PROBABILITÉS ET APPLICATIONS

Sandro VAIENTI

2023-2024

TABLE DES MATIÈRES

CHAPITRE 1

GÉNÉRALITÉS SUR LES PROBABILITÉS

1.1 Tribus

Définition 1.1.1 (σ-algèbre $\mathscr{A}$)**.** Ω ensemble. Les éléments de Ω constinuent une σ-algèbre $\mathscr{A}$:

1. $\Omega \in \mathscr{A}$;
2. Si $A \in \mathscr{A}$, alors $A^C \in \mathscr{A}$;
3. Si $\{A_n\}_{n=1}^{\infty}$ est une suite dénombrable dans $\mathscr{A}$, alors $\bigcup_n A_n \in \mathscr{A}$.

Exemple.

1. 2^{Ω}, ensemble de tous les ensembles de Ω, triviale ;
2. Grossière $\{\emptyset, \Omega\}$.

Définition 1.1.2 (σ-algèbre engendrée)**.** Dans Ω, on a une famille d'ensembles $\mathscr{S}$. On appelle $\sigma(\mathscr{S})$ la σ-algèbre engendrée par $\mathscr{S}$ qui est la plus petite σ-algèbre qui contient $\mathscr{S}$.

$$\sigma(\mathscr{S}) = \bigcap_{\substack{\mathscr{A}\,\sigma\text{ -algèbre} \\ \mathscr{A}_\alpha \in \mathscr{S}}} \mathscr{A}_\alpha$$

Exemple. $\Omega = \{a, b, c\}, \mathscr{S} = \{a, b\}$
Construire $\sigma(\mathscr{S})$.

Exemple. $\Omega = \{1, 2, 3, 4\}$ et $\mathscr{S} = \{\{1\}, \{2, 3\}\}$. Construire $\sigma(\mathscr{S})$.

Exemple. On a dans Ω deux ensembles A, B. Construire $\sigma(\{A, B\})$.
$\sigma(\{A, B\}) = \{\Omega, \emptyset, A, B, A^C, B^C, A \cup B, A \cup B^C, \dots\}$ (15 éléments).

Imaginons que $\Omega = \mathbb{R}$.
σ-algèbre de Borel (β). Il s'agit de la σ-algèbre engendrée par les intervalles ouverts.
$\mathscr{S} = \{(a, b), (-\infty, a), (b, +\infty)\}$.
On a $[a, b) \in \beta$, car $\bigcap_{k=1}^{\infty}(a - \frac{1}{k}, b) = [a, b)$.

Remarque (intersections dans une σ-algèbre). Si $A, B \in \mathscr{A}$, est-ce que $A \cap B \in \mathscr{A}$?
$A \cap B = (A^C \cup B^C)^C$.
$A - B = A \cap B^C$. $A - B \in \mathscr{A}$.
Les intersections dénombrables sont aussi dans $\mathscr{A}$.

Proposition 1.1.1. β est aussi engendrée par :

1. $\mathscr{S}_1 = \{[a, b)\}$;
2. $\mathscr{S}_2 = \{(a, b]\}$;
3. $\mathscr{S}_3 = \{[a, +\infty)\}$;
4. $\mathscr{S}_4 = \{(-\infty, a)\}$;
5. $\mathscr{S}_5 = \{(a, +\infty)\}$.

Démonstration. On montre $\mathscr{S}_1$.

$$[a, b) = \bigcap_{k=1}^{\infty} \left(a - \frac{1}{k}, b\right) \implies \sigma([a, b)) \subset \beta. \tag{1.1}$$

Montrons maintenant que $\beta \in \sigma([a, b))$.

$$\bigcup_{k=1}^{\infty} \left[a + \frac{1}{k}, b\right) = (a, b).$$

Donc $\beta \in \sigma([a, b))$.
Montrons $\mathscr{S}_3$. □

1.2 Probabilité

Définition 1.2.1 (Probabilité)**.** Soit $(\Omega, \mathscr{A})$ un espace mesurable.
On introduit une fonction d'ensemble $\mathbb{P} : \mathscr{A} \longrightarrow [0, 1]$ qu'on appelle **probabilité** et qui vérifie :

1. $\mathbb{P}(\Omega) = 1$;
2. Si (A_n) est une suite dénombrable dans $\mathscr{A}$ d'éléments deux à deux disjoints, alors

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

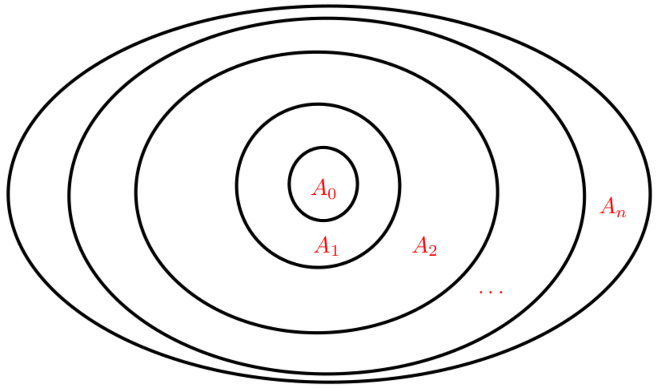


FIGURE 1.1 – On construit ainsi les couronnes

Si $\{A_n\}_{n=1}^{\infty}$ est telle que A_n ne sont pas deux à deux disjoints, alors

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Cette propriété s'appelle σ-sous additivité.

1.2.1 Continuité

Soit $\{A_n\}$ une suite croissante, i. e. $A_n \subset A_{n+1}$.
Est-ce que

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n)?$$

Soit $\{A_n\}_{n=1}^{\infty}$ une suite décroissante, ie $A_{n+1} \subset A_n$, alors

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Démonstration.

$$\bigcup_{n=1}^{\infty} A_n = A_1 \cup \left(\bigcup_{n=1}^{\infty} (A_{n+1} \setminus A_n)\right)$$

Donc par le deuxième axiome, on a

$$\begin{aligned}
\mathbb{P}(\bigcup_{n=1}^{\infty} A_n) &= \mathbb{P}(A_1) + \sum_{n=1}^{\infty} \mathbb{P}(A_{n+1} \setminus A_n) \\
&= \mathbb{P}(A_1) + \lim_{k\to+\infty} \sum_{n=1}^{k} \mathbb{P}(A_{n+1} \setminus A_n) \\
&= \mathbb{P}(A_1) + \lim_{k\to+\infty} \sum_{n=1}^{k} (\mathbb{P}(A_{n+1}) - \mathbb{P}(A_n)) \text{ (somme téléscopique)}
\end{aligned}$$

Or on a

$$\sum_{n=1}^{k} (\mathbb{P}(A_{n+1}) - \mathbb{P}(A_n)) = \mathbb{P}(A_{k+1}) - \mathbb{P}(A_1).$$

Donc

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \mathbb{P}(A_1) + \lim_{k\to\infty} [\mathbb{P}(A_{k+1}) - \mathbb{P}(A_1)]$$

et on obtient le résultat désiré. $\square$

$(\Omega, \mathscr{A}, \mathbb{P})$ espace mesuré ou de probabilité.
Dans le langage des probabilités, on appelle Ω l'univers et les éléments de $\mathscr{A}$ sont des événements.

1.3 Mesure de Dirac

Soit $\omega_0 \in \Omega$ quelconque.
$\delta_{w_0}(A), A \in \mathscr{A}$.
$\delta_{w_0}(A) = \begin{cases} 1 \text{ si } w_0 \in A \\ 0 \text{ sinon.} \end{cases}$
Montrer qu'il s'agit d'une probabilité.

Soit $\Omega = \{\omega_1, \omega_2, \dots, w_k, \dots\}$.
On associe à w_i un poids p_i tel que

$$\sum_{i=1}^{\infty} p_i = 1. \tag{1.2}$$

Si $A \subset \Omega$, on définit

$$\mathbb{P}(A) = \sum_{i,\omega_i \in A} p_i.$$

Si $card(\Omega) < \infty$, $\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$.
On associe à $\omega_j = \frac{1}{N}$.
Donc

$$\mathbb{P}(A) = \sum_{i,\omega_i \in A} \frac{1}{N} = \frac{card(\omega_i \text{ dans } A)}{card(\text{total de } \omega_i)} = \frac{\text{cas favorables}}{\text{cas possibles}}.$$

1.4 Mesure de Lebesgue

Elle est définie sur $\beta(\mathbb{R})$ et elle est la seule mesure qui se comporte comme ceci :
$\lambda((a, b]) = b - a$.
On a aussi

$$\lambda((a, b)) = \lambda([a, b]) = b - a. \tag{1.3}$$

$(a, b] = (a, b) \cup \{(b)\}$.
Il faut montrer que $\lambda(\{b\}) = 0$.
On a

$$\{b\} = \bigcap_{n=1}^{\infty} (b - \frac{1}{k}, b].$$

Donc

$$\begin{aligned}\lambda(\{b\}) &= \lambda(\bigcap_{k=1}^{\infty}(b-\frac{1}{k},b]) \text{ intersection décroissante}\\ &= \lim_{k\to\infty}\lambda((b-\frac{1}{k},b])\\ &= \lim_{k\to\infty}(b-b+\frac{1}{k}) = \lim_{k\to\infty}\frac{1}{k} = 0.\end{aligned}$$

1.4.1 Mesure de Lebesgue-Stieltjes

Soit F une fonction croissante bornée et continue à droite (i. e. $F(x_0) = \lim_{x\to x_0} F(x)$) et supposons que

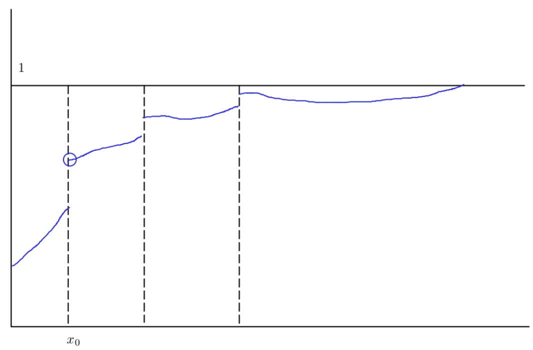


Figure 1.2 – Mesure de Lebesgue-Stieltjes

On définit une fonction d'ensemble $\nu((a,b]) = F(b) - F(a)$. ν devient une mesure de probabilité sur $\sigma((a,b]) = \beta$.

On a $\nu(\{x_d\}) \neq 0$.

Démonstration. $\{x_d\} = \bigcap_{k=1}^{\infty}(x_d - \frac{1}{k}, x_d]$.

Or

$$\begin{aligned}\nu(\{x_d\}) &= \lim_{k\to\infty}\nu((x_d-\frac{1}{k},x_d]) = \lim_{k\to\infty}F(x_d) - F(x_d-\frac{1}{k})\\ &= F(x_d) - \lim_{k\to\infty}F(x_d-\frac{1}{k}) = F(x_d)_+ - F(x_d)_-\\ &= \text{ différence entre la limite gauche et la limite droite } > 0.\end{aligned}$$

□

1.5 Fonctions mesurables

$(\Omega, \mathscr{A}) \to (\mathbb{R}, \beta),\ f : \Omega \to \mathbb{R}$.

Définition 1.5.1 (Mesurable)**.** On dit que f est mesurable si pour tout borélien $B \in \beta$, $f^{-1}(B) \in \mathscr{A}$.

Définition 1.5.2 (Equivalente)**.** $f^{-1}(\beta)$ est une sous σ-algèbre de $\mathscr{A}$.

Exercice 1. Montrer que $f^{-1}(\beta)$ est une σ-algèbre.

Démonstration.

1. $f^{-1}(\mathbb{R}) = \Omega$.
2. Si $A \in f^{-1}(\beta)$, est-ce que A^C est aussi dans $f^{-1}(\beta)$?
   Si $A \in f^{-1}(\beta)$, alors $\exists B \in \beta$ tel que $A = f^{-1}(B)$.
   $A^C = (f^{-1}(B))^C = f^{-1}(B^C)$.
3. Si $\{A_n\} \in f^{-1}(\beta)$, est-ce que $\bigcup_{n=1}^{\infty} A_n \in f^{-1}(\beta)$?

□

Proposition 1.5.1. $f : \Omega \to \mathbb{R}$.
Soit $\mathscr{C}$ une famille dans $\mathbb{R}$ telle que $\sigma(\mathscr{C}) = \beta$.
Si $f^{-1}(C), C \in \mathscr{C}$ est dans $\mathscr{A}$, alors f est mesurable.

Exemple. $f : \Omega$ (topologie) $\longrightarrow \mathbb{R}$ (topologie ouverts) continue.
f est mesurable.
Si $\mathcal{O}$ ouvert dans $\mathbb{R}$, $f^{-1}(\mathcal{O})$ est ouvert dans Ω.

$f : (\Omega, \mathscr{A}) \longrightarrow (\mathbb{R}, \beta)$. Pour tout $\omega \in \Omega$, $f(\omega) =$ constante.
On a deux cas :

1. $f^{-1}((a, b)) = \emptyset$ si la constante n'est pas dans (a, b) ;
2. $f^{-1}((a, b)) = \Omega$ si la constante est bien dans (a, b).

Exemple. Soient $f, g : (\Omega, \mathscr{A}) \to (\mathbb{R}, \beta)$ deux fonctions mesurables. Montrons que $f + g$ est mesurable.

Démonstration. On considère la famille $\{(a, \infty)\}$.
Si on veut montrer que $f + g$ est mesurable, il suffit de montrer que $(f + g)^{-1}((a, \infty)) \in \mathscr{A}$.
Donc il faut montrer que $\{\omega, f(\omega) + g(\omega) > a\} \in \mathscr{A}$, ie $\{\omega, f(\omega) > a - g(\omega)\}$.
Montrons d'abord que $a - g$ est mesurable.
$\omega \in (a - g)^{-1}((b, \infty))$.

$$a - g(\omega) > b \implies -g(\omega) > b - a$$
$$\text{Or } \{\omega, g(\omega) < a - b\} \in \mathscr{A},$$
$$\text{c'est à dire } g^{-1}((-\infty, a - b)) \in \mathscr{A}.$$

Notons $h = a - g$.
Si f, h sont mesurables, est-ce que $\{\omega, f(\omega) > h(w)\}$?
On a $\{\omega, f(\omega) > h(x)\} = \bigcup_{r \in \mathbb{Q}} \{f > r\} \cap \{h < r\}$.
Or $\{\omega, f(\omega) > r\} = f^{-1}(r, \infty) \in \mathscr{A}$ et $\{\omega, h(\omega) < r\} = h^{-1}((-\infty, r)) \in \mathscr{A}$.
On a donc montré que $f + g$ est mesurable. □

Proposition 1.5.2. On a aussi :

1. Si λ est un scalaire, λf est mesurable ;
2. $f \cdot g$ est mesurable ;
3. Si $g \neq 0$, $\frac{f}{g}$ est mesurable ;
4. Si on a une suite $\{f_n\}_{n \in \mathbb{N}}$ de fonctions mesurables, on a $\sup f_n, \inf f_n, \limsup f_n, \liminf f_n$ sont mesurables.

Soient $\Omega \in \mathbb{R}$ borné et $f : (\Omega, \beta) \longrightarrow (\mathbb{R}, \beta)$.

Soit $\mathcal{P}$ une partition de Ω, ie un ensemble $\mathcal{P} = \{P_i\}_{i=1}^{\infty}$ tel que

$$\bigcup_{i=1}^{\infty} P_i = \Omega \text{ et } P_i \cap P_j \neq 0 \text{ si } i \neq j.$$

Considérons par exemple la partition $\mathcal{P} = \{P_1, P_2, P_3, P_4\}$.

Comme $\mathcal{P}$ est une famille dans Ω, construisons $\sigma(\mathcal{P})$, ie la σ-algèbre engendrée par $\mathcal{P}$. Dans notre cas, $\sigma(P_1, P_2, P_3, P_4)$.

Cette σ-algèbre est composée par la réunion d'éléments de $\mathcal{P} \cup \{\emptyset\}$.

$$\sigma(\mathcal{P}) = \{\emptyset, \Omega, P_1, P_2, P_3, P_4, P_1^C = P_2 \cup P_3 \cup P_4, \ldots\}.$$

Si on a $B = \bigcup_{i=1}^{m} P_i$, alors $B^C = \bigcap_{i=1}^{m} P_i^C = \bigcap_{i=1}^{m} (\bigcup_{i=1}^{m} P_i)$.

Proposition 1.5.3. On sait exactement comment sont faites les fonctions mesurables dans ce cas. Il s'agit de fonctions constantes par morceaux.

Démonstration. Soit $\{p\} \in \mathbb{R}$, donc $f^{-1}(\{p\}) \in \sigma(\mathcal{P})$.

Supposons par exemple que $f^{-1}(\{p\}) = P_2 \cap P_3$. Si $x \in f^{-1}(\{p\})$, alors $f(x) = p$, mais $x \in P_2 \cup P_3$, donc sur $P_2 \cup P_3$ on aura en fait une valeur constante.

□

Définition 1.5.3 (Variable aléatoire)**.** Soit $X : (\Omega, \mathscr{A}, \mathbb{P}) \longrightarrow (\mathbb{R}, \beta)$.

