1. Introduction

Welcome to the series of E-learning modules on Linderberg Theorems. In this module, we are going to cover the concept of Linderberg conditions, statements of Linderberg -Levy Theorem, Linderberg- Lyapunov's Theorem and Linderberg- Feller Theorem and implications and applications of the theorems.

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

- Explain the classical approach of Central Limit Theorem
- Explain the Linderberg conditions
- State the Linderberg-Levy theorem, Lyapunov theorem and Linderberg-Feller Central Limit Theorem
- Explain the applications of the theorems

Whenever we come across the sequences of random variables, we are interested in behaviour of the functions of random variables such as means, variances, and proportions. However, for large samples, exact distributions can be difficult or impossible to obtain.

Limit Theorems can be used to obtain properties of estimators as the sample sizes tend to infinity. Let us recall the concepts we discussed in the previous topics.

- Convergence in probability gives limit of an estimator
- Convergence in distribution gives limit of a cumulative distribution function
- Weak Law of Large Numbers (WLLN): States that a sample mean converges in probability to the population mean mu

We know that if we take a sample of size n from a population whose elements have mean mu and standard deviation sigma, the sample mean x bar will have mean mu and standard deviation sigma by root n. One of the most important results of the probability theory known as CLT states that, for given certain conditions, the mean (thus also the sum) of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed.

The preceding version of the CLT is the most general form and it is shown as summation Xi, which will have an approximate normal distribution even in cases where the random variables Xi have different distributions. In fact, all the random variables tend to be roughly the same magnitude, so that none of them tends to dominate the value of the sum. It can be shown that the sum of the large number of independent random variables will have an approximately normal distribution.

CLT not only gives the method for approximating the distribution of the sum of the random variables but also it helps to obtain remarkable facts that the empirical frequencies of so many naturally occurring populations exhibit a bell shaped (that is normal) curve.

In fact, one of the first uses of the CLT was to provide the theoretical justification of the empirical fact that the measurement errors tend to be normally distributed. That is about an error in measurement as being composed of the sum of a large number of small independent errors. The CLT implies that it should be approximately normal.

For instance, an error of measurement in Astronomy can be regarded as being equal to the sum of the small errors caused by the following things:

- Temperature effects on the measuring device
- Bending of the device caused by the rays of the sun
- Elastic effects
- Air currents
- Air vibrations
- Human errors

Therefore, by CLT the total measurement errors will approximately follow normal distribution. From this, it follows that a histogram of errors resulting from a series of measurements of the same object will tend to follow a bell-shaped normal curve.
The CLT also partially explains why many data sets related to biological characteristics tends to be approximately normal. For instance consider a particular couple, name them Maria and Peter and consider the heights of their daughters (say when they are 20 years old).

Now, the height of the daughter may be composed of the sum of large number of roughly independent random variables – relating, among other things, to the random set of genes that the daughter received from her parents as well as environmental factors. Since each of these variables play only a small role in determining the total height and it is reasonably based on the CLT. Thus, the height of Peter's daughter would be Normally distributed.

If the Peter's family has many daughters, then the histogram of their heights should roughly follow a normal curve. (The same thing is true for sons of Peter and Maria, but the normal curve of the sons would have different parameters from one of the daughters). The CLT cannot be used to conclude that a plot of the heights of all the children would follow normal curve because the gender factor does not play a small role in determining height.

Thus, the CLT can be used to explain why the heights of the daughters of a particular pair of parents will follow a Normal curve. However, the theorem does not explain why a histogram of the heights of the collection of daughters from different parents will follow a Normal curve. To understand why the theorem does not explain the histogram, we will consider the same example with an addition of daughters of Henry and Catherine.

By the same argument given earlier the height of Catherine's daughter will be Normally distributed as well as the height of Maria's daughter. However, the parameters of these two normal distributions will be different. By the same reasoning, we can conclude that the heights of collection of many women form different families will come from different Normal distributions. Therefore, it is evident that a plot of those heights would itself follow a Normal curve.

The central limit theorem has a number of variants. In its common form, the random variables must be identically distributed. In variants, convergence of the mean to the normal distribution also occurs for non-identical distributions, given that they comply with certain conditions.

In more general probability theory, *a* **central limit theorem** is a set of weak-convergence theories. They all express the fact that a sum of many i.i.d. random variables, or alternatively, random variables with specific types of dependence, will tend distribute according to one of a small set of *attractor distributions*.
When the variance of the i.i.d. variables is finite, the attractor distribution is the normal distribution.

