MASTER

Coverage of lifetime confidence bounds for highly censored and few Weibull distributed data

Jongerius, Sabine C.

Award date:
2022

Coverage of lifetime confidence bounds for highly censored and few Weibull distributed data.

Sabine Jongerius
Supervised by Alessandro Di Bucchianico & Roberto Rocchetta

August 2022

Contents

Preface **3**

1 Introduction **4**
1.1 Original Problem Description . . . 4
1.2 Research Questions . . . 5
1.3 Outline of the report . . . 5

2 Mathematical Background **6**
2.1 Reliability analysis . . . 6
2.1.1 Failure time distribution function . . . 6
2.2 Censoring . . . 8
2.2.1 Type I censoring . . . 8
2.2.2 Type II censoring . . . 8
2.3 Weibull Distribution . . . 9
2.3.1 General form . . . 9
2.3.2 Alternative parametrization . . . 9

3 Literature Review **11**
3.1 Point Estimators . . . 11
3.2 Assumption in shape parameter . . . 11
3.3 Confidence Intervals . . . 12

4 Parameter Estimation Methods **13**
4.1 Likelihood for Weibull . . . 13

5 Confidence Interval and Hypothesis Testing Methods **17**
5.1 Fisher Matrix . . . 17
5.2 Normal approximate confidence bounds . . . 18
5.3 Nelson's method . . . 19
5.3.1 Parameter estimation Maximum Likeliood with given shape parameter and zero failures . . . 19
5.4 Likelihood ratio confidence bounds . . . 20
5.4.1 Bounds on parameter . . . 20
5.4.2 Bounds on Life Time and Reliability . . . 20
5.5 Hypothesis testing vs. confidence intervals . . . 21
5.6 Summary of the reviewed confidence bounds . . . 21
5.6.1 Zero failures . . . 21
5.6.2 Complementary methods . . . 22
5.7 Coverage probability . . . 22
5.8 Reliability testing: A simulation study . . . 22
5.8.1 The proposed Monte Carlo simulation procedure . . . 23

6 Analysis **25**
6.1 Effect of sample size and censoring rate 26
6.2 Effect of the assumed shape parameter 27
6.3 Bounds comparison for zero failures 27
6.4 Simulation results 28
6.4.1 Average of the lower confidence bounds for high censoring rates 28
6.4.2 Coverage probabilities 28

7 Conclusions **31**
7.1 Conclusions 31
7.2 Further research 32

Preface

This report is the result of my graduation project for the master program Industral and Applied Mathematics at the Eindhoven University of Technology (TU/e). The subject of this report was suggest by my supervisor Alessandro Di Bucchianico. The underlying actual industrial problem was part of the Studygroup Mathematics with Industry (SWI), proposed by SKF (Svenska Kullagerfabriken AB).

I would like to thank my supervisors Alessandro Di Bucchianico and Roberto Rocchetta for their immense support, patience and time in the process of my graduation. Alessandro, I very much appreciate your attention to detail, your personal guidance throughout this process and our talks about music and the university. Roberto, thank you very much for the intensive help with the code, numerous coffees and your time in general.

I would like to thank my parents for believing in me, no matter which choices I make. Their love and support is a constant throughout my life and education, making it much easier to deal with possible bumps in the road.

Dear friends, thank you for being there, with distracting activities or pep talks. Luuk, thank you for the unconditional support in the final phase of my education, always with fun and food.

And although it is a great cliché, put in a song by ABBA: "Thank you for the music". Without it, I would not be half as motivated, happy and distracted as I am now.

Chapter 1

Introduction

Acceptance tests are a set of experimental methods whose primary goal is to determine if new products meet specific expectations on their durability and reliability. Many companies perform mechanical, physical, and chemical tests to certify new products, like bearings and PCBs. Because well-designed engineering components seldom fail (due to their high durability), censoring is often inevitable. Although reliability acceptance tests have a solid theoretical foundation, very high censoring rates and data scarcity might complicate the analysis and prevent reliability engineers from delivering a final judgment on the acceptability of the products. Especially for highly reliable products, the lack of data uncertainty must be carefully addressed and formally quantified. This procedure is necessary to avoid: unjustified rejections of new reliable prototypes, the definition of wrong warranty contacts, and unexpectedly high maintenance costs.

The goal of this report is to look into reliability testing with a small number of items (less then 10) and high censoring rate, meaning two or less failures occur before termination of the test. This problem is inspired by an actual industrial problem, proposed in the 2018 edition of the Studygroup Mathematics with Industry (`www.swi-wiskunde.nl`). The problem was proposed by SKF (Svenska Kullagerfabriken AB). SKF is a global technology provider offering products and services related to bearings and units, seals, mechatronics, and lubrication systems. Its headquarters are located in Sweden. The company has around 165 production sites in 28 countries. SKF has several research centres, including one in the Netherlands in Nieuwegein.

1.1 Original Problem Description

Mechanical bearings are an important product of SKF. They are mechanical elements that constrain motions to desired motions only, and at the same time reduce friction between moving parts. There is a wide range of applications of bearings, including bicycles, cars, manufacturing machines, trains, wind turbines and airplanes. Sizes of bearings range from less than 10 mm to 14 m.

Since bearings are essential for the proper and safe functioning of machines and equipment, it is essential for SKF to give customers reliable information on the performance. The performance of mechanical bearings is expressed through their life, i.e., the amount of time or number of revolutions that a bearing is capable to reach within nominal functioning. Bearing performance is assessed via intensive life testing and physical modelling. To validate and communicate this guaranteed performance, customers may also require SKF to perform tailor-made acceptance life tests under specific operating conditions from their application. It is standard in this branch of industry to present the results of these acceptance tests in terms of parameters of the Weibull probability distribution. The precision of these Weibull parameter estimates depends strongly on the test strategy (number of bearings tested, test duration, number of observed failures, and replacement policy). Acceptance tests are due to high costs usually performed with a limited sample size meant to survive a test under severe conditions. To avoid being too conservative, acceptance tests may also be passed with very limited failures (1 or 2), which makes the statistical estimation of the life parameters (in particular the 10th life percentile) more challenging. It is thus necessary to have sound and transparent estimation methods that work well with practical sample sizes to be able to reach reliable acceptance decisions. Unfortunately, existing methods like the Normal approximation (Fisher Matrix) method do not perform well small sample sizes with high censoring rates. As a consequence, these methods cannot be used in industrial practice, in particular for expensive capital goods. It is not clear why these methods do not work well and how they can be improved or adapted.

1.2 Research Questions

Based on the problem description of Section 1.1, we identified the following research questions that we will address in this report:

1. Investigate the accuracy and sensitivity of the current reliability estimation methods for small data sets with few or no failures.

2. Investigate other estimation methods that do not suffer from the drawback of current reliability estimation methods for small data sets with few or no failures.

3. Study the effect of model assumptions on the coverage probability of reliability bounds.

1.3 Outline of the report

This report is structured as follows. In Chapter 2 we supply the mathematical background necessary to describe the research questions and existing reliability demonstration methods in precise mathematical terms. An overview of relevant literature can be found in Chapter 3. In Chapter 4 we describe the different estimation methods for parameters of the Weibull distribution in detail. Chapter 5 then proceeds with the different methods for construction of confidence intervals and hypothesis testing. To address the research question we propose and perform a simulation study in Chapter 6. In Chapter 7 we will present conclusions from the simulation study and literature review to answer the stated research questions and describe further research suggestions.

Chapter 2

Mathematical Background

In this chapter we will describe the necessary mathematical background to reliability analysis which we will use in the sequel of this report. Apart from introducing and discussing general notions, we will present the Weibull distribution in detail.

2.1 Reliability analysis

In this section we will briefly describe the concept of reliability analysis and its key aspects. Reliability analysis is the analysis of reliability data. Reliability analysis is similar to survival analysis in a mathematical context, where the term reliability analysis is used in a more industrial setting and survival analysis in a medical setting (Meeker and Escobar (1998)). We will focus on the more industrial setting with reliability analysis. In a time with rapid developing technology there is a need for reliable and safe products. Reliability is often defined as the probability that a product (or system, machine etc.) will perform as intended under operating conditions, for a specified time. Improving reliability is of high interest for developers of new products and systems, and for this reason analyses are designed to quantify the reliability. This involves collection of reliability data from different types of tests and monitoring.

We speak about reliability analysis when we analyze life and failure data to predict product life. For this analysis we need a notion of time. An important aspect of time is non-negativity. However, for reliability analysis we need to specify time even more. This is not just simply calendar time, we need the effective testing time or number of rotations or some other measure related to the effective performance period.

2.1.1 Failure time distribution function

In reliability we aim to express aspects of life time of a specific product. Lifetime data may be discrete, but for the purpose of this research we assume lifetime to be continuous. We assume that the lifetimes are from a nonnegative probability distribution with density f and distribution function F. We can express the relationship of these functions as shown in Table 2.1. So either of these can fully determine the probability distribution of the nonnegative continuous random variable. In lifetime data, we are naturally more interested in survival probabilities $P(T > t)$ than in the the probability $P(T < t)$, related to the cumulative distribution function. Therefore we use the reliability function R, defined as

$$R(t) = 1 - F(t) = P(T > t). \tag{2.1}$$

This clearly also completely determines the probability distribution of a nonnegative random variable, since $R = 1 - F$.

With reliability data, it is often of interest to consider survival probabilities given that the item at hand has already survived a certain time. Reliability functions are not suited for this purpose, and therefore another notion is being used.

Definition 2.1.1 *Let T be a nonnegative continuous random variable with density f and reliability function R. Then the failure rate λ (also known as failure intensity or hazard rate h) is for $t \geq 0$ defined as*

$$\lambda(t) = \frac{f(t)}{R(t)} \tag{2.2}$$

The cumulative failure rate Λ (also known as cumulative hazard function H) is defined through the relation $\Lambda(t) = \int_0^t \lambda(x)\,dx$.

With this definition we can fully express the failure time distribution. In reliability analysis a measure of high interest is the residual life after a certain amount of time, given that the item has not failed beforehand. We can express that as follows:

Definition 2.1.2 *The mean residual life m of a nonnegative random variable T is defined as the function*

$$m(t) = \mathbb{E}\,(T - t \mid T \geq t). \tag{2.3}$$

In industrial applications, this is often referred to as Remaining Useful Life (RUL).

We thus have different mathematical objects that fully characterize the probability distribution of a nonnegative continuous random variable. The following theorem shows that they are equivalent and fully characterize the probability distribution.

Theorem 2.1.3 *The probability distribution of nonnegative continuous random variable is fully determined by each of the following objects:*

1. *reliability function, R,*
2. *cumulative distribution, F,*
3. *density function, f,*
4. *failure rate, λ,*
5. *mean residual life, m.*

Every characterization fully determines the probability distribution of a nonnegative continuous variable, since we can express every characterization in terms of another one. In Table 2.1, from Di Bucchianico and Castro (2021), the relations between the different characterizations of lifetime distributions are given. For reliability demonstration, it is important to quantify the fraction of items that are

	$F(t)$	$f(t)$	$R(t)$	$\Lambda(t)$	$m(t)$
$F(t) =$	$-$	$\int_0^t f(x)\,dx$	$1 - R(t)$	$1 - e^{-\Lambda(t)}$	$1 - \frac{m(0)}{m(t)}\, e^{-\int_0^t \frac{1}{m(u)}\,du}$
$f(t) =$	$F'(t)$	$-$	$-R'(t)$	$\lambda(t)\, e^{-\Lambda(t)}$	$\frac{m(0)(m'(t)+1)}{m^2(t)}\, e^{-\int_0^t \frac{1}{m(u)}\,du}$
$R(t) =$	$1 - F(t)$	$\int_t^\infty f(x)\,dx$	$-$	$e^{-\Lambda(t)}$	$\frac{m(0)}{m(t)}\, e^{-\int_0^t \frac{1}{m(u)}\,du}$
$\lambda(t) =$	$\frac{F'(t)}{1-F(t)}$	$\frac{f(t)}{\int_t^\infty f(x)\,dx}$	$\frac{-S'(t)}{S(t)}$	$-$	$\frac{1+m'(t)}{m(t)}$
$m(t) =$	$\frac{\int_t^\infty (1-F(x))\,dx}{1-F(t)}$	$\frac{\int_t^\infty \int_x^\infty f(u)\,du\,dx}{\int_t^\infty f(x)\,dx}$	$\frac{\int_t^\infty S(x)\,dx}{S(t)}$	$e^{\Lambda(t)} \int_t^\infty e^{-\Lambda(x)}\,dx$	$-$

Table 2.1: Relationship between the different characterizations of lifetime distributions.

guaranteed to survive at least a given time period. Mathematically, we can state this as quantiles of a probability distribution. To be more precise, this can be stated in terms of a probability level p from the reliability function R. This is also known as the $100p$th quantile of the distribution which expresses that proportion $1 - p$ survives at least until time L_p (in industry this is directly expressed in percentages, so L_{10} is the 10th percentile). The expression is thus simply

$$L_p = R^{-1}(p) \qquad \text{with } p \in [0, 1] \tag{2.4}$$

In the remainder of the report we will mainly focus on L_{10} as this is a widely used measure of interest in industrial application.

