Probabilités et statistique

Cours de Hugo Duminil-Copin

Notes de Ekin Arikök

Automne 2021

Table des matières

Chapitre 0

Rappel de la théorie de la mesure

Cours 1

Définition 0.1. Soit E un ensemble. Une **tribu** (ou σ-algèbre) sur E est un sous ensemble $\mathcal{A} \subset \mathcal{P}(E)$ tel que :

1. $E \in \mathcal{A}$
2. $A \in \mathcal{A} \implies A^c \in \mathcal{A}$
3. $A \in \mathcal{A}^\omega \implies \bigcup\limits_{n\in\omega} A_n \in \mathcal{A}$

Définition 0.2. Soit $\mathcal{C} \subset \mathcal{P}(E)$ quelconque. Il existe une unique plus petite tribu contenant $\mathcal{C}$. Cette tribu, appelée **tribu engendrée par** $\mathcal{C}$, et noté $\sigma(\mathcal{C})$, est définie par $\sigma(\mathcal{C}) = \bigcap\limits_{\mathcal{A} \text{ tribu}, \mathcal{A}\supset\mathcal{C}} \mathcal{A}$

Exemple 0.3.

1. $\mathcal{P}(E)$ est une tribu. On utilise souvent (presque systématiquement) cette tribu lorsque E est fini ou Dénombrable. Dans ce cas, $\sigma(\{\{x\} : x \in X\}) \in \mathcal{P}(X)$
2. Si E est un espace topologique et $\mathcal{T}$ dénote l'ensemble des ouverts, alors $\mathcal{B}(E) = \sigma(\mathcal{T})$ est appelée **tribu de Lebegue**
3. Soient E_1 et E_2 deux espaces et $\mathcal{A}_1$ et $\mathcal{A}_2$ deux tribus sur E_1 et E_2 respectivement. La tribu $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{A}_1 \times \mathcal{A}_2) = \sigma(\{A \times B, A \in \mathcal{A}_1, B \in \mathcal{A}_2\})$ est appelée la **tribu produit**

Définition 0.4. Une **mesure (positive)** sur un espace mesurable $(E, \mathcal{A})$ est une fonction $\mu : \mathcal{A} \to [0, +\infty]$ vérifiant :

1. $\mu(\varnothing) = 0$
2. $\forall A \in \mathcal{A}^\omega$ disjoint, $\mu(\bigcup_{n\in\omega} A_n) = \sum_{n\in\omega} \mu(A_n)$

$(E, \mathcal{A}, \mu)$ est un **espace mesuré**.

Exemple 0.5.

1. Soit E fini ou dénombrable et $\mathcal{A} \in \mathcal{P}(E)$

$$\mu(A) = \text{card}(A), \forall A \in \mathcal{P}(E)$$

   μ est appelée mesure de comptage
2. Soit $x \in E$, la **mesure de Dirac** est définie par la formule

$$\delta_x(A) := \mathbb{1}_A(X)$$

   où $\mathbb{1}_A$ est la fonction caractéristique de A.

   Plus généralement, si $(x_n)_{n\in\omega}$ est une famille d'éléments de E et $\alpha \in [0, +\infty]^\omega$, alors on peut définir $\mu = \sum -n \in \omega\alpha_n\delta_{x_n}$
3. Considérons $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On appelle **mesure de Lebegue** λ l'unique mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\lambda((a, b)) = b - a$, $\forall a < b$.

Définition 0.6. Soient $(E, \mathcal{A})$ et $(F, \mathcal{B})$ deux espaces mesurables. Une fonction $f : E \to F$ est **mesurable** ssi $\forall B \in \mathcal{B}, f^{-1}(B) \in \mathcal{A}$.

Définition 0.7. Soit $f : (E, \mathcal{A}) \to (F, \mathcal{B})$ une fonction mesurable et μ une mesure sur $(E, \mathcal{A})$. La **mesure image** de μ par f est la mesure sur $(F, \mathcal{B})$ définie par $\forall B \in \mathcal{B}, f_*\mu(B) = \mu(f^{-1}(B))$

Théorème 0.8 (unicité)**.** *Soit* $(E, \mathcal{A})$ *un espace mesurable,* $\mathcal{C} \subset \mathcal{P}(E)$ *quelconque et* μ *et* ν *deux mesures sur* $(E, \mathcal{A})$*. Si*

1. $\mu(C) = \nu(C), \forall C \in \mathcal{C}$
2. $\mathcal{C}$ *est stable par intersection finie (i.e.* $A, B \in \mathcal{C} \implies A \cap B \in \mathcal{C}$*), et* $\sigma(\mathcal{C}) = \mathcal{A}$
3. *il existe une suite croissante* $(E_n)_{n\in\omega}$ *tq* $\bigcup_{n\in\omega} E_n = E$ *et* $\mu(E_n) = \nu(E_n) < \infty, \forall n \in \omega$

Alors $\mu = \nu$

Remarque.
- Le théorème implique l'unicité de la mesure de Lebesgue.
- Le théorème est une conséquence du théorème des classes monotones.

Soient μ_1 et μ_2 deux mesures finies sur $(E_1, \mathcal{A}_1)$ et $(E_2, \mathcal{A}_2)$ respectivement, il existe une unique mesure $\mu_1 \otimes \mu_2$, appelée mesure produit, sur $(E_1 \times E_2, \mathcal{A}_1 \otimes \mathcal{A}_2)$ t.q. $\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2), \forall A_1 \in \mathcal{A}_1, \forall A_2 \in \mathcal{A}_2$

Définition 0.9. Soit μ une mesure sur $(E, \mathcal{A})$ et f une fonction mesurable. On dit que f est **intégrable** si $\displaystyle\int |f(x)| d\mu(x) < \infty$

On définit $\displaystyle\int f(x) d\mu(x) := \int f(x) \mathbb{1}_{f^{-1}(\mathbb{R}_+)}(x) d\mu(x) - \int (-f)(x) \mathbb{1}_{f^{-1}(\mathbb{R}_-)}(x) d\mu(x)$

Théorème 0.10 (Fubini)**.** *Soit* $f : E \times F \to \mathbb{R}$ *intégrable pour* $\mu \otimes \nu$*. Alors*

$$\begin{array}{ccc} E & \to & \mathbb{R} \\ x & \mapsto & \int_F f(x,y) d\nu(y) \end{array} \quad et \quad \begin{array}{ccc} F & \to & \mathbb{R} \\ x & \mapsto & \int_E f(x,y) d\mu(x) \end{array}$$

sont respectivement intégrables sur E *et* F*, et*

$$\int_{E\times F} f(x,y) d\mu \otimes d\nu(x,y) = \int_E \left(\int_F f(x,y) d\nu(y) \right) d\mu(x) = \int_F \left(\int_E f(x,y) d\mu(x) \right) d\nu(y)$$

Remarque.
- Si $\int_{E\times F} |f(x,y)| (d\mu \otimes d\nu)(x,y) < \infty$, alors on peut échanger les intégrales.
- Si $f(x,y)$, ça marche ! (Tonelli)

Théorème 0.11 (Hölder)**.** *Soient* $p, q \in \mathbb{R}_+$ *tels que* $\frac{1}{p} + \frac{1}{q} = 1$*. Pour toutes fonctions mesurables de* E *dans* $\mathbb{R}$*,*

$$\int |fg| d\mu \leq \left(\int |f|^p d\mu \right)^{1/p} \left(\int |g|^q d\mu \right)^{1/q}$$

En particulier, si $|f|^p$ *et* $|g|^q$ *sont intégrables, alors* $|fg|$ *l'est aussi.*

Remarque.
- Si $p = q = 2$, on obtient Cauchy-Schwarz.
- $\|f\|_p = \int (|f|^p d\mu)^{1/p}$, alors $\|fg\|_1 \leq \|f\|_p \|g\|_q$

Théorème 0.12 (Convergence dominée)**.** *Soit* $(f_n)_{n\in\omega}$ *une suite de fonctions intégrables de* $(E, \mathcal{A}, \mu)$ *dans* $\mathbb{R}$*. Si*

1. *Il existe* f *mesurable de* $(E, \mathcal{A}, \mu)$ *dans* $\mathbb{R}$ *telle que* $\forall x \in E$, $\lim\limits_{n\to\infty} f_n(x) = f(x)$ μ*-presque partout*
2. *Il existe* g *de* $(E, \mathcal{A}, \mu)$ *dans* $\mathbb{R}_+$ *intégrable tel que* $|f_n(x)| \leq g(x)$ μ*-presque partout,* $\forall n \in \omega, \forall x \in E$

Alors f *est intégrable et* $\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$

Chapitre 1

Fondements de la théorie des probabilités

1.1 Espaces de probabilité

Cours 2

Définition 1.1. Une mesure $\mathbb{P}$ est dite **de probabilité** si $\mathbb{P}[\Omega] = 1$. Un espace mesurable $(\Omega, \mathcal{A})$ muni d'une mesure de probabilité sera appelé **espace de probabilité** $(\Omega, \mathcal{A}, \mathbb{P})$. Les éléments de $\mathcal{A}$ sont appelés **évènements**.

Remarque. Ω représente l'ensemble des possibles, en d'autres termes toutes les réalisations possibles du hasard. Un élément ω dans Ω est donc une **réalisation** du hasard. Les évènements sont exactement les ensembles de réalisations dont on veut calculer les **probabilités**.

Exemple 1.2 (espaces de probabilité)**.**

1. lancer de dé :
   $\Omega = \{1, 2, 3, 4, 5, 6\}$
   $\mathcal{A} = \mathcal{P}(\Omega)$
   $\mathbb{P} = \frac{\text{mesure de comptage}}{6}$
2. lancer de deux dés :
   $\Omega = \{1, 2, 3, 4, 5, 6\}^2$
   $\mathcal{A} = \mathcal{P}(\Omega)$
   $\mathbb{P}(\omega) = \frac{1}{36}$
3. lancer de fléchettes :
   $\Omega = \mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$
   $\mathcal{A} = \mathcal{B}(\mathbb{D})$
   $\mathbb{P} = \frac{dx}{\pi}$
4. main au poker :
   $\Omega = \{s \in \mathcal{P}(J) : |s| = 5\}$ où J est l'ensemble des cartes
   $\mathcal{A} = \mathcal{P}(\Omega)$
   $\mathbb{P}(s) = 1/\binom{52}{5}$
5. lancer de pièces (jusqu au premier Pile) :
   $\Omega = \{0, 1\}^{\mathbb{N}}$
   $\mathcal{A} = \sigma(\{\omega_1 = \xi_1, ..., \omega_k = \xi_k\}, \forall \xi \in \Omega, \forall k)$ (la tribu engendrée par les évènements dépendant d'un nombre fini de lancers.)
   $\mathbb{P}[\{\omega_1 = \xi_1, \omega_k = \xi_k\}] = \frac{1}{2^k}$
6. particule partant de 0 est voyageant de sommet en sommet sur $\mathbb{Z}^d$:
   $\Omega = \{\omega \in (\mathbb{Z}^d)^{\mathbb{N}}, \omega = 0\}$
   $\mathcal{A} = \sigma(\{X_k = k\}, \forall k \in \mathbb{N}, \forall x \in \mathbb{Z}^d) = \sigma(\{X_1 = x_1, ..., X_k = x_k\}, \forall k \in \mathbb{N}, \forall (x_1, ..., x_k) \in (\mathbb{Z}^d)^{\mathbb{N}})$
   Dans le cas où le marcheur choisit un voisin uniformément et se déplace vers lui
   $\mathbb{P}[\{X_1 = x_1, ..., X_k = x_k\}] = \frac{1}{(2d)^k}$
7. particule se déplaçant continuement dans $\mathbb{R}^d$:
   $\Omega = \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$
   $\mathcal{A} = \sigma(\{\omega \in U\}, \forall t \in \mathbb{R}_+, \forall U \text{ ouvert})$
   $\mathbb{P} =$ mesure de Wiener

8. graphe aléatoire (modèle de percolation) :
"On garde tous les sommets, et pour chaque arrête, on lance une pièce de monnaie truquée (tombant sur face avec prob $\phi \in [0,1]$) et si la pièce tombe sur face, on garde l'arrête, Sinon, on l'efface."
$\Omega = \mathcal{P}(E) = \{0,1\}^E = \{\omega = (\omega(e) = e \in E) \in \{0,1\}^E\}$
$\mathcal{A} = \sigma(\{\omega(e) = 1\}, e \in E) = \sigma(\{\omega(e_1) = \epsilon_1, ..., \omega(e_k) = \epsilon_k\}, \forall e_1, ..., e_k \in E, \forall \epsilon_1, ..., \epsilon_k \in \{0,1\})$
$\mathbb{P}[\{\omega(e_1) = \epsilon_1, ..., \omega(e_k) = \epsilon_k\}] = \prod_{i=1}^{k} p^{\epsilon_i}(1-p)^{1-\epsilon_i}$
($\omega(e) = 1$ si e dans le graphe, $\omega(e) = 0$ sinon.)

1.2 Variables aléatoires :

Fixons un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.

Motivation : Une variable X dont la valeur dépend de la réalisation ω.

Définition 1.3. Soit $(E, \mathcal{E})$ un espace mesurable. Une **variable aléatoire** (VA) à valeurs dans E est une application $X : \Omega \to E$ mesurable de (Ω, A) dans $(E, \mathcal{E})$

Remarque. Si $E = \mathbb{R}^d$, on prendra toujours $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$.

Exemple 1.4. On reprend les espaces de l'exemple 1.2 :

2. $X((i,j)) = i + j$ somme des points.
5. $X(\omega) = \inf\{i : \omega(i) = pile\}$, $(\inf \varnothing = \infty)$
6. $X(\omega) = \omega(i)$
8. $X(\omega) = \left|\{x \in \mathbb{Z}^d, x \text{ est connecté à } 0 \text{ dans } \omega\}\right|$

Définition 1.5. La **loi d'une variable aléatoire** X est la mesure (de probabilité sur $(E, \mathcal{E})$, noté $\mathbb{P}_X$ et donnée par la formule

$$\mathbb{P}_X = X_* \mathbb{P}$$

Remarque. On oubliera souvent ce que vaut $(\Omega, \mathcal{A}, \mathbb{P})$ et on se concentrera sur $(E, \mathcal{E}, \mathbb{P}_X)$. (Notation probabiliste :)

$$\mathbb{P}_X[B] = \mathbb{P}[\{\omega \in \Omega : X(\omega) \in B\}] =: \mathbb{P}[X \in B], \qquad B \in \mathcal{E}$$

Exemple 1.6.

