

DEVELOPMENT OF EXPONENTIATED GENERALIZED EXPONENTIATED EXPONENTIAL DISTRIBUTION

By

BUKOYE, ABDULWASIU

(13/68EP009)

M.Sc (Unilorin) 2015, B.Sc (Wudil) 2007

**A Thesis Submitted to the Department of Statistics, Faculty of
Physical Science in Partial Fulfillment of the Requirements for
the Award of the Degree of Doctor of Philosophy in Statistics**

DECEMBER, 2019

CERTIFICATION

This is to certify that this thesis has been read and approved as meeting the requirements of the Department of Statistics, University of Ilorin, Ilorin, Nigeria for the Award of Doctor of Philosophy in Statistics degree.

Prof. G. M. Oyeyemi
(Supervisor)

Date

Dr. O. Job
P.G. Coordinator

Date

Prof. A. O. Adejumo
(Head of Department)

Date

(External Examiner)

Date

DEDICATION

This research work is dedicated to my parents Late Alhaji and Alhaja Bukoye, my wife Mrs. Hafstat Damilola Bukoye and my children Abdulmateen, Mohammed and Mahdiya for their sacrifices and most importantly for their continuous believes in my dreams.

DECLARATION

I, BUKOYE ABDULWASIU with matriculation no 13/68EP009 hereby declares that this thesis entitled ON DEVELOPMENT OF FOUR-PARAMETERS EXPONENTIATED GENERALIZED EXPONENTIAL DISTRIBUTION is a record of my research work. It has neither been presented nor accepted in any previous application for a higher degree. All sources of information have been specifically acknowledged.

In addition, the research work has been ethically approved by the University Ethical Review Committee.

Bukoye, AbdulwasIU

Date

ACKNOWLEDGEMENTS

All forms praise belong to Almighty Allah, the lord of mankind. The Beneficient, the Merciful. HE taught man what he knew not. HE made everything possible.

My profound appreciation goes to my supervisor, Prof. G. M. Oyeyemi for his patient and encouragement, continued sacrifice of time and useful suggestion. I also want to acknowledge the Head of department Prof. O. Adejumo , P.G. coordinator Dr. O. Job, Prof. B. Oyejola, Prof. W. B Yahaya, Dr M.K. Garba, Prof A.A. Adewara, Prof R. Ipiyomi, Prof. E.T Jolayemi, Prof. L. Adeleke , Dr A. S. Idowu and the academic staff of Department of Statistics for their positive contribution in making these programme a success. May God reward you accordingly, s

My sincere appreciation goes to my beloved wife Mrs. Hafsat Damilola Bukoye for her support. Thanks for your mutual understanding and your spirit of “all for one and one for all” which has been pivotal not only in making sure that this work see the light of the day but also in all my endeavours of life and to the Bukoye’s , Abdulazeez’s, Yusuf’s, Bello’s, Elbaf Bashir and Afolabi’s families. Thank you all for your support.

Special regards to Mr T. Adeyemi, Mr O.Bada, Mr J. Afolabi, Mr O. Shaib Mr S. Akinyemi and all the staff of the Department of Statistics, Auchi Polytechnic Auchi as well as the entire Ph.D. students. I’ll cap this acknowledgement by appreciating everyone who has contributed in one capacity or the other towards the success of this work and the programme in general. You all hold a special spot in my heart.

ABSTRACT

Probability distributions have been used in describing and predicting real life phenomena. The inability of existing distributions to fit diverse nature of the data sets has necessitated the development of more flexible distributions for real life data set. Recent studies on developing new families of distributions from existing ones have helped in describing data better. The inability of exponential distributions to properly fit skewed data has been a serious concern. This study therefore aimed at developing a new tractable distribution that will fit data that are positively or negatively skewed. The objectives of the study were to: (i) derive a new distribution and its statistical properties. (ii) examine the effect of the shape and scale parameters on the distribution; and (iii) compare the performance of the proposed distribution with the existing ones using simulated and real-life data sets.

The generator employed in developing the proposed probability distribution function is the Exponentiated Generalized Family of distribution using the Exponentiated Exponential as the baseline distribution. The properties of the distribution such as mean, median and their deviations, quantile, Renyon entropy, skewness and kurtosis were obtained using analytical and simulation method. Simulated and real life data were used to assess and compare the performances of the proposed distribution with the existing ones using Akaike information criterion (AIC), Bayesian information criterion (BIC), log-likelihood function and likelihood ratio test.

The findings of the study were that:

- (i) the probability density function (pdf) and cumulative distribution functions(cdf) of the Exponentiated Generalized Exponentiated Exponential (EGEE) were derived together with their properties such as moments, quantiles, skewness, kurtosis, etc.;
- (ii) Exponential (E), Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions were obtain by setting the shape parameters value to one;
- (iii) the mean, standard deviation and median are increasing functions of the scale and shape parameters of the proposed distribution while the skewness and kurtosis were unaffected by scale parameter but only by the shape parameters; and
- (iv) the proposed EGEE competed very well with other existing distributions for small sample sizes and even performed better as the sample size increases.

The study therefore concluded that EGEE distribution can serve a better alternative distribution to the existing distributions (E, EE and EGE) in fitting skewed data. Therefore, the proposed EGEE distribution is preferred and recommended for fitting a skewed data.

TABLE OF CONTENTS

TITLE	i
CERTIFICATION	ii
DEDICATION	iii
DECLARATION	iv
ACKNOWLEDGEMENTS	v
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS AND NOTATIONS	ix-x
ABSTRACT	xi-xii

CHAPTER ONE: INTRODUCTION

1.1	Background to the Study	1-5
1.2	Statement of the Problem	5 - 6
1.3	Justification for the Study	6 - 7
1.4	Aim/Objectives of the Study	7
1.5	Significance of the Study	

CHAPTER TWO: REVIEW OF LITERATURE

2.1	Introduction	8
2.2	Statistical Distribution	8 -23
2.2	Appraisal of the Literature	23

CHAPTER THREE: METHODOLOGY / MATERIALS AND METHOD

3.1	Introduction	24
3.1	Theoretical Framework	24-25
3.2	Existing Distribution from Exponentiated Generalized Family	25
3.2.1	Exponentiated Generalized Frechet	25-26
3.2.2	Exponentiated Generalized Inverted Exponential	27-28
3.3	Proposed Distribution: Exponentiated Generalized Exponentiated Exponential (EGEE)	28-29
3.3.1	Proof of Proposed Distribution	29-30
3.3.1.1	Sub- Distributions of the Proposed Distribution	30-32
3.3.2	The Hazard Function	32
3.3.3	The Survival Function	32
3.3.4	The Reversed Hazard Function	33
3.3.5	The Odds Function of EGEE	33
3.3.6	Special Case of EGEE Distribution	33-36
3.4	Properties of Exponentiated Generalized Exponentiated Exponential (EGEE)	36-38
3.4.1	Quantile and Median of proposed EGEE Distribution	38-39
3.4.2	Moment of Proposed EGEE Distribution	39-42
3.4.3	Moment Generating Function of EGEE Distribution	42-43
3.4.4.1	Order statistics of Exponentiated Generalized Family (EGF) distribution	44

3.4.4.2	Order Statistics of EGEE	45-46
3.4.5	Skewness and Kurtosis of The EGEE Distribution	46-48
3.4.6	The Mean and Median Deviation of EGF	48
3.4.6.1	Mean Deviation of EGEE Distribution	48-49
3.4.7	Asymptotic Behavior of EGEE	50
3.4.8	Renyi Entropy of EGEE	50-51
3.4.9	Method of Maximum Likelihood	52-53
3.4.9.1	Maximum Likelihood of EGEE	54-56
3.5	Measure Statistics	56
3.5.1	Akaike Information Criterion (AIC)	56
3.5.2	Bayesian Information Criterion (BIC)	56
3.5.3	Likelihood Ratio Test (LRT)	57

CHAPTER FOUR: DATA ANALYSIS / RESULTS

4.1	Introduction	58
4.2.	Simulation Studies	58-63
4.2.1	Simulated Exponentiated Exponential for n=10	63-64
4.2.2	Simulated Exponentiated Exponential for n=50	65-66
4.2.3	Simulated Exponentiated Exponential for n=100	66-68
4.2.4	Simulated Exponentiated Exponential for n=200	68-70
4.2.5	Simulated Exponentiated Exponential for n=350	70-72
4.2.6	Simulated Exponential for n=10	73-74
4.2.7	Simulated Exponential for n=50	74-76
4.2.8	Simulated Exponential for n=100	76-78
4.2.9	Simulated Exponential for n=200	78-80

4.2.10 Simulated Exponential for $n=350$	80-82
4.2.11 Simulated Exponentiated Generalized Exponential for $n=10$	83-85
4.2.12 Simulated Exponentiated Generalized Exponential for $n=50$	85-87
4.2.13 Simulated Exponentiated Generalized Exponential for $n=100$	87-89
4.2.14 Simulated Exponentiated Generalized Exponential for $n=200$	89-91
4.2.15 Simulated Exponentiated Generalized Exponential for $n=350$	91-94
4.2.16 Simulated Beta distribution for $n=20$	95
4.2.17 Simulated Beta distribution for $n=50$	95-96
4.2.18 Simulated Beta distribution for $n=100$	96-97
4.3 Application	98
4.3.1 First Data set	98-101
4.3.2 Second Data Set	102-105
4.3.3: Third Data Set	105-108
3.3.4: Fourth Data Set	109-112
CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS	
5.1 Introduction	113
5.2 Summary	113- 118
5.3 Conclusion	118-120
5.4 Contribution to Knowledge	120-121
5.5 Recommendation	121
5.6 Suggestions for Further Studies	121
REFERENCES	122-129
APPENDIX	130-170

LIST OF TABLES

Table 3.1	Summary of EGEE and Sub-Distributions	36
Table 4.1	The Mean, Standard Deviation and Median of EGEE Distribution for $\Theta = 2, 4$ and 6 for $n=10$	61
Table 4.2	The Skewness and Kurtosis of EGEE Distribution for $\Theta = 2, 4$ and 6 for $n=10$	62
Table 4.3	The Mean, Standard Deviation and Median of EGEE Distribution for $\Theta = 2, 4$ and 6 for $n=100$	63
Table 4.4	The Skewness and Kurtosis of EGEE Distribution for $\Theta = 2, 4$ and 6 for $n=10$	64
Table 4.5	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 10	65
Table 4.6	Likelihood Ratio Test (LRT) Statistic for sample size 10	66
Table 4.7	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 50	67
Table 4.8	Table 4.8: Likelihood Ratio Test (LRT) Statistic for Sample size $n=50$	67
Table 4.9	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 100	68
Table 4.10	Likelihood Ratio Test (LRT) Statistic for Sample Size $n=100$	69
Table 4.11	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 200	70
Table 4.12	Likelihood Ratio Test (LRT) Statistic for sample size $n=200$	71
Table 4.13	Maximum Likelihood Estimate of Parameters, Standard Errors in Parenthesis,	

	Log-likelihood and Information Criteria of sample size 350	72
Table 4.14	Table 4.14: Likelihood Ratio Test (LRT) Statistic for sample size 350	73
Table 4.15	Summary Table on the Performance of Fitted Distributions to Simulated Data from Exponentiated Exponential Distribution	74
Table 4.16	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 10	75
Table 4.17	Likelihood Ratio Test (LRT) Statistic for sample size 10	75
Table 4.18	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 50	77
Table 4.19	Likelihood Ratio Test (LRT) Statistic for sample size 50	77
Table 4.20	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 100	79
Table 4.21	Likelihood Ratio Test (LRT) Statistic for sample size 100	79
Table 4.22	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 200	81
Table 4.23	Likelihood Ratio Test (LRT) Statistic for sample size 200	81
Table 4.24	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 350	83
Table 4.25	Likelihood Ratio Test (LRT) Statistic for sample size 350	83
Table 4.26	Summary Table on the Performance of Fitted Distributions to Simulate data from Exponential Distribution	85
Table 4.27	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 10	86
Table 4.28	Likelihood Ratio Test (LRT) Statistic for sample size 10	86
Table 4.29	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis,	

	Log likelihood and Information Criteria of sample size 50	88
Table 4.30	Likelihood Ratio Test (LRT) Statistic for sample size 50	88
Table 4.31:	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 100	90
Table 4.32	Likelihood Ratio Test (LRT) Statistic for sample size 100	90
Table 4.33	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 200	92
Table 4.34	Likelihood Ratio Test (LRT) Statistic for sample size 200	92
Table 4.35	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 350	94
Table 4.36	Likelihood Ratio Test (LRT) Statistic for sample size 350	94
Table 4.37	Summary Table on the Performance of Fitted Distributions to Simulated data from Exponentiated Generalized Exponential Distribution	94
Table 4.38:	Estimated log-likelihood values of simulated data from beta distribution at sample size $n=20$	95
Table 4.39:	Estimated log-likelihood values of simulated data from beta distribution at sample size $n=50$	96
Table 4.40:	Estimated log-likelihood values of simulated data from beta distribution at Sample size $n=100$	97
Table 4.41	Data on 100 Observations for Breaking Stress of Carbon Fibers (in Gba)	98
Table 4.42	Summary Statistics Carbon Fibers Data	99
Table 4.43	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criterion carbon Fibers	100

Table 4.44	Likelihood Ratio Test (LRT) Statistic for Carbon Fibers	101
Table 4.45	The Data Set Represent the Total Milk Production in the First Birth of 107 Cows from SINDI Race	102
Table 4.46	Summary Statistics Cow Data	103
Table 4.47	Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criterion of Cow Data	103
Table 4.48	Likelihood Ratio Test (LRT) Statistic for Cow Data	104
Table 4.49	The data relating to the strengths of 1.5cm glass fibres which was obtained by workers at UK National Physical laboratory	105
Table 4.50	Summary Statistics Glass Fibre Data	106
Table 4.51	Maximum Likelihood Estimate of Parameters ,Standard Errors in Parenthesis, Log likelihood and Information Criterion Glass fibres data	107
Table 4.52	Likelihood Ratio Test (LRT) Statistic for Glass Fibres	108
Table 4.53	Data set of 30 observations of March precipitation (in inches) Minneapolis/St.Paul	109
Table 4.54	Summary Statistics for precipitation in Minneapolis/St Paul	109
Table 4.55	Maximum Likelihood Estimate of Parameters, Standard Errors in Parenthesis, Log likelihood and Information Criterion on Precipitation data	110
Table 4.56	Likelihood Ratio Test (LRT) Statistic for Precipitation	111

LIST OF FIGURES

Figure 3.1	Plot of Cumulative Distribution Function for EGFT	27
Figure 3.2	Plot of Probability Density Function for EGFT	28
Table 3.3	Plot of Hazard Function for EGFT	28
Table 3.4	Plot of Cumulative Distribution Function for EGIE	29
Table 3.5	Plot of Probability Density Function for EGIE	30
Table 3.6	Plot of Hazard function for EGIE	30
Table 3.7	Probability of Density Function of EGEE, EE, EGE and E	36
Table 3.8	Distribution Function of EGEE, EE, EGE and E	36
Table 3.9	Hazard Function of EGEE, EE, EGE and E	37
Table 3.10	Survival Function of EGEE, EE, EGE and E	37
Table 3.11	Reversed Hazard Function of EGEE, EE, EGE and E	37
Table 3.12	Odds Function of EGEE, EE, EGE and E	38
Table 3.13	Plot for EGEE Skewness	50
Table 3.14	Plot for EGEE Kurtosis	50
Figure 4.1	Fitted Plot of EGEE, EGE, EE and E Models for sample size n=10	66
Figure 4.2	Fitted Plot of EGEE, EGE, EE and E Models for sample size n=50	68
Figure 4.3	Fitted Plot of EGEE, EGE, EE and E Models for sample size n=100	69

Figure 4.4	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=200$	71
Figure 4.5	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=350$	73
Figure 4.6	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=10$	76
Figure 4.7	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=50$	78
Figure 4.8	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=100$	80
Figure 4.9	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=200$	82
Figure 4.10	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=350$	84
Figure 4.11	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=50$	87
Figure 4.12	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=50$	89
Figure 4.13	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=100$	91
Figure 4.14	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=200$	93
Figure 4.15:	Fitted Plot of EGEE EGE, EE and E Models for sample size $n=350$	95
Table 4.16	Histogram and CDF Plots of an Empirical Distribution for Fibers Data	98
Table 4.17	Fitted Plot of EGEE EGE, EE and E Models on Fibers Data	100
Table 4.18	Histogram and CDF Plots of an Empirical Distribution for Cows Data	101
Table 4.19	Fitted Plot of EGEE EGE, EE and E Models on Cow Data	104
Table 4.20	Figure 4.5: Histogram and CDF Plots of an Empirical Distribution for Glass Fibres Data	105
Table 4.21	Fitted Plot of EGEE EGE, EE and E Models on Glass Fibres Data	107

Table 4.22	Histogram and CDF Plots of an Empirical Distribution for precipitation in Minneapolis/St Paul	108
Table 4.23	Fitted Plot of EGEE EGE, EE and E Models on precipitation data	111

LIST OF ABBREVIATIONS AND NOTATIONS

ABBREVIATIONS	MEANING
CDF	Cumulative Density Function
E	Exponential
EGF	Exponentiated Generalized Family
EGT	Exponential Geometry
EL	Exponential Logarithmic
EGFT	Exponentiated Generalized Frechet
EGIE	Exponentiated Generalized Inverted Exponential
EGEE	Exponentiated Generalized Exponentiated Exponential
EE	Exponentiated Exponential
EGE	Exponentiated Generalized Exponential
GLD	Generalized Lamda Distribution
GG	Gamma Generated
GBG	Generalized Beta Gamma
GE	Generalized Exponential
iid	Independent And Identically Distributed
LL-G	Log- Gamma Generated

PDF	Probability Density Function
PMF	Probability Mass Function
SND	Skewed Normal Distribution
Kw-G	Kumaraswamy-Generated
KBG	Kummer Beta Generalized
WG	Weibull Geometric
WP	Weibull Poisson
WNB	Weibull Negative Binomial
3EE	Three(3) Exponentiated Exponential

CHAPTER ONE

GENERAL INTRODUCTION

1.1 Background to the Study

Probability distribution functions generally are used in describing the real world phenomenal and are classified into discrete and continuous distribution functions, where the discrete is called probability mass function (PMF), and the continuous distribution is known as probability density function (PDF). The PMF and PDF of the probability distributions are used in obtaining probability of every observation when the unknown values for the parameter(s) of the PMF or PDF are known. Probability functions could be empirical or theoretical in nature (empirical when it obtains function from data while the theoretical function generates the data) and are generally probable not functional like other models. The theory of univariate continuous probability distributions has become one of the most extensively studied theories within the vanguard of statistical modeling. Univariate probability distributions have been used in recent decades in the modeling of real-life random processes. The inability of the classical univariate functions to adapt and fit some data sets lead to the development of flexible univariate probability functions, starting from the phenomenal work of Pearson (1895) down to the recent works of modern scholars.

The introduction of procedure for deriving density function with the use of the system of differential equation by Pearson, (1895) resulted in the formulation of some univariate continuous distributions. The inability of the Pearson method to give analytic solution to the formulated density function has led to several methods for generating either probability density functions or distribution functions as appeared in the literature. The complex nature of data structure leads to the breakdown of classical continuous univariate distributions, which possesses one or two parameters. Indeed, the development of new univariate PDF is

based on one fundamental premise – the need to arrive at probability distributions capable of fitting real-life data adequately. In the absence of constant hazard rate property arising from real-life processes, to capture the characteristics of such processes, it becomes inadequate to employ the exponential distribution. This circumstance resulted to the formulation of different generalizations of the exponential distribution so that it can handle processes with increasing and decreasing hazard rates. Generally, the development of new univariate PDF has been carried out either by adding extra -parameter(s) to an already existing or by merging more than one probability distributions to obtain a new one.

The techniques for generating univariate PDF was reviewed by Lee et al (2013) and were placed into two-time epochs, namely: the pre-1980 and the post-1980 periods. These methods include the approach based on systems of differential equations (Pearson,1895; Burr, 1942), quantiles (Hastings et al., 1947; Tukey, 1960) and transformation (Johnson, 1949); Other methods include the method for generating skewed distributions (Azzalini, 1985); method of adding parameter and generalization (Mudholkar and Srivastava; 1993; Marshall and Olkin, 1997; Shaw and Buckley, 2009), method based on generators (Cordeiro and de Castro, 2011); composite method (Cooray and Ananda, 2005) , Exponentiated Generalized method, Cadeiro et al. (2013) and the transformed – transformer method (Alzaatreh et al., 2013a; Aljarrah et al., 2014; Alzaatreh et al., 2014).

The enhancement of flexibility to probability distribution by adding extra parameter(s) to existing continuous univariate through generalization or by combining it with another probability distribution is motivated upon the fact that more parameters in a distribution enhance the shape flexibility of such a probability distribution and this, in turn, makes the probability distribution to be more adaptive in fitting several data types such as environmental pollution in environmental science; machine life cycle in engineering, survival times of patients after surgery in the medical science, failure rate of software in computer

science and average time from marriage to divorce in the social science which generates data sets with various degrees of skewness and kurtosis.

The negative exponential distribution simply referred to as exponential distribution is a distribution that illustrates the time interval among events in a Poisson process. The exponential distribution is a popular statistical distribution, most likely, is a parametric model widely used in numerous fields (Lemonte et al., 2013). The attractiveness of this distribution can be clearly explained by the simplicity of its cumulative function, which involves an unknown parameter $\theta > 0$, such that $G(x) = 1 - e^{-\frac{x}{\theta}}$ for $x > 0$ in addition to having constant hazard rate. The Exponential distribution plays a vital role in real-life applications, it is mostly used in reliability analysis and applied lifetime distributions. The distribution is found applicable in various areas where events occur independently, at random but with mean rate $\frac{1}{\theta}$ per unit of time, distance, volume e.tc, for illustration, in modeling:

- (i) the interval of time amongst emissions of radioactive material;
- (ii) the failure time of manufactured items;
- (iii) the length of queues at particular segments of the highway at different times;
- (iv) the inter-arrival times at ticket counters and
- (v) in insurance, the amount of insurance losses e.t.c.

The fact that the exponential distribution is the only continuous distribution with a memory less property makes it plays a crucial role in probability and statistics.

During the first half of the nineteenth century, Gompertz, (1825) and Verhulst, [(1838) and (1847)], formulated some cumulative distribution functions (CDF's) in order to compare well-known human mortality tables which signify mortality growth. One of the CDF's developed is

$$G(t) = (1 - \rho e^{-\lambda t})^\alpha \quad (1.1)$$

for $t > 1/\lambda \ln \rho$.

Where all parameters $(\rho, \lambda, \text{ and } \alpha)$ in (1.1) are all positive real numbers. Ahuja and Nash (1967), in twentieth-century, furthermore, considered (1.1) and made some additional generalization. The exponentiated exponential or generalized exponential distribution is defined as a particular case of the [Gompertz (1825), Verhulst, (1845) and (1847)] distribution function, when ρ is one. Therefore, X is a two-parameter exponentiated exponential random variable if it has the PDF $G(x; \alpha, \lambda) = (1 - e^{-\frac{x}{\theta}})^k$ and CDF. $g(x; \alpha, \theta) = \frac{k}{\theta} (1 - e^{-\frac{x}{\theta}})^{k-1} e^{-\frac{x}{\theta}}$ where θ and k play the role of the scale and shape of the distribution respectively. The generalized or exponentiated family of distribution express as $F(x) = (G(x))^k$ where $G(x)$ is a baseline cumulative distribution function and k serves as the added shape parameter to the existing distribution. The properties and application of Exponentiated Exponential (EE) Distribution was studied by Kundu and Gupta (1998) and they established that it can serve as a substitute distribution to both Gamma and Weibull distributions.

Furthermore, a class of distributions was proposed by Cadeiro et al., (2013) by adding two parameters to a univariate continuous distribution so as to extend the thought initially presented by Lehman, (1953) and thereafter considered by Nadarrah and Kotz, (2006). This technique prompts to another class of Exponentiated Generalized Family (EGF) of distribution that can be considered as a dual construction of Lehmann alternative. The EGF distributions broaden the exponentiated form of distribution and obtained some of its structural properties. Given the CDF $G(x)$, the CDF of class of EGF distributions is defined by

$$F(x) = [1 - \{1 - G(x)\}^\alpha]^\beta \quad (1.2)$$

and consequently the probability distribution function is obtained as;

$$f(x) = \alpha\beta\{1 - G(x)\}^{\alpha-1}[1 - \{1 - G(x)\}^\alpha]^{\beta-1}g(x) \quad (1.3).$$

where $\alpha > 0$ and $\beta > 0$ are added shape parameters. The Exponentiated Generalized Family has some properties that are tractable, particularly for simulations, since its quantile function can be expressed in a simple form:

$$x = Q_G\left(1 - \left(1 - u^{\frac{1}{\beta}}\right)^{\frac{1}{\alpha}}\right) \quad (1.4)$$

where $Q_G(u)$ is the existing quartile function. The EGF class of distributions generates new probability function and allows greater flexibility to the tail of the existing distribution that is widely applicable in real-life situations. The additional parameters α and β in the density function generate a new flexible distribution which can control both tail weights with heavier or lighter tail as appropriate. Also, the physical interpretation of the EGF distribution when α and β are positive integers are appealing (Cordeiro and Lemonte, 2014). Cordeiro et al., (2013) obtained several properties of the EGF class of distributions by illustrating its flexibility with some well-known univariate distribution such as Normal, Gamma, Gumbel and Frechet distributions. This therefore, provides motivations to adopt this generator.

1.2 Statement of the Problem

Different phenomena have been explained well by developing various classes of generalized distributions. The basic property of these distributions is that they have additional parameter to the existing ones. This property of generalized distributions has prompted the interests of researchers in developing a robust distribution which have remained active in the statistics profession (Alzaatreh et al., 2013). Existing classical distributions tend to perform poorly

when fitting data because of the diverse nature of the data. Johnson et al., (1994) affirmed that the application of a four-parameter distribution should be adequate for most realistic situations. The authors stated that at least four parameters are required, but they doubted any obvious improvement that can come up by including a fifth or sixth parameter. The failure of distributions to correctly describe both positive and negative skewed data is worrisome. Adding two parameters to the baseline distribution (Exponentiated Exponential) would generate a distribution that will serve as a better alternative distribution to Exponentiated Exponential, Exponentiated Generalized Exponential and Exponential distributions in fitting both negative and positive skewed data sets. This study therefore presented yet another flexible four parameters statistical distribution that will adequately capture skewness in data sets as compared to existing distributions.

1.3 Justification of the Study

A new method of adding two parameters to univariate continuous distribution was introduced by Cordeiro *et al.* (2013). The flexibility of the family was demonstrated by applying it to some well-known distributions. The Exponentiated Generalized Exponential (EGE), Exponentiated Exponential (EE) and Exponential (E) distributions which were developed over the years performed pretty better when fitting right-skewed data sets, but perform poorly when fitting left-skewed data set. The improved EE by Gupta and Kundu (1998) and EGE by Oguntunde *et al.* (2016) tend to have no significant difference when fitting some skewed data sets. This study therefore used Cordeiro *et al.* (2013) family of distribution to develop new probability distribution that will be suitable in fitting positively and negatively skewed datasets. The proposed distribution was then compared with other existing distributions using simulated and real life data sets. It is believed that the addition of two parameters to a baseline EE distribution which serves as substitute to Gamma and Weibull distributions will produce a new probability distribution with tractable properties that is appropriate to fit both

positively and negatively skewed real-life data. This study therefore presented yet a different flexible four parameters probability distribution that will fit both positively and negatively skewed data sets.

1.4 Aim and Objectives

This study aimed at developing a suitable distribution from an Exponentiated Generalized Family of distributions for modeling positively and negatively skewed data. The main objectives were to:

- (i) derive new statistical distribution and its properties that will be suitable to fit positively and negatively skewed unimodal data;
- (ii) examine the effect of the shapes and scale parameters on the distribution, and
- (iii) compare the performance of the proposed distribution with existing distributions using simulated data and suitable sets of real-life data.

1.5 Significance of the Study

Distribution theory has been widely used for modeling observed situations whose various characteristics as reflected by the data differ from those that would be anticipated under the simple component distribution. The improvement of existing classical distribution by adding one or two parameters with the help of an established generator has led to the flexibility and adaption of these classical distributions to different data sets. Thus, the inability of Exponential class of distributions to properly fit both left-skewed, right-skewed, and unimodal datasets makes the formulation of another distribution from EGF class of distribution a paramount task. It is believed that the research to development of a new distribution, which will serve as the best alternative distribution to existing distributions will provide a significant contribution in the area of study.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This chapter presents the review of the literature on methods of generating new univariate distribution and their respective drawbacks in obtaining their properties, expression, and data application.

2.2 Statistical Distributions

Eugene et al., 2002 stated that Statistical distributions are usually significant for parametric inference and application in describing real-life situation. Their theories have been extensively studied, and new probability distributions were generated from the existing ones. Some proposed procedures in eighteen century for generating continuous distributions include methods of differential equations (DE) proposed by Pearson (1895). Pearson (1895) did propose that for any member of PDF $f(x)$ satisfies DE;

$$\frac{1}{f(x)} \frac{df(x)}{dx} = \frac{a + x}{c_0 + c_1 + c_2 x^2} \quad (2.1)$$

where a, c_0, c_1 , and c_2 are the parameters which control the shape of the distribution.

The roots of the equation

$$c_0 + c_1 + c_2 x^2$$

determine the solution to form equation (2.1)

Pearson (1895) method was established in the literature for generating the Pearson type I – VI probability distributions, which were well known as beta, student t, generalized gamma or chi-square, inverse gamma, and beta prime or f-distributions respectively. The family of

Pearson (1895) subsumes some of the following distributions, the normal, exponential, and continuous uniform distributions.

Dunning and Hanson (1977) generalized the Pearson (1895) frame work in (2.1) by expanding the DE as

$$\frac{df(x)}{dx} = \frac{a_o + a_1x + a_2x^2 + \cdots \dots a_mx^m}{c_o + c_1x + c_2x^2 + \cdots \dots c_nx^n} \quad (2.2)$$

where $m > 0$ and $n > 0$. The generalized system proposed by Pearson in (2.2) becomes the Pearson (1895) system in (2.1) when $m = 1, n = 2, a_o = a$, and $a_1 = 1$. It appears from the literature that attention has been paid to the study of the family PDF's that can be derived as a solution to the generalized Pearson DE in (2.2), Chaudhry and Ahmad (1993), Cobb et al. Roy (1971) and Dunning and Hanson (1977), Lefevre *et al.* (2002), Sankaran *et al.* (2003), Stavroyiannis *et al.* (2007), Rossani and Scarfone (2009), Shakil *et al.* (2010), Hamedani (2011), Ahsanullah *et al.* (2013) and Shakil and Singh (2015) have all used the generalized Pearson to generate some distributions.

Burr (1942) proposed a new system of generating probability distribution based on differential equation and expressed it as:

$$dF = F(1 - F)g(x)dx \quad (2.3)$$

where $0 \leq F \leq 1$ and $g(x) > 0$

Base on some choice of $g(x)$, Burr, (1942) was able to arrive at twelve solutions to equation (2.3) in the form of CDF and became known as the Burr type 1-X11 distributions. The Burr (1942) system of the DE must satisfy the DE in (2.3) and base on some choice of $g(x)$. Fry, (1993) and Johnson *et al.* (1994) gave a detailed review of Burr system of distributions. Raqad, (1996) worked on the r^{th} order statistics of Burr X distribution, Mehment *et al.*,

(2017) illustrated the Burr III distribution in real life situations and finance literature etc. Mohan and Todd, (2013, 2014) introduced a method of simulating univariate and multivariate to Burr type III, VII and XII distributions through methods of L-moments and L-correlation. The method base on the differential equation has its shortcoming when it comes to generating a new distribution; some of them are:

- (i) the Pearson system does not include the logistic distribution for any of its solutions;
- (ii) obtaining solutions to the system of differential equations have proved to become a herculean task because many of the times analytical solutions do not exist for some chosen values of constants in the system; and
- (iii) apart from types I, III, X and XII of the Burr-distribution family, the properties of ~~the~~ some members are analytically not tractable hence their limited applications to real-life problems.

Johnson (1949) developed a method of generating distributions based on normalization transformation and defined its general form as:

$$z = \gamma + \delta f\left(\frac{x - \varepsilon}{\lambda}\right) \quad (2.4)$$

where $f(\cdot)$ is the transformation function, z is the standard normal random variable while γ, δ, λ , and ε are parameters. Johnson, (1949) assumed δ and λ to be positive and further defined three transformation functions as;

- i) the lognormal family

$$z = \gamma + \delta \ln(X - \varepsilon) \quad X \geq 0 \quad (2.4.1)$$

- ii) the bounded distribution family

$$z = \gamma + \delta \ln\left(\frac{x - \varepsilon}{\varepsilon + \lambda - X}\right) \quad \varepsilon \leq X \leq \varepsilon + \lambda \quad (2.4.2)$$

iii) the family of unbounded distribution

$$z = \gamma + \delta \sin^{-1} \left(\frac{x - \varepsilon}{\lambda} \right) \quad -\infty \leq X \leq \infty \quad (2.4.2)$$

Athayde, (2012) used the Johnson transformation to obtain the Birnbaum- Saender distribution.

Tukey, (1960) developed another method of transformation based on generating skew and heavy-tailed distribution known as the Tukey family. Fischer (2010) gave a more detailed review of the Tukey family as a generator. The general form of the Tukey family is expressed as

$$Y = Z.T^{\theta}(Z) \quad \theta \in \mathcal{R} \quad (2.5)$$

where the function $T(Z)$ is positive symmetric and strictly monotone increasing for $Z \geq 0$. Tukey (1977) introduced another family of distribution by using two nonlinear transformations called $g - h$ distribution which stated as

$$Y = Z.T_{g,h}(Z) = \frac{1}{g} (\exp\{gZ\} - 1) \exp \left\{ \frac{hZ^2}{2} \right\} \quad g \neq 0, h \in \mathcal{R} \quad (2.6)$$

where Z is a standard normal distribution. The Tukey's $g-h$ family of distributions has more useful features because of its ability to produce a good number of the Pearson, (1895) distributions and can also create most of the well-known distributions, Martinez and Iglewics, (1984). The two major limitations of methods based on transformation are:

- i) well defined CDF and PDF of generated distribution are not made available by the method (Headrick *et al.*, 2008); and
- ii) transformation function can only be positive symmetry and strictly monotone function.

Hastings *et al.*, (1947) and Tukey, (1960) worked on the generalized lambda families of distribution which is defined in term of quantile function and expressed its general form as

$$Q_{(p)} = Q(p; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2} \quad 0 \leq p \leq 1 \quad (2.7)$$

where λ_1 and λ_2 are location and scale parameters respectively while λ_3 and λ_4 are shape parameters, the corresponding pdf is;

$$f(x) = \frac{\lambda_2}{\lambda_3 p^{\lambda_3-1} + \lambda_4 (1-p)^{\lambda_4-1}} \quad x = Q_{(p)} \quad (2.8)$$

The required condition for (2.8) to be a true pdf is that λ_2 and $\lambda_3 p^{\lambda_3-1} + \lambda_4 (1-p)^{\lambda_4-1}$ must have the same sign for all p in $(0,1)$. The extended form of generalized lambda distribution (GLD) was done by Karian *et al.*, (1996) and incorporated the GLD and generalized beta distribution (GBD) with the pdf.

$$f(x) = \begin{cases} \frac{(x - \beta_1)^{\beta_3} (\beta_1 + \beta_2 - x)^{\beta_4}}{B(\beta_3 + 1, \beta_4 + 1) \beta_2 (\beta_3 + \beta_4 + 1)} & \beta_1 \leq x \leq \beta_1 + \beta_2 \\ 0, & \text{otherwise} \end{cases} \quad (2.9)$$

where $B(\dots)$ is complete beta function. A thorough review of the GLD and its extension was done by Karian and Dudewicz, (2000).

