9

Chapter Nine

Limit Theorems

Chapter Nine.

Limit Theorems

In Chapter Seven, we considered the joint distribution function of two random variables X and Y. Here we will examine some important properties of a large number of random variables $X_1, X_2, X_3, \ldots$ or equivalently, their distribution functions.

9.1. Sequence of Random Variables

Let us recall the definition of convergence of a sequence of real numbers $x_1, x_2, \ldots$ to a real number x. The sequence $\{x_n\}$ is said to be convergent to a limit x if for any $\varepsilon > 0$, chosen as small as possible, there is a positive integer m such that for n > m, we have

$$|x_n - x| < \varepsilon \text{ for all } n > m.$$

Now, let $X_1, X_2, \ldots$ be a sequence of random variables defined on the same probability space S . We shall assume that any finite collection $(X_1, X_2, \ldots, X_n), (n \geq 1)$ of n random variables are jointly distributed, that is , they have a joint distribution $F_n(x_1, x_2, \ldots, x_n)$ on R^n (n-Euclidean space) which is defined as

$$F_n(x_1, \ldots, x_n) = P\{X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n\}$$

where $x_1, x_2, \ldots, x_n$ are real numbers. The distribution function F_n satisfies the following properties

1. $0 \leq F_n(x_1, \dots, x_n) \leq 1$

2. $\lim\limits_{x_i \to -\infty} F_n(x_1, \dots, x_i, \dots, x_n) = 0$, for all $i = 1, 2, \dots, n$,

3. $\lim\limits_{\text{all } x_i \to \infty} F_n(x_1, \dots, x_n) = 1$, $i = 1, 2, \dots, n$.

4. $F_n(x_1, \dots, x_n)$ is right continuous in every variables.

In Probability theory, there are 4 kinds of convergence of the sequence $\{X_n\}$ of r.v.'s, which arc

1. Convergence in Probability (or Stochastic convergence)

A sequence of r.v.'s $\{X_n\}(n \geq 1)$ is said to be convergent in probability to a r.v.X (or weakly convergent), written as $X_n \xrightarrow{P} X$, , if for every $\varepsilon > 0$,

$$\lim_{n \to \infty} P\{|X_n - X| \geq \varepsilon\} = 0 \qquad \dots(9{\cdot}1{\cdot}1)$$

Equation (9.1.1) means that the probability of the event $\{|X_n - X| \geq \varepsilon\}$ tends to zero as $n \to \infty$. Also, it can be written as

$$P\{|X_n - X| < \varepsilon\} \to 1 \quad \text{as } n \to \infty.$$

In particular, the sequence $\{X_n\}$ converges stochastically to a constant a $(a \geq 0)$; written as $X_n \xrightarrow{P} a$, if for every $\varepsilon > 0$,

$$\lim_{n \to \infty} P\{|X_n - a| \geq \varepsilon\} = 0$$

If $X_n \xrightarrow{P} a$ and $Y_n \xrightarrow{P} b$ as $n \to \infty$, then we have

(i) $X_n \mp Y_n \to a \mp b$ as $n \to \infty$

(ii) $X_n Y_n \to a \cdot b$ as $n \to \infty$

(iii) $X_n / Y_n \to a / b$ as $n \to \infty$.

2. Convergence almost surely

A sequence of r.v.'s $\{X_n\}$ is said to be convergent almost surely (or strongly) to a r.v. X, written as $X_n \xrightarrow{a.s} X$ if

$$P\{\lim_{n \to \infty} X_n = X\} = 1$$

In this case, we write $\lim_{n \to \infty} X_n = X$ with probability one (simply w.p.1.)

In other words, for given $\varepsilon > 0$ if for almost all $w \in S$ (sample space) on which the r.v.'s $\{X_n\}$ and X are defined, we have

$$P\left[\bigcup_{n=N}^{\infty} \{|X_n(w) - X(w)| > \varepsilon\}\right] \to 0$$

as N $N \to \infty$. That is, the set of all $w \in S$ for which $X_n(w)$ does not converge to X(w) as $n \to \infty$ is an event and has probability zero, then the sequence $\{X_n\}$ a.s X.

