UNIVERSITY FOR DEVELOPMENT STUDIES

POWER SERIES INVERTED KUMARASWAMY DISTRIBUTION WITH APPLICATIONS TO LIFETIME DATA

GODWIN DZAKPASU

2019

UNIVERSITY FOR DEVELOPMENT STUDIES

POWER SERIES INVERTED KUMARASWAMY DISTRIBUTION WITH APPLICATIONS TO LIFETIME DATA

GODWIN DZAKPASU (B.Sc. MATHEMATICAL SCIENCE (STATISTICS OPTION))

UDS/MAS/0004/17

THESIS SUBMITTED TO THE DEPARTMENT OF STATISTICS, FACULTY OF MATHEMATICAL SCIENCES, UNIVERSITY FOR DEVELOPMENT STUDIES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF PHILOSOPHY DEGREE IN APPLIED STATISTICS

NOVEMBER, 2019

DECLARATION

Student

I hereby declare that this thesis is the result of my own original work and that no part of it has been presented for another degree in this University or elsewhere. Related works by others which served as a source of knowledge have been duly referenced.

Candidate's Signature: …………….....…………… Date: ……………..

Name: Godwin Dzakpasu

Supervisors'

I hereby declare that the preparation and presentation of the thesis was supervised in accordance with the guidelines on supervision of thesis laid down by the University for Development Studies.

Supervisor's Signature: ………………………… Date: …………….

Name: Dr. Solomon Sarpong

ABSTRACT

The inverted Kumaraswamy distribution has a drawback in modeling lifetime data which show non-monotone failure rates. Thus, a new class of distributions called power series inverted Kumaraswamy distribution was developed by compounding the inverted Kumaraswamy distribution with zero truncated power series distribution with the goal of making it more flexible. The new class of distributions were developed using the stochastic representation $Z = \min\ (T_1, T_2, ...T_N)$ which is the time to the first failure of a system of identical components that are in a series. The statistical properties such as quantile, moments, moment generating function, stochastic ordering property and order statistics were developed for the new class of distributions. The maximum likelihood method was employed to develop estimators for the parameters. Special sub-distributions namely, Poisson inverted Kumaraswamy distribution, geometric inverted Kumaraswamy distribution, binomial inverted Kumaraswamy distribution and logarithmic inverted Kumaraswamy distribution were developed from the new class of distributions. The failure rate of the special distributions can be increasing, decreasing, bathtub and upside-down bathtub-shaped. Monte Carlo simulations were performed to examine the behavior of the estimators. The applications of the special distributions were illustrated using two lifetime data sets and the results revealed that among all the special distributions, the Geometric inverted Kumaraswamy distribution performs better.

ACKNOWLEDGEMENTS

I express my whole hearted gratitude to the Almighty God for his grace and mercies and for a successful completion. I express my sincere appreciation to my supervisor Dr. Solomon Sarpong under whose guidance, constructive criticisms and suggestions this thesis has become successful. I am very thankful to the Principal of the Navrongo Campus Professor Albert Luguterah for his advice and support.

I equally profess my profound gratitude to my dear friend Dr. Suleman Nasiru for his immense support, advice, kindness, benevolence and contributions throughout the entire project. Words are not enough to describe how grateful I am. My sincere appreciation also goes to all the lecturers of the Statistics Department for their meaningful contributions.

My special thanks also goes to my friend Lea Anzagrah for her support, kindness and contribution to the work. To all my colleagues I say a big thank you. Last but not the least, my deepest gratitude goes to my family for their encouragement, financial support and prayers.

DEDICATION

To the Almighty God and my family for their support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS

AB	Average Bias
AE	Average Estimate
AIC	Akaike Information Criterion
AICc	Akaike Information Criterion Corrected
BIC	Bayesian Information Criterion
BIKUM	Binomial Inverted Kumaraswamy Distribution
CDF	Cumulative Density Function
CP	Coverage Probability
GIKUM	Geometric Inverted Kumaraswamy Distribution
IKUM	Inverted Kumaraswamy Distribution
LIKUM	Logarithmic Inverted Kumaraswamy Distribution
MGF	Moment Generating Function
PDF	Probability Density Function
PIKUM	Poisson Inverted Kumaraswamy Distribution.
PSIKUM	Power Series Inverted Kumaraswamy Distribution
RSME	Root Mean Square Error
SO	Stochastic Orders
MLEs	Maximum Likelihood Estimators/Estimates

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

In statistics, the study of data is the foundation of its analysis. Data can be analyzed with visible techniques or graphical representations such as histograms, bar plots, bar charts, and so on. The above mentioned falls under descriptive statistics which has a limitation. Its drawback is that it is unable to make a generalization to other people or objects. This is where probability distribution comes into play. The use of probability distribution to model real data enhances the characterization of the variability and unpredictability or uncertainty in a data set by determining the patterns of variation. In as much as statistical probability distributions summarize the observations into a brief but comprehensive mathematical form containing a few parameters, they also provide means to analyze the basic structure that govern the data generating mechanism.

To define the probability distribution of a stochastic variable Y, we make use of the following concepts: cumulative distribution function (CDF), $(F(y) = \Pr(Y \leq y))$ probability density function (PDF), $(f(y) = F'(y))$, quantile function $(Q(p) = F^{-1}(p))$, quantile density function, density quantile, $(f_p(p) = f_Y(Q(p))$. In classical statistics, CDF and PDF are considered the most popular techniques of defining most distributions in statistical theory and practice. In statistical modeling, the main objective is finding appropriate probability distributions that adequately describe a data set obtained by surveys, experiments and so on. To achieve this, there are two extensive methods that is, deriving theoretical models from basic assumptions using relations underlying the data and empirical modeling. The latter method is data dependent and is convenient in situations where there is lack of understanding in the data generating process. The underlying goal of this method is determining the best distributional approximation to the data by focusing on versatile families of distributions with sufficient parameters that are capable of producing various shapes and characteristics that match the properties exhibited by the available observations. The former method makes assumptions about the physical characteristics that govern the data generating process and subsequently obtain an appropriate model that satisfies such assumptions or adapting existing models from other disciplines (Atem, 2018).

The challenge in empirical statistical modeling is finding a distribution function's parameter estimates that are as close as possible to the true values of the theoretical model parameters. Based on the degree of accuracy desired, diverse modeling procedures that ensure closeness in the estimate and the true parameter value may be used.

However, we cannot find any statistical distribution that is suitable to the various types of data and so the need to generalize existing distributions or develop new ones. Current research works focus on determining new families of distributions that encompass classical distributions and also provide great versatility in modeling data. As a result, diverse techniques to develop modern models by the addition of extra parameters have been introduced. Most of these generalizations are formed as a clarification of the data generating process and providing a better fit.

In statistics, different data can be analyzed in different ways. The problem with data is making correct inferences about the data. The knowledge of the appropriate distribution of a data set or determining the right distribution of a specified data set enables one to make accurate inferences concerning the data. Classical statistical distributions may not be able to model correctly all the various data sets that exist. This is the motivation for generalizing existing distributions or creating new ones. Statisticians are coming out with new models and families of models that generalizes well-known distributions and also helps in modelling different types of data for instance lifetime data.

Lifetime data modelling has become prominent in the field of survival analysis. Currently, more distributions have been introduced to model different forms of data. Kumaraswamy (1980) developed a distribution with similar traits to the beta distribution but with some advantages, including closed-form CDF. This distribution can be applied to several natural phenomena whose results have lower and upper bounds, such as atmospheric temperatures, individual heights, rainfall data and so on (Kumaraswamy, 1980; Jones, 2009; Golizadeh *et al*., 2011; Sindhu *et al*., 2013; El-Deen *et al*., 2014). The inverted Kumaraswamy distribution proposed by Al-Fattah *et al*. (2017) is an inverse form of the

Kumaraswamy distribution proposed by Kumaraswamy (1980). Recently, extensions of the inverse Kumaraswamy (IKum) distribution have been proposed in literature. These include the Marshall-Olkin IKum distribution (Tomy and Gillariose, 2018) and the generalized IKum distribution (Iqbal *et al*., 2017). Thus, this study seeks to advance another modification of the IKum distribution called the power series IKum distribution with the aim of making it more flexible.

1.2 Problem Statement

The Ikum distribution has been used in modeling lifetime data. However, in many applied instances, the IKum distribution fails to give adequate fits to lifetime data such as the life cycle of machines, human mortality and biomedical data which show non-monotone failure rates.

Current statistical improvements focus on developing new generalizations of the IKum distribution to make it more flexible in modelling data from diverse fields of study. Thus, this study proposes another extension of the IKum distribution with the goal of making it more flexible with regards to modeling lifetime data.

1.3 General Objective

To develop and study the properties of power series IKum distribution and apply it to survival data.

1.4 Specific Objectives

1. To develop a new power series IKum distribution.
2. To derive the statistical properties of the new distribution.
3. To develop estimators for the parameters of the new distribution using maximum likelihood method.
4. To assess the performance of the estimators using simulation.
5. To demonstrate the application of the power series IKum distribution using real data sets.

1.5 Significance of the Study

Developing the Power series inverted Kumaraswamy distribution is premised on a number of motivations which can be applied in some practical situations given below.

1. Due to the stochastic illustration $Z = \min(T_1, T_2, ..., T_N)$ the power series IKum distribution can be applied in many industrial applications and biological organisms.
2. The power series IKum distribution can be used to model adequately the time to the first failure of a system of identical components that are in a series.
3. The power series IKum distribution exhibits some interesting characteristics with non-monotonic failure rates such as bathtub, upside bathtub and increasing-decreasing-

increasing failure rates which are more likely to be encountered in real life situations.

1.6 Thesis Outline

The thesis is organized into five chapters including this one. Chapter two presents literature on modified distributions. Chapter three presents the methodology of the study. Chapter four presents the results and discussions. Finally, chapter five presents the summary, conclusion and recommendations of the study.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter presents several modifications and generalizations of distributions which aim at making them more flexible. This is because this research work is an extension of an existing model.

2.2 Modified Distributions

A lot of research has been done on modifications of previously developed distributions to allow for more flexibility so as to cater for the increasing number of data sets.

Quite recently, Nasiru *et al*. (2018) developed a modern family of models by name the exponentiated generalized power series class. They defined some special distributions of the family and their applications were demonstrated with real sets of data. The results of the simulation revealed the parameters of the distributions were good with regards to the estimation techniques.

Nasiru *et al*. (2018) worked on the Poisson exponentiated Erlang truncated exponential model and the statistical characteristics of this current model were studied and expressed. The potency of the current model was exhibited using real data set and the empirical results obtained showed that the new model was better than the other models it was compared with in terms of goodness-of-fit. The bivariate Poisson exponentiated Erlang-truncated exponential distribution was also proposed.

Tomy and Gillariose (2018) pioneered a new class of continuous model called the Marshall-Olkin IKum Distribution using Marshall-Olkin extended method. Their new method has increasing, decreasing and unimodal pdf as some of its shapes. The new distribution can become a better substitute for the Kumaraswamy distribution and the IKum distribution. Their proposed model enfolded some special sub models and also got some existing distributions by adding appropriate transformations.

Alizadeh *et al*. (2018) advanced a modern family of continuous-models by name the complementary generalized transmuted Poisson-G class. They provided certain characteristics of the model. The importance of the model was shown using applications to two real sets of data.

Jamal *et al*. (2018) proposed the "Generalized Inverted Kumaraswamy-*G*" class of distributions. Once again, after careful assessment and examination, it was realized that the new class of models indicated good flexibility. It out-performed some of the already existing models in literature.

Elbatal *et al*. (2018) came up with a current model which is known by the name Kumaraswamy extension exponential model and it is practically based on the Kumaraswamy distribution. Sub groups of the new distribution were presented in the study. A number of its statistical characteristics were also expressed. The estimators were developed and the performance of the model was exhibited. Finally, how well the model fits real data sets was shown. The new model was compared with three other models.

Unal *et al*. (2018) proposed the "Alpha Power Inverted Exponential" distribution. They provided some of the statistical characteristics which include hazard rate function. The shapes of the latter were determined. The importance of the model was displayed with real data applications. The results obtained indicated that the model provided good fit.

Elgarhy *et al*. (2018) examined and proposed the exponential generalized Kumaraswamy distribution. Parameter estimation was discussed in the work and certain characteristics were derived. The strength of the new model was exhibited by the use of real sets of data and the results clearly showed that the model performed very well.

Elbatal *et al*. (2017) came up with a modern family of models known as the exponential Pareto power series models. They discussed some sub models and developed some characteristics of this model. Further studies revealed that the new model out-performed older ones.

Muhammad and Yahaya (2017) proposed and studied a distribution known as the half-logistic Poisson model. They provided several characteristics of the study and vividly expressed them in the study. It was asserted that their model gave better fits than other ones.

Fattah *et al*. (2017) proposed the exponential transmuted Weibull geometric distribution. The new model has 22 sub distributions. Its performance was examined and this was done with two real sets of data, censored and uncensored. The estimates were assessed using simulation. It was discovered that the model performed better than old ones in literature.

Dey *et al*. (2017) introduced a current model known as the α logarithmic transformed generalized exponential distribution. Special cases of the model were discussed and clearly expressed in the study. The α logarithmic transformed generalized exponential density function has the ability to produce different shapes. Also, the characteristics of the distribution were developed.

Usman *et al*. (2017) pioneered a current model known as the Kumaraswamy Half-logistic distribution. They examined some characteristics of the model and the importance of these features were clearly depicted in the study. Further studies revealed that the model can be a good substitute to previously developed ones.

Mohammed (2017) developed a new distribution by name the log-exponentiated Kumaraswamy model. The shapes of some of its functions were shown in the study and several characteristics were laid down with clarity. The data used to illustrate the performance of the new model were stratified into different working ages. The model was deemed to be a good fit for different kinds of data.

Chakrabarty and Chowdury (2016) spearheaded two probability distributions known as the compounded inverse Weibull distributions. Essential characteristics such as quantiles and moments were developed.

For both models, they exhibited their potentiality on three real sets of data. Their model thrived in comparison with other models.

Alkarni (2016) developed a modern distribution by name generalized extended Weibull power series class of models. He studied the characteristics of this group of models and expressed the process in the research work. The applicability of the models was displayed on a real set of data.

Silva *et al*. (2016) defined the generalized gamma power series family, a distribution with unique characteristics and good for lifetime data. These interesting characteristics were showcased in the study. Quite recently, the current model has become a subject worth exploring. Careful studies revealed that the model out-classed certain older ones developed.

Saboor *et al*. (2016) brought forth a new distribution by name the transmuted exponential weibull geometric distribution. It has about ten models as special cases. How well the model fits data sets was showcased in this research work. The study further revealed that the model performed better than existing ones.

Behairy *et al.* (2016) introduced the Kumaraswamy-Burr Type III distribution which has some special well-known sub-groups. Its characteristics were discussed and a careful study revealed that the new model out-performed other models. Their model has now become an interesting case for most researchers.

Abdelall (2016) advanced the odd generalized exponential modified-Weibull model by making use of the generator developed by Tahir et al. (2015). He studied certain characteristics of this model and compared this distribution with old ones to confirm the effectiveness of the new distribution and based on the result obtained, it was realized that the current model out-performed the others mentioned in the work.

Rodrigues *et al*. (2016) introduced a new distribution known as the exponentiated Kumaraswamy inverse Weibull, a modification of inverse Weibull distribution, which is more extensible than its predecessors and accommodate numerous special cases which include the inverse exponential, inverse Rayleigh and inverse Weibull models. The model parameters estimation was achieved via moments and maximum likelihood estimation methods.

Jafari and Tahmasebi (2015) came up with the Gompertz power series family of distribution which was developed for life data. They made use of certain techniques such as the expectation-maximization algorithm in their study. They also performed simulation studies and determined the importance of the model on real sets of data.