1. Soit μ un mesure de probabilité sur $(E, \mathcal{E})$.
   On prend
   $$\begin{array}{rccc} X: & E & \to & E \\ & \omega & \mapsto & \omega \end{array}$$
   X est mesurable de $(E, \mathcal{E})$ dans lui-même.
   $$\mathbb{P}_X[B] = \mu[\{\omega : X(\omega) = \omega \in B\}] = \mu[B]$$
2. Variables aléatoires discrètes :
   Lorsque E est dénombrable, la loi de X est donnée par
   $$\mathbb{P}_X = \sum_{x \in E} \mathbb{P}[X = x]\delta_X$$
   Dans ce cas, $\mathbb{P}_X$ est déterminé par les $\mathbb{P}[X = x]$
3. $\Omega = \{1, 2, ..., 6\}^{\mathbb{N}^*}$
   $X(\omega) = \inf\{i : \omega(i) = 6\}$, $(\inf \varnothing = \infty)$
   X est à valeurs dans $\mathbb{N}^* \cup \{\infty\}$.

$$
\begin{aligned}
\mathbb{P}_X[\{k\}] &= \mathbb{P}[X = k] \\
&= \mathbb{P}[\{\omega : X(\omega) = k\}] \\
&= \sum_{\xi_1,...\xi_k-1\in\{1,...,5\},\xi_k=6} \mathbb{P}[\omega(1) = \xi_1, ..., \omega(k) = \xi_k] \\
&= \frac{1}{6}\cdot\left(\frac{5}{6}\right)^{k-1}
\end{aligned}
$$

$$
\mathbb{P}_X[\{\infty\}] = 1 - \sum_{k\in\mathbb{N}^*} \mathbb{P}_X[\{k\}] = 1 - \sum_{k\in\mathbb{N}^*} \frac{1}{6}\left(\frac{5}{6}\right)^{k-1} = 0
$$

$$
\mathbb{P}_X[\{\infty\}] = \mathbb{P}_X[\bigcap_{k\in\mathbb{N}^*} [k,\infty)] = \mathbb{P}[\bigcap_{k\in\mathbb{N}^*} \{X \geq k\}] = \lim_{k\to\infty} \mathbb{P}[X \geq k] = \lim_{k\to\infty} (5/6)^{k-1} = 0
$$

Définition 1.7. Une variable aléatoire X à valeurs dans $\mathbb{R}^d$ est dite **à densité** si X est absolument continue par rapport à la mesure de Lebesgue, i.e. qu'il existe $p : \mathbb{R}^d \to \mathbb{R}_+$ telle que $\mathbb{P}_X[B] = \int_B p(x)\,dx$

Cours 3

Définition 1.8. Soit X une variable aléatoire à valeurs dans $(E, \mathcal{E})$. La **tribu engendrée par** X, notée $\sigma(X)$, est la tribu

$$
\sigma(X) := \{X^{-1}(B) : B \in \mathcal{E}\}
$$

Remarque. $\sigma(X)$ est la plus petite tribu sur Ω, rendant X mesurable.

Proposition 1.9. *Soit X une variable aléatoire à valeurs dans $(E, \mathcal{E})$, et Y une variable aléatoire réelle. Alors les deux propriétés suivantes sont équivalentes :*

1. *Y est $\sigma(X)$-mesurable.*
2. *il existe une fonction f mesurable $f : E \to \mathbb{R}$ mesurable de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, telle que $Y = f(X)$.*

Démonstration.

(2. $\implies$ 1.) Soit $A \in \mathcal{B}(\mathbb{R})$, $Y^{-1}(A) = X^{-1}(f^{-1}(A))$ puisque f est mesurable de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, $f^{-1}(A) \in \mathcal{E}$. Par définition de $\sigma(X)$, $X^{-1}(f^{-1}(A)) \in \sigma(X)$

(1. $\implies$ 2.) (étape 1 : "Y étagée") Supposons que $Y = \sum_{i=1}^n \lambda_i \mathbb{1}_{A_i}$, où $A_i \in \sigma(X)$. Puisque $A_i \in \sigma(X)$, il existe $B_i \in \mathcal{E}$ tel que $A_i = X^{-1}(B_i)$. Posons

$$
f := \sum_{i=1}^n \lambda_i \mathbb{1}_{B_i}
$$

f est mesurable de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et de plus

$$
f(X) = \sum_{i=1}^n \lambda_i \mathbb{1}_{B_i} \circ X = \sum_{i=1}^n \lambda_i \mathbb{1}_{A_i} = Y
$$

(étape 2 : "cas général) Il existe (Y_n) une suite de fonctions $\sigma(X)$-mesurables et étagées telle que $\lim_{n\to\infty} Y_n = Y$ presque sûrement. On sait par l'étape 1 qu'il existe (f_n) suite de fonctions mesurables de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $Y_n = f_n(X)$. Posons

$$
f(x) := \begin{cases} \lim_{n\to\infty} f_n(x) & \text{si la limite existe} \\ 0 & \text{sinon} \end{cases}
$$

f est mesurable. De plus, pour tout $\omega \in \Omega$ tel que $Y_n(\omega)$ converge vers $Y(\omega)$, on a

$$
Y(\omega) = \lim_{n\to\infty} Y_n(\omega) = \lim_{n\to\infty} f_n(X(\omega))
$$

Puisque la limite existe pour $x = X(\omega)$, on est dans le cas de la première ligne et on a bien

$$
\lim_{n\to\infty} f_n(X(\omega)) = f(X(\omega))
$$

□

1.3 Quelques lois classiques

Lois discrètes

1. **Loi uniforme :** Soit E un ensemble *fini* de cardinal n. Une variable aléatoire X est de loi uniforme si

$$\forall x \in E, \mathbb{P}[X = x] = \frac{1}{n}$$

2. **Loi de Bernoulli de paramètre p :** Soit $p \in [0, 1]$. Une variable aléatoire X à valeurs dans $\{0, 1\}$ est dite de Bernoulli si :

$$\mathbb{P}[X = 1] = p = 1 - \mathbb{P}[X = 0]$$

3. **Loi binomiale de paramètres n et p :** Soient $n \in \mathbb{N}^*$ et $p \in [0, 1]$. Une variable aléatoire X à valeurs dans $\{0, 1, ..., n\}$ est dite binomiale ssi

$$\forall k \in \{0, 1, ..., n\}, \mathbb{P}[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$$

4. **Loi géométrique de paramètre p :** Soit $p \in [0, 1]$. Une variable aléatoire X à valeurs dans $\mathbb{N}^*$ est dite géométrique ssi

$$\forall \in \mathbb{N}^*, \mathbb{P}[X = k] = p(1 - p)^{k-1}$$

5. **Loi de Poisson de paramètre λ :** Soit $\lambda \in (0, +\infty)$. Une variable aléatoire X à valeurs dans $\mathbb{N}$ est dite de Poisson ssi

$$P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}, \ \forall k \in \mathbb{N}$$

Lois continues

Dans les trois exemples suivants, X est une variable aléatoire réelle à densité $p(x)$.

1. **Loi uniforme sur $[a, b]$:** Soient $a < b$. La variable aléatoire X est uniforme sur $[a, b]$ ssi

$$p(x) = \frac{1}{b - a} \mathbb{1}_{[a,b]}(x)$$

2. **Loi exponentielle de paramètre λ :** Soit $\lambda \in (0, +\infty)$. La variable aléatoire X est dite exponentielle ssi

$$p(x) = \lambda e^{-\lambda x} \mathbb{1}_{[0,+\infty)}(x)$$

Remarque. Une propriété importante de la loi exponentielle : soient $a, b > 0$,

$$\mathbb{P}[X > a + b] = \mathbb{P}[X > a] \, \mathbb{P}[X > b]$$

3. **Loi gaussienne (normale) :** Soient $\mu \in \mathbb{R}$ et $\sigma > 0$. Une variable aléatoire X est de loi gaussienne de paramètres μ et σ (notée $\mathcal{N}(\mu, \sigma^2)$) ssi

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x - \mu)^2}{2\sigma^2} \right]$$

Remarque. C'est la loi la plus importante de la théorie des probabilités.

1.4 Espérance mathématique

Définition 1.10. Soit X une variable aléatoire réelle qui est soit *positive* soit *intégrable*. **L'espérance** de X est donnée par

$$\mathbb{E}[X] := \int_\Omega X(\omega) d\, \mathbb{P}[\omega]$$

Plus généralement, si X est à valeurs dans $\mathbb{R}^d$ ou dans $\mathbb{C}$, on définit

$$\mathbb{E}[X] := (\mathbb{E}[X_1], ..., \mathbb{E}[X_n]), \text{ où } X = (X_1, ..., X_n)$$

$$\mathbb{E}[X] := \mathbb{E}[X_1] + i\, \mathbb{E}[X_2], \text{ où } X = X_1 + iX_2, X_1 \text{ et } X_2 \text{ sont à valeurs réelles}$$

Remarque. $\mathbb{E}[X]$ est la valeur moyenne de la variable aléatoire X.

Soit $A \in \mathcal{A}$, $\mathbb{E}[\mathbb{1}_A] = \mathbb{P}[A]$.

Proposition 1.11 (formule de transfert)**.** *Soit X une variable aléatoire à valeurs dans $(E, \mathcal{E})$ et f une fonction mesurable de $(E, \mathcal{E})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,*

$$\mathbb{E}[f(X)] = \int_E f(x) d\,\mathbb{P}_X[x]$$

Démonstration. $\mathbb{P}_X = X_* \mathbb{P}$ + définition de $\mathbb{E}[f(X)]$ en fonction de $\mathbb{P}$. □

Exemple 1.12.

2. $X(\omega) = i + j$

$$\mathbb{E}[X] = \sum_{(i,j)} (i+j) \frac{1}{36} = 7$$

3. $X(\omega) = \inf\{i : X(i) = 6\}$

$$\mathbb{E}[X] = \int_E x d\,\mathbb{P}_X[x] = \sum_{k=1}^{+\infty} k\ \mathbb{P}_X[\{k\}] + \infty\ \mathbb{P}_X[\{\infty\}] = \sum_{k=1}^{+\infty} k (\frac{5}{6})^{k-1} \frac{1}{6} = 6$$

Espérance des lois :

Loi	**Espérence**
Bernoulli (p)	$p \cdot 1 + (1-p) \cdot 0 = p$
Binomiale (n, p)	$\sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np$
Géométrique (p)	$\sum_{k=0}^{+\infty} kp(1-p)^{k-1} = \frac{1}{p}$
Poisson (λ)	$\sum_{k=0}^{+\infty} ke^{-\lambda} \frac{\lambda^k}{k!} = \lambda$
Uniforme (a, b)	$\frac{1}{b-a} \int_a^b x dx = \frac{a+b}{2}$
Exponentielle (λ)	$\int_0^{+\infty} x\lambda e^{-\lambda x} dx = \frac{1}{\lambda}$
Gaussienne (μ, σ^2)	$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \mu$

Cours 4

1.5 Moments d'une variable aléatoire

Définition 1.13. Soit X une variable aléatoire réelle et $p > 0$. Supposons que $X \geq 0$ ou $\mathbb{E}[|X|^p] < \infty$. On appelle **moment d'ordre** p la quantité $\mathbb{E}[X^p]$.

Si $p \geq 1$, on peut définir l'espace $L^p(\Omega, \mathcal{A}, \mathbb{P})$, muni de la norme $\|X\|_p := \mathbb{E}[|X|^p]^{\frac{1}{p}}$. Dans ce contexte, l'inégalité de Hölder devient $\|XY\|_1 \leq \|X\|_p \|Y\|_q$, où $\frac{1}{p} + \frac{1}{q} = 1$.

Corollaire 1.14.

1. $\mathbb{E}[|X|^r] \leq \mathbb{E}[|X|^s]^{r/s}$*, où* $1 \leq r \leq s$
2. $\mathbb{E}[|XY|] \leq \mathbb{E}[X^2]^{1/2}\, \mathbb{E}[Y^2]^{1/2}$
3. $\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]$

Démonstration.

1. On applique Hölder à $|X|^r, Y = 1, p = \frac{s}{r}, \frac{1}{q} = 1 - \frac{r}{s}$
2. On applique Hölder pour $p = q = 2$

3. On applique le point *2.* à $Y = 1$: $\mathbb{E}[X]^2 \leq \mathbb{E}[|X|]^2 \leq \mathbb{E}[X^2]$

□

Remarque. $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$ par l'inégalité triangulaire.

Définition 1.15. Soit $X \in L^2$. La **variance** de X est la quantité

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2]$$

On appelle **écart-type** la quantité $\sigma_X := \sqrt{\mathrm{Var}[X]}$

Proposition 1.16. *Soit $X \in L^2$*

1. $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

2. $\mathrm{Var}(X) = \inf_{a \in \mathbb{R}} \mathbb{E}[(X-a)^2]$

3. $\mathrm{Var}(X) = 0 \implies \mathbb{P}[X = \mathbb{E}[X]] = 1$

Notation. Si $\mathbb{P}[A] = 1$, on dit que A a lieu **presque sûrement** (p.s.). Pour *3.*, on dit $X = \mathbb{E}[X]$ p.s.

Démonstration.

1.

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2 - 2X\,\mathbb{E}[X] + \mathbb{E}[X^2]] \\
&= \mathbb{E}[X^2] - 2\,\mathbb{E}[X]\,\mathbb{E}[X] + \mathbb{E}[X]^2 \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2
\end{aligned}$$

2. En reformulant, il suffit de montrer que $f(a) := \mathbb{E}[(X-a)^2]$ admet son minimum en $\mathbb{E}[X]$

$$f'(a) = -2(\mathbb{E}[X] - a)$$

En particulier, $\mathbb{E}[X]$ est la seule valeur réelle telle que $f'(a) = 0$. $f(a) \to \pm\infty$ en $\pm\infty$, donc elle admet un minimum qui est forcément $\mathbb{E}[X]$.

3. $\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2]$. Une fonction d'intégrale nulle vaut 0 p.s.

□

Exemple 1.17.

Bernoulli (p) : $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1-p)$

Binomiale (n,p) :

$$\mathbb{E}[X(X-1)] = \sum_{k=0}^{n} k(k-1)\binom{n}{k} p^k (1-k)^{n-k} = n(n-1)p^2$$

donc :

$$\begin{aligned}
\mathrm{Var}(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\
&= n(n-1)p^2 + np - (np)^2 \\
&= np(1-p)
\end{aligned}$$

Géométrique (p) :

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{p} \\
\mathbb{E}[X(X-1)] &= \sum_{k=0}^{\infty} pk(k-1)(1-p)^k = \frac{1}{p^2} \\
\mathrm{Var}(X) &= \frac{1}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p}
\end{aligned}$$

Poisson (λ) : $\operatorname{Var}(X) = \lambda$

Uniforme (a, b) :

$$\operatorname{Var}(X) = \frac{1}{b-a}\int_a^b (x - \frac{a+b}{2})^2 dx = \frac{(b-a)^2}{12}$$

Exponentielle (λ) :

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\lambda} \\
\mathbb{E}[X^2] &= \lambda \int_0^{+\infty} x^2 e^{-\lambda x} dx \\
&= \left[-x^2 e^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx \\
&= 0 + \frac{2}{\lambda}\mathbb{E}[X] = \frac{2}{\lambda^2} \\
\operatorname{Var}(X) &= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}
\end{aligned}$$

Gaussienne (μ, σ^2) :

$$\begin{aligned}
\mathbb{E}[X] &= \mu \\
\operatorname{Var}[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{+\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} y^2 e^{-\frac{y^2}{2}} dy \\
&= \frac{\sigma^2}{\sqrt{2\pi}} (\left[-y e^{-\frac{y^2}{2}}\right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy) \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{y^2}{2}} dy = \sigma^2
\end{aligned}$$

car

$$\frac{1}{2\pi}\int_{\mathbb{R}^2} e^{-\frac{y^2+x^2}{2}} dy dx = \int_0^{+\infty} r e^{-\frac{r^2}{2}} dr \int_0^{2\pi} \frac{d\theta}{2\pi} = 1 \cdot 1 = 1$$

Espérance et variance des lois :

Loi	**Espérence**	**Variance**
Bernoulli (p)	p	$p(1-p)$
Binomiale (n, p)	np	$np(1-p)$
Géométrique (p)	$\frac{1}{p}$	$\frac{1}{p}$
Poisson (λ)	λ	λ
Uniforme (a, b)	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentielle (λ)	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gaussienne (μ, σ^2)	μ	σ^2

