RAPPELS DE PROBABILITÉ

Espace de probabilité

Espace de probabilité est un triplet $\{\Omega, F, P\}$:

- Ω est un ensemble.

- F tribu et (ou σ algebre) sur Ω.

- P est une mesure de probabilité sur (Ω, F)

Définition

Une tribu sur Ω est une famille F de sous ensembles se Ω appelées evenement tq :

1- $\emptyset$, $\Omega \in F$.

2- $\forall A \in F, A^c \in F$.

3- $(A_n)_{n=1}^{\infty} \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in F$.

Définition

Soit $A = \{A_i, i \in I\}$ une famille de sous ensemble de Ω alors la tribu engendré par A est la plus petite tribu sur Ω qui contient tout les ensembles $A_i, i \in I$ elle est noté $\sigma(A)$. I n'est pas forcement denombrable.

Exemple

$\sigma(\{1\}) = \{\emptyset, \Omega, \{1\}, \{2, 3, 4, 5, 6\}\}$

Définition

Soit $\Omega = [0,1]$ la tribu borélienne sur $[0,1]$ est la tribu engendré par la famille $A = \{]a,b[, 0 \leq a \leq b \leq 1\}$:intervalles ouvert dans $[0,1]$ elle est noté par $B([0,1])$

Définition

Soit F une tribu sur Ω une mesure de probabilité sur (Ω, F) est une application $P : F \mapsto [0,1]$ telle que :$P(\emptyset) = 0$, $P(\Omega) = 1$.
$(A_n)_{n=1}^{\infty}$ disjointe (i.e : $A_n \bigcap A_m = \emptyset$, $\forall n \neq m$ $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$) en particulier : $A, B \in F$, et $A \bigcap B = \emptyset \Rightarrow P(A \bigcup B) = P(A) + P(B)$.
De plus :
1- Si $(A_n)_{n=1}^{\infty} \subset F$,$A_n \subset A_{n+1} \subset, \ldots, \bigcup_{n=1}^{\infty} A_n = A$ alors : $\lim_n P(A_n) = P(A)$.
2- Si $(A_n)_{n=1}^{\infty} \subset F$, $A_n \supset A_{n+1} \supset \ldots, \bigcap_{n=1}^{\infty} A_n = A$ alors : $\lim_n P(A_n) = P(A)$.

Définition

Soit $\Omega = [0,1]$ et $F = B([0,1])$ on appelle mesure de Lebsegue sur $[0,1]$ la mesure de probabilité définit par :

$$P(]0,1[) = b - a \ \forall 0 \leq a \leq b \leq 1$$

B n'est pas définit a priori que sur les intervales, mais est un uniquement extensible à tout ensemble borélien.
$B \in ([0,1])$ elle est notée $P(B) =\mid B \mid$ en utilisant la propriété (2) si dessus en déduit pour tout x, $x \in [0,1]$ $|\{x\}| = \lim_n \left|]x - \frac{1}{n}, x + \frac{1}{n}[\right| = \lim_n \frac{2}{n} = 0$

Généralisation à n dimension :

Soit $\Omega = [0,1]$ tribu borélienne $B(\Omega)$ où

$$\begin{aligned} A &= \{]a_1, b_1[\times]a_2, b_2[\times]a_3, b_3[, \ldots,]a_n, b_n[\} \\ &= (b_1 - a_1)(b_2 - a_2)(b_3 - a_3) \ldots (b_n - a_n) \end{aligned}$$

Notion de variable aléatoire

Définition

Soit (E, B) un espace mesurable : tout application mesurable :

$$X(\Omega, A) \mapsto (E, B)$$

est une variable aléatoire $\forall B \in B : X^{-1}(B) \in A$

Exemple

$\Omega = \{(1,1), (1,2), (1,3), \ldots, (6,6)\}$ $X(i,j) = i+j, E = \{2,3,4,5,6,7,8,9,10,11,12\}, X^{-1}(\{10\}) = \{(5,5), (4,6), (6,4)\}$

Définition

Soit $(\Omega, F, \mathbb{R})$ un espace de probabilité.

un variable aléatoire est une application : $X : \Omega \mapsto \mathbb{R}$

telle que : $\{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\} \in F, \forall B \in B(\mathbb{R})$

Les lois classiques directes

— La loi uniforme :
  E un ensemble fini, card E=n alors X sur la loi uniforme sur E si $P(X = x) = \frac{1}{n}$

— Loi de bernoulli :
  de paramêtre $p \in [0,1]$ la loi d'une variable aléatoire à valeur dans $[0,1]$ telle que : $P(X = 1) = p$, $P(X = 0) = 1 - p$, on interprête X comme le résultat de lance d'une piece de monaie qui trouve sur pile avec la probabilité P.

— La loi Binomial :
  $B(n,p), n \in \mathbb{N}, p \in [0,1]$, c'est la loi d'une variable aléatoire à valeur dans $\{1, \ldots, n\}$ tel que : $P(X = K) = C_n^k p^k (1-p)^{n-k}$ on interprète X comme le nombre de face obtenue rn n lancer.

— La loi géométrique de paramètre $p \in]0,1[$ est la loi d'une variable aléatoire X à valeur dans $\mathbb{N}$ telle que : $P(X = K) = (1-p)p^K$

  Définition

X est le nombre de piles obtenu premier face.

— La loi de poisson de paramètre $\lambda > 0$, X est à valeur dans $\mathbb{N}$

$$P(X = K) = \frac{\lambda^K}{K!}, \quad K \in \mathbb{N}$$

Les lois classiques continues

— La loi uniforme sur $[a, b]$, $a < b$

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

— La loi exponontielle de parametre $\lambda > 0$

$$f(x) = \lambda e^{-\lambda x}.\mathbb{1}_{\mathbb{R}^n}(x)$$

— La loi gaussien ou normal $\mathcal{N}(m, \sigma^2)$

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

la loi normale Poisson sont les lois plus importantes dans la théorie de probabilité.

