Statistics

Asymptotic Theory

Shiu-Sheng Chen

Department of Economics
National Taiwan University

Fall 2019

Asymptotic Theory: Motivation

- Asymptotic theory (or large sample theory) aims at answering the question: what happens as we gather more and more data?
- In particular, given random sample, $\{X_1, X_2, X_3, \ldots, X_n\}$, and statistic:

$$T_n = t(X_1, X_2, \ldots, X_n),$$

what is the limiting behavior of T_n as $n \longrightarrow \infty$?

Asymptotic Theory: Motivation

- Why asking such a question?
- For instance, given random sample $\{X_i\}_{i=1}^n \sim^{i.i.d.} N(\mu, \sigma^2)$, we know that
$$\bar{X}_n \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$
- However, if $\{X_i\}_{i=1}^n \sim^{i.i.d.} (\mu, \sigma^2)$ without normal assumption, what is the distribution of $\bar{X}_n$?
  - We don't know, indeed.
- Is it possible to find a good approximation of the distribution of $\bar{X}_n$ as $n \longrightarrow \infty$?
  - Yes! This is where the asymptotic theory kicks in.

Section 1

Preliminary Knowledge

Preliminary Knowledge

- Limit
- Markov Inequality
- Chebyshev Inequality

Limit of a Real Sequence

Definition (Limit)

If for every $\varepsilon > 0$, and an integer $N(\varepsilon)$,

$$|b_n - b| < \varepsilon, \quad \forall\ n > N(\varepsilon)$$

then we say that a sequence of real numbers $\{b_1, \ldots, b_n\}$ converges to a limit b.

- It is denoted by

$$\lim_{n \to \infty} b_n = b$$

Markov Inequality

Theorem (Markov Inequality)

Suppose that X *is a random variable such that* $P(X \geq 0) = 1$*. Then for every real number* $m > 0$,

$$P(X \geq m) \leq \frac{E(X)}{m}$$

Chebyshev Inequality

Theorem (Chebyshev Inequality)

Let $Y \sim (E(Y), Var(Y))$. Then for every number $\varepsilon > 0$,

$$P\Big(\big|Y - E(Y)\big| \geq \varepsilon\Big) \leq \frac{Var(Y)}{\varepsilon^2}$$

- Proof: Let $X = [Y - E(Y)]^2$, then

  $$P(X \geq 0) = 1$$

  and

  $$E(X) = Var(Y)$$

  Hence, the result can be derived by applying the Markov Inequality.

Section 2

Modes of Convergence

Types of Convergence

For a random variable, we consider three modes of convergence:

- Converge in Probability
- Converge in Distribution
- Converge in Mean Square

Converge in Probability

Definition (Converge in Probability)

Let $\{Y_n\}$ be a sequence of random variables and let Y be another random variable. For any $\varepsilon > 0$,

$$P(|Y_n - Y| < \varepsilon) \longrightarrow 1, \text{ as } n \longrightarrow \infty$$

then we say that Y_n converges in probability to Y, and denote it by

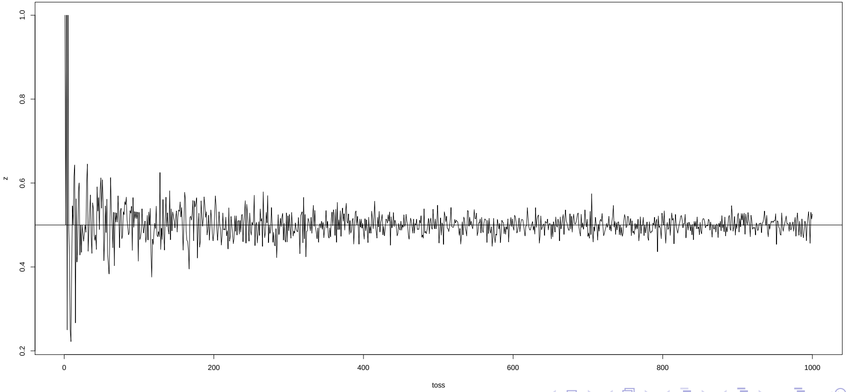
$$Y_n \xrightarrow{p} Y$$

- Equivalently,

$$P(|Y_n - Y| \geq \varepsilon) \longrightarrow 0, \text{ as } n \longrightarrow \infty$$

