Chapter 8
Limit Theorems

Lectures 35 - 40
Most important limit theorems in probability are ‘law of large numbers’ and ‘central limit theorems’.

Law of large numbers describes the asymptotic behavior of the averages \( \frac{X_1 + \cdots + X_n}{n} \), where \( \{X_n \mid n \geq 1\} \) is a sequence of random variables whereas central limit theorems describe the asymptotic behavior of appropriately scaled sum of random variables.

To describe the asymptotic behavior, for example in law of large numbers, one should define the meaning of

\[
\lim_{n \to \infty} \frac{X_1 + \cdots + X_n}{n}.
\]

i.e. one need to talk about convergence of random variables. There are multiple ways one can define convergence of sequence of random variables.

**Definition 7.1.** Let \( X_n, n \geq 1, X \) be random variables defined on a probability space. Then \( X_n \) is said to converges almost surely if

\[
P\{ \lim_{n \to \infty} X_n = X \} = 1.
\]

If \( X_n \) converges to \( X \) almost surely, we write \( X_n \to X \) a.s.

**Definition 7.2.** Let \( X_n, n \geq 1, X \) be random variables defined on a probability space. Then \( X_n \) is said to converge to \( X \) in Probability, if for each \( \epsilon > 0 \),

\[
\lim_{n \to \infty} P\{|X_n - X| > \epsilon\} = 0.
\]

Now we prove some very useful inequalities.

**Theorem 8.0.39** (Markov inequality) Let \( X \) be a non negative random variable with finite \( n \)th moment. Then we have for each \( \epsilon > 0 \)

\[
P\{X \geq \epsilon\} \leq \frac{EX^n}{\epsilon^n}.
\]

**Proof.** Set

\[
Y(\omega) = \begin{cases} 
0 & \text{if } X(\omega) < \epsilon \\
\epsilon^n & \text{if } X(\omega) \geq \epsilon.
\end{cases}
\]

Then \( Y \) is a non negative simple random variable such that \( Y \leq X^n \). Hence from Theorem 6.0.29, we have \( EY \leq EX^n \). Therefore

\[
\epsilon^n P\{X \geq \epsilon\} = EY \leq EX^n.
\]

This completes the proof.

As a corollary we have the Chebyshev’s inequality.
Chebyshev's inequality. Let $X$ be a random variable with finite mean $\mu$ and finite variance $\sigma^2$.

Then for each $\epsilon > 0$

$$P\{ |X - \mu| \geq \epsilon \} \leq \frac{\sigma^2}{\epsilon^2}$$

The proof of Chebyshev's inequality follows by replacing $X$ by $|X - \mu|$ in the Markov inequality.

There are two types of law of large numbers, weak law of large numbers and strong law of large numbers. Weak law of large numbers describe the asymptotic behavior of averages of random variables using convergence in probability and strong law of large numbers describe limit of averages of random variables using almost sure convergence.

**Theorem 8.0.40 (Weak law of large numbers)** Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean $\mu$ and finite variance $\sigma^2$. Then $\frac{S_n}{n}$ converges in probability to $\mu$. i.e., for each $\epsilon > 0$,

$$\lim_{n \to \infty} P\{ \left| \frac{S_n}{n} - \mu \right| \geq \epsilon \} = 0,$$

where $S_n = X_1 + \cdots + X_n$.

**Proof:** Note that

$$E\left( \frac{S_n}{n} \right) = \mu, \quad E\left( \frac{S_n}{n} - \mu \right)^2 = \frac{\sigma^2}{n}.$$

Hence applying Chebychev’s inequality to $\frac{S_n}{n}$ we get

$$P\{ \left| \frac{S_n}{n} - \mu \right| \geq \epsilon \} \leq \frac{\sigma^2}{n\epsilon^2}.$$

Therefore

$$\lim_{n \to \infty} P\{ \left| \frac{S_n}{n} - \mu \right| \geq \epsilon \} = 0.$$

