Universitetet
i Stavanger

FACULTY OF SCIENCE AND TECHNOLOGY

MASTER'S THESIS

Study programme/specialisation:

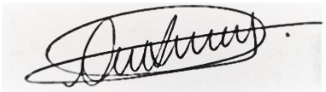
Risk Management

Spring, semester, 2019

Open

Author:

Anatoly Kurman Rivero



....
(signature of author)

Programme coordinator:

Supervisor(s): Jon Tømmerås Selvik   External Supervisor(s): Øystein Arild

Title of master's thesis:

Estimation of Expected Lifetime of Highly Reliable Systems using Bayesian Analysis

Credits: 30

Keywords:
Expected Lifetime
Reliability
Maximum Likelihood Estimation
Bayesian Analysis
Predictive Prior Distribution
Predictive Posterior Distribution
Highly Reliable Systems

Number of pages: 93

+ supplemental material/other: 17

Stavanger, June 14th 2019

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

Table A3. 2nd scenario with three different elicited expert opinions for prior parameters

GLOSSARY OF TERMS

AFR, Average failure rate

BA, Bayesian analysis

CDF, Cumulative density function.

CHF, Cumulative hazard function.

DF, Distribution function.

$IEEE$, Institute of Electrical and Electronics Engineers.

$l(x)$, Likelihood distribution.

MC, Monte Carlo algorithms.

MLE, Maximum likelihood estimation.

$MTTF$, Mean time to failure

NAN, Non numerical value.

PDF, Probability density function.

PfA, Frequentist probability for the occurrence of an event A.

$P(C \mid A)$, Conditional probability that event C will occur, given that event A occurs.

$p(\phi)$, Prior distribution.

t_f, Exact time at failure.

t_i, Time to inspection

.$XLSX$, Excel file format.

.BAK, Backup file format.

α, Alpha. Scale parameter of Weibull and Gamma distributions.

β, Beta. Shape parameter of Weibull and Gamma distributions.

λ, Lambda. Failure rate parameter of exponential function.

μ, Mu. Mean of a Lognormal distribution.

σ, Sigma, Standard deviation of a Lognormal distribution.

ϕ, Parameter(s) to be estimated with BA or MLE.

$\hat{\phi}$, Estimated parameter(s) of a given distribution.

$\hat{\Phi}_{prior}$, Estimated parameter(s) of a prior distribution.

$\hat{\Phi}_{MLE}$, Estimated parameter(s) of a likelihood distribution.

ABSTRACT

Nowadays, increasingly complex systems are critical due to the sectors and enterprises which they support. These are designed to be highly reliable and they are not expected to fail frequently. If a failure occurs, the safety, economical and operational consequences can be severe. Improvements and upgrades generate risk and uncertainty on their future performance. Therefore, there is a need for a procedure to estimate the expected lifetime of these highly reliable systems using a methodology based on available information.

The aim of this thesis is to obtain highly accurate reliability estimations for highly reliable systems using Bayesian analysis when few or no historical data is available. For this purpose, a model for reliability estimations of expected lifetime based on Bayesian analysis was created and tested. The model estimates the probability of survival, probability of failure, histograms and plots for four predefined statistical distributions. The estimations are based on available historical data of performance and elicited expert knowledge to create posterior sample data of the system using Montecarlo simulations. Some relevant examples are included to compare the results with another estimation method such as Maximum Likelihood Estimation.

Two main conclusions are derived; first, Bayesian analysis constitutes a powerful method to estimate the expected lifetime of highly reliable systems with high accuracy, compared to other methods such as Maximum Likelihood Estimation. Second, the model for reliability estimates provides decision support in a risk and operational context for maintenance or replacement.

ACKOWLEDGEMENTS

This report would not had been possible without the contribution and guidance of my supervisor Jon Tømmerås Selvik, professor Øystein Arild of the Oil & Energy institute of University of Stavanger and the personnel of AKRA Investments c.a. who provided the access to real data for a study case.

To my family, which has always been there for me. Gisela and Adriana you are the most reliable people I have ever met.

CHAPTER 1

1.1 Introduction

Nowadays there is an increase of aging and complex systems in many industries. This creates a need for them to be highly reliable under the variety of conditions in which they operate. Therefore, decisive steps to prevent failures that are considered critical must be taken.

One example is the increase of decommissioning and abandonment of wells in the oil and gas industry. Such systems, considered to be highly reliable, are aging and no monitoring or follow up of their performance has been systematically done. In addition, the negative consequences due to oil leakage or gas emission to the environment can be extensive.

These are the main reasons that indicate the need for a reliable method to estimate the expected lifetime of highly reliable systems. This method must combine available historical performance data and available expert knowledge. One method is Bayesian analysis.

Based on literature, there are several statistical distributions can be used for modelling the behaviour of most highly reliable systems. Among them are: Exponential, Weibull, Lognormal, Gamma, Beta, Poisson, Geometric, Normal and some more. However, this thesis is limited to the first four.

This thesis intends to give a relevant statistical background of estimation techniques and to explore the advantages and benefits of using Bayesian analysis as a method to estimate the expected lifetime of highly reliable systems. A computational tool to perform the estimation of expected lifetime of two case studies will be presented the corresponding practical model and results according to the characteristics of each case. The applicability of the methodology will be demonstrated on two examples with a discussion.

As a result of this report a model was created based on Bayesian analysis method and uses one of these four distribution methods in case of existing performance data.

1.2 Background

In the past decades, there has been an increase of complex processes and systems. Technological components, improved materials and the increasing demand for highly reliable systems require a methodology to predict the behaviour of such systems. In accordance with the latter, this report intends to obtain an innovative alternative for this methodology of statistical inference.

Common knowledge suggests that the expected lifetime of a highly reliable systems depends of several factors such as, operational techniques, quality of material and surrounding environment. However, in many cases a methodology to establish the expected lifetime of highly reliable systems that accounts for these factors has not been fully developed. Whereas such methodology relies in two fundamental pillars.

Previous expert knowledge of the system and data sets of historical performance. These pillars are the foundation for modelling behaviour of Bayesian analysis (BA) for systems like plugged of oil and gas bores, water dams, high voltage isolators in transmission lines or railroads that are treated as highly reliable.

Both prior knowledge and historical data contribute to determine a posterior statistical distribution through BA, that will help to predict the behaviour of any system modelled under these premises. Therefore, the use of BA instead of other methodologies such as Maximum Likelihood Estimation (MLE) when predicting the expected lifetime gains relevance for cases where historical data is not available, or the sample data is scarce to be considered representative.

1.3 Objectives

The main objective of this report is to develop and test a method to estimate the expected lifetime of highly reliable systems using Bayesian analysis. For that, the following activities must be performed.

1. Define the assumptions and scope, given the software tool and the type of reliability data available.
2. Develop the codes for both methods MLE and BA of expected lifetime estimation using the software tool Matlab R2018.

3. Obtain the expected lifetime and sensitivity analysis of a study case with MLE and BA using the developed codes in Matlab R2018 and real data available.
4. Discuss and compare the results of the example and study case.
5. Describe the possible applications of this thesis in real-life cases when estimating expected lifetime of highly reliable systems.

1.4 Content

This report is divided in six chapters, plus the appendix. The first chapter provides the background, objectives and content of the thesis. The second chapter presents a theoretical basis of statistical distributions, specifically for the four types of Probability Density Functions (PDF) handled throughout this thesis.

The third chapter covers the two estimation models used in this report BA and MLE, including an example of a random data set used to test the Matlab's codes. It also describes the types of reliability data including examples of a random data set and censored data set of failures times Moreover, a section for considerations of highly reliable systems that contains risk and uncertainty aspects.

In the fourth chapter, real study case of highly reliable systems is presented, with background, prior knowledge and historical data for estimation of expected lifetime using both methods BA and MLE. The results of this study case are presented and discussed in the fifth chapter, including a comparison between both methods and a sensitivity analysis.

The sixth and final chapter presents the conclusions and recommendations of the two study cases. Moreover, the applicability of BA to estimate the expected lifetime and other parameters for many real-life cases is discussed. Finally, several findings are presented for the reader to assess the relevance, accuracy and practicality of this method when estimating the expected lifetime and other relevant indicators for performance of highly reliable systems.

CHAPTER 2

2.1 Probability Models

To maximize system's performance and use the available resources in the most efficient way, it is important to predict the occurrence of failures (Modarres, 2010) Therefore, technologies and systems require lifetime prediction, often with only small samples available and a considerable degree of uncertainty. Hence, the need to use proven techniques to represent observations of the phenomena (Tobias, 2012). In addition to these observations, the use of numerical descriptive measures of a population called parameters, help to build a statistical function to predict the expected lifetime (Tobias, 2012)

Before discussing the MLE and BA approaches, it is useful to establish some previous definitions about probability models. Most of them are used for continuous sets of random variables. These are listed as follows:

- Probability density function
- Cumulative distribution function
- Reliability function
- Unreliability function
- Hazard function
- Cumulative Hazard function
- Average failure number
- Mean time to failure

One of the main definitions which is going to be used through this report is the PDF. A probability density function is that which satisfies these three conditions for a random set of continuous variables (Hamada, 2008).

1. $f(t) \geq 0$, the function must be greater or equal than zero.
2. $-\infty \leq t \leq \infty$, the random variable must be real.
3. $\int_{-\infty}^{\infty} f(t)dt = 1$, its integral must be equal to 1.

Another important definition is the CDF. The Cumulative Distribution Function expresses the cumulative probability of failure $F(t)$ and survival $R(t)$ (Hamada A., 2008). $F(t)$ is interpreted as the probability that a random element from a population fails by t hours. Hence that it takes a value equal or less to t (Hamada, 2008). The other interpretation is the fraction of all units in the population that fail by t hours (Tobias, 2012).

$$F(t) = P(T \leq t) = \int_{-\infty}^{t} f(s)ds \quad (1)$$

A Reliability function $R(t)$ estimates the probability that a system or component will survive by t hours. Hence, is the complement of $F(t)$. I.e. the fraction of all units in the population that will survive by t hours, see formula 2 (Tobias, 2012).

$$R(t) = 1 - F(t) = P(T > t) = \int_{t}^{\infty} F(s)ds \quad (2)$$

Hazard function $h(t)$, is interpreted as the propensity to fail in the next short interval of time, given that the system or component has survived to time t (Hamada A., 2008). The general formula for $h(t)$ is shown in formula 3.

$$h(t) = \frac{f(t)}{R(t)} \quad (3)$$

CHF, the Cumulative Hazard Function is the integral of the failure rate or conditional PDF of the component time to failure, given the component has survived to time $t\ hours$ (Modarres, 2010) (Formula 4).

$$H(t) = \int_{0}^{t} h(s)ds \quad (4)$$

AFR, the Average Failure Rate specifies failure-rate behaviour of a system over an interval of time (Tobias, 2012). It is expressed by formula 5.

$$AFR(t_{1,}t_{2,}) = \frac{H(t_{2,}) - H(t_{1,})}{t_{2,} - t_{1,}} \quad (5)$$

$MTTF$, the Mean Time to Failure expresses the expected time to failure. Hence the time which the system is expected to perform its function successfully (Modarres, 2010), See formula 6. $MTTF$ is also called the expected lifetime of the system.

$$MTTF = E(t) = \int_{0}^{\infty} tf(t)dt \quad (6)$$

All previous definitions are conditioned by $f(t) = 0$ for $t < 0$. Therefore, all the failure times censored or not must be positive. A summary including the relationships between them in the reliability setting is presented in Table 1 (Hamada, 2008).

Table 1. Relationships between probability density function, cumulative function, reliability function, hazard function and cumulative hazard function assuming

	$f(t)$	$F(t)$	$R(t)$	$h(t)$	$H(t)$
$f(t)$	$f(t)$	$\frac{d}{dt}f(t)$	$-\frac{d}{dt}R(t)$	$h(t).e^{\left[-\int_0^t h(s)ds\right]}$	$\left[\frac{d}{dt}H(t)\right].e^{[-H(t)]}$
$F(t)$	$\int_0^t f(s)ds$	$F(t)$	$1 - R(t)$	$1 - e^{\left[-\int_0^t h(s)ds\right]}$	$1 - e^{[-H(t)]}$
$R(t)$	$\int_t^\infty f(s)ds$	$1 - F(t)$	$R(t)$	$e^{\left[-\int_0^t h(s)ds\right]}$	$e^{[-H(t)]}$
$h(t)$	$\frac{f(t)}{\int_t^\infty f(s)ds}$	$\frac{\frac{d}{dt}F(t)}{[1 - F(t)]}$	$-\frac{d}{dt}log[R(t)]$	$h(t)$	$\frac{d}{dt}H(t)$
$H(t)$	$-log\left[1 - \int_0^t f(s)ds\right]$	$-log[1 - F(t)]$	$log[R(t)]$	$\int_0^t h(s)ds$	$H(t)$

From all the $PDF's$ available only four are going to be considered for this report. The selected $PDF's$ are Exponential, Weibull, Lognormal and Gamma. The main reasons for this choice are:

1. How well-fitted they are to model a highly reliable system regardless the type of variation through time it may have.
2. The flexibility in shape and scale parameters for any given historical data, if available, and prior expert knowledge.
3. The sensitivity when selecting initial values of parameters for predictive prior distributions.
4. The availability of built-in functions for these four statistical distributions in Matlab R2018 such as random generated samples, MLE estimators, Goodness of fitness with Kolmogorov-Smirnov, Anderson-Darling and Chi-square methods, acceptance-rejection sampling and iterative calculations with $PDF$and CDF.

An example of a random data set with 100 values within a time range from 0 to 5 years will be presented in chapter 2. This is used to test how the estimation of expected lifetime is done for each one of the four statistical distributions. The goal is to run the codes for MLE and BA of all four $PDF's$, and compare the obtained results.

2.1.1 Exponential Distribution

The exponential distribution is perhaps the most commonly used PDF in reliability assessments (Tobias, 2012). This can be attributed primarily to its simplicity and to the fact that it provides a constant failure rate, which is often the case for real life systems (Modarres, 2010). The notation for an exponential function is $E(\lambda, t)$, where λ represents the constant failure rate over time. The exponential PDF is shown in formula 7 (Tobias, 2012).

$$f(t) = \lambda e^{-\lambda . t} \qquad (7)$$

The exponential CDF is shown in formula 8 (Tobias, 2012).

$$f(t) = 1 - e^{-\lambda . t} \qquad (8)$$

For the exponential hazard function $h(t)$ see formula 9 (Tobias, 2012).

$$h(t) = \lambda \qquad (9)$$

From formula 9 it is obvious that the hazard function is equivalent to a constant failure rate over time. Hence, when modelling highly reliable systems with an exponential distribution, one considers them equally likely to fail as survival time increases and in the case of failure the system will be immediately replaced for a new one with the same properties (Modarres, 2010).

The exponential $MTTF$ is shown in formula 10 (Tobias, 2012).

$$MTTF = E(t) = \frac{1}{\lambda} \qquad (10)$$

The lack of memory property of the exponential function can be interpreted as a system that follows such function does not remember how long it has been operating. Therefore, the probability of failure for example in the next hour is the same for the system when it is considered new, recent or when it has been operating for a long time (Tobias, 2012)

In Matlab, there are two ways to define and generate an exponential PDF. It can be done either by using the prebuilt function **`y=exppdf(data,lambda)`**, (Magrab, 2011)or by writing the formula directly **`y=lambda.*exp(-lambda.*t)`** (Chapman, 2004). The latter also applies for CDF and $MTTF$.

A typical Exponential histogram with fixed parameter of failure rate λ is displayed in figure 1 for illustrative purposes.

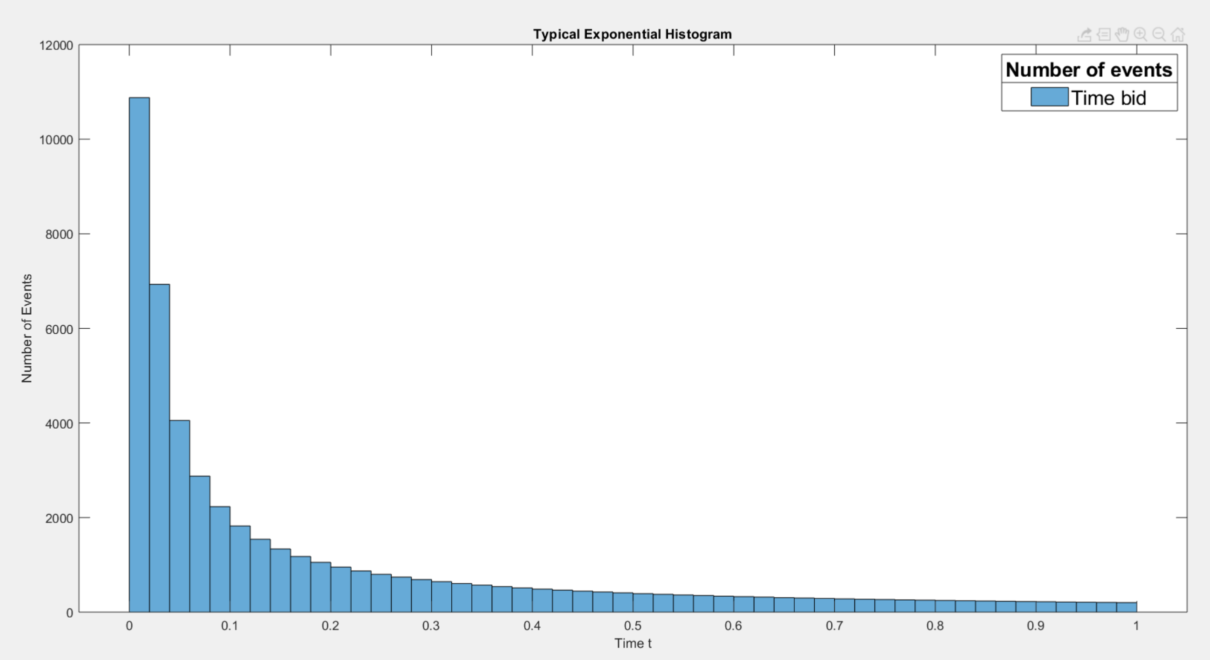


Figure 1. Typical histogram of an Exponential distribution.

In addition, a theoretical example of a random data set of 100 values between 0 to 5-time units, is presented in chapter 3. The idea is to test the MLE and BA codes for an Exponential PDF. Numerical and graphical representation of the results for comparison with other statistical functions in the scope of this report are also shown on chapter 3.

Typical Exponential plots with different parameters values of λ are displayed in Figure 2 for illustrative purposes. λ represents the failure rate for an exponential distribution.

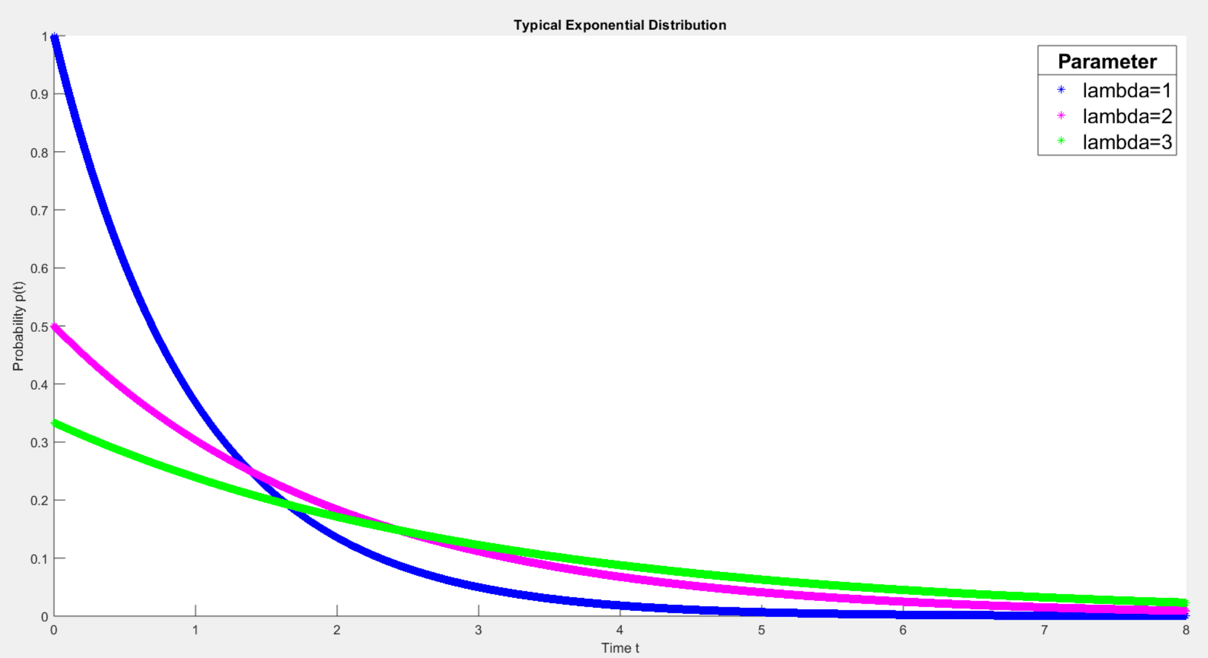


Figure 2. Typical plots of PDF for Exponential distributions

2.1.2 Weibull Distribution

Systems that follow this type of statistical functions have one common characteristic; a flexible failure rate over time. Hence, three types of failures rate are implicit in Weibull distributions (Tobias, 2012). These are increasing failure rate IFR, decreasing failure rate DFR, and the already known constant failure rate. Moreover, Weibull distribution is an extreme value distribution. The latter suggests its applicability when failure is due to the weakest link of many where failure can occur (Tobias, 2012).

The Weibull PDF is shown in formula 11 (Modarres, 2010).

$$f(t) = \frac{\beta . t^{\beta-1}}{\alpha^{\beta}} e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \qquad (11)$$

The Weibull CDF is shown in formula 12 (Modarres, 2010).

$$F(t) = 1 - e^{-\left(\frac{t}{\alpha}\right)^{\beta}} \qquad (12)$$

For the Weibull hazard function $h(t)$ see formula 13 (Tobias, 2012).

$$h(t) = \frac{\beta}{t}\left(\frac{t}{\alpha}\right)^{\beta} \qquad (13)$$

The Weibull $MTTF$ is shown in formula 14 (Modarres, 2010)

$$MTTF = E(t) = \alpha\Gamma\left(\frac{1+\beta}{\beta}\right) \qquad (14)$$

The parameter β sets the shape of the curve in the Weibull distributions, it has three main intervals to do so. For $0 < \beta < 1$ the PDF tends to infinite, therefore is decreasing as time increases. For $\beta > 1$ the PDF has an IFR up to a maximum value of $\alpha\left[1-\frac{1}{\beta}\right]^{\frac{1}{\beta}}$ (Tobias, 2012). Finally, for $\beta = 1$ the Weibull distribution is equivalent to an exponential distribution with a constant failure rate of $\lambda = \frac{1}{\alpha}$ (Tobias, 2012).

In Matlab there are two ways to define and generate a Weibull PDF. It can be done either by the using the prebuilt function **`y=wblpdf(data,lambda)`** (Magrab, 2011). Or by writing the Weibull formula directly, **`y=((beta.*t.^(beta-1))./(alpha.^beta)).*exp(-(t./alpha).^beta)`** (Chapman, 2004). This also applies for CDF and $MTTF$.

A typical Weibull histogram with fixed scale and shape parameters is displayed in figure 3 for illustrative purposes.

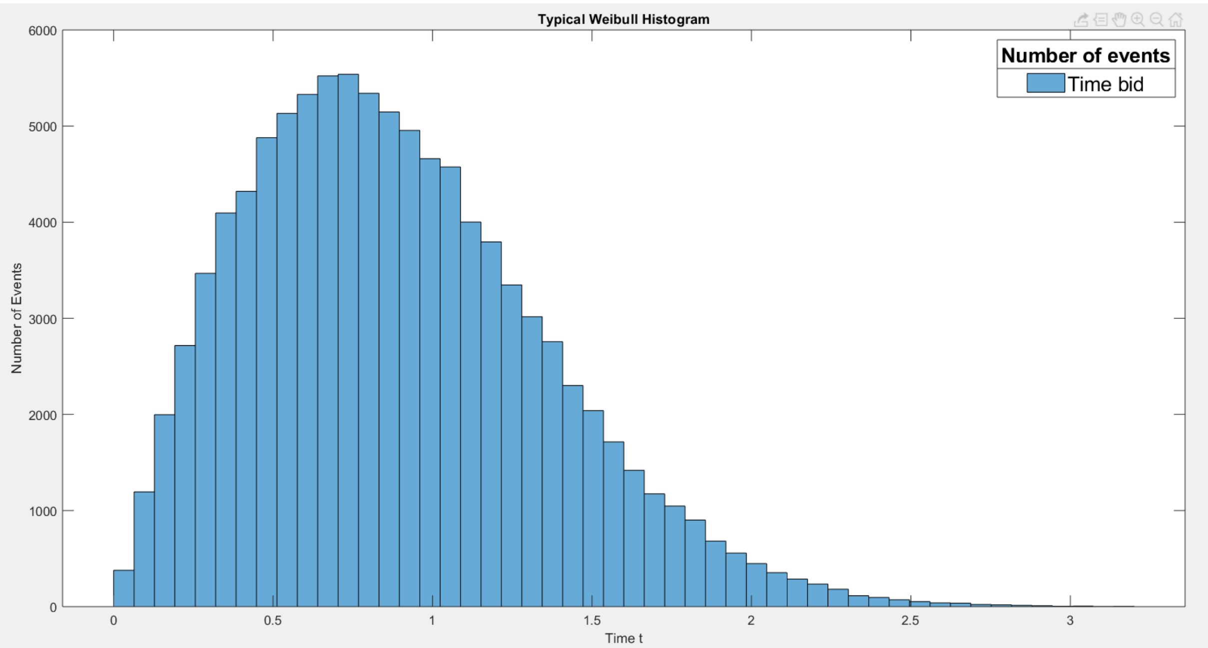


Figure 3. Typical histogram of a Weibull distribution

In addition, a theoretical example of a random data set of 100 values between 0-5 years is presented in chapter 3. The idea is to test the MLE and BA codes for a Weibull PDF. Numerical and graphical representation of the results for comparison with others statistical functions in the scope of this report are also shown on chapter 3.

Typical Weibull plots with different parameters values of α and β are displayed in Figure 2 for illustrative purposes, α and β represent the scale and shape parameters of a Weibull distribution.

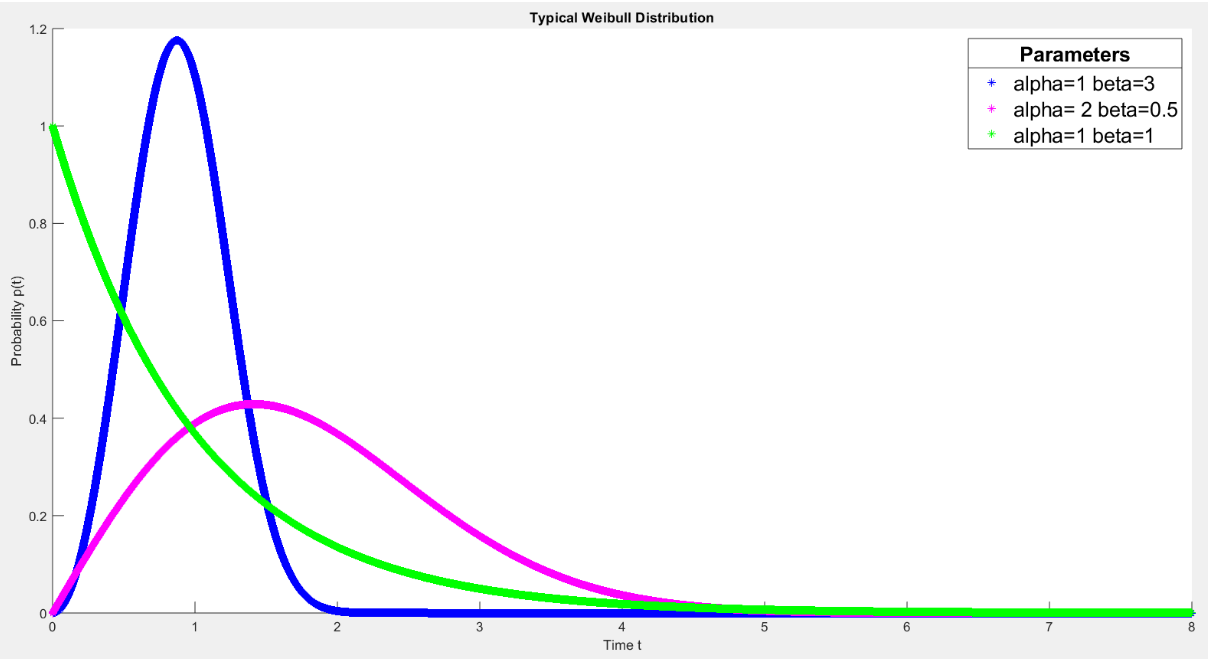


Figure 4. Typical plots of PDF for Weibull distributions

2.1.3 Lognormal Distribution

A simple way to understand the lognormal distribution is to compare with the normal. Where a random variable X has a normal distribution $N(\sigma, \mu, t)$, another random variable $T = e^X$ has a lognormal distribution $LogN(\sigma, \mu, T_{50}, t)$ (Hamada, 2008). The flexibility of the lognormal distribution for skewed data as time increases makes it useful. Moreover, its PDF shape resembles the Weibull distribution. Hence, a technique to determine which of these distributions fits better the data is based on the histogram of the logarithm of the data (Tobias, 2012).

The lognormal PDF is shown in formula 15 (Modarres, 2010).

$$f(t) = \frac{1}{\sigma_t . t . \sqrt{2\pi}} e^{\left[-\frac{1}{2.\sigma_t{}^2}.(ln\, t - \mu_t)^2\right]} \qquad (15)$$

The lognormal CDF is presented in formula 16 (Modarres, 2010).

$$F(t) = \frac{1}{\sigma_t t\sqrt{2\pi}} \int_0^t \frac{1}{\theta} e^{\left[-\frac{1}{2\sigma_t{}^2}(ln\, \theta - \mu_t)^2\right]} d\theta \qquad (16)$$

For the lognormal hazard function $h(t)$ see formula 17 (Modarres, 2010).

$$h(t) = \frac{\frac{1}{\sigma_t t\sqrt{2\pi}} e^{\left[-\frac{1}{2.\sigma_t{}^2}(lnln\, t - \mu_t)^2\right]}}{1 - \frac{1}{\sigma_t t\sqrt{2\pi}} \int_0^t \frac{1}{\theta} e^{\left[-\frac{1}{2\sigma_t{}^2}(lnln\, \theta - \mu_t)^2\right]} d\theta} \qquad (17)$$

The lognormal $MTTF$ is shown in formula 18 (Tobias, 2012).

$$MTTF = T_{50} e^{\frac{\sigma^2}{2}} \qquad (18)$$

The failure rate for this type of distribution increases over time and then decreases. It depends on the parameters μ_t and σ_t, the mean and standard deviation respectively (Modarres, 2010). On the other hand, the parameter T_{50} represents the median time of failure for a population of lognormal lifetimes (Tobias, 2012).

The lognormal PDF, is especially appropriate to model the time to failure of systems which early failure dominate its overall failure behaviour (Modarres, 2010). Moreover, it is a common model to represent Prior distributions.

In Matlab there are two ways to define and generate a Lognormal PDF. It can be done by using the prebuilt function `y=lognpdf(data,lambda)` (Magrab, 2011). Or by writing the Lognormal formula `y=((1./(sigma.*t.*sqrt(2*pi))).*exp((-1/(2.*sigma.^2)).*log((t-mu).^2)))` (Chapman, 2004). The latter also applies for CDF and $MTTF$.

