Original Research Article

CENTRAL LIMIT THEOREM AND ITS APPLICATIONS IN DETERMINING SHOE SIZES OF UNIVERSITY STUDENTS

Abstract : The Central limit theorem is a very powerful tool in statistical inference and Mathematics in general since it has numerous applications such as in topology and many other areas. For the case of probability theory, it states that, "given certain conditions, the sample mean of a sufficiently large number or iterates of independent random variables, each with a well-defined mean and well-defined variance, will be approximately normally distributed". In our research paper, we have given three different statements of our theorem (CLT) and thereafter proved it using moment generating functions and characteristic functions. We later showed vividly that the moment generating functions and the characteristic functions do exist for the normal distribution. This research paper has data regarding the shoe size and the gender of the of the university students. This paper is aimed at finding if the shoe sizes converges to a normal distribution as well as find the modal shoe size of university students and to apply the results of the two proofs of the central limit theorem to test the hypothesis if most university students put on shoe size seven. The Shoe sizes are typically treated as discretely distributed random variables, allowing the calculation of mean value and the standard deviation of the shoe sizes. The sample data which is used in this research paper belonged to different areas of Kibabii University which was divided into five strata. From two strata, a sample size of 15 respondents was drawn and from the remaining three strata, a sample of 14 students per stratum was drawn at random which totaled to a sample size of 72 respondents. By analyzing the data, using SPSS and Microsoft Excel, it was vivid that the shoe sizes are normally distributed with a well-defined mean and standard deviation. We also proved that most university students put on shoe size seven by testing our hypothesis using the p-value and the confidence interval. The modal shoe size for university students was found to be seven since it had the highest frequency (18/72) . This research was aimed at enlightening shoe investors, whose main market is the university students, on the shoe sizes that are on high demand among university students.

Keywords – Central Limit Theorem, Moment generating function, Characteristics function.

1.0 INTRODUCTION

The Central Limit Theorem has been around for over 280 years and many researchers in the field of mathematics have proved it in many different cases since it has many different versions according to different researchers in different areas of applications such as in probability theory and other areas. Its origin can be traced to The Doctrine of Chances by Abraham de Moivre 1738 [1]. In his book, he provided techniques for solving gambling problems, and also provided a statement of the Central Limit Theorem for Bernoulli trails as well as gave a proof for $p = \frac{1}{2}$. This was a very crucial invention during those early days which motivated many other researchers years later to look at his work and they continued to ascertain it for further cases. Many researchers had made several studies on the sums of independent random variables for many different error distribution before 1810 which had mostly led to very complicated formulas when Laplace released his first paper about the CLT. In 1812, Pierre Simon Laplace published his own book titled Theorie Analytique des Probabilities, where he generalized the theorem for $p \neq \frac{1}{2}$. He also gave a proof, although not a arduous one, for his finding [2]. Siméon Denis Poisson later published two articles (1824 and 1829) where he discussed the CLT with an idea that all procedures in the physical world are governed by distinct mathematical laws where he was trying to provide a more reliable mathematical analysis to Laplace's theorem. He provided a more rigorous proof for a continuous variable and also discussed the validity of the central limit theorem, mainly by providing a few counterexamples but he was unable to prove his general formula because he examined its validity for the special case of n=1.

Towards the end of 19 century, Dirichlet and Bessel followed the tracks of Laplace and Poisson in their proofs where they introduced the "discontinuity factor" in their proofs which enabled them to prove Poisson's equation for the general case. Probability theory was first considered as "pure " mathematics by Cauchy . He proved the CLT by first finding an upper bound to the difference between the exact value and the approximation and then specified conditions for this bound to tend to zero. Cauchy gives his proof for independent identically distributed variables $y_1 \ldots y_n$ with a symmetric density f(y), finite support [-a, a], variance $\sigma^2 > 0$ and a characteristic function $\psi(\theta)$.This proof finished the so called the first period of the central limit theorem (1810-1853) where the proofs presented in this period were not satisfactory in three respects namely, The theorem was not proved for distributions with infinite support, There were no explicit conditions, in terms of the moments, under which the theorem would hold, The rate of convergence for the theorem was not studied. These glitches were eventually solved by Chebyshev, Markov and Liapounov; the so-called "St. Petersburg School" between 1870 and 1910. Chebyshev's paper in 1887 is generally considered the beginning of rigorous proofs for the central limit theorem. In his paper, he considered a sequence of independent random variables each described by probability densities where he used the "method of moments", that he had earlier developed which he left incomplete. Markov later simplified and completed Chebyshev's proof of the CLT. In 1898, after Chebyshev's proof, Markov stated that: "a further condition needs to be added in order to make the theorem correct". He first proposed the following condition: i) B^2_n / n is uniformly bounded away from 0 which he later replaced by ii) $E(z^2_n)$ is bounded from 0 as $n \rightarrow \infty$. Liaupounov's proof, published in 1901, is considered the first "real" rigorous proof of the CLT where he considered a sequence of random variables with mean 0 and variance 1. At around 1901-1902 the Central Limit Theorem become more generalized and a complete proof was given by Aleksandr Lyapunov [3]. In 1922 Lindeberg gave a more generalized statement of CLT which states that, "the sequence of random variables need not be identically distributed, instead the random variables only need zero means with individual variances small compared to their sum" [4]. Numerous contributions to the statement of the Central Limit Theorem and different ways to prove the theorem began to appear around 1935, when both Levy and Feller published their own independent papers regarding the Central Limit Theorem[5]. Feller's paper of 1935 gives the necessary and sufficient conditions for the CLT, but the result was somewhat restricted which made it not to be the rigorous proof of the CLT. Feller considered an infinite sequence x_i of independent random variables. .In 1935, Lévy proved several things related to the central limit theorem:i) He gave necessary and sufficient conditions for the convergence of normed sums of independent and identically distributed random variables to a normal distribution ii) Lévy also gave the sufficient and necessary conditions for the general case of independent summands iii) He also tried to give the necessary and sufficient conditions for dependent variables, martingales. Lévy's proofs also was not satisfactory for the martingale case and therefore it did not stand a test of rigorousness since it relied on a hypothetical lemma.

