Igor Kortchemski

Cours 1 : Indépendance

⇝ c'est avec l'indépendance que la théorie des probabilités se distingue de la théorie de la mesure d'un point de vue "idée"

Plan : 1) Rappels (théorie mesure)
2) Indépendance (fini d'événements)
3) Indépendance (variables aléatoires, tribus)
4) Indépendance et intégration

1) Rappels de théorie de la mesure

$(\Omega, \mathcal{A}, \mathbb{P})$ espace de probabilité ($\mathcal{A}$: tribu sur Ω, $\mathbb{P}$ mesure de probabilité), sur lequel seront définies toutes les variables aléatoires

- Si $X : (\Omega, \mathcal{A}) \to (E, \mathcal{E})$ est une variable aléatoire (= formulation probabiliste de "fonction mesurable", on note souvent v.a.), sa loi $\mathbb{P}_X$ est une mesure de probabilité sur $(E, \mathcal{E})$ définie par

$$\mathbb{P}_X(A) = \mathbb{P}(X^{-1}(A))$$

$$= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\})$$

$$\boxed{= \mathbb{P}(X \in A)}$$

notation probabiliste

→ on essaye d'écrire le moins possible "ω"

pour $A \in \mathcal{E}$.

<u>théorie de la mesure</u> : $\mathbb{P}_X$ est la mesure image de $\mathbb{P}$ par X.

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

est la <u>tribu engendrée par X</u>
(plus petite tribu sur Ω
qui rend X mesurable)

<u>Intuition</u> : $\sigma(X)$ est
"l'information" obtenue en
observant X.

- Lorsque $E = (E, d)$ est un
espace métrique, on prendra
la tribu borélienne $\mathcal{E} = \mathcal{B}(E)$
($\boxed{\text{engendrée}}$ par les ouverts
ou fermés)

Si $\mathcal{R}$ est un ensemble de
sous-ensembles de E,
la tribu engendrée par $\mathcal{R}$

est par définition

$$\sigma(\mathcal{A}) = \bigcap_{\substack{\mathcal{F} \text{ tribu} \\ \mathcal{A} \subset \mathcal{F}}} \mathcal{F}$$

c'est la plus petite tribu contenant tous les éléments de $\mathcal{A}$.

<u>Exemple</u> $E = \{1, 2, 3\}$

$\mathcal{A} = \{ \{1\} \}$

Alors $\sigma(\mathcal{A}) = \{ \emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\} \}$

Caractérisation de lois

Comment montrer que

$X : \Omega \to (E, \mathcal{E})$

$Y : \Omega \to (E, \mathcal{E})$

ont même loi ?

Idée 1 : utiliser la définition, montrer que $\forall A \in \mathcal{E}$,

$\mathbb{P}_X(A) = \mathbb{P}_Y(A)$

c'est-à-dire

$$\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$$

PROBLÈME : souvent $\mathcal{E}$ est très gros et on ne connaît une "bonne" description des éléments de $\mathcal{E}$ (exemple : $\mathcal{B}(\mathbb{R})$)

<u>Idée 2</u> : montrer que

$$\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$$

pour tout $A \in \mathcal{C}$ avec $\mathcal{C}$ un

π-système générateur de $\mathcal{E}$.

stable par intersections finies

$\sigma(\mathcal{C}) = \mathcal{E}$

<u>Exemple</u> $\mathcal{C} = \{]a,b] : a < b \}$

$\mathcal{E} = \mathcal{B}(\mathbb{R})$.

AVANTAGE En pratique, $\mathcal{C}$ est bien connu

(idée de la preuve : lemme des classes monotones)

Exemple $\mathcal{C} = \{]-\infty, a] : a \in \mathbb{R} \}$

$\mathcal{E} = \mathcal{B}(\mathbb{R})$.

Conséquence : si $X, Y : \Omega \to \mathbb{R}$ v.a. telles que

$\mathbb{P}(X \leq x) = \mathbb{P}(Y \leq x)$

$\forall x \in \mathbb{R}$, alors

X et Y ont même loi.