2.2 Censoring

In reliability analysis it is often impossible to test the products of interest until they are all failed. In medical context one could imagine that patients cannot be followed after certain treatments until their death. This also holds in the case of testing bearings, which are designed to last multiple decades. These bearings are tested under accelerated life testing, i.e. testing under more severe conditions than normal in such a way that using physical degradation laws one can connect lifetimes under severe conditions to life times under normal conditions (see e.g., (Meeker and Escobar, 1998, Chapter 19)). However, even under this type of testing, not all bearings will fail after reasonable test time. Censored data is life data where not all failures occur within the time frame in which data is collected, due to removal of the subject, time constraints or other circumstances. In order not to lose information from the data, statistical techniques are applied to handle the censored data. There are several types of censoring. Here we state the two types relevant for this report. We restrict ourselves to single right censoring, where we know the starting time of the test but we do not know the end time or number of failures, due to termination of the experiment.

2.2.1 Type I censoring

Suppose we have n items, and we test until predetermined time T. We assume that the lifetime of a bearing $X_{(i)}$ is a nonnegative continuous random variable with density function f that depends on a parameter θ and F is the related cumulative distribution function. By $Y_{(i)}$ we denote the observable lifetime. Then

$$Y_i = \min(X_i, T) \quad \text{and} \quad \delta_i = \begin{cases} 1 & \text{if} \quad X_i \leq T \\ 0 & \text{if} \quad X_i > T \end{cases}$$

Thus we call it Type I censoring when the test is terminated after a fixed time T. The number of failed items r at time T is then a random variable R. This random variable R follows a binomial distribution with success probability F_θ. The likelihood function for type I censoring can be obtained by first conditioning on R:

$$\begin{aligned} L(x_{(1)}, \ldots, x_{(n)}; \theta) &= f_{\theta; X_{(1)}, X_{(2)}, \ldots, X_{(n)} | R=r}(x_{(1)}, \ldots, x_{(n)}) P_\theta(R = r) \\ &= r! \left(\prod_{i=1}^{r} \frac{f_\theta(x_{(i)})}{F_\theta(T)} \right) \binom{n}{r} F_\theta(T)^r (1 - F_\theta(T))^{n-r} \end{aligned}$$

For derivations of this and more complex censoring schemes, see (Bain and Engelhardt, 1991, Chapter 2). Simplifying the last expression results in the likelihood function for Type-I censoring:

$$L(x_{(1)}, \ldots, x_{(n)}; \theta) = \frac{n!}{(n-r)!} \left(\prod_{i=1}^{r} f_\theta(x_{(i)}) \right) (1 - F_\theta(T))^{n-r}, \tag{2.5}$$

which is defined for $0 \leq x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(k)}$, the ordered failure times.

2.2.2 Type II censoring

We speak of Type II censoring when we end the test after a fixed number of $k \leq n$ failures. Since k is deterministic we can now express the likelihood function as the joint density of the k first order statistics (see (Bain and Engelhardt, 1991, Chapter 2)). This type of censoring is thus mathematically easier to handle than type I censoring. The likelihood function for type II censoring equals

$$L(x_{(1)}, \ldots, x_{(n)}; \theta) = \frac{n!}{(n-k)!} \left(\prod_{i=1}^{k} f_\theta(x_i) \right) (1 - F_\theta(x_{k:n}))^{n-k} \tag{2.6}$$

which is defined for $0 \leq x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(k)}$, the ordered failure times. In this type of censoring, the number of failures is fixed, while the end time is random.

2.3 Weibull Distribution

The most widely used probability model for industrial reliability data is the Weibull distribution. Rinne (2008) is an extensive monograph discussing many details of the Weibull distribution. A reason for the popularity of the Weibull distribution in industrial reliability applications, is that it allows flexibility in modelling failure rates, since it can represent both increasing and decreasing failure rates (aging process or infant mortality). There is also a justification through a "weakest link" argument, which basically is a limit theorem for extremes rather than sums. The Weibull distribution is member of the family of extreme value distributions.

2.3.1 General form

A random variable X is two-parameter Weibull distributed with parameters $\beta > 0$ and $\alpha > 0$ if its probability density function is given by

$$f_X(x|\beta,\alpha) = \begin{cases} \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}} & x > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

Note that different parametrizations of the Weibull distribution are used in statistics, but in reliability this is the most common parametrization, next to the parametrization we will present in Section 2.3.2. The related distribution function equals

$$F(x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}} \tag{2.8}$$

where $\beta > 0$ is a shape parameter and $\alpha > 0$ is a scale parameter. The corresponding reliability function is

$$R(x) = e^{-\left(\frac{x}{\alpha}\right)^{\beta}}. \tag{2.9}$$

With these expressions we can see that when parameter $\beta = 1$ we have a special case that $\text{Wei}(\alpha, 1) \sim \text{Exp}(\alpha)$. If shape parameter β is known, then we have $X \sim \text{Wei}(\alpha, \beta)$ then $X^{\beta} \sim \text{Exp}(\alpha^{\beta})$. We can prove this using the reliability function. We have $X \sim \text{Wei}(\alpha, \beta)$ and then the reliability function equals $R(x) = e^{-\left(\frac{x}{\alpha}\right)^{\beta}} = e^{-\left(\frac{x^{\beta}}{\alpha^{\beta}}\right)}$, which is the reliability function for the exponential distribution for $X^{\beta} \sim \text{Exp}(\alpha^{\beta})$.

Another transformation that can be performed to a Weibull distribution is to an Extreme Value type I distribution. An Extreme Value type I distribution is defined as $X \sim EVI(\eta, \theta)$, then $P(X > x) = e^{-e^{\frac{x-\eta}{\theta}}}$. The transformation is done via a log-transformation. If $\log(X) \sim \text{Wei}(\alpha, \beta)$, then $X \sim \text{EVI}(\log(\alpha), 1/\beta)$. We will show this again with the reliability function. When we have $\log X \sim \text{Wei}(\alpha, \beta)$ we have

$$R(\log x) = e^{-\left(\frac{\log x}{\alpha}\right)^{\beta}} \tag{2.10}$$

$$= e^{-\left(e^{x-\log\alpha}\right)^{\beta}} \tag{2.11}$$

$$= e^{-\left(e^{\frac{x-\log\alpha}{1/\beta}}\right)}. \tag{2.12}$$

This is equal to the reliability function of the Extreme Value Type I Distribution with location parameter $\log\alpha$ and scale parameter $1/\beta$. This transformation has the advantage that it is now a location-scale distribution, which is convenient for constructing confidence intervals and hypothesis tests. Another advantage is that asymptotic convergence is faster for this type of transformation (see e.g., Meeker and Nelson (1976)).

2.3.2 Alternative parametrization

In industry a different parametrization of the Weibull distribution is commonly used. This parametrization replaces the scale parameter α by the 10%-quantile of the distribution (denoted in industry by

L_{10}), as presented in 2.4. The parameters are linked to each other through the relation

$$L_{10} = \alpha \left(-\log(9/10)\right)^{1/\beta}. \tag{2.13}$$

In this representation, the reliability function of the Weibull distribution has the form:

$$R_{L_{10},\beta}(x) = \frac{9}{10}\; e^{-\left(\frac{x}{L_{10}}\right)^{\beta}}. \tag{2.14}$$

Note that these two relations are specific to the 10%-quantile of the reliability distribution. More generally, the L_p lifetime can be expressed as follows

$$L_p = \alpha \left(-\log\left(1-p\right)\right)^{1/\beta}. \tag{2.15}$$

where $1 - p$ defines a target reliability level. Similarly, the reliability distribution function (Weibull) is given by:

$$R_{L_p,\beta}(x) = (1-p)\; e^{-\left(\frac{x}{L_p}\right)^{\beta}}. \tag{2.16}$$

Note that $p \in [0, 1]$ but it often indicated in percentage in the L_p subscript, i.e., selecting a $p = 0.1$ we indicate a lifetime L_{10}.

Chapter 3

Literature Review

In this chapter we present an overview of the literature on research that is either directly relevant or related research. Section 3.1 deals with point estimators for the parameters of the Weibull distribution, with emphasis on the case of zero or few failures. In section 3.2 we present literature that may provide more insight in assuming a shape parameter when estimating a Weibull Distribution. We conclude this chapter with an overview of literature on the different methods for constructing confidence intervals in Section 3.3.

3.1 Point Estimators

Several methods are being used for parameter estimation of Weibull data. Maximum likelihood estimation is known to be theoretically optimal (at least asymptotically), but in industrial practice regression methods are widely used. Maximum likelihood estimators have been derived in Cohen (1965) and can be shown to be unique when they exist as proven in Farnum and Booth (1997) (see also Section 4.1). Corrections for the bias in the parameters have been discussed in several papers, including McCool (1970b), Hirose (1999), Yang and Xie (2003), and Shen and Yang (2015). Bias correction for the regression estimation method has been discussed in Zhang et al. (2006).

Genschel and Meeker (2010) compare Maximum Likelihood estimation method for 2-parameter Weibull distribution with Median Rank Regression. This method is proposed as preferable to Maximum Likelihood in certain settings by different articles. However, in the comparison study of Genschel and Meeker (2010), there is no clear setting where Median Rank Regression outperforms MLE. It is also important to note that Median Rank Regression requires at least two failures, which makes it not suitable for the problem of this report.

In order to overcome the problem that Maximum Likelihood estimators do not exists for Weibull data with zero failures, Jiang et al. (2010) introduce a modified version of Maximum Likelihood estimation through a shrinkage factor. They also indicate how to estimate the reliability function, and indicate a way to determine an appropriate shrinkage factor.

Several papers propose a Bayesian approach to circumvent the problems in parametric inference when there are no failures (see e.g., Fan and Chang (2009), Han et al. (2014), Jiang et al. (2015), and Zhang (2021). However, it is not clear what the influence of the choice of prior distribution has on the outcome.

3.2 Assumption in shape parameter

Huang and Porter (1990) present and prove a lower bound on the Reliability or Life Limit (L_{10}) for censored data with few or no failures assuming a Weibull distribution. This is an important lower bound since it deals with the uncertainty of (unestimated) scale parameter β. It requires conditions on the total sample size and number of units with equal maximal test time. When these conditions hold, a minimum value of β exists for the reliability or life time function. With this minimal value a lower bound for the lower bound can be expressed, which holds for every (true) value of β. If the conditions on the data are not met, then a "worst case" β does not exist, and this method becomes useless. Since we deal with not only few failures, but also small sample size in the setting of this report, the conditions on data do not hold, and we cannot use these lower bounds to decrease the risk of the pre-assumed value of β.

McCool (1970a) expresses hypothesis tests based on Monte Carlo Methods with Maximum Likelihood Estimators, which are only dependent on the sample size and the degree of censoring. Besides two tests for comparison of two populations, it provides tests for the Weibull shape parameter and the Weibull percentile. These tests could be of interest when checking the validity of the estimated β. However, these methods are not suited for small datasets with low censoring, this will lead to very wide confidence intervals.

