CHAPTER 12

The Law of Large Numbers

12.1 INTRODUCTION

The law of large numbers attempts to provide a philosophical justification for all attempts to estimate a probability experimentally. It justifies the relative frequency theory of probability. Moreover, it uses the mean and variance of the random variable to extract further information about the random variable or more precisely about the sequence of random variables of similar type. In this chapter we will study some special cases of this law and will briefly discuss its various implications, particularly in terms of experimental measurements.

12.2 THE WEAK LAW OF LARGE NUMBERS

In 1837, Poisson published the first general formulation of a certain scientific law, which is now known as the empirical law of large numbers because it applies to the outcomes of a large number of trials of an experiment. Let us consider a probability space $(\Omega, \mathcal{A}, P(\bullet))$ and an infinite sequence $X_1, X_2, X_3, \ldots$ of random variables on Ω are given. We introduce the new random variables S_n and $\overline{X}_n$ for $n = 1, 2, \ldots$ We define,

$$S_n = X_1 + X_2 + \ldots + X_n$$

and

$$\overline{X}_n = \frac{1}{n} S_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$$

First we consider the special case in which $X_1, X_2, \ldots$ are independent and identically distributed random variables.

In this case all the random variables X_k have exactly the same distribution and hence the same mean μ and variance σ^2.

Thus, $E(S_n) = n\mu$ and $E(\overline{X}_n) = \mu$.

Also, as $X_1, X_2, \ldots$ are independent so,

$$\text{var}(X_i X_j) = 0 \text{ for } i \neq j, \text{ thus, var}\left(\overline{X}_n\right) = \frac{\sigma^2}{n}.$$

Under the above situation the weak law of large number holds, which is stated as follows :

Let $\{X_n\}$ be a sequence of *iid* variates, with $E(X_i) = \mu < \infty$, then

$$P\left[\left|\frac{S_n}{n}-\mu\right|>\in\right]\to 0 \quad \text{as} \quad n\to\infty$$

i.e.,
$$P\left[\left|\overline{X}_n-\mu\right|>\in\right]\to 0 \quad \text{as} \quad n\to\infty$$

So, $\overline{X}_n \xrightarrow{p} \mu$, *i.e.*, $\overline{X}_n$ converges to μ in probability. ...(12.1)

The theorem was first published by Khintchin in 1929 so sometimes also called as the Khintchin's law of large numbers.

Proof : Here, $S_n = X_1 + X_2 + \ldots + X_n$ and $\overline{X}_n = S_n/n$

Then using the addition theorem of characteristic function we have

$$\begin{aligned}\phi_{\overline{X}_n}(t) &= \phi_{\frac{S_n}{n}}(t) = \phi_{S_n}\left(\frac{t}{n}\right)\\ &= \left[\phi_{X_i}\left(\frac{t}{n}\right)\right]^n = \left[E\left(e^{i\frac{t}{n}X_i}\right)\right]^n\\ &= \left[1+i\mu\left(\frac{t}{n}\right)+0\left(\frac{t}{n}\right)\right]^n\end{aligned}$$

An, $n\to\infty$, the RHS tends to $e^{i\mu t}$, hence

$$\lim_{n\to\infty}\phi_{\overline{X}_n}(t)=e^{i\mu t}$$

Which is the characteristic function of a degenerate variate X such that $P(X=\mu)=1$

The *cdf* of which is given by

$$\begin{aligned}F(x) &= 1 \text{ if } X\geq\mu\\ &= 0 \text{ if } X<\mu\end{aligned}$$

The *cdf* is continuous everywhere except at $X=\mu$. Choosing $\in>0$, we can have

$$\begin{aligned}\lim_{n\to\infty}P\left\{\overline{X}_n\leq\mu-\in\right\} &= \lim_{n\to\infty}F_{\overline{X}_n}(\mu-\in)\\ &= F_X(\mu-\in)=0\end{aligned} \quad ...(12.2)$$

$$\begin{aligned}\lim_{n\to\infty}P\left\{\overline{X}_n\leq\mu+\in\right\} &= \lim_{n\to\infty}F_{\overline{X}_n}(\mu+\in)\\ &= F_X(\mu+\in)=1\end{aligned} \quad ...(12.3)$$

$\therefore$
$$\lim_{n\to\infty}P\left\{\overline{X}_n>\mu+\in\right\} = 0 \text{ by } (12.2)$$

Thus from (12.2) and (12.3) it follows that,

$$\lim_{n\to\infty}P\left\{\left|\overline{X}_n-\mu\right|>\in\right\} = 0$$

i.e.,
$$P\left[\left|\frac{S_n}{n}-\mu\right|>\in\right]\to 0 \text{ as } n\to\infty$$

12.3 A MORE GENERALIZED WEAK LAW

Chebyshev forwarded a more generalized definition to the weak law of large numbers. Let $X_1, X_2, \ldots$ be a sequence of random variables having expectations, $\mu_1, \mu_2, \ldots$. Further let,

$$V_n = \text{Var}(X_1 + X_2 + \ldots + X_n) \text{ if } \frac{V_n}{n^2} \to 0$$

as $n \to \infty$ then, given two positive quantities $\in$ and η however small we can find an n_0 depending on $\in$ and η such that

$$P\left[\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n}\right| \le \in\right] > 1 - \eta \qquad \ldots(12.4)$$

for all $n \ge n_0$

Proof : We have

$$E\left[\frac{X_1 + X_2 + \ldots + X_n}{n}\right] = \frac{1}{n} E(X_1 + X_2 + \ldots + X_n)$$

$$= \frac{1}{n}[E(X_1) + E(X_2) + \ldots + E(X_n)]$$

$$= \frac{1}{n}(\mu_1 + \mu_2 + \ldots + \mu_n)$$

and, $\text{var}\left(\frac{X_1 + X_2 + \ldots + X_n}{n}\right) = \frac{1}{n^2} V(X_1 + X_2 + \ldots + X_n) = \frac{V_n}{n^2}$

Now, by Chebychev's inequality we have

$$P[|X - \mu| \le t\sigma] > 1 - \frac{1}{t^2}$$

Thus, $P\left[\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n}\right| \le \frac{t\sqrt{V_n}}{n}\right] > 1 - \frac{1}{t^2}$

So, for any positive quantity t choosing it in such a way that, $\frac{t\sqrt{V_n}}{n} = \in$ we have

$$P\left[\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n}\right| \le \in\right] > 1 - \frac{V_n}{n^2 \in^2}$$

Since, $\frac{V_n}{n^2} \to 0$ as $n \to \infty$, given $n\in^2 > 0$, however small, one can find an n_0 depending on $\in$ and η, such that

$$\frac{V_n}{n^2} < \eta \in^2 \text{ for all } n \le n_0.$$

For such an n_0 we have

$$P\left[\left|\frac{X_1 + X_2 + \ldots + X_n}{n} - \frac{\mu_1 + \mu_2 + \ldots + \mu_n}{n}\right| \le \in\right] > 1 - \eta$$

whenever $n \ge n_0$.

12.4 BERNOULLI LAW OF LARGE NUMBERS

If the outcome A occurs n_A times during n identical trials of an experiment and if n is large enough, then the relative frequency n_A/n should be close to the probability p of A. Thus we may translate the concept of identical trails to independent repeated Bernoulli trials with probability p of success.

Thus, we write X_k equal to 1 or 0 according to whether the outcome of the kth trial is A or not, $p = P(X_k = 1)$ and $q = P(X_k = 0) = 1 - p$.

Then $\overline{X}_n$ is a random variable, which represents the relative frequency of A, $\mu = E(X_k) = E(\overline{X}_n) = p$ and so, using WLLN

$$\underset{n\to\infty}{\text{Lt}}\ P\left(\left|\overline{X}_n - p\right| > \epsilon\right) = 0 \quad \text{...(12.5)}$$

This is the Bernoulli law of large numbers.

