Three-parameters Gumbel distribution: Properties and Application

Abstract: In this research, we introduced a new three-parameter Gumbel distribution by adding a parameter to the traditional Gumbel distribution using the Marshall-Olkin method. This new distribution enhances flexibility and provides more efficient estimators for various data types, including normal, skewed, and extreme data. We derived the probability density function, cumulative distribution function, and other statistical properties of the new distribution. The parameters are estimated using the Maximum Likelihood Estimation (MLE) method, and thoroughly investigated the properties of the estimators, focusing on their asymptotic bias, consistency, and mean square error (MSE). Through simulation studies and real data applications, we demonstrate the superiority of the new distribution over existing models, evidenced by smaller Akaike Information Criterion (AIC) values and more efficient parameter estimates. We recommend the new distribution for future analyses, particularly for large sample sizes, and suggest further research to refine the location parameter for improved efficiency.

Keywords: Asymptotic,Unbiasedness , Mean Square Error, Consistency, Three parameters Gumbel distribution

1. Introduction

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 [3, 15, 18].

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[5, 6, 11].

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 [2, 4, 8, 19]. 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)[12, 19]. 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.

The research have found that adding a parameter to any existing distribution makes it 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 [13, 14]. However, from the reviewed literature, it was 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 about which we consider three parameters (that is, shape, location and dispersion) using Marshall Olkin (19997) proposed method. Since the extreme value analysis address the extreme deviations from the centre of probability distribution and it focus on limiting distributions which are distinct from normal distribution. 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. 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.

This research therefore developed a new distribution called a three parameters Gumbel distribution by adding the shape parameter to the already existing two parameter distribution with location and dispersion parameters. The new distribution was developed by applying the Marshall Olkin method for adding a new parameter to an existing distribution. To estimate a parameters of the three parameters Gumbel distribution, this study applied 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 [3, 6, 18, 20] 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 (ω), dispersion parameter (τ) and the shape parameter (δ). 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 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.

1.1. Three parameters Gumbel distribution

Suppose we have a random variable V, then the cdf and pdf functions for a three parameters Gumbel distribution obtained using Marshal Olkin method is as given in equations 1 and 2, respectively The corresponding cumulative distribution function is obtained as follows,

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

and a probability distribution function given as follows,

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

where δ is the introduced shape parameter

The process of estimating each of the three parameters in a three parameters distribution namely; the location parameters (ω), scale/dispersion parameter (τ) and the shape parameter (δ) using maximum likelihood estimation method. 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)[10, 16, 17].

Considering the probability distribution function given in equation (2), its likelihood function is given follows,

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

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

$$
R(v_i;\delta,\omega,\tau) = \big(\frac{\delta}{\tau}\big)^k \frac{exp\sum_{i=1}^{k}\big(-\frac{v_i-\omega}{\tau}-e^{\frac{-v_i-\omega}{\tau}}\big)}{\Big[exp\sum_{i=1}^{k} ln\big[1-(1-\delta)\big(1-exp\big(-e^{\frac{-v_i-\omega}{\tau}}\big)+exp(-e^{\frac{\omega}{\tau}})\big)\big]\Big]^2} \tag{3}
$$

The likelihood function R, can be expressed with a variable together with parameters to be estimated as shown in equation (3), and its log-likelihood function which maximizes the parameters becomes

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

$$
-2\sum_{i=1}^{k} ln\big[1-(1-\delta)\big(1-exp\big(-e^{\frac{-v_i-\omega}{\tau}}\big)+exp(-e^{\frac{\omega}{\tau}})\big)\big] \tag{4}
$$

Parameters estimation for δ, ω and τ are as presented in the following equations 5, 6 and 7 respectively.

$$
\hat{\delta} = \frac{\partial ln(L)}{\partial\delta} = \frac{k}{\delta} - 2\sum_{i=1}^{k}\frac{\big[1-exp\big(-e^{\frac{-v_i-\omega}{\tau}}\big)+exp(-e^{\frac{\omega}{\tau}})\big]}{p} = 0, \tag{5}
$$

$$
\begin{aligned}
\hat{\omega} &= \frac{\partial ln(L)}{\partial\omega} \\
&= \frac{1}{\tau}\sum_{i=1}^{k}\Big[1-e^{-\left(\frac{v_i-\omega}{\tau}\right)}\Big] + 2(1-\delta)\sum_{i=1}^{k}\left[\frac{exp\Big(-e^{\left(-\frac{v_i-\omega}{\tau}\right)}\Big).e^{\left(-\frac{v_i-\omega}{\tau}\right)}-e^{\frac{\omega}{\tau}}.exp(-e^{\frac{\omega}{\tau}})}{p\tau}\right] = 0
\end{aligned} \tag{6}
$$

$$
\hat{\tau} = \frac{\partial ln(L)}{\partial\tau} = \frac{-k}{\tau} + \frac{1}{\tau^2}\sum_{i=1}^{k}\Big[v_i-\omega\Big]\Big[1-e^{-\left(\frac{v_i-\omega}{\tau}\right)}\Big] + Q = 0 \tag{7}
$$

where,

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

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

2. Methodology

2.1. Asymptotic Properties of the Estimators

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)[1, 7, 9, 20], which we discuss below.