Shafiei *et al*. (2015) in their study came up with a current group of models known as the inverse Weibull distributions. They provided certain characteristics of the model such as moments. To show the flexibility of this distribution, they demonstrated it on real data sets. After comparing with older models, it was realized that the new model out-matches the rest.

Warahena-Liyanage and Pararai (2015) proposed a group of models popularly referred to as the Lindley power series model. They provided certain characteristics and touched on the use of these features. They also presented sub-groups and exhibited the accuracy of the estimates. The strength of the models was illustrated using real data applications.

Tahmsebi *et al*. (2015) pioneered a group of models known as the exponentiated G-power series models. They obtained certain characteristics of the model and also established that the model can be used to analyze different forms of data. The model was discovered to be better than some existing ones.

Abdul-Moniem (2015) introduced a new distribution called exponential Nadarajah and Haghighi's exponential distribution. This model exhibits certain desirable behaviors. By applying the new distribution to real data, it was seen that it can be quite effective to provide better fit than the Nadarajah and Haghighi's exponential distribution.

Nekoukhou and Bidram (2015) developed a trending model by name the exponentiated discrete Weibull model. The new distribution has sub-groups which are clearly expressed in the research work. They derived certain statistical characteristics of the model. The shapes of the model were exhibited and finally, they illustrated the relevance of the model using a real set of data.

Sankara and Anjana (2015) developed a new model which exhibits favorable shapes and characteristics for life data. They tested the effectiveness of the model and also derived some of its characteristics. Its importance was demonstrated on a real data set and the result proved that it provided better fit to the data.

Aryal and Elbatal (2015) proposed the Kumaraswamy modified inverse weibull distribution. They achieved this by extending the modified inverse Weibull distribution. Their work has provided a new model that is capable of modelling real life data. The performance of the new model was ascertained to be very good.

Oguntunde *et al*. (2015) derived a two parameter inverted weighted exponential distribution. Its various statistical characteristics were established. Real life applications were provided to assess the superiority of the inverted weighted exponential distribution over existing distributions. The model shows unimodal and decreasing failure rate as such, this distribution can be used to describe and model real life phenomena with unimodal or decreasing failure rates. For the real life applications provided, the inverted weighted exponential distribution performed better than the weighted exponential distribution.

Tahir *et al*. (2015) made an assertion that the generalized exponential model was applicable to lifetime data with monotonic hazard rate function. However, it is ineffective when some of its functions exhibit certain shapes. Based on these limitations mentioned above, they proposed a new class of models by name the generalized exponential family of distributions.

Several researchers have worked on the generator introduced by Tahir *et al*. (2015). Some studied the odd generalized exponential Gompertz and adopted the generator of the odd generalized exponential class of models. They derived Fisher's information matrix and compared the new distribution's performance with other already known models and it was established that the odd generalized Gompertz out-performed the rest.

Nadarajah *et al*. (2014) engineered a modern group of models referred to as the exponentiated-G geometric class. The flexibility of this model was illustrated and the results indicated that it fitted better than the other models it was compared with. The model exhibited some desirable shapes and certain characteristics were displayed.

Gui *et al*. (2014) introduced a new compound distribution, named the Lindley-Poisson distribution. Their aim of developing this new distribution was to apply it to lifetime data. At the end of their study, it was evident that the developed model contributes significantly to lifetime analysis and its parameters are quite good. It can also be used in place of other models.

Chung and Kang (2014) pioneered the exponentiated Weibull-geometric distribution. The process of compounding was used to develop the new model. Several characteristics were expressed in the work. The performance of the parameters were assessed. The importance of the new model was demonstrated and it was realized that the new model is more flexible.

Bidram and Nekoukhou (2013), came out with a modern group of models called the Double bounded Kumaraswamy-Power Series Class of Distributions. In their study, they showed the compounding process and the techniques used for establishing the characteristics of the new model. The potency of the model was also showcased in the research.

Alkarni (2013) introduced a group of models named a family of truncated binomial lifetime models and studied the characteristics of this class. The new family can generalize several distributions. After assessing its performance, it was shown to be quite good. The importance of the new distribution was also displayed in the study.

Nadarajah and Eljabri (2013) introduced the Kumaraswamy *GP* model. To obtain the new distribution, the generalized Pareto model was extended. Comparison to other distributions was done and from the results, it was clearly evident that the proposed model gave better fit than the other models.

Cordeiro *et al*. (2013) advanced the beta exponentiated Weibull model which is a modification of two other models mentioned in the literature. Their work showed some expressions of certain characteristics in closed form and others were represented by derived explicit expressions. They used the Akaike information criterion (AIC) and realized their model out-performed the two models it was compared to.

Sarhan *et al*. (2013) worked on the exponentiated generalized linear exponential model which has some sub-models. After some careful examinations, it was realized that the new model provided better flexibility than the other models to which it was compared with in the work. It is also evident that the model can be employed as a replacement to previously developed models.

Bakouch *et al*. (2012) came up with a new distribution called the exponentiated exponential binomial distribution. The performance of the model was compared to existing ones and it was evident that the new model performed better. Some of the statistical characteristics were showcased in the study.

Mahmoudi and Shiran (2012) developed a model popularly referred to as the exponentiated Weibull-geometric model. This new model has special cases which were discussed in the study. Certain functions of the model exhibits desirable shapes. They also developed some statistical characteristics of the new model.

Mahmoudi and Jafari (2012) pioneered the generalized exponential power series models with the aim of adding an extra dimension to the exponential power series in order to make it more flexible. Some characteristics of this family were explored and its ability to model different data sets was illustrated to great effect.

The Weibull power series family of distributions determined by Morais and Barreto-Souza (2011) contains alternate models such as the exponential power series models. Various shapes are exhibited by this family models. Also worth noting are the characteristics of the group which were clearly highlighted in the study.

Cordeiro and Castro (2011) developed the Kumaraswamy generated class for modifying other distributions by studying a stochastic variable X which has the baseline CDF $P(x)$. An idea akin to the above was used to consider the distribution functions of Weibull and Inverse Weibull distribution as candidates for $P(x)$ to achieve the Kumaraswamy Weibull model and Kumaraswamy inverse Weibull model.

Pascoa and Cordeiro (2011) introduced and studied the Kumaraswamy generalized gamma distribution, based on the Kumaraswamy distribution (Jones, 2009). Certain statistical characteristics were discussed. Two real data sets were analyzed with this distribution. A review of the research showed that the model is capable of providing good fits.

Chakhandi and Ganjali (2009) introduced the exponential power series class of models. This group of models was shown in the study to be very good and comprises of sub-models. They showcased the usefulness on different sets of data and the results obtained paint a good picture of the model. Their model is capable of replacing older ones.

Cho *et al*. (2009) came up with the acclaimed model known as the exponential extreme value distribution and reviewed critically some of its theoretical characteristics. Some derivations of this model were expressed with utmost clarity. The potency of the model was displayed using censored data.

Gupta and Kundu (1999) started and developed the generalized exponential models. More authors did further studies of this model. The model is known to exhibit favorable behaviors. The applications of the model was demonstrated using real data set.

In the past decades, a number of studies have been made with reference to distributions of power series, Patil (1962) studied certain features of extended power series models. He also considered estimation and other characteristics of the model. He ascertained the members of this family. He further studied the truncated versions and his research has become a headlight for statisticians investigating this area.

Khatri (1959) on some properties of power series distribution generalized what Noack (1950) did, to multivariate distributions. Further, he derived the multivariate modifications of power series distributions with the illustrations of multinomial distributions and extended it to truncated power series distributions.

Noack (1950) studied a family of stochastic variable alongside discrete distributions and delved into its moment and cumulant properties. He also determined that the special cases of these models with their moment and cumulant characteristics belong to this class.

CHAPTER THREE

METHODOLOGY

3.1 Introduction

The techniques for developing a new distribution, estimation of the parameters and the philosophy for the best estimation approach is presented. The focus of this study is on lifetime data and the procedure for modelling lifetime data including its analysis.

3.2 Data and Source

For purposes of the study, two lifetime secondary data sets were used to illustrate the applications of the developed distributions. The first data set comprises of 30 values for the failure time of repairable objects used by Murthy *et al*. (2004) and it is represented in Table 3.1.

Table 3.1: Failure time of repairable objects

1.43	3.46	4.36	0.70	0.63
0.11	2.46	0.40	1.06	1.23
0.71	0.59	1.74	1.46	1.24
0.77	0.74	4.73	0.30	1.97
2.63	1.23	2.23	1.82	1.86
1.49	0.94	0.45	2.37	1.17

The second data consists of 34 observations of vinyl chloride used for monitoring wells in mg/L. The data set presented in Table 3.2 was first used by Bhaumik *et al*. (2009).

Table 3.2: Vinyl chloride for monitoring wells in mg/L

5.1	8.0	2.0	2.5	0.1	6.8
1.2	0.8	0.5	2.3	0.1	1.2
1.3	0.4	5.3	1.0	1.8	0.4
0.6	0.6	3.2	0.2	0.9	0.2
0.5	0.9	2.7	0.1	2.0	
1.1	0.4	2.9	0.2	4.0	

3.3 Inverted Kumaraswamy Distribution

A random variable X is said to have a Kumaraswamy distribution (Kumaraswamy, 1980) if its PDF is given by

$$f\left(x;\alpha,\beta\right)=\alpha\beta x^{\alpha-1}\left(1-x^{\alpha}\right)^{\beta-1},\ 0<x<1,\ \alpha,\ \beta>0, \tag{3.1}$$

and its CDF is given by

$$F(x)=1-(1-x^{\alpha})^{\beta},\ 0<x<1,\ \alpha>0,\ \beta>0. \tag{3.2}$$

The IKum distribution which is the inverse form of the Kumaraswamy distribution is obtained using transformation $T=X^{-1}-1$. Hence, the random variable T is said to have the IKum distribution if its pdf is defined as

$$f(t;\alpha,\beta)=\alpha\beta(1+t)^{-(\alpha+1)}(1-(1+t)^{-\alpha})^{\beta-1},\ t>0;\ \alpha,\ \beta>0, \tag{3.3}$$

and CDF is defined as

$$F(t;\alpha,\beta)=(1-(1+t)^{-\alpha})^{\beta},\ t>0;\ \alpha,\ \beta>0. \tag{3.4}$$

3.4 Power Series Class of Distributions

Let N be a discrete random variable from a power series distribution (truncated at zero) and whose pdf is given as

$$P(N=n)=\frac{a_n\lambda^n}{C(\lambda)},\ n=1,2,3,..., \qquad (3.5)$$

where $C(\lambda)=\sum_{n=1}^{\infty}a_n\lambda^n$ and a_n depends on n and $\lambda\in(0,s)$, s can be ∞.

$C(\lambda)=\sum_{n=1}^{\infty}a_n\lambda^n$ is finite and its first, second and third derivatives with respect to λ are defined and given as $C'(.)$, $C''(.)$ and $C'''(.)$, respectively. The binomial, Poisson, geometric and logarithmic distributions are classified under the power series family of distribution. Table 1 depicts some useful quantities of the zero truncated power series distribution.

Table 3.3: Useful Quantities for Some Power Series Distributions

Distribution	$C(\lambda)$	$C'(\lambda)$	$C^{-1}(\lambda)$	a_n	s
Poisson	$e^{\lambda}-1$	e^{λ}	$\log(1+\lambda)$	$(n!)^{-1}$	$(0,\infty)$
Geometric	$\lambda(1-\lambda)^{-1}$	$(1-\lambda)^{-2}$	$\lambda(1+\lambda)^{-1}$	1	$(0,1)$
Logarithmic	$-\log(1-\lambda)$	$(1-\lambda)^{-1}$	$1-e^{-\lambda}$	n^{-1}	$(0,1)$
Binomial	$(1+\lambda)^m-1$	$\frac{m}{(1+\lambda)^{1-m}}$	$(\lambda+1)^{\frac{1}{m}}-1$	$\binom{m}{n}$	$(0,\infty)$

3.5 Maximum Likelihood Estimation

The maximum likelihood estimation (MLE) can be defined as a method of estimating a statistical model's parameters by choosing the set of values of the model parameters that maximizes the likelihood function. Let $X = (X_1, X_2, ..., X_n)^T$ be a vector of random variables from one of a class of distributions on $\mathbb{R}^n$ and indexed by p-dimensional parameters $\theta = (\theta_1, \theta_2, ..., \theta_p)^T$ where $\theta \in \Omega \subset \mathbb{R}^p$ and $p \le n$. Let $F(X/\theta)$ be the distribution function of X and that the joint density function $f(x_1, x_2, ...x_n / \theta)$ exists. Then the likelihood of θ is the function

$$L(\theta) = f(x_1, x_2, ..., x_n/\theta), \tag{3.6}$$

which is the probability of observing the given data as a function of θ. The values of θ that maximize the likelihood function are referred to as the maximum likelihood estimates of θ, i.e., the value(s) that make(s) the observed data the most probable. If $X = (X_1, X_2, ..., X_n)$ are independent and identically distributed, then the likelihood simplifies to

$$L(\theta) = \prod_{i=1}^{n} f(x_i/\theta). \tag{3.7}$$

Practically, it is often easy to solve the logarithms of the likelihood function, the log-likelihood function, given by

$$\ell(\theta) = \sum_{i=1}^{n} \log f(x_i / \theta).$$

Because logarithm is a monotone function when the likelihood function is maximized, the log-likelihood function is also maximized and vice versa.

The likelihood equations are obtained by setting the first partial derivatives of the log-likelihood function with respect to $\theta_1, \theta_2, ..., \theta_k$ to zero; that are $\frac{\partial \ell(\theta / x_1, x_2, ..., x_n)}{\partial \theta_i} = 0,\ i = 1, 2$, and solving the system of likelihood equations.

3.6 Methods of Evaluating Maximum Likelihood Estimators

Given that $X_1, X_2, ..., X_n$ stands for a random sample size of n from the sampling model $f(x/\theta)$ and θ is defined as an unknown parameter. An estimator of θ obtained by techniques such as method of moment and maximum likelihood estimation, is the sample's function, that is, a statistic $\hat{\theta} = T(X_1, X_2, ..., X_n)$. To study the quality of an estimator or asymptotic properties of the estimator, mean square error and bias (equivalently root mean square error and average bias) are used.

3.6.1 Mean Square Error of an Estimator

Let $\hat{\theta}$ represent the estimator of the unknown parameter θ from the random sample $X_1, X_2, ..., X_n$. Then the deviations from $\hat{\theta}$ to the true θ, $|\hat{\theta} - \theta|$ measures the quality or performance of the estimator. That is the mean square error (MSE) of an estimator $\hat{\theta}$ of a parameter θ is the function of θ defined as

$$\text{MSE}\,\hat{\theta} = \text{E}(\hat{\theta}-\theta)^2 = \text{var}(\hat{\theta}) + (\text{E}(\hat{\theta})-\theta)^2 = \text{var}(\hat{\theta}) + (Bias(\hat{\theta}))^2 \qquad (3.8)$$

The expectation in (3.7) corresponds to the random variables $X_1, X_2, ..., X_n$ since they are the only random components in the expression. The sequence of estimators $\{\hat{\Theta}_n\}$ is weakly consistent or equivalently MSE consistent if $\hat{\theta}_n \to \theta$ in probability as $n \to \infty$. That is, $\forall \in > 0$ if $n \to \infty$

$$\text{P}(|\hat{\theta}_n - \theta| > \in) \to 0. \qquad (3.9)$$

Equivalently, and a sequence of estimators $\hat{\theta}_n$ is weakly consistent if

$$\lim_{n\to\infty} MSE(\hat{\theta}_n) = 0. \qquad (3.10)$$

That is MSE descends to zero if the number of observations increase.

