MODELING A THREE PARAMETER GUMBEL DISTRIBUTION USING MARSHALL-OLKINS TECHNIQUE

OTIENO OKUMU KEVIN

A RESEARCH THESIS SUBMITTED TO SCHOOL OF PURE AND APPLIED SCIENCES IN FULFILLMENT FOR THE AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN APPLIED STATISTICS OF MAASAI MARA UNIVERSITY

2024

Declaration

This research thesis is my original work and has not been presented for a degree in any other university. No part of this thesis may be reproduced without the author and /or Maasai Mara University's authorization.

OTIENO OKUMU KEVIN

SP05/SP/MN/16305/2021

Signature................................Date.................

This research thesis has been submitted for examination with our approval as the University Supervisors.

Dr. Omondi Ouno

Department of Mathematics and Physical Sciences

Maasai Mara University

P . O Box 861-20500, Narok-Kenya.

Signature................................Date....................

Dr. Anthony Karanjah

Department of Mathematics

Multimedia University of Kenya

P . O Box 15653 – 00503, Nairobi - Kenya.

Signature................................Date...................

Dr. Samwel Muthiga

Department of Mathematics and Physical Sciences

Maasai Mara University

P . O Box 861-20500, Narok-Kenya.

Signature................................Date...................

Acknowledgment

To God be the glory for making me reach this far to accomplish this work. To my supervisors, Dr. Joseph Ouno, Dr. Anthony Karanjah and Dr. Muthiga Samuel, I deeply appreciate for your guidance and support throughout the entire journey of this research. I am also thankful to members of mathematics and physical sciences and the entire school of science and applied sciences for their significant support.

Dedication

This work is dedicated to my family especially my sons, friends and all members in mathematics, physical and applied science department.

Contents

Declaration i

Acknowledgment ii

Dedication iii

List Of Figures vi

List Of Tables vii

Abstract viii

1 Introduction **1**

1.1 Background Information 1

1.2 Basic concepts 4

1.2.1 Univariate and Generalized distributions 4

1.2.2 Mixture of Gumbel and other distributions 7

1.3 Problem statement 9

1.4 Objectives 10

1.5 Significance of the study 10

2 Literature Review **12**

2.1 Introduction 12

2.2 Extreme value distributions 12

2.3 Research gap 19

3 Methodology **20**

3.1 Marshall Olkin method of adding a parameter 20

3.2 Maximum Likelihood Estimation method (MLE) 21

3.3 Properties of the Estimators 23

3.4 Test of goodness of fit 26

3.4.1 Akaike's Information Criterion (AIC) 27

3.5 Distributions for comparison with their parameter estimates 27

4 Results and Discussion **33**

4.1 Three-parameters Gumbel distribution formulation 33

4.1.1 Three parameters Gumbel distribution formulation 33

4.1.2 Expected value of a three parameters Gumbel distribution . 39

4.1.3 The variance of a three parameters Gumbel distribution . . 42

4.2 Parameter Estimation 45

4.2.1 Maximum Likelihood Estimation for the parameters 45

4.2.2 Parameters estimation for δ 46

4.2.3 Parameters estimation for ω 47

4.2.4 Parameters estimation for τ 49

4.3 Properties of the Estimators 53

4.3.1 Investigating the symmetrical behaviour of a three-parameters Gumbel distribution using simulation 53

4.3.2 Asymptotic bias 60

4.3.3 Mean Square Error 66

4.3.4 Consistency 67

4.3.5 Asymptotic relationship between three parameters Gumbel distribution and other distributions 72

4.4 Application and Efficiency 78

4.4.1 Efficiency using simulated data 79

4.4.2 Efficiency Analysis using real data 80

5 Summary, Conclusion and Recommendations **92**

5.1 Summary 92

5.2 Conclusion 92

5.3 Recommendation 93

List of Figures

4.1 Distributions when ω and n are constant 57

4.2 Distributions when τ and n are constant 59

4.3 Distributions when δ are n are constant 60

4.4 Distributions when ω, τ and δ are constant and n is varying 61

4.5 Asymptotic distributions for ω 62

4.6 Asymptotic distributions for δ 64

4.7 Asymptotic distributions for τ 65

4.8 Asymptotic distributions for three parameters Gumbel and normal distributions 74

4.9 Asymptotic distributions for three parameters Gumbel and chi-square distributions 76

4.10 Asymptotic distributions for three parameters Gumbel and Weibull distributions 78

4.11 Data 1:Earthquake magnitude 84

4.12 Data 2: Avocado prices in dollars 88

4.13 Data3: Income prediction data 91

List of Tables

4.1 Mean and MSE of the parameters 67

4.2 Estimates and standard errors for normal data 68

4.3 Correlation between the estimators 68

4.4 Estimates and standard errors for skewed data 69

4.5 Correlation between the estimators 70

4.6 Estimates and standard errors for extreme data 71

4.7 Correlation between the estimators 71

4.8 AIC values for simulated data 79

4.9 Summary statistics for Earthquake magnitude 80

4.10 AIC values for Earthquake magnitude 81

4.11 Estimates and MSE for earthquake magnitude 82

4.12 Summary staistics for Avocado prices (in dollars) 85

4.13 AIC values for Avocado prices 85

4.14 Estimates and MSE for Avocado prices 86

4.15 AIC values for Income prediction 89

4.16 Estimates and MSE for income prediction 90

Abstract

This research introduced a new three-parameter Gumbel distribution by adding a parameter to the traditional Gumbel distribution using the Marshall-Olkin method. We derived the probability density function, cumulative distribution function, and other statistical properties of the new distribution. The parameters of the distribution are estimated using the Maximum Likelihood Estimation (MLE) method. The new distribution improved flexibility and provided more efficient estimators for a broader range of data types, including normal, skewed, and extreme data. The properties of the estimators are thoroughly investigated, including their asymptotic bias, consistency, and mean square error (MSE). Through simulation studies and real data applications, the research demonstrates the superiority of the new distribution over existing models, evidenced by smaller Akaike Information Criterion (AIC) values and more efficient parameter estimates. The research recommends the new distribution for future analyses, particularly for large sample sizes, and suggests further research to refine the location parameter, study some characteristics like quartile deviation, order statistics, and characteristic function, and apply different parameter estimation methods to improve the efficiency of a three-parameter Gumbel distribution.

CHAPTER ONE

Introduction

1.1 Background Information

Extreme value analysis is a branch of statistics dealing with the extreme deviations from the centre of probability distribution and it focuses on limiting distributions which are distinct from normal distribution. Extreme value studies originated majorly from the experts in astronomy who focused on analyzing the data observed from astronomical objects like comets, planets, moons, stars etc. The early papers on the extreme value theories focused both on methods of statistical analysis and on the application of the formulated extreme value distributions (Afify et al., 2018).

Over past years, extreme value theory has indicated that the world is gaining a better understanding of the statistical modeling and analysis of the extreme value concepts. The understanding of the behaviour of extreme event cases is useful for understanding the whole behaviour of such cases both under the ordinary and extra-ordinary circumstances. Therefore,it is a mistake to separate the extreme events from the other events when it comes to modeling and analysis(Khalil & Rezk, 2019).

Probabilistic extreme value analysis has many applications involving natural phenomena such as wind characteristics, air pollution, flood, corrosion, rainfall etc, that satisfies advance mathematical models and results on point estimation applications and regularly varying functions. This area of research initially attracted the interests of probabilists, mathematicians, engineers and economists, and relative recently of the statisticians (Khalil & Rezk, 2019). Extreme value analysis primarily deals with the stochastic behaviour of the minimum and the maximum independent and identically distributed random variables, meaning each random variable has the same probability distribution as the others and all are mutually independent. The distributional properties of extremes (minima and maxima), thresholds and order statistics are determined by lower and upper tails of the applied distribution. The tails of the distribution parameters may be investigated by means of statistical procedures based on extreme and intermediate order statistics. A frequently occurring problem in statistics is the model identification in statistical data analysis. In standard applications like categorical data analysis and regression analysis, model selection becomes easy since it may be related to the number of independent variables to have in the selected model. In the application of statistical analysis of extreme value data, convergence to some standard extreme value distributions like Gumbel, Frechet and Weibull is important and a decision has to be made occasionally between special cases of distributions and the more general versions of the distributions (Persson & Rydén, 2010)

For a number of years, the extreme values problems and analysis were known to be related to the research studies of Gumbel (1958), whose life and studies were affected by pre-World war II disruptions. According to (Gumbel, 1958), the theoretical development of the 1940s were dealing with: the practical application of the extreme value statistics in distribution of engineering materials, human lifetimes, strength of materials, radioactive emissions, rainfall analysis, flood analysis, earthquake, market analysis, hydrology to mention a few examples.

Gumbel was the first researcher to alert the statisticians and the engineers to possible applications of the formal extreme value study distributions which had in decades been treated empirically. The potential applicability of the Gumbel distribution to represent the cases of minima and maxima relates to the extreme value theory which implies that it is most likely to be applicable if the distribution of the underlying sample data is of the normal or exponential type (Persson

& Rydén, 2010).

Today, extreme point distributions have developed as one of the key statistical area for applied sciences. Analyzing extreme values therefore, requires parameter estimation and application of the probability of events that are more extreme than the previously experienced cases with the main goal of estimating the future expectations (Teamah et al., 2020). Extreme value analysis provide a framework that assists for this type of research work that deals with extreme data sets. Gumbel distribution is not only widely used in various application in extreme value studies but also referred as the mother to the extreme value distributions (that is, Frechet and Weibull distribution types)(Persson & Rydén, 2010; Teamah et al., 2020). Not many research have been published on the extensive study of the Gumbel distribution even with its ability to fit data from many different areas of the extreme value observation like engineering, physics, climate among others.

This study therefore developed a new distribution called three parameters Gumbel distribution to improve the flexibility of the already existing two parameter distribution. The new distribution was developed by applying the Marshall-Olkin method for adding a new parameter to an existing distribution. To estimate the parameters of the three parameters Gumbel distribution which we discuss in chapter three, this study intends to apply Maximum Likelihood Estimation method. This method of estimation was preferred over the other methods like Method of moments, Ordinary Least square, percentiles, Cramer-Von Mises etc because (Afify et al., 2018; Al-Subha & Alodatb, 2017; Hussein et al., 2021; Zhou et al., 2018) provide enough evidence supporting Maximum Likelihood Estimation as the best parameter estimation method since it provides better estimates for both small and large samples of data.

This research concentrates on three parameters namely; the location parameter, dispersion parameter and the shape parameter. The location parameter help in determining the shift of the distribution under study and as well tells us where the distribution is located/centered, the scale/dispersion parameter helps in describing how the distribution is scattered around the center or simply how the distribution is spread and the shape parameter guide us on the shape of the distribution depending on the value of the shape parameter, that is it changes the slope of the distribution.

1.2 Basic concepts

This section discuss the univariate and generalized distributions, mixture of Gumbel with other distributions(that is, gamma, beta, geometric, lomax, exponential etc), all of which are Gumbel family and have been used to model and analyze extreme values. The three primary parameters in discussion is, the location parameter which shifts the distribution along the horizontal axis. The scale parameter that stretches or compresses the distribution, and the shape parameter for determining the tail behavior of the distribution.

1.2.1 Univariate and Generalized distributions

Extreme value distributions are always considered to fall under three families namely; type one, two and three, whereby type one was introduced by Gumbel hence giving us the Gumbel distribution modeled as follows.

1. Type 1, (Gumbel distribution)

$$Pr[V \leq v] = exp\left[- exp^{\left(-\frac{v-\gamma}{\vartheta}\right)}\right], \tag{1.1}$$

   where V is the random variable with $i = 1, 2, ..., k$ observations, $\gamma > 0$ is the location parameter and $\vartheta > 0$ is the dispersion or scale parameter.

2. Type 2, (Frechet distribution)

   Type 2 distribution is called the Frechet distribution and its probability density function is given as follows.

$$Pr[V \leq v] = \begin{cases} 0, & v < \gamma \\ exp\left\{ -\left(\frac{v-\gamma}{\phi}\right)^{-\rho} \right\} & v \geq \gamma, \end{cases} \tag{1.2}$$

where $(\rho, \gamma, \phi) > 0$ are the three parameters namely shape, location and scale respectively.

3. Type 3, (Weibull distribution)

   Type 3 distribution is called the Weibull distribution. It is a distribution with three parameters and it is described as follows.

$$Pr[V \leq v] = \begin{cases} exp\left\{ -\left(\frac{v-\gamma}{\phi}\right)^{-\rho} \right\} & v \leq \gamma \\ 0, & v > \gamma \end{cases} \tag{1.3}$$

   where $\gamma > 0$ is the location parameter, $\rho > 0$ is the shape parameter and $\phi > 0$ is the scale/dispersion parameters and $-\infty < v < \infty$ is the random variable.

Of these three families of extreme value distribution, Gumbel type (type 1) is the most commonly used in the extreme value theory analysis. Indeed, majority of the scholars call Gumbel distribution the extreme value distribution. In view of this and again the fact that distribution of type 2 (Frechet distribution) and type 3 (Weibull distribution) can be transformed to Gumbel distribution by applying a simple transformation, that is $M = log(V - \gamma), M = -log(\gamma - V)$, respectively (Kotz & Nadarajah, 2000).

Univariate distributions include Gumbel, Frechet, and Weibull distributions, each characterized by specific parameters. Generalized distributions, such as the Generalized Extreme Value (GEV) distribution, combine these families and provide a more flexible framework for modeling extreme values.The foregoing means that generalized distributions have been modeled out of the three distributions of type 1, 2 and 3 as follows:

4. Generalized Extreme Value distribution (GEV) also known as Von Mises type extreme value distribution or Von Mises-Jenkinson type distribution.

   The generalized extreme value distributions was first developed by Jonsson(2014). This distribution was to be applied on various studies like sea levels, rainfall, air pollutants, river lengths and annual maxima among others. The cumulative distribution function of the GEV distribution is obtained as;

$$F(v) = \begin{cases} exp^{-(1+\beta(\frac{v-\lambda}{\theta}))^{\frac{-1}{\beta}}} & -\infty < v \leq \lambda - \frac{\theta}{\beta} \qquad for(\beta < 0) \\ & \lambda - \frac{\theta}{\beta} \leq v < \infty \qquad for(\beta > 0) \\ exp^{-exp^{-\frac{v-\lambda}{\theta}}} & -\infty < v < \infty \qquad for(\beta = 0) \end{cases} \tag{1.4}$$

   where the Generalized Extreme Value distribution include the Gumbel distribution when $\beta = 0$, Frechet distribution when $\beta > 0$ and Weibull distribution when $\beta < 0$ (Kotz & Nadarajah, 2000). And λ, β and θ are the location, shape and scale parameters respectively.

5. The Generalized Extreme Value (GEV) distribution

   This is also another combination of the three families of the extreme value distributions (Gumbel, Frechet and Weibull) distribution which was developed by Rychlik and Rydén(2006) and modeled as follows

$$F(v) = \begin{cases} exp\big(-\big(1 - \frac{\vartheta(v-\gamma)}{\phi}\big)^{\frac{1}{\vartheta}}\big), & \vartheta \neq 0 \\ exp\big(-exp\big(-\frac{(v-\gamma)}{\phi}\big)\big), & \vartheta = 0 \end{cases}, \tag{1.5}$$

   where γ, ϕ, ϑ are the location, dispersion and shape parameters respectively.

1.2.2 Mixture of Gumbel and other distributions

In this subsection, we present a mixture of Gumbel distribution with exponential, gamma, geometric, beta, Lomax and left truncated distributions known respectively as Exponentiated Gumbel, Gamma Gumbel, Gumbel geometric, beta Gumbel, Lomax Gumbel, Left truncated Gumbel and Odd exponentiated half logistic Gumbel represented as follows.

1. A generalization of the Gumbel distribution, referred to the exponential Gumbel distribution (EG) was introduced by (Nadarajah, 2006)

$$f_{EG}(v) = \frac{\phi}{\psi} exp\Big(-\frac{v-\rho}{\psi}\Big) exp\Big[-exp\Big(-\frac{v-\rho}{\psi}\Big)\Big]\Bigg[1-exp\Bigg\{-exp\Big(-\frac{v-\rho}{\psi}\Big)\Bigg\}\Bigg]^{\phi-1}, \tag{1.6}$$

   where, ρ is the location parameter, $\phi > 0$ is the shape parameter and $\psi > 0$ is the dispersion parameter.

2. The Gamma Gumbel distribution

   This type of distribution was introduced by Gholami et al. (2020). The probability density function of the Gamma Gumbel (GG) distribution is given by;

$$f(v) = \frac{1}{\theta\Gamma(\psi)} exp\Bigg\{-\frac{\psi(v-\rho)}{\theta} - exp\Big(-\frac{v-\rho}{\theta}\Big)\Bigg\}, v\in\Re, (\theta, \psi) > 0, \rho\in\Re \tag{1.7}$$

   where ρ is the location parameter, θ is the scale parameter and ψ gives the shape of the distribution. The Gumbel distribution is obtained from the Gamma Gumbel distribution as a submodel by setting $\psi = 1$.

3. The Gumbel Geometric distribution

   This distribution was introduced by Oseni and Okasha (2020). They stated that a random variable V, with the interval $(-\infty, \infty)$ is referred to a Gumbel geometric with parameters (ϕ, β, γ) with the probability density function given by;

$$\begin{aligned} f(v) &= \frac{(1-\gamma)}{\beta} exp\Bigg(-\frac{v-\phi}{\beta} - exp\Bigg(-\frac{v-\phi}{\beta}\Bigg)\Bigg) \times \\ & \Bigg(1-\gamma+\gamma exp\Bigg(-exp\Bigg(-\frac{v-\phi}{\beta}\Bigg)\Bigg)\Bigg)^{-2}, \end{aligned} \tag{1.8}$$

where $\phi\in(-\infty,\infty), \beta\in(0,\infty)$ and $\gamma\in[0,1)$ are the three parameters of the distribution representing the location parameter, scale parameter and the shape parameter respectively.

4. The Beta Gumbel distribution

According to Jonsson (2014), the probability density function of the Beta Gumbel distribution with four parameters, that is scale and location parameters, (z, w, ϕ, θ), is given by:

$$f(v) = \frac{1}{\phi B(z,w)} \rho exp^{z\rho}[1-exp^{-\rho}]^{w-1}, -\infty < v < \infty, \tag{1.9}$$

where $\rho = exp[-\frac{v-\gamma}{\phi}]$, $B(z, w)$ is the beta function and

$$-\infty < \gamma < \infty, (\phi, z, w) > 0$$

5. The Lomax Gumbel distribution

The Lomax Gumbel distribution was developed by Jaya et al. (2016). The study defined the cumulative density function of the Lomax-Gumbel distribution with four parameters as

$$F_{LG}(v; \vartheta, \gamma, \psi, \delta) = P\Bigg[1 - \big\{1 + \frac{1}{\gamma} exp\big\{ -exp\big(-\frac{v-\psi}{\delta}\big)\big\}\big\}^{-\vartheta}\Bigg] \tag{1.10}$$

where $P^{-1} = 1-(1+\gamma^{-1})^{-\vartheta}, (\vartheta, \gamma, \delta) > 0, -\infty < \psi < \infty, -\infty < v < \infty$. And ψ is the location parameter, (γ, δ) are scale parameters and ϑ is the shape parameter

6. Left truncated Gumbel distribution

   Suppose V is a random variable defined within the range $(-\infty, \infty)$, in order to have the probability density function of the random variable V defined on the interval $[Z, \infty]$, where z is the lower bound, then the probability density function of the random variable V is truncated from the left. Therefore, the left truncated Gumbel distribution for interval $(0, \infty)$ is given as (Neamah & Qasim, 2021)

$$f(v) = [1 - exp^{-exp\frac{p}{w}}]^{-1}\frac{1}{p}exp^{-exp^{\frac{-(v-p)}{w}}-\frac{(v-p)}{w}}; v \geq 0, \tag{1.11}$$

   where p is the location parameter and w is the scale parameter.

7. The odd exponentiated half-logistic.

   This is a distribution of Gumbel family with three parameters namely; γ is the shape parameter, δ is the scale parameter and ϕ is the location parameter (Afify et al., 2018).

$$F(v;\gamma,\delta,\phi) = \left\{\frac{1 - exp\left[\frac{-\delta G(v;\phi)}{G(v;\phi)}\right]}{1 + exp\left[\frac{-\delta G(v;\phi)}{G(v;\phi)}\right]}\right\}^{\gamma}, (\gamma, \delta, \phi > 0), \tag{1.12}$$

1.3 Problem statement

Extreme value analysis is developing more interest in applied sciences and therefore, there is need to develop more flexible distribution for analyzing/modeling the extreme data. Extreme value distributions are always viewed to include families of Gumbel, Frechet and Weibull distributions. Of the three distributions, Gumbel distribution is frequently used in the extreme value theory analysis because majority of the authors refer to Gumbel distribution as the mother to the extreme value distributions from the fact that the Frechet and Weibull distributions can be transformed to Gumbel distribution by applying a simple transformation.

Gumbel distribution with two parameters namely location and dispersion can be made more flexible to improve its reliability for modeling extreme data. Existing literature has shown that the addition of parameter to a distribution makes it robust and/or more flexible hence the study intends to improve the existing Gumbel distribution by making it more flexible through addition of shape parameter using Marshall and Olkin technique.

1.4 Objectives

The main objective of the study is to develop a three parameters Gumbel distribution using Marshall Olkins technique.
The specific objectives are as follows:

1. To derive a three parameter Gumbel distribution using Marshal-Olkins method
2. To estimate the parameter of a three parameters Gumbel distribution using Maximum Likelihood Estimation method
3. To evaluate the properties of the estimates of the three parameters Gumbel distribution using simulated data.
4. To apply a three parameter Gumbel distribution to simulated and real data sets.

1.5 Significance of the study

The development of a three parameters Gumbel distribution will significantly enhance the ability to model and predict extreme data, skewed data and normal data. This has profound implications to some lifetime studies like high temperature, wind characteristics, earthquakes, network designs, horse racing, queues in supermarket, insurance premium distributions, risk management, ozone concentration, flood, engineering and more recently financial concepts by improving the accuracy and reliability of the results through the provision of unbiased, sufficient, efficient and consistent estimators.

CHAPTER TWO

Literature Review

2.1 Introduction

This chapter gives the review of extreme value distributions and their extensions. It also explains the methods used to approximate the distribution's parameters, the findings and critics. The review of the literature is essential in that it helps in having a clear picture of the research gap.

2.2 Extreme value distributions

In the extreme value analysis, it was found that the maximum of many common distributions like lognormal and normal distributions converge to Gumbel distribution (Coles & Pericchi, 2003). In the literature of probability risk analysis, there has been debate over the decision to use Gumbel distribution contrast to Generalized Extreme Value distribution. Also, in the analysis of rainfall measurements by *optic*, the same discussions were made. Similar conclusions are given by the authors that Gumbel distribution provides narrower confidence interval than the Generalized Extreme Value distribution but has also the risk of under-estimating the parameters. Hence, the selection of the distribution is not significant (Coles et al., 2001). This shown that there is still a gap in the distribution modeling to come up with the more robust distribution which can be fulfilled by re-modeling the Gumbel distribution.

According to Neamah and Qasim (2021), left truncated Gumbel distribution is a suitable distribution for random variables within the interval $(0, \infty)$. In terms of the parameter estimation, the study found out that the best method to approximate the parameters of left truncated Gumbel distribution is the Maximum Likelihood Estimation method because this method provided the lowest standard mean squared error at all levels of the sample and for all cases of the default values for the parameters. This study support MLE as the greatest approximation method for approximating the parameters. This study did not look into modeling the original Gumbel distribution, it only truncated the observations of the random.

The study by Oseni and Okasha (2020), developed Gumbel geometric distribution and investigated its characteristics such as quantile function, moments and characteristics functions. The distribution was found to be quite flexible and the asymptotic properties of the distribution indicated that Mean Square Error decreases to Zero as $K \rightarrow \infty$, where K is the number of observations, and the bias either decreases or increases depending on the sign for each of the parameter. This study created a three parameters distribution from combination of two statistical distributions but not by adding a parameter to the type 1 (that is, Gumbel distribution).

The study by Jaya et al. (2016), introduced a new statistical distribution called Lomax Gumbel distribution by combining Gumbel and Lomax distributions. This was achieved by developing a new technique for constructing a family of statistical distributions by combining the cumulative density functions (cdf) of two known distributions. The study approximated the parameters by the method of Maximum Likelihood Estimation. This study further confirmed that the combined distribution is more robust than either the Loman distribution or the Gumbel distribution when applied individually. This study introduced the Lomax Gumbel distribution by combining the two distributions but not by adding a parameter to the univariate Gumbel distribution.

In a statistical analysis Gumbel distribution, Exponential Gumbel distribution and Generalized Extreme Value distribution (three parameter Generalized Extreme Value distribution) were applied on two types of data sets. The results showed that Exponential Gumbel distribution could serve as an alternative to Generalized Extreme Value distribution because it gives narrower confidence intervals than the Generalized Extreme Value distribution. It was also noted that the Exponential Gumbel distribution produced the smallest Anderson-Darling statistical value compared to the Gumbel and Generalized Extreme Value distribution. This research recommended further studies on estimation methodology and analysis using Exponential Gumbel distribution (Persson & Rydén, 2010). This study also combined exponential and Gumbel distribution, it did not add a parameter to the univariate Gumbel distribution using Marshall Olkin proposed method.

