# RELIABILITY ANALYSIS: ASSESSMENT OF HARDWARE AND HUMAN RELIABILITY

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

Master of Science

in

Mathematical Statistics

Department of Statistics

Rhodes University

by

Masakheke Mafu

(Student number: 611M7251)

March 2017

Supervisor: Dr. L. Raubenheimer

# Declaration

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

Masakheke Mafu | 14/12/2016

M. Mafu | Date

# Abstract

Most reliability analyses involve the analysis of binary data. Practitioners in the field of reliability place great emphasis on analysing the time periods over which items or systems function (failure time analyses), which make use of different statistical models. This study intends to introduce, review and investigate four statistical models for modeling failure times of non-repairable items, and to utilise a Bayesian methodology to achieve this. The exponential, Rayleigh, gamma and Weibull distributions will be considered. The performance of the two non-informative priors will be investigated. An application of two failure time distributions will be carried out. To meet these objectives, the failure rate and the reliability functions of failure time distributions are calculated. Two non-informative priors, the Jeffreys prior and the general divergence prior, and the corresponding posteriors are derived for each distribution. Simulation studies for each distribution are carried out, where the coverage rates and credible intervals lengths are calculated and the results of these are discussed. The gamma distribution and the Weibull distribution are applied to failure time data.

The Jeffreys prior is found to have better coverage rate than the general divergence prior. The general divergence shows undercoverage when used with the Rayleigh distribution. The Jeffreys prior produces coverage rates that are conservative when used with the exponential distribution. These priors give, on average, the same average interval lengths and increase as the value of the parameter increases. Both priors perform similar when used with the gamma distribution and the Weibull distribution.

A thorough discussion and review of human reliability analysis (HRA) techniques will be considered. Twenty human reliability analysis (HRA) techniques are discussed; providing a background, description and advantages and disadvantages for each. Case studies in the nuclear industry, railway industry, and aviation industry are presented to show the importance and applications of HRA.

Human error has been shown to be the major contributor to system failure.

***Keywords:*** Coverage; Credible interval; General divergence prior; Hardware reliability, Human reliability, Jeffreys prior, Metropolis-Hastings.

# Contents

# List of Figures

# List of Tables

# Acknowledgements

Thank you to my supervisor, Doctor Lizanne Raubenheimer for her continuous guidance in my research. I am very grateful for your kindness and support.

Thank you to my family for all the support and for believing in me. Another thank you goes to the Department of Statistics at Rhodes University for being my second home.

The financial assistance of the National Research Foundation (NRF) is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

# List of Abbreviations

| | |
|---|---|
| AEMA | Action Error Mode Analysis |
| AIC | Akaike Information Criterion |
| APJ | Absolute Probability Judgement |
| ATHEANA | A Technique for Human Error Analysis |
| CAHR | Connectionism Assessment of Human Reliability |
| CCR | Central Control Room |
| CDF | Cumulative Distribution Function |
| CESA | Commission Errors Search and Assessment |
| CL | Confidence Level |
| COCOM | Contextual Control Model |
| CODA | Conclusions from Occurrences by Descriptions of Actions |
| CPC | Common Performance Conditions |
| CP | Coverage Probability |
| CREAM | Cognitive Reliability and Error Analysis Method |
| DIC | Deviance Information Criterion |
| DNE | Direct Numerical Estimation |
| EPC | Error Producing Conditions |
| ERA | Error Reduction Analysis |
| ERM | Error Reduction Measure |

| | |
|---|---|
| FTA | Fault Tree Analysis |
| HAZOP | Hazard and Operability |
| HCR | Human Cognitive Reliability |
| HEART | Human Error Assessment and Reduction technique |
| HEAT | Human Error Action Taxonomy (concept in CREAM) |
| HEP | Human Error Probability |
| HFE | Human Failure Event |
| HRA | Human Reliability Analysis |
| HRAET | Human Reliability Analysis Event Tree |
| HRMS | Human Reliability Management System |
| IDA | Influence Diagrams Approach |
| JHEDI | Justified Human Error Data Information |
| MEMORS | Method d'Evaluation de la Realisation des Mission Operateur pour la Surete |
| MRL | Mean Residual Life |
| MTTF | Mean Time To Failure |
| NARA | Nuclear Action Reliability Assessment |
| NASA | National Aeronautics and Space Administration |
| PAF | Performance Affecting Factors |
| PC | Paired Comparisons |
| PDF | Probability Density Function |
| PIF | Performance Influencing Factors |
| PSA | Probabilistic Safety Assessment |
| PSF | Performance Shaping Factor |
| PHEA | Predictive Human Error Analysis |

| | |
|---|---|
| PRA | Probability Risk Assessment |
| RCM | Reliability Centered Maintenance |
| SARAH | System Approach to the Reliability Assessment of Humans |
| SHARP | Systematic Human Action Reliability Procedure |
| SLI | Success Likelihood Index |
| SLIM | Success-Likelihood Index Methodology |
| SPAR-H | Simplified Plant Analysis Risk Human Reliability Assessment |
| SME | Subject Matter Expert |
| TESEO | Technica Empirica Stima Errori Operatori |
| THERP | Technique for Human Error Rate Prediction |
| THORP | Thermal Oxide Reprocessing Plant |
| TQM | Total Quality Management |
| TRC | Time Reliability Correlation |

# Notation

| | |
|---|---|
| $t$ | observation from random variable $T$ |
| $f(t)$ | pdf of random variable $t$ |
| $F(t)$ | cdf of random variable $T$ |
| $R(t)$ | reliability function |
| $z(t)$ | failure rate function |
| $p(\theta)$ | prior distribution of $\theta$ |
| $p(\theta \mid \underline{t})$ | posterior distribution of $\theta$ |
| $L(\theta \mid \underline{t})$ | likelihood function |
| $I(\theta)$ | Fisher information |
| $E[T]$ | Expected value of $T$ |
| $var(T)$ | Variance of $T$ |
| $LF_I$ | Likelihood function for type I censoring |
| $LF_{II}$ | Likelihood function for type II censoring |

# Chapter 1

# Introduction

## 1.1 Overview

There is a great deal of overlap between reliability analysis and risk analysis. Reliability analysis tends to involve the study of components or systems that are exposed to non-repeated and repeated failures and the interaction of humans with these systems, as well as devising maintenance policies to solve the outcomes of such studies, French et al. (2009). Risk analysis, on the other hand is concerned with the study of once-off failures that can cause an enormous damage to the system. It is apparent that both analyses deal with predicting possible failures and assessing their likelihood.

Many reliability analyses are often concerned with the analysis of success/ failure data, that is, binary outcomes. However, in real life studies, the importance of analyzing the time periods over which items or systems function is often emphasized. These analyses are known as lifetime or failure time analyses, and involve the analysis of positive, continuous-valued quantities. Thus, these lifetime analysis require the use of different statistical models than analyses based on success/ failure data. Hamada et al. (2008) gives one of the advantages of using statistical models to carry out the reliability analysis of an item/ system, that a large portion of the information in experimental data may be lost when success/ failure analysis is used. For example, failures at 1 and 99 hours are regarded as equivalent if a component must operate for 100 hours.

On the other hand, some useful reliability analyses involve the study of methodologies for predicting and assessing the consequences of failures as a result of human error. These failures typically arise when humans operate within the system. According to French et al. (2009), human error is the major contributor to the risks and reliability of many system; with over 90% of reported cases in the nuclear industry, and over 80% in the chemical and petro-chemical industries. Therefore, any study of system risk or reliability also requires analysis of the potential for failure arising from human activities in operating and managing this.

## 1.2 Objectives

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

- Discuss the quantitative measures for the reliability of a non-repairable item and various types of censored data, together with the likelihood construction.
- Provide a thorough discussion of the Human Reliability Analysis (HRA) methodology. This includes background, advantages and disadvantages of the HRA methods.
- Discuss the case studies of the various industries where HRA is needed and applied. These involve the Nuclear, Railway, and aviation industries.
- Introduce various statistical models for modeling failure times of items. This includes a thorough discussion of the quantitative measures of the models.
- Develop Bayesian methodology for modeling the failure times of items. This includes a discussion of prior, posterior distributions and predictive reliability of the various statistical models.
- Perform simulation studies for the Bayesian methodology. This includes comparing the performance of the prior distributions and their predictive power.
- Provide applications of some of the statistical models to real life data sets. The performance of the different priors will be compared.
- Recommend a Bayesian approach to modeling failure time data.

## 1.3 Contributions

The contribution of the thesis can be summarised as follows:

- In this thesis general divergence prior has been derived for the four parametric models under study.
- The propriety of the posterior distributions that arise in this study, from using the general divergence prior, has been shown.
- Simulation studies, and applications using real failure time data, for the various models have been carried out to evaluate the performance of the general divergence prior on the four lifetime parametric models.

## 1.4 Outline

In **Chapter 2** a background to hardware reliability and human reliability is provided. In this chapter, quantitative measures of the reliability of a non-repairable item are given, together with the types of censored data. Furthermore, human error and the factors affecting human performance are discussed. In **Chapter 3** a review of four statistical parametric models for failure time data is discussed. These models include the exponential, Rayleigh, gamma and Weibull distributions. Bayesian statistics and its tools for modeling lifetime data is also explored. This includes the prior and posterior distributions and the predictive posteriors

**Chapter 4** provides simulation studies for the four parametric distributions and also gives a discussion of the Markov Chain Monte Carlo methods, Metropolis-Hastings algorithm, in particular. The performance of the two priors is discussed. In **Chapter 5** applications using the gamma and the Weibull distributions to insulating fluid data are shown. In this chapter, the performance and predictive reliability for the priors are compared. In **Chapter 7** a review HRA and of the many HRA methods/ techniques are provided. This chapter gives the background, description, and advantages and disadvantages of the various techniques.

**Chapter 6** provides a number of case studies from the industries where HRA is mostly needed and used. In this chapter, the importance of HRA in nuclear industry, railway industry, and the aviation industry is explored, including certain aspects of the HRA process. Concluding remarks and avenues for future research will be discussed in **Chapter 8**.

**Appendix A** gives the derivations of the Fisher information for the parametric models discussed. **Appendix B** contains Matlab® code of the simulation studies chapter and R® and OpenBugs® code of the application chapter.

# Chapter 2

# Literature Review

This chapter looks at the review of literature concerning a broad topic of reliability. There are three main branches of reliability found in literature: hardware reliability, software reliability, and human reliability.

## 2.1 Introduction to Hardware Reliability

This section focuses on the first of the three branches - the reliability of technical components and systems. Until the 1960s reliability was defined as “the probability that an item will perform a required function under stated conditions for a stated period of time”. There is nothing wrong with this definition; as a matter of fact some authors still prefer this definition, for example, Birolini (2007). Birolini (2007) further states that reliability can be viewed in two ways: qualitatively and quantitatively. From a qualitative point of view, reliability can be defined as the ability of the item to remain functional. Quantitatively, reliability specifies the probability that no operational interruptions will occur during a stated time interval. We will, however, in this section adopt the more general definition of reliability given by Hoyland & Rausand (2009).

Hoyland & Rausand (2009) defines reliability as the ability of an item to perform a required function, under given environmental and operational conditions and for a stated period of time. A close inspection of this definition reveals that the term ’item’ is used to denote any component, subsystem, or system that can be considered as an entity. Furthermore, a required function may be a single function or a combination of functions that is necessary to provide a specified service. Therefore, for a hardware item to be reliable, it must operate satisfactorily for a specified period of time in the actual application for which it is intended. Reliability, for technical systems, emerged with a technological meaning just after World War I and was then used in connection with comparing operational safety of one-, two-, and four-engine airplanes. The reliability was measured as the number of accidents per hour of flight time, Hoyland & Rausand (2009).

After World War II, the development of reliability continued throughout the world as more complicated products were produced, and the need for complicated and safety systems also became steadily more pressing. Between the 1950s and 1960s, competition between Russia and the United States of America (USA) to be the first nation to put men on the moon, led to an establishment of an association for engineers working with reliability questions. As a result, the first journal on the subject, IEEE Transactions on reliability was established in 1963, and a number of textbooks on the subject were published in the 1960s. In the 1970s interest grew, in the USA as well as in other parts of the world, in risk and safety aspects that involved the building and operation of nuclear power plants. In the majority of industries a lot of effort and emphasis is presently put on the analysis of risk and reliability problems.

In order to learn more about the importance of hardware reliability and why it is pursued here, we take a look at its applications. It is worth mentioning that the main objective of a reliability study should always be to provide information as a basis for decisions. Reliability technology has a wide range of application areas. Some of these areas are discussed below, and are from Hoyland & Rausand (2009).

1. **Risk analysis**

    - Identification and description of potential accidental events in the system. An accidental event is usually defined as a significant deviation from normal operating conditions that may lead to unwanted consequences. In oil/ gas processing plant a gas leak may, for example, be defined as an accidental event.
    - The potential causes of each accidental event are identified by a causal analysis. The causes are usually identified in a hierarchical structure starting with the main causes. Main causes, and sub-causes may be described by a tree structure called a fault tree.

2. **Environmental protection**

    - Reliability studies may be used to improve the design and operational regularity of antipollution systems like gas/ water cleaning systems.
    - Many industries have realized that the majority of the pollution from their plants is caused by production irregularities and that consequently the production regularity of the plant is the most important factor in order to reduce pollution. Reliability and regularity studies are among the most important tools to optimize production regularity.
    - An environmental risk analysis is carried out according to the same procedure as a standard risk analysis and has the same interface with reliability analysis.

3. **Quality**

- Quality management and assurance is increasingly focused, stimulated by the almost compulsory application of the ISO 9000 series of standards.
- The concepts of quality and reliability are closely connected. Reliability may in some respects be considered to be a quality characteristic. Complementary systems are therefore being developed and implemented for reliability management and assurance as part of a total quality management (TQM) system.

4. **Optimization of maintenance and operation**

   - Maintenance is carried out to prevent failures and to restore the system function when a failure has occurred. The prime objective of maintenance is thus to maintain or improve the system reliability and production/ operation regularity.
   - Many industries (e.g., the nuclear, aviation, defense, and the offshore industry) have fully realised the important connection between maintenance and reliability and have implemented the reliability centered maintenance (RCM) approach. The RCM approach is a main tool to improve the cost-effectiveness and control of maintenance in all types of industries, and hence to improve availability and safety.
   - Reliability assessment is also an important element of the following applications: life cycle cost, life cycle profit, logistic support, spare part allocation, and manning level analysis.

5. **Engineering design**

   - Reliability is considered to be one of the most important quality characteristics of technical products. Reliability assurance should therefore be an important topic during the engineering design process.
   - Many industries have realized this and integrated a reliability program in the design process. This is especially the case within nuclear power, the aviation, the aerospace, the automobile, and the offshore industries.

6. **Verification of quality/ reliability**

   - A number of official bodies require that the producer and/ or the user of technical systems are able to verify that their equipment satisfies specified requirements. Such requirement usually have a basis in safety and/ or environmental protection.
   - Reliability analyses and reliability demonstration testing are necessary tools in the verification process by the producers of equipment.

## 2.2 Quantitative measures for the reliability of a non-repairable item

Within hardware reliability we may use two different approaches:

- The physical approach;
- the actuarial approach.

The physical approach is mainly used for reliability analyses of structural elements, like beams and bridges. The approach is therefore often called structural reliability, Hoyland & Rausand (2009).

Further, in the actuarial approach, we describe all our information about the operating loads and the strength of the component in the probability distribution function $F(t)$ of the time to failure $T$. When several components are combined into a system, the analysis is called a system reliability analysis. This section will look at the actuarial approach and aims to present and discuss the terminology and quantitative measures used in hardware reliability studies.

We will do this by introducing four important measures for the reliability of a non-repairable item. These are:

- The reliability function $R(t)$;
- the failure rate function $z(t)$;
- the meantime to failure (MTTF);
- the mean residual life (MRL).

### 2.2.1 Lifetime/ Time to failure

Consider the length of life (lifetime) of a component/ system, which is the length of the time interval from the initial activation of the unit until its failure. Denote this lifetime (also called time to failure) by $T$. This variable $T$ is considered a positive random variable, since it is a lifetime and the length of life cannot be exactly predicted. The time to failure may be a discrete or continuous variable. It is noteworthy, however, that a discrete variable can be approximated by a continuous variable. We will assume here that the time to failure $T$ is continuously distributed with probability density function (PDF), $f(t)$. The cumulative (life) distribution function (CDF) of $T$, denoted by $F(t)$, is the probability that the lifetime does not exceed $t$, that is,

$$F(t) = P(T \leq t), \qquad 0 < t < \infty.$$

The CDF, $F(t)$, has the following properties:

- The function is non-decreasing, that is, if $t_1 < t_2$ then $F(t_1) \leq F(t_2)$;
- $\lim_{t \to \infty} F(t) = 1$;
- $\lim_{t \to 0} F(t) = 0$;
- $F(t)$ is right-continuous.

The probability density function of $T$, $f(t)$, is a non-negative function such that

$$F(t) = \int_0^t f(x)dx, \qquad 0 \leq t \leq \infty.$$

Mathematically, the PDF of $T$, $f(t)$, is defined as

$$\begin{aligned} f(t) &= \frac{d}{dt}F(t) \\ &= \lim_{\triangle t \to 0} \frac{F(t+\triangle t) - F(t)}{\triangle t} \\ &= \lim_{\triangle t \to 0} \frac{P(t < T < t+\triangle t)}{\triangle t}. \end{aligned}$$

This implies that when $\triangle t$ is small,

$$P(t < T \leq t + \triangle t) \approx f(t) \cdot \triangle t.$$

### 2.2.2 Reliability Function

The reliability function of a component or system having a life distribution $F(t)$ is defined as

$$R(t) = 1 - F(t) = P(T > t).$$

Thus $R(t)$ is the probability that the component or system does not fail in the time interval $(0, t]$, or in other words, the probability that the item survives the time interval $(0, t]$ and is still functioning at time $t$. Hence the reliability function $R(t)$ is also called the survival function in literature. The reliability function, $R(t)$, has the following properties

- The function is non-increasing, that is, if $t_1 < t_2$ then $R(t_1) \geq R(t_2)$;
- $\lim_{t \to \infty} R(t) = 0$;
- $\lim_{t \to 0} R(t) = 1$;
- $R(t)$ is right-continuous.

### 2.2.3 Hazard Rate/ Failure Rate Function

This is the instantaneous failure rate of an item which has survived $t$ units of time, that is, the probability that an item will fail in the time interval $(t, t+\triangle t]$ when we know that the item is functioning at time $t$, and the length of the interval $(t, t+\triangle t]$ approaches zero. Its mathematical definition is,

$$\begin{aligned} z(t) &= \lim_{\triangle t \to 0} \frac{P(t < T \leq t+\triangle t | T > t)}{\triangle t} \\ &= \lim_{\triangle t \to 0} \frac{F(t+\triangle t) - F(t)}{\triangle t} \frac{1}{R(t)} \\ &= \frac{f(t)}{R(t)}. \end{aligned} \tag{2.1}$$

This implies that when $\triangle t$ is small,

$$P(t < T \leq t+\triangle t | T > t) \approx z(t) \cdot \triangle t.$$

The hazard rate function $z(t)$ is non-negative, that is, $z(t) \geq 0$.

If we put a large number of identical items into operation at time $t = 0$, then $z(t) \cdot \triangle t$ will roughly represent the relative proportion of the items still functioning at time $t$, failing in $(t, t+\triangle t]$, Hoyland & Rausand (2009).

Now, since

$$\begin{aligned} f(t) &= \frac{d}{dt} F(t) \\ &= \frac{d}{dt}(1 - R(t)) \\ &= -R'(t) \end{aligned}$$

then from (2.1)

$$z(t) = \frac{-R'(t)}{R(t)} = -\frac{d}{dt} \ln R(t).$$

Since $R(0) = 1$, then

$$\int_0^t z(x)dx = -\ln R(t)$$

and

$$R(t) = \exp\left[-\int_0^t z(x)dx\right].$$

The function $H(t) = \int_0^t z(x)dx$ is called the cumulative hazard function. Thus, we see that the reliability function $R(t)$ and the distribution function $F(t)$ are **uniquely** determined by the failure rate function $z(t)$. It readily seen that the density function of $T$, $f(t)$, can be expressed by

$$f(t) = z(t) \cdot \exp(-\int_0^t z(x)dx) \quad \text{for } t > 0.$$

Figure 2.1 taken from http://www.weibull.com/hotwire/issue21/hottopics21.htm shows a bathtub curve of a failure rate.

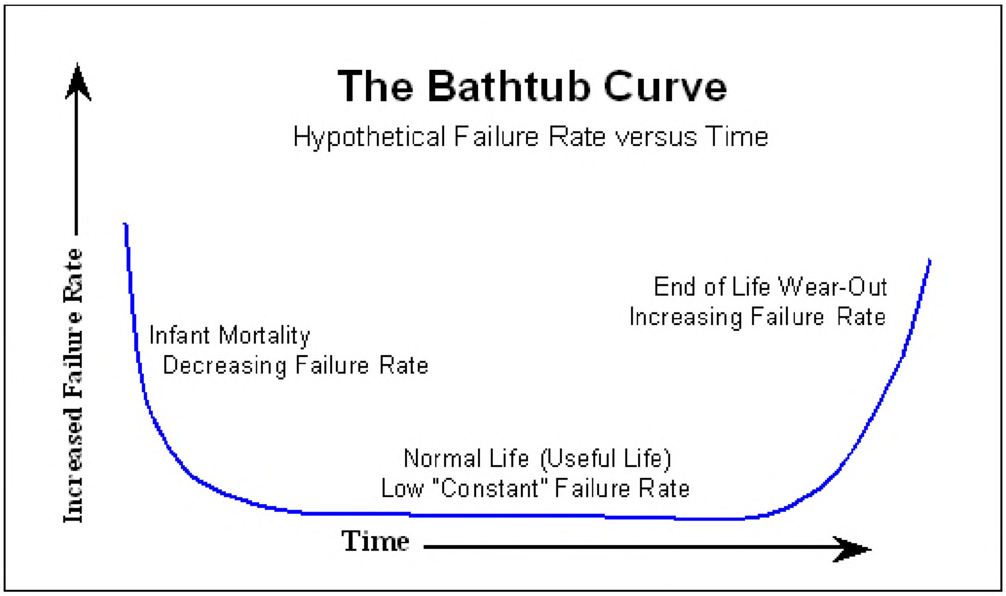


**Figure 2.1:** Bathtub example.

### 2.2.4 Measures of the center of a life distribution

#### 2.2.4.1 Mean Time To Failure (MTTF)

This is the average length of time until failure (the expected value of $T$). The general definition of the expected value of a lifetime random variable $T$ is

$$MTTF = E(T) = \int_0^\infty t f(t)dt.$$

Provided this integral is finite. Since $f(t) = R'(t)$,

$$MTTF = -\int_0^\infty tR'(t)dt.$$

By making use of integration by parts

$$MTTF = -[tR(t)]_0^\infty + \int_0^\infty R(t)dt.$$

If $MTTF < \infty$, then using the fact that $R(\infty) = 0$, it is easy to see that $[tR(t)]_0^\infty = 0$. In that case

$$MTTF = \int_0^\infty R(t)dt.$$

#### 2.2.4.2 Median Life

An alternative measure is the median life $t_m$, defined by

$$R(t_m) = 0.5.$$

The median divides the distribution in two halves. The item will fail time $t_m$ with 50% probability.

#### 2.2.4.3 Mode

The mode of a life distribution is the most likely failure time, that is, the time $t_m$ where the probability density function $f(t)$ attains its maximum:

$$f(t_{mode}) = \max_{0 \leq t < \infty} f(t).$$

### 2.2.5 Mean residual life (MRL)

This is expected remaining lifetime of a unit. It is plays a useful role in planning for maintenance and replacement, and in the context of health care, issues pertaining to the quality of life, Singpurwalla (2006). The mean residual life defined as

$$\begin{aligned} MRL &= \mu(t) = E[T - t \mid T > t] \\ &= \frac{\int_t^\infty (s-t) f(s) ds}{P(T > t)}. \end{aligned}$$

Integrating the numerator by parts gives

$$MRL = \mu(t) = \frac{-[(s-t)R(s)]_t^\infty + \int_t^\infty R(s)ds}{R(t)}$$

from which it can be seen that

$$\mu(t) = \frac{\int_t^\infty R(s)ds}{R(t)}.$$

Note that when $t = 0$ we obtain $\mu(0) = MTTF$, and

$$\mu'(t) = \frac{-R(t)R(t) + R'(t)\int_t^\infty R(s)ds}{(R(t))^2}$$

after some algebra the following is obtained

$$\mu'(t) = -1 + z(t)\mu(t).$$

Thus,

$$z(t) = \frac{1 + \mu'(t)}{\mu(t)}.$$

## 2.3 Complete and Censored data sets

This section will briefly look at the difference between complete data and censored data. Various types of censoring will be discussed. Throughout this section the censoring mechanism is assumed to satisfy the requirements of independent censoring. Furthermore, let $T_i$ denote the lifetime of item $i$ when the lifetime is considered as a random variable. The observed value of $T_i$ is denoted by $t_i$.

### 2.3.1 Complete Data Set

The data set is said to be complete when we are able to observe the real times to failure for all the $n$ items that are being studied. The data set takes the form $T_1, T_2..., T_n$, where $T_i$ denotes the time to failure of item $i$. The ordered data might be of interest, and one thus have to rearrange the data set in an increasing sequence:

$$T_{(1)} \leq T_{(2)} \leq \cdots \leq T_{(n)}$$

where $T_{(i)}$ is called the $i$th order statistic in the sample.

### 2.3.2 Censored data set

Often one has to be satisfied with incomplete data sets. This may be since it is impractical or too expensive to wait until all the items have failed, or because individual items are 'lost' for one reason or another, or because in recording lifetime we must make do with stating relatively large time intervals to which the lifetimes belong. In such situations the data is said to be censored, Hoyland & Rausand (2009). Thus, a censored data set will consist of items whose lifetimes are known to have occurred within certain intervals, and the remainder of the lifetimes will be known exactly.

#### 2.3.2.1 Type I Censoring

This type of censoring occurs when all items are activated at time $t = 0$ and followed until failure or until time $t_0$ when the experiment is terminated. This usually takes place in medical research. At the end of the experiment, only the lifetimes of those items that have failed before $t_0$ will be known exactly. The information in the data set obtained is of the form

$$T_{(1)} \leq T_{(2)} \leq \cdots \leq T_{(r)}$$

where $r \leq n$. Furthermore, we know that $n - r$ items have survived the time $t_0$, and this information should also be utilized. What is worth noting here is that the number of known lifetime(s) is stochastic, and that $t_0$, that denotes the time for the end of the experiment is fixed.

#### 2.3.2.2 Type II Censoring

In this type of censoring, the number of desired known lifetimes is fixed prior the experiment, that is, the design must allow for the experiment to terminate at the $r$th failure; $0 < r < n$. We still assume that the items are activated at time $t = 0$. As before, the information obtained through the test consist of the data set

$$T_{(1)} \leq T_{(2)} \leq \cdots \leq T_{(r)}$$

and $(n - r)$ items have survived the time $T_{(r)}$. Notice that here, we fixed the desired number of recorded lifetimes, but the time $T_{(r)}$is stochastic.

#### 2.3.2.3 Type III Censoring

This type of censoring is basically a combination of the first two types. In it the test terminates at the time that occurs first, that is, $t_0$ or the $r$th failure. This implies that both $t_0$ and $r$ must both be fixed beforehand.

#### 2.3.2.4 Type IV Censoring

In this type of censoring the $n$ numbered items are activated at different given point(s) in time. If it is found that the time for censoring of item $i$ is stochastic, the censoring is said to be type IV. We may set the activation at $t = 0$ for this censoring type.

### 2.3.3 Various Categories of Censoring

There are various categories of censoring that are worth mentioning, such as right censoring, left censoring, and interval censoring.

#### 2.3.3.1 Right Censoring

In this category, we know when an item was put into operation, but not necessarily when it fails. Here, the exact lifetime of the $i$th item will be known if, and only if, $T_i$ is less than of equal to $C_r$ (that is, $T_i \leq C_r$), where $C_r$ denotes 'right' censoring time. In the event that $T_i > C_r$, the item is said to be a survivor, and its event time is censored at $C_r$. The data from this experiment can be represented by pairs of random variables $(X, \delta)$, where $\delta$ indicates whether the lifetime $T_i$ corresponds to an event ($\delta = 1$) or is censored ($\delta = 0$), and $X$ is equal to $T_i$ if the lifetime is observed, and to $C_r$ if it is censored. Thus, $X = \min(T_i, C_r)$.

#### 2.3.3.2 Left Censoring

The data set is said to be left censored when we do not know when all the items were put into operation. In this category we know that an item is functioning when the observation period starts, but not necessarily how long the item has been functioning. In other words, a lifetime $T_i$ is considered to be left censored if it is less than a censoring time $C_l$(that is $T_i \leq C_l$), that is, the event of interest has already occurred for the individual before that person is observed in the study at time $C_l$. Here, $C_l$ denotes 'left' censoring time. The exact lifetime $T_i$ will be known if, and only if, $T_i$ is greater than or equal to $C_l$. The data from a test that involves left censoring can be represented by pairs of random variables $(X, \varepsilon)$, as in the previous subsection, where $X$ is equal to $T_i$ if the lifetime is observed and $\varepsilon$ indicates whether the exact is observed ($\varepsilon = 1$) or not ($\varepsilon = 0$). Hence, $X = \max(T_i, C_l)$.

#### 2.3.3.3 Interval Censoring

This is a more general category of censoring occurs when the lifetime is only known to occur within an interval. Such interval censoring occurs when patients in a clinical trial or longitudinal study have periodic follow-up and the patient's event time is only known to fall in an interval $(L_i, R_i]$ ($L$ for left endpoint and $R$ for right endpoint of the censoring interval), Klein & Moeschberger (2003).

#### 2.3.3.4 Truncation

This is another feature of many reliability and survival studies that is sometimes confused with censoring. As discussed in Klein & Moeschberger (2003), truncation occurs when only those items/individuals whose event time lies within a certain observational window $(Y_L, Y_R)$ are observed. An item whose event time is not in this interval is not observed and no information on this item is available to the investigator. This is in contrast to censoring where there is at least partial information on each item. Because we are only aware of items with event times in the observational window, the inference for truncated data is restricted to conditional estimation.

When $Y_R$ is infinite then we have left truncation. In this type of truncation we only observe those items whose event time $T_i$ exceeds the truncation time $Y_L$. That is we observe if $T_i$ if and only if$Y_L < T_i$. Right truncation occurs when $Y_L$is equal to zero. That is, we observe the lifetime $T_i$ only when $T_i \leq Y_R$.

### 2.3.4 Likelihood Construction for Censored Data

This section is concerned primarily with the likelihoods construction that is used for analyzing parametric models. They can also be used as a basis for determining the partial likelihoods used in the semi-parametric models regression methods. A critical assumption is that the lifetimes and censoring time are independent. If they are not independent, then specializes techniques must invoked. In constructing a likelihood function for censored data we nee to consider carefully what information gives us, Klein & Moeschberger (2003).

An observation corresponding to an exact event time provides information on the provides information on the probability that the event's occurring at this time, which is approximately equal to the density function of $T$ at this time. For a right censored observation all we know is that the event time is larger than this time, so the information is the reliability function evaluated at the on study time. Likewise, for a left censored observation, all we know is that the event has already occurred, so the contribution to the likelihood is the cumulative distribution function evaluated at the on study time. Lastly, for interval censored data we know only that the event occurred within the interval, so the information is the probability that the event is in this interval. It is worth mentioning that for truncated data, these probabilities are replaced by the appropriate conditional probabilities.

More precisely, the likelihoods for various types of censoring schemes may all be written by incorporating the following components:

| | | |
|---|---|---|
| exact lifetimes | - | $f(t)$ |
| right-censored items | - | $R(C_r)$ |
| left-censored items | - | $1-R(C_l)$ |
| interval-censored items | - | $[R(L)-R(R)]$ |
| left-truncated items | - | $f(t)/R(Y_L)$ |
| right-truncated items | - | $f(t)/[1-R(Y_R)]$ |
| interval-truncated items | - | $f(t)/[R(Y_L)-R(Y_R)]$ |

The likelihood function may be constructed by putting together the component parts as

$$LF \propto \prod_{i\in D} f(t_i) \prod_{i\in R} R(C_r) \prod_{i\in L} (1-R(C_l)) \prod_{i\in I} [R(L_i)-R(R_i)],$$

where $D$ is the set of death/ failure times, $R$ the set of right censored items, $L$ the set of left censored items, and $I$ the set of interval censored items. if each item has a different failure distribution, as might be the case when regression techniques are used, the likelihood becomes

$$LF \propto \prod_{i \in D} f_i(t_i) \prod_{i \in R} R_i(C_r) \prod_{i \in L} (1 - R_i(C_l)) \prod_{i \in I} [R_i(L_i) - R_i(R_i)]$$

In Klein & Moeschberger (2003) the details of constructing the likelihood function for type I censoring for the right censoring category are provided as follows. For $\delta = 0$, it can be seen that

$$\begin{aligned} P(X, \delta = 0) &= P(X = C_r \mid \delta = 0) P(\delta = 0) \\ &= P(\delta = 0) \\ &= P(T > C_r) = R(C_r) \end{aligned}$$

Also, for $\delta = 1$,

$$\begin{aligned} P(X, \delta = 1) &= P(X = T \mid \delta = 1) P(\delta = 1) \\ &= P(T = X \mid T \leq C_r) P(T \leq C_r) \\ &= \left[\frac{f(t)}{1 - R(C_r)}\right] [1 - R(C_r)] = f(t). \end{aligned}$$

We can combine these expressions into the single expression

$$P(t, \delta) = [f(t)]^{\delta} [R(t)]^{1-\delta}.$$

Thus if we have a random sample of pairs $(X_i, \delta_i)$, $i = 1, ..., n$, the likelihood function is

$$LF_I = \prod_{i=1}^{n} P(t_i, \delta_i) = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [R(t_i)]^{1-\delta_i}.$$

For type II censoring, the data consist of the $r$th smallest lifetimes $T_{(1)} \leq T_{(2)} \leq \cdots \leq T_{(r)}$ out of a random sample of $n$ lifetimes from the assumed distribution, the likelihood is given by

$$LF_{II} = \frac{n!}{(n-r)!} \left[\prod_{i=1}^{r} f(t_{(i)})\right] \left[R(t_{(r)})\right]^{n-r}.$$

## 2.4 Introduction to Human Reliability

There are a number of definitions of human reliability provided in the literature, we will only give two in this review. As defined by Swain & Guttmann (1983) human reliability is the probability that a person correctly performs some system-required activity in a required time period (if time is a limiting factor); and performs no extraneous activity that can degrade the system. Human reliability analysis (HRA) is a method by which human reliability is estimated. When one undertakes to perform human reliability analysis, it is important to determine specifically those human actions that have the potential to bring about effect on system reliability. Groth & Swiler (2013) define human reliability analysis as the aspect of probabilistic risk assessment that is concerned with systematically identifying and analyzing the cause and consequence of human errors. From the two definitions we see that human reliability analysis pertains to methodologies for estimating and analyzing the effect of those failures which involve human action or inaction, and says nothing about the failure of some physical component.

According to French et al. (2009) human reliability grew up in the 1960s with the intention of modeling the likelihood and consequences of human error. This means that it is a relative new field. It has its roots in the early probabilistic risk assessments performed as part of the United States of America nuclear energy development programme. At the time, it treated humans as any other component in the system, French et al. (2009). Over the years, methods of different propensities for assessing human reliability have been developed. Currently, as it is stated in Groth & Swiler (2013), there are over 50 HRA methods that can be used to estimate the probability of human error, and development of new HRA methods continues to be a topic of research. As a result, these techniques are categorized in the literature into first, second, and emerging third generation HRA methodologies, respectively. It is noted in French et al. (2009) that context and cognition are the two features that second generation methods are supposed to contain - yet context is certainly modeled to some extent in first generation methods. Third generation methods contain more dynamic simulation, and have to be implemented on a computer.

Since humans have the ability to initiate and reduce accident occurrence, it becomes important that the influence of humans on the entire reliability of the system be taken into consideration in any total probabilistic risk analysis. Few systems operate completely independent of humans. Thus any study of system risk or reliability requires analysis of the potential for failure arising from human activities, French et al. (2009). Risks occur throughout the life cycle: during construction or installation, during operation, during modification and during decommissioning. In the event of failure at any stage there are risks associated with repair and recovery. In all cases HRA will inform risk management. When we want to thoroughly assess the risks attributable to human error and to come up with strategies of reducing system vulnerability to human error impact, we learn to appreciate the importance of HRA. According to Kirwan (1994) this is accomplished by its three principal functions of identifying what errors can occur (Human error identification), deciding how likely the errors are to occur (Human error quantification), and by reducing this error likelihood (Human error reduction). Human reliability analysis can also enhance the profitability and availability of systems via human error reduction/avoidance, although the main drive for the development and application of HRA techniques has so far came from the risk assessment and reduction domain, Kirwan (1994). The following discussion of the principal functions of HRA is presented in Kirwan (1994).

### Human error identification

This step typically follows task analysis (which defines how the task should be carried out), as it considers what can go wrong. It looks (but not limited to) at the following error types:

- error of omission - failing to carry out a required activity;
- error of commission - failing to carry out a required act adequately: act performed without required precision, or with too much or too little force; act performed at wrong time; acts performed in wrong sequence;
- extraneous act - unrequired act performed instead of, or in addition to, required act;
- error-recovery opportunities - acts which can recover previous errors.

Many errors can be identified in this step, and only those that are important (ones which can contribute to a degraded system state) for the study are integrated into risk analysis.

### Human error quantification

Upon the representation of the human error potential, the next step is to quantify the likelihood of the errors involved and then determine the effect of human error on system safety or reliability. Human reliability methods all quantify the human error probability (HEP), which is the metric of human reliability assessment.

The human error probability is defined as follows:

$$HEP = \frac{Number\ of\ errors\ occured}{Number\ of\ opportunities\ for\ error}.$$

### Error reduction analysis

Error reduction measure may be derived in a number of ways: according to the identified root causes of the error; from the defined factors (called Performance Shaping Factors, to be discussed shortly) that contribute to the error's HEP; or else from an assessment of the task in its system context, using ergonomics/ engineering judgement to identify how to prevent the error, or how to reduce either its likelihood or its system impact. If error reduction is necessary to reduce risk to an acceptable level, then following such error reduction measures, the system risk level will need to be recalculated.

There are many reasons given in French et al. (2009) as to when and why one might undertake an assessment of human reliability:

- In the design of a system one may be concerned with 'designing out' the potential for system failure. Part of this involves analyzing how human behaviour may affect the system in its potential both to compromise its reliability and to avoid the threat of imminent failure.
- Sometimes an organization wants to restructure and change its reporting structures. In such circumstances, it may wish to understand how its organizational design may affect the reliability and safety of its systems; and in turn that understanding may inform the development of its safety culture.
- During licensing discussion between a government regulator and the system operator there may be a need to demonstrate that a system meets a safety target.
- There may be a need to modify a system in which case there are needs to design the modification and the project to deliver the modification.
- There may be a need to choose which of several potential systems to purchase and the risk of system failure may be a potential differentiator between the options.

Therefore there are many contexts and reasons for conducting an HRA. This also explains why there are so many HRA methodologies.

So far in this chapter we have learned, among many other things, that the objective of HRA is to analyse and help in the prevention of negative effects of human error system performance and safety.We now focus on the application of HRA. Human reliability analysis is usually applied in the context of the risk assessment of complex and potentially hazardous systems such as nuclear power plants, chemical plants, offshore installations, and mass transportation systems, Kirwan (1997). In order to see how HRA is applied in these industry sectors, the following discussion as noted in Kirwan (1994) is undertaken.

#### Nuclear power plant

Human reliability analysis is used when looking at distributed emergency operation, control room misdiagnosis, nuclear power scenarios, research reactor scenarios.

#### Chemical plants

This sector make use of HRA in the the determination of chlorine loading, chemical process control, chemical sampling, hazardous substances control.

**Offshore**

Here, HRA is used when looking at offshore drilling, offshore lifeboat evacuation, offshore platform blowout, offshore platform depression I and II.

**Marine transport**

This industry sector applies HRA in the examination of ship-platform collision.

## 2.5 Human error

Human error is defined as any member of a set of actions that exceeds some limit of acceptability. Thus, an error is merely an out-of-tolerance action, where the limits of tolerable performance are defined by the system. Furthermore, human errors include intentional errors and unintentional errors. The former occur when the operator intends to perform some act that is incorrect but believes it to be correct or to represent a superior method of performance, Swain & Guttmann (1983).

Human error in complex and potentially hazardous systems involves human action (or inaction) in unforgiving systems. For human error to have a negative system impact, there must first be an opportunity for an error, or a requirement for reliable human performance-often in response to some event. The error must then occur and fail to be corrected or compensated for by the system, and it must have negative consequences, Kirwan (1994).

## 2.6 Performance Shaping Factors

Performance shaping factors (PSFs) encompass those influence that enhance or degrade human performance. PSFs are used within human reliability analysis (HRA) methods to identify contributors to human errors and to provide a basis for quantifying those contributors systematically, Boring et al. (2007). In the literature there are many different terms given to these factors - PSF (performance shaping factors), EPC (error producing condition), PIF (performance influencing factors), CPC (common performance conditions), PAF (performance affecting factors), and so on and so forth.