**Theorem 8.0.41 (Strong law of large numbers)** Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed random variables, each having finite mean $\mu$ and finite second moment. Then

$$\lim_{n \to \infty} \frac{S_n}{n} \to \mu \text{ a.s.}$$

**Proof:**
\begin{align*}
(S_n - n\mu)^4 &= \left( \sum_{i=1}^{n}(X_i - \mu) \right)^4 \\
&= \sum_{i=1}^{n}(X_i - \mu)^4 + 4 \sum_{i \neq j}^{n}(X_i - \mu)^3(X_j - \mu) \\
&\quad + 6 \sum_{i \neq j}^{n}(X_i - \mu)^2(X_j - \mu)^2 \\
&\quad + 12 \sum_{i,j,k \text{ distinct}}^{n}(X_i - \mu)^2(X_j - \mu)(X_k - \mu) \\
&\quad + 24 \sum_{i,j,k,l \text{ distinct}}^{n}(X_i - \mu)(X_j - \mu)(X_k - \mu)(X_l - \mu)
\end{align*}

Hence using the fact that \((X_i - \mu), (X_j - \mu), (X_k - \mu), (X_l - \mu)\) are independent for \(i, j, k, l\) distinct, we get

\[ E(S_n - n\mu)^4 = \sum_{i=1}^{n} E(X_i - \mu)^4 + 6 \sum_{i \neq j}^{n} E(X_i - \mu)^2 E(X_j - \mu)^2. \]  

(8.0.1)

Since \(X_1, X_2, \ldots\) are identically distributed, we get

\[ E(S_n - n\mu)^4 = n E(X_1 - \mu)^4 + 3n(n-1) E(X_1 - \mu)^2 E(X_2 - \mu)^2 \leq nK + 3n(n-1)K, \]  

(8.0.2)

where \(K = E(X_1 - \mu)^4\). Hence

\[ E \left[ \frac{S_n}{n} - \mu \right]^4 \leq \frac{K}{n^3} + \frac{3K}{n^2}. \]  

(8.0.3)

Therefore

\[ E \left[ \sum_{n=1}^{\infty} \frac{S_n}{n} - \mu \right]^4 = \sum_{n=1}^{\infty} E \left[ \frac{S_n}{n} - \mu \right]^4 \leq K \sum_{n=1}^{\infty} \frac{1}{n^3} + 3K \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty. \]  

(8.0.4)

Hence

\[ P \left\{ \sum_{n=1}^{\infty} \left[ \frac{S_n}{n} - \mu \right]^4 < \infty \right\} = 1. \]

Therefore

\[ P \left\{ \lim_{n \to \infty} \frac{S_n}{n} = \mu \right\} = 1. \]

This completes the proof.

As an application we show that any continuous function can be approximated by Bernstein polynomials.

**Example 8.0.45** Let \( f : [0, 1] \to \mathbb{R} \) be a continuous function. Consider the Bernstein polynomials
Fix \( x \in (0,1) \). Let \( X_1, X_2, \ldots \) be independent and identically distributed Bernoulli \( (x) \) random variables. Then using strong law of large numbers, we have
\[
\frac{S_n}{n} \to x \text{ a.s. as } n \to \infty.
\]
Now note that \( S_n \) is Binomial \( (n, x) \) random variable. Hence
\[
B_n(x) = E\left[f\left(\frac{S_n}{n}\right)\right].
\]
Set
\[
Y_n = f\left(\frac{S_n}{n}\right).
\]
Then \( Y_n \to f(x) \) a.s. as \( n \to \infty \) and \( |Y_n| \leq K \) where \( K \) such that \(-K \leq f(x) \leq K\).

Here we use the fact that every continuous function defined on \([0, 1]\) is bounded. Now apply the dominated convergence theorem (Theorem 6.0.31), we get
\[
\lim_{n \to \infty} EY_n = f(x).
\]
i.e.
\[
\lim_{n \to \infty} B_n(x) = f(x), \quad 0 < x < 1.
\]
For \( x = 0, x = 1 \), the proof follows by observing that \( B_n(0) = f(0) \) and \( B_n(1) = f(1) \).

**Theorem 8.0.42** (Central limit theorem) Let \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed random variables, each having finite mean \( \mu \) and finite non zero variance \( \sigma^2 \).
Then
\[
\lim_{n \to \infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] = N(x), \quad x \in \mathbb{R},
\]
where \( N(\cdot) \) is the standard normal distribution function.

**Proof:**
Set
\[
\bar{S}_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}.
\]
For \( t \geq 0 \),
\[
\Phi_{\bar{S}_n}(t) = e^{-i\mu\frac{t}{\sigma\sqrt{n}}} \Phi_{S_n}(\frac{t}{\sigma\sqrt{n}})
= e^{-i\mu\frac{t}{\sigma\sqrt{n}}} (\Phi_{X_1}(\frac{t}{\sigma\sqrt{n}}))^n,
\]
(8.0.5)
where the second inequality uses the fact that \( X_1, X_2, \ldots \) be a sequence of independent and identically distributed.