Définition 1.18. Soient X et Y dans L^2. La **covariance** de X et Y est la quantité

$$\operatorname{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

De plus, si $X = (X_1, ..., X_d)$, on appelle **matrice de covariance** la matrice de $\mathcal{M}_{d,d}(\mathbb{R})$

$$(K_X)_{ij} := \operatorname{Cov}(X_i, Y_j). 1 \leq i, j \leq d$$

Remarque. $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

Proposition 1.19 (quelques inégalités classiques)**.**

(Markov) Si $X \geq 0$ et $a > 0$, alors

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$$

(Bienaymé-Tchebychev) Soit X une variable aléatoire et $a > 0$,

$$\mathbb{P}[|X - \mathbb{E}[X]| > a] \leq \frac{\mathrm{Var}(X)}{a^2}$$

(Chernov) Soit $X \geq 0$ et $a > 0$,

$$\mathbb{P}[X \geq a] \leq e^{-a}\,\mathbb{E}[e^X]$$

Démonstration. Soit $f : \mathbb{R} \to \mathbb{R}_{>0}$, Montrons que $\mathbb{P}[f(X) \geq f(a)] \leq \frac{\mathbb{E}[f(X)]}{f(a)}$.

$$\begin{aligned}
\mathbb{E}[f(X)] &\geq \mathbb{E}[f(X)\mathbb{1}_{f(X)\geq f(a)}] \\
&\geq \mathbb{E}[f(a)\mathbb{1}_{f(X)\geq f(a)}] \\
&= f(a)\,\mathbb{P}[f(X) \geq f(a)]
\end{aligned}$$

Donc en particulier on prend :

- Markov : $f(x) = x$
- BT : $f(x) = x^2$
$$\mathbb{P}[|X - \mathbb{E}[X]| > a] = \mathbb{P}[(X - \mathbb{E}[X])^2 > a^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}[X])^2]}{a^2}$$
- Chernov : $f(x) = e^x$
$$\mathbb{P}[X \geq a] = \mathbb{P}[e^X \geq e^a] \leq \frac{\mathbb{E}[e^X]}{e^a}$$

□

Remarque. Plus on peut prendre f grand, meilleure est la borne (x^k)

1.6 Fonction de répartition d'une variable aléatoire à valeurs réelles

Cours 5

Définition 1.20. Soit X une variable aléatoire à valeurs réelles. On appelle **fonction de répartition** la fonction

$$\begin{aligned} F_X : \quad \mathbb{R} &\to [0,1] \\ x &\mapsto \mathbb{P}[X \leq x] \end{aligned}$$

Exemple 1.21.

1. Variable aléatoire uniforme sur $[a,b]$
$$F_X(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } x \in (a,b) \\ 1 & \text{sinon} \end{cases}$$
2. Variable aléatoire exponentielle de paramètre λ
$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-\lambda x} & \text{sinon} \end{cases}$$
3. Variable aléatoire δ_0
$$F_X(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 & \text{sinon} \end{cases}$$

Proposition 1.22. *Si F est la fonction de répartition d'une variable aléatoire, alors*

- *F est croissante*

— $\lim_{x\to-\infty} F(x) = 0$ *et* $\lim_{x\to+\infty} F(x) = 1$
— F *est continue à droite*

Démonstration.

1. si $x \leq y$, alors $\{X \leq x\} \subset \{X \leq y\}$
2. $A_n = \{X \leq n\}$. $\lim_{n\to\infty} F(-n) = \lim_{n\to\infty} \mathbb{P}[A_n]$. Comme (A_n) est une suite décroissante d'évènements et que $\bigcap A_n = \varnothing$, on en déduit que $\lim_{n\to\infty} F(-n) = 0$. Combiné à la croissance de F, cela donne $\lim_{x\to-\infty} F(x) = 0$.
3. Soit $x \in \mathbb{R}$, soit (x_n) une suite décroissante et convergeant vers x.

$$\lim F(x_n) = \lim \mathbb{P}[\{X \leq x_n\}] = \mathbb{P}\left[\bigcap_n \{X \leq x_n\}\right] = \mathbb{P}[X \leq x] = F(x)$$

□

Proposition 1.23. *Si F satisfait les propriétés de la proposition précédente, alors, il existe une variable aléatoire X telle que $F = F_X$.*

Démonstration. Posons $\Omega = [0,1]$, $\mathcal{A} = \mathcal{B}([0,1])$, $\mathbb{P} = \text{Lebesgue}$. Définissons $X(\omega) := \sup\{s \in \mathbb{R} : F(s) \leq \omega\}$. Alors

$$F_X(x) = \mathbb{P}[X(\omega) \leq x] = \mathbb{P}[\{\omega : F(x) \geq \omega\}] = F(x)$$

□

1.7 Fonction caractéristique

Dans cette section, les variables aléatoires sont à valeurs dans $\mathbb{R}^d$.

Définition 1.24. Si X est une variable aléatoire à valeurs dans $\mathbb{R}^d$. La **fonction caractéristique** ϕ_X de X est la fonction

$$\phi_X(\xi) := \mathbb{E}[e^{i\xi X}], \forall \xi \in \mathbb{R}^d$$

Remarque.

$$\phi_X(\xi) = \int e^{i\xi x} d\,\mathbb{P}_X[x] \text{ est la transformée de Fourier de } \mathbb{P}_X$$

Exemple 1.25. Soit X une variable aléatoire gaussienne $\mathcal{N}(0,\sigma^2)$, on a

$$\phi_X(\sigma) = \exp\left[-\frac{\sigma^2\xi^2}{2}\right]$$

Démonstration. On se ramène facilement à $\sigma^2 = 1$. On souhaite calculer

$$\phi_X(\xi) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{i\xi x} e^{-x^2/2} dx$$

On a alors

$$\begin{aligned}
\phi'_X(\xi) &= \frac{i}{\sqrt{2\pi}} \int_{\mathbb{R}} x e^{i\xi x} e^{-x^2/2} dx \\
&= \frac{i}{\sqrt{2\pi}} \left(\left[-e^{i\xi x} e^{-x^2/2}\right]_{-\infty}^{+\infty} + \int_{\mathbb{R}} i\xi e^{i\xi x} e^{-x^2/2} dx \right) \\
&= -\xi \phi_X(\xi)
\end{aligned}$$

Donc ϕ_X vérifie $\phi_X(0) = 1$ et $\phi'_X(\xi) = -\xi\phi_X(\xi)$. D'où $\phi_X(\xi) = e^{-\xi^2/2}$ □

Théorème 1.26. *Soient X et Y deux variables aléatoires telles que $\phi_X = \phi_Y$, alors X et Y on la même loi.*

Démonstration. On commence par $d = 1$. Pour une mesure μ, on note

$$\hat{\mu}(\xi) := \int e^{i\xi x} d\mu(x), \forall \xi \in \mathbb{R}$$

Introduisons g_σ pout la densité de la variable aléatoire $\mathcal{N}(0,\sigma^2)$, i.e :

$$\forall x \in \mathbb{R}, g_\sigma(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2}\right]$$

Enfin,

$$\forall x \in \mathbb{R}, g_\sigma * \mu(x) := \int g_\sigma(x-y)d\mu(y)$$

Étape 1 : $g_\sigma * \mu$ est determinée par $\hat{\mu}$:

$$\begin{aligned}
g_\sigma * \mu(x) &= \int_{\mathbb{R}} g_\sigma(x-y)d\mu(y) \\
&\overset{1.25}{=} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{i(x-y)\xi} g_{1/\sigma}(\xi)d\xi \right) d\mu(y) \\
&\overset{\text{Fubini}}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{ix\xi} g_{1/\sigma}(\xi) \left(\int_{\mathbb{R}} e^{-iy\xi} d\mu(y) \right) d_\xi \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{ix\xi} g_{1/\sigma}(\xi)\hat{\mu}(-\xi)d\xi
\end{aligned}$$

Étape 2 : μ est déterminée par les fonctions $g_\sigma * \mu$, $(\sigma > 0)$

μ est caractérisée par les quantités $\int \varphi(x)d\mu(x), \forall \varphi$ continue bornée. On va donc montrer que

$$\lim_{\sigma \to 0} \int_{\mathbb{R}} \varphi(x) g_\sigma * \mu(x)dx = \int \varphi(x)d\mu(x)$$

En effet,

$$\int_{\mathbb{R}} \varphi(x) g_\sigma * \mu(x)dx = \int_{\mathbb{R}} \varphi(x) \int_{\mathbb{R}} g_\sigma(x-y)d\mu(y)dx \overset{\text{Fubini}}{=} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \varphi(x) g_\sigma(x-y)dx \right) d\mu(y)$$

Or $\int_{\mathbb{R}} g_\sigma(x-y)dx = 1$ et $\forall \epsilon > 0, \displaystyle\int_{|x-y|>\epsilon} g_\sigma(x-y)dx \xrightarrow{\sigma \to 0} 0$. On obtient, puisque φ est continue bornée,

$$\lim_{\sigma \to 0} \int_{\mathbb{R}} \varphi(x) g_\sigma(x-y)dy = \varphi(y), \forall y \in \mathbb{R}$$

puisque φ est bornée par $\|\varphi\|_\infty$, on peut vérifier que

$$\left| \int_{\mathbb{R}} \varphi(x) g_\sigma(x-y)dx \right| < \|\varphi\|_\infty$$

Par le théorème de convergence dominée, on obtient le résultat.

Étape 3 : d quelconque

On utilise la même preuve avec

$$g_\sigma^{(d)}(x_1, ..., x_d) = g_\sigma(x_1)...g_\sigma(x_d)$$

□

Proposition 1.27. *Soit $X = (X_1, ..., X_d)$ une variable aléatoire à valeurs dans $\mathbb{R}^d$ et dans L^2. Alors ϕ_X est de classe $\mathcal{C}^2$ et*

$$\phi_X(\xi) = 1 + i \sum_{i=1}^{d} \mathbb{E}[X_i]\xi_i - \frac{1}{2} \sum_{i,j=1}^{d} \mathbb{E}[X_i X_j]\xi_i \xi_j + o(\|\xi\|^2)$$

où $\xi = (\xi_1, ... \xi_d)$.

Remarque. On peut récupérer les espérences en dérivant. Si $d = 1$,

$$\phi_X(\xi) = 1 + i\,\mathbb{E}[X]\xi - \frac{1}{2}\,\mathbb{E}[X^2]\xi^2 + o(\xi^2)$$

Démonstration. En dérivant sous le signe intégrale ($\left|iX_j e^{i\xi x}\right| \leq |X_j|$), on obtient

$$\frac{\partial \phi_X}{\partial \xi_j}(\xi) = i\,\mathbb{E}[X_j e^{i\xi\cdot X}]$$

De même, puisque $\mathbb{E}[\left|X_j X_k e^{i\xi\cdot X}\right|] \leq \mathbb{E}[|X_j||X_k|] \leq \sqrt{\mathbb{E}[X_j^2]\,\mathbb{E}[X_k^2]} < \infty$,

$$\frac{\partial^2 \phi_X}{\partial \xi_j \partial \xi_k}(\xi) = -\,\mathbb{E}[X_j X_k e^{i\xi\cdot X}]$$

Par le théorème de continuité sous le signe intégrale, ces quantités sont continues, donc ϕ_X est de classe $\mathcal{C}^2$. L'égalité est juste le développement d'ordre 2 de Taylor (qui existe puisque ϕ_X est de classe $\mathcal{C}^2$). □

1.8 Fonction génératrice

Dans le cas d'une variable aléatoire à valeurs entières, on utilise plutôt les fonctions génératrices que caractéristiques.

Définition 1.28. Soit X une variable aléatoire à valeurs entières, la **fonction génératrice** de X est donnée par

$$G_X(r) := \mathbb{E}[r^X] = \sum_{k=0}^{\infty} \mathbb{P}[X = k] r^k, \forall r \in [0,1]$$

Proposition 1.29. *La fonction génératrice de X satisfait.*

1. *G_X est continue sur $[0,1]$ et $G_X(0) = \mathbb{P}[X = 0]$ et $G_X(1) = 1$.*
2. *G_X caractérise la loi de X.*
3. *G_X est dérivable à gauche en 1 et $G'_X(1) = \mathbb{E}[X]$*

Remarque. On n'a pas besoin de supposer que X est intégrable pour avoir 3., c'est juste que $G'_X(1)$ peut valoir l'infini.

$$G_X^{(k)}(1) = \mathbb{E}[X(X-1)...(X-k+1)]$$

Démonstration. Puisque $G_X(1) = \sum_{k=0}^{\infty} \mathbb{P}[X = k] = 1 < \infty$, G_X est une série entière de rayon de convergence au moins 1.

1. suit par continuité sous le signe intégrale
2. par développement de Taylor $\left(\mathbb{P}[X = n] = \frac{G_X^{(n)}(0)}{n!}\right)$
3. par différentiation sous le signe intégrale

□

Chapitre 2

Indépendance

Cours 6

2.1 Évènements indépendants

Considérons un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$

Définition 2.1. $A, B \in \mathcal{A}$ sont **indépendants** si $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$

Remarque. $\mathbb{P}[A|B] := \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}$ (probabilité de A sachant B). Si A et B sont indépendants, $\mathbb{P}[A|B] = \mathbb{P}[A]$. "Le fait que B a lieu ne donne pas d'information sur le fait que A a lieu."

Exemple 2.2.

1. Lancers de deux dés : $\Omega = \{1, ..., 6\}^2, \mathbb{P}[\{\omega\}] = \frac{1}{36}$

$$A = \{6\} \times \{1, ..., 6\} \text{ et } B = \{1, ..., 6\} \times \{6\}$$

   Sont indépendants : $\mathbb{P}[A] \cap \mathbb{P}[B] = \mathbb{P}[\{(6, 6)\}] = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = \mathbb{P}[A]\,\mathbb{P}[B]$

2. Lancer d'un dé, $A = \{1, 2\}$ et $B = \{1, 3, 5\}$ sont indépendants

Définition 2.3. n évènements $A_1, ..., A_n$ sont **indépendants** si

$$\forall j_1, ...j_k \in \{1, ..., n\}, \mathbb{P}[A_{j_1} \cap ... \cap A_{j_k}] = \mathbb{P}[A_{j_1}] ... \mathbb{P}[A_{j_k}]$$

Remarque. Il ne suffit pas de vérifier $\mathbb{P}[A_1 \cap ... \cap A_n] = \mathbb{P}[A_1] ... \mathbb{P}[A_n]$

2.2 Variables aléatoires et tribus indépendantes

Définition 2.4. Les tribus $\mathcal{B}_1, ..., \mathcal{B}_n \subset \mathcal{A}$ sont **indépendantes** si

$$\forall B_1 \in \mathcal{B}_1, ..., \forall B_n \in \mathcal{B}_n, \mathbb{P}[B_1 \cap ... \cap B_n] = \mathbb{P}[B_1] ... \mathbb{P}[B_n]$$

Définition 2.5. Les variables aléatoires $X_1, ...X_n$ a valeurs dans $(E_1, \mathcal{E}_1), ..., (E_n, \mathcal{E}_n)$ sont **indépendantes** si $\sigma(X_1), ..., \sigma(X_n)$ sont indépendants.

Lemme 2.6.