(la fonction de répartition d'une v.a. réelle caractérise sa loi)

Passons à une caractérisation fonctionnelle des lois.

Ingrédient :

Théorème de transfert

Soit $X : \Omega \to E$ une v.a.

$f : E \to \mathbb{R}_+$ mesurable

Alors

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

Pourquoi "transfert" ?

Rappelons que par

définition, $\mathbb{E}$ est une intégrale sur Ω par rapport à $\mathbb{P}$.

$\rightsquigarrow$ le théorème "change" une intégrale sur Ω en une intégrale sur E

Ce résultat mène à :

Principe de la fonction muette

Soit $X, Y : \Omega \to E$ deux v.a

① On suppose $\forall f : E \to \mathbb{R}_+$ mesurable $\mathbb{E}[f(X)] = \mathbb{E}[f(Y)]$

Alors X et Y ont même loi (notation : $X \overset{(d)}{=} Y$)

② On suppose (E,d) espace métrique et que $\forall f : E \to \mathbb{R}_+$ continue bornée

$$\mathbb{E}[f(X)] = \mathbb{E}[f(Y)].$$

Alors $X \overset{(d)}{=} Y$.

Preuve :

① Soit $A \in \mathcal{E}$. On prend $f = \mathbb{1}_A$. Alors

$$\mathbb{E}[\mathbb{1}_A(X)] = \mathbb{E}[\mathbb{1}_A(Y)]$$

$$\mathbb{P}(X \in A) \qquad \mathbb{P}(Y \in A)$$

D'où le résultat.

② <u>Problème</u> : $\mathbb{1}_A$ n'est pas une fonction continue.

<u>Idée</u> : argument d'approximation

Soit $F \in \mathcal{E}$ fermé.

On montre d'abord que

$$\mathbb{P}_X(F) \overset{(*)}{=} \mathbb{P}_Y(F)$$

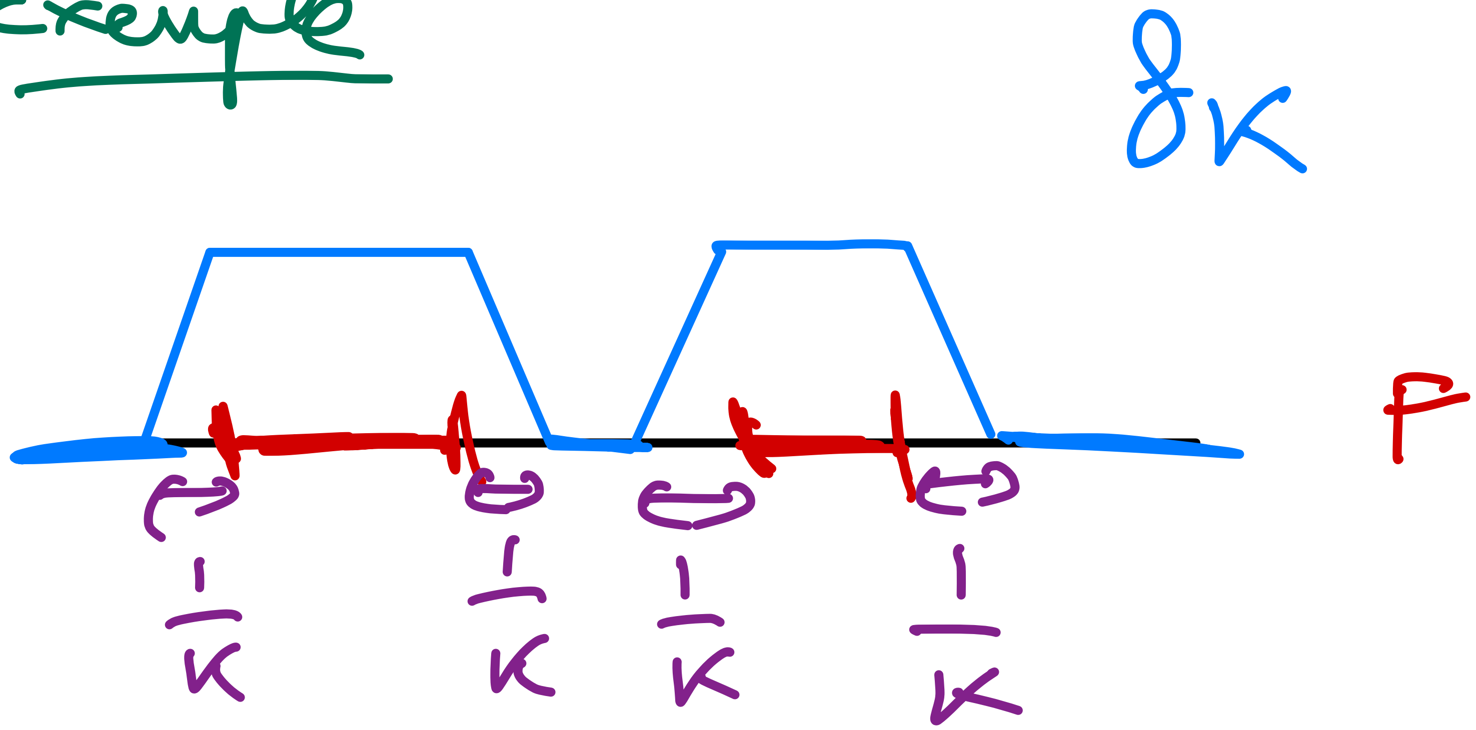
et on aura le résultat voulu d'après ce qu'on a vu avant (les fermés sont un π-système générateur de la tribu borélienne)