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 δ is used to represent a generic parameter of the population (for example μ, σ^2, p), and the symbol $\hat{\gamma}$ is used to represent the statistical estimator for γ. If the expected value of the estimator is equal to the parameter, that is , if:

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

the estimator is said to be unbiased. Otherwise, the estimator is said to be biased, that is $B = |E\left(\hat{\gamma}\right) - \gamma|$. 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;

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

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

$$MSE(\hat{\gamma}) = E[(\hat{\gamma} - \gamma)^2] = V(\hat{\gamma}) + [E(\hat{\delta}) - \gamma]^2 = V(\hat{\gamma}) + B^2 \tag{8}$$

Under minimum variance estimator, an estimator is said to be sufficient if the conditional distribution of the random samples given $\hat{\delta}$ does not depend on the parameter δ 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 $\hat{\gamma}$ be an estimator of γ 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,

$$lim_{K\to\infty} P_r\left(\left|\hat{\gamma} - \gamma\right| \geq \epsilon\right) = 0, \epsilon > 0 \tag{9}$$

Laws of large number are also used to induce if an estimator is consistent or not, that is, an estimator $\hat{\delta}$ is consistent for δ, for K observations if:
i bias$(\hat{\gamma}, \gamma) = 0$ as $K \to \infty$
ii MSE$(\hat{\gamma}, \gamma) = 0$ as $K \to \infty$
iii se$(\hat{\gamma}) = 0$ as $K \to \infty$
4. Asymptotic normality
Let $\hat{\gamma}$ be an estimator of γ based on random variable $\{V_i\}_{i=1}^{k}$. Then an estimator is said to be asymptotically normally if:

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

for large enough K, meaning that $f(\hat{\gamma})$ is known to be well approximated by normal distribution with mean γ and variance $se(\hat{\gamma})^2$
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 $E(V_i) = \mu$ and $var(V_i) = \beta^2$ is said to asymptotically normal with mean θ and variance $\frac{\beta^2}{K}$, and the cumulative density function of the standardized sample mean given as:

$$\frac{\bar{V} - \mu}{se(\bar{V})} = \frac{\bar{V} - \mu}{\frac{\sigma}{\sqrt{K}}} = \sqrt{K}(\frac{\bar{V} - \mu}{\sigma}),$$

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

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

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.

2.2. 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 below.

Akaike's Information Criterion (AIC)

The Akaike Information Criterion is computed as;

$$AIC = -2logP(W) + 2z, \tag{10}$$

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 [12]. The AIC technique was fitted on the modeled three parameters distribution and the other comparison distributions of the Gumbel family.

3. Properties and Application

3.1. Properties of the Estimators using simulated data

This study discussed the 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. It also investigate their Mean Square Error (MSE) 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). The study also discussed the consistency of the parameters to investigate the precision and reliability of the estimators as we increase the sample size. Finally, the study compared 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.

3.1.1. 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 1, 2 and 3 respectively.