Converge in Probability

- $\{X_i\}_{i=1}^n \sim^{i.i.d.}$ Bernoulli(0.5) and then compute $Y_n = \bar{X}_n = \frac{\sum_i X_i}{n}$
- In this case, $Y_n \xrightarrow{p} 0.5$



Converge in Distribution

Definition (Converge in Distribution)

Let $\{Y_n\}$ be a sequence of random variables with distribution function $F_{Y_n}(y)$, (denoted by $F_n(y)$ for simplicity). Let Y be another random variable with distribution function, $F_Y(y)$. If

$$\lim_{n\to\infty} F_n(y) = F_Y(y) \text{ at all } y \text{ for which } F_Y(y) \text{ is continuous}$$

then we say that Y_n converges in distribution to Y.

- It is denoted by

$$Y_n \xrightarrow{d} Y$$

- $F_Y(y)$ is called the limiting distribution of Y_n.

Converge in Mean Square

Definition (Converge in Mean Square)

Let $\{Y_n\}$ be a sequence of random variables and let Y be another random variable. If

$$E(Y_n - Y)^2 \longrightarrow 0, \quad \text{as } n \longrightarrow \infty.$$

Then we say that Y_n converges in mean square to Y.

- It is denoted by

$$Y_n \xrightarrow{ms} Y$$

- It is also called converge in quadratic mean.

Section 3

Important Theorems

Theorems

Theorem

$Y_n \xrightarrow{ms} c$ *if and only if*

$$\lim_{n\to\infty} E(Y_n) = c, \quad \textit{and} \quad \lim_{n\to\infty} Var(Y_n) = 0.$$

- Proof. It can be shown that

$$E(Y_n - c)^2 = E([Y_n - E(Y_n)]^2) + [E(Y_n) - c]^2$$

Theorems

Theorem

If $Y_n \xrightarrow{ms} Y$ *then* $Y_n \xrightarrow{p} Y$

- Proof: Note that $P(|Y_n - Y|^2 \geq 0) = 1$, and by Markov Inequality,

$$P(|Y_n - Y| \geq k) = P(|Y_n - Y|^2 \geq k^2) \leq \frac{E(|Y_n - Y|^2)}{k^2}$$

Weak Law of Large Numbers, WLLN

Theorem (WLLN)

Given a random sample $\{X_i\}_{i=1}^n$ *with* $\sigma^2 = Var(X_1) < \infty$. *Let* $\bar{X}_n$ *denote the sample mean, and note that* $E(\bar{X}_n) = E(X_1) = \mu$. *Then*

$$\bar{X}_n \xrightarrow{p} \mu$$

- Proof: (1) By Chebyshev Inequality (2) By Converge in Mean Square
- Sample mean $\bar{X}_n$ is getting closer (in probability sense) to the population mean μ as the sample size increases.
- That is, if we use $\bar{X}_n$ as a guess of unknown μ, we are quite happy that the sample mean makes a good guess.

WLLN for Other Moments

- Note that the WLLN can be thought as

$$\frac{\sum_{i=1}^{n} X_i}{n} = \frac{X_1 + X_2 + \cdots X_n}{n} \xrightarrow{p} E(X_1)$$

- Let $Y = X^2$, and by the WLLN,

$$\frac{\sum_{i=1}^{n} Y_i}{n} = \frac{Y_1 + Y_2 + \cdots Y_n}{n} \xrightarrow{p} E(Y_1)$$

- Hence,

$$\frac{\sum_{i=1}^{n} X_i^2}{n} = \frac{X_1^2 + X_2^2 + \cdots X_n^2}{n} \xrightarrow{p} E(X_1^2)$$

Example: An Application of WLLN

- Assume $W_n \sim \text{Binomial}(n, \mu)$, and let $Y_n = \frac{W_n}{n}$. Then

$$Y_n \xrightarrow{p} \mu$$

- Why?
- Since $W_n = \sum_i X_i$, $X_i \sim^{i.i.d.} \text{Bernoulli}(\mu)$ with $E(X_1) = \mu$, $Var(X_1) = \mu(1 - \mu)$, the result follows by WLLN.

