CHAPTER 3 :

KOLMOGOROV 0-1 LAW

Goals. :
- Independence reminders
- Importance of abstract theory :
  The elements of $\sigma(X_1, X_2, \dots)$ can be approximated by elements of $\sigma(X_1, \dots, X_n)$ for n large.
- Notion of tail-events.

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

. $(E, \mathcal{E})$ measured space.

We will consider general r.v.s $X : \Omega \to E$

1 TAIL σ-ALGEBRA.

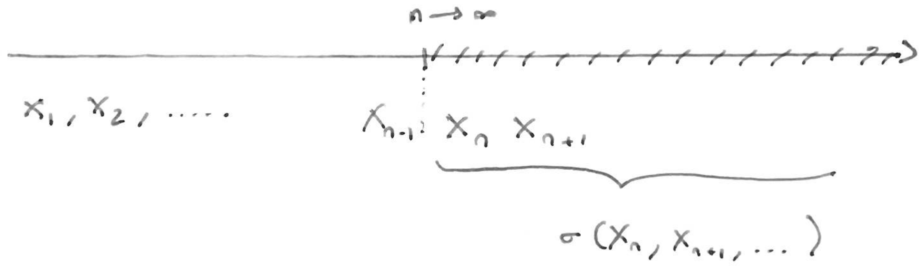
Def. Let $X_1, X_2, \dots$ r.v.s with values in E.

The tail σ-algebra associated to $X_1, X_2, \dots$ is defined by

$$\mathcal{T} = \bigcap_{n \geq 1} \sigma(X_n, X_{n+1}, \dots).$$

Rk: $\mathcal{T}$ is indeed a σ-algebra (as an intersection of σ-algebras)

Why "tail" ?



The tail σ-algebra contains the events A whose occurrence does not depend on any fixed finite subfamily of the X_i's.
To put it differently, for any n, we can change the values of $X_1, \ldots, X_n$ arbitrarily without changing whether A occurs or not.

Examples : $\emptyset, \Omega \in \mathcal{T}$

Are there interesting events in $\mathcal{T}$?

→ E = metric space (E, d), d distance.

Examples :

- $\{(X_n)_{n \geq 1}$ is a Cauchy sequence$\} \in \mathcal{T}$.
- $\{(X_n)_{n \geq 1}$ converges $\} \in \mathcal{T}$.
- $\{(X_n)_{n \geq 1}$ is bounded$\} \in \mathcal{T}$.

$$= \{\exists C < \infty \quad \forall n \geq 1 \quad d(x, X_n) \leq C\} \text{ where } x \text{ fixed in } E.$$

- $\{\forall n \geq 1 \quad d(x, X_n) \leq 10\} \notin \mathcal{T}$.

$\underline{E = \mathbb{R}}$ $\{\limsup X_n \leq 10\} \in \mathcal{T}$,

$\{\sup X_n \leq 10\} \notin \mathcal{T}$,

$\{\sup X_n < \infty\} \in \mathcal{T}$,

$\{\sum X_n \text{ cv}\} \in \mathcal{T}$.

2 KOLMOGOROV THEOREM.

Thm: Let $X_1, X_2, \dots$ be *independent* random variables with values in E. Let $\mathcal{T}$ be the associated tail σ-algebra.

$$\boxed{\forall A \in \mathcal{T} \quad P(A) \in \{0,1\}.}$$

"$X_1, X_2, \dots$ is tail-trivial"

Proof idea $A \in \mathcal{T}$

$X_1, \dots$ $\quad n >> 1 \quad X_n, \dots$

A "$\in$" $\sigma(X_1, X_2, \dots, X_n)$ for n large because $A \in \sigma\left(\bigcup_n \sigma(X_1, \dots X_n)\right)$

$A \in \sigma(X_{n+1}, \dots$ because $A \in \mathcal{T}$

$$P(A \cap A) = P(A) \times P(A)$$

Rk: Let $\mathcal{C} \subset \mathcal{F}$ be a class of sets.

$$\mathcal{C} \text{ independent of } \mathcal{C} \Rightarrow \forall A \in \mathcal{C} \quad \overbrace{P(A \cap A)}^{=P(A)} = P(A)^2$$

$$\Rightarrow \forall A \in \mathcal{C} \quad P(A) = P(A)^2$$

$\left(\text{if } x \in \mathbb{R} \qquad x^2 = x \iff x(1-x) = 0 \iff x \in \{0,1\} \right)$

Proof. Let $\mathcal{D}_n = \sigma(X_1, \dots, X_n)$, $n \geq 1$

By grouping, we have

$$\forall n \geq 1 \quad \mathcal{D}_n \text{ indep. of } \sigma(X_{n+1}, \dots)$$

