**Probability**

**Jan-April 2014**

**Chennai Mathematical Institute**

**Professor Nandini Kannan**

# Chapter 6

# Limit Theorems

## 6.1 Introduction

## 6.2 Convergence in Probability

**Definition:** A sequence of random variables $\{X_n\}$ is said to converge in probability to a random variable $X$, in symbols

$$X_n \xrightarrow{P} X$$

if for every $\epsilon > 0$

$$P\{|X_n - X| \geq \epsilon\} \longrightarrow 0 \qquad \text{as } n \to \infty$$

or equivalently

$$P\{|X_n - X| < \epsilon\} \longrightarrow 1 \qquad \text{as } n \to \infty.$$

1. The $X_i's$ need not be iid but must all be defined on the same probability space.
2. The definition of convergence given here does not have the same interpretation as convergence in real analysis. Here $X_n \xrightarrow{P} X$ does not imply that, given $\epsilon > 0$ we can find $N$ such that $|X_n - X| < \epsilon$ for $n \geq N$. The definition speaks only of the convergence of the sequence of probabilities $P\{|X_n - X| \geq \epsilon\}$ to 0.

**Example:** Let $\{X_n\}$ be a sequence of random variables with pmf

$$P(X_n = 1) = \frac{1}{n} \qquad\qquad P(X_n = 0) = 1 - \frac{1}{n}.$$

We have

$$P\{|X_n| > \epsilon\} = \begin{cases} P(X_n = 1) = \frac{1}{n}, & 0 < \epsilon < 1; \\ 0, & \epsilon \geq 1. \end{cases}$$

This implies

$$P\{|X_n| > \epsilon\} \longrightarrow 0 \qquad \text{as } n \to \infty,$$

i.e. $X_n \xrightarrow{P} 0$.

$\blacksquare$

**Theorem 6.2.1.** *Weak Law of Large Numbers: Let $X_1, X_2, \ldots,$ be iid random variables with $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 < \infty$. Then*

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} \mu.$$

**Proof:** For every $\epsilon > 0$, an application of Chebyshev's inequality yields the following:

$$\begin{aligned} P(|\bar{X}_n - \mu| \geq \epsilon) &= P[(\bar{X}_n - \mu)^2 \geq \epsilon^2] \\ &\leq \frac{E(\bar{X}_n - \mu)^2}{\epsilon^2} \\ &= \frac{Var(\bar{X})}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2}. \end{aligned}$$

The last term goes to 0 as $n \to \infty$.

$\blacksquare$

The Weak Law of Large Numbers (WLLN) says that under very general conditions, the sample mean approaches the population mean as $n \to \infty$. This is also the definition of ***consistency***.

**Example:** Let $\{X_n\}$ be a sequence of iid $B(1, p)$ (Bernoulli) random variables. We have

$$E(X_i) = p \qquad Var(X_i) = p(1-p) < \infty.$$

We have $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ as the sample proportion of successes in $n$ trials. The WLLN states that the sample proportion converges in probability to the population proportion as $n \to \infty$.

$\blacksquare$

**Theorem 6.2.2.** *Let $X_n \xrightarrow{P} X$ and $Y_n \xrightarrow{P} Y$. Then*

1. $a\ X_n \xrightarrow{P} a\ X$
2. $X_n + Y_n \xrightarrow{P} X + Y$
3. $X_n Y_n \xrightarrow{P} XY$

*4.* $\frac{X_n}{Y_n} \xrightarrow{P} \frac{X}{Y}$ *if* $P(Y_n = 0) = 0 \ \ \forall n$ *and* $P(Y = 0) = 0$.

