WEIBULL AND LOG-NORMAL DISTRIBUTIONS IN RELIABILITY ANALYSIS APPLICATIONS

Master of Science Thesis

Beldine Awuor OMONDI
Graduate School of Sciences
Eskisehir, 2017

WEIBULL AND LOG-NORMAL DISTRIBUTIONS IN RELIABILITY ANALYSIS APPLICATIONS

Beldine Awuor OMONDI

A thesis presented for the degree of
Master of Science, Statistics

Eskisehir
Anadolu University
Graduate School of Science
August, 2017

This thesis titled ”Weibull and Log-normal distributions in reliability analysis applications” has been prepared and submitted by Beldine Awuor Omondi in partial fullfillment of the requirements in “Anadolu University Directive on Graduate Education and Examination” for the Degree of Master of Science in Statistics Department has been examined and approved on …../…../……..

	Committee Members	**Signature**
Member (Supervisor)	**: Prof. Dr. Berna YAZİCİ**	**……………..**
Member	**: Assoc. Prof. Dr. Betül KAN KILINÇ**	**……………..**
Member	**: Assoc. Prof. Dr. Özlem ALPU**	**……………..**

…………

……………..
Date **Director Graduate School of Sciences**

ABSTRACT

WEILBULL AND LOG-NORMAL DISTRIBUTIONS IN RELIABILITY ANALYSIS APPLICATIONS

Beldine Awuor OMONDI

Anadolu University
Department of Statistics
Graduate School of Sciences, August, 2017

Supervisor: Prof. Dr. Berna Yazıcı

The life times of the components of an electrical device are often analyzed in quality-control process. Likewise the Design of Experiment is used mainly to achieve quality but its' application to life times are less common. Life times always associate with Reliability. In this study we focus on application of reliability analysis. The main purpose is to optimize the response by extending the lifetimes of a fan motor. An important point is to determine the stress factors subjected to different conditions which significantly effects the life times. Simulation design of weibull and log-normal distributions used to generate failure times and maximum likelihood (ML) applied in the estimation of parameters. A two level factorial experimental design is used to identify the factors that most influence the lifetimes.

Keywords: Reliability, Simulation, Weibull distribution, Log-normal distribution, Factorial design.

ÖZET

WEIBULL VE LOG-NORMAL GÜVENİRİLİK ANALİZİ UYUGULAMALARİ

Beldine Awuor OMONDI

Anadolu Üniversitesi
Istatistik Anabilim Dalı
Fen Bilimleri Enstitüsü, Agustos, 2017

Danışman: Prof Dr. Berna Yazıcı

Kalite kontrol süreçlerinde elektrikli cihazların ömür süreleri sıkça analiz edilmektedir. Kalite süreçlerinin çalışılmasında Deney Tasarımı büyük yer tutmaktadır, ancak yaşam sürelerin uygulaması daha azdır. Yaşam süreleri analizi her zaman güvenilirlik ile ilişkilendirilmektedir. Bu çalışmada, güvenilirlik analizinin iki parametreli weibull dağılımı ve log normal uygulamasına odaklanılmaktadır. Temel amaç, bir fan motorunun yaşam süresinin uzatılarak, yanıt değerinin optimize edilmesidir. Buradaki önemli nokta, zaman sürelerinin anlamlı biçimde etkileyen farklı koşulları göz önünde bulundurarak stres faktörlerinin belirlenmesidir. Weibull dağılımı ve log normal dağılımı, simülasyon tasarımı oluşturulması ve parametrelerin maximum likelihood (ML) tahminlerinin elde edilmesinde kullanılmıştır. Yaşam sürelerini etkileyen en önemli faktörlerin belirlenmesi amacıyla, 2 düzeyli faktöriyel deney tasarımı kullanılmıştır.

Anahtar Kelimeler: Güvenilirlik, Simülasyon ,Weibull dagılımları, Log-normal dagılımları.

ACKNOWLEDGEMENTS

I thank the almighty God for his grace and love. This report wouldn't have been a success if it weren't because of his grace. To him be the glory.

First of all, I would like to express my heartfelt gratitude to my supervisor Prof. Dr. Berna Yazıcı her advice, assistance, patience and encouragement during the two years of my study at Anadolu University and throughout this work.

My appreciation also goes to Assoc. Prof. Dr. Betül Kan Kılınç and Assoc. Prof. Dr. Özlem Alpu for their constructive comments.

I would like to extend my special thanks to my Mother and my sisters for their encouragement, financial support and prayers during my study period. Special thanks also to Kenneth Esekhaigbe for his love and support throughout my stud period and especially during this work. .

Finally, I am grateful to my friends Edward, Ayten and classmates Asli, Birsen, Huruy, Merve and Saed many others who supported me throughout this venture.

Beldine Omondi

August-2017

TABLE OF CONTENTS

page

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATIONS

The list includes some of the symbols and abbreviations used in this study. Some of the symbols and abbreviations have also been defined in the text at their appearance.

Symbol or Abbreviation	Meaning
AFR	Average Failure Rate
DOE	Design of Experiment
MLE	Maximum Likelihood Estimation
MTBF	Mean Time Between Failure
F (t)	Cumulative Distribution Function (C.D.F)
f (t)	Probability Density Function (P.D.F)
RSM	Response surface Methodology
PDF	Probability Density Function
RBD	Reliability Block Diagram
SEV	Smallest Extreme Value

1. INTRODUCTION TO RELIABILITY

1.1 Introduction

Technology advancement, budget constrains and restrictive testing time has translated to high reliability of products [10]. As a result manufactures strive to remain competitive and completely relevant to the global market by meeting consumer's demands in terms of quality expectations. Like wise, the product lifetime is equally an increasingly important characteristic [11] and manufactures need to understand the expected life times of their product under various operating conditions. As a result, manufacturing companies target new products to reduce the initial failures, minimize random failures and to increase product reliability.

Lifetime of a product is not only a common concern to engineers but also to statisticians and is a key characteristic of product quality and reliability. It characterizes the time-span during which a product can be expected to operate safely under certain conditions and meet specified standards of performance.

A lot of efforts have been made by researchers to develop new reliability models [12]. [13] developed reliability model of a complicated product with multiple failure modes under different stress conditions, [14] developed a new lifetime cost optimization model to predict product lifetime and [12] studied measures to reflect interaction between the reliability and performance characteristics of a product, [15] in the review article described the distribution of lifetimes and how lifetimes depend on set experimental (predictor) variables among others.

This study aims at determining factors that affect the lifetimes of a product and also increasing the lifetimes by conducting simulation with right censoring both for weibull and log-normal distribution. Simulation model is developed based on these factor levels. Maximum likelihood method is used to estimate the parameters of both weibull distribution and log-normal distribution. Finally we find that combination of input-factor values that optimizes the response.

1.2 What is Reliability?

Industrial statisticians plan experiments for a long period of time to improve product quality and reliability. This section explains what reliability is and why it is one of the basic components of quality.

Reliability is a word with many different implications. When applied on systems, it is the ability of the system to perform certain tasks according to a specified standards, consequently it is applied to a piece of equipment or a component of a system to mean the ability of that component or system to fulfill what is required of it [5]. The origin use of term was purely qualitative [5]. For instance, aerospace engineers recognized the desirability of having more than one engine on a plane without any precise measurement of failure rate.

Benavides [16] defines reliability as a mathematical relationship that models it when the stated conditions and the values of set parameters are given. The value of set parameter define the stated conditions of the product at any fixed time. The argument by [16] puts it that engineers need to consider reliability, lifetime, design parameters and operational parameter as an engineering design tool to relate stated parameters with design and operational parameters that define the stated conditions at a given time.

According to Khan and Islam [17] reliability is the percentage measure of degree to which product/system is in an operable and committable state at the point in time when it is needed. In other words, the availability of a product or a system.

As used today, reliability is more of a quantitative concept that conveys dependability, successful operations or performance and the absence of failure. The its quantitative concept therefore suggests the need for methods of measuring reliability.Several reasons are provided why reliability has to be quantitative. For instance; economic is considered as the most important aspect in that to improve reliability costs money, and this can only be justified if the cost of unreliable components are measured. On the other hand, different standards of reliability are required in different applications that is; relatively high failure rates may be considered acceptable in unmanned satellite. Reliability can be analyzed in two main headings, product reliability and system reliability. Product reliability is the probability that a product will operate properly for a specified period of time [5]. In the probability context, satisfactory performance is directly associated with the concept of failure and malfunctioning in that the equipment when in operation, is either operating satisfactorily or has failed or malfunctioned [18]. Therefore satisfactory performance is any other condition that is neither a failure nor malfunction.

Since reliability is an indicator of capability to perform within required limits of time when in operation, it normally involves a parameter which measures time. This may be any time unit which is preferable in cases where continuous operation is involved [18].

In this study, the ability of an item to perform a required function under stated environmental and operational conditions for a specified period of time will be adopted as the operational definition of reliability. This definition is generally accepted and used in [18], [5] and [3] Some of the key the key terminologies also used in this thesis are defined below:

- Failure - The termination of a component's ability to perform a required function.
- Function - The intended process for a component or a product that leads to a detectable but not necessarily an acceptable performance.
- Parts of a component - the smallest replaceable unit in a system.

It is equally important to note that reliability and lifetimes are used interchangeably throughout the study.

1.3 Importance of Reliability

Reliability is very important in that product and brand reputations are made or destroyed by their product reliability performance. Poor reliability (unreliability) causes adverse consequences and therefore a number of products or systems are serious threat. Unreliability may have implications for: Safety, reputation, good will, delays, profit margins, cost of repair and maintenance and competitiveness. Work to minimize failures, improve maintenance effectiveness, shorten repair times and meet customer and organization expectation has numerous benefits. Reliability has broad and important impact across the product life cycle. Therefore the cost of unreliability can damage a company's image.

The benefits of reliability analysis can be classified in the following ways:

- Unexpected failures cost time for both customers and organization to resolve the failures. Using reliability and availability concepts we can minimize failures and avoid wasting time.
- Products work under environmental and use conditions imposed by the customer. Creating a product that matches the expectations imposed by the customer permits the product to work as expected. Understanding the conditions allows the design to meet without over designing thus optimizing product cost and customer satisfaction. Expectation of consumers may not be fulfilled unless higher reliability

values are achieved because today consumers are conscious of how unreliability is more costly. Otherwise companies are faced with the loss of goodwill.

- Throughput Downtime for any reason reduces the system's throughput, downtime can be minimized by applying predictive and preventative maintenance programs.

  A well-maintained system minimizes operating expenses and maximizes throughput.

- Distribution Fewer failures and optimized maintenance implies fewer spare parts in the logistics system.

  This minimizes the distribution system costs for transportation, logistics, and storage for spare parts. This also minimizes service labor costs.

- Products that operate as expected without failure avoid being returned or serviced under warranty. Calls to service support, troubleshooting, product returns, failure analysis, and re-engineering all part of the cost of unreliability. The warranty provides customers insurance in case of failure and with reliability engineering techniques the costs are minimized.

- Some product failure cause unintended or unsafe conditions leading to loss of life or injury. Reliability engineering tools assist in identifying and minimizing safety risks.

- Product failures can cause the loss of property. Minimizing failures and mitigating the damage caused by any failure minimizes the exposure to liability for the property loss.

- Reliability studies reveal the types of failures experienced by components and systems and recommend design, research, and development efforts to minimize these failures. Enhancing the design team's reliability engineering capabilities through training and staffing of reliability professionals enables the entire team to make decisions fully considering the impact on product reliability. This reduces the need for expensive redesign or rework costs to address reliability related design errors.

1.4 System Reliability

Reliability relationship between products or systems and their components can be demonstrated through reliability block diagrams (RBD).

A system consists of components which determines whether it will function or not. As products become more complex(have more components), the chances that they will not function increases depending on the method of their arrangements. The method of arranging components affects the reliability of entire system. The RBD is used to determine the system or subsystem reliability of a design. RBD based reliability evaluation is useful when requirements dictate the level of design reliability or during component selection when each component has a different reliability [19]. For more complex systems these are useful as a usual tool to find out where failure occur. There are different ways of arranging components in a system or a product. Components can be arranged in series, parallel or combination of both.

1.4.1. Series system

A system is called series system when all the system functions only if every one of its components functions. In the same sense, a series circuit functions when there is an unbroken path through the components that form the system. Components arrangement in series system is shown in figure 1.1. As more and more components are added to the series, the system reliability decreases [18]. However, if the failure times of the components are statistically independent, then the system reliability is the product of the reliabilities of each components that comprises the system [2]. That is, let R_1, R_2....R_n denotes components of a system or product then reliability of this system R_{system} and R_i denote the i^{th} component reliability in the system.

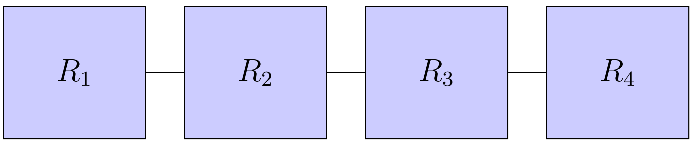

Figure 1.1. *A Series System Block Diagram with Four Components, adopted from [8]*

The system reliability will be determined by the equation 1.1 as follows:

$$R_{system} = R_1 * R_2 * ....R_n = \prod_{i=1}^{n} R_i \tag{1.1}$$

This shows that the reliability of a series system is always lower than that of the least reliable component in the system.

Here is an example of a series system. Assuming that life of the $i-th$ component follows a weibull distribution with a given scale η and shape β parameters the system reliability at time t becomes $R_{system}(t)$ where ;

$$R_{system}(t) = \prod_{i=1}^{n} exp\left[-\left(\frac{t}{\eta_i}\right)^{\beta}\right], \tag{1.2}$$

Which gives the equivalent scale parameter value and the mean [2] expressed in equation1.3 and equation 1.4 respectively.

$$\eta_e = \left\{ \sum_{i=1}^{n} \frac{1}{\eta_i^{\beta}} \right\}^{-\frac{1}{\beta}} \tag{1.3}$$

$$\mu = \eta\beta_1 = E(t_{system}) = \eta_e\beta_1. \tag{1.4}$$

suppose the life a certain component follows a weibull distribution with shape parameter β = 2 η_1 = 1000, η_2 = 500 and η_3 = 100, the system scale parameter is:

$$\eta_e = \left[\frac{1}{1000^2} + \frac{1}{500^2} + \frac{1}{100^2} \right] = 97.6hrs \tag{1.5}$$

the mean life of this series system is:

$$E(t_{system}) = \eta_e\Gamma\left(1 + \frac{1}{\beta}\right) = 97.6 * \Gamma\left(1 + \frac{1}{2}\right) = 86.5 \tag{1.6}$$

In case we are able to double the life of a with the longest lifetime, the system's scale parameter value will become:

$$\eta_e = \left[\frac{1}{2000^2} + \frac{1}{500^2} + \frac{1}{100^2} \right] = 97.9hrs \tag{1.7}$$

On the other hand when the component with the shortest lifetime is doubled, the result is as follows:

$$\eta_e = \left[\frac{1}{1000^2} + \frac{1}{500^2} + \frac{1}{200^2} \right] = 182.6hrs \tag{1.8}$$

This means that manufacturers should devote resources and effort to improve the lifetime of component with the smallest scale parameter in order to realize the greatest improvement in system life.