According to Jonsson (2014), who introduced Beta Gumbel distribution which had two additional parameters to permit for the study of variability of the tail weights and skewness and from which the simulated data provided the evidence that the Beta Gumbel distribution was more flexible than the Gumbel distribution. Providing the Maximum Likelihood Estimation for the introduced distribution was unsuccessful due to computational problems. This research further pointed out that Beta Gumbel distribution cannot be concluded as a better distribution to use and therefore, the study recommended for further research to be done to study the distribution and its applicability to numerous areas. From this study, there is evidence to prove that the work on the flexibility of the extreme distributions is still not complete since studies are concentrating on combination of distributions and not on the application of Marshal Olkin method of introducing a new parameter to a distribution.

The newly introduced odd exponential half-logistic Gumbel family distribution with two extra parameters had its model parameters approximated by the method of Maximum Likelihood Estimation making use of simulated data. The application of this model proved that it provided consistently better fits than other comparative models which were introduced using the Gumbel family (Afify et al., 2018). The introduced model was a result of combination of model and not additional of parameter to the Gumbel distribution

According to Al-Subha and Alodatb (2017), deciding a right distribution for analyzing a given data, is a very primary subject and when setting out a data set, it becomes a problem to select a correct distribution between Lognormal and Gumbel distributions with two parameters. The study used Monte Carlo simulations to simulate data. The finding was that on the basis of the Monte Carlo simulations the Probability of Correct Selection (PCS) for MLE estimation technique for the Gumbel distribution slightly outperforms the MLE and MOME for the logistic regression. This research supports Maximum Likelihood as the best method for parameter estimation both for lognormal and Gumbel distributions. This study did not model the Gumbel distribution, it compared the Gumbel and Lognormal distribution.

A study by Qaffou and Zoglat (2017), compared normal and Gumbel distribution after arguing that they are much alike in practical application in flood and engineering analysis. The researchers used the ratio of the Restricted Maximized Likelihood (RML) estimation as the test statistics under both normal and Gumbel (two parameter)distributions. Using Monte Carlo simulation, the probability of correct selection was compared with the asymptotic distribution results of the test statistics under the null hypothesis and it was found that ML can be used for differentiating between any two distributions of the scale and location parameter. This study showed Maximum Likelihood is the best estimation method for discriminating two or more distributions of the same family. This study compared the normal and Gumbel distribution, it did not add a parameter to the Gumbel distribution.

In the study of a new three parameter distribution called Burr X Frechet distribution which has the advantage of modeling aging and failure criteria. The Maximum Likelihood Estimation method was used to approximate the model parameters which were applied to assess the simulated data. The study proved the flexibility and the importance of the new Burr X Frechet distribution by compar-

ing it with other extensions of Frechet distribution. This study showed that it is indeed significant to modify a distribution to make it flexible and suitable for modeling real life events (Abouelmagd et al., 2018).
According to Khalil and Rezk (2019), a new extension of Frechet distribituion called Poison Burr X Frechet was introduced based on Zero Touch Provisioning (ZTP) model. The PDF of the poison Frechet model showed that it can be unimodal, left skewed or right skewed and therefore, more flexible than Fretch distribution. The method of Maximum Likelihood Estimation was used to approximate the parameters. This study combined distributions instead of adding a parameter to the univariate frechet distribution.
According to Hussein et al. (2021), a new flexible extension of the Frechet distributions referred to extended Weibull-Frechet distribution in which they believed provided more flexibility to model and analyze real life data than the other set of extreme value distributions like univarite Frechet and Weibull distributions. The researchers estimated the extended distribution parameters using Maximum Likelihood Estimators (MLEs), Maximum product Spacing Estimators (MPSEs), Anderson-Darling Estimators (ADEs), Least Square Estimators (LSEs), the Percentiles Estimators (PCEs), the Cramr-von Mises Estimators (CRVMESs). The estimation results showed that MLE and MPSE outperforms the other estimation methods. Also, the flexibility and importance of the extended four parameter Weibull-Frechet distribution was examined using real data sets from engineering and medicine field and the results showed that the Extended Weibull-Frechet distribution indicated an adequate fit compared to Frechet distribution extensions like beta-Frechet distribution, Exponential Frechet distribution, gamma extended Frechet distribution, transmuted exponential Frechet distribution and Frechet distribution. The study used goodness of fit measures like AIC (Akaike Information Criterion), CAIC (Consistent Akaike Information Criterion), BIC (Bayesian Information Criterion), Cramer-Von Mises, Anderson-Darling and KS (Kolmogorov-Smirnov). From this study, there is evidence that introducing a parameter to an

surviving distribution can make the extension more flexible and important for application. This study combined Fretch and Weibull distribution but it did apply the Marshall Olkin method of putting on a parameter to an existing distribution. According to Ramos et al. (2017), the proposed long term Frechet distribution gave out results towards the maximum likelihood parameter estimators and their asymptotic characteristics. The efficiency of the estimators was evidenced from the simulated data which generated acceptable outcomes even for small sample sizes. The study proposed application of other distribution extensions as well as other estimation methods like Bayesian estimation approach to improve the robustness of the distributions

A study by Teamah et al. (2020), established a new distribution called Frechet-Weibull distribution (FWD) which extended to Weibull distribution. Properties of the distribution was examined using moments, moments generating function, quantile function, mode and mean residual life function. Akaike Information Criterion and Bayesian Information Criterion were also applied to study the goodness of fit. The simulated data indicated that all the applied estimators perform well in terms of their bias. Real data application was used to investigate the potentiality and flexibility of the new distribution and the results proved that the new distribution (FWD) gave greater fits than other adjuncts of Frechet and Weibull distributions. This study implies that adding more parameters to a distribution makes it more flexible and important for application though it did not either add a parameter directly to a distribution or applied the Marshal Olkin method.

According to Mansour et al. (2018), who developed a three parameter Frechet distribution called Odd Lindney Frechet distribution (OLiFr) which extends the Frechet distribution, model estimates were obtained by Maximum Likelihood estimation methods. When they applied a two real data sets, they proved that three parameters Frechet distribution provide better fits than the Frechet distribution and other well known extensions of Frechet distribution (Kumaraswany Frechet, Exponentiated Frechet, gamma extended Frechet, beta Frechet, transmuted Frechet and transmuted Marshall-Olkin Frechet). This in general implies that a three parameter distribution more flexible to be used in any data. This study did not apply the Marshal Olkin proposed method for introducing a parameter to a distribution

A study by Willayat et al. (2022) used the marshall olkins method to add a parameter to the generalized form of the standard Gumbel type II distribution (also called Exponentiated Gumbel type II) and developed a distribution called Marshal-Olkins Extended Gumbel type II distribution with one scale parameter and three shape parameters. This study used maximum likelihood estimation method and concluded that it gave better estimates with increase in sample size. The study further studied properties like characteristic function, order statistics, Quartile function and Moment generating function among others. This research did not concentrate at introducing a new parameter to the original Gumbel distribution with two parameters namely scale and location.

Three methods of parameters estimation namely maximum likelihood estimation, method of moments and least square estimation and concluded that maximum likelihood estimation method provides greater estimates compared to method of moments and least square estimation (Salma & Abdelali, 2018).

A study by Johnson et al. (2011) compared methods of moments and maximum likelihood estimation techniques on gamma distribution and reached a conclusion that maximum likelihood estimation is better for parameter estimation since it gives superior estimates for continuous distributions.

According to Saleh et al. (2012) research on investigating different techniques for parameter estimation like power density method, maximum likelihood method, graphical estimation method and modified maximum likelihood method, the study concluded that maximum likelihood estimation method provides good estimators.

2.3 Research gap

From section 2,2, we have found that adding a parameter to some existing distribution makes them more flexible and important for modeling and analyzing both simulated and real life data sets. This is because the newly introduced parameters in a distribution provides better estimates and makes it more robust and/or efficient than the baseline distributions. However, from the reviewed literature, we realized that apart from combining Gumbel distribution with other distributions like exponential, gamma, geometric among others, no scholar have modeled a three parameters Gumbel distribution. For this study we wish to model a three parameters Gumbel distribution (3-PG) about which we consider three parameters (that is, shape, location and dispersion) using Marshall-Olkin (19997) proposed method.

CHAPTER THREE

Methodology

This chapter discusses how the three parameters Gumbel distribution was developed from the two parameters distribution, further it demonstrates how the parameters of the distribution are estimated. The chapter shows particular properties of the distribution to be investigated using simulated data. It also discuses how its efficiency compares with Gumbel distribution family. In developing the three parameter Gumbel distribution we rely on Marshall-Olkin technique of introducing a parameter to an existing distribution. Marshall-Olkin's technique is discussed in section 3.1.

3.1 Marshall Olkin method of adding a parameter

Marshall and Olkin (1997) proposed a procedure of introducing a new parameter to an existing distribution. This is because introducing a parameter to a well defined distribution is an honored procedure in that it helps in introducing a more flexible new family of the distribution for modeling various data types (Hassan et al., 2022). A new parameter for the main objective of this study is to be added to the following baseline probability distribution which is a two parameter distribution with ω representing location of the variables and τ representing scale or dispersion parameter.

$$f(v) = \frac{1}{\tau} exp\big(- \frac{v-\omega}{\tau} - e^{-\frac{v-\omega}{\tau}} \big) \tag{3.1}$$

where $v \in \Re$ and $\omega > 0, \tau > 0$ are location and dispersion/scale parameters respectively.

Marshall Olkin (1997) began with a parent survival function $\bar{F}(v)$ and considered a family of survival function given by:

$$\bar{S}(v) = \frac{\delta \bar{F}(v)}{1 - \bar{\delta}\bar{F}(v)} = \frac{\delta \bar{F}(v)}{F(v) + \delta \bar{F}(v)}, -\infty < v < \infty \qquad (3.2)$$

where $\delta > 0$, $\bar{\delta} = 1 - \delta$ and $\bar{F}(v) = 1 - F(v)$. For $\delta = 1, \bar{S}(v) = \bar{F}(v)$.

The two primary properties of the Marshall-Olkin family of distributions is that it has a stability property, that is if the method is put in application twice it returns back to the original distribution and the introduced distribution satisfies the property of the geometric extreme stability (Hassan et al., 2022).

After introducing the new distribution, the corresponding cumulative distribution function (cdf) and probability density function (pdf) are respectively obtained as given in the following equations:

$$T(v) = \frac{F(v)}{1 - \bar{\delta}\bar{F}(v)} \qquad (3.3)$$

and

$$t(v) = \frac{\delta f(v)}{\left(1 - \bar{\delta}\bar{F}(v)\right)^2} \qquad (3.4)$$

3.2 Maximum Likelihood Estimation method (MLE)

In this section discuss the method of Maximum Likelihood Estimation which we use in this study for the purpose of estimating parameters of the new three parameters Gumbel distribution to be modeled.

According to (Zhou et al., 2018), maximum likelihood method can be used in many problems since it has a strong instinctual appeal and it yield a better estimator(s). This method of maximum likelihood is widely put in application because it is more precise especially when dealing with large samples since it yields a more efficient

estimator(s) when the sample is large. This is an evidence from the literature review where we realized that Maximum Likelihood Estimation method gives the best parameter estimates compared to the other estimation methods like Minimum Distance Estimation, Method of Moments etc.
According to (Hurlin, 2013), if supposing we have say $\hat{Z}$ of P is a solution to the maximization problem given as

$$\hat{Z} = argMaxln(Z : v_1, v_2, ..., v_k), \tag{3.5}$$

where, $v_1, ..., v_k$ represents the data observations, then under suitable regularity conditions, where the first order condition is given as

$$\frac{\partial ln}{\partial Z}(Z : v_1, v_2, ..., v_k) = -k + \frac{1}{\hat{Z}}(\sum_{i=1}^{k} V_i) \tag{3.6}$$

These conditions are generally called the likelihood or log-likelihood equations. The first derivative or gradient of a condition (log-likelihood) which is solved at point $\hat{Z}$ need to satisfies the following equation

$$\frac{\partial ln(Z : v_1, v_2, ..., v_k)}{\partial Z} = \frac{\partial ln(\hat{Z} : v_1, v_2, ..., v_k)}{\partial Z} = 0 \tag{3.7}$$

The log-likelihood equation that coincide to linear or non-linear system of P equations with P unknown parameters $Z_1, Z_2, ..., Z_P$ with K observations is given by

$$\frac{\partial ln(Z : v_1, ..., v_k)}{\partial Z} = \left(\frac{\partial ln(Z : v_1, ..., v_k)}{\partial Z_1}\right) = ... = \left(\frac{\partial ln(Z : v_1, ..., v_k)}{\partial Z_P}\right) = 0 \tag{3.8}$$

Maximum Likelihood Estimation is a recommended technique for many distributions because it uses the values of the distribution parameters that makes the data more likely than any other parameters. This is achieved by maximizing the likelihood function of the parameters given the data. Some good features of maximum likelihood estimators is that they are asymptotically unbiased since the bias tends to zero as the sample size increases and also they are asymptotically efficient since they achieve the Cramer-Rao lower bound (CRLB) which states that for any unbiased estimator of population parameter P it gives a lower estimate for the variance of an unbiased estimator, as sample size approaches ∞ and lastly they are asymptotically normal (Gupta & Biswas, 2010; Hurlin, 2013).

3.3 Properties of the Estimators

This section considered the asymptotic properties of the estimators which are suppose to be tested on the parameters we estimate from the new distribution. We considered the bias of the individual estimators. Similarly, we expect to test the Mean Square Error (MSE) of the estimators. Further, we tested the estimators for the possibility of consistency and asymptotic normality. We used the Central Limit Theorem (CLT) to test if the distribution of the variable, V asymptotically converge to the standard normal.

Consider $\hat{\vartheta}$ as a random variable. In general, the probability density function (pdf) of $\hat{\vartheta}$, $f(\hat{\vartheta})$, depends on the pdf's of the random variables $\{V\}_{i=1}^{k}$. For the purpose of analysis, we investigate some properties like mean (expected value), variance and standard deviation (spread about the expected value). The expected value of an estimator is related to the concept of the estimator bias and the variance and/or standard deviation is related to the estimator correctness (Zhou et al., 2018).

The word "Asymptotic" means in an infinitely large sample (that is, as the sample size K tends to infinity the sample mean and variance tends to be normally distributed) and this means that, asymptotic results are only approximated in real world situations since getting very large sample is a challenge. The estimators bias and precision are finite sample features, meaning they are properties that hold only for a finite sample size K. Some times, we are focused in studying the properties of estimators when the sample size K gets very large. The very large sample leads to the property of consistency, asymptotic normality and Central Limit Theorem (CLM), which we discuss below (Gupta & Biswas, 2010; Zhou et al., 2018).

1. Bias

   Unbiasedness is a desirable property of any estimator of any given distribution under study, meaning that the estimators gives the correct answer "on average", where "on average" means over many hypothetical observations of the random variable $\{V_i\}_{i=1}^{k}$. The symbol ϑ is used to represent a generic parameter of the population (for example ρ, ϕ^2, p), and the symbol $\hat{\vartheta}$ is used to represent the statistical estimator for ϑ. If the expected value of the estimator is equal to the parameter, that is , if:

   $$E(\hat{\vartheta}) = \vartheta,$$

   the estimator is said to be unbiased. Otherwise, the estimator is said to be biased, that is $B = |E\left(\hat{\vartheta}\right) - \vartheta|$. The bias B is the absolute difference between the expected and the actual value of the parameter.

2. Mean Square Error

   A precise estimate is one in which the variability in the estimation error is small. This estimator is defined as the expected value of the square of the difference between the expected value and the parameter. That is;

   $$MSE(\hat{\vartheta}) = E[(\hat{\vartheta} - \vartheta)^2]$$

   If $E(\hat{\vartheta}) = \vartheta$ then the $MSE(\hat{\vartheta})$ reduces to the $V(\hat{\vartheta})$. This is because,

   $$MSE(\hat{\vartheta}) = E[(\hat{\vartheta} - \vartheta)^2] = V(\hat{\vartheta}) + [E(\hat{\vartheta}) - \vartheta]^2 = V(\hat{\vartheta}) + B^2 \qquad (3.9)$$

Under minimum variance estimator, an estimator is said to be sufficient if the conditional distribution of the random samples given $\hat{\vartheta}$ does not depend on the parameter ϑ for any V_i. And said to be more efficient than another estimator if it is more reliable and precise for the same sample size k.

3. Consistency

   Let $\hat{\vartheta}$ be an estimator of ϑ based on random variable $\{V_i\}_{i=1}^{k}$. An estimator is said to be consistent if the precision and reliability of its estimate improve with increase in sample size. That is, the bias approaches zero as the sample size approaches infinity. Precisely,

$$lim_{K\to\infty} P_r\left(\left|\hat{\vartheta}-\vartheta\right| \geq \epsilon\right) = 0, \epsilon > 0 \tag{3.10}$$

   Laws of large number are also used to induce if an estimator is consistent or not, that is, an estimator $\hat{\vartheta}$ is consistent for ϑ, for K observations if:

i bias$(\hat{\vartheta}, \vartheta) = 0$ as $K \to \infty$

ii MSE$(\hat{\vartheta}, \vartheta) = 0$ as $K \to \infty$

iii se$(\hat{\vartheta}) = 0$ as $K \to \infty$

4. Asymptotic normality

   Let $\hat{\vartheta}$ be an estimator of ϑ based on random variable $\{V_i\}_{i=1}^{k}$. Then an estimator is said to be asymptotically normally if:

$$\hat{\vartheta} \sim N(\vartheta, se(\hat{\vartheta})^2)$$

   for large enough K, meaning that $f(\hat{\vartheta})$ is known to be well approximated by normal distribution with mean ϑ and variance $se(\hat{\vartheta})^2$

5. Central Limit Theorem

The Central Limit Theorem (CLT) states that the sample averages of collection of independently and identically distributed random variables $V_1, V_2, ..., V_K$ with $E(V_i) = \rho$ and $var(V_i) = \psi^2$ is said to asymptotically normal with mean β and variance $\frac{\psi^2}{K}$, and the cumulative density function of the standardized sample mean given as:

$$\frac{\bar{V} - \rho}{se(\bar{V})} = \frac{\bar{V} - \rho}{\frac{\phi}{\sqrt{K}}} = \sqrt{K}(\frac{\bar{V} - \rho}{\phi}),$$

which converges to the cumulative density function of a standard normal random variable Z as $K \rightarrow \infty$, that is:

$$\sqrt{K}(\frac{\bar{V} - \rho}{\phi}) \sim Z \sim N(0, 1),$$

the CLT will helps us to understand the behaviour of the random variable V as the sample size approaches infinity, since it is expected that as the sample size shifts very large, the variance and the mean of the variable should tend to be normally distributed.

3.4 Test of goodness of fit

It is key to verify the suitability and the exactitude of the developed distribution by performing the test of goodness of fit which simply tell us how good is the distribution in comparison to the other families of the existing distributions like Gumbel distribution with two parameters, exponentiated Gumbel distribution, Gumbel geometric distribution, Weibull two parameters distribution and three parameters Frechet distribution. The test of goodness of fit statistic was examined using Akaike Information Criterion (AIC) for investigating the level of efficiency between the distributions under study. Both simulated and real life data was applied to help in understanding and reporting the results for the distribution efficiency. The test is discussed in sections (3.4.1).

3.4.1 Akaike's Information Criterion (AIC)

The Akaike Information Criterion is computed as;

$$AIC = -2logP(W) + 2z, \tag{3.11}$$

where; $logP(W)$ expound the value of the maximized log-likelihood objective function for a model with z parameters to k data points. A smaller AIC value constitute a superior fit, that is, greatest model for fitting the data (Persson & Rydén, 2010). The AIC technique was fitted on the modeled three parameters distribution and the other comparison distributions of the Gumbel family.

3.5 Distributions for comparison with their parameter estimates

For us to compare the estimates obtained from Maximum Likelihood Estimation method and to test efficiency between the new three parameter Gumbel distribution and the already existing distributions, we used the Mean Square Error (MSE) to compare the distributions. The decision was based on the values of bias and/or variance, whereby, the smaller Mean Square Error indicates the smaller error and a better estimator and hence a robust distribution. The following distributions and their parameter estimates were used for efficiency analysis .

1. Gumbel distribution

   Suppose $V_i, i = 1, 2, ..., k$ are random samples, then the Gumbel distribution is given as follows:

   $$f_G(v, \gamma, \vartheta) = \frac{1}{\vartheta} exp\left[- \frac{v - \gamma}{\vartheta} - exp^{\left(-\frac{v-\gamma}{\vartheta} \right)} \right],$$

   with its log-likelihood function given as:

   $$lnR_G(\gamma, \vartheta) = -klog(\vartheta) - \sum_{i=1}^{k} \frac{v - \gamma}{\vartheta} - \sum_{i=1}^{k} exp\big(- \frac{v - \gamma}{\vartheta} \big) \tag{3.12}$$

and its maximum likelihood estimators for location parameter (γ) and scale/dispersion parameter (ϑ) from equation (3.12) satisfies the following equations (Kotz & Nadarajah, 2000).

$$\hat{\gamma} = -\hat{\vartheta} log\left(\frac{1}{k}\sum_{i=1}^{k} exp^{-\frac{V_i}{\hat{\vartheta}}}\right)$$

and

$$\hat{\vartheta} = \bar{X} - \frac{\sum_{i=1}^{k} V_i exp^{-\frac{V_i}{\hat{\vartheta}}}}{\sum_{i=1}^{k} exp^{-\frac{V_i}{\hat{\vartheta}}}} \tag{3.13}$$

where, $\hat{\gamma}$ is the estimator for γ, and $\hat{\vartheta}$ is the estimator for ϑ

2. Exponentiated Gumbel distribution

   Suppose $V_i, i = 1, 2, ..., k$ are random samples, then the Exponentiated Gumbel distribution is given as follows:

$$f_{EG}(v) = \frac{\phi}{\psi} exp\Big(-\frac{v-\rho}{\psi}\Big) exp\Big[-exp\Big(-\frac{v-\rho}{\psi}\Big)\Big]\left[1-exp\left\{-exp\Big(-\frac{v-\rho}{\psi}\Big)\right\}\right]^{\phi-1},$$

   where $-\infty < v < \infty$, and its log-likelihood function represented as;

$$\begin{aligned} lnR(\phi,\psi,\rho) &= klog\phi - k\psi + (\phi-1)\sum_{i=1}^{k} log\Big[1 - exp\Big(-exp\Big(-\frac{v-\rho}{\psi}\Big)\Big)\Big] \\ &\quad -\sum_{i=1}^{k} -\frac{v-\rho}{\psi} - \sum_{i=1}^{k} exp\big(-\frac{v-\rho}{\psi}\big) \end{aligned} \tag{3.14}$$

   hence, the three parameters of the exponentiated Gumbel distribution namely shape parameter (ϕ), scale parameter (ψ) and the location parameter (ρ) are estimated from equation (3.14) as given in the following equation (Nadarajah, 2006).

$$\hat{\phi} = \frac{k}{\phi} + \sum_{i=1}^{k} log\big(1 - exp^{-e - \frac{v_i-\rho}{\psi}}\big),$$

$$\hat{\psi} = -\frac{k}{\psi}\sum_{i=1}^{k}\frac{v_i-\rho}{\psi^2}\big(1 - exp^{-\frac{v_i-\rho}{\psi}}\big) + \frac{\phi-1}{\psi^2}\sum_{i=1}^{k}\frac{(v_i-\rho)exp^{-\frac{v_i-\rho}{\psi}} - exp^{-exp-\frac{v_i-\rho}{\psi}}}{1 - exp^{-exp-\frac{v_i-\rho}{\psi}}}$$

and

$$\hat{\rho} = \frac{k}{\psi} - \frac{1}{\psi}\sum_{i=1}^{k} exp^{-\frac{v_i-\rho}{\psi}} + \frac{\phi-1}{\psi}\sum_{i=1}^{k} \frac{exp^{-\frac{v_i-\rho}{\psi}} - exp^{-exp-\frac{v_i-\rho}{\psi}}}{1-exp^{-exp-\frac{v_i-\rho}{\psi}}} \tag{3.15}$$

where, $\hat{\phi}, \hat{\psi}$ and $\hat{\rho}$ are the estimators.

3. Gumbel Geometric distribution

   Suppose $V_i, i = 1, 2, ..., k$ are random samples, then the probability density function for Gumbel geometric distribution is given as:

$$\begin{aligned} f(v) &= \frac{(1-\gamma)}{\beta} exp\Bigg(-\frac{v-\phi}{\beta} - exp\Bigg(-\frac{v-\phi}{\beta}\Bigg)\Bigg) \times \\ &\quad \Bigg(1-\gamma+\gamma exp\Bigg(-exp\Bigg(-\frac{v-\phi}{\beta}\Bigg)\Bigg)\Bigg)^{-2}, \end{aligned}$$

   and from the probability density function, the log-likelihood function is obtained as follows;

$$\begin{aligned} lnR(v,\gamma,\beta,\phi) &= klog(1-\gamma) - klog(\beta) - 2\sum_{i=1}^{k} log\big(1-\gamma+(\gamma exp(-p_i))\big) + \\ &\quad \sum_{i=1}^{k} log(p_i) - \sum_{i=1}^{k} p_i \end{aligned} \tag{3.16}$$

   by applying Maximum Likelihood Estimation method, that is taking partial derivatives from equation(3.16), the parameters for geometric Gumbel distribution namely shape parameter (γ), location parameter (ϕ) and scale parameter (β) are estimated follows:

$$\hat{\gamma} = -\frac{k}{1-\gamma} - 2\sum_{i=1}^{k} \big(exp(-p_i)-1\big)\big(1-\gamma+\gamma exp(-p_i)\big)^{-1}$$

$$\hat{\phi} = \frac{k}{\beta} - \frac{1}{\beta}\sum_{i=1}^{k} p_i + \frac{2\gamma}{\beta}\sum_{i=1}^{k} p_i exp(-p_i)\big(1-\gamma+\gamma exp(-p_i)\big)^{-1}$$

$$
\begin{aligned}
\hat{\beta} &= -\frac{k}{\beta} - \frac{1}{\beta}\sum_{i=1}^{k} log(p_i) + \frac{1}{\beta}\sum_{i=1}^{k} p_i log(p_i) \\
&\quad -\frac{2\gamma}{\beta}\sum_{i=1}^{k} p_i log(p_i) exp(-p_i)\big(1-\gamma+\gamma exp(-p_i)\big)^{-1} \qquad (3.17)
\end{aligned}
$$

where $p_i = exp\big(-\frac{v_i-\phi}{\beta}\big)$ and $\hat{\gamma}, \hat{\phi}$, and $\hat{\beta}$ are the estimators for shape, location and scale parameters respectively.

