Performance Evaluation and Networks

Refresher course in Probability

General framework : random experiments

Experiments & Randomness :

- System/Système (general meaning)
- Experiment/trial/expérience/tirage/épreuve
- Outcome/résultat/éventualité/réalisation
- Event/événement = set of outcomes, which is interesting and measurable

Usual working hypothesis :

- only information = list Ω of outcomes and a measure of their occurences via the measures of events.
- sometimes, no precise information about Ω, work focused on a reduced set of events of known measures.

General framework : random experiments

Experiments & Randomness :

- System/Système (general meaning) on which one can do
- Experiment/trial/expérience/tirage/épreuve which produces
- Outcome/résultat/éventualité/réalisation unknown a priori
- Event/événement = set of outcomes, which is interesting and measurable

Usual working hypothesis :

- only information = list Ω of outcomes and a measure of their occurences via the measures of events.
- sometimes, no precise information about Ω, work focused on a reduced set of events of known measures.

Formalisation

Definition (espace probabilisé/de probabilités/probability space)

A probability space is a triplet $(\Omega, \mathscr{F}, \mathbb{P})$ where :

- Ω set called univers/sample space
- $\mathscr{F}$ set of subsets of Ω (the "events")
- $\mathbb{P}$ function from $\mathscr{F}$ to $\mathbb{R}$

satisfying the following properties :

- $\mathscr{F}$ tribu/σ-algèbre/σ-field :
  - $\Omega \in \mathscr{F}$ and $\forall A \in \mathscr{F}$, $\overline{A} \in \mathscr{F}$
  - $\forall \{A_n, n \geq 0\}$ finite or countable family from $\mathscr{F}$, $\cup_{n \geq 0} A_n \in \mathscr{F}$.
- $\mathbb{P}$ probability measure :
  - $\mathbb{P}(\Omega) = 1$ and $\forall A \in \mathscr{F}$, $0 \leq \mathbb{P}(A) \leq 1$
  - for any finite or countable union of events $A_n \in \mathscr{F}$ pairwise disjoint, $\mathbb{P}(\cup_n A_n) = \sum_n \mathbb{P}(A_n)$.

Vocabulary & first remarks

Vocabulary :

- $\omega \in \Omega$ called outcome/réalisation.
- $A \in \mathscr{F}$ called event/événement.
- ω is a realisation/réalise/realizes A if $\omega \in A$.
- Almost sure event : $A \in \mathscr{F}$ such that $\mathbb{P}(A) = 1$
- Negligible event : $A \in \mathscr{F}$ such that $\mathbb{P}(A) = 0$

Remarks :

- Events of interest are usually defined in extenso (list of elements ω) or by properties
- Axioms $\Rightarrow$ $\emptyset \in \mathscr{F}$ and $\mathbb{P}(\emptyset) = 0$

⚠ you may encounter almost sure events (resp. negligible) different from Ω (resp. $\emptyset$)

Examples of probability spaces

Examples of σ-algebra ?

Examples of probability measures ?

Examples of probability spaces

Examples of σ-algebra :

- any Ω and $\mathscr{F} = \mathscr{P}(\Omega)$
- Borelian sets $\mathscr{B}(\mathbb{R})$: smallest σ-algebra on $\mathbb{R}$ containing open intervals.

Examples of probability measures :

- Case where $\mathscr{P}(\Omega)$ with finite or countable Ω : any function from Ω to $\mathbb{R}_+$ whose sum over Ω is 1 can be extended in a probability measure over $\mathscr{P}(\Omega)$.
- Case of borelian sets $\mathscr{B}(\mathbb{R})$: Lebesgue measure (1901).

⚠ Some borelian sets can not be obtained by a finite number of countable unions / intersections of open intervals.

