V Espérance conditionnelle

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

Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé et soit $\mathcal{U}$ une sous tribu de $\mathcal{F}$. Soit $X \in \mathcal{L}^2_{\mathbb{R}^d}(\Omega, \mathcal{F}, P)$ (l'espace des vecteurs aléatoires de dimension d à carré intégrable). On notera que le sous espace $L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)$ de $L^2_{\mathbb{R}^d}(\Omega, \mathcal{F}, P)$ constitué des classes d'équivalence des vecteurs aléatoires $\mathcal{U}-$mesurables, est un sous espace fermé et on a

$$L^2_{\mathbb{R}^d}(\Omega, \mathcal{F}, P) = L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P) \oplus \left(L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)\right)^{\perp},$$

où

$$\left(L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)\right)^{\perp} = \left\{Y \in L^2_{\mathbb{R}^d}(\Omega, \mathcal{F}, P) : \mathbb{E}(\langle Y, Z\rangle) = 0 \ \forall Z \in L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)\right\}$$

est l'orthogonal de $L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)$. La relation d'équivalence étant définie par:

$$U \mathcal{R} V \Leftrightarrow U = V \ p.s.$$

Ainsi X s'écrit d'une manière unique sous la forme:

$$X = Y + Z \qquad (*)$$

où $Y \in L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)$ et $Z \in \left(L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)\right)^{\perp}$.

Définition:

On appelle espérance conditionnelle de X *(on la note* $\mathbb{E}(X \mid \mathcal{U})$ *ou* $\mathbb{E}^{\mathcal{U}}(X)$*), sa projection*

orthogonale sur $L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)$, *i.e.*

$$\mathbb{E}(X \mid \mathcal{U}) = Y,$$

où Y *est défini par l'égalité* $(*)$.

On notera que $\mathbb{E}(X \mid \mathcal{U})$ est une variable aléatoire.

Remarque:

Il est clair que l'application $X \mapsto \mathbb{E}(X \mid \mathcal{U})$ est linéaire et que si X est $\mathcal{U}-$mesurable, alors

$$\mathbb{E}(X \mid \mathcal{U}) = X \ \ p.s.$$

En effet, X admet comme décomposition, la décomposition triviale $X = X + 0$, où

$X \in L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)$ et $0 \in \left(L^2_{\mathbb{R}^d}(\Omega, \mathcal{U}, P)\right)^{\perp}$.

Dans toute la suite, on ne considère que le cas où $d = 1$. Les résultats ainsi les démonstrations, dans le cas général, sont exactement les mêmes.

Proposition 1:

Les assertions suivantes sont équivalentes:

(i) $Y = \mathbb{E}(X \mid \mathcal{U})$ *p.s.*

(ii) Y *est* $\mathcal{U}-$*mesurable et* $\int_A X dP = \int_A Y dP$ *pour tout* $A \in \mathcal{U}$

(iii) Y *est* $\mathcal{U}-$*mesurable et* $\mathbb{E}(ZX) = \mathbb{E}(ZY)$ pour tout $Z \in \mathcal{L}^2(\Omega, \mathcal{U}, P)$

Remarque:

La propriété (iii) avec $Z=1$, entraine en particulier que

$$\mathbb{E}\left(\mathbb{E}\left(X \mid \mathcal{U}\right)\right)=\mathbb{E}\left(X\right).$$

Preuve:

$(i) \Leftrightarrow (iii)$ est une conséquence directe de la définition de la projection orthogonale. En effet; on a la décomposition évidente $X=Y+(X-Y)$ avec $Y\in L^2\left(\Omega,\mathcal{U},P\right)$ et $X-Y\in\left(L^2\left(\Omega,\mathcal{U},P\right)\right)^{\perp}$, d'où par liéarité de l'espérance,

$$\mathbb{E}\left(ZX\right)=\mathbb{E}\left(ZY\right)+\mathbb{E}\left(Z\left(X-Y\right)\right)=\mathbb{E}\left(ZY\right)$$

pour tout $Z\in\mathcal{L}^2\left(\Omega,\mathcal{U},P\right)$.

