UNIVERSITI PUTRA MALAYSIA

GENERALIZED BURR TYPE X DISTRIBUTION AND ITS PROPERTIES AND APPLICATIONS

MUNDHER ABDULLAH KHALEEL

FS 2017 38

GENERALIZED BURR TYPE X DISTRIBUTION AND ITS PROPERTIES AND APPLICATIONS

By

MUNDHER ABDULLAH KHALEEL

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Doctor of Philosophy

May 2017

DEDICATIONS

To my dear mom ,Taliah Ghanim Suliman , who has loved and taken care of me for all my life. She continues to support and encourage me throughout my education, as well as everything I pursue. She has always provided peaceful guidance during my difficult situations.
To my dad , Abdullah Khaleel Ibrahim , who has taught me to love others. He has always made me feel special and never failed to tell me how proud he is of me.
To my beloved wife , Nuha , for her exceptional support and encouragement.
To my son , Abdullah , and my daughters , Maria and Ruqaiah.
To my brothers , Farid, Thaeer and Maheer and sisters.

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Doctor of Philosophy

GENERALIZED BURR TYPE X DISTRIBUTION AND ITS PROPERTIES AND APPLICATIONS

By

MUNDHER ABDULLAH KHALEEL

May 2017

Chair: Professor Noor Akma Ibrahim, PhD
Faculty: Science

We develop three new models from the baseline Burr type X with two parameters distribution (BX) using the beta-G, gamma-G and Weibull-G families of distributions. Burr type X distribution is chosen because the probability density function (pdf) and the cumulative distribution function (cdf) are of closed form. As a consequence of this, it can be used very conveniently even for censored data. Unlike Weibull, Generalized Exponential and gamma distributions, BX has a nonmonotone hazard function, which is suitable for the applications for right skewed data. Thus, the new models are attractive in terms of flexibility in dealing with right skewed, left skewed and approximately symmetric data.

These three new distributions are known as beta Burr type X distribution (BBX), gamma Burr type X distribution (GBX) and Weibull Burr type X distribution (WBX). For each of the new distributions, the probability density function and failure rate function are found to be more flexible in accommodating different shapes. The new distributions have numerous sub-models as special cases and the expansions of pdf and cdf for each are obtained. We derived several mathematical properties of BBX, GBX and WBX distributions which include the quantile function, r^{th} moment, moment generating function, mean deviation, Rényi entropy and order statistics. The maximum likelihood method is employed for the estimation of the parameters of BBX, GBX and WBX distributions. Simulation studies are carried out for varying sample sizes and parameter values to assess the performance of BBX, GBX and WBX distributions. Real data sets are used to illustrate the flexibility of BBX, GBX and WBX distributions. From the results, each distribution has provided a better fit than its sub-models and non-nested models based on the criteria such as the Akaike information criterion (AIC), corrected Akaike information criterion ($CAIC$), Bayesian information criterion (BIC), Hannan-Quinn information criterion ($HQIC$) and Kolmogorov-Smirnov statistic ($K-S$). From

the real data analyses, WBX is found to fit left skewed and approximately symmetric data best while GBX fits the right skewed data best.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah

TABURAN BURR TERITLAK JENIS X DENGAN SIFAT DAN PENGGUNAAN

Oleh

MUNDHER ABDULLAH KHALEEL

Mei 2017

Pengerusi: Professor Noor Akma Ibrahim, PhD
Fakulti: Sains

Kami membangunkan tiga model baharu dari garis dasar taburan Burr jenis X dengan dua parameter (BX) menggunakan keluarga taburan beta-G, gamma-G dan Weibull-G. Taburan Burr jenis X dipilih kerana fungsi ketumpatan kebarangkalian (pdf) dan fungsi taburan kumulatif (cdf) adalah berbentuk tertutup. Akibat daripada ini, ia boleh digunakan dengan mudah walaupun untuk data tertapis. Tidak seperti taburan Weibull, Eksponen Teritlak dan gamma, BX mempunyai fungsi bahaya bukan-senada, yang sesuai untuk kegunaan bagi data pencong kekanan. Oleh itu, model baru ini menarik dari segi fleksibiliti dalam menangani data pencong ke kanan, pencong ke kiri dan data yang menghampiri simetri.

Ketiga-tiga taburan baharu ini dikenali sebagai taburan beta Burr jenis X (BBX), taburan gamma Burr jenis X (GBX) dan taburan Weibull Burr jenis X (WBX). Bagi setiap satu taburan ini, fungsi ketumpatan kebarangkalian dan fungsi kadar kegagalan didapati lebih fleksibel dalam menampung bentuk-bentuk berlainan. Taburan-taburan baharu ini mempunyai banyak sub-model sebagai kes-kes khas dan pengembangan pdf dan cdf untuk setiap satu diperolehi. Kami menerbitkan beberapa sifat matematik taburan BBX, GBX dan WBX yang termasuk fungsi kuantil, momen ke r, fungsi penjana momen, sisihan min, entropi Rényi dan statistik tertib. Kaedah kebolehjadian maksimum digunakan untuk anggaran parameter taburan BBX, GBX dan WBX. Kajian simulasi dijalankan untuk pelbagai saiz sampel dan nilai-nilai parameter untuk menilai prestasi taburan BBX, GBX dan WBX. Set data sebenar digunakan untuk menggambarkan fleksibiliti taburan BBX, GBX dan WBX. Daripada keputusan, setiap taburan telah memberi penyesuaian yang lebih baik daripada sub-model dan model bukan-tersarang berdasarkan kriteria seperti kriteria maklumat Akaike (AIC), kriteria maklumat Akaike diperbetulkan ($CAIC$), kriteria maklumat Bayesian (BIC) dan kriteria maklumat Hannan-Quinn ($HQIC$) dan Kolmogorov-Smirnov statistik ($K-S$). Dari-

pada analisis data sebenar, WBX adalan penyuaian terbaik dengan data pencong bagi pencong ke kiri dan data hampir simetri manakala GBX adalah penyuaian terbaik bagi data pencong ke kanan.

ACKNOWLEDGEMENTS

First of all, praise and thank to Allah, The merciful and the beneficent. I wish to express my deep respect and gratitude to Prof. Dr. Noor Akma Ibrahim, who has been a constant source of inspiration through out the preparation of my thesis. She has always been patient with my shortcomings and kept encouraging me to work in a better way. Without her help and support, perhaps, I would not have been able to complete this thesis. I am also grateful to my wonderful supervisory committee members: Dr. Mahendran Shitan, Prof. Dr. Adem Kilicman and Dr. Faton Merovci, for their contributions toward this work, and their willingness to help guide me through out the whole study process.

The writing of this thesis is not possible without the support of some very special people that made a difference in my education. I would like to express special thanks to my mother, Taliah Ghanim, and my father, Abdullah Khaleel - thank you for being such supportive parents and for teaching me about the power of learning and knowledge. I would like to express my deepest thanks to my brothers (Farid, Thaeer and Maheer) and sisters, who have provided me with their helping hands in each and every walk of my life. I would like to thank my lovely wife, Nuha Jabbar, and my sweet children (Abdullah, Maria and Ruqaiah) for always supporting me and encouraging me along with their best wishes.

Finally, I would like to offer my regards and blessings to my friends, Firas Adil, Mostafa Ibrahim, Aws Ali, Mohammed Jasim, Osama Ali, Welnis Jopen, Loh Yue Fang, Umar Mardaki who have somehow helped me during my study days.

Thank you for supporting me and enlightening me, I love you, and may Allah bless all of you.

I certify that a Thesis Examination Committee has met on 23 May 2017 to conduct the final examination of Mundher Abdullah Khaleel on his thesis entitled "Generalized Burr Type X Distribution and its Properties and Applications" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Habshah binti Midi, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Mohd Rizam bin Abu Bakar, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

Ibrahim Mohamed, PhD
Professor
Universiti Malaya
Malaysia
(External Examiner)

Olimjon Shukurovich Sharipov, PhD
Professor
Academy of Sciences of Uzbekistan
Uzbekistan
(External Examiner)

NOR AINI AB. SHUKOR, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 8 August 2017

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee were as follows:

Noor Akma Ibrahim, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Chairman)

Adem Kilicman, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

Mahendran Shitan, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

Faton Merovic, PhD
Associate Professor
Faculty of Natural Science and Mathematics
University of Mitrovica
(Member)

____

ROBIAH BINTI YUNUS, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ______________________ Date: ______________________

Name and Matric No: Mundher Abdullah Khaleel, GS38387

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ____________________________

Name of Chairman of Supervisory Committee

Professor Dr Noor Akma Ibrahim

Prof. Dr. Noor Akma Ibrahim
Pengarah
Institut Penyelidikan Matematik
Universiti Putra Malaysia
43400 UPM Serdang
SELANGOR DARUL EHSAN

