Econometrics: Lecture 5

Hang Miao

Rutgers University

January 21, 2021

Overview

1. Types of Convergence
2. Law of Large Number
3. Central Limit Theorem
4. Hypothesis Testing for Population Mean
5. Confidence intervals

Types of Convergence

Types of Convergence

Convergence of Random Variable

Throughout the following, we assume that $\{X_n\}$ is a sequence of random variables, and X is a random variable, and all of them are defined on the same probability space $(\Omega, \mathcal{F}, \Pr)$

Definition (Converge Almost Surely)

$\{X_n\}$ converge almost surely to X if $\Pr(\lim_{n\to\infty} X_n = X) = 1$,
or $\Pr\left(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\right) = 0$
Denoted as $X_n \xrightarrow{\text{a.s.}} X$

Definition (Converge in Probability)

$\{X_n\}$ converge in probability to X if $\lim_{n\to\infty} \Pr\left(|X_n - X| \geq \varepsilon\right) = 0$.
Denoted as $X_n \xrightarrow{P} X$, or $\underset{n\to\infty}{\text{plim}}\, X_n = X$.

Types of Convergence

Definition (Converge in Distribution)

$\{X_n\}$ converges in distribution, or converge weakly, or converge in law to X if

$$\lim_{n\to\infty} F_n(x) = F(x), \quad \forall x \in \mathbb{R}$$

Denoted as $X_n \xrightarrow{d} X$. where $F_n(\cdot)$ and $F(\cdot)$ are the cumulative distribution functions (CDF) of random variables $\{X_n\}$ and X, respectively.

Definition (Converge in Mean)

Given a real number $r \geq 1$, $\{X_n\}$ converges in the **r-th mean (or in the L^r-norm)** to X, if the r-th absolute moments $\mathrm{E}(| X_n |^r)$ and $\mathrm{E}(| X |^r)$ of $\{X_n\}$ and X exist, and

$$\lim_{n\to\infty} \mathrm{E}\left(|X_n - X|^r\right) = 0,$$

Denoted as $X_n \xrightarrow{L^r} X$.

Relationship Between Different Types of Convergence

- Almost sure convergence implies convergence in probability:

$$X_n \xrightarrow{\text{a.s.}} X \quad \Rightarrow \quad X_n \xrightarrow{p} X$$

- Convergence in probability implies there exists a sub-sequence k_n s.t X_{k_n} converges almost surely to X:

$$X_n \xrightarrow{p} X \quad \Rightarrow \quad X_{k_n} \xrightarrow{\text{a.s.}} X$$

- Convergence in probability implies convergence in distribution:

$$X_n \xrightarrow{p} X \quad \Rightarrow \quad X_n \xrightarrow{d} X$$

- Convergence in r-th order mean implies convergence in probability:

$$X_n \xrightarrow{L^r} X \quad \Rightarrow \quad X_n \xrightarrow{p} X$$

Example

Converge Almost Surely: Example 1

Consider an animal of some short-lived species. We record the amount of food that this animal consumes per day. This sequence of numbers will be unpredictable, but we may be quite certain that one day the number will become zero, and will stay zero forever after.

Converge Almost Surely: Example 2

Consider a man who tosses seven coins every morning. Each afternoon, he donates one pound to a charity for each head that appeared. The first time the result is all tails, however, he will stop permanently.

Let $X_1, X_2, \cdots$ be the daily amounts the charity received from him.

We may be almost sure that one day this amount will be zero, and stay zero forever after that.

However, when we consider any finite number of days, there is a nonzero probability the terminating condition will not occur.

Example

Converge in Probability: Example 1 (weak LLN)

Sample mean converge to the population mean:

$$\hat{\mu}_Y \xrightarrow{p} \mu_Y$$

Converge in Probability but not Almost Surely: Example 2

$$\begin{aligned} X_n &= \begin{cases} 1 & \text{with probability } \frac{1}{n} \\ 0 & \text{with probability } 1-\frac{1}{n} \end{cases} \\ X &= 0 \end{aligned}$$

- $X_n \xrightarrow{p} X$. Indicate 1's will get rarer and rarer as one looks ahead in the sequence.
- Almost surely says something about the entire tail of the sequence. That is $\exists N$ such that when $n \geq N$, 1's will completely go extinct (with probability 1). In this case, $X_n \not\to$ a.s. X

Example

Converge in Distribution: Example 1

Suppose a new dice factory has just been built. The first few dice come out quite biased, due to imperfections in the production process. The outcome from tossing any of them will follow a distribution markedly different from the desired uniform distribution.

As the factory is improved, the dice become less and less loaded, and the outcomes from tossing a newly produced die will follow the uniform distribution more and more closely.

Converge in Distribution: Example 2

Suppose Xi is an iid sequence. Then the (normalized) sums $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$ converge in distribution to Normal distribution. The result follows from Central Limit Theorem (CLT)

Example

Converge in Distribution: Example 3

$$X_n = \begin{cases} 1 & \text{with probability } \frac{1}{2} + \frac{1}{n} \\ 0 & \text{with probability } \frac{1}{2} - \frac{1}{n} \end{cases}$$

$$X = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ 0 & \text{with probability } \frac{1}{2} \end{cases}$$

- $X_n \xrightarrow{d} X$. Indicate the probability distribution of X_n gets closer and closer to that of X as n becomes large.
- Although the distribution of X_n and X become same, but the realized value for X_n and X will not be same for most cases, thus $X_n \not\to$ p X

