# STUDY ON GENERALIZED WEIBULL-RAYLEIGH DISTRIBUTION WITH ITS APPLICATIONS

BY

YOHANNA ALAKU ALOGALA

P13SCMT8098

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES, AHMADU BELLO UNIVERSITY, ZARIA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE DEGREE IN STATISTICS

DEPARTMENT OF STATISTICS,

FACULTY OF PHYSICAL SCIENCES,

AHMADU BELLO UNIVERSITY,

ZARIA, NIGERIA

AUGUST, 2017

## DECLARATION

I declare that the work in this dissertation titled "Study on Generalized Weibull-Rayleigh distribution with its applications" has been carried out by me in the Department of Statistics. The information derived from the literature has been duly acknowledged in the text and a list of references provided. No part of this dissertation was previously presented for another degree or diploma at this or any other institution.

_____ _____

YOHANNA Alaku Alogala Date

## CERTIFICATION

This dissertation titled STUDY ON GENERALIZED WEIBULL-RAYLEIGH DISTRIBUTION WITH ITS APPLICATIONS by YOHANNA Alaku Alogala meets the regulations governing the award of the degree of Master of Science of Ahmadu Bello University, and is approved for its contribution to knowledge and literary presentation.

| | | |
|---|---|---|
| _____ | _____ | _____ |
| Dr. Abubakar Yahaya<br>Chairman, Supervisory committee | Signature | Date |
| _____ | _____ | _____ |
| Prof. O. E Asiribo<br>Member, Supervisory committee | Signature | Date |
| _____ | _____ | _____ |
| External Examiner | Signature | Date |
| _____ | _____ | _____ |
| Dr. H. G. Dikko<br>Head of Department | Signature | Date |
| _____ | _____ | _____ |
| Prof. Kabir Bala<br>Dean, School of Postgraduate Studies | Signature | Date |

## DEDICATION

This work is dedicated to God Almighty and my beloved parents.

## ACKNOWLEDGEMENT

I first of all, like to give praise and thanks to God Almighty for the wisdom, understanding, knowledge and grace to start and end up this study peacefully.

I would like to express my sincere gratitude to my able supervisors Dr. Abubakar Yahaya and Prof. O.E. Asiribo, I appreciate so much the time and effort you invested in my work and prompt attention given to me at all times during this research work. I also like to extend this appreciation to the head of department Dr. H.G. Dikko , my able PG coordinator, Dr. Abubakar Yahaya , other lecturers and staff of the department for the guidance, encouragement and patience through-out my academic study.

My sincere gratitude goes to my parents, brothers, sisters, relatives, friends, and classmates for the moral contributions and supports directly or indirectly are thankfully acknowledged and may God continues to reward you people in abundance.

## ABSTRACT

The Rayleigh distribution enjoys a wide range of applications comprising life testing experiments, reliability analysis, applied statistics and clinical studies. The distribution is a special case of the two parameter Weibull distribution when the shape parameter takes on a value 2. In this thesis work, a Generalized Weibull-Rayleigh Distribution (*GWRD*) is presented. The respective definitions of its probability density function (*pdf*) and cumulative distribution function (*cdf*) using the generalized family proposed by earlier research were provided. Some properties of the new distribution such as moments, moment generating function, characteristics function, quantile function, survival function, hazard function and order statistics were studied. The estimation of the distribution's parameters was conducted using the method of maximum likelihood. The performance of the proposed probability distribution (*GWRD*) was compared to some other generalizations of Rayleigh distribution using three simulated datasets. The results obtained were compared using the values of the *AIC*, *CAIC*, *BIC* and *HQIC* of the fitted distributions and it showed that *GWRD* outperforms other similar distributions comprising Weibull-Rayleigh, Transmuted Rayleigh, Weibull and the conventional Rayleigh distributions.

TABLE OF CONTENTS

# Contents

DECLARATION ..... ii
CERTIFICATION ..... iii
DEDICATION ..... iv
ACKNOWLEDGEMENT ..... v
ABSTRACT ..... vi
LIST OF FIGURES ..... ix
LIST OF TABLES ..... x
CHAPTER ONE: INTRODUCTION ..... 1
1.1 Background to the study ..... 1
1.2 Statement of the problem ..... 2
1.3 Aim and objectives of the study ..... 2
1.4 Significance of the Study ..... 2
CHAPTER THREE: METHODOLOGY ..... 11
3.1 The Development of the Generalized Weibull-Rayleigh Distribution ..... 12
3.2 Some Properties of the Proposed Distribution ..... 17
..... 17
3.2.1 Model validity check
3.2.2 Moments ..... 20
3.2.3 Moment Generating Function (mgf) ..... 23
3.2.4 Characteristics Function ..... 24
3.2.5 Quantile Function ..... 25
3.2.6 Skewness and Kurtosis ..... 26
3.3 Reliability Analysis ..... 27
3.3.1 Survival Function ..... 27
3.3.2 Hazard Function ..... 28
3.4 Order Statistics ..... 30
3.5 Estimation of Parameter ..... 31
CHAPTER FOUR: ANALYSIS AND DISCUSSION ..... 34
CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS ..... 42

5.1 SUMMARY ..... 42
5.2 CONCLUSION ..... 42
5.3 RECOMMENDATIONS ..... 43
5.4 CONTRIBUTION TO KNOWLEDGE ..... 43
5.5 AREAS OF FURTHER RESEARCH ..... 43
REFERENCE ..... 44
APPENDIX ..... 46

## LIST OF FIGURES

Fig. 3.1: The graph of *pdf* of the *GWRD* …….....17

Fig. 3.2: The graph of *cdf* of the *GWRD* ……..…...17

Fig. 3.3: Survival function of the *GWRD*……. ……..…….28

Fig. 3.4: Hazard function of the *GWRD* .....……..…….…....30

Fig. 4.1: A histogram and density plot for data set I……..……..…….36

Fig. 4.2: A histogram and density plot for data set II……..…….……...…….…....37

Fig. 4.3: A histogram and density plot for data set III…….……...……..37

## LIST OF TABLES

Table 4.1: Summary of the dataset I……..……..35

Table 4.2: Summary of the dataset II……..……...……..35

Table 4.3: Summary of the dataset III……...……..36

Table 4.4: Performance of the distribution based on data set I……..…39

Table 4.5: Performance of the distribution based on data set II..…….....40

Table 4.6: Performance of the distribution based on data set III..…….....41

# CHAPTER ONE

# INTRODUCTION

## 1.1 Background to the study

Rayleigh (1980) derived the Rayleigh distribution from the amplitude of sound resulting from many important sources. The Rayleigh distribution has a wide range of applications including life testing experiments, reliability analysis, applied statistics and clinical studies. This distribution is a special case of the two parameter Weibull distribution with the shape parameter equal to 2. The origin and other aspects of this distribution can be found in Siddiqui (1962), Hirano (1986) as well as Howlader and Hossian (1995). Some generalizations of the Rayleigh distribution have led to the development of other distributions such as the generalized Rayleigh distribution by Kundu and Ragah (2005), Bivariate generalized Rayleigh distribution by Abdel-Hady (2013), Transmuted Rayleigh distribution by Merovci (2013) and the Transmuted Inverse Rayleigh distribution by Ahmad *et al*., (2014). These distributions have been found to be more flexible than the Rayleigh distribution when applied to real life data sets.

Many standard distributions for modeling data in different fields of study such as engineering, economics, finance, biological, environmental and medical sciences and so on and so forth have been extensively studied over the past years. However, extending these standard distributions has produced several compound models that are more adequate compared to the classical distributions. For this reason, several methods for generating new families of distributions have been studied. Therefore, this work will introduce a new extension of the Rayleigh distribution using the link function proposed by Cordeiro *et al*. (2015) and study some of its properties with applications to simulated data sets.

## 1.2 Statement of the problem

Researchers have proven that there are many distributions used in statistical modeling of data sets but, some of these data do not follow any of the existing standard probability distributions (models). Therefore this research work seeks to provide an alternative model that will compete favorably with the existing models and produce better fit compared to the other generalizations of the Rayleigh distribution.

## 1.3 Aim and objectives of the study

The aim of this research is to develop a Generalized Weibull-Rayleigh distribution and study some of its properties and applications to real life problems. The stated aim expected to be achieved through the following objectives; by

i. developing a Generalized Weibull-Rayleigh distribution function;
ii. deriving some statistical properties of the proposed distribution such as its' validity, moments, moment generating function, characteristics function, survival function, hazard function, quantile function and distribution of order statistics.
iii. estimating the parameters of the proposed distribution using the method of maximum Likelihood estimation (MLE).
iv. comparing the performance of the proposed model with the Weibull-Rayleigh, transmuted Rayleigh, Rayleigh and Weibull probability models using some datasets used in earlier studies.

## 1.4 Significance of the Study

This dissertation has introduced a new distribution called Generalized Weibull-Rayleigh distribution. It has successfully derived some important properties of the new distribution which have significant applications in various fields especially engineering and medicine. The study has also estimated the parameters of the new distribution as one of the major steps for fitting any

model. This study compared the new model to some existing generalizations of the Rayleigh distribution to identify it as the most fitted model using a real life data set.

# CHAPTER TWO

# LITERATURE REVIEW

In distribution theory, there have been several studies and propositions of compound probability distributions, most of which have been proven to perform better than the classical distributions.

There are so many extensions of the Rayleigh distribution found in literature, however, the purpose of this literature review is not to compare or identify limitations of existing probability models but to understand the theory and methodology behind generalization of probability distributions and be able to extend and study the Rayleigh distribution using an appropriate methodology.

The following are some of the generalization of the Rayleigh distribution and some other compound continuous probability distributions.

Kundu and Ragah (2005) proposed an extension of the Rayleigh distribution called the generalized Rayleigh; by studying some properties of the generalized Rayleigh distribution and concluding that the model performed better than the classical Rayleigh distribution and should be in various areas.

Merovci *et al.,* (2013) also added a single parameter to the Rayleigh distribution using the Quadratic rank transmutation map proposed by Shaw and Buckley (2007). This transmuted parameter introduce skewness into the Rayleigh distribution to prepare it in appropriately modelling skewed datasets. Their analysis also proved that the two parameter distribution (Transmuted Rayleigh Distribution) performed better than the classical Rayleigh distribution.

Ahmad *et al*., (2014) also followed the methodology of Shaw and Buckley (2007) to come up with the transmuted inverse Rayleigh distribution. This extension also performed better than the classical Rayleigh distribution and others based on adequate applications.

Bourguignon *et al.,* (2014) introduced a new wider family of distributions called the Weibull generalized (Weibull-G) family of distributions. They derived general mathematical properties of the new wider Weibull family of distributins. This mixture representation is important to derive several structural properties of this family in full generality. They provided some of this structural properties such as the ordinary and incomplete moments, Quantile function and order statistics. This work also recommended that for any Weibull-G family distributions, the estimation of model parameters should be done by method of maximum likelihood estimation. These authors also concluded that for each baseline distribution G, their methods and results should be adopted.

Merovci and Elbatal (2015) incorporated the Weibull generalized family of distributions developed by Bourguignon *et al.,* (2014) to propose and study the Weibull Rayleigh (WR) distribution. They derived expressions for the moments and the moments generating functions. Estimation of parameters of their model was done by the method of maximum likelihood estimation. Their model performance was compared to some baseline distributions such as beta Weibull, Exponentiated Weibull and the standard Weibull distribution. The results of this comparison using the likelihood ratio statistics showed that the Weibull Rayleigh distribution can be used quite effectively to provide better fit than the Weibull distribution.

Oguntunde *et al.,* (2015) also followed the Weibull-G family of distribution developed by Bourguignon *et al.,* (2014) to propose the Weibull Exponential distribution (*WED*). They also derived explicit expressions for some of its properties like moments, moment generating function, survival function, hazard function, limiting behavior and order statistics. The method of

maximum likelihood estimation was proposed for estimating its parameters and real life applications were provided to illustrate the flexibility of the new distribution over the standard exponential distribution.

Afify *et al.*, (2016) modeled a four parameter distribution called the Weibull Frechet distribution. They studied some of its properties. The applications of the new distribution with some generalized distributions showed that it is more fitted compared to kumaraswamy Frechet (*KFr*), exponentiated Frechet (*EFr*), beta Frechet (*BFr*), gamma extended Frechet (*GEFr*), transmuted marshallOlkin Frechet (*TMOFr*) and Frechet (*Fr*) distributions.