Vort, (1992) combined the Pearson and Tukey lines of approach to define a system of S-distribution by differential equation;

$$f(x) = \frac{dF(x)}{dx} = \alpha [(F(x))^g - (F(x))^h], \quad F(x_0) = F_o, \quad (2.10)$$

where α is a scale parameter g and h are shape parameters, the initial value x_o is equivalent to the F_o , quantile which is often taken to be 0.5 so that x_o is the median or the second quartile for $f(x)$ to be true density, it is required that $g < h$, if $g = h$ and $f(x) =$

0 for all x and $\alpha \neq 0$ and without loss of generality $\alpha > 0$ the logistic distribution ($g = 1, h = 2$) is a product of equation 2.10. Its limitation exists in the literature as:

- i) Freimer *et al.* (1988) did say that the GLD does not cover all skewness and kurtosis values;
- ii) according to Muino *et al.* (2006), the system of S- distribution only includes one symmetric distribution, the logistic distribution; and
- iii) the S- distribution system does not correctly accommodate distribution with a heavy tail, as well as the finite right tail, Muino *et al.*, (2006).

Generating a generalized distribution received less or no attention until the study of skewed symmetric distribution (SSD) and skewed normal distribution (SND) of Azzalini (1985) having probability distribution function (pdf).

$$f(x) = 2g(x)G(\lambda x), \quad -\infty < x < \infty \quad \lambda \in \mathcal{R} \quad (2.11)$$

where $g(\cdot)$ and $G(\cdot)$ are PDF and CDF respectively of standard normal variable, λ is the skew parameter interpreted as the parameter of asymmetry. Kotz and Vicari, (2005) and Lee *et al.* (2013) gave a detailed summary and characterization of generalized SND.

Azzalini (2005) improved equation 2.11 by defining the pdf as:

$$f(x) = 2g(x)G(w(x)), \quad x \in \mathcal{R} \quad (2.12)$$

where $g(x)$ is symmetric about 0, $G(x)$ and $w(x)$ are any odd functions. Arnold *et al.*, (2000) and Arnold and Beaver, (2002) used other distribution other than the normal distribution adopted by Azzalini. An alternative method for introducing skewness to continuous univariate unimodal symmetry distribution was proposed by Fernandez and Steel, (1998) and defined the pdf as:

$$g(x) = \begin{cases} cf(\alpha x), & x \geq 0 \\ cf(x/\alpha), & x < 0 \end{cases} \quad (2.13)$$

when $\alpha \neq 1$ the distribution becomes skewed and:

$$c = 2\alpha / (1 + \alpha^2)$$

is the normalization constant defined in order to make $g(x)$ a valid density function. Ferreira and Steel, (2006) proposed another framework for introducing skewness to symmetric distributions based on the inverse probability integral transformation and defined the family as:

$$g(x/f, p) = f(x)p(F(x)), \quad x \in R \quad (2.14)$$

if p is uniform $g = f$. The pdf $g(x/f, p)$ is a weighted version of $f(x)$, with the weight function given by $p(F(x))$. All the methods of adding skewness to symmetric distributions have their shortcomings, such as:

- (i) it only considers symmetric by adding skewness and not necessarily enhancing flexibility of shape of the distributions; and
- (ii) most of the developed (proposed) family of distributions using the method always produce distributions whose tails are thinner than that of normal distribution Azzalini (1985).

Mudholkar and Srivastava, (1993) used an approach based on *random variable* X with existing CDF $F(x)$ to classify a new family of distribution expressed as:

$$G(x) = (F(x))^\theta \quad x \in R \quad \theta > 0, \quad (2.15)$$

where θ is a shape and when $\theta = 1$ equation 2.15 reduces to CDF of the baseline distribution. Gupta and Kundu, (1998), and Mudholkar and Srivastava used these approaches to transform the Exponential and Weibull distributions respectively.

Another method for introducing a parameter to distribution within the context of lifetime analysis was proposed by Marshall and Olkin (1997) and is defined as:

$$G(x) = 1 - \frac{\alpha S(x)}{1 - (1 - \alpha)S(x)}, \quad x \in R \quad \alpha > 0 \quad (2.16)$$

where $S(x)$ being a survival function of the existing distribution and α is the additional parameter, when $\alpha = 1$ equation (2.16) change back to the existing distribution. The case of Weibull and exponential distribution was examined by Marshall and Olkin, (1997).

Shaw and Buckley, (2007) introduced quadratic rank transmutation map family which serves as an alternative approach and called it transmuted distribution defined by

$$G(x) = (1 + \lambda)F(x) - \lambda(F(x))^2 I \lambda I \quad (2.17)$$

when λ assumes the values $-1, 0$ and 1 equation (2.17) reduces to exponentiated family in equation (2.15) (with $\theta = 2$), the existing $F(x)$ and a quadratic form of the existing distribution $F(x)$ respectively. Shaw and Buckley, (2007), Das, (2015) and Bourguignon *et al.*, (2016) studied the general characteristics of the family. Tahir and Cordero (2016) listed some transmuted family as defined by various authors. There are majorly two limitations to these methods.

- i) to state the properties of the distribution in a simple and analytical form can be very tasking.
- ii) the methods only accommodate the introduction of only one parameter to the existing distributions which may not improve the flexibility required.

Adamidis and Loukas, (1998) pioneered the method for compounding continuous distribution and discrete family of distribution using conditional and marginal CDF. The method considers a continuous positive random variable X with CDF $F(x)$ and discrete random variable N with probability mass function pmf p_n , truncated at zero. Given N , let $[X_i]_{i=1}^N$ be a sequence of independent and identical distributed (iid) random variable from the distribution of X . Suppose $X_{(1)} = \min(X_{i=1}^N)$. The conditional CDF of $X_{(1)}/N = n$ is given by

$$Z_{x(1)/N=n}(x) = 1 - \prod_{i=1}^n [1 - F_i(x)] = 1 - [1 - F_i(x)]^n$$

Adamidis and Loukas, (1998) defined a new class of distribution using the marginal cumulative distribution function of the distributions of $X_{(1)}$ by the cumulative distribution function:

$$G(x) = \sum_{i=1}^n p_n Z_{x(1)/N=n}(x) \quad x > 0 \quad (2.18)$$

Several distributions have been compounded over the years. Adamidis and Loukas (1998) defined a 2-parameter Exponential-Geometric [EGT] by compounding the exponential geometric distributions. Kus, (2007), Tahmasbi and Rezaei, (2008), Barreto-Souza *et al.*, (2010a), Lu and Shi, (2011) and Rodrigues *et al.*, (2011) defined the Exponential Poisson (EP), Exponential-Logarithmic (EL), Weibull-Geometric (WG), Weibull Poisson (WP) and Weibull Negative Binomial (WNB) distributions respectively. Tahir and Cordeiro (2016) gave a comprehensive survey of the compounded class of distributions. The method based on compounding has following its limitations:

- (i) the method only accommodates continuous distribution with positive intervals;
- (ii) only zero truncated distribution can be compounded; and

(iii) results obtained can only be expressed in infinite series or special functions.

The family of the beta-generated class of distributions was proposed by Eugene *et al.*(2002). He utilized the property of random variable lies somewhere in the range of 0 and 1. The CDF of beta-generated distribution is expressed as follows;

$$G(x) = \int_a^{F(x)} b(t)dt \quad (2.19)$$

where $b(t) = t^{\alpha-1}(1-t)^{\beta-1}$ and $F(x)$ is the cumulative distribution of any random variable X. The PDF for beta-generated distribution is derived and expressed as follows:

$$g(x) = \frac{f(x)}{B(\alpha, \beta)} F^{\alpha-1}(x)(1-F(x))^{\beta-1} \quad \alpha > 0, \beta > 0 \quad (2.20)$$

where $B(\alpha, \beta)$ is the Beta-function. The equation (2.20) can generally be considered as order statistics (Jones, 2004, Eugene *et al.*, 2002). They further examined the beta-normal distribution and its properties as a member of the beta generated class of distributions.

Furthermore, the beta-normal distribution was shown by Famoye *et al.*, (2004) that it could be unimodal or bimodal, which adds more flexibility in fitting real-life data. Famoye *et al.*, (2005), defined the Beta-Weibull distribution and proved that the distribution is unimodal and find expressions for mean deviations and non-central moments. Lee *et al.*, (2007) did a further study to examine the Rényi entropy and hazard function of the distribution, and imposed censored data so as to compare the performance of Beta-Weibull distribution to that of Exponentiated Weibull and Weibull distributions. Nadarajah and Kotz (2004) formulated Beta-Gumbel distribution and studied extremes order statistics with the shapes and showed that the hazard function is always growing. The expressions for the non-central moments were derived and recommended the method of maximum likelihood (MLE) for estimation (MLE) of the parameters of the distribution.

Nadarajah and Kotz (2004) also defined the Beta-Exponential distribution and derived mathematical expressions for the first four raw moments for the probability distribution. Thereafter, they estimated Rényi and Shannon entropies, mean deviations, for the Beta-Exponential distribution, and recommended the MLE for estimating the parameter of the distribution. Several other Beta-Generated distributions have been proposed in the literature which have been very adaptive to several applications. Distributions in this class include Beta-Gamma (Kong *et al.*, 2007), Beta-Pareto (Akinsete *et al.*, 2008), Beta-Generalized Exponential (Barreto Souza *et al.*, 2010) and Beta-Generalized Pareto (Mahmoudi, 2011).

Following the heels of the developers of beta generated class of distribution, Cordeiro and de Castro (2011), and Jones (2009) proposed the Kumaraswamy-generated (Kw-G) class of distributions, using the Kumaraswamy as the generator just like the beta distribution was employed as the generator in the beta-generated member of the distributions. The general expressions for the density function, moments and density of order statistics of the Kw-G were derived by Cordeiro and de Castro, (2011) in terms of its PDF and CDF. The family of Kw-G PDF and CDF are respectively defined as;

$$g(x) = \alpha\beta f(x)F^{\alpha-1}(x)(1 - F(x))^{\beta-1} \quad a > 0, \beta > 0 \quad (2.21)$$

and

$$G(x) = 1 - (1 - F(x)^\alpha)^\beta \quad (2.22)$$

respectively. Some of the notable derived distributions under Kw-G family are Kumaraswamy-half normal by Cordeiro *et al.*, (2012), Kumaraswamy Pareto distribution by Nadaraja and Eljbri, (2013), Kumaraswamy linear Exponential distribution by Elbatal,(2013), Kumaraswamy Lomax distribution by Sham, (2013), Kumaraswamy Weibull distribution Cordeiro by *et al.*,(2014), Kumaraswamy generalized Rayleigh by Gomes *et al.*,(2014), Kumaraswamy Marshall- Olkin family of distribution by Alizadeh *et al.*,(2015)

Kumaraswamy Pareto IV Distribution by Tahir *et al* (2015) and Kumaraswamy-Exponential distributions among other. The Kumaraswamy class of distributions was used to compare the fit of other member of the existing distributions such as Weibull, Exponentiated Weibull and Exponentiated Exponential.

However using the methodology of Eugene *et al.*, (2002) and Cordeiro and de Castro, (2011) Pescim *et al.* (2012) introduced a new class of Kummer Beta Generalized (KBG) family of distributions that is suitable in fitting skewed data. From an arbitrary parent or baseline distribution with CDF $G(x)$ of random variable X and pdf $g(x)$, the KBG is defined by the following cdf:

$$F(x) = K \int_0^{G(x)} t^{\alpha-1} (1-t)^{\beta-1} e^{-pt} dt \quad (2.23)$$

where $\alpha > 0, \beta > 0$ are two shape parameters and $-\infty < p < \infty$ is a scale parameter.

The pdf of KBG is expressed by

$$f(x) = Kg(x)(G(x))^{\alpha-1} \{1 - G(x)\}^{\beta-1} e^{-pG(x)} \quad (2.24)$$

For $p = 0$ the KBG reduces to the beta generated family (see Pascim *et al.* (2012)).

The beta-generated class defined in (2.19) and its alternative in (2.23) required the use of a generator (beta and Kumaraswamy distributions) which are defined on $(0, 1)$.

The class of the gamma-generated family (GG) utilizes the gamma distribution $Gam(\alpha, 1)$ as the generator. Zografos and Balakrishnan, (2009) introduced a kind of GG family which is based on a baseline continuous or parent distribution $G(x)$ of a random variable X with parent survival function $\bar{G}(x)$ and parent density $g(x)$. They defined the cumulative density and probability density functions of the GG family respectively as:

$$F(x) = \frac{1}{\Gamma(\alpha)} \int_0^{-\log(1-G(x))} t^{\alpha-1} e^{-t} dt, \quad (2.25)$$

$$f(x) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, -\log(1 - G(x))), x \in R \quad \alpha > 0 \quad (2.26)$$

and

$$f(x) = \frac{1}{\Gamma(\alpha)} [-\log(1 - G(x))]^{\alpha-1} g(x) \quad (2.27)$$

where $\Gamma(\cdot)$ is the gamma function and $\gamma(\cdot, \cdot)$ is the incomplete lower gamma function. Following related procedure, Ristic and Balakrishnan, (2011) proposed a more robust gamma-generator which is shown by the CDF as follows:

$$F(x) = \frac{1}{\Gamma(\alpha)} \int_0^{-\log(G(x))} t^{\alpha-1} e^{-t} dt, \quad x \in R, \quad \alpha > 0 \quad (2.28)$$

while the resultant PDF is;

$$f(x) = \frac{1}{\Gamma(\alpha)} [-\log(1 - G(x))]^{\alpha-1} g(x) \quad (2.29)$$

The beta-class distributions were further extended by Alexander *et al.*, (2012), and the Kumaraswamy generated distributions by establishing a family of generalized beta generated (GBG) distributions using the density of a special case of GB1 as a generator. The beta-class of the Exponentiated distribution was compared with the GBG family. They further studied the general properties of the distribution and incorporated the proposition of expressions for the quantile functions and moments. They also used real-life data sets to compare the generalized-beta normal and the generalized-beta-Weibull distributions to other distributions. Alzaatreh *et al.*, (2013a) suggested another method for generating new distributions the $T - X$ class of distributions. They used a random variable T defined on the range $[a, b]$, $-\infty \leq a < b \leq \infty$ with CDF and PDF of $r(t)$ and $R(t)$ respectively and another random variable X with CDF $F(x)$ and PDF $f(x)$. By transforming $F(x)$ using $W(F(x))$, a new class $W(\cdot)$ of distributions was defined by the CDF of the form of;

$$G(x) = \int_a^{W(F(x))} r(t) dt \quad (2.30)$$

where $W(.)$ satisfies the conditions ;

- i.) $W(F(x)) \in [a, b]$
- ii.) W is differentiable and monotonically non-decreasing,
- iii.) $W(F(x)) \rightarrow a$ as $x \rightarrow -\infty$ and $W(F(x)) \rightarrow b$ as $x \rightarrow \infty$

They called the distribution in (2.30) the “ $T - X$ distribution.” Several functions $W(.)$ were defined by the authors, and a list of new families of distribution using $F(x)$ as the random generator variable were outlined. Observe that the $G(x)$ is a CDF since it satisfies the required conditions stated as follows for a distribution function:

- (i) G is non-decreasing,
- (ii) G is right-continuous,
- (iii) $G(x) \rightarrow 0$ as $x \rightarrow -\infty$ and $G(x) \rightarrow 1$ as $x \rightarrow \infty$

Furthermore, let $R(t)$ and $r(t)$ represent the PDF and CDF respectively of the continuous random variable T and let X represents any continuous random variable with CDF and PDF $F(x)$ and $f(x)$ respectively and also take $W(F(x))$ as the logit of the cdf $F(x)$ i.e.

$W(F(x)) = \ln\{F(x)/(1 - F(x))\}$, which also corresponds to considering it as the quantile function of the standard logistic distribution, then the cumulative distribution function on the basis of the logit function is presented as follows;

$$\int_a^{\ln\{F(x)/(1-F(x))\}} r(t)dt = R(\ln\{F(x)/(1 - F(x))\}), \quad (2.31)$$

The resulting PDF of the $T - X$ distribution in (2.32) is derived by taking a differentiation of (2.31) with respect to x to get

$$g(x) = r(\ln\{F(x)/(1 - F(x))\}) \frac{h(x)}{F(x)} \quad (2.32)$$

where $h(x) = \frac{f(x)}{(1-F(x))}$ is the hazard function of the random variable X with CDF $F(x)$.

Observe that $R(\ln\{F(x)/(1 - F(x))\}) = R(t)$

$$\Rightarrow \{F(x)/(1 - F(x))\} = t ,$$

Therefore,

$$x = F^{-1}(e^t/(1 + e^t))$$

hence ,

$$X = F^{-1}(e^T/(1 + e^T)) \quad (2.33)$$

The equation (2.33) is helpful for simulation of the random variable X with CDF $G(x)$ of the continuous distribution by simulating the random variable T which serves as the generator and thereafter transformed accordingly. Equation (2.33) can also be applied to obtain the r^{th} non-central moment of the $T - X$ random variable X . Alzaatreh *et al.*, (2013b) developed the Weibull-Pareto distribution as a part of the family. The limiting behavior, moments, shapes, and Shannon entropy for the Weibull-Pareto distribution were examined. Al-Aqtash *et al.*, (2014) also developed the Gumbel-Weibull distribution which is also a member of the T-X family. They study some properties of the distribution, such as, the modes and mean deviations. The maximum likelihood method was used ~~for~~ to estimate unknown parameters of the distribution. The distribution was applied to six different data sets to illustrate the flexible nature of the distribution when fitting bimodal and unimodal data sets. Osatohanmwon *et al.* (2017) proposed Gumbel-Burr XII distribution using T-X family framework, Numerous properties of the distribution were examined, and its relationship with other distributions was compared. The drawbacks associated with the T-X family are:

- (i) the cumulative density function related with the class in a few cases add scale parameter to the improved distributions.
- (ii) the failure of T-X family to add not more than one shape parameter to existing distributions had made it less flexible.

Amini *et al.*, (2014) introduced two new general families of continuous distributions called log gamma generated families of distributions (LL-G). For any continuous parent $G(x)$ of a

random variable X with corresponding parent pdf $g(x)$, Amini *et al.*, (2014) defined the two new LL-G families with the following two pdfs

$$f_1(x) = \frac{\sigma_1^c}{\Gamma(c)} [-\log \bar{G}(x)]^{c-1} [\bar{G}(x)]^{\sigma_1-1} g(x) \quad (2.34)$$

$$f_1(x) = \frac{\sigma_1^c}{\Gamma(c)} [-\log G(x)]^{c-1} [G(x)]^{\sigma_1-1} g(x) \quad (2.35)$$

Where c and σ_1 are two positive shape parameters. The drawback associated with LL- G family, are;

- (i) obtaining the CDF and simulating data from the family is very tasking.
- (ii) stating the properties of the distribution in a simple and analytical form can be tasking.

2.2 Appraisal of the Literature Reviewed

Some methods of generating new univariate continuous distribution have been reviewed along with their different shortcomings. Majorly, seven different approaches have been used in the past decades to generate new distribution, which are the method of differentiation, method of addition of parameter, method base on quantile function, method base on survival function, method base on composite function, method generated from probability density function to methods generated from cumulative distribution functions. However, among the methods studied since inception, it is only the method of Exponentiated Generalized Family of distribution that guarantees the following:

- (i) addition of two shape parameters with simplicity to baseline distributions;
- (ii) provides tractable properties, especially for simulation; and
- (iii) create greater flexibility for heavy tail nature of real-life data sets.

CHAPTER THREE

METHODOLOGY

3.0 Introduction

This chapter presents the theoretical structure of the Exponentiated Generalized Family (EGF) of distribution, existing distributions on EG, properties of proposed Exponentiated Generalized Exponentiated Exponential (EGEE) distribution and information criteria used to justify the additional parameters to baseline Exponentiated Exponential (EE) distribution.

3.1 Theoretical Framework

This study extended the work of Cadeiro *et al.*, (2013) who proposed a family of distributions by introducing two parameters to an existing continuous distribution, by improving the idea initially proposed by Lehman, (1953) and studied by Nadarrah and Kotz, (2006). This technique lead to a new family EGF of distribution that can be taken to be a double construction of Lehmann alternative. The distributions broaden the Exponentiated type of distribution and derived some of its theoretical properties. Given a cumulative distribution $G(x)$, we express the EGF family of distributions is expressed as follows;

$$F(x) = [1 - \{1 - G(x)\}^\alpha]^\beta \quad (3.1)$$

and

$$f(x) = \alpha\beta\{1 - G(x)\}^{\alpha-1}[1 - \{1 - G(x)\}^\alpha]^{\beta-1}g(x) \quad (3.2)$$

where $\alpha > 0$ and $\beta > 0$ are the two added shape parameters. The quantile function of EGF has tractable properties, particularly for simulation because it exists in a simple form;

$$x = Q_G \left(\left[1 - \left(1 - u^{\frac{1}{\beta}} \right)^{\frac{1}{\alpha}} \right] \right) \quad (3.3)$$

Where $Q_G(u)$ is the baseline quartile function.

3.2 Existing Distributions From Exponentiated Generalized Family (EGF)

This study examined two existing distributions from EGF class of distribution. Two parameters were added to Frechet and Inverted Exponential distribution by Cordeiro *et al.*, (2013) and Oguntunde *et al.*, (2014) respectively by using Exponentiated Generalized Family

3.2.1 Exponentiated Generalized Frechet (EGFT)

The CDF of the Frechet distribution (for $x > 0$) is;

$$G_{\sigma,\lambda}(x) = \exp \{-(\sigma/x)^\lambda\}$$

where $\lambda > 0$ and $\sigma > 0$.

The corresponding CDF of Exponentiated Generalized Frechet (EGFT) cumulative distribution (for $x > 0$) from (3.1) is express as;

$$F(x) = [1 - \{1 - \exp \{-(\sigma/x)^\lambda\}^\alpha]^\beta \quad (3.4)$$

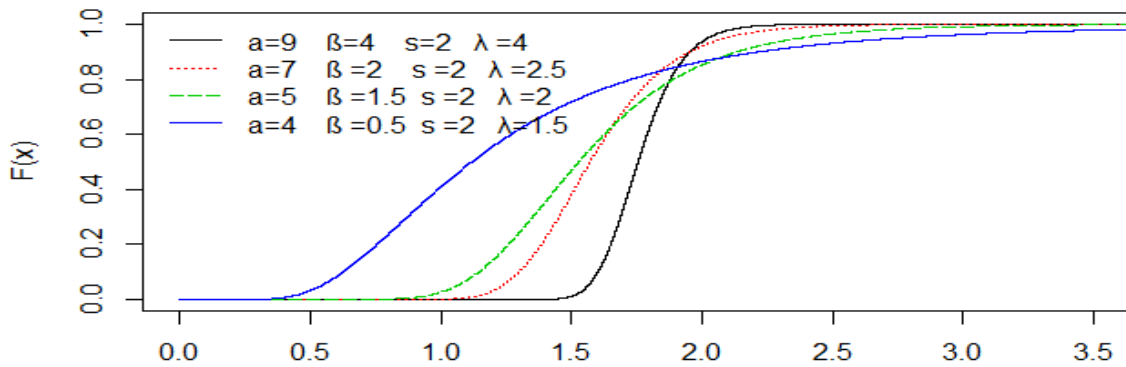


Figure 3.1: Cumulative Distribution Function of EGFT

where $\lambda > 0, \sigma = s > 0, \alpha = a > 0$ and $\beta > 0$.

The EGFT pdf is obtained from (3.4) as

$$f(x) = \alpha\beta\lambda\sigma^\lambda x^{-(\lambda+1)} \exp\{-(\sigma/x)^\lambda\} [1 - \exp\{-(\sigma/x)^\lambda\}]^{\alpha-1} \times (1 - [1 - \exp\{-(\sigma/x)^\lambda\}]^\alpha)^{\beta-1} \quad (3.5)$$

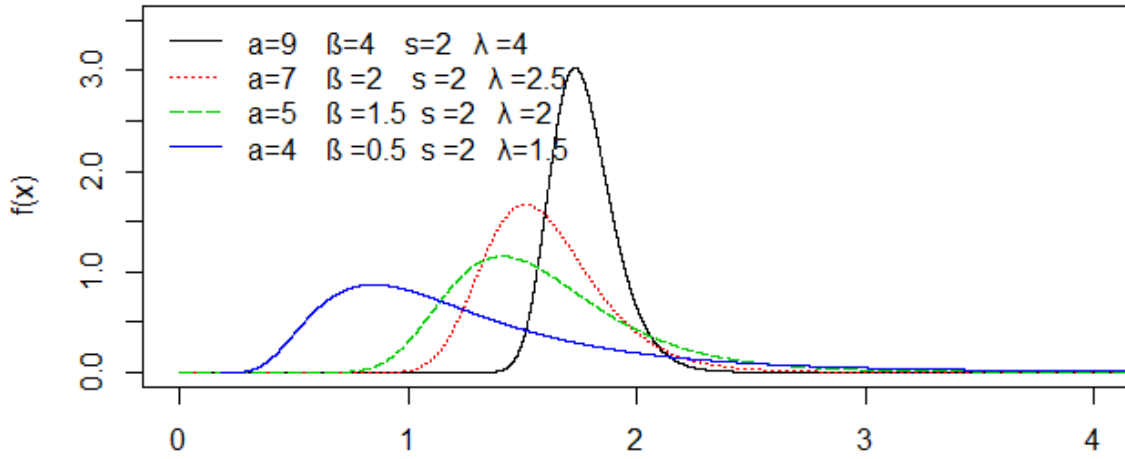


Figure 3.2: Probability Density Function for EGFT

The hazard function of EGFT

$$h(x) = \frac{\alpha\beta\lambda\sigma^\lambda x^{-(\lambda+1)} \exp\{-(\sigma/x)^\lambda\} [1 - \exp\{-(\sigma/x)^\lambda\}]^{\alpha-1} (1 - [1 - \exp\{-(\sigma/x)^\lambda\}]^\alpha)^{\beta-1}}{1 - [1 - \{1 - \exp\{-(\sigma/x)^\lambda\}\}^\alpha]^\beta} \quad (3.6)$$

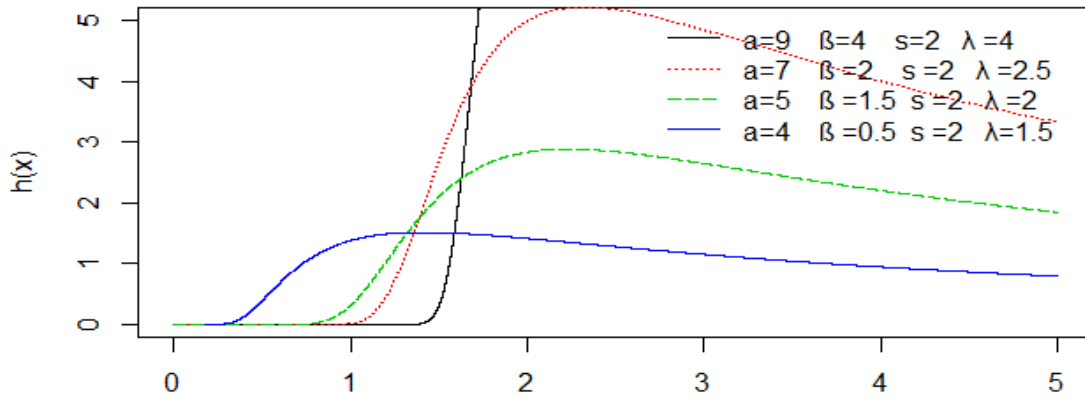


Figure 3.3: Hazard function of EGFT

3.2.2 Exponentiated Generalized Inverted Exponential (EGIE)

The PDF and CDF of the Inverted Exponential (IE) distribution were defined by Oguntunde *et al.*, (2014) as follows:

$$g(x) = \frac{\lambda}{x^2} \exp\left(-\frac{\lambda}{x}\right)$$

and

$$G(x) = 1 - \exp\left(-\frac{\lambda}{x}\right)$$

where $x > 0$, the scale parameter $\lambda > 0$. Therefore, the EGIE distribution is obtained by substituting CDF of IE distribution in equation (3.1) to obtain:

$$F(x) = [1 - \{1 - \exp -(\lambda/x)\}^\alpha]^\beta \quad (3.7)$$

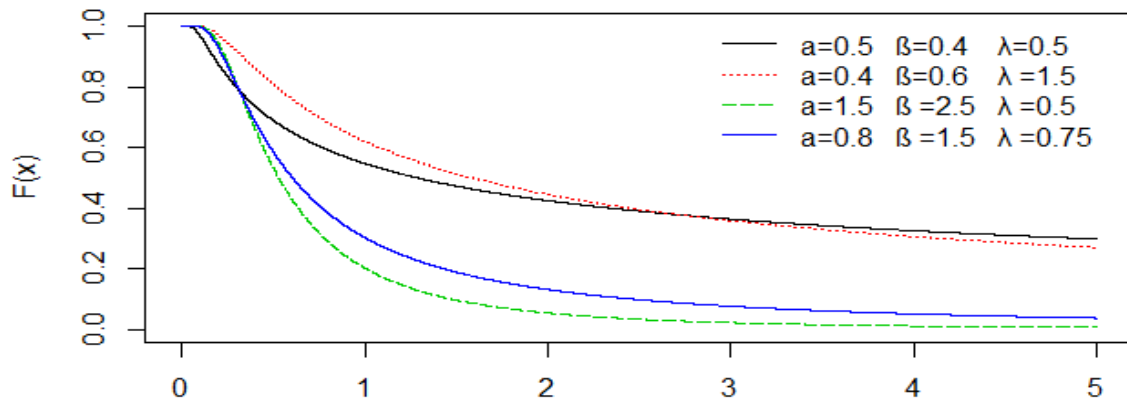


Figure 3.4: Cumulative Distribution Function of EGIE

The equivalent PDF is

$$f(x) = \alpha\beta\lambda x^2 \exp -(\lambda/x) \{\exp -(\lambda/x)\}^{\alpha-1} [1 - \{\exp -(\frac{\lambda}{x})\}^\alpha]^{\beta-1} \quad (3.8)$$

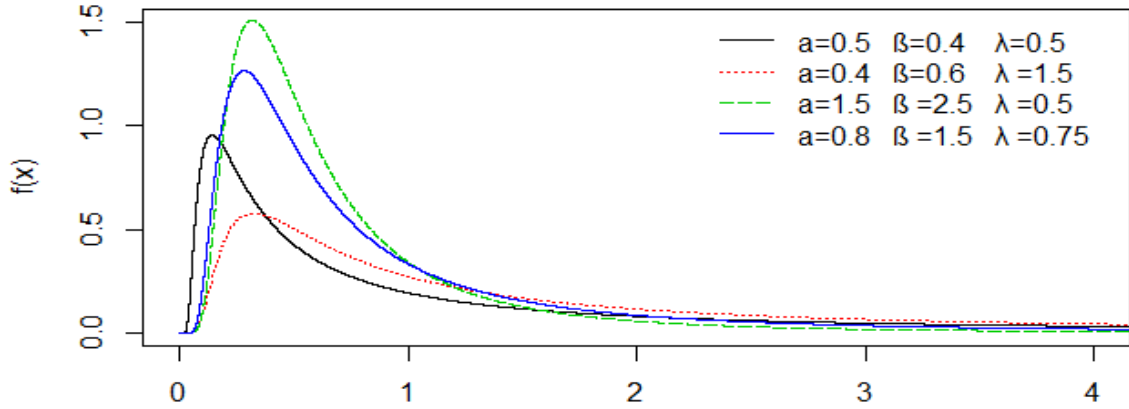


Figure 3.5: Probability Density Function of EGIE

where $x > 0, \alpha = a > 0, \beta > 0, \lambda > 0$

Hazard function of EGIE is expressed as;

$$h(x) = \frac{\alpha\beta\lambda x^2 \exp -(\lambda/x) \{\exp -(\lambda/x)\}^{\alpha-1} [1 - \{\exp -(\frac{\lambda}{x})\}^\alpha]^{\beta-1}}{1 - [1 - \{\exp -(\lambda/x)\}^\alpha]^\beta} \quad (3.9)$$

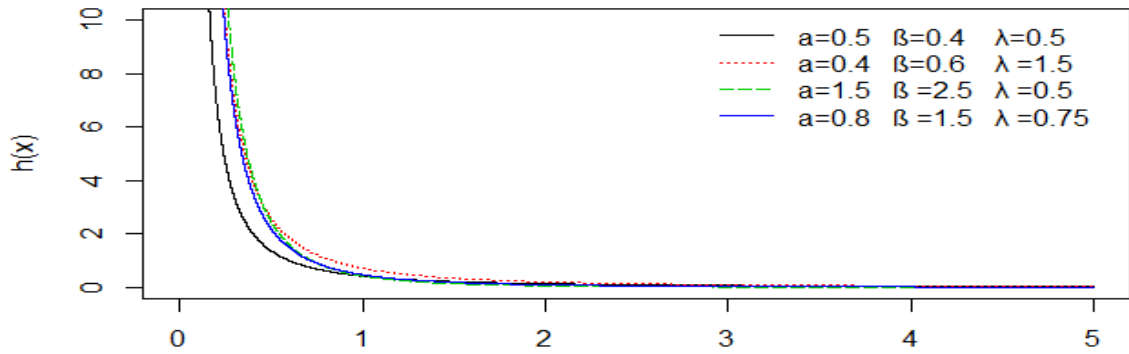


Figure 3.6: Hazard function of EGIE

3.3 Proposed Distribution: Exponentiated Generalized Exponentiated Exponential (EGEE)

The CDF and PDF of EE were defined by Gupta and Kundu (1998) in equations (3.10) and (3.11) respectively:

$$G(x) = (1 - e^{-\frac{x}{\theta}})^k \quad (3.10)$$

$$g(x, \theta, k) = k/\theta(1 - e^{-\frac{x}{\theta}})^{k-1}e^{-\frac{x}{\theta}} \quad (3.11)$$

the (EGEE) cumulative distribution from (3.1) is defined as;

$$F(x) = [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^\beta \quad (3.12)$$

By inserting (3.10) in (3.1) and the corresponding p.d.f from (3.1) is

$$f(x) = \frac{\alpha\beta k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^{\beta-1} \quad (3.13)$$

where $x > 0, \alpha > 0, \beta > 0, k > 0$ and $\theta > 0$

3.3.1 Proof of Proposed EGEE Distribution

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

$$\int_0^\infty \frac{\alpha\beta k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^{\beta-1} dx = 1 \quad (3.14)$$

$$\text{let } u = 1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha$$

$$\text{then } du = \left[-\alpha \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} \right] \times \left[-k(1 - e^{-\frac{x}{\theta}})^{k-1} \right] \times \left[-\left(-\frac{1}{\theta} e^{-\frac{x}{\theta}}\right) \right] dx$$

$$du = \left[-\alpha \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} \right] \times \left[-k(1 - e^{-\frac{x}{\theta}})^{k-1} \right] \times \left[\frac{1}{\theta} e^{-\frac{x}{\theta}} \right] dx$$

$$du = \frac{\alpha k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} dx$$

$$dx = \left(\frac{\alpha k}{\theta} \right)^{-1} \left[e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} \right]^{-1} du$$

Substitute for dx in 3.14

$$\int_0^\infty \frac{\alpha\beta k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^{\beta-1} \times \left(\frac{\alpha k}{\theta}\right)^{-1} \left[e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1}\right]^{-1} du = 1$$

$$\int_0^\infty \beta [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^{\beta-1} du$$

$$\text{Recall } u = 1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha$$

$$\int_0^\infty \beta [u]^{\beta-1} du$$

$$\beta \left[\frac{[u]^{\beta-1+1}}{\beta-1+1} \right]_0^\infty$$

$$\beta \left[\frac{[u]^\beta}{\beta} \right]_0^\infty$$

$$[u]_0^\infty$$

$$\left[1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha \right]_0^\infty$$

$$\left[1 - \{1 - (1 - e^{-\frac{\infty}{\theta}})^k\}^\alpha \right] - \left[1 - \{1 - (1 - e^{-\frac{0}{\theta}})^k\}^\alpha \right]$$

$$[1 - \{1 - (1 - 0)^k\}^\alpha] - [1 - \{1 - (1 - 1)^k\}^\alpha] = [1 - \{0\}^\alpha] - [1 - \{1\}^\alpha] = 1$$

Therefore the p.d.f is a true p.d.f.