In 1936, Cramér proved the lemma as a theorem and the matter of both Lévy' and Feller was settled. In 1937 they returned and refined their proofs using Cramérs result and thus, CLT was proved with both necessary and sufficient conditions. The Central Limit Theorem had unlimited impact and continues to have the same in the field of mathematics because the theorem is being used in topology, and other fields in mathematics and not limited to probability theory only.

1.1 Statement of the problem

The Central Limit Theorem is the dominating theorem in statistical inference. It permits us to make assumptions about a population and states that a normal distribution will occur regardless of what the initial distribution looks like for a su□ciently large sample size n. This theorem is used to make sound assumptions regarding the population since it is difficult to make such assumptions when the population isn't normally distributed and the shape of the distribution is unknown. The goal of this research project is to focus on the Central Limit Theorem and its applications in statistical inference, as well as to know the importance of central limit theorem, how to prove it and how to apply the theorem in shoe sizes data of Kibabii University students.

1.2 Significance of the study

The analysis of the shoe size data of kibabii University students will help the shoe investors around the University with the knowledge of the shoe size to stock more because of the high demand and as a result improve their sales and profit.

2. METHODOLOGY

2.1 Sample Size

Statistical inference refers to the process of making inference about the population characteristics of interest using a sample from the population. It is normally difficult , time consuming and costly to obtain information from the entire population of interest. Therefore, researchers collect samples which are subsets of the population in order to make inference on certain parameters of scientific interest. This makes sampling an important feature of statistical study. However deciding on the size of the sample that will represent the population well is a challenge. In this work, a formula recommended by **Creswell (2013)** was used.

$$n=\frac{Z^2 pqN}{e^2(N-1)+Z^2 pq}$$

where: n is the sample size, N is the population, and p is the population reliability which is considered as p= 0.05 and Z is 1.96 at a significance level of 0.05 and e is the standard error of 5%.

Therefore;

$$n = \frac{Z^2 pqN}{e^2(N-1)+Z^2 pq}$$

$$= \frac{1.96^2 \times 0.05 \times 0.95 \times 8000}{0.05^2(8000-1)+1.96^2 \times 0.05 \times 0.95}$$

$$= 72.33943192 \approx 72$$

2.2 Data

This study was conducted though a closed and open-ended questionnaire where 3 questions were related to the personal data and 3 questions related to the subject study totaling to 6 questions. This researcher selected 72 Kibabii University students which formed the required sample size.
The shoe size, height, body weights, gender, year of study and age data for students was collected in the following areas of Kibabii University.

AREA NUMBER	AREA NAME
1	Tuuti
2	Booster
3	Lavington
4	Butieli
5	Institution Area

2.3Statements of the Central Limit Theorem

Since many researchers have done many research works on the Central Limit Theorem, they have come up with many proofs which are all accepted. Let's first state Abraham de Moivre-Laplace Theorem which states as follows.

Theorem 2.3.1[1]. Consider a sequence of Bernoulli trials with probability p of success, where 0 < p < 1. Let S_n denote the number of successes in the first n trials, $n \geq 1$. For any $a, b \epsilon \mathbb{R} \cup \{\pm\infty\}$ $with\ a \angle b$

$$\lim_{n\to\infty}\left(a \leq \frac{S_{n-np}}{\sqrt{np(1-p)}} \leq b\right) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{\frac{-z^2}{2}}\,dz$$

Thereafter Lypunov gave the second statement of the Central Limit Theorem as:

Theorem 2.3.2

$Suppose\ X_{n,}\ n\ 1\ are\ indendent\ random\ variables\ with\ mean\ 0\ and\ \sum_{k=1}^{n}\frac{|x_k|^{\delta}}{s_n{}^{\delta}} \to 0\ for\ some\ \delta > 2.$ Then, $\frac{S_n}{s_n} \xrightarrow{distr} N(0,1), where\ S_n = X_1 + X_2 + \cdots + X_{n,}\ ,\ s_{n=}\sum_{k=1}^{n} E(X_k{}^2), n \geq 1\ and\ \xrightarrow{distr}$ represents convergence in distribution.

An independent and identically distributed random variable is defined as follows:

Definition 2.0.A sequence of random variables is said to be **independent and identically distributed** if all random variables are mutually independent, and if each random variable has the same probability distribution.

Now, The third and final statement of the central limit theorem states that:

Theorem2.3.3.