First properties

Proposition (complementary)

$$\forall A \in \mathcal{F},\ \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$$

Proposition (inclusion)

$$\forall A, B \in \mathcal{F},\ \textit{si } A \subseteq B \textit{ alors } \mathbb{P}(A) \leq \mathbb{P}(B)$$

Proposition (inclusion/exclusion)

$$\forall A, B \in \mathcal{F},\ \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

Proposition (generalised inclusion/exclusion)

$\forall A_1, \dots, A_n \in \mathcal{F},\ \mathbb{P}(\cup_{i=1}^{n} A_i) = \sum_i \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} \mathbb{P}(A_1 \cap \dots \cap A_n)$

First properties

Proposition (sub-additivity)

$\forall \{A_n, n \in \mathbb{N}\}$ *family from* $\mathscr{F}$, $\mathbb{P}(\cup_{n\in\mathbb{N}} A_n) \leq \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$

Proposition (continuity)

Let A_n, $n \geq 0$ *be a sequence in* $\mathscr{F}$ *such that* $A_0 \subseteq A_1 \subseteq \cdots \subseteq A_n \subseteq A_{n+1} \subseteq \cdots$, *let us denote* $A = \cup_{n\geq 0} A_n$ *its limit, we have* $\mathbb{P}(A) = \lim_{n\to+\infty} \mathbb{P}(A_n)$.

Proposition (law of total probabilities)

Let $A \in \mathscr{F}$ *and* $\{B_n, n \geq 0\}$ *a finite or countable family from* $\mathscr{F}$ *which partitions* Ω, *then* $\mathbb{P}(A) = \sum_{n\geq 0} \mathbb{P}(A \cap B_n)$.

Dependance between events

Definition (conditional probabilities)

Let $A, B \in \mathscr{F}$ with $\mathbb{P}(B) > 0$, the probability of A knowing/given B is defined by $\mathbb{P}(A|B) \stackrel{\text{def}}{=} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$.

Definition (independance of evenments)

- $A, B \in \mathscr{F}$ are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$
- $\{A_n, n \in \mathbb{N}\}$ is a family of independent events of $\mathscr{F}$ if for all $I \subseteq \mathbb{N}$ finite, $\mathbb{P}(\cap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i)$.

Proposition (law of total probabilities, conditional version)

Let $A \in \mathscr{F}$ and $\{B_n, n \geq 0\}$ be a finite or countable family finie of $\mathscr{F}$ which partitions Ω, then $\mathbb{P}(A) = \sum_{n \geq 0} \mathbb{P}(A|B_n)\mathbb{P}(B_n)$, with the convention that if $\mathbb{P}(B_n) = 0$, the corresponding term is 0.

Random variables (r.v.)

Definition (general r.v.)

A random variable with values in E is a function X from Ω to E, where $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space and E is equipped with the σ-algebra $\mathscr{B}$, such that $\forall B \in \mathscr{B}$,
$\{X \in B\} \stackrel{\text{def}}{=} \{\omega \in \Omega | X(\omega) \in B\} = X^{-1}(B) \in \mathscr{F}$.

Definition (real r.v.)

A real random variable is a r.v. X from Ω to $\mathbb{R}$ equipped with borelians, that is $\forall x \in \mathbb{R}$,
$\{X \leq x\} \stackrel{\text{def}}{=} \{\omega \in \Omega | X(\omega) \leq x\} = X^{-1}(]-\infty, x]) \in \mathscr{F}$.

Fonction de répartition / cumulative distribution function

Definition (cumulative distribution function of a real r.v.)

The cumulative distribution function for the r.v. X is the function F_x from $\mathbb{R}$ to $[0,1]$ defined by $F_X(x) = \mathbb{P}(X \leq x)$.

Proposition (regularity of cumulative distrib. functions for real r.v.)

F cumulative distribution function if and only if :

- $\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to+\infty} F(x) = 1$,
- *F non decreasing,*
- *F right continuous ($\forall x \in \mathbb{R}$, $\lim_{h\to0,h>0} F(x+h) = F(x)$).*

Connaissance de $F_X \rightarrow \mathbb{P}(X > x)$, $\mathbb{P}(a < X \leq b)$, $\mathbb{P}(X = x)$, ...

Discrete / continuous real random variable

Definition (discrete real r.v.)

A real r.v. X is called discrete if it gets its values from a finite or countable set $\{x_n, n \geq 0\}$ in $\mathbb{R}$. The function $f(x) = \mathbb{P}(X = x)$ is called mass/law/distribution (discret).

Definition (continuous real r.v.)

A real r.v. X is called continuous if its cumulative distrib function F satisties $F(x) = \int_{-\infty}^{x} f(u)du$ where f from $\mathbb{R}$ dans $[0, +\infty[$ is integrable, f is called density/loi/distribution (continuous).