1.4.2. Parallel system

A system is referred to as parallel system when there are many parallel paths through reliability block diagram and components as well. System will function as long as long as at least one of its component function. The reliability diagram in figure 1.2 shows a parallel system having four components. The parallel systems have two advantages over the series system: The system with more components is more reliable, the reliability of a parallel system is greater than the reliability of the most reliable component(best component) in the system [2]. Therefore the probability that a system will fail equals the probability that all components fail. Reliability of a parallel system is calculated as: 1-Prob[system fails]:

$$R_{system}(t) = 1 - \prod_{i=1}^{n} (1 - R_i). \tag{1.9}$$

Such that;

$$R_{system} = 1 - (1 - R_1)(1 - R_2(1 - R_3)(1 - R_4) \tag{1.10}$$

In figure 1.2 suppose R_1 =0.70, R_2 =0.80 and $R_3 = R_4 = 0.85$,then the reliability of the corresponding parallel system is:

R_{system} = 1-$(0.30 * 0.20 * 0.15 * 0.15)$ = 0.99865

It can be seen that the overall reliability is greater that the most powerful component in the system, in this example R_3 andR_4. Alternatively, given reliability function $R_{system}(t)$, simulation via a spreadsheet can be used to determine the system life distribution and its mean. In this case the system lifetime t_{system} is the life of the longest live component [2].

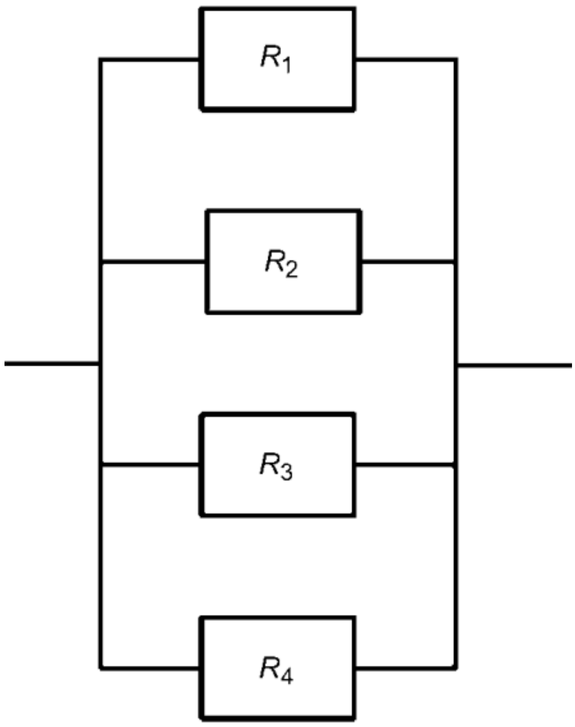

Figure 1.2. *A parallel system block diagram with four components [2]*

1.4.3. Combination of series and parallel arrangement

A combination of series and parallel is another common configuration of components.An example of such system is shown in figure 1.3. It consists of sub-systems that the components are either parallel or series. In order to determine the reliability of the whole system, the system is broken down into subsystems that are in parallel or in series. Suppose R_1,R_2,...R_9 represents reliability of components R_1 toR_9 respectively.

Given that, $R_1 = R_2 = 0.65$, $R_3 = R_4 = R_5 = 0.8, = R_6, R_7 = R_8 = R_9 = 0.73$

$R_{1,2}$= $1-(1-0.65)^2 = 0.8775$

$R_{3,4,5} = 1-(1-0.8)^3 = 0.992$

$R_{6,7,8,9} = 1-(1-0.73)^4 = 0.9947$

$R = R_{1,2} * R_{3,4,5} * R_{6,7,8,9} = 0.8659$

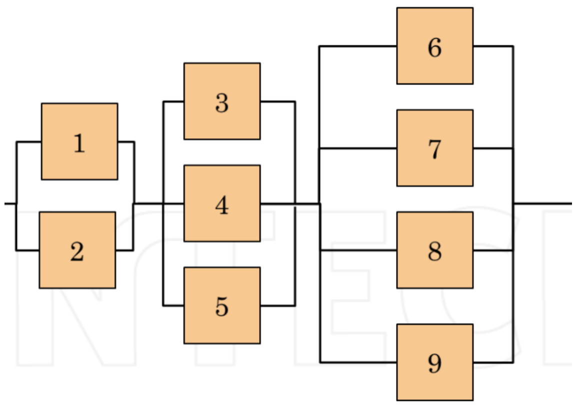

Figure 1.3. *A combination of series and parallel system. The picture shows 9 components [2]*

1.5 Thesis Outline

The thesis is organized as follows: Chapter one which is an introductory chapter that initially describe reliability background information definition, importance of reliability and reliability block diagrams. This aims to explain the importance of reliability and the effect of components arrangement on the reliability of a product. Chapter two presents the concept of reliability and parameter estimations of lifetimes. This is discussed as the theory and implementation of maximum likelihood estimation, which is used as a method of parameter estimation for lifetimes distribution. Chapter three introduces the RSM as the methodology of the study. It describes the theory and implementation of the reliability models that are used to model the lifetime distributions and the impact of different variables on the lifetime. Chapter four discusses the simulation results of the study. Chapters five concludes with a summary of the study, recommendation and future research.

2. RELIABILITY CONCEPTS

This chapter discusses literature review about the main topic of the dissertation.

2.1 Reliability testing

Many manufactures have adopted reliability testing to meet and exceed the demands of customers [20]. Considering the reliability definition by [5] as "probability of performing without failure a specified function under given conditions for a specified period of time." Therefore reliability testing usually involves simulation of conditions under which the item will be used during its lifespan.

Reliability does not compare the product to some predefined specifications, such as the case with quality assurance, but rather investigates the performance over a predefined period of time. For example, mobile phone devices can undergo an accelerated life test. In this case, devices are exposed to events that simulate real life situations that happen to mobile phones like drops, spills, or excessive heating. The goal of this test is to find out whether the produced items meet the specified minimum reliability requirement.

To reduce the cost of reliability testing, manufacturers apply sampling schemes to select items that represent all produced devices (population). As in any sampling, it is assumed that appropriate inference about the true population characteristics can be made based on rightfully selected samples.

Formulating reliability sampling in terms of testing a statistical hypothesis can be as follows:

- H_0: Mean lifetime is greater or equal to specified time t.
- H_1: Mean lifetime is less than specified time

Rejecting H_0 when it is true (type I error) would result to producer's risk while failing to reject H_0 while it is false (type II error) would result to consumer's risk. It can be seen that the concept of Producer's risks and consumer's risks are similar type I and type II errors respectively.

The producer's risk can be defined as the the probability of failing satisfactory products. In this case, It is associated with the level of reliability which has a high probability of acceptance, and, therefore, low fraction of non-conforming units. On the other hand, consumer's risk is the probability of passing faulty products.

When developing a sampling plan and method for reliability sampling, there is a need to answer the following inquiries :

- Is the testing procedure representative of real life events?
- Does the criteria to pass/fail act in accordance with consumer and producer risks?
- What sample size should be drawn?

Besides that, it is important to decide a prior what constitutes a failure, what units of measurement will be used, and when the test will be terminated.

In reliability testing, there are two types of tests that can be used [21]. Method of attributes: Noting the presence (or absence) of some characteristic or attribute in each of the items in the group under consideration and counting how many items do (or do not) possess the attribute, or how many such events occur in the item, group, or area.

Method of variables: Measuring and recording the numerical magnitude of a characteristic for each of the items in the group under consideration; this involves reference to a continuous scale of some kind. When using the method of attributes, a reliability engineer can specify the maximum number of non-conforming units that are allowable in the sample. It should be noted that the manufacturer should not knowingly produce any number of defective products and it is always better to have zero non-conforming items. Variables sampling plans are more often used for reliability sampling. Time to failure often described by Weibull distribution.

The test specification may be written in terms of: Mean life hazard rate at(number of items failing at time t), reliable life (some point of time beyond which predefined proportion of items will survive).

2.2 Reliability function

Lifetime distributions are theoretical population models used to describe the lifetimes of a product . For instance, if someone is interested in a particular type of a computer, then the population is considered to be all the lifetimes obtainable from computers of this type. Alternatively, it may be restricted to computers from one particular manufacturer made during a set period of time. Therefore, a lifetime distribution is the cumulative distribution function (CDF) for the population [3]. Tobias and Trindade [3] gives two useful interpretations of the CDF represented by F(t) as:

- F(t) is the probability that a random component selected from the population fails by t.
- F(t) is the fraction of all components in the population that fail by t.

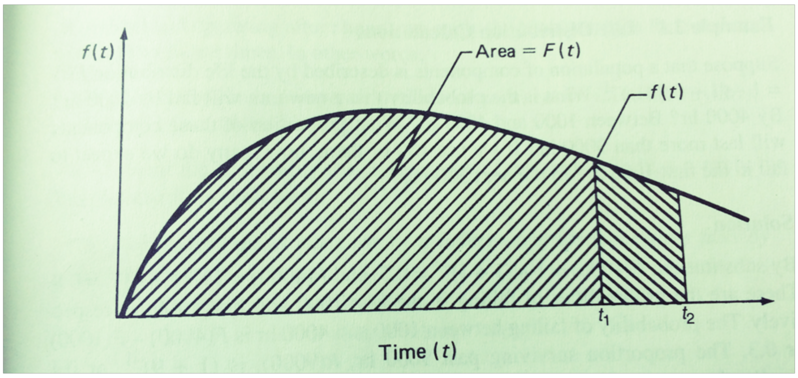

Figure 2.1. *Cumulative distribution function [3]*

Cumulative distribution function, F(t) can be plotted on time against the probability density function, f(t) as shown in figure 2.1 which is the area under probability density function, f(t) to the left of time (t). F(t) (area of the shaded region) is the probability that a new component will fail by time of operation t. the area defined by f(t) between two vertical lines at time t_1 and t_2 equals to the probability of a new product surviving to time t_1 and the failing in the interval between t_1 and t_2. The area can be obtained by subtracting the area to the left of t_1 from the area to the left of t_2, $F(t_2) - F(t_1)$ which is the probability that a new unit survives to t_1 but fails before time t_2. This is also the fraction of the entire population that fails in that interval.

When the attention is on the un-failed components (survivors), then the reliability function (survivor function) is defined by:

$$R(t) = 1 - F(t) \tag{2.1}$$

F(t) could also be called unreliability function may be discussed in two ways as:

- The probability a random unit selected from the population will still be operating after time t
- The fraction of all units in the population that will survive at least time t

2.2.1. The failure rate (Hazard function)

Distribution of lifetime data can be modeled with the help of probability density function (PDF), cumulative density function (CDF), reliability function and failure (hazard) rate function. PDF and CDF are very known terms since they are broadly used in statistical manner but failure rate function has a particular property in reliability study and because of this it is not widely known. PDF or f(t) stands for failure probability density function in reliability. From statistics it is familiar that f(t) is the derivative of F(t) with respect to t. It is the probability of failure in the interval t to $t + dt$ in which δt is an instant of time. Failure or hazard rate function is the instantaneous rate of failure for the survivors to time t during the next instant of time Tobias and Trindade. Failure rate function is expressed as units of failures per unit time. It is not a probability and can take values greater than 1. Failure rate is denoted by either z(t) or h(t). To see how failure rate is calculated, a little probability statistics must be used:

2.2.2. Average failure rate function

It is sometimes useful to define average rate over an interval of time that averages the failure rates in that interval. AFR (t_1, t_2) stands for the average failure rate between time t_1 to time t_2. The simplest way to specify AFR is to integrate the failure rate over the internal and divide by duration of the interval.

$$P(failure\,in\,next\,\delta t|survive\,to\,t) = \frac{F(t + \delta t) - F(t)}{R(t)} \tag{2.2}$$

we divide the equation by δ t in order to covert it into rate as:

$$\frac{F(t + \delta t) - F(t)}{R(t)\delta t} \tag{2.3}$$

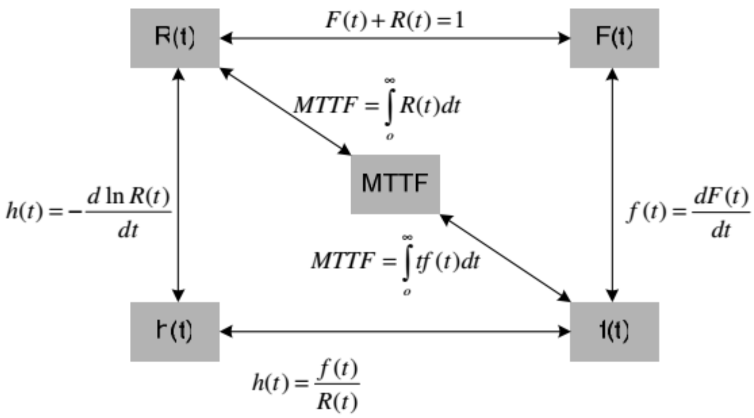

Figure 2.2. *Relating F(t) and others*

2.3 The Phases of Bathtub Curve

Traditionally, the lifetime of a product can be demonstrated by three different patterns of failures over time. These patterns are combined to produce a bath-tab curve as shown in Fig 2.3. The curves not only represent reliability performance of components or non repaired items but also observes the reliability performance of a large sample of homogeneous items entering the field at some start time (usually zero). If we observe the items over their lifetime without replacement then we can observe three distinct patterns or shapes.

The three distinct patterns or periods are mathematically categorized as; a decreasing rate of failure, a constant rate of failure and an increasing rate of failure which in practice are known as infant mortality, useful life and wear out as shown in Fig. 2.3. During the infant mortality phase, the weaker components are removed during manufacturing process, pre-delivery testing or when an item comes into service. At the useful phase, new component has an equal probability of failure as the old component. Finally the wear out failure which has increasing failure rate caused by weariness, fatigue and degradation [3]. As reliability (or part) of a product improves, cases of failed parts becomes less frequent in the field [22]. When a new product is introduced, a company ensures that initial failure is minimized, random failures are reduced during the expected working period of a product and finally to extend the product's lifetime.

