda)
# MCMC sampler function
BS_chain <- function(t, T=5000, sigma=0.1, alpha0=1, beta0=1) {
n <- length(t)
alpha <- alpha0; beta <- beta0
samps <- matrix(NA, nrow=T, ncol=2)
colnames(samps) <- c("alpha","beta") l
logpost_beta <- function(b, alpha) {
  if (b <= 0) return(-Inf)
val <- -(n + a_1 + 1) * log(b) - b_1 / b +
sum(log((b / t)^0.5 + (b / t)^1.5)) -
(1 / (2 * alpha^2)) * sum(t / b + b / t)
return(val) }
#3. MCMC loop
```

```
for (k in seq_len(T))
{   # Gibbs for α
 S <- sum(t / beta_1 + beta_1 / t - 2)
u <- rgamma(1, shape =  (n)/2 , rate = S)
alpha <- 1/sqrt(u)
# MH for beta
logb_prop <- rnorm(1, log(beta), sigma)
beta_prop <- exp(logb_prop) l
log_r <- logpost_beta(beta_prop, alpha) - logpost_beta(beta, alpha)
if (log(runif(1)) < log_r) beta <- beta_prop
samps[i,] <- c(alpha, beta)
}
mcmc(window(mcmc(samps), start=T/3)) # discard first half }
# Function to compute coverage & interval length per simulation
single_run <- function(alpha_true, beta_true, n=10) {
 rBS <- function(n, alpha, beta) {
Z <- rnorm(n)
Y <- (beta / 4) * (alpha * Z + sqrt((alpha * Z)^2 + 4))^2
return(Y) }
t <- rBS(n, alpha_true=0.5, beta_true=3)
post <- BS_chain(t, T=10000, sigma=0.35,
alpha0=alpha_true, beta0=beta_true)
ci_alpha <- HPDinterval(post[,"alpha"], prob=0.95)
ci_beta <- HPDinterval(post[,"beta"], prob=0.95)
cover_alpha<-(alpha_true >= ci_alpha[1,"lower"]&&
alpha_true<=ci_alpha[1,"upper"])
cover_beta <- (beta_true >= ci_beta[1,"lower"] &&
beta_true<=ci_beta[1,"upper"])
len_alpha <- ci_alpha[,"upper"] - ci_alpha[,"lower"]
len_beta <- ci_beta[,"upper"] - ci_beta[,"lower"]
return(c(cover_alpha, cover_beta, len_alpha, len_beta)) }
# Grid of true values
alphas_true <- c(0.1, 0.25, 0.5)
betas_true <- c(0.2, 1, 0.75)
set.seed(42)
reps <- 10000 # number of datasets per scenario
results <- expand.grid(alpha_t=alphas_true, beta_t=betas_true)
```

```
out <- data.frame(results, coverage_alpha=0, coverage_beta=0,
mean_len_alpha=0, mean_len_beta=0)
# Simulate for each scenario
for (i in seq_len(nrow(results))) {
a0 <- results$alpha_t[i]
b0 <- results$beta_t[i]
sim_res <- replicate(reps, single_run(a0, b0))
out$coverage_alpha[i] <- mean(sim_res[1,])
out$coverage_beta[i] <- mean(sim_res[2,])
out$mean_len_alpha[i] <- mean(sim_res[3,])
out$mean_len_beta[i] <- mean(sim_res[4,]) }
print(out)
Note that, when using  the second prior you change,
u to be  u <- rgamma(1, shape= a_1+(n+1)/2, rate=b_1+S/2) and
val<- -(n  + 1) * log(b)  +     sum(log((b / t)^0.5 + (b / t)^1.5)) -
  (1 / (2 * alpha^2)) * sum(t / b + b / t)  in the code above.
```

C.3 Birnbaum-Saunders Code: Simulation using Objective Priors

```
## Birnbaum-Saunders + reference prior π(α,β) ∝ 1/(αβ)
# Bivariate RW-MH in (η, ζ) = (log α, log β)
# Normal RW proposals only (no t, no mixtures)
# Plots shown separately #
set.seed(1)
suppressPackageStartupMessages({
 if (!requireNamespace("coda", quietly = TRUE))
stop("install.packages('coda')") }) library(coda)
# ---------- Stable helpers ----------
rbs <- function(n, alpha, beta) {
z <- rnorm(n); term <- (alpha * z) / 2
  beta * (term + sqrt(term^2 + 1))^2
}
# log f = log1p(x/β) - log(2 α √(2π))
- 1.5 log x - (x/β + β/x - 2)/(2 α^2)
bs_loglik_stable <- function(x, alpha, beta) {
```