Asymptotic biasedness for ω

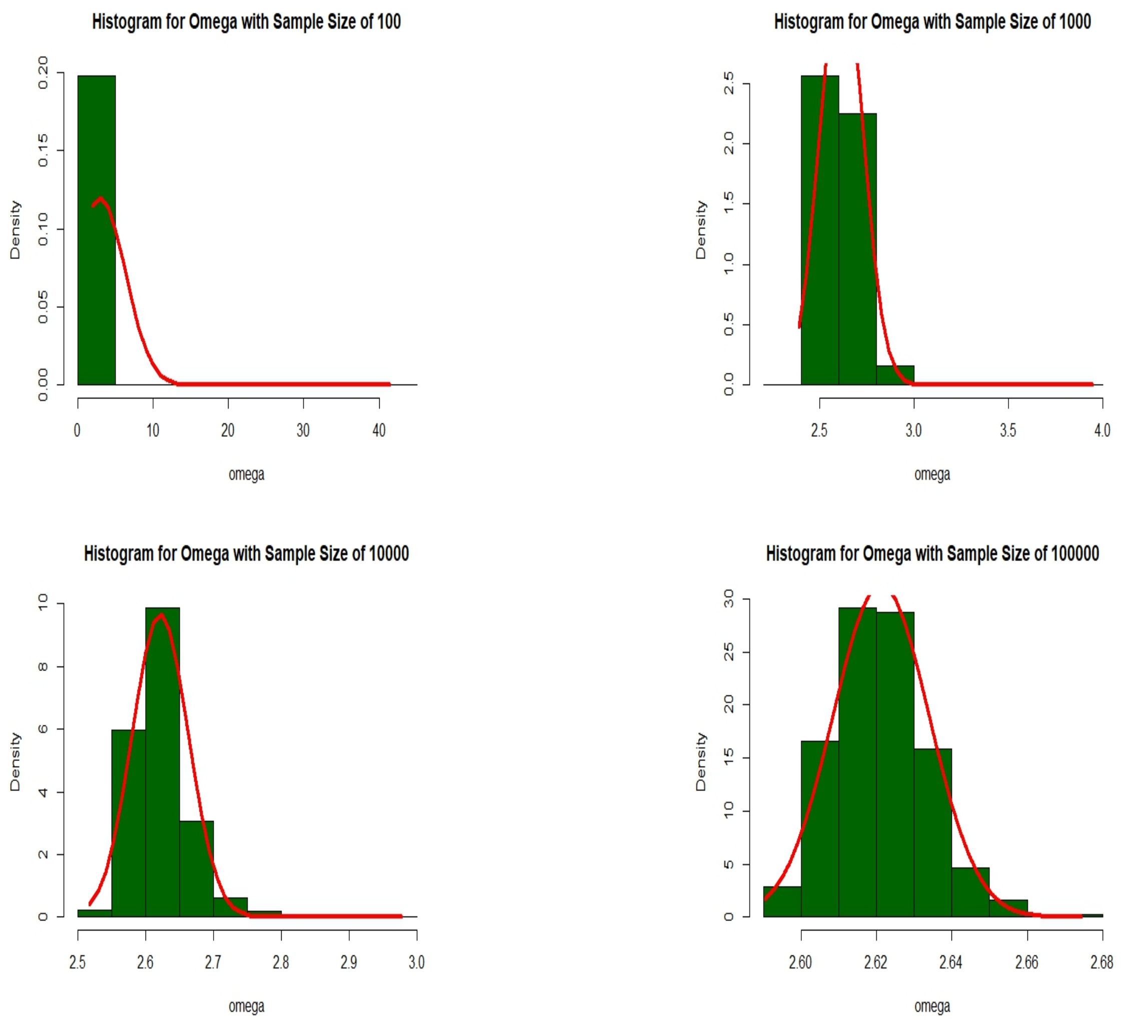

Figure 1. *Asymptotic distributions for ω*

From figure 1, it can be observed that as the sample size increases, the distribution for parameter ω approaches normal distribution. This means that the 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 2 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 $k \to \infty$.