Remarques : Let X a real r.v.,

- If X discrete, $f(x) = \mathbb{P}(X = x)$ fully characterizes F.
- If X continuous, F is continuous and $\forall x \in \mathbb{R}$, $\mathbb{P}(X = x) = 0$.
- There exists other types of real r.v. (singular, some mixes ...)

Discrete / continuous real random variable

Proposition (usual utilisation of laws)

Let X real r.v. discrete/continuous of mass/density f, then for any borelian B of $\mathbb{R}$,

- $\mathbb{P}(X \in B) = \sum_{x \in B} f(x)$ *in the discrete case,*
- $\mathbb{P}(X \in B) = \int_{x \in B} f(x)dx$ *in the continuous case.*

Remark : those formulas also apply to random vectors $X = (X_1, \ldots, X_n)$ with B borelian of $\mathbb{R}^n$, by putting multiple sums/integrals (cf next slides about random vectors).

Random vector & joint distribution

Definition (cumulative distribution fct of a random vector)

Let $X_1,\dots,X_n$ be real r.v. over the same set Ω, the cumulative distrib fct of vector $X = (X_1,\dots,X_n)$ is defined from $\mathbb{R}^n$ to $\mathbb{R}$ by $F(x_1,\dots,x_n) \stackrel{\text{def}}{=} \mathbb{P}(X_1 \leq x_1,\dots,X_n \leq x_n)$.

Definition (discrete joint distribution)

If X takes a finite or countable nb of values, F is characterized by its joint distribution $f(x_1,\dots,x_n) \stackrel{\text{def}}{=} \mathbb{P}(X_1 = x_1,\dots,X_n = x_n)$.

Definition (continuous joint distribution)

The r.v. $X_1,\dots,X_n$ are said conjointly continuous if it exists f from $\mathbb{R}^n$ to $\mathbb{R}$, integrable and called joint distribution, such that $F(x_1,\dots,x_n) = \int_{u_1=-\infty}^{x_1}\cdots\int_{u_n=-\infty}^{x_n} f(u_1,\dots,u_n)du_1\dots du_n$.

Independence of r.v.

Definition (independence of r.v.)

$X_1, \ldots, X_n$ real r.v. over the same Ω are said independent if the cumulative distrib fct of the vector satisfies $\forall x_1, \ldots, x_n$, $F(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n)$ with the marginal distrib $F_{X_i}(x_i) \stackrel{\text{def}}{=} \mathbb{P}(X_i \leq x_i) = F(\infty, \ldots, x_i, \ldots, \infty)$.

Proposition (independence for discrete/continuous cases)

$X_1, \ldots, X_n$ real discrete/continuous r.v. over the same Ω with masses/densities $f_1, \ldots, f_n$ are independent iff $\forall x_1, \ldots, x_n$, the joint distribution satisfies $f(x_1, \ldots, x_n) = f_1(x_1) \cdots f_n(x_n)$ (at pts where $F_{(X_1, \ldots, X_n)}$ differentiable in the continuous case).

Espérance / moyenne / expectation / mean

Definition (expectation of a discrete real r.v.)

The expectation of a discrete real r.v. X of mass f is

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \sum_{x \in \mathbb{R}} x f(x)$$

(finite or countable nb fini of non null terms) on condition that this sum is absolutely convergent (i.e. $\sum_{x \in \mathbb{R}} |x f(x)| < +\infty$).

Definition (expectation of a continuous real r.v.)

The expectation of a continuous real r.v. X of density f is

$$\mathbb{E}(X) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} x f(x) dx$$

on condition that this integral is Lebesgue integrable (i.e. $\int_{-\infty}^{+\infty} |x f(x)| dx < +\infty$).

Expectation & composition of functions

Proposition (composition for discrete real r.v.)

Let X discrete r.v. of mass f, and g function from $\mathbb{R}$ to $\mathbb{R}$, then $Y = g(X)$ is a discrete real r.v. and $\mathbb{E}(g(X)) = \sum_x g(x)f(x)$, on condition that this sum is absolutely convergent.

Proposition (composition for continuous real r.v.)

Let X continuous r.v. of density f, and g function from $\mathbb{R}$ to $\mathbb{R}$ such that $Y = g(X)$ is a continuous r.v., then $\mathbb{E}(g(X)) = \int_x g(x)f(x)dx$, on condition that it is Lebesgue integrable.