```
if (!is.finite(alpha) || !is.finite(beta) || alpha <= 0 || beta <= 0 ||
any(x <= 0)) return(-Inf)   cst <- -log(2 * alpha * sqrt(2*pi))
xb  <- x / beta
 val <- log1p(xb) + cst - 1.5 * log(x) - (xb + beta/x - 2) / (2 * alpha^2)
s <- sum(val); if (!is.finite(s)) -Inf else s }
# Reference prior ⇒ Jacobian cancels in (η,ζ)
log_target_eta_zeta <- function(eta, zeta, x) {
  a <- exp(eta); b <- exp(zeta)
bs_loglik_stable(x, a, b)
}
# ---------- RW-MH (Normal proposals only)
rw_mh_bs_ref <- function(x, steps, s_eta, s_zeta, init_alpha,
init_beta, seed) {   set.seed(seed)
eta  <- log(init_alpha); zeta <- log(init_beta)
lt   <- log_target_eta_zeta(eta, zeta, x)
if (!is.finite(lt)) stop("Initial state invalid;
try different init_alpha/init_beta.")
a_tr <- numeric(steps); b_tr <- numeric(steps); acc <- 0L
 for (t in seq_len(steps)) {
    eta_p  <- eta  + rnorm(1, sd = s_eta)   # Normal RW in
  zeta_p <- zeta + rnorm(1, sd = s_zeta)  # Normal RW in ζ
        lt_p <- log_target_eta_zeta(eta_p, zeta_p, x);
if (!is.finite(lt_p)) lt_p <- -Inf     if (log(runif(1)) < (lt_p - lt)) {
eta <- eta_p; zeta <- zeta_p; lt <- lt_p; acc <- acc + 1L }
 a_tr[t] <- exp(eta); b_tr[t] <- exp(zeta)
}
list(alpha = a_tr, beta = b_tr, acc_rate = acc/steps)
}
running_mean <- function(v) cumsum(v) / seq_along(v)
# optional helper: try to open a new device (ignored if not available)
open_plot_device <- function() { try(grDevices::dev.new(),
silent = TRUE); invisible(NULL) }
# ---------- Driver for a given n ----------
# Plots are produced one-by-one (separately).
run_case <- function(n,
steps,
burn_frac,
```

```
s_eta_base,
 s_eta_multipliers = c(1, 1.2, 1.4, 1.7, 2.0, 2.3),
                   s_zeta = 0.25,
                inits = list(c(0.3,0.3), c(0.8,0.4), c(1.5,0.5),
c(3.0,0.8), c(20,2.0), c(40,0.02)),
seed_base = 1000 + n) {
# Data (replace with your own x if desired)
x <- rbs(n, alpha = 0.6, beta = 10.0)
# Build 6 chains with ONLY Normal proposals, but different η scales
 stopifnot(length(inits) == length(s_eta_multipliers))
chains <- vector("list", length(inits))
for (j in seq_along(inits)) {
chains[[j]] <- rw_mh_bs_ref(
   x, steps = steps,
    s_eta  = s_eta_base * s_eta_multipliers[j],
    s_zeta = s_zeta,
init_alpha = inits[[j]][1],
  init_beta  = inits[[j]][2],
 seed = seed_base + j     )
 }
 cat("\n=============== n =", n, "===============\n")
cat("Acceptance rates (chains 1..6):\n",
     paste(sprintf("  c%-2d: %.3f", 1:6, sapply(chains,
‘[[‘, "acc_rate")), collapse = "\n"),       "\n")
# Diagnostics (interpret cautiously for an improper target)
mcmc_list <- mcmc.list(lapply(chains, function(ch)
mcmc(cbind(
   logalpha = log(ch$alpha), logbeta = log(ch$beta)
 ))))
cat("\nGelman-Rubin PSRF (full run):\n")
print(round(suppressWarnings(gelman.diag(mcmc_list,
autoburnin = FALSE, multivariate = FALSE)$psrf), 3))
 cat("\nGeweke z (per chain):\n"); print(lapply(mcmc_list,
geweke.diag))   cat("\nEffective sample size (ESS, full run):\n");
print(round(effectiveSize(mcmc_list), 1))
 # ======= Separate plots (no par(mfrow)) =======
 # A) log(alpha) traces (all chains)
```

```
  open_plot_device()
 matplot(sapply(chains, function(ch) log(ch$alpha)), type = "l",
lty = 1,
col = 1:6,
xlab = "Iteration", ylab = "log(alpha)",
  main = bquote(paste("Traces: log ", alpha, " (n=", .(n), ")")))
legend("topleft", legend = paste0("c", 1:6), col = 1:6,
lty = 1, bty = "n")
# B) log(beta) traces (all chains)
open_plot_device()
 matplot(sapply(chains, function(ch) log(ch$beta)),
type = "l", lty = 1, col = 1:6,
 xlab = "Iteration", ylab = "log(beta)",
   main = bquote(paste("Traces: log ",
beta, " (n=", .(n), ")")))
legend("topleft", legend = paste0("c", 1:6), col = 1:6,
lty = 1, bty = "n")
# C) Sequential R-hat for log(alpha) over growing prefixes
burn    <- floor(steps * burn_frac)
seq_pts <- seq(max(burn + 10000, 50000), steps,
by = max(50000, steps %/% 20))
rhat_seq <- sapply(seq_pts, function(k) {
mlk <- mcmc.list(lapply(mcmc_list, function(mc) window(mc,
end = k)))
 suppressWarnings(gelman.diag(mlk, autoburnin = FALSE,
multivariate = FALSE)$psrf["logalpha", "Point est."])
})
open_plot_device()
plot(seq_pts, rhat_seq, type = "b",
 xlab = "Iterations used",
ylab = expression(R[hat]~"(log "*alpha*")"),
main = bquote(paste("Sequential ", R[hat], " (log ", alpha, "), n=",
.(n))))
abline(h = 1.01, lty = 2, col = "gray40"); abline(h = 1.05, lty = 2,
col = "gray40")
invisible(list(chains = chains, mcmc_list = mcmc_list,
seq_points = seq_pts, rhat_seq = rhat_seq)
```