Signature: ____________________________

Name of Member of Supervisory Committee

Professor Dr Adem Kilicman

Prof. Dr. Adem Kilicman
Pensyarah
Jabatan Matematik
Universiti Putra Malaysia
43400 UPM, Serdang
Selangor Darul Ehsan

Signature: b/p ____________________________

Name of Member of Supervisory Committee

Associate Professor Dr Mahendran Shitan

Prof. Dr. Noor Akma Ibrahim
Pengarah
Institut Penyelidikan Matematik
Universiti Putra Malaysia
43400 UPM Serdang
SELANGOR DARUL EHSAN

Signature: b/p ____________________________

Name of Member of Supervisory Committee

Associate Professor Dr Faton Merovic

Prof. Dr. Noor Akma Ibrahim
Pengarah
Institut Penyelidikan Matematik
Universiti Putra Malaysia
43400 UPM Serdang
SELANGOR DARUL EHSAN

TABLE OF CONTENTS

	Page
ABSTRACT	i
ABSTRAK	iii
ACKNOWLEDGEMENTS	v
APPROVAL	vi
DECLARATION	viii
LIST OF TABLES	xiii
LIST OF FIGURES	xiv
LIST OF ABBREVIATIONS	xvi

CHAPTER

1	**INTRODUCTION AND BASIC DEFINITIONS**		1
	1.1	Introduction	1
		1.1.1 Beta Distribution	1
		1.1.2 Gamma Distribution	4
		1.1.3 Weibull Distribution	8
	1.2	Basic Definitions	10
		1.2.1 The Density Function	10
		1.2.2 The Cumulative Distribution Function	11
		1.2.3 The Survival Function	11
		1.2.4 The Hazard Rate Function	11
		1.2.5 The Cumulative Hazard Rate Function	12
	1.3	Useful Functions	12
		1.3.1 Incomplete Gamma Functions	12
		1.3.2 The Digamma Function	13
	1.4	Relevant Facts	13
		1.4.1 The Quantile Function	14
		1.4.2 Kurtosis	14
		1.4.3 Skewness	15
		1.4.4 The Non - Central Moments	15
		1.4.5 Moment Generating Function	16
		1.4.6 Mean Deviation	16
		1.4.7 Rényi Entropy	17
		1.4.8 Order Statistics	18
	1.5	Maximum Likelihood Estimation (MLE)	19
	1.6	Goodness-of-Fit Statistics for Model Selection	20
	1.7	Scope of the Study	21
	1.8	Objectives of the Study	21
	1.9	Organization of Thesis	22

2 **LITERATURE REVIEW** 23
2.1 Introduction 23
2.2 Burr Type X Distribution 23
2.3 Beta-Generator (Beta - G) Family of Distributions 25
2.4 Gamma-Generator (Gamma-G) Family of Distributions 30
2.5 Weibull- Generated (Weibull-G) Family of Distributions 36

3 **BETA BURR TYPE X DISTRIBUTION** 44
3.1 Introduction 44
3.2 Beta Burr type X 45
3.3 Expansion of pdf of BBX Distribution 49
3.4 Some Properties of the BBX Distribution 51
3.4.1 Limit Behavior 51
3.4.2 Quantile Function, Skewness and Kurtosis 52
3.4.3 Sub Models 56
3.4.4 Moment and the Moment Generating Function 56
3.4.5 Mean Deviations 59
3.4.6 Rényi Entropy 63
3.4.7 Order Statistics 64
3.5 Parameter Estimation 66
3.6 Simulation Study 69
3.7 Conclusion 71

4 **GAMMA BURR TYPE X DISTRIBUTION** 72
4.1 Introduction 72
4.2 Gamma Burr type X 74
4.3 Properties of the Distribution 75
4.3.1 Sub-Models of GBX 75
4.3.2 Limit Behavior 77
4.3.3 Quantile Function 78
4.3.4 Expansion for pdf and cdf of GBX 80
4.3.5 Moments 82
4.3.6 Moment Generating Function 85
4.3.7 Rényi Entropy 86
4.3.8 Order Statistics 88
4.4 Maximum Likelihood Function 90
4.5 Simulation Study 91
4.6 Conclusion 94

5 **WEIBULL BURR TYPE X DISTRIBUTION** 95
5.1 Introduction 95
5.2 Weibull Burr type X 96
5.3 Expansion of the pdf and cdf of WBX 97
5.4 Mathematical Properties 103
5.4.1 Sub-Models of WBX 103
5.4.2 Quantile Function 103

5.4.3 Moments 105
5.4.4 Moment Generating Function 108
5.4.5 Rényi Entropy 109
5.4.6 Order Statistics 112
5.5 Maximum Likelihood Function 117
5.6 Simulation Study 120
5.7 Conclusion 122

6 **APPLICATIONS OF REAL DATA** 123
6.1 Introduction 123
6.2 Time-to-Failure of the Turbochargers Diesel Engines (T/1000 h) Data Set. 123
6.3 Waiting Times (minutes) of 100 Bank Customers Data Set. 124
6.4 The Strengths of 1.5 cm Glass Fibers Data Set 128
6.5 Conclusion 131

7 **SUMMARY, CONCLUSION AND FUTURE WORK** 132
7.1 Summary 132
7.2 Conclusion 133
7.3 Future Work 133

BIBLIOGRAPHY 135
BIODATA OF STUDENT 140
LIST OF PUBLICATIONS 142

LIST OF TABLES

Table **Page**

3.1 Moments of different BBX distribution for some selected values of parameters ϑ and η. 59

3.2 Average of MLEs (AvE), Bias and root mean square errors (RMSE) for different parameter values. 70

4.1 Moments of different GBX distribution for some selected values of parameters ϑ and η. 84

4.2 Average of MLEs (AvE), Bias and root mean square errors (RMSE) for different parameter values. 93

5.1 Moments of different WBX distribution for some selected values of parameters ϑ and η. 108

5.2 Average of MLEs (AvE), Bias and root mean square errors (RMSE) for different parameter values. 121

6.1 The work time by 1000 hours of turbochargers diesel engines failure data set. 124

6.2 Descriptives statistics of turbochargers failure data set. 124

6.3 The ML estimates, $(K-S)$, P-value, $-l$, AIC , $CAIC$, BIC and $HQIC$ of turbochargers failure data set 125

6.4 The waiting times (minutes) of 100 bank customers data set. 125

6.5 Descriptives statistics of waiting times (minutes) of 100 bank customers data set. 125

6.6 The ML estimates,$(K-S)$, P-value, $-l$, AIC , $CAIC$, BIC and $HQIC$ of waiting times (minutes) of 100 bank customers data set 127

6.7 The breaking strengths of 1.5 cm glass fibers data set. 129

6.8 Descriptives statistics of breaking strengths data set. 129

6.9 The ML estimates,$(K-S)$, P-value, $-l$, AIC , $CAIC$, BIC and $HQIC$ of breaking strengths data set. 129

LIST OF FIGURES

Figure		Page
3.1	Plot of pdf of BBX with different parameter values of ν, ω, η and ϑ.	47
3.2	Plot of cdf of BBX with different parameter values of ν, ω, η and ϑ.	48
3.3	Plot of hazard function of BBX with different parameter values of ω.	48
3.4	Plot of hazard function of BBX with different parameter values of ν.	49
3.5	Plot of kurtosis of BBX with different values of ν.	54
3.6	Plot of kurtosis of BBX with different values of ω.	54
3.7	Plot of skewness of BBX with different values of ν.	55
3.8	Plot of skewness of BBX with different values of ω.	55
4.1	Plot of pdf of GBX with different parameter values.	75
4.2	Plot of hazard function of GBX with different parameter values of ν.	76
4.3	Plot of hazard function of GBX with different parameter values of ϑ.	76
4.4	Plot of hazard function of GBX with different parameter values of ν, ϑ and η	77
5.1	Plot of pdf of WBX with different parameter values of ν, ω, η and ϑ.	98

5.2 Plot of hazard function of WBX with different parameter values of ω and η. 98

5.3 Plot of hazard function of WBX with different parameter values of ν and η. 99

5.4 Plot of hazard function of WBX with different parameter values of ν and ϑ. 99

6.1 Histogram and plots of fitted densities for time-to-failure of the turbochargers diesel engines (T/1000 h) data. 126

6.2 Empirical cdf and cdfs for the time-to-failure of the turbochargers diesel engines (T/1000 h) data. 126

6.3 Histogram and plots of fitted densities for waiting times (minutes) of 100 bank customers data 127

6.4 Empirical cdf and cdfs for the waiting times (minutes) of 100 bank customers data 128

6.5 Histogram and plots of fitted densities for strengths of 1.5 cm glass fibers data 130

6.6 Empirical cdf and cdfs for strengths of 1.5 cm glass fibers data 130

LIST OF ABBREVIATIONS

pdf	probability density function
cdf	cumulative distribution function
BX	Burr type X
BBX	beta Burr type X
GBX	gamma Burr type X
WBX	Weibull Burr type X
MLE	maximum likelihood estimation
Qf	quantile function
μ_r	The r^{th} Non - Central Moments
mgf	Moment Generating Function
i.i.d	independent and identically distribution
AIC	Akaike information criterion
$AICC$	corrected Akaike information criterion
BIC	Bayesian information criterion
$HQIC$	Hannan-Quinn information criterion
$K-S$	Kolmogorov Smirnov test

CHAPTER 1

INTRODUCTION AND BASIC DEFINITIONS

1.1 Introduction

Statistical distributions are very useful for describing and predicting real-world phenomena. Many distributions have been developed to be fit to models. Some statisticians are interested in seeking families of lifetime distributions, such as exponential, Weibull, gamma, beta, and others. Several life distributions may exhibit increasing failure rates, for example the Weibull and Burr type X distributions, increasing linear failure rates, decreasing failure rates, or constant failure rates. Some distributions are bathtub shaped. The distributions hazard function shape has an important role in deciding whether the respective distribution can be fit to a given data set. In recent years, many attempts have been made to generalize lifetime distributions to make them more flexible and also to achieve new distributions with more parameters. Subsequently, the most important generalized distributions used in the present work are defined.