Si X est une fonction mesurable, on dit que X est une **variable aléatoire**.

Remarque (Notations). On notera

$$\mathbb{P}(X \in B) = \mathbb{P}(\omega, X(\omega) \in B).$$

1.6 Loi de la variable aléatoire

Définition 1.6.1 (Loi de variable aléatoire)**.** Il s'agit d'une mesure de probabilité définie sur $\mathbb{R}$. On va la noter avec le symbole P_X.

Si $B \in \beta$,

$$P_X(B) = \mathbb{P}(X^{-1}(B)),$$

avec

$$X^{-1}(B) = \{\omega \mid X(\omega) \in B\}.$$

Définition 1.6.2 (Rappel : mesure (cf cours d'intégration de L3))**.**

Soit (E, τ) un ensemble mesurable. Alors une application $\mu : \tau \to \overline{\mathbb{R}^+}$ est une mesure sur E si :

1. $\mu(\emptyset) = 0$;

2. Pour toute suite (A_n) de τ disjointe, ie $A_n \cap A_m = \emptyset$ si $m \neq n$, alors

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n) \text{ (σ-additivité.)}$$

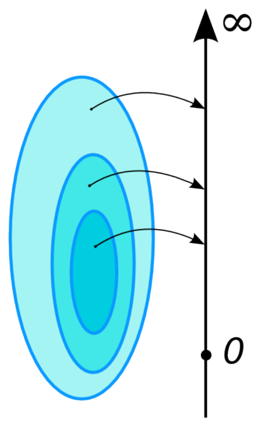


FIGURE 1.3 – De façon informelle, une mesure a la propriété d'être monotone : si l'ensemble E est un sous-ensemble de F, la mesure de E est inférieure ou égale à celle de F. De plus, on impose à la mesure de l'ensemble vide la valeur 0 (Wikipédia).

Exercice 2. Montrer que P_X est une mesure.

Démonstration. 1. $P_X(\mathbb{R}) = \mathbb{P}(X^{-1}(\mathbb{R})) = \mathbb{P}(\Omega) = 1$;

2. Si $\{B_n\}$ est une suite deux à deux disjointe,

$$P_X\left(\bigcup_{n=1}^{\infty} B_n\right) = \sum_{n=1}^{\infty} P_X(B_n),$$

car

$$\begin{aligned} P_X\left(\bigcup_{n=1}^{\infty} B_n\right) &= \mathbb{P}\left(X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right)\right) \\ &= \mathbb{P}\left(\bigcup_{n=1}^{\infty} X^{-1}(B_n)\right) \text{ (propriété de l'image réciproque)} \\ &= \sum_{n=1}^{\infty} \mathbb{P}(X^{-1}(B_n)) \text{ ($\mathbb{P}$ est une probabilité)} \\ &= \sum_{n=1}^{\infty} P_X(B_n). \end{aligned}$$

□

1.7 Intégrale

Définition 1.7.1 (Fonctions étagées/simples)**.** Soit X espace mesuré, $x \in X$. Une fonction h est appelée fonction étagée si h s'écrit de la manière suivante

$$h(x) = \sum_{k=1}^{M} c_k \mathbb{1}_{A_k}(x),$$

avec A_k ensemble mesurables (un élément de la σ-algèbre) tel que $A_k \cap A_j$ si $k \neq j$ et $c_k \geq 0$.

Théorème 1.7.1. *Soit $f : (X, \mathscr{A}, \mu)$ (espace mesuré) $\longrightarrow (\mathbb{R}^+, \beta)$ mesurable.*
Alors il existe une suite croissante $(\forall x, h_n(x) \leq h_{n+1}(x))$ de fonctions étagées $\{h_k\}_{n\in\mathbb{N}}$ telle que

$$f(x) = \lim_{n\to\infty} h_n(x).$$

Définition 1.7.2 (Intégrale de Lebesgue)**.**

1. *Première étape.* On considère une fonction simples

$$h(x) = \sum_{k=1}^{M} c_k \mathbb{1}_{A_k}(x),$$

alors

$$\int_X h d\mu = \int_X h(x) d\mu(x) = \sum_{k=1}^{M} c_k \mu(A_k).$$

2. *Deuxième étape.* Si f est mesurable non négative,

(a)

$$\int_X f d\mu = \sup\left(\int_X h d\mu, h \text{ simple}, h \leq f\right).$$

(b) Si $f = \lim_{k\to\infty} h_k, h_k$ simples non négatives,

$$\int_X f d\mu = \lim_{k\to\infty} \int h_k d\mu.$$

3. Si f mesurable de signe quelconque, on écrit $f(x) = f_+(x) - f_-(x)$, où $f_+ = \max(f, 0)$ et $f_- = \max(-f, 0)$.
(à suivre...)

Définition 1.7.3 (De classe $\mathscr{L}^1$)**.** On dira que f est de classe $\mathscr{L}^1$ si

$$\int_X |f| d\mu < +\infty.$$

Dans ce cas, $|f| = f_+ + f_-$.

Exemple. On considère la mesure de Dirac

$$\delta_a(A) = \begin{cases} 1 \text{ si } a \in A \\ 0 \text{ sinon.} \end{cases}$$

Alors

$$\int_X f(x) d\delta_a(x) = f(a). \tag{1.4}$$

Démonstration. Soit $f = \sum_{k=1}^{M} c_k \mathbb{1}_{A_k}$ une fonction simple.

$$\int \sum_{k=1}^{M} c_k \mathbb{1}_{A_k}(x) d\delta_a(x) = \sum_{k=1}^{M} c_k \delta_a(A_k) = c_{\overline{k}}$$

avec $\overline{k}$ le seul k tel que $a \in A_k$ car les A_k sont deux à deux disjoints.

Or $c_{\overline{k}} = f(a)$.

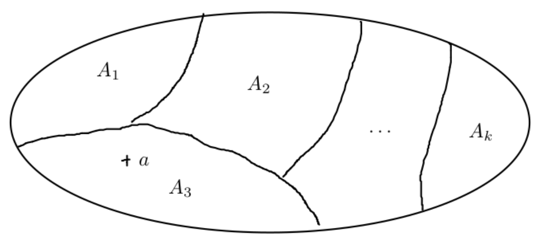


FIGURE 1.4 – a ne peut être que dans un seul des A_k.

On suppose maintenant que $f(x) = \lim_{k\to\infty} h_k(x)$.
Alors

$$\int_X f(x) d\delta_a(x) = \int_X \lim_{k\to\infty} h_k(x) d\delta_a(x) \overset{\text{Beppo-Levi}}{=} \lim_{k\to\infty} \int_X h_k d\delta_a(x) = \lim_{k\to\infty} h_k(a) = f(a).$$

□

Théorème 1.7.2 (Beppo-Levi)**.** *Si $\{h_k\}$ est une suite monotone non négative, alors*

$$\lim \int f_k = \int \lim f_k.$$

Remarque. Si $\mu = \sum_{k=1}^{M} p_k \delta_{a_k}$, avec $\sum_{k=1}^{M} p_k = 1$, alors

$$\int_X f d\left(\sum_{k=1}^{M} p_k \delta_{a_k}\right) = \sum_{k=1}^{M} p_k f(a_k).$$

Démonstration. Même démonstration que pour 1.4. □

1.7.1 Conséquences en probabilités

Soit $X : (\Omega, \mathscr{A}, \mathbb{P}) \longrightarrow (\mathbb{R}, \beta)$.

X est une variable aléatoire. Associée à X, il y a une mesure de Borel sur la droite qu'on appelle la loi notée P_X (cf 1.6.1).

Définition 1.7.4 (Espérance de X)**.** L'espérance de X se calcule comme suit :

$$\mathbb{E}[X] = \int_\Omega X(\omega) d\mathbb{P}(\omega). \tag{1.5}$$

Théorème 1.7.3 (De transfert)**.** *On a*

$$\int_\Omega f(X(\omega)) d\mathbb{P}(\omega) = \int_\mathbb{R} f(x) dP_X(x), \text{ avec } x \in \mathbb{R}. \tag{1.6}$$

Démonstration. Soit $f : \mathbb{R} \to \mathbb{R}$ fonction simple telle que $f(x) = \sum_{k=1}^{M} c_k \mathbb{1}_{A_k}(x)$.

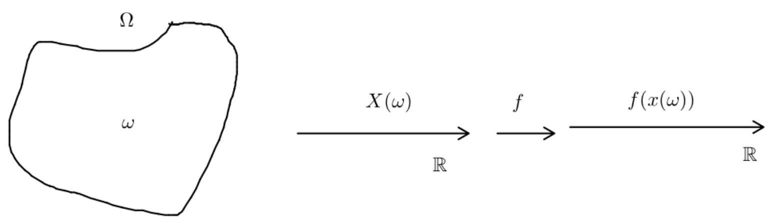


FIGURE 1.5 – Illustration théorème de transfert

$$\int_\Omega \sum_{k=1}^{M} c_k \mathbb{1}_{A_k}(X(\omega)) d\mathbb{P}(\omega) = \sum_{k=1}^{M} c_k \int_\Omega \mathbb{1}_{A_k}(X(\omega)) d\mathbb{P}(\omega). \tag{1.7}$$

Or

$$\int_\Omega \mathbb{1}_{A_k}(X(\omega)) d\mathbb{P}(\omega) = \int_\Omega \mathbb{1}_{X^{-1}(A_k)}(\omega) d\mathbb{P}(\omega) = \mathbb{P}(X^{-1}(A_k)) = P_X(A_k)$$

Donc 1.7 devient :

$$\sum_{k=1}^{M} c_k P_X(A_k) = \int_\mathbb{R} f(x) dP_X(x).$$

On considère que f est quelconque avec $f = \lim_{k\to\infty} h_k$, avec $\{h_k\}$ suite de fonctions simples.

$$\begin{aligned} \int f(X(\omega)) d\mathbb{P}(\omega) &= \int \lim_{k\to\infty} h_k(X(\omega)) d\mathbb{P}(\omega) = \lim_{k\to\infty} \int h_k(X(\omega)) d\mathbb{P}(\omega) \\ &= \lim_{k\to\infty} \int_\mathbb{R} h_k(x) dP_X(x) = \int_\mathbb{R} \lim_{k\to\infty} h_k(x) dP_X(x) = \int_\mathbb{R} f dP_X \end{aligned}$$

$\square$

1.8 Fonctions de répartition

Définition 1.8.1 (Fonction de répartition)**.** Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire. On définit

$$F(t) = \mathbb{P}(X \leq t), t \in \mathbb{R} \tag{1.8}$$
$$= \mathbb{P}(\omega, X(\omega) \leq t) = \mathbb{P}(X^{-1}((-\infty, t])) = P_X((-\infty, t]). \tag{1.9}$$

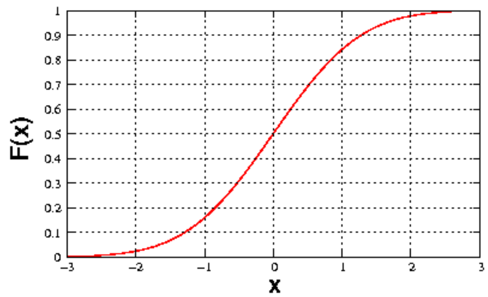


FIGURE 1.6 – Fonction de répartition

Proposition 1.8.1 (Propriétés de F)**.**

1. F est non négative et bornée entre 0 et 1.
2. F est croissante et continue à droite.
3. $$\begin{cases} \lim_{t\to+\infty} F(t) = 1 \\ \lim_{t\to-\infty} F(t) = 0. \end{cases}$$
4. F est discontinue dans au plus un nombre dénombrable de points.

Démonstration. 1. *Continuité à droite.* Si a est un point de discontinuité,

$$F(a) = \lim_{t\to a^+} F(t).$$

$$F(a) = P_X((-\infty, a]) = P_X\left(\bigcap_{k=1}^{\infty}\left(-\infty, \frac{1}{k}\right)\right) = \lim_{k\to\infty} P_X\left(\left(-\infty, a + \frac{1}{k}\right)\right) = \lim_{k\to\infty} F\left(a + \frac{1}{k}\right) = \lim_{t\to a^+} F(t),$$

car on écrit

$$(-\infty, a] = \bigcap_{k=1}^{\infty}\left(-\infty, a + \frac{1}{k}\right).$$

□

Soit $\pi(a)$ le saut dans un point de discontinuité de la fonction de répartition. On définit

$$A_n = \{t \in \mathbb{R}, \pi(t) \geq \frac{1}{n}\}.$$

Comme F est bornée entre 0 et 1, il peut y avoir au plus n éléments dans A_n.
Les points de discontinuité sont données par des points

$$t \in \bigcup_{n\geq 1} A_n \text{ (ensemble dénombrable).}$$

Comme F est continue et croissante à droite, elle définit une mesure de Lebesgue-Stieltjes

$$\nu((a,b]) = F(b) - F(a).$$

Démonstration. Montrons que $\nu = P_X$.

$$P_X((a,b]) = P_X((-\infty,b]) - P_X((-\infty,a]) = F(b) - F(a).$$

□

ν ne peut avoir que deux formes particulières. 18-09-2023

1. ν est une somme de masses de Dirac. Si A est un borélien de $\mathbb{R}$, alors

$$\nu(A) = \sum_{k=1}^{\infty} p_k \delta_{X_k}(A),$$

avec

$$\sum_{k=1}^{\infty} p_k = 1.$$

Alors

$$p_k = \mathbb{P}(X = \{x_k\}) = \mathbb{P}(\omega, X(\omega) = x_k).$$

Définition 1.8.2 (Variable aléatoire discrète)**.** On appelle **discrète** toute variable aléatoire X dont la loi a la forme :

$$P_X = \sum_{k=1}^{\infty} p_k \delta_{X_k}(A).$$

En particulier, nous écrirons toute variable aléatoire

$$X(\omega) = \sum_{k=1}^{\infty} x_k \mathbb{1}_{A_k}(\omega) \text{ et on aura } P_X(A) = \sum_{k=1}^{\infty} p_k \delta_{x_k}(A), \text{ avec } p_k = \mathbb{P}(A_k).$$

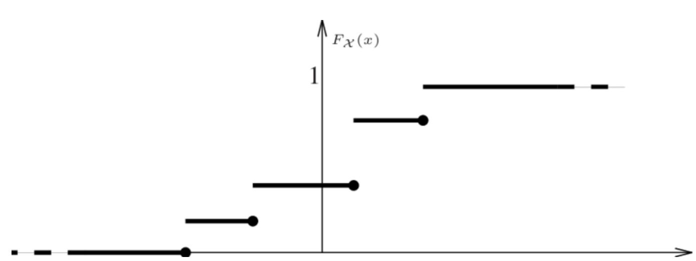


FIGURE 1.7 – Exemple de fonction de répartition d'une variable aléatoire discrète.

Remarque (Notations). Mesure de Lebesgue : $\begin{cases} \text{Leb} \\ dx \end{cases}$

2. Soit ν une mesure de probabilité de Borel sur $\mathbb{R}$. On dira que ν est **absolument continue** s'il existe une fonction non-négative $f \in L^1(\text{Leb})$ ($\int_{\mathbb{R}} f dx < \infty$) telle que, pour tout borelien B,

$$\nu(B) = \int_B f(x)dx.$$

On la dénote aussi $\nu \ll$ Leb.

En particulier si g est bornée ($L^\infty(\text{Leb})$), alors

$$\int_{\mathbb{R}} g d\nu = \int_{\mathbb{R}} g f dx.$$

f est la densité de f par rapport à Lebesgue (dérivée de Radon-Nykodym).

Dans notre cas, $\nu((a,b]) = F(b) - F(a)$.
Si F est la fonction de répartition de X, on a

$$P_X((a,b]) = F(b) - F(a) \overset{\text{Si la loi est AC}}{=} \int_{(a,b]} f(x)dx$$
$$= \int_{[a,b]} f(x)dx = \underbrace{\int_a^b f(x)dx}_{\text{Intégrale de Riemann}}$$

Définition 1.8.3 (Variable aléatoire à densité)**.** On dira qu'une variable aléatoire X est **à densité** si sa loi est absolument continue et on appelle f_X la densité de probabilité de X.

Si $\nu = c_1\delta_{X_1} + c_2\text{Leb}$, on écrit $\nu(A) = c_1\delta_{X_1}(A) + c_2\text{Leb}(A)$.
Il n'y a pas de saut dans la fonction de répartition d'une variable aléatoire absolument continue, car si $F(b) - F(a) = \int_a^b f(t)dt$, on a

$$\text{saut}(x_0) = F(x_0) - F(x_0^-) = 0.$$

C'est lié au fait que pour λ mesure de Lebesgue, on a $\lambda(\{b\}) = 0$ (cf 1.3).

Théorème 1.8.1. *Toute fonction de répartition F s'écrira (pour nous) de la forme*

$$F(t) = c_1F_d(t) + c_2F_{ac}(t),$$

où F_d est la fonction de répartition d'une variable aléatoire discrète et F_{ac} est la fonction de répartition d'une variable aléatoire absolument continue et $c_1 + c_2 = 1$.