3.6.2 Bias of an estimator

This can be defined as the difference between the estimator's expected value $\hat{\theta}$ and the true value of the parameter θ being estimated. That is,

$$Bias(\hat{\theta}) = \text{E}(\hat{\theta}) - \theta. \qquad (3.11)$$

An estimator is unbiased if $\text{E}(\hat{\theta}) = \theta, \forall \theta.$ For an unbiased estimator $\hat{\theta}$,

$$MSE\,\hat{\theta} = \text{E}(\hat{\theta}-\theta)^2 = \text{var}(\hat{\theta}) \qquad (3.12)$$

and so, an estimator is said to be unbiased if its MSE is equal to its variance. The sequence of estimators $\{\hat{\theta}\}$ is asymptotically unbiased if $E(\hat{\theta}) \to \theta$ as $n \to \infty$.

3.7 Model Comparison and Model Selection Criteria

To show how applicable and flexible our proposed model is, its performance is compared with other established competing models with reference to information lost. Basically, a comparison of various model-selection approaches' ability to discover a real model involves a trade-off between goodness-of-fit and model's parsimony. So, we tend to use information criteria techniques and goodness-of-fit statistics that correct model for complexity, to constrain the model from over fitting to assess the most effective model from a range of totally different models which can have different number of parameters. The Akaike information criteria (AIC), the corrected Akaike information criteria (AICc) and the Bayesian information criterion (BIC) are the most commonly used information criteria. The information criterion selects model with lesser values of AIC, AICc and BIC for a given set of candidate models and specified data set. The Akaike information criterion (AIC) (Akaike, 1974) measures statistical models' quality for an observed data set. It measures information lost when the data generating process is expressed as a statistical model by obtaining an equilibrium in the trade-off between goodness-of-fit of the model and its complexity. Assume we have a statistical model of some data x. Let p represent the number of estimated parameters in the model and $\hat{L}$ the model's likelihood function's maximum value, that is, $\hat{L} = p(x/\hat{\theta})$ where $\hat{\theta}$ are the values of the parameter that maximize the likelihood function. The AIC is then given by

$$AIC = 2p - 2\log(\hat{L}). \tag{3.13}$$

AIC rewards goodness-of-fit, but it also attaches a penalty (to minimize over fitting) that is an increasing function of the number of estimated parameters.

AICc (Hurvich and Tsai, 1989) is AIC with a correction for finite sample sizes defined as follows:

$$AICc = AIC + \frac{2p(p+1)}{(n-p-1)}, \tag{3.14}$$

where n is the number of observations, and p is the number of estimated parameters, that is, AICc is basically AIC with a bigger sanction for additional parameters. It is advisable to use AICc if the sample size is small or when the parameters of the model are too many (Anderson, 2002).

The Bayesian information criterion (BIC) (Schwarz, 1978) is a technique for model selection among a finite set of models. It is possible to increase the likelihood by adding parameters when fitting models but with trade-off for over fitting. Both BIC and AIC try to fix this problem by proposing a sanction term for the number of parameters in the model; the sanction term is larger in BIC than in AIC. The BIC is defined as

$$BIC = \log(n)p - 2\log(\hat{L}), \quad (3.15)$$

where $\hat{L}$ is the value that maximizes the model's likelihood function, n is the sample size and p is the number of parameters to be estimated.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.0 Introduction

The results and discussions of the study are presented in this chapter. The chapter is represented in six parts, which are: PSIKum distribution, statistical properties, estimation of parameters, special distributions, simulations and applications.

4.1 Power Series Inverted Kumaraswamy Distribution

Given the random variable N represents the number of failure causes, $n = 1, 2, ...,$ and the underlying distribution of N is the zero truncated power series distribution. Suppose that $T_1, T_2, ..., T_N$ is a sequence of independent and identically distributed, continuous random variables independent of N that follows IKum distribution with parameters α and β. These random variables denote the lifetimes associated with the failure causes. Usually N is the number of causes and the lifetime T_i is time to failure of the i^{th} subsystem. Let $T_{(1)}$ be the minimum failure time and is defined as:

$$T_{(1)} = \min\{T_1, T_2, ..., T_N\}. \tag{4.1}$$

The conditional CDF of $T_{(1)} \mid N = n$ is given by

$$
\begin{aligned}
F(t \mid N=n) &= 1-P(T>t \mid N) \\
&= 1-\prod_{i=1}^{n}[1-G(t)] \\
&= 1-[1-(1-(1+t)^{-\alpha})^{\beta}]^{n}. \qquad (4.2)
\end{aligned}
$$

Hence, the marginal CDF of $T_{(1)}$ is given by

$$
\begin{aligned}
F(t) &= \sum_{n=1}^{\infty}\frac{a_n\lambda^n}{C(\lambda)}\left\{1-[1-(1-(1+t)^{-\alpha})^{\beta}]^{n}\right\} \\
&= \sum_{n=1}^{\infty}\frac{a_n\lambda^n}{C(\lambda)}-\sum_{n=1}^{\infty}\frac{a_n}{C(\lambda)}\left\{\lambda[1-(1-(1+t)^{-\alpha})^{\beta}]\right\}^{n} \\
&= 1-\frac{C\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta}\right]}{C(\lambda)}, t>0, \alpha>0, \beta>0, \lambda>0. \qquad (4.3)
\end{aligned}
$$

The PSIKum distribution consists of a number of sub-distributions, which include the following: the Poisson IKum (PIKum), the geometric IKum (GIKum), the binomial IKum (BIKum) and the logarithmic IKum (LIKum) distributions. This proposed distribution can be referred to as the PSIKum class of distributions because it comprises of a number of sub-models. The importance of this newly developed model can be illustrated in the fields of medical, industrial and finance where minimum risk problems arise. For instance, in reliability analysis where identical components connected in series requires the failure of just one component for the entire system to shut down (Silva, 2013).

The associated PDF of the PSIKum distribution is obtained by differentiating the marginal CDF in equation (4.3) and is given by

$$f(t)=\lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\frac{C'\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta})\right]}{C(\lambda)},\ t>0, \quad (4.4)$$

where $\alpha>0,\ \beta>0$ are shape parameters and $\lambda>0$ is the scale parameter. Henceforth, we represent a random variable X that follows the PSIKum distribution as $X\sim$ PSIKum $(t;\alpha,\beta,\lambda)$.

The survival function plays a critical role in both biological and engineering studies. For example in the biological sciences and other related fields, it is used to study the average time to the occurrence of events. In the engineering sciences, it is used to estimate the reliability of a system. The survival function of the PSIKum can be expressed as

$$S(t)=1-F(t)$$
$$=\frac{C\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta}\right]}{C(\lambda)},t>0. \quad (4.5)$$

The hazard rate function of a random variable is useful when studying the failure rate of a component. It is the instantaneous rate at which events happen given no previous events (instantaneous failure rate). The hazard rate function of the PSIKum random variable is defined by

$$h(t)=\frac{f(t)}{S(t)}$$
$$=\lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\frac{C'\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta})\right]}{C\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta})\right]},\ t>0. \quad (4.6)$$

Proposition 4.1. The IKum converges to the PSIKum when $\lambda\to 0^{+}$.

Proof. Since $C(\lambda)=\sum_{n=1}^{\infty}a_n\lambda^n$, we have

$$F(t) = 1 - \frac{C\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta}\right]}{C(\lambda)},$$

$$= 1 - \frac{\sum_{n=1}^{\infty} a_n \lambda^n}{C(\lambda)} \left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^n.$$

Considering $\lambda \to 0^+$, we obtain

$$\lim_{\lambda \to 0^+} F(t) = 1 - \lim_{\lambda \to 0^+} \frac{\sum_{n=1}^{\infty} a_n \lambda^n \left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^n}{\sum_{n=1}^{\infty} a_n \lambda^n}.$$

By applying L'Hôpital's rule, we obtain

$$\lim_{\lambda \to 0^+} F(t) = 1 - \lim_{\lambda \to 0^+} \frac{\sum_{n=1}^{\infty} n a_n \lambda^{n-1} \left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^n}{\sum_{n=1}^{\infty} n a_n \lambda^{n-1}}$$

$$= 1 - \lim_{\lambda \to 0^+} \frac{a_1\left[1-(1-(1+t)^{-\alpha})^{\beta}\right] + \sum_{n=2}^{\infty} n a_n \lambda^{n-1} \left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^n}{a_1 + \sum_{n=2}^{\infty} n a_n \lambda^{n-1}}$$

$$= 1 - \frac{a_1\left[1-(1-(1+t)^{-\alpha})^{\beta}\right]}{a_1}$$

$$= (1-(1+t)^{-\alpha})^{\beta}.$$

This completes the proof.

Proposition 4.2. The PDF of the PSIKum distribution can be expressed as an infinite mixture of the density of the smallest order statistic of the IKum distribution with parameters α and β.

Proof. Using $C(\lambda) = \sum_{n=1}^{\infty} a_n \lambda^n$ and $C'(\lambda) = \sum_{n=1}^{\infty} n a_n \lambda^{n-1}$, the PDF of the PSIKum distribution can be expressed as

$$f(t) = \lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\frac{\sum_{n=1}^{\infty} na_n\left[\lambda(1-(1-(1+t)^{-\alpha})^{\beta})\right]^{n-1}}{C(\lambda)}$$

$$= \lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\frac{\sum_{n=1}^{\infty} na_n\lambda^{n-1}\left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^{n-1}}{C(\lambda)}$$

$$= \sum_{n=1}^{\infty}\frac{na_n\lambda^n}{C(\lambda)}\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^{n-1},$$

but

$$P(N=n) = \frac{a_n\lambda^n}{C(\lambda)}$$

$$\Rightarrow f(t) = \sum_{n=1}^{\infty} P(N=n)g_1(t),$$

where $g_{(1)}(t) = n\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\left[1-(1-(1+t)^{-\alpha})^{\beta})\right]^{n-1}$

is the density function of the smallest order statistic of the IKum. Thus the proof is complete.

4.2 Statistical Properties

Most at times, it is important to derive the statistical properties when new distributions are developed. This section presents statistical properties such as the quantile function, moments, moment generating function (MGF), stochastic ordering property and order statistics.

4.2.1 Quantile function

The quantile function or the inverse CDF of a random variable is very useful when generating random numbers from a given probability distribution. It can also be used to describe some properties of a distribution such as the median, kurtosis and skewness.

Proposition 4.3. The PSIKum quantile is given by

$$t_u = \left[1 - \left(1 - \frac{C^{-1}((1-u)C(\lambda))}{\lambda} \right)^{\frac{1}{\beta}} \right]^{\frac{-1}{\alpha}} - 1, u \in [0,1], \tag{4.7}$$

where C^{-1} is defined as the inverse of C.

Proof. By definition, the quantile function is given by

$$F^{-1}(u) = \inf \{ t_u : F(t_u) > u \}, 0 \leq u \leq 1. \tag{4.8}$$

If F is strictly increasing and continuous, then $F^{-1}(u)$ is the unique real number t_u such that $F(t_u) = u$. Thus, equating the CDF of the PSIKum distribution to u and solving for t_u yields

$$t_u = \left[1 - \left(1 - \frac{C^{-1}((1-u)C(\lambda))}{\lambda} \right)^{\frac{1}{\beta}} \right]^{\frac{-1}{\alpha}} - 1, \; u \in [0,1].$$

This completes the proof.

It is evident that the quantile function of the PSIKum class of distributions is tractable and can be used for generating random numbers from the distributions. Sometimes the data may contain outliers or extreme values and the median may be required as the most appropriate measure of central tendency rather than the mean. The median of the PSIKum class of distributions is obtained by substituting $u = 0.5$ into the quantile function. Thus, the median is given by

$$t_{0.5} = \left[1 - \left(1 - \frac{C^{-1}(0.5C(\lambda))}{\lambda} \right)^{\frac{1}{\beta}} \right]^{\frac{-1}{\alpha}} - 1. \qquad (4.9)$$

4.2.2 Moments

The moments of a random variable are essential in statistical inference. They are used to investigate important characteristics of a distribution such as the measures of central tendency, measures of dispersion and measures of shapes. In this subsection, the r^{th} non-central moment of the PSIKum random variable is derived.

Proposition 4.4. If T ~ PSIKum $(t;\alpha,\beta,\lambda)$, then the r^{th} non-central moment of T is given by

$$u_r^{'} = \beta \sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\sum_{j=0}^{r} nP(N=n)(-1)^{r+i-j}\binom{n-1}{i}\binom{r}{j}\times B\left(1-\frac{j}{\alpha}, \beta(i+1)\right), j<\alpha, r=1,2,..., \qquad (4.10)$$

where $B(a,b)=\int_0^1 y^{a-1}(1-y)^{b-1}dy$ is the beta function.

Proof. By definition, the r^{th} non-central moment of a continuous random variable T with the support $(0,\infty)$ is given by

$$u_r^{'} = E(T^r) = \int_0^{\infty} t^r f(t)dt.$$

From Proposition 4.2, we have

$$u_r' = \int_0^\infty t^r \sum_{n=1}^\infty P(N=n) g_{(1)}(t) dt$$

$$= \alpha\beta \sum_{n=1}^\infty nP(N=n) \int_0^\infty t^r (1+t)^{-\alpha-1} (1-(1+t)^{-\alpha})^{\beta-1} \left[1-(1-(1+t)^{-\alpha})^{\beta}\right]^{n-1} dt.$$

Using the binomial theorem $(1-z)^{\eta-1} = \sum_{i=0}^\infty (-1)^i \binom{\eta-1}{i} z^i, |z|<1$ and the fact that $0 < \left(1-(1+t)^{-\alpha}\right)^{\beta} < 1,$ the r^{th} non-central moment can further be expressed as

$$u_r' = \alpha\beta \sum_{n=1}^\infty nP(N=n) \int_0^\infty t^r (1+t)^{-\alpha-1} (1-(1+t)^{-\alpha})^{\beta-1} \sum_{i=0}^\infty (-1)^i \binom{n-1}{i} \left(1-(1+t)^{-\alpha}\right)^{\beta i} dt$$

$$= \alpha\beta \sum_{n=1}^\infty \sum_{i=0}^\infty nP(N=n)(-1)^i \binom{n-1}{i} \int_0^\infty t^r (1+t)^{-\alpha-1} \left(1-(1+t)^{-\alpha}\right)^{\beta(i+1)-1} dt.$$

Let $y = (1+t)^{-\alpha}$, by change of subject $t = y^{\frac{-1}{\alpha}} - 1$. As $t \to 0, y \to 1$ and as $t \to \infty, y \to 0$. Also, $-dy = \alpha(1+t)^{-\alpha-1} dt$. Hence,

$$u_r' = -\beta \sum_{n=1}^\infty \sum_{i=0}^\infty nP(N=n)(-1)^i \binom{n-1}{i} \int_1^0 \left(y^{\frac{-1}{\alpha}} - 1\right)^r (1-y)^{\beta(i+1)-1} dy.$$

Using the expansion $(x+a)^v = \sum_{j=0}^\infty \binom{v}{j} x^j a^{v-j}$. Thus,

$$u_r' = \beta \sum_{n=1}^\infty \sum_{i=0}^\infty \sum_{j=0}^r nP(N=n)(-1)^{r+i-j} \binom{n-1}{i} \binom{r}{j} \int_0^1 y^{\frac{-j}{\alpha}} (1-y)^{\beta(i+j)-1} dy$$

But

$$B(a,b)=\int_0^1 y^{a-1}(1-y)^{b-1}dy$$

$\Rightarrow$

$$u_r^{'}=\beta\sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\sum_{j=0}^{r}nP(N=n)(-1)^{r+i-j}\binom{n-1}{i}\binom{r}{j}B\left(1-\frac{j}{\alpha},\beta(i+1)\right), j<\alpha, r=1,2,....$$

This completes the proof for r^{th} non-central moment.

