SDS 387
Linear Models

Fall 2025

Lecture 6 - Thu, Sept 18, 2025

Instructor: Prof. Ale Rinaldo

☒ Last time: convergence in distribution or weak convergence
or vague convergence
or weak*-convergence

↓

- focus is on convergence of the <u>distribution</u> of X_n rather than the realizations of X_n.

<u>Example</u> $X, Y \sim$ Bernoulli(p) then $X \overset{d}{=} Y$ so, if $X_n \sim$ Bernoulli(p) and $X \sim$ Bernoulli(p)

$$X_n \xrightarrow{d} X$$

regardless of the joint distribution of the X_n's and X

- Let $X \in \mathbb{R}^d$ be a random vector and $x \in \mathbb{R}^d$. The function

$$x \in \mathbb{R}^d \longrightarrow F_X(x) = \mathbb{P}(X \le x)$$

is the cdf of X, where $\leq$ is intended to hold element-wise [i.e. $x = \begin{bmatrix} x(1) \\ \vdots \\ x(d) \end{bmatrix}$ and $y = \begin{bmatrix} y(1) \\ \vdots \\ y(d) \end{bmatrix}$ $x \leq y$ means $x(j) \leq y(j)$ all $j = 1, \dots, d$]

Properties of cdf:

i) it is non-decreasing (wrt partial order $\leq$)

$\hookrightarrow$ $x \leq y \Rightarrow F(x) \leq F(y)$

ii) it is right-continuous and has left limits

$\lim_{y \downarrow x} F(y) = F(x)$ $\qquad$ $\lim_{y \uparrow x} F(y)$ exists $= F(x^-)$

iii) $\lim_{y \to -\infty} F(y) = 0$ $\qquad$ $\lim_{y \to \infty} F(y) = 1$

When $d = 1$ let consider the interval

$\{x \in \mathbb{R} : a < x \leq b\}$ $\quad$ $A = (a, b]$ $\qquad -\infty < a < b < \infty$

Then $\mathbb{P}(X \in A) = F_X(b) - F_X(a)$

When $d = 1$ any function F with properties i), ii) and iii) defines a prob. distribution over $\mathbb{R}$.

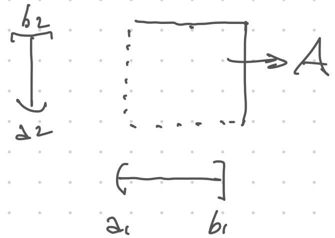
- When $d>1$ this is not the case! Last time we saw a counterexample of a cdf-like function that assigns negative probability to a set.

We need an extra assumption. Let

$$A = \prod_{j=1}^{d} (a_j, b_j] \qquad -\infty < a_j < b_j < \infty \quad \forall j = 1, \dots, d$$

$A \downarrow$ rectangle in d dim.

When $d=2$



Then

$$\mathbb{P}(X \in A) = F_X(b_1, b_2) - F_X(b_1, a_2) - F_X(a_1, b_2) + F_X(a_1, a_2)$$

$\downarrow$

For a cdf-like function satisfying properties i), ii) and iii) this may be negative.

- To avoid this issue, we need one more condition

For $A = \prod_{j=1}^{d} (a_j, b_j]$ let

$$\text{iv)} \quad \Delta_A F = \sum_{v \text{, vertex of } A} \operatorname{sgn}(v)\, F(v) \geq 0 \qquad \forall A \text{ rectangles}$$

$\Delta_A F \leftarrow$ "del of A" of F

where $\operatorname{sgn}(v) = (-1)^{\#\, a_j\text{'s in } v}$

- Properties i), ii), iii) and iv) guarantee that F is the cdf of a prob. distribution

- <u>Def</u>: A sequence of r.v.'s $\{X_n\}$ with cdf's $\{F_{X_n}\}$ converges in distribution to a r.v. X with cdf F_X when

$$\mathbb{P}(X_n \le x) \quad F_{X_n}(x) \longrightarrow F_X(x) = \mathbb{P}(X \le x)$$

pointwise convergence

for all continuity point x of F_X.

↳ necessary requirement

- <u>Theorem</u>: if $X_n \xrightarrow{p} X$ then $X_n \xrightarrow{d} X$

$\Rightarrow$ $\not\Leftarrow$

Pf/ $d=1$. $X_n \xrightarrow{p} X$ means that for any $\varepsilon > 0$ $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$ as $n \to \infty$

↳ small

So for any $c \in \mathbb{R}$,

$$\mathbb{P}(X \le c - \varepsilon) = \mathbb{P}\Big(\underbrace{\{X \le c-\varepsilon\} \cap \{|X_n - X| \le \varepsilon\}}_{\subseteq \{X_n \le c\}}\Big)$$

$$+ \ \mathbb{P}\Big(\underbrace{\{X \le c - \varepsilon\} \cap \{|X_n - X| > \varepsilon\}}_{\subset \{|X_n - X| > \varepsilon\}}\Big)$$

