ASYMPTOTIC PROPERTIES OF A THREE PARAMETERS GUMBEL DISTRIBUTION ESTIMATORS USING SIMULATED DATA

Abstract: Modeling extreme value theories is really gaining interest in the world with scientist working to improve the flexibility of the distributions by adding parameter(s). Extreme value distributions are always described to include families of Gumbel, Weibull and Frechet distributions. Of the three distributions, Gumbel distribution is the most commonly used in the extreme value theory analysis. Existing literature has shown that the addition of parameter to a distribution makes it robust and/or more flexible hence the study introduced a new three parameters Gumbel distribution using the Marshall and Olkin proposed method for introduucing a new estimator/parameter to an existing distribution. The developed distribution will be important to the applications in some life time studies like high temperature,earthquakes, network designs, horse racing, queues in supermarket, insurance, winds, risk management, ozone concentration, flood, engineering and financial concepts. The parameters for the developed distribution was estimated using Maximum Likelihood Estimation method and their asymptotic properties investigated using simulated data. The maximum likelihood estimates for the three parameters namely shape, location and dispersion are unbiased, efficient, sufficient and consistent as the sample size become large, making the distribution more flexible and better for application. The three parameters Gumbel distribution can be used in modeling and analysis of normal data, skewed data and extreme data since it will provide efficient, sufficient and consistent estimates more so as the sample size approaches infinity.

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

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

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. 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.

3. Properties of the Estimators using simulated data

This section discusses 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 subsection (3.2) 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 (3.3) to investigate the precision and reliability of the estimators as we increase the sample size. Finally, in subsection (3.4), we compare the asymptotic relationship between a three parameters Gumbel distribution and other distributions like normal, chisquare and Weibull respectively using simulated data sets of different sample sizes.

3.1. 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 on sub-sub-sections 3.1.1, 3.1.2 and 3.1.3 respectively.