In a discussion of PSFs in Swain & Guttmann (1983) it is noted that there are two kinds of PSFs that are normally considered in modeling human performance for PRA; the external PSFs that include the entire work environment, particularly the equipment design and written or oral instructions; and the internal PSFs that represent the individual characteristics of the person, his skills, motivations and expectations that influence his/her performance. One of the most influential PSFs is stress. According to Schuller (1997) performance shaping factors are hypothetical, since one does not know for certain that they will have a particular effect in a specific situation. Experimental work generally does not give the kind of data needed about performance shaping factors that is useful in human reliability quantification. The most important groups of performance shaping factors are:

- Operating environment
  - Physical work environment
  - Work pattern
- Task characteristics
  - Equipment design
  - Control panel design
  - Job aids and procedures
  - Training
- Operator characteristics
  - Experience
  - Personality factors
  - Physical condition and age
- Organizational and social factors
  - Team work and communications
  - Management policies.

Other kinds of PSFs are psychological and physiological stressors that result from a work environment in which the demands placed on the human by the system do not conform to his/her performance. Amongst others, some of the stressors are task speed, task load and fatigue, Zimolong (1992).

# Chapter 3

# Bayesian Inference for a Selection of Parametric Models for Hardware Reliability Analysis

## 3.1 Introduction

In this section a number of probability distributions that may be used to model the lifetime of a non-repairable item will be reviewed. These models differ in the number of parameters, which reflect shape, location, and scale, and for a particular application, provide a variety of hazard functions to choose from. Hence, it is appropriate and necessary to discuss these more widely used parametric models. These models are chosen, not only because of their popularity among researchers who analyze reliability data, but also because they offer insight into the nature of the various parameters and functions such as the hazard rate. The following important life distributions are discussed:

- The exponential distribution;
- the gamma distribution;
- the Rayleigh distribution;
- the Weibull distribution.

In the discussion, the following will be of interest:

- Density function;
- hazard function;
- expected lifetime (MTTF);
- reliability function.

## 3.2 Distributions

### 3.2.1 Exponential distribution

Consider an item that is put into operation at $t = 0$. The time to failure $T$ of the item is said to have an exponential distribution with parameter $\lambda$, written $T \sim exp(\lambda)$, if the density function is

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{for } t > 0,\, \lambda > 0 \\ 0 & \text{otherwise.} \end{cases}$$

The cumulative (life) distribution function of the item is

$$\begin{aligned} F(t) &= P(T \leq t) \\ &= \int_0^t \lambda e^{-\lambda u} du \\ &= 1 - e^{-\lambda t} \quad \text{for } t > 0. \end{aligned}$$

The reliability (survivor) function of the item is

$$\begin{aligned} R(t) &= 1 - F(t) \\ &= P(T > t) \\ &= \int_t^\infty \lambda e^{-\lambda u} du \\ &= e^{-\lambda t}. \end{aligned}$$

The mean time to failure is

$$\begin{aligned}
\text{MTTF} &= \mathrm{E}[T] \\
&= \int_0^\infty t\lambda e^{-\lambda t} dt \\
&= \int_0^\infty e^{-\lambda t} dt \\
&= \frac{1}{\lambda}
\end{aligned}$$

and the variance of $T$ is

$$\begin{aligned}
\mathrm{var}(T) &= \int_0^\infty t^2 \lambda e^{-\lambda t} dt - \left(\frac{1}{\lambda}\right)^2 \\
&= \frac{1}{\lambda^2}.
\end{aligned}$$

The mean residual life is

$$\begin{aligned}
\mathrm{MRL}(t) &= \mathrm{E}[T - t \mid T > t] \\
&= \frac{\int_t^\infty e^{-\lambda u} du}{e^{-\lambda t}} \\
&= \frac{1}{\lambda} \frac{e^{-\lambda t}}{e^{-\lambda t}} \\
&= \frac{1}{\lambda} = \text{MTTF}.
\end{aligned}$$

The failure rate function is

$$\begin{aligned}
z(t) &= \frac{f(t)}{R(t)} \\
&= \frac{\lambda e^{-\lambda t}}{e^{-\lambda t}} \\
&= \lambda = \text{constant},
\end{aligned}$$

showing that the failure rate of an item with an exponential life is independent of time. We also see that on average an item whose lifetime is exponentially distributed fails every $\frac{1}{\lambda}$ time units, and that this is the same as the item's expected remaining life.

When the data is complete the likelihood becomes

$$\begin{aligned} LF(\lambda \mid t_1, t_2, \ldots, t_n) &= \prod_{i=1}^{n} \lambda e^{-\lambda t_i} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i}. \end{aligned}$$

The likelihood function of the $n$ exponentially distributed censored lifetimes for type I censoring, when the data is incomplete, is given by

$$LF_I(\lambda) = \prod_{i\varepsilon D} \lambda e^{-\lambda t_i} \prod_{i\varepsilon R} e^{-\lambda C_r} \prod_{i\varepsilon L} (1 - e^{-\lambda C_r}) \prod_{i\varepsilon I} [e^{-\lambda L_i} - e^{-\lambda R_i}]$$

and for type II censoring

$$\begin{aligned} LF_{II}(\lambda) &= \frac{n!}{(n-r)!} \left[ \prod_{i=1}^{r} \lambda e^{-\lambda t_i} \right] \left[ e^{-\lambda t_r} \right]^{n-r} \\ &= \frac{n!}{(n-r)!} \lambda^r e^{-\lambda \left[ \sum_{i=1}^{r} t_i + (n-r) t_r \right]}. \end{aligned}$$

### 3.2.2 Gamma distribution

The lifetime $T$ of an item is said to be gamma distributed, written $T \sim gamma(\alpha, \lambda)$, with parameters $\lambda > 0$, and $\alpha > 0$, if the density function is

$$f(t) = \begin{cases} \dfrac{\lambda^{\alpha} t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha)} & \text{for } t > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ is the gamma function; $\lambda$ is the rate parameter, and $\alpha$ is the shape parameter. It is easy to see that when $\alpha = 1$, the gamma becomes an exponential distribution with rate parameter $\lambda$.

The mean time to failure is

$$
\begin{aligned}
&= E[T] \\
&= \int_0^\infty \frac{t\lambda^\alpha t^{\alpha-1} e^{-\lambda t} dt}{\Gamma(\alpha)} \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty t^\alpha e^{-\lambda t} dt \\
&= \frac{\alpha}{\lambda}
\end{aligned}
$$

and the variance of $T$ is

$$
\begin{aligned}
\text{var}(T) &= \int_0^\infty \frac{t^2\lambda^\alpha t^{\alpha-1} e^{-\lambda t} dt}{\Gamma(\alpha)} - \left(\frac{\alpha}{\lambda}\right)^2 \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}} - \left(\frac{\alpha}{\lambda}\right)^2 \\
&= \frac{\alpha}{\lambda^2}.
\end{aligned}
$$

The distribution function is

$$
\begin{aligned}
F(t) &= P(T \le t) \\
&= \int_0^t \frac{\lambda^\alpha s^{\alpha-1} e^{-\lambda s}}{\Gamma(\alpha)} ds \\
&= \frac{\lambda}{\Gamma(\alpha)} \int_0^t (\lambda s)^{\alpha-1} e^{-\lambda s} ds,
\end{aligned}
$$

and letting $\lambda s = x$, it is easy to see that

$$
\begin{aligned}
F(t) &= \frac{\int_0^{\lambda t} x^{\alpha-1} e^{-x} dx}{\Gamma(\alpha)} \\
&= \frac{\gamma(\alpha, \lambda t)}{\Gamma(\alpha)},
\end{aligned}
$$

where $\gamma(\alpha, \lambda t)$ is the lower incomplete gamma function. The reliability function is

$$\begin{aligned} R(t) &= P(T > t) \\ &= \int_t^\infty \frac{\lambda^\alpha s^{\alpha-1} e^{-\lambda s}}{\Gamma(\alpha)} ds \\ &= \frac{\int_{\lambda t}^\infty x^{\alpha-1} e^{-x} dx}{\Gamma(\alpha)} \\ &= \frac{\Gamma(\alpha, \lambda t)}{\Gamma(\alpha)}, \end{aligned}$$

where $\Gamma(\alpha, \lambda t)$ is the upper incomplete gamma function. The corresponding failure rate function is then given by

$$\begin{aligned} z(t) &= \frac{f(t)}{R(t)} \\ &= \frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t}}{\Gamma(\alpha, \lambda t)} \end{aligned}$$

When the shape parameter $\alpha$ takes only positive integers, the distribution function $F(t)$ can be written as

$$F(t) = 1 - \sum_{n=0}^{\alpha-1} \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

and the reliability function then becomes

$$R(t) = \sum_{n=0}^{\alpha-1} \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

The corresponding failure rate function is

$$z(t) = \frac{f(t)}{R(t)} = \frac{\lambda^\alpha t^{\alpha-1} e^{-\lambda t} / \Gamma(\alpha)}{\sum_{n=0}^{\alpha-1} \frac{(\lambda t)^n}{n!} e^{-\lambda t}}.$$

The likelihood function for the complete data from the gamma distribution is given by

$$\begin{aligned} L(\lambda \mid t_1, t_2, \ldots, t_n) &= \prod_{i=1}^n \frac{\lambda^\alpha t_i^{\alpha-1} e^{-\lambda t_i}}{\Gamma(\alpha)} \\ &= \frac{\lambda^{n\alpha}}{(\Gamma(\alpha))^n} \left[ \prod_{i=1}^n t_i^{\alpha-1} \right] e^{-\lambda \sum_{i=1}^n t_i}. \end{aligned}$$

When the data is incomplete from the gamma distribution, as a result of type I censoring, the likelihood becomes

$$LF_I(\alpha,\lambda)=\prod_{i\in D}\left[\frac{\lambda^{\alpha}t_i^{\alpha-1}e^{-\lambda t_i}}{\Gamma(\alpha)}\right]\prod_{i\in R}\left[\frac{\Gamma(\alpha,\lambda C_i)}{\Gamma(\alpha)}\right]\prod_{i\in L}\left[1-\frac{\Gamma(\alpha,\lambda C_i)}{\Gamma(\alpha)}\right]\prod_{i\in I}\left[\frac{\Gamma(\alpha,\lambda L_i)}{\Gamma(\alpha)}-\frac{\Gamma(\alpha,\lambda R_i)}{\Gamma(\alpha)}\right]$$

and for type II censoring the likelihood is

$$\begin{aligned} LF_{II}(\alpha,\lambda) &= \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r}\left(\frac{\lambda^{\alpha}t_{(i)}^{\alpha-1}e^{-\lambda t_{(i)}}}{\Gamma(\alpha)}\right)\right]\left[\frac{\Gamma(\alpha,\lambda t_{(i)})}{\Gamma(\alpha)}\right]^{n-r} \\ &= \frac{n!\lambda^{\alpha r}}{(n-r)!\,(\Gamma(\alpha))^{n}}\left[\prod_{i=1}^{r}t_{(i)}^{\alpha-1}e^{-\lambda t_{(i)}}\right]\left[\Gamma(\alpha,\lambda t_{(i)})\right]^{n-r}. \end{aligned}$$

### 3.2.3 Rayleigh distribution

This is a lifetime distribution named after Lord Rayleigh. It is a special case of the Weibull distribution, when the shape parameter, $\alpha$, is set equal to 2. It is widely used in communications theory, engineering, and in the physical sciences. The time to failure of an item is said to have a Rayleigh distribution with parameter $\lambda$ $(>0)$, written $T \sim Rayleigh(\lambda)$, if the probability density function is

$$f(t)=\begin{cases}2\lambda^2te^{-(\lambda t)^2} & \text{for } t>0\\ 0 & \text{otherwise .}\end{cases}$$

The corresponding distribution function is

$$F(t)=\begin{cases}1-e^{-(\lambda t)^2} & \text{for } t>0\\ 0 & \text{otherwise.}\end{cases}$$

The reliability function is given by

$$\begin{aligned} R(t) &= 1-F(t) \\ &= \Pr(T>t) \\ &= e^{-(\lambda t)^2}. \end{aligned}$$

The mean time to failure of the item is

$$\begin{aligned} \text{MTTF} &= \text{E}[T] \\ &= \frac{1}{\lambda}\Gamma\left(\frac{1}{2}+1\right) \end{aligned}$$

and the variance of $T$ is

$$\begin{aligned} \text{var}(T) &= \frac{1}{\lambda^2}\Gamma\left(\frac{2}{2}+1\right) - \left(\frac{1}{\lambda}\Gamma\left(\frac{1}{2}+1\right)\right)^2 \\ &= \frac{1}{\lambda^2} - \left(\frac{1}{\lambda}\Gamma\left(\frac{1}{2}+1\right)\right)^2. \end{aligned}$$

The failure rate function is

$$\begin{aligned} z(t) &= \frac{f(t)}{R(t)} \\ &= \frac{2\lambda^2 t e^{-(\lambda t)^2}}{e^{-(\lambda t)^2}} \\ &= 2\lambda^2 t. \end{aligned}$$

When the data is complete the likelihood becomes

$$\begin{aligned} L(\lambda \mid t_1, t_2, ..., t_n) &= \prod_{i=1}^{n} 2\lambda^2 t_i^{2-1} e^{-(\lambda t_i)^2} \\ &= 2^n \lambda^{2n} \left[\prod_{i=1}^{n} t_i\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^2}. \end{aligned}$$

The likelihood function of the $n$ Rayleigh distributed censored lifetimes for type I censoring, when the data is incomplete, is given by

$$LF_I(\lambda) = \prod_{i\in D}[2\lambda^2 t_i^{2-1} e^{-(\lambda t_i)^2}] \prod_{i\in R} e^{-(\lambda C_i)^2} \prod_{i\in L}(1 - e^{-(\lambda C_i)^2}) \prod_{i\in I}[e^{-(\lambda L_i)^2} - e^{-(\lambda R_i)^2}]$$

and for type II censoring the likelihood is

$$
\begin{aligned}
LF_{II}(\lambda) &= \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} 2\lambda^2 t_i^{2-1} e^{-(\lambda t_i)^2}\right]\left[e^{-(\lambda t_r)^2}\right]^{n-r} \\
&\propto 2^r \lambda^{2r}\left[\prod_{i=1}^{r} t_i^{2-1}\right] e^{-\sum_{i=1}^{r}(\lambda t_i)^2-(n-r)(\lambda t_r)^2}.
\end{aligned}
$$

### 3.2.4 Weibull distribution

This is one of the most widely used life distributions in reliability analysis. It is named after professor Waloddi Weibull (1887-1979) who developed the distribution for modeling the strength of materials. It is very flexible, and can model many types of failure rate behaviors. The lifetime $T$ of an item is said to be Weibull distributed, written $T \sim Weibull(\alpha, \lambda)$, with positive parameters $\alpha$ and $\lambda$ if the probability density function i given by

$$
f(t) = \begin{cases} \alpha\lambda(\lambda t)^{\alpha-1} e^{-(\lambda t)^{\alpha}} & \text{for } t > 0 \\ 0 & \text{otherwise .} \end{cases}
$$

The corresponding distribution function is

$$
F(t) = \begin{cases} 1 - e^{-(\lambda t)^{\alpha}} & \text{for } t > 0 \\ 0 & \text{otherwise,} \end{cases}
$$

where $\lambda$ is a scale parameter, and $\alpha$ is called the the shape parameters. It is easy to see that when $\alpha = 1$, the Weibull distribution is equal to the exponential distribution.

The reliability function is given by

$$
\begin{aligned}
R(t) &= 1 - F(t) \\
&= \Pr(T > t) \\
&= e^{-(\lambda t)^{\alpha}} \text{ for } t > 0.
\end{aligned}
$$

The mean time to failure of the MTTF is

$$\begin{aligned}
\text{MTTF} &= \mathrm{E}[T] \\
&= \int_0^\infty t\alpha\lambda\,(\lambda t)^{\alpha-1} e^{-(\lambda t)^\alpha} dt \\
&= \alpha\lambda^\alpha \int_0^\infty t^\alpha e^{-(\lambda t)^\alpha} dt \\
&= \lambda^\alpha \int_0^\infty u^{\frac{1}{\alpha}} e^{-\lambda^\alpha u} du \\
&= \lambda^\alpha \frac{\Gamma\left(\frac{1}{\alpha}+1\right)}{(\lambda^\alpha)^{\frac{1}{\alpha}+1}} \\
&= \frac{1}{\lambda}\Gamma\left(\frac{1}{\alpha}+1\right)
\end{aligned}$$

where the the substitution $u = t^\alpha$ was used to simplify the integral. The variance of $T$ is

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

The failure rate function is given by

$$\begin{aligned}
z(t) &= \frac{f(t)}{R(t)} \\
&= \frac{\alpha\lambda\,(\lambda t)^{\alpha-1} e^{-(\lambda t)^\alpha}}{e^{-(\lambda t)^\alpha}} \\
&= \alpha\lambda\,(\lambda t)^{\alpha-1} \text{ for } t > 0.
\end{aligned}$$

Note that when $\alpha = 1$, the failure rate is independent of time (exponential distribution); when $\alpha > 1$, the failure rate function is increasing; and when $\alpha < 1$, $z(t)$ is decreasing.

When the data is complete the likelihood becomes

$$L(\alpha, \lambda \mid t_1, t_2, \ldots, t_n) = \prod_{i=1}^{n} \alpha\lambda^\alpha t_i^{\alpha-1} e^{-(\lambda t_i)^\alpha} = \alpha^n \lambda^{\alpha n} \left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^\alpha}.$$

The likelihood function of the $n$ Weibull distributed censored lifetimes for type I censoring, when the data is incomplete, is given by

$$LF_I(\alpha,\lambda) = \prod_{i\varepsilon D}\left[\alpha\lambda^{\alpha}t_i^{\alpha-1}e^{-(\lambda t_i)^{\alpha}}\right]\prod_{i\varepsilon R}e^{-(\lambda C_i)^{\alpha}}\prod_{i\varepsilon L}(1-e^{-(\lambda C_i)^{\alpha}})\prod_{i\varepsilon I}[e^{-(\lambda L_i)^{\alpha}}-e^{-(\lambda C_i)^{\alpha}}]$$

and for typeII censoring

$$\begin{aligned} LF_{II}(\alpha,\lambda) &= \tfrac{n!}{(n-r)!}\left[\prod_{i=1}^{r}\alpha\lambda^{\alpha}t_i^{\alpha-1}e^{-(\lambda t_i)^{\alpha}}\right]\left[e^{-(\lambda t_r)^{\alpha}}\right]^{n-r} \\ &\propto \alpha^r\lambda^{\alpha r}\left[\prod_{i=1}^{r}t_i^{\alpha-1}\right]e^{-\sum_{i=1}^{r}(\lambda t_i)^{\alpha}-(n-r)(\lambda t_r)^{\alpha}}. \end{aligned}$$

## 3.3 Bayesian Statistics

### 3.3.1 Introduction to Bayesian Statistics

Bayesian Statistics is a branch of Statistics that offers a rationalist theory of personalistic beliefs in contexts of uncertainty, with the central aim of characterizing how an individual belief should act in order to avoid certain kinds of undesirable behavioral inconsistencies. The goal is to establish rules and procedures for individuals concerned with disciplined uncertainty accounting, Bernado & Smith (2001). In Bayesian Statistics the parameter(s) of interest is considered to be stochastic in order to represent all the various sources of uncertainty that we have in specifying its possible values.

The importance of Bayesian statistics is noted in Hamada et al. (2008) when they mention that consistency and efficiency are properties of estimators most relevant for estimation when sample sizes are large. However, since sample sizes available for estimations are never infinite, inference for small-to-moderate sample sizes is also important. The large sample properties of the maximum likelihood estimator do not pertain in more complicated settings. Take, for example, the hierarchical models, or when the parameter values fall close to the boundary of the parameter space. Lastly, deriving expressions for the maximum likelihood estimator is often difficult in high dimensional settings, and obtaining analytical expression for the information matrix required for inference may not be feasible. On the foundation of Bayesian Statistics lies Bayes' Theorem.

### 3.3.2 Bayes Theorem

Definitive records of Thomas Bayes' birth do not seem to exist. That his name lives on in the characterization of modern statistical methodology is a consequence of the publication of "An Essay towards Solving a Problem in the Doctrine of Chances", attributed to Bayes and community to the Royal Society after Bayes' death by Richard Price in 1763. The technical result at the heart of the essay is what we now know as Bayes' theorem. In its simplest form, if H denotes an hypothesis and D denotes data, the theorem states that

$$P(H \mid D) = P(D \mid H) \times P(H) / P(D)$$

with $P(H)$ regarded as a probabilistic statement of belief about $H$ before obtaining data $D$, the left hand side $P(H \mid D)$ becomes a probabilistic statement of belief about $H$ after obtaining $D$. Having specified $P(D \mid H)$ and $P(D)$ the mechanism of the theorem provides a solution to the problem of how to learn from data. Actually, Bayes only stated his result for a uniform prior. It was Laplace - apparently unaware of Bayes' work - who stated the theorem in its general (discrete) form, Bernado & Smith (2001). Mathematically, Bayes Theorem can be expressed in the general continuous form as

$$p(\theta \mid \underline{t}) = \frac{L(\theta \mid \underline{t})\, p(\theta)}{\int_{\theta} L(\theta \mid \underline{t})\, p(\theta)\, d\theta}$$

where the function $p(\theta \mid \underline{t})$ is called the posterior density, $p(\theta)$ is called the prior density of $\theta$, $L(\theta \mid \underline{t})$ is the sampling density of the data, $\int L(\theta \mid \underline{t})\, p(\theta)\, d\theta$ is the marginal density of the data, and $L(\theta \mid \underline{t})$ is the likelihood function. Thus we see that Bayes' Theorem provides the mathematical means of combining information and data, in the context of a probabilistic model, in order to update a prior state of knowledge. However, Bayes rule does not tell us what our beliefs should be, it tells us how they should change after seeing new information. The only problem in the use of Bayes' Theorem is that one has to be certain about what it is one is uncertain about.

According to Hoff (2009) Bayesian methods provide:

- parameter estimates with good statistical properties;
- parsimonious descriptions of observed data;
- predictions for missing data and forecasts of future data;
- a computational framework for model estimation, selection and validation.

### 3.3.3 Objective Priors

The density that expresses what one believes about the occurring value of $\theta$, before any observation has been taken, that is a priori, $p(\theta)$ is called the prior density of $\theta$. Objective priors will be considered in this thesis. These are prior distributions that may be dispersed when used to reflect prior information, indicating that little is known about the parameter of interest. In other words, objective priors are priors that are characterized by giving no preference to any of the possible parameter values. These priors are sometimes called diffuse, non-informative, or vague. We use vague priors when we feel that we have very little prior knowledge about the model parameters. The following objective priors will be considered:

- Jeffreys prior;

- General divergence prior.

#### 3.3.3.1 Jeffreys prior

This prior results from the use of Jeffreys' rule, described by Jeffreys (1946). Suppose that we have a one-to-one transformation of our parameter $\phi = h(\theta)$. Jeffreys' rule states that any rule for determining a prior distribution should yield the same prior distribution for $\phi$ whether we transform from a prior on $\theta$ or determine a prior directly for $\phi$. Define expected Fisher information as

$$I(\theta) = -E\left[\frac{d^2 \log L(\theta \mid \underline{t})}{d\theta^2}\right].$$

Jeffreys' rule defines a vague prior as

$$p_J(\theta) \propto \sqrt{|\, I(\theta) \,|}.$$

According to Syversveen (1998) there are problems, discussed in Box & Tiao (1992), that arise when Jeffreys' rule is generalized to multidimensional parameters $\theta$. The prior that maximizes missing information in an experiment, known as the reference prior is usually used in this case.

#### 3.3.3.2 A General divergence prior

The General divergence prior is defined in Gosh et al. (2011) as

$$p_G(\theta) \propto \sqrt[4]{|\, I(\theta) \,|}.$$

That is, the divergence prior is proportional to the positive fourth root of the Fisher information, or to the positive square root of the Jeffreys prior. This prior is proved, in Ghosh (2011), to possess the in-variance property under the one-to-one reparameterization. This prior has not received too much attention in literature; it is going to be compared with the Jeffreys prior.

**Table 3.1:** Fisher information for the four distributions.

| **Distribution** | **Fisher information** |
|---|---|
| Exponential | $\frac{n}{\lambda^2}$ |
| Rayleigh | $\frac{4n}{\lambda^2}$ |
| Gamma | $\begin{bmatrix} n\varphi'(\alpha) & -\frac{n}{\lambda} \\ -\frac{n}{\lambda} & \frac{n\alpha}{\lambda^2} \end{bmatrix}$ |
| Weibull | $\begin{bmatrix} \frac{n}{\alpha^2}+\frac{n(2\gamma_1+\gamma_2)}{\alpha^2} & \frac{n}{\lambda}+\frac{n\gamma_1}{\lambda} \\ \frac{n}{\lambda}+\frac{n\gamma_1}{\lambda} & \frac{n\alpha^2}{\lambda^2} \end{bmatrix}$ |

Table 3.1 gives the the Fisher information for each of the four distributions. The derivations of the given Fisher information is provided in Appendix A.1 to A.4.

Table 3.2 shows the two priors for the different distributions which will be used in this thesis.

**Table 3.2:** Priors.

| **Distribution** | **Jeffreys** | **General Divergence** |
|---|---|---|
| Exponential | $\frac{1}{\lambda}$ | $\frac{1}{\sqrt{\lambda}}$ |
| Rayleigh | $\frac{1}{\lambda}$ | $\frac{1}{\sqrt{\lambda}}$ |
| Gamma | $\frac{\sqrt{\alpha\varphi'(\alpha)-1}}{\lambda}$ | $\frac{(\alpha\varphi'(\alpha)-1)^{1/4}}{\sqrt{\lambda}}$ |
| Weibull | $\frac{1}{\lambda}$ | $\frac{1}{\sqrt{\lambda}}$ |

### 3.3.4 Posterior distribution

Bayes theorem tells us that the probability distribution for $\theta$ posterior to the data $x$ is proportional to the product of the distribution for $\theta$ prior to the data and the likelihood for $\theta$ given $x$. That is,

$$\text{posterior distribution} \propto \text{likelihood} \times \text{prior distribution}.$$

The likelihood plays a very important role in Bayes' formula, by being the function through which the data $x$ modifies prior knowledge of $\theta$. It can therefore be regarded as representing information about $\theta$ coming from the data. it is defined up to a multiplicative constant, Box & Tiao (1992).

In terms of probability density functions then,

$$p(\theta \mid \underline{t}) \propto L(\theta \mid \underline{t})\, p(\theta)\,.$$

The posterior distributions for the four parametric models, for complete data will now be considered. A similar approach can be adopted in the presence of censored data.

#### 3.3.4.1 Exponential distribution

Suppose we have a random sample $T_1, T_2, ..., T_n$ of lifetimes from an exponential distribution with rate parameter $\lambda$, and let $\underline{t} = (t_1, t_2, ..., t_n)$ denote the observed values of the random sample. The likelihood function of $\lambda$ given the observations $t$ is

$$\begin{aligned} L(\lambda \mid \underline{t}) &= \prod_{i=1}^{n} L(\lambda \mid t_i) \\ &= \prod_{i=1}^{n} \lambda e^{-\lambda t_i} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i}. \end{aligned}$$

The Jeffreys prior for $\lambda$ is given by

$$p_J(\lambda) \propto \sqrt{\mid -\frac{n}{\lambda^2} \mid} \propto \frac{1}{\lambda}.$$

The posterior distribution for $\lambda$ is

$$\begin{aligned} p_J(\lambda \mid \underline{t}) &\propto \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i} \times \frac{1}{\lambda} \\ &= \lambda^{n-1} e^{-\lambda \sum_{i=1}^{n} t_i}. \end{aligned}$$

Thus,

$$\lambda \mid \underline{t} \sim gamma\left(n, \sum_{i=1}^{n} t_i\right)$$

The general divergence prior for $\lambda$ is

$$p_G(\lambda) \propto \mid -\frac{n}{\lambda^2} \mid^{1/4} \propto \frac{1}{\sqrt{\lambda}}.$$

And the posterior distribution is

$$\begin{aligned} p_G(\lambda \mid \underline{t}) &\propto \lambda^n e^{-\lambda \sum_{i=1}^{n} t_i} \times \frac{1}{\sqrt{\lambda}} \\ &\propto \lambda^{n-1/2} e^{-\lambda \sum_{i=1}^{n} t_i}. \end{aligned} \tag{3.1}$$

Thus,

$$\lambda \mid \underline{t} \sim gamma\left(n + \frac{1}{2}, \sum_{i=1}^{n} t_i\right).$$

#### 3.3.4.2 Rayleigh distribution

Consider a random sample $T_1, T_2, ..., T_n$ of lifetimes from a Rayleigh distribution with scale parameter $\lambda$, and let $\underline{t} = (t_1, t_2, ..., t_n)$ denote the observed values of the random sample. The likelihood function of $\lambda$ given the observations $t$ is

$$\begin{aligned} L(\lambda \mid \underline{t}) &= \prod_{i=1}^{n} L(\lambda \mid t_i) \\ &= \prod_{i=1}^{n} 2\lambda^2 t_i e^{-(\lambda t_i)^2} \\ &= 2^n \lambda^{2n} \left[\prod_{i=1}^{n} t_i\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^2}. \end{aligned}$$

The Jeffreys prior for $\lambda$ is

$$p_J(\lambda) \propto \left| \frac{4n}{\lambda^2} \right|^{1/2} \propto \frac{1}{\lambda}.$$

The posterior distribution for $\lambda$ is given by

$$\begin{aligned} p_J(\lambda \mid \underline{t}) &\propto 2^n \lambda^{2n} \left[\prod_{i=1}^{n} t_i\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^2} \times \frac{1}{\lambda} \\ &\propto \lambda^{2n-1} e^{-\sum_{i=1}^{n}(\lambda t_i)^2}. \end{aligned}$$

The general divergence prior for $\lambda$ is

$$p_G(\lambda) \propto \left| \frac{4n}{\lambda^2} \right|^{1/4} \propto \frac{1}{\sqrt{\lambda}}.$$

The posterior distribution is

$$\begin{aligned} p_G(\lambda \mid \underline{t}) &\propto 2^n \lambda^{2n} \left[\prod_{i=1}^{n} t_i\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^2} \times \frac{1}{\sqrt{\lambda}} \\ &\propto \lambda^{2n-1/2} e^{-\sum_{i=1}^{n}(\lambda t_i)^2}. \end{aligned} \quad (3.2)$$

#### 3.3.4.3 Weibull distribution

Let $T_1, T_2, ..., T_n$ be a random sample from a Weibull distribution with parameters $\alpha > 0$ and $\lambda > 0$, and let $\underline{t} = (t_1, t_2, ..., t_n)$ be the observed values of this random sample. The joint likelihood function of $\alpha$ and $\lambda$ given $\underline{t}$ is

$$\begin{aligned} L(\alpha,\lambda \mid \underline{t}) &= \prod_{i=1}^{n} L(\lambda \mid t_i) \\ &= \prod_{i=1}^{n} \alpha\lambda\,(\lambda t_i)^{\alpha-1}\, e^{-(\lambda t_i)^{\alpha}} \\ &= (\alpha\lambda^{\alpha})^{n}\left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^{\alpha}}. \end{aligned}$$

The Jeffreys prior for $(\alpha,\lambda)$ is given by

$$p_J(\alpha,\lambda) \propto \sqrt{\mid \frac{n^2}{\lambda^2}\left(\gamma_2-\gamma_1^2\right) \mid} \propto \frac{1}{\lambda}.$$

The posterior distribution of $(\alpha,\lambda)$ is given by

$$\begin{aligned} p_J(\alpha,\lambda \mid \underline{t}) &\propto (\alpha\lambda^{\alpha})^{n}\left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^{\alpha}} \times \frac{1}{\lambda} \\ &\propto \alpha^{n}\lambda^{\alpha n-1}\left[\prod_{i=1}^{n} t_i^{\alpha}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^{\alpha}}. \end{aligned}$$

The general divergence prior of $(\alpha,\lambda)$ is

$$p_G(\alpha,\lambda \mid \underline{t}) \propto \mid \frac{n^2}{\lambda^2}\left(\gamma_2-\gamma_1^2\right) \mid^{1/4} \propto \frac{1}{\sqrt{\lambda}}.$$

The posterior distribution of $(\alpha,\lambda)$ is

$$\begin{aligned} P_G(\alpha,\lambda \mid \underline{t}) &\propto (\alpha\lambda^{\alpha})^{n}\left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^{\alpha}} \times \frac{1}{\sqrt{\lambda}} \\ &\propto \alpha^{n}\lambda^{\alpha n-1/2}\left[\prod_{i=1}^{n} t_i^{\alpha}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^{\alpha}}. \end{aligned} \tag{3.3}$$

#### 3.3.4.4 Gamma distribution

Suppose that we have a random sample $T_1, T_2, \ldots, T_n$ from a gamma distribution with parameters $\alpha > 0$ and $\lambda > 0$, and let $\underline{t} = (t_1, t_2, \ldots, t_n)$ be the observed values of this random sample. The joint likelihood function for $(\alpha, \lambda)$ is

$$\begin{aligned} L(\alpha, \lambda \mid \underline{t}) &= \prod_{i=1}^{n} L(\alpha, \lambda \mid t_i) \\ &= \prod_{i=1}^{n} \frac{\lambda^{\alpha} t_i^{\alpha-1} e^{-\lambda t_i}}{\Gamma(\alpha)} \\ &= \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^{n} \left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\lambda \sum_{i=1}^{n} t_i} \end{aligned}$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ is the gamma function.

The Jeffreys prior for $(\alpha, \lambda)$ is

$$p_J(\alpha, \lambda) \propto \frac{\sqrt{\alpha \varphi'(\alpha) - 1}}{\lambda}$$

The posterior distribution of $(\alpha, \lambda)$ is

$$\begin{aligned} p_J(\alpha, \lambda \mid \underline{t}) &\propto \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^{n} \left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\lambda \sum_{i=1}^{n} t_i} \times \frac{\sqrt{\alpha \varphi'(\alpha) - 1}}{\lambda}. \\ &\propto \frac{\lambda^{\alpha n - 1}}{(\Gamma(\alpha))^{n}} \left[\prod_{i=1}^{n} t_i^{\alpha}\right] e^{-\lambda \sum_{i=1}^{n} t_i} \times \sqrt{\alpha \varphi'(\alpha) - 1}. \end{aligned}$$

The general divergence prior is

$$\begin{aligned} p_G(\alpha, \lambda) &\propto \frac{\sqrt{n}\,(\alpha \varphi'(\alpha) - 1)^{1/4}}{\lambda^2} \\ &\propto \frac{(\alpha \varphi'(\alpha) - 1)^{1/4}}{\sqrt{\lambda}} \end{aligned}$$

and the posterior is

$$p_G(\alpha, \lambda \mid \underline{t}) \propto \left(\frac{\lambda^{\alpha}}{\Gamma(\alpha)}\right)^{n} \left[\prod_{i=1}^{n} t_i^{\alpha-1}\right] e^{-\lambda \sum_{i=1}^{n} t_i} \times \frac{(\alpha \varphi'(\alpha) - 1)^{1/4}}{\sqrt{\lambda}}. \tag{3.4}$$

In Bayesian statistics it is important, that the posterior distribution is proper. Improper posterior distributions yield invalid results. By a proper posterior distribution, we mean the integral of the posterior is finite. We will now investigate this element on the posterior distributions for the divergence prior that we have derived, as this has not yet been done in the literature reviewed thus far. The posterior distributions for the Jeffreys prior have been reported to be proper in, Yang & Berger (1996).

### 3.3.5 Properness of the resulting posterior distributions

When using the divergence prior and the exponential distribution, the posterior distribution from (3.1) is proper. This will be shown in Theorem 3.1.

**Theorem 3.1.** *The posterior distribution from (3.1),* $p_G(\lambda \mid \underline{t}) \propto \lambda^{n-\frac{1}{2}} e^{-\lambda \sum_{i=1}^{n} t_i}$*, is proper.*

*Proof.* Since

$$\int_0^\infty \lambda^{n-\frac{1}{2}} e^{-\lambda \sum_{i=1}^{n} t_i} d\lambda = \frac{\Gamma\left(n+\frac{1}{2}\right)}{\left(\sum_{i=1}^{n} t_i\right)^{n+\frac{1}{2}}},$$

and

$$\lambda \mid \underline{t} \sim gamma\left(n+\frac{1}{2}, \sum_{i=1}^{n} t_i\right).$$

□

When using the divergence prior and the Rayleigh distribution, the posterior distribution from (3.2) is proper. This will be shown in Theorem 3.2.

**Theorem 3.2.** *The posterior distribution from (3.2),* $p_G(\lambda \mid \underline{t}) \propto \lambda^{2n-1/2} e^{-\sum_{i=1}^{n} (\lambda t_i)^2}$*, is proper.*

*Proof.* To show that (3.2) is proper, the following should be true:

$$\int_0^\infty c\lambda^{2n-1/2} e^{-\sum_{i=1}^{n} (\lambda t_i)^2} d\lambda = 1,$$

where $c$ is the normalising constant.

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

$$\int_0^\infty c\lambda e^{-\sum_{i=1}^{n} (\lambda t_i)^2} d\lambda < \infty.$$

Note that

$$\int_0^\infty p_G(\lambda \mid \underline{t})\, d\lambda \quad = \quad \int_0^\infty \lambda^{2n-1/2} e^{-\sum_{i=1}^{n} (\lambda t_i)^2} d\lambda,$$

and by letting $x = \lambda^2$, and noting that $x \to \infty$ as $\lambda \to \infty$, and $\frac{d\lambda}{dx} = \frac{x^{\frac{1}{2}-1}}{2}$, we have

$$\begin{aligned}\int_0^\infty \lambda^{2n-1/2} e^{-\sum_{i=1}^n (\lambda t_i)^2} d\lambda &= \frac{1}{2}\int_0^\infty x^{\frac{1}{2}(2n+1/2)-1} e^{-x\sum_{i=1}^n t_i^2} dx \\ &= \frac{\Gamma\left(\frac{1}{2}(2n+1/2)\right)}{2\left(\sum_{i=1}^n t_i^2\right)^{\frac{1}{2}(2n+1/2)}} \\ &< \infty.\end{aligned}$$

Also, notice that

$$c^{-1} > \frac{\Gamma\left(\frac{1}{2}(2n+1/2)\right)}{2\left(\sum_{i=1}^n t_i^2\right)^{\frac{1}{2}(2n+1/2)}}$$

which completes the proof. □

When using the divergence prior and the Weibull distribution, the posterior from (3.3) is proper. This will be shown in Theorem 3.3.

**Theorem 3.3.** *The posterior distribution from (3.3),* $P_G(\alpha, \lambda \mid \underline{t}) \propto \alpha^n \lambda^{\alpha n - 1/2} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\sum_{i=1}^n (\lambda t_i)^\alpha}$, *is proper.*

*Proof.* To show that (3.3) is proper, the following should be true:

$$\int_0^\infty \int_0^\infty c\alpha^n \lambda^{\alpha n-1/2} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\sum_{i=1}^n (\lambda t_i)^\alpha} d\lambda d\alpha = 1,$$

where $c$ is the normalising constant.

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

$$\int_0^\infty \int_0^\infty \alpha^n \lambda^{\alpha n-1/2} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\sum_{i=1}^n (\lambda t_i)^\alpha} d\lambda d\alpha < \infty.$$

Carrying out the inner integration, by letting $x = \lambda^\alpha$, gives

$$\int_0^\infty \alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \int_0^\infty x^{n+1/2\alpha-1} e^{-x\sum_{i=1}^n t_i^\alpha} dx d\alpha$$

which simplifies to

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

But, it is well known that

$$\prod_{i=1}^{n} t_i^{\alpha} \leq \left(n^{-1}\sum_{i=1}^{n} t_i^{\alpha}\right)^{n},$$

which implies that

$$\begin{aligned}
\int_0^{\infty} \alpha^{n-1}\left[\prod_{i=1}^{n} t_i^{\alpha}\right]\Gamma\left(n+1/2\alpha\right)\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{-(n+1/2\alpha)} d\alpha &\leq \int_0^{\infty} \alpha^{n-1}\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{n} \frac{\Gamma\left(n+1/2\alpha\right)}{n^n\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{-(n+1/2\alpha)}} d\alpha \\
&= n^{-n}\int_0^{\infty} \alpha^{n-1}\Gamma\left(n+1/2\alpha\right)\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{-1/2\alpha} d\alpha.
\end{aligned}$$

Now, from Abramowitz & Stegun (1964), we obtain the asymptotic result

$$\Gamma\left(n+1/2\alpha\right) \sim \sqrt{2\pi}e^{1/2\alpha}\left(1/2\alpha\right)^{1/2\alpha+n-1/2},$$

thus, as $\alpha \to \infty$,

$$\begin{aligned}
\alpha^{n-1}\Gamma\left(n+1/2\alpha\right)\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{-1/2\alpha} &\sim 2^{-n+1}\sqrt{\pi}\alpha^{-1/2-1/2\alpha}\left(\frac{e}{2\sum_{i=1}^{n} t_i^{\alpha}}\right)^{1/2\alpha} \\
&= O\left(\frac{1}{W^{\alpha}}\right),\ W > 0.
\end{aligned}$$

Therefore,

$$\int_0^{\infty}\int_0^{\infty} \alpha^{n}\lambda^{\alpha n-1/2}\left[\prod_{i=1}^{n} t_i^{\alpha}\right] e^{-\sum_{i=1}^{n}(\lambda t_i)^{\alpha}} d\lambda d\alpha < \infty,$$

and that completes the proof. □

When using the divergence prior the gamma distribution, the posterior from (3.4) is proper. This will be shown in Theorem 3.4.