In Figure 5, a typical histogram of a Lognormal distribution $LogN(\sigma, \mu, T_{50}, t)$ with fixed parameters σ and μ over time t is displayed for illustrative purposes.

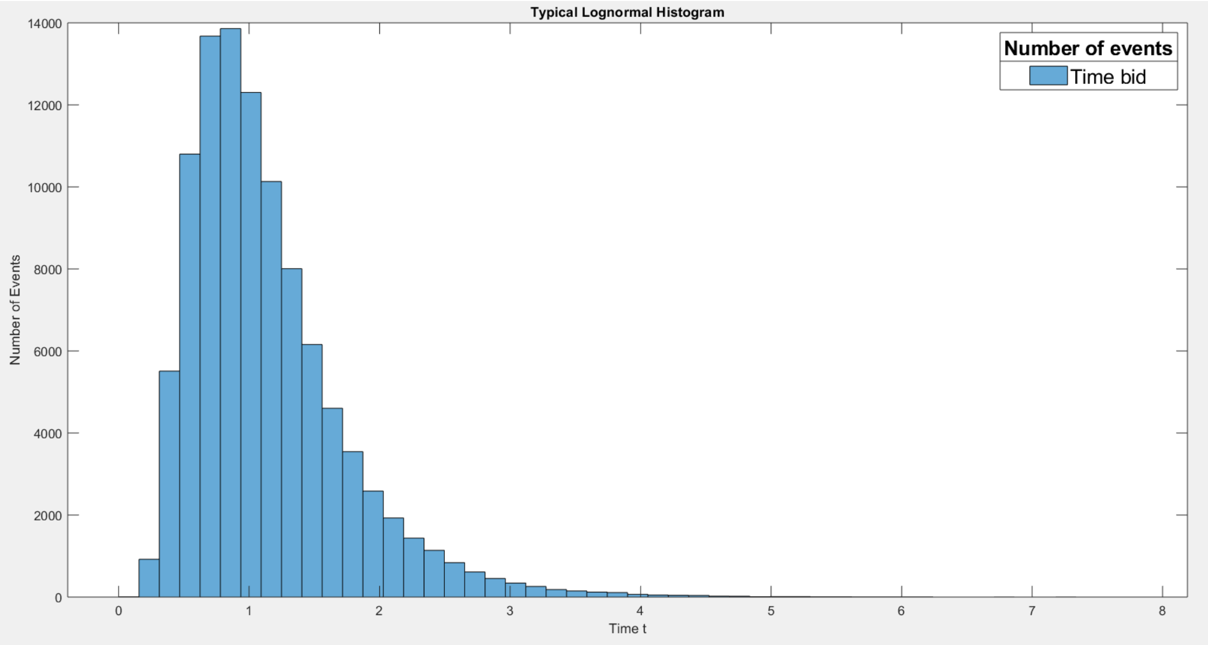


Figure 5. Typical Histogram of a Lognormal distribution

In addition, a theoretical example of a random data set of 100 values between 0-5 years is presented in chapter 3. The idea is to test the MLE and BA codes for a Lognormal PDF. Numerical and graphical representation of the results for comparison with others statistical functions in the scope of this report are also shown on chapter 3.

Typical Lognormal plots with different parameters values of σ and μ are displayed in figure 6 for illustrative purposes. σ and μ represent the mean and standard deviation parameters of a Lognormal distribution.

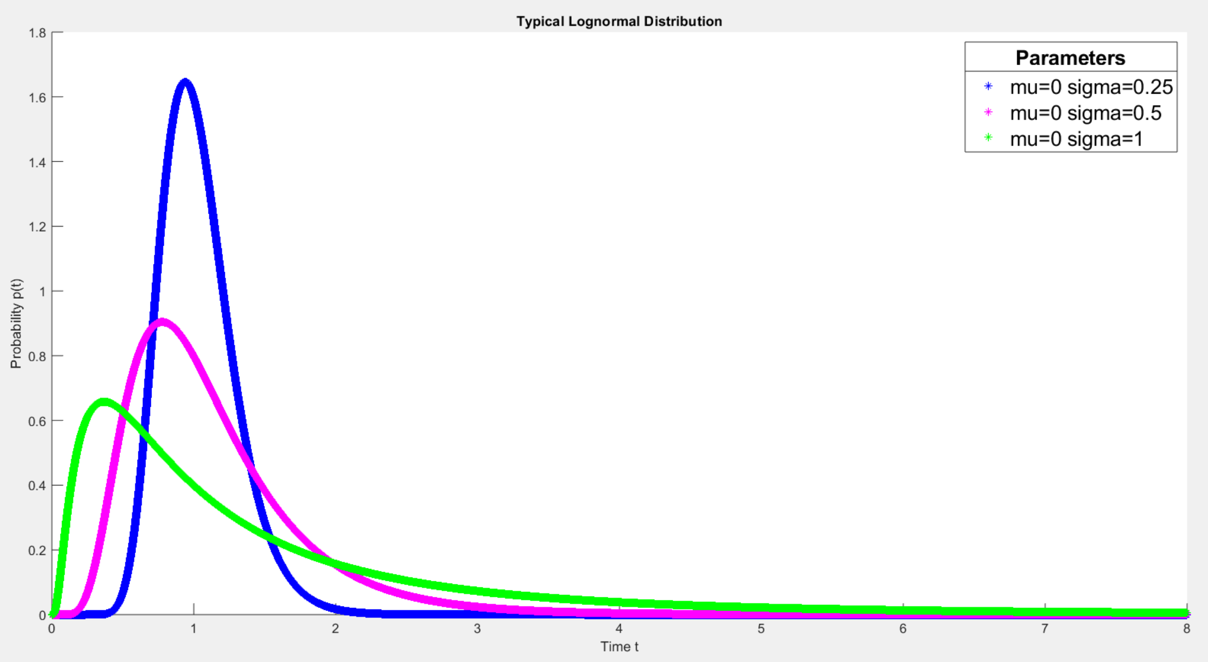


Figure 6. Typical plots of PDF for a Lognormal distributions

2.1.4 Gamma Distribution

This type of distribution is considerably flexible with regards to its parameters β and α. It contains the Gamma, Weibull and lognormal distributions as special cases (Ye, 2017). Moreover, from the generalized form when $\beta = 1$ and $\mu = 0$, it behaves as a Weibull distribution. The shape of the curve is given by β and the failure rate by α (Hamada, 2008). The main application in Bayesian Analysis when defining the prior distribution is related to systems with standby components. Also, for failure times between maintenance, recalibration and in (Modarres, 2010). In other words, the gamma distribution is often used as a distribution for waiting times and service times. Nevertheless, there are no closed-form expressions for the PDF estimators (Ye, 2017). To obtain the estimators for Gamma distributions statisticians rely in the moment estimators and techniques such as Newton Rapson (Ye, 2017). An additional feature of a gamma PDF is its skewness to the left side of the curve. Hence to lower time intervals of survival (Hamada, 2008).

The gamma PDF is shown in formula 19 for $\alpha, \beta, t > 0$ (Hamada, 2008).

$$f(t) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} t^{\alpha-1} e^{\left(-\frac{t}{\beta}\right)} \qquad (19)$$

The gamma CDF is shown in formula 20 (Modarres, 2010).

$$F(t) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^t y^{\alpha-1} e^{\left(-\frac{y}{\beta}\right)} dy \qquad (20)$$

For the gamma hazard function $h(t)$ see formula 21 (Modarres, 2010).

$$h(t) = \frac{t^{\alpha-1} e^{\left(-\frac{t}{\beta}\right)}}{\beta^{\alpha}\left[\Gamma(\alpha) - \int_0^t y^{\alpha-1}\, e^{\left(-\frac{y}{\beta}\right)} dy\right]} \qquad (21)$$

The gamma $MTTF$ is shown in formula 22 (Ye, 2017).

$$MTTF = \beta\alpha \qquad (22)$$

In Matlab there are two ways to define and generate a Gamma PDF. It can be done by the using the prebuilt function **`y=gampdf(data,lambda)`** (Magrab, 2011)**.** Or just by writing the Gamma formula, **`y=(1./(beta.^alpha).*gamma(alpha)).*(t).^(alpha-1).*exp(-(t./beta)))`** (Chapman, 2004). Both apply for CDF and $MTTF$.

The Figure 7 for a typical lognormal distribution $Gamma(\beta, \alpha, t)$ with fixed parameters α and β over time t is displayed for illustrative purposes.

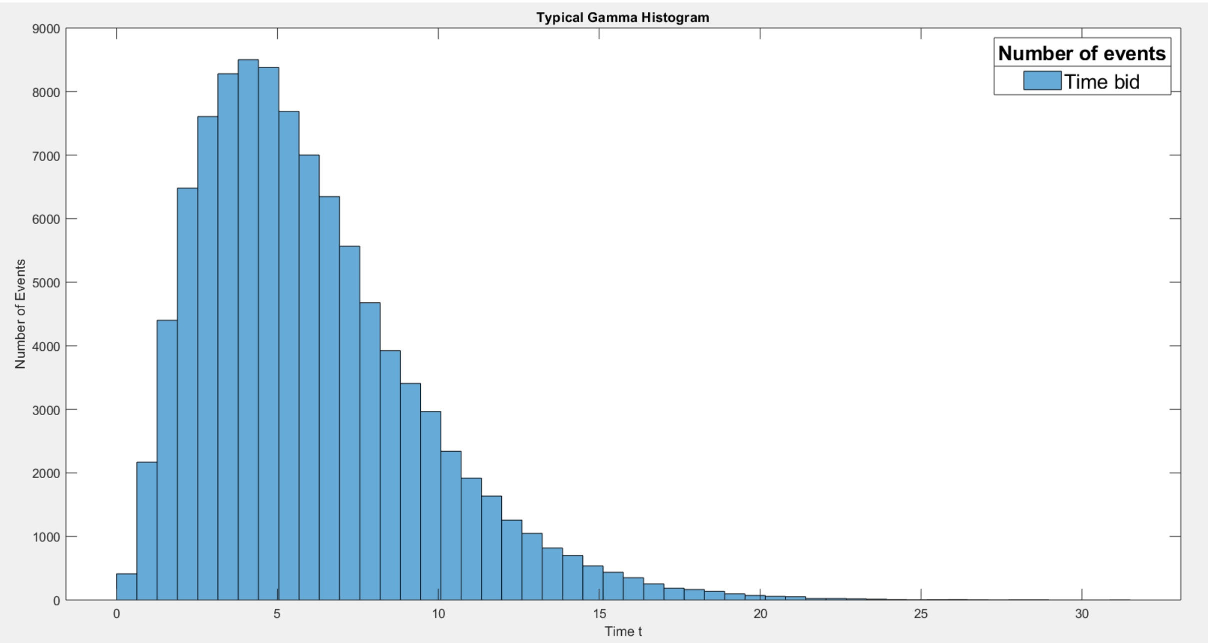


Figure 7. Typical histogram of a Gamma distribution

In addition, a theoretical example of a random data set of 100 values between 0-5 years is presented in chapter 3. The idea is to test the MLE and BA codes for a Gamma PDF. Numerical and graphical representation of the results for comparison with others statistical functions in the scope of this report are also shown on chapter 3.

Typical Gamma plots with different parameters values of are displayed in Figure 6 for illustrative purposes, α and β represent the scale and shape parameters of a Weibull distribution.

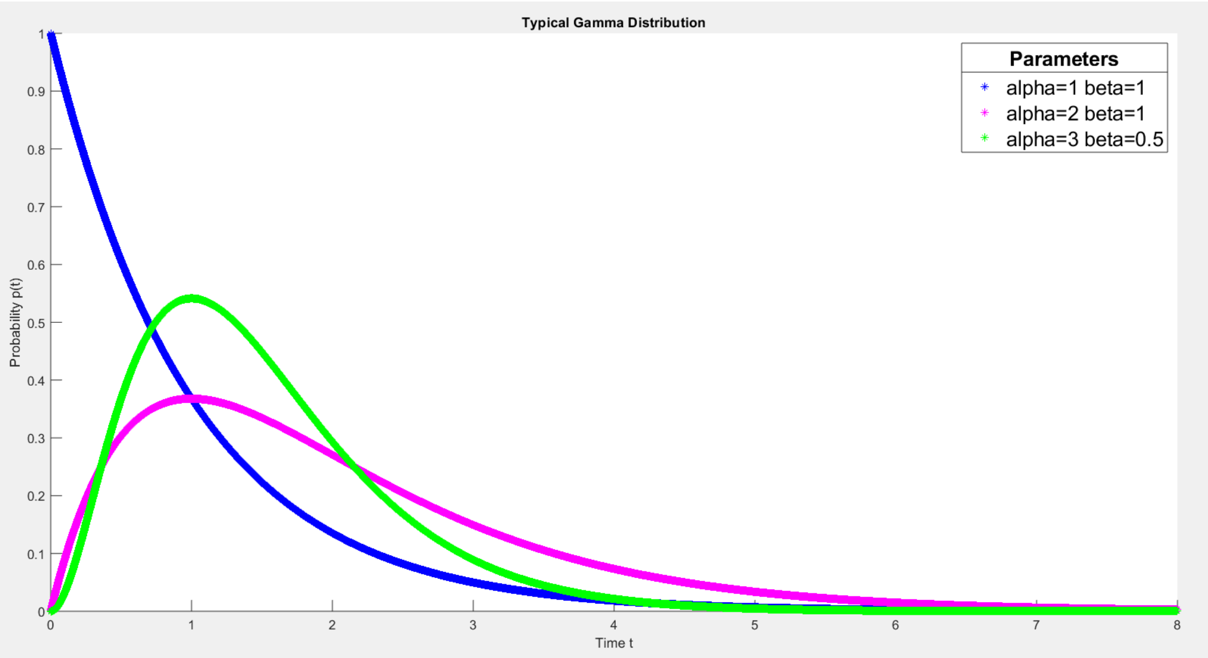


Figure 8. Typical plots of PDF for Gamma distributions

2.2 Estimation Models

The relevance and amount of historical data determines how accurate the MLE can be. Whereas expert knowledge and the techniques for elicit it to set a prior distribution determines the accuracy of Bayesian analysis. Moreover, the selection of prior distributions based on expert knowledge has an important effect on the posterior distributions (Mickelsson, 2015).

Most estimation models are made by using MC simulations. It is defined as "a process that generates random number inputs for uncertain values, which are then processed by a mathematical model, so that many scenarios can be evaluated" (Skinner, 2009).

There are several approaches to determine the parameters and probability distributions of a phenomena, and they are used according to the elements available (Modarres, 2010). In this study two main approaches are going to be tested and compared. These approaches to determine the parameter ϕ, PDF and CDF, are the classical approach based on Maximum Likelihood Estimation and the Bayesian Analysis approach. These are based on a combination of prior knowledge and likelihood data. This will be discussed further in this chapter.

Furthermore, to select the proper DF, it is necessary to verify the goodness of fit for the function with regards of the data if available (Hamada, 2008). There are several feasible techniques to test the goodness of fit for the selected statistical functions in the scope of this report. Among them are the following:

- Kolmogorov-Smirnov test
- Anderson-Darling test
- Chi-Square test

For this report the goodness of fit is done with Kolmogorov-Smirnov test for three main reasons. First, it is one of the easiest tests as it gives a probability value p which indicates how likely it is for the likelihood data to fit a specific PDF or CDF if needed. Second, it is well supported in Matlab as a built-in function (Casella, 2002). Third, after many trials of fitting random data sets to the four PDFs of this report, the outcomes for the indicators of Chi-square and Anderson Darling tests were not as precise as the Kolmogorov ones. Moreover, in some of the trials NAN outcomes for these indicators were obtained. A brief explanation for Kolmogorov-Smirnov test follows in the next section.

2.2.1 Kolmogorov-Smirnov Test

To determine the goodness of fit of a DF on a sample data set it is necessary to measure how well its predicted values match the observed data (James, 2014). In other words, it is necessary to quantify the extent to which the predicted response value of a given observation is close to the outcome value of the selected DF. Kolmogorov-Smirnov proposed a measure of deviation between these two values (James, 2014).

Given a sample DF defined by formula 23.

$$S_n(x) = \left\{ \begin{matrix} 0, & x < x_{(1)} \\ \frac{r}{n}, & x_{(r)} \le x < x_{(r+1)} \\ 1, & x_{(n)} \le x \end{matrix} \right\} \qquad (23)$$

Where x, represents time values of the x axis, $x_{(1)}$ to $x_{(n)}$ are the n observation of the sample data and r the step from one sample value to the next. Then, $S_n(x)$ is the proportion of observations from the sample data that do not exceed the DF value of x (Kendall, 1987).

For the Kolmogorov-Smirnov test of goodness of fit see the formula 24.

$$\lim_{n\to\infty} P\{S_n(x) = F_0(x)\} = 1 \qquad (24)$$

As most methods for testing the goodness of fit, Kolmogorov-Smirnov provides better results when large amount of iterative random sampling is performed. For such cases the probability value of the sample DF will be closer to 1 (Kendall, 1987).

Matlab calculates the Kolmogorov-Smirnov goodness of fit using two prebuilt functions. The first one is **`h=kstest(x)`**, which returns a test decision for the null hypothesis when the sample $x_{(1)}$ to $x_{(n)}$ comes from a specified distribution. Using the one-sample Kolmogorov-Smirnov test, the result of h is 1 if the test rejects the null hypothesis at the 5% significance level, h is 0 if otherwise.

The second one is **`[h1,p]=kstest(values,'CDF',cd1,'alpha',0.01)`**, (Chapman, 2004), which also returns a probability value $P\{S_n(x)\}$ of the hypothesis test, using any of the input arguments such as the type CDF and the degree of confidence in this value. Notice that by default the degree of confidence is 1%.

An example of the Kolmogorov-Smirnov test for goodness of fit with a typical Gamma CDF and a random Gamma sample data set of a thousand values is presented in Figure 9.

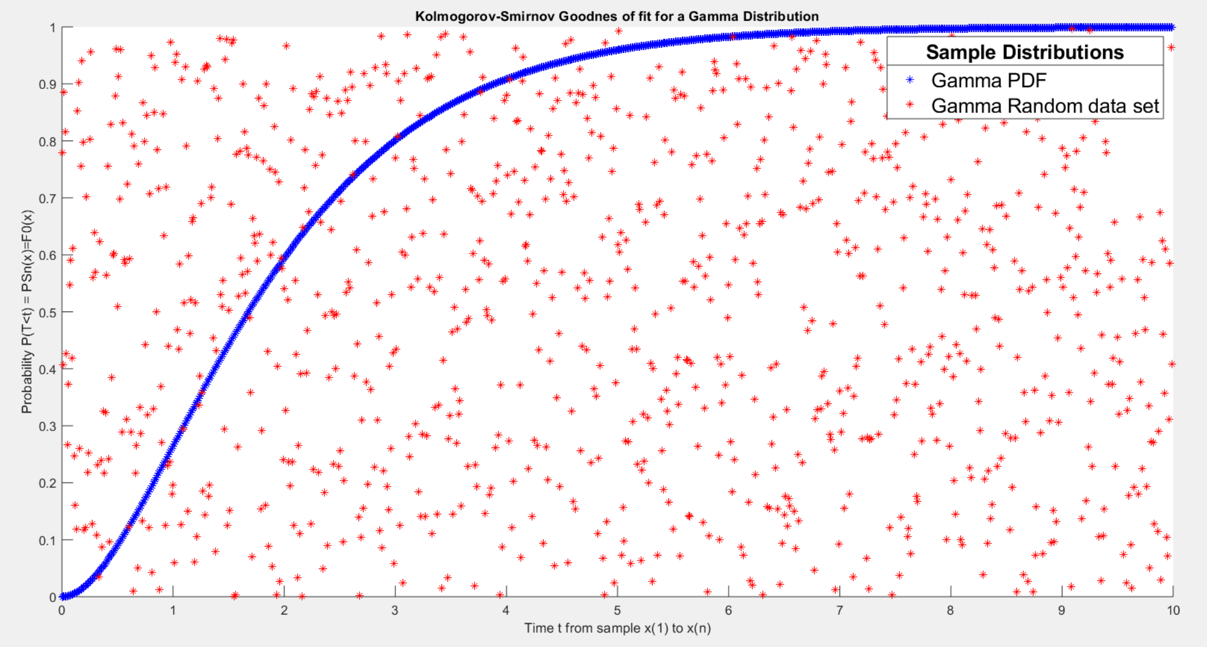


Figure 9. Kolmogorov-Smirnov test for Goodness of Fit for Gamma CDF

For this report the probability values of goodness of fit vary from relatively low values ($1x10^{-02}$) for the random data set presented in subchapter 3.1.1, to high values according to other data sets presented in the study cases, which behave more like a predefined CDF.

2.2.2 Maximum Likelihood Estimation

The classical approach of Maximum Likelihood Estimation provides a statistical framework for assessing wisely the information available in historical data (Eliason, 1993). The goal is to produce a point estimate ϕ_i of some population parameter ϕ such that it maximizes the likelihood of observing the historical data of the phenomena to occur (Eliason, 1993). This point estimate (estimator) is not expected to estimate $\hat{\phi}$ without error. However, it is expected to be close enough (Walpole, 2016). The procedure to determine the maximum likelihood estimator of a random variable given historical data is based on the following seven steps:

1. Define the set of independent and identically distributed variables $Y = [Y_1, Y_2, \dots Y_n]$ previously observed.
2. Establish the joint F $f(y_1; \phi)$, supposed to represent the phenomena.
3. Set the compound product of marginal distributions see formula 25. This is the likelihood function $L(\phi)$.

$$L(\phi) = \prod_{i=1}^{n} f(y_1, \phi) \qquad (25)$$

4. Apply the logarithm to the likelihood function to reduce computational difficulties and obtain the log-likelihood function $l(\phi)$ see formula 26.

$$l(\phi) = ln[L(\phi)] = ln\left[\prod_{i=1}^{n} f(y_1, \phi)\right] = \sum_{i}^{n} f(y_1, \phi) \qquad (26)$$

5. Maximize the value of the log-likelihood function by taking the 1st derivative $\frac{d\{ln[L(\phi)]\}}{\phi}$ and setting it equal to zero see formula 27.

$$\frac{d\{ln[L(\phi)]\}}{\phi} = 0 \qquad (27)$$

6. Solve the equation to find the maximum likelihood estimator $\hat{\phi}$.
7. Take the 2nd derivative of the log-likelihood function and verify that it is a maximum estimator by setting it lower than zero see formula 28.

$$\frac{d^2\{l(\phi)\}}{\phi^2} < 0 \qquad (28)$$

With the maximum likelihood estimator, any point estimate for the set of independent distributed variables can be determined with a relatively high degree of accuracy. However, the latter will depend on the amount and quality of historical data available as the desirable properties of the MLE are justified only in situations with large sample (Eliason, 1993). Otherwise, an assessment can be skewed.

The main result of MLE is a function of the unknown parameter, called the likelihood function (Pawitan, 2001). The parameter itself is called maximum likelihood estimator $\hat{\phi}$. (Eliason S, 1993). The information provided from the likelihood function tells where $\hat{\phi}$ is likely to fall with an inherent degree of uncertainty, conveyed (Pawitan, 2001). A formal definition for the likelihood is provided by (Pawitan, 2001), "*Assuming a statistical model parameterized by a fixed and unknown* ϕ, *the likelihood* $L(\phi)$ *is the probability of the observed data* x *considered as a function of* ϕ*".*

From the statistical perspective, the likelihood function is a tool for an objective analysis with available data. Moreover, when such data incorporates uncertainty due to the limited or restricted amount of information that it provides (Pawitan, 2001). In addition, $\hat{\phi}$ represents a way to simplify an analysis when real data analysis is done. One can prefer to handle many $\hat{\phi}'s$ instead of their related likelihood functions (Kendall, 1987).

The procedure to obtain $\widehat{\phi}$ of all four PDFs discussed previously is the same. Therefore, only the exponential $\hat{\phi}$ is going to be shown for illustrative purposes. A summary table with the results of all the PDFs used in this report is presented in Table 2.

A theoretical example on how to determine the MLE $\widehat{\phi}$ according to the procedure previously discussed is presented as follows:

1. The theoretical set of independent and identically distributed variables $Y = [Y_1, Y_2, \dots Y_n]$
2. Establish the joint $f(t,\lambda)$, supposed to represent the phenomena (Formula 29).

$$f(t,\lambda) = \lambda e^{-\lambda . t} \qquad (29)$$

3. The likelihood function $L(\lambda)$ (Formula 30).

$$L(\lambda) = \prod_{i=1}^{n} \lambda e^{-\lambda . t} = \lambda^n e^{-\lambda \sum_1^n t_i} \qquad (30)$$

4. The log-likelihood function $l(\lambda)$ (Formula 31).

$$l(\lambda) = ln\left[\lambda^n e^{-\lambda \sum_1^n t_i}\right] = nln(\lambda) - \lambda \sum_1^n t_i \qquad (31)$$

5. The 1st derivative $\frac{d\{ln[L(\phi)]\}}{\phi}$ and set it equal to zero. See formula 32

$$\frac{d\{l(\lambda)\}}{d\lambda} = \frac{n}{\lambda} - \sum_1^n t_i = 0 \qquad (32)$$

6. The maximum likelihood estimator $\hat{\lambda}$ (Formula 33).

$$\hat{\lambda} = \frac{\sum_1^n t_i}{n} \qquad (33)$$

7. Verification that it is a maximum estimator by setting it lower than zero. (Formula 34).

$$\frac{d^2\left\{\frac{n}{\lambda} - \sum_1^n t_i\right\}}{d\lambda} < 0 \qquad (34)$$

Table 2 presents a summary of MLE estimators for each PDF (Hamada, 2008).

The MLE codes in Matlab for each PDF follow the same procedure described above. Furthermore, the general MLE code applies for any type of data set as it evaluates the goodness of fit for each type of PDF, ranks them and selects the best fitted one to perform MLE.

Table 2. Summary of MLE estimators for different PDF's

Type of distribution	PDF	$MLE\ \hat{\Phi}$
Exponential	$f(t) = \lambda e^{-\lambda . t}$	$\hat{\lambda} = \frac{\sum_1^n t_i}{n}$
Weibull	$f(t) = \frac{\beta t^{\beta-1}}{\alpha^\beta} e^{-\left(\frac{t}{\alpha}\right)^\beta}$	$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} t_i^{\hat{\beta}^{\frac{1}{\hat{\beta}}}}$ $\hat{\beta} = \left[\frac{\sum_{i=1}^n t_i^{\hat{\beta}} ln(t_i)}{\sum_{i=1}^n t_i^{\hat{\beta}}} - \underline{ln(t)}\right]^{-1}$
Lognormal	$f(t) = \frac{1}{\sigma_t t\sqrt{2\pi}} e^{\left[-\frac{1}{2\sigma_t^2}(ln ln\, t - \mu_t)^2\right]}$	$\hat{\mu} = \frac{\sum_{i=1}^n ln(t_i)}{n}$ $\hat{\sigma}^2 = \frac{\sum_{i=1}^n [ln(t_i) - \hat{\mu}]^2}{n}$
Gamma	$f(t) = \frac{1}{\beta^\alpha \Gamma(\alpha)} t^{\alpha-1} e^{\left(-\frac{t}{\beta}\right)}$	$\hat{\beta} = \frac{1}{n\hat{\alpha}}\sum_{i=1}^{n} t_i$

Matlab performs MLE with built-in functions presented next:

Steps 1 and 2:

```
Lambda=rand()*3;
Y=exppdf(data,lambda)
```

Steps 3 to 6:

```
phat=mle(data,'pdf',@(data,lambda)exppdf(data,lambda),'start',[lambda]);
```

Where $phat$ is the estimator according to the selected PDF, in this case is Exponential. Lambda is the initial parameter drawn from a uniform distribution for the MC iterative calculations and $data$ constitutes the sample data from which phat is going to be estimated.

Step 7: is not necessary as the maximum likelihood estimator already meet the condition.

Given a real set of sample data, the maximum likelihood estimators for each PDF is obtained according to Table 3 (Magrab, 2011). In addition, point estimates can be easily found using a combination of available data and estimator(s) of the selected PDF.

Table 3. Summary of MLE estimators in MatlabR2018 of different PDF's

Type of distribution	$\boldsymbol{MLE}$ **estimators** $\hat{\phi}$ $(\boldsymbol{phat})$
Exponential	`phat = mle(data,'distribution','exponential');`
Weibull	`phat = mle(data,'distribution','weibull');`
Lognormal	`phat = mle(data,'distribution','lognormal');`
Gamma	`phat = mle(data,'distribution','gamma');`

In addition to maximum likelihood estimators, the classical approach requires an interval of confidence for such parameters, since the probability (Formula 35) that the random variable will be in this interval is unknown (Hogg, 2006). This interval is called confidence interval and it is denoted as $100(1-\alpha)\%$ (Casella, 2002). Where α accounts, to some extent, for the uncertainty related to the estimated parameter. Therefore, the larger α is, the smaller the confidence interval will be, and vice versa (Hogg, 2006). Hence, given two random variables A and B such that $A \leq B$ with probability equal to one, the confidence interval of an estimator $\hat{\phi}$ will be $[A, B]$ with a $(1-\alpha)$ probability (Hasting, 1997).

$$P\big[\phi \in [A, B]\big] = 1 - \alpha \qquad (35)$$

The method used to find a confidence interval for the different PDF is based on finding its boundaries. It consists on setting a reference and an estimated margin of error for the PDF. Then assume upper and lower bounds for the estimator $\hat{\phi}$ according to the PDF. In many cases these bounds are symmetrical around $\hat{\phi}$ (Hasting, 1997). Moreover, the true value will be within the confidence interval obtained from a function that includes both estimator and parameter (Formula 36).

$$-\hat{\phi}_{\alpha/2} \leq \phi \leq +\hat{\phi}_{\alpha/2} \qquad (36)$$

Where the boundaries are $= -\phi_{\alpha/2}$, $B = +\phi_{\alpha/2}$ and according to Formula 11. The probability of the estimator being within the confidence interval for symmetrical PDFs is described in formula 37 (Hasting, 1997).

$$P\left[-Z_{\alpha/2} \leq h(\phi, \hat{\phi}), \leq +Z_{\alpha/2}\right] = 1 - \alpha \qquad (37)$$

Where the bounds $-Z_{\alpha/2}$ and $+Z_{\alpha/2}$ are the quantiles normally distributed for the confidence interval, since $\hat{\phi}$ is asymptotically normally distributed. Hence, the function is presented in formula 38 (Walpole, 2016).

$$h(\phi, \hat{\phi}) = \frac{\hat{\phi} - \phi}{\sqrt{Var(\hat{\phi})}} \qquad (38)$$

For PDFs where the bounds are not symmetrical to the estimator different types of intervals can be set. The most common one is shown in formula 39 (Walpole, 2016).

$$P\left[W_{1-\alpha/2} \leq h(\phi, \hat{\phi}), \leq W_{\alpha/2}\right] = 1 - \alpha \qquad (39)$$

Where the bounds $W_{1-\alpha/2}$ and $W_{\alpha/2}$ are the quantiles of the confidence interval.

A summary of all the pre-built functions in Matlab to calculate confidence intervals and setting the degree of confidence in percentage given a scalar input is presented in Table 4 (Chapman, 2004).