$(iii)\Rightarrow(ii)$ car (iii) avec $Z=\mathbf{1}_A$, où $A\in\mathcal{U}$ donne (ii). En effet,

$$\mathbb{E}\left(ZX\right)=\int\limits_A XdP \text{ et } \mathbb{E}\left(ZY\right)=\int\limits_A YdP.$$

$(ii)\Rightarrow(iii)$

Si on a (ii), alors on a pour tout $\varphi=\sum\limits_i c_i\mathbf{1}_{A_i}$ variable aléatoire étagée,

$$\mathbb{E}\left(\varphi X\right)=\sum_i c_i\mathbb{E}\left(X\mathbf{1}_{A_i}\right)=\sum_i c_i\int\limits_{A_i}XdP=\sum_i c_i\int\limits_{A_i}YdP=\mathbb{E}\left(\varphi Y\right).$$

Si maintenant Z est une variable aléatoire positive. Alors il existe une suite croissante de variables aléatoires étagées (φ_n) qui converge vers Z, d'où d'aprés le théorème de la convergence monotone,

$$\mathbb{E}\left(ZX\right)=\lim_{n\to\infty}\mathbb{E}\left(\varphi_nX\right)=\lim_{n\to\infty}\mathbb{E}\left(\varphi_nY\right)=\mathbb{E}\left(ZY\right).$$

Si Z est de signe quelconque, alors on a $Z=Z^++Z^-$, où $Z^+=\max\left(Z,0\right)\geq 0$ et $Z^-=\max\left(-Z,0\right)\geq 0$, d'où

$$\mathbb{E}\left(ZX\right)=\mathbb{E}\left(Z^+X\right)+\mathbb{E}\left(Z^-X\right)=\mathbb{E}\left(Z^+Y\right)+\mathbb{E}\left(Z^-Y\right)=\mathbb{E}\left(ZY\right),$$

qui signifie que (iii) est vraie.■

Exemple:

Si $\mathcal{U}=\{\phi,\Omega\}$ est la plus petite sous tribu, alors

$$\mathbb{E}\left(X\mid\mathcal{U}\right)=\mathbb{E}\left(X\right)\ p.s.$$

En effet, $\mathbb{E}\left(X\right)$ est une constante donc $\mathcal{U}-$ mesurable et on a

$$\int\limits_{\phi}XdP=0=\int\limits_{\phi}YdP \text{ et } \int\limits_{\Omega}XdP=\mathbb{E}\left(X\right)=\int\limits_{\Omega}\mathbb{E}\left(X\right)dP\ ,$$

d'où l'affirmation d'aprés (ii). On notera que les seules variables aléatoires $\mathcal{U}-$mesurables sont les constantes.

5.2. Espérance conditionnelle sur $L^1(\Omega, \mathcal{F}, P)$

Dans ce qui suit, on admettra le théorème suivant:

Théorème:

Pour qu'une variable aléatoire $Y, \mathcal{U}-$*mesurable soit positive p.s., il faut et il suffit que*

$\int_A Y dP$ *soit positive pour tout* $A \in \mathcal{U}$

On a alors le théorème suivant:

Théorème:

L'application linéaire $\mathbb{E}(. \mid \mathcal{U})$ *est positive (i.e.* $X \geq 0$ *p.s.* $\Rightarrow \mathbb{E}(X \mid \mathcal{U}) \geq 0$ *p.s.), continue*

pour la norme de $L^1(\Omega, \mathcal{F}, P)$ *et est de norme* ≤ 1.

Remarque:

La positivité de l'espérance conditionnelle est équivalente à la propriété suivante:

$$X \geq Z \text{ } p.s. \Rightarrow \mathbb{E}(X \mid \mathcal{U}) \geq \mathbb{E}(Z \mid \mathcal{U}) \text{ } p.s.$$

Démonstration:

La positivité provient directement de (ii) de la dernière proposition.

Soit $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$. On a $-|X| \leq X \leq |X|$, d'où en utilisant la linéarité et la positivité de l'espérance conditionnelle,

$$-\mathbb{E}(|X| \mid \mathcal{U}) \leq \mathbb{E}(X \mid \mathcal{U}) \leq \mathbb{E}(|X| \mid \mathcal{U}).$$

Il s'en suit que

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

d'où par positivité de l'espérance,

$$\|\mathbb{E}(X \mid \mathcal{U})\|_1 = \mathbb{E}(|\mathbb{E}(X \mid \mathcal{U})|) \leq \mathbb{E}(\mathbb{E}(|X| \mid \mathcal{U})) = \mathbb{E}(|X|) = \|X\|_1,$$

qui signifie que $\mathbb{E}(. \mid \mathcal{U})$ est Lipschitzienne pour la norme $\|.\|_1$ donc continue. De plus, on a

$$\|\mathbb{E}(. \mid \mathcal{U})\| := \sup_{\|X\|_1 \leq 1} \|\mathbb{E}(X \mid \mathcal{U})\|_1 \leq 1.$$