Statement : Let $\{X_n\}$ be a sequence of independent and identically distributed Bernoulli random variables with $P(X_i = 1) = p$, $P(X_i = 0) = 1 - p = q$ with $0 < p < 1$. Then for every $\epsilon > 0$, we have

$$\lim_{n\to\infty} P\left(\left|\overline{X}_n - p\right| < \epsilon\right) = 1$$

Proof : We have

$$E(\overline{X}_n) = E\left(\frac{X_1 + X_2 + ... + X_n}{n}\right)$$

$$= \frac{p + p + ... n \text{ times}}{n} = \frac{np}{n} = p$$

$$\text{Var}(\overline{X}_n) = \text{Var}\left(\frac{X_1 + X_2 + ... + X_n}{n}\right) = \frac{1}{n^2}\ \text{var}\,(X_1 + X_2 + ... + X_n)$$

$$= \frac{npq}{n^2} = \frac{pq}{n}$$

Now, by Chebychev's inequality for every $\epsilon > 0$, we have

$$P\left(\left|\frac{\overline{X}_n - p}{\sqrt{pq/n}}\right| > \epsilon\right) \le \frac{1}{\epsilon^2}$$

and hence, $P\left[\left|\overline{X}_n - p\right| \ge \epsilon\right] \le \dfrac{pq}{n\,\epsilon^2}$

Thus, $\underset{n\to\alpha}{\text{Lim}}\ P\left[\left|\overline{X}_n - p\right| \ge \epsilon\right] = 0$

Since, in practice the unknown p has to be estimated empirically, the above theorem asserts that in order to have an agreement between p and its estimate $\overline{X}_n$ we need to have a large number of observations.

12.5 THE STRONG LAW OF LARGE NUMBERS

The strong law is more important than the weak law of large numbers, because, average of a sequence of independent random variables having a common distribution will, with probability 1, converge to the mean of that distribution.

Statement : Let $X_1, ..., X_n$ be independent random variables such that $E(X_k) = 0$ and var $(X_k) = \sigma_k^2 < \infty, k = 1, 2, ..., n$. Let

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

Then for all $\epsilon > 0$, we have

$$P\left[\max_{1\le k<n} |S_k| \ge \epsilon\right] \le \frac{1}{\epsilon^2}\sum_{k=1}^{n} \sigma_k^2 \quad ...(12.6)$$

Proof : Consider

$$\sum_{k=1}^{n} \sigma_k^2 = 1$$

So,

$$\sum_{k=1}^{n} \sigma_k^2 = \sum_{k=1}^{n} E\left(X_k^2\right) = E\left[\sum_{k=1}^{n} X_k^2\right]$$

$$= E\left[\left(\sum_{k=1}^{n} X_k^2\right) + 2\sum\sum_{k<k'} X_k X_{k'}\right]$$

$$= E\left(\sum_{k=1}^{n} X_k^2\right) = E\left(S_n^2\right) \quad ...(12.7)$$

Consider the event, $[\max |S_k| \ge \epsilon]$

$$= \bigcup_{k=1}^{n} (|S_k| \ge \epsilon)$$ $[\max |S_k| \le \epsilon]$ in (12.6), we have

$$= \bigcup_{k=1}^{n} A_k, \text{ where } A_k = [|S_k| \ge \epsilon]$$

$$= A_1 \cup A_1^c A_2 \cup A_1^c A_2^c A_3 \cup ...... \cup A_1^c A_2^c .... A_{n-1}^C A_n$$

$$= [|S_1| \ge \epsilon] \cup [|S_1| < \epsilon, |S_2| \ge \epsilon] \cup [|S_1| < \epsilon, |S_2| < \epsilon, |S_3| \ge \epsilon]$$
$$\cup ..... \cup [|S_1| < \epsilon, ..., |S_{n-1}| < \epsilon, |S_n| \ge \epsilon]$$

$$= \bigcup_{k=1}^{n} B_k, B_k = [|S_1| < \epsilon, |S_2| < \epsilon, ..., |S_{k-1}| < \epsilon, |S_k| \ge \epsilon]$$

So, B_k's are disjoint events, therefore, we have

$$\max_{1 \le k \le n} [|S_k| \ge \epsilon] = \bigcup_{k=1}^{n} [|S_n| \le k] = \sum_{k=1}^{n} B_k$$

where,

$B_k = [\,|S_1| < \epsilon, \ldots, |S_{k-1}| < \epsilon, |S_k| \ge \epsilon]$

So, B_k's are disjoint events. ....(12.8)

Consider, now,

$$\sum_{k=1}^{n} \sigma_k^2 = E\left[S_n^2\right] = \int S_n^2 \, dG(x)$$

where G is the distribution function of S_n

$$= \int_{(\max |S_k| \ge \epsilon)} S_n^2 \, dG(x) + \int_{(\max |S_k| \le \epsilon)} S_n^2 \, dG(x) \ge \int_{(\max |S_k| \ge \epsilon)} S_n^2 \, dG(x)$$

$$\Rightarrow \quad \sum_{k=1}^{n} \sigma_k^2 \le E\left[S_n^2 I_{\left[\max\limits_{1 \le k \le n} |S_k| \ge \epsilon\right]}\right] = E\left[S_n^2 I_{\sum\limits_{k=1}^{n} B_k}\right] \text{ using (12.8)} \quad \ldots(12.9)$$

or $$E\left[S_n^2 \left(I_{B_1} + I_{B_2} + \ldots + I_{B_n}\right)\right] = \sum_{k=1}^{n} E\left[S_n^2 I_{B_k}\right]$$

Consider,

$$\begin{aligned} E\left[S_n^2 I_{B_k}\right] &= E\left[\left(\overline{S_n - S_k} + S_k\right)^2 I_{B_k}\right] \\ &= E\left[\left\{(S_n - S_k)^2 + S_k^2 + 2(S_n - S_k)S_k\right\} I_{B_k}\right] \\ &\ge 2\,E[(S_n - S_k)S_k I_{B_k}] + E\left[S_k^2 I_{B_k}\right] \end{aligned}$$

$$\Rightarrow \quad E\left[S_n^2 I_{B_k}\right] \ge 0 + E\left[S_k^2 I_{B_k}\right]$$

Since,

$$\begin{aligned} S_n - S_k &= (X_1 + X_2 + \ldots + X_n) - (X_1 + X_2 + \ldots + X_k) \\ &= X_{k+1} + X_{k+2} + \ldots + X_n \end{aligned}$$

and $S_k = X_1 + X_2 + \ldots + X_k$ are independent

$$\begin{aligned} \therefore \quad E\left[(S_n - S_k)S_k I_{B_k}\right] &= E[S_n - S_k]\, E(S_k I_{B_k}) \\ &= E[(X_{k+1} + X_{k+2} + \ldots + X_n)\, S_k I_{B_k}] = 0 \end{aligned}$$

Since, $E(X_i) = 0 \; \forall \; i$.

Hence,

$$E\left[S_n^2 I_{B_k}\right] = E\left[S_k^2 I_{B_k}\right]$$

Using this result in (12.9), we get

$$\sum_{k=1}^{n} \sigma_k^2 \geq \sum_{k=1}^{n} E\left[S_k^2 I_{B_k}\right] \geq \epsilon^2 \sum E\left(I_{B_k}\right)$$

$$= \epsilon^2 \sum_{k=1}^{n} P(B_k) = \epsilon^2 P\left(\sum_{k=1}^{n} B_k\right)$$

So, $$P\left[\max_{1 \leq k < n} |S_k| \geq \epsilon\right] \leq \frac{1}{\epsilon^2} \sum_{k=1}^{n} \sigma_k^2$$

Note : If we take $k = 1$ in $S_k = \sum_{i=1}^{k} X_i$ then the SLLN reduces to,

$P[|X_1| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}$ which is the Chebyshev's inequality.

So, for $k = 1$, we get the Chebyshev's inequality.

