Reliability in fatigue

On the choice of distributions in the load-strength model

Tomas Torstensson

9th January 2004

Reliability in fatigue

On the choice of distributions in the load-strength model

Abstract

In this thesis the influence of the choice of distributions in the load-strength model is considered. Accurate predictions of the failure probability is very useful when aiming at the most cost effective design of a component. Two distributions for load and strength are evaluated, the lognormal distribution and the Weibull distribution. From the load-strength model the failure probability can be determined which is the probability that the component in question fails within a specific time. The main conclusion is that the lognormal distribution should be used rather than the Weibull distribution, especially when the data available is limited. A possible way of updating the model with observed failure rates using Bayesian methods is also suggested.

Tillförlitlighet inom utmattning

Val av fördelningar i last-styrka-modellen

Sammanfattning

I detta exjobb undersöks vilken påverkan olika fördelningsval har på last-styrka-modellen. Noggranna förutsägelser av felsannolikheten är mycket användbara för kostnadseffektiv dimensionering av komponenter. Två fördelningar för lasten och styrkan studeras, lognormalfördelningen och Weibullfördelningen. I last-styrka-modellen kan felsannolikheten beräknas, d.v.s. sannolikheten att den aktuella komponenten går sönder inom en viss tid. Huvudslutsatsen är att lognormalfördelningen bör användas snarare än Weibullfördelningen, i synnerhet vid begränsad tillgång på data. Ett möjligt sätt att uppdatera modellen med felutfall med hjälp av Bayesianska metoder föreslås också.

Acknowledgments

I would like to thank my supervisors Pär Johannesson, Jacques de Maré and Thomas Svensson for their extensive counseling and encouragement. The frequent meetings have made it possible to get feedback and advice in order to make progress with the project. I wish to thank my examiner Gunnar Englund. I am also grateful to Bengt Johannesson at Volvo Trucks and Bertil Jonsson at Volvo Articulated Haulers. I had access to their reports and measurement data and my Master's thesis could not have been carried out without that material. I wish to thank Fraunhofer-Chalmers Research Centre for giving me this opportunity to do my Master's thesis and it has been a pleasure to work here. A special thanks to Marlo who reviewed the manuscript. Finally I would like to thank Ebba for her love and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The load-strength model is a tool for reliability analysis in fatigue.The damage that an external cyclic load causes to a material is called the fatigue of the material. The load itself can be characterized by its local maxima and minima. A good way to examine the load is the rainflow count method, see Johannesson [3], which gives the load amplitudes. Some definitions are needed in order to describe the model

$$\begin{aligned} C &= \text{Capacity (strength)}\,, \\ D &= \text{Duty (load)}\,. \end{aligned}$$

The capacity can, e.g., represent the strength of a vehicle component and the duty is then the load that the component is exposed to. The basis of the model is that failure occurs (the component breaks) if the duty exceeds the capacity, i.e. $D > C$. In lecture notes from a course for Swedish industry, see Johannesson and de Maré [2], the load-strength model and several applications are described. The duty naturally depends on the time or distance the component has been used. The idea is that both capacity, C, and duty, D, are modelled as random variables. The scatter in the strength of the components is modelled by C and the scatter in the load is modelled by D. The scatter in the strength is easy to understand (material and manufacturing properties), but the scatter in the load is more complex. It consists of several parts. Duty varies based on how the vehicle is driven, the road conditions, and so on. Therefore it is much harder to find and motivate a suitable model for the duty than the capacity. The failure probability is the probability that failure occurs

$$P_f \;=\; P(D > C)\,.$$

The model can be used in several ways. One way of using the model is in the case where there is a restriction on the failure probability for a critical component. The objective could also be to find the minimum life cycle cost. The total cost is a sum of the manufacturing and operation cost. When the failure probability decreases (stronger component) the manufacturing cost increases and for the operation cost the relationship is reversed. This means that an optimal design failure probability can be found which minimizes the life cycle cost for the component. Then the component can be adjusted by changing the manufacturing procedure in order to satisfy that condition.

The load-strength model have been used by PSA Peugeot Citroën, see Thomas et al. [7], and Volvo Construction Equipment, see Olsson [5] and Samuelsson [6]. They use different assumptions for the distributions of the capacity and load. Volvo Construction Equipment uses a model in which capacity and duty is based on the three parameter Weibull distribution. PSA Peugeot Citroën models both capacity and duty with the normal distribution.

In Chapter 2 the background of the load-strength model is described and the model and definitions are introduced. The use of the load-strength model in industry is described in Chapter 3. In Chapter 4 the estimation of parameters in the lognormal and Weibull distribution are examined. A couple of methods for estimation of the parameters in a three parameter Weibull distribution are examined. The basis of the estimation is data from Volvo Trucks. Also the quantitative differences that depends on the choice of distribution are studied. The model of Volvo Articulated Haulers is analyzed in Chapter 5. In Chapter 6 feedback using Bayesian methods is examined. Finally in Chapter 7 conclusions are drawn and proposals for how to use the load-strength model in the future are suggested.

Chapter 2

Background

The capacity, C, and the duty, D, are assumed to be continuous random variables. A continuous random variable X is defined by its density function, $f_X(x)$, and its distribution function, $F_X(x)$.

$$\begin{aligned} P(a \leq X \leq b) &= \int_a^b f_X(x)\,dx = F_X(b) - F_X(a) \\ f_X(x) &= \frac{d}{dx} F_X(x) \end{aligned}$$

In this thesis principally two distributions will be considered, the lognormal distribution and the three parameter Weibull distribution (used by Volvo). The reason why these distributions are used is discussed in Section 2.6. In some examples in this chapter the normal distribution (used by PSA Peugeot Citroën) will also be considered.

2.1 Normal distribution

If $X \sim N(\mu, \sigma^2)$ the density function is

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), \quad -\infty < x < \infty\,.$$

The distribution function can not be determined explicitly

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\,dy\,.$$

If $X \sim N(0,1)$ (standard normal) then the distribution function is denoted by $\Phi(x)$.

2.2 Lognormal distribution

If $X \sim LN(\mu, \sigma^2)$ it means that $\log X \sim N(\mu, \sigma^2)$, where log is the natural logarithm. The density function is

$$f_X(x) = \frac{1}{x\sqrt{\sigma^2}\sqrt{2\pi}} \exp\left(-\frac{(\log x - \mu)^2}{2\sigma^2}\right), \quad x > 0\,.$$

Also for this distribution the distribution function can not be expressed explicitly

$$F_X(x) = \int_0^x f_X(y)\,dy\,.$$

2.3 Three parameter Weibull distribution

If $X \sim W(\beta, \eta, \gamma)$ then the density function is

$$f_X(x) = \frac{\beta}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right), \quad \beta, \eta > 0,\ x > \gamma\,.$$

For this distribution there is a explicit expression for the distribution function

$$F_X(x) = \int_\gamma^x f_X(y)\,dy = 1 - \exp\left(-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right)\,.$$

In case $\gamma = 0$ the distribution is called a two parameter Weibull.

2.4 A target customer

The $100p\%$ customer z_p is defined as a quantile in the duty distribution, see Johannesson and de Maré [2]

$$\begin{aligned} P(D \leq z_p) &= p \\ P(D > z_p) &= 1 - p \end{aligned}$$

where $1 - p$ is the probability of finding a customer more extreme than z_p. How to calculate an extreme customer depends of course on the distribution of the duty.

2.4.1 Duty based on normal distribution

Assume that the duty is normally distributed with mean value $m_D = 300$ MPa and standard deviation $\sigma_D = 60$ MPa. The 90% customer is sought, and $p = 0.90$ yields

$$z_{0.90} = m_D + \lambda_{0.10}\,\sigma_D = m_D + 1.28\,\sigma_D = 377 \text{ MPa}.$$

2.4.2 Duty based on lognormal distribution

If the duty is lognormal the $100p\%$ customer is found in the following way

$$\begin{aligned} P(D \le z_p) &= P(\log D \le \log z_p) = P\left(\frac{\log D - \mu}{\sigma} \le \frac{\log z_p - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{\log z_p - \mu}{\sigma}\right) = p \end{aligned}$$

which gives

$$\frac{\log z_p - \mu}{\sigma} = \lambda_p$$
$$z_p = e^{\mu + \lambda_p\,\sigma}\,.$$

Assuming that the mean and variance is the same as in the example with the normal distribution yields $z_{0.90} = 379$ MPa.

2.4.3 Duty based on Weibull distribution

Assume that the duty is inversely proportional to a Weibull distributed random variable, i.e. $D = 1/Y$ where $Y \sim W(\beta, \eta, 0)$.

$$F_Y(y) = 1 - \exp\left(-\left(\frac{y}{\eta}\right)^{\beta}\right), \quad y \ge 0, \quad \eta, \beta > 0\,.$$

The $100p\%$ customer can be determined directly from the definition

$$\begin{aligned} P(D \le z_p) &= P\left(\frac{1}{Y} \le z_p\right) = P\left(Y \ge \frac{1}{z_p}\right) \\ &= 1 - F_Y\left(\frac{1}{z_p}\right) = \exp\left(-\left(\frac{1}{\eta\,z_p}\right)^{\beta}\right) = p\,. \end{aligned}$$

Extracting z_p gives the explicit expression

$$z_p = \frac{1}{\eta\,(-\log p)^{1/\beta}}\,.$$

If the mean value and the standard deviation is the same as in the example with the normal distribution, then $z_{0.90} = 372$ MPa.

2.5 Failure probability

There are different kinds of failure probabilities. One of them is the probability that a failure occurs considering the entire population. Another is the probability that failure occurs for the $100p\%$ customer.

2.5.1 Entire population

The probability that a failure occurs for the entire population is denoted by P_f and is calculated as

$$P_f = P(D > C).$$

2.5.2 The target customer

Given the $100p\%$ customer the duty $D = z_p$ which implies that

$$P_{f,p} = P(z_p > C).$$

This means that once the $100p\%$ customer is known only the distribution of the capacity is needed.

In case that the capacity is normally distributed it is rather simple to calculate the failure probability. Let $z_{0.90} = 400$ MPa and assume that $C \sim N(500, 50^2)$. The failure probability for the target customer is then

$$\begin{aligned} P_f &= P(C < z_p) = P\left(\frac{C - m_C}{\sigma_C} < \frac{z_p - m_C}{\sigma_C}\right) = P\left(Z < \frac{400 - 500}{50}\right) \\ &= P(Z < -2), \quad \text{where } Z \sim N(0,1) \\ &= \Phi(-2) = 1 - \Phi(2) = 1 - 0.977 = 0.023 = 2.3\%. \end{aligned}$$

2.6 Distributions for duty and capacity

Until now some different distributions have been assumed for duty and capacity. The choice in a specific case depends on the application. There are three different models that are natural in the fatigue context.

1.	D=normal	C=normal
2.	$1/D$=Weibull	C=Weibull
3.	D=lognormal	C=lognormal

1. PSA Peugeot Citroën uses a model where both duty and capacity are normally distributed. The capacity is interpreted as the fatigue limit which explains the assumption of normal distribution.

2. Volvo Construction Equipment uses a Weibull distribution with three parameters. The choice of the Weibull distribution is justified by the ”weakest link” principle. The Weibull distribution also fits well to observations but does not give explicit formulas to calculate the failure probability.

3. The capacity is often assumed to be lognormally distributed which can be justified by analyzing the failures in the SN-graph. If the duty can also be assumed to be lognormally distributed, it gives explicit formulas for the failure probability.

2.7 Models for duty and capacity

According to the method that Volvo Construction Equipment uses the damage of a component that is accumulated over time must be transformed to a scalar value and one way to do that is to use the Palmgren-Miner hypothesis for accumulated fatigue damage

$$d = \sum_{i=1}^{M} \frac{1}{N(\Delta S_i)}$$

where d is the damage. The function that describes the number of cycles to failure for the component in interest is denoted by $N(\cdot)$. The different load amplitudes are ΔS_i, $i = 1, \ldots, M$. Furthermore damage 1 corresponds to failure.

Basquins equation is frequently used to describe $N(\cdot)$

$$N = C\,(\Delta S)^{-k}$$

where k is the Wöhler exponent. The capacity, C, is assumed to be a random variable. Now the damage can be written as

$$d = \sum_{i=1}^{M} \frac{(\Delta S_i)^k}{C} = \frac{D}{C}$$

and thus defines the duty as

$$D = \sum_{i=1}^{M} (\Delta S_i)^k \,.$$

Note that failure occurs when $d > 1$ which is consistent with the previous definition of failure $D > C$.

2.7.1 Estimation of capacity

By performing experiments with varying load amplitudes ΔS_i, a sequence of a total of M_C load cycles before the component breaks is received. Since failure corresponds to $d = 1$, C can be extracted

$$
\begin{aligned}
1 &= \sum_{i=1}^{M_C} \frac{1}{C\,\Delta S_i^{-k}} \\
C &= \sum_{i=1}^{M_C} \Delta S_i^k \,.
\end{aligned}
$$

Both the mean and the standard deviation of C can be estimated by repeating the experiment a number of times.