Table 4. Summary of pre-built functions in Matlab to calculate confidence intervals

Type of distribution	**Confidence Interval** $P[\phi \in [A, B]] = 1 - \alpha$
Exponential	`[muhat,muci]=expfit(data,degree_value)`
Weibull	`[muhat,muci]=wblfit(data,degree_value)`
Lognormal	`[muhat,muci]=lognfit(data,degree_value)`
Gamma	`[muhat,muci]=wblfit(data,degree_value)`

Notice that the input in Matlab of the degree for the confidence interval must be a decimal value, because the code is set to use **`degree_value=degree/100`**.

The specific codes for Exponential, Weibull, Lognormal and Gamma PDF are shown in the appendix. Furthermore, the general steps that the software follows are:

1. Import the likelihood data set from an external file. The codes only read Microsoft Excel (.xlsx) format files
2. Generate random values for the initial parameter(s) from a uniform distribution this can be $\hat{\lambda}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\mu}$ or $\hat{\sigma}$, according to the type of distribution
3. Evaluate the goodness of fit for each PDF by the Kolmogorov Smirnov test and rank them to select the best fitted one, i.e. the highest p value among them
4. Estimate the parameter(s) of the selected PDF
5. Asks for the $1 - \alpha$ degree of confidence intervals for the parameter(s)
6. Calculates the confidence interval for each parameter
7. Asks for a desired time $tsurvival$ to evaluate the reliability of the system given the likelihood data set
8. Calculates the reliability $R(t)$
9. Calculates the unreliability $F(t)$
10. Calculates the $MTTF$
11. Present a summary of the outcomes in a table
12. Generates the plot of $R(t)$
13. Generates the histogram of $R(t)$

2.2.3 Bayesian Analysis

Bayesian analysis is based on subjective probability (Hogg, 2006). It can yield more information from small size samples where classical statistical approaches such as MLE cannot (Tobias, 2012). Furthermore, it relies on two types of information available to provide a reliable prediction for future outcomes (Ayyub, 2003). These types of information can be:

- Objective information based on experimental results, observations or historical data
- Subjective information based on prior experience, intuition, similar situation previously encountered and problem knowledge

To perform BA, four main steps need to be done following the classical approach referred to as “analytical” in this thesis (Gelman, 2014).

1. Assign a prior distribution
2. Determine a likelihood distribution
3. Obtain the posterior distribution
4. Find the Bayesian estimator $\hat{\phi}_{Bayes}$

The other approach to perform BA is by “sampling” large amounts of data points between both prior and likelihood distributions to obtain the posterior. This is the procedure that is going to be followed in this thesis.

The Bayesian methodology implies that the PDF of a parameter to model a system`s behaviour can be obtained from the prior and posterior distributions (Modarres, 2010). In addition, more assumptions must be made for BA than a classical approach, especially to obtain the prior distribution (Tobias, 2012).

The basis of BA is Bayes theorem, which defines conditional probabilities (Formula 40) (Aven T. , Risk Analysis, 2012).

$$P(C \mid A) = \frac{P(A \mid C).P(A)}{P(C)} \qquad (40)$$

Where $P(C \mid A)$ is the conditional probability that event C will occur, given the occurrence of event A. Hence, $P(A)$ characterizes the opposite and will represent the prior distribution. $P(C)$ characterizes the events that already occurred. Hence, the likelihood distribution (Aven T. , Risk Analysis, 2012).

BA heavily relies on the accuracy of the prior and likelihood probability distributions to produce a suitable posterior distribution. Moreover, the outcome of the posterior depends on elicited knowledge from experts to assign the prior and the availability of relevant data to produce the likelihood (Hayakawa, 2001). In general, the probability of a theory becoming true will increase as supporting data increases (Lindley, 2014).

Once a posterior distribution is obtained, several outcomes are of high interest for risk and reliability assessments. These are $\hat{\phi}_{Bayes}$, expected lifetime of the system and credibility interval $\hat{\phi}_{Bayes}$ for analytical approach. Nevertheless, from the sampling perspective the predictive prior, likelihood and predictive posterior distributions are the relevant ones. Based on this all the required indicators such as $R(t)$, $F(t)$ and MTTF can be obtained.

The BA general framework to estimate the expected lifetime is shown in Figure 9.

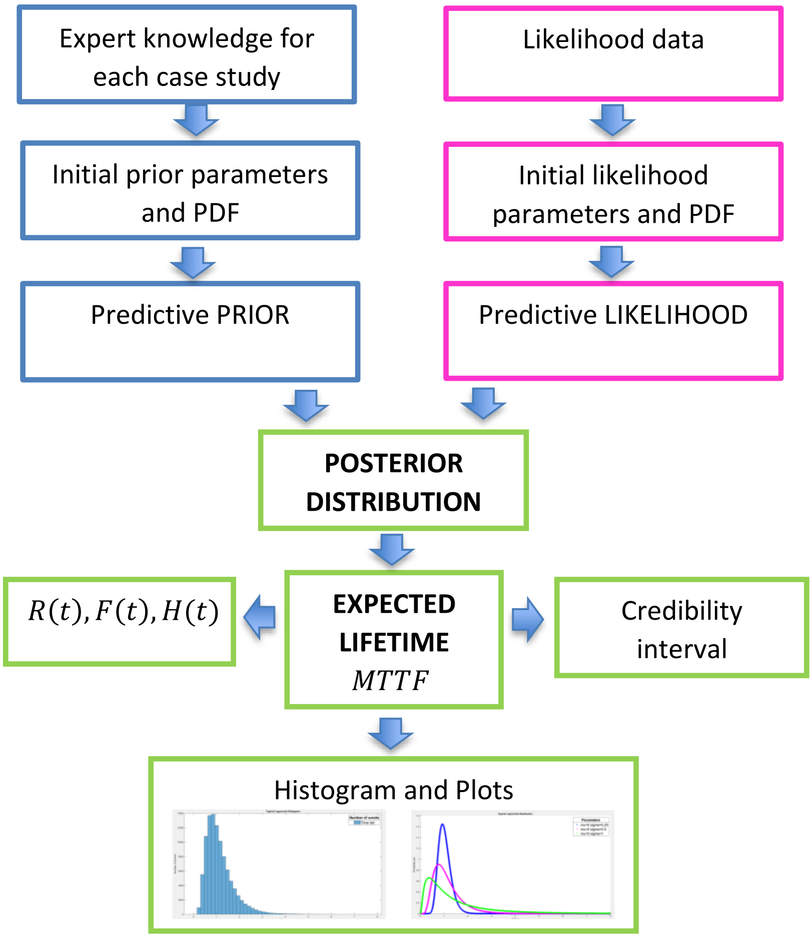


Figure 10. General framework to estimate expected lifetime using Bayesian Analysis

A way to interpret BA is that the prior distributions are assigned based on the previous knowledge with regards to the phenomena. Moreover, BA requires the selection of a PDF and its initial parameters based on the prior expert knowledge. Likelihood distributions are obtained from the available historical data. This does not require expert knowledge. Posterior distributions are established via the data sampling from Prior vs. Likelihood to model the expected survival and failure behaviours of the system (Tobias, 2012).

The weight that the likelihood and prior distributions will have on the posterior depends on how relevant and accurate they are. In other words, the posterior will be located closer to the likelihood or prior distribution according to the weight each one has (Mickelsson, 2015). See Figure 11 for a typical combination of a prior, likelihood and posterior distributions of a Normal PDF.

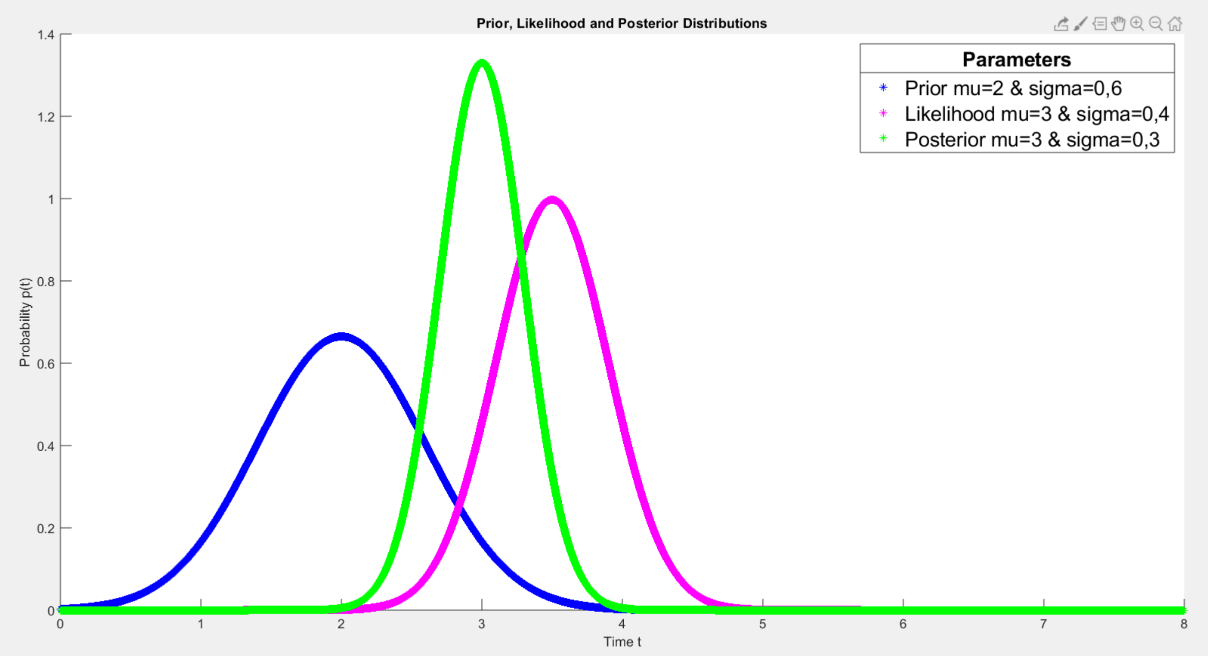


Figure 11. Combination of Prior, Likelihood and Posterior distributions of a Normal PDF

A brief discussion on the predictive prior distribution, elicitation principles to obtain it and other related issues will be presented in the next section.

Prior Distributions

Prior distributions are the foundation of Bayesian statistical inference (Gelman, 2014). They represent the degree of belief that a system will behave according to an assigned PDF based on the previous knowledge experts can have (Hayakawa, 2001). In other words, it expresses the degree of uncertainty related to the value of a parameter from which the PDF is based on. Suppose one needs to establish the relevance of the prior knowledge in BA, then Bayes theorem can be written in general according to formula 41 (Hayakawa, 2001).

$$P(K) = \int_0^\infty p(\phi, K).p(K) \qquad (41)$$

Where K expresses the prior knowledge, $p(K)$ represents the prior distribution and $p(\phi)$ represents the likelihood distribution.

There are three main categories of prior distributions in BA, these are conjugate, proper and informative, each one has pros and cons (Jurčiček, 2014). The most common category is the conjugate prior. Nevertheless, for the scope of this study BA is not going to be bounded only to these, as any type of prior could be useful to determine a relevant posterior distribution based on limited data and limited knowledge.

Conjugate prior distributions are distributions with similar shape and form than the likelihood PDF. Therefore, they form a conjugate family if the posterior density is the one shown in formula 42 (Lee, 2012). In this case, the posterior PDF is proportional to the conjugate prior.

$$P(x) \propto p(\phi).l(x) \qquad (42)$$

The previous is an advantage of a conjugate prior distribution, because a posterior distribution can be used as a new prior when updated data is available as the proportionality between both distributions remains (Jurčiček, 2014).

Informative prior distributions are formed by a population of possible parameter values, from which the parameter of interest is drawn (Gelman, 2014). Instead of including all possible values for a parameter, informative priors seek to concentrate in more truthful values around the true parameter (Gelman, 2014).

Improper prior distributions are extremely small over an infinite range of possible parameter values and they can be ambiguous with regards to the true parameter value. The uniform distribution (Formula 43) is a good example of a common distribution to account for prior knowledge when there is a large degree of uncertainty (Berger, 1985). In addition, it gives almost no valuable information to increase the probability for the parameter to be in the interval of the posterior.

$$p(\phi) = 1 \qquad (43)$$

In this thesis posterior distributions are obtained by using Monte Carlo simulations (MC). The procedure for generating a predictive prior of an exponential distribution in Matlab is presented as an example:

1. Determine the number of iterations for MC calculations of prior

```
nprior=input('Number of iterations for prior Parameters ')
```

2. Selecting the type of distribution according to the prior knowledge

```
for i=1:nprior
```

3. Draw an initial value for the parameter(s) of the PDF from the selected prior

```
lambda=rand()*5;
```

4. Generate a distributed random data set given the initial parameter(s)

```
estimator=exprnd(lambda,1,1);
```

5. Generate the predictive prior data by repeting step 3 with $nprior$ iterative calculations and the value of $\hat{\phi}_{prior\ Bayes}$

```
t(i)=exprnd(estimator,1,1);

end
```

The same procedure applies to the Weibull, Lognormal and Gamma distributions.

Elicitation Principles

Although there is no generic approach to elicit expert knowledge, the elicitation process should be structured, unbiased and aligned with the statistical model that will be used to incorporate expert information (Kuhnert, 2010). Several issues must be considered when eliciting knowledge with regards to how useful this knowledge will be to BA (Albert, 2012). Moreover, several challenges rise when eliciting knowledge from different experts . This is fundamental to ensure the closeness of $\hat{\phi}_{Prior}$ to $\hat{\phi}_{Bayes}$ and the true parameter value.

Any prior will push the estimate towards the mean of the prior distribution (Mickelsson, 2015). See Figure 12 to observe the influence of priors and likelihood PDF on the posterior.

Among such issues when eliciting expert knowledge is how to determine if the expert knowledge can be qualified as strong (Aven T. , Risk Analysis, 2012). To evaluate the strength of knowledge there is a list of five conditions that must be met so the expert knowledge is considered strong (Aven T. , Risk Analysis, 2012). Thus, the elicited information relevant. The conditions to be met for a strong knowledge are:

- The assumptions made are very reasonable
- Large amount of relevant/reliable data/information is available
- There is a broad agreement among experts
- The phenomena involved are well understood; the models used are known to give prediction with the required accuracy

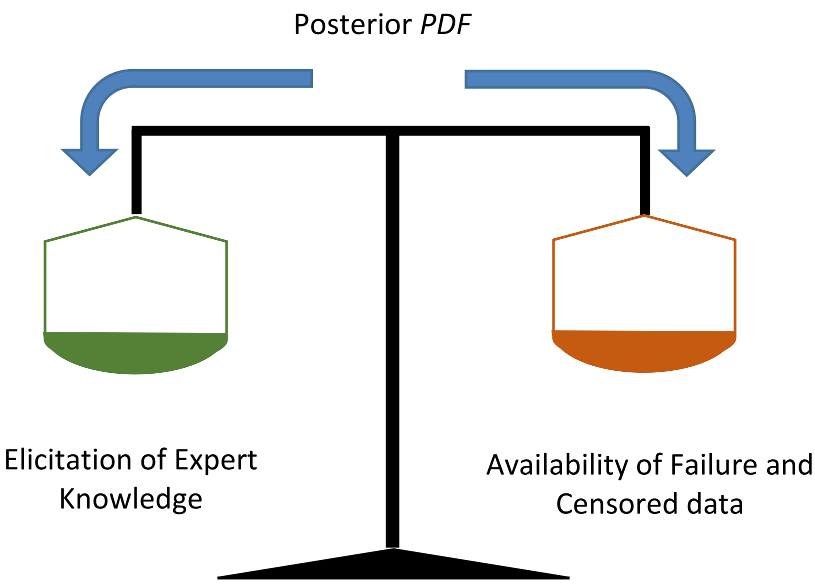


Figure 12. Illustrative influence of Prior and Likelihood on Posterior distribution

When there is no broad agreement between several experts and yet all experts' opinions seem to be relevant due to their own contributions, an alternative methodology is required (Samaniego, 2010). Assuming different opinions from k experts where the weight of each opinion is denoted by a factor a which meets the condition seen on formula 44,

$$\sum_{1}^{k} a_i = 1 \qquad (44)$$

Then the proposed $\hat{\phi}_{Prior}$ for the prior distribution is denoted by formula 45 (Samaniego, 2010).

$$\hat{\phi}_{Prior} = \sum_{i=1}^{k} a_i \hat{\phi}_i \qquad (45)$$

This constitute a relatively simple way to approach the consensus problem among experts when the elicited prior is not necessarily conjugate of the posterior and each expert only consider one prior for the phenomenon, not a combination of two or more. An example with two scenarios of expert opinions is provided in the appendix A6.

In practice many opinions for prior of real systems are more complex to examine as the weight for each one will tend to be similar. Therefore, the closer $\hat{\phi}_{Prior}$ will be to a standard mean (Samaniego, 2010). In addition, when eliciting expert knowledge for a few reliable sources, is not common to have such difference between the opinions with regards of prior parameter(s). The procedure used to estimate $\hat{\phi}_{Prior}$ for each of the four distributions is based on drawing $\hat{\phi}$

from a normally random generated DF. Then $\hat{\phi}_{Prior}$ from its corresponding PDF. Table 5 presents a summary of the procedure used in each case of PDF to generate $\hat{\phi}_{Prior}$. The initial values $\hat{\phi}$ can be modified easily in the code, according to expert knowledge of the specific study case.

Table 5. Summary of procedures to generate $\hat{\phi}_{Prior}$ for each PDF

Type of distribution	Procedure to generate $\hat{\phi}_{Prior}$
Exponential	`lambda=rand()*5;` `estimator=exprnd(lambda,1,1);`

Weibull	`alpha=normrnd(1,0.05);`
	`beta=normrnd(2,0.05);`
	`estimator=wblrnd(alpha,beta,1,2);`

Lognormal	`mu=0;`
	`sigma=0.25;`
	`estimator=lognrnd(mu,sigma,1,2);`

Gamma	`alpha=rand()*4;`
	`beta=rand()*2;`
	`estimator=gamrnd(alpha,beta,1,2);`

Posterior Distributions

To determine the posterior distribution directly from Bayes theorem requires computing the integral for the likelihood distribution (Kruschke, 2015). This can be challenging as the expression for the posterior distribution can be extremely difficult and time consuming. One method to address this challenge is using conjugate priors which are similar to posterior and easier to calculate (Kruschke, 2015).

In most cases, the type of posterior will not be strictly defined for different combinations of priors vs. likelihood distributions, i.e. the posterior will not behave the same way as a known PDF as its mathematical form will be similar but not equal to any of them. Therefore, the most recommended method to obtain the posterior is to generate a predictive posterior with random sampling. Otherwise, if one follows the analytical approach, the results are impractical and difficult to process since the possible combinations of priors vs. likelihoods are too many.

As previously mentioned, there are two major disadvantages when obtaining the posterior distribution following the analytical approach instead of using sampling.

- The types of PDFs for a posterior distribution given all possible combinations of prior vs likelihoods is too large to include in a single programming code. Hence, to produce and sample a posterior for reliability calculations can be time consuming and may require extensive computational capacity.

- The posterior DFs most likely will not follow any of the know PDF such as Exponential, Weibull, Lognormal, Gamma, etc. Moreover, it may result in undesired outcomes such as infinite, negative or extremely large values.

Based on the previous statement, it is better to obtain a predictive posterior distribution by randomly sampling a large amount of combinations of values with the pre-estimated parameters. This requires computer algorithms to do the iterative calculations, which are commonly referred as Monte Carlo simulations MC (Rossi, 2005). The algorithm used in Matlab is a pre-build function that generates random values which follow the pre-established condition between $p(\phi)$ and $l(x)$. The acceptance rejection sampling is useful to manage any type of prior other than conjugate in order to obtain a posterior (James, 2014).

In this thesis posterior distributions are obtained by the sampling approach using the prebuilt function of Matlab called acceptance rejection sampling. This function generates a specific type of posterior population formed by a fixed amount of data points by sampling the prior and likelihood distributions (Hanselman, 2012). In other words, the user can define the type of distribution to randomly generate the data points, the number of iterations or size of this posterior population and the criteria to accept or reject into the posterior each evaluated data point.

A code for BA has been written in Matlab following MC. The following is the procedure for BA that the software performs to obtain the predictive posterior distribution.

1. Presents four options of $PDFs$, Exponential, Weibull, Lognormal and Gamma, for the user to select one according to Prior Knowledge
2. Generate random values for the initial parameter(s) from a uniform distribution this can be $\hat{\lambda}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\mu}$ or $\hat{\sigma}$, according to the type of distribution. Asks for the number of iterations $nprior$ to generate prior distribution $p(\phi)$
3. Generates $p(\phi)$ and prior parameters
4. Generates the plot of $p(\phi)$
5. Generates the histogram of $p(\phi)$
6. Imports the likelihood data set from an external file. The codes only read Microsoft Excel (.xlsx) format files

7. Evaluates the goodness of fit for each PDF by the Kolmogorov-Smirnov test and rank them to select the best fitted one, i.e the highest p value among them
8. Generates random values for the initial parameter(s) from a uniform distribution this can be $\hat{\lambda}$, $\hat{\alpha}$, $\hat{\beta}$, $\hat{\mu}$ or $\hat{\sigma}$, according to the selected type of distribution
9. Asks for the number of iterations $nlikelihood$ to generate likelihood distribution $l(x)$
10. Estimates the parameter(s) of the selected likelihood function $l(x)$
11. Asks for the $1-\alpha$ degree of confidence intervals for the $l(x)$ parameter(s)
12. Calculates the confidence interval for each parameter
13. Performs an acceptance rejection sampling between $p(\phi)$ and $l(x)$. Here the code runs a predetermined number of iterations $nposterior$, which for accuracy purposes is equal to $1x10^6$
14. Generates the posterior PDF
15. Asks for a desired time $tsurvival$ to evaluate the reliability of the system given the posterior data set.
16. Calculates the reliability $R(t)$
17. Calculates the unreliability $F(t)$
18. Calculates the $MTTF$
19. Presents a summary of the results in tree tables. The first for types of PDFs, the second for number of iterations of prior, likelihood, posterior, results of $R(t)$, $P(t)$ and $MTTF$. The third table for the details of the $MTTF$ for years as the time unit from the date the programs is run
20. Generates the plot of $R(t)$
21. Generates the histogram of $R(t)$

A closer look into step 14^{th} of the procedure for BA that the Matlab is necessary, as this covers the acceptance rejection sampling for the predictive posterior distribution. The key factor to obtain a representative predictive posterior are the size of the population data sets, the criteria to generate the semirandom iterative values and the criteria to accept or reject such value into the predictive posterior itself. See appendix A8 for the code of the acceptance rejection sampling function of Matlab. Figure 13 presents a general diagram of the acceptance rejection sampling function.

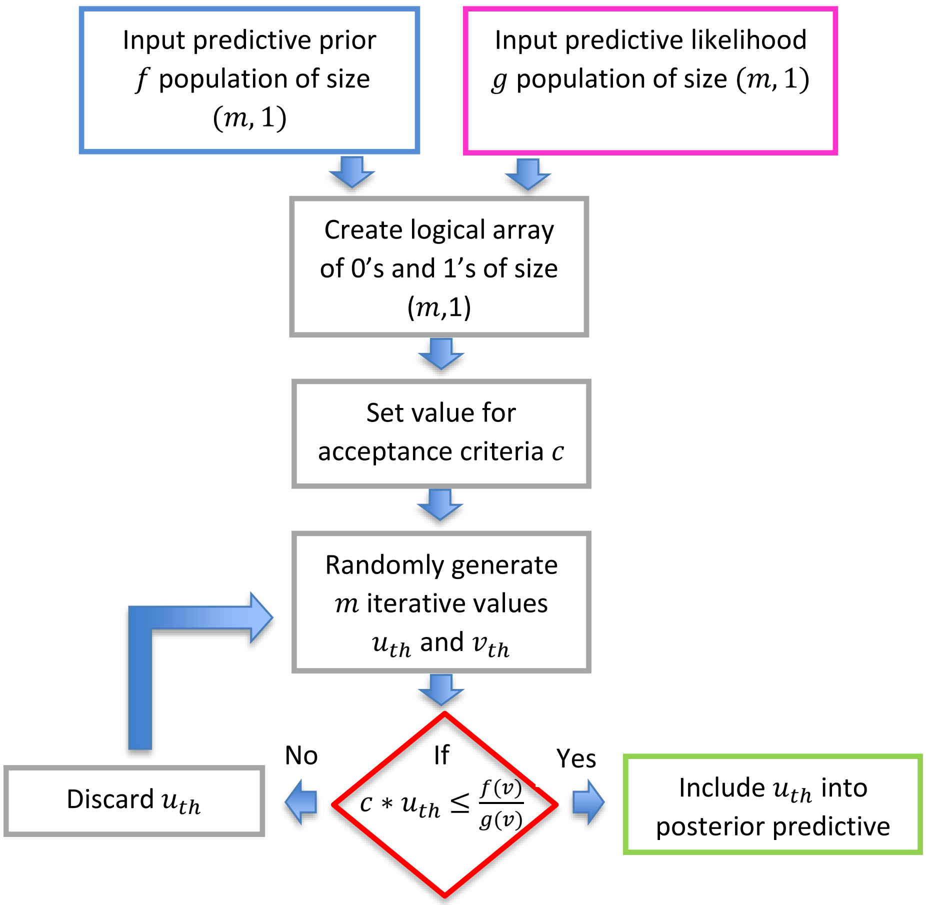


Figure 13. Diagram of the acceptance rejection sampling function in Matlab R2018.

Next is Figure 14 with a layout of the result tables for predictive posterior obtained with the *BA* Matlab code.

typeresults =

Prior	Likelihood
'Gamma'	'Weibull'

valueresults =

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+05	1e+06	1e+06	3	0.054628	0.94537	1.4297

mttfresults =

mttf	years	months	days	hours
1.4297	1	5	4	16.38

Figure 14. Layout of Results tables for predictive posterior with BA Matlab code.

CHAPTER 3

3.1 Reliability Data

Reliability is generally defined as the ability of a system, component or equipment to operate under specified conditions during a designated period of time or number of cycles (Modarres, 2010). The concept of reliability applies to every element that must perform an activity within an expected lifetime.

It is feasible to have different degrees of reliability based on the deterioration level of an element. Deterioration occurs gradually due to different physical and chemical variables which the elements are exposed to (Lydersen, 1988). Some of these variables for mechanical and electronical components are listed below.

- Temperature
- Mechanic Stress
- Torque
- Friction
- Water depth
- Pressure
- PH level
- Dust
- Moisture
- Number of operations

To include the level of degradation in a reliability assessment requires a refined analysis of the multiple variables to be considered and the criticality of each one towards the element/system's performance (Lydersen, 1988).

In the scope of this study the main assumption is that system performance is either satisfactory or unsatisfactory. This proposes a binary measurement system; if the system performs its required functions it is considered 100% reliable (Tobias, 2012), regardless the conditions under it operates or the degree of degradation that it may have. Therefore, all the data sets in this study provide failure times. These express the completely inability of the system to operate as it is supposed to.

An overview of failure states and degradation levels for reliability assessment is presented in Figure 15.

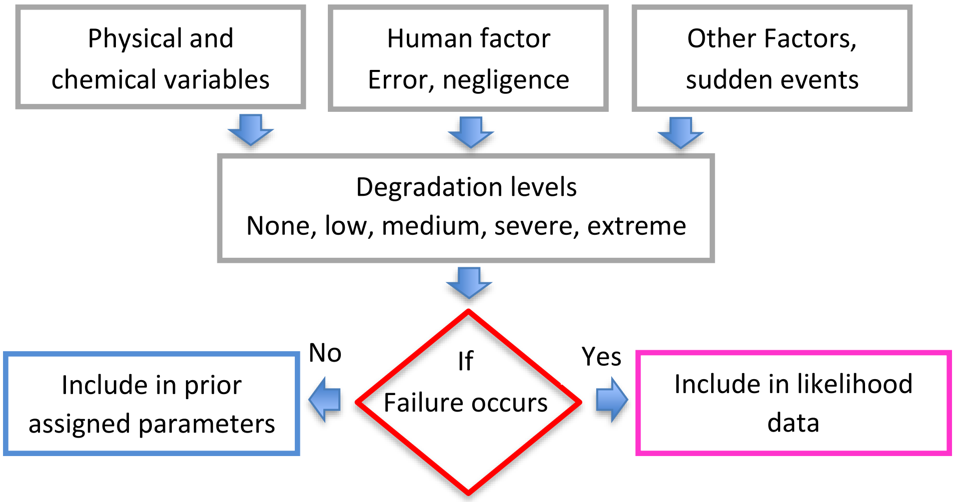


Figure 15. Overview of failure states and degradation levels for reliability assessment.

A list of assumptions as follows for BA:

1. No replacement is considered. Hence, data sets show failure times or similar components before being replaced for the first time. See the expected lifetime without replacement in Figure 16.
2. The time unit for failures is not fixed, i.e. the software codes do not differentiate between, hours, days, years, etc. unless required by the user.
3. For the study case of chapter 3, a large amount of data contains several types of events that generate the failure of the systems. These events are not categorized nor classified; therefore, they are treated in the same way.
4. Since there is no replacement the time to fix or replace a failed component is not considered.
5. The initial values for prior parameters are pseudorandom as they will be the product of prior knowledge and reasonable numerical intervals. Otherwise, when selecting $R(t)$, $F(t)$ and $MTTF$, the outcomes can be extremely large and, in some cases, unrealistic.
6. The number of iterations for prior, likelihood and posterior DF must be large to increase the accuracy of the results. The minimum recommended value is 1×10^4.

Once the assumptions are established, it is necessary to quantitatively assess the reliability of a system. Hence, previous data about its behaviour is required. Failure times, types of failure, number of failed units, conditions for system's operations and more constitute the reliability data (Modarres, 2010).

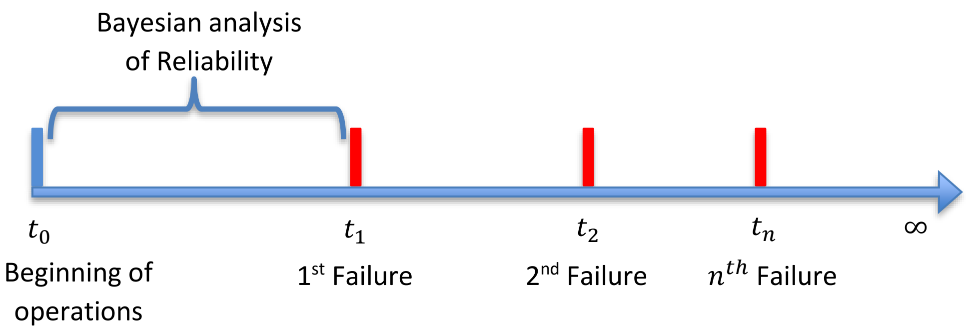


Figure 16. Expected lifetime without replacement.

Reliability data may have different formats to express the previous performance of the systems under analysis (Meeker, 1998). Among these formats are:

- Histograms
- Plots
- Tables
- Subsets of data
- Databases

As mentioned before, the time unit of the reliability data is not fixed as the program is flexible to handle any unit. In other words, failure times can be hours, days, months, years, number of cycles, number of revolutions, etc. Hence, the result's time unit will be the same as the data input. Nevertheless, the default time unit for the results is years. $MTTF$ is shown in years, months, days and hours.

Historical data is required to estimate the expected lifetime through maximum likelihood parameters of a PDF. However, fitting a given dataset to a distribution model may represent a challenge (Renyan, 2015). The relevance and importance of such data relies on the amount available and the main assumption that conditions of operations are maintained through time. In addition, censoring also influences reliability data (Meeker, 1998).

