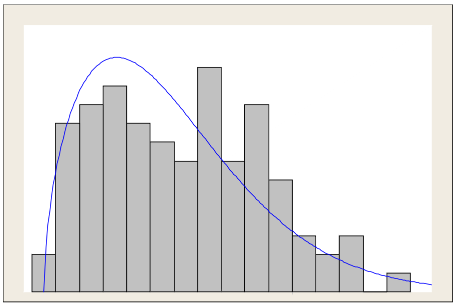
NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF NAVAL ARCHITECTURE & MARINE ENGINEERING

LABORATORY FOR MARINE TRANSPORT



RELIABILITY ANALYSIS/ MAPPING FOR MARINE VESSELS:

RESULTS AND CONCLUSIONS

AUTHOR: GEORGIOS KARADIMAS

SUPERVISOR: NIKOLAOS P. VENTIKOS

ATHENS JULY 2010

This page intentionally left blank

TABLE OF CONTENTS

TABLE OF FIGURES

This page intentionally left blank

Σύνοψις

Η παρούσα διπλωματική εργασία, αποσκοπεί στην ανάλυση αξιοπιστίας σε πλοία (συγκεκριμένα, σε πλοία μεταφοράς εμπορευματοκιβωτίων) βασιζόμενη στην στατιστική ανάλυση πραγματικών στοιχείων από βλάβες πάνω σε πλοία ίδιου τύπου.

Τα προς ανάλυση στοιχεία αφορούν σε βλάβες Κύριας Μηχανής, Ηλεκτρογεννητριών και Βοηθητικών Μηχανημάτων (Time To Failure), που συνελέγησαν από ελληνική ναυτιλιακή εταιρία.

Το πρώτο μέρος της μελέτης είναι κυρίως θεωρητικής φύσεως, καθώς επικεντρώνεται σε βασικούς ορισμούς, απαραίτητους για την βαθύτερη κατανόηση σημαντικών θεμάτων που αφορούν στις μεθοδολογίες αξιοπιστίας. Στη συνέχεα παρουσιάζονται και αναπτύσσονται χρήσιμοι μαθηματικοί ορισμοί και περιγράφονται στατιστικές μεθοδολογίες, ουσιώδους σημασίας για την περαιτέρω μελέτη, κατανόηση και ανάλυση των συλλεχθέντων δεδομένων.

Στο δεύτερο μέρος, παρουσιάζονται τα πραγματικά δεδομένα, και εφαρμόζονται οι μεθοδολογίες που αναπτύχθηκαν εκτενώς στα προηγούμενα κεφάλαια. Αρχικά λοιπόν, προχωρούμε σε κατανομή των μηχανημάτων τα οποία υπέστησαν τις υπό μελέτη βλάβες. Αυτός ο διαχωρισμός γίνεται με βάση το είδος του μηχανήματος (Κύρια Μηχανή, Ηλεκτρογεννήτριες και Βοηθητικά) και τις ώρες που προτείνονται από τον κατασκευαστή του κάθε μηχανήματος για επιθεώρηση. Στη συνέχεια, υπολογίζονται οι στατιστικές κατανομές που περιγράφουν τις πραγματικές βλάβες ανά κατηγορία. Ακολούθως, υπολογίζονται και παράγονται διαγράμματα που απεικονίζουν καμπύλες αξιοπιστίας. Ειδικότερα, εξάγονται διαγράμματα επιβίωσης, δεσμευμένων πιθανοτήτων βλάβης, ρυθμού βλάβης. Οι ίδιες καμπύλες παρουσιάζονται για παραμετρικές και μη παραμετρικές μεθόδους, όπως επίσης και αντίστοιχα πινακοποιημένα αποτελέσματα.

Στο τρίτο μέρος της εργασίας, γίνεται μια προσπάθεια υπολογισμού της συνολικής κατανομής βλαβών για όλο το πλοίο, με την χρήση της μεθόδου των δέντρων σφαλμάτων. Έτσι επιτυγχάνεται η σύμπτυξη των επιμέρους κατανομών σε μια συνολική, για λογούς απλοποίησης των υπολογισμών. Μ’ αυτόν τον τρόπο παράγονται καμπύλες αξιοπιστίας, αντίστοιχες με αυτές του δευτέρου μέρους, που όμως αφορούν το πλοίο ως σύνολο.

Τέλος, με βάση επικαιροποιημένες τιμές κόστους για τα ανταλλακτικά που χρησιμοποιούνται στις απαιτούμενες επισκευές και το αντίστοιχο εργατικό κόστος, υπολογίζουμε τη συνολική δαπάνη, για τις αναγκαίες επισκευές και τη συντήρηση σύμφωνα με τις συστάσεις του κατασκευαστή για επιθεώρηση με βάση τις ώρες λειτουργίας. Για το σκοπό αυτό χρησιμοποιούνται τόσο θεωρητικές/ στατιστικές μέθοδοι όσο και στοιχεία από την πραγματική λειτουργία των μηχανημάτων.

Ως συμπέρασμα οδηγούμαστε μέσω μελέτης ευαισθησίας στην εύρεση βέλτιστης σχέσης μεταξύ συχνότητας επιθεωρήσεων και απαιτούμενου κόστους.

ABSTRACT

The aim of the hereby exposed thesis is the analysis of the reliability on vessels (specifically on containerships) , based on the statistical analysis of real life data of failures on ships of this kind.

The under examination data regard failures of Main Engine, Diesel Generators and Auxiliary Equipment (Time To Failure) , collected from the archives of a greek shipping company.

The first part of this dissertation research is mainly of theoretical nature, as it is focused on basic definitions which are necessary for the deeper comprehension of essential issues related to reliability methods. Afterwards, useful mathematic definitions are presented and developed and statistical methods are described, which are of great importance to the further study, comprehension and analysis of the collected data.

In the second part, we present the real life data and apply methods, which were developed extensively in the previous chapters. Thus we first proceed to a categorization of the machinery which suffered the failures under examination. This categorization is made depending on the type of machinery (Main Engine, Diesel Generators, Auxiliaries) and the amount of hours recommended by the manufacturer of each part of the machinery for inspection. Then, we calculate the statistical distributions which describe the real life failures depending on a category basis. The next step is the calculation and the production of diagrams picturing reliability plots. More specifically, there are exported survival plots, conditional probabilities of failure and failure rate. The same plots as well as the corresponding tabulated results are presented for both parametric and non-parametric methods.

In the third part of this study, we make an attempt of calculating the joint distribution of failures for a vessel as a whole, by applying the fault tree method. In this way there is achieved the integration of the distributions for each subcategory to one joint distribution, which simplifies our calculations. Reliability plots are produced by this method, which correspond to the ones of the second part, but refer to the vessel as a whole.

Finally, based on updated cost for the spare parts used for the required repairs and the corresponding labor cost, we reach the calculation of the total expense for the required remedial actions, according to the recommendations of the manufacturer about repairs depending on operation hours. To this purpose both theoretical/statistical methods and data from the real life operation of the machinery are used.

In conclusion, through the application of sensitivity analysis, we are led to finding the optimum relationship between the frequency of inspections and the demanded cost.

This page intentionally left blank

CHAPTER 1 : THEORETICAL APPROACH

1.1 BASIC CONCEPTS

1.1.1 Reliability

The main concept of this thesis is reliability. However, concerning the broadest and more general definition of reliability, there is a considerable controversy.

Until the 1960s, reliability was defined as ‘’ the probability that an item will perform a required function under stated conditions for a stated period of time.’’ Some authors still prefer this definition ((1), (2)).

A more general definition is given in standards like ISO 8402 and BS 4778:

‘’the ability of an item to perform a required function, under given environmental and operational conditions and for a stated period of time’’ (ISO 8402).

From a qualitative point of view, reliability can be defined as the ability of the item to remain functional. Quantitatively, reliability specifies the probability that no operational interruptions will occur during a stated time interval.

- An item is a functional or structural unit of arbitrary complexity (e.g. component, assembly, equipment, subsystem, system) that can be considered as an entity for investigations.
- The required function specifies the item's task. It could be a single function or a combination of functions that is necessary to provide a specified service.

1.1.2 Availability

is a broad term, expressing the ratio of delivered service to expected service. It is often designated by A and used for the stationary & steady-state value of the point and average availability (PA = AA). Point availability (PA(t)) is a characteristic of an item expressed by the probability that the item will perform its required function under given conditions at a stated instant of time t. From a qualitative point of view, point availability can be defined as the ability of the item to perform its required function under given conditions at a stated instant of time (dependability).

For a given item, the point availability *PA(t)* can be defined as:

$$PA(t) = \frac{MTTF}{MTTF + MTTR}$$

where MTTF (mean time to failure) denotes the mean functioning time of the item and MTTR (mean time to repair) denotes the mean downtime after a repair.

1.1.3 Failure

‘’failure is the termination of the ability of an item to perform a required function’’ (IEC 50).

‘’equipment fails, if it is no longer able to carry out its intended function under the specific operational conditions for which it was designed.’’ (3)

1.1.4 Reliability theory

Deals with the interdisciplinary use of probability, statistics, and stochastic modeling, combined with engineering insights into the design and the scientific understanding of the failure mechanisms, to study the various aspects of reliability.

1.2 BRIEF HISTORY OF RELIABILITY

It is commonly accepted that probability and statistics are the essential ingredients without which Reliability Engineering as a technical discipline could not have emerged. They are the seminal ideas in the history of analytical thought upon which entire scientific and engineering constructs rest, such as Reliability Engineering.

Apart from these two essential pillars, we assume that the idea and practice of mass production—the manufacture of goods in large quantities from standardized parts—is another fundamental ingredient in the development of Reliability Engineering.

Mass production techniques emerged in the early years of 20th century, however a more quantitative/mathematical method and formal approach to reliability grew out of the demands of modern technology and particularly out of the experiences in World War II with complex military systems. The catalyst came in the form of an electronic component, the vacuum tube which for all practical purposes initiated the electronic revolution, enabled a series of applications such as the radio, television, radar and others. The vacuum tube was also the main source of equipment failure. Tube replacements were required five times as often as all other equipments. (4)

It is this experience with the vacuum tubes that prompted the US department of defense to initiate a series of studies for looking into these failures after the war; these efforts eventually consolidated and gave birth to a new discipline, Reliability Engineering.

The next development that provided the necessary ingredient to coalesce all the efforts into a new technical discipline was reliability prediction: if quantitative reliability requirements were going to be specified, there would be a need to estimate and predict component reliability before equipment was built and tested (5).

Then, in late 1950s, came the foundational AGREE report. It was built on many previous efforts in reliability and ‘provided all the armed services with the assurance that reliability could be specified, allocated, and demonstrated; i.e. that a reliability engineering discipline existed’ from that point and onwards, reliability engineering was rapidly developed. There was a trend from component-level reliability to system-level attributes (system reliability, effectiveness, availability, etc.).

Also, there was an increased specialization in statistical techniques, on the actual physics of failure, as well as in specifying, predicting and testing reliability.

Over the next decade, the evolution in reliability was mainly focused on system level reliability and safety of complex engineering system, such as nuclear power plants.

There were introduced techniques such as Probabilistic Risk Assessment and event trees. The second characteristic of the 1970s in the development of Reliability Engineering is the focus on software reliability. (6)

Over the last years, a number of software statistical packages were developed, providing more convenient analysis of life data as well as the extraction of accurate results.

CHAPTER 2

2.1 VARIOUS RELIABILITY ANALYSIS APPROACHES

In this chapter we will examine and analyze briefly various reliability analysis approaches proposed through time.

It is well understood that such a wide subject can be seen in different perspectives. In references the following material is presented in much greater detail.

2.1.1 Life data collection and terminology

Ascher and Feingold (7) evinced clearly the differences in reliability of parts and systems. As ‘parts’, are defined the non repairable items, in contrary to repairable systems. This basic difference results in different reliability and maintainability strategies, even when inferred from two different data sets, which contain exactly the same set-of –failure numbers. Hereby as failure numbers, are denoted parameters like MTTF, FOM (Force Of Mortality) or ROCOF (Rate Of Occurrence of failures).
Despite an almost 50-year evolution and development of reliability theory, Ascher highlights the misunderstanding in the use of FOM, ROCOF and failure rate which are usually confused.
In general, FOM is defined as the propensity that a non repairable item will fail in the next small interval of time.
ROCOF is defined as the mean rate of failures per (repairable) system time.
Summarizing Ascher’s work, it is of essential importance to understand the situation which generated a set of life data, as it will need different approaches to improve reliability. Also it is necessary for engineers and theorists to commence a strict terminology and notation (T&N) use.

2.1.2 Markov approach or Markov modeling

In order to evaluate the operational reliability of a system, a mathematical model based on continuous Markov chains can be used. (8)
Theoretically, for the basic Markov approach to be applicable, the behavior of the system must be characterized by a lack of memory, meaning that the future state of the system is independent of all past states except the present state. Additionally the process must be stationary, meaning that the behavior of the system must be the same at all points of time.
Due to this Markovian hypothesis this approach is applicable to those systems the behavior of which can be described by a distribution with constant hazard rate, such as exponential distribution.
There had been attempts to fit various distributions and obtain a more flexible manner in the Markovian hypothesis.
Such a process is called a semi Markov process.
In a Markovian approach, two states of a system are given: system operative and system failed. For each state, there are given the so called state transition rates:
λ=failure rate
μ=repair rate
The above rates are called state transition since they represent the state at which the system transits from one state to another.

Assuming a constant failure and repair rate, i.e. they are characterized by the exponential distribution, and after performing a few calculations that can be found in (9) it is concluded that the system availability:

$$A(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} e^{-(\lambda+\mu)t}$$

We obtained a reliability function using the Markov process, assuming exponential hazard rates. In following chapters this method will be generalized in order to include semi Markov processes.

2.1.3 Bayesian networks method

Bayesian networks (BN) have been proposed as an alternative to traditional reliability estimation approaches. From this perspective, a BN can be considered as an approach to represent the interactions among the components in a system from probabilistic perspective. This model is widely known as directed acyclic graph (DAG), where the nodes represent system components (variables) and the links between each pair of nodes represent relationship among them. This interaction of the components is leading to system ‘’success’’ or ‘’failure’’. In general the influence among nodes is uncertain; therefore we insert probabilities distribution to each of the links joining the different nodes.

Also it is necessary to assign conditional probabilities between two components, forming a child and parent relationship. The probabilities in the child nodes are calculated through the probability values assigned to the parent nodes.

It is clear that in order to represent all relationship among each node, our model can become substantially large, as the parent- child relationship leads to exponential growth of the conditional probabilities. Also throughout a system’s life, it is common to add or remove obsolete components. Thus the original BN may not be accurate through system’s life.

This complexity leads to the necessity for an expert with adequate knowledge of each specific system to build the BN.

As an alternative/ solution to these problems, a number of special algorithms was introduced and used in BN analysis. Such algorithms were designed to reduce the search space of the best structure that fits adequately each system.

One of the most used algorithms is the K2 algorithm which was defined (10) as a greedy heuristic search method.

Another approach in defining the best structure can be obtained using the MWST (Maximum Weight Spanning Tree) method, combined with standard algorithms, as the Kruskal algorithm.

The proposed methods use the following steps:

- Use of a dataset that contains observations (failure/full functionality) on a system
- Find association between system’s components
- Calculate the degrees of these associations
- Build the associated BN
- Use the BN to estimate overall system reliability.

In references (11) , (12) there can be found analytical calculations as well a full presentation of an applicable BN method.

2.1.4 OREDA analysis

OREDA (13) is a data collection program that was launched in the early 80s. The essential advantage of OREDA is the large number of failure data collected, since a number of important oil companies under the management of DNV and the cooperation of the Norwegian Institute of Technology assure for the importance of the project.
'' OREDA's main purpose is to collect and exchange reliability data among the participating companies and act as The Forum for co-ordination and management of reliability data collection within the oil and gas industry'', as it is clearly stated.
For that purpose reliability data have been collected for about 24.000 offshore equipment units, involving approximately 33.000 failures. The equipments that are covered in this database include rotating machinery (combustion engines, compressors etc), mechanical equipment (cranes, boilers), control and safety units and subsea equipment. It is obvious that all equipments are of highly importance for the marine industry.
The necessary mathematics to analyze the collected failure data are presented in the next chapter. Briefly, after the statistical analyses of failures, based on Markov theory, we can extract useful numerical data for reliability functions (MTTF, etc).
Such results had lead to important applications, such as estimate probabilities of critical events, quantitative risk assessment, reliability centered maintenance, reliability based inspection, life cycle cost, production availability, safety integrity level (SIL), spare parts storage, manning resources, FMEA-analysis, benchmarking/KPI assessment, root cause analysis.
The importance of this project is obvious as a number of International Standards are based upon these analyses (ISO 14 224: "Petroleum, petrochemical and natural gas industries – Collection and exchange of reliability and maintenance data for equipment").
Each attempt for a reliability analysis based on life data faces major difficulties in data collection. This is the advantage of OREDA database, the continuous inflow of such data.

2.2 COLLECTION AND ANALYSIS OF FAILURE DATA

To perform a reliability analysis using statistical/ probabilities methods, a set of data is necessary. Most cases concern data on time to failure. The so called ''failure data'', ''lifetime data'', ''life data'' and so forth. (14)
The key to application of reliability techniques is acquisition, interpretation and analysis of such data. As a result the field of statistics plays a major role in reliability applications.
One can detect that various types of data can be collected. For example operational data are those collected under actual operating conditions. Such experimental data, regarding the marine industry, are seldom available.

To implement and complete this thesis, we have contacted a major Greek shipping company and collected failure (operational) data of the containership's fleet.
These data contains time to failure of various major components, auxiliary items and machinery of 7 containerships out of a fleet of 11 ships and for a period of 2 years.
These failure data were mainly collected from maintenance record. This means that both component specific failures (primary failures) and common cause failures are included.
Repair times and necessary manpower for the repairs are recorded whenever possible.

CHAPTER 3 INTRODUCTION TO RELIABILITY MATHEMATICS

The methods used to quantify reliability are the mathematics of probability and statistics. Since we are dealing with life data it is necessary to understand and define basic descriptive statistics.
A parameter is defined to be a population characteristic; we have already seen the MTTF. The corresponding quantity in a sample is called a statistic.
The objective of descriptive statistics is to calculate appropriate statistics for purposes of description and summarization of the information in a set of data, thus, effectively and efficiently inferences concerning parameters could be determined.

3.1 HISTOGRAMS

A histogram is a graphical representation in bar chart form of a frequency table or frequency distribution.
For example, from the life data we have collected, and regarding the failures in a group of sister vessel's liners of Main Engine, for a period of 2 years, the following table can be obtained.

Table 1: sample data for failures in M/E liners (operational hours until failure)

1650	5723	17372
1924	5741	
2063	7745	
2676	8822	
3361	16505	
5093	17089	

To construct the histogram for the above data, we need to divide the range of the data into equally spaced class intervals covering all the data points. The range that the data will be divided is important, since too small range would not reveal the shape of the data, and too large a number would result in a flat appearing distribution.
Sturges (15) showed that the best visual graphical representation achieved when determining the number of data range into *k* equally spaced classes where

$$k=1+3.332\log(n).$$

Where n is the sample size.
Thus, the following table and diagram can be produced.

*Table 2**:** Frequency table of Liners data*

Cell boundaries	number in cell
0-3600	5
3600-7200	3
7200-10800	2
10800-14400	0
14400-18000	3

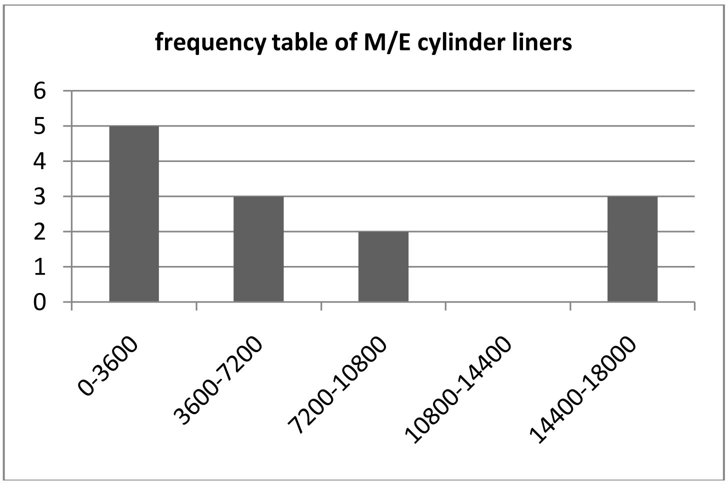


Figure 1: Histogram of measurements

3.2 FREQUENCY DISTRIBUTION

A frequency distribution is a graphical or numerical description of an entire set of data. The objective is to present the data information in a concise form and in such a way that, if possible, the general shape of the distribution is displayed.
For the above numerical example the well known cumulative distribution function has the following form.

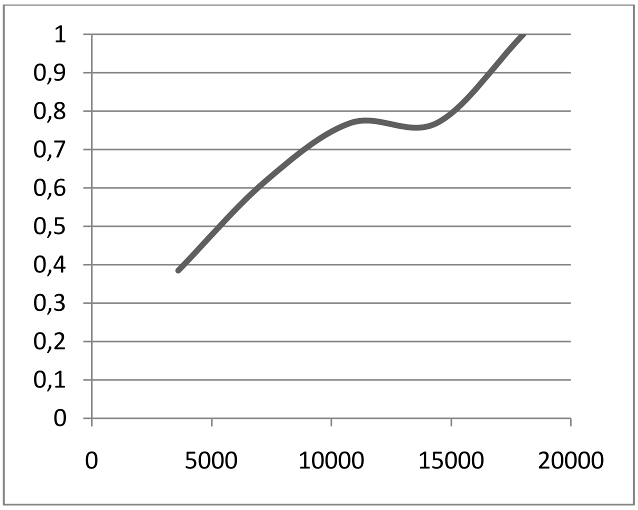


Figure 2: Plot of cumulative distribution function

Cumulative distribution gives the number of values in the data set that are at or below a given value. In our analysis the value of a population cumulative distribution function at a given time is the population fraction failing by that time.

3.3 DISTRIBUTIONS AND DENSITY FUNCTIONS

The model corresponding to the frequency distributions is the probability density function (PDF), denoted by f(x), where x is any value of interest (in our case, the time between failures).
In reliability approaches, the cumulative distribution function (CDF) is related to the PDF via the following relationship

$$F(x) = \int_0^t f(y)\,dy, \qquad 0 \le t < \infty$$

In reliability engineering we are concerned with the probability that an item will survive for a stated interval, meaning that there is no failure in the interval (0 to t). This is given by the reliability function R(x):

$$R(x) = 1 - f(x) = \int_x^\infty f(x)\,dx = 1 - \int_{-\infty}^x f(x)\,dx$$

The hazard function or hazard rate h(x) is the conditional probability of failure in the interval x to (x+dx), given that there was no failure by x:

$$h(x) = \frac{f(x)}{R(x)} = \frac{f(x)}{1 - F(x)}$$

The cumulative hazard function H(x) is given by:

$$H(x) = \int_{-\infty}^x h(x)\,dx = \int_{-\infty}^x \frac{f(x)}{1 - F(x)}\,dx$$

3.4 SAMPLE ESTIMATES OF POPULATION PARAMETERS

In order to complement the visual impression given by the various diagrams, a number of parameters are defined:

- The sample mean, often referred as the expected value:

$$\mu = \int_0^\infty t f(t)\,dt$$

- The variance, is an estimator of the spread of the data:

$$V(t) = \sigma^2 = \int_0^\infty (t - \mu)^2 f(t)\,dt$$

- The square root of the variance is called the standard deviation, also known as scale parameter.

The mean, variance, and other sample estimates are often referred to as nonparametric point estimators. They are nonparametric because they may be evaluated without knowing the population distribution from which the sample was drawn, and they are point estimators because they yield a single number.

3.5 CONFIDENCE INTERVALS

The uncertainty in the estimation of various probabilistic functions parameters can be approached through the so called confidence intervals.

How well the sample statistic estimates the underlying population value is always an issue. A confidence interval addresses this issue because it provides a range of values which is likely to contain the population parameter of interest. Confidence intervals are constructed at a confidence level, such as 95%. It means that if the same population is sampled on numerous occasions and interval estimates are made on each occasion, the resulting intervals would bracket the true population parameter in approximately 95% of the cases. A confidence stated at a 1-a level can be thought of as the inverse of a significance level, a.

For a given set of life data, if the parent distribution is known, then the point and interval estimates of the distribution parameters become the center of attention (16).

For large sample sizes, point estimates, and confidence intervals for distribution parameters may be expressed in elementary terms; then the sampling distributions approach the normal form, enabling the confidence intervals to be expressed in terms of the standard normal CDF.

That property yields directly from the Central Limit Theorem:
Provided that the sample size is sufficiently large (for reliability analysis, N>30 is considered adequate) the sampling distribution for mean becomes normal.

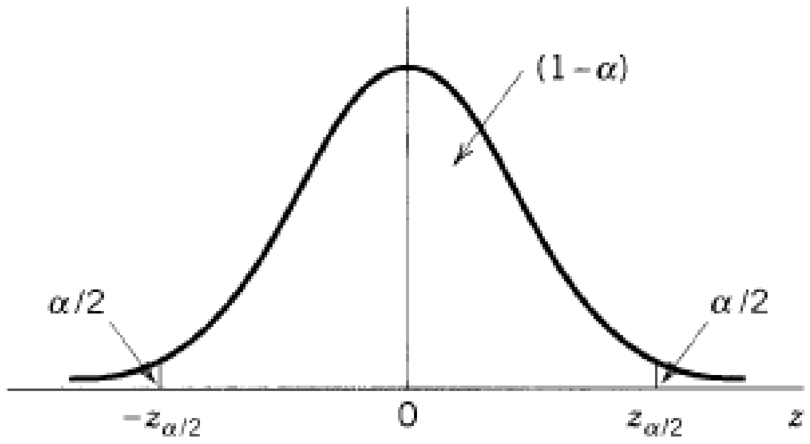


Figure 3: Standard Normal Distribution

In relative statistical references (17) can be found point estimates and associated confidence intervals for nearly all distributions.

The following figure displays the survival plot, with the 95% confidence limits for a given set of life data.
The use and importance of survival plots will be evinced in Chapter 5, where such figures will be used consecutively for the calculation of the number of units that have survived to a given time. Also the survival function leads to some important conclusion, i.e. the type and shape of the appropriate distribution that fits adequately well the real life data.

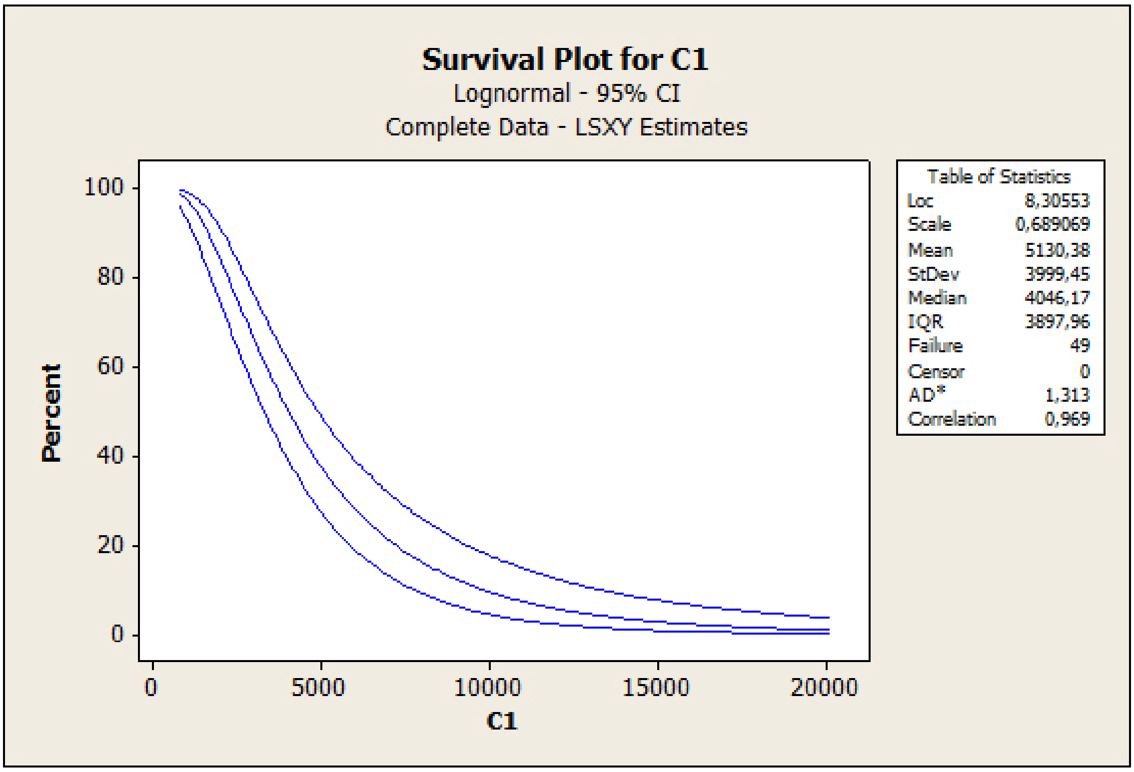


Figure 4: Survival plot with confidence limits

CHAPTER 4 GRAPHICAL ANALYSIS AND PROBABILITY PLOTTING

4.1 PARAMETRIC ANALYSIS

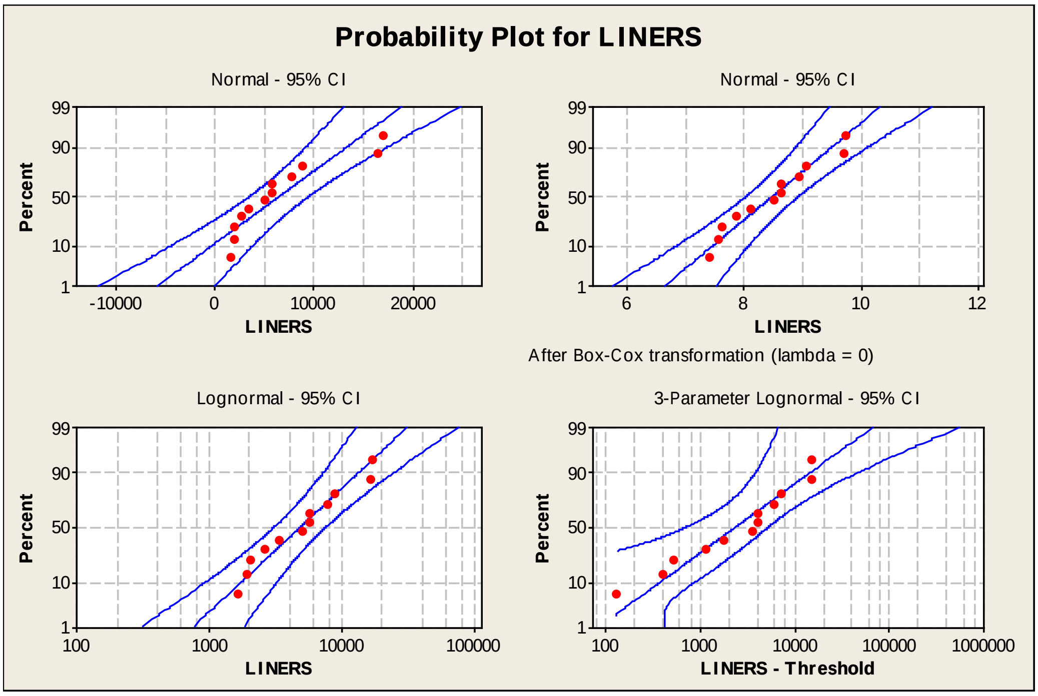
It is of great importance in reliability engineering to determine which distribution best fits a set of data and to derive interval estimates of the distribution parameters. Graphical estimation methods can greatly ease this task and probability plotting papers have been developed for this purpose. These are based upon the cumulative distribution function of each distribution. The axes of probability plotting papers are transformed in such a way that the true c.d.f. plots as a straight line. Therefore if the plotted data can be fitted by a straight line, the data fit the appropriate distribution (18).

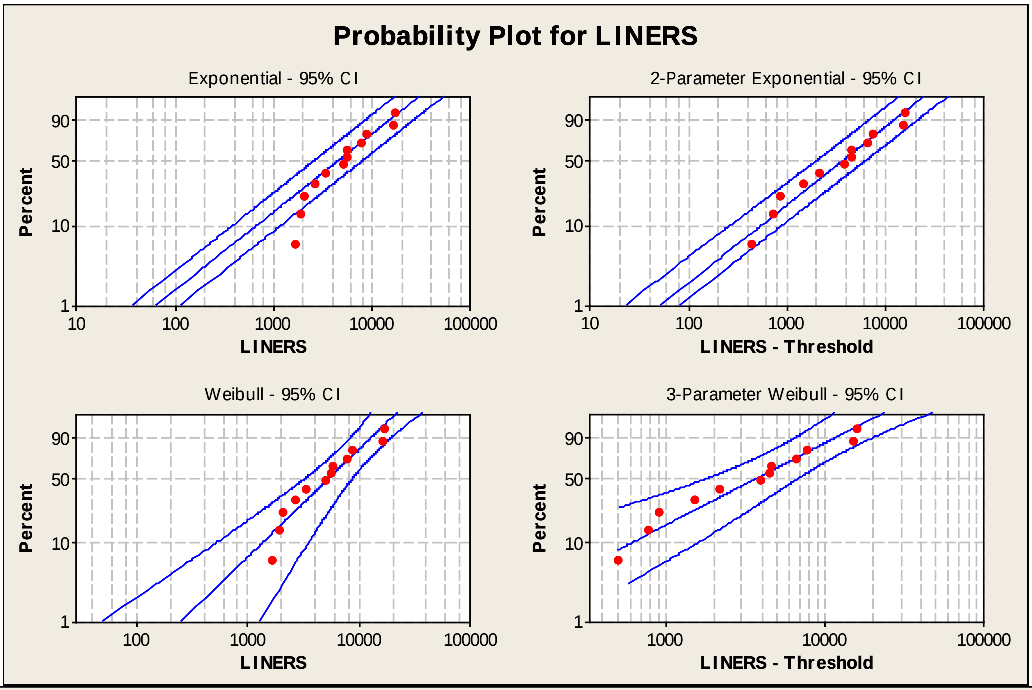
Now, recalling example 1, with the time to failure of the cylinder liners.

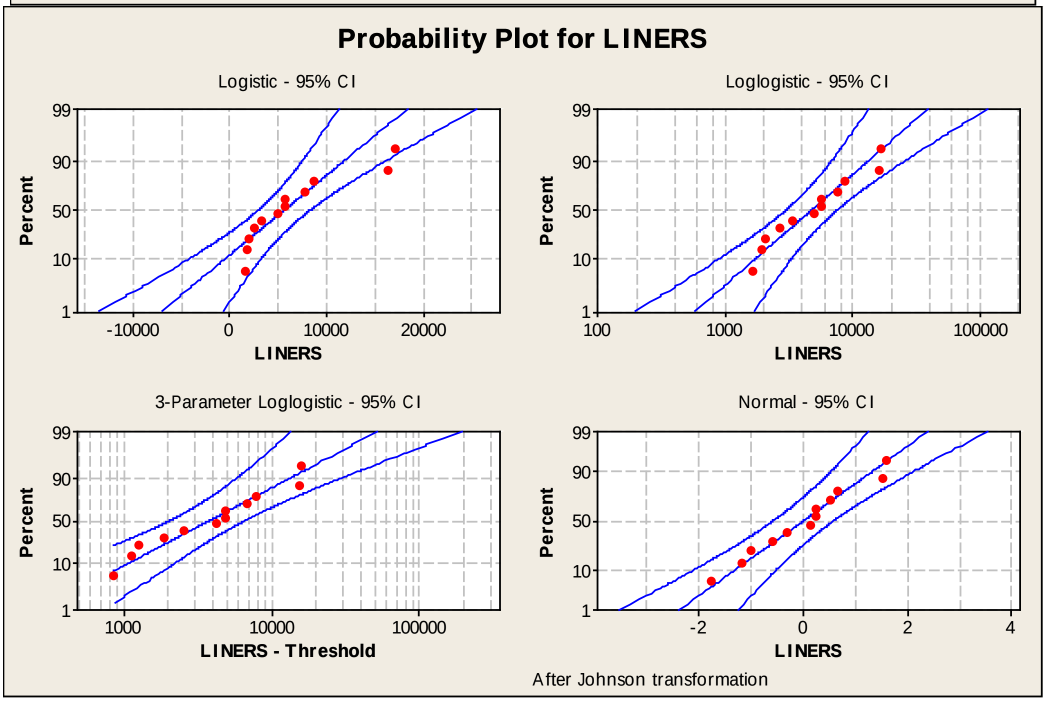
Table 3***: sample data for failures in M/E liners (working hours until failure)***

1650	5723	17372
1924	5741	
2063	7745	
2676	8822	
3361	16505	
5093	17089	

In our effort to plot these data in proper plotting papers, the following diagrams were produced.




Probability Plot for LINERS
Exponential - 95% CI
2-Parameter Exponential - 95% CI
Weibull - 95% CI
3-Parameter Weibull - 95% CI
Percent
LINERS
LINERS - Threshold


Probability Plot for LINERS
Logistic - 95% CI
Loglogistic - 95% CI
3-Parameter Loglogistic - 95% CI
Normal - 95% CI
Percent
LINERS
LINERS - Threshold
After Johnson transformation

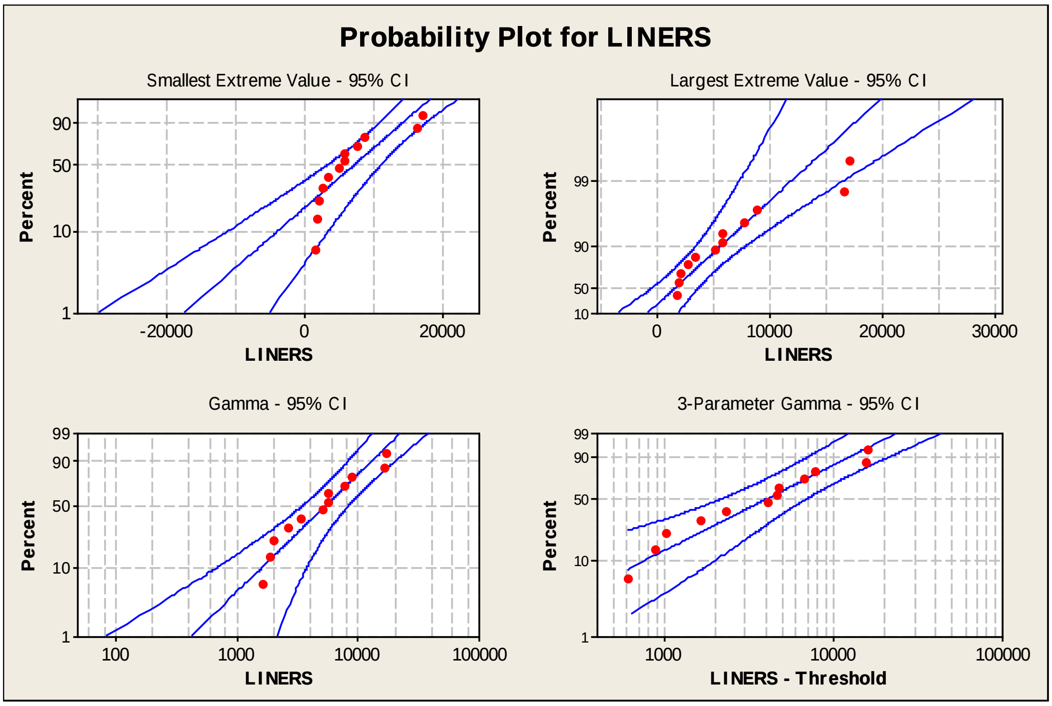


Figure 5: *Individual distribution identification*

Because of the small sample size, we can't export reliable conclusions for the best fitting distribution, but with a closer look we can take useful lessons.

4.2 NON PARAMETRIC ANALYSIS

In cases where no assumption is made for the form of the underlying distribution, there have been developed methods for measuring and comparing statistical variables. These methods called non parametric or distribution free statistical methods. They are slightly less powerful than parametric in terms of the accuracy of the inferences derived for assumed known distributions. The basic advantage is that they are simple to use and can be very useful in reliability studies, provided that the data which were being used are independently and identically distributed (i.i.d.) (19).

4.2.1 Non parametric estimates

For a given set of data the following sample measures may be derived.

- The sample mean

$$\bar{\mathbf{x}} = \frac{1}{N}\sum_{k=1}^{N}\mathbf{x}_k .$$

- The sample median

- The sample standard deviation

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}.$$

- The sample variance

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

- The sample coefficient of variation

$$\mathrm{CV}=\bar{x}/\sigma$$

4.2.2 The empirical distribution and survival function

For a set of n lifetimes data the empirical distribution function is defined as

$$F_n(t) = \frac{Number\ of\ lifetimes \leq t}{n}$$

The corresponding empirical survivor function is

$$R_n(t) = 1 - F_n(t) = \frac{Number\ of\ lifetimes > t}{n}$$

Usually such observations are distinct; therefore R_n is a step function that decreases by 1/n just before each observed failure time.

For the previously mentioned example we got the following figure:

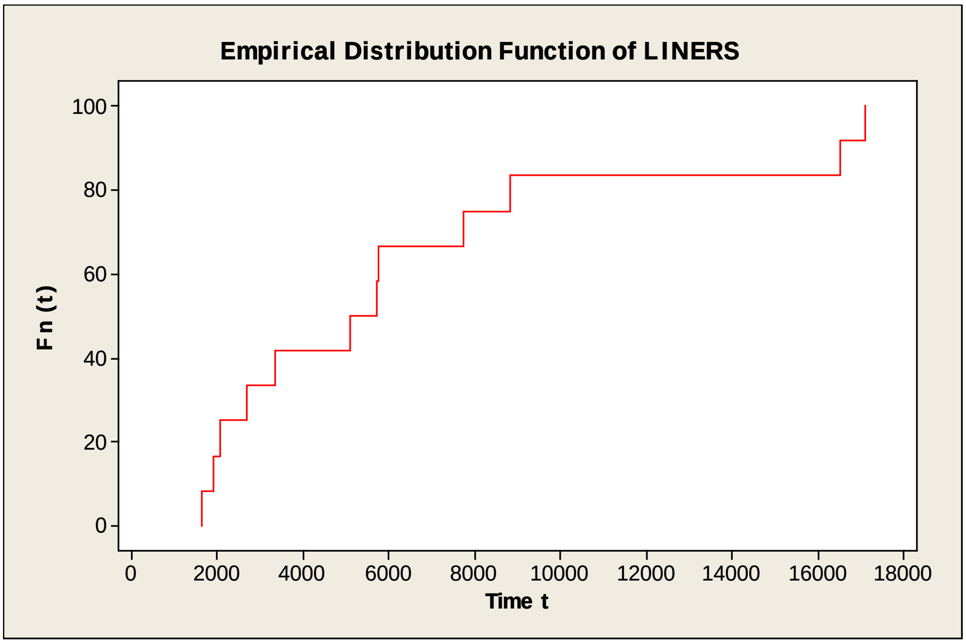


Figure 6: Empirical distribution function

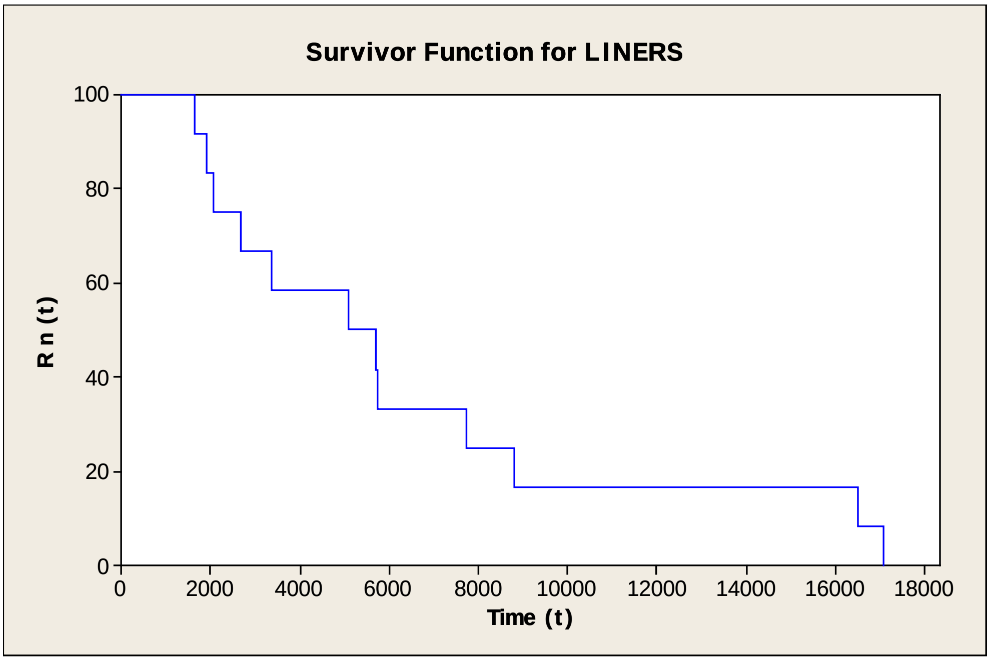


Figure 7: Survivor function

4.2.3 Kaplan Meier estimators

The Kaplan- Meier estimators provide simple estimates of the reliability (survival function) or the Cumulative Distribution Function for complete or censored sample data without assuming a particular distribution model (20).

A general expression for the K-M estimates can be written. Assume we have n units on test and order the observed times for these n units from t_1 to t_n. Some of these are actual failure times and some are running times for units taken off test before they fail. Keep track of all the indices corresponding to actual failure times. Then the K-M estimator of R(t) are given by:

$$\hat{R}(t) = \prod_{j=0}^{n} \frac{n_{j-1}}{n_j}$$

Where $j \in S$ and t_j less than or equal to t_i means we only form products for indices j that are in S and also correspond to times of failure less than or equal to t_i.

Once values for $R\ (t_i)$ are calculated, the CDF estimates are

$$F\ (t_i) = 1 - R\ (t_i)$$

A natural estimator of the cumulative failure rate Z(t) is deducted from R(t):

$$\hat{Z}(t) = -ln \prod_{j=0}^{n} \frac{n_{j-1}}{n_j}$$

If the points (t,Z(t)) are plotted, from the shape of the plot we can extract a number of conclusions.

a) A convex Z(t) indicates an increasing failure rate
b) A concave Z(t) indicates a decreasing failure rate.

The cumulative failure rate diagram is also referred as Nelson plot since it was suggested by Nelson (17).

4.2.4 Boxplots

In non parametric analysis, boxplot is a very useful tool.
Boxplots are an excellent tool for conveying location and variation information in data sets, particularly for detecting and illustrating location and variation changes between different groups of data.
Boxplots are formed by vertical axis, which is the response variable and horizontal axis, which is the factor of interest.
To produce a boxplot the following actions are performed.

- Calculate the median and the quartiles (the lower quartile is the 25th percentile and the upper quartile is the 75th percentile).
- Plot a symbol at the median (or draw a line) and draw a box (hence the name box plot) between the lower and upper quartiles; this box represents the middle 50% of the data- the "body" of the data.
- Draw a line from the lower quartile to the minimum point and another line from the upper quartile to the maximum point. Typically a symbol is drawn at these minimum and maximum points, although this is optional.

Thus the boxplot identifies the middle 50% of the data, the median, and the extreme points.

There is a useful variation of the boxplot that more specifically identifies outliers. To create this variation (19)

- Calculate the median and the lower and upper quartiles.
- Plot a symbol at the median and draw a box between the lower and upper quartiles.
- Calculate the interquartile range (the difference between the upper and lower quartile) and call it IQ.
- Calculate the following points (also called fences):

$$L1 = \text{lower quartile} - 1.5*IQ$$
$$L2 = \text{lower quartile} - 3.0*IQ$$
$$U1 = \text{upper quartile} + 1.5*IQ$$
$$U2 = \text{upper quartile} + 3.0*IQ$$

The line from the lower quartile to the minimum is now drawn from the lower quartile to the smallest point that is greater than L1. Likewise, the line from the upper quartile to the maximum is now drawn to the largest point smaller than U1.

Potential outliers points are indicated with an asterisk in the boxplot. It is obvious that none of our sample data is an outlier as it is clearly indicated in Figure 8. However, it must be noticed that an outlier point, as defined in next part, it is not necessarily ignored. Careful investigation must be carried out to determine the further use of outlier points.

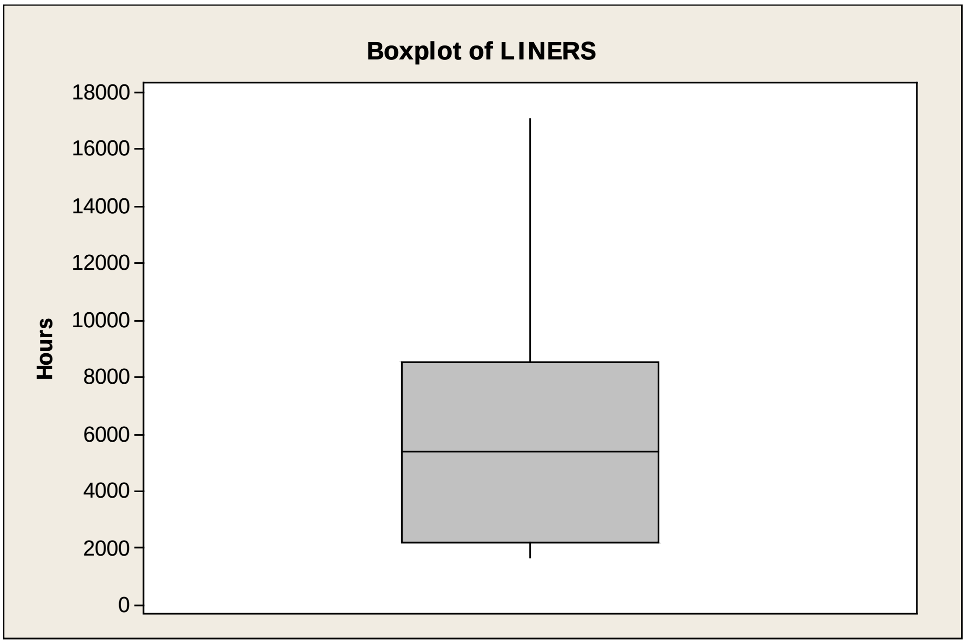


Figure 8: Boxplot for liners

4.3 OUTLIERS

An *outlier* is an observation that lies an abnormal distance from other values in a random sample from a population.
‘’ An outlying observation, or outlier, is one that appears to deviate markedly from other members of the sample in which it occurs.’’ (22)
A point beyond an inner fence on either side is considered a mild outlier. A point beyond an outer fence is considered an extreme outlier.
Outliers should be investigated carefully. Often they contain valuable information about the process under investigation or the data gathering and recording process. Before considering the possible elimination of these points from the data, one should try to understand why they appeared and whether it is likely that similar values will continue to appear.

4.4 GOODNESS OF FIT

In analyzing statistical data we need to determine how well the data fit an assumed distribution. The goodness of fit can be tested statistically to provide a level of s-significance that the null hypothesis (that the data indeed fit well an assumed distribution) is rejected. Goodness of fit testing can be considered an extension of s-significance testing in which the sample cdf is compared with the real cdf.
A number of methods are available for testing how closely a set of data fits an assumed distribution. Hereunder are presented the methods that are used in this thesis.

4.4.1 The x^2 goodness of fit

The chi-square test is used to test if a sample of data came from a population with a specific distribution.

An attractive feature of the chi-square goodness-of-fit test is that it can be applied to any univariate distribution for which you can calculate the cumulative distribution function. The chi-square goodness-of-fit test is applied to binned data (i.e., data put into classes). This is actually not a restriction since for non-binned data you can simply calculate a histogram or frequency table before generating the chi-square test. However, the value of the chi-square test statistic is dependent on how the data is binned. Another disadvantage of the chi-square test is that it requires a sufficient sample size in order for the chi-square approximation to be valid.

The chi-square test is defined for the hypothesis:
H_0 : The data follow a specified distribution.
H_1 : The data do not follow the specified distribution
Test statistic: For the chi-square goodness-of-fit computation, the data are divided into k bins and the test statistic is defined as

$$x^2 = \sum_{i=1}^{k} (O_i - E_i)^2 / E_i$$

Where O_i is the observed frequency for bin i and E_i is the expected frequency for bin i. The expected frequency is calculated by

$$E_i = N(F(Y_u) - F(Y_l))$$

where F is the cumulative Distribution function for the distribution being tested, $\boldsymbol{Y_u}$ is the upper limit for class $\boldsymbol{i}$, $\boldsymbol{Y_l}$ is the lower limit for class $\boldsymbol{i}$, and N is the sample size.
This test is sensitive to the choice of bins. There is no optimal choice for the bin width (since the optimal bin width depends on the distribution). Most reasonable choices should produce similar, but not identical, results.
The test statistic follows, approximately, a chi-square distribution with (k - c) degrees of freedom where k is the number of non-empty cells and c = the number of estimated parameters (including location and scale parameters and shape parameters) for the distribution + 1. For example, for a 3-parameter Weibull distribution, c = 4.

Therefore, the hypothesis that the data are from a population with the specified distribution is rejected if

$$x^2 > x^2_{(a,k-c)}$$

where $x^2_{(a,k-c)}$ is the chi-square percent point function with k - c degrees of freedom and a significance level of a.
In the above formulas for the critical regions, the convention that is followed is that x^2 is the upper critical value from the chi-square distribution and x^2_{1-a} is the lower critical value from the chi-square distribution.

4.4.2 The Kolmogorov- Smirnov test

Another goodness of fit test commonly used in statistics and reliability is the Kolmogorov-Smirnov (K-S) test. It is rather simpler to use than the x^2 test and can give better results with a relatively small number of life data.

The Kolmogorov–Smirnov statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution, or between the empirical distribution functions of two samples. The null distribution of this statistic is calculated under the null hypothesis that the samples are drawn from the same distribution (in the two-sample case) or that the sample is drawn from the reference distribution (in the one-sample case). In each case, the distributions considered under the null hypothesis are continuous distributions but are otherwise unrestricted.

The two-sample KS test is one of the most useful and general nonparametric methods for comparing two samples, as it is sensitive to differences in both location and shape of the empirical cumulative distribution functions of the two samples.

The Kolmogorov-Smirnov test is defined by:
H_0 : The data follow a specified distribution.
H_a : The data do not follow the specified distribution
Test statistic: the Kolmogorov-Smirnov test statistic is defined as

$$D = \max_{1 \le i \le N} (F(Y_i) - \frac{i-1}{N}, \frac{i}{N} - F(Y_i))$$

where $\boldsymbol{F}$ is the theoretical cumulative distribution of the distribution being tested which must be a continuous distribution (i.e., no discrete distributions such as the binomial or Poisson), and it must be fully specified (i.e., the location, scale, and shape parameters cannot be estimated from the data).
The hypothesis regarding the distributional form is rejected if the test statistic, $\boldsymbol{D}$, is greater than the critical value obtained from a table. There are several variations of these tables in the literature that use somewhat different scalings for the K-S test statistic and critical regions. These alternative formulations should be equivalent, but it is necessary to ensure that the test statistic is calculated in a way that is consistent with how the critical values were tabulated.

4.4.3 The Anderson- Darling test

The Anderson-Darling test is used to test if a sample of data came from a population with a specific distribution. It is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test. The K-S test is distribution free in the sense that the critical values do not depend on the specific distribution being tested. The Anderson-Darling test makes use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution. Currently, tables of critical values are available for the normal, lognormal, exponential, Weibull, extreme value type I, and logistic distributions.

The Anderson-Darling test is defined by:
H_0 : The data follow a specified distribution.
H_a : The data do not follow the specified distribution
Test statistic: the Anderson-Darling test statistic is defined as

$$A^2 = -N - S,$$
$$S = \sum_{i=1}^{N} \frac{2i-1}{N} (ln\, F(Y_i) + \ln\,(1 - F(Y_{N+1-i}))$$

where ***F*** is the cumulative distribution function of the specified distribution. Note that the Y_i are the ordered data.

The critical values for the Anderson-Darling test are dependent on the specific distribution that is being tested. Tabulated values and formulas have been published (23) for a few specific distributions (normal, lognormal, exponential, Weibull, logistic, extreme value type 1). The test is a one-sided test and the hypothesis that the distribution is of a specific form is rejected if the test statistic, A, is greater than the critical value.

4.4.4 P- value

The so called p- value determines the appropriateness of rejecting the null hypothesis in a hypothesis test. P-values range from 0 to 1. The smaller the p- value, the smaller the probability that rejecting the null hypothesis is a mistake. Before proceeding with the reliability analyses it is of great importance to determine the s-significance level. A commonly used value is 0.05. If the p- value of a test statistic is less than the s- level, the null hypothesis may be rejected.

Because of their indispensable role in hypothesis testing, p- values are used in many areas of statistics including basic statistics, linear models, reliability, and multivariate analysis among many others. The key is to understand what the null and alternate hypotheses represent in each test and then use the p-value to aid in the critical decision to reject the null hypothesis.

4.4.5 Examples of goodness of fit tests

For example, for the above mentioned set of life data, the goodness of fit test gives the following graphs. The assumed hypothesis is that the data following a normal distribution.

This test compares the empirical cumulative distribution function of the sample data with the distribution expected if the data were normal. If this observed difference is sufficiently large, the test will reject the null hypothesis of population normality.

If the p-value of this test is less than the chosen s-level, we can reject the null hypothesis and conclude that the population is non-normal.