$suppose\ that\ X_1, X_{2,\dots,} X_{n,}\ are\ independent\ and\ identically\ distributed\ with\ mean\ \mu\ and\ variance\ \delta^2 > 0. Then, \frac{S_n - n\mu}{\sqrt{n\delta^2}} \xrightarrow{distr} N(0,1), where\ S_n = X_1 + X_2 + \cdots + X_{n,}\ ,\ n \geq 1\ and$

$\xrightarrow{distr}$ represents convergence in distribution.

2.4 Proofs of Central Limit Theorem

In this work, we considered only two proves of the theorem, using the moment generating functions and using the characteristic functions .

2.4.1 Proof of Central Limit Theorem Using Moment Generating Functions

Here are some crucial aspects of moment generating functions we need to discuss before we look at the proof of the moment generating functions. These includes some definitions, remark and the properties of the moment generating functions where we are going to start with the definitions. [8].

Definition2.4.2 The moment-generating function (MGF) of a random variable X is defined to be

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x), & if\ X\ is\ discrete, \\ \int_{-\infty}^{+\infty} e^{tx} f(x)dx, & if\ X\ is\ continous. \end{cases}$$

Moments can also be found by di□erentiation.

Theorem2.4.3 Let X be a random variable with moment-generating function $M_X(t)$.We have$\frac{d^r M_X(t)}{dt^r}|_{t=0} = \mu_r{}'\quad where\ \mu_r{}' = E(X^r).$

Remark2.4.4

$E(X^r)$ describes the rth moment about the origin of the random variable X. We can see then that $\mu_1{}' = E(X)$ and $\mu_2{}'$ $= E(X^2)$ which therefore allows us to write the mean and variance in terms of moments.

Properties of Moment generating functions

Theorem 2.4.5 $M_{a+bX}(t) = E\left(e^{t(a+bX)}\right) = e^{at}M_X(bt).$

proof; $M_{a+bX}(t) = E\left\{e^{t(a+bX)}\right\} = E(e^{at}).E\left(e^{t(bX)}\right) = e^{at}E\left(e^{(bt)X}\right) = e^{at}M_X(bt).$

Theorem 2.4.6 Let X and Y be random variables with moment − generating functions $M_X(t)$ and $M_Y(t)$ respectively. Then $M_{X+Y}(t) = M_X(t).M_Y(t)$.

proof. $M_{X+Y}(t) = E\left(e^{t(X+Y)}\right) = E(e^{tX}.e^{tY}) = E(e^{tX}).E(e^{tY})$ *(by independence of random variables)*

$= M_X(t).M_Y(t)$

corollary 2.4.7 Let $X_1, X_2, \dots, X_n$ be random variables then, $M_{X_1+X_2+X_3+\dots+X_n}(t) = M_{X_1}.M_{X_2}.\ M_{X_3}.M_{X_4}\dots.M_{X_n}(t)$ *This prooof is nearly identical to the proof of the previous theorem.*

To proof the central limit theorem, it is necessary to know the moment generating function of the normal distribution.

Lemma 2.4.9 The moment geneerating function (MGF) of the normal random variable X with mean μ and

Variance δ^2, $(i.e., X \sim N(\mu, \delta^2))$ *is* $M_X(t) = e^{\mu t + \frac{\delta^2 t^2}{2}}$

proof; First we will find the MGF for the normal distribution with mean 0 and variance 1, i.e, N(0,1).

If $Y \sim N(0,1)$, *then;*

$M_Y(t) = E(e^{tY})$

$$= \int_{-\infty}^{+\infty} e^{ty} f(y)\,dx = \int_{-\infty}^{+\infty} e^{ty}\left(\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}y^2}\right)dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{ty}\,e^{\frac{-1}{2}y^2}\,dy$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{(ty\ \frac{-1}{2}y^2)}dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{(\frac{1}{2}t^2 + (\frac{-1}{2}(y^2+2ty+t^2)))}dy$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{1}{2}t^2}\,e^{\frac{-1}{2}(y^2-2ty+t^2)}dy$$

$$= e^{\frac{1}{2}t^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{-1}{2}(y-t)^2}\,dy.$$

But $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{\frac{-1}{2}(y-t)^2}\,dy$ this is the probability distribution function of the normal distribution. So;

$M_Y(t) = e^{\frac{1}{2}t^2}$. *Now, if* $X \sim N(\mu, \delta^2)$, *and*

$$M_X(t) = M_\mu + \delta Y(t) = e^{\mu t} M_Y(\delta t) = e^{\mu t} e^{(\frac{1}{2}\delta^2 t^2)} = e^{\left(\mu t + \frac{\delta^2 t^2}{2}\right)}$$

Let’s write the Taylor series formula before we start our proof because it’s of great significance in our proof

Lemma 3.8 $e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots, \ -\infty < x < \infty$ (*Taylor series*).

Now let us prove a special case of where $M_X(t)$ exists in a neighborhood of 0.