$A \subseteq B$ implies
$\mathbb{P}(A) \le \mathbb{P}(B)$

$$\le \mathbb{P}(X_n \le c) + \mathbb{P}(|X_n - X| > \varepsilon)$$

$$\Downarrow$$

$$(*) \quad \mathbb{P}(X \le c-\varepsilon) - \mathbb{P}(|X_n - X| > \varepsilon) \le \mathbb{P}(X_n \le c)$$

Similarly

$$(**) \quad \mathbb{P}(X_n \le c) \le \mathbb{P}(X \le c+\varepsilon) + \mathbb{P}(|X_n - X| > \varepsilon)$$

Recall that $\liminf_n \mathbb{P}(|X_n - X| > \varepsilon) = 0$

$\limsup_n \mathbb{P}(|X_n - X| > \varepsilon) = 0$

because $X_n \xrightarrow{P} X$. So $(*)$ and $(**)$

$$\underbrace{\mathbb{P}(X \le c-\varepsilon)}_{F_X(c-\varepsilon)} \le \liminf \mathbb{P}(X_n \le c) \le \limsup_n \mathbb{P}(X_n \le c) \le \underbrace{\mathbb{P}(X \le c+\varepsilon)}_{F_X(c+\varepsilon)}$$

Next let $\varepsilon \downarrow 0$ then

$$\underbrace{F_X(c-\varepsilon) \uparrow F_X(c^-)}_{\text{left limit}} \quad \text{and} \quad \underbrace{F_X(c+\varepsilon) \downarrow F_X(c)}_{\text{right continuity}}$$

$$\Downarrow$$

$$F_X(c-) \le \liminf F_{X_n}(c) \le \limsup F_{X_n}(c) \le F_X(c)$$

If c is a continuity point of F_X, then

$F_X(c-) = F_X(c)$, which implies

$$\lim_n F_{X_n}(c) = F_X(c)$$

$\blacksquare$

- Taxonomy of stochastic convergence:

$$\begin{tikzcd}
X_n \xrightarrow{wp1} X \arrow[dr] & & X_n \xrightarrow{L_p} X \arrow[dl] \\
& X_n \xrightarrow{p} X \arrow[d, shift right] & \\
& X_n \xrightarrow{d} X \arrow[u, shift right, "\text{if } X \text{ is degenerate}"'] &
\end{tikzcd}$$

- <u>Theorem</u> If $X_n \xrightarrow{d} c$, $c \in \mathbb{R}^d$,

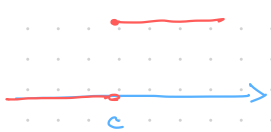
Then $X_n \xrightarrow{p} c$

Pf/ $d=1$. We want to show that

$$\mathbb{P}(|X_n - c| \ge \varepsilon) \to 0 \quad \text{for every } \varepsilon > 0$$
$\hookrightarrow$ small

$$\mathbb{P}(|X_n - c| \ge \varepsilon) \le \underbrace{\mathbb{P}(X_n \le c - \varepsilon)}_{F_{X_n}(c-\varepsilon) \to 0} + \underbrace{\mathbb{P}(X_n \ge c + \varepsilon)}_{1 - \mathbb{P}(X_n < c+\varepsilon)}$$

cdf of c



$$\underbrace{1 - \mathbb{P}(X_n < c+\varepsilon)} \le 1 - \underbrace{\mathbb{P}\left(X_n \le c + \tfrac{\varepsilon}{2}\right)}_{F_{X_n}(c+\varepsilon/2)}$$

$\to 0$

$\to 0$ as $n \to \infty$ ∎

- <u>Remarks</u>: • Suppose $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ Y \end{bmatrix}$. Then we can conclude that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$

  i.e. convergence of the joint implies convergence of the marginals

  • What if we only know that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$?

  Q can we conclude that:

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ Y \end{bmatrix} \ ?$$

No Example: Let $U \sim \text{Uniform}(0,1)$

$X_n = U$ all n and

$$Y_n = \begin{cases} U & n \text{ odd} \\ 1-U & n \text{ even} \end{cases}$$

Then $X_n \xrightarrow{d} U$ $Y_n \xrightarrow{d} U$ but

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \not\xrightarrow{d}$$

- Let F_n be the cdf of Uniform $[-n, n]$

Let $X_n \sim F_n$. Does $X_n \xrightarrow{d}$?

Similarly if

$$X_n = \begin{cases} c_n & \text{wp } 1 \\ 0 & \text{oth} \end{cases} \quad \text{where } c_n \to \infty$$

$X_n \xrightarrow{d}$? No!

$\downarrow$

We need a notion of $\overline{\overline{\text{boundedness}}}$ in probability, to avoid such issues