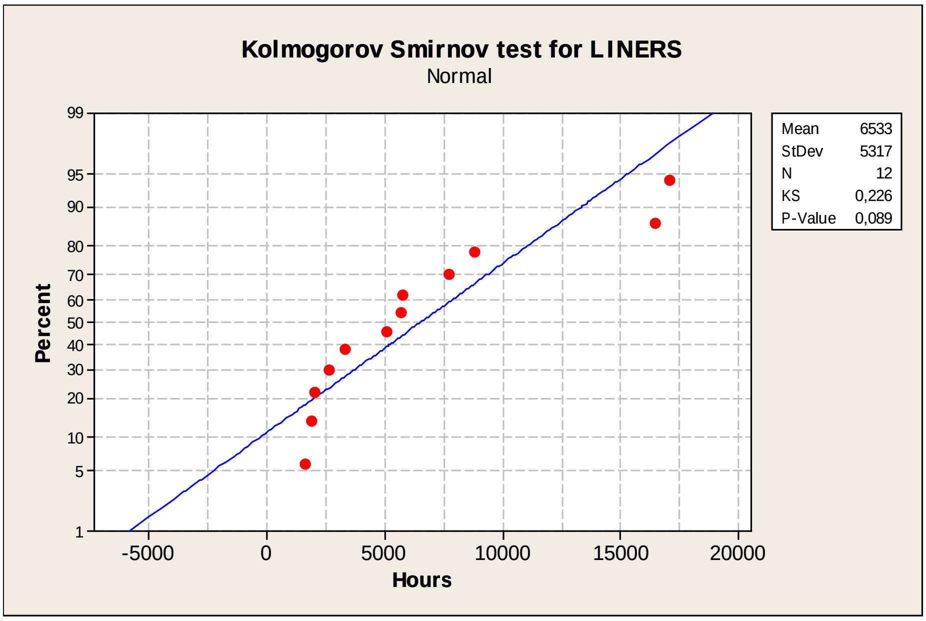


Figure 9: Kolmogorov- Smirnov test

The respective graph for Anderson-Darling test, concluding more details for the calculated parameters is the following:

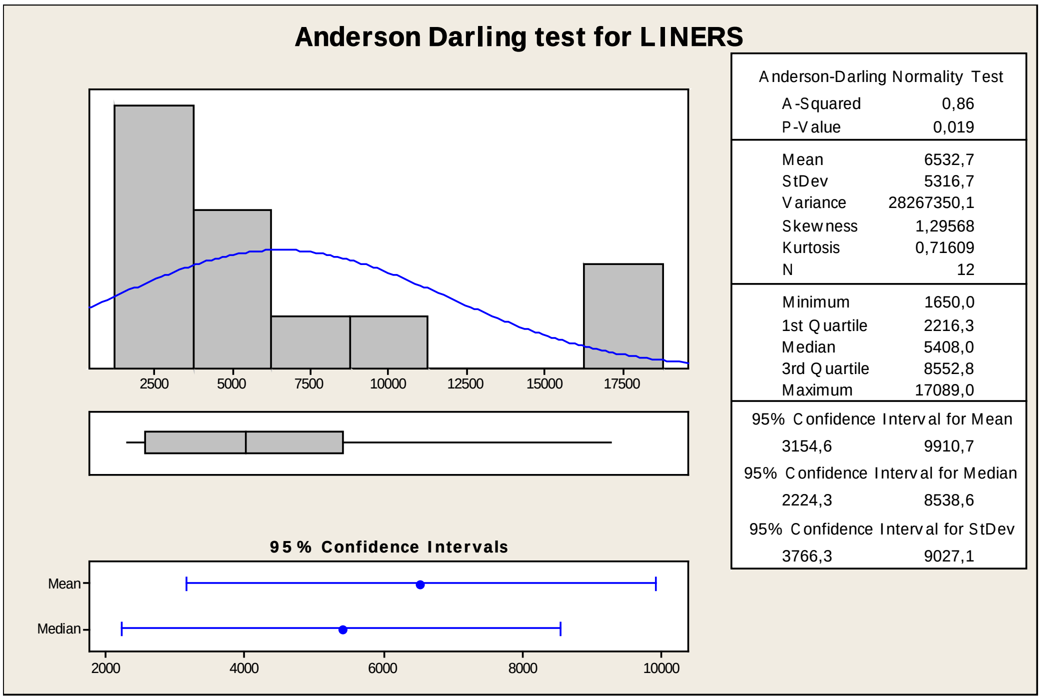


Figure 1.10: *Anderson- Darling test*

The graphical summary includes four graphs: histogram of data with an overlaid normal curve, boxplot, 95% confidence intervals for μ, and 95% confidence intervals for the median.

CHAPTER 5 EXPERIMENTAL RESULTS

The preceding theoretical analysis, was nothing more than an introduction, so that the below experimental results to become deeply understandable.

It is well understood that all vessels in order to avoid unexpected repairs and failures follow a preventive maintenance plan for all equipment. The time to inspection depends on each manufacturer.

PROBABILITY PLOTTING

The field data that where collected, regarding failures of onboard equipments in marine vessels can be initially analyzed through data plots. These data were divided into three categories, failures in Main Engine, failures in Diesel Generators and failures in Auxiliary equipment.

Failures in Main Engine concern parts which, according to the manufacturer, must be inspected so that reliable operation is ensured, every 4000 and 8000 working hours. Therefore a logical approach is to further divide Main Engine failures in two categories, i.e. failures to parts in which the manufacturer suggests 4000 and 8000 working hours between inspections.

Following the same approach for the auxiliary equipment, these were divided into three subcategories, parts with inter-inspection time of 4000, 6000 and 8000 working hours.

On this basis the following tables can be obtained.

Table 4*: M/E 4000h*

540	2553	4146
564	3088	4212
688	3098	4237
1092	3725	4256
1487	3860	4317
1634	3970	6374
2076	4043	9548
2134	4102	10042
2447	4137	

Table 5: *M/E 8000h*

357	2704	4650	6491
746	3022	5190	6565

1650	3073	5382	6608
1924	3095	5389	6790
1935	3103	5460	6899
2063	3300	5723	6947
2112	3348	5726	7549
2132	3473	5740	7745
2161	4436	5741	8330
2676	4465	5790	8864
2685	4480	5887	131399

Table 6***: Diesel Generators (inter inspection time of 4000hours)***

55	2100	4238
183	2527	5062
629	2546	5609
844	2971	5849
1100	2987	6058
1123	3121	6157
1300	3150	6485
1397	3180	6593
1512	3401	7690
1751	4128	

Table 7: *AUX 4000h*

552	1428	2650
651	1463	2800
1031	1892	3425
1300	2480	4256
1400	2490	

Table 8: *AUX 6000h*

196	1950	3723
1500	2540	3840
1605	3340	3856
1612	3700	5380
1700	3720	

Table 9: AUX 8000h

345	2630	5520	8148
350	3577	5835	8235
400	4000	5921	8542
475	4172	5922	9990
1444	4367	6334	15544
1461	4480	6650	19624
1462	5312	7800	
2250	5360	7900	
2300	5400	8020	

For each one of the grouped data, a statistical analysis, both non parametric and parametric, can be applied, so that the basic descriptive statistics are interpreted and understood and it is concluded whether these follow a specific distribution.

These calculations and diagrams will be repeated for all groups.

5.1 M/E 4000 h

5.1.1 Outlier points

This analysis will begin by checking the life data for possible outlier point. It is reminded that an outlier is not necessary a false measure. Often they contain valuable information about the process under investigation of the data gathering and recording process. Before considering the possible elimination of these points from the data, one should try to understand why they appeared and whether it is likely similar values will continue to appear.

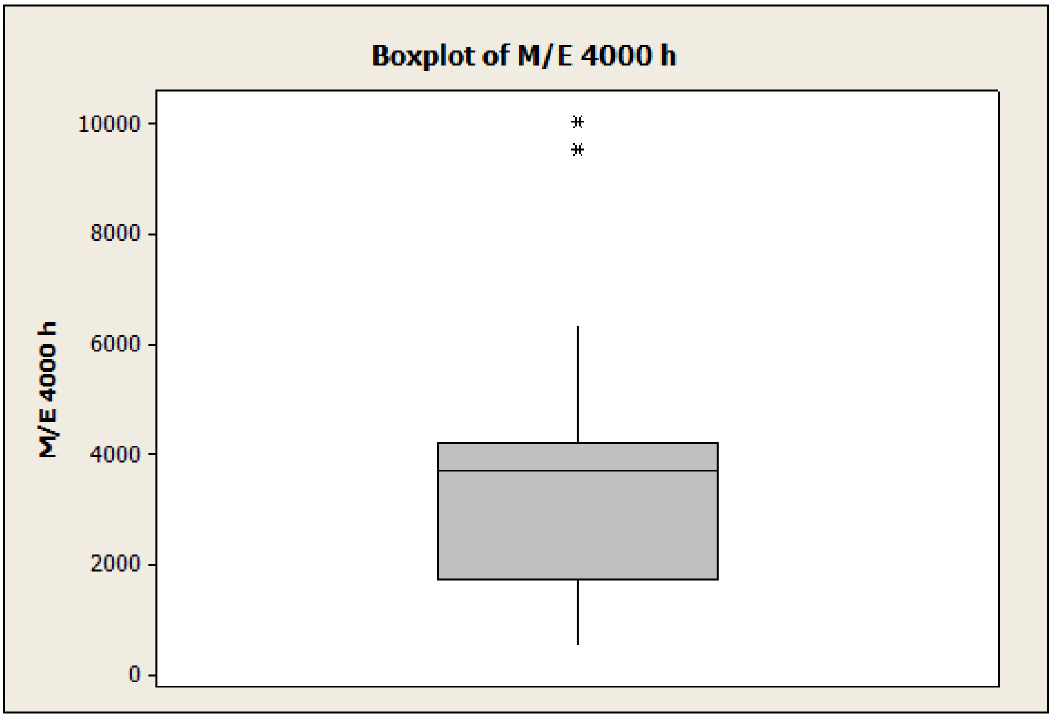


Figure 10: Boxplot of M/E 4000

In the above graph basic parameters are the following:

Table 10: Basic parameters

Median	Q1	Q3	IQR	lower outer fence	upper outer fence
3725	1744.5	4230.75	2486.25	540	6374

Two points are indicated as possible outliers. These correspond to failures in 9548 and 10042 working hours. Since there are two points in this area and after searching in the failure database, it is established that both concern failures of exhaust valves. It was decided to keep these points and continue with the analysis.

Hereby Q1, Q3 are the 25% and 75% percentile respectively. It is self understood that the median coincide with the 50% percentile of the under analysis data.

5.1.2 Basic statistics calculations

Subject data can underlie under a few basic statistics calculations such as mean, Standard Error of mean, standard deviation variance, coefficient of variation first quartile, etc.

Table 11: Basic statistics calculation

Mean	Standard error of the Mean	Standard Deviation	Variance	Coefficient of Variation	Q1	Median	Q3
3548,6	496,62	2432,9	5919158,	68,56	1744,50	3725,00	4230,7

IQR	Sum	Minimum	Maximum	Range	Sum of Squares	Skewness	Kurtosis
2486,25	85167,00	540,00	10042,00	9502,00	438366393,00	1,31	2,12

A graphical summary of the above can be plotted which provides through the histogram an approach of the appropriate distribution that describes well our data.

Also it is useful for better control and inspection to plot statistical parameters and confidence intervals.

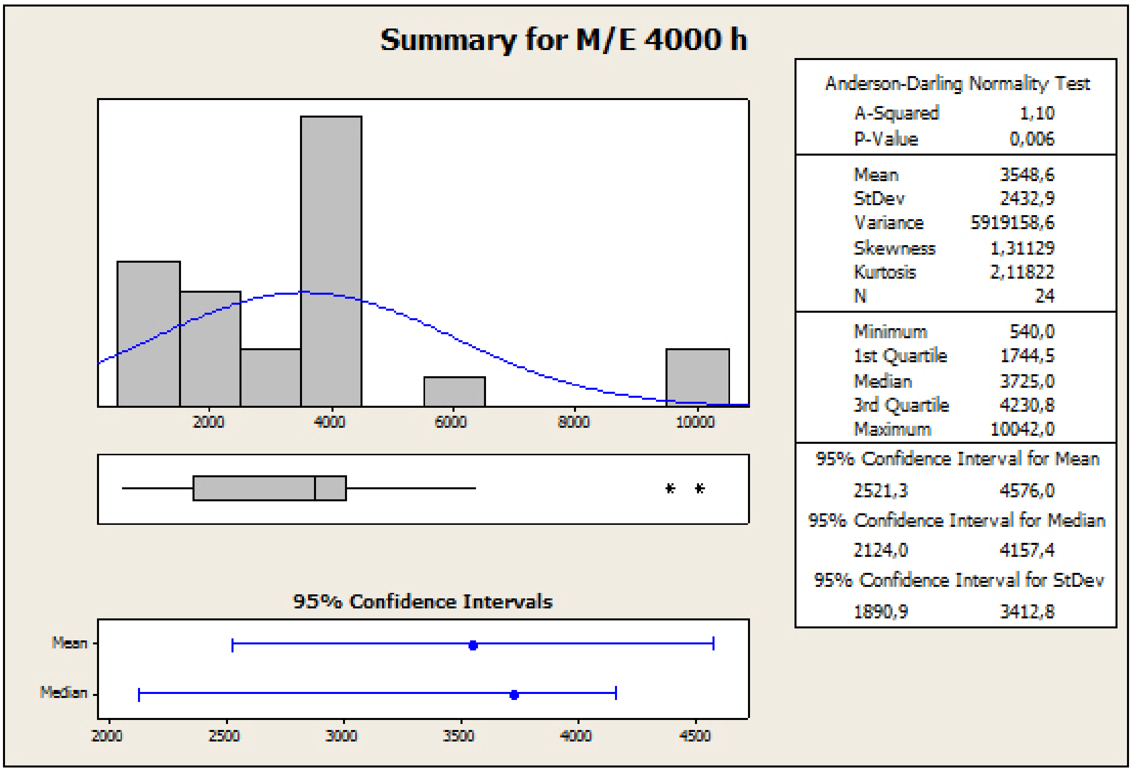


Figure 11: *Graphical summary for M/E 4000*

Additionally we can see 95% confidence intervals for the mean, median and standard deviation.

5.1.3 Histogram

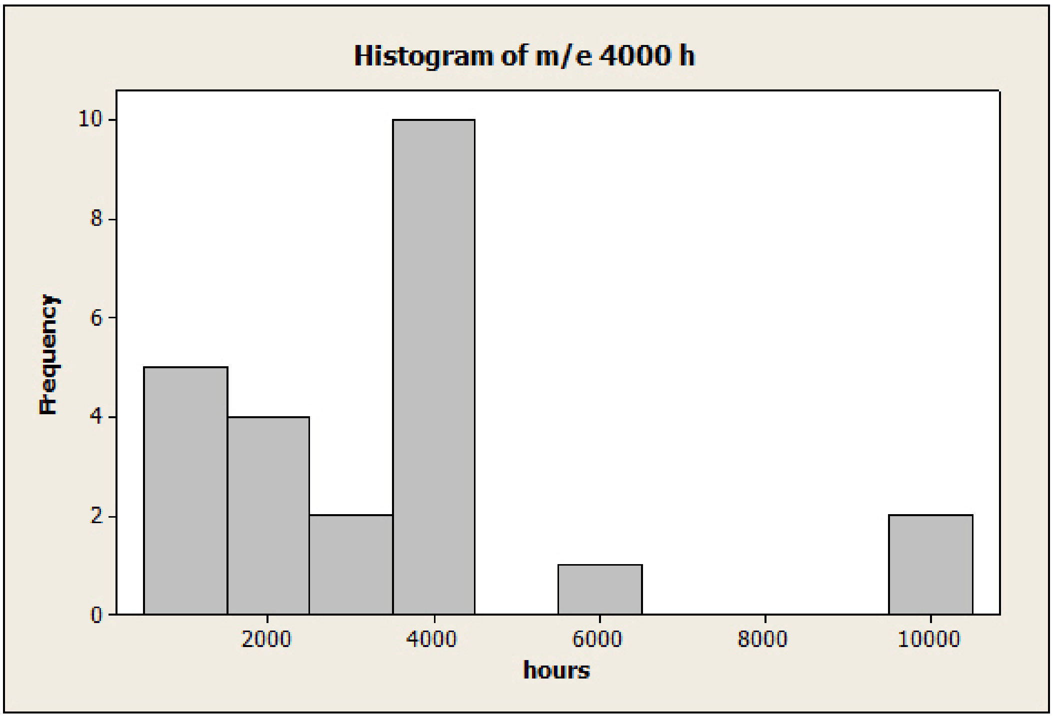


Figure 12: *Histogram of M/E 4000*

A histogram corresponds to the probability density function of a theoretical distribution and it is not as informative as a probability plot.

5.1.4 Cumulative distribution function

Next step is to calculate the empirical cumulative distribution function.

This function is associated with the empirical measure of our sample. As can be seen clearly from the plot, it is a step function that jumps for 1/n at each of the n data points.

The empirical cdf estimates the true underlying cdf of the points in the sample. The Kolmogorov Smirnov test, which was analyzed in Ch. 4.4.2 can be used to measure the discrepancy between the empirical distribution and the hypothesized distribution.

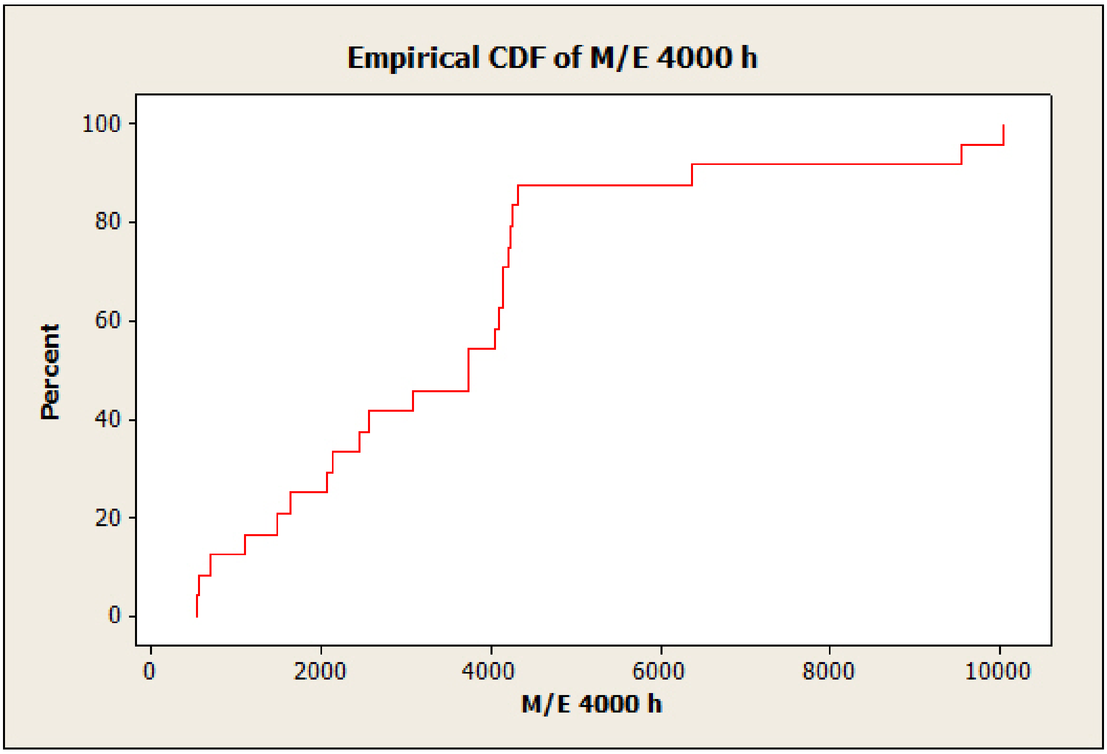


Figure 13: *Empirical CDF of M/E 4000*

From the above graph a number of useful conclusions can be extracted, since it is a consistent unbiased estimator of the population cdf.

For example in 4000 hours the 40% approximately of the Main Engine parts under consideration, will have a failure.

To evaluate further our data, a probability plot is necessary. Probability plot includes:

- An Anderson- Darling statistic and p- value test to verify whether the assumed distribution fits data well.
- Confidence intervals for estimated percentiles.

To proceed in these steps, a theoretical distribution must be fitted in the data. Then the theoretical curve estimates the population cdf.

5.1.5 Individual distribution identification

To decide which distribution follows well the failure data, we follow the method described in previous chapter. Life data are plotted in special plotting paper for each assumed distribution. The distribution plotting paper, for which data are plotted as a straight line, can be considered that fit the data.

Using the statistical software Minitab, the following graphs were obtained.

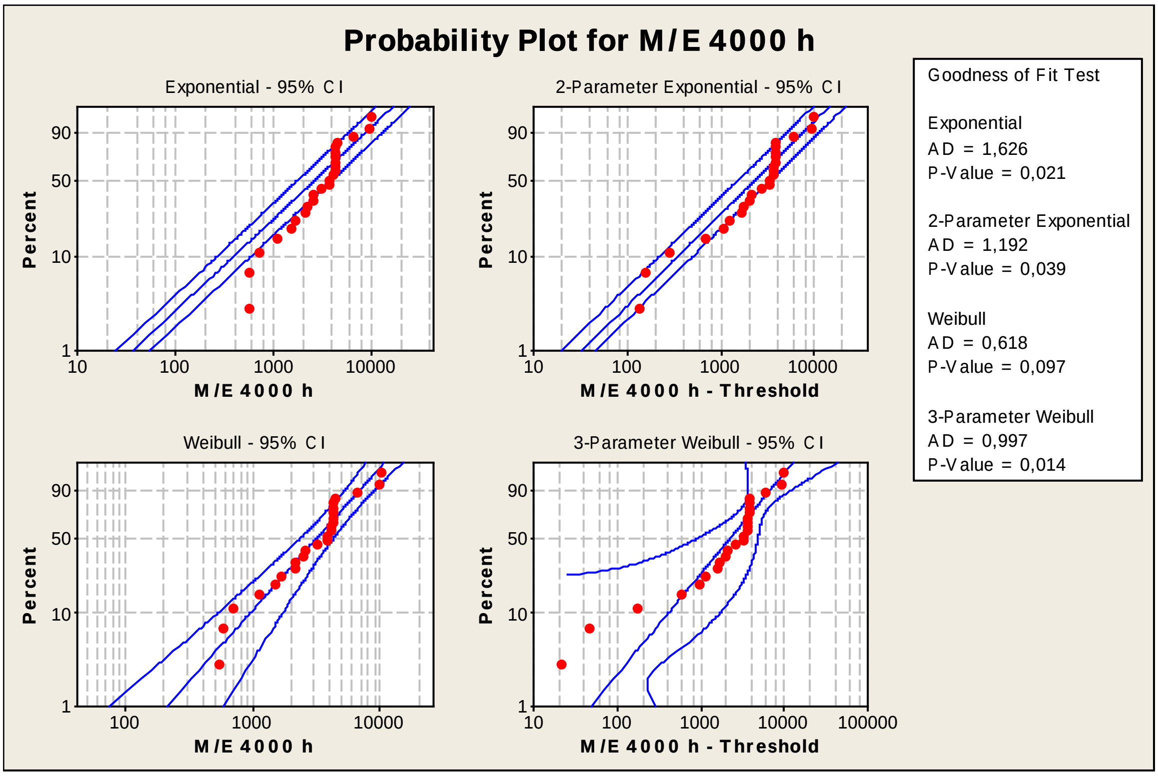


Figure 14: *Distribution identification for M/E 4000*

We have already presented an example with the use of such plots.

Beyond the good eyeball fit of the data to a straight line, more reliable results are provided through various statistical tests.

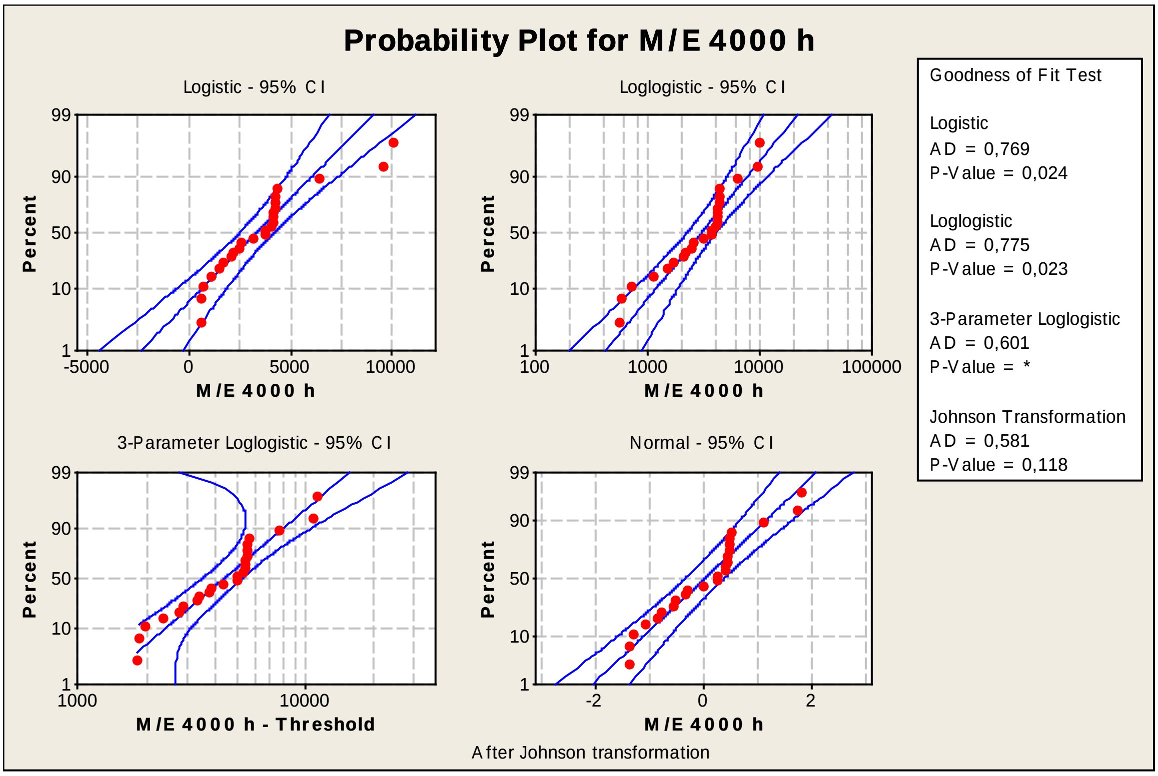


Figure 15: *Distribution identification for M/E 4000*

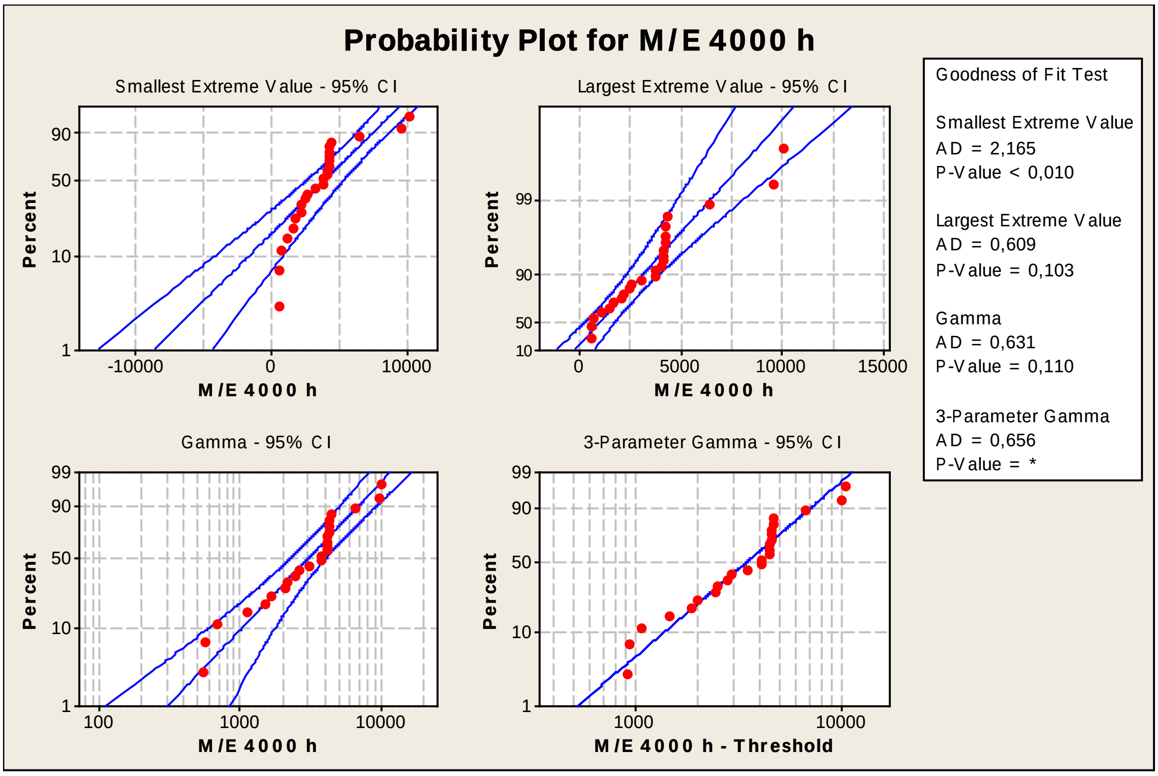


Figure 16: *Distribution identification for M/E 4000*

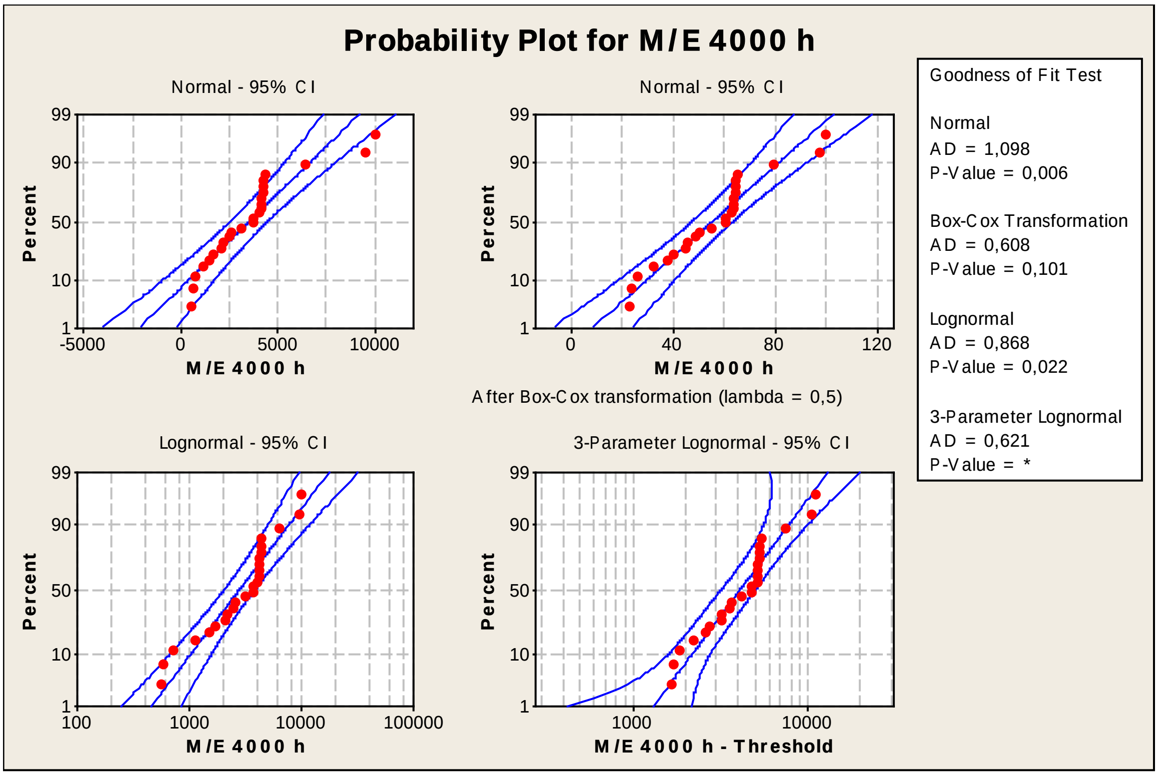


Figure 17: *Distribution identification for M/E 4000*

It is necessary to compare various goodness of fit test to decide which distribution is appropriate for this set of data.

Table 12: Comparison of goodness of fit tests for M/E 4000

Distribution	AD	P	LRT P
Normal	1,098	0,006	
Box-Cox Transformation	0,608	0,101	
Lognormal	0,868	0,022	
3-Parameter Lognormal	0,621	*	0,253
Exponential	1,626	0,021	
2-Parameter Exponential	1,192	0,039	0,015
Weibull	0,618	0,097	
3-Parameter Weibull	0,997	0,014	0,219
Smallest Extreme Value	2,165	<0,010	
Largest Extreme Value	0,625	0,103	
Gamma	0,631	0,110	
3-Parameter Gamma	0,656	*	0,185
Logistic	0,769	0,024	
Loglogistic	0,775	0,023	
3-Parameter Loglogistic	0,631	*	0,367
Johnson Transformation	0,581	0,118	

The Box-Cox transformations uses a lambda of 0.05 and the Johnson transformation function is -0,748045 + 1,29142 * Asinh((X - 1710,31) / 2304,44)
Minitab also includes a p- value for Likelihood Ratio Test (LRT), which tests whether a 2-parameter distribution would fit the data equally well compared to its 3-parameter counterpart.

Since it is preferable to use normal models instead of non-normal, previous transformations will be ignored.

Checking the Anderson Darling and p-value test, it is concluded that the Weibull distribution fits adequately well our data. In order to define subject distribution, it is necessary to estimate its parameters.

According to Nelson (17), the scale parameter α of a Weibull distribution is the 63rd percentile. Calculating the 63rd percentile, the scale parameter for this distribution is α= 3955.87.

The slope of the fitted line in a Weibull plot corresponds to the shape parameter β. Using Weibull plot, shape parameter is calculated as β=1.55328.

Summarizing, the fitted distribution for the first group of failures, is fully defined.

Weibull distribution of times to failure:

- CDF: $F(x) = 1 - \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right], 0 \leq x \leq \infty$

- PDF: $f(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right], 0 \leq x \leq \infty$

Where α=3955.87
β=1,55328.

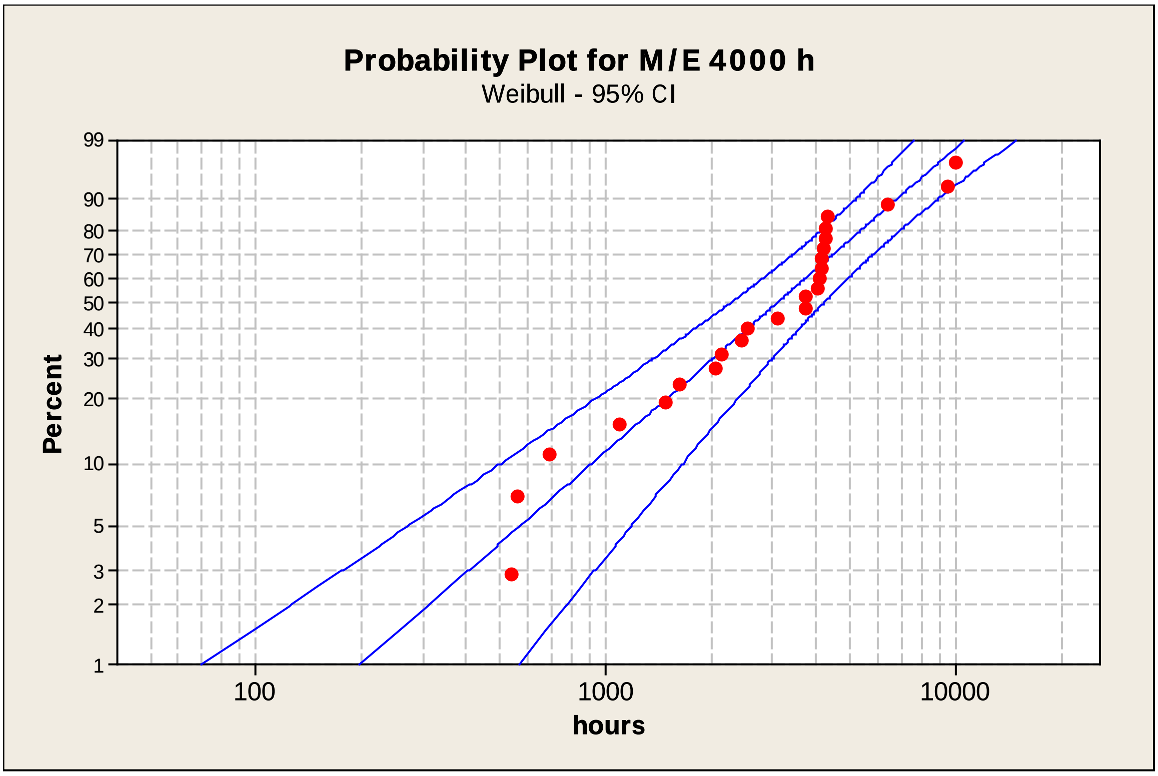


Figure 18: *Plot of Weibull fit to the data*

5.1.6 Reliability concepts

To understand clearly issues as reliability, failure rate, and survival function, the next important step is to extract relevant plots, given that the Weibull distribution fits adequately well the life data.

In addition to these plots, it is also of great assistance to extract tabulated values for possibilities of failures, etc in regards to time.

According to earlier symbols, it is clear that $F(t_2)$- $F(t_1)$ is the probability that a part survives to time t_1 but fails before t_2 . It is also the fraction of the entire population that fails in that interval.

Since it is useful to focus attention on the unfailed units, the reliability function or survival function is: R (t)= 1- F(t)

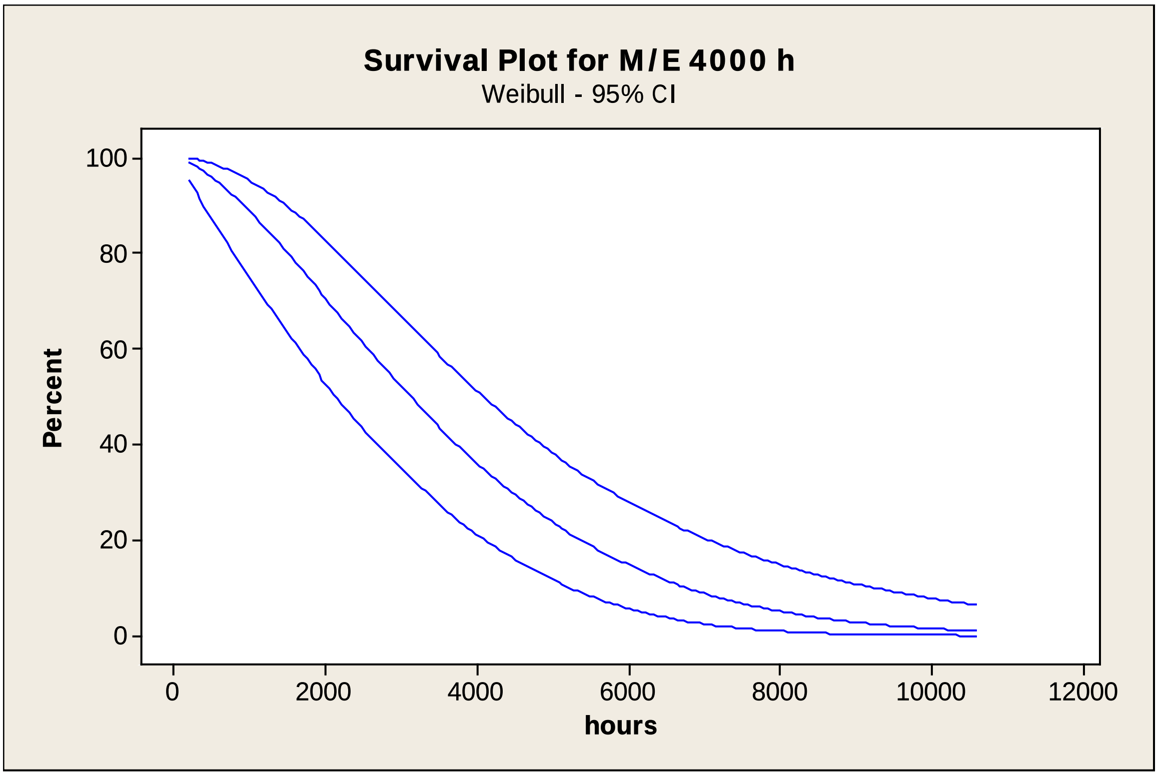


*Figure 19***:** *Plot of reliability function with 95% confidence intervals*

Table 13***: table of survival probabilities***

		95% confidence intervals	
time	probability	lower	upper
1000	0,888590645	0,75141978	0,952353538
2000	0,707046992	0,52884325	0,828089411

3000	0,521649861	0,34921057	0,668630713
4000	0,361539855	0,21244054	0,512642357
5000	0,237202679	0,11670455	0,381456047
6000	0,148100144	0,05706026	0,27977857
7000	0,088340731	0,02464218	0,203929802
8000	0,05049672	0,00937149	0,148225601
9000	0,027729244	0,00313563	0,107559004
10000	0,014658238	0,00092293	0,077943904
12000	0,003678923	5,4448E-05	0,040765882
13000	0,001751668	1,0915E-05	0,029419614
14000	0,000807534	1,9256E-06	0,021200586
15000	0,000360832	2,9891E-07	0,015255503

The table above shows the percentage of the items under investigation that will survive at a given time. The survival curve is surrounded by two outer lines - the approximate 95.0% confidence interval for the curve, which provides a range of reasonable values for the "true" survival function at each point.

The conditional survival probabilities gives the reliability for a new mission of *t* duration, having already accumulated *T* hours of operation up to the start of this new mission, and the units are checked out to assure that they will start the next mission successfully. It is called conditional because you can calculate the reliability of a new mission based on the fact that the unit or units already accumulated *T* hours of operation successfully.

For the Weibull distributed category M/E 4000 the conditional probabilities for a range of 1000 to 10000 hours is plotted in the following graph.

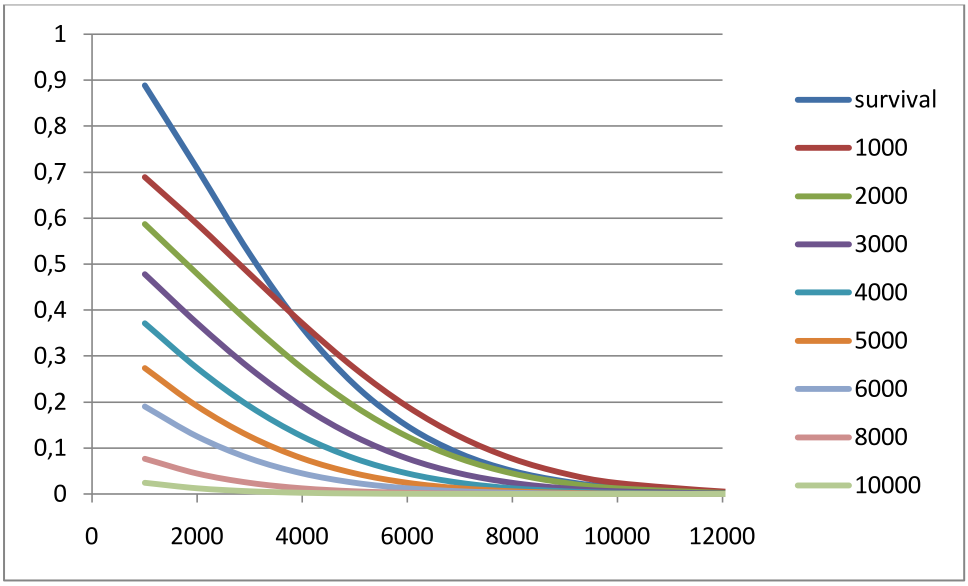


Figure 20: Conditional probabilities

For comparison the survival plot (as in Figure 24: Survival plot for M/E 4000 can be seen).

To understand deeply this figure, for equipment that has survived 4000 hours, the conditional probabilities to survive t hours are:

Table 14: Conditional probabilities for M/E 4000

t	P
1000	0,371234
2000	0,273451
3000	0,190452
4000	0,125061
5000	0,077241
6000	0,044782
8000	0,012373
10000	0,002611

For example, given that equipment has survived 4000 hours, the probability to survive additionally 2000 hours is 0.273 or 27.3%.

To visualize this example, a pie chart can be drawn which introduces the conditional probabilities, given that equipment has already survived for 4000 operational hours. Clearly can be seen the reduction of survival probabilities as the additional requested time to failure is increasing.

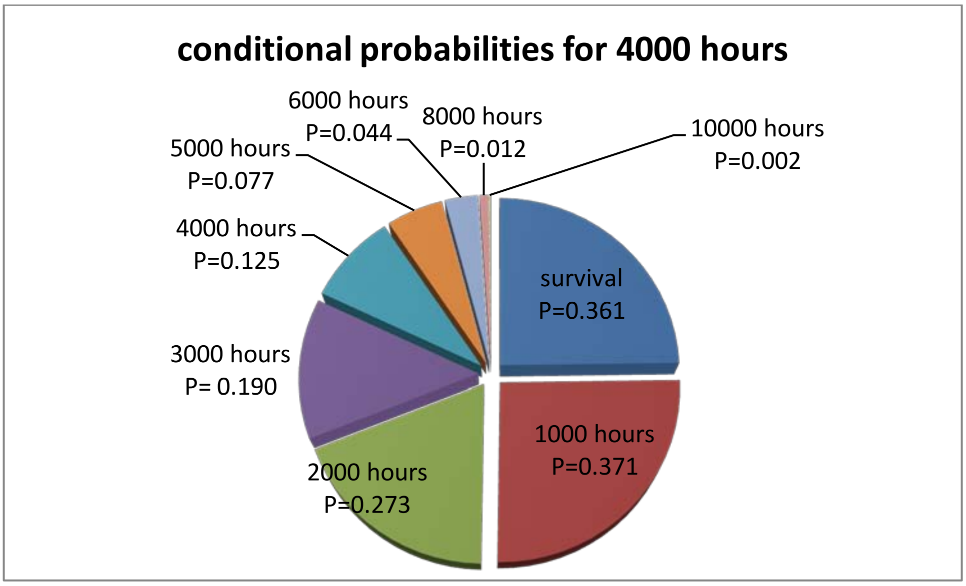


Figure 21: Conditional probabilities for M/E 4000 at 4000 hours

5.1.7 Failure rate plot

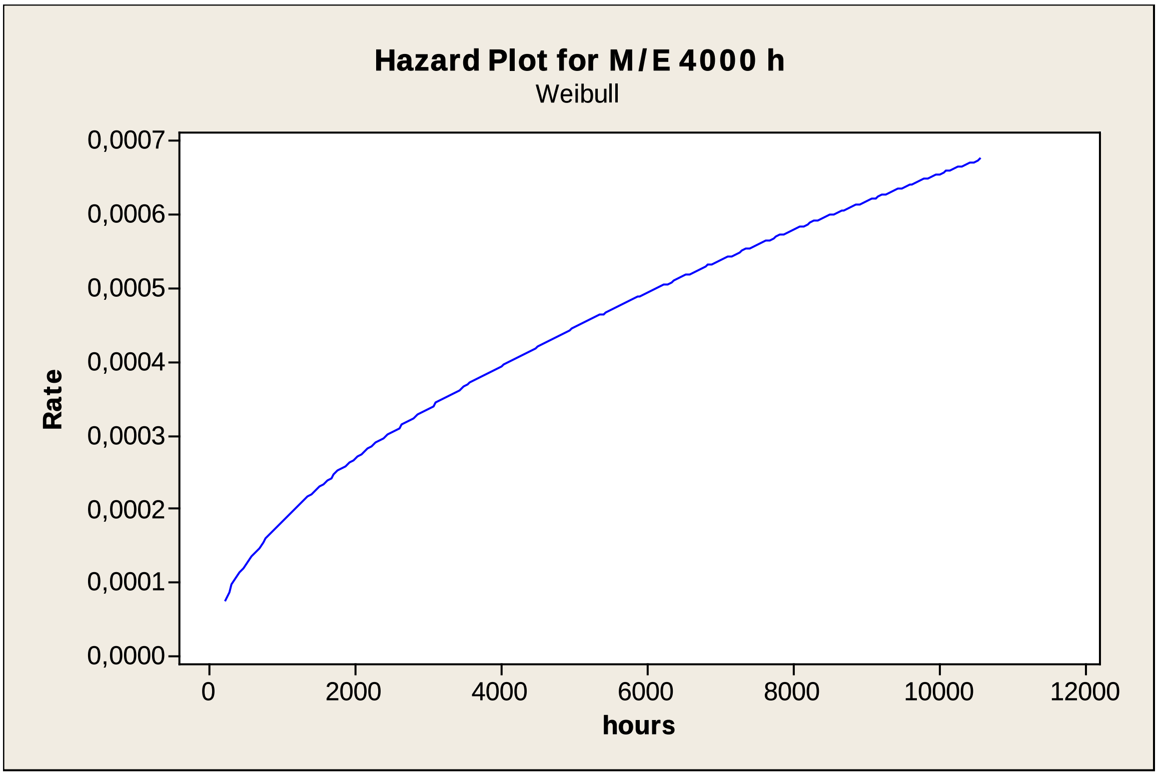


Figure 22: *Hazard plot for M/E 4000*

It is already defined in a previous chapter but it is also reminded here that the failure rate plot (hazard plot) is of great importance. The units of this rate are failures per unit time. It is the failure rate of the survivors to time t in the very next instant following t. It is not a probability and it can have values greater than 1.

In cases with constant failure plot, we can have a quick calculation of the MTTF (Mean Time To Failure) since it is the reciprocal of the assumed constant failure rate. However in Weibull distribution the failure rate is a function of time. As it was written in previous chapter, the concave shape of the curve indicates a decreasing failure rate.

5.1.8 Cumulative failure plot

We can integrate the hazard function to obtain the cumulative failure function. Each plot point represents the cumulative percentage of units failing at time t.

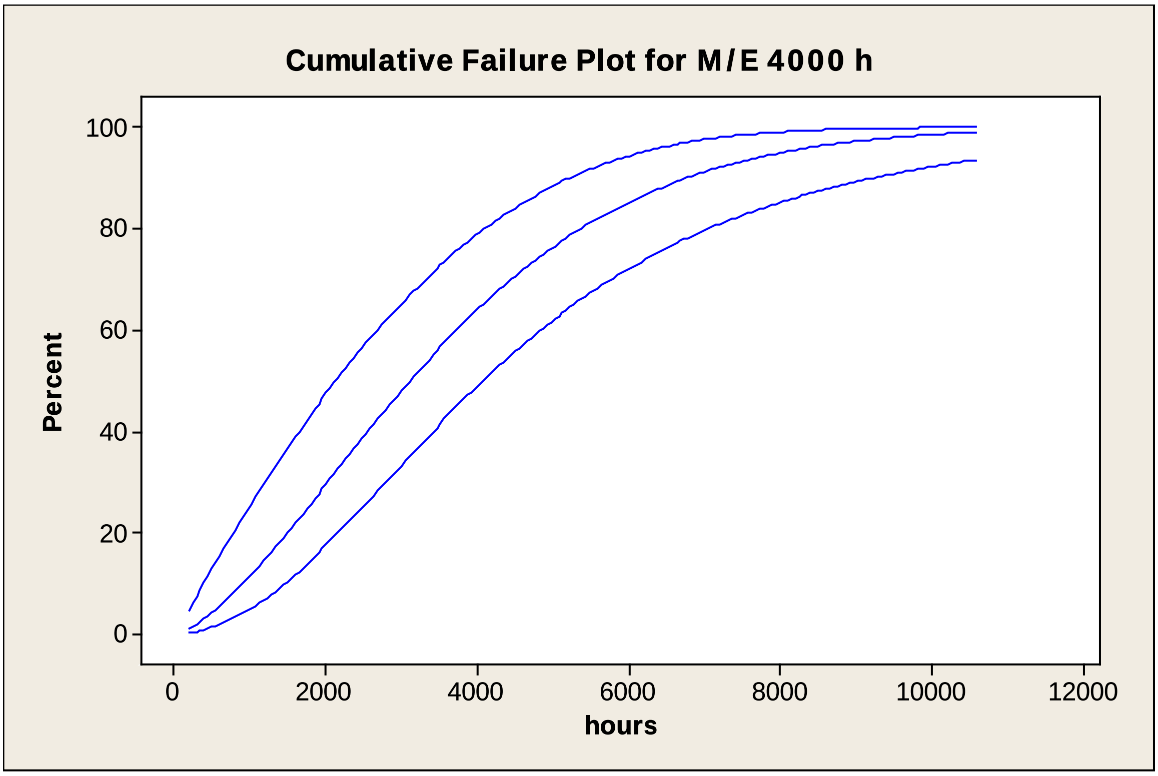


Figure 23: *Cumulative failure plot for M/E 4000*

Table 15: *table of cumulative failure probabilities*

		95% confidence intervals	
time	probability	lower	upper
1000	0,111409355	0,0476465	0,248580215
2000	0,292953008	0,1719106	0,471156745
3000	0,478350139	0,3313693	0,650789428
4000	0,638460145	0,4873576	0,787559461
5000	0,762797321	0,618544	0,883295446
6000	0,851899856	0,7202214	0,942939741
7000	0,911659269	0,7960702	0,975357825
8000	0,94950328	0,8517744	0,990628511
9000	0,972270756	0,892441	0,99686437
10000	0,985341762	0,9220561	0,999077074
12000	0,996321077	0,9592341	0,999945552
13000	0,998248332	0,9705804	0,999989085
14000	0,999192466	0,9787994	0,999998074
15000	0,999639168	0,9847445	0,999999701

The cumulative failure curve is surrounded by two outer lines - the approximate 95.0% confidence interval for the curve, which provides reasonable values for the "true" cumulative failure function.

5.1.9 Non parametric analysis

To use nonparametric methods, one does not assume a parametric form for a distribution. Nonparametric comparisons are less sensitive than parametric ones and are usually used in engineering applications because they are often adequate and yield more information for small samples.

Table 16: *Non parametric estimates*

	Standard	95,0% Normal CI	
Mean(MTTF)	Error	Lower	Upper
3548,63	496,620	2575,27	4521,98

Reliability plots can be produced also with a nonparametric approach.

A distribution-free estimate of the reliability at age y is the sample fraction that survive an age y. that is, if X of the n times to failure are beyond age y, then the estimate of reliability is:

$$R(y)=X/n$$

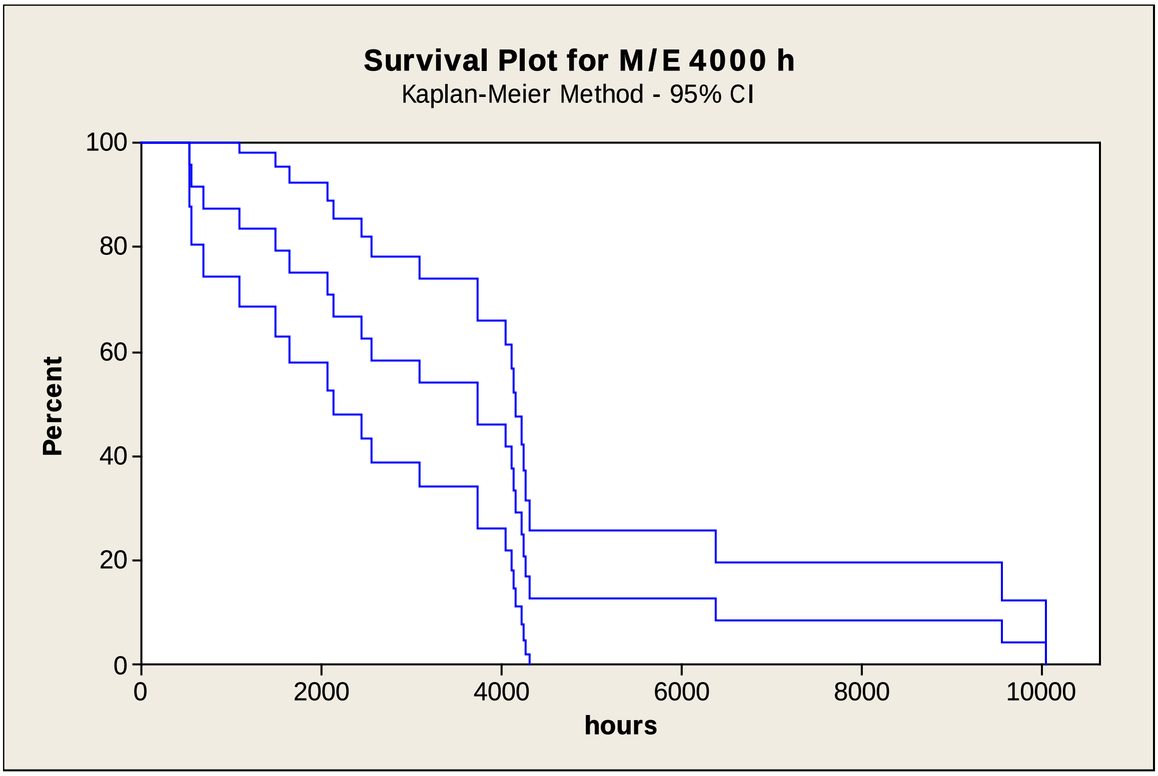


Figure 24: *Survival plot for M/E 4000*

Table 17**:** *kaplan meier estimates for survival probabilities*

		95% confidence intervals	
time	survival probability	lower	upper
540	0,95833	0,878388	1,000000
564	0,91667	0,806092	1,000000
688	0,87500	0,742687	1,000000
1092	0,83333	0,684234	0,982433
1487	0,79167	0,629189	0,954144
1634	0,75000	0,576762	0,923238
2076	0,70833	0,526487	0,890180
2134	0,66667	0,478069	0,855264
2447	0,62500	0,431314	0,818686
2553	0,58333	0,386093	0,780573
3088	0,54167	0,342324	0,741009
3725	0,45833	0,258991	0,657676
4043	0,41667	0,219427	0,613907
4102	0,37500	0,181314	0,568686
4137	0,33333	0,144736	0,521931
4146	0,29167	0,109820	0,473513
4212	0,25000	0,076762	0,423238
4237	0,20833	0,045856	0,370811
4256	0,16667	0,017567	0,315766
4317	0,12500	0,000000	0,257313
6374	0,08333	0,000000	0,193908
9548	0,04167	0,000000	0,121612
10042	0,00000	0,000000	0,000000

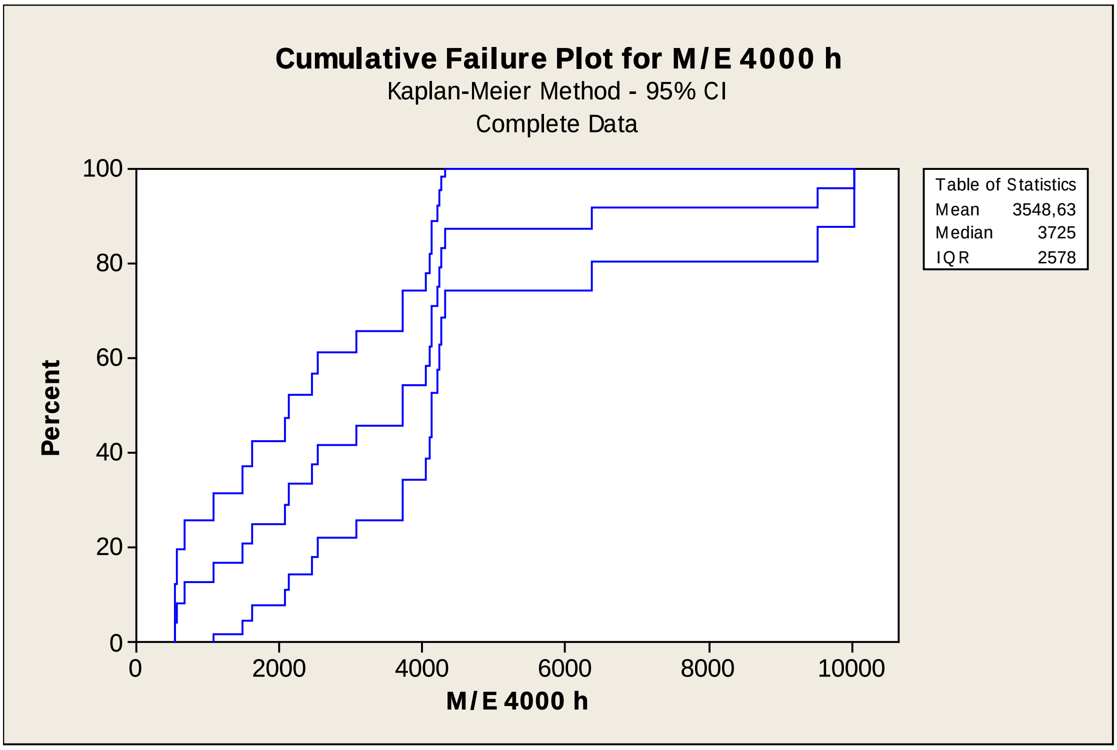


Figure 25: *Cumulative failure plot for M/E 4000*

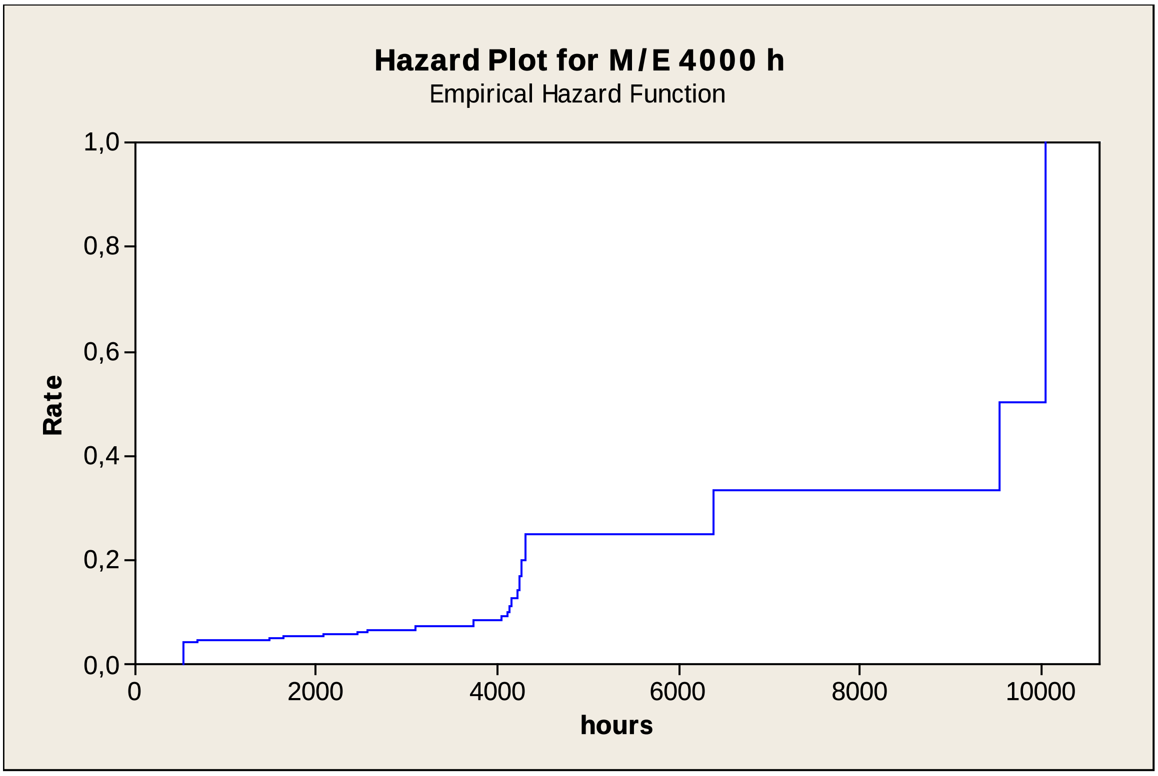


Figure 26: *Hazard plot for M/E 4000*

Table 18: cumulative failure probabilities and hazard rates

		95% confidence intervals		
time	cumulative failure probability	lower	upper	hazard rates
540	0,04167	0,000000	0,121612	0,04167
564	0,08333	0,000000	0,193908	0,04348
688	0,12500	0,000000	0,257313	0,04545
1092	0,16667	0,017567	0,315766	0,04762
1487	0,20833	0,045856	0,370811	0,05000
1634	0,25000	0,076762	0,423238	0,05263
2076	0,29167	0,109820	0,473513	0,05556
2134	0,33333	0,144736	0,521931	0,05882
2447	0,37500	0,181314	0,568686	0,06250
2553	0,41667	0,219427	0,613907	0,06667
3088	0,45833	0,258991	0,657676	0,07143
3725	0,54167	0,342324	0,741009	0,08333
4043	0,58333	0,386093	0,780573	0,09091
4102	0,62500	0,431314	0,818686	0,10000
4137	0,66667	0,478069	0,855264	0,11111
4146	0,70833	0,526487	0,890180	0,12500
4212	0,75000	0,576762	0,923238	0,14286
4237	0,79167	0,629189	0,954144	0,16667
4256	0,83333	0,684234	0,982433	0,20000
4317	0,87500	0,742687	1,000000	0,25000
6374	0,91667	0,806092	1,000000	0,33333
9548	0,95833	0,878388	1,000000	0,50000
10042	1,00000	1,000000	1,000000	1,00000

Another advantage of the non parametric method is that we can approach the Mean Residual Life of the machinery under investigation (24).

