# Lecture 4

Casella and Berger

Section 5.5

## Convergence Concepts

- convergence in probability
- almost sure (a.s.) convergence
- convergence in probability

**Def 4.1** A sequence of random variables, $X_1, X_2, \ldots$, *converges in probability* to a random variable $X$ if, for every $\epsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| \geq \epsilon) = 0 \text{ or, equivalently, } \lim_{n\to\infty} P(|X_n - X| < \epsilon) = 1.$$

**Theorem 4.1 (Weak Law of Large Numbers)** Let $X_1, X_2, ...$ be i.i.d. random variables with $EX_i = \mu$ and $VarX_i = \sigma^2 < \infty$. Then, for every $\epsilon > 0$,

$$\lim_{n\to\infty} P(|\bar{X}_n - \mu| < \epsilon) = 1;$$

that is $\bar{X}_n \xrightarrow{P} \mu$.

**Theorem 4.2 (Continuous Mapping Theorem)** Suppose that $X_1, X_2, ...$ converges in probability to a random variable $X$ and that $h$ is a continuous function. Then $h(X_1), h(X_2), ...$ converges in probability to $h(X)$.

**Def 4.2** A sequence of random variables, $X_1, X_2, ...$, converges almost surely to a random variable $X$ if, for every $\epsilon > 0$,

$$P(\lim_{n\to\infty} |X_n - X| < \epsilon) = 1.$$

**Theorem 4.3 (Strong Law of Large Numbers)** $X_1, X_2, ...$ be i.i.d. random variables with $EX_i = \mu$ and $VarX_i = \sigma^2 < \infty$. Then, for every $\epsilon > 0$,

$$P(\lim_{n\to\infty} |\bar{X}_n - X| < \epsilon) = 1;$$

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

**Def 4.3** A sequence of random variables, $X_1, X_2, ...$, converges in distribution (or weakly converges) to a random variable $X$ if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at all points $x$ where $F_X(x)$ is continuous.

**Theorem 4.4 (Central Limit Theorem)** Let $X_1, X_2, ...$ be a sequence of i.i.d. random variables whose mgfs exist in a neighborhood of 0. Let $EX_i = \mu$ and $Var X_i = \sigma^2 > 0$. Then $\sqrt{n}(\bar{X}_n - \mu)/\sigma$ has a limiting standard normal distribution, i.e., $\sqrt{n}(\bar{X}_n - \mu)/\sigma \xrightarrow{\mathcal{D}} N(0, 1)$.

Relationships among $\xrightarrow{P}$, $\xrightarrow{a.s.}$, and $\xrightarrow{\mathcal{D}}$

$X_n \xrightarrow{a.s.} X$ implies $X_n \xrightarrow{P} X$ and $X_n \xrightarrow{P} X$ implies $X_n \xrightarrow{\mathcal{D}} X$

**Theorem 4.5 (Slutsky's Theorem)** If $X_n \xrightarrow{\mathcal{D}} X$ and $Y_n \xrightarrow{P} a$, then

a. $Y_n X_n \xrightarrow{\mathcal{D}} aX$.

b. $X_n + Y_n \xrightarrow{\mathcal{D}} X + a$.

**Delta Method** Let $Y_n$ be a sequence of random variables that satisfies $\sqrt{n}(Y_n - \theta) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$. For a given function $g$ and specific value of $\theta$, suppose that $g'(\theta)$ exists and is not 0. Then

$$\sqrt{n}[g(Y_n) - g(\theta)] \xrightarrow{\mathcal{D}} N(0, \sigma^2[g'(\theta)]^2).$$

**Second-order Delta Method** Let $Y_n$ be a sequence of random variables that satisfies $\sqrt{n}(Y_n - \theta) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$. For a given function $g$ and specific value of $\theta$, suppose that $g'(\theta) = 0$ and $g''(\theta)$ exists and is not 0. Then

$$\sqrt{n}[g(Y_n) - g(\theta)] \xrightarrow{\mathcal{D}} \sigma^2 \frac{g''(\theta)}{2} \chi_1^2$$

**Multivariate Delta Method** Let $\mathbf{X}_1, ..., \mathbf{X}_n$ be a random sample with $E(X_{ij}) = \mu_i$ and $Cov(X_{ik}, X_{jk}) = \sigma_{ij}$. For a given function $g$ with continuous first partial derivatives and a specific value of $\boldsymbol{\mu} = (\mu_1, ..., \mu_p)$ for which $\tau^2 = \sum \sum \sigma_{ij} \frac{\partial g(\boldsymbol{\mu})}{\partial \mu_i} \cdot \frac{\partial g(\boldsymbol{\mu})}{\partial \mu_j} > 0$,

$$\sqrt{n}[g(\bar{X}_1, ..., \bar{X}_p) - g(\mu_1, ..., \mu_p)] \xrightarrow{N} (0, \tau^2).$$