**Theorem 3.4.** *The posterior distribution from (3.4),*

$$P_G(\alpha,\lambda \mid \underline{t}) \propto \frac{\lambda^{n\alpha-\frac{1}{2}}}{\left(\Gamma(\alpha)\right)^n}\left(\alpha\varphi'(\alpha)-1\right)^{1/4}\left[\prod_{i=1}^{n} t_i^{\alpha}\right] e^{-\lambda\sum_{i=1}^{n} t_i}$$

*is proper.*

*Proof.* Taking a similar approach to that of Liseo (1990), first note that

$$\begin{aligned}
p_G(\alpha \mid \underline{t}) &\propto \int_0^\infty \frac{\lambda^{n\alpha-\frac{1}{2}}}{(\Gamma(\alpha))^n}\left(\alpha\varphi'(\alpha)-1\right)^{1/4}\left[\prod_{i=1}^n t_i^\alpha\right] e^{-\lambda\sum_{i=1}^n t_i} d\lambda \\
&= \frac{\left(\alpha\varphi'(\alpha)-1\right)^{1/4}}{(\Gamma(\alpha))^n}\left[\prod_{i=1}^n t_i^\alpha\right]\frac{\Gamma\left(n\alpha+\frac{1}{2}\right)}{\left(\sum_{i=1}^n t_i\right)^{n\alpha+\frac{1}{2}}} \\
&\propto \left(\alpha\varphi'(\alpha)-1\right)^{1/4}\left(\frac{\prod_{i=1}^n t_i}{\left(\sum_{i=1}^n t_i\right)^n}\right)^\alpha \frac{\Gamma\left(n\alpha+\frac{1}{2}\right)}{(\Gamma(\alpha))^n}
\end{aligned}$$

where

$$\varphi'(\alpha) \sim \frac{1}{\alpha}+\frac{1}{2\alpha^2}+\frac{1}{6\alpha^3}-\frac{1}{30\alpha^5}+\frac{1}{42\alpha^7}\cdots$$

so that

$$\alpha\varphi'(\alpha)-1 \sim \frac{1}{2\alpha}+\frac{1}{6\alpha^2}-\frac{1}{30\alpha^4}+\frac{1}{42\alpha^6}\cdots$$

and

$$\left(\alpha\varphi'(\alpha)-1\right)^{1/4} = O\left(\frac{1}{\sqrt[4]{\alpha}}\right) \text{ as } \alpha\to\infty$$

that is,

$$\left(\frac{(\alpha\varphi'(\alpha)-1)}{1/\alpha}\right)^{1/4}$$

is bounded for large values of $\alpha$.

From Abramowitz & Stegun (1964), we have that

$$\Gamma\left(n\alpha+\frac{1}{2}\right) \sim \sqrt{2\pi}\left(\frac{\alpha}{e}\right)^{n\alpha} n^{n\alpha}(n\alpha)^{\frac{1}{2}-\frac{1}{2}}$$

and

$$\Gamma(\alpha) \sim \sqrt{\frac{2\pi}{\alpha}}\left(\frac{\alpha}{e}\right)^\alpha\left[1+\frac{1}{12\alpha}+\frac{1}{288\alpha^2}-\frac{139}{51840\alpha^3}-\frac{571}{2488320\alpha^4}+\cdots\right].$$

Therefore

$$\begin{aligned}
\frac{\Gamma\left(n\alpha+\frac{1}{2}\right)}{(\Gamma(\alpha))^n} &\sim \frac{\sqrt{2\pi}\left(\frac{\alpha}{e}\right)^{n\alpha} n^{n\alpha}(n\alpha)^{\frac{1}{2}-\frac{1}{2}}}{\left(\sqrt{\frac{2\pi}{\alpha}}\left(\frac{\alpha}{e}\right)^\alpha\left[1+\frac{1}{12\alpha}+\frac{1}{288\alpha^2}-\frac{139}{51840\alpha^3}-\frac{571}{2488320\alpha^4}+\cdots\right]\right)^n} \\
&\sim (2\pi)^{\frac{1-n}{2}} n^{n\alpha}\alpha^{\frac{n}{2}} \quad \text{as } \alpha\to\infty.
\end{aligned}$$

Now consider the geometric mean $\sqrt[n]{\prod_{i=1}^{n} t_i}$ and the arithmetic mean $\frac{1}{n}\sum_{i=1}^{n} t_i$. It is a well known result that

$$\frac{\sqrt[n]{\prod_{i=1}^{n} t_i}}{\frac{1}{n}\sum_{i=1}^{n} t_i} \leq 1$$

$$\iff \left(\frac{\sqrt[n]{\prod_{i=1}^{n} t_i}}{\frac{1}{n}\sum_{i=1}^{n} t_i}\right)^n \leq 1$$

Thus, it follows that

$$\left(\frac{\prod_{i=1}^{n} t_i}{\left(\sum_{i=1}^{n} t_i\right)^n}\right)^{\alpha} \frac{\Gamma\left(n\alpha+\frac{1}{2}\right)}{\left(\Gamma\left(\alpha\right)\right)^n} \sim (2\pi)^{\frac{1-n}{2}} \alpha^{\frac{n}{2}} \left(\frac{n^n \prod_{i=1}^{n} t_i}{\left(\sum_{i=1}^{n} t_i\right)^n}\right)^{\alpha}$$

$$= O\left(\frac{1}{M^{\alpha}}\right), \quad M > 1$$

and at least one $t_i \neq 1$.

Therefore, as $\alpha \to \infty$,

$$P\left(\alpha \mid t\right) \propto O\left(\frac{1}{\sqrt[4]{\alpha}}\right) \cdot O\left(\frac{1}{M^{\alpha}}\right), \quad M > 1$$

is integrable for every value of $\left(\sum_{i=1}^{n} t_i, \prod_{i=1}^{n} t_i\right)$, where at least one $t_i \neq 1$, and thus

$$p_G\left(\alpha, \lambda \mid \underline{t}\right) \propto \frac{\lambda^{n\alpha-\frac{1}{2}}}{\left(\Gamma\left(\alpha\right)\right)^n} \left(\alpha\varphi'\left(\alpha\right)-1\right)^{1/4} \left[\prod_{i=1}^{n} t_i^{\alpha}\right] e^{-\lambda\sum_{i=1}^{n} t_i}$$

is proper. □

### 3.3.6 Conditional and marginal posterior distributions

In this section the conditional and marginal distributions of the gamma and Weibull distributions will be investigated, as they are two-parameter distributions among the distributions under consideration. The results of this section will be used in the simulation studies section. The conditional posterior distribution of $\lambda$ is found by treating $\alpha$ as a given constant in $p\left(\alpha, \lambda \mid \underline{t}\right)$, that is, the conditional posterior distribution of $\lambda$ is,

$$p(\lambda \mid \alpha, \underline{t}) = \frac{p(\alpha, \lambda \mid \underline{t})}{p(\alpha \mid \underline{t})}.$$

The marginal posterior distribution of $\alpha$, $p(\alpha \mid \underline{t})$, is found by simply integrating out all the other nuisance parameters from the joint posterior $p(\alpha, \lambda \mid \underline{t})$ (in our case, this is $\lambda$). That is,

$$p(\alpha \mid \underline{t}) = \int p(\alpha, \lambda \mid t)\, d\lambda.$$

#### 3.3.6.1 Gamma distribution

When using the general divergence prior, the marginal posterior for $\alpha$ will be

$$\begin{aligned}
p_G(\alpha \mid \underline{t}) &= \int_0^\infty p_G(\alpha, \lambda \mid \underline{t})\, d\lambda \\
&\propto \int_0^\infty \frac{\lambda^{n\alpha - \frac{1}{2}}}{(\Gamma(\alpha))^n} \left(\alpha\varphi'(\alpha) - 1\right)^{1/4} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\lambda \sum_{i=1}^n t_i} d\lambda \\
&= \frac{\left(\alpha\varphi'(\alpha) - 1\right)^{1/4} \left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n} \left[\int_0^\infty \lambda^{n\alpha - \frac{1}{2}} e^{-\lambda \sum_{i=1}^n t_i} d\lambda\right] \\
&= \frac{\left(\alpha\varphi'(\alpha) - 1\right)^{1/4} \left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n} \frac{\Gamma\left(\alpha n + \frac{1}{2}\right)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n + \frac{1}{2}}}
\end{aligned}$$

$$\therefore \quad p_G(\alpha \mid \underline{t}) \propto \frac{\left(\alpha\varphi'(\alpha) - 1\right)^{1/4} \left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n} \frac{\Gamma\left(\alpha n + \frac{1}{2}\right)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}. \tag{3.5}$$

When using the Jeffreys prior, the marginal posterior for $\alpha$ will be

$$\begin{aligned}
p_J(\alpha \mid \underline{t}) &\propto \int_0^\infty \frac{\lambda^{n\alpha - 1}}{(\Gamma(\alpha))^n} \left(\alpha\varphi'(\alpha) - 1\right)^{1/2} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\lambda \sum_{i=1}^n t_i} d\lambda \\
&= \frac{\left(\alpha\varphi'(\alpha) - 1\right)^{1/2} \left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n} \frac{\Gamma(\alpha n)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}.
\end{aligned} \tag{3.6}$$

When using the general divergence prior, the conditional posterior for $\lambda$ will be

$$
\begin{aligned}
p_G(\lambda \mid \alpha, \underline{t}) &= \frac{p_G(\alpha, \lambda \mid \underline{t})}{p_G(\alpha \mid \underline{t})} \\
&\propto \frac{\frac{\lambda^{n\alpha-\frac{1}{2}}}{(\Gamma(\alpha))^n}\left(\alpha\varphi'(\alpha)-1\right)^{1/4}\left[\prod_{i=1}^n t_i^\alpha\right] e^{-\lambda\sum_{i=1}^n t_i}}{\frac{(\alpha\varphi'(\alpha)-1)^{1/4}\left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n}\frac{\Gamma\left(\alpha n+\frac{1}{2}\right)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}} \\
\therefore \quad p_G(\lambda \mid \alpha, \underline{t}) &\propto \frac{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}{\Gamma\left(\alpha n+\frac{1}{2}\right)}\lambda^{n\alpha-\frac{1}{2}} e^{-\lambda\sum_{i=1}^n t_i} \qquad (3.7) \\
\Rightarrow \lambda \mid \alpha, \underline{t} &\sim gamma\left(\alpha n+\frac{1}{2}, \sum_{i=1}^n t_i\right).
\end{aligned}
$$

When using the Jeffreys prior, the conditional posterior for $\lambda$ will be

$$
\begin{aligned}
p_J(\lambda \mid \alpha, \underline{t}) &\propto \frac{\frac{\lambda^{n\alpha-1}}{(\Gamma(\alpha))^n}\left(\alpha\varphi'(\alpha)-1\right)^{1/2}\left[\prod_{i=1}^n t_i^\alpha\right] e^{-\lambda\sum_{i=1}^n t_i}}{\frac{(\alpha\varphi'(\alpha)-1)^{1/4}\left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n}\frac{\Gamma(\alpha n)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}} \\
&= \frac{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}{\Gamma(\alpha n)}\lambda^{n\alpha-1} e^{-\lambda\sum_{i=1}^n t_i} \qquad (3.8) \\
\therefore \quad \lambda \mid \alpha, \underline{t} &\sim gamma\left(\alpha n, \sum_{i=1}^n t_i\right)
\end{aligned}
$$

#### 3.3.6.2 Weibull distribution

When using the general divergence prior, the marginal posterior for $\alpha$ will be

$$
\begin{aligned}
p_G(\alpha \mid \underline{t}) &= \int_0^\infty p(\alpha, \lambda \mid \underline{t})\, d\lambda \\
&\propto \int_0^\infty \alpha^n \lambda^{\alpha n - 1/2} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\sum_{i=1}^n (\lambda t_i)^\alpha} d\lambda \\
&= \alpha^n \left[\prod_{i=1}^n t_i^\alpha\right] \int_0^\infty \lambda^{\alpha n - 1/2} e^{-\sum_{i=1}^n (\lambda t_i)^\alpha} d\lambda
\end{aligned}
$$

and by letting $x = \lambda^\alpha$, it follows that

$$
\begin{aligned}
p_G(\alpha \mid \underline{t}) &\propto \alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \int_0^\infty x^{n+\frac{1}{2\alpha}-1} e^{-x\sum_{i=1}^n t_i^\alpha} dx \\
&= \alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \frac{\Gamma\left(n+\frac{1}{2\alpha}\right)}{\left(\sum_{i=1}^n t_i^\alpha\right)^{n+\frac{1}{2\alpha}}}. \qquad (3.9)
\end{aligned}
$$

When using the Jeffreys prior, the marginal posterior for $\alpha$ will be

$$
\begin{aligned}
p_J(\alpha \mid \underline{t}) &= \int_0^\infty p_J(\alpha, \lambda \mid \underline{t})\, d\lambda \\
&\propto \alpha^n \left[\prod_{i=1}^n t_i^\alpha\right] \int_0^\infty \lambda^{\alpha n - 1} e^{-\sum_{i=1}^n (\lambda t_i)^\alpha} d\lambda \\
&= \alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \frac{\Gamma(n)}{\left(\sum_{i=1}^n t_i^\alpha\right)^n} \\
&\propto \alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \left(\sum_{i=1}^n t_i^\alpha\right)^{-n}. \qquad (3.10)
\end{aligned}
$$

When using the general divergence prior, the conditional posterior for $\lambda$ will be

$$\begin{aligned} p_G(\lambda \mid \alpha, \underline{t}) &= \frac{p_G(\alpha, \lambda \mid \underline{t})}{p_G(\alpha \mid \underline{t})} \\ &\propto \frac{\alpha^n \lambda^{\alpha n - 1/2} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\sum_{i=1}^n (\lambda t_i)^\alpha}}{\alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \frac{\Gamma\left(n+\frac{1}{2\alpha}\right)}{\left(\sum_{i=1}^n t_i^\alpha\right)^{n+\frac{1}{2\alpha}}}} \\ \therefore \quad p_G(\lambda \mid \alpha, \underline{t}) &\propto \frac{\alpha \left(\sum_{i=1}^n t_i^\alpha\right)^{n+\frac{1}{2\alpha}}}{\Gamma\left(n+\frac{1}{2\alpha}\right)} \lambda^{\alpha n - 1/2} e^{-\sum_{i=1}^n (\lambda t_i)^\alpha}. \end{aligned} \tag{3.11}$$

When using the Jeffreys prior, the conditional posterior for $\lambda$ will be

$$\begin{aligned} p_J(\lambda \mid \alpha, \underline{t}) &= \frac{p_J(\alpha, \lambda \mid \underline{t})}{p_J(\alpha \mid \underline{t})} \\ &\propto \frac{\alpha^n \lambda^{\alpha n - 1} \left[\prod_{i=1}^n t_i^\alpha\right] e^{-\sum_{i=1}^n (\lambda t_i)^\alpha}}{\alpha^{n-1} \left[\prod_{i=1}^n t_i^\alpha\right] \left(\sum_{i=1}^n t_i^\alpha\right)^{-n}} \\ &= \alpha \left(\sum_{i=1}^n t_i^\alpha\right)^n \lambda^{\alpha n - 1} e^{-\sum_{i=1}^n (\lambda t_i)^\alpha}. \end{aligned} \tag{3.12}$$

### 3.3.7 Predictive Reliability

According to Hoff (2009) the predictive distribution of a random variable $T_{n+1}$ is a probability distribution for $T_{n+1}$ such that

- known quantities have been conditioned on;
- unknown quantities have been integrated out.

A predictive distribution that integrates over unknown parameters but is not conditional on observed data is called a **prior predictive distribution**. Furthermore, such a distribution can be useful in evaluating if a prior distribution for $\theta$ actually translates into reasonable prior beliefs for observable data $T_{n+1}$. After we have observed a sample $T_1, \ldots, T_n$ from the population, the relevant predictive distribution for a new observation becomes

$$\begin{aligned} P(T_{n+1} = t_{n+1} \mid T_1 = t_1, \ldots, T_n = t_n) &= \int_\theta p(t_{n+1} \mid \theta, t_1, \ldots, t_n)\, p(\theta \mid t_1, \ldots, t_n)\, d\theta \\ &= p(t_{n+1} \mid t_1, \ldots, t_n) \\ &\propto \int_\theta p(t_{n+1} \mid \theta)\, p(\theta \mid t_1, \ldots, t_n)\, d\theta. \end{aligned}$$

This is called a **posterior predictive distribution**, because it conditions on an observed data set. Thus we see that $p(t_{n+1} \mid t_1, ..., t_n)$ is the posterior expectation of $p(t_{n+1} \mid \theta)$. We can generalize this result to cover the case of $m$ future observations, $T_{n+1}, ..., T_{n+m}$ in light of $t_1, ..., t_n$. It can be readily verified that

$$p(t_{n+1}, ..., t_{n+m} \mid t_1, ..., t_n) = \int_\theta \prod_{n+1}^{m} p(t_i \mid \theta) \, p(\theta \mid t_1, ..., t_n) \, d\theta.$$

# Chapter 4

# Simulation studies for the parametric models

## 4.1 Introduction

When one speaks of simulation studies they are essentially referring to the act of taking a model generally expressed as a computer program and running it a number of times in order to obtain results that can be compared. In this chapter we will perform simulation studies to compare the performance of Jeffreys and general divergence prior distributions for the four parametric distributions under study - exponential distribution, gamma distribution, Rayleigh distribution, and Weibull distribution. The performance of the priors will be evaluated by looking at coverage rate and interval lengths.

- **Credibility intervals**

  Credibility intervals are the Bayesian analogues of classical confidence intervals. They are constructed by using a central $(1-\alpha)\times 100\%$ interval, which is a a range of values having $(\alpha/2)\times 100\%$ of the posterior probability above and below the endpoints.

- **Coverage probability**

  An interval $\left(\theta_{\left(\frac{\alpha}{2}n\right)}, \theta_{\left(\left(1-\frac{\alpha}{2}\right)n\right)}\right)$, based on the observed data $\underline{T}=\underline{t}$, has 95% Bayesian coverage probability for $\theta$ if

$$P\left(\theta_{\left(\frac{\alpha}{2}n\right)} < \theta < \theta_{\left(\left(1-\frac{\alpha}{2}\right)n\right)} \mid T=t\right) = 0.95$$

  where $n$ is the number of samples generated by a simulation. Hoff (2009) interpret this interval by saying that it describes our information about the location of the true value of $\theta$ after we have observed $\underline{T}=\underline{t}$. We will look at 95% credibility intervals. Therefore, if the credibility intervals are correct, we should get a coverage probability of 0.95. But, since we are dealing with a finite number of simulations, our credibility intervals are not all going to have an exact coverage probability of 0.95. Before we start with the simulation, we first look at some important concepts that surround coverage, as discussed in Burton et al. (2006).

- Over-coverage occurs when the coverage rates are above the nominal confidence level, $CL = 1-\alpha$, and suggests that the results are too conservative as more simulations will not find a significant result when there is a true effect; this leads to a loss of statistical power with too many type II error. Confidence intervals whose coverage probabilities are greater than the nominal confidence level are said to be conservative.
- Under-coverage occurs when the coverage rates are lower than the nominal confidence level; this indicates over-confidence in the estimates since more simulations will incorrectly detect a significant result, which leads to higher than expected type I errors.

An important question that arises out of this is: what criterion is used to measure the acceptability of the coverage rate being good or bad? Well it turns out that, as noted in Burton et al. (2006), a possible criterion for acceptability of the coverage is that the coverage should not fall outside of approximately two standard errors of the nominal confidence level ($CL$), SE $(CL) = \sqrt{CL(1-CL)/n}$. The average interval lengths of the intervals will also be investigated. The average length of the intervals are calculated using the following formula, where $I_i$ is the interval length and $n^*$ the number of intervals:

$$\text{average length} \quad = \quad \frac{1}{n^*}\sum_{i=1}^{n^*} I_i.$$

The ideal interval, is the interval with the shortest length and coverage closest to the nominal level.

The general simulation study for the two priors and posteriors will be done, in MATLAB®, as follows:

1. For a starting (given) value of the unknown parameter, simulate data from the distribution of interest.

2. Simulate values for the unknown parameter, from the posterior distribution. Repeat this for large number of times, say, 10000.

3. Order the 10000 values of the parameter from the smallest to biggest. The 95% credibility interval will then be $\left(\theta_{(250)}, \theta_{(9750)}\right)$.

4. Simulate 10000 data sets, and repeat Step 1 to Step 3. Then determine how many of the credibility interval contain the initial value of the unknown parameter. This will give us the coverage probability for the prior being used.

The code for all the simulations performed in this chapter are given in Appendix B.

## 4.2 Simulation for the exponential distribution

This section looks at the simulation study where the life distribution is the exponential distribution. Following the general procedure given at the beginning of this chapter, the simulation for this distribution was done as follows:

1. 300 different values for the unknown parameter, $\lambda$, will be considered.

2. 10000 values of the unknown parameter were simulated from the posterior distributions:

$$p_J(\lambda \mid \underline{t}) \propto \lambda^{n-1} e^{-\lambda \sum_{i=1}^{n} t_i}$$

$$p_G(\lambda \mid \underline{t}) \propto \lambda^{n-1/2} e^{-\lambda \sum_{i=1}^{n} t_i}.$$

3. The resulting parameter values from posteriors are ordered in ascending order. The 95% credibility interval $\left(\lambda_{(250)}, \lambda_{(9750)}\right)$ is then determined.

4. 10000 data sets were simulated, and Step 1 through Step 3 were repeated. The number of times a credibility interval in Step 3 contained the initial value of the unknown parameter was calculated. This number divided by 10000 gives the coverage probability for the prior distribution used.

Each of the 10000 simulated coverage probabilities from the credibility intervals and average interval lengths are recorded in Tables 4.1 to 4.5.

**Table 4.1:** Coverage rates and average interval lengths for $\lambda = 0.1 : 0.1 : 6.4$, for the exponential distribution.

| | $\lambda$ | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.08% | 94.99% | 95.04% | 95.06% | 94.87% | 94.81% | 95.16% | 94.56% |
| | **Average L** | 0.0566 | 0.1130 | 0.1695 | 0.2261 | 0.2832 | 0.3390 | 0.3953 | 0.4516 |
| **G** | **Coverage** | 94.86% | 94.9% | 94.95% | 94.96% | 94.71% | 94.8% | 95.06% | 94.61% |
| | **Average L** | 0.0569 | 0.1135 | 0.1704 | 0.2272 | 0.2847 | 0.3407 | 0.3973 | 0.4538 |
| | $\lambda$ | **0.9** | **1** | **1.1** | **1.2** | **1.3** | **1.4** | **1.5** | **1.6** |
| **J** | **Coverage** | 94.95% | 94.68% | 95.29% | 94.61% | 94.82% | 95.24% | 95.01% | 94.98% |
| | **Average L** | 0.5077 | 0.5647 | 0.6208 | 0.6773 | 0.7334 | 0.7903 | 0.8452 | 0.9053 |
| **G** | **Coverage** | 94.85% | 94.66% | 95.11% | 94.42% | 94.61% | 95.24% | 94.86% | 94.85% |
| | **Average L** | 0.5103 | 0.5674 | 0.6238 | 0.6807 | 0.7370 | 0.7942 | 0.8493 | 0.9098 |
| | $\lambda$ | **1.7** | **1.8** | **1.9** | **2** | **2.1** | **2.2** | **2.3** | **2.4** |
| **J** | **Coverage** | 95.36% | 94.88% | 95.02% | 94.9% | 95.24% | 95.26% | 94.98% | 95.04% |
| | **Average L** | 0.9567 | 1.0170 | 1.0713 | 1.1314 | 1.1857 | 1.2422 | 1.2963 | 1.3569 |
| **G** | **Coverage** | 95.39% | 94.6% | 94.92% | 94.76% | 95.17% | 95.25% | 94.84% | 94.89% |
| | **Average L** | 0.9615 | 1.0221 | 1.0767 | 1.1371 | 1.1915 | 1.2482 | 1.3024 | 1.3638 |
| | $\lambda$ | **2.5** | **2.6** | **2.7** | **2.8** | **2.9** | **3** | **3.1** | **3.2** |
| **J** | **Coverage** | 95.1% | 94.74% | 94.97% | 94.85% | 95% | 94.54% | 95.05% | 95.17% |
| | **Average L** | 1.4092 | 1.4706 | 1.5235 | 1.5796 | 1.6397 | 1.6919 | 1.7517 | 1.8063 |
| **G** | **Coverage** | 95% | 94.49% | 95.01% | 94.91% | 94.91% | 94.4% | 94.85% | 95.03% |
| | **Average L** | 1.4163 | 1.4778 | 1.5307 | 1.5873 | 1.6480 | 1.7004 | 1.7604 | 1.8153 |
| | $\lambda$ | **3.3** | **3.4** | **3.5** | **3.6** | **3.7** | **3.8** | **3.9** | **4** |
| **J** | **Coverage** | 95% | 95.39% | 95.35% | 95.1% | 95.18% | 94.63% | 95.19% | 94.71% |
| | **Average L** | 1.8646 | 1.9209 | 1.9841 | 2.0347 | 2.0919 | 2.1505 | 2.2039 | 2.2539 |
| **G** | **Coverage** | 94.84% | 95.3% | 95.15% | 95.06% | 95.05% | 94.47% | 95.04% | 94.64% |
| | **Average L** | 1.8741 | 1.9303 | 1.9940 | 2.0447 | 2.1019 | 2.1614 | 2.2144 | 2.2651 |
| | $\lambda$ | **4.1** | **4.2** | **4.3** | **4.4** | **4.5** | **4.6** | **4.7** | **4.8** |
| **J** | **Coverage** | 95.05% | 95.27% | 94.87% | 94.74% | 95.18% | 95.44% | 95.08% | 94.72% |
| | **Average L** | 2.3127 | 2.3757 | 2.4283 | 2.4838 | 2.5413 | 2.5960 | 2.6561 | 2.7066 |
| **G** | **Coverage** | 94.87% | 95.05% | 94.77% | 94.73% | 95.06% | 95.39% | 94.83% | 94.61% |
| | **Average L** | 2.3240 | 2.3877 | 2.4407 | 2.4960 | 2.5538 | 2.6087 | 2.6695 | 2.7198 |
| | $\lambda$ | **4.9** | **5** | **5.1** | **5.2** | **5.3** | **5.4** | **5.5** | **5.6** |
| **J** | **Coverage** | 94.91% | 94.94% | 94.96% | 94.92% | 94.81% | 94.95% | 95.26% | 94.99% |
| | **Average L** | 2.7717 | 2.8352 | 2.8799 | 2.9451 | 2.9991 | 3.0523 | 3.1062 | 3.1646 |
| **G** | **Coverage** | 94.88% | 94.83% | 94.83% | 94.93% | 94.71% | 94.8% | 95.31% | 94.73% |
| | **Average L** | 2.7860 | 2.8493 | 2.8945 | 2.9599 | 3.0138 | 3.0669 | 3.1225 | 3.1806 |
| | $\lambda$ | **5.7** | **5.8** | **5.9** | **6** | **6.1** | **6.2** | **6.3** | **6.4** |
| **J** | **Coverage** | 94.69% | 94.95% | 94.6% | 94.93% | 95.23% | 95.14% | 94.92% | 95.18% |
| | **Average L** | 3.2187 | 3.2780 | 3.3372 | 3.3902 | 3.4447 | 3.5099 | 3.5571 | 3.6150 |
| **G** | **Coverage** | 94.74% | 94.88% | 94.54% | 94.81% | 95.16% | 95.02% | 94.93% | 95.02% |
| | **Average L** | 3.2352 | 3.2945 | 3.3540 | 3.4067 | 3.4614 | 3.5276 | 3.5756 | 3.6330 |

**Table 4.2:** Coverage rates and average interval lengths for $\lambda = 6.5 : 0.1 : 12.8$, for the exponential distribution.

| | $\lambda$ | **6.5** | **6.6** | **6.7** | **6.8** | **6.9** | **7** | **7.1** | **7.2** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.29% | 95.05% | 94.87% | 95.14% | 95.19% | 95.14% | 95.42% | 95.02% |
| | **Average L** | 3.6661 | 3.7301 | 3.7883 | 3.8375 | 3.8806 | 3.9527 | 4.0173 | 4.0597 |
| **G** | **Coverage** | 95.11% | 95% | 94.67% | 95.14% | 95.04% | 95.02% | 95.12% | 95.01% |
| | **Average L** | 3.6846 | 3.7478 | 3.8070 | 3.8569 | 3.9010 | 3.9723 | 4.0380 | 4.0811 |
| | $\lambda$ | **7.3** | **7.4** | **7.5** | **7.6** | **7.7** | **7.8** | **7.9** | **8** |
| **J** | **Coverage** | 94.97% | 95.06% | 95.06% | 94.94% | 95.08% | 94.76% | 94.74% | 95.21% |
| | **Average L** | 4.1251 | 4.1790 | 4.2400 | 4.2941 | 4.3475 | 4.4021 | 4.4675 | 4.5186 |
| **G** | **Coverage** | 94.92% | 94.99% | 94.95% | 94.82% | 94.88% | 94.72% | 94.66% | 95.08% |
| | **Average L** | 4.1455 | 4.1991 | 4.2623 | 4.3149 | 4.3697 | 4.4244 | 4.4898 | 4.5417 |
| | $\lambda$ | **8.1** | **8.2** | **8.3** | **8.4** | **8.5** | **8.6** | **8.7** | **8.8** |
| **J** | **Coverage** | 94.94% | 94.91% | 94.78% | 95.25% | 94.76% | 95.36% | 94.65% | 95.07% |
| | **Average L** | 4.5660 | 4.6298 | 4.6870 | 4.7394 | 4.7980 | 4.8620 | 4.9115 | 4.9696 |
| **G** | **Coverage** | 94.85% | 94.71% | 94.69% | 95.34% | 94.86% | 95.11% | 94.47% | 95.07% |
| | **Average L** | 4.5890 | 4.6519 | 4.7114 | 4.7631 | 4.8220 | 4.8852 | 4.9358 | 4.9947 |
| | $\lambda$ | **8.9** | **9** | **9.1** | **9.2** | **9.3** | **9.4** | **9.5** | **9.6** |
| **J** | **Coverage** | 94.83% | 95.02% | 95.13% | 94.93% | 94.66% | 95.14% | 94.82% | 94.9% |
| | **Average L** | 5.0272 | 5.0803 | 5.1390 | 5.1955 | 5.2517 | 5.3100 | 5.3761 | 5.4137 |
| **G** | **Coverage** | 94.75% | 94.79% | 94.98% | 94.7% | 94.52% | 94.94% | 94.65% | 94.77% |
| | **Average L** | 5.0531 | 5.1048 | 5.1650 | 5.2222 | 5.2788 | 5.3358 | 5.4034 | 5.4409 |
| | $\lambda$ | **9.7** | **9.8** | **9.9** | **10** | **10.1** | **10.2** | **10.3** | **10.4** |
| **J** | **Coverage** | 95.2% | 95.24% | 95.26% | 95.15% | 95.04% | 95.11% | 95.03% | 95.11% |
| | **Average L** | 5.4668 | 5.5360 | 5.5924 | 5.6417 | 5.6903 | 5.7436 | 5.8190 | 5.8892 |
| **G** | **Coverage** | 95.07% | 95.06% | 94.99% | 95.04% | 94.98% | 95.03% | 94.81% | 94.73% |
| | **Average L** | 5.4946 | 5.5644 | 5.6191 | 5.6707 | 5.7195 | 5.7727 | 5.8475 | 5.9179 |
| | $\lambda$ | **10.5** | **10.6** | **10.7** | **10.8** | **10.9** | **11** | **11.1** | **11.2** |
| **J** | **Coverage** | 94.54% | 94.68% | 95.03% | 95.03% | 94.55% | 94.64% | 95.09% | 94.88% |
| | **Average L** | 5.9215 | 5.9965 | 6.0461 | 6.0970 | 6.1481 | 6.2076 | 6.2474 | 6.3174 |
| **G** | **Coverage** | 94.46% | 94.7% | 94.99% | 94.88% | 94.46% | 94.56% | 95.05% | 94.8% |
| | **Average L** | 5.9525 | 6.0258 | 6.0764 | 6.1273 | 6.1776 | 6.2384 | 6.2784 | 6.3489 |
| | $\lambda$ | **11.3** | **11.4** | **11.5** | **11.6** | **11.7** | **11.8** | **11.9** | **12** |
| **J** | **Coverage** | 95.09% | 94.87% | 94.79% | 95.05% | 94.74% | 95.13% | 94.8% | 95.33% |
| | **Average L** | 6.3857 | 6.4435 | 6.4889 | 6.5487 | 6.6052 | 6.6600 | 6.7170 | 6.7649 |
| **G** | **Coverage** | 94.98% | 94.9% | 94.8% | 94.9% | 94.68% | 95.08% | 94.73% | 95.39% |
| | **Average L** | 6.4173 | 6.4753 | 6.5214 | 6.5808 | 6.6383 | 6.6953 | 6.7503 | 6.7994 |
| | $\lambda$ | **12.1** | **12.2** | **12.3** | **12.4** | **12.5** | **12.6** | **12.7** | **12.8** |
| **J** | **Coverage** | 95.44% | 94.62% | 95.18% | 94.67% | 94.87% | 95.16% | 95.17% | 95.17% |
| | **Average L** | 6.8400 | 6.8808 | 6.9625 | 6.9861 | 7.0492 | 7.1124 | 7.1673 | 7.2229 |
| **G** | **Coverage** | 95.2% | 94.6% | 95.13% | 94.6% | 94.89% | 95% | 95.08% | 95.06% |
| | **Average L** | 6.8737 | 6.9151 | 6.9977 | 7.0201 | 7.0851 | 7.1483 | 7.2031 | 7.2606 |

**Table 4.3:** Coverage rates and average interval lengths for $\lambda = 12.9 : 0.1 : 18.4$, for the exponential distribution.

| | $\lambda$ | **12.9** | **13** | **13.1** | **13.2** | **13.3** | **13.4** | **13.5** | **13.6** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.93% | 95.02% | 94.78% | 94.74% | 94.87% | 94.89% | 95.31% | 95.23% |
| | **Average L** | 7.2838 | 7.3533 | 7.3940 | 7.4506 | 7.4974 | 7.5808 | 7.6016 | 7.6726 |
| **G** | **Coverage** | 95.08% | 94.8% | 94.8% | 94.66% | 94.83% | 94.69% | 95.32% | 95.21% |
| | **Average L** | 7.3189 | 7.3884 | 7.4301 | 7.4873 | 7.5355 | 7.6196 | 7.6385 | 7.7101 |

| | $\lambda$ | **13.7** | **13.8** | **13.9** | **14** | **14.1** | **14.2** | **14.3** | **14.4** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.14% | 94.99% | 94.75% | 95.11% | 95.22% | 95.11% | 95.34% | 94.92% |
| | **Average L** | 7.7425 | 7.7821 | 7.8443 | 7.9179 | 7.9603 | 8.0267 | 8.0590 | 8.1330 |
| **G** | **Coverage** | 95.07% | 94.94% | 94.48% | 95% | 95.02% | 94.97% | 95.23% | 94.69% |
| | **Average L** | 7.7809 | 7.8221 | 7.8824 | 7.9591 | 7.9994 | 8.0670 | 8.0984 | 8.1742 |

| | $\lambda$ | **14.5** | **14.6** | **14.7** | **14.8** | **14.9** | **15** | **15.1** | **15.2** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.83% | 94.84% | 94.76% | 95.06% | 94.9% | 95.11% | 94.54% | 95.03% |
| | **Average L** | 8.1857 | 8.2521 | 8.2857 | 8.3585 | 8.4060 | 8.4674 | 8.5255 | 8.5945 |
| **G** | **Coverage** | 94.84% | 94.84% | 94.65% | 94.86% | 94.9% | 95.06% | 94.35% | 95.06% |
| | **Average L** | 8.2271 | 8.2929 | 8.3302 | 8.3998 | 8.4473 | 8.5088 | 8.5672 | 8.6388 |

| | $\lambda$ | **15.3** | **15.4** | **15.5** | **15.6** | **15.7** | **15.8** | **15.9** | **16** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.89% | 95.05% | 95.09% | 95.04% | 95.28% | 94.88% | 94.64% | 95.33% |
| | **Average L** | 8.6707 | 8.7128 | 8.7698 | 8.7979 | 8.8531 | 8.9229 | 9.0017 | 9.0429 |
| **G** | **Coverage** | 94.6% | 95.09% | 94.93% | 95.06% | 95.16% | 94.84% | 94.6% | 95.19% |
| | **Average L** | 8.7141 | 8.7563 | 8.8143 | 8.8403 | 8.8979 | 8.9673 | 9.0488 | 9.0904 |

| | $\lambda$ | **16.1** | **16.2** | **16.3** | **16.4** | **16.5** | **16.6** | **16.7** | **16.8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.71% | 94.87% | 94.63% | 95.27% | 95% | 94.84% | 95.28% | 95.04% |
| | **Average L** | 9.1071 | 9.1585 | 9.2104 | 9.2750 | 9.3184 | 9.3857 | 9.4322 | 9.4738 |
| **G** | **Coverage** | 94.51% | 94.73% | 94.72% | 95.31% | 95.04% | 94.72% | 95.04% | 94.96% |
| | **Average L** | 9.1528 | 9.2034 | 9.2548 | 9.3237 | 9.3673 | 9.4310 | 9.4796 | 9.5199 |

| | $\lambda$ | **16.9** | **17** | **17.1** | **17.2** | **17.3** | **17.4** | **17.5** | **17.6** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.05% | 95.09% | 94.88% | 94.98% | 94.65% | 95% | 94.87% | 95.26% |
| | **Average L** | 9.5537 | 9.6090 | 9.6518 | 9.7126 | 9.7699 | 9.8112 | 9.8634 | 9.9407 |
| **G** | **Coverage** | 94.99% | 95.13% | 94.91% | 94.89% | 94.44% | 94.84% | 94.86% | 95.16% |
| | **Average L** | 9.6022 | 9.6567 | 9.7008 | 9.7578 | 9.8180 | 9.8604 | 9.9128 | 9.9915 |

| | $\lambda$ | **17.7** | **17.8** | **17.9** | **18** | **18.1** | **18.2** | **18.3** | **18.4** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.85% | 95.39% | 95.13% | 94.68% | 95.24% | 94.61% | 94.97% | 95.43% |
| | **Average L** | 9.9730 | 10.0643 | 10.1019 | 10.1608 | 10.2191 | 10.2941 | 10.3513 | 10.3915 |
| **G** | **Coverage** | 94.95% | 95.32% | 95.06% | 94.75% | 95.27% | 94.6% | 94.72% | 95.31% |
| | **Average L** | 10.0219 | 10.1138 | 10.1520 | 10.2126 | 10.2711 | 10.3461 | 10.4011 | 10.4461 |

**Table 4.4:** Coverage rates and average interval lengths for $\lambda = 18.5 : 0.1 : 24$, for the exponential distribution.

| | $\lambda$ | **18.5** | **18.6** | **18.7** | **18.8** | **18.9** | **19** | **19.1** | **19.2** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.86% | 94.99% | 94.94% | 95.27% | 95.03% | 94.96% | 94.66% | 95.02% |
| | **Average L** | 10.4577 | 10.4714 | 10.5521 | 10.6083 | 10.6581 | 10.7210 | 10.7822 | 10.8059 |
| **G** | **Coverage** | 94.71% | 94.96% | 94.99% | 95.16% | 95.05% | 94.77% | 94.61% | 95% |
| | **Average L** | 10.5064 | 10.5249 | 10.6056 | 10.6595 | 10.7128 | 10.7746 | 10.8373 | 10.8599 |
| | $\lambda$ | **19.3** | **19.4** | **19.5** | **19.6** | **19.7** | **19.8** | **19.9** | **20** |
| **J** | **Coverage** | 94.8% | 94.87% | 94.95% | 94.88% | 95.16% | 94.84% | 94.7% | 95.15% |
| | **Average L** | 10.9017 | 10.9385 | 11.0111 | 11.0607 | 11.1206 | 11.1588 | 11.2376 | 11.2897 |
| **G** | **Coverage** | 94.77% | 94.68% | 94.93% | 94.9% | 95.24% | 94.85% | 94.66% | 94.92% |
| | **Average L** | 10.9563 | 10.9913 | 11.0654 | 11.1167 | 11.1758 | 11.2162 | 11.2924 | 11.3489 |
| | $\lambda$ | **20.1** | **20.2** | **20.3** | **20.4** | **20.5** | **20.6** | **20.7** | **20.8** |
| **J** | **Coverage** | 95.18% | 94.71% | 94.61% | 94.94% | 94.24% | 95.09% | 95.02% | 95.22% |
| | **Average L** | 11.3721 | 11.4226 | 11.4718 | 11.5154 | 11.5980 | 11.6115 | 11.6866 | 11.7571 |
| **G** | **Coverage** | 95.13% | 94.58% | 94.54% | 94.8% | 94.22% | 95.01% | 94.92% | 95.01% |
| | **Average L** | 11.4286 | 11.4801 | 11.5292 | 11.5710 | 11.6577 | 11.6704 | 11.7442 | 11.8152 |
| | $\lambda$ | **20.9** | **21** | **21.1** | **21.2** | **21.3** | **21.4** | **21.5** | **21.6** |
| **J** | **Coverage** | 94.41% | 95.16% | 95.18% | 94.75% | 94.66% | 95.24% | 94.7% | 95.05% |
| | **Average L** | 11.7659 | 11.8546 | 11.9276 | 11.9508 | 12.0497 | 12.0965 | 12.1597 | 12.2070 |
| **G** | **Coverage** | 94.28% | 94.93% | 95.05% | 94.77% | 94.61% | 95.1% | 94.46% | 95% |
| | **Average L** | 11.8264 | 11.9156 | 11.9869 | 12.0100 | 12.1076 | 12.1564 | 12.2219 | 12.2684 |
| | $\lambda$ | **21.7** | **21.8** | **21.9** | **22** | **22.1** | **22.2** | **22.3** | **22.4** |
| **J** | **Coverage** | 95.09% | 94.83% | 95.28% | 94.9% | 94.83% | 94.73% | 95.36% | 95.18% |
| | **Average L** | 12.2535 | 12.2986 | 12.3700 | 12.4259 | 12.4563 | 12.5520 | 12.6061 | 12.6443 |
| **G** | **Coverage** | 94.92% | 94.91% | 95.23% | 94.78% | 94.63% | 94.7% | 95.27% | 95.17% |
| | **Average L** | 12.3124 | 12.3589 | 12.4305 | 12.4890 | 12.5186 | 12.6170 | 12.6703 | 12.7054 |
| | $\lambda$ | **22.5** | **22.6** | **22.7** | **22.8** | **22.9** | **23** | **23.1** | **23.2** |
| **J** | **Coverage** | 94.83% | 94.94% | 94.88% | 94.75% | 94.84% | 95.07% | 95.11% | 95.41% |
| | **Average L** | 12.7044 | 12.7516 | 12.8366 | 12.8965 | 12.9445 | 12.9937 | 13.0661 | 13.0910 |
| **G** | **Coverage** | 94.76% | 95.01% | 94.55% | 94.65% | 94.8% | 95.04% | 95.06% | 95.24% |
| | **Average L** | 12.7686 | 12.8160 | 12.9028 | 12.9613 | 13.0099 | 13.0602 | 13.1315 | 13.1572 |
| | $\lambda$ | **23.3** | **23.4** | **23.5** | **23.6** | **23.7** | **23.8** | **23.9** | **24** |
| **J** | **Coverage** | 95.29% | 95.04% | 95.05% | 94.86% | 94.75% | 95.31% | 94.55% | 94.9% |
| | **Average L** | 13.1753 | 13.1828 | 13.2616 | 13.3394 | 13.3482 | 13.4431 | 13.5463 | 13.5589 |
| **G** | **Coverage** | 95.26% | 94.9% | 95.02% | 94.73% | 94.67% | 95.19% | 94.4% | 94.84% |
| | **Average L** | 13.2403 | 13.2490 | 13.3312 | 13.4058 | 13.4154 | 13.5088 | 13.6178 | 13.6286 |

