African Journal of Mathematics and Statistics Studies

ISSN: 2689-5323

Volume 6, Issue 4, 2023 (pp. 77-80)

www.abjournals.org

APPLICATION OF THREE PROBABILITY DISTRIBUTIONS TO JUSTIFY CENTRAL LIMIT THEOREM

Okoro Ifeanyichukwu*, Uka Christian O. and Ogbara Chidi Obed

Department of Statistics, Federal College of Agriculture Ishiagu, Ebonyi State, Nigeria

*Corresponding Email: okoroifeanyichukwu6f@gmail.com; Phone: +2347065192515,

Cite this article:

Okoro I., Uka C.O., Ogbara C.O. (2023), Application of Three Probability Distributions to Justify Central Limit Theorem. African Journal of Mathematics and Statistics Studies 6(4), 77-80. DOI: 10.52589/AJMSS-LHCUQZLF

Manuscript History

Received: 29 Aug 2023

Accepted: 19 Oct 2023

Published: 22 NOv 2023

ABSTRACT: *This paper focused on the use of three probability distributions to justify the central limit theorem (CLT). The aim was to use the moment generating function (MGF) to prove (CLT) and also to portray the shape of different sample sizes (15, 30 and 100) of distributions of sample means on a histogram. The population distributions studied were: Normal, Gamma and Exponential distributions. In addition, sampling distribution of the mean table was constructed for better understanding of CLT. The study used simulation to simulate population distribution of Normal, Gamma and Exponential. Five hundred (500) distributions of sample means were drawn from each of the simulated population distributions at three different sample sizes (n): 15, 30 and 100. The shape of the simulated population distribution and sampling distribution of mean were presented on a histogram. The mean and standard deviation of each population distribution together with distribution of sample means at different sample sizes were also presented on the histogram plotted. The findings showed that under normal distribution, the sampling distribution of mean produced a shape like normal distribution irrespective of the sample size. Conversely, the shape of sampling distribution of mean under non-normal distributions gradually converges to normal distribution as sample size tends to infinity, while the variability of each sampling distribution decreases as the sample size increases. Therefore, CLT holds for large sample size ($n \geq 30$).*

KEYWORDS: Central limit theorem, sampling distribution of mean, normal distribution, exponential distribution, Gamma distribution, histogram.

Article DOI: 10.52589/AJMSS-LHCUQZLF

DOI URL: https://doi.org/10.52589/AJMSS-LHCUQZLF

INTRODUCTION

In the field of statistics, data type plays an important role in determining the validity of the statistical outcome. The data collected can follow different probability distributions such as Normal, Uniform, Binomial, Exponential, etc depending on the purpose of the study. However, the nature of the population distribution from which data is sourced may not be known by the researcher. In scenarios like this, the Central Limit Theorem (CLT) is applied to solve the problem of unknown population distribution in the circle of statistics. CLT was discovered by French mathematician Abraham de Moivre, who held that the binomial distribution approaches normal distribution as the number of trials increases (Winters & Kinneth, 2022). In other words, as the sample size (n) increases, the distribution of binomial converges to normal distribution. CLT is one of the most important theorems in statistics; it has many applications such as normality assumption in parametric hypothesis testing of mean. It is also suitable in measuring One Sample T Test, Two Samples T Test, ANOVA, Linear Regression, Confidence Interval Estimates, etc. CLT is also applied to all probability distributions except Cauchy distribution because of its infinite variance. Furthermore, it is applied to independent identically distributed variables, where the value of one observation does not depend on the value of another observation (Sutanapong & Louangrath, 2015). However, CLT depends largely on sampling distribution of mean for its proper application. Hence, this paper examines in detail the application of three probability distributions to justify central limit theorem (CLT).

Sampling Distribution of Mean

In statistics, population is the entire set of items or groups of individuals to be studied. In practice, studying all data from a population is impractical. Therefore, random samples are studied instead of the entire population. The samples are used to make statistical inferences about a population. Random sample is a randomly selected subset of a population. Sampling distribution of mean is the distribution of all possible mean values computed from each sample of the same size randomly drawn from the same population (Sawada, 2021). The distribution of sample mean is constructed using the following steps outlined: (a) Randomly draw all possible samples of size n from a finite population of size N (b) Compute the sample mean from each sample and (c) Present the distribution of sample means obtained in tabular format. The table below presents a constructed distribution of sample means for sample size m and number of sample n.