Cordeiro *et al.,* (2015) developed a new generalized Weibull family of distributions with two extra positive parameters. This family was able to generate any continuous probability distributions in full generality. This family also provide explicit expressions for the ordinary and incomplete moments, generating function, mean deviations, probability weighted moments, entropies, reliability and distributions of order statistics. The work stated that, the parameters of this distributions family can be estimated by the method of maximum likelihood estimation (MLEs).

Afify *et al.,* (2016) modeled a new four-parameter model, named the exponentiated Weibull-Pareto (EWP) distribution, which extends the Weibull- Pareto (WP) distribution introduced by Tahir *et al.,* (2015). Some properties of the distribution were provided. Explicit expressions for moments, generating function, R´enyi and q-entropies were also derived. The density function of the order statistics and their moments were also obtained. The proposed distribution was applied to a real dataset and the EWP provides a better fit than several other nested and non-nested models. They, therefore, concluded that the proposed model will attract wider application in

areas such as engineering, survival and lifetime data, meteorology, hydrology, economics and others.

Pescim *et al*., (2015) extended the normal distribution to have the Kummer beta normal distribution. They defined the pdf of the new distribution which allowed them to derive some of its mathematical properties like its ordinary and incomplete moments, mean deviations and order statistics. The estimation of parameters of the model was done by the method of maximum likelihood and Bayesian analysis. The performance of the model was illustrated by an application to a real data set. Their results showed that the Kummer beta normal distribution outperforms all the distributions considered in their study. So they concluded that the proposed distribution can yield better fits than the normal, beta normal, gamma normal, McDonald normal and the modified beta normal distribution and therefore may be an interesting alternative to these distributions for modeling skewed data sets.

Nadarajah *et al*., (2014) also developed a new family of skewed distribution called modified beta distributions. They studied some properties of the new family including estimation procedures. They also studied a member of this family called the modified beta normal distribution. A real data application and simulation studies were used to show superior performance of this new model versus other known models.

Lima *et al*., (2015) developed a new three parameter distribution called the gamma normal distribution which extends the normal distribution.The study provide explicit expressions for the ordinary and incomplete moments, Quantile and generating functions, mean deviations, Renyi entropy, Shannon entropy, order statistics and their moments. They estimated the model parameters using the method of maximum likelihood estimation. This study also derived the observed information matrix of the proposed distribution. The potentiality of the proposed model was evaluated based on three criteria and the results showed that the new distribution provides a

better fit than the beta normal distribution, kumaraswamy-normal and the standard normal distribution.

Alzaatreh *et al*., (2014) developed a new family of distribution named the Gamma – X family. They studied some properties of the Gamma – X family and studied a member of the family; the gamma – normal distribution. They provided the limiting behaviors, moments, mean deviations, dispersions, and Shannon entropy for the gamma-normal distribution. They also obtained bounds for the non-central moments. The method of maximum likelihood estimation was used for estimating the parameters of the gamma-normal distribution. Two real data sets were used to illustrate the application of the gamma-normal distribution. The result proved that the gamma-normal distribution gives a good fit to each data set compared to other existing generalized distributions.

Merovci and Puka (2014) studied a new model called the transmuted Pareto distribution, which extends the Pareto distribution, in the analysis of data with real support. An obvious reason for extending a standard distribution is that the generalized form provides greater flexibility in modeling real life data, they derived expansions for the expectation, variance, moments and the moment generating function. The estimation of parameters was done by the method of maximum likelihood, also the information matrix was given, they considered the likelihood ratio statistics to compare the model with its baseline model. An application of the transmuted Pareto distribution to real data showed that, the new distribution can be used quite effectively to provide better fits than the Pareto distribution.

Cordeiro *et al.,* (2012) proposed a five parameter life time model named the McDonald normal (*MCN*) distribution, which includes as special cases most of the commonly used distribution in the lifetime literature such as the normal, skew-normal, exponentiated normal, beta normal and

kumaraswamy normal distributions, among others. They obtained its ordinary moments, moment generating function and mean deviation. They also derived the ordinary moments of the order statistics. They considered the method of maximum likelihood to fit the new distribution and evaluated its performance with three applications to real data. Their results showed that the new model is much more flexible than the exponentiated normal, beta normal, skew normal and kumaraswamy normal models proposed recently.

Salem *et al., (2014)* proposed the Generalized Weibull-Exponential model. Some properties of GWED such as moments, variance, quantile function, Shannon entropy, skewness and kurtosis were studied. Three real data were fitted for the *GWED* and compared with four known distributions. The results proved that the *GWED* is a relatively better model to fit data than the other four distributions and the *GWED* characterized with highest p-value and skewness among the other. In simulation study, the mean and variance of *GWED* are increasing when the parameters and increase. And are fixed the mean of *GWED* was increase while the mean and variance of *GWED* were decreased.

Kareema *et al*., (2013) developed a new distribution that is dependent on the Exponential and Pareto distribution and proposed some properties such as the moment generated function, mean, mode, median, variance, the r-th moment about the mean, the r-th moment about the origin, reliability, hazard functions, coefficients of variation, skewness and kurtosis. The parameters of the model were estimated by the method of maximum likelihood estimation.

Alzaatreh *et al*., (2012) studied the gamma-Pareto distribution. Many properties of the gamma-Pareto distribution were investigated, such as moments, deviations from the mean and median, hazard function, unimodality, entropies and Fisher information matrix. Results of the uniformly minimum variance unbiased estimator was obtained for one of the shape parameters of the gamma-Pareto distribution. Three real data sets were fitted to the gamma-Pareto distribution and

compared with other known distributions. Results show that the gamma-Pareto distribution provides a good fit to each data set and suggests that the gamma-Pareto distribution can be a good model to fit data with a reversed J-shape, approximately symmetric and long right tail characteristics.

Bourguignon *et al*., (2012) proposed a new model named the Kumaraswany Pareto distribution. They developed and express the *pdf* of the distribution as a linear combination of the exponentiated densities, which allowed them to derive some of its properties like the moments, generating function; they provide the density function of the order statistics and obtain their moments. The method of maximum likelihood was used for estimating the model parameters and the observed information matrix was derived. An application to a real data set showed that the fit of the new model is superior to the fits of its other models.

Mahmoudi (2011) developed the Beta Generalized Pareto (*BGP*) distribution to extend the Pareto distribution. They derived important properties of the new distribution and obtained closed-form expressions for its moments. Applications of the BGP distribution to three real data sets were given to show that the new distribution provides consistently better fits than the Gamma Pareto, Beta Pareto, Weibull and Pareto distributions. Result revealed a better fits over the other distributions.

Therefore, this work will introduce a new extension of the Rayleigh distribution using the link function proposed by Cordeiro *et al*. (2015) and study some of its properties with applications to simulated data sets.

# CHAPTER THREE
# METHODOLOGY

## Introduction

This chapter presents statistical functions and methods that are useful for this research work. These among others include the following:

The Weibull distribution is a widely used statistical model for studying fatigue and endurance life in engineering devices and materials.

If a random variable $X$ has the Weibull distribution with scale parameter $\alpha > 0$ and shape parameter $\beta > 0$, then its cdf and pdf are respectively, given by

$$F(x) = 1 - e^{-\alpha x^{\beta}} \tag{3.1}$$

and

$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}} \tag{3.2}$$

For $x > 0, \alpha > 0, \beta > 0$ where $\alpha$ *and* $\beta$ are the scale and shape parameters respectively.

Similarly, a random variable $X$ is said to have a Rayleigh distribution with parameter $\theta$ if its probability density function *(pdf)* is given by:

$$g(x) = \theta x e^{-\frac{\theta}{2}x^2} \tag{3.3}$$

And the corresponding cumulative distribution function (*cdf*) is given as

$$G(x) = 1 - e^{-\frac{\theta}{2}x^2} \tag{3.4}$$

for $x > 0, \theta > 0$ where $\Theta$ is the scale parameter.

**3.1 The Development of the Generalized Weibull-Rayleigh Distribution.**

Following the methodology of Cordeiro *et al*, (2015), for any distribution with *cdf* $G(x;\eta)$ and pdf $g(x;\eta)=\partial G(x;\eta)/\partial x$, based on a parameter vector, $\eta$ the Generalized Weibull ("GW" for short) family of distributions is defined by the *cdf*

$$F(x)=\int_0^{-\log[1-G(x;\eta)]}\alpha\beta t^{\beta-1}e^{-\alpha t^{\beta}}dt \qquad (3.5)$$

Equation (3.1) yields

$$F(x;\alpha,\beta,\eta)=1-\exp\left\{-\alpha\left(-\log\left[1-G(x;\eta)\right]\right)^{\beta}\right\} \qquad (3.6)$$

The corresponding probability density function (*pdf*) for this generalize Weibull family is given by differentiating (3.6) with respect to x

$$f(x;\alpha,\beta,\eta)=\frac{\alpha\beta g(x;\eta)}{\left[1-G(x;\eta)\right]}\left\{-\log\left[1-G(x;\eta)\right]\right\}^{\beta-1}\exp\left\{-\alpha\left(-\log\left[1-G(x;\eta)\right]\right)^{\beta}\right\} \qquad (3.7)$$

Where $g(x;\eta)$ and $G(x;\eta)$ are the *pdf* and *cdf* of any continuous distribution respectively which depend on parameter vector, $\eta$ $\alpha>0$ and $\beta>0$ are the scale and shape parameters respectively.

A random variable *X* with *pdf* given by equation (3.7) is denoted by *X*~ Generalized Weibull-G $G(x;\alpha,\beta,\eta)$.

Note that equation (3.6) and (3.7) are products of equation (3.5) and the proof is shown as follow below:

**Proof**:

$$F(x) = \int_0^{-\log[1-G(x)]} \alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}}\, dt$$

We now solve the integral above using integration by substitution as follows:

Let

$$u = e^{-\alpha t^{\beta}} = e^{-v}$$

$$v = \alpha t^{\beta}$$

Hence,

$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial v} * \frac{\partial v}{\partial t}$$

$$\frac{\partial u}{\partial v} = -e^{-v} = -e^{-\alpha t^{\beta}}$$

$$\frac{\partial v}{\partial t} = \alpha\beta t^{\beta-1}$$

Which implies that

$$\frac{\partial u}{\partial t} = -\alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}}$$

$$dt = -\frac{du}{\alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}}}$$

Now substituting for *dt* above, we have:

$$F(x) = -\int_0^{-\log[1-G(x)]} du$$

$$F(x) = -\left[u\right]_0^{-\log[1-G(X)]}$$

Recall that

$$u = e^{-\alpha t^{\beta}}$$

$$F(x) = -\left[e^{-\alpha t^{\beta}}\right]_0^{-\log[1-G(X)]}$$

$$F(x) = -\left\{e^{-\alpha(-\log[1-G(X)])^{\beta}} - e^{-\alpha(0)^{\beta}}\right\}$$

$$F(x) = -\left\{e^{-\alpha(-\log[1-G(X)])^{\beta}} - 1\right\}$$

$$F(x) = 1 - e^{-\alpha(-\log[1-G(X)])^{\beta}} \qquad (3.8)$$

Also recall that for any continuous probability distribution,

$$f(x) = \frac{dF(x)}{dx}$$

Hence, let

$$F(x) = y = 1 - e^{-\alpha(-\log[1-G(X)])^{\beta}}$$

We now differentiate y using chain rule as follows:

Let

$$y = 1 - e^{-\alpha(-\log[1-G(X)])^{\beta}} = 1 - e^{-u}$$

$$u = \alpha\left(-\log\left[1 - G(X)\right]\right)^{\beta} = \alpha t^{\beta}$$

$$t = -\log\left[1 - G(X)\right]$$

Hence,

$$\frac{\partial F(x)}{\partial x} = \frac{\partial y}{\partial x} = \frac{\partial y}{\partial u} \times \frac{\partial u}{\partial t} \times \frac{\partial t}{\partial x}$$

$$\frac{\partial y}{\partial u} = e^{-u} = e^{-\alpha(-\log[1-G(X)])^{\beta}}$$

$$\frac{\partial u}{\partial t} = \alpha\beta\left(-\log\left[1 - G(X)\right]\right)^{\beta-1}$$

$$\frac{\partial t}{\partial x} = \frac{g(x)}{1 - G(x)}$$

Which implies that

$$f(x) = \frac{\partial F(x)}{\partial x} = \frac{\partial y}{\partial x} = \frac{\alpha\beta g(x)}{1 - G(x)}\left(-\log\left[1 - G(X)\right]\right)^{\beta-1} e^{-\alpha(-\log[1-G(X)])^{\beta}} \tag{3.9}$$

A random variable X is said to have the Rayleigh distribution with parameter $\Theta$ if its probability density function (*pdf*) is given by

$$g(x) = \theta x e^{-\frac{\theta}{2}x^2} \tag{3.10}$$

And the corresponding cumulative distribution function (*cdf*) is given as

$$G(x) = 1 - e^{-\frac{\theta}{2}x^2} \tag{3.11}$$

for $x > 0, \theta > 0$ where $\Theta$ is the scale parameter.