3.3 Confidence Intervals

An extensive overview and comparison of different methods for constructing confidence bounds is done by Jeng and Meeker (2000) with estimation methods of likelihood ratio and different bootstrapping methods. However, in their simulation studies they only present results for data with more than one failure. This is done since ML estimators do not exist for zero failures and some confidence bounds do not exist for one failure. From these simulations we can see that bootstrapping methods might be very suitable for small data with few failures for confidence bounds on the scale parameter. However, these results cannot be directly translated to confidence intervals for lifetime lower bounds. They suggest several methods to improve estimation for data with less than 10 expected failures. However, these include redesigning the experiment or extending censoring time. If this is not possible, they suggest, if possible, nonparametric methods, but as is well-known, nonparametric methods require more data than parametric methods.

Probably the most relevant research for this report was done by Nelson (1985). He proposes confidence intervals for the scale parameter and measures of interest for Weibull data with few or no failures based on fixed assumed shape parameter, based on the exponential transformation, as described in Section 2.3. These confidence intervals are especially interesting data with no failures, since other estimation methods lead to unreasonable estimates. Keats et al. (2000) recommended not to use this method, especially in the case of zero failures, but that study was quite limited in scope and mainly focused on errors in the estimation of the scale parameter. As Xie et al. (2000) explained, the poor estimation properties of the scale parameter are to be expected since the estimators of the scale and shape parameters are correlated. It does not automatically mean that estimators for the mean life or percentiles perform badly (the last issue is discussed in Section 4 of Xie et al. (2000)).

Keats et al. (2000) perform a simulation study on these confidence intervals for the Weibull parameters after exponential transformation. They investigate the effect of the misspecification of shape parameter β on the estimate. From their simulation study they conclude that confidence intervals for Weibull using the exponential transformation should not be used in any case. A similar comparison study is performed by Yuan and Rai (2011) for different confidence bounds.

Magalhaes and Gallardo (2020) propose the Bartlett correction factor for the likelihood statistic with censored Weibull distributed data. They perform a Monte Carlo simulation to study the performance of this factor. This factor could also be very useful for small the confidence bounds in the few failure setting of this report. In the simulation study of Jeng and Meeker (2000) this correction factor performs well. The machinery to obtain this factor is rather involved, when there is censored data, simulation is needed to determine the factor.

Chapter 4

Parameter Estimation Methods

In this chapter we will give a detailed description of the Maximum Likelihood estimation method for parameters of the Weibull distribution with right censoring. Afterwards we will demonstrate the uniqueness of these ML estimators.

4.1 Likelihood for Weibull

In this section we will first derive Maximum Likelihood estimates for the parameters for the Weibull distribution with right-censored data. Furthermore, we will show whether these equations have a solution, and if that solution is unique. We will explain in detail the results of Farnum and Booth (1997). That paper concisely demonstrates the uniqueness of Maximum Likelihood Estimates for Weibull, even in the case of censoring. We will follow their steps in higher detail, to concisely show the uniqueness of these estimates. An important side-result is that the derivative of the equation for the β parameter has a monotonicity property, so that there are no numerical issues: any optimisation algorithm will easily find the maximum likelihood estimate.

Recall the likelihood function for right censored data from 2.5:

$$L(x_{(1)}, \ldots, x_{(n)}; \theta) = \frac{n!}{(n-r)!} \left(\prod_{i=1}^{r} f_\theta(x_{(i)}) \right) \prod_{i=r+1}^{n} (1 - F_\theta(x_{(i)})), \tag{4.1}$$

which is defined for $0 \leq x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(r)}$, the ordered failure times. Note that the likelihood functions for type I and type II censoring are similar, but we have to be careful to keep in mind which type of censoring we are dealing with. For type I censoring, the number of failures r is random and the censoring time for the $n-r$ non-failure items is deterministic, whereas for type II censoring, the number of failures is deterministic and the censoring time X_r is random. The following expressions hold for both types of censoring, if we interpret r as the number of non-censored observations (random or deterministic). Throughout this part we will use X_T as the censoring time, for both type I and type II censoring. For type I this is trivial, namely the preset end-time T, such that for all i that is non-failure at time T, $X_{(i)} = X_T$. For type II censoring $X_T = X_r$, where r is the number of failures after which we end the test. First recall the distribution function and density function for the Weibull distribution

$$F(x) = 1 - e^{-(\frac{x}{\alpha})^\beta} \tag{4.2}$$

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha} \right)^{\beta-1} e^{-(\frac{x}{\alpha})^\beta}. \tag{4.3}$$

In the following steps we will express the Maximum Likelihood Equations and use them to derive a function $h(\beta)$ in order to estimate β. Since this will be a function independent of α, we can use this function to show uniqueness of both estimates. The Likelihood for Weibull with censoring equals

$$L = \frac{n!}{(n-r)!} \prod_{i=1}^{r} \frac{\beta}{\alpha} \left(\frac{X_{(i)}}{\alpha}\right)^{\beta-1} \prod_{i=r+1}^{n} e^{-\left(\frac{X_T}{\alpha}\right)^{\beta}} \tag{4.4}$$

$$= \frac{n!}{(n-r)!} \left(\frac{\beta}{\alpha}\right)^{r} \prod_{i=1}^{r} \left(\frac{X_{(i)}}{\alpha}\right)^{\beta-1} e^{-\sum_{i=1}^{r}\left(\frac{X_{(i)}}{\alpha}\right)^{\beta} - (n-r)\left(\frac{X_T}{\alpha}\right)^{\beta}}. \tag{4.5}$$

From this point on, we will ignore the factor $\frac{n!}{(n-r)!}$, for convenience of notation. This is allowed, because this factor would not influence the Maximum Likelihood estimator as it is independent of the parameters. We take the logarithm of the likelihood function to express the log-likehood and simplify the result.

$$\begin{aligned}\log(L) = C &+ r\log\beta - r\log\alpha - r(\beta-1)\log(\alpha) + (\beta-1)\sum_{i=1}^{r}\log\left(X_{(i)}\right) - \sum_{i=1}^{r}\left(\frac{X_{(i)}}{\alpha}\right)^{\beta} \\ &- (n-r)\left(\frac{X_T}{\alpha}\right)^{\beta} \\ &= C + r\left(\log\beta - \beta\log\alpha\right) + (\beta-1)\sum_{i=1}^{r}\log(X_{(i)}) - \sum_{i=1}^{r}\left(\frac{X_{(i)}}{\alpha}\right)^{\beta} - (n-r)\left(\frac{X_T}{\alpha}\right)^{\beta}\end{aligned}$$

Then we express the derivatives of the log-likelihood with respect to α and β to obtain the MLE.

$$\frac{\partial \log L}{\partial \alpha} = \frac{-r\beta}{\alpha} + \beta\left(\frac{1}{\alpha}\right)^{\beta+1}\left(\sum_{i=1}^{r} X_{(i)}^{\beta} + (n-r)X_T^{\beta}\right) \tag{4.6}$$

$$\frac{\partial \log L}{\partial \beta} = \frac{r}{\beta} - r\log\alpha + \sum_{i=1}^{r}\log X_{(i)} - \left(\frac{1}{\alpha}\right)^{\beta}\left(\sum_{i=1}^{r} X_{(i)}^{\beta}\log\left(\frac{X_{(i)}}{\alpha}\right) + (n-r)X_T^{\beta}\log\left(\frac{X_T}{\alpha}\right)\right) \tag{4.7}$$

We set the derivative to α from 4.6 equal to 0 and simplify to obtain

$$\alpha^{\beta} = \frac{\sum_{i=1}^{r} X_{(i)}^{\beta} + (n-r)X_T^{\beta}}{r}. \tag{4.8}$$

This leads to the estimator

$$\alpha = \left(\frac{\sum_{i=1}^{r} X_{(i)}^{\beta} + (n-r)X_T^{\beta}}{r}\right)^{\frac{1}{\beta}}. \tag{4.9}$$

Note that this estimator for α is not suitable for a data set with zero failures, as one would divide by zero. We will present a possible alternative in Section 5.3.1.

In the following part we will look into the uniqueness of the estimates following from MLE. We set the derivative to β from 4.7 equal to 0, substitute the expression from 4.8 for α^{β} and rearrange the terms. This leads to

$$0 = \frac{1}{\beta} + \frac{1}{r}\sum_{i=1}^{r}\log X_{(i)} - \frac{\sum_{i=1}^{r} X_{(i)}^{\beta}\log X_{(i)} + (n-r)X_T^{\beta}\log X_T}{\sum_{i=1}^{r} X_{(i)}^{\beta} + (n-r)X_T^{\beta}}. \tag{4.10}$$

Then we arrive at the function $h(\beta) = \frac{1}{\beta}$:

$$h(\beta) = \frac{\sum_{i=1}^{r} X_{(i)}^{\beta}\log X_{(i)} + (n-r)X_T^{\beta}\log X_T}{\sum_{i=1}^{r} X_{(i)}^{\beta} + (n-r)X_T^{\beta}} - \frac{1}{r}\sum_{i=1}^{r}\log X_{(i)}. \tag{4.11}$$

We will now show that h is an increasing function of β, and with properties of h we will prove uniqueness of the maximum likelihood estimator.

First we want to show h is an increasing function of β. We compute the derivative

$$h'(\beta) = \frac{\sum_{i=1}^r X_{(i)}^\beta \left(\log X_{(i)}\right)^2 + (n-r)X_T^\beta \left(\log X_T\right)^2}{\sum_{i=1}^r X_{(i)}^\beta + (n-r)X_T^\beta} - \frac{\left(\sum_{i=1}^r X_{(i)}^\beta \log X_{(i)} + (n-r)X_T^\beta \log X_T\right)^2}{\left(\sum_{i=1}^r X_{(i)}^\beta + (n-r)X_T^\beta\right)^2}$$

Putting these two terms under a common denominator $\left(\sum_{i=1}^r X_{(i)}^\beta + (n-r)X_T^\beta\right)^2$, we see that the numerator equals

$$\left(\sum_{i=1}^r X_{(i)}^\beta + (n-r)X_T^\beta\right) \cdot \left(\sum_{i=1}^r X_{(i)}^\beta \left(\log X_{(i)}\right)^2 + (n-r)X_T^\beta \left(\log X_T\right)^2\right) - \left(\sum_{i=1}^r X_{(i)}^\beta \log X_{(i)} + (n-r)X_T^\beta \log X_T\right)^2.$$

Since the denominator $\left(\sum_{i=1}^r X_{(i)}^\beta + (n-r)X_T^\beta\right)^2$ of $h'(\beta)$ is positive, it suffices to that the numerator of $h'(\beta)$ is positive to conclude h is an increasing function. When we further expand the numerator we get

$$\sum_{i=1}^r X_{(i)}^\beta \sum_{i=1}^r X_{(i)}^\beta \log^2 X_{(i)} + (n-r)X_T^\beta \log^2 X_T \sum_{i=1}^r X_{(i)}^\beta + (n-r)X_T^\beta \sum_{i=1}^r X_{(i)}^\beta \log^2 X_{(i)}$$
$$+(n-r)X_T^{2\beta} \log^2 X_T - \left(\sum_{i=1}^r X_{(i)}^\beta \log X_{(i)}\right)^2 - (n-r)X_T^2 \log^2 X_T - 2(n-r)X_T^\beta \log X_t \sum_{i=1}^r X_{(i)}^\beta \log X_{(i)}.$$

Then, rearranging the terms we get:

$$\begin{aligned}
&\sum_{i=1}^r X_{(i)}^\beta \sum_{i=1}^r X_{(i)} \log^2 X_{(i)} - \left(\sum_{i=1}^r X_{(i)}^\beta \log X_{(i)}\right)^2 \\
&+(n-r)X_T^\beta \sum_{i=1}^n X_{(i)}^\beta \left(\log^2 X_T + \log^2(X_{(i)} - 2\log(X_T)\log(X_{(i)})\right) \\
&= \sum_{i=1}^r \left(X_{(i)}^{\beta/2}\right)^2 \sum_{i=1}^r \left(X_{(i)}^{\beta/2}\right)^2 \log^2(X_{(i)}) - \left(\sum_{i=1}^r X_{(i)}^{\beta/2}\left(X_{(i)}^{\beta/2}\log(X_{(i)})\right)\right)^2 \\
&+(n-r)X_T^\beta \sum_{i=1}^r X_{(i)}^\beta \left(\log(X_T) + \log(X_{(i)})\right)^2.
\end{aligned}$$

With the Cauchy-Schwarz inequality, we see that the first term is bigger or equal to the second term. Since the third term is positive, we now know that the numerator of $h'(\beta)$ is positive and thus $h(\beta)$ is increasing.