**Table 4.5:** Coverage rates and average interval lengths for $\lambda = 24.1 : 0.1 : 30$, for the exponential distribution.

| | $\lambda$ | **24.1** | **24.2** | **24.3** | **24.4** | **24.5** | **24.6** | **24.7** | **24.8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.87% | 95.01% | 94.6% | 94.78% | 94.89% | 95.1% | 95.32% | 95.19% |
| | **Average L** | 13.6385 | 13.6801 | 13.7086 | 13.7739 | 13.7970 | 13.9177 | 13.9479 | 14.0075 |
| **G** | **Coverage** | 94.72% | 94.87% | 94.66% | 94.8% | 94.78% | 94.87% | 95.13% | 95.09% |
| | **Average L** | 13.7080 | 13.7449 | 13.7757 | 13.8423 | 13.8640 | 13.9923 | 14.0198 | 14.0764 |

| | $\lambda$ | **24.9** | **25** | **25.1** | **25.2** | **25.3** | **25.4** | **25.5** | **25.6** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.85% | 94.59% | 94.7% | 94.63% | 95.09% | 95.04% | 94.88% | 94.75% |
| | **Average L** | 14.0548 | 14.1439 | 14.1968 | 14.2251 | 14.2894 | 14.3730 | 14.3831 | 14.4754 |
| **G** | **Coverage** | 94.69% | 94.4% | 94.63% | 94.7% | 95.02% | 94.73% | 94.81% | 94.54% |
| | **Average L** | 14.1234 | 14.2143 | 14.2710 | 14.2958 | 14.3603 | 14.4439 | 14.4547 | 14.5460 |

| | $\lambda$ | **25.7** | **25.8** | **25.9** | **26** | **26.1** | **26.2** | **26.3** | **26.4** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.8% | 95.08% | 95.37% | 95.38% | 94.89% | 94.77% | 95.07% | 94.98% |
| | **Average L** | 14.5004 | 14.5882 | 14.6143 | 14.6660 | 14.7319 | 14.8190 | 14.8177 | 14.8599 |
| **G** | **Coverage** | 94.76% | 95.16% | 95.35% | 95.05% | 94.77% | 94.75% | 95.05% | 94.84% |
| | **Average L** | 14.5766 | 14.6614 | 14.6866 | 14.7389 | 14.8066 | 14.8913 | 14.8949 | 14.9318 |

| | $\lambda$ | **26.5** | **26.6** | **26.7** | **26.8** | **26.9** | **27** | **27.1** | **27.2** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.05% | 95.19% | 94.85% | 94.5% | 95.5% | 94.68% | 94.52% | 95.04% |
| | **Average L** | 14.9477 | 15.0572 | 15.0826 | 15.1354 | 15.1941 | 15.2226 | 15.2922 | 15.3486 |
| **G** | **Coverage** | 94.95% | 95.12% | 94.68% | 94.53% | 95.38% | 94.58% | 94.52% | 94.91% |
| | **Average L** | 15.0219 | 15.1313 | 15.1594 | 15.2131 | 15.2697 | 15.2953 | 15.3684 | 15.4254 |

| | $\lambda$ | **27.3** | **27.4** | **27.5** | **27.6** | **27.7** | **27.8** | **27.9** | **28** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.95% | 94.7% | 95.11% | 95.19% | 95.08% | 95% | 94.81% | 94.91% |
| | **Average L** | 15.4039 | 15.4747 | 15.5221 | 15.5771 | 15.6180 | 15.7056 | 15.7479 | 15.8352 |
| **G** | **Coverage** | 94.84% | 94.64% | 95.05% | 95.17% | 94.95% | 94.74% | 94.58% | 94.73% |
| | **Average L** | 15.4804 | 15.5549 | 15.5966 | 15.6595 | 15.6971 | 15.7816 | 15.8315 | 15.9134 |

| | $\lambda$ | **28.1** | **28.2** | **28.3** | **28.4** | **28.5** | **28.6** | **28.7** | **28.8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.37% | 95.08% | 95.05% | 95.11% | 94.46% | 95.18% | 95.18% | 95% |
| | **Average L** | 15.8546 | 15.9299 | 16.0060 | 16.0112 | 16.1517 | 16.1447 | 16.1777 | 16.2256 |
| **G** | **Coverage** | 95.19% | 94.83% | 94.93% | 94.97% | 94.38% | 95.16% | 95.12% | 94.9% |
| | **Average L** | 15.9310 | 16.0095 | 16.0849 | 16.0942 | 16.2351 | 16.2256 | 16.2570 | 16.3066 |

| | $\lambda$ | **28.9** | **29** | **29.1** | **29.2** | **29.3** | **29.4** | **29.5** | **29.6** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.07% | 95.32% | 95.16% | 94.71% | 94.78% | 94.74% | 94.55% | 94.64% |
| | **Average L** | 16.2947 | 16.3600 | 16.4334 | 16.4900 | 16.4980 | 16.5789 | 16.6707 | 16.7084 |
| **G** | **Coverage** | 95.11% | 95.34% | 95.06% | 94.61% | 94.68% | 94.73% | 94.49% | 94.72% |
| | **Average L** | 16.3726 | 16.4407 | 16.5198 | 16.5708 | 16.5827 | 16.6621 | 16.7526 | 16.7909 |

| | $\lambda$ | **29.7** | **29.8** | **29.9** | **30** |
|---|---|---|---|---|---|
| **J** | **Coverage** | 95.19% | 94.85% | 94.94% | 95.2% |
| | **Average L** | 16.7337 | 16.8736 | 16.8885 | 16.9211 |
| **G** | **Coverage** | 94.98% | 94.45% | 94.83% | 95.18% |
| | **Average L** | 16.8185 | 16.9595 | 16.9711 | 17.0046 |

**Table 4.6:** Overall average coverage rates and mean interval lengths for the exponential distribution, when $\lambda = 0.1 : 0.1 : 30$.

| | | **Average** |
|---|---|---|
| **J** | **Coverage** | 94.98% |
| | **Average L** | 8.4982 |
| **G** | **Coverage** | 94.88% |
| | **Average L** | 8.5407 |

A close inspection of the tables for the exponential distribution reveals a good coverage probability performance for both prior distributions under study, as there is no sign of poor coverage - all the coverage probabilities are within the range [94.22%, 95.5%]. However, in Table 4.6 it can be seen that on average, the Jeffreys prior produces better coverage probability than the general divergence prior. The average credibility interval lengths for both priors are approximately equal, and they increase as the value of lambda increases. The average interval lengths produced by the Jeffreys prior are narrower. This is an important observation as we desire narrow credibility intervals. On average, the Jeffreys prior produces narrower average credibility interval lengths than the general divergence prior.

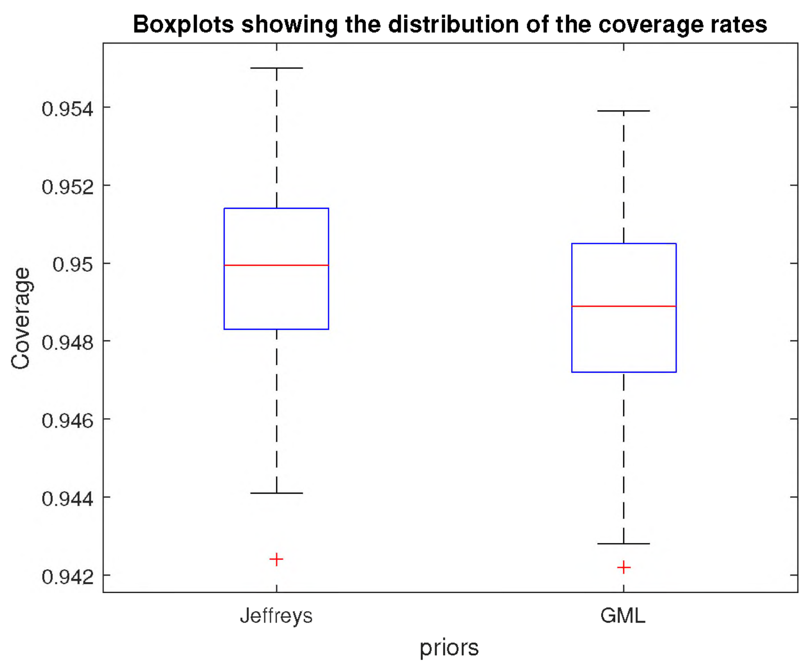


**Figure 4.1:** Box-plots for the exponential distribution showing the distribution of the coverage rates.

Figure 4.1 shows the distribution of the coverage probabilities for the two priors. The two distributions have approximately the same spread. We see that the Jeffreys prior has a higher median coverage probability than the general divergence prior. In fact, approximately half of the coverage probabilities from the Jeffreys prior slightly exceed 0.95 - a sign of over-coverage. The general divergence prior, on the other hand, has more than half of its coverage probabilities less than 0.95 - this might indicate under-coverage.

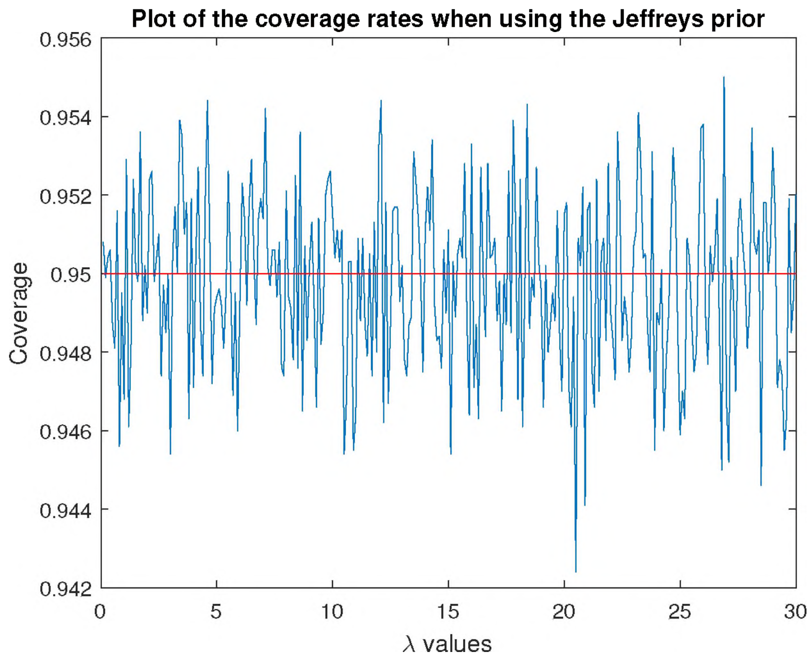


**Figure 4.2:** Coverage rates plots for exponential distribution using the Jeffreys prior

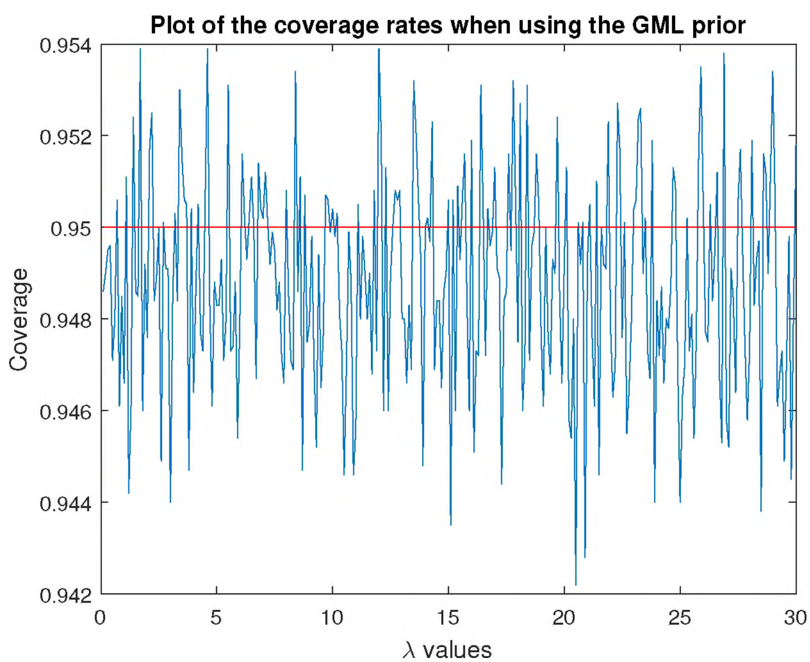


**Figure 4.3:** Coverage rates plots for exponential distribution using the general divergence prior.

Figure 4.2 is consistent with the results of Figure 4.1, that the Jeffreys prior produces higher coverage probabilities than the general divergence prior. From Figure 4.3 it can be seen that the general divergence prior has more than half of its coverage probabilities less than 0.95 - a sign of under-coverage.

## 4.3 Simulation for the Rayleigh distribution

This section looks at the simulation study where the life distribution is the Rayleigh distribution. Following the general procedure given at the beginning of this chapter, the simulation for this distribution was done as follows:

1. 300 different values for the unknown parameter, $\lambda$, will be considered.

2. 10000 values of the unknown parameter were simulated from the posterior distributions:

$$p_J(\lambda \mid \underline{t}) \propto \lambda^{2n-1} e^{-\sum_{i=1}^{n}(\lambda t_i)^2}$$

$$p_G(\lambda \mid \underline{t}) \propto \lambda^{2n-1/2} e^{-\sum_{i=1}^{n}(\lambda t_i)^2}.$$

3. The resulting parameter values from posteriors are ordered in ascending order. The 95% credibility interval $\left(\lambda_{(250)}, \lambda_{(9750)}\right)$ is then determined.

4. 10000 data sets were simulated, and Step 1 through Step 3 were repeated. The number of times a credibility interval in Step 3 contained the initial value of the unknown parameter was calculated. This number divided by 10000 gives the coverage probability for the prior distribution used.

The results are shown in Tables 4.7 to 4.11.

**Table 4.7:** Coverage rates and average interval lengths for $\lambda = 0.1 : 0.7.2$, for the Rayleigh distribution.

| | $\lambda$ | **0.1** | **0.2** | **0.3** | **0.4** | **0.5** | **0.6** | **0.7** | **0.8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.12% | 94.72% | 94.99% | 95.14% | 95.09% | 95.35% | 95.22% | 94.84% |
| | **Average L** | 0.0279 | 0.0558 | 0.0836 | 0.1115 | 0.1395 | 0.1673 | 0.1953 | 0.2234 |
| **G** | **Coverage** | 95.09% | 94.68% | 94.87% | 95.02% | 95.03% | 95.34% | 95.12% | 94.75% |
| | **Average L** | 0.0279 | 0.0558 | 0.0836 | 0.1115 | 0.1395 | 0.1673 | 0.1953 | 0.2233 |
| | $\lambda$ | **0.9** | **1** | **1.1** | **1.2** | **1.3** | **1.4** | **1.5** | **1.6** |
| **J** | **Coverage** | 94.76% | 94.76% | 95.28% | 94.54% | 94.75% | 94.97% | 94.85% | 95.01% |
| | **Average L** | 0.2507 | 0.2786 | 0.3064 | 0.3346 | 0.3626 | 0.3909 | 0.4178 | 0.4460 |
| **G** | **Coverage** | 94.64% | 94.71% | 95.3% | 94.44% | 94.79% | 94.87% | 94.71% | 94.98% |
| | **Average L** | 0.2508 | 0.2786 | 0.3064 | 0.3345 | 0.3627 | 0.3909 | 0.4179 | 0.4460 |
| | $\lambda$ | **1.7** | **1.8** | **1.9** | **2** | **2.1** | **2.2** | **2.3** | **2.4** |
| **J** | **Coverage** | 95.27% | 94.88% | 94.73% | 94.82% | 94.26% | 94.84% | 95% | 94.79% |
| | **Average L** | 0.4740 | 0.5018 | 0.5297 | 0.5572 | 0.5856 | 0.6128 | 0.6411 | 0.6687 |
| **G** | **Coverage** | 95.25% | 94.83% | 94.68% | 94.65% | 94.14% | 94.83% | 94.97% | 94.66% |
| | **Average L** | 0.4741 | 0.5019 | 0.5298 | 0.5571 | 0.5856 | 0.6127 | 0.6410 | 0.6687 |
| | $\lambda$ | **2.5** | **2.6** | **2.7** | **2.8** | **2.9** | **3** | **3.1** | **3.2** |
| **J** | **Coverage** | 95.46% | 94.96% | 95.06% | 95.1% | 94.86% | 95.09% | 94.83% | 94.81% |
| | **Average L** | 0.6966 | 0.7251 | 0.7533 | 0.7809 | 0.8078 | 0.8364 | 0.8635 | 0.8920 |
| **G** | **Coverage** | 95.32% | 94.82% | 94.93% | 94.76% | 94.77% | 95.1% | 94.85% | 94.76% |
| | **Average L** | 0.6967 | 0.7251 | 0.7533 | 0.7807 | 0.8079 | 0.8361 | 0.8636 | 0.8920 |
| | $\lambda$ | **3.3** | **3.4** | **3.5** | **3.6** | **3.7** | **3.8** | **3.9** | **4** |
| **J** | **Coverage** | 94.72% | 94.84% | 94.66% | 95.05% | 95.16% | 95.12% | 94.69% | 94.58% |
| | **Average L** | 0.9193 | 0.9480 | 0.9765 | 1.0045 | 1.0310 | 1.0594 | 1.0865 | 1.1152 |
| **G** | **Coverage** | 94.6% | 94.71% | 94.58% | 94.9% | 95.09% | 95.02% | 94.63% | 94.6% |
| | **Average L** | 0.9193 | 0.9479 | 0.9765 | 1.0045 | 1.0310 | 1.0595 | 1.0866 | 1.1154 |
| | $\lambda$ | **4.1** | **4.2** | **4.3** | **4.4** | **4.5** | **4.6** | **4.7** | **4.8** |
| **J** | **Coverage** | 94.82% | 95.06% | 94.98% | 94.7% | 94.94% | 94.65% | 95.1% | 94.64% |
| | **Average L** | 1.1439 | 1.1703 | 1.1985 | 1.2273 | 1.2530 | 1.2823 | 1.3102 | 1.3386 |
| **G** | **Coverage** | 94.82% | 94.92% | 94.83% | 94.46% | 94.85% | 94.64% | 94.85% | 94.5% |
| | **Average L** | 1.1440 | 1.1700 | 1.1985 | 1.2273 | 1.2531 | 1.2823 | 1.3101 | 1.3386 |
| | $\lambda$ | **4.9** | **5** | **5.1** | **5.2** | **5.3** | **5.4** | **5.5** | **5.6** |
| **J** | **Coverage** | 94.71% | 94.9% | 94.74% | 94.97 | 94.78% | 95.14% | 94.99% | 94.96% |
| | **Average L** | 1.3672 | 1.3941 | 1.4219 | 1.4492 | 1.4781 | 1.5073 | 1.5333 | 1.5626 |
| **G** | **Coverage** | 94.53% | 94.91% | 94.66% | 94.92% | 94.71% | 94.93% | 94.89% | 94.74% |
| | **Average L** | 1.3669 | 1.3940 | 1.4217 | 1.4489 | 1.4781 | 1.5072 | 1.5330 | 1.5620 |
| | $\lambda$ | **5.7** | **5.8** | **5.9** | **6** | **6.1** | **6.2** | **6.3** | **6.4** |
| **J** | **Coverage** | 94.85% | 94.87% | 94.81% | 94.81% | 95.05% | 94.87% | 94.75% | 95.35% |
| | **Average L** | 1.5878 | 1.6186 | 1.6424 | 1.6729 | 1.6990 | 1.7278 | 1.7556 | 1.7852 |
| **G** | **Coverage** | 94.69% | 94.76% | 94.78% | 94.78% | 94.88% | 94.81% | 94.58% | 95.15% |
| | **Average L** | 1.5877 | 1.6182 | 1.6421 | 1.6729 | 1.6992 | 1.7279 | 1.7557 | 1.7852 |
| | $\lambda$ | **6.5** | **6.6** | **6.7** | **6.8** | **6.9** | **7** | **7.1** | **7.2** |
| **J** | **Coverage** | 94.69% | 94.97% | 94.92% | 95.15% | 94.81% | 94.79% | 94.57% | 95.2% |
| | **Average L** | 1.8128 | 1.8399 | 1.8660 | 1.8959 | 1.9231 | 1.9533 | 1.9801 | 2.0060 |
| **G** | **Coverage** | 94.59% | 94.77% | 94.8% | 95.22% | 94.72% | 94.7% | 94.48% | 95.13% |
| | **Average L** | 1.8130 | 1.8396 | 1.8663 | 1.8958 | 1.9233 | 1.9534 | 1.9806 | 2.0063 |

**Table 4.8:** Coverage rates and average interval lengths for $\lambda = 7.3 : 0.1 : 14.4$, for the Rayleigh distribution.

| | $\lambda$ | **7.3** | **7.4** | **7.5** | **7.6** | **7.7** | **7.8** | **7.9** | **8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.99% | 94.85% | 94.71% | 94.75% | 95.16% | 94.99% | 94.87% | 94.96% |
| | **Average L** | 2.0344 | 2.0621 | 2.0906 | 2.1178 | 2.1456 | 2.1747 | 2.2012 | 2.2319 |
| **G** | **Coverage** | 94.83% | 94.74% | 94.65% | 94.74% | 94.97% | 94.8% | 94.77% | 94.8% |
| | **Average L** | 2.0345 | 2.0623 | 2.0906 | 2.1180 | 2.1452 | 2.1749 | 2.2011 | 2.2317 |
| | $\lambda$ | **8.1** | **8.2** | **8.3** | **8.4** | **8.5** | **8.6** | **8.7** | **8.8** |
| **J** | **Coverage** | 94.99% | 95.04% | 95.41% | 95.03% | 95.22% | 94.72% | 94.92% | 95.2% |
| | **Average L** | 2.2596 | 2.2851 | 2.3110 | 2.3428 | 2.3716 | 2.3984 | 2.4285 | 2.4517 |
| **G** | **Coverage** | 94.95% | 94.99% | 95.27% | 94.94% | 95.09% | 94.67% | 94.95% | 95.24% |
| | **Average L** | 2.2597 | 2.2854 | 2.3110 | 2.3429 | 2.3717 | 2.3979 | 2.4288 | 2.4516 |
| | $\lambda$ | **8.9** | **9** | **9.1** | **9.2** | **9.3** | **9.4** | **9.5** | **9.6** |
| **J** | **Coverage** | 95.11% | 94.45% | 94.67% | 94.69% | 94.7% | 94.81% | 94.72% | 94.95% |
| | **Average L** | 2.4809 | 2.5053 | 2.5375 | 2.5661 | 2.5941 | 2.6182 | 2.6469 | 2.6772 |
| **G** | **Coverage** | 95.06% | 94.45% | 94.63% | 94.51% | 94.66% | 94.75% | 94.65% | 94.93% |
| | **Average L** | 2.4811 | 2.5051 | 2.5377 | 2.5658 | 2.5944 | 2.6186 | 2.6474 | 2.6773 |
| | $\lambda$ | **9.7** | **9.8** | **9.9** | **10** | **10.1** | **10.2** | **10.3** | **10.4** |
| **J** | **Coverage** | 95.19% | 94.68% | 95.24% | 94.91% | 94.98% | 95.23% | 94.96% | 95.15% |
| | **Average L** | 2.7109 | 2.7356 | 2.7591 | 2.7880 | 2.8180 | 2.8430 | 2.8714 | 2.8977 |
| **G** | **Coverage** | 95% | 94.55% | 95.07% | 94.89% | 94.88% | 95.15% | 94.84% | 94.98% |
| | **Average L** | 2.7098 | 2.7360 | 2.7590 | 2.7884 | 2.8181 | 2.8430 | 2.8718 | 2.8984 |
| | $\lambda$ | **10.5** | **10.6** | **10.7** | **10.8** | **10.9** | **11** | **11.1** | **11.2** |
| **J** | **Coverage** | 94.55% | 94.94% | 95.09% | 94.75% | 94.83% | 95% | 95.2% | 95.19% |
| | **Average L** | 2.9301 | 2.9530 | 2.9832 | 3.0127 | 3.0374 | 3.0669 | 3.0927 | 3.1229 |
| **G** | **Coverage** | 94.49% | 95.06% | 94.87% | 94.61% | 94.79% | 94.9% | 95.13% | 95.09% |
| | **Average L** | 2.9297 | 2.9527 | 2.9836 | 3.0121 | 3.0375 | 3.0673 | 3.0922 | 3.1226 |
| | $\lambda$ | **11.3** | **11.4** | **11.5** | **11.6** | **11.7** | **11.8** | **11.9** | **12** |
| **J** | **Coverage** | 95.12% | 95.04% | 94.69% | 95.09% | 95.3% | 94.61% | 94.71% | 95.21% |
| | **Average L** | 3.1533 | 3.1762 | 3.2109 | 3.2306 | 3.2609 | 3.2902 | 3.3186 | 3.3417 |
| **G** | **Coverage** | 94.98% | 94.96% | 94.61% | 94.9% | 95.29% | 94.53% | 94.54% | 95.18% |
| | **Average L** | 3.1537 | 3.1761 | 3.2104 | 3.2312 | 3.2610 | 3.2905 | 3.3186 | 3.3418 |
| | $\lambda$ | **12.1** | **12.2** | **12.3** | **12.4** | **12.5** | **12.6** | **12.7** | **12.8** |
| **J** | **Coverage** | 95.22% | 95.36% | 95.11% | 94.82% | 95.02% | 95.23% | 94.53% | 94.73% |
| | **Average L** | 3.3741 | 3.4028 | 3.4302 | 3.4602 | 3.4854 | 3.5095 | 3.5406 | 3.5697 |
| **G** | **Coverage** | 94.97% | 95.09% | 94.97% | 94.64% | 94.93% | 95.19% | 94.54% | 94.61% |
| | **Average L** | 3.3747 | 3.4028 | 3.4299 | 3.4599 | 3.4852 | 3.5098 | 3.5403 | 3.5698 |
| | $\lambda$ | **12.9** | **13** | **13.1** | **13.2** | **13.3** | **13.4** | **13.5** | **13.6** |
| **J** | **Coverage** | 94.89% | 94.83% | 94.81% | 94.92% | 94.95% | 95% | 95.21% | 94.92% |
| | **Average L** | 3.5996 | 3.6246 | 3.6535 | 3.6804 | 3.7053 | 3.7359 | 3.7625 | 3.7964 |
| **G** | **Coverage** | 94.76% | 94.8% | 94.76% | 94.77% | 94.87% | 94.96% | 95.12% | 94.74% |
| | **Average L** | 3.5995 | 3.6253 | 3.6533 | 3.6804 | 3.7061 | 3.7355 | 3.7615 | 3.7964 |
| | $\lambda$ | **13.7** | **13.8** | **13.9** | **14** | **14.1** | **14.2** | **14.3** | **14.4** |
| **J** | **Coverage** | 94.78% | 94.58% | 95.36% | 94.89% | 94.95% | 94.69% | 95% | 94.64% |
| | **Average L** | 3.8207 | 3.8433 | 3.8721 | 3.9053 | 3.9310 | 3.9578 | 3.9848 | 4.0162 |
| **G** | **Coverage** | 94.77% | 94.47% | 95.22% | 94.78% | 94.96% | 94.63% | 94.9% | 94.52% |
| | **Average L** | 3.8211 | 3.8431 | 3.8722 | 3.9065 | 3.9318 | 3.9576 | 3.9843 | 4.0153 |

**Table 4.9:** Coverage rates and average interval lengths for $\lambda = 14.5 : 0.1 : 20.8$, for the Rayleigh distribution.

| | $\lambda$ | **14.5** | **14.6** | **14.7** | **14.8** | **14.9** | **15** | **15.1** | **15.2** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.86% | 94.81% | 94.87% | 94.71% | 94.9% | 94.55% | 95.14% | 94.55% |
| | **Average L** | 4.0432 | 4.0708 | 4.0921 | 4.1217 | 4.1524 | 4.1838 | 4.2105 | 4.2404 |
| **G** | **Coverage** | 94.85% | 94.71% | 94.89% | 94.56% | 94.85% | 94.41% | 94.95% | 94.46% |
| | **Average L** | 4.0434 | 4.0708 | 4.0924 | 4.1215 | 4.1528 | 4.1848 | 4.2100 | 4.2411 |
| | $\lambda$ | **15.3** | **15.4** | **15.5** | **15.6** | **15.7** | **15.8** | **15.9** | **16** |
| **J** | **Coverage** | 95.17% | 94.91% | 94.65% | 95.18% | 94.79% | 94.99% | 94.6% | 95.13% |
| | **Average L** | 4.2637 | 4.2959 | 4.3271 | 4.3479 | 4.3756 | 4.4080 | 4.4372 | 4.4555 |
| **G** | **Coverage** | 95.02% | 94.79% | 94.51% | 95.02% | 94.54% | 94.83% | 94.52% | 94.96% |
| | **Average L** | 4.2641 | 4.2959 | 4.3273 | 4.3480 | 4.3749 | 4.4080 | 4.4365 | 4.4558 |
| | $\lambda$ | **16.1** | **16.2** | **16.3** | **16.4** | **16.5** | **16.6** | **16.7** | **16.8** |
| **J** | **Coverage** | 94.9% | 94.78% | 95.23% | 95.03% | 95.06% | 94.84% | 94.76% | 94.55% |
| | **Average L** | 4.4850 | 4.5180 | 4.5429 | 4.5748 | 4.6039 | 4.6234 | 4.6570 | 4.6853 |
| **G** | **Coverage** | 94.87% | 94.7% | 95.07% | 94.89% | 94.88% | 94.67% | 94.8% | 94.48% |
| | **Average L** | 4.4846 | 4.5190 | 4.5424 | 4.5743 | 4.6042 | 4.6232 | 4.6582 | 4.6851 |
| | $\lambda$ | **16.9** | **17** | **17.1** | **17.2** | **17.3** | **17.4** | **17.5** | **17.6** |
| **J** | **Coverage** | 94.54% | 94.73% | 94.66% | 95.16% | 94.87% | 94.59% | 95.11% | 94.92% |
| | **Average L** | 4.7136 | 4.7408 | 4.7697 | 4.7995 | 4.8209 | 4.8555 | 4.8709 | 4.9080 |
| **G** | **Coverage** | 94.44% | 94.72% | 94.65% | 95.14% | 94.8% | 94.41% | 95% | 94.84% |
| | **Average L** | 4.7134 | 4.7407 | 4.7705 | 4.7985 | 4.8215 | 4.8556 | 4.8718 | 4.9085 |
| | $\lambda$ | **17.7** | **17.8** | **17.9** | **18** | **18.1** | **18.2** | **18.3** | **18.4** |
| **J** | **Coverage** | 94.64% | 94.49% | 95.08% | 95.21% | 94.93% | 94.99% | 95.25% | 94.68% |
| | **Average L** | 4.9304 | 4.9655 | 4.9830 | 5.0160 | 5.0474 | 5.0796 | 5.0999 | 5.1286 |
| **G** | **Coverage** | 94.49% | 94.43% | 94.98% | 95.14% | 94.9% | 94.82% | 95.2% | 94.74% |
| | **Average L** | 4.9313 | 4.9664 | 4.9824 | 5.0165 | 5.0471 | 5.0804 | 5.1001 | 5.1288 |
| | $\lambda$ | **18.5** | **18.6** | **18.7** | **18.8** | **18.9** | **19** | **19.1** | **19.2** |
| **J** | **Coverage** | 95.08% | 94.7% | 95.21% | 94.74% | 95.57% | 94.72% | 94.56% | 94.89% |
| | **Average L** | 5.1595 | 5.1817 | 5.2151 | 5.2331 | 5.2680 | 5.2970 | 5.3260 | 5.3585 |
| **G** | **Coverage** | 95.1% | 94.73% | 95.16% | 94.66% | 95.39% | 94.67% | 94.3% | 94.75% |
| | **Average L** | 5.1596 | 5.1808 | 5.2155 | 5.2324 | 5.2685 | 5.2971 | 5.3268 | 5.3593 |
| | $\lambda$ | **19.3** | **19.4** | **19.5** | **19.6** | **19.7** | **19.8** | **19.9** | **20** |
| **J** | **Coverage** | 95.11% | 94.91% | 94.89% | 94.69% | 94.94% | 94.87% | 94.73% | 94.92% |
| | **Average L** | 5.3825 | 5.4030 | 5.4343 | 5.4663 | 5.4860 | 5.5210 | 5.5478 | 5.5775 |
| **G** | **Coverage** | 95.13% | 94.78% | 94.78% | 94.66% | 94.81% | 94.74% | 94.59% | 94.79% |
| | **Average L** | 5.3817 | 5.4037 | 5.4345 | 5.4652 | 5.4853 | 5.5211 | 5.5475 | 5.5773 |
| | $\lambda$ | **20.1** | **20.2** | **20.3** | **20.4** | **20.5** | **20.6** | **20.7** | **20.8** |
| **J** | **Coverage** | 94.87% | 95.19% | 94.97% | 94.83% | 94.58% | 94.55% | 94.39% | 94.7% |
| | **Average L** | 5.6036 | 5.6245 | 5.6595 | 5.6822 | 5.7164 | 5.7494 | 5.7815 | 5.8026 |
| **G** | **Coverage** | 94.82% | 95.17% | 94.92% | 94.82% | 94.5% | 94.47% | 94.35% | 94.66% |
| | **Average L** | 5.6044 | 5.6248 | 5.6596 | 5.6829 | 5.7163 | 5.7489 | 5.7802 | 5.8031 |

**Table 4.10:** Coverage rates and average interval lengths for $\lambda = 20.9 : 0.1 : 26.4$, for the Rayleigh distribution.

| | $\lambda$ | **20.9** | **21** | **21.1** | **21.2** | **21.3** | **21.4** | **21.5** | **21.6** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.92% | 94.88% | 95.12% | 95.17% | 94.49% | 95.1% | 95.16% | 94.67% |
| | **Average L** | 5.8344 | 5.8523 | 5.8907 | 5.9163 | 5.9407 | 5.9673 | 5.9933 | 6.0200 |
| **G** | **Coverage** | 94.84% | 94.75% | 95.05% | 94.97% | 94.29% | 94.88% | 95.09% | 94.52% |
| | **Average L** | 5.8347 | 5.8523 | 5.8914 | 5.9183 | 5.9409 | 5.9674 | 5.9923 | 6.0203 |
| | $\lambda$ | **21.7** | **21.8** | **21.9** | **22** | **22.1** | **22.2** | **22.3** | **22.4** |
| **J** | **Coverage** | 95.06% | 94.6% | 94.54% | 94.81% | 94.74% | 95.12% | 94.59% | 94.92% |
| | **Average L** | 6.0479 | 6.0839 | 6.1021 | 6.1307 | 6.1623 | 6.1928 | 6.2101 | 6.2431 |
| **G** | **Coverage** | 94.92% | 94.47% | 94.48% | 94.76% | 94.7% | 95.05% | 94.55% | 94.8% |
| | **Average L** | 6.0464 | 6.0832 | 6.1012 | 6.1300 | 6.1616 | 6.1938 | 6.2085 | 6.2429 |
| | $\lambda$ | **22.5** | **22.6** | **22.7** | **22.8** | **22.9** | **23** | **23.1** | **23.2** |
| **J** | **Coverage** | 94.85% | 94.49% | 95.15% | 94.75% | 94.75% | 94.72% | 95.1% | 94.83% |
| | **Average L** | 6.2811 | 6.3068 | 6.3281 | 6.3559 | 6.3848 | 6.4130 | 6.4443 | 6.4629 |
| **G** | **Coverage** | 94.7% | 94.41% | 95% | 94.62% | 94.72% | 94.59% | 95.05% | 94.78% |
| | **Average L** | 6.2806 | 6.3066 | 6.3289 | 6.3571 | 6.3863 | 6.4113 | 6.4435 | 6.4636 |
| | $\lambda$ | **23.3** | **23.4** | **23.5** | **23.6** | **23.7** | **23.8** | **23.9** | **24** |
| **J** | **Coverage** | 94.55% | 95.01% | 95% | 94.63% | 94.76% | 94.79% | 94.76% | 94.66% |
| | **Average L** | 6.4999 | 6.5324 | 6.5519 | 6.5743 | 6.6088 | 6.6389 | 6.6580 | 6.6928 |
| **G** | **Coverage** | 94.38% | 94.95% | 94.91% | 94.52% | 94.74% | 94.67% | 94.54% | 94.52% |
| | **Average L** | 6.5005 | 6.5332 | 6.5496 | 6.5738 | 6.6083 | 6.6396 | 6.6594 | 6.6917 |
| | $\lambda$ | **24.1** | **24.2** | **24.3** | **24.4** | **24.5** | **24.6** | **24.7** | **24.8** |
| **J** | **Coverage** | 94.28% | 94.98% | 94.41% | 94.99% | 95.05% | 94.76% | 94.8% | 94.97% |
| | **Average L** | 6.7186 | 6.7437 | 6.7764 | 6.8051 | 6.8333 | 6.8527 | 6.8868 | 6.9085 |
| **G** | **Coverage** | 94.07% | 94.95% | 94.31% | 94.87% | 95.05% | 94.75% | 94.73% | 94.92% |
| | **Average L** | 6.7188 | 6.7441 | 6.7758 | 6.8050 | 6.8333 | 6.8524 | 6.8859 | 6.9084 |
| | $\lambda$ | **24.9** | **25** | **25.1** | **25.2** | **25.3** | **25.4** | **25.5** | **25.6** |
| **J** | **Coverage** | 94.85% | 94.68% | 94.63% | 95.29% | 94.7% | 95.19% | 94.85% | 94.67% |
| | **Average L** | 6.9451 | 6.9738 | 6.9998 | 7.0300 | 7.0574 | 7.0819 | 7.1096 | 7.1370 |
| **G** | **Coverage** | 94.84% | 94.59% | 94.53% | 95.06% | 94.64% | 94.99% | 94.79% | 94.69% |
| | **Average L** | 6.9441 | 6.9727 | 7.0011 | 7.0296 | 7.0572 | 7.0832 | 7.1088 | 7.1378 |
| | $\lambda$ | **25.7** | **25.8** | **25.9** | **26** | **26.1** | **26.2** | **26.3** | **26.4** |
| **J** | **Coverage** | 95.15% | 95.2% | 95.03% | 95.04% | 94.78% | 94.57% | 94.59% | 94.63% |
| | **Average L** | 7.1651 | 7.1851 | 7.2273 | 7.2538 | 7.2719 | 7.3044 | 7.3285 | 7.3600 |
| **G** | **Coverage** | 95.06% | 95.12% | 94.87% | 95% | 94.6% | 94.57% | 94.53% | 94.63% |
| | **Average L** | 7.1649 | 7.1846 | 7.2285 | 7.2541 | 7.2722 | 7.3051 | 7.3272 | 7.3597 |

**Table 4.11:** Coverage rates and average interval lengths for $\lambda = 26.5 : 0.1 : 30$, for the Rayleigh distribution.

| | $\lambda$ | **26.5** | **26.6** | **26.7** | **26.8** | **26.9** | **27** | **27.1** | **27.2** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.04% | 94.75% | 95.03% | 95.13% | 94.66% | 95.17% | 95.23% | 95.17% |
| | **Average L** | 7.3928 | 7.4182 | 7.4456 | 7.4706 | 7.4924 | 7.5177 | 7.5512 | 7.5855 |
| **G** | **Coverage** | 94.91% | 94.66% | 94.95% | 95.06% | 94.71% | 95.05% | 95.13% | 95.04% |
| | **Average L** | 7.3940 | 7.4177 | 7.4460 | 7.4718 | 7.4908 | 7.5188 | 7.5504 | 7.5860 |

| | $\lambda$ | **27.3** | **27.4** | **27.5** | **27.6** | **27.7** | **27.8** | **27.9** | **28** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.99% | 95.34% | 94.83% | 94.9% | 94.77% | 94.96% | 94.59% | 94.93% |
| | **Average L** | 7.6079 | 7.6393 | 7.6708 | 7.6842 | 7.7135 | 7.7489 | 7.7698 | 7.8158 |
| **G** | **Coverage** | 94.86% | 95.1% | 94.74% | 94.93% | 94.59% | 94.93% | 94.57% | 94.85% |
| | **Average L** | 7.6064 | 7.6387 | 7.6698 | 7.6851 | 7.7146 | 7.7485 | 7.7717 | 7.8156 |

| | $\lambda$ | **28.1** | **28.2** | **28.3** | **28.4** | **28.5** | **28.6** | **28.7** | **28.8** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 94.54% | 95.29% | 95.03% | 95.05% | 94.97% | 95.26% | 94.96% | 94.99% |
| | **Average L** | 7.8325 | 7.8618 | 7.8920 | 7.9129 | 7.9378 | 7.9667 | 8.0029 | 8.0393 |
| **G** | **Coverage** | 94.38% | 95.23% | 94.91% | 94.84% | 94.9% | 95.31% | 95.04% | 94.8% |
| | **Average L** | 7.8329 | 7.8620 | 7.8945 | 7.9116 | 7.9370 | 7.9686 | 8.0015 | 8.0375 |

| | $\lambda$ | **28.9** | **29** | **29.1** | **29.2** | **29.3** | **29.4** | **29.5** | **29.6** |
|---|---|---|---|---|---|---|---|---|---|
| **J** | **Coverage** | 95.22% | 94.94% | 95.3% | 94.99% | 94.89% | 94.92% | 95.06% | 95.13% |
| | **Average L** | 8.0646 | 8.0862 | 8.1103 | 8.1377 | 8.1616 | 8.1943 | 8.2300 | 8.2550 |
| **G** | **Coverage** | 95.08% | 94.91% | 95.28% | 94.98% | 94.79% | 94.87% | 94.96% | 95.01% |
| | **Average L** | 8.0649 | 8.0863 | 8.1112 | 8.1375 | 8.1632 | 8.1959 | 8.2294 | 8.2559 |

| | $\lambda$ | **29.7** | **29.8** | **29.9** | **30** |
|---|---|---|---|---|---|
| **J** | **Coverage** | 95.12% | 94.5% | 95.02% | 94.39% |
| | **Average L** | 8.2723 | 8.3084 | 8.3260 | 8.3653 |
| **G** | **Coverage** | 95.07% | 94.57% | 94.85% | 94.27% |
| | **Average L** | 8.2715 | 8.3065 | 8.3268 | 8.3653 |

**Table 4.12:** Overall average coverage rate and mean interval lengths for the Rayleigh distribution, when $\lambda = 0.1 : 0.1 : 30$.

| | | **Overall Average** |
|---|---|---|
| **J** | **Coverage** | 94.90% |
| | **Average L** | 4.1958 |
| **G** | **Coverage** | 94.81% |
| | **Average L** | 4.1958 |

A similar result is noticeable on the tables for the Rayleigh distribution. We notice a good coverage probability performance for both prior distributions, and there is no sign of poor coverage - all the coverage probabilities are in the range [94.07%]. However, we see in Table 4.12 that on average, the Jeffreys prior produces better coverage rates than the general divergence prior. The average credibility interval lengths for both priors are approximately equal, and they increase as the value of lambda increases. On average, both priors prior produce about the same average credibility interval lengths.