Useful formulas to compute $\mathbb{E}(Y)$ without knowing the discrete or continuous law f_Y of Y ("Law of the Unconscious Statistician")

Expectation & composition for random vectors

Proposition (composition for discrete joint distribution)

Let $X = (X_1, \dots, X_n)$ r.v. of discrete joint distrib f, and g function from $\mathbb{R}^n$ to $\mathbb{R}$, then $Y = g(X)$ is a discrete r.v. and $\mathbb{E}(g(X)) = \sum_{x_1} \cdots \sum_{x_n} g(x_1, \dots, x_n) f(x_1, \dots, x_n)$, on condition that this sum is absolutely convergent.

Proposition (composition for continuous joint distribution)

Let $X = (X_1, \dots, X_n)$ r.v. of continuous joint distrib f, and g function from $\mathbb{R}^n$ to $\mathbb{R}$ such that $Y = g(X)$ is a continuous r.v., then $\mathbb{E}(g(X)) = \int_{x_1} \cdots \int_{x_n} g(x_1, \dots, x_n) f(x_1, \dots, x_n) dx_1 \cdots dx_n$, on condition that it is Lebesgue integrable.

Simple extension of the case of real random variables (same proofs).

First properties of expectation/mean

Lemma ("telescope")

Let X real r.v.,

- *If X discrete with values in $\mathbb{N}$, $\mathbb{E}(X) = \sum_{x=0}^{+\infty} \mathbb{P}(X > x)$.*
- *If X continuous of null density over $\mathbb{R}_-^*$,* $\mathbb{E}(X) = \int_{x=0}^{+\infty} \mathbb{P}(X > x)dx$.

Proposition (monotony/linearity/constants/decorrelation)

Let X, Y real r.v. discrete or continuous,

- *If $X \geq 0$, $\mathbb{E}(X) \geq 0$.*
- *If $a, b \in \mathbb{R}$, $\mathbb{E}(aX + bY) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y)$,*
- $\mathbb{E}(\mathbb{1}_\Omega) = 1$,
- *X, Y independent $\Rightarrow$ $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ (decorrelated r.v.)*

Moments of a real r.v.

Definitions & vocabulary : let X real r.v. and an integer $k \geq 1$

- Moment of order k of X : $m_k(X) \stackrel{\text{def}}{=} \mathbb{E}(X^k)$.
- Centered moment of order k of X : $\sigma_k(X) \stackrel{\text{def}}{=} \mathbb{E}((X - \mathbb{E}(X))^k)$.
- Variance of X : $var(X) \stackrel{\text{def}}{=} \sigma_2(X)$ ("dispersion" around the mean).
- Ecart-type/standard deviation of X : $\sqrt{var(X)}$ (often denoted σ).

Proposition (properties of variance)

- $var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
- $var(aX + b) = a^2 var(X)$.
- X *and* Y *independent* $\Rightarrow var(X + Y) = var(X) + var(Y)$.

Events seen as r.v.

Proposition (event → real r.v.)

If A is an event, then its indicator function $\mathbb{1}_A$ is a real r.v. such that $\mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$.

A useful translation :

- one can work on events by computing some expectations
- compatibility between useful definitions like independence
- transfer of results from r.v. to events

Generating functions associated with a real r.v.

Definition (Generating functions associated with a real r.v.)

Let X a real r.v., one can define the next series :

- probabilities $G_X(s) \stackrel{\text{def}}{=} \mathbb{E}(s^X) \underset{\text{dans } \mathbb{N}}{\stackrel{\text{à valeurs}}{=}} \sum_n \mathbb{P}(X=n)s^n$
- moments $M_X(t) \stackrel{\text{def}}{=} \mathbb{E}(e^{tX}) \underset{\text{autour de } 0}{\stackrel{\text{si } <+\infty}{=}} \sum_n \frac{\mathbb{E}(X^n)}{n!} t^n$
- characteristic $\Phi_X(t) \stackrel{\text{def}}{=} \mathbb{E}(e^{itX}) \underset{\text{dans } \mathbb{N}}{\stackrel{\text{à valeurs}}{=}} \sum_n \mathbb{P}(X=n)e^{itn}$

Useful tool both from math and algo points of view.