MRL of an item at age t, is the expected remaining life of the item. It is sometimes of interest to study the function

$$g(t)=MRL(t)/MTTF$$

when an item has survived up to time t, then g(t) gives the MRL(t) as a percentage of the initial MTTF.

It is another approach/ expression of the failure rate function, as can be seen in the application of this function on the real life data.

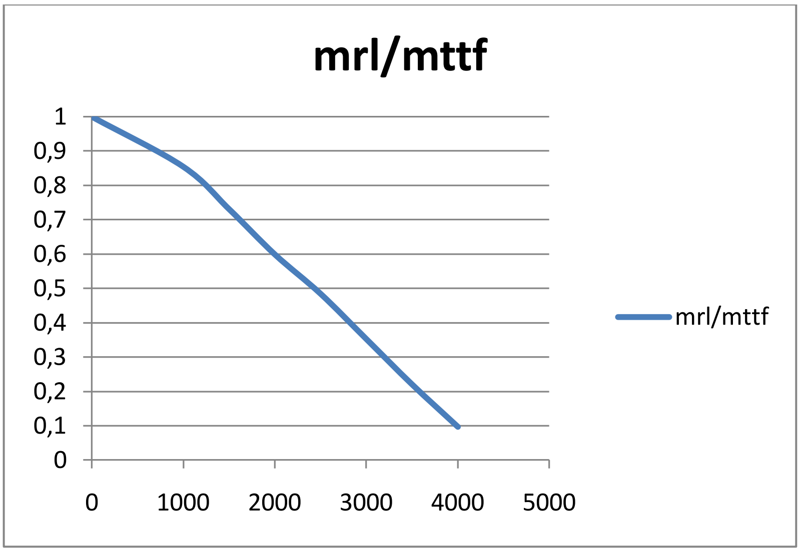


Figure 27: MRL/MTTF for M/E 4000h

In full compliance with the failure rate function, at every instant of time the remaining life of the items is decreasing. For example in 3000 hours is 35% of mean residual life at time 0.

5.2 M/E 8000 hours

Following exactly the same procedure, the data that concerns failures of Main Engine that the manufacturer suggests 8000 working hours between inspections, can be analyzed.

5.2.1 Outlier points

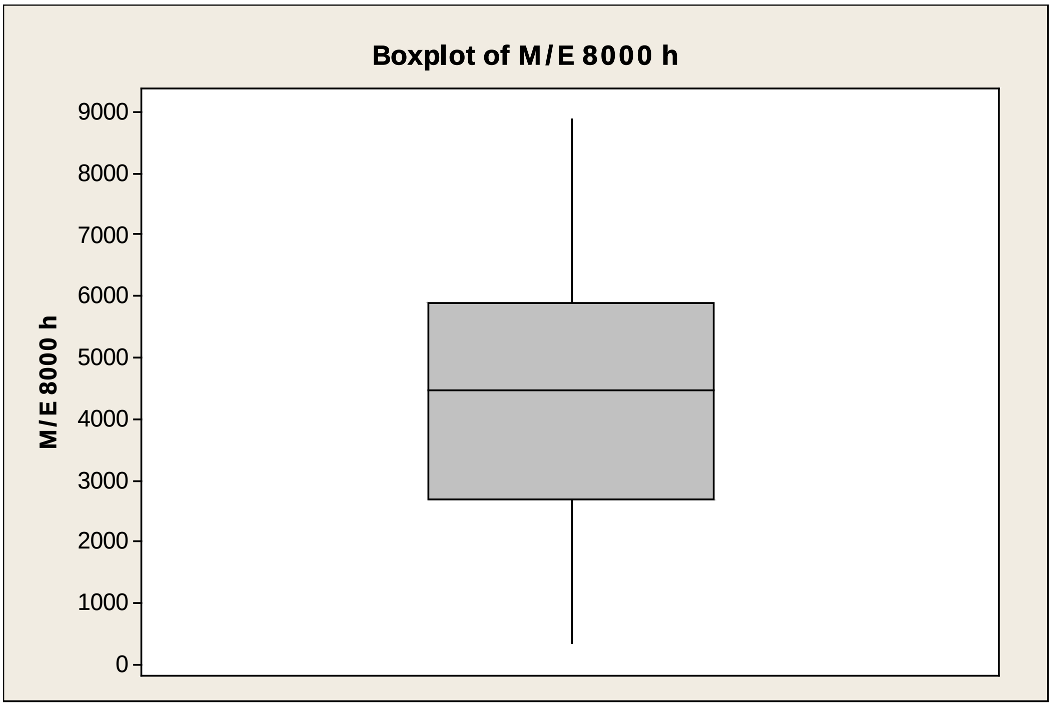


*Figure 28***:** *Boxplot of M/E 8000h*

Regarding this group of data, we can a priori exclude the measurement of 131399 since it is obvious an outlier point. It regards time to failure of one cylinder liner.

Table 19: *Basic parameters*

Median	Q1	Q3	IQR	lower outer fence	upper outer fence
4480	2685	5887	3202	357	8864

5.2.2 Basic statistics calculations

Some basic statistics calculations such as mean, Standard Error of mean, standard deviation variance, coefficient of variation first quartile, etc can be found in table 5.14

Table 20: Basic statistics calculation

Mean	Standard error of Mean	Standard Deviation	Variance	Coefficient of Variation	Q1	Median	Q3
4474,55	326,86	2143,4	4594262,2	47,9024505	2685	4480	5887
IQR	Sum	Minimum	Maximum	Range	Sum of Squares	Skewness	Kurtosis
3202	192406	357	8864	8507	1,05E+09	0,07319	-0,9168

A graphical summary of the above can be found in figure 5.13

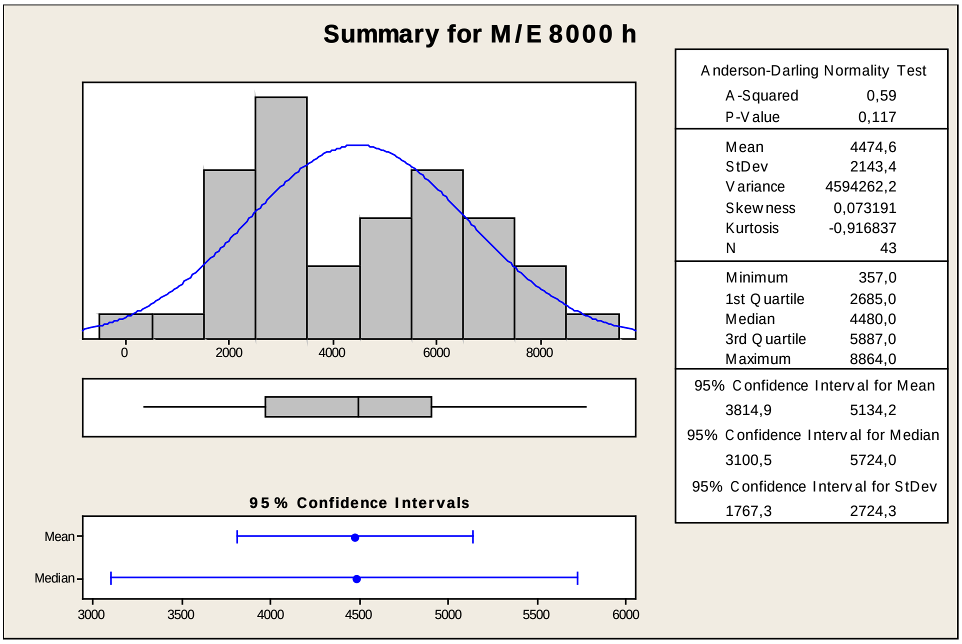


Figure 29: *Graphical summary for M/E 8000h*

As it was explained, the software package Minitab, that was used to produce these plots, assumes that the under examination data follows normal distribution and based on this assumption are calculated parameters as the mean, median, etc. As we proceed with our analysis, the proper distribution that fits adequately well our model will be used; therefore a possible variance between subject parameters is expected.

5.2.3 Histogram

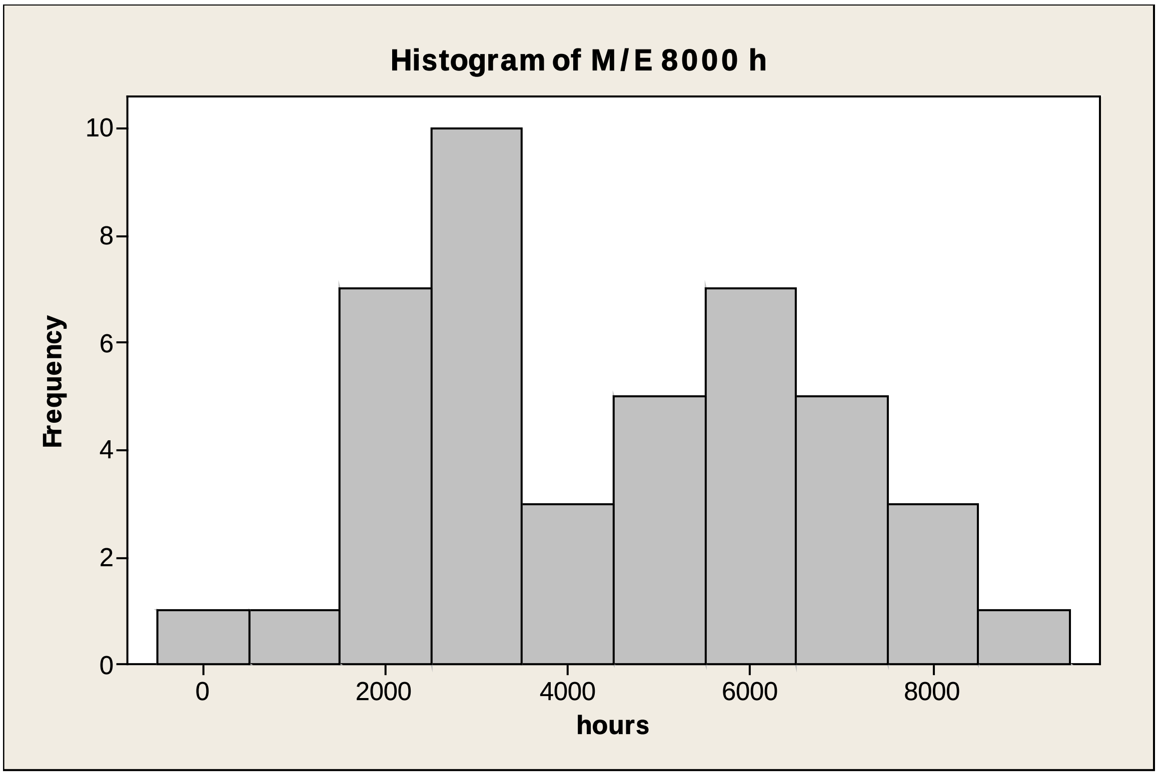


Figure 30: *Histogram of M/E 8000h*

5.2.4 Cumulative distribution function

As described in previous chapter, the empirical cdf is an estimator of the cumulative distribution function. To further define the proper distribution which fits the data well, these have to be plotted in special plotting paper.

Therefore an indication of the percentage of items that have failed until time t can be obtained from Figure 31.

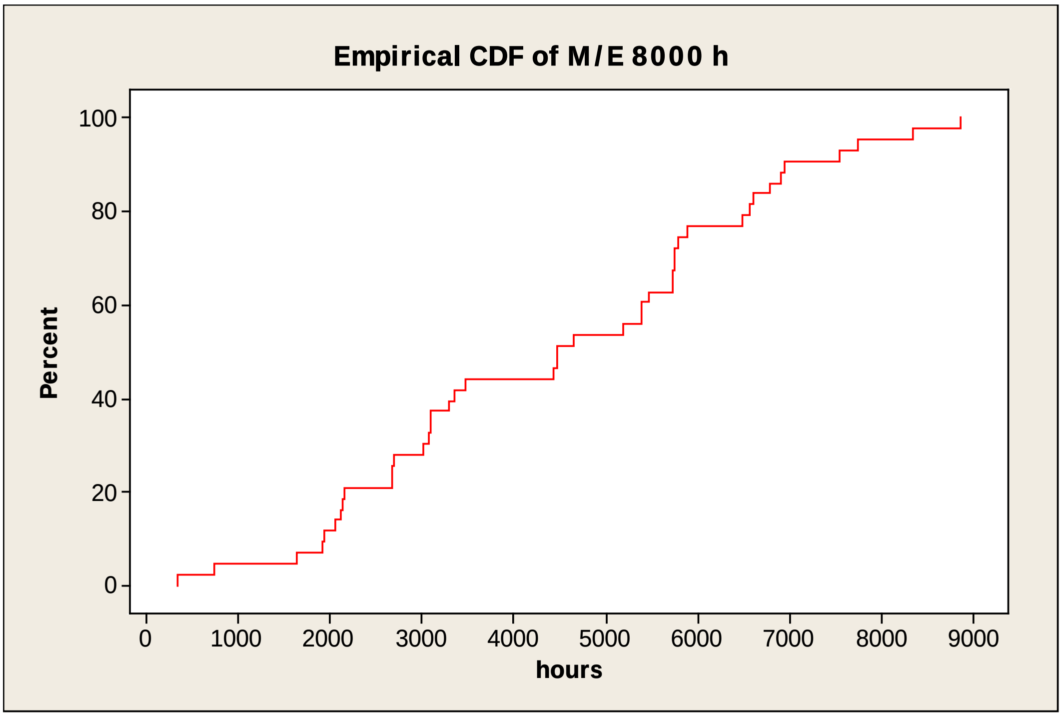


Figure 31: *Empirical CDF of M/E 8000h*

5.2.5 Individual distribution identification

Using the statistical software Minitab, the following graphs were obtained.

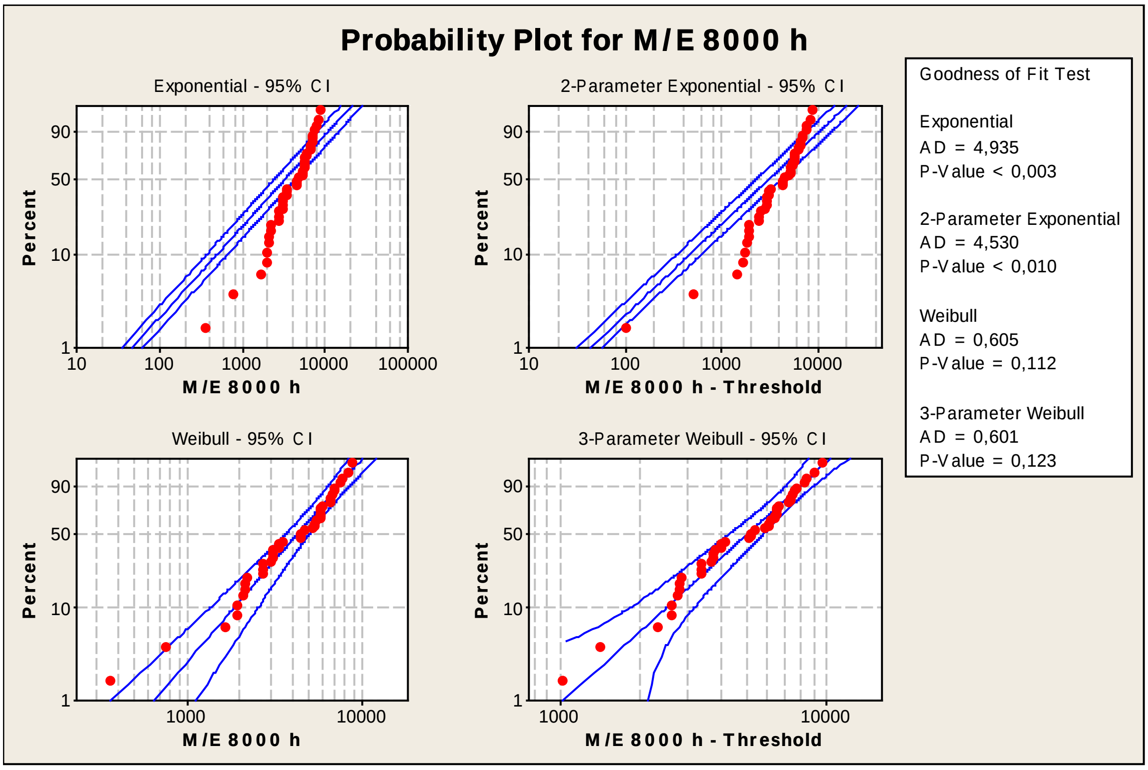


Figure 32: *Distribution identification for M/E 8000h*

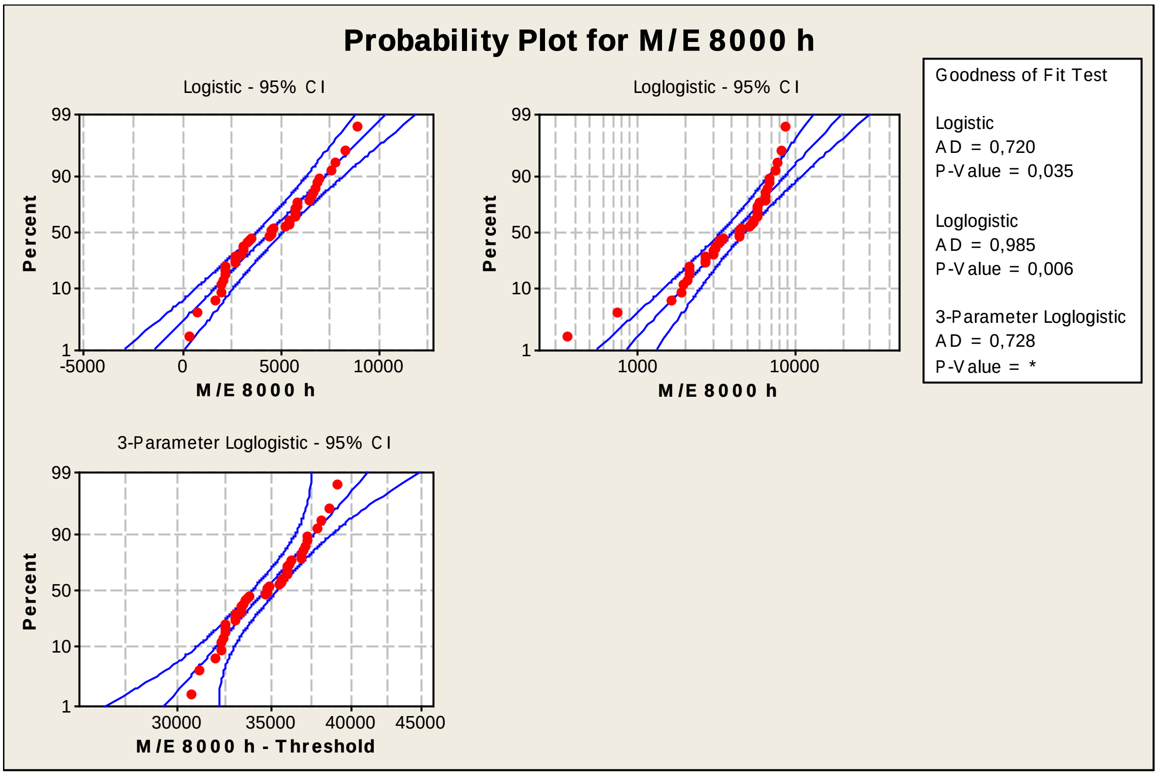


Figure 33: *Distribution identification for M/E 8000h*

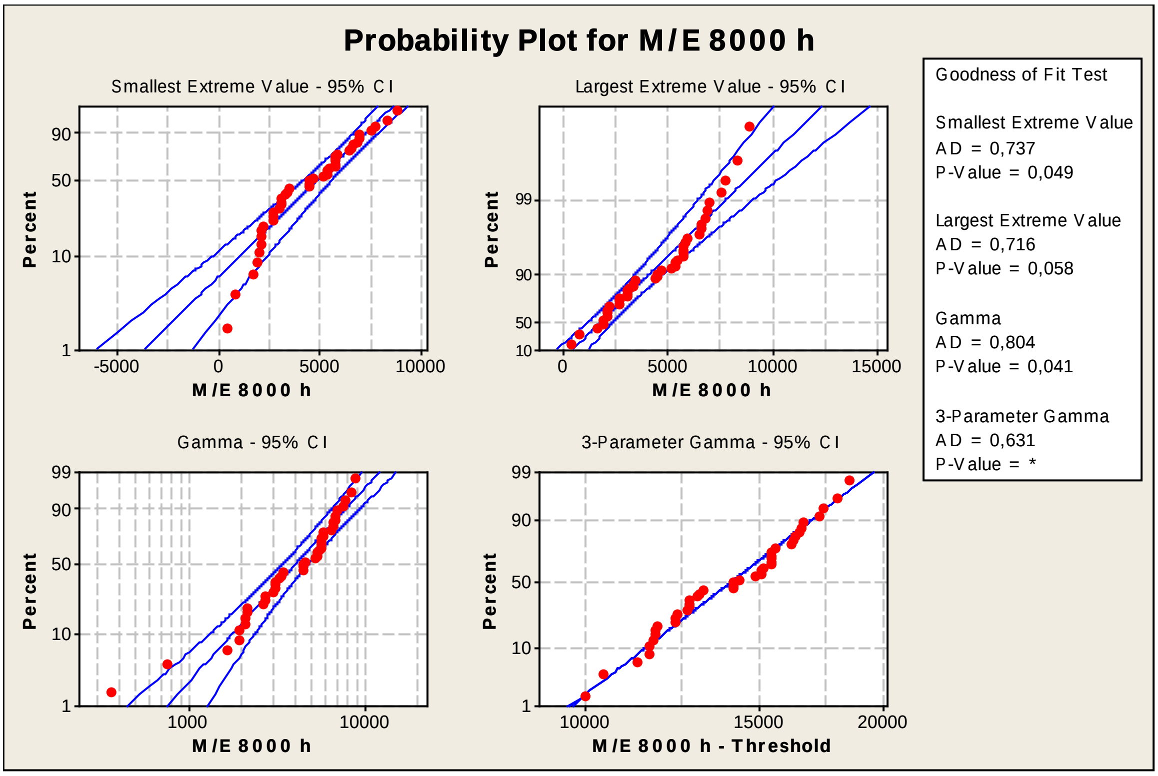


Figure 34: *Distribution identification for M/E 8000h*

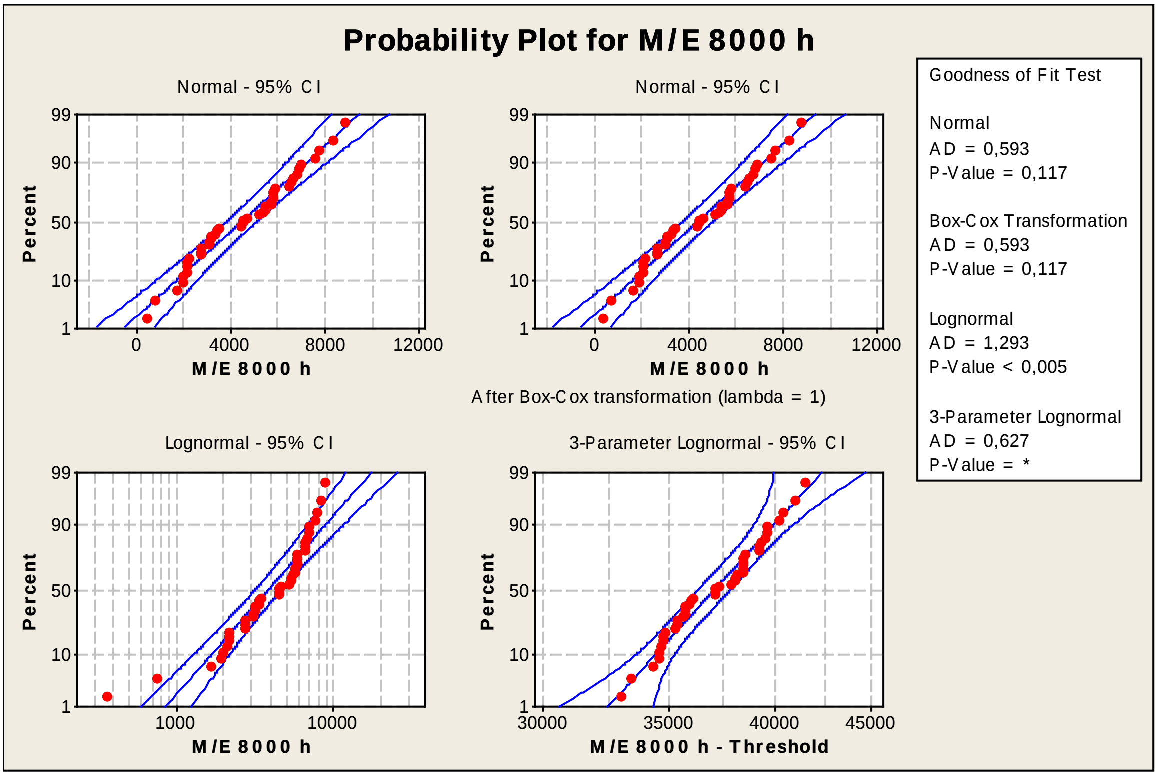


Figure 35: *Distribution identification for M/E 8000h*

It is necessary to compare various goodness of fit test to decide which distribution is appropriate for this set of data.

Table 21: *Comparison of goodness of fit tests for M/E 8000h*

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	0,593	0,117	
Box-Cox Transformation	0,593	0,117	
Lognormal	1,293	<0,005	
3-Parameter Lognormal	0,627	*	0,000
Exponential	4,935	<0,003	
2-Parameter Exponential	4,530	<0,010	0,024
Weibull	0,605	0,112	
3-Parameter Weibull	0,601	0,123	0,437
Smallest Extreme Value	0,737	0,049	
Largest Extreme Value	0,716	0,058	
Gamma	0,804	0,041	
3-Parameter Gamma	0,631	*	0,062
Logistic	0,720	0,035	
Loglogistic	0,985	0,006	
3-Parameter Loglogistic	0,728	*	0,019

Checking the Anderson Darling and p-value test, it is concluded that the normal distribution fits adequately well our data.

In order to define subject distribution it is necessary to estimate its parameters.

For the Normal distribution we need the mean and a standard deviation.

For this distribution:

Mean = 4474.55

Standard deviation = 2118.35

Summarizing, the fitted distribution for the first group of failures, is fully defined.

Normal distribution of times to failure:

- CDF: $F(x) = \frac{1}{2}[1 + \text{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)], x \in R$

- PDF:$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in R$

Where μ=4474,55
σ=2118,35

where erf is a function sometimes called the error function which can't be expressed in terms of finite additions, subtractions, multiplications, and root extractions, and so must be either computed numerically or otherwise approximated.

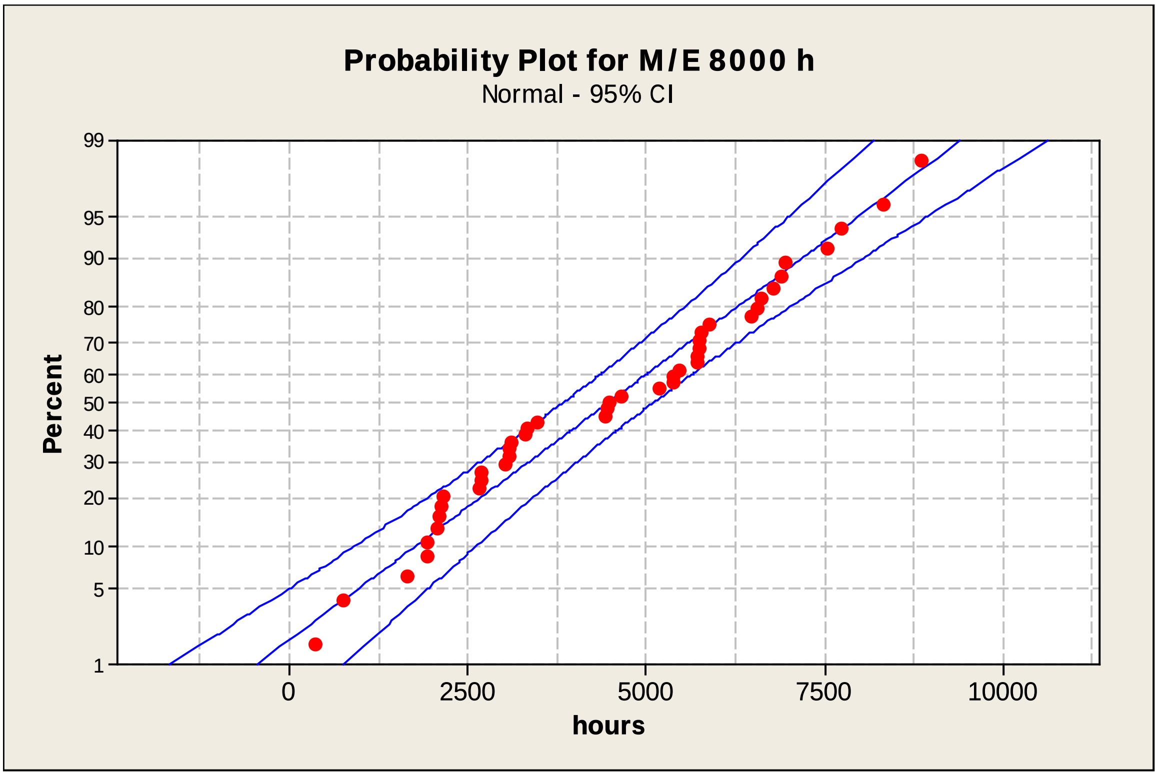


Figure 36: *Plot of Normal fit to the data*

As it was described earlier, because of the statistical nature of the process, it is necessary to use statistical intervals, which gives the range of plausible values for the process parameters based on the data and the underlying assumptions about the process. However, the intervals cannot always be guaranteed to include the true process parameters and still be narrow enough to be useful. Instead the intervals have a probabilistic interpretation that guarantees coverage of the true process parameters a specified proportion of the time. In order for these intervals to truly have their specified probabilistic interpretations, the form of the distribution of the random errors must be known. Although the form of the probability distribution must be known, the parameters of the distribution can be estimated from the data.

5.2.6 Reliability concepts

Relevant plots for this group of data can be produced, similarly to previous chapter.

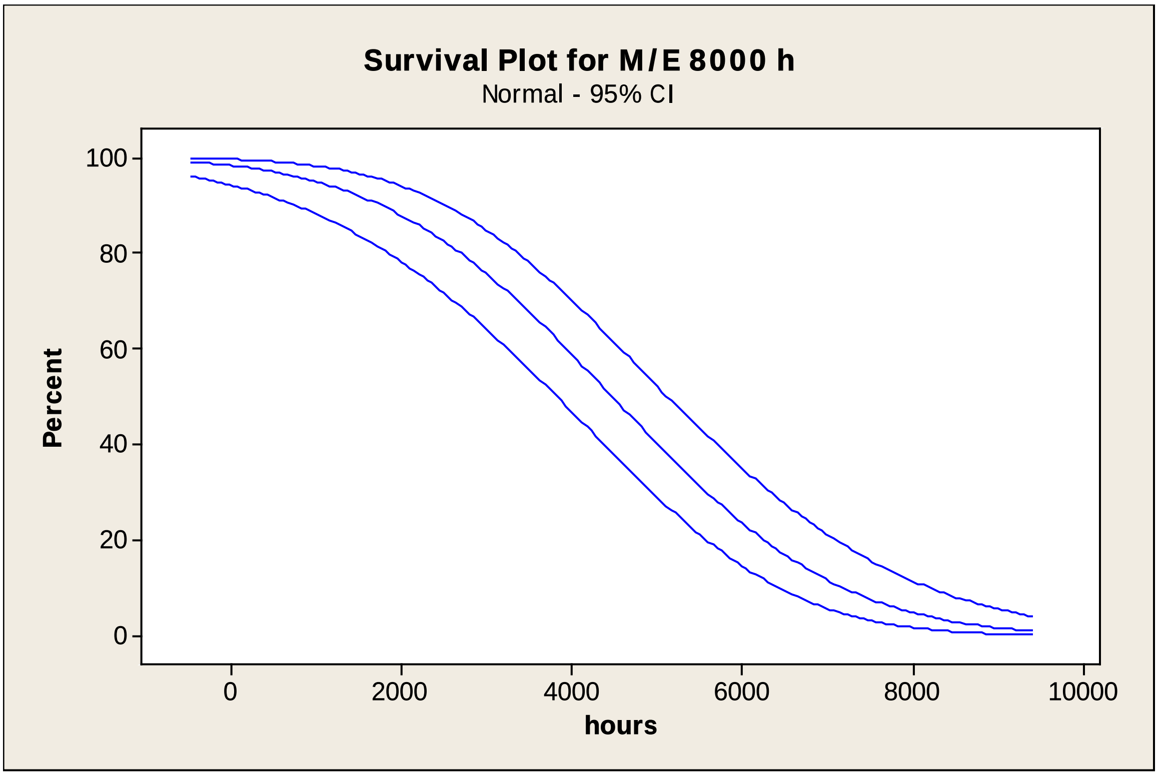


Figure 37: Plot of reliability function with 95% confidence intervals

Table 22*: table of survival probabilities*

		95% confidence intervals	
time	probability	lower	upper
1000	0,94951996	0,881495738	0,982044631
2000	0,878627271	0,782445855	0,940125005
3000	0,756812863	0,641679284	0,84831259
4000	0,588629998	0,468676873	0,700778455
5000	0,402050703	0,290646555	0,522094711

6000	0,235729394	0,1455946	0,35023035
7000	0,116596822	0,05669969	0,21149406
8000	0,048032508	0,01675938	0,114553816
9000	0,016327395	0,003705657	0,055375121
10000	0,004548767	0,000607245	0,023779993
12000	0,000190797	6,48481E-06	0,003031298
13000	2,85408E-05	4,19365E-07	0,000895117
14000	3,452E-06	1,97722E-08	0,000232299
15000	3,37031E-07	6,78394E-10	5,29068E-05

The table above shows the percentage of the items under investigation that will survive at a given time. The survival curve is surrounded by two outer lines - the approximate 95.0% confidence interval for the curve, which provides a range of reasonable values for the "true" survival function at each point.

The conditional survival probabilities gives the reliability for a new mission of *t* duration, having already accumulated *T* hours of operation up to the start of this new mission, and the units are checked out to assure that they will start the next mission successfully. It is called conditional because you can calculate the reliability of a new mission based on the fact that the unit or units already accumulated *T* hours of operation successfully.

For the Normal distributed category M/E 8000 the conditional probabilities for a range of 1000 to 10000 hours is plotted in the following graph.

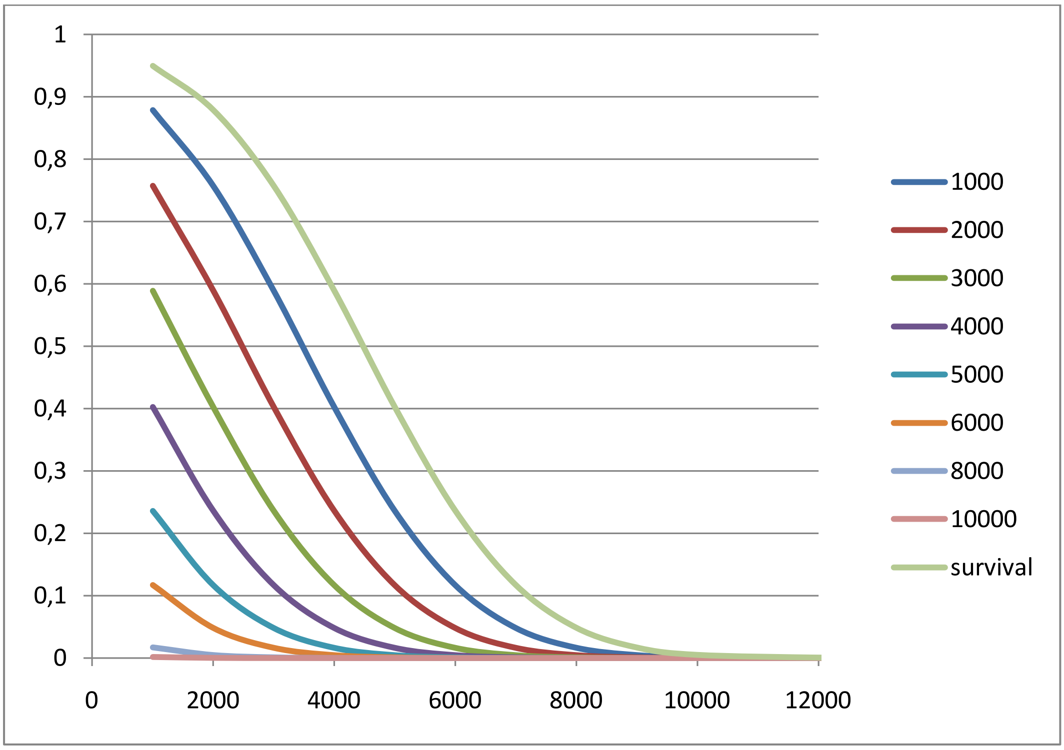


Figure 38: Conditional probabilities for M/E 8000h

To understand deeply this figure, for equipment that has survived 4000 hours, the conditional probabilities to survive t hours are:

Table 23:Conditional probabilities at 4000 hours

t	P
1000	0,402049087
2000	0,235727912
3000	0,116595754
4000	0,048031899
5000	0,016327121
6000	0,004548669
8000	0,000190791
10000	3,45184E-06

To visualize this table the following graph can be obtained.

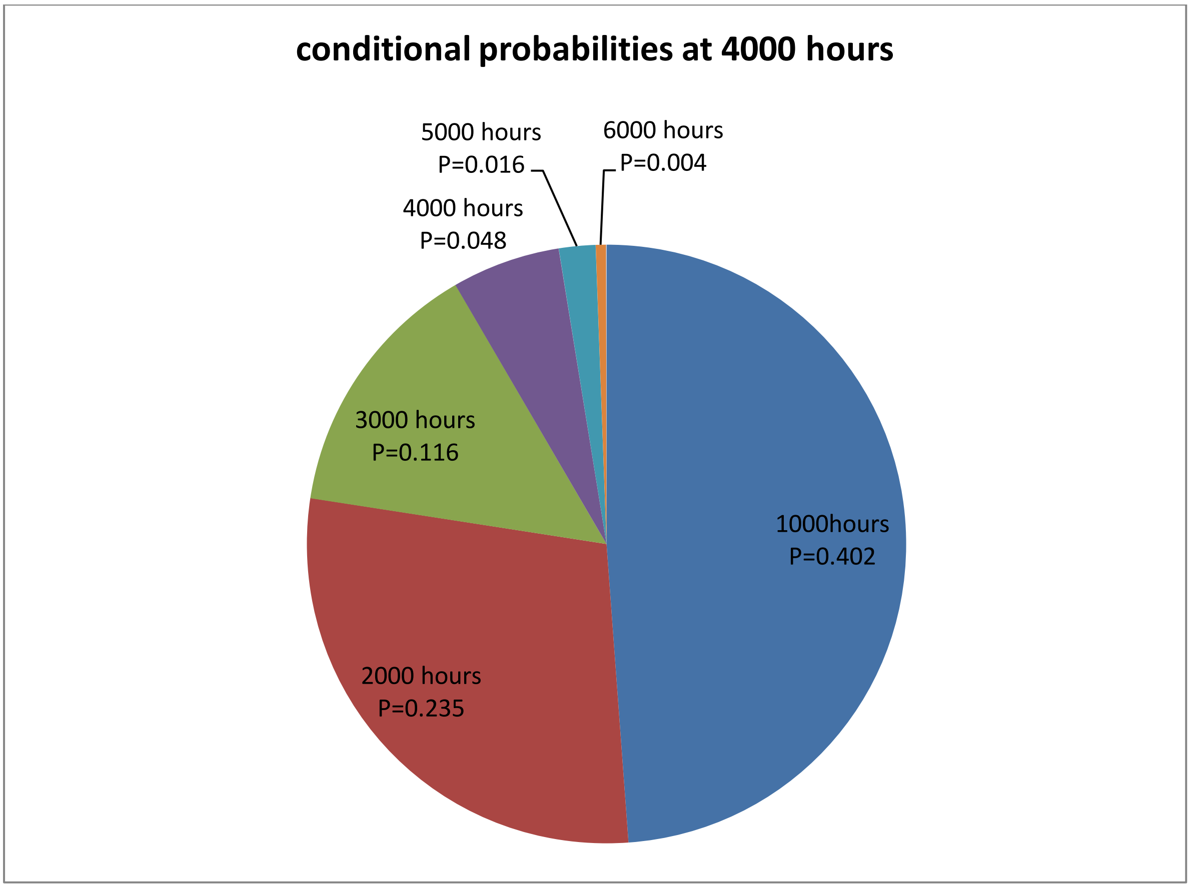


Figure 39: conditional probabilities at 4000 hours

5.2.7 Failure rate plot

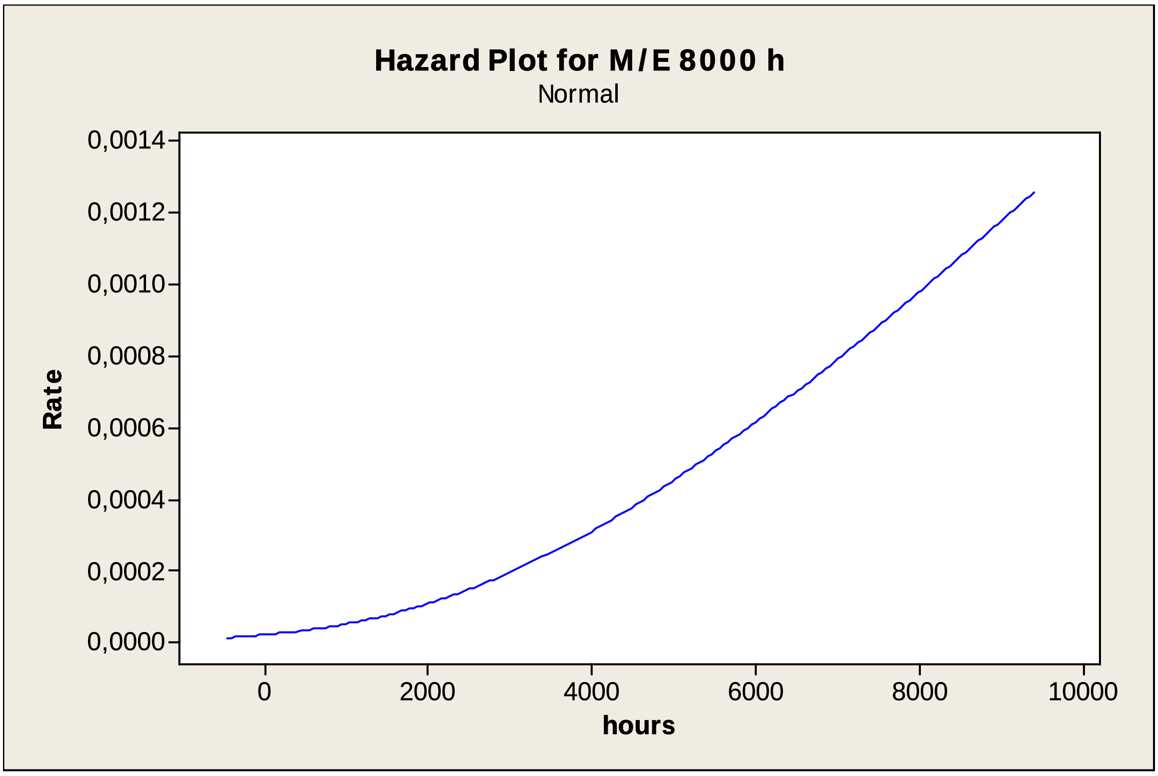


Figure 40: *Hazard plot for M/E 8000h*

It is reminded that regarding the shape of the failure rate curve is that it determines the frequency of failures. Nelson (25) suggested that if the failure rate curve is convex, then there will be an increasing failure rate life distribution. Correspondingly if the failure rate curve is concave, then there will be a decreasing failure rate life distribution.

5.2.8 Cumulative failure plot

The following table, indicates the percentage of items that will fail until time t. For example, regarding this group of data, at 4000 hours the 41.1% of the items will have failed.

Table 24*: table of cumulative failure probabilities*

		95% confidence intervals	
time	probability	lower	upper
1000	0,05048004	0,017955369	0,118504262
2000	0,121372729	0,059874995	0,217554145
3000	0,243187137	0,15168741	0,358320716
4000	0,411370002	0,299221545	0,531323127
5000	0,597949297	0,477905289	0,709353445
6000	0,764270606	0,64976965	0,8544054

7000	0,883403178	0,78850594	0,94330031
8000	0,951967492	0,885446184	0,98324062
9000	0,983672605	0,944624879	0,996294343
10000	0,995451233	0,976220007	0,999392755
12000	0,999809203	0,996968702	0,999993515
13000	0,999971459	0,999104883	0,999999581
14000	0,999996548	0,999767701	0,99999998
15000	0,999999663	0,999947093	0,999999999

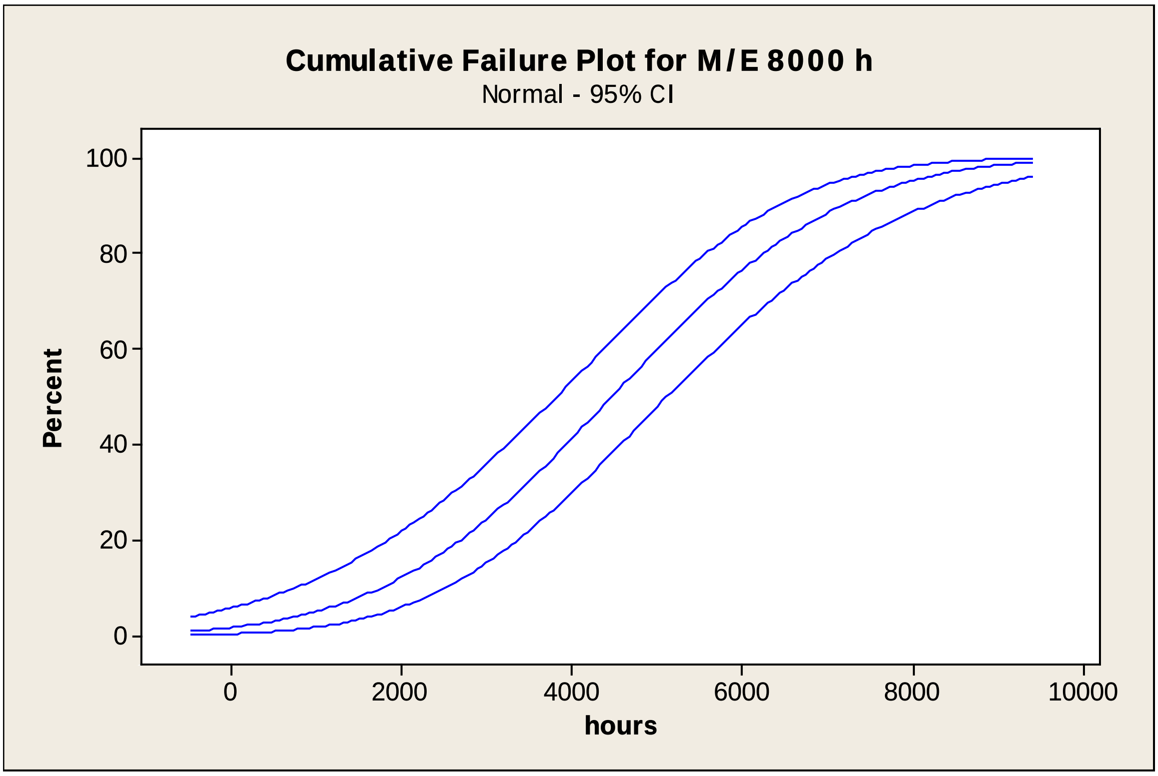


Figure 41: Cumulative failure plot for M/E 8000h

5.2.9 Non parametric analysis

In non parametric analysis, there is no assumption or fit of a known distribution, but only with statistical tools, the necessary statistics/ estimators are calculated.

Table 25: *Non parametric estimates*

```
             Standard   95,0% Normal CI
Mean(MTTF)      Error     Lower     Upper
   4474,56    326,869   3833,91   5115,21
```

A distribution free estimate of useful plots can be produced.

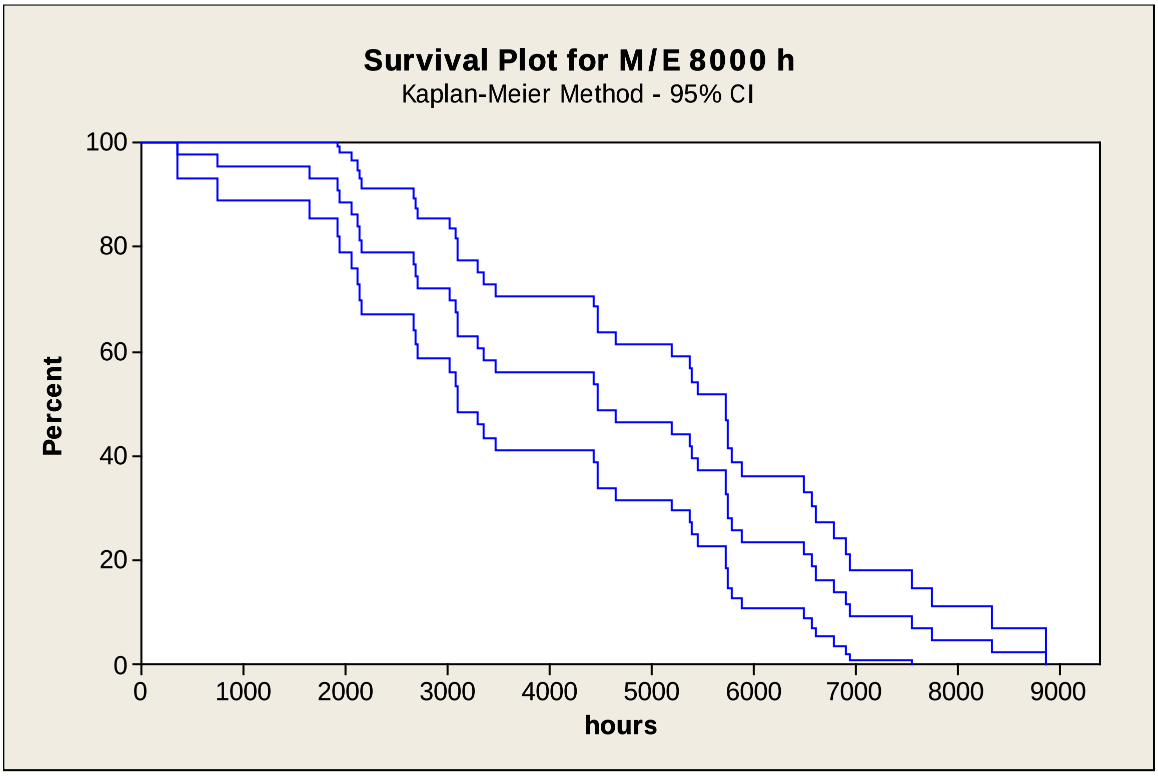


Figure 42: Survival plot for M/E 8000h

Table 26*: kaplan meier estimates for survival probabilities*

		95% confidence intervals	
time	survival probability	lower	upper
357	0,976744186	0,022983807	0,931696753
746	0,953488372	0,032114728	0,890544663
1650	0,930232558	0,038849724	0,854088498
1924	0,906976744	0,044295501	0,820159158
1935	0,88372093	0,048884831	0,787908422
2063	0,860465116	0,052841338	0,756897997
2112	0,837209302	0,056298599	0,726866076
2132	0,813953488	0,059343934	0,697641515
2161	0,790697674	0,062038034	0,669105361
2676	0,76744186	0,064424979	0,641171222
2685	0,744186047	0,066537832	0,613774292
2704	0,720930233	0,068401998	0,58686478
3022	0,697674419	0,070037336	0,560403762
3073	0,674418605	0,071459559	0,534360442
3095	0,651162791	0,072681179	0,508710298
3103	0,627906977	0,073712169	0,48343378

3300	0,604651163	0,074560438	0,458515389
3348	0,581395349	0,075232167	0,43394301
3473	0,558139535	0,075732054	0,409707436
4436	0,534883721	0,076063487	0,385802025
4465	0,511627907	0,076228663	0,362222472
4480	0,488372093	0,076228663	0,338966658
4650	0,465116279	0,076063487	0,316034584
5190	0,441860465	0,075732054	0,293428366
5382	0,418604651	0,075232167	0,271152313
5389	0,395348837	0,074560438	0,249213064
5460	0,372093023	0,073712169	0,227619827
5723	0,348837209	0,072681179	0,206384717
5726	0,325581395	0,071459559	0,185523233
5740	0,302325581	0,070037336	0,165054925
5741	0,279069767	0,068401998	0,145004315
5790	0,255813953	0,066537832	0,125402198
5887	0,23255814	0,064424979	0,106287501
6491	0,209302326	0,062038034	0,087710012
6565	0,186046512	0,059343934	0,069734538
6608	0,162790698	0,056298599	0,052447471
6790	0,139534884	0,052841338	0,035967765
6899	0,11627907	0,048884831	0,020466562
6947	0,093023256	0,044295501	0,00620567
7549	0,069767442	0,038849724	0
7745	0,046511628	0,032114728	0
8330	0,023255814	0,022983807	0
8864	0	0	0

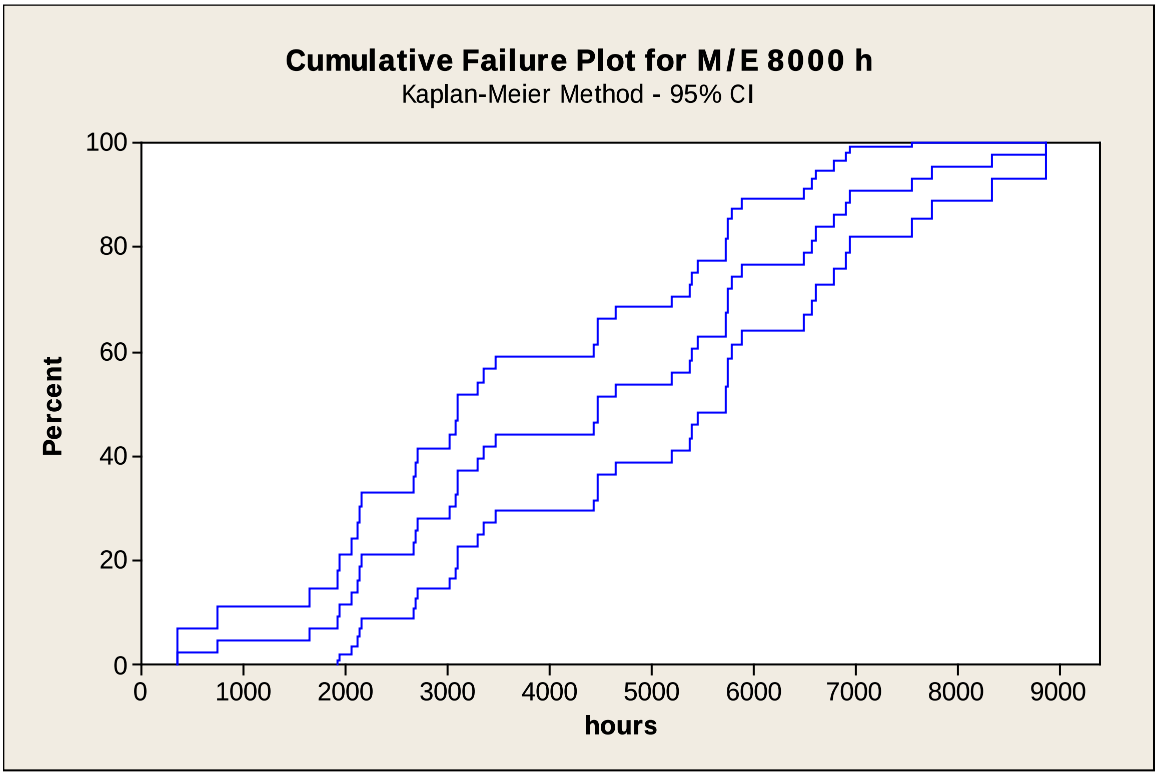


Figure 43: *Cumulative failure plot for M/E 8000h*

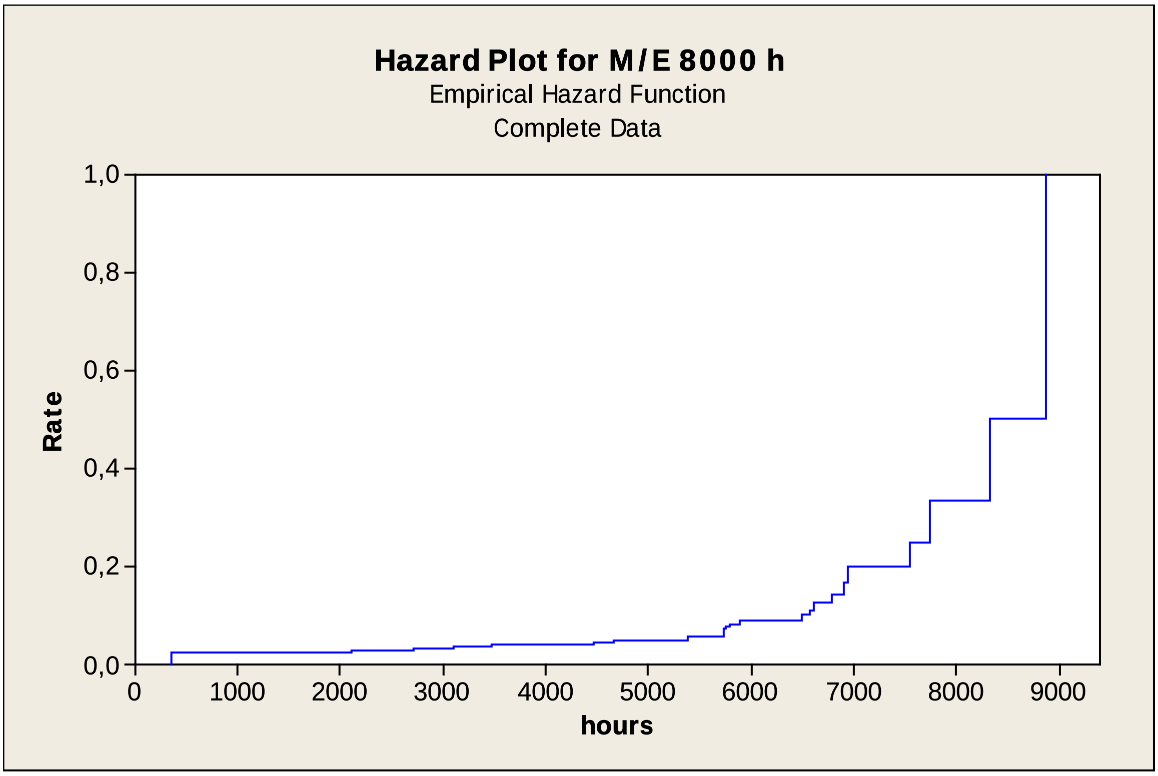


Figure 44: *Hazard plot for M/E 8000h*

Table 27*: cumulative failure probabilities and hazard rates*

		95% confidence intervals		
time	cumulative failure probability	lower	upper	hazard rates
357	0,023255814	0	0,068303247	0,023255814
746	0,046511628	0	0,109455337	0,023809524
1650	0,069767442	0	0,145911502	0,024390244
1924	0,093023256	0,00620567	0,179840842	0,025
1935	0,11627907	0,020466562	0,212091578	0,025641026
2063	0,139534884	0,035967765	0,243102003	0,026315789
2161	0,209302326	0,087710012	0,330894639	0,028571429
2676	0,23255814	0,106287501	0,358828778	0,029411765
2685	0,255813953	0,125402198	0,386225708	0,03030303
2704	0,279069767	0,145004315	0,41313522	0,03125
3022	0,302325581	0,165054925	0,439596238	0,032258065
3073	0,325581395	0,185523233	0,465639558	0,033333333
3095	0,348837209	0,206384717	0,491289702	0,034482759
3103	0,372093023	0,227619827	0,51656622	0,035714286
3300	0,395348837	0,249213064	0,541484611	0,037037037
3348	0,418604651	0,271152313	0,56605699	0,038461538
3473	0,441860465	0,293428366	0,590292564	0,04
4465	0,488372093	0,338966658	0,637777528	0,043478261
4480	0,511627907	0,362222472	0,661033342	0,045454545
4650	0,534883721	0,385802025	0,683965416	0,047619048
5389	0,604651163	0,458515389	0,750786936	0,055555556
5460	0,627906977	0,48343378	0,772380173	0,058823529
5723	0,651162791	0,508710298	0,793615283	0,0625
5726	0,674418605	0,534360442	0,814476767	0,066666667
5740	0,697674419	0,560403762	0,834945075	0,071428571
5741	0,720930233	0,58686478	0,854995685	0,076923077
5790	0,744186047	0,613774292	0,874597802	0,083333333
5887	0,76744186	0,641171222	0,893712499	0,090909091
6491	0,790697674	0,669105361	0,912289988	0,1
6565	0,813953488	0,697641515	0,930265462	0,111111111
6608	0,837209302	0,726866076	0,947552529	0,125
6899	0,88372093	0,787908422	0,979533438	0,166666667
7549	0,930232558	0,854088498	1	0,25
7745	0,953488372	0,890544663	1	0,333333333
8330	0,976744186	0,931696753	1	0,5
8864	1	1	1	1

In accordance to Figure 27: MRL/MTTF for M/E 4000h, we can plot the same function for M/E 8000 (24).

