UNIVERSITY OF TORONTO | Statistical Sciences

DoSS Summer Bootcamp Probability Module 7

Miaoshiqi (Shiki) Liu

University of Toronto

July 24, 2022

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

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

$a_1 \quad a_2 \quad a_3 \quad \cdots \quad a_{N(\varepsilon)}, \quad a_{N(\varepsilon)+1} \cdots\cdots$

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

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

UNIVERSITY OF TORONTO

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

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

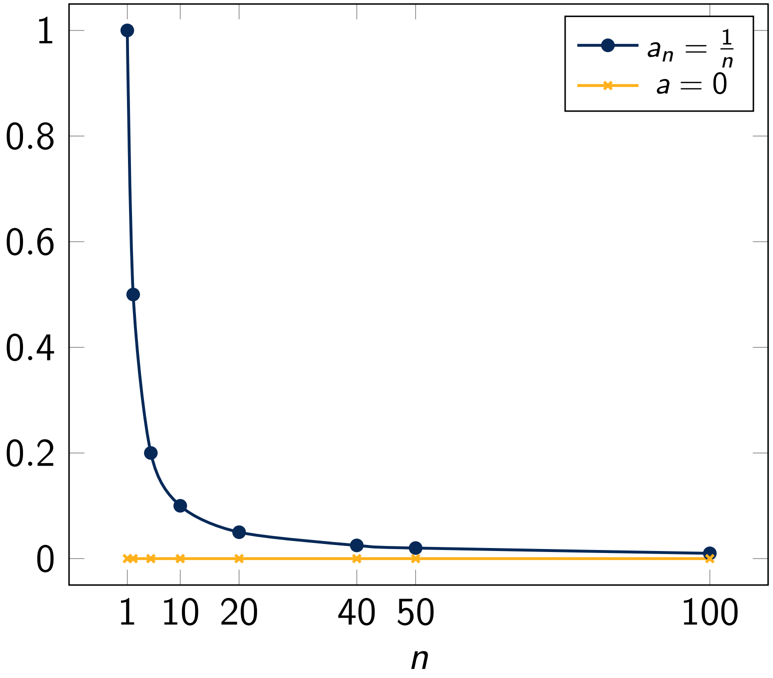
Example: $a_n = \frac{1}{n}$, $\forall \epsilon > 0$, take $N(\epsilon) = \lceil \frac{1}{\epsilon} \rceil$, then for $\underline{n > N(\epsilon)}$,

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

$$\frac{1}{n} < \varepsilon \qquad \frac{1}{n} < \frac{1}{N(\varepsilon)} \leq \varepsilon.$$

$$N(\varepsilon) = \lceil \frac{1}{\varepsilon} \rceil \geq \frac{1}{\varepsilon}$$

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

X_i: $X_1 \; X_2 \; \cdots \; X_n \; \cdots$

$\bar{X}_n$:

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 $\underline{\mathbb{E}(X_i) = \mu}$, $Var(X_i) = \sigma^2$, 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:

$$E(\bar{x}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \mu.$$

$$Var(\bar{x}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \left(\frac{1}{n}\right)^2 \sum_{i=1}^{n} Var(x_i)$$

$$= \frac{1}{n^2} \times n\sigma^2 = \frac{\sigma^2}{n}.$$

UNIVERSITY OF TORONTO

Stochastic Convergence

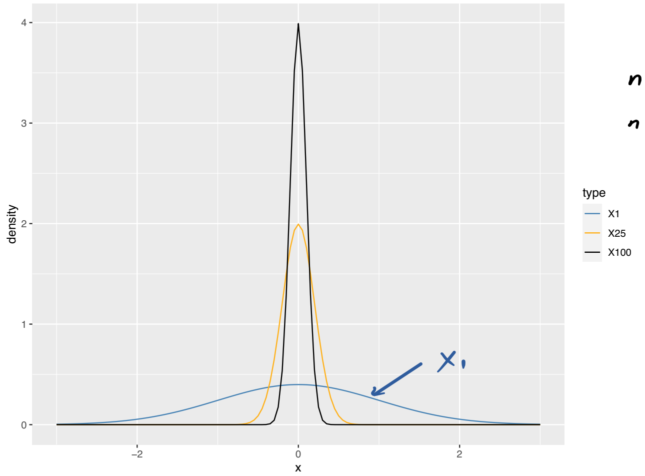
Example:

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 $\bar{X}$.

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 $\bar{X}$.



$\mu = 0$

$\sigma^2 = 1$

$n = 1.\quad \bar{x} \sim N(0, 1)$

$n = 25\quad \bar{x} \sim N(0, \frac{1}{25})$

Figure: Probability density curve of sample mean of normal distribution

Stochastic Convergence

Intuition:

- Series of numbers a_n $\Rightarrow$ Series of random variables X_n;
- Limit a $\Rightarrow$ Limit X; $|a_n - a|$
- 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

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

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.

UNIVERSITY OF TORONTO

Stochastic Convergence

X_n. X.

Alternatives of describing the closeness:

- Utilize CDF: $F_{X_n}(x) - F_X(x)$; $P(X \le x) = \mu((-\infty, x]) = F_X(x)$.
- Utilize probability of an event: $\mathbb{P}(|X_n - X| > \epsilon)$; $|X_n - X| > \varepsilon$.
- 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

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

X. Bernoulli 0, 1

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

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

$x \neq 0, 1$.