2. Classical and Linderberg CLT

Classical CLT
Let X1, up to Xn be a random sample of size n, that is, a sequence of i.i.d random variables drawn from distributions of expected values given by mu and finite variances given by sigma square.

Suppose we are interested in the sample average of these random variables. Then, by the law of large numbers, the sample averages converge in probability and is almost closer to the expected value mu as n tends to infinity. The classical central limit theorem describes the size and the distributional form of the stochastic fluctuations around the deterministic number mu during this convergence.

More precisely, it states that as *n* gets larger, the distribution of the difference between the sample average Xn bar and its limit mu, when multiplied by the factor root n (that is (root n into X n bar minus mu)), approximates the normal distribution with mean 0 and variance sigma square. For large enough n, the distribution of Sn is close to the normal distribution with mean n mu and variance *n sigma square.* The usefulness of the theorem is that the distribution of root n (Xn bar *minus* *mu*) approaches normality regardless of the shape of the distribution of the individual X_i's.

The first turning point for the Central Limit problem was the popular Lyapunov'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 necessary by Feller too.

It is the Linderberg-Feller Theorem, which makes the statement of CLT precise in providing the sufficient and necessary Linderberg condition whose satisfaction accounts for the smooth appearance of the bell shaped normal curve.
Here we provide a simpler, equivalent, and more easily interpretable probabilistic formulation of the Linderberg condition and demonstrate its sufficiency and partial necessity in the CLT using more elementary means.

The seeds of the Central Limit Theorem lie in the work of Abraham de Moivre, who, in seventeen thirty three, not being able to secure himself an academic appointment, supported himself consulting on problems of probability, and gambling. He approximated the limiting probabilities of the Binomial distribution, the one which governs the behaviour of the number Sn of success in an experiment, which consists of n independent trials, each one having the same probability p belongs to (0, 1) of success.

Linderberg proves sufficiency in Nineteen twenty two and necessity by Feller in Nineteen thirty five. Linderberg-Feller CLT is one of the most far-reaching results in probability theory. Nearly all generalizations of various types of central limit theorems spin from Linderberg-Feller CLT, such as, for example, CLT for martingales, for renewal processes, or for weakly dependent processes. The insights of the Linderberg condition are that the wild values of the random variables, compared with Sn, the standard deviation of Sn as the normalizing constant, are insignificant and can be truncated off without affecting the general behaviour of the partial sum Sn.

The following gives a self-contained treatment of the central limit theorem. It is based on Lindeberg's method. We assume that $X_{1,}$ up to X_n are independent random variables with means 0 and respective variances unity.

Linderberg CLT
For every epsilon greater than 0
Limit as n tends to infinity of 1 by sn square summation Expected value of (Xi minus mu i whole square into I for all modulus of (Xi minus mu i greater than epsilon into sn is equal to zero
Where I is the indicator function. Then the distribution of the standardized sums 1 by sn into summation (Xi minus mu i) converges towards the standard normal distribution with mean 0 and variance 1

3. Linderberg – Levy Theorem (CLT) and Lyapunov's Theorem

Linderberg – Levy Theorem (CLT)
Let {Xk} be the sequence of independent and identically distributed random variables with Expected value of Xk is equal to mu and variance of Xk is equal to sigma square which is finite, then
Summation k runs from 1 to n (Xi minus mu by sigma into root n) is asymptotically Normal with mean 0 and variance 1

In other words, as *n* approaches infinity, the random variables root n into (Xn bar minus mu) converge in distribution to a Normal, that is Normal with mean 0 and variance sigma square. That is
Root n into (summation Xi by n minus mu) converges in distribution to Normal distribution with mean 0 and variance sigma square.

In the case sigma greater than 0, convergence in distribution means that the cumulative distribution functions of root n into (Xn bar minus mu) converge point wise to the cumulative distribution function of the *Normal distribution with* mean 0 and variance sigma square. For every real number *z*,
Limit as n tends to infinity Probability of root n into (Xn bar minus mu) less than or equal to z is equal to Phi of z by sigma
Where, Phi of x is the standard normal density function evaluated at *x*. Note that the convergence is uniform in *z* that is
Limit as n tends to infinity Supremum as z belongs real R Probability of root n into (Xn bar minus mu) less than or equal to z minus Phi of z by sigma is equal to 0
Where, supremum denotes the least upper bound (or supremum) of the set.