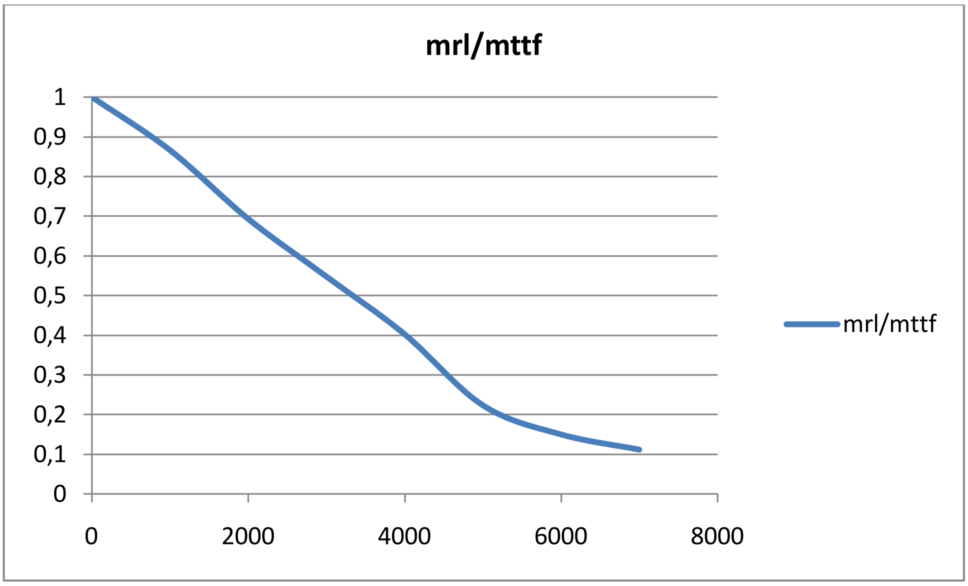


Figure 45: MRL/MTTF for M/E 8000

An increasing failure rate is also noticed in this category of items. For example at 4000 hours the remaining life is about 40% of that in time 0.

5.3 DIESEL GENERATORS

Following exactly the same procedure, the data that concerns failures in Diesel Generators, as can be found in Table 6**:** *Diesel Generators (inter inspection time of 4000hours)*, can be analyzed.

5.3.1 Outlier points

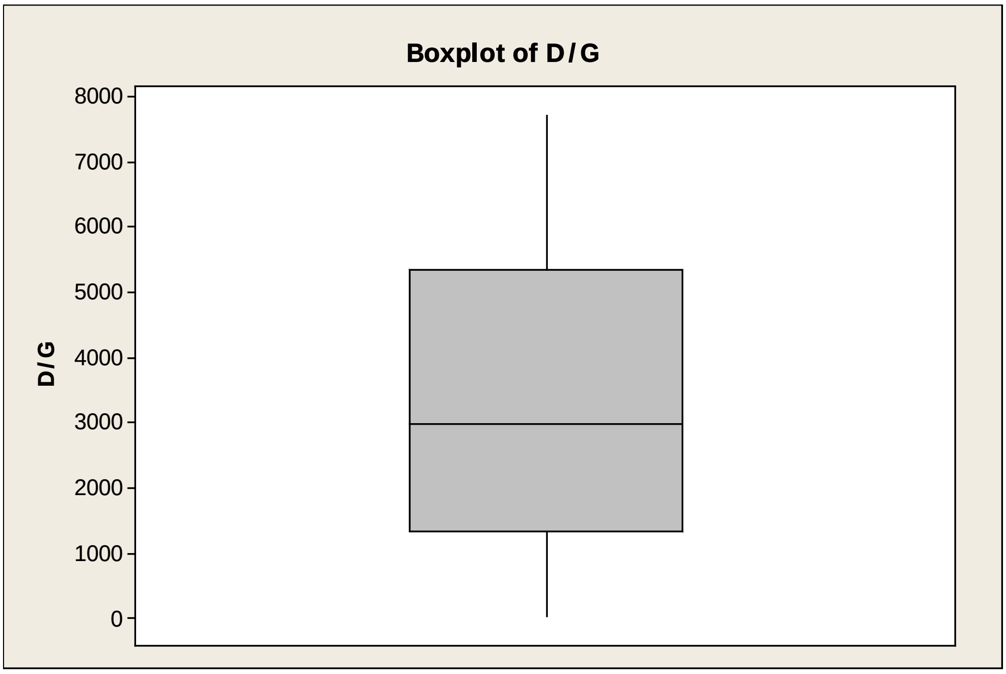


*Figure 46***:** *Boxplot of D/G*

From this plot we conclude that no outlier points exist in this group of data and the following parameters are calculated.

Table 28: *Basic parameters*

Median	Q1	Q3	IQR	lower outer fence	upper outer fence
2987	1348.5	5335.5	3987	55	7690

5.3.2 Basic statistics calculations

Some basic statistics calculations such as mean, Standard Error of mean, standard deviation variance, coefficient of variation first quartile, etc can be found in table 5.22

Table 29: *Basic statistics calculation*

Mean	Standard error of Mean	Standard Deviation	Variance	Coefficient of Variation	Q1	Median	Q3
3232,62	402,4201	2167,09	4696316,1	67,038	1348,5	2987	5335,5

IQR	Sum	Minimum	Maximum	Range	Sum of Squares	Skewness	Kurtosis
3987	93746	55	7690	7635	4,35E+08	0,425678	-0,91999

A graphical summary of the above can be found in figure 47.

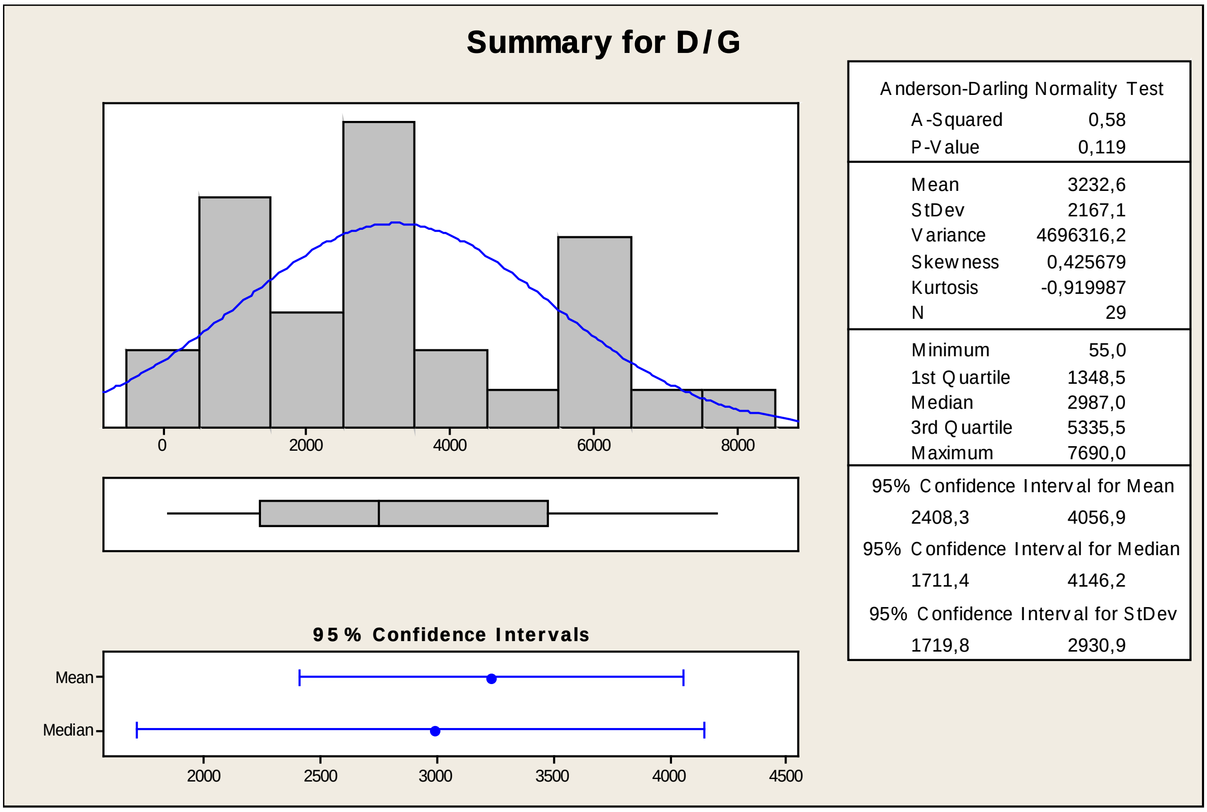


Figure 47: *Graphical summary for D/G*

5.3.3 Histogram

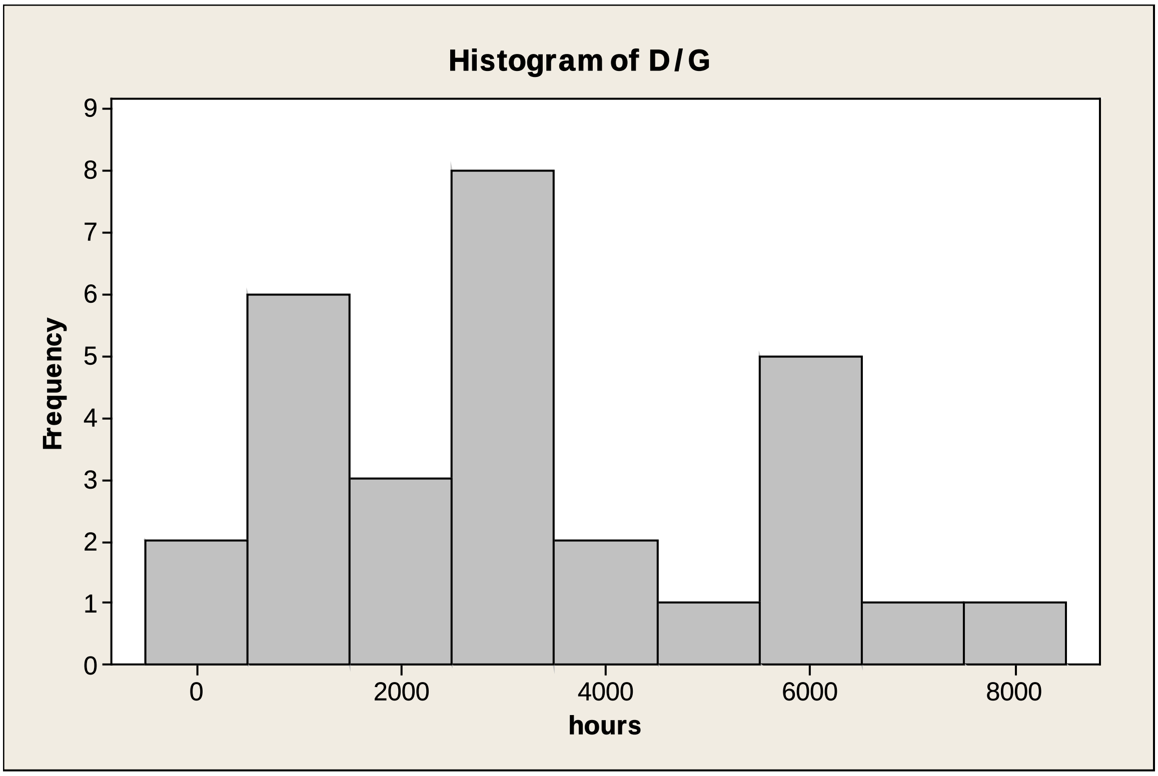


Figure 48: *Histogram of D/G*

5.3.4 Cumulative distribution function

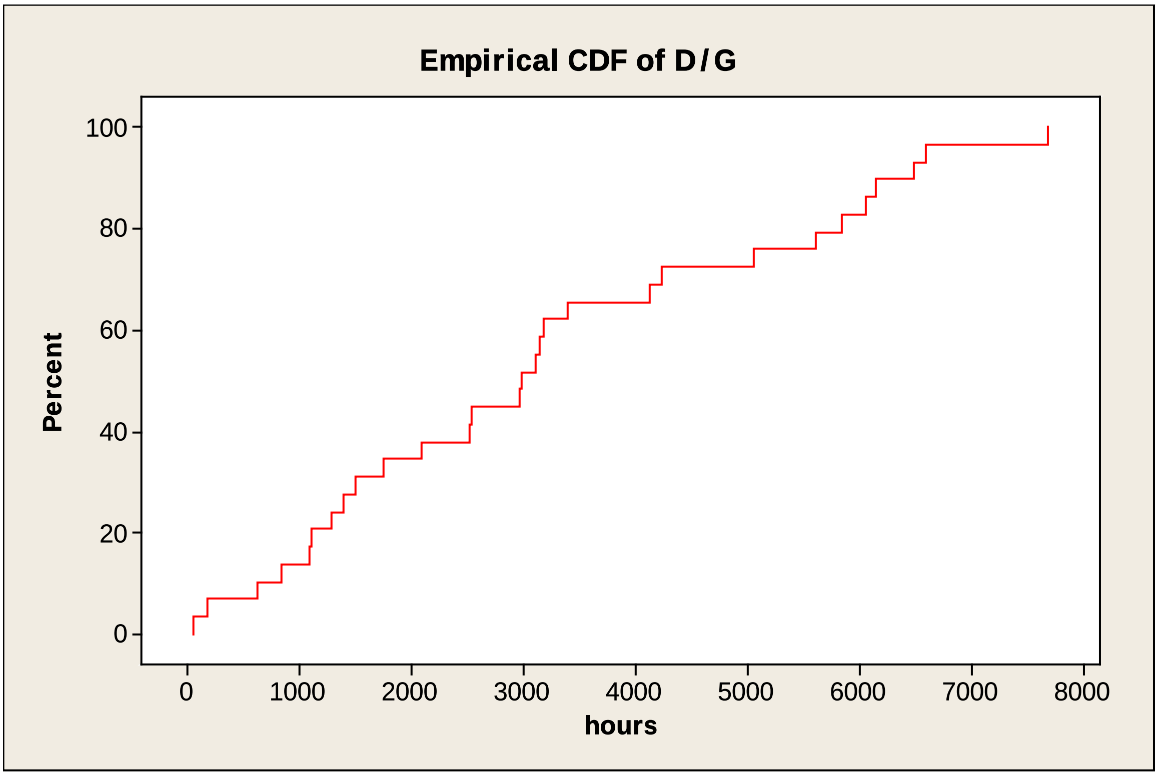


Figure 49: *Empirical CDF of D/G*

As described in previous chapter, the empirical cdf is an estimator of the cumulative distribution function. To further define the proper distribution which fits the data well, these have to be plotted in special plotting paper.

5.3.5 Individual distribution identification

Using the statistical software Minitab, the following graphs were obtained. It can be seen that the Weibull distribution describes adequately well the life data.

The Weibull distribution is widely used in reliability, since by altering the shape and scale parameters a wide range of fitted data can be described. For example, in case that the shape parameter is greater than 3, the distribution that arises is approximately normal.

In this case, where the shape parameter is greater than 1, this particular distribution appears to have an Increasing Failure Rate.

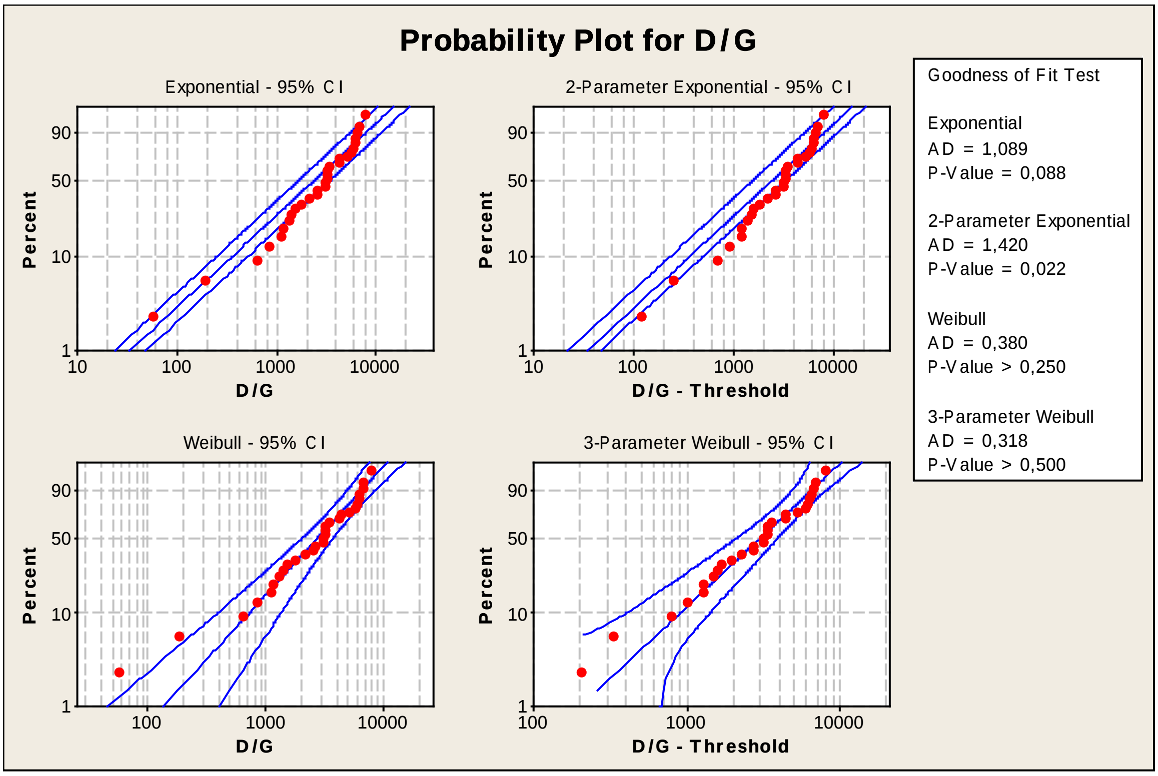


Figure 50: *Distribution identification for D/G*

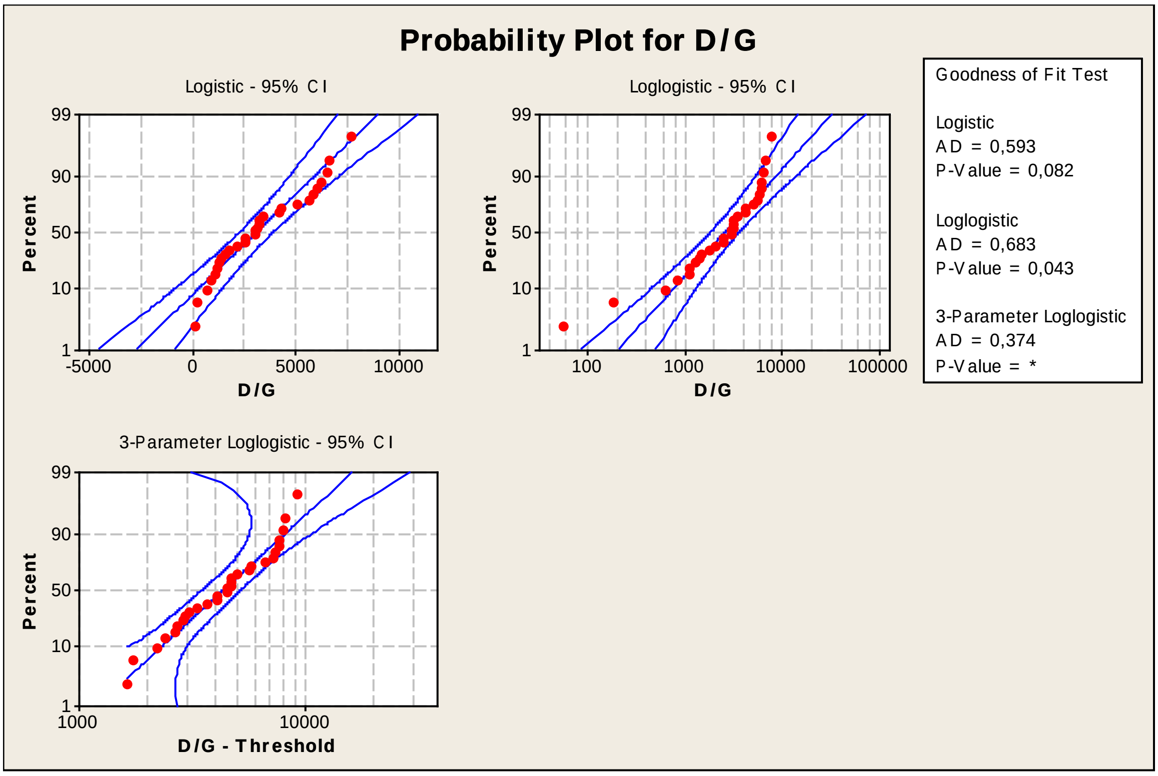


Figure 51: *Distribution identification for D/G*

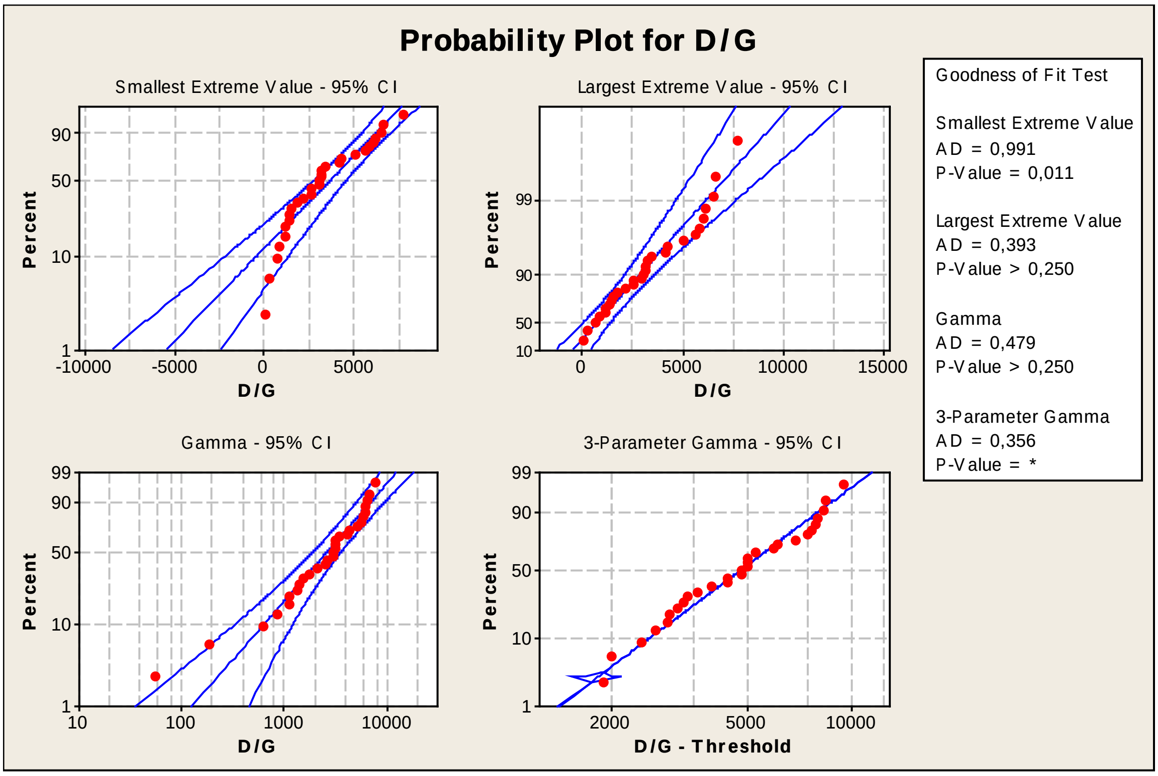


Figure 52: *Distribution identification for D/G*

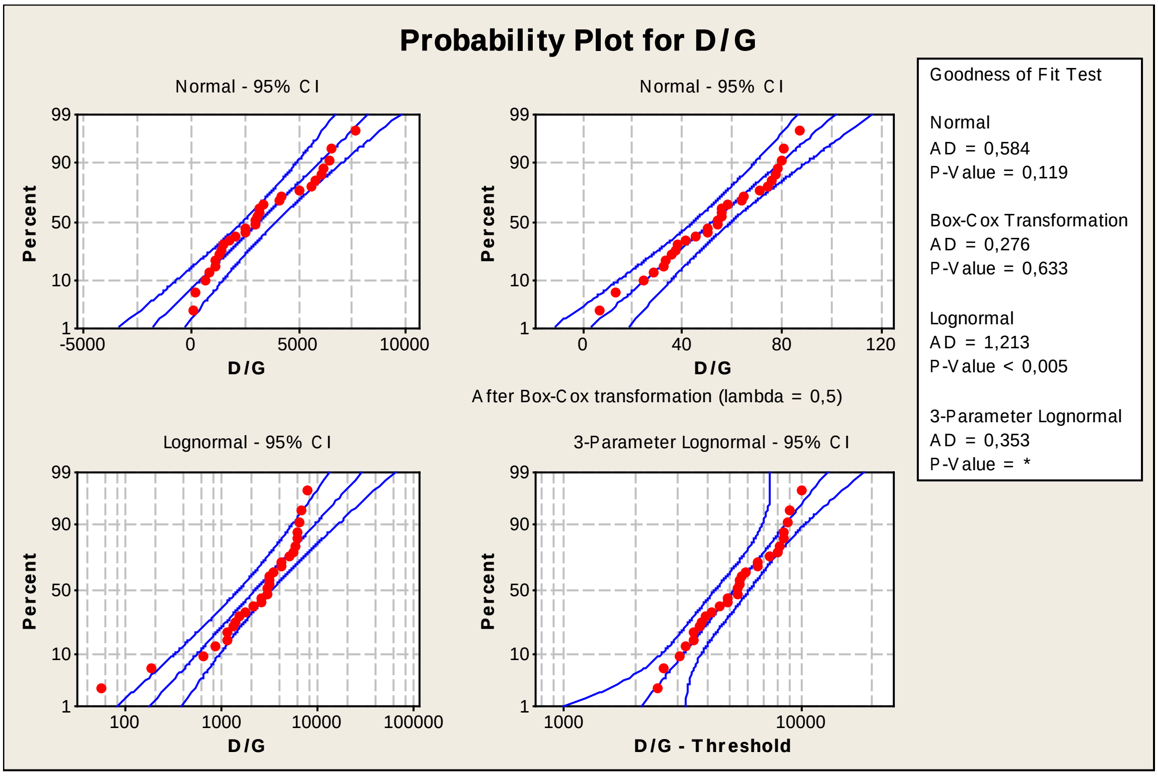


Figure 53: *Distribution identification for D/G*

It is necessary to compare various goodness of fit test to decide which distribution is appropriate for this set of data.

Table 30: *Comparison of goodness of fit tests for D/G*

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	0,584	0,119	
Box-Cox Transformation	0,276	0,633	
Lognormal	1,213	<0,005	
3-Parameter Lognormal	0,353	*	0,002
Exponential	1,089	0,088	
2-Parameter Exponential	1,420	0,022	1,000
Weibull	0,380	>0,250	
3-Parameter Weibull	0,318	>0,500	0,551
Smallest Extreme Value	0,991	0,011	
Largest Extreme Value	0,393	>0,250	
Gamma	0,479	>0,250	
3-Parameter Gamma	0,356	*	0,829
Logistic	0,593	0,082	
Loglogistic	0,683	0,043	
3-Parameter Loglogistic	0,374	*	0,052

Checking the Anderson Darling and p-value test, it is concluded that the Weibull distribution fits adequately well our data.

In order to define subject distribution it is necessary to estimate its parameters.

Following the same steps as in the other groups of data we need to calculate the shape and scale parameter for this distribution.

- Scale parameter α=3517.47
- Shape parameter β=1.40425

Summarizing, the fitted distribution for the first group of failures, is fully defined.

Weibull distribution of times to failure:

- CDF: $F(x) = 1 - \exp\,[-\left(\frac{x}{\alpha}\right)^{\beta}], 0 \leq x \leq \infty$
- PDF:$f(x) = \frac{\beta}{\alpha}\,(\frac{x}{\alpha})^{\beta-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right], 0 \leq x \leq \infty$

Where α, β have been already calculated.

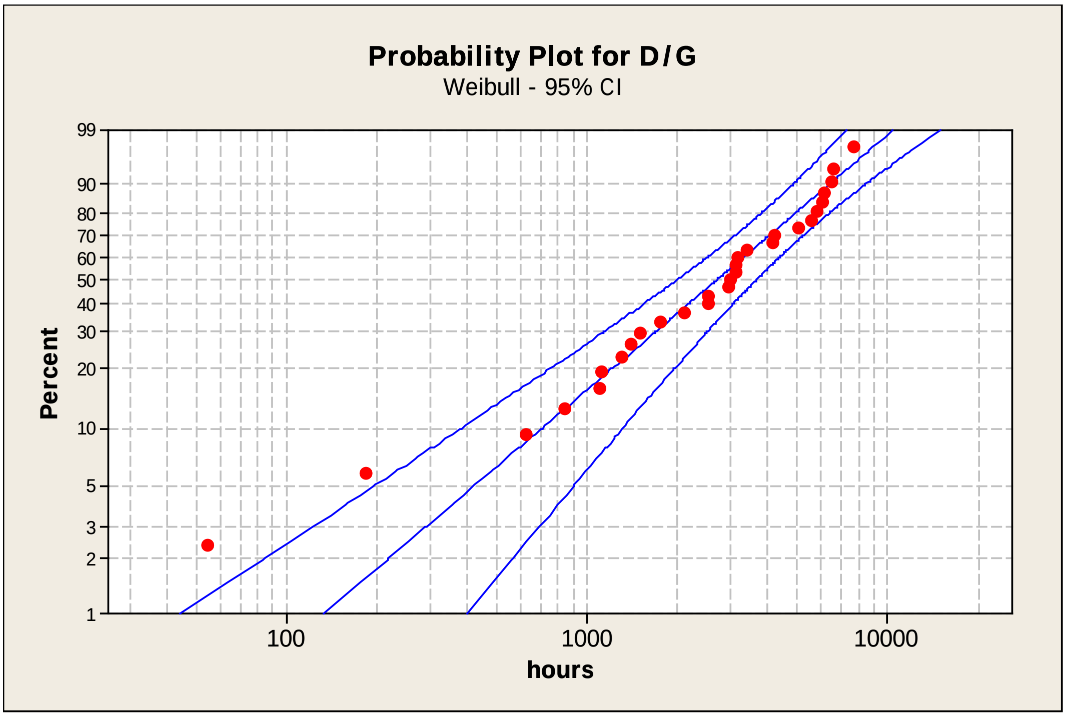


*Figure 54***:** *Plot of Weibull fit to the data*

In the plotted data in Weibull plotting paper, it is understood that this distribution fits well the collected life data.

5.3.6 Reliability concepts

Relevant plots for this group of data can be produced, similarly to previous chapter.

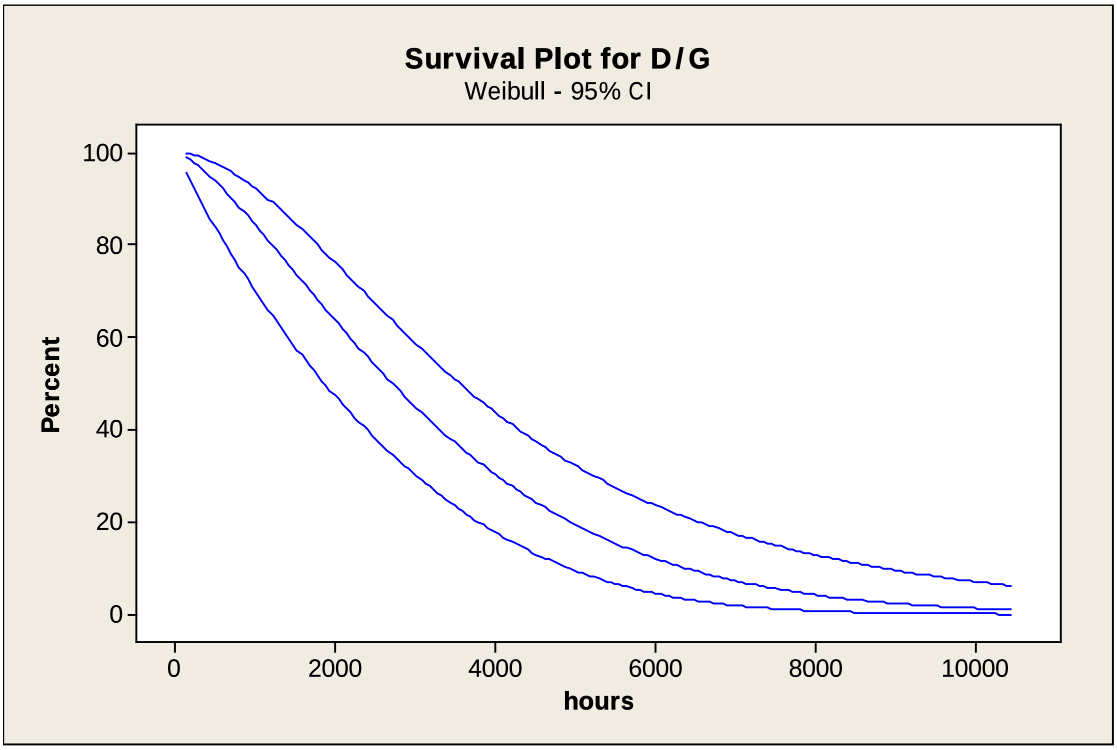


Figure 55: Plot of reliability function with 95% confidence intervals

Table 31*: table of survival probabilities*

		95% confidence intervals	
time	probability	lower	upper
1000	0,842834281	0,699137952	0,92156171
2000	0,635996853	0,472579419	0,76090491
3000	0,449442278	0,301450117	0,58662197
4000	0,301846413	0,177087075	0,43655613
5000	0,19424434	0,094157597	0,32096339
6000	0,120420337	0,045000398	0,23577378
7000	0,07219769	0,019328394	0,17366855
8000	0,041985698	0,007478481	0,1283315
9000	0,023738409	0,002613884	0,09509234
10000	0,013073913	0,000827452	0,07061856
12000	0,003688102	6,21458E-05	0,03913607
13000	0,00189387	1,47985E-05	0,0291867
14000	0,000952243	3,21541E-06	0,02178629
15000	0,000469236	6,38343E-07	0,01627437

The table above shows the percentage of the items under investigation that will survive at a given time. For example at 6000 hours the 12% of the items will have survive.

The conditional survival probabilities for this category, for a range of 1000 to 10000 hours is plotted in Figure 56: Conditional probabilities for D/G

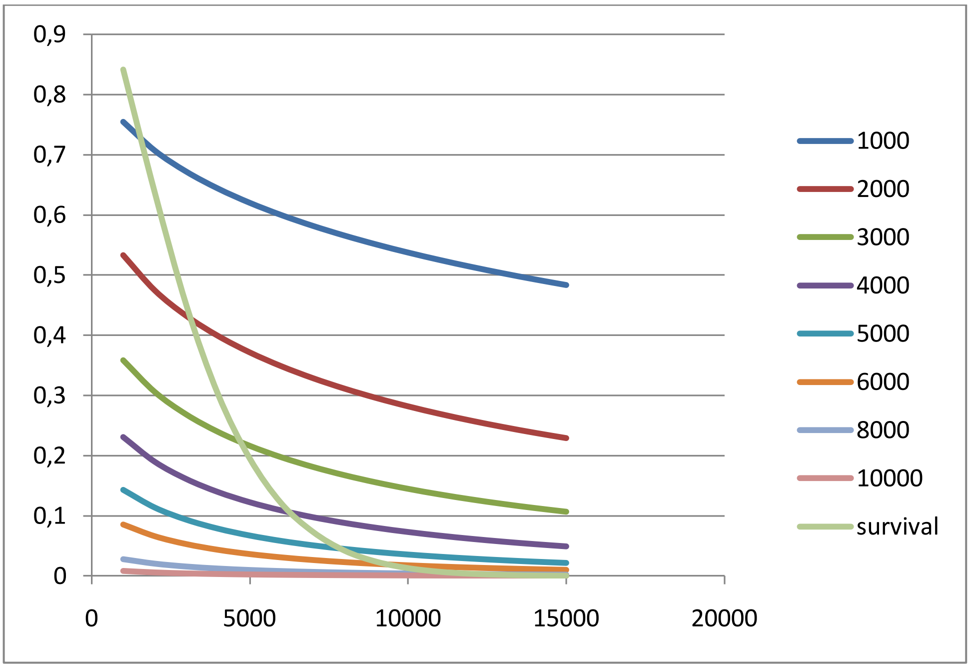


Figure 56: Conditional probabilities for D/G

To understand deeply this figure, for equipment that has survived 4000 hours, the conditional probabilities to survive t hours are:

Table 32: Conditional probabilities for D/G at 4000 hours

t	P
1000	0,643519
2000	0,398943
3000	0,239184
4000	0,139094
5000	0,078642
6000	0,043312
8000	0,012218
10000	0,003155

For example, given that an item of this group has survived 4000 hours, the probability to survive additionally 5000 hours is 0.078.

To visualize this example, a pie chart can be drawn.

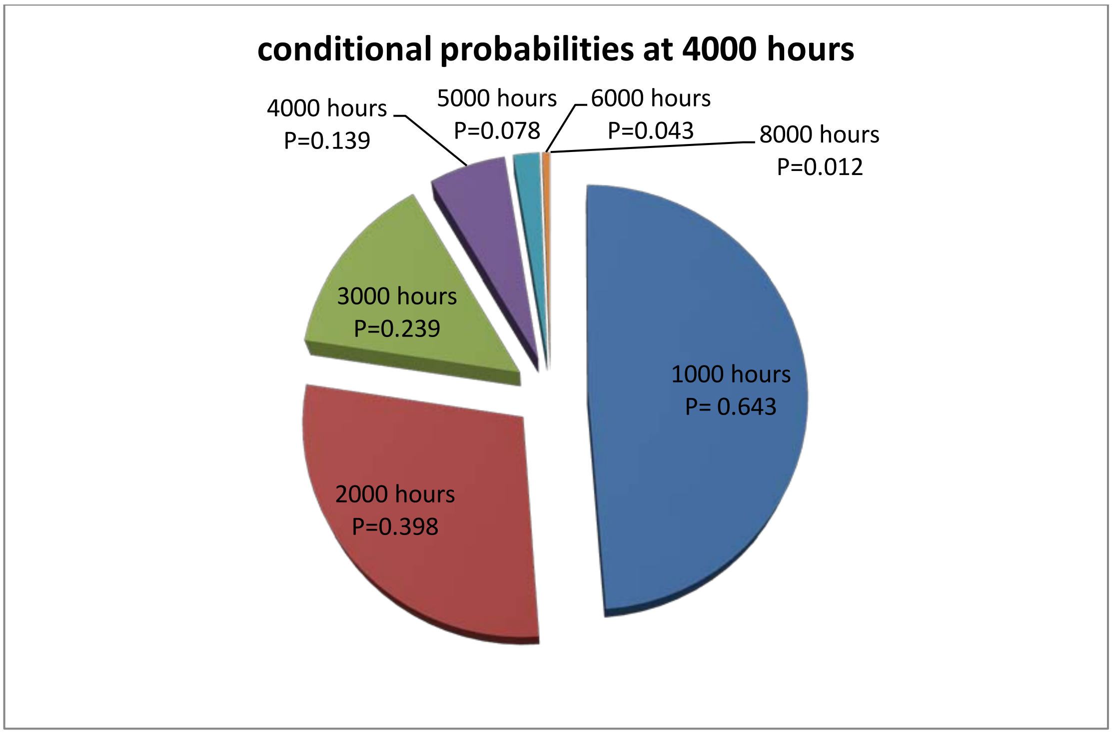


Figure 57: Conditional probabilities for D/G at 4000 hours

5.3.7 Failure rate plot

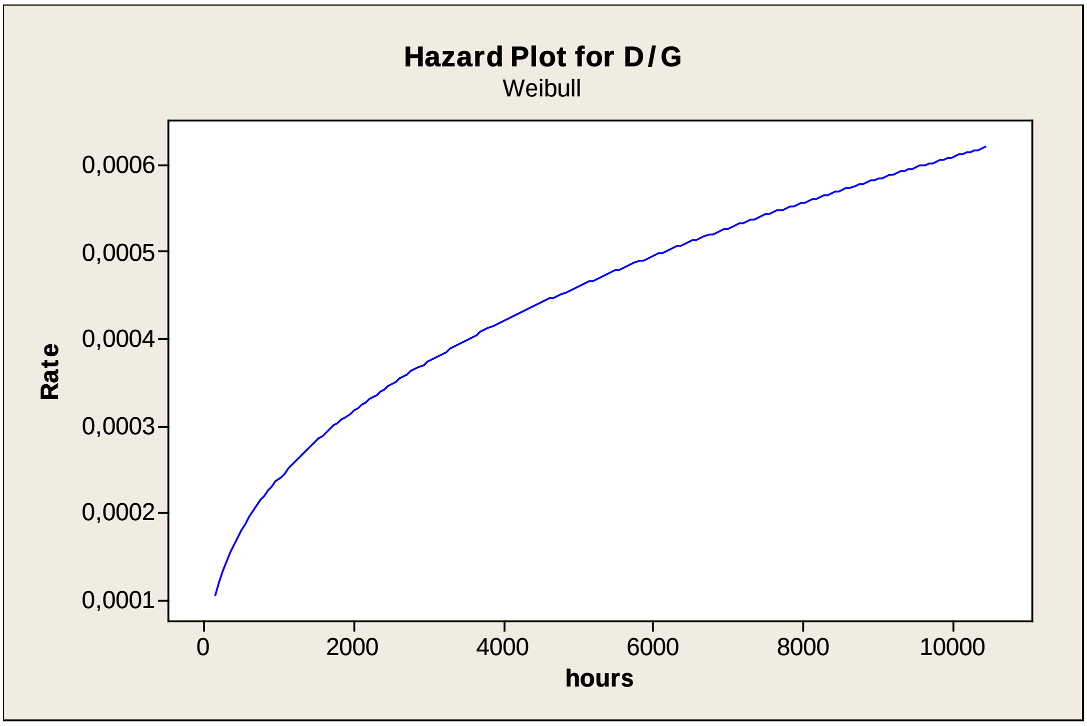


Figure 58: *Hazard plot for D/G*

5.3.8 Cumulative failure plot

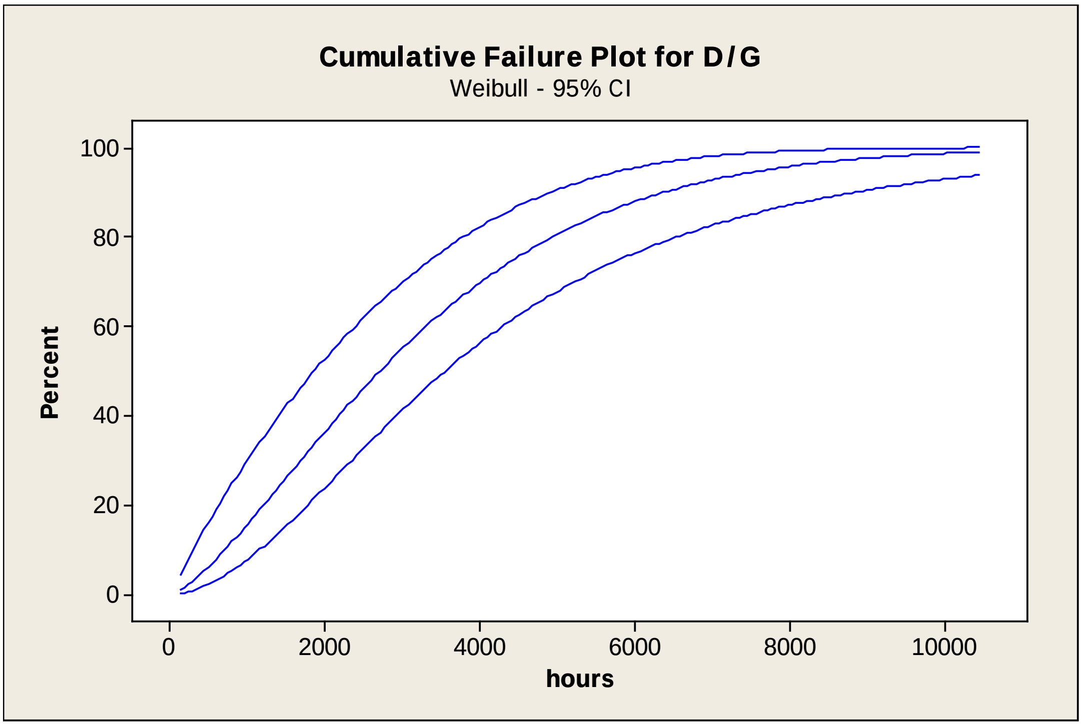


Figure 59: *Cumulative failure plot for D/G*

Table 33: *table of cumulative failure probabilities*

		95% confidence intervals	
time	probability	lower	upper
1000	0,157165719	0,078438295	0,300862048
2000	0,364003147	0,239095095	0,527420581
3000	0,550557722	0,413378027	0,698549883
4000	0,698153587	0,563443868	0,822912925
5000	0,80575566	0,679036608	0,905842403
6000	0,879579663	0,764226221	0,954999602
7000	0,92780231	0,826331446	0,980671606
8000	0,958014302	0,871668495	0,992521519
9000	0,976261591	0,90490766	0,997386116
10000	0,986926087	0,929381438	0,999172548
12000	0,996311898	0,960863933	0,999937854
13000	0,99810613	0,970813296	0,999985202
14000	0,999047757	0,978213712	0,999996785
15000	0,999530764	0,983725626	0,999999362

5.3.9 Non parametric analysis

Table 34: *Non parametric estimates*

```
             Standard   95,0% Normal CI
Mean(MTTF)      Error     Lower     Upper
   3232,62    402,420   2443,89   4021,35
```

A distribution free estimate of useful plots can be produced.

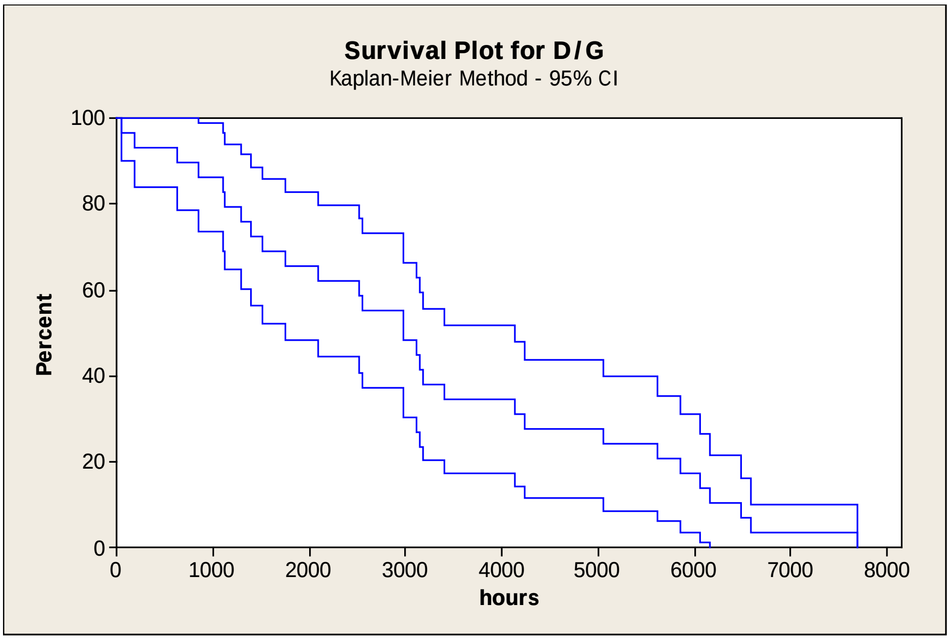


Figure 60: *Survival plot for D/G*

Table 35: kaplan meier estimates for survival probabilities

		95% confidence intervals	
time	survival probability	lower	upper
55	0,965517241	0,899107757	1
183	0,931034483	0,838809629	1
629	0,896551724	0,785711245	1
844	0,862068966	0,736566836	0,98757109
1100	0,827586207	0,690105512	0,9650669
1123	0,793103448	0,645671828	0,94053507
1300	0,75862069	0,602876643	0,91436474
1397	0,724137931	0,56146858	0,88680728
1512	0,689655172	0,521276397	0,85803395
1751	0,655172414	0,482179667	0,82816516
2100	0,620689655	0,444092506	0,7972868
2527	0,586206897	0,406954002	0,76545979
2987	0,482758621	0,300888757	0,66462848
3121	0,448275862	0,267274117	0,62927761
3150	0,413793103	0,234540209	0,593046
3180	0,379310345	0,202713195	0,55590749
3401	0,344827586	0,171834839	0,51782033
4128	0,310344828	0,141966052	0,4787236
4238	0,275862069	0,113192718	0,43853142
5062	0,24137931	0,085635264	0,39712336
5609	0,206896552	0,059464932	0,35432817
5849	0,172413793	0,034933099	0,30989449
6157	0,103448276	0	0,21428876
6485	0,068965517	0	0,16119037
6593	0,034482759	0	0,10089224
7690	0	0	0

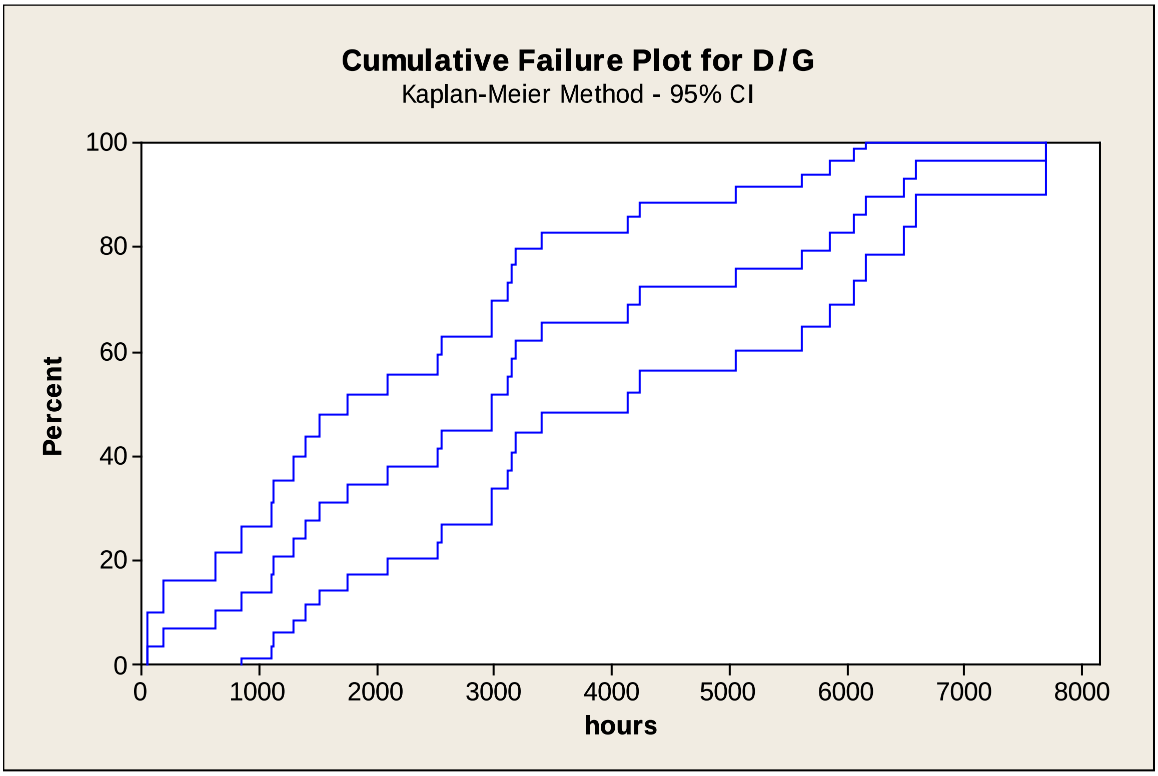


Figure 61: *Cumulative failure plot for D/G*

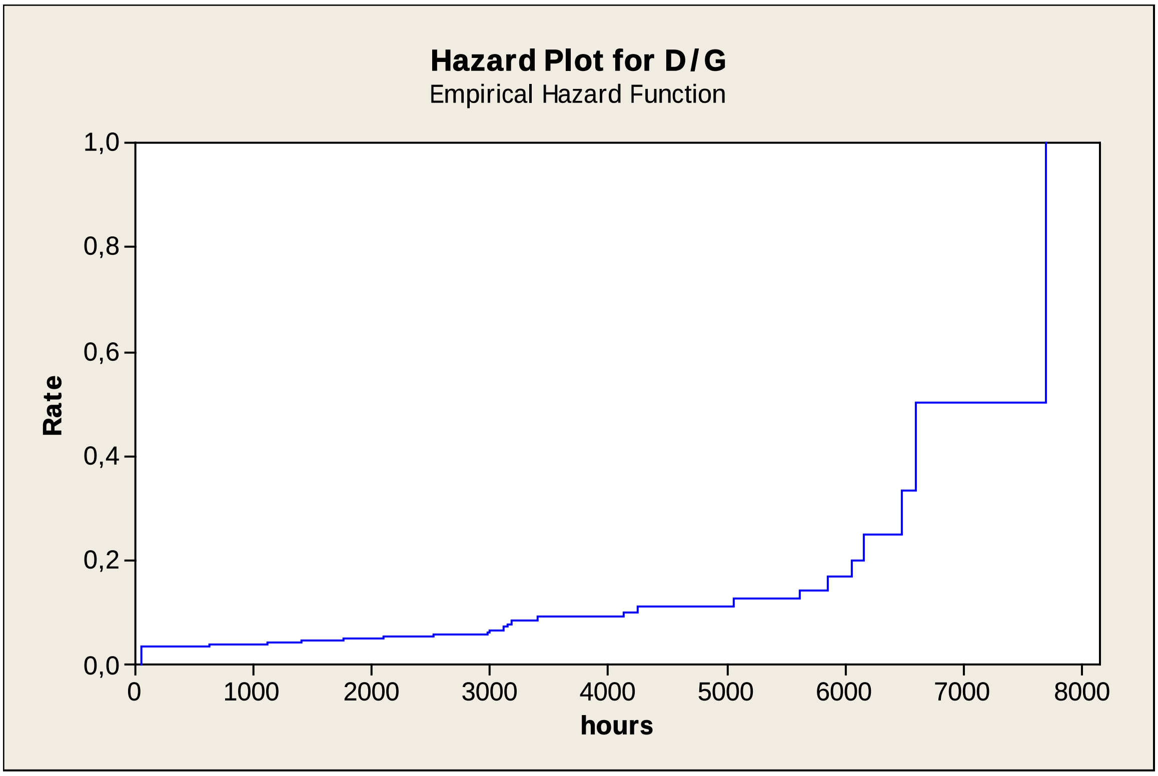


Figure 62: *Hazard plot for D/G*

Table 36*: cumulative failure probabilities and hazard rates*

		95% confidence intervals		
time	cumulative failure probability	lower	upper	hazard rates
55	0,034482759	0	0,100892243	0,034482759
183	0,068965517	0	0,161190371	0,035714286
629	0,103448276	0	0,214288755	0,037037037
844	0,137931034	0,012428905	0,263433164	0,038461538
1100	0,172413793	0,034933099	0,309894488	0,04
1123	0,206896552	0,059464932	0,354328172	0,041666667
1300	0,24137931	0,085635264	0,397123357	0,043478261
1397	0,275862069	0,113192718	0,43853142	0,045454545
1512	0,310344828	0,141966052	0,478723603	0,047619048
1751	0,344827586	0,171834839	0,517820333	0,05
2100	0,379310345	0,202713195	0,555907494	0,052631579
2527	0,413793103	0,234540209	0,593045998	0,055555556
2987	0,517241379	0,335371516	0,699111243	0,066666667
3121	0,551724138	0,370722393	0,732725883	0,071428571
3150	0,586206897	0,406954002	0,765459791	0,076923077
3180	0,620689655	0,444092506	0,797286805	0,083333333
3401	0,655172414	0,482179667	0,828165161	0,090909091
4128	0,689655172	0,521276397	0,858033948	0,1
4238	0,724137931	0,56146858	0,886807282	0,111111111
5062	0,75862069	0,602876643	0,914364736	0,125
5609	0,793103448	0,645671828	0,940535068	0,142857143
5849	0,827586207	0,690105512	0,965066901	0,166666667
6157	0,896551724	0,785711245	1	0,25
6485	0,931034483	0,838809629	1	0,333333333
6593	0,965517241	0,899107757	1	0,5
7690	1	1	1	1

In accordance to Figure 27: MRL/MTTF for M/E 4000h, we can plot the same function for D/G (24).

