UNIVERSITY OF TORONTO | Statistical Sciences

DoSS Summer Bootcamp Probability Module 7

Ichiro Hashimoto

University of Toronto

July 24, 2023

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

an is close to a by ε

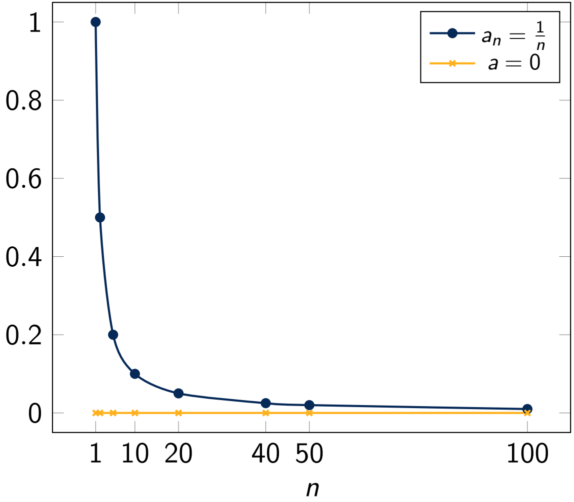
$$|a_n - a| < \epsilon, \quad \forall n > N(\epsilon).$$

when n is sufficiently large

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

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

Proof:

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

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

UNIVERSITY OF TORONTO

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

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

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

$$= \frac{1}{n^2}\sum_{i=1}^{n}E(X_i-\mu)^2 + \frac{1}{n^2}\sum_{i\neq j}\underbrace{E(X_i-\mu)(X_j-\mu)}_{\downarrow\ \text{independen}}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\underbrace{Var(X_i)}_{=\sigma^2} + \frac{1}{n^2}\sum_{i\neq j}\underbrace{E(X_i-\mu)\cdot E(X_j-\mu)}_{=0}.$$

$$= \frac{n\sigma^2}{n^2} + 0 = \frac{\sigma^2}{n}.$$

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 $\bar{X}$.

By the previous slide we know $E\bar{X} = \mu$, $Var(\bar{X}) = \frac{\sigma^2}{n}$.

In addition, when X_i are independent normal, then

$\sum_{i=1}^{n} X_i$ is also normally distributed.

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 $\bar{X}$.

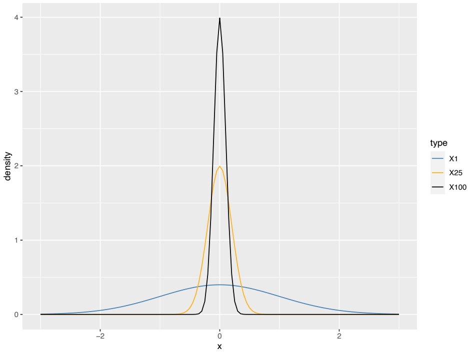


Figure: Probability density curve of sample mean of normal distribution

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

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

for each point w

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

using CDF to quantify "closeness"

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

or weak convergence.

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

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:

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

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 $Z \sim \mathcal{N}(0,1)$, then

Let Φ be CDF of Z..

Note that Φ is continuous.

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

new random variable which could be independent from Z

Proof:

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

Since Φ is continuous, $\lim_{n\to\infty} \Phi(x - \frac{1}{n}) = \Phi(x) = \mathbb{P}(Z \leq x)$

So, $\lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(Z \leq x)$ $\therefore X_n \xrightarrow{d} Z$

2) This is due to $-Z \sim N(0,1)$

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

$\therefore \lim_{n \to \infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(Z \leq x) = \mathbb{P}(-Z \leq x)$.

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 $\forall \varepsilon > 0$. $\mathbb{P}(|X_n - Z| > \varepsilon) = \mathbb{P}(\frac{1}{n} > \varepsilon) = 0$ if $n \geq \varepsilon^{-1}$.

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

- 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:

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

  $\longrightarrow 0$ as $n \to \infty$

almost always = a.s.

Stochastic convergence

a.s.

a.e.

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

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

pointwise.

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

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:

  $$\lim_{n\to\infty} X_n = \lim_{n\to\infty} \left(Z + \frac{1}{n}\right) = Z + \lim_{n\to\infty} \frac{1}{n} = Z$$

- 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: No, it does not converge almost surely.

To have $X_n \xrightarrow{a.s.} Z$, we must have $Y_n \xrightarrow{a.s.} 0$.

However, there is a counter example.

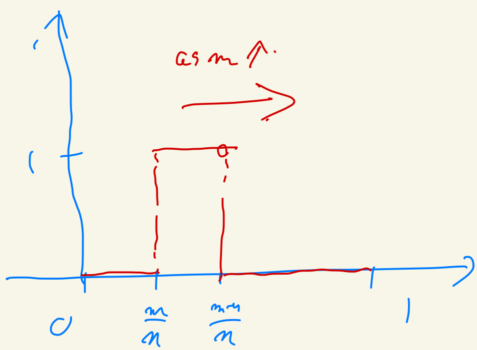
e.g.) $\Omega = (0,1)$. $\mathbb{P} \sim \text{Unif}(\Omega)$.

Define.

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

for $0 \leq m \leq n-1$. $\therefore \mathbb{P}(Y_{n,m} = 1) = \frac{1}{n}$

So, $E\, Y_{n,m} = \frac{1}{n}$.



For each n, we run m from $0 \sim n-1$

So, the interval run all points on $\Omega = (0,1)$.

Then, that means, for any point $w \in \Omega$,

$\lim_{n \to \infty} Y_n(w)$ does not exist.

Stochastic convergence

Convergence in L^p

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

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

$p > 0$

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

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 $X_n \xrightarrow{L^p} Z$.

  Proof:

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

- 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:

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

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

UNIVERSITY OF TORONTO

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.