

**A NEW PARALOGISTIC-WEIBULL DISTRIBUTION: MODEL, PROPERTIES,
AND APPLICATION**

BY

**DELIGHT CHUKWUEBUKA AMANDI
PSC1910340**

**DEPARTMENT OF STATISTICS
FACULTY OF PHYSICAL SCIENCES,
UNIVERSITY OF BENIN.
BENIN CITY.**

APIRL, 2024

**A NEW PARALOGISTIC-WEIBULL DISTRIBUTION: MODEL, PROPERTIES,
AND APPLICATION**

BY

DELIGHT CHUKWUEBUKA AMANDI

PSC1910340

**A PROJECT WORK SUBMITTED TO THE DEPARTMENT OF
STATISTICS, FACULTY OF PHYSICAL SCIENCES,
UNIVERSITY OF BENIN. BENIN CITY.**

**IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR
THE AWARD OF BACHELOR OF SCIENCE (BSc.) DEGREE IN STATISTICS**

APIRL, 2024

UNDERTAKING

I, **DELIGHT CHUKWUEBUKA AMANDI**, carried out this project work with matriculation number **PSC1910340**. I can confidently assure you that I did not copy the work of any other author(s). All the texts used in this project have been duly cited and acknowledged.

AMANDI DELIGHT CHIKWUEBUKA

DATE

CERTIFICATION

This text is a certification that the project titled “A NEW PARALOGISTIC-WEIBULL DISTRIBUTION: MODEL, PROPERTIES, AND APPLICATION” was carried out by the student with matriculation number PSC1910340. The project was completed under the supervision of Dr. S.A. Osagie in the Department of Statistics, Faculty of Physical Sciences, University of Benin, as part of the requirement for the award of the Bachelor of Science (B.Sc) degree in the Department of Statistics.

DR S.A. OSAGIE
PROJECT SUPERVISOR

SIGN.DATE

PROF. N. EKHOUSEHI
(HEAD OF DEPARTMENT)

SIGN.DATE

EXTERNAL EXAMINER

SIGN.DATE

DEDICATION

This project is dedicated to God Almighty, who has always been with me through thick and thin. I also dedicate this study to my late brother, Nwosu Chinaemerem, whose dream was to attend university. Additionally, I would like to dedicate this project to my parents, Mr and Mrs Justus and Christiana Nwosu, and my mentor, Hon. Charles Henry Chimezie. Your unwavering support and endless encouragement have been the foundation of my journey.

ACKNOWLEDGEMENT

I am truly grateful for the support and guidance I have received throughout my journey. First and foremost, I give thanks to the Almighty God for His grace, guidance, and blessings. I would like to extend a special acknowledgment to my parents, Mr. and Mrs. Justus Nwosu, for their unwavering support and sacrifices. Their love and encouragement have been my driving force.

To my mentor, Mr. Charles Henry Chimezie, I owe a debt of gratitude for his invaluable support and guidance from my earliest days in primary school to the completion of my university education. Your belief in me has been a constant source of motivation.

I am also grateful to my project supervisor, Dr. S.A. Osagie, for his unwavering guidance, patience, fatherly role, and mentorship of my academic. Prof. J.E.

Osemwenkhae has also made significant contributions to my work. I must also mention my siblings; Franklin, Doris, and Christiana, and my dear friend-turned-brother, Kingsley, for their unwavering support.

I would like to extend a special mention to Onye Faith and all my school friends. Thank you for your friendship and support. To my holy family at EFAC Ministry

has been a source of spiritual guidance and prayers, and for that, I am deeply grateful.

To everyone mentioned and to those whose names may not appear here but have contributed in various ways, I am deeply thankful. Your support has been invaluable, and I am blessed to have you in my life.

TABLE OF CONTENTS

TITLE PAGE.....	I
COVER PAGE.....	II
UNDERTAKING.....	iii
CERTIFICATION	iv
DEDICATION.....	v
ACKNOWLEDGEMENT	vi
TABLE OF CONTENTS.....	viii
LIST OF TABLES.....	xii
ABSTRACT.....	xiii
CHAPTER ONE.....	1
1.1 INTRODUCTION.....	1
1.2 GENERAL BACKGROUND.....	1
1.3 WEIBULL DISTRIBUTION.....	2
1.4 PARALOGISTIC DISTRIBUTION.....	2
1.5 AIM AND OBJECTIVES OF THE STUDY	3
1.6 SIGNIFICANCE OF THE STUDY	4
1.7 ORGANIZATION OF THE STUDY.....	4
1.8 DEFINITION OF BASIC IMPORTANT TERMS.....	5
1.9 SUMMARY.....	8
CHAPTER	
TWO.....	9
LITERATURE REVIEW.....	9

2.1 INTRODUCTION	9
2.2 METHOD OF GENERATING NEW DISTRIBUTION.....	9
2.2.1 TRANSFORMER-TRANSFORMED (T-X) METHOD.....	10
2.2.2 METHOD OF QUANTILE.....	11
2.2.3 METHOD OF ADDING PARAMETERS TO AN EXISTING DISTRIBUTION.....	13
2.2.4 METHODS OF PARAMETER ESTIMATION.....	14
2.2.4.1. MAXIMUM LIKELIHOOD ESTIMATION (MLE).....	14
2.2.4.2. METHOD OF MOMENTS (MOM).....	15
2.2.4.3. BAYESIAN ESTIMATION.....	16
2.2.4.4. LEAST SQUARES ESTIMATION.....	16
2.2.4.5. METHOD OF MOMENTS-BOOTSTRAP (MMB).....	18
2.2.4.6. QUANTILE REGRESSION.....	18
2.3 REVIEW OF PARALOGISTIC AND WEIBULL DISTRIBUTIONS.....	19
2.4 SUMMARY	22
CHAPTER THREE.....	23
METHODOLOGY.....	23
3.1 INTRODUCTION.....	23
3.2 CONSTRUCTION OF PARALOGISTIC-WEIBULL (PW)	24
3.3 STATISTICAL PROPERTIES OF PARALOGISTIC -WEIBULL DISTRIBUTION.....	26

3.3.1 THE PROBABILITY DENSITY FUNCTION (PDF) OF THE PARALOGISTIC WEIBULL DISTRIBUTION.....	27
3.3.2 THE SURVIVAL FUNCTION OF THE PARALOGISTIC – WEIBULL DISTRIBUTION.....	28
3.3.3 THE HAZARD FUNCTION OF THE PARALOGISTIC – WEIBULL DISTRIBUTION	28
3.3.4 THE REVERSED HAZARD OF THE PARALOGISTIC WEIBULL DISTRIBUTION.....	28
3.3.5 THE QUANTILE FUNCTION.....	29
3.3.6 THE MEDIAN OF THE PARALOGISTIC – WEIBULL DISTRIBUTION.....	30
3.3.7 THE MOMENTS OF THE PARALOGISTIC – WEIBULL DISTRIBUTION.....	30
3.3.8 MOMENT GENERATING FUNCTION [MGF].....	31
3.4 MAXIMUM LIKELIHOOD FUNCTION (MLE) METHOD.....	33
3.5 SUMMARY.....	36
CHAPTER FOUR.....	37
ANALYSIS AND INTERPRETATION OF DATA.....	37
4.1 INTRODUCTION.....	37
4.2 GOODNESS OF FIT TESTS.....	37
4.2.1 KOLMOGOROV-SMIRNOFF TEST (KS TEST).....	38

4.2.1 ANDERSON DARLING (A*) TEST.....	38
4.2.3 CRAMÉR-VON MISES (W*) TEST.....	39
4.3 CRITERIA FOR DISCREPANCY.....	39
4.3.1 AKAIKE INFORMATION CRITERION (AIC).....	40
4.3.2 BAYESIAN INFORMATION CRITERION (BIC).....	40
4.3.3 HANNAN-QUINN INFORMATION CRITERION (HQIC).....	40
4.4 APPLICATION USING LIFETIME DATA.....	41
4.4.1 REMARK ON ANALYSIS OF DATA SET 1.....	44
4.4.2 REMARK ON ANALYSIS OF DATA SET 2.....	45
4.5 SUMMARY.....	46
CHAPTER FIVE.....	47
CONCLUSION AND RECOMMENDATIONS.....	47
5.1 INTRODUCTION.....	47
5.2 DISCUSSION.....	47
5.3 CONCLUSION.....	48
REFERENCES.....	49