This function indicates an increasing failure rate, since the mean remaining life is constantly decreasing as can be clearly seen in Figure 63.

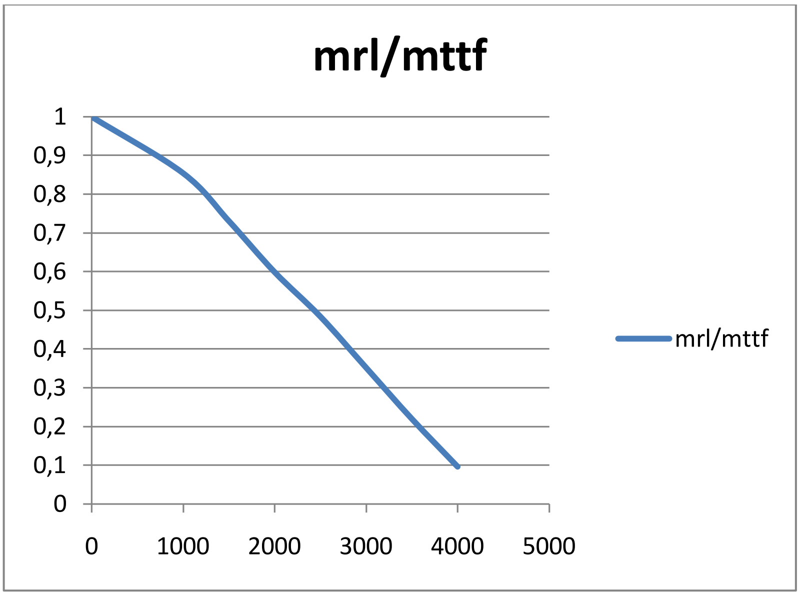


Figure 63: MRL/MTTF for D/G

5.4 AUX 4000

We can analyze the failures which regard auxiliary equipment. We begin with that equipment that the manufacturer suggests 4000 working hours between inspections to take place.

5.4.1 Outlier points

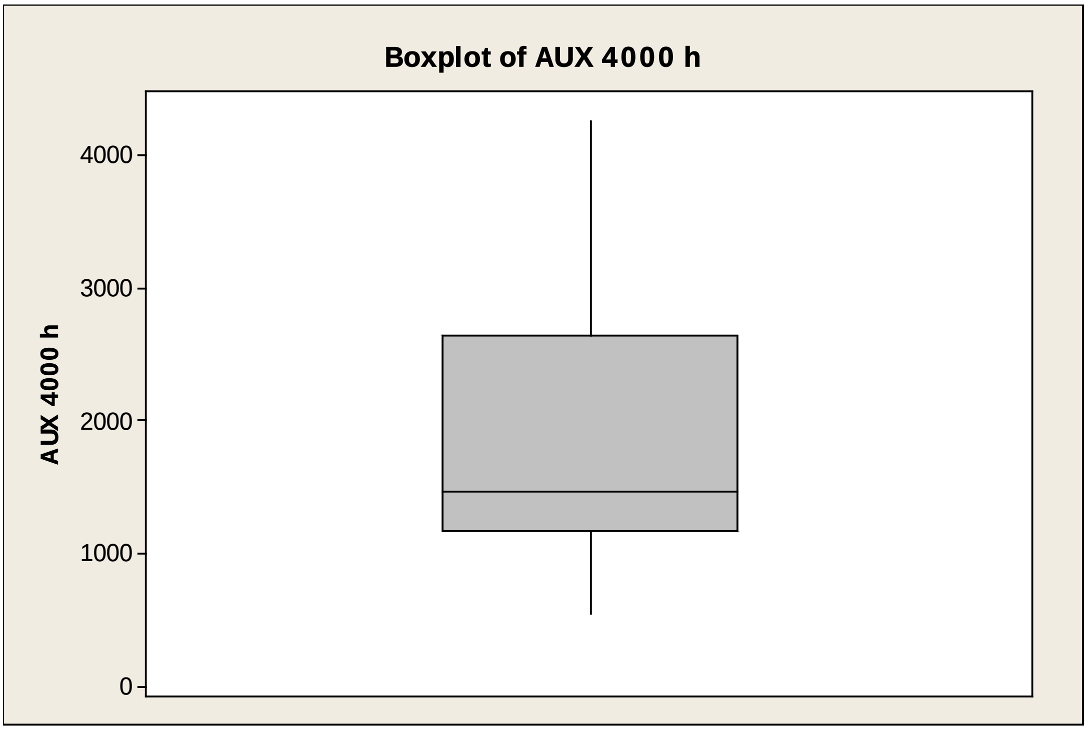


Figure 64:Boxplot of AUX 4000

From this plot we conclude that no outlier points exist in this group of data and the following parameters are calculated.

Table 37:Basic parameters

Median	Q1	Q3	IQR	lower outer fence	upper outer fence
1463	1165.5	2645	1479.5	552	4256

5.4.2 Basic statistics calculations

Table 38: Basic statistic calculations

Mean	Standard error of Mean	Standard Deviation	Variance	Coefficient of Variation	Q1	Median	Q3
1936	304,9810	1099,62	1209174	56,79880	1165,5	1463	2645

IQR	Sum	Minimum	Maximum	Range	Sum of Squares	Skewness	Kurtosis
1479,5	25168	552	4256	3704	63235344	0,7852907	0,038491

A graphical summary of the above:

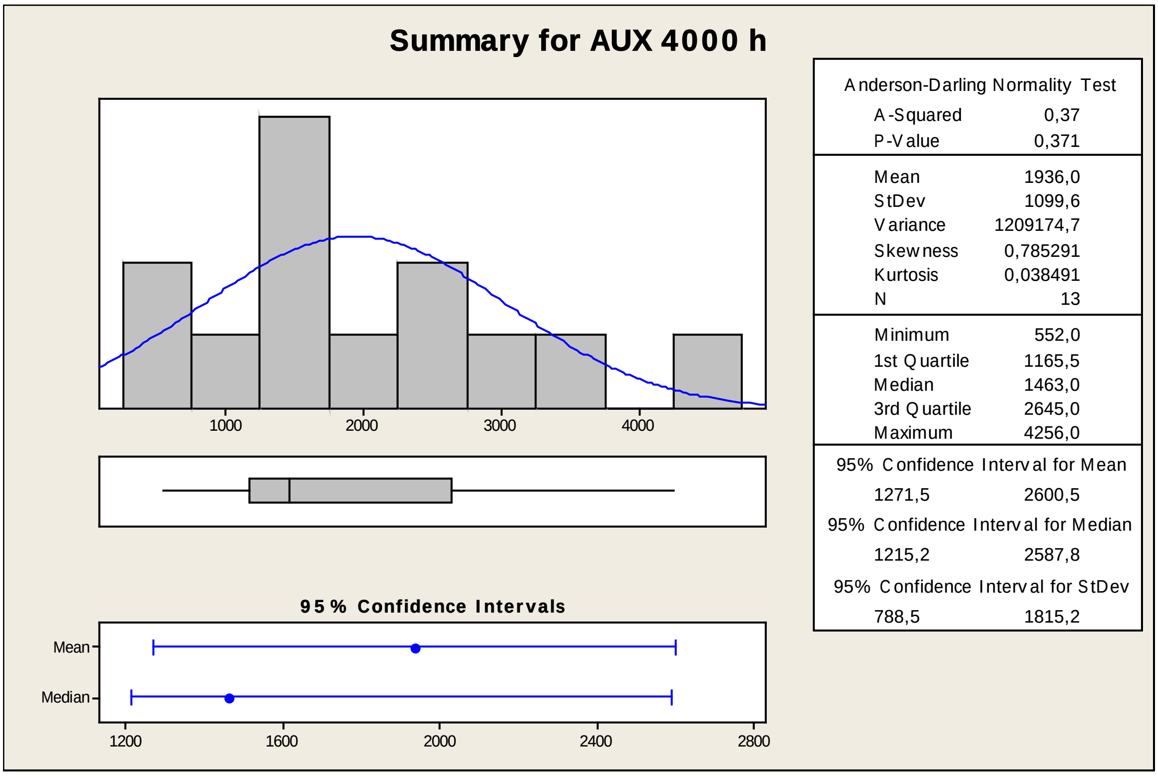


Figure 65: Graphical summary for AUX 4000

5.4.3 Histogram

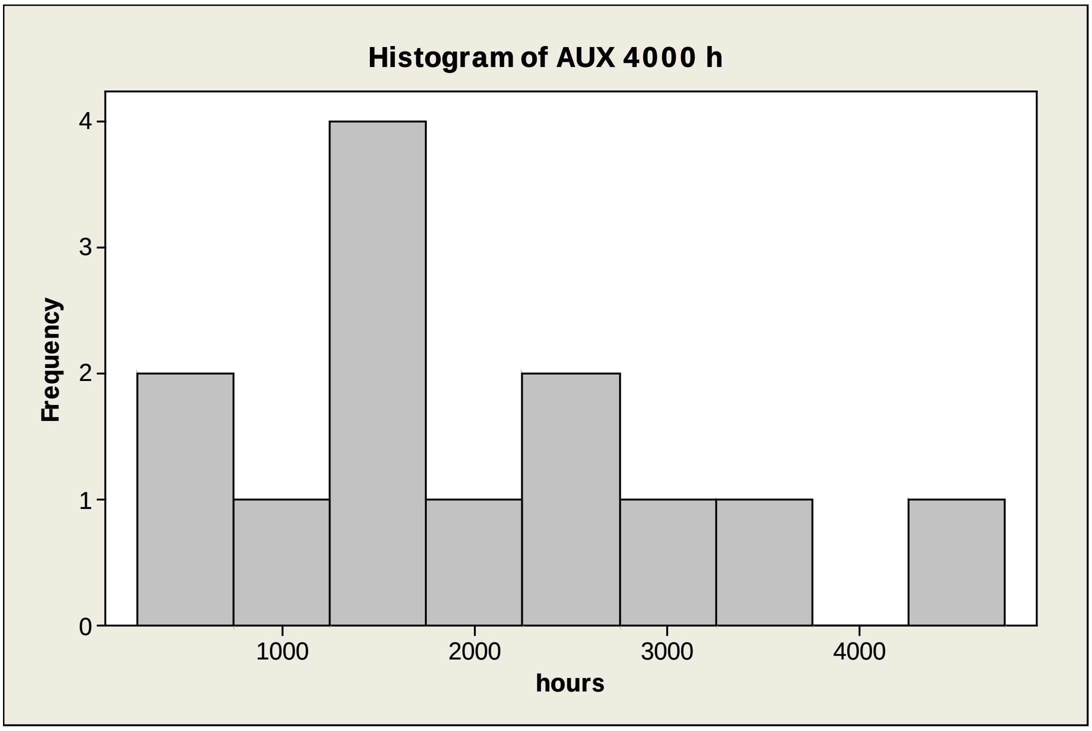


Figure 66: Histogram of AUX 4000

5.4.4 Cumulative distribution function

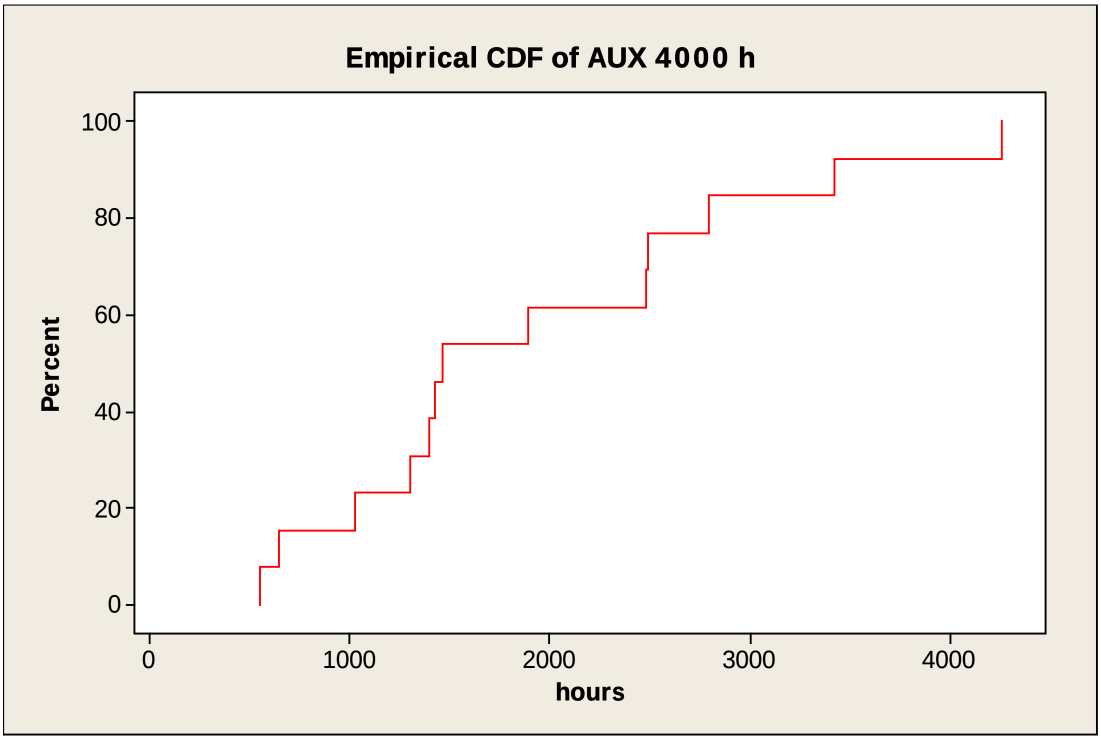


Figure 67: Empirical CDF of AUX 4000

As described in previous chapter, the empirical cdf is an estimator of the cumulative distribution function. To further define the proper distribution which fits the data well, these have to be plotted in special plotting paper. The next chapter describes the procedure for this category of data.

5.4.5 Individual distribution identification

Using the statistical software Minitab, the following graphs were obtained.

As it is concluded after the goodness of fit tests, the Loglogistic distribution describes well our data and therefore is used for the farther calculation of useful results.

It is noted that at Appendix A, the statistical properties for each used distribution are presented, as an assist for further examination.

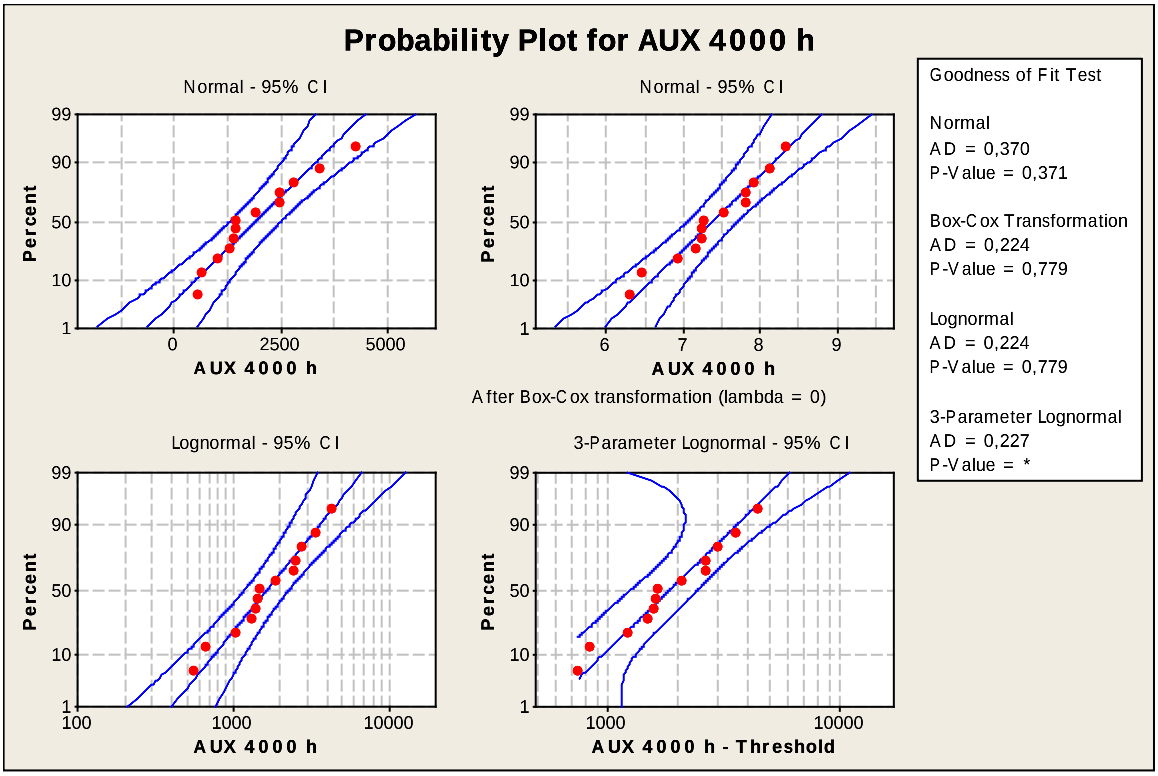


Figure 68: Distribution identification for AUX 4000

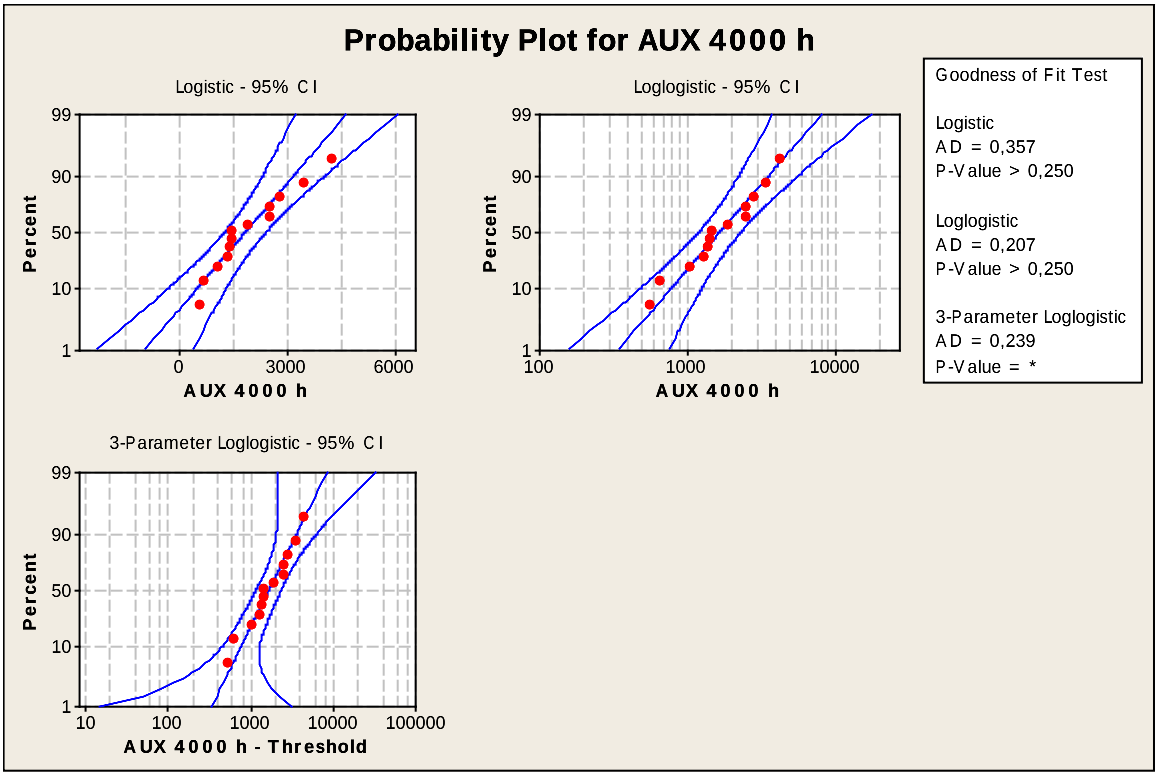


Figure 69: Distribution identification for AUX 4000

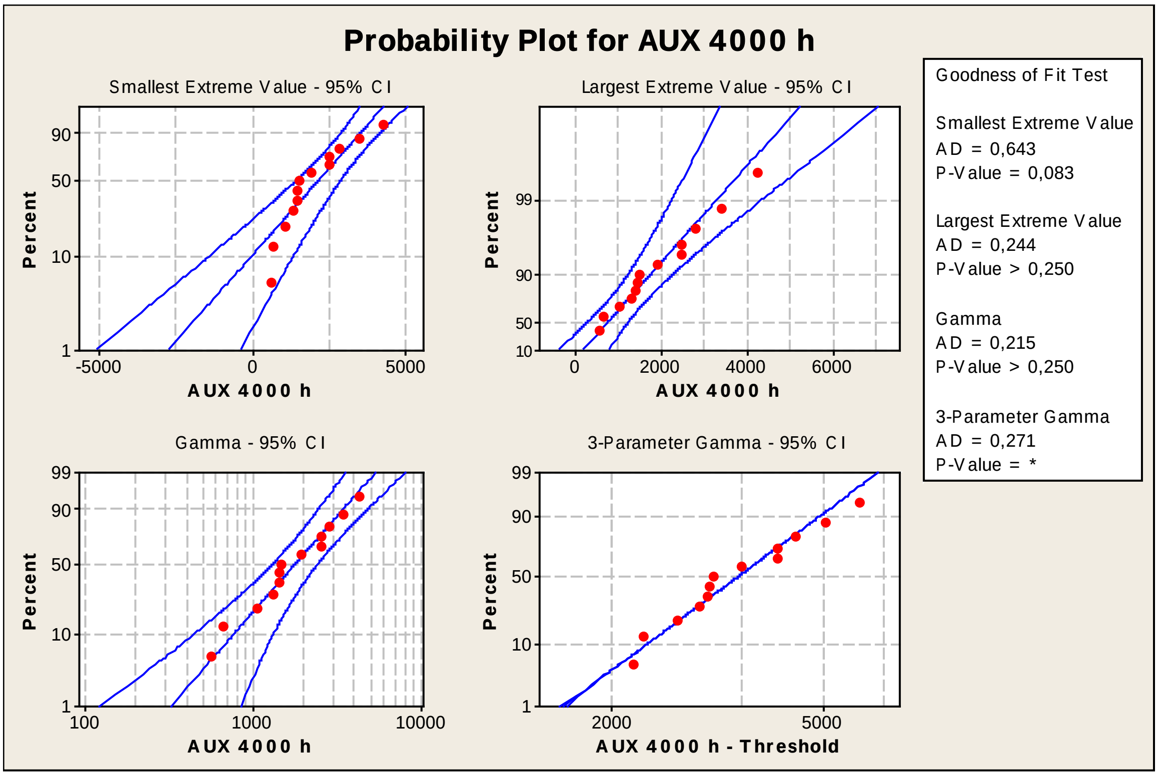


Figure 70: Distribution identification for AUX 4000

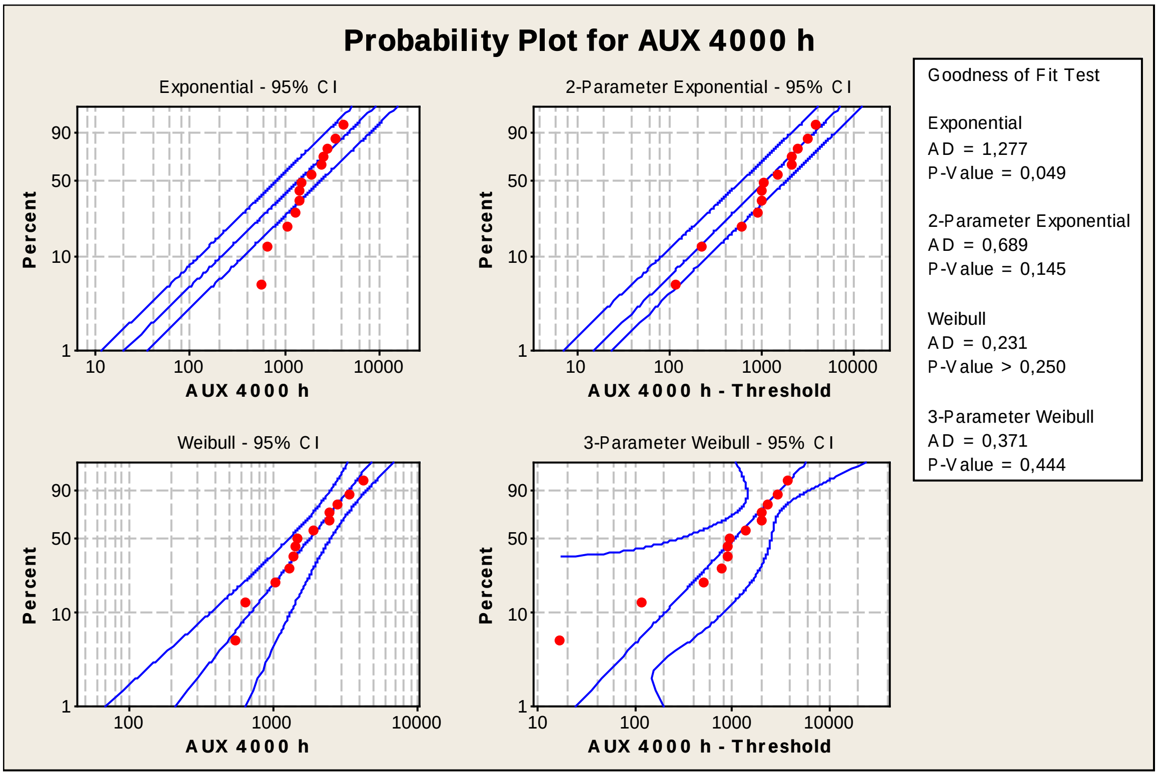


Figure 71: Distribution identification for AUX 4000

It is necessary to compare various goodness of fit test to decide which distribution is appropriate for this set of data.

Table 39: Comparison of goodness of fit tests for AUX 4000

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	0,370	0,371	
Box-Cox Transformation	0,224	0,779	
Lognormal	0,224	0,779	
3-Parameter Lognormal	0,227	*	0,831
Exponential	1,277	0,049	
2-Parameter Exponential	0,689	0,145	0,010
Weibull	0,231	>0,250	
3-Parameter Weibull	0,371	0,444	0,214
Smallest Extreme Value	0,643	0,083	
Largest Extreme Value	0,244	>0,250	
Gamma	0,215	>0,250	
3-Parameter Gamma	0,271	*	1,000
Logistic	0,357	>0,250	
Loglogistic	0,207	>0,250	
3-Parameter Loglogistic	0,239	*	0,978

Checking the Anderson Darling and p-value test, it is concluded that the Loglogistic distribution fits adequately well our data.

In order to define subject distribution it is necessary to estimate its parameters.

For the Loglogistic distribution to be fully defined, we must calculate the scale and the shape parameter.

The scale parameter for this distribution is also the median, for the shape parameter, we use the method of Maximum Likelihood. Further details can be found in (14).

Summarizing, the fitted distribution is:

Log- logistic distribution of times to failure

- CDF: $F(x) = \frac{1}{1+(x/\alpha)^{-\beta}}, 0 \leq x < \infty$
- PDF: $f(x) = \frac{\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1}}{[1+(x/\alpha)^{\beta}]^2}, 0 \leq x < \infty$

Where α=scale parameter= 1682.8

β=shape parameter=0.3419

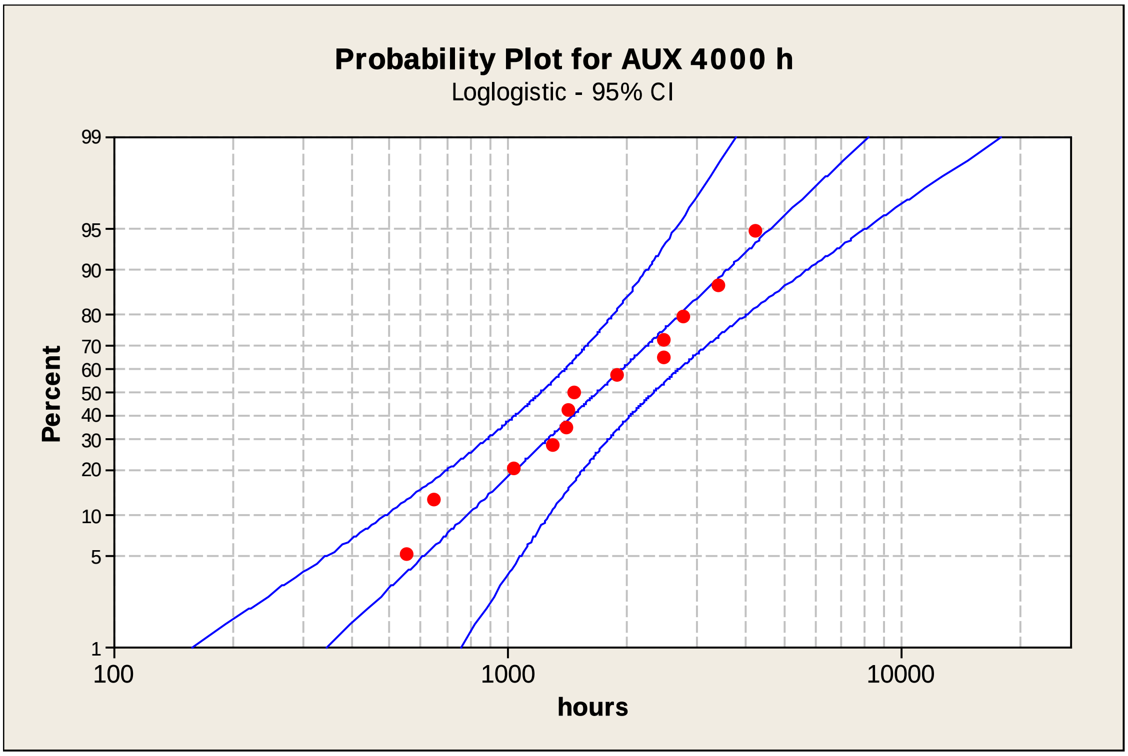


Figure 72: Plot of Loglogistic fit to the data

In the plotted data in Log- logistic plotting paper, it is understood that this distribution fits well the collected life data.

5.4.6 Reliability concepts

Relevant plots for this group of data can be produced, similarly to previous chapter.

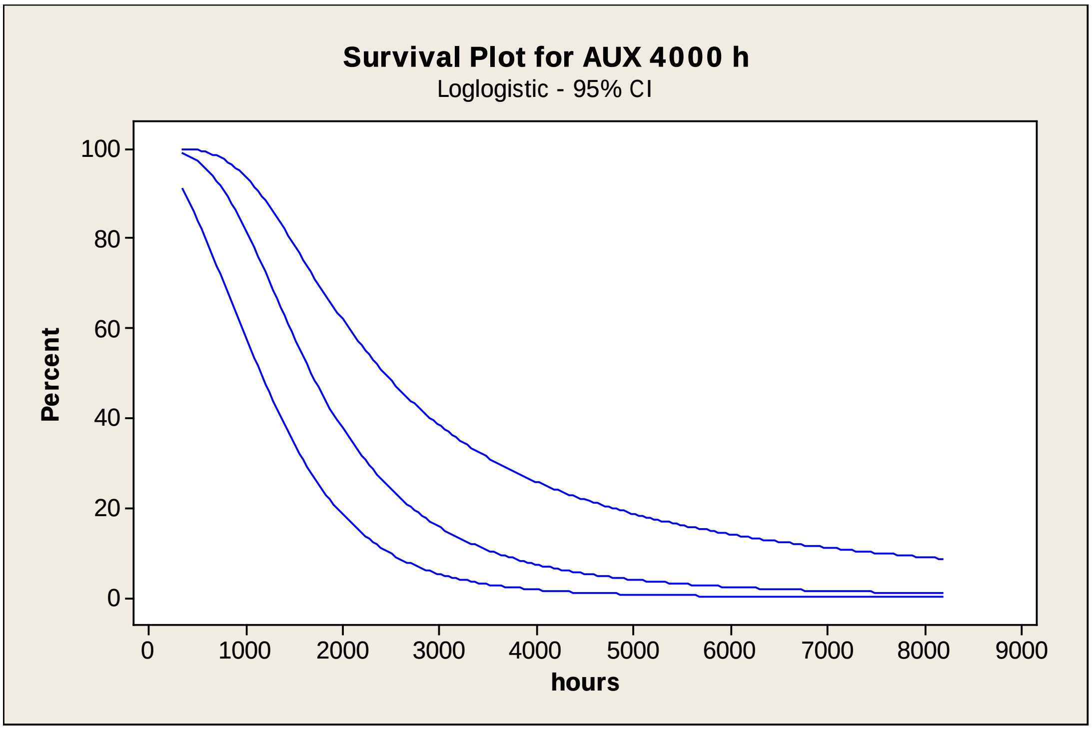


*Figure 73:*Plot of reliability function with 95% confidence intervals

Table 40 : table of survival probabilities

		95% confidence intervals	
time	probability	lower	upper
1000	0,819369214	0,581291261	0,936795115
2000	0,377137512	0,185411744	0,616961707
3000	0,157133767	0,0529529	0,383321058
4000	0,074777896	0,018390359	0,2585237
5000	0,040550605	0,007723819	0,186650179
6000	0,024280386	0,003737323	0,1416842
7000	0,015652331	0,002007439	0,111666277
8000	0,010673007	0,001166957	0,090592175
9000	0,007603596	0,000721449	0,075196582
10000	0,005609857	0,000468506	0,063582935
12000	0,003310626	0,00022135	0,047468293
13000	0,002625514	0,000159121	0,04172597
14000	0,002118017	0,000117171	0,037020823
15000	0,001733982	8,8094E-05	0,033112191

For the Loglostic distributed category AUX 4000 the conditional probabilities for a range of 1000 to 10000 hours is plotted in the following graph

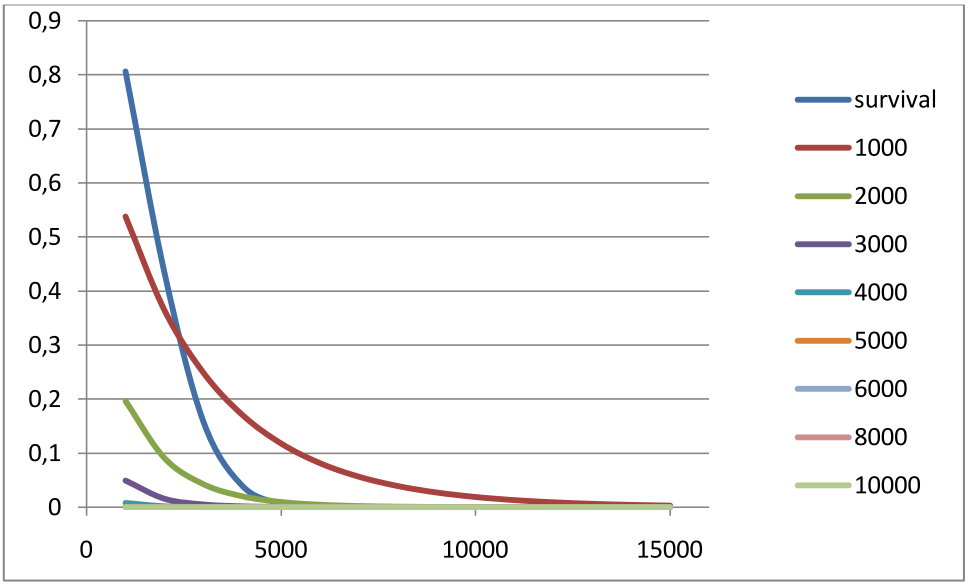


Figure 74: Conditional probabilities for AUX 4000

For example the conditional probabilities for an item to survive t hours, given that it has survived 2000 hours are:

Table 41: Conditional probabilities for AUX 4000 at 2000 hours

t	P
1000	0,3650
2000	0,0910
3000	0,0156
4000	0,0018
5000	0,0002
6000	0,0000
8000	0,0000
10000	0,0000

For this category the expected values of probabilities is particularly low. This can be explained partly due to the relatively small number of observations that were recorded for this items.

As can be seen in Table 7***:*** AUX 4000h, the 92% of the items that were recorded, have failed before the suggested by the manufacturer limit of 4000 hours for inspection.

This can be fixed in an oncoming collection of same life data, where in that case the available samples for analysis will be larger and therefore will produce more reliable results.

For the actual data, we can extract the following plot.

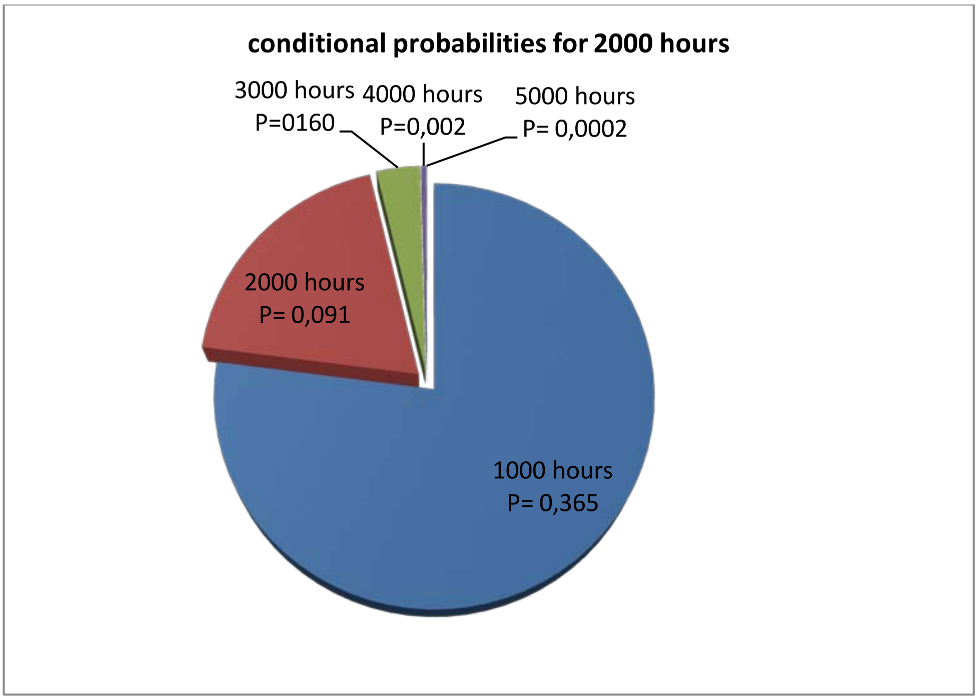


Figure 75: Conditional probabilities for AUX 4000 at 2000 hours

5.4.7 Failure rate plot

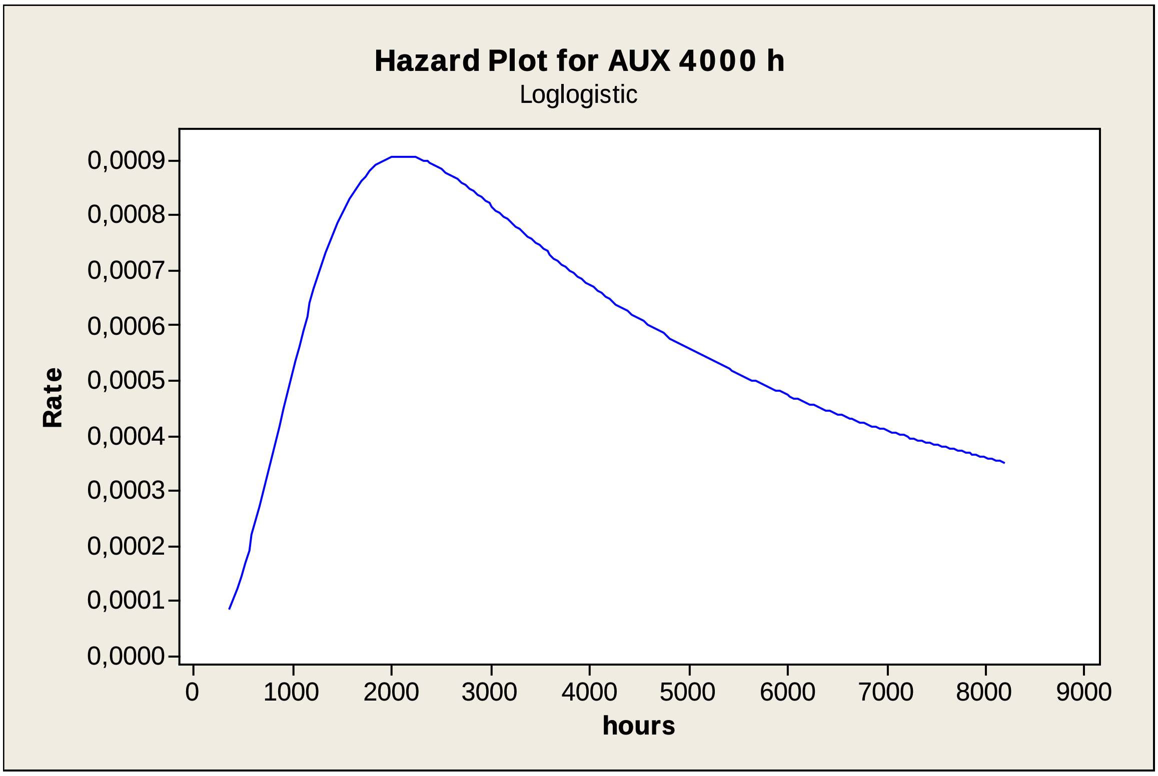


Figure 76: Hazard plot for AUX 4000

The shape of the failure rate curve indicates an increasing and then on/ about 1800 hours a decrease of the failure rate.

5.4.8 Cumulative failure plot

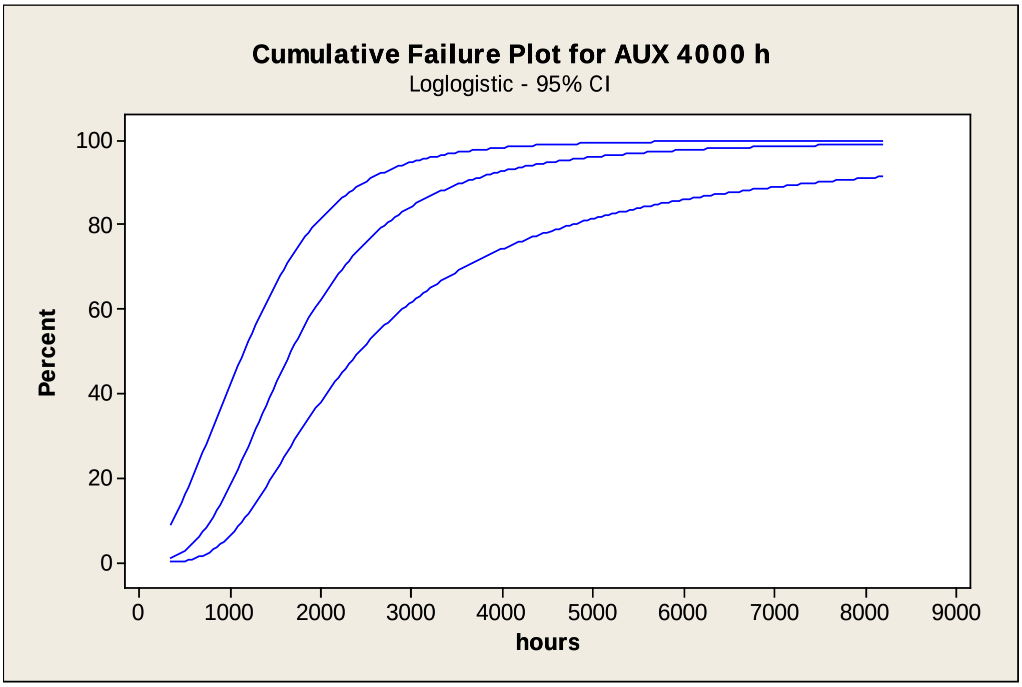


Figure 77: Cumulative failure plot for AUX 4000

Table 42 : table of cumulative failure probabilities

		95% confidence intervals	
time	probability	lower	upper
1000	0,180630786	0,063204885	0,418708739
2000	0,622862488	0,383038293	0,814588256
3000	0,842866233	0,616678942	0,9470471
4000	0,925222104	0,7414763	0,981609641
5000	0,959449395	0,813349821	0,992276181
6000	0,975719614	0,8583158	0,996262677
7000	0,984347669	0,888333723	0,997992561
8000	0,989326993	0,909407825	0,998833043
9000	0,992396404	0,924803418	0,999278551
10000	0,994390143	0,936417065	0,999531494
12000	0,996689374	0,952531707	0,99977865
13000	0,997374486	0,95827403	0,999840879
14000	0,997881983	0,962979177	0,999882829
15000	0,998266018	0,966887809	0,999911906

5.4.9 Non parametric analysis

Table 43: Non parametric estimates

```
              Standard    95,0% Normal CI
Mean(MTTF)       Error     Lower     Upper
      1936     304,981   1338,25   2533,75
```

A distribution free estimate of useful plots can be produced.

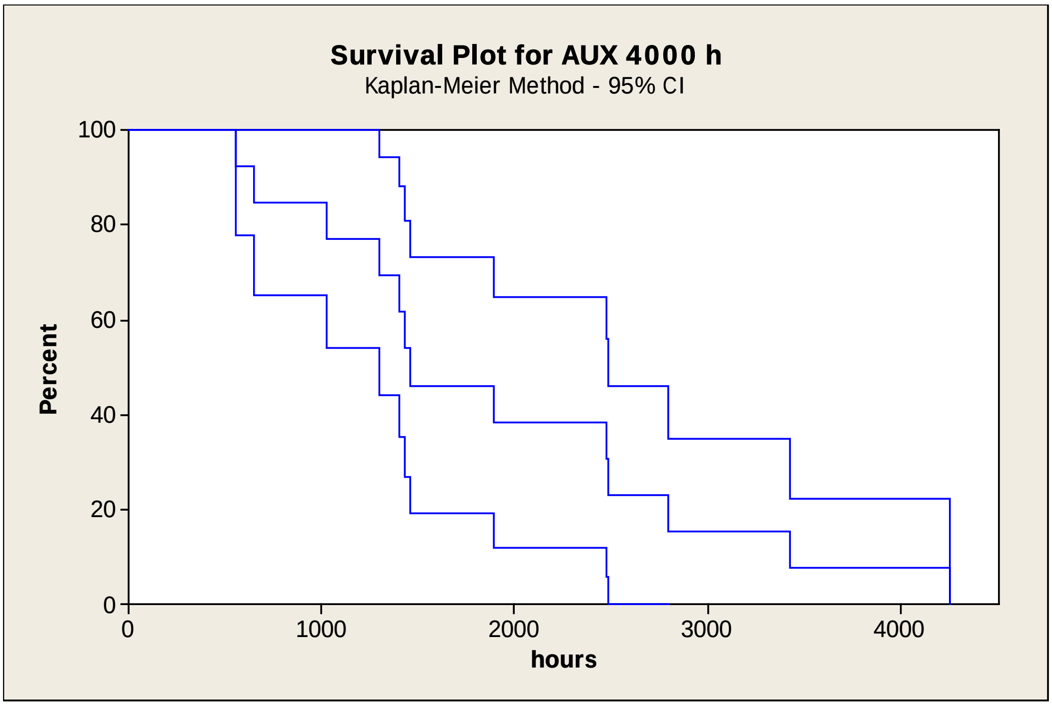


Figure 78: Survival plot for AUX 4000

Table 44 : kaplan meier estimates for survival probabilities

		95% confidence intervals	
time	survival probability	lower	upper
552	0,923076923	0,778225193	1
651	0,846153846	0,650023677	1
1031	0,769230769	0,540200075	0,998261464
1300	0,692307692	0,441417137	0,943198248
1400	0,615384615	0,350922749	0,879846482
1428	0,538461538	0,267468766	0,809454311
1463	0,461538462	0,190545689	0,732531234
1892	0,384615385	0,120153518	0,649077251
2480	0,307692308	0,056801752	0,558582863
2490	0,230769231	0,001738536	0,459799925
2800	0,153846154	0	0,349976323
3425	0,076923077	0	0,221774807
4256	0	0	0

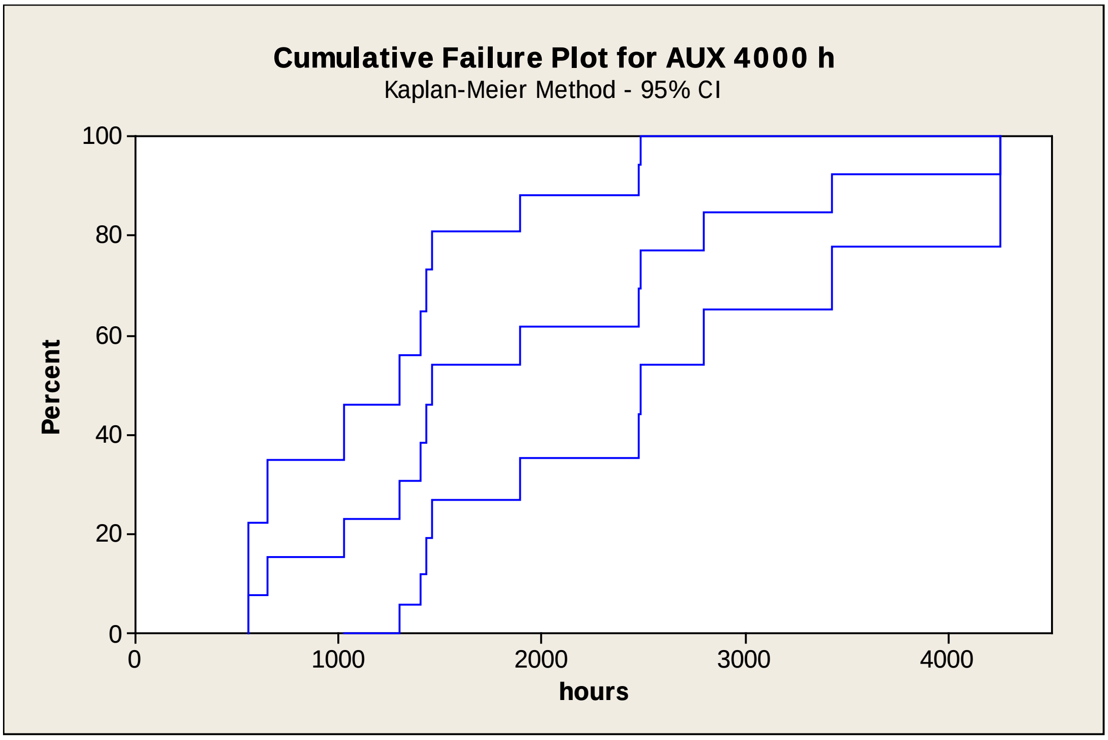


Figure 79: Cumulative failure plot for AUX 4000

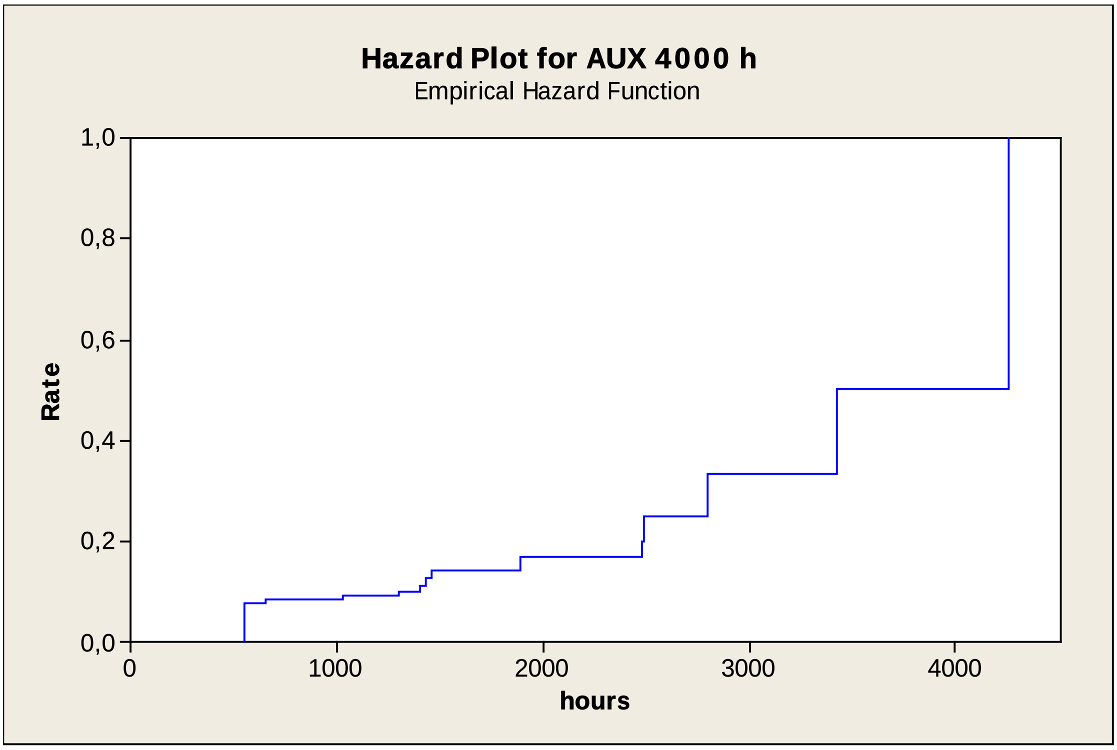


Figure 80: Hazard plot for AUX 4000

Table 45 : cumulative failure probabilities and hazard rates

		95% confidence intervals		
time	cumulative failure probability	lower	upper	hazard rates
552	0,076923077	0	0,221774807	0,076923077
651	0,153846154	0	0,349976323	0,083333333
1031	0,230769231	0,001738536	0,459799925	0,090909091
1300	0,307692308	0,056801752	0,558582863	0,1
1400	0,384615385	0,120153518	0,649077251	0,111111111
1428	0,461538462	0,190545689	0,732531234	0,125
1463	0,538461538	0,267468766	0,809454311	0,142857143
1892	0,615384615	0,350922749	0,879846482	0,166666667
2480	0,692307692	0,441417137	0,943198248	0,2
2490	0,769230769	0,540200075	0,998261464	0,25
2800	0,846153846	0,650023677	1	0,333333333
3425	0,923076923	0,778225193	1	0,5
4256	1	1	1	1

Regarding the mean remaining life of this category of items the function MRL/MTTF can be calculated.

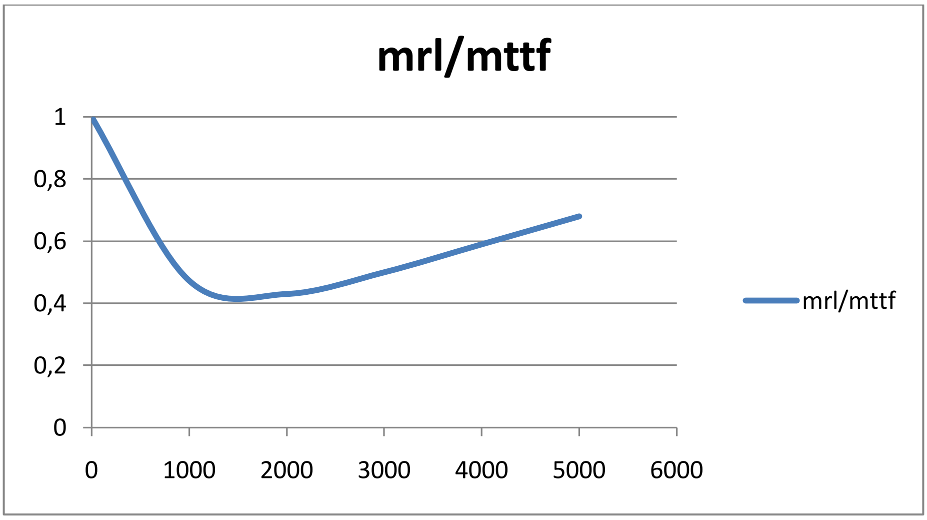


Figure 81: MRL/MTTF for AUX 4000

In accordance to the failure rate function, Figure 76: Hazard plot for AUX 4000, it is noted an increase in failure rate until about 1500 hours and then a slight decrease.

5.5 AUX 6000

With the same procedure, we can analyze failures in auxiliary equipment, in which the manufacturer suggests 6000 working hours between inspections.

5.5.1 Outlier points

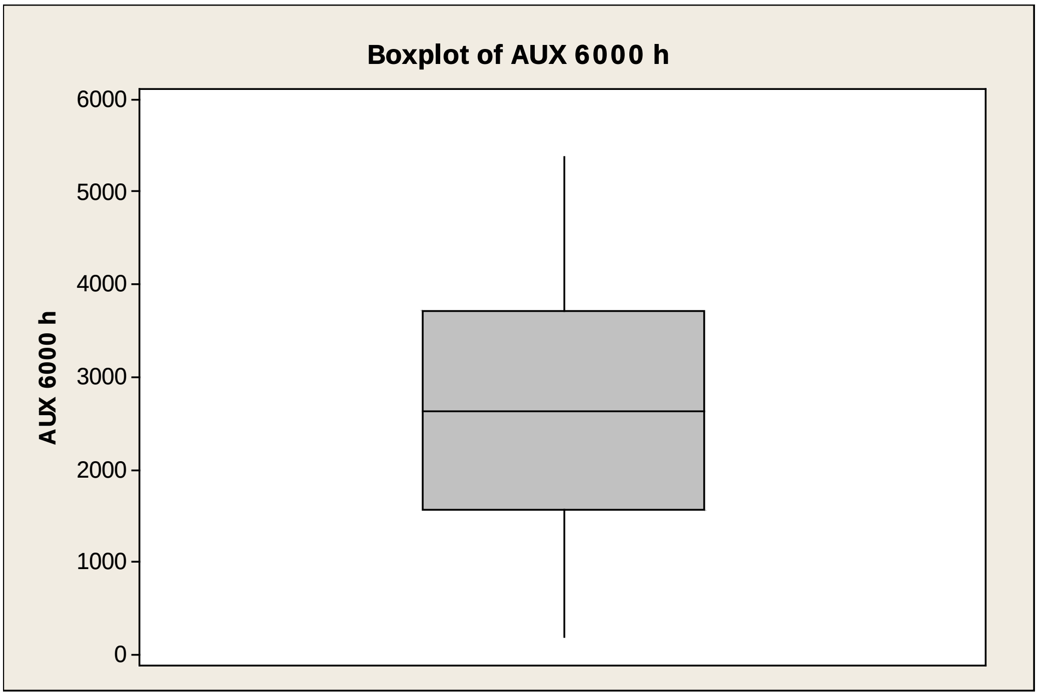


Figure 82: Boxplot of AUX 6000

From this plot we conclude that no outlier points exist in this group of data and the following parameters are calculated.