3.1.1 Random data set

Random data sets do not follow any pattern or statistical distribution as they are generated in an entirely aleatory process . They are used to represent uncertainty in models, MC simulations, iterative calculations, acceptance rejection sampling and others. For practical uses random sets are limited to upper and lower boundaries.

The **rand** function is used to generate a random data set in (Hanselman, 2012), where, size of the array or number of elements to be generated must be specified. By default, the boundaries of a **rand** function are within 0 and 1. However, by multiplying a scalar value one can expand such boundaries to any given interval (Hanselman, 2012).

```
a=rand(m,n)*scalar
```

a represents the data set to be created, m the number of columns, n the number of rows and $scalar$ the upper limit of the data set.

Pseudo random data sets can also be generated in Matlab. These types of data sets follow a given pattern such as a PDF. The following is an example of a Lognormal data set b that randomly generates $(100,1)$ elements within a 0 to 4 interval.

```
b=lognrnd(100,1)*4
```

Table 6 shows a theoretical example of a random data set, used to compare the results of using MLE of each PDF used in this report, and BA. The goal is to determine the $MTTF$, $R(t)$ and $F(t)$ in each case using Matlab.

A program was created to settle a random distributed data, evaluate it according to the PDF, obtain the parameters, corresponding confidence intervals, probability of survival $P(T \geq t)$, probability of failure $P(T \leq t)$, produce the histogram and finally $MTTF$.

Table 6. Random data set for MLE and BA example

t time values randomly generated within 0 to 5 units interval								
1.31	0.96	0.43	4.60	3.71	1.95	0.83	4.34	3.78
4.70	2.31	2.76	4.79	4.29	0.14	2.99	3.21	4.04
4.54	2.88	0.42	3.05	4.69	3.93	3.77	3.51	2.52
4.90	3.07	0.03	0.71	3.59	1.99	3.26	4.47	3.65
3.54	4.33	1.54	1.91	2.79	3.24	2.83	3.89	4.78
0.78	3.91	3.08	3.34	0.48	1.46	1.55	4.98	0.48
4.66	2.12	3.17	0.79	0.75	3.95	0.17	4.74	4.83
3.09	4.22	3.41	0.13	2.29	0.22	4.55	4.88	0.06
0.17	4.53	1.58	0.08	2.00	2.52	4.56	0.90	0.46
2.60	2.03	3.63	1.93	2.09	1.12	3.24	0.83	0.04
3.72	4.92	1.59	3.88	2.96	2.58	1.99	3.94	1.31

Exponential Distribution

The results for the example for MLE of Exponential distribution are presented in the summary Table 7.

Table 7. Summary Table for MLE of Exponential random data set using Matlab.

Data	**Parameter**		**Confidence Interval**			**T**	**Expected Lifetime**		
n	**λ**	**$Muhat$ $\hat{\lambda}$**	**deg %**	**$Muci$**		**$tsur$**	**$R(t)$**	**$F(t)$**	**$MTTF$**
Fixed n and variable λ, deg and t									
100	1.04	**2.651**	5	2.200	3.259	1	**0.070**	**0.929**	**0.377**
	2.22		10	2.266	3.151	2	**0.005**	**0.995**	
	0.09		15	2.311	3.084	3	**$3x10^{-3}$**	**0.998**	
	2.01		20	2.346	3.033	4	**$2.47x10^{-5}$**	**0.999**	
	1.85		25	2.376	2.992	5	**$1.74x10^{-6}$**	**1**	

Table 7, n is the number of observations for each case of $tsurvival$. $Landa$ (λ) is initial value of the estimated parameter. The failure rate for the exponential PDF. $Muhat$ is the estimated parameter $\hat{\lambda}$. $Muci$ is the confidence interval with the upper and lower bounds for the estimated parameter. deg (α) is the % of freedom for the confidence interval. $R(t)$ represents the reliability or survival function and $F(t)$ the unreliability given a time value t.

To interpret the results of Table 7 let us take row 2 for example. The number of elements from the historical data sample is 100. The initial random value generated by the program for the parameter λ is 1.04. The estimated value of parameter λ is 2.6518 with a 90% of confidence that the real value will be within 2.2666 and 3.1517. For a selected time of 2 units the system has a probability of survival of 0.5% and a probability of failure of 99.5% with a fixed $MTTF$ of 0.3771.

Table 7 shows the variation of boundaries for the confidence interval of the parameters as the % of freedom decreases. The higher the % of freedom, the more accurate is the estimated parameter. In addition, it also shows how $R(t)$ and $F(t)$ complement each other to the total value of 1. Therefore, whilst one decreases the other increase in the exact proportion. Finally, the results show a constant value of $MTTF$ regardless of the selected t and estimated parameter.

The results for the example of *BA* with exponential prior distribution are presented in the summary Table 8.

Table 8. Summary Table for BA with Exponential Prior and random data set using Matlab.

Data	**Number of Iterations**			**T**	**Expected Lifetime**		
$\boldsymbol{n}$	$\boldsymbol{nprio}$	$\boldsymbol{nlikeli}$	$\boldsymbol{nposte}$	$\boldsymbol{tsur}$	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	$\boldsymbol{MTTF}$
Fixed n, $nprior$, $nlikelihood$ vs. variable λ and $tsurvival$							
100	1x10^4	1x10^4	1x10^4	1	**0.2495**	**0.7505**	**0.7152**
				2	**0.1468**	**0.8532**	**0.7086**
				3	**0.0796**	**0.9204**	**0.7187**
				4	**0.0392**	**0.9608**	**0.7171**
				5	**0.0143**	**0.9857**	**0.7021**

Several observations can be drawn from the comparison of results on Tables 7 and 8. First, the number of iterations for prior and likelihood influence the $R(t)$ and $F(t)$ values as the sample is random generated. Hence the larger $nprior$ $nlikelihood$ and $nposterior$ the more accurate will $R(t)$ and $F(t)$ be. Second, $R(t)$ and $F(t)$ tend to decrease as time increases. Third, the influence of predictive prior on the $MTTF$. Which now is around 0.71 years instead of 0.377 as shown in Table 7. Fourth, $MTTF$ values in BA for each case are not identical as they are in

MLE. Nevertheless, all five values are very close to each other. Figure 17 presents the estimation of $R(t)$, $F(t)$ and $MTTF$ when $tsurvival = 5$.

Prior	Likelihood	Posterior
'Exponential'	'Weibull'	'Bayesian'

valueresults =

1×6 table

nprior	nposterior	tsurvival	psurvival	pfailure	mttf
10000	10000	5	0.0143	0.9857	0.70201

mttfresults =

1×5 table

mttf	years	months	days	hours
0.70201	0	8	12	17.356

Figure 17. BA with Exponential Prior, random data set for tsurvival=5

Weibull Distribution

The results of the example for MLE with Weibull distribution are presented in the summary table 9.

Table 9. Summary Table for MLE of Weibull random data set using Matlab.

Data	Parameters & Estimators				Confidence Intervals			T	Expected Lifetime		
n	**α**	**β**	**$Muhat$** $\hat{\alpha}$	$\hat{\beta}$	**deg %**	**$Muci$**		**$tsur$**	**$R(t)$**	**$F(t)$**	**$MTTF$**
Variable t and deg, fixed n, α, β											
100	1.9	1.2	**2.886**	**1.497**	5	[2.519, 3.305]	[1.261, 1.778]	1	**0.815**	**0.184**	**2.605**
	0.1	0.5			10	[2.575, 3.234]	[1.296, 1.730]	2	**0.561**	**0.438**	
	1.9	1.2			15	[2.612, 3.188]	[1.320, 1.699]	3	**0.346**	**0.653**	
	0.8	1.8			20	[2.641, 3.153]	[1.338, 1.6758]	4	**0.195**	**0.804**	
	1.2	0.1			25	[2.665, 3.125]	[1.354, 1.656]	5	**0.102**	**0.897**	

In Table 9, n is the number of observations for each case of $tsurvival$. $Alpha$ (α) and $beta$ (β) are the initial values of the estimated parameters. The scale and shape for the Weibull PDF. $Muhat$ are the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$, respectively. $Muci$ are the confidence intervals with the upper and lower bounds for the estimated parameters. deg (α) is the % of freedom for the confidence intervals. R represents the survival function $R(t)$ given a time value t.

Table 9 shows the variation of boundaries for confidence interval of the parameters as the % of freedom decreases. The higher the % of freedom, the more accurate the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$. In addition, Table 9 also shows that $R(t)$ and $F(t)$ complement each other to the total value of 1. Therefore, whilst one decreases the other increase in the exact proportion. Finally, it shows a constant value of $MTTF$ regardless of the selected ts and estimated parameter.

To interpret the results of Table 9 let us take row three for example. The number of elements from the historical data sample is 100. The initial random value generated by the program for the parameters α and β are 1.96 and 1.25 respectively. The estimated values by the program are 2.886 and 1.497 with a 85% of confidence that real values will be within [2.612, 3.188] and [1.320, 1.699] accordingly. For a selected time of 2 units the system has a probability of survival of 34.6% and a probability of failure of 65.3% with a fixed $MTTF$ of 2.605.

The results for the example of *BA* with Weibull prior distribution are presented in the summary Table 10.

Table 10. Summary Table for BA with Weibull Prior and random data set using Matlab.

Data Set	**Number of Iterations**			**T**	**Expected Lifetime**		
$\boldsymbol{n}$	$\boldsymbol{nprio}$	$\boldsymbol{nlikeli}$	$\boldsymbol{nposte}$	$\boldsymbol{t}$	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	$\boldsymbol{MTTF}$
Fixed n, $nprior$, $nlikelihood$ vs. variable α, β and $tsurvival$							
100	1x10^4	1x10^4	1x10^4	1	**0.0730**	**0.9227**	**0.2173**
				2	**0.0383**	**0.9617**	**0.2131**
				3	**0.0150**	**0.9850**	**0.1988**
				4	**0.0082**	**0.9918**	**0.2066**
				5	**0.0038**	**0.9962**	**0.2212**

Several observations can be drawn from the comparison of results on Tables 9 and 10. First, the number of iterations for prior and likelihood influence the $R(t)$ and $F(t)$ values as the sample is randomly generated. Therefore, the larger $nprior$ $nlikelihood$ and $nposterior$ the more accurate will $R(t)$ and $F(t)$ will be. These are fixed to 1x10^4 iterations for this trial. Second, $R(t)$ and $F(t)$ tend to decrease for as time increases. Third, the influence of predictive prior on the $MTTF$, which now is around 0.21 years instead of 2.60 as shown in table 9. Fourth, $MTTF$ values in BA for each case are not identical as they are in MLE. Nevertheless, all five values are close enough to each other, so an estimation is feasible.

Figure 18 presents the estimation of $R(t)$, $F(t)$ and $MTTF$ when $tsurvival = 5$.

```
    'Weibull'    'Weibull'     'Bayesian'

valueresults =

  1×7 table

    nprior    nlikelihood    nposterior    tsurvival    psurvival    pfailure    mttf
    ______    ___________    __________    _________    _________    ________    ______

    10000        10000          10000          5          0.003        0.997     0.2097

mttfresults =

  1×5 table

     mttf     years    months    days    hours
    ______    _____    ______    ____    ______

    0.2097      0        2        15     11.796

MTTF is estimated from todays date
```

Figure 18. BA with Weibull Prior, random data set for tsurvival=5

Lognormal Distribution

The results of the example for MLE of Lognormal distribution are presented in the summary Table 11.

Table 11. Summary Table for MLE of Lognormal random data set using Matlab

Data	Parameters				Confidence Intervals			T	Expected Lifetime		
n	μ	σ	***Muhat*** $\hat{\mu}$	$\hat{\sigma}$	***deg %***	***Muci***		***tsur***	$R(t)$	$F(t)$	***MTTF***
Variable t and deg, fixed n, α, β											
100	1.0	0.7	**2.856**	**0.933**	5	[2.669, 3.042]	[0.823, 1.089]	1	**0.998**	**0.001**	**4.415**
	0.7	0.4			10	[2.700, 3.011]	[0.841, 1.063]	2	**0.989**	**0.010**	
	0.2	0.7			15	[2.719, 2.992]	[0.852, 1.0469]	3	**0.970**	**0.029**	
	0.6	2.8			20	[2.735, 2.977]	[0.861, 1.034]	4	**0.942**	**0.057**	
	1.5	0.2			25	[2.747, 2.964]	[0.869, 1.024]	5	**0.909**	**0.090**	

Table 11, n is the number of observations for each case of $tsurvival$. $Mu\ (\mu)\ and\ Sigma\ (\sigma)$ are the parameters to be estimated. The median and standard deviation for the Lognormal PDF

are the initial values of the estimates $\hat{\mu}$ and $\hat{\sigma}$, respectively. $Muci$ are the confidence intervals with the upper and lower bounds for the estimated parameters. deg (α) is the % of freedom for the confidence intervals. $R(t)$ represents the reliability or survival function and $F(t)$ the unreliability given a time value t.

Table 11 illustrate the variation of boundaries for the confidence intervals of the parameters as the % of freedom decreases, The higher the % of freedom, the more accurate the estimated parameters $\widehat{\mu}$ and $\widehat{\sigma}$. Moreover, Table 11 also shows that $R(t)$ and $F(t)$ complement each other to the total value of 1. Therefore, whilst one decreases the other increase in the exact proportion. Finally, table 11 also shows the fixed value for $MTTF$ regardless of the selected ts and estimated parameter.

To interpret the results of Table 11 let us take row four as an example. The number of elements from the historical data sample is 100. The initial random value generated by the program for the parameters μ and σ are 0.64 and 2.85 respectively. The estimated values by the program for these parameters are 2.856 and 0.933 with an 80% of confidence that real values will be within [2.735, 2.977] and [0.861, 1.034] accordingly. For a selected time of two units the system has a probability of survival of 34.6% and a probability of failure of 65.3% with a fixed $MTTF$ of 2.605.

The results for the example of *BA* with Lognormal prior distribution are presented in the summary Table 12.

Table 12. Summary Table for BA with Lognormal Prior and random data set using Matlab.

Data	Number of Iterations			T	Expected Lifetime		
n	$nprio$	$nlikel$	$nposte$	$tsur$	$R(t)$	$F(t)$	$MTTF$
Fixed n, $nprior$, $nlikelihood$ vs. variable $tsurvival$							
100	$1x10^4$	$1x10^4$	$1x10^4$	1	**0.4591**	**0.5409**	**1.3265**
				2	**0.2722**	**0.7278**	**1.3199**
				3	**0.1526**	**0.8474**	**1.3394**
				4	**0.0748**	**0.9252**	**1.3327**
				5	**0.0320**	**0.9680**	**1.3192**

By comparing the results obtained in Tables 11 and 12, one can draw the following conclusions, first, $R(t)$ values for BA are lower to its corresponding MLE pair. Second, the tendency of $R(t)$ to decrease as time increases is more visible for BA, for example the 3rd row for time 3 years $R(t)$ is 0.1526. Therefore, the number of iterations for prior and likelihood influence the $R(t)$ and $F(t)$ values as the sample is randomly generated. Hence, the larger $nprior$ $nlikelihood$ and $nposterior$ the more accurate will $R(t)$ and $F(t)$ will be. These are fixed to $1\text{x}10^4$ iterations for this trial. Third, the influence of predictive prior on the $MTTF$, which now is around 1.32 years instead of 4.41 as shown in Table 9. Fourth, $MTTF$ values in BA for each case are not identical as they are in MLE. Nevertheless, all five values are close enough to each other, so an estimation is feasible.

See Figure 19 for estimation of $R(t)$, $F(t)$ and $MTTF$ when $tsurvival = 5$.

Prior	Likelihood	Posterior
'Lognormal'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
10000	10000	10000	5	0.032	0.968	1.3192

mttfresults =

1×5 table

mttf	years	months	days	hours
1.3192	1	3	24	21.839

MTTF is estimated from todays date

Figure 19. BA with Lognormal Prior, random data set for tsurvival=5

Gamma Distribution

The results for the example for MLE of Gamma distribution are presented in the summary Table 13. Where n represents the observations for each case of the gamma random distributed data set. $Alpha\ (\alpha)\ and\ beta\ (\beta)$ are the initial values of the estimates. The scale and shape for the gamma PDF are the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$. $Muci$ are the confidence intervals with the upper and lower bounds for the estimated parameters. deg (α) is the % of freedom for the confidence intervals. R is the survival function $R(t)$ for a time value t.

To interpret the results of Table 13 let us take row 5 for example. The number of elements from the historical data sample is the same as previous cases (100). The initial random value generated by the program for the parameters α and β are 1.40 and 0.49 respectively.

Table 13. Summary Table for MLE of Gamma random data set using Matlab

Data	Parameters				Confidence Intervals			T	Expected Lifetime		
n	α	β	*Muhat* $\hat{\alpha}$	$\hat{\beta}$	*deg* %	*Muci*		*tsur*	$R(t)$	$F(t)$	*MTTF*
Variable t and deg, fixed n, α, β											
100	1.21	0.33	**3.446**	**0.905**	5	[2,641, 4,4982]	[0.679, 1.206]	1	**0.203**	**0.796**	**3.122**
	1.85	0.89			10	[2.204, 5.389]	[0.559, 1.465]	2	**0.109**	**0.890**	
	1.70	0.25			15	[2.239, 5.305]	[0.569, 1.440]	3	**0**	**1**	
	0.57	0.77			20	[2.265, 5.244]	[0.576, 1.422]	4	**0**	**1**	
	1.40	0.49			25	[2.286, 5.197]	[0.582, 1.408]	5	**0**	**1**	

Table 13 shows the variation of boundaries for confidence interval of the parameters as the % decreases, which means that to be more confidence of the estimated parameters $\hat{\alpha}$ and $\hat{\beta}$, scale and shape for the Weibull distribution, the interval must enhance. In addition, Table 13 also shows that $R(t)$ and $F(t)$ complement each other to the total value of 1. Therefore, whilst one decreases the other increase in the exact proportion. Finally, Table 13 displays the fixed value of $MTTF$ independently of the selected ts and estimated parameter.

The estimated parameters $\hat{\alpha}$ and $\hat{\beta}$ by the program are 3.446 and 0.905, respectively, with a 75% of confidence that real values will be within [2.286, 5.197] and [0.582, 1.408], accordingly. For a selected time of 2 units the system has a probability of survival of approximately 11% and a probability of failure of 89% with a fixed $MTTF$ of 3.122.

The results for the example of *BA* with Gamma prior distribution are presented in the Table 14.

Table 14. Summary Table for BA with Gamma Prior and random data set using Matlab

Data	Number of Iterations			T	Expected Lifetime		
n	*nprio*	*nlikel*	*nposte*	*tsur*	$R(t)$	$F(t)$	*MTTF*
Fixed n, $nprior$, $nlikelihood$ vs. variable $tsurvival$							
100	$1x10^4$	$1x10^4$	$1x10^4$	1	**0.3323**	**0.6677**	**1.3774**
				2	**0.2722**	**0.7278**	**1.3535**
				3	**0.2005**	**0.7995**	**1.3680**
				4	**0.1367**	**0.8633**	**1.3242**
				5	**0.0880**	**0.9120**	**1.3405**

By comparing the results in Tables 13 and 14, one can draw the following conclusions. First, $R(t)$ values for BA are lower to its corresponding MLE pair. Second, the tendency of $R(t)$ to decrease as time increases is more visible for BA, for example the 5th row for a $tsurvival$ of 5 years $R(t)$ is equal to 0.0888. Which means that this system has an extremely low possibility to survive five years. Third, the influence of predictive prior on the $MTTF$, which now is around 1.35 years instead of 3.12 as shown in Table 13. Fourth, $MTTF$ values in BA for each case are not identical as they are in MLE. Nevertheless, all five values are close enough to each other, so an estimation is feasible.

Figure 20 shows the results for estimation of $R(t)$, $F(t)$ and $MTTF$ when $tsurvival = 5$.

Further discussions about the pros and cons of both estimation methods will be presented in following sections of this thesis.

```
    'Gamma'    'Weibull'    'Bayesian'

valueresults =

  1×7 table

    nprior    nlikelihood    nposterior    tsurvival    psurvival    pfailure    mttf
    ______    ___________    __________    _________    _________    ________    ______

    10000        10000         10000           5          0.088        0.912     1.3405

mttfresults =

  1×5 table

     mttf     years    months    days    hours
    ______    _____    ______    ____    ______

    1.3405      1        4        2      14.086

MTTF is estimated from todays date
```

Figure 20. BA with Gamma Prior, random data set for tsurvival=5

3.1.2 Censored data set

In this study which includes reliability modelling and expected lifetime estimation the data sets can be considered censored. Definitions for each type of censored data set are discussed in the succeeding paragraphs. In addition, an illustrative example is presented.

Censored data partially contains the value of observation (Renyan, 2015). There are three types of censored data: left censored, right censored and interval censored (Renyan, 2015), considering the exact time at failure t_f and the time to inspection t_i.

- Left censored data occurs when failures are not obvious and can only be detected by inspections. Hence, $0 < t_f < t_1$

- Right censored data takes place when the observation process stops before the failure occurs. Hence, $t_1 < t_f < \infty$
- Interval censored data occurs when observational times are planned at some specific time during inspections. Hence, $t_1 < t_f < t_2$

The following is an illustrative example of a data set from a standard system with the different types of censoring, according to different times of inspection t_i in years and number of failures n_i.

Table 15. Illustrative example of dataset with different types of censoring

Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Occurred Failures	0	0	0	0	0	n_1	0	0	n_2	n_3
Observed Failures 1st Case	0	0	0	0	0	$n_1 - 1$	0	0	n_2	n_3
Observed Failures 2nd Case	0	0	0	0	0	n_1	0	0	n_2	$-$
Observed Failures 3rd Case	$-$	$-$	$-$	$-$	$-$	$-$	$-$	0	n_2	$-$

From Table 15 the 1st case represents a left censored data set with a total number of observed failures equal to $n_1 - 1 + n_2 + n_3$ and observation period from year 1 to 10. The 2nd case represents a right censored data set with a total of observed failures equal to $n_1 + n_2$ and observation period from year 1 to 9. Finally, the 3rd case shows an interval censored data set with a total number of observed failures equal to n_2 and observation period of two years, the 8th and 9th year. The influence of censoring in the data set can be measured, given random values of the number of failures n_1, n_2 and n_3 and assigning an PDF to determine the expected lifetime of such system. See appendix A1.

From the appendix Table A1 it is evident that different failure rates will be provide the censoring of the data set. In this case the interval censoring from the 3rd Case has the lowest estimated parameter $\widehat{\Phi}_{3^{rd}}$.

In the MLE and BA codes, censoring is expressed by the 2nd type, i.e., right censoring from a given t_i according to the theoretical example and study case.

Table 16 presents the failure times of a component in hours. It will also be used to test the program for MLE and BA. The outcomes will be the estimated $PDF's$ parameters, expected lifetime, reliability, probability of failure at any given time, $MTTF$, and the histogram of the failures occurred (Meeker, 1998).

Table 16. Integrated Circuit Failure Times in Hours

Failure times t per unit tested					
0.10	0.10	0.15	0.60	0.80	0.80
1.20	2.50	3.00	4.00	4.00	6.00
10.00	10.00	12.50	20.00	20.00	43.00
43.00	48.00	48.00	54.00	74.00	84.00
94.00	168.00	263.00	593.00	-	-
The test ended at 1370 hours, there were 4128 non-failed units.					

MLE and BA results are presented to compare them based on the method of estimation and the selected prior distribution. The number of iterations for predictive prior, likelihood and posterior are fixed to $1x10^6$.

Table 17. MLE and BA results for Censored data set

Method	**Number of Iterations**			**T**	**Expected Lifetime**		
MLE	***nprio***	***nlikel***	***nposte***	***tsur***	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	***MTTF***
Fixed n, $nprior$, $nlikelihood$ vs. variable and $tsurvival$							
Gamma	$1x10^6$	$1x10^6$	$1x10^6$	1	0.0430	0.9561	0.1573
BA	***nprio***	***nlikel***	***nposte***	***tsur***	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	***MTTF***
Exponential	$1x10^6$	$1x10^6$	$1x10^6$	1	0.0780	0.9219	0.9609
Weibull					0	1	0.1986
Lognormal					0.1492	0.8508	1.9593
Gamma					0.1276	0.8723	2.2904

The Gamma PDF has the best goodness of fit. Therefore, it was chosen as the likelihood distribution used for all the BA cases. In addition, $tsurvival$ is expressed in years. Hence, the data values were converted from hours to years. (Figure 21), the $MTTF$ is estimated to be 1 month, 26 days and 15 hours to first failure. In the other hand, for BA, the $MTTF$ is estimated to be 2 years, 3 moths, 14 days and 13 hours, (Figure 22).

Reliability values for both estimations also vary. The most optimistic case takes place with the Gamma distribution with and $R(t)$ of 14.92%. Whilst the worst case is the product of Weibull distribution, where $R(t)$ has such a small value that is considered zero. In other words, the integrated circuit units will not survive a year of operations.

```
psurvival =
    0.0439
Probability of Failure at a given tsurvival
pfailure =
    0.9561
Mean time to failure of MLE Predictive Distribution
mttf =
    0.1573
Histogram of Weibull Distribution
mttfresults =
  1×5 table
```

mttf	years	months	days	hours
0.15731	0	1	26	15.194

Figure 21. MLE results for tsurvival $= 1$ *year of censored dataset.*

The difference between both $MTTF$ is based on the influence of the predictive prior distribution, as it will modify the scale and shape parameters of the Gamma PDF in this example. Moreover, the likelihood distribution and parameters are the same for both estimations. Another factor that influences the estimation of expected lifetime is the number of elements in the censored data. For this example, 4128 from a total of 4156 failures are right censored. This represents a source of uncertainty for the obtained results because most values are censored. Hence, the sample data is not completely representative of the component's behaviour.

Prior	Likelihood	Posterior
'Gamma'	'Gamma'	'Bayesian'

```
valueresults =
  1×7 table
```

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	1	0.12764	0.87236	2.2904

```
mttfresults =
  1×5 table
```

mttf	years	months	days	hours
2.2904	2	3	14	13.41

```
MTTF is estimated from todays date
```

Figure 22. BA results for tsurvival $= 1$ *year of censored dataset.*

3.2 Highly Reliable Systems

A highly reliable systems is a critical system which retain its reliability for prolonged periods of time as it evolves without incurring in prohibitive costs (Hinchey, 2010).

When discussing highly reliable systems, it is necessary to define several related terms. First and most important, what is reliability? When do we consider a system reliable? What are the requirements needed to consider a system highly reliable? And what are the most common probability distributions to predict the expected lifetime of a system?

These are the questions intended to be answered so the reader can have a clear picture with regards to highly reliable systems and the methodology used to determine the expected lifetime of such systems in the domain of this thesis.

First, the term reliability has several definitions depending on the author and context in which is referred to. One of the most standard definitions is the following: "The ability of an item to perform a required function under given environmental and operating conditions and for a stated period of time" (Hamada, 2008). A similar definition about the cycles of operations states, "The ability of an item (a product or a system) to operate under designated operating conditions for a designed period of time or number of cycles" (Modarres, 2010).

Second, to quantitatively measure the reliability of an item given a random variable T mathematically the reliability function $R(t)$ of a component or systems is defined by formula 11 (Hamada, 2008).

In general, highly reliable systems are those which do not require neither maintenance nor replacement over a long period of time (Pham, 2001). However, certain criteria for types of systems can help limit the scope of highly reliable systems:

1. Those in which configurations of its components are such that more than one level of redundancy is present
2. Those in which the specification of its components is oversized to the average requirements
3. Those designed to avoid bottlenecks
4. Those which have layers of protection against extreme external elements
5. Those designed to outlast the expected time for operations by a considerable amount of time

6. Those which allow follow up in a practical and quick way to detect needed maintenance

Consider a theoretical example of a manufacturing process that requires to produce 200 units per hour with two alternative systems A and B. System A has 4 identical components with capacity of 100 units per hour disposed in parallel and a monitoring equipment M. Meanwhile system B has two units of 50 and 150 units per hour. Given the previous criteria System A is considered to be the highly reliable one as it has two back up components, all its components are identical, it has a monitoring system and the configuration of components allows redundancy, (Figure 23).

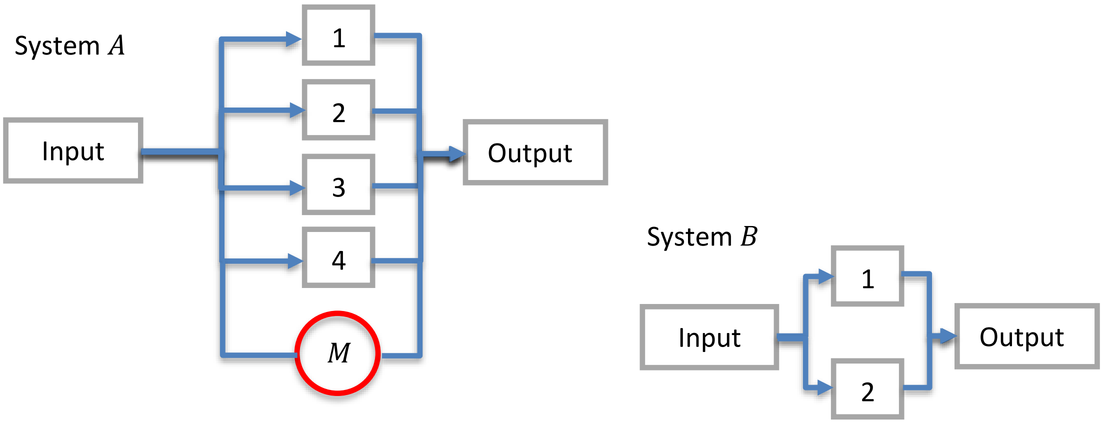


Figure 23. Theoretical example of Highly reliable systems for Manufacturing process

3.2.1 Risk and Uncertainty in Highly Reliable Systems

All processes and systems have a certain degree of risk and uncertainty in terms of their performance. The possibility of failure is ever present. Therefore, it becomes a key aspect when assessing the reliability of such systems. Furthermore, failure probability must be determined accounting for expert knowledge about the systems, its characteristics, strengths and weaknesses (Aven T. , Risk Analysis, 2012). Nevertheless, in most cases historical data is also necessary to complete a reliability assessment (Tobias, 2012).

When to consider a system robust? How likely is that a failure occurs at a specific time? What consequences derive when the system has a failure? What measures should be taken to prevent the occurrence of a failure? These among other questions are relevant to perform a reliability assessment. This thesis focuses on the 2nd, as the goal is to produce and verify a methodology for statistical inference of expected lifetime of highly reliable systems. For that purpose, the following assumptions need to be made.

- All types and magnitudes of failures from historical data are equal. Hence, partial or total failures that prevent the system from performing its purposes have the same value and are equally included in the analysis.
- The negative consequences of a failure are not considered for the analysis. Only the times in which they occur if observed, or the censored time if not.
- Maintenance times are not included in the analysis, for mainly the expected time to failure is of relevance for the study.
- Maintenance is assumed to be perfect. Therefore, the system is considered "as good as new ".
- The degree of uncertainty in the elicited expert knowledge is proportional to the overall uncertainty of the results. Therefore, a high level of accuracy is necessary.
- The risk of an unpredicted failure in a highly reliable system is influenced by the amount of failure and censored data available when assessing its reliability.
- The occurrence of events that can lead to a failure but given some specific circumstances did not, are not included in the reliability assessment. In other words, "near misses" are not accounted for.