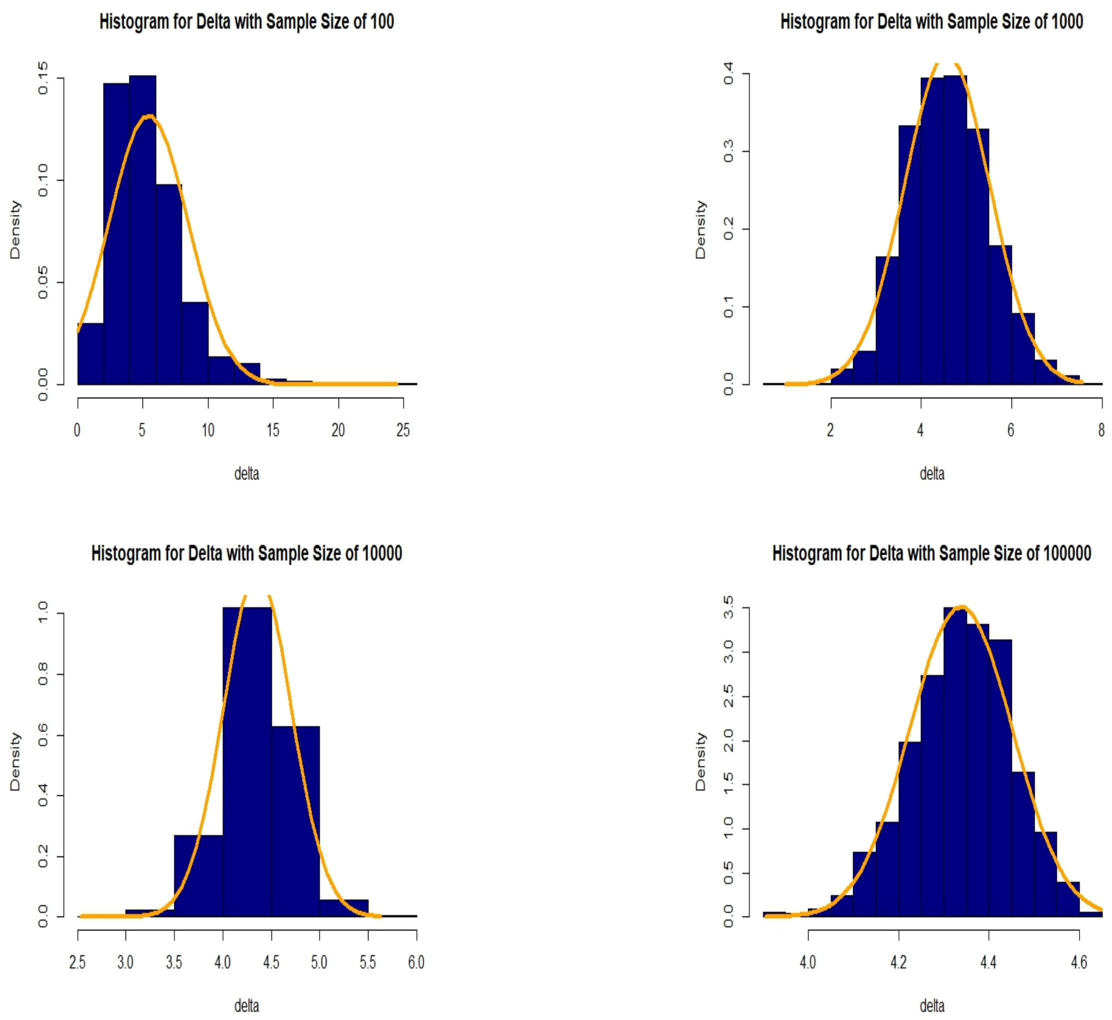

Figure 2. *Asymptotic distributions for δ*

Asymptotic biasedness for τ

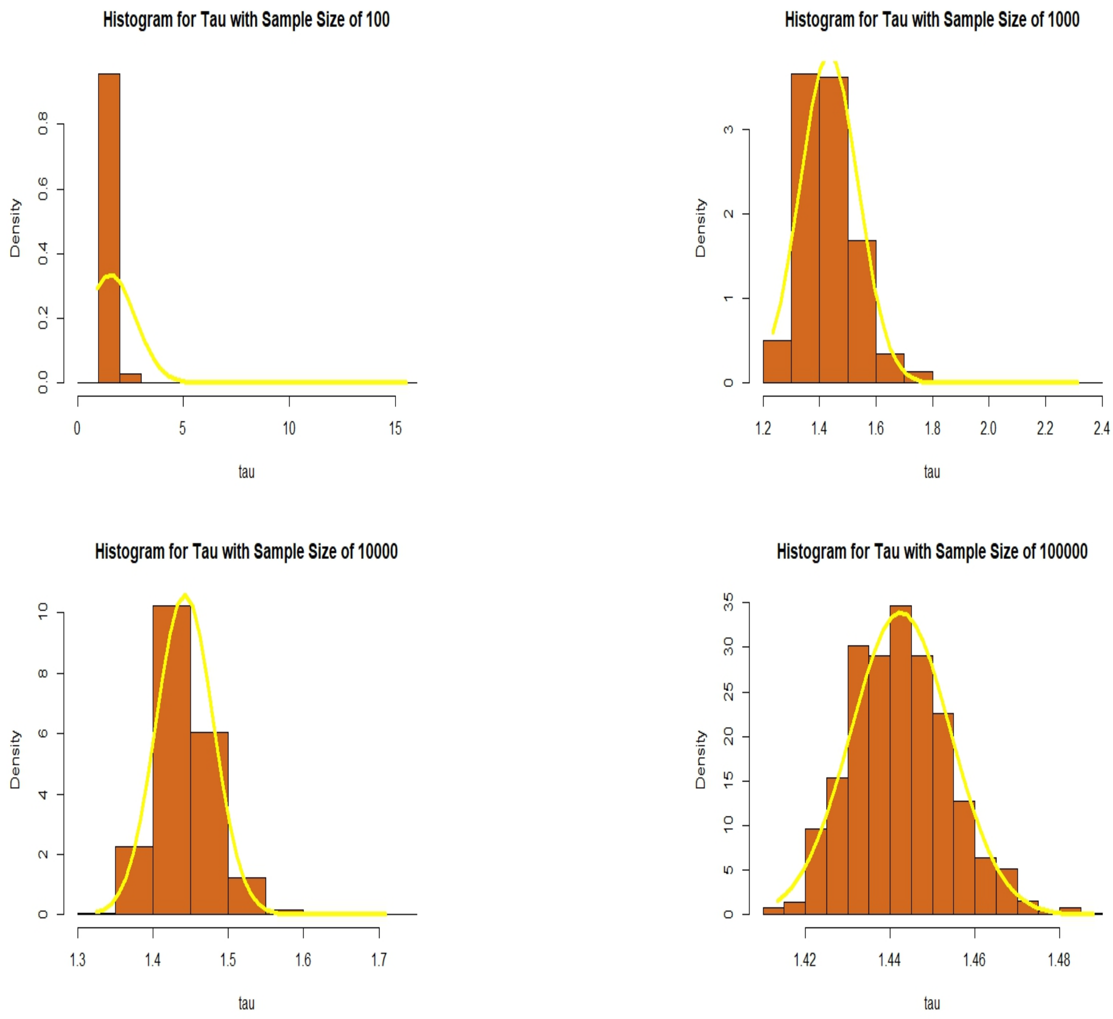

Figure 3. *Asymptotic distributions for τ*

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 3 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 3, 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 $k \rightarrow \infty$.

3.1.2. Mean Square Error

This study 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 1.

Table 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

From table 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 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.

3.1.3. Consistency

The study checked 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. 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. Asymptotic bias 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.

Table 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

From table 2, it can be seen that as the sample size increase from 100 to 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 $s.e(\hat{\omega}, \hat{\delta}, \hat{\tau}) \rightarrow 0$ as $k \rightarrow \infty$.

For the purpose of the asymptotic behaviuor of the estimates, it is important to determine the relationship between the estimates. Table 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.

In table 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

Table 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 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 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. *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 5 which shows that there is a small variation in the correlation coefficient between the estimates for any change in the sample size.

Table 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 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 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 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 7 which shows that there is a small variation in the correlation coefficient between the estimates for any change in the sample size.

Table 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 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.

3.1.4. Asymptotic relationship between three parameters Gumbel distribution and other distributions

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 as follows:

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 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.