The Linderberg–Levy central limit theorem relates to the statistic, root n into (Xn bar minus mu) and not directly to *X*. There are numerous CLTs. They differ in the assumptions required for their use. The Linderberg–Levy CLT is a particularly simple one and sufficient for most of the analysis.

Thus, in the case of i.i.d 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. But 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 Xk. Such sufficient conditions were provided by Linderberg and Lyapunov.

Lyapunov's Theorem: CLT
Let {Xk} be the sequence of independent random variables with with Expected value of Xk is equal to mu k and variance of X k is equal to sigma k square and Expected value of modulus of Xk minus mu k to the power 2 plus delta is finite, then summation (Xk minus mu

k by Cn) is asymptotically Normal with mean 0 and variance 1 provided
Limit as n tends to infinity of 1 by Cn to the power 2 plus delta summation Expected value of modulus of Xk minus mu k to the power 2 plus delta is equal to zero , for some delta , 0 less than delta less than or equal to 1 where Cn square is equal to summation sigma k square

Sometimes a particular case of the above theorem with delta equal to 1is also stated as Lyapunov's theorem. However, these theorems are not satisfactory, because the moments of higher order are used. A sufficient condition which is almost necessary is given by the Linderberg – Feller theorem.

4. Linderberg – Feller Theorem CLT

Let {Xk} be the sequence of independent random variables with with Expected value of Xk is equal to mu k and variance of Xk is equal to sigma k square which is finite and Fk be the distribution function of Xk then,

- Summation (Xk minus mu k by Cn) is asymptotically Normal with mean 0 and variance 1 and
- Limit as n tends to infinity maximum for all k less than or equal to n sigma k by Cn is equal to zero, holds if and only if for all epsilon greater than 0

bn of epsilon is equal to 1 by Cn square summation k runs from 1 to n, Integral over modulus of X minus mu k greater than or equal to epsilon Cn of X minus mu k whole square d F k of x tends to zero as n tends to infinity. Call this as (1)
In the above statement, Cn square is equal to summation sigma k square.

In other words, Linderberg-Feller Central Limit Theorem can also be stated as
If the random variates X1, X2 etc satisfy the Linderberg condition, then for all a less than b, Limit as n tends to infinity Probability of (a less than Sn minus E of Sn by sn less than b) is equal to Phi of b minus Phi of a
where Phi is the normal distribution function.

We may note that Lyapunov's condition implies Linderberg condition, but Lyapunov's condition is easy to verify in actual practice. Given a sequence of random variables, we first try to see whether Lyapunov's condition is satisfied and if the answer is in negative, one proceeds with the verification of Feller's condition. As stated earlier, Linderberg Fellers condition is sufficient for CLT which implies condition of Linderberg Feller theorem also.

In statistics, a theory stating that the sample size of identically distributed random numbers approaching infinity, it is more likely that the distribution of the numbers will approximate normal distribution.
That is, the mean of all samples within that universe of numbers will be roughly the mean of the whole sample.

5. Applications

- *If the* X_k's are Gaussian*, then* the sum S is Gaussian and as n tends to infinity, S is again Gaussian
- If the X_k's are Binomial, *S* is Binomial and as n tends to infinity, S is Gaussian
- If the X_k's are Poisson, *S* is Poisson and as n tends to infinity, S is Gaussian
- If the X_k's are Gamma, *S* is Gamma and as n tends to infinity, S is Gaussian
- If the X_k's are Negative Binomial, *S* is Negative Binomial and as n tends to infinity, S is Gaussian

However, CLT is not applied for Cauchy distribution
If the X_k's are Cauchy, *S* is Cauchy but as n tends to infinity, S is still Cauchy and not Gaussian or (Normal)

The Central Limit Theorem is one of the most striking and useful results in probability and statistics. It explains why the normal distribution appears in areas as diverse as gambling, measurement error, sampling, and statistical mechanics. In essence, the Central Limit Theorem states that the normal distribution applies whenever one is approximating probabilities for a quantity, which is a sum of many independent contributions all of which are roughly the same size.

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

- The importance of Central Limit Theorem
- The classical approach of the theorem
- The statement of Linderberg –Levy, Lyapunov, Linderberg- Feller Central Limit Theorems
- The implications and applications of the theorems