2.7.2 Estimation of duty

The duty, D, is estimated from a load process and an observation is determined by the formula

$$D = \sum_{j=1}^{M_D} \Delta S_j^k$$

where M_D is the total number of load cycles. The mean and the standard deviation of D can be estimated from load measurements on different customers.

2.8 Applications

In a typical application only the relation between C and D is of interest. Furthermore, a value for the Wöhler exponent is chosen, e.g. Volvo Construction Equipment (VCE) often uses $k = 3$ since the components are welded. In that case C and D are determined from the expressions

$$C_{VCE} = \sum_{i=1}^{M_C} \Delta S_i^3 \quad , \quad D_{VCE} = \sum_{j=1}^{M_D} \Delta S_j^3 .$$

2.9 Duty intensity

Since the duty is accumulated linearly it is possible to determine the duty intensity. If e.g. observations of the duty are given but they correspond to different distances they must obviously be normalized in some way. The observations are transformed to the duty intensity, $\tilde{D}$ [duty/km], to make it possible to use them for further inference. The duty intensity is calculated as the duty divided by the driven distance. The duty at a certain distance, s, is just $D = s \cdot \tilde{D}$. If the duty is observed at different times the duty intensity have the unit [duty/h] . Then the duty at a certain time, t, is calculated as $D = t \cdot \tilde{D}$. The duty intensity will be explained more carefully in the context in which it is used in the thesis.

Chapter 3

Applications in industry

3.1 Volvo Trucks

Volvo Trucks has made extensive trials in order to predict the life distribution by means of the load strength model.

3.1.1 Load-strength model

1. *Analysis of fatigue data regarding strength.*
   The analysis includes a description of fatigue data variation and fatigue modelling by the life distribution. The capacity, C, is calculated from rig tests and is determined by the formula

   $$C = \sum_i n_i \left(\Delta S_i\right)^k$$

   where the n_i is the number of cycles until a failure occurs at the load level ΔS_i and $\sum\limits_i n_i$ is the total number of cycles until failure occurs.

2. *Analysis of operational loading.*
   The accumulated duty, D, is calculated from the load spectrum and is determined by the formula

   $$D = \sum_i n_i \left(\Delta S_i\right)^k.$$

   Duty values are divided by the distance in km to get a comparable unit [duty/km], i.e. the duty intensity $\tilde{D}$. The Wöhler exponent $k = 4$ are used in the calculations which differs from the value $k = 3$ that Volvo Construction Equipment uses. That value is justified from examination of the Wöhler curve.

3. *Life predictions and modelling.*
   The values for duty/km and capacity are examined and fitted inference is based on a three parameter Weibull distribution. Then the life can be predicted by means of simulation techniques. There are two more assumptions except those already stated.

   - The duty and the capacity are assumed to be independent. Duty is actively accumulated until it reaches capacity (failure occurrence).
   - The duty of the component is based on the lateral loading only, even though loading may exist in vertical and longitudinal direction as well.

3.1.2 Load-strength simulations

In order to estimate the duty, a number of field measurements have been performed at different markets, e.g Sweden, Norway, Germany and Brazil. The distribution of the capacity has been examined from rig tests and a three parameter Weibull inference has been made. The results are shown in diagrams with “Level crossing spectrum”, “Range spectrum” and “Power spectrum”. Some other types of diagrams are plotted too, e.g. accumulated failure rate against driven distance and accumulated driven distance against time [year].

3.2 Volvo Articulated Haulers

Bertil Jonsson at Volvo Articulated Haulers has written a report in which he uses the load strength model in a slightly different way. The load is estimated by driving for an hour, and the strength is estimated from the Paris equation.

3.2.1 Estimation of capacity

In order to estimate the capacity the Paris equation is used

$$\frac{da}{dN} = B\,(\Delta K)^k$$

where a is the crack length. The number of cycles is denoted by N. The crack growth rate is $\frac{da}{dN}$. The random variable B corresponds to material properties and crack geometry of the component. The range of the stress intensity factor is denoted by ΔK and $k = 3$ is a constant. Later it will turn out that it is equal to the Wöhler exponent.

If the load amplitude is denoted by ΔS, and using the fact that

$$\Delta K = \Delta S \cdot f(a)$$

where $f(\cdot)$ is a function that is determined by measuring the stress intensity factor for different crack lengths, the expression can be written in the form

$$\frac{da}{dN} = B \cdot (\Delta S \cdot f(a))^k.$$

Rewriting and integration the expression above yields

$$\begin{aligned} \int_{a_0}^{a_c} \frac{da}{B\,(\Delta S)^k\, f(a)^k} &= \int_0^{N_c} dN \\ \frac{1}{B} \int_{a_0}^{a_c} \frac{da}{f(a)^k} &= N_c\,(\Delta S)^k \end{aligned}$$

where a_0 is the initial crack length and a_c (20 mm here) is the crack length defined as failure. The number of cycles until failure is denoted by N_C. Extracting N_c gives

$$N_c = \frac{1}{B} \int_{A_0}^{a_c} \frac{da}{f(a)^k} \cdot (\Delta S)^{-k}$$

where A_0 is with capital letter since it is considered to be a random variable. Comparing this expression to the other one for N

$$N = C\,(\Delta S)^{-k}$$

implies that C can be identified and written as

$$C = g(B, A_0) = \frac{1}{B} \int_{A_0}^{a_c} \frac{da}{f(a)^k}$$

i.e. a function $g(\cdot)$ of the random variables B and A_0.

The entities B and A_0 are considered to be normally distributed and experiments have been made in order to estimate mean and standard deviation. The distribution of C can then be estimated by making computer simulations. Volvo Articulated Haulers then fits a three parameter Weibull distribution to these capacity values.

3.2.2 Estimation of duty

The duty is determined by driving in different ways (normal and forced driving) for one hour and after that making an assumption of how common the different types of driving styles are. Once that is made an estimation of the distribution of the duty intensity, $\tilde{D}$, can be determined. In these calculations the Wöhler exponent $k = 3$ which is a value that is often chosen in this context. The duty is assumed to be based on a three parameter Weibull distribution which is fitted to the observations.

Chapter 4

Inference based on data from Volvo Trucks

In this section estimation of parameters in the lognormal and the Weibull distribution will be examined. The basis of the examination will be the data from Volvo Trucks. The different estimation methods have been implemented in Matlab. All numerical calculations in this thesis have been carried out through the use of Matlab.

4.1 The lognormal distribution

Assume that the random variables X_1, X_2, ..., X_n $\sim$ i.i.d. (independent and identically distributed) $LN(\mu, \sigma^2)$. Let $\boldsymbol{\theta} = (\theta_1, \theta_2) = (\mu, \sigma^2)$. The density function for X_i will then be

$$f_X(x, \boldsymbol{\theta}) = \frac{1}{x\sqrt{\theta_2}\sqrt{2\pi}} \exp\left(-\frac{(\log x - \theta_1)^2}{2\theta_2}\right), \quad x > 0\,.$$

Given that x_1, x_2, ..., x_n is a random sample from X_1, X_2, ..., X_n, the parameter vector $\boldsymbol{\theta}$ can be estimated. Since $\log X_i$ is normally distributed the MLE (maximum likelihood estimate) of $\boldsymbol{\theta}$ will be the same as for the normal distribution, but with the logarithm of the observations, i.e.

$$\begin{aligned} \theta_1^* &= \frac{1}{n}\sum_{i=1}^{n} \log x_i\,, \\ \hat{\theta}_2 &= \frac{1}{n}\sum_{i=1}^{n} (\log x_i - \theta_1^*)^2\,. \end{aligned}$$

The estimate $\hat{\theta}_2$ is slightly biased, but the estimate

$$\theta_2^* = \frac{1}{n-1}\sum_{i=1}^{n}(\log x_i - \theta_1^*)^2$$

is unbiased. The unbiased estimate, θ_2^*, is used further on in this thesis.

4.2 The Weibull distribution

Assume that the random variables $X_1,\ X_2,\ \ldots,\ X_n\ \sim$ i.i.d. $\mathrm{W}(\beta, \eta, \gamma)$ are given. $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3) = (\beta, \eta, \gamma)$. The density function for X_i is

$$f_X(x, \boldsymbol{\theta}) = \frac{\beta}{\eta}\left(\frac{x-\gamma}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right), \quad \beta, \eta > 0,\ x > \gamma$$

where β is the shape parameter, η is the scale parameter, and γ is the location parameter. Note that γ is a threshold. Given a sample $x_1, x_2, \ldots, x_n$ from $X_1, X_2, \ldots, X_n$ the task is to find an estimator of $\boldsymbol{\theta}$. The likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f_X(x_i) = \left(\frac{\beta}{\eta}\right)^n \prod_{i=1}^{n}\left(\frac{x_i-\gamma}{\eta}\right)^{\beta-1} \prod_{i=1}^{n} \exp\left(-\left(\frac{x_i-\gamma}{\eta}\right)^{\beta}\right)$$

Normally L or $\log L$ is maximized in order to find the MLE of $\boldsymbol{\theta}$, but the problem with the three parameter Weibull distribution is that it has a threshold which means that $L(\boldsymbol{\theta})$ can be singular due to the factor

$$(x_k - \gamma)^{\beta-1} \quad \text{where}\ \ x_k = \min_{1\le i\le n} x_i$$

Since $\gamma > x_i\ \ \forall i$ it is only the factor with the minimum x_i that has to be examined. The expression can be examined by letting γ tend to x_k

$$\begin{aligned}
\beta < 1: \quad & \lim_{\gamma\to x_k^-}(x_k-\gamma)^{\beta-1} &=& \ \infty \\
\beta = 1: \quad & \lim_{\gamma\to x_k^-}(x_k-\gamma)^{\beta-1} &=& \ 1 \\
\beta > 1: \quad & \lim_{\gamma\to x_k^-}(x_k-\gamma)^{\beta-1} &=& \ 0
\end{aligned}$$

This irregularity causes difficulties in the determination of the parameters and thus it would be an advantage to use other methods to estimate at least one of them. It turns out that it is often most convenient to estimate the threshold γ first.

4.2.1 Estimators based on percentiles

The distribution function of a three parameter Weibull distributed variable is

$$F_X(x) = 1 - \exp\left(-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right), \quad x > \gamma .$$

The $100p\%$ population percentile, x_p is determined from the relationship

$$F_X(x_p) = p$$

which gives

$$\begin{aligned} 1 - \exp\left(-\left(\frac{x_p-\gamma}{\eta}\right)^{\beta}\right) &= p \\ \left(\frac{x-\gamma}{\eta}\right)^{\beta} &= -\log(1-p) \\ x_p &= \gamma + \eta\,(-\log(1-p))^{1/\beta} . \end{aligned}$$

Let $y_1, y_2, \ldots, y_n$ be the ordering of a sample $x_1, x_2, \ldots, x_n$, i.e. $y_i = x_{(i)}$. Then the sample distribution function that corresponds to the ith ordered observation is

$$p_i = \frac{i}{n+1}$$

and then the corresponding $100p_i$ percent sample percentile t_i is given by

$$t_i = y_{\lceil np_i \rceil}$$

where $\lceil x \rceil$ is the smallest integer that is larger than or equal to x.
The idea is that three sample percentiles yield a system of three equations with the three unknown parameters. Let the estimation of the parameters be $\tilde{\beta}$, $\tilde{\eta}$ and $\tilde{\gamma}$. Then the system is given by

$$t_s = \tilde{\gamma} + \tilde{\eta}\,[-\log(1-p_s)]^{1/\tilde{\beta}} = y_{\lceil np_s \rceil} \quad s = i, j, k$$

where $0 < p_i < p_j < p_k < 1$.
If p_j is chosen such that

$$-\log(1-p_j) = \{[-\log(1-p_i)][-\log(1-p_k)]\}^{1/2}$$

the estimation of β will be, see Zanakis [8]

$$\tilde{\beta} = \log\left(\frac{-\log(1-p_k)}{-\log(1-p_i)}\right) \Big/ \log\left(\frac{t_k-\tilde{\gamma}}{t_i-\tilde{\gamma}}\right).$$

where the estimate $\tilde{\gamma}$ also is found with help from the sample percentiles

$$\tilde{\gamma} = \frac{t_i t_k - t_j^2}{t_i + t_k - 2t_j}.$$

If this estimated value exceeds the smallest observed value, y_1, then it is not permissible and $\gamma = y_1$ should be used instead, but the probability that such a situation occurs is very small.
The values for p_i and p_k can be chosen such that the asymptotic variance of the estimator $\tilde{\beta}$ is minimized, see Dubey [1], which yields

$$p_i = 0.16731,$$
$$p_k = 0.97366.$$

A efficient estimator of γ can be found by using the 1st, 2nd and nth ordered observation of the sample. When the estimation of γ is determined it is an easy task to find the estimation of η. Finally the estimators are

$$\begin{aligned}
\tilde{\gamma} &= \frac{y_1 y_n - y_2^2}{y_1 + y_n - 2y_2}, \\
\tilde{\eta} &= -\tilde{\gamma} + y_{\lceil 0.63n \rceil}, \\
\tilde{\beta} &= \log\left(\frac{-\log(1-p_k)}{-\log(1-p_i)}\right) \Big/ \log\left(\frac{t_k - \tilde{\gamma}}{t_i - \tilde{\gamma}}\right).
\end{aligned}$$

The advantage of these estimators are both their simplicity and accuracy, especially when n is small which suits the load-strength application since it often involves small samples. These estimates will be denoted by $\tilde{\beta}$, $\tilde{\eta}$ and $\tilde{\gamma}$.

