Table des matières

Introduction

L'espérance conditionnelle est un concept important en probabilités notamment utilisé dans des domaines tels que l'étude des martingales est l'intégration stochastique.

Chapitre 1

Espérance Conditionnelle

On se place dans le cas général d'un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$.

1.1 Conditionnement par un événement :

On se donne B un événement (*i.e.* $B \in \mathcal{A}$) tel que $\mathbb{P}(B) > 0$. Pour tout $A \in \mathcal{A}$, on pose :

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

Définition 1.1.1 *Q_B est la mesure de probabilité sachant l'événement B. On la note souvent $\mathbb{P}(./B)$*

Proposition 1.1.1 *Soit X une variable aléatoire. Si X est $\mathbb{P}$-intégrable i.e. $\mathbb{E}|X| = \int |X| d\mathbb{P} < +\infty$ alors X est aussi $\mathbb{P}(./B)$-intégrable. De plus, on a :*

$$\begin{aligned}\mathbb{E}(X/B) = \int X d\mathbb{P}(./B) &= \frac{1}{\mathbb{P}(B)} \int_B X d\mathbb{P} \\ &= \frac{1}{\mathbb{P}(B)} \mathbb{E}[X \mathbb{I}_B] \qquad (1.1)\end{aligned}$$

$\mathbb{E}(X/B)$ est l'espérance de X sachant (ou conditionnelle à) l'événement B.

DÉMONSTRATION. Ce résultat se démontre par étapes, en commençant par une fonction indicatrice, ce qui permet de passer au fonctions étagées puis aux fonctions mesurable positives, pour conclure sur les fonctions intégrables.

*É*tape 1 : $X = \mathbb{I}_A$

$\int \mathbb{I}_A d\mathbb{P}(./B) = \mathbb{P}(A/B) = \dfrac{\mathbb{P}(B \cap A)}{\mathbb{P}(B)} = \dfrac{1}{\mathbb{P}(B)} \displaystyle\int \mathbb{I}_{B\cap A} d\mathbb{P} = \dfrac{1}{\mathbb{P}(B)} \displaystyle\int \mathbb{I}_A d\mathbb{P}$

*É*tape 2 : $X = \displaystyle\sum_{j\in J} \lambda_j \mathbb{I}_{A_j}$ o*ù* J est un ensemble fini et les (A_j) sont des événements disjoints. On obtient la relation (1.1) grâce à la linéarité de l'intégrale et l'étape précédente.

*É*tape 3 : X est mesurable positive. Il existe une suite croissante de variables $(X_n)_{n\in\mathbb{N}}$ étagées positives qui converge vers X. D'aprés l'étape 2, on a :

$\forall n \in \mathbb{N}$: $\int X_n d\mathbb{P}(./B) = \dfrac{1}{\mathbb{P}(B)} \displaystyle\int_B X_n d\mathbb{P}$

On applique en suite le théorème de Beppo Levi aux deux intégrales.

*É*tape 4 : X intégrable. La variable aléatoire X se décompose sous la forme $X = X^+ - X^-$ avec : $X^+ = \max(X, 0)$ et $X^- = \max(-X, 0)$.

Les variables X^+ et X^- sont positives et $\mathbb{P}$-intégrables. L'étape 3 assure que :

$\int X^+ d\mathbb{P}(./B) = \dfrac{1}{\mathbb{P}(B)} \displaystyle\int_B X^+ d\mathbb{P} < +\infty$ et

$\int X^- d\mathbb{P}(./B) = \dfrac{1}{\mathbb{P}(B)} \displaystyle\int_B X^- d\mathbb{P} < +\infty$

donc $|X| = X^+ + X^-$ est $\mathbb{P}(./B)$-intégrable et on obtient (1.1).

