CHAPTER 1 :

ALMOST SURE CONVERGENCE
CONVERGENCE IN PROBABILITY.

Goals : • definition of a.s. cv / cv in probability.

• relations between the two types of cv.

Setup : • $(\Omega, \mathcal{F}, P)$ fixed probability space

• (E, d) metric space.

$d : E \times E \longrightarrow \mathbb{R}_+$ with

• $\forall x, y \in E \quad d(x,y) = 0 \iff x = y$

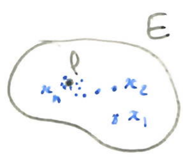
• $\forall x, y \in E \quad d(x,y) = d(y,x)$

• $\forall x, y, z \in E \quad d(x,z) \leq d(x,y) + d(y,z)$

I INTRO / MOTIVATION

In E, we have a notion of convergence : a sequence $(x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$ converges to $\ell \in E$ if

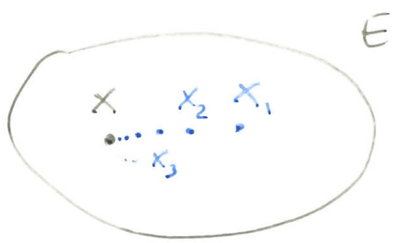
$$\lim_{n \to \infty} d(x_n, \ell) = 0$$



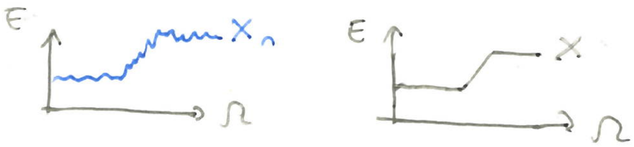
ie.

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \quad \forall n \geq N \quad d(x_n, \ell) \leq \varepsilon$$

How does this notion of convergence extend to random elements of E? Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables with values in E. Intuitively, X_n is a random point in E, and we want to express that X_n is "close to" a point X (possibly random as well).



Formally, $X_n : \Omega \to E$ is a function from Ω to E, and there are several ways to define convergence towards a function $X : \Omega \to E$.



Eg (1) Pointwise convergence: $\forall \omega \in \Omega \quad \lim\limits_{n\to\infty} X_n(\omega) = X(\omega)$

(2) convergence of sup: $\sup\limits_{\omega\in\Omega} \; d(X_n(\omega)), X(\omega)) \underset{n\to\infty}{\longrightarrow} 0$

(3) convergence a.s. (P- a.e.) : $P\left(\{\omega \in \Omega : \lim\limits_{n\to\infty} X_n(\omega) = X(\omega)\}\right) = 1$

(4) convergence in probability: $\forall \varepsilon > 0 \quad P(d(X_n, X) \geq \varepsilon) \underset{n\to\infty}{\longrightarrow} 0$

(5) If $E = \mathbb{R}$ $\to$ cv in L^p $\quad \int_{\mathbb{R}} |X_n - X|^p \, dP \underset{n\to\infty}{\longrightarrow} 0$

...

- ① and ② are not suitable for probability, because random variables are defined a.s. in general.
- In this chapter, we study ③ and ④
- ⑤ will be analyzed in Chapter 3.
- In general, the study of convergence of random variables is related to functional analysis. Giving sense to $X_n \xrightarrow[n\to\infty]{} X$ corresponds to the choice of a space S of functions where the two random variables X_n and X "lives" and equip S with a topology.

2 ALMOST SURE CONVERGENCE

Def. Let $(X_n)_{n\geq 1}$, X be r.v.'s with values in E.

We write

$$\lim_{n\to\infty} X_n = X \quad \text{a.s.}$$

"(X_n) converges towards X a.s."

if $P(\{w : \lim_{n\to\infty} X_n(w) = X(w)\}) = 1$.

Rk: we also write $X_n \xrightarrow[n\to\infty]{a.s.} X$.

Rk: $\{\omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}$ is measurable

because it can be written as

$$\bigcap_{A\geq 1} \bigcup_{N\geq 1} \bigcap_{n\geq N} \left\{\omega : d(X_n(\omega), X(\omega)) < \frac{1}{A}\right\}.$$

- d is continuous on $E\times E$ (because 1-Lipschitz) and we can consider the random variable $d(X_n, X)$. (in $\mathbb{R}$)

We have

$$\underbrace{\left(\lim_{n\to\infty} X_n = X \text{ a.s}\right)}_{\text{a.s cv. in } E} \iff \underbrace{\left(\lim_{n\to\infty} d(X_n, X) = 0 \text{ a.s.}\right)}_{\text{a.s. cv. in } (\mathbb{R}, |\cdot|)}$$

Examples.