Box-plots and some coverage plots for the two priors are shown in Figure 4.4 and Figure 4.5.

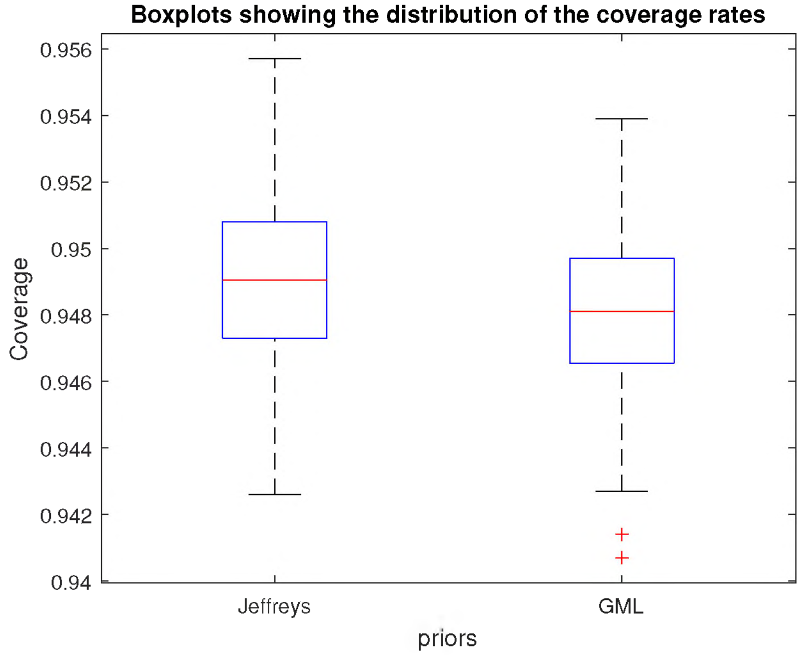


**Figure 4.4:** Box-plots for the Rayleigh distribution showing the distribution of the coverage rates.

Figure 4.4 shows the distribution of the coverage probabilities for the two priors. The Jeffreys prior has more than 50% of its coverage displaying under-coverage. Also, we see that the Jeffreys prior has a higher median coverage probability than the general divergence. The general divergence prior, has more than 75% of its coverage showing under-coverage. We notice a decrease in the coverage of both priors for the Rayleigh distribution compared to those of the exponential distribution.

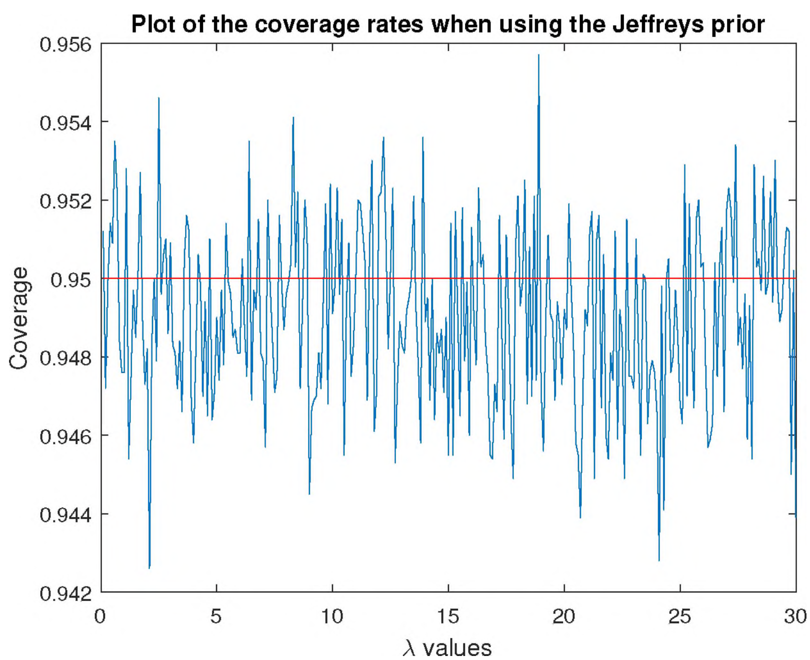


**Figure 4.5:** Coverage plots for the Rayleigh distribution using the Jeffreys prior.

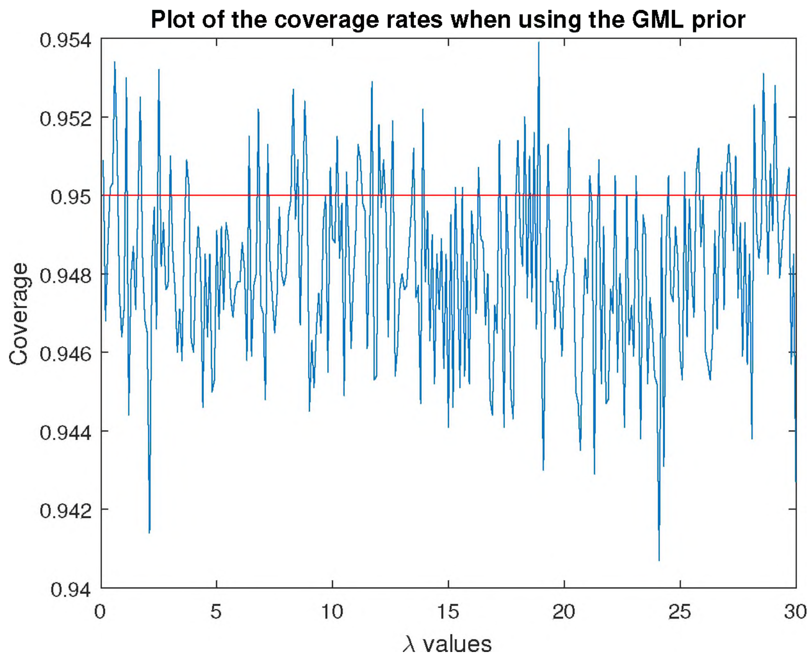


**Figure 4.6:** Coverage plots for the Rayleigh distribution using the general divergence prior.

In Figure 4.5 there is a decrease in the coverage performance for both priors. However, the Jeffreys prior shows a better coverage than the general divergence prior.

## 4.4 Metropolis-Hastings sampler

Markov Chain Monte Carlo (MCMC) methods are a general class of computational methods used to produce samples from posterior distributions. Since their introduction in the 1990s, they are said to have been successfully applied to literally thousands of applications. The main goal of an MCMC algorithm is to simulate values from the posterior distribution of a parameter vector. Inference about likely parameter values, or functions of parameter values, is then based on these simulated values, Hamada et al. (2008). One of the general categories of MCMC algorithm are the Metropolis-Hastings algorithms.

Metropolis-Hastings algorithms provide a simple, generic prescription for drawing from a posterior distribution when full conditional distributions for a parameter is not available. As noted in Hamada et al. (2008), the Metropolis-Hastings algorithm can be explained as follows:

1. Suppose that $\theta$ is a $q-$dimensional real-valued parameter vector and $\theta^{(j)}$ denote the $j$th component of $\theta$. It is a property of MCMC algorithms that the distribution of the $j$th iterate in the sequence of sampled values converges to a random sample drawn from the posterior distribution as $j$ becomes large. The first step is to generate a candidate point, $\theta^*$. A common method for generating the candidate point is to add a mean-zero normal deviate to a single component of $\theta^{(j-1)}$,

say $\theta_i^{(j-1)}$, to get $\theta_i^* = \theta_i^{(j-1)} + sZ$, where $Z$ is a standard normal deviate and $s$ is an arbitrary constant. For continuous-valued components of the parameter vector, let $f\left(\theta^* \mid \theta^{(j-1)}\right)$denote the proposal density used to generate $\theta^*$ from $\theta^{(j-1)}$. In theory, any density or mass function can serve as the proposal density as long as it satisfies three conditions:

- the proposal density must allow us to move from any subset of the parameter space to any other subset of the parameter space in a finite number of moves.
- the proposal density cannot be periodic - any moves to any subset of the parameter space can occur at any time.
- the rule used to specify the proposal density satisfies

$$0 < \frac{f\left(\theta^* \mid \theta^{(j-1)}\right)}{f\left(\theta^{(j-1)} \mid \theta^*\right)} < \infty,$$

  for all values of $\theta^{(j-1)}$and $\theta^*$.

2. The second step of the Metropolis-Hastings algorithm involves the calculation of the *acceptance probability*, denoted by $r$. This is the probability that the candidate value will be accepted as the next simulated value in the sequence, and is defined as

$$r = \min\left(1, \frac{p\left(\theta^* \mid \text{data}\right)}{p\left(\theta^{(j-1)} \mid \text{data}\right)} \frac{f\left(\theta^{(j-1)} \mid \theta^*\right)}{f\left(\theta^* \mid \theta^{(j-1)}\right)}\right).$$

   Note that if the proposal density is symmetric, then $f\left(\theta^* \mid \theta^{(j-1)}\right) = f\left(\theta^{(j-1)} \mid \theta^*\right)$, and

$$r = \min\left(1, \frac{p\left(\theta^* \mid \text{data}\right)}{p\left(\theta^{(j-1)} \mid \text{data}\right)}\right).$$

3. The third step of the Metropolis-Hastings algorithm involves a decision to either accept or reject the candidate point with probability equal to $r$. In order to reach this decision we first draw a $uniform(0,1)$ random variable, say $u$, and compare $u$ to $r$. If $r \geq u$, then accept the candidate value and set $\theta^{(j)} = \theta^*$. On the other hand, if $r < u$, then reject the candidate value and set $\theta^{(j)} = \theta^{(j-1)}$. This process is repeated for each component of $\theta$.

**Algorithm 4.1.** ***Metropolis-Hastings***

*For given* $\theta^{(j-1)}$,

*1. Generate* $\theta^* \sim f\left(\theta^* \mid \theta^{(j-1)}\right)$

2. *Calculate* $r = \min\left(1, \frac{p(\theta^*|data)}{p(\theta^{(j-1)}|data)} \frac{f(\theta^{(j-1)}|\theta^*)}{f(\theta^*|\theta^{(j-1)})}\right)$ *and draw* $u$ *from Uniform*$(0, 1)$
3. *Set*

$$\theta^{(j)} = \begin{cases} \theta^* & \textit{with probability} \quad r, \textit{ if } r \geq u \\ \theta^{(j-1)} & \textit{with probability} \quad 1-r, \textit{ if } r < u \end{cases}.$$

In general, successive draws from the posterior are correlated, but this correlation tends to die out as the interval between draws increases. Thus, the period during which the successive draws are correlated until the time the correlation has died out is known as the *burn-in* period; the iteration from the burn in period do not represent samples from the posterior distribution.

## 4.5 Simulation for the gamma distribution

This section looks at the simulation study where the life distribution is the gamma distribution. Following the general procedure given at the beginning of this chapter, the simulation for this distribution was done as follows:

1. Consider $\alpha = 1$ and $\lambda = 1$, and also $\alpha = 2$ and $\lambda = 1$. Other values for $\alpha$ and $\lambda$ can also be considered, but was not considered in this study, due to time constraints.

2. 1000 values of the unknown parameter, $\alpha$, were simulated from the posterior distributions:

$$p_J(\alpha \mid \underline{t}) \propto \frac{(\alpha\varphi'(\alpha) - 1)^{1/2} \left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n} \frac{\Gamma(\alpha n)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}$$

$$p_G(\alpha \mid \underline{t}) \propto \frac{(\alpha\varphi'(\alpha) - 1)^{1/4} \left[\prod_{i=1}^n t_i^\alpha\right]}{(\Gamma(\alpha))^n} \frac{\Gamma\left(\alpha n + \frac{1}{2}\right)}{\left(\sum_{i=1}^n t_i\right)^{\alpha n}}.$$

   The Metropolis-Hastings algorithm was used to do this, where the built-in function *'mhsample'* in MATLAB® was used. A gamma distribution was used as the proposal distribution.

3. 1000 values of the unknown parameter, $\lambda$, were simulated from the posterior distributions:

$$
\begin{aligned}
p_J(\lambda \mid \alpha, \underline{t}) &= \frac{\left(\sum_{i=1}^{n} t_i\right)^{\alpha n}}{\Gamma(\alpha n)} \lambda^{n\alpha - 1} e^{-\lambda \sum_{i=1}^{n} t_i} \\
\Rightarrow \lambda \mid \alpha, \underline{t} &\sim gamma\left(\alpha n, \sum_{i=1}^{n} t_i\right) \\
p_G(\lambda \mid \alpha, \underline{t}) &\propto \frac{\left(\sum_{i=1}^{n} t_i\right)^{\alpha n}}{\Gamma\left(\alpha n + \frac{1}{2}\right)} \lambda^{n\alpha - \frac{1}{2}} e^{-\lambda \sum_{i=1}^{n} t_i} \\
\Rightarrow \lambda \mid \alpha, \underline{t} &\sim gamma\left(\alpha n + \frac{1}{2}, \sum_{i=1}^{n} t_i\right).
\end{aligned}
$$

4. The resulting parameter values from posteriors are ordered in ascending order. The 95% credibility interval for $\alpha$ is then determined: $\left(\alpha_{(25)}, \alpha_{(975)}\right)$. The 95% credibility interval for $\lambda$ is then determined: $\left(\lambda_{(25)}, \lambda_{(975)}\right)$.

5. 1000 data sets were simulated, and Step 1 through Step 3 were repeated. The number of times a credibility interval in Step 3 contained the initial value of the unknown parameter was calculated. This number divided by 1000 gives the coverage probability for the prior distribution used.

The results are shown in Table 4.13. This simulation study is very limited. For $\lambda = 1$ and $\alpha = 1$, both priors give coverage below the nominal level. For $\alpha = 2$, the coverage obtained when using the divergence prior is the closest to the nominal level, and the interval is also shorter than that of the Jeffreys prior.

**Table 4.13:** Coverage rates and mean interval lengths for the gamma distribution, when $\lambda = 1$ and $\alpha = 1$, and $\lambda = 1$ and $\alpha = 2$ with sample size 50.

| | | $\alpha = 1$ | $\lambda = 1$ | $\alpha = 2$ | $\lambda = 1$ |
|---|---|---|---|---|---|
| **J** | **Coverage** | 90.60% | 94.80% | 96.40% | 96.40% |
| | **Average L** | 1.1193 | 0.5658 | 1.4857 | 0.3959 |
| **G** | **Coverage** | 92.10% | 94.90% | 95.40% | 93.80% |
| | **Average L** | 0.4733 | 0.5663 | 1.3373 | 0.3970 |

## 4.6 Simulation for the Weibull distribution

This section looks at the simulation study where the life distribution is the Weibull distribution. Following the general procedure given at the beginning of this chapter, the simulation for this distribution was done as follows:

1. Consider $\alpha = 1$ and $\lambda = 1$. Other values for $\alpha$ and $\lambda$can also be considered, but was not considered in this study, due to time constraints.

2. 1000 values of the unknown parameter, $\alpha$, were simulated from the posterior distributions:

$$p_J(\alpha \mid \underline{t}) \propto \alpha^{n-1}\left[\prod_{i=1}^{n} t_i^{\alpha}\right]\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{-n} \tag{4.1}$$

$$p_G(\alpha \mid \underline{t}) \propto \alpha^{n-1}\left[\prod_{i=1}^{n} t_i^{\alpha}\right]\frac{\Gamma\left(n+\frac{1}{2}\right)}{\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{n+\frac{1}{2}}}. \tag{4.2}$$

   The Metropolis-Hastings algorithm was used to do this, where the built-in function *'mhsample'* in MATLAB® was used. A gamma distribution was used as the proposal distribution. The equations (4.1) and (4.2) are not the same as those given in Section 3.3.6. The equations used here, are for the reparameterized Weibull, with density $f(t) = \alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}}$, this has been done, since this is the form used in MATLAB®.

3. 1000 values of the unknown parameter, $\lambda$, were simulated from the posterior distributions:

$$p_J(\lambda \mid \alpha, \underline{t}) = \frac{\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{n}}{\Gamma(n)}\lambda^{n-1}e^{-\lambda\sum_{i=1}^{n} t_i^{\alpha}} \tag{4.3}$$

$$\Rightarrow \lambda \mid \alpha, \underline{t} \sim gamma\left(n, \sum_{i=1}^{n} t_i^{\alpha}\right)$$

$$p_G(\lambda \mid \alpha, \underline{t}) \propto \frac{\left(\sum_{i=1}^{n} t_i^{\alpha}\right)^{n+\frac{1}{2}}}{\Gamma\left(n+\frac{1}{2}\right)}\lambda^{n-1/2}e^{-\lambda\sum_{i=1}^{n} t_i^{\alpha}}. \tag{4.4}$$

$$\Rightarrow \lambda \mid \alpha, \underline{t} \sim gamma\left(n+\frac{1}{2}, \sum_{i=1}^{n} t_i^{\alpha}\right).$$

   The equations (4.3) and (4.4) are not the same as those given in Section 3.3.6. The equations used here, are for the reparameterized Weibull, with density $f(t) = \alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}}$, this has been done, since this is the form used in MATLAB®.

4. The resulting parameter values from posteriors are ordered in ascending order. The 95% credibility interval for $\alpha$ is then determined: $\left(\alpha_{(25)}, \alpha_{(975)}\right)$. The 95% credibility interval for $\lambda$ is then determined: $\left(\lambda_{(25)}, \lambda_{(975)}\right)$.

5. 1000 data sets were simulated, and Step 1 through Step 3 were repeated. The number of times a credibility interval in Step 3 contained the initial value of the unknown parameter was calculated. This number divided by 1000 gives the coverage probability for the prior distribution used.

The results are shown in Table 4.14. This simulation study is very limited. For $\lambda = 1$ and $\alpha = 1$, both priors give coverage below the nominal level for $\alpha$. For this limited study, the general divergence prior performed better.

**Table 4.14:** Coverage rates and mean interval lengths for the Weibull distribution, when $\lambda = 1$ and $\alpha = 1$, with sample size 50.

| | | $\alpha$ | $\lambda$ |
|---|---|---|---|
| **J** | **Coverage** | 89.60% | 94.90% |
| | **Average L** | 0.7095 | 0.5637 |
| **G** | **Coverage** | 90.40% | 95.10% |
| | **Average L** | 0.7095 | 0.5653 |

# Chapter 5

# Applications using the Gamma and Weibull Distributions

## 5.1 Introduction

In this chapter the gamma and the Weibull distributions will be applied to a lifetime data set. The Jeffreys prior and the general divergence priors will be used in the analysis. The outcome of the analysis will then be used to compare the performance of these priors, and which distribution is the preferred one. In this section an example from Lawless (1982) will be considered, this example was also considered in Moala et al. (2013) and is originally from Nelson (1972). In this life test experiment specimens of a type of electrical insulating fluid were subjected to a constant voltage stress. The length of life until each specimen failed, was observed. Nelson (1972) considered seven groups of specimens, which were tested ate voltages ranging from 26 kV to 38 kV. As in Moala et al. (2013), the specimen of size 19 at voltage level 34 kV will only be considered. The experiment was run long enough to observe the failure of all the insulation specimens tested, there are thus no censored values present. The values given in Table 5.1 represent the lifetime (in minutes) to failure.

**Table 5.1:** Times to failure (in minutes) at voltage level 34 kV.

| | | | | | |
|---|---|---|---|---|---|
| 0.96 | 4.15 | 0.19 | 0.78 | 8.01 | 31.75 |
| 7.35 | 6.50 | 8.27 | 33.91 | 32.52 | 16.03 |
| 4.85 | 2.78 | 4.67 | 1.31 | 12.06 | 36.71 |
| 72.89 | | | | | |

## 5.2 Gamma distribution Application

Due to computational challenges in the posterior, OpenBugs® will be used. The gamma density given in Chapter 3 is in the same form as the gamma density used in OpenBugs®, i.e. if $T \sim gamma(\alpha, \lambda)$, the density function is

$$f(t) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} \text{ for } t > 0.$$

From Chapter 3, the Jeffreys prior was given as

$$p_J(\alpha, \lambda) \propto \frac{\sqrt{\alpha\varphi'(\alpha) - 1}}{\lambda}. \tag{5.1}$$

The prior given in (5.1) can be written as

$$\begin{aligned} p_J(\alpha, \lambda) &\propto \left(\alpha\varphi'(\alpha) - 1\right)^{\frac{1}{2}} \lambda^{-1} \\ &= p_J(\alpha) p_J(\lambda). \end{aligned}$$

Since OpenBugs® requires that a full probability model is defined and all prior distribution should be proper, we will approximate $p_J(\lambda)$ by a $gamma(0.0001, 0.0001)$ and $p_J(\alpha)$ by a $uniform(0, 16)$.

From Chapter 3, the divergence prior was given as

$$p_G(\alpha, \lambda) \propto \frac{\left(\alpha\varphi'(\alpha) - 1\right)^{1/4}}{\sqrt{\lambda}}. \tag{5.2}$$

The prior given in (5.2) can be written as

$$\begin{aligned} p_G(\alpha, \lambda) &\propto \left(\alpha\varphi'(\alpha) - 1\right)^{\frac{1}{4}} \lambda^{-\frac{1}{2}} \\ &= p_G(\alpha) p_G(\lambda). \end{aligned}$$

Since OpenBugs® requires that a full probability model is defined and all prior distribution should be proper, we will approximate $p_G(\lambda)$ by a $gamma(0.5, 0.0001)$ and $p_G(\alpha)$ by a $uniform(0, 4)$.

**Table 5.2:** Statistics on the posteriors for the parameters $\alpha$ and $\lambda$.

| | Jeffreys prior | | | | Divergence prior | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Median | MC error | Mean | SD | Median | MC error |
| $\alpha$ | 0.7491 | 0.2054 | 0.7288 | 0.001127 | 0.7769 | 0.2084 | 0.757 | 0.001179 |
| $\lambda$ | 0.0498 | 0.0191 | 0.0476 | 1.051E-4 | 0.0534 | 0.0195 | 0.0512 | 1.107E-4 |

In Table 5.2, statistics on the posteriors for the two parameters are shown. Both priors produce approximately equal statistics values for the parameters. The MC error quantifies the variability in the estimates that is due to Markov chain variability. It is very small for both priors, as desired; it is much smaller than the SD ($< 5\%$ of SD). The MC errors are lower than the SDs - the estimated posterior means were estimated with high precision.

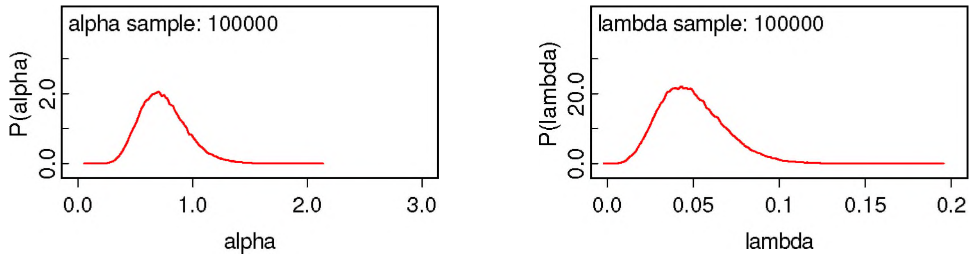


**Figure 5.1:** Posterior density plots of $\alpha$ and $\lambda$ using the Jeffreys prior for the gamma application.

The posterior density estimates plots for $\alpha$ and $\lambda$ are shown in Figure 5.1, when using the Jeffreys prior. The posterior density estimate for $\lambda$ and $\alpha$ both appear to be positively skewed.

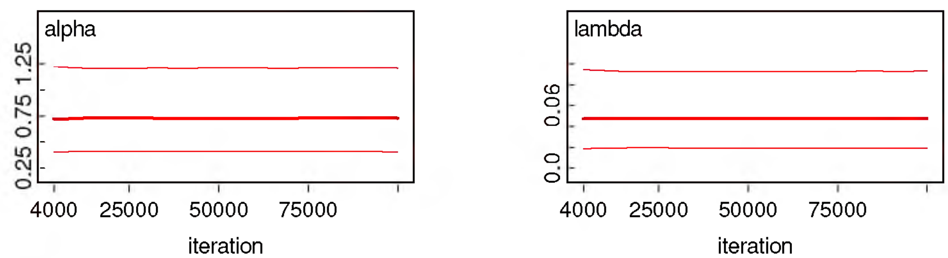


**Figure 5.2:** Quantile plots of $\alpha$ and $\lambda$ using the Jeffreys prior for the gamma application.

In Figure 5.2, quantile plots of the two parameters using the Jefrreys prior for the gamma distribution are shown. The quantiles have stabilized as the iterations increase, indicating convergence of the algorithm, in terms of the two parameters, $\alpha$ and $\lambda$.

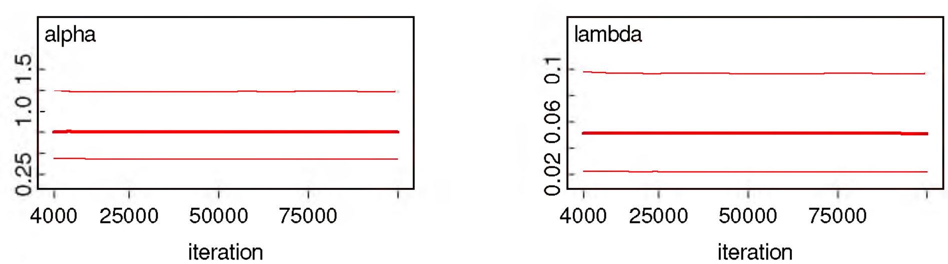


**Figure 5.3:** Quantile plots of $\alpha$ and $\lambda$ using the divergence prior for the gamma application.

Figure 5.3 shows quantile plots of the two parameters using the general divergence prior for the gamma distribution application. Since there is a stabilized pattern displayed by these quantiles, the algorithm has converged for the two parameters.

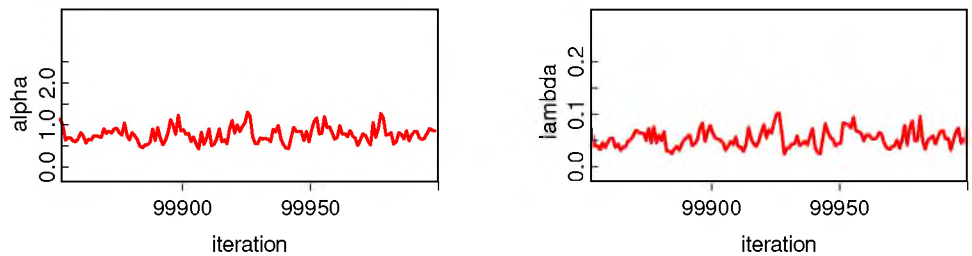


**Figure 5.4:** Trace plots when using the Jeffreys prior for the gamma application.

Figure 5.4 gives dynamic trace plots when using the Jeffreys prior for the gamma distribution application. It shows the successive samples from the algorithm that are joined together. There are no flat patterns, which would indicate multiple successive rejections.

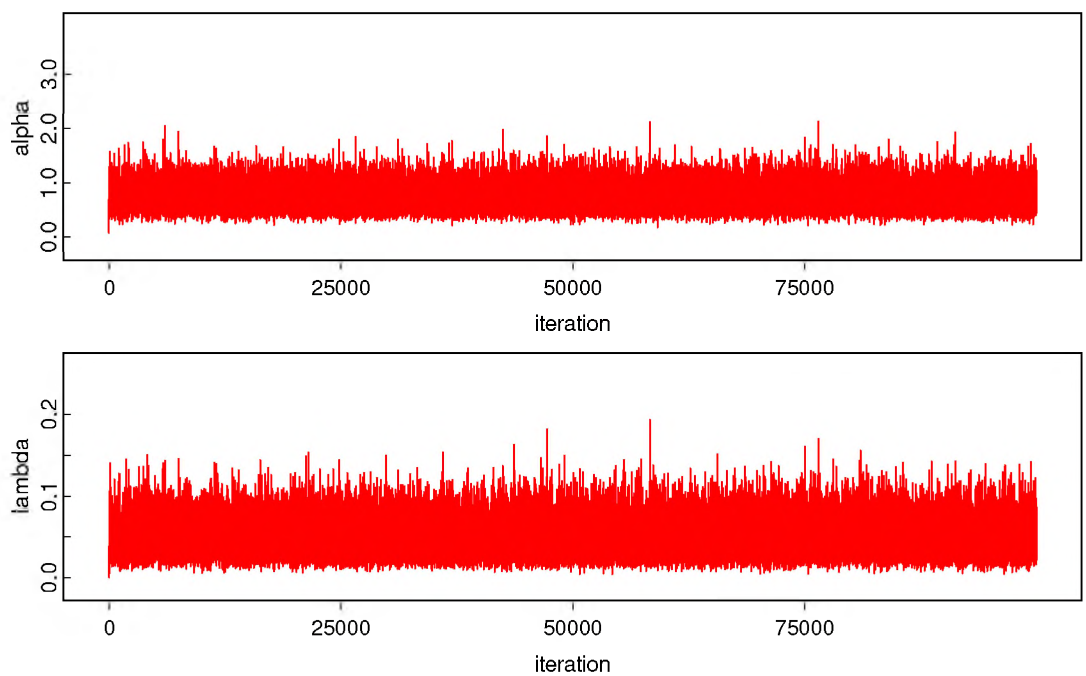


**Figure 5.5:** History plots when using the Jeffreys prior for the gamma application.

Figure 5.5 shows time series plots or history plots for the gamma application when using the Jeffreys prior. Since there are no irregularities present, these plots suggest that the Markov chains have converged for both priors.

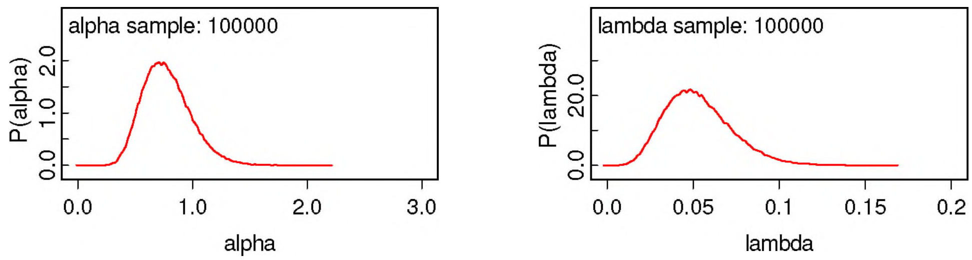


**Figure 5.6:** Posterior density plots of $\alpha$ and $\lambda$ using the divergence prior for the gamma application.

Posterior density estimate plots of $\alpha$ and $\lambda$ using the general divergence prior for the gamma application are shown in5.6. The posterior density estimate for $\alpha$ and $\lambda$ look positively skewed.

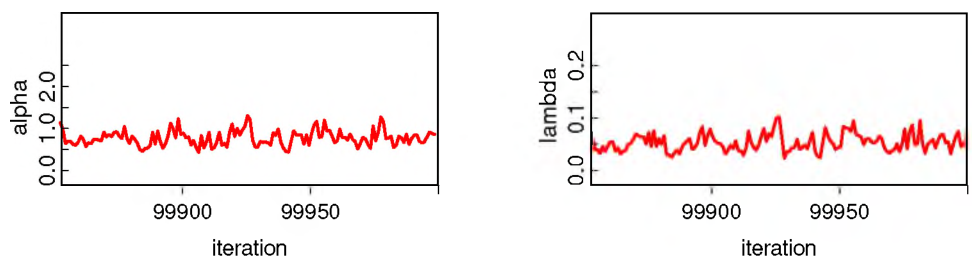


**Figure 5.7:** Trace plots when using the divergence prior for the gamma application.

Dynamic trace plots for $\alpha$ and $\lambda$ when using the general divergence prior are shown in Figure 5.7. Lack of flat patterns in both plots indicates that there are no multiple successive rejections.

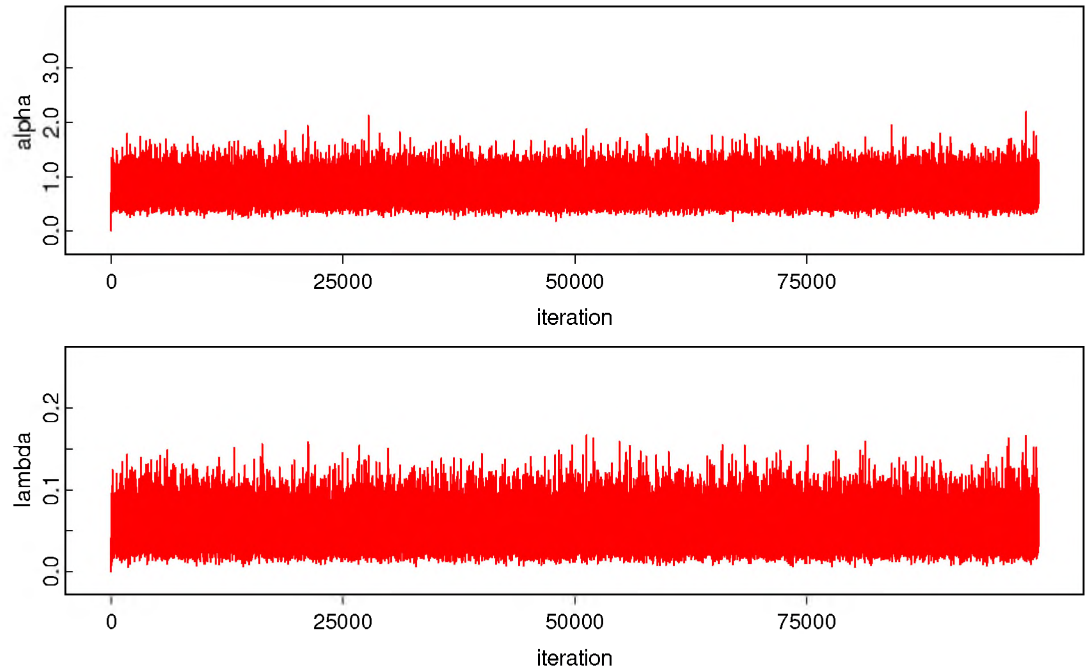


**Figure 5.8:** History plots when using the divergence prior for the gamma application.

History plots of the parameters are shown in Figure 5.8, when the general divergence prior is used for the gamma application. There are no signs of irregularities, indicating that the Markov chain has converged.

### Deviance Information Criterion (DIC)

The deviance information criterion (DIC) is a statistic used for model comparison. Spiegelhalter et al. (2002) introduced it as a means to compare models of arbitrary structure (complex hierarchical models). The construction of the DIC involves two two parameters: $p_D$ and $\overline{D(\theta)}$.

Spiegelhalter et al. (2002) defines $p_D$ as the complexity measure for the number of effective parameters in a model, and is given as

$$p_D = \overline{D(\theta)} - D\left(\bar{\theta}\right). \tag{5.3}$$

The first term of the right hand side of (5.3) is known as the posterior mean deviance, and the second term is called the deviance of the means. Also, for the maximised likelihood value over the unknown parameters, $L(\theta \mid data)$, and some function of the data alone, $f(data)$,

$$D(\theta) = -2\log(L(\theta \mid data)) + 2\log(f(data))$$

is known as the Bayesian deviance. As a result,

$$\begin{aligned}\overline{D(\theta)} &= E[D(\theta)] \\ &= E[-2\log(L(\theta \mid data)) + 2\log(f(data))]\end{aligned}$$

and

$$D(\bar{\theta}) = -2\log(L(E[\theta \mid data])) + 2\log(f(data)).$$

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

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

and from (5.3), the following holds

$$DIC = 2p_D + D(\bar{\theta}).$$

The model with smallest DIC value is the desired model.

**Table 5.3:** Model comparison.

| | DIC | $p_D$ | $\bar{D}$ | $\hat{D}$ |
|---|---|---|---|---|
| Jeffreys prior | 143.4 | 1.971 | 141.4 | 139.5 |
| Divergence prior | 143.4 | 1.907 | 141.4 | 139.5 |

Table 5.3 shows the plug-in deviance, $\hat{D}$, the mean posterior deviance, $\bar{D}$, the measure of the number of effective parameters, $p_D$, and the deviance information criterion (DIC) using the gamma distribution. These are used for the purpose of model comparison. The $\hat{D}$ and $\bar{D}$ values are the equal for both prior distributions, indicating equal measure of fit for the two models produced by these priors. As a result, $p_D$ has approximately equal values. Also, and $p_D \approx 2$, the number of parameters, implying that both parameters are effective for both priors. With a difference of zero in the DIC values, it safe to say that both priors result in models that can make good short-term predictions.

## 5.3 Weibull distribution Application

In this section we will consider two applications.

Due to computational challenges in the posterior, OpenBugs® will be used. The Weibull density given in Chapter 3 is not in the same form as the Weibull density used in OpenBugs®, i.e. $T \sim Weibull(\alpha, \lambda)$, with positive parameters $\alpha$ and $\lambda$ if the probability density function is given by

$$f(t) = \alpha\lambda(\lambda t)^{\alpha-1} e^{-(\lambda t)^{\alpha}} \text{ for } t > 0.$$

In OpenBugs® the following form of the Weibull distribution is used

$$f(t) = \alpha\lambda t^{\alpha-1}e^{-\lambda t^{\alpha}} \text{ for } t > 0. \tag{5.4}$$

The Jeffreys prior for the reparameterized Weibull distribution for $(\alpha, \lambda)$ is given by

$$p_J(\alpha, \lambda) \propto \frac{1}{\alpha\lambda}. \tag{5.5}$$

The prior given in (5.5) can be written as

$$\begin{aligned} p_J(\alpha, \lambda) &\propto \alpha^{-1}\lambda^{-1} \\ &= p_J(\alpha)\, p_J(\lambda). \end{aligned}$$

Since OpenBugs® requires that a full probability model is defined and all prior distribution should be proper, we will approximate $p_J(\lambda)$ by a *gamma* $(0.0001, 0.0001)$ and $p_J(\alpha)$ by a *gamma* $(0.0001, 0.0001)$.