3. Convergence in Mean

A sequence of r.v.'s $\{X_n\}$ is said to be convergent in r-th mean to a r.v. X if

$$EX_n^r < \infty, \ EX^r < \infty \text{ and}$$
$$\lim_{n \to \infty} [E|X_n - X|^r] = 0, \ r > 0$$

We use the notation $X_n \xrightarrow{r.m.} X$.

The case r=2 is known as convergence in quadratic mean, or mean square convergence, in this case we write $X_n \xrightarrow{m.s.} X$.

Theorem 9.1.1.

If a sequence of r.v.'s $\{X_n\}$ converges in mean square to X, then it also converges in probability to X.

Proof:

Recall the Chebyshev's inequality we have for any $\varepsilon > 0$

$$P\{|X_n - X| \geq \varepsilon\} \leq \frac{E[|X_n - X|^2]}{\varepsilon^2}$$

and, for fixed $\varepsilon > 0$, we have

$$\lim_{n \to \infty} P\{|X_n - X| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} \lim_{n \to \infty} E[|X_n - X|^2]$$

Therefore, $X_n \xrightarrow{P} X$.

4. Convergence in Distribution

The sequence $\{X_n\}$ of r.v.'s is said to be convergent in distribution (or in law) to the r.v.X if the distribution function $F_n(x)$ of X_n converges to the distribution function F(x) of X at every continuity point x of F. It is written as $X_n \xrightarrow{d} X$.

Theorem 9.1.2.

If $X_n \xrightarrow{p} X$ then the distribution function $F_n(x)$ of X_n tends to the distribution function F(x) of X at every continuity point of F(.)

(For the proof see Srinivasan and Mehata, p. 234.)

Definition 9.1.1

Let $X_1, X_2, X_3, \dots, X_n, \dots$ be a sequence of random variables with distribution function $F_n(x)$ (n=1,2, ...). We said that the sequence of distribution functions $\{F_n(x)\}$ is convergent if there exists a distribution function F(x) such that, at every continuity point of F(x), the relation

$$\lim_{n \to \infty} F_n(x) = F(x)$$

is hold.

To aviod the difficulty in finding the limit distribution function, specially that this limit is not defined at the discontinuity points, we use another approach which will be clarified by the following theorem (also for the proof of this theorem, see Srinivasan and Mehata, pp. 238-242.

Theorem 9.1.3.

Let $\phi_n(t)$ be the characteristic function of X_n. If $X_n \xrightarrow{d} X$ then $\phi_n(t) \to \phi(t)$, where $\phi(t)$ is the c.f. of X. If $\phi_n(t) \to \phi(t)$ and the limit function is continuous at t=0, then $X_n \xrightarrow{d} X$

9.2. The Law of Large Numbers

In many experiments, empirical data appear to obey a certain general law, that is, if an experiment is repeated n times and if an event E occurs with probability p, then with probability one the proprotion of successes of E approach p. This results is known as the strong law of large numbers.

Bernoulli's Law of Large Numbers

An n independent Bernoulli trials with probability of success p is performed.

If X is the number of successes in these n trials, then E(X)=np and Var (X)=npq (since X has a binomial distribution with parameters n and p). The variable X/n represents the proprotion of successes, we have

$$E\left(\frac{X}{n} \right) = \frac{E(X)}{n} = \frac{np}{n} = p$$

and

$$Var\left(\frac{X}{n} \right) = \frac{Var(X)}{n^2} = \frac{npq}{n^2} = pq/n .$$

Then, for given $\varepsilon > 0$,

$$P\left\{ \left| \frac{X}{n} - p \right| \geq \varepsilon \right\} \to 0 \text{ as } n \to \infty \qquad \dots (9{\cdot}2{\cdot}1)$$

