IPS Journal of Physical Sciences
ISSN: 3115-5715
IPS J. Phys. Sci., 3(1): 160-165 (2026)
DOI: https://doi.org/10.54117/ijps.v3i1.26

Fitting a Gamma Distribution to the Weight of Akwa Ibom State University Students Using Chi – Square Method

Itoro T. Michael[1], Ukeme P. Akra[1], Efiong A. Adakama[2], Anietie I. Akpan[1], Ndiana J. Sam[1]

[1]Department of Statistics, Akwa Ibom State University, Nigeria.
[2]Department of General Studies, Maritime Academy of Nigeria, Oron.

*Corresponding author- itoromichael@aksu.edu.ng

Abstract

This paper fits a gamma distribution to the weights of Students of Akwa Ibom State University using chi-square technique. The weights of 617 students were collected from the Medical Centre of the Akwa Ibom State University, Main Campus, Ikot Akpaden, Akwa Ibom State. A chi-square test is used to ascertain whether or not the weights of Students are gamma distributed. The gamma distribution parameters were estimated using maximum likelihood approach. The graphical displays of the simulated and real data with the same parameter value are presented. It was observed from the results that the gamma distribution fits the weights of students of the Akwa Ibom State University at the significance level $\alpha = 0.91\%$. The graphs of the weights of students, the simulated weights of students and the gamma densities values also showed a great parity.

Keywords: Chi-square Test; Gamma Distribution; Weights of Students; Maximum Likelihood Estimates.

Article History
Received: 26 Mar 2026
Accepted: 30 Apr 2026
Published: 18 May 2026

Scan QR code to view•

How to cite this paper: Michael, I. T., Akra, U. P., Adakama, E. A., Akpan, A. I., & Sam, N. J. (2026). Fitting a Gamma Distribution to the Weights of Akwa Ibom State University Students Using Chi-Square Method. *IPS Journal of Physical Sciences*, *3*(1), 160–165. https://doi.org/10.54117/ijps.v3i1.26

1. Introduction

In statistics, several probability distributions exist for modelling real world phenomenon. Among the distributions are the gamma, Weibull, Normal, Pareto, Logistic, exponential distributions, these distributions are inherent with so much information about real data sets obtained from all areas of studies; Engineering, Medical, Sciences, art etc and to aid in extracting useful information from the datasets. The data sets possess varieties of properties such as skewness, unimodality, bimodality and a couple of others.

Gamma distribution is a well-known probability model in statistics with the scale and shape parameters and very useful in modelling skewed, non-negative data like lifespan, insurance claims and product of the two parameters gives the mean. The weight is a non-negative real data set. Michael et al. (2017, 2022, 2025) and Michael and Iseh (2025) fitted the normal, log-normal and logistic distributions to the weights and heights of students of the Akwa Ibom State University using the Chi-square approach by splitting the students' weights into different cells to obtain the observed values and using the raw data for the maximum likelihood estimation of normal and logistic models' parameters; the mean and standard deviation, thereafter, calculating the cells probability and the chi-squared value. Whereas, the normal and log-normal fitted, the logistic distribution did not fit the heights of students even though logistic shares a similarity to the normal distribution.

It was useful to carry out a goodness of fit tests to indicate whether or not it is reasonable to assume that a random sample comes from a specific probability distribution. Any probability model that does not fits a data set appropriately if used for further statistical analysis can give misleading and non-reliable results. According to Michael et al. (2019), measures of goodness of fit typically summarize the discrepancy between observed and expected values under the model considered and such measures can be used in statistical hypothesis testing to test for; normality of residuals, whether two samples are drawn from identical distributions or whether outcome frequencies follow a specified distribution and others.

Authors like Anderson and Dar-ling (1952) introduced the Anderson-Darling test, a statistical test of whether a given sample data is drawn from a given probability distribution with no parameter to be estimated. Shapiro and Wilk (1965) introduced the Shapiro-Wilk test to test the null hypothesis that the random samples constituting a random variable comes from a normally distributed population. D'Agostino (1970) introduced the D'Agostino's K2 test, a goodness of fit measure of departure from normality; the test aims to establish whether or not the given sample comes from a normally distributed population. Pearson (1900) investigated the properties of Pearson's chi-squared test. Pearson chi-squared test tests a null hypothesis that the frequency distribution of certain events observed in a sample is consistent with a particular theoretical distribution. Lilliefors (1967) introduced the Lilliefors test, a normality test based on the Kolmogorov-Smirnov test. It is used to test the null hypothesis that data come from a normally distributed population, when the null hypothesis does not specify which normal distribution.