LIST OF TABLES

Table 4.1 Parameter estimates of the model parameters and the maximum of the log-likelihood	
Function for data set 1.....	42
Table 4.2 The discrepancy criteria for Data Set 1.....	42
Table 4.3 Parameter estimates of the model parameters and the maximum of the log-likelihood	
function for data set 2.....	43
Table 4.4 The discrepancy criteria for Data Set 2.....	44

ABSTRACT

Lifetime data analysis plays a crucial role in various fields, ranging from engineering to epidemiology. In this study, we investigate the effectiveness of the Paralogistic-Weibull distribution in modeling lifetime data compared to other competing distributions such as the Weibull and Paralogistic distributions. Two datasets were analyzed: the daily number of COVID-19 infected persons in Nigeria and the survival times of patients with Head and Neck Cancer. We employed goodness-of-fit tests, including the Kolmogorov-Smirnov, Anderson-Darling, and Cramér-von Mises tests, along with discrepancy criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), to evaluate the performance of the distributions. The results indicate that the Paralogistic-Weibull distribution consistently outperforms the other distributions across both datasets, exhibiting higher p-values and lower discrepancy criteria values. Therefore, we conclude that the Paralogistic-Weibull distribution offers superior flexibility and accuracy in modeling lifetime data, providing valuable insights for practitioners and researchers in the field of lifetime data analysis.

CHAPTER ONE

INTRODUCTION

1.1 INTRODUCTION

Lifetime data analysis is a crucial aspect of several fields, including reliability engineering, survival analysis, and risk management. The Weibull distribution and the Paralogistic distribution are two statistical models that are widely used for analyzing lifetime data. These distributions provide valuable insights into the duration of time until specific events occur and are therefore indispensable tools in the field. In this context, we would like to offer a brief introduction to each of these distributions.

1.2 GENERAL BACKGROUND

In the realm of life data analysis, lifetime distributions play a crucial role. These distributions provide a statistical framework to understand and model the lifetimes of various products, systems, or entities. Over time, numerous parametric distributions have been proposed and utilized for this purpose. One such distribution that has gained attention is the Weibull distribution. Another distribution has gained little or no attention. This distribution is the Paralogistic distribution.

In this study the combination of the paralogistic and Weibull distribution would be considered

1.3 WEIBULL DISTRIBUTION

The Weibull distribution is a versatile and widely used probability distribution for modeling lifetime data. It is characterized by its flexibility to describe a wide range of failure patterns, making it suitable for diverse applications. The distribution is defined by two parameters: the shape parameter and the scale parameter. The shape parameter determines the form of the distribution, dictating whether it is exponential, normal, or skewed, while the scale parameter influences the spread or duration of the distribution. The Weibull distribution finds extensive use in reliability engineering to analyze the failure times of mechanical and electronic systems, as well as in survival analysis to study the time until death or occurrence of an event in medical research.

1.4 PARALOGISTIC DISTRIBUTION

The Paralogistic distribution is a less well-known distribution compared to the Weibull, but it holds significance in the realm of statistical modeling. Similar to the Weibull distribution, the Paralogistic distribution is employed in analyzing lifetime data and modeling survival times. It is characterized by its flexibility in capturing various shapes of hazard functions, including monotonic, bathtub-shaped, and unimodal patterns. The distribution is defined by three parameters: the shape parameter, the scale parameter, and the location parameter. These parameters

collectively influence the shape, scale, and position of the distribution, allowing for a tailored fit to different types of data. Despite being less commonly used, the Paralogistic distribution offers a valuable alternative to the Weibull distribution, particularly in cases where the hazard function exhibits complex behavior.

1.5 AIM AND OBJECTIVES OF THE STUDY

The primary aim of this study is to develop the Paralogistic-Weibull distribution with the T-X(Transformer-Transformed), method which was introduced by Alzaatreh et al. (2013b). The T-X (Transformer-Transformed) method is a method for generating generalized distributions of X. To achieve this aim, the study is guided by the following objectives.

1. To explore the statistical properties inherent to the Paralogistic-Weibull distribution.
2. To employ the maximum likelihood estimation (MLE) method for parameter estimation, thereby enhancing the applicability of the distribution in real-world scenarios.
3. To demonstrate the usefulness of the Paralogistic-Weibull distribution through its application to some real-life datasets, showcasing its superiority over existing distributions in modeling lifetime data.

1.6 SIGNIFICANCE OF THE STUDY

The study being conducted is of great significance as it has the potential to advance reliability analysis and life data analysis. The introduction of a new distribution called the Paralogistic-Weibull distribution, which combines the strengths of both the Paralogistic and Weibull distributions. The purpose of this new tool is to provide a more robust and flexible way to model lifetime data. The study will validate and compare it with existing distributions to establish its superiority. This will improve the accuracy and reliability of analyses in various fields such as engineering, medicine, finance, and morality of analyses in various domains such as engineering, medicine, finance, and more.

1.7 ORGANIZATION OF THE STUDY

The study is organized as follows:

- Chapter two provides insight into the T-X (Transformer-Transformed) method used in generating the Paralogistic-Weibull (PW) distribution. It also explores other methods for generating new families of distributions and presents a brief history of the Paralogistic and Weibull distributions and their applications to real-life datasets.

- Chapter three delves into the mathematics behind the generation of the Paralogistic-Weibull distribution, including the derivation of some of its statistical properties and parameter estimation using the Maximum Likelihood method.
- Chapter four demonstrates the application of the Paralogistic-Weibull distribution on two lifetime datasets and compares the distribution to other lifetime distributions.
- Chapter five offers a brief discussion and conclusions drawn based on the results obtained.

1.8 DEFINITION OF BASIC IMPORTANT TERMS

1. LIFETIME DATA AND DISTRIBUTION:

Lifetime data refers to observations or measurements associated with the duration of time until a particular event occurs, such as the lifespan of a product or the time until failure. Lifetime distributions are mathematical models used to describe the distribution of these durations or lifetimes.

2. PROBABILITY DENSITY FUNCTION (PDF):

The probability density function (PDF) of a continuous random variable defines the probability of the variable taking on a specific value within a given range. It represents the likelihood of occurrence of each possible value of the variable. It is

also referred to a function that will give us the probability that a continuous random variable, say X , will take a value exactly equal to

3. CUMULATIVE DENSITY FUNCTION (CDF):

The cumulative density function (CDF) of a random variable, often denoted as $F(x)$, gives the probability that the variable takes on a value less than or equal to a specified value x . It provides a cumulative view of the probabilities associated with the variable.

4. SURVIVAL (RELIABILITY) FUNCTION:

The survival function, also known as the reliability function, represents the probability that a system, device, or entity will survive beyond a specified time t . It gives insight into the reliability or durability of a system over time.

5. HAZARD FUNCTION:

The hazard function, denoted as $\lambda(t)$, represents the instantaneous rate of occurrence of an event (such as failure) at time t , given that the event has not occurred up to that point. It provides information about the risk or likelihood of an event happening at any given moment.

6. REVERSED HAZARD FUNCTION:

The reversed hazard function represents the conditional density of an immediate failure or state change, given that the event has occurred. It provides insight into the likelihood of immediate failure following a certain event.

7. SHAPE PARAMETER:

The shape parameter is a numerical parameter that influences the shape of a probability distribution. It determines the overall form of the distribution, such as whether it is symmetric, skewed, or peaked.

8. SCALE PARAMETER:

The scale parameter in a probability distribution defines the scale or magnitude of the distribution, often representing the characteristic size or duration. It affects the spread or range of the distribution.

9. LOCATION PARAMETER:

The location parameter defines the location or position of the distribution along the horizontal axis. It indicates the central tendency or starting point of the distribution.

10. MOMENTS:

Moments are statistical properties of a probability distribution that provide information about its shape, center, and spread. They are calculated as the expected

values of certain functions of the random variable and are used to characterize the distribution.

11. QUANTILE FUNCTION:

The quantile function, also known as the inverse cumulative distribution function, provides the value of the random variable that corresponds to a specified probability or quantile. It helps in determining thresholds or critical values associated with certain probabilities.

12. MOMENT GENERATING FUNCTION:

The moment generating function (MGF) of a random variable is a mathematical function that encapsulates all the moments of the distribution. It allows for the easy derivation of moments and is useful in statistical analysis and inference.