1.1.1 Beta Distribution

A random variable Y is said to have standard beta distribution with parameter υ,ω if the pdf is:

$$\phi(y) = \frac{y^{\upsilon-1}(1-y)^{\omega-1}}{\beta(\upsilon,\omega)} \quad 0<y<1,\ \upsilon>0,\ \omega>0, \tag{1.1}$$

where $\beta(\upsilon,\omega) = \int_0^1 t^{\upsilon-1}(1-t)^{\omega-1}dt$ denote the beta function. Many authors have proposed generalizations of the probability density function. Libby and Novick (1982) considered a generalized beta function whose pdf is defined as:

$$\phi(y) = \frac{c^{\upsilon} y^{(\upsilon-1)}(1-y)^{(\omega-1)}}{\beta(\upsilon,\omega)\left[1-(1-c)y\right]^{(\omega+\upsilon)}} \quad 0<y<1\ ,\ \upsilon,\ \omega,c>0. \tag{1.2}$$

McDonald (1984) proposed two generalized beta distributions, namely generalized beta of the first kind ($GB1$) and the generalized beta of second kind ($GB2$). The pdf of $GB1$ and $GB2$ respectively are:

$$GB1:\phi(y) = \frac{c\, y^{(c\upsilon-1)}}{\beta(\upsilon,\omega)\, b^{c\upsilon}}[1-(\frac{y}{b})^{c}]^{(\omega-1)} \quad 0\le y\le b,\ \upsilon,\omega,c,b>0. \tag{1.3}$$

$$GB2: \phi(y) = \frac{c}{\beta(\upsilon,\omega)\, b^{c\upsilon}} \frac{y^{(c\upsilon-1)}}{[1-(\frac{y}{b})^c]^{(\omega+c)}} \quad y \geq 0\ ,\ \upsilon,\omega,c,b > 0. \tag{1.4}$$

Gordy (1998) proposed a new generalized beta distribution called the confluent hypergeometric distribution. The pdf is given as:

$$\phi(y) = \frac{y^{\upsilon-1}(1-y)^{\omega-1}\, e^{-\gamma y}}{\beta(\upsilon,\omega)\ {}_1F_1(\upsilon;\upsilon+\omega;-\gamma)}\ 0 < y < 1,\ \upsilon,\omega > 0, -\infty < \gamma < \infty \tag{1.5}$$

where

$${}_1F_1(\upsilon;\upsilon+\omega,-\gamma) = \sum_{k=0}^{\infty} \frac{(\upsilon)_k\, (-\gamma)^k}{(\omega)_k\, k!}$$

and

$$(u)_k = u(u+1)(u+2)...(u+k-1).$$

Nadarajah and Kotz (2006c) proposed and studied four new generalizations of the standard beta distribution. The first generalization is called a compound beta distribution. This method is based on the characteristic that if X and Y are independent gamma random variables, then the ratio $\frac{X}{X+Y}$ has the pdf of the standard beta if X and Y have the pdfs for $x > 0$ and $y > 0$, respectively as follows,

$$\phi(x,a,b) = \frac{x^{(a-1)}(1+x)^{-(a+b)}}{\beta(a,b)}, \tag{1.6}$$

$$\phi(y,\nu,\omega) = \frac{y^{(\nu-1)}(1+y)^{-(\nu+\omega)}}{\beta(\nu,\omega)}. \tag{1.7}$$

They considered the distribution of $W = \frac{X}{X+Y}$ and the pdf is given by:

$$\phi(w) = \frac{C(1-w)^{b-1}}{w^{b+1}(1-2w)} \sum_{k=0}^{\infty} \frac{(-1)^k (a+b)_k B(k+b-\beta, a)}{k!} \times \left(\frac{1-2w}{w}\right)^k \Big[(2b+k)w - b - k\Big] \; a, b, \alpha, \beta > 0,$$

where C is a constant and is given by:

$$C = \frac{{}_3F_2(\beta, 1-\alpha, b-\beta; \beta-1, b+a-\beta; 1)}{\beta\, B(a,b)\, B(\alpha,\beta)}.$$

The second beta distribution generalization is called power beta. It is similar to the first method except the pdfs of X and Y for $x > 0$ and $y > 0$ respectively are as follows,

$$\phi(x, a, \beta) = \frac{x^{(a-1)} e^{(\frac{-x}{\beta})}}{\beta^a\, \Gamma(a)}, \tag{1.8}$$

$$\phi(y, b, \beta) = \frac{y^{(b-1)} e^{(\frac{-y}{\beta})}}{\beta^b\, \Gamma(b)} \; \beta > 0, \tag{1.9}$$

They considered the distribution of $W = \frac{X^c}{X^c + Y^c}$ when $c > 0$. The pdf of w is given by:

$$\phi(w) = \frac{w^{-(1+\frac{b}{c})} \left(1-w\right)^{(\frac{b}{c}-1)}}{b\,c\,B(a,b)} \Big\{ b\; {}_2F_1\big(b, a+b; b+1; -\big(\frac{1-w}{w}\big)^{(\frac{1}{c})}\big) - \Big(\frac{1-w}{w}\Big)^{(\frac{1}{c})} {}_2F_1\big(b+1, a+b+1; b+2; -\big(\frac{1-w}{w}\big)^{(\frac{1}{c})}\big) \Big\} \; a, b > 0.$$

The third generalization is the hypergeometric beta distribution and its pdf is:

$$\phi(y) = \frac{\omega\, \beta(\nu, \omega)\, y^{(\nu+\omega-1)}\, {}_2F_1(1-\gamma, \nu; \upsilon+\omega; y)}{\beta(\nu, \omega+\gamma)} \; 0 < y < 1 \; \nu, \gamma, \omega > 0, \tag{1.10}$$

where

$$ {}_2F_1(a,b;c;y) = \sum_{k=0}^{\infty} \frac{(a)_k\,(b)_k\,y^k}{(c)_k\,k!} \;\; a,b,c>0 $$

denotes the ascending factorial. The last generalization is the trigonometric beta (TB) distribution as it involves trigonometric functions. The authors proposed four pdfs for the TB distribution, two of which involve the cosine function and the other two involve the sine function.

A new technique to generalized family of beta class distribution was introduced by Eugene et al. (2002) who used the logit of beta distribution. The beta class cdf is given by:

$$ \Phi(y,\nu,\omega) = \int_0^{F(y)} \frac{1}{\beta(\nu,\omega)} t^{\nu-1}(1-t)^{\omega-1}\,dt,\; \nu,\omega>0, \tag{1.11} $$

where $F(y)$ is the cdf of any continuous random variable and the pdf corresponding to the cdf is:

$$ \phi(y,\nu,\omega) = \frac{f(y)}{\beta(\nu,\omega)} \Big[F(y)\Big]^{\nu-1} \Big[1-F(y)\Big]^{\omega-1}. \tag{1.12} $$

Alexander et al. (2012) provided a generalized beta distribution called generalized beta generated. The pdf of the new generalized beta is:

$$ \phi_{gB}(y,\upsilon,\omega,c) = \frac{c}{\beta(\upsilon,\omega)}(t)^{\upsilon c-1}\left(1-t^c\right)^{\omega-1} \;\; 0<t<1,\; \upsilon,\omega,c>0. \tag{1.13} $$

1.1.2 Gamma Distribution

A random variable Y has a gamma distribution if its pdf is given as:

$$ \phi(y,\nu,\omega) = \frac{1}{\Gamma(\nu)\omega^{\nu}}\, y^{(\nu-1)}\, e^{-(\frac{y}{\omega})} \quad y>0,\; \nu,\omega>0. \tag{1.14} $$

Amoroso (1925) proposed a generalization of the gamma distribution and obtained the following density function:

$$\phi(y,\nu,\omega,p) = \frac{p}{\Gamma(\nu)\omega^{\nu}}\, y^{(\nu p-1)}\, e^{-(\frac{y}{\omega})^{p}} \quad y>0,\ \nu,\omega,p>0. \tag{1.15}$$