The relevance for the estimation of expected lifetime of highly reliable systems from the risk perspective is based on the potential negative consequences once a failure has occurred and the complexity of restoring the system's functionality back to normal (Aven T. , Risk Analysis, 2012). Therefore, the more accurate methodology for assess the reliability of such systems the better. In addition, analysts can account for uncertainty to some extent by taking a conservative approach when eliciting prior knowledge about key factors (Albert, 2012). Among them, the parameters of the prior's PDF (Albert, 2012). Furthermore, by providing confidence and credibility intervals for MLE and BA. These allows the analyst and decision makers to have additional information about the provided values for the system's reliability so they can take a better-informed decision.

3.2.2 Uncertainty when Estimating Expected Lifetime

There are several issues to account when estimating the expected lifetime of highly reliable systems. Particularly when historical data is available. Otherwise, the uncertainty embedded in the estimation can be too large (Aven T. , Risk Analysis, 2012). The first issue is related to highly reliable systems with extensive lifetime. In years or decades. For those, the conditions through their lifetime are not the same. Factors such as human training, technological modifications, safety measures, quality of materials and more are very likely to have changed over time. Thus, to assume that the historical data shows exactly the previous behaviour of the same system is a bold assumption.

The second issue with regards of estimation of expected lifetime for highly reliable systems refers to biases in the elicitation of expert knowledge. Different biases can influence the outcome of the expert knowledge for the prior distribution, such as:

- Personal preferences due to extensive experience on the phenomenon.
- Economic reasons to favour unnecessary replacements at shorter times.
- Misleading reports of previous failures. Specifically, over reporting or under reporting.
- Hidden assumptions from experts regarding conditions of operations and degradation.
- Tendency to use the expected outcome approach without considering unexpected events.

In relation to the latter, unexpected sudden events are frequently dismissed as their probability of occurrence is believed to be extremely low. In the next lines a brief description of these events referred to as "black swans" is presented.

A black swan is a surprising event with extreme consequences that is not believed to occur due to its low probabilities (Taleb, 2010). A black swan can disrupt the performance of highly reliable systems regardless of their resilience. It is convenient to take a conservative approach when estimating expected lifetime of such systems. In BA this can be done through selection of parameters and prior distributions (Aven T. , Risk Analysis, 2012).

As mentioned in section 3.1, no replacement is considered for estimation of expected lifetime. Therefore, the uncertainty is focused on the estimation methods, their strengths and weaknesses. BA has an advantage as a tool to provide decision support, given the input of prior distributions into the assessment. The latter constitutes the main reason to perform BA to the study case presented in the next chapter and at the same time compare the results with a MLE of the same case.

Different factors contribute to uncertainty for the estimation of expected lifetime. These can be divided according to their origin (Taleb, 2010). Furthermore, they are present in most risk assessments and reliability studies. To identify which factors, contribute to uncertainty for the scope of this thesis see Figure 24.

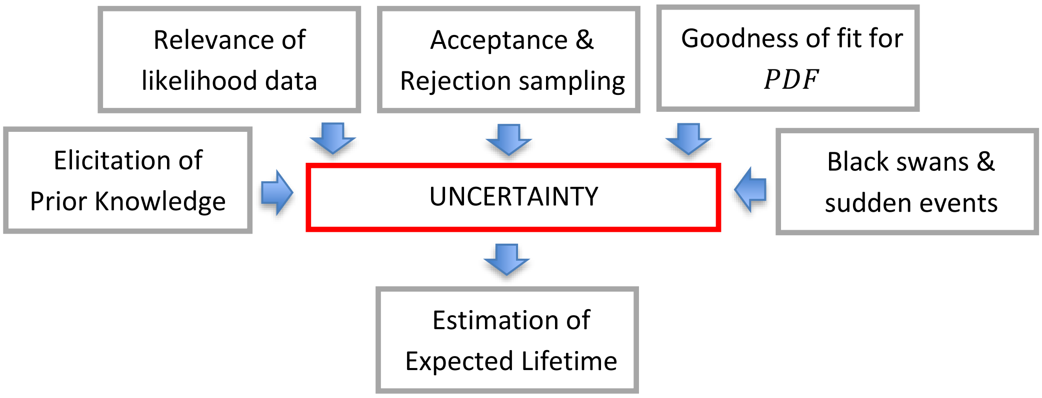


Figure 24. Contributing factors to uncertainty in estimation of expected lifetime with BA

Relevance of likelihood data has special importance among all other factors, because features like the availability, presence and type of censorship, size of sample and its degree of randomness. An estimation of expected lifetime with the wrong data set or no data at all means a higher degree of uncertainty in the results of $R(t)$, $F(t)$ and $MTTF$, regardless of how good the elicitation of prior knowledge and BA themselves may be.

This has a relevant significance when it comes to take decisions based on such results. Hence the need to reduce as much as possible all sources of uncertainties present in the analysis. Furthermore, whenever possible, uncertainty in all results must be recognized (Lindley, 2014).

3.2.3 Modelling using MatlabR2018

The software application used to analyse the data was Matlab version R2018. There are several reasons for selecting this program, among them are the following:

1. Easy to obtain, install and operate, compared to similar programs.
2. Author's previous experience with the program.
3. Availability to customize the results for better understanding and illustrative purposes of sensitivity analysis.
4. Simplicity for importing and exporting data sets to other applications such as Microsoft excel and Power BI.

The study cases to be modelled are systems considered highly reliable, as the time frame of performance is at least 5 years. The aim of modelling them is to represent in the most accurate way possible the behaviour of such systems, so an analyst can predict future outcomes with a high degree of credibility. For the latter, the used methodology was based on two methods to estimate the expected lifetime, MLE and BA.

The list of initial assumptions which define the scope of the four codes for MLE of each type of statistical distribution are:

- Initial data set is generated randomly with initial predefined values for parameters and number of elements, every time the code is run.
- Only right censoring of the data was considered. Therefore, data sets with censored data after a specific time are possible to compute. The time for censor is set by the user.
- The degree of confidence (deg) for the interval of the estimated parameter is set by the user in scalar number, not in terms of %. For example, to define a 95% of confidence interval the user must input a value of 0.05. See formula 35.
- The time for probability of survival is defined by the user, it must be a value that meets the condition $0 < t < 100$.
- The results are presented in a histogram and plots for PDF and $R(t)$ with fixed time scale of x axis.

The following, the list of initial assumptions which define the scope of the code for BA of all statistical distributions:

- Initial data for Prior is generated randomly based on the type of distribution selected by the user, with initial values for parameters every time the code is run.
- The number of elements (*size*) of the prior distribution can be randomly selected by the user, Nevertheless, it is recommended that it is equal to the number of the data set imported if possible. Hence, $nprior$ should be equal to $nlikelihood$.
- Prior parameters and data set are generated according to the predefined initial values for the parameter of the corresponding distribution.
- The preliminary result of prior predictive is presented as a histogram and plot with fixed time scale.
- For posterior distribution a real data set can be obtained from a specific file in excel which has to be called “sample.xlsx”, “samplecensored.xlsx” or “samplestudycase1.xlsx” according to the data set being analysed so the software will search in the “documents/MATLAB” folder and read it (Chapman, 2004).
- To obtain the posterior the real data set is combined with the prior predictive in a rejection sampling routine, where all the time values of the data set are given a random value of probability with a predefined maximum which is draw from a uniform distribution between 0 and 1.

- The outcomes of prior predictive, rejection sampling and posterior predictive are presented in independent plots for illustrative purposes.
- The time for probability of survival given the posterior distribution is defined by the user, it must be a value that meets the condition $0 < t < 100$.
- Sensitivity analysis of the outcomes can be done by changing the values of initial parameter of the Prior predictive, time for probability of survival ($tsurvival$), or the number of elements randomly generated in the predictive prior ($nprior$), likelihood, ($nlikelihood$)and predictive posterior ($nposterior$).

Furthermore, for both methods of estimation the $MTTF$ values are shown in two formats based on a year: The scalar format and the date format. The scalar format indicates a fraction of a year(s). For example, an $MTTF = 0.5$ represent half a year until the failure occurs. Whilst the date format indicates how many years, months, days, and hours are remaining for the occurrence of a failure from the exact moment the analysis is done. For example: $MTTF = 0.5$ represents 0 years, 6 months, 15 days and 12 hours for the failure to take place.

In addition, all samples data randomly generated and obtained from the study cases are assumed to be reporting only failure and censored times. The type and magnitude of failure is not part of the scope, i.e. all the input values are treated equally, if they represent a failure that will prevent the system itself to operate. The latter indicates that for the software, there is no hierarchy of failure according to their origin, type or consequence.

Another relevant issue when modelling with Matlab is the acceptance-rejection prebuilt function. This has a numerical value which allow the user to adjust the criteria to accept or reject every sample value randomly generated. For all the codes this value is fixed at two. In other words, for an *ith* iteration to produce an accepted value for the predictive posterior, it has to be equal or larger than twice the reference value of the sample distribution.

CHAPTER 4

4.1 Server-database for Commercial Use Study Case

By presenting a brief background of the company which provided the data, one can understand its nature and the reason for the reliability assessment, moreover the scope for estimation of expected lifetime of its main equipment.

AKRA Investments c.a. is a logistics company which supplies over 300 clients in 20 different states in Venezuela. The main target is retail companies of high consumption products. It provides a platform for inventory, transport, budgeting, ordering and customer service through an online system which is the study case of concern.

The study case refers to a combination of a main server which handles all cloud, storage and service operations. The unit works in a semi restricted environment where only authorised users can access it and the database to perform different actions such as maintenance, update, repairs and modifications.

The server has several features that allow administrators and users to perform daily operations for the company. Among them are the following:

- Maintain backup of all inputs and modifications to the main database
- Channel all incoming and outgoing files to cloud platforms and apps such as e-mails, data transfer between user, transfer of data between company employees and customers
- Frame remote access to the company's computers and terminals
- Prevent unauthorised access from external and malicious entities

These are critical operations that require a reliable server-database for the company. Furthermore, the ability to estimate failure time is critical to avoid negative consequences.

A log for all events occurred is available in the database's backup files. This allows to filter the data and elicit failure events from all others. Different types of failures were encountered. Nevertheless, from the reliability assessment they are all going to be treated as common failure. Where the time of occurrence is the main concern for this study case. These types of failure and their corresponding code and possible causes are listed in Table 18, based on the log data and company's criteria.

Table 18. Types of failure events for Server-Database study case

Type of Failure	Code	Criticality	Possible Causes
Partial loss of data	PLD	4	Virus or Bug
			Electrical
			Other
Total loss of data	TLD	5	Virus or Bug
			Electrical
Virus or Bug	VBF	4	Unauthorised access
			Firewall down
Electrical	ELF	3	Power grid blackout
			Extreme voltage fluctuations
Update	UPD	2	Firewall down
			No Internet connection
Other	OTH	2	Human error
			No internet connection
			Physical damage to server

Table 18, shows that, one failure can be the cause of others. The criteria for the criticality of failure is based on a subjective scale from 1 to 5, where 5 represents the most critical type of failure in terms of its consequences for the company.

In principle a failure can be ranked for its derived consequences. In other words, on how it impacts the organization when this type of event occurs (Aven T. , Foundations of Risk Analysis, 2012). For AKRA Investments c.a. there are four main variables to rank types of failure events.

1. Amount of lost data
2. Server's down time
3. Delayed time for supplying customers
4. Difficulty and resources needed to repair the failure

4.1.1 Modelling with MLE

There are two main actions required for the estimation of expected lifetime of the Server-database with MLE. The first one is to filter the entire data base by some criterion that allows to retrieve only failure events. The criterion is based on failure events described on Table 18. Second, once the filtered data is available, one must convert its entries from date format and create an .*XLSX* file which represent the likelihood data for MLE.

The historical record for daily operations and failure events was provided by the company in a $.BAK$ file. Therefore, to perform both MLE and BA in Matlab R2018, it was modified to a $.XLSX$ format which is the one supported by the written codes. Figure 25 presents AKRA Investments Server-database failure entries in $.XLSX$ format.

As previously mentioned, all recorded events have a standard date format (dd/mm/yyyy) including hours, minutes and seconds of occurrence. Hence, all failure times where obtained by converting those values to a decimal format. This allows to interpret the data and results in terms of years instead of more large values such as hours. Take rows 1 and 2 of Figure 25 for example, it shows that the 38th failure occurred on the January the 8th of 2019 at 18 hours, 12 minutes and 00 sec. Then 0.2708 years or 2.372 hours after failure 39th took place.

Table 19 shows the modified log for failure times from the original $XLSX$ format.

Table 19. Modified log with failure times of Server-database study case

t time values in years								
0.1322	0.4771	0.4797	0.9233	1.0744	1.7699	1.9969	2.6798	3.3370
3.6158	4.0630	4.1664	4.2030	4.2949	4.5668	4.7171	4.9087	5.2217
5.4426	5.7553	6.3116	6.4590	6.8103	6.8500	7.0774	7.5751	7.6908
7.9202	8.1342	8.4548	9.0708	9.3316	9.7728	10.2183	10.5120	10.9838
11.2546	-	-	-	-	-	-	-	-
The Server-database's log is ongoing at the time this report was finished.								

In order to reduce the uncertainty of the results of MLE from Table 19, the time values in years are included with four decimals, the variable $nlikelihood$ is fixed at $1x10^6$, the confidence intervals for estimated parameters is fixed to 95% and both histograms and Plots are displayed to compare with numerical values.

Company:		AKRA Investments c.a.		Report date:	31/05/2019
Equipment:		HP Pavilion 570-p038 PC		Initial record:	07/01/2009
RECORD OF FAILURE EVENTS FOR SERVER-DATABASE					
#	**Failure Event**	**Failure Date (dd/mm/yyyy)**	**Failure Time (hh:mm:ss)**	**Hours Between Failure**	**Years Between Failure**
1	OTH	16/01/2008	10:16:20	1158	0.1322
2	UPD	04/03/2008	16:18:32	3021	0.3449
3	ELF	08/07/2008	13:20:53	23	0.0026
4	UPD	09/07/2008	13:01:02	3886	0.4436
5	VBF	18/12/2008	11:24:17	1324	0.1511
6	PLD	11/02/2009	16:03:04	6092	0.6954
7	ELF	23/10/2009	12:18:22	1989	0.2271
8	OTH	14/01/2010	09:32:06	5982	0.6829
9	PLD	20/09/2010	16:03:04	1579	0.1803
10	OTH	25/11/2010	11:52:10	4178	0.4769
11	ELF	18/05/2011	14:04:05	2442	0.2788
12	VBF	28/08/2011	09:03:39	3918	0.4473
13	VBF	07/02/2012	15:18:22	906	0.1034
14	PLD	16/03/2012	10:11:09	320	0.0365
15	OTH	29/03/2012	18:29:13	805	0.0919
16	VBF	02/05/2012	08:02:07	2382	0.2719
17	OTH	09/08/2012	14:14:12	1317	0.1503
18	ELF	03/10/2012	11:52:50	1678	0.1916
19	OTH	14/12/2012	10:29:43	2742	0.3130
20	TLD	05/04/2013	17:05:32	1935	0.2209
21	PLD	25/06/2013	08:47:00	2739	0.3127
22	VBF	17/10/2013	12:04:10	4874	0.5564
23	OTH	08/05/2014	14:45:59	1291	0.1474
24	OTH	01/07/2014	10:23:37	3077	0.3513
25	ELF	06/11/2014	16:02:20	348	0.0397
26	OTH	21/11/2014	04:23:02	1992	0.2274
27	UPD	12/02/2015	17:16:58	4360	0.4977
28	PLD	13/08/2015	09:45:40	1013	0.1156
29	OTH	24/09/2015	14:51:03	2010	0.2295
30	OTH	17/12/2015	09:04:53	1875	0.2140
31	VBF	04/03/2016	13:02:04	2808	0.3205
32	PLD	29/06/2016	13:10:11	5396	0.6160
33	OTH	09/02/2017	10:07:56	2285	0.2608
34	ELF	15/05/2017	15:25:02	3865	0.4412
35	VBF	23/10/2017	16:49:04	3902	0.4454
36	UPD	04/04/2017	07:03:17	2573	0.2937
37	PLD	20/07/2018	13:01:19	4133	0.4718
38	VBF	08/01/2019	18:12:00	2372	0.2708
39	ELF	17/04/2019	15:03:18	*	*

Figure 25. AKRA Investments Server-database failure entries in .XLSX format.

The best fitted PDF to the likelihood data from Table 19 according to Kolmogorov-Smirnov ranking is Weibull. Therefore, the estimated $R(t)$, $F(t)$ and $MTTF$ are based on $\hat{\alpha}$ and $\hat{\beta}$ the corresponding scale and shape parameters. These estimations represent the default values for comparison between MLE and BA. In other words, how does the estimated values are influenced by the weight of likelihood data in comparison to prior expert knowledge. MLE results are presented on Table 20.

Table 20. Summary Table for MLE of Weibull random data set using Matlab.

Data	Parameters & Estimators				Confidence Intervals			T	Expected Lifetime		
n	α	β	$Muhat$ $\hat{\alpha}$	$\hat{\beta}$	deg %	$Muci$		$tsur$	$R(t)$	$F(t)$	$MTTF$
Variable t, α and β and fixed deg and n											
38	1.3	0.7	**6.136**	**1.661**	5	[4.307, 8.741]	[1.028, 2.683]	1	**0.952**	**0.047**	**5.484**
	0.2	1.1						2	**0.856**	**0.143**	
	0.9	0.2						3	**0.737**	**0.262**	
	1.1	1.4						4	**0.611**	**0.388**	
	0.8	0.7						5	**0.490**	**0.509**	

Table 20 shows an estimated $MTTF$ of 5.484 years given $\hat{\alpha}$ and $\hat{\beta}$ equal to 6.136 and 1.661 respectively for survival times from one to five years. Moreover, based on the likelihood data $R(t)$ and $F(t)$ values are decreasing as $tsurvival$ increase. For example, the reliability of the system is approximately 95% at after one year and 50% after five years of operations.

The histograms for each $tsurvival$ are presented in Figure 26, 27, 28, 29 and 30. It is visible that all of them have a typical Weibull form with small variations between each time bid. Moreover, the maximum number of events is around 3.700 and the extreme values from the tail of the distribution are not beyond 20 years. This is a consequence of the randomly generated time values for iterations of likelihood distribution.

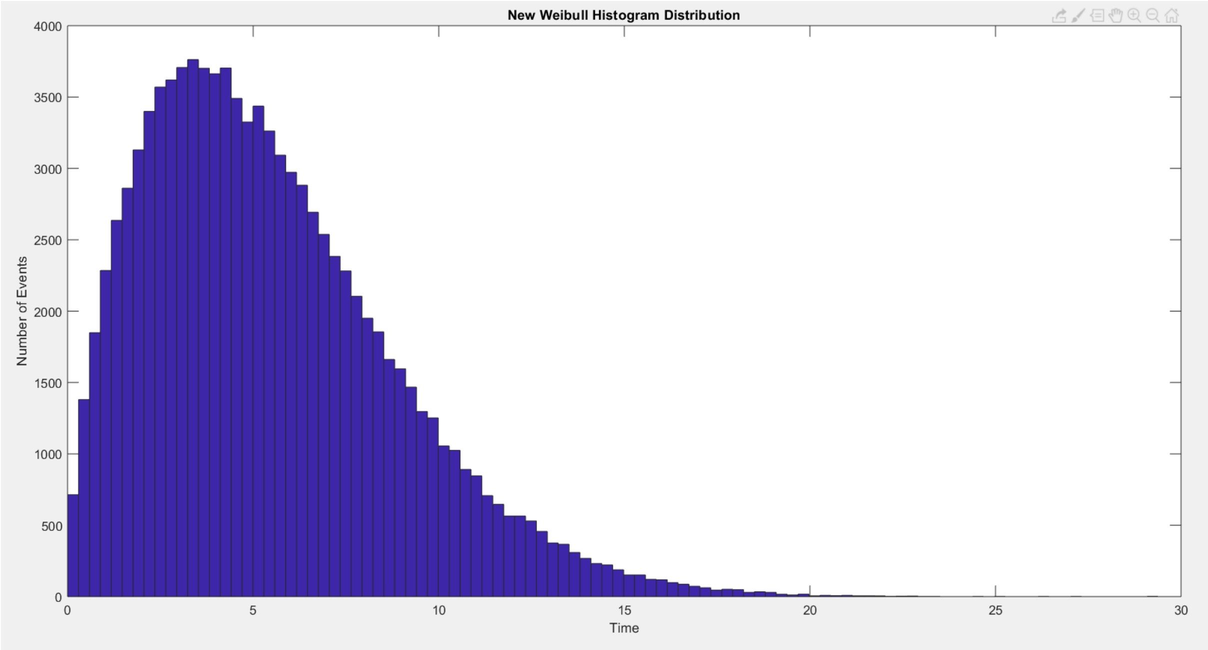


Figure 26. MLE Histogram for tsurvival of 1 year of AKRA Investments Server-database

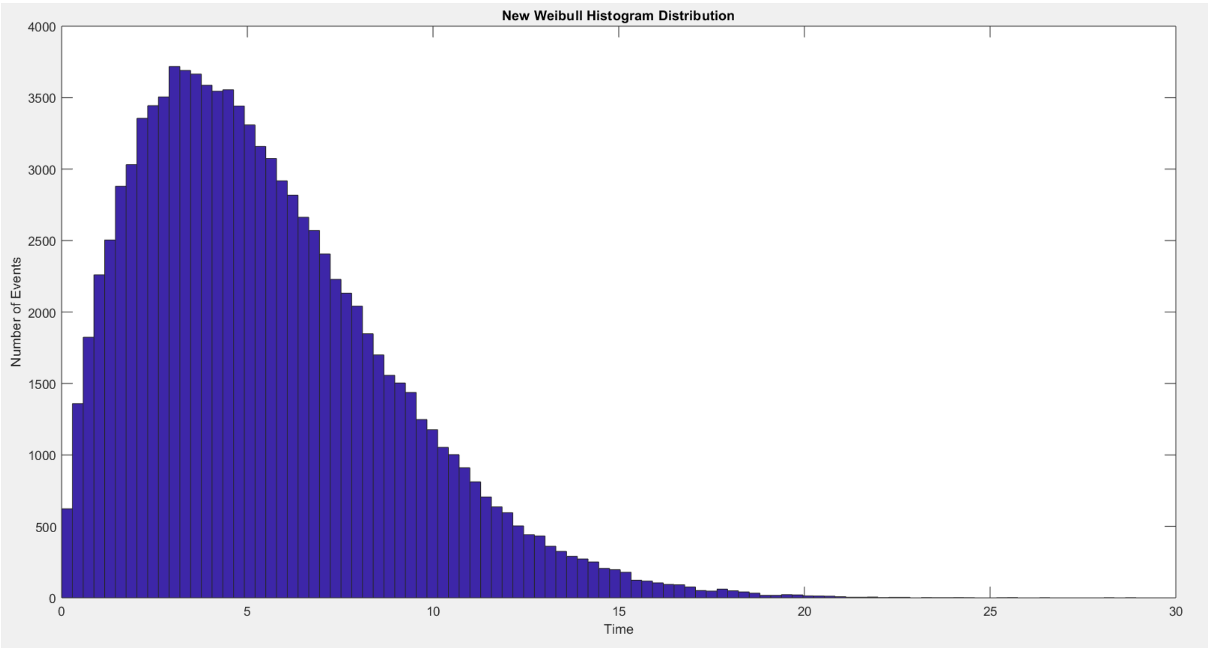


Figure 27. MLE Histogram for tsurvival of 2 years of AKRA Investments Server-database

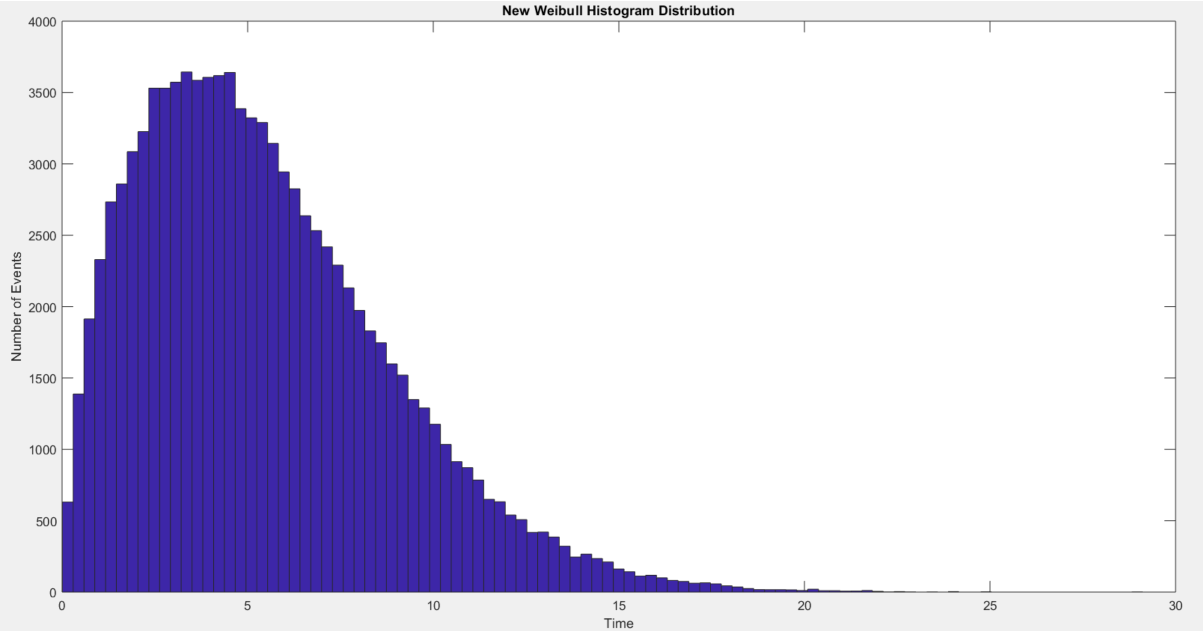


Figure 28. MLE Histogram for tsurvival of 3 years of AKRA Investments Server-database

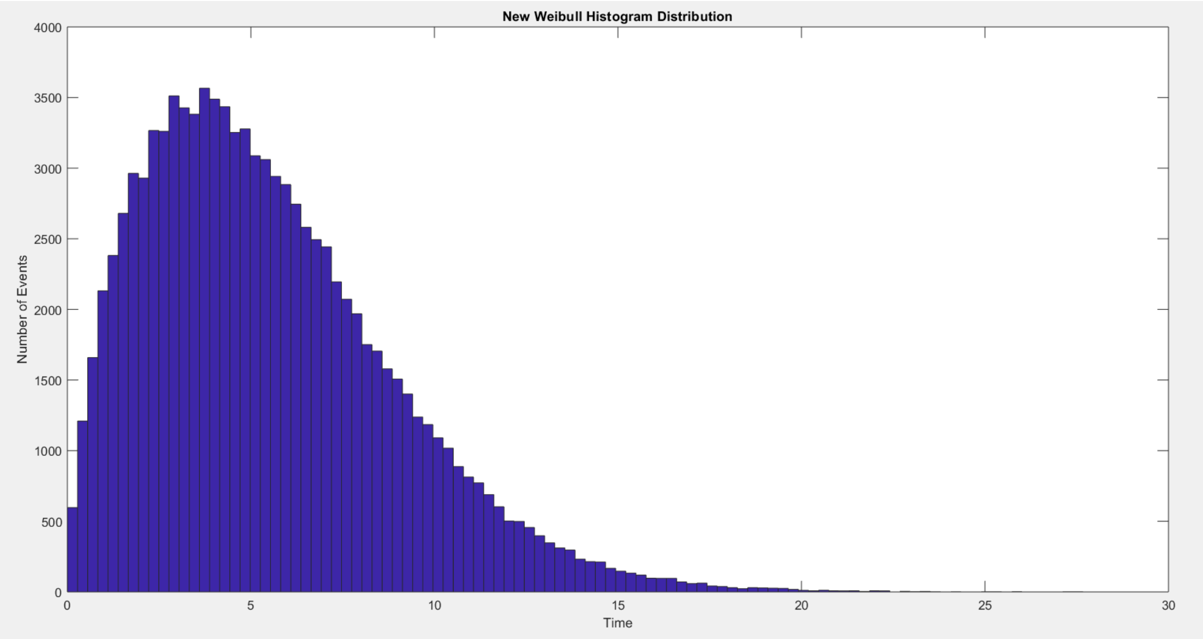


Figure 29. MLE Histogram for tsurvival of 4 years of AKRA Investments Server-database

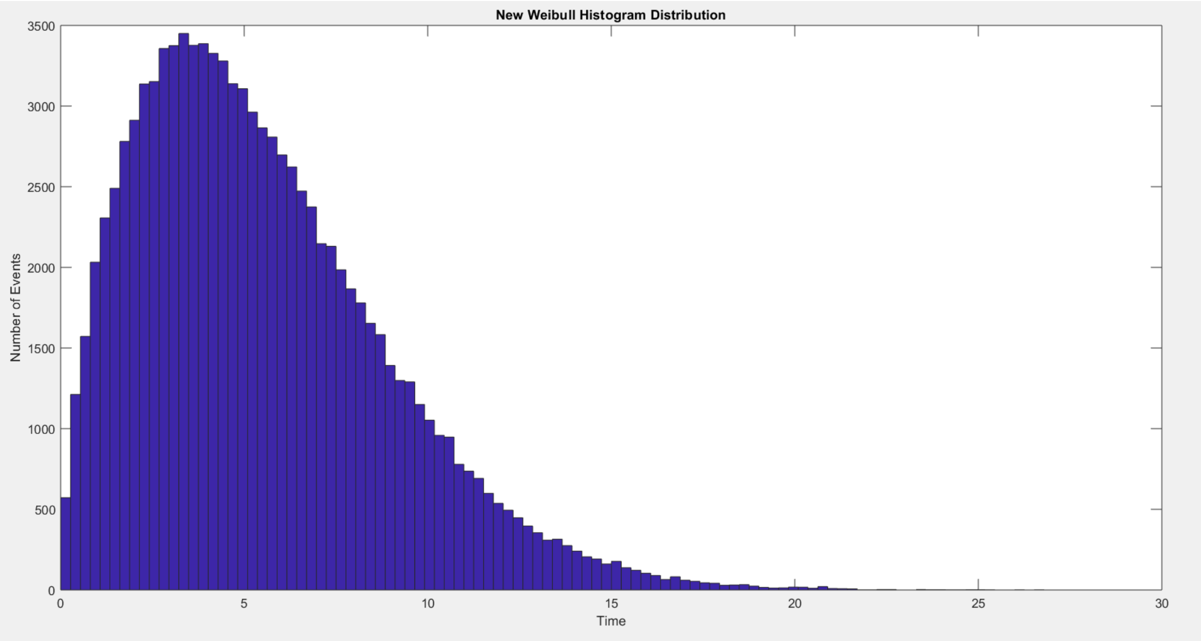


Figure 30. MLE Histogram for tsurvival of 5 years of AKRA Investments Server-database

Notice that all 5 histograms of MLE are very similar in shape and scale, both x and y axis are the same in most cases except for figure which has a max value of event for the y axis is equal to 3.500 instead of 4.000 as the previous histograms. Furthermore, the only real difference between these five histograms is the edges from bid to bid, which vary slightly near the mean values and the tail of the distributions. As mentioned before, the reason for this is the randomly generated time values for iterations of likelihood distribution. Hence, it seems logical to assume that the larger $nlikelihood$ is, the smoother these difference between time bid would be.

The plot for the PDF based on the MLE given the likelihood data of the study case is presented in Figure 31.