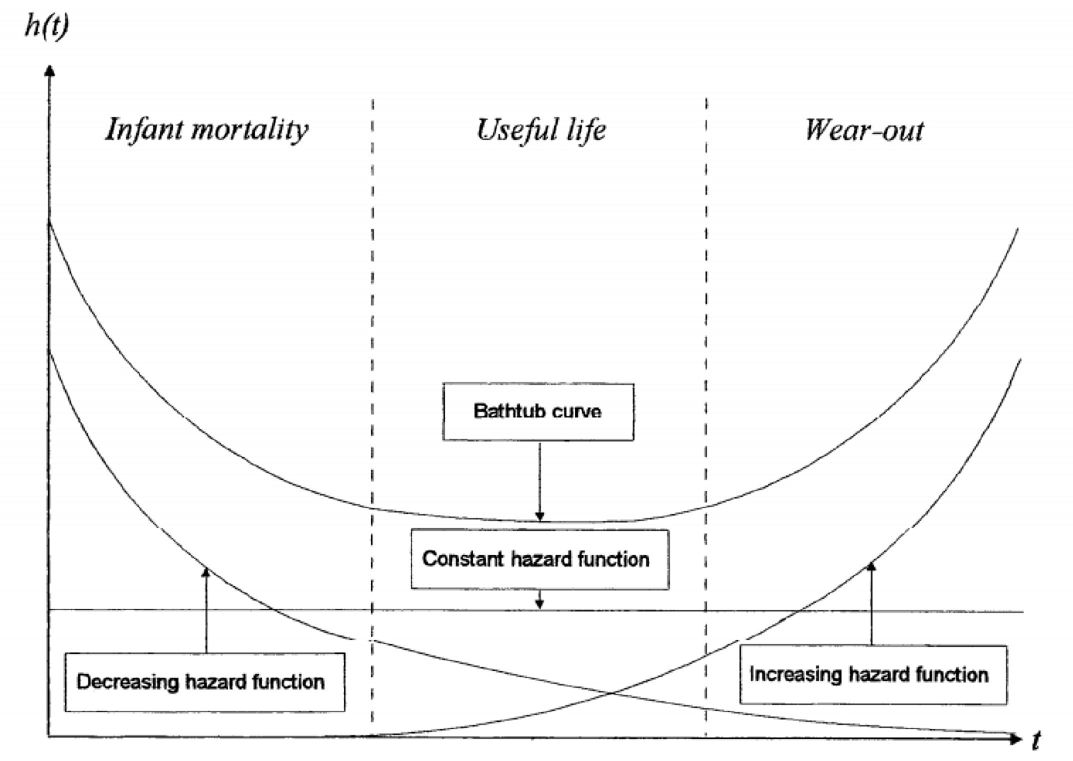

Figure 2.3. *Bathtub curve showing different types of failures. Time can be hours, months or years [3]*

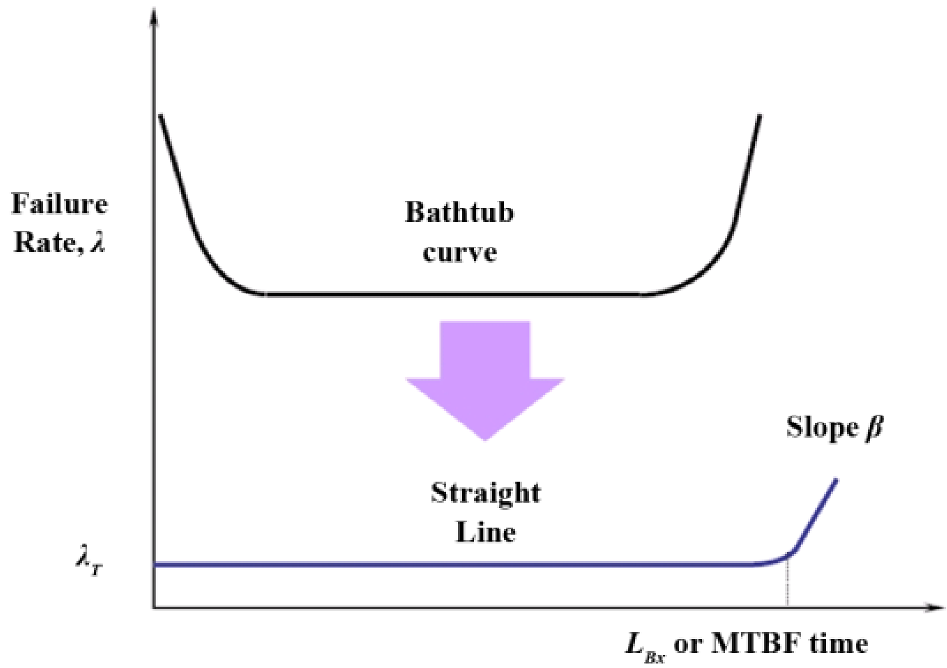

Figure 2.4. *The Bathtub Curve and Straight Line with Slope β [3]*

2.4 Sources Reliability Data and Reliability Data Classification

2.4.1. Sources of Reliability Data

The components or product's population history of failure statistics is the basis for reliability models [23]. Consequently, there are other sources that may be considered in order to predict the product's future behavior. Reliability analysis depends on historical data [24]. Ruhi, [23] argues that the availability of actual usage conditions and valuable information on the performance of any product makes the field data superior to experiment, and is directly linked to the financial aspect of the product. Borgia, De Carlo, Fanciullacci, and Tucci, [25] claims that the performance of a new product can only be estimated both from experimental tests and historical analysis of the previous similar products.

There are many sources of collecting field product reliability data [23]. Warranty claim which may not contain adequate information related to reliability data since most systems are designed to track finances and not the product performance [23], sales and forecasting which gives the population of units in the field being used at any given time and is vital to perform reliability oriented calculation, field service data connects with field services like repairing a failed product. However it is only powerful if the system is in place to gather necessary information during the service call. Others are customer support data and return parts or failure analysis data. The different sources mentioned here are summarized as:

- Inspection records.
- Tests:field demonstration,environmental qualification and field installation.
- Failure reporting systems from customers
- Records generated during development phase
- Previous experiences with similar product
- Warranty claims
- Repair facility records
- Customers' failure reporting systems

Censoring, data aggregation (pooled or combined) and data with small failure events are some of the challenges in reliability data collection. Most cases, researchers are faced with several repairable similar products which may have different performance [24]. This is because, units may be installed in different environment and may operate under different conditions and maintained by different policies. In other words, different operating conditions (temperature, humidity, pressure) may change the product's failure patterns.

2.4.2. Reliability data classification

2.4.2.1 *Complete data*

Life data is referred to as complete data when all failure data are available at the end of an experiment [4]. That is, the failure data contain failure times of all components in the sample. McCool [2] describes as one for which the exact value of the variable is observed for each member of the sample.

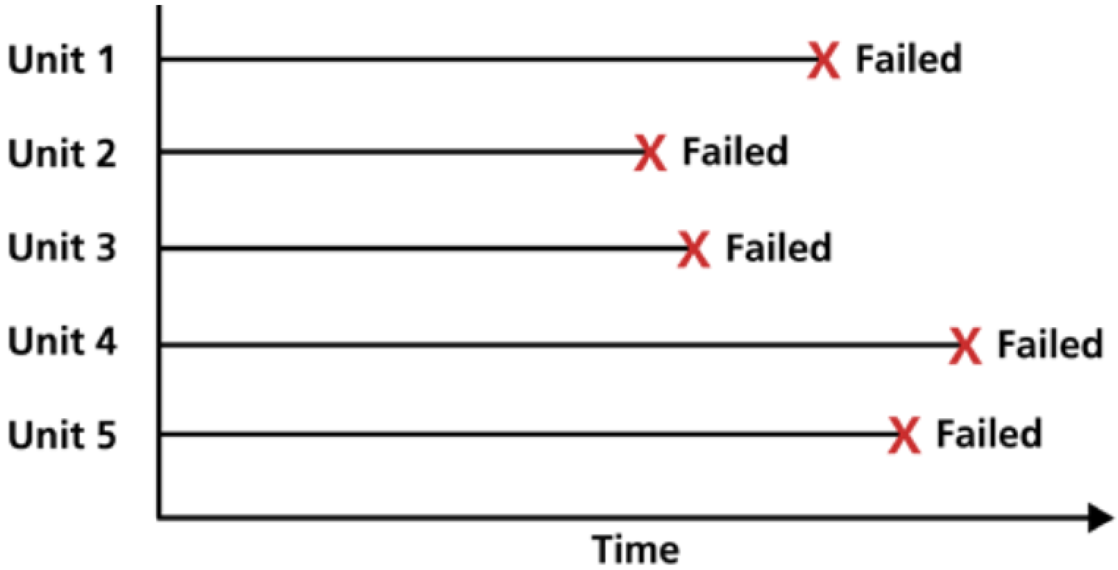

Figure 2.5. *complete data. [4]*

2.4.2.2 *Censored data*

Generally, the data used in failure time analysis are censored. Censoring is common in reliability experiments and occurs when it is impossible to observe failure times of all components.In some cases, It is also impractical to wait for all units to fail. Many products are designed to last longer period before failure, but manufactures are impatient to wait that long to understand the lifetime distributions. Censoring occurs when the experimenter does not observe the exact time of failure. Yan-Qiao and Shi-Liang [4] describing it as failure times of failed units and running times of the units that have not failed by the end of experiment.

There are three ways in which reliability data can be censored: Right censoring occurs when the test stops before all units fail. In type I right censoring, an experiment terminates at a predetermined time. On the other hand, type II right censoring occurs when the experiment stops after a pre-specified number of failures. Experiments with type I and type II censoring may also contain interval, and/or left censoring. Left censoring occurs when the failure occurs before a known time (i.e. the unit failed before the first inspection at week one).

Engineers often perform reliability experiments with censoring to determine how different factors affect product life. For instance, if engineers believe operating temperature affects the life of a product they may design an experiment with different temperature levels and operate the products at each temperatures.

The experimenters implement type II right censoring in order to complete the experiment in a timely fashion, while ensuring a specific number of failures.In this case, the researchers record the times at which a given number of products fail for each temperature level. The ones that have not failed by the end of specified period are censored observations. Interval censored data, units are found to have failed between two inspections. Interval censoring occurs when one knows a failure occurs between two times but does not know the exact failure time. For example, a unit failed sometime after it was inspected at ten weeks, but before it was inspected at eleven weeks.

An example, In a failure analysis test, let n be the number of units. During time C, r failures are observed, where $0 <= r <= n$. Since the times of failure for $n - r$ units are unknown, but it is known that these times to failure are larger than C, the data are right censored. Another example that explains both the censored and complete data is shown in figure 2.6.

In reliability analysis it is commonly assumed that the variables C and T are independent. This is called non-informative censoring. The distribution of survival times of units that are censored at a particular time is no different from that of units that are still observed at this time. One common type of independent censoring is the simple type I censoring, where all subjects in the study are censored at the same, fixed time.

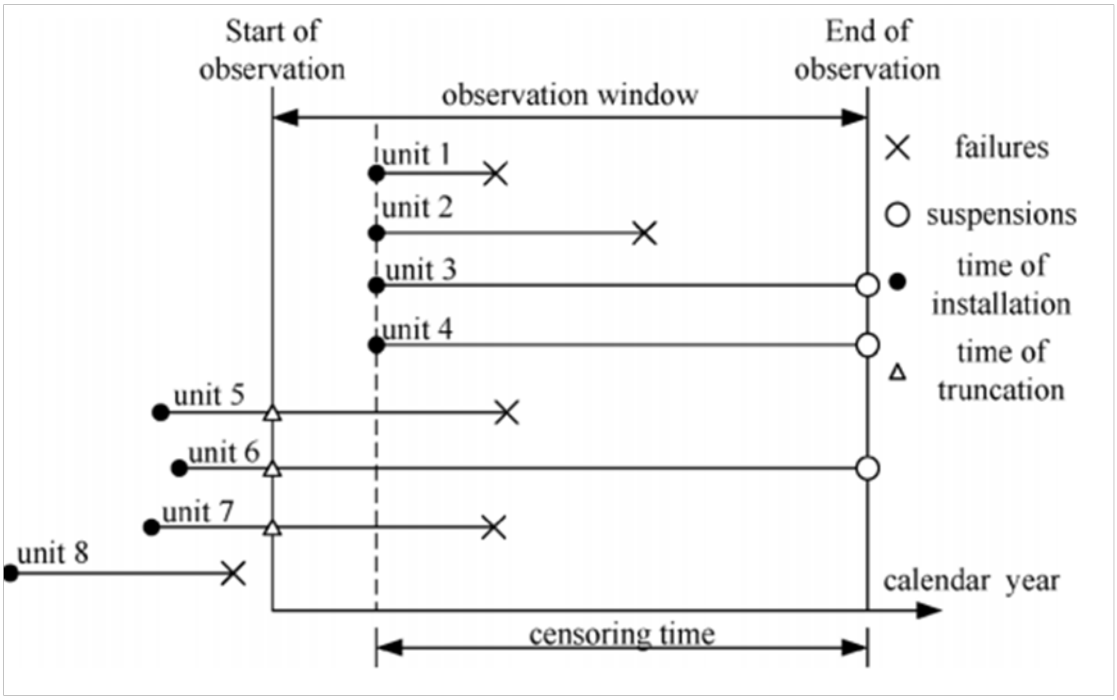

Figure 2.6. *An illustration of complete and censored data. [5]*

2.5 Lifetime Modeling and Parameter Estimation

This section describes the models that can be used for lifetime modeling and reliability estimation.

2.5.1. Lifetime modeling

Whenever design of experiment (DOE) is used for life testing, the response is life or failure time [26]. Failure time distribution is the most widely used measure for reliability of a product. This distribution is constructed based on failure time data of a product. The failure time or lifetime of a product is described as a continuous random variable T. Its probability distribution function is characterized by cumulative density function, probability density function, reliability function or hazard function. The reliability function gives the probability of a product surviving up to time t while the hazard function also known as the failure rate function describes the probability of failure at the smallest interval (t, δ +t), given survival up to.

2.5.2. Distributions associated with lifetimes

Design and quality engineers considerably apply design of experiment (DOE) which is a statistical technique. Consequently, these techniques cannot be used to test plan and analyze reliability experiment because the distribution of lifetimes is usually skewed to the right(non-normal)[15]. As a result, the F tests from analysis of variance are not valid. However, weibull distribution commonly used in reliability analysis does not belong to the exponential family [27]. In addition, censoring may occur in lifetime data. When censoring exists, DOE that based on least square method cannot be used to improve the product reliability.

2.5.2.1 ***Weibull distribution***

The Weibull distribution invented by Waloddi Weibull is a generalization of the exponential distribution that is appropriate for modeling the lifetimes having constant, strictly increasing and decreasing hazard functions [6].

It is the leading distribution in the world for fitting and analyzing life data. Weibull distribution has proved to be a successful model for many product failure mechanisms because it is a flexible distribution with a wide variety of failure rate curve shapes. Hence it is capable of describing various failure rate conditions by adjustment of parameters and so different models can be derived from Weibull distribution.