Table 1: Illustrate random sample (n) with sample size (m) and their corresponding sample means

	Random (1)	Random (2)	.	.	Random (m)	Sample Mean
Sample (1)	x_{11}	x_{12}	.	.	x_{1m}	$\underline{x}_1$
Sample (2)	x_{21}	x_{22}	.	.	x_{2m}	$\underline{x}_2$
.	.	.	.	.	.	
.	.	.	.	.	.	
.	.	.	.	.	.	
Sample (n)	x_{n1}	x_{n2}	.	.	x_{nm}	$\underline{x}_n$

Source: *Computed by the Researcher*

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 6, Issue 4, 2023 (pp. 77-80)
www.abjournals.org

When sampling is from a normally distributed population, the distribution of the sample mean has the following properties: (a) the distribution of sample mean will be normal (b) the mean of sample mean is equal to the population mean from which the samples were drawn (c) the variance of sample mean is equal to the population variance divided by sample size. Alternatively, when sampling is from a non-normally distributed population, Central Limit Theorem (CLT) is applied.

According to Islam and Mohammed (2018), the distribution of sample mean will follow normal distribution when sampling is from a normally distributed population or from a non-normally distributed population, provided that the sample is large. It is applied when sampling is from a population whose functional form is unknown, provided that sample size is large enough (n ≥ 30).

CENTRAL LIMIT THEOREM (CLT)

This theorem states that irrespective of population distribution from which samples were drawn, the distribution of sample means ($\underline{x}$) calculated from each sample will be normally distributed with mean (population mean) and variance ($\frac{population\ variance}{sample\ size\ (n)}$) provided that sample size is large (n ≥ 30) (Shige, 2019). The standardized variate for the sampling distribution of mean is given by:

$$Z = \frac{\underline{x} - mean}{\frac{variance}{(n)}} \quad (1.1)$$

where

$\underline{x}$ = sample mean

$mean$ = population mean

$variance$ = population variance

n = sample size

Mathematically, this theorem is represented below as:

z =) ~ N(0,1) (1.2)

That is, as $n \rightarrow \infty$ the distribution of the sample means approaches the Normal distribution with mean (0) and variance (1). This theorem provides an avenue to sample non-normally distributed populations with a guarantee of getting approximately the same results as would have been obtained if the population were normally distributed provided that sample size is large. However, for small sample sizes, the shape of the sampling distribution of mean is less than the parent population from which samples were drawn (Fatheddin, 2015). On the other hand, the shape looks more like a normal distribution as sample size gets larger. As an illustration of CLT, let x be a population distribution from Binomial with mean (np) and variance (npq), then the CLT states that the distribution of computed sample means gotten from Binomial distribution will be approximately Normal with mean (mean of Binomial

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

distribution) and variance ($\frac{variance\ of\ Binomial\ distribution}{sample\ size\ (n)}$) provided that sample size is large (n ≥ 30).

Population Distributions Studied

This section focuses on three population distributions used to justify CLT. The population distributions studied were presented in details below:

Normal Distribution

A normal distribution is known by its bell-like shape and it is symmetric about the mean. The mean, mode and median values of the normal distribution are equal. The probability density function (pdf) of Normal is given below:

$$f(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \quad (2.1)$$

where

x is the random variable, μ is the mean and σ is the standard deviation.

Gamma Distribution

Gamma distribution is defined by the probability density function (pdf) given below:

$$f(x) = \frac{\pi}{\Gamma(a)\theta^2} x^{a-1} e^{-\frac{x}{\theta}} \quad \text{if } 0 < x < \infty \quad (2.2)$$

where a $>$ 0 and $\theta > 0$, a is the shape parameter and θ is the scale parameter. The distribution is characterized by mean (a θ) and variance (a θ^2).

Exponential Distribution

Exponential distribution is known by the probability density function (pdf) given below:

$$f(x) = \lambda e^{-\lambda x} \quad \text{if } 0 < x < \infty \quad (2.3)$$

where $\lambda > 0$ is the rate parameter. The distribution is characterized by mean ($\frac{1}{\lambda}$) and variance ($\frac{1}{\lambda^2}$).

Proof of Central Limit Theorem (CLT) Using Normal Distribution

This is to prove that Moment Generating Function (MGF) of standardized sample mean (z) from normal distribution converges to MGF of standardized normal distribution ($e^{\frac{t^2}{2}}$) with mean (0) and variance (1) as n goes to infinity (n → ∞).

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 6, Issue 4, 2023 (pp. 77-80)
www.abjournals.org

That is $m_{\hat{z}}(t) \;= m_z(t)$ (2.4)

$$= e^{\frac{t^2}{2}} \quad (2.5)$$

where z ~ N (0,1)

Proof: Let $x_1,\ x_2,\ x_3, \ldots x_n$ be independent identically distributed (iid) normal distribution with mean (μ) and variance (σ^2).