■

**Theorem 6.2.3.** *If $f(x)$ is a continuous, real valued function and $X_n \xrightarrow{P} X$, then*

$$f(X_n) \xrightarrow{P} f(X).$$

■

**Convergence in $r - th$ mean** A sequence of random variables $\{X_n\}$ is said to converge in the $r - th$ mean, in symbols

$$X_n \xrightarrow{r} X$$

if

$$E|X_n - X|^r \to 0 \qquad as \ \ n \to \infty.$$

**Theorem 6.2.4.** $X_n \xrightarrow{r} X \Rightarrow X_n \xrightarrow{P} X$. *If the $X_n$'s are almost surely bounded, the converse is true.*

**Counterexample:** Let

$$X_n = \begin{cases} n, & \text{with probability } 1/n; \\ 0, & \text{with probability } 1 - 1/n. \end{cases}$$

We have $P\{|X_n - 0| > \epsilon\} = \frac{1}{n}$ which tends to 0 as $n \to \infty$. This implies

$$X_n \xrightarrow{P} 0.$$

However, $E(X_n) = 1 \neq E(X) = 0$.

## 6.3 Convergence in Distribution

**Definition:** Let $\{X_n\}$ be a sequence of random variables and let $X$ be a random variable. Let $F_{X_n}(.)$ be the cdf of $X_n$, i.e.

$$F_{X_n}(x) = P(X_n \leq x),$$

and let $F_X(.)$ be the cdf of $X$. We say that $X_n$ ***converges in distribution (or law)*** to $X$, denoted by

$$X_n \xrightarrow{D} X$$

if

$$F_{X_n}(x) \to F_X(x) \qquad \text{as} \;\; n \to \infty$$

at all continuity points of $F_X$.

The critical assumption is that $F(-\infty) = 0$ and $F(\infty) = 1$, i.e. no probability mass escapes to $\pm\infty$.

**Example:** Consider the sequence of distribution functions

$$F_n(x) = \begin{cases} 0, & x < n; \\ 1, & x \geq n. \end{cases}$$

$F_n(x)$ is the cdf of the random variable $X_n$ degenerate at $n$. $F_n(x)$ converges to a function $F$ that is identically equal to 0 and hence is not a cdf.

■

**Example:** Let $X_1, \ldots, X_n$ be iid with common pdf

$$f(x) = \begin{cases} \frac{1}{\theta}, & 0 < x < \theta; \\ 0, & \text{o.w.} \end{cases}$$

Let $Y_n = max(X_1, \ldots, X_n)$. The cdf of $Y_n$ is

$$F_n(y) = \begin{cases} 0, & y < 0; \\ (y/\theta)^n, & 0 \leq y \leq \theta; \\ 1, & y > \theta. \end{cases}$$

As $n \to \infty$, we have

$$F_n(y) \to F(y) = \begin{cases} 0, & y < \theta; \\ 1, & y \geq \theta. \end{cases}$$

which is a cdf.
Let $Z_n = n(\theta - Y_n)$. What is the limiting distribution of $\{Z_n\}$?

■

**Theorem 6.3.1.** $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$.

■

To study the approximate distribution of a random variable $Y_n$, it is often necessary to study the behaviour of a normalized or rescaled version of $Y_n$, say

$$\frac{Y_n - b_n}{a_n}$$

for constants $a_n$ and $b_n$.

**Theorem 6.3.2.** *Slutsky's Theorem. If $X_n \xrightarrow{D} X$, and $A_n$ and $B_n$ tend in probability to constants $a$ and $b$ respectively, then*

$$A_n + B_n X_n \xrightarrow{D} a + bX.$$

■

**Theorem 6.3.3.** *Continuity Theorem: Let $\{F_n\}$ be a sequence of cdf's with corresponding mgf's $\{M_n\}$. Suppose $M_n(t)$ exists for $|t| \leq t_0$ for every $n$. If there exists a cdf $F$ with corresponding mgf $M$ which exists for $|t| \leq t_1 < t_0$ such that $M_n(t) \to M(t)$ as $n \to \infty$ for every $t \in [-t_1, t_1]$, then $F_n \to F$.*