4. Weibull two parameters distribution

   Suppose $V_i, i = 1, 2, ..., k$ are random samples, then the probability distribution for Weibull distribution with two parameters namely shape parameter (ϕ) and scale/dispersion parameter (γ) is given as follows:

$$
f(v) = \frac{\phi v^{\phi-1}}{\gamma^{\phi}} exp\Big[-\big(\frac{v}{\gamma}\big)^{\phi}\Big]; (v, \phi, \gamma) > 0,
$$

   From the probability density function, the parameters are estimated by maximizing the likelihood function of the distribution from which we obtain the log likelihood function given as:

$$
lnR(f(v)) = klog\phi - k\phi log\gamma + (\phi - 1)\sum_{i=1}^{k} log(v_i) - \sum_{i=1}^{k}\big(\frac{v_i}{\gamma}\big)^{\phi} \qquad (3.18)
$$

   and the Maximum Likelihood estimates of ϕ and γ are obtained by differentiating the log-likelihood function with respect to the parameters and equating to zero. Given the log-likelihood function as in equation (3.18), the parameters are estimated as indicated in the following equations (Gholami et al., 2020).

$$
\hat{\gamma} = -\frac{k\phi}{\gamma} + \frac{\phi}{\gamma}\sum_{i=1}^{k}\big(\frac{v_i}{\gamma}\big)^{\phi} = \left(\frac{\sum_{i=1}^{k} {v_i}^{\phi}}{k}\right)^{\frac{1}{\phi}},
$$

$$
\hat{\phi} = \sqrt{-\frac{1}{\phi^2} - \big(\frac{v_i}{\gamma}\big)^{\phi}\Big(log\big(\frac{v_i}{\gamma}\big)\Big)^2} \qquad (3.19)
$$

where, $\hat{\gamma}$ and $\hat{\phi}$ are the estimators for scale and shape parameters respectively.

5. Three parameters Frechet distribution

   Suppose $V_i, i = 1, 2, ..., k$ are random samples, then the three parameters Frechet distribution density function is given as (Mansour et al., 2018)

$$f(v) = \frac{\theta\sigma^{\theta}\psi^{2}v^{-\theta-1}exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}{(1+\psi)\Big\{1-exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]\Big\}^{3}}exp\left\{\frac{-\psi exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}{1-exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}\right\}$$

   From the probability distribution function above, the log likelihood function is obtained as:

$$\begin{aligned} lnR &= k\big(log\theta + \theta log\sigma + 2log\psi - log(1+\psi)\big) - (\theta+1)\sum_{i=1}^{k} log(v_i) \\ &\quad -\sum_{i=1}^{k}\big(\frac{\sigma}{v_i}\big)^{\theta} - 3\sum_{i=1}^{k} log\Big\{1-exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]\Big\} \\ &\quad -\psi\sum_{i=1}^{k}\frac{exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}{1-exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]} \end{aligned} \tag{3.20}$$

   and the maximum likelihood estimation of the parameters scale parameter (σ), location parameter (ψ) and the shape parameters (θ) for the three parameters Frechet distribution are given in the following equations

$$\hat{\sigma} = \frac{k\theta}{\sigma} - \frac{\theta}{\sigma}\sum_{i=1}^{k}\big(\frac{\sigma}{v_i}\big)^{\theta} + (\psi-3)\frac{\theta}{\sigma}\sum_{i=1}^{k}\left\{\frac{\big(\frac{\sigma}{v_i}\big)^{\theta}exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}{1-exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}\right\},$$

$$\hat{\psi} = \frac{2k}{\psi} - \frac{k}{(1+\psi)} - \sum_{i=1}^{k}\frac{exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]}{1-exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]},$$

and

$$
\begin{aligned}
\hat{\theta} \quad = \quad & \frac{k}{\theta} + klog\sigma - \sum_{i=1}^{k} log(v_i) - \sum_{i=1}^{k} \Big(\frac{\sigma}{v_i}\Big)^{\theta} log\Big(\frac{\sigma}{v_i}\Big) \\
& +(\psi - 3) \sum_{i=1}^{k} \left\{ \frac{\big(\frac{\sigma}{v_i}\big)^{\theta} exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big] log\big(\frac{\sigma}{v_i}\big)}{1 - exp\Big[-\big(\frac{\sigma}{v_i}\big)^{\theta}\Big]} \right\}, \qquad (3.21)
\end{aligned}
$$

where, $\hat{\sigma}, \hat{\psi}$ and $\hat{\theta}$ are the estimators for scale, location and shape parameters respectively.

CHAPTER FOUR

Results and Discussion

This chapter explains the results obtained from the specific objectives. It explains in details the step by step procedures for deriving the new three parameter Gumbel distribution, estimating its parameters using MLE, describing the parameter properties (bias, MSE, consistency and efficiency) using simulated data, and the application of the distribution to three sets of real life data.

4.1 Three-parameters Gumbel distribution formulation

This section provides a step by step procedure of formulating a three-parameters Gumbel distribution and proving that it is a probability density function (pdf) as discussed in subsection 4.1.1. This chapter also provides information about some characteristic of the three-parameters Gumbel distribution like mean and variance of the distribution as discussed in subsection 4.1.2 and 4.1.3 respectively.

4.1.1 Three parameters Gumbel distribution formulation

Model assumptions

The following assumptions were assumed for a three parameters Gumbel distribution.

1. The new parameters $\omega, \tau, \delta > 0$
2. The parameters (ω, τ, δ) are less than ∞
3. All the parameters are members of real number i.e $(\omega, \tau, \delta \in R)$

4. The distribution obtained is a pdf

As discussed in section 3.1, the formulation of a three parameters Gumbel distribution is done using Marshal Olkin method as given in equation (3.2). Consider a random variable V, then the cdf is given as follows;

$\bar{F}(v) = 1 - F(v)$ whereby $F(v) = \int_{v=0}^{k} f(v)dv$, where $f(v)$ is pdf of the random variable V.

Therefore, having that,

$$F(k) = \int_{v=0}^{k} f(v)dv = \frac{1}{\tau}exp\big(-\frac{v-\omega}{\tau} - exp^{-\frac{v-\omega}{\tau}}\big)dv \tag{4.1}$$

where,

ω is the location parameter

τ is the scale/dispersion parameter

Then using the laws of indices on the expression for $f(v)$ results to,

$$F(k) = \int_{v=0}^{k} exp\big(exp^{-\frac{v-\omega}{\tau}}\big)\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv \tag{4.2}$$

Now, by letting,

$$p = -exp^{-\frac{v-\omega}{\tau}}$$

gives,

$$dp = \frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv$$

Using this substitution in equation (4.2), gives,

$$F(k) = \int_{v=0}^{k} exp^{p}dp = exp^{p}$$

$$= exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big|_{v=0}^{k}$$

$$F(k) = exp\big(-exp^{-\frac{k-\omega}{\tau}}\big) - exp\big(-exp^{\frac{\omega}{\tau}}\big). \tag{4.3}$$

Therefore given that;

$$\bar{F}(v) = 1 - F(v)$$

$$\bar{F}(v) = 1 + exp\big(- exp^{\frac{\omega}{\tau}}\big) - exp\big(- exp^{-\frac{v-\omega}{\tau}}\big) \tag{4.4}$$

With $\bar{F}(v)$, the survival function for the random variable V as illustrated in equation (3.2) is given by:

$$\bar{S}(v) = \frac{\delta\big[1 + exp\big(- exp^{\frac{\omega}{\tau}}\big) - exp\big(- exp^{-\frac{v-\omega}{\tau}}\big)\big]}{1 - \Big\{(1-\delta)\big[1 + exp\big(- exp^{\frac{\omega}{\tau}}\big) - exp\big(- exp^{-\frac{v-\omega}{\tau}}\big)\big]\Big\}} \tag{4.5}$$

with the corresponding cumulative distribution function (cdf) obtained from equation (3.3) as follows,

$$F(v) = \frac{exp\big(- exp^{-\frac{v-\omega}{\tau}}\big) - exp\big(- exp^{\frac{\omega}{\tau}}\big)}{1 - \Big\{(1-\delta)\big[1 + exp\big(- exp^{\frac{\omega}{\tau}}\big) - exp\big(- exp^{-\frac{v-\omega}{\tau}}\big)\big]\Big\}} \tag{4.6}$$

and a probability distribution function (pdf) given as follows (as from equation (3.4),

$$f(v) = \frac{\frac{\delta}{\tau}exp\big(- \frac{v-\omega}{\tau} - exp^{-\frac{v-\omega}{\tau}}\big)}{\Big[1 - \Big\{(1-\delta)\big\{1 + exp\big(- exp^{\frac{\omega}{\tau}}\big) - exp\big(- exp^{-\frac{v-\omega}{\tau}}\big)\big\}\Big\}\Big]^2} \tag{4.7}$$

where δ is the introduced shape parameter and $(\omega, \tau, \delta) > 0$

Three-parameters Gumbel distribution obtained in equation (4.7) is mathematically different from the Generalized Extreme Value distribution in equation (1.4). Also, its shape parameter is estimable ($\delta > 0$) compared to the Generalized Extreme Value distribution where if the shape parameter β equals zero the distribution to fall back to Gumbel distribution with two parameters scale and location. Otherwise if $\beta \neq 0$, then the Generalized Extreme Value distribution fall back to either Weibull or Frechet distribution.

Proving that the developed three-parameters Gumbel distribution is a pdf:

To show that the expression in equation (4.7) is a pdf (that is $\int_{v=0}^{\infty} f(v)dv = 1$), simplify the function by letting

$$q = exp(-exp^{\frac{\omega}{\tau}}).$$

which makes the denominator of the function in equation (4.7) to become,

$$\Big[1-(1-\delta)\big\{1+q-exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}\Big]^2,$$

from which it gives,

$$1-\Big[1-\big\{exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}+q-\delta+\delta\big\{exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}-q\delta\Big]^2,$$

$$\Big[\big\{exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}-q+\delta-\delta\big\{exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}+q\delta\Big]^2,$$

$$\Big[\big\{exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}(1-\delta)+\delta+q\delta-q\Big]^2$$

Therefore, expression in equation (4.7) can be written as:

$$f(v)=\frac{\delta}{\tau}\left[\frac{exp^{-\frac{v-\omega}{\tau}}exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)}{\big[exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)(1-\delta)+\delta+q\delta-q\big]^2}\right], \tag{4.8}$$

Hence getting,

$$\int f(v)dv=\delta\int\frac{exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}}{\big[exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)(1-\delta)+\delta+q\delta-q\big]^2} \tag{4.9}$$

From equation (4.9), by letting

$$p=-exp^{-\frac{v-\omega}{\tau}},$$

and,

$$dp = \frac{1}{\tau} exp^{-\frac{v-\omega}{\tau}} dv,$$

for ease of integration, gives

$$\int f(v)dv = \delta \int \frac{exp^p dp}{\left[e^p(1-\delta) + \delta + q\delta - q\right]^2} \tag{4.10}$$

By rearranging the denominator in equation (4.10), gives,

$$\int f(v)dv = \delta \int \frac{exp^p dp}{\left[(1-\delta)(exp^p - q) + \delta\right]^2} \tag{4.11}$$

from which again if by letting

$$z = (1-\delta)(exp^p - q) + \delta$$

then,

$$dz = (1-\delta)exp^p dp,$$

implying that

$$exp^p dp = \frac{dz}{1-\delta}.$$

Therefore, from equation (4.11), it gives,

$$\int f(v)dv = \frac{\delta}{1-\delta} \int \frac{dz}{z^2} = \frac{\delta}{(\delta-1)z}$$

whose solution becomes,

$$\frac{\delta}{(\delta-1)z} = \frac{\delta}{(\delta-1)\left[(1-\delta)(exp^p - q) + \delta\right]},$$

and by taking the integration from $v = 0$ to k, gives

$$= \frac{\delta}{(\delta-1)\big[(1-\delta)(exp^p - q) + \delta\big]}\Bigg|_{v=0}^{k} \tag{4.12}$$

The foregoing means that taking $v = \infty$ and also $v = 0$, then for $v = \infty$, $exp^p = exp\big(-exp^{-\frac{v-\omega}{\tau}}\big) = 1$ and for $v = 0$, $exp^p = exp\big(-exp^{-\frac{v-\omega}{\tau}}\big) = exp\big(-exp^{\frac{\omega}{\tau}}\big) = q$. Hence, the value for the integral from equation (4.12) becomes:

$$\frac{\delta}{\delta-1}\Big[\frac{1}{(1-\delta)(1-q)+\delta} - \frac{1}{(1-\delta)(q-q)+\delta}\Big],$$

$$\frac{\delta}{\delta-1}\Big[\frac{1}{(1-\delta)(1-q)+\delta} - \frac{1}{\delta}\Big],$$

$$\frac{\delta}{\delta-1}\Big[\frac{\delta-1+q-q\delta}{\big[(1-\delta)(1-q)+\delta\big]\delta}\Big],$$

Since, $\delta - 1 + q - q\delta$ can be factored as $\delta - 1 + q(1-\delta) = 1(\delta-1) - q(\delta-1) = (1-q)(\delta-1)$. Also, $(1-\delta)(1-q)+\delta$ in the denominator can be simplified as $1 - q - \delta + q\delta + \delta = 1 - q + q\delta$, then it gives,

$$\frac{\delta}{\delta-1}\Big[\frac{(1-q)(\delta-1)}{\delta[1-q+q\delta]}\Big] = \frac{1-q}{1-q+q\delta}. \tag{4.13}$$

But, knowing that,

$$q = exp(-exp^{\frac{\omega}{\tau}})$$

and since, $\omega > \tau$, then it follows that $q \to 0$. And therefore, equation (4.13) approaches 1, that is,

$$\frac{1-q}{1-q+q\delta} \to 1,$$

hence $f(v)$ is a pdf as required.

4.1.2 Expected value of a three parameters Gumbel distribution

In this subsection, we are deriving the expected value of the three parameters Gumbel distribution. This is because we cannot assume that the location parameter is the mean since the measures of location are mean, mode and median.

Given $f(v)$ as the probability distribution function, the expected value $(E(V))$ is obtained as;

$$E(v) = \int vf(v)dv$$

$$E(v) = \int_{-\infty}^{\infty} \frac{v\frac{\delta}{\tau}exp\big(-\frac{v-\omega}{\tau} - exp^{-\frac{v-\omega}{\tau}}\big)dv}{\Big[1 - \Big\{(1-\delta)\big\{1 + exp\big(-exp^{\frac{\omega}{\tau}}\big) - exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}\Big\}\Big]^2} \quad (4.14)$$

$$E(v) = \delta\int_{-\infty}^{\infty} \frac{v.exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv}{\Big[1 - \Big\{(1-\delta)\big\{1 + exp\big(-exp^{\frac{\omega}{\tau}}\big) - exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}\Big\}\Big]^2}, \quad (4.15)$$

by letting

$$z = exp^{-\frac{v-\omega}{\tau}} \Longrightarrow ln(z) = -\frac{v-\omega}{\tau} = \frac{-v}{\tau} + \frac{\omega}{\tau} \Longrightarrow v = \omega - \tau ln(z)$$

then it gives,

$$dz = -\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv$$

by substituting v, z, and dz in equation (4.15), gives

$$E(v) = -\delta\int \frac{(\omega - \tau ln(z)exp^{-z}dz}{\big[1 - \big((1-\delta)(1 - exp^{-z} + q)\big)\big]^2} \quad (4.16)$$

where

$$q = exp\big(-exp^{\frac{\omega}{\tau}}\big)$$

Therefore,

$$E(v) = -\delta\int \frac{\omega exp^{-z}dz}{\big[1 - \big((1-\delta)(1 - exp^{-z} + q)\big)\big]^2} + (\delta\tau)\int \frac{ln(z)exp^{-z}dz}{\big[1 - \big((1-\delta)(1 - exp^{-z} + q)\big)\big]^2} \quad (4.17)$$

The function $E(v)$ have two terms which are integrated as follows, for the first term from equation (4.17), that is;

$$-\delta\int\frac{\omega exp^{-z}dz}{\left[1-\left((1-\delta)(1-exp^{-z}+q)\right)\right]^2}$$

letting $p=(1-\delta)(exp^{-z}-q)+\delta$ then, $dp=(1-\delta)(-exp^{-z})=(\delta-1)exp^{-z}dz$ implying that $exp^{-z}dz=\frac{dp}{\delta-1}$. This gives,

$$-\delta\int\frac{\omega exp^{-z}dz}{\left[1-\left((1-\delta)(1-exp^{-z}+k)\right)\right]^2}=\frac{-(\delta\omega)}{\delta-1}\int\frac{dp}{p^2}$$

$$=\frac{-(\delta\omega)}{\delta-1}\left[\frac{1}{p}\right]=\frac{-(\delta\omega)}{\delta-1}\left[\frac{1}{1-\delta)(exp^{-z}-q)+\delta}\right]$$

which is given as follows after re-substituting z

$$=\frac{-(\delta\omega)}{\delta-1}\left[\frac{1}{(1-\delta)(exp^{-exp^{-\frac{v-\omega}{\tau}}}-q)+\delta}\right]\Big|_{-\infty}^{\infty}$$

this gives the following result after taking the integral with the limits of v from $-\infty$ to ∞

$$\begin{aligned}&=\frac{-(\delta\omega)}{\delta-1}\left[\frac{1}{(1-\delta)(1-q)+\delta}-\frac{1}{-q(1-\delta)+\delta}\right]\\&=\frac{-(\delta\omega)}{\delta-1}\left[\frac{-q(1-\delta)+\delta-[(1-\delta)(1-q)+\delta]}{[(1-\delta)(1-q)+\delta][-q(1-\delta)+\delta]}\right]\end{aligned}$$

the integral for the first term finally becomes

$$\frac{\delta\omega}{(\delta q+\delta-q)(1-q+\delta q)} \tag{4.18}$$

For the second terms in equation (4.17), which is given as,

$$(\delta\tau)\int\frac{ln(z)exp^{-z}dz}{\left[1-\left((1-\delta)(1-exp^{-z}+q)\right)\right]^2}=(\delta\tau)\int\frac{ln(z)exp^{-z}}{[(1-\delta)(exp^{-z}-q)+\delta]^2}\Big]$$

Integration by parts is used to integrate the function. By applying the integration by parts method, it follows that, $wp - \int pdw$ is applied to solve the function. Let $w = ln(z)$ giving $dw = \frac{dz}{z}$. Also by letting

$$dp = \frac{exp^{-z}dz}{\big[(1-\delta)(exp^{-z} - q) + \delta\big]^2}$$

implying that

$$p = \int \frac{exp^{-z}dz}{\big[(1-\delta)(exp^{-z} - q) + \delta\big]^2} = \frac{-1}{(\delta - 1)[(1-\delta)(exp^{-z} - q) + \delta]}$$

this gives the integral result as

$$\int \frac{ln(z)exp^{-z}}{[(1-\delta)(exp^{-z} - q) + \delta]^2}\Big] = wp - \int pdw$$

$$= \frac{-ln(z)}{(\delta - 1)[(1-\delta)(exp^{-z} - q) + \delta]} + \frac{1}{\delta - 1}\int \frac{dz}{z[(1-\delta)(exp^{-z} - q) + \delta]} \tag{4.19}$$

Equation (4.19) shows that the function still requires integration by parts for the second term, that is

$$\frac{1}{\delta - 1}\int \frac{dz}{z[(1-\delta)(exp^{-z} - q) + \delta]}.$$

By letting

$$w = \frac{1}{[(1-\delta)(exp^{-z} - q) + \delta]}$$

implying that

$$dw = -[(1-\delta)(exp^{-z} - q) + \delta]^{-2}(\delta - 1)exp^{-z}.$$

Also, by taking $dp = \frac{dz}{z}$ gives $p = ln(z)$. Therefore, getting

$$\frac{1}{\delta - 1}\Big[\frac{ln(z)}{[(1-\delta)(exp^{-z} - q) + \delta]} + (\delta - 1)\int \frac{ln(z)exp^{-z}}{[(1-\delta)(exp^{-z} - q) + \delta]^2}\Big] \tag{4.20}$$

Replacing equation (4.20) in equation (4.19), gives 0 implying that the function is indeterminate and therefore it is undefined. The expected value is therefore given as;

$$E(v) = \frac{\delta\omega}{(\delta q + \delta - q)(1 - q + \delta q)}$$
$$= \frac{\delta\omega}{[\delta exp(-e^{\frac{\omega}{\tau}}) + \delta - exp(-e^{\frac{\omega}{\tau}})][1 - exp(-e^{\frac{\omega}{\tau}}) + \delta exp(-e^{\frac{\omega}{\tau}})]} \tag{4.21}$$

4.1.3 The variance of a three parameters Gumbel distribution

This subsection discuss the variance of the three parameters Gumbel distribution. Knowing that, $Var(v) = E(v^2) - [E(v)]^2$. The $E(v)$ is already obtained in subsection 4.1.2. Now $E(v^2)$ is obtained as follows.

$$E(v^2) = \delta \int_{-\infty}^{\infty} \frac{v^2.exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv}{\Big[1 - \Big\{(1-\delta)\big\{1 + exp\big(-exp^{\frac{\omega}{\tau}}\big) - exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\big\}\Big\}\Big]^2}$$

$$E(v^2) = \delta \int_{-\infty}^{\infty} \frac{v^2.exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv}{\Big[(1-\delta)\big(exp^{(-exp^{\frac{-v-\omega}{\tau}})} - q\big) + \delta\Big]^2}$$

In subsection 4.1.2, letting $q = exp\big(-exp^{\frac{\omega}{\tau}}\big)$ and $z = exp^{-\frac{v-\omega}{\tau}} \Longrightarrow v = \omega - \tau ln(z)$ hence, $v^2 = (\omega - \tau ln(z))^2 \Longrightarrow v^2 = \omega^2 - 2\omega\tau ln(z) + \tau^2[ln(z)]^2$. From z function, it gives

$$dz = -\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}dv$$

Therefore, getting,

$$\begin{aligned} E(v^2) &= \delta \int_{-\infty}^{\infty} \frac{v^2.exp\big(-exp^{-\frac{v-\omega}{\tau}}\big)\frac{1}{\tau}exp^{-\frac{v-\omega}{\tau}}}{\Big[(1-\delta)\big(exp^{(-e^{\frac{-v-\omega}{\tau}})} - q\big) + \delta\Big]^2} \\ &= -\delta \int \frac{\big(\omega^2 - 2\omega\tau ln(z) + \tau^2[ln(z)]^2\big)exp^{-z}dz}{[(1-\delta)(exp^{-z} - q) + \delta]^2} \end{aligned}$$

$$
\begin{aligned}
E(v^2) &= -\delta\omega^2 \int \frac{exp^{-z}dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} + 2\delta\omega\tau \int \frac{exp^{-z}ln(z)dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} \\
&\quad -\delta\tau^2 \int \frac{exp^{-z}[ln(z)]^2dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} \qquad (4.22)
\end{aligned}
$$

As evidence in subsection 4.1.2, the first term of $E(v^2)$ from equation (4.22) which is

$$
\int \frac{exp^{-z}dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} = \frac{1}{(\delta q+\delta-q)(1-q+q\delta)}.
$$

Also, the integral for the second term, that is

$$
\delta\omega\tau \int \frac{exp^{-z}ln(z)dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} = 0
$$

as evidence from equations (4.19) and (4.20).