4.2.3 Moment Generating Function

The MGFs are special functions employed to establish the moments if they exist for a random variable and functions of moments such as mean and variance, kurtosis and skewness in a much simpler way.

Proposition 4.5. If T ~ PSIKum $(t;\alpha,\beta,\lambda)$, then the MGF is

$$M_T(z)=\beta\sum_{r=0}^{\infty}\sum_{n=1}^{\infty}\sum_{i=0}^{\infty}\sum_{j=0}^{r}\frac{(-1)^{r+i-j}z^r nP(N=n)}{r!}\binom{n-1}{i}\binom{r}{j}\times B\left(1-\frac{j}{\alpha},\beta(i+1)\right), j<\alpha. \quad (4.11)$$

Proof. If the MGF premised on the support $(0,\infty)$, exist, then

$$M_T(z)=E\left(e^{zT}\right)=\int_0^{\infty}e^{zt}f(t)dt.$$

Employing Taylor series expansion, the MGF can be expressed as

$$M_T(z)=\sum_{r=0}^{\infty}\frac{z^r}{r!}u_r^{'},$$

where $u_r^{'}$ is the r^{th} non-central moment. Hence, substituting the r^{th} non-central moment gives the MGF as

$$M_T(z) = \beta \sum_{r=0}^{\infty} \sum_{n=1}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{r} \frac{(-1)^{r+i-j} z^r nP(N=n)}{r!} \binom{n-1}{i} \binom{r}{j} B\left(1 - \frac{j}{\alpha}, \beta(i+1)\right), j < \alpha.$$

This completes the proof.

4.2.4 Stochastic Ordering Property

Stochastic orders (SO) are very applicable in several fields of applied probability and statistics. In the fields of reliability and maintainability theory, SO have important applications in many areas. Some examples are defining notions of positive and negative aging, bounding system reliabilities and availability, and comparing maintenance policies (Ohnishi, 2002). Stochastic ordering is the common way of ordering mechanism in lifetime distributions. A random variable T_1 is said to be greater than a random variable T_2 in likelihood ratio order if $f_{T_1}(t) / f_{T_2}(t)$ is an increasing function of t.

Proposition 4.6. Let $\mathrm{T}_1 \sim \mathrm{PSIKum}\ (t; \alpha, \beta, \lambda)$ and $\mathrm{T}_2 \sim \mathrm{IKum}\ (t; \alpha, \beta)$, then T_1 is smaller than T_2 in likelihood ratio order $(T_1 \leq_{lr} T_2)$ provided $\lambda > 0$.

Proof. The ratio of the densities T_1 and T_2 is

$$\frac{f_{T_1}(t)}{f_{T_2}(t)} = \frac{\lambda C'\left[\lambda\left(1-(1-(1+t)^{-\alpha})^{\beta}\right]}{C(\lambda)}.$$

Taking the first derivative of $f_{T_1}(t) / f_{T_2}(t)$ with respect to t, we have

$$\frac{d}{dt}\frac{f_{T_1}(t)}{f_{T_2}(t)} = -\lambda^2\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}\frac{C''\left[\lambda\left(1-(1-(1+t)^{-\alpha})^{\beta}\right)\right]}{C(\lambda)}.$$

Since $\frac{d}{dt}\frac{f_{T_1}(t)}{f_{T_2}(t)} < 0$ for all t, T_1 is smaller than T_2 in likelihood ratio order. That is $T_1 \leq_{lr} T_2$. This completes the proof.

4.2.5 Order statistics

Order statistics are very important tools in non-parametric statistics and inference. They are derived from transformation that involves the ordering of an entire set of observations on a random variable. Since order statistics have hordes of applications in several areas of statistics, it is imperative to derive some common order statistics for the PSIKum class of distributions. Suppose $T_1, T_2, ..., T_n$ are independent identically distributed random sample of size n from PSIKum class of distributions with CDF $F(x)$ and PDF $f(x)$. Let $T_{1:n} \leq T_{2:n} \leq T_{3:n} \leq ... \leq T_{n:n}$ represent the order statistics obtained from the sample. The PDF of the k^{th} order statistic, for $k = 1, 2, ..., n$ is given by

$$f_{k:n}(t) = \frac{n!}{(k-1)!(n-k)!} f(t)[F(t)]^{k-1}[1-F(t)]^{n-k}.$$

Since $0 < 1-F(t) < 1$, using the binomial expansion

$$f_{k:n}(t) = \frac{n!}{(k-1)!(n-k)!}\sum_{i=0}^{n-k}(-1)^i\binom{n-k}{i}\left[F(t)\right]^{K+i-1} f(t) \quad . \tag{4.13}$$

Substituting the CDF and PDF of the PSIKum into equation (4.13), we have

$$f_{k:n}(t)=\frac{n!\lambda\alpha\beta(1+t)^{-\alpha-1}\left(1-(1+t)^{-\alpha}\right)^{\beta-1}}{(k-1)!(n-k)!}\times$$

$$\sum_{i=0}^{n-k}(-1)^{i}\binom{n-k}{i}\left[1-\frac{C\left(\lambda\left(1-(1-(1+t)^{-\alpha})\right)^{\beta}\right)}{C(\lambda)}\right]^{k+i-1}\times$$

$$\frac{C^{'}\left[\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]}{C(\lambda)}. \qquad (4.15)$$

4.3 Parameter Estimation

To illustrate the applications of the developed distribution with regards to modeling real data sets, it is vital to develop estimators for estimating the parameters of the distribution. In this section, estimators are developed for estimating the parameters of the PSIKum class of distributions. Suppose $t_1, t_2, \ldots, t_n$ are possible outcomes of a random sample of size n from $T \sim$ PSIKum $(t;\alpha,\beta,\lambda)$, then the total log-likelihood function is given by

$$\ell=n\log(\lambda\alpha\beta)-(\alpha+1)\sum_{i=1}^{n}\log(1+t_i)+(\beta-1)\sum_{i=1}^{n}\log(1-(1+t_i)^{-\alpha})$$
$$-n\log\left(C(\lambda)\right)+\sum_{i=1}^{n}\log\left[C^{'}\left[\lambda\left(1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right)\right]\right]. \qquad (4.16)$$

The first derivatives of the total log-likelihood function with respect to the parameters are;

$$\frac{\partial\ell}{\partial\lambda}=\frac{n}{\lambda}-\frac{nC^{'}(\lambda)}{C(\lambda)}+\sum_{i=1}^{n}\frac{\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]C^{''}\left[\lambda\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]\right]}{C^{'}\left[\lambda\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]\right]}. \qquad (4.17)$$

$$\frac{\partial \ell}{\partial \beta}=\frac{n}{\beta}+\sum_{i=1}^{n}\log\left(1-(1+t_i)^{-\alpha}\right)$$
$$-\sum_{i=1}^{n}\frac{\left[\lambda\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]C''\left[\lambda\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]\right]\log\left(1-(1+t_i)^{-\alpha}\right)}{C'\left[\lambda\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]\right]}. \quad (4.18)$$

$$\frac{\partial \ell}{\partial \alpha}=\frac{n}{\alpha}-\sum_{i=1}^{n}\log(1+t_i)+(\beta-1)\sum_{i=1}^{n}\frac{(1+t_i)^{-\alpha}\log(1+t_i)}{\left(1-(1+t_i)^{-\alpha}\right)}$$
$$-\sum_{i=1}^{n}\frac{\lambda\beta\left(1-(1+t_i)^{-\alpha}\right)^{\beta-1}(1+t_i)^{-\alpha}C''\left[\lambda\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]\right]\log(1+t_i)}{C'\left[\lambda\left[1-\left(1-(1+t_i)^{-\alpha}\right)^{\beta}\right]\right]}. \quad (4.19)$$

The normal equations that need to be solved simultaneously to obtain the maximum likelihood estimates of the parameters are obtained by equating equations (4.17), (4.18) and (4.19) to zero. That is, $\frac{\partial \ell}{\partial \lambda}=0, \frac{\partial \ell}{\partial \beta}=0, \frac{\partial \ell}{\partial \alpha}=0.$ The resulting normal equations do not have closed form and so the maximum likelihood estimates are obtained by solving the equations using numerical methods.

4.4 Special Distributions

In this section, we present the CDF, PDF and the hazard functions of the special sub-models of the PSIKum class of distributions. These are: the PIKum, GIKum, BIKum and LIKum distributions.

4.4.1 Poisson Inverted Kumaraswamy Distribution

The zero truncated Poisson distribution is a special case of the power series distribution with $a_n = \frac{1}{n!}$, $C^{'}(\lambda) = e^{\lambda}$, and $C(\lambda) = e^{\lambda} - 1$, $(\lambda > 0)$.

From equation (4.3), the CDF of the PSIKum distribution is

$$F(t) = \frac{e^{\lambda} - e^{\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)}}{e^{\lambda} - 1}, t > 0, \qquad (4.21)$$

where $\alpha > 0, \beta > 0$ are shape parameters and $\lambda > 0$ is the scale parameter. Figure 4.1 shows the plot of the CDF of the PIKum distribution for certain chosen values of the parameters.

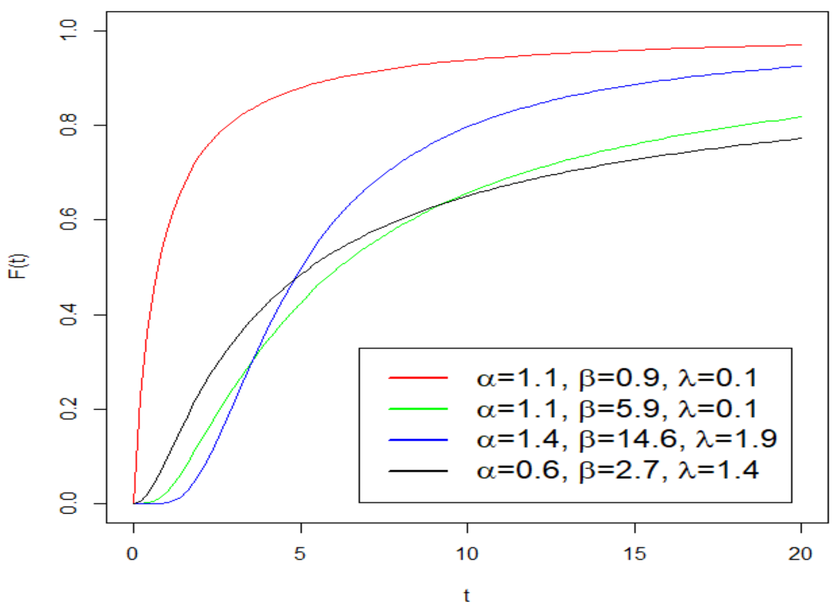

Figure 4.1: Plot of CDF of PIKum

The PIKum distribution PDF is given by

$$f(t)=\lambda\alpha\beta(1+t)^{-\alpha-1}\left(1-(1+t)^{-\alpha}\right)^{\beta-1}\frac{e^{\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)}}{e^{\lambda}-1}, t>0. \qquad (4.22)$$

Figure 4.2 depicts the PDF of the PIKum distribution for some chosen values of the parameters. It is shown clearly from the plot that for certain chosen values of the parameters, the PDF can be approximately decreasing, upside-down bathtub and right skewed.

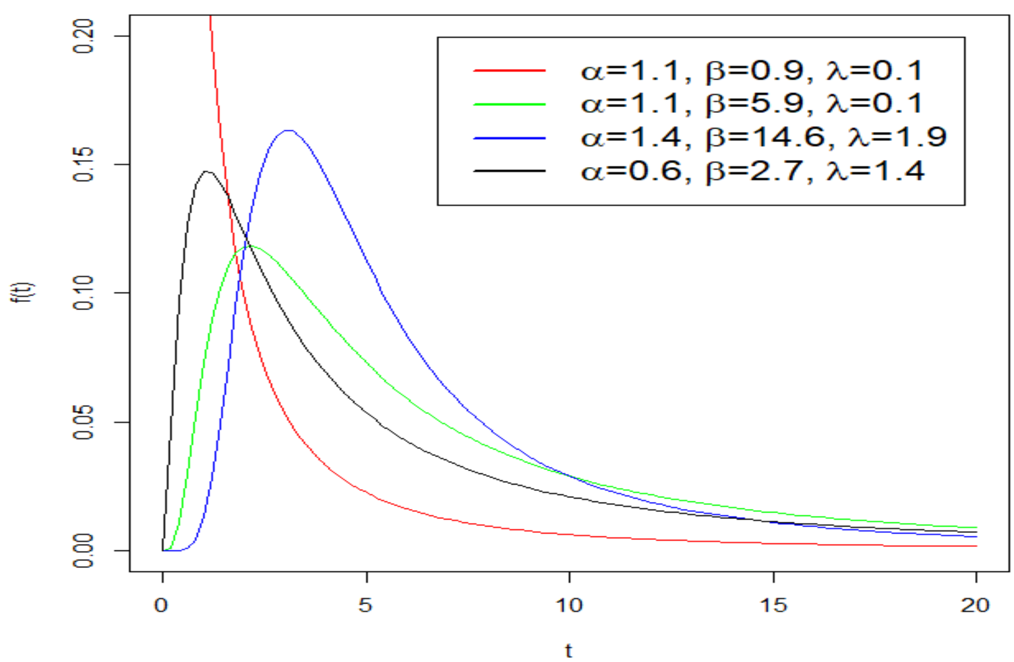

Figure 4.2: Plot of PDF of PIKum

The PIKum distribution hazard function is given by

$$h(t)=\lambda\alpha\beta(1+t)^{-\alpha-1}\left(1-(1+t)^{-\alpha}\right)^{\beta-1}\frac{e^{\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)}}{e^{\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)}-1}, t>0. \qquad (4.23)$$

The plot of the PIKum hazard function is given in Figure 4.3. It can be seen that the hazard rate function shows decreasing, right skewed and upside-down bathtub failure rates.

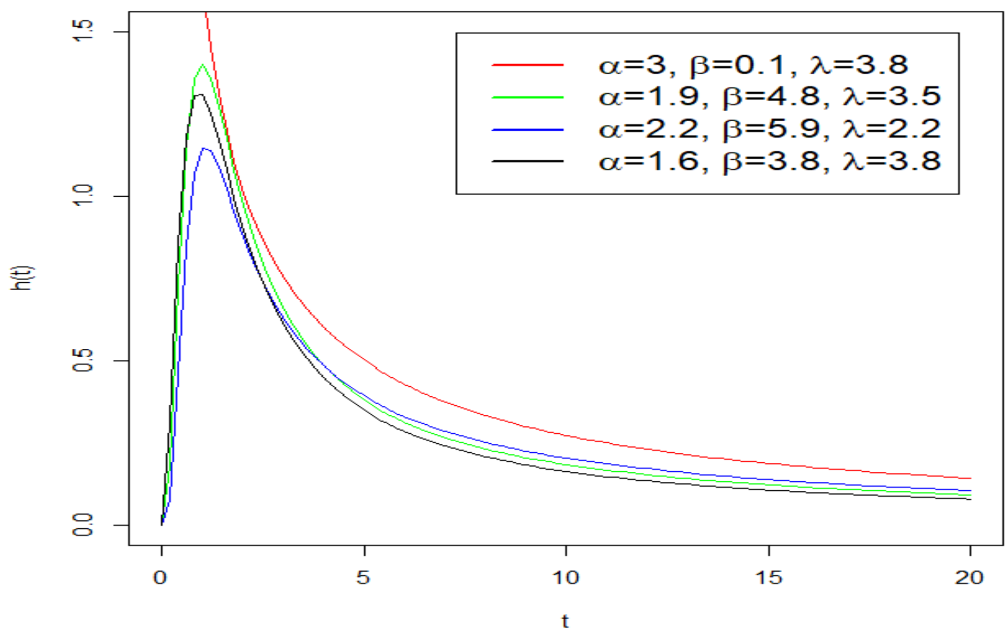

Figure 4.3: Plot of the PIKum hazard function

4.4.2 Geometric Inverted Kumaraswamy Distribution

The zero truncated geometric distribution is a special case of the power series distribution with $a_n = 1,\ C(\lambda) = \dfrac{\lambda}{1-\lambda}$, and $C^{'}(\lambda) = \dfrac{1}{(1-\lambda)^2}$, $(0 < \lambda < 1)$. From equation (4.3), the CDF of the PSIKum distribution is given by

$$F(t) = 1 - \frac{(1-\lambda)\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)}{1-\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)},\ t > 0, \tag{4.24}$$

where $\alpha > 0, \beta > 0$ are shape parameters and $0 < \lambda < 1$ is a scale parameter. It is important to note that λ is also valid for $(-\infty, 1)$. Figure 4.4 shows the CDF of the GIKum distribution for certain parameter values.