Proof; Let $Y_i = \frac{X_i - \mu}{\delta}$ for $i = 1,2,3, \ldots$ and $R_n = Y_1 + Y_2 + \cdots + Y_n$

$$\frac{S_n - n\mu}{\sqrt{n\delta^2}} = \frac{Y_1 + Y_2 + \cdots + Y_n}{\sqrt{n}} = \frac{R_n}{\sqrt{n}},$$

So $\frac{S_n - n\mu}{\sqrt{n\delta^2}} = \frac{R_n}{\sqrt{n}} = Z_n$

Since R_n is the sum of independent random variables, we see that its moment generating function is:

$$M_{R_n}(t) = M_{Y_1}(t)\, M_{Y_2}(t) \ldots M_{Y_n}(t)$$

$$= [M_Y(t)]^n$$

We now note that this is true because each Y_i is independent and identically distributed. Now,

$$M_{Z_n}(t) = M_{\frac{R_n}{\sqrt{n}}}(t) = E\left(e^{\frac{R_n}{\sqrt{n}}}\right) = E\left\{e^{(R_n)(\frac{t}{\sqrt{n}})}\right\} = M_{R_n}\left(\frac{t}{\sqrt{n}}\right) = \{M_Y(\frac{t}{\sqrt{n}})\}^n.$$

Taking the natural logarithm of each side,

$$l_n M_{Z_n}(t) = n l_n M_Y(\frac{t}{\sqrt{n}})$$

But we know that:

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = E\left(e^{\frac{t}{\sqrt{n}}Y}\right).$$

$$= E\left\{1 + \frac{tY}{\sqrt{n}} + \frac{(\frac{t^2 Y}{\sqrt{n}})^2}{2} + O\left(\frac{1}{n^{\frac{3}{2}}}\right)\right\}.$$

$$= 1 + \frac{t^2}{n}\frac{E(Y^2)}{2} + O\left(\frac{1}{n^{\frac{3}{2}}}\right).$$

$$= 1 + \frac{t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right).$$

Where $O\left(\frac{1}{n^{\alpha}}\right)$ *stands for* $\lim Sup_{n\to\infty} \frac{|O(\frac{1}{n^{\alpha}})|}{\frac{1}{n^{\alpha}}}$

Then $l_n M_{Z_n}(t) = n ln\left\{1 + \frac{t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right)\right\}.$

$$lnM_{Z_n}(t) = n\left\{\frac{t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right)\right\}$$

$$= \left\{\frac{t^2}{2} + O\left(\frac{1}{n^{\frac{3}{2}}}\right)\right\}$$

$$l_n M_{Z_n}(t) = \frac{t^2}{2} + O\left(\frac{1}{n^{\frac{1}{2}}}\right),$$

So we have that, $M_{Z_n}(t) \to e^{\frac{t^2}{2}}$ *as* $n \to \infty$.

$$Thus, Z_n \to N(0,1), i.e, \quad \frac{S_n - n\mu}{\sqrt{n\delta^2}} \to N(0,1).$$

2.4.2 Proof of Central Limit Theorem Using Characteristic Functions

Let us now prove the Central Limit Theorem using the characteristic functions. This is because the moment generating functions do not exist for all distributions when the moments of a given distribution are not finite. In such a situation when the moments are not finite, we generally look at the characteristic functions because they exist for every given distribution. [8].

Defination2.4.2.1The Characteristic function of a continuous random variable X

$$C_X(t) = E(e^{itX}) = \int_{-\infty}^{+\infty} e^{itx} f(x)\,dx, \quad \textit{where } t \textit{ is a real valued function, and } i = \sqrt{-1}.$$

$C_X(t)$ will always exist because e^{itx} is a bounded function, i.e, $|e^{itx}| = 1 \ \forall\ t,x \in \mathbb{R}$, and so the integral exists.

The characteristic function also has many similar properties to moment generating functions.

Let us look at the characteristic function of the normal distribution before we prove the central limit theorem.

Lemma 2.5.2 Let $R_n, n \geq 1$ be a sequence of random variables.

$If, n \to \infty, C_{R_n}(t) = E\left(e^{iR_n t}\right) \to e^{\frac{-t^2}{2}} \quad \forall t \epsilon(-\infty, \infty), then\ R_n \to N(0,1).$

We can now prove the central limit theorem using characteristics functions.

$PROOF: Similar\ to\ the\ proof\ using\ moment\ generating\ funtions\ let\ Y_i = \dfrac{X_i - \mu}{\delta}\ for\ i = 1,2,3,\dots and\ let\ R_n = Y_1 + Y_2 + \cdots + Y_n\ so,$

$$\frac{S_n - n\mu}{\sqrt{n\delta^2}} = \frac{R_n}{\sqrt{n}} = Z_n, \qquad where\ S_n = X_1 + X_2 + \cdots + X_n.$$

Now we note that

$R_n\ is\ the\ sum\ of\ independent\ random\ variables\ , so\ we\ see\ that\ the\ characteristic\ function\ of\ R_n\ is:$

$$C_Y(t) = C_{Y_1}(t)C_{Y_2}(t)\dots C_{Y_n}(t)$$

$$= [C_Y(t)]^n$$

Since all $Y_i's\ are\ independent\ and\ identically\ distributed. Now,$

$$C_{Z_n}(t) = C_{\frac{R_n}{\sqrt{n}}}(t)$$

$$= E\{e^{\frac{iR_1 t}{\sqrt{n}}}\}$$

$$= E\{e^{i}(R_n)(\frac{t}{\sqrt{n}})\}$$

$$= C_{R_n}(\frac{t}{\sqrt{n}})$$

$$= [C_Y(\frac{t}{\sqrt{n}})]^n.$$

Taking the natural logarithm on each side,

$$lnC_{Z_n}(t) = nlnC_Y\left(\frac{t}{\sqrt{n}}\right).$$

We can note from the previous proof with some modifications that:

$$C_Y(t) = 1 - \frac{-t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right).$$

Then we have,

$$lnC_{Z_n}(t) - nln(1 - \frac{t^2}{2n} + O\left(\frac{1}{n^{\frac{3}{2}}}\right))$$

