BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024

Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Measure and Probability Theory

Course Code: 23ST02CC

Unit-V

Central Limit Theorem

Dr. T. Jai Sankar
Associate Professor and Head
Department of Statistics

Ms. S. Soundarya
Guest Faculty
Department of Statistics

UNIT – V

LIMIT THEOREM

Central limit theorem

The central limit theorem (CLT) plays an important role in statistical theory. It is one of the usual assumptions in statistics to assume that the underlying observations follow the normal distribution at least approximately. Also, the theory of errors/residuals used by physicists or astronomers can be justified on the basic of central limit theorem.

The first version of central limit theorem was given by the French born mathematician Abraham De Moivre in 1733. Later Laplace extended De Moivre‟s finding by approximating the binomial distribution with the normal distribution. The contribution of central limit theorem was first made by Laplace in 1812 and a theoretical prove based on general condition was given by Liapounoff in 1901. Thus central limit theorem is also named as second fundamental theorem of probability.

Definition

It states that "The distribution of means of random samples taken from a population having mean μ and finite variates σ^2 approaches the normal distribution with mean μ and variance σ^2/n, as $n \to \infty$"(or)

The sum of the large number of independent random variables has a distribution that is approximately normal.

Uses of Central Limit Theorem

- The central limit theorem explains the sort of relationship between the shape of the population distribution and the mean of sampling distribution.
- It is great use of statistical inference, where it provides the use of sample statistics to make inferences about the population parameters.

- It helps to explain the remarkable fact of the empirical frequencies of many natural populations exist bell shape normal curves.
- Many of the distribution of sample statistic approaches to normality for large samples d such as they can be studied with the help of normal curves.
- The normal distribution has larger applications in statistical quality control in industry for setting control limits and taking decisions.

Statement of CLT

If X_i $(i = 1, 2, ..., n)$ be independent random variables such that $E(X_i) = \mu_i$ and $V(Xi) = \sigma_i^2$ then under certain very general conditions, the random variables $S_n = X_1 + X_2 + ... + X_n$ is asymptotically normal with mean μ and standard deviation σ where,

$$\mu = \sum_{i=1}^{n} \mu_i \text{ and } \quad \sigma^2 = \sum_{i=1}^{n} \mu_i^2$$

Central Limit theorem for (independent and identically distributed) variables was proved by Linderberg and Levy. If $X_1, X_2, ...X_n$ are independent and i.i.d random variables with $E(X_i) = \mu_i$ and $V(X_i) = \sigma_i^2$ $i = 1,2,3 .....n$ then the sum $S_n = X_1 + X_2 + ... + X_n$ is asymptotically normal with mean $\mu = n\mu_1$ and variance $\sigma_i^2 = n\sigma_i^2$.

Demoivre's Laplace Central Limit theorem

Statement:

A particular case of central limit theorem is De-Moivre's theorem which states as follows:

If Xi = 1 with probability p, 0 with probability q

then the distribution of the random variable $S_n = X_1 + X_2 + ... + X_n$, where X_i's are independent, is asymptotically normal as $n \to \infty$.

Proof.:

M.G.F. of X_i, is given by: $Mx_i(t) = E(e^{tx_i}) = e^{t1}p + e^{t\,0}q = (q + pe^t)$

M.G.F. of the sum $S_n = X_1 + X_2 + ... + X_n$, is given by

$Ms_n, (t) = M_{x1+X2+-...+xn,} (t) = Mx_1(t)\ Mx_2(t) .... Mx_n(t) = [Mx_i (t)]^n = (q+pe^t)$",

which is the M.G.F. of a binomial variate with parameters n and p.

$$E(S_n) = np = \mu \text{ (say), and } V (S_n) = npq = \sigma^2, \text{ (say).}$$

Let

$$Z = S_n - \mu / \sigma$$

$$M_z(t) = [1+t^2/2n + O(n^{-3/2})]^n$$

Where $O(n^{-3/2})$] represents terms involving 13/2 and higher powers of n in the denominator.

$$\text{Lim } M_Z(t) = e^{t2/2}$$

which is the M.G.F. of a standard normal variate. Hence by the uniqueness theorem of M.G.F.'s

$$Z = S_n - \mu / \sigma$$

Hence, $S_n = X_1 + X_2 + ... + X_n$, is asymptotically $N(\mu , \sigma^2)$ as $n \to \infty$.

