



UNIVERSITY OF
TORONTO

Statistical Sciences

DoSS Summer Bootcamp Probability Module 8

Ichiro Hashimoto

University of Toronto

July 22, 2026

Recap

Learnt in last module:

- Stochastic convergence
 - ▷ Convergence in distribution
 - ▷ Convergence in probability
 - ▷ Convergence almost surely
 - ▷ Convergence in L^p
 - ▷ Relationship between convergences

Outline

- Convergence of functions of random variables
 - ▷ Slutsky's theorem
 - ▷ Continuous mapping theorem
- Laws of large numbers
 - ▷ WLLN
 - ▷ SLLN
 - ▷ Glivenko-Cantelli theorem
- Central limit theorem

Convergence of functions of random variables

Recall: Stochastic convergence If $X_n \rightarrow X$, $Y_n \rightarrow Y$ in some sense, how is the limiting property of $f(X_n, Y_n)$?

Convergence of functions of random variables

Recall: Stochastic convergence If $X_n \rightarrow X$, $Y_n \rightarrow Y$ in some sense, how is the limiting property of $f(X_n, Y_n)$?

Convergence of functions of random variables (a.s.)

Suppose the probability space is complete, if $X_n \xrightarrow{a.s.} X$, $Y_n \xrightarrow{a.s.} Y$, then for any real numbers a, b ,

- $aX_n + bY_n \xrightarrow{a.s.} aX + bY$;
- $X_n Y_n \xrightarrow{a.s.} XY$.

Remark:

- Still require all the random variables to be defined on the same probability space

Convergence of functions of random variables

Proof of $X_n + Y_n \xrightarrow{a.s.} X + Y$:

Convergence of functions of random variables

Convergence of functions of random variables (probability)

Suppose the probability space is complete, if $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, then for any real numbers a, b ,

- $aX_n + bY_n \xrightarrow{P} aX + bY$;
- $X_n Y_n \xrightarrow{P} XY$.

Remark:

- Still require all the random variables to be defined on the same probability space

Convergence of functions of random variables

Proof of $X_n + Y_n \xrightarrow{P} X + Y$:

Convergence of functions of random variables

Convergence of functions of random variables (L^p)

Suppose the probability space is complete, if $X_n \xrightarrow{L^p} X$, $Y_n \xrightarrow{L^p} Y$, then for any real numbers a, b ,

- $aX_n + bY_n \xrightarrow{L^p} aX + bY$;

Remark:

- Still require all the random variables to be defined on the same probability space

Convergence of functions of random variables

Proof of $X_n + Y_n \xrightarrow{L^p} X + Y$:

Convergence of functions of random variables

Remark: Convergence in distribution is different.

Slutsky's theorem

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} c$ (c is a constant), then

- $X_n + Y_n \xrightarrow{d} X + c$;
- $X_n Y_n \xrightarrow{d} cX$;
- $X_n / Y_n \xrightarrow{d} X/c$, where $c \neq 0$.

Convergence of functions of random variables

Remark: Convergence in distribution is different.

Slutsky's theorem

If $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{P} c$ (c is a constant), then

- $X_n + Y_n \xrightarrow{d} X + c$;
- $X_n Y_n \xrightarrow{d} cX$;
- $X_n / Y_n \xrightarrow{d} X / c$, where $c \neq 0$.

Remark:

- The theorem remains valid if we replace all the convergence in distribution with convergence in probability.

Convergence of functions of random variables

Remark: The requirement that $Y_n \xrightarrow{P} c$ (c is a constant) is necessary.

Convergence of functions of random variables

Remark: The requirement that $Y_n \xrightarrow{P} c$ (c is a constant) is necessary.

Examples:

$X_n \sim \mathcal{N}(0, 1)$, $Y_n = -X_n$, then

- $X_n \xrightarrow{d} Z \sim \mathcal{N}(0, 1)$, $Y_n \xrightarrow{d} Z \sim \mathcal{N}(0, 1)$;
- $X_n + Y_n \xrightarrow{d} 0$;
- $X_n Y_n = -X_n^2 \xrightarrow{d} -\chi^2(1)$;
- $X_n / Y_n = -1$.

Convergence of functions of random variables

Continuous mapping theorem

Let X_n, X be random variables, if $g(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ satisfies $\mathbb{P}(X \in D_g) = 0$, then

- $X_n \xrightarrow{a.s.} X \Rightarrow g(X_n) \xrightarrow{a.s.} g(X) ;$
- $X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X) ;$
- $X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X) ;$

where D_g is the set of discontinuity points of $g(\cdot)$.

Convergence of functions of random variables

Continuous mapping theorem

Let X_n, X be random variables, if $g(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$ satisfies $\mathbb{P}(X \in D_g) = 0$, then

- $X_n \xrightarrow{a.s.} X \Rightarrow g(X_n) \xrightarrow{a.s.} g(X) ;$
- $X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X) ;$
- $X_n \xrightarrow{d} X \Rightarrow g(X_n) \xrightarrow{d} g(X) ;$

where D_g is the set of discontinuity points of $g(\cdot)$.

Remark:

- If $g(\cdot)$ is continuous, then ...
- If X is a continuous random variable, and D_g only include countably many points, then ...

Law of large numbers

Weak Law of Large Numbers (WLLN)

If $X_1, X_2, \dots, X_n$ are i.i.d. random variables, $\mu = \mathbb{E}(X_i)$, and $\mathbb{E}|X_i| < \infty$, then

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{P} \mu.$$

Remark:

A more easy-to-prove version is the L^2 weak law, where an additional assumption $\text{Var}(X_i) < \infty$ is required.

