SDS 387
Linear Models

Fall 2024

Lecture 5 - Tue, Sep 10, 2024

Instructor: Prof. Ale Rinaldo

■ Convergence in distribution (weakest form of stochastic convergence)

A sequence of random vectors in $\mathbb{R}^d$ $\{X_n\}$ converges in distribution (or weakly) to a random vector X when, $\forall x \in \mathbb{R}^d$ s.t. F_X is continuous at x,

F_X: cdf of X

$$F_{X_n}(x) \rightarrow F_X(x) \text{ as } n \rightarrow \infty.$$

F_{X_n}: cdf of X_n

Same:
$X_n \xrightarrow{d} X$
$X_n \Rightarrow X$
$X_n \rightsquigarrow X$

[pointwise convergence of $\{F_{X_n}\}$ at all continuity points of F_X]

Remarks

1) This definition does not impose any restriction on the joint distribution of the X_n's and of X

For example $X_n = (-1)^n Z$ $\quad Z \sim N(0,1)$

1

and $X \sim N(0,1)$ then $X_n \xrightarrow[\text{wp1}]{d} X$

but $X_n \not\xrightarrow{P} X$ nor $X_n \not\xrightarrow{\text{wp1}} X$

ii) The fact that X_n and X have similar cdf's does not mean they take on similar values. In above example $X_n \sim N(0,1)$ and $X \sim N(0,1)$ but in general $|X_n - X|$ can be large

<u>Examples</u> : Let Φ be the cdf of $N(0,1)$ and let

i)
$$F_n(x) = \begin{cases} 0 & x < -n \\ \dfrac{\Phi(x) - \Phi(-n)}{\Phi(n) - \Phi(-n)} & -n \le x < n \\ 1 & x \geq n \end{cases}$$

Then $F_n \to \Phi$ pointwise

ii)
$$F_n(x) = \begin{cases} 0 & x < -n \\ \Phi(x) & -n \le x < n \\ 1 & x \geq n \end{cases}$$

neither continuous nor discrete $F_n \to \Phi$

The restriction that converge takes place at all continuity points of F_X only is necessary!

Example

$$X_n = \begin{cases} \left.\begin{matrix} 1-\frac{1}{n} & \text{wp} & 1/2 \\ 0-\frac{1}{n} & \text{wp} & 1/2 \end{matrix}\right\} & n \text{ odd} \\ \left.\begin{matrix} 1+\frac{1}{n} & \text{wp} & 1/2 \\ 0+\frac{1}{n} & \text{wp} & 1/2 \end{matrix}\right\} & n \text{ even} \end{cases}$$

It is clear that, for large n, $X_n \approx \text{Bernoulli}(1/2)$

So we would like to say that $X_n \xrightarrow{d} \text{Bernoulli}(1/2)$

That is in fact the case. But for $x=1$

$$F_{X_n}(1) = \begin{cases} 1 & n \text{ odd} \\ 1/2 & n \text{ even} \end{cases}$$

↳ $F_{X_n}(1)$ does not converge!!

That is ok: $x=1$ is a point of discontinuity of the cdf Bernoulli $(1/2)$

Result: If $X_n \xrightarrow{P} X$ then $X_n \xrightarrow{d} X$.

Pf/ $X_n \xrightarrow{P} X$ means that, $\forall \varepsilon > 0$ (→ small),

$$\mathbb{P}(|X_n - X| > \varepsilon) \to 0 \text{ as } n \to \infty.$$

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

$$\mathbb{P}(X \le x-\varepsilon) = \mathbb{P}(\overbrace{\{X \le x-\varepsilon\} \cap \{|X_n - X| \le \varepsilon\}}^{\subseteq \{X_n \le x\}}) +$$

$$\mathbb{P}(\underbrace{\{X \le x-\varepsilon\} \cap \{|X_n - X| > \varepsilon\}}_{\subseteq \{|X_n - X| > \varepsilon\}})$$

$P(A) = P(A \cap B) + \mathbb{P}(A \cap B^c)$

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

↳

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

Similarly,

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

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

inequalities (*) and (**) imply

$$\mathbb{P}(X \le x - \varepsilon) \le \liminf_n \mathbb{P}(X_n \le x) \le \limsup_n \mathbb{P}(X_n \le x) \le \mathbb{P}(X \le x + \varepsilon)$$