Espérance variable aléatoire :

Construction de l'espérance (l'inégalité de Lebesgue) on prosède à 3 étapes : - Soit $X(\omega) = \sum_{i=0}^{\infty} x_i.\mathbb{1}_{A_i}(\omega)$, $x_i \geq 0$, $A_i \in F$, on définit l'espérance de telles variable aléatoire (diter sipmle) comme suit : $E(X) = \sum_{i=0}^{\infty} x_i P(A_i) \in [0, +\infty[$.
Attention!! $\mathbb{E}(X)$ peut prendre la valeur $+\infty$.

Terminologie :

1. Si $\mathbb{E}(X) = 0$ alors on dit que X est une variable aléatoire centré.
2. Si $\mathbb{E}(X) < \infty$ alors on dit que X est une variable intégrable.

3. Si $\mathbb{E}(X^2) < \infty$ on dit que X est une variable aléatoire de carré intégrable.
   On dit que X est une variable aléatoire bornée $\exists$ une $K > 0$ telle que $|X(\omega)| \leq K$, $\forall \omega \in \Omega$

Remarque

X borné $\Rightarrow \mathbb{E}(X^2) < \infty \Rightarrow \mathbb{E}(|X|)$.

Proposition

Soient X une variable aléatoire et $g : \mathbb{R} \to \mathbb{R}$ une fonction borélienne tel que $\mathbb{E}(|g(x)|) < \infty$ alors :

- Si X est une variable aléatoire discrète (a valeur dans D dénombrable), alors : $\mathbb{E}(g(x)) = \sum_{n \in D} g(x) P(X = x)$
- Si X est une variable aléatoire continue (avec densité f_x) alors :$\mathbb{E}(g(x)) = \int_{\mathbb{R}} g(x).f_x(x)dx$.
  Ceci s'applique en particulier si $g(x) = x$.
  Variance variable aléatoire : Soient X,Y deux variable aléatoire de carré intégrable, on pose :

$$Var(X) = \mathbb{E}((X - \mathbb{E}(X)^2)) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geq 0$$

$$Cov(X, Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Terminologie :

- Un évènement $A \in F$ est dit négligeable si $P(A) = 0$
- Un évènement $A \in \mathcal{F}$ est dit presque sûr (souvent abrégé p.s) si $P(A) = 1$, i.e., si A est négligeable.

Proposition

Si $(A_n)_{n=1}^{\infty} \subset \mathcal{F}$ est une famille d'événements négligeables (i.e. $P(A_n) = 0, \forall n$), alors $\bigcup_{n=1}^{\infty} A_n$ est négligeable.

Preuve

$P\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} P(A_n) \leq 0.$

Propriétés de $\mathbb{E}(X)$

Soient X, Y deux variables aléatoires intégrables.
- **Linéarité** : $\mathbb{E}(cX + Y) = c\mathbb{E}(X) + \mathbb{E}(Y)$, $c \in \mathbb{R}$. X et Y variables aléatoires intégrables.
- **Positivité** : Si $X \geq 0$ p.s alors $\mathbb{E}(X) \geq 0$.
- **Positivité stricte** : Si $X \geq 0$ p.s et $\mathbb{E}(X) = 0$ alors $X = 0$ p.s.
- **Monotonie** : Si $X \geq Y$ p.s alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$.

Inégalités importantes

Inégalité de Cauchy-Schwarz

Soient X, Y deux variables aléatoires de carré intégrable. Alors :

(i) XY est intégrable,

(ii) $\left(\mathbb{E}(|XY|)\right)^2 \leq \mathbb{E}(X^2)\mathbb{E}(Y^2)$.

En posant $Y = 1$, on trouve que $\left(\mathbb{E}(|X|)\right)^2 \leq \mathbb{E}(X^2)$ (donc $\mathbb{E}(|X|) < 1$ si $\mathbb{E}(X^2) < 1$.

Inégalité triangulaire

Soient X, Y deux variables aléatoires intégrables. Alors :

$$\mathbb{E}(|X + Y|) \leq \mathbb{E}(|X|) + \mathbb{E}(|Y|).$$

Inégalité de Jensen

Soient X une variable aléatoire et $\phi : \mathbb{R} \to \mathbb{R}$ une fonction borélienne et convexe telle que $\mathbb{E}(|\phi(X)|) < \infty$. Alors :

$$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X)).$$

En particulier, $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$.

Démonstration

Vu que ϕ est convexe, on a :

$$\phi(x) = \sup_{a,b: ax+b \leq \phi(y), \forall y \in \mathbb{R}} (ax+b),$$

et donc :

$$\phi(\mathbb{E}(X)) = \sup_{a,b:\cdots}(a\mathbb{E}(X)+b) = \sup_{a,b:\cdots} \mathbb{E}(aX+b) \leq \sup_{a,b:\cdots} \mathbb{E}(\phi(X)) = \mathbb{E}(\phi(X)).$$

Inégalité de Markov

Soient X une variable aléatoire et $\psi : \mathbb{R} \to \mathbb{R}_+$ une fonction borélienne et croissante sur $\mathbb{R}_+$, avec $\psi(a) > 0$ pour tout $a > 0$ et $\mathbb{E}(\psi(X)) < \infty$. Alors :

$$P(\{X \geq a\}) \leq \frac{\mathbb{E}(\psi(X))}{\psi(a)}, \quad \forall a > 0.$$

Démonstration

Du fait que ψ est croissante sur $\mathbb{R}_+$, on a :

$$\mathbb{E}(\psi(X)) \geq \mathbb{E}(\psi(X)\mathbf{1}_{\{X \geq a\}}) \geq \mathbb{E}(\psi(a)\mathbf{1}_{\{X \geq a\}}) = \psi(a)\mathbb{E}(\mathbf{1}_{\{X \geq a\}}) = \psi(a)P(\{X \geq a\}).$$

Comme $\psi(a) > 0$, ceci permet de conclure.

Conditionnement sur un évènement

Définition

Soient A et B deux évènements de $\mathcal{F}$ (i.e. $A, B \in \mathcal{F}$). Alors, la probabilité conditionnelle de A sachant que B est donnée par :

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

pour tout B tel que $P(B) \neq 0$.

Propriétés

$P(\cdot \mid B)$ est une nouvelle probabilité sur l'espace $(\Omega, \mathcal{F})$.

