Lecture 17

giovedì 13 novembre 2025 16:25

CONVERGENCE of RANDOM VARIABLES

$X_n \to X$ $\quad\quad \lim_n X_n$

$X : \Omega \to \mathbb{R}$, sequence $X_n : \Omega \to \mathbb{R}$ $\quad n = 1, 2, \dots$

$X_n \xrightarrow{n\to\infty} X$ in what sense ?

$(f_n)_n$ sequence of functions :

$f_n \xrightarrow[n\to\infty]{} f$??

• POINTWISE CONVERGENCE

$f_n \to f$ POINTWISE $\quad\quad$ domain of functions

if $\lim_{n\to\infty} f_n(x) = f(x) \quad \forall x \in D$

for RANDOM VARIABLES

$X_n \to X$ pointwise if $X_n(\omega) \xrightarrow{n\to\infty} X(\omega) \quad \forall \omega \in \Omega$

$X \stackrel{A.E.}{=} Y$ almost everywhere (ALMOST SURELY)

if $P(X = Y) = 1$

$= P(\{\omega \in \Omega : X(\omega) = Y(\omega)\}) = 1$

Definitions of convergence of random variables :

1) ALMOST SURE CONVERGENCE

$(X_n)_n$ sequence of r.v. , $X_n : \Omega \to \mathbb{R}$

we say that X_n converges almost surely to X,

and we write $\quad X_n \xrightarrow[n\to\infty]{a.s.} X$

if $P(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$

2) CONVERGENCE IN PROBABILITY

$(X_n)_n, X \qquad X_n, X : \Omega \to \mathbb{R}$

$X_n \xrightarrow[n\to\infty]{P} X$ if

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

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

$|X_n(\omega) - X(\omega)| > \varepsilon$

$X_n(\omega) > X(\omega) + \varepsilon$ or $X_n(\omega) < X(\omega) - \varepsilon$

Note $X_n(\omega) \to X(\omega)$ if $\varepsilon(\omega)$

s.t. $|X_n(\omega) - X(\omega)| < \varepsilon(\omega)$ for n large

(this is pointwise convergence)

3) CONVERGENCE in L^p - NORM

recall $X \in L^p(\Omega)$ if $E|X|^p < \infty$

$p \geq 1$

$X_n \xrightarrow[n\to\infty]{L^p} X$ if $X_n \in L^p \ \forall n$, $X \in L^p$

and $\lim_n E[|X_n - X|^p] = 0$

4) CONVERGENCE in DISTRIBUTION

$X_n \xrightarrow[n\to\infty]{D} X$

if $\lim_{n\to\infty} E[f(X_n)] = E[f(X)]$

for any $f : \mathbb{R} \to \mathbb{R}$ bounded and continuous

Theorem

- $X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{D} X$

  The opposite implications are not true
- $X_n \xrightarrow{L^1} X \Rightarrow X_n \xrightarrow{P} X$
- $X_n \xrightarrow{P} X$ and $\sup_{n\in\mathbb{N}} E|X_n|^2 < \infty$ Then $X_n \xrightarrow{L^1} X$

- $X_n \to X$ and $\sup_{n \in \mathbb{N}} E|X_n| < \infty$ then $X_n \to X$
- $X_n \xrightarrow{P} X$ then there exist a subsequence $(n_k)_k$ s.t. $X_{n_k} \xrightarrow[k\to\infty]{a.s.} X$

MARKOV INEQUALITY

$Y : \Omega \to [0, +\infty)$ r.v, $Y \geq 0$ and $\delta > 0$. Then

$$P(Y \geq \delta) \leq \frac{E[Y]}{\delta}$$

Proof $\quad Y \geq Y \mathbb{1}_{\{Y \geq \delta\}} \geq \delta \mathbb{1}_{\{Y \geq \delta\}} \quad \forall \omega$

$$\mathbb{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases} \qquad = \begin{cases} Y(\omega) & \text{if } Y(\omega) \geq \delta \\ 0 & \text{otherwise} \end{cases}$$