This work fits the gamma distribution to the weights of Akwa Ibom State University Students using the Chi-Squared test and graphical method. The weights of 617 students of the Akwa Ibom State University were collected from the Medical Centre, Main Campus, Ikot Akpaden, Akwa Ibom State, Nigeria.

2. Methodology

This paper uses two methods for testing or verifying if gamma distribution fits the weights of Akwa Ibom State University Students; the graphical method and the chi-square methods.

2.1 The graphical method

The graphical methods involve the use of graphical tools to display box plots, histogram and density plot of the given data sets and comparing same with that of the theoretical distribution. In this research work, we display the gamma density plots for the raw and the simulated datasets and the weights of students of the Akwa Ibom State University

2.2 The chi-squared method

According to Wackerly et al. (2008), Karl Person in 1900 proposed the following test statistics, which is a function of the deviations of the observed counts from their expected values, weighted by the reciprocals of their expected values. Thus,

$$\chi^2_{k-1} = \sum_{i=1}^{k} \frac{[n_i - E(n_i)]^2}{E(n_i)} = \sum_{i=1}^{k} \frac{[n_i - np_i]^2}{np_i} \quad (1)$$

called the Pearson chi-square test and denoted by χ^2_{k-1} with $k-1$ degrees of freedom.

Where:

$\mathrm{n_i}$ = an observed frequency (i.e. count) for $\mathrm{n_i}$

$\mathrm{E(n_i)} = \mathrm{np_i}$ an expected frequency for $\mathrm{n_i}$ asserted by the null hypothesis.

n = the sample size

2.3 The Gamma Probability Distribution Model

According to Hogg et al. (2013), a random variable X is said to have a gamma distribution with parameter α *and* β if and only if the density function of X is

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}} & ,0 < x < \infty, 0 < \alpha, \beta \\ 0, & otherwise \end{cases} \quad (2)$$

Consider a random sample $X_1, X_2, \ldots, X_k$ of size n from this distribution and If we let denote the frequency of n_i , i = 1,2,3, … , k, so that $\mathrm{n_1} + \mathrm{n_2} + \cdots + \mathrm{n}_k = n$, then the random χ^2_{k-1} variable in (1) cannot be computed once $\mathrm{n_1}, \mathrm{n_2}, \cdots, \mathrm{n}_k$ have been observed, since each pi , and hence χ^2_{k-1}, is a function of α *and* β. The value of α *and* β that minimize χ^2_{k-1} are difficult to compute therefore, their maximum likelihood estimates are used to evaluate pi and χ^2_{k-1}. Using maximum likelihood estimates of the parameters in place of minimum chi-square estimates tend to lead to the rejection of the null hypothesis since the χ^2_{k-1} value is not minimized by maximum likelihood estimates, and as such the computed value is somewhat greater than it would be if minimum chi-square estimates are used.

2.4 Research hypothesis

The Null hypothesis ($\mathrm{H_0}$): The weight of students follows a gamma distribution.

The Alternative Hypothesis ($\mathrm{H_1}$): The weight of students does not follow a gamma distribution.

2.5 Log-Likelihood function of the Gamma Model Parameters

The likelihood function of the gamma distribution is defined as

$$L = \prod_{i=1}^{\mathrm{n}} f(x_i; \alpha, \beta)$$

And the log-likelihood function is given as

$$\ell = \sum_{i=1}^{\mathrm{n}} \log(f(x_i; \alpha, \beta))$$

by applying equations (2) and (3) to one, we have the likelihood function of the gamma distribution as

$$L = \prod_{i=1}^{\mathrm{n}} \left(\frac{1}{\Gamma(\alpha)\beta^{\alpha}} x_i^{\alpha-1} e^{-\frac{x_i}{\beta}} \right)$$

and the log-likelihood function of the gamma distribution is

$$\ell(\alpha, \beta) = -\sum_{i=1}^{\mathrm{n}} \left(-\alpha \log(\beta) - \log\big(\Gamma(\alpha)\big) + (\alpha - 1)\log(x_i) - x_i/\beta\right) \quad (3)$$