3.3.1.1 Sub-Distribution of the Proposed Distribution

Some well-known existing distributions are Special cases of EGEE, if $\alpha = \beta = 1$, the distribution decreases to EE distribution developed by Kundu and Gupta, (1998) with CDF and PDF in equation (3.10) and (3.11) respectively as;

$$F(x) = G(x)$$

$$F(x) = (1 - e^{-\frac{x}{\theta}})^k$$

$$f(x) = k/\theta(1 - e^{-\frac{x}{\theta}})^{k-1}e^{-\frac{x}{\theta}}$$

which is the baseline distribution.

For $k = 1$ distribution reduced to EGE distribution developed by Oguntunde *et al.*, (2016), whose CDF and PDF are derived as;

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

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

For $\beta = 1$ the distribution decreases to GE distribution whose CDF and PDF are expressed respectively as;

$$F(x) = 1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha$$

$$f(x) = \frac{\alpha k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1}$$

For $\alpha = 1$ the distribution reduced to 3 parameters Exponentiated Exponential (3EE) distribution with CDF and PDF expressed respectively as;

$$F(x) = (1 - e^{-\frac{x}{\theta}})^{\beta k}$$

$$f(x) = k\beta/\theta(1 - e^{-\frac{x}{\theta}})^{\beta k-1}e^{-\frac{x}{\theta}}$$

Lastly, for $\alpha = \beta = k = 1$ the distribution reduces to Exponential (E) distribution CDF and PDF as presented below:

$$F(x) = 1 - e^{-\frac{x}{\theta}}$$

$$f(x) = 1/\theta e^{-\frac{x}{\theta}}$$

3.3.2 The Hazard Function

The function of measuring the lowest or highest chance of an event surviving a certain time based on its past survival time t is called hazard function. By definition, the hazard function or failure rate of a random variable X with PDF $f(x)$ and a cumulative distribution function $F(x)$ is given by;

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

where $F(x)$ and $f(x)$ are given by equation (3.12) and (3.13), respectively. Using equations (3.12) and (3.13) the hazard function of EGEE is stated as follows;

$$h(x) = \frac{\frac{\alpha\beta k}{\theta} e^{-x/\theta} (1 - e^{-x/\theta})^{k-1} \{1 - (1 - e^{-x/\theta})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^{\beta-1}}{1 - [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^\beta} \quad (3.15)$$

3.3.3 The Survival Function

The survival function determines the probability that a device, patient or other objects will survive beyond any time t and is given as;

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

It implies that $s(x)$ is

$$s(x) = 1 - [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^\beta \quad (3.16)$$

3.3.4 The Reversed Hazard Function

The reversed hazard function in recent time has drawn the attention of the researcher by providing additional information to that of hazard function with an interesting and useful feature, especially in reliability analysis. It is defined as the ratio of pdf over CDF and expressed as:

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

It implies that $r(x)$ for EGEE distribution is

$$r(x) = \frac{\frac{\alpha\beta k}{\theta} e^{-x/\theta} (1 - e^{-x/\theta})^{k-1} \{1 - (1 - e^{-x/\theta})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^{\beta-1}}{[1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^\beta} \quad (3.17)$$

3.3.5 The Odds Function.

Odds function of EGEE distribution can be obtained by dividing CDF $F(x)$ with the survival function $s(x)$

$$O(x) = \frac{F(x)}{s(x)}$$

Its implies that $O(x)$

$$O(x) = \frac{[1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^\beta}{1 - [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^\beta} \quad (3.18)$$

3.3.6 Special Cases of EGEE Distribution.

- The E distribution is a special case of EGEE when $\alpha = \beta = k = 1$.
- When $\alpha = \beta = 1$, the EGEE gives an EE distribution
- When $k = 1$, the EGEE gives a member of EGF which EGE distribution.

The PDF, CDF, hazard, survival, reverse-hazard, and odd functions plots of EGEE distribution along with its sub-distribution are presented in the following figures

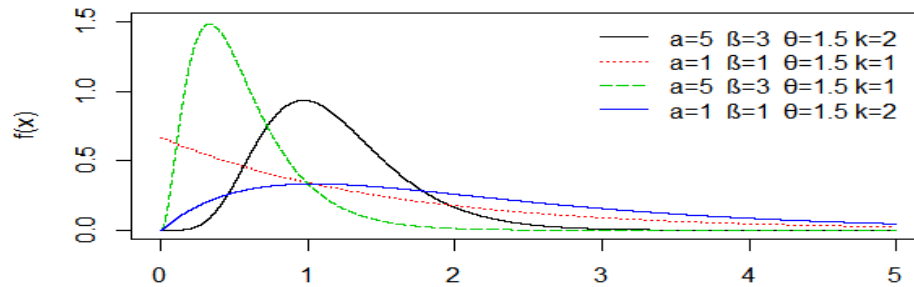


Figure 3.7: PDF of EGEE, EE, EGE and E

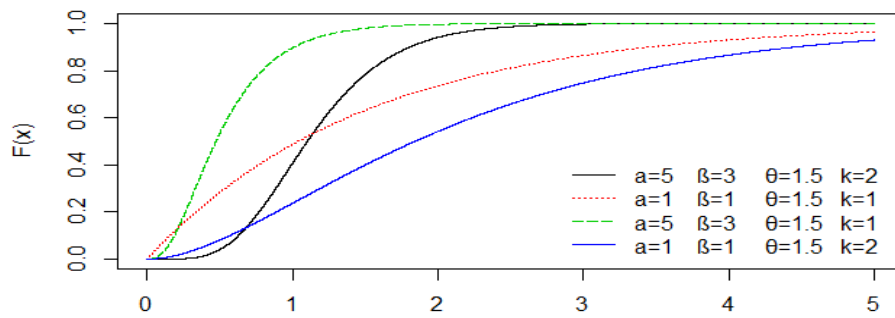


Figure 3.8: CDF of EGEE, EE, EGE, and E

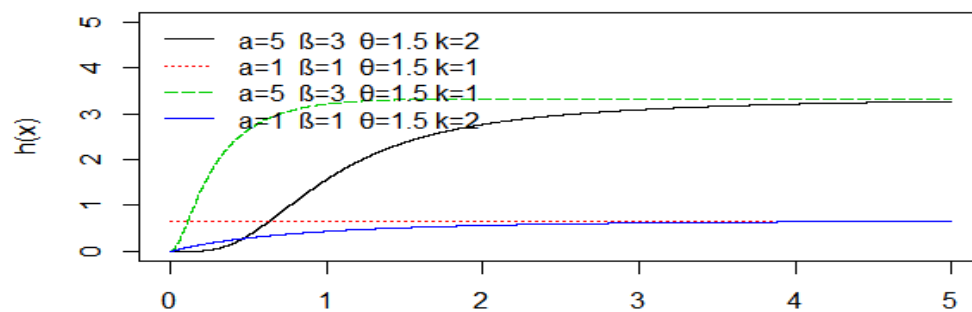


Figure 3.9: Hazard Function of EGEE, EE, EGE, and E

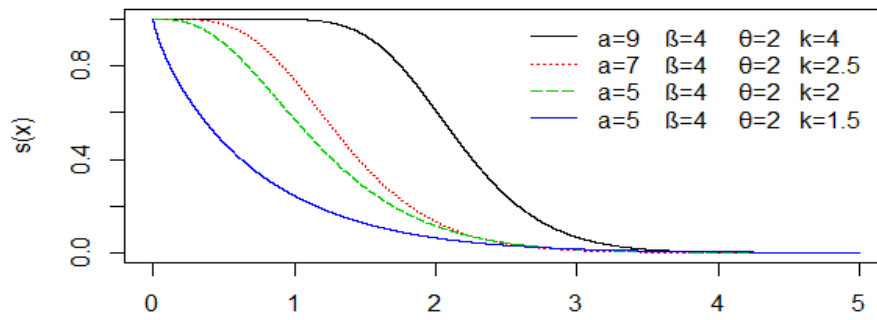


Figure 3.10: Survival Function of EGEE, EE, EGE, and E

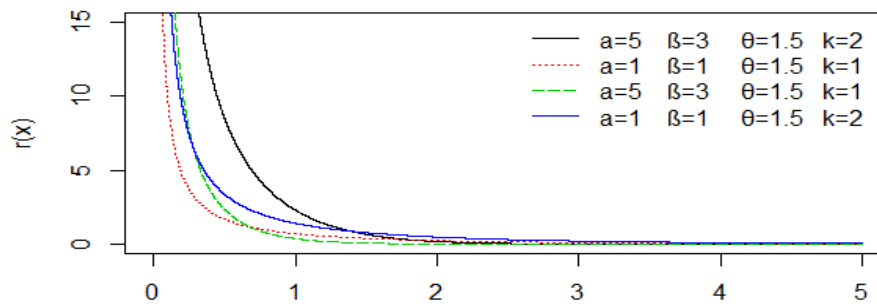


Figure 3.11: Reversed Hazard Function of EGEE, EE, EGE, and E

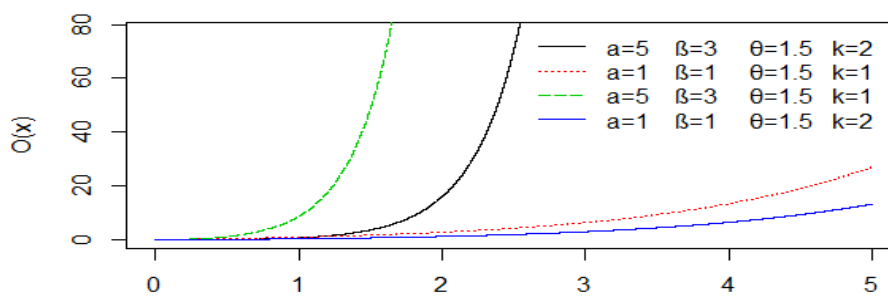


Figure 3.12: Odds Function of EGEE, EE, EGE, and E

Table 3.1**Summary of EGEE and Sub-Distributions**

Distribution	β	α	k	θ
EGEE	β	α	k	θ
EGE	β	α	1	θ
EE	1	1	k	θ
E	1	1	1	θ
GE or EG	1	α	k	θ
3EE	β	1	k	θ

Table 3.1 shows the possible distributions that can be obtained from the proposed distribution. The occurrence of all unknown parameters gives the proposed distribution EGEE while nonexistence of a shape parameter resulted to the EGE distribution, the absence of two shape parameters leads to EG distribution while the absence of all the three shape parameters gives an exponential (E) distribution.

3.4 Properties of Exponentiated Generalized Exponentiated Exponential (EGEE) Distribution

The CDF and PDF of EE are stated respectively in (3.10) and (3.11) as.

$$G(x) = (1 - e^{-\frac{x}{\theta}})^k$$

$$g(x, \theta, k) = \frac{k}{\theta} (1 - e^{-\frac{x}{\theta}})^{k-1} e^{-\frac{x}{\theta}}$$

The properties of the proposed distribution are derived from the EGF distribution in (3.1) and (3.2).

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

$$\alpha > 0, \beta > 0, x \in \mathbb{R}$$

$$f(x) = \alpha\beta\{1 - G(x)\}^{\alpha-1}[1 - \{1 - G(x)\}^\alpha]^{\beta-1}g(x)$$

By expanding the density function, we have

$$(1 - z)^{\beta-1} = \sum_{\lambda=0}^{\infty} \frac{(-1)^\lambda \Gamma(\beta)}{\Gamma(\beta - \lambda) \lambda!} z^\lambda \quad |z| < 1, \beta \in \mathbb{R} \quad (3.19)$$

Hence, using related expansion on (3.10)

$$F(x) = \sum_{\lambda=0}^{\infty} \frac{(-1)^\lambda \Gamma(\beta + 1)}{\Gamma(\beta - \lambda + 1) \lambda!} \{1 - G(x)\}^{\alpha\lambda}$$

considering

$$\{1 - G(x)\}^{\alpha\lambda} = \sum_{j=0}^{\infty} \frac{(-1)^j \Gamma(\alpha\lambda + 1)}{\Gamma(\alpha\lambda - j + 1) j!} G(x)^j$$

$$F(x) = \sum_{j=0}^{\infty} \sum_{\lambda=0}^{\infty} \frac{(-1)^{\lambda+j} \Gamma(\beta + 1) \Gamma(\alpha\lambda + 1)}{\Gamma(\beta - \lambda + 1) \Gamma(\alpha\lambda - j + 1) j! \lambda!} G(x)^j \quad (3.20)$$

$$F(x) = \sum_{j=0}^{\infty} w_j G(x)^j \quad (3.21)$$

where $w_j = w_j(\alpha, \beta) = \sum_{\lambda=0}^{\infty} \frac{(-1)^{\lambda+j} \Gamma(\beta+1) \Gamma(\alpha\lambda+1)}{\Gamma(\beta-\lambda+1) \Gamma(\alpha\lambda-j+1) j! \lambda!}$

Differentiating equation (3.20) with respect to x gives the pdf

$$f(x) = \sum_{j=0}^{\infty} j w_j g(x) (G(x))^{j-1} \quad (3.22)$$

Putting equation (3.10) and (3.11) into equation (3.22)

$$f(x) = \sum_{j=0}^{\infty} j w_j \frac{k}{\theta} (1 - e^{-\frac{x}{\theta}})^{k-1} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k(j-1)}$$

$$f(x) = \sum_{j=0}^{\infty} j w_j \frac{k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{kj-k+k-1}$$

$$f(x) = \frac{k}{\theta} \sum_{j=0}^{\infty} j w_j e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{kj-1} \quad (3.23)$$

3.4.1 Quantile and Median of EGEE

Quantile functions are used in theoretical aspects of probability theory, statistics applications, and simulation. Simulation methods utilize quantile function to produce simulated random variables for classical and new continuous distributions. Therefore, the quantile function for EGF distribution is presented in (3.3) as $x = Q_G([1 - (1 - u^{1/\beta})^{1/\alpha}])$. Hence, the quantile function of EGEE distribution is expressed as $Q(p) = F^{-1}(p)$

$$p = [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^\beta$$

After simplification, the form the quartile function is;

$$x = -\theta \left(\ln \left[1 - \left(1 - \left(1 - (p)^{\frac{1}{\beta}} \right)^{\frac{1}{\alpha}} \right)^{\frac{1}{k}} \right] \right) \quad (3.24)$$

when $p = 0.75, 0.5$ and 0.25 , the third quantile, median, and first quartile respectively of the proposed -distribution (EGEE) are derived.

Specifically, $Q(0.5) = F^{-1}(u)$ is the median of the probability distribution expressed as

$$F(x) = p_r(X \leq m) = \int_0^m f(x) dx = 0.5 \quad (3.25)$$

equating equation (3.12) to 0.5 gives the median of EGEE Distribution as follows

$$F(x) = [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^\beta$$

$$[1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^\alpha]^\beta = 0.5$$

$$Median = -\theta \left(\ln \left[1 - \left(1 - \left(1 - (0.5)^{\frac{1}{\beta}} \right)^{\frac{1}{\alpha}} \right)^{\frac{1}{k}} \right] \right) \quad (3.26)$$

3.4.2 The Moment of the Proposed (EGEE) Distribution

The r th moment of random variable X for the proposed (EGEE) is derived as follow

$$\mu^r = \int_0^{\infty} x^r f(x) dx \quad (3.27)$$

Using the expansion of equation (3.23), the r th moment of EGEE is expressed as

$$\mu^r = \int_0^{\infty} x^r \frac{k}{\theta} \sum_{j=0}^{\infty} j w_j e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{kj-1} dx$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} j w_j \int_0^{\infty} x^r e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{kj-1} dx$$

$$(1 - e^{-\frac{x}{\theta}})^{kj-1} = \sum_{m=0}^{\infty} (-1)^m \binom{kj-1}{m} e^{-\frac{xm}{\theta}} = \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} e^{-\frac{xm}{\theta}}$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} j w_j \int_0^{\infty} x^r e^{-\frac{x}{\theta}} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} e^{-\frac{xm}{\theta}} dx$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \int_0^{\infty} x^r e^{-\frac{x}{\theta}} e^{-\frac{xm}{\theta}} dx$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \int_0^{\infty} x^r e^{-x(\frac{m+1}{\theta})} dx$$

$$let y = x \left(\frac{m+1}{\theta} \right)$$

$$x = \frac{\theta y}{m+1}$$

$$\frac{dx}{dy} = \frac{\theta}{m+1}$$

$$dx = \frac{\theta dy}{m+1}$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \int_0^{\infty} \left(\frac{\theta y}{m+1} \right)^r e^{-y} \frac{\theta dy}{m+1}$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \int_0^{\infty} \left(\frac{\theta}{m+1} \right)^r y^r e^{-y} \frac{\theta dy}{m+1}$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \left(\frac{\theta}{m+1} \right)^{r+1} \int_0^{\infty} y^r e^{-y} dy$$

$$\int_0^{\infty} y^r e^{-y} dy = \Gamma(r+1)$$

$$\mu^r = \frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \left(\frac{\theta}{m+1} \right)^{r+1} \Gamma(r+1)$$

$$\mu^r = \frac{k\theta^r}{(m+1)^{r+1}} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \Gamma(r+1) \quad (3.28)$$

The two infinite series in (3.28) is convergent for every $\beta > 0, \alpha > 0, \theta > 0, k > 0$ and $x > 0$ (Nasiru *et al.*, 2017)

The mean is obtained by making μ^r moment equal to one ($r = 1$)

$$\mu_j^1 = \frac{k\theta^1}{(m+1)^{1+1}} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \Gamma(1+1)$$

$$\mu_j^1 = \frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \Gamma(2)$$

$$\mu_j^1 = \frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \quad (3.29)$$

when $r = 2$

$$\mu_j^2 = \frac{k\theta^2}{(m+1)^{2+1}} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \Gamma(2+1)$$

$$\mu_j^2 = \frac{k\theta^2}{(m+1)^3} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \Gamma(3)$$

$$\mu_j^2 = \frac{2k\theta^2}{(m+1)^3} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} j w_j \quad (3.30)$$

The variance of is respectively derived from (3.29) and (3.30):

$$\sigma^2 = \mu_j^2 - (\mu_j^1)^2 \quad (3.31)$$

$$\begin{aligned}
& \frac{2k\theta^2}{(m+1)^3} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j - \left(\frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j \right)^2 \\
& \frac{2k\theta^2}{(m+1)^3} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j - \frac{(k\theta)^2}{(m+1)^4} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j \\
& \frac{k\theta^2}{(m+1)^3} \left[2 \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j - \frac{k}{(m+1)} \left(\sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j \right)^2 \right] \\
\sigma^2 &= \frac{k\theta^2}{(m+1)^3} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j \left(2 - \frac{k}{(m+1)} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m)m!} jw_j \right) \quad (3.32)
\end{aligned}$$

3.4.3 The Moment Generating Function of Proposed EGEE Distribution

$$M_x(t) = E(e^{tx})$$

$$E(e^{tx}) = \int_0^{\infty} e^{tx} f(x) dx \quad (3.33)$$

$$\int_0^{\infty} e^{tx} \frac{k}{\theta} \sum_{j=0}^{\infty} jw_j e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{kj-1} dx$$

Using binomial theorem,

$$(1 - e^{-\frac{x}{\theta}})^{kj-1} = \sum_{m=0}^{\infty} (-1)^m \binom{kj-1}{m} e^{-\frac{xm}{\theta}}$$

$$\int_0^{\infty} e^{tx} \frac{k}{\theta} \sum_{j=0}^{\infty} jw_j \sum_{m=0}^{\infty} (-1)^m e^{-\frac{x}{\theta}} \binom{kj-1}{m} e^{-\frac{xm}{\theta}} dx$$

$$\sum_{j=0}^{\infty} j \mathbf{w}_j \sum_{m=0}^{\infty} (-1)^m \int_0^{\infty} \frac{k}{\theta} \binom{kj-1}{m} e^{-\frac{xm}{\theta} - \frac{x}{\theta} + tx} d\mathbf{x}$$

$$\frac{k}{\theta} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} (-1)^m j \mathbf{w}_j \int_0^{\infty} \binom{kj-1}{m} e^{-x \left(\frac{m+1-\theta t}{\theta} \right)} d\mathbf{x}$$

$$\text{let } y = x \left(\frac{m+1-\theta t}{\theta} \right)$$

$$x = \frac{y\theta}{m+1-\theta t}$$

$$dx = \frac{\theta dy}{m+1-\theta t}$$

$$\frac{k}{\theta} \frac{\theta}{(m+1-\theta t)} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} (-1)^m \binom{kj-1}{m} j \mathbf{w}_j \int_0^{\infty} e^{-y} dy$$

$$M_x(t) = \frac{k}{(m+1-\theta t)} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} (-1)^m \binom{kj-1}{m} j \mathbf{w}_j \quad (3.34)$$

To get the first moment from the mgf we differentiate (3.34) with respect to t .

$$\boldsymbol{\mu}_j^1 = M'_x(t)$$

$$M'_x(t) = \frac{\theta k}{(m+1-\theta t)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} (-1)^m \binom{kj-1}{m} j \mathbf{w}_j$$

$$M'_x(0) = \frac{\theta k}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} (-1)^m \binom{kj-1}{m} j \mathbf{w}_j$$

3.4.4 Order Statistics

Order statistics appears in various fields of statistics, such as reliability and life testing situation. Let $Y_1, Y_2, \dots, Y_n$ be independently and identically distributed (i.i.d) random variable with their corresponding cumulative function (CDF) $F(x)$. Let $X_{(1)}$ be the smallest (first order-statistic) of $Y_1, Y_2, \dots, Y_n$, $X_{(2)}$ is the second order-statistics larger than $X_{(1)}$ and $X_{(n)}$ is the greatest order-statistics. Thus, the order-statistics $X_{(1)} < X_{(2)} \dots < X_{(n)}$ corresponds to random variables $Y_1, Y_2, \dots, Y_n$ (Kapur and Sexena, 1960).

The pdf $f_{n:i}(x)$ of the i th order-statistics, for $i = 1, \dots, n$, from iid random variable $Y_1 \dots Y_n$ is expressed as

$$f_{n:i}(x) = \frac{f(x)}{B(i, n-i+1)} F(x)^{i-1} (1-F(x))^{n-i} \quad 3.35$$

where $B(i, n-i+1)$ is the Beta function.

3.4.4.1 Order-Statistics of EGF distribution

Substitute p.d.f. and c.d.f. of EGF distribution in equation 3.35 produces

$$\begin{aligned} f_{n:i}(x) &= \frac{f(x)}{B(i, n-i+1)} F(x)^{i-1} (1-F(x))^{n-i} \\ f_{n:i}(x) &= \frac{\alpha \beta g(x) (1-G(x))^{\alpha-1} (1-(1-G(x))^\alpha)^{\beta-1}}{B(i, n-i+1)} (1-(1-G(x))^\alpha)^{\beta(i-1)} \times (1-(1-(1-G(x))^\alpha)^\beta)^{n-i} \\ f_{n:i}(x) &= \frac{\alpha \beta g(x) (1-G(x))^{\alpha-1}}{B(i, n-i+1)} (1-(1-G(x))^\alpha)^{\beta(i-1)} \times (1-(1-(1-G(x))^\alpha)^\beta)^{n-i} \end{aligned}$$

From binomial theorem, it becomes

$$(1 - (1 - (1 - G(x))^\alpha)^\beta)^{n-i} = \sum_q^{n-i} (-1)^q \binom{n-i}{q} (1 - (1 - G(x))^\alpha)^{\beta q}$$

$$\frac{\alpha \beta g(x) (1 - G(x))^{\alpha-1}}{B(i, n-i+1)} \sum_q^{n-i} (-1)^q \binom{n-i}{q} (1 - (1 - G(x))^\alpha)^{\beta q} (1 - (1 - G(x))^\alpha)^{\beta i-1}$$

$$f_{n:i}(x) = \frac{\alpha \beta g(x) (1 - G(x))^{\alpha-1}}{B(i, n-i+1)} \sum_q^{n-i} (-1)^q \binom{n-i}{q} (1 - (1 - G(x))^\alpha)^{\beta(q+i)-1}$$

From binomial theorem, it becomes

$$(1 - (1 - G(x))^\alpha)^{\beta(q+i)-1} = \sum_p^\infty (-1)^p \binom{\beta(q+i)-1}{p} (1 - G(x))^{\alpha p}$$

$$f_{n:i}(x) = \frac{\alpha \beta g(x)}{B(i, n-i+1)} \sum_q^{n-i} \sum_p^\infty (-1)^{q+p} \binom{n-i}{q} \binom{\beta(q+i)-1}{p} (1 - G(x))^{\alpha(p+1)-1}$$

From binomial theorem, it becomes $(1 - G(x))^{\alpha(p+1)-1} = \sum_l^\infty (-1)^l \binom{\alpha(p+1)-1}{l} G(x)^l$

$$f_{n:i}(x) = \frac{\alpha \beta g(x)}{B(i, n-i+1)} \sum_{q=0}^{n-i} \sum_{p=0}^\infty \sum_{l=0}^\infty (-1)^{q+p+l} \binom{n-i}{q} \binom{\beta(q+i)-1}{p} \binom{\alpha(p+1)-1}{l} G(x)^l$$

$$\text{Let } s_l = \sum_q^{n-i} \sum_p^\infty (-1)^{q+p+l} \binom{n-i}{q} \binom{\beta(q+i)-1}{p} \binom{\alpha(p+1)-1}{l}$$

$$\frac{\alpha \beta}{B(i, n-i+1)} \sum_{l=0}^\infty s_l g(x) G(x)^l \quad (3.36)$$

3.4.4.2 Order-Statistics of EGEE

Order-statistics of EGEE distribution can be derived by substituting the CDF and PDF of EE into (3.36)

$$\frac{\alpha\beta}{B(i, n-i+1)} \sum_{l=0}^{\infty} s_l \frac{k}{\theta} e^{-\frac{x}{\theta}} \left(1 - e^{-\frac{x}{\theta}}\right)^{k-1} \left(\left(1 - e^{-\frac{x}{\theta}}\right)^{(k-1)}\right)^l$$

$$\frac{\alpha\beta}{B(i, n-i+1)} \sum_{l=0}^{\infty} s_l \frac{k}{\theta} e^{-\frac{x}{\theta}} \left(1 - e^{-\frac{x}{\theta}}\right)^{k-l+kl-1}$$

From binomial theorem, it becomes

$$\left(1 - e^{-\frac{x}{\theta}}\right)^{k-l+kl-1} = \sum_{d=0}^{\infty} (-1)^d \binom{k-l+kl-1}{d} e^{-\frac{xd}{\theta}}$$

$$\frac{\alpha\beta}{B(i, n-i+1)} \frac{k}{\theta} \sum_{l=0}^{\infty} \sum_{d=0}^{\infty} (-1)^d \binom{k-l+kl-1}{d} s_l e^{-x\left(\frac{d+1}{\theta}\right)}$$

$$\frac{\alpha\beta}{B(i, n-i+1)} \frac{k}{\theta} \sum_{l=0}^{\infty} \sum_{d=0}^{\infty} (-1)^d \frac{\Gamma(k(l+1)-l)}{\Gamma(k(l+1)-l-d)d!} s_l e^{-x\left(\frac{d+1}{\theta}\right)}$$

Therefore the order-statistics for EGEE distribution is given by

$$\begin{aligned} & \frac{\alpha\beta}{B(i, n-i+1)} \frac{k}{\theta} \\ & * \sum_{q=0}^{n-i} \sum_{p=0}^{\infty} \sum_{l=0}^{\infty} \sum_{d=0}^{\infty} (-1)^{p+q+l+d} \frac{\Gamma(n-i+1)\Gamma(\beta(q+i))\Gamma(\alpha(p+1))\Gamma(k(l+1)-l)}{\Gamma(n-i+q+1)\Gamma(\beta(q+i)-p)\Gamma(\alpha(p+1)-l)\Gamma(k(l+1)-l-d)d!q!p!l!} e^{-x\left(\frac{d+1}{\theta}\right)} \end{aligned} \quad (3.37)$$

The four infinite series in (3.37) is convergent for $\beta > 0, \alpha > 0, \theta > 0, k > 0$ and $x > 0$

3.4.5 Skewness and Kurtosis of the EGEE Distribution

Two methods are used to estimate the Skewness and Kurtosis of the EGEE distribution.

These methods for the measures of both Skewness (SK) and Kurtosis (KU) are moments and quantiles based methods. In the moments based approach,

$$S.K = \frac{\mu'_3 - 3\mu\mu'_2 + 2\mu^2}{(\mu'_2 - \mu^2)^{3/2}} \quad (3.38)$$

and
$$K.U = \frac{\mu'_4 - 4\mu\mu'_3 + 6\mu^2\mu'_2 - 3\mu^4}{(\mu'_2 - \mu^2)^2} \quad (3.39)$$

Galton (1883) and Moore (1988) both proposed some quantile approaches for obtaining skeweness (SK) and kurtosis (KU) of a distribution that its quantile function exists in closed or simple analytical form. The expressions for SK and KU are expressed using the following relation respectively.

$$S.K = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)} \quad (3.40)$$

$$K.U = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)} \quad (3.41)$$

Since (3.24) of the proposed EGEE distribution exists in simple analytical form, then (3.40) and (3.41) can be used to obtain the SK and KU of the proposed EGEE distribution.

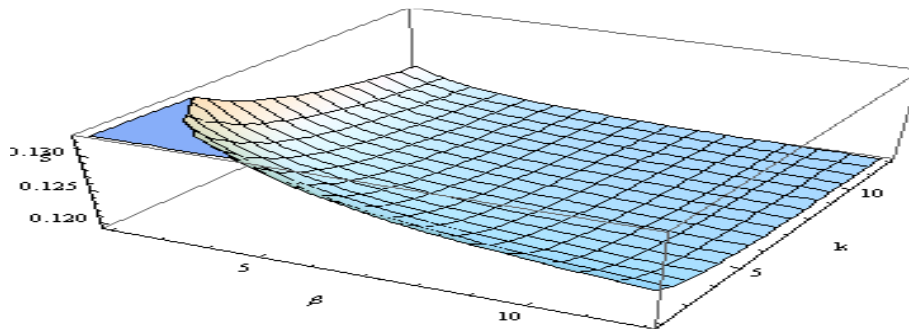


Figure 3.13 Plot for EGEE Skewness

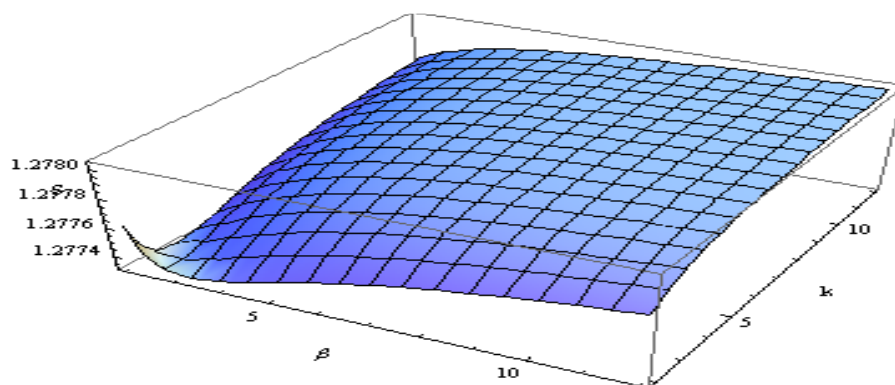


Figure 3.14 Plot for EGEE Kurtosis

The quantile function of EGEE distribution was used in plotting Skweness and Kurtosis of 3D plots in figures 3.13 and 3.34, for $\alpha = \theta = 1$ while $\beta = k$ takes the values from 2 to 12.

3.4.6 The Mean and Median Deviation of EGF

The deviation from the mean (for symmetric distributions) or deviation from the median (in the case of skewed distributions) can be use as a measure of spread in the population. Let X be an EGEE random variable with mean $\mu = E(X)$ and median M . The mean deviation from both mean ($D(\mu)$) and median ($D(M)$) are respectively as express follows

$$\begin{aligned}
 D(\mu) &= E\{|X - \mu|\} = \int_{-\infty}^{\infty} |x - \mu| f_x dx \\
 &= \int_{-\infty}^{\mu} (\mu - x) f_x dx + \int_{\mu}^{\infty} (x - \mu) f_x dx \\
 &= 2\mu'_1 F_x(\mu) - 2 \int_{-\infty}^{\mu} x f_x(x) dx
 \end{aligned} \tag{3.42}$$

Where $F_x(\mu)$, μ'_1 and $\int_{-\infty}^{\mu} x f_x(x) dx$, are cumulative function, mean and incomplete moment respectively

$$\begin{aligned}
 D(M) &= E\{|X - M|\} = \int_{-\infty}^{\infty} |x - M| f_x dx \\
 &= \mu'_1 - 2 \int_{-\infty}^M x f_x(x) dx
 \end{aligned} \tag{3.43}$$

here M , μ'_1 and $\int_{-\infty}^M x f_x(x) dx$ are median, mean and incomplete moment respectively (Cordeiro et al., 2013).