■

Conséquence:

Comme $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ est dense dans $\mathcal{L}^1(\Omega, \mathcal{F}, P)$, alors l'application $\mathbb{E}(. \mid \mathcal{U})$ se prolonge par continuité sur $\mathcal{L}^1(\Omega, \mathcal{F}, P)$. Ainsi, si $X \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$, il existe une suite (X_n) de $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, qui converge vers X. Dans cas, par définition

$$\mathbb{E}(X \mid \mathcal{U}) = \lim_{n \to \infty} \mathbb{E}(X_n \mid \mathcal{U}) \text{ pour la norme } \|.\|_1.$$

Remarque:

Par un passage à la limite, la proposition 1 reste valable si $X \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ au lieu de $\mathcal{L}^2(\Omega, \mathcal{F}, P)$ et si Z est une variable aléatoire bornée et $\mathcal{U}-$mesurable (au lieu de $Z \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$) car l'ensemble de ces variables aléatoires est dense dans $\mathcal{L}^2(\Omega, \mathcal{F}, P)$.

Définition:

Soient $\mathcal{U}$ et $\mathcal{V}$ deux sous tribus de $\mathcal{F}$.

On dira qu'une variable aléatoire X est indépendante de $\mathcal{U}$ si elle est indépendante avec

toute variable aléatoire $\mathcal{U}-$mesurable.

On dira que $\mathcal{U}$ et $\mathcal{V}$ sont indépendantes, si toute variable aléatoire $\mathcal{U}-$mesurable est

indépendante avec toute variable aléatoire $\mathcal{V}-$mesurable.

On notera que si $\mathcal{U} = \{\phi, \Omega\}$, alors toute variable aléatoire X est indépendante de $\mathcal{U}$, car les seules variables aléatoires $\mathcal{U}-$mesurables sont les constantes.

Proposition 2:(propriétés de l'espérance conditionnelle)

Pour tout $X \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$, on a:

(i) $\mathbb{E}(X \mid \mathcal{U}) = X$ *p.s. si X est $\mathcal{U}-$mesurable*

(ii) $\mathbb{E}(X \mid \mathcal{U}) = \mathbb{E}(X)$ *p.s. si X est indépendante de $\mathcal{U}$. En particulier si $\mathcal{U} = \{\phi, \Omega\}$, on a*

$$\mathbb{E}(X \mid \{\phi, \Omega\}) = X \text{ } p.s.$$

(iii)*Si $\mathcal{U}$ et $\mathcal{V}$ deux sous tribus de $\mathcal{F}$ telles que $\mathcal{V} \subset \mathcal{U}$, alors on a*

$$\mathbb{E}(X \mid \mathcal{V}) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{U}) \mid \mathcal{V}) \text{ } p.s.$$

(Propriété d'emboitement)

(iv) $\mathbb{E}(\mathbb{E}(X \mid \mathcal{U})) = \mathbb{E}(X)$

(v) $Z\mathbb{E}(X \mid \mathcal{U}) = \mathbb{E}(ZX \mid \mathcal{U})$ *pour toute variable aléatoire bornée et $\mathcal{U}-$mesurable*

(vi) $\mathbb{E}(ZX) = \mathbb{E}(Z\mathbb{E}(X \mid \mathcal{U}))$ *pour toute variable aléatoire bornée et $\mathcal{U}-$mesurable*

Preuve:

Soit (X_n) une suite de $\mathcal{L}^2(\Omega, \mathcal{F}, P)$, qui converge vers X.