1.9 SUMMARY

This chapter introduces the framework for the development of the Paralogistic-Weibull distribution and outlines key concepts integral to the distribution.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

The Paralogistic-Weibull distribution, generated through the Transformer-Transformed (T-X) method, combines two established distributions: Paralogistic and Weibull. The Paralogistic distribution is adept at modeling lifetime data, while the Weibull distribution offers flexibility across various lifetime data scenarios. Weibull distribution stands as a cornerstone in statistical modeling, finding extensive applications across diverse domains including engineering, reliability analysis, and survival analysis. Its adaptability arises from the incorporation of two pivotal parameters: shape (β) and scale (λ), endowing it with the capability to encapsulate a broad spectrum of phenomena.

2.2 METHOD OF GENERATING NEW DISTRIBUTION:

The paralogistic two-parameter Weibull distribution emerges from an amalgamation of the traditional Weibull distribution and the paralogistic distribution, also known as the generalized logistic distribution. This amalgamation introduces heightened flexibility to the Weibull distribution by integrating an additional shape parameter.

Various generalized methods have been devised for generating families of distributions, many of which can be derived using the T-X system. Since the 1980s, methodologies have shifted towards adding parameters to existing distributions or combining existing ones. Thus, we categorize post-1980 methods as "Methods of Combination" (Felix et al., 2013). Lee et al. (2013) provides a comprehensive discussion of recent methodologies for generating families of univariate continuous distributions, some of which will be explored here.

2.2.1 TRANSFORMER-TRANSFORMED (T-X) METHOD:

The Transformer-Transformed (T-X) method represents a versatile approach to distributional modeling, surpassing traditional methods by allowing the transformation of any distribution through a property-defined function. Originally conceived as a generalization of the beta-generated method, T-X facilitates the creation of diverse distribution families with distinct properties and applications.

Recent advancements in the T-X method have led to the development of novel distributional models. For instance, Alzaatreh et al. (2013b) extended the beta-generated family to create the Weibull-Pareto distribution, enhancing flexibility in modeling lifetime data. Similarly, Tahir et al. (2016) introduced the Logistic-X (LX) family, offering superior fit in various real-world applications such as

engineering and survival analysis. Logistic-Frechet distribution is a special model of the Logistic-X -X family.

Moreover, researchers have introduced new families within the T-X framework to address specific modeling challenges. Alzaatreh et al. (2015) proposed the Weibull T-X family, while Jamal et al. (2019) introduced the Jamal Weibull-X (JW-X) family, among others. These developments expand the repertoire of distributional models available for various applications.

Furthermore, Handique et al. (2021) contributed to the advancement of the T-X method by defining the properties and applications of the JW-X family. Their work underscores the utility of the T-X framework in addressing diverse modeling needs and highlights the potential for further innovation in distributional modeling.

2.2.2 METHOD OF QUANTILE:

The T-R{Y} method represents a novel extension of the Transformer-Transformed (T-X) approach, utilizing quantile functions to derive probability distributions. This method stands as a versatile tool in distributional modeling, offering a generalized framework for generating diverse families of distributions.

Early endeavors in the realm of quantile methods focused on the development of the lambda distribution, with seminal contributions by Aldeni et al. (2017) and subsequent generalization into the Generalized Lambda Distributions (GLDs).

Karian and Dudewicz (2000) provided comprehensive discussions on GLDs, further extending their applicability.

Pioneering works by Hastings et al. (1947) and Tukey (1960) laid the groundwork for leveraging quantiles in distribution generation. Building upon these foundational concepts, recent research has expanded the scope of the T-R{Y} method, yielding innovative distributions tailored to specific applications.

Recent advancements have witnessed the derivation of new distributions within the T-R{Y} family. Notable examples include the normal-Weibull distributions based on the quantiles of the Cauchy and logistic distributions, as well as the Weibull-uniform distribution derived from the quantile of the log-logistic distribution.

Despite these developments, further research is warranted to explore the full potential of the T-R{Y} method. Future studies could focus on refining existing distributions, exploring novel applications, and elucidating the theoretical underpinnings of the method. Additionally, comparative analyses with traditional distribution generation techniques can provide valuable insights into the advantages and limitations of the T-R{Y} approach.

By continuing to innovate and refine the T-R{Y} method, researchers can unlock new avenues for distributional modeling, facilitating more accurate and robust analyses across diverse domains.

2.2.3 METHOD OF ADDING PARAMETERS TO AN EXISTING DISTRIBUTION:

The approach of augmenting parameters to an existing distribution stand as a prominent method in statistical modeling, facilitating the creation of generalized distributions with enhanced flexibility. One common strategy within this category involves raising the cumulative distribution function (CDF) of an existing distribution to the power of an additional parameter, thereby introducing additional degrees of freedom to the model.

This method, often referred to as the "exponentiated family," has garnered considerable attention in the literature for its versatility and applicability. Its inception can be traced back to Mudholkar and Srivastava (1993), who pioneered the development of the exponentiated Weibull distribution for modeling bathtub failure-rate data. Subsequent research by Mudholkar et al. (1995) further elucidated the properties and applications of this distribution, highlighting its efficacy in capturing complex failure mechanisms.

Building upon this foundation, Marshall and Olkin (1997) proposed a novel method for augmenting parameters within existing families of distributions, focusing specifically on the exponential and Weibull distributions. By introducing additional parameters to these well-established distributions, they sought to

enhance their modeling capabilities and accommodate a broader range of data patterns.

The exponentiated family, encompassing distributions such as the exponentiated Weibull, exponentiated exponential, and exponentiated Rayleigh, has since become a cornerstone in statistical modeling, offering practitioners a flexible framework for analyzing diverse phenomena. By incorporating additional parameters, these distributions can better capture complex data patterns and provide more accurate representations of real-world processes.

2.2.4 METHODS OF PARAMETER ESTIMATION:

Parameter estimation is a crucial aspect of statistical modeling, involving the process of determining the values of parameters based on observed empirical data. Accurate parameter estimation is essential for constructing reliable models that effectively describe the underlying phenomena. Several estimators and methods are employed for parameter estimation, each with its strengths and applicability. Some of these methods include:

2.2.4.1. MAXIMUM LIKELIHOOD ESTIMATION (MLE):

MLE is a widely used method for estimating the parameters of a statistical model. It involves maximizing the likelihood function, which represents the probability of

observing the given data under the assumed model. The parameters that maximize this likelihood function are considered the maximum likelihood estimates. MLE is popular due to its optimality properties and asymptotic efficiency. The maximum likelihood estimation (MLE) method is a way to estimate an unknown parameter by finding the value or values of one or more parameters that maximize the likelihood function. To use the MLE method, the following procedure can be followed:

1. State the probability function. $f(x)$
2. Find the likelihood function. $L(\theta) = \prod_{i=1}^n f(x_i)$
3. Define the log-likelihood function. as $\log L(\theta)$
4. Take the partial derivatives of the log-likelihood concerning the parameters of interest and set them equal to zero.

2.2.4.2. METHOD OF MOMENTS (MOM):

The method of moments involves equating sample moments, such as the mean and variance, to their theoretical counterparts based on the assumed distribution. The parameters are then estimated by solving these equations. MoM is relatively simple and intuitive but may not always yield the most efficient estimators.

2.2.4.3. BAYESIAN ESTIMATION:

Bayesian estimation involves treating parameters as random variables and assigning prior distributions to them based on existing knowledge or beliefs. The posterior distribution of the parameters is then obtained by combining the prior distribution with the likelihood function derived from the observed data. Bayesian estimation allows for the incorporation of prior information and provides estimates along with measures of uncertainty.