2.6 Estimation of the Gamma Model Parameters Using MaxLik package in R

Henningsen and Toomet (2009) introduced the maxLik package in R for maximum likelihood estimation of a model's parameter. The package is very helpful in estimating the parameters of a given distribution and a data set. Micheal et al. (2025a) and Michael et al (2025b) used the maxLik package in R for estimating the parameters of the gamma-Weibull and the two-parameter Burr Type (X) distributions. In this paper will make use of the maxLik package for estimating the gamma distribution parameters.

```
X=WEIGHT of Students
library(maxLik)
loggamma<-function(statistics){
alpha<-statistics[1]
beta<-statistics[2]
sum(-(alpha*log(beta))-log(gamma(alpha))+((alpha-1)*log(X))-(X/beta))
}
mlegamma1<-maxLik(logLik=loggamma, start=c(alpha=20, beta=2))
summary(mlegamma1)
coef(mlegamma1)
alpha<-coef(mlegamma1)[1] #α= 44.0188
beta<-coef(mlegamma1)[2] # β=1.3353
```

3. Results and Discussion

3.1 Graphical displays

Figure 1 is a plot of the simulated data obtained from the gamma distribution for a sample 617 with parameters $\hat{\alpha} = 44.0188, \hat{\beta} = 1.3353$. The graph of the simulated data points resembles a bell-shaped which shares similarity with that of the normal distribution

Figure 2 is a plot of the gamma distribution density values and the weights of students obtained from the Medical Centre of the Akwa Ibom State University. The graph shows a great parity with that of the simulated indicating that the data follows the gamma probability distribution.

```
y<-((X^(alpha-1))*exp(-X/beta))/((beta^(alpha))*gamma(alpha))
plot(X,y,main="Gamma Density Plot of the Weights of Students",
ylab="Gamma Density",xlab=" Weights of Students")
```