Then,

Mean of sample mean ($\underline{x}$) is given by:

$$E(\underline{x}) = E(\frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}) \quad (2.6)$$

$$= \frac{1}{n}(E(x_1) + \ E(x_2) + \ E(x_3) + \ldots + E(\ x_n)) \quad (2.7)$$

$$= \frac{1}{n}(\ \mu + \ \mu + \ \mu + \ldots + \mu \quad (2.8)$$

$$E(\underline{x}) = \mu \quad (2.9)$$

Variance of sample mean ($\underline{x}$) is given by:

$$var(\underline{x}) = var(\frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}) \quad (2.10)$$

$$= \frac{1}{n^2}(var(x_1) + \ var(x_2) + \ var(x_3) + \ldots + var(\ x_n)) \quad (2.11)$$

$$= \frac{1}{n^2}(\ \sigma^2 + \ \sigma^2 + \ \sigma^2 + \ldots + \sigma^2) \quad (2.12)$$

$$var(\underline{x}) \ = \frac{\sigma^2}{n} \quad (2.13)$$

The standardized normal variate corresponding to the sample means (z) is defined as:

$$z = \frac{\underline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (2.14)$$

Simplification of z

$$z = \frac{\frac{\sum_{i=1}^{n} \ x_i}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (2.15)$$

$$= \frac{\sqrt{n}(\sum_{i=1}^{n} \ (x_i - n\mu)}{\sigma n} \quad (2.16)$$

Note $\sum_{i=1}^{n} \quad \mu = n\mu$ (2.17)

$$= \frac{\sqrt{n}\sum_{i=1}^{n} \ (x_i - \mu)}{\sigma n} \quad (2.18)$$

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

$= \frac{(\sum_{i=1}^{n} \quad (x_i - \mu)}{\sigma\sqrt{n}}$ (2.19)

Let y = $\frac{x_i - \mu}{\sigma\sqrt{n}}$ (2.20)

The MGF of y is given below:

$m_y(t) \ = \text{E}(\ e^{yt})$ (2.21)

$= \frac{\sum_{i=0}^{n} \quad E(y^i)t^i}{i!}$ (2.22)

$= \frac{\sum_{i=0}^{n} \quad E(x_i - \mu)^i}{i!} (\frac{t}{\sigma\sqrt{n}})^i$ (2.22)

$= \frac{E(x_0 - \mu)^0}{0!} (\frac{t}{\sigma\sqrt{n}})^0 + \frac{E(x_1 - \mu)^1}{1!} (\frac{t}{\sigma\sqrt{n}})^1 + \frac{E(x_2 - \mu)^2}{2!} (\frac{t}{\sigma\sqrt{n}})^2 + \ldots .$ (2.23)

Note that $(x - \mu)^1 = 0$ and $\text{E}(x - \mu)^2 = \sigma^2$ (2.24)

$= 1 + 0 + \frac{t^2}{2n} + \frac{\sum_{i=0}^{n} \quad E(x_3 - \mu)^3}{3!} (\frac{t}{\sigma\sqrt{n}})^3 + \ldots .$ (2.25)

Then the MGF of z variable is given by:

$m_z(t) \ = m_{\sum_{i=1}^{n} \quad (y_i)}(t)$ (2.26)

$= (m_y(t))^n$ (2.27)

Since y_i is iid normal distribution

$m_z(t) \ = (1 + \frac{t^2}{2n} + \frac{\sum_{i=3}^{n} \quad E(x_3 - \mu)^3}{3!} \left(\frac{t}{\sigma\sqrt{n}}\right)^3 + \ldots)^n$ (2.28)

The limit as n → ∞ on the both sides

$m_{\hat{z}}(t) \ = (1 + \frac{t^2}{2n} + \frac{\sum_{i=3}^{n} \quad E(x_3 - \mu)^3}{3!} (\frac{t}{\sigma\sqrt{n}})^3)^n$ (2.29)

$= e^{\frac{t^2}{2}}$

Hence, by the Uniqueness Theorem, the MGF of standardized sample mean (z) from normal distribution converges to MGF of standard normal distribution ($e^{\frac{t^2}{2}}$) with mean (0) and variance (1) as n goes to infinity (n → ∞).

Proof of Central Limit Theorem (CLT) Using Gamma Distribution

This is to prove that MGF of standardized sample mean (z) sample from Gamma distribution converges to MGF of standardized normal distribution ($e^{\frac{t^2}{2}}$) with mean (0) and variance (1) as n goes to infinity (n → ∞).