3.1.1. Asymptotic biasedness for ω

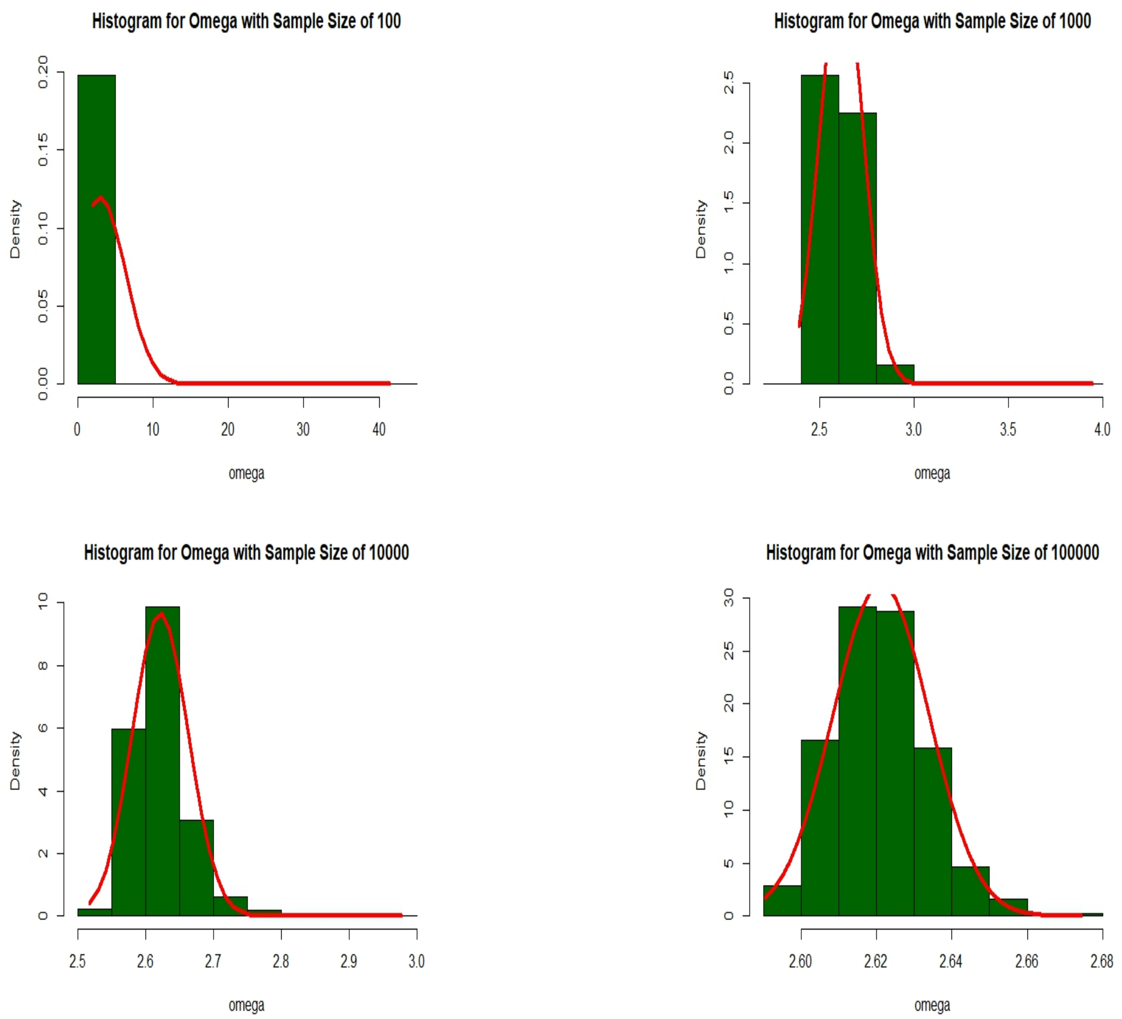

Figure 1. *Asymptotic distributions for* ω

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, $bias(\hat{\omega}, \omega) \rightarrow 0$ as $k \rightarrow \infty$ (that is, as k, the sample size becomes larger and larger).

3.1.2. 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, $bias(\hat{\delta}, \delta) \rightarrow 0$ as $k \rightarrow \infty$.