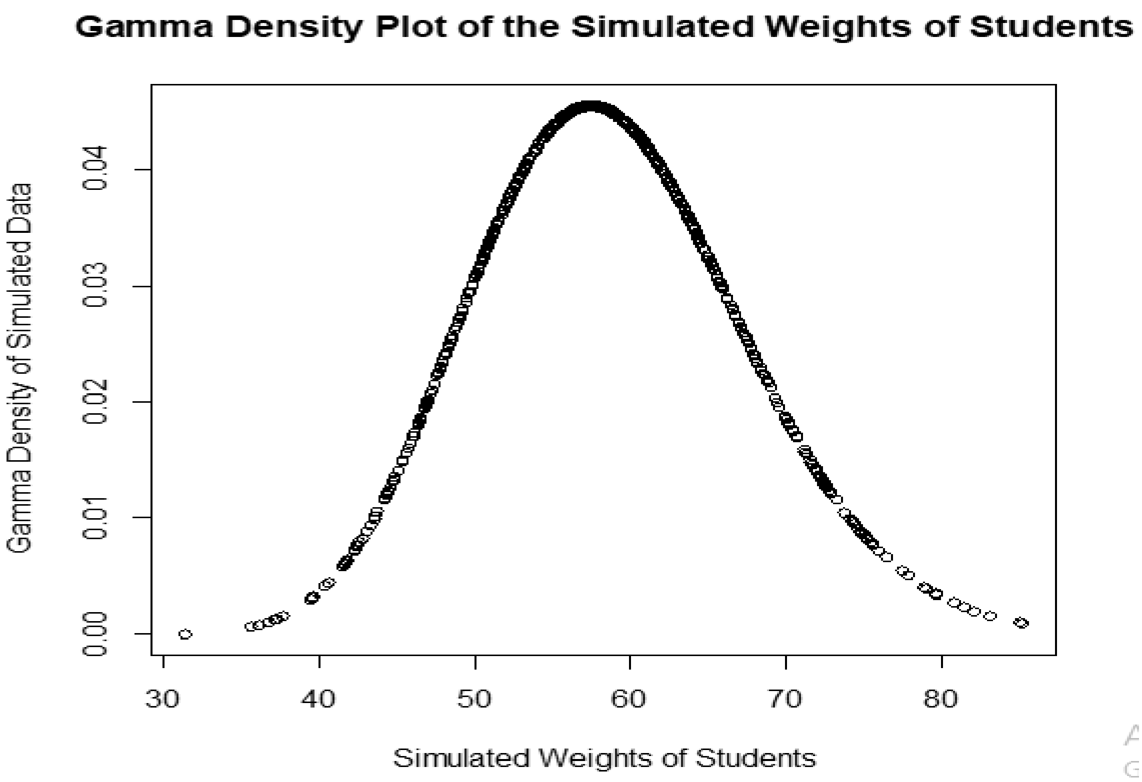

Figure 1: Gamma Density Plot of the Simulated Weights of Students

```
X1<-rgamma(617,alpha,scale=beta)
y1<-((X1^(alpha-1))*exp(-X1/beta))/((beta^(alpha))*gamma(alpha))
plot(X1,y1,main="Gamma Density Plot of the Simulated Weights of Students",
ylab="Gamma Density of Simulated Data",xlab="Simulated Weights of Students")
```

