UNIVERSITY OF TORONTO | **Statistical Sciences**

DoSS Summer Bootcamp Probability Module 7

Ichiro Hashimoto

University of Toronto

July 20, 2026

Recap

Learnt in last module:

- Covariance
  - Covariance as an inner product
  - Correlation
  - Cauchy-Schwarz inequality
  - Uncorrelatedness and Independence
- Concentration
  - Markov's inequality
  - Chebyshev's inequality
  - Chernoff bounds

Outline

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

Stochastic Convergence

Recall: Convergence

Convergence of a sequence of numbers

A sequence $a_1, a_2, \cdots$ converges to a limit a if

$$\lim_{n\to\infty} a_n = a.$$

That is, for any $\epsilon > 0$, there exists an $N(\epsilon)$ such that

$$\underline{|a_n - a| < \epsilon}, \quad \underline{\forall n > N(\epsilon)}.$$

For any $n > N(\epsilon)$

a_n stays within ϵ from a.

Stochastic Convergence

Recall: Convergence

Convergence of a sequence of numbers

A sequence $a_1, a_2, \cdots$ converges to a limit a if

$$\lim_{n\to\infty} a_n = a.$$

That is, for any $\epsilon > 0$, there exists an $N(\epsilon)$ such that

$$\frac{1}{\epsilon} \leq \lceil \tfrac{1}{\epsilon} \rceil$$

$$|a_n - a| < \epsilon, \quad \forall n > N(\epsilon).$$

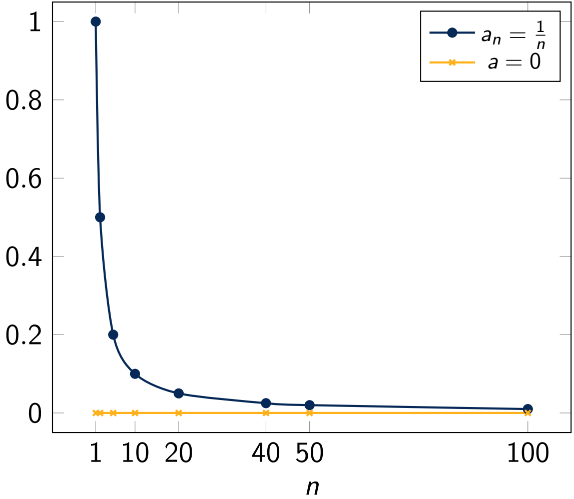
Example: $a_n = \frac{1}{n}$, $\forall \epsilon > 0$, take $N(\epsilon) = \lceil \frac{1}{\epsilon} \rceil$, then for $n > N(\epsilon) = \lceil \frac{1}{\epsilon} \rceil$

$$\Leftrightarrow \frac{1}{n} < \frac{1}{\lceil \frac{1}{\epsilon} \rceil} \leq \epsilon$$

$$|a_n - 0| = \underset{\frac{1}{n}}{a_n} < \epsilon, \quad \lim_{n\to\infty} a_n = 0.$$

$$\frac{1}{n} < \frac{1}{\lceil \frac{1}{\epsilon} \rceil} \leq \epsilon$$

Stochastic Convergence



- Capture the property of a series as $n \to \infty$;
- The limit is something where the series concentrate for large n;
- $|a_n - a|$ quantifies the closeness of the series and the limit.

Stochastic Convergence

Observation: closeness of random variables

Sample mean of i.i.d. random variables

For i.i.d. random variables $X_i, i = 1, \cdots, n$ with $\mathbb{E}(X_i) = \mu$, $Var(X_i) = \sigma^2$, then for the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$,

$$\mathbb{E}(\bar{X}) = \mu, \quad Var(\bar{X}) = \frac{\sigma^2}{n}.$$

Proof: $\mathbb{E}\bar{X} = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}X_i = \frac{1}{n}\sum_{i=1}^n \mu = \mu$

$$Var(\bar{X}) = \mathbb{E}(\bar{X}-\mu)^2 = \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^n X_i - \mu\right)^2$$

$$= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n (X_i-\mu)\right]^2$$

UNIVERSITY OF TORONTO

$$= \mathbb{E}\left[\left(\frac{1}{n}\right)^2 \cdot \left(\sum_{i=1}^{n} (X_i-\mu)^2 + \sum_{i\neq j} (X_i-\mu)(X_j-\mu)\right)\right]$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} \underbrace{\mathbb{E}(X_i-\mu)^2}_{\substack{=\mathrm{Var}(X_i) \\ =\sigma^2}} + \frac{1}{n^2}\sum_{i\neq j} \underbrace{\mathbb{E}\left[(X_i-\mu)(X_j-\mu)\right]}_{\substack{\text{by independence} \\ = \mathbb{E}(X_i-\mu)\cdot\mathbb{E}(X_j-\mu) \\ = 0\cdot 0 = 0}}$$

$$= \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$

Stochastic Convergence

Example:

