A REVIEW OF SOME LIFETIME DISTRIBUTIONS:

PROPERTIES AND APPLICATIONS

BY

TAWA FAITH ISHOLA

PSC2003883

DEPARTMENT OF STATISTICS

FACULTY OF PHYSICAL SCIENCES

UNIVERSITY OF BENIN

BENIN CITY

FEBRUARY, 2025

A REVIEW OF SOME LIFETIME DISTRIBUTIONS: PROPERTIES AND APPLICATIONS

BY

TAWA FAITH ISHOLA

PSC2003883

A PROJECT WRITTEN AND SUBMMITED TO THE DEPARTMENT OF STATISTICS, IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF BACHELOR OF SCIENCE (BSc.) DEGREE IN STATISTICS, FACULTY OF PHYSICAL SCIENCES, UNIVERSITY OF BENIN, BENIN CITY, NIGERIA

FEBRUARY 2025

UNDERTAKING

This project was carried out by **TAWA FAITH ISHOLA** with matriculation number **PSC2003883**. I have neither copied nor duplicated the work of any other author(s). All works used have been duly cited and acknowledged.

________________________________ ____________________________

ISHOLA TAWA FAITH **DATE**

CERTIFICATION

This is to certify that the project titled, "A Review of Some Lifetime Distributions: Properties and Applications" submitted by TAWA FAITH ISHOLA, MAT NO PSC2003883 in partial fulfillment of the requirements for the degree of B.SC , Bachelor of Science in Statistics , is a bonafide record of the work carried out under the supervision of MR C.O. ODIJIE, has satisfied the regulations governing the award of the Bachelor of science degree in the department of statistics, University of Benin, Benin City, Edo State, Nigeria.

SUPERVISOR: MR C.O. ODIJIE

SIGN & DATE

H.O.D: PROF. A. IDUSERI

SIGN & DATE

EXTERNAL EXAMINAL:

SIGN & DATE

DEDICATION

I dedicate this project work to Almighty God for his strength, guidance and knowledge through course of doing this project.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all those who contributed to the successful completion of this project. First and foremost, I extend my heartfelt thanks to my project supervisor, MR C.O. ODIJIE for his invaluable guidance, insightful feedback, and unwavering support throughout the course of this research. Their expertise and encouragement have been instrumental in shaping this work. I am also grateful to the faculty members of the Department of Statistics, who provided me with the necessary knowledge and resources to undertake this project. Special thanks to my colleagues and friends for their constructive discussions, motivation, and assistance during challenging phases of this work. I would like to acknowledge the National Physical Laboratory, England, and Birnbaum & Saunders for providing the datasets used in this study. Their contributions have been essential in enabling the practical application of the theoretical concepts discussed in this project.

Lastly, I extend my sincere appreciation to my family for their constant encouragement, patience, and belief in my abilities. Their support has been a source of strength and inspiration throughout this journey.

TABLE OF CONTENTS

ABSTRACT

Lifetime distributions are fundamental tools in statistics and probability theory, widely used to model the time until an event of interest occurs, such as the failure of a system, the death of an organism, or the occurrence of a financial default. This study focuses on reviewing three key lifetime distributions: The Exponential, Weibull, and Gamma distributions with an emphasis on their mathematical formulations, properties, and practical applications. The Exponential distribution, characterized by its memoryless property and constant hazard rate, is simple but limited in its applicability to more complex scenarios. The Weibull distribution, with its flexibility in modeling increasing, decreasing, or constant hazard rates, is highly versatile and widely used in reliability engineering. The Gamma distribution, capable of modeling multi-phase processes, is particularly useful for analyzing skewed data and complex lifetime behaviors. This study contributes to the theoretical understanding of lifetime distributions and provides practical guidelines for their application in fields such as reliability engineering, survival analysis, and risk assessment. Recommendations for future research include exploring additional distributions, analyzing larger datasets, and applying advanced estimation methods to further enhance the accuracy and applicability of lifetime distribution models.

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND OF THE STUDY

The study of failure and survival pattern mechanism underlying the time-to-event phenomena is crucial in various fields, including engineering, healthcare and social sciences. The analysis of time dependent data has become increasingly important in predicting failure rate, optimizing system performance and improving decision making. Statistical models that capture the probability of failure or survival over time are essential tool in this endeavor. The lifetime data analysis Land scape has transformed dramatically since 1970 in methodology theory and fields of application. Lifetime distribution also known as survival distribution model the time - to -failure or time-to-event phenomena. They describe the probability distribution of the time until a specific event occurs, such as equipment failure, component failure, or patient survival. In modeling lifetime data, probability distributions are often used, because they offer insight into the nature of various parameters and function. Examples of widely used distributions include Exponential, Weibull, Gamma distributions, as well as several others.

1.2 PROBLEM STATEMENT

In recent years, the increasing complexity of systems and processes has highlighted the need for reliable and accurate lifetime distribution models. Advances in data collection and analysis have also created opportunities for exploring new lifetime distribution models and application. However, despite the importance and abundance of the distributions, lifetime distribution face challenges including limited understanding of their properties and inadequate application and identifying a suitable distribution for a specific application. The Exponential, Weibull, and Gamma distributions are commonly used, but each has specific assumptions and properties that may or may not align with the characteristics of the data. For example:

- The Exponential distribution assumes a constant hazard rate, which may not be suitable for systems with aging components or increasing failure rates.

- The Weibull distribution offers flexibility in modeling varying hazard rates but requires careful parameter estimation.

- The Gamma distribution is useful for modeling multi-phase processes but can be computationally intensive and less intuitive to interpret.

This study of time-to- event data has gained significant attention in recent years due to its application in; Reliability engineering: predicting component failure times and optimizing maintenance schedule. Survival analysis: analyzing patient survival rate disease progression and treatment efficacy. Biomedical research: understanding disease mechanism, identifying risk factor and develop prognostic models.

1.3 AIM AND OBJECTIVES OF THE STUDY

The purpose of this study is to provide a comprehensive review of lifetime. The objectives include to:

- examine their statistical properties such as probability density function, cumulative distribution function, survival function and hazard function.
- explore the properties of each distribution such as their scale, shape and moment, and how these properties influence their applicability. Since each distribution has unique characteristics that makes it suitable for specific application.
- explore the practical applications of these distributions in various fields such as reliability engineering, medical survival analysis and actuarial science.

- provide guidelines for selecting the appropriate lifetime distributions for a given dataset.

The statistical analysis of lifetime distribution provides insight into how different systems behave over time and under varying conditions. By the distribution of lifetimes, we can make informed predictions. optimize system performance and guide decision-making in risk management.

In this project work, we study some lifetime distributions that are used in modeling survival times. Studies include the behavior of density and hazard function for many class of pdf. We also look at their various plots at different level to see their relationship and differences.

1.4 DEFINITION OF TERMS

CUMULATIVE DESTRIBUTION FUNCTION (CDF): The cumulative distribution function, gives the probability that a random variable is less than or equal to a specific value.

MOMENT GENERATING FUNCTION (MGF): The moment generating function is a function that generates the moments (e.g., mean, variance) of a random variable

MOMENT: In statistics and probability theory, a moment is a quantitative measure that describes the shape or characteristics of a distribution.

CENSORING: In statistics and data analysis, censoring refers to the phenomenon where some data points are not fully observed or recorded typically due to limitations in measurements or observations.

SURVIVAL FUNCTION: The survival function f(x) represents the probability that the event of interest (e.g. failure, death has not occurred by time x. It is the complement of the CDF. The actual survival time of an individual, x can be regarded as the value of a variable **X** which can take any non-negative value. The different values that **X** can take have a probability distribution, and we call **X** the random variable associated with the survival time. Now suppose that the random variable **X** has a probability distribution with underlying density function f(x).