Dans ce cas, $P_X = c_1(\text{masses de Dirac}) + c_2(\text{mesure absolument continue})$.

Soit X une variable à densité f_X. Soit $g : \mathbb{R} \to \mathbb{R}$. On considère la même variable aléatoire $g \circ X$. Quelle est sa loi ?

Exemple. Soit une variable aléatoire X avec la loi exponentielle. $X : \Omega \longrightarrow \mathbb{R}$,

$$f_X(x) = e^{-x}\mathbb{1}_{[0,\infty)}(x).$$

Trouver la loi de $X^2 = g \circ X$, avec $g : x \to x^2$.

Soit h une fonction bornée positive quelconque (fonction test).

$$\int_\Omega h(g(X))d\mathbb{P} = \int_\Omega (h\circ g)(X)d\mathbb{P} \overset{\text{transfert}}{=} \int_{\mathbb{R}} (h\circ g)(x)dP_X(x) \tag{1.10}$$
$$= \int_{\mathbb{R}} (h\circ g)(x)f_X(x)dx = \int_A h(g(x))f_X(x)dx. \tag{1.11}$$

On fait un changement de variable en posant $y = g(x), dy = g'(x)dx$.
Dans ce cas, 1.11 devient

$$\int h(y)f_X(g^{-1}(y))dy.$$

Or

$$\int h(y)c(y)dy = \int_{g(A)} h(y)f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}dy,$$

avec c la densité associée à $y = g(x)$ et $(g^{-1})' = \dfrac{1}{g'(g^{-1})}$.
Comme h est quelconque,

$$c(y) = f_X(g^{-1}(y))\frac{1}{g'(g^{-1}(y))}\mathbb{1}_{g(A)}(y).$$

Dans notre exemple, X a la densité $f_X(x) = e^{-x}\mathbb{1}_{[0,\infty)}(x)$ et $g(x) = x^2$, $A = [0,\infty)$.
On a

$$\begin{aligned} y &= x^2 \\ x &= \sqrt{y}. \end{aligned}$$

On obtient

$$\begin{aligned} \overbrace{c(y)}^{\text{densité} X^2} &= e^{-\sqrt{y}}\mathbb{1}_{[0,\infty)}(y)\frac{1}{2\sqrt{y}}\mathbb{1}_{g([0,\infty))}(y) \\ &= \frac{e^{-\sqrt{y}}}{2\sqrt{y}}\mathbb{1}_{[0,\infty)}(y), \end{aligned}$$

car $g([0,\infty)) = [0,\infty)$.

Exemple. X suit la loi uniforme entre $[-1,1]$.
$f_X = \frac{1}{2}\mathbb{1}_{[-1,1]}$, car $\dfrac{1}{2}\displaystyle\int_{-1}^{1} dt = 1$. Calculer X^2.

$$\begin{aligned} \int h(g(X))d\mathbb{P} &= \int (h\circ g)(X)d\mathbb{P} = \int_{-1}^{1} (h\circ g)(x)f_X(x)dx \\ &= \int_{-1}^{0} (h\circ g)(x)f_X(x)dx + \int_0^1 (h\circ g)(x)f_X(x)dx \\ &= \int_0^1 h(y)f_X(-\sqrt{y})\frac{1}{2\sqrt{y}}dy + \int_0^1 h(y)f_X(\sqrt{y})\frac{1}{2\sqrt{y}}dy \\ &= \int_0^1 h(y)\frac{1}{2\sqrt{y}}[f_X(-\sqrt{y}) + f_X(\sqrt{y})]dy. \end{aligned}$$

On a

$$\int h(g(X))d\mathbb{P} = \int h(y)c_Y(y)dy,$$

avec

$$c_Y(y) = \frac{1}{2\sqrt{y}}[f_X(-\sqrt{y}) + f_X(\sqrt{y})]\mathbb{1}_{[0,1]}(y).$$

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $B \in \mathscr{A}, \mathbb{P}(B) > 0$.
On peut définir $\mathbb{P}(A \mid B)$, la probabilité conditionnelle de A sachant B, définie comme suit :

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Exercice 3. Si l'on fixe B, montrer que $\mathbb{P}(\cdot \mid B) \to [0,1]$ est une probabilité.

On dira que deux événements A et B sont **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \iff \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, avec $\mathbb{P}(B) > 0$.

1.9 Système complet d'événements

Définition 1.9.1. Soit

$$\Omega = \bigcup_{i=1}^{\infty} A_i,$$

avec $A_i \cap A_j = \emptyset$ si $i \neq j$ et $\forall i \geq 1, \mathbb{P}(A_i) > 0$.
On appelle $\{A_i\}_{i=1}^{\infty}$ un **système complet d'événements** (SCE).

Théorème 1.9.1 (Formule des probabilités totales)**.** *Si $B \in \mathscr{A}$ quelconque, alors*

$$\mathbb{P}(B) = \sum_{i=1}^{\infty} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{\infty} \mathbb{P}(B \mid A_i)\mathbb{P}(A_i).$$

22-09-2023

CHAPITRE 2

VECTEURS ALÉATOIRES

On considère maintenant des variables aléatoires $X : (\Omega, \mathscr{A}, \mathbb{P}) \longrightarrow \mathbb{R}^n$ (Borel engendrée par les ouverts).

On étudiera surtout les cas où $n = 2$. On dénote un vecteur aléatoire $X \equiv (X_1, X_2)$ (parfois (X, Y)). Parfois on écrit, pour $i = 1, \ldots, n$,

$$X_i = \pi_i \circ X, \text{ où } \pi_i : \mathbb{R}^n \to \mathbb{R}, \pi_i(x_1, \ldots, x_n) = x_i.$$

π_i est une **projection**.

Définition 2.0.1 (Loi d'un vecteur aléatoire)**.** Soit B un borélien de $\mathbb{R}^n$. Alors la loi de X (loi conjointe) est définie ainsi :

$$P_X(B) = \mathbb{P}(\omega \in \Omega, X(w) = (X_1(\omega), \ldots, X_n(\omega)) \in B) = \mathbb{P}(X^{-1}(B)).$$

Est-il possible de calculer la loi de X_1 ?

Si on connaît la loi du couple de variables aléatoires, il est possible de calculer les lois des deux variables. Par contre, on ne peut pas avoir la loi du couple en connaissant la loi des deux variables aléatoires.

Définition 2.0.2 (Loi marginale)**.** Si D est un borélien de $\mathbb{R}$, la loi de X est

$$\underbrace{P_{X_1}(D)}_{\text{loi marginale}} = \mathbb{P}(X_1 \in D) = \mathbb{P}(\pi_1 \circ X \in D) = \mathbb{P}(X \in \pi_1^{-1}(D)) = P_X(\underbrace{\pi_1^{-1}(D)}_{\text{borélien}}).$$

Remarque. $\pi_i^{-1}(D)$ est borélien, car π_i est continue.

2.1 Fonction de répartition

Dans le cas où $n = 2$, $X = (X_1, X_2)$,

$$F_X(t_1, t_2) = \mathbb{P}(X_1 \leq t_1, X_2 \leq t_2) = P_X((-\infty, t_1] \times (-\infty, t_2]).$$

Définition 2.1.1. On dira que le vecteur aléatoire X est **absolument continu** s'il existe une fonction f mesurable et **non-négative** définie comme $f : \mathbb{R}^2 \to \mathbb{R}$ telle que, pour tout borélien B de $\mathbb{R}^2$, on a

$$P_X(B) = P_{(X_1,X_2)}(B) = \iint_B \underbrace{f(x_1,x_2)}_{\text{densité du couple}} dx_1 dx_2$$

et si $B = (-\infty, t_1] \times (-\infty, t_2]$, alors

$$F_X(t_1,t_2) = \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} f(x_1,x_2)dx_1dx_2 = \iint_{\mathbb{R}^2} f(x_1,x_2)\mathbb{1}_{(-\infty,t_1]}(x_1)\mathbb{1}_{(-\infty,t_2]}(x_2)dx_1dx_2.$$

Proposition 2.1.1.

$$F_{X_1}(t_1) = \lim_{t_2\to\infty} F_{(X_1,X_2)}(t_1,t_2) \text{ et } F_{X_2} = \lim_{t_1\to\infty} F_{(X_1,X_2)}(t_1,t_2).$$

Démonstration.

$$F_{X_1}(t_1) = \mathbb{P}(X_1 \le t_1) = \mathbb{P}(X_1 \le t_1, \underbrace{X_2 \in \mathbb{R}}_{\text{toujours vrai}}),$$

car

$$\{X_1 \le t_1\} = \{\omega, X_1(\omega) \le t_1\} = \{\omega, X_1(\omega) \le t_1\} \cap \Omega = \{\omega, X_1(\omega) \le t_1\} \cap \underbrace{X_2^{-1}(\mathbb{R})}_{X_2\in\mathbb{R}}).$$

Donc

$$F_{X_1}(t) = \mathbb{P}(X_1 \le t_1, X_2 \in \mathbb{R}) = P_{X_1,X_2}((-\infty,t_1] \times \mathbb{R}) = \lim_{k\to\infty} P_{X_1,X_2}((-\infty,t_1] \times (-\infty,k]),$$

car

$$\left((-\infty,t_1] \times \bigcup_{k=1}^{\infty}(-\infty,k)\right) = (-\infty,t_1] \times \mathbb{R}.$$

Ainsi

$$\lim_{k\to\infty} P_{X_1,X_2}((-\infty,t_1] \times (-\infty,k]) = \lim_{k\to\infty} F_{(X_1,X_2)}(t_1,k), k \in \mathbb{R}.$$

□

Théorème 2.1.1 (De transfert pour un vecteur aléatoire)**.** *Soit g une fonction mesurable. Alors on a*

$$\int_\Omega g(X)d\mathbb{P} = \int_\Omega g((X_1,X_2))d\mathbb{P} = \iint_{\mathbb{R}^2} g(x_1,x_2)dP_X(x_1,x_2) = \iint_{\mathbb{R}^2} g(x_1,x_2)f_X(x_1,x_2)dx_1dx_2.$$

Proposition 2.1.2. Soit $X = (X_1, X_2)$ absolument continu, donc il existe une densité $f_X(x_1, x_2)$. Alors on a

$$\begin{cases} f_{X_1}(x_1) = \int_{\mathbb{R}} f_{(X_1,X_2)}(x_1, x_2)dx_2, \\ f_{X_2}(x_2) = \int_{\mathbb{R}} f_{(X_1,X_2)}(x_1, x_2)dx_1. \end{cases}$$

Démonstration. Si X_1 a une densité, cela signifie

$$F_{X_1}(t_1) = \int_{-\infty}^{t_1} f_{X_1}(x_1)dx_1.$$

Ce résultat est-il vrai ?

Par le théorème précédent,

$$\begin{aligned} F_{X_1}(t_1) &= \lim_{t_2\to\infty} F_{(X_1,X_2)}(t_1, t_2) = \lim_{t_2\to\infty} \int_{-\infty}^{t_1} \int_{-\infty}^{t_2} f_{(X_1,X_2)}(u_1, u_2)du_1 u_2 \\ &= \lim_{t_2\to\infty} \iint_{\mathbb{R}^2} f_{(X_1,X_2)}(u_1, u_2)\mathbb{1}_{(-\infty,t_1]}(u_1)\mathbb{1}_{(-\infty,t_2]}(u_2)du_1 du_2 \\ &= \lim_{t_2\to\infty} \int_{\mathbb{R}} \mathbb{1}_{(-\infty,t_1]}(u_1) \left[\int_{\mathbb{R}} f_{(X_1,X_2)}(u_1, u_2)\mathbb{1}_{(-\infty,t_2]}(u_2)du_2\right] du_1 \\ &= \int_{\mathbb{R}} \mathbb{1}_{(-\infty,t_1]}(u_1) \int_{\mathbb{R}} f_{(X_1,X_2)}(u_1, u_2)du_2 du_1 = \int_{-\infty}^{t_1} \left(\int_{\mathbb{R}} f_{(X_1,X_2)}(u_1, u_2)du_2\right) du_1. \end{aligned}$$

□

CHAPITRE 3

INDÉPENDANCE

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.

3.1 Evénements indépendants

Définition 3.1.1. Deux événements $A, B \in \mathscr{A}$ sont indépendants si

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

ou

$$\mathbb{P}(A \mid B) = \mathbb{P}(A) \text{ si } \mathbb{P}(B) > 0.$$

3.2 Mutuellement indépendants

Définition 3.2.1. Soit $\{A_n\}$ une suite dénombrable d'événements. On dira que cette suite est **mutuellement indépendante** si pour toute sous-suite d'événements $A_{i_1}, \ldots, A_{i_k}$, on a

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \ldots \mathbb{P}(A_{i_k}).$$

3.3 Classes d'événements indépendantes

Définition 3.3.1. Deux classes d'événements $\mathscr{C}_1$ et $\mathscr{C}_2$ sont dites indépendantes si $\forall A_1 \in \mathscr{C}_1$ et $A_2 \in \mathscr{C}_2$, A_1 et A_2 sont indépendants.

Cela se généralise à n classes $\mathscr{C}_1, \ldots, \mathscr{C}_n$.

Définition 3.3.2. Soit $\mathscr{C}_1$ et $\mathscr{C}_2$ deux classes d'événements indépendants et en plus qui sont **stables par intersection finie** (π-système). Alors les σ-algèbres engendrées par $\mathscr{C}_1$ et $\mathscr{C}_2$ sont indépendantes.

Définition 3.3.3. Soit $\mathscr{C}$ une classe. Soient $C_1, C_2, \dots, C_n \in \mathscr{C}$. On dira que $\mathscr{C}$ est stable par intersection finie si

$$C_1 \cap C_2 \cap \cdots \cap C_n \in \mathscr{C}.$$

Exemple. 1. Dans $\mathbb{R}$, on considère la classe $\mathscr{C}_1 = \{(a,b], a,b \in \mathbb{R}\}$. C'est un π-système, car pour tout $(a,b], (c,d]$, on aura

$$(a,b] \cap (c,d] = (\max\{a,c\}, \min\{b,d\}].$$

Dans $\mathbb{R}^2$, $\mathscr{C} = \{(a_1,b_1] \times (a_2,b_2]\}$. Alors $\sigma(\mathscr{C}) = \sigma$-algèbre de Borel dans $\mathbb{R}^2$.

Définition 3.3.4. Soit $X : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire telle que $X^{-1}(\beta) \subset \Omega$. On appelle $\sigma(X)$ la plus petite σ-algèbre qui rend X mesurable.

Proposition 3.3.1.

$$\sigma(X) = X^{-1}(\beta),$$

où β est la σ-algèbre de Borel dans $\mathbb{R}$.

Remarque (Personnelle). Soient E_1, E_2 ensembles et $f : E_1 \to E_2$. On rappelle que si τ_2 est une tribu sur E_2, alors

$$f^{-1}(\tau_2) = \{f^{-1}(A), A \in \tau_2\}$$

est une tribu sur E_1, appelée **la tribu réciproque**.
Dans le cas de probabilités, on aura

$$X^{-1}(\beta) = \{X^{-1}(B), B \text{ borélien}\}.$$

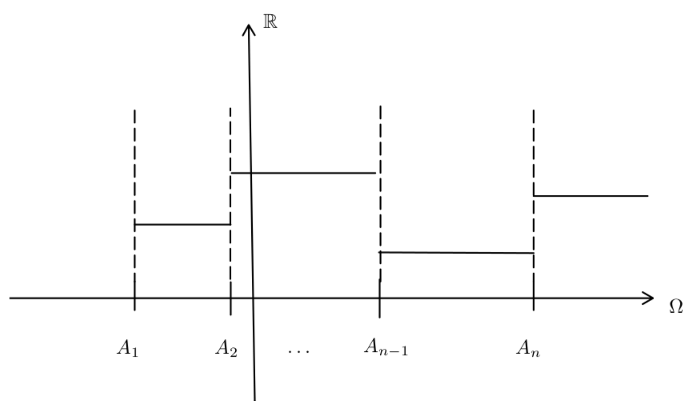


Figure 3.1 – Dans ce cas, $\sigma(X) = \sigma$-algèbre engendrée par la partition $\mathscr{A}$.

Remarque (Personnelle). On rappelle que pour la figure 3.1, si $A_1, \ldots, A_n$ est une partition de l'ensemble X, on définit

$$\tau = \left\{ \bigcup_{i \in J} A_i, J \subset \{1, \ldots, n\} \right\}.$$

On démontre que τ est une tribu.

3.4 Indépendance de variables aléatoires

Définition 3.4.1. Deux variables aléatoires X_1, X_2 sont indépendantes si $\sigma(X_1) = X_1^{-1}(\beta)$ et $\sigma(X_2) = X_2^{-1}(\beta)$ sont deux **familles indépendantes** (cf définition 3.3.2).