Lindeberg-Levy's Central Limit Theorem

Statement:

If $X_1, X_2, ..., X_n$ are independently and identically distributed random variables with

$$E (X_i) = \mu_1$$

$$V (X_i) = \sigma_1^2, i= 1, 2, ..., n$$

then the sum $S_n = X_1+X_2+ ... + X_n$, is asymptotically normal with mean $\mu = \eta\mu_1$ and variance $\sigma^2 = \sigma_1^2$

Liapounov forms of Central Limit Theorem

Let $X_1, X_2, ..., X_n$ be independent random variables such that

$E(X_i) = \mu_i$

$V(X_i) = \sigma_i^2$, i = 1, 2, 3, ..., n

Let us suppose that 3^{rd} absolute moment, say P_i^3 of X_i about its mean exists & is finite.

$P_i^3 = E\{|X_i - \mu_i|^3\}$, i = 1, 2, 3, ..., n

Further, let $P^3 = \sum P_i^3$, i = 1, 2, 3, ..., n

If lim (n→∞) (P/σ) = 0, the sum $X = X_1 + X_2 + ... + X_n$ is asymptotically $N(\mu, \sigma^2)$, where

$\mu = \sum \mu_i$, i = 1, 2, 3, ..., n

$\sigma^2 = \sum \sigma_i^2$, i = 1, 2, 3, ..., n

Lindeberg Feller's Central Limit Theorem

Let $X_1, X_2, ..., X_n$ be iid non-degenerate r.v's with the distribution function $F_1, F_2, ..., F_n$ with means $\mu_1, \mu_2, ..., \mu_n$ and finite variance $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$ respectively. Let $B_n^2 = \sigma_1^2 + \sigma_2^2 + ... , + \sigma_n^2$ and $S_n = X_1 + X_2 + ... + X_n$.

If F_i are absolutely continuous with pdf $f_i(x)$ and if the condition

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^{n} \int_{A_n} (x - \mu_i)^2 f_i(x) dx = 0 \text{ where } A_n = |X - \mu_i| > \in B_n \text{ holds for all } \varepsilon > 0. \text{ Then}$$

$$\lim_{n\to\infty} P\left\{ \frac{S_n - E(S_n)}{B_n} \le a \right\} = \phi(a): \text{ cdf of } N(0, 1).$$

Law of Large Numbers (LLN):

A law of large numbers is a proposition given a set of conditions that are sufficient to guarantee the convergence of the sample mean $\bar{X}n$ to a constant as the sample size n increases.

Weak law of large numbers (WLLN)

Strong law of large numbers (SLLN)

1. Strong law of large numbers:

If the sequence $\{X_n\}$ converges almost surely, the law of large numbers is called a strong law of large numbers.

Let $X_1, X_2, \ldots$ be independent and identically distributed with $E[|X_1|] < \infty$ (In fact, pairwise independence suffices)

Let $S_n = \Sigma X_k$ and $\mu = E[X_1]$. Then,

$$S_n/n \to \mu \text{ a.s}$$

If instead $E[X_1] = \infty$, then

$$P(\lim S_n \text{ exists} \in (-\infty,+\infty)) = 0$$

2. Weak law of large numbers (WLLN):

If the sequence $\{X_n\}$ converges in probability, then the Law of large numbers is called a weak law of large numbers.

Weak law of large numbers:-

Let $(X_n)_n$ be independent and identically distributed and $S_n = \Sigma X_k$

A necessary and sufficient condition for the existence of constants $(M_n)_n$ such that

$$(S_n - M_n)/n \to P\ 0$$

is $n\ P\ [|X_1| > n] \to 0$

In that case, the choice

$$M_n = E[X_1\ I\ |X_1| \leq n].$$

Khintchine's Weak Law of Large Number

Khintchine's Theorem (WLLN):

(If Xi's are not identically and independently distributed the only condition necessary for the law of large number to hold is that E(Xi), i= 1,2,... exists.