$S(t) = P(T_i > t) = 1 - F(t)$.

PROBABILITY DENSITY FUNCTION (PDF)

The PDF describes the likelihood of a random variable falling within a specific range of values. In the case of a continuous random variable, the probability taken by X on some given value x is always 0. probability density function of the nonnegative random variable T is $f(t) = -S(t)\ t \geq 0$,

where S(t) is the survivor function and its derivative exists.

SCALE PARAMETER

A scale parameter is a parameter of a probability distribution that determines the spread or dispersion of the distribution. It is a measure of how stretched or compressed the distribution is. Examples of scale parameters include the standard deviation (σ) in the normal distribution and the scale parameter (β) in the exponential distribution.

SHAPE PARAMETER

A shape parameter is a parameter of a probability distribution that determines the shape of the distribution. It is a measure of how skewed or symmetric the distribution is. Examples of shape parameters include the shape parameter (k) in the Weibull distribution and the shape parameter (α) in the gamma distribution.

MEAN

The mean, also known as the arithmetic mean or average, is a measure of the central tendency of a set of data. It is calculated by summing up all the values in the dataset and dividing by the number of data.

VARIANCE

The variance is a measure of the spread or dispersion of a set of data. It represents how much individual data points deviates from the average value.

STANDARD DEVIATION

Standard deviation is a measure of the amount of variation or dispersion in a set of values. It quantifies how much the individual data points differ from the mean (average) of the dataset. A low standard deviation indicates that the data points are close to the mean, while a high standard deviation indicates that the data points are spread out over a wider range.

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

Maximum likelihood estimation (MLE) is a statistical method used to estimate the parameters of a probability distribution. The method finds the values of the parameters that maximize the likelihood function, which is the probability of observing the data given the parameters.

SKEWNESS

Skewness is a measure of the asymmetry of a probability distribution. It is a measure of how much the distribution deviates from symmetry. A distribution with zero skewness is symmetric, while a distribution with positive skewness is skewed to the right and a distribution with negative skewness is skewed to the left.

KURTOSIS

Kurtosis is a measure of the "tailed ness" of a probability distribution. It is a measure of how much the distribution deviates from normality in terms of its tails. A distribution with zero kurtosis is normal, while a distribution with positive kurtosis is leptokurtic (has heavier tails) and a distribution with negative kurtosis is Platykurtic (has lighter tails).

Akaike Information Criterion (AIC)

AIC is used to compare and select the best model among a set of candidate models. It balances the trade-off between the goodness of fit of the model and its complexity (number of parameters).

Bayesian Information Criterion (BIC)

BIC is another criterion for model selection, but it imposes a stronger penalty for models with more parameters compared to AIC. It is derived from a Bayesian perspective.

1. 5 CONCLUSION

The rest of this project is organized as follows: Chapter Two reviews some literature on lifetime distributions, their distributions, reliability functions, and moments.

Chapter Three examines the properties and applications of lifetime distributions, including their hazard and survival functions, and their maximum likelihood estimation (MLE) concentrations.

Chapter Four presents simulation studies of some lifetime distributions, focusing on the interpretation of their main features, limitations and implications. This chapter has introduced the importance of lifetime distributions in modeling and analyzing time-to-event data. It has outlined the problem statement, objectives, significance, and scope of the study, as well as the organization of the remaining chapters.

CHAPTER TWO
LITERATURE REVIEW

2 .1 INTRODUCTION

In this chapter, we review relevant literature on lifetime models, focusing on their application as population models for failure times distribution across diverse products and failure mechanism. A crucial aspect of this process is identifying the underlying lifetime distribution. Probability distributions help capture data uncertainty by recognizing patterns of variation. By condensing observations into a mathematical framework with few unknown parameters, distributions provide valuable insights into the underlying data generation mechanism. Lifetime distributions, which probabilistically describe the lifespan of a device, are heavily influenced by the device's failure mode. Nair (2018) .The concept of "lifetime distribution" in statistics originated with the study of reliability and failure analysis in engineering, where the exponential distribution was initially used to model the lifespan of components, with significant developments later including the Weibull distribution, which offered greater flexibility in capturing different aging patterns across various materials and systems; researchers like Lieblein, Zelen, and Kao played key roles in identifying and applying these distributions to real-world data analysis.

The exponential distribution was the first widely used lifetime distribution in the early 20th century due to its simple mathematical form and the property of a constant hazard rate, making it suitable for modeling situations where failures occur at a constant rate throughout the lifespan.

The Swedish engineer, Weibull in 1951 introduced the Weibull distribution, which proved more versatile as it allowed for modeling different aging patterns, including increasing, decreasing, or constant hazard rates, significantly expanding the applicability of lifetime analysis. Sheldon (2009) addressed life testing challenges, where the exponential distribution takes precedence over the normal distribution. It examines a population of items with independent lifetimes, sharing a common distribution with an unknown parameter. The goal is to estimate this parameter using available data. The concept of hazard rate function, a valuable tool in engineering, is introduced to specify lifetime distributions. This chapter provides an overview of some lifetime distributions, specifically the Exponential, Weibull, and Gamma distributions. The review includes their practical applications, key properties, and relevant characteristics, offering a comprehensive understanding of these fundamental distributions in reliability and survival analysis.

Additionally, it develops a hypothesis test to compare the means of two exponentially distributed populations and discusses two methods for estimating Weibull distribution parameters. Two-parameter gamma and two-parameter Weibull are the most popular distributions for analyzing any lifetime data. Gamma has a long history and it has several desirable properties, see Johnson, Balakrishnan (1994) for the different properties of the two-parameter gamma distribution. It has lots of applications in different fields other than lifetime distributions, some of the references can be made to Alexander (1962), Jackson (1963), Klinken (1961) and Kuroiwa (1952). The two parameters of a gamma distribution represent the scale and the shape parameters and because of the scale and shape parameters, it has quite a bit of flexibility to analyze any positive real data. It has increasing as well as decreasing failure rate depending on the shape parameter, which gives an extra edge over exponential distribution, which has only constant failure rate. Since sum of independent and identically distributed (i.i.d.) gamma random variables has a gamma distribution, it has a nice physical interpretation also. If a system has one component and n-spare parts and if the component and each spare parts have i.i.d. gamma lifetime distributions, then the lifetime distribution of the system also follows a gamma distribution. Another interesting property of the family of gamma distributions is that it has likelihood ratio ordering, with respect to the shape parameter, when the scale parameter remains constant. It naturally implies the ordering in hazard rate as well as in distribution. But one major disadvantage of the gamma distribution is that the distribution function or survival function cannot be expressed in a closed form if the shape parameter is not an integer. Since it is in terms of an incomplete gamma function, one needs to obtain the distribution function, survival function or the failure rate by numerical integration. This makes gamma distribution little bit unpopular compared to the Weibull distribution, which has a nice distribution function, survival function and hazard function. Weibull distribution was originally proposed by Weibull (1939), a Swedish physicist, and he used it to represent the distribution of the breaking strength of materials. Weibull distribution also has the scale and shape parameters. In recent years the Weibull distribution becoming very popular to analyze lifetime data mainly because in presence of censoring it is much easier to handle, at least numerically, compared to a gamma distribution. It also has increasing and decreasing failure rates depending on the shape parameter. Physically it represents a series system, because the minimum of i.i.d. Weibull distributions also follows a Weibull distribution. Several applications of the Weibull distribution can be found in Plait (1962) and Johnson (1968) although some of the negative points of the Weibull distribution can be found in Gorski (1968). One of the disadvantages can be pointed that the asymptotic convergence to normality for the distribution of the maximum likelihood estimators is very slow (Bain,1976). Therefore most of the asymptotic inferences (for example asymptotic unbiasedness or asymptotic confidence interval) may not be very accurate unless the sample size is very large. Some ramifications of this problem can be found in Bain (1976). It also does not enjoy any ordering properties like gamma distribution.

