| | |
|---|---|
| **Modern Statistics** | **Lecture 7 - 3/10/2025** |

# Lecture 7 Convergence of R.V.

*Lecturer:Xiangyu Chang* *Scribe: Ruishi Wang,Mengyao Wang*

*Edited by: Zhihong Liu*

## 1 Recall

- $X_n \xrightarrow{P} X,\ \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$
- $X_n \xrightarrow{d} X,\ \lim_{n\to\infty} F_n(t) = F_X(t)$ , for all $t$ for which $F_X$ is continuous.
- $X_n \xrightarrow{L^2} X,\ \lim_{n\to\infty} \mathbb{E}[(X_n - X)^2] = 0.$

## 2 Some Proof about Types of Convergence

### 2.1 Classical Example

$$X_n \sim N(0, \frac{1}{n})$$

**To prove**:

- ① $X_n \xrightarrow{d} X,\ P(X = 0) = 1.$

**Proof**:

$$\begin{aligned} F_n(t) &= \mathbb{P}(X_n \le t) \\ &= \mathbb{P}(\sqrt{n}X_n \le \sqrt{n}t) \\ &= \Phi(\sqrt{n}t), \end{aligned}$$

where $\sqrt{n}X_n \sim N(0,1)$

$$\lim_{n\to\infty} F_n(t) = 1, \quad \text{for } t > 0.$$

$$\lim_{n\to\infty} F_n(t) = 0, \quad \text{for } t < 0.$$

$$\lim_{n\to\infty} F_n(t) = \frac{1}{2}, \quad \text{when } t = 0.$$

$$F_X(t) = \begin{cases} 1 & \text{if } t \geq 0, \\ 0 & \text{if } t < 0. \end{cases}$$

$$\begin{aligned} &\text{For } t > 0: \quad \lim_{n\to\infty} F_n(t) = F_X(t) = 1, \\ &\text{For } t < 0: \quad \lim_{n\to\infty} F_n(t) = F_X(t) = 0, \\ &\text{At } t = 0: \quad \lim_{n\to\infty} F_n(0) = \frac{1}{2} \neq F_X(0) = 1 \quad \text{(Point of discontinuity)}. \end{aligned}$$

$$\therefore \quad X_n \xrightarrow{d} X.$$

Supplementary Theorem

**Theorem 1** (Markov Inequality)**.** *If random variable $X > 0$, then for all $\epsilon > 0$,*

$$\mathbb{P}(X > \epsilon) \leq \frac{\mathbb{E}(X)}{\epsilon}.$$

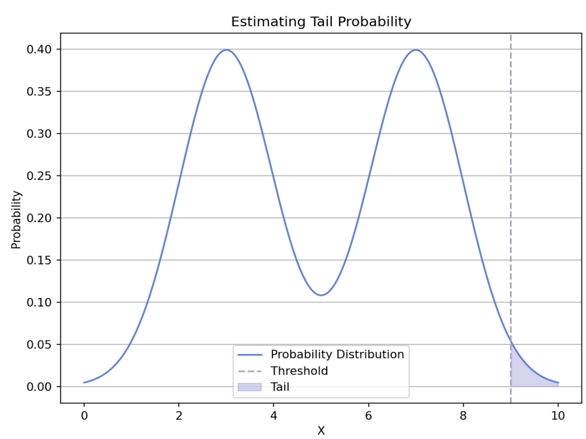


图 1: Markov Inequality Example Figure

**Proof**:

Let $y_1 = \epsilon I(x > \epsilon)$ and $y_2 = x(x \geq \epsilon)$.

Since

$$y_1 \leq y_2,$$

then

$$\mathbb{E}(y_1) \le \mathbb{E}(y_2),$$

$$\epsilon\mathbb{E}[I(X \ge \epsilon)] \le \mathbb{E}(X),$$

$$\epsilon\mathbb{P}(X \ge \epsilon) \le \mathbb{E}(X),$$

$$\mathbb{P}(X \ge \epsilon) \le \frac{\mathbb{E}(X)}{\epsilon}.$$

**Theorem 2** (Chebyshev's Inequality)**.** *Given a sequence* $\{X_n\}$, $\mathbb{E}(X_n) = \mu$, $Var(X_n) = \sigma^2$.

$$\Rightarrow \mathbb{P}(|X_n - \mu| \ge \epsilon) \le \frac{\sigma^2}{\epsilon^2}.$$

**Proof**:

$$\begin{aligned}\mathbb{P}(|X_n - \mu| \ge \epsilon) &= \mathbb{P}((X_n - \mu)^2 \ge \epsilon^2) \\ &\le \frac{\mathbb{E}[(X_n - \mu)^2]}{\epsilon^2} \\ &= \frac{\sigma^2}{\epsilon^2}.\end{aligned}$$

**Tips**.
(1) Concentration Inequality
(2)

$$\mathbb{P}(\exp(X_n - \mu) \ge \exp(\epsilon)) \le \frac{\mathbb{E}[\exp(X_n - \mu)]}{\exp(\epsilon)}.$$

Supplementary Theorem

- ② $X_n \xrightarrow{P} 0$.