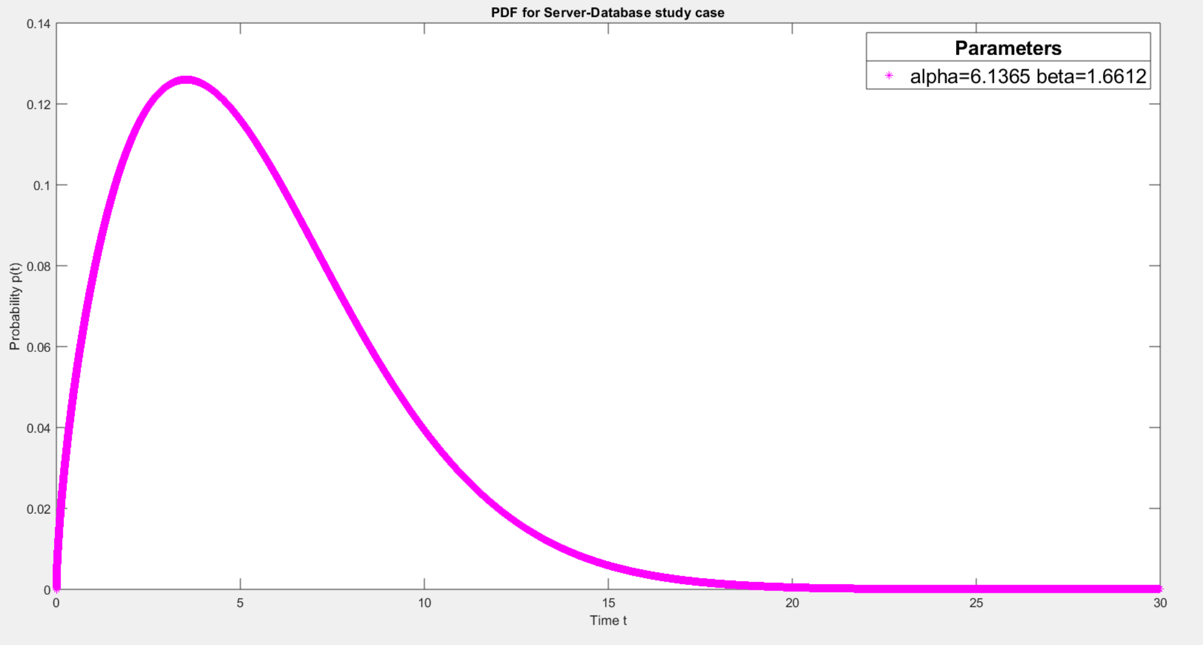


Figure 31. PDF based on MLE given the likelihood data of the study case

The plot for $R(t)$ based on the MLE given the likelihood data of the study case is presented in Figure 32.

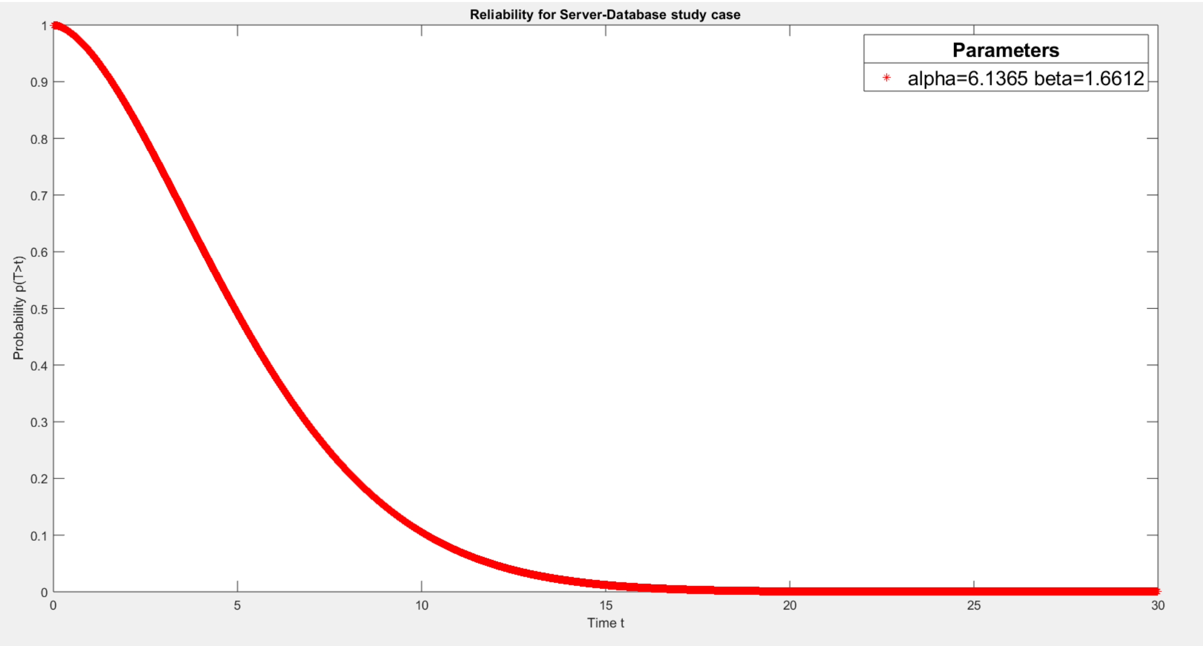


Figure 32. Reliability based on MLE given the likelihood data of the study case

Figure 32 confirms the results of Table 20 for reliability for different values of $tsurvival$ from 1 to 5 years. For example, at the 1st year $R(t)$ is close to 95% and at the 5th year $R(t)$ is approximately 50%. As mentioned before these results are based only on the estimated parameter of a Weibull distribution given the likelihood data of the study case. For further assessments BA is presented in the next section.

4.1.2 Modelling with BA

In addition to the likelihood data of failures times provided in Table 19, all four PDFs will be used as predictive prior distributions for BA, based on the same prior parameters used for estimations performed according to Table 5 for the random and censored data sets. A sensitivity analysis for the most suitable result is presented with three variations of these parameters to account for the three most common approaches: cautionary, average and risky.

The parameters for each approach are obtained by simulating disagreement among experts on common prior parameters with different weight for each one. Following the procedure shown on appendix A6.

For an exponential predictive prior distribution, the results of BA are presented for $tsurvival$ equal to 1 and 5 years in Table 21. Furthermore, Figures 33 and 34 show the predictive posterior PDF and $R(t)$ plots respectively. In addition, Figures 35 and 36 present the BA results for both values of $tsurvival$ as seen in Matlab's command window.

Table 21 confirms that reliability values decrease for larger times. It also shows that $MTTF$ values are equal to 5.484 years for all trials of MLE. Furthermore, for all trials of BA it is approximately 0.94 years. There is a very small difference between the two BA trials. In addition, $R(t)$ is considerable smaller for any trial of BA compared with its corresponding MLE trial. Take $tsurvival$ equal to 1 year for example, the $R(t)$ for MLE is 95%, whilst for BA is 20%. This is equivalent to less than a quarter. The opposite occurs with $F(t)$ as it is a complementary function.

Table 21. Results of MLE and BA with Exponential prior predictive for study case

Method of Estimation		**Number of Iterations**			**T**	**Expected Lifetime**		
$\boldsymbol{MLE}$		$\boldsymbol{nprio}$	$\boldsymbol{nlikel}$	$\boldsymbol{nposte}$	$\boldsymbol{tsur}$	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	$\boldsymbol{MTTF}$
Fixed n, $nprior$, $nlikelihood$ vs. variable and $tsurvival$								
Prior	**Likelihood**	1x10^6	1x10^6	1x10^6	1	**0.952**	**0.047**	**5.484**
None	Weibull				5	**0.490**	**0.509**	**5.484**
$\boldsymbol{BA}$		$\boldsymbol{nprio}$	$\boldsymbol{nlikel}$	$\boldsymbol{nposte}$	$\boldsymbol{tsur}$	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	$\boldsymbol{MTTF}$
Prior	**Likelihood**	1x10^6	1x10^6	1x10^6	1	**0.2061**	**0.7938**	**0.9409**
Exponential	Weibull				5	**0.0707**	**0.9292**	**0.9360**

Figure 33 shows the predictive posterior PDF in a fixed time scale of 20 years. Notice the period from 0 to 2 years where the maximum probability of failure value is aproximately 0.028 to occur in half a year of operations.

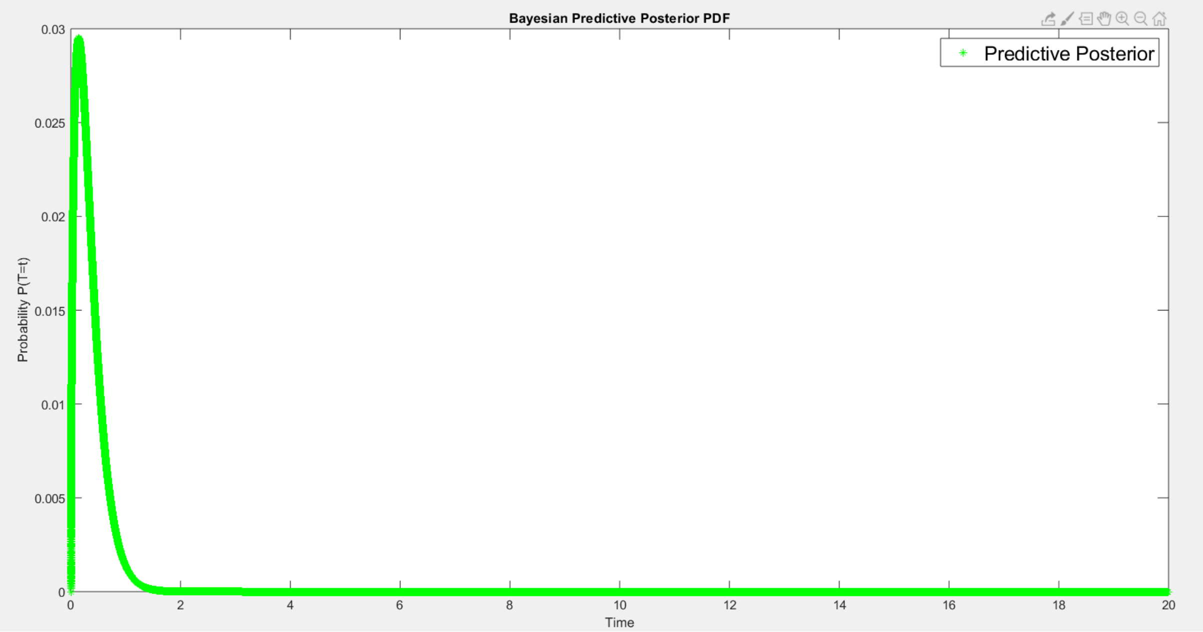


Figure 33. PDF based on BA given an Exponential predictive prior and the likelihood data of the study case

Figure 34 shows $R(t)$ of a BA for the study case assuming an Exponential prior with the default failure rate. In addition, it confirms the results obtained in Tabe 21, Figures 35 and 36. Moreover, $R(t)$ after 1 and 5 years of operations is equal to 20% and 7% respectively.

Figure 35 and 36 confirm that $MTTF$ for the study case assuming an Exponential prior with the default failure rate is between 11 months 8 days and 18 hours and 11 months 6 days and 23 hours. This slight difference is a consequence of the difference between 0.9409 and 0.9360 years.

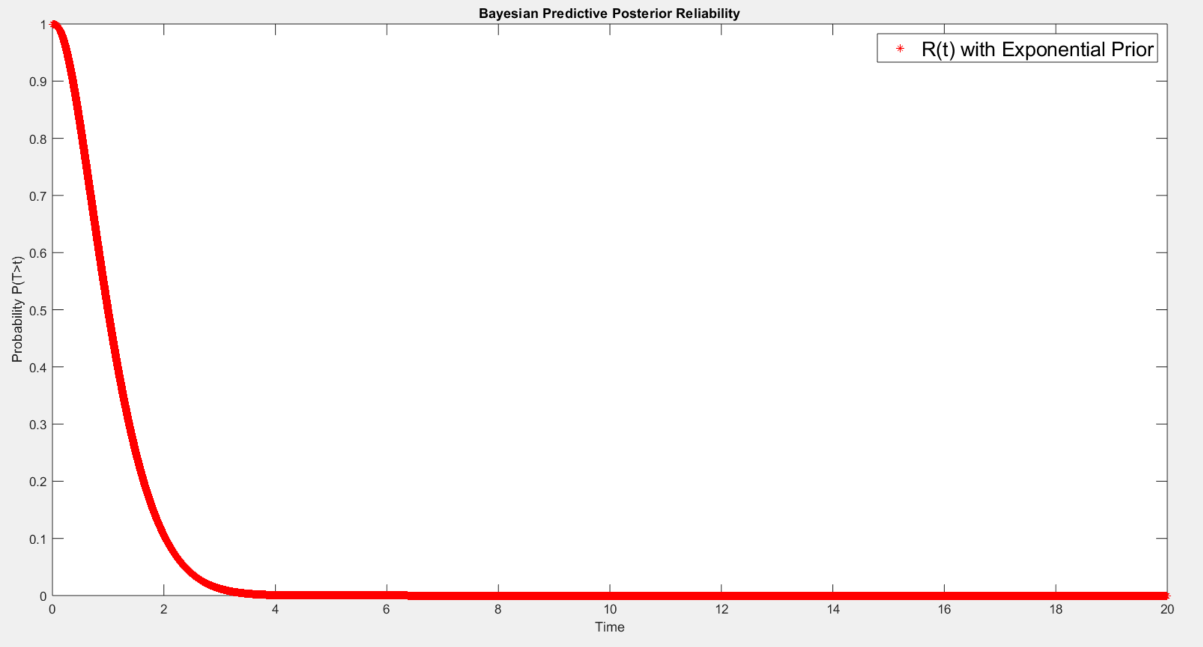


Figure 34. Reliability based on BA given an Exponential predictive prior and the likelihood data of the study case

Prior	Likelihood	Posterior
'Exponential'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	1	0.20614	0.79386	0.94098

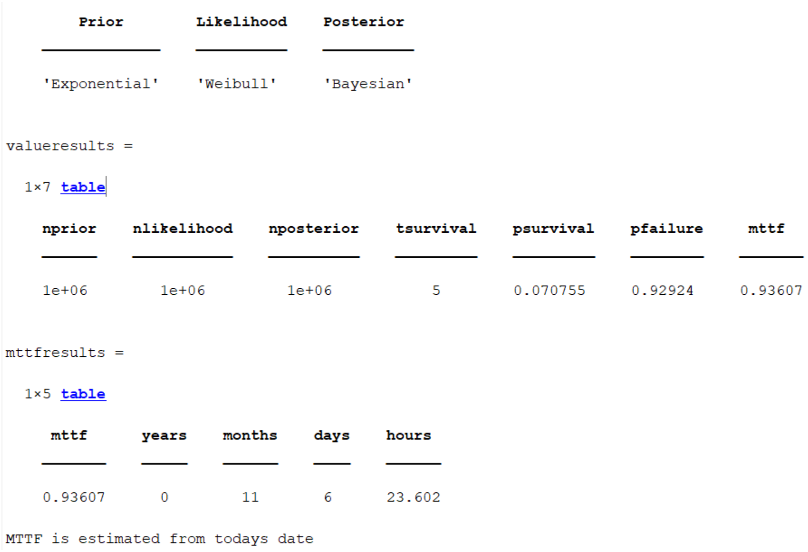
mttfresults =

1×5 table

mttf	years	months	days	hours
0.94098	0	11	8	18.093

MTTF is estimated from todays date

Figure 35. BA results for Exponential predictive prior and $tsurvival = 1$ *year as seen in Matlab's command window*



```
     Prior         Likelihood    Posterior
  _____________    __________    __________

  'Exponential'    'Weibull'     'Bayesian'

valueresults =

  1×7 table

    nprior    nlikelihood    nposterior    tsurvival    psurvival    pfailure     mttf
    ______    ___________    __________    _________    _________    ________    _______

    1e+06        1e+06         1e+06           5        0.070755     0.92924     0.93607

mttfresults =

  1×5 table

     mttf      years    months    days    hours
    _______    _____    ______    ____    ______

    0.93607      0        11       6      23.602

MTTF is estimated from todays date
```

Figure 36. BA results for Exponential predictive prior and $tsurvival = 5$ *years as seen in Matlab's command window*

For a Weibull predictive prior distribution, the results of BA are presented for $tsurvival$ equal to 1 and 5 years in Table 22. Furthermore, Figures 37 and 38 show the predictive posterior PDF and $R(t)$ plots respectively. In addition, Figures 39 and 40 present the BA results for both values of $tsurvival$ as seen in Matlab's command window.

Table 22. Results of MLE and BA with Weibull prior predictive for study case

Method of Estimation		**Number of Iterations**			**T**	**Expected Lifetime**		
MLE		***npri***	***nlikel***	***nposte***	***tsur***	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	***MTTF***
Fixed n, $nprior$, $nlikelihood$ vs. variable and $tsurvival$								
Prior	**Likelihood**	$1x10^6$	$1x10^6$	$1x10^6$	1	**0.952**	**0.047**	**5.484**
None	Weibull				5	**0.490**	**0.509**	**5.484**
BA		***npri***	***nlikel***	***nposte***	***tsur***	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	***MTTF***
Prior	**Likelihood**	$1x10^6$	$1x10^6$	$1x10^6$	1	**0.0529**	**0.9471**	**0.2174**
Weibull	Weibull				5	**0.0142**	**0.9857**	**0.2175**

Table 22 confirms that reliability values decrease for larger times and shows that $MTTF$ values are equal to 5.484 years for all trials of MLE. While for BA it is approximately 0.21 years. There is a very small difference between the two BA trials. Also, $R(t)$ is considerable smaller for any trial of BA compared with its corresponding MLE trial. Take $tsurvival$ equal to 1 year for example, the $R(t)$ for MLE is 95%, whilst for BA is 5%. Here the influence of prior parameters is notorious. The opposite occurs with $F(t)$ values as it is a complementary function.

Figure 37 shows the predictive posterior *PDF* in a fixed time scale of 1 year. This is due to the extremely low time for to reach the maximum probability value. Notice the period from 0 to 0.1 years where the maximum probability of failure value is aproximately 0.042 to occur in half a year of operations.

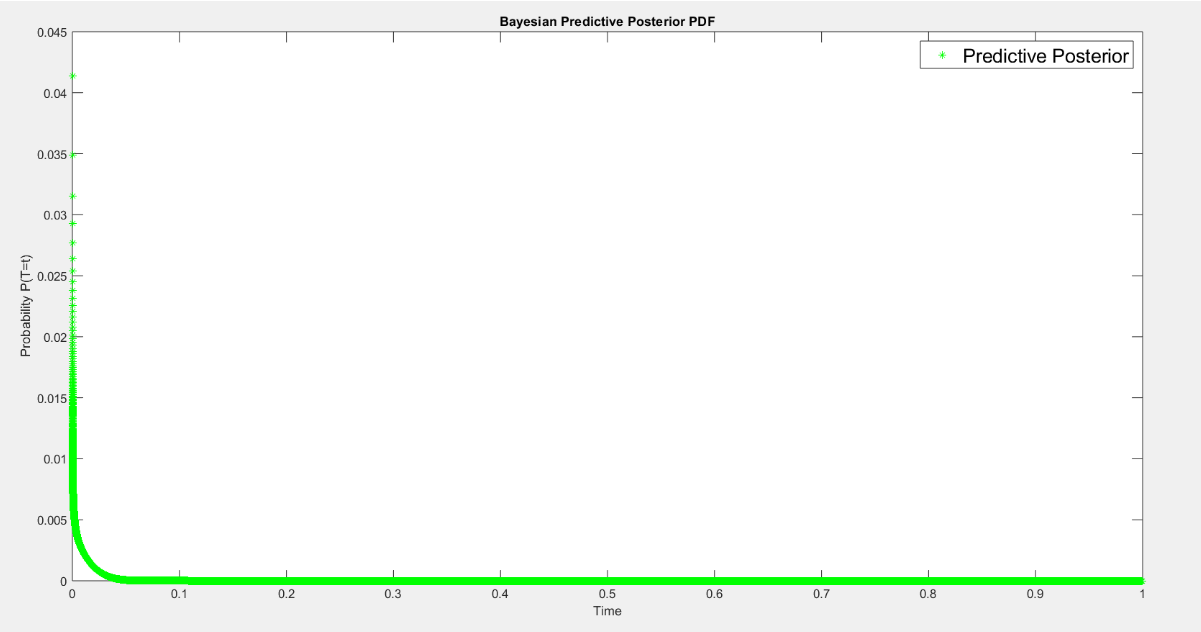


Figure 37. PDF based on BA given a Weibull predictive prior and the likelihood data of the study case

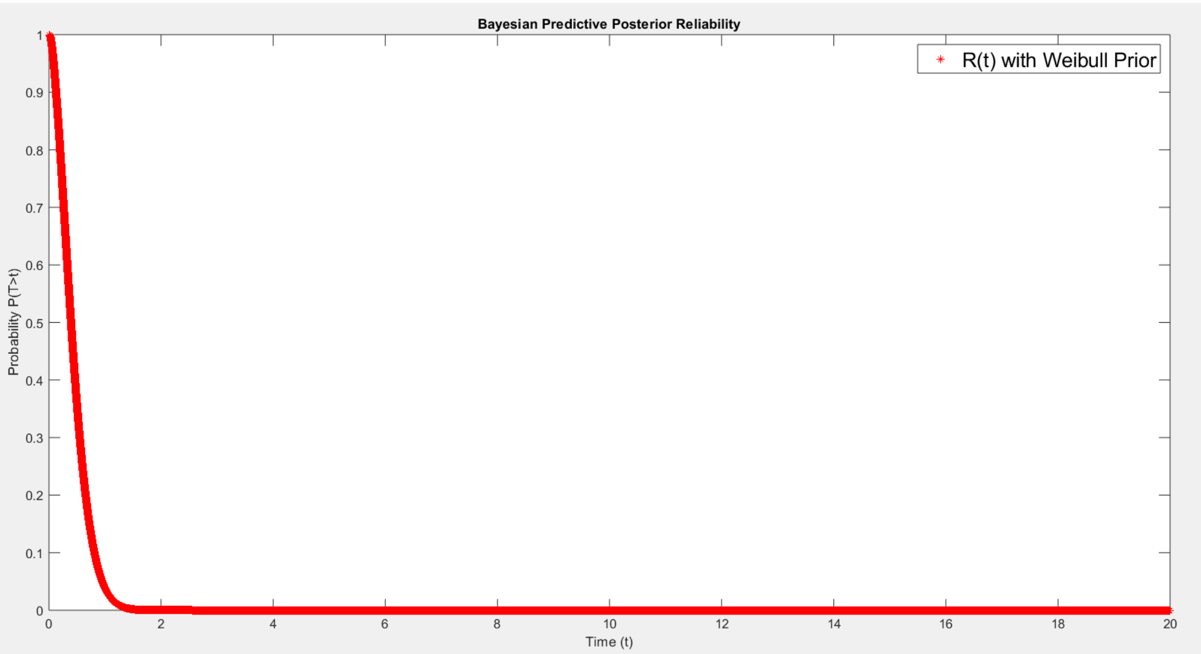


Figure 38. Reliability based on BA given a Weibull predictive prior and the likelihood data of the study case

Figure 38 shows $R(t)$ of a BA for the study case assuming an Weibull prior with the default failure rate. In addition, it confirms the results obtained in Table 22, Figures 39 and 40. Moreover, $R(t)$ after 1 and 5 years of operations is equal to 5% and 1,5 % respectively.

Figures 39 and 40 confirm that $MTTF$ for the study case assuming a Weibull prior with the default failure rate is between 2 months 18 days and 4 hours and 2 months 18 days and 7 hours. This slight difference is a consequence of the difference between 0.2172 and 0.2175 years.

Prior	Likelihood	Posterior
'Weibull'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	1	0.052687	0.94731	0.21722

mttfresults =

1×5 table

mttf	years	months	days	hours
0.21722	0	2	18	4.7391

Figure 39. BA results for Weibull predictive prior and $tsurvival = 1$ $year$ as seen in Matlab's command window

Prior	Likelihood	Posterior
'Weibull'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	5	0.014288	0.98571	0.21751

mttfresults =

1×5 table

mttf	years	months	days	hours
0.21751	0	2	18	7.2637

MTTF is estimated from todays date

Figure 40. BA results for Weibull predictive prior and $tsurvival = 5$ $years$ as seen in Matlab's command window

For a Lognormal predictive prior distribution, the results of BA are also presented for $tsurvival$ equal to 1 and 5 years in Table 23. Furthermore, Figures 41 and 42 show the corresponding predictive posterior PDF and $R(t)$ plots. In addition, Figures 43 and 44 present the results of BA for both values of $tsurvival$ as seen in Matlab's command window.

Table 23. Results of MLE and BA with Lognormal prior predictive for study case

Method of Estimation		**Number of Iterations**			**T**	**Expected Lifetime**		
$\boldsymbol{MLE}$		$\boldsymbol{nprio}$	$\boldsymbol{nlikel}$	$\boldsymbol{nposte}$	$\boldsymbol{tsur}$	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	$\boldsymbol{MTTF}$
Fixed n, $nprior$, $nlikelihood$ vs. variable and $tsurvival$								
Prior	**Likelihood**	1x10^6	1x10^6	1x10^6	1	**0.952**	**0.047**	**5.484**
None	Weibull				5	**0.490**	**0.509**	**5.484**
$\boldsymbol{BA}$		$\boldsymbol{nprio}$	$\boldsymbol{nlikel}$	$\boldsymbol{nposte}$	$\boldsymbol{tsur}$	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	$\boldsymbol{MTTF}$
Prior	**Likelihood**	1x10^6	1x10^6	1x10^6	1	**0.3907**	**0.6092**	**1.802**
Lognormal	Weibull				5	**0.1374**	**0.8625**	**1.797**

Table 23 also confirms that reliability values decrease for larger times and unreliability values increase. Moreover, it shows that $MTTF$ values are equal to 5.484 years for all trials of MLE, and approximately 1.8 years for BA. There is a very small difference between the two BA trials. In addition, $R(t)$ is smaller for any trial of BA compared with its corresponding MLE value. For $tsurvival$ equal to 1 year, the $R(t)$ for MLE is 95%, whilst for BA is 39%. Here the influence of prior parameters is notorious. The opposite occurs with $F(t)$ values as it is a complementary function.

Figure 41 shows the predictive posterior PDF in a fixed time scale of 1 year. This is due to the extremely low time for to reach the maximum probability value. Notice the period from 0 to 0.1 years where the maximum probability of failure value is aproximately 0.025 to occur in less than half a year of operations.

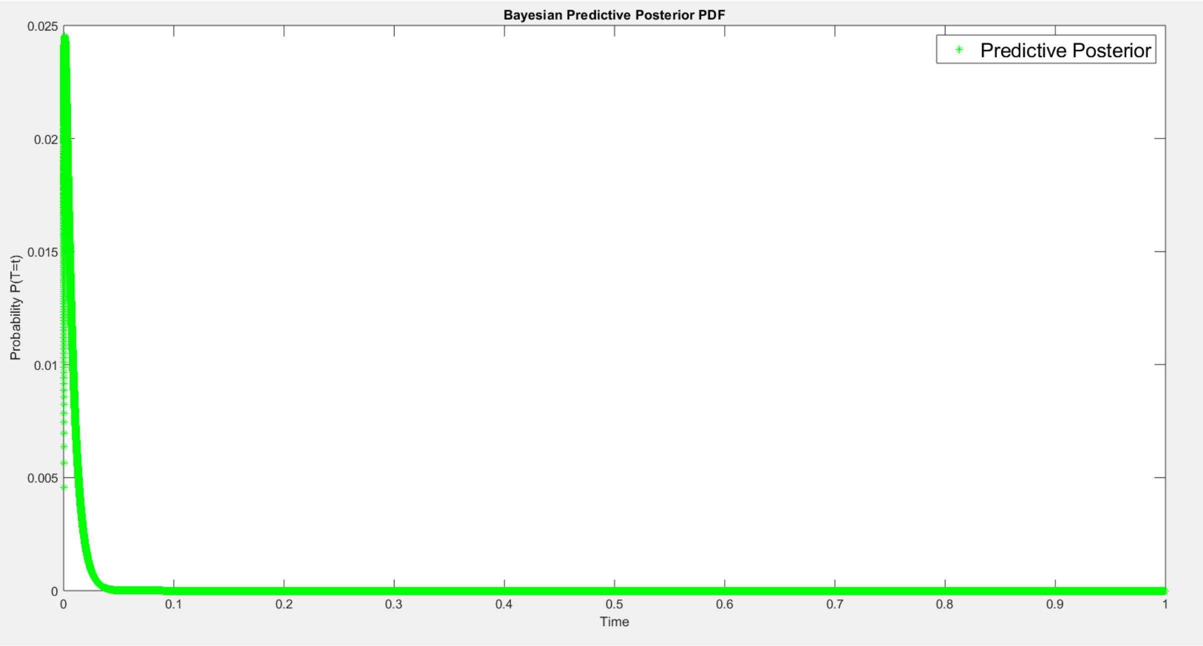


Figure 41. PDF based on BA given a Lognormal predictive prior and the likelihood data of the study case

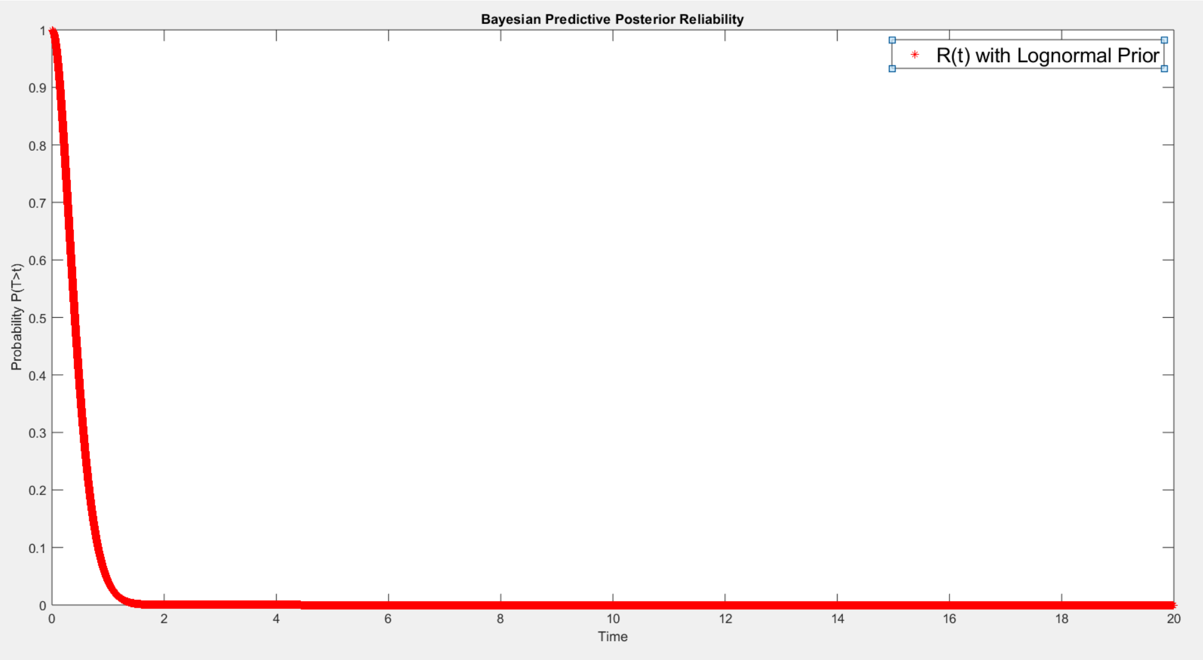


Figure 42. Reliability based on BA given a Lognormal predictive prior and the likelihood data of the study case

Figure 42 shows $R(t)$ of a BA for the study case assuming an Lognormal prior with the default failure rate. In addition, it confirms the results obtained in Table 23. Figures 43 and 44. Moreover, $R(t)$ after 1 and 5 years of operations is equal to 39% and 13 % respectively.