2.2 EXPONENTIAL DISTRIBUTION

The Exponential distribution is a memoryless, continuous probability distribution that only takes positive values, meaning the likelihood of an event occurring is independent of the time elapsed since the last event and it can assume any value within a defined range. It is characterized by a constant hazard rate, ensuring that the probability of an event happening within a specific time interval remains unchanged over time. This distribution is widely applied in various fields, such as modeling the time between system failures in reliability analysis, analyzing patient survival times in medical research, predicting customer arrival intervals in queueing systems, and representing the time between events in financial markets and biological systems. Additionally, it is used to model the time between machine failures in manufacturing processes, making it a versatile tool for understanding and predicting time-based events across diverse domains. It is the unique model, for continuous T, with the so-called 'memory-less property', that is the probability for the event to take place in a future interval does not depend on the current age of the item or individual considered. This distribution models the random times between events in homogeneous poisson process.

In the 18th century, Thomas Simpson (1757) laid the groundwork for exponential decay in probability theory. Later, Pierre-Simon Laplace (1781) built upon this concept, deriving the exponential distribution from the binomial distribution. In the 20th century, researchers applied the exponential distribution to various fields: Wallodi Weibull (1939) used it to model mechanical component failures. Frederick (1952) analyzed electronic component reliability. Ronald Fisher (1922) developed a statistical method for estimating exponential distribution parameters. The exponential distribution also found applications in medicine and complex systems: Ernest Greenwood (1946) modeled cancer patient survival times. Proschan (1963) analyzed the reliability of complex systems.

2.3 WEIBULL DISTRIBUTION

The Weibull distribution is a very widely used probability distribution in reliability, and it has two parameters: scale parameter $\alpha > 0$ and shape parameter $\beta > 0$. For $\beta = 1$ this distribution is an Exponential distribution, whereas $\beta > 1$ leads to an increasing hazard rate, and hence can be considered to model wear-out, as often deemed appropriate for mechanical units, and $\beta < 1$ leads to decreasing hazard rate, hence modelling wear-in of a product as often advocated for electronic units. One Weibull distribution cannot model both wear-in and wear-out at different stages, it cannot resemble a 'bath-tub shaped hazard rate', but different Weibull distributions are often advocated for the different stages of a units' life. Although its form is mathematically straightforward, statistical inference based on the Weibull distribution is less trivial due to the fact that there are no reduced-dimensional sufficient statistics for the shape parameter β.

Weibull distributions are also frequently used in regression-type inferences in reliability, enabling information from covariates to be taken into account. There is also a useful relationship between the Weibull distribution and an Extreme Value distribution (also called the Gumbel distribution), which occurs as a limiting distribution for extreme value of a sample under some assumptions on the tail of the sample distribution: if T has a Weibull distribution, then ln T has such an Extreme Value distribution.

Waloddi Weibull pioneered the Weibull distribution, introducing it in 1939 as a model for material breaking strength. He later published a comprehensive paper in 1951, outlining the distribution's properties and applications. The Weibull distribution soon found applications in various fields: Davis (1952) used it to model electronic component failure times. A.E. Green and A.J. Bourne (1964) applied it to analyze mechanical system reliability. Meanwhile, statisticians Mann and D.R. Whitney (1947) developed methods to estimate the distribution's parameters, further expanding its utility.

2.4 GAMMA DISTRIBUTION

The Gamma distribution is a versatile probability model that describes the time gaps between events occurring in a Poisson process. Key characteristics include: Being a continuous distribution, allowing for any value within a range Only taking positive values, making it suitable for modeling time and duration. Offering flexibility to accommodate various data patterns, from exponential to skewed distributions. The Gamma distribution has numerous practical applications: In reliability engineering, it models the time until systems fail. In survival analysis, it examines the duration of patient survival. In finance, it analyzes the time between events in financial markets. In engineering, it models the time gaps between machine failures. The Gamma distribution is also a useful lifetime model in many reliability applications, and has two parameters: scale parameter $\lambda > 0$ and shape parameter $\kappa > 0$.

The Gamma distribution's evolution is marked by significant contributions from renowned mathematicians and statisticians. Abraham de Moivre's (1733) introduction of the Gamma function laid the foundation. Pierre-Simon Laplace (1781) later built upon this, deriving the Gamma distribution from the binomial distribution. The 20th century saw major advancements:

Pearson (1915) pioneered parameter estimation using the method of moments. Fisher (1922) introduced maximum likelihood estimation, enhancing accuracy. Jerzy Neyman (1937) developed confidence intervals, providing a framework for uncertainty assessment. Practical applications flourished: Ernest Greenwood (1946) successfully modeled cancer patient survival times. Frank Proschan (1963) leveraged the Gamma distribution to analyze complex system reliability. The comprehensive book by Norman L. Johnson and Samuel Kotz (1970) unified the distribution's properties and applications, cementing its importance in statistical analysis.

COMPARING HAZARD FUNCTION FOR THE THREE DISTRIBUTIONS

Parameters	Gamma	Weibull	Exponential
α=1	constant	constant	constant
α>1	increasing	increasing	Increasing
	from 0 to λ	from 0 to ∞	from 0 to λ
α <1	Decreasing	Decreasing	Decreasing
	from ∞ to λ	from ∞ to 0	from ∞ to λ

Therefore, the hazard function of the EE distribution behaves like the hazard function of the gamma distribution, which is quite different from the hazard function of the Weibull distribution. For the Weibull distribution if a > 1, the hazard function increases from zero to 1 and if a < 1, the hazard function decreases from 1 to zero. Many authors point out (see Bain, 1976) that since the hazard function of a gamma distribution (for a > 1) increases from zero to a finite constant, the gamma may be more appropriate as a population model when the items in the population are in a regular maintenance program. The hazard rate may increase initially, but after some times the system reaches a stable condition because of maintenance. The same comments hold for the EE distribution also. Therefore, if it is known that the data are from a regular maintenance environment, it may make more sense to fit the gamma distribution of Exponential distribution than the Weibull distribution.

CHAPTER THREE

3. 1 RESEARCH METHODOLOGY

In fields like reliability engineering, survival analysis, and risk assessment, understanding lifetime distributions is crucial. These statistical models help analyze time-to-event data, such as how long machinery lasts, patient survival rates, or unemployment durations. Understanding the properties and applications of these distributions is essential for accurate modeling, prediction, and decision-making. This chapter reviews some of the most widely used lifetime distributions, highlighting their mathematical properties, key characteristics, and practical applications.

3.2 WEIBULL DISTRIBUTION

The Weibull distribution is a versatile lifetime distribution that can model a wide range of behaviors, including increasing, decreasing, and constant failure rates. Sometimes an experiment is characterized by a continuous random variable whose probability distribution is best fit by a function involving two or more parameters, which allows greater freedom in fitting the experimental results. The Weibull distribution is most easily introduced in terms of its cumulative distribution function, which has the form

$$F(x, \beta, \alpha) = 1 - e^{-\left(\frac{x}{\alpha}\right)\beta} \quad for\ x \geq 0 \qquad 3.2.1$$

$$= 0\ for \leq 0$$

where the two parameters α and β are positive. This function is easily seen to be a generalization of the cumulative distribution function for the exponential distribution, which corresponds to the choice of β = 1. The probability density function for the Weibull distribution can then be obtained as the derivative of F(x ; β,α) with respect to x:

$$f(x, \beta, \alpha) = \beta\alpha^{-\beta}\, x^{\beta-1} e^{-(x/\alpha)\beta}, \quad for\ x \geq 0 \qquad 3.2.2$$

$= 0 \quad for\ x \leq 0$

When β=1, the Weibull distribution reduces to the exponential distribution with the identification of the exponential parameter λ with 1/α. The parameter β is a shape parameter affecting the shape of the distribution, while α is a scaling parameter affecting the scale. It is also possible to change the location of the Weibull distribution by replacing the variable x by x-x_0.