3.4.6.1 The Mean and Median Deviation of EGEE Distribution

Equation (3.42) is used to obtain the mean deviation of the proposed EGEE distribution

$$D(\mu) = 2\mu'_1 F_x(\mu) - 2\mu'_1 \int_0^{\mu} x f_x(x) dx$$

$F_x(\mu)$ is obtained from equation (3.12) as

$$F(u) = [1 - \{1 - (1 - e^{-\mu/\theta})^k\}^\alpha]^\beta$$

while

$$\int_0^\mu x f_x(x) dx$$

is gotten from equation (3.28) as:

$$\frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m) m!} j w_j \gamma(2, \mu)$$

where $\gamma(2, \mu)$ is known as the lower incomplete gamma function.

$$\begin{aligned} & 2 \frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m) m!} j w_j [1 - \{1 - (1 - e^{-\mu/\theta})^k\}^\alpha]^\beta \\ & - 2 \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m) m!} j w_j \gamma(2, \mu) \end{aligned} \quad (3.44)$$

The median deviation of EGEE distribution is derived from (3.43)

$$\mu'_1 - 2 \int_{-\infty}^m x f_x(x) dx$$

where $\int_0^m x f_x(x) dx$ is gotten from (3.28) by replacing μ for m

$$\frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m) m!} j w_j \gamma(2, m)$$

M is the median of the EGEE distribution derived in equation (3.26)

$$\frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m) m!} j w_j - 2 \frac{k\theta}{(m+1)^2} \sum_{j=0}^{\infty} \sum_{m=0}^{\infty} \frac{(-1)^m \Gamma(kj)}{\Gamma(kj-m) m!} j w_j \gamma(2, m)$$

(3.45)

3.4.7 Asymptotic Behavior

The study of the behavior of the EGEE distribution was investigated using the expression in

(3.13), as $x \rightarrow 0$ and as $x \rightarrow \infty$. This involves considering $\lim_{x \rightarrow 0} f(x)$ and $\lim_{x \rightarrow \infty} f(x)$

$$\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} \left[\frac{\alpha\beta k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-x/\theta})^{k-1} \{1 - (1 - e^{-x/\theta})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^{\beta-1} \right] = 0$$

$$\left[\frac{\alpha\beta k}{\theta} e^{-0} (1 - e^0)^{k-1} \{1 - (1 - e^0)^k\}^{\alpha-1} [1 - \{1 - (1 - e^0)^k\}^\alpha]^{\beta-1} \right]$$

$$\left[\frac{\alpha\beta k}{\theta} (1 - 1)^{k-1} \{1 - (1 - 1)^k\}^{\alpha-1} [1 - \{1 - (1 - 1)^k\}^\alpha]^{\beta-1} \right]$$

$$\left[\frac{\alpha\beta k}{\theta} (0)^{k-1} \{1 - (0)^k\}^{\alpha-1} [1 - \{1 - (0)^k\}^\alpha]^{\beta-1} \right] = 0$$

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} \left[\frac{\alpha\beta k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-x/\theta})^{k-1} \{1 - (1 - e^{-x/\theta})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-x/\theta})^k\}^\alpha]^{\beta-1} \right] = 0$$

$$\left[\frac{\alpha\beta k}{\theta} e^{-\infty} (1 - e^{-\infty})^{k-1} \{1 - (1 - e^{-\infty})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-\infty})^k\}^\alpha]^{\beta-1} \right]$$

$$\left[\frac{\alpha\beta k}{\theta} 0 (1 - 0)^{k-1} \{1 - (1 - 0)^k\}^{\alpha-1} [1 - \{1 - (1 - 0)^k\}^\alpha]^{\beta-1} \right]$$

$$\left[\frac{\alpha\beta k}{\theta} 0 (1)^{k-1} \{1 - 1\}^{\alpha-1} [1 - \{1 - 1\}^\alpha]^{\beta-1} \right] = 0$$

The above result shows that EGEE distribution exists with mode (Oguntunde *et al.*, 2014).

3.4.8 Renyi Entropy of the Proposed EGEE

The entropy of random variable X is a measure of the change in an uncertainty. Several entropy measures have been studied and discussed in the literature, but the most popular of them is the Renyi entropy, (Renyi 1961). The Renyi entropy of random variable X with proposed EGEE probability density function is expressed as:

$$I_{R(\rho)} = \frac{1}{(1 - \rho)} \log \left(\int_0^\infty g(x)^\rho dx \right) \quad (3.47)$$

where $\rho > 0$ and $\rho \neq 1$

Inserting equation 3.23 into 3.47

$$\begin{aligned}
I_{R(\rho)} &= \frac{1}{(1-\rho)} \log \left(\int_0^\infty \left(\frac{k}{\theta} \sum_{j=0}^\infty j \mathbf{w}_j e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{kj-1} \right)^\rho dx \right) \\
I_{R(\rho)} &= \frac{1}{(1-\rho)} \log \left(\frac{k^\rho}{\theta^\rho} \sum_{j=0}^\infty (j \mathbf{w}_j)^\rho \int_0^\infty e^{-\frac{\rho x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{\rho(kj-1)} dx \right) \\
\int_0^\infty e^{-\frac{\rho x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{\rho(kj-1)} dx &= \sum_{b=0}^\infty \binom{\rho(kj-1)}{b} (-1)^b e^{-\frac{\rho x}{\theta}} e^{-\frac{xb}{\theta}} \\
&= \sum_{b=0}^\infty \binom{\rho(kj-1)}{b} (-1)^b e^{-x(\frac{\rho+b}{\theta})} \\
\int_0^\infty \sum_{b=0}^\infty \binom{\rho(kj-1)}{b} (-1)^b e^{-x(\frac{\rho+b}{\theta})} dx & \\
\text{let } y = x(\frac{\rho+b}{\theta}) & \\
dx = \frac{\theta dy}{\rho+b} & \\
\binom{\rho(kj-1)}{b} (-1)^b \frac{\theta}{\rho+b} \int_0^\infty e^{-y} dy & \\
\sum_{b=0}^\infty \binom{\rho(kj-1)}{b} (-1)^b \frac{\theta}{(\rho+b)} &
\end{aligned}$$

Substitute in the expression above

$$I_{R(\rho)} = \frac{1}{(1-\rho)} \log \left(\frac{k^\rho}{\theta^\rho} \sum_{j=0}^\infty (j \mathbf{w}_j)^\rho \binom{\rho(kj-1)}{b} (-1)^b \frac{\theta}{(\rho+b)} \right) \quad (3.48)$$

3.4.9 Method of Maximum Likelihood

In this section, the estimates of the parameters for the proposed EGEE distribution are derived using the Maximum Likelihood Estimates (MLE). The concept of the Maximum Likelihood Estimates (MLE) is to obtain the values of the unknown parameters that maximize the probability of the sample data (Mood *et al.*, 1974). It starts with writing a mathematical expression known as the likelihood function of the sample data. The likelihood of a data is the probability of obtaining that particular set of data given the chosen probability distribution. The expression contains the unknown parameters of the distribution. The ML estimate is the value of these unknown parameters that maximize the sample likelihood (Elgarhy, 2017).

For a random sample $X_1, X_2, \dots, X_n$ of size n , from a distribution with PDF $f(x, \theta)$, the likelihood function $L(\theta)$ is expressed as:

$$L(\theta, \underline{x}) = f(x_1, x_2, \dots, x_n; \theta), \quad \theta \in \Omega, \quad (3.49)$$

where Ω is the bound for the value of θ . Where $f(x_1, x_2, \dots, x_n; \theta)$ is the value of the joint PDF of the random variable $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$. Hence, the objective of the *ML* is to maximize the likelihood function (LF) with respect to θ . The value *ML* estimator of θ is the solution of equation $\frac{dL(\theta)}{d(\theta)} = 0$.

If the likelihood function contains p parameters, it implies, the likelihood function is expressed as :

$$L(\theta, \underline{x}) = f(x_1, x_2, \dots, x_n; \theta_1, \theta_2, \dots, \theta_p), \quad \theta_1, \theta_2, \dots, \theta_p \in \Omega,$$

The *ML* estimators of the unknown parameters $\theta_1, \theta_2, \dots, \theta_p$ is the value $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p$ in Ω which maximize $L(\theta_1, \theta_2, \dots, \theta_p)$. These values are obtained by simultaneously solving the following p equations (Mood *et al.*, (1974)):

$$\frac{\partial L(\theta_1, \theta_2, \dots, \theta_p)}{\partial(\theta_i)} = 0 \quad i = 1, 2, \dots, p$$

Since the maximum value $L(\hat{\theta})$ will occur at the same points as the maximum value $\ln L(\hat{\theta})$, it will be easier to use the logarithm of the likelihood function.

The major importance of the *ML* estimator is that the logarithm of the likelihood function is been utilized in constructing the Fisher information matrix or simply the information matrix, I , which is defined as follows:

Let $X_1, X_2 \dots X_n$ be mutually independent and identically distributed (i.i.d.) random variables having a pdf $f(x, \theta)$, where θ is a vector of parameters having p components. Then the $(ij)^{th}$ element of the information matrix I_{ij} is

$$I_{ij} = -E \left(\frac{\partial^2 \log L(\theta; X)}{\partial \theta_i \partial \theta_j} \right) \Big|_{\theta = \hat{\theta}} \quad (3.50)$$

It is known that the Fisher information matrix is required to obtain the variance-covariance matrix of the estimators. The inverse matrix I^{-1} , with the $(ij)^{th}$ element denoted by I^{ij} , is the variance - covariance matrix of the estimators $\hat{\theta}$'s, so that $Var(\hat{\theta}_i) = I^{ii}$ and $cov(\hat{\theta}_i, \hat{\theta}_j) = I^{ij}$. Finally, the asymptotic variance-covariance matrix of the estimators $\hat{\theta}$'s is expressed as

$$I^{-1} = \begin{pmatrix} Var(\hat{\theta}_1) & cov(\hat{\theta}_1, \hat{\theta}_2) & \dots & cov(\hat{\theta}_1, \hat{\theta}_p) \\ cov(\hat{\theta}_2, \hat{\theta}_1) & Var(\hat{\theta}_2) & \dots & cov(\hat{\theta}_2, \hat{\theta}_p) \\ \vdots & \vdots & \dots & \vdots \\ cov(\hat{\theta}_p, \hat{\theta}_1) & cov(\hat{\theta}_p, \hat{\theta}_2) & \dots & Var(\hat{\theta}_p) \end{pmatrix}$$

Where $cov(\hat{\theta}_i, \hat{\theta}_j)$ is the covariance of $\hat{\theta}_i$ and $\hat{\theta}_j$ while $Var(\hat{\theta}_i)$ is the variance of the estimator $\hat{\theta}_i$.

The form $\hat{\theta} \xrightarrow{d} N(\theta, I^{-1}(\theta))$ has a special interest in finding the confidence interval (CI) for the unknown parameters. A $(1 - \alpha)100\%$ confidence interval (CI) for θ_i is:

$$C.I(\theta_i) = \hat{\theta}_i \pm Z_{\frac{\alpha}{2}} \sqrt{I_{ii}^{-1}} \quad 3.51$$

where I_{ii}^{-1} is the i th diagonal of $I^{-1}(\theta)$. (Cohen (1965)).

3.4. 9.1 Maximum Likelihood of the proposed EGEE

The log-likelihood function of four parameters EGEE is expressed at $X_1 = x_1, X_2 = x_2 \dots X_n = x_n$. Consequently, the objective of the *ML* is to maximize likelihood function with respect to the unknown parameters related with proposed EGEE distribution α, β, k and θ is given by

$$\begin{aligned}
 L &= \sum_{i=1}^n \ln(fx) = \sum_{i=1}^n \ln \left(\frac{\alpha \beta k}{\theta} e^{-\frac{x}{\theta}} (1 - e^{-\frac{x}{\theta}})^{k-1} \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha-1} [1 - \{1 - (1 - e^{-\frac{x}{\theta}})^k\}^{\alpha}]^{\beta-1} \right) \\
 &\quad n \ln \alpha + n \ln \beta + n \ln k - n \ln \theta - \sum_{i=1}^n \frac{x_i}{\theta} + \\
 &\quad (k-1) \sum_{i=1}^n \ln \left(1 - e^{-\frac{x_i}{\theta}} \right) + (\alpha-1) \sum_{i=1}^n \ln \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\} + \\
 &\quad (\beta-1) \sum_{i=1}^n \ln [1 - \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha}] \\
 \frac{\partial L}{\partial \alpha} &= \frac{n}{\alpha} + \sum_{i=1}^n \ln \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\} \\
 &\quad + \frac{(\beta-1) \sum_{i=1}^n [1 - \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha}] \ln \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}}{[1 - \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha}]} \\
 \frac{\partial L}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^n \ln [1 - \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha}] \\
 \frac{\partial L}{\partial k} &= \frac{n}{k} + \sum_{i=1}^n \ln \left(1 - e^{-\frac{x_i}{\theta}} \right) - (\alpha-1) \sum_{i=1}^n \frac{(1 - e^{-\frac{x_i}{\theta}})^k \ln \left(1 - e^{-\frac{x_i}{\theta}} \right)}{\{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}} + \\
 &\quad \alpha(\beta-1) \sum_{i=1}^n \frac{\{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha-1} (1 - e^{-\frac{x_i}{\theta}})^k \ln \left(1 - e^{-\frac{x_i}{\theta}} \right)}{[1 - \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha}]}
 \end{aligned}$$

$$\begin{aligned} \frac{\partial L}{\partial \theta} = & \frac{-n}{\theta} + \sum_{i=1}^n \frac{x_i}{\theta^2} - (k+1) \sum_{i=1}^n \frac{x_i e^{-\frac{x_i}{\theta}}}{\theta^2 (1 - e^{-\frac{x_i}{\theta}})} + (\alpha-1) \sum_{i=1}^n \frac{k(1 - e^{-\frac{x_i}{\theta}})^{k-1} x_i e^{-\frac{x_i}{\theta}}}{\theta^2 (1 - \{1 - e^{-\frac{x_i}{\theta}}\}^k)} - \\ & (\beta-1) \sum_{i=1}^n \frac{\alpha k (1 - e^{-\frac{x_i}{\theta}})^{k-1} x_i e^{-\frac{x_i}{\theta}} \{1 - (1 - e^{-\frac{x_i}{\theta}})^k\}^{\alpha-1}}{\theta^2 \left(1 - \left\{1 - \left\{1 - e^{-\frac{x_i}{\theta}}\right\}^k\right\}^\alpha\right)} \end{aligned}$$

Solving the nonlinear system of the equation of $\frac{\partial L}{\partial \alpha} = 0, \frac{\partial L}{\partial \beta} = 0, \frac{\partial L}{\partial k} = 0, \text{ and } \frac{\partial L}{\partial \theta} = 0$ gives the ML estimates of the parameters $\alpha, \beta, k, \text{ and } \theta$, respectively. The observed 4×4 information matrix for the proposed EGEE is obtained through,

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \\ \hat{k} \\ \hat{\theta} \end{pmatrix} \sim N \left[\begin{pmatrix} \alpha \\ \beta \\ k \\ \theta \end{pmatrix}, \begin{pmatrix} \hat{V}_{\alpha\alpha} & \hat{V}_{\alpha\beta} & \hat{V}_{\alpha k} & \hat{V}_{\alpha\theta} \\ \hat{V}_{\beta\alpha} & \hat{V}_{\beta\beta} & \hat{V}_{\beta k} & \hat{V}_{\beta\theta} \\ \hat{V}_{k\alpha} & \hat{V}_{k\beta} & \hat{V}_{kk} & \hat{V}_{k\theta} \\ \hat{V}_{\theta\alpha} & \hat{V}_{\theta\beta} & \hat{V}_{\theta k} & \hat{V}_{\theta\theta} \end{pmatrix} \right]$$

$$V^{-1} = -E \begin{pmatrix} V_{\alpha\alpha} & V_{\alpha\beta} & V_{\alpha k} & V_{\alpha\theta} \\ V_{\beta\alpha} & V_{\beta\beta} & V_{\beta k} & V_{\beta\theta} \\ V_{k\alpha} & V_{k\beta} & V_{kk} & V_{k\theta} \\ V_{\theta\alpha} & V_{\theta\beta} & V_{\theta k} & V_{\theta\theta} \end{pmatrix}$$

where

$$\begin{aligned} V_{\alpha\alpha} &= \frac{\partial^2 L}{\partial \alpha^2}, & V_{\beta\beta} &= \frac{\partial^2 L}{\partial \beta^2}, & V_{kk} &= \frac{\partial^2 L}{\partial k^2}, & V_{\theta\theta} &= \frac{\partial^2 L}{\partial \theta^2} \\ V_{\alpha\beta} &= V_{\beta\alpha} = \frac{\partial^2 L}{\partial \alpha \partial \beta}, & V_{\alpha k} &= V_{k\alpha} = \frac{\partial^2 L}{\partial \alpha \partial k}, & V_{\alpha\theta} &= V_{\theta\alpha} = \frac{\partial^2 L}{\partial \alpha \partial \theta}, \\ V_{\beta k} &= V_{k\beta} = \frac{\partial^2 L}{\partial \beta \partial k}, & V_{\beta\theta} &= V_{\theta\beta} = \frac{\partial^2 L}{\partial \beta \partial \theta}, & V_{k\theta} &= V_{\theta k} = \frac{\partial^2 L}{\partial k \partial \theta} \end{aligned}$$

The solution to the above inverse dispersion matrix yields the asymptotic variance and covariance of the maximum likelihood estimators $\hat{\alpha}, \hat{k}, \hat{\beta}$, and $\hat{\theta}$. The confidence interval for $\alpha, \beta, k, \text{ and } \theta$ is expressed as:

$$\hat{\alpha} \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{V}_{\alpha\alpha}}, \quad \hat{\beta} \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{V}_{\beta\beta}}, \quad \hat{k} \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{V}_{kk}}, \quad \hat{\theta} \pm Z_{\frac{\alpha}{2}} \sqrt{\hat{V}_{\theta\theta}}$$

where $Z_{\frac{\alpha}{2}}$ is the α^{th} percentiles of the standard normal distribution.

3.5 Measure Statistics

The statistical measures used in comparing the proposed probability distribution and the existing sub-distribution are the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Likelihood Ratio Test (LRT).

3.5.1 The Akaike Information Criterion (AIC)

One of the widely accepted information criteria, the AIC was introduced by Akaike (1973). The idea is to select the distribution that minimizes the negative likelihood penalized by the number of parameters as specified by the expression below:

$$AIC = 2k - 2\ln L \quad 3.52$$

where k and L are numbers of parameters and likelihood function respectively of the fitted distribution.

3.5.2 The Bayesian Information Criterion (BIC)

The most commonly used information criterion is the BIC which was studied by Schwarz (1978). The BIC unlike the AIC is an estimate of the Bayes factor for two competing distributions. The BIC is expressed as:

$$BIC = k\ln(n) - 2\ln L \quad (3.53)$$

where n in (3.53) is known as the random sample size.

Specifically, AIC is aimed at obtaining the best-approximating distribution to the unknown correct data generating process. Apparently, BIC differs from AIC only in the first term, which depends on sample size n . Distributions that give the minimum BIC are selected. From a given data set, the BIC helps in choosing the most probable distribution with best description of the data set.

3.5.3 Likelihood Ratio Test (LRT)

The superiority of the proposed EGEE distribution with other existing sub-distributions is compared using LRT statistics. The values of the restricted and unrestricted log-likelihood are used in constructing the likelihood ratio test statistics. The hypothesis of the test statistics is given as

$$H_0: \Theta = \Theta_o \text{ vs } H_1: \Theta \neq \Theta_o$$

The LR statistics use for testing H_0 vs H_1 is

$$\omega = 2 \left(l(\hat{\Theta}) - l(\hat{\Theta}_o) \right) \quad (3.54)$$

Where $\hat{\Theta}$ and $\hat{\Theta}_o$ are the log-likelihood under H_1 and H_0 , respectively. The statistic ω is asymptotically distributed as χ_k^2 , where k is the dimension of the subset Ω of interest and degree of freedom. H_0 is rejected at a significance p – values.

CHAPTER FOUR

4.1 Introduction

The estimation of parameters for the proposed EGEE distribution and other related existing distributions for both simulated and real- life data sets are presented in this chapter. The best distribution was also determined using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) , Likelihood Ratio Test (LRT), Log-Likelihood (LL) and fitted plot.

4.2 Simulation Study

A simulation study was carried out using the quantile function of the proposed EGEE distribution in (3.24) with the help of R-statistics package. Data were generated from sample sizes $n = 10$ and 100 with different shapes ($\beta = 0.5, 5, \alpha = 0.5, 0.8, 5$ and 8 and $k = 0.5, 0.8, 5$ and 8) and scale ($\theta = 2$, and 6) parameters combinations.

The mean, standard deviation and median of the proposed EGEE distribution for different parameter values are shown in Tables 4.1 and 4.3, while the skweness and kurtosis values obtained for the same combination of parameters are given in Table 4.2 and 4.4 respectively. As part of the simulation study, sample sizes of $n = 10, n = 50, n = 100, n = 200$ and 350 were generated from Exponentiated Exponential (EE), Exponential (E) and Exponentiated Generalized Exponential (EGE) distributions and fitted on the proposed Exponentiated Generalized Exponentiated Exponential (EGEE), EGE, EE and E distributions respectively. Tables 4.5, 4.7, 4.9, 4.11, 4.13, 4.16, 4.18, 4.20, 4.22, 4.24, 4.27, 4.29, 4.31, 4.33 and 4.35 show the estimated values obtained for the parameters and measured statistics for EGEE, EGE, EE and E distributions.

Tables 4.6, 4.8, 4.10, 4.12 and 4.14, 4.17, 4.19, 4.21, 4.23, 4.25, 4.28, 4.30, 4.32, 4.34 and 4.36 contained the likelihood ratio test results obtained from the aforementioned distributions, while Figures 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.9, 4.10, 4.11, 4.12, 4.13, 4.15, 4.17, 4.19, 4.21 and 4.23 shows the histograms of the simulated data sets along with the fitted plots.

**Table 4.1: The Mean, Standard Deviation and Median of EGEE Distribution
for $\Theta = 2, 4$ and 6 for $n=10$**

B	α	K	$\Theta=2$			$\Theta=4$			$\Theta=6$		
			Mean	SD	MD	Mean	SD	MD	Mean	SD	MD
0.5	0.5	0.5	0.9881	1.1960	0.9494	1.9763	2.3920	1.8988	2.9645	3.5880	2.8483
		0.8	1.4591	1.4541	1.5719	2.9182	2.9095	3.1438	4.3772	4.3642	4.7158
		5	4.0798	2.3439	4.7558	8.1596	4.6879	9.5117	12.2394	7.0317	14.2675
		8	4.9063	2.4738	5.6598	9.8126	4.9477	11.3196	14.7189	7.4215	16.9794
	0.8	0.5	1.7873	1.6779	1.9866	3.5747	3.3557	3.9731	5.3619	5.0336	8.9649
		0.8	2.3864	1.9472	2.7638	4.7729	3.8943	5.5276	7.1592	5.8416	8.2914
		5	5.3494	2.6859	6.1927	10.6988	5.3719	12.3854	16.0482	8.0577	18.5782
		8	6.2219	2.7717	7.1155	12.4438	5.5433	14.2309	18.6656	8.3149	21.3464
	5	0.5	6.9609	3.3771	8.1485	13.9219	6.7543	16.2971	20.8829	10.1314	24.4456
		0.8	7.8523	3.4522	9.0821	15.7045	6.9045	18.1642	23.5568	10.3567	27.2464
		5	11.4456	3.5659	12.7383	22.8912	7.1319	25.4766	34.3369	10.6978	38.2149
		8	12.3803	3.5744	13.6777	24.7607	7.1489	27.3553	37.141	10.7233	41.0329
	8	0.5	8.6479	3.5899	9.9517	17.2959	7.1798	19.9033	25.9437	10.7695	29.855
		0.8	9.5646	3.6279	10.8891	19.1292	7.2559	21.7782	28.6938	10.8839	32.6672
		5	13.1962	3.6832	14.5506	26.3925	7.3663	29.1012	39.5887	11.0495	43.6518
		8	14.1338	3.6871	15.4904	28.2676	7.3744	30.9807	42.4015	11.0616	46.4711
5	0.5	0.5	0.0219	0.0341	0.0166	0.0437	0.0683	0.0333	0.0656	0.1024	0.0499
		0.8	0.0988	0.1052	0.1025	0.1977	0.2104	0.205	0.2966	0.3156	0.3075
		5	1.6329	0.7460	1.9305	3.2659	1.4920	3.8609	4.8989	2.2382	5.7914
		8	2.3349	0.8861	2.7023	4.6698	1.7721	5.4046	7.0047	2.6582	8.1069
	0.8	0.5	0.0449	0.0540	0.0431	0.0899	0.1081	0.0861	0.1349	0.1621	0.1292
		0.8	0.1672	0.1449	0.1891	0.3344	0.2898	0.3781	0.5016	0.4346	0.5672
		5	1.9778	0.7475	2.2819	3.9555	1.4951	4.5638	5.9332	2.2426	6.8457
		8	2.7302	0.8514	3.0851	5.4603	1.7028	6.1701	8.1905	2.5542	9.2552
	5	0.5	0.2658	0.1644	0.3100	0.5316	0.3288	0.6200	0.7973	0.4932	0.9301
		0.8	0.6080	0.2884	2.1169	1.2160	0.5769	1.4113	1.8241	0.8654	2.1169
		5	3.2184	0.6224	3.4699	6.4368	1.2447	6.9398	9.6552	1.8671	10.4097
		8	9.6552	1.8671	10.4097	8.1439	1.3114	8.6781	12.2159	1.9671	13.0171
	8	0.5	0.3567	0.1928	0.4144	0.7134	0.3856	0.8289	1.0701	0.5784	1.2434
		0.8	0.7549	0.3111	0.8642	1.5098	0.6223	1.7284	2.2647	0.9334	2.5926
		5	3.5005	0.5911	3.7378	7.0001	1.1829	7.4756	10.5014	1.7734	11.2135
		8	4.3669	0.6174	4.6174	8.7339	1.2348	9.2329	13.1008	1.8522	13.8492

Source: Computations from the simulated data on EGEE Distribution

Table 4.2: The Skewness and Kurtosis of EGEE Distribution for $\Theta = 2, 4$ and 6 for $n=10$

B	α	k	$\Theta=2$		$\Theta=4$		$\Theta=6$	
			SK	KT	SK	KT	SK	KT
0.5	0.5	0.5	1.5700	1.6535	1.5700	1.6535	1.5700	1.6535
		0.8	1.1411	0.6917	1.1411	0.6917	1.1411	0.6917
		5	0.1097	-0.7757	0.1097	-0.7757	0.1097	-0.7757
		8	0.0036	-0.8161	0.0036	-0.8161	0.0036	-0.8161
	0.8	0.5	0.9649	0.3352	0.9649	0.3352	0.9649	0.3352
		0.8	0.6261	-0.2490	0.6261	-0.2490	0.6261	-0.2490
		5	-0.0394	-0.8538	-0.0394	-0.8538	-0.0394	-0.8538
		8	-0.1001	-0.8582	-0.1001	-0.8582	-0.1001	-0.8582
	5	0.5	-0.2396	-0.9347	-0.2396	-0.9347	-0.2396	-0.9347
		0.8	-0.2827	-0.9132	-0.2827	-0.9132	-0.2827	-0.9132
		5	-0.3461	-0.8721	-0.3461	-0.8721	-0.3461	-0.8721
		8	-0.3507	-0.8687	-0.3507	-0.8687	-0.3507	-0.8687
	8	0.5	-0.3333	-0.8971	-0.3333	-0.8971	-0.3333	-0.8971
		0.8	-0.3541	-0.8816	-0.3541	-0.8816	-0.3541	-0.8816
		5	-0.3837	-0.8574	-0.3837	-0.8574	-0.3837	-0.8574
		8	-0.3858	-0.8556	-0.3858	-0.8556	-0.3858	-0.8556
5	0.5	0.5	1.9758	2.7099	1.9758	2.7099	1.9758	2.7099
		0.8	1.3206	1.08311	1.3206	1.08311	1.3206	1.08311
		5	-0.4770	-0.9557	-0.4770	-0.9557	-0.4770	-0.9557
		8	-0.6179	-0.8201	-0.6179	-0.8201	-0.6179	-0.8201
	0.8	0.5	1.5949	1.7345	1.5949	1.7345	1.5949	1.7345
		0.8	0.8203	0.0853	0.8203	0.0853	0.8203	0.0853
		5	-0.5867	-0.8213	-0.5867	-0.8213	-0.5867	-0.8213
		8	-0.6843	-0.7056	-0.6843	-0.7056	-0.6843	-0.7056
	5	0.5	0.2204	-0.6939	0.2204	-0.6939	0.2204	-0.6939
		0.8	-0.1743	-0.6876	-0.1743	-0.6876	-0.1743	-0.6876
		5	-0.6454	-0.6876	-0.6454	-0.6876	-0.6454	-0.6876
		8	-0.6735	-0.6565	-0.6735	-0.6565	-0.6735	-0.6565
	8	0.5	0.05313	-0.3008	0.05313	-0.3008	0.05313	-0.3008
		0.8	-0.3109	-0.4021	-0.3109	-0.4021	-0.3109	-0.4021
		5	0.7379	-0.1575	0.7379	-0.1575	0.7379	-0.1575
		8	-0.7636	-0.1295	-0.7636	-0.1295	-0.7636	-0.1295

Source: Computations from the simulated data on EGEE Distribution

**Table 4.3: The Mean, Standard Deviation and Median of EGEE Distribution
for $\Theta = 2, 4$ and 6 for $n=100$**

B	α	k	$\Theta=2$			$\Theta=4$			$\Theta=6$		
			Mean	SD	MD	Mean	SD	MD	Mean	SD	MD
0.5	0.5	0.5	1.1563	2.1610	0.1596	2.3126	4.3220	0.3192	3.4689	6.4830	0.4789
		0.8	1.5416	2.4109	0.4486	3.0832	4.8219	0.8972	4.6248	7.2329	1.3438
		5	3.9936	3.1106	2.9703	7.9913	6.2212	5.9407	11.9870	9.3318	8.9111
		8	4.8134	3.1957	3.8174	9.6308	6.3915	7.6348	14.4463	9.5873	11.4523
	0.8	0.5	0.6126	1.2268	0.0684	1.2253	2.4537	0.1369	1.838	3.6806	0.2034
		0.8	0.9176	1.4645	0.2538	1.8352	2.929	0.5116	2.7528	4.3936	0.7675
		5	3.1623	2.2042	2.4921	6.3246	4.4084	4.9843	9.4869	6.6127	7.4765
		8	3.9588	2.3000	3.3108	7.9177	4.6000	6.6216	11.8767	6.9000	9.9325
	5	0.5	0.0339	0.0898	0.0020	0.0679	0.1796	0.004	0.1019	0.2694	0.0061
		0.8	0.1099	0.1881	0.0271	0.2199	0.3763	0.0343	0.3299	0.5645	0.0814
		5	1.5641	0.8074	1.3946	3.1282	1.6148	2.7893	4.6924	2.4222	4.1839
		8	2.2385	0.9178	2.1002	4.3170	1.8337	4.2005	6.7755	2.7536	6.3008
	8	0.5	0.0147	0.0410	0.0008	0.0294	0.0821	0.0016	0.0441	0.1231	0.0024
		0.8	0.0625	0.1083	0.0131	0.125	0.2166	0.0303	0.1875	0.3249	0.0454
		5	1.3301	0.6638	1.2234	2.7002	1.3277	2.4469	4.0503	1.9916	3.6704
		8	2.0177	0.7734	1.9011	4.0354	1.5469	3.8022	6.0531	2.3203	5.7034
0.8	0.5	0.5	1.8222	2.6498	0.6224	3.6445	5.2997	1.2449	5.4668	7.9496	1.8674
		0.8	2.3493	2.8795	1.1544	4.6986	5.759	2.3089	7.0480	8.6385	3.4634
		5	5.2278	3.4056	4.1820	10.4537	6.8112	8.3641	15.6835	10.2168	12.5462
		8	6.0986	3.4589	5.0735	12.1973	6.9179	10.1470	18.2960	10.3768	15.2205
	0.8	0.5	0.9738	1.5198	0.2869	1.9517	3.0396	0.5738	2.9275	4.5594	0.8608
		0.8	1.4030	1.7560	0.6688	2.8061	3.512	1.3377	4.2091	5.268	2.0066
		5	4.0622	2.3501	3.4042	8.1244	4.7002	6.8084	12.1867	7.0503	10.2126
		8	4.912	2.4137	4.2707	12.1312	7.0634	10.2736	14.7361	7.2413	12.8123
	5	0.5	0.0539	0.1143	0.0098	0.1118	0.2287	0.0197	0.1678	0.3431	0.0296
		0.8	0.1698	0.2285	0.0736	0.3397	0.457	0.1472	0.5095	0.6855	0.2208
		5	1.9175	0.7853	1.7728	3.8350	1.5706	3.5457	5.7525	2.3559	5.3186
		8	2.6653	0.8668	2.5279	5.3307	1.7336	5.0559	7.9960	2.6004	7.5838
	8	0.5	0.0243	0.0525	0.0039	0.0486	0.1050	0.0079	0.0729	0.1575	0.0118
		0.8	0.0967	0.1318	0.0413	0.1934	0.2636	0.0826	0.2901	0.3954	0.1238
		5	1.6431	0.6347	1.5382	3.2862	1.2694	3.0765	4.9294	1.9042	4.6147
		8	2.3631	0.7163	2.2644	4.7261	1.4325	4.5288	7.0892	2.1488	6.7932
5	0.5	0.5	6.7432	4.0122	5.6622	13.4863	8.0244	11.3243	20.2294	12.0566	16.9865
		0.8	7.6388	4.0533	6.5795	15.2776	8.1066	13.1391	22.9164	12.1599	19.7387
		5	11.2400	4.1140	10.2129	22.4801	8.228	20.4259	33.7201	12.342	30.6389
		8	12.1754	4.1185	11.1507	24.3508	8.237	22.3014	36.5262	12.3553	33.4321
	0.8	0.5	3.8049	2.4336	3.1149	7.6098	4.8672	6.2299	11.4147	7.3008	9.3448
		0.8	4.6351	2.5138	3.9690	9.2703	5.0277	7.9381	13.9054	7.5416	11.9072
		5	8.1386	2.6378	7.5118	16.2773	5.2756	15.0236	24.4160	7.9133	22.5354
		8	9.0667	2.6471	8.4429	18.1335	5.2943	16.8859	27.2003	7.9415	25.3289
	5	0.5	0.2615	0.2278	0.1833	0.5230	0.4556	0.3666	0.7846	0.6834	0.5500
		0.8	0.5918	0.3512	0.4925	1.1337	0.7025	0.9830	1.7756	1.0538	1.4776
		5	3.1786	0.6473	3.0637	6.3573	1.2947	6.1275	9.536	1.9420	9.1913
		8	4.0303	0.6754	3.9154	8.0607	1.3509	7.8309	12.0911	2.0263	11.7464
	8	0.5	0.1172	0.1087	0.0791	0.2345	0.2174	0.1582	0.3517	0.3261	0.2573
		0.8	0.3401	0.2054	0.2812	0.6802	0.4108	0.5625	1.0203	0.6163	0.8438
		5	2.6454	0.4920	2.5642	5.2908	0.9841	5.1283	7.9362	1.4762	7.6925
		8	3.4695	0.5219	3.3876	6.939	1.0438	6.7753	10.4085	1.5658	10.1630

Source: Computations from the simulated data on EGEE Distribution

Table 4.4: Skewness and Kurtosis at sample size 100 EGEE Distribution for $\Theta = 2, 4$ and 6 for $n=100$

β	α	K	$\theta=2$		$\theta=4$		$\theta=6$	
			SK	KT	SK	KT	SK	KT
0.5	0.5	0.5	3.455	16.6414	3.455	16.6414	3.455	16.6414
		0.8	2.9251	12.2422	2.9251	12.2422	2.9251	12.2422
		5	1.844	5.3403	1.844	5.3403	1.844	5.3403
		8	1.7359	4.8308	1.7359	4.8308	1.7359	4.8308
	0.8	0.5	3.8469	20.4196	3.8469	20.4196	3.8469	20.4196
		0.8	3.0237	13.0736	3.0237	13.0736	3.0237	13.0736
		5	1.5348	3.6594	1.5348	3.6594	1.5348	3.6594
		8	1.3982	3.1088	1.3982	3.1088	1.3982	3.1088
	5	0.5	5.7922	42.2041	5.7922	42.2041	5.7922	42.2041
		0.8	3.4497	16.9675	3.4497	16.9675	3.4497	16.9675
		5	0.7109	0.3459	0.7109	0.3459	0.7109	0.3459
		8	0.528	0.0460	0.528	0.0460	0.528	0.0460
	8	0.5	6.1811	46.9722	6.1811	46.9722	6.1811	46.9722
		0.8	3.5372	17.8128	3.5372	17.8128	3.5372	17.8128
		5	0.5932	0.0521	0.5932	0.0521	0.5932	0.0521
		8	0.4048	-0.1832	0.4048	-0.1832	0.4048	-0.1832
0.8	0.5	0.5	2.616	9.7831	2.616	9.7831	2.616	9.7831
		0.8	2.2731	7.5198	2.2731	7.5198	2.2731	7.5198
		5	1.6133	4.1072	1.6133	4.1072	1.6133	4.1072
		8	1.5528	3.862	1.5528	3.862	1.5528	3.862
	0.8	0.5	2.9067	12.0859	2.9067	12.0859	2.9067	12.0859
		0.8	2.3469	8.0295	2.3469	8.0295	2.3469	8.0295
		5	1.3670	2.9446	1.3670	2.9446	1.3670	2.9446
		8	1.2824	2.6424	1.2824	2.6424	1.2824	2.6424
	5	0.5	4.5766	28.4804	4.5766	28.4804	4.5766	28.4804
		0.8	2.6914	10.6886	2.6914	10.6886	2.6914	10.6886
		5	0.6153	0.2882	0.6153	0.2882	0.6153	0.2882
		8	0.4817	0.0893	0.4817	0.0893	0.4817	0.0893
	8	0.5	5.0456	33.6027	5.0456	33.6027	5.0456	33.6027
		0.8	2.8079	11.5959	2.8079	11.5959	2.8079	11.5959
		5	0.5075	0.1026	0.5075	0.1026	0.5075	0.1026
		8	0.3661	-0.0556	0.3661	-0.0556	0.3661	-0.0556
5	0.5	0.5	1.2706	2.5145	1.2706	2.5145	1.2706	2.5145
		0.8	1.2308	2.4016	1.2308	2.4016	1.2308	2.4016
		5	1.1722	2.2454	1.1722	2.2454	1.1722	2.2454
		8	1.1678	2.2343	1.1678	2.2343	1.1678	2.2343
	0.8	0.5	1.3568	2.8333	1.3568	2.8333	1.3568	2.8333
		0.8	1.2529	2.4782	1.2529	2.4782	1.2529	2.4782
		5	1.0994	2.0164	1.0994	2.0164	1.0994	2.0164
		8	1.0881	1.9853	1.0881	1.9853	1.0881	1.9853
	5	0.5	2.1308	7.2064	2.1308	7.2064	2.1308	7.2064
		0.8	1.4232	3.2443	1.4232	3.2443	1.4232	3.2443
		5	0.6491	0.6401	0.6491	0.6401	0.6491	0.6401
		8	0.6029	0.5456	0.6029	0.5456	0.6029	0.5456
	8	0.5	2.3714	8.9285	2.3714	8.9285	2.3714	8.9285
		0.8	1.4675	3.4686	1.4675	3.4686	1.4675	3.4686
		5	0.5613	0.4339	0.5613	0.4339	0.5613	0.4339
		8	0.5106	0.3435	0.5106	0.3435	0.5106	0.3435

Source: Computations from the simulated data on EGEE Distribution

Tables 4.1 and 4.3 show an increase in values of mean, standard deviation and media, when the value of scale parameter was increased ($\theta = 2, 4$, and 6) while other parameters were fixed. Increase in the shape parameters (α, β and k) also increases the values of the mean, standard deviation and median when the scale parameter θ is held constant. Increasing a shape parameter and holding others shape parameters and scale parameter constant increases the mean, standard deviation and median. From Tables 4.2 and 4.4 it was observed that the

skweness and kurtosis values remained constant as the value of scale parameter was increased ($\theta = 2, 4$, and 6) and that of the shape parameters (α, β and k) were fixed. The values of Skweness and Kurtosis from the simulated data reduce as the values of shape parameters (α, β and k) increase. An increase in one of the shape parameters while others and scale parameter are fixed reduces the values of skweness and kurtosis.