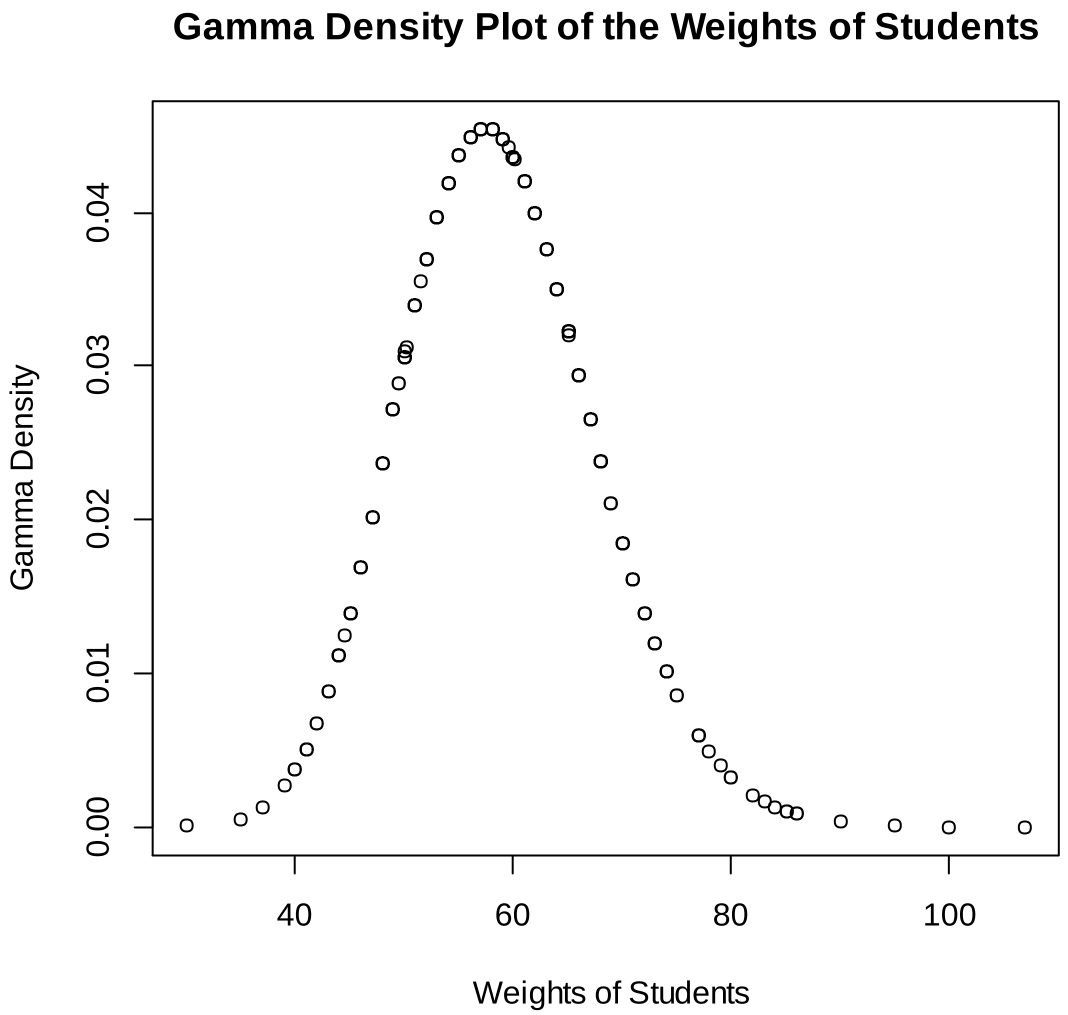

Figure 2: Gamma Density Plot of the Weights of Students

3.2 Computation of the Gamma Distribution Cells Probabilities

The random variable X, denoting the weights of students is partitioned into the following $k = 10$ mutually disjoint sets: $A_1 = \{0 < X \leq 40\}, A_2 = \{40 < X \leq 45\}, A_3 = \{45 < X \leq 50\}, A_4 = \{50 < X \leq 55\}, A_5 = \{55 < X \leq 60\}, A_6 = \{60 < X \leq 65\}, A_7 = \{65 < X \leq 70\}, A_8 = \{70 < X \leq 75\}, A_9 = \{75 < X \leq 80\}, A_{10} = \{80 < X \leq 85\}, A_{11} = \{85 < X \leq \infty\}$

Let $p(A_i) = p_i, i = 1,2, \ldots, k$, where p_i is the probability that the outcome of the random experiment is an element of the set A_i from the gamma probability distribution. The probabilities are obtained as follows:

$$p_i = \int_a^{\mathrm{b}} \left(\frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-\frac{x_i}{\beta}} \right) dx \text{ , } i = 1,2,3, \ldots .,10 \tag{4}$$

Where a and b are the lower and upper limits for each M_i, $i = 1,2,3, \ldots .,10$

The Table 1 shows the calculated probabilities, Observed Frequencies and Expected Frequencies of the Weights of Students and the gamma Distribution obtained from (4)

Table 1. The calculated probabilities, Observed Frequencies and Expected Frequencies

Cells(i)	Sets (A_i)	Observed Frequencies (X_i)	Probabilities (p_i)	Expected Frequencies (np_i)
1	(0, 40]	7	0.0094	5.7998
2	(40,45]	28	0.0405	24.9885
3	(45,50]	66	0.1102	67.9934
4	(50,55]	129	0.1899	117.1683
5	(55,60]	146	0.2243	138.3931
6	(60,65]	122	0.1927	118.8959
7	(65,70]	70	0.1262	77.8654
8	(70,75]	28	0.0654	40.3518
9	(75,80]	10	0.0277	17.0909
10	(80,85]	5	0.0098	6.0466
11	(85,∞)	6	0.004	2.468

The Test Statistic

$$\chi^2_{k-2} = \sum_{i=1}^{n} \frac{(n_i - np_i)^2}{np_i} \quad (5)$$

The test statistic in (5) where n_i and np_i denote the observed and expected frequencies respectively with $k - 2 =$ the degree of freedom.

Table 2: Computed Chi-square Value

Cells (i)	Observed Frequencies (X_i)	Expected Frequencies (np_i)	$X_i - np_i$	$(X_i - np_i)^2$	$\frac{(X_i - np_i)^2}{np_i}$
1	7	5.7998	1.2002	1.4405	0.24837
2	28	24.9885	3.0115	9.0691	0.36293
3	66	67.9934	-1.9934	3.9736	0.05844
4	129	117.1683	11.8317	139.9891	1.19477
5	146	138.3931	7.6069	57.8649	0.41812
6	122	118.8959	3.1041	9.6354	0.08104
7	70	77.8654	-7.8654	61.8645	0.79451
8	28	40.3518	-12.3518	152.5670	3.78092
9	10	17.0909	-7.0909	50.2809	2.94197
10	5	6.0466	-1.0466	1.0954	0.18115
11	6	2.468	3.532	12.4750	5.05471
Total					15.11693

$$\chi^2_{k-2} = \sum_{i=1}^{n} \frac{(n_i - np_i)^2}{np_i} = 15.11693 \quad (6)$$

3.3 Significant levels and critical values

The degree of freedom $(df) = k - s - 1 = 11 - 2 - 1 = 8$, where k is the number of cells and s, the number of parameters estimated. Some significance levels and their corresponding critical values are presented in Table 3.

Table 3: Significance Levels and Corresponding Critical Values for $df = 8$

Significance Level	Critical Values	Degree of Freedom
0.0001	31.82763	8
0.0011	25.88259	8
0.0021	24.22610	8
0.0031	23.21408	8
0.0041	22.48008	8
0.0051	21.90236	8
0.0061	21.42507	8
0.0071	21.01791	8
0.0081	20.66254	8
0.0091	20.34700	8

The Decision Rule

Reject H_o if $\chi^2_{k-2} > \chi^2_{crit}$, where χ^2_{k-2} is the computed value of the test statistic and χ^2_{crit} is the critical value as given in table 3.