The mean and variance of the Weibull distribution have complicated forms:

$$E(X) = \alpha\Gamma\left(1 + \frac{1}{\beta}\right) and\ V(X) = \alpha^2 \left\{\Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\right\} \qquad 3.2.3$$

where

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

PDF

$$F(x) = \left(\frac{\beta}{\eta}\right)\left(\frac{x}{\eta}\right)^{\beta-1} e^{\left(-\frac{x}{\eta}\right)^{\beta}} \text{ For } x \geq 0 \qquad 3.2.4$$

CDF

$$F(x) = 1 - e^{\left(\frac{-x}{\eta}\right)^{\beta}} \qquad 3.2.5$$

Mean: 3.2.6

$\mu = \eta\, \Gamma(1 + \frac{1}{\beta})$

Variance: 3.2.7

$\sigma^2 = \eta^2$ [Γ(1 + $\frac{2}{\beta}$) – (Γ (1 + $\frac{1}{\beta}$))^2]

Standard Deviation:

$\sigma = \sqrt{\eta^2}\ [\Gamma(1 + \frac{2}{\beta} - (\Gamma\left(1 + \frac{1}{\beta}\right))^2]$

Harvard Function

$$\mu(x) = \left(\frac{\beta}{\eta}\right)\left(\frac{x}{\eta}\right)^{\beta-1} \qquad 3.2.8$$

Where

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

is the gamma function.

The Weibull distribution is widely used in modeling failure times, because a great variety of shapes of probability curves can be generated by different choices of the two parameters, β and α. Weibull distributions range from exponential distributions to curves resembling the normal distribution. Weibull distributions correspond to distributions of failure times that are peaked at certain ages and skewed in different fashions. It is used as a model with diverse types of items such as ball bearing automobile components Lieblein and Zelien (1956), and electrical insulators. It is also used in biological and medical applications. In most practical cases, the hazard rate is non - constant, contradicting the exponential distribution's assumption.

According to Collette 2003, a more general form of the hazard function is such that

h(t) = $\lambda\gamma t^{\beta-1}$, for $0 \le t \le \infty$,

is a function that depends on two parameters λ and β which are both greater than zero. We notice that in a particular case where λ = 1, the hazard function takes a constant value λ which is equivalent to equation 2.0. For other values of γ, the hazard function increases (decreases) monotonically. The shape of hazard function depends critically in value of γ, and so γ is known as the scale parameter.

it's density function is given b. The weibull distribution has been successfully employed to analyze failure data from; fatigue failure (weibull 1939) vacuum tube failure (kao 1956), ball-bearing failure (Lieblein and Zelien 1956). For the weibull distribution denoted WEI (λ , α), the mean (Expected value) of random variable T that has this distribution can be shown to be given by ;

$$E(T) = \lambda^{-\frac{1}{\alpha}}\Gamma(\alpha^{-1} + 1),$$

$$\text{Where} \quad \Gamma(x) = \int_0^\infty u^{x-1}e^{-u}\, du. \qquad 3.2.9$$

Research has identified limitations in weibull model applicability for lifetime data, specifically issues with maximum likelihood estimation when the location parameter is zero, as discussed by Bain (1978).

APPLICATION: The Weibull distribution is a versatile tool with numerous applications. It helps predict product lifespan and material durability in reliability analysis, and is also used in weather forecasting, medical research, and industrial engineering to understand and anticipate failures, ensuring more reliable systems and processes.

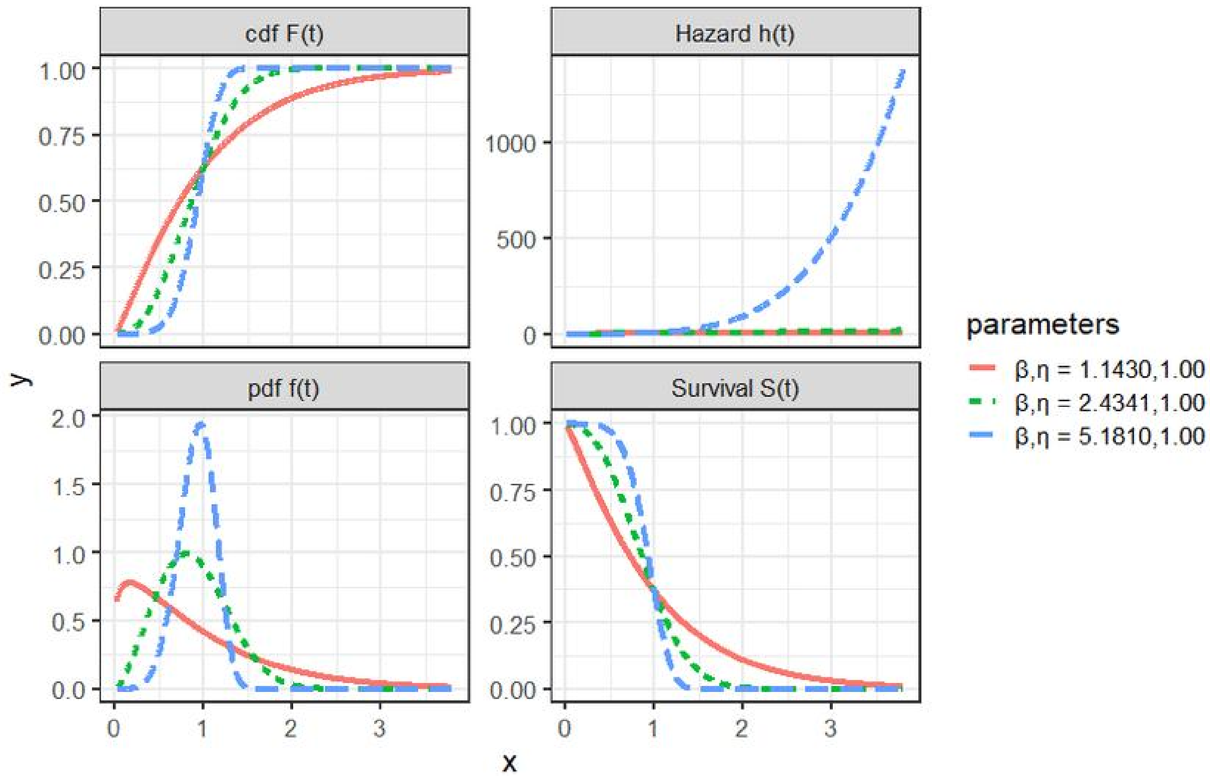


Fig 1: pdf, cdf, hazard , survival fit . lawless (2003).

3.3 FITTING THE DISTRIBUTION

The likelihood function is maximized to obtain estimates for shape, scale, and rate parameters.

The likelihood Function L is defined as

$$L = \prod_{i=1}^{n} [f(x_i;\boldsymbol{\theta})\,] \qquad (3.3.1)$$

where $f(x_i;\boldsymbol{\theta})$ is the probability density function (PDF) of the chosen distribution, and θ represents the parameters to be estimated.