Preuve

1. On a :
$$P(\Omega \mid B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1.$$

2. Si $A_1, A_2, \ldots, A_n$ sont n-événements deux à deux disjoints ($A_i \cap A_j = \emptyset$ pour $i \neq j$), alors :
$$P\left(\bigcup_{i=1}^{n} A_i \mid B\right) = \frac{P\left(\left(\bigcup_{i=1}^{n} A_i\right) \cap B\right)}{P(B)}.$$
Or,
$$\bigcup_{i=1}^{n} (A_i \cap B) = \left(\bigcup_{i=1}^{n} A_i\right) \cap B,$$
et comme les événements A_i ($i = 1, \ldots, n$) sont deux à deux disjoints, les $A_i \cap B$ ($i = 1, \ldots, n$) sont également deux à deux disjoints. D?après la définition d?une probabilité, on en déduit que :
$$P\left(\bigcup_{i=1}^{n} A_i \mid B\right) = \frac{\sum_{i=1}^{n} P(A_i \cap B)}{P(B)}.$$
Ainsi :
$$P\left(\bigcup_{i=1}^{n} A_i \mid B\right) = \sum_{i=1}^{n} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{n} P(A_i \mid B).$$

Définition

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et soit X une variable aléatoire définie sur cet espace.

Considérons le cas de X à valeurs dans $\{x_1, x_2, \ldots, x_n\}$. Soit B un événement de $(\Omega, \mathcal{F}, P)$ fixé et $Q(A) := P(A \mid B)$.

L'espérance de X par rapport à Q est donnée par :

$$\mathbb{E}_Q(X) = \sum_{j=1}^{n} x_j Q(X = x_j) = \sum_{j=1}^{n} x_j \frac{P(X = x_j \cap B)}{P(B)} = \frac{1}{P(B)} \sum_{j=1}^{n} x_j P(X = x_j \cap B).$$

On sait que :

$$P(A) = \int_A dP, \quad \forall A \in \mathcal{F},$$

et

$$P(\{X = x_j\} \cap B) = \int_{(X=x_j)\cap B} dP = \int_B 1_{(X=x_j)}(\omega)\, dP, \quad (1)$$

où

$$1_{(X=x_j)}(\omega) = \begin{cases} 1 & \text{si } \omega \in \{X = x_j\}, \\ 0 & \text{si } \omega \notin \{X = x_j\}. \end{cases}$$

On peut écrire :

$$X(\omega) = \sum_{j=1}^{n} x_j 1_{(X=x_j)}(\omega), \quad (2)$$

et d'après (1) et (2), on a :

$$\mathbb{E}_Q(X) = \frac{1}{P(B)} \sum_{j=1}^{n} x_j P(X = x_j \cap B) = \frac{1}{P(B)} \sum_{j=1}^{n} x_j \int_B 1_{(X=x_j)}\, dP = \frac{1}{P(B)} \int_B \sum_{j=1}^{n} x_j 1_{(X=x_j)}\, dP = \frac{1}{P(B)} \int_B X\, dP.$$

Cela implique que :

$$\mathbb{E}_Q(X) P(B) = \int_B X\, dP,$$

ce qui implique :

$$\int_B \mathbb{E}_Q(X)\, dP = \int_B X\, dP,$$

où B est un événement fixé.

Ainsi, on note :

$$\mathbb{E}_Q(X) = \mathbb{E}(X \mid B).$$

Espérance conditionnelle par rapport à une tribu

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et soit X une variable aléatoire définie sur cet espace. Soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$.

Définition

On appelle l'espérance conditionnelle de la variable aléatoire X sachant $\mathcal{G}$, et on la note $\mathbb{E}(X \mid \mathcal{G})$, l'unique variable aléatoire telle que :

1. $\mathbb{E}(X \mid \mathcal{G})$ est $\mathcal{G}$-mesurable.
2. Pour tout $A \in \mathcal{G}$, on a :
$$\int_A \mathbb{E}(X \mid \mathcal{G})\, dP = \int_A X\, dP.$$

C'est l'unique variable $\mathcal{G}$-mesurable (à égalité P-presque sûre près) telle que :

$$\mathbb{E}[\mathbb{E}(X \mid \mathcal{G})Y] = \mathbb{E}(XY),$$

pour toute variable Y, $\mathcal{G}$-mesurable et bornée.

Propriétés de l'espérance conditionnelle

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité donné, et soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$. Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathcal{F}, P)$.

1. Soit a et b deux constantes telles que :
$$\mathbb{E}(aX + bY \mid \mathcal{G}) = a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}) \quad \text{(Linéarité)}.$$
2. Soit X et Y deux variables aléatoires telles que $X \leq Y$, alors :
$$\mathbb{E}(X \mid \mathcal{G}) \leq \mathbb{E}(Y \mid \mathcal{G}) \quad \text{(Croissance)}.$$
3. Si X est $\mathcal{G}$-mesurable, alors :
$$\mathbb{E}(X \mid \mathcal{G}) = X.$$

4. Si Y est $\mathcal{G}$-mesurable, alors :

$$\mathbb{E}(XY \mid \mathcal{G}) = Y\mathbb{E}(X \mid \mathcal{G}).$$

5.

$$\mathbb{E}[\mathbb{E}(X \mid \mathcal{G})] = \mathbb{E}(X).$$

6. Si X est indépendante de $\mathcal{G}$, alors :

$$\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(X).$$

7. Si X est une variable aléatoire telle que $X \in L^p(\Omega, \mathcal{F}, P)$, avec $p \geq 1$, alors :

$$||\mathbb{E}(X \mid \mathcal{G})||_{L^p(\Omega,\mathcal{F},P)} \leq ||X||_{L^p(\Omega,\mathcal{F},P)}.$$

8. Si $\mathcal{G}$ et $\mathcal{H}$ sont deux tribus telles que $\mathcal{H} \subseteq \mathcal{G}$, alors :

$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{H}) \mid \mathcal{G}) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \mid \mathcal{H}) = \mathbb{E}(X \mid \mathcal{H}).$$

9. Si φ est une application convexe et mesurable, alors :

$$\mathbb{E}[\varphi(X) \mid \mathcal{G}] \leq \varphi(\mathbb{E}(X \mid \mathcal{G})) \quad \text{(Inégalité de Jensen).}$$

Preuve

Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathcal{F}, P)$ et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$.