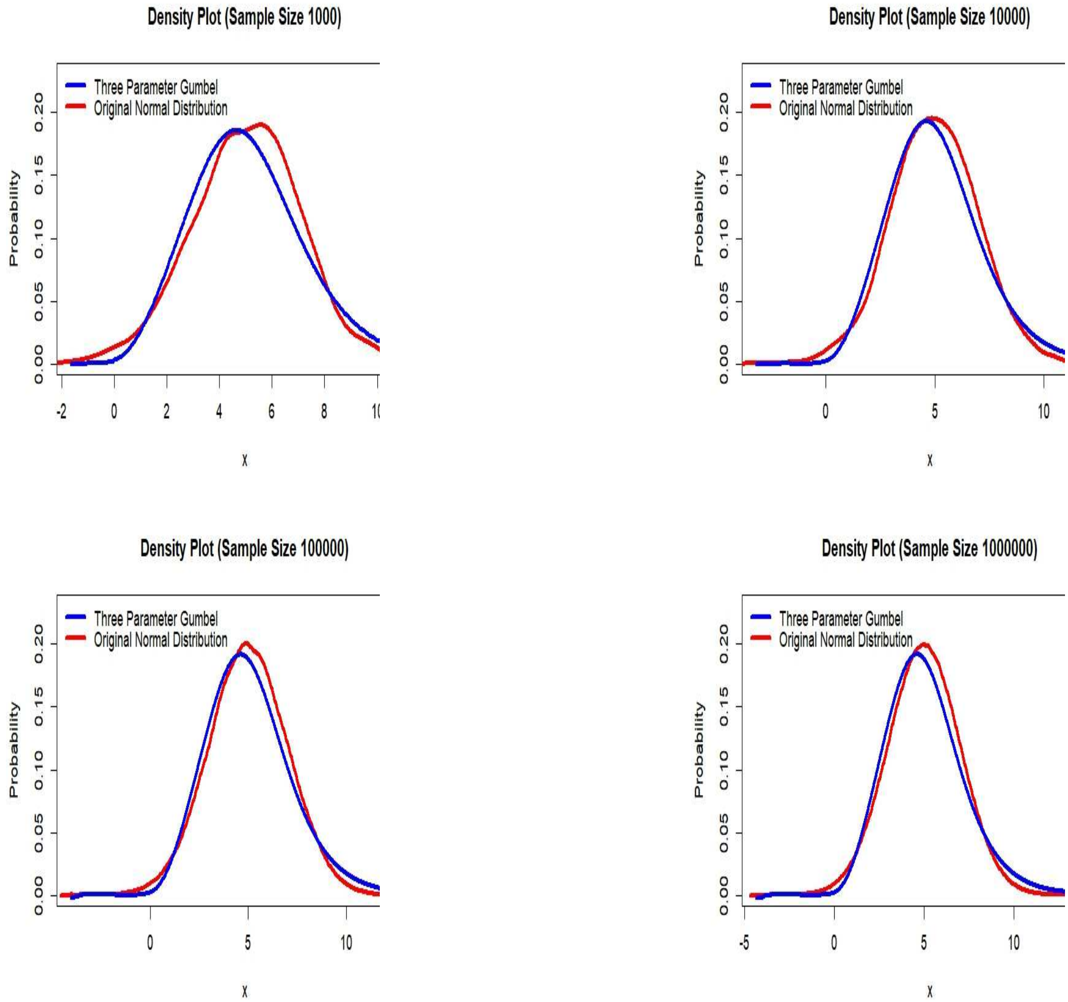

Figure 4. *Asymptotic distributions for three parameters Gumbel and normal distributions*

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 5 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.