We will now show, using V for notation purposes, that

$$\lim_{\beta\to\infty} h(\beta) = \log(X_T) - \frac{1}{r}\sum_{i=1}^r \log(X_{(i)}) = V. \tag{4.12}$$

First recall $h(\beta)$ from 4.11 and rewrite as

$$h(\beta) = \frac{\sum_{i=1}^r \left(\frac{X_{(i)}}{X_T}\right)^\beta \log X_{(i)} + (n-r)\log X_T}{\sum_{i=1}^r \left(\frac{X_{(i)}}{X_T}\right)^\beta + (n-r)} - \frac{1}{r}\sum_{i=1}^r \log X_{(i)} \tag{4.13}$$

We will now distinguish two cases, first $X_{(i)} \leq X_r < X_T$, which is the case for type I censoring. Then

$$\lim_{\beta\to\infty} h(\beta) = \frac{(n-r)\log(X_T)}{n-r} - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)}) = \log(X_T) - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)}) = V \tag{4.14}$$

The second case is type II censoring where $X_r = X_T$ by definition. Then

$$\lim_{\beta\to\infty} h(\beta) = \frac{\sum_{i=1}^{r}\left(\frac{X_{(i)}}{X_T}\right)^{\beta}\log X_{(i)} + (n-r)\log X_T}{\sum_{i=1}^{r}\left(\frac{X_{(i)}}{X_T}\right)^{\beta} + (n-r)} - \frac{1}{r}\sum_{i=1}^{r}\log X_{(i)} \tag{4.15}$$

$$= \frac{\left(\frac{X_T}{X_T}\right)\log(X_T) + (n-r)\log(X_T)}{1+n-r} - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)}) = \log(X_T) - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)}) = V. \tag{4.16}$$

Then the final thing we need to show is $h(0) = (1-\frac{r}{n})V$:

$$h(0) = \frac{\sum_{i=1}^{r}\log(X_{(i)}) + (n-r)\log(X_T)}{r+(n-r)} - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)}) \tag{4.17}$$

$$= \frac{\sum_{i=1}^{r}\log(X_{(i)}) + (n-r)\log(X_T)}{n} - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)}) \tag{4.18}$$

$$= \left(\frac{1}{n} - \frac{1}{r}\right)\sum_{i=1}^{r}\log(X_{(i)}) + \frac{n-r}{n}\log(X_T) \tag{4.19}$$

$$= -\frac{1}{r}\left(1-\frac{r}{n}\right)\sum_{i=1}^{r}\log(X_{(i)}) + \left(1-\frac{r}{n}\right)\log(X_T) \tag{4.20}$$

$$= \left(1-\frac{r}{n}\right)\left(\log(X_T) - \frac{1}{r}\sum_{i=1}^{r}\log(X_{(i)})\right) = \left(1-\frac{r}{n}\right)V. \tag{4.21}$$

Along the lines of Farnum and Booth (1997) we now know three aspects of our Maximum Likelihood Estimates for α and β. First, if $V > 0$, since $h(\beta)$ is an increasing function for $\beta > 0$ and $\frac{1}{\beta}$ is decreasing, $\widehat{\alpha}$ and $\widehat{\beta}$ are unique. Secondly, if $V > 0$, then $\widehat{\beta} > 1/V$ and $\widehat{\alpha} > \left(n^{-1} \cdot \sum_{i=1}^{n} t_i^{1/V}\right)^V$. And thirdly, if and only if $V = 0$, all failure times are equal. For clarification in intuition, see Farnum and Booth (1997) for a visualisation of this derivation. Also recall that since $h(\beta)$ is monotone, so that there are no numerical issues: any optimisation algorithm will easily find the maximum likelihood estimate.

Chapter 5

Confidence Interval and Hypothesis Testing Methods

In the previous chapter, we derived point estimates for the Weibull distribution. In this chapter, we extend the mathematical background by reviewing and describing various methods for constructing confidence intervals. The derivation of confidence bounds depends on assumptions used for the derivations. Unfortunately, unwarranted assumptions like asymptotic behaviour, symmetry of two-sided bounds, and other model assumptions, may not hold in practice. This is especially true for lack of knowledge/data situations when the samples are limited and heavily censored. This report will look into different methods for computing confidence bounds and will target specific problems with severely censored data. For synthesis' sake, we will restrict our focus to a small number of methods of which some are specific to the Weibull distribution family. Moreover, we will only focus on one-sided confidence bounds, worst cases, which are often applied in the industry to tackle reliability qualification tasks. In the remainder of the report we will focus on estimating methods with known or assumed value for the shape parameter β, as suggested by for example Nelson (1985). Methods that are suited for simultaneously estimating both parameters are presented first presented in general. We will briefly look into the relationship between hypothesis testing and confidence intervals, which will be useful to study reliability demonstration hypothesis tests.

5.1 Fisher Matrix

In order to construct normal approximate confidence bounds, also known as Fisher's bounds, we need estimates for the variance of α and β based on the Fisher matrix. The observed variance-covariance matrix for a 2-parameter Weibull model can be defined from the inverse of the observed information matrix as follows, see e.g. Meeker and Escobar (1998):

$$\Sigma_{\widehat{\alpha},\widehat{\beta}} = \begin{bmatrix} \mathrm{Var}(\widehat{\alpha}) & \mathrm{Cov}(\widehat{\alpha},\widehat{\beta}) \\ \mathrm{Cov}(\widehat{\alpha},\widehat{\beta}) & \mathrm{Var}(\widehat{\beta}) \end{bmatrix} = \mathbb{E}\begin{bmatrix} -\frac{\partial^2\mathcal{L}(\alpha,\beta)}{\partial\alpha^2} & -\frac{\partial^2\mathcal{L}(\alpha,\beta)}{\partial\alpha\partial\beta} \\ -\frac{\partial^2\mathcal{L}(\alpha,\beta)}{\partial\beta\partial\alpha} & -\frac{\partial^2\mathcal{L}(\alpha,\beta)}{\partial\beta^2} \end{bmatrix}^{-1} = \frac{1}{\mathcal{I}(\alpha,\beta)}$$

where the $\mathrm{Var}(\cdot)$ and $\mathrm{Cov}(\cdot)$ are the variance and covariance operators, $\mathcal{I}(\alpha,\beta)$ is the observed Fisher information matrix for the shape and scale parameters (α,β), and $\mathcal{L}$ is the log-likelihood function. The factors $\widehat{\beta},\widehat{\alpha}$ are, respectively, estimators for the scale and shape parameters. Note that the Fisher information matrix quantifies the amount of information carried by an observable random variable about the unknown parameters α and β.

Specifically for the Weibull distribution family, the second derivative of the log-likelihood with respect to α is given by:

$$\frac{\partial^2\mathcal{L}(\alpha,\beta)}{\partial\alpha^2} = \sum_{i=1}^{r}\left[\frac{\beta}{\alpha^2} - \left(\frac{\beta+\beta^2}{\alpha^2}\right)\left(\frac{X_{(i)}}{\alpha}\right)^{\beta}\right] - (n-r)\left(\frac{\beta+\beta^2}{\alpha^2}\right)\left(\frac{X_T}{\alpha}\right)^{\beta}.$$

Note that, assuming β is known and fixed, the partial derivative $\frac{\partial \mathcal{L}(\alpha,\beta)}{\partial \beta} = 0$ and the information matrix reduces to $I(\alpha) = \mathbb{E}[-\frac{\partial^2 \mathcal{L}(\alpha,\beta)}{\partial \alpha^2}]$. For this case, the variance of the scale parameter can be simply expressed as follows, see e.g. Meeker and Escobar (1998),

$$\mathrm{Var}(\widehat{\alpha}) = \frac{1}{r}\left(\frac{\widehat{\alpha}}{\beta}\right)^2. \tag{5.1}$$

This expression will be used to construct normal approximate confidence bounds in the following section.

5.2 Normal approximate confidence bounds

The normal approximate confidence bounds are computed from asymptotic distributions, so the goodness of these bound generally rely on the availability of large sample sizes. However, under a lack of data, asymptotic confidence bounds may fail. Since it is a commonly used method, we will present it here. We can express normal approximate confidence bounds for a parameter θ in general, see e.g., LLoyd and Lipow (1977):

$$\theta_U = \widehat{\theta} \cdot e^{\frac{K_{\frac{1-\delta}{2}} \sqrt{\mathrm{Var}(\widehat{\theta})}}{\widehat{\theta}}} \text{ (Two-sided upper)}, \tag{5.2}$$

$$\theta_L = \frac{\widehat{\theta}}{e^{\frac{K_{\frac{1-\delta}{2}} \sqrt{\mathrm{Var}(\widehat{\theta})}}{\widehat{\theta}}}} \text{ (Two-sided lower)}. \tag{5.3}$$

Here K_p is the quantile function of the Standard Normal distribution, so $K_p = \Phi^{-1}(1-p)$, where $\Phi(x)$ is the distribution function of a Standard Normal random variable. Note that we use δ to indicate the significance level in order to avoid confusion with our parameter α (which is often used in the literature to express a confidence level). From equation (5.3) and replacing the variance of α with 5.1 the following lower confidence bound for scale parameter holds:

$$\alpha_L = \frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}} \cdot K_{\frac{1-\delta}{2}}}} \qquad r \geq 1. \tag{5.4}$$

Then we can transform this into a lower bound on the time for a specific reliability at level p. First recall the Reliability function, as presented in Section 2.3.2:

$$L_p = \alpha \left(-\log(R)\right)^{1/\beta}. \tag{5.5}$$

where $R = 1 - p \in [0, 1]$ stands for the targeted reliability level. In order to obtain a lower bound for L_p, we will substitute α in this expression with α_L. This leads to:

$$L_{p,L} = \frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}} \cdot K_{\frac{1-\delta}{2}}}} (-\log(R))^{\frac{1}{\beta}} \qquad r \geq 1 \tag{5.6}$$

Hence, by selecting a reliability level $R = 0.9$ and experession for the the $L_{10,L}$ can be derived as follows:

$$L_{10,L} = \frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}} \cdot K_{\frac{1-\delta}{2}}}} (-\log(9/10))^{\frac{1}{\beta}} \qquad r \geq 1 \tag{5.7}$$

These approximate normal confidence intervals are not suited for situations with a very limited number of failures or without failures. These bounds are not defined when there are zero failures, since we would divide by zero. Therefore, in the next section we present the method proposed by Nelson (1985), specifically aimed at few and zero failure data.