4.2.2 Estimators based on Maximum Likelihood

Consider the two parameter Weibull distribution, which corresponds to $\gamma = 0$. The distribution function is

$$F_X(x) = 1 - \exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right), \quad x > 0.$$

Let $X \sim W(\beta, \eta, 0)$ and $Y \sim W(\beta, \eta, \gamma)$. Assume that γ is known or estimated. Then it is possible to transform the three parameter Weibull distributed random variable into a two parameter such one. The advantage of this approach is that there will be no singularity problem since the threshold has already been estimated.

Let $Z = Y - \gamma$ and examine the distribution function of Z

$$F_Z(z) = P(Z \leq z) = P(Y - \gamma \leq z) = P(Y \leq z + \gamma) = F_Y(z + \gamma) = F_X(z)$$

which implies that Z and X are identically distributed.

Since the MLE can be determined for the two parameter Weibull distribution, a way to estimate the parameters could be to first estimate γ according to the rule

$$\tilde{\gamma} = \frac{y_1 y_n - y_2^2}{y_1 + y_n - 2y_2}$$

and then determine the MLE of β and η by maximizing the likelihood function for the sample values $\{x_i - \tilde{\gamma}\}$, $i = 1, \ldots, n$. The estimates of the parameters according to this method will be denoted by $\hat{\beta}$, $\hat{\eta}$ and $\tilde{\gamma}$.

4.2.3 Estimators based on an iterative procedure

A natural estimation of the location parameter γ is $\gamma = x_{(1)}$. The disadvantage with this estimation is that it will always exceed the true value. It is therefore of interest to examine the expectation of $X_{(1)}$

$$E(X_{(1)}) = \int_{\gamma}^{\infty} x\, f_{X_{(1)}}(x) dx\,.$$

Consequently the distribution of $X_{(1)}$ must be determined

$$\begin{aligned}
F_{X_{(1)}}(x) &= P(X_{(1)} \leq x) = 1 - P(X_{(1)} > x) = 1 - P(X_1 > x, X_2 > x, \ldots, X_n > x) \\
&= \{\text{independent}\} = 1 - P(X_1 > x) \cdot P(X_2 > x) \cdot \ldots \cdot P(X_n > x) \\
&= 1 - (P(X > x))^n = 1 - (1 - F_X(x))^n \\
&= 1 - \left(1 - \left(1 - \exp\left(-\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right)\right)\right)^n \\
&= 1 - \exp\left(-n\left(\frac{x-\gamma}{\eta}\right)^{\beta}\right) = 1 - \exp\left(-\left(\frac{x-\gamma}{\breve{\eta}}\right)^{\beta}\right)
\end{aligned}$$

where

$$n\frac{1}{\eta^{\beta}} = \frac{1}{\breve{\eta}^{\beta}} \Rightarrow \breve{\eta} = \frac{\eta}{n^{1/\beta}}.$$

The calculations above show that $X_{(1)} \sim W(\beta, \breve{\eta}, \gamma)$. The expectation of $X_{(1)}$ can be determined since the expectation of a two parameter Weibull distributed variable is known. Assume that $Y \sim W(\beta, \eta, 0)$. Then it follows that

$$\mathrm{E}(Y) = \eta\, \Gamma\!\left(\frac{\beta+1}{\beta}\right)$$

where $\Gamma(\cdot)$ is the gamma function that is defined by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt\,.$$

The random variable X is simply a translation of Y, γ units to the right which implies that

$$\mathrm{E}(X) = \mathrm{E}(Y) + \gamma\,.$$

Finally the expectation of $X_{(1)}$ can be determined

$$\mathrm{E}(X_{(1)}) = \gamma + \breve{\eta}\,\Gamma\!\left(\frac{\beta+1}{\beta}\right) = \gamma + \frac{\eta}{n^{1/\beta}}\,\Gamma\!\left(\frac{\beta+1}{\beta}\right).$$

This means that the bias of the estimate $\gamma = X_{(1)}$ is

$$\frac{\eta}{n^{1/\beta}}\,\Gamma\!\left(\frac{\beta+1}{\beta}\right).$$

Therefore

$$\gamma^* = x_{(1)} - \frac{\eta}{n^{1/\beta}}\,\Gamma\!\left(\frac{\beta+1}{\beta}\right)$$

could be an appropriate bias corrected estimate of γ.

Therefore it would be possible to estimate the parameters iteratively according to the scheme

1. Starting guesses for β and η are determined by a direct method.
2. An estimation of γ is determined by the rule above.
3. The parameters β and η are determined by a two parameter Maximum Likelihood distribution fit to the values $\{X_i - \gamma^*\}, \quad i = 1, \ldots, n$.

By looping over point 2 and 3 the parameters will eventually converge and the loop is terminated when they change less than a certain tolerance in one iteration.

4.2.4 Estimators based on elimination

Since an estimate of γ can be determined

$$\gamma^* = x_{(1)} - \frac{\eta}{n^{1/\beta}}\,\Gamma\!\left(\frac{\beta+1}{\beta}\right)$$

a reasonable method to calculate the parameters would be to substitute γ in the likelihood function. Now there is only two free parameters left which can be estimated with ML. When that is done γ is determined by the expression above using the MLE of β and η.

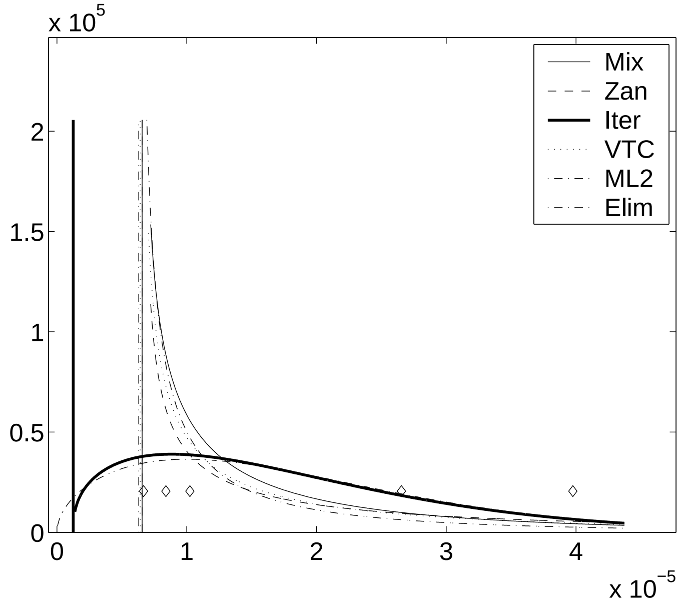

Figure 4.1: Probability density $f_X(x)$, City.

4.2.5 Estimators based on Maximum Likelihood with two parameters

If the threshold $\gamma = 0$ the distribution is the usual Weibull distribution. The likelihood function can be maximized directly, for estimating β and η in this case, and no singularity problem occurs.

4.2.6 Density and distribution fits

Since real observations were available from Volvo Trucks it is interesting to apply the proposed methods to this data. Note that Volvo Trucks uses the assumption that $1/\tilde{D}$ is three parameter Weibull distributed. One way to visualize the result from the inference is to plot the fitted density or distribution function for each of the methods (see Figures 4.1-4.4). (The resulting Weibull density and distribution functions have been plotted.) The observations are represented by diamonds ($\diamond$) in the figures. The estimated parameter values for the different methods are tabulated in Appendix A.

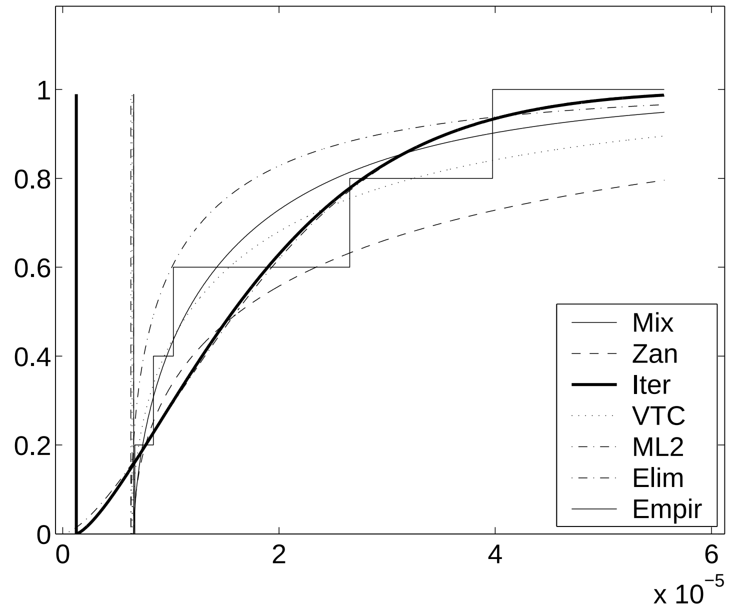

Figure 4.2: Cumulative distribution $F_X(x)$, City

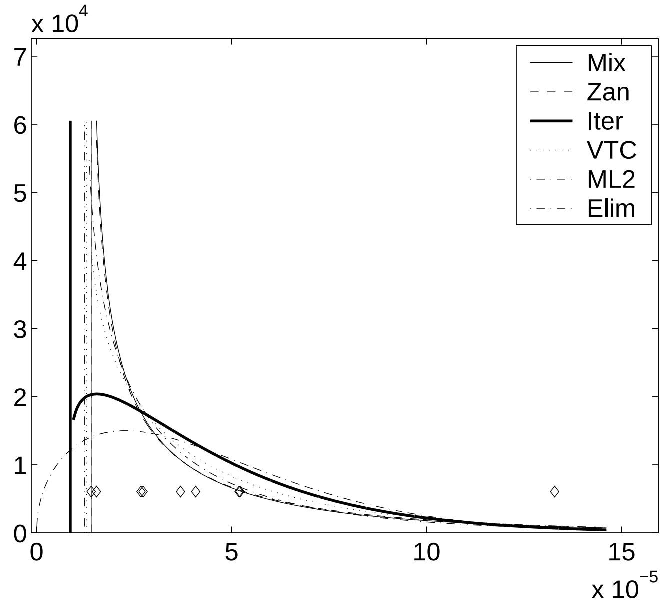

Figure 4.3: Probability density $f_X(x)$, Highway

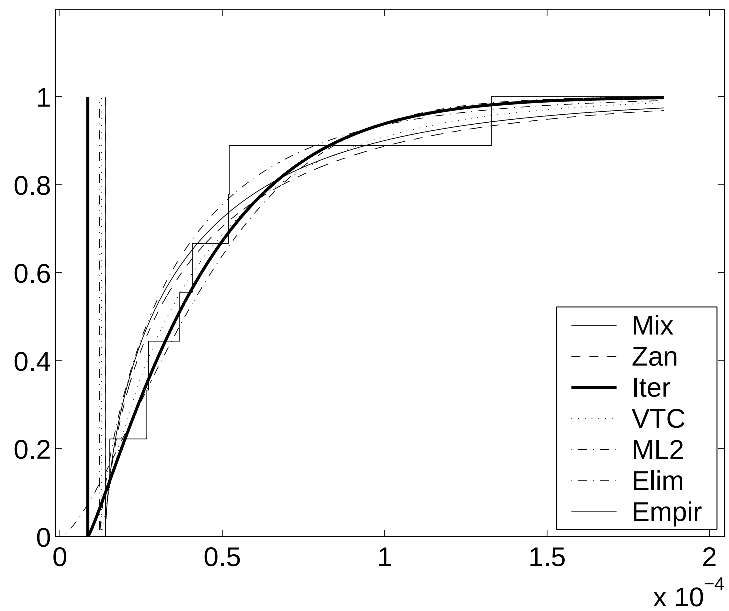

Figure 4.4: Cumulative distribution $F_X(x)$, Highway

4.3 Comparing the estimation methods

Assume θ^* is an estimate of θ. The usual way to estimate the accuracy of an estimate is to calculate the mean squared error (m.s.e.)

$$m.s.e.(\theta^*) = E((\theta^* - \theta)^2) = V(\theta^*) + (E(\theta^* - \theta))^2 = V(\theta^*) + b(\theta)^2.$$

4.3.1 Simulations

The different estimation methods can be compared by simulations. The capacity $C \sim W(2, 2.06 \cdot 10^{11}, 1.56 \cdot 10^{11})$ where the parameters are taken from the Volvo Trucks report. In these simulations only the duty intensity, $\tilde{D}$, is varied. The design distance (a reasonable distance for the life of a vehicle) is set to 1 000 000 km which means that $s = 1\,000\,000$ and the duty $D = s \cdot \tilde{D}$ as usual. This means that $D = s \cdot \tilde{D} = \frac{s}{Y}$ where $Y \sim W(\beta, \eta, \gamma)$. In order to compare the methods four different sets of parameters for Y have been chosen.