Un élément de $X^{-1}(\beta)$ s'écrit comme $X^{-1}(B_1), B_1 \in \beta$.
On écrira alors

$$\mathbb{P}(X_1^{-1}(B_1) \cap X_2^{-1}(B_2)) = \mathbb{P}(X_1^{-1}(B_1))\mathbb{P}(X_2^{-1}(B_2)).$$

On aura par conséquent, pour $(-\infty, b]$ qui engendrent β ($\{(-\infty, b]\}$ est un π-système),

$$\mathbb{P}(X_1^{-1}((-\infty, b_1]) \cap X_2^{-1}(-\infty, b_2]) = \mathbb{P}(X_1^{-1}(-\infty, b_1])\mathbb{P}(X_2^{-1}(-\infty, b_2]),$$

donc

$$\mathbb{P}(X_1 \in (-\infty, b_1], X_2 \in (-\infty, b_2]) = \mathbb{P}(X_1 \in (-\infty, b_1])\mathbb{P}(X_2 \in (-\infty, b_2]).$$

On obtient ainsi

$$\mathbb{P}(X_1 \le b_1, X_2 \le b_2) = \mathbb{P}(X_1 \le b_1)\mathbb{P}(X_2 \le b_2). \tag{3.1}$$

Le résultat 3.1 est exactement la définition de l'indépendance des deux variables aléatoires. 25-09-2023
On peut écrire 3.1 en terme de lois de probabilités :

$$\mathbb{P}(\overbrace{X_1 \le b_1, X_2 \le b_2}^{\subset \Omega}) = P_X((-\infty, b_1] \times (-\infty, b_2]) \overset{\text{si indép}}{=} P_{X_1}((-\infty, b_1])P_{X_2}((-\infty, b_2]). \tag{3.2}$$

Proposition 3.4.1. Deux variables aléatoires X_1, X_2 définies sur l'espace $(\Omega, \mathscr{A}, \mathbb{P})$ sont indépendantes si et seulement si

$$P_{(X_1, X_2)} = P_{X_1} P_{X_2}.$$

Démonstration.

1. *Partie nécessaire.* On l'a démontré dans 3.2.
2. *Partie suffisante.* On suppose que

$$\overbrace{P_{(X_1, X_2)}}^{\text{loi du couple}}(B_1 \times B_2) = P_{X_1}(B_1)P_{X_2}(X_2).$$

On prend $B_1 = (-\infty, b_1]$ et $B_2 = (-\infty, b_2]$.
Ainsi

$$P_{(X_1, X_2)}((-\infty, b_1] \times (-\infty, b_2]) = \mathbb{P}(X_1 \le b_1, X_2 \le b_2).$$

□

Exemple. Considérons un vecteur aléatoire (X_1, X_2) de loi

$$\begin{cases} P_{X_1} = \frac{1}{2}\delta_{a_1} + \frac{1}{2}\delta_{a_2} = f_1(y)dy \\ P_{X_2} = f_2(x)dx. \end{cases}$$

On veut trouver la loi de (X_1, X_2).

On calcule :

$$\int e^{X_1+X_2} d\mathbb{P} = \iint e^{x_1+x_2} dP_{(X_1,X_2)} P(x_1, x_2) = \iint e^{x_1+x_2} d\left(\frac{1}{2}\delta_{a_1} + \frac{1}{2}\delta_{a_2}\right) f_2 dx_2 = \dots$$

Corollaire. *Si X_1, X_2 sont indépendantes, la même chose est vraie pour $f(X_1)$ et $g(X_2)$ avec f et g réelles et mesurables.*

Proposition 3.4.2. Les variables aléatoires X_1 et X_2 sont indépendantes si et seulement si pour toutes f, g non-négatives, on a

$$\mathbb{E}[f(X_1)g(X_2)] = \mathbb{E}[f(X_1)]\mathbb{E}[g(X_2)]$$

et la même chose pour f et g réelles et bornées.

Démonstration.

1. *Partie nécessaire.*

$$\int_\Omega f(X_1)g(X_2)d\mathbb{P} = \iint_{\mathbb{R}^2} f(x_1)g(x_2)d_{X_1X_2}P(x_1, x_2). \tag{3.3}$$

Or si les variables sont indépendantes, la loi $P_{X_1X_2}$ se factorise dans les marginales.
Donc 3.3 devient :

$$\iint_{\mathbb{R}^2} f(x_1)g(x_2)dP_{X_1}(x_1)dP_{X_2}(x_2) = \int_{\mathbb{R}} f(x_1)dP_{X_1}(x_1) \int_{\mathbb{R}} g(x_2)dP_{X_2}(x_2) = \mathbb{E}[X_1]\mathbb{E}[X_2].$$

2. *Partie suffisante.* On considère $f = \mathbb{1}_{(-\infty,b_1]}$ et $g = \mathbb{1}_{(-\infty,b_2]}$.
On a

$$\int \mathbb{1}_{(-\infty,b_1]} \circ X_1 \mathbb{1}_{(-\infty,b_2]} \circ X_2 d\mathbb{P} = \mathbb{E}[f(X_1)g(X_2)] = \mathbb{E}[f(X_1)]\mathbb{E}[g(X_2)] = \mathbb{P}(X_1 \leq b_1)\mathbb{P}(X_2 \leq b_2).$$

□

Proposition 3.4.3 (Indépendance et fonctions de répartition)**.** Les variables aléatoires X_1 et X_2 sont indépendantes si et seulement si

$$F_{X_1,X_2}(t_1, t_2) = F_{X_1}(t_1)F_{X_2}(t_2).$$

Démonstration. 1. *Partie nécessaire.* Si les variables aléatoires X_1, X_2 sont indépendantes, alors

$$F_{X_1,X_2}(t_1,t_2) = \mathbb{P}(X_1 \leq t_1, X_2 \leq t_2) = \mathbb{P}(X_1 \leq t_1)\mathbb{P}(X_2 \leq t_1) = F_{X_1}(t_1)F_{X_2}(t_2).$$

2. *Partie suffisante.* Point de départ : on écrit la condition

$$\mathbb{P}(X_1 \leq t_1, X_2 \leq t_2) = F_{X_1X_2}(t_1,t_2) = F_{X_1}(t_1)F_{X_2}(t_2) = \mathbb{P}(X_1 \leq t_1)\mathbb{P}(X_2 \leq t_2),$$

avec t_1, t_2 quelconques et on obtient que X_1 et X_2 sont indépendantes comme c'est dit dans la définition 3.1.

□

Proposition 3.4.4 (Indépendance et densités)**.**

1. Soient X_1, X_2 deux variables aléatoires avec les densités f_{X_1} et f_{X_2} et supposons que X_1 et X_2 sont indépendantes. Alors le couple (X_1, X_2) et sa densité vérifient :

$$f_{X_1,X_2}(x_1,x_2) = f_{X_1}(x_1)f_{X_2}(x_2).$$

2. Supposons que (X_1, X_2) admet la densité $f_{X_1,X_2}(x_1,x_2)$ qui est le produit entre deux fonctions intégrables non-négatives $\tilde{f}_1(x_1), \tilde{f}_2(x_2)$. Alors $\tilde{f}_1(x_1)$ et $\tilde{f}_2(x_2)$ sont à des facteurs multiplicatifs près, les densités de X_1 et X_2 sont **indépendantes**.

Démonstration.

1. Loi de $X_1 \implies f_{X_1}(x_1)dx_1 : P_{X_1}$ et loi de $X_2 \implies f_{X_2}(x_2)dx_2 : P_{X_2}$.

$$f_{X_1,X_2}(x_1,x_2)dx_1dx_2 = P_{X_1X_2} = P_{X_1}P_{X_2} = f_{X_1}(x_1)f_{X_2}(x_2)dx_1dx_2.$$

2. On sait que

$$f_{X_1}(x_1) = \int f_{X_1X_2}(x_1,x_2)dx_2 = \tilde{f}_{X_1}(x_1)\int \tilde{f}_{X_2}(x_2)dx_2$$
$$f_{X_2}(x_2) = \int f_{X_1X_2}(x_1,x_2)dx_1 = \tilde{f}_{X_2}(x_2)\int \tilde{f}_{X_1}(x_1)dx_1.$$

Ensuite,

$$\iint_{\mathbb{R}^2} f_{X_1X_2}(x_1,x_2)dx_1dx_2 = 1 = \iint_{\mathbb{R}^2} \tilde{f}_{X_1}(x_1)\tilde{f}_{X_2}(x_2)dx_1dx_2 = \int_{\mathbb{R}} \tilde{f}(x_1)dx_1 \int_{\mathbb{R}} \tilde{f}_{X_2}(x_2)dx_2.$$

On a

$$f_{X_1}(x_1)f_{X_2}(x_2) = \tilde{f}_{X_1}(x_1)\tilde{f}_{X_2}(x_2)\underbrace{\int \tilde{f}_{X_2}(x_2)dx_2 \int \tilde{f}_{X_1}(x_1)dx_1}_{=1}.$$

En conclusion, on a

$$f_{X_1}(x_1)f_{X_2}(x_2) = \tilde{f}_{X_1}(x_1)\tilde{f}_{X_2}(x_2).$$

Pour que f_{X_1} et f_{X_2} deviennent des densités, on pose :

$$f_{X_1}(x_1) = \frac{\tilde{f}_{X_1}(x_1)}{\int f_{X_1}(x_1)dx_1} \text{ et } f_{X_2}(x_2) = \frac{\tilde{f}_{X_2}(x_2)}{\int f_{X_2}(x_2)dx_2}.$$

□

26-09-2023

3.5 Changement de variables

On considère (X_1, X_2) un vecteur aléatoire à densité $f_{X_1X_2}(x_1, x_2)$. On construit deux autres variables aléatoires U_1, U_2 telles que

$$\begin{cases} U_1 = g_1(X_1, X_2) \\ U_2 = g_2(X_1, X_2). \end{cases}$$

On veut trouver la loi de U_1, U_2, à savoir la densité du couple.
Soit $h : \mathbb{R}^2 \to \mathbb{R}$ une fonction mesurable positive quelconque que l'on appelle "fonction test".

$$\int_\Omega h(U_1, U_2) d\mathbb{P} = \int_\Omega h(g_1(X_1, X_2), g_2(X_1, X_2)) d\mathbb{P} \tag{3.4}$$

$$\overset{\text{transfert}}{=} \iint_{\mathbb{R}^2} h(g_1(x_1, x_2), g_2(x_1, x_2)) dP_{X_1X_2} \tag{3.5}$$

$$= \iint_{\mathbb{R}^2} h(g_1(x_1, x_2), g_2(x_1, x_2)) f_{X_1,X_2}(x_1, x_2) dx_1 dx_2. \tag{3.6}$$

On pose

$$\begin{cases} u_1 = g_1(x_1, x_2) \\ u_2 = g_2(x_1, x_2). \end{cases}$$

On sait que (X_1, X_2) ne sont pas forcément définies sur $\mathbb{R}^2$, mais sur un certain domaine $A \subset \mathbb{R}^2$. On pose $G = (g_1, g_2) : \mathbb{R}^2 \to \mathbb{R}^2$.

On doit avoir $G^{-1} : \begin{cases} x_1 = d_1(u_1, u_2) \\ x_2 = d_2(u_1, u_2) \end{cases}$, donc il faut que G soit inversible. De plus, on devrait calculer la matrice jacobienne

$$J_{U_1U_2} = \begin{pmatrix} \frac{\partial d_1}{\partial u_1} & \frac{\partial d_1}{\partial u_2} \\ \frac{\partial d_2}{\partial u_2} & \frac{\partial d_2}{\partial u_2} \end{pmatrix}.$$

3.6 devient :

$$\iint_{\mathbb{R}^2 \cap G(A)} h(u_1, u_2) f_{X_1X_2}(d_1(u_1, u_2), d_2(u_1, u_2)) |\det(J_{U_1U_2}(u_1, u_2))| du_1 du_2.$$

Comme h est quelconque, on peut choisir $h(x) = \mathbb{1}_{\{g>f\}}(x)$ et puis $h(x) = \mathbb{1}_{\{g<f\}}$.
On obtient la formule de changement de variable :

Théorème 3.5.1 (Formule de changement de variables)**.**

$$f_{U_1U_2}(u_1, u_2) = f_{X_1X_2}(d_1(u_1, u_2), d_2(u_1, u_2)) |\det J_{U_1U_2}(u_1, u_2)| \mathbb{1}_{G(A)}(u_1, u_2).$$

Si on a deux variables aléatoires X_1 et X_2, quelle est la loi de $(X_1 + X_2)$?
On introduit

$$\begin{cases} U_1 = X_1 + X_2, U_2 = X_2 \end{cases} \implies \begin{cases} X_1 = U_1 - U_2 \\ X_2 = U_2 \end{cases}.$$

De plus,

$$J_{U_1U_2} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

On calcule

$$f_{U_1U_2}(u_1, u_2) = f_{X_1X_2}(u_1 - u_2, u_2)$$

et on obtient

$$f_{U_1}(u_1) = \int f_{U_1,U_2}(u_1, u_2)du_2 = \int f_{X_1X_2}(u_1 - u_2, u_2)du_2 = \int_{\mathbb{R}} f_{X_1}(u_1 - u_2)f_{X_2}(u_2)du_2.$$

Il s'agit du **produit de convolution**.

CHAPITRE 4

CONVERGENCE DES VARIABLES ALÉATOIRES

On a une suite de variables aléatoires $\{X_n\}_{n=1}^{\infty}$ toutes définies sur $(\Omega, \mathscr{A}, \mathbb{P})$. 02-10-2023
Comment calculer $\lim_{n\to\infty} X_n$?

1. Convergence “en probabilité” ;
2. Convergence “en moyenne” ;
3. Convergence “presque sûre” ;
4. Convergence “en loi” $\longrightarrow$ Théorème central limite.

4.1 Convergence en probabilité

Définition 4.1.1. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires définies toutes sur $(\Omega, \mathscr{A}, \mathbb{P})$. On dira que X_n converge en probabilité vers la variable aléatoire X et on écrira

$$X_n \xrightarrow{\mathbb{P}} X$$

si

$$\forall \varepsilon > 0, \mathbb{P}(|X_n - X| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0, \text{ i.e. } \lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

4.2 Convergence en moyenne L^p

Définition 4.2.1. On dira que X_n converge en moyenne L^p vers X si

$$\lim_{n\to\infty} \|X_n - X\|_{L^p} = 0 \text{ ou}$$

$$\lim_{n\to\infty} \left(\int |X_n(\omega) - X(\omega)|^p d\mathbb{P} \right).$$

Proposition 4.2.1. La convergence en moyenne L^p entraîne la convergence en probabilité.

Démonstration. Soit $\varepsilon > 0$. Calculons

$$\mathbb{P}(|X_n - X| > \varepsilon) \overset{4.2.2}{\leq} \frac{1}{\varepsilon^p} \underbrace{\mathbb{E}[|X_n - X|^p]}_{\underset{n\to\infty}{\longrightarrow} 0}.$$

□

Proposition 4.2.2 (Inégalité de Tchebychev)**.** Pour $\alpha > 0$ et $p \geq 1$, on a

$$\mathbb{P}(|X|^p > \alpha) \leq \frac{1}{\alpha}\mathbb{E}[|X|^p].$$

Démonstration.

$$\begin{aligned} \mathbb{E}[|X|^p] &= \int_\Omega |X|^p d\mathbb{P} = \int_{\{|X|^p > \alpha\}} |X|^p d\mathbb{P} + \int_{\{|X|^p < \alpha\}} |X|^p d\mathbb{P} \\ &\geq \mathbb{P}(|X|^p > \alpha) + \underbrace{\int_{\{|X|^p \leq \alpha\}} |X|^p d\mathbb{P}}_{\geq 0} \geq \mathbb{P}(|X|^p > \alpha). \end{aligned}$$

□

Proposition 4.2.3 (Propriétés de la convergence en probabilité)**.**

1. Supposons que le vecteur aléatoire $(X_n, Y_n)_{n=1}^\infty$ est tel que

$$\begin{cases} \mathbb{P}(|X_n - X| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0 \\ \mathbb{P}(|Y_n - Y| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0. \end{cases}$$

   Si $h : \mathbb{R}^2 \to \mathbb{R}$ est continue sur $\mathbb{R}^2$, alors la suite $h(X_n, Y_n)$ converge en probabilités vers $h(X, Y)$, i. e.

$$\mathbb{P}(|h(X_n, Y_n) - h(X, Y)| > \varepsilon) \underset{n\to\infty}{\longrightarrow} 0.$$

2. Si maintenant on a une suite $X_n \xrightarrow{\mathbb{P}} X$ et h est continue sur $\mathbb{R}$, alors $h(X_n) \xrightarrow{\mathbb{P}} h(X)$.
3. Si $(X_n, Y_n) \xrightarrow{\mathbb{P}} (X, Y)$, alors $X_n + Y_n \xrightarrow{\mathbb{P}} X + Y$, $X_nY_n \xrightarrow{\mathbb{P}} XY$, $\frac{X_n}{Y_n} \xrightarrow{\mathbb{P}} \frac{X}{Y}$ si $\mathbb{P}(Y = 0) = 0$.

4.3 Convergence presque sûre

Définition 4.3.1. On dira que la suite $\{X_n\}_{n=1}^\infty$ converge presque sûrement vers la variable aléatoire X et on écrira $X \xrightarrow{\text{P.S.}} X$ s'il existe un ensemble négligeable $N \in \mathscr{A}$ ($\mathscr{A}$ est la σ-algèbre sur Ω), $\mathbb{P}(N) = 0$ telle que

$$\forall \omega \in \Omega \setminus N, \text{ on a } X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega).$$

Il s'agit de la convergence entre nombres.