This means that the only remaining integral solution is for the third term of $E(v^2)$ in equation (4.22) which is given as

$$
-\delta\tau^2 \int \frac{exp^{-z}[ln(z)]^2dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2}
$$

to integrate this function, apply integration by part again (that is, $wp - \int pdw$). By letting $w = (ln(z))^2 \Longrightarrow dw = \frac{2ln(z)dz}{z}$. Again, letting,

$$
dp = \frac{exp^{-z}dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2}
$$

implying that

$$
p = \int \frac{exp^{-z}dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} = \frac{-1}{(\delta-1)[(1-\delta)(exp^{-z}-q)+\delta]}
$$

Therefore, leading to the $wp - \int pdw$ for the function given as;

$$
\begin{aligned}
\int \frac{exp^{-z}[ln(z)]^2dz}{[(1-\delta)(exp^{-z}-q)+\delta]^2} &= \frac{-[ln(z)]^2}{(\delta-1)[(1-\delta)(exp^{-z}-q)+\delta]} \\
&\quad + \frac{2}{\delta-1} \int \frac{ln(z)dz}{z[(1-\delta)(exp^{-z}-q)+\delta]} \qquad (4.23)
\end{aligned}
$$

Using integration by parts again in the second term of equation (4.23), by letting

$$w = \frac{1}{[(1-\delta)(exp^{-z} - q) + \delta]} \Longrightarrow dw = -[(1-\delta)(exp^{-z} - q) + \delta]^2(\delta - 1)exp^{-z}.$$

And, also letting

$$dp = \frac{ln(z)dz}{z} \Longrightarrow p = \frac{[ln(z)]^2}{2}.$$

This gives,

$$\begin{aligned}\frac{2}{\delta-1}\int \frac{ln(z)dz}{z[(1-\delta)(exp^{-z} - q) + \delta]} &= \frac{2}{\delta-1}\Big[\frac{[ln(z)]^2}{2[(1-\delta)(exp^{-z} - q) + \delta]} \\ &+ \frac{\delta-1}{2}\int \frac{[ln(z)]^2 exp^{-z}}{[(1-\delta)(exp^{-z} - q) + \delta]^2}\Big]\end{aligned} \tag{4.24}$$

Replacing equation (4.24) in equation (4.23), gives 0 implying that the function is indeterminate and therefore it is undefined. The $E(v^2)$ is therefore given as;

$$E(v^2) = \frac{\delta\omega^2}{(\delta q + \delta - q)(1 - q + q\delta)} \tag{4.25}$$

Therefore, $Var(v) = E(v^2) - [E(v)]^2$ is obtained from equations (4.21) and (4.25) as folows;

$$\begin{aligned} Var(v) &= \frac{\delta\omega^2}{(\delta q + \delta - q)(1 - q + q\delta)} - \Big[\frac{\delta\omega}{(\delta q + \delta - q)(1 - q + \delta q)}\Big]^2 \\ &= \frac{\delta\omega^2}{(\delta q + \delta - q)(1 - q + q\delta)} - \frac{\delta^2\omega^2}{[(\delta q + \delta - q)(1 - q + \delta q)]^2} \\ &= \frac{\delta\omega^2[(\delta q + \delta - q)(1 - q + \delta q)] - \delta^2\omega^2}{[(\delta q + \delta - q)(1 - q + \delta q)]^2} \\ &= \frac{\delta\omega^2\big([(\delta q + \delta - q)(1 - q + \delta q)] - \delta\big)}{[(\delta q + \delta - q)(1 - q + \delta q)]^2} \end{aligned}$$

Replacing q gives us,

$$Var(v) = \frac{\delta\omega^2\big([(\delta exp(-e^{\frac{\omega}{\tau}}) + \delta - exp(-e^{\frac{\omega}{\tau}}))(1 - exp(-e^{\frac{\omega}{\tau}}) + \delta exp(-e^{\frac{\omega}{\tau}}))] - \delta\big)}{[(\delta exp(-e^{\frac{\omega}{\tau}}) + \delta - exp(-e^{\frac{\omega}{\tau}}))(1 - exp(-e^{\frac{\omega}{\tau}}) + \delta exp(-e^{\frac{\omega}{\tau}}))]^2} \tag{4.26}$$

4.2 Parameter Estimation

This section demonstrates Maximum Likelihood Estimation technique for the purpose of estimating the parameters of a three-parameter Gumbel distribution. The process of estimating each of the three parameters namely; the location parameters (ω), scale/dispersion parameter (τ) and the shape parameter (δ) using Maximum Likelihood Estimation method in subsections 4.2.2, 4.2.3 and 4.2.4 respectively. The Maximum Likelihood Estimation method involves three steps, that is, getting the likelihood function, the log of the likelihood function and the derivative with respect to the required parameter.

4.2.1 Maximum Likelihood Estimation for the parameters

Considering the probability distribution function given in equation (4.7), its likelihood function is given follows,

$$\begin{aligned} R &= \prod_{i=1}^{k} f(v_i) \\ &= \prod_{i=1}^{k} \frac{\delta}{\tau}\left[\frac{exp\big(-\frac{v_i-\omega}{\tau} - exp^{\frac{-v_i-\omega}{\tau}}\big)}{\big[1-(1-\delta)\big(1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big)\big]^2}\right] \end{aligned}$$

where R is the symbol used in this study to represent likelihood function

$$R(v_i;\delta,\omega,\tau) = \big(\frac{\delta}{\tau}\big)^k \frac{exp\sum_{i=1}^{k}\big(-\frac{v_i-\omega}{\tau} - exp^{\frac{-v_i-\omega}{\tau}}\big)}{\Big[exp\sum_{i=1}^{k} ln\big[1-(1-\delta)\big(1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big)\big]\Big]^2} \tag{4.27}$$

The likelihood function R, can be expressed with a variable together with parameters to be estimated as shown in equation (4.27), and its log-likelihood function which maximizes the parameters becomes

$$ln(R) = k\big\{ln(\frac{\delta}{\tau})\big\} + \sum_{i=1}^{k}\Big[-(\frac{v_i-\omega}{\tau}) - exp^{-\left(\frac{v_i-\omega}{\tau}\right)}\Big]$$

$$-2\sum_{i=1}^{k} ln\big[1-(1-\delta)\big(1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big)\big] \tag{4.28}$$

4.2.2 Parameters estimation for δ

With the log likelihood function, the estimate for δ is obtained by performing the partial derivative for the log likelihood function with respect to δ and equating the result to zero (that is, computing $\frac{\partial ln(R)}{\partial\delta}$). This is done by considering the terms in equation (4.28) containing δ because it is known that terms independent of δ will definitely give a derivative result of zero. In equation (4.28), the first term (that is, $k\big\{ln\big(\frac{\delta}{\tau}\big)\big\}$, which is differentiated with respect to δ as follows:

$$d = k\big\{ln(\frac{\delta}{\tau})\big\}$$

then,

$$\frac{\partial d}{\partial\delta} = \frac{\tau}{\delta}.\frac{k}{\tau} = \frac{k}{\delta} \tag{4.29}$$

And letting c to represent the second term with δ then,

$$c = -2\sum_{i=1}^{k} ln\big[1-(1-\delta)\big(1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big)\big],$$

which is differentiated with respect to δ as follows:

by letting,

$$p = 1-(1-\delta)\big[1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big]$$

$$= 1-1\big[1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big]+\delta\big[1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big]$$

Therefore, it gives,

$$\frac{dp}{\partial\delta} = \big[1 - exp\big(- exp^{\frac{-v_i - \omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big]$$

Now differentiating c with respect to p, gives,

$$\frac{dc}{dp} = \sum_{i=1}^{k}(lnp) = 2\sum_{i=1}^{k}\frac{1}{p} \tag{4.30}$$

This leads to,

$$\frac{\partial c}{\partial\delta} = 2\sum_{i=1}^{k}\frac{\big[1 - exp\big(- e^{\frac{-v_i - \omega}{\tau}}\big) + exp(-e^{\frac{\omega}{\tau}})\big]}{p} \tag{4.31}$$

Thus we have the estimate of δ obtained from equations(4.29) and (4.31) as,

$$\hat{\delta} = \frac{\partial ln(L)}{\partial\delta} = \frac{k}{\delta} - 2\sum_{i=1}^{k}\frac{\big[1 - exp\big(- exp^{\frac{-v_i - \omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big]}{p} = 0, \tag{4.32}$$

where,

$$p = 1 - (1 - \delta)\big[1 - exp\big(- exp^{\frac{-v_i - \omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big]$$

4.2.3 Parameters estimation for ω

This subsection shows the process of estimating the location parameter ω by the use of maximum likelihood estimate. By using the log likelihood function in equation (4.28) and working out the partial derivative while equating to zero, it is assumed that the term in the equation without ω have their partial derivatives equals to zero hence leaving the study to work out partial derivatives of only terms having ω as follows;

For the second term whereby

$$\sum_{i=1}^{k}\Big[- \big(\frac{v_i - \omega}{\tau}\big) - exp^{-\left(\frac{v_i - \omega}{\tau}\right)}\Big]$$

by letting

$$r = -\big(\frac{v_i - \omega}{\tau}\big) = -\frac{v_i}{\tau} + \frac{\omega}{\tau}$$

This results to,

$$\frac{\partial r}{\partial \omega} = \frac{1}{\tau} \tag{4.33}$$

Also, by letting

$$s = exp^{-\left(\frac{v_i - \omega}{\tau}\right)}$$

and,

$$r = -\big(\frac{v_i - \omega}{\tau}\big)$$

then

$$\frac{\partial s}{\partial \omega} = \frac{1}{\tau} exp^{r} = \frac{1}{\tau} exp^{-\left(\frac{v_i - \omega}{\tau}\right)} \tag{4.34}$$

Combining $\frac{\partial r}{\partial \omega}$, $\frac{\partial s}{\partial \omega}$, results to the derivative of the second term with respect to ω given as;

$$\begin{aligned} &\sum_{i=1}^{k} \Big[\frac{1}{\tau} - \frac{1}{\tau} exp^{-\left(\frac{v_i - \omega}{\tau}\right)}\Big] \\ &\frac{1}{\tau}\sum_{i=1}^{k} \Big[1 - exp^{-\left(\frac{v_i - \omega}{\tau}\right)}\Big] \end{aligned} \tag{4.35}$$

For the third term letting,

$$t = -2\sum_{i=1}^{k} ln\big[1 - (1-\delta)\big(1 - exp\big(-exp^{\frac{-v_i - \omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big)\big]$$

and also by letting,

$$p = 1 - (1-\delta)\big[1 - exp\big(-exp^{\frac{-v_i - \omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big],$$

obtains,

$$\frac{\partial t}{\partial \omega} = 2(1-\delta)\sum_{i=1}^{k}\left[\frac{exp\Big(-e^{\left(-\frac{v_i-\omega}{\tau}\right)}\Big).e^{\left(-\frac{v_i-\omega}{\tau}\right)} - e^{\frac{\omega}{\tau}}.exp(-e^{\frac{\omega}{\tau}})}{p\tau}\right] \tag{4.36}$$

Combining $\frac{\partial r}{\partial \omega}$, $\frac{\partial s}{\partial \omega}$ and $\frac{\partial t}{\partial \omega}$ as given in equations (4.33), (4.34) and (4.36) respectively gives the maximum likelihood estimate for ω as;

$$\hat{\omega} = \frac{\partial ln(L)}{\partial \omega} = \frac{1}{\tau}\sum_{i=1}^{k}\left[1 - exp^{-\left(\frac{v_i-\omega}{\tau}\right)}\right]$$
$$+2(1-\delta)\sum_{i=1}^{k}\left[\frac{exp\Big(-e^{\left(-\frac{v_i-\omega}{\tau}\right)}\Big).e^{\left(-\frac{v_i-\omega}{\tau}\right)} - e^{\frac{\omega}{\tau}}.exp(-e^{\frac{\omega}{\tau}})}{p\tau}\right] = 0 \tag{4.37}$$

where,

$$p = 1 - (1-\delta)\big[1 - exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big]$$

4.2.4 Parameters estimation for τ

This subsection explains the process of obtaining the maximum likelihood estimator for τ by partially differentiating the log likelihood function given below with respect to τ and equating to zero as follows.

$$ln(R) = k\big\{ln\big(\frac{\delta}{\tau}\big)\big\} + \sum_{i=1}^{k}\left[-\big(\frac{v_i-\omega}{\tau}\big) - exp^{-\left(\frac{v_i-\omega}{\tau}\right)}\right]$$
$$-2\sum_{i=1}^{k} ln\big[1-(1-\delta)\big(1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big)\big]$$

The first term of the likelihood function is differentiated using chain rule as follows, let,

$$p = k\big\{ln\big(\frac{\delta}{\tau}\big)\big\}$$

therefore,

$$\frac{\partial p}{\partial \tau} = \frac{-k}{\tau} \tag{4.38}$$

For the second term of the log likelihood function, letting

$$m = -\big(\frac{v_i - \omega}{\tau}\big),$$

$$\frac{\partial m}{\partial \tau} = (v_i - \omega)\tau^{-1} = (v_i - \omega)\tau^{-2} = \frac{v_i - \omega}{\tau^2} \tag{4.39}$$

Again by letting
$r = exp^{-\left(\frac{v_i-\omega}{\tau}\right)}$, and first dealing with the term in the parenthesis, by letting this term be $m = -\left(\frac{v_i-\omega}{\tau}\right)$.
results to,

$$\frac{\partial r}{\partial \tau} = (exp^r).\frac{v_i - \omega}{\tau^2} = \frac{v_i - \omega}{\tau^2}.exp^{-\left(\frac{v_i-\omega}{\tau}\right)} \tag{4.40}$$

Therefore, the derivative of the second term with respect to τ becomes,

$$\begin{aligned} &\sum_{i=1}^{k}\Big[\frac{v_i - \omega}{\tau^2} - \frac{v_i - \omega}{\tau^2}.exp^{-\left(\frac{v_i-\omega}{\tau}\right)}\Big] \\ &\frac{1}{\tau^2}\sum_{i=1}^{k}\Big[v_i - \omega\Big]\Big[1 - exp^{-\left(\frac{v_i-\omega}{\tau}\right)}\Big] \end{aligned} \tag{4.41}$$

The third term of the function is,

$$t = -2\sum_{i=1}^{k} ln\big[1 - (1-\delta)\big(1 - exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big)\big],$$

about which letting,

$$p = 1 - (1-\delta)\big(1 - exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big) + exp(-exp^{\frac{\omega}{\tau}})\big)$$

The function for t has two terms with parameter τ which are differentiated with respect to τ as follows.

For,

$$g = -exp\big(- exp^{\frac{-v_i - \omega}{\tau}}\big),$$

letting

$$h = -exp^{\frac{-v_i - \omega}{\tau}},$$

and,

$$j = \frac{-v_i - \omega}{\tau},$$

then,

$$\frac{dh}{d\tau} = \frac{dh}{dj}.\frac{dj}{d\tau} = -exp^t.\frac{v_i - \omega}{\tau^2} = -exp^{\frac{-v_i - \omega}{\tau}}.\frac{v_i - \omega}{\tau^2}$$

Hence,

$$\begin{aligned}
\frac{dg}{d\tau} &= \frac{dg}{dh}.\frac{dh}{d\tau} \\
&= -exp^h. - exp^{\frac{-v_i - \omega}{\tau}}.\frac{v_i - \omega}{\tau^2} \\
&= exp\big(- exp^{\frac{-v_i - \omega}{\tau}}\big).exp^{\frac{-v_i - \omega}{\tau}}.\frac{v_i - \omega}{\tau^2} \qquad (4.42)
\end{aligned}$$

The last term in the t function is $exp(-exp^{\frac{\omega}{\tau}})$ for which is written by letting $h = -exp^{\frac{\omega}{\tau}}$ and the power is written as $j = \frac{\omega}{\tau}$.

From which,

$$\frac{dh}{d\tau} = \frac{d}{dj}.\frac{dj}{d\tau} = -exp^t. - \frac{\omega}{\tau^2} = \frac{\omega}{\tau^2}exp^{\frac{\omega}{\tau}},$$

Hence,

$$\begin{aligned}
\frac{d}{dh}.\frac{dh}{d\tau} &= exp^h.\frac{\omega}{\tau^2}exp^{\frac{\omega}{\tau}} \\
&= \frac{\omega}{\tau^2}exp(-exp^{\frac{\omega}{\tau}})exp^{\frac{\omega}{\tau}} \qquad (4.43)
\end{aligned}$$

Using equations (4.42) and (4.43), the derivative for the t function with respect to τ becomes,

$$\frac{\partial p}{\partial \tau} = -(1 - \delta)\Big[\frac{v_i - \omega}{\tau^2}exp\big(- e^{\frac{-v_i - \omega}{\tau}}\big).e^{\frac{-v_i - \omega}{\tau}} + \frac{\omega}{\tau^2}exp(-e^{\frac{\omega}{\tau}})e^{\frac{\omega}{\tau}}\Big]$$

Because of equation (4.30), the final derivative for the third term in the log likelihood function with respect to τ is

$$\frac{\partial t}{\partial \tau} = -2(1-\delta)\sum_{i=1}^{k} \frac{(v_i-\omega)exp\big(-e^{\frac{-v_i-\omega}{\tau}}\big).e^{\frac{-v_i-\omega}{\tau}} + \omega.exp(-e^{\frac{\omega}{\tau}})e^{\frac{\omega}{\tau}}}{\tau^2 p} = Q \quad (4.44)$$

where,

$$p = 1-(1-\delta)\big[1-exp\big(-exp^{\frac{-v_i-\omega}{\tau}}\big)+exp(-exp^{\frac{\omega}{\tau}})\big].$$

Therefore, the estimator for τ is given as follows from equations (4.38), (4.41) and (4.44) above

$$\hat{\tau} = \frac{\partial ln(L)}{\partial \tau} = \frac{-k}{\tau} + \frac{1}{\tau^2}\sum_{i=1}^{k}\Big[v_i-\omega\Big]\Big[1-exp^{-\left(\frac{v_i-\omega}{\tau}\right)}\Big] + Q = 0 \quad (4.45)$$

where Q is defined in equation (4.44)

4.3 Properties of the Estimators

This section discusses the symmetrical behaviour of a three-parameter Gumbel distribution, asymptotic properties of the estimators with a view to investigating their asymptotic bias, whereby we investigate if each of the three parameters are asymptotically unbiased as the sample size becomes large. We also investigate their Mean Square Error (MSE) in subsection (4.3.3) in order to ascertain if in each case the MSE tends to zero as the sample size approaches infinity (that is, the sample size becomes very large). We also discuss the consistency of the parameters in subsection (4.3.4) to investigate the precision and reliability of the estimators as we increase the sample size. Finally, in subsection (4.3.5), we compare the asymptotic relationship between a three parameters Gumbel distribution and other distributions like normal, chi-square and Weibull respectively using simulated data sets of different sample sizes.

4.3.1 Investigating the symmetrical behaviour of a three-parameters Gumbel distribution using simulation

Knowing the type of distribution is very important before using the distribution for modelling or analyzing any type of data. The well known types of distribution categories are discrete or continuous. The introduced three-parameters Gumbel distribution is continuous distribution like the mother two parameters Gumbel distribution. A distribution can be continuous but symmetrical, left/negatively skewed or right skewed/positively skewed.

To simulate data, the study used inverse transformation technique by making v the subject from the cdf function of a three-parameters Gumbel distribution. The cdf function is given as follows from equation 4.6.

$$F(v) = \frac{exp\Big(-e^{-\frac{v-\omega}{\tau}}\Big) - exp\Big(-e^{\frac{\omega}{\tau}}\Big)}{1 - \Big\{(1-\delta)\Big[1 + exp\Big(-e^{\frac{\omega}{\tau}}\Big) - exp\Big(-e^{-\frac{v-\omega}{\tau}}\Big)\Big]\Big\}}$$

let

$$U = F(v) = \frac{exp\big(-e^{-\frac{v-\omega}{\tau}}\big) - exp\big(-e^{\frac{\omega}{\tau}}\big)}{1 - \Big\{(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big) - exp\big(-e^{-\frac{v-\omega}{\tau}}\big)\big]\Big\}} \tag{4.46}$$

where U follows a uniform distribution. Then it is solved as follows to make v the subject with respect to U from equation 4.46.

$$U\Big[1 - \Big\{(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big) - exp\big(-e^{-\frac{v-\omega}{\tau}}\big)\big]\Big\}\Big] = exp\big(-e^{-\frac{v-\omega}{\tau}}\big) - exp\big(-e^{\frac{\omega}{\tau}}\big)$$

$$U - U(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big) - exp\big(-e^{-\frac{v-\omega}{\tau}}\big)\big] = exp\big(-e^{-\frac{v-\omega}{\tau}}\big) - exp\big(-e^{\frac{\omega}{\tau}}\big)$$

$$U - U(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big)\big] + U(1-\delta)\big[exp\big(-e^{-\frac{v-\omega}{\tau}}\big)\big] = exp\big(-e^{-\frac{v-\omega}{\tau}}\big) - exp\big(-e^{\frac{\omega}{\tau}}\big) \tag{4.47}$$

collecting the like terms in equation 4.47 gives,

$$U(1-\delta)\big[exp\big(-e^{-\frac{v-\omega}{\tau}}\big)\big] - exp\big(-e^{-\frac{v-\omega}{\tau}}\big) = -U + U(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big)\big] - exp\big(-e^{\frac{\omega}{\tau}}\big)$$

factoring the LHS gives,

$$exp\big(-e^{-\frac{v-\omega}{\tau}}\big)\big[U(1-\delta)\big] = -U + U(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big)\big] - exp\big(-e^{\frac{\omega}{\tau}}\big)$$

hence,

$$exp\big(-e^{-\frac{v-\omega}{\tau}}\big) = \frac{U(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big)\big] - U - exp\big(-e^{\frac{\omega}{\tau}}\big)}{\big[U(1-\delta)\big]} \tag{4.48}$$

To remove the exponential from equation 4.48, logarithm is applied twice as follows:

Taking the first logarithm gives,

$$-exp^{-\frac{v-\omega}{\tau}} = ln\Big[\frac{U(1-\delta)\big[1 + exp\big(-e^{\frac{\omega}{\tau}}\big)\big] - U - exp\big(-e^{\frac{\omega}{\tau}}\big)}{\big[U(1-\delta)\big]}\Big]$$

To remove the negative in the LHS, both sides are diveded by -1 which results to:

$$exp^{-\frac{v-\omega}{\tau}} = ln\Big[\frac{\big[U(1-\delta)\big]}{U(1-\delta)\big[1+exp\big(-e^{\frac{\omega}{\tau}}\big)\big]-U-exp\big(-e^{\frac{\omega}{\tau}}\big)}\Big] \tag{4.49}$$

Taking the second logarith from equation 4.49, results to;

$$-\frac{v-\omega}{\tau} = ln\left[ln\Big[\frac{\big[U(1-\delta)\big]}{U(1-\delta)\big[1+exp\big(-e^{\frac{\omega}{\tau}}\big)\big]-U-exp\big(-e^{\frac{\omega}{\tau}}\big)}\Big]\right]$$

$$\frac{v+\omega}{\tau} = -ln\left[ln\Big[\frac{\big[U(1-\delta)\big]}{U(1-\delta)\big[1+exp\big(-e^{\frac{\omega}{\tau}}\big)\big]-U-exp\big(-e^{\frac{\omega}{\tau}}\big)}\Big]\right]$$

$$v+\omega = -\tau ln\left[ln\Big[\frac{\big[U(1-\delta)\big]}{U(1-\delta)\big[1+exp\big(-e^{\frac{\omega}{\tau}}\big)\big]-U-exp\big(-e^{\frac{\omega}{\tau}}\big)}\Big]\right]$$

$$v_i = -\omega-\tau ln\left[ln\Big[\frac{\big[U(1-\delta)\big]}{U(1-\delta)\big[1+exp\big(-e^{\frac{\omega}{\tau}}\big)\big]-U-exp\big(-e^{\frac{\omega}{\tau}}\big)}\Big]\right] \tag{4.50}$$

where;

ω = location parameter > 0

τ = scale/dispersion parameter > 0

δ = shape parameter > 0

U = is the transformation which follows uniform distribution.

To comment on the nature of symmetry of a three-parameter Gumbel distribution, inverse transformation method was used to simulate data using equation 4.50 and the results are as discussed below:

When the location parameter (ω) is constant and n = 10000

Setting the location parameter constant and varying scale and shape parameters at n = 10000, a three-parameters Gumbel distribution appears as given figure 4.1. Figure 4.1, shows that if a three-parameters Gumbel distribution is used to simulate data with a fixed location parameter, the distribution appeared to the skewed to the right (positively skewed). The two-parameter Gumbel distribution which was the baseline for modelling the new three-parameters Gumbel distribution is a distribution known for analyzing/modelling the extreme data sets, meaning that it is skewed distribution. This makes it reasonable to support the results in Figure 4.1 which are showing that a three-parameters Gumbel distribution is right skewed distribution.

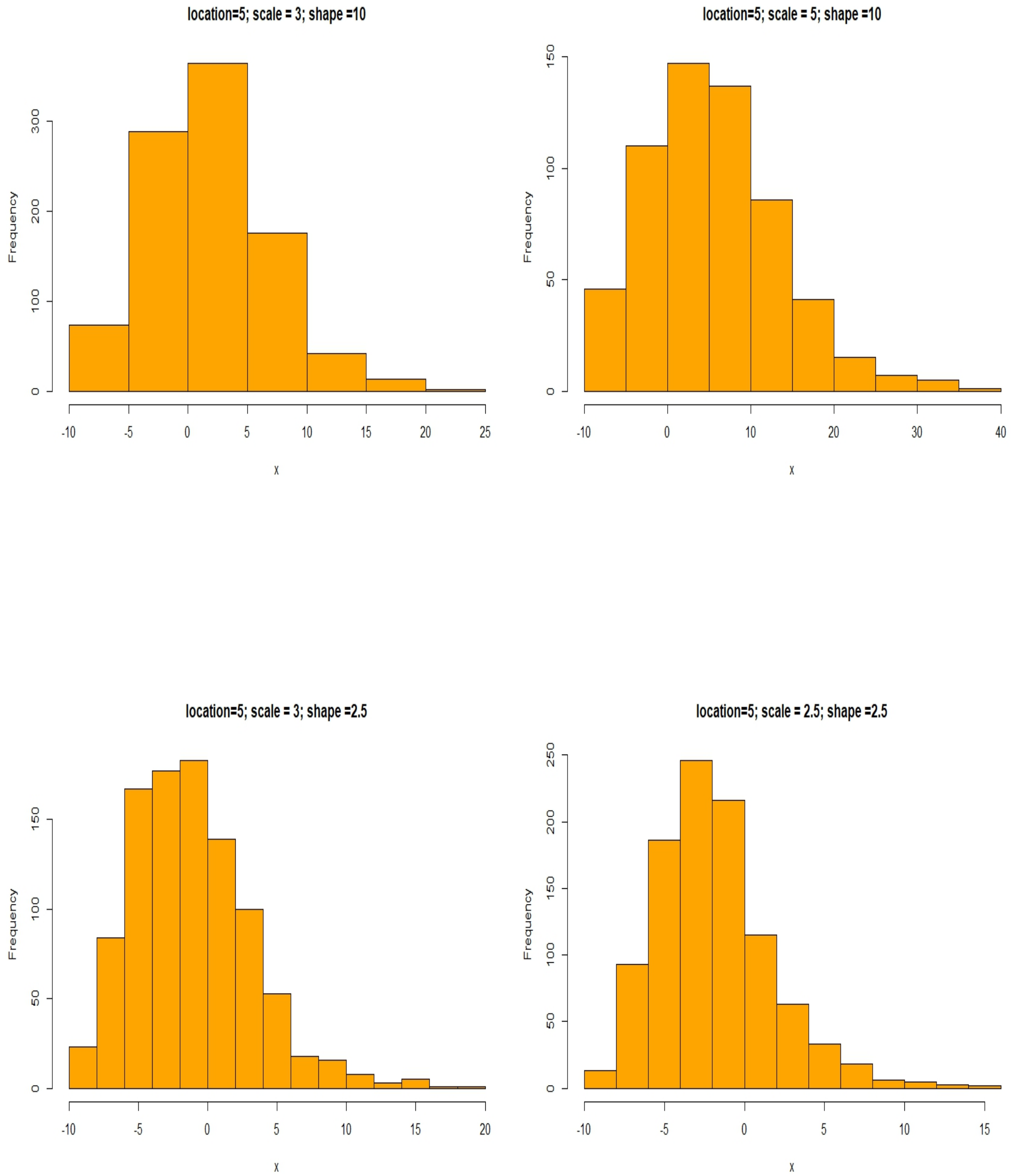

Figure 4.1: Distributions when ω and n are constant

When the scale parameter (τ) is constant and n = 10000

Figure 4.2 shows that when the scale/dispersion parameter is held constant at any given sample size and vary the values for location parameter and the shape parameter when simulating data, a three parameters Gumbel distribution still appears to be skewed to the right. A property that makes the distribution good for modelling/analyzing extreme data sets and skewed data sets.

