# Large Sample Theory

- Convergence
  - Convergence in Probability
  - Convergence in Distribution
- Central Limit Theorems
- Asymptotic Distribution
- Delta Method

# Convergence in Probability

- A sequence of random scalars $\{z_n\} = (z_1, z_2, \ldots)$ **converges in probability** to z (a constant or a random variable) if, for any $\varepsilon>0$, $\lim_{n\to\infty} \text{Prob}(|z_n - z|>\varepsilon) = 0$. z is the **probability limit** of $z_n$ and is written as: $\text{plim}_{n\to\infty} z_n = z$ or $z_n \to_p z$.
- Extension to a sequence of random vectors or matrices: element-by-element convergence in probability, $\mathbf{z}_n \to_p \mathbf{z}$.

# Convergence in Probability

- A special case of convergence in probability is mean square convergence: if $E(z_n) = \mu_n$ and $Var(z_n) = \sigma_n^2$ such that $\mu_n \rightarrow z$ and $\sigma_n^2 \rightarrow 0$, then $z_n$ converges in mean square to z, and $z_n \rightarrow_p z$.
- What is the difference between $E(z_n)$ and plim $z_n$ or $z_n \rightarrow_p z$?
- Mean square convergence is sufficient (not necessary) for convergence in probability.

# Convergence in Probability

- Example: $x \sim (\mu, \sigma^2)$
  - Sample: $\{x_1, x_2, \cdots\}$
  - Sample Mean: $\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$
  - Sequence of Sample Means: $\{\bar{x}_n\}, n \to \infty$

$$E(\bar{x}_n) = \mu \to \mu \qquad \Rightarrow \quad \bar{x}_n \to_p \mu$$

$$Var(\bar{x}_n) = \frac{\sigma^2}{n} \to 0$$

# Almost Sure Convergence

- A sequence of random scalars $\{z_n\} = (z_1, z_2, \ldots)$ **converges almost surely** to z (a constant or a random variable) if, $\text{Prob}(\lim_{n\to\infty} z_n = z) = 1$. Write: $z_n \to_{as} z$.
- If a sequence converges almost surely, then it converges in probability. That is, $z_n \to_{as} z \Rightarrow z_n \to_p z$.
- Extension to a sequence of random vectors or matrices: element-by-element almost sure convergence. In particular, $\mathbf{z}_n \to_{as} \mathbf{z} \Rightarrow \mathbf{z}_n \to_p \mathbf{z}$.

# Laws of Large Numbers

- Let $\bar{z}_n = \frac{1}{n}\sum_{i=1}^{n} z_i$
- LLN concern conditions under which the sequence $\{\bar{z}_n\}$ converges in probability.
- Chebychev's LLN:

$$[\lim_{n\to\infty} E(\bar{z}_n) = \mu, \lim_{n\to\infty} Var(\bar{z}_n) = 0] \Rightarrow \bar{z}_n \to_p \mu$$

# Laws of Large Numbers

- Kolmogorov's LLN: Let $\{z_i\}$ be i.i.d. with $E(z_i) = \mu$ (the variance does not need to be finite). Then

  $$\bar{z}_n \rightarrow_{as} \mu$$

- This implies $\bar{z}_n \rightarrow_{p} \mu$

# Speed of Convergence

- Order of a Sequence
  - 'Little oh' o(.): Sequence $z_n$ is $o(n^{\delta})$ (order less than $n^{\delta}$) if and only if $n^{-\delta}z_n \rightarrow 0$.
    - Example: $z_n = n^{1.4}$ is $o(n^{1.5})$ since $n^{-1.5}\, z_n = 1/n^{.1} \rightarrow 0$.
  - 'Big oh' O(.): Sequence $z_n$ is $O(n^{\delta})$ if and only if $n^{-\delta}z_n \rightarrow$ a finite nonzero constant.
    - Example 1: $z_n = (n^2 + 2n + 1)$ is $O(n^2)$.
    - Example 2: $\Sigma_i x_i^2$ is usually $O(n^1)$ since this is n×the mean of $x_i^2$ and the mean of $x_i^2$ generally converges to $E[x_i^2]$, a finite constant.
- What if the sequence is a random variable? The order is in terms of the variance.
  - Example: What is the order of the sequence $\bar{x}_n$ in random sampling? Because $\text{Var}[\bar{x}_n] = \sigma^2/n$ which is $O(1/n)$

# Convergence in Distribution

- Let $\{z_n\}$ be a sequence of random scalars and $F_n$ be the c.d.f. of $z_n$. $\{z_n\}$ **converges in distribution** to a random scalar z if the c.d.f. $F_n$ of $z_n$ converges to the c.d.f. $F$ of z at every continuity point of $F$. That is, $z_n \to_d z$. $F$ is the **asymptotic** or **limiting distribution** of $z_n$.
- $z_n \to_p z \Rightarrow z_n \to_d z$, or $z_n \to_a F(z)$

# Convergence in Distribution

- The extension to a sequence of random vectors: $\mathbf{z}_n \rightarrow_d \mathbf{z}$ if the joint c.d.f. $F_n$ of the random vector $\mathbf{z}_n$ converges to the joint c.d.f. $F$ of $\mathbf{z}$ at every continuity point of $F$. However, element-by-element convergence does not necessarily mean joint convergence.