Prior	Likelihood	Posterior
'Lognormal'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	1	0.39071	0.60929	1.802

mttfresults =

1×5 table

mttf	years	months	days	hours
1.802	1	9	18	17.067

MTTF is estimated from todays date

Figure 43. BA results for Lognormal predictive prior and $tsurvival = 1$ *year as seen in Matlab's command window*

Figure 43 and 44 confirm that $MTTF$ for the study case assuming a Lognormal prior with the default failure rate is between 1 year 9 months 18 days and 17 hours and 1 year 9 months 16 days and 23 hours. This slight difference is a consequence of the difference between 1.802 and 1.797 years.

Prior	Likelihood	Posterior
'Lognormal'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	5	0.13745	0.86255	1.7972

mttfresults =

1×5 table

mttf	years	months	days	hours
1.7972	1	9	16	23.795

MTTF is estimated from todays date

Figure 44. BA results for Lognormal predictive prior and $tsurvival = 5$ *years as seen in Matlab's command window*

Finally, for a Gamma predictive prior distribution, the results of BA are presented for $tsurvival$ equal to 1 and 5 years in Table 24. Furthermore, Figures 45 and 46 show the corresponding predictive posterior PDF and $R(t)$ plots. In addition, Figures 47 and 48 display the results of BA for both values of $tsurvival$ as seen in Matlab's command window.

Table 24. Results of MLE and BA with Gamma prior predictive for study case

Method of Estimation		**Number of Iterations**			**T**	**Expected Lifetime**		
MLE		***nprio***	***nlikel***	***nposte***	***tsur***	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	***MTTF***
Fixed n, $nprior$, $nlikelihood$ vs. variable and $tsurvival$								
Prior	**Likelihood**	$1x10^6$	$1x10^6$	$1x10^6$	1	**0.952**	**0.047**	**5.484**
None	Weibull				5	**0.490**	**0.509**	**5.484**
BA		***nprio***	***nlikel***	***nposte***	***tsur***	$\boldsymbol{R(t)}$	$\boldsymbol{F(t)}$	***MTTF***
Prior	**Likelihood**	$1x10^6$	$1x10^6$	$1x10^6$	1	**0.2361**	**0.7638**	**2.2048**
Gamma	Weibull				5	**0.2102**	**0.7897**	**2.2054**

Table 24 confirms the decrease of reliability values for larger times. Moreover, it shows that $MTTF$ values are equal to 5.484 years for all trials of MLE. It is approximately 2.205 years for BA. There is a very small difference between the two BA trials.

In addition, $R(t)$ is smaller for any trial of BA compared with its corresponding MLE value. Take $tsurvival$ equal to 1 year for example, the $R(t)$ for MLE is 95%, whilst for BA is 23 %. Here the influence of prior parameters is notorious. The opposite occurs with $F(t)$ values as it is a complementary function.

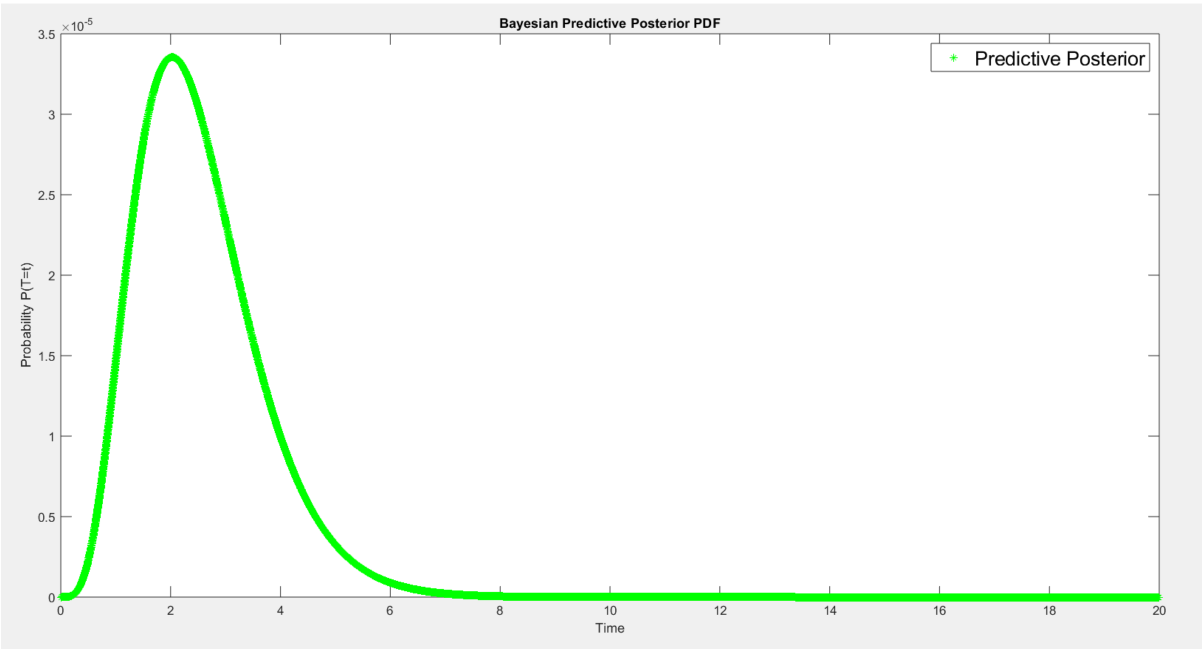


Figure 45. PDF based on BA given a Gamma predictive prior and the likelihood data of the study case

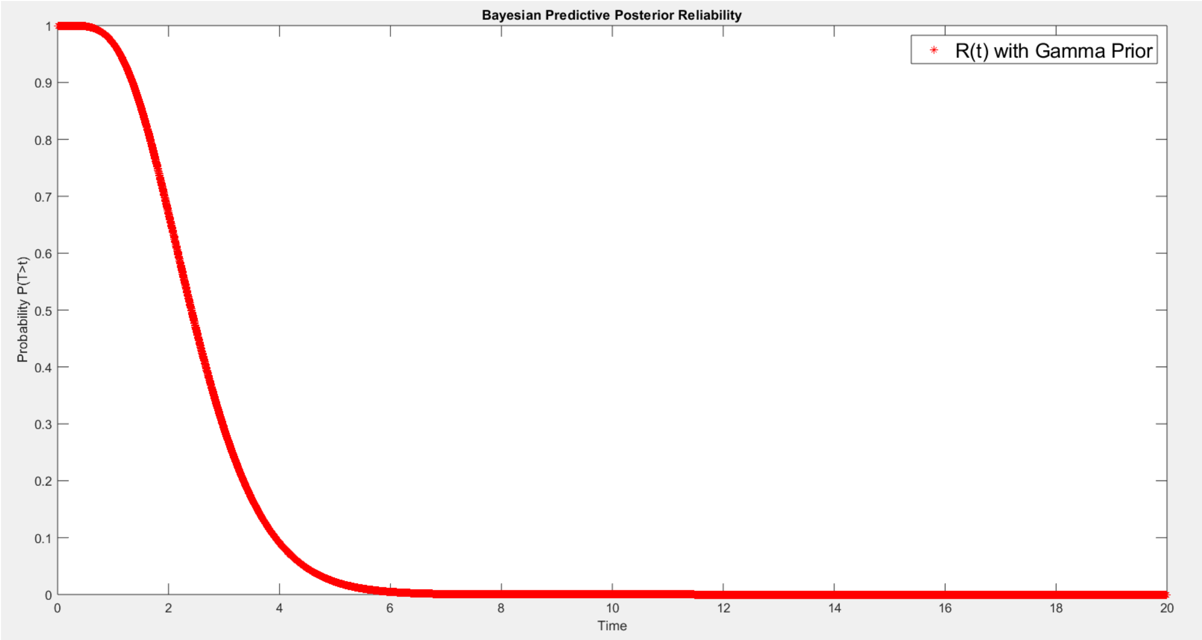


Figure 46. Reliability based on BA given a Gamma predictive prior and the likelihood data of the study case

Prior	Likelihood	Posterior
'Gamma'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	1	0.23616	0.76384	2.2048

mttfresults =

1×5 table

mttf	years	months	days	hours
2.2048	2	2	13	17.256

MTTF is estimated from todays date

Figure 47. BA results for Gamma predictive prior and $tsurvival = 1\ year$ as seen in Matlab's command window

Figure 43 and 44 confirm that $MTTF$ for the study case assuming a Gamma prior with the default failure rate is between 2 years 2 months 13 days and 17 hours and 2 years 2 months 13 days and 17 hours. This slight difference is a consequence of the difference between 2.2048 and 2.2054 years.

Prior	Likelihood	Posterior
'Gamma'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	5	0.21024	0.78976	2.2054

mttfresults =

1×5 table

mttf	years	months	days	hours
2.2054	2	2	13	22.381

MTTF is estimated from todays date

Figure 48. BA results for Gamma predictive prior and $tsurvival = 5\ years$ *as seen in Matlab's command window*

Once the four PDFs have been evaluated as predictive priors for the same likelihood data, the most suitable is the Gamma predictive prior. According to the values of $MTTF$, $R(t)$ and $F(t)$, the predictive posterior for a Gamma predictive prior and the likelihood data has the larger values. Hence, the highest reliability among all others.

4.2 Sensitivity Analysis

Three different initial parameters for Gamma predictive prior are used to account for the three most common approaches: cautionary, average and risky. The idea is to have an overview of the influence that the elicited initial parameters of predictive priors have on the predictive posterior distribution. In other words, how do they affect the values of $MTTF$, $R(t)$ and $F(t)$ for the Server-database study case in combination with the likelihood data.

The elicitation of initial parameters is done for a conservative and risky perspective, and the default initial parameters and used on the previous result are going to be taken as the average perspective. Table 25 shows the expected lifetime for the Server-database study case according to different approaches for initial parameters of the Gamma predictive prior for sensitivity analysis.

Table 25. Expected lifetime of Server-database according to different approaches for initial parameters of the Gamma predictive prior

Approach	Method of Estimation *BA*		Initial Parameters of Predictive Prior α and β	Expected Lifetime		
	Prior	Likeli.		$R(t)$	$F(t)$	$MTTF$
To generate nprior times $\hat{\phi}_{Prior}$ `estimator=gamrnd(alpha,beta,1,2);`						
Cautionary	Gamma	Weibull	`alpha=rand()*2;` `beta=rand()*1;`	**0.0211**	**0.9788**	**0.2165**
Average			`alpha=rand()*4;` `beta=rand()*2;`	**0.2106**	**0.7893**	**2.2089**
Risky			`alpha=rand()*6;` `beta=rand()*4;`	**0.5031**	**0.4968**	**5.4351**

Table 25 has a fixed number of $1x10^6$ iterations, $tsurvival\ of\ 5\ years$ and procedure to generate $\hat{\phi}_{Prior}$. Under these conditions the three approaches show very different results for the three indicators of expected lifetime such as $R(t)$, $F(t)$ and $MTTF$. For example, the cautionary approach has a probability of surviving 5 years of 2% and a mean time to failure of 0.21 years, or 2 months, 18 days and 1 hour. Whilst the risky approach has a probability of surviving 5 years of 50% and its mean time to failure is 5 years, 3 months, 3 days and 10 hours. The results for cautionary and average approach as seen in Matlab's command window are available on appendix A9 and A10 respectively.

The latter indicates the influence of initial parameters of prior predictive distributions on the predictive posterior even in presence of likelihood data, which is the case for the Server-database study case. Furthermore, it seems logical to assume that this influence will be even greater when there is no likelihood data available as all the weight of the estimation of expected lifetime relies on the elicited expert knowledge. This scenario is presented next.

A scenario to perform sensitivity analysis for the Server-database study case with no likelihood data is convenient to determine the influence of such historical data on the predictive posterior and the deviation of the results from each approach of the elicited expert knowledge regarding the initial parameter of the predictive prior.

Table 26 shows the expected lifetime for the Server-database study case according to different approaches for initial parameters of the Gamma predictive prior for sensitivity analysis, for the case when there is no likelihood data available.

Table 26. Expected lifetime of Server-database according to different approaches for initial parameters of the Gamma predictive prior with no likelihood data

Approach	**Method of Estimation *BA***		**Initial Parameters of Predictive Prior α and β**	**Expected Lifetime**		
	Prior	**Likeli.**		$R(t)$	$F(t)$	$MTTF$
To generate nprior times $\hat{\phi}_{Prior}$ `estimator=gamrnd(alpha,beta,1,2);`						
Cautionary	Gamma	None	`alpha=rand()*2;` `beta=rand()*1;`	**0.0182**	**0.9817**	**0.4428**
Average			`alpha=rand()*4;` `beta=rand()*2;`	**0.2857**	**0.7174**	**7.1328**
Risky			`alpha=rand()*6;` `beta=rand()*4;`	**0.6007**	**0.3992**	**63.988**

Table 26 has also a fixed number of 1x10^6 iterations, $tsurvival\ of$ 5 $years$ and procedure to generate $\hat{\phi}_{Prior}$. With BA done based only on predictive prior distribution, the three approaches show completely different results for the three indicators of expected lifetime such as $R(t)$, $F(t)$ and $MTTF$. For example, the cautionary approach has a probability of surviving 5 years of 1.8% and a mean time to failure of 0.44 years or 5 months, 9 days and 9 hours. The Average approach has a probability of surviving the 5th year of 28% a mean time to failure of 7.1328 years which is equivalent to 7 years, 1 month, 17 days and 19 hours.

The risky approach has a probability of surviving 5 years of 60% and its mean time to failure is an unrealistic 63.988 or 63 years, 11 months, 25 days and 17 hours. The results as seen in Matlab's command window are available on appendix A10, A11 and A12 respectively.

CHAPTER 5

5.1 Discussion of Bayesian Analysis Results

As a basis for the discussion, it is important to mention the reasons why the Server-database study case was selected in chapter 4.

1. The size of the historical data is relatively small, specifically thirty-nine failure times. Therefore, it has some influence into the predictive posterior but not enough to override the predictive prior influence.

2. The historical data contains only failure times of one system. Hence, the frequency of different types of failure occurred in such system does not have to be accounted for. In other words, they are treated as the same type.

3. The system has been operating for nearly 11 years, without major modifications or replacement. Thus, it can be considered highly reliable.

4. There are several factors to be considered when eliciting expert knowledge about the prior distribution.

5. Since the Server-database it is composed of electronic equipment, the selected PDF's for this report are suitable to some extent to model the lifetime according to its characteristics.

6. The results can be considerable useful to the company for establishing an improved maintenance program. Moreover, decision support for future upgrading or replacement.

The discussion is based on two main perspectives: the reliability perspective and risk and uncertainty perspective.

5.1.1 From the Reliability Perspective

An empirical way of estimating the expected time of failure based on historical data is to calculate the average of such data, then assume that the system will behave exactly in the future as it did in the past. However, easy and quick it is to implement this measure does not reflect entirely the "reality" of the system and its surrounding environment. Nevertheless, it can be used a reference to compare estimation of expected lifetime with more advanced methods, especially when it comes to highly reliable systems.

The average of the study case data presented on Figure 25 for years between failure is 5.555 years. At the same time the estimated $MTTF$ according to the MLE method is 5.484 years. The difference is considerably small; 0.071 years which is the equivalent to 0 years, 0 months, 25 days and 13 hours as presented to Figure 49.

```
mttfresults =

  1×5 table

    mttf     years    months    days    hours
    _____    _____    ______    ____    _____

    0.071      0        0        25     13.44

MTTF is estimated from todays date
```

Figure 49. Difference between Average time to failure and MLE's MTTF

Based on BA results of Table 25, the $MTTF$ has three values according to the approach of elicited prior knowledge and likelihood data. These are 0.2165 years, 2.2089 years and 5.4351 years for the corresponding cautionary, average and risky approach. Notice that the closest result to the MLE's $MTTF$ and the average failure time is the $MTTF$ of the risky approach. The difference between these two is only 0.049 years according to the displayed result of figure 50.

```
mttfresults =

  1×5 table

    mttf     years    months    days    hours
    _____    _____    ______    ____    _____

    0.049      0        0        17     15.36

MTTF is estimated from todays date
```

Figure 50. Difference between Average time to failure and risky approach for BA's MTTF

The previous indicates that both estimation methods produced very similar results for the Server-database study case, when considering a risk approach when eliciting expert knowledge regarding the initial parameter of predictive prior.

Whit regards to the reliability and unreliability results, BA method is sensitive to relatively small changes in the initial parameters of predictive prior and to the availability of relevant likelihood data. This is evident after observing the differences of $R(t)$ and $F(t)$ between MLE and BA estimates. For example, when comparing their values presented in the average approach of Tables 25 and 26 for a fixed $tsurvival$ of 5 years. $R(t)$ varies from 21% to 28.5%, $F(t)$ varies from 78.9% to 71.7% and even the $MTTF$ varies from 2.2 years to 7.13 years.

The previous results appear to be a direct consequence of the variation of the initial parameters for the Weibull predictive prior α and β. Furthermore, a consequence of likelihood data availability for the Server-database study case.

Another factor to consider when an analyst needs to interpret MLE and BA results from the reliability perspective is the censorship of the data. For the study case, the company provided the full set of data. Therefore, no censorship was considered. Nevertheless, the results for the censored data set of presented on Table 17 indicate a significant difference between MLE and BA according to the code for each estimation of expected lifetime. That case produced a MLE's $R(t)$ equal to 4.3% and BA's $R(t)$ equal to 12.7% for the Gamma distribution, which was the one that best fitted the censored data set of Integrated Circuit Failure Times.

Furthermore, the censored data set presents a difference between the MLE's $MTTF$ equal to 0.1573 years and the BA's $MTTF$ equal to 2.2904 for a Gamma predictive prior. This is equivalent to 2.1331 years or 2 years, 1 month, 17 days and 22 hours (Figure 51).

```
mttfresults =

  1×5 table

    mttf      years    months    days    hours 
   ______     _____    ______    ____    ______

   2.1331       2        1        17     21.984

MTTF is estimated from todays date
```

Figure 51. Difference between MLE's MTTF and BA's MTTF for Gamma predictive prior of censored data set

The final issue to consider from the reliability perspective is the nature of the predictive posterior as it is not a fixed known distribution. Predictive posteriors are the result of a combination between the attributes of predictive prior and predictive likelihoods. Therefore, according to the nature of the latter, predictive posterior can take different forms as known statistical distributions or completely different ones. This is accounted with rejection sampling in this report. Otherwise the statistical and computational complexity would had been considerable, as there would had been a unique distribution for every combination of priors and likelihood.

Through acceptance and rejection sampling the codes generate randomly a predefined number of values to create a representative sample of the predictive posterior values, which meet the acceptance conditions. As mentioned in chapter 3, This sample is based on a $nposterior$ number of iterations.

5.1.2 From the Risk and Uncertainty Perspective

From the risk and uncertainty perspective, there are several factors to be aware of when estimating expected lifetime using Bayesian analysis. Moreover, when is done with a software such as Matlab for this report. These are classified according to their context.

For prior distributions

There is uncertainty on the elicited initial parameters, since s they may not represent completely the true nature on the systems behaviour. Taking the sensitivity analysis of section 3.2.3 it is normal that variations in these parameters affect considerably the results. Furthermore, they can be misleading if not elicited properly.

In addition, uncertainty is present when performing acceptance-rejection sampling, as it does not provide 100% accurate results every time it is performed. The latter is especially true if $nprior$ is relatively small or the time range from which random values are generated is too wide.

Based on results presented on Table 24, it is visible that $MTTF$ values are not identical, for $tsurvival$ of 1 year is equal to 2.2048 and for $tsurvival$ of 5 years is 2.2054, but they are very similar. That small deviation from value to the other is a consequence of the acceptance-rejection sampling criteria and the size of $nprior$, which in the study case is fixed to $1x10^6$ iterations. For demanding maintenance programs larger difference can represent a large source of uncertainty.

For prior and likelihood distributions

A source of risk and uncertainty can be the method used to test the goodness of fitness for the CDF to the likelihood data. For the study cases the probability value of the selected fitting the data is 0.7423. In other words, approximately 74% of the data fits the scale and shape of a Weibull CDF.

As shown in Figure 52 the rank probability values for goodness of fit as seen in Matlab's command window, vary for each CDF, from 0.0308 for an Exponential, 0.0720 for a Lognormal and 0.3127 for a Gamma. Therefore, the higher value of the variable "rank" is selected. The lower the probability for goodness of fit is, the higher the risk of using and inappropriate CDF.

```
rank =

    0.0308    0.7423    0.0720    0.3127


likelihood =

    0.7423

Number of iterations for likelihood Parameters
```

Figure 52. Rank probability values for goodness of fit as seem in the Matlab's command window

A different source of risk and uncertainty is the probability of an extreme event to occur. Most of the time this factor is not reflected into the predictive likelihood data. Moreover, for predictive likelihood with a small size of historical data. The estimated parameters may not be entirely representative of the systems previous behaviour. The reason is that a small historical may not record a sudden event as it is unlikely to occur in a frequent basis. Nevertheless, a large $nlikelihood$ will increase the confidence in the intervals for such estimated parameters and to some extent compensate for the implied uncertainty.

The results present for the study case had a 95% of confidence in the estimated likelihood parameters. Compare to the results present on Tables 7, 9, 11 and 13 where the confidence interval changed from 95% to 75%, precisely to demonstrate the variation of such intervals for the estimated parameters.

Finally, it is necessary to clarify how to interpret the results in a risk context to avoid uncertainty. All $R(t)$ and $F(t)$ results are probability values, and they are interpreted as conditional probabilities according to the knowledge behind them. Based on the results presented in Table 24, it is understood that the Server-database system has a 23.61% of surviving the 1st year given the previous knowledge of the system and that it has already survived. The same applies for the probability of failure before the 1st year.

If an analyst interprets these results differently, for example, by taking a frequentist interpretation, it may be misleading when presented to the decision makers. See appendix A14.

5.2 Usefulness of Estimation of Expected Lifetime with Bayesian Analysis

To estimate the expected lifetime and other reliability indicators is considerable relevant. Given the results for the different examples of data sets and the study case presented in this report, an analyst may be able to assess the system behaviour with a higher degree of confidence. Moreover, it represents a powerful tool for statistical inference, particularly for cases where non relevant historical data is available, cases where the degree of uncertainty on the available historical data is extremely significant or for cases where strong expert knowledge is available.

Referring to the Server-database study case, results such as the ones presented in Tables 21, 22, 23 and 24 constitute a useful way to estimate critical indicators of system considered to be highly reliable. Especially, when the consequences can be severe in terms of economic losses, human safety, technological performance and environmental damage. For those situations it is necessary to estimate with a high degree of accuracy the expected lifetime, probability of surviving and probability of failure at any given time. BA provides such required estimations.

According to Table 18 most of the failures represent critical events for AKRA Investments c.a. Hence, their necessity for an updated maintenance and replacement program. With highly accurate estimates of $MTTF$, $R(t)$ and $F(t)$. An unpredicted failure in the Server-database can create considerable economical losses, logistical problems, reduction or interruption of daily operations, unnecessary risks to personnel and loss of data among others. Finally, MLE and BA codes are suitable to estimate the expected lifetime of different types of systems besides highly reliable and the results valid if some predefine conditions are met.

CHAPTER 6

6.1 Conclusions

Several conclusions can be made after carrying out MLE and BA for the examples and study case presented in this report. They are listed as follows:

1. BA becomes relevant to estimate expected lifetime when there is no likelihood data, or the available historical data of a system's performance is limited.

2. BA represents a powerful method to estimate the expected lifetime of highly reliable systems with a higher degree of accuracy in comparison to MLE and other estimation methods.

3. The acceptance-rejection sampling constitutes a valid method to create predictive posterior based on large samples of randomly generated data for BA.

4. To obtain more accurate estimations, extensive and strong expert knowledge is required to elicit the parameters for prior distributions.

5. In BA through MC methods, the number of iterations for sampling data of predictive prior, likelihood and posterior must be large enough, at least $1\text{x}10^6$, to have accurate estimates.

6. BA estimates of $R(t)$, $F(t)$ and $MTTF$ provide decision support in a risk and operational context, to perform the required maintenance or replacement when needed.

7. There are more than one sources of uncertainty in BA estimates. These can influence negatively the relevance of such estimates. The most critical are the lack of historical failure data, extensive censorship in the historical failure data and the occurrence of a black swan in the future.

6.2 Recommendations for Future Estimations

This section provides some recommendations to improve the estimations of expected lifetime and other performance indicators of highly reliable systems presented in this study.

- Include additional prebuilt function from the software to account for different initial parameters of predictive prior distributions following the methodology presented in the elicitation principles section. Specifically, by implementing formula 45.
- Edit the software to present the results of $R(t)$, $F(t)$ and $MTTF$ in a % format instead of decimals. For example, instead of $R(t) = 0.2857$ the result should be displayed as $R(t) = 28.57\%$ In that way it will be more practical to interpret for decision makers.
- Increase the number of PDFs available for prior and likelihood distributions with the following: Normal, Beta, Inverse Gaussian and Birnbaum-Sanders distributions.
- Asses the feasibility to incorporate replacement into the BA, so the estimates can be produced including that option, which is common in other types of systems than highly reliable.
- Incorporate the option of estimating the time for a specific $R(t)$ and $F(t)$ predefined value, required by a decision maker. For example, to estimate the expected time t for the system to have a given value or $R(t)$. This can be particularly useful for entities with predefined maintenance and replacement politics. In which given that the system has a $R(t)$ equal or lower than certain value, the system will be replaced regardless its condition.
- Compare the estimations of reliability indicators of different study cases with other computational codes if available. In addition, a sensitivity analysis is recommended in such codes as well.
- Perform a cost-benefit analysis of estimating the expected lifetime of highly reliable systems with the code for BA used in this thesis. Moreover, it should be focused on the code's accuracy, simplicity to implement and relevance of results in terms of maintenance and reliability policies.

REFERENCES

Albert, I. D.-J.-C. (2012). *Combining Expert Opinion in Prior Elicitiation* (503-532 ed.). International society for Bayesian Analysis.

Aven, T. (2012). *Foundations of Risk Analysis* (2nd edition ed.). New York: John Wiley & Sons.

Aven, T. (2012). *Risk Analysis* (3rd edition ed.). Stavanger: Wiley.

Aven, T. R. (2013). How to define and interpret a probability in a risk and safety setting. *Safety Science*, 223-231.

Ayyub, B. M. (2003). *Probability, statistics and reliability for engineers and scientists* (2nd edition ed.). Boca Raton: Chapman & Hall.

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis* (2nd ed.). Durham: Springer.

Casella, G. B. (2002). *Statistical Inference* (Second edition ed.). Brooks/Cole.

Chapman, S. (2004). *MATLAB Programming for Engineers* (Third edition ed.). Toronto: Thomson.

Eliason, S. (1993). *MAXIMUM LIKELIHOOD ESTIMATION, Logic and Practice.* Newbury Park: Sage Publications.

Gelman, A. (2014). *Bayesian data analysis 2014* (3rd edition ed.). Boca Raton: CRC Press.

Hamada, M. (2008). *Bayesian reliability.* New York: Springer.

Hanselman, D. L. (2012). *Mastering MATLAB* (International edition ed.). Harlow: Pearson.

Hasting, K. (1997). *Probability and Statistics.* New York: Addison-Wesley.

Hayakawa, I. (2001). *System and Bayesian Reliability. Essays in Honor Richard E. Barlow on his 70th birthday* (5th edition ed.). New Jersey: World Scientific.

Hinchey, M. C. (2010). Evolving Critical Systems: a Research Agenda for Computer-Based Systems. *17th IEEE International Conference and Workshops on Engineering of Computer-Based Systems*, 430.

Hogg, T. (2006). *Probability and statistical inference.* Upper Saddle River: Pearson Prentice Hall.

James, G. W. (2014). *An introducton to statistical learning: with applications in R.* New York: Springer.

Jurčiček, F. (2014). Types of Priors. *Institute of Formal and Applied linguistics, Charles University in Prague*.

Kendall, M. S. (1987). *Kendall's advanced theory of statistics* (5th edition ed.). New York: Oxford University Press.

Kruschke, J. (2015). *Doing Bayesian Data Analysis a Tutorial with R, JAGS and Stan* (2nd edition ed.). Oxford: Elsevier.

Kuhnert, M. (2010). A Guide to Eliciting and using expert knowledge in Bayesian ecological models. *Ecology Letters*, 900-914.

Lee, P. (2012). *Bayesian Statistics* (4th edition ed.). West Sussex: Wiley.

Lindley, D. (2014). *Understanding Uncertainty* (Revised edition ed.). New Jersey: Jhon Wiley and Sons, Inc.

Lydersen, S. (1988). Reliability Testing based on deterioration measurements. *Fakultet for Informasjonsteknologi, Matematikk og Elektroteknikk, NTNU*.

Magrab, E. A. (2011). *An Engineer's guide to MATLAB* (Third edition ed.). New Jersey: Pearson.

Meeker, W. E. (1998). *Staatistical Methods for Reliability Data.* New York: Wiley.

Mickelsson, G. (2015). Estimation of DSGE models: Maximum Likelihood vs. Bayesian Methods. *Departmanet of Economics, Uppsala University*.

Modarres, M. K. (2010). *Reliability engineering and risk analysis : a practical guide* (Third edition. ed.). Boca Raton: CRC Press,Taylor & Francis Group.

Pawitan, Y. (2001). *In All Likelihood Statistical Model and Inference Using Likelihood.* Stockholm: Clarendon Press Oxford.

Pham, H. (2001). *Recent advantages in Reliability and Quality Engineering.* New Jersey: World Scientific.

Renyan, J. (2015). *Introduction to Quality and Reliability Engineering.* Berlin: Springer.

Rossi, P. A. (2005). *Bayesian Statistics and Marketing.* New Jersey: Wiley.

Samaniego, F. (2010). *A Comparison of the Bayesian and Frequentist Approaches to Estimation.* London: Springer.

Skinner, D. (2009). *Introduction to Decision analysis: The practitioners guide to improve decision quality.* Sugar Land: Probabilistic Publishing.

Taleb, N. (2010). *The Black Swan: The impact of the highly improbable* (2nd edition ed.). New York: Random House.

Tobias, P. A. (2012). *Applied reliability* (Vol. 3rd). Boca Raton: CRC/Taylor & Francis.

Walpole, M. (2016). *Probability & Statistics for Engineers & Scientists* (International edition ed.). Boston: Pearson.

Ye, C. (2017). Maximum Likelihood-Like for the Gamma Distribution. *Department of Industrial and Systems Engineering, National University of Singapur*.

APPENDIX

Appendix A1

Example of $\hat{\Phi}_{MLE}$ estimators from theoretical censored data set.

