1. Introduction

Welcome to the series of E-learning modules on De Moivre, Laplace and Levy Theorems. In this module, we are going to cover the concept of De Moivre, Laplace and Levy Theorems, implication of the theorems, contribution of De Moivre, Laplace to Central Limit theorem and applications of the theorem.

By the end of this session, you will be able to:

- Explain the De Moivre, Laplace and Levy Central Limit Theorem
- Explain the aassumptions and implication of the theorem
- Explain the contribution of De Moivre and Laplace to Central Limit Theorem
- Explain the applications of the theorem

We have seen that the sum (or mean) of any number of independent normally distributed random variables is also normally distributed. The following result is also true:
If the sum Sn is equal to summation x k of n independent random variables is normally distributed, then each X_k itself is normally distributed.

We have also seen that for a sequence of independent random variables,
Xn bar minus mu bar tends to zero in the weak or strong sense of probability under certain conditions. However, it does not help us in approximating the distribution of Xn bar for large samples.

In this topic, we consider the conditions under which the distribution of Sn or Xn bar converges to normal distribution.

Now, we present the definition of normal convergence or the Central limit theorem given by De Moivre, Laplace and Levy.
If the distribution of a random variable Y depends on a parameter n and if there exists two quantities mu or sigma (which may or may not depend on n) such that
Limit as n tends to infinity Probability of [Y minus mu by sigma less than or equal to t] is equal to 1 by root 2 phi integral from minus infinity to t exponential (minus x square by 2) dx

Then, we say that Y is asymptotically normally distributed with mean mu and variance sigma square. We always use the statement in Italics as synonymous with the above. We also say that (Y minus mu by sigma) follows the central limit law or normal convergence if the above is true. Above may be true even when the random variable Y has no moments for any n. Consequently, it is not implied that limit of n tends to infinity Expected value of Y is equal to mu nor that limit of n tends to infinity Variance of Y is equal to sigma square

The important implication of the Normal convergence is that for large n, each percentage point of the distribution of Y will be close to the value computed from a Normal distribution with mean mu and variance sigma square, independent of whether Y has these moments or not.
In Normal convergence, the limiting distribution is always continuous (in fact Normal) no matter whether Y is continuous or not.

In the following cases of Central Limit Theorem, we shall consider the sequence of independent random variables $\{X_k\}$ with Expected value of (X_k) is equal to mu k and Variance of (X_k) is equal to sigma k square, such that under given conditions
Sn minus Expected value of Sn by root of Variance of Sn is asymptotically Normal with mean 0 and variance 1 where Sn is equal to summation Xk

Historically, the first CLT was obtained in the first part of eighteenth century by De Moivre who sharpened Bernoulli's theorem and stated that the variables of a Bernoulli sequence obey the Central limit law. The present form is due to Laplace and hence it goes by the name of both being called De Moivre- Laplace Theorem.

Laplace working for about 20 years in this problem was the first to suggest that the Normal distribution is the limiting law of a large variety of cases. The first turning point for the Central Limit problem was the popular Liapounov's theorem given in Nineteen zero one. There have been many studies of the problem, since then all aimed at improving it.

The next significant step in this direction came in nineteen twenty two when Linderberg gave the sufficient condition, which was later in nineteen forty five shown by Feller to be necessary too. The case with dependent variables has also been studied. However, here we shall be concerned with the classical problem, where variables are independent.

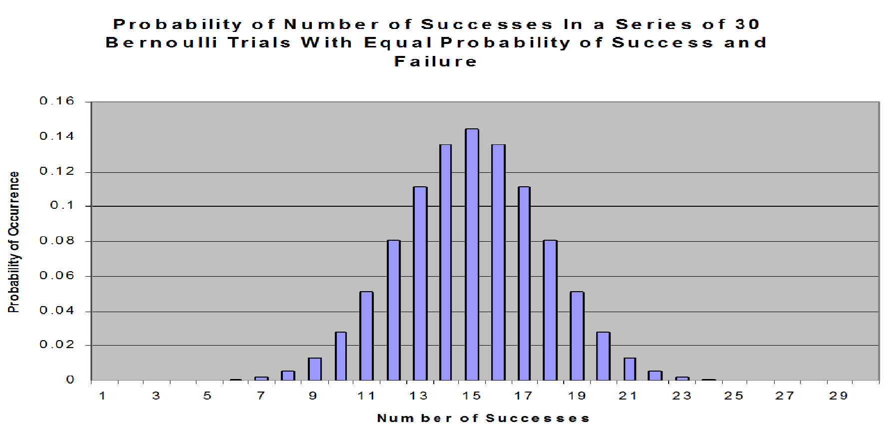
In probability theory, the De Moivre–Laplace theorem is a normal approximation to the binomial distribution. It is a special case of the central limit theorem. It states that the binomial distribution of the number of "successes" in *n* independent Bernoulli trials with probability *p* of success on each trial is approximately a normal distribution with mean *n into p* and standard deviation square root of n into p into q, if *n* is very large and some conditions are satisfied.
The theorem appeared in the second edition of The Doctrine of Chances by Abraham De Moivre, which was published in seventeen thirty eight. The "Bernoulli trials" were not so-called in that book, but rather De Moivre wrote about the probability distribution of the number of times "heads" appears when a coin is tossed three thousand six hundred times.