# Central Limit Theorems

- CLT concern about the limiting behavior of $\bar{z}_n - \mu$ blown up by $\sqrt{n}$.

  Note: $\mu = E(\bar{z}_n) = E(z_i)$ if $z_i$ is i.i.d.

- Lindeberg-Levy CLT (multivariate):
  Let $\{z_i\}$ be i.i.d. with $E(z_i) = \mu$ and $Var(z_i) = \mathbf{\Sigma}$. Then

$$\sqrt{n}(\bar{z}_n - \mu) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(z_i - \mu) \to_d N(0, \mathbf{\Sigma})$$

# Central Limit Theorems

- Lindeberg-Levy CLT (univariate):
  If $z \sim (\mu,\sigma^2)$, and $\{z_1,z_2,\ldots,z_n\}$ are random sample.

$$Define\ \bar{z}_n = \frac{1}{n}\sum_{i=1}^{n} z_i,\ then$$

$$\sqrt{n}(\bar{z}_n - \mu) \to_d N(0,\sigma^2)$$

# Central Limit Theorems

- Lindeberg-Feller CLT (univariate):
  If $z_i \sim (\mu_i, \sigma_i^2)$, i=1,2,…,n.
  $Let\ \bar{\mu}_n = \frac{1}{n}\sum_{i=1}^{n}\mu_i,\ and\ \bar{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\sigma_i^2 \to \bar{\sigma}^2$
  If no single term dominates this average variance, then

$$\sqrt{n}(\bar{z}_n - \bar{\mu}_n) \to_d N(0, \bar{\sigma}^2)$$

# Asymptotic Distribution

- An asymptotic distribution is a finite sample approximation to the true distribution of a random variable that is good for large samples, but not necessarily for small samples.
- Stabilizing transformation to obtain a limiting distribution: Multiply random variable $x_n$ by some power, a, of n such that the limiting distribution of $n^a x_n$ has a finite, nonzero variance.
  - Example, $\overline{x}_n$ has a limiting variance of zero, since the variance is $\sigma^2/n$. But, the variance of $\sqrt{n}\overline{x}_n$ is $\sigma^2$. However, this does not stabilize the distribution because $E(\sqrt{n}\overline{x}_n) = \sqrt{n}\mu$. The stabilizing transformation would be $\sqrt{n}(\overline{x}_n - \mu)$

# Asymptotic Distribution

- Obtaining an asymptotic distribution from a limiting distribution:
  - Obtain the limiting distribution via a stabilizing transformation.
  - Assume the limiting distribution applies reasonably well in finite samples.
  - Invert the stabilizing transformation to obtain the asymptotic distribution.

# Asymptotic Distribution

- Example: Asymptotic normality of a distribution.

From $\sqrt{n}(\bar{x} - \mu) / \sigma \rightarrow_d N[0,1]$

$$\sqrt{n}(\bar{x} - \mu) \rightarrow_a N[0, \sigma^2]$$

$$(\bar{x} - \mu) \rightarrow_a N[0, \sigma^2 / n]$$

$$\bar{x} \rightarrow_a N[\mu, \sigma^2 / n]$$

Asymptotic distribution.

$\sigma^2 / n =$ asymptotic variance of $\bar{x}$.

# Asymptotic Efficiency

- Comparison of asymptotic variances
- How to compare consistent estimators? If both converge to constants, both variances go to zero.
- Example: Random sampling from the normal distribution,
  - Sample mean is asymptotically $N[\mu,\sigma^2/n]$
  - Median is asymptotically $N[\mu,(\pi/2)\sigma^2/n]$
  - Mean is asymptotically more efficient.

# Convergence: Useful Results

- Multivariate Convergence in Distribution
  Let $\{\mathbf{z}_n\}$ be a sequence of K-dimensional random vectors. Then:
  $\mathbf{z}_n \rightarrow_d \mathbf{z} \Leftrightarrow \boldsymbol{\lambda}'\mathbf{z}_n \rightarrow_d \boldsymbol{\lambda}'\mathbf{z}$
  for any K-dimensional vector of $\boldsymbol{\lambda}$ real numbers.