1) Soient a et b deux constantes, alors on a

$$\mathbb{E}(aX + bY \mid \mathcal{G}) = a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}).$$

On a $\mathbb{E}(X \mid \mathcal{G}) < \infty$ et $\mathbb{E}(Y \mid \mathcal{G}) < \infty$, tel que :

$$\int_A \mathbb{E}(aX + bY \mid \mathcal{G})\, dP = \int_A (aX + bY)\, dP, \quad \forall A \in \mathcal{G}$$

$$= a \int_A X\, dP + b \int_A Y\, dP$$

$$= a \int_A \mathbb{E}(X \mid \mathcal{G})\, dP + b \int_A \mathbb{E}(Y \mid \mathcal{G})\, dP$$

$$= \int_A (a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}))\ dP.$$

On déduit que :

$$\int_A [\mathbb{E}(aX + bY \mid \mathcal{G}) - (a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}))]\, dP = 0, \quad \forall A \in \mathcal{G}.$$

On a $\mathbb{E}(aX + bY \mid \mathcal{G})$ est une variable aléatoire $\mathcal{G}$-mesurable, et $a \in \mathbb{R}$ et $\mathbb{E}(X \mid \mathcal{G})$ v.a. $\mathcal{G}$-mesurables, donc $b\mathbb{E}(Y \mid \mathcal{G})$ v.a. $\mathcal{G}$-mesurable. Ce qui implique $\mathbb{E}(aX + bY \mid \mathcal{G}) - (a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}))$ est une v.a. $\mathcal{G}$-mesurable. Donc, on obtient :

$$\mathbb{E}(aX + bY \mid \mathcal{G}) - (a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G})), = 0P\text{-presque sûr.}$$

$$\mathbb{E}(aX + bY \mid \mathcal{G}) = a\mathbb{E}(X \mid \mathcal{G}) + b\mathbb{E}(Y \mid \mathcal{G}), \quad P\text{-presque sûr.}$$

Remarque Si α_i sont des constantes et $(X_i)_i$ sont des variables aléatoires, alors on a

$$\mathbb{E}\left[\sum_{i=0}^{n} \alpha_i X_i \mid \mathcal{G}\right] = \sum_{i=0}^{n} \alpha_i \mathbb{E}(X_i \mid \mathcal{G}).$$

2) Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathcal{F}, P)$. Si $X \leq Y$, alors $\mathbb{E}(X \mid \mathcal{G}) \leq \mathbb{E}(Y \mid \mathcal{G})$.

On a $\mathbb{E}(X \mid \mathcal{G}) < +\infty$ et $\mathbb{E}(Y \mid \mathcal{G}) < +\infty$. D'après $X \leq Y$, on a

$$\int_A X\, dP \leq \int_A Y\, dP, \quad \forall A \in \mathcal{G}.$$

Donc, pour $A \in \mathcal{G}$,

$$\int_A \mathbb{E}(X \mid \mathcal{G})\, dP = \int_A X\, dP \leq \int_A Y\, dP = \int_A \mathbb{E}(Y \mid \mathcal{G})\, dP.$$

Cela implique que

$$\int_A [\mathbb{E}(Y \mid \mathcal{G}) - \mathbb{E}(X \mid \mathcal{G})]\, dP \geq 0, \quad \forall A \in \mathcal{G}.$$

Ainsi, $\mathbb{E}(X \mid \mathcal{G}), \mathbb{E}(Y \mid \mathcal{G})$ est une variable aléatoire $\mathcal{G}$-mesurable.Donc $\mathbb{E}(Y \mid \mathcal{G}) - \mathbb{E}(X \mid \mathcal{G})$ une v.a.$\mathcal{G}$-mesurable.
On déduit que

$$\mathbb{E}(Y \mid \mathcal{G}) - \mathbb{E}(X \mid \mathcal{G}) \geq 0, \quad P\text{-presque sûr.}$$

D'où

$$\mathbb{E}(Y \mid \mathcal{G}) \geq \mathbb{E}(X \mid \mathcal{G}), \quad P\text{-presque sûr.}$$

3) Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$.
Si X est $\mathcal{G}$-mesurable, alors $\mathbb{E}(X \mid \mathcal{G}) = X$.

On sait que $\mathbb{E}(X \mid \mathcal{G})$ est une variable aléatoire $\mathcal{G}$-mesurable telle que :

$$\int_A \mathbb{E}(X \mid \mathcal{G})\, dP = \int_A X\, dP, \quad \forall A \in \mathcal{G}.$$

Cela implique que

$$\int_A [\mathbb{E}(X \mid \mathcal{G}) - X]\, dP = 0, \quad \forall A \in \mathcal{G}.$$

On sait que $\mathbb{E}(X \mid \mathcal{G})$: une v.a. $\mathcal{G}$-mesurable et X une v.a. $\mathcal{G}$-mesurable (par hypothèse) implique :

$$\mathbb{E}(X \mid \mathcal{G}) - X : \text{une v.a. } \mathcal{G}\text{-mesurable.}$$

Ce qui implique :

$$\mathbb{E}(X \mid \mathcal{G}) - X = 0, \quad P\text{-p.s.}$$

Donc :

$$\mathbb{E}(X \mid \mathcal{G}) = X, \quad P\text{-p.s.}$$

D'où, $\mathbb{E}(X \mid \mathcal{G}) = X, \quad P$-p.s.

4) Soient X et Y deux variables aléatoires définies sur $(\Omega, \mathcal{F}, P)$ et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$. Si Y est $\mathcal{G}$-mesurable, alors $\mathbb{E}(XY \mid \mathcal{G}) = Y\mathbb{E}(X \mid \mathcal{G})$ presque sûrement.

L'objectif est de démontrer que :

$$\int_A \mathbb{E}(XY \mid \mathcal{G})\, dP = \int_A Y\mathbb{E}(X \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

Cela revient à prouver que :

$$\int_A (\mathbb{E}(XY \mid \mathcal{G}) - Y\mathbb{E}(X \mid \mathcal{G}))\, dP = 0, \quad \forall A \in \mathcal{G}.$$

On a $\mathbb{E}(XY \mid \mathcal{G})$ est une variable aléatoire $\mathcal{G}$-mesurable par définition, et Y est $\mathcal{G}$-mesurable par hypothèse. De plus, $\mathbb{E}(X \mid \mathcal{G})$ est $\mathcal{G}$-mesurable. Ainsi, $Y\mathbb{E}(X \mid \mathcal{G})$ est également $\mathcal{G}$-mesurable.