  **Proof**:

$$\begin{aligned}\mathbb{P}(|X_n - 0| \ge \epsilon) &= \mathbb{P}((X_n - 0)^2 \ge \epsilon^2) \\ &= \mathbb{P}((X_n - \mathbb{E}(X_n))^2 \ge \epsilon^2) \\ &\le \frac{Var(X_n)}{\epsilon^2} \quad (by \quad Theorem 2) \\ &= \frac{1}{n\epsilon^2} \to 0.\end{aligned}$$

## 2.2 Relationship between types of convergence

**Theorem 3.** ①

$$X_n \xrightarrow{L^2} X \Rightarrow X_n \xrightarrow{P} X.$$

②

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X.$$

③

$$X_n \xrightarrow{d} X, \mathbb{P}(X = c) = 1 \Rightarrow X_n \xrightarrow{P} X.$$

Summary:

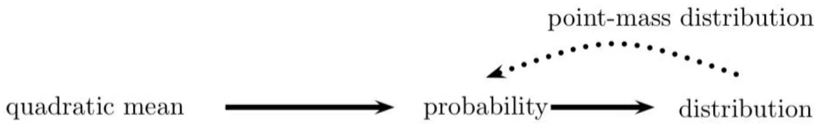


图 2: Relationship between Types of Convergence.

**Proof**:

①

$$\begin{aligned}\mathbb{P}(|X_n - X| \geq \epsilon) &= \mathbb{P}((X_n - X)^2 \geq \epsilon^2) \\ &\leq \frac{\mathbb{E}[(X_n - X)^2]}{\epsilon^2} \overset{\because X_n \overset{L^2}{\to} X}{\to} 0.\end{aligned}$$

②

$$\begin{aligned}F_n(x) &= \mathbb{P}(X_n \leq x) \\ &= \mathbb{P}(X_n \leq x, X > \ x + \epsilon) + \mathbb{P}(X_n \leq x, X \leq x + \epsilon) \\ &\leq \mathbb{P}(X_n \leq x, X > \ x + \epsilon) + \mathbb{P}(X \leq x + \epsilon) \\ &\leq \mathbb{P}(|X - X_n| \geq \epsilon) + F_X(x + \epsilon).\end{aligned}$$

$$\begin{aligned}F_X(x - \epsilon) &= \mathbb{P}(X \leq x - \epsilon) \\ &= \mathbb{P}(X \leq x - \epsilon, X_n \leq x) + \mathbb{P}(X \leq x - \epsilon, X_n > \ x) \\ &\leq \mathbb{P}(X_n \leq x) + \mathbb{P}(X_n > X + \epsilon) \\ &\leq F_n(x) + \mathbb{P}(|X_n - X| > \epsilon).\end{aligned}$$

Therefore,

$$F_X(x - \epsilon) - \mathbb{P}(|X_n - X| > \epsilon) \leq F_n(x) \leq F_X(x + \epsilon) + \mathbb{P}(|X_n - X| \geq \epsilon).$$

As $n \to \infty$, $F_n(x) \to F_X(x)$ at $x$ which is a continuous point of $F_X$.

③

$$X_n \xrightarrow{d} X, \mathbb{P}(X = c) = 1 \Rightarrow X_n \xrightarrow{P} X.$$

$$
\begin{aligned}
\mathbb{P}(|X_n - c| > \varepsilon) &= \mathbb{P}(X_n \geq c + \varepsilon) + \mathbb{P}(X_n \leq c - \varepsilon) \\
&= F_n(c - \varepsilon) + 1 - F_n(c + \varepsilon).
\end{aligned}
$$

As $n \to \infty$,

$$
\begin{aligned}
&F_X(c - \varepsilon) + 1 - F_X(c + \varepsilon) \\
&= 0 + 1 - 1 \\
&= 0
\end{aligned}
$$

# 3 Main Theorems

## 3.1 Weak Law of Large Numbers (WLLN)

**Theorem 4.** *If $X_1, \ldots, X_n$ are IID with $\mathbb{E}X_i = \mu$, then $\overline{X}_n \xrightarrow{\text{P}} \mu$.*

**Proof**:
By Chebyshev's inequality:

$$\mathbb{P}(|\overline{X}_n - \mu| > \epsilon) \leq \frac{\mathrm{V}(\overline{X}_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \to 0 \quad (n \to \infty).$$