Then;

$$lnC_{Z_n}(t) = -\frac{t^2}{2} + O\left(\frac{1}{n^{\frac{1}{2}}}\right),$$

$$So, as\ n \to \infty, \qquad lnC_{Z_n}(t) \to \frac{-t^2}{2}\ and$$

$$C_{Z_n}(t) \to -e^{\frac{t^2}{2}}\ as\ n \to \infty$$

We therefore conclude that;

$$Z_n = \frac{S_n - n\mu}{\sqrt{n\delta^2}} \to N(0,1).$$

3. RESULTS AND DISCUSSION

Here, we discuss the results that we have found from our analysis as well as the significance of our research work. These results will help in devising the appropriate conclusion and the recommendations. Before we start our analysis, let's first say something about our theorem;

Central Limit Theorem is one of the most great and worthwhile ideas in all of Statistics and there are two alternative forms of the theorem, and both describe the center, spread and shape of a certain sampling distribution. We have considered the two case in our analysis. We define the sampling distribution of a statistic as the distribution of values of that statistic when all possible samples of the same size are taken from the same population. Sampling distributions form the foundation for almost all methods in inferential statistics, and the Central Limit Theorem allows us to explicitly describe the sampling distribution for a sample mean x. We have discussed these two cases i.e. sampling distribution for the sample means and sample sums below.

3.1 Sampling distribution for the sample mean

We have provided the results and the discussion of the distribution of the sample means below.

SAMPLE SUMS	SAMPLE AVERAGES	FREQUENCIES
202	6.73	1
203	6.77	2
209	6.97	3

210	7.00	4
211	7.03	3
213	7.10	2
215	7.17	1

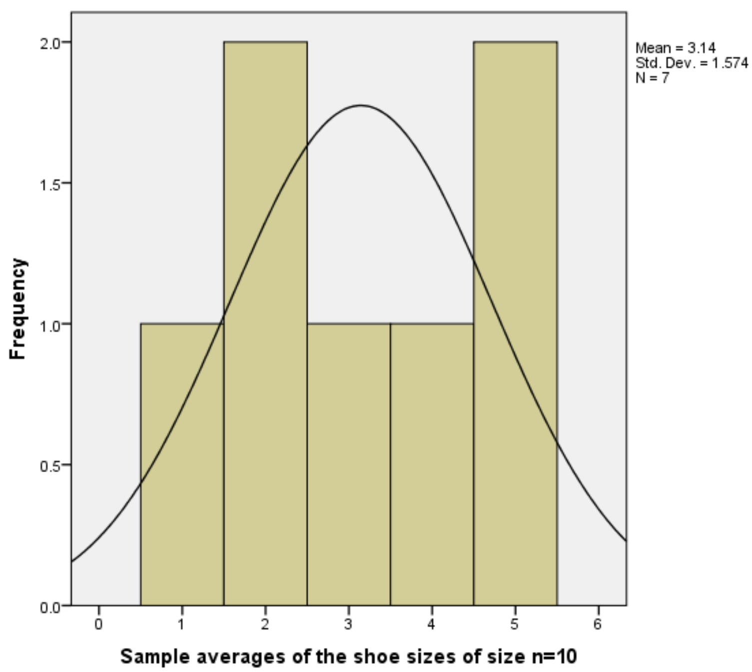


Fig 3.0 Sampling distribution of the mean shoe sizes of samples of size n=10

From the above figure, we have the samples of size 10 which does not give us a pretty idea of convergence to a normal distribution. This is because the samples so drawn did not meet the condition of the central limit theorem which states that the sample size n should be sufficiently large for a normal distribution convergence. It is also vivid that most of the sample means are not even close to the population mean which should be the case for the data of the shoe sizes to converge to the normal distribution where we expect that the sample mean should be close or even equal to the population mean. The graph also does not seem to resemble a normal curve and there comes the need of a sufficiently large sample size n. This has a well-defined mean of 3.63 and standard deviation of 1.991 but fails to be normally distributed simply because of a small sample sizes.