Stacy (1962) proposed another generalization of a gamma distribution, which is a special case of Amoroso′s generalization and the pdf is given by:

$$\phi(y,\nu,\omega,p,d) = \frac{p}{\Gamma(\frac{d}{p})\omega^{\nu}}\, y^{(\nu-1)}\, e^{-(\frac{y}{\omega})^{p}} \quad y>0,\ \nu,\omega,p,d>0. \tag{1.16}$$

Stacy′s generalization has many special cases, such as Chi-square, gamma, Weibull, exponential, and Chi distributions. Gupta et al. (1998) proposed and studied a new gamma distribution generalization called the exponentiated gamma distribution. The general form of their exponentiated gamma distribution is defined as:

$$G_{\nu}(y) = \left[\Phi(y)\right]^{\nu}, \tag{1.17}$$

where ν is the exponentiated parameter. The pdf of the exponentiated distribution is defined as:

$$g_{\nu}(y) = \nu\,[\Phi(y)]^{\nu-1}\,\phi(y). \tag{1.18}$$

The cdf of the exponentiated gamma distribution can be written as:

$$\Phi(y,c) = \left[1-e^{y}(y+1)\right]^{c},\ \ y>0, c>0. \tag{1.19}$$

Agarwal and Al-Saleh (2001) provided another gamma distribution generalization whose pdf is defined as:

$$\phi(y,m,n,c,\delta) = \frac{c^{m-\delta}y^{(m-1)}e^{-(cy)}}{\Gamma_{\delta}(m,n)(\frac{y+n}{c})^{\delta}} \quad y\geq 0, m,n,c>0, \delta=\eta-1, \eta\geq 0, \tag{1.20}$$

where

$$\Gamma_{\delta}(m,n) = \int_0^{\infty} \frac{y^{(m-1)} e^{-(y)}}{(y+n)^{\delta}}\, dy.$$

They showed that failure rate can be monotonic or bathtub shaped. Nadarajah and Kotz (2006b) proposed a generalization of the gamma distribution by using Gupta et al. (1998) method. This is called the generalized gamma distribution and its cdf is given as:

$$\Phi(y) = \left[\Psi_1(m,y)\right]^{\nu} \quad y > 0, \nu > 0, m > 0, \tag{1.21}$$

where $\Psi_1(.,.)$ is the incomplete gamma function and defined as:

$$\Psi_1(m,y) = \frac{\psi(m,y)}{\Gamma(m)},$$

with

$$\psi(m,y) = \int_0^{y} t^{m-1}\, e^{-t}\, dt,$$

is the incomplete gamma function and

$$\Gamma(m) = \int_0^{\infty} t^{m-1}\, e^{-t}\, dt,$$

is the complete gamma function.

Zografos and Balakrishnan (2009) introduced and studied a new generalized gamma distribution for univariate continuous distribution (generalized gamma type 1) using a particular case of Stacy′s generalized gamma distribution. The cdf of the family of distributions generated by gamma random variables can be defined by:

$$\Phi(y,m) = \frac{1}{\Gamma(m)} \int_0^{-\log\left[1-\Phi(y)\right]} t^{m-1} e^{-t} \, dt \qquad (1.22)$$
$$= \frac{1}{\Gamma(m)} I_{-\log\left[1-\Phi(y)\right]}(m), m > 0,$$

The pdf corresponding to the cdf is given by:

$$\phi(y,m) = \frac{\phi(y)}{\Gamma(m)} \left\{ -\log\left[1-\Phi(y)\right] \right\}^{m-1} \quad y > 0, m > 0. \qquad (1.23)$$

Cordeiro and de Castro (2011) also proposed and studied a generalized gamma distribution called exponentiated generalized gamma distribution. The pdf of this new distribution is given as:

$$\phi(y,\nu,\omega,c,k) = \frac{\nu c (\frac{y}{\omega})^{\nu k-1} e^{-(\frac{y}{\omega})^c}}{\omega\Gamma(k)} \left[\Psi_1\left(k, (\frac{y}{\omega})^c\right)\right]^{\nu-1}, \; y > 0, \nu, c, \omega, k > 0. \qquad (1.24)$$

Ristić and Balakrishnan (2012) introduced an alternative to the family defined by Zagrofose and Balakrishnan. The pdf of the new generalized gamma distribution (generalized gamma type 2) is given by:

$$\phi(y,\delta) = \frac{\phi(y)\left\{ -\log\left[\Phi(y)\right] \right\}^{\delta-1}}{\Gamma(\delta)}, \; y > 0, \delta > 0. \qquad (1.25)$$

The cdf corresponding to the pdf is given by:

$$\Phi(y,\delta) = 1 - \frac{\gamma\left\{ \delta, -\log\left[\Phi(y)\right] \right\}}{\Gamma(\delta)}, \; y > 0, \delta > 0, \qquad (1.26)$$

where $\gamma\{.,.\}$ is incomplete gamma function.

1.1.3 Weibull Distribution

Swedish Physicist Waloddi Weibull (1939) introduced the Weibull distribution (Johnson and Kotz, 1994). A random variable Y has the Weibull distribution with three parameters a, b, c if the pdf is:

$$\phi(y,a,b,c) = \frac{c}{b}\left(\frac{y-a}{b}\right)^{c-1} e^{-(\frac{y-a}{b})^c} \; y \geq a, \; a,b,c > 0, \tag{1.27}$$

where a is location parameter, or shift parameter, b is scale parameter , and c is the shape parameter. The two-parameter Weibull distribution is obtained when $a = 0$ and we shall refer to it as the standard Weibull. Its pdf is:

$$\phi(y,b,c) = \frac{c}{b}\left(\frac{y}{b}\right)^{c-1} e^{-(\frac{y}{b})^c} \; y \geq 0, \; b,c > 0. \tag{1.28}$$

A different form in which the two-parameter Weibull distribution can be written is:

$$\phi(y,\eta,b) = b\eta^b y^{b-1} e^{-(\eta y)^b} \; y \geq 0, \; \eta, b > 0. \tag{1.29}$$

Mudholkar and Srivastava (1993) proposed a generalized form of the standard Weibull distribution by introducing an additional shape parameter (exponentiated Weibull). The cdf is given by:

$$\Phi(y,b,\eta,k) = \left[1 - e^{-(\eta y)^b}\right]^k \; y \geq 0, \; \eta, b, k > 0. \tag{1.30}$$

The pdf of exponentiated Weibull is:

$$\phi(y,\eta,b,k) = kb\eta^b \, y^{b-1} \, e^{-(\eta y)^b} \left[1 - e^{-(\eta y)^b}\right]^{k-1} \; y \geq 0, k, \eta, b > 0. \tag{1.31}$$

A new technique was proposed by Marshall and Olkin (1997) to generalize the standard Weibull distribution by adding one parameter $a > 0$. The new technique is the extended Weibull distribution and its cdf is given by:

$$\Phi(y,a,b,\eta) = 1 - \frac{a\, e^{-(\eta y)^b}}{1-(1-a)e^{-(\eta y)^b}}\ \eta, a, b > 0. \tag{1.32}$$

A new generalized Weibull distribution, i.e. the modified Weibull distribution was introduced by Lai et al. (2003) . The cdf of the modified Weibull distribution is:

$$\Phi(y,\eta,\nu,b) = 1 - e^{-(\eta\, y^b\, e^{\nu\, y})}\ \ y \geq 0,\ \eta, \nu, b > 0. \tag{1.33}$$

Mudholkar et al. (1996) proposed another generalization called the generalized Weibull family. The cdf of this new model is given by

$$\Phi(y,\eta,\nu,b) = 1 - [1-\nu\ (\eta\ y)^b]^{\frac{1}{\nu}}\ \ \ y \geq 0,\ \eta, b > 0, -\infty < \nu < \infty. \tag{1.34}$$

Xie et al. (2002) proposed and study a new generalization of the standard Weibull distribution and named it Weibull extension. They extended the modified Weibull proposed by Lai et al. (2003). The cdf of the new model is given by:

$$\Phi(y,\nu,\omega,\eta) = 1 - e^{\eta\,\nu\,(1-e^{(\frac{y}{\nu})^{\omega}})}\ \ y \geq 0,\ \nu, \omega, \eta > 0. \tag{1.35}$$

A further generalized Weibull family was introduced and assessed by Haghighi and Nikulin (2006). This is the generalized power Weibull family and its cdf and pdf are respectively:

$$\Phi(y,\eta,\omega,\vartheta) = 1 - e^{\left\{1-\left[1+(\eta\, y)^{\omega}\right]^{\vartheta-1}\right\}}\ \ y \geq 0, \eta, \omega, \vartheta > 0, \tag{1.36}$$

and

$$\phi(y,\eta,\omega,\vartheta) = \vartheta\,\omega\,\eta^{\omega}\left[1+(\eta\, x)^{\omega}\right]^{\vartheta} e^{\left\{1-\left[1+(\eta\, y)^{\omega}\right]^{\vartheta-1}\right\}}. \tag{1.37}$$

Cooray (2006) also proposed a generalized Weibull family, which is the odd Weibull family. The pdf is given by:

$$\phi(y,\eta,\nu,\vartheta) = \frac{\nu\,\vartheta\,\eta\,y\,e^{(\eta\,y)^{\nu}}\left[e^{(\eta\,y)^{\nu}}-1\right]^{\vartheta-1}}{y}\left\{1+\left[e^{(\eta\,y)^{\nu}}-1\right]^{\vartheta}\right\}^{-2} \; y \geq 0, \eta, \nu, \vartheta > 0. \tag{1.38}$$

Alzaatreh et al. (2013) introduced and studied a Weibull generalization named the Weibull-X family of distributions by using a new technique. Its cdf is:

$$\Phi(y) = \int_0^{W(G(y))} r(y)\, dy\,, \tag{1.39}$$

where $r(y)$ is the standard Weibull distribution, $W(G(y)) = -\log[1-G(y)]$ and $G(y)$ is any random distribution.