Remarque (Notation). Nous allons étudier $X \xrightarrow{\text{P.S.}} 0$. Sinon $Y_n = X_n - X$.

Nous allons introduire deux objets.

Définition 4.3.2. Soit $\{A_n\}_{n=0}^{\infty}$ une suite d'événements dans $\mathscr{A}$ et on définit

$$\lim_{n\to\infty} \inf A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$$
$$\lim_{n\to\infty} \sup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Si $\omega \in \lim_{n\to\infty} \sup A_n$, alors $\forall k \geq n, \omega \in \bigcup_{k=n}^{\infty}$.

Lemme (De Borel-Cantelli). *Si $\{A_n\}_{n=0}^{\infty}$ est une suite d'événements telle que*

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty,$$

alors

$$\mathbb{P}\left(\lim_{n\to\infty} \sup A_n\right) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) = 0,$$

c'est-à-dire l'ensemble des points qui sont dans une infinité des $\{A_n\}$ à probabilité 0 si et seulement si presque tout point est dans un nombre fini d'éléments.

Démonstration.

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right) \leq \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \sum_{k=n}^{\infty} \mathbb{P}(A_k) \underset{n\to\infty}{\longrightarrow} 0.$$

En effet, on a

$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) < \infty \implies \sum_{k=1}^{n-1} \mathbb{P}(A_k) + \sum_{k=n}^{\infty} \mathbb{P}(A_k) < \infty.$$

□

03-10-2023

Lemme (De Borel-Cantelli, deuxième). *Si $\{A_n\}$ sont* ***indépendantes*** *et $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty$, alors*

$$\mathbb{P}\left(\lim_{n\to\infty} \sup A_n\right) = 1.$$

Exemple. $\Omega = [0,1], \beta$ BOREL sur $[0,1], \mathbb{P}$ mesure de Lebesgue sur $[0,1]$,

$$A_n = \left(0, \frac{1}{n}\right].$$

Montrer que $\lim_{n\to\infty} \sup A_n = \emptyset$, mais $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = +\infty$.

Démonstration. En TD. On tâchera de montrer que les événements A_n sont **ne** sont **pas** indépendants. □

Exemple. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires telles que $\mathbb{P}(X_n = 0) = \frac{1}{n^2}$. Donc par Borel-Cantelli on sait que

$$\sum_{n=1}^{\infty} \mathbb{P}(X_n = 0) = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

et $\mathbb{P}\left(\lim_{n\to\infty} \sup\{X_n = 0\}\right) = 0$. Interpréter le résultat.

Solution. La probabilité que l'événement en question se produise une infinité de fois est 0 (Borel-Cantelli). Dans notre cas, presque sûrement (avec probabilité 1) $X_n \neq 0$ pour tout n sauf un nombre fini de n. □

Revenons à la convergence presque sûre. Etudions $X_n \xrightarrow{\text{P.S.}} 0$. Fixons $\varepsilon > 0$. On introduit

$$E_n = \{|X_n| > \varepsilon\}.$$

Notons

$$E(\varepsilon) = \lim_{n\to\infty} \sup E_n(\varepsilon) = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\},$$

$$D = \bigcup_{\varepsilon \geq 0} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}.$$

D **est l'ensemble des points qui ne convergent pas**.

Proposition 4.3.1. $X_n \xrightarrow{\text{P.S.}} 0$ si et seulement si $\mathbb{P}(D) = 0$ et si et seulement si

$$\forall \varepsilon > 0, \mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}\right) = 0.$$

Remarque. Quand $\varepsilon \longrightarrow 0$, la suite

$$\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}$$

est croissante.

Proposition 4.3.2. Si la suite $\{X_n\}_{n=1}^{\infty}$ vérifie $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \varepsilon) < \infty$, alors $X_n \xrightarrow{\text{P.S.}} 0$.

Démonstration. On utilise le lemme de Borel-Cantelli (cf 4.3). Alors on a

$$\mathbb{P}\left(\lim_{n\to\infty} \sup\{X_n > \varepsilon\}\right) = 0,$$

i. e.

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \{|X_k| > \varepsilon\}\right) = 0.$$

□

Exemple. Soit $([0,1], \beta\ (\text{Borel}), \lambda\ (\text{Leb}))$. On note $X_n(x) = \mathbb{1}_{[0,\frac{1}{n}]}, x \in [0,1]$.

Montrer que $X_n \xrightarrow{\text{P.S.}} 0$, mais $\sum_{n=1}^{\infty} \mathbb{P}(|X_n| > \varepsilon) = +\infty$.

Solution. Sur feuille. □

Proposition 4.3.3. Si $\{X_n\}$ converge presque sûrement, elle converge en probabilité. Par contre, si $\{X_n\}$ converge en probabilité, il existe une sous-suite qui converge presque sûrement.

Démonstration.

1. Soit $X_n \xrightarrow{\text{P.S.}} X$. L'ensemble de divergences est

$$D = \bigcup_{\frac{1}{l}} \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{|X_k| > \frac{1}{l}\right\}.$$

On voit que $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{|X_k| > \frac{1}{l}\right\} \subset D$. Alors

$$\begin{aligned} 0 = \mathbb{P}(D) \geq \mathbb{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} \left\{|X_k| > \frac{1}{l}\right\}\right) &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k=n}^{\infty} \left\{|X_k| > \frac{1}{l}\right\}\right) \\ \geq \lim_{n\to\infty} \mathbb{P}\left(\left\{|X_{k(n)}| > \frac{1}{l}\right\}\right) &\iff \lim_{n\to\infty} \mathbb{P}\left(\left\{|X_{k(n)}| > \frac{1}{l}\right\}\right) = 0, \end{aligned}$$

c'est-à-dire X_n converge en probabilité.

2. Soit $\varepsilon > 0$ et $\{\eta_k\}_{k=1}^{\infty}$ suite de nombres positifs tels que $\sum_{k=1}^{\infty} \eta_k < \infty$. Comme on a convergence en probabilité, $\forall k \geq 1$, il existe $n_k \geq 1$ tel que

$$\mathbb{P}(|X_{n_k}| > \varepsilon) < \eta_k.$$

Donc

$$\mathbb{P}\left(\sup_{k\geq l}\{|X_{n_k} > \varepsilon|\}\right) = \mathbb{P}\left(\bigcup_{k\geq l}\{|X_k| > \varepsilon\}\right) \leq \sum_{k=l}^{\infty} \mathbb{P}\left(\{|X_{n_k} > \varepsilon|\}\right) \leq \sum_{k=l}^{\infty} \eta_k \xrightarrow[l\to\infty]{} 0.$$

□

Lemme.

$$\bigcup_{k\geq n}\{|X_k|>\varepsilon\}=\{\sup_{k\geq n}|X_k|>\varepsilon\}.$$

04-10-2023

4.4 Loi des grands nombres

Il existe deux "versions" de la loi des grands nombres :
- ⋆ faible ;
- ⋆ forte.

Si $\{X_n\}$ est une suite de variables aléatoires, on considère $S_n=\dfrac{1}{n}\displaystyle\sum_{k=1}^{n}X_k$ (somme de Cèsaro).

Cadre On a une suite de variables aléatoires $\{X_n\}_{n=1}^{\infty}$ sur $(\Omega,\mathscr{A},\mathbb{P})$. On considère $S_n=\dfrac{1}{n}\displaystyle\sum_{k=1}^{n}X_k$.

4.4.1 Loi faible des grands nombres

Définition 4.4.1. Supposons que les variables aléatoires X_n soient *centrées*, i. e. X_n devient $X_n-\mathbb{E}[X_n]$ (donc les X_n auront l'espérance nulle). On dira que la suite $\{X_n\}$ satisfait **la loi faible des grands nombres** si la suite S_n converge vers 0 **en probabilité**, c'est-à-dire

$$\forall\varepsilon>0,\mathbb{P}(|S_n|>\varepsilon)\underset{n\to\infty}{\longrightarrow}0.$$

Proposition 4.4.1. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires de $L^2(\mathbb{P})$ qui sont deux-à-deux non-corrélées et centrées. Pour tout n, on pose :

$$\mathbb{V}(X_n)=\sigma_n^2<\infty$$

et on suppose que

$$\frac{1}{n^2}\sum_{k=1}^{n}\sigma_k^2\underset{n\to\infty}{\longrightarrow}0.$$

Alors

$$\frac{1}{n}\sum_{k=1}^{n}X_k\xrightarrow{L^2(\mathbb{P})}0$$

et donc elle converge aussi en probabilités.

Définition 4.4.2. Soient X_1,X_2 deux variables aléatoires. On introduit

$$\mathrm{Cov}(X_1,X_2)=\mathbb{E}[X_1X_2]-\mathbb{E}[X_1]\mathbb{E}[X_2].$$

Remarque. Si X_1,X_2 sont indépendantes, on a que $\mathbb{E}[X_1X_2]=\mathbb{E}[X_1]\mathbb{E}[X_2]$.

Définition 4.4.3. On dira que deux variables aléatoires X_1, X_2 sont non-corrélées si $\mathrm{Cov}(X_1, X_2) = 0$.

Deux variables indépendantes sont non-corrélées, mais le contraire est en général faux.

Remarque (Rappel). Si X est une variable aléatoire, on définit :

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

Proposition 4.4.2. 1. $\mathbb{V}(\alpha X) = \alpha^2 \mathbb{V}(X)$.

Remarque.

$$\mathbb{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathbb{V}(X_k) + 2 \sum_{1 \leq j \leq k \leq n} \mathrm{Cov}(X_j, X_k).$$

Démonstration de la proposition 4.4.1. On rappelle la définition de la convergence L^2.

$$\left\| \frac{1}{n} \sum_{k=1}^{n} X_k \right\|_{L^2}^2 = \left(\int \frac{1}{n^2} \left(\sum_{k=1}^{n} X_k \right)^2 \right)^{\frac{1}{2}}.$$

On a

$$\left\| \frac{1}{n} \sum_{k=1}^{n} X_k \right\|_{L^2}^2 = \mathbb{V}\left(\frac{1}{n} \sum_{k=1}^{n} X_k \right) = \frac{1}{n^2} \mathbb{V}\left(\sum_{k=1}^{n} X_k \right) = \frac{1}{n^2} \sum_{k=1}^{n} \mathbb{V}(X_k),$$

car les variables aléatoires sont non-corrélées.
De plus, on a

$$\frac{1}{n^2} \sum_{k=1}^{n} \mathbb{V}(X_k) = \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 \underset{n \to \infty}{\longrightarrow} 0.$$

On vient de montrer que

$$S_n = \frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{L^2} 0.$$

□

Proposition 4.4.3. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires $L^2(\mathbb{P})$ deux-à-deux non-corrélées. Posons $\mu_n = \mathbb{E}[X_n]$ et supposons que

1. $\frac{1}{n} \sum_{k=1}^{n} \mu_k \underset{n \to \infty}{\longrightarrow} \mu$;
2. $\frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 \underset{n \to \infty}{\longrightarrow} 0$.

Alors

$$\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{L^2} \mu.$$

Démonstration. On utilise les théorèmes précédents. On écrit :

$$\frac{1}{n}\sum_{k=1}^{n}\overbrace{(X_k-\mu_k)}^{\text{centrée}}$$

et on applique la proposition 4.4.1. □

Proposition 4.4.4. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires de L^2 centrées **qui ont la même loi** et sont deux-à-deux non-corrélées. Alors

$$\frac{1}{n}\sum_{k=1}^{n}X_k \xrightarrow{L^2} 0.$$

Démonstration. On applique 4.4.1. □

Proposition 4.4.5. ⚠ Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires de classe $L^2(\mathbb{P})$ i. i. d. (indépendantes de même loi, *independent identically distributed*). Alors

$$\frac{1}{n}\sum_{k=1}^{n}X_k \xrightarrow{L^2} \mu, \text{ où } \mu = \mathbb{E}[X_1].$$

Remarque. La dernière version reste vraie pour des variables aléatoires i. i. d. aussi si les variables sont de classe $L^1(\mathbb{P})$.

Remarque. On peut remplacer L^2 par L^1.

4.4.2 Loi forte des grands nombres

Théorème 4.4.1. *Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires de classe L^1 centrées et i.i.d. Alors*

$$\frac{1}{n}\sum_{k=1}^{n}X_k \xrightarrow{P.S.} 0.$$

Remarque. Si les variables ne sont pas centrées, alors

$$\frac{1}{n}\sum_{k=1}^{n}X_k \xrightarrow{\text{P.S.}} \mathbb{E}[X_1].$$

Démonstration. Cf polycopié. □

4.4.3 Lien entre convergences

4.5 Convergence en loi

4.5.1 Fonction caractéristique

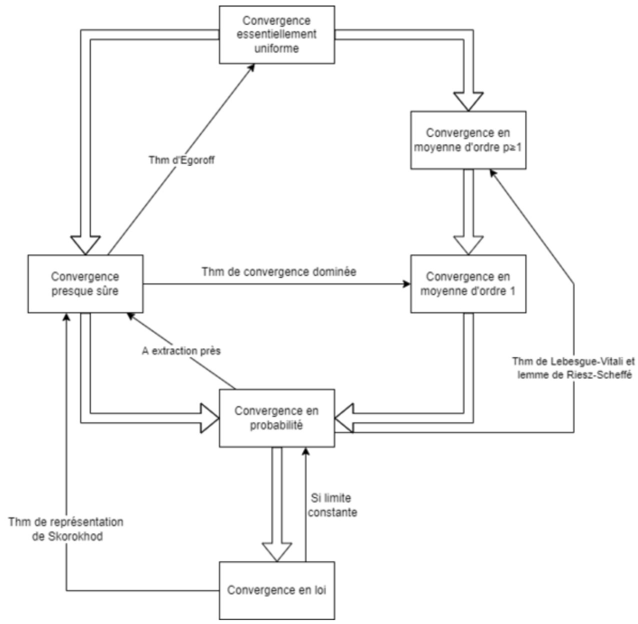


FIGURE 4.1 – Schéma résumant les liens entre convergences de variables (plus compliqué que celui donné en cours). La double flèche est une implication et la flèche simple est une implication dans certains cas.

Définition 4.5.1. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. On définit la fonction caractéristique de la variable aléatoire réelle X de la manière suivante :

$$\varphi_t(X) = \mathbb{E}\left[e^{itX}\right] = \int_\Omega e^{itX} d\mathbb{P} = \int_\mathbb{R} e^{itx} dP_X(x).$$

Remarque (Cas particulier)*.* Si X est à densité $f_X(x)$, alors :

$$\varphi_X(t) = \int_\mathbb{R} e^{itx} f_X(x) dx.$$

Il s'agit de la transformée de Fourier.

Proposition 4.5.1 (Propriétés de la fonction caractéristique)**.**

1. φ_X est définie et continue $\forall t \in \mathbb{R}$, en particulier elle est **uniformément continue**.
2. $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$;
3. Si $\{X_i\}_{i=1}^\infty$ sont indépendantes et on écrit $S_n = \displaystyle\sum_{i=1}^n X_i$, alors

$$\varphi_{X_n}(t) = \varphi_{X_1}(t) - \varphi_{X_n}(t).$$

Si en plus les $\{X_i\}$ ont la même loi, alors

$$\varphi_{S_n}(t) = \varphi_{X_1}(t)^n.$$

Démonstration.

1. Il faut que $\varphi_X \in L^1(\mathbb{P})$, i. e. $\varphi_X \in L^1(\mathbb{R}, P_X)$. Alors

$$\int_\Omega \left|e^{itX}\right| d\mathbb{P} = \int d\mathbb{P} = 1.$$

2. Montrons que φ_X est uniformément continue.
On a

$$\begin{aligned}\varphi_X(t) - \varphi_X(t_0) = \left|\int \left(e^{itX} - e^{it_0X}\right)\right| &= \int \left|iXe^{i\xi X}\right| |t - t_0| \, d\xi \\ &\leq |t - t_0| \int |X| \, d\mathbb{P},\end{aligned}$$

ce qui veut dire que la fonction est lipschitzienne, donc uniformément continue.

3. On a

$$\varphi_{S_n} = \int_\Omega e^{it}(X_1 + \cdots + X_n) d\mathbb{P} = \int_\Omega e^{itX_1} \dots e^{itX_n} d\mathbb{P} = \mathbb{E}[e^{itX_1} \dots e^{itX_n}] = \prod_{i=1}^n \mathbb{E}[e^{itX_i}] = \prod_{i=1}^n \varphi_{X_i}(t).$$

□

Exemple (Calculs des fonctions caractéristiques).