Generating functions : properties

Proposition (characterization of a law via series)

Let X, Y real r.v. discrete or continuous, X and Y have the same law iff their characteristic series satisfies $\Phi_X(t) = \Phi_Y(t)$ (thanks to Fourier transformation).

⚠ also true with moments series if finite around 0, otherwise there exists examples where $F_X \neq F_Y$ although $\forall k \geq 1,\ m_k(X) = m_k(Y)$ (cf. log-normal laws).

Proposition (series for sums of independent r.v.)

Let X, Y real r.v. over the same Ω and independent, then the series associated with the sum satisfy $G_{X+Y}(s) = G_X(s)G_Y(s)$, $M_{X+Y}(t) = M_X(t)M_Y(t)$, $\Phi_{X+Y}(t) = \Phi_X(t)\Phi_Y(t)$.

Classical discrete laws

Definition

Let X discrete real r.v., it is said :

- uniform if $\mathbb{P}(X=i)=1/n$ for $1 \leq i \leq n$
- Bernoulli if $X = \begin{cases} 1 & \text{with proba } p \\ 0 & \text{with proba } 1-p \end{cases}$
- binomial if $\mathbb{P}(X=i)=\binom{n}{i}p^i(1-p)^{n-i}$ for $0 \leq i \leq n$
- geometric if $\mathbb{P}(X=i)=p(1-p)^{i-1}$ for $i \geq 1$
- Poisson if $\mathbb{P}(X=i)=e^{-\lambda}\lambda^i/i!$ for $i \geq 0$

Classical continuous laws

Definition

Let X continuous real r.v. of density f, it is said :

- uniform if $f(x) = 1/(b-a)$ for $a \leq x \leq b$
- exponential if $f(x) = \lambda e^{-\lambda x}$ for $x \geq 0$
- normal if $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-\frac{(x-\mu)^2}{2\sigma^2})$ over $\mathbb{R}$ (denoted $\mathcal{N}(\mu, \sigma^2)$)
- log-normal if $f(x) = \frac{1}{x\sqrt{2\pi}} exp(-\frac{(\log x)^2}{2})$ for $x > 0$

An art of inequalities (I)

Proposition (large deviations : an inequality about distribution tails)

Let h function from $\mathbb{R}$ to $\mathbb{R}_+$ such that $h(X)$ remains a real r.v., then for all $a > 0$, $\mathbb{P}(h(X) \geq a) \leq \frac{\mathbb{E}(h(X))}{a}$.

Corollary (Markov inequality)

For all $a > 0$, $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}|X|}{a}$.

Corollary (Bienaymé-Tchebychev inequality)

For all $a > 0$, $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{var(X)}{a^2}$.

An art of inequalities (II)

Proposition (Jensen inequality)

Let h convex function from $\mathbb{R}$ *to* $\mathbb{R}$ *and* X *real r.v. with* $\mathbb{E}(X) < +\infty$, *then* $\mathbb{E}(h(X)) \geq h(\mathbb{E}(X))$.

Proposition (Hölder inequality)

Let $p, q \geq 1$ *real nbs such that* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$\mathbb{E}|XY| \leq (\mathbb{E}|X^p|)^{1/p} (\mathbb{E}|X^q|)^{1/q}.$$

Proposition (Minkowski inequality)

Let $p \geq 1$ *real nb, then* $[\mathbb{E}(|X+Y|^p)]^{1/p} \leq (\mathbb{E}|X^p|)^{1/p} + (\mathbb{E}|Y^p|)^{1/p}$.

An art of inequalities (III)

Proposition (Chernoff inequality)

Let $X_1, \ldots, X_n$ independent real r.v. with values in $\{0,1\}$, let $X = \sum_{i=1}^n X_i$ and $\mu = \mathbb{E}(X)$, then for all $\delta > 0$,

$$\mathbb{P}(X > (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu$$

Proposition (Hoeffding inequality)

Let $X_1, \ldots, X_n$ independent real v.a. a.s bounded with $\mathbb{P}(X_i \in [a_i, b_i]) = 1$ for $1 \leq i \leq n$, i.e. $\overline{X} = (\sum_{i=1}^n X_i)/n$ their empirical mean, then

$$\mathbb{P}(|\overline{X} - \mathbb{E}(\overline{X})| \geq t) \leq 2exp\left(-\frac{2t^2n^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

Convergence modes

Let $(X_n)_{n\in\mathbb{N}}, X$ real r.v. on the same probability space $(\Omega, \mathscr{F}, \mathbb{P})$,

Definition (convergence in law / in distribution)

$X_n \xrightarrow[n\to+\infty]{loi/D} X$ if $\forall x$ pt of continuity of F_X, $F_{X_n}(x) \underset{n\to+\infty}{\longrightarrow} F_X(x)$.