Using equation (3.10) and (3.11) in (3.6) and (3.7) and simplifying, we obtain the cdf and pdf of the Generalized Weibull-Rayleigh distribution as follows:

$$F(x) = 1 - e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{3.12}$$

and

$$f(x) = \frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{3.13}$$

respectively. Hence equation (3.12) and (3.13) are the cdf and pdf of the Generalized Weibull-Rayleigh distribution.

The *pdf* and *cdf* of the G*WRD* using some parameter values are displayed in **figures 3.1** and **3.2** as follows.

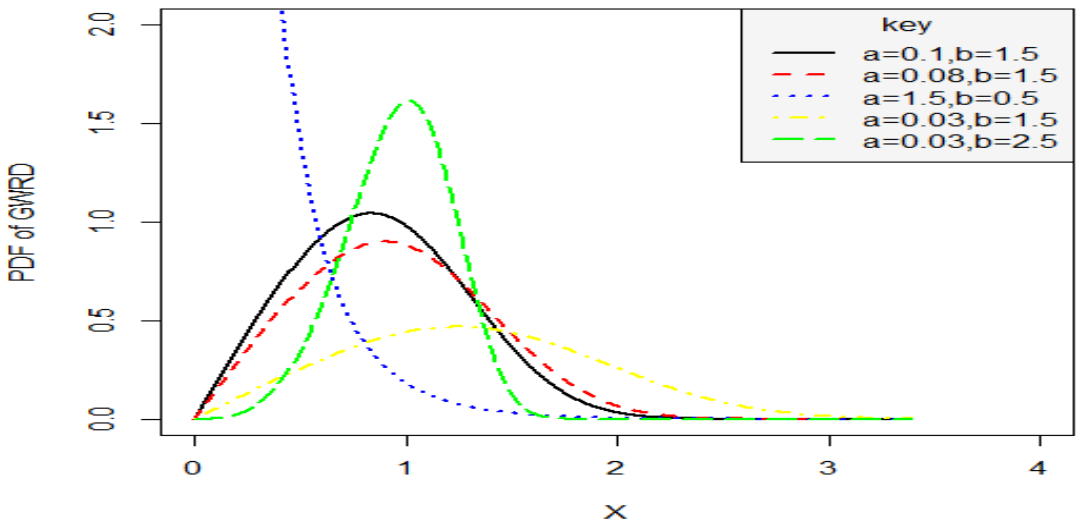



**Figure 3.1:** The graph of *pdf* of the G*WRD* using different parameter values where $a=\alpha, b=\beta \, for \theta=6.5$.

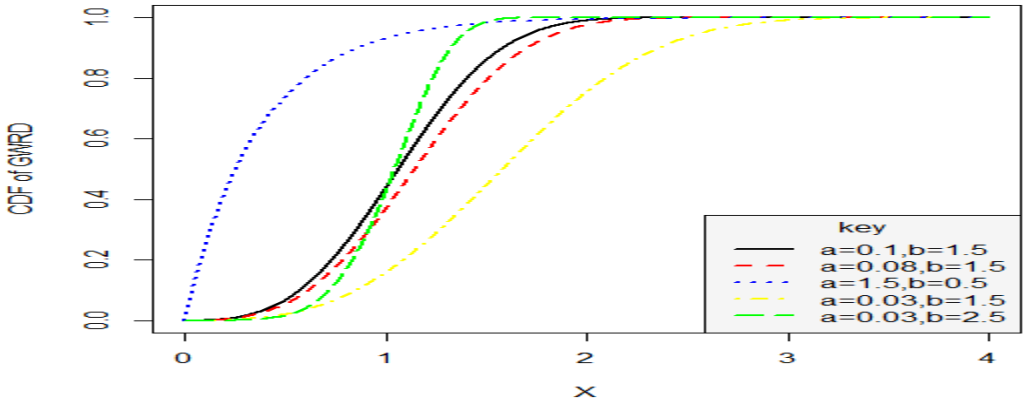



**Figure 3.2:** The graph of *cdf* of the G*WRD* at different parameter values where $a=\alpha, b=\beta \, for \theta=6.5$.

The plot for the *pdf* reveals that the *GWRD* is skewed and therefore will be a good model for modeling skewed data sets.

### 3.2 Some Properties of the Proposed Distribution

#### 3.2.1 Model validity check

$$\int_{-\infty}^{\infty} f(x)dx = 1 \tag{3.14}$$

Recall that for any valid continuous probability distribution, $f(x)$, the relation in (3.14) above holds,

**Proof**

Considering the *pdf* of the Generalized Weibull-Rayleigh distribution, where

$$f(x)=\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

$$\int_0^{\infty} f(x)dx=1$$

$$\int_0^{\infty}\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}dx$$

$$\int_0^{\infty}\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}dx \qquad (3.15)$$

Now let

$$y=e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}=e^{-u}$$

$$u=\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}=\alpha t^{\beta}$$

$$t = -\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]$$

where

$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial u} \times \frac{\partial u}{\partial t} \times \frac{\partial t}{\partial x}$$

$$\frac{dy}{dx} = -\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

$$dx = -\frac{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]dy}{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}}$$

Substituting for *dx* in (3.15) and simplifying the resulting expression, we obtain

$$-\int_0^{\infty} dy = -\left[y\right]_0^{\infty}$$

But recall that

$$y = e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

Hence,

$$\int_0^{\infty} f(x)dx = -\left\{e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}\right\}_0^{\infty}$$

$$\int_0^{\infty} f(x)dx = -\left\{ e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}(\infty)^2}\right)\right]\right)^{\beta}} - e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}(0)^2}\right)\right]\right)^{\beta}} \right\}$$

$$\int_0^{\infty} f(x)dx = -\{0-1\}$$

$$\int_0^{\infty} f(x)dx = 1$$

### 3.2.2 Moments

Let X denote a continuous random variable, the $n^{th}$ moment of X is given by;

$$\mu_n^{'} = E\left(X^n\right) = \int_0^{\infty} x^n f(x)dx \tag{3.16}$$

where *f(x)* the *pdf* of the Generalized Weibull-Rayleigh distribution is as given in equation (3.13).

$$\mu_n^{'} = E\left(X^n\right) = \int_0^{\infty} x^n f(x)dx$$

but

$$f(x) = \frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

Expansion and simplification of the *pdf* gives:

$$f(x) = \frac{\alpha\beta\theta^{\beta}}{2^{\beta-1}} x^{2\beta-1} e^{-\alpha\left(\frac{\theta}{2}x^2\right)^{\beta}} \tag{3.17}$$

Hence,

$$\mu_n^{'} = E\left(X^n\right) = \int_0^{\infty} x^n f(x)dx = \int_0^{\infty} \frac{\alpha\beta\theta^{\beta}}{2^{\beta-1}} x^{n+2\beta-1} e^{-\alpha\left(\frac{\theta}{2}x^2\right)^{\beta}} dx$$

$$\mu_n^{'} = E\left(X^n\right) = \frac{\alpha\beta\theta^{\beta}}{2^{\beta-1}} \int_0^{\infty} x^{n+2\beta-1} e^{-\alpha\left(\frac{\theta}{2}x^2\right)^{\beta}} dx \tag{3.18}$$

Using integration by substitution in (3.18) above we have,

Let

$$y = \alpha\left(\frac{\theta}{2}x^2\right)^{\beta}$$

where

$$\frac{dy}{dx} = \frac{2\alpha\beta\theta^{\beta}x^{2\beta-1}}{2^{\beta}}$$

$$\Rightarrow dx = \frac{2^{\beta}dy}{2\alpha\beta\theta^{\beta}x^{2\beta-1}}$$

Therefore, substituting for *dx* in equation (3.18) and simplifying, we obtain

$$\mu_n^{'} = E\left(X^n\right) = \int_0^{\infty} x^n e^{-y} dy \tag{3.19}$$

Recall from the previous statements that x can be express in terms of y as follows:

$$y = \alpha\left(\frac{\theta}{2}x^2\right)^{\beta}$$

$$\Rightarrow x = \frac{2^{\frac{1}{2}}y^{\frac{1}{2\beta}}}{\alpha^{\frac{1}{2\beta}}\theta^{\frac{1}{2}}} = \left(\frac{y}{\alpha}\right)^{\frac{1}{2\beta}}\left(\frac{2}{\theta}\right)^{\frac{1}{2}}$$

Substituting for x in (3.19), we have:

$$\mu_n^{'} = E\left(X^n\right) = \int_0^{\infty} \left( \left(\frac{y}{\alpha}\right)^{\frac{1}{2\beta}} \left(\frac{2}{\theta}\right)^{\frac{1}{2}} \right)^n e^{-y} dy$$

$$\mu_n^{'} = E\left(X^n\right) = \left( \left(\frac{1}{\alpha}\right)^{\frac{n}{2\beta}} \left(\frac{2}{\theta}\right)^{\frac{n}{2}} \right) \int_0^{\infty} y^{\frac{n}{2\beta}} e^{-y} dy$$

Note that:

$$\int_0^{\infty} y^{\frac{n}{2\beta}} e^{-y} dy = \int_0^{\infty} y^{\frac{n}{2\beta}+1-1} e^{-y} dy = \Gamma\left(\frac{n}{2\beta}+1\right) \text{because } \Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy$$

Hence,

$$\mu_n^{'} = E\left(X^n\right) = \left(\frac{1}{\alpha}\right)^{\frac{n}{2\beta}} \left(\frac{2}{\theta}\right)^{\frac{n}{2}} \Gamma\left(\frac{n}{2\beta}+1\right) \tag{3.20}$$

**The Mean**

The mean of the *GWRD* can be obtained from the $n^{th}$ moment of the distribution when $n=1$ as follows:

$$E\left(X^n\right) = \left(\frac{1}{\alpha}\right)^{\frac{1}{2\beta}} \left(\frac{2}{\theta}\right)^{\frac{1}{2}} \Gamma\left(\frac{1}{2\beta}+1\right) \tag{3.21}$$

Also the second moment of the *GWRD* is obtained from the $n^{th}$ moment of the distribution when $n=2$ as

$$E\left(X^2\right) = \left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}} \left(\frac{2}{\theta}\right) \Gamma\left(\frac{1}{\beta}+1\right) \tag{3.22}$$

**The Variance**

The $n^{th}$ central moment or moment about the mean of $X$, say $\mu_n$, can be obtained as

$$\mu_n = E\left(X - \mu_1^{'}\right)^n = \sum_{i=0}^{n} (-1)^i \binom{n}{i} \mu_1^{'i} \mu_{n-i}^{'} \tag{3.23}$$

The variance of *X* for *GWRD* is obtained from the central moment when *n=2*, that is,

$$Var(X) = E\left(X^2\right) - \left\{E\left(X\right)\right\}^2 \tag{3.24}$$

$$Var(X) = \left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}} \left(\frac{2}{\theta}\right) \Gamma\left(\frac{1}{\beta}+1\right) - \left\{\left(\frac{1}{\alpha}\right)^{\frac{1}{2\beta}} \left(\frac{2}{\theta}\right)^{\frac{1}{2}} \Gamma\left(\frac{1}{2\beta}+1\right)\right\}^2$$

$$Var(X) = \left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}} \left(\frac{2}{\theta}\right) \left\{\Gamma\left(\frac{1}{\beta}+1\right) - \left[\Gamma\left(\frac{1}{2\beta}+1\right)\right]^2\right\} \tag{3.25}$$

The variation, skewness and kurtosis measures can also be calculated from the non-central moments using some well-known relationships.