1. *$X_1, ..., X_n$ indépendants si $\forall A_1 \in \mathcal{E}_1, ..., \forall A_n \in \mathcal{E}_n$,*

$$\mathbb{P}[\{X_1 \in A_1\} \cap ... \cap \{X_n \in A_n\}] = \mathbb{P}[X_1 \in A_1] ... \mathbb{P}[X_n \in A_n]$$

2. *Soient $\mathcal{B}_1, ..., \mathcal{B}_n$ n tribus indépendantes sur tout X_1 $\mathcal{B}_1$-mesurable,…,X_n $\mathcal{B}_n$-mesurable, alors $X_1, ..., X_n$ sont indépendants*

   Démonstration. $\sigma(X_i) \subset \mathcal{B}_i$ pour tout i □

3. *$A_1, ..., A_n$ sont indépendants si $\sigma(\{A_1\}), ..., \sigma(\{A_n\})$ sont indépendants.*

Exemple 2.7. $X = (X_1, ..., X_d)$ uniforme sur l'ensemble $\mathcal{N} = \{(\pm 1, ... \pm 1) \subset \mathbb{R}^d\}$. En particulier, pour tout $x \in \mathcal{N}, \mathbb{P}[X = x] = \frac{1}{|\mathcal{N}|} = \frac{1}{2^d}$.

Les variables aléatoires $X_1, ..., X_d$ sont indépendantes. $A_1 \subset \{\pm 1\}, ... A_n \subset \{\pm 1\}$.

$$\mathbb{P}[\{X_1 \in A_1\} \cap ... \cap \{X_n \in A_n\}] = \mathbb{P}[X \in A] \text{ où } A = \{x = (x_1, ..., x_d) \in \mathcal{N} : x_i \in A_i\}$$

$$\begin{aligned}
\mathbb{P}[X \in A] = \frac{|A|}{2^d} = \prod_{i=1}^d \frac{|A_i|}{2} &= \prod_{i=1}^d \frac{2^{d-1}|A_i|}{2^d} \\
&= \prod_{i=1}^d \mathbb{P}[X \in A_i] \text{ où } A_i = \{x \in \mathcal{N} : x_i \in A_i\} \\
&= \prod_{i=1}^d \mathbb{P}[X_i \in A_i]
\end{aligned}$$

Proposition 2.8. *Les variables aléatoires $X_1, ..., X_n$ sont indépendantes si et seulement si*

$$\mathbb{P}_{(X_1, ..., X_n)} = \mathbb{P}_{X_1} \otimes ... \otimes \mathbb{P}_{X_n}$$

En particulier, pour tout $f_1, ..., f_n$ intégrables par rapport à $\mathbb{P}_{X_1}, ..., \mathbb{P}_{X_n}$,

$$\mathbb{E}\left[\prod_{i=1}^n f_i(X_i)\right] = \prod_{i=1}^n \mathbb{E}\left[f_i(X_i)\right]$$

Démonstration. Soient $F_1 \in \mathcal{E}_1, ..., F_n \in \mathcal{E}_n$.

$$\begin{aligned}
\mathbb{P}_{(X_1, ..., X_n)}[F_1 \times ... \times F_n] &= \mathbb{P}[(X_1, ..., X_n) \in F_1 \times ... \times F_n] \\
&= \mathbb{P}[\{X_1 \in F_1\} \cap ... \cap \{X_n \in F_n\}] \\
&= \mathbb{P}[X_1 \in F_1] ... \mathbb{P}[X_n \in F_n] \\
&= \mathbb{P}_{X_1}[F_1] ... \mathbb{P}_{X_n}[F_n] \\
&= \mathbb{P}_{X_1} \otimes ... \otimes \mathbb{P}_{X_n}[F_1 \times ... \times F_n]
\end{aligned}$$

Ces identités concluent $\Longleftarrow$. Pour $\Longrightarrow$, on observe que $\mathbb{P}_{(X_1, ..., X_n)}$ et $\mathbb{P}_{X_1} \otimes ... \otimes \mathbb{P}_{X_n}$ coïncident sur $\mathcal{C} = \mathcal{E}_1 \times ... \times \mathcal{E}_n$. $\mathcal{C}$ et stable par intersection et cérifie $\sigma(\mathcal{C}) = \mathcal{E}_1 \otimes ... \otimes \mathcal{E}_n$, donc par le Théorème 0.8, on conclut $\mathbb{P}_{(X_1, ..., X_n)} = \mathbb{P}_{X_1} \otimes ... \otimes \mathbb{P}_{X_n}$

Pour la deuxième assertion :

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^n f_i(X_i)\right] &= \int \prod_{i=1}^n f_i(x_i) d\,\mathbb{P}_{(X_1, ..., X_n)}(x_1, ..., x_n) \\
&= \int \prod_{i=1}^n f_i(x_i) d\,\mathbb{P}_{X_1}(x_1) ... d\,\mathbb{P}_{X_n}(x_n)
\end{aligned}$$

si les f_i sont positifs, alors on a

$$\begin{aligned}
\int \prod_{i=1}^n f_i(x_i) d\,\mathbb{P}_{X_1}(x_1) ... d\,\mathbb{P}_{X_n}(x_n) &= \prod_{i=1}^n \int f_i(x_i) d\,\mathbb{P}_{X_i}(x_i) \\
&= \prod_{i=1}^n \mathbb{E}\left[f_i(X_i)\right]
\end{aligned}$$

si les f_i sont intégrables, alors

$$\mathbb{E}\left[\prod_{i=1}^n |f_i(X_i)|\right] = \prod_{i=1}^n \mathbb{E}\left[|f_i(X_i)|\right] < \infty$$

donc $\prod_{i=1}^n f_i(X_i)$ est intégrable. On peut donc appliquer Fubini pour obtenir le résultat. □

Remarque. Si X et Y sont intégrables et indépendantes, alors XY l'est aussi.

Remarque 2.9.

1. Soit X indépendant de X sur $(E, \mathcal{E})$. Alors $\mathbb{P}_X(A) \in \{0,1\}, \forall A \in \mathcal{E}$.
   En effet, $\mathbb{P}_X(A) = \mathbb{P}_X(A \cap A) = \mathbb{P}[\{x \in A\} \cap \{x \in A\}] = \mathbb{P}[X \in A]^2 = \mathbb{P}_X(A)^2$, donc $\mathbb{P}_X(A) \in \{0,1\}$.
   On en déduit qu'il existe $x \in E$ tel que $\mathbb{P}[X = x] = 1$.
2. $Y_1, ..., Y_n$ n variables aléatoires sur $(E_1, \mathcal{E}_1), ..., (E_n, \mathcal{E}_n)$, sur $\Omega = E_1 \times ... \times E_n$. $\mathcal{A} = \mathcal{E}_1 \otimes ... \otimes \mathcal{E}_n$, $\mathbb{P} = \mathbb{P}_{Y_1} \otimes ... \otimes \mathbb{P}_{Y_n}$. Alors les (Y_i) sont indépendantes sur $(\Omega, \mathcal{A}, \mathbb{P})$.
3. Soient $X_1, ..., X_n$ n variables alétoires indépendantes.

   $$\mathbb{E}[X_1 ... X_n] = \mathbb{E}[X_1] ... \mathbb{E}[X_n]$$

   En particulier, si X et Y sont indépendants et dans L^1,

   $$Cov(X,Y) = 0$$

   Remarque. $Cov(X,Y) = 0 \not\Rightarrow X$ et Y sont indépendants.
   $X \sim \mathcal{N}(0,1)$ et ϵ tq $\mathbb{P}[\epsilon = 1] = \mathbb{P}[\epsilon = -1] = \frac{1}{2}$. Supposons X et ϵ indépendants.

   $$\mathrm{Cov}(X, \epsilon X) = \mathbb{E}[\epsilon X^2] - \mathbb{E}[X]\,\mathbb{E}[\epsilon X] = \mathbb{E}[\epsilon](\mathbb{E}[X^2] - \mathbb{E}[X]^2) = 0$$

   Si X et ϵX sont indépendants, alors $|X|$ et $|\epsilon X|$ sont indépendants, et donc $|X|$ est indépendant de lui-même. Cela impliquerait que $|X|$ est constante presque partout, ce qui n'est pas le cas. On a donc que X et ϵX ne sont pas indépendants.
4. Si X, Y, Z sont indépendantes, alors $f(X)$ est indépendant de $g(Y,Z)$. Ce principe se généralise à n variables aléatoires indépendantes
5. Si $X_1, ..., X_n$ indépendantes et à densité $p_1, ..., p_n$, alors $(X_1, ..., X_n)$ est une variable aléatoire à densité, de densité égale à $p(x_1, ..., x_n) = p(x_1) ... p_n(x_n)$
6. En résumé, si les variables aléatoires sont à valeurs réelles, les propriétés suivantes sont équivalentes :
   (a) $X_1, ... X_n$ indépendantes
   (b) $\mathbb{P}[X_1 \leq a_1, ..., X_n \leq a_n] = \mathbb{P}[X_1 \leq a_1] ... \mathbb{P}[X_n \leq a_n] \forall a_1, ..., a_n \in \mathbb{R}$
   (c) $\forall f_1, ..., f_n, \mathbb{E}[f_1(X_1), ... f_n(X_n)] = \prod_{i=1}^n \mathbb{E}[f_i(X_i)]$
   (d) $\phi_{(X_1, ..., X_n)}(\xi_1, ..., \xi_n) = \phi_{X_1}(\xi_1) ... \phi_{X_n}(\xi_n)$

   Démonstration.
   (a) $\Rightarrow$ *(b)* $\mathbb{P}_{(X_1, ..., X_n)} = \mathbb{P}_{X_1} \otimes ... \otimes \mathbb{P}_{X_n}$
   (b) $\Rightarrow$ *(c)* Théorème 0.8 + Proposition 2.8
   (c) $\Rightarrow$ *(d)* $f_j(t) = e^{i\xi_j t}$
   (d) $\Rightarrow$ *(a)* Unicité de la fonction caractéristique car $\phi_{X_1}(\xi_1) ... \phi_{X_n}(\xi_n)$ est la fonction caractéristique de $\mathbb{P}_{X_1} \otimes ... \otimes \mathbb{P}_{X_n}$

   □

7. $\forall \mathcal{C}_1, ..., \mathcal{C}_n$ stables par intersection finie, $(\sigma(\mathcal{C}_1), ..., \sigma(\mathcal{C}_n))$ sont indépendantes si et seulement si

   $$\forall A_1 \in \mathcal{C}_1, ..., \forall A_n \in \mathcal{C}_n, \mathbb{P}[A_1 \cap ... \cap A_n] = \mathbb{P}[A_1] ... \mathbb{P}[A_n]$$

2.3 Loi du tout ou rien

Cours 7

Définition 2.10. Soit $(\mathcal{B}_i)_{i \in I}$ de sous-tribu de $\mathcal{A}$. On dit que les $\mathcal{B}_i$ sont **indépendantes** si et seulement si $\forall J \subset I$ fini, les $(\mathcal{B}_i)_{i \in J}$ sont indépendantes. Les variables aléatoires $(X_i)_{i \in I}$ sont **indépendantes** si et seulement si $(\sigma(X_i))_{i \in I}$ est une famille de tribus indépendantes.

Exemple 2.11 (construction d'une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires indépendantes de la loi $Ber(\frac{1}{2})$)**.**
$\Omega = [0,1), \mathcal{A} = \mathcal{B}([0,1)), \mathbb{P} =$ Lebesgue

Pour $\omega \in [0,1)$, on écrit $\omega := \sum_{n=0}^{\infty} \frac{X_n(\omega)}{2^n}$ où $X_n(\omega) \in \{0,1\}$, où les $(X_n(\omega))$ ne stationnent pas à 1. Les $(X_n)_{n \in \mathbb{N}}$ sont indépendants de loi $Ber(\frac{1}{2})$.

Rappel.

$$\mathbb{P}_X[A] \in \{0,1\}, \forall A \in \mathcal{E}$$
$$\mathbb{P}[X \in B] \in \{0,1\}, \forall B \in \mathcal{E}$$
$$\mathbb{P}[C] \in \{0,1\}, \forall C \in \sigma(X)$$

Définition 2.12. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes. La **tribu asymptotique** est définie par

$$\mathcal{B}_\infty := \cap_{n\in\mathbb{N}}\sigma(X_n, X_{n+1}, ...)$$

Remarque. Tout ce qui ne dépend d'un nombre fini de variables (X_n). Par exemple :

1. $X_n \sim Ber(p)$. indexons les arêtes de $\mathbb{Z}^2$ par $(e_1, ..., e_n, ...)$ et on considère le graphe ω composé des sommets de $\mathbb{Z}^2$, et des arrêtes e_i tq $X_i = 1$. $A = \{$il existe une composante connexe infinie dans $\omega\}$.
2. $(X_i) \sim Ber(\frac{1}{2})$,
$$A = \{\text{il existe une infinité de } i \text{ t.q. } X_i = 1\}$$
3. (X_i) à valeurs réelles.
$$A = \{\lim_n X_n \text{ existe et appartenant à } B\} \text{ où } B \in \mathcal{B}(\mathbb{R})$$
$$X = \liminf X_n$$

Théorème 2.13 (loi du tout ou rien, loi du $0-1$ de Kolmogorov)**.** *Soit (X_n) une famille de variables aléatoires indépendantes, alors*

$$\mathbb{P}[A] \in \{0,1\}, \forall A \in \mathcal{B}_\infty$$

Démonstration. Posons

$$\mathcal{B}_n := \sigma(X_{n+1}, ..., ...)$$
$$\mathcal{D}_n := \sigma(X_1, ..., X_n)$$

Étape 1 : $\mathcal{B}_n$ et $\mathcal{D}_n$ sont indépendants :

Soit $D \in \mathcal{D}_n$ et $B \in \sigma(X_{n+1}, ..., X_m)$ où $m > n$. Puisque $(X_1, ..., X_n)$ est indépendant de $(X_{n+1}, ..., X_m)$ par la remarque 2.9 (iv), donc $\mathbb{P}[B \cap D] = \mathbb{P}[B]\,\mathbb{P}[D]$. Puisque $\mathcal{C} = \bigcup_{m>n}\sigma(X_{n+1}, ..., X_m)$ est stable par intersection finie et que $\sigma(\mathcal{C}) = \mathcal{B}_n$, on déduit de la remarque 2.9 (vii) que $\mathcal{B}_n$ et $\mathcal{D}_n$ sont indépendants.

Étape 2 : $\mathcal{B}_\infty$ est indépendant de $\mathcal{B}_\infty$

On a que $\mathcal{D}_n$ est indépendant de $\mathcal{B}_n \supset \mathcal{B}_\infty$, En particulier, $\forall A \in \bigcup_n D_n, \forall B \in \mathcal{B}_\infty$, $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$. Puisque $\mathcal{C} = \bigcup_n \mathcal{D}_n$ est stable par intersection finie et que $\sigma(\mathcal{C}) \supset \mathcal{B}_\infty$, on déduit de la remarque 2.9 (iii) que $\forall A, B \in \mathcal{B}_\infty$, $\mathbb{P}[A \cap B] = \mathbb{P}[A]\,\mathbb{P}[B]$.

Étape 3 :

$\forall B \in \mathcal{B}_\infty, \mathbb{P}[B] \in \{0,1\}$. On applique à $A = B$. □

Exemple 2.14.

1. Soit (X_n) des variables aléatoires indépendantes. $X = \limsup \dfrac{X_1 + ... + X_n}{n}$ est $\mathcal{B}_\infty$ mesurable et donc constante presque sûrement.
2. Toute variable aléatoire B_∞ mesurable est constante presque sûrement.

Démonstration. Par classes monotones. □

2.4 Lemme de Borel-Cantelli

Lemme 2.15 (Borel-Cantelli)**.** *Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements.*

1. *Si $\sum_{n\in\mathbb{N}} \mathbb{P}[A_n] < \infty$, alors $\mathbb{P}[\exists\infty$ de n t.q A_n a lieu$] = 0$*
2. *Si $\sum_{n\in\mathbb{N}} \mathbb{P}[A_n] = \infty$, et (A_n) sont indépendants, alors $\mathbb{P}[\exists\infty$ de n t.q A_n a lieu$] = 1$*

Remarque. Dans *2.*, il est important de supposer que les (A_n) sont indépendants. Posons $A_n = A, \forall n$, avec $\mathbb{P}[A] \in (0,1)$.