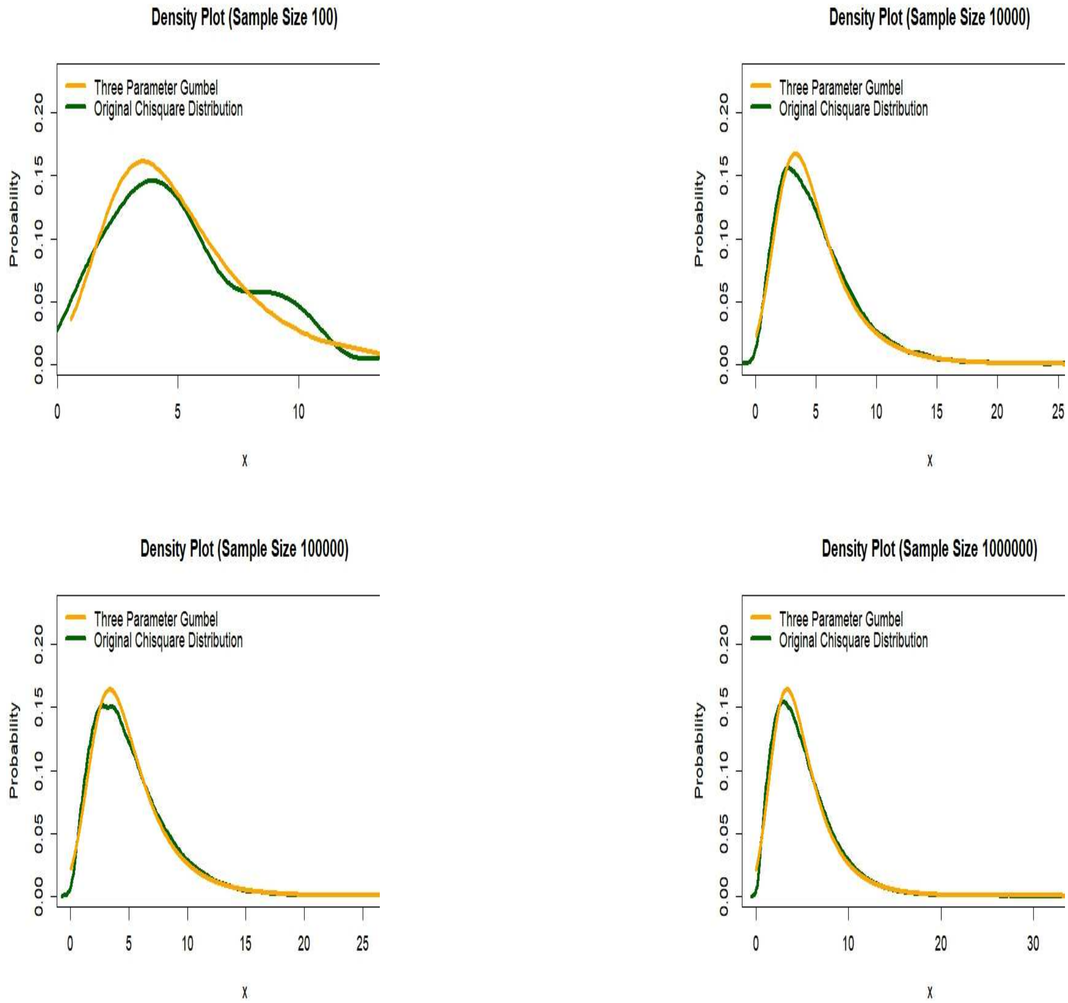

Figure 5. *Asymptotic distributions for three parameters Gumbel and chi-square distributions*

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 distribution 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 6 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.

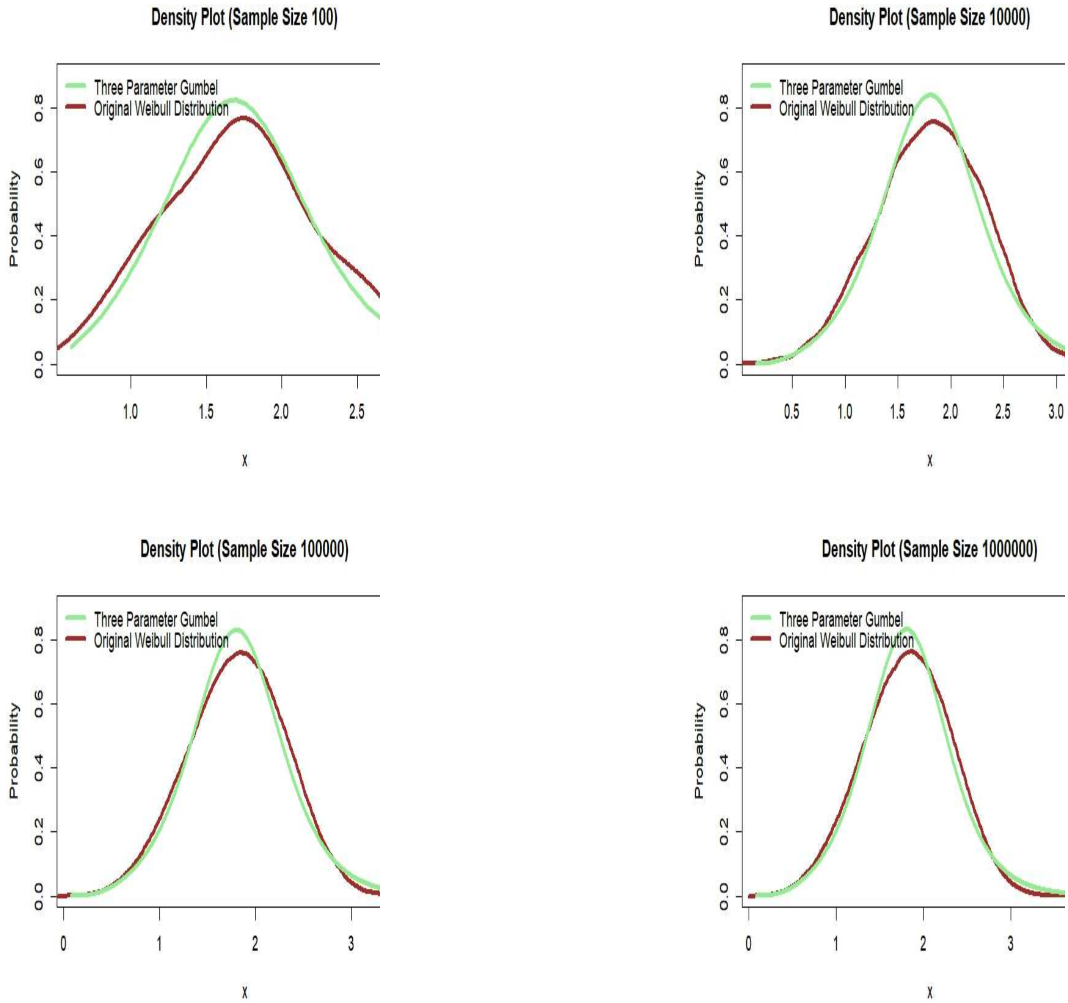

Figure 6. *Asymptotic distributions for three parameters Gumbel and Weibull distributions*

3.2. Application

The study investigated 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 section (6.1), we investigate the efficiency using three different types of simulated data sets and in section (6.2) 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.