Theorem:-

Let {Xn} be any sequence of rv's, write:

$$Yn = \{Sn-E(Sn)\}/n$$

where, $Sn=X_1+X_2+...+X_n$

A necessary and sufficient condition for sequence {Xn} to satisfy the WLLN is

$$E\{Yn^2/(1+|Yn|)\} \to 0 \text{ as } n\to\infty \text{ ---->(1)}$$

Proof:-

Let us assume that (1) holds, we shall prove {Xn} satisfies WLLN for real numbers a,

b

$a\geq1$ $a=b=0$, we have $a\geq b \to a+ab=b+ab$ ----> (2)

Let us define the event A,

$$A = \{\eta \varepsilon \Omega \mid |Xn| \geq \varepsilon \}$$

we have

$$\sum P(|Yn| \geq \varepsilon) < \infty$$

Taking $a = y_n^2$ and $b = \varepsilon^2$ in (2), we define another event B as follows:

$$B = \{ y_n^2 / (1 + y_n^2) \geq \varepsilon^2 / (1 + \varepsilon^2) \} = \{ y_n^2 \geq \varepsilon^2 \}$$

Since $w \in A$

$\Rightarrow w \in B, A \subseteq B$

$\Rightarrow P(A) \leq P(B)$

$P(|Y_n| \geq \varepsilon) \leq P(y_n^2 / (1 + y_n^2) \geq \varepsilon^2 / (1 + \varepsilon^2)) \leq E \{ y_n^2 / (\varepsilon^2(1 + y_n^2)) \}$

$\Rightarrow 0$ as $n \to \infty$

$P(|Xn| \geq \varepsilon) \to 0$ as $n \to \infty$

lim P{ | Sn - E(Sn) | ≥ ε } → 0.

n→∞ n

Then WLLN holds for the sequence {Xn} of r.v's conversely

of {Xn} satisfies WLLN, we shall establish (1)

Let us assume that Xi's are continuous and let Xn have

pdf fn(y). Then,

$$E(y_n^2 /1 + y^2) = \int y_n^2/1 + y^2 f_n (y)dy$$

$$(1 + y^2) fn(y) dy.$$

where $A = \{y \mid y^2 / (1 + y^2) \leq \varepsilon\}$ and $A^c = \{y \mid y^2 / (1 + y^2) > \varepsilon\}$

$$E(Y_n^2 / (1 + Y_n^2)) = \int y^2 / (1 + y^2) f_n(y) dy$$

$$\leq \int_A f_n(y) dy + \int_{A^c} y^2 / (1 + y^2) f_n(y) dy$$

$$\leq P(A) + \varepsilon^2 \int_{A^c} f_n(y) dy$$

$$[\because\ y^2 / (1 + y^2) \leq 1 \text{ and } y^2 / (1 + y^2) > \varepsilon \text{ on } A^c]$$

$$= P(A) + \varepsilon^2 P(A^c) \leq P(A) + \varepsilon^2$$

$[\because \; P(A^c) \leq 1]$

But, since $\{X_n\}$ satisfies WLLN, we have,

$$\lim_{n\to\infty} P(|Y_n| \geq \varepsilon) \to 0$$

and since ε is arbitrarily small positive number,

we get on taking limits

$$\lim_{n\to\infty} (Y_n^{\alpha} / 1+Y_n^{\alpha}) \to 0$$

Hence the proof

Kolmogrove's inequality.

Let $X_1,X_2,...,X_n$ be a sequence of independent random variables with mean zero (i.e., $E[Xi]=0$ for all □) and finite variances $Var(Xi)=\sigma_i^2$. Define the partial sum:

$$S_k = \sum_{i=1}^{k} X_i \text{ for } k = 1, 2, ..., n$$

$$P(\max_{1\leq k\leq n} |S_k| \geq \lambda) \leq \frac{\sum_{i=1}^{n} \sigma_i^2}{\lambda^2}$$

This inequality provides a bound on the probability that the maximum absolute value of the partial sums exceeds a given threshold .

Sk exceeds a given threshold λ

Proof:

Consider the event

$A_k = \{|S_k| \geq \lambda\}$, where $|S_k| \geq \lambda$ for some $k \in N$

Define T as the smallest index k such that

$$|S_k| \geq \lambda.$$

Then $T \leq n$.

If the event $\max_{1 \le k \le n} |S_k| \ge \lambda$ occurs,

we use

$E[S_k^2] = \sum_{i=1}^{k} \sigma_i^2$ due to independence and zero-mean.

Let A be the event

$\max_{1 \le k \le n} |S_k| \ge \lambda$. Then,

$$\sigma^2 = E[S_n^2] = E[S_n^2|A]P(A) + E[S_n^2|A^c]P(A^c)$$

$$\sigma^2 \ge \lambda^2 P(A)$$

$$P(\max_{1 \le k \le n} |S_k| \ge \lambda) \le \sigma^2/\lambda^2$$

Hence proved.