More recently, Bourguignon et al. (2014) proposed a Weibull generalization called the Weibull - G family of distributions by using the same technique as Alzaatreh et al. (2013) as in (1.39). If $W(G(y)) = \frac{G(y)}{1-G(y)}$, the cdf of the Weibull-G family is:

$$\Phi(y,\Theta) = \int_0^{\frac{G(y,\Theta)}{1-G(y,\Theta)}} r(y)\, dy\,, \tag{1.40}$$

when $r(y)$ is the standard Weibull distribution.

1.2 Basic Definitions

In this section, important functions are defined, such as the density, cumulative distribution, survival, hazard rate, cumulative hazard rate, and reversed hazard rate functions. Each of these functions describes the distribution of a random variable.

1.2.1 The Density Function

The density function or probability density function (pdf) of continuous random variables satisfies $\phi(y) \geq 0\; \forall y$,

$$\int_{-\infty}^{+\infty} \phi(y)dy = 1.$$

1.2.2 The Cumulative Distribution Function

The cumulative distribution function (cdf) is also called the failure distribution. It is denoted by $\Phi(y)$ and can be derived using the density function as follows:

$$\Phi(y) = P_r(Y \leq y) = \int_{-\infty}^{y} \phi(x)dx.$$

1.2.3 The Survival Function

The survival function or reliability function denoted by $R(y)$ is complementary to the cumulative distribution function. It can be found using the density function or cumulative distribution function as follows:

$$R(y) = 1 - \Phi(y) = P_r(Y > y) = \int_{y}^{\infty} \phi(x)dx.$$

1.2.4 The Hazard Rate Function

The hazard rate function or rate function of Y is denoted by $h(y)$ and is given by:

$$\begin{aligned} h(y) &= \lim_{\Delta y \to 0} \frac{P_r(y < Y < y + \Delta y | Y > y)}{\Delta y} \\ &= \lim_{\Delta y \to 0} \frac{\Phi(y + \Delta y) - \Phi(y)}{\Delta y R(y)} = \frac{\phi(y)}{R(y)} = \frac{\phi(y)}{1 - \Phi(y)} \end{aligned}$$

Therefore, when Δy is very small and satisfies $h(y) \geq 0 \, \forall y$,

$$\int_{0}^{\infty} h(y)\, dy = \infty$$

1.2.5 The Cumulative Hazard Rate Function

The cumulative or integrated hazard rate function denoted by $H(y)$ is a significant function. Recently, many distributions have been generated using this function and it can be derived using the hazard rate function, survival function, or cumulative distribution function. It is given by:

$$H(y) = \int_0^y h(x)\,dx$$
$$H(y) = -\log(R(y)) = -\log(1-\Phi(y))$$

1.3 Useful Functions

This section presents some brief notes on gamma and special gamma functions, whose derivations will be used throughout this thesis. The primary reason the gamma function is useful in such contexts is the prevalence of type $\phi(u)e^{-\phi(u)}$ expressions, which describe processes that decay exponentially in time or space. Integrals of such expressions can occasionally be solved in terms of the gamma function when no elementary solution exists. The gamma function, denoted by $\Gamma(.)$, is defined as:

$$\Gamma(\nu) = \int_0^\infty u^{\nu-1}e^{-u}du\,;\nu>0$$

That is, if ν is a positive integer one has

$$\Gamma(\nu) = (\nu-1)!$$

1.3.1 Incomplete Gamma Functions

In mathematics, the upper incomplete gamma function $\Gamma(\nu,y)$, and lower incomplete gamma function $\gamma(\nu,y)$, are types of special functions that arise as solutions to various mathematical problems such as certain integrals. The upper incomplete gamma function is defined as:

$$\Gamma(\nu,y) = \int_y^\infty u^{\nu-1}e^{-u}du. \tag{1.41}$$

The lower incomplete gamma function is defined as:

$$\gamma(\nu, y) = \int_0^y u^{\nu-1} e^{-u} du. \tag{1.42}$$

Furthermore, both higher and lower incomplete gamma ratios denoted by $\Gamma_\nu(y)$ and $\gamma_\nu(y)$, respectively, can be defined as:

$$\Gamma_\nu(y) = \frac{\Gamma(\nu, y)}{\Gamma(\nu)},$$

and

$$\gamma_\nu(y) = \frac{\gamma(\nu, y)}{\Gamma(\nu)},$$

where

$$\Gamma(\nu) = \Gamma(\nu, y) + \gamma(\nu, y).$$

1.3.2 The Digamma Function

The logarithmic derivative of the gamma function is known as the digamma (or psi) function, $\psi(.)$, which is defined by:

$$\psi(\nu) = \frac{d}{d\nu} \ln \Gamma(\nu) = \frac{\Gamma'(\nu)}{\Gamma(\nu)}.$$

1.4 Relevant Facts

This section presents a number of basic relevant facts pertaining to this thesis. Most of these facts were presented by Hogg and Craig (1995), among others.

1.4.1 The Quantile Function

The quantile function is one way to determine a probability function. The quantile function Qf of the distribution is defined as the inverse of the cumulative distribution function. The quantile function takes the form below:

$$y = Qf(p) = \{y : P_r(Y \leq y)\} = p = \Phi^{-1}(y).$$

The quantile function is essential for the generation of observations in simulation studies. It also helps find the Moors kurtosis and Bowley skewness of any distribution when the distribution does not have moments or has a large skewness value.

1.4.2 Kurtosis

Kurtosis is an essential measure of a distribution′s shape. It is a measure of the relative peakiness of its frequency curve or flatness relative to a normal distribution. Thus, the central moment of the kurtosis (μ_n) and the cumulants (τ_n) of X are

$$\mu_n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k \mu_1^k \, \mu_{n-k},$$

and

$$\tau_n = \mu_n - \sum_{k=1}^{n-1} \binom{n-1}{k-1} (-1)^k \tau_k \, \mu_{n-k},$$

respectively, where $\tau_1 = \mu_1$, $\tau_2 = \mu_2 - \mu_1^2$,$\tau_3 = \mu_3 - 3\mu_2\,\mu_1 + 2\mu_1^3$, and $\tau_4 = \mu_4 - 4\mu_3\,\mu_1 - 3\,\mu_2^2 + 12\mu_2\,\mu_1^2 - 6\mu_1^4$. Pearson's measure of kurtosis (K_u) is determined as:

$$K_u = \frac{\mu_4}{\mu_2^2}, \tag{1.43}$$

where μ_2 , μ_4 are the second and fourth central moments from the mean. Based on these values, distributions with $K_u > 3$ are called leptokurtic, those with $K_u < 3$ are called platykurtic, while distributions with $K_u = 3$ are called mesokurtic. The last case is similar to the bell shape of a normal distribution. Another measure of kurtosis that is not dependent on the central moment is based on the quantile function of Moors kurtosis (Moors, 1988).