Pour toute variable aléatoire X (admettons qu'elle est à valeur dans $\mathbb{R}$, mais elle pourrait aussi être à valeurs dans un autre espace) ce qui précéde définit aussi la probabilité ,sachant B, que X appartienne à n'importe quel borélien de $\mathbb{R}$. Autrement dit on construit facilement la loi conditionelle de X et, par le théorème de transfére, l'éspérance conditionnelle se calcule aussi en utilisant cette loi :

Définition 1.1.2 *La loi de X sachant (ou conditionnelle à) l'événement B est la transportée de* $\mathbb{P}(./B)$ *par X. Soit, pour tout* $E \in \mathcal{B}(\mathbb{R})$ *:*
$\mathbb{P}_X(E/B) = \mathbb{P}(X^{-1}(E)/B) = \mathbb{P}(X \in E/B)$
et on a :

$$\mathbb{E}(X/B) = \int_{\mathbb{R}} x d\mathbb{P}_X(x/B)$$

Exemples 1.1.1 *Soit* (X_1, X_2) *un couple de variables aléatoires indépendantes et de* *même loi définie par :*

$$\mathbb{P}(X_1 = k) = \alpha^k(1-\alpha) \qquad \forall k \in \mathbb{N}$$

où α *est fixé dans* $]0,1[$. *(On note* $\tilde{G}(\alpha)$ *cette loi)*
On pose :

$$N = \begin{cases} 1 & si\ X_1 \leq X_2 \\ 2 & sinon \end{cases}$$

On va déterminer la loi de X_1 *sachant* $N = 1$.

Commençons par calculer la loi de la variable aléatoire N. La loi du couple (X_1, X_2) *étant symétrique, on a :*

$$\mathbb{P}(X_1 < X_2) = \frac{1}{2}(1 - \mathbb{P}(X_1 = X_2))$$

On en déduit :

$$\mathbb{P}(N = 1) = \mathbb{P}(X_1 = X_2) + \mathbb{P}(X_1 < X_2) = \frac{1}{2}(1 + \mathbb{P}(X_1 = X_2))$$

$$= \frac{1}{2}(1 + \sum_{k=0}^{+\infty} \mathbb{P}(X_1 = k, X_2 = k))$$

$$= \frac{1}{2}(1 + \sum_{k=0}^{+\infty} \mathbb{P}(X_1 = k)\mathbb{P}(X_2 = k))$$

car les variables aléatoires sont indépendantes.

$$\mathbb{P}(N = 1) = \frac{1}{2}(1 + (1-\alpha)^2 \sum_{k=0}^{+\infty} \alpha^{2k})$$

$$= \frac{1}{2}(1 + \frac{(1-\alpha)^2}{1-\alpha^2}) = \frac{1}{1+\alpha}$$

La loi conditionnelle de X_1 *sachant l'événement* $[N = 1]$ *est donnée par :*

$$\mathbb{P}(X_1 = k / N = 1) = \frac{\mathbb{P}(N = 1, X_1 = k)}{\mathbb{P}(N = 1)} = \frac{\mathbb{P}(X_2 \geq k, X_1 = k)}{\mathbb{P}(N = 1)}$$

$$= (1+\alpha)\mathbb{P}(X_1 = k) \sum_{l \geq k} \mathbb{P}(X_2 = l)$$

$$= (1+\alpha)\alpha^k(1-\alpha)\sum_{l\geq k}(1-\alpha)\alpha^l$$

$$= (1-\alpha^2)\alpha^{2k}, \qquad \forall k \in \mathbb{N}$$

La loi conditionnelle de X_1 sachant l'événement $[N=1]$ est une loi $\tilde{G}(\alpha^2)$. On en déduit que l'éspérance de X_1 sachant l'événement $[N=1]$ est égale à :

$$\mathbb{E}[X_1/N=1] = \sum_{k=0}^{+\infty} k\mathbb{P}(X_1 = k/N=1) = \frac{\alpha^2}{1-\alpha^2}$$

De la même façon, on obtient la loi de X_1 sachant $[N=2]$

$\mathbb{P}(X_1 = k/N=2) = (1-\alpha^2)\alpha^{k-1}(1-\alpha^k) \qquad \forall k \geq 1$

et l'éspérance conditionnelle de X_1 sachant $N=2$ est égale à $\dfrac{1+\alpha+\alpha^2}{1-\alpha^2}$