For **Maximum Likelihood Estimation (MLE),** we typically work with the **log-likelihood function**:

$$In\,L = \prod_{i=1}^{n} [In\,f(x_i;\boldsymbol{\theta})\,] \qquad 3.3.2$$

FITTING THE WEIBULL DISTRIBUTION

The Weibull distribution is defined as:

$$f(x: \lambda, k) = \left(\frac{k}{\lambda}\right)\left(\frac{x}{\lambda}\right)^{k-1} \ell^{-\left(\frac{x}{\lambda}\right)^{k}} \qquad 3.3.3$$

where:

- k is the **shape parameter**,
- λ is the **scale parameter**.

The **log-likelihood function** for X = {x1 ,x2 ,...,xn } is:

$$\text{In } L(k , \lambda) = n \text{ In } k - nk \text{ In } (\lambda) + (k , \lambda) \sum_{i=1}^{n} Inx_i - \sum_{i=1}^{n} \left(\frac{x_i}{\lambda}\right) ^k \qquad 3.3.4$$

The MLE estimates are found by solving:

$$\frac{n}{k} - n \text{ In } \lambda + \sum_{i=1}^{n} Inx_i = 0$$

$$\lambda = \left(\frac{1}{n} \sum_{i=1}^{n} {x^k}_i\right) ^\frac{1}{k}$$

It follows that $E[x] \quad = \frac{1}{\lambda}\Gamma\left(1 + \frac{r}{a}\right), and$

$$V[X] = E[X^2] - (E[X])^2$$

$$= \frac{1}{\lambda^2}\left[\Gamma\left(1 + \frac{2}{a}\right) - \Gamma^2\left(1 + \frac{1}{a}\right)\right].$$

3.4 EXPONENTIAL DISTRIBUTION

The exponential distribution is one of the simplest and most commonly used lifetime distributions. It is characterized by its constant hazard rate, which makes it memoryless. The assumption of ED is that it has constant hazard function overtime. i.e.;

$$h(t) = \lambda, \qquad for\ 0 \leq t \leq \infty, \lambda > 0$$

Thus, $$S(t) = exp\left[-\int_0^t h(t)dt\right],$$

This implies that;

$$S(t) = exp[-\lambda t] = e^{-\lambda t}, \quad 0 \leq t \leq \infty,$$

The P.D.F. can be derived from equation (222), i.e.,

$$\frac{-dR(t)}{dt} = f(t); \qquad R(t) = e^{-\lambda t},$$

$$\Rightarrow f(t) = \frac{-d}{dt}e^{-\lambda t} = -(-\lambda e^{-\lambda t})$$

$$= \lambda e^{-\lambda t}. \qquad 3.4.1$$

The cumulative distribution function is given by

$$F(x) = 1\text{-}exp^{(-\lambda x)} \qquad 3.4.2$$

Variance: 3.4.3

$\sigma^2 = \frac{1}{\sigma^2}$

Standard Deviation 3.4.4

$\sigma = \frac{1}{\lambda}$

Hazard Function

h(x) = λ Constant 3.4.5

Memoryless Property

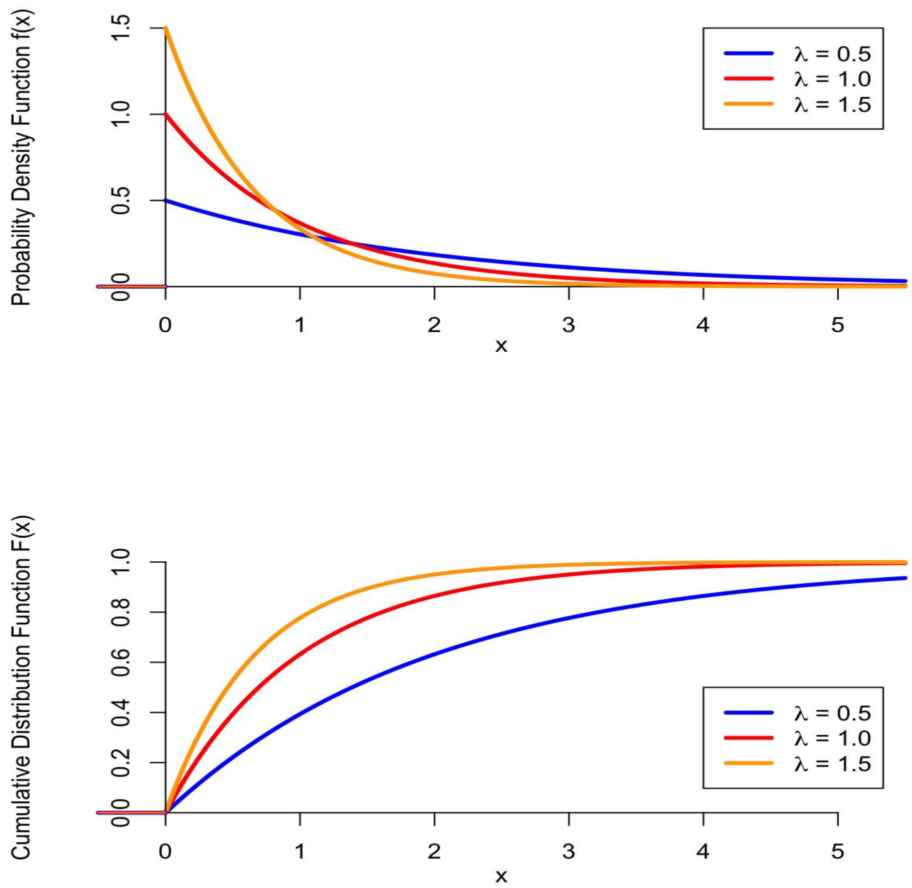
$P\left(\frac{X > S+t}{X > S}\right) = P(x > t)$ 3.4.6

The shape parameter α and the scale parameter β is determined by the shape of the distribution. This study examines the estimation of the parameter through model fitting to observed data. Because of Its constant failure rate property, exponential distribution is a perfect model for a long flat portion of the bathtub curve. Given that most components are systems that spend most of their lifetimes in this portion of the bathtub curve. The data shows historically that the exponential distribution was the first widely discussed lifetime distribution model. This was in part because of the availability of simple statistical method for it. The assumption of a constant hazard function is very restrictive, so the model application is fairly limited. In recent years, in order to overcome such a problem, a new class of models were introduced based on modifications of exponential distribution. Gupta and

Kundu (1999) proposed a general exponential distribution, which is able to accommodate data with increasing and decreasing failure rate functions.

APPLICATION: The exponential distribution has numerous practical applications. It's commonly used in reliability engineering to predict the time between system failures, assuming a steady failure rate. Additionally, it's applied in queuing theory to analyze gaps between arrivals and in survival analysis to model the time until an event occurs, provided the risk remains constant.

Fig 2: pdf, cdf of the exponential distribution. Schervish, 2012.



3.5 FITTING THE EXPONENTIAL DISTRIBUTION

The **Exponential distribution** is a special case of the Weibull distribution where k=1, and it is given by:

$$F(x\ ;\lambda)=\ \lambda e^{-\lambda x}, \qquad x\geq 0, \lambda>0 \qquad 3.5.1$$

where λ is the **rate parameter**.