2. De Moivre Laplace Theorem – CLT

The sum of n independently identically distributed Bernoulli Random variables with success probability p is a binomial Random variable. As n tends to infinity, the Central limit theorem also says that this sum approaches a Gaussian or (Normal) with mean np and variance npq.
As n grows large, for k in the neighbourhood of n p we can approximate
Nck into p to the power k into q to the power n minus k is approximately equal to 1 by square root of (2 into phi into n into p into q) into exponential minus (k minus n into p) whole square by 2 into n into p into q,
P plus q is equal to 1, p is greater than 0 and q is greater than 0
Here the ratio of the left-hand side to the right-hand side converges to 1 as n tends to infinity.

As *n* grows large, the shape of the binomial distribution begins to resemble the smooth Gaussian curve.

Figure 1



In other words, the De Moivre- Laplace Theorem–CLT can be stated as follows:
Let Xk be the indicator function of the event in the k^{th} Bernoulli trial with probability p of occurrence of the event. Then, Sn is equal to summation k is equal to 1 to n, Xk is asymptotically Normal with mean n into p and variance n into p into (1 minus p)
Expected value of (Xk) is equal to p and Variance of (Xk) is equal to p into (1 minus p) less than or equal to 1 by 4
Hence, the result follows from the Linderberg Levy theorem.

Thus, in the case of independently identically distributed random variables (the case of equal components as it is sometimes called) it is sufficient to assume that the common distribution function F of xk has a finite variance for the Central Limit Theorem to hold. However, if the Xk is not identically distributed we need some further conditions for the validity of the Central Limit Law. The purpose of additional conditions is to reduce the probability that an individual Xk will have a relatively large contribution to the sum summation k equal to 1 to n, X k. Such sufficient conditions were provided by Linderberg and Liapounov

3. CLT as a Generalization of Large Numbers

If Xn bar is equal to summation, k runs from 1 to n, Xk by n which is equal to Sn by n and sigma of Sn is equal to Standard deviation of Sn which is equal to sn
Suppose Xn bar minus expected value of Xn bar by root of variance of Xn bar is equal to Sn minus Expected value of Sn by sn is equal to Sn star, where sn is equal to Standard deviation of Sn is a Standard Normal variate with mean 0 and variance unity. If Xk is independent and identically distributed, Variance of (Sn) is equal to n into sigma square, where sigma square is equal to Variance of Xk, mu is equal to Expected value of Xk so that
Sn star is equal to Sn minus n into mu by sigma into root n which is equal to Xn bar minus mu by sigma by root n
Then, Probability of modulus of Sn star less than or equal to x implies Probability of modulus of Xn bar minus mu less than or equal to epsilon tends to 1.

Thus, Xn bar tends to mu in probability, which is the statement of the Weak Law of Large Numbers. Thus, CLT gives the probability bound for modulus of Xn bar minus mu while law of large numbers gives only the limiting value.
For example, Bernoulli Law of large numbers says that Sn by n tends to p in probability. De Moivre Laplace Central limit theorem gives us the bound for the values of S n minus n into p by square root of n into p into q
in terms of Normal probability law, thus providing an estimate of the rate of convergence.

When Xk's are independent but not identically distributed, Central limit theorem might hold but Law of large numbers might not hold or Law of large numbers might hold and Central Limit Theorem might not hold
Since we are concerned with standardized variables Sn star, whose mean and variance are 0 and unity respectively, there is no loss of generality in assuming that the Xk's have mean 0.