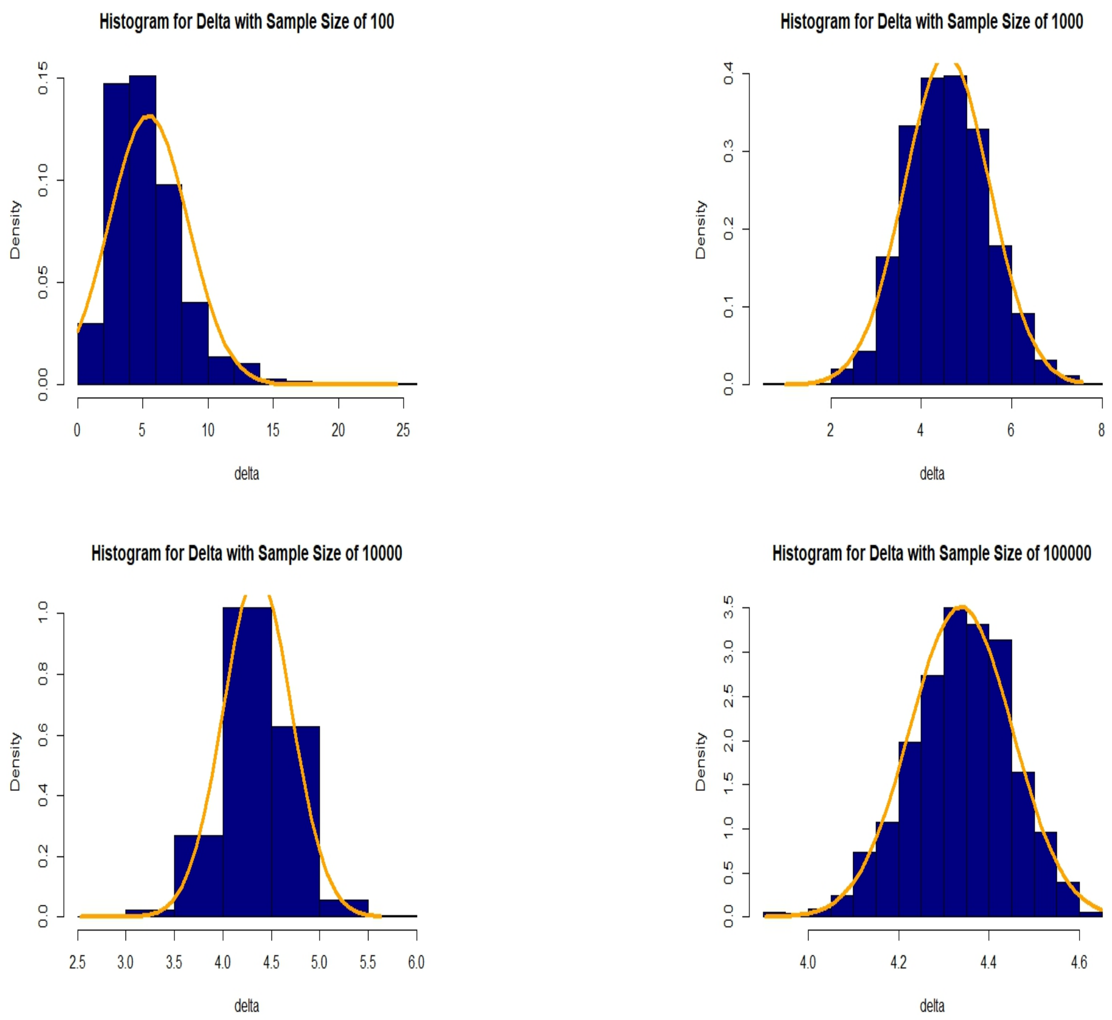

Figure 2. *Asymptotic distributions for* δ

3.1.3. Asymptotic biasedness for τ

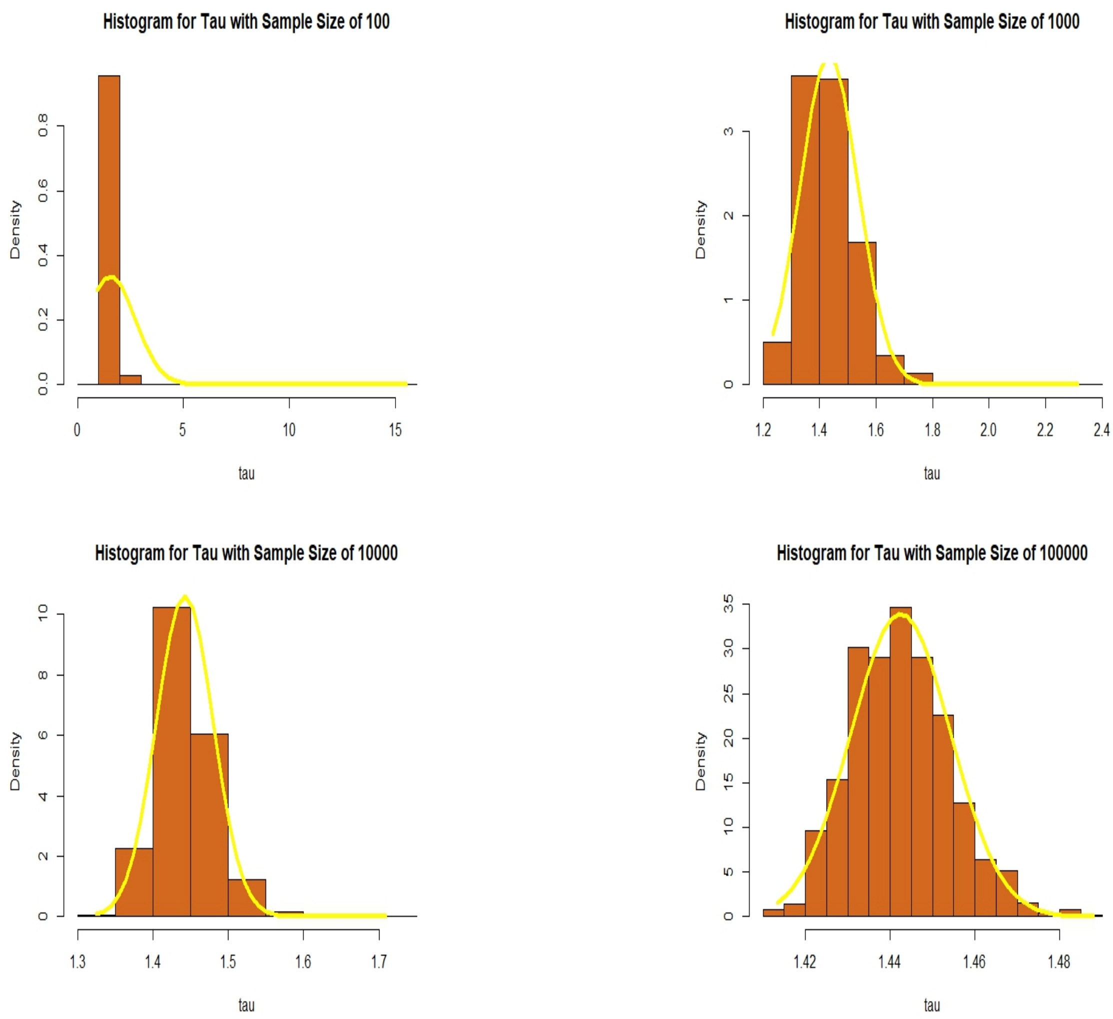

Figure 3. *Asymptotic distributions for τ*

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 $bias(\hat{\tau}, \tau) \to 0$ as $k \to \infty$.

3.2. Mean Square Error

This section 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.

3.3. Consistency

In this section, 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. 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. Section 3.1 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 in sub-sub-sections 3.3.1, 3.3.2 and 3.3.3 respectively.