4. Conclusion

It is observed from Tables 2 & 3 that $\chi^2_{k-2} = 15.11693 < \chi^2_{crit} = 20.347$ when the significance level $\alpha \geq 0.0091$ or 0.91% . Hence, the weights of students of Akwa Ibom State University follows a gamma distribution using the chi-squared test. This result is also supported by the graphs of the simulated and real data obtained from the Medical Centre of the Akwa Ibom State University.

Acknowledgements

The first author appreciates the Director of Health Services, Dr Geoffrey C. Udoh and the Head of Medical Laboratory Unit, Mr. Aniekan B. Okure of the Akwa Ibom state University Medical Centre for their help during the data collection process from the Medical Centre for this study.

References

[1] Shapiro, S. S. & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika.* 52 (3–4): 591–611. https://doi.org/10.1093/biomet/52.3-4.591.

[2] Michael, I. T., Usoro,A. E., Ikpang, I. N., George, E. U. & Bassey, D. I.(2019). Fitting a Gamma Distribution using a chi-squared approach to the Heights of students of Akwa Ibom State University, Nigeria. *International Journal of Advanced Statistics and Probability* 7(2) (2019) 42 – 46.

[3] Michael, I.T., Isaac, A. A. and Etim, A. Fitting a Normal Distribution to the Heights of Akwa Ibom State University Students Using Chi-square. Asian Journal of Probability 27(2) (2025), 42-49

[4] D’Agostino, R. B. (1970). Transformation to normality of the null distribution of g1. *Biometrika.* 57 (3): 679–681. JSTOR 2334794 https://doi.org/10.1093/biomet/57.3.679.

[5] Anderson, T. W. & Darling, D.A (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Annals of Mathe-matical Statistics.* 23: 193–212. https://doi.org/10.1214/aoms/1177729437.

[6] Henningsen, A. & Toomet, O. (2009). MaxLik: Tools for Maximum Likelihood Estimation. R package version 0.5, Retrieved September 10, 2017 from *http://CRAN.R-project.org.*

[7] Hogg, R. V., McKean, J. W. & Craig, A. T. (2013). *Introduction to Mathematical Statistics, 7th Ed.*, Boston: Pearson Education, Inc.

[8] Wackerly, D. D., Mendenhall, W. & Scheaffer, L. R. (2008). Mathematical statistics with applications, 7th Ed., USA: Thomson Higher Education, Inc.

[9] Michael, I. T., Ikpang, I. N. & Isaac, A. A (2017). Goodness of fit test: a chi-squared approach to fitting of a normal distribution to the weights of students of Akwa Ibom State University, Nigeria. *Asian Journal of Natural and Applied Sciences* 6(4) 107 – 113.

[10] Lilliefors, H. (1967). On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Asso-ciation*. 62; 399–402. https://doi.org/10.1080/01621459.1967.10482916.

[11] Pearson, K. (1990). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*. 50(302): 157-175. https://doi.org/10.1080/14786440009463897.

[12] Michael, I.T., Iseh, M.J., Ikpang, I. N, George, E. U. & BASSEY, D. I. A. Goodness of fit test: Using chi-squared approach to fitting of a log-normal probability model to the weights of students of Akwa Ibom State University, Nigeria. *International Journal of Applied Science a*nd Research 4(6) (2021), 146 – 152.

[13] Michael, I. T and Iseh, M. J. Utilizing the Chi-square Technique in Fitting a Logistic Distribution to the Heights of Students of Akwa Ibom State University. SCIREA JOURNAL OF MATHEMATICS 10(2) (2025), 29-39

[14] Michael, I. T., Iseh, M. J., and Effiong, U. A. (2025). Comparing Two-parameter Burr Type (X) and Gamma-Weibull Distributions Using Information Criteria for the Heights of Akwa Ibom State University Students. Science Journal of Applied Mathematics and Statistics 13(2)(2025), 56-62. Science Publishing Group.

Michael, I. T., Anthony, U. and Iseh, J. Matthew. Non-nested Test Statistic for Comparison of the Two-Parameter Burr Type(X) and Gamma-Weibull Distributions with Application to Heights of Students of Akwa Ibom State University. Asian Journal of Probability and Statistics 27(6), 122-137.

• *Thank you for publishing with us.*