Remarque 1.1.1 *L'éspérance de X_1 s'exprime à partir des espérances conditionnelles $\mathbb{E}(X_1/N=1)$ et $\mathbb{E}(X_1/N=2)$. En effet, on peut décomposer X_1 en :*

$$X_1 = X_1\mathbb{I}_{N=1} + X_1\mathbb{I}_{N=2}$$

et grâce à (1.1), on obtient :

$$\mathbb{E}[X_1] = \mathbb{E}[X_1\mathbb{I}_{N=1}] + \mathbb{E}[X_1\mathbb{I}_{N=2}]$$

$$= \mathbb{P}(N=1)\mathbb{E}(X_1/N=1) + \mathbb{P}(N=2)\mathbb{E}(X_1/N=2)$$

$$= \frac{\alpha}{1-\alpha}$$

1.2 Conditionnement par une partition au plus dénombrable d'événements :

Dans cette partie, $(B_n)_{n\in\mathbb{N}}$ désigne une partition de Ω telle que $\mathbb{P}(B_n) \neq 0$ pour tout n. On note $\mathcal{B}$ la tribu engendrée par cette partition $\mathcal{B} = \sigma(B_n, n \in \mathbb{N})$.

1.2.1 Definition et propriétés de la probabilité conditionnelle :

Soit A un événement . On définit la probabilité conditionnelle de A sachant B comme étant la variable aléatoire qui prend la valeur $\mathbb{P}(A/B_i)$ pour tout $\omega \in B_i$.On la note

$\mathbb{P}(A/\mathcal{B})$. On a donc pour tout $\omega \in \Omega$

$$\mathbb{P}(A/\mathcal{B})(\omega) = \sum_{n \in \mathbb{N}} \mathbb{P}(A/B_n)\mathbb{I}_{B_n}(\omega) \tag{1.2}$$

Proposition 1.2.1 *Soit A fixé, la variable aléatoire $\mathbb{P}(A/\mathcal{B})$ satisfait les propriétés suivantes :*

$$[\pi_1] : \begin{cases} \mathbb{P}(A/\mathcal{B}) \ est\ \mathcal{B} - mesurable \\ pour\ tout\ B \in \mathcal{B}, on\ a \displaystyle\int_B \mathbb{P}(A/\mathcal{B})d\mathbb{P} = \mathbb{P}(A \cap B) \end{cases}$$

DÉMONSTRATION. En tant que fonction constante sur chaque B_i, la variable aléatoire $\mathbb{P}(A/\mathcal{B})$ est mesurable. Soit $B \in \mathcal{B}$. Il s'écrit de la forme $B = \cup_{i \in I_0} B_i$ et pour toute fonction $\mathbb{P}$-intégrable f, on a :

$$\int_B f d\mathbb{P} = \sum_{i \in I_0} \int_{B_i} f d\mathbb{P}$$

On obtient

$$\begin{aligned}
\int_B \mathbb{P}(A/\mathcal{B})d\mathbb{P} &= \sum_{i \in I_0} \int_{B_i} \mathbb{P}(A/\mathcal{B})d\mathbb{P} \\
&= \sum_{i \in I_0} \int_{B_i} \sum_{j \in \mathbb{N}} \mathbb{P}(A/B_j)\mathbb{I}_{B_j} d\mathbb{P} \\
&= \sum_{i \in I_0} \sum_{j \in \mathbb{N}} \mathbb{P}(A/B_j) \int_{B_i \cap B_j} d\mathbb{P} \\
&= \sum_{i \in I_0} \mathbb{P}(A/B_i)\mathbb{P}(B_i) \quad \text{car } B_i \cap B_j = \emptyset \text{ si } j \neq i \text{ et } B_i \text{ sinon} \\
&= \mathbb{P}(A \cap \cup_{i \in I_0}) = \mathbb{P}(A \cap B)
\end{aligned}$$

Proposition 1.2.2 *À ω fixé, $\mathbb{P}(./\mathcal{B})$ est une mesure de probabilité.*

DÉMONSTRATION. En effet, si on fixe $\omega \in \Omega$, il existe un unique élément de la partition, noté B^ω tel que $\omega \in B^\omega$. On a alors

$$\mathbb{P}(./\mathcal{B})(\omega) = \sum_{j\in\mathbb{N}} \mathbb{P}(./B_j)\mathbb{I}_{B_j}(\omega) = \mathbb{P}(./B^\omega)$$

c'est bien une probabilité.