3.2.1. Efficiency using simulated data

For the efficiency part, 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 8. From (3.1.4), 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 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.

Table 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

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.

3.2.2. Efficiency Analysis using real data

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 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 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 10.

The results provided in table 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

Table 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

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 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
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

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 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 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 10.

Lastly, the shape parameter results from table 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 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 7. The probability distribution functions for the six distribution (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).

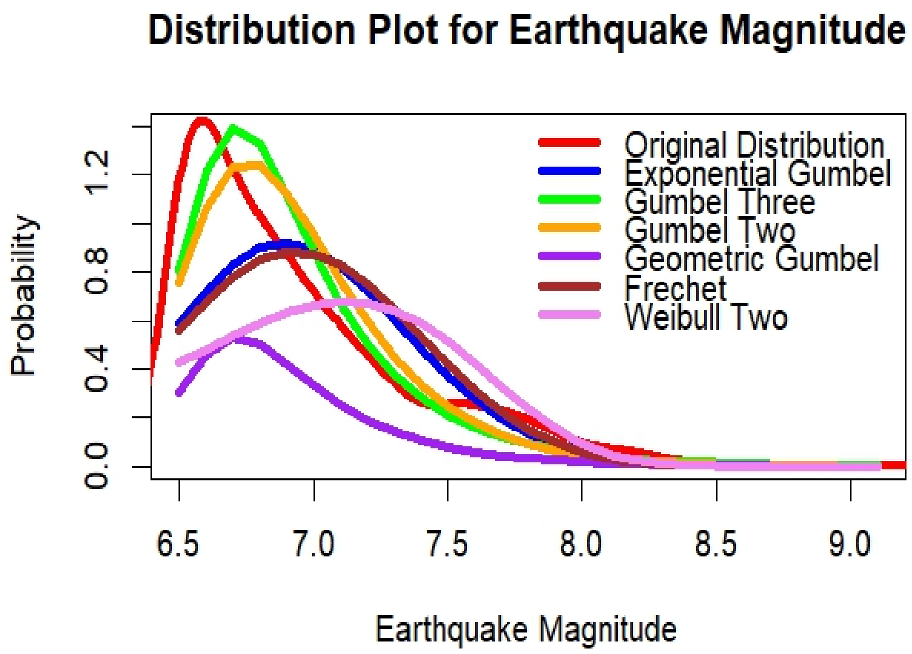

Figure 7. *Data 1:Earthquake magnitude*

The figure 7 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 magnitude 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 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.

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 13.

The results provided in table 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 48959.13, 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).

Table 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

Table 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	48959.13
Two parameter Weibull	53200.5

We 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 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
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

By applying the Cramer-Rao approach on the location parameter, it can be observed from table 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 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 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 8

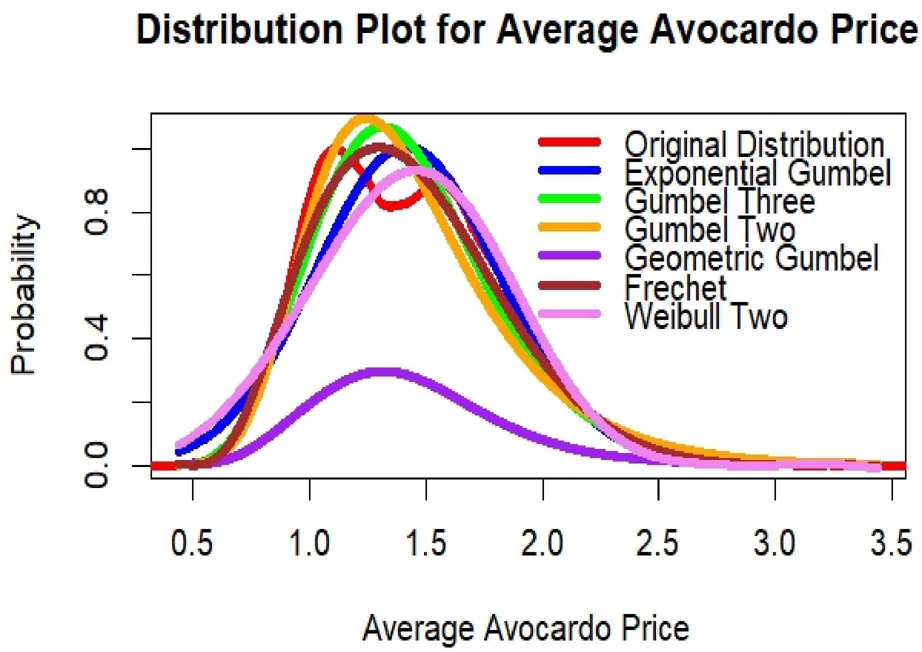

Figure 8. *Data 2: Avocado prices in dollars*

From figure 8, 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 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 efficient 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 15.

From table 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 16.

Table 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

Table 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

The results for the location parameter from table 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 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 indicate least efficient estimate for scale parameter with two two distributions displaying MSE value of 0.0010 and 0.0018 respectively.

Lastly from table 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 9, was used to display how the five distributions looks like when fitted to this data (income prediction data).

From figure 9, 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 15). From the analysis, the failure on this distribution is due to its shape parameter which 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.