Further suppose $X_i, i = 1, \cdots, n$ i.i.d. with distribution $\mathcal{N}(\mu, \sigma^2)$, then $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, so we can draw the probability density plot of $\bar{X}$.

$E\bar{X} = \mu$

$Var(\bar{X}) = \frac{\sigma^2}{n}$

UNIVERSITY OF TORONTO

Stochastic Convergence

Example:

Further suppose $X_i, i = 1, \cdots, n$ i.i.d. with distribution $\mathcal{N}(\mu, \sigma^2)$, then $\bar{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$, so we can draw the probability density plot of $\bar{X}$.

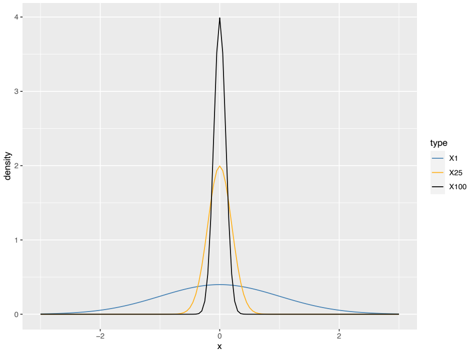


Figure: Probability density curve of sample mean of normal distribution

Stochastic Convergence

Intuition:

- Series of numbers $a_n \quad \Rightarrow \quad$ Series of random variables X_n;
- Limit $a \quad \Rightarrow \quad$ Limit X;
- How to quantify the closeness? ($|X_n - X|$?)

Stochastic Convergence

Intuition:

- Series of numbers $a_n \quad \Rightarrow \quad$ Series of random variables X_n;
- Limit $a \quad \Rightarrow \quad$ Limit X;
- How to quantify the closeness? ($|X_n - X|$?)

Pointwise convergence / Sure convergence

Suppose random variables X_n and X are defined over the same probability space, then we say X_n converges to X pointwise if

$$\lim_{n\to\infty} X_n(\omega) = X(\omega), \ \forall \omega \in \Omega.$$

convergence holds at each point ω

Stochastic Convergence

Intuition:

- Series of numbers $a_n \quad \Rightarrow \quad$ Series of random variables X_n;
- Limit $a \quad \Rightarrow \quad$ Limit X;
- How to quantify the closeness? $(|X_n - X|?)$

Pointwise convergence / Sure convergence

Suppose random variables X_n and X are defined over the same probability space, then we say X_n converges to X pointwise if

$$\lim_{n \to \infty} X_n(\omega) = X(\omega), \ \forall \omega \in \Omega.$$

Remark:

Incorporate probability measure in some sense.

Stochastic Convergence

Alternatives of describing the closeness:

- Utilize CDF: $F_{X_n}(x) - F_X(x)$;
- Utilize probability of an event: $\mathbb{P}(|X_n - X| > \epsilon)$;
- Utilize the probability over all ω: $\mathbb{P}(\lim_{n\to\infty} X_n(\omega) = X(\omega))$;
- Utilize mean/moments: $\mathbb{E}|X_n - X|^p$.

Stochastic Convergence

Quantify closeness by CDF

Convergence in distribution

A sequence $X_1, X_2, \cdots$ of real-valued random variables is said to converge in distribution, or converge weakly to a random variable X if

$$\lim_{n \to \infty} F_n(x) = F(x),$$

for every number $x \in \mathbb{R}$ at which $F(\cdot)$ is continuous. Here, $F_n(\cdot)$ and $F(\cdot)$ are the cumulative distribution functions of the random variables X_n and X, respectively.