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:

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

$F_n(x) = P(X_n \leq x)$ $\quad (\Omega, \mathcal{F}, P)$

Remark:

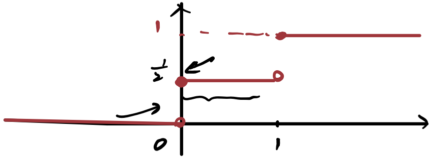
X_n and X do not need to be defined on the same probability space.

Stochastic Convergence

CDF right

$$\lim_{y \downarrow x} F(y) = F(x)$$

$$F(x) = P(X \le x)$$



Example:

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

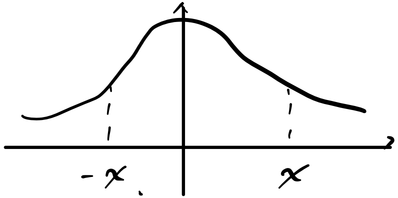
- $X_n \xrightarrow{d} Z$,
- $X_n \xrightarrow{d} -Z$,
- $X_n \xrightarrow{d} Y$, $Y \sim \mathcal{N}(0,1)$.

Proof: $F_{X_n}(x) = P(X_n \le x) = P(Z + \frac{1}{n} \le x) = P(Z \le x - \frac{1}{n})$

$$= \underbrace{F_Z(x - \tfrac{1}{n})}_{\Delta} .$$

$n \to \infty$. $\longrightarrow F_Z(x)$

$X_n \xrightarrow{d} X$.



$-Z \sim N(0,1)$

$F_{-Z}(x) = F_Z(x)$.

$Y \sim N(0,1)$

$F_{X_n}(x) = F_Z(x - \frac{1}{n}) = F_{-Z}(x - \frac{1}{n})$

$= F_Y(x - \frac{1}{n})$

Stochastic Convergence

Convergence in probability

A sequence X_n of random variables converges in probability towards the random variable X if for all $\epsilon > 0$, $|a_n - a| < \varepsilon \quad n \to \infty$

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

$\lim_{n \to \infty} \underbrace{F_{X_n}(x)} = \underbrace{F_X(x)}$ $\quad \underbrace{P(X \le x)}$

$P(X_n \le x)$

$P(\{w : X_n(w) \le x\})$

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

Remark: $\{w : |X_n(w) - X(w)| > \varepsilon\}$

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

Stochastic Convergence

Examples: $X_n(\omega) = Z(\omega) + \frac{1}{n}.$

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

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

Proof: $\forall \varepsilon > 0.$

$$P(|X_n - Z| > \varepsilon) = P(\underbrace{\tfrac{1}{n} > \varepsilon}_{\{\frac{1}{n} > \varepsilon\}}) \longrightarrow 0.$$

$n \to \infty.$

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

Proof: $\forall \varepsilon > 0$

$$P(|X_n - Z| > \varepsilon)$$

$$= \underline{P(|Y_n| > \varepsilon)} \le \frac{E|Y_n|}{\varepsilon} = \frac{1}{n\varepsilon} \longrightarrow 0.$$

$n \to \infty.$

Stochastic 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

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

$X_n(\omega)$ $X(\omega)$

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

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 $X_n \xrightarrow{a.s.} Z$.

Proof: $P(\{w : \lim_{n\to\infty} X_n(w) = Z(w)\})$

$= P(\{w : \lim_{n\to\infty} Z(w) + \frac{1}{n} = Z(w)\})$

$= P(\{w : \lim_{n\to\infty} \frac{1}{n} = 0\}) = 1.$

- 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: $\{w : \lim_{n\to\infty} X_n(w) = Z(w)\}$

$= \{w : \lim_{n\to\infty} Z(w) + Y_n(w) = Z(w)\}$

$= \{w : \lim_{n\to\infty} Y_n(w) = 0\}$

$Y_n = \text{Bernoulli}(\frac{1}{n})$

$= \begin{cases} 1. & \frac{1}{n} \\ 0. & 1 - \frac{1}{n}. \end{cases}$

$\{Y_n = 1\}$

$\sum_{n=1}^{\infty} P(Y_n = 1) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty.$

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

$$\lim_{n\to\infty} \underline{\mathbb{E}|X_n - X|^p} = 0 \qquad \frac{3}{2}$$

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: $E|X_n - Z|^p = E\underbrace{|\tfrac{1}{n}|^p}_{} = (\frac{1}{n})^p \longrightarrow 0.$

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

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

Stochastic convergence

Relationship between convergences (on complete probability space):

$$\xrightarrow{L^s} \underset{s>r\geq 1}{\Longrightarrow} \xrightarrow{L^r}$$

$$\Downarrow$$

$$\xrightarrow{\text{a.s.}} \Longrightarrow \xrightarrow{p} \Longrightarrow \xrightarrow{d}$$

Figure: relationship between convergences

Stochastic convergence

$\{\omega : |X_n - c| > \varepsilon\}$ $c(\omega')$

$= \{\omega : X_n - c > \varepsilon\} \cup \{\omega : X_n - c < -\varepsilon\}$

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:

$\{\omega : |X_n - X| > \varepsilon\}$

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

$X = c$.

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.