Démonstration.

1. $$\mathbb{E}[\sum_{n\in\mathbb{N}} \mathbb{1}_{A_n}] = \sum_{n\in\mathbb{N}} \mathbb{P}[A_n] < \infty$$

   Donc $\sum_{n\in\mathbb{N}} \mathbb{1}_{A_n}$ est finie presque sûrement, donc presque sûrement il existe un nombre fini de n t.q. A_n a lieu.

2. Fixons $k \in \mathbb{N}$,

   $$\begin{aligned}\mathbb{P}[\bigcup_{n=k}^{N} A_n] &= 1 - \mathbb{P}[\bigcap_{n=k}^{N} A_n^c] \\ &\overset{\text{ind.}}{=} 1 - \prod_{n=k}^{N} \mathbb{P}[A_n^c] \\ &= 1 - \prod_{n=k}^{N} (1 - \mathbb{P}[A_n])\end{aligned}$$

   en laissant N tendre vers l'infini, on obtient :

   $$\mathbb{P}[\bigcup_{n\geq k} A_n] = 1 - \prod_{n\geq k}(1-\mathbb{P}[A_n]) \overset{(*)}{=} 1$$
   $$\implies \mathbb{P}[\exists\infty \text{ de } n \text{ tq } A_n] = \mathbb{P}[\bigcap_k \bigcup_{n\geq k} A_n] = \lim_{k\to+\infty} \mathbb{P}[\bigcup_{n\geq k} A_n] = 1$$

   $(*)$ car si $\sum_{n\in\mathbb{N}} \mathbb{P}[A_n] = \infty$, comme $\log(1-x) \leq -x$, on a

   $$\log\left(\prod_{n\geq k}(1-\mathbb{P}[A_n])\right) = \sum_{n\geq k}\log(1-\mathbb{P}[A_n]) \leq \sum_{n\geq k} -\mathbb{P}[A_n] = -\infty$$

   et donc

   $$\prod_{n\geq k}(1-\mathbb{P}[A_n]) = 0$$

□

Exemple 2.16 (Un oiseau bourré ne retrouve pas son chemin)**.**

Soit $S_n = \sum_{i=1}^n X_i$ où les (X_i) sont indépendants de loi uniforme sur $F = \{-1,1\}^d$.

Par l'exemple 2.7, $X_i = (X_i^{(1)}, X_i^{(2)}, ..., X_i^{(d)})$, alors les $(X_i^{(j)})$ sont indépendants de loi $2Ber(\frac{1}{2}) - 1$.

$$\mathbb{P}[(S_n) \text{ visite une infinité de fois } (0,...,0)] = 0$$

pour $d \geq 3$

Démonstration. Montrons que $\mathbb{P}[S_{2n+1} = 0] = 0$ et $\mathbb{P}[S_{2n}] \sim \frac{1}{(\pi n)^{d/2}}$.

$$\begin{aligned}\mathbb{P}[S_{2n+1} = 0] &= 0 \text{ par parité} \\ \mathbb{P}[S_{2n} = 0] &= \mathbb{P}\left[\forall j = 1, ..., d, \sum_{i=1}^{2n} X_i^{(j)} = 0\right] \\ &= \prod_{j=1}^{d} \mathbb{P}\left[\sum_{i=1}^{2n} X_i^{(j)} = 0\right] \text{ par indépendence des } (X_i^{(j)}) \text{ pour les } j \\ &= \left(\binom{2n}{n}\frac{1}{2^{2n}}\right)^d \overset{\text{Stirling}}{\sim} \left(\frac{1}{\sqrt{\pi n}}\right)^d\end{aligned}$$

Par le lemme de Borel-Cantelli, $\mathbb{P}[\exists$ une infinité de n t.q. $S_n = 0] = 0$, puisque $\sum_{n=0}^{\infty} \frac{1}{(\sqrt{\pi n})^d} < \infty$ pour $d \geq 3$. □

2.5 Sommes de variables aléatoires indépendantes

Définition 2.17 (convolution)**.** Soient μ et ν deux mesures de probabilité sur $\mathbb{R}^d$. La **convolution** de μ et ν est la mesure $\mu * \nu$ vérifiant $\forall \varphi$ mesurable positive,

$$\int_{\mathbb{R}^d} \varphi(z) d\mu * \nu(z) = \int_{\mathbb{R}_d} \int_{\mathbb{R}_d} \varphi(x+y) d\mu(x) d\nu(x)$$

Cours 8

Remarque. Si $d\mu(x) = f(x)dx$ et $d\nu(x) = g(y)dy$, alors $d\mu * \nu(z) = f * g(z)dz$.

Proposition 2.18. *Soient X et Y deux variables aléatoires de $\mathbb{R}^d$* ***indépendantes****. Alors :*

1. $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$
2. $\phi_{X+Y} = \phi_X \phi_Y$
3. *Si X et Y sont dans L^2, $K_{X+Y} = K_X + K_Y$.*
   En particulier, si $d = 1$, $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

Démonstration. 1. $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$

$$\begin{aligned} \mathbb{E}[\varphi(X+Y)] &= \int\int_{\mathbb{R}^d} \varphi(x+y) d\,\mathbb{P}_{(X,Y)}(x,y) \\ &= \int\int_{\mathbb{R}^d} \varphi(x+y) d\,\mathbb{P}_X(x) d\,\mathbb{P}_Y(y) \\ &= \int_{\mathbb{R}^d} \varphi(z) d(\mathbb{P}_X * \mathbb{P}_Y)(z) \end{aligned}$$

Puisque ceci est vrai pour tout φ mesurable, borné, on déduit que $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$

2. $\varphi(z) = e^{i\xi z}$

$$\phi_{X+Y}(\xi) = \mathbb{E}[\varphi(X+Y)] = \mathbb{E}[\varphi(X)\varphi(Y)] \overset{\text{indépendance}}{=} \mathbb{E}[\varphi(X)]\,\mathbb{E}[\varphi(Y)] = \phi_X(\xi)\phi_Y(\xi)$$

3.

$$\mathrm{Cov}(X_i + Y_i, X_j + X_j) = \mathrm{Cov}(X_i, X_j) + \underbrace{\mathrm{Cov}(X_i, Y_j)}_{=0 \text{ par indépendance}} + \underbrace{\mathrm{Cov}(Y_i, X_j)}_{=0 \text{ par indépendance}} + \mathrm{Cov}(Y_i, Y_j)$$

□

Exemple 2.19.

1. $X_1 \sim \mathcal{N}(0, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(0, \sigma_2^2)$, et X_1 et X_2 indépendantes, alors $X_1 + X_2 \sim \mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$
2. $X_1 \sim Poisson(\lambda_1)$, $X_2 \sim Poisson(\lambda_2)$, et X_1 et X_2 indépendantes, alors $X_1 + X_2 \sim Poisson(\lambda_1 + \lambda_2)$

Théorème 2.20 (Loi faible des grands nombres)**.** *Soit (X_n) une suite de variables aléatoires indépendantes de même loi, et dans L^2. Alors*

$$\mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}[X_1]\right)^2\right] \overset{n\to\infty}{\longrightarrow} 0$$

Remarque. On dit que $S_n = \frac{1}{n}\sum_{i=1}^n X_i$ converge en norme L^2 vers la constante $\mathbb{E}[X_1]$.

Démonstration. $\mathbb{E}[X_i] = \mathbb{E}[X_1]$ pour tout i.

$$\begin{aligned} \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}[X_1]\right)^2\right] &= \mathbb{E}\left[\left(\frac{1}{n}\sum_{i=1}^n X_i - \mathbb{E}[X_i]\right)^2\right] \\ &= \frac{1}{n^2}\mathrm{Var}(X_1 + ... + X_n) \\ &= \frac{1}{n^2}(\mathrm{Var}(X_1) + ... + \mathrm{Var}(X_n)) = \frac{\mathrm{Var}(X_1)}{n} \end{aligned}$$

qui tend bien vers 0 quand $n \to \infty$ □

Chapitre 3

Convergence des variables aléatoires

3.1 Notions de convergence

Question : Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variable aléatoires. Que voulons nous dire par $X_n \overset{n\to\infty}{\longrightarrow} X$?

Par la suite, dans 3.1, on suppose toujours que les (X_n) et X sont définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et sont à valeurs dans $\mathbb{R}^d$.

Définition 3.1 (notions de convergence)**.**
- (X_n) **converge presque sûrement** ($X_n \overset{p.s}{\to} X$) si
$$\mathbb{P}[\{\omega \in \Omega, \lim_{n\to\infty} X_n(\omega) \text{ existe et vaut } X(\omega)\}] = 1$$
- (X_n) **converge en norme** $L^p, (p \geq 1)$ ($X_n \overset{L^p}{\to} X$) si
$$\|X_n - X\|_p := \mathbb{E}[|X_n - X|^p]^{1/p} \overset{n\to\infty}{\longrightarrow} 0$$
- (X_n) **converge en probabilité** ($X_n \overset{(p)}{\to} X$) si
$$\forall \epsilon > 0, \lim_{n\to\infty} \mathbb{P}[|X_n - X| > \epsilon] = 0$$

Notation. $x \wedge y := \min(x, y)$, $x \vee y := \max(x, y)$

Proposition 3.2. *Soit $L^0 = L^0(\Omega, \mathcal{A}, \mathbb{P})$ l'espace des variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d$. Notons $L^0 = L^0/\sim$ où $X \sim Y$ si et seulement si $\mathbb{P}[X = Y] = 1$, Alors $d(X, Y) := \mathbb{E}[|X - Y| \wedge 1]$ définit une distance sur L^0, pour laquelle L^0 est complet. De plus $X_n \overset{(p)}{\to} X$ si et seulement si $d(X_n, X) \overset{n\to\infty}{\longrightarrow} 0$*

Démonstration. d est clairement une distance. De plus, pour $\epsilon < 1$,
$$\epsilon\,\mathbb{P}[|X - Y| > \epsilon] \leq \mathbb{E}[|X - Y| \wedge 1] \leq \mathbb{P}[|X - Y| > \epsilon] + \epsilon\,\mathbb{P}[|X - Y| \leq \epsilon] = \mathbb{P}[|X - Y| > \epsilon] + \epsilon$$

Il suit (quasi-) immédiatement que $(X_n) \overset{(p)}{\to} X$ ssi $d(X_n, X) \to 0$.

Montrons maintenant que (L^0, d) est complet.

Soit (X_n) une suite de Cauchy.

Étape 1 : Construction du candidat limite
Soit (n_k) t.q. $Y_k = X_{n_k}$ vérifie $d(Y_{k+1}, Y_k) \leq \frac{1}{2^k}$. $\mathbb{E}[\sum_{k=0}^{\infty} |y_{k+1} - Y_k| \wedge 1] \leq \sum_{k=0}^{\infty} \frac{1}{2^k} < \infty$ On en déduit que $\sum_{k=0}^{\infty} \min(|Y_{k+1} - Y_k|, 1) < \infty$ presque sûrement. Du coup $\sum_{k=0}^{\infty} |Y_{k+1} - Y_k| < \infty$ presque sûrement. On peut donc définir $X(\omega) = X_0(\omega) + \sum_{k=0}^{\infty} Y_{k+1}(\omega) - Y_k(\omega)$. La convergence de la série est due au fait qu'elle est absolument convergente *et* que $\mathbb{R}^d$ est complet.

Étape 2 : $d(X_1, X) \overset{n\to\infty}{\longrightarrow} 0$
On a $d(Y_k, X) = \mathbb{E}\left[\left|\sum_{j=k}^{\infty} Y_{j+1} - Y_j\right| \wedge 1\right] \leq \sum_{j=k}^{\infty} d(Y_j, Y_{j+1}) \leq \frac{1}{2^{k-1}}$. Donc $(X_{n_k}) \to X$. De plus une suite de Cauchy ayant un point d'accumulation est convergente. □

Proposition 3.3. *1. Si $X_n \overset{p.s}{\to} X$, alors $X_n \overset{(p)}{\to} X$*

2. Si $X_n \overset{L^p}{\to} X$, alors $X_n \overset{(p)}{\to} X$

3. Si $X_n \overset{(p)}{\to} X$, alors $\exists (n_k) : X_{n_k} \overset{p.s}{\to} X$

4. Si $X_n \overset{(p)}{\to} X$ et $\forall n \in \mathbb{N}, \mathbb{E}[|X_n|^r] < C$, alors $X_n \overset{L^p}{\to} X, \forall p < r$

Démonstration.

1. $d(X_n, X) = \mathbb{E}[|X_n - X| \wedge 1] \to 0$ par convergence dominée.
2. $d(X_n, X)) = \mathbb{E}[|X_n - X| \wedge 1] \le \mathbb{E}[|X_n - X|] \overset{\text{Hölder}}{\le} \mathbb{E}[|X_n - X|^p]^{1/p} \overset{n\to\infty}{\longrightarrow} 0$.
3. Les (Y_k) de la preuve précédente vérifient $Y_k \overset{p.s}{\to} X$.
4. Par le lemme de Fatou, $\mathbb{E}[(\liminf |X_{n_k}|)^r] \le \liminf \mathbb{E}[|X_{n_k}|^r] \le C$. Hors si $X_n \overset{p.s}{\to} X$, alors $(X_{n_k}) \overset{p.s}{\to} X$ donc $\liminf |X_{n_k}|^r = |X|^r$ d'où $\mathbb{E}[|X|^r] \le C$. Soit $\epsilon > 0$, on a

$$\begin{aligned}\mathbb{E}[|X_n - X|^p] &\le \epsilon^p \, \mathbb{P}[|X_n - X| \le \epsilon] + \mathbb{E}[|X_n - X|^p \mathbb{1}_{|X_n - X| > \epsilon}] \\ &\le \epsilon^p + \mathbb{E}[|X_n - X|^r]^{p/r} \, \mathbb{E}[(\mathbb{1}_{|X_n - X|>\epsilon})]^{1-\frac{p}{r}} \\ &= \epsilon^p + \mathbb{E}[|X_n - X|^r]^{p/r} \, \mathbb{P}[|X_n - X| > \epsilon]^{1-p/r} \\ &\le \epsilon^p + (2c^{1/r})^p \, \mathbb{P}[|X_n - X| > \epsilon]^{1-p/r}\end{aligned}$$

d'où $\limsup \mathbb{E}[|X_n - X|^p] \le \epsilon^p$.

Cours 9 □

Exemple 3.4 (approximation polynomiale)**.** Soit $f \in \mathcal{C}^0([0,1])$, posons pour $n \in \mathbb{N}$,

$$f_n(x) = \sum_{k=0}^{n} f(\frac{k}{n}) \binom{n}{k} x^k (1-x)^{n-k} \qquad \text{(polynôme de degré } n\text{)}$$

On a $\|f_n - f\|_\infty \overset{n\to\infty}{\longrightarrow} 0$. En effet, soit $x \in [0,1]$ et $\epsilon > 0$, et $X_{n,x} \sim \frac{1}{n}\mathrm{Bin}(n,x) \sim \frac{Y_1 + ... + Y_n}{n}$, où $Y_i \overset{i.i.d}{\sim} \mathrm{Ber}(x)$, alors on a

$$\begin{aligned}|f_n(x) - f(x)| &= |\mathbb{E}[f(X_{n,x}) - f(x)]| \\ &\le \mathbb{E}\left[\left|f(\frac{Y_1 + ... + Y_n}{n}) - f(x)\right|\right] \\ &\le 2\|f\|_\infty \, \mathbb{P}\left[\left|\frac{Y_1 + ... + Y_n}{n} - x\right| > \epsilon\right] + \eta(f,\epsilon) \, \mathbb{P}\left[\left|\frac{Y_1 + ... + Y_n}{n} - x\right| \le \epsilon\right]\end{aligned}$$

où $\eta(f,\epsilon)$ est le module de continuité de f.