Remarque 1.2.1 *Supposons qu'il existe des éléments de la partition ayant une probabilité nulle. On note* $I = \{i : \mathbb{P}(B_i) \neq 0\}$. *La probabilité conditionnelle sachant* $\mathcal{B}$ *est alors définie par*

$$\mathbb{P}(A/B)(\omega) = \sum_{n\in I} \mathbb{P}(A/B_n)\mathbb{I}_{B_n}(\omega) \tag{1.3}$$

$\mathbb{P}(A/B)$ *est nulle sur l'ensemble* $\cup_{i\in I^c} B_i$.

1.2.2 Definition et propriétés de l'éspérance conditionnelle :

Définition 1.2.1 *Soit* X *une variable aléatoire dans* $\mathrm{L}^1(\Omega, \mathcal{A}, \mathbb{P})$, *l'éspérance de* X *conditionnellement à* $\mathcal{B}$ *est la variable définie par*

$$\mathbb{E}(X/\mathcal{B}) = \int X d\mathbb{P}(./\mathcal{B})(\omega) = \sum_{i\in\mathbb{N}} \mathbb{E}(X/B_i)\mathbb{I}_{B_i}(\omega) \tag{1.4}$$

pour tout $\omega \in \Omega$.

Remarque 1.2.2 *Pour tout* ω, *c'est tout simplement*

$$\mathbb{E}(X/\mathcal{B}) = \int X d\mathbb{P}(./B^\omega) = \mathbb{E}(X/B^\omega)$$

où B^ω *est l'unique élément de la partition contenant* ω .

Proposition 1.2.3 *Soit* (X, Y) *un couple de variables aléatoires intégrables par rapport à* $\mathbb{P}$ *et* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires* $\mathbb{P}$*-intégrables.*

À ω *fixé, l'éspérance conditionnelle vérifie les propriétés suivantes :*

(1) *pour tout couple* $(\lambda, \mu) \in \mathbb{R}^2, \mathbb{E}(\lambda X + \mu Y/\mathcal{B}) = \lambda\mathbb{E}(X/\mathcal{B})(\omega) + \mu\mathbb{E}(Y/\mathcal{B})(\omega)$

(2) *si* $X \leq Y$ *p.s. alors* $\mathbb{E}(X/\mathcal{B})(\omega) \leq \mathbb{E}(Y/\mathcal{B})(\omega)$

(3) *si* $(X_n)_n$ *est une suite croissante de variables aléatoires positives intégrables telle que* $\lim\limits_{n\to+\infty} \uparrow X_n = X$ *alors* $\lim\limits_{n\to+\infty} \uparrow \mathbb{E}(X_n/\mathcal{B})(\omega) = \mathbb{E}(X/\mathcal{B})(\omega)$

DÉMONSTRATION. Ces propriétés se déduisent immédiatement des propriétés de l'éspérance. En effet, à ω fixé $\mathbb{E}(./\mathcal{B})(\omega)$ est égale à l'éspérance par rapport à la mesure de probabilité $\mathbb{P}/B^\omega$ où B^ω est l'unique élément de la partition contenant ω.

Proposition 1.2.4 *Soit X une variable aléatoire $\mathbb{P}$-intégrable.*
(1) *L'application $\omega \mapsto \mathbb{E}(X/\mathcal{B})(\omega)$ satisfait*

$$[\pi_2] : \begin{cases} \mathbb{E}(X/\mathcal{B}) \; est \; \mathcal{B}-mesurable \\ pour\,tout\; B \in \mathcal{B}, on\, a \displaystyle\int_B \mathbb{E}(X/\mathcal{B})d\mathbb{P} = \int_B X d\mathbb{P} \end{cases}$$

En particulier en prenant $B = \Omega$, on obtient $\quad \mathbb{E}(\mathbb{E}(X/\mathcal{B})) = \mathbb{E}(X)$
i.e. les variables aléatoires X et $\mathbb{E}(X/\mathcal{B})$ ont même valeur moyenne.
(2) *Si Y et YX sont $\mathbb{P}$-intégrables et si X est $\mathcal{B}$-mesurable alors* $\mathbb{E}(XY/\mathcal{B}) = X\mathbb{E}(Y/\mathcal{B})$
(3) *Si X est indépendante de la tribu $\mathcal{B}$ alors* $\mathbb{E}(X/\mathcal{B}) = \mathbb{E}(X)$