When the shape parameter (δ) is constant and n = 10000

For simulation of 10000 samples using a three-parameters Gumbel distribution when the shape parameter is held constant, it can be observed from Figure 4.3 that manipulating the simulation values for the location parameter and the scale parameter retains the distribution as right skewed distribution. This means that the simulation results will still yield a right skewed distribution even if the sample size is increased or decreased.

When the sample size is varying

Figure 4.4 shows how a three-parameters Gumbel distribution behaves when the simulation samples are increased. When the samples are increased from 1000 to 1000000, at a given level for the location, scale and shape parameters, the distribution is skewed to the right/positively skewed. This leads to the conclusion that at any given sample size will produce samples which are positively skewed.

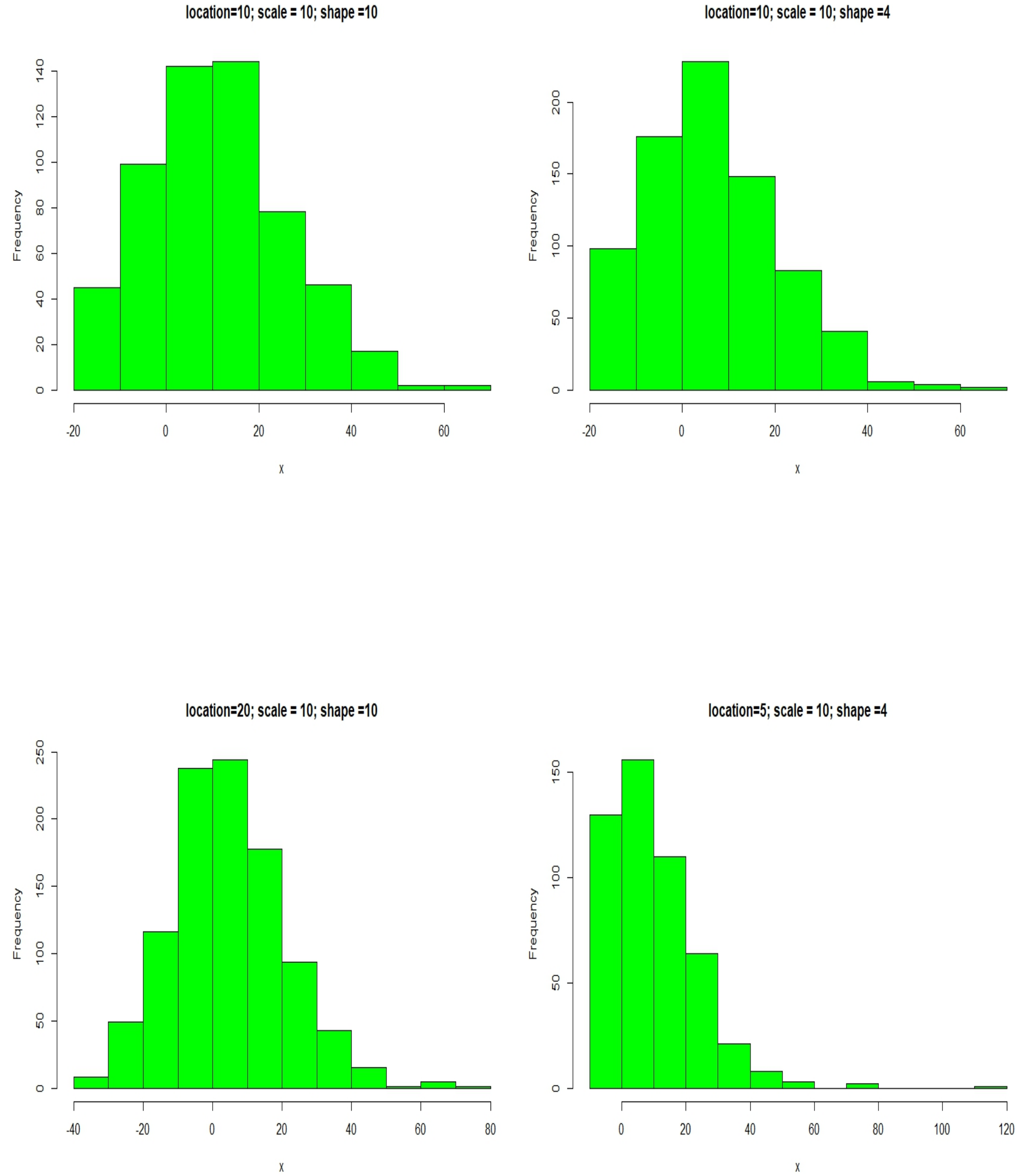

Figure 4.2: Distributions when τ and n are constant

Figures 4.1,4.2,4.3, and 4.4 demonstrates that a three-parameters Gumbel distribution is a positively skewed distribution (skewed to the right). Therefore, the distribution is applicable for analyzing positively skewed data sets and also extreme data (because extreme data are skewed data with high level of skeweness).

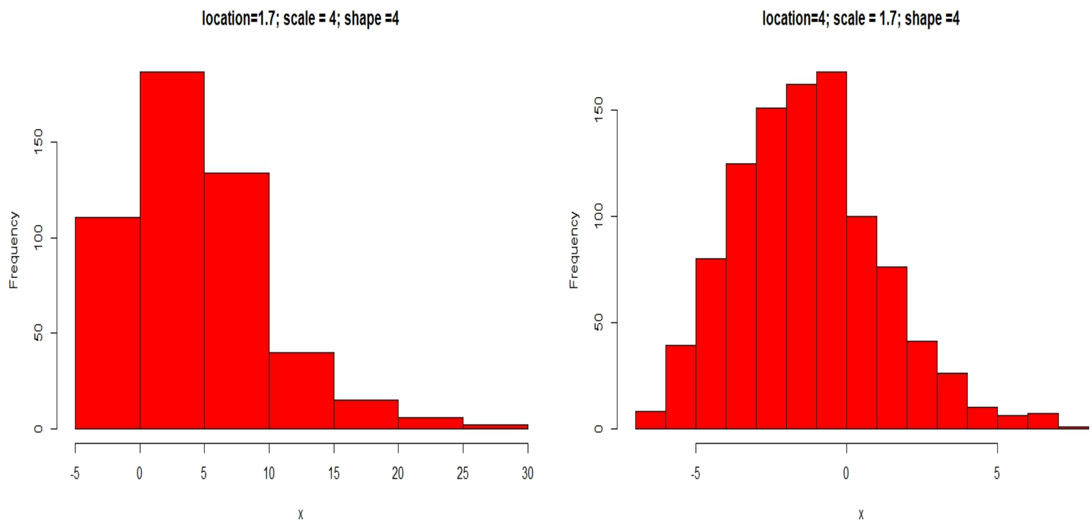

Figure 4.3: Distributions when δ are n are constant

4.3.2 Asymptotic bias

One important problem to the interpretation of quantitative data analysis and statistical modeling is the danger of bias of estimators, leading to inconsistent estimates in statistical analysis which do not tend to be close to the right answer asymptotically (as data sets approaches infinity). An asymptotically unbiased estimator is an estimator that is right on average as the sample size becomes very large (that is, as $k \to \infty$), bias converges to 0 and the fact is that all unbiased estimators are asymptotically unbiased. This was explained using asymptotic distributions as shown in Figures 4.5, 4.6 and 4.7 respectively.

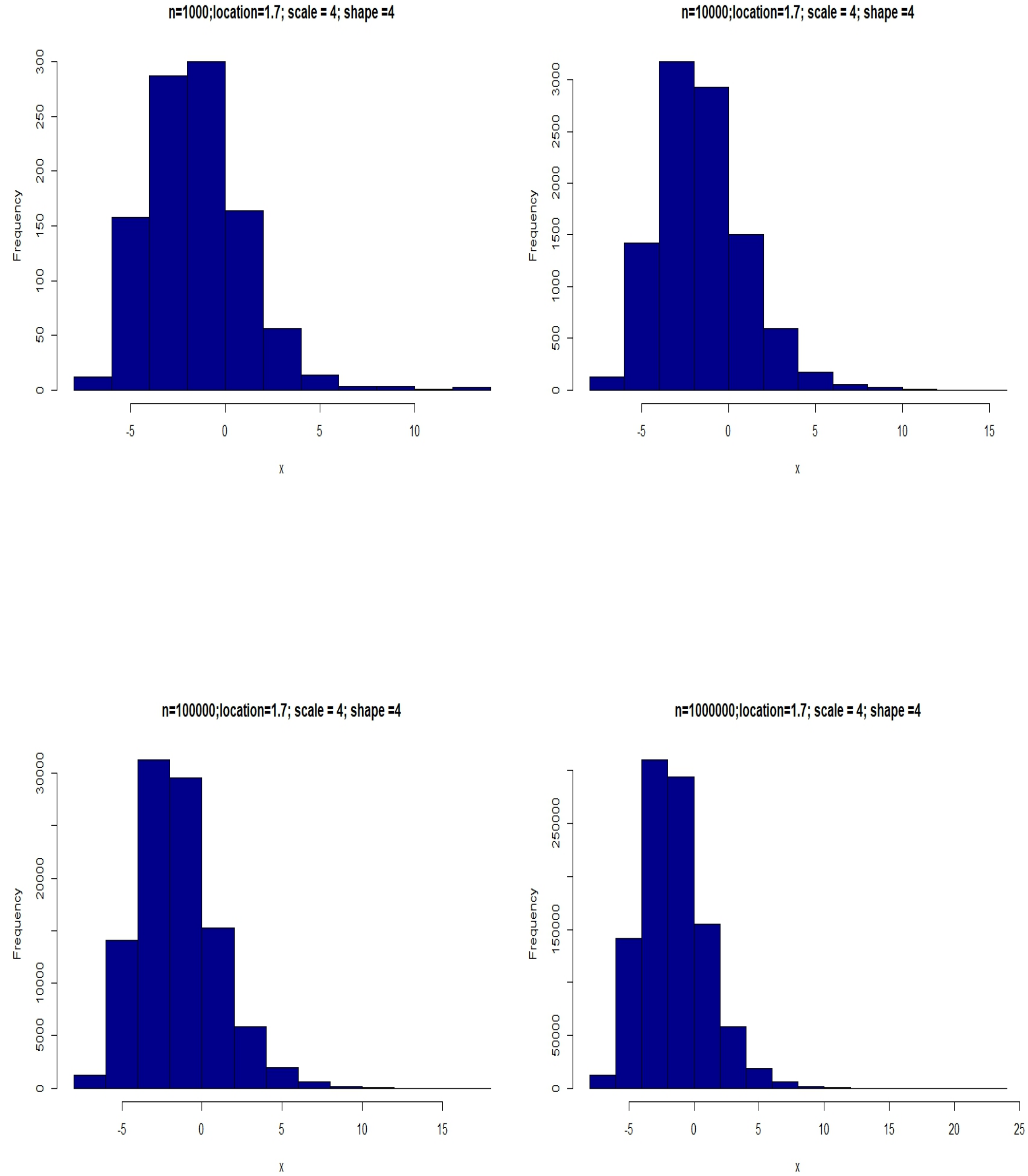

Figure 4.4: Distributions when ω, τ and δ are constant and n is varying

Asymptotic biasedness for ω

From Figure 4.5, it can be observed that as the sample size increases, the distribution for parameter ω approaches normal distribution. This means that the

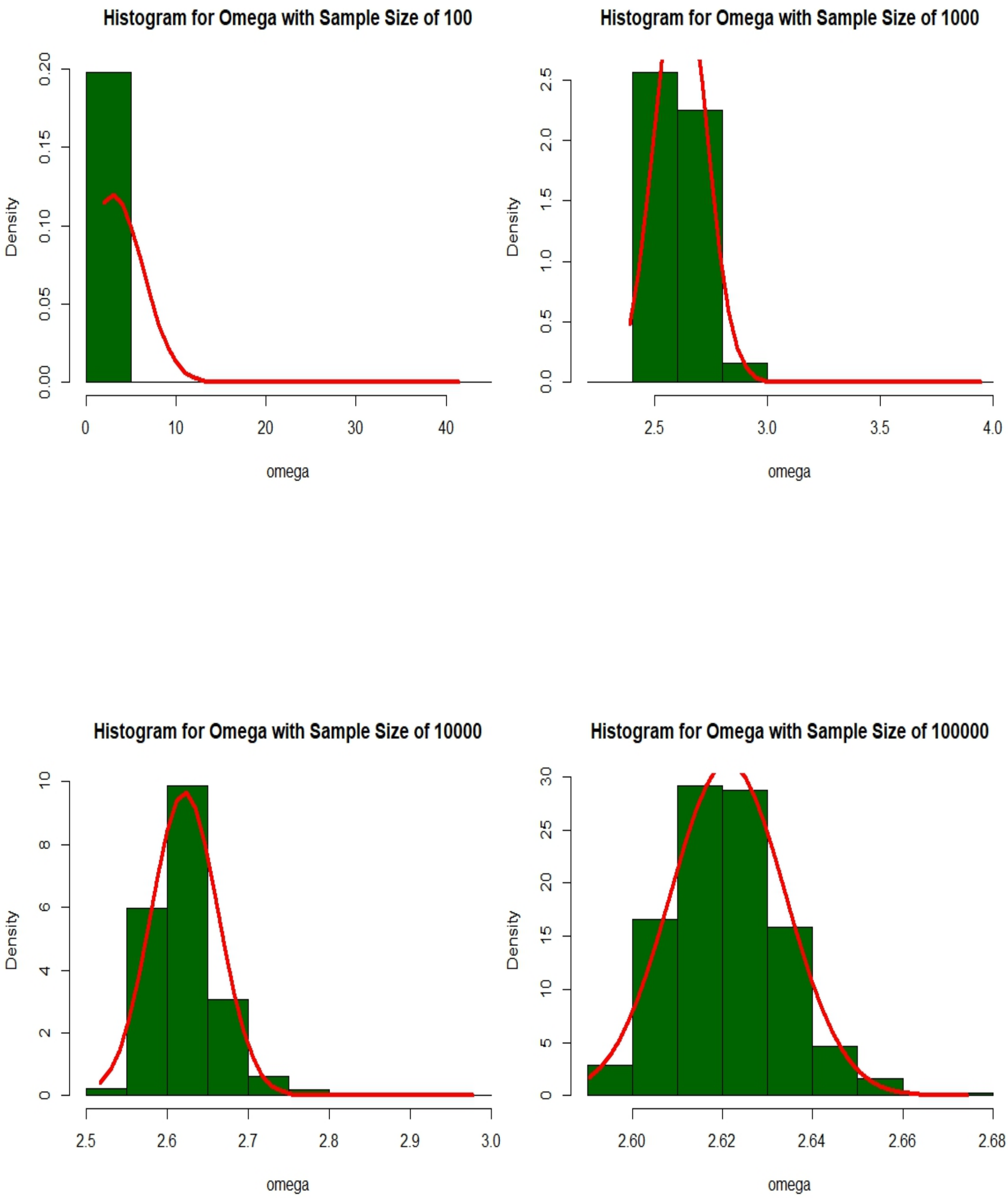

Figure 4.5: Asymptotic distributions for ω

information at the Maximum Likelihood Estimation, estimated the true value of ω (but unknown) as the sample size becomes large. This is clear since it is evidently supported by the asymptotic behaviuor of the ω as the sample size for the simulated data is increased from 100, 1000, 10000 and 100000 respectively. The normality approach as the sample size become large for the parameter indicates that bias is small enough to be definitely acceptable meaning that the estimate of the said parameter ω is suitably close to the true population value as the sample size becomes large and therefore, implying that the parameter ω is unbiased as the sample size approaches infinity, that is, $bias(\hat{\omega}, \omega) \to 0$ as $k \to \infty$ (that is, as k, the sample size becomes larger and larger).

Asymptotic biasedness for δ

Figure 4.6 below, show the distribution of parameter δ for simulated data sets with different sample sizes, that is, samples of sizes 100, 1000, 10000 and 100000 respectively. This shows the asymptotic behavour of the parameter as the sample size increases, indicating that as the sample size increases the distribution of δ approaches normal distribution. This means that the estimated value of the δ reflects the true value of parameter δ as the sample size becomes large. The normality approach as the sample size become large for the parameter indicates that bias is small enough to be tolerable meaning that the estimate of parameter δ is adequately close to the true population value as the sample size becomes large and therefore, implying that the said parameter δ is unbiased as the sample size approaches infinity,that is, $bias(\hat{\delta}, \delta) \to 0$ as $k \to \infty$.

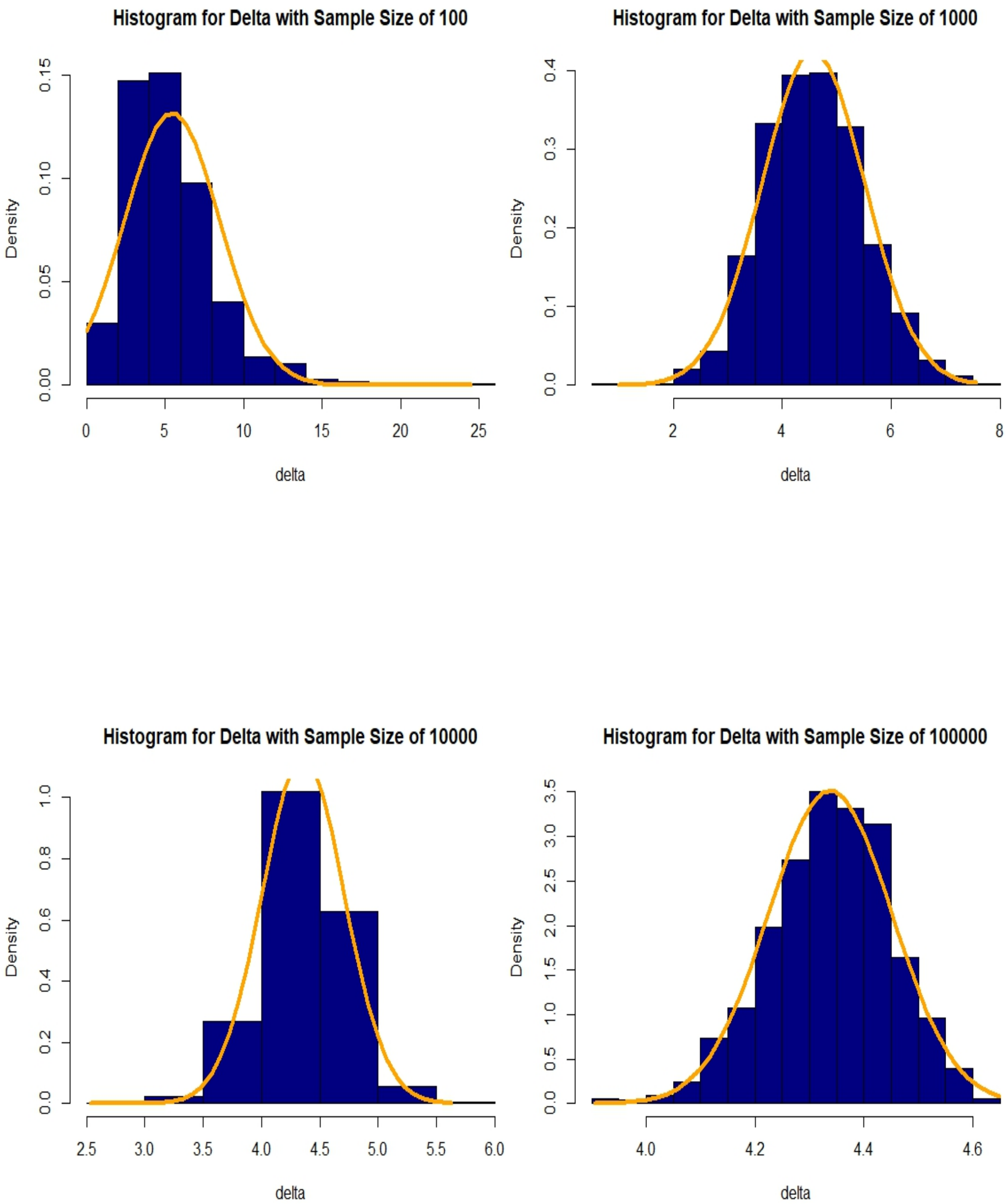

Figure 4.6: Asymptotic distributions for δ

Asymptotic biasedness for τ

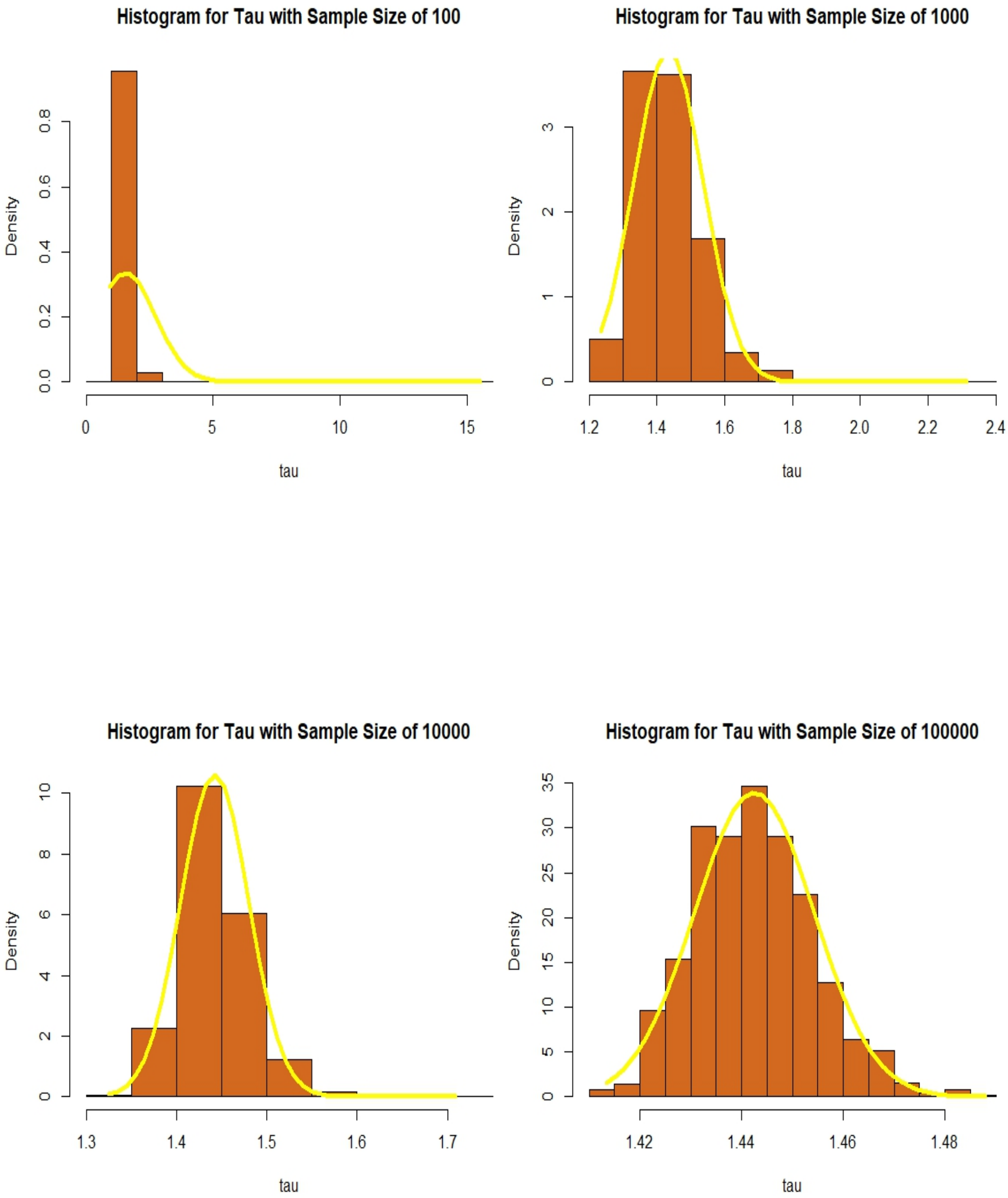

Figure 4.7: Asymptotic distributions for τ

The asymptotic behaviuor of the parameter τ was studied at different sample sizes of the simulated data which were samples of size 100, 1000, 10000 and 100000. For parameter τ, it can be observed from Figure 4.7 that for the sample sizes 100 and 1000, the distribution for the parameter τ is kind of skewed to the left.