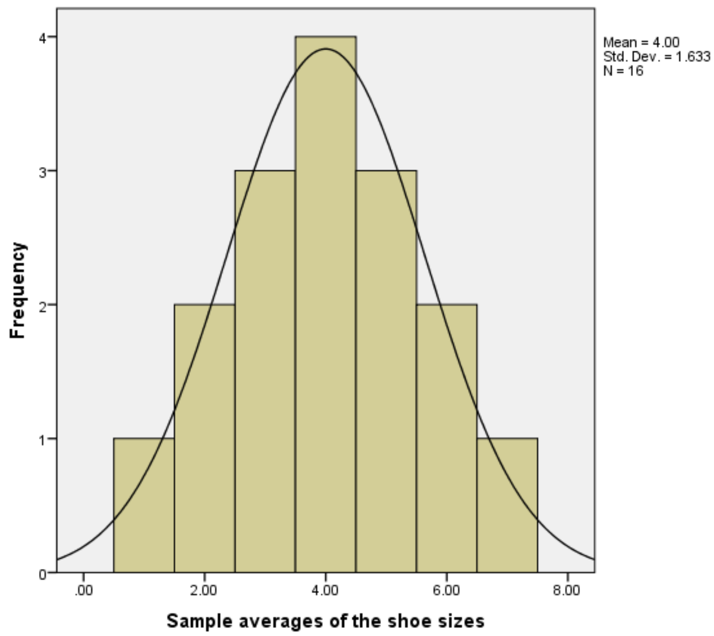


Fig 3.1 Sampling distribution for the mean shoe sizes of samples of size n=30

From this figure, we can see that the sampling distribution for the sample means of shoe sizes converges perfectly to the normal distribution. This is because the condition of drawing a reasonably large sample size was observed making the distribution to be symmetrical. This also indicates without doubt that most of the sample means are pretty close to the population mean, thus making the pdf of the distribution to approach zero as we move away from the center. We can also ascertain that the sample mean underestimates the population mean and so we have positive and negative deviations from the population mean which are almost similar thus making our distribution to be symmetrical or bell-shaped. Moreover, the mean of this sampling distribution is the mean of the population from which we sampled which is shoe size seven for our case. So this clearly indicates that most of the university students put on shoe size 7, with less people putting on shoe sizes 4 and 10. This distribution also shows a well- defined mean of 4.oo and a standard deviation of 1.663

3.2 Sampling distribution of the sample sums

The results for the distribution of the sample sums is discussed below,

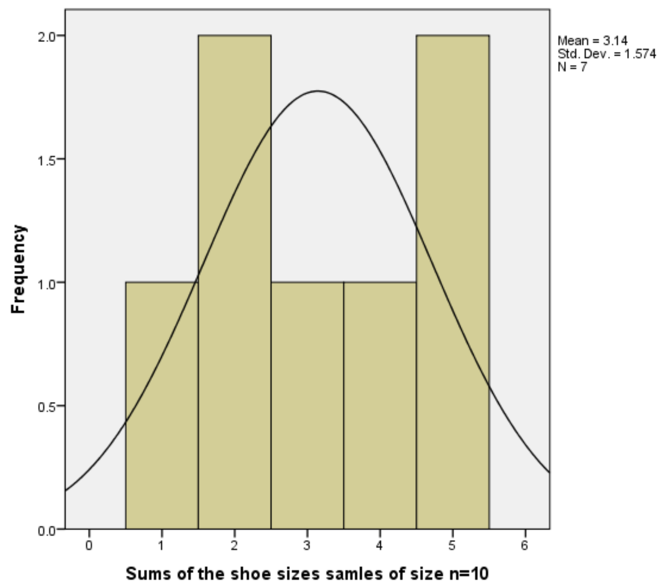


Fig 3.20 Sampling distribution for the sample sums of shoe sizes of samples of size n=10

We explains this using the second version of the central limit theorem which says that if the sample averages converges to a normal distribution, also the distribution for the sample sums will also be normally distributed. So since our sample were not normally distributed, so is the distribution for the sample sums. We can see that most of the means of the distribution are not concentrated to the center of our graph and so it isn't normal and its curve is not bell-shaped. This also is caused by drawing small sample sizes since the mean and the standard deviations are well-defined.

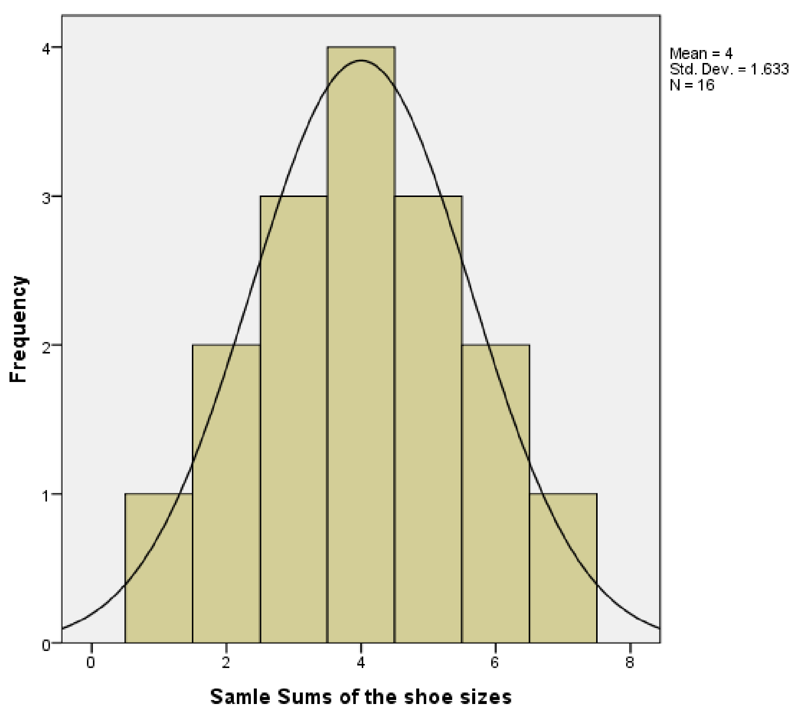
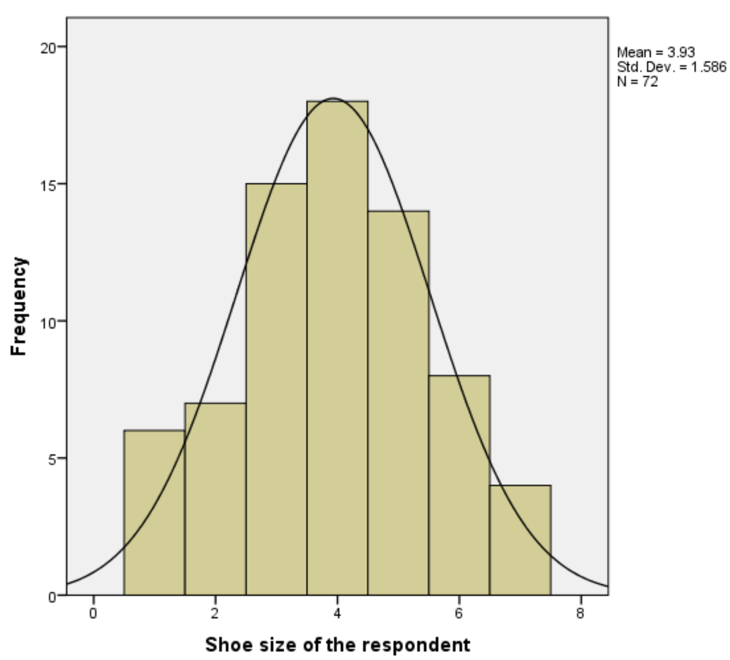