1. $\beta = 0.65$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$, $n = 5$
2. $\beta = 0.65$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$, $n = 10$
3. $\beta = 1.30$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$, $n = 5$

4. $\beta = 1.30$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$, $n = 10$

where n is the number of observations. Both the parameters and the number of observations have values that are similar to the ones that were determined from the real observations. This is very important since it means that the conclusions from the simulations are in some way also true for real cases. For each setting of parameters 10 000 iterations have been carried out to get an accurate estimate of the m.s.e. The results are found in tables in Appendix B. The column "$V(\beta^*)$ (%)" means the proportion in percent of the m.s.e, i.e. $100 \cdot \frac{V(\beta^*)}{m.s.e.(\beta^*)}$. The remaining part of the m.s.e. is due to the bias of the estimates and it is shown in the column "$b(\beta)^2$ (%)".

4.3.2 Conclusions

The direct method (Zan) and the method that first uses the direct method in order to estimate γ and then computes MLE of the other two parameters (Mix) are the best methods (smallest m.s.e.). For almost every setting of parameters these methods are first and second best. It is not possible to conclude which of the two methods is best when looking only at the simulations (see Appendix B). The direct method is simple to use which makes it more suitable for use in industrial applications.

4.4 Distribution of the duty

Since the duty depends on the driven distance a design distance must be chosen. A reasonable design distance is 1 000 000 km which means that $s =$ 1 000 000. The duty $D = s \cdot \tilde{D}$. The distribution of D can be determined since the distribution of $1/\tilde{D}$ is known. Let $\tilde{D} = 1/Y$ where $Y \sim W(\beta_Y, \eta_Y, \gamma_Y)$.

$$\begin{aligned} D &= s \cdot \tilde{D} \\ f_D(d) &= \frac{1}{s} f_{\tilde{D}}\left(\frac{d}{s}\right) \end{aligned}$$

$$\begin{aligned} F_{\tilde{D}}(\tilde{d}) &= P(\tilde{D} \leq \tilde{d}) = P\left(\frac{1}{Y} \leq \tilde{d}\right) = P\left(Y \geq \frac{1}{\tilde{d}}\right) = 1 - F_Y\left(\frac{1}{\tilde{d}}\right) \\ f_{\tilde{D}}(\tilde{d}) &= -f_Y\left(\frac{1}{\tilde{d}}\right) \cdot \frac{-1}{\tilde{d}^2} = \frac{1}{\tilde{d}^2} \cdot f_Y\left(\frac{1}{\tilde{d}}\right) \end{aligned}$$

which implies that

$$f_D(d) = \frac{\beta_Y}{\eta_Y} \frac{s}{d^2} \left(\frac{\frac{s}{d} - \gamma_Y}{\eta_Y}\right)^{\beta_Y - 1} \exp\left(-\left(\frac{\frac{s}{d} - \gamma_Y}{\eta_Y}\right)^{\beta_Y}\right), \quad 0 < d < \frac{s}{\gamma_Y}.$$

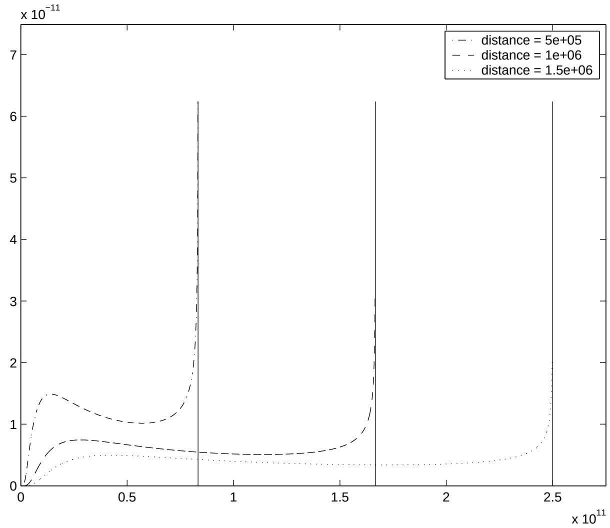

Figure 4.5: Density functions, f_D, when $\beta = 0.65$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$.

The density functions for two parameter sets are plotted in Figure 4.5 and Figure 4.6. Since the density function is plotted for different distances it is possible to see how it changes and especially how the upper limit $(\frac{s}{\gamma_Y})$ moves to the right as the distance increases.

4.5 Failure probability

The failure probability is determined in different ways depending on the distributions of capacity and duty. Below $C \sim W(\beta_C, \eta_C, \gamma_C)$ but the calculations would be similar for other distributions. If the duty is based on three parameter Weibull distribution, i.e. $D = \frac{s}{Y}$ where $Y \sim W(\beta_Y, \eta_Y, \gamma_Y)$ then the failure probability

$$\begin{aligned} P_f &= P(D > C) = P\left(\frac{s}{Y} > C\right) = P(CY < s) = \iint\limits_{cy<s} f_{C,Y}\, dc\, dy = \\ &= \int\limits_{c=\gamma_C}^{\frac{s}{\gamma_Y}} f_C(c) \left(\int\limits_{y=\gamma_Y}^{\frac{s}{c}} f_Y(y)\, dy \right) dc = \int\limits_{c=\gamma_C}^{\frac{s}{\gamma_Y}} f_C(c) \left[F_Y(y)\right]_{\gamma_Y}^{\frac{s}{c}} dc \\ &= \int\limits_{c=\gamma_C}^{\frac{s}{\gamma_Y}} f_C(c) \cdot F_Y\left(\frac{s}{c}\right) dc\,. \end{aligned}$$

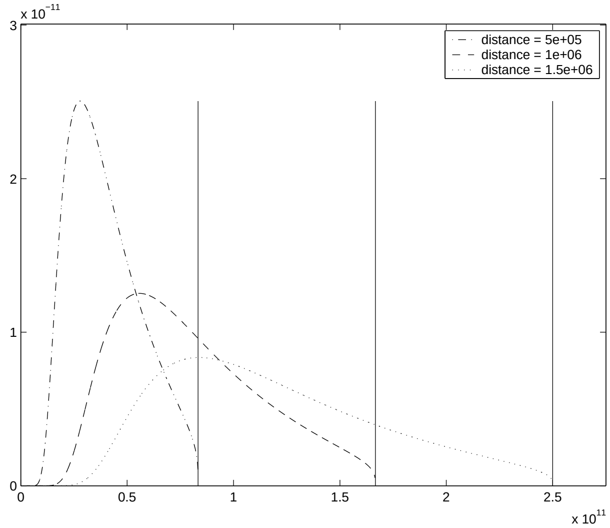

Figure 4.6: Density functions, f_D, when $\beta = 1.30$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$.

If the duty is lognormal, i.e. $D \sim LN(\mu, \sigma^2)$ then the calculation of the failure probability is slightly different.

$$\begin{aligned} P_f &= P(D > C) = \iint\limits_{c<\tau} f_{C,D}\, dc\, d\tau = \int\limits_{\tau=\gamma_C}^{\infty} f_D(\tau) \left(\int\limits_{c=\gamma_C}^{\tau} f_C(c)\, dc \right) d\tau \\ &= \int\limits_{\tau=\gamma_C}^{\infty} f_D(\tau) \left[F_C(c)\right]_{\gamma_C}^{\tau} d\tau = \int\limits_{\tau=\gamma_C}^{\infty} f_D(\tau) F_C(\tau)\, d\tau\,. \end{aligned}$$

In general none of the two integrals for the failure probability can be solved analytically and numerical methods, e.g. Simpson's rule which is used in this thesis, must be applied.

4.5.1 Simulations

According to the results of the simulations, see 4.3.2, the methods Zan and Mix are chosen for further examination. Suppose that the true failure probability is p and one of the methods gives the approximation p^*. A good measure of how close the estimate is to the real value is

$$E_{exp} = \lg \frac{p^*}{p}$$

where lg is the logarithm to the base 10. One problem is how to deal with the cases when $p^* = 0$ since then $E_{exp} = -\infty$. If those cases are dealt with separately it is possible to determine the accuracy of the two methods when $p^* \neq 0$. The probability that the estimated P_f is less or equal to zero is denoted by p_0. The mean (m_E) and standard deviation (s_E) of the error E_{exp} can be determined by simulating a number of times, in this case 10 000. When $\beta = 0.65$, $P_f = 1.4490 \cdot 10^{-4}$ and when $\beta = 1.30$, $P_f = 1.0634 \cdot 10^{-5}$.

It is also of interest to examine how the proportion p_0 changes when the

	n=5			n=10		
	m_E	s_E	p_0^*	m_E	s_E	p_0^*
Zan	0.105	1.032	0.4373	-0.5007	0.9605	0.2626
Mix	-0.0457	1.063	0.4386	-0.3364	0.9155	0.2627

Table 4.1: The entities m_E and s_E when $\beta = 0.65$.

	n=5			n=10		
	m_E	s_E	p_0^*	m_E	s_E	p_0^*
Zan	0.9168	1.6	0.7507	-0.7226	1.508	0.7808
Mix	0.5553	1.657	0.7619	-0.4021	1.463	0.7772

Table 4.2: The entities m_E and s_E when $\beta = 1.30$.

sample sizes increases. This relation is plotted in Figures C.1 and C.2 which are found in Appendix C. The direct method was used because it is very fast, especially for bigger samples.

4.5.2 Conclusions

It is hard to draw any conclusions from the simulations that are summarized in Tables 4.1 and 4.2 since the probability that $P_f = 0$ is so high. This measure of the error, E_{exp} would probably be better in a situation where the failure probability always is positive, e.g. when either the capacity or the duty is lognormal.

4.5.3 Sensitivity analysis

Since the accuracy in determining the failure probability P_f depends on the accuracy in the estimates of the parameters, it is of interest to examine how big that influence is. One way to do that is to vary one of the parameters and keeping the other two fixed and then calculate the failure probability for each combination. The figures are found in Appendix D. The m.s.e. for the direct method is included in the captions.

4.5.4 Conclusions

Since the estimations of the parameters are dependent the figures in Appendix D do not show the true relationship, but nevertheless they give some qualitative information. Anyhow it is clear that the failure probability depends mostly on the value of γ. Therefore the method of determining γ, if there should be a threshold at all, will have a big influence on the final result. Probably it would be better to use a more robust model.

4.6 Qualitative analysis of duty distribution

There are two properties that have to be examined in order to decide how to model the duty. The first of them, which will be discussed in this section is the qualitative property of the model. The other of the two properties are composed of the computational properties, such as stability and accuracy.

4.6.1 Duty based on three parameter Weibull distribution

Since the duty $D = \frac{s}{Y}$ where $Y \sim W(\beta_Y, \eta_Y, \gamma_Y)$ as a consequence there will be an upper limit for D which will be equal to $\frac{s}{\gamma_Y}$. The question is if it is reasonable to have an upper limit for the duty? This means that it is impossible to exceed a certain duty limit no matter how the driver drives or what the road conditions are. This seems strange, but maybe it could be justified if the upper limit is so high that the probability to cause a duty close to this upper limit would be very low. One way to examine this phenomenon more strictly mathematically is to determine certain quantiles for the duty distribution.

$$\begin{aligned} P(D \leq z_p) &= p, \quad \text{which gives} \\ z_p &= \frac{s}{\gamma_Y + \eta_Y \cdot (-\log p)^{1/\beta_Y}} \end{aligned}$$

In order to be able to compare the results for different duty distributions the observations from the Volvo Construction Equipment report have been used (City and Highway). The parameters that were estimated by Volvo

Construction Equipment (Weibull++) are used here, see Appendix A. Two important quantiles are the median $z_{0.50}$ and the 95% quantile $z_{0.95}$. The design distance s is 1 000 000 here. The quantiles and the upper limits are given below.

1. City: $z_{0.50} = 8.53 \cdot 10^{10}$, $z_{0.95} = 15.5 \cdot 10^{10}$, $\frac{s}{\gamma_Y} = 15.6 \cdot 10^{10}$
2. Highway: $z_{0.50} = 6.74 \cdot 10^{10}$, $z_{0.95} = 7.76 \cdot 10^{10}$, $\frac{s}{\gamma_Y} = 7.81 \cdot 10^{10}$

It is clear that if it were be an upper limit it should not be as close to $z_{0.95}$ as it is in this case. It is unreasonable that so many drivers have duty values that are close to the upper limit. Therefore this model seems inappropriate.

4.6.2 Duty based on lognormal distribution

In this case there is no upper limit. The quantiles in the lognormal distribution can be found numerically for the two cases.

1. City: $z_{0.50} = 6.98 \cdot 10^{10}$, $z_{0.95} = 25.0 \cdot 10^{10}$
2. Highway: $z_{0.50} = 2.83 \cdot 10^{10}$, $z_{0.95} = 8.77 \cdot 10^{10}$

4.6.3 Duty based on two parameter Weibull distribution

Here there is no upper limit. The quantiles are given by the same expression as for the three parameter Weibull distribution, but with $\gamma_Y = 0$.