The **log-likelihood function** is:

$$\text{In } L(\lambda) = \ n \text{ In } (\lambda) - (\lambda) \sum_{i=1}^{n} x_i \qquad 3.5.2$$

Setting the derivative to zero gives:

$$\hat{\lambda}=\frac{n}{\sum_{i=1}^{n} x_i}$$

3.6 GAMMA DISTRIBUTION

The random variable , X is said to have a gamma distribution with parameter λ and α , denoted by Gam (λ,α) if its density function is given by ;

$$F_{\lambda,\alpha}(x) = \frac{\lambda\alpha x^{\alpha-1}}{\Gamma(\alpha)} e^{-\lambda x}, \quad x > 0$$

PDF

$$F(x) = \left(\frac{1}{\Gamma(\alpha)}\right)\left(\frac{1}{\beta^{\alpha}}\right) X^{(\alpha-1)}\ e^{\left(\frac{-x}{\beta}\right)} \qquad 3.6.1$$

For $x \geq 0$

CDF

$$F(x) = \frac{\gamma\left(\alpha,\frac{X}{\beta}\right)}{\Gamma(\alpha)} \qquad 3.6.2$$

Mean:

$$M = \alpha\beta \qquad 3.6.3$$

Variance:

$$\sigma^2 = \alpha\beta^2$$

Standard Deviation:

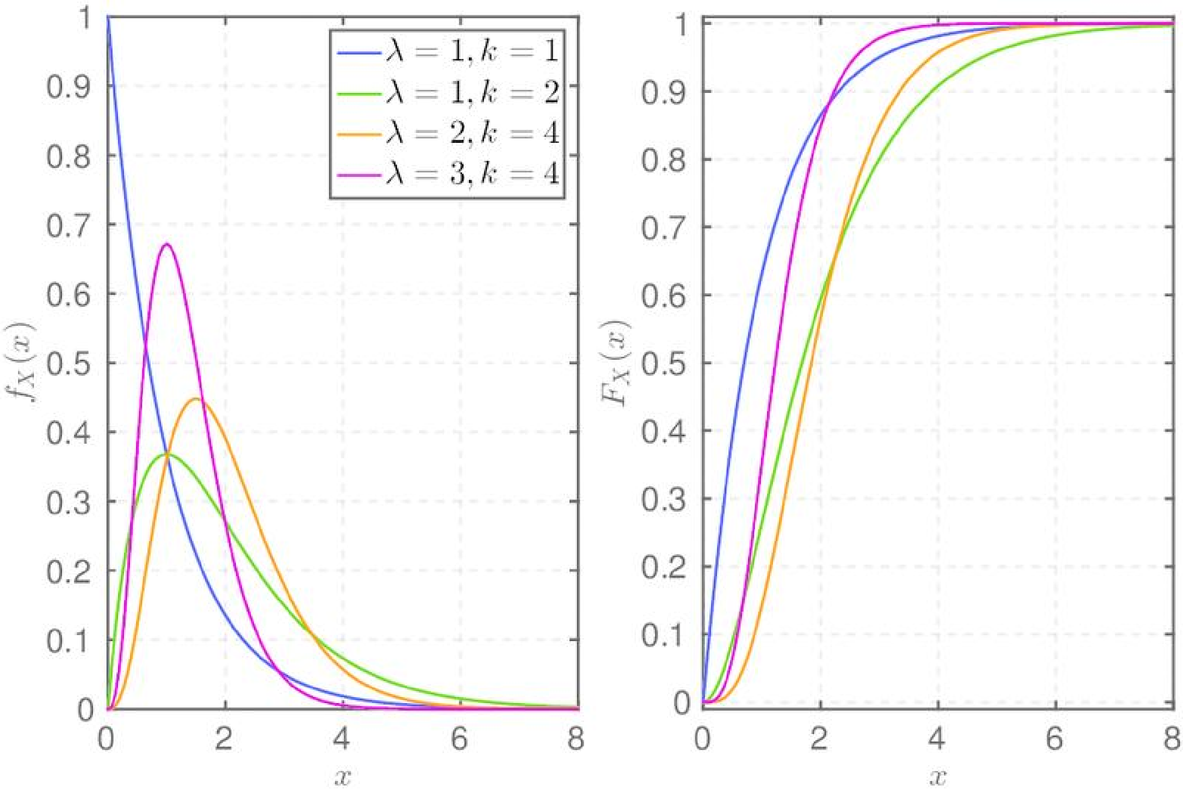
$$\sigma = \sqrt{\alpha\beta^2} \qquad 3.6.4$$

When α= 1, the distribution reduces to an exponential with constant failure rate λ.

The shape parameter α and the scale parameter β is determined by the shape of the distribution. Despite its versatility in modeling various lifetime data, the gamma distribution is relatively underutilized compared to exponential, lognormal, and log-logistic distributions in lifetime modeling applications. Nevertheless, it emerges naturally in certain contexts involving exponential distributions, particularly due to the established result that sums of independent and identically distributed (i.i.d.) exponential random variables follow a gamma distribution. A notable limitation of the gamma distribution, however, lies in the computational complexity of its distribution function and survival function when the shape parameter is non-integer, necessitating reliance on mathematical tables or specialized software (Gupta 2010).

APPLICATION: The gamma distribution has diverse applications across multiple disciplines. From optimizing queues and assessing insurance risks to modeling financial systems, it provides valuable insights. Additionally, it helps environmental scientists understand rainfall patterns and medical researchers analyze the timing of recurring diseases.

Fig 3: PDF, CDF of gamma distribution for various parameters. (Ross, 2014).



3.7 FITTING THE GAMMA DISTRIBUTION

The **Gamma distribution** is defined as:

$$f(x;\alpha,\beta) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \quad 3.7.1$$

where:

- α is the **shape parameter**, (k)
- β is the **rate parameter**
- x represents the failure times,
- Γ(α) is the **Gamma function**.

The **log-likelihood function** for a dataset X={x1, x2,..., xn} is:

$$\text{In } L(\alpha , \beta) = n\alpha \text{ In } \beta \ - n \text{ In } \Gamma(\alpha) \ + (\alpha - 1) \sum_{i=1}^{n} Inx_i \text{ -}\beta \sum_{i=1}^{n} x_i \qquad 3.7.2$$

The MLE estimates of α and β are found by solving:

$$\hat{\beta} \ = \ \frac{n\alpha}{\sum_{i=1}^{n} x_i} \qquad 3.7.3$$

$$\text{In}\hat{\alpha} \ - \frac{\Gamma'(\hat{\alpha})}{\Gamma(\hat{\alpha})} = \ \frac{1}{n}\sum_{i=1}^{n} Inx_i \text{ -In } \left(\frac{1}{n}\sum_{i=1}^{n} x^k{}_i\right)$$

3.8 METHODOLOGY FOR ANALYSIS

The methodology for analyzing the datasets using the three distributions is outlined for data 1 and data 2

PARAMETER ESTIMATION

The parameters of the Exponential, Weibull, and Gamma distributions were estimated using Maximum Likelihood Estimation (MLE). The log-likelihood functions were maximized to obtain the parameter estimates.

GOODNESS - OF - FIT MEASURES

The goodness-of-fit of the distributions was evaluated using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). These measures balance the fit of the model with its complexity, with lower values indicating a better fit.

This chapter provided an in-depth examination of key lifetime distributions, including exponential, Weibull, gamma and generalized extreme value distributions. Each distribution boasts distinct characteristics, rendering them suitable for specific uses in reliability engineering, survival analysis, and risk assessment. Mastering these distributions and their applications is vital for precise modeling and analysis of time-to-event data across various disciplines. The choice of distribution depends on the specific characteristics of the data and the distinct features of the phenomenon being studied. Future research opportunities lie in developing new distributions or refining existing ones to better capture the complexities of real-world data.

CHAPTER FOUR

DATA ANALYSIS AND INTERPRETATION

4.1 Introduction

This chapter presents the analysis and interpretation of the data collected for this research, which aims to review and compare the effectiveness of various lifetime distributions. Using two datasets (Data 1 and Data 2), the study evaluates the performance of three selected lifetime distributions: Exponential, Weibull, and Gamma.