CDF is right continuous

- CDF has at most countable discontinuities

Notation:

$X_n \xrightarrow{d} X, \quad X_n \xrightarrow{\mathcal{D}} X, \quad X_n \Rightarrow X.$

UNIVERSITY OF TORONTO

Stochastic Convergence

Convergence in distribution

A sequence $X_1, X_2, \cdots$ of real-valued random variables is said to converge in distribution, or converge weakly to a random variable X if

$$\lim_{n \to \infty} F_n(x) = F(x),$$

for every number $x \in \mathbb{R}$ at which $F(\cdot)$ is continuous. Here, $F_n(\cdot)$ and $F(\cdot)$ are the cumulative distribution functions of the random variables X_n and X, respectively.

Notation:

$X_n \xrightarrow{d} X, \quad X_n \xrightarrow{\mathcal{D}} X, \quad X_n \Rightarrow X.$

Remark:

X_n and X do not need to be defined on the same probability space.

Stochastic Convergence

Example:

Let $X_n = Z + \frac{1}{n}$, where $\underbrace{Z \sim \mathcal{N}(0,1)}$, then

symmetric.

- $X_n \xrightarrow{d} Z$,
- $X_n \xrightarrow{d} -Z$,
- $X_n \xrightarrow{d} Y$, $Y \sim \mathcal{N}(0,1)$.

Proof:

$$F(X_n \le x) = \mathbb{P}(X_n \le x) = \mathbb{P}(Z + \tfrac{1}{n} \le x) = \mathbb{P}(Z \le x - \tfrac{1}{n}) = F_Z(x - \tfrac{1}{n})$$

$$\underbrace{\lim_{n \to \infty} F_Z(x - \tfrac{1}{n})}_{\text{Left limit}} = F_Z(x)$$

by continuity of CDF of normal distribution.

$F(X_n \leq x) = \cdots\cdots = \mathbb{P}\left(Z \leq x - \frac{1}{n}\right)$ since Z is symmetric

$= \mathbb{P}\left(Z \geq -\left(x - \frac{1}{n}\right)\right)$

$= \mathbb{P}\left(-Z \leq x - \frac{1}{n}\right)$

$= F_{-Z}\left(x - \frac{1}{n}\right) \longrightarrow F_{-Z}(x)$

since $-Z \sim N(0,1)$, i.e. F_{-Z} is continuous everywhere.

$X_n \xrightarrow{d} Z$ and $X_n \xrightarrow{d} -Z$ at the same time is not strange

since Z and $-Z$ share the same CDF.

$X_n \xrightarrow{d} Y$ is obvious since the CDF of Y is same as that of Z.

Stochastic Convergence

Quantify the closeness using probability of an event

Convergence in probability

A sequence X_n of random variables converges in probability towards the random variable X if for all $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0.$$

event where X_n and X differs at least ε

Notation: $X_n \xrightarrow{p} X$, $\quad X_n \xrightarrow{P} X$.

Remark:

X_n and X need to be defined on the same probability space.

Probability of X_n and X differs $\to 0$

Stochastic Convergence

Examples:

- Let $X_n = Z + \frac{1}{n}$, where $Z \sim \mathcal{N}(0,1)$, then $X_n \xrightarrow{P} Z$.

  Proof: $\forall \varepsilon > 0$,

  $$\mathbb{P}(|X_n - Z| > \varepsilon) = \mathbb{P}\left(\frac{1}{n} > \varepsilon\right) = \begin{cases} 0 & \cdots\ n \geq \frac{1}{\varepsilon} \\ 1 & \cdots\ n < \frac{1}{\varepsilon} \end{cases}$$

  Thus $\lim_{n\to\infty} \mathbb{P}(|X_n - Z| > \varepsilon) = 0$.

- Let $X_n = Z + Y_n$, where $Z \sim \mathcal{N}(0,1)$, $\mathbb{E}(|Y_n|) = \frac{1}{n}$, then $X_n \xrightarrow{P} Z$.

  Proof: Let $\forall \varepsilon > 0$.

  by Markov

  $$\mathbb{P}(|X_n - Z| > \varepsilon) = \mathbb{P}(|Y_n| > \varepsilon) \leq \frac{\mathbb{E}|Y_n|}{\varepsilon} = \frac{1}{\varepsilon n}$$

  By taking $n \to \infty$, we conclude that

  $$\lim_{n\to\infty} \mathbb{P}(|X_n - Z| > \varepsilon) = 0.$$

UNIVERSITY OF TORONTO

Stochastic convergence

Very similar to pointwise convergence.