This measure is based on the octile of the distribution and is less sensitive to outlier data. It is defined as:

$$M_u = \frac{Qf(1/8)+Qf(3/8)+Qf(7/8)-Qf(5/8)}{Qf(3/4)-Qf(1/4)}. \tag{1.44}$$

1.4.3 Skewness

The skewness of a distribution is defined as the lack of symmetry. In a symmetrical distribution, the mean, median and mode are equal to each other, and the ordinate at the mean divides the distribution into two equal parts such that one part is a mirror image of the other. Karl Pearsons measure of skewness is based upon the means divergence from the mode in a skewed distribution. Pearsons measure of skewness is given by:

$$S_k = \frac{\mu_3}{\mu_2^{(\frac{3}{2})}}, \tag{1.45}$$

where μ_2 and μ_3 are the second and third central moments. When the distribution is symmetric about the mean, then $\mu_3 = 0 \Longrightarrow S_k = 0$. Similar to a normal distribution, it has a skewness of zero, meaning the shape is symmetric. It can also be noted that the measure of skewness S_k may take on negative or positive values depending on whether μ_3 is negative or positive, respectively. Hence, distributions with $S_k > 0$ are said to be positively skewed distributions, while those with $S_k < 0$ are said to be negatively skewed. Another measure of skewness is the Bowley skewness (Kenney and Keeping, 1954), which is based on quartiles and is one of the earliest skewness measures. It is defined by:

$$S_k = \frac{Qf(3/4)+Qf(1/4)-2Qf(1/2)}{Qf(3/4)-Qf(1/4)}. \tag{1.46}$$

1.4.4 The Non - Central Moments

The symbol μ_r refer to the r^{th} non - central moment (or the moment about the origin) of a continuous random variable Y having a distribution $\phi(y)$, for $r \geq 1$. The r^{th} non -

central moment is given by:

$$\mu_r = E(Y^r) = \int_{-\infty}^{\infty} y^r \, \phi(y) \, dy,$$

the first moment about zero is called the mean, and it is a measure of central tendency denoted by μ.

1.4.5 Moment Generating Function

The term $M_Y(t)$ refers to moment generating function (mgf) of a continuous random variable y with pdf $\phi(y)$. It is given by:

$$M_Y(t) = E(e^{tY}) = \int_{-\infty}^{\infty} e^{ty} \, \phi(y) \, dy = \sum_{r=0}^{\infty} \frac{t^r}{r!} \int_{-\infty}^{\infty} y^r \, \phi(y) \, dy = \sum_{r=0}^{\infty} \frac{\mu_r t^r}{r!}.$$

1.4.6 Mean Deviation

The amount of scatter in a population is measured to some extent by the totality of deviations from the mean and median. The mean deviation from the mean and median can be used to measure a population variation. The mean deviation from the mean is a robust statistic as it is more resilient to outliers in a data set than standard deviation. In standard deviation, the distance from the mean is squared. Hence, on average, large deviations are weighted more heavily, and thus outliers can heavily influence it. In the mean deviation from the mean, the magnitude of a small number of outliers distances is irrelevant. The mean deviation from the median is a measure of statistical dispersion. It is a more robust estimator of scale than sample variance or standard deviation. It thus behaves better with a distribution without a mean or variance such as the Cauchy distribution.

The mean deviation from the mean and mean deviation from the median are respectively defined by:

$$D(\mu) = E(Y - \mu)$$

$$D(M) = E(Y - M)$$

where $\mu = E(Y)$ and $M = \text{median}(Y)$. The measures $D(\mu)$ and $D(M)$ can be found using the following relationships:

$$D(\mu) = \int_{-\infty}^{\infty} |y - \mu|\phi(y)\,dy = 2\mu[\Phi(\mu)] - 2\mu + 2\int_{\mu}^{\infty} y\phi(y)\,dy, \tag{1.47}$$

or

$$D(\mu) = \int_{-\infty}^{\infty} |y - \mu|\phi(y)\,dy = 2\mu[\Phi(\mu)] - 2\int_{-\infty}^{\mu} y\phi(y)\,dy. \tag{1.48}$$

The $D(M)$ is found by:

$$D(M) = \int_{-\infty}^{\infty} |y - M|\phi(y)\,dy = -\mu + 2\int_{M}^{\infty} y\phi(y)\,dy. \tag{1.49}$$

These will be explained later in more detail.

1.4.7 Rényi Entropy

Rényi entropy is a measure of the variation or uncertainty of random variables. It is a very popular entropy measure in many fields of science, such as engineering, theory of communication and probability. The Rényi entropy for a random variable with any pdf of distribution can be determined from the definition:

$$I_R(\rho) = \frac{1}{1-\rho}\log\int_{-\infty}^{\infty}\phi^{\rho}(y)dy \quad \rho > 0,\ \ \rho \neq 1. \tag{1.50}$$

Statistical entropy is a probabilistic measure of uncertainty (or ignorance) about the outcome of a random experiment. It is also a measure of the reduction in that uncertainty. Abundant entropy and information is induced, among which the Rényi entropy has been developed and used in various disciplines and contexts (Rrnyi, 1961). Information-theoretic principles and methods have become integral parts of probability and statistics and are applied in various branches of statistics and related fields. Rényi entropy tends toward Shannon entropy, as $\rho \Longrightarrow 1$.

1.4.8 Order Statistics

Let $Y_1, Y_2, ..., Y_n$ denote n independent and identically distribution (iid) random samples from a distribution function $\Phi(y)$. Let variables Y_i be the arrangement of the sample value from the smallest to the largest and denoted as $Y_{1:n} \leq Y_{2:n} \leq ... \leq Y_{n:n}$, called corresponding order statistics. Here the $Y_{1:n}$ is the first order statistics denoting the smallest of Y_is

$$Y_{1:n} = minY_{i:n} \quad i = 1, 2, ..., n \ ,$$

$$Y_{2:n} = The\ second\ smallest\ of\ Y_{i:n} \quad i = 1, 2, ..., n \ ,$$

and $Y_{n:n}$ is the n^{th} order statistics denoting the largest of Y_is

$$Y_{n:n} = maxY_{i:n} \quad i = 1, 2, ..., n \ .$$

The r^{th} value of this order statistics is called the r^{th} order statistics of the sample. The cdf of the $Y_{r:n}$ is given by:

$$\begin{aligned} \Phi_{r:n}(y) =& P_r(Y_{r:n} \leq y) \\ =& \sum_{i=r}^{n} C_i^n \Big[\Phi(y)\Big]^i \Big[1 - \Phi(y)\Big]^{n-i} \ . \end{aligned} \tag{1.51}$$

From (1.51) we can find the cdfs of the $Y_{1:n}$ and $Y_{n:n}$ as follows,

$$\Phi_{1:n}(y) = 1 - \Big[1 - \Phi(y)\Big]^n \quad -\infty < y < \infty,$$

$$\Phi_{n:n}(y) = \Big[1 - \Phi(y)\Big]^n \quad -\infty < y < \infty.$$

The pdf for $Y_{r:n}$ is given by:

$$\phi_{r:n}(y) = \frac{n!}{(r-1)!(n-r)!}\Big[\Phi(y)\Big]^{r-1}\Big[1-\Phi(y)\Big]^{n-r}\phi(y)\ . \tag{1.52}$$

Substituting $r = 1$ in (1.52) gives the pdf of the $X_{1:n}$ as the sample minimum:

$$\phi_{1:n}(y) = n\,\phi(y)\Big[1-\Phi(y)\Big]^{n-1}\ .$$

Substituting $r = n$ in (1.52) gives the pdf of the $Y_{n:n}$ as the sample maximum:

$$\phi_{n:n}(y) = n\,\phi(y)\Big[\Phi(y)\Big]^{n-1}\ .$$

1.5 Maximum Likelihood Estimation (MLE)

There are several different methods of estimating model parameters. Maximum likelihood estimation (MLE) is one of the most widely used methods for estimating statistical model parameters. With a given statistical model, when MLE is applied on a data set, the method is able to estimate the parameters of the model. In other words, for a fixed data set with a given statistical model, the maximum likelihood approach selects a set of model parameter values that maximize the likelihood function. This method directly maximizes the agreement of the chosen model with the observed data. For discrete random variables MLE maximizes the probability of the observed data under the conditions of distribution (Sijbers et al., 1998). In applying the approach of maximum likelihood estimation, the first step is to determine the joint density function for all observations. The mathematical definition of MLE is presented below:

Let $Y_1, Y_2, ..., Y_n$ be an iid sample with pdf $\phi(y_i;\vartheta)$, where ϑ is a $(k\times 1)$ vector of parameters that characterize $\phi(y;\vartheta)$. The joint density of the sample is, by independence, equal to the product of the marginal densities as:

$$\phi(y_1, y_2, ..., y_n;\vartheta) = \prod_{i=1}^{n}\phi(y_i;\vartheta)\,.$$

Then the likelihood function denoted by $l(\phi)$ is defined as the joint density treated as a function of parameter ϑ:

$$l(\phi) = \phi(\vartheta |\, y_1, y_2, ..., y_n) = \prod_{i=1}^{n} \phi(y_i; \vartheta).$$

The MLE of ϑ is that value that maximizes $l(\phi)$. In other words, it is the value that makes the observed data the most probable. Rather than maximizing this product, which can be quite tedious, it is often considered that the logarithm is an increasing function so it will be equivalent to the maximized log likelihood denoted by:

$$L(\phi) = \log l(\phi) :$$

$$L(\phi) = \sum_{i=1}^{n} \log(\phi(y_i; \vartheta)).$$

1.6 Goodness-of-Fit Statistics for Model Selection

In model selection, we assume there can be any kind of data, such as real or generated data and a set of models. It is also assumed that statistical inference is model-based. Classically, it is assumed there is at least a single correct (or even true) best model, which suffices as the sole model for making inferences from the data. Although the identity (and parameter values) of that model is unknown, it can be estimated, sometimes very well. Therefore, classical inference often involves a data-based search over the model set for (i.e., the selection of) that single correct model but with estimated parameters. Then the inference is based on the fitted selected model as if it was the only model considered. Model selection uncertainty is ignored. This is considered justified because after all, the single best model has been found. However, many selection methods (e.g., classical stepwise selection) are not even based on an explicit criterion of what a best model is, see Burnham and Anderson (2004).