Pour $K > 0$, on pose :

$$f_K(x) = \max(1 - K\,d(x, F), 0)$$

continue bornée

Exemple



Alors

- $g_K \xrightarrow[K\to\infty]{\text{simplement}} \mathbb{1}_F$

  (car F fermé)

- $|g_K| \leq 1$

Donc par convergence dominée,

$$\begin{array}{ccc} \mathbb{E}[g_K(X)] & = & \mathbb{E}[g_K(Y)] \\ \downarrow K\to\infty & & \downarrow K\to\infty \\ \mathbb{E}[\mathbb{1}_F(X)] & & \mathbb{E}[\mathbb{1}_F(Y)] \end{array}$$

Donc $\mathbb{P}(X \in F) = \mathbb{P}(Y \in F)$

$\sim$

Remarque Intérêt de ②?
Pourquoi prendre des fonctions continues bornées ? Plusieurs raisons :

- régularité peut être utile (par exemple pour convergence dominée)
- la notion de convergence en loi, vue plus tard, utilise les fonctions continues bornées.

Espaces et mesure produits

(produit fini)

- Si $(E_i, \mathcal{E}_i)_{1 \leq i \leq n}$ sont des espaces mesurables, la tribu produit sur $E_1 \times \cdots \times E_n$ est

$$\mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_n = \sigma(\underbrace{A_1 \times \cdots \times A_n}_{\text{pavé élémentaire.}} : A_i \in \mathcal{A}_i)$$

(un π-système générateur)

• Si P_i est une mesure de probabilité sur $(E_i, \mathcal{E}_i)$, la mesure produit $P_1 \otimes \cdots \otimes P_n$

sur $(E_1 \times \dots \times E_n, \mathcal{E}_1 \otimes \dots \otimes \mathcal{E}_n)$ est l'unique mesure de probabilité telle que

$$\mathbb{P}_1 \otimes \dots \otimes \mathbb{P}_n (A_1 \times \dots \times A_n) = \mathbb{P}_1(A_1) \times \dots \times \mathbb{P}_n(A_n)$$

(<u>unicité</u> provient de ce qu'on a dit avant, <u>existence</u> plus compliqué)