3.3.1. 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

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

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.

3.3.2. 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.

3.3.3. 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.

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

This section 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 sub-sub-sections 3.4.1, 3.4.2 and 3.4.3 respectively.

3.4.1. 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.

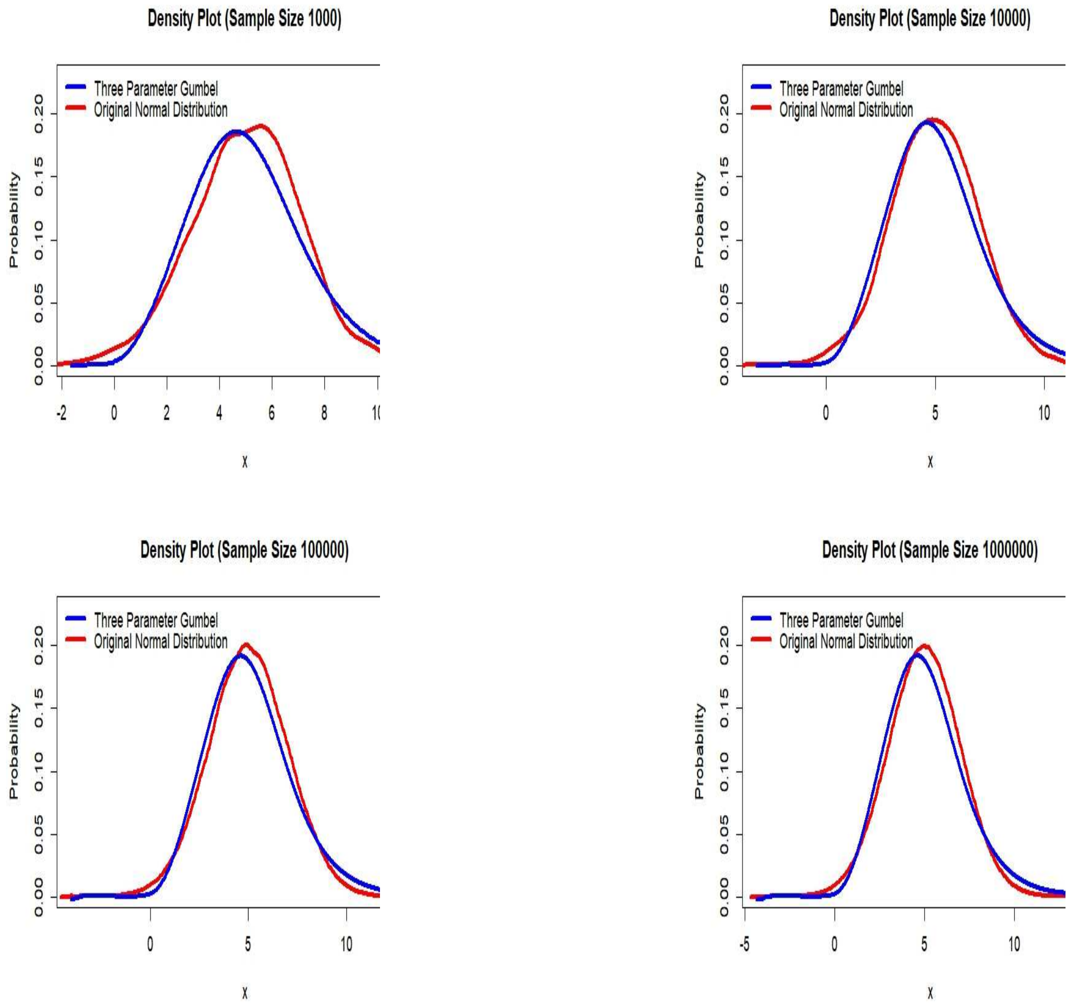

Figure 4. *Asymptotic distributions for three parameters Gumbel and normal distributions*

3.4.2. 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.