4.2.1 Simulated Exponentiated Exponential for $n=10$

A sample size of $n = 10$ was simulated using shape and scale parameters values of $k = 1.5$ and $\theta = 0.5$ respectively. Data so obtained were fitted on EGEE, EGE, EE and E distributions.

Table: 4.5: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 10

Distributions	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	3.93370	1.9366	1.4562	2.8926	- 6.1233	20.2467	21.4570
	(32.7117)	(16.1031)	(0.6130)	(19.4167)			
EGE	9.3791	1.6927		4.9781	- 6.1663	18.3326	19.2404
	(NA)	(NA)		(NA)			
EE			1.6928	0.5306	- 6.1664	16.3326	16.9378
			(0.7668)	(0.1918)			
E				1.3781	- 6.7929	15.5858	15.8884
				(0.43579)			

Source: Simulated data from Exponentiated Exponential Distribution

Table 4.5 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicates that the best model is Exponential distribution with values 15.5858 and 15.8884 as compare to EE, EGE, and EGEE with AIC values of 16.3326, 18.3326 and 20.2467, while the BIC values are 16.9378, 19.2404 and 21.4570 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.6 for LRT:

$$LR = 2(-6.1233 - (-6.1663)) = 0.086$$

$$LR = 2(-6.1233 - (-6.1664)) = 0.0862$$

$$LR = 2(-6.1233 - (-6.7929)) = 1.3392$$

Table: 4.6: Likelihood Ratio Test (LRT) Statistic for sample size 10

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs $H_1: \text{not } H_0$	0.086	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1$ vs $H_1: \text{not } H_0$	0.0862	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs $H_1: \text{not } H_0$	1.3392	$p > 0.05$

Source: Simulated data from Exponentiated Exponential Distribution

The computed p-values from Table 4.6 indicate that there no significant difference between the proposed **EGEE** and the sub distributions.

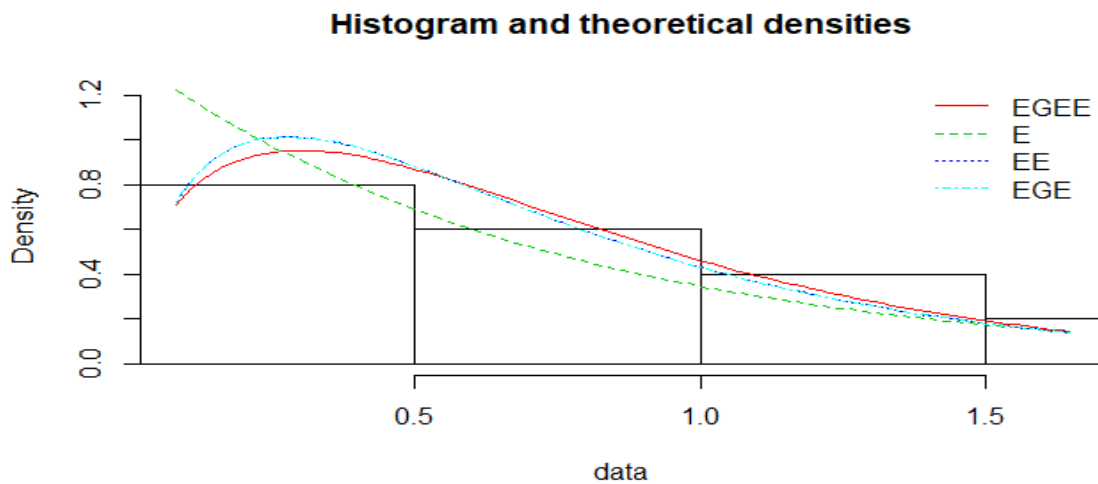


Figure 4.1: Fitted Plot of EGEE EGE, EE and E Models for sample size n=10

The histogram of the simulated data set at sample size $n = 10$ along with fitted distribution is shown in Figure 4.1. The plots showed that the Exponential (E) distribution fitted the data best than the other distributions.

4.2.2 Simulated Exponentiated Exponential for n=50

A sample size of $n=50$ was simulated from an Exponentiated Exponential (EE) distribution using shape and scale parameters values of $k = 1.5$ and $\theta = 0.5$, respectively. Data so obtained were fitted on EGEE, EGE and E distributions.

Table: 4.7: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 50

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	2.8108	3.0702	1.5419	2.6262	-24.6366	57.2732	64.9213
	(29.4391)	(32.1564)	(0.4332)	(10.5534)			
EGE	8.1520	1.8581		3.7300	-24.9069	55.8138	61.5499
	(NA)	(NA)		(NA)			
EE			1.8577	0.4575	-24.9069	53.8138	57.6378
			(0.3822)	(0.0721)			
E				1.5201	-29.0601	60.1202	62.0322
				(0.2149)			

Source: Simulated data from Exponentiated Exponential Distribution

Table 4.7 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicate that the best model is Exponentiated Exponential distribution with values 53.8138 and 57.6378 as compare to E, EGE and EGEE distribution with AIC values of 60.1202, 55.8138 and 57.2732, while the BIC values are 62.0322, 61.5499 and 64.9213 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.8 below for LRT:

$$LR = 2(-24.6366 - (-24.9069)) = 0.5406$$

$$LR = 2(-24.6366 - (-24.9069)) = 0.5406$$

$$LR = 2(-24.6366 - (-29.0601)) = 17.6940$$

Table 4.8: Likelihood Ratio Test (LRT) Statistic for Sample size n=50

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: not H_0$	0.5406	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: not H_0$	0.5406	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: not H_0$	17.6940	< 0.0001

Source: Simulated data from Exponentiated Exponential Distribution

The computed p-values from Table 4.8 indicates that there no significant difference between **EGEE vs. EGE** and **EGEE vs. EE**, while there is a significant difference between **EGEE vs. E** distributions.

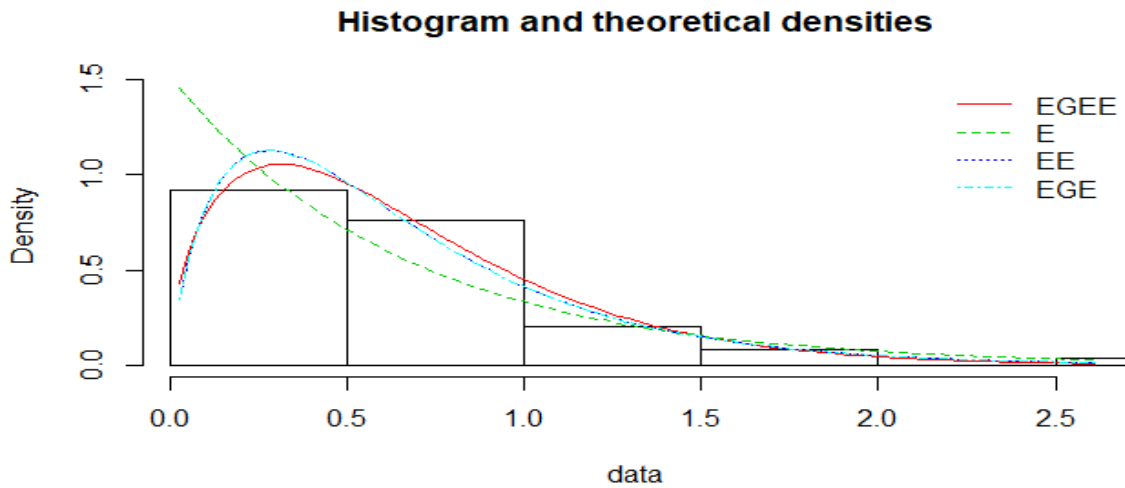


Figure 4.2: Fitted Plot of EGEE, EGE, EE and E Models for sample size $n=50$

The histogram of the simulated data set at sample size $n = 50$ along with fitted distribution is shown in Figure 4.2. The plots showed that the Exponentiated Exponential (EE) distribution fitted the data best than the other distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.3 Simulated Exponentiated Exponential for $n=100$

From an Exponentiated Exponential distribution a sample size of $n = 100$ was simulated using shape and scale parameters values of $k = 1.5$ and $\theta = 0.5$ respectively. Data so obtained were fitted on EGEE, EGE and E distributions.

Table: 4.9: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 100

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	1.2474	2.4461	1.7514	1.0279	-42.6800	93.3600	103.7808
	(7.4279)	(14.5685)	(0.4595)	(1.9253)			
EGE	8.7302	2.0534		3.6109	-42.8643	91.7287	99.5442
	(NA)	(NA)		(NA)			
EE			2.0531	0.4137	-42.8643	89.7287	94.9390
			(0.3066)	(0.0454)			
E				1.5927	-53.4543	108.9087	111.5139
				(0.1592)			

Source: Simulated Data from Exponentiated Exponential Distribution

Table 4.9 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicates that the best model is Exponentiated Exponential (EE) with values 89.7287 and 94.9390 as compare to E, EGE and EGEE distribution with AIC values of 108.9087, 91.7287 and 93.3600, while the BIC values are 111.5139, 99.5442 and 103.7808 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.10 below for LRT:

$$LR = 2(-42.6800 - (-42.8643)) = 0.3686$$

$$LR = 2(-42.6800 - (-42.8643)) = 0.3686$$

$$LR = 2(-42.6800 - (-53.4543)) = 21.5486$$

Table 4.10: Likelihood Ratio Test (LRT) Statistic for Sample Size n=100

Model	Hypothesis	LRT	<i>p – value</i>
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: not H_0$	0.3686	$p > 0.05$.
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: not H_0$	0.3686	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: not H_0$	21.5486	$p < 0.0001$

Source: Simulated data from Exponentiated Exponential Distribution

The computed p-values from Table 4.10 indicate no significant difference between **EGEE vs. EGE** and **EGEE vs. EGE** distributions, while **EGEE vs. E** is significantly different.

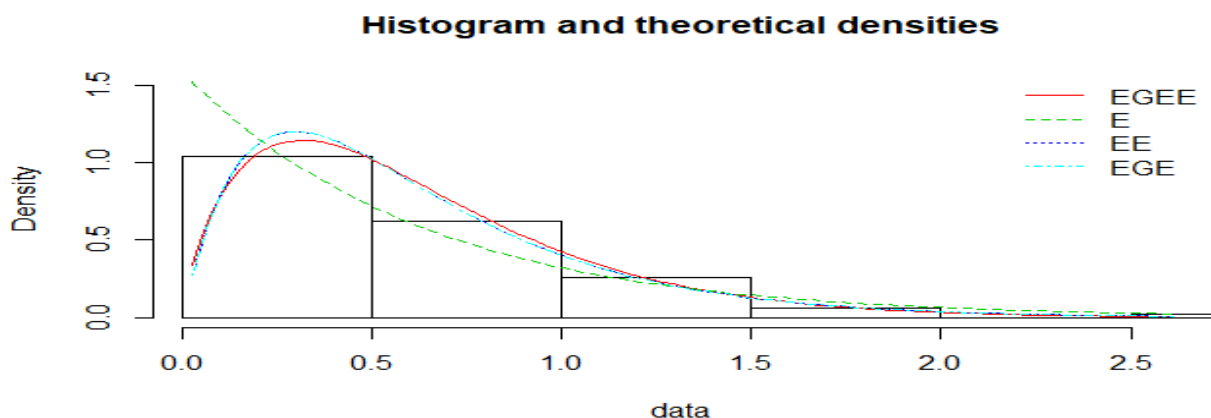


Figure 4.3: Fitted Plot of EGEE EGE, EE and E Models for sample size n=100

The histogram of the simulated data set at sample size $n = 100$ along with fitted distribution is shown in Figure 4.4. The plots showed that the Exponentiated Exponential (EE) distribution fitted the data best than the other models. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.4 Simulated Exponentiated Exponential for n=200

From an Exponentiated Exponential distribution, a sample size of $n = 200$ was simulated using shape and scale parameters values of $k = 1.5$ and $\theta = 0.5$, respectively. Data so obtained were fitted on EGEE, EGE and E distributions.

Table: 4.11: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 200

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	0.2791	2.9864	2.0506	0.3687	-92.1802	192.3605	205.5538
	(1.8801)	(20.1582)	(0.4828)	(0.2941)			
EGE	8.5003	1.9723		3.6603	-92.1988	190.3977	200.2927
	NA	NA		NA			
EE			1.9719	0.4307	-92.1988	188.3977	194.9943
			(0.2082)	(0.0340)			
E				1.5587	-111.2255	224.4509	227.7493
				(0.1102)			

Source: Simulated Data from Exponentiated Exponential Distribution

Table 4.11 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicates that the best model is Exponentiated Exponential (EE) distribution with values 188.3977 and 194.9943 as compare to E, EGE and EGEE distribution with AIC values of 224.4509, 190.3977 and 192.3605, while the BIC values are 227.7493, 200.2927 and 205.5538 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.12 below for LRT:

$$LR = 2(-92.1802 - (-92.1988)) = 0.0372$$

$$LR = 2(-92.1802 - (-92.1988)) = 0.0372$$

$$LR = 2(-92.1802 - (-111.2555)) = 38.1506$$

Table 4.12: Likelihood Ratio Test (LRT) Statistic for sample size n=200

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: \text{not } H_0$	0.0372	0.500
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: \text{not } H_0$	0.0372	0.500
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: \text{not } H_0$	38.1506	< 0.00001

Source: Simulated data from Exponentiated Exponential Distribution

The computed p-values from Table 4.12 indicate no significant difference between **EGEE vs. EGE** and **EGEE vs. EGE**, while **EGEE vs. E** distribution is significant.

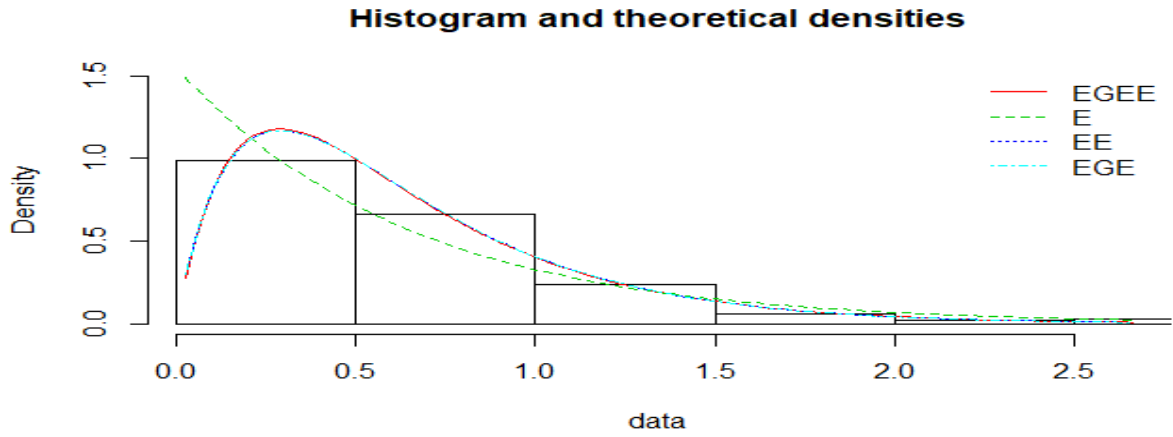


Figure 4.4: Fitted Plot of EGEE EGE, EE and E Models for sample size n=200

The histogram of the simulated data set at sample size $n = 200$ along with fitted distribution is shown in Figure 4.4. The plots showed that the Exponentiated Exponential (EE) distribution fitted the data best than any other. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL). It was observed that the performance of the fit of the proposed distribution increases gradually as the sample size increases.

4.2.5 Simulated Exponentiated Exponential for n=350

From an Exponentiated Exponential distribution ,a sample size of $n = 350$ was simulated using shape and scale parameters values of $k = 1.5$ and $\theta = 0.5$,respectively. Data so obtained were fitted on EGEE, EGE and E distributions.

Table 4.13: Maximum Likelihood Estimate of Parameters, Standard Errors in Parenthesis, Log-likelihood and Information Criteria of sample size 350

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	3.9942	0.0469	3.3777	0.0969	-157.4098	322.8195	338.2513
	(4.3533)	(0.0509)	(0.2866)	(0.01509)			
EGE	8.1194	1.7885		3.5873	-161.2381	328.4761	340.0499
	NA	NA		NA			
EE			1.7879	0.4418	-161.2381	326.4762	334.1920
			(0.1417)	(0.0272)			
E				1.5961	-186.3417	374.6834	378.5413
				0.0853			

Source: Simulated data from Exponentiated Exponential Distribution

Table 4.13 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC indicates that the best model is Exponentiated Generalized Exponentiated Exponential (EGEE) distribution with value 322.8195 as compare to E, EE and EGE distribution with AIC values of 374.6834, 326.4762 and 328.4761. However, the BIC value of 334.1920 shows that the best distribution is EE distribution as compared to EGEE, EGE, and E distributions of values 338.2513, 340.0499, and 378.5413, respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in 4.14 below for LRT:

$$LR = 2(-157.4098 - (-161.2381)) = 7.6566$$

$$LR = 2(-157.4098 - (-161.2381)) = 7.6566$$

$$LR = 2(-157.4098 - (-186.3417)) = 57.8638$$

Table 4.14: Likelihood Ratio Test (LRT) Statistic for sample size 350

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: H_0$ is false	7.6566	P<0.05
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: H_0$ is false	7.6566	P<0.05
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: H_0$ is false	57.8638	P<0.05

Source: Simulated data from Exponentiated Exponential Distribution

The computed p-values from Table 4.14 indicates significant differences between the proposed **EGEE vs. EGE**, **EGEE vs. EGE**, and **EGEE vs. E** distributions.

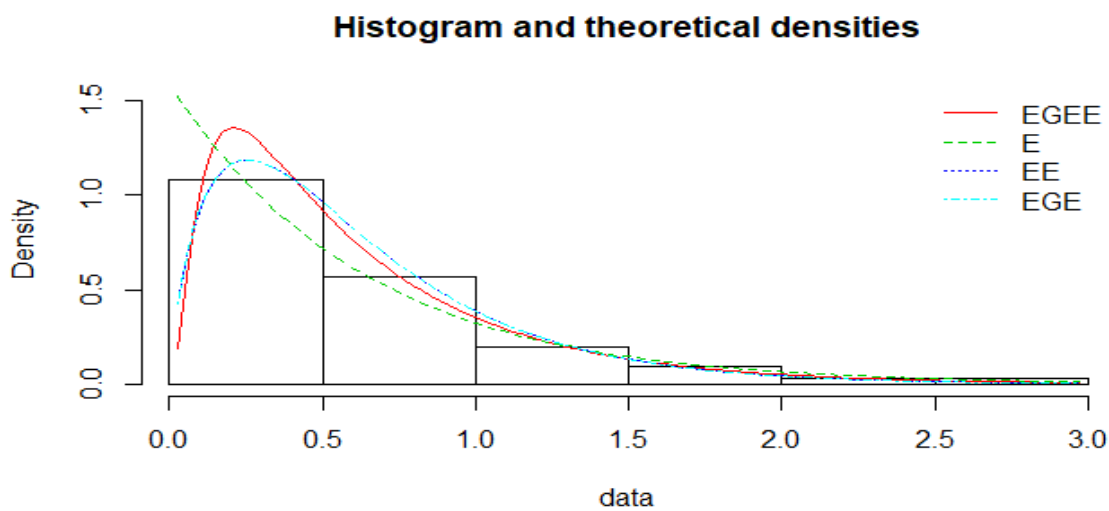
**Figure 4.5: Fitted Plot of EGEE EGE, EE and E Models for sample size n=350**

Figure 4.5 shows the histogram of the simulated data se at sample size $n=350$ along with the fitted distributions. The plots showed that the Exponentiated Generalized Exponentiated Exponential (EGEE) distribution fitted the data best than the other models. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

Table 4.15: Summary Table on the Performance of Fitted Distributions to Simulated data from Exponentiated Exponential Distribution

sample size		Rank		
	MEASURED STATISTIC	1 st	2 nd	3 rd
10	AIC	E	EGE	EGEE
	BIC	E	EGE	EGEE
	Log-likelihood	EGEE	EGE	E
50	AIC	EGE	EGEE	E
	BIC	EGE	E	EGEE
	Log-likelihood	EGEE	EGE	E
100	AIC	EGE	EGEE	E
	BIC	EGE	EGEE	E
	Log-likelihood	EGEE	EGE	E
200	AIC	EGE	EGEE	E
	BIC	EGE	EGEE	E
	Log-likelihood	EGEE	EGE	E
350	AIC	EGEE	EGE	E
	BIC	EGEE	EGE	E
	Log-likelihood	EGEE	EGE	E

Source: Simulated data from Exponentiated Exponential Distribution

Table 4.15 shows the ranking in performance of the fitted EGEE, EGE, and E distributions on simulated data from EE distribution. The ranks 1st, 2nd, and 3rd indicate the order of performance of the distributions at various sample sizes $n=10, 50, 100, 200$, and 350 . For sample size $n=10$ the reported AIC and BIC shows that the E distribution outperformed the EGE and EGEE distributions, while the log-likelihood indicate the best in EGEE distribution. AIC and BIC values for sample sizes of $50, 100$ and 200 show an order in performances as EGE, EGEE and E distributions, while the log-likelihood showed an order of EGEE, EGE, and E distributions respectively. Sample size $n= 350$ shows that the proposed EGEE distribution can serve as the best alternative distribution to data simulated from EE distribution, and these can be confirmed from the AIC, BIC, and log-likelihood values.

4.2.6 Simulated Exponential for $n=10$

The sixth simulated data of sample size $n=10$ was simulated from an Exponential (E) distribution using scale parameters value of $\theta = 0.5$. Data obtained were fitted on EGEE, EGE, and EE distributions.

Table 4.16: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 10

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	0.0555	1.7878	1.8506	0.1593	-14.5871	37.1742	38.3845
	(0.1590)	(5.1208)	(1.3619)	(0.0025)			
EGE	3.4536	1.3053		4.9227	-15.0420	36.0840	36.9917
	NA	NA		NA			
EE			1.3056	1.4253	-15.0420	34.0840	34.6892
			(0.5737)	(0.5576)			
E				0.5933	-15.2191	32.4382	32.7408
				(0.1876)			

Source: Simulated Data from Exponential Distribution

Table 4.16 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicates that the best model is Exponential distribution with values 32.4382 and 32.7408 as compare to EE, EGE and EGEE with AIC values of 34.0840, 36.0840 and 37.1742, while there BIC values are 34.6892, 36.9917 and 38.3845 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.17 below for LRT:

$$LR = 2(-14.5871 - (-15.0420)) = 0.9098$$

$$LR = 2(-14.5871 - (-15.0420)) = 0.9098$$

$$LR = 2(-14.5871 - (-15.2191)) = 1.264$$

Table 4.17: Likelihood Ratio Test (LRT) Statistic for sample size 10

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: not H_0$	0.9098	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: not H_0$	0.9098	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: not H_0$	1.264	$p > 0.05$

Source: Simulated data from Exponential Distribution

The computed p-values from Table 4.17 indicate no significant difference between the proposed **EGEE vs. EGE**, **EGEE vs. EGE**, and **EGEE vs. E** distributions.

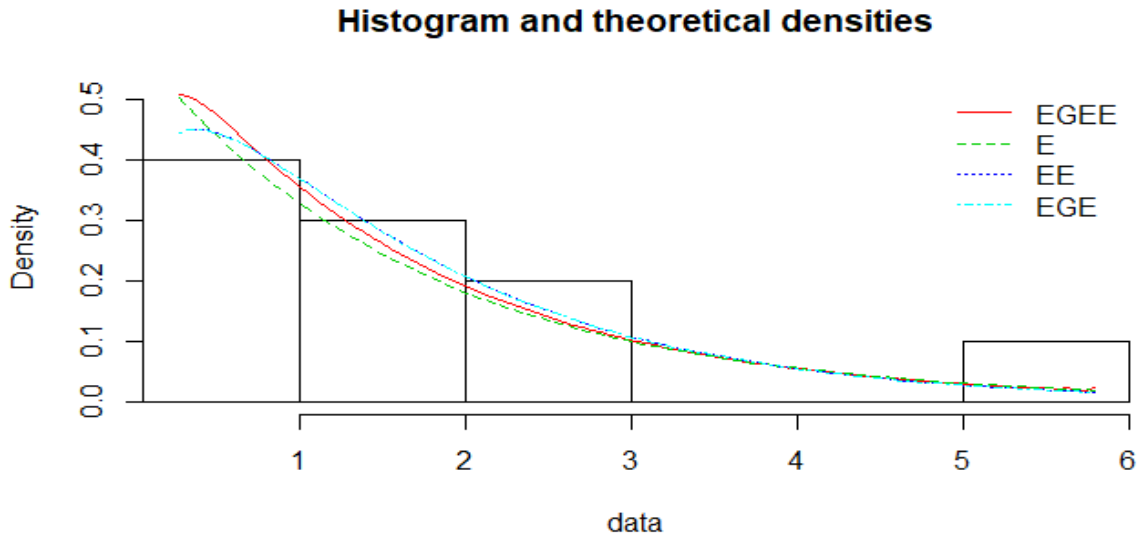


Figure 4.6: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=10$

The histogram of the simulated data set at sample size $n = 10$ along with fitted distribution is shown in Figure 4.6. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data equally, and this is confirmed by the values of the log-likelihood (LL). The plot indicates no significant differences in all the fitted distributions.

4.2.7 Simulated Exponential for $n=50$

The Seventh simulated data of sample size $n=50$ was simulated from an Exponential (E) distribution using scale parameters value of $\theta = 0.5$. Data obtained were fitted on EGEE, EGE, and EE distributions.

Table 4.18: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 50

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	0.1998	3.2989	1.5275	1.0691	-82.1833	172.3666	180.0147
	(2.5507)	(42.3245)	(0.5391)	(1.9918)			
EGE	3.1945	1.4506		5.0016	-82.2015	170.403	176.1391
	(NA)	(NA)		(NA)			
EE			1.4506	1.5650	-82.2015	168.403	172.2271
			(0.2888)	(0.2648)			
E				0.5081	-83.8530	169.706	171.618
				(0.0718)			

Source: Simulated Data from Exponential Distribution

Table 4.18 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC indicates that the best model is Exponentiated Exponential distribution (EE) with value 168.403 as compare to E, EGE, and EGEE with AIC values of 169.706, 170.403, and 172.3666. However, the BIC value (171.618) shows the best in Exponential (E) distribution compared to EE, EGE, and EGEE distributions with BIC values 172.2271, 176.1391, and 180.0147, respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.19 below for LRT:

$$LR = 2(-82.1833 - (-82.2015)) = 0.0364$$

$$LR = 2(-82.1833 - (-82.2015)) = 0.0364$$

$$LR = 2(-82.1833 - (-83.8530)) = 3.3394$$

Table 4.19: Likelihood Ratio Test (LRT) Statistic for sample size 50

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1 \text{ vs } H_1: \text{not } H_0$	0.0364	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1 \text{ vs } H_1: \text{not } H_0$	0.0364	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1 \text{ vs } H_1: \text{not } H_0$	3.3394	$p > 0.05$

Source: Simulated Data from Exponential Distribution

The computed p-values from Table 4.19 indicate no significant difference between the proposed **EGEE vs. EGE**, **EGEE vs. EE**, and **EGEE vs. E** distributions.

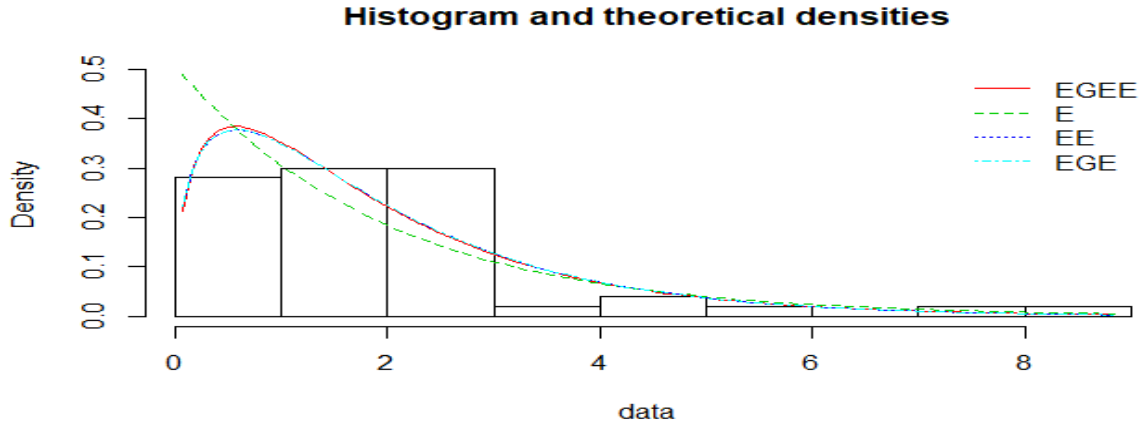


Figure 4.7: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=50$

The histogram of the simulated data set at sample size $n = 50$ along with fitted distribution is shown in Figure 4.7. The fitted plots showed no significant difference amongst Exponentiated Generalized Exponentiated Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.8 Simulated Exponential for $n=100$

The Eight simulated data of sample size $n=100$ was simulated from an Exponential (E) distribution using scale parameters value of $\theta = 0.5$. Data obtained were fitted on EGEE, EGE, and EE distributions.

Table 4.20: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 100

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	0.0371	6.7627	1.7381	0.2304	-100.2602	208.5203	218.941
	(0.1086)	(25.3728)	(0.9597)	(0.4631)			
EGE	4.914212	1.3806		4.1461	-100.4822	206.9643	214.7798
	(NA)	(NA)		(NA)			
EE			1.38017	0.8439	-100.4822	204.9643	210.1747
			(0.1928)	(0.1024)			
E				0.9702	-103.0215	208.0431	210.6482
				(0.0970)			

Source: Simulated Data from Exponential Distribution

Table 4.20 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria's AIC and BIC indicate that the best distribution is Exponentiated Exponential (EE) distribution with values 204.9643 and 210.1747 as compare to E, EGE, and EGEE with AIC values of 208.0431, 206.9643 and 208.5203, while the BIC values are 210.6482, 214.7798 and 218.941 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.21 below for LRT:

$$LR = 2(-100.2602 - (-100.4822)) = 0.444$$

$$LR = 2(-100.2602 - (-100.4822)) = 0.4440$$

$$LR = 2(-100.2602 - (-103.0215)) = 5.5226$$

Table 4.21: Likelihood Ratio Test (LRT) Statistic for sample size 100

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1 \text{ vs. } H_1: \text{not } H_0$	0.4440	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1 \text{ vs. } H_1: \text{not } H_0$	0.4440	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1 \text{ vs. } H_1: \text{not } H_0$	5.5226	$p > 0.05$

Source: Simulated data from Exponential

The computed p-values from Table 4.21 indicates no significant difference between the proposed **EGEE vs. EGE**, **EGEE vs. EGE**, and **EGEE vs. E** distributions.

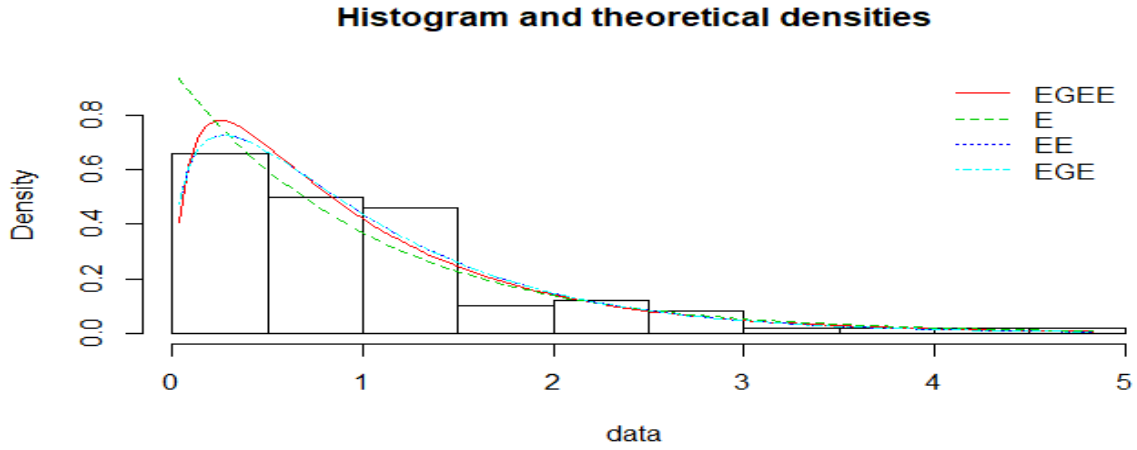


Figure 4.8: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=100$

The histogram of the simulated data set at sample size $n = 100$ along with fitted distribution is shown in Figure 4.8. The fitted plots showed closeness amongst Exponentiated Generalized Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.9 Simulated Exponential for $n=200$

The ninth simulated data of sample size $n=200$ was simulated from an Exponential (E) distribution using scale parameters value of $\theta = 0.5$. Data obtained were fitted on EGEE, EGE, and EE distributions.