4. History of CLT from De Moivre's to Laplace

- De Moivre investigated the limits of the binomial distribution as the number of trials increases without bound and found that the function exponential minus x square came up in connection with this problem
- The formulation of the normal distribution, 1 by root 2 into phi exponential minus x square by 2 came with Thomas Simpson
- This idea was expanded by the German mathematician Carl Friedrich Gauss, who then developed the principle of least squares
- Independently, the French mathematicians Pierre Simon de Laplace and Legendre also developed these ideas. It was with Laplace's work that the first inklings of the Central Limit Theorem appeared
- In France, the normal distribution is known as Laplacian Distribution; while in Germany it is known as Gaussian

Derivation of CLT

- Initial Work: Laplace was calculating the probability distribution of the sum of meteor inclination angles. He assumed that all the angles were random variables following a triangular distribution between 0 and 90 degrees

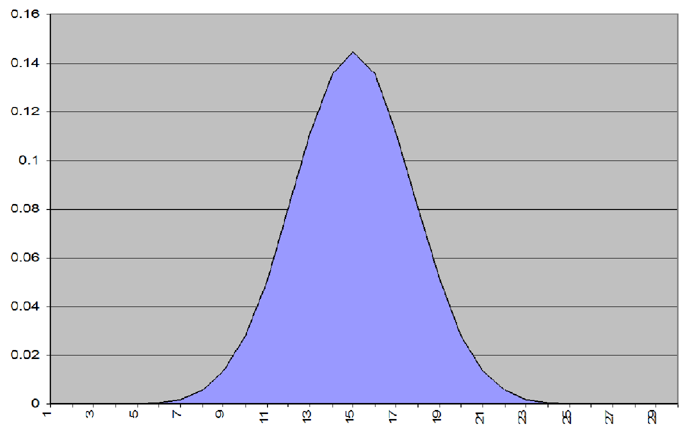
➢ Problems he faced were:

- The deviation between the arithmetic mean which was inflicted with observational errors, and the theoretical value
- The exact calculation was not achievable due to the considerable amount of celestial bodies
  Hence, he tried to obtain solution to the problems by finding an approximation.
- Laplace first introduced the moment generating function, which is known as Laplace Transform of f of x
- He then introduced the Characteristic function:
- Assuming that we have a discrete random variable x, that takes on the values
- Minus m, minus m plus 1, etc, 0, etc, m minus 1, m with probability p minus m, p minus m plus 1, up to, pm
- and Let Sn be the sum of the n possible errors
- It was Lyapunov's analysis that led to the modern characteristic function approach to the Central Limit Theorem

Bernoulli's idea

- In seventeen seventy-seven, Johann Bernoulli took a different approach
- He thought that scientists were more likely to arrive at a measurement near the middle of a reasonable range or error than near the outer edges

Figure 2



The true error curve

• The true curve representing errors is a normal curve discovered by French mathematician Abraham de Moivre in seventeen thirty three. However, it was not recognized as the curve representing errors until Gauss published it in eighteen zero nine

5. De Moivre’s Contribution, Laplace Contribution and Other Applications

De Moivre’s contribution

- Consider a Bernoulli Trial - An experiment with only 2 possible outcomes: success or failure
- De Moivre explored the number of successes that occurred in a given number of trials, focusing only on experiments that had equal chances of success and failure
- He looked at the probability of achieving 1, 2, up to n minus 1, n successes
- De Moivre discovered that as he graphed the probability of each number of successes, it formed a bell-shaped curve
- He then found the following function to represent the curve:

F of x is equal to 1 by sigma into root 2 into phi exponential minus 1 by 2 into sigma square into (x minus mu) whole square
Where, x is equal to the number of successes, mu is the mean, and sigma is the standard deviation

- De Moivre used this *continuous* function to approximate the probability of the *discrete* number of successes
- To do this, he had to integrate the function over the appropriate limits
- To make the integration easier, de Moivre standardized the variable and also needed to adjust the limits
- He standardized the variable using the following formula: X minus mu by sigma which is equal to X star is equal to X minus n by 2 divided by square root of n by 2

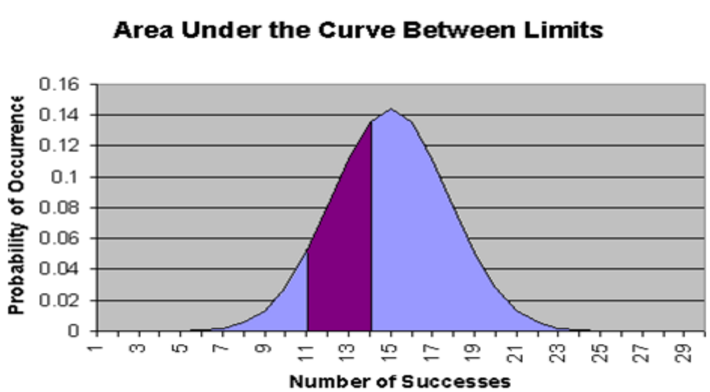
Where, n by 2 is the mean and square root of n by 2 is the standard deviation for the specific Bernoulli Trials he studied.