Cela implique que $\mathbb{E}(XY \mid \mathcal{G}) - Y\mathbb{E}(X \mid \mathcal{G})$ est une variable aléatoire $\mathcal{G}$-mesurable.

Étape 1 : On pose $Y = 1_C$, où C est un ensemble mesurable par $\mathcal{G}$, et on remarque que 1_C est aussi $\mathcal{G}$-mesurable. Par un simple calcul, on trouve :

$$\int_A 1_C \mathbb{E}(X \mid \mathcal{G})\, dP = \int_{A\cap C} \mathbb{E}(X \mid \mathcal{G})\, dP = \int_{A\cap C} X\, dP, \quad \forall A \in \mathcal{G}, \forall C \in \mathcal{G} \Rightarrow A \cap C \in \mathcal{G}.$$

Donc

$$\int_A 1_C \mathbb{E}(X \mid \mathcal{G})\, dP = \int_A 1_C X\, dP, \quad \forall A \in \mathcal{G}.$$

On conclut que

$$\int_A 1_C \mathbb{E}(X \mid \mathcal{G})\, dP = \int_A \mathbb{E}(1_C X \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

Ce qui implique

$$\int_A \mathbb{E}(XY \mid \mathcal{G})\, dP = \int_A Y\mathbb{E}(X \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

D'où :

$$1_C \mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(1_C X \mid \mathcal{G}), \quad P\text{-presque sûr.}$$

Donc, $Y\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(YX \mid \mathcal{G})$, $\quad P$-presque sûr.

Étape 2 : On pose que la variable aléatoire Y est écrit sous forme d'une fonction étagée :

$$Y = \sum_{i=1}^{n} \alpha_i 1_{C_i},$$

tels que C_i est $\mathcal{G}$-mesurable (i.e., Y est une variable aléatoire étagée).

Alors, on obtient par un calcul simple :

$$\int_A Y\mathbb{E}(X \mid \mathcal{G})\, dP = \int_A \left(\sum_{i=1}^{n} \alpha_i 1_{C_i} \right) \mathbb{E}(X \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

Cela donne :

$$= \sum_{i=1}^{n} \alpha_i \int_A 1_{C_i} \mathbb{E}(X \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

Puis, en utilisant le calcul précédent, on obtient :

$$= \sum_{i=1}^{n} \alpha_i \int_A 1_{C_i} X\, dP, \quad \forall A \in \mathcal{G}.$$

Ce qui revient à :

$$= \int_A \left(\sum_{i=1}^{n} \alpha_i 1_{C_i} \right) X\, dP, \quad \forall A \in \mathcal{G}.$$

Ainsi :

$$= \int_A YX\, dP, \quad \forall A \in \mathcal{G}.$$

Cela montre que :

$$\int_A Y\mathbb{E}(X \mid \mathcal{G})\, dP = \int_A \mathbb{E}(YX \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

On en déduit que :

$$\int_A [Y\mathbb{E}(X \mid \mathcal{G}) - \mathbb{E}(YX \mid \mathcal{G})]\, dP = 0, \quad \forall A \in \mathcal{G}.$$

Cela implique que :

$$Y\mathbb{E}(X \mid \mathcal{G}) - \mathbb{E}(YX \mid \mathcal{G}) = 0, \quad P\text{-presque sûr.}$$

D'où : $Y\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(YX \mid \mathcal{G})$, P-presque sûr.

Étape 3 : Si Y est une variable aléatoire positive $\mathcal{G}$-mesurable, alors il existe une suite de variables aléatoires étagées $(Y_n)_{n\geq 1}$ croissante et positive telle que :

$$\lim_{n\to+\infty} Y_n = Y.$$

D'après l'Étape 2, on a :

$$\int_A Y_n\mathbb{E}(X \mid \mathcal{G})\, dP = \int_A Y_n X\, dP, \quad \forall A \in \mathcal{G}.(*)$$

Cas 1 : Si X est une variable aléatoire positive telle que $Y_n \to Y$ quand $n \to +\infty$ et $(XY_n)_{n\geq 1} \to XY_n$ quand $n \to +\infty$, avec $(Y_n)_{n\geq 1}$ croissante et $X \geq 0$, alors $(XY_n)_{n\geq 1}$ est une suite croissante, et on a :

$$Y_n \leq Y_{n+1} \quad \text{implique} \quad XY_n \leq XY_{n+1}, \quad \forall n \in \mathbb{N} \text{ tel que } X \geq 0.$$

D'après le lemme de Beppo Lévy et $(*)$, on a :

$$\int_A XY_n\, dP \to \int_A XY\, dP, \quad \forall A \in \mathcal{G} \text{ quand } n \to +\infty,$$

$$\int_A Y_n \mathbb{E}(X \mid \mathcal{G})\, dP \to \int_A Y\mathbb{E}(X \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G} \text{ quand } n \to +\infty.$$

Alors,

$$\lim_{n\to+\infty} \int_A Y_n \mathbb{E}(X \mid \mathcal{G})\, dP = \lim_{n\to+\infty} \int_A Y_n X\, dP, \quad \forall A \in \mathcal{G}.$$

Ce qui implique :

$$\int_A Y\mathbb{E}(X \mid \mathcal{G})\, dP = \int_A YX\, dP, \quad \forall A \in \mathcal{G}.$$

C'est-à-dire :

$$= \int_A \mathbb{E}(XY \mid \mathcal{G})\, dP, \quad \forall A \in \mathcal{G}.$$

On conclut que :

$$\int_A [Y\mathbb{E}(X \mid \mathcal{G}) - \mathbb{E}(XY \mid \mathcal{G})]\, dP = 0, \quad \forall A \in \mathcal{G}.$$

On sait que Y une v.a. $\mathcal{G}$-mesurable et $\mathbb{E}(X \mid \mathcal{G})$ est v.a. $\mathcal{G}$-mesurable (par définition), donc $Y\mathbb{E}(X \mid \mathcal{G})$ est v.a. $\mathcal{G}$-mesurable et $\mathbb{E}(XY \mid \mathcal{G})$ est v.a. $\mathcal{G}$-mesurable.