Definition (convergence in proba)

$X_n \xrightarrow[n\to+\infty]{P} X$ if $\forall \varepsilon > 0$, $\mathbb{P}(|X_n - X| > \varepsilon) \underset{n\to+\infty}{\longrightarrow} 0$.

Definition (convergence almost sure)

$X_n \xrightarrow[n\to+\infty]{p.s./a.s.} X$ if $\mathbb{P}(\{\omega \in \Omega | X_n(\omega) \underset{n\to+\infty}{\longrightarrow} X(\omega)\}) = 1$.

Remark : "same proba space" not necessary for conv. in law

Comparison of convergences

Theorem (comparison of convergence modes)

Let $(X_n)_{n\in\mathbb{N}}, X$ *real r.v. on the same proba space* $(\Omega, \mathscr{F}, \mathbb{P})$,

then : $X_n \xrightarrow{p.s.} X \;\Rightarrow\; X_n \xrightarrow{P} X \;\Rightarrow\; X_n \xrightarrow{D} X.$

Beware of traps :

⚠ $X_n \xrightarrow{p.s.} X \;\not\Leftarrow\; X_n \xrightarrow{P} X \;\not\Leftarrow\; X_n \xrightarrow{D} X$

⚠ $X_n \xrightarrow{D} X \;\not\Rightarrow\; X_n - X \xrightarrow{D} 0$

⚠ $X_n \xrightarrow{p.s.} X \;\not\Rightarrow\; \mathbb{E}(X_n) \longrightarrow \mathbb{E}(X)$

A tip of integration :

- $X_n \geq 0$ a.s. and $X_n \leq X_{n+1}$ a.s. $\Rightarrow \mathbb{E}(X_n) \longrightarrow \mathbb{E}(X)$
- $\forall n,\ |X_n| \leq Y$ a.s. and $\mathbb{E}|Y| < \infty \Rightarrow \mathbb{E}(X_n) \longrightarrow \mathbb{E}(X)$

Convergences & recurrent events

Notation : let $(A_n)_{n\in\mathbb{N}}$ a sequence of events,
$\{A_n \ \infty \text{ often}\} \stackrel{\text{def}}{=} \{\omega \in \Omega | \omega \in A_n \text{ for } \infty \text{ many } A_n\} =$ with $\cup$ and $\cap$?

Theorem (CNS of convergence a.s.)

$X_n \xrightarrow{p.s.} X$ *iff* $\forall \varepsilon > 0,\ \mathbb{P}(|X_n - X| \geq \varepsilon \ \infty \ \textit{often}) = 0.$

Theorem (Borel-Cantelli)

Let $(A_n)_{n\in\mathbb{N}}$ *a sequence of events,*

- *If* $\sum_n \mathbb{P}(A_n) < \infty$, *then* $\mathbb{P}(A_n \ \infty \ \textit{often}) = 0.$
- *If* $\sum_n \mathbb{P}(A_n) = \infty$ *and* A_n *independent, then* $\mathbb{P}(A_n \ \infty \ \textit{often}) = 1.$

Convergences & recurrent events

Notation : let $(A_n)_{n\in\mathbb{N}}$ a sequence of events,
$\{A_n \infty \text{ often}\} \stackrel{\text{def}}{=} \{\omega \in \Omega | \omega \in A_n \text{ for } \infty \text{ many } A_n\} = \cap_{k\geq 0} \cup_{n\geq k} A_n$.

Theorem (CNS of convergence a.s.)

$X_n \xrightarrow{p.s.} X$ *iff* $\forall \varepsilon > 0$, $\mathbb{P}(|X_n - X| \geq \varepsilon \ \infty \text{ often}) = 0$.