$\Rightarrow$ $E[\cdot]$ monotone

$$E[Y] \geq E[\delta \mathbb{1}_{\{Y \geq \delta\}}] = \delta E[\mathbb{1}_{\{Y \geq \delta\}}] = \delta P(Y \geq \delta)$$

$$\Rightarrow P(Y \geq \delta) \leq \frac{E[Y]}{\delta}$$

Remark

This implies that $X_n \xrightarrow{L^1} X \Rightarrow X_n \to P$

Indeed: $\forall \varepsilon > 0$

$$P(\underbrace{|X_n - X|}_{\geq 0} > \varepsilon) \overset{\text{MARKOV}}{\leq} \frac{E[|X_n - X|]}{\varepsilon} \xrightarrow{n\to\infty} 0 \quad \text{as } X_n \xrightarrow{L^1} X$$

CHEBISCHEV INEQUALITY

$X \in L^2(\Omega)$

i.e. $E[|X|^2] < \infty$ square integrable r.v.

We apply Markov inequality to $Y = (X - E[X])^2$

$$P(|X - E[X]| \geq \varepsilon) = P(|X - E[X]|^2 \geq \varepsilon^2) \leq \frac{E[|X - E[X]|^2]}{\varepsilon^2} = \frac{Var(X)}{\varepsilon^2}$$

$$\leq \frac{E[|X-E[X]|^2]}{\varepsilon^2} = \frac{Var(X)}{\varepsilon^2}$$

$$P(|X-E[X]| \geq \varepsilon) \leq \frac{Var(X)}{\varepsilon^2} \qquad \forall \varepsilon > 0$$

LAW OF LARGE NUMBERS

$(X_i)_{i\in\mathbb{N}}$ sequence of random variables $X_i : \Omega \to \mathbb{R}$

$(X_i)_i$ IID = independent and identically distributed

$[\ \mu_{X_i} = \mu_{X_1} \quad \forall i \in \mathbb{N}\ ]$

$\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ EMPIRICAL MEAN

LLN $\Leftrightarrow$ $\overline{X}_n \to \mu$ $\qquad \mu = E[X_1]$

WEAK LLN $\overline{X}_n \xrightarrow{P} \mu$

STRONG LLN $\overline{X}_n \xrightarrow{a.s.} \mu$

Theorem Assume $(X_i)_i$ IID sequence of square integrable random variables. Let $\mu = E[X_i]$, $\sigma^2 = Var(X_i) < \infty$

$\forall \varepsilon > 0$ and $\forall n$

$$P(|\overline{X}_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}$$

In particular, $\overline{X}_n \xrightarrow{P} \mu$

Proof

$$E[\overline{X}_n] = E\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n \underbrace{E[X_i]}_{=\mu} = \frac{1}{n} n\mu = \mu$$

$$Var(\overline{X}_n) = Var\left(\frac{1}{n}\sum_{i=1}^n X_i\right) = \frac{1}{n^2} Var\left(\sum_{i=1}^n X_i\right)$$

$$\underset{\text{INDEP.}}{=} \frac{1}{n^2}\sum_{i=1}^n \underbrace{Var(X_i)}_{=\sigma^2} = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

We apply Chebychev inequality to $\overline{X}_n$

$$P(|\overline{X}_n - E[\overline{X}_n]| > \varepsilon) \leq \frac{Var(\overline{X}_n)}{}$$

$$P\left(\left|\overline{X}_n - E[\overline{X}_n]\right| > \varepsilon\right) \le \frac{\mathrm{Var}(\overline{X}_n)}{\varepsilon^2}$$

$$P\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) \le \frac{\sigma^2}{n\,\varepsilon^2} \qquad \boxtimes$$

Thm STRONG LAW of LARGE NUMBERS

$(X_n)_{n\in\mathbb{N}}$ IID sequence of r.v. in L^1

$(\ E|X_n| < \infty \quad \forall n\ ) \qquad \mu = E[X_n] \quad \forall n.$

Then $\overline{X}_n \xrightarrow[n\to\infty]{a.s.} \mu.$

Remark STRONG LLN $\Rightarrow$ WEAK LLN

CHERNOFF BOUNDS

$\overline{X}_n \xrightarrow{P} \mu$

Assuming more on the distribution of the random variables, we aim at getting a better rate of convergence of $P(|\overline{X}_n - \mu| \ge \varepsilon) \xrightarrow{n} 0 \quad \forall \varepsilon$

