UNIVERSITY OF TORONTO | Statistical Sciences

DoSS Summer Bootcamp Probability Module 7

Ichiro Hashimoto

University of Toronto

July 23, 2024

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

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

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

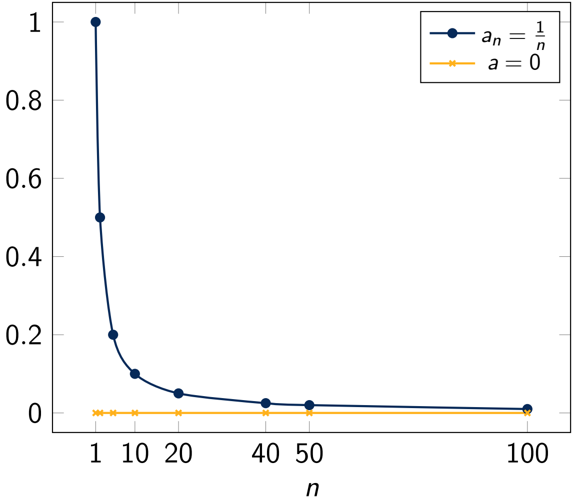
$$|a_n - a| < \epsilon, \quad \forall n > N(\epsilon).$$

a_n is concentrated on $(a-\varepsilon, a+\varepsilon)$ for $\forall n > N(\varepsilon)$

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

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

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 $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$,

independent and identically distributed.

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

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

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

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

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

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

$$= \frac{1}{n^2} \sum_{i=1}^{n} \underbrace{\mathbb{E}(X_i - \mu)^2}_{\substack{= \mathrm{Var}(X_i) \\ = \sigma^2.}} + \frac{1}{n^2} \sum_{i \neq j} \underbrace{\mathbb{E}(X_i - \mu)(X_j - \mu)}_{\substack{\text{Note that} \\ X_i \text{ and } X_j \text{ are} \\ \text{independent}}}$$

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

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

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 $\bar{X}$.

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 $\bar{X}$.

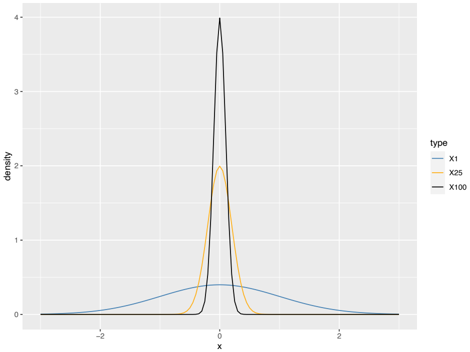


Figure: Probability density curve of sample mean of normal distribution

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

a point

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

Once ω is fixed $X_n(\omega)$ is just a deterministic sequence.

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

Use CDF to capture closeness of X_n and X.

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:

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

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.

Because it's about closeness of distributions

Stochastic Convergence

Example:

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

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

A new random variable which could be independent / defined on a different sample space.

Proof:

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

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

$= \Phi\left(x - \frac{1}{n}\right)$, where Φ is CDF of $\mathcal{N}(0,1)$

Since Φ is continuous on $\mathbb{R}$,

$$\lim_{n \to \infty} \Phi\left(x - \frac{1}{n}\right) = \Phi(x) = \mathbb{P}(Z \leq x).$$

2.) Since $Z \sim N(0,1)$ is symmetric,

$$\mathbb{P}(-Z \leq x) = \mathbb{P}(Z \leq x).$$

3.) Since $Y \sim N(0,1)$,

$$\mathbb{P}(Y \leq x) = \mathbb{P}(Z \leq x).$$

Stochastic Convergence

Convergence in probability

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

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

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.

Stochastic Convergence

Examples:

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

  Proof: Let $\forall \varepsilon > 0$. $\mathbb{P}(|X_n - Z| > \varepsilon) = \mathbb{P}(\frac{1}{n} > \varepsilon) = 0$

  for any $n > \varepsilon^{-1}$

  Thus $\mathbb{P}(|X_n - Z| > \varepsilon) \to 0$ and $X_n \xrightarrow{P} Z$.

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

  $\to 0$

UNIVERSITY OF TORONTO

Stochastic convergence

It looks like pointwise convergence but slightly different.

a.s. convergence allows $\lim_{n\to\infty} X_n(\omega) \neq X(\omega)$ on a set with probability 0.

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

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

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

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

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

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

- 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: we already know $X_n \xrightarrow{P} Z$

  No, X_n does not converge to Z a.s.

< Counter example. >

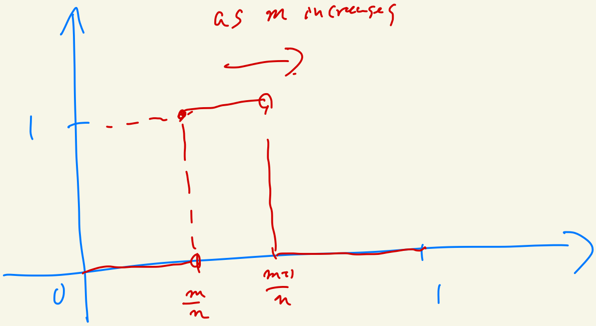
$$\Omega = (0,1), \quad \mathbb{P} \sim \mathrm{Unif}(\Omega)$$

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

for $0 \le m \le n-1$

$$\mathbb{P}(Y_{m,n} = 1) = \frac{1}{n}$$

Thus $\mathbb{E}|Y_{m,n}| = \frac{1}{n}$.



That means $\lim_{n\to\infty} Y_{m,n}(\omega)$ does not exist

It is impossible for $X_n(\omega) \to Z(\omega)$ without $Y_n(\omega) \to 0$.

Stochastic convergence

using pth moment to capture closeness

Convergence in L^p

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

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

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

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

UNIVERSITY OF TORONTO

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:

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

- 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:

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

Stochastic convergence

L^p norm

Recall: A random variable $X \in L^p$ if $\boxed{\|X\|_{L^p} = (E|X|^p)^{1/p}} < \infty$.

$X_n \to X$ 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: Lyapunov inequality

$$\underset{\|X\|_{L^p}}{\underset{\shortparallel}{\left(\mathbb{E}|X|^p\right)^{1/p}}} \leq \underset{\|X\|_{L^q}}{\underset{\shortparallel}{\left(\mathbb{E}|X|^q\right)^{1/q}}} \quad \text{if } 0<p<q.$$

By Lyapunov inequality

$$\left(\mathbb{E}|X_n - X|^p\right)^{1/p} \leq \left(\mathbb{E}|X_n - X|^q\right)^{1/q} \to 0 \text{ by assumption}$$

thus $X_n \xrightarrow{L^p} X$.

Stochastic convergence

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

L^p convergence implies Convergence in Probability

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

Proof: By Markov inequality

$$\mathbb{P}(|X_n - X| > \varepsilon) = \mathbb{P}(|X_n - X|^p > \varepsilon^p) \overset{\text{Markov}}{\leq} \frac{\mathbb{E}|X_n - X|^p}{\varepsilon^p} \to 0$$

since $X_n \to X$ in L^p

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

Stochastic convergence

Relationship between convergences (on complete probability space):

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

strong

weakest

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.