- Si $(X_1, \dots, X_n)$ est une v.a. à valeurs dans $E_1 \times \dots \times E_n$, sa loi $\mathbb{P}_{(X_1, \dots, X_n)}$ est donc une mesure de probabilité sur $(E_1 \times \cdots \times E_n, \mathcal{E}_1 \otimes \cdots \otimes \mathcal{E}_n)$, donnée par

$$\mathbb{P}_{(X_1, \dots, X_n)}(A_1 \times \cdots \times A_n)$$

$$= \mathbb{P}((X_1, \dots, X_n) \in A_1 \times \cdots \times A_n)$$

$$= \mathbb{P}(X_1 \in A_1, X_2 \in A_2, \dots, X_n \in A_n).$$

2) Indépendance d'un nombre fini d'événements

$(\Omega, \mathcal{A}, \mathbb{P})$: espace de proba.

$A, B \in \mathcal{A}$ sont dits <u>indépendants</u> si

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

Interprétation : Si $\mathbb{P}(B)>0$, cela revient à dire que

$$\mathbb{P}(A|B) \overset{\text{def}}{=} \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}$$

(proba conditionnelle)

est égal à $\mathbb{P}(A)$.

Intuitivement, le fait de savoir que B est réalisé ne donne pas d'information sur la probabilité de réalisation ou non de A.

Examples (1) Lancer de deux dés : $\Omega = \{1,..,6\}^2$

$$\mathbb{P}(\{\omega\}) = \frac{1}{36} \text{ pour } \omega \in \Omega$$

Alors $A = \{6\} \times \{1, \dots, 6\}$

$= \{(6,1), (6,2), \dots, (6,6)\}$

et $B = \{1, \dots, 6\} \times \{6\}$

$= \{(1,6), \dots, (6,6)\}$

sont indépendants car

$\mathbb{P}(A)\,\mathbb{P}(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

$\mathbb{P}(A \cap B) = \frac{1}{36}$

(2) Lancer d'un dé

$\Omega = \{1, 2, \dots, 6\}$

$\mathbb{P}(\{\omega\}) = \frac{1}{6}$ pour $\omega \in \Omega$.

$A = \{1, 2\}$

$B = \{1, 3, 5\}$

sont indépendants

car $\mathbb{P}(A) = \frac{1}{3}$

$\mathbb{P}(B) = \frac{1}{2}$

$A \cap B = \{1\}$

$\mathbb{P}(A \cap B) = \frac{1}{6}$

Définition On dit que n événements $A_1, \dots, A_n$ sont **indépendants** si pour tout sous-ensemble non vide $\{j_1, \dots, j_p\}$ de $\{1, \dots, n\}$ on a

$$\mathbb{P}(A_{j_1} \cap \dots \cap A_{j_p}) = \mathbb{P}(A_{j_1}) \cdots \mathbb{P}(A_{j_p})$$

(ou encore

$\forall J \subset \{1, \dots, n\}$, $\# J \geq 1$,

$$\mathbb{P}\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} \mathbb{P}(A_j)$$

Notation $(A_i)_{1 \leq i \leq n}$ sont

$\perp\!\!\!\perp$

⚠ ⚠ ⚠ Pour l'indépendance il ne suffit pas juste d'avoir

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \dots \mathbb{P}(A_n)$$

ni d'avoir

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j)$$

pour $i \neq j$

(indépendance 2 à 2, plus faible)

indépendance = indépendance mutuelle

Proposition Des évènements $A_1, \dots, A_n$ sont $\perp\!\!\!\perp$ ssi
$$P(B_1 \cap \dots \cap B_n) = P(B_1) \cdots P(B_n) \quad (*)$$
pour tous
$$B_i \in \sigma(\{A_i\}) = \{\emptyset, A_i, A_i^c, \Omega\}$$
pour $1 \leq i \leq n$.

Preuve :

$\boxed{\Leftarrow}$ Si $\{j_1, \dots, j_p\} \subset \{1, \dots, n\}$

Alors $P(A_{j_1} \cap \dots \cap A_{j_p}) = P(A_{j_1}) \cdots P(A_{j_p})$

simplement en prenant
$$B_i = A_i \text{ si } i \in \{j_1, \dots, j_p\}$$
$$B_i = \Omega \text{ sinon}$$
dans $(*)$.