Ainsi, $[Y\mathbb{E}(X \mid \mathcal{G}) - \mathbb{E}(XY \mid \mathcal{G})]$ est $\mathcal{G}$-mesurable.

Implique :

$$Y\mathbb{E}(X \mid \mathcal{G}) - \mathbb{E}(XY \mid \mathcal{G}) = 0 \quad P\text{-presque sûr.}$$

D'où :

$$Y\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(XY \mid \mathcal{G}) \quad P\text{-presque sûr.}$$

Rappel (Lemme de Beppo-Lévy) : Soit $(Y_n)_{n\geq 0}$ une suite de variables aléatoires monotones, et si

$$Y_n \xrightarrow{\text{p.s.}} Y,$$

alors,

$$\lim_{n\to+\infty} \mathbb{E}(Y_n) = \mathbb{E}(Y),$$

c'est-à-dire,

$$\lim_{n\to+\infty} \int_\Omega Y_n\, dP = \int_\Omega Y\, dP.$$

Cas 2 : Soit X une variable aléatoire quelconque.

On a : $X = X^+ - X^-$, où X^+ et X^- sont deux variables aléatoires positives. D'après le Cas 1, on a :

$$Y\mathbb{E}(X^+ \mid \mathcal{G}) = \mathbb{E}(X^+Y \mid \mathcal{G}),$$
$$Y\mathbb{E}(X^- \mid \mathcal{G}) = \mathbb{E}(X^-Y \mid \mathcal{G}),$$

ce qui implique :

$$Y\Big[\mathbb{E}(X^+ \mid \mathcal{G}) - \mathbb{E}(X^- \mid \mathcal{G})\Big] = \mathbb{E}(X^+Y - X^-Y \mid \mathcal{G}),$$

et ceci donne également :

$$Y\mathbb{E}(X^+ - X^- \mid \mathcal{G}) = \mathbb{E}(Y(X^+ - X^-) \mid \mathcal{G}).$$

D'où :

$$Y\mathbb{E}(X^\pm \mid \mathcal{G}) = \mathbb{E}(Y(X^\pm) \mid \mathcal{G}).$$

i.e

$$Y\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(XY \mid \mathcal{G}).$$

Étape 4 : Si Y est une variable aléatoire intégrable.

D'après l'Étape 3, on a :

$$Y \equiv Y^\pm = Y^+ - Y^-,$$

où Y^+ et Y^- sont deux variables aléatoires positives.

On a :

$$\mathbb{E}(XY^+ \mid \mathcal{G}) = Y^+\mathbb{E}(X \mid \mathcal{G}),$$
$$\mathbb{E}(XY^- \mid \mathcal{G}) = Y^-\mathbb{E}(X \mid \mathcal{G}).$$
$$\Rightarrow \mathbb{E}(XY^+ \mid \mathcal{G}) - \mathbb{E}(XY^- \mid \mathcal{G}) = Y^+\mathbb{E}(X \mid \mathcal{G}) - Y^-\mathbb{E}(X \mid \mathcal{G}).$$

D'après la linéarité de l'espérance conditionnelle, on a :

$$\mathbb{E}(X(Y^+ - Y^-) \mid \mathcal{G}) = (Y^+ - Y^-)\mathbb{E}(X \mid \mathcal{G}).$$

D'où :

$$\mathbb{E}(XY^{\pm} - XY^{-} \mid \mathcal{G})Y^{\pm}\mathbb{E}(X \mid \mathcal{G}),$$

et donc :

$$\mathbb{E}(XY \mid \mathcal{G}) = Y\mathbb{E}(X \mid \mathcal{G}).$$

5) Soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$ et X une variable aléatoire définie sur $(\Omega, \mathcal{F}, P)$. Montrons que $\mathbb{E}(\mathbb{E}(X \mid \mathcal{G})) = \mathbb{E}(X)$.

On sait que par définition, $\mathbb{E}(X \mid \mathcal{G})$ est une variable aléatoire $\mathcal{G}$-mesurable telle que :

$$\int_A \mathbb{E}(X \mid \mathcal{G})\, dP = \int_A X\, dP, \quad \forall A \in \mathcal{G}.$$

On pose $A = \Omega$, alors :

$$\int_\Omega \mathbb{E}(X \mid \mathcal{G})\, dP = \int_\Omega X\, dP.$$

D'où :

$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{G})) = \mathbb{E}(X).$$

6) Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité, X une variable aléatoire définie sur cet espace, et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$. Montrons que si X est indépendante de $\mathcal{G}$, alors $\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(X)$.

On sait que $\mathbb{E}(X)$ est une constante, et pour tout $A \in \mathcal{G}$, on a :

$$\int_A \mathbb{E}(X)\, dP = \mathbb{E}(X) \int_A dP, \quad \forall A \in \mathcal{G},$$

et par définition, on a :

$$\mathbb{E}(X) = \int_\Omega X\, dP.$$

Alors,

$$\left(\int_\Omega X\, dP\right)\left(\int_A 1\, dP\right); \quad \forall A \in \mathcal{G}, \quad \int_\Omega X\, dP \int_\Omega 1_A\, dP; \quad \forall A \in \mathcal{G}.$$

On a X est indépendant de $\mathcal{G}$; alors X est indépendant de 1_A ; $\forall A \in \mathcal{G}$.

Ce qui implique :

$$\int_A \mathbb{E}(X)\,dP = \left(\int_\Omega X\,dP\right)\left(\int_\Omega 1_A\,dP\right); \quad \forall A \in \mathcal{G},$$

$$= \int_\Omega X 1_A\,dP; \quad \forall A \in \mathcal{G},$$

$$= \int_A X\,dP; \quad \forall A \in \mathcal{G},$$

$$\overset{\text{def}}{=} \int_A \mathbb{E}(X \mid \mathcal{G})\,dP; \quad \forall A \in \mathcal{G}.$$

Ce qui implique :

$$\int_A (\mathbb{E}(X) - \mathbb{E}(X \mid \mathcal{G}))\ dP = 0; \quad \forall A \in \mathcal{G}.$$

On a $\mathbb{E}(X)$ une constante est $\mathcal{G}$-mesurable et $\mathbb{E}(X \mid \mathcal{G})$ une v.a. est $\mathcal{G}$-mesurable, donc $\mathbb{E}(X) - \mathbb{E}(X \mid \mathcal{G})$ une v.a. est $\mathcal{G}$-mesurable.