Table 4.22: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 200

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	1.6780	1.3262	1.2687	3.4983	-334.4808	676.9617	690.1549
	(4.9644)	(3.9226)	(0.1991)	(9.5631)			
EGE	3.1015	1.3267		5.2093	-334.5726	675.1453	685.0402
	(NA)	(NA)		(NA)			
EE			1.3270	1.6792	-334.5726	673.1453	679.7419
			(0.1288)	(0.1441)			
E				0.5000	-338.6263	679.2526	682.5509
				(0.0353)			

Source: Simulated Data from Exponential Distribution

Table 4.22 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicates the best model in Exponentiated Exponential (EE) distribution with values 673.1453 and 679.7419 as compare to E, EGE, and EGEE with AIC values of 679.2526, 675.1453 and 676.9617. While there BIC values are 682.5509, 685.0402, and 690.1549, respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.23 below for LRT:

$$LR = 2(-334.4808 - (-334.5726)) = 0.1836$$

$$LR = 2(-334.4808 - (-334.5726)) = 0.1836$$

$$LR = 2(-334.4808 - (-338.6263)) = 8.2890$$

Table 4.23: Likelihood Ratio Test (LRT) Statistic for sample size 200

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: H_0$ is not equal	0.1836	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: H_0$ is not equal	0.1836	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: H_0$ is not equal	8.2890	$p < 0.05$

Source: Simulated data from Exponential Distribution

The computed p-values from Table 4.23 indicate no significant difference between the proposed **EGEE vs. EGE**, **EGEE vs. EE** distributions, while **EGEE vs. E** is significantly different.

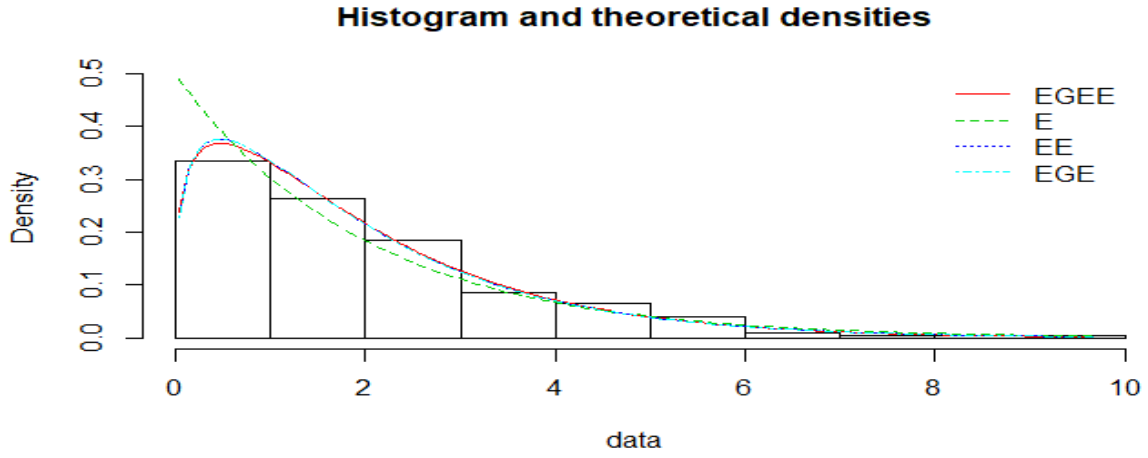


Figure 4.9: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=200$

The histogram of the simulated data set at sample size $n = 200$ along with fitted distribution is shown in Figure 4.9. The fitted plots showed closeness amongst Exponentiated Generalized Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.10 Simulated Exponential for $n=350$

The tenth simulated data of sample size $n=350$ was simulated from an Exponential (E) distribution using scale parameters value of $\theta = 0.5$. Data obtained were fitted on EGEE, EGE, and EE distributions.

Table 4.24: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 350

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	5.1863	0.4152	1.1337	3.755	-587.1266	1182.253	1197.685
	(10.9189)	(0.8675)	(0.0884)	(8.4598)			
EGE	3.0981	1.1545		5.6062	-587.2752	1180.55	1192.124
	(NA)	(NA)		(NA)			
EE			1.1548	1.8091	-587.2752	1178.55	1186.266
			(0.0826)	(0.1214)			
E				0.5048	-589.2473	1180.495	1184.353
				(0.0269)			

Source: Simulated Data from Exponential Distribution

Table 4.24 shows the estimated parameters values for all the aforementioned fitted distributions. The information criterion AIC indicates that the best distribution is Exponentiated Exponential (EE) distribution with value 1178.55 compare to EE, EGE, and EGEE with AIC values of 1180.495, 1180.55, and 1182.253. However, the BIC value for Exponential (E) distribution shows the best compare to EE (1186.266), EGE (1192.124) and EGEE (1197.685) distributions, respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.25 below for LRT:

$$LR = 2(-587.1266 - (-587.2752)) = 0.2972$$

$$LR = 2(-587.1266 - (-587.2752)) = 0.2972$$

$$LR = 2(-587.1266 - (-589.2473)) = 4.2414$$

Table 4.25: Likelihood Ratio Test (LRT) Statistic for sample size 350

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1 \text{ vs. } H_1: \text{not } H_0$	0.2972	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1 \text{ vs. } H_1: \text{not } H_0$	0.2972	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1 \text{ vs. } H_1: \text{not } H_0$	4.2414	$p > 0.05$

Source: Simulated data from Exponential

The computed p-values from Table 4.25 indicates no significant difference between the proposed **EGEE vs. EGE**, **EGEE vs. EGE**, and **EGEE vs. E** distributions.

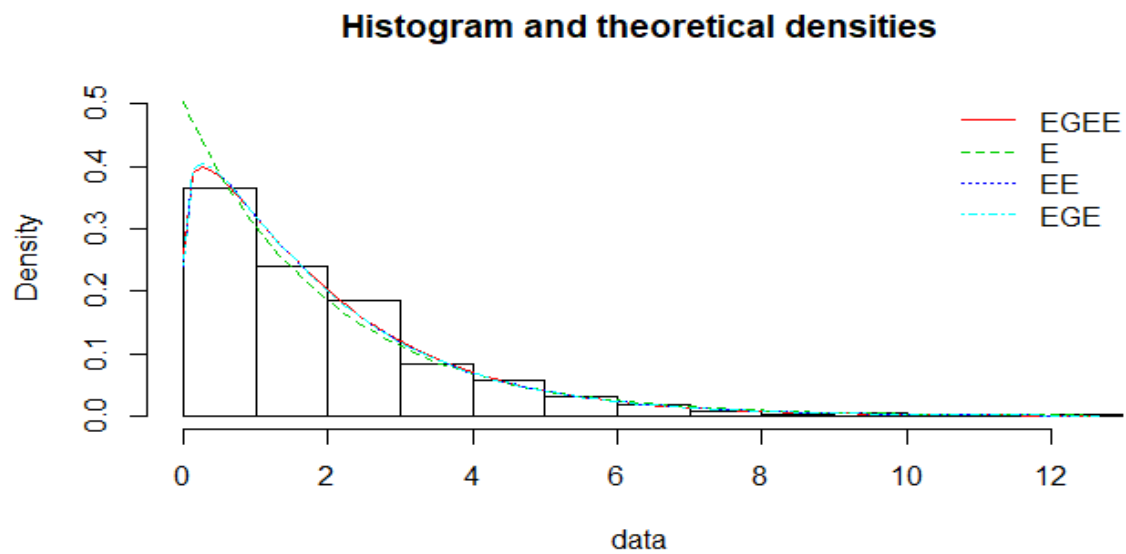


Figure 4.10: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=350$

The histogram of the simulated data set at sample size $n = 350$ along with fitted distribution is shown in Figure 4.10. The fitted plots showed closeness amongst Exponentiated Generalized Exponentiated Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

Table 4. 26: Summary Table on the Performance of Fitted Distributions to Simulated data from Exponential Distribution

sample size		Rank		
	measured statistics	1 st	2 nd	3 rd
10	AIC	EE	EGE	EGEE
	BIC	EE	EGE	EGEE
	Log-likelihood	EGEE	EE	EGE
50	AIC	EE	EGE	EGEE
	BIC	EE	EGE	EGEE
	Log-likelihood	EGEE	EE	EGE
100	AIC	EE	EGE	EGEE
	BIC	EE	EGE	EGEE
	Log-likelihood	EGEE	EE	EGE
200	AIC	EE	EGE	EGEE
	BIC	EE	EGE	EGEE
	Log-likelihood	EGEE	EE	EGE
350	AIC	EE	EGE	EGEE
	BIC	EE	EGE	EGEE
	Log-likelihood	EGEE	EE	EGE

Source: Simulated Data from Exponential Distribution Data

Table 4.26 shows the ranking in performance of the fitted EGEE, EGE, and EE distributions on simulated data from E distribution. The ranks 1st, 2nd, and 3rd indicate the order of performance of the distributions at various sample sizes n=10, 50, 100, 200, and 350. AIC and BIC values for sample sizes of 10, 50, 100, 200 and 350 show an order in performances as EE, EGE, and EGEE distributions, while the log-likelihood showed an order of EGEE, EE and EGE distributions respectively. The proposed EGEE distribution can serve as the best alternative distribution to data simulated from E distribution, and these can be confirmed from the log-likelihood values, while the AIC and BIC show the best alternative as EE distribution.

4.2.11 Simulated Exponentiated Generalized Exponential for n=10

The eleventh simulated data of sample size n=10 was simulated from an Exponentiated Generalized Exponential (EGE) distribution using a scale parameter value of = 0.5 and shape parameters $\alpha = 0.5$ and $\beta = 6.5$. Data obtained were fitted on EGEE, EE, and E distributions.

Table 4.27: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 10

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	7.3501	0.3267	4.6026	0.6303	-4.4882	16.9765	18.1868
	(125.252)	(5.5453)	(11.273)	(2.7635)			
EGE	5.4081	7.0568		1.8937	-4.5005	15.0010	15.9088
	(790.3244)	(4.4642)		(276.8193)			
EE			7.0578	0.3501	-4.5005	13.001	13.6062
			(4.4396)	(0.1024)			
E				1.1005	-9.0423	20.0846	20.3872
				(0.3480)			

Source: Simulated Data from Exponentiated Generalized Exponential Distribution

Table 4.27 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicate that the best distribution is Exponentiated Exponential distribution with values 13.001 and 13.6062 as compare to E, EGE, and EGEE with AIC values of 20.0846, 15.0010 and 16.9765, while the BIC values are 20.3872, 15.9088 and 18.1868 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.28 below for LRT:

$$LR = 2(-4.4882 - (-4.5005)) = 0.0246$$

$$LR = 2(-4.4882 - (-4.5005)) = 0.0246$$

$$LR = 2(-4.4882 - (-9.0423)) = 9.1026$$

Table 4.28: Likelihood Ratio Test (LRT) Statistic for sample size 10

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1 \text{ vs. } H_1: \text{not } H_0$	0.0246	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1 \text{ vs. } H_1: \text{not } H_0$	0.0246	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1 \text{ vs. } H_1: \text{not } H_0$	9.1026	$p < 0.05$

Source: Simulated data from Exponentiated Generalized Exponential

The computed p-values from Table 4.28 indicates no significant difference between the proposed **EGEE vs. EGE** and **EGEE vs. EGE** distributions, while **EGEE vs. E** is significantly different.

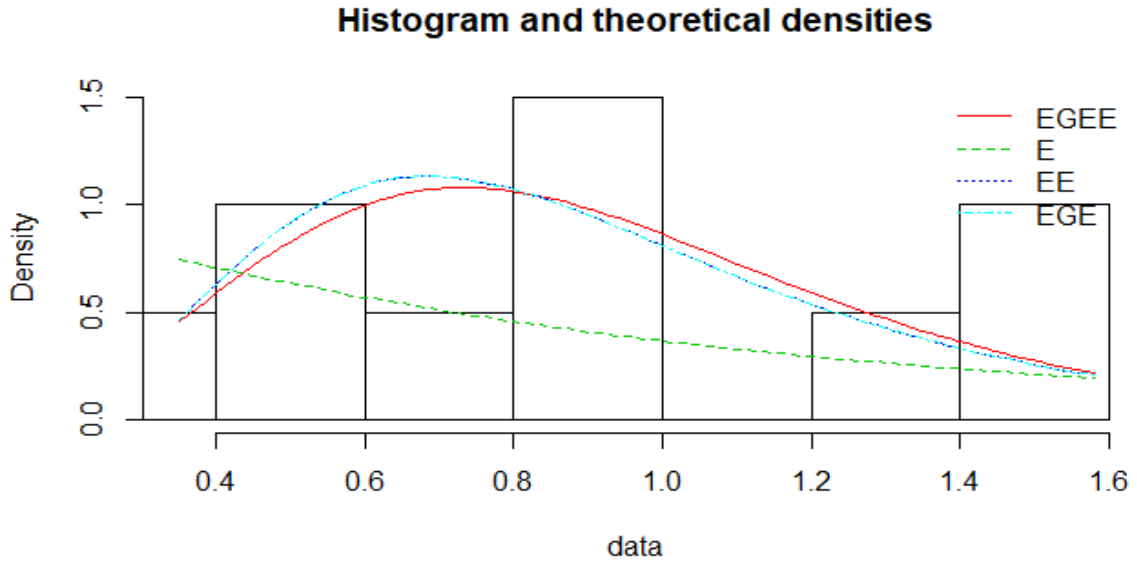


Figure 4.11: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=10$

The histogram of the simulated data set at sample size $n = 10$ along with fitted distribution is shown in Figure 4.11. The fitted plots showed little difference amongst Exponentiated Generalized Exponentiated Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL)

4.2.12 Simulated Exponentiated Generalized Exponential for $n=50$

The twelfth simulated data of sample size $n=50$ was simulated from an Exponentiated Generalized Exponential (EGE) distribution using a scale parameter value of $= 0.5$ and shape parameters $\alpha = 0.5$ and $\beta = 6.5$. Data obtained were fitted on EGEE, EGE, and E distributions.

Table 4.29: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 50

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	0.6043	3.8054	5.4831	0.5312	-16.9047	41.8094	49.4575
	(11.0214)	(69.4090)	(2.8245)	(0.3890)			
EGE	6.9700	8.2643		2.1964	-17.2633	40.5266	46.2626
	(506.1203)	(2.3496)		(159.500)			
EE			8.2680	0.3150	-17.2633	38.5266	42.3506
			(2.3505)	(0.0394)			
E				1.1610	-42.5320	87.0640	88.9761
				(0.1642)			

Source: Simulated Data from Exponentiated Generalized Exponential Distribution

Table 4.29 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicate that the best distribution is Exponential distribution with values 38.5266 and 42.3506 as compare to E, EGE, and EGEE with AIC values of 87.0640, 40.5266 and 41.8094, while the BIC values are 88.9761, 46.2626 and 49.4575 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.30 below for LRT:

$$LR = 2(-16.9047 - (-17.2633)) = 0.7172$$

$$LR = 2(-16.9047 - (-17.2633)) = 0.7172$$

$$LR = 2(-16.9047 - (-42.5320)) = 51.2546$$

Table 4.30: Likelihood Ratio Test (LRT) Statistic for sample size 50

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: \text{not } H_0$	0.7172	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: \text{not } H_0$	0.7172	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: \text{not } H_0$	51.2546	$p < 0.05$

Source: Simulated data from Exponentiated Generalized Exponential

The computed p-values from Table 4.30 indicates no significant difference between the proposed **EGEE vs. EGE** and **EGEE vs. EGE** distributions, while **EGEE vs. E** is significantly different.

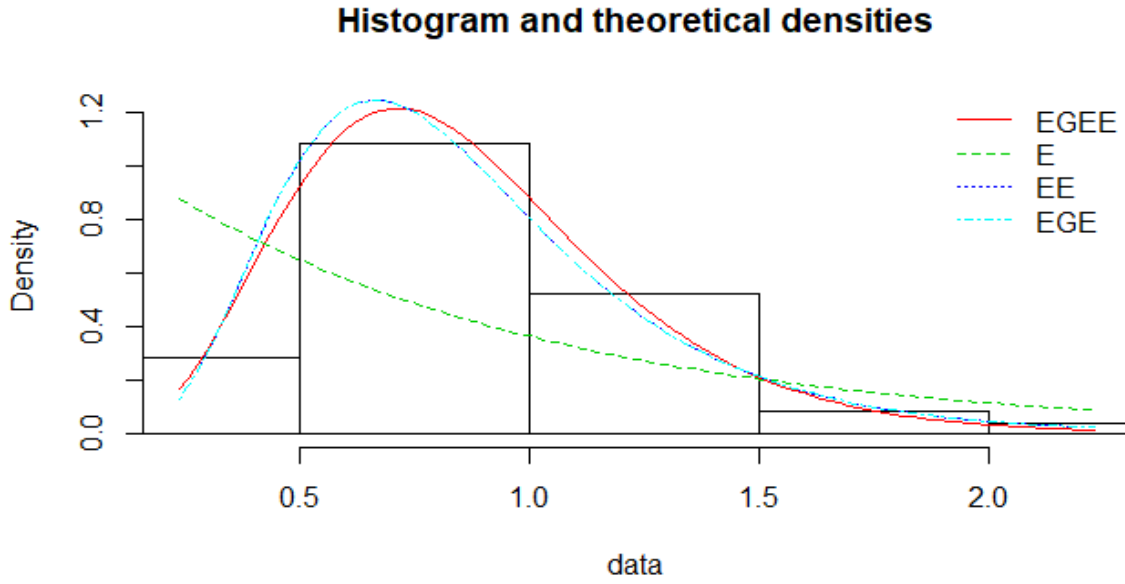


Figure 4.12: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=50$

The histogram of the simulated data set at sample size $n = 50$ along with fitted distribution is shown in figure 4.12. The fitted plots showed closeness amongst Exponentiated Generalized Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.13 Simulated Exponentiated Generalized Exponential for $n=100$

The twelfth simulated data of sample size $n=100$ was simulated from an Exponentiated Generalized Exponential (EGE) distribution using a scale parameter value of $= 0.5$ and shape parameters $\alpha = 0.5$ and $\beta = 6.5$. Data obtained were fitted on EGEE, EGE, and E distributions.

Table 4.31: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 100

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	7.4504	0.2302	7.0793	0.4072	-27.1091	62.2183	72.6389
	(29.1878)	(0.9008)	(3.2849)	(0.2092)			
EGE	-0.2334	9.0199		-0.0704	-27.5192	61.0384	68.8539
	(0.6021)	(2.9292)		(0.190)			
EE			9.9064	0.2887	-27.3936	58.7872	63.9976
			(2.1017)	(0.0252)			
E				1.1900	-82.60425	167.2085	169.8137
				(0.1190)			

Source: Simulated Data from Exponentiated Generalized Exponential Distribution

Table 4.31 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicate that the best distribution is Exponentiated Exponential distribution with values 58.7872 and 63.9976 as compare to E, EGE and EGEE with AIC values 167.2085, 61.0384 and 62.2183, while there BIC values are 169.8137, 68.8539 and 72.6389 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.32 below for LRT:

$$LR = 2(-27.1091 - (-27.5192)) = 0.8202$$

$$LR = 2(-27.1091 - (-27.3936)) = 0.569$$

$$LR = 2(-27.1091 - (-82.60425)) = 110.9903$$

Table 4.32: Likelihood Ratio Test (LRT) Statistic for sample size 100

Model	Hypothesis	LRT	$p - value$
EGEE vs. EGE	$H_0: \beta = 1 \text{ vs. } H_1: \text{not } H_0$	0.8202	$p > 0.005$
EGEE vs. EE	$H_0: \beta = k = 1 \text{ vs. } H_1: \text{not } H_0$	0.569	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1 \text{ vs. } H_1: \text{not } H_0$	110.9903	$p < 0.05$

Source: Simulated data from Exponentiated Generalized Exponential

The computed p-values from Table 4.32 indicates no significant difference between the proposed **EGEE vs. EGE** and **EGEE vs. EGE** distributions, while **EGEE vs. E** is significantly different.

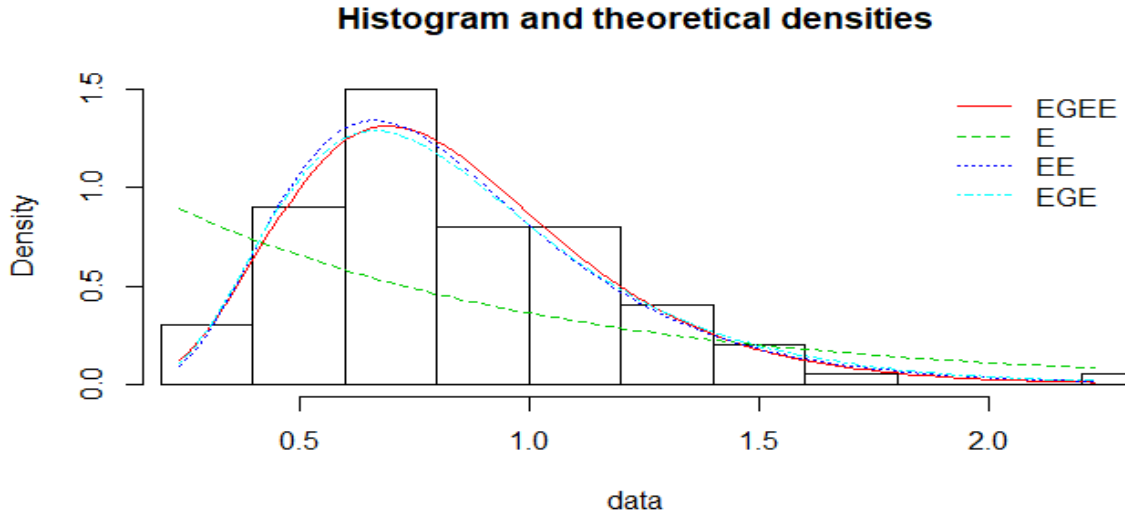


Figure 4.13: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=100$

The histogram of the simulated data set at sample size $n = 100$ along with fitted distribution is shown in figure 4.13. The fitted plots showed closeness amongst Exponentiated Generalized Exponentiated Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.14 Simulated Exponentiated Generalized Exponential for $n=200$

The fourteenth simulated data of sample size $n=200$ was simulated from an Exponentiated Generalized Exponential (EGE) distribution using a scale parameter value of $\theta = 0.5$ and shape parameters $\alpha = 0.5$ and $\beta = 6.5$. Data obtained were fitted on EGEE, EGE, and EE distributions.

Table 4.33: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 200

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	14.3193	0.0705	9.3459	0.2978	-60.9408	129.8817	143.0749
	(48.8454)	(0.2389)	(3.3453)	(0.0884)			
EGE	2.6635	9.4232		0.7876	-60.9410	127.8821	137.7770
	(33.590)	(1.4156)		(9.931)			
EE			9.4278	0.2956	-60.9410	125.8821	132.4787
			(1.4154)	(0.0186)			
E				1.1772	-167.3624	336.7248	340.0231
				(0.0832)			

Source: Simulated Data from Exponentiated Generalized Exponential Distribution

Table 4.33 shows the estimated parameters values for all the aforementioned fitted distributions. The information criteria AIC and BIC indicate that the best distribution is the Exponentiated Exponential distribution with values 125.8821 and 132.4787 as compare to E, EGE, and EGEE with AIC values of 336.7248, 127.8821 and 129.8817, while the BIC values are 340.0231, 137.777 and 143.0749 respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.34 below for LRT:

$$LR = 2(-60.9408 - (-60.9410)) = 0.002$$

$$LR = 2(-60.9408 - (-60.9410)) = 0.002$$

$$LR = 2(-60.9408 - (-167.3624)) = 212.8482$$

Table 4.34: Likelihood Ratio Test (LRT) Statistic for sample size 200

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs $H_1: not H_0$	0.002	p>0.05
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: not H_0$	0.002	p>0.05
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: not H_0$	212.8482	P<0.05

Source: Simulated data from Exponentiated Generalized Exponential

The computed p-values from Table 4.34 indicates no significant difference between the proposed **EGEE vs. EGE** and **EGEE vs. EGE** distributions, while **EGEE vs. E** is significantly different.

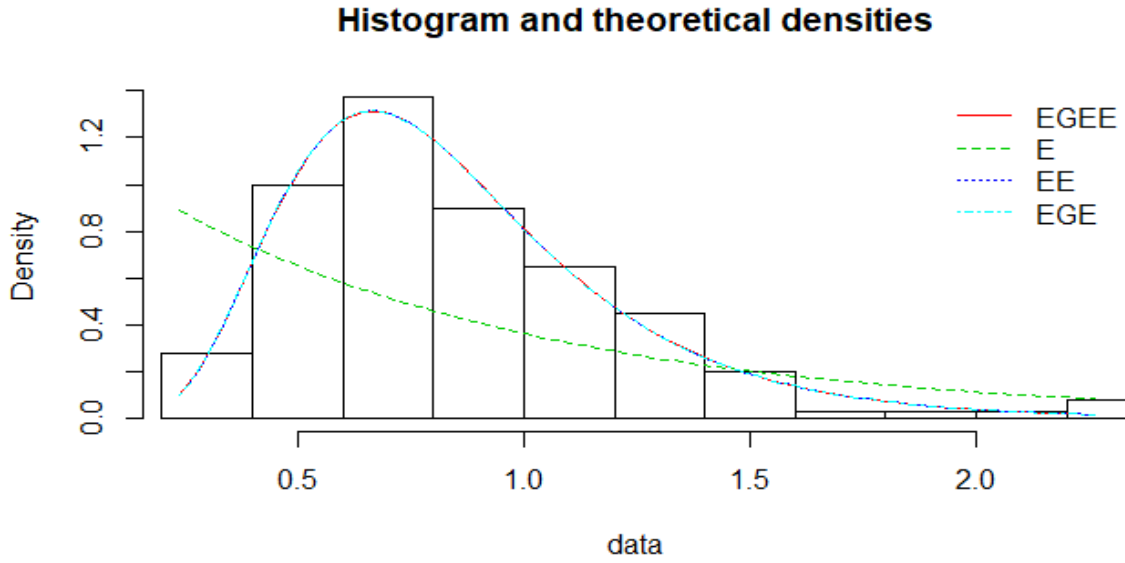


Figure 4.14: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=200$

The histogram of the simulated data set at sample size $n = 200$ along with fitted distribution is shown in figure 4.14. The fitted plots showed closeness amongst Exponentiated Generalized Exponentiated Exponential (EGEE) Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

4.2.15 Simulated Exponentiated Generalized Exponential for $n=350$

The Fifteenth simulated data of sample size $n=350$ was simulated from an Exponentiated Generalized Exponential (EGE) distribution using a scale parameter value of $\theta = 0.5$ and shape parameters $\alpha = 0.5$ and $\beta = 6.5$. Data obtained were fitted on EGEE, EGE and EE distribution

Table 4.35: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Log likelihood and Information Criteria of sample size 350

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	39.2362	0.0120	19.3337	0.1723	-108.9908	225.9817	241.4134
	(20.9545)	(0.0055)	(7.3835)	(0.0338)			
EGE	15.4973	9.4232		4.5994	-112.1637	230.3273	241.9011
	(NA)	(NA)		(NA)			
EE			8.6892	0.29956	-112.1636	228.3273	236.0432
			(0.9944)	(0.0146)			
E				1.1979	-286.792	575.584	579.4419
				(0.0640)			

Source: Simulated Data from Exponentiated Generalized Exponential Distribution

Table 4.35 shows the estimated parameters values for all the aforementioned fitted distributions. The information criterion AIC indicates that the best distribution is the Exponentiated Generalized Exponentiated Exponential distribution with values 225.9817 as compare to E, EE and EGE with AIC values of 575.584, 228.3273 and 230.3273, while the BIC value of EE indicates the best in EE (236.0432) distribution compared to E (579.4419), EGE (241.9011) and EGEE (241.4134) distributions respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as and presented in Table 4.36 below for LRT:

$$LR = 2(-108.9908 - (-112.1637)) = 6.3458$$

$$LR = 2(-108.9908 - (-112.1636)) = 6.3458$$

$$LR = 2(-108.9908 - (-286.792)) = 355.6024$$

Table 4.36: Likelihood Ratio Test (LRT) Statistic for sample size 350

Model	Hypothesis	LRT	<i>p - value</i>
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: not H_0$	6.3458	P<0.05
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: not H_0$	6.3458	P<0.05
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: not H_0$	355.6024	P<0.05

Source: Simulated data from Exponentiated Generalized Exponential

The computed p-values from Table 4.36 indicates significant differences between the proposed **EGEE vs. EGE**, **EGEE vs. EE**, and **EGEE vs. E** distributions.

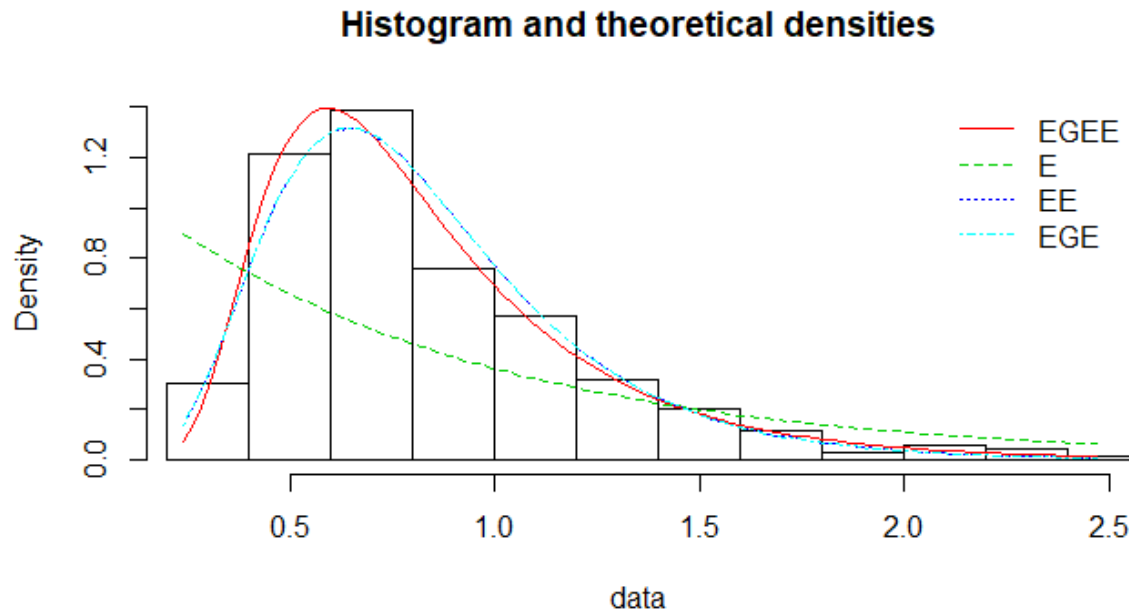


Figure 4.15: Fitted Plot of EGEE EGE, EE and E Models for sample size $n=350$

The histogram of the simulated data set at sample size $n = 350$ along with fitted distribution is shown in Figure 4.15. The fitted plots showed the best distribution in Exponentiated Generalized Exponentiated Exponential (EGEE) distribution with a very high peak compared to Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions. However, the fits of Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) fit the data very closely, and this is confirmed by the values of the log-likelihood (LL).

Table 4.37: Summary Table on the Performance of Fitted Distributions to Simulated data from Exponentiated Generalized Exponential Distribution

sample size		Rank		
	MEASURED STATISTIC	1 st	2 nd	3 rd
10	AIC	EE	EGEE	E
	BIC	EE	EGEE	E
	Log-likelihood	EGEE	EE	E
50	AIC	EE	EGEE	E
	BIC	EE	EGEE	E
	Log-likelihood	EGEE	EE	E
100	AIC	EE	EGEE	E
	BIC	EE	EGEE	E
	Log-likelihood	EGEE	EE	E
200	AIC	EE	EGEE	E
	BIC	EE	EGEE	E
	Log-likelihood	EGEE	EE	E
350	AIC	EGEE	EE	E
	BIC	EE	EGEE	E
	Log-likelihood	EGEE	EE	E

Source: Simulated data from Exponentiated Generalized Exponential Distribution

Table 4.37 shows the ranking in performance of the fitted EGEE, EGE, and E distributions on simulated data from EGE distribution. The ranks 1st, 2nd, and 3rd indicate the order of performance of the distributions at various sample sizes n=10, 50, 100, 200 and 350 using the values of AIC, BIC, and log-likelihood. For sample size n=10, 50, 100 and 200 the reported AIC and BIC show that the EE distribution outperformed the EGEE and E distributions, while the log-likelihood indicate the best in EGEE distribution. AIC and log-likelihood values for sample sizes of 350 show an order in performances as EGEE, EE and E distributions, while the BIC showed an order of EE, EGEE, and E distributions respectively. Sample size n= 350 shows that the proposed EGEE distribution can serve as the best alternative distribution to data simulated from EGE distribution, and these can be confirmed from the AIC and log-likelihood values.

4.2.16 Simulated Beta distribution for n=20

The sixteenth simulated data of sample size $n=20$ was simulated from a beta distribution using different combination of shape and scale parameters to obtain the skewness values of -2.8494, -1.2891, 0.0009, 0.8249 and 2.2318. Data obtained were fitted to EGEE, EGE, EE and E distribution.

Table 4.38: Estimated log-likelihood values of simulated data from beta distribution at sample size $n=20$

	SK= -2.8494	SK= -1.2891	SK= 0.0009	SK= 0.8249	SK= 2.2318
Distribution	LL	LL	LL	LL	LL
EGEE	-2.0751	16.8499	4.9630	14.3812	22.3024
EGE	-10.8340	10.5456	3.7605	13.7946	22.2518
EE	-10.8340	10.5456	3.7605	13.7946	22.2518
E	-18.3309	-17.9300	-6.2987	10.6527	22.0744

Table 4.38 above shows the estimated log-likelihood values of the fitted distributions to the data simulated from beta distribution at different level of skewness levels of -2.8494, -1.2891, 0.0009, 0.8249 and 2.2318. The reported values show that the proposed distribution gave the best estimated values as compared to the existing distributions.

4.2.17 Simulated Beta distribution for n=50

The seventeenth simulated data of sample size $n=50$ was simulated from a beta distribution using different combination of shape and scale parameters to obtain the skewness values of -4.7567, -1.1317, 0.1266, 1.7086 and 2.3126. Data obtained were fitted to EGEE, EGE, EE and E distribution.

Table 4.39: Estimated log-likelihood values of simulated data from beta distribution at sample size n=50

Distribution	SK=-4.7567	SK=-1.1317	SK= 0.1266	1.7086	2.3126
	LL	LL	LL	LL	LL
EGEE	53.6868	12.6996	12.1867	128.0601	NA
EGE	9.2858	0.8791	8.4056	127.4537	NA
EE	9.2858	0.8792	8.4057	127.4537	NA
E	-48.5694	-40.7264	-14.6947	78.84422	NA

Table 4.39 depicts the estimated values of sample size 50 generated from beta distribution at various levels of skweness. The estimated log-likelihood values shows that at negative skweness values of -4.7567 and -1.1317, there are significant differences among the distributions with the values of 53.6868, 9.2858, 9.2858 and -48.5694 for **EGEE**, **EGE**, **EE** and **E** distributions respectively. The large gap (53.6868 - 9.2858) shows that the proposed distribution has a significant likelihood value when fitting a negative skewed data.