Central Limit Theorem, CLT

Theorem (CLT)

Let $\{X_i\}_{i=1}^n$ *be a random sample, where* $E(X_1) = \mu < \infty$, $Var(X_1) = \sigma^2 < \infty$, *then*

$$Z_n = \frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{Var(\bar{X}_n)}} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} N(0,1)$$

- If a random sample is taken from any distribution with mean μ and variance σ^2, regardless of whether this distribution is discrete or continuous, then the distribution of the random variable Z_n will be approximately the standard normal distribution in large sample.

CLT

- Using notation of asymptotic distribution,

$$\frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \sim^A N(0,1),$$

Or

$$\bar{X}_n \sim^A N\left(\mu, \frac{\sigma^2}{n}\right),$$

- where $\sim^A$ represents asymptotic distribution, and A represents Asymptotically

An Application of CLT

- Example: Assume $\{X_i\} \sim^{\text{i.i.d.}}$ Bernoulli(μ), then

$$\frac{\bar{X}_n - \mu}{\sqrt{\frac{\mu(1-\mu)}{n}}} \xrightarrow{d} N(0,1).$$

- Why?
- Since $E(\bar{X}_n) = \mu$, and $Var(\bar{X}_n) = \frac{\sigma^2}{n} = \frac{\mu(1-\mu)}{n}$

Continuous Mapping Theorem

Theorem (CMT)

Given $Y_n \xrightarrow{p} Y$, *and* $g(\cdot)$ *is continuous, then*

$$g(Y_n) \xrightarrow{p} g(Y).$$

- Proof: omitted here.
- Examples: if $Y_n \xrightarrow{p} Y$, then
  - $\frac{1}{Y_n} \xrightarrow{p} \frac{1}{Y}$
  - $Y_n^2 \xrightarrow{p} Y^2$
  - $\sqrt{Y_n} \xrightarrow{p} \sqrt{Y}$

Theorem

Theorem

Given $W_n \xrightarrow{p} W$ *and* $Y_n \xrightarrow{p} Y$, *then*

- $W_n + Y_n \xrightarrow{p} W + Y$
- $W_n Y_n \xrightarrow{p} WY$

- Proof: omitted here.

Slutsky Theorem

Theorem

Given $W_n \xrightarrow{d} W$ *and* $Y_n \xrightarrow{p} c$, *where* c *is a constant. Then*

- $W_n + Y_n \xrightarrow{d} W + c$
- $W_n Y_n \xrightarrow{d} cW$
- $\frac{W_n}{Y_n} \xrightarrow{d} \frac{W}{c}$ *for* $c \neq 0$

- Proof: omitted here.

The Delta Method

Theorem

Given $\sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$. *Let* $g(\cdot)$ *be differentiable, and* $g'(\theta) \neq 0$ *exists, then*

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

- Proof: (sketch) Given 1st-order Taylor approximation

$$g(Y_n) \approx g(\theta) + g'(\theta)(Y_n - \theta),$$

then

$$\frac{\sqrt{n}(g(Y_n) - g(\theta))}{g'(\theta)} \approx \sqrt{n}(Y_n - \theta) \xrightarrow{d} N(0, \sigma^2)$$

Example

- Given $\{X_i\}_{i=1}^n \sim^{i.i.d.} (\mu, \sigma^2)$, find the asymptotic distribution of $\frac{\bar{X}_n}{1-\bar{X}_n}$.
  - Note that by CLT,

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} N(0, \sigma^2)$$

  - Hence, by the Delta method,

$$g(\bar{X}_n) = \frac{\bar{X}_n}{1-\bar{X}_n}, \qquad g(\mu) = \frac{\mu}{1-\mu}, \qquad g'(\mu) = \frac{1}{(1-\mu)^2}$$

$$\sqrt{n}\left(\frac{\bar{X}_n}{1-\bar{X}_n} - \frac{\mu}{1-\mu}\right) \xrightarrow{d} N\left(0, \frac{1}{(1-\mu)^4}\sigma^2\right)$$