Convergence almost surely

A sequence X_n of random variables converges almost surely or almost everywhere or with probability 1 or strongly towards X means that

$$\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = \mathbb{P}\left(\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right) = 1.$$

$\Updownarrow$

$\lim X_n = X$ with probability 1

$\Updownarrow$

$\lim X_n \neq X$ with probability 0

Notation: $X_n \xrightarrow{a.s.} X$.

Remark:
X_n and X need to be defined on the same probability space.

Converge a.s. allows $\lim X_n \neq X$ with probability 0

Stochastic convergence

Examples:

- Let $X_n = Z + \frac{1}{n}$, where $Z \sim \mathcal{N}(0,1)$, then $X_n \xrightarrow{a.s.} Z$.

  Proof: This is actually pointwise convergence

  For any $\omega \in \Omega$. $\lim_{n\to\infty} X_n(\omega) = \lim_{n\to\infty} \left(Z(\omega) + \frac{1}{n}\right)$

  $= Z(\omega)$.

- Let $X_n = Z + Y_n$, where $Z \sim \mathcal{N}(0,1)$, $\mathbb{E}(|Y_n|) = \frac{1}{n}$, do we have $X_n \xrightarrow{a.s.} Z$?

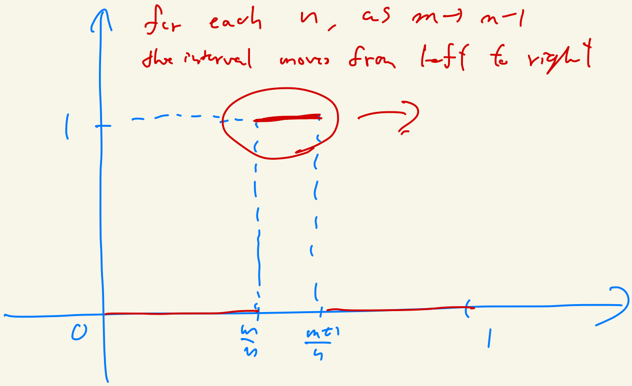
  Proof: Earlier we showed $X_n \xrightarrow{P} Z$

  No. $X_n \xrightarrow{a.s.} Z$ is false in general.

Counterexample.

$$Y_{m,n} : \mathbb{R} \to \mathbb{R} \qquad 0 \leq m \leq n-1$$

$$Y_{m,n}(\omega) = \begin{cases} 1 & \text{if } \omega \in \left[\frac{m}{n}, \frac{m+1}{n}\right] \\ 0 & \text{otherwise.} \end{cases}$$



The interval $\left[\frac{m}{n}, \frac{m+1}{n}\right]$ scans $[0,1]$ from left to right for every n.

Then, we have. $\lim_{n\to+\infty} Y_{m,n}(\omega)$ does not exist for $\forall\, \omega \in [0,1]$

However, we have $\mathbb{E}\, Y_{m,n}(\omega) = \frac{1}{n}$

Therefore,

$$X_{m,n} = Z + Y_{m,n} \xrightarrow{P} Z$$

while $X_{m,n} \xrightarrow{a.s.} Z$

Stochastic convergence

Quantify closeness by $\mathbb{E}|X_n - X|^p$, i.e. pth moment

L^p norm of X = $\|X\|_{L^p} = (\mathbb{E}|X|^p)^{1/p}$

Convergence in L^p

A sequence $\{X_n\}$ of random variables converges in L_p to a random variable X, $p \geq 1$, if

$$\lim_{n\to\infty} \mathbb{E}|X_n - X|^p = 0$$

$\iff \lim_{n\to\infty} \|X_n - X\|_{L^p} = 0$

Notation: $X_n \xrightarrow{L^p} X$.

L^p norm allows the triangle inequality

Remark:
X_n and X need to be defined on the same probability space.

$\|X + Y\|_{L^p} \leq \|X\|_{L^p} + \|Y\|_{L^p}$

Stochastic convergence

Examples:

- Let $X_n = Z + \frac{1}{n}$, where $Z \sim \mathcal{N}(0,1)$, then $X_n \xrightarrow{L^p} Z$.

  Proof: $\mathbb{E}|X_n - Z|^p = \mathbb{E}(\frac{1}{n})^p = \frac{1}{n^p} \to 0$

  Thus, $X_n \to Z$ in L^p.