4.2.18 Simulated Beta distribution for n=100

The eighteenth simulated data of sample size n=100 was simulate from a beta distribution using different combination of shape and scale parameters to obtain the skweness values of -4.9176, -1.9879, 0.2562, 1.8227 and 2.2448. Data obtained were fitted to EGEE, EGE, EE and E distribution.

**Table 4.40: Estimated log-likelihood values of simulated data from beta distribution at
Sample size n=100**

Distribution	-4.9176	-1.9879	0.2562	1.8227	2.2448
	LL	LL	LL	LL	LL
EGEE	213.5726	88.2975	30.0529	159.9254	215.1940
EGE	109.2245	42.7350	25.3219	159.9243	215.1532
EE	109.2246	42.7350	25.3219	159.9243	215.1532
E	-98.8102	-91.6580	-29.4668	120.2888	174.0229

Table 4.39 shows the estimated values of sample size 100 generated from beta distribution at various levels of skweness. The estimated log-likelihood values shows that at negative values of -4.9176 and -1.9879 there are significant differences among the distributions with the values of 213.5726, 109.2245, 109.2246 and -98.8102 for **EGEE**, **EGE**, **EE** and **E** distributions respectively. The large gaps (213.5726, 109.2245, 109.2245 and- -98.8102) shows that the proposed distribution has a significant likelihood value when fitting a negative skewed data.

4.3 Application

In this section, the proposed EGEE distribution is used to fit four (4) real-life data sets and estimates obtained were used to compare the fitted values of EGE, EE, and E distributions, respectively.

4.3.1 First Data set

The first data set is a Nichols and Padgett, (2006) uncensored data consisting of 100 observations as given in Table 4.41 on breaking stress of carbon fibrers (in Gba).

Table 4.41: 100 Observations on Breaking Stress of Carbon Fibers (in Gba)

3.70	4.42	3.75	3.15	3.31	1.41	2.17	1.59	2.48	2.03
2.74	2.41	2.43	2.35	3.31	3.68	1.17	2.00	0.85	1.80
2.73	3.19	2.95	2.55	2.85	2.97	5.08	1.22	1.61	1.57
2.50	3.22	2.97	2.59	2.56	1.36	2.48	1.12	2.79	1.08
3.60	1.69	3.39	2.38	3.56	0.98	1.18	1.71	4.70	2.03
3.11	3.28	2.96	2.81	2.81	2.76	3.19	1.84	3.51	1.61
3.27	3.09	2.53	4.20	2.77	4.91	1.57	3.65	2.17	2.12
2.87	1.87	2.67	3.33	2.17	3.68	0.81	2.05	1.69	1.89
1.47	3.15	2.93	2.55	2.83	1.84	5.56	0.39	1.25	2.88
3.11	4.90	3.22	3.39	1.92	1.59	1.73	3.68	4.38	2.82

Source: Nichols and Padgett (2006)

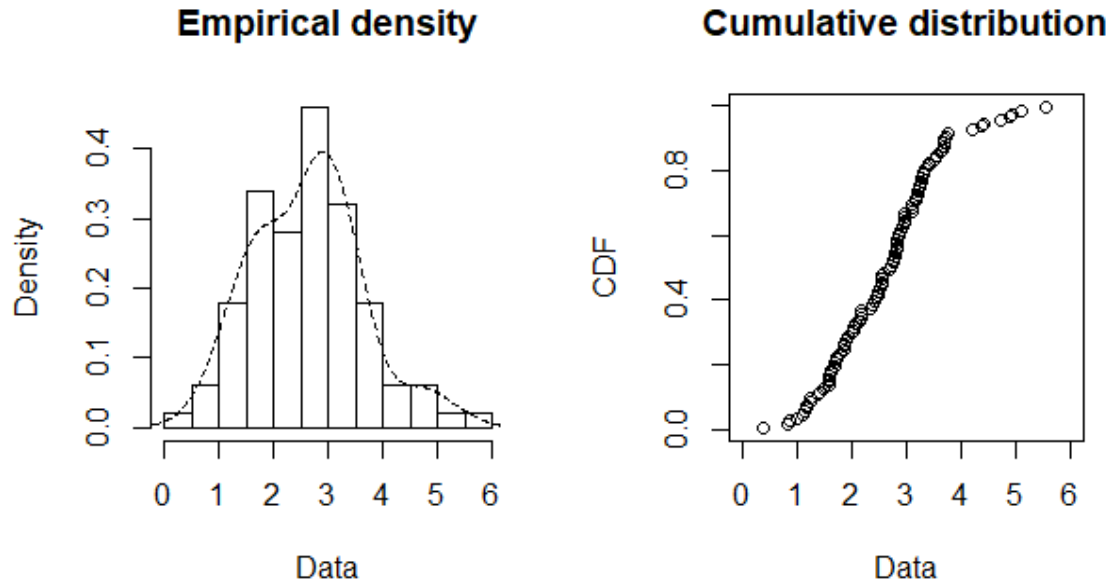


Figure 4.16: Histogram and CDF Plots of an Empirical Distribution for Fibers Data

Table 4.39: Summary Statistics for Carbon Fibers Data

Min	Max	Median	Mean	S.D.	Skewness	Kurtosis
0.3900	5.5600	2.700	2.6214	1.0139	0.3738	3.1729

The statistical summary table from Table 4.42 shows that the fibers data is under-disperse, non-symmetry, and right-skewed.

The Correlation matrix of the parameters of EGEE distribution estimated from the Carbon Fibres Data is:

$$\begin{pmatrix} 1.0000000 & -0.8855513 & -0.2215622 & 0.2383665 \\ -0.8855513 & 1.0000000 & -0.2214452 & 0.2382406 \\ -0.2215622 & -0.2214452 & 1.0000000 & -0.9543851 \\ 0.2383665 & 0.2382406 & -0.9543851 & 1.0000000 \end{pmatrix}$$

The above correlation matrix shows the relationship between the combine sets of parameters. The off-diagonal entries coefficients with values -0.8855513, -0.2215622 and 0.2383665 shows the relationship between parameter (α, β) , (α, k) and (α, θ) respectively, the values -0.2214452 and 0.2382406 depicts the relationship of β parameter with k and θ . The final entry of -0.9543851 indicates a negative relationship between parameters k and θ .

Table 4.40: Maximum Likelihood Estimate of Parameters Standard Errors in Parenthesis, Loglikelihood and Information Criteria carbon Fibers

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	11.3312	5.8340	3.2679	9.0031	-141.318	290.6361	301.0557
	(88.1057)	(45.3608)	(0.8815)	(15.6154)			
EGE	2.3577	7.7871		2.327	-146.182	298.3646	306.1801
	(102.0782)	(1.5042)		(100.7249)			
EE			7.79061	0.9868	-146.182	296.3646	301.575
			(1.4966)	(0.0851)			
E				0.3814	-196.371	394.7417	397.3469
				(0.0381)			

Table 4.43 shows the reported values of log-likelihood function (LL) estimates for EGEE, EGE, EE and E distributions obtained as -141.318, -146.182, -146.182 and -196.371 respectively with Akaike Information Criterion (AIC) values of 290.6361, 298.3646, 296.3646, and 394.7417. The Bayesian Information Criterion (BIC) indicates the values of 301.0557, 306.801, 301.575, and 397.3469, respectively, for all the distributions, as mentioned above. The estimated result confirmed that the proposed distribution EGEE with values of -141.318, 290.6361 and 301.0557 for LL, AIC, and BIC respectively gives a better distribution fit for the data set. The EGE and EE distribution performed equally on the data set as the likelihood values confirm it in Table 4.43.

The asymptotic variance-covariance matrix for the estimated parameters is presented below as:

$$I_{ij}^{-1} = \begin{pmatrix} 7762.6178 & -3539.1498 & -17.207477 & 327.94615 \\ -3539.1498 & 2057.60624 & -8.854517 & 168.75244 \\ -17.207477 & -8.854517 & 0.777025 & -13.13694 \\ 327.94615 & 168.75244 & -13.13694 & 243.84074 \end{pmatrix}$$

The variances of the distributions based on maximum likelihood estimate (MLE) for the parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{k}$ and $\hat{\theta}$ are 7762.6178, 2057.6062, 0.7770 and 243.8407 with

corresponding 95% confidence intervals values of (-161.356, 184.0184), (-83.0733, 94.7413), (1.5402, 4.9956) and (-21.6031, 39.6093) respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as:

$$LR = 2(-141.318 - (-146.182)) = 9.7800$$

$$LR = 2(-141.318 - (-146.182)) = 9.7800$$

$$LR = 2(-141.318 - (-196.37)) = 110.1040$$

Table 4.41: Likelihood Ratio Test (LRT) Statistic for Carbon Fibers

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: H_0$ is not equal	9.7800	< 0.00001
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: H_0$ is not equal	9.7800	< 0.00001
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: H_0$ is not equal	110.1040	< 0.00001

Table 4.44 depicts the estimated (LRT) statistic for EGEE vs EGE, EGEE vs EE and EGEE vs E distributions of reported values 9.7800, 9.7800 and 110.1040 respectively and the corresponding p-values of 0.001, 0.001 and 0.00001. The finding reveals that **EGEE** distribution is significant at 0.05 level of significance to EGE, EE, and E distributions.

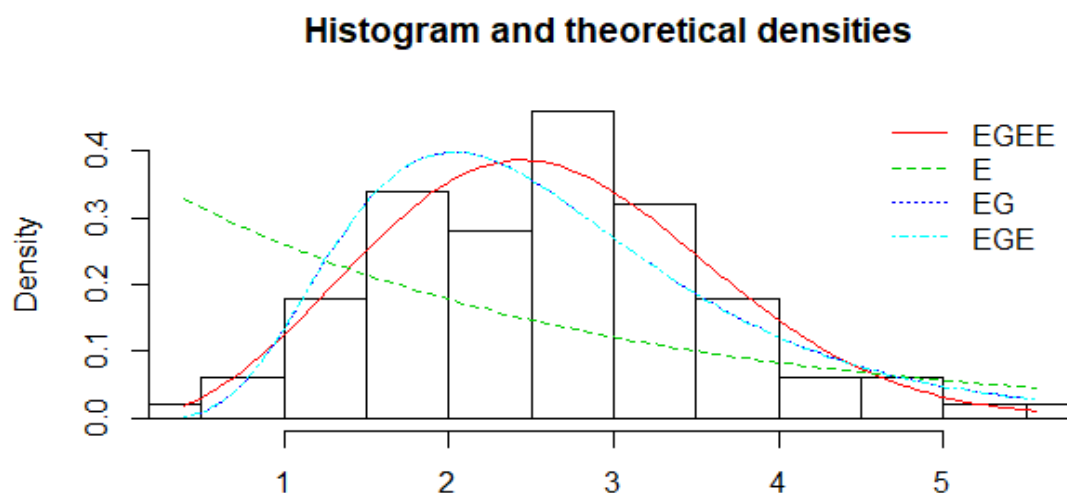


Figure 4.17: Fitted Plot of EGEE EGE, EE and E Models on Fibers Data

The histogram of the fibers data set at along with fitted distribution is shown in figure 4.17. The theoretical density of EGEE distribution depicts the best spread compared to EGE, EE, and E distributions on the data set.

4.3. 2 Second Data Set

The second data presented in Table 4.45 is on total milk production in the first birth of 107 cows from Sindi race and are the properties of Carnauba Farm which belong to Agropecuaria Manoel Danta Ltda (AMDA) located in Taperoa City Paraiba(Brazil) gotten from Cordeiro and Binto (2012).

Table 4.42: The Data Set Represent the Total Milk Production in the First Birth of 107 Cows from SINDI Race

0.43650	0.4260	0.5140	0.6907	0.7471	0.2605	0.6196	0.8781	0.4990	0.6058	0.7629
0.6891	0.5770	0.5394	0.1479	0.2356	0.6012	0.1525	0.5483	0.6927	0.7261	0.5941
0.3323	0.0671	0.2361	0.4800	0.5707	0.7131	0.5853	0.6768	0.5350	0.4151	0.6174
0.6789	0.4576	0.3259	0.2303	0.7687	0.4371	0.3383	0.6114	0.3480	0.4564	0.6860
0.7804	0.3406	0.4823	0.5912	0.5744	0.5481	0.1131	0.7290	0.0168	0.5529	0.0609
0.4530	0.3891	0.4752	0.3134	0.3175	0.1167	0.6750	0.5113	0.5447	0.4143	0.6488
0.5627	0.5150	0.0776	0.3945	0.4553	0.4470	0.5285	0.5232	0.6465	0.0650	0.2747
0.8492	0.8147	0.3627	0.3906	0.4438	0.4612	0.3188	0.2160	0.6707	0.6220	0.4741
0.5629	0.4675	0.6844	0.3413	0.4332	0.0854	0.3821	0.4694	0.3635	0.4111	0.3598
0.5349	0.3751	0.1546	0.4517	0.2681	0.4049	0.5553	0.5878			

Source: Cordeiro and Brito (2012)

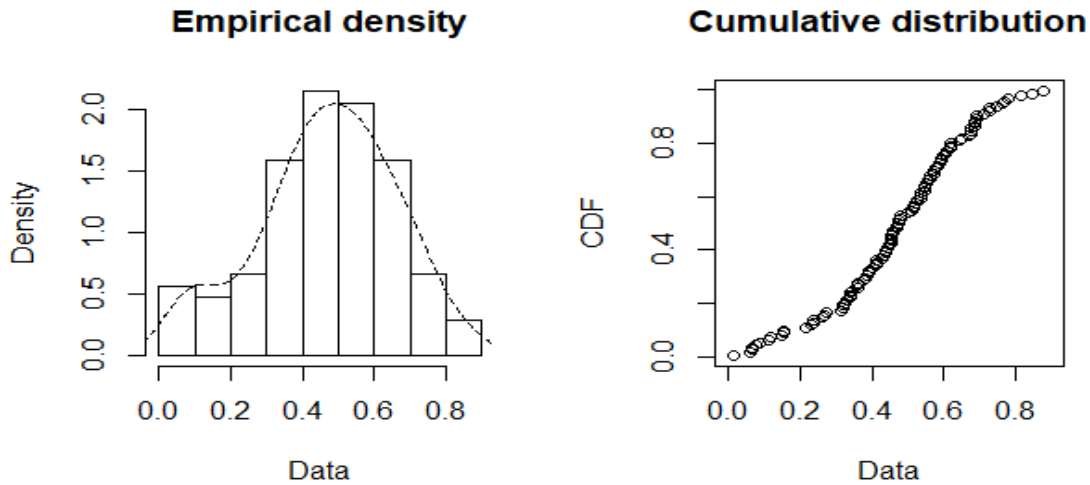


Figure 4.18: Histogram and CDF Plots of an Empirical Distribution for Cows Data

Table 4.43: Summary Statistics for Cow Data

Min	Max	Median	Mean	S.D.	Skewness	Kurtosis
0.0168	0.8781	0.4741	0.4689	0.1919	-0.3306	-0.3639

Table 4.46 presents the summary statistics of the cow data which is non-symmetry and left skewed.

The Correlation matrix of EGEE distribution estimated from the Cow Data is::

$$\begin{pmatrix} 1.000000 & 0.5508470 & -0.5424908 & 0.861898 \\ 0.5508470 & 1.000000 & -0.5423818 & 0.8617364 \\ -0.5424908 & -0.5423818 & 1.000000 & -0.7636886 \\ 0.861898 & 0.8617364 & -0.7636886 & 1.000000 \end{pmatrix}$$

Correlation matrix above shows the relationship between combine sets of parameters. The coefficients show both positive and negative values depending on the combine parameters. Moreover, this allows the measures of relationships amongst the parameters.

**Table 4.44: Maximum Likelihood Estimate of Parameters Standard Errors in
Parenthesis, Loglikelihood and Information Criterion of Cow data**

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	78.11548	302.417	2.5343	27.4904	21.0898	-34.1796	-23.4883
	(159.4893)	(617.053)	(0.295)	(15.6154)			
EGE	13.7707	3.7135		3.2788	5.03874	-4.07749	3.9409
	(NA)	(NA)		(NA)			
EE			3.7146	0.23801	5.038748	-6.07749	-0.73184
			(0.5657)	(0.0211)			
E				2.132872	-25.9508	53.9015	56.57438
				(0.20619)			

Table 4.47 presents the estimated values of EGEE, EGE, EE and E distribution for log-likelihood (LL) function obtained as 21.0898, 5.03874, 5.03875 and -25.9508 respectively with Akaike Information Criterion (AIC) values of -34.1796, -4.07749, -6.07749 and 53.9015. The Bayesian Information Criterion (BIC) indicates the values of -23.4883, 3.94091, -0.73184 and 56.5744, respectively, for all the distributions mentioned above. The estimated result confirmed that the proposed distribution EGEE with values of 21.0898, -34.1796 and -23.4883 for LL, AIC, and BIC respectively, therefore, give a better distribution fit for the data set.

The asymptotic variance covariance matrix for the estimated parameters is.

$$I_{ij}^{-1} = \begin{pmatrix} 25436.86163 & 54210.79981 & -25.56504711 & 6626.11145 \\ 54210.79981 & 380755.4731 & -98.88969 & 25631.17162 \\ -25.56504711 & -98.88969 & 0.0873062 & -10.87702 \\ 6626.11145 & 25631.17162 & -10.87702 & 2323.49581 \end{pmatrix}$$

The variances of the distributions based on maximum likelihood estimate (MLE) for the parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{k}$ and $\hat{\theta}$ are 25436.86163, 380755.4731, 0.0873062 and 2323.49581 with corresponding 95% confidence intervals values of (-234.44837, 390.7147) (-907.0086, 1511.8426), (1.9552, 3.11343) and (-66.9808, 121.973601) respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as:

$$LR = 2(21.0898 - (5.03875)) = 33.0100$$

$$LR = 2(21.0898 - (5.03875)) = 33.0100$$

$$LR = 2(21.0898 - (-25.9508)) = 94.0920$$

Table 4.48: Likelihood Ratio Test (LRT) Statistic for Cow Data

Model	Hypothesis	LRT	<i>p – value</i>
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: H_0$ is not equal	33.0100	< 0.00001
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: H_0$ is not equal	33.0100	< 0.00001
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: H_0$ is not equal	94.0920	< 0.00001

Table 4.48 depicts the estimated (LRT) statistic for EGEE vs. EGE, EGEE vs. EE and EGEE vs. E distributions of reported values 33.0100, 33.0100 and 94.0920 respectively with corresponding p-values of 0.0001, 0.0001 and 0.000001. The finding reveals that EGEE distribution is significant at 0.05 level of significance to EGE, EE, and E distributions.

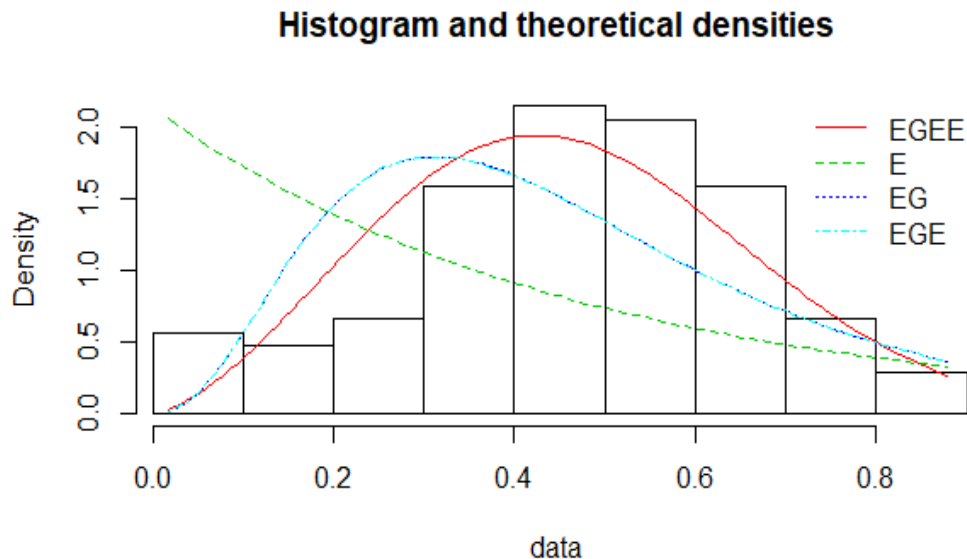


Figure 4.19: Fitted Plot of EGEE EGE, EE and E Distributions on Cow Data

Figure 4.19 shows the histogram of the cow data set along with fitted distributions. The theoretical density of EGEE distribution has a better spread compared to its sub-distribution (EGE, EE, and E) on the data set. The EGE and EE distribution performed equally on the data set as the likelihood values confirm it in Table 4.21.

4.3.3: Third Data Set

The third data were obtained by workers at UK National Physical laboratory on strengths of 1.5cm glass fibers. This data set has previously been used by Oguntunde et al. (2017), Smith and Naylor (1987), Bourguignon et al. (2014) and Merovei et al. (2013).

Table 4.49: The data relating to the strengths of 1.5cm glass fibers which was obtained by workers at UK National Physical laboratory

0.55	0.74	0.77	0.81	0.84	1.24	0.93	1.04	1.11	1.13	1.3
1.25	1.27	1.28	1.29	1.48	1.36	1.39	1.42	1.48	1.51	1.49
1.49	1.50	1.50	1.55	1.52	1.53	1.54	1.55	1.61	1.58	1.59
1.60	1.61	1.63	1.61	1.61	1.62	1.62	1.67	1.64	1.66	1.66
1.66	1.70	1.68	1.68	1.69	1.70	1.78	1.73	1.76	1.76	1.77
1.89	1.81	1.82	1s.84	1.84	2.00	2.01	2.24			

Source: Bourguignon et al. (2014).

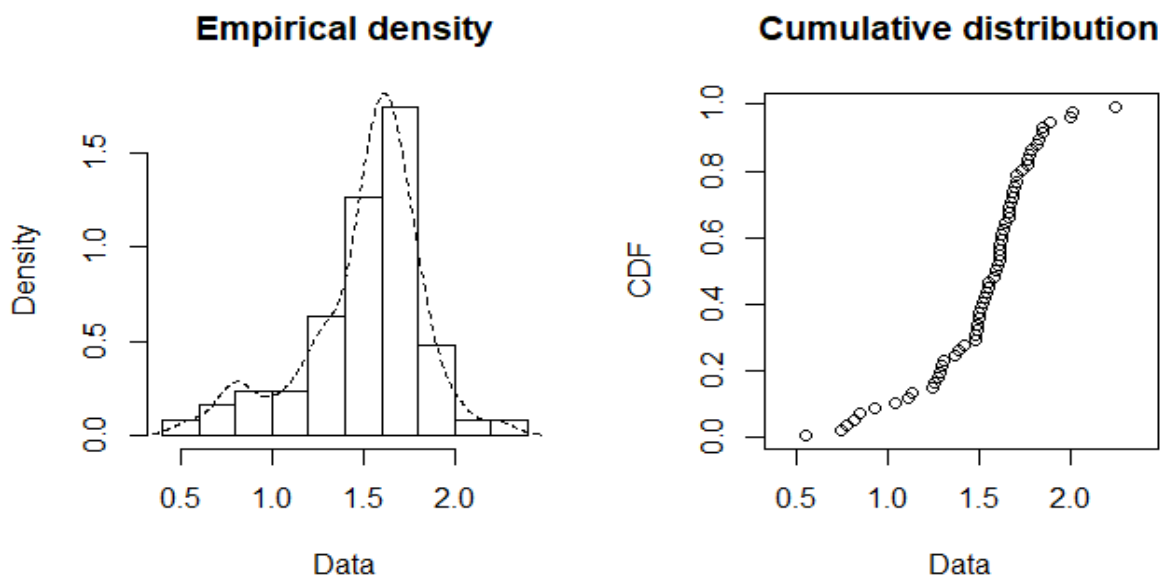


Figure 4.20: Histogram and CDF Plots of an Empirical Distribution for Glass Fibre Data

Table 4.47: Summary Statistics Glass Fibre

Min	Max	Median	Mean	S.D.	Skewness	Kurtosis
0.550	2.240	1.590	1.507	0.324	-0.878	0.800

The statistical summary table from table 4.50 shows that the glass fiber data is non-symmetry and left-skewed.

The Correlation matrix of EGEE distribution estimated from the Glass Fibres Data is:

$$\begin{pmatrix} 1.0000000 & -0.0940118 & -0.4069278 & 0.6071182 \\ -0.0940118 & 1.0000000 & -0.4069282 & 0.6071190 \\ -0.4069278 & -0.4069282 & 1.0000000 & -0.8851251 \\ 0.6071182 & 0.6071190 & -0.8851251 & 1.0000000 \end{pmatrix}$$

Correlation matrix above shows the relationship between combine sets of parameters. The coefficients exhibit both positive and negative values depending on the combine parameter, and this allows us to see which pair has a negative and positive correlation.

Table 4.48: Maximum Likelihood Estimate of Parameters, Standard Errors in Parenthesis, Loglikelihood and Information Criterion Glass fibers data

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	3.2594	45.663	7.710	2.188	-16.912	41.824	50.3971
	(2.705)	(37.900)	(1.6465)	(0.8228)			
EGE	64052.895	10.712		36621.219	-38.5517	83.1035	89.539
	(7278.081)	(3.1496)		(6444.711)			
EE			31.3627	0.3828	-31.3834	66.7669	71.0532
			(9.5330)	(0.0349)			
E				0.6636	-88.8303	179.6606	181.8038
				(0.0836)			

Table 4.48 presents the estimated values of EGEE, EGE, EE and E distribution for log-likelihood (LL) function obtained as -16.9120,-38.5517, -31.3834 and -88.8303 respectively with Akaike Information Criterion (AIC) values of 41.8240, 83.1035,66.7669 and 6606..

The Bayesian Information Criterion (BIC) indicates the values of 50.3971, 89.539, 71.0532, and 181.8038, respectively, for all the distributions, as mentioned earlier. The estimated result confirmed that the proposed distribution EGEE with values of -16.912, 41.824 and 50.3971 for LL, AIC, and BIC respectively and therefore gives the best fit for the data set.

The asymptotic variance covariance matrix for the estimated parameters is

$$I_{ij}^{-1} = \begin{pmatrix} 7.318780 & -9.639312 & -1.812658 & 1.3514472 \\ -9.639312 & 1436.44403 & -25.394599 & 18.9332283 \\ -1.812658 & -25.394599 & 2.711180 & -1.1991964 \\ 1.3514472 & 18.9332283 & -1.1991964 & 0.6770378 \end{pmatrix}$$

The variances of the distributions based on maximum likelihood estimate (MLE) for the parameters $\hat{\alpha}, \hat{\beta}, \hat{k}$, and $\hat{\theta}$ are 7.318780, 436.44403, 2.711180 and 0.6770378 with corresponding 95% confidence intervals values of (-2.0404, 8.5618), (-28.6219, 119.9479), (4.4827, 10.9373) and (0.5753, 3.8001) respectively were obtained.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as:

$$LR = 2(-16.912 - (-38.5517)) = 43.2794$$

$$LR = 2(-16.912 - (-31.3834)) = 28.9428$$

$$LR = 2(-16.912 - (-88.8303)) = 143.8366$$

Table 4.52: Likelihood Ratio Test (LRT) Statistic for Glass Fibres

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: H_0$ is not equal	43.2794	< 0.00001
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: H_0$ is not equal	28.9428	< 0.00001
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: H_0$ is not equal	143.8366	< 0.00001

Table 4.52 depicts the estimated (LRT) statistic for EGEE vs. EGE, EGEE vs. EE and EGEE vs. E distributions of reported values 43.2794, 28.9428 and 143.8366 respectively and the corresponding p-values of 0.0001, 0.0001 and 0.000001. The finding reveals that EGEE distribution is significant at 0.05 level of significance to EGE, EE, and E distributions.

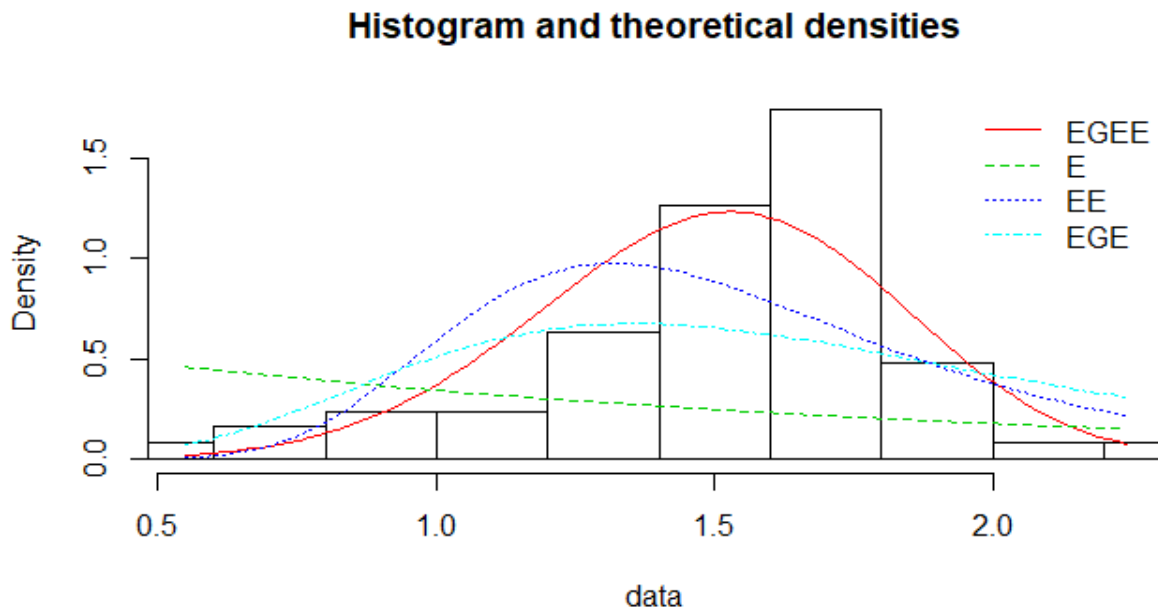


Figure 4.21: Fitted Plot of EGEE EGE, EE and E Models on Glass Fibres Data

The histogram of the glass fibers data set at along with fitted distribution is shown in figure 4.21. The spread of the theoretical density of EGEE distribution best described the data set than its sub-distributions.

4.3.4: Fourth Data Set

The Fourth data set is on 30 observations of precipitation in inches for March 1977 at Minneapolis/ St. Paul as obtained by Hinkley (1977) (2016)

Table 4.53: Data set of 30 observations of March precipitation (in inches) Minneapolis/St. Paul

0.77	1.74	0.81	1.20	1.95	1.20	0.47	1.43	3.37	2.20
3.00	3.09	1.51	2.10	0.52	1.62	1.31	0.32	0.59	0.81
2.81	1.87	1.18	1.35	4.75	2.48	0.96	1.89	0.90	2.05

Source: Hinkley, (1977)

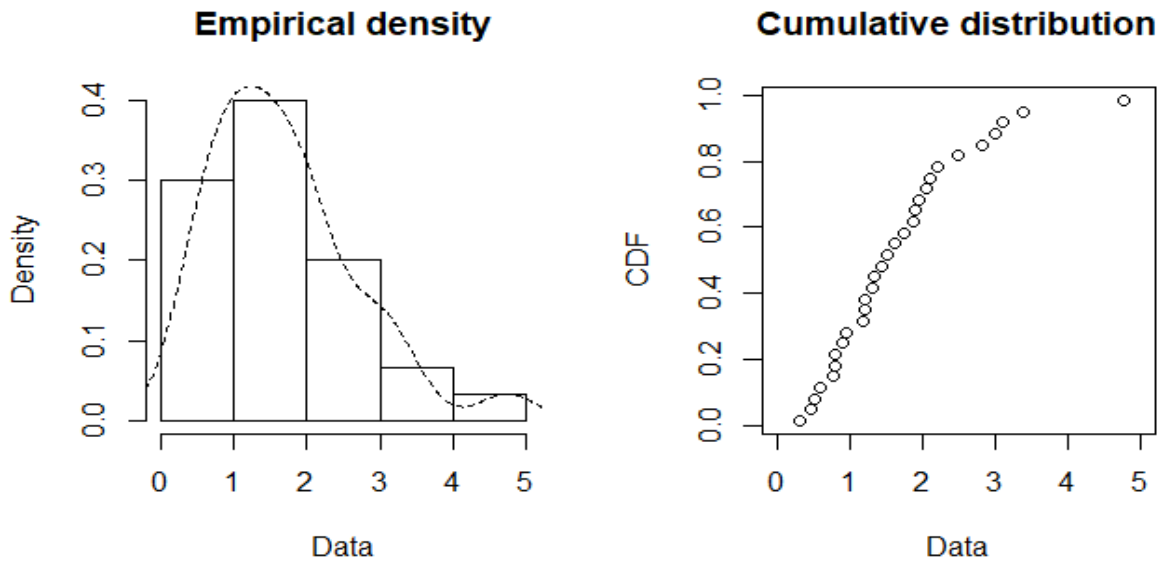


Figure 4.22: Histogram and CDF Plots of an Empirical Distribution for precipitation in Minneapolis/St Paul

Table 4.54: Summary Statistics for precipitation in Minneapolis/St Paul

Min	Max	Median	Mean	S.D.	Skewness	Kurtosis
0.320	4.750	1.470	1.675	1.000	1.0323	0.931

Table 4.54 present the summary statistics of the precipitation in Minneapolis/St Paul data which is -symmetry and right skewed.

The Correlation matrix of EGEE distribution estimated from the Precipitation Data is:

$$\begin{pmatrix} 1.0000000 & -0.9681910 & -0.08597274 & 0.09223604 \\ -0.9681910 & 1.0000000 & -0.1473026 & 0.15803392 \\ -0.08597274 & -0.1473026 & 1.0000000 & -0.95555314 \\ 0.09223604 & 0.15803392 & -0.95555314 & 1.0000000 \end{pmatrix}$$

Correlation matrix above shows the relationship between combine sets of parameters. The coefficients indicate both positive and negative values depending on the combine parameters, and this allows us to see which pairs of parameters are negative and positively correlated.

Table 4.52: Maximum Likelihood Estimate of Parameters, Standard Errors in Parenthesis, Log likelihood and Information Criterion on Precipitation data

Models	$\hat{\alpha}$	$\hat{\beta}$	$\hat{k}$	$\hat{\theta}$	LL	AIC	BIC
EGEE	0.1193253	9.191329	3.3437	0.9196	-38.0932	84.1863	89.7911
	(0.9536734)	(74.08602)	(2.6418)	(1.4285)			
EGE	3.63214	3.4608		3.11249	-38.09426	82.1852	86.3921
	(533.6998)	(1.0645)		(457.3723)			
EE			3.46051	0.3828	-38.0942	80.1885	82.9909
			(0.856786)	(0.15765)			
E				0.59701	-45.4743	92.9487	94.3499
				(0.1089)			

Table 4.55 presents the estimated values of EGEE, EGE, EE and E distribution of log-likelihood (LL) function obtained as -38.0932, -38.09426, -38.0942 and -45.4743 respectively, with Akaike Information Criterion (AIC) values of 84.1863, 82.1852, 80.1885 and 92.9487. The Bayesian Information Criterion (BIC) indicates the values of 89.7911, 86.3921, 82.9909, and 94.3499, respectively, for all the distributions, as mentioned earlier. The estimated result shows that the proposed EGEE distribution performed relative close with EGE and EE distributions with the fitted data set.