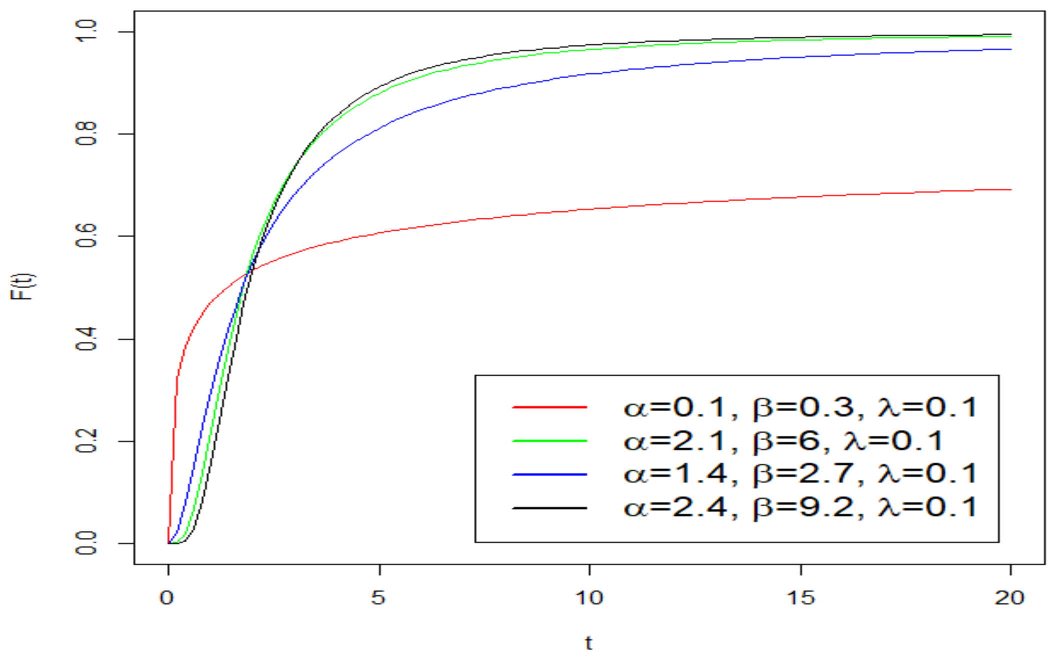

Figure 4.4: Plot of CDF of GIKum

The associated GIKum PDF is

$$f(t)=\frac{(1-\lambda)\alpha\beta(1+t)^{-\alpha}\left(1-(1+t)^{-\alpha}\right)^{\beta-1}}{\left[1-\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]^{2}},\ t>0. \tag{4.25}$$

Figure 4.5 is the plot of the GIKum distribution PDF for a few chosen values. It is clearly shown from the plot that the PDF of the GIKum can depict right skewed, upside down bathtub shapes and decreasing failure rates.

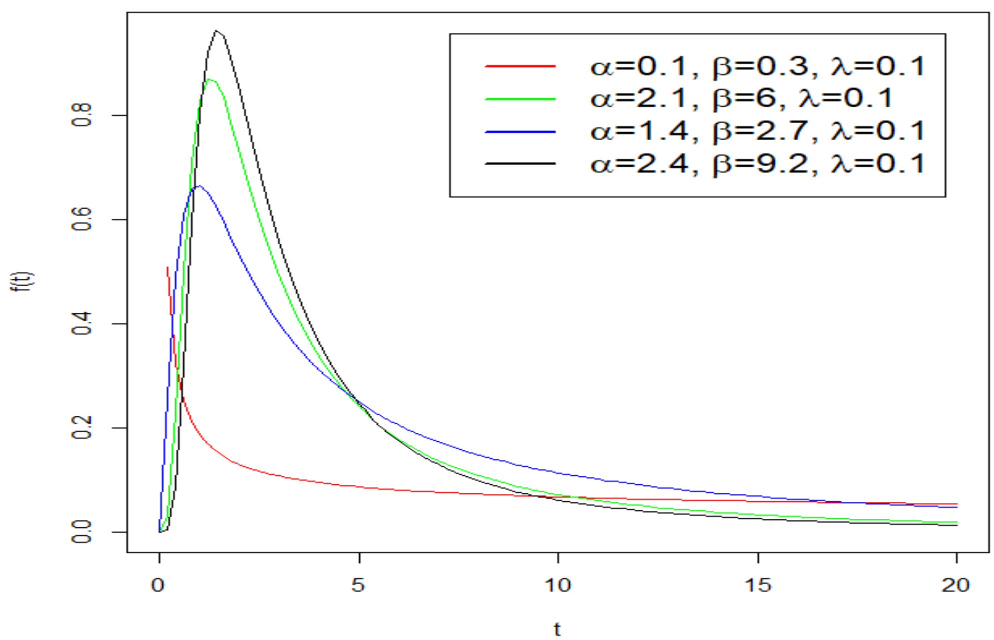

Figure 4.5: Plot of PDF of GIKum

The GIKum hazard function is

$$h(t)=\frac{\alpha\beta(1+t)^{-\alpha-1}\left(1-(1+t)^{-\alpha}\right)^{\beta-1}}{\left[1-\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]\left[\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]},\ t>0. \tag{4.26}$$

The plot of GIKum hazard function are shown in Figure 4.6.

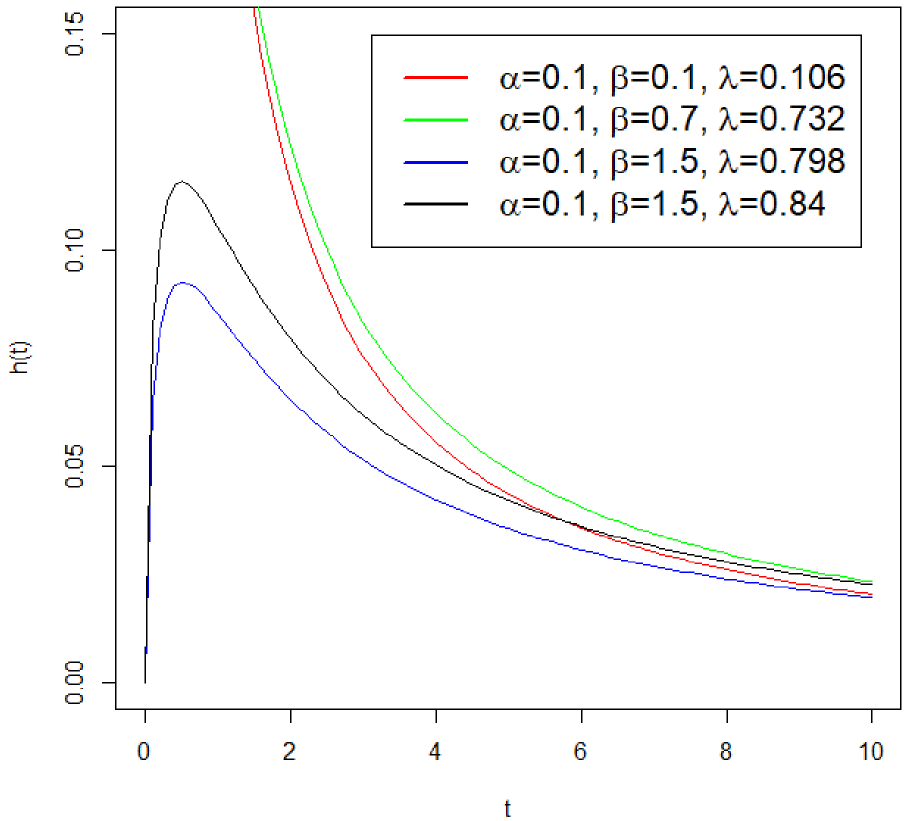

Figure 4.6: Plot of the GIKum hazard function

4.4.3 Binomial Inverted Kumaraswamy Distribution

The zero truncated binomial distribution is a special case of the power series distribution with $a_n = \binom{m}{n}$, $C(\lambda) = (1+\lambda)^m - 1$ and

$C^{'}(\lambda) = m(1+\lambda)^{m-1}$, $(\lambda > 0)$, where $m(n \le m)$ is the number of replicas and is a positive integer. Using equation (4.3), the CDF of the BIKum distribution is given by

$$F(t) = 1 - \frac{\left[1+\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]^{m} - 1}{(1+\lambda)^{m} - 1}, \ t > 0, \qquad (4.27)$$

where $\alpha > 0, \beta > 0$ are shape parameters and $\lambda > 0$ is the scale parameter. Figure 4.7 shows the CDF of the BIKum for some chosen parameter values and $m = 5$.

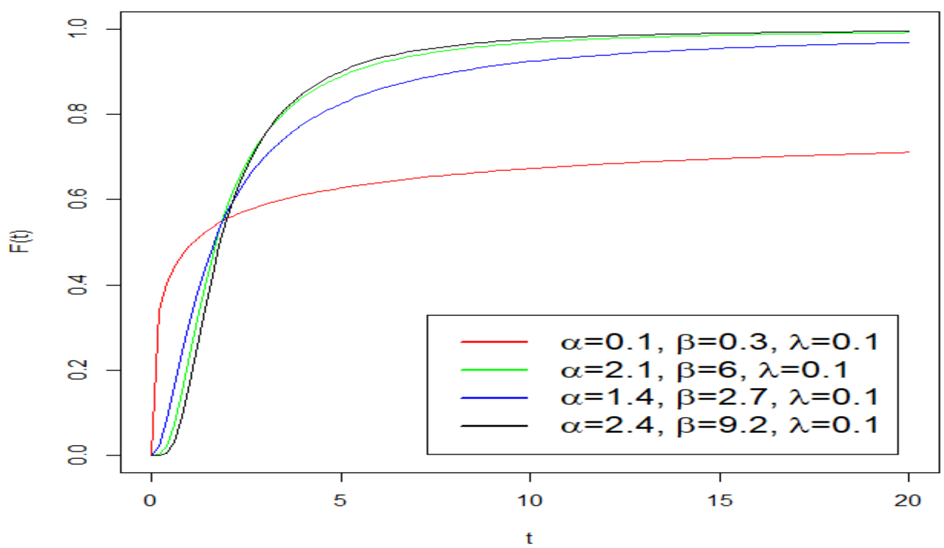

Figure 4.7: Plot of BIKum CDF

The associated BIKum PDF is

$$f(t) = m\lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1} \times \frac{[1+\lambda(1-(1-(1+t)^{-\alpha})^{\beta})]^{m-1}}{\left((1+\lambda)^{m}-1\right)}, \; t>0. \qquad (4.28)$$

The plot of PDF of the BIKum distribution for some selected values of the parameters exhibits decreasing failure rates, and right skewed shapes with varying degrees of kurtosis as shown in Figure 4.8.

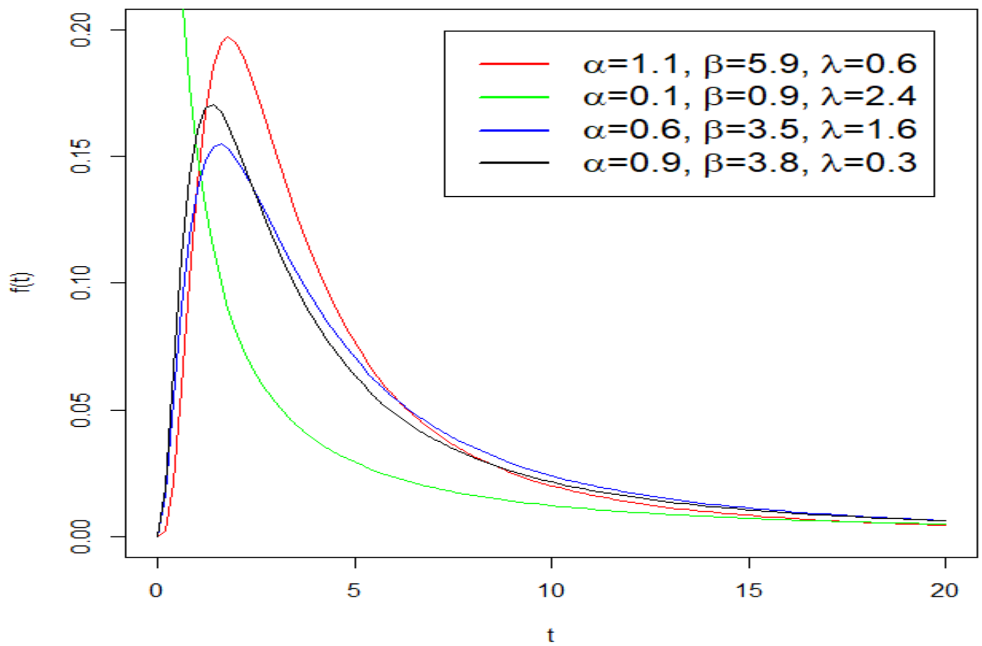

Figure 4.8: Plot of BIKum CDF

The BIKum hazard function is

$$h(t) = m\lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1} \times \frac{[1+\lambda(1-(1-(1+t)^{-\alpha})^{\beta})]^{m-1}}{\left[1+\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]^{m}-1}, \; t>0. \qquad (4.29)$$

The plot of the BIKum hazard function for $m = 5$ are shown in Figure 4.9. It can be seen that the hazard rate function exhibits decreasing, right skewed and upside-down bathtub failure rates.

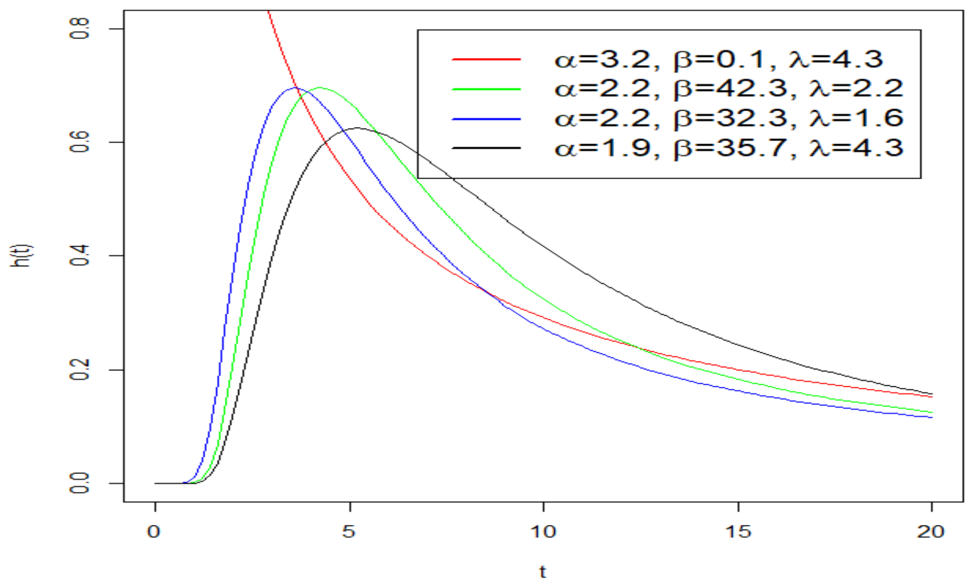

Figure 4.9: Plot of BIKum hazard rate function

4.4.4 Logarithmic Inverted Kumaraswamy Distribution

The zero truncated logarithmic distribution is a special case of the power series distribution with $a_n = \frac{1}{n}$ $, C(\lambda) = -\log(1-\lambda)$ and $C'(\lambda) = (1-\lambda)^{-1}$, $(0 < \lambda < 1)$. Using equation (4.3) the CDF of the LIKum distribution is given by

$$F(t) = 1 - \frac{\log\left[1-\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]}{\log(1-\lambda)}, t > 0, \tag{4.30}$$

where $\alpha > 0, \beta > 0$ are shape parameters and $0 < \lambda < 1$ is a scale parameter. It is worth noting that the parameter λ is also valid for $(-\infty, 1)$. Figure 4.10 shows the CDF of the LIKum distribution for some chosen values of the parameters.