Since $\mathcal{T} \subset \sigma(X_{n+1}, \dots)$, we deduce

$$\forall n \geq 1 \quad \mathcal{D}_n \text{ indep. of } \mathcal{T},$$

and therefore

$$\bigcup_{n \geq 1} \mathcal{D}_n \quad \text{indep. of } \mathcal{T}.$$

Both $\bigcup_{n \geq 1} \mathcal{D}_n$ and $\mathcal{T}$ are π-systems containing $\emptyset, \Omega$,

hence

$$\sigma\left(\bigcup_{n \geq 1} \mathcal{D}_n\right) \text{ indep of } \mathcal{T}.$$

Since $\mathcal{T} \subset \sigma\left(\bigcup_{n \geq 1} \mathcal{D}_n\right)$, we deduce that

$$\mathcal{T} \text{ indep. of } \mathcal{T} \qquad \blacksquare$$

Rk: independence is important. Consider

$X \sim Ber\left(\frac{1}{2}\right)$ $\quad X_n = X$ for every n.

$A = \{ \exists n_0 : \forall n \geq n_0 \; X_n = 1 \} \in \mathcal{T}$

but $P(A) = P(X = 1) = \frac{1}{2}$.

3 ALTERNATIVE APPROACH: CYLINDRICAL APPROXIMATION.

Prop. Let $X_1, X_2, \ldots$ be arbitrary r.v. with values in E.

For every $A \in \sigma(X_1, X_2, \ldots)$, for every $\varepsilon > 0$,

$$\exists n \geq 1 \; \exists A_\varepsilon \in \sigma(X_1, \ldots, X_n) \quad P(A \Delta A_\varepsilon) \leq \varepsilon$$

($A \Delta B := A \setminus B \cup B \setminus A$ symmetric difference)

Rk:
- It is a particular case of a more general result in measure theory. It is important here that P is σ-finite (because it is finite).
- It can be difficult to represent the elements of $\sigma(X_1, X_2, \ldots)$. The proposition above helps to understands such events: they are the events that can be well approximated by events of $\sigma(X_1, \ldots, X_n)$ for n large.

<u>Proof</u>. Let $\mathcal{C} = \bigcup_{n=1}^{\infty} \sigma(X_1, \dots, X_n)$.

NB: . $\mathcal{C}$ is stable under complement and finite union.

. $\sigma(\mathcal{C}) = \sigma(X_1, X_2, \dots)$

Define $\mathcal{G} = \{ A \in \mathcal{F} : \forall \varepsilon > 0 \ \exists A_\varepsilon \in \mathcal{C} \ P(A \Delta A_\varepsilon) \leq \varepsilon \}$

We have $\mathcal{C} \subset \mathcal{G}$ and $\mathcal{G}$ is a σ-algebra. Indeed.

- $A \in \mathcal{G} \Rightarrow A^c \in \mathcal{G}$ (because $A^c \Delta D^c = A \Delta D \ \forall A, D \in \mathcal{F}$)

- Let $A_1, A_2, \dots \in \mathcal{G}$. We claim that $A = \bigcup_{i=1}^{\infty} A_i$ is also in $\mathcal{G}$. Let $\varepsilon > 0$.

  By continuity of P, we can fix N s.t.

$$P\left(A \setminus \bigcup_{i=1}^{N} A_i\right) \leq \frac{\varepsilon}{2}.$$

  For $i = 1, \dots, N$, let $A_{\varepsilon,i} \in \mathcal{C}$ s.t. $P(A_i \Delta A_{\varepsilon,i}) \leq \frac{\varepsilon}{2^i}$.

  For $i > N$ let $A_{\varepsilon,i} = \emptyset$. Since $A_\varepsilon = \bigcup_{i \geq 1} A_{\varepsilon,i}$ is a finite union of elements of $\mathcal{C}$, we have $A_\varepsilon \in \mathcal{C}$ and

$$P(A \Delta A_\varepsilon) \overset{\circledast}{\leq} P\left(\left(\bigcup_{i=1}^{N} A_i \Delta A_{\varepsilon,i}\right) \cup \bigcup_{i>N} A_i \right)$$

$$\leq \underbrace{\sum_{i=1}^{N} \frac{\varepsilon}{2^i}}_{\leq \frac{\varepsilon}{2}} + \underbrace{P\left(\bigcup_{i>N} A_i\right)}_{\leq \frac{\varepsilon}{2}} \leq \varepsilon$$

Therefore, $\sigma(\mathcal{E}) \subset \mathcal{G}$, which concludes the proof. ∎

(*) Here we use $\forall I \quad \forall (A_i)_{i \in I} \ \forall (B_i)_{i \in I}$

$$\left(\bigcup_{i \in I} A_i\right) \Delta \left(\bigcup_{i \in I} B_i\right) \subset \bigcup_{i \in I} (A_i \Delta B_i). \qquad \text{(exercise)}$$

(if $x \in \left(\bigcup_i A_i\right) \setminus \left(\bigcup_i B_i\right)$, then $\exists i$ s.t. $x \in A_i \setminus B_i$)

We can now use this approximation result to deduce the Thm:

<u>Second proof of the Kolmogorov 0-1 law:</u>

Let $A \in \mathcal{T}$. Let $\varepsilon > 0$.