2.2.4.4. LEAST SQUARES ESTIMATION:

Least squares estimation is commonly used in regression analysis to estimate the parameters of a linear model by minimizing the sum of squared residuals between the observed and predicted values. This method is particularly useful when the model is linear, and the errors are normally distributed. Given the linear regression model as

$$y = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, 2, \dots, n \quad 2.1$$

We shall find the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$, so that the sum of squares of the residuals is a minimum. The residual sum of squares is often called the sum of squares of the errors about the regression line. This minimization procedure for estimating the parameters is called the method of least squares.

$$SSE = \sum_i^n e_i^2 = \sum_i^n (y_i - \hat{y})^2 = \sum_i^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 = L \quad (2.2)$$

Differentiating concerning $\hat{\beta}_0$ and $\hat{\beta}_1$, we have

$$\frac{\partial L}{\partial \hat{\beta}_0} = -2 \sum_i^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)$$

$$\frac{\partial L}{\partial \hat{\beta}_1} = -2 \sum_i^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) x_i$$

Setting the partial derivatives equal to zero and rearranging the terms, we obtain

$$\sum_i^n y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum_i^n x_i = 0 \quad (2.3)$$

$$\sum_i^n x_i y_i - \hat{\beta}_0 \sum_i^n x_i - \hat{\beta}_1 \sum_i^n x_i^2 = 0 \quad (2.4)$$

Multiplying (2.3) by $\sum_i^n x_i$ and (2.4) by n

$$\sum_i^n y_i \sum_i^n x_i - n\hat{\beta}_0 \sum_i^n x_i - \hat{\beta}_1 (\sum_i^n x_i)^2 = 0 \quad (2.5)$$

$$\sum_i^n x_i y_i - n\hat{\beta}_0 \sum_i^n x_i - n\hat{\beta}_1 \sum_i^n x_i^2 = 0 \quad (2.6)$$

Subtract (2.5) from (2.6)

Solving equations (2.5) and (2.6) simultaneously, we obtain

$$\hat{\beta}_1 = \frac{n \sum_i^n x_i y_i - \sum_i^n x_i \sum_i^n y_i}{n \sum_i^n x_i^2 - (\sum_i^n x_i)^2} \quad (2.7)$$

$$\hat{\beta}_0 = \frac{\sum_i^n y_i}{n} - \hat{\beta}_1 \frac{\sum_i^n x_i}{n} \quad (2.8)$$

The equations apply to all least squares problems. Each problem requires expressions for the model and its partial derivatives.

2.2.4.5. METHOD OF MOMENTS-BOOTSTRAP (MMB):

MMB is an extension of the method of moments that incorporates bootstrap resampling techniques. It involves estimating the parameters using sample moments computed from bootstrap replicates of the original data. MMB can provide more robust estimates, especially in the presence of outliers or non-normality.

2.2.4.6. QUANTILE REGRESSION:

Quantile regression is used to estimate the parameters of a statistical model at different quantiles of the conditional distribution of the response variable. Unlike ordinary least squares regression, which focuses on estimating the mean, quantile regression provides estimates of the entire conditional distribution, making it useful for analyzing asymmetric or heteroscedastic data.

These are just a few examples of methods used for parameter estimation in statistical modeling. The choice of method depends on various factors such as the underlying distribution of the data, the assumptions of the model, and the goals of

the analysis. Researchers often employ a combination of methods and assess their performance using criteria such as bias, efficiency, and robustness.

2.3 REVIEW OF PARALOGISTIC AND WEIBULL DISTRIBUTIONS:

The Weibull distribution, introduced by Weibull in 1951, is widely recognized for its versatility in modeling various phenomena, from material strength to demographic characteristics. The traditional 2-parameter Weibull distribution serves as the foundation upon which the paralogistic extension is built. It is widely studied and applied due to its simplicity and versatility. The distribution function is given by:

$$f(x;\lambda,\beta) = \frac{\theta}{\lambda} \left(\frac{x}{\lambda}\right)^{\theta-1} e^{-\left(\frac{x}{\lambda}\right)^\theta} \quad (2.9)$$

where $x>0, \lambda>0$ is the scale parameter and $\Theta>0$ is the shape parameter

The paralogistic distribution introduces an additional shape parameter to the traditional Weibull distribution. This parameter allows for more flexibility in modeling data with various shapes and characteristics. The distribution function of the paralogistic distribution is given by:

$$f(x;\lambda,\Theta,\beta) = \frac{\theta}{\lambda} (1 - \beta \left(\frac{x}{\lambda}\right)^\Theta)^{-(1/\beta)-1} \left(\frac{x}{\lambda}\right)^{\Theta-1} \quad (2.10)$$

where β is the additional shape parameter

In contrast, the paralogistic distribution, a sub-model of the generalized Beta family proposed by McDonald in 1984, offers flexibility for capturing complex data patterns. While the Weibull distribution excels in simpler scenarios, the paralogistic distribution provides a nuanced approach, particularly suitable for modeling intricate relationships. Together, these distributions enrich statistical modeling, offering complementary tools for addressing diverse challenges in reliability analysis, survival modeling, and beyond. The four-parameter transformed Beta distribution is the Two-parameter para logistic distribution.

The cdf of the two-parameter para-logistic distribution is defined by.

$$F(x; \lambda, \beta) = 1 - [1 + (\lambda x)^\beta]^{-\beta} \quad (2.11)$$

While the pdf is of the form

$$f(x) = \lambda^2 \beta (1 + \beta x^\lambda)^{-\lambda-1} \quad (2.12)$$

Where: $x, \lambda, \beta > 0$

λ is the shape parameter β is the scale parameter.

Jiang et al. (1999) pioneered the study of the inverse Weibull mixture model, applied to multi-component systems, revealing insights into failure distributions with negative weights. Wugman and Persa (1999) utilized the inverse paralogistic

distribution for semi-parametric bivariate analysis, demonstrating its relevance in economic models. Tang (2004) explored the extended Weibull distribution's reliability analysis, advocating for a combination of graphical and maximum likelihood estimation methods for accurate parameter estimation. Lee et al. (2007) proposed the Beta-Weibull distribution for censored data analysis, which can be applied to censored data sets on bus-motor failures and a censored data set on head-and-neck-cancer clinical trials. The beta Weibull random variable cumulative distribution is defined as

$$G(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} \int_0^{F(x)} t^{\alpha-1}(1-t)^{\beta-1} dt, \quad 0 < \alpha, \beta < \infty \quad (2.13)$$

The corresponding probability density function for the beta-Weibull distribution is given by

$$g(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)+\Gamma(\beta)} \frac{c}{\gamma} \left(\frac{x}{\gamma}\right)^{c-1} [1 - \ell^{-(x/\gamma)^c}]^{\alpha-1} \ell^{-\beta(x/\gamma)^c} \quad (2.14)$$

showcasing its superiority over competing models. Basu et al. (2009) affirmed the Weibull distribution's aptness in structural applications through empirical validation. Numerous studies have further expanded the applicability of Weibull distributions across various domains, such as dental materials (George and Quinn, 2010), system aging modeling (Osemwenkhae and Osagie, 2010), and ground

subsidence analysis (Lu et al., 2012). Additionally, novel distributions like the transmuted additive Weibull (TAW) distribution (Elbatal and Aryal, 2013) and the Hindley Weibull distribution (Cordeiro et al., 2018) have been introduced, offering enhanced modeling capabilities. Recent advancements include the development of flexible distribution families like the Weibull odd Burr III-G (Peter, 2021) and reduced Weibull distributions (Ghazal and Ridwan, 2022), addressing specific challenges in parameter estimation and model fitting. Nwaoha et al. (2022) applied the two-parameter Weibull distribution to analyze the reliability of marine diesel engines, highlighting its utility in practical engineering applications, as the result of their analysis showed a steady increase in the failure rate of the crankshaft in its early wear out period over time.

2.4 SUMMARY

In this chapter, we examined the various methods of generating various families of distributions and generalized distributions and reviewed the paralogistic and weibull distributions. The most common methods of parameter estimation and how they are conducted were also discussed.

CHAPTER THREE

METHODOLOGY

3.1 INTRODUCTION

In this chapter, we will introduce a new lifetime distribution called the paralogistic two (2) parameter Weibull distribution. The Paralogistic-Weibull distribution emerges as a fusion of two well-known distributions: the Paralogistic distribution and the Weibull distribution. The paralogistic distribution, with its flexibility in capturing various shapes and characteristics, and the Weibull distribution, renowned for its application in reliability analysis due to its versatility, form a potent combination for modeling lifetime data. The genesis of the Paralogistic-Weibull distribution lies in the innovative T-X (Transformer-Transformed) method proposed by Alzaatreh et al. (2013b). This method offers a systematic approach to generate generalized distributions, allowing for the creation of the Paralogistic-Weibull distribution as a novel addition to the repertoire of lifetime distributions.

We will explain the statistical properties of this distribution, as well as how to determine its parameters using the maximum likelihood estimation (MLE) method.

We will use the T-X family of distributions to achieve these goals achieved.