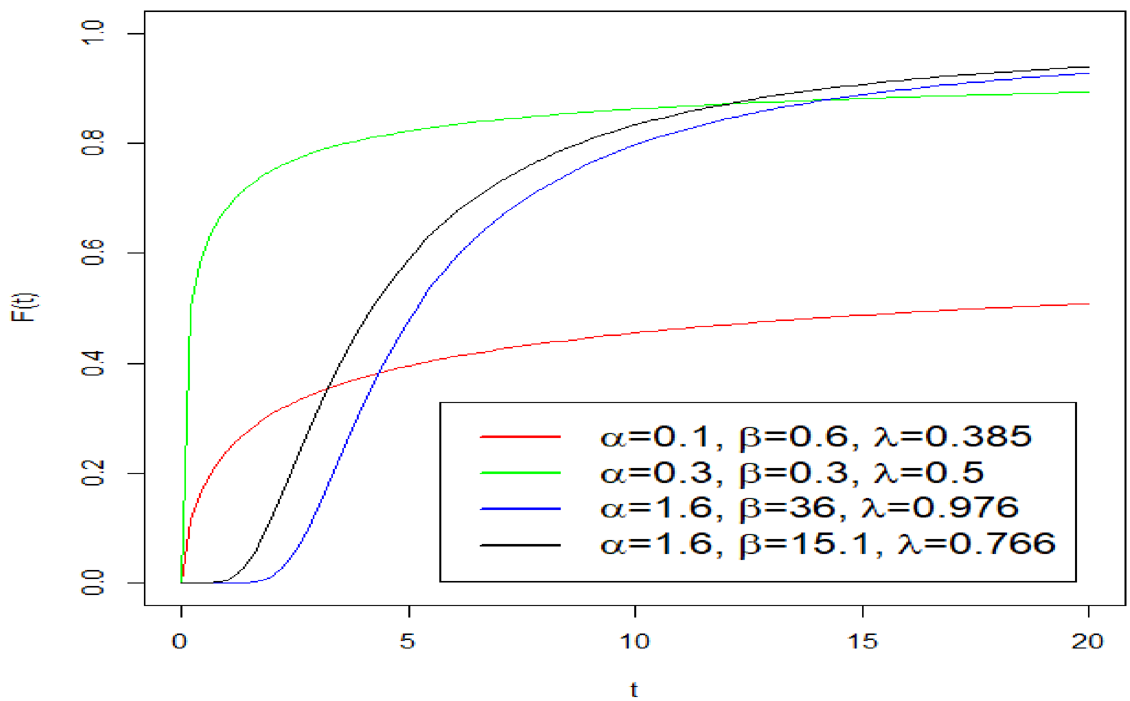

Figure 4.10: Plot of LIKum CDF

The associated LIKum PDF is

$$f(t) = \frac{\lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}}{\log(1-\lambda)\left[\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)-1\right]}, t > 0, \qquad (4.31)$$

Figure 4.11 exhibits the plot of the LIKum density function. It is visible that the PDF of the LIKum distribution shows right skewed upside-down bathtub and decreasing failure rates for certain chosen parameter values.

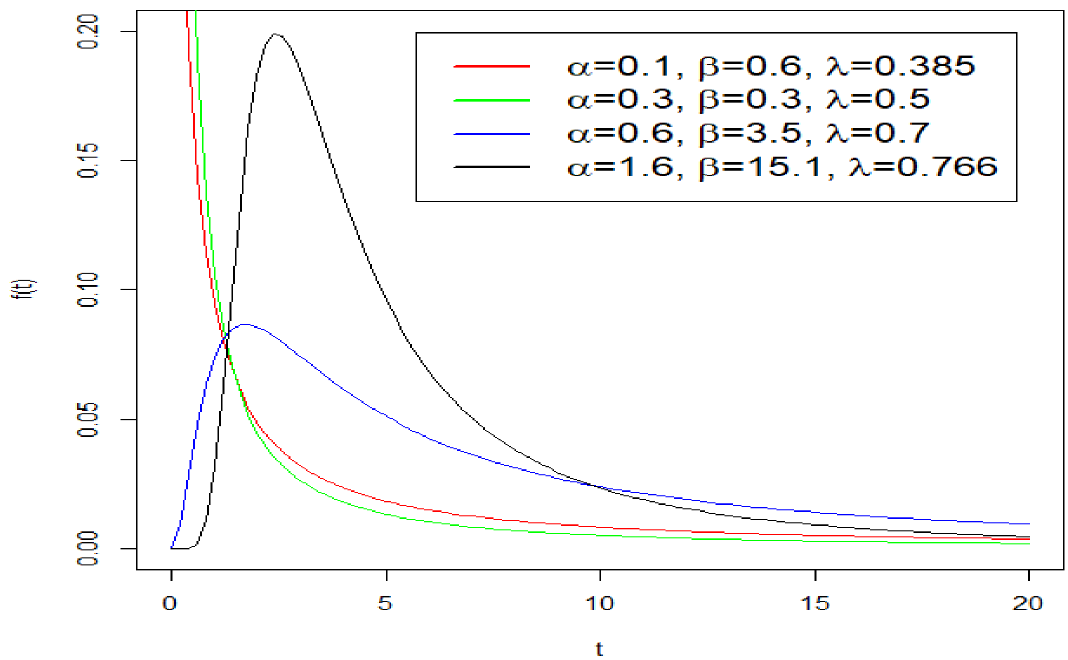

Figure 4.11: Plot of the LIKum PDF

The corresponding LIKum hazard function is

$$h(t)=\frac{\lambda\alpha\beta(1+t)^{-\alpha-1}(1-(1+t)^{-\alpha})^{\beta-1}}{\left[\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)-1\right]\log\left[1-\lambda\left(1-\left(1-(1+t)^{-\alpha}\right)^{\beta}\right)\right]},\ t>0. \quad (4.32)$$

The plot of LIKum hazard function is shown in Figure 4.12. The LIKum hazard function can exhibit decreasing, upside-down bathtub and right skewed failure rates as depicted in Figure 4.12.

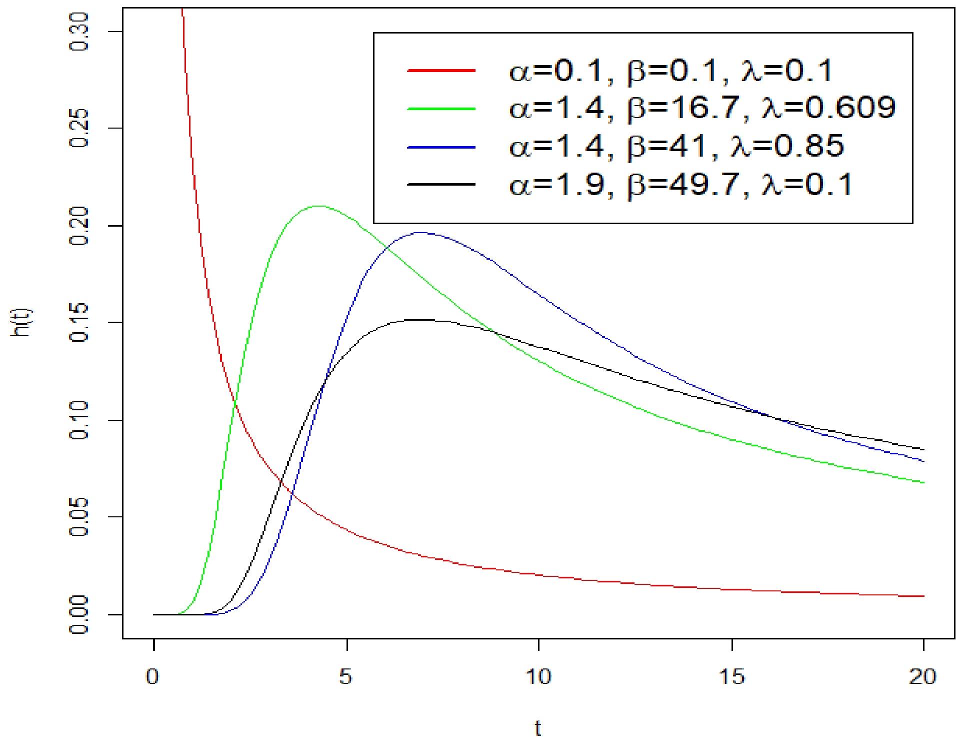

Figure 4.12: Plot of LIKum hazard function

4. 5 Monte Carlo Simulation

Monte Carlo simulation experiments were carried out to investigate the performance of the estimators of the parameters of the PSIKum model. For demonstration purposes, the PIKum distribution was employed for the experiment. The Monte Carlo simulation was replicated $1,000$ times for each sample size $n = 30, 60, 90, 150, 200, 300$ and 500. The following sets of parameter values $\text{I}: \alpha = 0.1, \beta = 0.8, \lambda = 3.1$ and $II: \alpha = 4.5, \beta = 3.8, \lambda = 8.1$ were used to derive random numbers from the

PIKum model. The average estimate (AE), average bias (AB), root mean square error (RMSE) and coverage probability (CP) for the estimators of the parameters are shown in Table 4.1. The AE as shown in Table 4.1 vary with respect to the sample sizes and approaches the actual value as the sample size increases. Looking at the absolute value of the ABs, it can be seen that the AB for the estimators decays towards zero as the sample size increases. Also, the RMSEs for the estimators generally decrease to zero as the sample size increases. This implies that the consistency property of the estimators can be achieved as $n \to \infty$. The 95% confidence interval CPs for the estimators are generally closer to the nominal value of 0.95. Thus, the estimators for the parameters estimate the parameters of the distribution well.

Table 4.1: Monte Carlo simulation results

		I				II			
Parameter	n	**AE**	**AB**	**RMSE**	**CP**	**AE**	**AB**	**RMSE**	**CP**
α	30	0.1546	0.0546	0.1008	0.982	6.6695	2.1695	5.2363	0.913
	60	0.1384	0.0384	0.0761	0.962	5.6402	1.1402	3.9844	0.909
	90	0.1325	0.0325	0.0699	0.917	5.1789	0.6789	3.4716	0.917
	150	0.123	0.023	0.0578	0.885	4.7494	0.2494	2.4403	0.94
	200	0.117	0.017	0.052	0.894	4.7177	0.2177	2.0413	0.953
	300	0.1109	0.0109	0.0434	0.879	4.5311	0.0311	1.435	0.952
	500	0.1058	0.0058	0.0365	0.87	4.5289	0.0289	0.8944	0.959
β	30	0.816	0.016	0.184	0.966	5.2899	1.4899	4.19	0.981
	60	0.7966	-0.0034	0.1198	0.959	4.5369	0.7369	2.1967	0.946
	90	0.789	-0.011	0.0937	0.972	4.2293	0.4293	1.5747	0.896
	150	0.7878	-0.0122	0.0753	0.957	3.9827	0.1827	1.0116	0.919
	200	0.7896	-0.0104	0.0653	0.955	3.9486	0.1486	0.8695	0.935
	300	0.7903	-0.0097	0.0509	0.949	3.8498	0.0498	0.5855	0.936
	500	0.7939	-0.0061	0.0382	0.959	3.84	0.04	0.3894	0.96
[illegible]	30	2.4831	-0.6169	2.279	0.986	24.8835	16.7835	64.6653	0.754
	60	2.5791	-0.5209	1.6547	0.943	18.9536	10.8536	43.5351	0.843
	90	2.748	-0.352	1.8798	0.887	18.2635	10.1635	40.2821	0.868
	150	2.8645	-0.2355	1.7003	0.879	13.662	5.562	25.0641	0.925
	200	3.0132	-0.0868	1.6121	0.892	10.6578	2.5578	11.3016	0.934
	300	3.1349	0.0349	1.6059	0.894	9.4379	1.3379	4.935	0.955
	500	3.1973	0.0973	1.3866	0.906	8.6515	0.5515	2.7021	0.961

4.6 Applications to Lifetime Data

The importance of the special distributions were displayed using two lifetime data sets. The performance of the special distributions was compared to each other by comparing their AIC, AICc and BIC values.

4.6.1 Failure Time of Repairable Objects Data

Table 4.2 presents the descriptive statistics of the failure time for repairable objects data. From Table 4.2, the minimum and maximum failure times of the repairable objects are given as 0.11 and 4.73 respectively. The mean failure time is given as 1.543 and the coefficient of skewness and excess kurtosis are given as 1.36 and 1.80 respectively. The last two values show that the failure time of the repairable objects is right skewed and more peaked than the normal curve.

Table 4.2: Descriptive statistics of failure time of repairable objects.

Minimum	Maximum	Mean	Skewness	Excess Kurtosis
0.110	4.730	1.543	1.36	1.80

The failure times were modeled using the PIKum, GIKum, BIKum and LIKum distributions. It is worth mentioning that $m = 5$ replicas were used to fit the BIKum distribution to the data set. Table 4.3 shows the maximum likelihood estimates for the parameters of the fitted distribution. For the PIKum and LIKum distribution, only the β parameter is significant at the 5% significance level. For the BIKum distribution, all three parameters are significant at the 5% significance level. The GIKum distribution also has only one parameter α being significant at 5% significance level

Table 4.3: Maximum likelihood estimates for the failure time of repairable objects data

Model	Estimate	standard error	z-value	P-value
PIKum	$\hat{\alpha} = 0.3401$	0.3112	1.0928	0.2745
	$\hat{\beta} = 2.5997$	0.5784	4.4949	6.96×10^{-6} *
	$\hat{\lambda} = 28.7383$	42.7294	0.6726	0.5012
GIKum	$\hat{\alpha} = 3.8016$	0.8158	4.6600	3.162×10^{-6} *
	$\hat{\beta} = 2.1907$	2.0339	1.0771	0.2814
	$\hat{\lambda} = -8.5820$	13.9504	–0.6152	0.5384
BIKum	$\hat{\alpha} = 0.8924$	0.1911	4.6688	3.030×10^{-6} *
	$\hat{\beta} = 2.9988$	0.6793	4.4145	1.012×10^{-5} *
	$\hat{\lambda} = 197.09$	1.8117×10^{-7}	1.0879	$< 2.2 \times 10^{-16}$ *
LIKum	$\hat{\alpha} = 4.9600$	2.7090	1.8309	0.0671
	$\hat{\beta} = 3.9362$	1.8706	2.1042	0.0354*
	$\hat{\lambda} = -175.6579$	827.6443	–0.2122	0.8319

* : means significant at 5% significance level.

Table 4.4 depicts the goodness-of-fit statistics for the fitted distributions. It can be seen from Table 4.4 that the GIKum distribution gives the best fit for the data since it has the highest log-likelihood value and the least values for the AIC, AICc, and BIC.