Since $A \in \sigma(X_1, X_2, \dots)$, there exists $n \geq 1$ and $A_\varepsilon \in \sigma(X_1, \dots, X_n)$ such that

$$P(A \Delta A_\varepsilon) \leq \varepsilon.$$

We have $A \in \sigma(X_{n+1}, \dots)$. Hence, by grouping,

$$P(A_\varepsilon \cap A) = P(A_\varepsilon)\, P(A)$$

Since $A_\varepsilon \subset A \cup (A_\varepsilon \setminus A) \subset A \cup (A \Delta A_\varepsilon)$ and we have $P(A_\varepsilon) \leq P(A) + \varepsilon$. Similarly, $P(A \cap A) \leq P(A_\varepsilon \cap A) + \varepsilon$.

Hence $\quad P(A) - \varepsilon \leq P(A)^2 + \varepsilon$.

Since ε is arbitrary, we conclude $P(A) \in \{0,1\}$ ∎

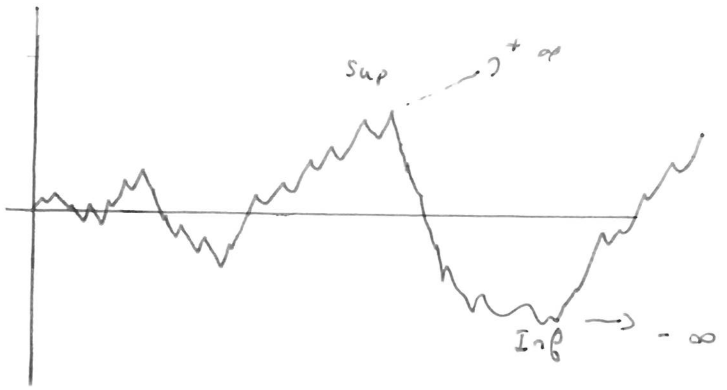
4 APPLICATION : RECURRENCE OF 1D RANDOM WALK.

Prop. Let $(X_n)_{n\geq 1}$ be a sequence of iid r.v.s in $\{-1,1\}$ s.t.

$$\forall n \quad P(X_n = -1) = P(X_n = +1) = \frac{1}{2} .$$

Setting $S_n = X_1 + \cdots + X_n$ for every $n \geq 1$, we have

$$\sup S_n = +\infty \text{ a.s and } \inf S_n = -\infty \text{ a.s.}$$



Corollary: Almost surely, we have $S_n = 0$ for infinitely many n.

(The corollary is left as an exercise)

Proof of the proposition :

Step 1: We first prove

$$\forall a \in \mathbb{N} \quad P(-a \leq \operatorname{Inf} S_n \leq \operatorname{Sup} S_n \leq a) = 0 . \quad (*)$$

Let $a \in \mathbb{N}$. By Borel Cantelli II, for $b > 2a+1$

$$P\left(\bigcup_{j \in \mathbb{N}} \{X_{jb+1} = +1, \dots X_{jb+b} = 1\}\right) = 1$$

Indeed, the events in the union are independent and

$$\sum_{j \in \mathbb{N}} \underbrace{P(X_{jb+1} = 1, \dots, X_{jb+b} = 1)}_{= \frac{1}{2^b} > 0 \text{ indep. of } j} = +\infty .$$

Since $\bigcup_{j \in \mathbb{N}} \{X_{jb+1} = 1, \dots, X_{jb+b} = 1\} \subset \{-a \leq \operatorname{Inf} S_n \leq \operatorname{Sup} S_n \leq a\}^c$,

we obtain $(*)$.

Step 2 Letting a tend to infinity in $(*)$ and taking the complement, we get

$$\begin{aligned} 1 &= P(\{\operatorname{Inf} S_n = -\infty\} \cup \{\operatorname{Sup} S_n = +\infty\}) \\ &\leq P(\operatorname{Inf} S_n = -\infty) + P(\operatorname{Sup} S_n = +\infty) \\ &\underset{\text{Sym.}}{=} 2 \times P(\operatorname{Sup} S_n = +\infty) \end{aligned}$$

Therefore $P(\sup S_n = +\infty) > 0$.

Since $\{\sup S_n = +\infty\} \in \mathcal{T}$

(indeed $\forall k \in \mathbb{N}$ $\{\sup S_n = +\infty\} = \{\sup_{n \geq k} X_k + \dots + X_n = +\infty\}$
$\in \sigma(X_k, \dots)$),

by the 0-1 law, we have

$$P(\sup S_n = +\infty) = 1,$$

and by symmetry $P(\operatorname{Inf} S_n = +\infty) = 1$. $\blacksquare$