5.3 Nelson's method

Nelson (1985) presents a method for construction of confidence intervals for data with few or zero failures. This method is based on the fact that β can be assumed from engineering knowledge and historical data. Other methods in this chapter allow for simultaneously estimating α and β, but this is not possible for this method. However, one may never assume β is known. After constructing confidence bounds with this method, a sensitivity analysis on the value of β should be done, to get more insight in the quality of the bounds. The estimate for α as presented in Nelson (1985) is $\widehat{\alpha} = \left(\sum_{i=1}^{n} X_{(i)}^{\beta}/r\right)^{1/\beta}$, note that this is equal to (4.9), although it provides less insight in the structure of the data. Then with this estimate for α, Nelson (1985) introduced a lower $1-\delta$ confidence bound on α:

$$\begin{aligned} \alpha_L &= \widehat{\alpha}\left(\frac{2r}{\chi^2(1-\delta;2r+2)}\right)^{1/\beta} && \text{for } r \geq 1, \\ &= \left(\frac{2\sum_{i=1}^{n} X_{(i)}^{\beta}}{\chi^2(1-\delta;2r+2)}\right)^{1/\beta} && \text{for } r \geq 0. \end{aligned} \tag{5.8}$$

where $\chi^2(1-\delta;2r+2)$ is the chi-squared distribution with probability parameter (confidence) $1-\delta$ and degrees of freedom $2r+2$. We can also express the lower confidence bound for L_p as follows

$$\begin{aligned} L_{p,L} &= \alpha_L\left(-\log(R)\right)^{1/\beta} \\ &= \left(\frac{-2\log(R)\sum_{i=1}^{n} X_{(i)}^{\beta}}{\chi^2(1-\delta;2r+2)}\right)^{1/\beta}. \end{aligned} \tag{5.9}$$

The basis of these bounds lies in the transformation to the exponential distribution, as described in Section 2.3. To repeat it, when parameter β is known, then we have $X \sim \text{Wei}(\alpha,\beta)$ then $X^{\beta} \sim \text{Exp}(\alpha^{\beta})$. When we use this transformation, we have an estimate for the mean of the corresponding exponential distribution, $\widehat{\theta} = \sum_{i=1}^{n} X_{(i)}^{\beta}/r$. The corresponding Poisson $1-\delta\%$ confidence lower bound equals:

$$\theta_L = 2r\widehat{\theta}/\chi^2(1-\delta;2r+2) = \frac{2\sum_{i=1}^{n} X_{(i)}^{\beta}}{\chi^2(1-\delta;2r+2)}.$$

We can use this expression and transform back to get an expression for α, as written in 5.8.

In order to have an estimate for α when there are no failures in the data, an alternative is needed for 4.9. We will describe a possible alternative in the following section, as presented in Nelson (1985).

5.3.1 Parameter estimation Maximum Likeliood with given shape parameter and zero failures

The described estimate on α as in 4.9 does not hold when no failures are in the data. But as described in e.g., Nelson (1985) and (Meeker and Escobar, 1998, Chapter 8), we can use the lower bound estimate with a 50% confidence bound as a substitute for the parameter estimate. So the estimate will be, just as the lower bound mentioned earlier

$$\widehat{\alpha} = \alpha_L^* = \left(\frac{2\sum_{i=1}^{n} X_{(i)}^{\beta}}{\chi^2_{(0.5;2)}}\right)^{1/\beta} = \left(\frac{\sum_{i=1}^{n} X_{(i)}^{\beta}}{-\log(0.5)}\right)^{1/\beta} \tag{5.10}$$

since $\chi^2_{(1-\delta;2)} = -2\log(\delta)$. Note that Nelson (1985) describes that this estimate might be conservative (low) and that its theoretical performance has not been studied. However, there is no real alternative available, since our MLE for α divides by the number of failures. In the article of Nelson (1985) it is mentioned in the introduction that for less than 2 failures, alternative estimation methods are needed. However, this is not true for methods based on Maximum Likelihood estimates. For this method a minimum of 1 failure is needed, as we are dividing by r.

5.4 Likelihood ratio confidence bounds

The Likelihood Ratio bounds (LRB) method is sometimes preferred over approximate normal (or FM) method, especially in situations where there are smaller sample sizes. See e.g. Meeker and Escobar (1998) for short technical derivation. This methods is based on the Likelihood ratio equation:

$$-2 \cdot \log\left(\frac{L(\theta)}{L(\widehat{\theta})}\right) \geq \chi^2_{1-\delta;k}. \tag{5.11}$$

In equality of this expression we find the confidence bounds on parameters (no closed form). Using this, we can numerically determine the bounds. The bounds we will first describe are also known as uncorrected in e.g. Doganaksoy and Schmee (1993), as we just use above expression, without any corrections. In Section 3.3 we describe a possible correction factor to improve this method. Note that this method can be quite heavy numerically. In specific instances it might be necessary to work with loglikelihood expression, to improve numerical stability.

5.4.1 Bounds on parameter

We recall the Likelihood function for right-censored data as in 4.5:

$$L(\beta, \alpha) = \left(\frac{\beta}{\alpha}\right)^r \prod_{i=1}^{r} \left(\frac{X_{(i)}}{\alpha}\right)^{\beta-1} e^{-\sum_{i=1}^{r}\left(\frac{X_{(i)}}{\alpha}\right)^{\beta} - (n-r)\left(\frac{X_T}{\alpha}\right)^{\beta}} \tag{5.12}$$

which is defined for $0 \leq x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(r)}$, the ordered failure times. Then we calculate $L(\widehat{\beta}, \widehat{\alpha})$ with our MLE estimates for our parameters (or β fixed and α estimated). And then we can calculate

$$L(\beta, \alpha) - L(\widehat{\beta}, \widehat{\alpha}) \cdot e^{\frac{-\chi^2_{1-\delta;1}}{2}} = 0 \tag{5.13}$$

where δ is the specified confidence level, and with that we can calculate the value for the chi-squared statistic. Then values for β and α need to be found that satisfy this equation, since $L(\widehat{\beta}, \widehat{\alpha})$ is a known quantity with our data and MLEs. For simultaneously finding both parameters this is an iterative process, but for the purpose of this report we have a fixed β and we can just calculate the value of α.

5.4.2 Bounds on Life Time and Reliability

In order to express confidence bounds on Life time and Reliability, we need to transform 5.12 into a function of β and t or R. We rewrite the Reliability function $R = e^{-(\frac{t}{\alpha})^{\beta}}$ to express $\alpha = \frac{t}{(-\log(R))^{\frac{1}{\beta}}}$. We replace α in 5.12 to obtain a Likelihood function in terms of β and t.

$$L(\beta, t) = \left(\frac{\beta}{\frac{t}{(-\log(R))^{\frac{1}{\beta}}}}\right)^r \prod_{i=1}^{r} \left(\frac{X_{(i)}}{\frac{t}{(-\log(R))^{\frac{1}{\beta}}}}\right)^{\beta-1} e^{-\left(\frac{X_{(i)}}{\frac{t}{(-\log(R))^{\frac{1}{\beta}}}}\right)^{\beta} - (n-r)\left(\frac{X_T}{\frac{t}{(-\log(R))^{\frac{1}{\beta}}}}\right)^{\beta}}$$

$$= \left(\frac{\beta\,(-\log(R))^{\frac{1}{\beta}}}{t}\right)^r \prod_{i=1}^{r} \left(\frac{X_{(i)}\,(-\log(R))^{\frac{1}{\beta}}}{t}\right)^{\beta-1} e^{-\left(\frac{X_{(i)}(-\log(R))^{\frac{1}{\beta}}}{t}\right)^{\beta} - (n-r)\left(\frac{X_T(-\log(R))^{\frac{1}{\beta}}}{t}\right)^{\beta}}$$

Then we use

$$L(\beta, t) - L(\widehat{\beta}, \widehat{\alpha}) \cdot e^{\frac{-\chi^2_{1-\delta;1}}{2}} = 0 \tag{5.14}$$

in a similar way as for parameter bounds to calculate the Lifetime lower bound with confidence level δ. It is important to note that when β is known, we do not have to do this transformation on the Likelihood function. Since we do not need to estimate the β simultaneously we can use 2.13 with the estimated lower bound on α to express the lower confidence bound on L_{10}. So we just plug the estimate in the following

$$L_{10,L} = \alpha_L \left(-\log(9/10)\right)^{1/\beta}.$$

Then we obtain a lower confidence bound on L_{10}.

5.5 Hypothesis testing vs. confidence intervals

From our confidence intervals we can construct hypothesis tests. We will do this for the example of a lower bound for L_{10}. As in 5.7 we have expressed this one-sided $100(1-\delta)\%$ lower confidence bound for L_{10} as follows

$$L_{10,L} = \frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}}\cdot K_{1-\delta}}}\left(-\log(9/10)\right)^{\frac{1}{\beta}} \tag{5.15}$$

where $K_{1-\delta}$ is the quantile function for a Standard Normal distribution for quantile $1-\delta$. Then with $L_{10,L}$ we can express our hypothesis test as follows:

$$H_0 \; : \; L_{10} \geq L_{10,L} \tag{5.16}$$

$$H_a \; : \; L_{10} < L_{10,L} \tag{5.17}$$

We will reject H_0 when $L_{10} < L_{10,L}$ with significance level δ. We can rewrite this expression as follows

$$L_{10} \geq \frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}}\cdot K_{1-\delta}}}\left(-\log(9/10)\right)^{\frac{1}{\beta}} \tag{5.18}$$

$$\frac{L_{10}}{\widehat{\alpha}\left(\log(9/10)\right)^{1/\beta}} \geq \frac{1}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}}\cdot K_{1-\delta}}} \tag{5.19}$$

$$e^{\frac{1}{\beta}\sqrt{\frac{1}{r}}\cdot K_{1-\delta}} \geq \frac{\widehat{\alpha}\left(\log(9/10)\right)^{1/\beta}}{L_{10}} \tag{5.20}$$

$$K_{1-\delta} \geq \log\left(\frac{\widehat{\alpha}\left(\log(9/10)\right)^{1/\beta}}{L_{10}}\right)\frac{\beta}{\sqrt{\frac{1}{r}}}. \tag{5.21}$$

This construction of hypothesis tests related to the confidence intervals (or vice versa) can be done for all different parameter estimates.

5.6 Summary of the reviewed confidence bounds

In Table 5.1 we give an overview of the bounds reviewed in this section. We will analyze and discuss tightness and coverage of the bounds in Chapter 6. Note that we do not have a closed form expression for the Likelihood ratio method, since these are not available. The Normal approximation and Likelihood ratio is not (directly) applicable for zero failures, since there is not ML estimate, so an alternative for estimating α is needed, as described in the following section.

5.6.1 Zero failures

As described in Section 5.3.1, an estimate for the shape parameter is necessary to apply the reviewed method. Unfortunately, for $r = 0$, the MLE cannot be commuted, making M1-M3 not directly applicable. For zero failure, we need an alternative estimate for the scale parameter we need an alternative estimate to the maximum likelihood estimator for the scale parameter α. The proposed solution is to replace the MLE with the median estimate of the confidence bound from Nelson's method. For a data set with zero failures, the lower confidence bounds only depend on sample size, censoring time and the guess for the shape parameter β.

ID	Name	α_L	$L_{10,L}$	No-failures
M1	Normal approximation	$\frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}}\cdot K_{\frac{1-\delta}{2}}}}$	$\frac{\widehat{\alpha}}{e^{\frac{1}{\beta}\sqrt{\frac{1}{r}}\cdot K_{\frac{1-\delta}{2}}}}(-\log(0.9))^{\frac{1}{\beta}}$	No
M2	Nelson's	$\left(\frac{2\sum_{i=1}^{n} X_{(i)}^{\beta}}{\chi^2(1-\delta;2r+2)}\right)^{1/\beta}$	$\left(\frac{-2\log(0.9)\sum_{i=1}^{n} X_{(i)}^{\beta}}{\chi^2(1-\delta;2r+2)}\right)^{1/\beta}$	Yes
M3	Likelihood ratio	$L(\beta,\alpha_L) = L(\widehat{\beta},\widehat{\alpha})\cdot e^{\frac{-\chi^2_{\delta;1}}{2}}$	$L(\beta,L_{10}) = L(\widehat{\beta},\widehat{\alpha})\cdot e^{\frac{-\chi^2_{\delta;1}}{2}}$	No

Table 5.1: A summary of the reviewed confidence bounds and their applicability to no-failures situations.

In addition to that, we need to check whether the method for the confidence bounds can be used when there are zero failures. The normal approximation method has no defined expression for the lower confidence bounds on α when there are no failures and consequently there is no bound on L_{10}. For the likelihood ratio method this problem does not occur, however in literature there is no mention of the use of this method when there are few or no failures. So there is no research on the performance of the bound with zero.