- Let $X_n = Z + Y_n$, where $Z \sim \mathcal{N}(0,1)$, $\mathbb{E}(|Y_n|^p) = \frac{1}{n}$, then $X_n \xrightarrow{L^p} Z$.

  Proof: $\mathbb{E}|X_n - Z|^p = \mathbb{E}|Y_n|^p = \frac{1}{n} \to 0.$

  Thus, $X_n \to Z$ in L^p

Stochastic convergence

Recall: A random variable $X \in L^p$ if $\|X\|_{L^p} = (E|X|^p)^{1/p} < \infty$.
$X_n \to X$ in L^p if $\lim_{n\to\infty} \|X_n - X\|_{L^p} = 0$

Monotonicity of L^p Convergence

If $q > p > 0$, L^q convergence implies L^p convergence

Proof: We use Lyapunov inequality:

If $q > p > 0$, then $\|X\|_{L^p} \le \|X\|_{L^q}$. ($L^q$ norm is bigger than L^p norm)

Applying this to $X_n - X$, we have.

$$\|X_n - X\|_{L^p} \le \|X_n - X\|_{L^q} \quad \cdots\cdots \quad (*)$$

Since $X_n \to X$ in L^q, we have $\lim_{n\to\infty} \|X_n - X\|_{L^q} = 0$.

Therefore, we have $\lim_{n\to\infty} \|X_n - X\|_{L^p} = 0$ by $(*)$.

Stochastic convergence

Recall: X_n converges to X in probability if for any $\epsilon > 0$ $\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$.

need to show this from assumption: $\|X_n - X\|_{L^p} \to 0$

L^p convergence implies Convergence in Probability

If $X_n \to X$ in L^p, then $X_n \to X$ in probability.

Proof:

For any $\varepsilon > 0$,

$$\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{P}(|X_n - X|^p > \varepsilon^p)$$

$$\leq \frac{\mathbb{E}|X_n - X|^p}{\varepsilon^p} \quad \text{by Markov inequality}$$

Since $X_n \to X$ in L^p, we have $\lim_{n\to\infty} \mathbb{E}|X_n - X|^p = 0$.

Therefore, we have $\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$

Stochastic convergence

Recall: X_n converges to X in probability if for any $\epsilon > 0$ $\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$.

a.s. Convergence implies Convergence in Probability

If $X_n \to X$ almost surely, then $X_n \to X$ in probability.

Proof:

omit

Stochastic convergence

Recall: X_n converges to X in distribution if for any continuity point x of $P(X \leq x)$, $\lim_{n\to\infty} P(X_n \leq x) = P(X \leq x)$ holds.

Convergence in Probability implies Convergence in Distribution

If $X_n \to X$ in probability, then $X_n \to X$ in distribution.

Proof: Omitted

UNIVERSITY OF TORONTO

Stochastic convergence

Relationship between convergences (on complete probability space):

$$
\begin{array}{ccccc}
\xrightarrow{L^s} & \underset{s>r\geq 1}{\Rightarrow} & \xrightarrow{L^r} & & \\
 & & \Downarrow & & \\
\xrightarrow{\text{a.s.}} & \Rightarrow & \xrightarrow{p} & \Rightarrow & \xrightarrow{d}
\end{array}
$$

Figure: relationship between convergences

Stochastic convergence

Highlights:

- 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 convergence in distribution:

$$X_n \xrightarrow{P} X \quad \Rightarrow \quad X_n \xrightarrow{d} X;$$

- If X_n converges in distribution to a constant c, then X_n converges in probability to c:

$$X_n \xrightarrow{d} c \quad \Rightarrow \quad X_n \xrightarrow{P} c, \quad \text{provided } c \text{ is a constant.}$$

Problem Set

Problem 1: Prove that on a complete probability space, if $X_n \xrightarrow{L^p} X$, then $X_n \xrightarrow{P} X$. (Hint: use Markov's inequality)

Problem 2: Let $X_1, \cdots, X_n$ be i.i.d. random variables with $Bernoulli(p)$ distribution, and $X \sim Bernoulli(p)$ is defined on the same probability space, independent with X_i's. Does X_n converge in probability to X?

Problem 3: Give an example where X_n converges in distribution to X, but not in probability.