Table 4.4: Goodness-of-fit statistics for the failure time of repairable objects data

Model	log-likelihood	AIC	AICc	BIC
PIKum	−39.9000	85.8026	86.7256	90.0062
GIKum	−14.2900∗	34.5738∗	35.4969∗	38.7774∗
BIKum	−40.1500	86.2964	87.2195	90.4999
LIKum	−39.9900	85.9818	86.7818	90.1854

*: means best based on the goodness-of-fit-statistic

Figure 4.13 depicts the empirical CDF of the failure time of repairable objects and the fitted CDFs. It can be seen that the fitted distributions mimic the empirical CDF of the failure time of repairable objects data.

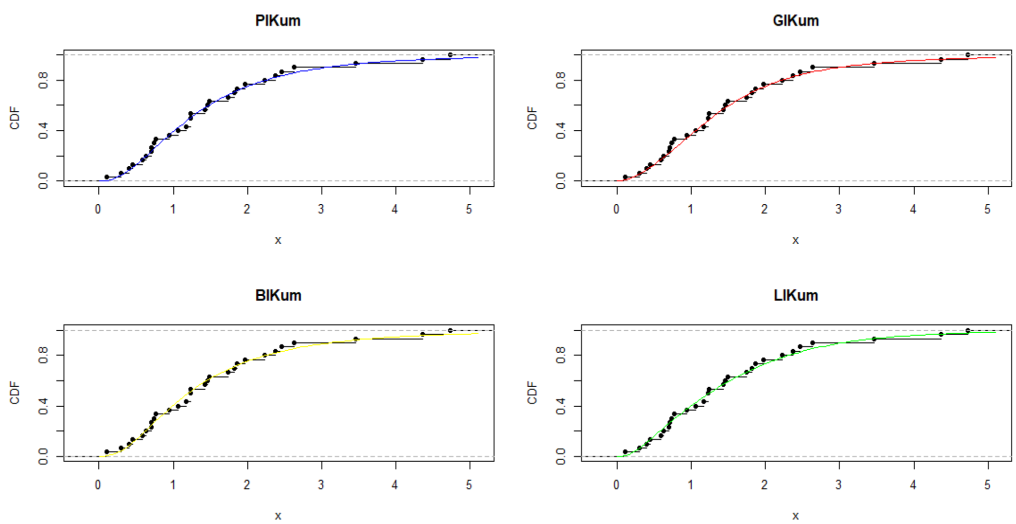

Figure 4.13: Plots of fitted CDFs for repairable objects data.

The probability-probability plots of the fitted distributions were plotted to investigate how well the distributions fit the given data set. From Figure 4.14, it can be clearly seen that all the special distributions gave good fit to the data set as the plot of their observed probability against the expected cluster along the diagonal. However, by comparing their AIC, AICc and BIC values, it is evident that the GIKum distribution provided the best fit.

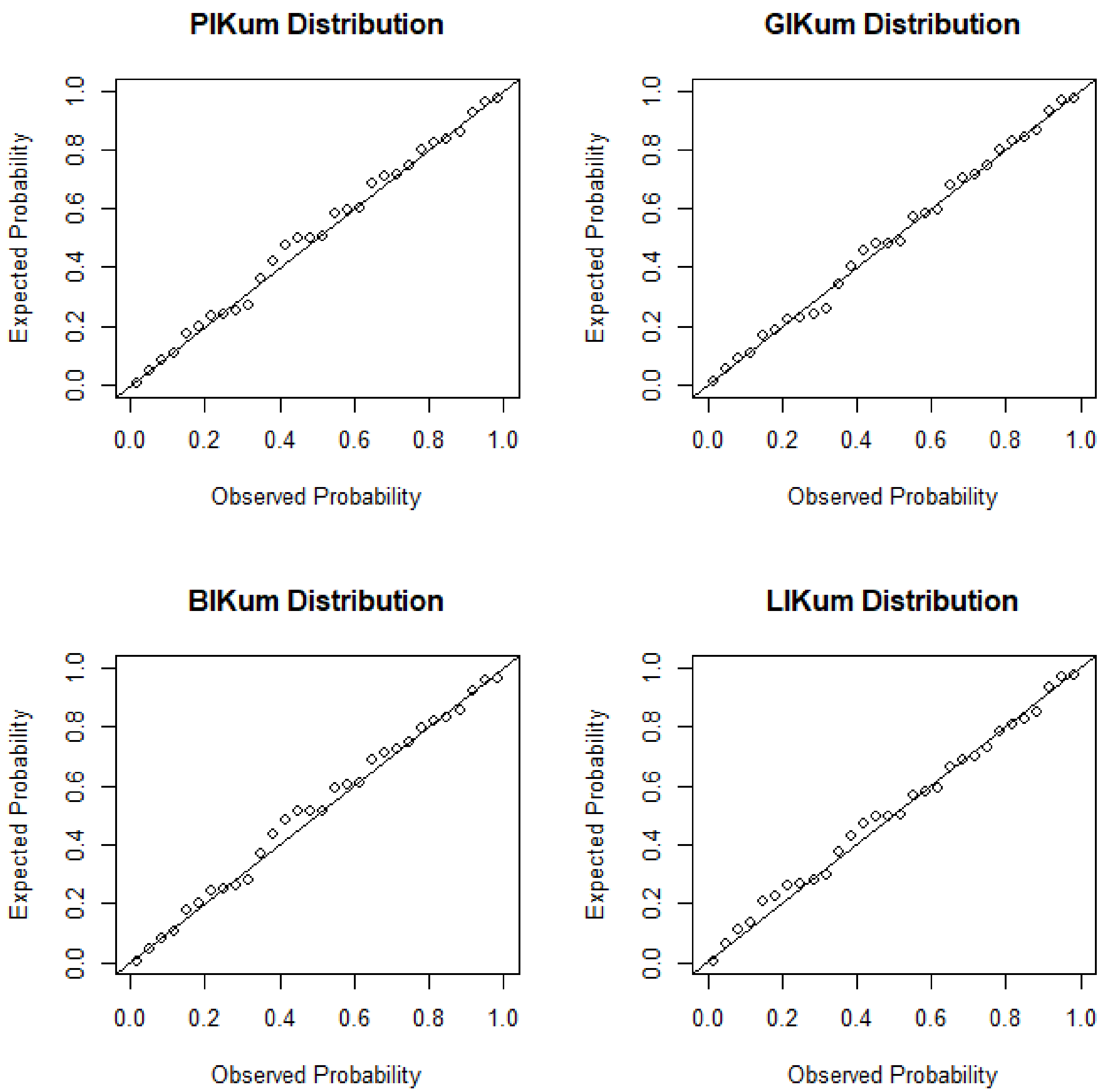

Figure 4.4: probability-probability plots of fitted distributions for repairable objects data

4.6.2 Vinyl chloride used for monitoring wells in mg/L data

The observed values in the data stand for vinyl chloride used for monitoring wells in mg/L. Table 4.5 shows the descriptive statistics for the data set. From Table 4.5, the minimum and maximum values in mg/L are given as 0.1 and 8 respectively. The mean value in mg/ L is 1.879. The skewness value of 1.68 shows that the data are right skewed. The excess kurtosis value of 2.53 shows that the distribution of the vinyl chloride data is more peaked than the normal curve and the observations are closely distributed around their average value.

Table 4.5 Descriptive statistics for vinyl chloride used for monitoring wells in mg/L data

Minimum	Maximum	Mean	Skewness	Excess Kurtosis
0.1000	8.0000	1.8790	1.6800	2.5300

The vinyl chloride data were modeled using the PIKum, GIKum, BIKum and LIKum distributions. It is worth noting that $m=5$ replicas were used to fit the BIKum distribution to the dataset. Table 4.6 shows the estimates for the parameters of the fitted distribution. For the PIKum distribution, all the parameters except λ are significant at the 5% significance level. For the remaining three distributions, all the parameters are significant at the 5% significance level.

Table 4.6: Maximum likelihood estimate for vinyl chloride in mg/L data

Model	Estimate	standard error	z-value	p-value
PIKum	$\hat{\alpha} = 4.3544 \times 10^{-2}$	1.9774×10^{-2}	2.2020	$0.0277*$
	$\hat{\beta} = 1.5954$	2.3238×10^{-1}	6.8654	$6.6320 \times 10^{-12} *$
	$\hat{\lambda} = 1.5931 \times 10^{2}$	3.0174×10^{-5}	5.2798×10^{6}	$< 2.2 \times 10^{-16} *$
GIKum	$\hat{\alpha} = 2.1196$	0.5716	3.7082	$2.087 \times 10^{-4} *$
	$\hat{\beta} = 1.7028$	0.8019	2.1235	$0.0337*$
	$\hat{\lambda} = -1.2472$	2.3560	−0.5294	0.5966
BIKum	$\hat{\alpha} = 0.5021$	1.3309	2.1118	$8.214 \times 10^{-5} *$
	$\hat{\beta} = 1.6485$	0.6412	2.5709	$1.036 \times 10^{-2} *$
	$\hat{\lambda} = -25.6923$	86.7895	−0.2960	0.8535
LIKum	$\hat{\alpha} = 2.8676$	1.4467	1.9822	$0.0475*$
	$\hat{\beta} = 1.6373$	0.6477	2.5280	$0.0115*$
	$\hat{\lambda} = -29.8536$	108.6257	−0.2748	0.7835

$*$: means significant at 5% significance level.

Table 4.7 depicts the goodness-of-fit statistics for the fitted distributions and this was used to compare their performance. It can be seen from Table 4.7 that the GIKum distribution provided the best fit for the data since it has the highest value of log-likelihood and the least values of AIC, AICc, and BIC.

Table 4.7 Goodness-of-fit statistics for vinyl chloride in mg/L data

Model	Log-likelihood	AIC	AICc	BIC
PIKum	−55.2800	116.5500	117.35	121.1291
GIKum	−25.5300∗	57.0506∗	57.8506∗	61.6297∗
BIKum	−55.4600	116.9216	117.7216	121.5006
LIKum	−55.1900	116.3774	117.1774	120.9565

*: means best based on the goodness-of-fit-statistic

Figure 4.15 depicts the empirical CDF of the vinyl chloride data and the fitted CDFs. It can be seen that the fitted distributions mimic the empirical CDF of the failure time of the vinyl chloride data.

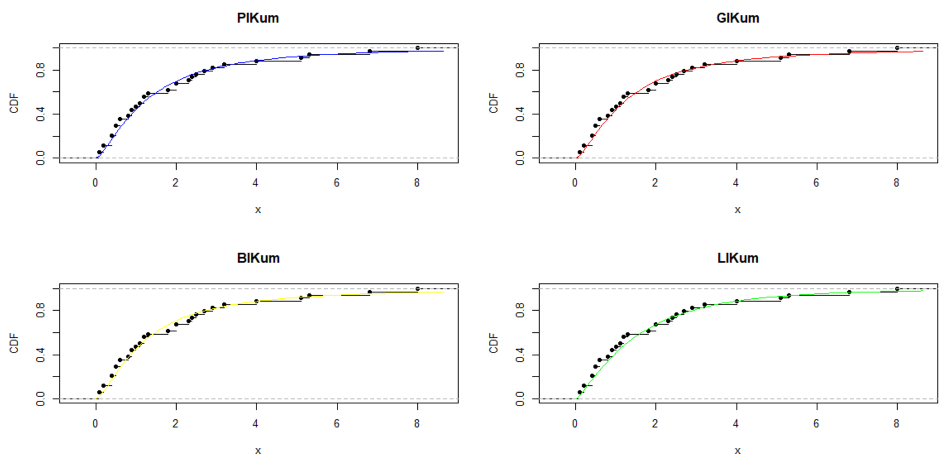

Figure 4.15: Plots of fitted CDFs for vinyl chloride data.

The probability-probability plots were used to determine how well the special distributions fit the data. From Figure 4.16, it is evident that all the distributions provided good fit to the data set and their difference cannot be easily distinguished. However, by comparing their AIC, AICc and BIC values, it is evident that the GIKum provided the best fit to the data set.

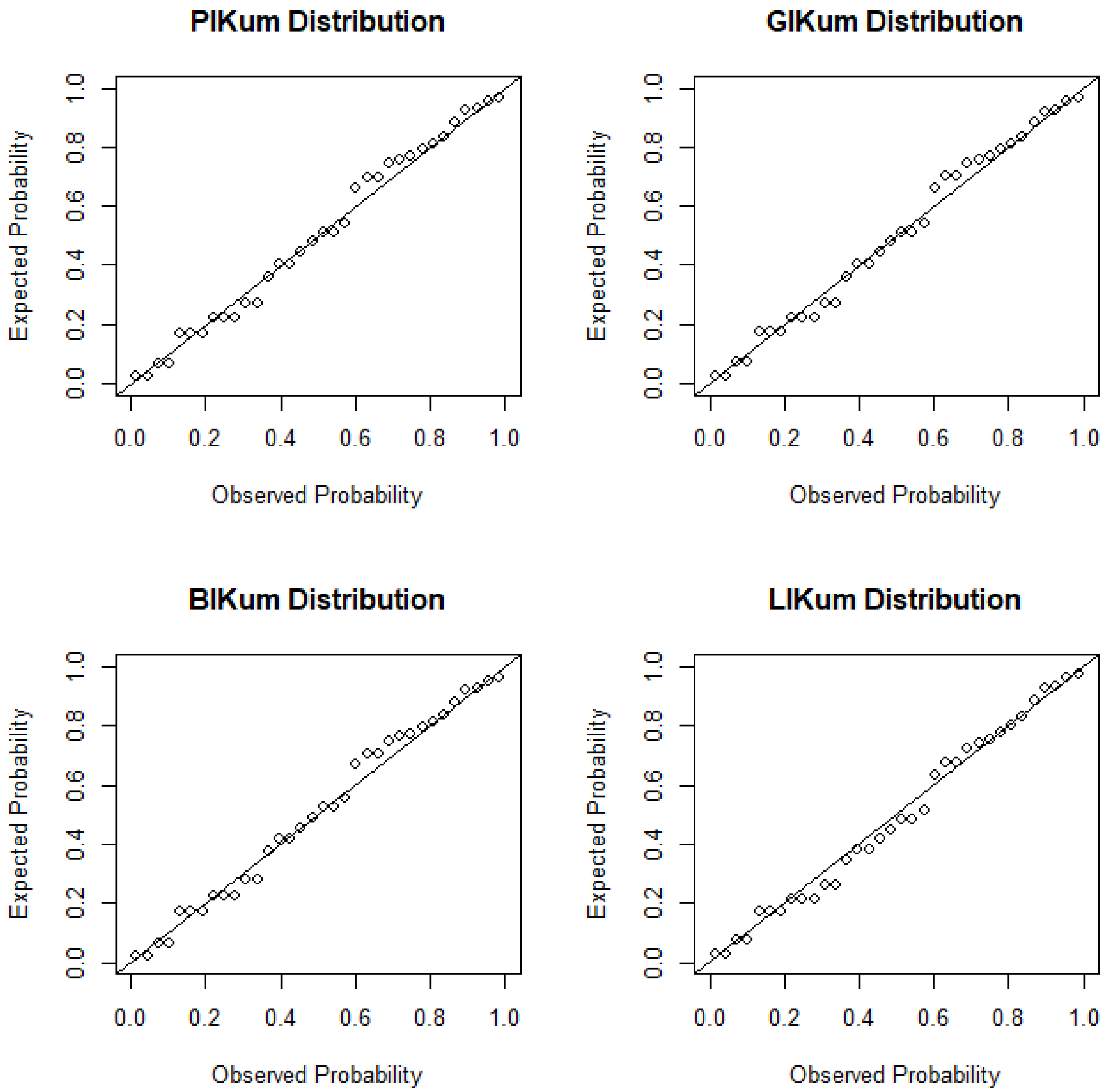

Figure 4.16: P-P plots of fitted distributions for vinyl chloride data

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.0 Introduction

The summary of the study, conclusions and recommendations are presented below.