Table 46: Basic parameters

Median	Q1	Q3	IQR	lower outer fence	upper outer fence
2645	1578.75	3720.75	2142	196	5380

5.5.2 Basic statistics calculations

Table 47: Basic statistic calculations

Mean	Standard error of Mean	Standard Deviation	Variance	Coefficient of Variation	Q1	Median	Q3
2681,4	484,73	1532,86	2349667	57,166	1578,75	2645	3720,7
IQR	Sum	Minimum	Maximum	Range	Sum of Squares	Skewness	Kurtosis
2142	26814	196	5380	5184	93046070	0,14113339	-0,4137

A graphical summary of the above:

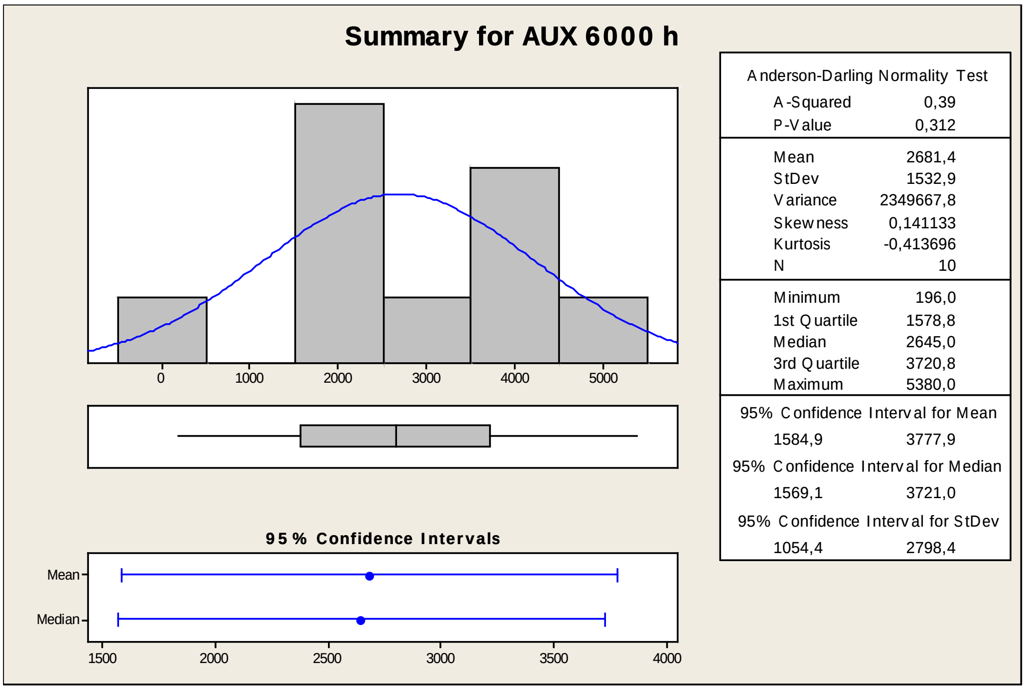


Figure 83: Graphical summary for AUX 6000

Additionally we can see 95% confidence intervals for the mean, median and standard deviation.

5.5.3 Histogram

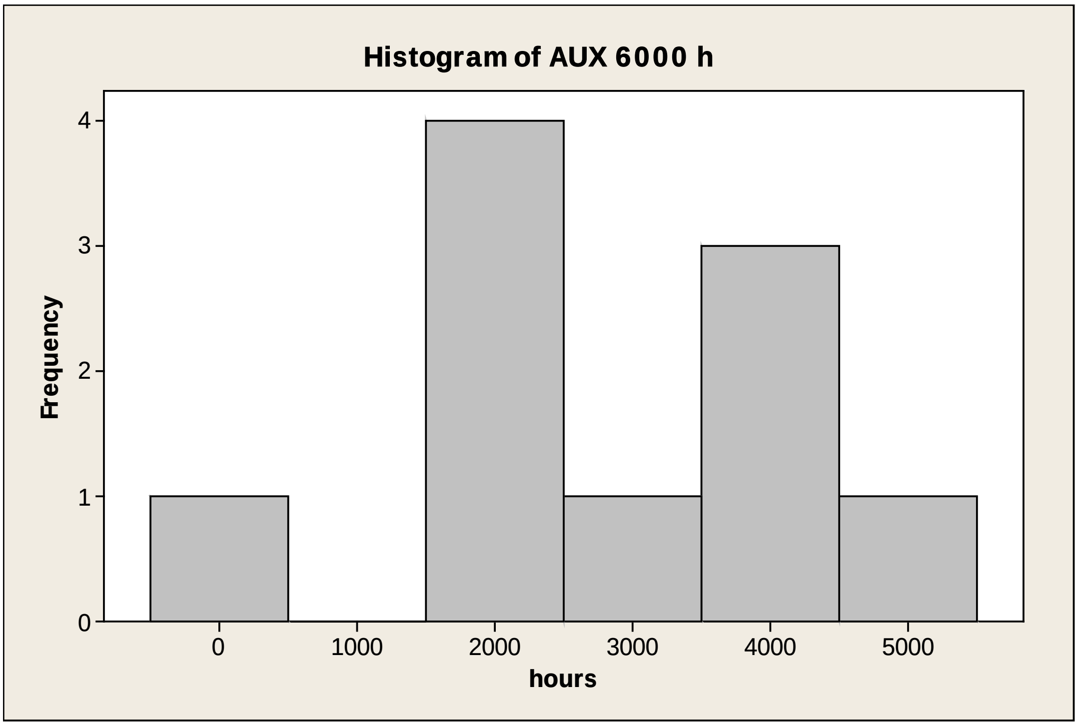


Figure 84: Histogram of AUX 6000

5.5.4 Cumulative distribution function

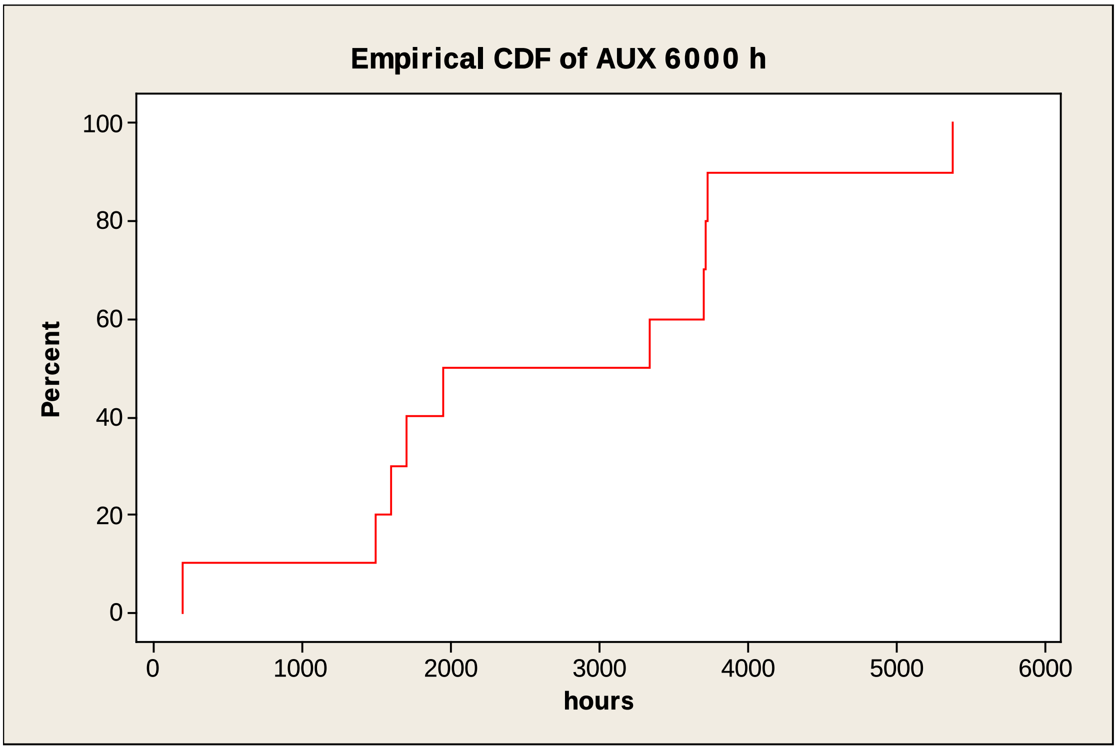


Figure 85: Empirical CDF of AUX 6000

As described in previous chapter, the empirical cdf is an estimator of the cumulative distribution function. To further define the proper distribution which fits the data well, these have to be plotted in special plotting paper.

5.5.5 Individual distribution identification

Using the statistical software Minitab, the following graphs were obtained for the identification of the proper distribution that describes this category of data. It is reminded that not only the good eye ball fit is necessary, but the various goodness of fit tests provide the final necessary information, that will allow to choose the appropriate distribution among all the experimentally tested.

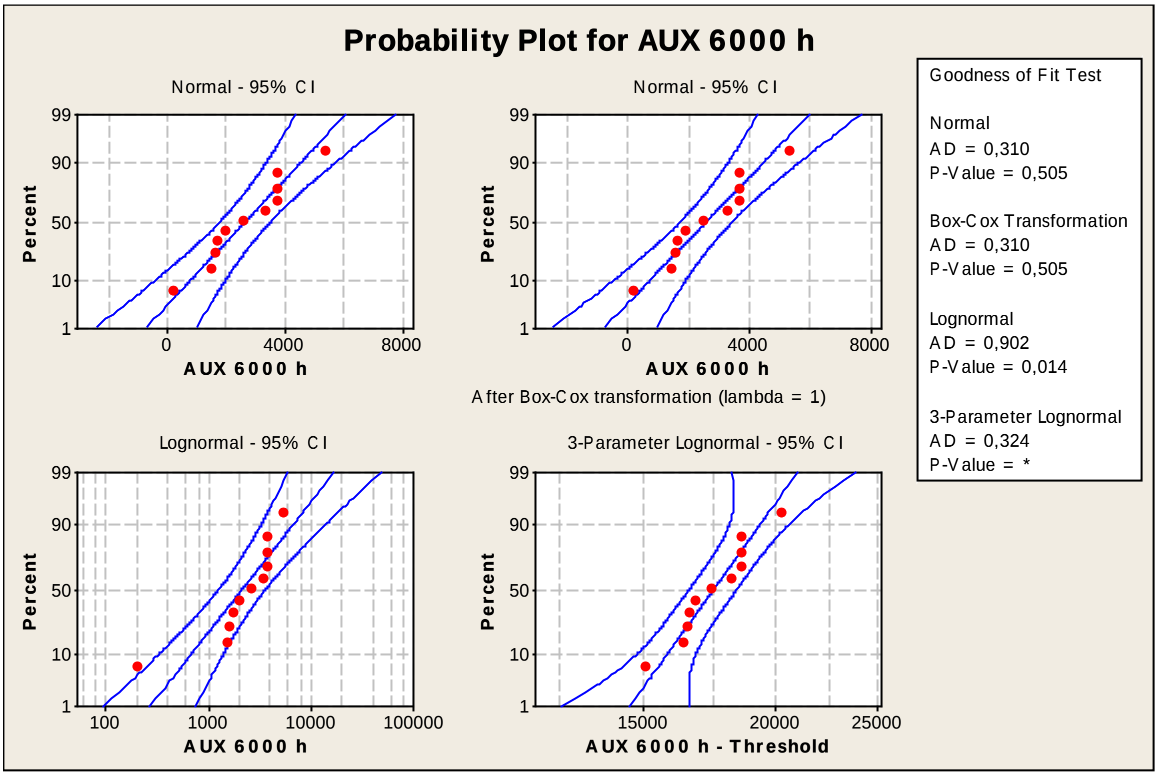


Figure 86: Distribution identification for AUX 6000

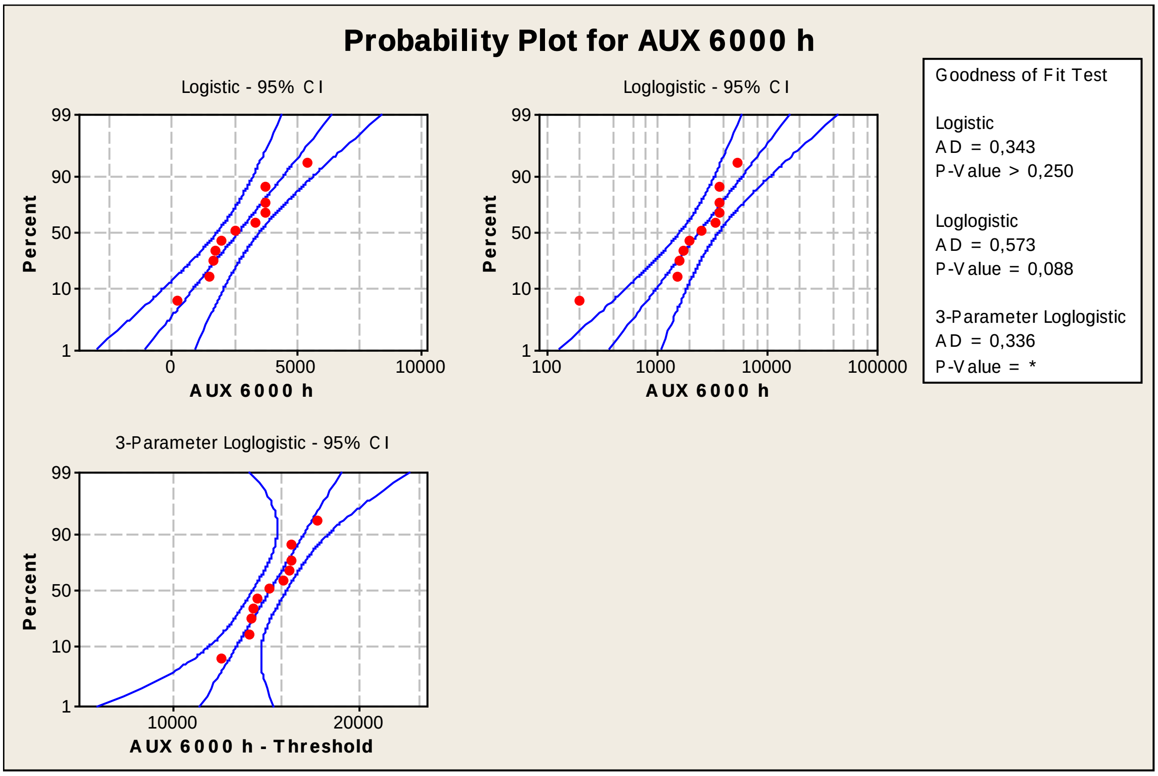


Figure 87: Distribution identification for AUX 6000

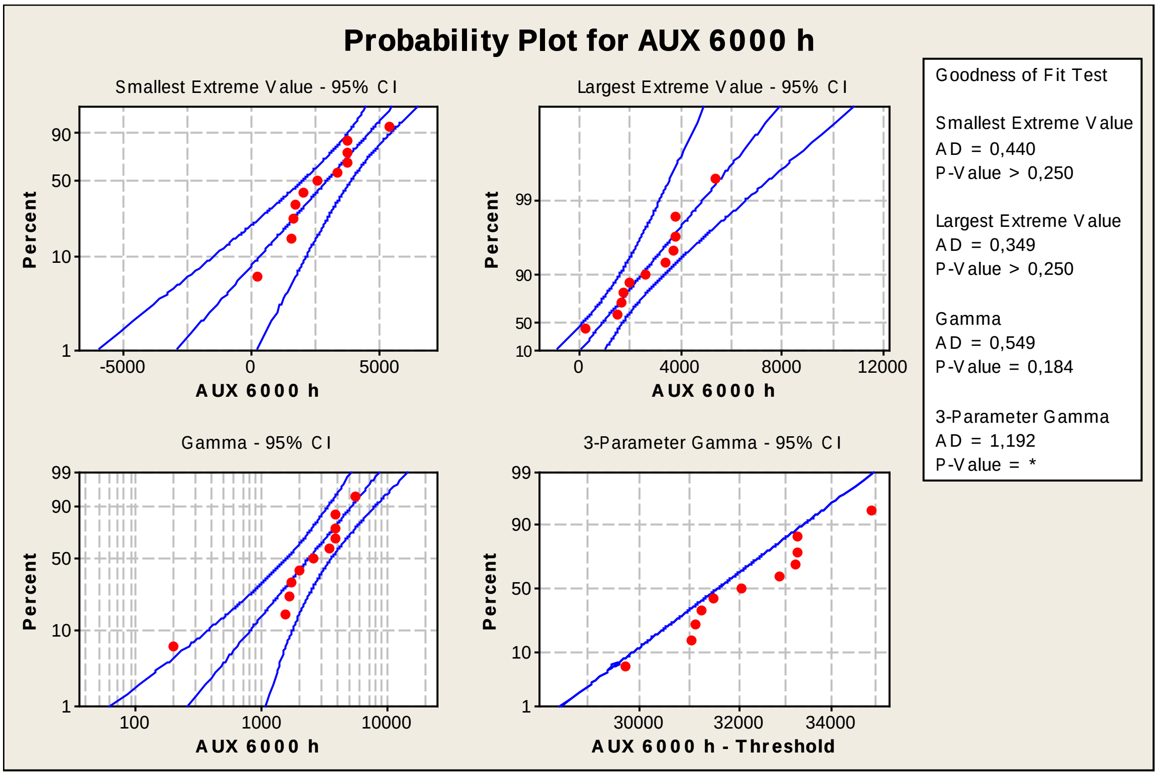


Figure 88: Distribution identification for AUX 6000

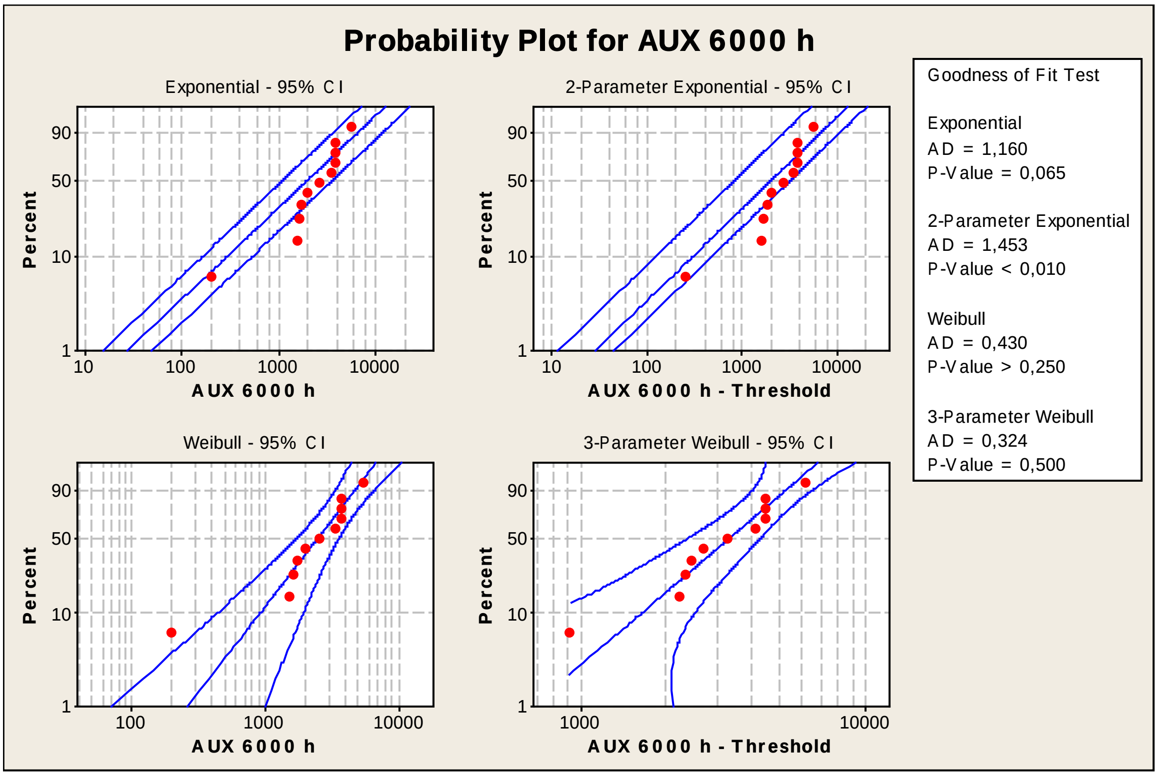


Figure 89: Distribution identification for AUX 6000

It is time to compare various goodness of fit test to decide which distribution is appropriate for this set of data.

Table 48: Comparison of goodness of fit tests for AUX 6000

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	0,310	0,505	
Box-Cox Transformation	0,310	0,505	
Lognormal	0,902	0,014	
3-Parameter Lognormal	0,324	*	0,017
Exponential	1,160	0,065	
2-Parameter Exponential	1,453	<0,010	1,000
Weibull	0,430	>0,250	
3-Parameter Weibull	0,324	0,500	0,424
Smallest Extreme Value	0,440	>0,250	
Largest Extreme Value	0,349	>0,250	
Gamma	0,549	0,184	
3-Parameter Gamma	1,192	*	1,000
Logistic	0,343	>0,250	
Loglogistic	0,573	0,088	
3-Parameter Loglogistic	0,336	*	0,081

Checking the Anderson Darling and p-value test, it is concluded that the normal distribution fits adequately well our data.

In order to define subject distribution it is necessary to estimate its parameters.

For the normal distribution we need the mean and a standard deviation.

For this distribution:

Mean = 2668.55

Standard deviation = 1387.12

Summarizing, the fitted distribution for the first group of failures, is fully defined.

Normal distribution of times to failure:

- CDF: $F(x) = \frac{1}{2}[1 + \operatorname{erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)], x \in R$
- PDF:f(x)=$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in R$

Where μ=2668.55
σ=1387.12

where erf is a function sometimes called the error function which can't be expressed in terms of finite additions, subtractions, multiplications, and root extractions, and so must be either computed numerically or otherwise approximated.

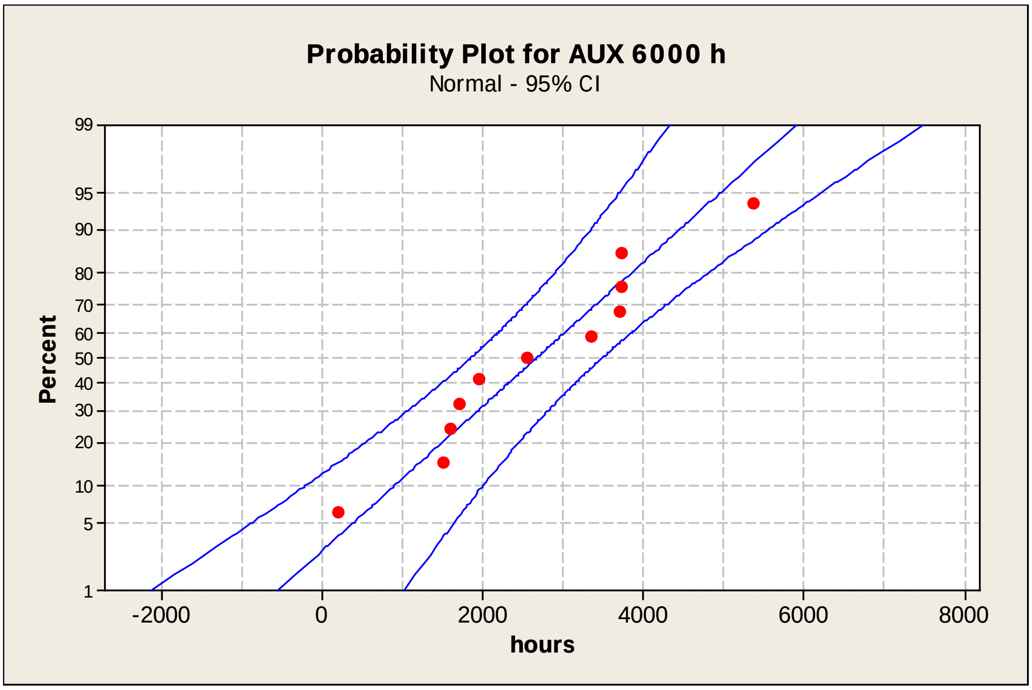


Figure 90: Plot of Normal fit to AUX 6000 data

In the plotted data in normal plotting paper, it is understood that this distribution fits well the collected life data.

5.5.6 Reliability concepts

Relevant plots for this group of data can be produced, similarly to previous chapter.

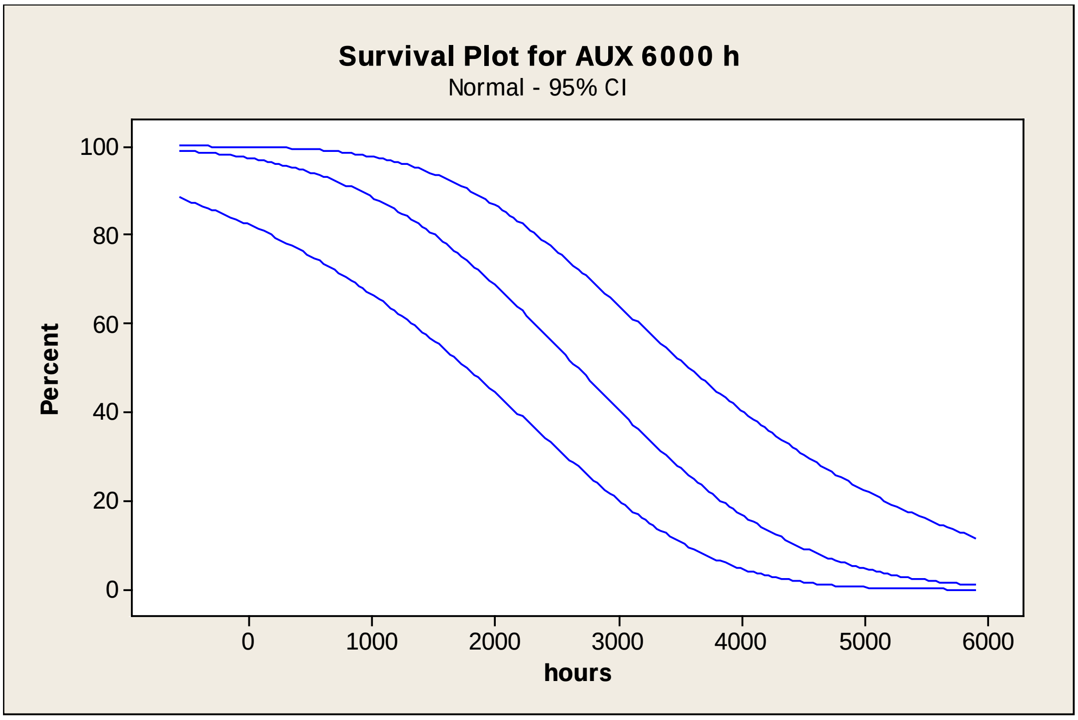


Figure 91 : Plot of reliability function with 95% confidence intervals

Table 49 : table of survival probabilities

		95% confidence intervals	
time	probability	lower	upper
1000	0,8854892	0,665338538	0,976074442
2000	0,6850849	0,443397334	0,865700111
3000	0,4055717	0,200937088	0,640717117
4000	0,1685608	0,047057221	0,40297492
5000	0,0464023	0,004679312	0,222760735
6000	0,0081595	0,000180991	0,108034221
7000	0,0008962	2,61882E-06	0,045494656
8000	6,064E-05	1,38895E-08	0,016495816
9000	2,504E-06	2,66824E-11	0,005119162
10000	6,273E-08	1,84297E-14	0,001353887
12000	8,648E-12	0	5,79377E-05
13000	4,73E-14	0	9,33627E-06
14000	1,11E-16	0	1,27118E-06
15000	0	0	1,46079E-07

For the Normal distributed category AUX 6000 the conditional probabilities for a range of 1000 to 10000 hours is plotted in the following graph.

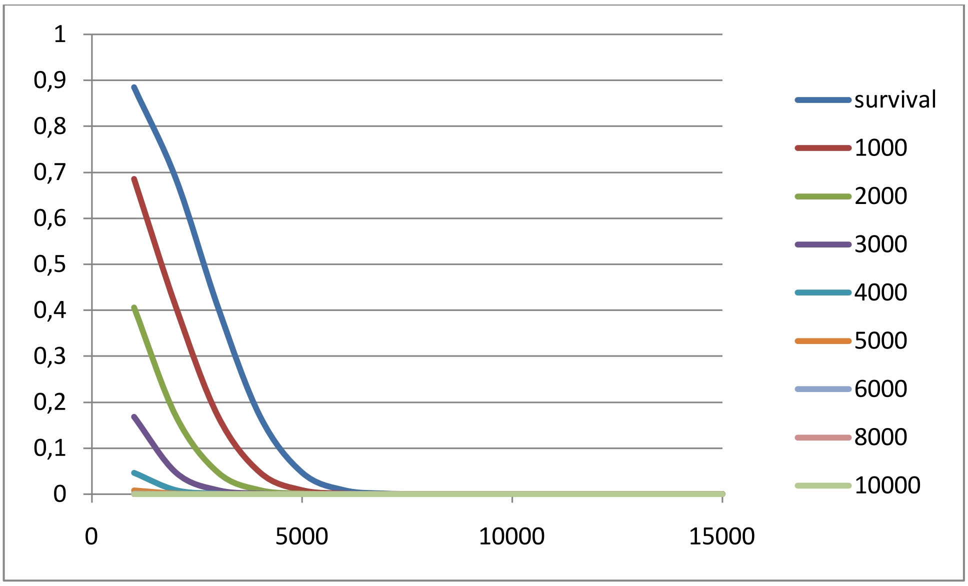


Figure 92: Conditional probabilities for AUX 6000

For an item that has survived 2000 hours, the conditional probabilities to survive additionally t hours are:

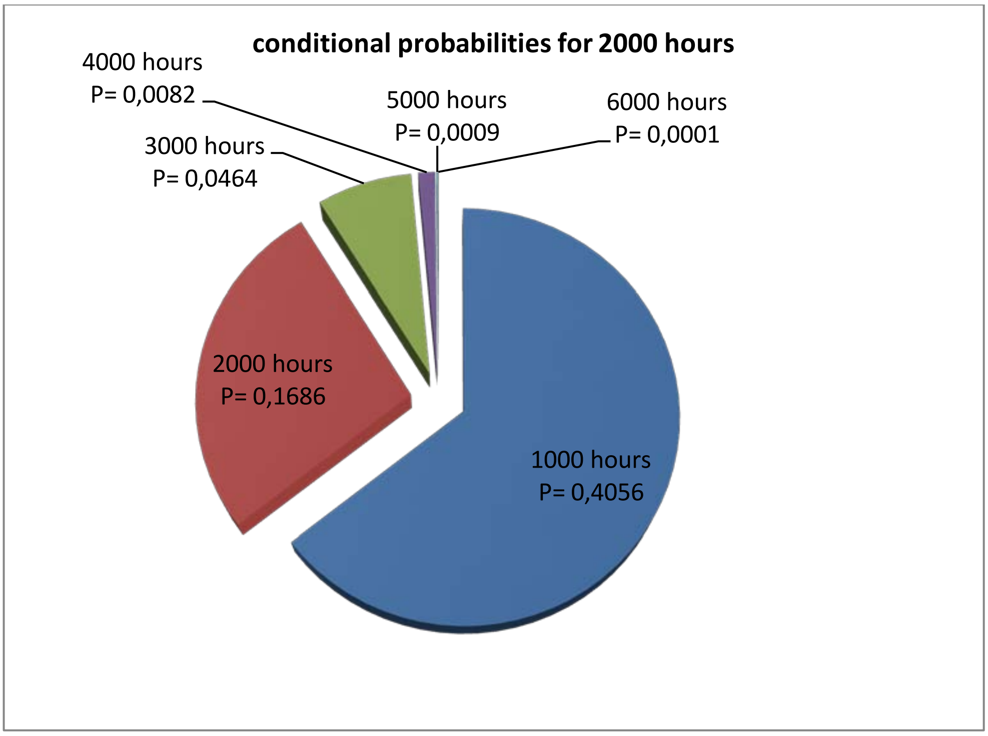


Figure 93: Conditional probabilities for AUX 6000 at 2000 hours

This plot in a tabulated view:

Table 50: Conditional probabilities for AUX 6000 at 2000 hours

t	P
1000	0,405572827
2000	0,168561208
3000	0,046402376
4000	0,00815948
5000	0,000896242
6000	6,06363E-05
8000	6,27236E-08
10000	8,64719E-12

5.5.7 Failure rate plot

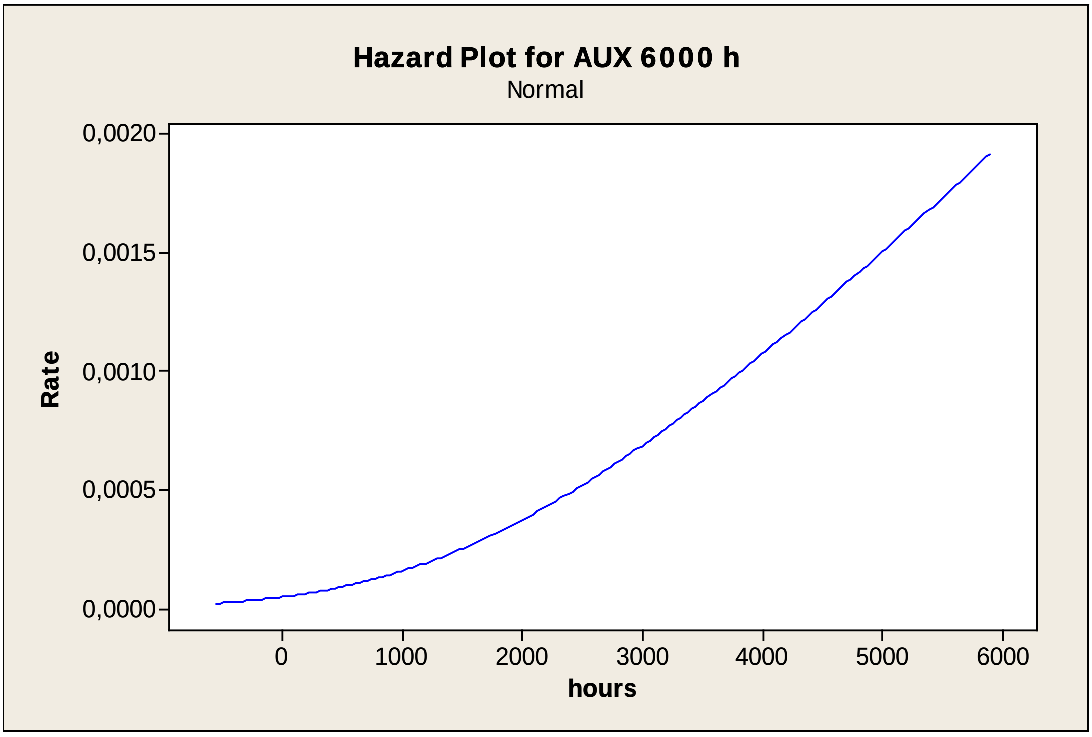


Figure 94: Hazard plot for AUX 6000

The units of this rate are failures per unit time. It is the failure rate of the survivors to time t in the very next instant following t. It is not a probability and it can have values greater than 1.

5.5.8 Cumulative failure plot

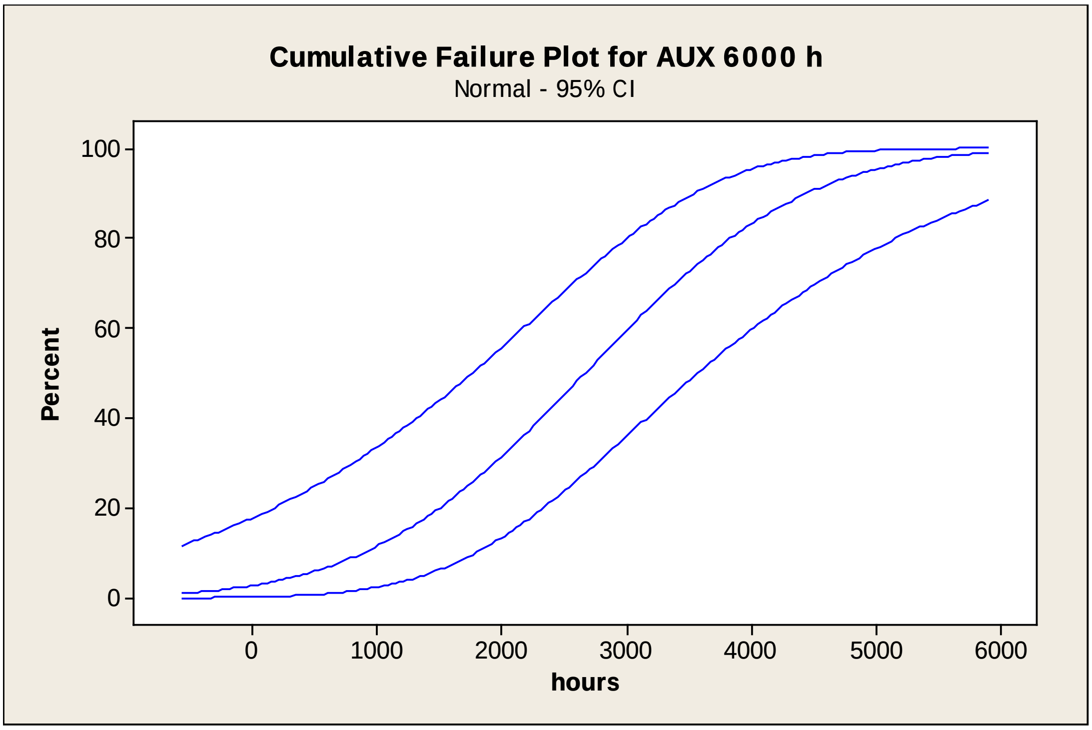


Figure 95: Cumulative failure plot for AUX 6000

Table 51: table of cumulative failure probabilities

		95% confidence intervals	
time	probability	lower	upper
1000	0,114510829	0,023925558	0,334661462
2000	0,314915117	0,134299889	0,556602666
3000	0,5944283	0,359282883	0,799062912
4000	0,831439241	0,59702508	0,952942779
5000	0,953597688	0,777239265	0,995320688
6000	0,99184051	0,891965779	0,999819009
7000	0,999103753	0,954505344	0,999997381
8000	0,999939363	0,983504184	0,999999986
9000	0,999997496	0,994880838	1
10000	0,999999937	0,998646113	1
12000	1	0,999942062	1
14000	1	0,999998729	1
15000	1	0,999999854	1

5.5.9 Non parametric analysis

Table 52: Non parametric estimates

```
            Standard   95,0% Normal CI
Mean(MTTF)     Error     Lower     Upper
   2668,55   438,647   1808,81   3528,28
```

A distribution free estimate of useful plots can be produced.

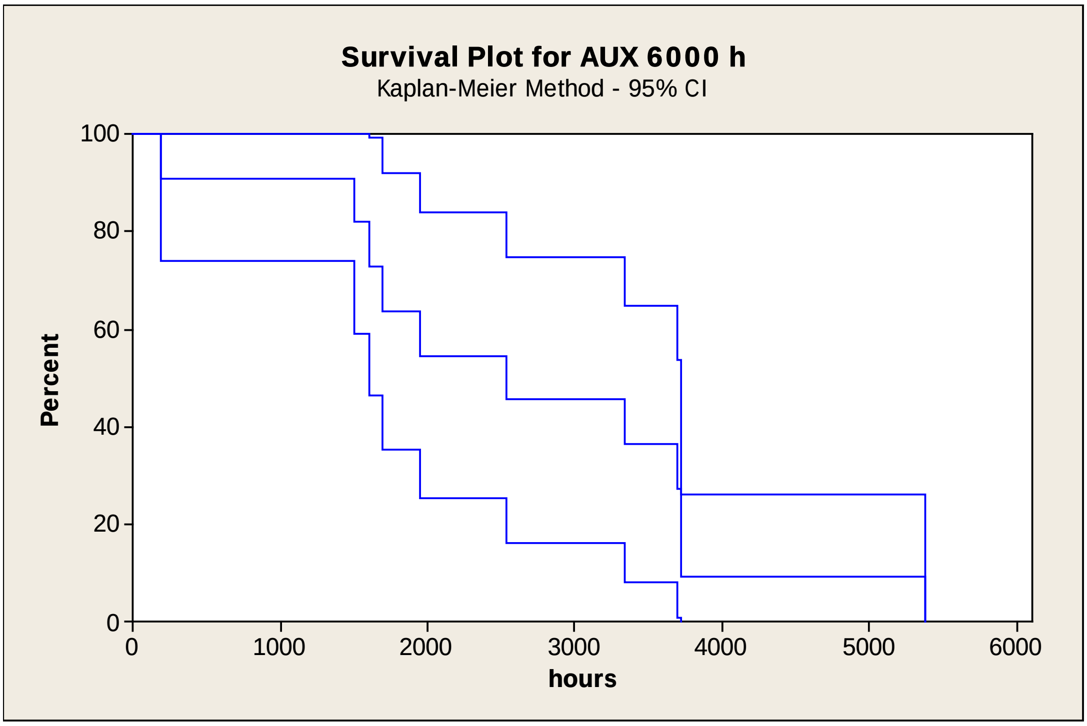


Figure 96: Survival plot for AUX 6000

Table 53: kaplan meier estimates for survival probabilities

		95% confidence intervals	
time	survival probability	lower	upper
196	0,9090909	0,739204333	1
1500	0,8181818	0,590255059	1
1605	0,7272727	0,464085576	0,990459879
1700	0,6363636	0,352089022	0,920638251
1950	0,5454545	0,251202364	0,839706727
2540	0,4545455	0,160293273	0,748797636
3340	0,3636364	0,079361749	0,647910978
3700	0,2727273	0,009540121	0,535914424
3720	0,1818182	0	0,409744941
3723	0,0909091	0	0,260795667
5380	0	0	0

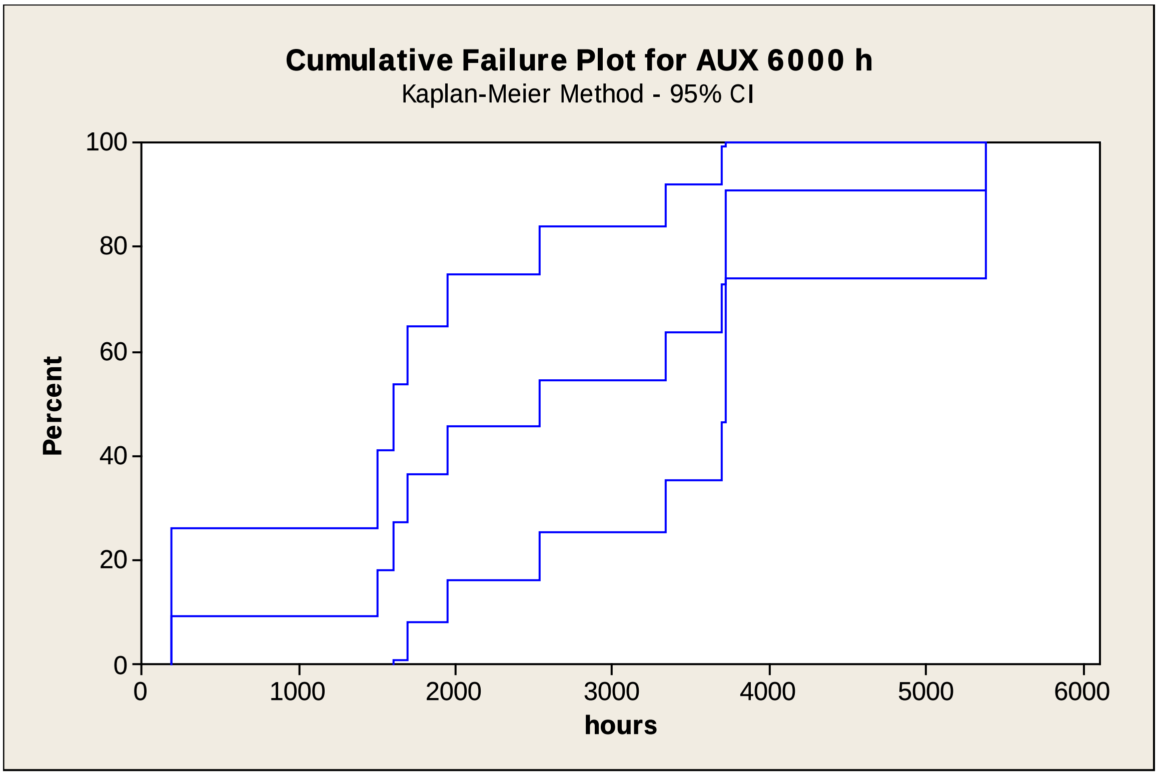


Figure 97: Cumulative failure plot for AUX 6000

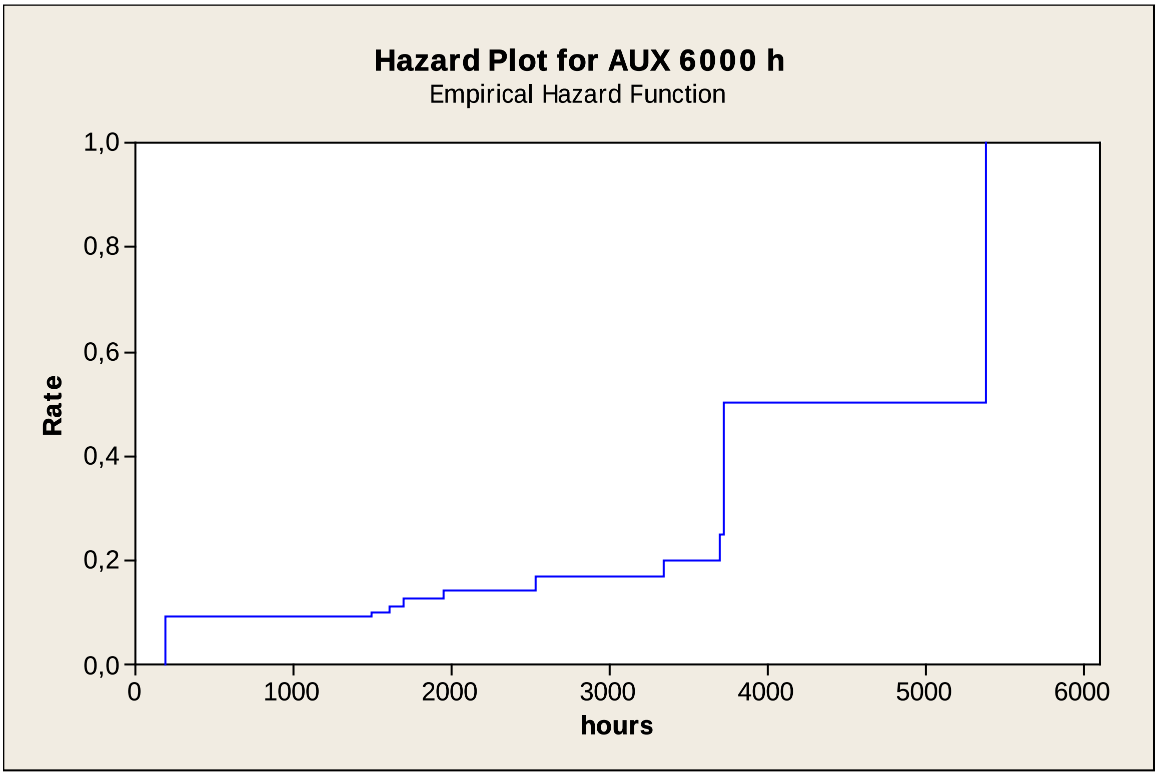


Figure 98: Hazard plot for AUX 6000

Table 54 : cumulative failure probabilities and hazard rates

		95% confidence intervals		
time	cumulative failure probability	lower	upper	hazard rates
196	0,090909091	0	0,260795667	0,090909091
1500	0,181818182	0	0,409744941	0,1
1605	0,272727273	0,009540121	0,535914424	0,111111111
1700	0,363636364	0,079361749	0,647910978	0,125
1950	0,454545455	0,160293273	0,748797636	0,142857143
2540	0,545454545	0,251202364	0,839706727	0,166666667
3340	0,636363636	0,352089022	0,920638251	0,2
3700	0,727272727	0,464085576	0,990459879	0,25
3720	0,818181818	0,590255059	1	0,333333333
3723	0,909090909	0,739204333	1	0,5
5380	1	1	1	1

The mean remaining life function, versus the MTTF is plotted in the next figure.

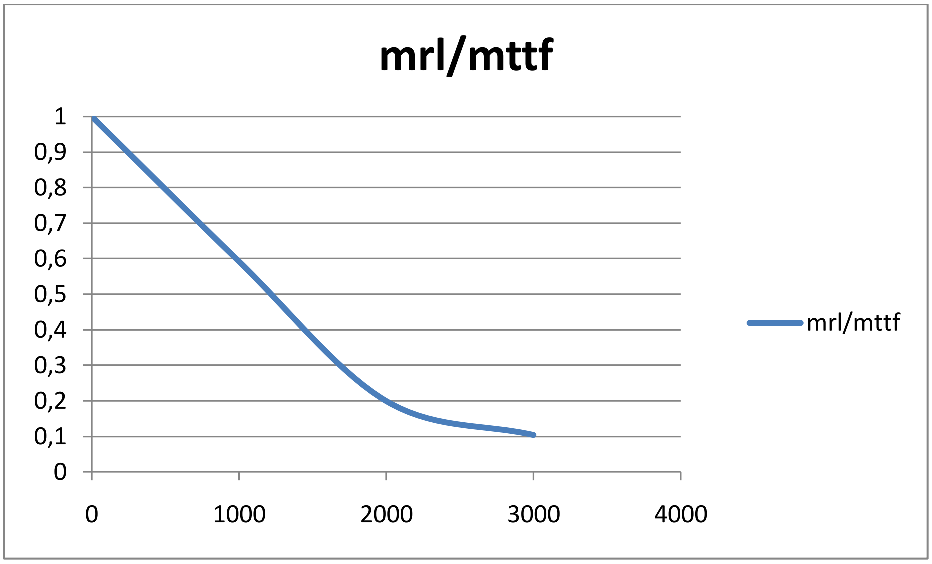


Figure 99: MRL/MTTF for AUX 6000

As it was noticed in the failure rate function, the mean remaining life is decreasing, for example at 2000 hours the remaining life of an item will be the 20% of the remaining life at time 0.

5.6 AUX 8000

The last group of data concerns failures in auxiliary equipment, in which the manufacturer suggests 8000 working hours between inspections.

5.6.1 Outlier points

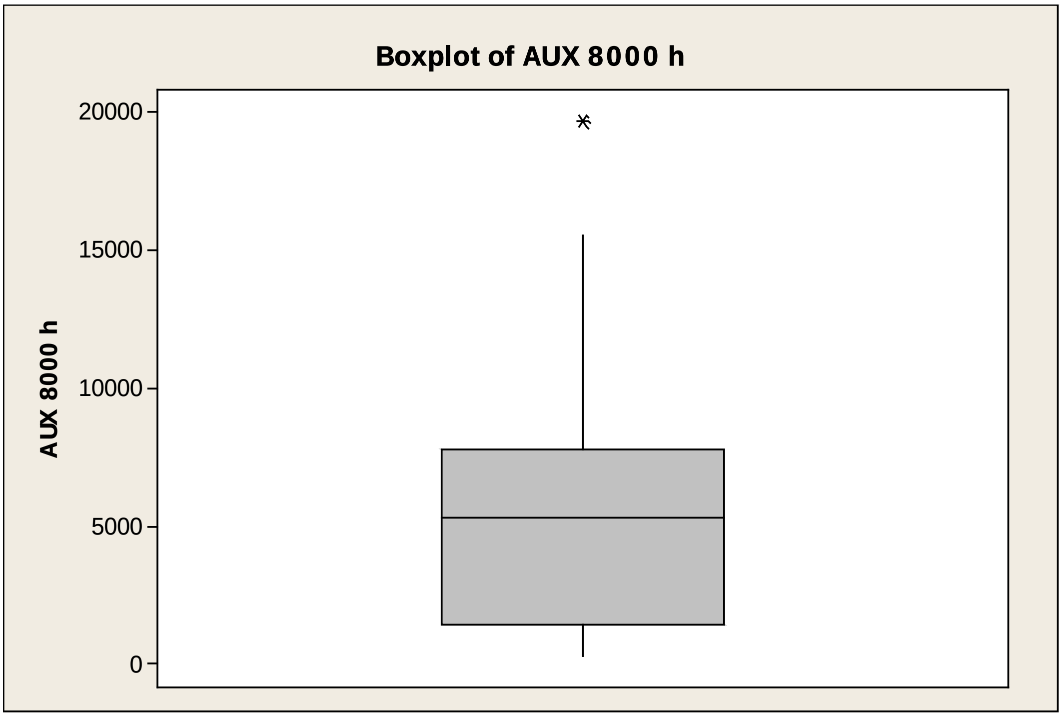


Figure 100: Boxplot of AUX 8000

It is obvious that there is an outlier point in this set of data. After checking our database, this failure in 19624 hours, concern broken ball bearings in the electric motor of hot water circulating pump and can't be ignored. The following parameters are calculated.

Table 55: Basic parameters

Median	Q1	Q3	IQR	lower outer fence	upper outer fence
5312	1462	7800	6338	345	15544

5.6.2 Basic statistics calculations

Some basic statistics calculations such as mean, Standard Error of mean, standard deviation variance, coefficient of variation first quartile, etc can be found in table 56

Table 56: Basic statistics calculations

Mean	Standard error of Mean	Standard Deviation	Variance	Coefficient of Variation	Q1	Median	Q3
5157,57	713,14	4219,03	17800262	81,802	1462	5312	7800
IQR	Sum	Minimum	Maximum	Range	Sum of Squares	Skewness	Kurtosis
6338	180515	345	19624	19279	1,54E+09	1,4739	3,3974

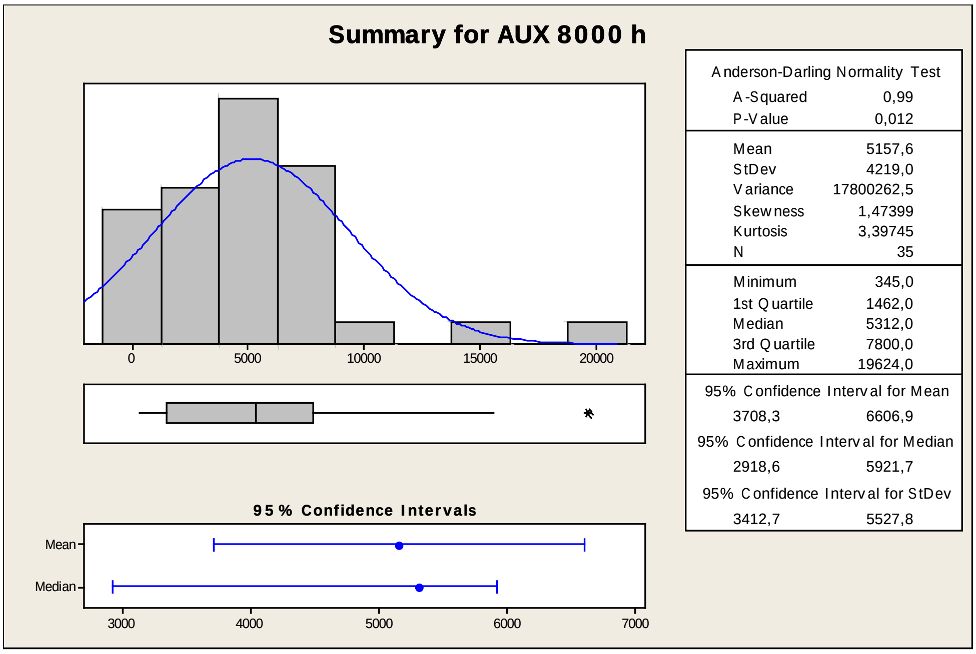


Figure 101: Graphical summary for AUX 8000

A significant concentration of failures in the area of 5000 operational hours is indicated in this graph, which affects the mean and median calculations.

5.6.3 Histogram

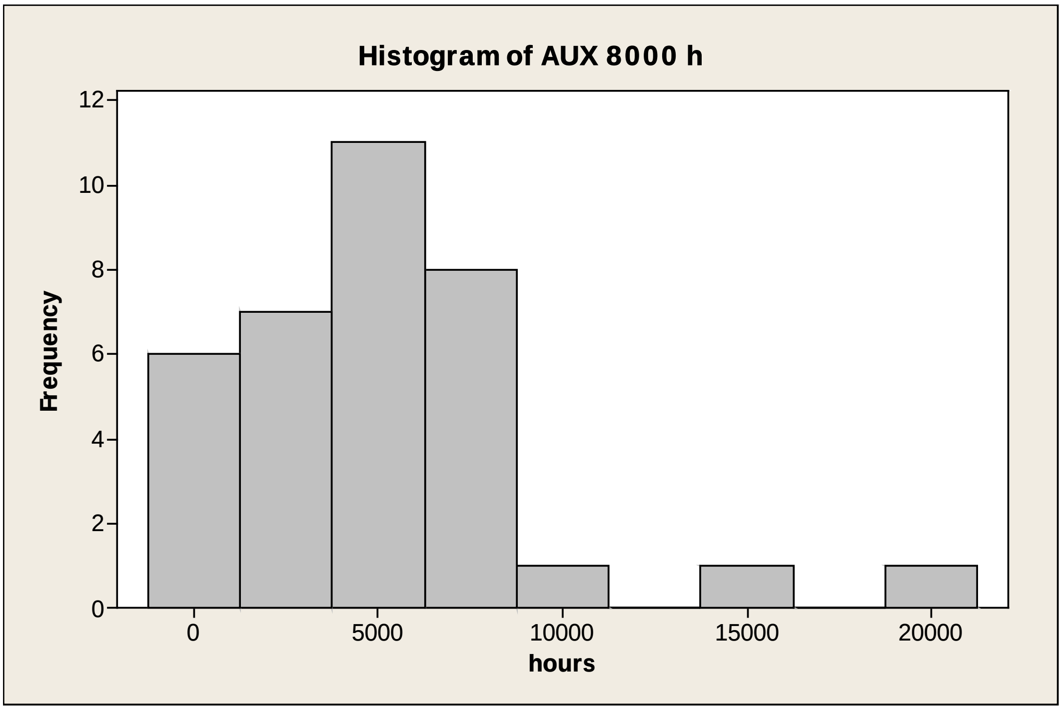


Figure 102: Histogram of AUX 8000

5.6.4 Cumulative distribution function

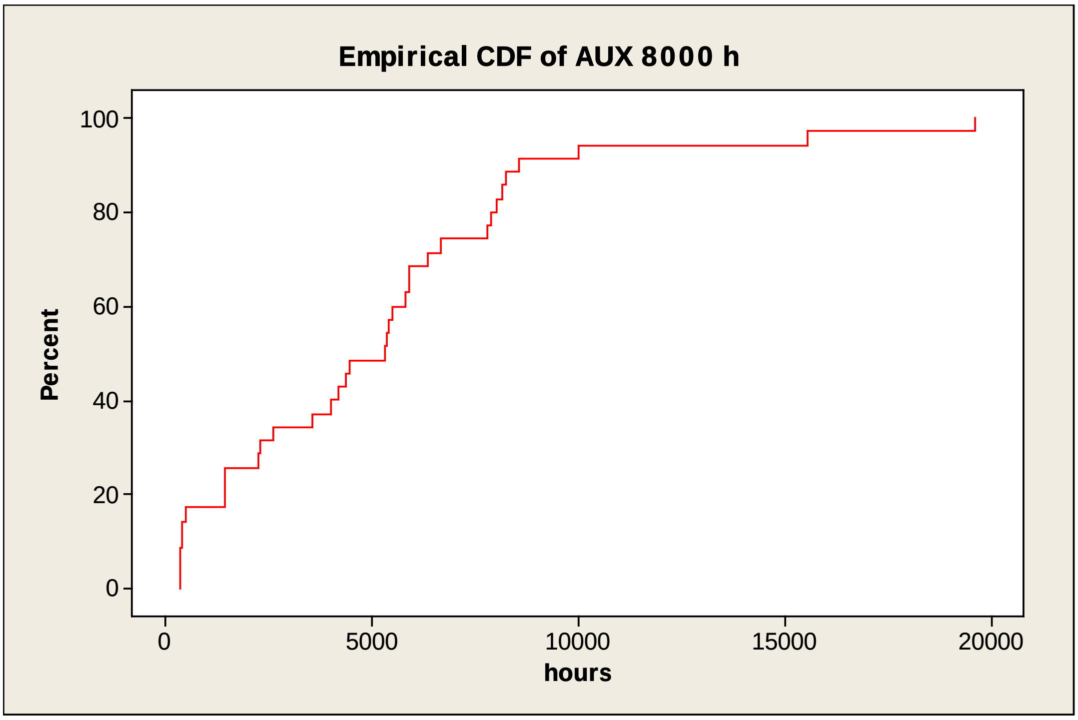


Figure 103: Empirical CDF of AUX 8000

As described in previous chapter, the empirical cdf is an estimator of the cumulative distribution function. To further define the proper distribution which fits the data well, these have to be plotted in special plotting paper.