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

African Journal of Mathematics and Statistics Studies
ISSN: 2689-5323
Volume 6, Issue 4, 2023 (pp. 77-80)
www.abjournals.org

Proof: Let $x_1,\ x_2,\ x_3, . . . x_n$ be independent identically distributed (iid) Gamma distribution with mean $(\theta\beta)$ and variance $(\theta^2\beta)$, where (θ) and (β) are shape and scale parameters.

Then,

Mean of sample mean $(\underline{x})$ is given by:

$$E(\underline{x}) = E(\frac{x_1+x_2+x_3+...+x_n}{n}) \quad (2.30)$$

$$= \frac{1}{n}(E(x_1) + E(x_2) + E(x_3) + ... + E(x_n)) \quad (2.31)$$

$$= \frac{1}{n}(((\theta\beta) + (\theta\beta) + (\theta\beta) + ... + (\theta\beta) \quad (2.32)$$

$$E(\underline{x}) = (\theta\beta) \quad (2.33)$$

Variance of sample mean $(\underline{x})$ is given by:

$$var(\underline{x}) = var(\frac{x_1+x_2+x_3+...+x_n}{n}) \quad (2.34)$$

$$= \frac{1}{n^2}(var(x_1) + var(x_2) + var(x_3) + ... + var(x_n)) \quad (2.35)$$

$$= \frac{1}{n^2}((\theta^2\beta) + (\theta^2\beta) + (\theta^2\beta) + ... + (\theta^2\beta)) \quad (2.36)$$

$$var(\underline{x}) = \frac{(\theta^2\beta)}{n} \quad (2.37)$$

The standardized normal variate corresponding to the sample means (z) is defined as:

$$z = \frac{\underline{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \quad (2.38)$$

Simplification of z

$$z = \frac{\frac{\sum_{i=1}^{n} x_i}{n} - (\theta\beta)}{\frac{\frac{\theta\sqrt{\beta}}{n}}{\sqrt{n}}} \quad (2.29)$$

$$= \frac{\sqrt{n}(\sum_{i=1}^{n} (x_i - n(\theta\beta))}{n(} \quad (2.40)$$

$$\text{Note } \sum_{i=1}^{n} \mu = n\mu \quad (2.41)$$

$$= \frac{\sqrt{n}\sum_{i=1}^{n} (x_i - (\theta\beta))}{n(\theta\sqrt{\beta})} \quad (2.42)$$

$$= \frac{(\sum_{i=1}^{n} (x_i - (\theta\beta))}{\sqrt{n}(\theta\sqrt{\beta})} \quad (2.43)$$

$$\text{Let } y = \frac{x_i - (\theta\beta)}{\sqrt{n}(\theta\sqrt{\beta})} \quad (2.44)$$

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

The MGF of y is given below:

$$m_y(t) = \mathrm{E}(e^{yt}) \tag{2.45}$$

$$= \frac{\sum_{i=0}^{n} E(y^i)t^i}{i!} \tag{2.46}$$

$$= \frac{\sum_{i=0}^{n} E\left(x_i - (\theta\beta)\right)^i}{i!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^i \tag{2.47}$$

$$= \frac{E\left(x_0 - (\theta\beta)\right)^0}{0!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^0 + \frac{E\left(x_1 - (\theta\beta)\right)^1}{1!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^1 + \frac{E\left(x_2 - (\theta\beta)\right)^2}{2!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^2 + \dots.$$

(2.48)

Note that $(x - (\theta\beta))^1 = 0$ and $\mathrm{E}(x - (\theta\beta))^2 = \sigma^2$ (2.49)

$$= 1 + 0 + \frac{t^2}{2n} + \frac{\sum_{i=0}^{n} E\left(x_3 - (\theta\beta)\right)^3}{3!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^3 + \dots. \tag{2.50}$$

Then the MGF of z variable is given by:

$$m_z(t) = m_{\sum_{i=1}^{n} (y_i)}(t) \tag{2.51}$$

$$= (m_y(t))^n \tag{2.52}$$

Since y_i is iid Gamma distribution

$$m_z(t) = \left(1 + \frac{t^2}{2n} + \frac{\sum_{i=3}^{n} E\left(x_3 - (\theta\beta)\right)^3}{3!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^3 + \dots\right)^n \tag{2.53}$$

The limit as n → ∞ on the both sides

$$m_{\hat{z}}(t) = \left(1 + \frac{t^2}{2n} + \frac{\sum_{i=3}^{n} E\left(x_3 - (\theta\beta)\right)^3}{3!} \left(\frac{t}{\sqrt{n}(\theta\sqrt{\beta})}\right)^3\right)^n \tag{2.54}$$

$$= e^{\frac{t^2}{2}}$$

Hence, by the Uniqueness Theorem, the MGF of standardized sample mean (z) from gamma distribution converges to MGF of standard normal distribution ($e^{\frac{t^2}{2}}$) with mean (0) and variance (1) as n goes to infinity (n → ∞).