Illustration 12.1 : Let $\{X_n\}$ be any sequence of random variables. We write $\overline{X}_n = \frac{1}{n} \sum_{i=1}^{k} X_i$. A necessary and sufficient condition for the sequence $\{X_n\}$ to satisfy WLLN is that

$$E\left[\frac{\overline{X}_n^2}{1+\overline{X}_n^2}\right] \to 0 \text{ as } n \to \infty$$

Proof : We first prove that,

$$E\left[\frac{Y_n^2}{1+Y_n^2}\right] \to 0 \text{ as } n \to \infty$$

$$\Rightarrow \qquad Y_n \xrightarrow{p} 0$$

For any two positive numbers, a, b, $a \geq b > 0$, we know that

$$\left(\frac{a}{1+a}\right)\left(\frac{b+1}{b}\right) \geq 1 \qquad \text{....(1)}$$

Let, $A = \{|Y_n| \geq \epsilon\}$. Then $w \in A \Rightarrow |Y_n|^2 \geq \epsilon^2$ and using (1), we see that for all $w \in A$.

$$\left(\frac{Y_n^2}{1+Y_n^2}\right)\left(\frac{1+\epsilon^2}{\epsilon^2}\right) \geq 1$$

Thus, $$|Y_n| \geq \epsilon \quad \Rightarrow \left(\frac{Y_n^2}{1+Y_n^2}\right)\left(\frac{1+\epsilon^2}{\epsilon^2}\right) \geq 1$$

$$\Rightarrow \qquad P[|Y_n| \geq \in] \leq P\left[\left(\frac{Y_n^2}{1+Y_n^2}\right)\left(\frac{1+\in^2}{\in^2}\right) \geq 1\right]$$

$$= P\left[\frac{Y_n^2}{1+Y_n^2} \geq \frac{1+\in^2}{\in^2}\right] \leq \frac{E\left[\frac{Y_n^2}{1+Y_n^2}\right]}{\left(\in^2/(1+\in^2)\right)}$$

Using Markov's inequality.

$$\Rightarrow \qquad \lim_{n\to\infty} P[|Y_n| \geq \in] \leq 0$$

$$\Rightarrow \qquad Y_n \xrightarrow{p} 0$$

We now prove that

$$Y_n \xrightarrow{p} 0 \qquad \Rightarrow E\left[\frac{Y_n^2}{1+Y_n^2}\right] \to 0 \text{ as } n \to \infty$$

For this we consider

$$E\left[\frac{Y_n^2}{1+Y_n^2}\right] = \int_{-\infty}^{\infty} \frac{y^2}{(1+y^2)} f_n(y)\, dy$$

$$= \int_{-\infty}^{-\in} \frac{y^2}{1+y^2} f_n(y)\, dy + \int_{-\in}^{\in} \frac{y^2}{1+y^2} f_n(y)\, dy + \int_{-\in}^{\infty} \frac{y^2}{1+y^2} f_n(y)\, dy$$

$$= \int_{|y_n| \geq \in} \left(\frac{y^2}{1+y^2}\right) f_n(y)\, dy + \int_{|y_n| < \in} \left(\frac{y^2}{1+y^2}\right) f_n(y)\, dy$$

$$= \text{I} + \text{II}$$

In the first integral, we have

$$|Y_n| \geq \in > 0, \quad \left(\frac{a}{a+1}\right)\left(\frac{1+b}{b}\right) > 1$$

$$\Rightarrow \qquad Y_n^2 \geq \in^2 > 0 \quad \Rightarrow \quad \frac{Y_n^2}{1+Y_n^2} \geq \frac{\in^2}{1+\in^2}$$

$$\Rightarrow \qquad \frac{\in^2}{1+\in^2} \leq \frac{Y_n^2}{1+Y_n^2} \leq 1$$

In the second integral we have

$$\frac{Y_n^2}{1+Y_n^2} \leq Y_n^2 \text{ where } |Y_n| < \in$$

$$\Rightarrow \qquad E\left[\frac{Y_n^2}{1+Y_n^2}\right] \leq \int_{|y_n| \geq \in} f_n(y)dy + \int_{|y_n| < \in} y^2 f_n(y)dy$$

$$\leq P[|Y_n| \geq \epsilon] + \epsilon^2 P[|Y_n| < \epsilon]$$

$$\Rightarrow \quad E\left[\frac{Y_n^2}{1+Y_n^2}\right] \leq P[|Y_n| \geq \epsilon] + \epsilon^2$$

$$\Rightarrow \quad \lim_{n\to\infty} E\left[\frac{Y_n^2}{1+Y_n^2}\right] \leq \epsilon^2$$

$$\Rightarrow \quad \lim_{n\to\infty} E\left[\frac{Y_n^2}{1+Y_n^2}\right] \leq 0$$

Illustration 12.2 : Let $X_1, X_2, \ldots$ be a sequence of *iid* random variables in U[0, 1]. For the $G_n = (X_1, X_2, \ldots, X_n)^{1/n}$, show that, $G_n \to C$ for some finite number C. Find C.

Solution : Let, $G_n = (X_1, X_2, \ldots, X_n)^{1/n}$

So, $\log G_n = \frac{1}{n} \Sigma \log X_i$

$$= -\frac{1}{n}\left[\sum_{i=1}^{n} -\log X_i\right]$$

To find the *pdf* of $Y = -\log X$ we proceed as follows

$$g(y) = \left|\frac{dx}{dy}\right| f(x)$$

where the right hand side is expressed in terms of Y

So, $\quad g(y) = e^{-y}$

and $\quad E(Y) = 1 \;\forall\; y,$

So, $\overline{Y}_n \xrightarrow{p} 1$ by WLLN

$$\Rightarrow \quad -\frac{1}{n}\sum_{i=1}^{n} \log X_i \xrightarrow{p} 1 \quad \Rightarrow -\log G_n \xrightarrow{p} 1$$

or $G_n \xrightarrow{p} e^{-1}$. So, $C = e^{-1}$

Illustration 12.3. Examine if the law of large numbers holds for the sequence of independent random variables $\{X_n\}$ with the distribution of X_n given by

$$f_n(x) = \frac{1}{|x|^3}, \; |x| > 1$$
$$= 0, \text{ otherwise}$$

Solution :

$$f_n(x) = \frac{1}{|x|^3}, \; |x| > 1$$
$$= 0, \text{ otherwise}$$

i.e.,
$$f_n(x) = \frac{1}{|x|^3}, x > 1 \text{ or } x < -1$$
$$= 0, \text{ otherwise}$$

$$E(X) = \int_x x f(x) dx$$

$$= \int_1^{\infty} x \frac{1}{|x|^3} dx + \int_{-\infty}^{-1} x \frac{1}{|x|^3} dx$$

$$= \int_1^{\infty} x \frac{1}{x^3} dx + \int_{-\infty}^{-1} x \frac{1}{(-x)^3} dx$$

$$= \int_1^{\infty} \frac{1}{x^2} dx - \int_{-\infty}^{-1} \frac{1}{x^2} dx$$

$$= \left(-\frac{1}{x}\right)_1^{\infty} - \left(-\frac{1}{x}\right)_{-\infty}^{-1}$$

$$= 0$$

$$E(X^2) = \int_x x^2 f(x) dx$$

$$= \int_1^{\infty} x^2 \frac{1}{|x|^3} dx - \int_{-\infty}^{-1} x^2 \frac{1}{|x|^3} dx$$

$$= \int_1^{\infty} x^2 \frac{1}{x^3} dx + \int_{-\infty}^{-1} x^2 \frac{1}{(-x)^3} dx$$

$$= \int_1^{\infty} \frac{1}{x} dx - \int_{-\infty}^{-1} \frac{1}{x} dx$$

$$= (\log x)_1^{\infty} - (\log x)_{-\infty}^{-1}$$
$$= (\log \infty - \log 1) - (\log(-1) - \log(-\infty))$$
$$= (\infty - 0) - (\log(-1) - \log(-\infty))$$

$\therefore$ $E(X^2)$ does not exist.

$\therefore$ $V(X)$ does not exist.

$\therefore$ law of large numbers does not hold for the sequence of independent random variables $\{X_n\}$.

Illustration 12.4. Show that the sequence $P[X_k = \pm 2^k] = \frac{1}{2}$ does not obey the weak law of large numbers.

Solution :

$$P[X_n = \pm 2^k] = \frac{1}{2}$$

$$\begin{aligned}
E(X_k) &= \Sigma X_k P(x_k) \\
&= 2^k P(X_k = 2^k) + (-2^k) P(x_k = -2^k) \\
&= 2^k \times \frac{1}{2} - 2^k \times \frac{1}{2} \\
&= 0
\end{aligned}$$

$$\begin{aligned}
E\left(X_k^2\right) &= \Sigma X_k^2 P(X_k) \\
&= (2^k)^2 P(X_k = 2^k) + (-2^k)^2 P(X_k = -2^k) \\
&= 2^{2k}\frac{1}{2} + 2^{2k}\frac{1}{2} \\
&= 2^{2k}
\end{aligned}$$

$$V(X_k) = E\left(X_k^2\right) - \{E(X_k)\}^2 = 2^{2k} - 0^2 = 2^{2k}$$

Now,

$$\begin{aligned}
B_n &= V\left(\sum_{k=1}^{n} X_k\right) \\
&= \sum_{k=1}^{n} V(X_k) \\
&= \left(\sum_{k=1}^{n} 2^{2k}\right) \\
&= 2^{2.1} + 2^{2.2} + 2^{2.3} + \dots + 2^{2n} \\
&= A + A^2 + A^3 + \dots + A^n \text{ if } A = 2^2 \\
&= \frac{A(A^n - 1)}{A - 1} \\
&= \frac{2^2\left[\left(2^2\right)^n - 1\right]}{2^2 - 1} \\
&= \frac{4}{3}[2^{2n} - 1]
\end{aligned}$$

$$\frac{B_n}{n^2} = \frac{4}{3n^2}[2^{2n} - 1]$$

$$\lim_{n\to\infty} \frac{B_n}{n^2} = \lim_{n\to\infty} \frac{4}{3n^2}[2^{2n} - 1] = \frac{4}{3}\lim_{n\to\infty}\left[\frac{2^{2n}}{n^2} - \frac{1}{n^2}\right]$$

$$= \frac{4}{3} \lim_{n \to \infty} \frac{2^{2n}}{n^2} \qquad \left[\because \lim_{n\to\infty} \frac{1}{n^2} = 0\right]$$