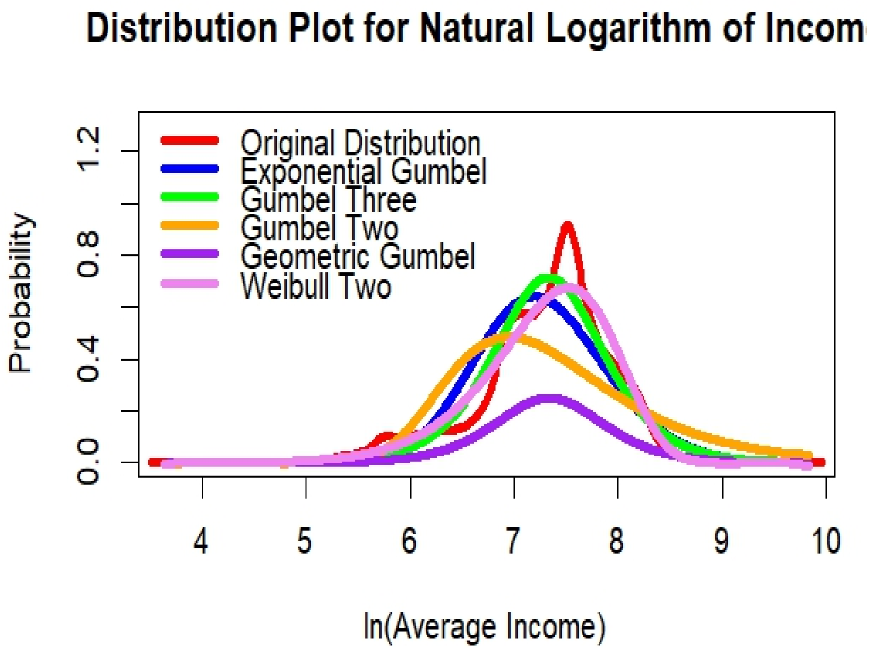

Figure 9. *Income prediction data*

4. Conclusion

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 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).

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. 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 parameters 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 parameters 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 parameters 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-parameters 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-parameters Gumbel distribution

Conflict of Interest

The authors declare no conflict of interest

References

[1] Abdelaziz, Q. and Zoglat, A. (2011).Discriminating between normal and Gumbel distributions.

[2] Abouelmagd, T. H. M., and Hamed, M. S. (2018). The Burr X Frechet distribution with its properties and applications. *Journal of Applied Probability and Statistics*, 13(1): 23 - 51.

[3] Afify, A.Z., and Zayed, M. (2018). The extended exponential distribution and its application. *Journal of Statistical Theory and Applications*, 17(2): 213 -229.

[4] Anwar, H., Dar, I.H., and Lone, M.A. (2022). A novel family of generating distributions based on Marshall Olkin transformation with an application to exponential distribution. *Pak. J. Statist*, 38(1) : 113 - 128.

[5] Coles, S. (2001).An introduction to statistical modeling of extreme values. *Springer-Verlag*.

[6] Ekramy, A.H., Aljohani, H.M., and Ahmed, Z.A. (2021). The extended Weibull-Frechet distribution: properties, inference, and application in medicine and engineering. *AIMS Mathematics*, 7(1): 225 - 246.

[7] Garg, M., Gupta, J., and Gupta, M., (2016). The Lomax Gumbel. *Palestine journal of Mathematics*, 5(1):35 - 42.

[8] Gumbel, E.J. (1958). Statistics of extreme, *Colombia University Press*.

[9] Gupta, R., and Biswas, A. (2010). Wind data analysis of Silchar (Assam India) by Rayleigh and Weibull methods: *Journal of mechanical engineering research*, 2: 10-24.

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

[11] Khalil, G.M., and Rezk, H. (2019). Extended poison Frechet distribution and its application, 15(4): 905 - 919.

[12] Klara, P., and Jesper, R. (2010). Exponentiated Gumbel distribution for estimation of return levels of significant wave heights.*Journal of environmental statistics*, 1(3).

[13] Lawan, S.M., Abidin, W.A.W.Z., Chai, W.Y., Baharum, A., and Masri, T., (2015). Statistical modelling of long-term wind speed data : *American journal of computer science and information technology*,13 : 79-121.

[14] Neamah, M.W., and Qasim, B.A., (2021). A new truncated Gumbel distribution: Properties and Estimation. *Journal of Physics: Conference series*

[15] Pedro, L. R., Nascimento, D., and Louzada, F. (2016). The Long Term Frechet distribution: Estimation, Properties and its Applications. *Institute of Mathematical Science and Computing, University of Sao Paulo*

[16] Pericchi, L., and Coles, S. (2003). Anticipating catastrophes through extreme value modeling. *Appl. Statist.* 52: 405 - 416.

[17] Phien, N.H., and Fang, E.T. (1989). Parameter estimation for the general extreme value distribution, 15(2).

[18] Subh, S.A., and Alodat, M.T. (2017). Discrimination between Logistic and Gumbel distribution. *International Journal of Sciences: Basic and Applied Research*, 36(3): 244 - 255.

[19] Teamah, A.M.A., Elbanna, A.A., and Gemeay, A.M. (2020). Frechet-Weibull distribution with applications to earthquakes data sets. *Pak. J. Statist*, 36(2): 135 - 147.

[20] Zheng, S., (2018). Maximum Likelihood Estimation. *Statistical theory II*..