### 3.2.3 Moment Generating Function (mgf)

This is a clever way of organizing all the moments into one mathematical object, and that object is called the *mgf*. In other words, the *mgf* generates the moments of *X* by differentiation i.e., for any real number say *k*, the $k^{th}$ derivative of $M_X(t)$ evaluated at $t = 0$ is the $k^{th}$ moment $\mu_k^{'}$ of *X*.

The *mgf* of a random variable *X* can be obtained by

$$M_x(t) = E\left(e^{tx}\right) = \int_0^{\infty} e^{tx} f(x) dx \tag{3.26}$$

$$M_x(t) = E\left(e^{tx}\right) = \int_0^{\infty} e^{tx} \left(\frac{\alpha\beta\theta^{\beta}}{2^{\beta-1}} x^{2\beta-1} e^{-\alpha\left(\frac{\theta}{2}x^2\right)^{\beta}}\right) dx$$

$$M_x(t) = E\left(e^{tx}\right) = \frac{\alpha\beta\theta^{\beta}}{2^{\beta-1}} \int_0^{\infty} x^{2\beta-1} e^{tx} e^{-\alpha\left(\frac{\theta}{2}x^2\right)^{\beta}} dx = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mu_n^{'} \tag{3.27}$$

where

$$\mu_n^{'} = E\left(X^n\right) = \left(\frac{1}{\alpha}\right)^{\frac{n}{2\beta}} \left(\frac{2}{\theta}\right)^{\frac{n}{2}} \Gamma\left(\frac{n}{2\beta}+1\right)$$

is as defined in equation (3.20) previously.

### 3.2.4 Characteristics Function

The characteristics function has many useful and important properties which give it a central role in statistical theory. Its approach is particularly useful for generating moments, characterization of distributions and in analysis of linear combination of independent random variables.

The characteristics function of a random variable $X$ is given by;

$$\varphi_x(t) = E\left(e^{itx}\right) = E\left[\cos(tx) + i\sin(tx)\right] = E\left[\cos(tx)\right] + E\left[i\sin(tx)\right] \qquad (3.28)$$

Recall from power series expansion that

$$\cos(tx) = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} x^{2n}$$

$$E\left[\cos(tx)\right] = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \mu'_{2n}$$

$$E\left(x^{2n}\right) = \mu'_{2n}$$

Since

And also that

$$\sin(tx) = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} x^{2n+1}$$

$$E\left[\sin(tx)\right] = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} \mu'_{2n+1}$$

Simple algebra and power series expansion proves that

$$\phi_x(t)=\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n}}{(2n)!}\mu'_{2n}+i\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{(2n+1)!}\mu'_{2n+1} \tag{3.29}$$

Where $\mu'_{2n}$ and $\mu'_{2n+1}$are the moments of $X$ for $n=2n$ and $n=2n+1$ respectively and can be obtained from $\mu'_n$ as

$$\mu'_{2n}=E\left(X^{2n}\right)=\left(\frac{1}{\alpha}\right)^{\frac{n}{\beta}}\left(\frac{2}{\theta}\right)^{n}\Gamma\left(\frac{n}{\beta}+1\right)$$

and

$$\mu'_{2n+1}=E\left(X^{2n+1}\right)=\left(\frac{1}{\alpha}\right)^{\frac{2n+1}{2\beta}}\left(\frac{2}{\theta}\right)^{\frac{2n+1}{2}}\Gamma\left(\frac{2n+1}{2\beta}+1\right)$$

respectively.

Thus

$$\phi_x(t)=\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n}}{(2n)!}\left(\frac{1}{\alpha}\right)^{\frac{n}{\beta}}\left(\frac{2}{\theta}\right)^{n}\Gamma\left(\frac{n}{\beta}+1\right)+$$

$$i\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{(2n+1)!}\left(\frac{1}{\alpha}\right)^{\frac{2n+1}{2\beta}}\left(\frac{2}{\theta}\right)^{\frac{2n+1}{2}}\Gamma\left(\frac{2n+1}{2\beta}+1\right)$$

### 3.2.5 Quantile Function

The Quantile function plays a very important role in the calculation of moments of random variables such as skewness and kurtosis. It is also used to obtain the median and for simulation of random numbers. Mathematically, it is obtained by inverting of the *cdf*.

Hence, quantile function, say *X=Q(u),* of the *GWRD* can be obtained by inverting its' cdf as follows:

$$F(x)=1-e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

Let *F(x)=u* such that the quantile function *Q(u)* is given as

$$F(x)=u$$

$$Q(u)=x=F^{-1}(u)$$

$$1-e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}=u$$

$$1-u=e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

$$1-u=\frac{1}{e^{\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}}$$

$$e^{\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}=\frac{1}{1-u}$$

$$\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}=\ln\left(\frac{1}{1-u}\right)$$

$$\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]=-\left(\frac{1}{\alpha}\ln\left(\frac{1}{1-u}\right)\right)^{\frac{1}{\beta}}$$

$$e^{-\frac{\theta}{2}x^2}=\exp\left\{-\left(\frac{1}{\alpha}\ln\left(\frac{1}{1-u}\right)\right)^{\frac{1}{\beta}}\right\}$$

$$\frac{\theta}{2}x^2=\left(\frac{1}{\alpha}\ln\left(\frac{1}{1-u}\right)\right)^{\frac{1}{\beta}}$$

$$x^2=\frac{2}{\theta}\left(\frac{1}{\alpha}\ln\left(\frac{1}{1-u}\right)\right)^{\frac{1}{\beta}}$$

$$X=Q(u)=\sqrt{\frac{2}{\theta}\left(\frac{1}{\alpha}\ln\left(\frac{1}{1-u}\right)\right)^{\frac{1}{\beta}}} \qquad (3.30)$$

Hence, the median of *X* from the *GWRD* is simply $X_{\frac{1}{2}}=Q(1/2)$ is derived by setting *u*=0.5 in equation (3.30). Furthermore, the lower and the upper quartile can also be derived from (3.30) by setting *u*=0.25 and *u*=0.75 respectively.

### 3.2.6 Skewness and Kurtosis

In this dissertation, the quantile based measures of skewness and kurtosis will employed due to non-existence of the classical measures in some cases. The Bowley's measure of skewness (Kennedy and Keeping, 1962.) based on quartiles is given by;

$$SK = \frac{Q\left(\frac{3}{4}\right) - 2Q\left(\frac{1}{2}\right) + Q\left(\frac{1}{4}\right)}{Q\left(\frac{3}{4}\right) - Q\left(\frac{1}{4}\right)} \tag{3.31}$$

And the Moores' (1998) kurtosis is on octiles and is given by;

$$KT = \frac{Q\left(\frac{7}{8}\right) - Q\left(\frac{5}{8}\right) - Q\left(\frac{3}{8}\right) + \left(\frac{1}{8}\right)}{Q\left(\frac{6}{8}\right) - Q\left(\frac{1}{4}\right)} \tag{3.32}$$

## 3.3 Reliability Analysis

### 3.3.1 Survival Function

Survival function is the likelihood that a system or an individual will not fail after a given time.

Mathematically, the survival function is given by:

$$S(x) = 1 - F(x) \tag{3.33}$$

Where $F(x)$ is *cdf* of the proposed Generalized Weibull-Rayleigh distribution.

$$F(x) = 1 - e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

$$S(x) = 1 - \left\{1 - e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}\right\}$$

$$S(x) = e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{3.34}$$

A plot for the survival function of the *GWRD* using different parameter values is as shown in Figure **3.3** below;

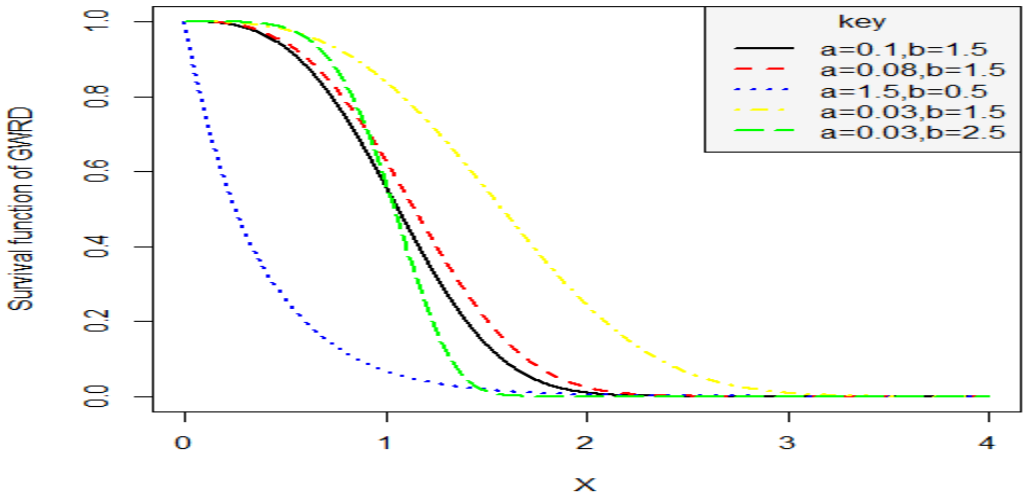



**Figure. 3.3**: Survival function of the *GWRD* at different parameter values where $a=\alpha, b=\beta \, for \theta = 6.5.$

The graph in figure **3.3** shows that the value of the survival function equals one (1) at initial time or early age and it decreases as*X* increases and equals zero (0) as *X* becomes larger.

### 3.3.2 Hazard Function

Hazard function is the probability that a component will fail or die for an interval of time. The hazard function is defined as;

$$h(x) = \frac{f(x)}{1-F(x)} \tag{3.35}$$

where *f(x)* and *F(x)* the *pdf* and *cdf* of the proposed Generalized Weibull-Rayleigh distribution are given as

$$f(x) = \frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

$$f(x) = \alpha\beta\theta x\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

and

$$F(x)=1-e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

respectively.

Substituting for *f(x)* and *F(x)* and simplifying gives

$$h(x)=\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}}$$

$$h(x)=\alpha\beta\theta x\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} \tag{3.36}$$

The following are some possible curves for the hazard rate at various values of the model parameters

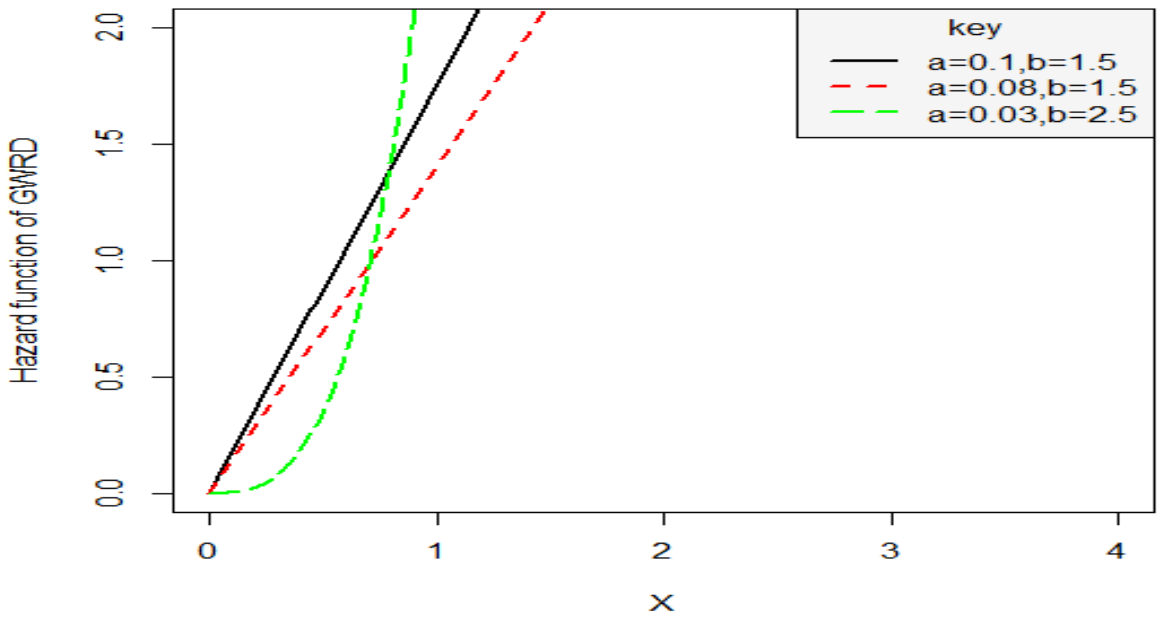


**Figure3.4**: Hazard function of the *GWRD* at different parameter values where $a=\alpha, b=\beta \, for \theta = 6.5.$

We can see from figure 3.4 that the value of the hazard function increases as*X* increases. It gets higher as the value of *X* increases. This means that the *GWRD* may be appropriate for modeling time dependent events, where risk or hazard increases with time or age.