5.6.2 Complementary methods

In general, more methods are available to estimate the lower confidence bounds on reliability data and specifically censored Weibull data. Some of these methods are studied in for example Jeng and Meeker (2000). However, few of these methods are suitable for few failures or few data points in general. For example all methods that involve bootstrapping require larger data sets with some failures. The selected methods for this report are the most used methods and the methods that are aimed at few failures.

5.7 Coverage probability

The coverage probability can be interpreted in terms of hypothetical repetitions of the entire data collection and statistical analysis procedure such that independent data sets are drawn from the same probability distribution and a confidence interval is computed from each of these data sets. The coverage probability is the fraction of these computed confidence intervals that include the desired but unknown and not observable parameter value. The coverage probability is the frequency of times that a confidence interval will contain the value of the target parameter and is defined as follows:

$$P_c = \mathbb{P}[\theta \geq \theta_L] \tag{5.22}$$

where θ_L is the lower confidence bound on an unknown parameter θ. For instance, the coverage probability of a one-sided lower confidence bound on the the L_{10} is $\mathbb{P}[L_{10} \geq L_{10,L}]$. Note that for CI the "nominal coverage probability" is often set at 0.95, e.g., the value of δ is set such that a 95 % confidence in the bounds is achieved. However, the true coverage probability is the actual probability that the confidence interval will contain the true target value. Note that, to estimate the coverage probability, the true L_{10} is required. This is never available in practice. A simulation study will be used to analyze coverage properties of different methods in a controllable and verifiable numerical environment. The interested reader is reminded to, e.g., Jeng and Meeker (2000); Kabaila and Leeb (2006); Diciccio and Romano (1988) and Brookmeyer and Crowley (1982), for further details on coverage probabilities.

5.8 Reliability testing: A simulation study

We will perform a simulation study to compare different estimation methods for Weibull distributed data with one failure and small sample size (6 samples) with type I censoring (time censoring on one specific point in time for all non-failure objects). We will perform a similar analysis to Jeng and

Meeker (2000), which investigated the performance of many different confidence interval methods, for censored data. In particular it looked at the coverage probability dependent on different censoring proportions and different sample sizes. They deliberately removed all observations with 1 or zero failures, as some of the methods behaved poorly in these cases. In line with the research question of this report, we will extend this analysis to cases where there are few failures, for several methods. We will compare the coverage probabilities of the L_{10} confidence bounds of the following methods:

M1: Normal-approximation,

M2: Nelson's method,

M3: Likelihood ratio.

Furthermore we will investigate the influence of a varying shape parameter, when this parameter is (wrongly) assumed to be known.

5.8.1 The proposed Monte Carlo simulation procedure

We propose a Monte Carlo (MC) simulator to generate time-to-failure and censoring indicators from a known Weibull distribution and assess the validity of the reviewed methods. A numerical simulator allows controlling the data generating mechanism, the censoring rates, experiment duration, and sample size. We will exploit this feature to study different experimental scenarios (low sample sizes, high-censoring rates) and issues related to model miss-specification. Moreover, because the underlying distribution generating lifetime data is known precisely, the true L_{10}, α, and β can be easily computed. We will use their reference values to check the goodness of statistical estimators, the reviewed confidence bounds and their coverage probabilities. In this work, we study three generating mechanisms and model errors:

(1) The assumed shape parameter, $\beta_{guess} = 1.5$,is exactly equal to the true value $\beta = 1.5$.

(2) The assumed shape parameter, $\beta_{guess} = 1.5$, overestimates the true value $\beta = 1.39$.

(3) The assumed shape parameter, $\beta_{guess} = 1.5$, underestimate the true value $\beta = 1.61$.

In both cases, the reliability distribution family selected to generate the data is the Weibull family. The scale parameter is set to $\alpha = 1283$. Note that for case (1) the true L_{10} is equal to 286.21 hours, for case (2) $L_{10} = 254.16$ hours, whilst for case (3) the L_{10} is only 317.09 hours. In other words, for case (3), about 10% of the components do not survive the 13 days of operations.

The numerical implementation of the MC simulator work as follows:

(i) **Data Generation**: First, we draw n samples from a known Weibull distribution with true $\alpha = 1283$ and β selected as case (1), (2) or (3). If a sample exceed the selected censoring time T (the duration of the experiment) their value is set equal to T and censoring indicator assigned to them. In this example, a T between 100 and 400 hours leads to high censoring rates, between approximately 82 % to 98% depending on the selected case.

(ii) **Estimation**: The L_{10} and α are estimated from the data generated in (i), see e.g. equation (2.13), and methods M1, M2 and M3 applied to derive $L_{10,L}$ and $\alpha_{10,L}$. We always apply a confidence level $\delta = 0.05$ to compute the bounds.

(iii) **Coverage**: An indicator function assigned to each interval as follows:

$$\mathbb{1}_A = \begin{cases} 1 & \text{if } A \text{ is true} \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\mathbb{1}_A$ is equal to 1 if the statement A is true and 0 otherwise. For instance if the statements of interest are $A := \{\alpha \geq \alpha_L\}$ and $A := \{L_{10} \geq L_{10,L}\}$, two indicator functions are assigned, one for each bound.

(iv) **Iterate and post-process the results:** Steps (i) to (iii) are repeated N times and the results processed to estimate coverage probabilities and average values of the lower confidence bounds. For instance, coverage probabilities are estimated as follows,

$$P_{c,i} = \mathbb{E}[\mathbb{1}_A] \approx \frac{\sum_{i=1}^{N} \mathbb{1}_{\{A\}}}{N}.$$

We repeat this procedure for various experimental set ups, for instance, by increasing sample sizes and test duration (and this decreasing censoring rates).

Chapter 6

Analysis

In this section, we will use the simulation tool described in Section 5.8 to investigate the sensitivity, coverage probabilities and tightness of the confidence bounds on α and L_{10} reviewed in this thesis. In the original problem description provided by SKF, a data set with six devices is available for the reliability qualification analysis. The duration of the reliability test is 300 hours, the Weibull distribution family assumed for the characterization of the components' reliability and a pre-determined shape parameter set to $\beta_{guess} = 1.5$. Among these six objects, five survived the test and only one failed. Hence, the lifetime of the former is right-censored (Type I censoring) at 300 hours, whilst the latter is assumed to fail at either 50, 100, 150, 200, and 250 hours.

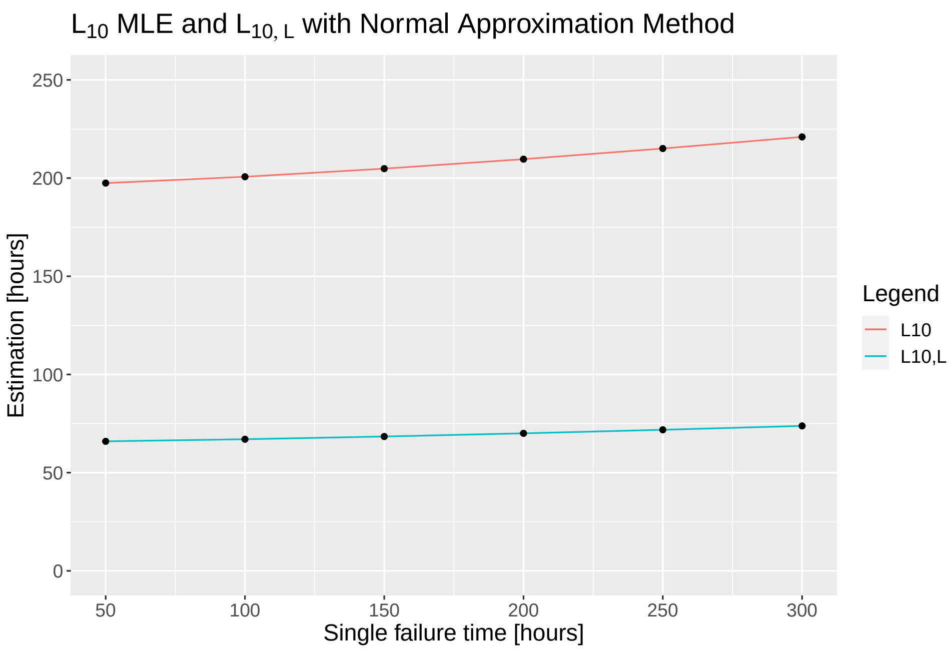

Figure 6.1: ML estimation L_{10} and normal approximation method for$L_{10,L}$ with $\delta = 0.05$ for original problem, for different single failure times, $\beta_{guess} = 1.5$.

Figure 6.1 presents the estimated reliability for the original problem as stated by SKF. The marked blue and red lines present, respectively, the maximum likelihood estimates of the, L_{10}, and the lower bound, $L_{10,L}$. The x-axis shows the failure time of the only failed component. Note that the lower bound on $L_{10,L}$ is computed using the variance of the Fisher information matrix and the normal approximation method and setting a confidence level $1 - \delta = 0.95$, a confidence level which we will use throughout this chapter. Both the estimate and its lower bounds increase for increasing the value of the non-censored failure time. However, one could argue that the estimates, especially the lower bound, are not very sensitive to different failure times.

6.1 Effect of sample size and censoring rate

To investigate this alleged lack of sensitivity, we study estimates for an increasing number of failed components (from 1 to 6) and for increasing sample sizes (from 6 to 12). Figure 6.2 presents the results of the first analysis, where solid lines represent the L_{10}, and the dashed lines the lower bounds. In this figure, we notice that increasing the number of failed components has a more significant effect on L_{10} than on $L_{10,L}$. These results suggest that $L_{10,L}$ might not be so sensitive to the number of failures in this specific setting. To further investigate the sensitivity of the different estimates, we also plot the L_{10} and $L_{10,L}$ for increasing sample sizes. In this setting, we again have just one failure, and all other objects are censored at a time of 300 hours. The results are shown in Figure 6.3. In this setting, we see an increase of data, but still, one failure only marginally changes the $L_{10,L}$.

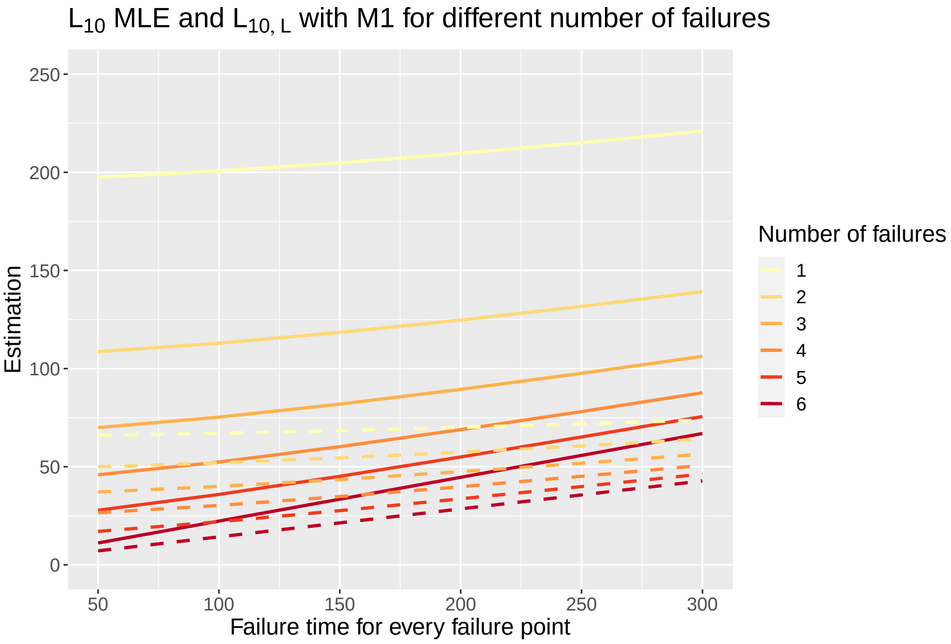

Figure 6.2: Estimates of L_{10} (solid lines)) and the lower bounds $L_{10,L}$ (dashed lines) with $\delta = 0.05$ for different number of failures $r = 1, ..., 6$, $n = 6$.