From (9.2.1) it follows that

$$\frac{X}{n} \xrightarrow{P} p \text{ as } n \to \infty .$$

Theorem 9.2.1. The Weak Law of Large Numbers

Let $X_1, X_2, \dots,$ be be a sequence of independent and identically distributed random variables, each having a finite mean $\mu = E(X_i), i = 1, 2, \dots$ Then, for any $\varepsilon > 0$ and $Var(X_i) = \sigma^2 < \infty$,

$$P\left\{ \left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \geq \varepsilon \right\} \to 0 \text{ as } n \to \infty \quad \dots (9{\cdot}2{\cdot}2)$$

If we write

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

then, from (9.2.2), it follows

$$\bar{X}_n \xrightarrow{p} \mu \text{ as } n \to \infty,$$

where

$$\mu = E(\bar{X}_n).$$

By using Chebyshev's inequality, it can be easily proved (9.2.2),

If X_i, i=1,2, ... are independent r.v.'s with a finite mean $E(Xi)= \mu_i$ for i=1,2, ... Then

$$\bar{X}_n \xrightarrow{p} \bar{\mu}_n \text{ as } n \to \infty,$$

where

$$\text{where } \bar{\mu}_n = E(\bar{X}_n) = \frac{\Sigma \mu_i}{n},$$

provided that

$$\lim_{n \to \infty} Var(\bar{X}_n) = 0.$$

Theorem 9.2.2. The Strong Law of Large Numbers

Let $X_1, X_2, \ldots, \ldots$ be independent and identically distributed with a finite mean $\mu = E(X_i)$ and finite fourth central moment $\mu_4 = E(X_i - \mu)^4$ for i=1,2, ... Then

$$\bar{X}_n \xrightarrow{a.s.} \mu.$$

That is

$$P\{\lim_{n \to \infty} \bar{X}_n = \mu\} = 1.$$

9.3. Central Limit Theorem

Consider a sequence of independent and identically distributed random variables $\{X_n\}$ with mean $E(X_n) = \mu$ and $Var(X_n) = \sigma^2$. Let

$$S_n = \sum_{i=1}^{n} X_i, \quad \bar{X}_n = S_n / n$$

Then

$$E(\bar{X}_n) = E\frac{(X_1 + \dots + X_n)}{n} = \frac{n\mu}{n} = \mu,$$

and

$$Var(\bar{X}_n) = Var\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{n\sigma^2}{n^2} = \sigma^2 / n.$$

It is clear that $Var(\bar{X}_n) \to 0$ as $n \to \infty$. This means that the density function of $\bar{X}_n$ is concentrated around the mean μ for large n. We shall see later that as $n \to \infty$, the distribution function of $\bar{X}_n$ tends to the normal distribution under certain conditions. This property is one of the forms of the central limit theorem, which is

Theorem 9.3.1.

Let $X_1, X_2, \dots$ be a sequence of i.i.d. r.v.'s each witl mean μ and variance σ^2. Then the distribution of

$$Z_n = \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}}$$

tends to the standard normal as $n \to \infty$. That is

$$P\{Z_n \le z\} = \phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt. \text{ as } n \to \infty.$$

Proof:-

From Theorem 9.1.3, it is sufficient to show that $\phi_{Z_n}(t) \to \phi(t)$ as $n \to \infty$, where $\phi_{Z_n}(t)$ and $\phi(t)$ are the c.f.'s of Z_n and Z (standard normal distribution), respectively. W(have

$$E(S_n) = nE(X_i) = n\mu$$

and $Var(S_n) = n\, Var(X_i) = n\sigma^2$. .

Also we have

$$\phi_{S_n}(t) = \phi_{X_1 + \cdots + X_n}(t) = \prod_{i=1}^{n} \phi_{X_i}(t)$$

$$= [\phi_X(t)]^n \text{ because } X_i \text{ are i.i.d.r.v.'s.}$$

Therefore ,
$$Z_n = \frac{S_n - n\mu}{\sqrt{n}\,\sigma}$$

$$\phi_{Z_n}(t) = e^{-itn\mu/\sqrt{n}\,\sigma} \cdot \phi_{S_n}\left(\frac{t}{\sqrt{n}\,\sigma} \right)$$

$$= e^{-it\sqrt{n}\,\mu/\sigma} \left[\phi_X\left(\frac{t}{\sqrt{n}\,\sigma} \right) \right]^n$$