Rappel. Le **module de continuité** d'une fonction f est défini par

$$\eta(f,\epsilon) := \sup\{\eta \in \mathbb{R} \mid \forall x, y \in [0,1], |x - y| < \eta \implies |f(x) - f(y)| < \epsilon\}$$

En résumé,

$$|f_n(x) - f(x)| \le 2\|f\|_\infty \, \mathbb{P}\left[\left|\frac{Y_1 + ... + Y_n}{n} - x\right| > \epsilon\right] + \eta(f,\epsilon)$$

Pour la loi faible des grands nombres, $\frac{Y_1 + ... + Y_n}{n}$ converge vers x en norme L^2 et donc en probabilité. Donc

$$\mathbb{P}\left[\left|\frac{Y_1 + ... + Y_n}{n} - x\right| > \epsilon\right] \overset{n\to\infty}{\longrightarrow} 0$$

Théorème 3.5 (loi forte des grands nombres)**.** *Soit (X_n) une suite de variables aléatoires indépendantes, identiquement distribuées et intégrables. Alors*

$$\frac{X_1 + ... + X_n}{n} \underset{n\to\infty}{\overset{p.s.}{\longrightarrow}} \mathbb{E}[X_1]$$

Remarque 3.6.

1. L'hypothèse est optimale puisque il faut que X_i soit intégrable pour parler d'espérance.
2. Si $X_1 \ge 0$ p.s. et $\mathbb{E}[X_1] = +\infty$, alors il est facile de montrer que $\frac{1}{n}(X_1 + ... + X_n) \underset{n\to\infty}{\overset{p.s.}{\longrightarrow}} \infty$

Démonstration. Étape 1 : préparation :
Pour $c > 0$, et $a < \mathbb{E}[X_1 \mathbb{1}_{X_1 < c}]$. Définissons $Y_n := (X_n - a)\mathbb{1}_{X_n < c}$ et $S_n := Y_1 + ... + Y_n$ et $S_0 := 0$. On montrera que $(*)$ $\inf S_n > -\infty$ p. s. (i.e. S_n est minorée presque sûrement) Remarquons que $(*)$ implique

$$\inf \sum_{k=1}^{n} (X_k - a)\mathbb{1}_{X_k < c} > -\infty \text{ p.s}$$

On en déduit $\liminf_n \frac{1}{n}\sum_{k=1}^{n} X_k \mathbb{1}_{X_k < c} \geq a$ p.s., donc $\liminf_n \frac{1}{n}\sum_{k=1}^{n} X_k \mathbb{1}_{X_k < c} - \mathbb{E}[X_1 \mathbb{1}_{X_1 > c}] \geq a - \mathbb{E}[X_1 \mathbb{1}_{X_1 > c}]$ p.s. En laissant tendre a vers $\mathbb{E}[X_1 \mathbb{1}_{X_1 > c}]$ (en fait (a_n) une suite, pour avoir la dénombrabilité), on obtient

$$\liminf_n \frac{1}{n} \sum_{k=1}^{n} X_k \mathbb{1}_{X_k < c} - \mathbb{E}[X_1 \mathbb{1}_{X_1 > c}] \geq 0 \text{ p.s}$$

d'où

$$\liminf_n \frac{1}{n} \sum_{k=1}^{n} X_k \geq \mathbb{E}[X_1 \mathbb{1}_{X_1 > c}] \text{ p.s}$$

En laissant $c = c_n$ tendre vers l'infini, on a

$$\liminf_n \frac{1}{n} \sum_{k=1}^{n} X_k \geq \mathbb{E}[X_1] \text{ p.s}$$

En appliquant à $(-X_i)$, on a

$$\limsup_n \frac{1}{n} \sum_{k=1}^{n} X_k \leq \mathbb{E}[X_1] \text{ p.s}$$

Étape 2 :
$\mathbb{E}[T] = \mathbb{E}[T']$ où $T := \inf\{n : S_n > 0\}$ et $T' := \#\{n : S_n = \min_{k \leq n} S_k\}$

Preuve (réversibilité des Marches Aléatoires)
Fixons $n \in \mathbb{N}$, définissons $S_0^{(n)} = 0$, $S_k^{(n)} = S_n - S_{n-k} = Y_n + ... + Y_{n-k+1}$. On a immédiatement que $(S_0, ..., S_n)$ a la même loi que $(S_0^{(n)}, S_1^{(n)}, ..., S_n^{(n)})$.
Maintenant,

$$\mathbb{P}[T(S^{(n)}) \geq n] = \mathbb{P}[S_n = \min_{k \leq n} S_k]$$

En sommant sur $n \in \mathbb{N}$,

$$\mathbb{E}[T] = \sum_{n \geq 0} \mathbb{P}[T \geq n] = \sum_{n \in \mathbb{N}} \mathbb{P}[S_n = \min_{k \leq n} S_k] = \sum_{n \in \mathbb{N}} \mathbb{P}[n \in \{m : S_m = \inf_{k \leq m} S_k\}] = \mathbb{E}[\#\{m : S_m = \min_{k \leq m} S_k\}] = \mathbb{E}[T']$$

Étape 3 : $\inf S_n > -\infty$ *p.s.*
Remarquons qu'il suffit de montrer $T' < \infty$ p.s. ou même $\mathbb{E}[T'] < \infty$, ou par l'étape 2, $\mathbb{E}[T] < \infty$.

Fixons $n \in \mathbb{N}$,

$$C \geq \mathbb{E}[S_{T \wedge n}] = \mathbb{E}[\sum_{k=1}^{T \wedge n} Y_k] = \mathbb{E}[Y_1]\, \mathbb{E}[T \wedge n]$$

Puisque $\mathbb{E}[Y_1] < \infty$, laisser n tendre vers l'infini donne

$$\mathbb{E}[T] \leq \frac{C}{\mathbb{E}[Y_1]} < \infty$$

□

3.2 Convergence en loi

Rappelons que $\mathcal{C}_b(\mathbb{R}^d)$ désigne l'ensemble des fonctions continues bornées de $\mathbb{R}^d$ dans $\mathbb{R}$. On le munit de la norme

$$\|\varphi\|_\infty := \sup_{x \in \mathbb{R}} |\varphi(x)|$$

Définition 3.7. Une suite (X_n) de variables aléatoires à valeurs dans $\mathbb{R}^d$ **converge en loi** vers X ($X_n \xrightarrow{\mathcal{L}} X$)

si

$$\forall \varphi \in \mathcal{C}_b(\mathbb{R}^d), \mathbb{E}[\varphi(X_n)] \overset{n\to\infty}{\longrightarrow} \mathbb{E}[\varphi(X)]$$

Remarque 3.8.

1. La convergence en loi concerne en fait *les lois* des variables aléatoires. Seuls les $\mathbb{P}_{X_n}$ et $\mathbb{P}_X$ importent !
2. L'espace des mesures de proba est inclus dans $(\mathcal{C}_b(\mathbb{R}^d))^*$. On peut donc voir la convergence en loi comme la convergence faible $*$ des mesures $\mathbb{P}_{X_n}$.

Exemple 3.9. 1. (X_n) à valeurs dans $\mathbb{Z}^d$ converge en loi vers X si $\forall k \in \mathbb{Z}$, $\mathbb{P}[X_n = k] \overset{n\to\infty}{\longrightarrow} \mathbb{P}[X = k]$

2. (X_n) à densité p_n et $p_n \overset{n\to\infty}{\longrightarrow} p$ p.p. et $\exists q \geq 0$ est intégrable telle que $|p_n| \leq q$ p.p., $\forall n$. Alors (X_n) converge en loi vers X.
3. X_n variable aléatoire uniforme sur $\{\frac{1}{n}, \frac{2}{n}, ..., 1\}$. Alors $(X_n) \overset{\mathcal{L}}{\to} \text{Unif}([0,1])$.

$$\mathbb{E}[\varphi(X_n)] = \frac{1}{n}(\varphi(\frac{1}{n}) + \varphi(\frac{2}{n}) + ... + \varphi(1)) \overset{n\to\infty}{\longrightarrow} \int_0^1 \varphi(x)dx, \quad \forall \varphi \in \mathcal{C}_b([0,1])$$

4. $X_n \sim \mathcal{N}(0, \sigma_n^2)$ et $\sigma_n^2 \overset{n\to\infty}{\longrightarrow} 0$, alors $X_n \to X$ où $X = 0$ p.s.
Soit $\varphi \in \mathcal{C}_b(\mathbb{R})$,

$$\begin{aligned}\mathbb{E}[\varphi(X_n)] &= \frac{1}{\sqrt{2\pi\sigma_n^2}} \int_{\mathbb{R}} \varphi(x) e^{-\frac{x^2}{2\pi\sigma_n^2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \varphi(\sigma_n y) e^{-\frac{y^2}{2}} dx\end{aligned}$$

$\varphi(\sigma_n y)e^{-\frac{y^2}{2}} \overset{n\to\infty}{\longrightarrow} \varphi 0)e^{-\frac{y^2}{2}}$ et est bornée par $\|\varphi\|_\infty e^{-\frac{y^2}{2}}$ Par le théorème de convergence dominée,

$$\mathbb{E}[\varphi(X_n)] \overset{n\to\infty}{\longrightarrow} \varphi(0)$$

Proposition 3.10.

1. $X_n \overset{(p)}{\to} X \implies X_n \overset{(\mathcal{L})}{\to} X$

2. Si $X_n \overset{(\mathcal{L})}{\to} X$ *et* X *est constante p.s., alors* $X_n \overset{(p)}{\to} X$

Cours 10

Démonstration. 1. Soit $(X_n) \overset{(p)}{\to} X$ et supposons que $X_n \overset{(\mathcal{L})}{\not\to} X$. Il existe donc $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$ telle que

$$\mathbb{E}[\varphi(X_n)] \not\to \mathbb{E}[\varphi(X)]$$

prenons (n_k) telle que :
— $\mathbb{E}[\varphi(X_{n_k})] \not\to \mathbb{E}[\varphi(X)]$
— $X_{n_k} \overset{p.s.}{\to} X$ (par *3.* de la Proposition 3.3)
Or, le théorème de convergence dominée et le fait que φ est *continue* et *bornée* impliquent que $\mathbb{E}[\varphi(X_{n_k})] \overset{k\to\infty}{\longrightarrow} \mathbb{E}[\varphi(X)]$ ce qui contredit la première propriété.

2. on suppose que $X_n \overset{(\mathrm{L})}{\to} a$. Soit $\epsilon > 0$. Considérons $\varphi_\epsilon \in \mathcal{C}_b(\mathbb{R}^d)$ telle que
— $\varphi_\epsilon(x) = 1$ si $\|x - a\| \leq \epsilon$
— $\varphi_\epsilon(x) = 0$ si $\|x - a\| \geq 2\epsilon$
— $0 \leq \varphi_\epsilon(x) \leq 1$

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}[|X_n - a| \leq 2\epsilon] &\geq \liminf \mathbb{E}[\varphi_\epsilon(X_n)] \\ &= \mathbb{E}[\varphi_\epsilon(a)] \text{ par convergence en loi} \\ &= 1\end{aligned}$$

D'où $\lim_{n\to\infty} \mathbb{P}[|X_n - a| > 2\epsilon] = 0, \forall \epsilon$

□

Remarque. Il n'est pas vrai que $X_n \overset{(\mathcal{L})}{\to} X$ implique $\lim \mathbb{P}[X_n \in \mathcal{B}] = \mathbb{P}[X \in B], \forall B$. Par exemple, $B = \{0\}$ et Ex 3.9 (d),

$$X_n \sim \mathcal{N}(0, \sigma_n^2) \overset{(\mathcal{L})}{\to} X = 0 \text{ p.s}$$

Proposition 3.11. *Soit* (X_n) *et* X *des variables aléatoires à valeurs dans* $\mathbb{R}^d$*. Les propriétés suivantes sont équivalentes :*

1. $(X_n) \overset{(\mathcal{L})}{\to} X$
2. $\forall$ *ouvert* U, $\liminf \mathbb{P}[X_n \in U] \geq \mathbb{P}[X \in U]$
3. $\forall$ *fermé* F, $\limsup \mathbb{P}[X_n \in F] \leq \mathbb{P}[X \in F]$
4. $\forall B$ *tel que* $\mathbb{P}[X \in \partial B] = 0$, $\lim \mathbb{P}[X_n \in B] = \mathbb{P}[X \in B]$

Démonstration. 1. $\Longrightarrow$ 2. Soit φ_ϵ tel que

- $0 \leq \varphi_\epsilon \leq \mathbb{1}_U$
- $(\varphi_\epsilon) \uparrow \mathbb{1}_U$

$$\begin{aligned}\liminf \mathbb{P}[X_n \in U] &\geq \sup_\epsilon \lim \mathbb{E}[\varphi_\epsilon(X_n)] \text{ (car } \varphi \leq \mathbb{1}_U) \\ &\geq \sup_\epsilon \mathbb{E}[\varphi_\epsilon(X)] \\ &= \mathbb{E}[\mathbb{1}_U(x)] = \mathbb{P}[X \in U]\end{aligned}$$

2. $\Longrightarrow$ 3. évident par passage au complémentaire

2. et 3. $\Longrightarrow$ 4.

$$\liminf \mathbb{P}[X_n \in B] \geq \liminf \mathbb{P}[X_n \in \mathring{B}] \geq \mathbb{P}[X \in \mathring{B}] \geq \mathbb{P}[X \in B] - \mathbb{P}[X \in \partial B] = \mathbb{P}[X \in B]$$

$$\limsup \mathbb{P}[X_n \in B] \leq \limsup \mathbb{P}[X_n \in \bar{B}] \leq \mathbb{P}[X \in \bar{B}] \leq \mathbb{P}[X \in B] + \mathbb{P}[X \in \partial B] = \mathbb{P}[X \in B]$$

4. $\Longrightarrow$ 1. Soit $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$,

$$\mathbb{E}[\varphi(X_n)] = \int_{\mathbb{R}^d} \varphi(x) d\,\mathbb{P}_{X_n}(x) \underset{\text{Fubini}}{=} \int_0^{\|\varphi\|_\infty} \mathbb{P}[\varphi(X_n) \leq t] dt$$

Il existe au plus un nombre dénombrable de t tel que $\mathbb{P}[\varphi(X) = t] > 0$. Si $\mathbb{P}[\varphi(X) = t] = 0$, alors par 4., $\mathbb{P}[\varphi(X_n) \geq t] \overset{n\to\infty}{\longrightarrow} \mathbb{P}[\varphi(X) \geq t]$.
Par convergence dominée, on obtient

$$\int_0^{\|\varphi\|_\infty} \mathbb{P}[\varphi(X_n) \geq t] dt \overset{n\to\infty}{\longrightarrow} \int_0^{\|\varphi\|_\infty} \mathbb{P}[\varphi(X) \geq t] dt = \mathbb{E}[\varphi(X)]$$

$\square$

Rappelons que $\mathcal{C}_c(\mathbb{R}^d)$ est l'espace des fonctions continues à support compact.