Proof of Central Limit Theorem (CLT) Using Exponential Distribution

This is to prove that Moment Generating Function (MGF) of standardized sample mean (z) from exponential distribution converges to MGF of standardized normal distribution ($e^{\frac{t^2}{2}}$) with mean (0) and variance (1) as n goes to infinity (n → ∞).

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

ISSN: 2689-5323
Volume 6, Issue 4, 2023 (pp. 77-80)

That is, $m_{\hat{z}}(t) \; = m_z(t)$ (2.55)

$$= e^{\frac{t^2}{2}}$$

where z ~ N (0,1)

Proof: Let $x_1, \; x_2, \; x_3, \; . \; . \; . \; x_n$ be independent identically distributed (iid) exponential distribution with mean $(\frac{1}{\lambda})$ and variance $(\frac{1}{\lambda^2})$.

Then,

Mean of sample mean ($\underline{x}$) is given by:

$$\mathrm{E}(\underline{x}) = \mathrm{E}(\frac{x_1 + x_2 + x_3 + . \, . \, . + x_n}{n}) \tag{2.51}$$

$$= \frac{1}{n}(E(x_1) + \; E(x_2) + \; E(x_3) + . \, . \, . + E(\, x_n)) \tag{2.52}$$

$$= \frac{1}{n}(\left(\frac{1}{\lambda}\right) + \left(\frac{1}{\lambda}\right) + \left(\frac{1}{\lambda}\right) + . \, . \, . + \left(\frac{1}{\lambda}\right)) \tag{2.53}$$

$$\mathrm{E}(\underline{x}) = (\frac{1}{\lambda}) \tag{2.54}$$

Variance of sample mean ($\underline{x}$) is given by:

$$\mathrm{var}(\underline{x}) = \mathrm{var}(\frac{x_1 + x_2 + x_3 + . \, . \, . + x_n}{n}) \tag{2.55}$$

$$= \frac{1}{n^2}(var(x_1) + \; var(x_2) + \; var(x_3) + . \, . \, . + var(\, x_n)) \tag{2.56}$$

$$= \frac{1}{n^2}(\, (\frac{1}{\lambda^2}) + \; (\frac{1}{\lambda^2}) + \; (\frac{1}{\lambda^2}) + . \, . \, . + (\frac{1}{\lambda^2})) \tag{2.57}$$

$$\mathrm{var}(\underline{x}) \; = (\frac{1}{\lambda^2}) \tag{2.58}$$

The standardized normal variate corresponding to the sample means (z) is defined as:

$$\mathrm{z} = \frac{\underline{x} - (\frac{1}{\lambda})}{\frac{(\frac{1}{\lambda})}{\sqrt{n}}} \tag{2.59}$$

Simplification of z

$$\mathrm{z} = \frac{\frac{\sum_{i=1}^{n} \; x_i}{n} - (\frac{1}{\lambda})}{\frac{(\frac{1}{\lambda})}{\sqrt{n}}} \tag{2.60}$$

$$= \frac{\sqrt{n}(\sum_{i=1}^{n} \; (x_i - n(\frac{1}{\lambda}))}{(\frac{1}{\lambda})n} \tag{2.61}$$

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

ISSN: 2689-5323
Volume 6, Issue 4, 2023 (pp. 77-80)

Note that $\sum_{i=1}^{n} \quad \mu = n\mu$ (2.62)

$$= \frac{\sqrt{n}\sum_{i=1}^{n} \quad (x_i - (\frac{1}{\lambda}))}{(\frac{1}{\lambda})n} \quad (2.63)$$

$$= \frac{(\sum_{i=1}^{n} \quad (x_i - (\frac{1}{\lambda}))}{(\frac{1}{\lambda})\sqrt{n}} \quad (2.64)$$

Let y = $\frac{x_i - (\frac{1}{\lambda})}{(\frac{1}{\lambda})\sqrt{n}}$ (2.65)

The MGF of y is given below:

$$m_y(t) \; = \mathrm{E}(\, e^{yt}) \quad (2.66)$$

$$= \frac{\sum_{i=0}^{n} \quad E(y^i)t^i}{i!} \quad (2.67)$$

$$= \frac{\sum_{i=0}^{n} \quad E\left(x_i - (\frac{1}{\lambda})\right)^i}{i!} (\frac{t}{(\frac{1}{\lambda})\sqrt{n}})^i \quad (2.68)$$

$$= \frac{E\left(x_0 - (\frac{1}{\lambda})\right)^0}{0!} (\frac{t}{(\frac{1}{\lambda})\sqrt{n}})^0 + \frac{E\left(x_1 - (\frac{1}{\lambda})\right)^1}{1!} (\frac{t}{(\frac{1}{\lambda})\sqrt{n}})^1 + \frac{E\left(x_2 - (\frac{1}{\lambda})\right)^2}{2!} (\frac{t}{(\frac{1}{\lambda})\sqrt{n}})^2 + \ldots. \quad (2.69)$$