By the Taylor series expansion of $\phi_X(t)$, we get

$$\phi_X(t) = 1 + it\,E(X) - \frac{t^2}{2!} E(X^2) + 0(t^2).$$

Therefore

$$\phi_X\left(\frac{t}{\sqrt{n}\,\sigma} \right) = 1 + \frac{it}{\sqrt{n}\,\sigma} E(X) - \frac{t^2}{2n\sigma^2} E(X^2) + 0\left(\frac{t^2}{n\sigma^2} \right),$$

and

$$\phi_{Z_n}(t) = e^{\frac{it\sqrt{n}\,\mu}{\sigma}} \left[1 + \frac{itE(X)}{\sqrt{n\sigma}} - \frac{t^2}{2n\sigma^2} E(X^2) + 0\left(\frac{t^2}{n\sigma^2} \right) \right]^n,$$

$$= e^{-\frac{it\sqrt{n}\,\mu}{\sigma}} \left[1 + \frac{it\mu}{\sqrt{n}\,\sigma} - \frac{t^2}{2n\sigma^2} (\sigma^2 + \mu^2) + 0\left(\frac{t^2}{n\sigma^2} \right) \right]^n$$

$$= e^{-\frac{it\sqrt{n}\,\mu}{\sigma}} \left[1 + \frac{it\mu}{\sqrt{n}\,\sigma} - \frac{t^2}{2n} + 0\left(\frac{t^2}{n\sigma^2} \right) \right]^n$$

We have

$$\left[1 + \frac{it\mu\sqrt{n}/\sigma - t^2/2}{n} + 0\left(\frac{t^2}{n\sigma^2} \right) \right]^n$$

$$\to e^{\frac{it\mu\sqrt{n}}{\sigma} - t^2/2} \qquad \text{as } n \to \infty .$$

Therefore

$$\lim_{n \to \infty} \phi_{Z_n}(t) = e^{-\frac{it\sqrt{n}\,\mu}{\sigma}} \cdot e^{it\mu\sqrt{n}/\sigma - t^2/2} \qquad = e^{-t^2/2}$$

$$= \phi_Z(t),$$

Where $\phi_Z(t) = e^{-t^2/2}$ is the c.f. of the standard normal distribution.

By Theorem 9.1.3, the distribution of Z_n approaches the standard normal distribution.

Example 9.3.1

The number of students that have enrolled in a Probability course is a Poisson random variable with mean 80. The professor in charge of the course has decided that if the number enrolling is 100 or more he will teach the course in two seperate sections, otherwise he will teach it in one section. What is the probability that the professor will teach the course in one section?

Solution:-

The required probability is

$$P(X < 100) = \sum_{k=0}^{99} e^{-80} \frac{(80)^k}{k!}$$

If we assume that the Poisson random variable with mean 80 is the sum of 80 independent Poisson random variables each with mean 1, then by central limit theorem, we have

$$P(X < 100) = P\left(\frac{X - 80}{\sqrt{80}} < \frac{100 - 80}{\sqrt{80}} \right)$$

$$= P\left(Z < \frac{20}{8{\cdot}344}\right) = \phi(2{\cdot}396) = 0{\cdot}9916 .$$

9.4. Problems

1. Show that if $X_n \xrightarrow{p} X$, then $X_n \xrightarrow{d} X$.

2. Let $\{X_n\}$ be a sequence of independent r.v.'s such tha X_n has a binomial distribution with parameters n an p. If $Z_n = \dfrac{X_n - E(X_n)}{\sqrt{Var(X_n)}}$, then use the central limit theorem show that Z_n has N(0,1) as $n \to \infty$

3. Repeat Problem (2) if X_n has a Poisson distributic with parameter λ

4. Let X be the number of successes in 5000 Bernoulli trials wit probability of success of 0·7 on a given trial. Use the central limi theorem to estimate

   1. $P(X < 4950)$;
   2. $P(3920 < X < 4240)$.