The cumulative distribution function is given by:

$$F(t) = 1 - exp^{-\eta t^{\beta}}, t > 0, \beta > 0, \eta > 0 \tag{2.4}$$

where β and η are both positive valued parameters. it is clear that $\beta = 1$ gives the exponential with mean $1/\beta$ hence it is viewed as generalization of the exponential distribution.

Abernethy[28] stated advantages of the distribution as follows:

- It has the ability to provide accurately reasonable failure analysis and forecast failures with extremely small samples.The Small samples thus allow the study to be cost effective.
- It provides a simple and useful graphical plot of the failure data as the plot is extremely important to managers and engineers.
- It exhibits wide range of distribution shapes which makes it the leading distribution.

Weibull distribution has either two or three parameters. The two parameter Weibull distribution consists of beta β and η. β is the shape or sometimes called Weibull gradient whereas η is the scale, feature or characteristic life parameter. Shape parameter controls the shape of the distribution while scale parameter fixes one point of the cumulative Distribution Function F(t), the 63.2 percentile or characteristic life point. 63.2 $\%$ of the population fails by the characteristic life point, independent of the value of the shape parameter. This is expressed in hours, days, cycles, etc. In most cases it it assumed that the value of the shape parameter (β) and the scale parameter (η) are unknown.

If it is known that an item will not fail until a specific time in service, then a third parameter can be added to Weibull distribution. This parameter is called location parameter and is symbolized by gamma (γ). It should be noted that there are several ways of symbolizing these parameters, here β, η and γ are used. Shape, scale, and location parameters must be greater than zero and the distribution is defined for only positive times.

This is normally associated with the bathtub curve, shape parameter is the most important issue because all the three distinct region of the bathtub curve can be modeled with Weibull distribution. If shape parameter is wrongly estimated, the lifetime model will be useless to find out the reliability or survival at one point of time.

The parameter takes three different types of values for the different periods or regions. The infant mortality period of lifetime is modeled with a $\beta < 1$. Failure rate is decreasing during infant mortality and Weibull distribution with a shape parameter of $\beta < 1$ also indicates the same curve of the failure rate function.

In the same way Weibull distribution with a shape parameter of $\beta = 1$ is used to model useful life period failures, because if $\beta = 1$, Weibull reduces to exponential distribution with a constant failure rate. For modeling wear-out period, Weibull's shape parameter takes a value of $\beta > 1$ since the failure rate is increasing.

Wrong estimation of the shape parameter affects results of the Weibull model, because wrong lifetime period of population could be taken up. The actual values could show different periods but the model could point out values of different period. That is why any reliability estimation at any time would be incorrect.

The weibull distribution is a popular distribution for modeling the failure times of a product because of its flexibility in being able to model the multiple types of failure mechanisms. For a random variable T, denoting the failure time, the commonly used parameterization of the two parameter weibull probability density function (PDF) is:

$$f\left(\frac{t}{\eta},\beta\right) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{(\beta-1)} exp\left(-(\frac{t}{\eta})\beta\right), \quad t > 0 \tag{2.5}$$

and the cumulative distribution function (CDF), the probability that an item will fail before time t is:

$$F\Big(\frac{t}{\eta},\beta\Big) = 1 - exp\Bigg(- (\frac{t}{\eta})^{\beta}\Bigg),\ \ \eta > 0,\quad \beta > 0 \tag{2.6}$$

η is the scale parameter and β is the shape parameter.

The hazard function is :

$$h(t/\eta,\beta) = \lim \delta t- > 0\,\frac{P(failure\,in\,t, t+\delta t)/T > t}{\delta t}$$

$$= \frac{f(t)}{1-F(t)} = \frac{\beta}{\eta}\Bigg(\frac{t}{\eta}\Bigg)^{\beta-1} \tag{2.7}$$

The general assumption is that the value of the shape parameter is constant regardless of the predictors (independent variables).

The mean and the variance of the distribution are:

$$E(T) = \eta\,\Gamma\Bigg(1+\frac{1}{\beta}\Bigg) \tag{2.8}$$

$$Var(T) = \eta^2\,\Gamma\Bigg(1+\frac{2}{\beta}\Bigg) - E\big(T\big)^2 \tag{2.9}$$

respectively. where Γ is the gamma function. The mean and the variance are both functions of weibull parameters η and β . Another important function of the two parameters is the weibull quantile function t_p;

$$t_p = \eta\,\Bigg(- log\big(1-p\big)\Bigg)^{\frac{1}{\beta}} \tag{2.10}$$

The lifetime quartile represents the time at which a specified proportion p of the population fails. For example if $\beta = 1$ and $\eta = 50$ then $t_.80 = 80.47$ is the time at which 80% of the population will fail. Notice too, $F(t_p) = p$ and $F(\hat{p}) = t_p$.

The censoring rate for the distribution is given by ;

$$p_c = exp\,\Bigg(- (\frac{C}{\eta})^{\beta}\Bigg) \tag{2.11}$$

Where C is the predetermined fixed time at which censoring occurs for any units that have not failed.

The scale parameter η has the same units as T and determines the spread of the distribution. It is sometimes referred to as the characteristic life because for any value of β, η is the time by which 63% of the units are expected to fail. The shape parameter β is a unitless number that reflects the specific failure mechanism.

The Weibull distribution models early failures or infant mortality failures when $\beta < 1$. Products that follow this early failure distribution typically fail due to a design flaw or a manufacturing defect. Infant mortality failures often arise with electronic components.

The Weibull distribution models wear out failures when $\beta > 1$. A wear out failure implies that a product's failure rate increases with time, quite common for mechanical systems. The Weibull distribution closely resembles the normal distribution for $\beta = 3$. One can also model random failures under the Weibull distribution using $\beta = 1$, which corresponds to the exponential distribution. Random failures are independent of system time and often due to external events. The effect of β can be translated into various modes of failure. This has been summarized in table 2.1.

Table 2.1. *Life characteristics for weibull distribution*

β value	Failure type	Implication
$\beta < 1$	early failures	high probability of failure at early stages(short life)
$\beta = 1$	random failures	failures are independent of time
$1 < \beta < 4$	early wear out	generic failure modes e.g. corrosion
$\beta > 4$	rapid wear out	wear out

It is sometimes more convenient to use an alternative parameterization for the Weibull distribution which is based on the relationship that exists between the Weibull distribution and the smallest extreme value (SEV) distribution, a location-scale distribution. For the random variable Y , the SEV pdf and cdf are;

$$f(y/\mu, \sigma) = \frac{1}{\sigma}\phi_{sev}\left(\frac{y-\mu}{\sigma}\right), \ \ -\infty < \mu < \infty, \ \ 0 < \sigma, \ \ -\infty < y < \infty, \tag{2.12}$$

$$\Phi_{sev}\left(\frac{y-\mu}{\sigma}\right); \ \ -1 < \mu < 1; \ \ 0 < \mu; \ \ -1 < y < 1; \tag{2.13}$$

where $\Phi_{sev}(z) = \exp\left[z - \exp(z)\right]$, $\Phi_{sev}(z) = 1 - \exp\left[-\exp(z)\right]$ and $z = y - \sigma\mu$. If T follows a Weibull(η; β) distribution, then the random variable $Y = \log(T)$ follows the $sev(\mu = \log(\eta); \sigma = \beta_1$) distribution. The alternative parameterization for the random variable T having a Weibull distribution, is

$$f(t/\mu;\beta) = \frac{\beta}{t}\Phi_{(sev)}(\beta\Big(\log(t) - \mu)\Big) = \frac{\beta}{t}\exp\left[z - \exp(z)\right] \tag{2.14}$$

$$F(t/\mu,\beta)) = \Phi_{(sev)}\beta\Big(\log(t) - \mu\Big) = 1 - \exp\left[-\exp(z)\right]; t > 0 \tag{2.15}$$

where, $z = \beta(\log(t) - \mu)$ and $\mu = \log(\eta)$. Under this parameterization the Weibull quantile function is $t_p = \exp\left[\mu + \frac{\Phi^{-p}_{(sev)}}{\beta}\right]$

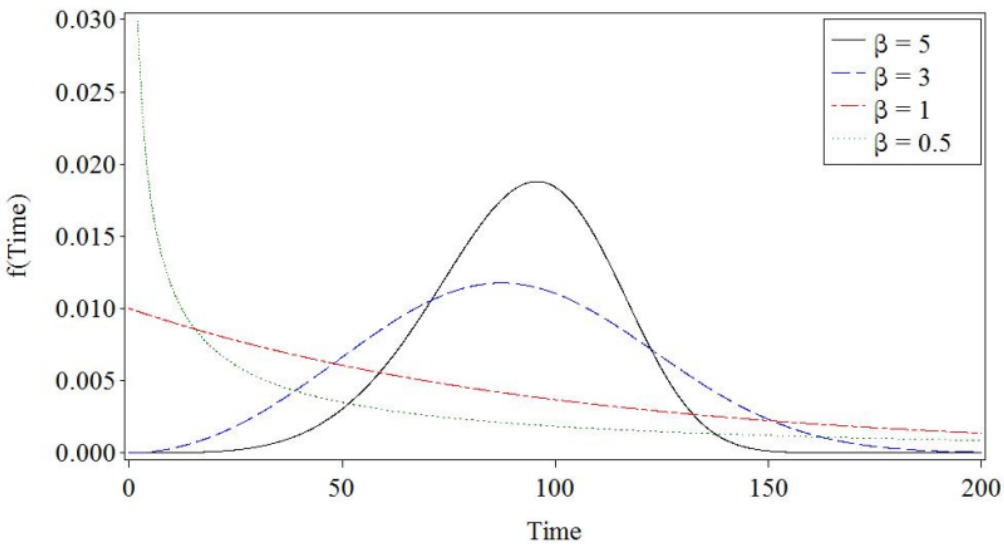

Figure 2.7. *Probability Distribution Function of the Weibull Distribution [6]*

Figure2.8 below shows the behavior of different shape parameters β of o.6, 1 and 3 i.e failure rate functions of weibull distribution also known as the decreasing rate, constant and increasing rates of failures respectively, with a scale parameter η with value of 1000 and different shape parameters

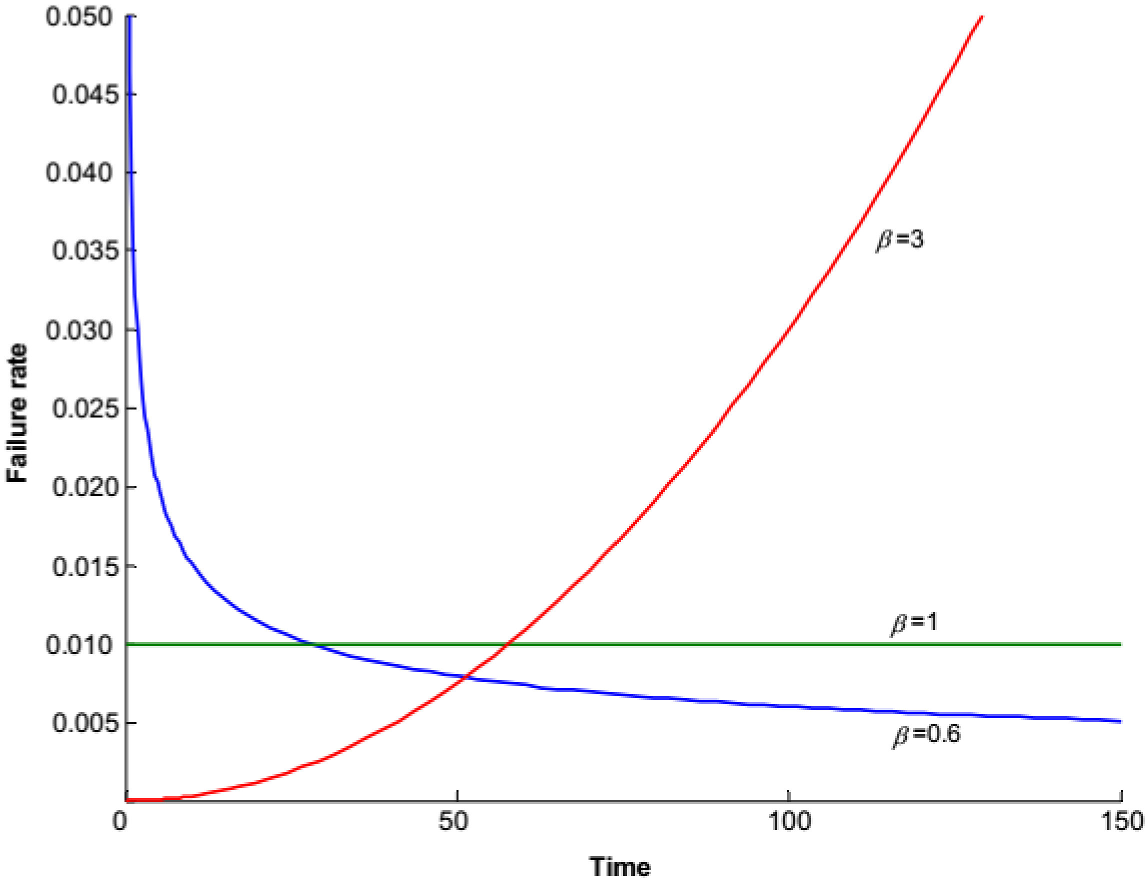

Figure 2.8. *Failure-rate Function with Scale Parameter η=1000 and Shape Parameters β [7]*

Table 2.2. *Summary of life characteristics for weibull distribution [1]*

Life characteristic	Equation
Proportion failing before time t	$F(t) = 1 - exp^{-\left(\frac{t}{\eta}\right)^{\beta}}$
Reliability	$R_t = exp^{-\left(\frac{t}{\eta}\right)^{\beta}}$
Mean life	$\mu = \eta\Gamma\left(1 + \frac{1}{\beta}\right)$
Hazard rate	$h(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1}$
Cumulative hazard	$H(t) = \left(\frac{t}{\eta}\right)^{\beta}$

2.5.2.2 *Exponential distribution*

The exponential distribution is the only distribution that exhibits the no aging property[6]. It has a probability density function (PDF)

$$f(t/\theta) = \frac{1}{\theta} exp\left(- \frac{t}{\theta} \right), t > 0 \tag{2.16}$$

and the cumulative density function (CDF)

$$F(t/\theta) = 1 - exp\left(- \frac{t}{\theta} \right), t > 0 \tag{2.17}$$

where $\theta > 0$. The hazard function, the (limit of the) probability of failure in a small interval divided by the length of the interval is constant thus having a constant failure rate as a characteristic property of the exponential [3].

The exponential distribution also has the memoryless property: the probability of surviving an additional t time units is the same regardless of the age of the system. The constant hazard function, along with the memoryless property, makes the exponential distribution rather unrealistic for most situations. Thus, through regression models can be developed for the exponential distribution, they are not particularly useful in practice. Exponential regression does, however, serve as a springboard to more complicated models.