1. X est une variable aléatoire binomiale $B(n, p)$. On rappelle que

$$\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}.$$

Alors

$$\varphi_X(t) = \int e^{itX} d\mathbb{P} = \int e^{itx} dP_X(x) = \sum_{k=0}^n e^{itk} C_n^k p^k (\underbrace{1 - p}_{=q})^{n-k} = (q + pe^{it})^n.$$

2. X est de Bernoulli $B(p)$, avec $\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$. Alors

$$\varphi_X(t) = e^{it} p + q.$$

3. Variable de Gauss avec X qui a la loi $\mathscr{N}(0, 1)$.

$$\varphi_X(t) = \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} e^{itx} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int (\cos(tx) + i \sin(tx)) e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_\mathbb{R} i \cos(tx) e^{-\frac{x^2}{2}} dx,$$

car

$$i \int \sin(tx) e^{-\frac{x^2}{2}} dx = 0 \text{ (fonction impaire)}.$$

On dérive $\varphi_X(t)$. Si

$$\varphi_X(t) = \int_\mathbb{R} \cos(tx) e^{-\frac{x^2}{2}} dx,$$

alors

$$\varphi_X'(t) = \int_{\mathbb{R}} -x\sin(tx)e^{-\frac{x^2}{2}}\,dx = \underbrace{\left[\sin(tx)e^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty}}_{=0} - \int_{-\infty}^{+\infty} t\cos(tx)e^{-\frac{x^2}{2}}\,dx = -t\varphi_X(t),$$

ce qui montre que $\varphi_X'(t) = -t\varphi_X(t)$, i. e.

$$\frac{d\varphi_X(t)}{dt} = -t\varphi_X(t).$$

C'est une équation différentielle du premier ordre. On utilise la méthode des quadratures pour la résoudre. On a :

$$\frac{d\varphi_X(t)}{\varphi_X(t)} = -tdt,$$

ce qui donne, quand on intègre,

$$\int_{\varphi_X(0)}^{\varphi_X(t)} \frac{d\varphi_X(t)}{\varphi_X(t)} = -\int_0^t tdt,$$

ce qui donne $\ln(\varphi_X(t)) = -\dfrac{t^2}{2}$, i. e. $\varphi_X(t) = e^{-\frac{t^2}{2}}$.

Proposition 4.5.2. Supposons que $\mathbb{E}[|X|^n] < \infty$ (moment d'ordre n fini). Alors $\varphi_X^{(r)}(t)$ existe pour $r \leq n$ où $\varphi_X^{(r)}$ dénote la dérivée d'ordre r de φ_X par rapport à t et :

$$\varphi_X^{(r)}(t) = \int_{\mathbb{R}} (ix)^r e^{itx} dP_X(x)$$

et

$$\mathbb{E}[X^r] = \frac{\varphi_X^{(r)}(0)}{i^r} \text{ pour } r \leq n.$$

De plus, on a :

$$\varphi(t) = \sum_{k=0}^{n} \frac{(it)^k}{k!}\mathbb{E}[X^k] + \underbrace{\frac{(it)^n}{n!}\varepsilon_n(t)}_{\text{erreur}},$$

avec $|\varepsilon_n(t)| \leq 3\mathbb{E}[|X|^n] \underset{t\to 0}{\longrightarrow} 0$.

Remarque. Si $\mathbb{E}[|X|^n] < \infty \implies \mathbb{E}[|X|^r] < \infty$, avec $r \leq n$.

Démonstration. Soit $n \geq 1$, on fait la preuve pour $r = 1$. On élimine l'indice X dans $\varphi_X(t)$. On doit contrôler

$$\lim_{h\to 0}\frac{\varphi(t+h)-\varphi(t)}{h}=\lim_{h\to 0}\frac{[e^{i(t+h)x}-e^{itx}]}{h}dP_X(t)=\int_{\mathbb{R}}\lim_{h\to 0}\left[\frac{e^{i(t+h)x}-e^{itx}}{h}\right]dP_X(x). \tag{4.1}$$

On développe la quantité $e^{ix(t+h)}$ dans le point 0 :

$$\frac{e^{ix(t+h)}-e^{itx}}{h}=\frac{e^{ixt}+e^{ixt}ixh+O(h^2)-e^{itX}}{h}\approx e^{ixt}ix+O(h).$$

L'égalité 4.1 devient :

$$4.1=\int_{\mathbb{R}} ixe^{itx}dP_X(x)<\infty.$$

□

20-10-2023 **Théorème 4.5.1.** *Soient X et Y deux variables aléatoires telles que $\varphi_X(t)=\varphi_Y,\forall t\in\mathbb{R}$. Donc $P_X=P_Y$.* ***Les fonctions caractéristiques caractérisent complétement la loi.***

Théorème 4.5.2 (Formule d'inversion de Levy)**.** *Soit X une variable aléatoire de fonction caractéristique*

$$\varphi_X(t)=\int_{\mathbb{R}} e^{itx}dP_X(x).$$

Soient x et y deux points où la fonction de répartition F_X est continue. Alors

$$F_X(b)-F_X(a)=\lim_{c\to+\infty}\frac{1}{2\pi}\int_{-c}^{c}\frac{e^{-ita}-e^{-itb}}{it}\varphi_X(t)dt \text{ (intégrale impropre de Riemann).}$$

On rappelle que la loi P_X est la mesure de Lebesgue-Stieltjes construite avec F_X, i. e.

$$P_X((a,b])=F_X(b)-F_X(a).$$

Théorème 4.5.3. *Soit $\varphi_X\in L^1(\mathbb{R})$, i. e. $\int_{\mathbb{R}}|\varphi_X(t)|\,dt<\infty$. Alors la densité de X existe et*

$$f_X(x)=\frac{1}{2\pi}\int_{\mathbb{R}} e^{itx}\varphi_X(t)dt.$$

CHAPITRE 5

CONVERGENCE EN LOI

5.1 Convergence de mesures bornées sur $\mathbb{R}^n$

On note $\mathscr{M}$ l'ensemble de mesures positives sur $\mathbb{R}^n$ muni de la tribu de Borel.

On note $\mathscr{M}(b)$ l'ensemble de mesures de $\mathscr{M}$ telles que si $\mu \in \mathscr{M}(b)$, alors $\mu(\mathbb{R}^n) \leq b$. Ce sont les mesures bornées par b.

Enfin on note $\mathscr{M}^1$ les mesures de probabilité $\mu \in \mathscr{M}^1$ telle que $\mu(\mathbb{R}^n) = 1$.

Avant de construire les topologies, on introduit deux classes de fonctions :

1. $\mathscr{C}_0(\mathbb{R}^n)$ qui sont les fonctions tendant vers 0 à l'infini, c'est-à-dire

$$\forall \varepsilon > 0, \exists K_\varepsilon \text{ compact tel que } |f(x)| < \varepsilon, \forall x \in (K_\varepsilon)^C.$$

2. $\mathscr{C}_b(\mathbb{R}^n)$, l'ensemble des fonctions bornées :

$$\forall x \in \mathbb{R}, |f(x)| \leq b.$$

Définition 5.1.1. Sur $\mathscr{M}$ on définit deux topologies, `faible` et `étroite`. Il s'agit des deux topologies moins fines qui rendent continues les applications :

$$\mu \longmapsto \int f d\mu, f \in \mathcal{C}_0(\mathbb{R}^n) \text{ (faible)},$$
$$\mu \longmapsto \int df\mu, f \in \mathscr{C}_b(\mathbb{R}^n) \text{ (étroite)}.$$

Pour $\mu_1, \mu_2 \in \mathscr{M}$, on définit

$$d(\mu_1, \mu_2) = \sum_{k=1}^{\infty} \frac{\left|\int f_k d\mu_1 - \int f_k d\mu_2\right|}{2^k},$$

où $\{f_k\}$ est une suite dense dans $\mathscr{C}_{0,b}(\mathbb{R}^n)$.

5.1.1 Convergence

Définition 5.1.2. Prenons une suite $\{\mu_n\} \in \mathscr{M}$.

On dira que μ_n converge faiblement vers $\mu \in \mathscr{M}$ si

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu, \forall f \in \mathscr{C}_0(\mathbb{R}^n).$$

Elle converge `étroitement` si :

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu, \forall f \in \mathscr{C}_b(\mathbb{R}^n).$$

Remarque. La convergence étroite implique la convergence faible.

En probabilité, on utilisera surtout la convergence étroite.

Proposition 5.1.1.

1. Sur $\mathscr{M}^1$, les deux topologies coïncident.
2. *"Théorème de Banach-Ailogu"* L'espace $\mathscr{M}(b)$ est compact pour la topologie faible, c'est-à-dire toute suite de mesures bornées (par b) admet une sous-suite convergente faiblement vers une mesure $\mu \in \mathscr{M}(b)$.
3. Soit $\{\mu_n\}$ et $\mu \in \mathscr{M}(b)$ telle que $\mu_n \longrightarrow \mu$ faiblement. Alors la suite $\{\mu_n\}$ converge étroitement vers μ si

$$\lim_{k\to\infty} \mu_k(\mathbb{R}^n) = \mu(\mathbb{R}^n).$$

La réciproque est aussi vraie.

Définition 5.1.3. Soit μ un élément de $\mathscr{M}$. Un borélien A est appelé de μ-continuité si

$$\mu(\underbrace{\partial A}_{\text{frontière}}) = 0.$$

Théorème 5.1.1. *Soit $\{\mu_n\}_{n=1}^{\infty}$ une suite de mesures dans $\mathscr{M}(b)$. Les deux assertions sont alors équivalentes :*

1. *μ_n converge étroitement vers μ ;*
2. *Pour tout borélien A de μ-continuité, on a*

$$\lim_{n\to\infty} \mu_k(A) = \mu(A).$$

On va parler maintenant de la convergence en loi.

Soit $\forall n \in \mathbb{N}$ une suite de variables aléatoires X_n définies sur les espaces $(\Omega_n, \mathscr{A}_n, \mathbb{P}_n)$. Soit X une autre variable aléatoire réelle à valeurs dans $\mathbb{R}^d$ et définie sur $(\Omega, \mathscr{A}, \mathbb{P})$.

Définition 5.1.4. La suite $\{X_n\}_{n=1}^{\infty}$ converge en loi (en distribution) vers X si la suite des lois P_{X_n} converge étroitement vers P_X, et on écrira :

$$P_{X_n} \xrightarrow{\mathscr{L}} P_X.$$

Définition 5.1.5. Soit $\{X_k\}_{k=1}^{\infty}$ une suite de variables aléatoires sur $(\Omega_k, \mathscr{A}_k, \mathbb{P}_k)$ et soit μ une probabilité sur $\mathbb{R}^d$. La suite $\{X_k\}_{k=1}^{\infty}$ converge étroitement vers μ si la suite P_{X_n} converge étroitement vers μ et on écrira

$$X_n \xrightarrow{\mathscr{L}} \mu.$$

Remarque. Ici, on ne connaît pas forcément la variable aléatoire limite. On connaît seulement la distribution limite.

Théorème 5.1.2 (Levy). *A nouveau on a la suite $\{X_k\}_{k=1}^{\infty}$ avec $X_k \in (\Omega_k, \mathscr{A}_k, \mathbb{P}_k)$.*

1. *Supposons que X_k converge en loi vers la variable aléatoire X définie sur $(\Omega, \mathscr{A}, \mathbb{P})$. Alors*

$$\lim_{k\to\infty} \varphi_{X_k}(t) = \varphi_X(t), \forall t \in \mathbb{R}.$$

2. *Supposons que $\varphi_{X_k}(t)$ converge simplement (pour tout $t \in \mathbb{R}$) vers une fonction $\varphi(t)$ qui est continue en 0. Alors φ est la fonction caractéristique d'une mesure de probabilité μ et P_{X_n} converge étroitement vers μ, c'est-à-dire on a*

$$P_{X_k} \xrightarrow{\mathscr{L}} \mu.$$

*En plus, il existe une variable aléatoire X **non unique** telle que*

$$X_k \xrightarrow{\mathscr{L}} X.$$

Démonstration. 23-10-2023

1.

$$\varphi_{X_k}(t) = \int_{\mathbb{R}} \underbrace{e^{itx_k}}_{\substack{\text{continue}\\ \text{bornée}}} dP_{X_k}(x_k) \underset{k\to+\infty}{\longrightarrow} \int_{\mathbb{R}} e^{itx} dP_X(x).$$

2. On dénote $P_{X_n} = P_n$. Par compacité, il existe une sous-suite n_k telle que $P_{n_k} \longrightarrow \mu$, avec μ une mesure bornée ($\in \mathscr{M}(b)$), et cette convergence est faible. L'objectif est de démontrer que μ est une mesure de probabilité. On y arrive si on montre que

$$\lim_{k\to+\infty} P_{n_k}(\mathbb{R}^d) = \mu(\mathbb{R}^d).$$

Grâce à ce résultat-là, on montrera aussi que la convergence est étroite.
Astuce Soit ν une mesure bornée quelconque et soit φ_ν sa fonction caractéristique. On a

$$\varphi_\nu(t) = \int_{\mathbb{R}^d} e^{itx} d\nu(x).$$

On construit (pour $\nu > 0$)

$$\frac{1}{u}\int_0^u \varphi_\nu(t)dt = \frac{1}{u}\int_0^u \left(\int_{\mathbb{R}^d} e^{itx}d\nu(x)\right) \overset{\text{Fubini}}{=} \frac{1}{u}\int_{\mathbb{R}} \left(\int_0^u e^{itx}dt\right) d\nu(x) = \frac{1}{u}\int_{\mathbb{R}} \frac{e^{ixu}-1}{iux} d\nu(x).$$

On applique ces calculs à la mesure μ :

$$\frac{1}{u}\int_0^u \varphi_\mu(t)dt = \int_{\mathbb{R}} \underbrace{\frac{e^{ixu}-1}{iux}}_{\substack{\text{continue,}\\ \in \mathscr{C}_b(\mathbb{R}^d)}} d\mu(x) = \lim_{k\to+\infty} \int_{\mathbb{R}^d} \frac{e^{iux}-1}{iux} dP_{n_k} = \lim_{k\to+\infty} \frac{1}{u}\int_0^u \varphi_{P_{n_k}}(t)dt. \quad (5.1)$$

On observe que $\left|\varphi_{P_{n_k}}(t)\right| \leq \varphi_{P_{n_k}}(0) = 1$, donc on peut passer à la limite dans l'intégrale :

$$5.1 = \frac{1}{u}\int_0^u \lim_{k\to\infty} \varphi_{P_{n_k}}(t)dt = \frac{1}{u}\int_0^u \varphi(t)dt.$$

On a trouvé que

$$\frac{1}{u}\int_0^u \underbrace{\varphi_\mu(t)dt}_{\substack{\text{cont. en 0}\\ \text{car fct car.}}} = \frac{1}{u}\int_0^u \underbrace{\varphi(t)}_{\substack{\text{cont. en 0}\\ \text{par hyp.}}} dt.$$

Par le théorème de la valeur moyenne,

$$\lim_{u\to 0} \frac{1}{u}\int_0^u \varphi_\mu(t)dt = \varphi_\mu(0),$$
$$\lim_{u\to 0} \frac{1}{u}\int_0^u \varphi(t)dt = \varphi(0).$$

Il faut se rappeler qu'on doit montrer que $\mu(\mathbb{R}^d) = 1$. En effet,

$$\mu(\mathbb{R}^d) = \int_{\mathbb{R}^d} e^{i0x}d\mu(x) = \varphi_\mu(0) = \varphi(0) \overset{\text{hyp}}{=} \lim_{k\to\infty} \varphi_{P_{n_k}}(0) = \lim_{k\to\infty} \int_{\mathbb{R}^d} e^{i0x}dP_{n_k}(x) = 1.$$

Donc P_{n_k} converge aussi étroitement vers μ (mesure de probabilité).

Par la première partie du théorème de Levy 1, la convergence étroite de $P_{n_k} \longrightarrow \mu$ entraîne que $\varphi_{P_{n_k}}(t) \longrightarrow \varphi_\mu(t)$.
Mais aussi $\varphi_{P_{n_k}}(t) \longrightarrow \varphi(t)$. Donc on a établi que $\varphi_\mu(t) = \varphi(t)$. Pour une autre sous-suite, on a $\varphi_{P_{n_j}} \longrightarrow \varphi_{\mu'}(t) = \varphi(t)$.
Donc toutes les fonctions caractéristiques correspondant à des μ différentes sont *égales*. Il y a une correspondance biunivoque entre φ et μ. On en déduit que toutes les "μ" sont égales, ce qui implique que toutes les sous-suites qu'on peut engendrer à partir de P_n ont la même limite. On a ainsi montré que

$$P_n \overset{\text{étroite}}{\longrightarrow} \mu.$$

□

Question *Y a-t-il une variable aléatoire sur un certain espace probabilisé qui a μ comme loi ?*
On prend $(\Omega, \mathscr{A}, \mathbb{P}) = (\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d), \mu)$ et $X : \Omega \longrightarrow \mathbb{R}^d$ est telle que $X = \text{id}$.