Let
$$U_n = \frac{2^{2n}}{n^2}$$

$$U_{n+1} = \frac{2^{2n+2}}{(n+1)^2}$$

$$\frac{U_{n+1}}{U_n} = \frac{2^{2n+2}}{2^{2n}} \frac{n^2}{(n+1)^2} = \frac{2^2}{\left(1+\frac{1}{n}\right)^2} = \frac{4}{\left(1+\frac{1}{n}\right)^2}$$

$$\therefore \quad \lim_{n\to\infty} \left|\frac{U_{n+1}}{U_n}\right| = 4 > 1$$

$\therefore \quad U_n \to \infty$ as $n \to \infty$.

$\therefore \quad \frac{B_n}{n^2} \to \infty$ as $n \to \infty$.

$\therefore$ The sequence $P[X_k = \pm 2^k] = \frac{1}{2}$ does not obey the weak law of large numbers.

Illustration 12.5. Let $X_1, X_2, \ldots$ be *iid* variates with *pmf*.

$f(x) = \alpha^x/(1-\alpha)$, $x = 0, 1, 2, \ldots, 0 < \alpha < 1$

Show that WLLN is followed by X_i's.

Solution : $f(x) = \frac{\alpha^x}{1-\alpha}$, $x = 0, 1, 2, \ldots, 0 < \alpha < 1$

$$E(X) = \sum_{x=0}^{\infty} x f(x)$$

$$\therefore \quad E(X) = \sum_{x=0}^{\infty} x \frac{\alpha^x}{1-\alpha}$$

$$= \frac{1}{1-\alpha} \sum_{x=0}^{\infty} x\alpha^x$$

$$= \frac{1}{1-\alpha} [0.\alpha^{\circ} + 1.\alpha^1 + 2.\alpha^2 + 3\alpha^3 + \ldots..]$$

$$= \frac{1}{1-\alpha} [\alpha + 2.\alpha^2 + 3\alpha^3 + \ldots..]$$

$$= \frac{\alpha}{1-\alpha} [1 + 2\alpha + 3\alpha^2 + \ldots..]$$

$$= \frac{\alpha}{1-\alpha}(1-\alpha)^{-2}$$

$$= \frac{\alpha}{(1-\alpha)^3}$$

$$E(X^2) = \sum_{x=0}^{\infty} x^2 \frac{\alpha^x}{1-\alpha}$$

$$= \frac{1}{1-\alpha}\left[\sum_{x=0}^{\infty} x^2 \alpha^x\right]$$

$$= \frac{1}{1-\alpha}[0^2.\alpha^0 + 1^2.\alpha^1 + 2^2.\alpha^2 + 3^2\alpha^3 + ...]$$

$$= \frac{1}{1-\alpha}[\alpha + 4\alpha^2 + 9\alpha^3 + ...]$$

$$= \frac{1}{1-\alpha}\alpha(1 + 4\alpha + 9\alpha^2 + ...)$$

$$= \frac{1}{1-\alpha}\alpha(1+\alpha)(1-\alpha)^{-3}$$

$$= \frac{\alpha(1+\alpha)}{(1-\alpha)^4}$$

$$V(X) = E(X^2) - \{E(X)\}^2$$

$$= \frac{\alpha(1+\alpha)}{(1-\alpha)^4} - \left\{\frac{\alpha}{(1-\alpha)^3}\right\}^2$$

$$= \frac{\alpha(1+\alpha)}{(1-\alpha)^4} - \frac{\alpha^2}{(1-\alpha)^6}$$

$$= \frac{\alpha}{(1-\alpha)^6}[\alpha^3 - \alpha^2 - 2\alpha + 1]$$

$$B_n = V\left(\sum_{i=1}^{n} X_i\right)$$

$$= \sum_{i=1}^{n} V(X_i)$$

$$= \sum_{i=1}^{n} \frac{\alpha}{(1-\alpha)^6}[\alpha^3 - \alpha^2 - 2\alpha + 1]$$

$$= \frac{n\alpha}{(1-\alpha)^6} [\alpha^3 - \alpha^2 - 2\alpha + 1]$$

$$\therefore \quad \frac{B_n}{n^2} = \frac{n\alpha}{(1-\alpha)^6} \frac{[\alpha^3 - \alpha^2 - 2\alpha + 1]}{n^2}$$

$$= \frac{\alpha}{n(1-\alpha)^6} [\alpha^3 - \alpha^2 - 2\alpha + 1]$$

$$\therefore \quad \frac{B_n}{n^2} \to 0 \text{ as } n \to \infty$$

$\therefore$ WLLN is followed by X_i, $(i = 1, 2, ..., n)$.

Illustration 12.6. A pack of cards numbered 1, 2, ..., n is shuffled and the cards are dealt one at a time. The variate X_i posses values 1 or 0 according as the ith card dealt has number i on it or not and each card is equiprobable to appear at the ith place. Show that WLLN holds for the sequence $\{X_n\}$.

Solution. $X_i = 1$ if the ith card dealt has the number i on it
$= 0$ otherwise

$$E(X_i) = 1\, P(X_i = 1) + 0\, P(X_i = 0) = P(X_i = 1) = \frac{1}{n}$$

$$E(X_i^2) = 1^2\, P(X_i = 1) + 0^2\, P(X_i = 0) = P(X_i = 1) = \frac{1}{n}$$

$$\therefore \quad V(X_i) = E(X_i^2) - \{E(X_i^2)\}^2 = \frac{1}{n} - \left(\frac{1}{n}\right)^2 = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2}$$

$$B_n = V\left(\sum_{i=1}^{n} X_i\right)$$

$$= \sum_{i=1}^{n} V(X_i)$$

$$= n \times \frac{n-1}{n^2}$$

$$= \frac{n-1}{n}$$

$$\therefore \quad \frac{B_n}{n^2} = \frac{n-1}{n.n^2}$$

$$= \frac{n-1}{n^3}$$

$$= \frac{1}{n^2} - \frac{1}{n^3}$$

$\therefore$ As $n \to \infty$, $\frac{B_n}{n^2} \to 0$

$\therefore$ WLLN holds for the sequence $\{X_n\}$

EXERCISE

1. State the conditions for the existence of the weak law of large numbers. Examine whether the weak law of large number holds for the mean of a sequence of independent variates X_k with,

   $P\{X_k = \pm \sqrt{\log k}\} = \frac{1}{2}$ [IAS, 2003]

2. Explain the concept of strong and week law of large numbers. [IAS, 2000]

3. Let $X_1, X_2, ..., X_n$ is a sequence of independent random variables with $P[X_k = \pm K^a] = \frac{1}{2}$ where $a > 0$. Prove that the law of large numbers holds for this sequence for $a < \frac{1}{2}$. [IAS, 1999]

4. Show that a sequence of random variables $\{X_K\}$ satisfies the week law of large numbers, if

   $$\frac{1}{n^2} \text{Var}\left[\sum_{k=1}^{n} X_k\right] \to 0 \text{ as } n \to \infty$$ [IAS, 1981]

5. Verify whether the strong law of large numbers holds for a sequence of random variables $\{X_n\}$ defined by, $P[X_n = -2^{-n}] = 1/2$ [IAS, 1987]

6. State and prove the weak law of large numbers for non-identically distributed random variables. [IAS, 1987]

7. Examine whether the law of large number holds for the following sequence of random variables :

   $P[X_n = \pm 2^n] = 2^{-(2n+1)} P[X_n = 0] = 1 - 2^{-2n}$ [IAS, 1987]

8. Let $\{X_n, n = 1, 2, ...\}$ be a sequence of random variables each with mean 0 and unit variance. Suppose that the correlation coefficient between X_m and X_n where $m \neq n$, equals to $\{-|m - n| \alpha\} > 0$. Show that the weak law of large numbers holds for the sequence $\{X_n, n = 1, 2, ...\}$

   [IAS, 1988]

9. Examine whether the law of large number holds for sequence of independent random variable $\{X_n, n = 1, 2, ...\}$ with the distribution of X_n given by,

   $$f_n(x) = \begin{cases} \frac{1}{|x|^3}, & |x| > 1 \\ 0, & \text{Otherwise} \end{cases}$$ [ISS, 1998]

□□

CHAPTER 13

The Central Limit Theorem

13.1 INTRODUCTION

The Central limit theorem is one of the most remarkable results in probability theory. Very roughly speaking, it states that the sum of a large number of independent random variables has a distribution that is approximately normal. Hence, it provides a simple way of computing approximate probabilities for sum of independent random variables. The theorem also provides means to inform that the empirical frequencies of so many natural populations exhibit the normal form.