Dataset 1: Strength measurements of 1.5 cm glass fibers (National physical laboratory, England)

0.55	**0.93**	**1.25**	**1.36**	**1.49**	**1.52**	**1.58**	**1.61**	**1.64**	**1.68**	**1.73**	**1.81**
2.00	0.74	1.04	1.27	1.39	1.49	1.53	1.59	1.61	1.66	1.68	1.76
1.82	2.01	0.77	1.11	1.28	1.42	1.50	1.54	1.60	1.62	1.66	1.69
1.76	1.84	2.24	0.81	1.13	1.29	1.48	1.51	1.55	1.61	1.62	1.66
1.70	1.77	1.84	0.84	1.24	1.30	1.48	1.51	1.55	1.61	1.63	1.67
1.70	1.78	1.89									

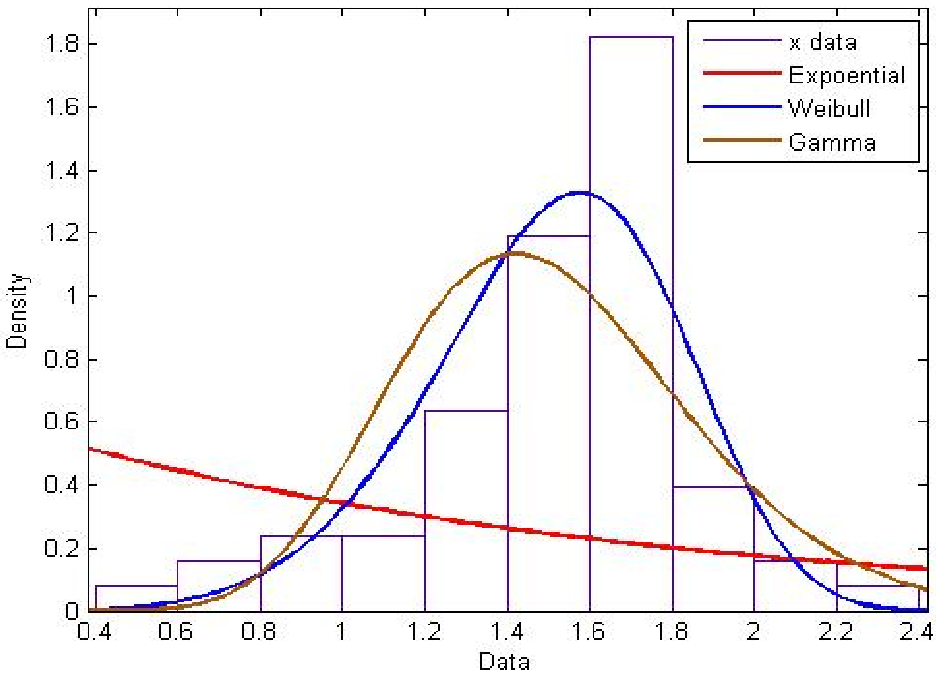
Table 4.1: Descriptive Statistics for Dataset 1

DESCRIPTIVE STATISTICS	
Mean	1.5070
Standard Error	0.0408
Median	1.5900
Mode	1.6100
Standard Deviation	0.3241
Sample Variance	0.1051
Kurtosis	1.1051
Skewness	-0.9236
Range	1.6900
Minimum	0.5500
Maximum	2.2400
Sum	94.9400
Count	63.0000
Confidence Level(95.0%)	0.0816

Table 4.2 Simulation Results for Dataset 1

Model	MLE		Log-Likelihood	AIC	BIC
	$\hat{k}$	$\hat{\lambda}$			
Exponential	-	0.5070 (0.1900)	-88.837	179.6700	181.8200
Weibull	1.62826 (0.0371)	5.7827 (0.5763)	- 15.1994	34.3990	38.6850
Gamma	17.4396 (3.0780)	0.0864 (0.0155)	- 23.9500	51.9000	56.1860

Standard error of estimates in parenthesis, $\lambda \equiv \beta$ for the Gamma model



Dataset 2: Fatigue Life of 6061-T6 Aluminum coupons (Birnbaum & Saunders, 1969)

5	25	31	32	34	35	38	39	39	40	42	43
43	43	44	44	47	47	48	49	49	49	51	54
55	55	55	56	56	56	58	59	59	59	59	59
63	63	64	64	65	65	65	66	66	66	66	66
67	67	67	68	69	69	69	69	71	71	72	73
73	73	74	74	76	76	77	77	77	77	77	77
79	79	80	81	83	84	84	86	86	87	90	91
92	92	92	92	93	94	97	98	98	99	101	103
105	109	136	147								

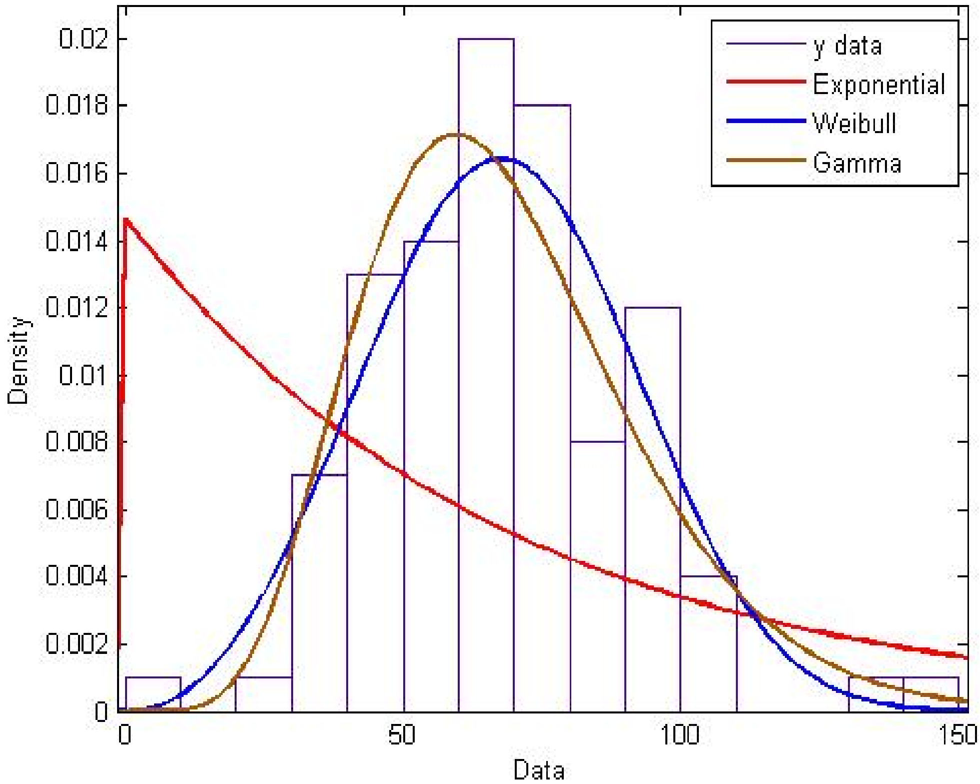
Table 4.3 Descriptive Statistics for dataset 2

DESCRIPTIVE STATISTICS	
Mean	68.3400
Standard Error	2.2419
Median	67.0000
Mode	77.0000
Standard Deviation	22.4194
Sample Variance	502.6307
Kurtosis	1.3798
Skewness	0.4152
Range	142.0000

Minimum	5.0000
Maximum	147.0000
Sum	6834.0000
Count	100.0000
Confidence Level(95.0%)	4.4485

Table 4.4 Simulation Results for Dataset 2

Model	MLE		Log-Likelihood	AIC	BIC	K-S	P-Value
	$\hat{k}$	$\hat{\lambda}$					
Exponential	-	68.34 (6.834)	-522.45	1046.9	1049.5		
Weibull	75.9232 (2.4913)	3.2103 (0.2370)	-454.29	912.58	917.79		
Gamma	7.6827 (1.0637)	8.8953 (1.2728)	-457.91	919.93	925.04		



4.2 INTERPRETATION OF THE RESULT

The results are interpreted below to determine the best-fitting distribution for each dataset

Dataset 1 (Glass Fiber Strength): The Weibull and Gamma distributions provide the best fit, with the Weibull distribution being slightly superior due to its flexibility in modeling varying failure rates.