2.5.2.3 *Log-normal distribution*

Log- normal distribution is one of the distributions commonly used for modeling lifetimes or reaction-times, cycles-to-failure in fatigue, material strengths and loading variables in probabilistic design and is particularly useful for modeling data which are positively skewed [29]. It has been extensively discussed by many authors including [30, 29, 8, 31] among others. There is a very close relationship between the normal and log-normal distribution. If $X = log(T)$ is normally distributed with mean μ and standard deviation σ , then the distribution of T becomes a two-parameter lognormal distribution with parameter $\theta = (\mu; \sigma)$. The probability density function of such a two-parameter lognormal distribution is given in equation 2.18.

When the natural logarithms of the times-to-failure are normally distributed, then we say that the data follow the log-normal distribution. Both log-normal and weibull distributions sometimes may fit a specific set of data equally well. They are not in the same mathematical family but they can be modeled to fit lifetime data with acceptable level of accuracy. It is argued that the log-normal distribution at earlier times predict lower average rates of failure than that of weibull distribution. In equation 2.18 , parameter μ can be interpreted as the mean of the random variable's logarithm, while the parameter σ may be interpreted as the standard deviation of the random variable's logarithm.

In addition, μ is said to be a scale parameter, while σ is said to be a shape parameter of the log-normal density function.Basak, Basak, and Balakrishnan [29] suggest that it is more appropriate to use $\beta = exp(\sigma^2)$ and $\eta = exp(\mu)$ as the shape and scale parameters of the log-normal variable T respectively. Log-normal distribution is summarized according to its probability density function (PDF).

The PDF for the log-normal distribution is given by:

$$f(t/\mu, \sigma^2) = \frac{1}{\sigma t \sqrt{2\pi}} exp\left(\frac{-\left(\log t - \mu\right)^2}{2\sigma^2} \right), \quad t > 0 \tag{2.18}$$

Where σ is the shape parameter and $\mu = T_{50}$, median (scale) parameter.

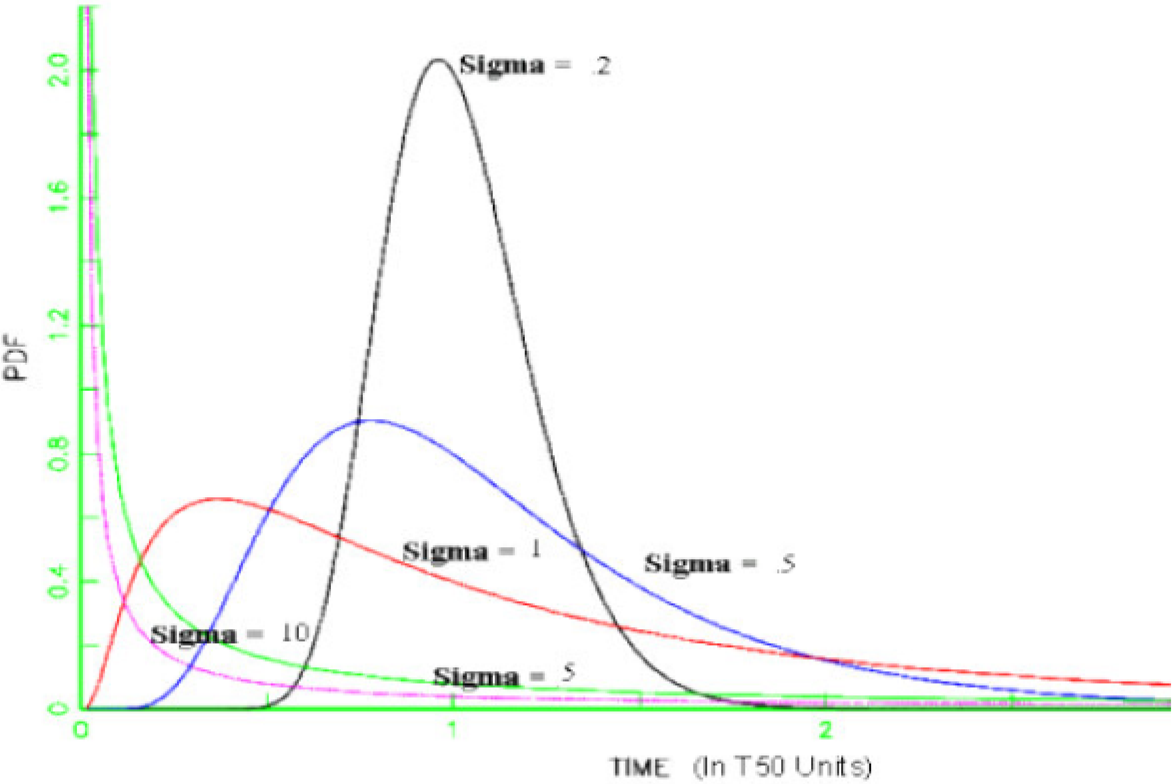

Figure 2.9. *Probability Density Function of log normal Distribution [8]*

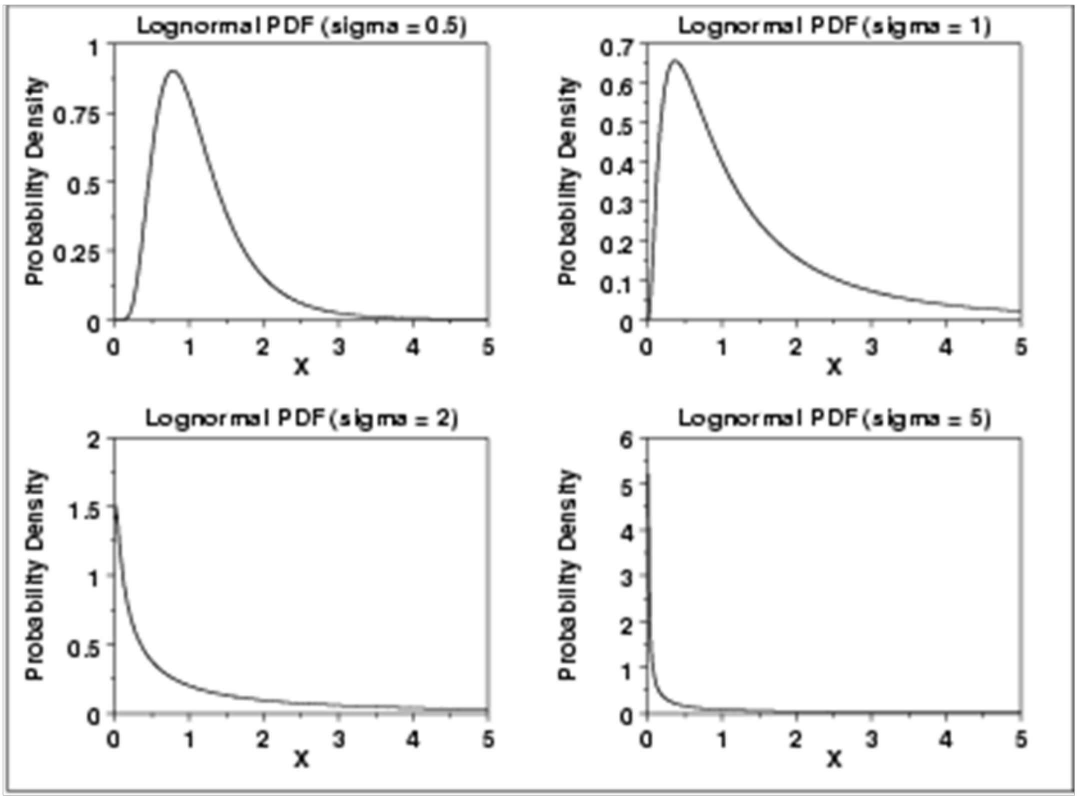

Figure 2.10. *Log-normal Probability Density Function for Different Values of σ. [8]*

Antle [32], stated that the multiplicative property is an important property of the log-normal distribution. The property states that if two independent random variables X_1 and X_2 distributed as log-normal (μ_1, σ_1^2) and log-normal(μ_2, σ_2^2) respectively, the the product of X_1 and X_2 is distributed as log-normal ($\mu_1\mu_2, \sqrt{\sigma_1^2 + \sigma_2^2}$).

Another important property of the log-normal distribution identified by [32] is the fact that for the very small values of σ ($\sigma < 0.3$), the log-normal is nearly not different from the normal distribution. However, unlike the normal distribution , the log-normal does not possess a moment generating function.

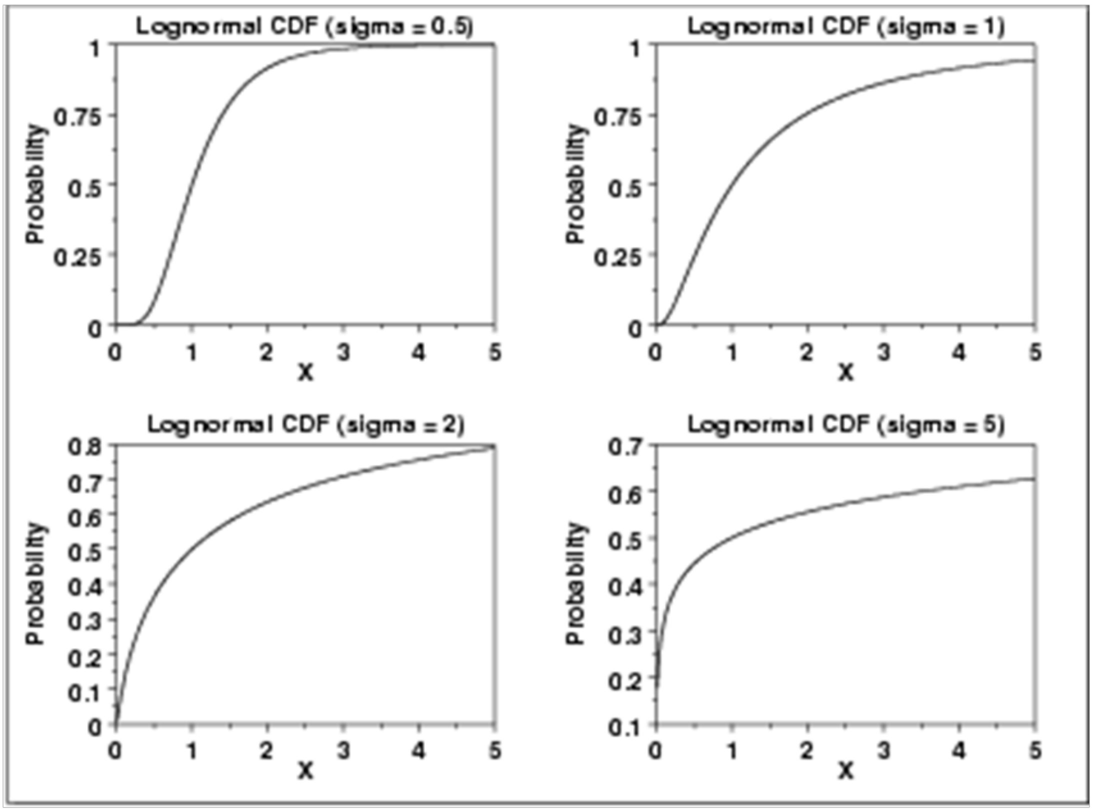

Figure 2.11. *Plot of the Log-normal Cumulative Distribution Function with values of σ [6]*

Table 2.3. *Life Characteristics for Log-normal Distribution*

Life characteristic	**Equation**
Proportion failing before time t	$F(t) = \Phi\left(\frac{lnt\frac{t}{T_{50}}}{\sigma}\right)$
Reliability	$R_t = 1 - F(t)$
Mean life	$E(t) = T_{50}exp^{\frac{\sigma^2}{2}}$
Hazard rate	$h(t) = \frac{f(t)}{R(t)}$
PDF f(t)	$\frac{1}{\sigma t\sqrt{2\pi}}e^{-\left(\frac{1}{2\sigma^2}\right)\left(lnt-lnT_50\right)^2}$
Median lifetime of 50% failure point	T_{50}

The two dimension α is used as the location parameter for the log-normal distribution. The table show comparison of weibull distribution and log-normal distribution when the location parameter is included. The table shows the comparison between the two distributions though they do not represent same conditions. This is because the scale parameter of the weibull distribution η represents the approximately 63% of components or product failing at time t, while the scale parameter of the log-normal distribution α represents 50% of the components expected to fail by time t. The mean life is also known as the mean time to failure for both weibull and log-normal distribution.

Table 2.4. *Comparison of weibull and log-normal distribution with location parameter*

Function	Weibull	Log-normal
Cumulative Density Function F(t)	$1 - exp\left[-\left(\frac{t-t_0}{\eta}\right)^{\beta}\right]$	$\Phi\left[ln\left(\frac{t-t_0}{\sigma}\right)^{\alpha}\right]$
Cumulative hazard function H(t)	$\frac{\left(t-t0\right)^{\beta}}{\eta}$	$ln\left(1-\Phi\left[in\frac{\left(t-t_0\right)}{\sigma}\right]^{\alpha}\right)$
Probability density function	$\frac{\beta}{\eta}\left(\frac{t-t_0}{\eta}\right)^{\beta-1} exp\left(-\left(\frac{t-t_0}{\eta}\right)\right)^{\beta}$	$\frac{exp-\frac{\left[ln\left(\frac{t}{\sigma}\right)^{\alpha}\right]^2}{2}}{\sqrt{2\pi\left(t-t_0\right)}}$
Reliability function R(t)	$exp\left(-\left(\frac{t-t_0}{\eta}\right)^{\beta}\right)$	$1-\Phi ln\frac{\left(t-t_0\right)^{\alpha}}{\sigma}$
Median life	$t_0 + \eta\left(ln2\right)^{\frac{1}{\beta}}$	$t_0 + \sigma$
mean life	$t_0 + \eta\gamma_1$	$t_0 + \sqrt[\sigma]{w}$
variance	$\eta^2\left(\gamma_2 - \gamma_1^2\right)$	$\sigma\left(w^2 - w\right)$

2.5.3. Estimating parameters of lifetimes distribution

There are different methods that can be used to estimate parameters of lifetimes data. These methods are categorized as graphical and statistical methods.

The graphical method of parameters estimate include the use of data transformation and plots of the transformed data to estimate the underlying model. These plots and transformation depends on the assumed model. The graphical method is simple and easy to find comparatively good estimates of parameters. From the graphics, the model fit can also be determined. Graphical method has been extensively discussed by several researchers. [8] stated the advantage of graphical method as the ability to visually test of the model i.e test data to the fitted model thus, used for model check (how well the points fit the model). In addition, the method is also considered to be useful in providing an initial estimation as the starting point for statistical methods. Other advantages include, the plot making visual sense and calculations can be done with help of a software package. The disadvantage of the graphical method is that the estimates obtained from this method is not accurate compared to the one concluded from statistical methods, lack of minimum variance for large samples and finally, parameters lack confidence interval, difficulty in dealing with model of the asymptotic properties.