1. City: $z_{0.50} = 6.22 \cdot 10^{10}$, $z_{0.95} = 34.7 \cdot 10^{10}$
2. Highway: $z_{0.50} = 2.60 \cdot 10^{10}$, $z_{0.95} = 15.3 \cdot 10^{10}$

4.6.4 Failure probability for different driven distances

A model that describes duty in a good way should give reasonable results for different driven distances (s). One way to visualize this is to determine P_f for different s and the resulting graphs are found in Figures 4.7 and 4.8. The curves that are denoted by "drivers" corresponds to the failure probability for one certain driver, i.e. one duty intensity value ($\tilde{d}_i$). From this the empirical distribution function for the duty is generated and it is determined by the relations

$$\begin{aligned} P(D < s \cdot \tilde{d}_i) &= 0 \\ P(D \geq s \cdot \tilde{d}_i) &= 1 \end{aligned}$$

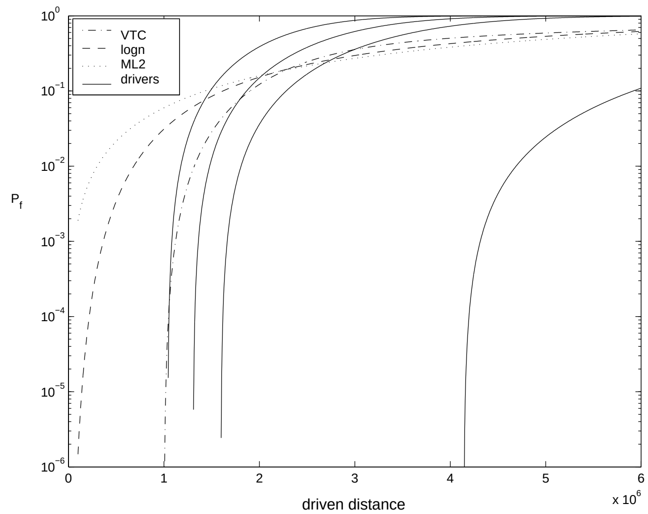

Figure 4.7: Failure probability P_f for different s, City.

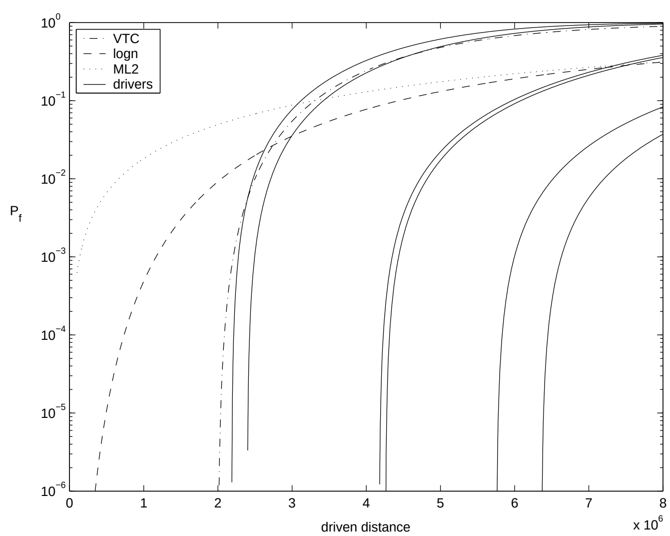

Figure 4.8: Failure probability P_f for different s, Highway.

In Figures 4.7 and 4.8 it is clear that the upper limit results in strange properties for the failure probability. It is zero until a certain distance and then it suddenly increases very fast. This is not reasonable because the failure probability should increase in a smoother way as it does for the other two distributions. It is also interesting to note how much this upper limit differs in the two cases.

There is also a big difference between the lognormal fit and the Weibull fit without upper limit (corresponds to two parameter Weibull distribution). This is due to the fact that the density function for the lognormal distribution decreases faster for larger values compared to the Weibull distribution, especially for smaller s. That is why the failure probability differs with orders of magnitude for small s. Consequently the load-strength model is very sensitive with respect to the choice of distribution. Therefore the model must be compared with real outcomes before it can be used.

Chapter 5

Analysis of Volvo Articulated Haulers model

The model by Volvo Articulated Haulers is described more precisely in Section 3.2. One interesting aspect of this approach is the model for calculating capacity. It is stated that

$$C = \frac{1}{B} \int_{A_0}^{a_c} \frac{da}{f(a)^k}$$

where the random variables and parameters are described in Section 3.2. The function $f(\cdot)$ has been determined by tests

$$f(a) \quad = \quad 0.1388\, a^2 + 0.35\, a + 5.4\,.$$

Since the assumption is that A_0, the initial crack length, and B, a material parameter, are normally distributed tests have been done in order to determine the mean and variance. The results of these tests are that

$$\begin{aligned} A_0 &\sim N(m_{A_0}, \sigma^2_{A_0}) = N(10, 1.78)\,, \\ B &\sim N(m_B, \sigma^2_B) = N(1.832 \cdot 10^{-13}, 2.098 \cdot 10^{-27})\,. \end{aligned}$$

5.1 Model properties

In the model by Volvo Articulated Haulers it is assumed that the initial crack, A_0, should be greater than zero and less than the crack length by failure, i.e.

$$0 \leq A_0 \leq 20\,.$$

Therefore it is of interest to examine this property for the model that is chosen. Since the initial crack is normally distributed it can happen that A_0 takes values outside the interval $[0, 20]$. Therefore the corresponding probabilities are of importance

$$
\begin{aligned}
P(A_0 < 0) &= P\left(\frac{A_0 - m_{A_0}}{\sigma_{A_0}} < \frac{-m_{A_0}}{\sigma_{A_0}}\right) = 1 - \Phi\left(\frac{m_{A_0}}{\sigma_{A_0}}\right) \\
&= 1 - \Phi\left(\frac{10}{\sqrt{1.78}}\right) = 3.31 \cdot 10^{-14}
\end{aligned}
$$

and by symmetry

$$P(A_0 > 20) \;=\; P(A_0 < 0).$$

Since the probability that $A \notin [0, 20]$ is negligible this model error is probably not a big problem. The material parameter, B, should be greater than zero. Since B is normally distributed it can take values less than zero and therefore that probability must be determined

$$
\begin{aligned}
P(B < 0) &= P\left(\frac{B - m_B}{\sigma_B} < \frac{-m_B}{\sigma_B}\right) = 1 - \Phi\left(\frac{1.832 \cdot 10^{-13}}{\sqrt{2.098 \cdot 10^{-27}}}\right) \\
&= 3.17 \cdot 10^{-5}.
\end{aligned}
$$

This means than on average 3 of 100 000 simulated values will be less than zero. Since a large number of observations can be obtained easily and quickly the probability that $B < 0$ is too high, therefore the distribution should be truncated in order to avoid such values.

5.2 General truncation of the normal distribution

Suppose that $X \sim N(m, \sigma^2)$, but it is known that X is restricted to the interval $[x_l, x_u]$ and the probability that X takes values outside this interval is so small that this restriction will have negligible influence on the mean and variance of X. If the original Gaussian shape is preserved then

$$f_X(x) \;=\; \alpha \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right), \quad x_l \le x \le x_u$$

where α is a constant that is determined by the relation

$$P(x_l \le X \le x_u) = 1$$

which gives

$$\begin{aligned} P(x_l \le X \le x_u) &= \int_{x_l}^{x_u} f_X(x)dx = \alpha \cdot \left(\Phi\left(\frac{x_u - m}{\sigma}\right) - \Phi\left(\frac{x_l - m}{\sigma}\right)\right) = 1 \\ \alpha &= \frac{1}{\Phi\left(\frac{x_u - m}{\sigma}\right) - \Phi\left(\frac{x_l - m}{\sigma}\right)}. \end{aligned}$$

For simulating from this distribution, values are generated from the original normal distribution and then values outside the interval $[x_l, x_u]$ are not used.

5.3 Simulations

In order to understand what the resulting distribution of the capacity is for this model 10 000 capacity values were generated ($n = 10\,000$). Then a three parameter Weibull distribution and a lognormal distribution were fitted to these artificial observations. Finally the fitted distributions could be compared with the empirical distribution which gives a lot of information because the number of observations is large. The distribution functions are plotted in Figure 5.1 and the density functions (histogram for the empirical density function) are located in Figure 5.2. Volvo Articulated Haulers uses the three parameter Weibull distribution. Thirty capacity values have been simulated. The Weibull parameters have been determined from a Weibull paper.

5.4 Conclusions

In the Figures 5.1 and 5.2 it is clear that the lognormal distribution gives a better fit than the Weibull distribution especially for low capacity values which is the most important part of the distribution. One reason for this difference is that the smallest observed value has a big influence on the Weibull distribution fit which means that the density function will be comparatively large for small capacity values. When the lognormal fit is carried out the smallest capacity value will not have that big of an influence. In this model the theoretical threshold for the capacity is zero (corresponds to $B \to \infty$ or $A_0 \to a_c$) which means that a distribution with a threshold can not be justified. It is also of interest to compare the Weibull distribution fit carried out by Volvo Articulated Haulers, from thirty capacity values, with the one in this thesis which is based on 10 000 capacity values. The result is found in Table 5.1. Since the threshold γ depends very much on the smallest capacity

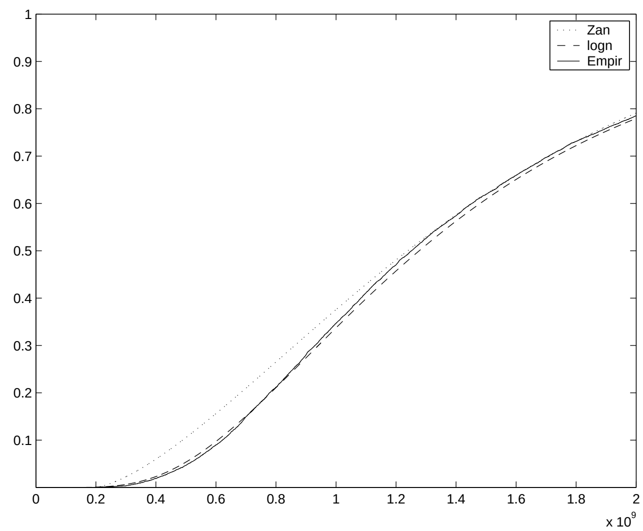

Figure 5.1: Distribution fits for C when $n = 10000$.

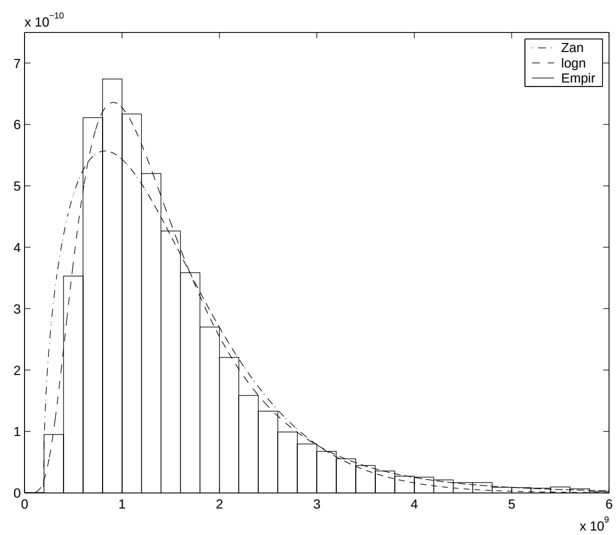

Figure 5.2: Density fits for C when $n = 10000$.

observation it is clear that γ will be smaller when $n = 10\,000$. That in turn has a influence on the estimation of β and η. The use of a threshold can be questioned since it is clear that the more observations (x_i) that are used the better, but at the same time the threshold will tend toward zero since $0 \leq \gamma \leq x_{(1)}$ and $\lim_{n\to\infty} x_{(1)} = 0$.

5.5 Failure probability

In the Volvo Articulated Haulers report a distribution for the duty (corresponds to 1 hours of driving), actually the duty intensity $\tilde{D}$ [duty/h], is found and $\tilde{D} = \frac{1}{Y}$ where $Y \sim W(\beta_Y, \eta_Y, \gamma_Y)$. The parameter values are $\beta_Y = 1.0$, $\eta_Y = 1.065 \cdot 10^{-5}$ and $\gamma_Y = 3.5 \cdot 10^{-7}$. Just three observations were used in order to determine the parameters which means that the uncertainty is very high. A reasonable choice of design life for the machine is 10 000 hours which corresponds to $t = 5\,000$ hours in use (half of the time the engine is idling) which means that the duty $D = t \cdot D = \frac{t}{Y}$. In this case the failure probability has been determined for different capacity distributions: Weibull, lognormal and empirical distribution (using 10 000 capacity values). The result is found in Table 5.2. The failure probabilities are all close to 30% which is the value that was calculated by Bertil Jonsson in the Volvo Articulated Haulers report. The reason that the values are similar is that the failure probability is comparatively high which means that the overlap of the distributions of capacity and duty is large which makes the calculation stable with respect to the choice of distribution. Since the failure probability of course depends on the distribution of the duty more extensive experiments with more duty

Parameter	*Volvo ART (n=30)*	*Zan (n=10000)*
β	2.0	1.49
η	$1.1 \cdot 10^9$	$1.34 \cdot 10^9$
γ	$4 \cdot 10^8$	$1.95 \cdot 10^8$

Table 5.1: Weibull parameter values for different distribution fits.