En conclut que :

$$\mathbb{E}(X) - \mathbb{E}(X \mid \mathcal{G}) = 0 \quad P\text{-p.s.}$$

D'où, $\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(X) \quad P$-p.s.

7) Si X une variable aléatoire telle que $X \in L^p(\Omega, \mathcal{F}, P)$, avec $p \geq 1$. Alors $\|\mathbb{E}(X \mid \mathcal{G})\|_{L^p(\Omega,\mathcal{F},P)} \leq \|X\|_{L^p(\Omega,\mathcal{F},P)}$.

On sait que $X \leq |X|$, donc :

$$\mathbb{E}(X \mid \mathcal{G}) \leq \mathbb{E}(|X| \mid \mathcal{G}),$$

d'après la croissance de l'espérance conditionnelle. Cela implique que :

$$|\mathbb{E}(X \mid \mathcal{G})|^p \leq \mathbb{E}(|X| \mid \mathcal{G})^p \leq \mathbb{E}(|X|^p \mid \mathcal{G}),$$

et donc :

$$||\mathbb{E}(X \mid \mathcal{G})||^p_{L^p(\Omega,\mathcal{F},P)} \leq \mathbb{E}(||X||^p_{L^p(\Omega,\mathcal{F},P)}),$$

ce qui donne :

$$||\mathbb{E}(X \mid \mathcal{G})||_{L^p(\Omega,\mathcal{F},P)} \leq ||X||_{L^p(\Omega,\mathcal{F},P)}.$$

8) Soit $\mathcal{G}, \mathcal{H}$ deux tribus de $\mathcal{F}$ telle que $\mathcal{H} \subset \mathcal{G}$, alors $\mathcal{E}(\mathbb{E}(X \mid \mathcal{H} \mid \mathcal{G})) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{G} \mid \mathcal{H})) = \mathbb{E}(X \mid \mathcal{H})$.

On sait que :

$$\mathbb{E}(\mathbb{E}(X \mid \mathcal{G}) \mid \mathcal{H}) = \mathbb{E}(X \mid \mathcal{G}),$$

telle que $\mathbb{E}(X \mid \mathcal{G})$ est une c.a. $\mathcal{G}$-mesurable (par definition). Donc $\mathbb{E}(X \mid \mathcal{G})$ une v.a. $\mathcal{G}$-mesurable et $\mathcal{H} \subset \mathcal{G}$
Ce qui implique $\mathbb{E}(X \mid \mathcal{H})$ est $\mathcal{G}$-mesurable, on obtient $\mathcal{E}(\mathbb{E}(X \mid \mathcal{H} \mid \mathcal{G})) = \mathbb{E}(X \mid \mathcal{H})$.

9) On sait que φ est une fonction convexe, alors il existe une fonction croissante k de $\mathbb{R}$ dans $\mathbb{R}$ et donc une fonction borélienne telle que, pour tout $x, b \in \mathbb{R}$:

$$\varphi(x) - \varphi(b) \leq k(b)(x - b).$$

Soit $W = \mathbb{E}(X \mid \mathcal{G})$, alors on a, pour tout $\omega \in \Omega$,

$$\varphi(X(\omega)) - \varphi(W(\omega)) \leq k(W(\omega))(X(\omega) - W(\omega)),$$

ce qui donne :

$$\varphi(X \mid \mathcal{G}) - \varphi(W \mid \mathcal{G}) \leq k(W)(X - W).$$

On aimerait intégrer cette inégalité sur un élément de $\mathcal{G}$, mais cela n'est pas possible car les variables aléatoires $\varphi(W)$ et $k(W)(X - W)$ peuvent ne pas être intégrables. Pour $p \in \mathbb{N}^*$, on introduit donc $A_p = \{|W| \leq p\}$ tel que les variables aléatoires $1_{A_p}k(W)(X - W)$ et $1_{A_p}\varphi(W)$ soient intégrables (on note que $k(W)$ est bornée sur A_p car k est croissante). On pose $A = \{\mathbb{E}(\varphi(X) \mid \mathcal{G})\} - \varphi(W) < 0\}$ et $B_p = A_p \cap A$.

Soit $p \in \mathbb{N}^*$, l'inégalité (2.2) donne

$$1_{B_p}(\varphi(X) - \varphi(W)) \geq 1_{B_p} k(W)(X - W),$$

et donc, en intégrant sur Ω,

$$\int_\Omega 1_{B_p}(\varphi(X) - \varphi(W))dP \geq \int_\Omega 1_{B_p} k(W)(X - W)dP.$$

Cela implique

$$\int_{B_p} \varphi(X) - \varphi(W)dP \geq \int_{B_p} k(W)(X - W)dP.$$

Comme W et $\mathbb{E}(\varphi(X) \mid \mathcal{G})$ sont $\mathcal{G}$-mesurables, on a donc 1_{B_p} est $\mathcal{G}$-mesurable car B_p est $\mathcal{G}$-mesurable (c'est-à-dire $B_p \in \mathcal{G}$). On a aussi que $k(W)$ est $\mathcal{G}$-mesurable car k est borélienne. Donc, $1_{B_p} k(W)$ est $\mathcal{G}$-mesurable.