The divergence prior for the reparameterized Weibull distribution for $(\alpha, \lambda)$ is given by

$$p_G(\alpha, \lambda) \propto \frac{1}{\sqrt{\alpha\lambda}}. \tag{5.6}$$

The prior given in (5.6) can be written as

$$\begin{aligned} p_G(\alpha, \lambda) &\propto \alpha^{-\frac{1}{2}}\lambda^{-\frac{1}{2}} \\ &= p_G(\alpha)\, p_G(\lambda). \end{aligned}$$

Since OpenBugs® requires that a full probability model is defined and all prior distribution should be proper, we will approximate $p_G(\lambda)$ by a *gamma* $(0.5, 0.0001)$ and $p_G(\alpha)$ by a *gamma* $(0.5, 0.0001)$.

The same data as in Section 5.2 will be considered in this section.

**Table 5.4:** Statistics on the posteriors for the parameters $\alpha$ and $\lambda$.

| | Jeffreys prior | | | | Divergence prior | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Median | MC error | Mean | SD | Median | MC error |
| $\alpha$ | 0.7895 | 0.1415 | 0.7833 | 0.001271 | 0.7711 | 0.1366 | 0.7655 | 0.001255 |
| $\lambda$ | 0.1421 | 0.06978 | 0.1296 | 5.901E-4 | 0.1536 | 0.0731 | 0.1407 | 6.316E-4 |

As can be seen in Table 5.4, the mean, median, standard deviation (SD), and the Monte Carlo (MC) error on the posteriors for the two parameters are approximately equal. However, MC error which measures the variation of the mean of the parameters due to the simulation, is lower than the SD for both priors; this indicates that the estimated posterior means were estimated with high precision.

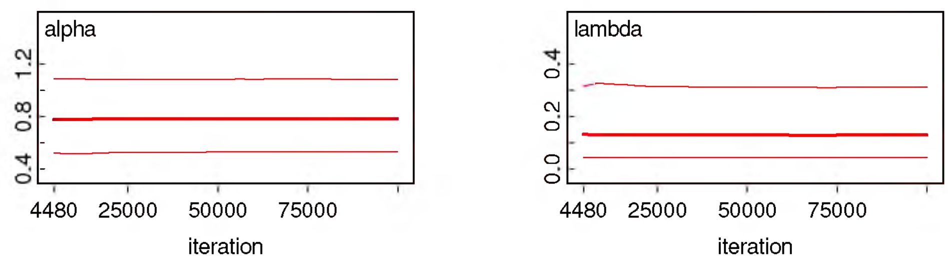


**Figure 5.9:** Quantile plots of $\alpha$ and $\lambda$ using the Jeffreys prior for the Weibull application.

Figure 5.9 shows the quantile plots of the parameters when using the Jeffreys prior, for the Weibull distribution. This figure indicates that the quantiles have stabilized, implying that the algorithm has converged, in terms of $\alpha$ and $\lambda$.

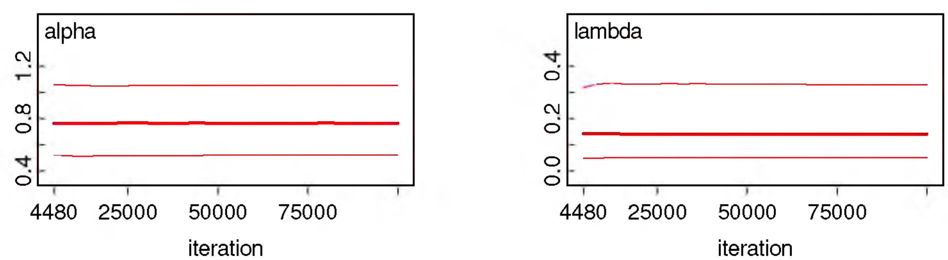


**Figure 5.10:** Quantile plots of $\alpha$ and $\lambda$ using the divergence prior for the Weibull application.

In Figure 5.10 the quantile plots of both parameters obtained from using the general divergence prior, for the Weibull distribution, show a stabilized pattern. This indicates convergence for the two parameters.

**Table 5.5:** Model comparison.

| | DIC | $p_D$ | $\bar{D}$ | $\hat{D}$ |
|---|---|---|---|---|
| Jeffreys prior | 143.1 | 1.880 | 141.2 | 139.3 |
| Divergence prior | 143.0 | 1.765 | 141.2 | 139.4 |

In Table 5.5 the candidates for model comparison - the plug-in deviance, $\hat{D}$, the mean posterior deviance, $\bar{D}$, the measure of the number of effective parameters, $p_D$, and the deviance information criterion (DIC) are shown, using the Weibull distribution. Both $\hat{D}$ and $\bar{D}$ are approximately equal for both prior distributions. Consequently, $p_D$ is approximately equal for the two priors. Also, $p_D \approx 2$, the number of parameters. This is not surprising because we used non-informative priors. The difference of 0.1 for the two DIC values indicates that both priors produce models that can be expected to make the best short-term predictions.

The predicted reliability function is defined as

$$\begin{aligned} R(t) &= P(T > t|Data) \\ &= \int_0^\infty \int_0^\infty P(T > t|\alpha,\lambda)p(\alpha,\lambda|Data)d\alpha d\lambda \\ &= \int_0^\infty \int_0^\infty exp\left[-(\lambda t)^\alpha\right]p(\alpha,\lambda|Data)d\alpha d\lambda \\ &= \frac{1}{N}\sum_{k=1}^{N} exp\left[-\left(\lambda^{(k)}t\right)^{\alpha^{(k)}}\right]. \end{aligned} \tag{5.7}$$

Using (5.7), the predictive reliability when the Jeffreys prior is used, is

$$R_J(t) = 0.0230$$

and the predictive reliability when using the general divergence prior, is

$$R_G(t) = 0.0195.$$

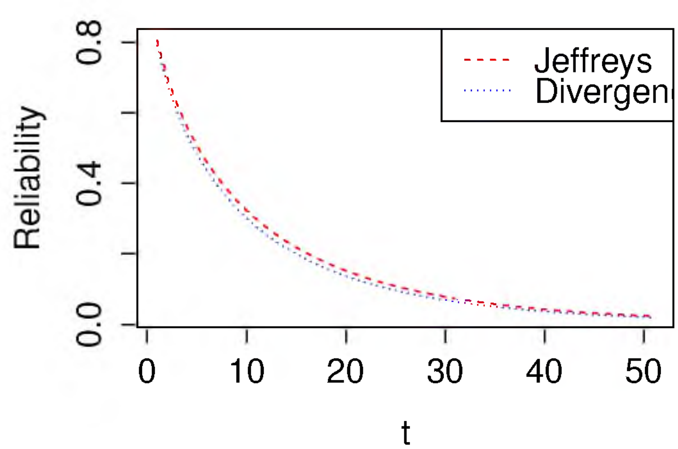


**Figure 5.11:** Reliability plot for the Jeffreys and divergence priors.

The reliability curves or plots generated using the Jeffreys prior and the divergence prior for the lifetime of the specimens is shown in Figure 5.11. The Jeffreys prior models the reliability of the specimens lifetimes better than the divergence prior. This is noticeable for the time period $t = 6$ to $t = 30$.

## 5.4 Weibull vs Gamma

In this application both the Weibull and gamma distributions were used. By looking at the DIC values from the posterior distributions in Table 5.3, for the gamma, and Table 5.5, for the Weibull, the smallest DIC value is obtained when using the Weibull distribution and the general divergence prior. By just looking at the data, the following question may arise: Which distribution is more appropriate, the gamma or the Weibull? Distribution fitting in R® was done, by using the *'fitdistrplus'* package. The Cullen and Frey graph, from Cullen & Frey (1999), is a skewness-kurtosis graph for the choice of distributions, and is given in Figure 5.12.

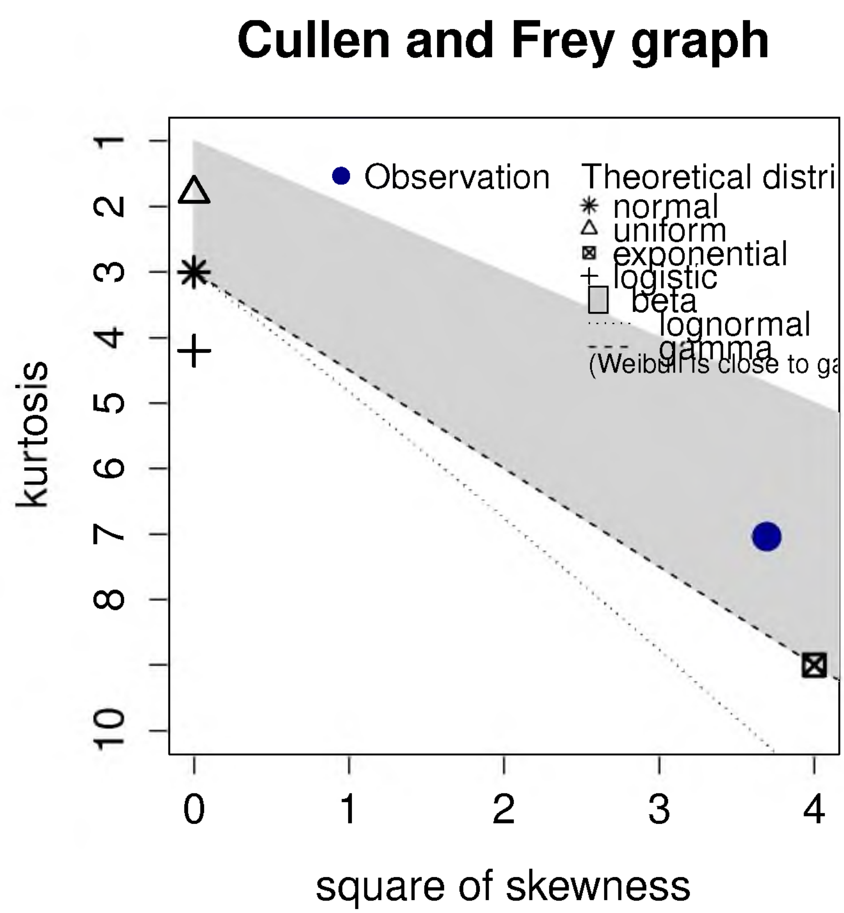


**Figure 5.12:** Cullen and Frey graph for insulating fluid data.

From Figure 5.12 it seems that the Weibull and gamma distributions are possible distributions for the data. The kurtosis and squared skewness of the sample is plotted as the blue point in the graph.

Figure 5.13 illustrates the plot of fitting the gamma distribution to the data, and Figure 5.14 gives the plot of fitting the Weibull distribution to the data.

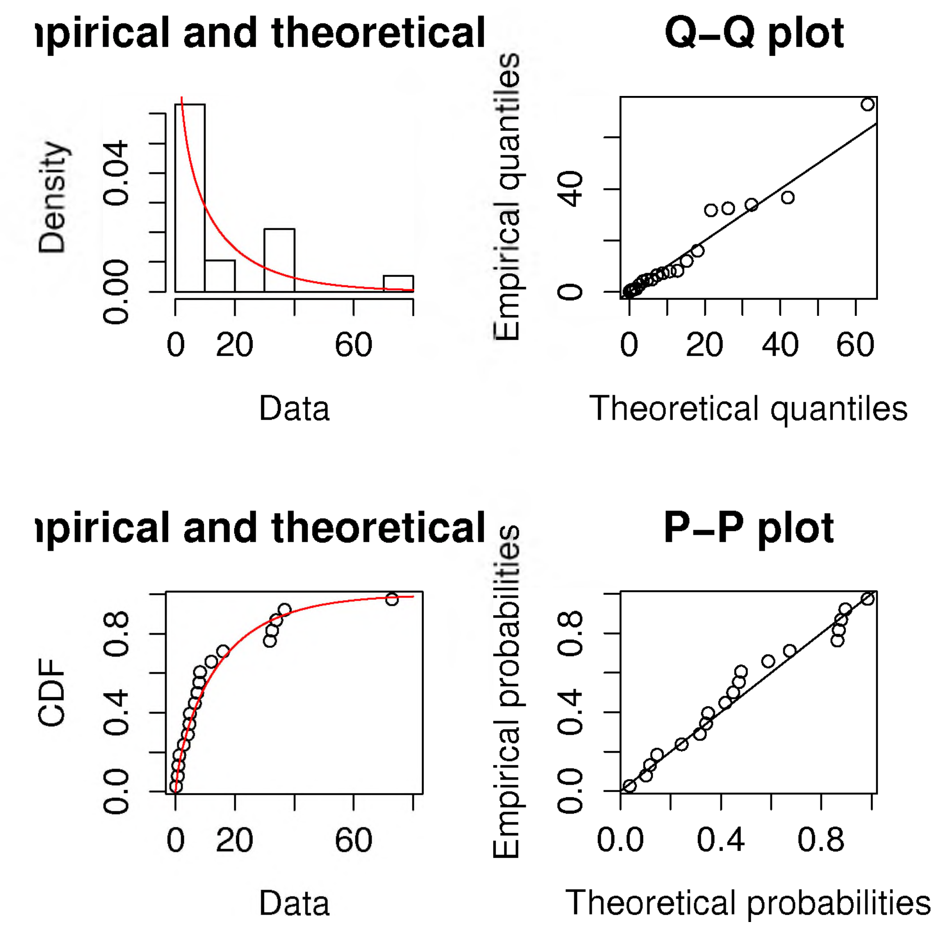


**Figure 5.13:** Fit of the gamma distribution to the data.

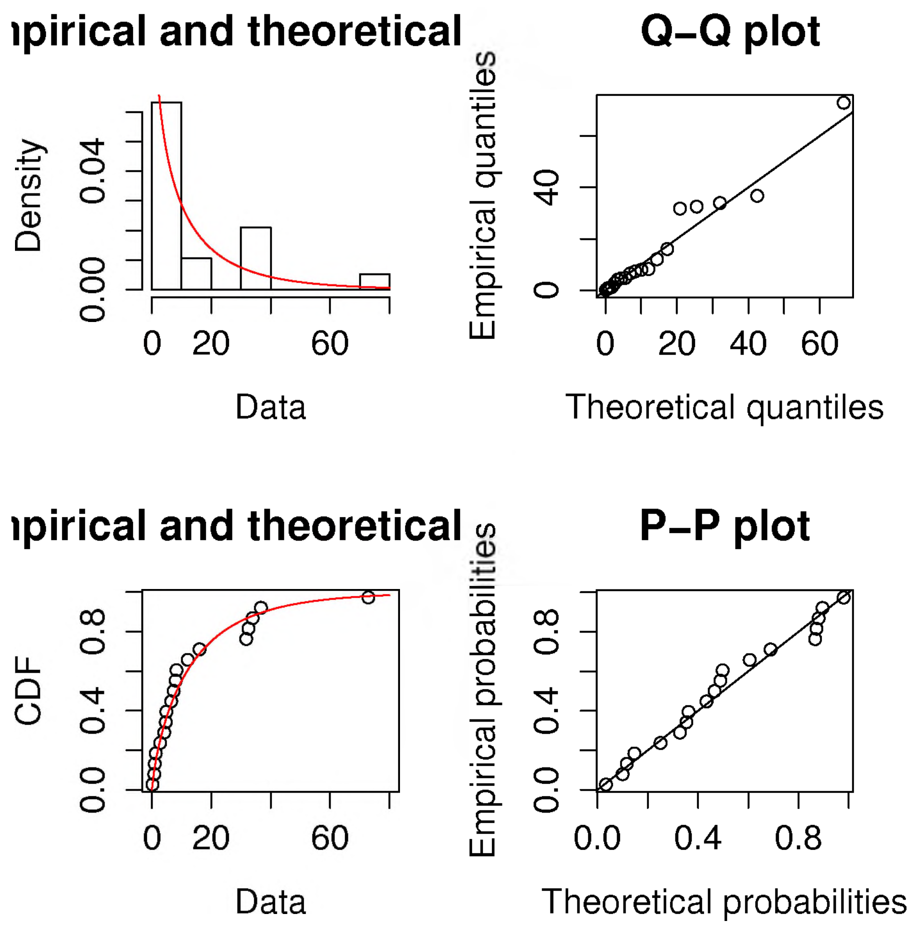


**Figure 5.14:** Fit of the Weibull distribution to the data.

Both the fits look good and similar, looking at the QQ-plot, the Weibull looks a bit better at the tails. The Akaike information criterion (AIC), from Akaike (1974), is well-known and is a measure of quality of statistical models for a data set. AIC is useful for model selection The AIC for the Weibull is slightly lower than that of the gamma. Where the AIC for the Weibull is 143.2018 and the AIC for the gamma is 143.4298.

# Chapter 6

# Case Studies

In this chapter we will first consider a number of HRA case studies as discussed in French et al. (2009), each focusing on the relevant developments in behavioural, cognitive, management and organizational sciences. Secondly, we will also show, whenever possible, the various HRA tools, methods and considerations in action in a range of industries (chemical, offshore, nuclear-power and chemical, and marine transport) as carried out in Kirwan (1994). In the second step, a number of HRA techniques are demonstrated through the case studies, and each case study focuses on certain aspects of the HRA process.

## 6.1 Nuclear

The nuclear industry offers some of the best recorded incidents in which human error either led to a system failure or had the potential to do so. As an example, we note briefly some events that indicate significant human potential for errors of commission or omission.

### 6.1.1 Chernobyl

The Chernobyl accident occurred on 26 April 1986 and involved an enormous release of radioactivity, International Atomic Energy Agency (1991). The accident itself was caused by a deliberate act, an experiment that caused an explosion and fire. The experiment involved running the plant outside its design parameters at very low power. The staff of the Chernobyl station were familiar with the experiment because they carried it out previously at the Chernobyl-3 reactor and the Kursk station in Russia, Marples (1997). However, the personnel responsible for the experiment had been working for some fifteen to twenty hours when the experiment started because of delays in handing the reactor over for the experiment. They were tired and under pressure to complete the experiment fast and return the reactor to normal energy production.

What must be noted here is that during the experiment, they deliberately turned off three separate safety systems and switched to manual control. This was against safety instructions, but probably they had done it frequently before. In this case, human failure on on-line operations was caused by the effect that seemingly 'freak infringement rules' which did not cause an accident in the past lead to more violation of rules in the future, Dorner (1996). Apparently the accident occurred without any component failure. However, the design of the reactor depended on the operators following certain safety instructions. As noted, the operating and regulatory regime in place was inadequate. Attempts by operators to recover the situation triggered flash boiling of water, which in turn resulted to a breach of fuel can or containment, and the exposure of the hot fuel element to water. Within seconds, a major chemical explosion occurred which destroyed the reactor and caused the worst nuclear catastrophe in history. While there is no doubt that deliberate acts were key in causing the accident, there is equally no doubt that the tiredness, stress and poor safety culture within the operating team were contributing factors.

### 6.1.2 Three Mile Island

The accident happened on 28 March 1979, at the Three Mile Island nuclear power plant near Harrisburg, Pennsylvania, Kemeny et al. (1979). Apparently there was no significant release of radiation, but a full core melt was only just averted. The causes of the accident continue to be debated to this day, but one thing is clear. The initiating event, the formation of the hydrogen bubble which forced down cooling water exposing the core, had not been anticipated in the reactor's design or safety studies.

The operators not only did recognize what was happening, but also had never anticipated it might. It was an incident beyond their experience and imagination: in a very sense outside of scientific and engineering knowledge as it stood then. A key learning point in relation to HRA is that the operators behaved entirely sensibly and in accordance with their mental models of what they believed was happening. There was no error in their behaviour in this respect, not at least in the sense of human error of HRA theory.

### 6.1.3 Sellafield pigeons

In 1998 it was discovered that pigeons were transferring radioactivity from the Sellafield nuclear site to the surrounding region. This potential route for contamination to be taken off-site had not been anticipated, despite many risk analyses at the plant over the years. A recent leak at Sellafield is particularly relevant because human and organizational behaviours interacted and led to the incident. We therefore give a thorough description of this incident.

On 20 April 2005 a leak was detected in the Sellafield Thermal Oxide Reprocessing Plant (THORP) after a video camera revealed that approximately eighty three thousands of liters of radioactive waste, or dissolver liquid had leaked into the base of the cell. Apparently a feed pipe to accountancy vessel V2217B had fractured. This is believed to have been due to fatigue stress induced by excessive movement of the vessel to which the pipe is connected. It is estimated that the pipe suffered major failure around the $15^{th}$ of January, 2005, but may have started to leak as early as July 2004.

Subsequent investigations indicated the following errors and behaviours all contributed to the leak itself and a failure to detect it sooner.

- The vessel and pipework design was changed during the construction of the plant. The original design may well not have fractured in the same way. However, during the preparation of the safety case for the plant , assumptions and precautions in relation to seismic activity were changed and the design changed to allow for this. So a design fault that had originally been engineered out was re-introduced in another phase of the design process.

- There were earlier indications that there might be a leak from accounting of the input and output to and from the cell. However, these were ignored because of a 'new plant' culture. Apparently there was a belief that such a modern plant could not suffer from leaks or other failures. Some of the written instructions were ambiguous, leaving too much to the interpretation of the operating staff. In the context of the 'new plant' culture and other management imperatives, it was too easy to ignore inconclusive but pertinent readings and observations. It is also noteworthy that this 'new plant' culture was implicated in two previous smaller incident elsewhere in THORP in 1998 and earlier in 2005.

- Misunderstandings among senior management led to operations being continued to meet production targets longer than they should have been, despite a decision that was taken to investigate the leak.

The point that the new plant culture was at the heart of many of the failings of the operators, management processes and organization is repeatedly made in Snelson (2005).

### 6.1.4 Doonreay Shaft

A 65 meters deep shaft at the Doonreay nuclear power plant, originally dug t remove rock from a pipe discharging treated water was used from 1958 to 1977 as a low level waste pit.In 1977 there was a major explosion in the shaft caused by the reaction of sodium potassium (NaK) alloy with water. Some radioactive waste was spread over a large area.

One can wonder about many things to do with the management of this pit, but one thing is clearly incredible: the deposition of NaK, although encased in cast iron, into the pit. The scientists and engineers involved would surely have known that: the shaft was wet - it linked to the discharge pipe and went below sea-level; NaK reacts explosively with water; and cast iron corrodes. Put these facts together and see that this was nothing but a recipe for disaster - the 'accident' was completely predictable.

### 6.1.5 Emergency Shutdown (ESD) Scenario

#### 6.1.5.1 Background

This assessment was for a nuclear chemical plant still at the detailed design stage. The plant had a comprehensive crash-shutdown system which was manually initiated from the central control room (CCR), or from a diverse backup location (in the event of a fire etc., in the central control room itself). However, a rare-event scenario was identified in which the shutdown system could only be partially effective. In this unlikely situation, operators would have to detect the partial failure and then manually locate and rectify the problem by determining which automatic ESD valve had failed to close and the sending an operator to close a manual, back-up ESD valve on the same pipeline.

This scenario was evaluated in two ways, firstly by the use of the technique THERP and secondly by using two independent APJ groups, operating in a group-discussion mode, whose individual estimates were aggregated into a geometric mean value. This case study therefore usefully demonstrates two independent methods for assessing the same system via two independent APJ teams and an independent THERP assessor.

#### 6.1.5.2 Problem definition

The question being posed by the study was a purely quantitate one: how reliable would the operators be in a partial-ESD-failure situation? The operators would have two hours to resolve the situation. No error-reduction analysis (ERA) was required, since if the results were found to be unacceptable, a design solution would be implemented to increase the reliability of the ESD system itself (which, it should be noted, was already highly reliable). Since the plant was at the detailed design stage, no operators were available and there was no similar plant in existence. However, operators from the intended site were available to participate in the study.

#### 6.1.5.3 Task analysis

The task analysis itself proceeded by researching the relevant documentation: the ESD-system description, the review of the ESD panel-instrumentation drawings, CCR layout and staffing plans, local panel-instrumentation plans (not at as an advanced design level as CCR) and general plant layout. It is noteworthy that the plant in question was in fact a very large plant with an integrated modular design, and as a result was therefore divided into, and controlled by means of, individual plant units. Once these details had been reviewed, discussions were held with ESD design personnel, safety-assessment personnel and future operations personnel as what the correct sequence of actions was to be.

#### 6.1.5.4 Human-error analysis

The following human errors were identified directly from the task analysis as requiring quantification:

- HEP 1.1: operator fails to initiate ESD within 20 minutes.
- HEP 1.2: given 1.1, supervisor fails to initiate ESD success within 20 minutes.
- HEP 1.3: operator fails to detect only a partial ESD success within 2 hours.
- HEP1.4: given 1.3, supervisor fails to detect only a partial ESD (from the operator's panel) within 2 hours.
- HEP1.5: operator fails both to identify appropriate equipment room and to communicate location of this room to an outside operator.
- HEP 1.6: outside operator fails to get to correct ESD actuator panel, to identify failed actuator(s) and to communicate this information to the control-room operator.
- HEP1.7: control room operator fails both to determine which valves must be closed, in order to achieve a total ESD, and to communicate this information to outside operator.
- HEP1.8: outside operator fails to close correct valves within 2 hours distributed control unit (DCU) failure/DCU-power failure.

#### 6.1.5.5 Representation and Quantification

An event tree was constructed in order to represent the scenario. The HEPs were quantified using the THERP and APJ approaches.

#### 6.1.5.6 Impact

The results from all the assessments made (in fact, the HEART and SLIM methods were also used to calculate HEPs for this scenario) suggested that the overall probability of failure was between 0,03 and 0.3. A further reliability analysis was undertaken to evaluate in more detail the likelihood of a partial failure and its potential consequences. From this further work, it was concluded that, given the the low likelihood of the initiating event, on the one hand, and the high reliability of the system on the other, and in the light of a consideration of any potential consequences, the system hardware did not need to be changed.

However, due to the relatively high values of the HEPs the ESD panels were, however, rendered more ergonomic, particularly with respect to the signals denoting whether the ESD had been effective or not. This redesign was extensive, affected every ESD panel in the CCR and backup ESD room and involved a hybrid team of human-factors personnel, designers and operators.

## 6.2 Railway

The case studies on Saxmundham Collision and Lambrigg Derailment are from French et al. (2009). The ramioning case studies on The Flixborough disaster and Ekofisk Bravo blowout are given in Kirwan (1994).

### 6.2.1 Saxmundham Collision (User worked crossing collision)

On $22^{nd}$ May 2006, a freight train collided with a car trying to cross at a User Worked Crossing (UWC) near Saxmundham, England. No one was injured as a result of the collision and the train was not derailed, however both vehicles did suffer minor damages at the area of impact. The concerned UWC is situated on a private road which from the north to south side, leads to private dwellings. Permitted users of the UWC include residents of the dwellings, farmers of the surrounding farmlands and users with authorized access from the residents including such parties as delivery vehicles.

The immediate cause of the incident was reported to be the fault of the driver of the vehicle, who failed to stop at the check point to observe for oncoming trains. In the incident report it was further disclosed that the authorized person who permitted the motorist to use the UWC did not give the motorist a briefing on how to use the crossing correctly i.e. in a safe manner. We now mention the additional causal factors which led to the occurrence of the incident as per the findings of the inquiry.

- The gates on both sides of the crossing were found to have been left open for a lengthy period of time despite requiring to be closed when the crossing is not in use. Apparently, the gates were unable to be closed due to the overgrown vegetation that had developed around them, and thus rendered the gate usable; this was due to inadequate maintenance of the gates.

- The motorist involved in the collision had used the crossing for 36 years and six times in the week leading up to the incident; in this time the driver had never come across a train on this section of the track. Due to this past experience he, along with other authorized users, became accustomed to leaving the gates open and were thus of the expectation that a train would come when the track was being crossed.

There were, of course, other factors concerning the crossing, which were not immediately apparent, that were also thought to have contributed to the incident:

- the short warning time to alert a motorist of an oncoming train;

- the signs which warn a user to stop and read how to cross safely had poor visibility due to shrouding by foliage and vegetation and, moreover, there were problems with their wording.

- no telephone number was provided at the crossing for contacting a railway employee in the event of an accident.

In the subsequent enquiry it was also disclosed that letters had been sent to the authorized users of this gate several times, reminding them of the rules regarding gate closure. When questioned, they could not explain why the gates were left open. It is therefore most likely that due to the low levels of the users became complacent about the safe use of the crossing. In summary, there was no physical breakdown that caused the accident, but a number of unwanted behaviours arose through poor information flows, an unchecked growth in poor practice and complacency.

### 6.2.2 Lambrigg Derailment

On the 23$^{rd}$ February 2007, a Virgin train traveling from London to Glasgow derailed between Preston and Carlisle, at Lambrigg Ground Frame crossover located near Cumbria, Rail Accident Investigation Branch (RAIB) (2007). Of the 108 passengers and four crew members traveling aboard, one fatality resulted with a further twenty two individuals requiring hospital treatment. The immediate cause of the derailment was identified as faulty points on the track; this was in turn the result of the fault in the stretcher bar of the points which consequently led to the left an right switch rails to become disconnected.

Two securing bolts were also detached from the stretcher, one of which was lying next to the points while the other was missing entirely. The nuts which secure the stretchers together, having being tightened with the incorrect equipment, became free due to dynamic loadings; this fault failed to be identified in a subsequent inspection due to unauthorized splitting of patrol groups. As such faults could have potentially been identified and corrected through remedial action, prior to derailment, and the incident could have thus been prevented as it was highlighted that the deterioration of the points had occurred some time before.

The causes highlighted above, which were the most apparent causes of the derailment, were further identified in the post-incident investigation, as having been caused by a number of underlying contributory factors based on human reliability and error.

1. There existed a number of deficiencies in the inspection and maintenance regime ultimately causing the points to fall into despair and the fault thus being unidentified. Such deficiencies included:

    - A breakdown in the local management structure responsible for inspection and maintenance. Inspection carried out were found to be non compliant with set standards and procedures and supervisors tended to reinforce this behaviour as acceptable by employing unsafe arrangements.
    - Track patrolling regime's systematically failed to inspect the area adequately. Routine inspections routinely cover a required mileage of track; due to management incompetence the area of the containing the fault was overlooked one week prior to the derailment; the

subsequent patrol report was nevertheless authorized, with a gap in the inspection going unnoticed.

2. Mandated standards were not communicated or executed in the required manner, with a lack of sample checking of the track to test inspection quality and arrangements. Between maintenance and track supervisory management there was evidence of a split 'them and us' culture which had consequential effects on the way in which operations were conducted.

3. Patrolling of the track was poorly managed; patrollers were allocated to random patrol lengths thus compromising understanding of certain areas of the track and many of the patrollers' certification of competence had lapsed with lack of evidence to suggest any assessment of monitoring; despite this lapse being highlighted to local management, it was ignored and this behaviour thus became acceptable. There was no review of patrols and there was no definite method by which defects on site were marked, with checklists to identify these being used inconsistently.

4. The quality assurance regime did not recognize failures in the reliability of inspection regimes or in the application of best practice. Personnel were not briefed about any new standards requiring compliance and staffs' competency in the following practice was not managed. Failure to follow rules and standards went unreported and not acted upon; such unacceptable behaviour was encouraged as it was enacted too by higher level supervisors.

While the accident had a clear immediate physical cause in the faulty points, it is clear that the real cause of the accident was human, managerial and organizational. No single individual failed. Rather, many human factors contributed to the accident ranging from the managerial to the cultural.

### 6.2.3 The Flixborough disaster

On 01 June 1974, a pipeline rapture at a Flixborough plant led to the explosion pf a cyclohexane vapour cloud, causing 28 deaths and very large amount of collateral damage to the surrounding populated area (a821 houses and 167 shops were damaged). The cause of this accident was primarily an inadequate design modification on Reactor Number 5, in the form of a bypass pipe. This modification was made in response to damage to the reactor, namely a $5'6''$ long split, whose cause was not fully investigated. The team appeared to assume that the installation of the bypass line was a relatively routine plumbing job, failing to realize the necessity for a proper technical assessment of safety considerations.

There was no pressure testing carried out for the modified system, and there was also a failure to site the explosive material away from the control room, or adequately protect the control room from any potential blast, despite the fact that the HMFI had apparently already suggested the installation of shatter-proof windows in the control room. The three main factors involved were therefore:

- The inadequate design/safety-testing of the bypass pipeline (latent failure).

- The inadequate training/engineering experience of the staff involved.
- The inappropriate siting of explosive material or a lack of protection facilities for the control room (or both).

Another contributory factor in this disaster appears to have been the desire to get the system back into service as soon as possible.

### 6.2.4 The Ekofisk Bravo blowout

The Ekofisk blowout incident, which occurred in 1977, caused a large degree of collateral damage. A blowout developed while an offshore well was a being worked on. Apparently, a blowout preventer should have been able to prevent this, but this had unfortunately been incorrectly installed. The events that led to the incident occurred over approximately a 36-hour period, although some of the deeper, root causes appear to relate to high-level failures to impose proper safety controls. The significant human-error forms elicited from this incident were:

- The failure of information to be properly transferred across shift boundaries.
- The fact that one individual, in control of the operation, had been on shift almost continuously for nearly 36 hours. In addition, a critical safety component was inserted upside down, making it ineffective. This mistake could have been rendered impossible by the design of the system; nor was the mistake obvious to the operators, since particular visual information that would have indicated which way up it should be inserted was missing.
- The failure on the part of government departments to impose high-level controls on the safety aspects of particular operations, and controls such as blowout preventers.

Of particular interest is the failure of important safety information to be transferred effectively across the shift boundary, as a result of inadequate shift-handover procedures. Some symptoms of an impending blowout had occurred (signs of a 'kick' developing), but, none of the shift workers appeared to make any effort to tell any of the oncoming shift's workers about these signs. There were problems with the operation, and one of the key personnel had been on shift for an exhausting amount of time. The lack of visual feedback as to the correct orientation or nature of a component is not uncommon in maintenance, and yet here we see the total disabling of a major safety barrier (preventing an oil/gas blowout) via one simple maintenance-type error.

## 6.3 Aviation

The work presented in this section is taken from Kirwan (1994).

### 6.3.1 The Paris air disaster

In 1973 a Turkish Airlines DC-10 crashed in a forest shortly after take-off from Paris, as a result of mishap involving a pressurized cargo door. The cargo door opened outwards, and its accidental opening was prevented by means of locking pins, Williams (1987). The 'operators' knew that the door would be secure if a handle, connected to the locking-pin mechanism, could be put in the closed position, hence allowing pressurization. However, whilst the operators believed that the successfully closed handle indicated a fully secured door, what they were unaware of was the case where, due to a source of resistance elsewhere (the door not properly closed), a rod in the mechanism could buckle and allow the handle to appear closed without actually locking the pins home - the situation which occurred on this day, leading to a tragic loss of life (via sudden depressurization during the late stages of take off-the door flew at altitude). The primary cause, aside from a simple 'design error' (the operator could not have foreseen in this event), was, as William states, the violation of a human-factors principle, namely:

- The feedback mechanism should have given direct (and hence reliable or true) feedback to the effect that the pins were lock home, which did not.

This lack of direct feedback recurred in the Three Mile Island incident, discussed earlier. This type of problem has also been mentioned with respect to the activation of a control component. For example, the operator should know not only that a control input to an actuator has been sent, or even simply that it is has been received, he or she should in fact know that the component has moved, or fired, or achieved whatever its mission was.

### 6.3.2 The crash of the BEA trident 1

The crash of the BEA Trident 1 at Staines on 18 June 1972 incurred a death toll of 118. Specifically, the aircraft stalled during take-off, due to the pilot's failing to maintain sufficient air-speed. Howland (1980) cites three main causes for this accident:

- The absence of a mechanical interlock preventing the retraction of slats at too low an air-speed.
- A lack of experience/training, on the part of the crew involved, for this type of emergency.
- The captain's ill health during take-off.

All three component causes are important. The use of interlocks is an important way of avoiding certain types of potentially fatal, or at least hazardous, errors. Interlocks may be omitted where it is felt that personnel have enough experience not to do 'incredible' things, but this strategy can still clearly fail - as in this case, where an inexperienced team was involved. However, too many interlocks may lead to operators or maintainers trying to use various methods to overcome them, for example, an incident occurred in which one maintenance crew overcame an interlock in a nuclear-reactor plant. This led to an incident in which, after they had gone to lunch leaving the interlock disabled, another operator attempted to operate the system, unaware of the status of the system and of the illegal interlock-override that had been carried out.

The lack of experience and training for a particular emergency is an often-cited problem. In any kind of emergency which causes high stress levels, it is somewhat optimistic to expect operators to handle an event which may be complex and totally new to them. A case in point was the Kegworth air crash, which presented the pilots with a complex scenario for which training and their interface had left them ill-prepared. This perception accounts for the large amount of simulator training which pilots traditionally undertake. Illness - or worse, the incapacitation of one member of a crew of a crew - is a potential problem which a high-risk plant must address via its staffing policy.

### 6.3.3 The Challenger Space Shuttle disaster

On 28 January 1986, the Space Shuttle Challenger was destroyed in an accident during take-off, together with its full crew of seven astronauts. Essentially, the accident was cause by a low-temperature-induced fuel-sealed (o-ring) failure on the booster rockets. Fuel leaking out eventually reached the ignition source at the bottom of the firing rocket boosters whereupon the boosters catastrophically failed and much of the rocket disintegrated, the shuttle itself plunging back towards the ocean. The following major causes were found to have contributed to the accident:

- An effective 'silent safety program' within NASA (this was the principal cause).
- The lack of clear corporate safety organization.
- The lack of adequate communications down the line and between lines (e.g., between project managers and flight-operations staff).
- An increased flight rate combined, on the other hand, with decreased resources.
- Maintenance-management problems.

All in all, the challenger disaster appeared to lack a clear and adequate safety-management structure, and this may reflect the difficulties that a largely technical body such as NASA experienced in trying to cope with commercial pressures with which it was not familiar.This accident also highlighted the use of risk analysis in determining operational safety in a pressurized situation. According to the inquiry accidents, the company producing the boosters was concerned at the intention to carry on with the launch following very cold temperatures, since they could not verify the safety of the o-rings (the devices which failed ) following such low temperatures. Thus, the engineers claimed that this matter fell outside of their design basis, and hence that they could not say if it was safe to launch. A management decision then overrode these concerns by arguing that the engineers could not at the same time prove that it was unsafe to launch - with hindsight a most unsatisfactory decision stance to adopt. This shows the effect of pressures on the management of technical systems, and demonstrates how good safety advice can sometimes be ignored.

### 6.3.4 Research reactor sensitivity analysis

#### 6.3.4.1 Background

This study concerned a human reliability analysis for a research reactor. In this analysis, which used the SLIM-MAUD technique, one particular error, related to a valve closure, was found to be contributing significantly to the estimated level of risk inherent in the system. It was therefore desirable to calculate how to reduce the HEP via error-reduction measures (ERMs), by using the SLIM system. This case study only details this error-reduction analysis/sensitivity-analysis part of the study, so as to exemplify how such an error-reduction analysis (ERA) may be carried out, whether by means of the SLIM, HEART or HRMS systems or by means of other such systems with ERA capabilities T. Warers and D. Embrey were the co-assessor.

#### 6.3.4.2 Problem definition

The problem facing the assessor was the quantification of a failure to close a valve in specified period of time, occurring within a loss-of-coolant type of accident scenario. If the resultant HEPs were found to be too high, then an ERA would be required for a reduction of the HEP to an acceptable value.

#### 6.3.4.3 Task, human error, representation and quantification analyses

We do not give a comprehensive detail of these here. A detailed task analysis was, however, undertaken - via interviews, walk-throughs, observations and documentation reviews - by the client. A fault tree had also been developed by the client, in which the particular error of concern was required to be calculated. A small panel of expert judges was then convened, and using the SLIM-MAUD system, the HEP was then calculated.

#### 6.3.4.4 Impact assessment

The resultant HEP was found to be 0.10. Although this is not an uncommon value for a time- pressured emergency response made with little procedural support, there was a desire, nevertheless, to reduce this figure to a value of 0.001 - a two-orders-of-magnitude change, and a significantly large error reduction. If this could be achieved, then the results of the analysis would indicate a tolerable level of safety.

#### 6.3.4.5 Error reduction analysis

The ERA proceeded with two members of the original expert group. Using a part of the SLIM-MAUD software system called SLIM-SARAH, a series of alternative PSF profiles were generated, profiles which could help to reduce the HEP. It is important to stress that this was not a purely numerical exercise - in fact, far from from it. The exercise proceeded by first determining which PSF-rating/-weighting combination was contributing most to the HEP value. In this case it was the *procedures* PSF, with an undesirably high PSF rating value of 8 - which in this case signified that procedures were virtually non-existent for this particular scenario. Ways of improving procedures were then discussed. Such procedural improvements would amount to the provision of a symptom-based flowchart procedure with check-points, etc. The specific improvements were noted down as assumptions underpinning the to be-revised figure. The two experts agreed what new rating value would be gained if such improvements were implemented (in this case a rating of 2).

This new value was substituted in to the SARAH system, and the HEP was recalculated, giving a new figure of 0.03, and an error-reduction factor of 3. Since this was not enough of a reduction, an alternative PSF was tried, namely that of *training*. Various retraining improvements were considered which altogether could reduce the rating to a value of 3. However, when the original HEP was now recalculated, the results yielded a reduction factor of only 1.4 and a new HEP of only 0.07. Other combinations of improvements were considered, including the provision of a new alarm in the interface. This ERM, together with the greatest degree of improvement in training and procedures that could be realistically achieved within the client's budget, yield a value of 0.001 in iteration number 6. Several more iterations were tried, but only iteration 6 produced the desired reduction factor. Its assumptions in terms of changes to be made to training, procedures and the interface, were also accepted for implementation by the client.