DÉMONSTRATION.
Preuve de (1). $\mathbb{E}(X/\mathcal{B})$ est $\mathcal{B}$-mesurable car elle est constante sur chaque B_i.
Pour tout $B \in \mathcal{B}$, on démontre l'égalité $\displaystyle\int_B \mathbb{E}(X/\mathcal{B})d\mathbb{P} = \int_B Xd\mathbb{P}$ par la méthode standard.
*É*tape 1 : $X = \mathbb{I}_A$ avec $A \in \mathcal{A}$ le resultat est immédiat car on retrouve la relation $[\pi_1]$. Il suffit d'écrire la définition de $\mathbb{E}(X/\mathcal{B})$ i.e. $\mathbb{E}(\mathbb{I}_A/\mathcal{B}) = \mathbb{P}(A/\mathcal{B})$
*É*tape 2 : $X = \displaystyle\sum_{j \in J} \lambda_j \mathbb{I}_{A_j}$ où J est un sous ensemble fini. Le résultat découle de l'étape précédente en utilisant la linéarité de l'intégrale et de la linéarité de l'éspérance conditionnelle.
*É*tape 3 : X est mesurable positive. Il existe une suite de variables aléatoires étagées positives telle que $lim \uparrow X_n = X$. L'étape précédente assure que pour tout n, on a

$$\int_B \mathbb{E}(X_n/\mathcal{B})d\mathbb{P} = \int_B X_n d\mathbb{P}$$

D'aprés (1) et (2) de la proposition (1.2.3) pour obtenir le resultat.
*É*tape 4 : X est intégrable. On décompose $X = X^+ - X^-$ avec $\mathbb{E}(X^\pm) < \infty$.
L'étape précédente assure que $\int \mathbb{E}(X^\pm/\mathcal{B})d\mathbb{P} = \int_B X^\pm d\mathbb{P}$ pour tout $B \in \mathcal{B}$.

On peut ensuite conclure en utilisant les propriétés de linéarité.
Preuve de (2). On a

$$\mathbb{E}(XY/\mathcal{B}) = \sum_{i\in I} \mathbb{E}(XY/B_i)\mathbb{I}_{B_i} = \sum_{i\in I} \int XY d\mathbb{P}(./B_i)\mathbb{I}_{B_i} = \sum_{i\in I} \frac{1}{\mathbb{P}(B_i)} \int_{B_i} XY d\mathbb{P}\, \mathbb{I}_{B_i}$$

Comme X est $\mathcal{B}$-mesurable donc X s'écrit de la forme $\sum_{i\in I} x_i \mathbb{I}_{B_i}$ avec $I \subset \mathbb{N}$. On obtient

$$\mathbb{E}(XY/\mathcal{B}) = \sum_{i\in I} \frac{1}{\mathbb{P}(B_i)} \int_{B_i} XY d\mathbb{P}\, \mathbb{I}_{B_i} = \sum_{i\in I} \frac{1}{\mathbb{P}(B_i)} \sum_{j\in\mathbb{N}} x_j \int_{B_i} \mathbb{I}_{B_j} Y d\mathbb{P}\, \mathbb{I}_{B_i}$$

Or

$$\mathbb{I}_{B_j\cap B_i} : \begin{cases} \mathbb{I}_{B_j} & si\ i = j \\ 0 & sinon \end{cases}$$

d'o$\grave{u}$

$$\mathbb{E}(XY/\mathcal{B}) = \sum_{i\in I} x_i \frac{1}{\mathbb{P}(B_i)} \int_{B_i} Y d\, \mathbb{P}\mathbb{I}_{B_i} = \sum_{i\in I} x_i \mathbb{E}(Y/B_i)\mathbb{I}_{B_i}$$

D'autre part

$$X\mathbb{E}(Y/\mathcal{B}) = \sum_{i\in I} x_i \mathbb{I}_{B_i} \sum_{j\in I} x_i \mathbb{E}(Y/\mathcal{B}_j)\mathbb{I}_{B_j} = \sum_{i,j\in I} x_i \mathbb{E}(Y/B_j)\mathbb{I}_{B_j\cap B_i} = \sum_{i\in I} x_i \mathbb{E}(Y/B_i)\mathbb{I}_{B_i}$$

d'o$\grave{u}$ l'égalité annoncée.
Preuve de (3). X étant indépendante de $\mathcal{B}$,elle est aussi indépendante des variables aléatoires $\mathbb{I}_{B_j}$. On a donc

$$\mathbb{E}(X/\mathcal{B}) = \sum_{i\in I} \frac{1}{\mathbb{P}(B_i)} \mathbb{E}(X\mathbb{I}_{B_i})\mathbb{I}_{B_i} = \sum_{i\in I} \frac{1}{\mathbb{P}(B_i)} \mathbb{E}(X)\mathbb{E}(\mathbb{I}_{B_i})\mathbb{I}_{B_i} = \mathbb{E}(X) \sum_{i\in I} \mathbb{I}_{B_i} = \mathbb{E}(X)$$