(i)Comme X est $\mathcal{U}-$ mesurable, alors les variables aléatoires X_n le sont aussi, d'où

$$\mathbb{E}(X_n \mid \mathcal{U}) = X_n \text{ } p.s.,$$

d'où par passage à la limite L^1

$$\mathbb{E}(X \mid \mathcal{U}) = X \text{ } p.s.$$

(ii)On a pour tout $A \in \mathcal{U}$

$$\int_A \mathbb{E}(X)\, dP = \mathbb{E}(X) P(A) = \mathbb{E}(X)\mathbb{E}(\mathbf{1}_A) \overset{ind}{=} \mathbb{E}(X\mathbf{1}_A) = \int_A X dP,$$

d'où

$$\mathbb{E}(X \mid \mathcal{U}) = \mathbb{E}(X) \text{ } p.s.$$

(iii)La propriété d'enboitement étant vraie pour $X \in \mathcal{L}^2(\Omega, \mathcal{F}, P)$ (voir série $N°4$ TD), alors on peut écrire que

$$\mathbb{E}(X_n \mid \mathcal{V}) = \mathbb{E}(\mathbb{E}(X_n \mid \mathcal{U}) \mid \mathcal{V}) \text{ } p.s.$$

D'autre part, la continuité de l'espérance conditionnelle pour la norme de L^1 entraine que

$$\lim_{n\to\infty}\mathbb{E}\left(X_n\mid\mathcal{V}\right)=\mathbb{E}\left(X\mid\mathcal{V}\right)\text{ et }\lim_{n\to\infty}\mathbb{E}\left(\mathbb{E}\left(X_n\mid\mathcal{U}\right)\mid\mathcal{V}\right)=\mathbb{E}\left(\mathbb{E}\left(X\mid\mathcal{U}\right)\mid\mathcal{V}\right),$$

d'où par unicité de la limite

$$\mathbb{E}\left(X\mid\mathcal{V}\right)=\mathbb{E}\left(\mathbb{E}\left(X\mid\mathcal{U}\right)\mid\mathcal{V}\right)\ p.s.$$

(iv)Provient directement de la propriété (iii) avec $\mathcal{V}=\{\phi,\Omega\}$.
(v)Soit $A\in\mathcal{U}$. Alors on a

$$\begin{aligned}\int_A Z\mathbb{E}\left(X\mid\mathcal{U}\right)dP &= \mathbb{E}\left(\mathbf{1}_A Z\mathbb{E}\left(X\mid\mathcal{U}\right)\right)\\ &= \mathbb{E}\left(\mathbf{1}_A ZX\right),\text{ d'aprés }(iii)\text{ de la pro.1 puisque }\mathbf{1}_A Z\text{ est }\mathcal{U}-\text{mes}\\ &= \int_A ZXdP\\ &= \int_A \mathbb{E}\left(ZX\mid\mathcal{U}\right)dP\text{ d'aprés }(ii)\text{ de la pro.1,}\end{aligned}$$

d'où l'affirmation
(vi)En appliquant $\mathbb{E}$ aux deux membres de (v) et en utilisant (iv), on obtient

$$\mathbb{E}\left(ZX\right)=\mathbb{E}\left(Z\mathbb{E}\left(X\mid\mathcal{U}\right)\right).$$

■

Proposition 3:(inégalité de Jensen au conditionnel)
Soit $X\in\mathcal{L}^1(\Omega,\mathcal{F},P)$ *et soit* $\varphi:\mathbb{R}\to\mathbb{R}$ *une fonction convexe telle que* $\varphi\circ X\in\mathcal{L}^1(\Omega,\mathcal{F},P)$.
Alors on a

$$\varphi\circ\left(\mathbb{E}\left(X\mid\mathcal{U}\right)\right)\leq\mathbb{E}\left(\varphi\circ X\mid\mathcal{U}\right)\ p.s.$$

La démonstration se conduit exactement de la même manière que celle du cas classique.

Cas particuliers importants
En utilisant la notation $X^+=\max(X,0)$, on a

$$\begin{aligned}\mathbb{E}\left(X\mid\mathcal{U}\right)^+ &= \mathbb{E}\left(X^+\mid\mathcal{U}\right),\\ \left|\mathbb{E}\left(X\mid\mathcal{U}\right)\right| &\leq \mathbb{E}\left(|X|\mid\mathcal{U}\right),\\ \left|\mathbb{E}\left(X\mid\mathcal{U}\right)\right|^p &\leq \mathbb{E}\left(|X|^p\mid\mathcal{U}\right)\text{ pour tout }p\in[1,+\infty[.\end{aligned}$$

Corollaire:
On a, pour tout $p\in[1,+\infty[$

$$\left\|\mathbb{E}\left(X\mid\mathcal{U}\right)\right\|_p\leq\left\|X\right\|_p$$

et donc $\mathbb{E}(. \mid \mathcal{U})$ *est continue de* $L^p(\Omega, \mathcal{F}, P)$ *dans* $L^p(\Omega, \mathcal{F}, P)$.