# Chapter 7

# Review of HRA methods

## 7.1 Introduction

Human reliability assessment (HRA) involves the use of qualitative and quantitative methods to assess the human contribution to risk. There are many and varied methods available for HRA, with some high hazard industries developing 'bespoke' industry focused methods. HRA techniques generally fall into two categories, those utilizing a database, and those using expert opinion. The origins of HRA methods dates from the 1960s with the objective of modeling the likelihood and consequences of human error. Over the years, methods of varying tendency to fail have been developed and this has resulted in over a dozen of HRA techniques. These HRA methodologies are categorized in literature in terms of their underlying concepts, models and applicability. They are divided into three categories: first, second, and third generation.

The first generation methods were the first to be developed to help risk assessors predict and quantify the likelihood of human error. They tend to be atomistic in nature; they encourage the assessor to break a task into component parts and then consider the potential impact of PSFs, Bell & Holroyd (2009). In the early 1990s, the need to improve HRA methods approaches to focus on cognitive human behaviour interested a number of important research and development activities around the world, Di Pasquale et al. (2013). This gave birth to second generation methods. They consider context and errors of commission in human error prediction. New techniques are now emerging based on earlier first generation methods, and are being referred in literature as third generation methods. According to French et al. (2009) these tools go further modelling a range of behaviours in order to recover and avert failure; they contain more dynamic simulation, and have to be implemented on a computer.

## 7.2 HRA methods

The following HRA methods are reviewed with the use of the following references: Bedford & Cooke (2001), Bell & Holroyd (2009), French et al. (2009), Kirwan (1994), and Kirwan (1997), except for the areas where indicated.

### 7.2.1 Absolute Probability Judgement (APJ)

#### 7.2.1.1 Background

APJ, also known as Direct Numerical Estimation (DNE), is an expert judgement-based approach which involves using the beliefs of experts (e.g. front-line staff, process engineers etc.) to estimate HEPs. The technique has two forms: group methods and single expert methods. The particular technique used for expert opinion combination is not fixed. Group methods tend to be more popular and widely used as they are more robust and can be used to generate a consensus opinion. In the group approach, the outcome of aggregating individual knowledge and opinions is said to be more reliable, in the sense that it allows for different expertise to be represented in the expert group. APJ has been used in the nuclear and offshore industries.

#### 7.2.1.2 Description of the technique

The APJ approach is conceptually the most straightforward HRA quantification approach. It simply assumes that people can estimate directly the likelihood of an event. There are different APJ approaches that can be applied to determine human reliability. A 'single expert APJ' would require one expert to make their own judgements on the chances of a human error. There are four main group methods by which APJ can be conducted.

- Aggregated individual method. This is where individuals make their estimates individually and then a geometric mean of these estimates is calculated.
- Delphi method. For this method, individuals make their estimates independently of each other, but the assessments are then shared, allowing the experts to reassess their own estimates based on the new information. Then a geometric mean of the HEP scores is calculated.
- Nominal group technique. This method is very similar the Delphi method, the difference is that the experts are given the opportunity to discuss their estimates and confidentially re-evaluate their assessment. these scores are then statistically aggregated.
- Consensus group method. In this method, the experts meet and discuss their estimates, following which a consensus an an agreed estimate must be reached. If this is not possible then a statistical

aggregation of the individual estimates is calculated, or it may be necessary to revert to one of the other group APJ methods.

#### 7.2.1.3 Advantages and Disadvantages

*Advantages of APJ*

- The method is relatively quick and straightforward to employ, and allows as much detailed discussion as the experts think fit; this can often itself be qualitatively useful.
- The technique has been shown to give accurate estimates in a wide variety of fields (e.g. weather forecasting).
- APJ is not restricted to be specialised for use in a particular field; it is applicable to an HRA on any industrial sector thus making it a generic technique for use in a wide range applications.
- Useful suggestions may result from discussion as to ways in which a reduction in errors can be achieved.

*Disadvantages of APJ*

- The APJ is prone to certain biases, as well as to personality/group problems and conflicts, which, if not effectively countered, can significantly undermine the validity of the technique.
- Locating suitable experts for the APJ exercise is a difficult stage of the process, more so due to the ambiguity with which the term 'expert' can be defined.
- Since the technique is often likened to 'guessing', it enjoys a somewhat low degree of apparent, or 'face' validity. That is, there is no means by which guesses can be validated.

### 7.2.2 A Technique for Human Error Analysis (ATHEANA)

#### 7.2.2.1 Background

ATHEANA is both a retrospective and prospective HRA methodology developed by the US nuclear industry regulatory commission in 2000. It is a second generation tool. It was developed in the hope that certain types of human behaviour in nuclear plants and industries, which use similar processes, could be represented in a way in which they could be more easily understood. It seeks to provide a robust psychological framework to evaluate and identify PSFs -including organisational/ environmental factors - which have driven incidents involving human factors, primarily with the intention of suggesting process improvement. There is no evidence of it being applied in other domain.

#### 7.2.2.2 Description of the technique

There are ten steps in the ATHEANA methodology.

- Define and interpret the issue under consideration.
- Detail the required scope of the analysis.
- Describe the Base case scenario for a given initiating event, including the norm of operations within the environment, considering actions and procedures.
- Define Human Failure Events (HFEs) and/ or unsafe action (UAs) which may affect the task in question.
- Identify potential vulnerabilities in operator's knowledge base.
- Search for deviations from the base case scenario for which UAs are likely.
- Identify and evaluate complicating factors and links to PSFs.
- Evaluate recovery potential.
- Quantify HFE probability.
- Incorporate results into the PRA.

#### 7.2.2.3 Advantages and Disadvantages

*Advantages of ATHEANA*

- It provides a much richer and more holistic understanding of the context concerning the Human Factors known to be the cause of the incident, as compared with most first generation methods.
- It increases the guarantee that the key risks associated with the HFEs in question have been identified, and focuses on the important issues of context and cognition.
- By making use of ATHEANA, it is possible to estimate HEPs considering a variety of differing factors and combinations.
- It can be used to develop detailed qualitative insights into conditions that may cause problems.
- It allows for the consideration of a much wider range of PSFs and also does not require that these be treated as independent.

*Disadvantages of ATHEANA*

- The method is very Cumbersome and presumable very costly. The guidance is too complex and depends too much on subject matter experts.
- The qualitative results are good, but these might have been obtained in other ways, perhaps more efficiently.
- The quantification method is weak, and the quantitative results are unsubstantiated. The quantification is excessively dependent on expert judgement, hence possibly has low reliability as a method.
- The method is not described in sufficient detail that one could be sure that different teams would produce the same results.

### 7.2.3 Conclusions from Occurrences by Descriptions of Actions (CODA)

#### 7.2.3.1 Background

Reer (1997) first outlined CODA in his conference paper. The CODA method uses an open list of guidelines based on insights from previous retrospective analyses. The general approach is to compile a short story that includes all visual occurrences and their essential context without excessive technical details. The analysis should then focus on the potential major occurrences first. This tool has been used in the nuclear industry.

#### 7.2.3.2 Description of the technique

- The method presents a list of criteria (i.e. items for data acquisition) that are easy to obtain (i.e., objective as free from judgement as possible) and which have been proved to be useful for causal analysis.
- For example, for each incorrect human response that has happened the analyst will look for: the critical action; the underlying goal or plan if it is self-evident; the anticipated correct response and its consequence; the underlying task and sub-task; the underlying sequence of events.
- Many guidelines are given for the causal analysis of each situation; these are mainly holistic, comparative and generalizing in nature.
- There are a number of event cases in the literature that have been used to demonstrate that CODA is able to identify cognitive tendencies as typical attitudes or habits in human decision-making.

#### 7.2.3.3 Advantages and Disadvantages

*Advantages of CODA*

- CODA treats human interventions in a neutral and flexible manner.
- It incorporate three search processes: actions, system-failures and scenarios. Through integrating these approaches the search procedure may be optimised towards low analytical effort and high completeness.

*Disadvantages of CODA*

- The necessary details for carrying out an assessment using CODA are not available.

### 7.2.4 Connection Assessment of Human Reliability (CAHR)

#### 7.2.4.1 Background

CAHR was developed at the Technical University of Munich and the Gesellschaft fur Anlagen-und Reaktorsicherheit (GRS) between 1992 and 1997. This method is being developed for use in Air Traffic Management. It is a second generation tool that combines event analysis and assessment in order to use past experience as the basis for HRA. It is a database system used for analysing 'operational disturbances', which are caused by inadequate human actions or organisational factors. CAHR has a generic underlying model that is applicable to all observable events and to allow the collection of all information on human error events. This tool is used in the nuclear industry, and has been applied to various areas including occupational health and safety, shipping, car industry, aviation safety and software ergonomics.

#### 7.2.4.2 Description of the technique

There are three key elements to the tool:

- A framework for structured data collection (both retrospective and prospective information).
- A method for quantitative analysis.
- A method for HRA (quantitative analysis).

The philosophy underlying this technique is:

- The focus of analysing is the work system and not the human.
- Human error results from the interrelation of several situational and causal factors of the working system.

- The uses a fixed structure but no fixed taxonomy.
- Strict differentiation between observable information (phenotypes) and causes (genotypes) in the event analysis and description.

#### 7.2.4.3 Advantages and Disadvantages

*Advantage of CAHR*

- This tool is one of a number of tools that attempt to collate information from previous events and build an extensive database. It is an attempt to move towards analysing error of commission and to capture the complexity of human behaviour.

*Disadvantage of CAHR*

- There has not been a proven, practical methodology for analysing potential errors of commission.

### 7.2.5 Cognitive Reliability and Error Analysis Method (CREAM)

#### 7.2.5.1 Background

Hollnagel (1998) was the first to develop CREAM, following an analysis of the methods for HRA already in place. It is the most widely utilized second generation HRA technique and is based on three primary areas of work; task analysis, opportunities for reducing errors and possibility to consider human performance with regards to overall safety of a system. There are two version of the technique, the basic and the extended version. Both versions have the ability to identify the importance of human performance in a given context and a helpful cognitive model and associated framework, usable for both prospective and retrospective analysis. CREAM has been applied in the nuclear industry and to rail crash scenario, but there is no evidence of extensive use.

#### 7.2.5.2 Description of the technique

- Task Analysis. For the purpose of HRA (the CREAM basic method) the first step is a task analysis. Based on this a list of operator activities is produced, from which a CPC analysis is carried out.
- Context description. The intention of the basic CREAM method is to use it as a screening technique with aim of identifying processes which require a deeper level of analysis; this analysis may then be carried out by the extended CREAM method.
- Specification of Initiating Events. CPCs are assessed according to the descriptors in order to judge their expected effect on performance.

- Error Prediction. The assessment of the CPCs the require to be adjusted according to some specified rules in order to take account of synergistic effects.
- Finally, a simple count is performed of the number of CPCs that are causing an improvement in reliability and those that which are reducing it. From this number the probable control mode is determined.
- The extended version of the CREAM methodology operates in a slightly different manner. Following the initial task analysis, a refinement is then provided in terms of the cognitive activities which are involved in the considered task. To these activities Contextual Control Model (COCOM) function is ascribed so that a cognitive demand profile may be established.
- Following this stage, the probable cognitive function failures are identified, based on a knowledge of the specified tasks, yet following a set of generic cognitive functions failures associated to the COCOM functions. Each of the generic failures is associated with a nominal probability which is based on a table given in CREAM. However these probabilities are adjusted according to the particular mode.

Hence in the extended version of the CREAM methodology the control mode acts in the role of a performance shaping factor with the task of performing adjustments to a nominal probability, French et al. (2009).

#### 7.2.5.3 Advantages and Disadvantages

*Advantages of CREAM*

- It allows for the direct quantification of HEPs.
- The resultant model is highly integrate-able into the primary safety process in use.
- The approach is very concise, well structured and follows a well laid out system of procedure.
- It also allows the assessor to specifically tailor the use of the technique to the contextual situation.
- The approach uses the same principle for retrospective and predictive analyses.

*Disadvantages of CREAM*

- The technique requires a high level of resource use, including lengthy time periods for completion.
- It also requires an initial expertise in the field of human factors in order to use the technique successfully and may therefore appear rather complex for an inexperienced user.
- CREAM does not put forth potential means by which the identified errors can be reduced.

### 7.2.6 Commission Errors Search and Assessment (CESA)

#### 7.2.6.1 Background

Reer & Dang (2007) proposed an errors of commission identification method, which gave rise to the CESA method . What it does is, it integrate aspects from the search schemes of some HRA methods (e.g. ATHEANA) with concepts from CODA (to be discussed). CESA has been applied in the nuclear industry.

#### 7.2.6.2 Description of the technique

- The CESA method is strongly based on importance screening. The identification process prioritises plant systems with a high achievement worth and scenarios are prioritised based on the size of the contribution to the core damage frequency.
- A trade-off is made between the scenarios with a high safety impact against the completeness of the search.
- The intention is to bias the search towards Errors Of Commission (EOC) situations that are risk-significant and credible.
- The first step in the CESA methodology is to catalogue key action response to the plant events to be reviewed .
- This catalogue is then used in a systematic search of context-action combinations, to obtain a set of situations with EOC opportunities; these situations are then analysed in detail.

#### 7.2.6.3 Advantages and Disadvantages

*Advantages of CESA*

- The CESA identification process is feasible and effective; it is able to identify plausible situations in which EOCs may occur.

*Disadvantages of CESA*

- The quantification of the risk contribution using CESA is uncertain and this uncertainty is larger than would be typical of other HRA methods.

### 7.2.7 Human error HAZOP (hazard and operability) study

#### 7.2.7.1 Background

According to Crawley & Tyler (2015) the HAZOP study method was developed by ICI in the 1960s and its use and development was encouraged by the Chemical Industries Association (CIA) Guide published in 1977. As mentioned in Ellis & Holt (2009) the objective is to identify human failures during an activity and assess the potential for operator recovery or other risk reduction measures that stop this failure escalating to a major accident. Since then it has become the technique of choice for many of those involved in the design of new processes and operations. In addition to its power in identifying safety, health, and environmental (SHE) hazards, a HAZOP study can also be used to search for potential operating problems. The HAZOP study is a well-established technique in process-design audits and risk assessments in engineering. It is primarily applied to process and instrument diagrams.

#### 7.2.7.2 Description of the technique

In Ellis & Holt (2009) the different steps in carrying out follows this technique are discussed as follows.

- Identify 'safety critical' activities- these are defined as operating or maintenance procedures with the potential to cause or limit the escalation of major accident hazards (MAH).
- Hierarchical task analysis- a task analysis is used to list the key steps in the activity that will be used for human failure identification stage. The level of detail at each step in the procedure is related to the specific hazardous event under assessment.
- Identify potential human failures - a team of knowledgeable and experienced staff from the plant are required to carry out the hazard identification study.
- Assess consequences - for all credible human failures the team assesses the initial and ultimate consequences assuming that there is no recovery and that other non-passive risk reduction measures fail to prevent escalation.
- Assess potential for recovery - for failures with significant consequences, the team assesses the potential for human recovery from the initial failure. The recovery process generally follows three phases: detection of the error, diagnosis of what went wrong and how, and correction of the problem.
- Assess risk reduction measures - the team identifies the 'engineered' risk reduction measures currently in place, including inherent, passive and active protection systems. The team should consider improvement options to eliminate or reduce the risk associated with human failure; if the risk of human failure can be significantly reduced by recommendations for improvements no further assessment is required, otherwise the team consider how likelihood oh human failure can be reduced in the next step.
- Improvements to prevent human failures - it is said that based on the causes of human failure and the potential for recovery , the likelihood is assessed by the team. The objective is to assess

whether events or near misses are occurring more frequently than would be anticipated from generic data. Factors that could affect the likelihood of the identified human failure are assessed by the team.

- Record of 'human-HAZOP' - the results of the HRA are recorded on a 'human-HAZOP' record table with the following columns:

  1. Step: Description of task by a person carrying out activity
  2. Human Failure: Description of credible human failures when carrying out this task based on guide diagram.
  3. Consequences/ Severity: Ultimate consequences of the human failure if there is a failure to recover or mitigate the event.
  4. Potential to recover/ Likelihood: Potential for human recovery involving detection of problem, diagnosis and correction.
  5. Risk Reduction Measure: Risk reducing measures to prevent escalation of the incident that do not involve human intervention.
  6. Recommendations: Practical actions to reduce the potential for failure based on the PIFs or for further risk reduction measures based on the hierarchy of measures.

#### 7.2.7.3 Advantages and Disadvantages

*Advantages of Human error HAZOP*

- It is usually applied in the early stages of the design of a system or subsystem. It also translates the experience both of HAZOP chairman and of selected system-design and operational personnel into a powerful analysis of a new design.
- This technique also identifies errors in a system-knowledge-environment.

*Disadvantages of Human error HAZOP*

- An extension of the traditional HAZOP technique may only be useful for specific, largely human-operated situation. This technique therefore requires further development of its range of applications.

### 7.2.8 Human Cognitive Reliability Correlation (HCR)

#### 7.2.8.1 Background

HCR is a psychology/cognitive modelling approach to HRA developed by Hannaman et al. (1984) in 1984. This technique comprises a set of three Time Response Curves (TRCs), where time is plotted on one axis, and probability of diagnostic non-response is plotted on the other. The method uses

Rasmussen's idea of rule-based, skill-based, and knowledge-based decision making to determine the likelihood of failing a given task, as well as considering the PSFs of operator experience, stress and interface quality, see Rasmussen (1983). HCR is the most famous of the Time Response Curve approaches. It was originally developed for use within the nuclear industry, primarily to assess failure to respond in time in emergency decision-making situations, and is not applicable to situations outside this domain.

#### 7.2.8.2 Description of the technique

The methodology for HCR is broken down into a sequence of steps as given below:

- The first step is to determine the situation in need of a human reliability assessment. Then an appropriate decision making is determined.
- From the relevant literature, the appropriate HCR mathematical model or graphical curve is then selected.
- The median response time to determine the task in question is thereafter determined.This is commonly done by expert judgement, operator view or simulator experiment. This time is sometimes referred to as the nominal response time in literature, and denoted $T_{1/2}$.
- The median time is adjusted to make specific to the situational context. This is done by means of the PSFs coefficients $K_1$ (operator experience), $K_2$ (stress level) and $K_3$ (quality of operator/plant interface) given in the literature and using the following formula

$$T_{1/2adjusted} = T_{1/2nominal}(1+K_1)(1+K_2)(1+K_3).$$

- For the action being assessed the time window ($T$) should the be calculated, which is the time in which the operator must take action to resolve correctly the situation.
- To obtain the non-response probability, the time window ($T$) is divided by $T_{1/2}$, the median time. This gives the Normalised Time Value. The probability of non-response can then be found by referring to the HCR curve selected earlier.

This non-response probability may be integrated into a fuller HRA; a complete HEP can only be reached in conjunction with other methods as non-response is not the sole source of human error.

#### 7.2.8.3 Advantages and Disadvantages

*Advantages of HCR*

- It is a fairly quick technique to carry out and has a relative ease of use.

- The method explicitly models the time-dependent nature of HRA.
- The three models of decision-making proposed by Rasmussen, are all modelled.

*Disadvantages of HCR*

- The method is very sensitive to changes in the estimate of the median time. Thus, any inaccuracy in this estimate will make the estimation of HEP suffer.
- Only three PSFs are included in the methodology; there are several other PSFs that could affect performance which are unaccounted for.
- It is highly resource intensive to collect all the required date for the HCR methodology, particularly due to the necessity of all new situations which require an assessment.
- The HEP produced by HCR is not complete; it does not give any regard to misdiagnoses or rule violations.
- There is no sense of output from the model that indicates in any way of how human reliability could be adjusted to allow for improvement or optimization to meet required goals of performance.
- The same probability curves are used to model non-detection and slow response failures. These are very different processes, and it is unlikely that identical curves could model their behaviour.
- The rules for judging Knowledge-based, Skill-based and Rule-based behaviour are not exhaustive.Wrong assignment of behaviour to a task can mean difference of up to two orders of magnitude in the HEP.

### 7.2.9 Human Error Assessment and Reduction Technique (HEART)

#### 7.2.9.1 Background

HEART was first developed by Williams (1986). It is a first generation HRA technique and is still widely used in the UK. This method is based upon the principle that every time a task is performed there is a possibility of failure and that the probability of this is affected by one or more Error Producing Conditions (EPCs) This technique is based on the technique author's analysis of a substantial section of the Ergonomics literature, and uses a set of basic error probabilities modified by the assessor by structured PSF considerations. It is a general method applicable to any situation or industry where human reliability is important. It has been successfully applied in many industries including nuclear, chemical, aviation, rail and medical.

#### 7.2.9.2 Description of the technique

HEART is designed to be a quick and simple method for quantifying the risk of human error. The following stages are used in this method.

- The first stage of the process is to identify the full range of sub-tasks that a system operator would be required to complete within a given task.
- Once this task description has been constructed a nominal human reliability score for the particular task is then determined, usually by consulting local experts. Based around this calculated point, a $5^{th} - 95^{th}$ percentile confidence interval range is established.
- The EPCs, which are potentially relevant for the given situation, are then considered and the extent to which each EPC applies to the task in question is discussed and agreed, again with local experts.
- A final estimate of the HEP is then calculated using the EPC scores.

#### 7.2.9.3 Advantages and Disadvantages

*Advantages of HEART*

- HEART is versatile, very quick and straightforward to use.
- The technique provides the user with useful suggestions as to how to reduce the occurrence of errors.
- It provides ready linkage between Ergonomics and Process Design, with reliability improvement measure being a direct conclusion which can be drawn from the assessment procedure.
- Requires relatively limited resources to complete an assessment.

*Disadvantages of HEART*

- Error dependency modelling is not included.
- Requires greater clarity of description to assist user when discriminating between generic tasks and their associated EPCs; there is potential for two assessors to calculate very different HEPs for the same task.
- Potential for double counting (some elements of EPCs are implicit in the task description).
- Lack of information about the extent to which tasks should be decomposed for analysis.
- It has a subjective nature of determining the assessed proportion of affect.

### 7.2.10 Human Reliability Management System (HRMS)

#### 7.2.10.1 Background

HRMS was developed primarily by Kirwan & James (1989) to inform the design process for BNFL THORP (British Nuclear Fuels Ltd., Thermal Oxide Reprocessing Plant). It can be used to carry out task analysis , error analysis, and performance shaping factor-based quantification. The assessors choose the error descriptor which most resembles the characteristics of the HEP being assessed, and the system's algorithms extrapolate from this datum to the required HEP based on the assessors answers to the PSF questions. It can only be used to quantify HEPs for tasks related to its contextual origins, namely nuclear chemical plant operation.

#### 7.2.10.2 Description of the technique

- The methodology is based on other techniques that were available at the time of development. Specifically, the quantification system (PHOENIX, the Prediction of Human Operator Error using Numerical Index eXtrapolation) is based on the Success Likelihood Index Method (SLIM), Influence Diagrams Approach (IDA) and HEART.
- PHOENIX comprises
  1. Seven Task types categories - the assessor selects the task type that most closely resembles the task being assessed, and within each task category there are a number of error types that the assessor could choose from.
  2. Six PSFs - about which the assessor is asked up to fifty questions to determine the strength of impact (this removes the need for the subjective judgements by assessors)
- The HEP associated with the selected task type is multiplied with the PSF value, however the model allows for interactions between PSFs and with the task types, something that previous tools were unable to offer.

#### 7.2.10.3 Advantages and Disadvantages

*Advantages of HRMS*

- The method is useful for scenario that need to be assessed in depth.
- It deals with the whole HRA process from task analysis to error reduction and documentation.
- PSF rating questions are factual rather than judgmental.
- HRMS has the potential to derive or sometimes learn extrapolation rules.

- It is founded on industrial data.

*Disadvantages of HRMS*

- The method is resource-intensive.
- It is not empirically validated but has been used successfully in the nuclear industry.
- It is viewed to be a HRA expert's tool, rather than as a general tool for reliability experts.

### 7.2.11 Influence Diagrams Approach (IDA)

#### 7.2.11.1 Background

An Influence Diagram (ID) is essentially a graphical representation of the probabilistic dependencies between PSFs, the factors which influence the failure probability in the performance of a task. The approach originates from the field of decision analysis and uses expert judgement to formulate and often to quantify the models. It was first outlined by Howard and Matheson, and then developed specifically for the nuclear industry by Philips. It is a first generation tool, and explicitly considers the inter-dependency of operator and organisational PSFs.

#### 7.2.11.2 Description of the technique

The IDA methodology is conducted in a series of 10 steps as follows:

- Description of all relevant conditioning events by experts who posses sufficient knowledge of the situation under evaluation.
- Refine the target event definition. The target event must be defined as tightly as possible.
- Balance of evidence. Select a middle-level event in the situation and using each of the bottom level influence, assess the weight of evidence, also known as the 'balance of evidence'.
- Assess the weight of evidence for this middle-level influence, which is conditional on bottom-level influences.
- Repeat step 3 and 4 for the remaining middle-level and bottom-level influences.
- Assess probabilities of target event conditional on middle-level influences.
- Calculate the unconditional probability of target event and unconditional weight of evidence of middle-level influences. For the various combinations of influences that have been considered, the experts identify direct estimates of the likelihood of either success or failure.

- Compare these results to the holistic judgements of HEPs by the assessors. Revise to reduce discrepancies.
- Repeat above steps until assessors are finished refining their judgements. The process is stopped when all participants reach a consensus that any misgivings about the discrepancies are resolved.
- Perform sensitivity analyses when individual experts are unsure of the discrepancies. Conducting a cost-benefiting analysis is also possible at this stage of the process.

#### 7.2.11.3 Advantages and Disadvantages

*Advantages of IDA*

- Data requirements are very low.
- Dependence between PSFs is explicitly acknowledged and modelled.
- PSFs and other influence creating error producing conditions are prioritised and if desired, the less significant ones may be ignored.
- It can be used in a strategic overview or in a very fine breakdown of a task element.
- Sensitivity analysis is possible with use of this technique.
- PSFs are precisely defined and their influence is explored in depth.

*Disadvantages of IDA*

- Eliciting HEPs requires further research with regards to their accuracy and justification.
- Building IDAs is highly resource-intensive in terms of organising and support an extensive group session a suitable range of experts.

### 7.2.12 Intent

#### 7.2.12.1 Background

- Intent was presented by Gertman et al. (1992) as a method for estimating human error probabilities for decision based errors.
- It was developed for nuclear industry; there is no evidence of it being applied to other sectors.
- In this methodology, errors of intention are viewed as an important subset of error of commission because they are related to cognitive functions (e.g. problem solving).

- Since errors of intention can result from a wide range of factors, it is difficult to model and quantify them. This tool aims to incorporate errors of intent into probabilistic safety assessment.
- This is was done by compiling a number of potential errors of intention pertinent to nuclear power plants, by using a variety of analytical techniques.

#### 7.2.12.2 Description of the technique

Four categories of error of intention were identified:

- Action consequence - these errors consider the relationship between consequences and decision making.
- Crew response set - these errors look at the influence that inhibition, experience and training have on performance.
- Attitudes leading to circumvention - these are errors that are rooted in the manner in which individuals view the world.
- Resource dependencies - this category is made up of internal resources (e.g. memory capacity) and external resources (e.g. operating procedures).

The INTENT user is offered a choice of twenty nominal errors from these four categories.

- For each error, INTENT gives lower bound and upper bound estimates of the occurrence probability, which are based upon expert opinion.
- This tool also includes a set of eleven (very brief and general, e.g. workload) PSFs whose weighting factors were also determined by expert estimates.
- Once the errors have been selected from the twenty nominal errors, estimates for the eleven PSFs are then given on a five-point scale.
- Taking the weighting factors into account it is then possible to calculate a reliability index from these estimates.

#### 7.2.12.3 Advantages and Disadvantages

*Advantages of INTENT*

- It is a relatively easy to use tool, whether by a novice or a experienced practitioner.
- This tool is claimed to provide data that can account for rare, high consequence failures due to errors of intention.

*Disadvantages of INTENT*

- This tool relies on the skill of the assessor, and the degree to which the assessor understands the task being assessed.
- The generated list of errors in INTENT may not be exhaustive.

### 7.2.13 Justified Human Error Data Information (JHEDI)

#### 7.2.13.1 Background

JHEDI was primarily developed by Kirwan & James (1989) to inform the design process for BNFL THORP (British Nuclear Fuels Ltd., Thermal Oxide Reprocessing Plant). It was developed to provide a faster screening technique than that of its 'parent', the HRMS approach, of which it is a derivative. Due to the nature of the underpinning data, JHEDI is only applicable to the UK nuclear chemical industry and, specifically, reprocessing tasks. JHEDI system is relatively rapid, and may only take up half a day for a scenario with several HEPs, assuming that the assessor knows the task requirements reasonably well.

#### 7.2.13.2 Description of the technique

- Both HRMS and JHEDI can be used to carry out task analysis, error analysis, and performance shaping factor-based quantification, but JHEDI involves a less detailed assessment than HRMS.
- The HEPs have been made more conservative to allow for simplicity within JHEDI (i.e. account for a measure of dependence) and also requires less PSF questions to be answered.
- JHEDI uses less multipliers and extrapolation rules than HRMS, making it conceptually similar to THERP and HEART, in that both of these techniques have a nominal HEP which is increased according to levels of PSF identified to be present in the situation, via multiplication.
- In JHEDI, the PSF rating process is arguably more straightforward, as it asks factual questions which which can be substantiated, rather than asking for subjective judgement by the assessor.
- Like HRMS, JHEDI is based on actual industry data, which is context specific, and supplemented with expert judgement by the tool's authors.

#### 7.2.13.3 Advantages and Disadvantages

*Advantages of JHEDI*

- This tool is relatively quick and requires little training to apply.

- JHEDI uses real data rather than simulated data.
- It is computerised tool and the data can therefore be easily recorded and audited.
- Models such as HEART (which has validity) are the basis for the approach.

*Disadvantages of JHEDI*

- This is a BNFL proprietary tool and therefore is not publicly available.
- It has less functionality than HRMS.

### 7.2.14 Methode d'Evaluation de la Realisation des Missions Operateur pour (MEMORS)

#### 7.2.14.1 Background

In English this means: Assessment method for the performance of safety operation. It was developed for Electricite de France (EdF) in 1999. MERMOS is a second-generation HRA method that is reported to be an improvement of EdF's previous methods. It is designed to guide EdF's analysts in taking human factors aspects into account in the 'level 1' PRA for units of the N4 series (the most recent type of French reactors). The method only consider emergency operation during the four hours after the incident initiator, as it is assumed that four hours after the initiator the crisis support team will prevent or recover any human failure. The main underlying concept of MERMOS is the 'human factor mission'. For each initiator, a functional analysis will determine the 'missions' that have to be performed to recover or mitigate the accident. The human factors mission refers to safety critical actions that the operating system has to initiate and carry out the situation. MERMOS has only been applied to the nuclear industry.

#### 7.2.14.2 Description of the technique

- MERMOS considers that the performance of the human factors mission is the responsibility of what is termed the 'emergency operations system' (EOS) - this comprise the operating crew, operating procedures, the man-machine interface, the formal organization and the workplace.
- Instead of assuming that human error is a decisive element of failure, in MERMOS it is embedded within the EOS as one of the determinants of inadequate performance.

The method is divided into two modules:

1. **Module 1 -** Identification and definition of the HF mission through functional analysis. From a functional analysis of the state of the plant after the initiator, the analyst describes the characteristics of each HF mission and their context in a standard form.

2. **Module 2 -** Qualitative and quantitative analysis of the HF missions.The qualitative analysis aims to identify as many scenarios as possible leading to the HF mission failure. The mission failure event occurs if one of the scenarios described in the quantitative analysis occurs and leads to failure. The failure scenarios of a mission are a set of events and these sets are exclusive. The probability of a mission failure is therefore the sum of the probabilities of occurrence of the failure of each of the scenarios described for a mission. The probability of occurrence of the mission failure, according to a given scenario, can be worked out from the probability of occurrence of each event in the scenario.

To make the method more user friendly to the analysts, the developers are building a database of HF mission and failure scenarios.

#### 7.2.14.3 Advantages and Disadvantages

*Advantages of MERMOS*

- It attempts to deal with important underlying concepts of HRA, as it moves away from focusing on individual error, and instead considers the operating system as a whole.
- EdF report good results from using MERMOS.

*Disadvantages of MERMOS*

- It is currently only for emergency operations and does not propose a normal operation model, to assess the performance of operators in non-emergency situations.
- Its validity and reliability have yet to be established as it is still a new tool.
- This tool is not freely available for review and use.

### 7.2.15 Nuclear Action Reliability Assessment (NARA)

#### 7.2.15.1 Background

NARA was developed for the nuclear power company, British Energy by a consortium of HRA experts. It has been developed using the HEART methodology as its basis, using more recent data, and tailored to the needs of UK Nuclear Power Plant Probabilistic Safety Assessments and HRAs. Like HEART, NARA consists of Generic Task Types (GTTs) and Error Producing Conditions (EPCs). It is a British Energy proprietary tool and is a nuclear specific modified version of HEART; it is not applicable to other domains.

#### 7.2.15.2 Description of the technique

The key elements of the NARA approach are:

- Classify the task analysis into one of the Generic Task Types and assign the nominal Human Error Probability to the task.
- Decide which EPCs may affect task reliability and then consider the assessed proportion of affect for each EPC. Then calculate the task HEP.

The same formula, used in HEART, is used in NARA for deriving the HEP. Furthermore, the list of GTTs in NARA is partly a subset of the original HEART GTTs. There are fourteen GTTs and eighteen EPCs quantified in NARA.

#### 7.2.15.3 Advantages and Disadvantages

*Advantages of NARA*

- Has a prototype approach to error of commission quantification as the new feature
- Features an approach to quantifying operator reliability in relation to long time-scale events.

*Disadvantages of NARA*

- NARA is not publicly available and is nuclear specific - not suitable for other sectors.

### 7.2.16 Paired Comparisons (PC)

#### 7.2.16.1 Background

The PC method originated in 1960s. It is a means of defining preferences between items (human errors) and asks experts to make judgements, albeit of a relatively simple variety. In this method, subject matter experts make simple comparative judgements rather than absolute judgements. Each expert compares all possible pairs of error descriptions and decides, in each case, which of the two errors is more probable. For n tasks, each expert makes n(n-1)/2 comparisons. From combining the comparison made by different experts a relative scaling of error likelihood can be constructed. This is then calibrated using using a logarithmic calibration equation, which requires that the human error probabilities be known for at least two of the errors within the task set. The PC method has been applied to the transport and nuclear industries.

#### 7.2.16.2 Description of the technique

The sixteen steps to be followed when using this tool given in Kirwan (1994) are:

- Define the tasks involved. The tasks should be defined simply, unambiguously and comprehensively.
- Incorporate the calibration tasks. At least two of the descriptions. for which the HEPs are known, should be included in the task set.
- Select the expert judges. The subject matter experts should have experience of the tasks being assessed.
- Prepare the exercise. Each pair of the tasks should be presented on its own, so that the expert only considers pair at any one time.
- Brief the experts. The subject matter experts should be briefed on the purpose of the study and nature of the task to be assessed.
- Carry out paired comparisons. When the subject matter experts are carrying out the paired comparisons it is useful for them to have the support of the analyst so that they are able to seek and receive clarification of the tasks.
- Derive the raw frequency matrix. Each cell in the matrix indicates the number of subject matter experts tat considered one event as being more likely than an alternative event.
- Derive the proportion matrix. The next step is to normalise the scores by determining the proportion of subject matter experts that have concluded that one event was more likely than the other.
- Derive transformation X-matrix. The next step is to convert these scores into their equivalent unit normal deviate, using normal distribution tables.
- Calculate the column-difference Z-matrix. This is a simple calculation of the difference between the adjacent column values.
- Calculate the scale values. The average column differences are converted into a linear scale by setting the most preferred task to zero.
- Estimate the calibration points. Ideally error probabilities for the calibration of tasks should be estimated form the frequencies obtained via actual observations, otherwise the APJ can be adopted.

- Transform the scale values into probabilities. The scale values are transformed into human error probabilities (HEPs) via a method of simultaneous equations and using the logarithmic relationship:

$$HEP = ax + b.$$

- Determine the within-judge level of consistency - Experts can exhibit internal consistencies that need to be accounted for.
- Determine the inter-judge level of consistency. As with APJ, to determine the level of consistency between subject matter experts an analysis of variance could be performed.
- Estimate the uncertainty bounds.

#### 7.2.16.3 Advantages and Disadvantages

*Advantages of PC*

- The technique makes it possible to determine whether or not individual judges are qualified to make judgements about a particular datum or data set.
- Even without calibration, the tool provides a useful means of deriving a measure of the relative importance of different human errors or human events.
- With a small number of tasks and a set of rapidly available experts, the technique can be applied fairly quickly, when carried out on a computer.
- Subjective knowledge can be profitably extracted from comparative judgements, provided that the assumptions of the methods are upheld and the value of the human judgement proves greater than the value of any confusion arrived at via direct numerical assessments.

*Disadvantages of PC*

- The tasks being considered may be too complex to allow an easy comparison.
- The comparisons made may not be independent of each other.
- The tasks may not be homogeneous.
- If the number of comparisons is large, the judges may become tired and therefore, carry out later comparisons differently from earlier ones.

### 7.2.17 Success Likelihood Index Method (SLIM)

#### 7.2.17.1 Background

SLIM was first developed by Embrey & Kirwan (1983) for use in the US Nuclear Regulatory Commission. It is a decision-analytic approach to HRA which uses expert judgement to quantify PSFs. Such factors are used to derive a Success Likelihood Index (SLI), a form of preference index, which is calibrated against existing data to derive a final HEP. Significant PSFs for the context under study are are selected experts. In this method, relative success likelihoods are established for a range of tasks, and then calibrated using a logarithmic transformation.

The technique consists of two modules:

- Multi-Attribute Utility Decomposition (MAUD) is a computer based procedure which scales the relative success likelihood in performing a range of tasks;
- Systematic Approach to the Reliability Assessment of Humans (SARAH) which calibrates these success scores with tasks with known HEP values to provide an overall figure.

SLIM has been used in the nuclear and chemical industries.

#### 7.2.17.2 Description of the technique

- The basic rationale underlying SLIM is that the likelihood of an error occurring in a particular situation depends on the combined effects of a relatively small set of PSFs.
- It is assumed that an expert judge (or judges) is able to assess the relative importance (or weight) of each PSF with regard to its effect on reliability in the task being evaluated.
- It is also assumed that, independent of the assessment of relative importance, the judge(s) can make a numerical rating of how good or how bad the PSFs are in the task under consideration, where 'good' or 'bad' mean that the PSF will either enhance or degrade reliability.
- The relative importance of weights are then multiplied together for each PSF and the resulting products are then summed to give the SLI. The SLI is a quantity, which represents the overall belief of the judge(s), regarding the positive or negative effects of the PSFs on the likelihood of success for the task under consideration.
- The SLI for each task is calculated using the following formula:

$$SLI_j = \sum_{i=1}^{x} R_{ij} W_i$$

where $SLI_j$ is the SLI for task $j$ ; $W_i$ is the importance weight for the $i^{th}$ PSF; $R_{ij}$ is the scaled rating of task $j$ on the $i^{th}$ PSF; and $x$ represents the number of PSSFs considered.

- The SLIs previously calculated need to be transformed to HEPs as they are only relative measures of the likelihood of success of each of the ordered tasks. The relationship suggested is a logarithmic relationship of the form,

$$Log(P) = aSLI + b$$

where $P$ is the probability of success and $a$ and $b$ are empirically derived constants.

- Uncertainty bounds can be estimated using expert judgement methods such as APJ.

- By considering the PSFs which may be altered, the degree to which they can be changed and the importance of the PSFs, it is possible to conduct a cost-benefit analysis to determine how worthwhile suggested improvements may be, i.e. what-if analysis, the optimal means by which the calculated HEPs can be reduced.

#### 7.2.17.3 Advantages and Disadvantages

*Advantages of SLIM*

- The technique is highly visible and audit-able and is also sophisticated and well developed

- It allows gross cost-benefit evaluations to take place.

- The method has a sound basis in decision theory and a reasonably high level of theoretical validity.

- Sensitivity analysis is relatively straightforward to execute.

- It can be used to evaluate HEPs for discrete tasks as well as at a higher, more holistic level.

*Disadvantages of SLIM*

- Extensive use of expert judgement is required.

- It is prone to biases, which can significantly undermine the validity of the technique.

- The method by which PSFs are selected is somewhat arbitrary.

- The validity of the logarithmic transformation has not been established i.e it requires some empirical and experimental justification.

- SLIM is relatively resource-intensive to carry out compared to other HRA methods.

- The absolute HEPs depend on the two calibration events used to establish the linear scaling parameters. Hence errors in these two HEPs will induce errors in the other calculated HEPs.

### 7.2.18 Simplified Plant Analysis Risk Human Reliability Assessment (SPAR-H)

#### 7.2.18.1 Background

SPAR-H was developed for the US Nuclear Research Commission, Office of Regulatory Research in 1999. SAPR-H is a revision of the Accident Sequence Precursor (ASP) HRA screening method; it can be used as both a screening method and a detailed analysis method. The technique works by telling the analyst to decompose each task into either a diagnosis or an action subtask. Action tasks - carrying out one or more activities indicated by diagnosis, operating rules or written procedures. for example, operating equipment, performing line-ups starting pumps etc. Diagnosis tasks - reliance on knowledge and experience to understand existing conditions, planning and priotising activities, and determining appropriate courses of action. It includes worksheets that allow the analyst to provide complete descriptions of the tasks and capture task data in a standard format. This tool requires the analyst to determine the system activity type and then provides HEPs for the four combinations of the error type and system activity type. It has been successfully applied to risk informed regulatory sectors, and no evidence was found of the method being used in other sectors.

