# Introduction to Estimation Theory

Wei-Ning Chen

May 25, 2016

# OUTLINE

# Convergence in Distribution

**Definition (Convergence in Distribution)**

Let random variables $X_n$ and $X$ have distributions $F_n(\cdot)$ and $F(\cdot)$ respectively. Then $X_n$ is said to converge in distribution to $X$, if

$$F_n \Rightarrow F$$

or equivalently,

$$\lim_{n\to\infty} Pr[X_n \leq x] = Pr[X \leq x]$$

for evey $x$ such that $Pr[X = x] = 0$.

Remark: We only require $F_n(\cdot)$ to converge at every continuous point!

### EXAMPLE

Take $p_{n,k} = \frac{\lambda}{n}$

$$\mu\{k\} = \binom{n}{k}(\frac{\lambda}{n})^k(1-\frac{\lambda}{n})^(n-k), \text{ for } 0 \leq k \leq k$$

Then

$$\mu_n \Rightarrow \text{Poisson}(\lambda).$$

### Theorem

$X_n \xrightarrow{\mathcal{P}} X$ *implies* $X_n \Rightarrow X$

### Proof.

Obseve that $Pr[A \leq a] - Pr[|A - B|] \leq Pr[B \leq a + b]$.
Then

$$Pr[X \leq x - \epsilon] - Pr[|X_n - X| > \epsilon] \leq Pr[X_n \leq (x - \epsilon) + \epsilon] = Pr[X_n \leq x],$$

and

$$Pr[X_n \leq x] - Pr[|X_n - X| > \epsilon] \leq Pr[X \leq x + \epsilon].$$

Hence we get

$$\begin{aligned} Pr[X \leq x - \epsilon] - Pr[|X_n - X| > \epsilon] &\leq Pr[X_n \leq x] \\ &\leq Pr[X \leq x + \epsilon] + Pr[|X_n - X| > \epsilon] \end{aligned}$$

**CON'D.**

which implies that

$$\begin{aligned} Pr[X \leq x - \epsilon] &\leq \liminf Pr[X_n \leq x] \\ &\leq \limsup Pr[X_n \leq x] \\ &\leq Pr[X \leq X + \epsilon] \end{aligned}$$

Therefore, for each continuous point, we have

$$\lim_{n \to \infty} Pr[X_n \leq x] = Pr[X \leq x]$$

□

Remark: $X_n \Rightarrow X$ does not imply $X_n \xrightarrow{\mathcal{P}} X$

**EXAMPLE**

$X \perp\!\!\!\perp Y$ and $X \sim$ Ber(p), $Y \sim$Ber(p). Let $X_n = X$, for all $n$. Then we have

$$X_n \Rightarrow Y, \text{ but } X_n \not\to Y$$

Remark: Convergence in distribution gives NO information between the correlation of each random variables!

# Fundamental Theorem

**Theorem (Skorogod's Theorem)**

*Suppose $\mu_n$ and $\mu$ are probability measures on $(\mathbb{R}, \mathcal{B})$, and $\mu_n \Rightarrow \mu$. Then there exist random variables $Y_n$ and $Y$ such that:*

1. *they are both defined on common probability space $(\Omega, \mathcal{F}, \mathcal{P})$*
2. $Pr[Y_n \leq y] = \mu_n(-\infty, y]$, *for all* $y$
3. $Pr[Y \leq y] = \mu(-\infty, y]$, *for all* $y$
4. $\lim_{n\to\infty} Y_n(\omega) = Y(\omega)$, *for every* $\omega$ *in* $\Omega$

Remark: This implies that cdfs are sufficient; we do not need to rely on the inherited probability space.

### Historical Aspects

- CLT concerns the situation that the limit distribution of the normalized sum is normal
- As an example, for i.i.d. zero-mean sequence $X_1, X_2, ...$

$$\frac{X_1 + ... + X_n}{n\mathbb{E}[X_i^2]} \Rightarrow N$$

  where $N$ has standard normal distribution
- **Question:** What is the rare of convergence of normalized sum distribution to standard normal distrubution?

- The first convergence rate estimates in the CLT were obtainrd by A.M, Lyapounov in 1900-1901
- In the beginning of 1940s, the classic Berry-Esseen estimate came to the light

# Berry-Esseen Theorem

**Theorem (Berry-Esseen Theorem (i.i.d. case))**

$$\sup_{x \in \mathbf{R}} |F_n(x) - \Phi(x)| \leq C \frac{\beta_3}{\sigma^3 \sqrt{n}}$$

*where*

- $F_n$ *is the cdf of* $\dfrac{X_1 + \ldots + X_n}{n\mathbb{E}[X_i^2]}$
- $\Phi$ *is the standard normal cdf*
- $\beta_3 = \mathbb{E}[|X - \mathbb{E}X|^3]$
- $\sigma^2 = \mathbb{E}[|X - \mathbb{E}X|^2]$
- $C$ *is a universal constant, independent of* $n$, $F_n$

# Berry-Esseen Theorem

**Theorem (Berry-Esseen Theorem (independent case))**

$$\sup_{x\in\mathbf{R}} |Pr[\frac{X_1+...+X_n}{s_n} \leq a] - \Phi(x)| \leq C\frac{r_n}{s_n^3}$$

*where*

- $\Phi$ *is the standard normal cdf*
- $r_n = \mathbb{E}[|(X_1+...+X_n) - \mathbb{E}(X_1+...+X_n)|^3]$
- $s_n^2 = \mathbb{E}[|(X_1+...+X_n) - \mathbb{E}(X_1+...+X_n)|^2]$
- $C$ *is a universal constant, independent of* $n$, $F_n$

Remark:

- Lower bound of $C$: $C \geq \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \approx 0.40973$ (Esseen (1956))
- Upper bound of $C$: $C \leq 0.4785$ (Tyurin (2010))