5.6.5 Individual distribution identification

Using the statistical software Minitab, the following graphs were obtained.

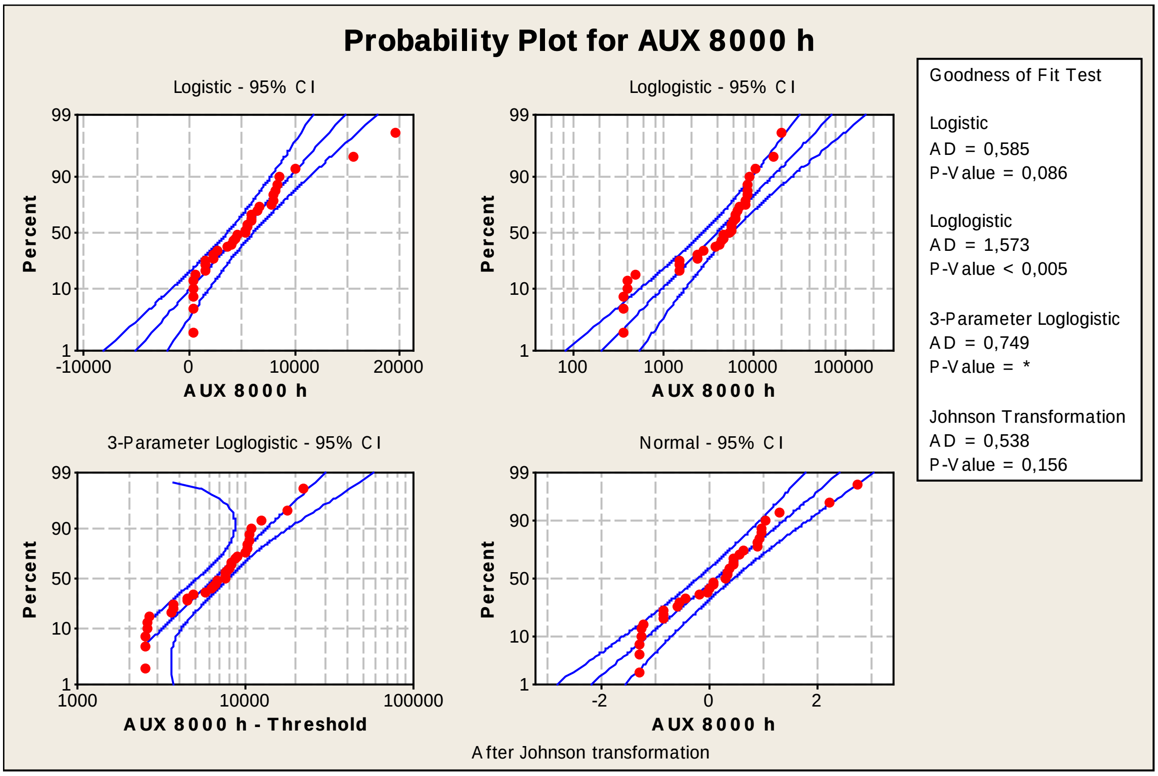


Figure 104: Distribution identification for AUX 8000

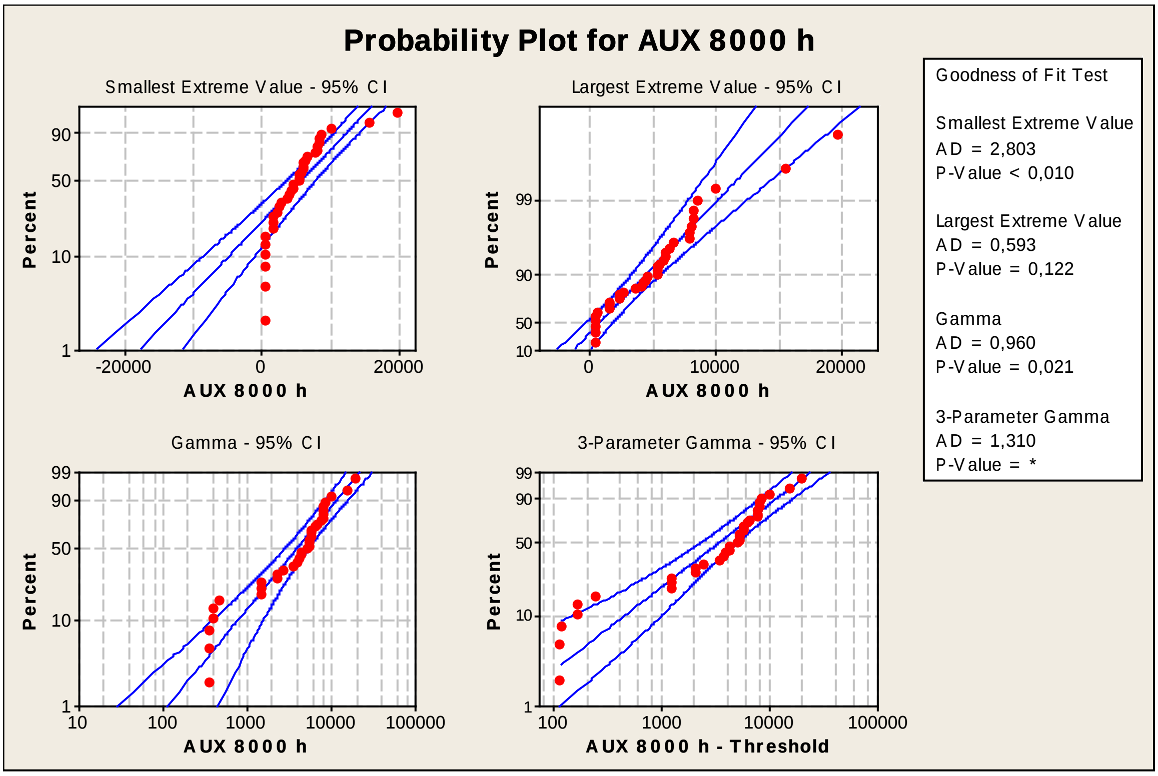


Figure 105: Distribution identification for AUX 8000

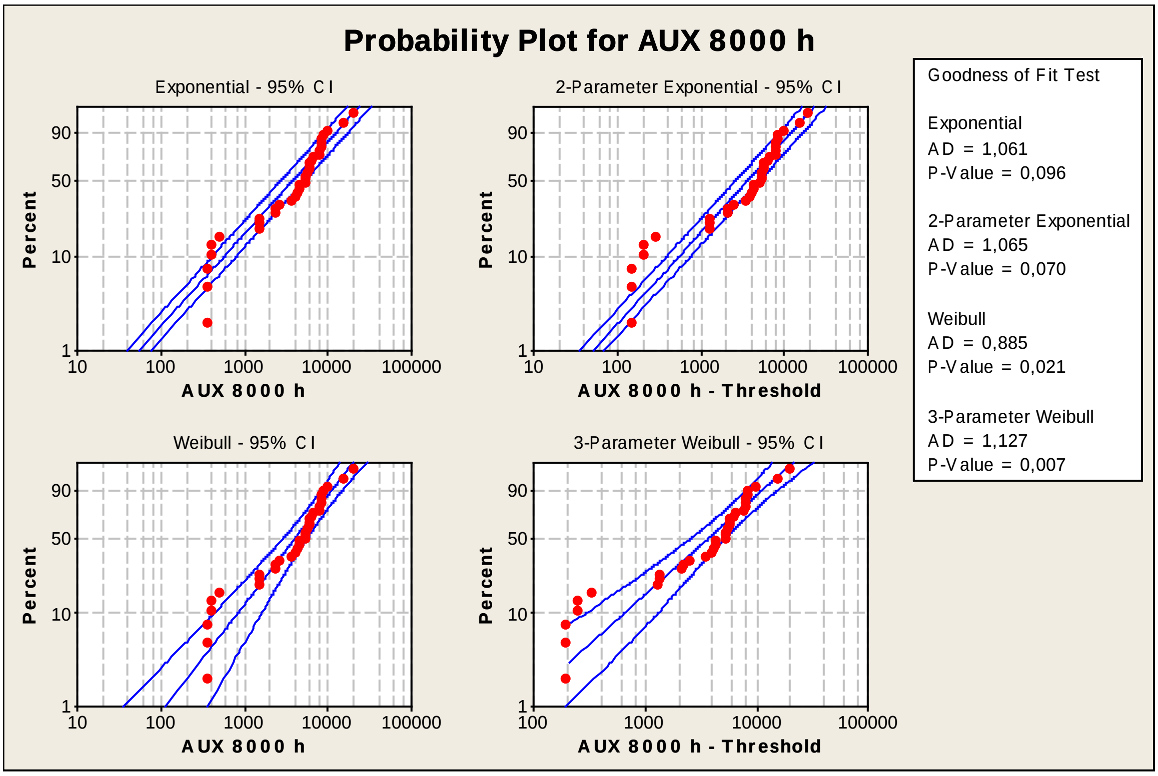


Figure 106: Distribution identification for AUX 8000

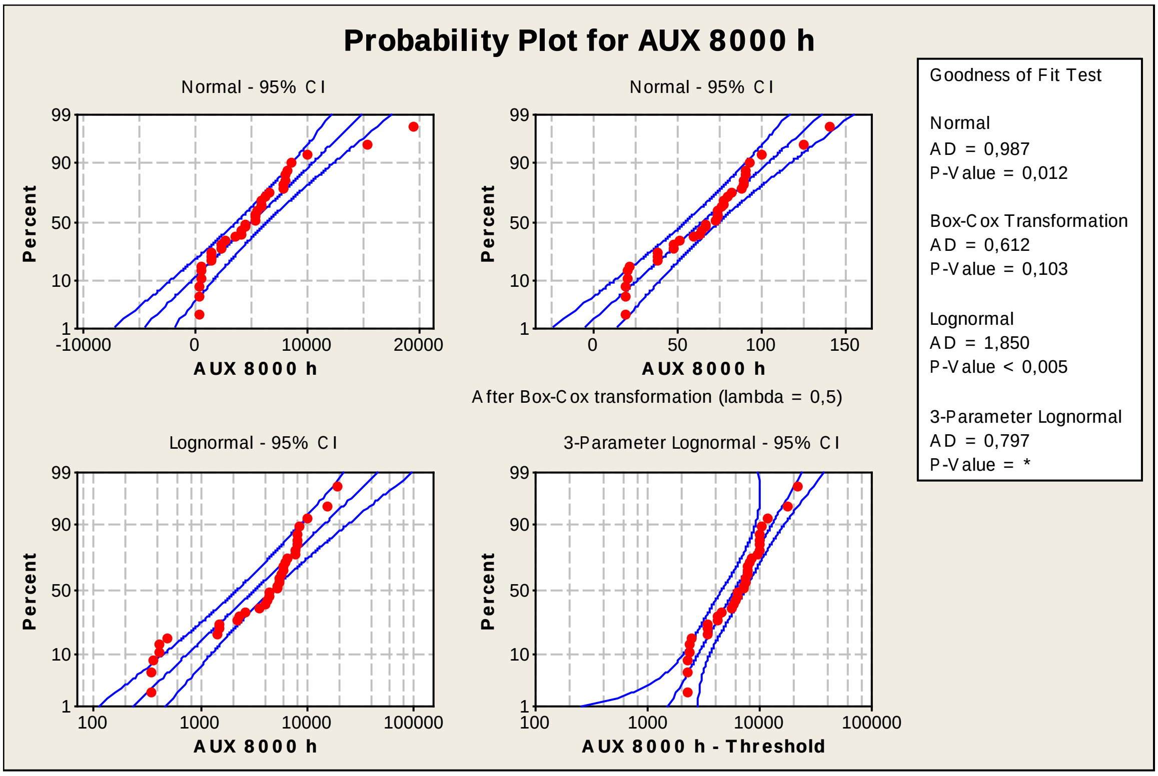


Figure 107: Distribution identification for AUX 8000

Table 57: Comparison of goodness of fit tests for AUX 8000

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	0,987	0,012	
Box-Cox Transformation	0,612	0,103	
Lognormal	1,850	<0,005	
3-Parameter Lognormal	0,797	*	0,078
Exponential	1,061	0,096	
2-Parameter Exponential	1,065	0,070	0,093
Weibull	0,885	0,021	
3-Parameter Weibull	1,127	0,007	0,292
Smallest Extreme Value	2,803	<0,010	
Largest Extreme Value	0,593	0,122	
Gamma	0,960	0,021	
3-Parameter Gamma	1,310	*	0,113
Logistic	0,585	0,086	
Loglogistic	1,573	<0,005	
3-Parameter Loglogistic	0,749	*	0,189
Johnson Transformation	0,538	0,156	

Checking the Anderson Darling and p-value test, it is concluded that the logistic distribution fits adequately well our data.

In order to define subject distribution it is necessary to estimate its parameters.

The logistic distribution has no shape parameter. This means that the logistic *pdf* has only one shape, the bell shape, and this shape does not change. The shape of the logistic distribution is very similar to that of the normal distribution.

The location parameter is also the median and the mean and therefore can be easily calculated.

The scale parameter

Summarizing, the fitted distribution is:

Logistic distribution of times to failure

- CDF: $F(x) = \frac{1}{1+e^{-(x-\mu)/s}}, -\infty < x < \infty$
- PDF: $f(x) = \frac{e^{-(x-\mu)/s}}{s[1+e^{-(x-\mu)/s}]^2}, -\infty < x < \infty$

Where s=scale parameter= 2171

μ=location parameter=4773.35

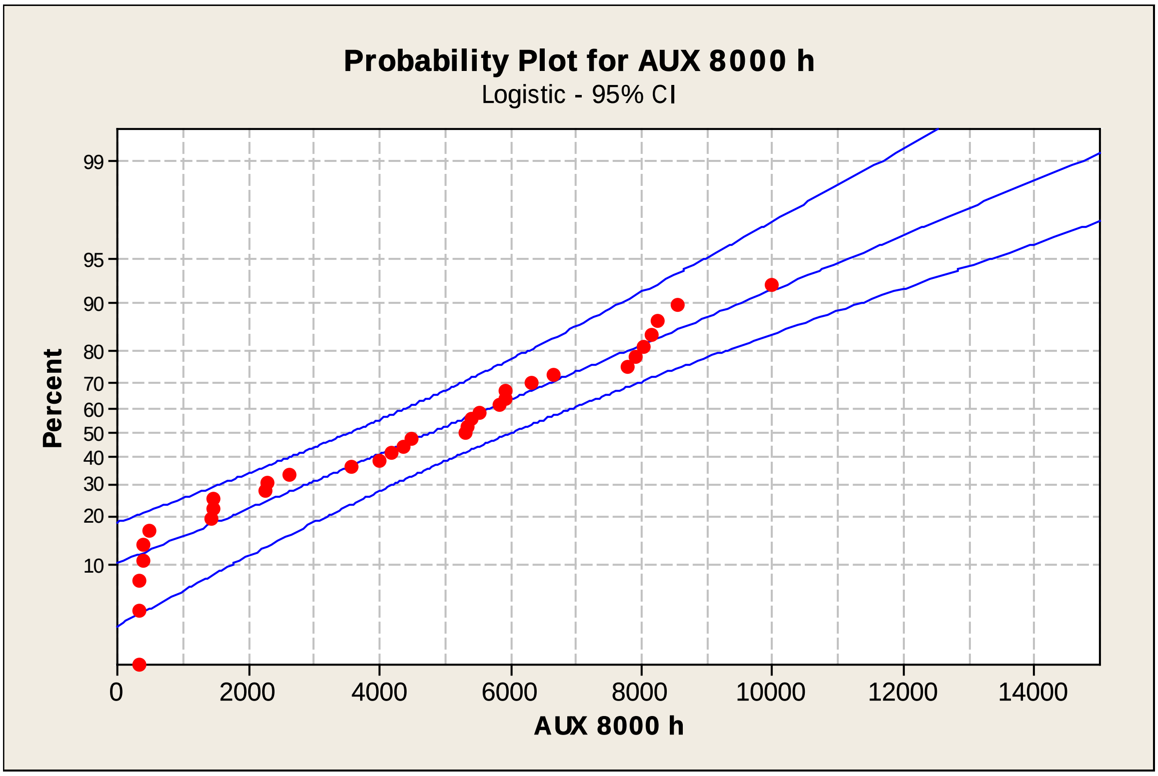


Figure 108: Plot of logistic fit to the data

In the plotted data in logistic plotting paper, it is understood that this distribution fits well the collected life data.

5.6.6 Reliability concepts

Relevant plots for this group of data can be produced, similarly to previous chapter.

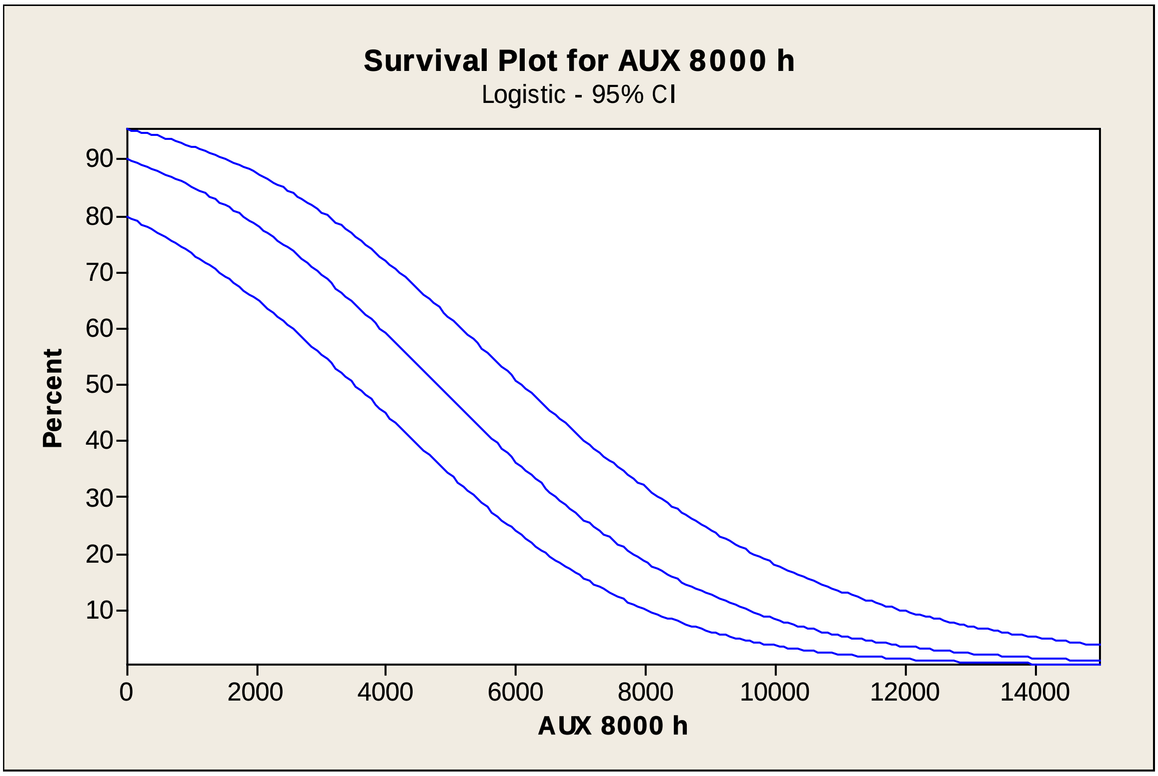


Figure 109: Survival plot for AUX 8000

Plot of reliability function with 95% confidence intervals

Table 58: table of survival probabilities			
		95% confidence intervals	
time	probability	lower	upper
1000	0,850441121	0,731631147	0,922242695
2000	0,782015203	0,649153093	0,874305896
3000	0,693563594	0,552182671	0,805991599
4000	0,588124252	0,445273374	0,717525467
5000	0,473923766	0,337067283	0,61481357
6000	0,362387691	0,238509148	0,507709099
7000	0,263931809	0,158405954	0,405855036
8000	0,184485525	0,099828383	0,315751974
9000	0,124895673	0,060495136	0,240317678
10000	0,082604046	0,035673851	0,179763678
12000	0,034599198	0,011819846	0,096971177
13000	0,022110825	0,006710207	0,070353928
14000	0,014064387	0,003789522	0,050778371
15000	0,008919451	0,002132455	0,036517032

the conditional probabilities for AUX 8000 are calculated and plotted according to the functions that can be found in Appendix A.

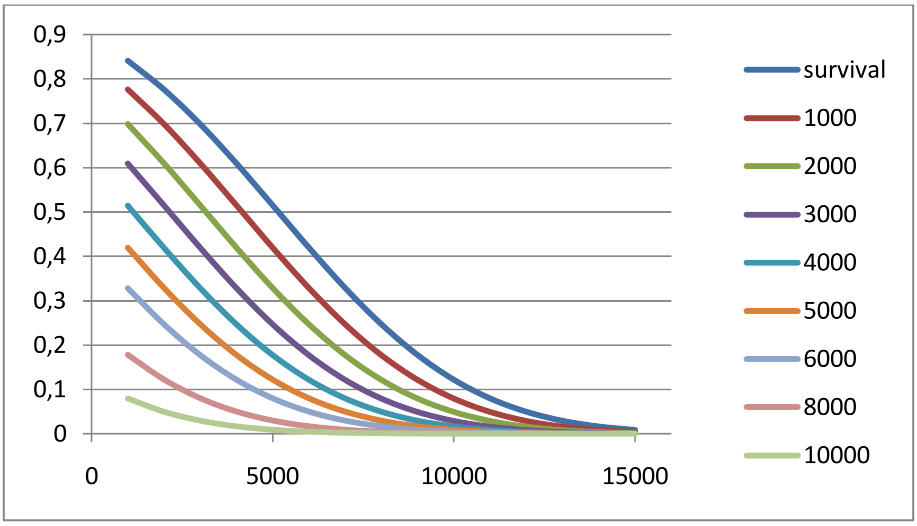


Figure 110: Conditional probabilities for AUX 8000

Given that an item has survived at 4000 hours, the probability to survive additionally time t

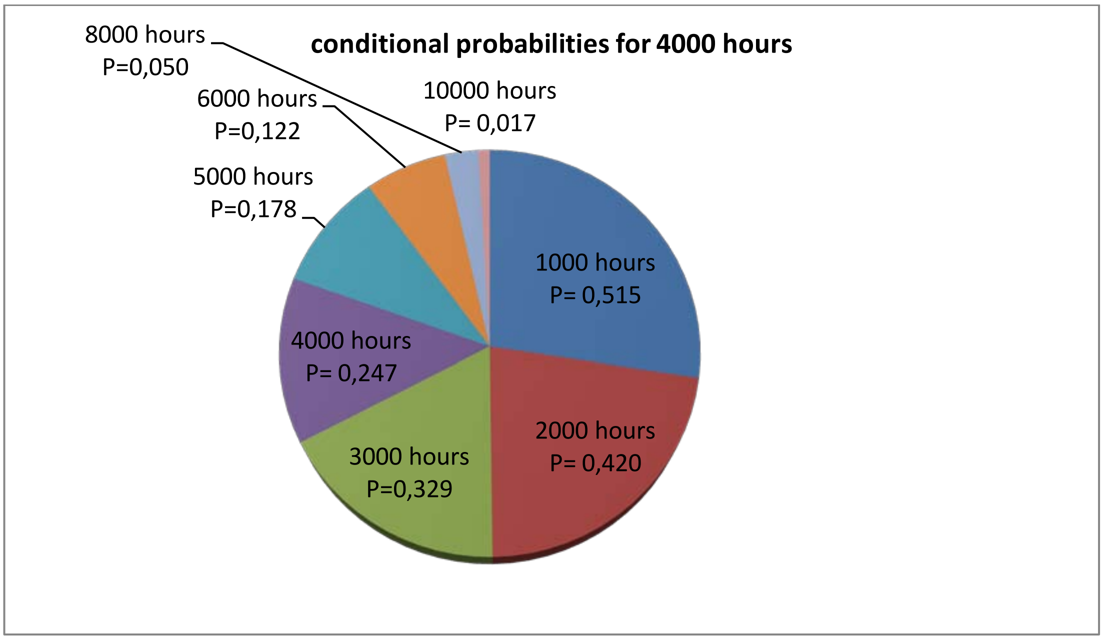


Figure 111: Conditional probabilities for AUX 8000 at 4000 hours

The tabulated values of this figure are:

Table 59: Conditional probabilities for AUX 8000 at 4000 hours

t	P
1000	0,51506
2000	0,419668

3000	0,328795
4000	0,247069
5000	0,17768
6000	0,122062
8000	0,049908
10000	0,016721

5.6.7 Failure rate plot

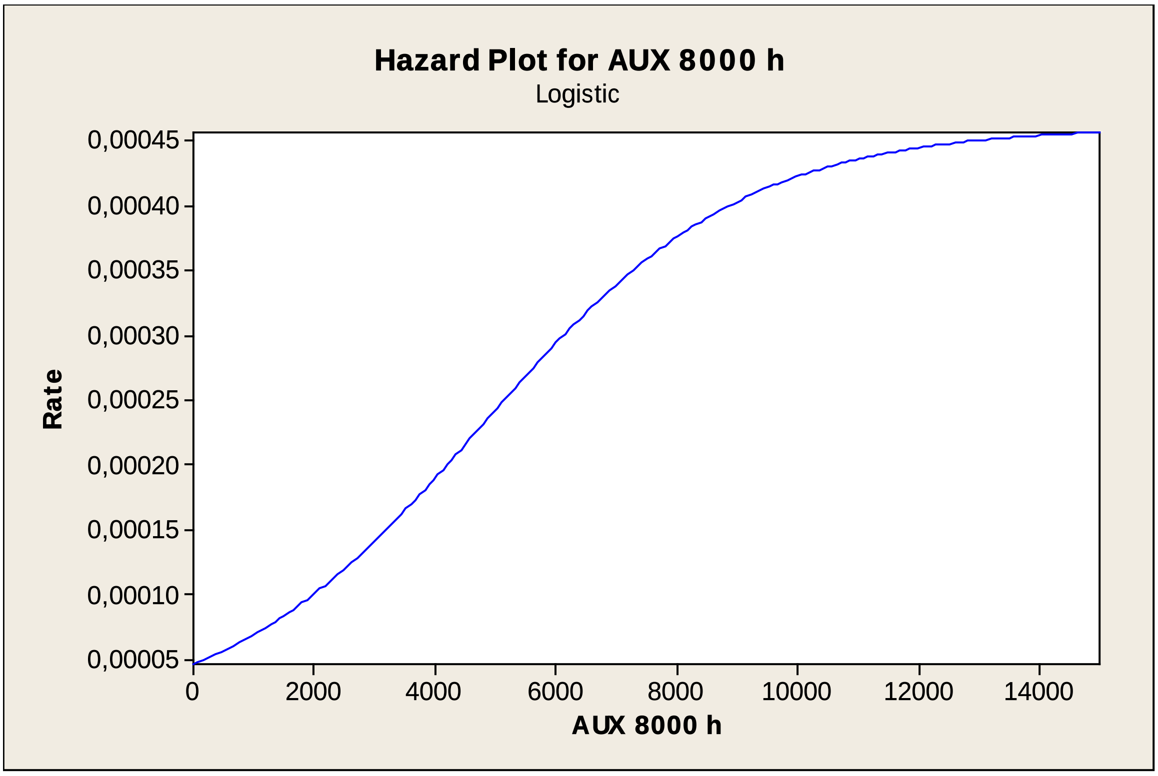


Figure 112: Failure rate plot for AUX 8000

5.6.8 Cumulative failure plot

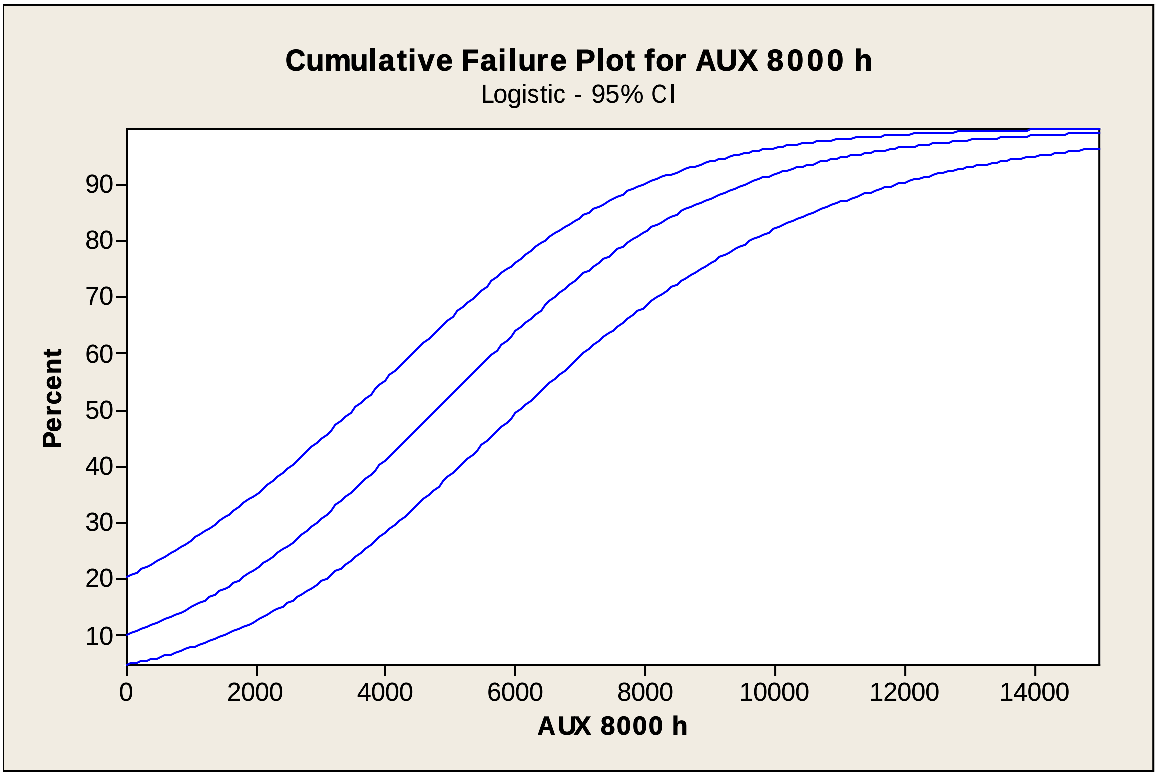


Figure 113: Cumulative failure plot for AUX 8000

Table 60: table of cumulative failure probabilities			
		95% confidence intervals	
time	probability	lower	upper
1000	0,149558879	0,077757305	0,268368853
2000	0,217984797	0,125694104	0,350846907
3000	0,306436406	0,194008401	0,447817329
4000	0,411875748	0,282474533	0,554726626
5000	0,526076234	0,38518643	0,662932717
6000	0,637612309	0,492290901	0,761490852
7000	0,736068191	0,594144964	0,841594046
8000	0,815514475	0,684248026	0,900171617
9000	0,875104327	0,759682322	0,939504864
10000	0,917395954	0,820236322	0,964326149
12000	0,965400802	0,903028823	0,988180154
13000	0,977889175	0,929646072	0,993289793
14000	0,985935613	0,949221629	0,996210478
15000	0,991080549	0,963482968	0,997867545

5.6.9 Non parametric analysis

Table 61: Non parametric estimates

```
             Standard    95,0% Normal CI
Mean(MTTF)      Error      Lower     Upper
   5157,57    713,147    3759,83   6555,31
```

A distribution free estimate of useful plots can be produced.

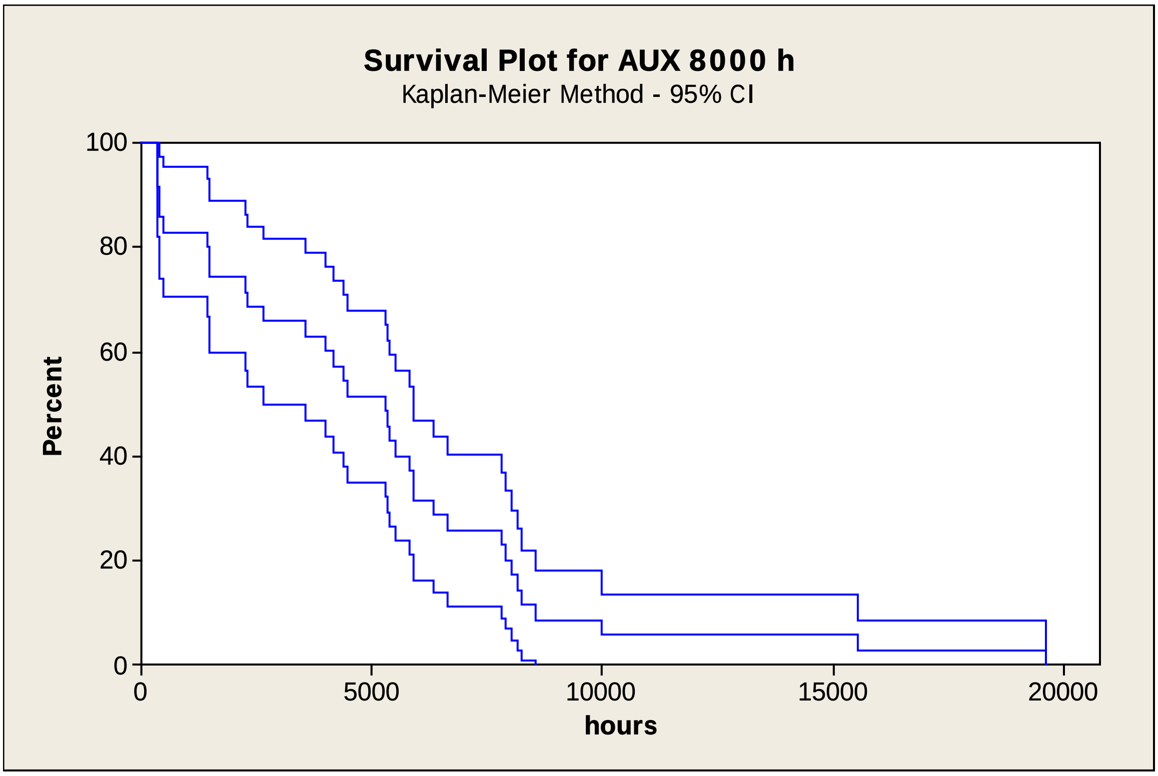


Figure 114: Survival plot for AUX 8000

Table 62: kaplan meier estimates for survival probabilities

		95% confidence intervals	
time	survival probability	lower	upper
345	0,942857143	0,865958621	1
350	0,914285714	0,821542607	1
400	0,857142857	0,741213973	0,973071741
475	0,828571429	0,7037122	0,953430658
1444	0,8	0,667482248	0,932517752
1461	0,771428571	0,632313911	0,910543232
1462	0,742857143	0,598062014	0,887652272
2250	0,714285714	0,564622169	0,86394926
2300	0,685714286	0,531917242	0,83951133

2630	0,657142857	0,499889268	0,814396446
3577	0,628571429	0,46849439	0,788648467
4000	0,6	0,437699563	0,762300437
4172	0,571428571	0,407480372	0,735376771
4367	0,542857143	0,377819574	0,707894711
4480	0,514285714	0,348706149	0,679865279
5312	0,485714286	0,320134721	0,651293851
5360	0,457142857	0,292105289	0,622180426
5400	0,428571429	0,264623229	0,592519628
5520	0,4	0,237699563	0,562300437
5835	0,371428571	0,211351533	0,53150561
5921	0,342857143	0,185603554	0,500110732
5922	0,314285714	0,16048867	0,468082758
6334	0,285714286	0,13605074	0,435377831
6650	0,257142857	0,112347728	0,401937986
7800	0,228571429	0,089456768	0,367686089
7900	0,2	0,067482248	0,332517752
8020	0,171428571	0,046569342	0,2962878
8148	0,142857143	0,026928259	0,258786027
8235	0,114285714	0,008881769	0,21968966
8542	0,085714286	0	0,178457393
9990	0,057142857	0	0,134041379
15544	0,028571429	0	0,083764617
19624	0	0	0

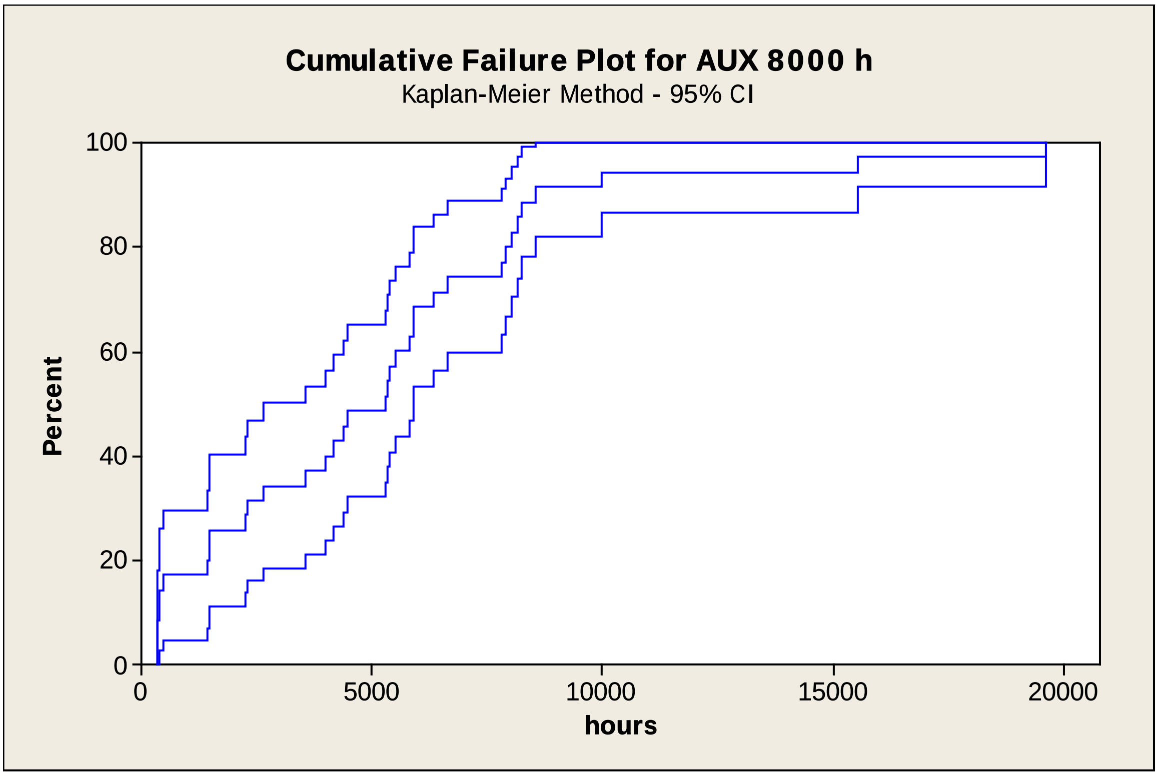


Figure 115: Cumulative failure plot

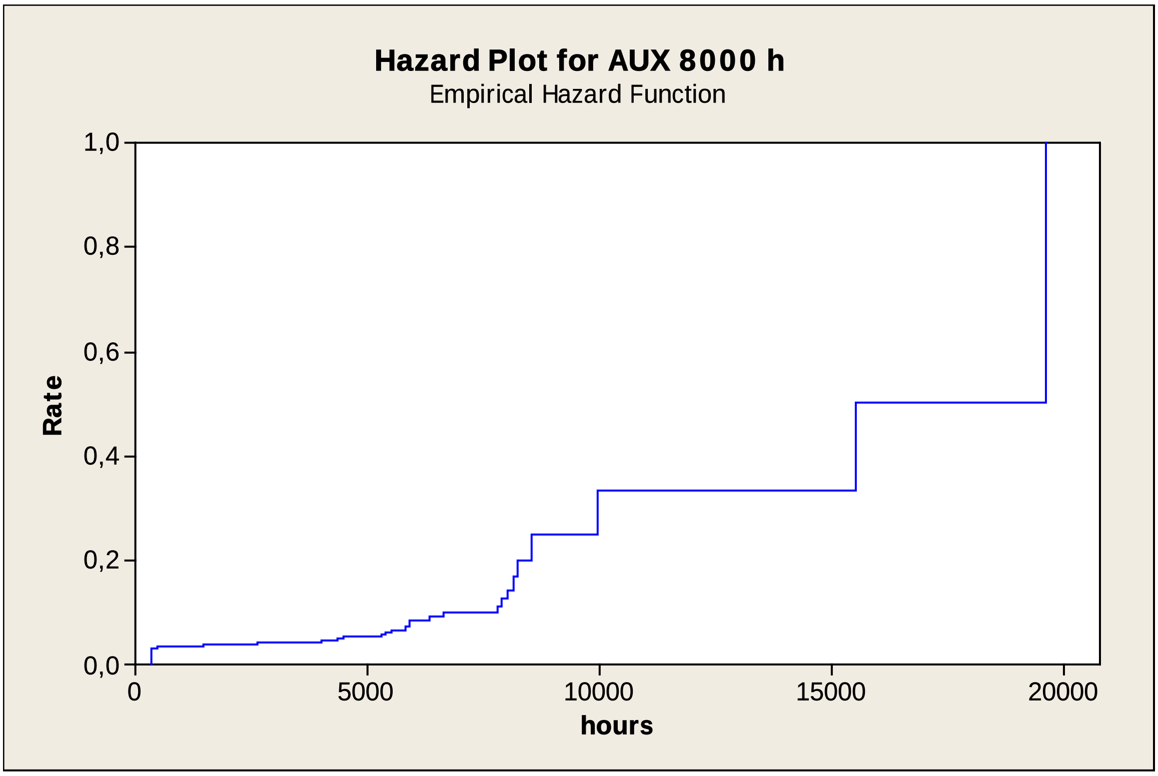


Figure 116: Hazard plot for AUX 8000

Table 63: cumulative failure probabilities and hazard rates

		95% confidence intervals		
time	cumulative failure probability	lower	upper	hazard rates
345	0,057142857	0	0,134041379	0,029411765
350	0,085714286	0	0,178457393	0,03030303
400	0,142857143	0,026928259	0,258786027	0,032258065
475	0,171428571	0,046569342	0,2962878	0,033333333
1444	0,2	0,067482248	0,332517752	0,034482759
1461	0,228571429	0,089456768	0,367686089	0,035714286
1462	0,257142857	0,112347728	0,401937986	0,037037037
2250	0,285714286	0,13605074	0,435377831	0,038461538
2300	0,314285714	0,16048867	0,468082758	0,04
2630	0,342857143	0,185603554	0,500110732	0,041666667
3577	0,371428571	0,211351533	0,53150561	0,043478261
4000	0,4	0,237699563	0,562300437	0,045454545
4172	0,428571429	0,264623229	0,592519628	0,047619048
4367	0,457142857	0,292105289	0,622180426	0,05
4480	0,485714286	0,320134721	0,651293851	0,052631579
5312	0,514285714	0,348706149	0,679865279	0,055555556
5360	0,542857143	0,377819574	0,707894711	0,058823529
5400	0,571428571	0,407480372	0,735376771	0,0625
5520	0,6	0,437699563	0,762300437	0,066666667
5835	0,628571429	0,46849439	0,788648467	0,071428571
5921	0,657142857	0,499889268	0,814396446	0,076923077
5922	0,685714286	0,531917242	0,83951133	0,083333333
6334	0,714285714	0,564622169	0,86394926	0,090909091
6650	0,742857143	0,598062014	0,887652272	0,1
7800	0,771428571	0,632313911	0,910543232	0,111111111
7900	0,8	0,667482248	0,932517752	0,125
8020	0,828571429	0,7037122	0,953430658	0,142857143
8148	0,857142857	0,741213973	0,973071741	0,166666667
8235	0,885714286	0,78031034	0,991118231	0,2
8542	0,914285714	0,821542607	1	0,25
9990	0,942857143	0,865958621	1	0,333333333
15544	0,971428571	0,916235383	1	0,5
19624	1	1	1	1

The mean remaining life function, versus the MTTF can be plotted.

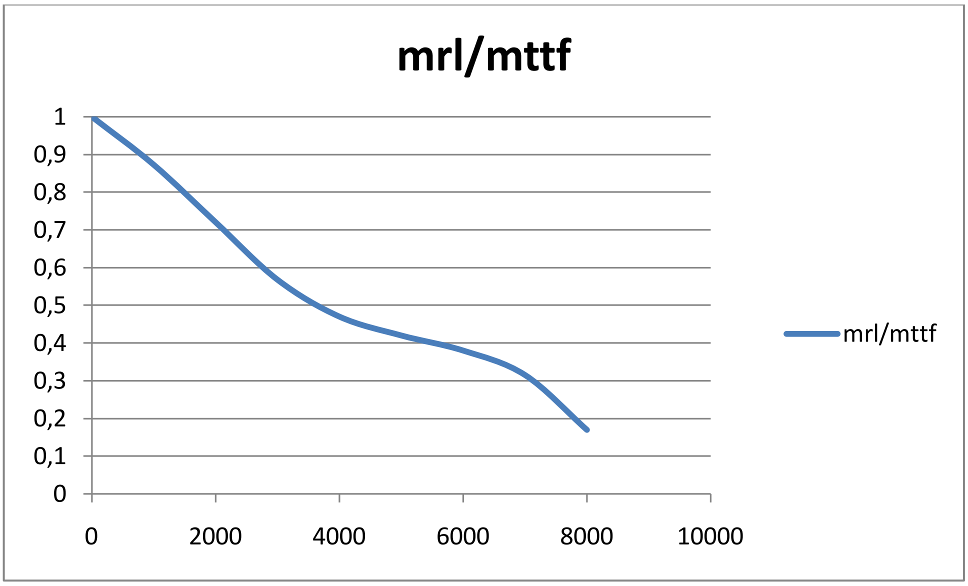


Figure 117: MRL/MTTF for AUX 8000

The transition window between 4000-6000 hours that where noticed in failure rate plot is clearly visible in this plot as well.

CHAPTER 6 JOINT DISTRIBUTION

In this thesis so far, we have calculated and produced useful plots and results for the reliability analysis of a marine vessel.

Both parametric and non parametric methods were used to produce more reliable results.

The next step in this thesis is to calculate the joint distribution of all 6 categories of machinery, in order to export the total distribution of failures and relevant plots and results for a vessel as a whole.

In cases such as this, it can be useful to characterize the system's behavior by fitting a distribution to the overall system and calculating parameters for this distribution. (Note: This is particularly useful in system simulation, since it significantly reduces simulation time.) This is equivalent to fitting a single distribution to describe $R_S(t)$. In essence, it is like reducing the entire system to a component in order to simplify calculations.

For this reason, a qualitative analysis must be performed.

The main objective of system reliability is the construction of a model (life distribution) that represents the times-to-failure of the entire system based on the life distributions of the system's elements. These elements can be components' assemblies, sub-systems etc.

6.1 APPROACHES OF SYSTEM RELIABILITY

In theory and in praxis there are two basic approaches (categories of approaches):

- Analytical calculations
    1. Static analytical calculations
    2. Time-dependent calculations
- Simulation calculations

Two types of analytical calculations can be performed using RBD (Reliability Block Diagrams) or FTA (Fault Tree Analysis): static reliability calculations and time-dependent reliability calculations. Systems can contain static blocks, time-dependent blocks or a mixture of the two (27).

Static analytical calculations are performed on RBD or failure trees that contain static blocks. A static block can be interpreted either as a block with a reliability value that is known only at a given time (but the block's entire distribution is unknown) or as a block with a probability of success that is constant with time. Static calculations can only be performed in the analytical mode and not in the simulation calculations.

Time-dependent analysis approaches reliability as a function of time. That means that a known failure distribution is assigned to each component. The time scale can be any quantifiable time measure, such as years, months, hours, minutes or seconds, and also units that are not directly related to time.

If one includes information on the repair and maintenance characteristics of the components and resources available in the system, other information can also be analyzed/obtained, such as i.e. system availability, maintability etc. This can be accomplished through discrete event simulation.

In simulation, random failure times from each component's failure distribution are generated. These failure times are then combined in accordance with the way the components are reliability-wise arranged within the system. The overall results are analyzed in order to determine the behavior of the entire system.

6.2 FAULT TREE ANALYSIS, RELIABILITY BLOCK DIAGRAMS

Block diagrams are widely used in engineering in many different forms. Fault trees and reliability block diagrams are both symbolic analytical logic techniques that can be applied to analyze system reliability and related characteristics. They can also be used to describe the interrelation between the components and to define the system.

When blocks are connected with direction lines, that represent the reliability relationship between these blocks, this is referred as reliability block diagram (RBD) (28).

A fault tree diagram follows a top-down structure and represents a graphical model of the pathways within a system that can lead to a foreseeable, undesirable loss event (or a failure). The pathways interconnect contributory events and conditions using standard logic symbols (AND, OR, etc.). Fault tree diagrams consist of gates and events connected with lines.

The most fundamental difference between fault tree diagrams and reliability block diagrams is that in an RBD, you work in the "success space", while in a fault tree you work in the "failure space". In other words, the RBD searches for success combinations while the fault tree searches for failure combinations. In addition, fault trees have traditionally been used to analyze fixed probabilities (*i.e.* each event that comprises the tree has a fixed probability of occurring) while RBDs may include time-varying distributions for the success (reliability equation) and other properties, such as repair/restoration distributions. In general (and with some specific exceptions), a fault tree can be easily converted to an RBD. However, it is generally more difficult to convert an RBD into a fault tree, especially if one allows for highly complex configurations.

6.3 BLOCKSIM MODEL OF RELIABILITY BLOCK DIAGRAM

For the calculation of the total joint distribution, the software package Reliasoft BlockSim was used.

This is a specialized software tool for system reliability, availability and related analyses.

It also supports an array of reliability block diagram configuration and fault tree analysis capabilities, which turns to be a very useful tool for the needs of this thesis.

Briefly this tool provides：

- Identification of Critical Components
- Reliability Optimization
- System Maintainability Analysis (Determine Optimum Preventive Maintenance Intervals, etc.)
- System Availability Analysis (Calculate Uptime, Downtime, Availability, etc.)
- Throughput Calculation (Identify Bottlenecks, Estimate Production Capacity, etc.)
- Resource Allocation for Maintenance Planning
- Life Cycle Cost Analysis

6.3.1 BUILDING THE JOINT DISTRIBUTION

We have already calculated in Chapter 5 the distributions that fit adequately well each group of components.

Our goal is to build a fault tree that will allow calculating the joint distribution of failures.

Briefly, fault tree analysis is a deductive technique where we start with a specified system failure. The system failure is called the TOP event of the fault tree. The immediate casual events that either alone or in combination may lead to the TOP event are identified and connected to the TOP event through a logic gate. This procedure is continued deductively until we reach a suitable level of detail. The events in the lowest level are called the basic events of the fault tree (29).

Fault tree analysis is a binary analysis. All events are assumed to occur or not to occur.

Steps in a fault tree analysis (30)

The following outline describes the steps to be taken in a fault tree analysis.

Step 1: Define the TOP event.

In our fault tree, the TOP event is considered to be any kind of failure, i.e. failure in Main Engine, Diesel Generators or Auxiliaries. In this thesis the performed functions and failures are not classified according to their severity or importance. Any kind of failure is considered as the actualization of the TOP event.

Step 2: Construction of the fault tree

The fault tree construction always starts with the TOP event. We must thereafter try to identify all fault events that are the causes that result in the TOP event.
In the fault tree that describes a failure in a vessel as a whole, the fault events are the failures distributions in all 6 subcategories that are already defined.
As the object of this research is the probability of any failure, the TOP event is connected with the basic events with an OR gate. In this case, any failure that occurs in each subcategory, leads to a TOP event failure.
Therefore the categories M/E 4000, M/E 8000, D/G, AUX 4000, AUX 6000, AUX 8000 are connected under OR gates.

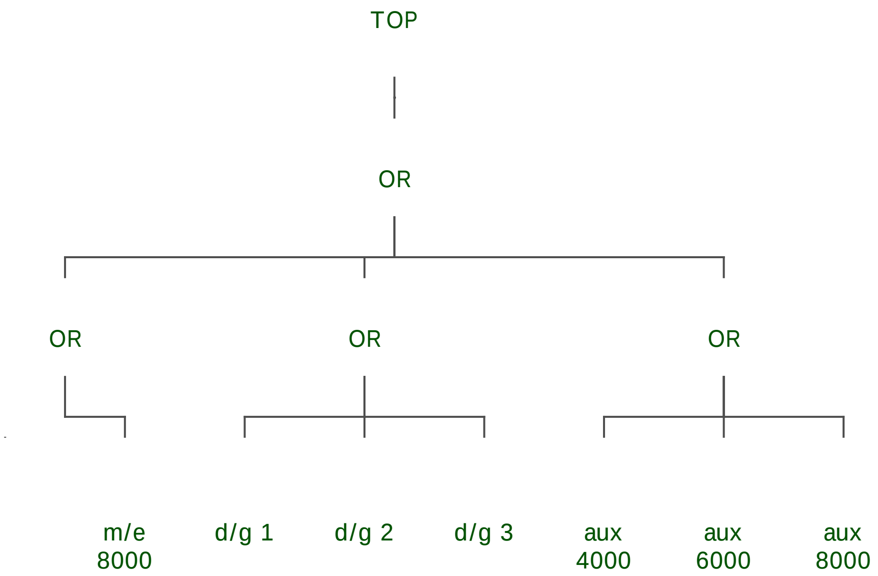
Step 3: Qualitative analysis- Minimal cut sets

A combination of fault events that will result in the TOP event is called a cut set. A cut set is said to be minimal, if the set cannot be reduced without losing its status as a cut set.
In our fault tree we can trace possible cut sets that lead to a vessel's failure.

Step 4: Quantitative analysis

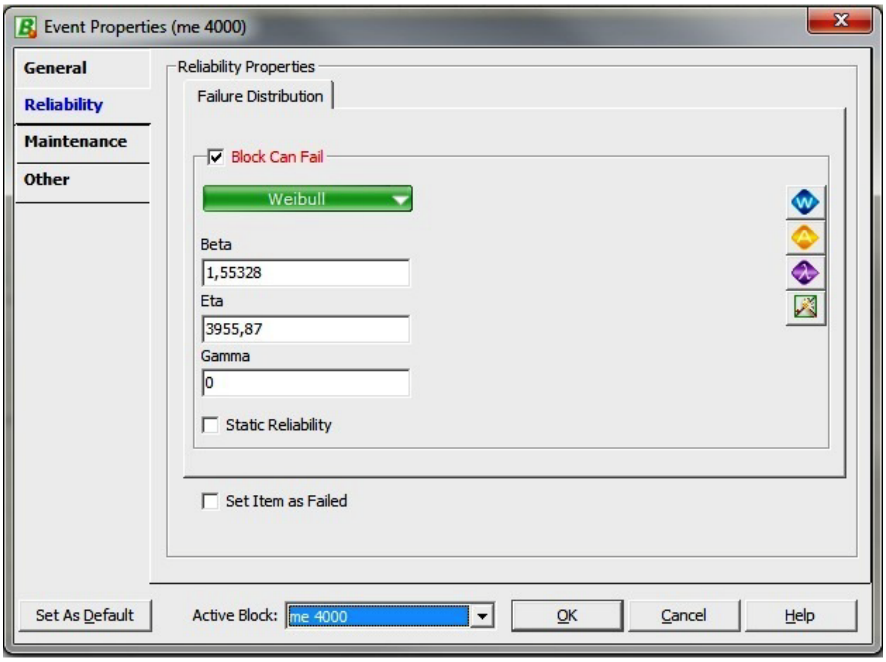
Quantitatively, we can calculate the probability of the top event, rank basic events by importance, extract reliability plots, such as failure rate, availability plots, etc.

Following the steps described above, the following fault tree can be produced.



It is reminded that as top event is considered any failure that occurs, whether it comes from the Main Engine, Diesel Generators or the auxiliaries.

The properties of the basic events have been defined according to the results of chapter 5. The Blocksim software was used for the significant advantages over similar fault tree analysis packages. In the properties of each event, the failure distribution can be defined accurately.



For each event the failure distribution is used as input, therefore there is a greater accuracy in the probabilities calculation.

Also, the inspection interval can be defined; therefore each category has a priori a known inspection interval according to the manufacturer's suggestion.

Regarding the group of auxiliary equipment, there is an important detail that must be taken into account.

During normal operation of a marine vessel, the Main Engine is, obviously, operating, so that the vessel moves. The electrical loads of a vessel are covered from the diesel generators. Usually a ship has 3 or 4 Diesel Generators and an emergency generator, which is capable to cover only the electric demands of safety equipment, such as radars, steering gear, etc.

As a matter of good engineering practice, in order to achieve maximum efficiency of Diesel Generators, usually the operation of one Generator is capable to cover the normal electrical load of a vessel, and the second Generator to be in stand-by mode.

For the operation of the auxiliary equipment, which we focus on, in this thesis, it is necessary that at least one Diesel Generator operates, so that to provide the necessary electrical energy.

Therefore for the modeling of this modification in a fault tree, the Generators should operate for the auxiliaries to be able to operate.

Based on this solution, the three Generators should be modeled under an AND gate (parallel operation), in series with the auxiliaries.

Nevertheless, the above modification with three Generators in parallel operation gave the following results:

Table 64: Failure probabilities for D/Gs

t	probability of failure for diesel generators in parallel operation
1000	0,0091
2000	0,0629
3000	0,1659
4000	0,2966
5000	0,4319
6000	0,5563
7000	0,6623
8000	0,7479
9000	0,8147
10000	0,8654
11000	0,9032
12000	0,9309
13000	0,951
14000	0,9654
15000	0,9757

Table 65: Statistical parameters

MTTF	Mean availability	Expected Number of Failures
8900 hours	0,8938	1,69

The above tables, which have been produced after simulating the fault tree with 3 Generators in parallel operation for 15.000 operating hours, indicate that the Mean Time To Failure is 8900 operating hours.