As the sample size increases as can be observed under sample size of 10,000 and 100,000 as shows in Figure 4.7, it can be observed that, the distribution for parameter τ approaches normal distribution. The normality approach as the sample size become large for the parameter τ indicates that bias is small enough to be definitely acceptable as the same case with parameters ω and δ, meaning that the estimate of parameter τ is clearly close to the true population value of parameter τ as the sample size becomes large and therefore, implying that the parameter τ is unbiased as the sample size approaches infinity, that is $bias(\hat{\tau}, \tau) \rightarrow 0$ as $k \rightarrow \infty$.

4.3.3 Mean Square Error

This subsection discusses the Mean Square Error of the estimators as k becomes large for the purpose of understanding precision the estimators. A precise estimate is one in which the variability in the estimation error is small. This estimator is defined as the expected value of the square of the difference between the expected value and the parameter or the average square difference between the estimated value and the actual value of the parameter. The Mean Square Error is usually positive but not zero as the errors approaches zero because the estimators does not comprise of information that could lead to a completely accurate estimate. This shows that the Mean Square Error for a good estimator should tend to zero as the sample size approaches infinity. Because the estimators are unbiased, the Mean Square Error is the variance of the estimator. The results for the Mean Square Error for simulated data sets of sizes 100, 1000, 10000 and 100000 are given in Table 4.1.

From Table 4.1, it can be clearly observed that as the sample size increases for all the three parameters (ω, δ and τ), the Mean Square Errors (MSE) of their

Table 4.1: Mean and MSE of the parameters

Statistic	n = 100	n = 1000	n = 10,000	n = 100,000
$\bar{\omega}$	3.0047	2.6535	2.6211	2.6220
MSE($\hat{\omega}$)	11.6756	1.2334	0.0014	0.0002
$\bar{\delta}$	5.4234	4.5177	4.3591	4.3312
MSE($\hat{\delta}$)	8.8773	0.8509	0.1116	0.0129
$\bar{\tau}$	1.5397	1.4492	1.4419	1.4429
MSE($\hat{\tau}$)	1.4910	0.1554	0.0012	0.0001

estimators approaches zero. This is a clear indication that as the sample size becomes very large the errors approaches zero hence leading to the provision of accurate estimate values, that is, the estimate values that nearly approaches the true value of the parameter.

4.3.4 Consistency

In this subsection, we now check if the estimators are precise and reliable as sample size,k, becomes large which is proved by considering the consistency of the estimators. An estimator is said to be consistent if the precision and reliability of its estimate improve with increase in sample size as discussed in section 3.3. From the concept of law of large numbers, an estimator is termed as consistent if its bias and standard error tends to zero as the sample size tends to infinity. Subsection 4.3.2 confirmed that the estimators for parameters ω, δ and τ are unbiased as the sample sizes approaches infinity. This section therefore, presents the consistency of the estimators for normally distributed data, skewed data and extreme data as discussed below.

Consistency analysis using normally distributed data

The normally distributed data was generated using the normal concept. The data was simulated for different sample sizes of 100, 1000, 10000, 100000 and 1000000. The simulated data was used to demonstrate if the estimators are consistent as the sample size increase.

From Table 4.2, it can be seen that as the sample size increase from 100 to

Table 4.2: Estimates and standard errors for normal data

Parameter	n = 100	n = 1000	n = 10,000	n = 100,000	n = 1,000,000
$\hat{\omega}$	2.5279	2.6361	2.5997	2.6234	2.6202
$s.e(\hat{\omega})$	0.3140	0.1031	0.0329	0.0106	0.0033
$\hat{\delta}$	5.4317	4.2028	4.4488	4.3123	4.3508
$s.e(\hat{\delta})$	2.7317	0.6197	0.2121	0.0651	0.0208
$\hat{\tau}$	1.3267	1.4917	1.4313	1.4462	1.4424
$s.e(\hat{\tau})$	0.1587	0.0556	0.0166	0.0053	0.0017

1,000,000 the standard errors for estimated values of ω, δ and τ tends to zero. This provides sufficient evidence to conclude that as the sample size of the normally distributed data approaches infinity, the estimators of the parameters are consistent since their respective standard errors approach zero, that is $s.e(\hat{\omega}, \hat{\delta}, \hat{\tau}) \to 0$ as $k \to \infty$.

For the purpose of the asymptotic behaviuor of the estimates, it is important to determine the relationship between the estimates. Table 4.3 shows that there is a slight variation in the relationship between the estimates, that is, the relationship between the estimates is almost same irrespective of the sample size.

Table 4.3: Correlation between the estimators

Sample size	Estimator	ω	δ	τ
n = 100	ω	1.0000	-0.8872	0.7994
	δ	-0.8872	1.0000	-0.8952
	τ	0.7994	-0.8952	1.0000
n = 1,000,000	ω	1.0000	-0.8903	0.8218
	δ	-0.8903	1.0000	-0.8932
	τ	0.8218	-0.8932	1.0000

In Table 4.3, the correlation coefficient between ω and δ for a sample size of 100 is -0.8872 and for a sample size of 1,000,000 the correlation coefficient is -0.8903. This shows that there is a strong negative relationship between ω and δ implying that for any given sample size a decrease/increase in the estimated value for ω leads to increase/decrease in the estimated value for δ (with a stronger relationship as the sample size becomes large).

The relationship between ω and τ is 0.7994 for sample size of 100 and 0.8218 for sample size of 1,000,000. This means that there is almost a strong positive relationship between the estimated values of ω and τ, implying that increase/decrease in the estimated values of ω results to increase/decrease in the estimated values of τ for any given sample size for normally distributed data sets, but with more strength of correlation as the sample size approaches infinity.

Lastly, τ and δ provided a correlation coefficients of -0.8952 and -0.8932 for the sample size of 100 and 1,000,000 respectively. This shows that there is a strong negative relationship between τ and δ implying that for any given sample size a decrease/increase in the estimated value for τ leads to increase/decrease in the estimated value for δ.

Consistency analysis using skewed data

The skewed data was generated using the chi-square distribution which is known to be a right/positively skewed distribution. The data was simulated for different sample sizes (100, 1000, 10000, 100000 and 1000000). The simulated data was used to demonstrate if the estimators are consistent as the sample size increase for skewed data. The results provided in Table 4.4 shows that as the sample size increase from 100 to 1,000,000 the standard errors for estimated values of ω, δ and τ approaches zero. This supports the conclusion that as the sample size of skewed data sets approaches infinity, the estimators of the parameters are consistent since their respective standard errors approach zero.

Table 4.4: Estimates and standard errors for skewed data

Parameter	n = 100	n = 1000	n = 10,000	n = 100,000	n = 1,000,000
$\hat{\omega}$	4.3908	4.8235	4.9245	4.9538	4.9501
$s.e(\hat{\omega})$	1.0781	0.4320	0.1466	0.0465	0.0147
$\hat{\delta}$	0.6108	0.3989	0.3925	0.4046	0.4011
$s.e(\hat{\delta})$	0.4182	0.0953	0.0325	0.0105	0.0033
$\hat{\tau}$	2.6300	2.9614	2.9034	2.9335	2.9356
$s.e(\hat{\tau})$	0.5201	0.2114	0.0689	0.0219	0.0069

Further, it is important to investigate the relationship between the estimated values of the parameters for skewed data sets. This is presented in Table 4.5

which shows that there is a small variation in the correlation coefficient between the estimates for any change in the sample size.

Table 4.5: Correlation between the estimators

Sample size	Estimator	ω	δ	τ
n = 100	ω	1.0000	-0.9665	0.9197
	δ	-0.9665	1.0000	-0.8976
	τ	0.9197	-0.8976	1.0000
n = 1,000,000	ω	1.0000	-0.9766	0.9470
	δ	-0.9766	1.0000	-0.9165
	τ	0.9470	-0.9165	1.0000

In Table 4.5, the correlation coefficient between ω and δ for a sample size of 100 is -0.9665 and for a sample size of 1,000,000 the correlation coefficient is -0.9766 which demonstrates that as the sample size increases the relationship gets stronger. This further shows that there is a strong negative relationship between ω and δ implying a decrease/increase in the estimated value for ω leads to increase/decrease in the estimated value for δ.

Next, the relationship between ω and τ is 0.9197 and 0.9470 for sample size of 100 and 1,000,000 respectively, indicating that the relationship is somehow developing stronger as the sample size approaches infinity. The strong positive relationship between the estimated values of ω and τ, implying that increase/decrease in the estimated values of ω results to increase/decrease in the estimated values of τ for skewed data sets.

Lastly, τ and δ provided a correlation coefficients of -0.8976 and -0.9165 for the sample size of 100 and 1,000,000 respectively, meaning that the relationship is developing stronger as the sample size approaches infinity. Hence, the strong negative relationship between τ and δ implying that a decrease/increase in the estimated value for τ leads to increase/decrease in the estimated value for δ for the skewed data sets.

Consistency analysis using extreme data

One of the best distribution for modeling extreme data sets is the Weibull distribution. This leads to the simulation of extreme data sets for different sample sizes (that is, samples of size 100, 1000, 10000, 100000 and 1000000) using the Weibull distribution. The simulated data was applied to study if the estimators are consistent as the sample size for extreme data are increased. The results provided in Table 4.6 shows that as the sample size increase from 100 to 1,000,000 the standard errors for estimated values of ω, δ and τ approaches zero as the sample sizes tends to infinity. This supports the conclusion that as the sample size of extreme data sets approaches infinity, the estimators of the parameters are consistent since their respective standard errors approach zero.

Table 4.6: Estimates and standard errors for extreme data

Parameter	n = 100	n = 1000	n = 10,000	n = 100,000	n = 1,000,000
$\hat{\omega}$	0.9382	0.8004	0.7549	0.7633	0.7618
$s.e(\hat{\omega})$	0.1985	0.0614	0.0161	0.0052	0.0016
$\hat{\delta}$	10.8587	23.2746	32.6220	30.1165	30.7807
$s.e(\hat{\delta})$	9.6346	6.3402	2.5801	0.7565	0.2455
$\hat{\tau}$	0.3179	0.3156	0.3028	0.3065	0.3054
$s.e(\hat{\tau})$	0.0351	0.0098	0.0029	0.0009	0.0003

Because the estimators are consistent, it is important to comment on the relationship between the estimated values of the parameters for extreme data sets as sample size becomes large. The results for the relationship is presented in Table 4.7 which shows that there is a small variation in the correlation coefficient between the estimates for any change in the sample size.

Table 4.7: Correlation between the estimators

Sample size	Estimator	ω	δ	τ
n = 100	ω	1.0000	-0.9442	0.5758
	δ	-0.9442	1.0000	-0.7594
	τ	0.5758	-0.7594	1.0000
n = 1,000,000	ω	1.0000	-0.8997	0.4777
	δ	-0.8997	1.0000	-0.7614
	τ	0.4777	-0.7614	1.0000

In Table 4.7, the correlation coefficient between ω and δ for a sample size of 100 is -0.9442 and for a sample size of 1,000,000 the correlation coefficient is -0.8797 which demonstrates that as the sample size increases the relationship gets slightly weaker. This further shows that there is a negative relationship between ω and δ implying a decrease/increase in the estimated value for ω leads to increase/decrease in the estimated value for δ.

Secondly, the relationship between ω and τ is 0.5758 and 0.4777 for sample size of 100 and 1,000,000 respectively, indicating that the relationship is becoming weaker as the sample size approaches infinity. The positive relationship between the estimated values of ω and τ, implying that increase/decrease in the estimated values of ω results to increase/decrease in the estimated values of τ.

Lastly, τ and δ provided a correlation coefficients of -0.7594 and -0.7614 for the sample size of 100 and 1,000,000 respectively, meaning that the relationship is becoming slightly stronger as the sample size approaches infinity. Hence, the moderate negative relationship between τ and δ implying that a decrease/increase in the estimated value for τ leads to increase/decrease in the estimated value for δ for the extreme data sets.

4.3.5 Asymptotic relationship between three parameters Gumbel distribution and other distributions

This subsection discusses the asymptotic relationship between a three parameter Gumbel distribution and other common distributions like the normal distribution, chi square distribution and Weibull distribution. The three distributions namely normal, chi-square and Weibull are well known for the purpose of modeling of data which is either normally distributed data or right(positively) skewed data or extreme data respectively. The asymptotic relationship between three parameters Gumbel distribution and the three distributions (that is, normal, chi-square and Weibull) is discussed on below.

Relationship between a three parameters Gumbel distribution and normal distribution

Asymptotic normality states that as the sample size becomes large, the distribution tends to be normal. This is also supported by the Central Limit Theorem (CLT) which clarifies that the behaviour of the random variable V as the sample size approaches infinity, is expected to shift the variance and the mean of the variable making it tend to be normally distributed.

Figure 4.8 below shows the behaviour of the three parameters Gumbel distribution when the sample size for normally distributed data is increased. The figure illustrates that as sample size for normal data increases, the three parameters Gumbel distribution is kind of similar to the normal distribution. The displayed graphs shows that for a small sample size(n=1000) there is a big variation on the values of a three parameters Gumbel distribution and the normal distribution since the location, dispersion and shape parameters of the graphs indicates no similarity. As the sample size becomes large (n = 1,000,000), the three parameter Gumbel distribution approaches the normal distribution and it can be observed that the spread and shape parameters are almost same with a small difference on the location parameter. Therefore, the three parameters Gumbel distribution is recommended for analyzing and/or modeling the large sample sized normal data sets.

Further, it can be observed from the figure that as the sample size becomes large, the mean of the three parameters Gumbel distribution becomes closer to the mean of the normal distribution as required by the Central Limit Theorem.

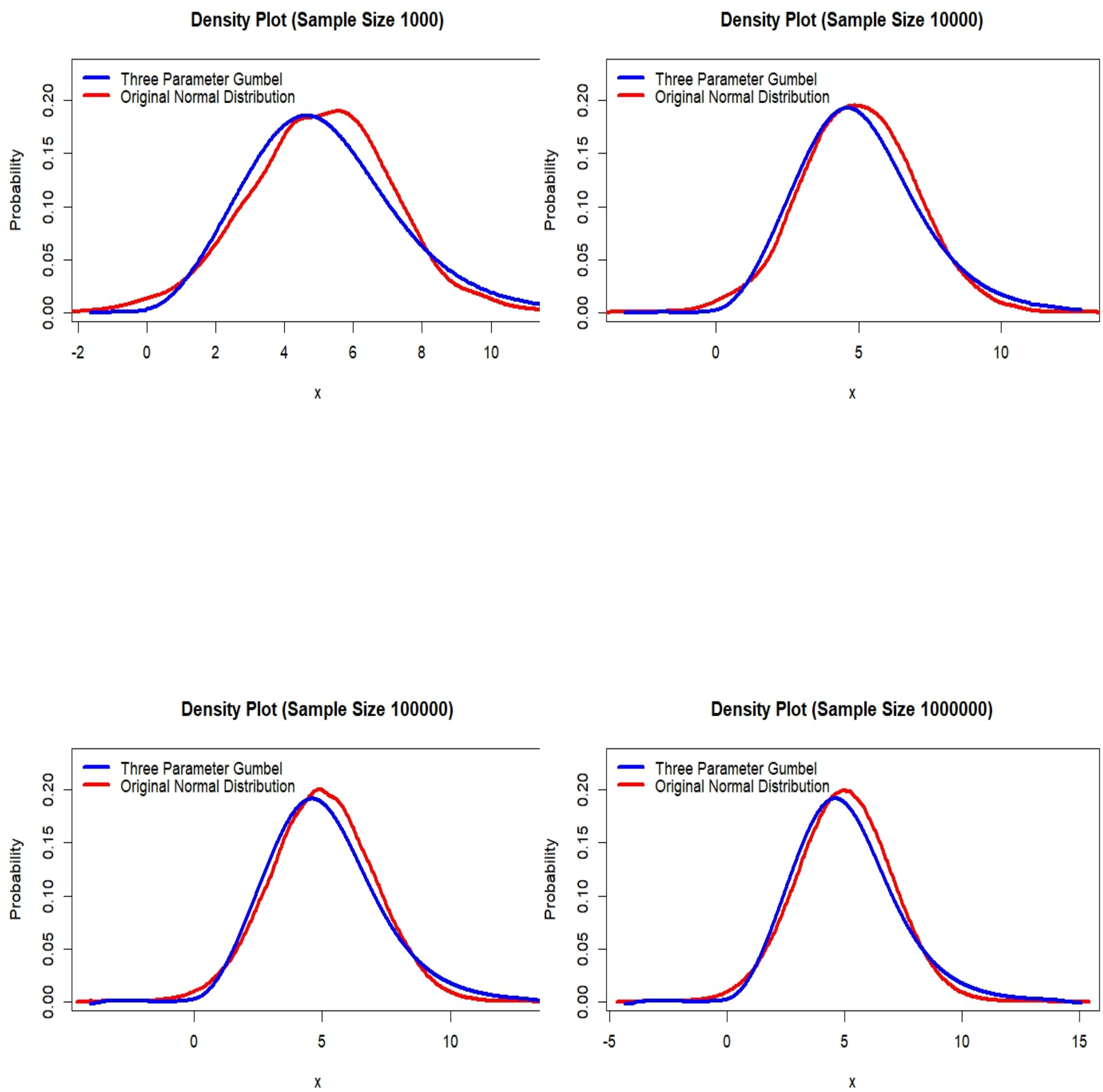

Figure 4.8: Asymptotic distributions for three parameters Gumbel and normal distributions

Relationship between a three parameters Gumbel distribution and Chi-square distribution

Chi-square is a right skewed distribution and therefore, we are investigating if a three parameters Gumbel distribution can also be used to fit skewed data sets by studying its relationship with skewed distribution (Chi-square distribution), for samples of size 100, 10000, 100000, and 1000000.

Figure 4.9 shows the behaviour of the three parameters Gumbel distribution and chi-square distribution when the sample size for skewed data is increased.The displayed graphs shows that for a small sample size(n=100) there no similarity in terms of the shape, spread and the mean of a three parameters Gumbel distribution and the chi-square distribution. As the sample size becomes large (n = 1,000,000) for the skewed data sets, a three parameter Gumbel distribution is almost same as the chi-square distribution. This is supported by the clear observation showing similarity on the spread and shape of the distributions with a small difference on the location/mean of the distributions. This implies that as the sample approaches infinity, three parameters Gumbel distribution becomes similar to chi-square distribution, hence, a three parameters Gumbel distribution is also useful in modeling and/or analyzing the large sample sized skewed data sets.

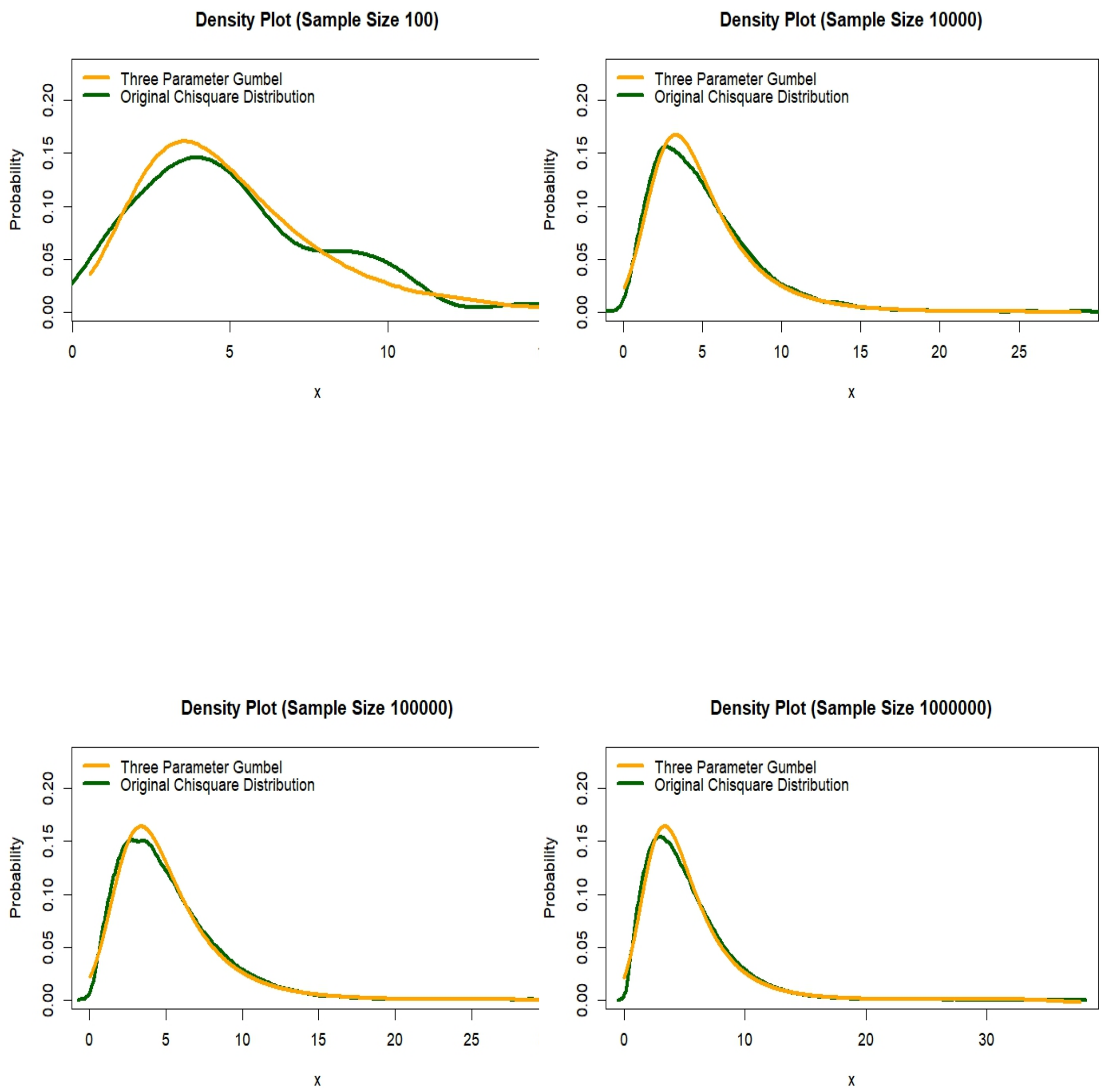

Figure 4.9: Asymptotic distributions for three parameters Gumbel and chi-square distributions

Relationship between a three parameters Gumbel distribution and Weibull distribution

Weibull distribution is known for analyzing and/or modeling the extreme events. We are therefore interested on investigating if a three parameters Gumbel distri-

bution can also be used to analyze/model the extreme data sets. This was by studying the relationship between a three parameters Gumbel distribution with Weibull distribution (for samples of size 100, 10000, 100000, and 1000000).
Figure 4.10 below, shows the behaviour of a three parameters Gumbel distribution and Weibull distribution when the sample size for extreme data is increased.The graphical results shows that for a small sample size(n=100) there no similarity in terms of the shape, spread and the mean of a three parameters Gumbel distribution and the Weibull distribution. As the sample size becomes large (n = 1,000,000) for the extreme data sets, the three parameter Gumbel distribution is almost same as the Weibull distribution. This is supported by the clear observation showing similarity on the spread and shape of the distributions with a small difference on the location of the distributions. This implies that as the sample approaches infinity, a three parameters Gumbel distribution becomes similar to Weibull distribution, hence, three parameters Gumbel distribution is also useful in modeling and/or analyzing the large sample sized extreme data sets.

In general, this chapter therefore, concludes from the asymptotic properties discussed that a three parameters Gumbel distribution provides unbiased, efficient and consistent estimates as the sample size gets large (approaches infinity). The chapter further demonstrates that a three parameters Gumbel distribution is flexible for fitting normal data, skewed data and extreme data sets with provision of unbiased and efficient estimators(an advantage that majority of the distributions cannot achieve).