From graphical plots, suitable models for analyzing the failure data can be differentiated. Consequently, a more complex estimation method, such as maximum likelihood estimation, one of the most frequently used statistical estimation methods, can be used together to find the parameter estimates, and the asymptotically inferences of MLE base on large sample theory can be applied to gain more insight of the failure data. The commonly used graphical methods include methods include: Weibull probability plot , Cumulative distribution Plot, Hazard rate plot [33] [34] [8] among others. A here is a simple demonstration of using a graphic plot in parameter estimation. A weibull Probability Plotting is used for brief explanation. Weibull probability plot was developed in the early 1970's, is a graphical method by sorting and transforming the observed data. It is a simple procedure for estimating parameters. The first step is to transform the CDF of the distribution in this case weibull distribution as;

$$Y = ln\{-R(t)\},\ and\, x = ln(t) \tag{2.19}$$

that will lead to an equation of a straight line from equation 2.19 thus, obtaining the relationship:

$$Y = \beta x + \beta\alpha \tag{2.20}$$

y which is a function of $F(x_i)$ is plotted against $X = (lnx_i)$ using the following procedure: Rank the data x_i in ascending order, estimate F(x) of the i^{th} rank and finally plot Y_i against X_i If the outliers are not present in the model, the plotting points tend to form a straight line. which means that the model well fits the data. [8] provides a detailed explanation of estimated value of β which is the slope of the line and estimated value of α.

In contrast, statistical methods are based on the theory and inference of mathematics and statistics. It is generally applicable to different data types and various model formulations. Below is a demonstration of how graphical method is used in parameter estimation.

The estimators obtained from statistical methods are considered to be more accurate than graphical estimators, and the asymptotic properties of the estimators are well developed. However, most of statistical methods equations that are difficult and more complex to solve. Therefore the knowledge of statistics is required. Fortunately, statistical estimates could easily be obtained from various statistical softwares and programming tools. Methods of moments, Maximum likelihood estimation, interval estimation among others are the some of the commonly used statistical methods of parameter estimation. Their properties enable one to identify which method would be suitable in a particular data. In this study, maximum likelihood Estimation (MLE) technique is used.

2.5.3.1 *Maximum likelihood estimation of lifetimes*

In this section we discuss how to compute the MLEs of the model parameters based on the both weibull and log-normal distributions discussed earlier in this section. Maximum likelihood (ML) methods are commonly used for various types of data sets including lifetimes data.The MLE performs very well in every simulated parameter combination. It is the most dependable among all the estimators. Some characteristics of maximum likelihood estimation include asymptotically unbaised, that is the bias tends to zero as the sample size n increases; asymptotically efficient in that they achieve the Cramer-Rao lower bound as n approaches infinity and they are asymptotically normal among others. Maximum likelihood method easily accounts for censored data and non-normal data therefore it is are widely used to provide parameter estimates and other interesting values such the confidence interval (CI) of a lifetime models.

Maximum likelihood estimate (MLE) can be described as follows: Given the model and its parameters, the MLE function is the probability (density) of a sample data seen as a function of the model parameters. Estimates of the model parameters are obtained by maximizing logarithm of the likelihood. Estimates are the values that maximizes probability of the sample data. Its main aim is to find combination of parameters β and η that maximize the probability of a given data. MLE estimates tend to predict long life with small samples.

2.5.3.2 *Maximum likelihood estimation for weibull distribution*

To estimate parameters η and β, assume $\theta = (\beta, \eta)$. Let t_1, t_2,....t_n such that t_i is Weib(η, β) be a random sample of size n drawn from $f(t, \theta)$ where θ is unknown.

$$f(\eta, \beta, (t)) = \left\{ \frac{\beta}{\eta} \left(\frac{t}{\eta} \right)^{\beta - 1} exp - \left(\frac{t}{\eta} \right)^{\beta} \right\}, t >= 0. \tag{2.21}$$

$$L = \prod_{i=1}^{n} ft_i(t_i, \theta) \tag{2.22}$$

$$L(t_1...t_n, \beta, \eta) = \prod_{i=1}^{n} f(\eta, \beta t_i) = \prod_{i=1}^{n} \frac{\beta}{\eta} \left(\frac{t_i}{\eta} \right)^{\beta - 1} exp - \left(\frac{t}{\eta} \right)^{\beta} \tag{2.23}$$

$$\frac{dlnL}{d\beta} = \frac{n}{\eta} + \sum_{i=1}^{n} lnt_i - \frac{1}{\eta} \sum_{i=1}^{n} t^{\beta} lnt_i = 0 \tag{2.24}$$

$$\frac{dlnL}{d\eta} = \frac{-n}{\eta} + \frac{1}{\eta^2}\sum_{i=1}^{n} t^{\beta} = 0 \tag{2.25}$$

Joint density function that would describe the likelihood function for weibull parameters for n failed items is given by:

$$\Pi\left(\frac{\beta}{\eta}\right)^{\beta-1} exp\left[-\frac{t_i}{\eta}\right]^{\beta} \tag{2.26}$$

Suppose t_1,t_2,....t_n are samples consisting of both censored and complete data, the ML for weibull estimates will be:

$$L(\eta;\beta/t) = C\prod_{i=1}^{n}\left[f\left(t_i/\eta,\beta\right)\right]^{r_i}\left[1-F\left(t_i/\eta;\beta\right)\right]^{1-r_i} \tag{2.27}$$

$$= \prod_{i=1}^{n}\left[\frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} exp\left(-\left(\frac{t}{\eta}\right)^{\beta}\right)\right]^{r}\left[exp\left(-\left(\frac{t}{\eta}\right)^{\beta}\right)\right]^{1-r_i} \tag{2.28}$$

Where r_i= 1 means failure, r_i= 0 is censored observation and C is constant independent of the model parameters generally =1 The simplified likelihood equation is:

$$L(\eta;\beta/t) = \sum_{i=1}^{n} r_i log\left(\frac{\beta}{\eta}\left(\frac{t_i}{\eta}\right)^{\beta-1}\right) - \sum_{i=1}^{n}\left(\frac{t_i}{\eta}\right)^{\beta} \tag{2.29}$$

$$\frac{\sum_{i=1}^{n} t^{\beta} log(t_i)}{\sum_{i=1}^{n} t_i^{\beta}} - \frac{1}{\beta} - \frac{1}{r}\sum_{i=1}^{n} r_i log(t_i) = 0 \tag{2.30}$$

simplifying further;

$$\eta^{\beta} - \frac{1}{r}\sum_{i=1}^{n} t_i^{\beta} = 0 \tag{2.31}$$

$$\eta = \left(\frac{1}{r}\sum_{i=1}^{n} t_i^{\beta}\right)^{\frac{1}{\beta}} \tag{2.32}$$

β must be known.

2.5.3.3 ***Likelihood ratio test***

Likelihood ratio tests are widely applicable tests related to maximum likelihood estimation. The LRT is defined by a ratio of the likelihood under null hypothesis to the likelihood under the alternative hypothesis. The test works by setting a cut-off value for the ratio between the two likelihoods. If the ratio is less than that cutoff then the test rejects the null hypothesis.

It is defined as:

$$LR(effect\,k) = -2ln\frac{L(effect\,k\,removed)}{L(full\,model)} \tag{2.33}$$

Where L() is the likelihood value. Likelihood ratio (LR) follows a chi-square distribution if the effect k is not significant. Removing effect k from the model of equation 2.33 will not have effect on the likelihood value. A very large likelihood value means that the effect of the parameter k is significant [35]. A detailed explanation on the methods of estimation, testing and function for these computation is explained by [35]

2.5.3.4 *Maximum likelihood estimation for log-normal distribution*

Assuming that the lifetimes of a product random variable T follows a log-normal distribution with the PDF as shown in equation 2.34 , to compute the Maximum likelihood estimators, we start with the likelihood function. The likelihood function of the log-normal distribution for: $t_i s$ (i=1,2,3...n) is determined by taking the product of probability density of individual lifetimes (time to fail) t_is

$$f(t;\mu;\sigma) = \frac{1}{\sqrt{2\pi\sigma t}}exp - \frac{(logt-\mu)^2}{2\sigma^2}; t > 0; -\infty < \mu < \infty; \sigma > 0 \tag{2.34}$$

Taking the natural logarithm ;

$$L(\mu;\sigma/T) = \prod_{i=1}^{n}\left[f\left(T_i/\mu,\sigma\right)\right]$$

$$= \prod_{i=1}^{n}\left(\left(2\pi\sigma^2\right)^{-\frac{1}{2}}T_i^{-1}exp\left[\frac{-(log(T_i)-\mu)^2}{2\sigma^2}\right]\right)$$

$$= ln\left(\left(2\pi\sigma^2\right)^{\frac{-n}{2}}\prod_{i=1}^{n}T_i^{-1}exp\left[\sum_{i=1}^{n}\frac{-ln(T_i)-\mu)^2}{2\sigma^2}\right]\right)$$

$$= -\frac{n}{2}ln(2\pi\sigma^2) - \sum_{i=1}^{n}ln(T_i) - \frac{\sum_{i=1}^{n}\left(In(T_i)-\mu\right)^2}{2\sigma^2}$$

$$= -\frac{n}{2}(2\pi\sigma^2) - \sum_{i=1}^{n}In(T_i) - \frac{\sum_{i=1}^{n}\left[In(T_i)^2 - 2In(T_i)\mu + \mu^2\right]}{2\sigma^2}$$

$$= -\frac{n}{2}In(2\pi\sigma^2) - \sum_{i=1}^{n} In(T_i) - \frac{\sum_{i=1}^{n} In(T_i)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n} 2In(T_i)\mu}{2\sigma^2} - \frac{\sum_{i=1}^{n} \mu^2}{2\sigma^2}$$

$$= -\frac{n}{2}In(2\pi\sigma^2) - \sum_{i=1}^{n} In(T_i) - \frac{\sum_{i=1}^{n} In(T_i)^2}{2\sigma^2} + \frac{\sum_{i=1}^{n} In(T_i)\mu}{\sigma^2} - \frac{n\mu^2}{2\sigma^2} \tag{2.35}$$

To find the values of parameters that maximizes $L(\mu;\sigma/T)$, we take the derivative of equation 2.35 with respect to μ and σ^2 and set it equal to 0. Thus the likelihood estimators are:

$$\mu^* = \frac{\sum_{i=1}^{n} In(T_i)}{n} \tag{2.36}$$

and

$$\sigma^* = \frac{\sum_{i=1}^{n} \left(In(T_i) - \frac{\sum_{i=1}^{n} In(X_i)^2}{n}\right)}{n} \tag{2.37}$$

2.6 Fan motor of a refrigerator

Generally, household appliances have been the fastest growing end product markets for electronic motor drives [36]. The demand for such motor drives is expected to more than double in the next five years. However, industries have desirable specifications of products to improve the quality of its products and processes. Currently, manufactures are favoring the use of specific motors for refrigerator because of their efficiency, low power density and low cost. Accurate prediction of parameters at the design stage is clearly critical since inappropriate design may lead to excessive noise emission and catastrophic failure. Recently, in addition to electric efficiency, sound comfort has become equally important to consumers. An acceptable noise level is becoming a major marketing point for many products. The main sources of noise of a household refrigerator are compressors and fans. The role of a fan is to blow the refrigerated air to a freezer chamber and cool chamber of refrigerators.

In this section we discuss the general structure of a fan motor and main requirements it has to fulfill in its selection.

2.6.1. Fan structure

A rotating fan and electric motor are the key elements of a fan as shown in figure 2.12. The electric motor is composed of a stator situated in the fan housing and a rotor, while the blades are attached on the rotor. The structure when disassembled is shown in figure

2.13 and figure 2.14. When the fan is in use, the magnet in the rotor interacts with force generated by the stator causing the rotor to rotate [9].

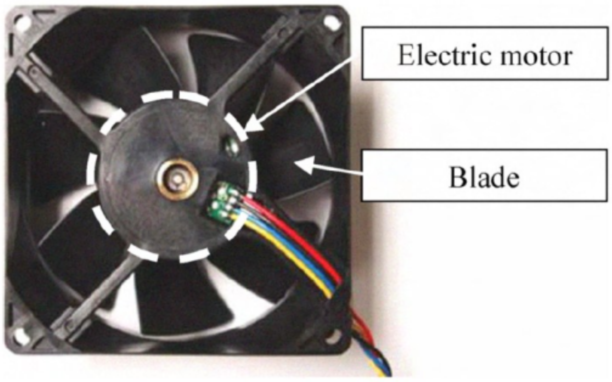

Figure 2.12. *The General Structure of a Condenser Fan [9]*

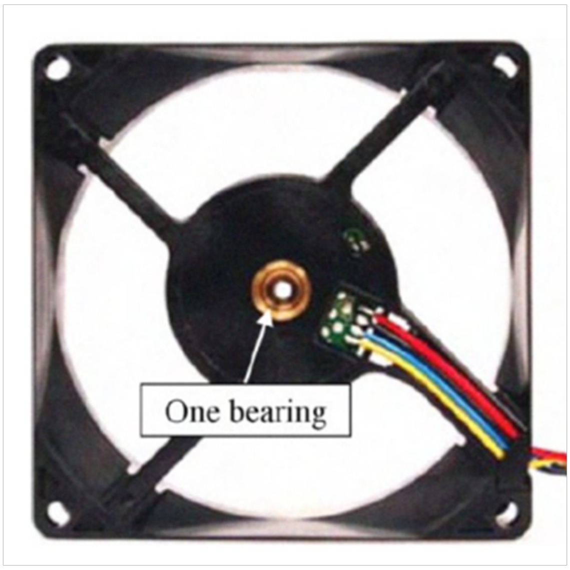

Figure 2.13. *(a) Stator in a Fan Housing [9]*

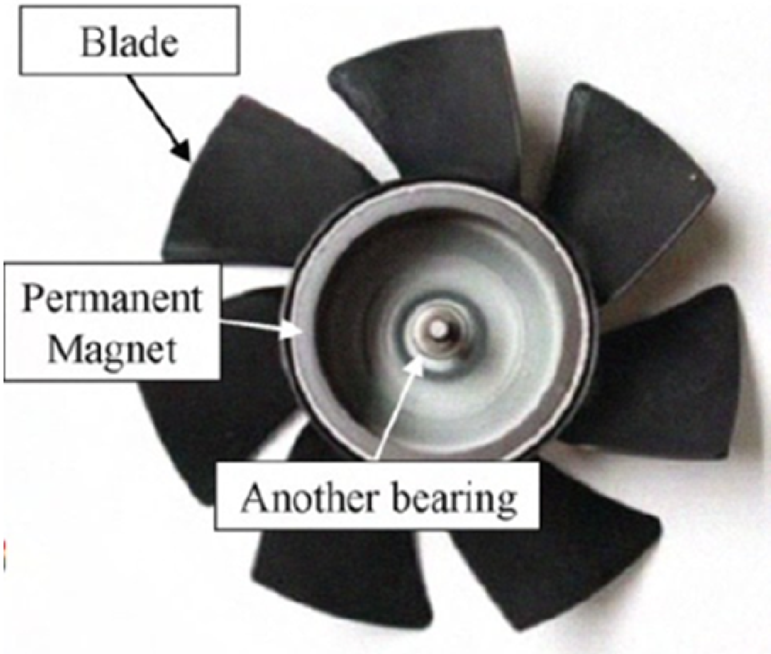

Figure 2.14. *(b) Rotor with Blades [9]*

A fan motor is located next to the compressor. It moves air across the condenser coils to help cool the hot refrigerant coming out the compressor. Refrigerator fans serve several purposes: Cooling the compressor, circulate cold air from the evaporator to the freezer and the cold chamber, and finally making cakes of ice for the external ice dispenser [37].