For each fixed \( t > 0 \), for sufficiently large \( n \), \( \Phi_{X_1}(\frac{t}{\sigma \sqrt{n}}) \) is close to \( 1 \). Hence for sufficiently large values of \( n \), we have from (8.0.5),

\[
\Phi_{S_n}(t) = e^{n[\ln(\Phi_{X_1}(\frac{t}{\sigma \sqrt{n}})) - i\mu(\frac{t}{\sigma \sqrt{n}})]},
\]

(8.0.6)

Hence for \( t \neq 0 \),

\[
\lim_{n \to \infty} n[\ln(\Phi_{X_1}(\frac{t}{\sigma \sqrt{n}})) - i\mu(\frac{t}{\sigma \sqrt{n}})] = \frac{t^2}{\sigma^2} \lim_{n \to \infty} \frac{\ln(\Phi_{X_1}(\frac{t}{\sigma \sqrt{n}})) - i\mu(\frac{t}{\sigma \sqrt{n}})}{(\frac{t}{\sigma \sqrt{n}})^2}
\]

(8.0.7)

Consider

\[
\lim_{n \to \infty} \frac{\ln(\Phi_{X_1}(\frac{t}{\sigma \sqrt{n}})) - i\mu(\frac{t}{\sigma \sqrt{n}})}{(\frac{t}{\sigma \sqrt{n}})^2} = \lim_{x \to 0} \frac{\ln(\Phi_{X_1}(x)) - i\mu x}{x^2} = \frac{\Phi'_{X_1}(x) - i\mu \Phi_{X_1}(x)}{x} = \frac{\Phi''_{X_1}(0) - i\mu \Phi'(0)}{2} = -\frac{\sigma^2}{2}.
\]

The last two equalities follow from l'Hospital's rule in view of Theorem 7.0.35. Combining (8.0.6), (8.0.6) and (8.0.6), we get, for each \( t \neq 0 \),

\[
\lim_{n \to \infty} \Phi_{S_n}(t) = e^{-\frac{t^2}{2}}.
\]

For \( t = 0 \), the above limit follows easily. Hence for each \( t \in \mathbb{R} \), we have

\[
\lim_{n \to \infty} \Phi_{S_n}(t) = e^{-\frac{t^2}{2}}.
\]

(8.0.9)

Hence using Theorem 7.0.38, we have from (7.0.38), we complete the proof.

As a corollary to Central limit theorem, one can get DeMoivre-Laplace limit theorem.

**Theorem 8.0.43** Let \( S_n \) be Bernoulli \((n, p)\) random variable. Then

\[
\lim_{n \to \infty} P \left\{ \frac{S_n - np}{\sqrt{np(1 - p)}} \leq x \right\} = N(x), \quad x \in \mathbb{R}.
\]

**Proof:** Observe that if \( X_1, X_2, \ldots \) are independent and identically distributed Binomial \( (p) \) random variables, then \( X_1 + \ldots + X_n \) is Bernoulli \((n, p)\) random variable. Now apply Central limit theorem, we get DeMoivre-Laplace limit theorem.

One can use central limit theorem to get the normal approximation formula. Under the hypothesis of central limit theorem, we have

\[
P \left\{ \frac{S_n - n\mu}{\sigma \sqrt{n}} \leq x \right\} \equiv N(x).
\]

i.e.
\[ P\{S_n \leq x\} \equiv N\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right). \]  \hspace{1cm} (8.0.10)

Here \( \equiv \) means "approximately equal to".

**Example 8.0.46**  Let \( X_1, X_2, \ldots \) be an independent and identically distributed Bernoulli (p) random variables. Then using (8.0.10), we get for \( x \) a non negative integer,

\[
P\{S_n = x\} = P\left\{ x - \frac{1}{2} < S_n \leq x + \frac{1}{2} \right\}
\equiv N\left(\frac{x + (1/2) - n\mu}{\sigma\sqrt{n}}\right) - N\left(\frac{x - (1/2) - n\mu}{\sigma\sqrt{n}}\right),
\]

where

\[ \mu = p, \quad \sigma^2 = p(1 - p). \]