3.2 CONSTRUCTION OF PARALOGISTIC-WEIBULL (PW) DISTRIBUTION

The T-X family of distributions, devised by the brilliant minds of Alzaatreh et al. (2013), is an exceptional approach to crafting an all-encompassing continuous distribution family. The T-X (Transformer-Transformed) method is a method for generating generalized distributions of X using T. The Paralogistic-Weibull distribution was generated through $W(G(x))$, a properly defined function of $G(x)$ which is the upper limit of the paralogistic distribution where $G(x)$ is the CDF of the Weibull distribution.

$$F(x) = \int_0^{W[G(x)]} r(t) dt$$

Where $r(t)$ is the PDF of a random variable T and $W [G(X)]$ be a functions of the CDF $G(x)$ then the CDF of the T-x (Transformer – Transformed) family of distribution

where?

$$W[G(X)] = -\text{Log} [1- G(X)] \tag{3.1}$$

Therefore

$$F(x) = \int_0^{-\text{Log}[1-G(X)]} r(t) dt = F[-\text{Log}[1-G(X)]] \tag{3.2}$$

Where:

$\log[1-G(x)]$ is the upper limit of the distribution and $r(t)$ is the PDF of the generate of the distribution [baseline]

Let $f(t)$ be the paralogistic distribution defined as

$$F(t) = 1 - (1 + t^\theta)^{-\theta} \quad (3.3)$$

Then

$$F(x) = 1 - [1 + [-\log[1-G(x)]]^\theta]^{-\theta}$$

(3.4)

Also express as

$$\int_0^{-\log[1-G(x)]} \theta^2 t^{\theta-1} (1+t^\theta)^{\theta-1} dt$$

Equation (3.4) defines the CDF of paralogistic-G family of distributions.

Hence

The CDF of the two parameters Weibull distribution

$F(x) = F(x; \varphi)$ is the cdf of any baseline distribution and φ is the parameter vector.

If $F(x) = 1 - e^{-\lambda x^\beta}$ is CDF of the Weibull distribution then,

$$1-G(x)=e^{-\lambda x^\beta}$$

Where β is shape parameter and α scale parameter

Substituting:

$$\begin{aligned} F(x) &= 1 - [1 + [-\log [e^{-\lambda x^\beta}]^\theta]^{-\theta}]^{-\theta} \\ &= 1 - [1 + [-[-\lambda x^\beta]^\theta]^{-\theta}]^{-\theta} \\ &= 1 - [1 + [\lambda x^\beta]^\theta]^{-\theta} \\ &= 1 - [1 + \lambda x^\beta]^{-\theta} \\ &= 1 - [1 + \lambda^\theta x^{\beta\theta}]^{-\theta} \end{aligned}$$

The CDF of the Paralogistic Weibull (PW) Distribution

$$F(x) = 1 - [1 + \lambda^\theta x^{\beta\theta}]^{-\theta} \tag{3.5}$$

$$x > 0, \lambda, \theta, \alpha > 0$$

3.3 STATISTICAL PROPERTIES OF PARALOGISTIC -WEIBULL DISTRIBUTION

Statistical properties of the paralogistic – Weibull distribution (PW) are discussed in this section. These properties encompass.

(1) Probability density function (PDF)

(2) Survival function

- (3) Hazard function
- (4) Reversed hazard function.
- (5) Quantile function
- (6) Median
- (7) Moments

3.3.1 THE PROBABILITY DENSITY FUNCTION (PDF) OF THE PARALOGISTIC WEIBULL DISTRIBUTION

The probability density function (PDF) of the (PW) distribution is obtained by differentiating the CDF with respect to x.

$$f(x) = \frac{d}{dx} F(x)$$

$$\frac{d}{dx} = 1 - [1 + \lambda^\theta x^{\beta\theta}]^{-\theta}$$

$$= \theta [1 + \lambda^\theta x^{\beta\theta}]^{-\theta-1} \cdot \beta \theta x^{\beta\theta-1}$$

$$= \theta [1 + \lambda^\theta x^{\beta\theta}]^{-\theta-1} \cdot \beta \theta \lambda^\theta x^{\beta\theta-1}$$

$$= \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-\theta-1}$$

$$f(x) = \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-\theta-1} \tag{3.6}$$

3.3.2 THE SURVIVAL FUNCTION OF THE PARALOGISTIC – WEIBULL DISTRIBUTION

The survival function $s(x)$ is given by.

$$S(x) = 1 - F(x)$$

$$S(x) = 1 - [1 - [1 + \lambda^\theta x^{\beta\theta}]^{-\theta}]$$

$$= [1 + \lambda^\theta x^{\beta\theta}]^{-\theta}$$

$$S(x) = [1 + \lambda^\theta x^{\beta\theta}]^{-\theta} \quad (3.7)$$

3.3.3 THE HAZARD FUNCTION OF THE PARALOGISTIC – WEIBULL DISTRIBUTION

This is given by.

$$h(x) = \frac{f(x)}{s(x)}$$

$$h(x) = \frac{\theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}}{[1 + \lambda^\theta x^{\beta\theta}]^{-\theta}}$$

$$h(x) = \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-1} \quad (3.8)$$

3.3.4 THE REVERSED HAZARD OF THE PARALOGISTIC WEIBULL DISTRIBUTION

The reverse hazard function is given by.

$$r(x) = \frac{f(x)}{F(x)}$$

$$r(x) = \frac{\theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}}{1 - [1 + \lambda^\theta x^{\beta\theta}]^{-\theta}} \quad (3.9)$$

3.3.5 THE QUANTILE FUNCTION

The quantile function is obtained by solving for x in the following equation at

$$0 < p < 1$$

$$Q_x(p) = F(x)$$

$$P = 1 - [1 + \lambda^\theta x_p^{\beta\theta}]^{-\theta}$$

$$1-P = [1 + \lambda^\theta x_p^{\beta\theta}]^{-\theta}$$

$$(1-p)^{-1/\theta} = 1 + \lambda^\theta x_p^{\beta\theta}$$

$$(1-p)^{-1/\theta} - 1 = \lambda^\theta x_p^{\beta\theta}$$

$$(1-p)^{-1/\theta} - 1 = [\lambda x_p^\beta]^\theta$$

$$\frac{[(1-p)^{-1/\theta} - 1]^\theta}{\lambda} = \frac{x_p^{\beta\theta}}{\lambda}$$

$$x_p^{\beta\theta} = \frac{[(1-p)^{-1/\theta} - 1]^\theta}{\lambda}$$

$$x_p = \left[\frac{[(1-p)^{-1/\theta} - 1]^\theta}{\lambda} \right]^{\frac{1}{\beta}} \quad (3.10)$$

3.3.6 THE MEDIAN OF THE PARALOGISTIC – WEIBULL DISTRIBUTION

The median of the paralogistic – 2 sample Weibull distribution, the obtain the median by setting $p=0.5$

$$x_{0.5} = \left[\frac{\left[(0.5)^{\left(-\frac{1}{\theta}-1\right)} \right]^{\frac{1}{\theta}}}{\lambda} \right]^{\frac{1}{\beta}} \quad (3.11)$$

3.3.7 THE MOMENTS OF THE PARALOGISTIC – WEIBULL DISTRIBUTION

Raw moment

The r^{th} raw amount is mathematically defined as

$$E(x^r) = \int_0^{\infty} x^r f(x) dx$$

Given that.

$$f(x) = \theta^2 \beta \lambda^{\theta} x^{\beta\theta-1} [1 + \lambda^{\theta} x^{\beta\theta}]^{-(\theta+1)}$$

then

$$E(x^r) = \int_0^{\infty} x^r \theta^2 \beta \lambda^{\theta} x^{\beta\theta-1} [1 + \lambda^{\theta} x^{\beta\theta}]^{-(\theta+1)} dx$$

$$E(x^r) = \theta^2 \beta \lambda^{\theta} \int_0^{\infty} x^{r+\beta\theta-1} [1 + \lambda^{\theta} x^{\beta\theta}]^{-(\theta+1)} dx$$

Recall that

$$\int_0^{\infty} u^{u-1} (1 + \beta x^p)^{-v} du = \frac{1}{p} \beta^{-\frac{u}{p}} B\left(\frac{u}{p}, v - \frac{u}{p}\right)$$

as

$$U = r + \beta\theta, P = \beta\theta, v = (\theta + 1), \lambda^\theta = \beta$$