### 3.4 Order Statistics

Sample values such as the smallest, largest, or middle observation from a random sample provide important information. For example, the highest rainfall, flood or minimum temperature recorded during past years might be useful when planning for future emergencies. Let $X_{(1)}$ denote the smallest of $X_1, X_2, ...., X_n$, $X_{(2)}$ denote the second smallest of $X_1, X_2, ...., X_n$, and similarly $X_{(i)}$ denote the $i^{th}$ smallest of $X_1, X_2, ...., X_n$.Then the random variables $X_{(1)}, X_{(2)}, ...., X_{(n)}$, called the order statistics of the sample $X_1, X_2, ...., X_n$, has probability density function of the $i^{th}$ order statistic, $X_{(i)}$, as:

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} f(x) F(x)^{i-1} \left[1 - F(x)\right]^{n-i} \tag{3.37}$$

Where $f(x)$ and $F(x)$ are the *pdf* and *cdf* of the proposed distribution respectively.

Using (3.8) and (3.9), the *pdf* of the $i^{th}$ order statistics $X_{i:n}$, can be expressed from (3.37) as;

$$f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} \left[ \alpha\beta\theta x \left( -\log\left[ 1 - \left( 1 - e^{-\frac{\theta}{2}x^2} \right) \right] \right)^{\beta-1} e^{-\alpha\left( -\log\left[ 1-\left( 1-e^{-\frac{\theta}{2}x^2} \right) \right] \right)^{\beta}} \right]$$
$$\left[ 1 - e^{-\alpha\left( -\log\left[ 1-\left( 1-e^{-\frac{\theta}{2}x^2} \right) \right] \right)^{\beta}} \right]^{i+k-1} \tag{3.38}$$

Hence, the *pdf* of the minimum order statistic $X_{(1)}$ and maximum order statistic $X_{(n)}$ of the *GWRD* are given by;

$$f_{1:n}(x) = n\sum_{k=0}^{n-1}(-1)^k\binom{n-1}{k}\left[\alpha\beta\theta x\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}\right] X$$

$$\left[1-e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}\right]^k \tag{3.39}$$

and

$$f_{n:n}(x) = n\left[\alpha\beta\theta x\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}\right] X$$

$$\left[1-e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}\right]^{n-1} \tag{3.40}$$

respectively.

### 3.5 Estimation of Parameter

Let $X_{1,\ldots,}X_n$ be a sample of size 'n' independently and identically distributed random variables from the *GWRD* with unknown parameters *α, β, a,* and *b* defined previously. The *pdf* of the *GWRD* is given as

$$f(x) = \frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}}$$

The likelihood function is given by;

$$L(X_1, X_2, \ldots\ldots, X_n/\alpha,\beta,\theta) = \frac{(\alpha\beta\theta)^n\sum_{i=1}^{n}x_i}{\sum_{i=1}^{n}\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]} e^{-\frac{\theta}{2}\sum_{i=1}^{n}x^2}\sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1}$$

$$\times \exp \sum_{i=1}^{n} \left\{ -\alpha \left( -\log \left[ 1-\left( 1-e^{-\frac{\theta}{2}x^2} \right) \right] \right)^{\beta} \right\} \tag{3.41}$$

Let the log-likelihood function, $l(\varphi) = \log L(X_1, X_2, ....., X_n / \alpha, \beta, \theta)$, therefore

$$l(\varphi) = n\log\alpha + n\log\beta + n\log\theta + \sum_{i=1}^{n}\log x_i - \frac{\theta}{2}\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n}\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]$$
$$+(\beta-1)\sum_{i=1}^{n}\log\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right) - \alpha\sum_{i-1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta} \tag{3.42}$$

Differentiating $l(\varphi)$ partially with respect to $\alpha$, $\beta$ and $\theta$ respectively gives;

$$\frac{\partial l(\varphi)}{\partial \alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta} \tag{3.43}$$

$$\frac{\partial l(\varphi)}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^{n}\log\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)$$
$$-\alpha\sum_{i=1}^{n}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}\log\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right) \tag{3.44}$$

$$\frac{\partial l(\phi)}{\partial \theta} = \frac{n}{\theta} + \frac{(\beta-1)}{2}\sum_{i=1}^{n}\left\{\frac{x_i^2 e^{-\frac{\theta}{2}x^2}}{\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)\left(1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right)}\right\}$$
$$-\frac{\alpha\beta}{2}\sum_{i=1}^{n}\left\{\frac{\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1}\left(x_i^2 e^{-\frac{\theta}{2}x^2}\right)}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\right\} \tag{3.45}$$

The solution of the non-linear system of equations of $\frac{dl}{d\alpha} = 0$, $\frac{\partial l}{\partial \beta} = 0$, and $\frac{\partial l}{\partial \theta} = 0$ will give us the maximum likelihood estimates of parameters $\alpha, \beta$ and $\theta$. However, the solution cannot be gotten analytically except numerically with the aid of suitable statistical software like R, SAS, e.t.c when data sets are available. Hence, we used the optimization (optim) function in R software to estimate the parameters of the model in the process of applications.

# CHAPTER FOUR

# ANALYSIS AND DISCUSSION

This chapter presents three simulated datasets, their descriptive statistics, graphics and applications to some selected generalizations of the Rayleigh distribution. We have compared the adequacy of the *GWRD* to those of three generalizations of the Rayleigh model including the Weibull-Rayleigh distribution (*WRD*), the Transmuted Rayleigh distribution (*TRD*), Weibull distribution (*WD*) and the Rayleigh distribution (*RD*). The density functions of these distributions are given as follows;

**The Generalized Weibull-Rayleigh Distribution (*GWRD*)**

The *pdf* of the *GWRD* distribution is given as;

$$f(x;\theta,\alpha,\beta)=\frac{\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]}\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta-1} e^{-\alpha\left(-\log\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]\right)^{\beta}} \tag{4.1}$$

**The Weibull-Rayleigh Distribution (*WRD*)**

The *pdf* of the *WRD* is given as;

$$f(x;\alpha,\beta,\mu,\sigma)=\alpha\beta\theta x e^{-\frac{\theta}{2}x^2}\frac{\left[1-e^{-\frac{\theta}{2}x^2}\right]^{\beta-1}}{\left[1-\left(1-e^{-\frac{\theta}{2}x^2}\right)\right]^{\beta+1}}\exp\left\{-\alpha\left[\frac{1-e^{-\frac{\theta}{2}x^2}}{1-\left(1-e^{-\frac{\theta}{2}x^2}\right)}\right]^{\beta}\right\} \tag{4.2}$$

**The Transmuted Rayleigh Distribution (*TRD*)**

The *pdf* of the *TRD* is given by;

$$f(x;\theta,\lambda)=\theta x e^{-\frac{\theta}{2}x^2}\left[1+\lambda-2\lambda\left(1-e^{-\frac{\theta}{2}x^2}\right)\right] \tag{4.3}$$

**The Rayleigh Distribution (*RD*)**

The *pdf* of the Rayleigh distribution is given by

$$f(x,\theta) = \theta x e^{-\frac{\theta}{2}x^2} \tag{4.4}$$

**The Weibull Distribution (*WD*)**

The *pdf* of the Weibull distribution is given by

$$f(x) = \alpha\beta x^{\beta-1} e^{-\alpha x^{\beta}} \tag{4.5}$$

We also considered three simulated data sets for fitting the above selected models and the descriptive statistics for these data sets are also provided in Tables **4.1, 4.2** and **4.3** as follows;

**Data set I**: This data set represents 25 identically and independently distributed random samples from the proposed distribution using its quantile function at some selected values of the parameters. Its summary is given as follows:

**Table 4.1: Summary Statistics for data set I**

| parameters | N | Minimum | $Q_1$ | Median | $Q_3$ | Mean | Maximum | Variance | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Dataset I | 25 | 0.5528 | 1.1940 | 1.5868 | 1.8600 | 1.5691 | 2.6015 | 0.2882 | 0.0743 | -0.6851 |

**Data set II**: This data set represents 75 random samples from the Generalized Weibull-Rayleigh distribution obtained by using the quantile function derived from the distribution. These sample values are summarized as follows:

**Table 4.2: Summary Statistics for Data set II**

| parameters | n | Minimum | $Q_1$ | Median | $Q_3$ | Mean | Maximum | Variance | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Dataset II | 75 | 0.5036 | 1.1310 | 1.5694 | 1.9820 | 1.5978 | 3.3210 | 0.3349 | 0.2134 | -0.1689 |

**Data set III**: The third dataset represents 125 samples simulated from the generalized Weibull-Rayleigh distribution using the quantile function of the distribution. The descriptive statistics for this data are as follows:

**Table 4.3: Descriptive statistics for data set III**

| parameters | n | Minimum | $Q_1$ | Median | $Q_3$ | Mean | Maximum | Variance | Skewness | Kurtosis |
|---|---|---|---|---|---|---|---|---|---|---|
| Dataset III | 125 | 0.3986 | 1.1080 | 1.5021 | 1.9990 | 1.5661 | 3.1115 | 0.3295 | 0.2656 | -0.5144 |

We also provide some histograms and densities for the three simulated data sets as shown in **Figures 4.1, 4.2** and **4.3** below respectively.