From the dice experiment in the probability chapter, it is seen that the stability of the proportions of the outcomes is observed, when the experiment is repeated under the constant conditions. The same can also be seen in the coin tossing experiment. A coin is tossed a large number of times unde the constant conditions, the probability of occurrence of head approaches to 0.5 for an unbased coin. The possible outcome of the coin tossing experiment is two and the sample space is $w = \{H, T\}$, where H and T represent Head and Tail. We denote the probability of occurrence of Head as $P(H) = 0.5$ and the probability of occurrence of Tail as $P(T) = 0.5$. The repeated trials are independent. These repeated independent trials are called *Bernoulli trials*, which is defined as follows.

In general, for a Bernoulli trial there will be only two outcomes. Let A be one of the occurrence of the outcome and $\bar{A}$, the non-occurrence of the outcome. The sample space is denoted by $w = \{A, \bar{A}\}$. Let $P(A) = p$ and $P(\bar{A}) = q$, such that $p + q = 1$.

Here, we are interested to know the effect of considering large number of trials, which leads to *Central Limit Theorem*. We define the *Central Limit Theorem* in the following theorem.

Theorem. Consider a sample space $w = \{A, \bar{A}\}$ with $P(A) = p$ and $P(\bar{A}) = q$, such that $p + q = 1$. Given n independent ramdon variables $X_1, X_2, ..., X_n$, each assumes the value

$$X_i = \begin{cases} 1, & \text{for } A \\ 0, & \text{for } \bar{A} \end{cases} \qquad i = 1, 2, ..., n$$

Let $$S_n = \sum_{i=1}^{n} X_i$$

and $$S_n^* = \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}}$$

Then for any real number a and b,

$$\lim_{n\to\infty} P(a \le S_n^* \le b) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx$$

For any sequence of events for any n, S_n^* must be some number. Thus for all n,

$$P(-\infty \le S_n^* \le \infty) = 1$$

Thus for large n, the integral $P(-\infty \le S_n^* \le \infty)$, *i.e.*, $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right)dx$ must be one, which is the main interest for the *central limit theorem*. If the integral is not equal to one, the *central limit theorem* is not true.

Let $X_1, X_2, ..., X_n$ be a sequence of jointly distributed random variables with finite means and variances, $0 < var(X_i) < \infty$. Then the sequence of is said to follow the Central Limit Theorem (CLT) iff

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} \xrightarrow{L} N(0, 1)$$

where, $S_n = \sum_{i=1}^{n} X_i$

The meaning of central limit theorem is that repeated samples of size 'n' even from a non-normal population generate an approximately, normal distribution of either sums or means. For, example if $X_i = (i = 1, 2, ..., n)$ be a sample from uniform (0, 1) distribution (say), with n being sufficiently large then $\sum X_i$ will generate an approximately normal distribution.

13.2 LINDEBERG - LEVY CENTRAL LIMIT THEOREM

Let $X_1, X_2, ..., X_n$ be *iid* variates, with mean μ and variance σ^2, $0 < \sigma^2 < \infty$. Let

$$F_n(x) = P\left[\left(\frac{\bar{x} - \mu}{\sigma/\sqrt{n}}\right) \le x\right]$$

and

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

Then, $F_n(x) \to \Phi(x)$ as $n \to \infty$ for all $x \in R$.

Proof : Let,

$$Z_n = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} = \frac{n\bar{X} - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}} \frac{n\frac{\sum X_i}{n} - n\mu}{\sigma}$$

$$= \frac{1}{\sqrt{n}} \frac{\sum (X_i - \mu)}{\sigma} = \frac{1}{\sqrt{n}} \sum\left(\frac{X_i - \mu}{\sigma}\right)$$

$$= \frac{1}{\sqrt{n}} \sum Y_i \quad (i = 1, 2, ..., n)$$

where, $Y_i = \frac{X_i - \mu}{\sigma}$ are *iid* variates with $E(Y_i) = 0$, $var(Y_i) = 1$.

The characteristic function of Z_n is ϕ and is given by,

$$\phi_{\sum Y_i/\sqrt{n}}(t) = \phi_{\sum Y_i}(t/\sqrt{n}) = \left[\phi_{Y_i}(t/\sqrt{n})\right]^n \quad ...(13.1)$$

Now, according to Maclaurin's expansion of $\phi_Y(t/\sqrt{n})$ up to second order terms, we have,

$$\phi_Y\left(\frac{t}{\sqrt{n}}\right) = \phi_Y(0) + \frac{t}{\sqrt{n}} \phi'_Y(0) + \frac{t^2}{2n} \phi''_Y(0) + 0\left(\frac{t^2}{n}\right)$$

From the properties of characteristic function we have,

$\phi_Y(0) = 1$, $\phi'_Y(0) = i \Rightarrow E(Y) = 0$

and $\phi''_Y(0) = i^2 \Rightarrow E(Y^2) = -1$

So, $$\phi_Y\left(\frac{t}{\sqrt{n}}\right) = 1 - \frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right) \quad ...(13.2)$$

Substituting (2) in (1) considering $Y_i \doteq Y$ gives,

$$\phi_{\frac{\sum Y_i}{\sqrt{n}}}(t) = \left[1 - \frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]^n$$

Now taking limit $n \to \infty$, we have,

$$\underset{n\to\infty}{Lt}\, \phi_{\frac{\sum Y_i}{\sqrt{n}}}(t) = \underset{n\to\infty}{Lt}\left[1 - \frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]^n$$

$$= e^{-t^2/2}$$

which is the characteristic function of N (0, 1).

Thus from the uniqueness property of charactistic function we can have $\frac{\sum_{i=1}^{n} Y_i}{\sqrt{n}}$ follows standard normal distribution as $n \to \infty$.

i.e., $Z_n \sim N(0, 1) \Rightarrow \frac{n(\bar{X} - \mu)}{\sigma} \sim N(0, 1)$

i.e., $\frac{(\bar{X} - \mu)}{\sigma\sqrt{n}} \sim N(0, 1) \quad \Rightarrow \quad \frac{S_n - E(S_n)}{Var(S_n)} \sim N(0,1)$

Thus, $\{X_n\}$ follows the central limit theorem.

13.3 THE DE-MOIVRE'S LAPLACE CENTRAL LIMIT THEOREM

Let $X_1, X_2, ..., X_n$ be *iid* Bernoulli variates with common probability p of success. If $X = X_1 + X_2 + ... + X_n$, then

$$\underset{n \to \infty}{\text{Lim}} \; P\left[\frac{X_n - np}{\sqrt{npq}} \le x\right] = \Phi(x)$$

In other words, if $Y_n = \frac{X_n - np}{\sqrt{npq}}$ then

$F_n(y) \xrightarrow{L} F(y)$, where $F(y)$ refers to the c.d.f. of a standard normal variable.