Distribution of C	*Failure probability*
Weibull	0.3353
lognormal	0.3319
Empirical	0.3103

Table 5.2: Failure probability for different capacity distributions.

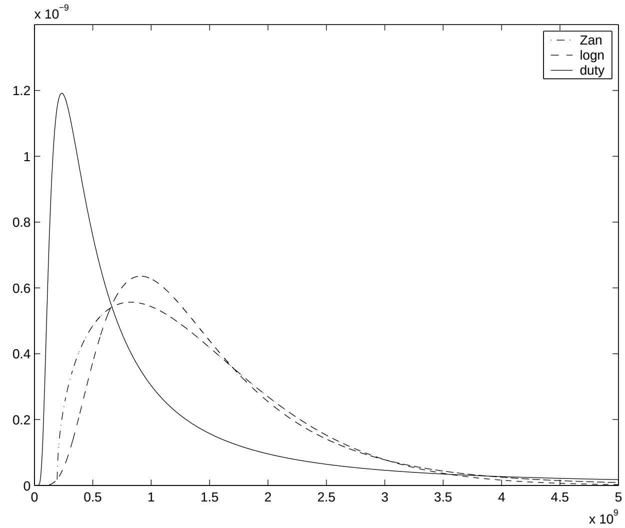

Figure 5.3: Density functions for C and D.

values and other types of distribution fits must be carried out before further conclusions can be drawn. According to this model approximately 30% of the machines are broken after 10 000 hours which can be compared to the real outcome which is about 2%. This seems to be a rather large difference but actually it is not so large if it is considered how sensitive the calculation of the failure probability is. Regarding that only three duty values were used in order to determine its distribution the estimated failure probability is actually close to the real failure probability. Nevertheless the accuracy must be better than that in order to be able to use the model in industrial applications.

5.6 Density functions for capacity and duty

One way to visualize the result is to plot the density functions for capacity and duty. The curves are found in Figure 5.3. Since the density functions for capacity and duty overlap to a rather large extent it is natural that the failure probability will be as big as 30%.

Chapter 6

Feedback using Bayesian estimates

In the Volvo Articulated Haulers' model the failure rate (proportion of failures in the field) was known, but it has not been used in the model. A way to improve the estimates by using the failure rate is to use Bayesian methods. The Bayesian method will here be used for a simple update of the model but it could be generalized to a more sophisticated update.

6.1 Distributions of capacity and duty

Let it be assumed that both capacity and duty are lognormal. Since the data from the Volvo Articulated Haulers' report is being used the distribution of the duty intensity, which in that model is based on a three parameter Weibull distribution, will be transformed to a lognormal distribution with the same mean and variance. This gives $\tilde{D} \sim LN(\mu_{\tilde{D}}, \sigma^2_{\tilde{D}})$ where $\mu_{\tilde{D}} = E(\log \tilde{D}) = 11.97$ and $\sigma^2_{\tilde{D}} = V(\log \tilde{D}) = 1.139$. The distribution of the capacity is determined from the extensive simulations which gives $C \sim LN(\mu_C, \sigma^2_C)$ where $\mu_C = 20.97$ and $\sigma^2_C = 0.3386$. The design life of the machine, t, is assumed to be 5 000 h which gives the duty $D = t \cdot \tilde{D}$. Then it will hold that $D \sim LN(\mu_{\tilde{D}} + \log t, \sigma^2_{\tilde{D}})$. The failure probability, P_f, can be determined

$$P_f = P(D > C) = P(\log D - \log C > 0) = \Phi\left(\frac{\mu_{\tilde{D}} + \log t - \mu_C}{\sqrt{\sigma^2_{\tilde{D}} + \sigma^2_C}}\right)$$

and with the numerical values

$$P_f = \Phi\left(\frac{11.97 + \log 5\,000 - 20.97}{\sqrt{1.139 + 0.3386}}\right) = \Phi(-0.3972) = 0.346.$$

This value can be compared to the failure probability when the duty was based on a three parameter Weibull distribution which gave $P_f = 0.335$ and this means that the transformation to a lognormal duty seems reasonable.

6.2 Bayesian estimates

For each machine included in the set of observed machines a random variable is associated

$$\Lambda_i \;=\; 1_{\{D_i > C_i\}}, \quad i = 1, 2, \ldots, n.$$

This means that machine number i is broken if $\Lambda_i = 1$, and it works if $\Lambda_i = 0$. The total number of machines that are broken after a certain time can then be expressed in the following way

$$S_n \;=\; \sum_{i=1}^{n} \Lambda_i \,.$$

Since the probability that $\Lambda_i = 1$ simply is equal to the failure probability it will hold that

$$S_n \;\sim\; Bin(n, p), \quad p = P_f.$$

In the Volvo Articulated Haulers' report the total number of machines is $n = 917$ and the number of broken machines after 5 000 h is 18 which corresponds to $1.96\% \approx 2\%$ of the population. In this case this would yield the observation $s_{917} = 18$.
Now the Bayesian method can be applied. First, assume that the distribution of C is much more accurate than the distribution of D. This is a reasonable assumption since in general, more information about the capacity than the duty is available. Therefore the distribution of C is kept fixed, but the distribution of D is modified. Since it is easier if only one parameter is free the new model will be

$$\begin{aligned} C &\sim LN(20.97, 0.3386) \\ \tilde{D} &\sim LN(\mu, 1.139) \end{aligned}$$

where the parameter μ is assumed to have normal prior distribution. Suppose that

$$\mu \;\sim\; N(11.97, 1)$$

The prior mean value of μ is just the estimation of $\mu_{\tilde{D}}$ that was determined from the duty observations. The prior variance of μ has here been set to 1, which seems reasonable. It is not evident what variance to use. One idea could be to take into account the spread in the estimate of $\mu_{\tilde{D}}$, but this is hard to carry out since the lognormal distribution of the duty was transformed from a Weibull distribution and not fit from data. The issue of choosing the variance is actually a question about how much influence the prior distribution will have on the posterior distribution, i.e. the final model. A small prior variance means that the duty observations will have a great influence on the model, and a large variance means that the failure rate will have a greater influence. Therefore this issue must be carefully examined in developing this type of Bayesian methods for the load-strength model.

Due to the Bayesian method a reasonable estimation of $\mu_{\tilde{D}}$ would be $E(\mu|s_{197} = 18)$. In general

$$E(\mu|S_n = s_n) \quad = \quad \int \tau \, f_{\mu|S_n}(\tau|s_n) \, d\tau \, .$$

The density function $f_{\mu|S_n}(\tau|s_n)$ can be found via Bayes theorem, see Lindgren [4]

$$f_{\mu|S_n}(\tau|s_n) \quad = \quad \frac{f_{S_n|\mu}(s_n|\tau) \, f_\mu(\tau)}{\int f_{S_n|\mu}(s_n|\tau) \, f_\mu(\tau) d\tau} \, .$$

Since S_n is binomial it will hold that

$$f_{S_n|\mu}(s_n|\tau) \quad = \quad P(S_n = s_n|\mu = \tau) = \binom{n}{s_n} p^{s_n}(1-p)^{n-s_n}, \quad s_n = 0, 1, \ldots, n$$

where

$$p \quad = \quad \Phi\left(\frac{\tau + \log 5\,000 - 20.97}{\sqrt{1.139 + 0.3386}}\right) .$$

The normalization integral can then be solved numerically

$$\int_{-\infty}^{\infty} f_{S_n|\mu}(s_n|\tau) f_\mu(\tau) d\tau = \int_{-\infty}^{\infty} \binom{917}{18} \cdot \Phi\left(\frac{\tau - 12.45}{1.216}\right)^{18} \left(1 - \Phi\left(\frac{\tau - 12.45}{1.216}\right)\right)^{899}$$
$$\cdot \frac{1}{\sqrt{2\pi \cdot 1}} \exp\left(-\frac{(\tau - 11.97)^2}{2 \cdot 1}\right) d\tau = 0.001427 = C^{-1} \, .$$

Now the posterior distribution of μ can be determined

$$f_{\mu|S_n}(\tau|s_n) \quad = \quad C \cdot \binom{917}{18} \cdot \Phi\left(\frac{\tau - 12.45}{1.216}\right)^{18} \left(1 - \Phi\left(\frac{\tau - 12.45}{1.216}\right)\right)^{899}$$
$$\cdot \frac{1}{\sqrt{2\pi \cdot 1}} \exp\left(-\frac{(\tau - 11.97)^2}{2 \cdot 1}\right) .$$

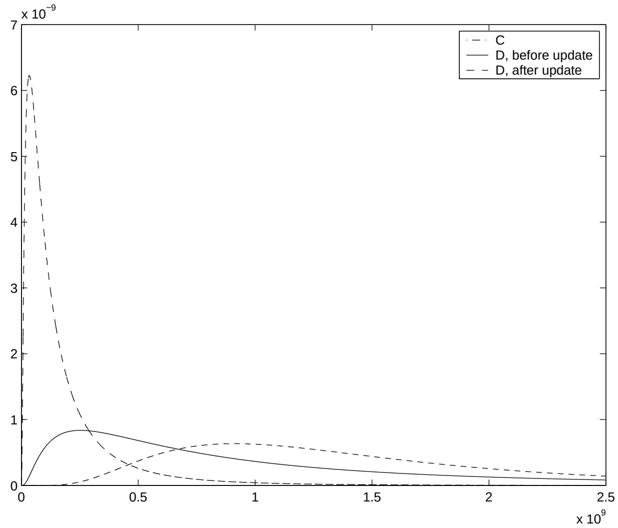

Figure 6.1: Density functions for C and D before and after update.

Since the posterior distribution is known both the mean and variance can be determined numerically

$$\begin{aligned} E(\mu|S_{917} = 18) &= 9.963 \\ V(\mu|S_{917} = 18) &= E(\mu^2|S_{917} = 18) - (E(\mu|S_{917} = 18))^2 = 99.267 - 9.9626^2 \\ &= 0.0132. \end{aligned}$$

Let $\tilde{\mu}$ be the estimation of $\mu_{\tilde{D}}$. The prior estimate $\tilde{\mu} = E(\mu) = 11.97$ and the posterior estimate $\tilde{\mu} = E(\mu|S_{917} = 18) = 9.96$ with variance 0.0132. This means that the parameter $\mu_{\tilde{D}}$ in the load-strength model has decreased from 11.97 to 9.963 due to the failure rate. It is of interest to compare the distribution before and after the update. In Figure 6.1 the density functions are plotted and it can be observed that the distribution of the duty has moved to the left which means that the duty values in general are much lower after the update. The model after the update gives the new failure probability

$$P_f = \Phi\left(\frac{9.963 - 12.45}{1.216}\right) = \Phi(-2.0455) = 0.0204.$$

which means that the failure probability has decreased from 35% to 2%. The reason that the new distribution fits well to the observation of the actual failure probability is a combination of the fact that the number of machines that is observed ($n = 917$) is large which gives a high accuracy in the proportion 2% and that the variance where set to 1. This means that the load-strength model adapts almost completely to the failure rate. Therefore it would be interesting to see how much the result would differ if it instead was one out of 50 machines that were broken. This is still 2% but the accuracy is much lower.

Assume that the prior distribution of μ is the same as before, i.e. $\mu \sim N(11.97, 1)$ and that the observation $s_{50} = 1$ is given. Then the normalization integral is

$$\int_{-\infty}^{\infty} f_{S_n|\mu}(s_n|\tau) f_\mu(\tau) d\tau = \int_{-\infty}^{\infty} \binom{50}{1} \cdot \Phi\left(\frac{\tau - 12.45}{1.216}\right)^1 \left(1 - \Phi\left(\frac{\tau - 12.45}{1.216}\right)\right)^{49}$$
$$\cdot \frac{1}{\sqrt{2\pi \cdot 1}} \exp\left(-\frac{(\tau - 11.97)^2}{2 \cdot 1}\right) = 0.02690.$$

In this case the posterior mean and variance of μ will be

$$\begin{aligned} E(\mu|S_{50} = 1) &= 10.228 \\ V(\mu|S_{50} = 1) &= E(\mu^2|S_{50} = 1) - (E(\mu|S_{50} = 1))^2 = 104.76 - 10.228^2 \\ &= 0.156. \end{aligned}$$

This means that the estimate $\tilde{\mu}$ has decreased from 11.97 to 10.23. This updated parameter value corresponds to the failure probability

$$P_f = \Phi\left(\frac{10.23 - 12.45}{1.216}\right) = \Phi(-1.826) = 0.0339.$$

This value is somewhat larger than 2% which was the result in the other case. Therefore the observation $s_{50} = 1$ had less influence on the model than the observation $s_{917} = 18$, just as predicted.