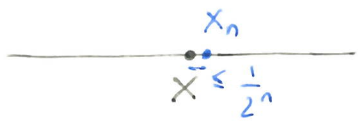
- Let $(a_n)_{n\geq 1}$, ℓ be elements of E, and

$X_n = a_n$ a.s. .

$$\lim_{n\to\infty} X_n = \ell \text{ a.s} \iff \lim_{n\to\infty} a_n = \ell$$

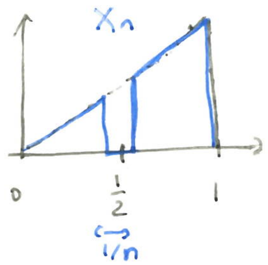
- X real r.v.. $X_n = \min\left(n, \frac{\lfloor 2^n X \rfloor}{2^n}\right)$

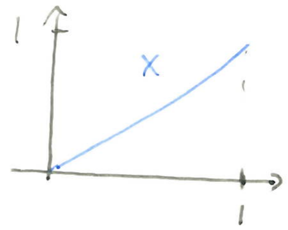
$$\lim_{n\to\infty} X_n = X \text{ a.s.}$$



- $\Omega = [0,1]$ $P = \mathcal{L}$ (Lebesgue). Define

$$X_n(\omega) = \omega \times 1_{|\omega - \frac{1}{2}| \geq \frac{1}{2n}} \qquad X(\omega) = \omega$$





$$\forall \omega \in [0,1] \setminus \{\tfrac{1}{2}\} \qquad \lim_{n\to\infty} X_n(\omega) = X(\omega)$$

Since $P([0,1]\setminus\{\frac{1}{2}\}) = 1$ we have $\lim_{n\to\infty} X_n = X$ a.s.

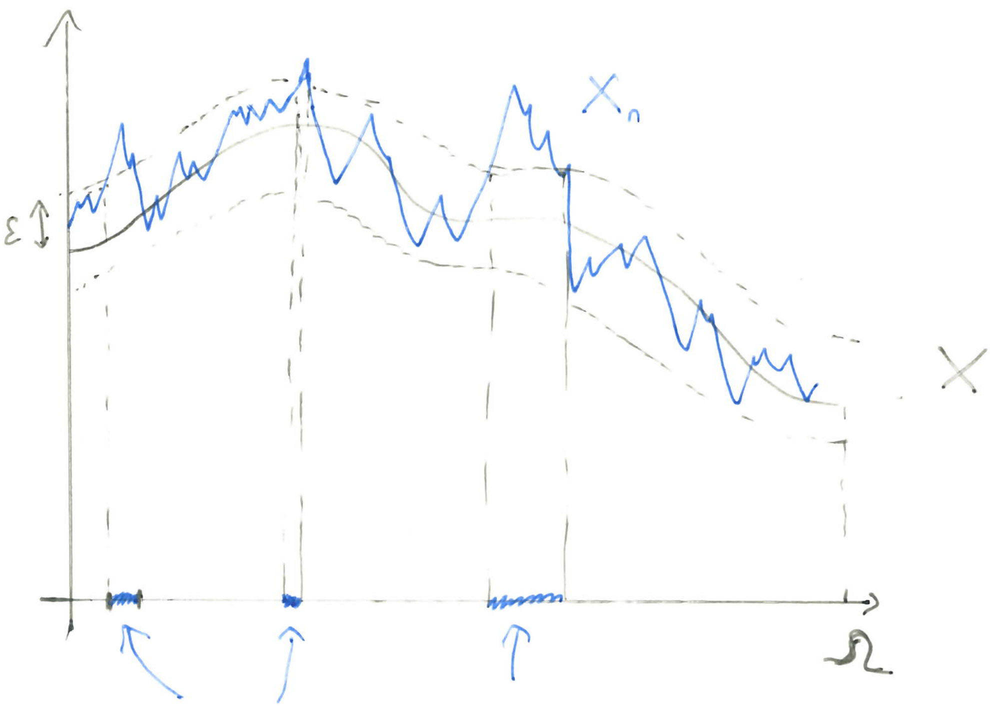
"We do not care what happens when $\omega = \frac{1}{2}$, because it does not happen".

3 CONVERGENCE IN PROBABILITY.

Def. Let $(X_n)_{n\geq 1}$, X r.v. with values in E. We say that $(X_n)_{n\geq 1}$ <u>converges towards X in probability</u> (written $X_n \xrightarrow{P} X$) if

$$\boxed{\forall \varepsilon > 0 \quad \lim_{n\to\infty} P\big(d(X_n, X) > \varepsilon\big) = 0.}$$

Illustration cv. in probability.



$$\{\omega \in \Omega : d(X_n, X) \geq \varepsilon\}$$

The probability of this set is small for n large.

<u>Proposition</u>: Let $(X_n)_{n\geq 1}$, X be rvs with values in E.