Proof : The characteristic function of X_n is given by

$$\phi_{X_n}(t) = E\left[e^{itX_n}\right]$$
$$= (q + pe^{it})^n \qquad ...(13.3)$$

Then the characteristic function of $Y_n = \frac{X_n - np}{\sqrt{npq}}$

Now

$$\phi_{Y_n}(t) = E\left[e^{itY_n}\right] = E\left[e^{it(X_n - np)/\sqrt{npq}}\right]$$

$$= e^{-\frac{inpt}{\sqrt{npq}}} E\left[e^{itX_n/\sqrt{npq}}\right]$$

$$= e^{-\frac{inpt}{\sqrt{npq}}}\left(q + pe^{\frac{it}{\sqrt{npq}}}\right)^n \qquad \text{[using (13.3)]}$$

$$= \left(qe^{-\frac{pit}{\sqrt{npq}}} + pe^{\frac{qit}{\sqrt{npq}}}\right)^n \qquad ....(13.4)$$

We have,

$$e^{iz} = \sum_{k=0}^{\infty} \frac{(iz)^k}{k!} + 0(z^k) \qquad ...(13.5)$$

Using (13.5) we can write

$$qe^{-\frac{pit}{\sqrt{npq}}} = q - it\sqrt{\frac{pq}{n}} - \frac{t^2 p}{2n} + 0\left(\frac{t^2}{n}\right) \qquad ...(13.6)$$

and

$$pe^{\frac{qit}{\sqrt{npq}}} = p + it\sqrt{\frac{pq}{n}} - \frac{qt^2}{2n} + 0\left(\frac{t^2}{n}\right) \qquad ...(13.7)$$

Replacing (13.6) and (13.7) in (13.4) and using $p + q = 1$,
We have

$$\phi_{Y_n}(t) = \left[1 - \frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]^n$$

$$\log \phi_{Y_n}(t) = n \log\left[1 - \frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]$$

$$= n\log(1 + Z), \text{ where } Z = -\frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)$$

where

$$|Z| = \left|\frac{-t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right| < 1 \;\forall\; t \text{ in } R$$

So, for sufficiently large n, one can write

$$\log \phi_{Y_n}(t) = n\,Z$$

$$= n\left(-\frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right)$$

$$= -\frac{t^2}{2} + 0\left(\frac{t^2}{n}\right)$$

Thus,

$$\underset{n \to \infty}{\text{Lt}} \log \phi_{Y_n}(t) = -\frac{t^2}{n}$$

$$\Rightarrow \underset{n \to \infty}{\text{Lt}} \phi_{Y_n}(t) = e^{-t^2/2}$$

which is the characteristic function of a standard normal variate.

Thus, from the uniqueness property of characteristic function we can say that the standard binomial variate converges to the standard normal variate for large N.

13.4 LIAPOUNOV'S CENTRAL LIMIT THEOREM

Let $X_1, X_2, ..., X_n$ be n independent random variables with μ_r and σ_r as the mean and standard deviation of X_r $(r = 1, 2, ..., n)$.

Let us suppose that the third absolute moment of X_r about its mean, *i.e.*,

$$\gamma_r^3 = E[|X_r - \mu_r|^3] \text{ exist for all } r.$$

If,

$$\gamma^3 = \gamma_1^3 + \gamma_2^3 + ... + \gamma_n^3$$

$$\Lambda = \mu_1 + \mu_2 + ... + \mu_n,$$

$$\Sigma^2 = \sigma_1^2 + \sigma_2^2 + ... + \sigma_n^2$$

and,

$$\underset{n \to \infty}{\text{Lt}} \frac{\gamma}{\Sigma} = 0 \text{ then,}$$

$X = \sum_{r=1}^{n} X_r$ is asymptotically normally distributed with mean Λ and standard deviation Σ.

Proof : Let $\phi_r(t)$ be the characteristic function of $(X_r - \mu_r)$ and

$\phi_{\frac{X-\mu}{\Sigma}}(t) = \phi_Z(t)$ be the characteristic function of $\frac{X - \Lambda}{\Sigma}$.

Now,

$$\frac{X-\Lambda}{\Sigma} = \frac{X_1 + X_2 + ... + X_n - \mu_1 - \mu_2 - ..... - \mu_n}{\Sigma}$$

$$= \frac{\sum_{i=1}^{n}(X_i - \mu_i)}{\Sigma} = \sum_{i=1}^{n}\left(\frac{X_i - \mu_i}{\Sigma}\right)$$

Now, we have

$$\phi_{\frac{X-\mu}{\Sigma}}(t) = \phi_Z(t) = E\left[\exp\left\{it\sum_{r=1}^{n}\left(\frac{X_r - \mu_r}{\Sigma}\right)\right\}\right]$$

$$= \prod_{r=1}^{n} E\left[\exp\left\{it\left(\frac{X_r - \mu_r}{\Sigma}\right)\right\}\right]$$

$$\Rightarrow \quad \phi_Z(t) = \prod_{r=1}^{n} \phi_r\left(\frac{t}{\Sigma}\right), \text{ where } \phi_r\left(\frac{t}{\Sigma}\right) = E\left[\exp\left\{it\left(\frac{X_r - \mu_r}{\Sigma}\right)\right\}\right] \quad ...(13.8)$$

Now,

$$\phi_r(t) = E[\exp(it(X_r - \mu_r))]$$

$$= 1 + it\,E(X_r - \mu_r) - \frac{t^2}{2}E(X_r - \mu_r)^2 + \frac{\theta t^3}{6}E[|X_r - \mu_r|^3]$$

Using Euler-Maclaurin's expansion where $|\theta| < 1$

$$\Rightarrow \quad \phi_r(t) = 1 - \frac{t^2}{2}\sigma_r^2 + \theta\frac{t^3}{6}\gamma_r^3 + ... \quad ...(13.8)$$

From the assumption, we have

$$\lim_{n\to\infty}\frac{\Lambda}{\Sigma} = 0 \quad \Rightarrow \quad \frac{\gamma_p}{\Sigma} \le \frac{\gamma}{\Sigma} < 1 \quad ...(13.9)$$

While observing $\sigma_r \le \gamma_r\ \forall r$. If γ_k represents the Kth absolute moment from origin.

Now, putting K = 2, 3, we have

$$(\gamma_2')^{1/2} \le (\gamma_3')^{1/3}$$

Also, $(\gamma_2)^{1/2}$ and $(\gamma_3)^{1/3}$ for the rth population $(r = 1, 2, ..., n)$ are equal to σ_r and γ_r respectively.

Again,

$$(\gamma_r')^{3/2} = (\gamma_r)^{3/2} = E[|X_r - \mu_r|^3]^{1/2} \ge E[|X_r - \mu_r|^2]^{1/2} = \sigma_r \quad ...(13.10)$$

Thus from (13.9) and (13.10), we have

$$\sigma_r \le \gamma_r,\ r = 1, 2, 3, ... \quad ....(13.11)$$

Also,

$$\log\phi_r\left(\frac{t}{\Sigma}\right) = \log\left(1 - \frac{\sigma_r^2 t^2}{2\Sigma^2} + \frac{\theta\, t^3\gamma_r^3}{3!\,\Sigma^3}\right)$$

$$= \log(1 + \lambda) \quad ...(13.11)$$

where,

$$\lambda = -\frac{\sigma_r^2 t^2}{2\Sigma^2} + \frac{\theta\, t^3 \gamma_r^3}{3!\,\Sigma^3} \quad ....(13.12)$$

$$\ge -\frac{\theta\gamma_r^2 t^2}{2\Sigma^2} + \frac{\theta\gamma_r^3 t^3}{6\Sigma^3} \text{ (using } \sigma_r \le \gamma_r)$$

Thus, $\lambda \to 0$ as $n \to \infty$.

So, $|\lambda| < \frac{1}{2}$ holds for all sufficiently large n.

Thus, for $|\lambda| < \frac{1}{2}$, we can have

$$\log(1+\lambda) = \lambda - \frac{\lambda^2}{2} + \frac{\lambda^3}{2} - \frac{\lambda^4}{4} + ...$$

$$= \lambda - \frac{\lambda^2}{2}\left(1-\frac{2}{3}\lambda+\frac{2}{4}\lambda^2-...\right)$$

$$= \lambda - \frac{1}{2}\theta\left(1+\frac{1}{2}+\frac{1}{2^2}+...\right)\lambda^2$$

$$= \lambda + \frac{1}{2}\theta\lambda^2\left(1-\frac{1}{2}\right)^{-1}$$

$$= \lambda + \theta\lambda^2.$$

Again from (13.11) and (13.12), we have

$$\log \phi_r\left(\frac{t}{\Sigma}\right) = \log\left(1-\frac{\sigma_r^2 t^2}{2\Sigma^2}+\frac{\theta\gamma_r t^3}{6\Sigma^3}\right)$$

$$\Rightarrow \quad \log(1+\lambda) = -\frac{\sigma_r^2 t^2}{2\Sigma^2}+\frac{\theta\gamma_r^3 t^2}{6\Sigma^3}+\frac{\theta\gamma_r^4}{\Sigma^4}\left(\frac{t^2}{2}+\frac{|t|^3}{6}\right)^2$$

$$\Rightarrow \quad \log\phi_r\left(\frac{t}{\Sigma}\right) = -\frac{\sigma_r^2}{\Sigma}.\frac{t^2}{2}+\frac{\theta\gamma_r^3}{\Sigma}\left(\frac{1}{6}|t|^3+\left\{\frac{1}{2}t^2+\frac{1}{6}|t|^3\right\}^2\right)$$