6.3 Conclusions

Bayesian estimation is a powerful tool for this kind of application since it makes it possible to improve the model as new data becomes available. Furthermore it can be adjusted so that the observations that are most accurate have a major influence on the resulting model. It would be interesting to have failure rates at different times as then the time development of the load-strength model could be studied. Probably this sort of feedback is one of the things that can improve the model enough to make it useful in industrial applications.

Chapter 7

Conclusions and discussion

The main focus of this report is examining the properties of the load-strength model with respect to the choice of distributions for capacity and duty. Two distributions for duty and capacity have been examined, the Weibull and the lognormal distribution. The estimation methods were evaluated based on the data measurements from Volvo Trucks. Since it was not obvious how to estimate the parameters if an entity is modelled as three parameter Weibull distributed, different methods were considered and the effectiveness was determined by extensive simulations. A very simple direct method was one of the two most effective ones and therefore it is recommendable to use that one. Since the estimation method is general this method could also be useful in other applications where the three parameter Weibull distribution is used.

It is also important to study the properties of the distributions in the context of the load-strength model. The assumption that one over the duty intensity $1/\tilde{D}$ is three parameter Weibull distributed leads to some strange properties for the distribution of the duty, D, if the shape parameter $\beta < 1$. By plotting the density function when $\beta < 1$ one can see that a rather high proportion of the probability mass is close to the upper limit which is obviously unreasonable. Results suggest that a reasonable condition when using the three parameter Weibull distribution, for modelling D, is that the shape parameter $\beta > 1$. Furthermore, the Weibull assumption implies an upper limit for the duty and a lower limit for the capacity which means that a safe distance is established. If one were to drive less than this distance the failure probability is zero. Then, the failure probability increases rather rapidly as you drive on past that safe distance. In contrast if both capacity and duty are assumed to be lognormal there will always be an overlap of the distributions, also for very short distances, which means that the failure probability will increase more smoothly as the distance increases.

The failure probability depends on the upper tail of the duty distribution and the lower tail of the capacity distribution. Many observations are needed in order to determine the tail of a distribution with high accuracy. Therefore, the number of observations must be increased considerably in order to obtain reasonable accuracy in the calculation of the failure probability. The fact that the failure probability depends mostly on the tails means that the choice of distribution has a great impact on the final result. This is true even if there is no threshold involved. For example, the upper tail of the duty distribution is in general significantly thinner if the lognormal distribution is used compared to the tail if it is based on a two parameter Weibull distribution.

In the Volvo Articulated Haulers' report a model for determining the distribution of the capacity is used in which no rig test is needed. It is our opinion that it would be very interesting to examine this method more rigorously. From this model one can easily obtain a large number of capacity values and then a distribution fit can be carried out. The lognormal distribution seems to describe the capacity distribution better than the three parameter Weibull distribution in this case. We think that the distributions of the initial crack length a_0 and the parameter B could be determined more precisely which in turn would improve the model. Further on we think the model must be compared to rig tests on the same type of components in order to find out if it works properly. If the model were to give a good description of the capacity it would be very useful since it is cheaper than a rig test and it is easier to apply.

Feedback has been carried out in order to use the failure rate measured in the field. A natural approach is the Bayesian method. Even though only a simple example has been examined in this report the result of this example shows how powerful this method is. In this example the model adapts very closely to the failure rate. This is good since this is a direct observation of what we want to predict. If capacity and duty values are observed they are only indirect observations of the failure probability. Observations of the failure rate could be useful in many ways. First, the model could be updated with the Bayesian model. Secondly, this information could be used for the improvement of the original load-strength model. Then it can, e.g., be examined if the load-strength model is unbiased with respect to the failure probability. Examinations of these kind of properties can be the basis for how to weigh the capacity and duty observations in comparison to the failure rates. In addition, if the times or distances when components fail were observed it would be very useful in determining the distributions for capacity and duty. Then time or distance properties of the model could also be further examined and the ability to predict the future failure rate could be checked.

It is concluded based on the research presented here that the load-strength model must be used with great care and the user must be aware of the fact that subjectivity in the choice of distribution has a huge impact on the result, especially if few observations are available. Even if many observations are available the result can differ significantly due to the fact that the relevant tails have different properties depending on the choice of distribution. According to the study in this report the lognormal distribution should be used rather than the three parameter Weibull distribution for modelling both the capacity and duty.

7.1 Future research

When using the model it is our opinion that the accuracy in the estimation of the distribution parameters should be included which would give a confidence interval for the calculated failure probability. Doing this would gain insight into what extent you can trust the result.

In general the model can be compared to the use of safety factors. Safety factors are tools for handling component design. They are not very accurate, but they will continue to be used as long as there is no other method that works better. Hopefully it will turn out that the load-strength model, if used properly, gives more precise and accurate results. Since statistics are involved in the load-strength model it is very important that accuracy in the predictions are rigorously examined. We think that it is a great challenge to further develop this model. The fact that recent progress in computer technology makes it possible to collect huge amounts of data from field measurements creates a situation where statistical methods can be increasingly useful.

Appendix A

Estimated parameters with different methods

The Weibull distributions fits are carried out on the duty observations from City and Highway driving and the resulting parameter values are tabulated for the different estimation methods.

Estimates based on Maximum Likelihood and direct method (Mix)

	β^*	η^*	γ^*
City	0.634	$8.81 \cdot 10^{-6}$	$6.57 \cdot 10^{-6}$
Highway	0.669	$2.47 \cdot 10^{-5}$	$1.40 \cdot 10^{-5}$

Estimates based on percentiles, direct method (Zan)

	β^*	η^*	γ^*
City	0.51	$2.00 \cdot 10^{-5}$	$6.57 \cdot 10^{-6}$
Highway	0.670	$2.68 \cdot 10^{-5}$	$1.40 \cdot 10^{-5}$

Estimates based on iteration (Iter)

	β^*	η^*	γ^*
City	1.393	$1.89 \cdot 10^{-5}$	$1.25 \cdot 10^{-6}$
Highway	1.160	$3.77 \cdot 10^{-5}$	$8.62 \cdot 10^{-6}$

Estimates determined by VTC, Weibull++

	β^*	η^*	γ^*
City	0.53	$1.06 \cdot 10^{-5}$	$6.41 \cdot 10^{-6}$
Highway	0.85	$3.12 \cdot 10^{-6}$	$1.28 \cdot 10^{-5}$

Estimates based on ML with 2 parameters (ML2)

	β^*	η^*
City	1.514	$2.05 \cdot 10^{-5}$
Highway	1.469	$4.94 \cdot 10^{-5}$

Estimates based on elimination (Elim)

	β^*	η^*	γ^*
City	0.508	$4.50 \cdot 10^{-6}$	$6.30 \cdot 10^{-6}$
Highway	0.793	$2.46 \cdot 10^{-5}$	$1.23 \cdot 10^{-5}$

Appendix B

Results from Weibull inference

The different estimation methods are compared with respect to the mean squared error (m.s.e.). Two different parameter sets are used

1. $\beta = 0.65$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$
2. $\beta = 1.30$, $\eta = 10^{-5}$, $\gamma = 6 \cdot 10^{-6}$

and then the m.s.e. is determined for the estimation of β, η and γ, respectively. For each combination 10 000 samples have been used in order to give a high accuracy in the calculation of the m.s.e.

	m.s.e.(β^*)	$V(\beta^*)$ (%)	$b(\beta)^2$ (%)	*Rank*
Zan	0.1315	95.9	4.1	1
Mix	0.1773	97.5	2.5	2
Iter	0.3742	64.6	35.4	4
ML2	1.343	22.1	77.9	5
Elim	0.1852	99.9	0.1	3

Table B.1: Estimation of β when $\beta = 0.65$ and $n = 5$.

	m.s.e.(η^*)	$V(\eta^*)$ (%)	$b(\eta)^2$ (%)	*Rank*
Zan	$2.092 \cdot 10^{-10}$	86.5	13.5	4
Mix	$5.391 \cdot 10^{-11}$	96.9	3.1	2
Iter	$2.046 \cdot 10^{-10}$	77.4	22.6	3
ML2	$2.431 \cdot 10^{-10}$	35.9	64.1	5
Elim	$5.317 \cdot 10^{-11}$	62.3	37.7	1

Table B.2: Estimation of η when $\beta = 0.65$ and $n = 5$.

	m.s.e.(γ^*)	$V(\gamma^*)$ (%)	$b(\gamma)^2$ (%)	*Rank*
Zan	$4.453 \cdot 10^{-12}$	87.6	12.4	2
Mix	$4.412 \cdot 10^{-12}$	87.1	12.9	1
Iter	$6.473 \cdot 10^{-12}$	85.5	14.5	4
Elim	$4.721 \cdot 10^{-12}$	98.4	1.6	3

Table B.3: Estimation of γ when $\beta = 0.65$ and $n = 5$.

	m.s.e.(β^*)	$V(\beta^*)$ (%)	$b(\beta)^2$ (%)	*Rank*
Zan	0.0571	99.5	0.5	2
Mix	0.04224	95.7	4.3	1
Iter	0.2085	59.1	40.9	4
ML2	0.9158	21.6	78.4	5
Elim	0.08776	99.0	1.0	3

Table B.4: Estimation of β when $\beta = 0.65$ and $n = 10$.

	m.s.e.(η^*)	$V(\eta^*)$ (%)	$b(\eta)^2$ (%)	*Rank*
Zan	$6.219 \cdot 10^{-11}$	94.4	5.6	4
Mix	$2.551 \cdot 10^{-11}$	95.1	4.9	1
Iter	$5.601 \cdot 10^{-11}$	85.3	14.7	3
ML2	$1.679 \cdot 10^{-10}$	24.7	75.3	5
Elim	$3.422 \cdot 10^{-11}$	74.1	25.9	2

Table B.5: Estimation of η when $\beta = 0.65$ and $n = 10$.

	m.s.e.(γ^*)	$V(\gamma^*)$ (%)	$b(\gamma)^2$ (%)	*Rank*
Zan	$5.461 \cdot 10^{-13}$	76.2	23.8	1
Mix	$5.488 \cdot 10^{-13}$	76.4	23.6	2
Iter	$2.229 \cdot 10^{-12}$	70.2	29.8	4
Elim	$7.973 \cdot 10^{-13}$	99.9	0.1	3

Table B.6: Estimation of γ when $\beta = 0.65$ and $n = 10$.

	m.s.e.(β^*)	$V(\beta^*)$ (%)	$b(\beta)^2$ (%)	*Rank*
Zan	0.448	59.6	40.4	4
Mix	0.3782	84.7	15.3	1
Iter	0.391	97.2	2.8	2
ML2	1.234	13.1	86.9	5
Elim	0.4347	80.1	19.9	3

Table B.7: Estimation of β when $\beta = 1.30$ and $n = 5$.

	m.s.e.(η^*)	$V(\eta^*)$ (%)	$b(\eta)^2$ (%)	*Rank*
Zan	$2.508 \cdot 10^{-11}$	99.9	0.1	2
Mix	$2.574 \cdot 10^{-11}$	58.2	41.8	3
Iter	$2.484 \cdot 10^{-11}$	92.1	7.9	1
ML2	$7.611 \cdot 10^{-11}$	15.6	84.4	5
Elim	$3.672 \cdot 10^{-11}$	50.0	50.0	4

Table B.8: Estimation of η when $\beta = 1.30$ and $n = 5$.

	m.s.e.(γ^*)	$V(\gamma^*)$ (%)	$b(\gamma)^2$ (%)	*Rank*
Zan	$9.955 \cdot 10^{-12}$	64.1	35.9	4
Mix	$9.93 \cdot 10^{-12}$	61.1	38.9	3
Iter	$7.911 \cdot 10^{-12}$	99.7	0.3	1
Elim	$9.264 \cdot 10^{-12}$	83.1	16.9	2

Table B.9: Estimation of γ when $\beta = 1.30$ and $n = 5$.

	m.s.e.(β^*)	$V(\beta^*)$ (%)	$b(\beta)^2$ (%)	*Rank*
Zan	0.2424	83.8	16.2	2
Mix	0.2298	66.2	33.8	1
Iter	0.3711	71.7	28.3	4
ML2	1.278	10.3	89.7	5
Elim	0.2774	99.0	1.0	3

Table B.10: Estimation of β when $\beta = 1.30$ and $n = 10$.

	m.s.e.(η^*)	$V(\eta^*)$ (%)	$b(\eta)^2$ (%)	*Rank*
Zan	$1.133 \cdot 10^{-11}$	94.0	6.0	1
Mix	$1.278 \cdot 10^{-11}$	53.0	47.0	2
Iter	$1.399 \cdot 10^{-11}$	91.7	8.3	3
ML2	$6.335 \cdot 10^{-11}$	9.9	90.1	5
Elim	$1.558 \cdot 10^{-11}$	84.5	15.5	4

Table B.11: Estimation of η when $\beta = 1.30$ and $n = 10$.

	m.s.e.(γ^*)	$V(\gamma^*)$ (%)	$b(\gamma)^2$ (%)	*Rank*
Zan	$3.624 \cdot 10^{-12}$	46.3	53.7	1
Mix	$3.631 \cdot 10^{-12}$	47.6	52.4	2
Iter	$4.502 \cdot 10^{-12}$	91.7	8.3	4
Elim	$3.793 \cdot 10^{-12}$	99.4	0.6	3

Table B.12: Estimation of γ when $\beta = 1.30$ and $n = 10$.