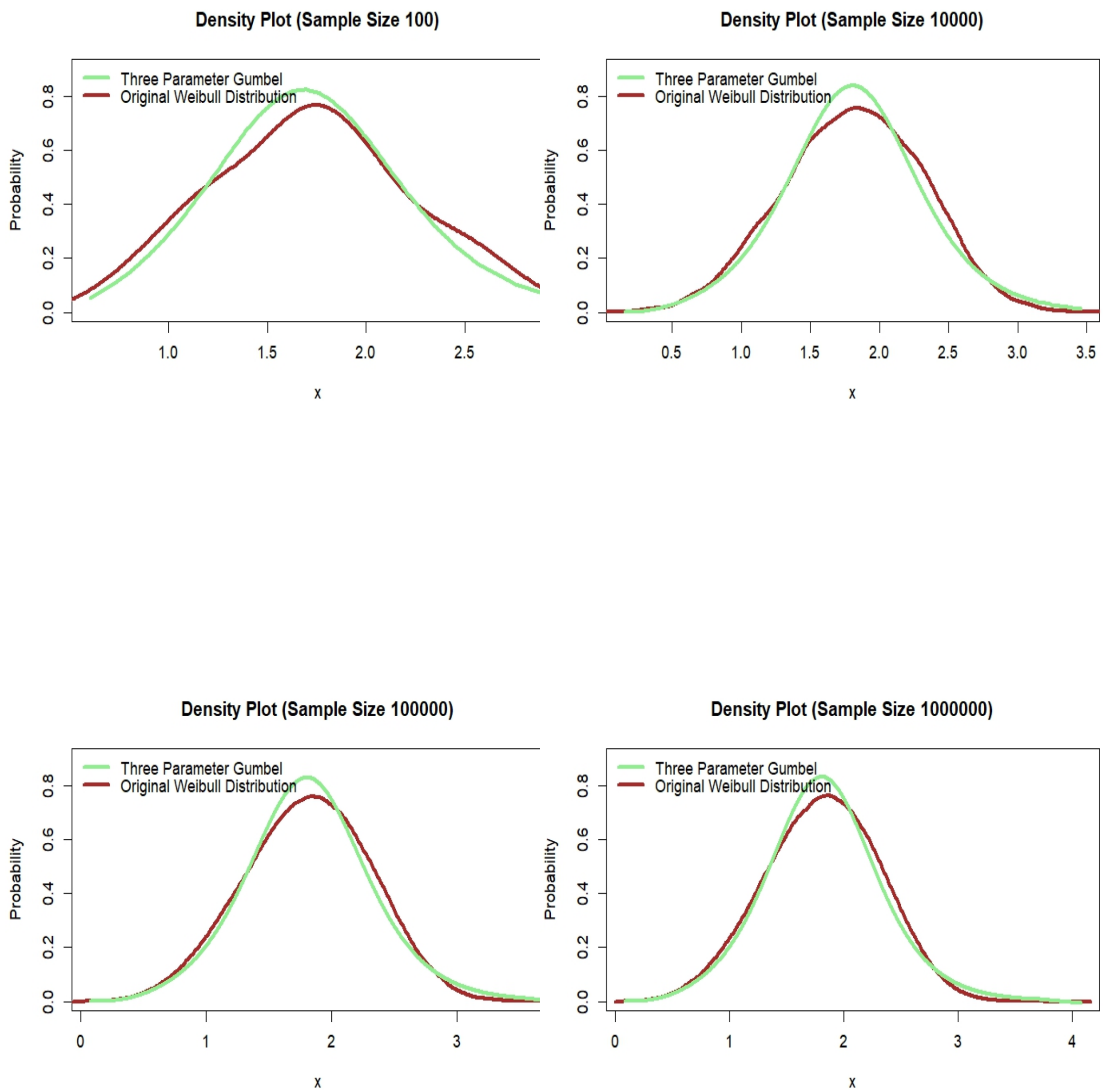

Figure 4.10: Asymptotic distributions for three parameters Gumbel and Weibull distributions

4.4 Application and Efficiency

This section investigates the efficiency of a three parameters Gumbel distribution using Mean Square Error (MSE) and/or Akaike's Information Criterion (AIC) by applying simulated and real data . In subsection (4.4.1), we investigate the efficiency using three different types of simulated data sets and in subsection (4.4.2) we examined the efficiency by applying real life data sets to compare the three parameters Gumbel distribution with other existing distributions like the two parameters Gumbel distribution, exponentiated Gumbel distribution, Gumbel geometric distribution, two parameters Weibull distribution and three parameters Frechet distribution.

4.4.1 Efficiency using simulated data

In this subsection, we ascertain the type of data in which a three parameter Gumbel distribution is more efficient in terms of application. Three different types of simulated data sets, that is, normal data which was simulated using the normal distribution, extreme data simulated using Weibull distribution and skewed data simulated using chi-square distribution are applied in this section to investigate the efficiency using Akaike's Information Criterion. These three types of data are simulated for different sample sizes (that is, for sample sizes of 100, 1000, 10000, 100000 and 1000000). The data sets were fitted using three parameters Gumbel distribution and the AIC values obtained for the different types of data at different sample sizes are as given in Table 4.8.

Table 4.8: AIC values for simulated data

Sample size	Normal	Skewed	Extreme
n = 100	434.5441	490.9186	153.5914
n = 1000	4408.551	4870.835	1549.833
n = 10,000	43180.57	48350.63	14917.24
n = 100,000	433335.1	486704.6	151138.3
n = 1,000,000	4331444	4863754	1506155

From subsection (4.3.5), it was concluded that as the sample size becomes large, the formulated three parameters Gumbel distribution approaches the normal distribution, chi-square distribution and Weibull distribution for normal, skewed and extreme data respectively. This further showed that the formulated three parameters Gumbel distribution is flexible for fitting the normal, skewed and extreme data sets. From Table 4.8, we can observe that the extreme value data sets yielded smaller AIC value, followed by the normal data sets and lastly, the skewed data sets. This indicates that the three parameters Gumbel distribution is superior for fitting extreme data, normal data and lastly skewed data. Making it more superior for modeling extreme data than other sets of data, hence it is in line with the improved mother distribution which was for analyzing the extreme data sets. This section therefore, concludes that as much as three parameters Gumbel distribution fits extreme, skewed and normal data, it is best for fitting extreme data, followed by normal data and lastly skewed data.

4.4.2 Efficiency Analysis using real data

In this subsection, three different types real life data sets namely Earthquake magnitude (data 1), average Avocado prices (data 2) and Income expectation (data 3) were used to investigate the efficiency of a three parameter Gumbel distribution in comparison to other existing distributions (the two parameters Gumbel distribution, exponentiated Gumbel distribution, Gumbel geometric distribution, two parameters Weibull distribution and three parameters Frechet distribution). The data was downloaded from **www.kaggle.com**.

Data 1: Earthquake magnitude

The earthquake magnitude data measured on Moment Magnitude Scale (MMS) was collected from 1995 - 2023 and comprise of 1000 observations.

Table 4.9: Summary statistics for Earthquake magnitude

Statistic	Value
Mean	6.9402
Standard deviation	0.4381
Median	6.8
Mode	6.5
Minimum	6.5
Maximum	9.1

From Table 4.9, the minimum magnitude value is 6.5 with a maximum value of 9.1. The summary statistics reports that the mean of the 1000 observations is

6.9402 with a standard deviation of 0.4381. This data was further fitted using six different distribution including the new three parameter Gumbel distribution for the purpose of determining the most efficient distribution for application in data modeling and parameter estimation. The AIC results is applied to determine the best distribution for fitting the earthquake magnitude data and the results are as shown in Table 4.10.

Table 4.10: AIC values for Earthquake magnitude

Distribution	AIC value
Exponentiated Gumbel	3115.843
Three parameter Gumbel	794.5079
Two parameter Gumbel	832.3538
Gumbel geometric	974.5083
Three parameter Frechet	1165.169
Two parameter Weibull	1589.9

The results provided in Table 4.10 shows that three parameter Gumbel distribution is best distribution for fitting the earthquake magnitude data because it is having the smallest AIC value of 794.5079. The second best distribution is the Gumbel geometric distribution with an AIC value of 794.5083, followed by the two parameters Gumbel distribution which gave the AIC value of 832.3538. Three parameter Frechet distribution, two parameter Weibull distribution and Exponentiated Gumbel distribution gave the largest AIC statistic of 1165.169, 1589.9 and 3115.843 respectively, implying that they are not best for fitting the earthquake magnitude data.

For the purpose of uniformity, we investigate the efficiency using the three parameter distributions (that is, three parameters Gumbel, Exponentiated Gumbel, Gumbel geometric and three parameter Frechet) using MSE results. This is because variability determines efficiency in that an estimator is said to more efficient than another estimator, that is if it is more precise and reliable for the same sample size. Also, based on Cramer-Rao Inequality, when there are several unbiased estimators of the same parameter, then the one with least variance is the more efficient.

Table 4.11: Estimates and MSE for earthquake magnitude

Distribution	Parameter	Estimate	MSE
Exponentiated Gumbel	Location	7.5798	0.0562
	Scale	0.8604	0.0291
	Shape	5.5640	0.5076
Three parameter Gumbel	Location	6.9820	0.0481
	Scale	0.3783	0.0219
	Shape	0.2841	0.0617
Two parameter Gumbel	Location	6.7522	0.0497
	Scale	0.2933	0.0278
Gumbel geometric	Location	6.9824	0.0482
	Scale	0.3783	0.0219
	Shape	0.7165	0.0619
Three parameter Frechet	Location	7.0215	1.1367
	Scale	7.8344	0.1233
	Shape	6.9246	0.3530
Two parameter Weibull	Scale	7.1613	0.0483
	Shape	13.1882	0.2720

The three parameters (location, scale and shape) are estimated using four different distributions. By applying the Cramer-Rao approach on the location parameter, it can be observed from Table 4.11 that three parameter Gumbel give the least MSE value of 0.0481, followed by Gumbel geometric which yield a MSE value of 0.0482, then exponentiated Gumbel with MSE of 0.0562 and lastly, the three parameters Frechet with MSE value of 1.1367. This demonstrates that the three parameter Gumbel gives the most efficient estimate for the location parameter. This is further supported by the closeness of the estimated value for the location parameter (6.9820) using three parameter Gumbel distribution to the mean value of the data (6.9402) obtained from Table 4.9.

For the scale parameter, both three parameter Gumbel distribution and Gumbel geometric distribution provided the MSE value of 0.0219, followed by exponentiated Gumbel with MSE of 0.0291 and lastly, three parameters Frechet with MSE of 0.1233. This indicates that both three parameter Gumbel and Gumbel geometric provide efficient estimate for scale parameter. But the three parameter Gumbel distribution is more flexible than the Gumbel geometric due to the fact that three parameters Gumbel had smaller AIC value than Gumbel geometric distribution as evidence in Table 4.10.

Lastly, the shape parameter results from Table 4.11 indicates that three parameter Gumbel distribution gives the least MSE value of 0.0617, followed by the Gumbel geometric distribution with MSE value of 0.0619, then three parameter Frechet with MSE static of 0.3530 and the exponentiated Gumbel with the highest MSE statistic of 0.5076. This results together with the AIC results from Table 4.10 demonstrates that three parameters Gumbel distribution is the best and provide efficient estimate for the shape parameter.

The data was also fitted by graphical analysis as shown in Figure 4.11. The probability distribution functions for the six distribution (a three parameter Gumbel, exponentiated Gumbel, Gumbel distribution, Gumbel geometric, three parameter Frechet and two parameter Weibull) was fitted to the earthquake magnitude data. A probability density function is defined as a mathematical function that demonstrates a continuous probability distribution. It indicates the probability density of each value of a variable, which can be greater than one but cannot be non negative. In a graph form, a probability density function is always a curve, which may give a value greater than one for some values of v_i, because it is not the value of $f(v)$ but the area under the curve that represent probability. The total area under the whole curve is always exactly one for all probability distributions because it is certain that an observation i will fall somewhere in the variable's interval/range (https://www.scribbr.com/statistics/probability-distribution).

The Figure 4.11 showed that three parameter Gumbel distribution is the best for fitting the earthquake magnitude since it is more closer to the original distribution of the data compared to other distribution. The second best distribution is the two parameter Gumbel distribution as observed from the figure. On the side, the two parameters Weibull and Gumbel geometric distributions are not the best among the listed distribution for fitting this data.

Gumbel geometric provided the second least AIC value for the earthquake mag-

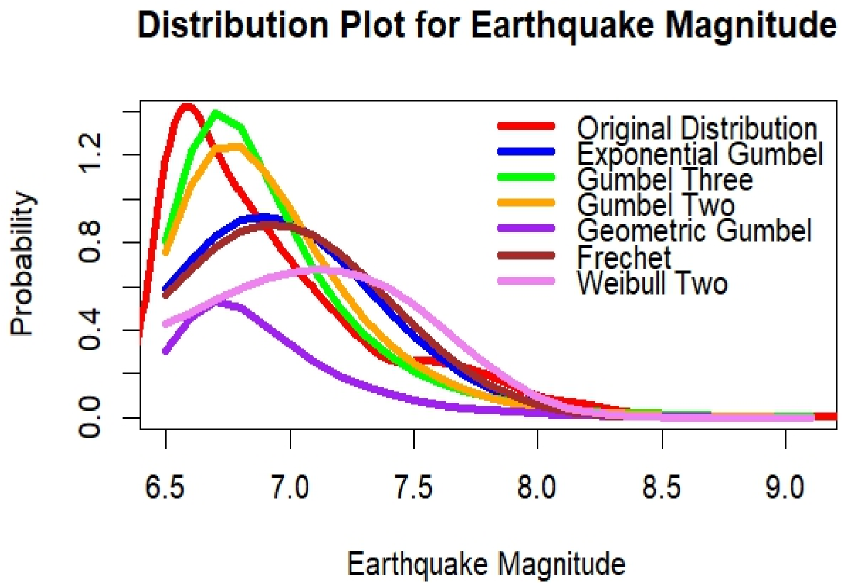

Figure 4.11: Data 1:Earthquake magnitude

nitude data with a slight difference from the AIC value for the three parameter Gumbel distribution but graphically it displayed the worst graph for the data. This could be due to the difference on the efficiency of the shape parameter estimated by Gumbel geometric distribution.

The application of data 1: earthquake magnitude proves that three parameters Gumbel distribution is the best for fitting this data. This is because the three parameters Gumbel distribution provides the smallest AIC value with more efficient estimates for the three parameters namely location, scale and shape and also the graphical analysis proves that three parameter Gumbel is the best distribution for fitting this data.

Data 2: Avocado prices (in dollars)

The average avocado price data is having 53,415 observation which was collect in USA from 2015 to 2023. The summary statistics given in Table 4.12 shows that the data have a mean price of 1.4289 dollars with a standard deviation of 0.3931 dollars. The minimum average reported price of avocado in USA is 0.44 dollars and maximum price of 3.44 dollars.

Table 4.12: Summary staistics for Avocado prices (in dollars)

Statistic	value
Mean	1.4289
Standard deviation	0.3931
Median	1.4
Mode	1.26
Minimum value	0.44
Maximum value	3.44

To comment on the best distribution for fitting this data, AIC statistic was applied to give us the answer to the best distribution and the results are as given in Table 4.13.

Table 4.13: AIC values for Avocado prices

Distribution	AIC value
Exponentiated Gumbel	158199.7
Three parameter Gumbel	48974.15
Two parameter Gumbel	49782.18
Gumbel geometric	48974.16
Three parameter Frechet	49959.13
Two parameter Weibull	53200.5

The results provided in Table 4.13 shows that three parameter Gumbel distribution is best distribution for fitting the data sets2 because it is having the smallest AIC value of 48974.15. The second best distribution is the Gumbel geometric distribution with an AIC value of 48974.16 which is slightly small than the three parameters Gumbel distribution AIC value. The third best is the three parameters Frechet distribution which have the AIC value of 49959.13, followed by two parameter Gumbel distribution with AIC statistic value of 49782.18. The two parameter Weibull distribution and Exponentiated Gumbel distribution gave the largest AIC statistic values of 53200.5 and 158199.7 respectively, indicating that they are not best for fitting the data (avocado price data).

The study further investigate the efficiency between the three parameter distributions using MSE and the Cramer-Rao Inequality approach which states that when there are several unbiased estimators of the same parameter, then the one with least variance is the more efficient.

Table 4.14: Estimates and MSE for Avocado prices

Distribution	Parameter	Estimate	MSE
Exponentiated Gumbel	Location	2.3459	0.0090
	Scale	0.9472	0.0040
	Shape	9.4421	0.1404
Three parameter Gumbel	Location	1.0625	0.0058
	Scale	0.2771	0.0017
	Shape	2.7583	0.0949
Two parameter Gumbel	Location	1.2417	0.0075
	Scale	0.3356	0.0021
Gumbel geometric	Location	1.0626	0.0058
	Scale	0.2770	0.0017
	Shape	-1.7557	0.0946
Three parameter Frechet	Location	1.8309	0.0681
	Scale	1.4169	0.0168
	Shape	2.3052	0.0214
Two parameter Weibull	Scale	1.5774	0.0018
	Shape	3.8512	0.1123

By applying the Cramer-Rao approach on the location parameter, it can be observed from Table 4.14 that three parameter Gumbel distribution and the Gumbel geometric distribution all give the least MSE value of 0.0058, followed by exponentiated Gumbel with MSE of 0.0090 and finally,the three parameters Frechet with MSE value of 0.0681. This demonstrates that the three parameter Gumbel distribution gives the most efficient estimate for the location parameter.

Secondly, for the scale parameter, both three parameter Gumbel distribution and Gumbel geometric distribution provided the MSE value of 0.0017, followed by exponentiated Gumbel with MSE of 0.0040 and lastly, three parameters Frechet with MSE of 0.0168. This indicates that both three parameter Gumbel and Gumbel geometric provide efficient estimate for scale parameter.

Lastly, the shape parameter results from Table 4.14 indicates that the three parameter Frechet distribution have the smaller MSE value of 0.0214, followed by the Gumbel geometric distribution with MSE value of 0.0946 but with a negative estimate value of -1.7557. The three parameter Gumbel distribution is the third best with MSE static value of 0.0949 and lastly, the exponentiated Gumbel with the highest MSE statistic value of 0.1404.

Therefore, the three parameter Gumbel distribution is more flexible than the Gumbel geometric distribution and three parameters Frechet distribution due to the fact that three parameters Gumbel had smaller AIC value than Gumbel geometric distribution and Frechet distribution as evidence in Table 4.13.

To support the conclusion that the three parameters Gumbel distribution is flexible for fitting this data, graphical analysis was conducted in support and the result displayed in Figure 4.12

From Figure 4.12, it can be clearly observed that three parameter Gumbel distribution, Frechet distribution and the two parameter Gumbel distribution can fit the data well. And because of the AIC technique, three parameter Gumbel distribution is termed the best since it had the smallest AIC value.

Also, it can be observed from the graph that the worst distribution for fitting this data is the Gumbel Geometric distribution. This goes against its AIC value which emerged to be the second smallest value (from Table 4.13). From the analysis, the failure on this distribution is due to its shape parameter which seems to be poorly estimated. For example in the case of this data, the Gumbel geometric distribution gave a negative estimate to the shape parameter yet the fitted data are prices which cannot be negative in nature. This disadvantage makes the three parameter Gumbel distribution more superior and flexible for fitting this data.

The application of data 2: avocado prices demonstrates that three parameters Gumbel distribution is the best for fitting this data. This is because the three parameters Gumbel distribution provides the smallest AIC value with more ef-

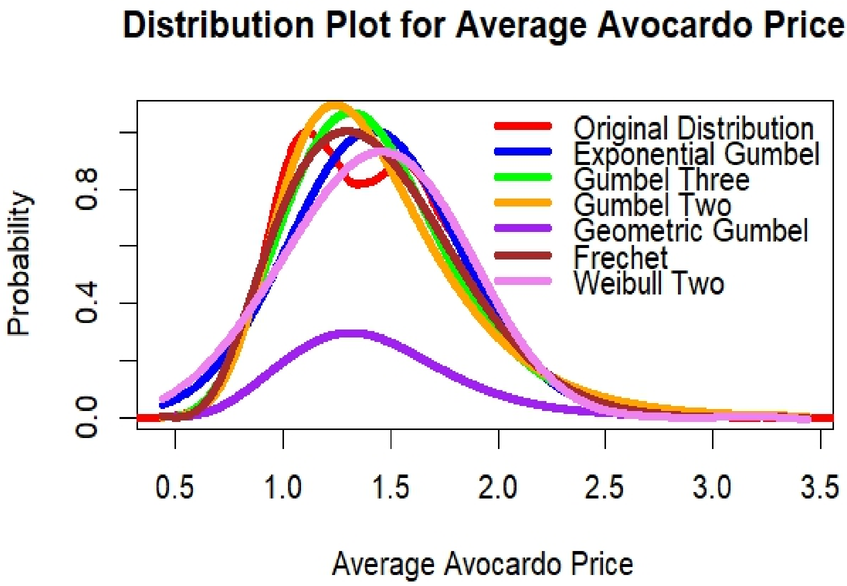

Figure 4.12: Data 2: Avocado prices in dollars

ficient estimates for the three parameters and also the graphical analysis proves that three parameter Gumbel is among the best distributions for fitting this data.

Data 3: Income prediction

The income prediction data was collected from 299,285 USA citizens. The minimum expected income from the USA citizens was 37.87 dollars with a maximum expected income was 18656.30 dollars per week. The mean value for the data is 1740.10 dollars with standard deviation of 994.1443 dollars. For the purpose of normality, the data was transformed by taking logarithm. The summary statistics after transformation showed a mean of 3.1644 dollars with a standard deviation of

0.2769. The data was fitted using five distributions namely exponentiated Gumbel distribution, three parameters Gumbel distribution, two parameters Gumbel distribution, Gumbel geometric distribution and two parameter Weibull distribution. The AIC values for the distributions are given in Table 4.15.

Table 4.15: AIC values for Income prediction

Distribution	AIC value
Exponentiated Gumbel	1219008
Three parameter Gumbel	569797.1
Two parameter Gumbel	701777.8
Gumbel geometric	569513.5
Two parameter Weibull	552879.8

From Table 4.15, the two parameters Weibull is the best based on AIC result since it have the smallest AIC value of 552879.8. The second best distribution is Gumbel geometric distribution with AIC value of 56513.5 and the third best distribution is the three parameters Gumbel distribution with AIC value of 569797.1. The results show that exponentiated Gumbel distribution is the worst for fitting this data with AIC value of 1219008, followed by the two parameters Gumbel with the AIC statistic value of 701777.8.

A flexible distribution is one which is superior for a given data (one with smaller AIC values) and with efficient estimators. For the efficiency of the estimators, the results are given in Table 4.16.

The results for the location parameter from Table 4.16, shows that two parameter Gumbel provided the smallest MSE value of 0.0015 meaning its more efficient for the location parameter, followed by exponentiated Gumbel with MSE value of 0.0046. The third distribution which provided smaller MSE for the location parameter is three parameters Gumbel distribution with a value of 0.0096 and lastly, Gumbel geometric with MSE value of 0.0117.

The efficiency of the scale/dispersion parameter as observed from Table 4.16 shows that three parameters Gumbel distribution and Gumbel geometric distribution provide more efficient scale parameter estimate with both having MSE value of 0.0005. The two parameter Weibull and exponentiated Gumbel distributions in-

Table 4.16: Estimates and MSE for income prediction

Distribution	Parameter	Estimate	MSE
Exponentiated Gumbel	Location	8.1134	0.0046
	Scale	1.1985	0.0018
	Shape	5.0839	0.0297
Three parameter Gumbel	Location	4.6480	0.0096
	Scale	0.3501	0.0005
	Shape	2134.6490	59.1504
Gumbel geometric	Location	4.4749	0.0117
	Scale	0.3506	0.0005
	Shape	-3482.2666	114.6687
Two parameter Gumbel	Location	6.9465	0.0015
	Scale	0.7595	0.0009
Two parameter Weibull	Scale	7.5612	0.0010
	Shape	13.8270	0.0193

dicate least efficient estimate for scale parameter with two two distributions displaying MSE value of 0.0010 and 0.0018 respectively.

Lastly from Table 4.16, for the shape parameter, two parameters Weibull distribution and exponentiated Gumbel distribution provides least MSE value of 0.0193 and 0.0297 respectively. The given estimated value for the shape parameter by the Gumbel geometric distribution is -3482.2666. The negative coefficient indicates that this estimate is meaningless because the predicted income cannot be negative.

The graphical application is also important on making the decision to the best distribution for fitting any given data set. For this reason, Figure 4.13 was used to display how the five distributions looks like when fitted to this data (income prediction data).

From Figure 4.13, it can be clearly observed that three parameter Gumbel distribution and the two parameter Weibull distribution can fit the data well compared to other three distributions . Also, the worst distribution for fitting this data is the Gumbel Geometric distribution and two parameters Gumbel distribution. The graphical display of the Gumbel geometric distribution goes against its AIC value which emerged to be the second smallest value (from Table 4.15). From the analysis, the failure on this distribution is due to its shape parameter which

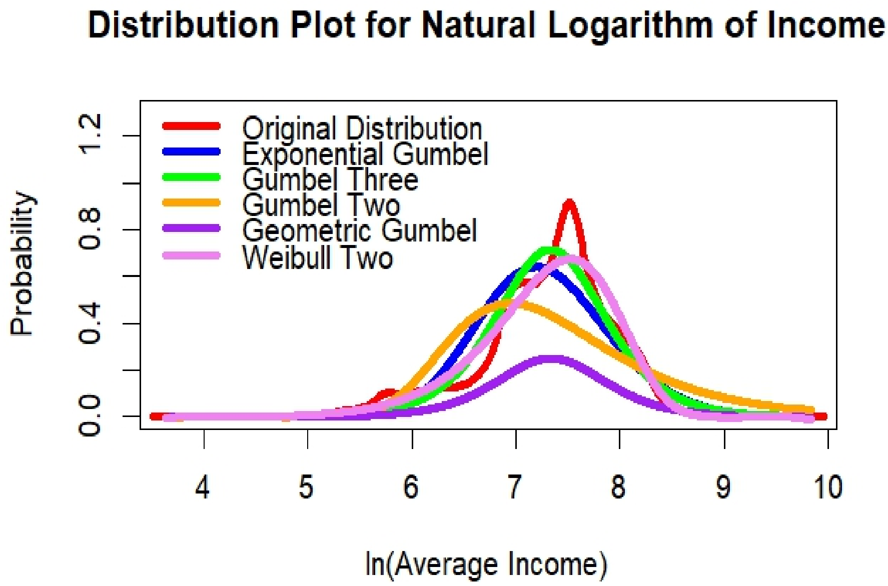

Figure 4.13: Data3: Income prediction data

seems to be poorly estimated (-3482.2666).

Therefore, from the analysis using data 3: income prediction data, the best distributions for fitting this data are the three parameters Gumbel distribution and the two parameters Weibull distribution. This is evidenced from the AIC results, efficiency test using MSE and the graphical analysis.

In general, this chapter therefore, concludes that three parameters Gumbel distribution is flexible for modeling or fitting any type of data. The application results from three data sets (earthquake magnitude data, avocado price data and income prediction data) demonstrates that this distribution is among the best and provides efficient estimates.

CHAPTER FIVE

Summary, Conclusion and Recommendations

This chapter summarizes the analysis based on the objectives to give conclusion and recommendation from the research. The summary is explained in section (5.1), conclusion of the research is discussed in section (5.2) and the research recommendation discussed in section (5.3) respectively.