■

**Lemma 6.3.4.** *Let $a_1, a_2, \ldots$ be a sequence of numbers converging to $a$, i.e. $\lim_{n\to\infty} a_n = a$. Then*

$$\lim_{n\to\infty} \left(1 + \frac{a_n}{n}\right)^n = e^a.$$

■

**Example:** Let $X_1, \ldots, X_n$ be iid $B(1, p)$. Then

$$\frac{S_n - np}{\sqrt{npq}} \xrightarrow{L} N(0, 1),$$

where $S_n = \sum_{i=1}^n X_i$.
**Proof:** We know that $S_n \sim Bin(n, p)$. The mgf of $S_n$ is

$$M_{S_n}(t) = (pe^t + q)^n.$$

What is the mgf of $Y_n = \frac{S_n - np}{\sqrt{npq}}$? Find the limit of the mgf of $Y_n$.

---

$\blacksquare$

**Theorem 6.3.5.** *Central Limit Theorem. Let $X_1, X_2, \ldots$ be a sequence of iid random variables whose mgf's exist in a neighbourhood of 0, i.e. $M_{X_i}(t)$ exists for $|t| < h$ for some positive $h$. Let $E(X_i) = \mu$ and $Var(X_i) = \sigma^2 > 0$. (Since the mgf exists, both $\mu$ and $\sigma^2$ are finite.)*

*Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, and let $G_n(x)$ denote the cdf of $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma}$. Then for any $x \in \mathcal{R}$, we have*

$$\lim_{n\to\infty} G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy,$$

*i.e.*

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{D} N(0,1).$$

**Proof:** Let

$$Y_i = \frac{X_i - \mu}{\sigma}.$$

Let $M_Y(t)$ denote the common mgf of the $Y_i's$ which exists for $|t| < \sigma h$. We have

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i.$$

Therefore

$$\begin{aligned}
M_{\frac{\sqrt{n}(\bar{X}_n-\mu)}{\sigma}}(t) &= M_{\frac{1}{\sqrt{n}}\sum_{i=1}^n Y_i}(t) \\
&= M_{\sum_{i=1}^n Y_i}\left(\frac{t}{\sqrt{n}}\right) \\
&= \left[M_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n.
\end{aligned}$$

Expanding $M_Y(t/\sqrt{n})$ in a Taylor series around 0, we have

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = \sum_{k=0}^{\infty} M_Y^{(k)}(0)\frac{(t/\sqrt{n})^k}{k!},$$

where

$$M_Y^{(k)}(0) = \frac{d^k}{dt^k}M_Y(t)\mid_{t=0}.$$

The expansion is valid if $t < \sqrt{n}\sigma h$. We have

$$M_Y^{(0)} = 1 \qquad M_Y^{(1)} = 0 \qquad M_Y^{(2)} = 1.$$

Substituting these values, we have

$$M_Y\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + R_Y\left(\frac{t}{\sqrt{n}}\right).$$

Using Taylor's theorem, for fixed $t \neq 0$, the remainder term is such that

$$\lim_{n\to\infty}\frac{R_Y(t/\sqrt{n})}{(t/\sqrt{n})^2} = 0.$$

Since $t$ is fixed, we also have

$$\lim_{n\to\infty}\frac{R_Y(t/\sqrt{n})}{(1/\sqrt{n})^2} = \lim_{n\to\infty} nR_Y(t/\sqrt{n}) = 0,$$

which is also true at $t = 0$, since

$$R_Y(0/\sqrt{n}) = 0.$$

Therefore, for any fixed $t$, we have

$$\begin{aligned}
\lim_{n\to\infty}\left[M_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n &= \lim_{n\to\infty}\left[1 + \frac{1}{n}\left(\frac{t^2}{2} + nR_Y\left(\frac{t}{\sqrt{n}}\right)\right)\right]^n \\
&= e^{t^2/2}.
\end{aligned}$$

The limit is the mgf of a standard normal random variable. Using the continuity theorem, we have the result.

■

**Example:** Let $X_1, \ldots, X_n$ be iid $\chi^2(1)$. Then

$$\frac{S_n - n}{\sqrt{2n}} \xrightarrow{D} N(0,1),$$

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

■