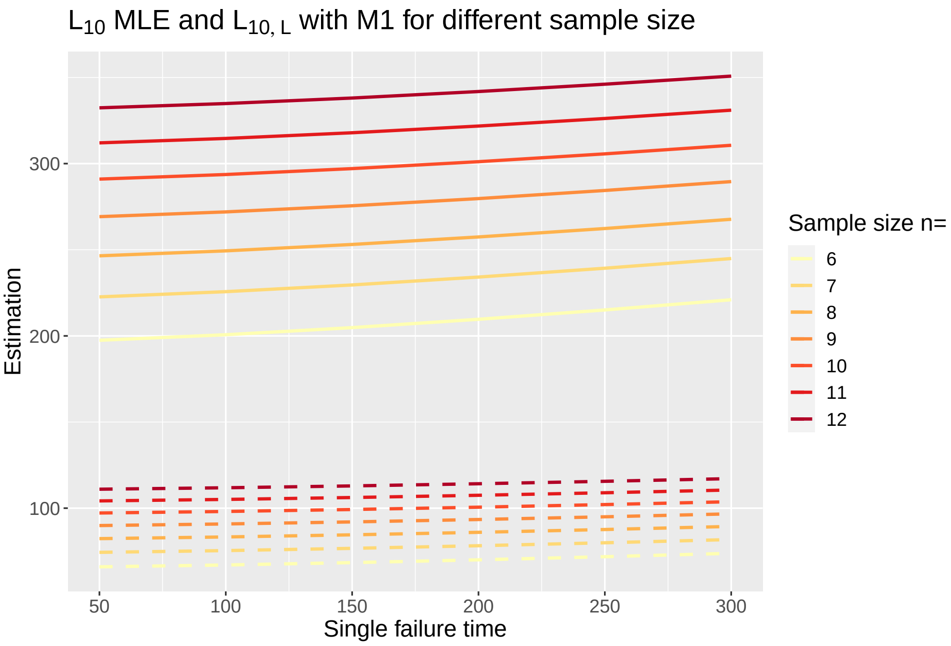

Figure 6.3: Estimates of L_{10} (solid lines) and lower bounds $L_{10,L}$ (dashed lines) with $\delta = 0.05$ for sample sizes $n = 6, .., 12$ and for single failure times with different number of objects.

6.2 Effect of the assumed shape parameter

In practice, when only few data points are available, the shape parameter β is guessed beforehand. This model assumption can lead to poor quality of the estimated α parameter and L_{10} since β is part of the estimates. To investigate the effect of β, we repeated the estimation procedure for several values of $\beta \in [1, 2]$. In Figure 6.4 we see what a variation in β has on the estimated L_{10} in our original problem of $n = 6$ data points with $r = 1$ failure.

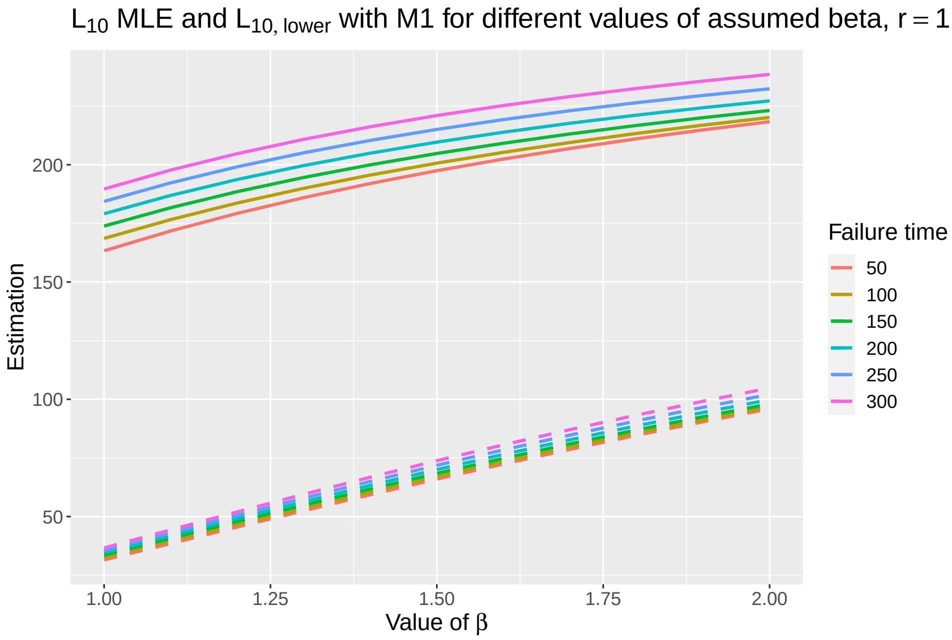

Figure 6.4: Estimates of L_{10} (solid lines) and lower bounds $L_{10,L}$ (dashed lines) with $\delta = 0.05$ for different values of β with single failure times between 50 and 300 hours.

6.3 Bounds comparison for zero failures

In this section, we study the lower bounds for L_{10} when there is data with zero failures. The unavailability of failures is a probable outcome of reliability tests on high-quality and expensive samples given a limited testing time (e.g. due to budget constraints). When this happens, the estimate only depends on the sample size, the censoring time and the assumed value of the shape parameter β, assuming the lifetime data is Weibull distributed. In Table 6.1, we provide the lower confidence bounds on α and L_{10} for different values of a guess shape parameter β and sample sizes. We only present Nelson's method (M2), and the likelihood ratio method (M3) since the normal approximation method (M1) is not directly applicable when there are no failures. See Section 5.6.1 for further details.

Note that, in the results of Table 6.1, the Nelson's estimate for $\hat{\alpha}$ substitute the MLE and is applied to derive the bounds for both M2 and M3. This procedure was necessary to compute the bounds since the MLE for $\hat{\alpha}$ is not computable for zero failure data. The Nelson's estimate $\hat{\alpha}$ for sample sizes $n = [6, 10, 20]$ and $\beta = 1.39$ are [1146.3, 1574.3, 2421.5]. For a higher $\beta = 1.61$, the estimated $\hat{\alpha}$ increases significantly to [1417.2, 2046.7, 3369.9] for the three sample sizes whilst for $\beta = 1.5$ the results is [1264.7, 1777.8, 2822.2].

From the results in Table 6.1 we can observe that the values of the bounds increase when the sample size gets larger, without significant differences between M2 and M3. Increasing the value of β, the lower bound also increases. In general, M2 and M3 behave largely the same when samples increase or different β_{guess} are used. However, the estimates of M3 are higher in general, so M2 is more conservative for zero failure data.

	$L_{10,L}$ for zero failures					
	M2			M3		
data size	$\beta = 1.39$	$\beta = 1.5$	$\beta = 1.61$	$\beta = 1.39$	$\beta = 1.5$	$\beta = 1.61$
n = 6	97.95	106.33	114.14	108.05	116.4	124.23
n = 10	141.45	149.47	156.76	156.0	163.7	170.61
n = 20	232.91	237.27	241.11	256.9	259.85	262.4
	α_L for zero failures					
	M2			M3		
data size	$\beta = 1.39$	$\beta = 1.5$	$\beta = 1.61$	$\beta = 1.39$	$\beta = 1.5$	$\beta = 1.61$
n = 6	494.5	476.6	461.8	545.4	522.0	502.6
n = 10	714.1	670.1	634.3	787.6	733.8	690.3
n = 20	1175.7	1063.6	975.5	1296.9	1164.8	1061.7

Table 6.1: The lower bounds computed with M2 and M3 on L_{10} for data without failures. Different sample sizes are investigated and different assumed values for the shape parameter are used, with a censoring time of 300, $\delta = 0.05$.

6.4 Simulation results

In the following section we will present the results of a simulation study on the three selected methods, namely the normal approximation method (M1), Nelson's method (M2) and the likelihood ratio method (M3). We look at the performance for the different methods for confidence intervals with different censoring rates. Furthermore, we will investigate the implications of under- and overestimating shape parameter β. Moreover we will investigate coverage probabilities for the different methods.

6.4.1 Average of the lower confidence bounds for high censoring rates

In Table 6.2 we present the results of the simulation study for the different confidence bounds for L_{10} and α with different censoring rates and for different values of β with $\delta = 0.05$. For this simulation we simulated 2000 data sets of size 100 for every scenario. The data sets that contained zero failures ($r = 0$), so fully censored data sets, where omitted, since this would give a skewed view on the normal approximation method, which cannot handle zero failure data.

6.4.2 Coverage probabilities

In Table 6.5 the coverage probabilities are presented from the simulation study of the different confidence bounds on L_{10} and α for different censoring rates and different values of β. We computed the coverage probabilities estimates from 2000 independent realization of random data set of size 30 and 100. We see that for all simulation with 30 samples, the coverage probability is very high, higher than the 0.95, which was an expected result since the lower bound computed have been computed for a confidence parameter $\delta = 0.05$, see Section 5.7. For data with 100 samples, we see coverage probabilities closer to 0.95, especially when censoring is lower than 90%. The results suggest that when less data is available (either because of the data set size or because of higher censoring rates), the reviewed methods become more conservative.

For Table 6.5, we can observe that (small) errors in the guessed parameter β do not significantly impact the coverage probabilities. Nevertheless, the average magnitude of the estimated bounds (and therefore its "closeness" to the true estimate) can change substantially for different methods. In other words, the probability that the confidence bounds on the lifetime given by the three are correct is always very high and quasi-identical. On the other hand, the impact of the statement on the minimum lifetime L_{10} (which is the average magnitude of the bound) can be quite different, e.g., for $\beta = 1.61$ and 100 h test, M1 guarantees an L_{10} of at least 140 hours whilst M3 almost 149 hours. Note an inversion in the order of magnitudes which occurs for lower censoring rates, where method M3 gives

$L_{10,true} = 254.16$ [h]		Overestimated shape, β =1.39		
Censoring	T	M1	M2	M3
82 %	400 h	207.57	206.67	200.0
86 %	320 h	195.99	195.02	192.0
90 %	240 h	182.41	181.47	178.93
94 %	170 h	164.05	163.89	162.11
97 %	100 h	126.06	133.29	128.4
$L_{10,true} = 286.21$ [h]		Exact shape, β =1.5		
Censoring	T	M1	M2	M3
84.0%	400 h	221.94	220.92	215.89
88.2%	320 h	211.00	212.14	206.40
92.3 %	240 h	199.04	198.08	194.59
95.3 %	170 h	181.21	181.58	180.75
97.8 %	100 h	133.41	136.36	138.91
$L_{10,true} = 317.09$ [h]		Underestimated shape, β =1.61		
Censoring	T	M1	M2	M3
85.8 %	400 h	238.80	237.70	232.36
89.8 %	320 h	233.01	231.81	227.61
93.4 %	240 h	220.07	219.35	216.36
96.2 %	170 h	196.3	197.9	198.46
98.3 %	100 h	140.8	145.1	148.9
$\alpha_{true} = 1283$		Overestimated shape, β =1.39		
Censoring	T	M1	M2	M3
82 %	400 h	930.50	926.46	896.6
86 %	320 h	878.61	874.26	860.8
90 %	240 h	817.71	813.49	802.1
94 %	170 h	735.88	734.68	726.7
97 %	100 h	589.20	597.53	575.8
$\alpha_{true} = 1283$		Exact shape, β =1.5		
Censoring	T	M1	M2	M3
84.0%	400 h	1002.0	997.38	967.81
88.2%	320 h	955.93	951.01	925.26
92.3 %	240 h	898.79	894.58	872.30
95.3 %	170 h	807.98	809.43	810.26
97.8 %	100 h	599.95	613.59	622.73
$\alpha_{true} = 1283$		Underestimated shape, β =1.61		
Censoring	T	M1	M2	M3
85.8 %	400 h	1078.38	1073.12	1038.29
89.8 %	320 h	1038.49	1033.06	1020.35
93.4 %	240 h	979.91	981.63	961.76
96.2 %	170 h	860.98	905.24	884.56
98.3 %	100 h	631.44	650.62	668.52

Table 6.2: Average values for the lower bounds on L_{10} and α computed according to M1, M2 and M3 and neglecting $r = 0$ scenarios, $\beta_{true} = 1.5$. Different censoring rates are investigated and two cases where the shape parameter is under and over estimated, respectively. Note that the values reported in the table are averages over 2000 randomized data set of size 100.

shorter guarantees on the lifetime compared to M1 and M2, e.g., for $\beta = 1.61$ and 400 h test, M3 ensures an L_{10} of at least 232 hours whilst M1 and M2 are closer to 238 hours.