## 3.2 Central Limit Theorem (CLT)

**Theorem 5.** *Let $X_1, \ldots, X_n$ be IID with mean $\mu$ and variance $\sigma^2$. Define $\overline{X}_n = n^{-1} \sum_{i=1}^n X_i$. Then:*

$$Z_n \equiv \frac{\overline{X}_n - \mu}{\sqrt{\mathrm{V}(\overline{X}_n)}} = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \xrightarrow{\text{d}} Z \sim N(0, 1).$$

## 3.3 Slutsky's Theorem

**Theorem 6.** *Let $X_n, Y_n$ be random sequences:*

*(a)* $X_n \xrightarrow{\text{P}} X,\, Y_n \xrightarrow{\text{P}} Y \implies X_n + Y_n \xrightarrow{\text{P}} X + Y.$

*(b)* $X_n \xrightarrow{L^2} X,\, Y_n \xrightarrow{L^2} Y \implies X_n + Y_n \xrightarrow{L^2} X + Y.$

*(c)* $X_n \xrightarrow{\text{d}} X,\, Y_n \xrightarrow{\text{d}} c \implies X_n + Y_n \xrightarrow{\text{d}} X + c.$

*(d)* $X_n \xrightarrow{\text{d}} X,\, Y_n \xrightarrow{\text{d}} c \implies X_n Y_n \xrightarrow{\text{d}} c \cdot X.$

*(e)* $X_n \xrightarrow{\text{P}} X,\, Y_n \xrightarrow{\text{P}} Y \implies X_n Y_n \xrightarrow{\text{P}} XY.$

## 3.4 Continuous Mapping Theorem (CMT)

**Theorem 7.** *Let $g$ be a continuous function:*

*(1)* $X_n \xrightarrow{\text{P}} X \implies g(X_n) \xrightarrow{\text{P}} g(X)$.

*(2)* $X_n \xrightarrow{\text{d}} X \implies g(X_n) \xrightarrow{\text{d}} g(X)$.

## 3.5 Multivariate Central Limit Theorem

**Theorem 8.** *Let $X_1, \dots, X_n$ be IID random vectors:*

$$X_i = \begin{pmatrix} X_{1i} \\ X_{2i} \\ \vdots \\ X_{ki} \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{pmatrix}, \quad \Sigma = Var(X_i).$$

*Define* $\overline{X} = \begin{pmatrix} \overline{X}_1 \\ \overline{X}_2 \\ \vdots \\ \overline{X}_k \end{pmatrix}$. *Then:*

$$\sqrt{n}(\overline{X} - \mu) \xrightarrow{d} N(0, \Sigma).$$

## 3.6 Delta Method

**Theorem 9.** *Suppose* $\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$ *and $g$ is differentiable.*
*Then:*

$$\frac{\sqrt{n}(g(\overline{X}_n) - g(\mu))}{\sigma} \xrightarrow{d} N\left(0, [g'(\mu)]^2\right).$$

**Proof**:
By Taylor expansion:

$$g(\overline{X}_n) - g(\mu) = g'(\mu)(\overline{X}_n - \mu) + \frac{1}{2} g''(\xi)(\overline{X}_n - \mu)^2$$
$$\frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma} = \frac{\sqrt{n}(g(\overline{X}_n) - g(\mu))}{\sigma \cdot g'(\mu)} - \underbrace{\frac{\sqrt{n} g''(\xi)(\overline{X}_n - \mu)^2}{2\sigma g'(\mu)}}_{R_n}.$$

Analysis

1. Left-hand side: $\xrightarrow{d} N(0,1)$.
2. $|X_n - \mu| = O_p(n^{-1/2}) \implies R_n = O_p(n^{-1/2}) \to 0$ (in probability).

# 4 Applications

**Problem** Prove that $\frac{\sqrt{n}(\overline{X}_n-\mu)}{S_n} \xrightarrow{d} N(0,1)$.

**Proof**:
expand the sample variance:

$$\begin{aligned}
S_n^2 &= \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X}_n)^2 \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\overline{X}_n^2\right] \\
&= \frac{n}{n-1}\left[\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}_n^2\right].
\end{aligned}$$

Key Steps

1. $X_n^2 \xrightarrow{P} \mu^2$ (CMT).
2. $\frac{1}{n}\sum X_i^2 \xrightarrow{P} \mathbb{E}(X^2) = \sigma^2 + \mu^2$ (WLLN).
3. $\frac{n}{n-1} \xrightarrow{n\to\infty} 1$.
4. $S_n^2 \xrightarrow{P} \sigma^2 \implies S_n \xrightarrow{d} \sigma \quad$ (CMT).

Conclusion

$$\frac{\sqrt{n}(\overline{X}_n-\mu)}{S_n} = \frac{\sqrt{n}(\overline{X}_n-\mu)}{\sigma}\cdot\frac{\sigma}{S_n} \xrightarrow{d} N(0,1)\cdot 1 \quad \text{(CLT \& Slutsky)}.$$