If $X_n \xrightarrow[n\to\infty]{a.s} X$ then $X_n \xrightarrow[n\to\infty]{P} X$.

<u>Proof</u>. Let $\varepsilon > 0$. If $X_n \xrightarrow[n\to\infty]{a.s.} X$, then

$$1_{d(X_n,X)\geq\varepsilon} \xrightarrow[n\to\infty]{a.s.} 0 \quad \left(\text{because } d(X_n,X) \xrightarrow[n\to\infty]{a.s} 0\right)$$

By dominated convergence $\left(1_{d(X_n,X)\geq\varepsilon} \leq 1\right)$, we have

$$\lim_{n\to\infty} \underbrace{E\left(1_{d(X_n,X)\geq\varepsilon}\right)}_{P(d(X_n,X)\geq\varepsilon)} = 0 \qquad \blacksquare$$

4 CV IN PROBABILITY DOES NOT IMPLY A.S. CV

<u>Example 1</u> $E = \mathbb{R}$ $(d = |\cdot|)$

Let $(X_n)_{n\geq 1}$ be independent rvs with

$$X_n \sim \mathrm{Ber}\left(\frac{1}{n}\right)$$

For every $\varepsilon > 0$, we have

$$P(|X_n| \geq \varepsilon) = P(X_n = 1) = \frac{1}{n} \xrightarrow[n\to\infty]{} 0$$

Hence $X_n \xrightarrow[n\to\infty]{P} 0$.

We have $\sum_{n \geq 1} P(X_n = 1) = \sum_{n \geq 1} \frac{1}{n} = +\infty$

Since the events $\{X_n = 1\}$ are independent, by Borel-Cantelli Lemma II, we have $\limsup_{n \to \infty} X_n = 1$ a.s. and therefore

$$X_n \not\xrightarrow[n \to \infty]{a.s} 0 .$$

"most of the time we have $X_n = 0$ but sometimes $X_n = 1$"

$$(X_n)_{n \geq 1} = 0100100000000010000000010000000000000 \cdots$$

Example 2. "typesetter example"

$\Omega = [0,1]$ $\quad \mathcal{F} = \mathcal{B}([0,1])$ $\quad P = \mathcal{L}$

$X_1 = 1$

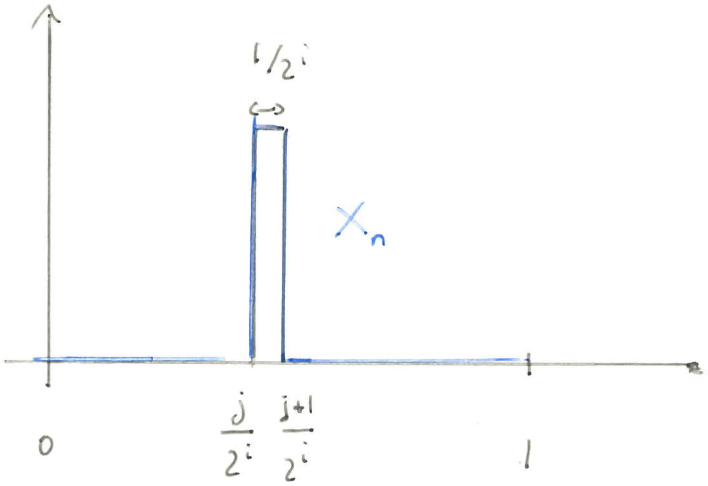
$X_2 = 1_{[0, \frac{1}{2}]}$ $\quad X_3 = 1_{[\frac{1}{2}, 1]}$

$X_4 = 1_{[0, \frac{1}{4}]}$ $\quad X_5 = 1_{[\frac{1}{4}, \frac{1}{2}]}$ $\quad X_6 = 1_{[\frac{1}{2}, \frac{3}{4}]}$ $\quad X_7 = 1_{[\frac{3}{4}, 1]}$

$\vdots$

$X_{2^i} = 1_{[0, \frac{1}{2^i}]}$ $\quad \cdots \quad X_{2^i + j} = 1_{[\frac{j}{2^i}, \frac{j+1}{2^i}]}$ $\quad \cdots \quad X_{2^{i+1}-1} = 1_{[1-\frac{1}{2^i}, 1]}$

$\vdots$



$n = 2^i + j$

$\forall \varepsilon \quad P(|X_n| \geq \varepsilon) \leq \frac{2}{n} \xrightarrow[n\to\infty]{} 0$

But $\forall\, \omega \in [0,1] \quad X_n(\omega) \not\xrightarrow[n\to\infty]{} 0$ (it takes the value 1 infinitely often)

$X_n \xrightarrow[n\to\infty]{P} 0$ but $X_n \not\xrightarrow[n\to\infty]{a.s} 0$.