It is reminded that for the parts of Diesel Generators that we examine, the manufacturer suggests 4000 operating hours between inspections. Therefore, it is obvious that before an expected failure occurs, two inspections will have already taken place.

It is concluded that the auxiliary equipment can be modeled in the fault tree under an OR gate, in series with the Main Engine and Diesel Generator.

It was mentioned that a simulation for 15.000 hours was performed in the fault tree.

To illustrate the simulation process, assume a single block with a failure and a repair distribution. The first event, E_{F_1}, would be the failure of the component. Its first time-to-failure would be a random number drawn from its failure distribution, T_{F_1}. Thus, the first failure event, E_{F_1}, would be at T_{F_1}. Once failed, the next event would be the repair of the component, E_{R_1}. The time to repair the component would now be drawn from its repair distribution, T_{R_1}. The component would be restored by time $T_{F_1} + T_{R_1}$. The next event would now be the second failure of the component after the repair, E_{F_2}. This event would occur after a component operating time of T_{F_2} after the item is restored (again drawn from the failure distribution), or at $T_{F_1} + T_{R_1} + T_{F_2}$. This process is repeated until the end time. It is important to note that each run will yield a different sequence of events due to the probabilistic nature of the times. To arrive at the desired result, this process is repeated many times and the results from each run (simulation) are recorded. In other words, if we were to repeat this 1,000 times, we would obtain 1,000 different values for E_{F_1}, or $\left[E_{F_{1_1}}, E_{F_{1_2}}, \ldots, E_{F_{1_{1,000}}}\right]$. The average of these values, $\left(\frac{1}{1000}\sum_{i=1}^{1,000} E_{F_{1_i}}\right)$, would then be the average time to the first event, E_{F_1}, or the mean time to first failure (MTTFF) for the component. Obviously, if the component were to be 100% renewed after each repair, then this value would also be the same for the second failure, etc (31).

6.3.2 BLOCKSIM RESULTS

Following the described procedure, after simulating the fault tree for 15.000 operating hours, the following results were obtained.

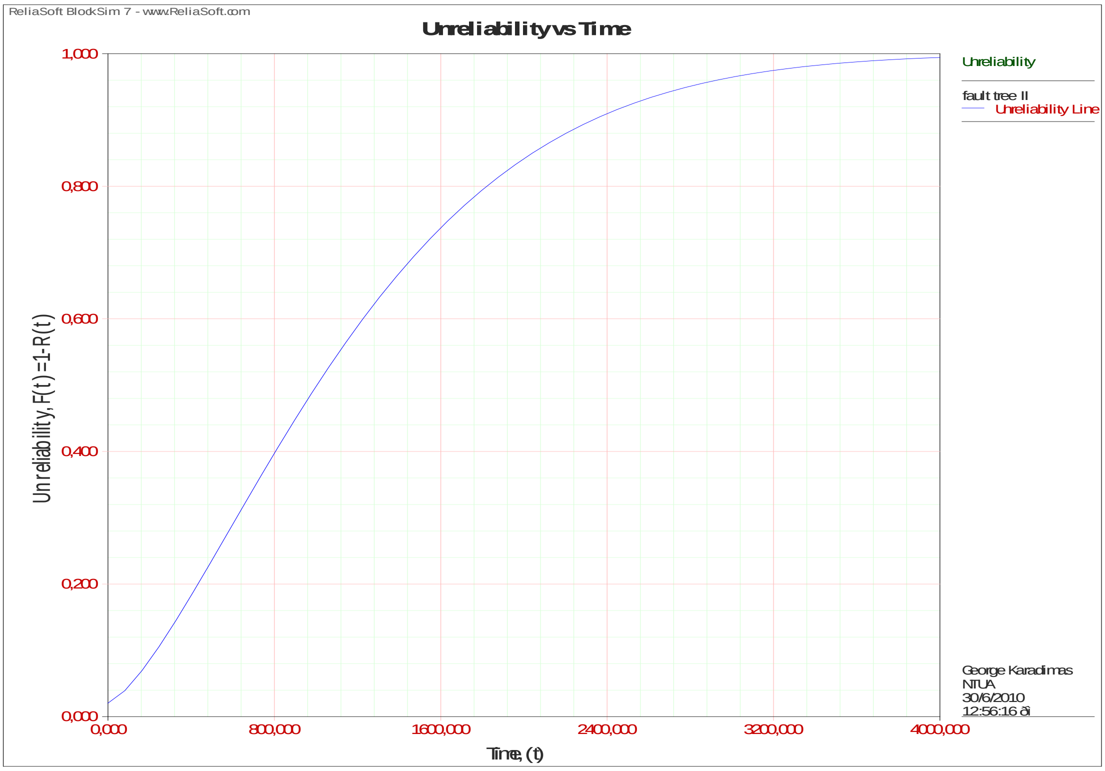


Figure 118: Unreliability function of joint distribution

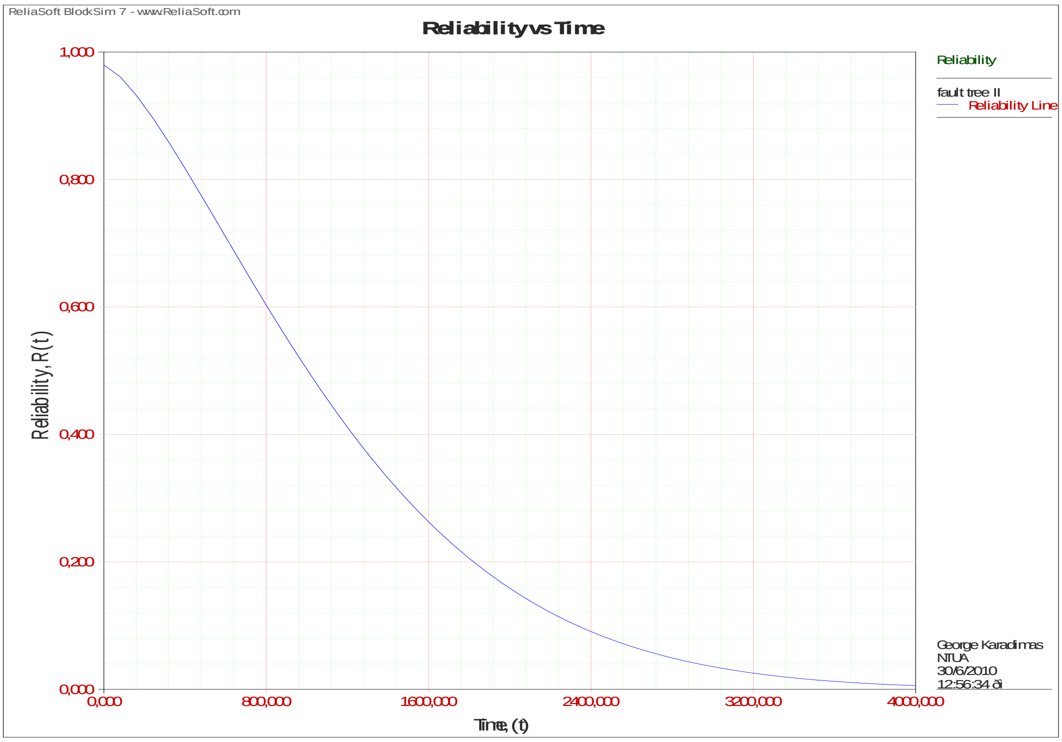


Figure 119: Reliability function of joint distribution

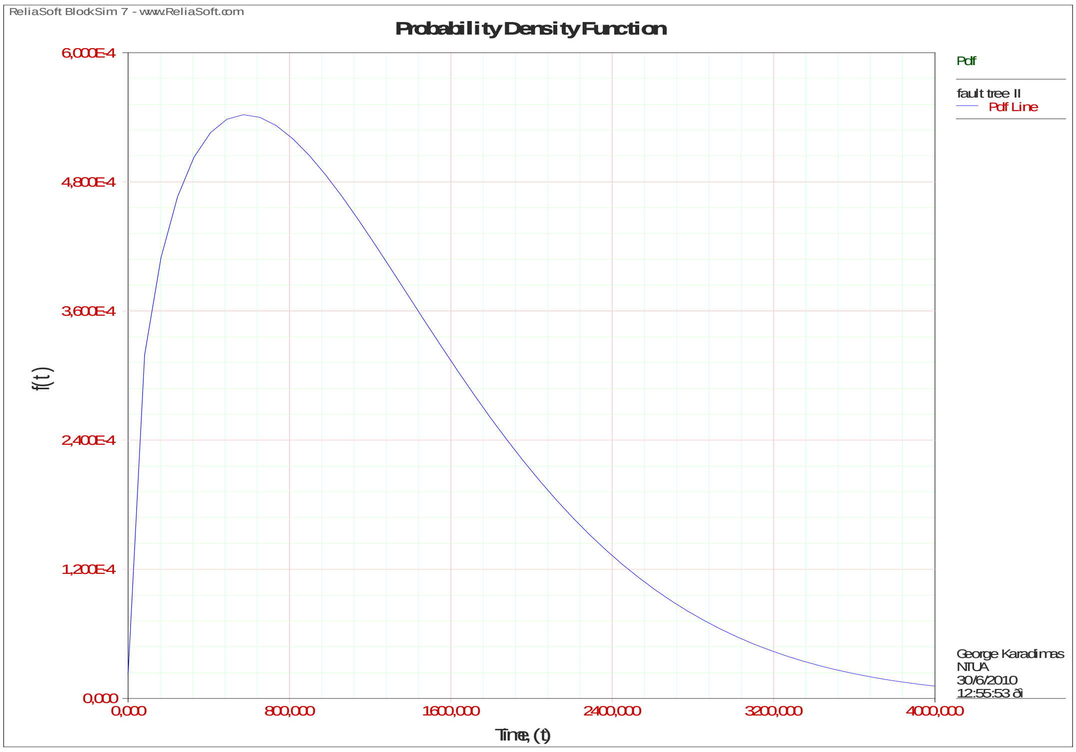


Figure 120: PDF of the joint distribution

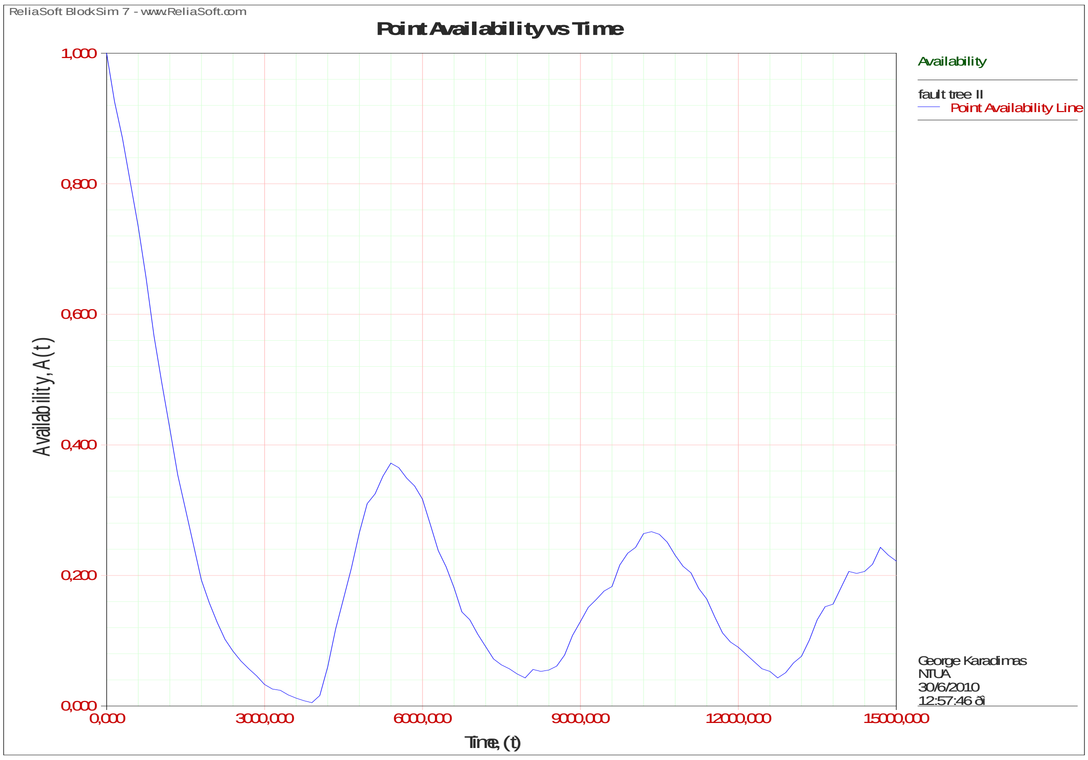


Figure 121: Point availability function

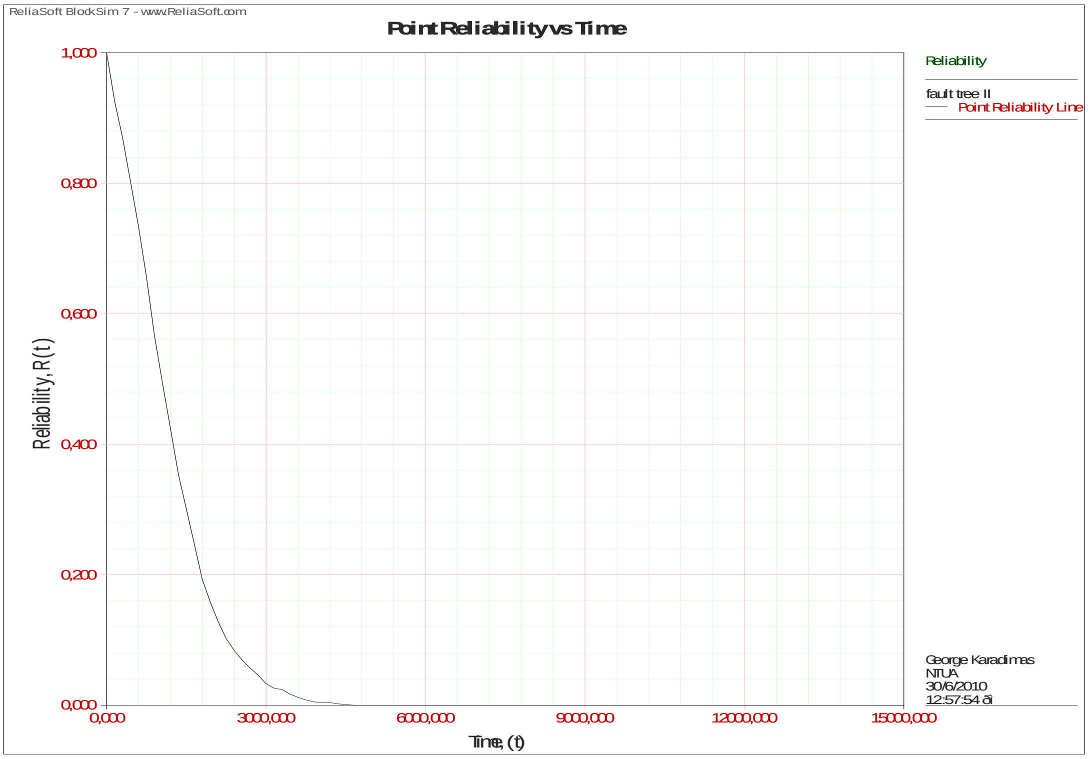


Figure 122: Point reliability function

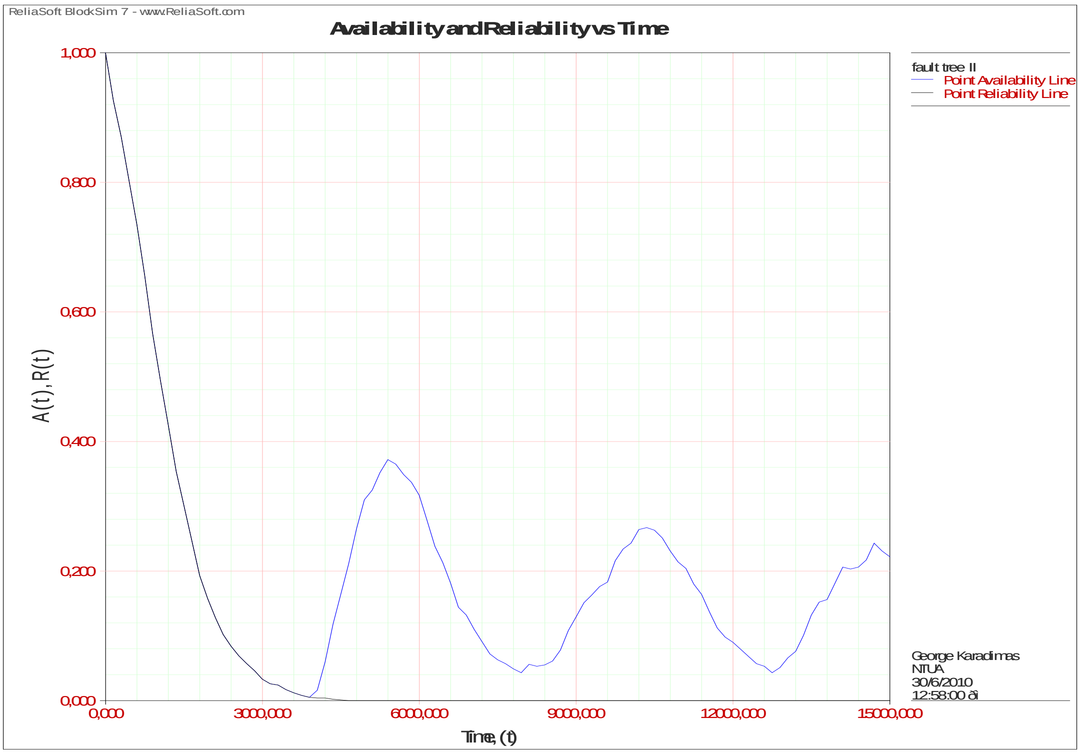


Figure 123: Point Availability/ Reliability

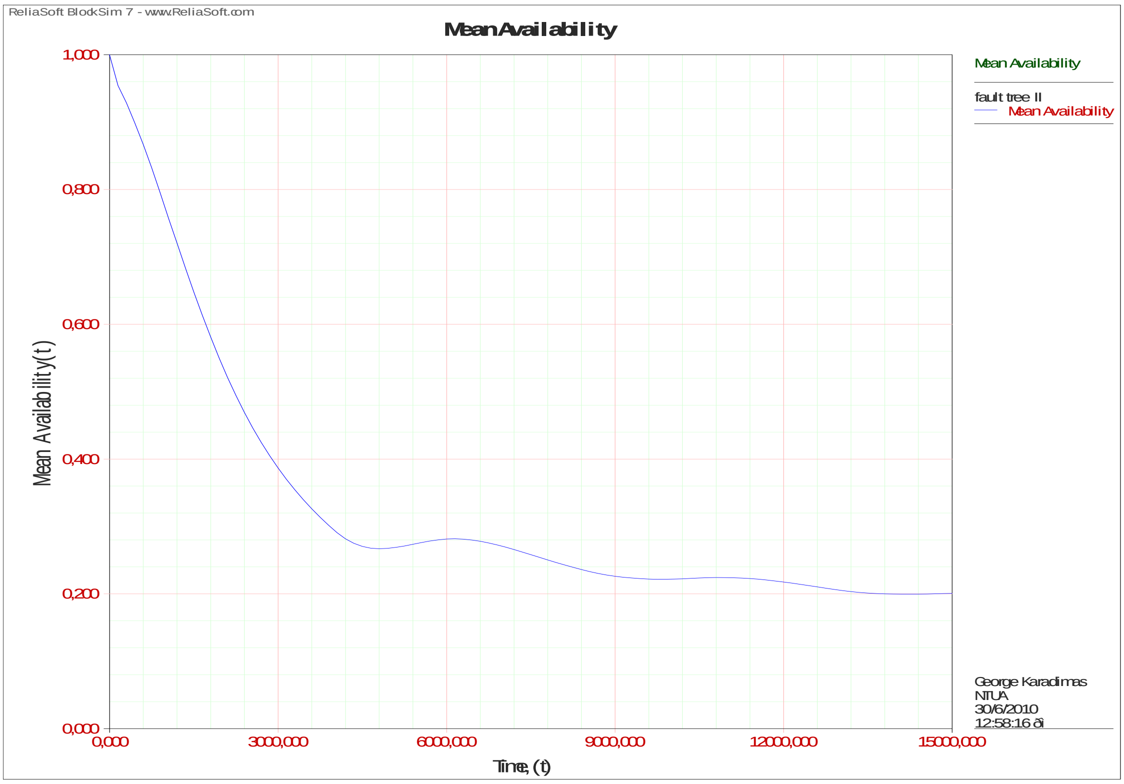


Figure 124: Mean availability function

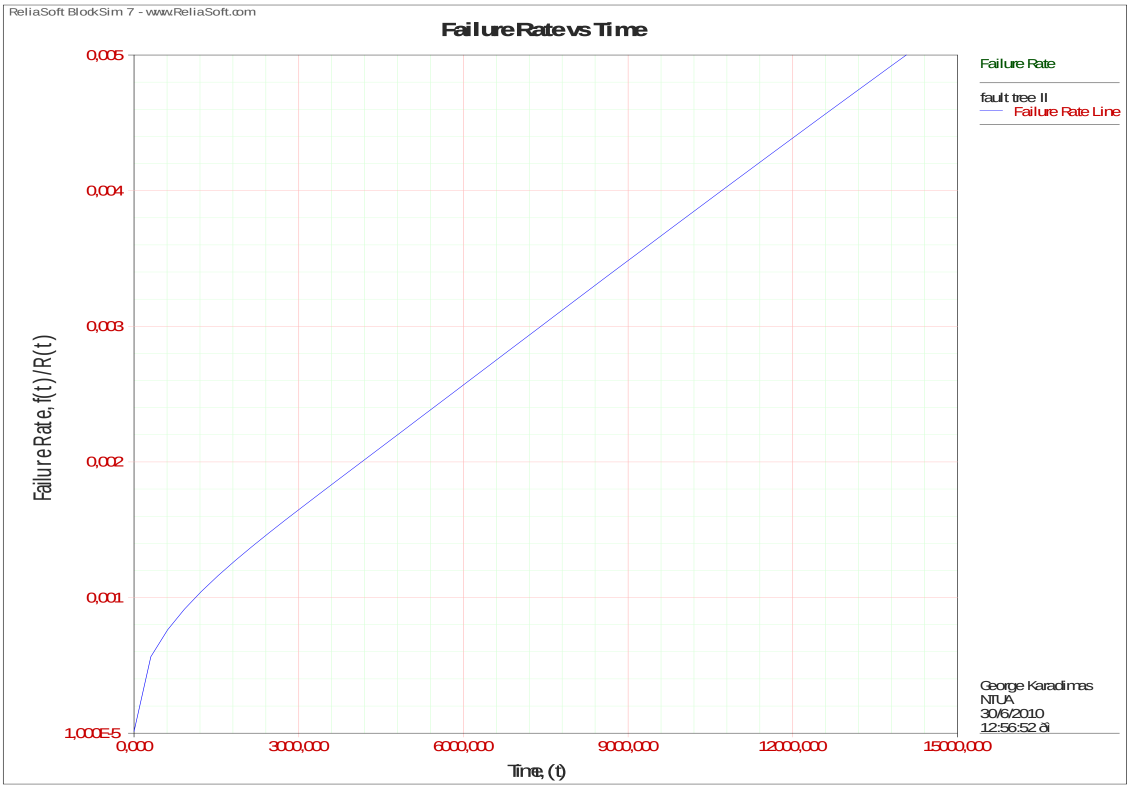


Figure 125: Failure rate of the joint distribution

6.4 DEFINITIONS

The plotted parameters are defined as following.

6.4.1 Mean Availability (All Events)

This is the mean availability due to all downing events, which can be thought of as the operational availability. It is the ratio of the system uptime divided by the total simulation time (total time).

6.4.2 Point Availability (All Events), $A(t)$

This is the probability that the system is up at time *t*. As an example, to obtain this value at *t* = 300, then a special counter would need to be utilized during the simulation. This counter is incremented by one every time the system is up at 300 hours. Thus, the point availability at 300 would be the times the system was up at 300 divided by the number of simulations.

6.4.3 Point Reliability (Fail Events), $R(t)$

This is the probability that the system has not failed by time *t*. This is similar to point availability with the major exception that it only looks at the probability that the system did not have a single failure. Other (non-failure) downing events are ignored. During the simulation, a special counter again must be utilized. This counter is incremented by one (once in each simulation) if the system has had at least one failure up to 300 hours. Thus, the point reliability at 300 would be the number of times the system did not fail up to 300 divided by the number of simulations.

6.4.4 Expected Number of Failures, N_F

This is the average number of system failures. The system failures (not downing events) for all simulations are counted and then averaged.

6.4.5 Failure rate

A function that describes the number of failures that can be expected to take place over a given unit of time. The failure rate function has the units of failures per unit time among surviving units, i.e. one failure per month.

In the availability figure, at 4000, 6000, 8000 hours an inspection/ corrective maintenance is carried out, if necessary. For this reason the availability at subject hours is increasing.

The figures for reliability and availability, tabulated, are shown in Table 66: Reliabity/ Availability for joint distribution

Table 66: Reliabity/ Availability for joint distribution

T	R(t)	A(t)
1000	0,5183	0,5183
2000	0,148	0,148
3000	0,033	0,033
4000	0,0043	0,0123
5000	5,40E-06	0,315
6000	8,50E-07	0,317
7000	0	0,1173
8000	0	0,0473
9000	0	0,129
10000	0	0,24

For the total joint distribution of failures, it is very useful to calculate the conditional probabilities of failures.

Given that a vessel has run a failure-free time of T1 hours, what is the probability that a failure occurs in the next T2 hours?

Using the conditional probabilities theory (Appendix A) the following table is produced.

Table 67: conditional probabilities calculations

t	probability of failure
500/500	0,6574
1000/1000	0,6844
2000/1000	0,7758
2000/2000	0,9632
3000/1000	0,836
3000/2000	0,9802

3000/3000	0,9982
4000/1000	0,8792
4000/2000	0,9892
4000/3000	0,9993
4000/4000	0,99995

For example, if a vessel had operated 2000 hours without a failure, then the probability that a failure occurs in the next 1000 hours is 0.7758 or 77.6%.

6.5 EXPANDED ALGEBRAIC SOLUTION

Blocksim provides the complete algebraic solution for the calculation of the joint equations.

In the case that the basic events are modified in series and the component failure characteristics can be described by distributions, the system reliability is actually time-dependent (system reliability theory page 119). In this case, the system reliability can be written:

$$R_S(t) = R_1(t) \cdot R_2(t) \cdot R_3(t)$$

The reliability of the system for any mission time can now be estimated.

In the under examination fault tree, system reliability is:

$R_{System} = + R_{me\ 4000} \cdot R_{m/e\ 8000} \cdot R_{d/g\ 1} \cdot R_{d/g\ 2} \cdot R_{d/g\ 3} \cdot R_{aux\ 4000} \cdot R_{aux\ 8000} \cdot R_{aux\ 6000}$

Block Failure Distribution Legend

m/e 8000: Normal μ=4474,56; σ=2182,48
aux 4000: Loglogistic μ=2103,9; σ=2155,06
aux 6000: Logistic μ=2681,4; σ=1715,5
aux 8000: Logistic μ=5157,5; σ=4198,6
d/g 3: Weibull β=1,40425; η=3517,47; γ=0
d/g 2: Weibull β=1,40425; η=3517,47; γ=0
d/g 1: Weibull β=1,40425; η=3517,47; γ=0

me 4000: Weibull β=1,55328; η=3955,87; γ=0

Using the same procedure, the probability function and the failure rate equation can be calculated.

Pdf function:

$$f_{System} = + f_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,8000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,8000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{d/g}\,1} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,8000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{d/g}\,2} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,8000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{d/g}\,3} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,8000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{aux}\,4000} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,8000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{aux}\,8000} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,6000} + f_{\mathrm{aux}\,6000} \cdot R_{\mathrm{me}\,4000} \cdot R_{\mathrm{m/e}\,8000} \cdot R_{\mathrm{d/g}\,1} \cdot R_{\mathrm{d/g}\,2} \cdot R_{\mathrm{d/g}\,3} \cdot R_{\mathrm{aux}\,4000} \cdot R_{\mathrm{aux}\,8000}$$

Failure rate function:

$$fr_{System} = + fr_{\mathrm{me}\,4000} + fr_{\mathrm{m/e}\,8000} + fr_{\mathrm{d/g}\,1} + fr_{\mathrm{d/g}\,2} + fr_{\mathrm{d/g}\,3} + fr_{\mathrm{aux}\,4000} + fr_{\mathrm{aux}\,8000} + fr_{\mathrm{aux}\,6000}$$

where f_i , R_i and fr_i are, correspondingly, the probability density function, reliability function and failure rate function of each subgroup that have been calculated in chapter 5.

6.6 ADVANTAGES OF THE ANALYTICAL METHOD

The primary advantage of the analytical solution is that it produces a mathematical expression that describes the reliability of the system. Once the system's reliability function has been determined, other calculations can then be performed to obtain metrics of interest for the system. Such calculations include:

- Determination of the system's *pdf*.
- Determination of warranty periods.
- Determination of the system's failure rate.
- Determination of the system's MTTF.

Some useful parameters for the total distribution can be found in table 68.

Table 68 : Basic parameters for joint distribution

MTTF	Mean availability
1178	0,2007

Summarizing, the MTTF for a vessel as a whole is 1178 hours. This mean that in average every 1178 hours a failure will occur.

The theoretical calculated MTTF for the joint distribution of failures, is lower even than the MTTF of the least reliable subcategory of equipments. This can be explained due to the probabilistic nature of the calculations/ simulation in the fault tree analysis.

CHAPTER 7 REPLACEMENT AND INSPECTION COST

So far we have handled with probabilities of failure, reliability and relevant plots. In order that useful qualitative calculations are made, it is necessary for us to work on the cost that these failures lead to.

Nowadays, the effort to reduce operating cost of a vessel is very important. The proper preventive maintenance policy can help to avoid unexpected failures and therefore unexpected cost.

To calculate the consequences of failures in a vessel, this effort commenced with the collection of data for the cost of each failure that appears in the collected life data. These costs where categorized according to each subgroup of equipment that constitutes a vessel.

Next, through a research in a shipping company, there were collected the cost for spare parts, the required time to complete each repair and the required number of manpower to carry out subject repair.

The below table indicates a sample of repairs that regard the Main Engine.

Table 69: Indicative repairs and cost for M/E

PART	COST/UNIT ($)	REQUIRED TIME FOR REPAIRS (hours)	TOTAL NUMBER OF MANPOWER
T/C BLOWER SIDE BEARING	9300,00	5	2
AIR STARTING VALVE	4000,00	1	1
T/C PEDESTAL ROTOR	14000,00	12	2
CYLINDER COVER	26000,00	10	3
STUFFING BOX	25000,00	4	2
CYLINDER LINER	19500,00	14	3
PISTON CROWN	14300,00	14	3
EXHAUST VALVE	16000,00	3	2
PUNCTURE VALVE	3000,00	2	1
PISTON RINGS	640,00	12	3

It is noted that for the replacement of broken equipment it is assumed that a failure had occurred and the replacement of subject equipment is necessary to bring the system back to operating condition. Additionally, in a case of a simple inspection, for example as required from the manufacturer's planned maintenance system, the cost that the repair includes is the labor cost for a repair/ inspection.

For each one of the 6 subgroups of a vessel, the average cost for replacement of the group's equipment, along with the inspection cost were calculated. It is noted that the labor cost was calculated based on the realistic assumption of 50$/hour for each employer.

The following table is clearly indicative.

Table 70: average cost for repairs/ inspection (all values in USD)

category	M/E 4000	M/E 8000	D/G	AUX 4000	AUX 6000	AUX 8000
average replacement cost	7280	17136	3060	3950	3950	3950
average inspection cost	733	1121	400	150	150	150

In chapter 5 the following values for the expected MTTF for each category were calculated.

Table 71: MTTF for each category

category	M/E 4000	M/E 8000	D/G	AUX 4000	AUX 6000	AUX 8000
MTTF	3548	4474	3232	1936	2668	5157

7.1 COST CALCULATION THROUGH THE STATISTICAL RESULTS

For a period of 1 year, the overall operational days for a typical containership are about 280 to 320, the expected cost for replacement and maintenance, along with 95% confidence intervals for these values can be found in Table 72.

Table 72: replacement cost

category	**M/E 4000**	**M/E 8000**	**D/G**	**AUX 4000**	**AUX 6000**	**AUX 8000**	**sum**
failures per year	2,001	1,587	2,197	3,667	2,661	1,377	
cost per replacement	7280	17135,71	3060	3950	3950	3950	
cost per inspection	733	1121	400	150	150	150	
replacement cost	14568	27193	6722	14486	10512	5438	**78920**
lower CI	4883	14815	196	1192	865	447	**22398**
upper CI	32618	39572	13249	27780	20158	10429	**143806**

Therefore the expected cost for replacement of failed equipment, based on the statistical analysis that was carried out, is about 79.000$, for a period of 1 year.

7.2 INTERPRETING THE RESULTS FOR REAL- LIFE DATA

The life data that were collected concern a fleet of 11 containerships, of which we examine and analyze the failures that occurred to the 7 vessels of them.

In this attempt, the failures for each category will be sorted by the time of occurrence in connection to the suggested by manufacturer inspection time. For example, for M/E 4000 hours' category, the following graph indicates that 42% of the incidents occurred at time less than the suggested by the manufacturer, 46% of the incidents occurred within the time span and 12% of the equipment was inspected at least one time before failure.

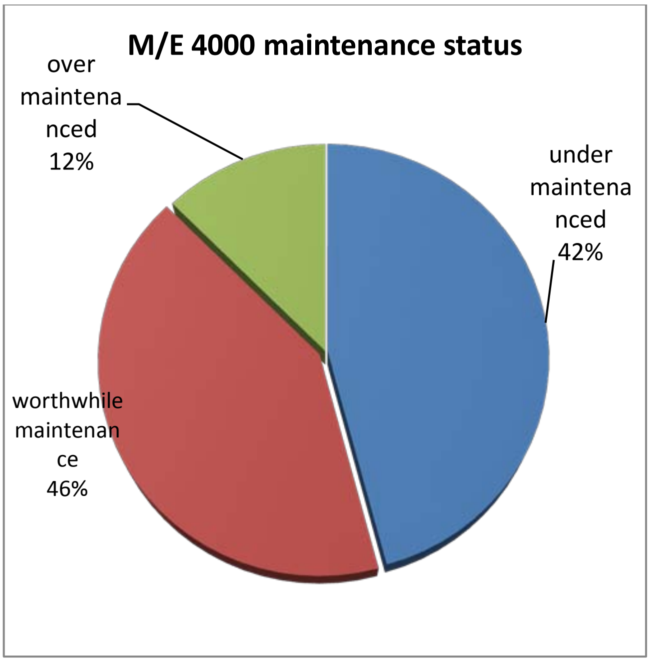


Figure 126: M/E 4000 maintenance status

The same graphs for the other categories can be calculated. In the following graphs can be seen that the majority of the equipments under investigation, have at least one failure prior even of the first suggested inspection.

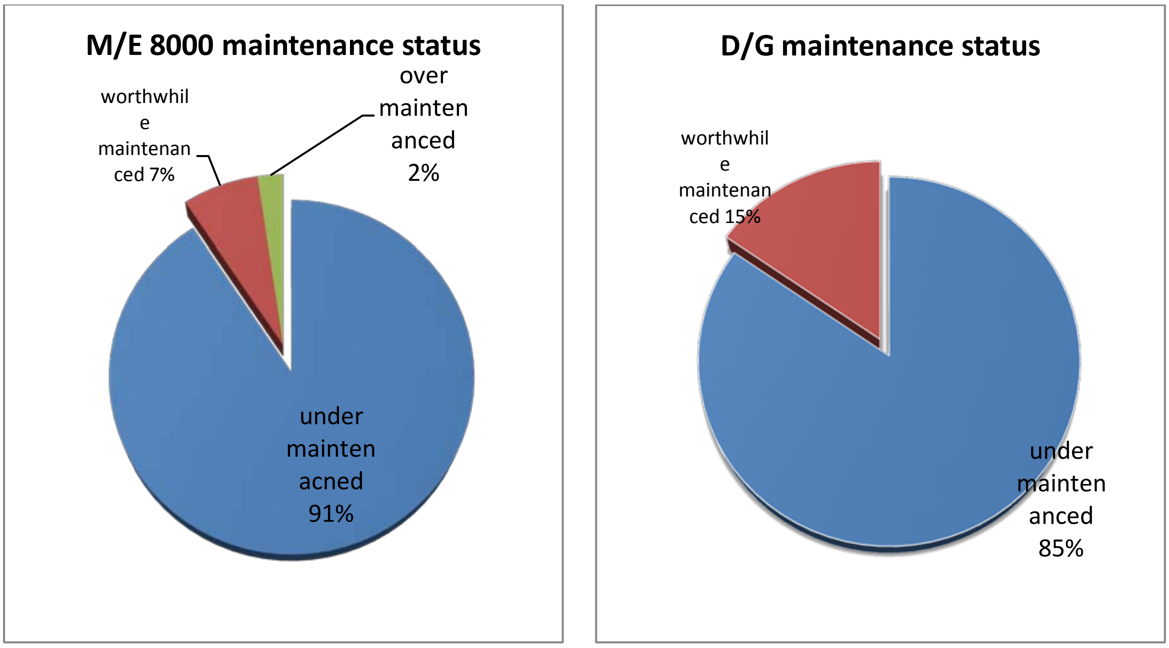


Figure 127: M/E 8000, D/G maintenance status

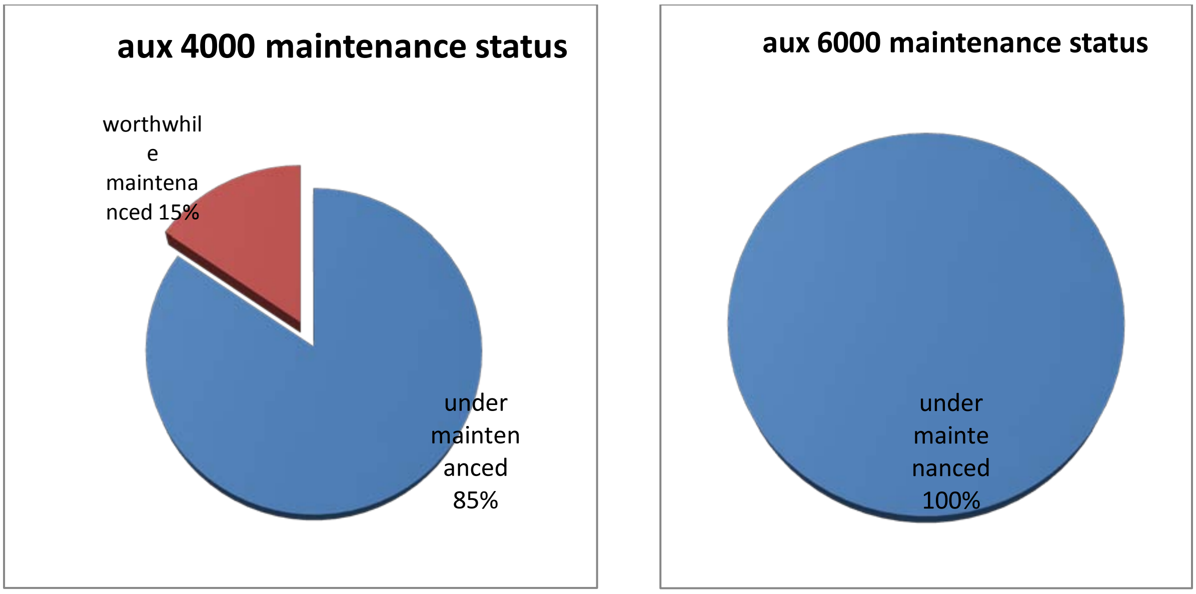


Figure 128: AUX 4000, AUX 6000 maintenance status

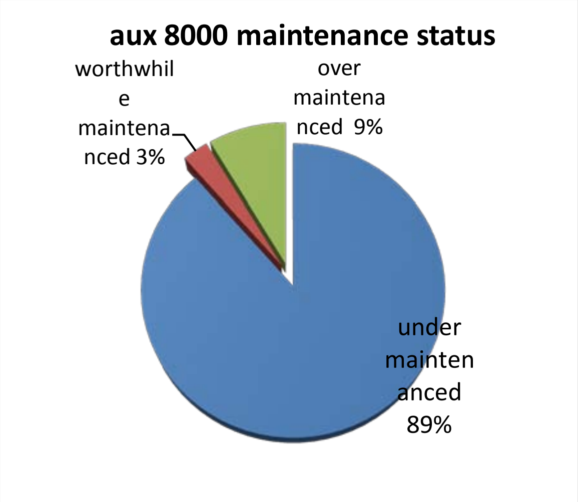


Figure 129: AUX 8000 maintenance status

Based on this analysis, the actual cost for repairs, for the under investigation period of 2 years, for the 7 containerships:

Table 73: Actual repair costs

category	M/E 4000	M/E 8000	D/G	AUX 4000	AUX 6000	AUX 8000	**sum**
under maintenanced replacement	101920	719700	61200	51350	43450	126400	**1104020**
lower CI	26840	364093	979	3575	3575	10075	**409137**
upper CI	179300	972493	66341	83325	83325	234825	**1619609**
worthwile, overmaintenanced inspection	9533	4486	800	300	0	600	**15719**

Using the same assumption, but cutting down the suggested inspection time to the 75% of the proposed by the manufacturer time, we can recalculate these costs.

Table 74: Repair cost

category	M/E 4000	M/E 8000	D/G	AUX 4000	AUX 6000	AUX 8000	**sum**
under maintenanced replacement	72800	565478	45900	43450	39500	94800	**861929**
lower CI	24400	308078	1335	3575	3250	7800	**348439**
upper CI	163000	822878	90465	83325	75750	181800	**1417219**
worthwile, overmaintenanced inspection	10266	11214	5600	300	150	1650	**29181**

The replacement cost per ship per year in the first case was 78.800$/ship/year and in the second case was 61.500$/ship/year.

In proportion, the inspection cost was raised from 1150$/ship/year to 2100$/ship/year.

If we recalculate these costs, but by cutting down the inspection time to the half of the proposed of the manufacturer we get:

Table 75: Repair cost

category	M/E 4000	M/E 8000	D/G	AUX 4000	AUX 6000	AUX 8000	**sum**
under maintenanced replacement	36400	325578	33660	27650	19750	47400	**490439**
lower CI	12200	177378	979	2275	1625	3900	**198358**
upper CI	81500	473778	66341	53025	37875	90900	**803420**
worthwile, overmaintenanced inspection	13933	26914	7200	900	900	3450	**53298**

The replacement cost per ship per year in this case reduced to 35000$/ship/year and the inspection cost was raised to 3800$/ship/year.

A plot that describes the relation between inspection and replacement cost , in regards to the reduction of inter inspection time can be produced.

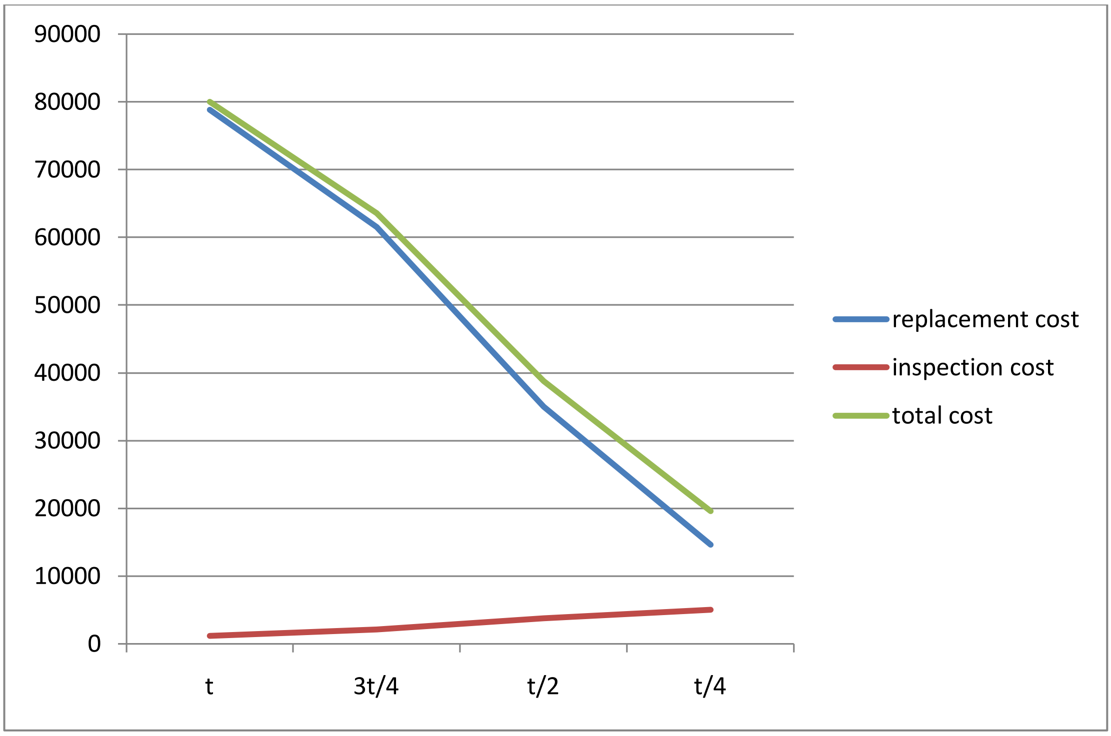


Figure 130: Replacement vs. inspection cost

In Figure 130, t is the time for inspection for each component as it is proposed by the manufacturer. We can see how the replacement cost decreases as the inspection is carried out more frequently.

Although the inspection cost increases, almost more than 3 times in the case of cutting down inspection time to half, since the absolute value is small, compared to the replacement cost, the overall cost is drastically decreasing.

Obviously the ideal case would be to perform a continuous inspection, but this is impossible, due to the large amount of equipment and machinery, that has to be inspected. For reference each one of the containerships has onboard at least 210 electric motors of different sizes. It is clear, that even if we reduce the inspection time of the electric motors from 4000 to 2000 operating hours, it will become very difficult to perform this type of remedial actions.

CONCLUSIONS

In this thesis, so far, we have dealt with reliability analysis of real life data which concerns failures of Main Engine, Diesel Generators and Auxiliary Equipment, taken from a fleet of containerships. These failures where categorized according to the suggested by the manufacturers, inspection time. Hereafter the distributions that describes the failures where calculated and relevant plots where produced.

In order to simplify calculations, the joint distribution of failures, for a vessel as a whole was calculated. Finally it was calculated the cost for necessary repairs based on the theoretical analysis and this cost was compared to the cost from the real life data.

This calculation of the cost could be considered as a qualitative comparison of the methods that were used in this research. The difference between the expected cost based on theoretical expected number of failures and the real cost that was calculated according to real- in situ data, was less than 1%. For the specific data of failures the present analysis could be considered as the proper approach to this subject.

However in real life situations, availability is often determined more by spares holdings and administrative times than by predictable factors such as mean repair times. Therefore predictions and models of system reliability and availability should be used as a form of design review or in the marine industry as a form of annual budget and planning review.

The consequences of failures in a vessel, apart from the repairs cost, could be the loss of hire in the case of a significant failure which could immobilize the vessel. The target is to plan the repairs to be carried out during a scheduled dry-docking or during anchorage.

This research can be considered as a guide for step by step reliability analysis, for a future collection of similar data. In this way and according to the primary laws of statistics, as the sample number increase, more reliable results could be extracted and expanded for other types of ships.

The major difficulty in similar attempts, is the collection of failures data. For example, the OREDA database includes information for 24000 offshore equipment units, involving more than 33000 failures. More analytical calculations could be performed, in case that a larger amount of such data was available. For example, a Failure Mode, Effects and Criticality Analysis could be carried out. That could lead to the calculation of severity of consequences for various failures.

Nevertheless, the creation of a national database, under the supervision of a University or any other authority, in which all the concerned could register failures data for marine industry, is a significant step for the actualization of such analyses. The tremendous number of vessels, operating under Greek flag or under Greek management, provides a guarantee that a future effort of this kind will be successful, even if only the support of the domestic shipping companies is granted.

APPENDIX A

Distributions statistical properties

For the distributions we used in this thesis, in this appendix can be found the statistical properties.

Weibull distribution

The Mean or MTTF

The mean, $\overline{T}$, (also called *MTTF* or *MTBF* by some authors) of the Weibull *pdf* is given by:

$$\overline{T} = \gamma + \eta \cdot \Gamma\left(\frac{1}{\beta} + 1\right)$$

where $\Gamma\left(\frac{1}{\beta} + 1\right)$ is the gamma function evaluated at the value of $\left(\frac{1}{\beta} + 1\right)$. The gamma function is defined as:

$$\Gamma(n) = \int_0^\infty e^{-x} x^{n-1} dx$$

The Median

The median, $\breve{T}$, is given by:

$$\breve{T} = \gamma + \eta(\ln 2)^{\frac{1}{\beta}}$$

The Mode

The mode, $\tilde{T}$, is given by:

$$\tilde{T} = \gamma + \eta\left(1 - \frac{1}{\beta}\right)^{\frac{1}{\beta}}$$

The Standard Deviation

The standard deviation, σ_T, is given by:

$$\sigma_T = \eta \cdot \sqrt{\Gamma\left(\frac{2}{\beta} + 1\right) - \Gamma\left(\frac{1}{\beta} + 1\right)^2}$$

The Weibull Reliability Function

The equation for the three-parameter Weibull cumulative density function, *cdf*, is given by:

$$F(T) = 1 - e^{-\left(\frac{T-\gamma}{\eta}\right)^{\beta}}$$

Recalling that the reliability function of a distribution is simply one minus the *cdf*, the reliability function for the three-parameter Weibull distribution is given by:

$$R(T) = e^{-\left(\frac{T-\gamma}{\eta}\right)^{\beta}}$$

The Weibull Conditional Reliability Function

The three-parameter Weibull conditional reliability function is given by:

$$R(t|T) = \frac{R(T+t)}{R(T)} = \frac{e^{-\left(\frac{T+t-\gamma}{\eta}\right)^{\beta}}}{e^{-\left(\frac{T-\gamma}{\eta}\right)^{\beta}}}$$

or:

$$R(t|T) = e^{-\left[\left(\frac{T+t-\gamma}{\eta}\right)^{\beta}-\left(\frac{T-\gamma}{\eta}\right)^{\beta}\right]}$$

These equations gives the reliability for a new mission of t duration, having already accumulated T hours of operation up to the start of this new mission, and the units are checked out to assure that they will start the next mission successfully. It is called conditional because you can calculate the reliability of a new mission based on the fact that the unit or units already accumulated T hours of operation successfully.

The Weibull Reliable Life

The reliable life, T_R, of a unit for a specified reliability, starting the mission at age zero, is given by:

$$T_R = \gamma + \eta \cdot \{-\ln[R(T_R)]\}^{\frac{1}{\beta}}$$

This is the life for which the unit will be functioning successfully with a reliability of $R(T_R)$. If $R(T_R) = 0.50$ then $T_R = \breve{T}$, the median life, or the life by which half of the units will survive.

The Weibull Failure Rate Function

The Weibull failure rate function, $\lambda(T)$, is given by:

$$\lambda(T) = \frac{f(T)}{R(T)} = \frac{\beta}{\eta}\left(\frac{T-\gamma}{\eta}\right)^{\beta-1}$$

Normal distribution

The Normal Mean, Median and Mode

The normal mean or MTTF is actually one of the parameters of the distribution, usually denoted as μ. Since the normal distribution is symmetrical, the median and the mode are always equal to the mean, $\mu = \tilde{T} = \breve{T}$.

The Normal Standard Deviation

As with the mean, the standard deviation for the normal distribution is actually one of the parameters, usually denoted as $\sigma_{T.}$

The Normal Reliability Function

The reliability for a mission of time T for the normal distribution is determined by:

$$R(T) = \int_T^\infty f(t)dt = \int_T^\infty \frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_T}\right)^2} dt$$

There is no closed-form solution for the normal reliability function. Solutions can be obtained via the use of standard normal tables. Since the application automatically solves for the reliability, we will not discuss manual solution methods.

The Normal Conditional Reliability Function

The normal conditional reliability function is given by:

$$R(t|T) = \frac{R(T+t)}{R(T)} = \frac{\int_{T+t}^\infty \frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_T}\right)^2} dt}{\int_T^\infty \frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_T}\right)^2} dt}$$

Once again, the use of standard normal tables for the calculation of the normal conditional reliability is necessary, as there is no closed form solution.

The Normal Reliable Life

Since there is no closed-form solution for the normal reliability function, there will also be no closed-form solution for the normal reliable life. To determine the normal reliable life, one must solve:

$$R(T) = \int_{T}^{\infty} \frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_T}\right)^2} dt$$

for *T*.

The Normal Failure Rate Function

The instantaneous normal failure rate is given by:

$$\lambda(T) = \frac{f(T)}{R(T)} = \frac{\frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{T-\mu}{\sigma_T}\right)^2}}{\int_{T}^{\infty} \frac{1}{\sigma_T\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma_T}\right)^2} dt}$$

Log logistic distribution

Mean, Median and Mode

The mean of the loglogistic distribution, $\overline{T}$, is given by:

$$\overline{T} = e^{\mu}\Gamma(1+\sigma)\Gamma(1-\sigma)$$

Note that for $\sigma \geq 1$, $\overline{T}$ does not exist.
The median of the loglogistic distribution, $\breve{T}$, is given by:

$$\hat{T} = e^{\mu}$$

The mode of the loglogistic distribution, $\tilde{T}$, if $\sigma < 1$, is given by:

$$\widetilde{T} = e^{\mu+\sigma\ln\left(\frac{1-\sigma}{1+\sigma}\right)}$$

The Standard Deviation

The standard deviation of the loglogistic distribution, σ_T, is given by:

$$\sigma_T = e^{\mu}\sqrt{\Gamma(1+2\sigma)\Gamma(1-2\sigma) - (\Gamma(1+\sigma)\Gamma(1-\sigma))^2}$$

Note that for $\sigma \geq 0.5$, the standard deviation does not exist.

The Loglogistic Reliability Function

The reliability for a mission of time *T*, starting at age 0, for the loglogistic distribution is determined by:

$$R = \frac{1}{1+e^{z}}$$

where:

$$z = \frac{T' - \mu}{\sigma}$$

$$T' = \ln(t)$$

The unreliability function is:

$$F = \frac{e^z}{1 + e^z}$$

The Loglogistic Reliable Life

The loglogistic reliable life is:

$$T_R = e^{\mu + \sigma[\ln(1-R) - \ln(R)]}$$

The Loglogistic Failure Rate Function

The loglogistic failure rate is given by:

$$\lambda(T) = \frac{e^z}{\sigma T(1 + e^z)}$$

Logistic distribution

The logistic mean or MTTF is actually one of the parameters of the distribution, usually denoted as μ. Since the logistic distribution is symmetrical, the median and the mode are always equal to the mean, $\mu = \tilde{T}$

The Logistic Standard Deviation

The standard deviation of the logistic distribution, σ_T, is given by:

$$\sigma_T = \sigma \pi \frac{\sqrt{3}}{3}$$

The Logistic Reliability Function

The reliability for a mission of time T, starting at age 0, for the logistic distribution is determined by:

$$R(T) = \int_T^{\infty} f(t)dt$$

or:

$$R(T) = \frac{1}{1 + e^z}$$

The unreliability function is:

$$F = \frac{e^z}{1 + e^z}$$

where:

$$z = \frac{T - \mu}{\sigma}$$

The Logistic Conditional Reliability Function

The logistic conditional reliability function is given by:

$$R(t/T) = \frac{R(T + t)}{R(T)} = \frac{1 + e^{\frac{T-\mu}{\sigma}}}{1 + e^{\frac{t+T-\mu}{\sigma}}}$$

The Logistic Reliable Life

The logistic reliable life is given by:

$$T_R = \mu + \sigma[\ln(1 - R) - \ln(R)]$$

The Logistic Failure Rate Function

The logistic failure rate function is given by:

$$\lambda(T) = \frac{e^z}{\sigma(1 + e^z)}$$

1. **David, Smith.** *Reliability, Maintainability and Risk - Practical Methods for Engineers.* s.l. : Elsevier, 1997.

2. **R.T.Anderson, A.A. Lakner.** *Reliability Engineering for Nuclear and Other High Technology Systems: A practical guide.* s.l. : Chapman & Hall.

3. **Nieuwhof, G.** *The concept of failure in reliability engineering.* 1984.

4. **Brennan, R.** *An essay on the history and future of reliability from the perspective of replication.* s.l. : Journal of Educational Measurements, 2001.

5. **A., Coppola.** *Reliability engineering of electronic equipment.*

6. **J. Saleh, K. Marais.** *Highlights from the early (and pre-) history of reliability engineering.* s.l. : Department of Aeronautics and Astronautics, MIT, 2004.

7. **Feingold, Ascher and.** *Repairable System Reliability: modeling inference, misconceptions and their causes.* 1984.

8. **R.Billinton, R.Allan.** *Reliability evaluation of engineering systems.*

9. **Limnios, N.** *Semi- Markov processes and reliability.* s.l. : Birkhauser.

10. **Herskovits, Cooper &.** *A bayesian method for the induction of probabilistic networks from data.* 1992.

11. **B. Navtig, H. Eide.** *Baeysian Estimation of System Reliability.* s.l. : University of Oslo, 1987.

12. **O Doguc, J Ramirez.** *A generic method for estimating system reliability using Bayesian networks.*

13. **Langseth, H.** *Analysis of OREDA Data for Maintenance Optimisation.* s.l. : SINTEF.

14. *Lifetime Data Analysis, An International Journal Devoted to Statistical Methods and Applications for Time-to-Event Data.* s.l. : Springer.

15. **Sturges, H.** *The choice of a class interval," Journal of American Statisticians Association, vol. 21, 65-66, 1926.*

16. **Prince, J.** *Engineering ststistics handbook, Chapter 7.* s.l. : SEMATEC.

17. **Nelson, Wayne.** *Applied Life Data Analysis.* s.l. : Wiley, 1982.

18. **Deshpande, J.** *Life- Time Data Statistical models and methods.* s.l. : Technometrics.

19. **Cambell, M. J.** *Statistics at Square One, Chapter3 Nonparametric tests.*

20. **M. Rausand, A. Hoyland.** *System Reliability Theory Models, Statistical Methods and Applications p.474.* s.l. : Wiley.

21. *Engineering Statistic Handbook, http://www.itl.nist.gov/div898/handbook/index.htm.*

22. **Hoaglin, Boris Iglewicz and David.** *"Volume 16: How to Detect and Handle Outliers", The ASQC Basic References in Quality Control: Statistical Techniques.* 1993.

23. **Stephen, M.A.** *Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes.* 1952.

24. **W. BLISCHKE, D.N. MURTHY.** *Reliability, Modelling, Prediction and Optimization, pp121.* s.l. : WILEY.

25. **Nelson.** *Reliability Data Analysis and Accelerated Life Testing.* 1969.

26. **G.Srinivasa Rao, R.R. Kantam et al,.** *Reliability estimation in log- logistic distribution from censored samples.*

27. **Chovanec, A.** *Simulation of FTA in Simulink.* s.l. : R&RATA #4, 2008.

28. **W. Wang, J. Loman.** *Reliability Block Diagram Simulation Techniques Applied to the IEEE STD.493 Standard Network.*

29. **Barlow, Richard.** *Engineering Reliability.*

30. **Stamatelatos, M.** *Fault tree handbook with Aerospace Applications.*

31. *Repairable System Analysis Through Simulation.* s.l. : RELIASOFT CORPORATION.