Fan and motor breakdown contributes nearly 30% of refrigerator failure [4].

3. EXPERIMENTS FOR LIFETIMES DATA

3.1 Transfer Function Identification and Optimization Using Response Surface Methodology

The Response surface methodology (RSM) was introduced by Box and Wilson [38] to emphasize on the advantages associated with it's sequential and immediacy nature. RSM is a collection of statistical and mathematical techniques useful for developing, improving and optimizing processes. The methodology summarized by the combination of design of experiment, fitting the model and optimization processes. It also has important applications in the design, development and formulation of new products as well as in the improvement of existing product design. The method is extensively used in the industrial sector especially when several input variables influence performance of a product or process.

In a standard RSM, the important factors are studied with an aim of eliminating unimportant factors to the response (screening process) in order to produce efficient results with few runs. Method of steepest ascend is applied to find the optimum response. Next, a second-order experiment is run to fully characterize the response surface. The use of sequential experimentation process lends it well to industrial applications which focus on regression model (a response surface) as the center of analysis . This sequential nature allows industrial researchers to take advantage of significant cost savings, especially when little is known about the nature of the response surface [39]. The RSM utilizes Taylor series approximations to find a parametric model for response prediction over a finite experimental region as its primary goal. A two factor levels and fractional factorial designs are the most commonly used designs in the support of first-order model, model including main effect and interaction.

The relationship between the input variables ξ_1, ξ_2....ξ_k that affects the response y of a product or a system of is as follows:

$$y_i = f(\xi_1, \xi_2, ...\xi_k) + \epsilon \tag{3.1}$$

Where ϵ represents the error observed in the response y. This also represents other sources of variability not accountable for in the function. Thus ξ accounts for effects of measurement error on the response, other sources of variation that are inherent in the process or system and the effect of other variables. Assuming ϵ is normally distributed with mean zero, the expected response is denoted by $E(y) = f(\xi_1, \xi_2, ...\xi_k) = \eta$, then surface represented by η is the response surface.

The variables represented by ξ_1, ξ_2,....ξ_k are normally referred as natural variables because they are expressed in the natural units of measurement [40].These natural variables are always converted into coded vaiables X_1, X_2....X_k. When the response is modeled by a linear function [41] of independent variables as shown in equation 3.2, then the function is first-order model.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ...\beta_k X_k + \epsilon \tag{3.2}$$

3.2 Two Level Fractional Factorial Design for Variable Screening

A complete factorial designs can become large even at two levels of each factors [42]. An experiment with eight factors would require $2^5 = 32$ runs. In the case of reliability experiment, each run may be measured as in hours, days, months etc and may take several thousand hours, which would be practically infeasible to conduct 32 runs. Thus, more efficient methods of conducting experiments is needed.

Fractional factorial designs can reduce the number of runs significantly by choosing a subset or fraction of the complete factorial design. A 2^{5-2} factorial design would reduce the number of runs by $2^{-2} = 1/4$ (2^{5-2}) to eight runs. Although this provides an advantage of reducing the number of runs significantly, the disadvantage is that many of the effects are hidden or confounded by the main effect factors that the experimenter deems the most important. This assumes that many of the confounded effects are not significant and have very minimal effect or do not affect the response at all. Fractional factorial designs are particularly useful if it can be estimated which main effects and interactions are significant so that the remaining effects can be confounded [42].

3.3 Experiments for Improving Lifetimes of a Product

Manufacturing industries have made significant efforts to consider reliability prediction in early phases of a project. The application of build-in reliability and design for reliability have theoretically influenced the manufacturing industries to make significant efforts to consider reliability prediction in the earlier phases of a project. For instance, Designers and engineers have extensively used computer simulation(Computer supported information generating method and analysis). The aim of reliability simulation is to help the designer achieve the reliability equipment while minimizing the resources used.

The lifetimes of a product could be affected by many potential parameters (factors). However, some of these factors could be more important than others .i.e, they have greater influence on the lifetimes of a product as their values are changed. Experiments can be used to identify these important factors by deliberately changing their values and examining the lifetime behavior of that particular product. Other than identifying these important factors, the values of these important factors that yield lifetime gains can be recommended. A well designed experiment provides a systematic and an efficient plan to achieve these goals. Several factors can be studied by using as few resources as possible. Design of experiments have been used widely to improve quality characteristics of products and therefore can be adopted for the lifetime improvement hence reliability.

Some terminology used to describe the plan will be helpful. The plan of experimentation is referred to as experimental design or design. The design consist of a list of runs where a run is a combinations of values (levels) at which factors in the experiment are set. Run size is the number of runs in the experimental design. Treatment factors does not suggest drugs as in medical experiment, it means items whose effect on the response is to be studied. For example, this could be temperature, humidity or teaching method. While levels are the specific amounts of treatment factors to be used in the experiment, such as selected temperature settings in the range of interest. Experimental units are the "material" to which the levels of treatment factor(s) are applied [43]. For example, in agriculture these would be different plots of land. These units should be representative of both the material and conditions to which the conclusions will be applied.

The experiment involves setting units(fan motor) according to the specified run conditions by the design and these factors are observed until failure occurs. The two parameter β and η weibull were calculated using likelihood estimation method as the most preferred method [25] for censored data and small sample size.

3.3.1. Choice of the factors

In service conditions, the failure process of an electrical component is driven by several working conditions acting together such as dirt, moisture, temperature, aggressive chemicals and radiation among others. In rotating machines, over-voltage occur at the motor terminals and it an lead is not homogeneously distributed to the whole component. The reduction or over-voltage can lead to partial discharges occurring between the components of the product. In this study only two parameters are chosen to determine their corresponding influence on the lifetimes, which are:

- Voltage (V)
- Temperature (C)

In this study these parameters will be called factors. The factors will be studied in levels of high and low.

3.3.2. Simulation assumptions

The assumptions of the model are:

- Regression model (a response surface) as the basis of analysis
- The experiment involves m treatment combinations $X_i = (xa, xb, xc......Xm)$, each combination has n items. the X_i's are independent with specific probability distribution.
- failure time t follows both weilbull and log-normal distributions.
- ϵ_{ij} are independent and identically distributed.
- Each component is assumed to have a lifetime X_i and a fixed censoring time C

For weibull distribution, scale parameter denoted by η is known as the "life characteristic" and β is the shape parameter. Generally the shape parameter of weibull distribution is given as β. We have reserved β for the shape parameter as generally known and used α for the parameters that relate the prdedictor variables to the expected lifetime($\alpha_0, \alpha_1......\alpha_n$).

For an L-level factor, it will have L-1 independent effect if the zero sum constant is used. The life characteristics can be treated as a function of these independent effects. The scale parameter which is the life characteristic is assumed to be a function of effects, the link function between life and the two factors effect with interaction for weibull distribution is expressed as:

$$1n(\eta) = \alpha_0 + \alpha_1 Temperature + \alpha_2 Voltage + \alpha_{12}(Temperature)(Voltage)$$

$$\eta = exp(\alpha_0 + \alpha_1 Temperature + \alpha_2 Voltage + \alpha_{12}(Temperature)(Voltage)) \quad (3.3)$$

we assume the first order model,

$$T_i/X_i = Weibull\left(\eta = exp\left(\alpha_0 + \alpha_1 + \alpha_2\right), \beta\right)$$

where temperature and voltage are the two independent effects of materials. The model that involve interaction is not be considered. If it is level 1, temperature = 1 and voltage = 0; if it is level 2, temperature = 0 and voltage = 1; if it is level 3, temperature = -1 and voltage =-1. Therefore, for each level, there is a different life characteristic value. Using this link function, the log likelihood function for the failure data at level i will be:

$$ln(LKV_I) = \sum_{j=1}^{F}\left[1n\beta - \beta ln\eta_i + (\beta - 1)1nt_j - \left(\frac{t}{\eta_i}\right)^{\beta}\right] - \sum_{j=1}^{S}\left(\frac{t_j}{\eta_i}\right)^{\beta} \tag{3.4}$$

In this case F is the failure time and S is suspension time. The overall likelihood function will be:

$$Ln(LKV) = \sum_{i=1}^{L} 1n(LKV_i) \tag{3.5}$$

Where t_{ij} are the life times,$X_i = I(x_a, x_b, ...x_n)$ the $\{X_i\}$ are the corresponding vector covariate values, β or of the parameter. ϵ is the vector of the parameters and σ is the scale parameter. ϵ_{ij} are iid (Independent and Identical distribution) whose survival (S.F) and PDF are $exp(exp(-w))$ and $exp(w - (exp(w)))$ respectively. Lawless [8] expressed the experimental factors as a linear function of the log lifetime.

3.3.3. Algorithm characterization

The following notations are used in the characterization of the lifetime data.

The exact failure time will be known if and only if X is less than or equal to C. When X is greater than C, the units become suspensions with the lifetime data censored at time C. The simulation was replicated 2000 times.

4. APPLICATION AND RESULTS

4.1 Simulation Study

In this section lifetimes of a component is modeled by simulation based on the two input factors which includes environmental load such as temperature and product operational load, voltage. Lifetime experiments can also be used to address the root cause of failure and accelerating forces responsible for product failures [44]

4.1.1. Weibull simulation study

In order to establish the data matrix, a two levels (low=-1 and high=1 value) for each factor is used in the study. The levels are identified with the real conditions as follows: Voltage (105 , 120) and temperature (38 , 43) for the lower and higher ones respectively. The design supports the first order model with no interaction. To include censoring, the values of first q failures per run resulting in failure or censoring are recorded. The failure times of each tests strand were simulated using a weibull distribution with an assumed shape parameter $\beta = 3$ and model representing main effects, scale parameter as:

$$\mu = log(\eta_i) = \alpha_0 + \alpha_1(Temp_i) + \alpha_2(Volt_i) \tag{4.1}$$

Where $Temp_i$ and $Volt_i$ are the temperature and voltage respectively of 2^2 factorial design.

In this study, two level, two factorial experiment design is adopted to evaluate the effects of independent variables (temperature and voltage) on the response (lifetimes). Equation 4.2, a quadratic (2nd- order) model used in order to optimize the response with few experimental runs.

$$Y = \alpha_0 + \alpha_1 X_A + \alpha_2 X_B + \alpha_3 X_A X_B + \alpha_4 X_A^2 + \alpha_5 X_B^2 \tag{4.2}$$

where Y is the response; α_0 to α_5 are the regression coefficients and temperature and voltage are the factors. This model is used to estimate the relationship between lifetimes Y, and the two independent factors temperature, and voltage,. The values, α_1 and α_2 coefficients denote the main effect of factors temperature and voltage, respectively. Besides, α_3 denotes the interaction between factors temperature, voltage; Finally, α_4 and α_5 denote the quadratic effect of factors temperature and voltage, respectively.

The factor levels are coded for low and high settings as -1 and 1 respectively. Table 4.1 shows the coded values of main effect.

Table 4.1. *Factors and factor levels used in the design*

Factor	Low (-1)	High (1)
Temp	38	43
Volt	105	120

Table 4.2. *Layout of* 2^2 *design matrix*

Run	I	A	B	AB	Response
1	1	+	−	−	R_1
2	1	−	+	−	R_2
3	1	+	+	+	R_3
4	1	−	−	+	R_4

In order to include the factor levels for temperature and voltage, a model that represents main effects and interaction specified in equation 4.1 is used. The equivalence of equation 4.1 above can be written as:

$$R_1 = \alpha_0 + \alpha_1(1) + \alpha_2(-1) + \alpha_3(1)(-1)$$

$$R_2 = \alpha_0 + \alpha_1(-1) + \alpha_2(1) + \alpha_3(-1)(1)$$

$$R_3 = \alpha_0 + \alpha_1(1) + \alpha_2(1) + \alpha_3(1)(1)$$

$$R_4 = \alpha_0 + \alpha_1(-1) + \alpha_2(-1) + \alpha_3(-1)(-1)$$

- Factor levels were used to determine scale parameters for weibull and log-normal distributions.

- Failure are generated using the modeled scale parameters and assumed shape parameter independently.
- Censoring is included in the generated time to fail.
- MLE applied to find the the estimates of the parameters.
- 2000 replications were carried out. The assumed model contains temperature and voltage with no interaction.

The results from using survreg in the survival package of R statistical program version 3.4.0 was used in the analysis. Table 4.3 and table 4.4 shows the simulation results for weibull distribution and log-normal distributions respectively.

Table 4.3. *Weibull distribution analysis results*

Likelihood ratio test-table					
Model	**Effect**	**DF**	**Ln (Likelihood Value)**	**Likelihood Value**	**P-Value**
Reduced	Temp	1	-1378.3	6.884	8E-06
Reduced	Volt	1	-1281	-	6.4-E04
Full	-	4	-1377		
Estimate of Parameters					
Predictor	**Coef**	**S.d Error**	**P-value**		
Intercept	4.8139	1.647	0.068		
Temperature	0.0304	0.013	1.15E-02		
Voltage	-0.0466	0.257	7.17E-03		

4.1.2. Log-normal simulation study

Table 4.4. *Log-normal distribution analysis results*

Likelihood ratio test-table					
Model	**Effect**	**DF**	**Ln (Likelihood Value)**	**Likelihood Value**	**P-Value**
Reduced	Temp	1	-1425	6.884	0
Reduced	Volt	1	-1654	-	0.021
Full	-	4	-2144		
Estimate of Parameters					
Predictor	**Coef**	**S.d Error**	**P-value**		
Intercept	5.0753	0.0432	6.71E-3		
Temperature	0.0998	0.077	2.309E-02		
Voltage	-0.1006	0.0266	1.58E-04		

4.2 Discussion

The two factor, two levels factorial design was used to generate lifetime data by monte-carlo simulation method. In reliability Design of experiment, The scale parameter, η is assumed to be a function of the factors which is expressed in this case as:

$$ln(\eta_i)=\alpha_0+\alpha_1(Temp_i) +\alpha_2(Volt_i)$$

$$\eta_i= \exp\left(\alpha_0+\alpha_1(Temp_i) +\alpha_2(Volt_i)\right)$$

Where α_0 is the intercept, α_1 is the effect coefficient of temperature and α_2 is the effect coefficient of voltage. Previous studies indicate that interaction does not affect the life of products therefore it was not included in the model. The simulated data is censored with many failures. MLE was used to generate parameters of each distribution, However, it is not suitable to apply ML method to estimate both the scale and the shape parameters

[45]. Therefore an assumed shape parameter is suitable [46]. An assumed Shape parameter was used which was later estimated as β=1.85, indicating failure due to product wear-out and therefore as β increases number of failure times also increases [1].