Proposition 3.12. *Soit H un sous-ensemble de $\mathcal{C}_b(\mathbb{R}^d)$ dont l'adhérence contient $\mathcal{C}_c(\mathbb{R}^d)$. Alors les propositions suivantes sont équivalentes :*

1. $X_n \overset{(\mathcal{L})}{\to} X$
2. $\mathbb{E}[\varphi(X_n)] \overset{n\to\infty}{\longrightarrow} \mathbb{E}[\varphi(X)], \forall \varphi \in H$

Démonstration.

1. $\Longrightarrow$ 2. est évident.

2. $\Longrightarrow$ 1. Soit $\varphi \in \mathcal{C}_b(\mathbb{R}^d)$, donnons nous $(f_k) \in \mathcal{C}_c(\mathbb{R}^d)$ tel que $0 \leq f_k \leq 1$ et $f_k \uparrow 1$ et $\varphi_k \in H$ telle que $\|\varphi_K - f_k\varphi\|_\infty \leq \frac{1}{k}$ (On utilise la densité de H et le fait que $f_k\varphi \in \mathcal{C}_c(\mathbb{R}^d)$)

$$\begin{aligned}|\mathbb{E}[\varphi(X_n)] - \mathbb{E}[\varphi(X)]| \leq\ & \overbrace{|\mathbb{E}[\varphi(X_n) - (f_k\varphi)(X_n)]|}^{\leq \|\varphi\|_\infty \mathbb{E}[(1-f_k)(X_n)]} + \overbrace{|\mathbb{E}[(f_k\varphi)(X_n) - \varphi_k(X_n)]|}^{\leq \|f_k\phi - \phi_k\|_\infty \leq \frac{1}{k}} \\ & + \underbrace{|\mathbb{E}[\varphi_k(X_n) - \varphi_k(X)]|}_{\to 0 \text{ par 2.}} + \underbrace{|\mathbb{E}[\varphi_k(X)| - (f_k\varphi)(X)]}_{\leq \|f_k\phi - \phi_k\|_\infty \leq \frac{1}{k}} + \underbrace{|\mathbb{E}[(f_k\varphi)(X) - \varphi(X)]|}_{\leq \|\varphi\|_\infty \mathbb{E}[(1-f_k)(X)]^{\text{conv. dominée}} \to 0}\end{aligned}$$

Appliqué a $\varphi \in \mathcal{C}_c(\mathbb{R}^d)$, on peut choisir k assez grand pour avoir $f_k\phi = \phi$, on obtient

$$\lim_{n\to\infty} \mathbb{E}[\varphi(X_n)] = \mathbb{E}[\varphi(X)]$$

Dans le cas général, le premier terme vérifie

$$\begin{aligned}\limsup |\mathbb{E}[\varphi(X_n) - (f_k\phi)(X_n)]| &\leq \|\phi\|_\infty \lim_{n\to\infty} \mathbb{E}[(1-f_k)(X_n)] \\ &= \|\varphi\|_\infty \mathbb{E}[(1-f_k)(X)]\end{aligned}$$

□

Proposition 3.13. *Soient (X_n) et X à valeurs réelles. On a équivalence entre les deux assertions suivantes :*

1. $X_n \overset{(\mathcal{L})}{\to} X$
2. $F_{X_n}(x) \to F_X(x)$ *pour tout point de continuité de* F_X.

Démonstration. 1. $\implies$ 2. Si F_X est continue en x, alors $\mathbb{P}[X \in \{x\}] = 0$, Donc 4. de la Prop 3.11 implique le résultat.

2. $\implies$ 1. Soit $a < b$,

$$\begin{aligned}\liminf \mathbb{P}[X_n \in (a,b)] &= \liminf(\mathbb{P}[X_n < b] - \mathbb{P}[X_n \leq a]) \\ &= \liminf F_{X_n}(b^-) - F_{X_n}(a)\end{aligned}$$

où $F_{X_n}(b^-) := \sup_{c<b} F_{X_n}(c)$. Il suffit donc de montrer que $\liminf F_{X_n}(b^-) \geq F_X(b^-)$ et $\limsup F_{X_n}(a) \leq F_X(a)$, ce qui impliquera $\geq F_X(b^-) - F_X(a) = \mathbb{P}[X \in (a,b)]$.
En prenant $(x_k) \uparrow b$ par valeurs inférieures $F_{X_n}(x_k) \overset{n\to\infty}{\longrightarrow} F_X(x_k)$, donc $\forall k, \limsup F_{X_n}(b^-) \geq \lim_{n\to\infty} F_{X_n}(x_k) = F_X(x_k)$. En faisant tendre k vers l'infini, on obtient

$$\limsup F_{X_n}(b^-) \geq \sup_k F_X(x_k) = F_X(b^-)$$

De même en a.
En résumé, $\liminf \mathbb{P}[X_n \in (a,b)] \leq \mathbb{P}[X \in (a,b)]$. Si U est un ouvert quelconque, $U = \bigsqcup_k (a_k, b_k)$.

$$\liminf \mathbb{P}\left[X_n \in \bigsqcup_k (a_k,b_k)\right] \geq \liminf \mathbb{P}\left[X_n \in \bigsqcup_{k\leq K}(a_k,b_k)\right] \geq \sum_{k\leq K} \mathbb{P}\left[X \in (a_k,b_k)\right]$$

On obtient le même résultat en faisant tendre vers l'infini.

□

Théorème 3.14 (Lévy)**.** *Soient (X_n) et X des variables aléatoires à valuers dans $\mathbb{R}^d$. Les deux propriétés suivantes sont équivalentes :*

1. $X_n \overset{(\mathcal{L})}{\to} X$
2. $\phi_{X_n}(\xi) \to \phi_X(\xi), \forall \xi \in \mathbb{R}^d$.

Démonstration.

1. $\implies$ 2. $\varphi(x) = e^{ix\xi} \in \mathcal{C}_b(\mathbb{R}^d)$, donc $\mathbb{E}[e^{iX_n\xi}] \to \mathbb{E}[e^{iX\xi}]$
2. $\implies$ 1. Posons $g_\sigma(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{x^2}{2\sigma^2}\right]$ et $H := \{g_\sigma(x) * f : f \in \mathcal{C}_c(\mathbb{R}^d), \sigma > 0\}$

Étape 1 : H est dense dans $\mathcal{C}_c(\mathbb{R}^d)$

Cours 11 On a déjà vu $g_\sigma * f \underset{\sigma>0}{\to} f$ p.p.. En fait, la convergence est uniforme.

Étape 2 : $\forall \varphi \in H, \mathbb{E}[\varphi(X_n)] \to \mathbb{E}[\varphi(X)]$

$$\mathbb{E}[\varphi(X_n)] = \int \varphi(x) d\,\mathbb{P}_{X_n}(x) \underset{\varphi\in H, \exists f, \varphi = g_\sigma * f}{=} \int g_\sigma * f(x) d\,\mathbb{P}_{X_n}(x)$$

En se souvenant que

$$\begin{aligned}g_\sigma * d\,\mathbb{P}_{X_n}(x) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}^d} e^{i\xi x} g_{1/\sigma}(\xi)\phi_{X_n}(-\xi)d\xi \\ &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{1}{\sqrt{2\pi\sigma^2}} f(x) e^{i\xi x} g_{1/\sigma}(\xi)\phi_{X_n}(-\xi)d\xi dx\end{aligned}$$

avec $\left|\frac{1}{\sqrt{2\pi\sigma^2}} f(x)e^{i\xi x} g_{1/\sigma}(\xi)\phi_{X_n}(-\xi)\right| \leq \frac{\|f\|_\infty}{\sqrt{2\pi\sigma^2}}\left|g_{1/\sigma}(\xi)\right|$ intégrable. Par convergence dominée,

$$\underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \frac{1}{\sqrt{2\pi\sigma^2}} f(x)e^{i\xi x} g_{1/\sigma}(\xi)\phi_X(-\xi)d\xi dx = \mathbb{E}[\varphi(X)]$$

□

3.3 Le théorème central limite

Théorème 3.15 (TCL)**.** *Soit (X_n) des variables aléatoires à valeurs dans $\mathbb{R}$, indépendantes, identiquement distribuées, et de carré intégrable. Alors*

$$\frac{1}{\sqrt{n}}(X_1 + ... + X_n - n\,\mathbb{E}[X_1]) \overset{(\mathcal{L})}{\to} \mathcal{N}(0, \sigma^2)$$

où $\sigma^2 = \mathrm{Var}(X_1)$

Remarque. $\mathrm{Var}(X_1 + ... + X_n - n\,\mathbb{E}[X_1]) = \mathrm{Var}(X_1 + ... + X_n) = n\,\mathrm{Var}(X_1) = n\sigma^2$

Démonstration. On peut supposer que $\mathbb{E}[X_1] = 0$. Soit $\xi \in \mathbb{R}$,

$$\begin{aligned}
\phi_{\frac{X_1+...+X_n}{\sqrt{n}}}(\xi) &= \mathbb{E}\left[\exp\left[i\xi\frac{X_1 + ... + X_n}{\sqrt{n}}\right]\right] \\
&\overset{\text{Prop 2.17}}{=} \mathbb{E}\left[e^{i\frac{\xi}{\sqrt{n}}X_1}\right]^n \\
&= \phi_{X_1}(\frac{\xi}{\sqrt{n}})^n \\
&\overset{\text{Prop 2.18}}{=} \left[1 + i\,\mathbb{E}[X_1]\frac{\xi}{\sqrt{n}} - \frac{1}{2}\,\mathbb{E}[X_1^2](\frac{\xi}{\sqrt{n}})^2 + o((\frac{\xi}{\sqrt{n}})^2)\right]^n \\
&= \left(1 - \frac{1}{2}\sigma^2\frac{\xi^2}{n} + o((\frac{\xi}{\sqrt{n}})^2)\right)^n \\
&= \exp\left[-(1 + o(1))\frac{1}{2}\sigma^2\xi^2\right] \\
&\overset{n\to\infty}{\longrightarrow} e^{-\frac{1}{2}\sigma^2\xi^2} = \phi_{\mathcal{N}(0,1)}(\xi)
\end{aligned}$$

Par le théorème de Lévy,

$$\frac{1}{\sqrt{n}}(X_1 + ... + X_n) \overset{(\mathcal{L})}{\to} \mathcal{N}(0, \sigma^2)$$

□

Définition 3.16. Soit C une matrice $d \times d$ symétrique positive. Une variable aléatoire X à valeurs dans $\mathbb{R}^d$ et un **vecteur gaussien de covariance** C, noté $\mathcal{N}(0, C)$, si

$$\forall \xi \in \mathbb{R}^d, \phi_X(\xi) = \exp\left[-\frac{1}{2}\xi^T C\xi\right]$$

Remarque 3.17.
1. $d = 1$, $C = [\sigma^2]$ et $\mathcal{N}(0, C) = \mathcal{N}(0, \sigma^2)$
2. Soit C matrice symétrique positive de dimension d, $\exists$ une vecteur gaussien X de covariance C

Théorème 3.18 (TCL multidimensionnel)**.** *Soit (X_n) de variables aléatoires à valeurs dans $\mathbb{R}^d$, indépendantes, identiquement distribuées, et de carré intégrable. Alors*

$$\frac{1}{\sqrt{n}}(X_1 + ... + X_n - n\,\mathbb{E}[X_1]) \overset{(\mathcal{L})}{\to} \mathcal{N}(0, K_{X_1})$$

où K_{X_1} est la matrice de covariance de X_1

Démonstration. Même preuve. □

Chapitre 4

Conditionnement

4.1 Conditionnement discret

Considérons $(\Omega, \mathcal{A}, \mathbb{P})$ un espace de probabilité.

Définition 4.1. Soit $B \in \mathcal{A}$ tel que $\mathbb{P}[B] > 0$. On définit **l'espérance conditionnelle**

1. de $A \in \mathcal{A}$ sachant B par la formule :

$$\mathbb{P}[A|B] := \frac{\mathbb{P}[A \cap B]}{\mathbb{P}\, B}$$

2. d'une variable aléatoire X (avec $X \geq 0$ ou $X \in L^1$) sachant B :

$$\mathbb{E}[X|B] := \frac{\mathbb{E}[X \mathbb{1}_B]}{\mathbb{P}[B]}$$

Définition 4.2. Soit $X \geq 0$ ou dans L^1 et un variable aléatoire Y à valeurs dans un espace E *dénombrable*. **L'espérence conditionnelle de** X **sachant** Y est la *variable aléatoire*

$$\mathbb{E}[X|Y] := \varphi(Y)$$

où $\varphi(y) := \begin{cases} \mathbb{E}[X|Y=y] & \text{si } \mathbb{P}[Y=y] > 0 \\ 0 & \text{sinon} \end{cases}$

Remarque 4.3.
1. La valeur de $\varphi(y)$ sur $E' = \{y \in E, \mathbb{P}[Y=y] = 0\}$ n'est pas importante
2. $\mathbb{E}[X|Y]$ est une variable aléatoire : $\mathbb{E}[X|Y](\omega) = \mathbb{E}[X|y]$ si $\omega \in \{Y=y\}$
3. $\mathbb{E}[X|Y]$ est $\sigma(Y)$-mesurable
4. $\mathbb{E}[X|f(Y)] = \mathbb{E}[X|Y]$ pour f bijective

Exemple 4.4.
1. Lancer d'un dé $\Omega = \{1, ..., 6\}$. $Y = \mathbb{1}_{\omega \text{impair}}$, $X = \omega$

$$\mathbb{E}[X|Y] = \begin{cases} 3 & \text{si } \omega \text{ impair} \\ 4 & \text{si } \omega \text{ pair} \end{cases}$$

2. $S_n = X_1 + ... + X_n$ où les X_i sont indépendantes, $\mathbb{P}[X_i = 1] = \mathbb{P}[X_i = -1] = \frac{1}{2}$ Soit $N \geq n$. $\mathbb{E}[S_N|S_n] = S_n$.
   En effet, pour $\omega \in \{S_n = k\}$.

$$\begin{aligned}
\mathbb{E}[S_N|S_n](\omega) = \mathbb{E}[S_N|S_n = k] &= \frac{\mathbb{E}[S_N \mathbb{1}_{S_n=k}]}{\mathbb{P}[S_n = k]} \\
&= \frac{S_n \mathbb{1}_{S_n=k}}{\mathbb{P}[S_n = k]} + \frac{\mathbb{E}[(X_{n+1} + ... + X_N)\mathbb{1}_{S_n=k}}{\mathbb{P}[S_n = k]} \\
&= k + \mathbb{E}[\underbrace{X_{n+1} + ... + X_N}_{=0}] \frac{\mathbb{P}[S_n = k]}{\mathbb{P}[S_n = k]} = k = S_n(\omega)
\end{aligned}$$

4.2 Définition de l'espérance conditionnelle

4.2.1 Cas de $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$

Théorème-Définition 4.5. *Soit $\mathcal{B}$ une sous-tribu de $\mathcal{A}$ et $X \in L^2(\Omega, \mathcal{A}, \mathbb{P})$. Il existe une unique variable aléatoire dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$, notée $\mathbb{E}[X|\mathcal{B}]$, telle que :*

$$(*) \forall Z \in L^2(\Omega, \mathcal{B}, \mathbb{P}), \mathbb{E}[Z\,\mathbb{E}[X|B]] = \mathbb{E}[ZX]$$

Démonstration. Existence : $L^2(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de Hilbert et $L^2(\Omega, \mathcal{B}, \mathbb{P})$ est un sous-espace vectoriel de $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Définissons $\hat{X}$ la **projection orthogonale** de X sur $L^2(\Omega, \mathcal{B}, \mathbb{P})$.