		Generative model with β =1.39 α =1283.					
		100 samples			30 samples		
Censoring	T	M1	M2	M3	M1	M2	M3
82 %	400 h	0.94	0.94	0.96	0.98	0.98	0.97
86 %	320 h	0.95	0.96	0.96	1.00	1.00	0.97
90 %	240 h	0.96	0.96	0.96	1.00	1.00	1.00
94 %	170 h	0.98	0.98	0.98	1.00	1.00	1.00
97 %	100 h	1.00	1.00	1.00	1.00	1.00	1.00

Table 6.3: Coverage probability of the lower bound on L_{10}, where $P_{c,i}$ is estimated for Models M_i with $i = 1, .., 3$. We consider the cases when the shape parameter β is overestimated ($\beta = 1.5$ when the true value is 1.39). We present two set of results, for samples sizes of 100 and 30. The probabilities are estimated for 5×10^3 data sets drawn from the same reliability distribution but for increasing censoring rates, $\delta = 0.05$.

		Generative model with β =1.61 α =1283.					
		100 samples			30 samples		
Censoring	T	M1	M2	M3	M1	M2	M3
85.8 %	400 h	0.96	0.97	0.98	1.0	1.0	0.97
89.8 %	320 h	0.96	0.97	0.97	1.0	1.0	1.0
93.4 %	240 h	0.96	0.96	0.96	1.0	1.0	1.0
96.2 %	170 h	1.0	1.0	1.0	1.0	1.0	1.0
98.3 %	100 h	1.0	1.0	1.0	1.0	1.0	1.0

Table 6.4: The coverage probability, $P_{c,i}$ of the lower bound on L_{10} for Models M_i with $i = 1, .., 3$ and under estimated β (the true value is 1.61). We present two set of results, for samples sizes of 100 and 30. The probabilities are estimated for 5×10^3 data sets drawn from the same reliability distribution but for increasing censoring rates, $\delta = 0.05$.

		Generative model with β =1.5 α =1283.					
		100 samples			30 samples		
Censoring	T	M1	M2	M3	M1	M2	M3
84.0%	400 h	0.96	0.96	0.96	0.98	0.96	0.96
88.2%	320 h	0.95	0.95	0.97	1.0	1.0	1.0
92.3 %	240 h	0.96	0.96	0.96	1.0	1.0	1.0
95.3 %	170 h	1.0	0.96	0.96	1.0	1.0	1.0
97.8 %	100 h	1.0	1.0	1.0	1.0	1.0	1.0

Table 6.5: The coverage probability, $P_{c,i}$ of the lower bound on L_{10} for Models M_i with $i = 1, .., 3$ for an exact guess of the shape parameter, $\beta_{guess} = \beta$. We present two set of results, for samples sizes of 100 and 30. The probabilities are estimated for 5×10^3 data sets drawn from the same reliability distribution but for increasing censoring rates, $\delta = 0.05$.

Chapter 7

Conclusions

In this chapter we will present the conclusions following the literature review and analysis from the simulation study as performed in this report. In Section 7.2 we present some possibilities for future research on the topic of this report.

7.1 Conclusions

In this thesis, we investigated reliability acceptance tests and methods used in the industry to estimate the reliability of new products and systems. We analyzed statistical confidence bounds for data sets having small sizes and few or no failures and studied their applicability to no-failure situations, their coverage probability, and average tightness. We considered both lower bounds on service life and reliability model parameters. For synthesis' sake, we restricted our analysis to Weibull lifetime models, type I censored data, and scenarios characterized by high censoring rates and low sample sizes. Note that the reviewed bounds only apply to the formerly mentioned scenarios and do not generalize to other distribution families. Nevertheless, lifetime data and small sample sizes are not uncommon in the engineering and manufacturing industry, where tests on new products can be expensive and seldom lead to failures (because of highly reliable components). Moreover, Weibull reliability models are prevalently used in practice, making this investigation of immediate importance for practitioners in the field of reliability qualification and reliability acceptance testing.

We started this work with a concise but comprehensive mathematical introduction to reliability and survival analysis. Then, we presented a literature review of some of the most commonly adopted confidence bounds for reliability qualification. We showed uniqueness for MLE for Weibull distributed data with right-censoring. Then, we focused on three specific bounds: (1) the normal approximate bounds, (2) Nelson's confidence bounds, and (3) bounds derived through the likelihood ratio methods. These bounds are well-known in the reliability, and this work focused on investigating their advantages, shortcomings and applicability to high-censoring and low sample size scenarios. Issues related to sample size and censoring might arise due to the specific set-up of the experiment and underlying (and unknown) failure-generating mechanism. Besides these issues, we also focused on model errors and, specifically, errors arising from a wrong guess of the shape parameter.

To investigate the coverage probability of the three bounds, we designed a simulation model where we generate data, estimate the parameters α and L_{10} and their lower confidence bounds and compute the coverage probability for different underlying parameters and sample sizes. A separate data study for zero failure data is done for Nelson's method and the likelihood ratio method.

The results suggest that higher censoring rates give more conservative confidence bounds, for all the investigated methods. Increasing the sample size in general would improve the bounds. In the comparison of the three methods, we see that for higher censored cases the Likelihood Ratio method provides the highest lower bounds, whereas the normal approximation method gives higher lower bounds for lower censoring cases. In the results for the zero failure data, we see that Nelson's method provides more conservative lower confidence bounds in general. Increasing the sample size with data with zero failures greatly increases the confidence bound.

Another effect we see in the results is that under estimated shape parameter leads to a higher coverage probability, leading to conservative confidence bounds and vice versa. Note that when estimating lifetime, conservative confidence bounds can be undesirable with for example expensive components.

7.2 Further research

In addition to the studies on the different methods for the confidence bounds, a study on the quality of different estimates for the parameters of the Weibull distribution for data sets with zero failures could give more insight, since the methods depend on the quality of the estimate. For example the alternative MLE as suggested by Jiang et al. (2010) could be compared to the median estimate as suggested by Nelson. To extend this simulation study, more bounds could be included that are assumed to behave nicely for few failure data. As described by Magalhaes and Gallardo (2020), it might prove beneficial to correct the likelihood ratio method. Research in the performance of this correction are necessary especially in the case of few or zero failures.

Bibliography

L.J. Bain and M. Engelhardt. *Statistical Analysis of Reliability and Life-Testing Models.* Marcel Dekker, New York, 1991.

R. Brookmeyer and J. Crowley. A confidence interval for the median survival time. *Biometrics*, pages 29–41, 1982.

A.C. Cohen. Maximum Likelihood estimation in the Weibull distribution based on complete and on censored samples. *Technometrics*, 7(4):579–588, 1965.

A. Di Bucchianico and R.M. Castro. *Lecture notes for 2MMS90 Sequential and Nonparametric Statistics.* Eindhoven University of Technology, 2021.

T.J. Diciccio and J.P. Romano. A review of bootstrap confidence intervals. *Journal of the Royal Statistical Society: Series B (Methodological)*, 50(3):338–354, 1988.

N. Doganaksoy and J. Schmee. Comparisons of approximate confidence intervals for distributions used in life-data analysis. *Technometrics*, 35(2):175–184, 1993.

T.-H. Fan and C.-C. Chang. A Bayesian zero-failure reliability demonstration test of high quality electro-explosive devices. *Quality and Reliability Engineering International*, 25(8):913–920, 2009.

N. R. Farnum and P. Booth. Uniqueness of maximum likelihood estimators of the 2-parameter Weibull distribution. *IEEE Transactions on Reliability*, 46(4):523–525, 1997.

U. Genschel and W.Q. Meeker. A comparison of maximum likelihood and median-rank regression for Weibull estimation. *Quality engineering*, 22(4):236–255, 2010.

L. Han, P. Jiang, Y. Yu, and B. Guo. Bayesian reliability evaluation for customized products with zero-failure data under small sample size. In *2014 10th International Conference on Reliability, Maintainability and Safety (ICRMS)*, pages 904–907. IEEE, 2014.

H. Hirose. Bias correction for the maximum likelihood estimates in the two-parameter Weibull distribution. *IEEE Transactions on Dielectrics and Electrical Insulation*, 6(1):66–68, 1999.

Z. Huang and A.A.. Porter. Lower bound on reliability for Weibull distribution when shape parameter is not estimated accurately. *IEEE Transactions on Reliability*, 1(26), 1990.

S. Jeng and W.Q. Meeker. Comparisons of approximate confidence interval procedures for type I censored data. *Technometrics*, 42(2):135–148, 2000.

P. Jiang, J.-H. Lim, M.J. Zuo, and B. Guo. Reliability estimation in a Weibull lifetime distribution with zero-failure field data. *Quality and Reliability Engineering International*, 26(7):691–701, 2010.

P. Jiang, Y. Xing, X. Jia, and B. Guo. Weibull failure probability estimation based on zero-failure data. *Mathematical Problems in Engineering*, 2015, 2015.

P. Kabaila and H. Leeb. On the large-sample minimal coverage probability of confidence intervals after model selection. *Journal of the American Statistical Association*, 101(474):619–629, 2006.

J.B. Keats, P.C. Nahar, and K.M. Korbel. A study of the effects of mis-specification of the Weibull shape parameter on confidence bounds based on the Weibull-to-exponential transformation. *Quality and Reliability Engineering International*, 16(1):27–31, 2000.

D.K. LLoyd and M. Lipow. *Reliability: Management, Methods and Mathematics.* The American Society for Quality Control, 1977.

T.M. Magalhaes and D.I. Gallardo. Bartlett and Bartlett-type corrections for censored data from a Weibull distribution. *SORT-Statistics and Operations Research Transactions*, 44(1):127–140, 2020.

J.I. McCool. Inference on Weibull percentiles and shape parameter from maximum likelihood estimates. *IEEE Transactions on Reliability*, R-19(2-9):523–525, 1970a.

J.I. McCool. Evaluating Weibull endurance data by the method of maximum likelihood. *ASLE transactions*, 13(3):189–202, 1970b.

W.Q. Meeker and L.A. Escobar. *Statistical Methods for Reliability Data.* John Wiley & Sons, New York, 1998.

W.Q. Meeker and W. Nelson. Weibull percentile estimates and confidence limits from singly censored data by maximum likelihood. *IEEE Transactions on Reliability*, 25(1):20–24, 1976.

W. Nelson. Weibull analysis of reliability data with few or no failures. *Journal of Quality Technology*, 17(3):140–146, 1985.

H. Rinne. *The Weibull Distribution: A Handbook.* CRC Press, Boca Raton, Florida, 2008.

Y. Shen and Z. Yang. Bias-correction for Weibull common shape estimation. *Journal of Statistical Computation and simulation*, 85(15):3017–3046, 2015.

M. Xie, Z. Yang, and O. Gaudoin. More on the mis-specification of the shape parameter with Weibull-to-exponential transformation. *Quality and Reliability Engineering International*, 16(4):281–290, 2000.

Z. Yang and M. Xie. Efficient estimation of the Weibull shape parameter based on a modified profile likelihood. *Journal of Statistical Computation and Simulation*, 73(2):115–123, 2003.

X. Yuan and S.N. Rai. Confidence intervals for survival probabilities: A comparison study. *Communications in Statistics—Simulation and Computation*, 40(7):978–991, 2011.

C.W. Zhang. Weibull parameter estimation and reliability analysis with zero-failure data from high-quality products. *Reliability Engineering & System Safety*, 207:107321, 2021.

L.F. Zhang, M. Xie, and L.C. Tang. Bias correction for the least squares estimator of Weibull shape parameter with complete and censored data. *Reliability Engineering & System Safety*, 91(8):930–939, 2006.