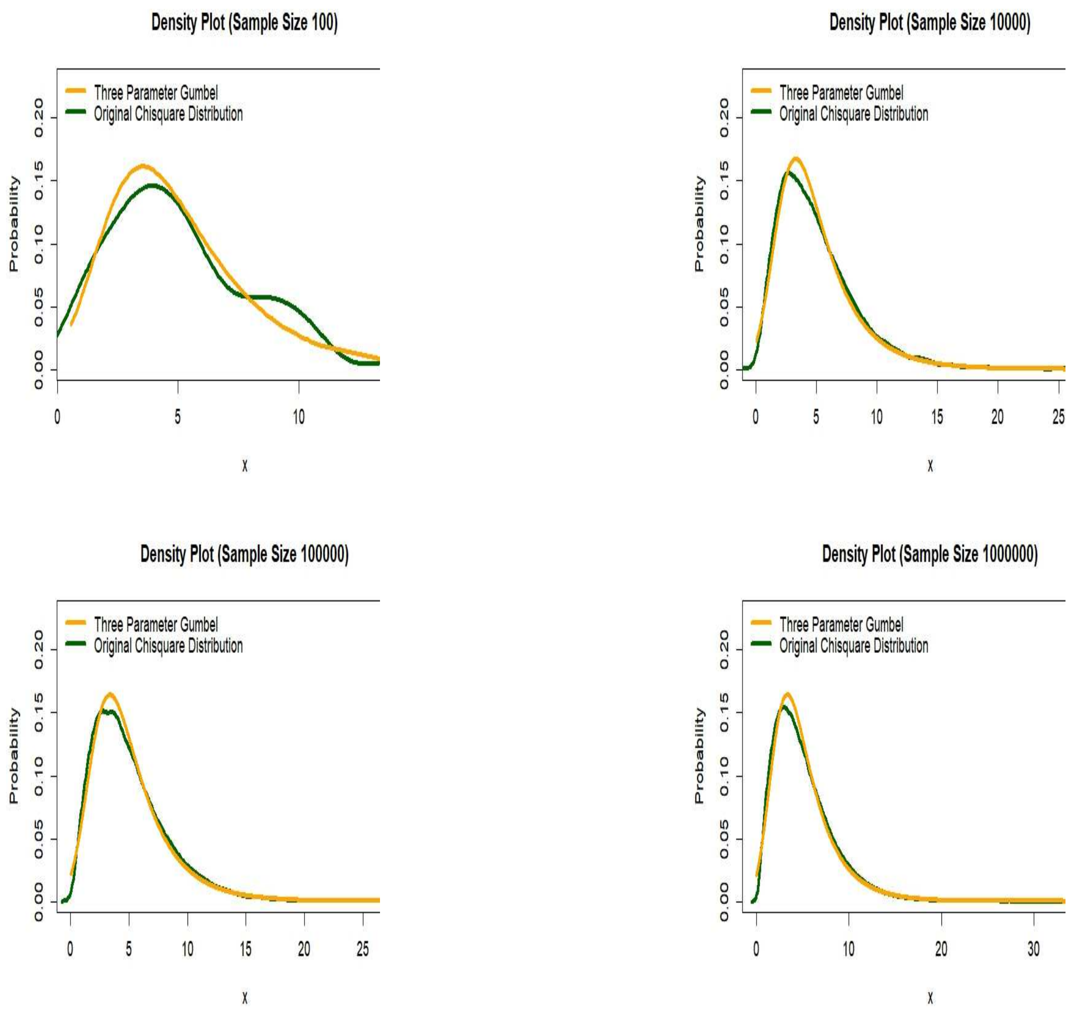

Figure 5. *Asymptotic distributions for three parameters Gumbel and chi-square distributions*

3.4.3. 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.