Fig 3.21 Sampling distribution for the sample sums of shoe sizes of samples of size n=10

From this figure, we can see that the sampling distribution for the sample means of shoe sizes converges to the normal distribution symmetrically and with a bell- shaped curve since we had drawn large sample sizes necessary for any distribution or non- parametric distribution to converge to a normal distribution. The symmetrical distribution means that the sample mean is pretty close to the population mean, thus making the pdf of the distribution to approach zero as we move away from the center. We can also ascertain that the sample mean underestimates the population mean and so we have positive and negative deviations from the population mean which are almost similar thus making our distribution to be symmetrical or bell-shaped as from our case above. Moreover, the mean of this sampling distribution is the mean of the population from which we sampled which is shoe size seven for most university students... So this clearly indicates that most of the university students put on shoe size 7, with less people putting on shoe sizes 4 and 10. This distribution also shows a well- defined mean of 4.oo and a standard deviation of 1.663

3.2 THE DISTRIBUTION OF SHOE SIZES OF THE RESPONDENTS

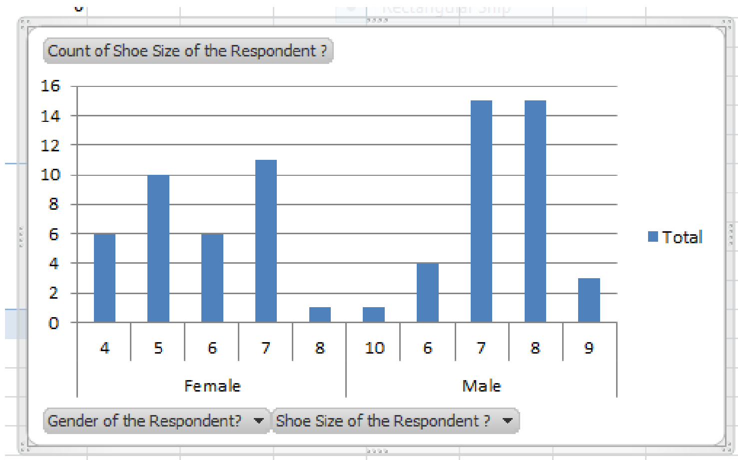
This histogram suggests that the shoe sizes of university students are normally distributed with a well-defined expected value of 3.93 and a well-defined standard deviation of 1.586 .For this case we have not used the concept of our theorem but we have just drawn a graph of the shoe sizes to see how they are distributed for only 72 respondents. From our graph we can see that most of university students put on shoe size 7 and a few people put on shoe size 4 and 10.This undoubtedly shows that a shoe investor needs to stock more on shoe size 7 followed by 6,8and 9, 5, 4and stock less on shoe size 10. So we notice that if ones happen to ask enough people about their shoe sizes, the distribution of the shoe sizes is normally distributed with a bell-shaped curve.



3.3 DISTRIBUTION OF SHOE SIZES ACCORDING TO THE GENDER OF THE RESPONDENTS

This graph compares shoe sizes and the gender of the respondents which indicates that shoe sizes differ with the gender of the respondent.

We see that most ladies put on small shoes i.e. shoe size 5 and 7 with the minority of ladies putting on shoe 4, 6 and 8. For the case of men, most respondents had shoe sizes 7 and 8 and a few had shoe sizes 10, 6, and 9.This apparently shows that most men put on big shoe size as compared to the case of ladies which also means that most ladies put on small shoe sizes as compared to men.



Since Central Limit Theorem has many applications in probability theory and statistical inference, we have limited our research paper on hypothesis testing using shoe size data of Kibabii University students. Before we begin to compute if most people put on shoe size seven, we must first satisfy four conditions;

(i) independence condition(assumption): This condition for our case states that, each respondent's shoe size that the researcher is going to meet is independent of the shoe size of the next respondents.

(ii) random condition: Since we have many students in Kibabii University totaling to almost 8,000, taking just 72 students to observe the data will account for our randomization condition.

(iii) 10% condition: In this condition, the sample size n, should be less than 10% of the population size. For our case, our sample size n=72 which is less than 10% of the total population which is 800.Therefore 72 ∠800 and so our sample holds true the 10% condition.

(iv) success/failure: This simply state that the population size multiplied by our proportion in our hypothesis must be greater than 10.Since our proportion p=0.5, we can proof the condition by multiplying the two. i.e.np= 8,000*0.5=4,000.So the success/failure condition also holds true because 4,000>10.