$\boxed{\Rightarrow}$ On suppose $A_1, \dots, A_n$ ⫫

<u>But</u> : montrer $(*)$.

On peut supposer $B_i \neq \emptyset$ $\forall 1 \leq i \leq n$

(sinon $(*)$ donne $0 = 0$)

Si $\{j_1, \dots, j_p\} = \{i : B_i \neq \Omega\}$

il s'agit de montrer que

$$\mathbb{P}(B_{j_1} \cap \dots \cap B_{j_p}) = \mathbb{P}(B_{j_1}) \cdots \mathbb{P}(B_{j_p})$$

dès que $B_{j_k} = A_{j_k}$ ou $A_{j_k}^c$.

Ainsi, il suffit de montrer

que si $C_1, \dots, C_p$ sont ⫫

alors $C_1^c, C_2, \dots, C_p$ sont ⫫

Pour cela, pour

$\{i_1, \dots, i_q\} \subset \{2, \dots, p\}$

on écrit :

$$\mathbb{P}(C_1^c \cap C_{i_1} \cap \dots \cap C_{i_q})$$

$$= \mathbb{P}(C_{i_1} \cap \dots \cap C_{i_q}) - \mathbb{P}(C_1 \cap \dots \cap C_{i_q})$$

$$= \mathbb{P}(C_{i_1}) \dots \mathbb{P}(C_{i_q}) - \mathbb{P}(C_1) \dots \mathbb{P}(C_{i_q})$$

par $\perp\!\!\!\perp$

$$= \mathbb{P}(C_1^c)\, \mathbb{P}(C_{i_1}) \dots \mathbb{P}(C_{i_q})$$

(on utilise $\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$)

Ce résultat amène naturellement à la notion de <u>tribus</u> $\perp\!\!\!\perp$ qui est <u>le</u> bon cadre pour l'indépendance.

3) Indépendance d'un nombre fini de tribus et de v.a.

Définitions

- Soit $\mathcal{B}_1, \ldots, \mathcal{B}_n \subset \mathcal{A}$ des sous tribus de $\mathcal{A}$. Elles sont indépendantes (notation $\perp\!\!\!\perp$) si :

$$\forall A_1 \in \mathcal{B}_1, \ldots, \forall A_n \in \mathcal{B}_n,$$
$$\mathbb{P}(A_1 \cap \ldots \cap A_n) = \mathbb{P}(A_1) \ldots \mathbb{P}(A_n)$$

- Soit $X_1, \ldots, X_n$ des v.a. à valeurs dans $(E_1, \mathcal{E}_1), \ldots, (E_n, \mathcal{E}_n)$. Elles sont indépendantes si les tribus $\sigma(X_1), \ldots, \sigma(X_n)$ sont $\perp\!\!\!\perp$

Remarques

- Par définition $X_1, \dots, X_n$ sont $\perp\!\!\!\perp$ ssi $\forall A_1 \in \mathcal{E}_1, \dots, \forall A_n \in \mathcal{E}_n$,

$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \mathbb{P}(X_1 \in A_1) \cdots \mathbb{P}(X_n \in A_n)$$

(en effet $\sigma(X_i) = \{X_i^{-1}(A_i) : A_i \in \mathcal{E}_i\}$

- Les événements $A_1, \dots, A_n$ sont $\perp\!\!\!\perp$ ssi les tribus $\sigma(\{A_1\}), \dots, \sigma(\{A_n\})$ sont $\perp\!\!\!\perp$ d'après la proposition précédente.

- ⚠ La définition d'$\perp\!\!\!\perp$ dépend de $\mathbb{P}$: si on change $\mathbb{P}$ on peut perdre l'$\perp\!\!\!\perp$.

• Pour parler d'$\perp\!\!\!\perp$ de v.a. elles doivent être définies sur le même espace de probabilité.