It is thought that the first step to improve inference under model selection would be to establish a selection criterion, such as the Akaike information criterion (AIC), corrected Akaike information criterion ($AICC$), Bayesian information criterion (BIC) and Hannan-Quinn information criterion ($HQIC$). However, first step is to establish a philosophy about models and data analysis and then find a suitable model selection criterion. In fact, all AIC,$AICC$, BIC, and $HQIC$ depend on the log- likelihood function evaluated by the maximum likelihood estimates (l). The following formulas present the method of calculating all above-mentioned efficiency measures:

$$AIC = -2l + 2k, \qquad AICC = AIC + \frac{2k(k+1)}{n-k-1},$$

and

$$BIC = -2l + k\log(n), \qquad HQIC = 2\log\Big\{\log(n)\big[n-2l\big]\Big\},$$

where l denotes the log-likelihood function evaluated by the maximum likelihood estimates, k is the number of parameters, and n is the sample size. In order to quantify the performance of the current distribution, we adopt the goodness-of-fit statistic for discriminating both the empirical and fitted densities. We use the Kolmogorov-Samirnov $(K-S)$ test test along with the corresponding P-value. The best distribution corresponds to smallest value of AIC,$AICC$,BIC, $HQIC$ and $(K-S)$ values.

1.7 Scope of the Study

The focus of this study is on the generalization of the Burr type X distribution with two parameters by adding one or two parameters to the baseline distribution. The mathematical properties of these new distributions are derived. Simulation studies are carried out to evaluate the performance of the estimators. Real data sets are used to illustrate the flexibility and potentiality of the models. This study has a number of advantages. For one, new distributions with new parameters are presented, and owing to their flexibility it is possible to apply these new distributions in the different fields of science. Moreover, the new distributions have several sub-models. The new models offer better fit (for various data sets) compared to the sub-models based on various criteria. The real data sets used do not involved censored observation

1.8 Objectives of the Study

The most significant goal in this research is to generate three distributions in different ways for the Burr type X distribution. Thus, the objectives of this thesis are as follows:

1. Introduce three new distributions: Beta Burr type X (BBX), Gamma Burr type X (GBX) and Weibull Burr type X (WBX).
2. Investigate and study the statistical properties of the three new distributions.
3. Estimate the unknown parameters in these distributions by using maximum likelihood estimation.

4. Perform simulation studies with different sample sizes and parameter values to assess the performance of models.

5. Application of real data sets to illustrate how the proposed new distributions fit the data sets better than their sub-models and non-nested models.

1.9 Organization of Thesis

The thesis is divided into seven Chapters. Following this introductory chapter is Chapter 2, which presents an overview of baseline distribution and also works related to this research. This chapter presents the history of Burr type X distribution and related research for the three families beta-G , gamma-G and Weibull-G.

Chapter 3 deals with some facts about beta-G family. We develop and define the pdf and cdf of beta Burr type (BBX) distribution and we plot the pdf and hazard rate function. Then, we provide a useful expansion for the pdf of BBX model. Accordingly, we provide several mathematical properties such as limit behavior, quantile function etc. The maximum likelihood method is used to estimate the distribution parameters. Finally, algorithm and Monte Carlo simulation study is carried out to assess the performance of the maximum likelihood estimators the conclusion.

In Chapter 4 we introduce the pdf and cdf of gamma-G family and using this to develop the gamma Burr type X (GBX) distribution. A comprehensive mathematical properties of GBX model is provided like limit behavior, moment etc. Then the method of maximum likelihood is used for estimating the GBX parameter. A Monte Carlo simulation study is used to assess the performance of the maximum likelihood estimators.

In Chapter 5 the development of Weibull Burr type X (WBX) distribution is introduced. The pdf and hazard functions are also plotted. Then the expansion of pdf and cdf are provided. Several mathematical properties of WBX distribution such as moment and, Rényi entropy etc. are obtained. The parameters are estimated using the maximum likelihood method. A Monte Carlo simulation study is achieved to assess the performance the maximum likelihood estimators.

In Chapter 6 the usefulness of the new three models is illustrated by using to the three real data sets. Then, we compare the results to the baseline distribution. Lastly, a brief conclusion is given about the three models and data sets.

Chapter 7 presents the overall summary of the current study, the conclusion and future work.

BIBLIOGRAPHY

Abbasi, B., Hosseinifard, S. Z., and Coit, D. W. (2010). A neural network applied to estimate burr xii distribution parameters. *Reliability Engineering & System Safety*, 95(6):647–654.

Abd, E.-B. A., El-Adll, M. E., and ALOafi, T. A. (2015). Estimation under burr type x distribution based on doubly type ii censored sample of dual generalized order statistics. *Journal of the Egyptian Mathematical Society*, 23(2):391–396.

Afify, A. Z., Yousof, H. M., Cordeiro, G. M., M. Ortega, E. M., and Nofal, Z. M. (2016). The weibull fréchet distribution and its applications. *Journal of Applied Statistics*, pages 1–19.

Agarwal, S. K. and Al-Saleh, J. A. (2001). Generalized gamma type distribution and its hazard rate function. *Communications in Statistics-Theory and Methods*, 30(2):309–318.

Ahmad, K., Fakhry, M., and Jaheen, Z. (1997). Empirical bayes estimation of p(y $<$ x) and characterizations of burr-type x model. *Journal of Statistical Planning and Inference*, 64(2):297–308.

Ahmad Sartawi, H. and Abu-Salih, M. S. (1991). Bayesian prediction bounds for the burr type x model. *Communications in Statistics-Theory and Methods*, 20(7):2307–2330.

Ahsanullah, M. and Albin, J. (1994). Record values from univariate distributions.

Alexander, C., Cordeiro, G. M., Ortega, E. M., and Sarabia, J. M. (2012). Generalized beta-generated distributions. *Computational Statistics & Data Analysis*, 56(6):1880–1897.

Ali Mousa, M. (2001). Inference and prediction for the burr type x model based on records. *Statistics: A Journal of Theoretical and Applied Statistics*, 35(4):415–425.

Aljarrah, M. A., Famoye, F., and Lee, C. (2015). A new weibull-pareto distribution. *Communications in Statistics-Theory and Methods*, 44(19):4077–4095.

Alzaatreh, A. and Ghosh, I. (2015). On the weibull-x family of distributions. *Journal of Statistical Theory and Applications*, 14(2):169–183.

Alzaatreh, A., Ghosh, I., and Said, H. (2014). On the gamma-logistic distribution. *Journal of Modern Applied Statistical Methods*, 13(1):5.

Alzaatreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, 71(1):63–79.

Amoroso, L. (1925). Ricerche intorno alla curva dei redditi. *Annali di matematica pura ed applicata*, 2(1):123–159.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1992). *A first course in order statistics*, volume 54. Siam.

Balakrishnan, N., Ahsanullah, M., and Chan, P. (1995). On the logistic record values and associated inference. *J. Appl. Statist. Sci*, 2(3):233–248.

Barreto-Souza, W., Cordeiro, G. M., and Simas, A. B. (2011). Some results for beta fréchet distribution. *Communications in StatisticsTheory and Methods*, 40(5):798–811.

Bourguignon, M., Silva, R. B., and Cordeiro, G. M. (2014). The weibull–g family of probability distributions. *Journal of Data Science*, 12(1):53–68.

Burnham, K. P. and Anderson, D. R. (2004). Multimodel inference understanding aic and bic in model selection. *Sociological methods & research*, 33(2):261–304.

Burr, I. W. (1942). Cumulative frequency functions. *The Annals of Mathematical Statistics*, 13(2):215–232.

Castellares, F. and Lemonte, A. J. (2014). A new generalized weibull distribution generated by gamma random variables. *Journal of the Egyptian Mathematical Society*, 23:382–390.

Castellares, F., Santos, M. A., Montenegro, L. C., and Cordeiro, G. M. (2015). A gamma-generated logistic distribution: Properties and inference. *American Journal of Mathematical and Management Sciences*, 34(1):14–39.

Cooray, K. (2006). Generalization of the weibull distribution: the odd weibull family. *Statistical Modelling*, 6(3):265–277.

Cordeiro, G. M. and de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, 81(7):883–898.

Cordeiro, G. M., Gomes, A. E., da Silva, C. Q., and Ortega, E. M. (2013). The beta exponentiated weibull distribution. *Journal of Statistical Computation and Simulation*, 83(1):114–138.

Cordeiro, G. M., Ortega, E. M., and Popović, B. V. (2014). The gamma-linear failure rate distribution: theory and applications. *Journal of Statistical Computation and Simulation*, 84(11):2408–2426.