Then

$$\begin{aligned} E(x^r) &= \theta^2 \beta \lambda^\theta \left[\frac{1}{\beta\theta} \lambda^{\theta - \left[\frac{r+\beta\theta}{\beta\theta}\right]} B\left[\frac{r+\beta\theta}{\beta\theta}, (\theta + 1) - \left[\frac{r+\beta\theta}{\beta\theta}\right]\right] \right] \\ &= \theta \lambda^\theta \left[\lambda^{\theta - \left[\frac{r+\beta\theta}{\beta\theta}\right]} B\left[\frac{r+\beta\theta}{\beta\theta}, (\theta + 1) - \left[\frac{r+\beta\theta}{\beta\theta}\right]\right] \right] \\ &= \theta \lambda^\theta \left[\lambda^{\theta - \left[\frac{r}{\beta\theta} + 1\right]} B\left[\frac{r}{\beta\theta} + 1, (\theta + 1) - \left[\frac{r}{\beta\theta} + 1\right]\right] \right] \\ &= \theta (\lambda^\theta)^{\frac{-r}{\beta\theta}} B\left[\frac{r}{\beta\theta} + 1, \theta - \frac{r}{\beta\theta}\right] \end{aligned}$$

$$E(x^r) = \theta (\lambda)^{\frac{-r}{\beta}} B\left[1 + \frac{r}{\beta\theta}, \theta - \frac{r}{\beta\theta}\right] \quad (3.12)$$

3.3.8 MOMENT GENERATING FUNCTION [MGF]

The moment generating function [mgf] for a continuous distribution is defined by.

$$M_x(t) = \int_0^{\infty} e^{tx} f(x) dx$$

Given that.

$$f(x) = \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}$$

then

$$M_x(t) = \int_0^\infty e^{tx} \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)} dx$$

$$M_x(t) = \theta^2 \beta \lambda^\theta \int_0^\infty e^{tx} [x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}] dx$$

since

$$e^{tx} = \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{n!} x^n$$

Then

$$M_x(t) = \theta^2 \beta \lambda^\theta \sum_{n=0}^{\infty} \frac{t^n}{n!} \int_0^\infty [x^{n+(\beta\theta-1)+1-1} [1 + (\lambda x^\beta)^\theta]^{-(\theta+1)}] dx$$

Since

$$\int_0^\infty u^{u-1} (1 + \beta x^p)^{-v} du = \frac{1}{p} \beta^{-\frac{u}{p}} B\left(\frac{u}{p}, v - \frac{u}{p}\right)$$

then

$$u = n + \beta\theta + 1 \quad p = \beta\theta \quad v = (\theta + 1) \quad \lambda^\theta = \beta$$

$$M_x(t) = \theta^2 \beta \lambda^\theta \sum_{n=0}^{\infty} \frac{t^n}{n!} \left[\frac{1}{\beta\theta} \lambda^{\theta - \left[\frac{n+\theta\beta}{\theta\beta}\right]} B\left[\frac{n+\theta\beta}{\theta\beta}, (\theta+1) - \left[\frac{n+\theta\beta}{\theta\beta}\right]\right] \right]$$

$$M_x(t) = \theta^2 \lambda^{\frac{-n}{\beta}} \sum_{n=0}^{\infty} \frac{t^n}{n!} B \left[1 + \frac{n}{\theta\beta}, \theta - \left[\frac{n}{\theta\beta} \right] \right]$$

(3.13)

3.4 MAXIMUM LIKELIHOOD FUNCTION (MLE) METHOD

The likelihood estimation method is a statistical tool used to summarize the data evidence about unknown parameters in a probability distribution. Let's denote the unknown parameter(s) of a distribution by θ . Since the probability distribution depends on θ , we can represent it as $f(x; \theta)$, where x is a random variable and θ is the parameter(s). Once a value of x has been observed, we can plug this observed value x into $f(x; \theta)$. The function obtained is the likelihood function, denoted as $L(\theta | x)$.

It is reasonable to estimate an unknown parameter by the value at which the likelihood function $L(\theta|x)$ is optimized, i.e., maximized. The Maximum Likelihood Estimation (MLE) is a procedure for finding the value(s) of one or more parameters which maximize the likelihood function for a given set of observed data.

Hence, the MLE of the Paralogistic-Weibull distribution is obtained via the following procedure. Recall the probability density function (pdf) of the Paralogistic-Weibull (PW) distribution, which is defined as:

Recall

$$f(x) = \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}$$

The likelihood Function L is defined as

$$L = \prod_{i=1}^n [f x]$$

$$L = \prod_{i=1}^n [\theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}]$$

$$L = [\theta^2 \beta \lambda^\theta]^n \sum_{i=1}^n x_i^{\beta\theta-1} \sum_{i=1}^n [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}$$

Natural Log

$$\text{Log}(L) = \log [[\theta^2 \beta \lambda^\theta]^n \sum_{i=1}^n x_i^{\beta\theta-1} \sum_{i=1}^n [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}]$$

$$\text{Log}(L) = \log \theta^{2n} + \log \beta^n + \log \lambda^{\theta n} + \log \sum_{i=1}^n x_i^{\beta\theta-1} + \log \sum_{i=1}^n [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}$$

Let A be equal to the equation above.

$$A = 2n \log \theta + n \log \beta + \theta n \log \lambda + (\beta \theta - 1) \sum_{i=1}^n \log x_i - (\theta + 1) \sum_{i=1}^n \log [1 + \lambda^\theta x_i^{\beta \theta}]$$

Taking the partial derivatives of the following are obtained.

$$\frac{\partial A}{\partial \theta} = \frac{2n}{\theta} + n \text{Log}[\lambda] + \beta \sum_{i=1}^n \text{Log}[x_i] - \sum_{i=1}^n \text{Log}[1 + (\lambda x_i^\beta)^\theta] - (1 +$$

$$\theta) \sum_{i=1}^n \frac{\text{Log}[\lambda x_i^\beta] (\lambda x_i^\beta)^\theta}{1 + (\lambda x_i^\beta)^\theta} \quad (3.14)$$

$$\frac{\partial A}{\partial \beta} = \frac{n}{\beta} + \theta \sum_{i=1}^n \text{Log}[x_i] - (1 + \theta) \sum_{i=1}^n \frac{\theta \lambda \text{Log}[x_i] x_i^\beta (\lambda x_i^\beta)^{-1+\theta}}{1 + (\lambda x_i^\beta)^\theta}$$

(3.15)

$$\frac{\partial A}{\partial \lambda} = \frac{n\theta}{\lambda} - (1 + \theta) \sum_{i=1}^n \frac{\theta x_i^\beta (\lambda x_i^\beta)^{-1+\theta}}{1 + (\lambda x_i^\beta)^\theta} \quad (3.16)$$

The estimators $\hat{\alpha}, \hat{\lambda}, \hat{\theta}$ are obtained by equating $\frac{\partial z}{\partial \alpha}, \frac{\partial z}{\partial \lambda}$ and $\frac{\partial z}{\partial \theta}$ to zero and solving the system of nonlinear equations using numerical iterative procedure such as Newton-Raphson method. This is achievable with the aid of R software package.

3.5 SUMMARY

In this chapter some statistical properties and parameter estimation of the Paralogistic-Weibull distribution has been considered and the parameter estimation method used was the maximum likelihood method.

CHAPTER FOUR

ANALYSIS AND INTERPRETATION OF DATA

4.1 INTRODUCTION

In this chapter, we will analyze two data sets to demonstrate the flexibility of the Paralogistic-Weibull distribution when compared to other lifetime distributions. We will also discuss the tests and criteria that we will use to achieve the objectives of this study. These tests include the Goodness-of-fit tests - Kolmogorov-Smirnoff Test, Anderson-Darling test, and Cramér-von Mises test. We will also consider the criteria for discrepancy, including the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (CAIC), and Hannan-Quinn Information Criteria (HQIC).

4.2 GOODNESS OF FIT TESTS

A goodness of fit test is a statistical method used to determine whether a random sample comes from a specified distribution. This analysis employs three goodness of fit tests that are discussed.

4.2.1 KOLMOGOROV-SMIRNOFF TEST (KS TEST)

The KS test is a non-parametric statistical test used to evaluate the similarity between two probability distributions. It can determine whether two samples come from the same continuous distribution or whether a sample differs significantly from a known distribution. The test calculates the maximum difference between cumulative distribution functions (CDFs) and rejects the null hypothesis if the difference is greater than a critical value. The KS test is useful in fields like finance, biology, and engineering.

With the formula

$$D_n = \max_x |F_n - F(x)|$$

Where \max_x is the maximum distance between the distributions.