Summing over all $r = 1, 2, 3, ..., n$ and employing

$$\phi_z(t) = \prod_{r=1}^{n}\phi_r\left(\frac{t}{\Sigma}\right) \text{ we have,}$$

$$\log\phi_z(t) = \sum_{r=1}^{n}\log\phi_r\left(\frac{t}{\Sigma}\right)$$

$$= -\frac{\sum_{r=1}^{n}\sigma_r^2}{\Sigma^2}\times\frac{t^2}{2}+\theta\frac{\sum_{r=1}^{n}\gamma_r^2}{\Sigma^3}\left(\frac{1}{6}|t|^3+\left\{\frac{1}{2}t^2+\frac{1}{6}|t|^3\right\}^2\right)$$

$$= -\frac{t^2}{2}+\frac{\theta\gamma^3}{\Sigma^3}\left(\frac{1}{6}|t|^3+\left\{\frac{1}{2}t^2+\frac{1}{6}|t|^3\right\}^2\right)$$

Since, $\frac{\gamma}{\Sigma} \to 0$ as $n \to \infty$.

$$\Rightarrow \quad \log\phi_z(t) \to -\frac{t^2}{2} \;\forall\, t$$

$$\Rightarrow \quad \phi_z(t) \to \exp\left(-\frac{t^2}{2}\right)$$

which proves the Liapounov's theorem.

13.5 THE LINDEBERG-FELLER CENTRAL LIMIT THEOREM

Statement : Let $X_1, X_2, ..., X_n$ be independent non-degenerate variables with distribution functions $F_1, F_2, ..., F_n$ and means $\mu_1, \mu_2, ..., \mu_n$ and finite variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$ respectively. Let

$$B_n^2 = \sigma_1^2+\sigma_2^2+...+\sigma_n^2, \quad S_n = X_1+X_2+...+X_n$$

If F_i are absolutely continuous with *pdf* f_i and if the condition,

$$\underset{n\to\infty}{\text{Lt}}\frac{1}{B_n^2}\sum_{i=1}^{n}\int_{A_n}(x-\mu_i)^2 f_i(x)\,dx = 0$$

holds for all $\in > 0$, then

$$\underset{n\to\infty}{\text{Lt}}\, P\left[\frac{S_n - E(S_n)}{B_n} \le a\right] = \phi(a) \text{ is the C.D.F. of } N(0,1)$$

Here $\{A_n = |X-\mu_i| > \in B_n\}$

(Proof is omitted)

Illustration 13.1. Examine if Central Limit Theorem holds for the following

$$P(X_k = \pm 2^k) = 2^{-(2k+1)}, \quad P(X_k = 0) = 1 - 2^{-2k}$$

Solution : $P(X_k = \pm 2^k) = 2^{-(2k+1)}, \quad P(X_k = 0) = 1 - 2^{-2k}$

$$\mu_k = E(X_k) = \sum X_k P(X_k)$$
$$= 2^k . P(X_k = 2^k) + (-2^k) P(X_k = -2^k) + 0\, P(X_k = 0)$$
$$= 2^k . 2^{-(2k+1)} - 2^k . 2^{-(2k+1)} + 0.(1-2^{-2k})$$
$$= 0$$

$$E(X_k^2) = \sum X_k^2 P(X_k)$$
$$= (2^k)^2 P(X_k = 2^k) + (-2^k)^2 P(X_k = -2^k) + 0^2 P(X_k = 0)$$
$$= 2^{2k} . 2^{-(2k+1)} + 2^{2k} . 2^{-(2k+1)} + 0.(1-2^{-2k})$$
$$= 2.2^{2k}\, 2^{-(2k+1)}$$
$$= 2^{1+2k-2k-1}$$
$$= 1$$

Thus, $\sigma_k^2 = E(X_k^2) - [E(X_k)]^2 = 1 - 0 = 1$

Using Liapounov's Theorem, we know that central limit theorem holds if

$$\lim_{n\to\infty}\frac{\rho}{\sigma} = 0$$

where $\rho^3 = \sum_{k=1}^{n} E|X_k - \mu_k|^3$

$$\sigma = \sqrt{\sum_{k=1}^{n} \sigma_k^2}$$

$$\mu_k = E(X_k)$$

Now,
$$E[|X_k - \mu_k|^3] = E[|X_k|^3]$$
$$= (2^k)^3 . 2^{-(2k+1)} + (2^k)^3 2^{-(2k+1)} + 0^3 (1 - 2^{-2k})$$
$$= 2^{3k}.2^{-(2k+1)} + 2^{3k}.2^{-(2k+1)}$$
$$= 2.2^{3k}.2^{-(2k+1)}$$
$$= 2^{1+3k-2k-1}$$
$$= 2^k$$

$$\rho^3 = \sum_{k=1}^{n} E\left[|X_k - \mu_k|^3\right] = \sum_{k=1}^{n} 2^k = 2 + 2^2 + ... + 2^n = \frac{2(2^n - 1)}{2-1} = 2(2^n - 1)$$

$\therefore$
$$\rho = \{2(2^n - 1)\}^{1/3} = 2^{1/3} (2^n - 1)^{1/3}$$

$\therefore$
$$\sigma = \sqrt{\sum_{k=1}^{n} \sigma_k^2} = \sqrt{\sum_{k=1}^{n} (1)} = \sqrt{n} = n^{1/2}$$

$\therefore$
$$\lim_{n\to\infty} \frac{\rho}{\sigma} = \lim_{n\to\infty} \frac{2^{1/3}(2^n - 1)^{\frac{1}{3}}}{n^{1/2}}$$

$$= 2^{1/3} \lim_{n\to\infty} \frac{(2^n - 1)^{1/3}}{n^{1/2}}$$

$$= 2^{1/3} \lim_{n\to\infty} \frac{\left[\left(2^n\right)^{1/3} - \frac{1}{3}\left(2^n\right)^{1/3-1} + \frac{1/3(1/3-1)}{\underline{|2}}\left(2^n\right)^{1/3-2} - ...\right]}{n^{1/2}}$$

$$= 2^{1/3} \lim_{n\to\infty} \left[\frac{2^{n/3}}{n^{1/2}} - \frac{1}{3}\frac{\left(2^n\right)^{-\frac{2}{3}}}{n^{1/2}} + ...\right]$$

$$= 2^{1/3} \lim_{n\to\infty} \frac{2^{n/3}}{n^{1/2}}$$ [$\because$ Limit of the remaining terms tends to 0 as $n \to \infty$]

Let
$$u_n = \frac{2^{n/3}}{n^{1/2}}$$

$\therefore$
$$u_{n+1} = \frac{2^{(n+1)/3}}{(n+1)^{1/2}}$$

$\therefore$
$$\frac{u_{n+1}}{u_n} = \frac{2^{1/3}}{\left(1+\frac{1}{n}\right)^{1/2}}$$

$\therefore$
$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \mathop{Lt}_{n\to\infty} \frac{2^{1/3}}{\left(1+\frac{1}{n}\right)^{1/2}} = 2^{1/3} > 1$$

$\therefore$
$$u_n \to \infty \text{ as } n \to \infty \text{ so } \frac{\rho}{\sigma} \to \infty \text{ as } n \to \infty$$

Hence, CLT does not hold.