```
) }
# =================== Run scenarios (examples) ===================
# Tune steps and step sizes as you like; below are just example settings.
# n = 10
out10 <- run_case(
 n = 10,
steps = 3000000,
 burn_frac = 0.10,
  s_eta_base = 50,
s_eta_multipliers = c(1.0, 1.4, 1.8, 2.2, 2.6, 7.0),
s_zeta = 0.1
)
# n = 40
out40 <- run_case(
n = 40,
steps = 3000000,
burn_frac = 0.10,
s_eta_base = 20,
s_eta_multipliers = c(1.0, 1.3, 1.6, 1.9, 2.2, 8.6),
s_zeta = 0.20
)
# n = 200
out200 <- run_case(
  n = 200,
steps = 1000000,
burn_frac = 0.06,
  s_eta_base = 3.5,
s_eta_multipliers = c(1.0, 1.25, 1.5, 1.8, 4.1, 7.0),
  s_zeta = 0.10 )
```

C.4 OpenBUGS Code for Weibull Distribution

```
model {
for (i in 1:N) {
dweib(shape, lambda)
data[i] ~ dweib(alpha, lambda)   }
  # Relationship between BUGS lambda and Weibull scale parameter beta
```

```
lambda <- pow(1 / beta, alpha)
  # ---- Inverse-Gamma priors implemented via reciprocals
  # alpha ~ Inv-Gamma(a_shape, b_shape)
 inv_alpha ~ dgamma(0.0001, 0.001)
alpha<- 1 / inv_alpha
  # beta ~ Inv-Gamma(0.0001, 0.0001)
inv_beta ~ dgamma(0.0001, 0.0001)
beta <- 1 / inv_beta
  }
list(N=10, data=c(152.7, 172.0, 172.5, 173.3, 193.0, 204.7,
216.5, 234.9, 262.6, 422.6))
```

C.5 OpenBUGS Code for Birnbaum-Saunders Distribution

```
model {
for (i in 1:N) {
 data[i] ~ dbs( alpha,beta) }
   inv_alpha ~ dgamma(0.0001,0.0001 )
inv_beta  ~ dgamma(0.0001,0.0001 )
alpha <- 1 / inv_alpha
beta  <- 1 / inv_beta
 }
 list(N=10, data=c(152.7, 172.0, 172.5, 173.3, 193.0,
204.7, 216.5, 234.9, 262.6, 422.6))
```

Appendix D: Research Ethics Declaration

1

RHODES UNIVERSITY
Where leaders learn

Department of Statistics

RESEARCH ETHICS DECLARATION

This statement is to be included in the appendices of research papers/dissertations/theses submitted for postgraduate examination where research did **not** involve interaction with human participants, or the use of animal subjects, and therefore did not require research ethics approval. Candidates whose research did require ethics clearance must include their ethics approval letter in the appendix of their examination submission.

Name of Candidate: Sinoyolo Ntanjana
Name of Supervisor(s): Dr Sharkay Izally & Prof Lizanne Raubenheimer
Degree: MSc
Title of research: A comparison of life distributions in Bayesian reliability theory

DECLARATION

I declare that my research did not require ethical clearance because (tick all that apply):

I used previously collected data that had already received ethics clearance.	
I analysed documents/open-access digital texts that are freely available in the public domain.	✓
I did a literature review/analysis of theoretical or secondary material only.	
I used human datasets of non-sensitive information that are either anonymous (identifiers were never collected) or have been deidentified (identifiers have been completely removed).	
I used commercially produced human biological material (e.g. established human cell lines).	
I observed people in public spaces and natural environments where they had no reasonable expectation of privacy and I did not interact with them or intervene in any way.	
I used non-living animal materials (eg bones of already deceased organisms or fossils) while complying with any custody and/or jurisdiction requirements.	
I did a content analysis of public media (newspapers, advertisements, and social media posts).	
I did a simulation study with no real-world consequences and does not involve disturbing or distressing content.	✓
I observed flora, fauna, and ecosystems without interfering with or disturbing their natural state while complying with any jurisdiction requirements.	
Other (Please provide details):	

Signature of Candidate: Date: 07/11/2025

Signature of Supervisor: Date: 07/11/2025