5.1 Summary

In summary, three parameters Gumbel distribution is a new distribution is a positively skewed distribution with better parameter estimators and therefore, it implies that this new distribution can be applied in analyzing extreme data sets as evidenced from the improved distribution (two parameter Gumbel distribution). The distribution is also able to analyze normal and skewed data sets but only if the data sets are very large since from the discussion it is evidenced that as the sample size approaches infinity, the Gumbel distribution becomes similar to the normal and chi square distributions. This is one of the superior achievement that the new three Gumbel distribution have made.

5.2 Conclusion

A positively skewed three parameters Gumbel distribution provided unbiased estimators under the property of asymptotic bias. The distribution further shows that its estimators are more efficient and more consistent as the sample size becomes bigger (approaching infinity).

Flexibility of a three-parameters Gumbel distribution: From the simulated data, a three parameters Gumbel distribution demonstrates that it is flexible for modeling or fitting any type of data, including normal, skewed, and extreme data sets, and provides more efficient and consistent estimates for these data sets as the sample size approaches infinity. On the same, it was found that three-parameters Gumbel distribution is best for fitting the extreme data, followed normal data and lastly, skewed data.

Bias and Efficiency: The new three parameters Gumbel distribution provided unbiased estimators under the property of asymptotic bias. The distribution further shows that its estimators are more efficient and more consistent as the sample size becomes larger (approaching infinity).

Consistency and MSE: As the sample size of any data sets becomes large, three parameter Gumbel distribution provides more reliable and precise estimators with their standard errors tending to zero.

Comparison with other distributions: By applying real life data (earthquake magnitude data and avocado prices data), a three parameters Gumbel distribution is found to be more flexible than the original two parameters Gumbel distribution and other distributions like exponentiated Gumbel, three parameters Frechet, two parameters Weibull and Gumbel geometric. This is evidenced by small Akaike's Information Criterion values and more efficient estimators when compared with these distribution.

5.3 Recommendation

This research made the following specific and future research recommendations which are significant for guiding future researchers on the application, modeling and parameter estimation:

Specific Recommendations

1. Three parameter Gumbel distribution is more flexible and is recommended for future analysis since it will provide unbiased, more efficient and consistent

estimators especially for large sample sized data sets.

2. Three parameter Gumbel distribution can be applied to the analysis of both extreme data, normal data and skewed data sets and is best in all the data cases (extreme, normal and skewed) when the sample size is approaching infinity

Recommendation for further research

1. The graphical analysis shows that location parameter of a three parameter Gumbel distribution is not that close to the normal, Chi-square or Weibull location parameters. Therefore, it is recommended that a future research can me done to modify the location parameter of a three parameters Gumbel distribution to make it more efficient.

2. Future researchers to apply other parameter estimation methods like Minimum Distance Estimation, Bayesian Estimation, Method of Moments and Least Square Estimation to investigate if any of the above stated methods can provide better estimates for a three parameter Gumbel distribution than the applied method of Maximum Likelihood Estimation. This future research can help in improving the level of efficiency of the parameter estimation using a three parameter Gumbel distribution

3. Researchers to study some characteristics of a three parameter Gumbel distribution like quartile deviation, order statistics and characteristic function.

4. To examine the efficiency of the estimators of a three parameter Gumbel distribution and estimators of well known distributions like normal, exponential among others.

REFERENCES

Abouelmagd, T., Hamed, M., Afify, A. Z., Al-Mofleh, H., & Iqbal, Z. (2018). The burr x fréchet distribution with its properties and applications. *Journal of Applied Probability and Statistics*, *13*(1), 23–51.

Afify, A. Z., Zayed, M., & Ahsanullah, M. (2018). The extended exponential distribution and its applications. *Journal of Statistical Theory and Applications*, *17*(2), 213–229.

Al-Subha, S., & Alodatb, M. (2017). Discrimination between logistic and gumbel distribution. *International Journal of Sciences: Basic and Applied Research*, *36*, 244–255.

Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208). Springer.

Coles, S., & Pericchi, L. (2003). Anticipating catastrophes through extreme value modelling. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *52*(4), 405–416.

Gholami, G., Pourdarvish, A., MirMostafaee, S. M. T., Alizadeh, M., & Nashi, A. N. (2020). On the gamma gumbel distribution. *Applied Mathematics E-Notes*, *20*, 142–154.

Gumbel, E. J. (1958). *Statistics of extremes*. Columbia university press.

Gupta, R., & Biswas, A. (2010). Wind data analysis of silchar (assam, india) by rayleigh's and weibull methods. *Journal of mechanical engineering research*, *2*(1), 010–024.

Hassan, A., Dar, I., & Lone, M. (2022). A novel family of generating distributions based on marshall olkin transformation with an application to exponential distribution. *Pakistan Journal of Statistics*, *38*(1).

Hurlin, C. (2013). Maximum likelihood estimation. *Advanced econometrics. The University of Orleans*.

Hussein, E. A., Aljohani, H. M., & Afify, A. Z. (2021). The extended weibull–fréchet distribution: Properties, inference, and applications in medicine and engineering. *AIMS Mathematics*, *7*(1), 225–246.

Jaya, Garg, M., & Gupta, M. (2016). The lomax-gumbel distribution. *Palestine Journal of Mathematics*, *5*(1), 35–42.

Johnson, R. W., Kliche, D. V., & Smith, P. L. (2011). Comparison of estimators for parameters of gamma distributions with left-truncated samples. *Journal of applied meteorology and climatology*, *50*(2), 296–310.

Jonsson, F. (2014). Statistical studies of the beta gumbel distribution. *Statistica Applicata - Italian journal of Applied Statistics*.

Khalil, M. G., & Rezk, H. (2019). Extended poisson fréchet distribution and its applications. *Pakistan Journal of Statistics and Operation Research*, 905–919.

Kotz, S., & Nadarajah, S. (2000). *Extreme value distributions: Theory and applications*. world scientific.

Mansour, M. M., Elrazik, E., Altun, E., Afify, A. Z., & Iqbal, Z. (2018). Pak. j. statist. 2018 vol. 34 (6), 441-458 a new three-parameter fréchet distribution: Properties and applications. *Pak. J. Statist*, *34*(6), 441–458.

Nadarajah, S. (2006). The exponentiated gumbel distribution with climate application. *Environmetrics: The official journal of the International Environmetrics Society*, *17*(1), 13–23.

Neamah, M., & Qasim, B. (2021). A new left truncated gumbel distribution: Properties and estimation. *Journal of Physics: Conference Series*, *1897*(1), 012015.

Oseni, B., & Okasha, H. (2020). Gumbel-geometric distribution: Properties and applications. *Gazi University Journal of Science*, *33*(4), 925–941.

Persson, K., & Rydén, J. (2010). Exponentiated gumbel distribution for estimation of return levels of significant wave height. *Journal of Environmental Statistics*, *1*(3), 1–12.

Qaffou, A., & Zoglat, A. (2017). Discriminating between normal and gumbel distributions. *REVSTAT-Statistical Journal*, *15*(4), 523–536.

Ramos, P. L., Nascimento, D., & Louzada, F. (2017). The long term fr\'echet distribution: Estimation, properties and its application. *arXiv preprint arXiv:1709.07593*.

Rychlik, I., & Rydén, J. (2006). Probability and risk analysis.

Saleh, H., Aly, A. A. E.-A., & Abdel-Hady, S. (2012). Assessment of different methods used to estimate weibull distribution parameters for wind speed in zafarana wind farm, suez gulf, egypt. *Energy*, *44*(1), 710–719.

Salma, O., & Abdelali, A. (2018). Comparing maximum likelihood, least square and method of moments for tas distribution. *Journal of Humanities and Applies science*, *11*, 1–19.

Teamah, A.-E. A., Elbanna, A. A., & Gemeay, A. M. (2020). Fréchet-weibull distribution with applications to earthquakes data sets. *Pakistan Journal of Statistics*, *36*(2).

Willayat, F., Saud, N., Ijaz, M., Silvianita, A., & El-Morshedy, M. (2022). Marshall–olkin extended gumbel type-ii distribution: Properties and applications. *Complexity*, *2022*(1), 2219570.

Zhou, W., Xie, J., Zhang, B., & Li, G. (2018). Maximum likelihood detector in gamma-distributed sea clutter. *IEEE Geoscience and Remote Sensing Letters*, *15*(11), 1705–1709.

Appendix

R-codes

Simulation using a three-parameter Gumbel distribution

```
gumbel.three¡-function(n,location,scale1,shape){
u=runif(n,0,1)
x1=(u*(1-shape))-1
x2=(u*(1-shape))*(1+exp(-exp(location/scale1)))-u-exp(-exp(location/scale1))
x3=log(x1/x2)
x=-location-(scale1*log(x3))
return(x)
}
x¡-gumbel.three(1000,4,5,10)
hist(x,col = "orange")
```

Normal data sets

```
set.seed(100)
library(bbmle)
#Normal Sample size 100
x=rnorm(100,5,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fitn1< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method="Nelder-Mead")
stats4::summary(fitn1)
stats4::vcov(fitn1)
stats4::confint(fitn1)
cov2cor(vcov(fitn1))
stats4::AIC(fitn1)

prob.fun1< −function(x,omega=coef(fitn1)['omega'],delta=coef(fitn1)['delta'],
tau=coef(fitn1)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
```

```
}

dx< −density(x)
plot(x,prob.fun1(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 100)”)
#Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun1(x))
#Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit2, col = ”blue”, lwd = 4)
legend(”topleft”, legend = c(”Three Parameter Gumbel”, ”Original Normal Dis-
tribution”),
col =c(”blue”,”red”), lwd = 4, bty = ”n”)

#Normal Sample size 1000
x=rnorm(1000,5,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)[’omega’],delta=coef(fit)[’delta’],
tau=coef(fit)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 1000)”)
#Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
```

```
#Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit2, col = ”blue”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Normal Distribution”),
col =c(”blue”,”red”), lwd = 4, bty = ”n”)

    #Normal Sample size 10000
x=rnorm(10000,5,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

    fit<-stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

    prob.fun< −function(x,omega=coef(fit)[’omega’],delta=coef(fit)[’delta’],
tau=coef(fit)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom= (1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

    dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 10000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit2, col = ”blue”, lwd = 4)
legend(”topleft”, legend = c(”Three Parameter Gumbel”, ”Original Normal Dis-
tribution”),
col =c(”blue”,”red”), lwd = 4, bty = ”n”)
```

```
#Normal Sample size 100000
x=rnorm(100000,5,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}
```

```
fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)
```

```
prob.fun¡-function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'],
tau=coef(fit)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}
```

```
dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 100000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit2, col = ”blue”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Normal Distribution”),
col =c(”blue”,”red”), lwd = 4, bty = ”n”)
```

```
#Normal Sample size 1000000
x=rnorm(1000000,5,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
```

```
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

    fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

    prob.fun< −function(x,omega=coef(fit)[’omega’],delta=coef(fit)[’delta’],
tau=coef(fit)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

    dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 1000000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit2, col = ”blue”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Normal Distribution”),
col =c(”blue”,”red”), lwd = 4, bty = ”n”)
```

Skewed data sets

```
#Chisquare
set.seed(100)
#sample size 100
x=rchisq(100,5)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau
) f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
```

return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1), method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)[’omega’],delta=coef(fit)[’delta’], tau=coef(fit)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=$(1-((1-delta)*(1+exp(-exp(omega/tau))-exp(-exp(k)))))^2$
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”, main=”Density Plot (Sample Size 100)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”darkgreen”)
lines(fit2, col = ”orange”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Chisquare Distribution”),
col =c(”orange”,”darkgreen”), lwd = 4, bty = ”n”)

#sample size 1000
x=rchisq(1000,5)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1), method=”Nelder-Mead”)
summary(fit)
vcov(fit)

```
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'],
tau=coef(fit)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type="n",ylim=c(0,0.23), ylab="Probability",
main="Density Plot (Sample Size 1000)")
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = "darkgreen")
lines(fit2, col = "orange", lwd = 4)
legend("topleft",
legend = c("Three Parameter Gumbel", "Original Chisquare Distribution"),
col =c("orange","darkgreen"), lwd = 4, bty = "n")

#sample size 10000
x=rchisq(10000,5)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method="Nelder-Mead")
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'],
tau=coef(fit)['tau']){
k=-((x-omega)/tau)
```

```
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 10000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”darkgreen”)
lines(fit2, col = ”orange”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Chisquare Distribution”),
col =c(”orange”,”darkgreen”), lwd = 4, bty = ”n”)

#sample size 100000
x=rchisq(100000,5)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)[’omega’],delta=coef(fit)[’delta’],
tau=coef(fit)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

dx< −density(x)
```

```
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 100000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”darkgreen”)
lines(fit2, col = ”orange”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Chisquare Distribution”),
col =c(”orange”,”darkgreen”), lwd = 4, bty = ”n”)

    #sample size 1000000
x=rchisq(1000000,5)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

    fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

    prob.fun< −function(x,omega=coef(fit)[’omega’],delta=coef(fit)[’delta’],
tau=coef(fit)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

    dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.23), ylab=”Probability”,
main=”Density Plot (Sample Size 1000000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”darkgreen”)
lines(fit2, col = ”orange”, lwd = 4)
```

legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Chisquare Distribution”),
col =c(”orange”,”darkgreen”), lwd = 4, bty = ”n”)

Extreme data

```
#Weibul Distribution
#sample size 100
x=rweibull(100,4,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'],
tau=coef(fit)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
```
denom=$(1-((1-delta)*(1+exp(-exp(omega/tau))-exp(-exp(k)))))^2$
```
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.9), ylab=”Probability”,
main=”Density Plot (Sample Size 100)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”brown”)
lines(fit2, col = ”lightgreen”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Weibull Distribution”),
col =c(”lightgreen”,”brown”), lwd = 4, bty = ”n”)
```

```
#Weibul Distribution
#sample size 1000
x=rweibull(1000,4,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'],
tau=coef(fit)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.9), ylab=”Probability”,
main=”Density Plot (Sample Size 1000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”brown”)
lines(fit2, col = ”lightgreen”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Weibull Distribution”),
col =c(”lightgreen”,”brown”), lwd = 4, bty = ”n”)

#Weibul Distribution
#sample size 10000
x=rweibull(10000,4,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
```

k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1), method=”Nelder-Mead”)
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun< −function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'], tau=coef(fit)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=$(1-((1-delta)*(1+exp(-exp(omega/tau))-exp(-exp(k)))))^2$
p=num/denom
return(p)
}

dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.9), ylab=”Probability”, main=”Density Plot (Sample Size 10000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”brown”)
lines(fit2, col = ”lightgreen”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Weibull Distribution”),
col =c(”lightgreen”,”brown”), lwd = 4, bty = ”n”)

#Weibul Distribution
#sample size 100000
x=rweibull(100000,4,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)

}

fit¡-stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1), method="Nelder-Mead")
summary(fit)
vcov(fit)
confint(fit)
cov2cor(vcov(fit))
stats4::AIC(fit)

prob.fun$< -$function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'], tau=coef(fit)['tau']){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=$(1-((1-delta)*(1+exp(-exp(omega/tau))-exp(-exp(k)))))^2$
p=num/denom
return(p)
}

dx$< -$density(x)
plot(x,prob.fun(x), type="n",ylim=c(0,0.9), ylab="Probability",
main="Density Plot (Sample Size 100000)")
# Fit a smooth spline to the data fit2 $< -$ smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = "brown")
lines(fit2, col = "lightgreen", lwd = 4)
legend("topleft",
legend = c("Three Parameter Gumbel", "Original Weibull Distribution"),
col =c("lightgreen","brown"), lwd = 4, bty = "n")

#Weibul Distribution
#sample size 1000000
x=rweibull(1000000,4,2)
gambel.three.link¡-function(delta, omega, tau){
n=length(x)
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}

fit$< -$stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1), method="Nelder-Mead")
summary(fit) vcov(fit) confint(fit) cov2cor(vcov(fit)) stats4::AIC(fit)
prob.fun$< -$function(x,omega=coef(fit)['omega'],delta=coef(fit)['delta'], tau=coef(fit)['tau']){

```
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}
```

```
dx< −density(x)
plot(x,prob.fun(x), type=”n”,ylim=c(0,0.9), ylab=”Probability”,
main=”Density Plot (Sample Size 1000000)”)
# Fit a smooth spline to the data
fit2 < − smooth.spline(x, prob.fun(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”brown”)
lines(fit2, col = ”lightgreen”, lwd = 4)
legend(”topleft”,
legend = c(”Three Parameter Gumbel”, ”Original Weibull Distribution”),
col =c(”lightgreen”,”brown”), lwd = 4, bty = ”n”)
```

Efficiency Analysis

```
# Exponentiated Gumbel ex.gumbel.link< −function(beta, sigma, mu){
w = length(x)
q = ((x - mu)/sigma)
z1 = (w*log(beta)) - (w*sigma)
z2 = (beta-1)*(sum(log(1-exp(-exp(-q)))))
z3 = (sum(q))
z4 = sum(exp(-q))
ex.gumbel.link1=z1+z2-z3-z4
return(-ex.gumbel.link1)
}
efitn1< −stats4::mle(ex.gumbel.link,start=list(beta=1, sigma=1, mu=1),
method=”Nelder-Mead”)
stats4::summary(efitn1)
stats4::AIC(efitn1)
expgprob.fun1< −function(x,beta=coef(efitn1)[’beta’],
sigma=coef(efitn1)[’sigma’],mu=coef(efitn1)[’mu’]){
k=(-(x-mu)/sigma)
a1=(beta/sigma)*exp(k)*exp(-exp(k))
a2=(1 − exp(−exp(k)))^(beta − 1)
p=a1*a2
return(p)
}
```

```
# Three parameter Gumbel
gambel.three.link< −function(delta, omega, tau){
n=length(x)
```

```
k=(-(x-omega)/tau)
f1=(n*log(delta))-(n*log(tau))
f2=sum(k-exp(k))
f3=2*sum(log(1-((1-delta)*(1-exp(-exp(k))+exp(-exp(omega/tau))))))
gambel.link=f1+f2-f3
return(-gambel.link)
}
g3fitn1< −stats4::mle(gambel.three.link,start=list(delta=1, omega=1, tau=1),
method=”Nelder-Mead”)
stats4::summary(g3fitn1)
stats4::AIC(g3fitn1)
prob.fun1< −function(x,omega=coef(g3fitn1)[’omega’],delta=coef(g3fitn1)[’delta’],
tau=coef(g3fitn1)[’tau’]){
k=-((x-omega)/tau)
num=(delta/tau)*exp(k-exp(k))
denom=(1 − ((1 − delta) ∗ (1 + exp(−exp(omega/tau)) − exp(−exp(k)))))^2
p=num/denom
return(p)
}

# Gumbel distribution gumbel1.link< −function(mu1,theta1){
n = length(x)
k=(-(x-mu1)/theta1)
c1= -n*log(theta1)
c2=sum(k-exp(k))
gumbel1.link1=c1+c2
return(-gumbel1.link1)
}
gfitn1< −stats4::mle(gumbel1.link,start=list(mu1=1, theta1=1),
method=”Nelder-Mead”)
stats4::summary(gfitn1)
stats4::AIC(gfitn1)
g2prob.fun1< −function(x,mu1=coef(gfitn1)[’mu1’],
theta1=coef(gfitn1)[’theta1’]){
k=-((x-mu1)/theta1)
a1=1/theta1
a2=exp(k-exp(k))
p=a1*a2
return(p)
}

# Gumbel geometric
gumbelgeo.link< −function(mu2, sigma2, lambda2){
n = length(x)
p=exp(-(x-mu2)/sigma2)
a1=n*log(1-lambda2)
a2=n*log(sigma2)
a3=2*sum(log(1-lambda2+(lambda2*exp(-p))))
```

a4=sum(log(p))-sum(p)
gumbelgeo.link1=a1-a2-a3+a4
return(-gumbelgeo.link1)
}
ggfitn1< −stats4::mle(gumbelgeo.link,start=list(mu2=0.5, sigma2=0.5, lambda2=0.5),
method="Nelder-Mead")
stats4::summary(ggfitn1)
stats4::AIC(ggfitn1)
ggprob.fun1< −function(x,mu2=coef(ggfitn1)['mu2'],sigma2=coef(ggfitn1)['sigma2'],
lambda2=coef(ggfitn1)['lambda2']){
k=-((x-mu2)/sigma2)
a1=(1-lambda2)
a2=exp(k-exp(k))
a3=$(1 - lambda2 + (lambda2 * exp(-exp(k))))^(- 2)$
p=a1*a2*a3
return(p)
}

# Frechet
fretchet.like< −function(alpha2,beta2,theta2){
n=length(x)
k2=$(alpha2/x)^{beta2}$
a1=n*(log(beta2)+(beta2*log(alpha2))+(2*log(theta2))-log(1+theta2))
a2=-(beta2+1)*sum(log(x))
a3=-sum(k2)
a4=-3*sum(log(1-exp(-k2)))
a5=-theta2*sum(exp(-k2)/(1-exp(-k2)))
fretchet.like1=a1+a2+a3+a4+a5
return(-fretchet.like1)
}

# Weibull
weibul.like< −function(tau,beta){
n=length(x)
a1=n*log(beta)-(n*beta*log(tau))
a2=(beta-1)*sum(log(x))
a3=$sum((x/tau)^beta a)$
weibul.like1=a1+a2-a3
return(-weibul.like1)
}

web2fitn1< −stats4::mle(weibul.like,start=list(tau=2, beta=2)
method="Nelder-Mead")
stats4::summary(web2fitn1)
stats4::AIC(web2fitn1)
webprob.fun1< −function(x,beta=coef(web2fitn1)['beta'],
tau=coef(web2fitn1)['tau']){
a1=$(beta * (x^(beta - 1)))/(tau^b eta)$

a2=$exp(-((x/tau)^beta a))$
p=a1*a2
return(p)
}

dx< −density(x)
plot(x,prob.fun1(x), type=”n”,ylab=”Probability”,
main=”Distribution Plot for Earthquake Magnitude”, xlab=”Earthquake Magnitude”)
# Fit a smooth spline to the data
fit1 < − smooth.spline(x, expgprob.fun1(x))
fit2 < − smooth.spline(x, prob.fun1(x))
fit3 < − smooth.spline(x, g2prob.fun1(x))
fit4 < − smooth.spline(x, ggprob.fun1(x))
fit5 < − smooth.spline(x, freprob.fun1(x))
fit6 < − smooth.spline(x, webprob.fun1(x))

# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit1, col = ”blue”, lwd = 4)
lines(fit2, col = ”green”, lwd = 4)
lines(fit3, col = ”orange”, lwd = 4)
lines(fit4, col = ”purple”, lwd = 4)
lines(fit5, col = ”brown”, lwd = 4)
lines(fit6, col = ”violet”, lwd = 4)

legend(”topright”,
legend = c(”Original Distribution”, ”Exponential Gumbel”,
”Gumbel Three”,”Gumbel Two”,”Geometric Gumbel”,”Frechet”,”Weibull Two”),
col =c(”red”,”blue”,”green”,”orange”,”purple”,”brown”,”violet”),
lwd = 4, bty = ”n”)

dx< −density(x) plot(x,prob.fun1(x), type=”n”,ylab=”Probability”,
main=”Distribution Plot for Average Avocardo Price”,
xlab=”Average Avocardo Price”)
# Fit a smooth spline to the data
fit1 < − smooth.spline(x, expgprob.fun1(x))
fit2 < − smooth.spline(x, prob.fun1(x))
fit3 < − smooth.spline(x, g2prob.fun1(x))
fit4 < − smooth.spline(x, ggprob.fun1(x))
fit5 < − smooth.spline(x, freprob.fun1(x))
fit6 < − smooth.spline(x, webprob.fun1(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit1, col = ”blue”, lwd = 4)
lines(fit2, col = ”green”, lwd = 4)
lines(fit3, col = ”orange”, lwd = 4)
lines(fit4, col = ”purple”, lwd = 4)

```
lines(fit5, col = ”brown”, lwd = 4)
lines(fit6, col = ”violet”, lwd = 4)

    legend(”topright”,
legend = c(”Original Distribution”, ”Exponential Gumbel”,
”Gumbel Three”,”Gumbel Two”,”Geometric Gumbel”,”Frechet”,”Weibull Two”),
col =c(”red”,”blue”,”green”,”orange”,”purple”,”brown”,”violet”),
lwd = 4, bty = ”n”)

    dx< −density(x)
plot(x,prob.fun1(x), type=”n”,ylim=c(0,1.3),ylab=”Probability”,
main=”Distribution Plot for Natural Logarithm of Income”,
xlab=”ln(Average Income)”)
# Fit a smooth spline to the data
fit1 < − smooth.spline(x, expgprob.fun1(x))
fit2 < − smooth.spline(x, prob.fun1(x))
fit3 < − smooth.spline(x, g2prob.fun1(x))
fit4 < − smooth.spline(x, ggprob.fun1(x))
fit6 < − smooth.spline(x, webprob.fun1(x))
# Add the smooth curve to the plot
lines(dx, lwd = 4, col = ”red”)
lines(fit1, col = ”blue”, lwd = 4)
lines(fit2, col = ”green”, lwd = 4)
lines(fit3, col = ”orange”, lwd = 4)
lines(fit4, col = ”purple”, lwd = 4)
lines(fit6, col = ”violet”, lwd = 4)

    legend(”topleft”,
legend = c(”Original Distribution”, ”Exponential Gumbel”,
”Gumbel Three”,”Gumbel Two”,”Geometric Gumbel”,”Weibull Two”),
col =c(”red”,”blue”,”green”,”orange”,”purple”,”violet”), lwd = 4, bty = ”n”)
```