Illustration 13.2. Let $X_1, X_2, ..., X_n$ be a sequence of *iid* random variables. Show that CLT holds for the following *pdf*

$$f(x) = \frac{1}{2} e^{-|x|}, \quad -\infty < x < \infty$$

Solution :$f(x) = \frac{1}{2} e^{-|x|}, \quad -\infty < x < \infty$

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

$$= \int_{-\infty}^{\infty} x.\frac{1}{2} e^{-|x|} dx$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} x e^{-|x|} dx$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x e^{-|x|} dx + \int_{0}^{\infty} x e^{-|x|} dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x e^{-(-x)} dx + \int_{0}^{\infty} x e^{-x} dx\right] \quad \left[\because |x| = x \text{ if } x>0 \atop = -x \text{ if } x<0\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x e^{x} dx + \int_{0}^{\infty} e^{-x} x^{2-1} dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x e^{x} dx + \underline{|2}\right]$$

$$= \frac{1}{2}\left[\int_{\infty}^{0}(-y)e^{-y}(-dy) + \lfloor 2-1 \right] \quad \text{[Let } x = -y,\ dx = -dy \text{ as } x = -\infty,\ y = \infty,\ x = 0,\ y = 0]$$

$$= \frac{1}{2}\left[-\int_{0}^{\infty} y^{2-1}e^{-y}dy + 1\right]$$

$$= \frac{1}{2}\left[-\lfloor 2 + 1\right]$$

$$= \frac{1}{2}[-1+1]$$

$$= 0$$

$$E(X^2) = \left[\int_{-\infty}^{\infty} x^2 f(x)dx\right]$$

$$= \int_{-\infty}^{\infty} x^2 \frac{1}{2} e^{-|x|}dx$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} x^2 e^{-|x|}dx$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x^2 e^{-|x|}dx + \int_{0}^{\infty} x^2 e^{-|x|}dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x^2 e^{-|x|}dx + \int_{0}^{\infty} x^2 e^{-x}dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} x^2 e^{x}dx + \int_{0}^{\infty} x^2 e^{-x}dx\right] \quad ...(4)$$

Now, $\int_{-\infty}^{0} x^2 e^{-x}dx = \int_{\infty}^{0}(-y)^2 e^{-y}(-dy)$

[Let $x = -y$, $dx = -dy$, $x = -\infty$, $y = \infty$, $x = 0$, $y = 0$]

$$= -\int_{\infty}^{0} y^2 e^{-y}dy$$

$$= \int_{0}^{\infty} e^{-y} y^{3-1}dy$$

$$= \lceil 3 = 2 \quad ...(5)$$

Again, $\int_{0}^{\infty} x^2 e^{-x}dx = \int_{0}^{\infty} e^{-x}x^{3-1}dx = \lceil 3 = \lfloor 3-1 = \lfloor 2 = 2$ ...(6)

Using (5) and (6) in (4) we have

$$E(X^2) = \frac{1}{2}[2+2] = \frac{1}{2}\times 4 = 2$$

$\therefore$ $V(X) = E(X^2) - [E(X)]^2 = 2 - 0 = 2$

Also, moment generating function of X is

$$M_X(t) = E(e^{tX})$$

$$= \int_{-\infty}^{\infty} e^{tx}\frac{1}{2}e^{-|x|}dx$$

$$= \frac{1}{2}\left[\int_{-\infty}^{\infty} e^{tx}.e^{-|x|}dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} e^{tx}e^{-|x|}dx + \int_{0}^{\infty} e^{tx}e^{-|x|}dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} e^{tx}e^{-(-x)}dx + \int_{0}^{\infty} e^{tx}e^{-x}dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} e^{tx}e^{x}dx + \int_{0}^{\infty} e^{tx}e^{-x}dx\right]$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} e^{(t+1)x}dx + \int_{0}^{\infty} e^{(t-1)x}dx\right] \quad ...(7)$$

$$= \frac{1}{2}\left[\int_{-\infty}^{0} e^{(t-1)x}dx\right]$$

Now, $\int_{-\infty}^{0} e^{(t+1)x}dx = \int_{-\infty}^{0} e^{(1+t)y}(-dy)$

[Let $x = -y \Rightarrow dx = -dy,\ x = -\infty, y = \infty,\ x = 0, y = 0$]

$$= \int_0^{\infty} e^{-(1+t)y}\, y^{1-1}\, dy$$

$$= \frac{\sqrt{1}}{1+t}$$

$$= \frac{1}{1+t} \quad \text{...(8)}$$

Also, $$\int_0^{\infty} e^{(t-1)x} x^{1-1}\, dx = \int_0^{\infty} e^{-(1-t)x} x^{1-1}\, dx$$

$$= \frac{\sqrt{1}}{1-t}$$

$$= \frac{1}{1-t} \quad \text{...(9)}$$

Using (8) and (9) in (7) we get

$$M_x(t) = \frac{1}{2}\left[\frac{1}{1+t} + \frac{1}{1-t}\right] = \frac{1}{2}\left[\frac{1-t+1+t}{(1+t)(1-t)}\right] = \frac{1}{1-t^2}$$

Now, $$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = n \times 0 = 0$$

$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = n \,.\, 2 = 2n$$

$$Z = \frac{\sum X_i - E(\sum X_i)}{\sqrt{V(\sum X_i)}} = \frac{\sum X_i - 0}{\sqrt{2n}} = \frac{\sum X_i}{\sqrt{2n}} \sim N(0, 1)$$

$$M_Z(\theta) = M_{\frac{\sum_{i=1}^{n} X_i}{\sqrt{2n}}}(t)$$

$$= \prod_{i=1}^{n} M_{\frac{X_i}{\sqrt{2n}}}(t)$$

$$= \left[M_{X_i}\left(\frac{t}{\sqrt{2n}}\right)\right]^n$$

$$= \left[\frac{1}{1-\left(\frac{t}{\sqrt{2n}}\right)^2}\right]^n \qquad \left[\because M_X(\theta) = \frac{1}{1-t^2}\right]$$

$$= \left(\frac{1}{1-\frac{t^2}{2n}}\right)^n$$

$$= \left[\left(1-\frac{t^2}{2n}\right)^{-1}\right]^n$$

$$= \left(1-\frac{t^2}{2n}\right)^{-n}$$

$$= \left(1-\frac{t^2/2}{n}\right)^{-n}$$

$\therefore$ $$\lim_{n\to\infty} M_Z(t) = \lim_{n\to\infty}\left(1-\frac{t^2/2}{n}\right)^{-n}$$

$$= e^{t^2/2}$$ which is the *mgf* of the standard normal variate

Hence CLT holds for the *pdf*

$$f(x) = \frac{1}{2} e^{-|x|}, \qquad -\infty < x < \infty$$

Illustration 13.3. A random sample of size 100 is taken from a population whose mean is 60 and variance is 400. Using CLT, with what probability can we assent that the mean of the sample will not differ from $\mu = 60$ by more than 4?

Solution : Given, sample size, $n = 100$

Population mean, $\mu = 60$

Variance, $\sigma^2 = 400$

$\therefore$ $\sigma = 20$

We know, by CLT, if $\{X_n\}$ is a sequence of independently and identically distributed random variables then

$$\frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$ is asympotically N(0, 1)

$$\text{Now, } P[(\bar{x}-\mu)>4] = P[(\bar{x}-60)>4]$$

$$= P\left[\frac{\bar{x}-60}{\sigma/\sqrt{n}} > \frac{4}{\sigma/\sqrt{n}}\right]$$

$$= P\left[\frac{\bar{x}-60}{20/10} > \frac{4}{20/10}\right]$$

$$= P[Z>2] \quad \left[\text{using CLT, } Z = \frac{\bar{x}-60}{20/10} \sim N(0, 1)\right]$$

$$= P(Z>0) - P(0 \le Z \le 2)$$

$$= P(Z>0) - \frac{1}{2} P(-2 \le Z \le 2)$$

$$= 0.5 - \frac{1}{2}(0.9544)$$

$$= 0.0228$$

$\therefore$ Probability that the mean of the sample will not differ from $\mu = 60$ by more than 4

$$= 1 - P\left[(\bar{X}-\mu)>4\right]$$

$$= 1 - 0.0228$$

$$= 0.9772$$

EXERCISE

1. State Liapounov's form of central limit theorem. Let $\{X_k\}$ be a sequence of independent Bernoulli variate with

   $P(X_k = 1) = p_k = 1 - P(X_k = 0)$

   Show that $\{X_k\}$ obeys central limit theorem. [IAS, 2002]

2. Explain the concept of central limit theorem. Let $\{X_n\}$ be a sequence of uniformly distributed random variables over $(-\beta n^{\lambda}, \beta n^{\lambda})$, $\beta > 0, \lambda > 0$. Test if strong law of large numbers, weak law of large numbers and central limit theorem holds. [IAS, 2000]

3. State and prove Lindeberg-Levi central limit theorem. [IAS, 1981]

4. Prove that central limit theorem holds for independent and identically distributed random variables with finite variance. [IAS, 1981]

5. State Lindeberg-Levy condition for the central limit theorem to hold. Prove with the usual notation that this condition implies:

   (*i*) $\left(\begin{matrix}\max \sigma_k \\ l \le k \le n\end{matrix}\right)/S_n \to 0$

   (*i*) $S_n \to \infty$ as $n \to \infty$

   [IAS, 1982]