Appendix C

Proportion p_0 for different n

The entity p_0 is the proportion of cases when the numerical calculation of the failure probability, P_f, gives the result zero even though the actual value is not zero. This reflects how sensitive the calculation of P_f is when the distribution of the duty has an upper limit and the distribution of the capacity has a lower limit (threshold).

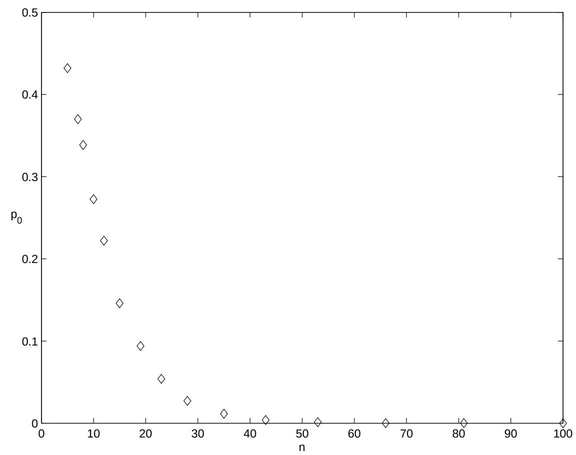

Figure C.1: Proportion p_0 for different n when $\beta = 0.65$.

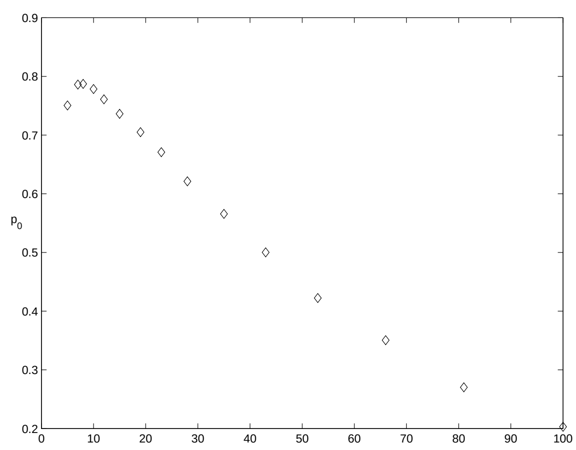

Figure C.2: Proportion p_0 for different n when $\beta = 1.30$.

Appendix D

Sensitiveness in failure probability

The failure probability is determined by varying one of the parameters at the time and keeping the other two fixed. Since the parameter estimates are not independent this does not give the true variation but nevertheless some qualitative information. The parameters that are varied corresponds to the distribution of the duty.

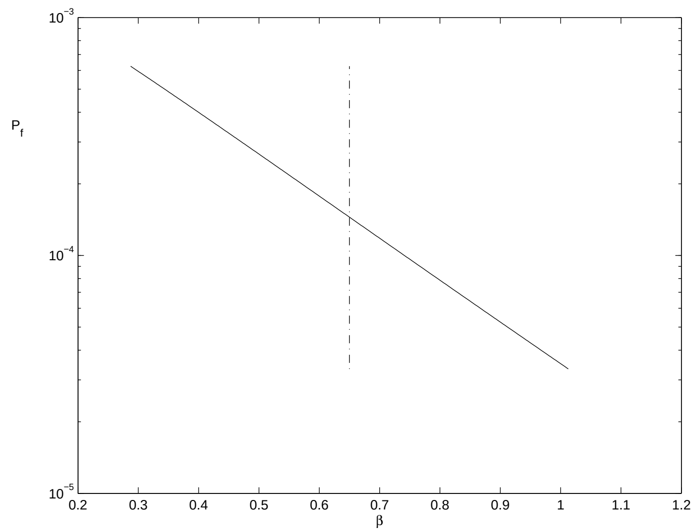

Figure D.1: Failure probability P_f for different β, varying around $\beta = 0.65$ Zan: m.s.e.(β^*)=0.1315 ($n = 5$), m.s.e.(β^*)=0.0571 ($n = 10$).

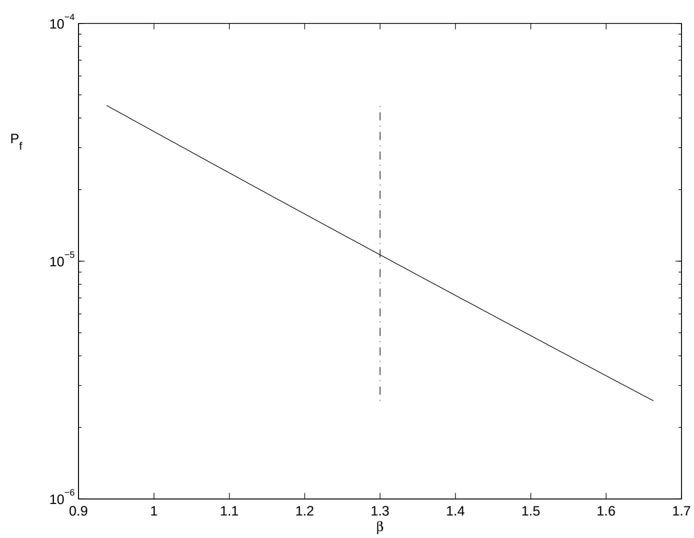

Figure D.2: Failure probability P_f for different β, varying around $\beta = 1.30$ Zan: m.s.e.(β^*)=0.448 ($n = 5$), m.s.e.(β^*)=0.2424 ($n = 10$).

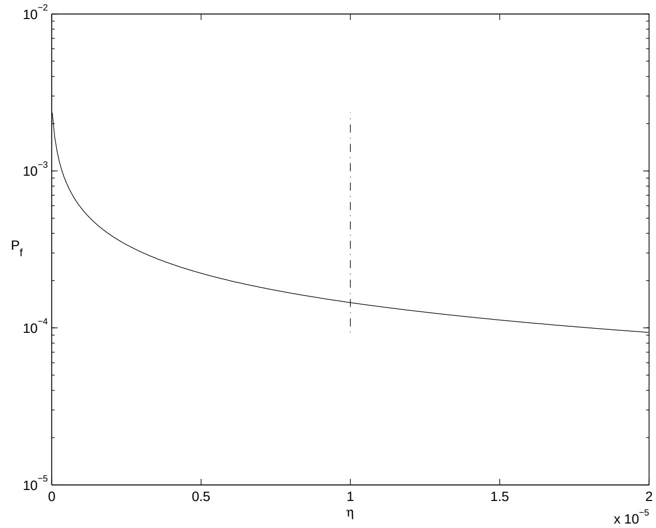

Figure D.3: Failure probability P_f for different η when $\beta = 0.65$
Zan: m.s.e.(η^*)=$2.092 \cdot 10^{-10}$ ($n = 5$), m.s.e.(η^*)=$6.219 \cdot 10^{-11}$ ($n = 10$).

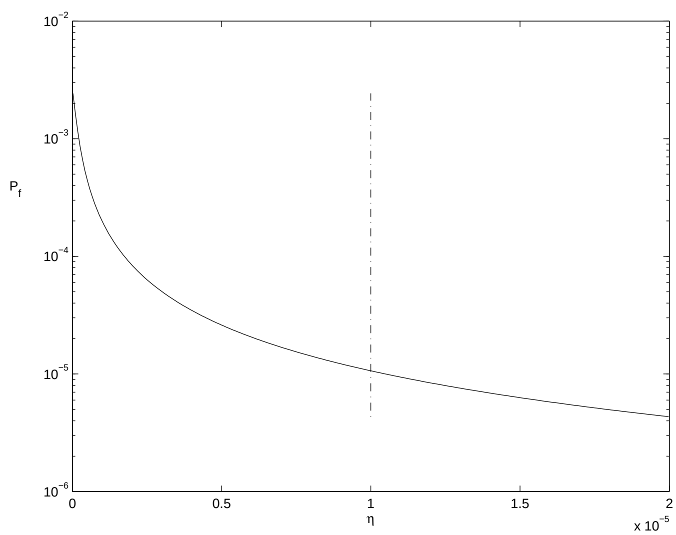

Figure D.4: Failure probability P_f for different η when $\beta = 1.30$
Zan: m.s.e.(η^*)=$2.508 \cdot 10^{-11}$ ($n = 5$), m.s.e.(η^*)=$1.133 \cdot 10^{-11}$ ($n = 10$).

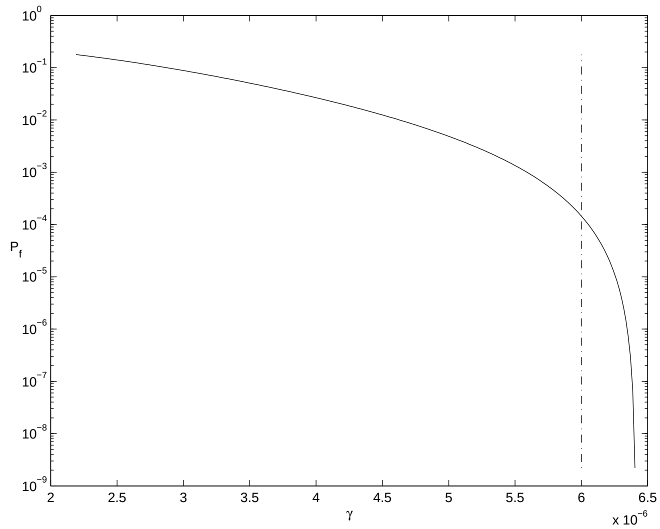

Figure D.5: Failure probability P_f for different γ when $\beta = 0.65$
Zan: m.s.e.(γ^*)=$4.453 \cdot 10^{-12}$ ($n = 5$), m.s.e.(γ^*)=$5.461 \cdot 10^{-13}$ ($n = 10$).

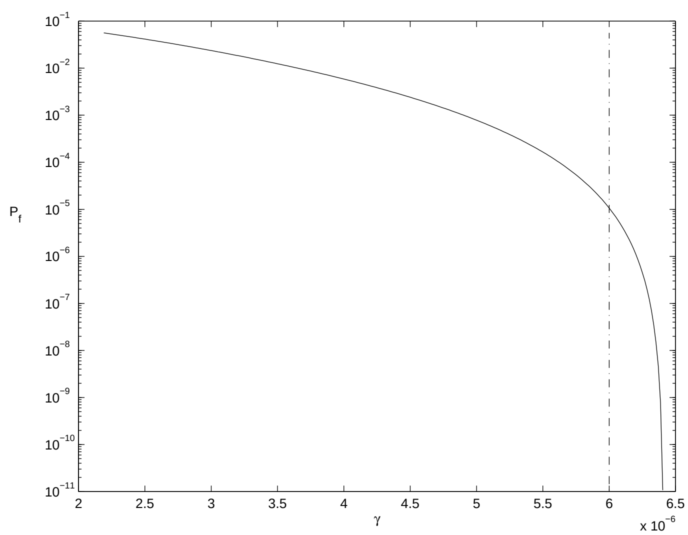

Figure D.6: Failure probability P_f for different γ when $\beta = 1.30$
Zan: m.s.e.(γ^*)=$9.955 \cdot 10^{-12}$ ($n = 5$), m.s.e.(γ^*)=$3.624 \cdot 10^{-12}$ ($n = 10$).

Bibliography

[1] Dubey, S.D. (1967). Some percentile estimators for Weibull parameters, *Technometrics* 9.

[2] Johannesson, P. and de Maré, J. (2002), (2003-12-19). *Problemdriven statistikkurs, Utmattning, belastning och tillförlitlighet*, <http://www.fcc.chalmers.se/~pj/UTMIS/ProblemdrivenStatistikkurs/>

[3] Johannesson, P. (1999). *Rainflow analysis of switching markov loads*, Doctoral Thesis in Mathematical Sciences, Lund Institute of Technology, ISBN: 91-628-3784-2.

[4] Lindgren, B.W. (1998). *Statistical theory*, Fourth edition, Florida, Chapman & Hall, ISBN: 0-412-04181-2.

[5] Olsson, K.E. (1989). Fatigue reliability prediction, *Scandinavian Journal of Metallurgy* 18.

[6] Samuelsson, J. (1997). Fatigue design of construction equipment, *Volvo Technology Report*.

[7] Thomas, J.J., Perroud, G., Bignonnet, A. and Monnet, D. (1999). Fatigue design and reliability in the automotive industry, In *Fatigue Design and Reliability*, ESIS publication 23, edited by Marquis, G. and Solin, J.

[8] Zanakis, S.H. (1979). A Simulation Study of Some Simple Estimators for the Three-Parameter Weibull Distribution, *J. Statist. Comput. Simul.*, Vol. 9.