1. $\hat{X} \in L^2(\Omega, \mathcal{B}, \mathbb{P})$
2. pour $Z \in \mathrm{L}^2(\Omega, \mathcal{B}, \mathbb{P})$, $0 = \mathbb{E}[Z(X - \hat{X})] = \mathbb{E}[ZX] - \mathbb{E}[Z\hat{X}]$

Unicité : Soient $\hat{X}$ et $\check{X}$ deux variables aléatoires dans $L^2(\Omega, \mathcal{B}, \mathbb{P})$ vérifiant $(*)$.

$$\mathbb{E}[(\hat{X} - \check{X}) \underbrace{\mathbb{1}_{\hat{X} > \check{X}}}_{\mathcal{B}\text{-mesurable}}] \overset{(*)}{=} 0$$

donc $\mathbb{P}[\hat{X} > \check{X}] = 0$. Par symétrie on a également $\mathbb{P}[\hat{X} < \check{X}] = 0$, et donc $\mathbb{P}[\hat{X} \neq \check{X}] = 0$ $\square$

Remarque 4.6.
1. On peut remplacer $(*)$ par $\forall B \in \mathcal{B}, \mathbb{E}[X\mathbb{1}_B] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]\mathbb{1}_B]$
2. Si $\mathcal{B} = \sigma(Y)$, on note $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$
3. Si $A \in \mathcal{A}$, $\mathbb{P}[A|\mathcal{B}] := \mathbb{E}[\mathbb{1}_A|\mathcal{B}]$

Proposition 4.7. *Les propriétés suivantes sont vérifiées :*

1. *Si $X \geq 0$, alors $\mathbb{E}[X|\mathcal{B}] \geq 0$ p.s.*
2. *Si $X \geq X'$, alors $\mathbb{E}[X|\mathcal{B}] \geq \mathbb{E}[X'|B]$ p.s.*
3. *L'application $X \mapsto \mathbb{E}[X|\mathcal{B}]$ est linéaire*
4. *Si X est $\mathcal{B}$-mesurable, $\mathbb{E}[X|\mathcal{B}] = X$ p.s.*
5. *Si $\sigma(X)$ est indépendant de $\mathcal{B}$, $\mathbb{E}[X|\mathcal{B}] = \mathbb{E}[X]$ p.s.*
6. *$\mathbb{E}[\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[X]$*
7. *$\|\mathbb{E}[X|\mathcal{B}]\|_2 \leq \|X\|_2$*

Démonstration.
1. $0 \underset{(X \geq 0)}{\leq} \mathbb{E}[X\mathbb{1}_{\mathbb{E}[X|\mathcal{B}]<0}] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]\mathbb{1}_{\mathbb{E}[X|\mathcal{B}]<0}] \leq 0$, donc $\mathbb{P}[\mathbb{E}[X|\mathcal{B}] < 0] = 0$
2. 1. appliqué à $X - X' \geq 0$ nous donne $\mathbb{E}[X - X'|\mathcal{B}] \geq 0$. Par 3., on a $\mathbb{E}[X|\mathcal{B}] - \mathbb{E}[X'|\mathcal{B}] \geq 0$
3. $a\,\mathbb{E}[X|\mathcal{B}] + b\,\mathbb{E}[Y|B]$ vérifie $(*)$ pour $aX + bY$ et vaut donc $\mathbb{E}[aX + bY|\mathcal{B}]$.
4. X vérifie $(*)$
5. $\mathbb{E}[X]$ vérifie $(*)$
6. $(*)$ appliqué à $\mathbb{1}_\Omega$
7. $\mathbb{E}[X|\mathcal{B}]$ est une projection de X.

$\square$

Cours 12

4.2.2 Cas de $X \geq 0$

Théorème-Définition 4.8. *Soit $X \geq 0$, il existe une unique variable aléatoire, notée $\mathbb{E}[X|\mathcal{B}]$, positive, $\mathcal{B}$-mesurable, telle que*

$$(*)_{\geq 0}, \forall Z \geq 0 \mathcal{B}\text{-}mesurable, \mathbb{E}[Z\,\mathbb{E}[X|\mathcal{B}]] = \mathbb{E}[ZX]$$

Démonstration. Existence : Définissons $\mathbb{E}[X|B]$ comme la limite p.s. de la suite $\mathbb{E}[X \wedge n|\mathcal{B}]$ (cette espérence conditionnelle existe puisque $X \wedge n$ est bornée et donc dans L^2). La suite $(\mathbb{E}[X \wedge n|\mathcal{B}])_n$ converge bien car pour tout n, $X \wedge n \leq X \wedge (n+1)$ et donc par la Prop 4,7 (2.), $\mathbb{E}[X \wedge n|\mathcal{B}] \leq \mathbb{E}[X \wedge (n+1)|\mathcal{B}]$ p.s.

Vérifions $(*)_{\geq 0}$. Soit $Z \geq 0$,

$$\begin{aligned}
\mathbb{E}[Z\,\mathbb{E}[X|\mathcal{B}]] &= \mathbb{E}[Z \lim_{n\to\infty} \mathbb{E}[X\wedge n|\mathcal{B}]] \\
&= \lim_{n\to\infty} \mathbb{E}[Z\,\mathbb{E}[X\wedge n|\mathcal{B}] \\
&= \lim_{n\to\infty} \mathbb{E}[(Z\wedge n)\,\mathbb{E}[X\wedge n|\mathcal{B}]] \\
&= \lim_{n\to\infty} \mathbb{E}[(Z\wedge n)(X\wedge n)] \\
&= \mathbb{E}[ZX]
\end{aligned}$$

Unicité : Soient X' et X'' vérifiant les hypothèses du théorème. Soient $a < b$ rationnels. On applique $(*)_{\geq 0}$ à $Z = \mathbb{1}_{X' \leq a < b \leq X''}$

$$a\,\mathbb{P}[X' \leq a < b \leq X''] \geq \mathbb{E}[X'Z] = \mathbb{E}[X''Z] \geq b\,\mathbb{P}[X' \leq a < b \leq X'']$$

d'où $\mathbb{P}[X' \leq a < b \leq X''] = 0$

On en déduit $\mathbb{P}[X' < X''] = \mathbb{P}[\bigcup_{a<b\in\mathbb{Q}}\{X' \leq a < b \leq X''\}] = 0$. De même $\mathbb{P}[X' > X''] = 0$ □

Remarque 4.9. $\mathbb{E}[X|\mathcal{B}]$ peut valoir l'infini sur un évènement de proba positive *même* si $X < \infty$ p.s. En effet, posons $(\Omega, \mathcal{A}, \mathbb{P}) = ((0,1], \mathcal{B}((0,1]), dx)$ et
— $X(\omega) = \frac{1}{\omega}$
— $\mathcal{B} = \sigma((\frac{i-1}{n}, \frac{i}{n}], 1 \leq i \leq n)$
(Exercice) $\mathbb{E}[X|\mathcal{B}] = \infty \mathbb{1}_{(0,\frac{1}{n}]} + \sum_{i=2}^{n} n \log\left(\frac{i}{i-1}\right) \mathbb{1}_{(\frac{i-1}{n},\frac{i}{n}]}$

Proposition 4.10. *Les propriétés suivantes sont vérifiées.*

1. $X, X' \geq 0$, *et* $a, b \geq 0$, $\mathbb{E}[aX + bX'|\mathcal{B}] = a\,\mathbb{E}[X|B] + b\,\mathbb{E}[X'|\mathcal{B}]$
2. *Si* X *est* $\mathcal{B}$*-mesurable,* $\mathbb{E}[X|B] = X$ *p.s.*
3. *Si* X *est indépendant de* $\mathcal{B}$, $\mathbb{E}[X|B] = \mathbb{E}[X]$ *p.s.*
4. *Si* $X_n \uparrow X$ *p.s.,* $\mathbb{E}[X_n|B] \uparrow \mathbb{E}[X|B]$ *p.s.*
5. *Si* $X_n \geq 0$, $\mathbb{E}[\liminf X_n|\mathcal{B}] \leq \liminf \mathbb{E}[X_n|\mathcal{B}]$ *p.s.*

.

Démonstration. 1. à 3. : par unicité de l'espérance conditionnelle.
4. On a fait la preuve dans l'existence.
5. $Y_n = \inf_{k\geq n} X_k$ est croissante, 5. suit de 4. □

4.2.3 Cas de X intégrable

Théorème-Définition 4.11. *Pour* $X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$*, Il existe une unique variable aléatoire intégrable et* $\mathcal{B}$*-mesurable, notée* $\mathbb{E}[X|\mathcal{B}]$*, telle que :*

$$(*) \forall Z \text{ mesurable bornée}, \mathbb{E}[Z\,\mathbb{E}[X|B]] = \mathbb{E}[ZX]$$

Démonstration. Existence : Posons $X = X_+ - X_-$, avec $X_\pm \geq 0$ et

$$\mathbb{E}[X|\mathcal{B}] = \mathbb{E}[X_+|\mathcal{B}] - \mathbb{E}[X_-|B]$$

On a $\mathbb{E}[X|\mathcal{B}] \in L^1(\Omega, \mathcal{B}, \mathbb{P})$ par la Prop 4.10 (4.). On vérifie $(*)$ facilement.

Unicité : $Z = \mathbb{1}_{X'<X''}$, même preuve qu'avant. □

Remarque 4.12. On peut définir directement l'espérance conditionnelle dans le cas L^1 en utilisant le théorème de Radon-Nikodym.

Proposition 4.13. *Les propriétés suivantes sont vérifiées :*

1. *les mêmes qu'avant.*

2. Si $(X_n) \overset{p.s.}{\to} X$ et il existe Z intégrable tel que $\forall n, |X_n| < Z$ p.s., alors

$$\mathbb{E}[X_n|\mathcal{B}] \overset{p.s.}{\to} \mathbb{E}[X|\mathcal{B}]$$

3. Si φ est convexe, positive et X intégrable,

$$\mathbb{E}[\varphi(X)|\mathcal{B}] \geq \mathbb{E}[X|\mathcal{B}]$$

Démonstration. 2. Appliquer Prop 4.10 (5.) à $X_n + Z$ et $Z - X_n$
3. Soit $E_\varphi = \{(a,b) \in \mathbb{Q}^2 : \forall x \in \mathbb{R}, \varphi(x) \geq ax + b\}$. Observons que $\varphi(x) = \sup_{(a,b)\in E_\varphi} ax + b$. Donc

$$\begin{aligned}\mathbb{E}[\varphi(X)|\mathcal{B}] &= \mathbb{E}[\sup_{(a,b)\in E_\varphi} aX + b|\mathcal{B}] \\ &\geq \sup_{(a,b)\in E_\varphi} \mathbb{E}[aX + b|\mathcal{B}] \\ &= \sup_{(a,b)\in E_\varphi} (a\,\mathbb{E}[X|\mathcal{B}] + b) = \varphi(\mathbb{E}[X|\mathcal{B}])\end{aligned}$$

□

4.3 Autres propriétés de l'espérance conditionnelle

Proposition 4.14. *Soit X une variable aléatoire réelle et Y $\mathcal{B}$-mesurable, réelle. On a $\mathbb{E}[XY|\mathcal{B}] = Y\,\mathbb{E}[X|\mathcal{B}]$ Lorsque ces quantités ont du sens.*

Démonstration. Si $X, Y \geq 0$. On observe que $Y\,\mathbb{E}[X|\mathcal{B}]$ est ≥ 0 et $\mathcal{B}$-mesurable pour $z \geq 0$,

$$\mathbb{E}[Z(Y\,\mathbb{E}[X|\mathcal{B}])] = \mathbb{E}[(ZY)\,\mathbb{E}[X|\mathcal{B}]] \overset{(*)\geq 0}{=} \mathbb{E}[ZYX]$$

□

Proposition 4.15. *Si $\mathcal{B}' \subset \mathcal{B}$, alors pour tout X $(\geq 0, L^1, L^2)$,*

$$\mathbb{E}[\mathbb{E}[X|\mathcal{B}]|\mathcal{B}'] = \mathbb{E}[X|\mathcal{B}']$$

Remarque. Si $\mathcal{B}' \subset \mathcal{B}$, $\mathbb{E}[X|\mathcal{B}]$ est $\mathcal{B}'$-mesurable, donc égal à son espérance conditionnelle :

$$\mathbb{E}[\mathbb{E}[X|\mathcal{B}]|\mathcal{B}'] = \mathbb{E}[X|\mathcal{B}]$$

Démonstration. Si $\mathcal{B} \subset \mathcal{B}'$, et supposons $X \geq 0$, soit Z $\mathcal{B}'$-mesurable,

$$\mathbb{E}[Z\,\mathbb{E}[\mathbb{E}[X|\mathcal{B}]|\mathcal{B}']] \overset{(*)\geq 0 \text{ pour } \mathbb{E}[X|\mathcal{B}]}{=} \mathbb{E}[Z\,\mathbb{E}[X|\mathcal{B}] \overset{(*)\geq 0 \text{ pour } X}{=} \mathbb{E}[ZX]$$

□

Proposition 4.16. *Deux tribus $\mathcal{B}$ et $\mathcal{B}'$ sont indépendantes si et seulement si $\forall X$ $\mathcal{B}$-mesurable, $\mathbb{E}[X|\mathcal{B}'] = \mathbb{E}[X]$ p.s.*

Démonstration. $\Rightarrow$ est trivial car X est indépendant de $\mathcal{B}'$.
$\Leftarrow$ Pour $A \in \mathcal{B}$ et $A' \in \mathcal{B}'$,

$$\mathbb{P}[A \cap A'] = \mathbb{E}[\mathbb{1}_A \mathbb{1}_{A'}] \overset{(*)}{=} \mathbb{E}[\mathbb{E}[\mathbb{1}_A|\mathcal{B}']\mathbb{1}_{A'}] = \mathbb{E}[\mathbb{P}[A]\mathbb{1}_{A'}] = \mathbb{P}[A]\,\mathbb{P}[A']$$

□

Proposition 4.17. *Soient X et Y deux variables aléatoires à valeurs dans E et F respectivement, et $\mathcal{B}$ une tribu. Si*

1. X est indépendant de $\mathcal{B}$

2. Y est $\mathcal{B}$-mesurable,

alors pour tout $g : E \times F \to \mathbb{R}_+$ mesurable,

$$\mathbb{E}[g(X,Y)|\mathcal{B}] = \int_E g(x,Y)d\,\mathbb{P}_X(x)$$

Démonstration. $\int_E g(x,Y)d\,\mathbb{P}_X(x) = \varphi(Y)$ pour $\varphi : y \to \int_E g(x,y)d\,\mathbb{P}_X(x)$ est bien $\mathcal{B}$ mesurable positive.

Soit $Z \geq 0$. $\mathcal{B}$-mesurable,

$$\begin{aligned}
\mathbb{E}[Zg(X,Y)] &= \int zg(x,y)d\,\mathbb{P}_{(X,Y,Z)}(x,y,z) \\
&= \int zg(x,y)d\,\mathbb{P}_X(x)d\,\mathbb{P}_{(Y,Z)}(y,z) \text{car } (Y,Z) \text{ } \mathcal{B}\text{-mesurable et indépendant de } X \\
&\overset{\text{Fubini}}{=} \int z(\underbrace{\int g(x,y)d\,\mathbb{P}_X(x)}_{=\varphi(y)})d\,\mathbb{P}_{(Y,Z)}(y,z) = \mathbb{E}[Z\varphi(Y)]
\end{aligned}$$

□