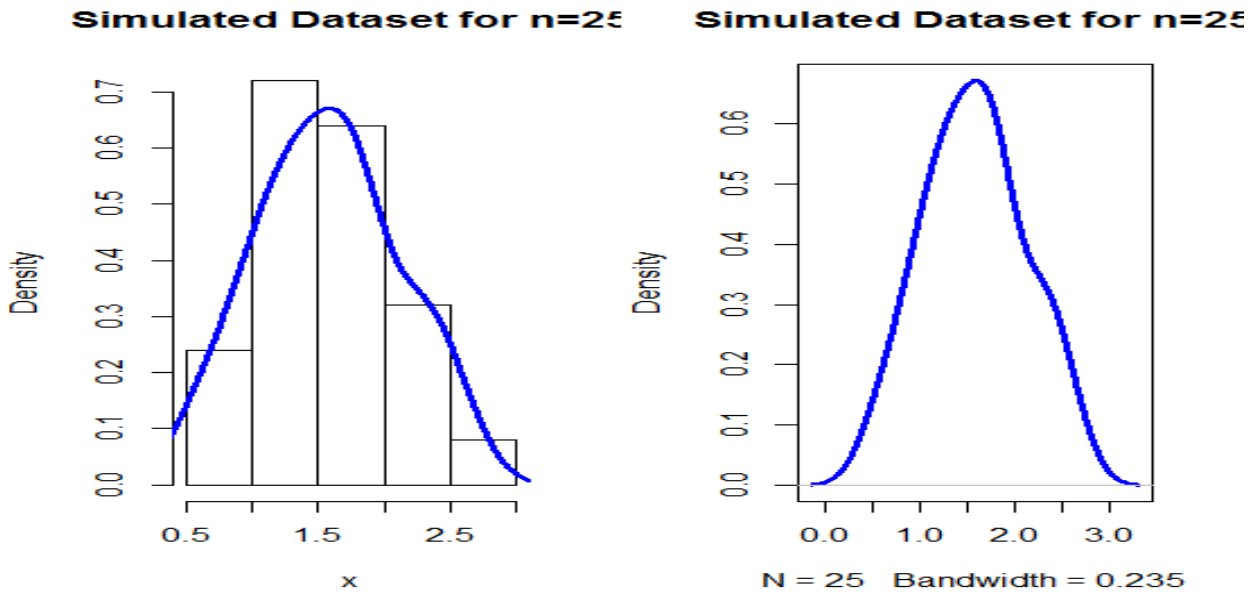



**Figure 4.1: A histogram and density plot for the simulated data when n=25 (Data set I)**

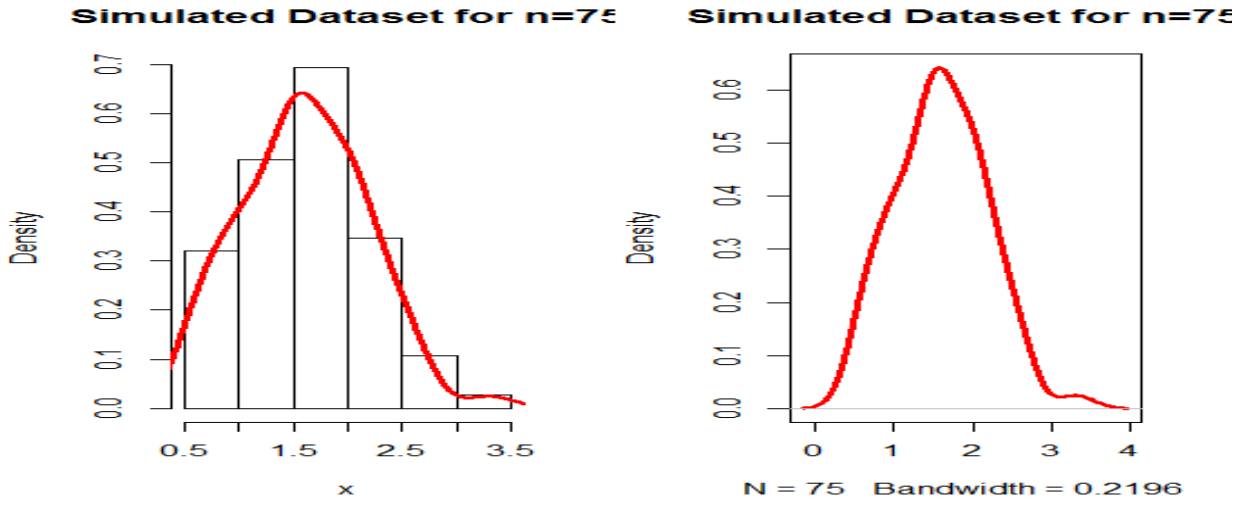



**Figure 4.2: A Histogram and density plot for Simulated data when n=75 (Data set II)**

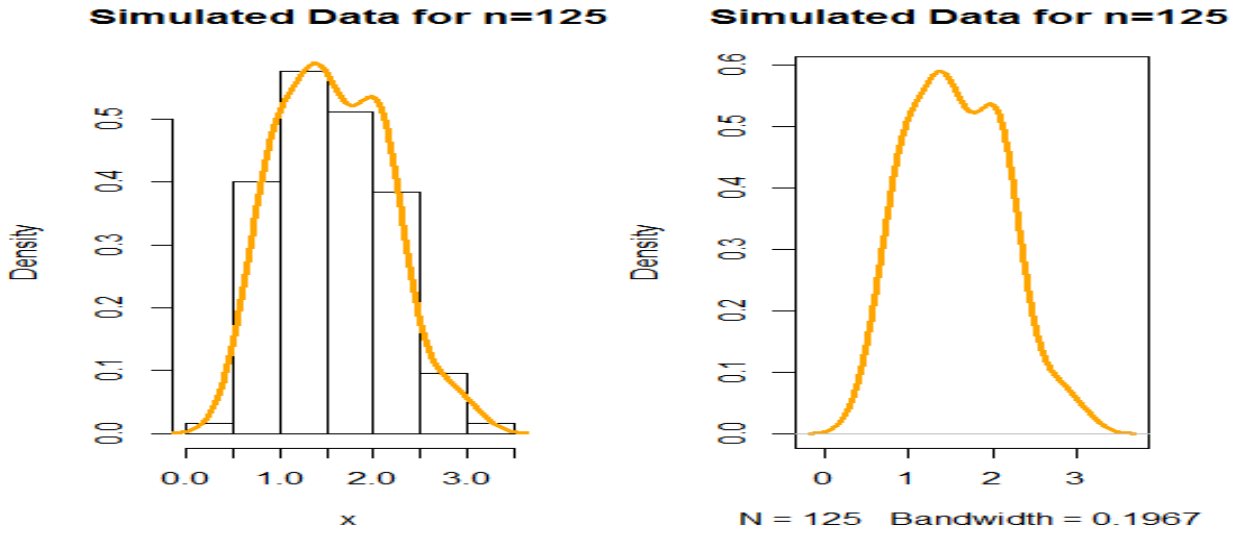



**Figure 4.3: A Histogram and density plot for the simulated data when n=125 (Data set III)**

From the descriptive statistics in tables 4.1, 4.2 and 4.3 as well as the histograms and densities shown above in figures 4.1, 4.2 and 4.3 for the three simulated data sets respectively, we observed that the first data set is normally distributed, the second and third data sets are positively skewed though not highly skewed and therefore suitable for distributions that are symmetry and skewed to the left respectively. Besides the models listed above, these three data sets could also be analyzed using a distribution which has various shapes depending on the

values of the parameters and which the generalized Weibull-Rayleigh distribution is a special case.

For us to assess the models listed above, we made use of some criteria: the *AIC* (Akaike Information Criterion), *CAIC* (Consistent Akaike Information Criterion), *BIC* (Bayesian Information Criterion) and *HQIC* (Hannan Quin information criterion). The formulas for these statistics are given as follows:

$$AIC = -2ll + 2k$$

$$BIC = -2ll + k\log(n),$$

$$CAIC = -2ll + \frac{2kn}{(n-k-1)} \text{ and}$$

$$HQIC = -2ll + 2k\log\left[\log(n)\right]$$

Where ll denotes the log-likelihood function evaluated at the *MLEs*, k is the number of model parameters and n is the sample size.

Decision bench mark: The model with the lowest values of these statistics would be chosen as the best model to fit the data.

**Table 4.4:** Performance of the distribution using the *AIC, CAIC, BIC and HQIC* values of the models based on simulated dataset I when n=25.

| Distributions | Parameter estimates | -ll=(-log-likelihood value) | *AIC* | *CAIC* | *BIC* | *HQIC* | Ranks of models performance |
|---|---|---|---|---|---|---|---|
| *GWRD* | $\hat{\theta}$=0.3158<br>$\hat{\alpha}$=0.8474<br>$\hat{\beta}$=7.8454 | -20.2866 | -34.5732 | -33.4303 | -36.3794 | -39.7003 | **1** |
| *WRD* | $\hat{\theta}$=0.1215<br>$\hat{\alpha}$=10.7175<br>$\hat{\beta}$=1.5130 | 19.1265 | 44.2530 | 45.3959 | 42.4468 | 39.1259 | **4** |
| *TRD* | $\hat{\theta}$=1.1977<br>$\hat{\lambda}$=-1.2865 | 19.2442 | 42.4884 | 43.0339 | 41.2843 | 39.0704 | **3** |
| *RD* | $\hat{\theta}$=0.7303 | 23.2090 | 47.4180 | 48.5919 | 47.8159 | 46.7090 | **5** |
| *WD* | $\hat{\alpha}$=0.1759<br>$\hat{\beta}$=3.1379 | 19.2037 | 42.4074 | 42.9529 | 41.2033 | 38.9894 | **2** |

**Table 4.4** also presents the parameter estimates and the values of *-ll*, *AIC*, *BIC*, *CAIC* and *HQIC* for the five fitted models using simulated data when the sample size is 25 (n=25). The values in the above table also reveal that the *GWRD* has better performance with the lowest values of *AIC*, *CAIC*, *BIC* and *HQIC* compared to the other four models.

**Table 4.5:** Performance of the distribution using the *AIC, CAIC, BIC and HQIC* values of the models based on simulated dataset II when n=75.

| Distributions | Parameter estimates | -ll=(-log-likelihood value) | *AIC* | *CAIC* | *BIC* | *HQIC* | Ranks of models performance |
|---|---|---|---|---|---|---|---|
| *GWRD* | $\hat{\theta}$=0.3095<br>$\hat{\alpha}$=0.9126<br>$\hat{\beta}$=10.5550 | -81.8963 | -157.7926 | -157.4546 | -158.1674 | -162.1545 | **1** |
| *WRD* | $\hat{\theta}$=0.2849<br>$\hat{\alpha}$=1.9242<br>$\hat{\beta}$=1.3905 | 58.2104 | 122.4208 | 122.7588 | 122.0460 | 118.0589 | **3** |
| *TRD* | $\hat{\theta}$=1.1060<br>$\hat{\lambda}$=-1.2643 | 60.4441 | 124.8882 | 125.0549 | 124.6383 | 121.9803 | **4** |
| *RD* | $\hat{\theta}$=0.6805 | 72.1753 | 146.3506 | 146.4054 | 146.2257 | 144.8966 | **5** |
| *WD* | $\hat{\alpha}$=0.1425<br>$\hat{\beta}$=3.2958 | 58.9808 | 181.9616 | 122.1283 | 121.7117 | 119.0537 | **2** |

Again, **Table 4.5** also gives the parameter estimates and the values of all the performance measures for the five fitted distributions when the sample is increased to 75 (n=75). Based on the values in the above table, we also conclude that the *GWRD* fits the simulated data better than the other four distributions.

**Table 4.6:** Performance of the distribution using the *AIC, CAIC, BIC and HQIC* values of the models based on simulated dataset III when n=125.

| Distributions | Parameter estimates | -ll=(-log-likelihood value) | *AIC* | *CAIC* | *BIC* | *HQIC* | Ranks of models performance |
|---|---|---|---|---|---|---|---|
| *GWRD* | $\hat{\theta}$=0.2767<br>$\hat{\alpha}$=0.7021<br>$\hat{\beta}$=6.5166 | -61.4170 | -116.8340 | -116.6357 | -116.5433 | -120.9045 | **1** |
| *WRD* | $\hat{\theta}$=0.2668<br>$\hat{\alpha}$=2.1384<br>$\hat{\beta}$=1.1972 | 107.2665 | 220.5330 | 220.7313 | 220.8237 | 216.4625 | **4** |
| *TRD* | $\hat{\theta}$=1.1067<br>$\hat{\lambda}$=-1.0401 | 107.1828 | 218.3656 | 218.4978 | 218.5594 | 215.6519 | **3** |
| *RD* | $\hat{\theta}$=0.7303 | 119.2528 | 239.5056 | 240.5381 | 240.6025 | 239.1488 | **5** |
| *WD* | $\hat{\alpha}$=0.1977<br>$\hat{\beta}$=2.9330 | 106.0427 | 216.0854 | 216.1838 | 216.2792 | 213.3717 | **2** |

Lastly, we can as well see from the last table, **Table 4.6** that the proposed distribution has smaller values of *-ll, AIC, BIC, CAIC* and *HQIC* compared to the other four distributions. The values in the above table also provide evidence for us to agree that the *GWRD* fits the simulated data better than the other four models. This therefore implies that the *GWRD* could be used to model all kinds of data sets both real and simulated.

# CHAPTER FIVE: SUMMARY, CONCLUSION AND RECOMMENDATIONS

## 5.1 SUMMARY

One important and obvious fact about compound distributions is that they perform better than their classical form. In this dissertation, we have developed a compound distribution known as Generalized Weibull-Rayleigh distribution. Some properties of the proposed distribution such as its validity, ordinary moments, moment generating function, characteristics function, reliability analysis, quantile function and order statistics are being derived. The estimation of parameters has been done using the method of maximum likelihood estimation. The performance of the new distribution is tested by means of some applications to three simulated datasets.