Note that $(x \; - \; (\frac{1}{\lambda}))^1 = 0$ and $\mathrm{E}(x \; - \; (\frac{1}{\lambda}))^2 = (\frac{1}{\lambda^2})$ (2.70)

$$= 1 + 0 + \frac{t^2}{2n} + \frac{\sum_{i=0}^{n} \quad E\left(x_3 - (\frac{1}{\lambda})\right)^3}{3!} (\frac{t}{(\frac{1}{\lambda})\sqrt{n}})^3 + \ldots. \quad (2.71)$$

Then the MGF of z variable is given by:

$$m_z(t) \; = m_{\sum_{i=1}^{n} \quad (y_i)}(t) \quad (2.72)$$

$$= (m_y(t))^n$$

Since y_i is iid exponential distribution

$$m_z(t) \; = (1 + \frac{t^2}{2n} + \frac{\sum_{i=3}^{n} \quad E\left(x_3 - (\frac{1}{\lambda})\right)^3}{3!} \left(\frac{t}{(\frac{1}{\lambda})\sqrt{n}}\right)^3 + \ldots)^n \quad (2.73)$$

The limit as n → ∞ on the both sides

$$m_{\hat{z}}(t) \; = (1 + \frac{t^2}{2n} + \frac{\sum_{i=3}^{n} \quad E\left(x_3 - (\frac{1}{\lambda})\right)^3}{3!} (\frac{t}{(\frac{1}{\lambda})\sqrt{n}})^3)^n \quad (2.74)$$

$$= e^{\frac{t^2}{2}}$$

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

Hence, by the Uniqueness Theorem, the MGF of standardized sample mean (z) from exponential distribution converges to MGF of standard normal distribution ($e^{\frac{t^2}{2}}$) with mean (0) and variance (1) as n goes to infinity (n → ∞).

DATA ANALYSIS

In this section, R version 3.5.1 was used to simulate 500 distributions of sample means drawn from three population distributions: Normal, Gamma and exponential distributions, at three different sample sizes: 15, 30 and 100. From the simulated distribution of sample means, the histogram, mean and standard deviation were calculated. The histogram was used to represent the shape of sampling distribution of means at three different sample sizes. The mean and standard deviation were used to measure central tendency and variability for each sampling distribution of means. In addition, mean, standard deviation and histogram of population distribution was also presented to illustrate the effect of different sample sizes. The images below represent the histogram of three distributions under study.

Histogram of Normal Distribution

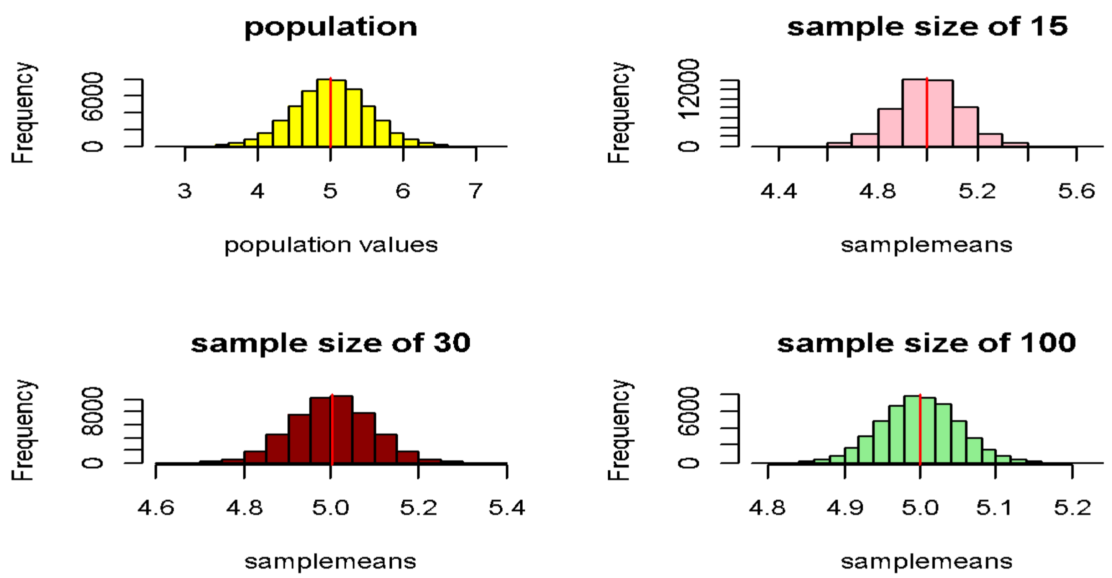

Figure 1: Histogram of population and sampling distribution of mean from Normal distribution