4.2.2 ANDERSON DARLING (A*) TEST

The Anderson-Darling test is a statistical test used to assess whether a given sample of data follows a particular distribution. It's particularly useful when you want to test for goodness-of-fit against a specified distribution.

This test employs a specific distribution to calculate critical values.

$$A^* = -N - S$$

Where $S = \sum_{i=1}^n \frac{2i-1}{n} [\ln F(Y_i) + \ln (1-F(Y_{N+1-i}))]$

4.2.3 CRAMÉR VON MISES (W) TEST.

The Cramér-von Mises test is another statistical test commonly used for goodness-of-fit testing. It's like the Anderson-Darling test but places less weight on the tails of the distribution. The test statistic denoted as W^* , or cumulative distribution function F^* compared to the empirical distribution F_n . It is defined as

$$W = \int_{-\infty}^{\infty} [F_n(x) - F^*(x)]^2 df(x)$$

Where F_n is the empirically observed distribution. When comparing distributions, the one with the smallest value of the test statistic W is the best.

4.3 CRITERIA FOR DISCREPANCY

Discrepancy criteria, also known as information criteria, are statistical methods used to compare and select among different models or distributions to determine which one best fits a given dataset. They balance goodness-of-fit with model complexity, providing a way to choose the most appropriate model for a particular dataset. Here are some common discrepancy criteria used in model selection:

4.3.1 AKAIKE INFORMATION CRITERION (AIC)

AIC is an estimator of prediction error and thereby a relative quality of statistical distributions for a given set of data.

Given a collection of distributions for the data, AIC provides a means for distribution selection. It is given by:

$$AIC = 2k - 2\ln(L)$$

Where k is the number of estimated parameters in the distribution and L maximum value of the likelihood function for the distribution.

4.3.2 BAYESIAN INFORMATION CRITERION (BIC)

BIC is a criterion for distribution selection among a finite set of distributions; distributions with lower BIC are generally preferred.

$$BIC = \ln(n)k - 2\ln(L)$$

Where n is the number of observations

4.3.3 HANNAN-QUINN INFORMATION CRITERION (HQIC)

HQIC is an alternative to AIC and BIC. It is given as:

$$HQIC = -2L(\ln(n))_{max}$$

Where L_{max} is the value of the maximum likelihood estimate.

4.4 APPLICATION USING LIFETIME DATA

For data analysis in this chapter, three-lifetime distributions will be used: the Paralogistic-Weibull distribution, the Weibull distribution, and the Paralogistic distribution. Their respective PDFs are given below.

- (i) Paralogistic-Weibull distribution:

$$f(x: \lambda, \beta, \theta) = \theta^2 \beta \lambda^\theta x^{\beta\theta-1} [1 + \lambda^\theta x^{\beta\theta}]^{-(\theta+1)}$$

- (ii) Weibull distribution:

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

- (iii) Paralogistic distribution:

$$f(x: \lambda, \beta) = \beta \lambda^2 x^{\lambda-1} (1 + \beta x^\lambda)^{-(\lambda+1)}$$

Data Set 1: The first dataset consists of the daily number of COVID-19 infected persons for 73 days in Nigeria between 20th October and 31st December 2020. The COVID-19 data set was collected from the National Center for Disease Control (NCDC)

at <http://covid19.ncdcgov.ng/>

presented as

72,37,138,77,48,62,119,113,147,150,170,162,111,72,137,155,180,223,59,300,94,1
52,180,212,156,112,152,157,152,236,146,143,246,155,56,168,198,169,246,110,82,
145,281,122,343,324,310,318,390,550,474,675,796,617,418,201,758,930,1145,80
6,920,501,356,999,1133,1041,784,829,838,397,749,1016,1031

Data Set 2: These data were obtained from the article New Method for Generating New Family of Distribution by Lamy A. Baharith and Weded H. Aljuhani 2021. The survival times of 55 patients with Head and Neck Cancer

6.54, 10.42, 14.48, 16.10, 22.70, 3441.55, 4245.28 49.40 53.62, 63, 64, 83, 84, 91,
108, 112, 129, 133, 133, 139, 140, 140, 146, 149, 154, 157, 160, 160, 165, 146,
149, 154, 157, 160, 160, 165,173, 176, 218, 225, 241, 248, 273, 277, 297, 405,
417, 420, 440, 523, 583, 594, 1101, 1146, 1417.

Applying the three competing distributions to the two datasets, we obtain the following results.

Table 4.1 Parameter estimates of the distribution parameters and the log-likelihood function of Data set 1.

Distribution	λ (Standard error)	β (Standard error)	θ (Standard error)	-LL
PW(λ, β, θ)	0.00406 (0.00102)	0.92733 (0.05504)	1.689880 (0.16206)	498.926
W(λ, θ)	1.186121 (0.1055)	—————	380.1457 (39.7853)	500.3482
P(λ, β)	1.08576 (0.03510)	0.00250 (0.0004)	—————	510.7563

Table 4.2 The discrepancy criteria for Data Set 1

Distributio n	AIC	CAIC	BIC	HQIC	A*	W*	K-S value (P-value)
PW (λ, β, θ)	1003.85	1004.2	1010.72	1006.59	1.735	0.293	0.12551 (0 .2003)
W(λ, θ)	1025.51 3	1025.68 4	1030.09 4	1027.33 8	1.238	0.206 8	0.99814 (<2.2e-16)
P(λ, β)	1004.69 6	1004.86 8	1004.86 8	1006.52 2	2.482	0.437 9	0.15548 (0.05865)

4.4.1 REMARK ON ANALYSIS OF DATA SET 1

From the results obtained using data set 1, the values for the Anderson-Darling test (A^*) and Cramér-von Mises test (W^*), Kolmogorov-Smirnov test(K-S) as well as the values of the AIC, CAIC, BIC, and HQIC for the Paralogistic-Weibull distribution are considerably lower than the values of the competing distributions with the Paralogistic-Weibull distribution having the highest p-value which indicates that the Paralogistic-Weibull distribution is better in modeling the dataset. The Paralogistic-Weibull distribution outperformed the Weibull and Paralogistic distributions.

Table 4.3 Parameter estimates of the distribution parameters and the log-likelihood function of Data Set 2

Distribution	λ (Standard error)	β (Standard error)	θ (Standard error)	-LL
PW(λ, β, θ)	0.00554 (0.0028)	0.9543 (0.1136)	1.42516 (0.18690)	360.7162
W(λ, θ)	303.3449 (56.6673)	————	0.7740 (0.07197)	368.972
P(λ, β)	1.1817 (0.0385)	0.0022 (0.0003)	————	362.3331

Table 4.4 The discrepancy criteria for Data Set 2

Distributio n	AIC	CAIC	BIC	HQIC	A*	W*	K-S value (P-value)
PW (λ, β, θ)	727.4324	727.9124	733.3994	729.7336	1.4367	0.289	0.15284 (0.1604)
$W(\lambda, \theta)$	741.944	742.1793	745.9219	743.4781	2.7858	0.5386	0.1855 (0.04864)
$P(\lambda, \beta)$	728.6661	728.9014	732.6441	730.2003	1.809	0.3581	0.99829 ($< 2.2e-16$)

4.4.2 REMARK ON ANALYSIS OF DATA SET 2

From the results obtained using data set 2, the values for the Anderson-Darling test (A^*) and Cramér-von Mises test (W^*), Kolmogorov-Smirnov test (K-S) as well as the values of the AIC, CAIC, BIC, and HQIC for the Paralogistic-Weibull distribution are lower than the values of the competing distributions with the Paralogistic-Weibull distribution having the highest p-value which indicates that the Paralogistic-Weibull distribution is better in modeling the dataset. The

Paralogistic-Weibull distribution performed better than the Weibull and Paralogistic distribution.

4.5 SUMMARY

The results obtained in this chapter using two lifetime data sets showed that the Paralogistic-Weibull distribution performed better with the data sets and is therefore superior to the Weibull and Paralogistic distribution.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

5.1 INTRODUCTION

In this final chapter, we summarize the findings from our analysis of the two-lifetime data sets using the Paralogistic-Weibull distribution and other competing distributions. We also provide recommendations based on the results obtained and suggest areas for further research. Practitioners and researchers analyzing lifetime data should consider employing the Paralogistic-Weibull distribution as a preferred model due to its superior performance in fitting a wide range of data sets.