Lemme (principe de composition)

Soit $X_1, \dots, X_n$ des v.a à valeurs dans $(E_1, \mathcal{E}_1), \dots, (E_n, \mathcal{E}_n)$ $\perp\!\!\!\perp$. Soit $f_i : (E_i, \mathcal{E}_i) \to (F_i, \mathcal{F}_i)$ mesurables. Alors

$f_1(X_1), \dots, f_n(X_n)$

sont $\perp\!\!\!\perp$

Preuve Pour $B_i \in \mathcal{F}_i$, on a

$(f_i(X_i))^{-1}(B_i) = X_i^{-1}(\underbrace{f_i^{-1}(B_i)}_{\in \mathcal{E}_i})$

car f_i mesurable.

Donc $\sigma(f_i(X_i)) \subset \sigma(X_i)$

Comme $\sigma(X_1), \dots, \sigma(X_n)$ sont $\perp\!\!\!\perp$

on en déduit

$\sigma(f_1(X_1)), \dots, \sigma(f_n(X_n))$

sont $\perp\!\!\!\perp$.

<u>Exemple</u> : Si X, Y, Z sont des v.a. réelles $\perp\!\!\!\perp$, alors

$X^2, \dfrac{Y^2}{1+Y^2}, \sin(Z)$

sont $\perp\!\!\!\perp$.

Remarque

Soit $X = (X_1, \dots, X_n)$ une v.a. dans $E_1 \times \dots \times E_n$.

La loi $\mathbb{P}_X$ est appelée <u>loi jointe</u> de $(X_1, \dots, X_n)$.

La loi $\mathbb{P}_{X_i}$ de X_i est appelée <u>loi marginale</u> de X_i

Puisque

$$\mathbb{P}_{(X_1,\dots,X_n)}\left(E_1 \times \dots \times E_{i-1} \times A_i \times E_{i+1} \times \dots E_n\right)$$

$$= \mathbb{P}(X_i \in A_i) = \mathbb{P}_{X_i}(A_i),$$

les lois marginales sont déterminées par la loi jointe.

Réciproque <u>FAUSSE</u> en général

les lois marginales, en général, ne permettent pas de déterminer la la loi jointe.

Il se trouve que c'est vrai pour des v.a. $\perp\!\!\!\perp$!

Proposition Soit $X_1, \dots, X_n$ des v.a. à valeurs dans $(E_1, \mathcal{E}_1), \dots, (E_n, \mathcal{E}_n)$.
Elles sont $\perp\!\!\!\perp$ ssi

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

(lien $\perp\!\!\!\perp$ $\leftrightarrow$ mesures produit)

Preuve $\boxed{\Leftarrow}$ Pour $B_i \in \mathcal{E}_i$,

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n)$$

$$\left[= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) \in B_1, \dots, X_n(\omega) \in B_n\})\right]$$

$$= \mathbb{P}_{(X_1, \dots, X_n)}(B_1 \times \dots \times B_n)$$

$= \mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n}(B_1 \times \cdots \times B_n)$

$= \mathbb{P}_{X_1}(B_1) \cdots \mathbb{P}_{X_n}(B_n)$

$= \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_n \in B_n)$

$\Rightarrow$ D'après le même calcul,

$\mathbb{P}_{(X_1, \dots, X_n)}$ et $\mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n}$ coincident sur

$\{ B_1 \times \cdots \times B_n : B_i \in \mathcal{E}_i \}$,

qui forme un π-système générateur de la tribu produit

On conclut que

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

4) Indépendance et intégration

Il existe une caractérisation fonctionnelle très utile de l'$\perp\!\!\!\perp$

Théorème Soit $X_1, \dots, X_n$ des v.a. dans $(E_1, \mathcal{E}_1), \dots, (E_n, \mathcal{E}_n)$. Elles sont $\perp\!\!\!\perp$ ssi $\forall f_i : (E_i, \mathcal{E}_i) \to \mathbb{R}_+$ mesurables on a