Preuve:

On a, d'aprés le troisième cas important,

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

d'où l'affirmation.■

L'espérance étant un cas particulier de l'espérance conditionnelle, on a la proposition suivante (sans démonstration).

Proposition 4:

Soit (X_n) *une suite de variables aléatoires numériques intégrables. Alors on a:*

1.*Si les* X_n *sont positives et si* $X_n \uparrow X$ *intégrable, alors*

$$\mathbb{E}(X_n \mid \mathcal{U}) \text{ converge vers } \mathbb{E}(X \mid \mathcal{U}) \text{ p.s. et dans } L^1(\Omega, \mathcal{F}, P)$$

"Théorème de Beppo-Lèvy au conditionnel"

2.*Si les* X_n *sont positives, alors on a:*

a)Si $\underline{\lim} X_n$ *est intégrable, alors*

$$\mathbb{E}(\underline{\lim} X_n \mid \mathcal{U}) \leq \underline{\lim} \mathbb{E}(X_n \mid \mathcal{U})$$

b)Si $\overline{\lim} X_n$ *est intégrable, alors*

$$\mathbb{E}\left(\overline{\lim} X_n \mid \mathcal{U}\right) \geq \overline{\lim} \mathbb{E}(X_n \mid \mathcal{U})$$

"Lemme de Fatou au conditionnel"

3.*On suppose que la variable aléatoire* $\sup_n |X_n|$ *est intégrable et que* (X_n) *converge vers* X

p.s., alors on a

$$\mathbb{E}(X_n \mid \mathcal{U}) \text{ converge vers } \mathbb{E}(X \mid \mathcal{U}) \text{ p.s. et dans } L^1(\Omega, \mathcal{F}, P).$$

Si de plus $\sup_n |X_n| \in L^p(\Omega, \mathcal{F}, P)$, *alors la convergence a lieu dans* $L^p(\Omega, \mathcal{F}, P)$.

"Théorème de la convergence dominée".

Définition:

On appelle probabilité conditionnelle d'un événement $A \in \mathcal{F}$ *par rapport à* $\mathcal{U}$, *l'espérance*

conditionnelle

$$P(A \mid \mathcal{U}) := \mathbb{E}(\mathbf{1}_A \mid \mathcal{U}).$$

Remarque:

Si de plus $\mathcal{U} = \{\phi, B, B^c, \Omega\}$ *la tribu de Bernoulli associée à* $B \in \mathcal{F}$, *alors* $P(A \mid \mathcal{U})$ *n'est*

autre que la probabilité conditionnelle de A *sachant* B.

En effet, comme les fonctions $\mathcal{U}-$mesurables sont constantes sur B et sur B^c, alors par caractérisation de l'espérance cnditionnelle, on a $P(A \mid \mathcal{U}) = \alpha$ sur B, d'où

$$\alpha P(B) = \int_B P(A \mid \mathcal{U})\, dP = \int_B \mathbf{1}_A dP = P(A \cap B),$$

par suite

$$\alpha = \frac{P(A \cap B)}{P(B)}, \text{ si } P(B) \neq 0,$$

soit

$$P(A \mid \mathcal{U}) = P(A \mid B).$$

5.3 Espérance conditionnelle par rapport à une variable aléatoire

Soient $(\Omega, \mathcal{F}, P)$ un espace probabilisé, $(E, \mathcal{E})$ un espace probabilisable et $U : (\Omega, \mathcal{F}) \to (E, \mathcal{E})$ une variable aléatoire.

Définition:

Soit X une variable aléatoire à valeurs dans $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$ et soit $\sigma(U)$ la tribu engendrée par

U. On appelle espérance conditionnelle de X par rapport à U, l'espérance conditionnelle

$$\mathbb{E}(X \mid U) := \mathbb{E}(X \mid \sigma(U)).$$

Remarque:

Comme les fonction $\sigma(U)-$mesurables s'écrivent sous la forme $\varphi(U)$, où

$\varphi : (E, \mathcal{E}) \to (\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$, alors $\mathbb{E}(X \mid U) = \varphi(U)$. En pratique, on calcule d'abord

$$\mathbb{E}(X \mid U = u) = \varphi(u).$$

Dans ce cas

$$\mathbb{E}(X \mid U) = \varphi(U)$$