The results from weibull distribution in table 4.3 show that temperature and voltage are significant at 0.05 level of significance . From table 4.3 and table 4.4, both temperature and voltage are significant at 0.05 level of significance. It is appropriate to apply standard regression techniques to Lifetimes experiments where the exact failure times are observed [47]. However, when some of the data are censored i.e all failure times are not observed, application of regression technique is inadequate [35] [47]. In this case, the likelihood based approach is considered. Unlike ANOVA whose P values are obtained from F ratio test, in experiments involving lifetimes, the p values are based on the likelihood ratio test [48]. The ML estimates of the effect coefficients corresponding to temperature and voltage are 0.0304 and - 0.0466 for weibull distribution. Based on these coefficients, the appropriate settings for these effects to optimize the response are:

- To set temperature at a higher level of 1 since its coefficient is positive.
- To set voltage at lower level -1 since its coefficient is negative.

This is suggested by Bajzer, Therneau, Sharp, and Prendergast [49] who also used maximum likelihood method as a tool to determine failures. Using the ML method estimate of coefficients from the weibull distribution the scale parameter is predicted as;

$$\ln \eta = 4.8139 + 0.0304(Temp) - 0.0466(Volt)$$
$$= 4.8139 + 0.0304(1) - 0.0466(-1)$$
$$\ln\eta = 4.89, \eta = exp(4.89)$$
$$\eta = 133$$

let lifetimes be denoted by Time then,

Time ~ Weibull (β = 1.85, η = 133).

In the log-normal distribution temperature and voltage also have the similar effect on the lifetimes as that of weibull distribution. Table 4.4 shows the log-normal simulation results. The log-normal distribution is also considered as a useful and flexible model for reliability analysis [3] like weibull distribution, log-normal also consist of scale (median T_{50}) parameter and shape parameter. The results from table 4.4 shows that temperature and voltage are significant and therefore they are useful in the lifetimes of a component. The ML estimates of the coefficients based on the simulation are 0.0998 for temperature and -0.1006 for voltage. Based on these results, temperature is set at a high level and voltage at low level to optimize the response.

Log-normal distribution is considered to be successful in modeling failures due to chemical reactions such as corrosion,material movement because of diffusion and it is with this reason that it is considered as the most popular distribution [3].

5. CONCLUSIONS

Instant results and sequentiality make agricultural experiments different from industrial experiments. For most industrial experiments results are always available instantly (days, hours e.t.c) and the results from each group can be acted upon to be used in the next experiments while in agricultural experiments, processes are always restricted during growing seasons. In addition, normal distribution which characterizes most experimental designs is not a logical distribution for lifetimes due to censoring.

Due to censoring, analysis of variance (ANOVA) and (least square method (LSM) can not be used to improve reliability [27]. In this case, ANOVA method can only be applicable if suspensions are treated as failures and midpoint of interval data used as failure times. This approximation give wrong results and lead to wrong conclusions [50]. Consequently, the use of ANOVA method on lifetimes data violates the normal distribution assumption among others [26].

In addition, the commonly used ML estimate approach is considered to have an "estimablity" problem. Testing for important effects in the model cannot be done by comparing the ML estimates with their corresponding standard errors because the ML estimates may be infinite see [47]. Consider an example given by [47] of a one factor two levels experiment with censored observation and failure time data given according to the factor levels. In this example it is observed that as the parameter tends toward infinity the likelihood function increase to the maximum and therefore concluded that ML estimate for the main effect tends to be infinite when the true factor effect is large. Further explanation of the ML estimate can be found in [51]. Thus, the use of Likelihood ratio test is suggested by [52].

The study aims at increasing the lifetimes of a product by simulation. The important factors are identified through the screening process and a second order (quadratic) response surface methodology used to determine the optimal values. Weibull and log-normal models with an assumed scale parameter was used in the model simulation and maximum likelihood estimation for the parameter estimation. Temperature and voltage were found to be significant and therefore they play an important role in the lifetimes of a fan motor.

From log-likelihood values, weibull distribution provides larger likelihood values as compared to the log-normal. Once the researcher decides the value of shape parameter and estimate of the parameters, it is better to use weibull distribution to model lifetimes. In the contrary, [3] argues that the choice of distribution to model lifetimes depends greatly on the theoretical justification of the model based on the failure patterns or mechanisms under investigation.

In both weibul distribution and log-normal distribution, a constant scale parameter was assumed but it was not tested whether it is constant or not. For future studies, the assumption of constant scale parameter can also be tested and consequences of assuming a constant scale parameter be investigated. In addition, a quadratic model(second order model) can be used to optimize. The median value of the weibull distribution can also be used because reliability data is usually skewed. Other than 2^2 factorial design, it would also be interesting to use more factors in the design formation since these are not the only factors thats have influence the lifetimes of a fan motor.

REFERENCES

[1] L. J. Freeman and G. G. Vining, "Reliability data analysis for life test designed experiments with sub-sampling," *Quality and Reliability Engineering International*, vol. 29, no. 4, 2013.

[2] J. I. McCool, *Using the Weibull distribution: reliability, modeling and inference.* John Wiley & Sons, 2012, vol. 950.

[3] P. A. Tobias and D. Trindade, *Applied reliability.* CRC Press New York, 2011.

[4] J. Yan-Qiao and W. Shi-Liang, "Statistical analysis of reliability of container refrigeration units," *International journal of refrigeration*, vol. 19, no. 6, pp. 407–413, 1996.

[5] M. Crowder, A. Kimber, R. Smith, and T. Sweeting, *Statistical concepts in reliability.* Springer US, 1991.

[6] J. Deshpande and S. G. Purohit, *Life time Data: Statistical Models and methods*, 1st ed. World Scientific publishing, 2005.

[7] L. J. Freeman and G. G. Vining, "Reliability data analysis for designed experiments laura j. freeman and g. geoffrey vining department of statistics, virginia tech, blacksburg, va 24061-0439," no. 1, pp. 1–8, 2003.

[8] J. F. Lawless, *Statistical models and methods for lifetime data.* John Wiley & Sons, 2011, vol. 362.

[9] S. Lee, S. Heo, and C. Cheong, "Prediction and reduction of internal blade-passing frequency noise of the centrifugal fan in a refrigerator," *International Journal of Refrigeration*, vol. 33, no. 6, pp. 1129–1141, 2010.

[10] C. W. Zhang, T. Zhang, D. Xu, and M. Xie, "Analyzing highly censored reliability data without exact failure times: an efficient tool for practitioners," *Quality Engineering*, vol. 25, no. 4, pp. 392–400, 2013.

[11] J. L. K. Kensler, L. J. Freeman, and G. G. Vining, "A practitioner's guide to analyzing reliability experiments with random blocks and subsampling," *Quality Engineering*, vol. 26, no. 3, pp. 359–369, 2014.

[12] M. D. Beaudry, "Performance-related reliability measures for computing systems," *IEEE Transactions on Computers*, vol. 27, no. 6, pp. 540–547, 1978.

[13] Z. Yu, Z. Ren, J. Tao, and X. Chen, "Accelerated testing with multiple failure modes under several temperature conditions," *Mathematical Problems in Engineering*, vol. 2014, 2014.

[14] Z. Hu and X. Du, "Lifetime cost optimization with time-dependent reliability," *Engineering Optimization*, vol. 46, no. 10, pp. 1389–1410, 2014.

[15] S. E. Rigdon, B. R. Englert, I. a. Lawson, C. M. Borror, D. C. Montgomery, and R. Pan,

"Experiments for reliability achievement," *Quality Engineering*, vol. 25, no. 1, pp. 54–72, 2012.

[16] E. M. Benavides, "General model for reliability-based engineering design," *Communications in Statistics-Theory and Methods*, vol. 43, no. 10-12, pp. 2342–2356, 2014.

[17] M. A. Khan and H. M.-u. Islam, "Bayesian analysis of system availability with half-normal life time," *Quality Technology & Quantitative Management*, vol. 9, no. 2, pp. 203–209, 2012.

[18] I. Bazovsky, *Reliability theory and practice.* Courier Corporation, 2004.

[19] M. Soleimani, M. Pourgol-Mohammad, A. Rostami, and A. Ghanbari, "Design for reliability of complex system: Case study of horizontal drilling equipment with limited failure data," *Journal of Quality and Reliability Engineering*, vol. 2014, 2014.

[20] B.-J. Lwo, M.-S. Lin, K.-H. Huang *et al.*, "Tsv reliability model under various stress tests." IEEE, 2014, pp. 620–624.

[21] N. J. Salkind, *Encyclopedia of measurement and statistics.* Sage, 2007, vol. 1.

[22] S.-w. Woo and D. L. O. Neal, "Improving the reliability of a domestic refrigerator compressor subjected to repetitive loading," no. March, pp. 99–115, 2016.

[23] S. Ruhi, "Application of mixture models for analyzing reliability data: A case study," *OALib*, vol. 02, no. 09, pp. 1–8, 2015.

[24] A. Garmabaki, A. Ahmadi, J. Block, H. Pham, and U. Kumar, "A reliability decision framework for multiple repairable units," *Reliability Engineering & System Safety*, vol. 150, pp. 78–88, 2016.

[25] O. Borgia, F. De Carlo, N. Fanciullacci, and M. Tucci, "Accelerated life tests for new product qualification: a case study in the household appliance," *IFAC Proceedings Volumes*, vol. 46, no. 7, pp. 269–274, 2013.

[26] H. Guo and A. Mettas, "Design of experiments and data analysis," in *2010 Reliability and Maintainability Symposium, San Jose, CA, USA*, 2010.

[27] G. Wang, Z. He, L. Xue, Q. Cui, S. Lv, and P. Zhou, "Bootstrap analysis of designed experiments for reliability improvement with a non-constant scale parameter," *Reliability Engineering & System Safety*, vol. 160, pp. 114–121, 2017.

[28] R. B. Abernethy, "An overview of weibull analysis," *Abernethy RB, Florida, pp1-11*, 2002.

[29] P. Basak, I. Basak, and N. Balakrishnan, "Estimation for the three-parameter lognormal distribution based on progressively censored data," *Computational Statistics & Data Analysis*, vol. 53, no. 10, pp. 3580–3592, 2009.

[30] H. L. Harter and A. H. Moore, "Local-maximum-likelihood estimation of the parameters of three-parameter lognormal populations from complete and censored samples," *Journal of the American statistical association*, vol. 61, no. 315, pp. 842–851, 1966.

[31] A. Munro and R. Wixley, "Estimators based on order statistics of small samples from a three-parameter lognormal distribution," *Journal of the American Statistical Association*, vol. 65, no. 329, pp. 212–225, 1970.

[32] C. E. Antle, "Lognormal distribution," *Encyclopedia of statistical sciences*, 1985.

[33] W. B. Nelson, *Applied life data analysis.* John Wiley & Sons, 2005, vol. 577.

[34] J. H. Kao, "A graphical estimation of mixed weibull parameters in life-testing of electron tubes," *Technometrics*, vol. 1, no. 4, pp. 389–407, 1959.

[35] L. A. Escobar and W. Q. Meeker, "A review of accelerated test models," pp. 552–577, 2006.

[36] Y.-H. Kim, B.-S. Yang, and C.-J. Kim, "Noise source identification of small fan-bldc motor system for refrigerators," *International Journal of Rotating Machinery*, vol. 2006, 2006.

[37] T. He, C. Mei, and J. P. Longtin, "Thermosyphon-assisted cooling system for refrigeration applications," *International Journal of Refrigeration*, vol. 74, pp. 163–174, 2017.

[38] J. Box and W. Wilson, "Central composites design," *JR Stat Soc*, vol. 1, pp. 1–35, 1951.

[39] R. H. Meyers and D. C. Montgomery, "Response surface methodology," *Process and Product Optimisation Using Design Experiments, second ed, Willey, New York, NY*, 2002.

[40] R. Myer and D. C. Montgomery, "Response surface methodology: process and product optimization using designed experiment," *John Wiley and Sons, New York*, pp. 343–350, 2002.

[41] D. C. Montgomery, *Design and analysis of experiments.* John Wiley & Sons US, 2008.

[42] G. E. Box, J. S. Hunter, and W. G. Hunter, *Statistics for experimenters: design, innovation, and discovery.* Wiley-Interscience New York, 2005, vol. 2.

[43] A. M. Dean and D. Voss, *Design and Analysis of Experiments.* Springer New York, 1999, vol. 1.

[44] F. N. Nwobi and C. A. Ugomma, "A comparison of methods for the estimation of weibull distribution parameters," *Metodoloski zvezki*, vol. 11, no. 1, p. 65, 2014.

[45] C. W. Zhang, T. Zhang, D. Xu, and M. Xie, "Analyzing highly censored reliability data without exact failure times: an efficient tool for practitioners," *Quality Engineering*, vol. 25, no. 4, pp. 392–400, 2013.

[46] R. G. Miller Jr, *Survival analysis.* John Wiley & Sons, 2011, vol. 66.

[47] C. J. Wu and M. S. Hamada, *Experiments: planning, analysis, and optimization.* John Wiley & Sons, New York, 2011, vol. 552.

[48] C. J. Wu and M. Hamada, "Experiments planning analysis and parameter design optimization. jhon wiley and sons," *Inc., Singapore*, 2000.

[49] Ž. Bajzer, T. M. Therneau, J. C. Sharp, and F. G. Prendergast, "Maximum likelihood method for the analysis of time-resolved fluorescence decay curves," *European biophysics journal*, vol. 20, no. 5, pp. 247–262, 1991.

[50] H. Guo and A. Mettas, "Reliability assessment using a likelihood ratio test," *International Journal of Performability Engineering*, vol. 4, no. 2, pp. 196–198, 2008.

[51] M. J. Silvapulle and J. Burridge, "Existence of maximum likelihood estimates in regression models for grouped and ungrouped data," *Journal of the Royal Statistical*

Society. Series B (Methodological), pp. 100–106, 1986.

[52] D. B. Clarkson and R. I. Jennrich, "Computing extended maximum likelihood estimates for linear parameter models," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 417–426, 1991.