## 5.2 CONCLUSION

In this dissertation, a new distribution has been proposed. Some mathematical and statistical properties of the proposed distribution have been studied appropriately. The derivations of some expressions for its validity, moments, moment generating function, characteristics function, survival function, hazard function, quantile function and ordered statistics has been done appropriately. Some plots of the distribution revealed that it can take any shape depending on values of the parameters. The model parameters have been estimated using the method of maximum likelihood estimation. The implications of the plots for the survival function indicate that the Generalized Weibull-Rayleigh distribution could be used to model age-dependent events or variables whose survival decreases as time grows or where survival rate decreases with time. The results of the three applications to simulated datasets showed that the proposed distribution (Generalized Weibull-Rayleigh distribution) performs better than the Weibull-Rayleigh, transmuted Rayleigh, Weibull and the Rayleigh distributions irrespective of the nature of the data sets and the sample sizes. This implies that the proposed distribution can be used when we have skewed and symmetric datasets. Also, all the compound distributions perform better than their

baseline distribution (Rayleigh distribution) which is in line with the fact that generalizing any continuous distribution provides a compound distribution with better performance.

## 5.3 RECOMMENDATIONS

We recommend based on the findings of this dissertation that the proposed distribution should be used to model both skewed and symmetric data sets with moderate peaks and tails. We also recommend the use of the new model for analyzing age dependent variables.

## 5.4 CONTRIBUTION TO KNOWLEDGE

In this research, we have developed a new distribution which is useful for modeling both skewed and symmetric data sets with moderate peaks and tails. We also provided expressions for some of its properties which are useful for studying the shape characteristics of the distribution as well as the reliability functions which are useful in the field of engineering. The model parameters are also estimated clearly by using the method of Maximum Likelihood Estimation.

Last but not the least, the performance of the new distribution has been assessed in comparison with some existing distributions and we discovered that it is better than other models mostly when dealing with skewed data sets with moderate peaks as well as symmetric data sets.

## 5.5 AREAS OF FURTHER RESEARCH

Subsequent studies in this area of statistics should look at estimation of confidence intervals for the parameters of the proposed distribution. Also, others can estimate the parameters of the new distribution using another classical or Bayesian method.

## REFERENCE


Abdel-Hady, D. H. (2013). Bivariate Generalized Rayleigh Distribution. *Journal of Applied Sciences Research, 9* (9): 5403-5411

Afify, A. Z., Yousof, H. M., Hamedani, G. G, and Aryal, G. (2016). The Exponentiated Weibull-Pareto Distribution with Application. https://www.researchgate.net/publication/296259233

Afify, M. Z., Yousof, H. M., Cordeiro, G. M., Ortega, E. M. M. and Nofal, Z. M. (2016).The Weibull Frechet Distribution and Its Applications. *Journal of Applied Statistics*.1-22.

Ahmad, A., Ahmad, S. P. and Ahmed, A. (2014).Transmuted Inverse Rayleigh distribution: a generalization of the Inverse Rayleigh distribution. *Mathematical Theory and Modeling,* 4(7): 90 98

Alzaatreh, A., Famoye, F. and Lee, C. (2014). The Gamma-Normal Distribution: Properties and Applications, *Computational Statistics and Data Analysis*,69: 67-80.

Bourguignon, M., Silva, R. B. and Cordeiro, G. M. (2014). The Weibull-G Family of Probability Distributions. *Journal of Data Science,* 12: 53-68.

Bourguignon. M, R.B. Silva.L. M.and G. M. Cordeiro, (2012), The Kumaraswamy Pareto distribution, *AMS 2000 Subject Classifications:* 62E10, 62E15, 62E99.

Cordeiro, G. M., Ortega, E. M. M. and Ramires, T. G. (2015), A new generalized Weibull family of distributions: Mathematical properties and applications; *Journal of Statistical Distributions and Applications*, 2:13DOI 10.1186/s40488-015-0036-6

Cordeiro, G.M., Cintra, R.J., RˆEgo, L.C. and Ortega, E.M.M. (2012).The Mcdonald Normal Distribution. *Pakistan Journal of Statistics and Operations Research,* 8: 301-329.

Hirano, K. (1986). Rayleigh Distributions, *New York: John Wiley*.

Howlader, H. A. and Hossain, A. (1995). On Bayesian estimation and prediction from Rayleigh distribution based on type-II censored data. *Communication Statistics: Theory and Methods*, 24(9): 2249-2259

Kareema, A. A. and Mohammed, A. B. (2013). Exponential Pareto distribution. *Mathemetical Theory and modeling,* 5: 2224-5804

Kenney, J. F. & Keeping, E. S. (1962).Mathematics of Statistics, 3 edn, *Chapman & Hall Ltd, New Jersey*.

Kundu, D., and Raqab, M. Z. (2005). Generalized Rayleigh Distribution: Different methods of estimations. *Computational Statistics and Data Analysis, 49*: 187 – 200.

Lima, M. S., Cordeiro, G. M. and Ortega, E. M. M. (2015). A New Extension of The Normal Distribution. *Journal of Data Science,* 13: 385-408.

Merovci, F. (2013). The transmuted rayleigh distribution. *Australian Journal of Statistics*, 22(1): 21–30

Merovci, F. and Puka, I. (2014). Transmuted pareto distribution. *Probability Statistics Forum,* 07:1–11

Merovci, F., and Elbatal, I. (2015).Weibull Rayleigh Distribution: Theory and Applications. *Applied Mathematics and Information Science*, 9(5): 1-11.

Mhmoudi, E. (2011). Beta generalized Pareto distribution (BGPD), *Mathematics and Computer in Simulation*, 81: 2414-2430

Moors, J. J. (1988).A quantile alternative for kurtosis.*Journal of the Royal Statistical Society*: *Series D*, 37: 25–32.

Nadarajah, S., Teimouri, M. and Shih, S. H. (2014). Modified Beta Distributions: *The Indian Journal of Statistics*, B, 76: 19-48.

Oguntunde, P. E., Balogun, O. S, Okagbue, H. I, and Bishop, S. A. (2015). The Weibull-Exponential Distribution: Its Properties and Applications. *Journal of Applied Science.*15(11): 1305-1311.

Pescim, R. R. and Nadarajah, S. (2015). The Kummer Beta Normal: A New Skew Model. *Journal of Data Science,* 13: 509-532.

Rayleigh, J. (1980). On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, *Philosophy and Management*, 10: 73–78

Shaw, W. T. and Buckley, I. R. (2007). The alchemy of probability distributions: beyond Gram-Charlier expansions and a skew-kurtotic-normal distribution from a rank transmutation map. *Research report*

Siddiqui, M. M. (1962). Some problems connected with Rayleigh distributions, *Journal of Research Nat. Bur. Stand*, 60D, 167–174

# APPENDIX

## R COMMANDS USED IN THIS DISSERTATION

## A-PLOTS

```
a=0.1

> b=1.5

> t=2.5

> x=seq(0,4,0.01)

>GWRD=function(x,a,b,t){(a*b*t*exp(-0.5*t*x^2)*((-log(1-(1-exp(-0.5*t*x^2))))^(b-1))*exp(-a*(-log(1-(1-exp(-
0.5*t*x^2))))^b)/(1-(1-exp(-0.5*t*x^2))))}

> plot(x,GWRD(x,0.1,1.5,2.5),type="n",ylim=c(0,2),xlab="X",ylab="PDF of GWRD")

> lines(x,GWRD(x,0.1,1.5,6.5),col="black",lwd=2,lty=1)

> lines(x,GWRD(x,0.08,1.5,6.5),col="red",lwd=2,lty=2)

> lines(x,GWRD(x,1.5,0.5,6.5),col="blue",lwd=2,lty=3)

> lines(x,GWRD(x,0.03,1.5,6.5),col="yellow",lwd=2,lty=4)

> lines(x,GWRD(x,0.03,2.5,6.5),col="green",lwd=2,lty=5)

>legend("topright",cex=1,title="key",c("a=0.1,b=1.5","a=0.08,b=1.5","a=1.5,b=0.5","a=0.03,b=1.5","a=0.03,b=2.5"),horiz=FALSE,
lty=c(1,2,3,4,5),lwd=c(2,2,2,2,2),col=c("black","red","blue","yellow","green"),bg="grey96")

> CDFGWRD=function(x,a,b,t){1-exp(-a*(-log(1-(1-exp(-0.5*t*x^2))))^b)}

> plot(x,CDFGWRD(x,0.1,1.5,6.5),type="n",ylim=c(0,1),xlab="X",ylab="CDF of GWRD")

> lines(x,CDFGWRD(x,0.1,1.5,6.5),col="black",lwd=2,lty=1)

> lines(x,CDFGWRD(x,0.08,1.5,6.5),col="red",lwd=2,lty=2)

> lines(x,CDFGWRD(x,1.5,0.5,6.5),col="blue",lwd=2,lty=3)

> lines(x,CDFGWRD(x,0.03,1.5,6.5),col="yellow",lwd=2,lty=4)

> lines(x,CDFGWRD(x,0.03,2.5,6.5),col="green",lwd=2,lty=5)

>legend("topright",cex=1,title="key",c("a=0.1,b=1.5","a=0.08,b=1.5","a=1.5,b=0.5","a=0.03,b=1.5","a=0.03,b=2.5"),horiz=FALSE,
lty=c(1,2,3,4,5),lwd=c(2,2,2,2,2),col=c("black","red","blue","yellow","green"),bg="grey96")

> SFGWRD=function(x,a,b,t){exp(-a*(-log(1-(1-exp(-0.5*t*x^2))))^b)}

> plot(x,SFGWRD(x,0.1,1.5,6.5),type="n",ylim=c(0,1),xlab="X",ylab="Survival function of GWRD")

> lines(x,SFGWRD(x,0.1,1.5,6.5),col="black",lwd=2,lty=1)

> lines(x,SFGWRD(x,0.08,1.5,6.5),col="red",lwd=2,lty=2)

> lines(x,SFGWRD(x,1.5,0.5,6.5),col="blue",lwd=2,lty=3)

> lines(x,SFGWRD(x,0.03,1.5,6.5),col="yellow",lwd=2,lty=4)

> lines(x,SFGWRD(x,0.03,2.5,6.5),col="green",lwd=2,lty=5)
```

```
>legend("topright",cex=1,title="key",c("a=0.1,b=1.5","a=0.08,b=1.5","a=1.5,b=0.5","a=0.03,b=1.5","a=0.03,b=2.5"),horiz=FALSE,
lty=c(1,2,3,4,5),lwd=c(2,2,2,2,2),col=c("black","red","blue","yellow","green"),bg="grey96")

> HFGWRD=function(x,a,b,t){(a*b*t*exp(-0.5*t*x^2)*((-log(1-(1-exp(-0.5*t*x^2))))^(b-1))/(1-(1-exp(-0.5*t*x^2))))}

> plot(x,HFGWRD(x,0.1,1.5,6.5),type="n",ylim=c(0,2),xlab="X",ylab="Hazard function of GWRD")

> lines(x,HFGWRD(x,0.1,1.5,6.5),col="black",lwd=2,lty=1)

> lines(x,HFGWRD(x,0.08,1.5,6.5),col="red",lwd=2,lty=2)

> lines(x,HFGWRD(x,1.5,0.5,6.5),col="blue",lwd=2,lty=3)

> lines(x,HFGWRD(x,0.03,1.5,6.5),col="yellow",lwd=2,lty=4)

> lines(x,HFGWRD(x,0.03,2.5,6.5),col="green",lwd=2,lty=5)

>legend("bottomright",cex=1,title="key",c("a=0.1,b=1.5","a=0.08,b=1.5","a=1.5,b=0.5","a=0.03,b=1.5","a=0.03,b=2.5"),horiz=F
ALSE,lty=c(1,2,3,4,5),lwd=c(2,2,2,2,2),col=c("black","red","blue","yellow","green"),bg="grey96")
```