- So for his approximation, de Moivre used the following formula:

Limit as n tends to infinity Probability of [a star less than or equal to X star less than or equal to b star] is equal to 1 by root 2 into phi exponential minus t square by 2 dt
Where, a star and b star are the appropriate adjusted limits of integration

The graph represents the probability or area under the Normal curve between two limits. In this case, area highlighted is between the eleventh and fourteenth successes.

Figure 3



Laplace contribution

- He expanded De Moivre's approximation to Bernoulli trials with any probability of success
- More importantly, the first Central Limit Theorem was credited to him
- This theorem allowed scientists to find how many samples they needed to make their average reasonably accurate
- Laplace recognized that the sum of independent, identically distributed random variables follows a normal curve as the number of trials, or summands, increases
- For our purposes, "independent, identically distributed random variables" can simply translate to repeating an identical experiment

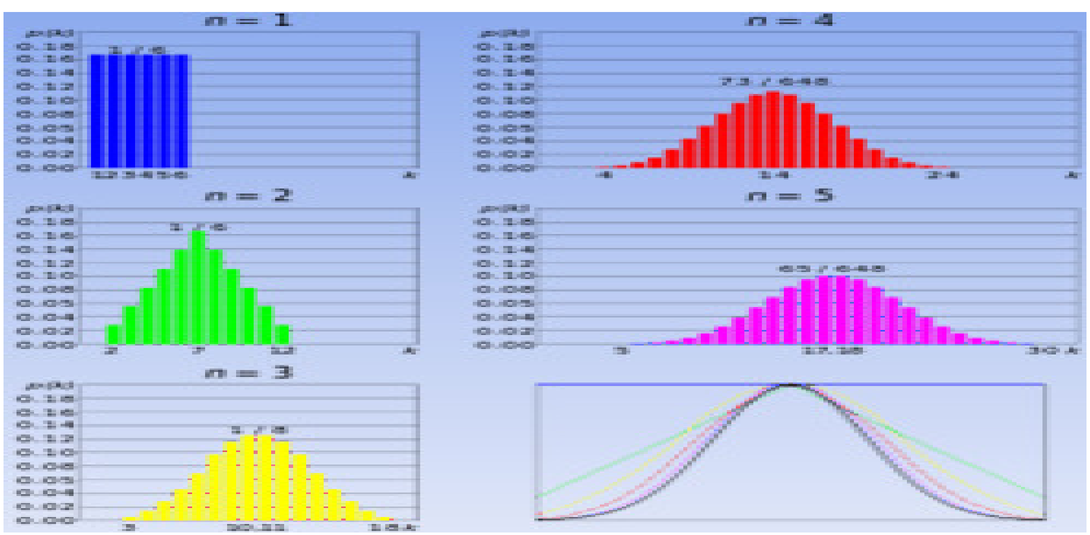
Examples:

- Playing repeated games of blackjack at a casino
- Taking many measurements of the distance between Earth and a given star
- Performing quality tests on many samples of a product at a factory

A simple example of the central limit theorem is rolling a large number of identical, unbiased dice. The distribution of the sum (or average) of the rolled numbers will be well approximated by a normal distribution. Since real-world quantities are often the balanced sum of many unobserved random events, the central limit theorem also provides a partial explanation for the prevalence of the normal probability distribution. It also justifies the approximation of large-sample statistics to the normal distribution in controlled experiments.

Comparison of probability density functions, *p of* (*k*) for the sum of *n* fair 6-sided dice to show their convergence to a normal distribution with increasing *n*, in accordance to the central limit theorem is presented here.

Figure 4



In the bottom-right graph, smoothed profiles of the previous graphs are rescaled, superimposed and compared with a normal distribution (black curve).

Other applications

- Scientists can also use this idea to figure out how many samples they need to ensure ninety-five percent accuracy in their experiments

- Casinos and gamblers can use this idea to figure out their chances of winning or losing a given amount of money after playing a game n number of times

The central limit theorem applies in particular to sums of independent and identically distributed discrete random variables. A sum of discrete random variables is still a discrete random variable, so that we are confronted with a sequence of discrete random variables whose cumulative probability distribution function converges towards a cumulative probability distribution function corresponding to a continuous variable (namely that of the normal distribution).

This means that if we build a histogram of the realisations of the sum of *n* independent identical discrete variables, the curve that joins the centres of the upper faces of the rectangles forming the histogram converge towards a Gaussian or Normal curve as *n* approaches infinity, this relation is known as De Moivre – Laplace Theorem.

Here's a summary of our learning in this session, where we understood:

- The concept of De Moivre – Laplace and Levy Central Limit Theorem
- The implication of the theorem
- The history of Central limit Theorem - De Moivre's to Laplace
- The contribution of De Moivre and Laplace to Central Limit Theorem
- The applications of De Moivre – Laplace and Levy theorem