Dataset 2 (Fatigue Life of Aluminum Coupons): The Gamma distribution provides the best fit, followed closely by the Weibull distribution. The Exponential distribution is not suitable for this dataset. while the Exponential distribution performs poorly. The Gamma distribution shows moderate performance, suggesting that it may be suitable for certain types of data but is generally outperformed by the Weibull distribution. These findings highlight the importance of selecting appropriate lifetime distributions for accurate reliability analysis and failure prediction.

4.3 CONCLUSION

The Gamma distribution shows moderate performance, with higher AIC and BIC values compared to the Weibull distribution, suggesting it may be suitable for specific data types but is generally outperformed. The Exponential distribution performs poorly for both datasets, indicating that the assumption of a constant failure rate is not suitable for this type of data. This chapter has demonstrated the practical application of the Exponential, Weibull, and Gamma distributions to real-world datasets. The results highlight the importance of selecting the appropriate distribution based on the characteristics of the data. In the next chapter, we will summarize the findings of this study and provide recommendations for future research

CHAPTER FIVE

SUMMARY, RECOMMENDATION AND CONCLUSION

5.1 SUMMARY

This study has reviewed three important lifetime distributions—the Exponential, Weibull, and Gamma distributions—focusing on their mathematical formulations, properties, and applications. Through theoretical analysis and practical case studies, we have demonstrated how these distributions can be used to model real-world lifetime data. The Exponential distribution, while simple and memoryless, is limited by its constant hazard rate and provided a poor fit for both datasets analyzed. The Weibull distribution, with its flexibility in modeling varying hazard rates, performed well for the glass fiber strength data, while the Gamma distribution, ideal for multi-phase processes, excelled in modeling the fatigue life of aluminum coupons. Both Weibull and Gamma distributions outperformed the Exponential distribution, highlighting their versatility. The study's contributions include a theoretical review, practical applications, and guidelines for distribution selection. Limitations include the focus on only three distributions and the use of relatively small datasets. Future research should explore additional distributions, larger datasets, advanced estimation methods, and applications in new fields. Overall, this study underscores the importance of selecting the appropriate lifetime distribution based on data characteristics, providing valuable insights for researchers and practitioners in reliability engineering, survival analysis, and related fields.

5 .2 RECOMMENDATION

The findings of this study highlight the superiority of the Weibull distribution for modeling lifetime data due to its flexibility in capturing varying failure rates, making it the most suitable choice for applications such as glass fiber strength and fatigue life analysis. The Gamma distribution offers moderate performance and can be a useful alternative for datasets with specific characteristics, such as skewness, but is generally outperformed by the Weibull distribution. The Exponential distribution. with its assumption of a constant failure rate, is often inadequate for real-world data and should be avoided unless there is strong evidence supporting its use. For future research, it is recommended to explore additional distributions, such as the Log-Normal and Generalized Gamma, and to validate findings using larger and more diverse datasets. Practitioners should prioritize the Weibull distribution for reliability analysis, while also considering the Gamma distribution for specific cases. However, this study has limitations, including its focus on only three distributions, the relatively small sample sizes of the datasets, and assumptions made during data preparation, which may have influenced the results. These limitations suggest the need for further research to improve the generalizability and robustness of the findings.

5. 3 CONCLUSION

In conclusion, this study demonstrated that the Weibull distribution is the most suitable model for analyzing lifetime data, while the Exponential distribution is often inadequate due to its restrictive assumptions. The Gamma distribution offers moderate performance but is generally outperformed by the Weibull distribution. These findings emphasize the importance of selecting appropriate lifetime distributions for accurate reliability analysis and failure prediction. Future research should build on these findings to explore additional distributions and validate the results using larger datasets.

REFERENCES

1. Lawless, J. F. (2003). Statistical Models and Methods for Lifetime Data (2nd ed.). Wiley.
2. Meeker, W. Q., & Escobar, L. A. (1998). Statistical Methods for Reliability Data. Wiley.
3. Klein, J. P., & Moeschberger, M. L. (2003). Survival Analysis: Techniques for Censored and Truncated Data (2nd ed.). Springer.
4. Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). Continuous Univariate Distributions (Vol. 1, 2nd ed.). Wiley.
5. Nelson, W. (2004). Applied Life Data Analysis. Wiley.
6. Cox, D. R., & Oakes, D. (1984). Analysis of Survival Data. Chapman & Hall.
7. Weibull, W. (1951). "A Statistical Distribution Function of Wide Applicability." Journal of Applied Mechanics, 18(3), 293–297.
8. Birnbaum, Z. W., & Saunders, S. C. (1969). "A New Family of Life Distributions." Journal of Applied Probability, 6(2), 319–327.
9. Gupta, R. C., & Kundu, D. (1999). "Generalized Exponential Distributions." Australian & New Zealand Journal of Statistics, 41(2), 173–188.
10. Bain, L. J. (1978). Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods . Marcel Dekker.
11. Collett, D. (2003). *Modelling Survival Data in Medical Research* (2nd ed.). Chapman & Hall/CRC.
12. DeGroot, M. H., & Schervish, M. J. (2012). Probability and statistics (4th ed.). Pearson.
13. Khan Academy. (2023). Exponential distribution. Retrieved from https://www.khanacademy.org
14. Lawless, J. F. (2003). Statistical models and methods for lifetime data (2nd ed.). Wiley.
15. Ross, S. M. (2014). Introduction to probability models (11th ed.). Academic Press.
16. Weibull.com. (2023). Weibull distribution. Retrieved from https://www.weibull.com
17. Wikipedia contributors. (2023). Gamma distribution. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org
18. National Institute of Standards and Technology (NIST). (2023). Engineering Statistics Handbook. Retrieved from

https://www.itl.nist.gov/div898/handbook/

19. MathWorks. (2023). Statistics and Machine Learning Toolbox Documentation. Retrieved from https://www.mathworks.com/help/stats/
20. Lieblein, J., & Zelen, M. (1956). "Statistical Investigation of the Fatigue Life of Deep-Groove Ball Bearings." Journal of Research of the National Bureau of Standards, 57(5), 273–316.
21. Kao, J. H. K. (1956). "A New Life-Quality Measure for Electron Tubes." IRE Transactions on Reliability and Quality Control, 5(1), 1–12.
22. Greenwood, E. (1946). "The Natural Duration of Cancer." Reports on Public Health and Medical Subjects, 33, 1–26.
23. Proschan, F. (1963). "Theoretical Explanation of Observed Decreasing Failure Rate." Technometrics ,5(3), 375–383.
24. D'Agostino, R. B., & Stephens, M. A. (1986). Goodness-of-Fit Techniques. Marcel Dekker.
25. Casella, G., & Berger, R. L. (2002). Statistical Inference (2nd ed.). Duxbury.
26. National Physical Laboratory, England. (n.d.). Glass Fiber Strength Dataset. Retrieved from https://www.npl.co.uk/
27. Birnbaum, Z. W., & Saunders, S. C. (1969). Fatigue Life of Aluminum Coupons Dataset. Retrieved from https://www.jstor.org/