#### 7.2.18.2 Description of the technique

SPAR-H has been reported to do the following:

- Decompose probability into contributions from diagnosis failures and action failures.;
- Accounts for the context associated with human failure events (HFEs) by using PSFs, and dependency assignment to adjust a base-cease HEP;
- Uses pre-defined base-case HEPs and PSFs, together with guidance on how to assign the appropriate value of the PSF;
- Employs a beta distribution for uncertainty analysis, which can mimc normal and log normal distributions, but it has the advantage that probabilities calculated with this approach range from 0 to 1 and;
- Uses designated worksheets to ensure analyst consistency.

Eight PSFs were identified as being capable of influencing human performance and are accounted for in the SPAR-H quantification process. They are: Available time; Stress and Stressors; Experience and training; complexity; Ergonomics (and Human interface); Procedures; Fitness for duty; Work processes.

#### 7.2.18.3 Advantage and Disadvantages

*Advantages of SPAR-H*

- A simple underlying model makes SPAR-H relatively easy to use and results are traceable.
- The THERP-like dependence model can be used to address both subtask and event sequence dependence.
- The eight PSFs included cover many situations where more detailed analysis is not required.

*Disadvantages of SPAR-H*

- This tool does not provide guidelines for task decomposition, the analyst has the responsibility to identify how many diagnosis and/ or action activities should be considered for a given task. This affects the HEP of the task.
- The degree of solution of the PSFs may be inadequate for detailed analysis.
- The method may not be more appropriate where more realistic, detailed analysis of diagnosis errors is needed.

### 7.2.19 Technica Empirica Stima Errori Operatori (TESEO)

#### 7.2.19.1 Background

TESEO was developed by Bello & Columbari (1980) with the intention of using it for the purpose of conducting HRA of process industries. The methodology is relatively straightforward and is easy to use but is also limited; It is useful quick overview HRA assessments as opposed to those which are highly detailed and in-depth. Authors of review papers question the theoretical background of this method, and it is not considered to be accurate.

#### 7.2.19.2 Description of the technique

This is a time based model which describes the probability of a system operator's failure as a multiplicative function of five main factors. These factor are as follows:

1. $K_1$ : The type of task to be executed
2. $K_2$ : The time available to the operator to complete the task
3. $K_3$ : The operator's level of experience/characteristics
4. $K_4$ : The operator's state of mind

5. $K_5$ : The environmental and ergonomic conditions prevalent

Through the use of these figures, an overall HEP can be calculated with the formulation provided below:

$$K_1 \times K_2 \times K_3 \times K_4 \times K_5.$$

When putting this technique into practice, it is necessary for the designated HRA assessor to thoroughly consider the task requiring assessment and therefore also consider the value for $K_n$ that applies in the context. Once this value has been decided upon, the standard tables that take account of the method in which the HEP is derived, are then consulted from ch a related value for each of the identified factors is found in order to allow the HEP to be calculated French et al. (2009).

#### 7.2.19.3 Advantages and Disadvantages

*Advantages of TESEO*

- TESEO is typically quick and straightforward in comparison to other HRA tools, not only in producing a final result, but also in sensitivity analysis.
- It is widely applicable to various control room designs or with procedures with varying characteristics.

*Disadvantages of TESEO*

- There is limited work published with regards to the theoretical foundations of this technique, in particular relating to the justification of the five factor methodology. It uses only five factors to describe the full range of error producing conditions, which fails to be highly realistic.
- The values of the fives factors are unsubstantiated and the suggested multiplicative relationship has no sufficient theoretical or empirical evidence for justification purposes.

### 7.2.20 Technique for Human Error Rate Prediction (THERP)

#### 7.2.20.1 Background

THERP is a first generation methodology which was developed by Swain over a lengthy period of time. Swain & Guttmann (1983) then prepared the THERP handbook for US Nuclear Regulatory Commission. Essentially, the THERP handbook presents tabled entries of HEPs that can be modified by the effects of plant specific PSFs, using other tables. It was developed in the Sandia Laboratories for the US Nuclear Regulatory Commission. THERP methodology relies on a large human reliability database containing HEPs which is based upon both plant data and expert judgements. It is referred to as a 'decomposition' approach in that its descriptions of task, have a higher degree of resolution than many other techniques. It is also a logical approach and one that puts a larger degree of emphasis on error recovery than most other techniques. It was developed for probabilistic risk assessments of nuclear power plants but has been applied to other sectors such as offshore and medical.

#### 7.2.20.2 Description of the technique

The methodology for the THERP technique is broken down into five main stages:

- Define the system failures of interest which include functions of the system in which human error has a greater likelihood of influencing the probability of a fault, and those which are of interest to the risk assessor.
- List and analyse the related human operations, and identify human errors that can occur and relevant human error recovery modes. This stage must involve a task and human error analysis.
- Assess the relevant error probabilities, which includes entering HEPs for each sub-task into the tree and making sure that all failure branches have probabilities. Human Reliability Analysis Event Trees (HRAETs) provide the function of breaking down the primary operator tasks into finer steps which are represented in the form of success and failures.
- Estimate the effects of human error on the system failure events by inserting the HEPs into the full system's fault event tree which allows human factors to be considered within the context of the full system.
- Recommend changes to the system and recalculate the system failure probabilities by incorporating error recovery paths into the event tree. This will help the assessor to identify errors that can be reduced.

When all the HEPs are $0.01$ or smaller, the exact failure equation can be approximated by summing only the primary failure paths, ignoring all the success limbs, Bell & Holroyd (2009).

#### 7.2.20.3 Advantages and Disadvantages

*Advantages of THERP*

- It has a powerful methodology that can be audited, i.e it can be tailored to the requirements of a particular assessment.
- It is compatible with PRA; the methodology of the technique means that it can be readily integrated with fault tree reliability methodologies.
- It is transparent, structured and provides a logical review of the human factors considered in a risk assessment

- It is founded on a database of information that is included in the THERP handbook.
- THERP is a unique methodology in the way that it highlights error recovery and it also quantitatively models a dependency relation between the various actions or errors.
- It is usable within a wide range of human reliability domains and has a degree of face validity.

*Disadvantages of THERP*

- THERP is very resource intensive and time consuming.
- It does not offer enough guidance on modelling scenarios and the impact of PSFs on performance.
- As a first generation HRA tool, it does not take account of context.
- Large discrepancies have been found in practice with regards to different analysts assessment of the risk associated with the same task.
- The THERP HRAETs implicitly assume that each sub-task's HEP is independent from all others.

The major problem faced by the HRA community is that, despite the many methodologies that exist, there has not been a single method that can can be used for any context within PRA. This is noted in Forester et al. (2004) when they mention that the ATHEANA methodology provides a detailed search process for identifying important human actions and the contexts that can lead to either their success or failure. However, an accepted model of human behaviour suitable for formally supporting the quantification of human actions does not exist. It is also argued in French et al. (2009) that the variety of tasks that HRA is called upon to perform and the range of contexts in which it is applied are so great that it would be optimistic in the extreme to expect one method to be sufficient to meet these requirements. They believe that what is needed is a portfolio of HRA methods, and that the appropriateness of any HRA method may depend on the decision context that is being assessed. Therefore, to assess the appropriateness of any HRA method to a decision context, it is necessary to understand how each models such cognitive, behavioural and organizational effects and their strengths and weaknesses in doing so.

Validation of some of the HRA techniques, which refers to the degree of agreement between HRA technique estimates of HEPs, and real known values (unknown to the appliers of the technique(s)), has been done in Kirwan (1997). In Kirwan (1997) major criteria for validating HRA techniques and categories for classifying validation studies are highlighted.

# Chapter 8

# Concluding Remarks

## 8.1 Conclusion

The Jeffreys and a general divergence priors for the following distributions were derived: exponential, Rayleigh, gamma and Weibull. The resulting posterior distributions were derived, and it was shown theoretically that these posteriors are proper. For the gamma and Weibull distributions, the marginal and conditional posterior distributions were derived, since there are two parameters involved in these distributions. This thesis discussed the results obtained from performing simulation studies. The simulation studies were carried out to compare the performance of the Jeffreys prior and the general divergence prior. Coverage rates and average interval lengths for the priors were obtained, for all models. Both priors were found to produce good coverage rates for the exponential and the Rayleigh distributions. The two priors gave relative good coverage for the gamma and Weibull distributions. The coverage rate produced by the Jeffreys prior tend to be conservative. The interval lengths obtained from the priors were shown to increase as value of the parameter increase. Both priors produce, on average, the same average interval lengths. Based on the simulations studies, the divergence prior is the preferred one. A Metropolis-Hastings algorithm was introduced to simulate from the marginal posterior distributions, in the case of the gamma and Weibull distributions.

An application of the gamma and Weibull distribution to lifetime data was considered. The analysis was carried out in OpenBugs®. Statistics on the posteriors for the parameters were obtained. The performance of the priors was found to be similar in estimating the parameters. Both priors produce kernel posterior densities that are skewed for the gamma and Weibull distribution. Both priors had history/ trace plots that showed convergence of the MCMC. Model comparison for the gamma and the Weibull model showed that both prior distributions produce models that can be used to make good short-term predictions for the gamma and the Weibull distribution. The model resulting from the Weibull distribution and the general divergence prior yielded the smallest DIC. The case studies on HRA were presented. They showed that human error is the major contributor to failure of the system.

## 8.2 Short Comings

Only two vague priors were considered. A more in-depth comparison could be done, where various other vague priors are also considered. The simulation studies for the gamma and Weibull distributions were very limited.

## 8.3 Future Research

This study investigated the performance of two non-informative priors - the Jeffreys prior and the general divergence prior. For future research, the following priors could be included: probability matching prior, reference prior and maximal data information prior. Also, this study used the M-H algorithm to simulate values from the posteriors. The case where the full conditional posterior distribution of the parameter is known (the gamma is a an example), the Gibbs sampler can be used or a weighted Monte-Carlo method. Furthermore, the study focused only on four lifetime distributions. There are other lifetime distributions that can be used to model failure times. The lognormal distribution is an example of such distributions.

This study only focussed on vague priors, and used a textbook example to illustrate the usefulness of the models. It would be interesting to get involved in a real-life study where actual subjective information could be used to determine the priors.

The chapter on human reliability only focussed on a review of current models, and then some case studies. In future research focus will be on improving some of these current models and developing some new models. Also, to get involved in an actual case study applicable to cases and studies in South Africa. For example the news headlines in February 2015: “Human error trips Koeberg’s power” and “Koeberg human error cost economy R7.5bn”. Koeberg is a nuclear power station in South Africa, currently the only one in South Africa. It is owned by the only national electricity supplier, Eskom.

# References


Abramowitz, M. & Stegun, I. A. (1964). *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Washington DC: U.S. Government Printing Office, 1st edition.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716 – 723.

Bedford, T. & Cooke, R. (2001). *Probabilistic Risk Analysis: Foundations and Methods*. Cambridge: Cambridge University Press, 1st edition.

Bell, J. & Holroyd, J. (2009). *Review of human reliability assessment methods*. Derbyshire: Health and Safety Laboratory, 1st edition.

Bello, G. C. & Columbari, V. (1980). The human factors in risk analyses of process plant: The control room model 'TESEO'. *Reliability Engineering*, 1(1), 3 – 14.

Bernado, J. M. & Smith, A. F. M. (2001). *Bayesian theory*. England: John Wiley & Sons, 1st edition.

Birolini, A. (2007). *Reliability Engineering Theory and Practice*. Berlin: Springer, 5th edition.

Boring, R. L., Griffith, C. D., & Joe, J. (2007). The measure of human error: Direct and performance shaping factors. *IEEE 8th Human Factors and Power Plants and HPRCT 13th Annual Meeting*, 170 - 176.

Box, G. E. & Tiao, G. C. (1992). *Bayesian Inference in Statistical Analysis*. England: John Wiley & Sons, 1st edition.

Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of simulation studies in medical statistics. *Satistics in medicine*, 25(24), 4279 – 4292.

Crawley, F. & Tyler, B. (2015). *HAZOP: Guide to best practice*. Amsterdam: Elsevier, 3rd edition.

Cullen, A. C. & Frey, H. C. (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs*. New York: Plenum Press, 1st edition.

Di Pasquale, V., Iannone, R., Miranda, S., & Riemma, S. (2013). *Operations Management*, chapter An overview of human reliability analysis techniques in manufacturing operations. InTech: Croatia, 332 - 337.

Dorner, D. (1996). *The logic of failure: Recognizing and avoiding error in complex situations*. Cambridge: Perseus Books, 1st edition.

Ellis, G. R. & Holt, A. (2009). A practical application of human - HAZOP for critical procedures. *Hazards XXI Symposium Series*, 155, 434 – 439.

Embrey, D. E. & Kirwan, B. (1983). A comparative evaluation of three subjective human reliability quantification techniques. *Proceedings of the annual ergonomics society conference*, 137 - 142.

Forester, J., Bley, D., Cooper, S., Lois, E., Siu, N., & Kolaczkowski, A.and Wreathall, J. (2004). Expert elicitation approach for performing ATHEANA quantification. *Reliability Engineering & System Safety*, 83, 207 – 220.

French, S., Adhikari, S., Bayley, C., Bedford, T., Busby, J., Cliffe, A., Devgun, G., Eid, M., Keshvala, R., Pollard, S., Soane, E., Tracy, D., & Wu, S. (2009). *Human reliability analysis: A review and critique*. Technical Report 589, The University of Manchester: Manchester Business School.

Gertman, D. I., Blackman, H. S., Haney, L. N., Seidler, K. S., & Hahn, H. A. (1992). INTENT: a method for estimating human error probabilities for decisionbased errors. *Reliability Engineering & System Safety*, 35(2), 127 – 136.

Ghosh, M. (2011). Objective priors: An introduction for frequentists. *Statistical Science*, 26(2), 187 – 202.

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

Groth, K. M. & Swiler, L. P. (2013). Bridging the gap between HRA research and HRA practice: A Bayesian network version of SPAR-H. *Reliability Engineering and System Safety*, 115, 33 – 42.

Hamada, M. S., Wilson, A., Reese, C. S., & Martz, H. (2008). *Bayesian Reliability*. New York: Springer-Verlag, 1st edition.

Hannaman, G. W., Spurgin, A. J., & Lukic, Y. D. (1984). Human cognitive reliability model for PRA analysis. *Electric Power Research Institute*.

Hoff, P. D. (2009). *A first course in Bayesian Statistical methods*. New York: Springer-Verlag, 1st edition.

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method (CREAM)*. Elsevier Science, 1st edition.

Howland, A. H. (1980). Hazard analysis and the human element. *Preprints EFCE 3rd International Symposium Loss Prevention and Safety Promotion in the Process Industry, Basel, Swiss*.

Hoyland, A. & Rausand, M. (2009). *System Reliability Theory: Models and Statistical Methods*. New Jersey: John Wiley & Sons, 2nd edition.

International Atomic Energy Agency (1991). The international chernobyl project: Technical report. *IAEA: Vienna*.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 186(1007), 453 – 461.

Kemeny, J. G., Babbitt, B., McPherson, H. C., Haggerty, P. E., Peterson, R. W., Lewis, C. Pigford, T. H., Marks, P. A., Taylor, T. B., Marrett, C. B., Trunk, A. D., & McBride, L. (1979). The accident at Three Mile island. *US GPO: Washington DC*, 1 - 179.

Kirwan, B. (1994). *A Guide to Practical Human Reliability Assessment*. London: Taylor /& Francis, 1st edition.

Kirwan, B. (1997). Validation of human reliability assessment techniques: Part 1 - Validation issues . *Safety Science*, 27(1), 25 – 41.

Kirwan, B. & James, N. J. (1989). A human reliability management system.

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

Lawless, J. F. (1982). *Statistical Model and Methods for Lifetime Data*. New York: John Wiley & Sons, 1st edition.

Liseo, B. (1990). *Elimination of Nuisance Parameters with Reference Noninformative Priors*. Technical Report 90-58C, Purdue University.

Marples, D. R. (1997). Nuclear power in the former USSR: historical and contemporary perspectives. *Nuclear energy and security in the former Soviet Union*.

Moala, F. A., Ramos, P. L., & Achcar, J. A. (2013). Bayesian inference for two-parameter gamma distribution assuming different noninformative priors. *Revista Colombiana de Estadisticia*, 36(2), 321 – 338.

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

Rail Accident Investigation Branch (RAIB) (2007). Progress report: Derailment at grayrigg. *Cumbria, Department of Transport.*

Rasmussen, J. (1983). Skills, rules, knowledge; signals, signs and symbols and other distinctions in human performance models. *IEEE Transactions on Systems, Man and Cybernetics*, 13(3).

Reer, B. (1997). Conclusions from occurrences by descriptions of actions (CODA). *Studies of risk and hazard. Annual meeting of the society for risk-analysis Europe, New Risk Frontiers, Center for risk research. Stockholm*, 370 - 379.

Reer, B. & Dang, V. N. (2007). *The commission errors search and assessment (CESA) method*. Technical Report 07-03, Paul Scherrer Institut.

Schuller, J. C. H. (1997). *Methods for Determining and Processing Probabilities*. The Hague: Committee for the Prevention of Disasters, 2 edition.

Singpurwalla, N. D. (2006). *Reliability and Risk: A Bayesian Perspective*. England: John Wiley & Sons, 1st edition.

Snelson, B. (2005). Fractured pipe with loss of primary containment in the THORP feed clarification cell. *Board of Inquiry Report: British Nuclear Fuels Limited*, 35(5), 1 – 34.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measure of model comcomplex and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583 – 639.

Swain, A. D. & Guttmann, H. E. (1983). *Handbook of Human-Reliability Analysis with Emphasis on Nuclear Power Plant Applications*. USA: Sandia National Laboratories, 1st edition.

Syversveen, A. R. (1998). *Noninformative Bayesian priors. Interpretation and problems with construction and applications*. Technical Report Statistics 3, Department of Mathematical Sciences, NTNU, Trondheim.

Williams, J. C. (1986). HEART - A proposed method for assessing and reducing human error. *Proceedings of the 9th Advances in Reliability Technology Symposium, Birmingham, AL.*

Williams, J. C. (1987). Paris air disaster. *Safety and Reliability Directorate, UKAEA, Warrington.*

Yang, R. & Berger, J. O. (1996). A catalog of noninformative priors. *Institute of Statistics and Decisions Science, Duke University.*

Zimolong, B. (1992). Empirical evaluation of THERP, SLIM and ranking to estimate HEPs. *Reliability Engineering & System Safety*, 35(1), 1 – 11.

# Appendix A: Derivation of the Fisher Information

This is defined, for a single parameter $\theta$, as

$$I(\theta) = -\mathrm{E}\left[\frac{\partial^2 \ln L(\theta \mid t)}{\partial \theta^2}\right]$$

## A.1 Fisher information for the exponential

The log-likelihood is given by

$$l(\lambda \mid \underline{t}) = n \ln \lambda - \lambda \sum_{i=1}^{n} t_i$$

the first and the second derivative of the log-likelihood function for $\lambda$ are

$$\frac{\partial l(\lambda \mid \underline{t})}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i$$

$$\frac{\partial^2 l(\lambda \mid \underline{t})}{\partial \lambda^2} = -\frac{n}{\lambda^2}$$

And

$$\begin{aligned} I(\lambda) &= -\mathrm{E}\left[\frac{\partial^2 l(\lambda \mid \underline{t})}{\partial \lambda^2}\right] \\ &= \frac{n}{\lambda^2}. \end{aligned}$$

## A.2 Fisher information for the Rayleigh

The log-likelihood function is

$$l(\lambda \mid \underline{t}) = n\ln 2 + 2n\ln\lambda + \sum_{i=1}^{n}\ln t_i - \lambda^2\sum_{i=1}^{n}t_i^2$$

and the first and second derivative of the log-likelihood function are

$$\frac{\partial l(\lambda \mid \underline{t})}{\partial\lambda} = \frac{2n}{\lambda} - 2\lambda\sum_{i=1}^{n}t_i^2$$

$$\frac{\partial^2 l(\lambda \mid \underline{t})}{\partial\lambda^2} = -\frac{2n}{\lambda^2} - 2\sum_{i=1}^{n}t_i^2$$

Therefore, since $\mathrm{E}[T^r] = \frac{1}{\lambda^r}\Gamma(2) = \frac{1}{\lambda^r}$, we have

$$\begin{aligned} I(\lambda) &= -\mathrm{E}\left[-\frac{2n}{\lambda^2} - 2\sum_{i=1}^{n}t_i^2\right] \\ &= \frac{2n}{\lambda^2} + 2\sum_{i=1}^{n}\mathrm{E}\left[t_i^2\right] \\ &= \frac{4n}{\lambda^2}. \end{aligned}$$

## A.3 Fisher information for the Weibull

The log-likelihood function for $(\alpha, \lambda)$ is given by

$$l(\alpha, \lambda \mid \underline{t}) = n\ln\alpha + \alpha n\ln\lambda + \sum_{i=1}^{n}(\alpha-1)\ln t_i - \sum_{i=1}^{n}(\lambda t_i)^{\alpha}$$

and the first and second partial derivatives of $l(\alpha, \lambda \mid \underline{t})$ are

$$\frac{\partial l(\alpha, \lambda \mid \underline{t})}{\partial \alpha} = \frac{n}{\alpha} + n \ln \lambda + \sum_{i=1}^{n} \ln t_i - \sum_{i=1}^{n} (\lambda t_i)^{\alpha} \ln(\lambda t_i)$$

$$\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \alpha^2} = -\frac{n}{\alpha^2} - \sum_{i=1}^{n} (\lambda t_i)^{\alpha} (\ln(\lambda t_i))^2$$

$$\frac{\partial l(\alpha, \lambda \mid \underline{t})}{\partial \lambda} = \frac{n\alpha}{\lambda} - \frac{\alpha}{\lambda} \sum_{i=1}^{n} (\lambda t_i)^{\alpha}$$

$$\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \lambda^2} = -\frac{n\alpha}{\lambda^2} - \frac{\alpha(\alpha-1)}{\lambda^2} \sum_{i=1}^{n} (\lambda t_i)^{\alpha}$$

$$\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \alpha \partial \lambda} = \frac{n}{\lambda} - \left[\frac{1}{\lambda} \sum_{i=1}^{n} (\lambda t_i)^{\alpha} + \frac{\alpha}{\lambda} \sum_{i=1}^{n} (\lambda t_i)^{\alpha} \ln(\lambda t_i)\right]$$

And

$$\mathrm{E}\left[\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \alpha^2}\right] = -\frac{n}{\alpha^2} - \sum_{i=1}^{n} \mathrm{E}\left[(\lambda t_i)^{\alpha} (\ln(\lambda t_i))^2\right]$$

$$\mathrm{E}\left[\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \lambda^2}\right] = -\frac{n\alpha}{\lambda^2} - \frac{\alpha(\alpha-1)}{\lambda^2} \sum_{i=1}^{n} \mathrm{E}\left[(\lambda t_i)^{\alpha}\right]$$

$$\mathrm{E}\left[\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \alpha \partial \lambda}\right] = \frac{n}{\lambda} - \frac{1}{\lambda} \sum_{i=1}^{n} \mathrm{E}\left[(\lambda t_i)^{\alpha}\right] - \frac{\alpha}{\lambda} \sum_{i=1}^{n} \mathrm{E}\left[(\lambda t_i)^{\alpha} \ln(\lambda t_i)\right]$$

Consider the random variable $Y=(\lambda T)^{\alpha}$, where $T\sim Weibull\,(\alpha,\lambda)$. Since

$$
\begin{aligned}
F_Y(y) &= \Pr(Y\le y)\\
&= \Pr\left((\lambda T)^{\alpha}\le y\right)\\
&= \Pr\left(T\le \frac{1}{\lambda}y^{1/\alpha}\right)\\
&= F_T\left(\frac{1}{\lambda}y^{1/\alpha}\right)\\
& \quad 1-e^{-y}\\
\Rightarrow \quad Y &\sim exp\,(\lambda=1)
\end{aligned}
$$

we have that

$$\int_0^{\infty}(\ln y)^r e^{-y}dy = r\text{th moment of } \ln Y=\gamma_r.$$

Thus,

$$
\begin{aligned}
\alpha \mathrm{E}\left[(\lambda t_i)^{\alpha}\ln(\lambda t_i)\right] &= \mathrm{E}\,[y\ln y]\\
&= \int_0^{\infty} y\ln(y)\,e^{-y}dy\\
&= -e^{-y}y\ln y\,|_0^{\infty}+\int_0^{\infty}\left(\ln y+\frac{y}{y}\right)e^{-y}dy\\
&= \int_0^{\infty}e^{-y}dy+\int_0^{\infty}(\ln y)\,e^{-y}dy\\
&= 1+\gamma_1
\end{aligned}
$$

Similarily,

$$\begin{aligned}
\alpha^2 \mathrm{E}\left[(\lambda t_i)^{\alpha}(\ln(\lambda t_i))^2\right] &= \mathrm{E}\left[y(\ln y)^2\right] \\
&= \int_0^{\infty} y(\ln(y))^2 e^{-y} dy \\
&= -e^{-y} y(\ln y)^2 \mid_0^{\infty} + \int_0^{\infty} \left((\ln y)^2 + 2\ln y\right) e^{-y} dy \\
&= 2\int_0^{\infty} (\ln y) e^{-y} dy + \int_0^{\infty} \left((\ln y)^2\right) e^{-y} dy \\
&= 2\gamma_1 + \gamma_2.
\end{aligned}$$

It can be readily seen that

$$\mathrm{E}\left[\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \alpha^2}\right] = -\frac{n}{\alpha^2} - \frac{n(2\gamma_1 + \gamma_2)}{\alpha^2}$$

$$\mathrm{E}\left[\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \lambda^2}\right] = -\frac{n\alpha^2}{\lambda^2}$$

$$\mathrm{E}\left[\frac{\partial^2 l(\alpha, \lambda \mid \underline{t})}{\partial \alpha \partial \lambda}\right] = -\frac{n}{\lambda} - \frac{n\gamma_1}{\lambda}$$

Therefore,

$$\begin{aligned}
I(\alpha, \lambda) &= -\mathrm{E}\begin{bmatrix} \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial \alpha^2} & \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial \alpha \partial \lambda} & \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial \lambda^2} \end{bmatrix} \\
&= \begin{bmatrix} \frac{n}{\alpha^2} + \frac{n(2\gamma_1+\gamma_2)}{\alpha^2} & \frac{n}{\lambda} + \frac{n\gamma_1}{\lambda} \\ \frac{n}{\lambda} + \frac{n\gamma_1}{\lambda} & \frac{n\alpha^2}{\lambda^2} \end{bmatrix}
\end{aligned}$$

and

$$\det(I(\alpha,\lambda)) = \frac{n^2}{\lambda^2} + \frac{n^2(2\gamma_1+\gamma_2)}{\lambda^2} - \frac{(n+n\gamma_1)^2}{\lambda^2}$$
$$= \frac{n^2\left(\gamma_2-\gamma_1^2\right)}{\lambda^2}$$

where $\gamma_2-\gamma_1^2 = var(\ln Y) = \text{constant}$.

## A.4 Fisher information for the Gamma

The log-likelihood function is

$$l(\alpha,\lambda \mid \underline{t}) = n\alpha\ln\lambda - n\ln\Gamma(\alpha) + \sum_{i=1}^{n}(\alpha-1)\ln t_i - \sum_{i=1}^{n} t_i$$

the first and second partial derivatives of $l(\alpha,\lambda \mid \underline{t})$ are

$$\frac{\partial l(\alpha,\lambda \mid \underline{t})}{\partial\alpha} = n\ln\lambda - n\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} + \sum_{i=1}^{n}\ln t_i$$
$$\frac{\partial l(\alpha,\lambda \mid \underline{t})}{\partial\lambda} = \frac{n\alpha}{\lambda} - \sum_{i=1}^{n} t_i$$

$$\frac{\partial^2 l(\alpha,\lambda \mid \underline{t})}{\partial\alpha^2} = n\left[\frac{(\Gamma'(\alpha))^2-\Gamma''(\alpha)\Gamma(\alpha)}{(\Gamma(\alpha))^2}\right]$$
$$\frac{\partial^2 l(\alpha,\lambda \mid \underline{t})}{\partial\lambda^2} = -\frac{n\alpha}{\lambda^2}$$

$$\frac{\partial^2 l(\alpha,\lambda \mid \underline{t})}{\partial\alpha\partial\lambda} = \frac{n}{\lambda}$$

where $\Gamma'(\alpha) = \frac{\partial\Gamma(\alpha)}{\partial\alpha}$ and $\Gamma''(\alpha) = \frac{\partial^2\Gamma(\alpha)}{\partial\alpha^2}$.

Now, let $\varphi(\alpha) = \frac{\Gamma'(\alpha)}{\Gamma(\alpha)}$ so that $\varphi'(\alpha) = \frac{\mathrm{d}\left(\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}\right)}{\mathrm{d}\alpha}$. Therefore,

$$
\begin{aligned}
I(\alpha, \lambda) &= -\mathrm{E}\begin{bmatrix} \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial\alpha^2} & \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial\alpha\partial\lambda} \\ \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial\alpha\partial\lambda} & \frac{\partial^2 l(\alpha,\lambda|\underline{t})}{\partial\lambda^2} \end{bmatrix} \\
&= \begin{bmatrix} n\varphi'(\alpha) & -\frac{n}{\lambda} \\ -\frac{n}{\lambda} & \frac{n\alpha}{\lambda^2} \end{bmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\det(I(\alpha, \lambda)) &= \frac{n^2\alpha\varphi'(\alpha)}{\lambda^2} - \frac{n^2}{\lambda^2}. \\
&= \frac{n^2(\alpha\varphi'(\alpha) - 1)}{\lambda^2}
\end{aligned}
$$

# Appendix B: Code for simulation studies

## B.1 MATLAB® Code for the exponential distribution

The code is written in MATLAB® and given below:

```
clc
clear
lambda1 = 0.1:0.1:30;

k1 = length(lambda1);
matrixJ = [];
AVGJ = [];
COUNTJ = [];
matrixG = [];
AVGG = [];
COUNTG = [];

for i=1:k1
countJ = 0;
countG = 0;
n = 10000;
IiJ=[];
THETAJ = [];
IiG=[];
THETAG = [];
for k=1:n
aJ = 0; %Jef
bJ = 0;
m = 50;
aG = 0.5; %GML
```

```
bG = 0;
x1 = exprnd(1/lambda1(i),1,m);
X1 = sum(x1);

%Jef
lambda_1J = gamrnd(m+aJ,1/(X1+bJ),1,n);
thetaJ = lambda_1J;
THETAJ = [thetaJ];
tempJ = sort(THETAJ);
lowerJ = tempJ((0.025*n));
upperJ = tempJ((0.975*n));
lengJ = upperJ-lowerJ;

IiJ = [IiJ lengJ];
if lowerJ < lambda1(i)
    if upperJ > lambda1(i)
        countJ = countJ + 1;
end
end

%GML
lambda_1G = gamrnd(m+aG,1/(X1+bG),1,n);
thetaG = lambda_1G;
THETAG = [thetaG];
tempG = sort(THETAG);
lowerG = tempG((0.025*n));
upperG = tempG((0.975*n));
lengG = upperG-lowerG;

IiG = [IiG lengG];
if lowerG < lambda1(i)
    if upperG > lambda1(i)
        countG = countG + 1;
end
end
end
avgJ = sum(IiJ)/n;
```

```
AVGJ = [AVGJ avgJ];
COUNTJ = [COUNTJ countJ];
avgG = sum(IiG)/n;
AVGG = [AVGG avgG];
COUNTG = [COUNTG countG];
end
matrixJ = [matrixJ; COUNTJ./n; AVGJ;];
matrixG = [matrixG; COUNTG./n; AVGG;];

Jeffreys = COUNTJ./n;
GML = COUNTG./n
J = Jeffreys';
G = GML';
X = [J G];
boxplot(X,'labels',{'Jeffreys','GML', })
title({'Boxplots showing the distribution of the coverage rates'})
ylabel('Coverage')
xlabel('priors')
```

## B.2 MATLAB® Code for the Rayleigh distribution

The code is written in MATLAB® and given below:

```
clc
clear
lambda1 = 0.1:0.1:30;

k1 = length(lambda1);
matrixJ = [];
AVGJ = [];
COUNTJ = [];
matrixG = [];
AVGG = [];
COUNTG = [];

for i=1:k1
countJ = 0;
```

```
countG = 0;
n = 20000;
IiJ =[];
THETAJ = [];
IiG =[];
THETAG = [];
for k=1:n
aJ = 0.5; %Jef
bJ = 0;
m = 50;
aG = 0.75; %GML
bG = 0;
x1 = raylrnd(1/((sqrt(2))*lambda1(i)),1,m);
X1 = sum(x1.^2);

%Jef
lambda2_1J = gamrnd(m+aJ,1/(X1+bJ),1,n);
lambda_1J = lambda2_1J.^0.5;
thetaJ = lambda_1J;
THETAJ = [thetaJ];
tempJ = sort(THETAJ);
lowerJ = tempJ((0.025*n));
upperJ = tempJ((0.975*n));
lengJ = upperJ-lowerJ;

IiJ = [IiJ lengJ];
if lowerJ < lambda1(i)
    if upperJ > lambda1(i)
        countJ = countJ + 1;
end
end

%GML
lambda2_1G = gamrnd(m+aG,1/(X1+bG),1,n);
lambda_1G = lambda2_1G.^0.5;
thetaG = lambda_1G;
THETAG = [thetaG];
```

```
tempG = sort(THETAG);
lowerG = tempG((0.025*n));
upperG = tempG((0.975*n));
lengG = upperG-lowerG;

IiG = [IiG lengG];
if lowerG < lambda1(i)
    if upperG > lambda1(i)
        countG = countG + 1;
end
end
end
avgJ = sum(IiJ)/n;
AVGJ = [AVGJ avgJ];
COUNTJ = [COUNTJ countJ];
avgG = sum(IiG)/n;
AVGG = [AVGG avgG];
COUNTG = [COUNTG countG];
end
matrixJ = [matrixJ; COUNTJ./n; AVGJ;];
matrixG = [matrixG; COUNTG./n; AVGG;];

Jeffreys = COUNTJ./n;
GML = COUNTG./n
J = Jeffreys';
G = GML';
X = [J G];
boxplot(X,'labels',{'Jeffreys','GML', })
title({'Boxplots showing the distribution of the coverage rates'})
ylabel('Coverage')
xlabel('priors')
```

## B.3 MATLAB® Code for the gamma distribution

The code is written in MATLAB® and given below:

```
clc
```

```
clear
rng default;
lambda1 = 1;
alpha1 = 2;

AVG_a = [];
COUNT_a = [];
AVG_l = [];
COUNT_l = [];
count_a = 0;
count_l = 0;
n = 1000;
Ii_a =[];
Ii_l = [];
THETA_a = [];
THETA_l = [];
for i=1:n
    m = 50;
    x1 = gamrnd(alpha1,lambda1,1,m);
    X1 = sum(x1);
    nsamples = n;
    shape = 0.25;
    %shape = 0.5; %Jef
    rate = log((((prod(x1))).^(-1)).*((sum(x1)).^m));
    pdf_ALPHA=@(x)((((x*psi(1,x)-1)^0.25)*(prod(x1.^x))*(gamma(x*m+0.5)))/(((
    %pdf_ALPHA=@(x)((((x*psi(1,x)-1)^0.5)*(prod(x1.^x))*(gamma(x*m)))/(((gamm
    proppdf_ALPHA= @(x,y) gampdf(x, shape,rate);
    proprnd_ALPHA = @(x) gamrnd(shape,rate);
    [MH_ALPHA,accept_ALPHA] = mhsample(0.01,nsamples,'pdf',pdf_ALPHA,'proppdf
    Palpha = [MH_ALPHA];
    ALPHA_post = mean(MH_ALPHA);
    sortAlpha = sort(Palpha);
    theta_a = Palpha;
    THETA_a = [theta_a];
    temp_a = sort(THETA_a);
    lower_a = temp_a((0.025*n));
    upper_a = temp_a((0.975*n));
```

```
        leng_a = upper_a-lower_a ;

        Ii_a = [Ii_a leng_a];
        if lower_a < alpha1
            if upper_a > alpha1
                count_a = count_a + 1;
            end
        end

        LAMBDA = gamrnd(alpha1.*m+0.5,1/X1,1,n); %LAMBDA = gamrnd(ALPHA_post.*m+(
        %LAMBDA = gamrnd(alpha1.*m,1/X1,1,n); %Jef
        theta_l = LAMBDA;
        THETA_l = [theta_l];
        temp_l = sort(THETA_l);
        lower_l = temp_l((0.025*n));
        upper_l = temp_l((0.975*n));
        leng_l = upper_l-lower_l ;

        Ii_l = [Ii_l leng_l];
        if lower_l < lambda1
            if upper_l > lambda1
                count_l = count_l + 1;
            end
        end
end
avg_a = sum(Ii_a)/n;
AVG_a = [AVG_a avg_a];
COUNT_a = [COUNT_a count_a];
avg_l = sum(Ii_l)/n;
AVG_l = [AVG_l avg_l];
COUNT_l = [COUNT_l count_l];
```

## B.4 MATLAB® Code for the Weibull distribution

The code is written in MATLAB® and given below:

```
clc
```

```
clear
rng default;
lambda1 = 0.5;
alpha1 = 2;

AVG_a = [];
COUNT_a = [];
AVG_l = [];
COUNT_l = [];
count_a = 0;
count_l = 0;
n = 1000;
Ii_a=[];
Ii_l = [];
THETA_a = [];
THETA_l = [];
for i=1:n
    m = 50;
    x1 = wblrnd(alpha1,lambda1,1,m);
    X1 = sum(x1);
    nsamples = n;
    shape = 1;
    rate = abs(log((prod(x1)).^(-1)));
    pdf_ALPHA=@(x) (x.^(m-1)).*(prod(x1.^x)).*((sum(x1.^x)).^(-m-0.5));
    %pdf_ALPHA=@(x) (x.^(m-1)).*(prod(x1.^x)).*((sum(x1.^x)).^(-m)); %Jef
    proppdf_ALPHA= @(x,y) gampdf(x, shape,rate);
    proprnd_ALPHA = @(x) gamrnd(shape,rate);
    [MH_ALPHA,accept_ALPHA] = mhsample(0.01,nsamples,'pdf',pdf_ALPHA,'proppdf
    Palpha = [MH_ALPHA];
    ALPHA_post = mean(MH_ALPHA);
    sortAlpha = sort(Palpha);
    xxhat = cumsum(MH_ALPHA.^2)./(1:nsamples)';
    plot(1:nsamples,xxhat)

    theta_a = Palpha;
    THETA_a = [theta_a];
    temp_a = sort(THETA_a);
```

```
    lower_a = temp_a((0.025*n));
    upper_a = temp_a((0.975*n));
    leng_a = upper_a-lower_a ;

    Ii_a = [Ii_a leng_a];
    if lower_a < alpha1
       if upper_a > alpha1
          count_a = count_a + 1;
       end
    end
    LAMBDA = gamrnd(m+0.5,1/(sum(x1.^alpha1)),1,n);
    %LAMBDA = gamrnd(m,1/(sum(x1.^alpha1)),1,n); %Jef
    theta_l = LAMBDA;
    THETA_l = [theta_l];
    temp_l = sort(THETA_l);
    lower_l = temp_l((0.025*n));
    upper_l = temp_l((0.975*n));
    leng_l = upper_l-lower_l ;

    Ii_l = [Ii_l leng_l];
    if lower_l < lambda1
       if upper_l > lambda1
           count_l = count_l + 1;
       end
    end
end
avg_a = sum(Ii_a)/n;
AVG_a = [AVG_a avg_a];
COUNT_a = [COUNT_a count_a];
avg_l = sum(Ii_l)/n;
AVG_l = [AVG_l avg_l];
COUNT_l = [COUNT_l count_l];
```

## B.5 OpenBugs® Code for the application in Chapter 5

The code is written in OpenBugs® and given below:

```
model

{

for ( i in 1:N){

TI[i]~dgamma(alpha ,lambda)

}

alpha~dunif(0,4)

lambda~dgamma(0.5,0.0001)

}

Data

list(TI=c(0.96, 4.15, 0.19, 0.78, 8.01, 31.75, 7.35, 6.50, 8.27, 33.91, 32.52

16.03, 4.85, 2.78, 4.67, 1.31, 12.06, 36.71, 72.89), N=19)

Inits

list(lambda = 0.01, alpha = 0.01)
```

## B.6 R® Code for the application in Chapter 5

The code is written in OpenBugs® and given below:

```
library(fitdistrplus)
library(logspline)

x <- c(37.50,46.79,48.30,46.04,43.40,39.25,38.49,49.51,40.38,36.98,40.00,
       38.49,37.74,47.92,44.53,44.91,44.91,40.00,41.51,47.92,36.98,43.40,
       42.26,41.89,38.87,43.02,39.25,40.38,42.64,36.98,44.15,44.91,43.40,
       49.81,38.87,40.00,52.45,53.13,47.92,52.45,44.91,29.54,27.13,35.60,
```

```
      45.34,43.37,54.15,42.77,42.88,44.26,27.14,39.31,24.80,16.62,30.30,
      36.39,28.60,28.53,35.84,31.10,34.55,52.65,48.81,43.42,52.49,38.00,
      38.65,34.54,37.70,38.11,43.05,29.95,32.48,24.63,35.33,41.34)

descdist(x, discrete = FALSE)
fit.weibull <- fitdist(x, "weibull")
fit.gamma <- fitdist(x, "gamma")

plot(fit.weibull)
plot(fit.gamma)
```