5 CRITERION FOR A.S. CV.

Proposition. Let (X_n), X r.v.s with values in E. If

$$\forall \varepsilon > 0 \quad \sum_{n \geq 1} P(d(X_n, X) \geq \varepsilon) < \infty$$

Then $X_n \xrightarrow[n\to\infty]{a.s} X$.

Proof. By Borel - Cantelli Lemma I,

$$\forall k \geq 1 \quad \limsup_{n\to\infty} |X_n - X| < \frac{1}{k} \quad \text{a.s.}$$

(because $\{ |X_n - X| \geq \frac{1}{k} \}$ occurs for finitely many n a.s.)

It follows that $\limsup_{n\to\infty} |X_n - X| = 0$ a.s.

$$\left(\{ \limsup_{n\to\infty} |X_n - X| = 0 \} = \bigcap_{k \geq 1} \{ \limsup |X_n - X| < \frac{1}{k} \} \right)$$

Application $(U_n)_{n \geq 1}$ iid $\mathcal{U}([0,1])$

$$X_n := \min(U_1, \dots, U_n)$$

$$\lim_{n\to\infty} X_n = 0 \quad \text{a.s.}$$

Proof. Let $\varepsilon > 0$. $P(|X_n| \geq \varepsilon) = P(U_1 \geq \varepsilon)^n = (1-\varepsilon)^n$.

Hence $\sum_{n\to\infty} P(|X_n - 0| \geq \varepsilon) < \infty$ ∎

6. CHARACTERISATION OF CV IN PROBABILITY.

Thm. Let (X_n), X r.v with values in E

$$\left(X_n \xrightarrow[n\to\infty]{P} X\right) \iff \left(\lim_{n\to\infty} \mathbf{E}\left(d(X_n,X)\wedge 1\right) = 0\right)$$

Rk: If X, Y are r.v., $d(X,Y)\wedge 1$ is a real random variable with values in $[0,1]$. Hence $E(d(X,Y)\wedge 1)$ is well defined and lies in $[0,1]$.

Proof. $\boxed{\Rightarrow}$ Let $\varepsilon \in (0,1)$, $n \geq 1$

$$E\left(d(X_n,X)\wedge 1\right) = E\left((d(X_n,X)\wedge 1)\,1_{d(X_n,X)<\varepsilon}\right) + E\left((d(X_n,X)\wedge 1)\,1_{d(X_n,X)\geq\varepsilon}\right)$$

$$\leq E\left(\varepsilon\, 1_{d(X_n,X)<\varepsilon}\right) + E\left(1_{d(X_n,X)\geq\varepsilon}\right)$$

$$\leq \varepsilon + P\left(d(X_n,X)\geq\varepsilon\right)$$

Hence $\limsup_{n\to\infty} E\left(d(X_n,X)\wedge 1\right) \leq \varepsilon$.

$\boxed{\Leftarrow}$ Let $\varepsilon \in (0,1]$,

$$P\left(d(X_n,X)\geq\varepsilon\right) = P\left(1\wedge d(X_n,X)\geq\varepsilon\right)$$

$$\overset{\text{Markov}}{\leq} \frac{1}{\varepsilon} E\left(1\wedge d(X_n,X)\right) \xrightarrow[n\to\infty]{} 0$$

If $\varepsilon \geq 1$ we have $P(d(X_n,X)\geq\varepsilon) \leq P(d(X_n,X)\geq 1) \xrightarrow[n\to\infty]{} 0$ ■

7. CONVERGING SUBSEQUENCE

> **Prop.** Let $(X_n)_{n\geq 1}$, X r.v.s with values in E.
> If $(X_n)_{n\geq 1}$ cv to X in probability, then there is a subsequence $(X_{n(k)})_{k\geq 1}$ that cv a.s. to X.

Proof. Assume $X_n \xrightarrow[n\to\infty]{P} X$. Since $E(1 \wedge d(X_n, X)) \xrightarrow[n\to\infty]{} 0$,

we can construct $n(1) < n(2) < \dots$ such that

$$\forall k \in \mathbb{N} \quad E(1 \wedge d(X_{n(k)}, X)) \leq \frac{1}{2^k}.$$

For every $\varepsilon > 0$ we have

$$P(d(X_{n(k)}, X) \geq \varepsilon) \overset{\text{Markov}}{\leq} \frac{1}{\varepsilon} \times \frac{1}{2^k}.$$

Therefore $\sum_{k\geq 1} P(d(X_{n(k)}, X) \geq \varepsilon) < \infty$, which

implies that $X_{n(k)} \xrightarrow[k\to\infty]{a.s} X$. ∎

Example: . For the typesetter example, X_{2^k} converges to 0 a.s. (exercise)

. For the Bernoulli example X_{k^2} converges to 0 a.s. (exercise)

CCL: Relation between a.s cv and cv in probability