Proposition 5.1.2. Supposons que $X_n \xrightarrow{\mathscr{L}} X$ (en loi). Soit f une fonction continue de $\mathbb{R}^d$ à valeurs dans $\mathbb{R}$. Alors

$$f(X_n) \xrightarrow{\mathscr{L}} f(x).$$

Théorème 5.1.3 (La convergence en probabilité implique la convergence en loi)**.** *Soit $\{X_n\}$ une suite de variables aléatoires définies sur le même espace $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d$. Supposons que X_n converge en probabilité vers la variable aléatoire X. Alors X_n converge aussi en loi vers X.*

Démonstration. Si l'on veut montrer que $X_n \xrightarrow{\mathscr{L}} X$, il suffit, pour $f \in \mathscr{C}_b(\mathbb{R}^d)$, de montrer que

$$\int f dP_{X_n} \underset{n \to +\infty}{\longrightarrow} \int f dP_X \text{ (étroite).}$$

En effet,

$$\begin{aligned}
\left|\int f dP_{X_n} - \int f dP_X\right| &\overset{\text{transfert}}{=} \left|\int_\Omega f \circ X_n d\mathbb{P} - \int_\Omega f \circ X_n d\mathbb{P}\right| = \left|\int [f \circ X_n - f \circ X] d\mathbb{P}\right| \\
= \int_{\{|f\circ X_n - f \circ X| > \varepsilon\}} &|f \circ X_n - f \circ X|\, d\mathbb{P} + \int_{\{|f\circ X_n - f\circ X| < \varepsilon\}} |f \circ X_n - f \circ X|\, d\mathbb{P} \\
&\leq \varepsilon + 2\,\|f\|_\infty \underbrace{\mathbb{P}(|f \circ X_n - f \circ X| > \varepsilon)}_{\longrightarrow 0 \text{ quand } n \longrightarrow \infty}.
\end{aligned}$$

□

Théorème 5.1.4. *Soit $\{X_n\}$ une suite de variables aléatoires toutes définies sur un même $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}^d$. Supposons que $X_n \xrightarrow{\mathscr{L}}$ constante $= c$. Alors $X_n \xrightarrow{\mathbb{P}} c$.*

Démonstration. $\mu_n \xrightarrow{\mathscr{L}} \mu$ si et seulement si pour tout borélien A de μ-continuité, $\mu_n(A) \longrightarrow \mu$.

On sait que $P_{X_n} \xrightarrow{\mathscr{L}} \delta_{\{c\}}$ (loi de la variable aléatoire limite $X = c$). On dénote avec $B(c, \varepsilon)$ la boule fermée de centre c et de rayon ε. On note que $\delta_{\{c\}}(\partial B(c, \varepsilon)) = 0$. Donc $B(c, \varepsilon)$ est un ensemble de $\delta_{\{c\}}$-continuité. Donc par la définition 5.1.3, on sait que

$$\lim_{n \to \infty} P_{X_n}(B(c, \varepsilon)) = \delta_{\{c\}}(B(c, \varepsilon)) = 1,$$

donc

$$\lim_{n \to \infty} P_X(x, |x - c| > \varepsilon) = 0.$$

□

Exemple. Sur $(\Omega, \mathscr{A}, \mathbb{P})$ on a une variable aléatoire de Bernoulli X de paramètre $\dfrac{1}{2}$. On pose $X_n = X$. On a que $X_n \xrightarrow{\mathscr{L}} X$. On considère $Y = 1 - X$. Elle est encore de Bernoulli de paramètre $\dfrac{1}{2}$ et $X_n \xrightarrow{\mathscr{L}} Y$. Montrer que X_n ne converge pas en probabilité vers Y.

Démonstration. $\forall f \in \mathscr{C}_b$, on doit montrer que $\int f dP_{X_n} \longrightarrow \int f dP_Y$. On a

$$\underbrace{f(0)\frac{1}{2} + f(1)\frac{1}{2}}_{X} = \underbrace{f(1)\frac{1}{2} + f(0)\frac{1}{2}}_{Y},$$

donc X_n converge en probabilité vers Y.
On a

$$\mathbb{P}(|X_n - Y| > \varepsilon) = \mathbb{P}(|X - Y| > \varepsilon) = \mathbb{P}(|X - 1 + X| > \varepsilon) = \mathbb{P}(|2X - 1| > \varepsilon).$$

On a $\mathbb{P}(|2X - 1| = 1) = 1$, donc X_n ne converge pas en probabilité vers Y. □

Théorème 5.1.5. *Soit $\{X_n\}$ une suite de variables aléatoires définies sur des espaces $(\Omega_n, \mathscr{A}_n, \mathbb{P}_n)$ avec chacune une fonction de répartition F_{X_n}. Soit X une autre variable aléatoire définie sur $(\Omega, \mathscr{A}, \mathbb{P})$ avec fonction de répartition F_X. La suite $\{X_n\}$ converge en loi vers X si et seulement si $F_{X_n}(t)$ converge vers $F_X(t)$ dans les points de continuité de cette dernière.*

Démonstration.

1. *Partie nécessaire.* On suppose que $X_n \xrightarrow{\mathscr{L}} X$. Soit t un point où F_X est continue. On veut montrer que $F_{X_n}(t) \longrightarrow F_X(t)$.
   Prenons l'ensemble $(-\infty, t]$. Cet ensemble-là est de P_X-continuité, car

   $$P_X(\{t\}) = F_X(t) - F_{X-0}(t) = 0.$$

   Donc

   $$\lim_{n\to\infty} F_{X_n}(t) = \lim_{n\to\infty} P_{X_n}((-\infty, t]) = P_X((-\infty, t]) = F_X(t).$$

2. *Partie suffisante.* Par hypothèse, $F_{X_n}(t) \longrightarrow F_X(t)$ et F_{X_n} est continue en t. On va montrer que $X_n \xrightarrow{\mathscr{L}} X$.
   On sait que les points où F_X n'est pas continue sont au plus dénombrables. Etant donné $f \in \mathscr{C}_b(\mathbb{R})$, on introduit une fonction en escalier $g : \mathbb{R} \longrightarrow \mathbb{R}$ de type

   $$g(x) = \sum_{j=1}^{n} \alpha_j \mathbb{1}_{[a_j, b_j]}, a_j < b_j \leq a_{j+1} < b_{j+1}$$

   et les points $\{a_j, b_j\}$ sont de continuité par F_X et en plus

   $$\|F_X(t) - g(t)\|_\infty < \varepsilon.$$

   On a

   $$\int_{\mathbb{R}} g(t) dP_{X_n}(t) = \sum_{j=1}^{k} \alpha_j P_{X_n}([a_j, b_j]) = \sum_{j=1}^{k} \alpha_j [F(b_j) - F(a_j)].$$

   On a alors :

   $$\begin{aligned}\lim_{n\to\infty} \int_{\mathbb{R}} g(t) dP_{X_n}(t) &= \lim_{n\to\infty} \sum_{j=1}^{k} \alpha_j [F_{X_n}(b_j) - F_{X_n}(a_j)] \\ &= \sum_{j=1}^{k} \alpha_j [F_X(b_j) - F_X(a_j)] = \int_{\mathbb{R}} g(t) dP_X(t).\end{aligned}$$

Soit $f \in \mathscr{C}_b(\mathbb{R})$. On écrit

$$\left|\int f dP_{X_n} - \int f dP_X\right| \leq \left|\int (f-g) dP_{X_n}\right| + \left|\int g dP_{X_n} - \int g dP_X\right| + \left|\int (g-f) dP_X\right|$$
$$\leq 2\left\|f-g\right\|_\infty + \left|\int g dP_{X_n} - \int g dP_X\right| \leq 2\varepsilon + \left|\int g dP_{X_n} - \int g dP_X\right|.$$

□

Lemme (De Schiffe, pour les variables à densité)**.** *Soit $\{X_n\}$ une suite de variables aléatoires définies sur $(\Omega, \mathscr{A}_n, \mathbb{P}_n)$ et avec des densités f_{X_n}. Si la suite f_{X_n} converge Lebesgue-presque partout vers une fonction f telle que $\int f d\,\mathrm{Leb}$ est croissante, alors la suite X_n converge en loi vers la mesure $f d\,\mathrm{Leb}$.*

5.1.2 Variables discrètes

Proposition 5.1.3. Soit $\{X_n\}$ suite de variables discrètes sur le même espace $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs dans $\mathbb{Z}$. On a l'équivalence entre les assertions suivantes :

1. $X_n \xrightarrow{\mathscr{L}} X$;
2. $\lim\limits_{n\to\infty} \mathbb{P}(\underset{r\in\mathbb{Z}}{X_n = r}) = \mathbb{P}(X = r)$.

Théorème 5.1.6 (Central limite)**.** *Soit $\{X_n\}$ une suite de variables aléatoires indépendantes et de même loi et de classe $L^2(\mathbb{P})$. Posons $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \mathbb{V}(X_1)$ $(0 < \sigma)$. Soit $S_n = \sum_{j=1}^{n} X_j$. On écrit :*

$$\frac{S_n - \mu_n}{\sigma\sqrt{n}} = \frac{\sum_{j=1}^{n} X_j - n\mu}{\sigma\sqrt{n}}.$$

Alors

$$\mathbb{P}\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \leq t\right) = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{t} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx.$$

On aura

$$\frac{S_n - \mu n}{\sigma\sqrt{n}} \xrightarrow{\mathscr{L}} \mathscr{N}(\mu, \sigma^2).$$

Démonstration. On introduit

$$\overline{X_n} = \frac{X_n}{n} = \frac{1}{n}\sum_{j=1}^{n} X_j,$$
$$Y_n = \frac{S_n - \mu n}{\sigma\sqrt{n}} = \frac{\overline{X_n} - \mu}{\frac{\sigma}{\sqrt{n}}}.$$

Considérons la variable centrée $X_n - \mu = X_1 - \mu$. On a

$$\varphi_{X_1-\mu}(t) = 1 - \frac{t^2\sigma^2}{2} + o(t^3).$$

L'espérance de $X_1 - \mu$ est nulle. La variance de X_1 est σ^2.
On a

$$Y_n = \frac{\sum_{j=1}^n (X_j - \mu)}{\sigma\sqrt{n}}.$$

La fonction caractéristique de cette variable vaut :

$$\varphi_{Y_n}(t) = \varphi_{X_1-\mu}\left(\frac{t}{\sigma\sqrt{n}}\right)^n = \left[1 - \frac{1}{2}\frac{\sigma^2 t^2}{\sigma^2 n} + o(t^2)\right] = \left[1 - \frac{1}{2}\frac{t^2}{n} + o(t^2)\right] \longrightarrow e^{-\frac{t^2}{2}}.$$

$\square$

CHAPITRE 6

VARIABLES GAUSSIENNES

6.1 Généralités

27-10-2023

Soit $X^T = (X_1, \dots, X_n)$ (transposée) un vecteur aléatoire, avec $X_1, \dots, X_n$ définies sur le même espace $(\Omega, \mathscr{A}, \mathbb{P})$.

Définition 6.1.1. On dira que le vecteur $X = (X_1, \dots, X_n)$ est gaussien (vecteur aléatoire gaussien, VAG) si toute combinaison linéaire des $X_i, i \in \{1, \dots, n\}$ est une variable gaussienne.

Donc en particulier, chaque X_i est une variable gaussienne.

Définition 6.1.2 (Rappel : variable gaussienne)**.** Une variable Y est dite gaussienne si c'est une variable à densité, et la densité a la forme suivante :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2},$$

où $m = \mathbb{E}[Y]$ et $\sigma^2 = \mathbb{V}(Y)$.

On utilisera souvent la gaussienne standart de loi $\mathscr{N}(0,1)$, dans ce cas on aura

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}.$$

L'objectif est de trouver la loi du vecteur, $P_{X_1,\dots,X_n}(x_1, \dots, x_n)$ et de savoir si elle est à densité. On veut aussi savoir s'il existe un critère d'indépendance des $\{X_i\}_{i=1}^n$. Enfin, on cherchera à calculer la fonction caractéristique de X.

6.2 Fonctions caractéristiques

Pour la gaussienne standart, la fonction caractéristique est

$$\varphi_Y(t) = e^{-\frac{1}{2}t^2}.$$

Dans le cas général, la fonction caractéristique se calcule comme suit :

$$\varphi_Y(t) = e^{itm} e^{-\frac{1}{2}\sigma^2 t^2}.$$

Proposition 6.2.1. La fonction caractéristique du vecteur gaussien X se calcule de la façon suivante :

$$\begin{aligned}\varphi_{X_1,\dots,X_n}(t_1,\dots,t_n) &= \int e^{i[t_1X_1+\cdots+t_nX_n]}d\mathbb{P} = \int e^{i(t,X)}d\mathbb{P} \\ &\overset{\text{transfert}}{=} \int e^{i(t,X)}dP_X(x_1,\dots,x_n) = \int e^{itY}d\mathbb{P} = \varphi_Y(1) = e^{im}e^{-\frac{1}{2}\sigma^2},\end{aligned}$$

où $Y = \sum_{i=1}^n t_iX_i$. C'est une variable gaussienne. Lorsqu'on introduit Y, on fait les calculs comme si c'était une variable gaussienne. On rappelle que $m = \mathbb{E}[Y], \sigma^2 = \mathbb{V}(Y)$.

Remarque. $(t_1,\dots,t_n)$ est le vecteur ligne tandis que $\begin{pmatrix}X_1\\ \vdots\\ X_n\end{pmatrix}$ est le vecteur colonne. (t,X) est le produit scalaire euclidien.

Remarque. L'espérance de Y est facile à calculer. En effet,

$$m = \mathbb{E}[Y] = \mathbb{E}\left[\sum_{j=1}^n t_jX_j\right] = \sum_{j=1}^n t_i\mathbb{E}[X_j] = (t, M_X),$$

avec $M_X = \begin{pmatrix}\mathbb{E}[X_1]\\ \vdots\\ \mathbb{E}[X_n]\end{pmatrix}$.

On a de plus

$$\sigma_Y^2\mathbb{V}(Y) = \mathbb{E}[(Y-\mathbb{E}[Y])^2] = \mathbb{E}\left[\sum_{j=1}^n \{t_jX_j - t_jm_{X_j}\}^2\right] = t\Sigma_Xt^T,$$

où Σ_X est la matrice de dispersion. On la calcule ainsi

$$\Sigma_X = \begin{pmatrix}\sigma_{X_1X_1} & \dots & \sigma_{X_1X_n}\\ \vdots & \ddots & \vdots\\ \sigma_{X_nX_1} & \dots & \sigma_{X_nX_n},\end{pmatrix}$$

où $\sigma_{X_iX_j} = \mathrm{Cov}(X_iX_j) = \mathbb{E}[X_iX_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$. Les éléments diagonaux de la matrice Σ_X sont en fait les variances de $X_1,\dots,X_n$. C'est une matrice **symétrique**.

On peut écrire de manière plus concise $m_Y = t\cdot M_X$ et $\sigma_Y^2 = t\Sigma_Xt^T$.

Le calcul de la fonction caractéristique devient donc

$$\varphi_{X_1,\dots,X_n}(t_1,\dots,t_n) = \varphi_Y(1) = e^{im_Y}e^{-\frac{1}{2}\sigma_Y^2} = e^{itm_X}e^{-\frac{1}{2}t\Sigma_Xt^2}.$$

On suppose que Σ_X est **inversible** (critère de non-dégénérescence). On sait que Σ_X est symétrie et positive, elle est donc diagonalisable, c'est-à-dire il existe une matrice A inversible telle que

$$A^{-1}\Sigma_XA = \mathrm{Id} = \begin{pmatrix}1 & \dots & 0\\ \vdots & \ddots & \vdots\\ 0 & \dots & 1\end{pmatrix} (?????)$$

On fait un changement de variable $Z = A^{-1}(X - m_X)^T = (z_1, \ldots, z_n)^T$. La matrice de dispersion de Z est donnée par

$$\Sigma_Z = A^{-1}\Sigma_X A = \mathrm{Id}\,.$$

La fonction caractéristique de Z est donc

$$\varphi_Z(t) = e^{-\frac{1}{2}t\Sigma_Z t^T} = e^{-\frac{1}{2}tt^T} = \prod_{i=1}^{n} e^{-\frac{1}{2}t_j^2} = \prod_{i=1}^{n} \varphi_{Z_i}(t).$$

Proposition 6.2.2. $Z_1, \ldots, Z_n$ sont indépendantes si et seulement si les fonctions caractéristiques se factorisent, autrement dit

$$\varphi_{Z_1 \ldots Z_n}(t_1, \ldots, t_n) = \varphi_{Z_1}(t_1) \ldots \varphi_{Z_n}(t_n).$$

Observons que l'on a les densités de $Z_1, \ldots, Z_n$, à savoir $f_Z(z_i) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}}$.

Donc

$$f_{Z_1 \ldots Z_n}(z_1, \ldots, z_n) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}} = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\frac{1}{2}zz^T},$$

avec $Z = (z_1, \ldots, z_n)$. On remplace Z avec X. La densité de X est égale à

$$f_{X_1 \ldots X_n}(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{\frac{n}{2}} \det(\Sigma_X)^{\frac{1}{2}}} e^{-\frac{1}{2}[(x - m_X)\Sigma_X^{-1}(x - m_X)^T]}.$$

Proposition 6.2.3. Un vecteur gaussien est indépendant si et seulement si il est non-corrélé.