Source: *Computed by the Researcher*

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

Figure 1 presents the histogram, mean and standard deviation of both population distribution and sampling distribution of means. The vertical red line in the above plot indicates the mean and standard deviation of the distribution. The different shapes of histogram created above were obtained from stimulated 500 sample means at different sample sizes: 10, 30 and 100 drawn from Normal distribution.

Histogram of Exponential Distribution

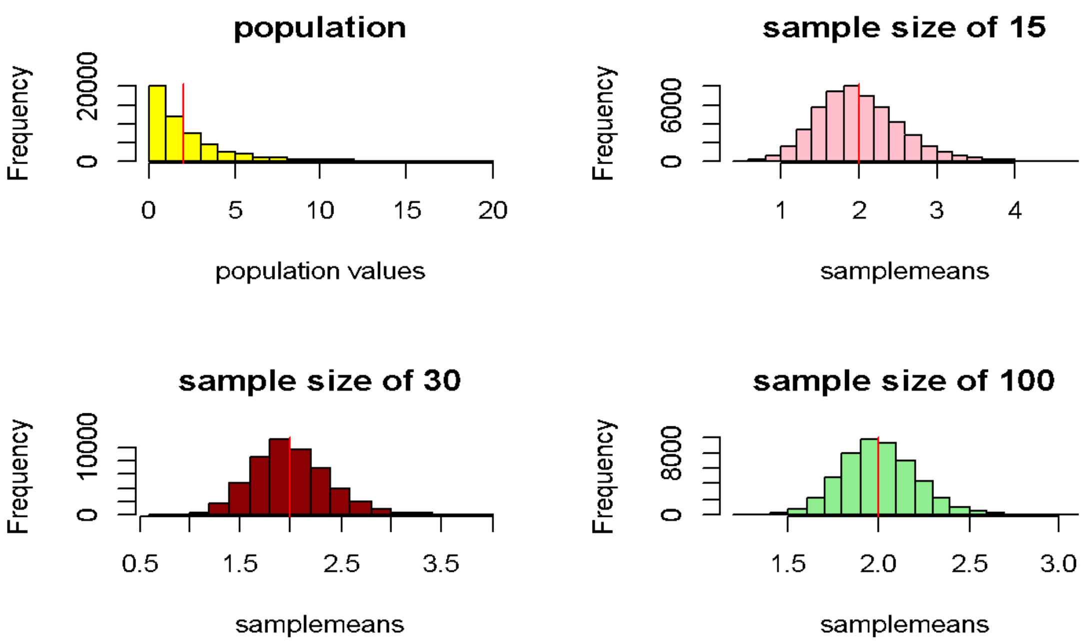

Figure 2: Histogram of population and sampling distribution of mean from exponential distribution

Source: *Plotted by the Researcher*

Figure 2 shows the histogram, mean and standard deviation of both population distribution and sampling distribution of means drawn from Exponential distribution with rate 0.5. The shapes of histogram portrayed above were obtained from stimulated 500 sample means of different sample sizes: 10, 30 and 100 drawn from Exponential distribution.