Law of large numbers

A generalization of the theorem: triangular array

Triangular array

A triangular array of random variables is a collection $\{X_{n,k}\}_{1 \leq k \leq n}$.

$$X_{1,1}$$

$$X_{2,1}, X_{2,2}$$

$$X_{3,1}, X_{3,2}, X_{3,3}$$

$$\vdots$$

$$X_{n,1}, X_{n,2}, \dots, X_{n,n}$$

Remark: We can consider the limiting property of the row sum S_n .

Law of Large Numbers

L^2 weak law for triangular array

Suppose $\{X_{n,k}\}$ is a triangular array, $n = 1, 2, \dots$, $k = 1, 2, \dots, n$. Let $S_n = \sum_{k=1}^n X_{n,k}$, $\mu_n = \mathbb{E}(S_n)$, if $\sigma_n^2/b_n^2 \rightarrow 0$, where $\sigma_n^2 = \text{Var}(S_n)$ and b_n is a sequence of positive real numbers, then

$$\frac{S_n - \mu_n}{b_n} \xrightarrow{P} 0.$$

Remark:

The L^2 weak law for i.i.d. random variables is a special case of that for triangular array.

Law of large numbers

Proof:

Law of large numbers

Proof:

Remark:

A more generalized version incorporates truncation, then the second-moment constraint is relieved.

Law of large numbers

Strong Law of Large Numbers (SLLN)

Let $X_1, X_2, \dots$ be an i.i.d. sequence satisfying $\mathbb{E}(X_i) = \mu$ and $\mathbb{E}(|X_i|) < \infty$, then

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{\text{a.s.}} \mu.$$

Remark: The proof needs Borel-Cantelli lemma.

Law of large numbers

Strong Law of Large Numbers (SLLN)

Let $X_1, X_2, \dots$ be an i.i.d. sequence satisfying $\mathbb{E}(X_i) = \mu$ and $\mathbb{E}(|X_i|) < \infty$, then

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} \xrightarrow{a.s.} \mu.$$

Remark: The proof needs Borel-Cantelli lemma.

Glivenko-Cantelli theorem

Let $X_i, i = 1, \dots, n$ i.i.d. with distribution function $F(\cdot)$, and let $F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x)$, then as $n \rightarrow \infty$,

$$\sup_{x \in \mathbb{R}} |F(x) - F_n(x)| \rightarrow 0, \quad a.s.$$

Law of large numbers

Proof of Glivenko-Cantelli theorem:

Limit Theorems and Counterexamples

Recall: For the law of large numbers to hold, the assumption $E|X| < \infty$ is crucial.

Law of Large Numbers fail for infinite mean i.i.d. random variables

If $X_1, X_2, \dots$ are i.i.d. to X with $E|X_i| = \infty$, then for $S_n = X_1 + \dots + X_n$,
 $P(\lim_{n \rightarrow \infty} S_n/n \in (-\infty, \infty)) = 0$.

Proof: Omitted

Central Limit Theorem

What is the limiting distribution of the sample mean?

Classic CLT

Suppose $X_1, \dots, X_n$ is a sequence of i.i.d. random variables with $\mathbb{E}(X_i) = \mu$, $\text{Var}(X_i) = \sigma^2 < \infty$, then

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

Remark:

- The proof involves characteristic function.
- A more generalized CLT is referred to as “Lindeberg CLT”.

Central Limit Theorem

Example:

Suppose $X_i \sim \text{Bernoulli}(p)$, i.i.d., consider $Z_n = \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}}$, then by CLT, $Z_n \sim \mathcal{N}(0, 1)$ asymptotically.

Monotone Convergence Theorem

Monotone Convergence Theorem

If $X_n \geq c$ and $X_n \nearrow X$, then $EX_n \nearrow EX$

Usage:

Dominated Convergence Theorem

Dominated Convergence Theorem

If $X_n \rightarrow X$ a.s. and $|X_n| \leq Y$ a.s. for all n and Y is integrable, then $EX_n \rightarrow EX$

Usage:

Delta Method

More about CLT: Delta method

Suppose X_n are i.i.d. random variables with $EX_n = 0$, $VAR(X_n) = \sigma^2 > 0$. Let g be a measurable function that is differentiable at 0 with $g'(0) \neq 0$. Then

$$\sqrt{n} \left(g \left(\frac{\sum_{k=1}^n X_k}{n} \right) - g(0) \right) \rightarrow N(0, \sigma^2 g'(0)^2) \quad \text{weakly.}$$

Proof under stronger assumption: Here, we suppose g is continuously differentiable on $\mathbb{R}$. If you are interested in a general proof refer to Robert Keener's *Theoretical Statistics*.

Delta Method

Proof Cont'd:

Problem Set

Problem 1: Prove that on a complete probability space, if $X_n \xrightarrow{a.s.} X$, $Y_n \xrightarrow{a.s.} Y$, then $X_n + Y_n \xrightarrow{a.s.} X + Y$.

Problem 2: Prove that on a complete probability space, if $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y$, then $X_n + Y_n \xrightarrow{P} X + Y$.

Problem 3: A bank teller serves customers standing in the queue one by one. Suppose that the service time X_i for customer i has mean $\mathbb{E}(X_i) = 2$ (minutes) and $\text{Var}(X_i) = 1$. We assume that service times for different bank customers are independent. Let Y be the total time the bank teller spends serving 50 customers. Find $\mathbb{P}(90 < Y < 110)$.