## B-DATA ANALYSIS

### SIMULATED DATASETS

```
> theta=0.5

> alpha=1.5

> beta=1.5

> u=runif(25,0,1)

> x=((2/theta)*((1/alpha)*log(1/(1-u)))^(1/beta))^0.5

> loglike1=function(p){-sum(log(p[1]*x*exp(-0.5*p[1]*x^2)))}

> a=optim(c(0.5),fn=loglike1,hessian=T)

Warning message:

In optim(c(0.5), fn = loglike1, hessian = T) :

  one-dimensional optimization by Nelder-Mead is unreliable:

use "Brent" or optimize() directly

> vca=solve(a$hessian)

> sea=sqrt(diag(vca))

> a$convergence

[1] 0

> a

$par

[1] 0.7302734

$value

[1] 23.20897
```

```
$counts
function gradient
     26       NA
$convergence
[1] 0
$message
NULL
$hessian
         [,1]
[1,] 46.87817
> loglike2=function(p){-sum(log(p[1]*x*exp(-0.5*p[1]*x^2)*(1+p[2]-2*p[2]*(1-exp(-0.5*p[1]*x^2)))))}
> b=optim(c(0.008,0.0079),fn=loglike2,hessian=T)
> vcb=solve(b$hessian)
> seb=sqrt(diag(vcb))
> b$convergence
[1] 0
> b
$par
[1]  1.197686 -1.286481

$value
[1] 19.24424
$counts
function gradient
    119       NA
$convergence
[1] 0
$message
NULL
$hessian
         [,1]     [,2]
[1,] 40.67258 20.98783
```

[2,] 20.98783 25.45169

```
> loglike3=function(p){-sum(log(p[2]*p[3]*p[1]*x*exp(-0.5*p[1]*x^2)*(1-exp(-0.5*p[1]*x^2))^(p[3]-1)*(1-(1-exp(-
0.5*p[1]*x^2)))^-(p[3]+1)*exp(-p[2]*((1-exp(-0.5*p[1]*x^2))/(1-(1-exp(-0.5*p[1]*x^2))))^p[3]))}

> c=optim(c(0.007,0.005,0.008),fn=loglike3,hessian=T)

> vcc=solve(c$hessian)

> sec=sqrt(diag(vcc))

> c$convergence

[1] 1

> c

$par

[1]  0.1214589 10.7175257  1.5129527

$value

[1] 19.12654

$counts

function gradient

    502       NA

$convergence

[1] 1

$message

NULL


$hessian

          [,1]      [,2]       [,3]

[1,] 5056.69870 33.0306131 -436.137635

[2,]   33.03061  0.2176461   -2.992727

[3,] -436.13764 -2.9927266   59.086063

> loglike4=function(p){-sum(log(p[2]*p[3]*p[1]*((-log(1-(1-exp(-0.5*p[1]*x^2))))^(p[2]-1))*exp(-p[2]*(-log(1-(1-exp(-
0.5*p[1]*x^2))))^p[3]))}

> d=optim(c(0.01,0.01,0.01),method="CG",fn=loglike4,hessian=T)

There were 22 warnings (use warnings() to see them)

> vcd=solve(d$hessian)

> sed=sqrt(diag(vcd))

> d$convergence
```

```
[1] 1

> d

$par

[1] 0.3157934 0.8474068 7.8453807

$value

[1] -20.28657

$counts

function gradient

    364      101

$convergence

[1] 1

$message

NULL

$hessian

          [,1]       [,2]       [,3]

[1,] 1489.629006 -9.63838129  5.90181772

[2,]   -9.638381 34.81427693 -0.07643936

[3,]    5.901818 -0.07643936  0.45074856


> pdfwd=function(p,x){-sum(log(p[1]*p[2]*x^(p[2]-1)*exp(-p[1]*x^p[2])))}

> resultwd=optim(c(1,0.5),method="CG",fn=pdfwd,hessian=T,x=x)

There were 45 warnings (use warnings() to see them)

> vcwd=solve(resultwd$hessian)

> sewd=sqrt(diag(vcwd))

> resultwd$convergence

[1] 1

> resultwd

$par

[1] 0.1759246 3.1379269

$value

[1] 19.20374

$counts
```

```
function gradient
    390      101
$convergence
[1] 1
$message
NULL
$hessian
         [,1]     [,2]
[1,] 807.82021 96.23876
[2,]  96.23876 15.45950
> theta=0.5
> alpha=1.5
> beta=1.5
> u=runif(75,0,1)
> x=((2/theta)*((1/alpha)*log(1/(1-u)))^(1/beta))^0.5
> loglike1=function(p){-sum(log(p[1]*x*exp(-0.5*p[1]*x^2)))}
> a=optim(c(0.5),fn=loglike1,hessian=T)
Warning message:
In optim(c(0.5), fn = loglike1, hessian = T) :
  one-dimensional optimization by Nelder-Mead is unreliable:
use "Brent" or optimize() directly
> vca=solve(a$hessian)
> sea=sqrt(diag(vca))
> a$convergence
[1] 0
> a
$par
[1] 0.6804688
$value
[1] 72.17526
$counts
function gradient
```

```
     22       NA

$convergence

[1] 0

$message

NULL

$hessian

         [,1]

[1,] 161.9745

> loglike2=function(p){-sum(log(p[1]*x*exp(-0.5*p[1]*x^2)*(1+p[2]-2*p[2]*(1-exp(-0.5*p[1]*x^2)))))}

> b=optim(c(0.008,0.0079),fn=loglike2,hessian=T)

> vcb=solve(b$hessian)

> seb=sqrt(diag(vcb))

> b$convergence

[1] 0

> b

$par

[1]  1.106013 -1.264333


$value

[1] 60.4441

$counts

function gradient

    113       NA

$convergence

[1] 0

$message

NULL

$hessian

         [,1]     [,2]

[1,] 139.49631 65.27558

[2,]  65.27558 73.93520

> loglike3=function(p){-sum(log(p[2]*p[3]*p[1]*x*exp(-0.5*p[1]*x^2)*(1-exp(-0.5*p[1]*x^2))^(p[3]-1)*(1-(1-exp(-
0.5*p[1]*x^2)))^-(p[3]+1)*exp(-p[2]*((1-exp(-0.5*p[1]*x^2))/(1-(1-exp(-0.5*p[1]*x^2))))^p[3]))))}
```

```
> c=optim(c(0.007,0.005,0.008),fn=loglike3,hessian=T)

> vcc=solve(c$hessian)

> sec=sqrt(diag(vcc))

> c$convergence

[1] 0

> c

$par

[1] 0.2849324 1.9242209 1.3904522

$value

[1] 58.21035

$counts

function gradient

   176       NA

$convergence

[1] 0

$message

NULL

$hessian

         [,1]      [,2]      [,3]

[1,] 3399.94825 259.648634 -20.080390

[2,]  259.64863  20.255904  -6.461175

[3,]  -20.08039  -6.461175  64.394388

> loglike4=function(p){-sum(log(p[2]*p[3]*p[1]*((-log(1-(1-exp(-0.5*p[1]*x^2))))^(p[2]-1))*exp(-p[2]*(-log(1-(1-exp(-
0.5*p[1]*x^2))))^p[3]))))}

> d=optim(c(0.01,0.01,0.01),method="CG",fn=loglike4,hessian=T)

Warning messages:

> vcd=solve(d$hessian)

> sed=sqrt(diag(vcd))

> d$convergence

[1] 1

> d

$par

[1]  0.3094852  0.9125937 10.5549888
```

```
$value
[1] -81.89625
$counts
function gradient
    370      101
$convergence
[1] 1
$message
NULL
$hessian
          [,1]     [,2]      [,3]
[1,] 7803.958141  9.256573 18.3913969
[2,]    9.256573 90.054901 -0.1082290
[3,]   18.391397 -0.108229  0.7455748

> pdfwd=function(p,x){-sum(log(p[1]*p[2]*x^(p[2]-1)*exp(-p[1]*x^p[2])))}
> resultwd=optim(c(1,0.5),method="CG",fn=pdfwd,hessian=T,x=x)
There were 50 or more warnings (use warnings() to see the first 50)
> vcwd=solve(resultwd$hessian)
> sewd=sqrt(diag(vcwd))
> resultwd$convergence
[1] 1
> resultwd
$par
[1] 0.1425054 3.2957597
$value
[1] 58.98075
$counts
function gradient
    444      101
$convergence
[1] 1
```

```
$message
NULL
$hessian
        [,1]     [,2]
[1,] 3693.527 372.85095
[2,]  372.851  48.13109
> u=runif(125,0,1)
> x=((2/theta)*((1/alpha)*log(1/(1-u)))^(1/beta))^0.5
> loglike1=function(p){-sum(log(p[1]*x*exp(-0.5*p[1]*x^2)))}
> a=optim(c(0.5),fn=loglike1,hessian=T)
Warning message:
In optim(c(0.5), fn = loglike1, hessian = T) :
  one-dimensional optimization by Nelder-Mead is unreliable:
use "Brent" or optimize() directly
> vca=solve(a$hessian)
> sea=sqrt(diag(vca))
> a$convergence
[1] 0
> a
$par
[1] 0.7302734
$value
[1] 119.2528
$counts
function gradient
     24       NA
$convergence
[1] 0
$message
NULL
$hessian
        [,1]
```

```
[1,] 234.3908

> loglike2=function(p){-sum(log(p[1]*x*exp(-0.5*p[1]*x^2)*(1+p[2]-2*p[2]*(1-exp(-0.5*p[1]*x^2)))))}

> b=optim(c(0.008,0.0079),fn=loglike2,hessian=T)

Warning message:

In log(p[1] * x * exp(-0.5 * p[1] * x^2) * (1 + p[2] - 2 * p[2] *  :

  NaNs produced

> vcb=solve(b$hessian)

> seb=sqrt(diag(vcb))

> b$convergence

[1] 0

> b

$par

[1]  1.106694 -1.040069

$value

[1] 107.1828

$counts

function gradient

    151       NA

$convergence

[1] 0

$message

NULL

$hessian

         [,1]     [,2]

[1,] 189.71135 76.35996

[2,]  76.35996 93.69704

> loglike3=function(p){-sum(log(p[2]*p[3]*p[1]*x*exp(-0.5*p[1]*x^2)*(1-exp(-0.5*p[1]*x^2))^(p[3]-1)*(1-(1-exp(-
0.5*p[1]*x^2)))^-(p[3]+1)*exp(-p[2]*((1-exp(-0.5*p[1]*x^2))/(1-(1-exp(-0.5*p[1]*x^2))))^p[3]))))}

> c=optim(c(0.01,0.01,0.01),method="CG",fn=loglike3,hessian=T)

There were 20 warnings (use warnings() to see them)

> vcc=solve(c$hessian)

> sec=sqrt(diag(vcc))

> c$convergence
```

```
[1] 1

> c

$par

[1] 0.2668402 2.1384490 1.1971988

$value

[1] 107.2665

$counts

function gradient

    384      101

$convergence

[1] 1

$message

NULL

$hessian

         [,1]     [,2]     [,3]

[1,] 4753.08400 349.98290 -63.64054

[2,]  349.98290  27.33458 -15.63510

[3,]  -63.64054 -15.63510 162.04490

> loglike4=function(p){-sum(log(p[2]*p[3]*p[1]*((-log(1-(1-exp(-0.5*p[1]*x^2))))^(p[2]-1))*exp(-p[2]*(-log(1-(1-exp(-
0.5*p[1]*x^2))))^p[3]))))}

> d=optim(c(0.01,0.01,0.01),method="CG",fn=loglike4,hessian=T)

Warning messages:

> vcd=solve(d$hessian)

> sed=sqrt(diag(vcd))

> d$convergence

[1] 1

> d

$par

[1] 0.2767335 0.7020754 6.5166013

$value

[1] -61.417

$counts

function gradient
```

```
    377     101

$convergence

[1] 1

$message

NULL

$hessian

         [,1]     [,2]      [,3]

[1,] 8170.08216  50.16855 145.466716

[2,]   50.16855 253.59712   5.527510

[3,]  145.46672   5.52751   5.194688

> pdfwd=function(p,x){-sum(log(p[1]*p[2]*x^(p[2]-1)*exp(-p[1]*x^p[2])))}

> resultwd=optim(c(1,0.5),method="CG",fn=pdfwd,hessian=T,x=x)

There were 50 or more warnings (use warnings() to see the first 50)

> vcwd=solve(resultwd$hessian)

> sewd=sqrt(diag(vcwd))

> resultwd$convergence

[1] 1

> resultwd

$par

[1] 0.1977308 2.9329662

$value

[1] 106.0427

$counts

function gradient

    437     101

$convergence

[1] 1

$message

NULL

$hessian

         [,1]     [,2]

[1,] 3197.3012 438.43632
```

[2,] 438.4363 83.53867