We will get exponential rate $\cdots \le e^{-n\,g(\varepsilon)}$

Let $X_1, \dots, X_n$ be IID r.v. and set

$m(t) = E[e^{tX_1}]$ moment generating function of X_1 (and also of $X_2, \dots, X_n$)

We assume that $m(t) < \infty \quad \forall t \in (-\varepsilon, \varepsilon)$ for a suitable $\varepsilon > 0$ small.

$$P(|\overline{X}_n - \mu| \ge \varepsilon) = P\left(\{\overline{X}_n - \mu \ge \varepsilon\} \cup \{\overline{X}_n - \mu \le -\varepsilon\}\right)$$
$$= \underbrace{P\left(\overline{X}_n - \mu \ge \varepsilon\right)}_{\text{UPPER TAIL}} + \underbrace{P\left(\overline{X}_n - \mu \le -\varepsilon\right)}_{\text{LOWER TAIL}}$$

$$P(\overline{X}_n \geq \mu+\varepsilon) = P(t(X_1+\dots+X_n) \geq tn(\mu+\varepsilon))$$

Let $t>0$

$$= P(\underbrace{\exp(t(X_1+\dots+X_n))}_{\geq 0} \geq \exp(tn(\mu+\varepsilon)))$$

use MARKOV INEQ

$$\leq \frac{E[\exp(t(X_1+\dots+X_n))]}{\exp(tn(\mu+\varepsilon))}$$

$$= \frac{E[e^{tX_1}\cdots e^{tX_n}]}{e^{tn(\mu+\varepsilon)}} \overset{\text{INDEPENDENT}}{=} \frac{\overbrace{E[e^{tX_1}]}^{=m(t)}\cdots \overbrace{E[e^{tX_n}]}^{m(t)}}{e^{tn(\mu+\varepsilon)}}$$

$$= \frac{(m(t))^n}{e^{tn(\mu+\varepsilon)}} \quad \text{for any } t \text{ small}$$

$$\frac{(m(t))^n}{e^{tn(\mu+\varepsilon)}} = \frac{\exp(n\log m(t))}{\exp(tn(\mu+\varepsilon))}$$

$\exp(\log a^b) = \exp(b\log a)$

$$= \exp(n\log m(t) - tn(\mu+\varepsilon))$$

So we get

$$P(\overline{X}_n \geq \mu+\varepsilon) \leq \exp(-n\,g(t))$$

where $g(t) = t(\mu+\varepsilon) - \log(m(t))$

t is arbitrary $\Rightarrow$ look for $\bar{t}$ to get the best bound

We search for $\bar{t}$ s.t. $g(\bar{t})>0$

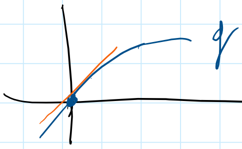
Such t always exists:

Claim $\exists\ \bar{t}>0$ s.t. $g(\bar{t})>0$

Proof $g(0) = -\log m(0) = -\log 1 = 0$

$m(t) = E[e^{tX_1}] \Rightarrow m(0) = 1$

If $g'(0)>0$ then $\exists\ \bar{t}$ s.t. $g(\bar{t})>0$



$$g'(\varepsilon) = \mu + \varepsilon - \frac{m'(\varepsilon)}{m(\varepsilon)}$$

recall $m'(\varepsilon)\big|_{\varepsilon=0} = E[X_1] = \mu$

$$g'(0) = \mu + \varepsilon - \frac{\mu}{1} = \varepsilon > 0$$ $\boxtimes$

UPPER TAIL ESTIMATE

$P(\overline{X}_n \geq \mu + \varepsilon) \leq e^{-n g(\varepsilon)}$ $\forall n$

$\exists\ \bar{\varepsilon}$ small s.t. $g(\bar{\varepsilon}) > 0$