From Table 15, assuming random values to the number of failures $n_1 = 4$, $n_2 = 6$ and $n_3 = 2$ and an exponential PDF. The maximum likelihood estimator $\hat{\Phi}_{MLE}$ for each case is presented in Table 15. According to the formula.

$$\hat{\Phi}_{MLE} = \hat{\lambda} = \frac{\sum_1^n t_i}{n}$$

Table A1. Example of MLE estimators from different types of censored data

Censored Data Set	Estimator
Observed Failures 1st Case	$\hat{\Phi}_{1^{st}} = \hat{\lambda}_{1^{st}} = \frac{n_1 - 1 + n_2 + n_3}{n_1 + n_2 + n_3} = \frac{11}{12} = 0{,}9166$
Observed Failures 2nd Case	$\hat{\Phi}_{2^{nd}} = \hat{\lambda}_{2^{nd}} = \frac{n_1 + n_2}{n_1 + n_2 + n_3} = \frac{10}{12} = 0{,}8333$
Observed Failures 3rd Case	$\hat{\Phi}_{3^{rd}} = \hat{\lambda}_{3^{rd}} = \frac{n_2}{n_1 + n_2 + n_3} = \frac{6}{12} = 0{,}50$

$$\hat{\Phi}_{3^{rd}} < \hat{\Phi}_{2^{nd}} < \hat{\Phi}_{1^{st}}$$

Appendix A2

Code for MLE of an Exponential random data set in Matlab R2018.

```
%UNIVERSITY OF STAVANGER
%RISK MANAGEMENT
%MLE OF EXPECTED LIFETIME FOR EXPONENTIAL DATA SET
%ANATOLY KURMAN RIVERO
clear
clc
disp('MLE for Exponential Data')
%Random Data set Exponentially distributed
disp('Press any key to adquire the data set    ')
pause
data=xlsread('sample.xlsx')
%Parameter Estimate
disp('Press any key for parameter estimates')
pause
lambda=rand()*3;
y=exppdf(data,lambda);
phat=mle(data,'pdf',@(data,lambda)exppdf(data,lambda),'start',[lambda]);
mleexp=phat
%Confidence Intervals
disp('Confidence intervals 100(1-alpha)%')
```

```
degree=input('Set the degree in % for the 100(1-degree) Confidence
Intervals     ')
if degree < 0
    disp('The degree for Confidence interval must be a positive integer')
    degree=input('Set the % for the 100(1-degree) Confidence Intervals   ')
end
degree_value=degree/100;
[muhat,muci]=expfit(data,degree_value);
muci
%New Exponential Distribution
t=rand(1000,1)*2;
mleexpdata = mleexp.*exp(-mleexp.*t);
%Probability of Survival P(T>tsurvival) and Failure P(T<=tsurvival) at any
given time tsurvival
disp('Reliability at a given time tsurvival')
tsurvival=input('Set the time tsurvival for Reliability psurvival     ')
T = mleexpdata>mleexpdata(tsurvival);
survival=mleexpdata(T);
length(survival);
pfailure = 1-exp(-mleexp.*tsurvival);
disp('Probability of Survival at a given tsurvival');
psurvival=1-pfailure;
disp('Mean time to failure of MLE Predictive Distribution    ');
mttfexp=1/mleexp;
disp('Press any key to see the table of results');
pause
results = table(lambda,mleexp,tsurvival,psurvival,pfailure,mttfexp)
filename = 'MLEExponentialEstimation.xlsx';
writetable(results,filename,'Sheet',1,'Range','A1')
reliability = exp(-mleexp.*t);
unreliability = 1-exp(-mleexp.*t);
%Plotting the distributions and Histogram
plot(t,unreliability,'b*');
title('Unreliability p(t<=tsurvival) Distribution')
ylabel('Probability')
xlabel('Time')
pause
clf
plot(t,reliability,'b*');
title('Reliability p(t>tsurvival) Distribution')
ylabel('Probability')
xlabel('Time')
pause
hist(mleexpdata,10);
title('New Exponential Histogram Distribution')
ylabel('Number of Events')
xlabel('Time')
disp('End of script')
```

Appendix A3

Code for MLE of a Weibull random data set in Matlab R2018.

```
%UNIVERSITY OF STAVANGER
%RISK MANAGEMENT
%MLE OF EXPECTED LIFETIME FOR WEIBULL DATA SET
%ANATOLY KURMAN RIVERO
clear
clc
disp('MLE for Weibull Data')
```

```
%Random Data set Weibull distributed
disp('Press any key to adquire the data set     ')
pause
data=xlsread('samplestudycase1.xlsx')
%Parameter Estimate
disp('Press any key for parameter estimates')
pause
alpha=rand()*2;
beta=rand()*2;
y=wblpdf(data,alpha,beta);
phat = mle(data,'distribution','weibull')
mlewbl1=phat(1,1);
mlewbl2=phat(1,2);
%Confidence Intervals
disp('Confidence intervals 100(1-alpha)%')
degree=input('Set the degree in % for the 100(1-degree) Confidence
Intervals    ')
if degree < 0
    disp('The degree for Confidence interval must be a positive integer')
    degree=input('Set the % for the 100(1-degree) Confidence Intervals
')
end
degree_value=degree/100;
[muhat,muci]=wblfit(data,degree_value);
muci
%New Weibull Distribution
t=rand(100000,1);
mlewbldata = ((phat(1,2).*t.^(phat(1,2)-1))./(phat(1,1).^phat(1,2))).*exp(-
(t./phat(1,1)).^phat(1,2));
mlewblplot = wblrnd(phat(1,1),phat(1,2),100000,1);
%Probability of Survival P(T>tsurvival) and Failure P(T<=tsurvival) at any
given time tsurvival
disp('Reliability at a given time tsurvival')
tsurvival=input('Set the time tsurvival for Reliability psurvival     ')
T = mlewbldata>mlewbldata(tsurvival);
survival=mlewbldata(T);
length(survival);
pfailure = 1-exp(-(tsurvival./phat(1,1)).^phat(1,2));
disp('Probability of Survival at a given tsurvival');
psurvival=1-pfailure;
disp('Mean time to failure of MLE Predictive Distribution    ');
mttfwbl=phat(1,1).*gamma((1+phat(1,2))./phat(1,2));
disp('Press any key to see the table of results');
pause
results =
table(alpha,beta,mlewbl1,mlewbl2,tsurvival,psurvival,pfailure,mttfwbl)
filename = 'MLEExponentialEstimation.xlsx';
writetable(results,filename,'Sheet',1,'Range','A1')
reliability = exp(-(t./phat(1,1)).^phat(1,2));
unreliability = 1-exp(-(t./phat(1,1)).^phat(1,2));
%Plotting the distributions and Histogram
plot(t*1000,unreliability,'b*');
title('Unreliability p(t<=tsurvival) Distribution')
ylabel('Probability')
xlabel('Time')
pause
clf
plot(t*10,mlewbldata,'b*');
title('Reliability p(t>tsurvival) Distribution')
ylabel('Probability')
xlabel('Time')
```

```
pause
hist(mlewblplot,100);
title('New Weibull Histogram Distribution')
ylabel('Number of Events')
xlabel('Time')
disp('End of script')
```

Appendix A4

Code for MLE of a Lognormal random data set in Matlab R2018.

```
%UNIVERSITY OF STAVANGER
%RISK MANAGEMENT
%MLE OF EXPECTED LIFETIME FOR LOGNORMAL DATA SET
%ANATOLY KURMAN RIVERO
clear
clf
clc
disp('MLE for Lognormal Data')
%Random Data set lognormally distributed
disp('Press any key to adquire the data set    ')
pause
data=xlsread('sample.xlsx')
%Parameter Estimate
disp('Press any key for parameter estimates')
pause
mu=round(rand()*5);
sigma=round(rand()*5);
phat=mle(data,'pdf',@(data,mu,sigma)lognpdf(data,mu,sigma),'start',[mu,sigm
a]);
mlelogn=phat
%Confidence Intervals
disp('Confidence intervals 100(1-alpha)%')
degree=input('Set the % for the 100(1-degree) Confidence Intervals    ')
if degree < 0
    disp('The degree for Confidence interval must be a positive integer')
    degree=input('Set the % for the 100(1-degree) Confidence Intervals
')
    end
degree_value=degree/100;
[muhat,muci]=lognfit(data,degree_value);
muci
%New Lognormal Distribution
t=rand(1000000,1)*20;
mlelogndata=((1./(mlelogn(1,2).*t.*sqrt(2*pi))).*exp((-
1/(2.*mlelogn(1,2).^2)).*log((t-mlelogn(1,1)).^2)));
%Probability of Survival P(T>tsurvival) and Failure P(T<=tsurvival) at any
given time tsurvival
disp('Reliability at a given time tsurvival')
tsurvival=input('Set the time for Reliability p(T>tsurvival)    ')
T = mlelogndata>tsurvival;
survival=mlelogndata(T);
length(survival);
psurvival=length(survival)/length(mlelogndata)
pause
disp('Probability of Failure at a given tsurvival')
pfailure=1-psurvival
pause
disp('Mean time to failure of MLE Predictive Distribution    ')
mttflogn=mlelogn(1,1)*exp((mlelogn(1,2)^2)/2)
```

```
disp('Histogram of Lognormal Distribution')
pause
%Ploting the histogram
hist(mlelogndata,50)
title('New Lognormal Histogram Distribution')
ylabel('Events')
xlabel('time')
pause
plot(t,mlelogndata,'b*')
title('New Lognormal Plot Distribution')
ylabel('Probability')
xlabel('time')
disp('End of script')
```

Appendix A4

Code for *MLE* of a Gamma random data set in Matlab R2018.

```
%UNIVERSITY OF STAVANGER
%RISK MANAGEMENT
%MLE OF EXPECTED LIFETIME FOR GAMMA DATA SET
%ANATOLY KURMAN RIVERO
clear
clf
clc
disp('MLE for Gamma Data')
%Random Data set Gamma distributed
disp('Press any key to adquire the data set    ')
pause
data=xlsread('sample1.xlsx')
%Parameter Estimate
disp('Press any key for parameter estimates')
pause
alpha=round(rand()*5);
beta=round(rand()*5);
[phat,pci] = gamfit(data,0.05);
mlegam=phat
%Confidence Intervals
disp('Confidence intervals 100(1-alpha)%')
degree = input('Set the % for the 100(1-degree) Confidence Intervals    ')
if degree < 0
    disp('The degree for Confidence interval must be a positive integer')
    degree=input('Set the % for the 100(1-degree) Confidence Intervals
')
    end
degree_value=degree/100;
[muhat,muci]=gamfit(data,degree_value);
muci
%New Gamma Distribution
t=rand(10000,1)*5;
mlegamdata=(1./(mlegam(1,2).^mlegam(1,1)).*gamma(mlegam(1,1))).*(t).^(mlega
m(1,1)-1).*exp(-(t./mlegam(1,2)));
%Probability of Survival P(T>tsurvival) and Failure P(T<=tsurvival) at any
given time tsurvival
disp('Reliability at a given time tsurvival')
tsurvival=input('Set the time for Reliability p(T>tsurvival)    ')
T = mlegamdata>tsurvival;
survival=mlegamdata(T);
length(survival);
psurvival=length(survival)/length(mlegamdata)
```

```
pause
disp('Probability of Failure at a given tsurvival')
pfailure=1-psurvival
pause
disp('Mean time to failure of MLE Predictive Distribution    ')
mttf=mlegam(1,1)*mlegam(1,2)/(365)
disp('Histogram of Weibull Distribution')
pause
%Ploting the histogram
hist(mlegamdata,50)
title('New Gamma Histogram Distribution')
ylabel('Events')
xlabel('time')
pause
plot(t,mlegamdata,'b*')
title('New Gamma Plot Distribution')
ylabel('Probability')
xlabel('time')
years=fix(mttf);
months=fix((mttf-fix(mttf))*12);
days=fix((((mttf-fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30);
hours=((((mttf-fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30-fix((((mttf-
fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30))*24;
mttfresults=table(mttf,years,months,days,hours)
disp('MTTF is estimated from todays date')
today=datetime('today')
disp('End of script')
```

Appendix A6

Example of elicited Example of $\widehat{\phi}_{prior}$ estimators from different expert opinions.

Table A2 contains the 1st scenario has three different elicited expert opinions for the parameter of the same exponential PDF. To estimate $\widehat{\phi}_{prior}$by using formula #.

Table A2. 1st scenario with three different elicited expert opinions for prior parameters

Expert	Expert's opinion weight a_i	Expert's Estimator	$\lvert\hat{\lambda}_{Prior} - \hat{\phi}_i\rvert$
1st	0.5	2	0.6
2nd	0.3	3	0.4
3rd	0.2	3.5	0.9

$$\hat{\phi}_{Prior} = \sum_{i=1}^{3} a_i * \hat{\phi}_i$$

$$\hat{\lambda}_{Prior} = a_1 * \hat{\lambda}_1 + a_2 * \hat{\lambda}_2 + a_3 * \hat{\lambda}_3$$

$$\hat{\lambda}_{Prior} = 0.5 * 2 + 0.3 * 3 + 0.2 * 3.5$$

$$\hat{\phi}_{Prior} = 2.6$$

Table contains the 2nd scenario has also different expert opinions for the parameter of the same exponential PDF. To estimate $\widehat{\phi}_{prior}$by using formula #.

Table A3. 2nd scenario with three different elicited expert opinions for prior parameters

Expert	Expert's opinion weight a_i	Expert's Estimator	$\lvert\hat{\lambda}_{Prior} - \hat{\phi}_i\rvert$
1st	0.4	5	0.24
2nd	0.3	4.2	0.56
3rd	0.3	4.6	0.16

$$\hat{\lambda}_{Prior} = 0.4 * 5 + 0.3 * 4.4 + 0.3 * 4.6$$

$$\hat{\phi}_{Prior} = 4.76$$

Notice that the values of $\hat{\lambda}_{Prior} - \hat{\phi}_i$ is smaller as the weights of the expert's opinions and the elicited prior parameters are closer to each other.

Appendix A7

Code for Bayesian Analysis of predictive posterior in Matlab R2018.

```
%UNIVERSITY OF STAVANGER
%RISK MANAGEMENT
%ESTIMATION OF EXPECTED LIFETIME OF HIGHLY RELIABLE SYSTEMS USING BAYESIAN
ANALYSIS
%ANATOLY KURMAN RIVERO
clear
clc
%AVAILABILITY OF DATA
%If There is data available press 1 for YES
%If there is no data available press 0 for NO
disp('ESTIMATION OF EXPECTED LIFETIME OF HIGHLY RELIABLE SYSTEMS WITH
BAYESIAN ANALYSIS');
disp('Is there any likelihood data available    ');
decision=input('1 Yes  0 No    ');
if decision == 0
disp('There is NO likelihood data only PRIOR');
%PRIOR DISTRIBUTION
disp('Select the prior PDF acccording to Expert Knowledge')
prior=input('1 Exponential  2 Weibull  3 Lognormal  4 Gamma    ');
switch prior
    case 1
    %Exponential Prior
    Prior={'Exponential'};
    disp('Prior PDF is Exponential');
    % Assume that prior is Exponential with lambda a random number between
0-2
    % Assume that the data is exponential
    nprior=input('Number of iterations for prior Parameters    ');
    for i=1:nprior
    lambda=rand()*5;
    estimator=exprnd(lambda,1,1);
    t(i)=exprnd(estimator,1,1);
```

```
    end
    case 2
    %Weibull Prior
    Prior={'Weibull'};
    disp('Prior PDF is Weibull');
    % Assume that prior is Weibull with alpha and beta random numbers
between 0-2
    % Assume that the data is Weibull
    nprior=input('Number of iterations for prior Parameters    ');
    for i=1:nprior
    alpha=normrnd(1,0.05);
    beta=normrnd(2,0.05);
    estimator=wblrnd(alpha,beta,1,2);
    t(i)=wblpdf(estimator(1,1),estimator(1,2));
    end
    case 3
    %Lognormal Prior
    Prior={'Lognormal'};
    disp('Prior PDF is Lognormal');
    % Assume that prior is Lognormal with mu and sigma random numbers
between 0-2
    % Assume that the data is Lognormal
    nprior=input('Number of iterations for prior Parameters    ');
    for i=1:nprior
    mu=0;
    sigma=0.25;
    estimator=lognrnd(mu,sigma,1,2);
    t(i)=lognrnd(estimator(1,1),estimator(1,2),1,1);
    end
    otherwise 4
    %Gamma Prior
    Prior={'Gamma'};
    disp('Prior PDF is Gamma');
    % Assume that prior is Gamma with alpha and beta random numbers between
0-2
    % Assume that the data is Gamma
    nprior=input('Number of iterations for prior Parameters    ');
    for i=1:nprior
    alpha=rand()*4;
    beta=rand()*2;
    estimator=gamrnd(alpha,beta,1,2);
    t(i)=gamrnd(estimator(1,1),estimator(1,2),1,1);
    end
end
%Probability of Survival P(T>tsurvival) and Failure P(T<=tsurvival) at any
given time tsurvival
disp('Reliability at a given time tsurvival')
tsurvival = input('set the time for evaluate the reliability p(T>tsurvival)
')
T = t > tsurvival;
survival=t(T);
length(survival);
psurvival=length(survival)/length(t);
disp('Probability of Failure at a given tsurvival');
pfailure=1-psurvival;
disp('Mean time to failure of Bayesian Predictive Distribution    ');
mttf=mean(t);
disp('Press any key to display the Histogram, PDF, Reliability and
Unreliability funtions')
pause
histogram(t,100)
```

```
title('Bayesian Prior Predictive Histogram')
ylabel('Number of Events')
xlabel('Data')
%RESULTS BAYESIAN PRIOR PREDICTIVE
nposterior=0;
%%
Likelihood={'None'};
Posterior={'Bayesian'};
typeresults=table(Prior,Likelihood,Posterior)
valueresults=table(nprior,nposterior,tsurvival,psurvival,pfailure,mttf)
years=fix(mttf);
months=fix((mttf-fix(mttf))*12);
days=fix((((mttf-fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30);
hours=((((mttf-fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30-fix((((mttf-
fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30))*24;
mttfresults=table(mttf,years,months,days,hours)
disp('MTTF is estimated from todays date')
today=datetime('today')
end
%LIKELIHOOD PRIOR AND POSTERIOR DISTRIBUTIONS ****************************
if decision == 1
disp('There is likelihood data and PRIOR');
%POSTERIOR DISTRIBUTION
%Imput of Data sets
disp('Press any key to adquire data set');
pause
data=xlsread('samplestudycase1.xlsx')
values=data(:,1);
disp('Press any key to evaluate the goodness of fit of each PDF for the
data sample')
pause
cd1=fitdist(data,'exponential');
[h1,fitexp] = kstest(values,'CDF',cd1,'alpha',0.01)
cd2=fitdist(data,'weibull');
[h2,fitwbl] = kstest(values,'CDF',cd2,'alpha',0.01)
cd3=fitdist(data,'lognormal');
[h3,fitlogn] = kstest(values,'CDF',cd3,'alpha',0.01)
cd4=fitdist(data,'gamma');
[h4,fitgam] = kstest(values,'CDF',cd4,'alpha',0.01)
rank=[fitexp,fitwbl,fitlogn,fitgam]
likelihood=max(rank)
nlikelihood=input('Number of iterations for likelihood Parameters    ');
nposterior=input('Set in the number of iterations for Posterior Parameters
');
switch likelihood

%***************************************************************************
For Exponential Likelihood
    case rank(1,1);
    disp('The selected PDF for the likelihood data is Exponential');
    Likelihood={'Exponential'};
    disp('Select the prior PDF acccording to Expert Knowledge')
    prior=input('1 Exponential  2 Weibull  3 Lognormal  4 Gamma    ');
    nprior=input('Number of iterations for prior Parameters    ');
        switch prior
        case 1
        Prior={'Exponential'};
        for i=1:nprior
        lambda=rand()*5;
        estimator=exprnd(lambda,1,1);
        t(i)=exprnd(estimator,1,1);
```

```
        end
        phat = mle(data,'distribution','exponential');
        for i=1:nlikelihood
        x(i)=exprnd(phat,1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        case 2
        Prior={'Weibull'};
        for i=1:nprior
        alpha=normrnd(1,0.05);
        beta=normrnd(2,0.05);
        estimator=wblrnd(alpha,beta,1,2);
        t(i)=wblpdf(estimator(1,1),estimator(1,2));
        end
        phat = mle(data,'distribution','exponential');
        for i=1:nlikelihood
        x(i)=exprnd(phat,1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        case 3
        Prior={'Lognormal'};
        for i=1:nprior
        mu=0;
        sigma=0.25;
        estimator=lognrnd(mu,sigma,1,2);
        t(i)=lognrnd(estimator(1,1),estimator(1,2),1,1);
        end
        phat = mle(data,'distribution','exponential');
        for i=1:nlikelihood
        x(i)=exprnd(phat,1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        otherwise 4
        Prior={'Gamma'};
        for i=1:nprior
        alpha=rand()*4;
        beta=rand()*2;
        estimator=gamrnd(alpha,beta,1,2);
        t(i)=gamrnd(estimator(1,1),estimator(1,2),1,1);
        end
        phat = mle(data,'distribution','exponential');
        for i=1:nlikelihood
        x(i)=exprnd(phat,1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        end

%***************************************************************************
For Weibull data
    case rank(1,2);
    disp('The selected PDF for the likelihood data is Weibull') ;
    Likelihood={'Weibull'};
    disp('Select the prior PDF acccording to Expert Knowledge')
```

```
prior=input('1 Exponential  2 Weibull  3 Lognormal  4 Gamma    ');
nprior=input('Number of iterations for prior Parameters    ');
 switch prior
    case 1
    Prior={'Exponential'};
    for i=1:nprior
    lambda=rand()*5;
    estimator=exprnd(lambda,1,1);
    t(i)=exprnd(estimator,1,1);
    end
    phat = mle(data,'distribution','weibull');
    for i=1:nlikelihood
    x(i)=wblrnd(phat(1,1),phat(1,2),1,1);
    end
    f=@(t)t;
    g=@(x)x;
    B = acceptance(t,x,2,1,nposterior);
    case 2
    Prior={'Weibull'};
    for i=1:nprior
    alpha=normrnd(1,0.05);
    beta=normrnd(2,0.05);
    estimator=wblrnd(alpha,beta,1,2);
    t(i)=wblpdf(estimator(1,1),estimator(1,2));
    end
    phat = mle(data,'distribution','weibull');
    for i=1:nlikelihood
    x(i)=wblrnd(phat(1,1),phat(1,2),1,1);
    end
    f=@(t)t;
    g=@(x)x;
    B = acceptance(t,x,2,1,nposterior);
    case 3
    Prior={'Lognormal'};
    for i=1:nprior
    mu=0;
    sigma=0.25;
    estimator=lognrnd(mu,sigma,1,2);
    t(i)=lognrnd(estimator(1,1),estimator(1,2),1,1);
    end
    phat = mle(data,'distribution','weibull');
    for i=1:nlikelihood
    x(i)=wblrnd(phat(1,1),phat(1,2),1,1);
    end
    f=@(t)t;
    g=@(x)x;
    B = acceptance(t,x,2,1,nposterior);
    otherwise 4
    Prior={'Gamma'};
    for i=1:nprior
    alpha=rand()*4;
    beta=rand()*2;
    estimator=gamrnd(alpha,beta,1,2);
    t(i)=gamrnd(estimator(1,1),estimator(1,2),1,1);
    end
    phat = mle(data,'distribution','weibull');
    for i=1:nlikelihood
    x(i)=gamrnd(phat(1,1),phat(1,2),1,1);
    end
    f=@(t)t;
    g=@(x)x;
```

```
        B = acceptance(t,x,2,1,nposterior);
     end

%**************************************************************************
For Lognormal data
    case rank(1,3);
    disp('The selected PDF for the likelihood data is Lognormal');
    Likelihood={'Lognormal'};
    disp('Select the prior PDF acccording to Expert Knowledge')
    prior=input('1 Exponential  2 Weibull  3 Lognormal  4 Gamma    ');
    nprior=input('Number of iterations for prior Parameters    ');
     switch prior
        case 1
        Prior={'Exponential'};
        for i=1:nprior
        lambda=rand()*5;
        estimator=exprnd(lambda,1,1);
        t(i)=exprnd(estimator,1,1);
        end
        phat = mle(data,'distribution','lognormal');
        for i=1:nlikelihood
        x(i)=lognrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        case 2
        Prior={'Weibull'};
        for i=1:nprior
        alpha=normrnd(1,0.05);
        beta=normrnd(2,0.05);
        estimator=wblrnd(alpha,beta,1,2);
        t(i)=wblpdf(estimator(1,1),estimator(1,2));
        end
        phat = mle(data,'distribution','lognormal');
        for i=1:nlikelihood
        x(i)=wblrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        case 3
        Prior={'Lognormal'};
        for i=1:nprior
        mu=0;
        sigma=0.25;
        estimator=lognrnd(mu,sigma,1,2);
        t(i)=lognrnd(estimator(1,1),estimator(1,2),1,1);
        end
        phat = mle(data,'distribution','lognormal');
        for i=1:nlikelihood
        x(i)=lognrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        otherwise 4
        Prior={'Gamma'};
        for i=1:nprior
        alpha=rand()*4;
        beta=rand()*2;
```

```
        estimator=gamrnd(alpha,beta,1,2);
        t(i)=gamrnd(estimator(1,1),estimator(1,2),1,1);
        end
        phat = mle(data,'distribution','lognormal');
        for i=1:nlikelihood
        x(i)=gamrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
     end

%***************************************************************************
For Gamma data
    otherwise rank(1,4);
    disp('The selected PDF for the likelihood data is Gamma') ;
    Likelihood={'Gamma'};
    disp('Select the prior PDF acccording to Expert Knowledge')
    prior=input('1 Exponential  2 Weibull  3 Lognormal  4 Gamma    ');
    nprior=input('Number of iterations for prior Parameters    ');
     switch prior
        case 1
        Prior={'Exponential'};
        for i=1:nprior
        lambda=rand()*5;
        estimator=exprnd(lambda,1,1);
        t(i)=exprnd(estimator,1,1);
        end
        phat = mle(data,'distribution','gamma');
        for i=1:nlikelihood
        x(i)=gamrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        case 2
        Prior={'Weibull'};
        for i=1:nprior
        alpha=normrnd(1,0.05);
        beta=normrnd(2,0.05);
        estimator=wblrnd(alpha,beta,1,2);
        t(i)=wblpdf(estimator(1,1),estimator(1,2));
        end
        phat = mle(data,'distribution','gamma');
        for i=1:nlikelihood
        x(i)=wblrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        case 3
        Prior={'Lognormal'};
        for i=1:nprior
        mu=0;
        sigma=0.25;
        estimator=lognrnd(mu,sigma,1,2);
        t(i)=lognrnd(estimator(1,1),estimator(1,2),1,1);
        end
        phat = mle(data,'distribution','gamma');
        for i=1:nlikelihood
        x(i)=gamrnd(phat(1,1),phat(1,2),1,1);
```

```
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
        otherwise 4
        Prior={'Gamma'};
        for i=1:nprior
        alpha=rand()*4;
        beta=rand()*2;
        estimator=gamrnd(alpha,beta,1,2);
        t(i)=gamrnd(estimator(1,1),estimator(1,2),1,1);
        end
        phat = mle(data,'distribution','gamma');
        for i=1:nlikelihood
        x(i)=gamrnd(phat(1,1),phat(1,2),1,1);
        end
        f=@(t)t;
        g=@(x)x;
        B = acceptance(t,x,2,1,nposterior);
     end
end
 %Probability of Survival P(T>tsurvival) and Failure P(T<=tsurvival) at any
given time tsurvival
    disp('Reliability at a given time tsurvival');
    tsurvival=input('set the time for evaluate the reliability
p(T>tsurvival)     ');
    T =  B > tsurvival;
    survival=B(T);
    length(survival);
    psurvival=length(survival)/length(B);
    pfailure=1-psurvival;
    mttf=mean(B);
    disp('Press any key to display the Histogram, PDF, Reliability and
Unreliability funtions')
    pause
    histogram(B,100)
    title('Bayesian Posterior Predictive Histogram')
    ylabel('Number of events')
    xlabel('Time')

%RESULTS BAYESIAN POSTERIOR PREDICTIVE
Posterior={'Bayesian'};
typeresults=table(Prior,Likelihood,Posterior)
valueresults=table(nprior,nlikelihood,nposterior,tsurvival,psurvival,pfailu
re,mttf)
years=fix(mttf);
months=fix((mttf-fix(mttf))*12);
days=fix((((mttf-fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30);
hours=((((mttf-fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30-fix((((mttf-
fix(mttf))*12)-fix((mttf-fix(mttf))*12))*30))*24;
mttfresults=table(mttf,years,months,days,hours)
disp('MTTF is estimated from todays date')
today=datetime('today')
end
disp('End of Script')
```

Appendix A8

Acceptance rejection sampling function in Matlab R2018. Where *c* is the variable to adjust the acceptance or rejection limit.

```
function B = acceptance(t,x,c,m,n)
 B = zeros(m,n); % Preallocate memory
 for i=1:m*n
     accept = false;
     u=rand();
     if x(i)*u*c <= t(i)
        B(i)=x(i);
     accept = true;
     end
 end
mttf=mean(B)
```

Appendix A9

Figure A1 presents the results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with a cautionary approach to elicit expert knowledge about the initial parameters of predictive prior.

Prior	Likelihood	Posterior
'Gamma'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	5	0.02112	0.97888	0.21662

mttfresults =

1×5 table

mttf	years	months	days	hours
0.21662	0	2	17	23.595

MTTF is estimated from todays date

Figure A1. Results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with a cautionary approach to elicit expert knowledge about the initial parameters of predictive prior with likelihood data

Appendix A10

Figure A2 presents the results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with an average approach to elicit expert knowledge about the initial parameters of predictive prior.

Prior	Likelihood	Posterior
'Gamma'	'Weibull'	'Bayesian'

valueresults =

1×7 table

nprior	nlikelihood	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	1e+06	1e+06	5	0.50245	0.49755	5.4262

mttfresults =

1×5 table

mttf	years	months	days	hours
5.4262	5	5	3	10.036

MTTF is estimated from todays date

Figure A2. Results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with an average approach to elicit expert knowledge about the initial parameters of predictive prior with Weibull likelihood data

Appendix A11

Figure A3 presents the results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with a average approach to elicit expert knowledge about the initial parameters of predictive prior and no likelihood data available.

Prior	Likelihood	Posterior
'Gamma'	'None'	'Bayesian'

valueresults =

1×6 table

nprior	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	0	5	0.28257	0.71743	7.1328

mttfresults =

1×5 table

mttf	years	months	days	hours
7.1328	7	1	17	19.482

MTTF is estimated from todays date

Figure A3. Results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with an average approach to elicit expert knowledge about the initial parameters of predictive prior

Appendix A12

Results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with a risky approach to elicit expert knowledge about the initial parameters of predictive prior and no likelihood data available.

Prior	Likelihood	Posterior
'Gamma'	'None'	'Bayesian'

valueresults =

1×6 table

nprior	nposterior	tsurvival	psurvival	pfailure	mttf
1e+06	0	5	0.60078	0.39922	63.988

mttfresults =

1×5 table

mttf	years	months	days	hours
63.988	63	11	25	17.905

MTTF is estimated from todays date

Figure A4. Results as seen on Matlab's command window for estimation of expected lifetime of Server-database study case with a risky approach to elicit expert knowledge about the initial parameters of predictive prior

Appendix A13

A frequentist probability of an event A is denoted PfA and it is defined as the fraction of times an event A occurs if the trail is repeated hypothetically an infinite number of times (Aven T. R., 2013). PfA is equal to the limit of the ratio between the number of times the event A occurs n. divided by the total amount of times the trial is performed under similar conditions. It is assumed that the output is a good estimate of the true probability of the event to occur.

Given that an infinite number of trials under similar conditions is impossible to conduct, the frequentist interpretation for probabilities is thus a mind constructed quantity.