Cordeiro, G. M., Ortega, E. M., and Popović, B. V. (2015). The gamma-lomax distribution. *Journal of Statistical Computation and Simulation*, 85(2):305–319.

Eugene, N., Lee, C., and Famoye, F. (2002). Beta-normal distribution and its applications. *Communications in Statistics-Theory and methods*, 31(4):497–512.

Famoye, F., Lee, C., and Olumolade, O. (2005). The beta-weibull distribution. *Journal of Statistical Theory and Applications*, 4(2):121–136.

Ghitany, M., Atieh, B., and Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4):493–506.

Gomes, A. E., da Silva, C. Q., Cordeiro, G. M., Ortega, E. M., et al. (2013). The beta burr iii model for lifetime data. *Brazilian Journal of Probability and Statistics*, 27(4):502–543.

Gómez-Déniz, E., Calderín-Ojeda, E., and Sarabia, J. M. (2013). Gamma-generalized inverse gaussian class of distributions with applications. *Communications in Statistics-Theory and Methods*, 42(6):919–933.

Gordy, M. B. (1998). Computationally convenient distributional assumptions for common-value auctions. *Computational Economics*, 12(1):61–78.

Gradshteyn, I. (2007). Im ryzhik table of integrals. *Series, and Products, Alan Jeffrey and Daniel Zwillinger (eds.), Seventh edition (Feb 2007)*, 885.

Gupta, R. C., Gupta, P. L., and Gupta, R. D. (1998). Modeling failure time data by lehman alternatives. *Communications in Statistics-Theory and methods*, 27(4):887–904.

Haghighi, F. and Nikulin, M. (2006). A chi-squared test for power generalized weibull family for the head-and-neck cancer censored data. *Journal of Mathematical Sciences (New York)*, 133(3):1333–1341.

Hogg, R. V. and Craig, A. T. (1995). *Introduction to mathematical statistics.(5"" edition)*. Upper Saddle River, New Jersey: Prentice Hall.

Jaheen, Z. (1996). Empirical bayes estimation of the reliability and failure rate functions of the burr type x failure model. *Journal of Applied Statistical Science*, 3(4):281–288.

Jaheen, Z. and Al-Matrafi, B. (2002). Bayesian prediction bounds from the scaled burr type x model. *Computers & Mathematics with Applications*, 44(5):587–594.

Jaheen, Z. F. (1995). Bayesian approach to prediction with outliers from the burr type x model. *Microelectronics Reliability*, 35(1):45–47.

Jeffrey, A. and Zwillinger, D. (2007). *Table of integrals, series, and products*. Academic Press.

Johnson, K. and Kotz, S. (1994). Balakrishnan (1994), continuous univariate distributions. *New York: VViley*.

Kenney, J. F. and Keeping, E. S. (1954). Mathematics of statistics-part one.

Lai, C., Xie, M., and Murthy, D. (2003). A modified weibull distribution. *IEEE Transactions on reliability*, 52(1):33–37.

Lemonte, A. J. et al. (2014). The beta log-logistic distribution. *Brazilian Journal of Probability and Statistics*, 28(3):313–332.

Libby, D. L. and Novick, M. R. (1982). Multivariate generalized beta distributions with applications to utility assessment. *Journal of Educational and Behavioral Statistics*, 7(4):271–294.

Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and weibull families. *Biometrika*, 84(3):641–652.

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica: journal of the Econometric Society*, pages 647–663.

Merovci, F. and Elbatal, I. (2015). Weibull rayleigh distribution: Theory and applications. *Appl. Math*, 9(4):2127–2137.

Moors, J. (1988). A quantile alternative for kurtosis. *The statistician*, pages 25–32.

Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, 42(2):299–302.

Mudholkar, G. S., Srivastava, D. K., and Kollia, G. D. (1996). A generalization of the weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, 91(436):1575–1583.

Murthy, D. P., Xie, M., and Jiang, R. (2004). *Weibull models*, volume 505. John Wiley & Sons.

Nadarajah, S., Cordeiro, G. M., and Ortega, E. M. (2015). The zografos–balakrishnan-g family of distributions: Mathematical properties and applications. *Communications in Statistics-Theory and Methods*, 44(1):186–215.

Nadarajah, S. and Kotz, S. (2004). The beta gumbel distribution. *Mathematical Problems in Engineering*, 2004(4):323–332.

Nadarajah, S. and Kotz, S. (2006a). The beta exponential distribution. *Reliability engineering & system safety*, 91(6):689–697.

Nadarajah, S. and Kotz, S. (2006b). The exponentiated type distributions. *Acta Applicandae Mathematica*, 92(2):97–111.

Nadarajah, S. and Kotz, S. (2006c). Some beta distributions. *BULLETIN-BRAZILIAN MATHEMATICAL SOCIETY*, 37(1):103–125.

Nadarajah, S. and Kotz, S. (2007). Multitude of beta distributions with applications. *Statistics*, 41(2):153–179.

Nadarajah, S. and Pal, M. (2008). Explicit expressions for moments of gamma order statistics. *Bulletin of the Brazilian Mathematical Society, New Series*, 39(1):45–60.

Nasiru, S. and Luguterah, A. (2015). The new weibull-pareto distribution. *Pakistan Journal of Statistics and Operation Research*, 11(1).

Nevzorov, V. B. (2001). *Records: mathematical theory*. American Mathematical Soc.

Ng, K. and Kotz, S. (1995). Kummer-gamma and kummer-beta univariate and multivariate distributions. Technical report, Technical Report 84, Department of Statistics, The University of Hong Kong, Hong Kong.

Oguntunde, P., Owoloko, E., and Balogun, O. (2016). On a new weighted exponential distribution: Theory and application. *Asian Journal of Applied Sciences*, 9(1):1–12.

Raqab, M. Z. and Kundu, D. (2005). Comparison of different estimators of p$[y < x]$for a scaled Burr type x distribution. *Communications in Statistics: Simulation and Computation®*, 34(2):465–483.

Ristić, M. M. and Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, 82(8):1191–1206.

Rrnyi, A. (1961). On measures of entropy and information. In *Fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 547–561.

Sijbers, J., den Dekker, A. J., Scheunders, P., and Van Dyck, D. (1998). Maximum-likelihood estimation of rician distribution parameters. *IEEE Trans. Med. Imaging*, 17(3):357–361.

Sindhu, T. and Aslam, M. (2014). On the parameter of the Burr type x under bayesian principles. *matrix*, 1:0.

Smith, J., Wong, A., and Zhou, X. (2015). Higher order inference for stress–strength reliability with independent Burr-type x distributions. *Journal of Statistical Computation and Simulation*, 85(15):3092–3107.

Smith, R. L. and Naylor, J. (1987). A comparison of maximum likelihood and bayesian estimators for the three-parameter Weibull distribution. *Applied Statistics*, pages 358–369.

Soliman, A. A., Ellah, A. H. A., Abou-Elheggag, N. A., and El-Sagheer, R. M. (2015). Inferences using type-ii progressively censored data with binomial removals. *Arabian Journal of Mathematics*, 4(2):127–139.

Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, pages 1187–1192.

Surles, J. and Padgett, W. (2001). Inference for reliability and stress-strength for a scaled Burr type x distribution. *Lifetime Data Analysis*, 7(2):187–200.

Surles, J. and Padgett, W. (2005). Some properties of a scaled Burr type x distribution. *Journal of Statistical Planning and Inference*, 128(1):271–280.

Tahir, M., Cordeiro, G. M., Mansoor, M., and Zubair, M. (2015). The Weibull-lomax distribution: properties and applications. *Hacettepe Journal of Mathematics and Statistics*.

Tahir, M., Zubair, M., Mansoor, M., Cordeiro, G. M., Alizadeh, M., and Hamedani, G. (2014). A new Weibull-g family of distributions. *Hacet. J. Math. Stat.(2015b). forthcoming*.

Torabi, H. and Hedesh, N. M. (2012). The gamma-uniform distribution and its applications. *Kybernetika*, 48(1):16–30.

Xie, M., Tang, Y., and Goh, T. N. (2002). A modified Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, 76(3):279–285.

Xu, K., Xie, M., Tang, L. C., and Ho, S. (2003). Application of neural networks in forecasting engine systems reliability. *Applied Soft Computing*, 2(4):255–268.

Zea, L. M., Silva, R. B., Bourguignon, M., Santos, A. M., and Cordeiro, G. M. (2012). The beta exponentiated pareto distribution with application to bladder cancer susceptibility. *International Journal of Statistics and Probability*, 1(2):8.

Zografos, K. and Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6(4):344–362.

Zoraghi, N., Abbasi, B., Niaki, S. T. A., and Abdi, M. (2012). Estimating the four parameters of the Burr iii distribution using a hybrid method of variable neighborhood search and iterated local search algorithms. *Applied Mathematics and Computation*, 218(19):9664–9675.