These are the two methods which are used to test the hypothesis.

1. **THE CONFIDENCE INTERVAL**

$the\ right-sided\ 100(1-\alpha)\%\ confidence\ interval\ for\ p\ for\ a\ large\ sample\ which\ is\ given\ by;$

$$\hat{p}\ - Z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p \leq 1$$

Since there are 26 respondents having shoe size 7, we gets that;

$\hat{p} = \frac{26}{72} = 0.36$ and $Z_{0.05} = 1.645$ and n=72. So by applying the above formula we get that;

$0.36 - 1.645\sqrt{\frac{0.36(1-0.36)}{72}}$=0.2669 0.2669<p$\leq 1$.

Since $0.5 \in (0.2669{,}1]$, we cannot reject $H_0: p = 0.5$ in favor of $H_1: p < 0.5$ at the 0.05 level of significance. This is because we have enough evidence from our data to support that most university students put on shoe size seven.

2. **THE P-VALUE**

Now we will use the p-value approach to test our hypothesis. We must find the z-value for testing our observed value. We use the following equation to do so;

$$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.36-0.5}{\sqrt{\frac{0.5*0.5}{72}}} = -2.3759$$

```
/VARIABLES=ShoeSizeoftheRespondent
/CRITERIA=CI(.95).
```

T-Test

```
[DataSet1] C:\Users\admin\Desktop\new morris.sav
```

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Shoe Size of the Respondent ?	72	6.68	1.372	.162

One-Sample Test

	Test Value = 7					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Shoe Size of the Respondent ?	-1.976	71	.052	-.319	-.64	.00

This corresponds to a p value of 0.52. Since 0.52>0.05 we cannot reject H_0 in favor of our alterative hypothesis because we have enough evidence from our data to support that most university students put on shoe size seven. Therefore from the two cases i.e. using the p-value and the confidence interval, it's clear that most university students put on shoe size seven since we have not rejected the null hypothesis for both cases due to presence of enough evidence from our data.

4.0 CONCLUSIONS

It is now clear from our data that the shoe sizes of university students converge to a normal distribution using the proof of the central limit theorem by considering the moment generating functions as well as the characteristic functions. Using the shoe sizes data so collected, we were able to prove that most students put on shoe size 7 by testing our hypothesis using the p-value and the confidence interval. This is because for both cases, we have enough evidence from our data to show that most students put on shoe size seven. By finding the mode also, we found that most university students put on shoe size seven because it had the highest frequency.

4.1 RECOMMENDATIONS

Since most university students put on shoe size seven, we recommend shoe investors around the institutions of higher learning to be stocking more of shoe size seven because it's the shoe size with majority of the students. Followed by shoe sizes 5, 6 and 8 and doing so, they will curb the big problem of so much dead stock that they face day in day out.

In future, it may be interesting to use my applications on other areas such as sports, finding the distribution of the change people carry in their pockets, although we must make sure that we have a sufficiently large sample size to have accurate results of a smooth convergence in normal distribution since some of the distributions are heavily skewed as well as when testing the hypothesis. Other applications of the Central Limit Theorem, as well as other properties such as convergence rates may also be interesting areas of study for the future.

REFERENCES

1. Dunbar, Steven R. "The de Moivre-Laplace Central Limit Theorem."
2. Laplace, Pierre-Simon. Pierre-Simon Laplace Philosophical Essay on Probabilities: Translated from the fifth French edition of 1825 With Notes by the Translator. Vol. 13. Springer Science & Business Media, 2012.
3. Adams, William J. The life and times of the central limit theorem. Vol. 35. American Mathematical Soc., 2009.
4. Linnik, Ju V. "An information-theoretic proof of the central limit theorem with Lindeberg conditions." Theory of Probability & Its Applications 4.3 (1959): 288-299.
5. Le Cam, Lucien. "The central limit theorem around 1935." Statistical science (1986): 78-91.
6. Hogg, Robert V., and Allen T. Craig. Introduction to mathematical statistics.(5"" edition). Upper Saddle River, New Jersey: Prentice Hall, 1995.
7. Chen, Yanling. "Introduction to probability theory." Lecture Notes on Information Theory, Duisburg-Essen Univ., Duisburg, Germany (2010).
8. Filmus, Yuval. "Two proofs of the central limit theorem." Recuperado de http://www. cs. toronto. edu/yuvalf/CLT. pdf(2010).
9. Grinstead, Charles Miller, and James Laurie Snell. Introduction to probability. American Mathematical Soc., 2012.
10. Grinstead, C.M. and Snell, J.L., 1997. Central limit theorem. Introduction to probability, pp.325-355.
11. Anderson, Nicole. "Central Limit Theorem and Its Applications to Baseball." (2014).
12. Cryer, Jon, and Jeff Witmer. "Introduction to the Theory of Inference."
13. Tufts, D. W., et al. "Detection of signals in the presence of strong, signal-like interference and impulse noise." Topics in Non-Gaussian Signal Processing. Springer, New York, NY, 1989. 184-196.
14. Everitt, Brian S. "Mixture Distributions—I." Encyclopedia of statistical sciences 7 (2004).
15. Walpole, Ronald E., et al. Probability and statistics for engineers and scientists. Vol. 5. New York: Macmillan, 1993..

447