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

Histogram of Gamma Distribution

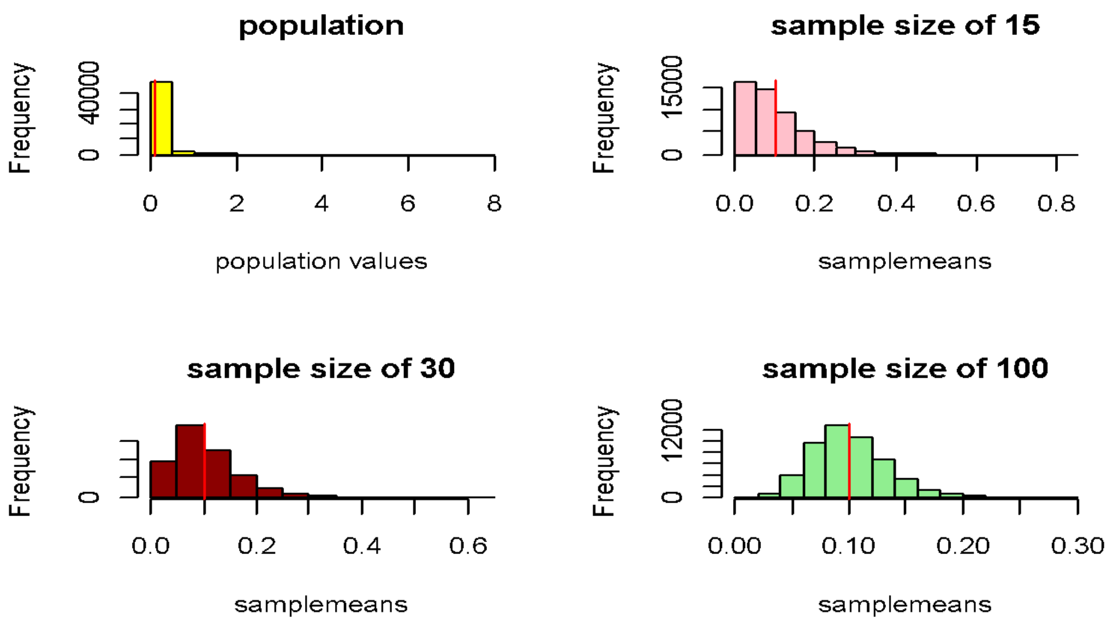

Figure 3: Histogram of population and sampling distribution of mean from Gamma distribution

Source: *Plotted by the Researcher*

Figure 3 presents the histogram, mean and standard deviation of both population and sample means distribution. These distributions were drawn from Gamma distribution with shape 0.1. Furthermore, these shapes represented 500 stimulated sample means drawn from Gamma distribution.

RESULT AND DISCUSSION

From the histogram plotted in Figure 1, both the population distribution and sampling distribution of mean drawn from a normal distribution produced a bell-like shape in all the sample sizes (15, 30 and 100). The red vertical line in Figure 1 indicates the mean of the Normal distribution. This red vertical line also shows that population mean is equal to the mean of sample means. Furthermore, the variability of each sampling distribution decreases as the sample size increases. Figure 2 displays a positive skewed shape for exponential distribution. Its sampling distribution gradually transforms to normal distribution as the sample size increases. The plot in Figure 2 above also shows that population mean is equal to the mean of sample means. Also, as the sample size increases, the variability of the sampling distribution reduces. Figure 3 indicates that Gamma distribution has a positive skewed shape. But the sampling distribution of mean drawn from Gamma distribution portrays a normal-like shape as the number of sample sizes gradually increases to infinity. The red vertical line in Figure 3 denoted the mean of both the population and sample mean distributions. The plot shows that

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx

the mean of the Gamma distribution is equal to the mean of the sampling distribution. Similarly, the variability of each sampling distribution decreases as the sample size increases.

CONCLUSION

Three parameters were used to justify CLT: shape, mean and variance. From the findings, it can be concluded that under normal distribution, the sampling distribution of mean produced a shape like normal distribution irrespective of sample size with mean equal to the mean of Normal distribution and variance equal to the $\frac{variance\ of\ normal\ distribution}{sample\ size\ (n)}$. Conversely, the shape of sampling distribution of mean under non-normal distributions gradually converges to Normal distribution with mean equal to the mean of non-normal distribution and variance equal to the $\frac{variance\ of\ non-normal\ distribution}{sample\ size\ (n)}$ as sample size goes to infinity ($n \to \infty$). More so, the variability of each sampling distribution decreases as the sample size increases.

REFERENCES

Fatheddin, P. (2015). Central Limit Theorem for a Class of SPDEs. Journal of Applied Probability, 52(3), 786 -796.

Islam, k. and Mohammed, R. (2018). Sample Size and Its Role in Central Limit Theorem (CLT). International Journal of Physics and Mathematics, vol. 1, pp. 37 – 47.

Sawada, T. (2021). Conditions of the Central Limit Theorem Are Rarely Satisfied in Empirical Psychological Studies. Front. Psychol.12:762418.doi:10.3389/fpsyg.2021.762418

Shige, P. (2019). Law of Large Numbers and Central Limit Theorem Under Nonlinear Expectations. Probability Uncertainty and Quantitative Risk, 4:4. doi: 10.1186/s4156-019-003-2.

Sutanapong, C. and Louangrath, P. (2015). Central Limit Theorem And Its Applications. Part 3 of 4 ” Inter.j. Res. Methodol. Soc Sci., Vol., 1, No, 3: pp 30 -41. (Jul – Sep, 2015).

Winters, J. and Kinneth, S.(2022). Application of the Central Limit Theorem. Selected Honors Theses.158.

Article DOI: 10.52589/xxxxxx
DOI URL: https://doi.org/10.52589/xxxxxxxxx