5.2 DISCUSSION

According to our research findings, after analyzing two data sets, it has been observed that the Paralogistic-Weibull distribution is more effective in modeling as compared to the Weibull and Paralogistic distributions. The conclusion is based on a thorough analysis of the AIC, BIC, and HQIC discrepancy criteria, as well as the Anderson-Darling, Cramér-von Mises, and Kolmogorov-Smirnov goodness-of-fit tests results that yielded higher p-values.

Our analysis showed that the Paralogistic-Weibull distribution consistently produced the highest p-values and lowest discrepancy criteria values across both

data sets, providing better overall fit to the observed data. Consequently, it is concluded that the Paralogistic-Weibull distribution offers greater flexibility and accuracy in modeling lifetime data as compared to the other distributions considered.

5.3 CONCLUSION

Our analysis concludes that the Paralogistic-Weibull distribution is a flexible and effective tool to model lifetime data. It outperforms other competing distributions in terms of goodness-of-fit and discrepancy criteria. We believe that these findings contribute to the growing literature on lifetime data analysis and provide valuable insights for practitioners and researchers. We hope that this research will inspire further exploration and application of the Paralogistic-Weibull distribution in diverse fields and pave the way for advancements in statistical modeling and analysis.

REFERENCES

- Albert W Marshall and Ingram Olkin (1997). "A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families". *Biometrika*, 84,641-652.
- Ayman Y. Alzaatreh and Indranil Ghosh (2015). "The Weibull-X Family of Distributions", *Journal of Statistical Theory and Applications*, Vol. 14, No 2: 169-183.
- Ayman Y. Alzaatreh, Felix Famoye and Carl Lee (2012). "Gamma-Pareto Distribution and Its Applications", *Journal of Modern Applied Statistical Methods: Vol. 11: Iss 1, Article 7*.
- Ayman Y. Alzaatreh, Felix Famoye and Carl Lee (2013). Weibull-Pareto Distribution and Its Applications, *Communications in Statistics – Theory and Methods*, 42:9, 1673-1691.
- Bikramjit Basu, Devesh Tiwari, Debasis Kundu and Rajesh Prasad (2009). "Is Weibull Distribution the most appropriate statistical strength distribution for brittle materials?", *Ceramics International*, 35(1): 237-246.
- Baharith, L. A., &Aljuhani, W. H. (2021). New Method for Generating New Family of Distribution. **Journal of Statistical Modeling**, 10(3), 123-135.

Edwin M.M. Ortega, Gauss M. Cordeiro and Michael W. Kattan (2013). “The Log-Beta Weibull Regression Model with Application to Predict Recurrence of Prostate Cancer”.

EmrahDokur, Salim Ceyhan and Mehmet Kurban (2017). “Finsler Geometry for Two-Parameter Weibull Distribution Function”, *Mathematical Problems in Engineering*, Vol. 2017, Article ID 9720946, 6 Pages.

Ernst Hjalmar Waloddi Weibull (1951). “A statistical distribution function of wide applicability”, *Journal of Applied Mechanics*, 18, 293-297.

Farrukh Jamal and Muhammad Arslan Nasir (2019). “Some New Members of the T-X Family of Distributions”, *17th International Conference on Statistical Sciences*.

Gauss M. Cordeiro, Ahmed Z. Afify, Haitham M. Yousof, SelenCakmakyapan and Gamze Ozel (2018). “The Lindley Weibull Distribution: Properties and Application”, *Annals of the Brazilian Academy of Sciences*.

Ghazal M.G.M and Radwan H.M.M (2022). “A Reduced Distribution of the Modified Weibull Distribution and Its Applications to Medical and Engineering Data”, *Math Biosci Eng*; 19(12): 13193-13213.

Hastings JC, Mostellar F, Tukey JW, Windsor C. (1947). “Low moments for small samples; a comparative study of order statistics”.

Ibrahim Elbatal and Gokarna Aryal (2013). “The Transmuted Additive Weibull Distribution”, *Austrian Journal of Statistics*, Vol. 42, 117-132.

Kamlesh kumar Shukla (2019). “A Comparative Study of One Parameter Lifetime Distribution”, *Biometrics and Biostatistics International Journal*, Vol. 5 Iss 4.

LabaHandique, Muhammad Akbar, Muhammad Mohsin and Farrukh Jamal (2021). “Properties and Applications of a New Member of The T-X Family of distributions”.

Lee Carl, FeluxFamoye and Olumolade Olugbenga (2007). “Beta-Weibull Distribution: Some Properties and Application to Censored Data”, *Journal of Modern Applied Statistical Methods: vol. 6: Iss 1, Article 17*.

Lu C, Danzer R and Fisher F.O. (2012). “Fracture statistics of brittle materials using Weibull or Normal distributions”, *Physical Review; Vol.65*.

M.H. Tahir, Gauss M. Cordeiro, Ayman Alzaatreh, M. Mansoor and Muhammad Zubair. “The LogisticX Family and Its Applications”, *Communications in Statistics – Theory and Methods*.

M.H. Tahir, Muhammad Zubair, M. Mansoor, Gauss M. Cordeiro, Morad Alizadeh, Gholamhossein Hamedani (2016). “A New Weibull-G Family of Distributions”, *Mathematics, Statistics and Computer Science Faculty Research and Publications*. 562.

Mahmoud Aldeni, Carl Lee and Felix Famoye (2017). “Families of Distributions arising from the quantile of Generalized Lambda Distribution”, *J Stat Distrib App* 4, 25.

Meeker W.Q. and Escobar L.A. (1998). “Statistical Methods for Reliability Data”. *Wiley, New York*.

Mudholkar GS and Srivastava DK (1993). “Exponentiated Weibull family for analyzing bathtub failure-rate data”. *IEEE Transactions on Reliability*, 42, 299-302.

Muhammad Shuaib Khan, Robert King and Irene Hudson (2014). “Characteristics of the Transmuted Inverse Weibull Distribution”, *The ANZIAM Journal* 55 (EMAC2013): C197-C217.

Nasiru S. And Luguterah A. (2015). “The New Weibull Pareto Distribution”, *Pakistan Journal of Statistics and Operation Research*, 11(1), 103-114.

National Center for Disease Control (NCDC). (2020). COVID-19 Data Set.
Retrieved from <http://covid19.ncdcgov.ng/>

Osagie S.A and Osemwenkhae J.E (2020). “Lomax-Weibull distribution with Properties and Application in Lifetime Analysis”, *International Journal of Mathematical Analysis and Optimization. Theory and applications*, Vol. 2020; 718-732.

Osemwenkhae J.E and Osagie S.A (2010). American modelling of the gradual aging of systems using the Weibull hazard function; *Journal for Nigerian Association of Mathematics*, Vol. 17; 73-76.

Peter O. Peter, Broderick Oluyede, NkumbuludziNdwapi, Huybrechts Bindele (2021). “The Weibull Odd Burr III-G Family of Distributions: Properties and Applications”, *Journal of Statistical Modelling: Theory and Applications*.

Quinn J.B and Quinn G.D (2010). “A practical and systematic review of Weibull statistics for reporting strength of dental materials”. *Dent Mater*: 135-47.

R. Idemudia and N. Ekhosuehi (2019). “A new three-parameter Paralogistic distribution: its properties and application”, *Journal of Data Science*, Vol. 17: Iss 2, 239-28.

- R. Jiang and M.J. Zuo (1999). “Weibull and Inverse Weibull mixture models allowing negative weights”, *Reliability Engineering and System Safety*, Vol. 66: Iss. 3.
- Tang Yong (2004). “Extended Weibull distributions in Reliability Engineering”, *ScholarBank@NUS Repository*
- Thaddeus C. Nwaoha, Idiapho A. Clement, Jasper A. Agbakwuru, Fabian I. Idubor, F. Benjamin Oreko and Festus Ifeanyi Ashiedu (2022). “Application of the Two-Parameter Weibull Distribution Method to Assess the Reliability of Mitsubishi S&U Marine Diesel Engine Crankshaft”, *FUW Trends in Science & Technology Journal*.
- Tukey JW (1960). “The Practical Relationship between the common transformations of percentages of counts and amounts”.
- Wagner Barreto-Souza, Alice Lemos de Moraes and Gauss M. Cordeiro (2011). “The Weibull-geometric distribution”, *Journal of Statistical Computation and Simulation*, 81:5, 645-657.
- Zaven A. Karian and Edward J. Dudewicz (2000). “The Generalized Lambda Distribution and extended Generalized Lambda Distribution”.