Let $\varepsilon \downarrow 0$ to get:

as $\varepsilon \downarrow 0$ goes to

as $\varepsilon \downarrow 0$ goes to

$$\underbrace{\mathbb{P}(X < x)}_{F_X(x^-)} \le \liminf_n \mathbb{P}(X_n \le x) \le \limsup_n \mathbb{P}(X_n \le x) \le \underbrace{\mathbb{P}(X \le x)}_{F_X(x)}$$

If x is a continuity point of F_X, $F_X(x^-) = F_X(x)$

so,

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

∎

Schematic of stochastic convergence:

$$\begin{tikzcd}
X_n \xrightarrow{wp\,1} X \arrow[dr] & & X_n \xrightarrow{L_p} X \arrow[dl] \\
& X_n \xrightarrow{p} X \arrow[d] & \\
& X_n \xrightarrow{d} X &
\end{tikzcd}$$

Result: If $X_n \xrightarrow{d} X$ where X is degenerate (non-random)

then $X_n \xrightarrow{p} X$

Pf/ Let call X x [i.e. $\mathbb{P}(X=x)=1$] Then we want to show that $\mathbb{P}(|X_n - x| \geq \varepsilon) \to 0$ $\forall \varepsilon > 0$

Fix $\varepsilon > 0$ ($\to$ small). Then

$$\mathbb{P}(|X_n - x| \geq \varepsilon) \leq \mathbb{P}(X_n \leq x-\varepsilon) + \underbrace{\mathbb{P}(X_n \geq x+\varepsilon)}_{1-\mathbb{P}(X_n < x+\varepsilon)}$$

$$\leq \underbrace{\mathbb{P}(X_n \leq x-\varepsilon)}_{\to 0} + 1 - \underbrace{\mathbb{P}\left(X_n \leq x+\tfrac{\varepsilon}{2}\right)}_{\to 1}$$

$$\to 0 \quad \blacksquare$$

<u>Remarks</u>:

i) If $\begin{bmatrix} X_n \\ Y_n \end{bmatrix} \xrightarrow{d} \begin{bmatrix} X \\ Y \end{bmatrix}$ then we can conclude that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$ HW!

joint weak convergence implies weak convergence of marginals but not the other way around

However, if $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$, in general we cannot conclude that

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

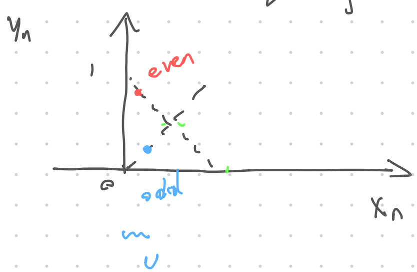
<u>Example</u> Let $U \sim \text{Uniform}(0,1)$

$X_n = U$ all n

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

Clearly $X_n \xrightarrow{d} U$ $\quad Y_n \xrightarrow{d} U$

But $\begin{bmatrix} X_n \\ Y_n \end{bmatrix}$ does not converge at all!



HW

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

Does F_n converge? No!

The issue is that the mass is too

About v): when $d = 1$ take $A = (-\infty, x]$

dimension

then $\mathbb{P}(X_n \in A) = F_n(x)$ and

$$\mathbb{P}(X \in \partial A) = \mathbb{P}(X = x)$$

which is $= 0$ when x is a continuity point of F_X

More generally: X_n is a discret uniform on $\{0, 1, \dots, n-1\}$

Then $\frac{X_n}{n} \xrightarrow{d} U$ where $U \sim$ Uniform $(0,1)$

$A = [0,1] \cap \mathbb{Q}_n$ then $\mathbb{P}(X_n \in A) = 1$

but $\mathbb{P}(U \in A) = 0$

Q: explain why condition v) is not violated!

HW

Continuous mapping theorem:

Let $f: \mathbb{R}^d \to \mathbb{R}$ and let X be a random vector in $\mathbb{R}^d$ s.t. $\mathbb{P}(X \in C) = 1$ where C is the set of continuity points of f. Then

$$X_n \xrightarrow{*} X \implies f(X_n) \xrightarrow{*} f(X)$$

where $*$ means wp 1 or in probability or in distribution.

Note: if f is continuous this is trivially true