Theorem (Borel-Cantelli)

Let $(A_n)_{n\in\mathbb{N}}$ *a sequence of events,*

- *If* $\sum_n \mathbb{P}(A_n) < \infty$, *then* $\mathbb{P}(A_n \ \infty \text{ often}) = 0$.
- *If* $\sum_n \mathbb{P}(A_n) = \infty$ *and* A_n *independent, then* $\mathbb{P}(A_n \ \infty \text{ often}) = 1$.

Limit theorems

- $(X_n)_{n\geq 1}$ i.i.d. r.v.= defined on the same probability space, independent, identically distributed (same law).
- Empirical mean $\overline{X_n} \stackrel{\text{def}}{=} \frac{1}{n}(X_1 + \cdots + X_n)$.

Theorem (weak law of large numbers, simple proof when $\sigma_2 < \infty$)

Let $(X_n)_{n\geq 1}$ i.i.d. where $\mu = \mathbb{E}(X_1)$ finite, then $\overline{X_n} \xrightarrow{P} \mu$.

Theorem (strong law of large numbers, simple proof when $\sigma_4 < \infty$)

Let $(X_n)_{n\geq 1}$ i.i.d. where $\mu = \mathbb{E}(X_1)$ finite, then $\overline{X_n} \xrightarrow{p.s.} \mu$.

Theorem (central limite theorem)

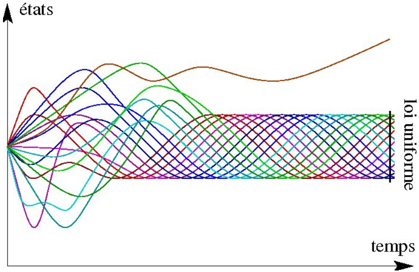
Let $(X_n)_{n\geq 1}$ i.i.d. where $\mu = \mathbb{E}(X_1)$ and $\sigma^2 = var(X_1)$ finite, then

$$\frac{\sqrt{n}}{\sigma}(\overline{X_n} - \mu) \xrightarrow{D} \mathcal{N}(0,1).$$

Illustrations of convergence modes

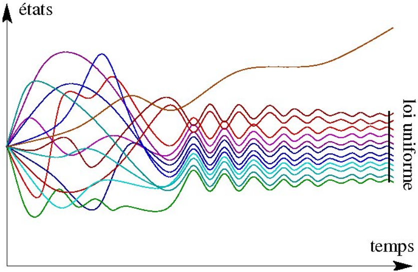
Vocabulary : stochastic process : evolution of r.v. formalised by a sequence $(X_t)_{t\in\mathbb{N} \text{ ou } \mathbb{R}_+}$, of r.v. over the same space $(\Omega,\mathscr{F},\mathbb{P}) \to$ a trajectory/réalisation : the sequence $(X_t(\omega))_t$ for a fixed $\omega \in \Omega$.

cv in law



global repartition of trajectories becomes invariant

cv a.s.

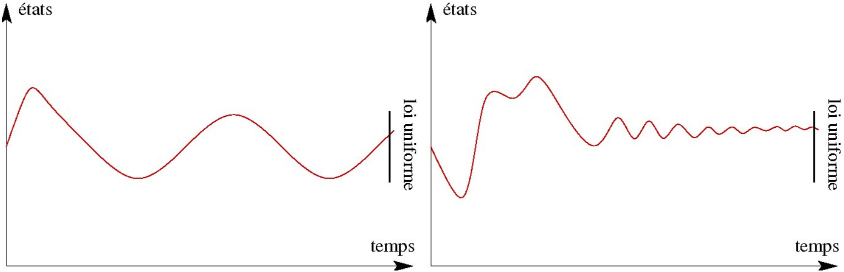


each trajectory converges individually (a.s.)

Illustrations of convergence modes

Vocabulary : stochastic process : evolution of r.v. formalised by a sequence $(X_t)_{t \in \mathbb{N} \text{ ou } \mathbb{R}_+}$, of r.v. over the same space $(\Omega, \mathscr{F}, \mathbb{P})$ → a trajectory/réalisation : the sequence $(X_t(\omega))_t$ for a fixed $\omega \in \Omega$.

cv in law | **cv a.s.**



global repartition of trajectories becomes invariant

each trajectory converges individually (a.s.)