On conclut que

$$\int_{B_p} (\varphi(X) - \varphi(W))dP = \int_{B_p} (1_{B_p}(\varphi(X) - \varphi(W)))dP = \mathbb{E}(1_{B_p}(\mathbb{E}(\varphi(X) \mid \mathcal{G}) - \varphi(W))).$$

Et

$$\int_{B_p} k(W)(X-W)dP = \int_\Omega (1_{B_p} k(W)(X-W))dP = \mathbb{E}(1_{B_p} k(W)(X-W)) = 0 \quad (\text{car } W \in \mathbb{E}(X \mid \mathcal{G})).$$

D'après (2.3), on conclut que

$$\int_{B_p} (\mathbb{E}(\varphi(X) \mid \mathcal{G}) - \varphi(W))dP \geq 0,$$

et donc

$$\int_{B_p} \mathbb{E}(\varphi(X) \mid \mathcal{G})dP \leq \int_{B_p} \varphi(W)dP.$$

Comme $\mathbb{E}(\varphi(X) \mid \mathcal{G}) - \varphi(W) < 0$ sur B_p car $B_p \subset A$, on a donc $P(B_p) = 0$ et $P(A) = P(\bigcup_{p \in \mathbb{N}^*} B_p) = 0$, d'où

$$\mathbb{E}(\varphi(X) \mid \mathcal{G}) \geq \varphi(W) \quad \text{p.s..}$$

Ainsi, on obtient

$$\mathbb{E}(\varphi(X) \mid \mathcal{G}) \geq \varphi(\mathbb{E}(X \mid \mathcal{G})) \quad \text{p.s..}$$

Espérance conditionnelle dans $L^2(\Omega, \mathcal{F}, P)$

Si X est une variable aléatoire de carré intégrable et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$, alors $\mathbb{E}(X \mid \mathcal{G})$ est la projection de X sur l'espace des variables aléatoires $\mathcal{G}$-mesurables de carré intégrable. Autrement dit, $\mathbb{E}(X \mid \mathcal{G})$ est une variable aléatoire qui minimise $\mathbb{E}((X-Y)^2)$ parmi les variables aléatoires Y, $\mathcal{G}$-mesurables.

Preuve

Soit Y une variable aléatoire telle que $\mathbb{E}(|Y|^2) < \infty$ et $\mathcal{G}$-mesurable.
Soit W une variable aléatoire $\mathcal{G}$-mesurable, alors :

$$\mathbb{E}(XW \mid \mathcal{G}) = W\mathbb{E}(X \mid \mathcal{G}),$$

(car W est $\mathcal{G}$-mesurable) ce qui implique :

$$\mathbb{E}[\mathbb{E}(XW \mid \mathcal{G})] = \mathbb{E}[W\mathbb{E}(X \mid \mathcal{G})],$$

ce qui donne :

$$\mathbb{E}(XW) = \mathbb{E}[W\mathbb{E}(X \mid \mathcal{G})]. \quad (3)$$

Ainsi, on a :

$$\mathbb{E}(W(X - \mathbb{E}(X \mid \mathcal{G}))) = \mathbb{E}(WX) - \mathbb{E}[W\mathbb{E}(X \mid \mathcal{G})].$$

D'après (3), cela implique :

$$\mathbb{E}[W\mathbb{E}(X \mid \mathcal{G})] - \mathbb{E}[W\mathbb{E}(X \mid \mathcal{G})] = 0,$$

alors :

$$\mathbb{E}(W(X - \mathbb{E}(X \mid \mathcal{G}))) = 0. \quad (4)$$

Posons $Y = \mathbb{E}(X \mid \mathcal{G}) + W$, ce qui implique que Y est $\mathcal{G}$-mesurable parce que $\mathbb{E}(X \mid \mathcal{G})$ est v.a. $\mathcal{G}$-mesurable (par définition) et W est $\mathcal{G}$-mesurable.

Alors :

$$\begin{aligned}\mathbb{E}((X-Y)^2) &= \mathbb{E}[(X-\mathbb{E}(X \mid \mathcal{G})-W)^2] = \mathbb{E}[(X-\mathbb{E}(X \mid \mathcal{G}))^2 + W^2 - 2W(X-\mathbb{E}(X \mid \mathcal{G}))] \\ &= \mathbb{E}[(X-\mathbb{E}(X \mid \mathcal{G}))^2 + \mathbb{E}(W^2) - 2\mathbb{E}(W(X-\mathbb{E}(X \mid \mathcal{G})))].\end{aligned}$$

D'après (4), on a $\mathbb{E}[(X-\mathbb{E}(X \mid \mathcal{G}))W] = 0$, donc :

$$\mathbb{E}((X-Y)^2) = \mathbb{E}[(X-\mathbb{E}(X \mid \mathcal{G}))^2] + \mathbb{E}(W^2).$$

Alors, on conclut $\mathbb{E}((X-Y)^2)$ est minimise si :

$$\mathbb{E}(W^2) = 0,$$

ce qui implique $W^2 = 0$ et

$$W = 0$$

Donc, on déduit que la valeur $\mathbb{E}((X-Y)^2)$ est minimum si

$$Y = \mathbb{E}(X \mid \mathcal{G}) + W$$

. D'où $Y = \mathbb{E}(X \mid \mathcal{G})$. Ce qui termine la preuve.

Variance conditionnelle

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et soit Y une variable aléatoire définie sur cet espace. Soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$. Alors, on définit la variance conditionnelle comme suit :

$$\mathrm{Var}(Y \mid \mathcal{G}) = \mathbb{E}(Y^2 \mid \mathcal{G}) - \mathbb{E}^2(Y \mid \mathcal{G}).$$

En vertu de l'inégalité de Jensen : Soit F une fonction convexe, on a :

$$\mathbb{E}[F(Y) \mid \mathcal{F}] \geq F(\mathbb{E}[Y \mid \mathcal{F}]).$$

Densité conditionnelle

Si $f_{X,Y}$ est la densité jointe de (X, Y), alors :

- La densité conditionnelle de Y sachant $X = x$ est définie par :

$$f_{Y|X}(x, y) = f(Y \mid X) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_X(x)} & \text{si } f_X(x) > 0, \\ 0 & \text{si } f_X(x) = 0. \end{cases}$$

- La densité conditionnelle de X sachant $Y = y$ est définie par :

$$f_{X|Y}(x, y) = f(X \mid Y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{si } f_Y(y) > 0, \\ 0 & \text{si } f_Y(y) = 0. \end{cases}$$

BIBLIOGRAPHIE

[1] M.Jeanblanc. (2006). M.Jeanblanc. (2006). Cours de calcul stochastique. Master 2IF EVRY. Lecture Notes.

[2] J. Yong and X.Y. Zhou (1999), Stochastic controls, Hamiltonian Systems and HJB Equations. Springer Verlag.

[3] Polycopie, Espérance conditionelle, HAFAYED Mokhtar, Université de Biskra.