5.1 Summary

Generalizing most well-known classical distributions and finding an alternative to them has become an area of importance to statisticians in academic research due to their inability to model correctly all the different forms of data sets that exist.

This thesis presented a current family of three-parameter distributions by name the power series inverted Kumaraswamy distribution which is a modification of the inverted Kumaraswamy distribution by incorporating an extra parameter thereby making it more flexible. The CDF, PDF, hazard and survival rate functions of the PSIKum model were derived. The PDF of the PSIKum was stated as an infinite mixture of the density of the smallest order statistic of the IKum distribution to facilitate the derivation of the statistical properties of the PSIKum distribution. Some statistical properties such as the moments, moment generating functions, quantiles, stochastic ordering property and order statistics of the PSIKum distribution were derived. Four special sub-distributions of the PSIKum were derived and these are the PIKum, GIKum, BIKum and LIKum distributions. For each of these sub distributions, the CDF, PDF, survival and hazard rate functions were derived. The sub distributions can take several shapes such as unimodal, bathtub, upside-down bathtub and right skewed shapes with varying degrees of kurtosis and this was exhibited through the plots of the PDF and hazard rate function. This shows that the new family of distributions can model both monotonic and non-monotonic failure rates.

The parameters of the new distribution were estimated by the use of maximum likelihood estimation method. The properties of the estimators developed for the PSIKum distribution were assessed by performing Monte Carlo simulation experiments and it was shown that the estimators are capable of estimating the parameters well.

To demonstrate the flexibility of the new models, two sets of lifetime data were used. The first data comprises of 30 values for the failure time of repairable objects employed by Murthy *et al.* (2004) and the second data first used by Bhaumik *et al.* (2009) is represented by vinyl chloride used for monitoring wells in mg/L. Descriptive statistics for both data sets shows that they are right skewed and more peaked than the normal distribution. For both data sets, the GIKum distribution was adjudged the best model because it had the least AIC, AICc and BIC values.

To test how well the new distributions fit the data sets, the empirical CDFs of the data sets and the fitted CDFs were plotted. It was discovered that all four distributions fit the given data sets very well. The P-P plots of the fitted distribution were also plotted to investigate how well the distributions fit the given data set and it was realized that they all provide good fit as the plot of their observed probability against the expected cluster along the diagonal.

5.2 Conclusions

This study developed the PSIKum distribution by compounding the zero truncated power series distribution with the IKum distribution. The mathematical and statistical properties of this new model were studied. The model's hazard function and PDF exhibits different flexible behaviors: increasing, decreasing, right skewed and upside-down bathtub. These characteristics make them suitable for modeling lifetime data sets which exhibit such failure rates.

Estimators for estimating the parameters were developed. Monte Carlo simulation studies were conducted to examine the stability of the estimates of the parameters in terms of the average biases and root mean square errors and it was revealed that the estimates were asymptotically consistent and unbiased.

Applications of the special distributions using two sets of lifetime data showed their flexibility and usefulness in modeling different sets of data. It was determined that for both data sets, the GIKum provided the best fit as compared to the other sub-distributions and this was done by comparing their AIC, AICc and BIC values.

5.3 Recommendations for Further Studies

i. This study focused on scenarios where the least number of events has to occur and so we used the stochastic representation $T_{(1)} = \min\left(T_1, T_2, ..., T_N\right)$ to develop the PSIKum distribution. However, one can also study the maximum case by considering the stochastic representation $T_{(n)} = \max\left(T_1, T_2, ..., T_N\right)$ to obtain a new class of lifetime distributions for the parallel system.

ii. In this thesis, all the data sets provided are complete sets of lifetime data. However, survival times of some individuals might be censored due to various reasons. Therefore, subsequent research can be extended to cater for censored survival data. Estimation of the parameters of the model can be carried out using the maximum likelihood method through the expectation-maximization algorithm and the Bayesian approach can also be considered.

REFERENCE

Abdelall, Y. Y. (2016). The Odd generalized exponential modified Weibull distribution. *International Mathematical Forum*, ***11****(19)*: 943-959.

Abdul-Moniem, I. B. (2015). Exponentiated Nadarajah and Haghighi's exponential distribution. *International Journal of Mathematical Analysis and Applications*, **2**(5): 68-73.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**(6): 716-723.

Al-Fattah, A. M., El-Helbawy, A. A., and Al-Dayian, G. R., (2017). Inverted Kumaraswamy Distribution: Properties and Application. *Pakistan Journal of Statistics*, **33**(1): 37-61.

Alizadeh, M., Yousof, H. M., Afify, A. Z., Cordeiro, G. M., and Mansoor, M. (2018). The complementary generalized transmuted Poisson-G family of distributions. *Austrian Journal of Statistics*, **47**(4): 51-71.

Alkarni, S. H. (2013). A class of truncated binomial lifetime distributions. *Open Journal of Statistics*, **3**(5): 305-311.

Alkarni, S. H. (2016). Generalized extended Weibull power series family of distribution. *Journal of Data Science*, **14**(3): 415-440.

Anderson, D. R. (2002). Model Selection and multi-model inference: a practical information- theoretic approach. *Springer*.

Aryal, G. and Elbatal, I. (2015). Kumaraswamy modified inverse weibull distribution: theory and application. *Applied Mathematics and information sciences*, **9**(2): 651-660.

Atem, B. A. M. (2018). On the odd Kumaraswamy inverse Weibull distribution with application to survival data. *Unpublished M.Sc. Thesis*, Pan African University, Institute for Basic Sciences, Technology and Innovation.

Bakouch, S. H., Ristić, M. M., Asgharzadeh, A., Esmaily, L., Al-Zahrani, B. M. (2012). An exponentiated exponential binomial distribution with application. *Statistics and Probability letters*, **82**(1):1067-1081.

Behairy, S. M., Al-Dayian, G. R., and El-Helbawy, A. A. (2016). The Kumaraswamy-Burr type III distribution: properties and estimation. *British Journal of Mathematics and Computer Science*, **14**(2): 1-21.

Bhaumik, D. K., Kapur, K., Gibbons, R. D. (2009). Testing parameters of a gamma distribution for small samples. *Technometrics* **51**(3): 326-334.

Bidram, H., and Nekoukhou, V. (2013). Double bounded Kumaraswamy-power series class of distributions. *Statistics and Operations Research Transactions*, **37**(2): 211-230.

Chahkandi, M., and Ganjali, M. (2009). On some lifetime distributions with decreasing failure rate. *Computational Statistics and Data Analysis*, **53**(12): 4433-4440.

Chakrabarty, J. B., and chowdury, S. (2016). Compounded inverse Weibull distributions: properties, inference and applications. *IIMK working paper*

Cho, Y. S., Kang, S. B., and Han, J. T. (2009). The exponentiated extreme value distribution. *Journal of the Korean Data and Information Science Society*, **20**(4): 719-731.

Chung, Y., and Kang, Y. (2014). The exponentiated Weibull-geometric distribution: properties and estimations. *Communications for Statistical Applications and Methods*, **21**(2):147-160.

Cordeiro, G. M., and 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. M. (2013). The beta exponentiated Weibull distribution. *Journal of Statistical Computation and Simulation*, **83**(1): 114-138.

Dey, S., Nassar, M., and Kumar, D. (2017). Alpha logarithmic transformed family of distributions with application. *Annals of Data Science*, DOI 10.1007/s40745-017-0115-2.

Elbatal, I., Louzada, F., and Granzotto, D. C. T. (2018). The Kumaraswamy extension exponential distribution. *Open Access (Biostatistics and Bioinformatics)*, **2**(1): 1-9.

Elbatal, I., Zayed, M., Rasekhi, M., and Butt, N. S. (2017). The exponential Pareto power series distribution: Theory and Applications. *Pakistan Journal of Statistics and Operation Research*, **13**(3): 603-615.

El-Deen, M. S., Al-Dayian, G. R, and El-Helbawy, A. (2014). Statistical inference for Kumaraswamy distribution based on generalized order statistics with applications. *British Journal of Computer Science*, **4**(12): 1710-1743.

Elgarhy, M., Haq, M. A. and Ain, Q. (2018). Exponentiated generalized Kumaraswamy distribution with applications. *Annals of Data Science*, **5**(2): 273-292.

Fattah, A. A., Nadarajah, S., and Ahmed, A. H. N. (2017). The exponentiated transmuted Weibull geometric distribution with application in survival analysis. *Communications in Statistics-Simulation and Computation*, **46**(6): 1-20

Golizadeh, A., Sherazi, M. A., and Moslamanzadeh, S. (2011). Classical and Bayesian estimation on Kumaraswamy distribution using grouped and ungrouped data under difference of loss functions. *Journal of Applied Sciences*, **11**(12): 2154-2162.

Gui, W., Zhang, S., and Lu, X. (2014). The Lindley-Poisson distribution in lifetime analysis and its properties. *Journal of Mathematics and Statistics*, **43**(6): 1063-1077.

Gupta, R. D., and Kundu, D. (1999). Generalized exponential distribution. *Australian and New Zealand Journal of Statistics*, **41**(2): 173-188.

Hurvich, C. M., and Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika*, **76**(2): 297-307.

Iqbal, Z., Tahir, M. M., Riaz, N., Ali, S. A., and Ahmad, M. (2017). Generalized inverted Kumaraswamy distribution: properties and applications. *Open Journal of Statistics*, **7**(4): 645-662.

Jafari, A. A., and Tahmasebi, S. (2015). Gompertz-power series distributions. *Communications in Statistics-Theory and Methods*, **45**(13): 3761-3781.

Jamal, F., Elgarhy, M., Nasir, Ozel, G., and Khan, N. M. (2018). Generalized inverted Kumaraswamy generated family of distributions: theory and applications. *hal-01907258 https:// hal.archives-ouvertes.fr/hal-01907258.*

Jones, M. C. (2009). Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Journal of statistical methodology*, **6**(1): 70-81.

Khatri, C. G. (1959). On the mutual independence of certain statistics. *The Annals of Mathematical Statistics*, **30**(4): 1258-1262.

Kumaraswamy, P. (1980). A generalized probability density function for double-bounded random processes. *Journal of Hydrology*, **46**(1): 79-88.

Mahmoudi, E., and Jafari, A. A. (2012). Generalized exponential power series distributions. *Computational Statistics and Data Analysis*, **56**(12): 4047-4066.

Mamoudi, E., and Shiran, M. (2012). Exponentiated Weibull-geometric distribution and its applications. arXiv:1206.4008vi [stat.ME].

Mohammed, H. F. (2017). Inference on the log-exponentiated Kumaraswamy distribution. *International Journal of Contemporary Mathematical Sciences*, **12**(4): 165-179.

Morais, A. L., and Barreto-Souza, W. (2011). A compound family of Weibull and power series distributions. *Computational Statistics and Data Analysis*, **55**(3): 1410-1425.

Muhammad, M., and Yahaya, M. A. (2017). The half logistic Poisson distribution. *Asian Journal of Mathematics and Applications*, **2017**: 1-15.

Murthy, D. P., Xie, M., Jian, R. (2004). Weibull models, vol. 505. *John Wiley and sons.*

Nadarajah, S. and Eljabri, S. (2013). The Kumaraswamy GP distribution. *Journal of data science*, 11: 739-766.

Nadarajah, S., Cordeiro, G. M., and Ortega, E. (2014). The exponentiated G geometric family of distributions. *Journal of Statistical Computation and simulation*, **85**(8): 1-17.

Nasiru, S., Atem, B. A. M., and Nantomah, K. (2018). Poisson exponentiated erlang-truncated exponential distribution. *Journal of Statistics Applications and Probability*, **7**(2): 1-17.

Nasiru, S., Mwita, P. N., Ngesa, O. (2018). Exponentiated generalized power series family of distributions. *Annals of Data Science*. https://doi.org/10.1007/s40745-018-0170-3.

Nekoukhou, V. and Bidram, H. (2015). The exponentiated discrete Weibull distribution. *SORT* **39**(1): 127-146.

Noack, A. (1950). A class of random variables with discrete distributions. *The Annals of Mathematical Statistics*, **21**(1): 127-132.

Oguntunde, P. E., Ilori, K. A., and Okagbue, H. I. (2018). The inverted weighted exponential distribution with applications. *International Journal of Advanced and Applied Sciences*, **5**(11): 46-50.

Ohnishi, M. (2002). Stochastic orders in reliability theory. In: Osaki S. (eds) Stochastic Models in Reliability and Maintenance. Springer, Berlin, Heidelberg.

Pascoa, M. A. R., and Cordeiro, G. M. (2011). The Kumaraswamy generalized distribution with application in survival analysis. *Statistical Methodology*, **8**(5): 411-413.

Patil, G. P. (1962). Certain properties of the generalized power series distribution. *Annals of the Institute of Statistical Mathematics*, **14**(1): 179-182.

Rodrigues, J. A., Silva, A. P. C. M., Hamedani, G. G. (2016). The exponentiated Kumaraswamy inverse Weibull distribution with application in survival analysis. *Journal of Statistical Theory and Applications*, **15**(1): 8-24.

Saboor, A., Elbatal, I., and Cordeiro, G. M. (2016). The transmuted exponential Weibull geometric distribution: theory and applications. *Hacettepe Journal of Mathematics and statistics,* **45**(3):973-987.

Sankaran, P. G., and Anjana, S. (2015). Parametric analysis of lifetime data with multiple causes of failure using cause specific reversed hazard rates. *Calcutta Statistical Association Bulletin*, **67**(267-268): 129-142.

Sarhan, A., Abd El-Baset, A. A., and Ibtesam, A. A. (2013). Exponentiated generalized linear exponential distribution. *Applied Mathematical modelling*, **37**(5): 2838-2849.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, **6**(2): 461-464.

Shafiei, S., Darijani, S., and Saboori, H. (2015). Inverse Weibull power series distributions: properties and applications, *Journal of statistical computation and simulation*, **86**(6): 1-26.

Silva, R. B., Bourguignon, M., and Cordeiro, G. M. (2016). A new compounding family of distributions: The generalized gamma power series distributions. *Journal of Computational and Applied Mathematics*, **303**: 119-139.

Silva, R. B., Bourguignon, M., Dias, C. R. B., and Cordeiro, G. M. (2013). The compound family of extended Weibull power series distributions. *Computational Statistics and Data Analysis*, **58**: 352-367.

Sindhu, T. N, Feroze, N., and Aslam, M. (2013). Bayesian analysis of the Kumaraswamy distribution under failure censoring sampling scheme. *International Journal of Advanced Science and Technology*, **51**: 39-58.

Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M., Zubair, M., and Hamedani, G.G. (2015). The odd generalized exponential family of distributions with applications. *Journal of Statistical Distributions and Applications*, **2**(1): 1-28.

Tahmsebi, S., Jafari, A. A., and Gholizadeh, B. (2015). The exponentiated G family of power series distributions. 46th Annual Iranian Mathematics Conference (AIMC 46), At Yazd University, volume: **1**.

Tomy, L., and Gillariose, J. (2018). The Marshall-Olkin IKum Distribution. *Biometrics and Biostatistics International Journal*, **7**(1): 10-15.

Unal, C., and Cakmakyapan, S. Ozel, G. (2018). Alpha power inverted exponential distribution: properties and application. *Gazi University Journal of Science*, **31**(3): 954-965.

Usman, R. M., Ul Haq, M. A., and Talib, J. (2017). Kumaraswamy half-logistic distribution: properties and applications. *Journal of statistics applications and probability*, **6**(3), 597-609.

Warahena-Liyanage, G., and Pararai, M. (2015). The Lindley power series class of distributions: model, properties and applications. *Journal of computations and modeling*, **5**(3): 35-80.