The asymptotic variance covariance matrix for the estimated parameters is.

$$I_{ij}^{-1} = \begin{pmatrix} 0.9094930 & 68.40644 & -0.216602 & 0.125660 \\ 68.40644 & 5488.73809 & -28.8303310 & 16.725724 \\ -0.216602 & -28.8303310 & 6.9792054 & -3.6061709 \\ 0.125660 & 16.725724 & -3.6061709 & 2.0407827 \end{pmatrix}$$

The variances of the distributions based on maximum likelihood estimate (MLE) for the parameters $\hat{\alpha}$, $\hat{\beta}$, $\hat{k}$ and $\hat{\theta}$ are 0.909493, 5488.73809, 6.9792054 and 2.0407827 with corresponding 95% confidence intervals values of (-1.7512, 1.9811), (136.0173, 154.3999), (-1.8340, 8.5237) and (-1.880396, 3.7196) respectively.

The LRT for the proposed EGEE distribution as against EGE, EE and E distributions are computed as:

$$LR = 2(-38.0932 - (-38.09426)) = 0.00212$$

$$LR = 2(-38.0932 - (-38.09426)) = 0.00212$$

$$LR = 2(-38.0932 - (45.4743)) = 14.7622$$

Table 4.56: Likelihood Ratio Test (LRT) Statistic for Precipitation

Model	Hypothesis	LRT	<i>p</i> – value
EGEE vs. EGE	$H_0: \beta = 1$ vs. $H_1: H_0$ is not equal	0.00212	$p > 0.05$
EGEE vs. EE	$H_0: \beta = k = 1$ vs. $H_1: H_0$ is not equal	0.00212	$p > 0.05$
EGEE vs. E	$H_0: \beta = k = \alpha = 1$ vs. $H_1: H_0$ is not equal	14.7622	$P < 0.05$

The LRT statistic in Table 4.56 shows that the EGEE vs. EGE and EGEE vs. EE distributions are all suited for the data set, since the p-values are not significant, while the EGEE vs. E distribution indicates a significant value.

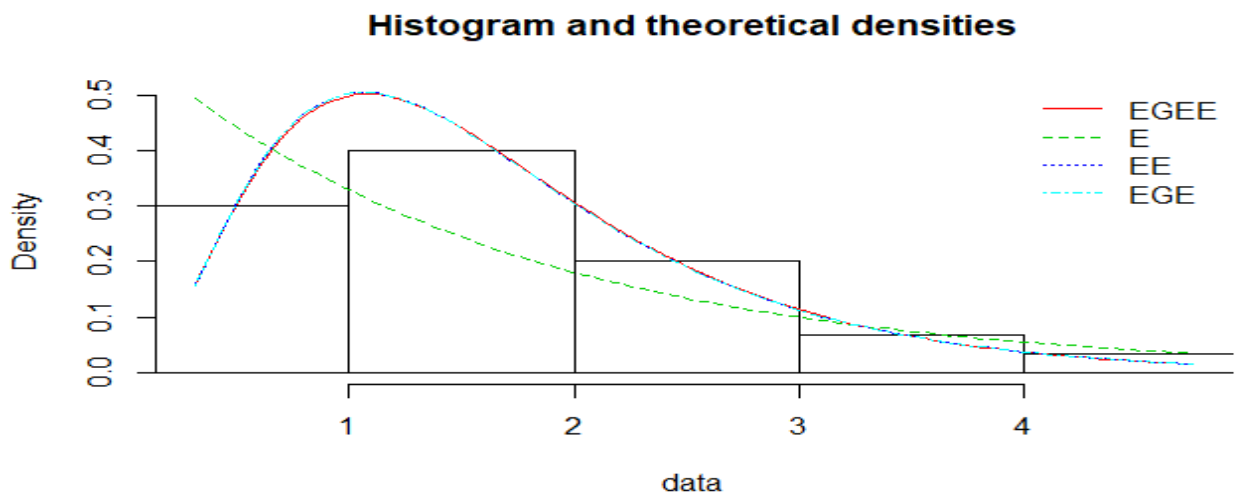


Figure 4.23: Fitted Plot of EGEE EGE, EE and E Models on precipitation data

The histogram of the precipitation data set along with fitted distribution is shown in Figure 4.23. The theoretical density of EGEE, EGE, and EE distributions show that the three distributions show no significant difference when fitting some right-skewed data sets.

CHAPTER FIVE

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

5.1 Introduction

This chapter focuses on the overview of the methodology and results of the findings, conclusion, and recommendations. Suggestions for further research and contribution to knowledge are also presented. Besides, the significance of the results obtained from the simulated and real-life data applications are discussed in detail.

5.2 Summary

In this study, an Exponentiated Generalized Family (EGF) of distribution was used to derive a four-parameter Exponentiated Generalized Exponentiated Exponential (EGEE) distribution. The proposed Kundu and Gupta (1998), Exponentiated Exponential (EE) distribution was used as the baseline distribution to obtain a new cumulative distribution function (CDF) and a related probability density function (PDF) was derived from the new CDF of the proposed distribution, which provides three shape parameters and one scale parameter. The PDF, CDF, and hazard function plots in figure 3.7, 3.8 and 3.9, confirms that the proposed EGEE distribution is unimodal and had an increasing hazard rate depending on the shape parameters. Secondly, the expressions for the general moment for EGEE distribution was derived and used to obtain the mean, variance, and standard deviation. Also, the cumulative distribution function of EGEE distribution was used in deriving its median and quantile function. The Galton and Moore expressions for skewness and kurtosis were used to obtain the skewness and kurtosis of the EGEE distribution since its quantile function exists in closed form. Other properties such as mean deviation, median deviation, and renyi entropy were also obtained in this study.

Some known continuous distributions such as the Exponential (E), Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions were found to be sub-distributions of the new proposed distribution when some parameters were relaxed such as for E distribution, for EE distribution and for EGE distribution.

Simulation studies were carried out by varying the scale parameter θ while the shapes parameters β, α and k were respectively held constant. The reported values indicates that the proposed distribution exhibits flexibility by returning positive and negative values for skewness and kurtosis as depicted in Tables 4.2 and 4.4, respectively. Data sets of sample size $n = 10, 50, 100, 150, 200$ and 350 were also simulated from an EE distribution with shape $k = 1.5$ and scale $\theta = 0.5$, E distribution with scale $\theta = 0.5$ and EGE distribution shapes $\beta = 6.5, \alpha = 1.5$ and scale $\theta = 0.5$ respectively. The data sets simulated were fitted with EGEE, EGE, E, and EE distributions.

The First sets of estimated values obtained from simulated Exponentiated Exponential distribution data sets in Tables 4.5 for sample size $n=10$ showed that the best distribution is the Exponential distribution with minimum values of AIC, BIC and maximum values of log-likelihood respectively as compared to other competing distributions (EGEE, EGE, and EE). The estimated values of sample sizes $n=50, 100$ and 200 generated from Exponentiated Exponential distribution and fitted with EGEE, EGE, EE, and E distributions presented in Tables 4.7, 4.9 and 4.11 showed that the best distribution is EE distribution with minimum AIC and BIC values respectively. The last sample size $n=350$ generated from EE distribution in table 4.13 showed an improvement in the proposed EGEE distribution outperforming the EGE, EE and E distributions with minimum values of AIC and BIC as compared to EGE, EE and E distributions respectively.

Tables 4.6, 4.8, 4.10 and 4.12 show the reported Likelihood Ratio Test (LRT) obtained from the simulated EE distribution with no significant differences between EGEE vs EGE and EGEE vs EE distributions respectively, while Tables 4.10 and 4.12 indicated that there are significant differences between EGEE and E distributions. The EGEE distribution tends to improve in performance compared to EGE, EE, and E distributions as the sample sizes of data sets simulated from EE distribution increases.

The Exponential distribution was also used in simulating data sets of sample sizes $n=10, 50, 100, 200$, and 350 . The estimated values obtained for sample size $n=10$ shown in Table 4.16 depicted that the best distribution is Exponential distribution with minimum values of AIC, BIC and maximum value log-likelihood respectively when compared to AIC, BIC and log-likelihood values of other competing distributions (EGEE, EGE and EE). The estimated values of sample size $n=50$ generated from Exponential distribution and fitted with EGEE, EGE, EE, and E distributions presented in Table 4.18 showed that the best distribution is EE distribution with minimum AIC value. However, the BIC value for E distribution reported the minimum.

The estimated values obtained in Tables 4.20, 4.22 and 4.24 for sample sizes $n=100, 200$, and 350 of simulated data sets from Exponential distribution showed that the best distribution is the Exponentiated Exponential distribution with minimum values of AIC, BIC and maximum Log-likelihood value when compared to the values of EGEE, EGE, and E distributions respectively.

Tables 4.17, 4.19, 4.21, 4.23, and 4.25 show the reported Likelihood Ratio Test (LRT) obtained from the simulated E distribution. From the test results, there were no significant differences between EGEE vs EGE and EGEE vs EE distributions in Tables 4.17, 4.19, 4.21

and 4.23 respectively, while the result in Table 4.25 indicated a significant difference between EGEE and E distributions.

The EGEE distribution tends to be more competing as the sample sizes of data sets generated from exponential distribution increases.

The second to last sets of simulated data were generated from Exponentiated Generalized Exponential (EGE) distribution for sample sizes $n=10, 50, 100, 200$ and 350 respectively. The simulated data sets were fitted with EGEE, EGE, EE and E distributions and the estimated values obtained in Tables 4.27, 4.29, 4.31 and 4.33 for sample sizes $n=10, 50, 100$ and 200 showed that the best distribution is the Exponentiated Exponential distribution with minimum values of AIC, BIC and maximum value of log-likelihood when compared to the values obtained by other competing distributions (EGEE, EGE and EE). The last sample size $n=350$ generated from EGE distribution in Table 4.35 showed an improvement in the proposed EGEE distribution with a reported minimum value of AIC and maximum log-likelihood value as compared to EGE, EE and E distributions respectively.

Tables 4.28, 4.30, 4.32, and 4.34 show the reported Likelihood Ratio Test (LRT) obtained from the simulated EGE distribution. The result showed that there is no significant difference between EGEE vs EGE and EGEE vs EE distributions respectively, while Tables 4.28, 4.30, 4.32 and 4.34 depicted that there is a significant difference between EGEE and E distributions. It was observed that the EGEE distribution tend to improve in performance compared EE and E distributions as the sample sizes of data sets simulated from EGE distribution increases.

The final simulated data set were generated from Beta distribution at different skweness levels and sample sizes $n=20, n=50$ and $n=100$ respectively. The simulated data were fitted to EGEE, EGE, EE and E distributions. the estimated log-likelihood values in tables 4.38, 4.39

and 4.40 shows that the proposed EGEE distribution fitted the data better than the existing distribution for both negatively and positively skewed data set generated for all sample sizes.

Finally, the Maximum Likelihood Estimator (MLE) was used to estimate the unknown parameters. Similarly, the information matrix was used in obtaining confidence intervals for the proposed distribution parameters. The proposed distribution was applied to fit four different real-life data sets which are known to be positively and negatively skewed, as shown in Figures 4.16, 4.18, 4.20, and 4.22 respectively. The estimates obtained from the first real-life data set presented in Table 4.43 were used to compare the proposed distribution with existing sub-distributions. The log-likelihood (-141.314), AIC (290.6361) and BIC (301.0557) values obtained from the proposed distribution when compared to that of EGE, EE and E distributions with log-likelihood (-146.182, -146.182 and -196.371), AIC (298.3646, 296.3646 and 394.7417) and BIC (306.1801, 301.575 and 397.3469) values respectively, showed overwhelming superiority of the proposed distribution over existing distributions. The Likelihood Ratio Test in Table 4.44 showed that there are significant differences in the proposed distribution compared with that of the existing sub-distributions with p-values of 0.0001, 0.00001 and 0.00001 for EGEE vs EGE, EGEE vs EE and EGEE vs E distribution respectively.

The second data set, which is left-skewed was fitted with the proposed and existing distributions with estimates presented in Table 4.47. The results reported a log-likelihood (21.0898), AIC (-34.1796) and BIC (-23.4883) values for the proposed distribution and that from the EGE, EE and E distributions with log-likelihoods (5.03874, 5.03875 and -25.9508), AIC (-4.07749, -6.30775 and 53.9015) and BIC (3.9409, -0.7318 and 56.5744) values respectively. The reported maximum value for log-likelihood and minimum values for AIC and BIC indicated that the proposed distribution best fitted the data than the other existing

distributions. The Three (3) Likelihood Ratio Tests obtained in Table 4.48 were significant in favor of the proposed distribution.

The log-likelihood (-16.912), AIC (41.824) and BIC (50.3971) estimates obtained by the proposed distribution as presented in Table 4.51 from the third data set are compared to the values obtained by the existing distribution which are; log-likelihood, AIC and BIC of EGE (-38.5517, 83.1035 and 89.5390), EE (-31.3834, 66.7669 and 71.0532) and E (-88.8303, 179.6606 and 181.8038) distributions respectively. The reported values showed that the best distribution is the proposed EGEE distribution.

The Likelihood Ratio Test shown in Table 4.52 depicts the significance of the proposed distribution to the existing sub distributions with p-values 0.0001, 0.00001, and 0.00001.

The fourth data set which is known to be right skewed with estimated values from all the distributions are presented in Table 4.53 and the results showed the relative closeness of EGEE, EGE and EE distributions with their log-likelihoods (-38.0932, -38.0943 and -38.0942) and p-values (0.5008 and 0.5008) indicating there are no significant differences in fitting the distribution to some right-skewed data sets. Hence they can serve as an alternative distribution in some areas of research.

5.3 Conclusion

The theory of Probability distribution plays a vital role in scientific research. Often times fitting distributions to data sets results into poor fit and to improve the flexibility of these existing distributions, several probability generators have been proposed in the literature to generalize them. Therefore, this study has been able to define and derive the four parameters EGEE distribution using EGF as the generator and EE as the baseline distribution. The proposed EGEE and various sub-distributions density plot in figure 3.7 shows that the

proposed EGEE can be left-skewed, right-skewed, monotonically decreasing (reversed J shape) and unimodal depending on the shape parameters α , β , and γ . In Figure 3.8, EGEE's cumulative distribution function (CDF) indicates a satisfactory level of CDF not exceeding 1 on the y-axis. The values of the shape parameters have a strong influence on the graphs. The graphical hazard function plot shown in Figure 3.9, shows that the proposed distribution can have an increasing and constant hazard rate depending on the parameters value. In Figure 3.10, the graph of the survival function shows initial constant values before decreasing at the end of the graph.

Since the quantile function of the proposed EGEE distribution occurs in closed form, Mathematica was used in plotting the 3D plots for Skewness and Kurtosis as shown in Figures 3.13 and 3.14. The plots confirm high level of skewness and kurtosis, which might take any shape (platykurtic or leptokurtic) depending on the parameters. The mean, standard deviation and median are increasing function of the scale parameters when the other parameters are held constant as observed in Tables 4.1 and 4.3.

Increasing the scale parameter increases the mean, standard deviation and median for fixed shape α . The mean, standard deviation and mean deviation are increasing functions of the shape parameters for fixed scale. Specifically, an increase in one of the shape parameters when others are held constant increases the values of the mean, standard deviation and median of EGEE distribution.

In Tables 4.2 and 4.4, skewness and kurtosis remain constant as the scale parameter increases when the shape parameters are held constant. An increase in the shape parameters reduces the values of skewness and kurtosis when the scale parameter is held constant.

The estimated fitted data simulated data sets from an Exponentiated Exponential distribution in Tables 4.5, 4.7, 4.9, 4.11, and 4.13 show the closeness of the proposed distribution with the

existing distributions. As the sample size n increases, the existing distribution tend to perform poorly as compared with the proposed EGEE distribution. Figures 4.1, 4.2, 4.3, 4.4, and 4.5 show consistency in the improvement of the peak and starting point of the proposed distribution.

The second data sets simulated from Exponential distribution reported in Tables 4.16, 4.18, 4.20, 4.22, and 4.24 show consistent improvement in the estimates of the proposed EGEE distribution as the sample sizes increases.

The last simulated data sets from EGE distribution reported the best distribution in the proposed EGEE distribution at $n = 350$, as shown in table 4.35.

The application of the EGEE and its sub-distributions on fibres, cow and glass data in Tables 4.41, 4.45 and 4.49 shows that the proposed distribution serves as a best distribution when compared with its sub-distributions using the log likelihood, information criterion (AIC and BIC), likelihood ratio test and the fit distribution curve, while the data on precipitation in Table 4.50 and its fitted plot on Figure 4.23, show that EGEE distribution can serve as an alternative and better distribution where the sub-distributions are applied.

5.4 Contribution to Knowledge

Distribution theory has been widely used for modelling observed situations whose various characteristics as reflected by the data differ from those that would be anticipated under the simple component distribution. The Exponential (E), Exponentiated Exponential (EE) and Exponentiated Generalized Exponential (EGE) distributions developed over the years performed pretty well when fitting a short tailed right skewed data sets and perform very poorly when fitting a long tailed right and left-skewed data sets.

Therefore, this study proposed a distribution (EGEE) which has the ability to accommodate asymmetric datasets for both left and right skewed data better than the existing EGE, EE and E distributions, by given better fitted values. This has been shown using statistical measures like Bayesian Information Criterion (BIC), Akaike Information Criterion (AIC), Log-likelihood and Likelihood Ratio Test (LRT) for large simulated and real-life data sets.

5. 5 Recommendations

The generator, exponential generalized family of distributions by Cordeiro (2013), serves as one of the best generator in formulating a unimodal and flexible distribution because of the additional two shape parameters that is guaranteed from the generator as shown in the thesis. The EGEE distribution so proposed can serve as an alternative distribution in fitting real data sets with asymmetry properties. It is therefore recommended that the EGEE distribution can be applied not only to data from production industries but can also be applied to data emanating from other fields.

5.6 Suggestions for Further Research or Study

There are several areas which are researchable within the context of this research work, and they are as follows;

- i) comparative study with other four parameters models can be developed for real-life application in the complex structure.
- ii) extension of the EGEE model can be done by using any of the established generating families of distribution in adding a shape or scale parameter as it may suit the data sets.
- iii) adoption of the distribution as a generator for other distributions is also researchable.

REFERENCES

- Adamidis, K., & Loukas, S. (1998). A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, 39(1), 35-42.
- Ahuja, J. C., & Nash, S. W. (1967). The generalized Gompertz-Verhulst family of distributions. *Sankhyā: The Indian Journal of Statistics, Series A*, 141-156.
- Akaike, H. (1973). Information theory and extension of the maximum likelihood principle, In: B.N. Petrov and F. Csaki (eds), *Second edition, international Symposium on Information Theory*: 267-87 Budapest Akademiai Kiado.
- Akinsete, A., Famoye, F. and Lee, C. (2008). The beta-Pareto distribution, *Statistics*, 42(6), 547 – 563.
- Aljarrah, M. A., Lee, C., & Famoye, F. (2014). On generating TX family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*, 1(1), 2.
- Al-Aqtash, R., Lee, C., & Famoye, F. (2014). Gumbel-weibull distribution: Properties and applications. *Journal of Modern applied statistical methods*, 13(2), 11.
- Alexander, C., Cordeiro, G.M., Ortega, E.M.M. and Sarabia, J.M. (2012). Generalized beta-generated distributions, *Computational Statistics and Data Analysis*, 56, 1880 - 1897.
- Alzaatreh, A., Famoye, F. and Lee, C. (2013b). Weibull-Pareto distribution and its applications, *Communications in Statistics-Theory & Methods*, 42, 1673 - 1691.
- Alzaatreh, A., Lee, C., and Famoye, F. (2013). A new method for generating families of continuous distributions, *Metron*, 71(1), 63-79.
- Arnold, B.C., Beaver, R.J., Groeneveld, R.A., and Meeker W.Q. (1993). The marginal of a truncated bivariate normal distribution. *Psychometrika*, 58, 471 – 488.
- Athayde, E., Azevedo, C., Leiva, V., & Sanhueza, A. (2012). About Birnbaum–Saunders distributions based on the Johnson system. *Communications in Statistics-Theory and Methods*, 41(11), 2061-2079.
- Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian journal of statistics*, 171-178.
- Barreto-Souza, W., Santos, A.H.S. and Cordeiro, G.M. (2010). The beta generalized exponential distribution, *Journal of Statistical Computation and Simulation*,

80(2), 159 - 172.

- Bourguignon, M., Silva, R. B., Zea, L. M., & Cordeiro, G. M. (2013). The kumaraswamy Pareto distribution. *Journal of Statistical Theory and Applications*, 12(2), 129-144.
- Bourguignon, M. Silva, R.B. and Cordeiro, G.M. (2014), The Weibull G family of probability distributions,” *Journal of Data Science*, 12, 53–68.
- Bourguignon, M., Ghosh, I., & Cordeiro, G. M. (2016). General results for the transmuted family of distributions and new models. *Journal of Probability and Statistics*, 2016.
- Burr, I. W. (1942). Cumulative frequency functions. *The Annals of mathematical statistics*, 13(2), 215-232.
- Chaudhry, M. A., & Ahmad, M. (1993). On a probability function useful in size modelling. *Canadian Journal of Forest Research*, 23(8), 1679-1683.
- Cooray, K. and Ananda, M. M. A. (2008). A generalization of half normal distribution with application to life time data. *Communication in statistics theory and method* 37; 1323-1337
- Cordeiro, G.M., Ortega, E.M.M., and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*, 347, 1399 - 1429.
- Cordeiro, G.M. and Brito, D. (2012). The Beta Power distribution: *Brazilian Journal of Probability and Statistics*, 26, 88-112.
- Cordeiro, G. M., & Lemonte, A. J. (2014). The exponentiated generalized Birnbaum–Saunders distribution. *Applied Mathematics and Computation*, 247, 762-779.
- Cordeiro, G. M., De Andrade, T. A., Bourguignon, M., & Gomes-Silva, F. (2017). The exponentiated generalized standardized half-logistic distribution. *International Journal of Statistics and Probability*, 6(3), 24.
- Cordeiro, G.M., Pescim, R.R., and Ortega, E.M.M. (2012). The Kumaraswamy generalized half-normal distribution for skewed positive data, *Journal of Data Science*, 10, 195 - 224.
- Cordeiro, G. M., Gomes, A. E., & da-Silva, C. Q. (2014). Another extended Burr III model: some properties and applications. *Journal of Statistical Computation and Simulation*, 84(12), 2524-2544..
- Cordeiro, G. M., Ortega, E. M., & da Cunha, D. C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, 11(1), 1-27.
- Cordeiro, G.M. and de Castro, M. (2011). A new family of generalized distributions, *Journal of Statistical Computation and Simulation*, 81(7), 883 - 898.

- da Silva, R. V., Gomes-Silva, F., Ramos, M. W. A., & Cordeiro, G. M. (2015). A new extended gamma generalized model. *International Journal of Pure and Applied Mathematics*, 100(2), 309-335.
- Das, K.K. (2015). On some generalized transmuted distributions. *International Journal Science Engineering Res* 6,1686 – 1691.
- De Andrade, T.A.N., Bourguignon, M. and Cordeiro, G.M. (2016). The exponentiated generalized extended exponential distribution. *Journal of Data Science*, 14, 393-414.
- De Andrade, T.A.N., Rodrigues, H. and Cordeiro, G.M. (2015). The exponentiated generalized Gumbel distribution. *Revista Colombiana de Estadística*, 38, 123-143.
- Dunning, K., and Hanson, J.N. (1977). Generalized Pearson system of distributions and linear programming. *Journal of Statistics and Computer Simulation*, 6, 115 - 128.
- Elbatal, I. (2013). Kumaraswamy linear exponential distribution. *Pioneer Journal Theoretical Applied Statistics*, 5, 59-73.
- Elbatal, I. and Muhammed, H.Z. (2014). Exponentiated generalized inverse Weibull distribution. *Applied Mathematical Sciences*, 8, 3997-4012.
- Eugene, N., Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications, *Communications in Statistics-Theory and Methods*, 31(4), 497 - 512.
- Famoye, F., Lee, C. and Eugene, N. (2004). Beta-normal distribution: Bimodality properties and applications, *Journal of Modern Applied Statistical Methods*, 3(1), 85 - 103.
- Famoye, F., Lee, C. and Olumolade, O. (2005). The beta-Weibull distribution, *Journal of Statistical Theory and Applications*, 4(2), 121 - 136.
- Ferreira, J. T. S., & Steel, M. F. J. (2006). A constructive representation of univariate skewed distributions. *Journal of the American Statistical Association*, 101(474), 823-829.
- Fernandez, C., and Steel, M.F.J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93, 357 – 371.
- Fischer, M. (2010). Generalized Tukey-type distributions with applications to finance and telegraphic data. *Statistical Papers*, 51, 41-56.
- Freimer, M., Kollia, G., Mudholkar G.S., and Lin, C.T. (1988). A study on the generalized Tukey lambda family. *Communication in Statistical Theory & Methods*, 17, 3547 – 3567.
- Fry, T.R.L. (1993). Univariate and Multivariate Burr distributions: A survey. *Pakistan Journal of Statistics Ser A*, 9, 1 – 25.
- Galton, F. (1883). *Enquiries into Human Faculty and its Development*, Macmillan and Company, London.

- Gomez, Y.M., Bolfarine, H. and Gomez, H.W. (2014). A new extension of the exponential distribution. *Revista Colombiana de Estadística*, 37(1), 25-34.
- Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. *Philosophical Transactions of the Royal Society London*, 115, 513-585.
- Gupta, R.C., Gupta, P.L. and Gupta, R.D. (1998). Modeling failure time data by Lehman alternatives. *Communications in Statistics, Theory and Methods*, 27, 887-904
- Gupta, R.D. and Kundu, D. (1999). Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, 41(2), 173-188.
- Hastings, J.C., Mosteller, F., Tukey, J.W., and Windsor, C. (1947). Low moments for small samples: a comparative study of order statistics. *Annal of Statistics*, 18, 413 – 426.
- Headrick, T.C., Kowalchuk, R.K., and Sheng Y. (2008). Parametric probability densities and distributions functions for Tukey g-and-h transformations and their use for fitting data. *Applied Mathematical Science*, 2, 449 – 462.
- Johnson, N.L., (1949). System of frequency curves generated by method of translation. *Biometrika* 36, 149-176
- Johnson, N.L., Kotz, S. and Balakrishnan, N (1994). Continuous Univariate Distributions, 1, 2nd Edn. Wiley, New York.
- Jones, M.C. (2009). Kumaraswamy's distribution: A beta-type distribution with tractability advantages, *Statistical Methodology*, 6, 70 - 81.
- Karian, Z.A., Dudewicz, E.J., and McDonald, P. (1996). The extended generalized lambda distribution systems for fitting distributions to data: History, completion of theory, tables application, the “final word” on moments fits. *Commun Stat Comput Simul*, 25, 611 - 642.
- Karian, Z.A., and Dudewicz, E.J. (2000). Fitting Statistical Distributions – The Generalized lambda Distribution and Generalized Bootstrap Methods. *Boca Raton, FL: Chapman & Hall/CRC Press*.
- Kapur, J.N. and Sexena, H.C. (1960). First Edition S. Chand and Company Ltd. Ram Nagar, New Delhi. *Mathematical Statistics*.
- Kong, L., Lee, C. and Sepanski, J.H. (2007). On the properties of beta-gamma distribution, *Journal of Modern Applied Statistical Methods*, 6(1), 187 - 211.
- Kumaraswamy, P. (1980). A generalized probability density functions for double-bounded random processes, *Journal of Hydrology*, 46, 79 - 88.

- Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27(1), 31-53.
- Kuş, C. (2007). A new lifetime distribution. *Computational Statistics & Data Analysis*, 51(9), 4497-4509.
- Lee, C., Famoye, F. and Olumolade, O. (2007). Beta-Weibull distribution: Some properties and applications to censored data. *Journal of Modern Applied Statistical Methods*, 6(1), 173 - 186.
- Lee, C., Famoye, F., and Alzaatreh, A. (2013). Methods for generating families of univariate continuous distributions in the recent decades. *WIREs Computational Statistics*, 5, 219 - 238.
- Lehmann, E.L. (1953). The power of rank tests. *Annals of Mathematical Statistics*, 24, 23-43.
- Lemonte, A.J., Barreto-Souza, W. and Cordeiro, G.M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of probability and Statistics*, 27(1), 31-53
- Lu, W., & Shi, D. (2012). A new compounding life distribution: the Weibull–Poisson distribution. *Journal of applied statistics*, 39(1), 21-38.
- Mahmoudi, E. (2011). The beta generalized Pareto distribution with application to lifetime data, *Mathematics and Computers in Simulation*, 81, 2414 - 2430.
- Marshall, A.W., and Olkin, I. (1997). A new method for adding parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, 84, 641 – 652.
- McDonald J.B. (1984). Some generalized functions for the size distribution of income, *Econometrica*, 52(3), 647 - 663.
- Merovci, F., Khaleel, M.A., Ibrahim, N.A. and Shitan M. (2016), “The beta type X distribution: properties with application,” *SpringerPlus*, 5, 697.
- Merovci, F. (2013). Transmuted exponentiated exponential distribution. *Mathematical Sciences and Applications E-Notes*, 1(2), 112-122.
- Mohan, D. P. and Todd, C. H (2014). A method for simulating Burr Type VII. *Journal of statistics and Econometric method* . 3(3) , 23-63
- Mood, A.M., Graybill, F.A. and Boes, D. (1974). *Introduction to theory of Statistics*, Third edition, McGraw-Hill, New York.
- Moor, J.J. (1988). A quantile alternative for Kurtosis. *The Statistician*, 37, 25-32.
- Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing

- bathtub failure rate data. *IEEE Transactions on Reliability*, 42(2), 299-302.
- Muino, J. M., Voit, E. O., & Sorribas, A. (2006). GS-distributions: A new family of distributions for continuous unimodal variables. *Computational statistics & data analysis*, 50(10), 2769-2798.
- Nadarajah, S., & Kotz, S. (2004). The beta Gumbel distribution. *Mathematical Problems in engineering*, 2004(4), 323-332.
- Nadarajah, S., and Gupta, A.K. (2004). The beta-Frechet distribution. *Far East Journal of theoretical Statistics*, 14, 15 – 24.
- Nadarajah, S. and Kotz, S. (2006). The beta exponential distribution, *Reliability Engineering and System Safety*, 91(6), 689-697.
- Nadarajah, S. and Zografos, K. (2003). Formulas for R nyi information and related measures for univariate distributions, *Information Sciences*, 155, 119 - 138.
- Nichols, M.D. and Padgett, W.J. (2006). A bootstrap control chart for Weibull percentiles. *Quality and Reliability Engineering International*, 22(2), 141-151.
- Nasiru, S., Mwita, P. N. and Ngesa, O. (2017). Exponentiated Generalized Exponential Dagun Distribution *Journal of King Saud University science* (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)
- Oguntunde, P.E., Adejumo, A.O. and Balogun, O.S. (2014). Statistical properties of the exponentiated generalized inverted exponential distribution. *Applied Mathematics*, 4, 47-55
- Oguntunde P. E., Adejumo A. O., and Kazeem A. A.(2016). Assessing the Flexibility of the Exponentiated Generalized Exponential Distribution. *The Pacific Journal of Science and Technology* . Number 1. May 2016 (Spring)
- Oguntunde, P. E., Khaleel, M. A., Ahmed, M. T., Adejumo, A. O., & Odetunmbi, O. A. (2017). A New Generalization of the Lomax Distribution with Increasing, Decreasing, and Constant Failure Rate. *Modelling and Simulation in Engineering*, 2017.
- Osatohanmwun, P., Oyegun, F. O., & Ogbonmwun, S. M. (2017). A New Member from the T– X Family of Distributions: the Gumbel-Burr XII Distribution and Its Properties. *Sankhya A*, 1-25.
- Pant, M. D., & Headrick, T. C. (2013). A Method for Simulating Burr Type III and Type XII Distributions through-Moments and-Correlations. *ISRN Applied Mathematics*, 2013.
- Pearson, K. (1895). X. Contributions to the mathematical theory of evolution.—II. Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society of London.(A.)*, (186), 343-414.

- Pescim, R. R., Cordeiro, G. M., Demetrio, C. G. B., Ortega, E. M. M. and Nadarajah, S. (2012). The new class of Kummer beta generalized distributions. *Statistics and Operations Research Transactions*, 36, 153-180.
- Progi F. I (2016) Exponential Generalized Beta Distribution *Journal of Geolocation, Geo-information, and Geo-intelligence*.
- Rasekhi, M., Alizadeh, M., Altun, E., Hamedani, G. G., Afify, A. Z., & Ahmad, M. (2017). The Modified Exponential Distribution With Applications. *Pakistan Journal of Statistics*, 33(5).
- Rényi, A. (1961). On measures of information and entropy: *In Proceedings of the 4th Berkely Symposium on Mathematical Statistics and Probability*. First Edition, University of California Press Barkley
- Ristic, M. M. and Balakrishnan, N. (2011). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, 82, 1191-1206.
- Rodrigues, C., Cordeiro, G.M., Demetrio, C.G.B., and Ortega, E.M.M. (2011). The Weibull negative binomial distribution. *Advances Applied Statistics*, 22, 25-55.
- R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistic*, 6, 461-464.
- Shams, T.M, (2013). The kumaraswamy-generalized Lomax distribution. *Middle East Journal of Scientific Research* 17(5),641-646
- Shakil, M., Kibria, B.M.G., and Singh, J.N. (2010). A new family of distributions based on generalized Pearson differential equation with some applications. *Australia Journal os Statistics*, 39, 259 - 278.
- Shaw, W. T., & Buckley, I. R. (2007). The alchemy of probability distributions: Beyond gram-charlier & cornish-fisher expansions, and skew-normal or kurtotic-normal distributions. *Submitted, Feb*, 7, 64.
- Smith R. L. and Naylor J. C. (1987), “A comparison of maximum likelihood and Bayesian Estimators for the three-parameter weibull distribution,” *Applied Statistics*, 36. 258–369,.
- Soliman, A.H., Elgarhy, M.A.E. and Shakil, M. (2017). Type II half logistics generated family of distribution with applications. *Pakistan Journal of Statistics and Operation Research*, 13, 245-264.
- 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.
- Tukey, J.W. (1960): The practical relationship between the common transformations of percentages of counts and amounts. *Technical Report 36, Statistical Techniques Research Group, Princeton University, Princeton, NJ*
- Verhulst, P.F. (1838). Notice sur la loi la population suit dans son accroissement. *Correspondence mathématique et physique, publiée L.A.J. Quetelet*, 10, 113-121.
- Verhulst, P. F. (1845). La loi d'accroissement de la population. *Nouv. Mem. Académic. Royal. Society. Belle-lettr. Bruxelles*, 18(1).
- Verhulst, P. (1847). Deuxième mémoire sur la loi d'accroissement de la population. *Nouv. Mémoires del' Académie Royale des Science. et Belles-Lettres de Bruxelles* 20, 1.
- Vicari, S. K. D., & Kotz, S. (2005). Survey of developments in the theory of continuous skewed distributions. *Metron*, 63(2), 225-261.
- Vort, E.O. (1992). The S-distribution: a tool for approximation and classification of univariate, unimodal probability distributions. *Biometrical Journal*, 7, 855 – 878.
- Yakubu, A. and Doguwa, S.I. (2017). On the Properties of the Weibull-Burr III Distribution and its Application to Uncensored and Censored Survival Data. *CBN Journal Applied Statistics*, 8, 91-116.
- Zografos, K. and Balakrishnan, N. (2009). On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, 6, 344- 362.