DÉNOMBREMENT

On a 3 objets $\{a, b, c\}$. 15-09-2023

6.3 Dispositions sans répétition

Eléments	Combinaisons sans ordre	Combinaisons avec ordre
1 à 1	$\{a\}, \{b\}, \{c\}$	$\{a\}, \{b\}, \{c\}$
2 à 2	$\{a, b\}, \{a, c\}, \{b, c\}$	$\{a, b\}, \{b, a\}, \{a, c\}, \{c, a\}, \{b, c\}, \{c, b\}$
3 à 3	$\{a, b, c\}$	$\{a, b, c\}, \{a, c, b\}, \{b, a, c\}, \{c, a, b\}, \{c, b, a\}, \{b, c, a\}$

6.4 Dispositions avec répétition

Eléments	Dispositions sans ordre	Dispositions avec ordre
1 à 1	$\{a\}, \{b\}, \{c\}$	$\{a\}, \{b\}, \{c\}$
2 à 2	$\{a, a\}, \{b, a\}, \{b, b\}, \{a, c\}, \{b, c\}, \{c, c\}$	$\{a, a\}, \{a, b\}, \{b, a\}, \{b, b\}, \{a, c\}, \{c, a\}, \{b, c\}, \{c, b\}\{c, c\}$
3 à 3	$\{a, a, a\}, \{a, a, c\}, \dots$ (10 éléments)	$\{a, a, a\}, \{a, a, c\}, \{a, c, a\}, \{c, a, a\}, \dots$ ($3^3 = 27$ éléments)

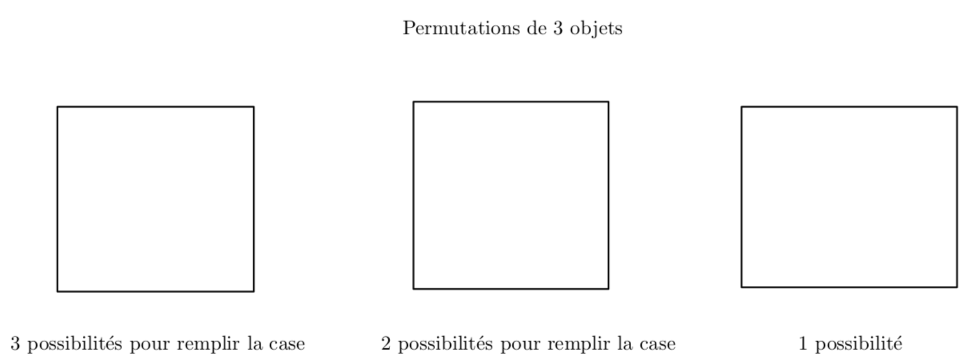


Figure 6.1 – Permutations dans le cas de 3 objets (dispositions sans répétition et avec ordre)

Arrangements de n objets pris k à k (dispositions avec ordre et sans répétition) Si on a 4 objets pris 2 à 2 : $\{a, b, c, d\}$.

Si a et b sont fixés, alors $\{a, b\}$ engendrera $\{a, b, c, d\}$ et $\{a, b, d, c\}$.

On a

$$A_n^k = \frac{n!}{(n-k)!}$$

Combinaisons de n objets pris k à k (dispositions sans ordre et sans répétition)

$$C_n^k = \frac{n!}{k!(n-k)!} = \binom{n}{k}. \tag{6.1}$$

En fait,

$$C_n^k = \frac{A_n^k}{k!},$$

avec $k!$ le nombre de permutations des k éléments.
Si on a n objets à combiner k à k avec répétition, mais sans ordre, il y a

$$C_k^{n+1-k} = \frac{(n+1-k)!}{k!(n-1)!} \text{ combinaisons possibles.} \tag{6.2}$$

6.5 Tirage des urnes sans remise

N boules de type N_a, N_b tels que $N_a + N_b = N$.
On tire $n < N$ boules.
Soit E l'événement suivant : $\{k$ boules parmi n sont de type $a\}$.
Calculer la probabilité de E.

$$\mathbb{P}(E) = \frac{C_{N_a}^k C_{N_b}^{n-k}}{C_N^n} \text{ (formule hypergéométrique).}$$

Démonstration. $\Omega =$ combinaisons de N objets pris n à n sans les répéter.
On a $\sharp(\Omega) = C_N^n$.
Cas favorables : $C_{N_a}^k C_{N_b}^{n-k}$. □

Cette formule est utilisée pour calculer la probabilité de gagner au loto. On a une grille de 49 numéros et on tire 6 numéros.
$N_a =$ les 6 numéros cochés par le joueur, $N_b = 49 - 6 = 43$ et $n = N_a$.

$$\mathbb{P}(\text{avoir 3 numéros gagnants}) = \frac{C_6^3 C_{43}^3}{C_{49}^6} \approx 0.018$$
$$\mathbb{P}(\text{avoir 6 numéros gagnants}) = \frac{C_6^6 C_{43}^0}{C_{49}^6} \approx 7,15 \times 10^{-8}.$$

6.6 Tirage des urnes avec remise

N boules, N_a de type a, N_b de type b. On en tire n (avec n quelconque) et

$$E = \{k \text{ boules parmi les } n \text{ tirées sont de type } a\}.$$

On a $\sharp(\Omega) = N^n$.
Cas favorables : $N_a^k N_b^{n-k} \binom{n}{k}$.
Donc

$$\mathbb{P}(E) = \frac{N_a^k N_b^{n-k} \binom{n}{k}}{N^n} = \frac{N_a^k N_b^{n-k} \binom{n}{k}}{N^k N^{n-k}} = \left(\frac{N_a^k}{N}\right)^k \left(\frac{N_b}{N}\right)^{n-k} \binom{n}{k} = p_a^k p_b^{n-k} \binom{n}{k},$$

où p_a et p_b sont les pourcentages de a et de b.
Il s'agit de la **loi binomiale**.

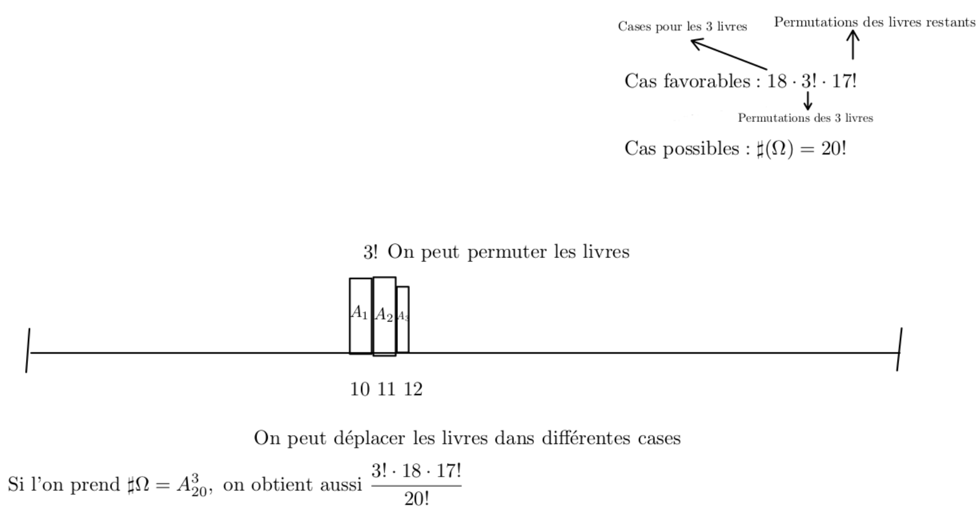
TRAVAUX DIRIGÉS

Dans ce document, vous ne trouverez que les énoncés des exercices, les corrections étant manuscrites. La numérotation dans les feuilles manuscrites correspond à celle dans ce document.

Exercice 4. Dans une bibliothèque, il y a n livres sur une étagère repartis au hasard. Parmi ces n livres, k sont d'un même auteur A, les autres d'auteurs différents. Calculer la probabilité qu'au moins p livres de A se retrouvent côte à côte dans les cas suivants :

1. $n = 20, k = 3, p = 3$;
2. $n = 20, k = 5, p = 2$ (**au moins** 2 livres).

Démonstration.



1. ing

Figure 6.2 – Solution pour (1)

□

Exercice 5. On lance 10 fois une pièce de monnaie. Calculer la probabilité qu'au cinquième lancer on obtient pile en sachant que le nombre total des piles obtenus est 3.

Exercice 6. On a deux variables aléatoires X_1 et X_2 de même densité $f(x) = \frac{1}{X^2}\mathbb{1}_{[1,\infty)}(x)$ et elles sont indépendantes. On pose

$$U = X_1 X_2$$
$$V = \frac{X_1}{X_2}.$$

1. Calculer la loi du couple (U, V).
2. U et V sont-elles indépendantes ?

Exercice 7. On place six boules de manière aléatoire et indépendante dans 3 boîtes. Calculer la probabilité que la première boîte contienne deux boules.

Exercice 8. Un robinet a été installé le jour J et il fuit, les fuites se produisent chaque heure de manière indépendante et avec la probabilité p.

1. Calculer la loi de la variable aléatoire F égale au nombre de fuites qui se sont produites tout au long de la journée (en 24 heures), ie $\mathbb{P}(F = k), k \in \{0, \ldots, 24\}$.
2. On dénote $T(1)$ la variable aléatoire égale à l'heure de la première fuite. Calculer la loi de $T(1)$, ie $\mathbb{P}(T(1) = k), k = 1, \ldots, 24$.
3. On dénote $T(2)$ la variable aléatoire égale à l'heure de la deuxième fuite. Montrer que $T(1)$ et $T(2) - T(1)$ sont indépendantes et ont la même loi.

Exercice 9. Sur l'espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$, on considère le vecteur aléatoire (X, Y) à valeurs dans $\mathbb{R}^2$ et de loi P_X à densité

$$f_{XY}(x, y) = \alpha(1 - x^2)\mathbb{1}_{[0,1]}(x) y e^{-3y} \mathbb{1}_{[0,\infty)}(y),$$

où $\alpha \in \mathbb{R}$.

1. Déterminer α.
2. Déterminer les lois marginales.
3. Calculer $\mathbb{P}(0 < X \leq 2, Y \geq 1)$.
4. Calculer la matrice de dispersion D de (X, Y) :

$$D = \begin{pmatrix} \mathbb{V}(X) & \mathrm{Cov}(X, Y) \\ \mathrm{Cov}(Y, X) & \mathbb{V}(Y) \end{pmatrix}.$$

Exercice 10. Soient X et Y deux variables aléatoires indépendantes et de même loi $e^{-x}\mathbb{1}_{[0,\infty)}(x)$. Déterminer la loi conjointe des nouvelles variables $U = X + Y$ et $V = \dfrac{X}{X+Y}$ et montrer que V a une loi uniforme sur $[0, 1]$.

Exercice 11. Soient X et Y deux variables aléatoires dont Y est de densité $f_Y(y)$ et posons $Z = X + Y$.

1. Prenons X et Y indépendantes. Montrer que Z est à densité indépendante de la nature de X. *On peut essayer avec X qui suit une loi binomiale.*
2. Soit maintenant X une variable gaussienne $\mathscr{N}(0, 1)$ et $Y = XB$ où $\mathbb{P}(B = 1) = \mathbb{P}(X = -1) = \frac{1}{2}$. Montrer que X et Y sont dépendantes et que $Z = X + Y$ n'est pas à densité.

Exercice 12. Deux variables aléatoires indépendantes avec densités :

$$f_X(x) = \frac{1}{\pi} \frac{1}{\sqrt{1 - x^2}}, |x| < 1,$$

$$f_Y(y) = \frac{y}{\sigma^2} e^{-\frac{y}{2\sigma^2}}.$$

Soit $Z = X \cdot Y$. Montrer que Z a la loi $\mathscr{N}(0, \sigma^2)$, avec

$$F_{\mathscr{N}}(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}}.$$

Exercice 13. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables indépendantes de loi

$$\mathbb{P}(X_n = 0) = 1 - \frac{1}{n},$$

$$\mathbb{P}(X_n = 1) = \frac{1}{n}.$$

1. Montrer que la suite $\{X_n\}$ converge en probabilité vers 0.
2. Montrer que la suite $\{X_n\}$ diverge presque sûrement.
   Idée : utiliser le deuxième Borel-Cantelli 4.3.
3. On a $\{X_n\}_{n=1}^{\infty}$ de variables aléatoires indépendantes telles que

$$\mathbb{P}(X_n = 0) = 1 - \frac{1}{n^2}\ ;\ \mathbb{P}(X_n = n^2) = \frac{1}{n^2}.$$

   Montrer que X_n converge presque sûrement vers 0.
   Montrer que $\{X_n\}$ ne converge pas vers 0 dans L^1.

Exercice 14.
Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires indépendantes de loi exponentielle $\mathscr{E}(2) = 2e^{-2x}$. Montrer que la suite

$$\frac{1}{\ln(n)} \max_{1\leq k\leq n} X_k$$

converge en probabilités vers $\frac{1}{2}$.

Exercice 15. Soit $\{X_n\}_{n=1}^{\infty}$ une suite de variables aléatoires de carré intégrable et de même espérance μ. On pose $S_n = \displaystyle\sum_{k=1}^{n} X_k$ et on suppose que

$$\frac{1}{n^\beta}\mathbb{V}(S_n) \underset{n\to\infty}{\longrightarrow} 0 \text{ avec } \beta > 0.$$

Pour quelles valeurs de β la suite $\frac{S_n}{n} \longrightarrow \mu$ en norme $L^2(\mathbb{P})$ (et donc en probabilité) ?

Exercice 16. On lance plusieurs fois et de manière indépendante deux dés. Soit X le nombre de lancers de premier dé nécessaires pour obtenir 1 et Y la variable aléatoire qui donne le nombre de lancers du deuxième dé nécessaires pour obtenir 5 ou 6.

1. Calculer les lois de X et Y.
2. On pose $Z = \max(X, Y)$. Déterminer $F_Z(n) = \mathbb{P}(Z \leq n)$.
3. Calculer $\mathbb{P}(Z = n)$.
   Idée : exprimer $\mathbb{P}(Z = n)$ en terme de $F_Z(n)$ et $F_Z(n-1)$.
4. Calculer la loi de $W = \min(X, Y)$.
5. Calculer $\mathbb{P}(X \geq Y)$.
   Idée : l'événement $X \geq Y$ est donnée par la réunion des événements $x \geq y$ pris sur les valeurs y de Y.

Exercice 17. Soit $f_n, n \geq 1$ la fonction définie par

$$f_n(x) = \mathbb{1}_{\mathbb{R}^+}(x) n^2 e^{-\frac{n^2x^2}{2}}.$$

1. Montrer que f_n est une densité de probabilité.
2. Montrer que la suite de variables aléatoires X_n qui est de densité f_n converge en probabilités vers une variable aléatoire à déterminer.

Exercice 18. Soit X une variable aléatoire de Gauss et soit U une variable aléatoire de Bernoulli avec $\mathbb{P}(U = 1) = \mathbb{P}(U = -1) = \dfrac{1}{2}$, avec U et X indépendantes.

1. Montrer que $Y = UX$ a la même loi que X.
2. X et Y ne sont pas corrélées.
3. X et Y ne sont pas indépendantes.
4. Trouver cette combinaison linéaire.

Examen de l'année dernière

Exercice 19. Soit $\{X_n\}_{n\geq 1}$ une suite de variables aléatoires indépendantes et de même loi exponentielle $\lambda e^{-\lambda x}\mathbb{1}_{[0,\infty)}, \lambda > 0$. On pose

$$Y_n = \frac{1}{n}\sum_{i=1}^{n}\frac{1+X_i}{2}.$$

1. Calculer $\mathbb{E}[Y_n]$ et $\mathbb{V}(Y_n)$.
2. Montrer que Y_n converge en probabilité vers $\mathbb{E}[Y_n]$.
3. Calculer la fonction caractéristique de Y_n et montrer qu'elle donne un résultat cohérent avec le point 2, c'est-à-dire calculer la limite de la fonction caractéristique pour $n \longrightarrow \infty$.

Exercice 20. Soit $\{U_n\}_{n=1}^{\infty}$ une suite de variables aléatoires uniformes sur $[0,1]$ et indépendantes. Soit ensuite $\{X_n\}_{n=1}^{\infty}$ une autre suite de variables aléatoires indepéndantes des U_n de loi $B(n, \frac{1}{2})$. Considérons maintenant la suite $Z_n = \min\limits_{0\leq i\leq X_n} U_i$.

Vers quelle variable aléatoire en loi la suite des Z_n converge ?

Pour arriver à cela, on doit répondre aux questions suivantes :

1. Calculer $\mathbb{P}(\min\limits_{0\leq i\leq k} U_i \leq t)$.
2. Etudier $\mathbb{P}(Z_n \leq t)$ en conditionnant par rapport à X_n.
3. Trouver la limite en loi des Z_n.

Exercice 21. Considérons la suite $\{X_n\}_{n=1}^{\infty}$ chacune de densité

$$f_{X_n}(x) = \begin{cases} \frac{n}{x^{n+1}}, x \geq 1 \\ 0 \text{ sinon.} \end{cases}$$

Posons $Y_n = (n-1)X_n - n$.

1. Etudier la convergence en loi de Y_n (utiliser les fonctions de répartition).
2. Etudier la convergence avec la fonction caractéristique.