1.2.3 Conditionnement par une variable aléatoire discr$\grave{e}$te :

Soit Y une variable aléatoire , $Y(\Omega) = \{y_i, i \in I \subset \mathbb{N}\}$. On peut appliquer les résultats précédents à la tribu engendrée par Y, i.e. la tribu engendrée par la partition constituée des événements $B_i = Y^{-1}(\{y_i\}), i \in I$. On note $\sigma(Y)$ cette tribu. On définit l'éspérance de X sachant Y par

$$\mathbb{E}(X/\sigma(Y)) = \mathbb{E}(X/Y)$$

et de même la loi de X sachant Y

$$\mathbb{P}_X(./\sigma(Y)) = \mathbb{P}(./Y)$$

En particulier , on a

$$\mathbb{E}(X/Y) = \sum_{i\in\mathbb{N}} \mathbb{E}(X/Y = y_i)\mathbb{I}_{[Y=y_i]} = g(Y)$$

D'aprés $[\pi_2]$,

$$\int_B g(Y)d\mathbb{P} = \int_B Xd\mathbb{P}$$

Pour tout $B \in \sigma(Y)$. Par conséquent pour tout partie C de $\{y_i, i \in I\}$,on a $Y^{-1}(C) \in \sigma(Y)$ et la relation précédent s'écrit

$$\sum_{y\in C} g(y)\mathbb{P}(Y = y) = \int_{Y^{-1}(C)} Xd\mathbb{P}$$

1.2.4 Quelques remarques :

On sait désormais calculer l'éspérance conditionnelle par rapport à une tribu engendrée par une partition dénombrable ou par une variable aléatoire discréte.
Il est important de remarquer que les propriétés $[\pi_2]$ définissent l'éspérance conditionnelle presque s$\hat{u}$rement. En effet,

Proposition 1.2.5 *Si deux variables aléatoires Y_1 et Y_2 vérifient les propriétés $[\pi_2]$ alors elles sont égales $\mathbb{P}$-presque sûrement*

DÉMONSTRATION. On a pour tout $B \in \mathcal{B}$,

$$\int_B (Y_1 - Y_2)d\mathbb{P} = 0 \tag{1.5}$$

Posons $B_1 = \{Y_1 - Y_2 > 0\}$ et $B_2 = \{Y_1 - Y_2 < 0\}$. Ces deux événements sont dans la tribu $\mathcal{B}$ car $Y_1 - Y_2$ est $\mathcal{B}$-mesurable. En appliquant (1.5) aux événement B_1 et B_2, on obtient qu'ils sont $\mathbb{P}$-négligeables. Par conséquent, on a bien $Y_1 - Y_2 = 0$ $\mathbb{P}$-presque s$\hat{u}$rement.
Les conséquences de ce resultat sont les suivantes :

(1) Si on modifie la variable aléatoire $\sum \mathbb{E}(X/B_i)\mathbb{I}_{B_i}$ sur un ensemble $\mathcal{B}$-mesurable et de probabilité nulle alors elle vérifie toujours les propriétés $[\pi_2]$.
(2) Si l'on trouve une variable aléatoire qui vérifie les propriétés $[\pi_2]$ alors elle est presque sûrement égale à $\sum \mathbb{E}(X/B_i)\mathbb{I}_{B_i}$
Dans ce cas on parlera de **version de l'éspérance conditionnelle.**

1.3 Conditionnement par une tribu :

Définition 1.3.1 *Soit $\mathcal{B}$ une tribu et X une variable aléatoire. L'éspérance conditionnelle de X sachant $\mathcal{B}$ est la classe des variables aléatoires $\mathbb{E}(X/\mathcal{B})$ qui vérifient $[\pi_2]$ i.e.*

$$\begin{cases} \mathbb{E}(X/\mathcal{B}) \ est\ \mathcal{B}-mesurable \\ pour\ tout\ B \in \mathcal{B}, on\ a \displaystyle\int_B \mathbb{E}(X/\mathcal{B})d\mathbb{P} = \int_B X d\mathbb{P} \end{cases}$$

On appelle version de l'éspérance conditionnelle un élément de cette classe.