# Convergence: Useful Results

- Slutsky Theorem: Suppose a(.) is a scalar- or vector-valued continuous function that does not depend on n:
  $z_n \rightarrow_p \alpha \Rightarrow a(z_n) \rightarrow_p a(\alpha)$
  $\mathbf{z}_n \rightarrow_d \mathbf{z} \Rightarrow a(\mathbf{z}_n) \rightarrow_d a(\mathbf{z})$
- $\mathbf{x}_n \rightarrow_d \mathbf{x}, \mathbf{y}_n \rightarrow_p \alpha \Rightarrow \mathbf{x}_n + \mathbf{y}_n \rightarrow_d \mathbf{x} + \alpha$
  $\mathbf{x}_n \rightarrow_d \mathbf{x}, \mathbf{y}_n \rightarrow_p 0 \Rightarrow \mathbf{y}_n' \mathbf{x}_n \rightarrow_p 0$

# Convergence: Useful Results

- Slutsky results for matrices:
  $\mathbf{A}_n \rightarrow_p \mathbf{A}$ (plim $\mathbf{A}_n = \mathbf{A}$),
  $\mathbf{B}_n \rightarrow_p \mathbf{B}$ (plim $\mathbf{B}_n = \mathbf{B}$),
  (element by element)
  $\Rightarrow$
  plim $(\mathbf{A}_n^{-1}) = [\text{plim } \mathbf{A}_n]^{-1} = \mathbf{A}^{-1}$
  plim $(\mathbf{A}_n\mathbf{B}_n) = (\text{plim}\mathbf{A}_n)(\text{plim}\mathbf{B}_n) = \mathbf{AB}$

# Convergence: Useful Results

- $\mathbf{x}_n \to_d \mathbf{x}$, $\mathbf{A}_n \to_p \mathbf{A} \Rightarrow \mathbf{A}_n\mathbf{x}_n \to_d \mathbf{A}\mathbf{x}$
  In particular, if $\mathbf{x} \sim N(\mathbf{0},\mathbf{\Sigma})$, then
  $\mathbf{A}_n\mathbf{x}_n \to_d N(\mathbf{0},\mathbf{A\Sigma A}')$
- $\mathbf{x}_n \to_d \mathbf{x}$, $\mathbf{A}_n \to_p \mathbf{A} \Rightarrow \mathbf{x}_n'\mathbf{A}_n^{-1}\mathbf{x}_n \to_d \mathbf{x}'\mathbf{A}^{-1}\mathbf{x}$

# Delta Method

- Suppose $\{\mathbf{x}_n\}$ is a sequence of K-dimensional random vector such that $\mathbf{x}_n \rightarrow \boldsymbol{\beta}$ and $\sqrt{n}(\mathbf{x}_n - \boldsymbol{\beta}) \rightarrow_d \mathbf{z}$. Suppose $\mathbf{a}(.)$: $R^K \rightarrow R^r$ has continuous first derivatives with $\mathbf{A}(\boldsymbol{\beta})$ defined by $$\mathbf{A}(\boldsymbol{\beta}) = \frac{\partial \mathbf{a}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}'}$$

  Then $\sqrt{n}[\mathbf{a}(\mathbf{x}_n) - \mathbf{a}(\boldsymbol{\beta})] \rightarrow_d \mathbf{A}(\boldsymbol{\beta})\mathbf{z}$

# Delta Method

- $\sqrt{n}(\mathbf{x}_n - \boldsymbol{\beta}) \to_d N(\mathbf{0}, \boldsymbol{\Sigma})$
  $\Rightarrow$
  $\sqrt{n}[\mathbf{a}(\mathbf{x}_n) - \mathbf{a}(\boldsymbol{\beta})] \to_d N(\mathbf{0}, \mathbf{A}(\boldsymbol{\beta})\boldsymbol{\Sigma}\mathbf{A}(\boldsymbol{\beta})')$

# Delta Method

- Example

$$\bar{x}_n \xrightarrow{a} N[\mu, \sigma^2 / n]$$

What is the asymptotic distribution of

$f(\bar{x}_n)=\exp(\bar{x}_n)$ or $f(\bar{x}_n)=1/\bar{x}_n$

(1) Normal since $\bar{x}_n$ is asymptotically normally distributed

(2) Asymptotic mean is $f(\mu)=\exp(\mu)$ or $1/\mu$.

(3) For the variance, we need $f'(\mu) = \exp(\mu)$ or $-1/\mu^2$

$\text{Asy.Var}[f(\bar{x}_n)] = [\exp(\mu)]^2\sigma^2 / n$ or $[1/\mu^4]\sigma^2 / n$