Example

Converge Almost Surely but not In Mean

$$
\begin{aligned}
X_n &= \begin{cases} 3X_{n-1} & \text{with probability } \frac{1}{2} \\ 0 & \text{with probability } \frac{1}{2} \end{cases}, \quad X_0 = 1,\ n = 1, \cdots, \infty \\
X &= 0
\end{aligned}
$$

- Imagine someone start with \$1 and each day you play a game in which you either lose all your money or triple it with equal probability. X_n be a random variable that represents the amount of money you have after n days.
- We can easily verify that ultimately the person will end up with 0 almost surly with probability 1.
- While the sequence does not converge in mean to 0 (nor to any other constant)

$$\lim_{n\to\infty} \mathrm{E}_n\left(|X_n - 0|\right) = \lim_{n\to\infty} \mathrm{E}_n\left(X_n\right) = \lim_{n\to\infty} \mathrm{E}_{n-1}\left(\frac{3}{2}X_{n-1}\right) = \lim_{n\to\infty}\left(\frac{3}{2}\right)^n$$

Digression: Law of Iterative Expectation

Definition (Law of Total Expectation)

The Law of Total Expectation is also called as the law of iterated expectations, the tower rule, Adam's law, and the smoothing theorem. States that if X is a random variable whose expected value $\mathrm{E}(X)$ is defined, and Y is any random variable on the same probability space, then

$$\mathrm{E}(X) = \mathsf{E}_Y\left[\mathsf{E}_X\left[X \mid Y\right]\right]$$

Example in the Previous Slides

$$\begin{aligned}
\mathrm{E}[X_n] &= \mathrm{E}_{n-1}\Big[\mathrm{E}_n\left[X_n \mid X_{n-1}\right]\Big] \\
&= \mathrm{E}_{n-1}\left[3X_{n-1} \cdot 1/2 + 0 \cdot 1/2\right] \\
&= \frac{3}{2}\mathrm{E}_{n-1}\left[X_{n-1}\right] = \frac{3}{2}\frac{3}{2}\mathrm{E}_{n-2}\left[X_{n-2}\right] = \cdots = \left(\frac{3}{2}\right)^n \mathrm{E}_0\left[X_0\right]
\end{aligned}$$

Law of Large Number

Law of Large Number

Assumption

Suppose $\{X_n\}$ is a sequence of random variables. The sample mean of $\{X_n\}$ is defined as follow

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

Definition (Weak Law of Large Numbers (Chebychev's LLN))

$$\mathrm{E}\left[X_i\right] = \mu,\ \lim_{n\to\infty} \mathrm{Var}\left(\overline{X}_n\right) = 0 \implies \overline{X}_n \xrightarrow{p} \mu$$

Definition (Strong Law of Large Numbers (Kolmogorov's LLN))

$$\mathrm{E}\left[X_i\right] = \mu,\ \{X_n\}\text{i.i.d.} \implies \overline{X}_n \xrightarrow{a.s} \mu$$

Central Limit Theorem

Central Limit Theorem

Assumption

The central limit theorem (CLT) establishes that, in some situations, when independent random variables are added, their properly normalized sum tends toward a normal distribution (informally a "bell curve") even if the original variables themselves are not normally distributed.

Definition (Lindeberg–Levy CLT)

Suppose $\{X_n\}$ is a sequence of i.i.d. random variables with $\mathrm{E}[X_i] = \mu$ and $\mathrm{Var}[X_i] = \sigma^2 < \infty$. Then as $n \to \infty$,

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

Or equivalently,

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

Hypothesis Testing for Population Mean

Hypothesis Testing for Population Mean with Known Variance σ^2

Motivation

Equipped with Central Limit Theorem we could make the statistical inference about the population mean from sample mean by Hypothesis Testing introduced before.

Prerequisite Setting

Suppose there is a sequence of i.i.d. observed sample $\{X_n\}$ from population. The variance of the underlying population is known as σ^2. The population mean μ is unknown.

Hypothesis Testing

We could propose certain value μ_0 for μ as null hypothesis $H_0 : \mu = \mu_0$ then the alternative hypothesis should be $H_1 : \mu \neq \mu_0$.
The object is to see how credible H_0 is and how likely we make a type I mistake if we reject the H_0 while in favor of the alternative H_1

Hypothesis Testing for Population Mean with Known Variance σ^2

Statistics

According to the central limit theorem (CLT), the Z-score statistic $Z_n = \frac{\overline{X}_n - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ converge in distribution to standard normal $N(0, 1)$. If the sample size is large enough, Z_n approximately follows standard normal distribution.

Prerequisite Setting

Suppose we have a sequence of i.i.d. observed sample $\{X_n\}$ from population. The variance of the underlying population is known as σ^2. The population mean μ is unknown.

Hypothesis Testing

We could propose certain value μ_0 for μ as null hypothesis $H_0 : \mu = \mu_0$ then the alternative hypothesis should be $H_1 : \mu \neq \mu_0$.

We would like to see how credible H_0 is and how likely we make a type I

Confidence intervals

Confidence Intervals

Definition (Confidence intervals)

Confidence Interval (CI) is a type of interval estimate, computed from the statistics of the observed data, that might contain the true value of an unknown population parameter.

Confidence Intervals for the Population Mean

$$99\% \text{ confidence interval for } \mu_X = \left[\overline{X}_n \pm 2.58 \times \frac{\sigma}{\sqrt{n}}\right]$$

$$95\% \text{ confidence interval for } \mu_X = \left[\overline{X}_n \pm 1.96 \times \frac{\sigma}{\sqrt{n}}\right]$$

$$90\% \text{ confidence interval for } \mu_X = \left[\overline{X}_n \pm 1.64 \times \frac{\sigma}{\sqrt{n}}\right]$$