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

Preuve

$\boxed{\Leftarrow}$ Pour $B_1 \in \mathcal{E}_1, \dots, B_n \in \mathcal{E}_n$

On prend $f_i = \mathbb{1}_{B_i}$

Alors

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

$$= \mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n)$$

et $\prod_{i=1}^{n} \mathbb{E}[\mathbb{1}_{B_i}(X_i)] = \prod_{i=1}^{n} \mathbb{P}(X_i \in B_i)$

$\boxed{\Rightarrow}$ On applique le théorème de transfert sur $E_1 \times \cdots \times E_n$ avec $\mathbb{P}_{(X_1, \ldots, X_n)}$ et $(x_1, \ldots, x_n) \mapsto \prod_{i=1}^{n} f_i(x_i)$:

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

$$\int_{E_1 \times \cdots \times E_n} f_1(x_1) \ldots f_n(x_n) \, \mathbb{P}_{(X_1, \ldots, X_n)}(dx_1 \ldots dx_n)$$

$$= \int_{E_1 \times \dots \times E_n} f_1(x_1) \cdots f_n(x_n) \, \mathbb{P}_{X_1}(dx_1) \otimes \cdots \otimes \mathbb{P}_{X_n}(dx_n)$$

$\boxed{\text{(Fubini)}}$

$$= \prod_{i=1}^{n} \int_{E_i} f_i(x_i) \, \mathbb{P}_{X_i}(dx_i)$$

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

Remarque si les fonction f_i sont de signe quelconque, l'égalité

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

$(*)$ $(**)$

reste vraie à condition

que $\mathbb{E}[|f_i(X_i)|] < \infty$

$\forall 1 \leq i \leq n$

(on utilise le théorème de transfert, version intégrable)

On a alors que si

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

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

ce qui justifie l'existence de $(\circledast)$ et $(\circledast\circledast)$

- En particulier si $X_1, \dots, X_n$ sont $\in L^1$ et $\perp\!\!\!\perp$, alors

$X_1 \cdot \ldots \cdot X_n \in L^1$ et

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

⚠ ⚠ ⚠ ⚠

En général $X, Y \in L^1$

$$\not\Rightarrow \quad XY \in L^1$$

<u>Rappel</u> : $X : \Omega \to \mathbb{R}$ est dans $L^p = L^p(\Omega, \mathcal{A}, \mathbb{P})$ si $\mathbb{E}[|X|^p] < \infty$.

<u>Corollaire</u> Si X, Y sont deux v.a. $\perp\!\!\!\perp$ dans L^2, alors

$$\mathrm{Cov}(X, Y) := \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$$

<u>Preuve</u> on a

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$$

car $X, Y \perp\!\!\!\perp$ et $X \in L^2 \Rightarrow X \in L^1$

∞

⚠ ⚠ Réciproque fausse

En effet soit X une v.a. de loi $\mathcal{N}(0,1)$, de densité

$$\frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Soit ε tq $\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = \frac{1}{2}$

avec $X \perp\!\!\!\perp \varepsilon$.

On pose $Y = \varepsilon X$.

Alors:

$Cov(X,Y)$
$= \mathbb{E}[\varepsilon X^2] - \underbrace{\mathbb{E}[\varepsilon X]\,\mathbb{E}[X]}_{=0}$
$= \mathbb{E}[\varepsilon]\,\mathbb{E}[X^2]$
$= 0$

Mais X et Y ne sont pas $\perp\!\!\!\perp$

En effet, par l'absurde on suppose $X \perp\!\!\!\perp Y$.

Alors par principe de composition, $|X| \perp\!\!\!\perp |Y|$

Or $|Y| = |X|$

Donc $|X| \perp\!\!\!\perp |X|$. Donc

$$\mathbb{P}(|X| \leq 1) = \mathbb{P}(|X| \leq 1, |X| \leq 1) \overset{\perp\!\!\!\perp}{=} \mathbb{P}(|X| \leq 1)^2$$

Donc $\mathbb{P}(|X| \leq 1) = 0$ ou 1.

<u>Absurde</u>.

Petit exercice : X v.a. réelle

Montrer que $X \perp\!\!\!\perp X$ ssi X est constante