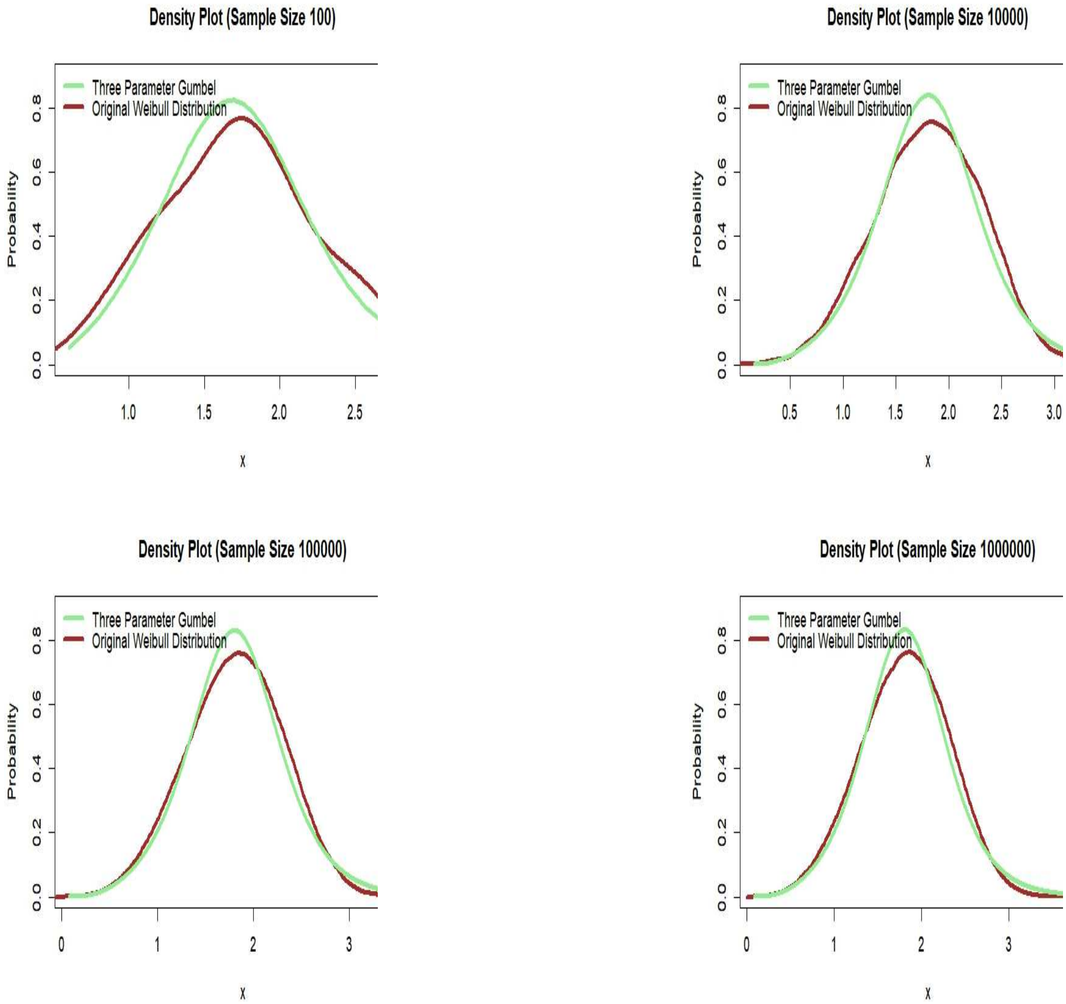

Figure 6. *Asymptotic distributions for three parameters Gumbel and Weibull distributions*

4. Conclusion

This study 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).
Introducing a new parameter to an existing distribution makes it more flexible and robust for application. In this research, a new parameter called the shape parameter was introduced to the existing two parameter Gumbel distribution. The introduced three parameters Gumbel distribution is a probability distribution function which can be used in modelling statistical data since it is more flexible. The maximum likelihood estimates for the three parameters namely shape (δ), location (ω) and dispersion (τ) are efficient, sufficient and consistent and this makes the function more flexible and better for application

5. Recommendation

The three parameters Gumbel distribution can be used in modeling and analysis of normal data, skewed data and extreme data since it will provide efficient, sufficient and consistent estimates. For the purpose of future improvement on a three parameters Gumbel distribution, future researches can investigate the behaviour of its location parameter because it can be evidenced from figure 4,5, and 6 that the measurer of location for a three parameters Gumbel distribution is showing a slight difference with the location for normal, Weibull and Chi-square distributions.

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*..