La suite de cette section s'organise de la façon suivante :
(1) On montre que $\mathbb{E}(X/\mathcal{B})$ vérifie les propriétés obtenues dans la section précédente pour l'éspérance conditionnelle (section 1.3.1)
(2) On montre l'existence de $\mathbb{E}(X/\mathcal{B})$ sous certaines hypothéses (section 1.3.2)

1.3.1 Propriétés de $\mathbb{E}(X/\mathcal{B})$:

Soit X une variable aléatoire $\mathbb{P}$-intégrable. On suppose qu'il existe une variable aléatoire $\mathbb{E}(X/\mathcal{B})$ qui satisfait les propriétés $[\pi_2]$ .Sous ces conditions, la variable aléatoire $\mathbb{E}(X/\mathcal{B})$ vérifie les propriétés suivantes :
Propriété A. $\mathbb{E}(./\mathcal{B})$ est linéaire.
Propriété B. Si X est positive alors $\mathbb{E}(X/\mathcal{B})$ l'est aussi. Sinon, il existe $B \in \mathcal{B}$ de probabilité non nulle telle que $\int_B \mathbb{E}(./\mathcal{B})d\mathbb{P} < \infty$. D'où $\int_B X d\mathbb{P} < 0$ ce qui contredit l'hypothése X positive.
Propriété C.[théoréme de Beppo Levi pour l'éspérance conditionnelle] Soit $(X_n)_{n\in\mathbb{N}}$ une suite croissante de variables aléatoires positives qui converge vers X alors

$lim \uparrow \mathbb{E}(X_n/\mathcal{B}) = \mathbb{E}(X/\mathcal{B})$.
Notons d'abord que $(\mathbb{E}(X_n/\mathcal{B}))_n$ est une suite croissante de variables aléatoires positives et $\mathcal{B}$-mesurable .De plus elle est majorée par $\mathbb{E}(./\mathcal{B})$ donc elle converge et sa limite (notée Y) est $\mathcal{B}$-mesurable. De plus, en utilisant Beppo Levi, on a :

$$\int_B Y d\mathbb{P} = lim \int_B \mathbb{E}(X_n/\mathcal{B}) d\mathbb{P} = lim \int_B X_n d\mathbb{P} = \int_B X d\mathbb{P}$$

donc Y vérifie $[\pi_2]$.
Propriété D.[Lemme de fatou pour l'éspérance conditionnelle] Si $(X_n)_{n\in\mathbb{N}}$ est une suite positive, alors

$$\mathbb{E}(\liminf X_n/\mathcal{B}) \leq \liminf \mathbb{E}(X_n/\mathcal{B})$$

Propriété E.[Théoréme de la convergence dominée pour l'éspérance conditionnelle] Si $(X_n)_{n\in\mathbb{N}}$ est une suite qui converge vers X et s'il existe une variable aléatoire Y intégrable telle que pour tout n, $X_n \leq Y$ alors

$$\lim_{n\to\infty} \mathbb{E}(|X_n - X|/\mathcal{B}) = 0$$

Proposition 1.3.1 *Soit (X, Y) un couple de variables aléatoires. Si X est $\mathcal{B}$-mesurable alors $\mathbb{E}(XY/\mathcal{B}) = X\mathbb{E}(Y/\mathcal{B})$ presque sûrement.*

DÉMONSTRATION. $X\mathbb{E}(Y/\mathcal{B})$ est le produit de variables aléatoires $\mathcal{B}$-mesurables donc elle l'est aussi. On montre maintenant par la méthode standard l'égalité : pour tout $B \in \mathcal{B}$

$$\int_B X\mathbb{E}(Y/\mathcal{B}) d\mathbb{P} = \int_B XY d\mathbb{P}$$

Fixons $B \in \mathcal{B}$.
Étape (1) : $X = \mathbb{I}_A$ avec $A \in \mathcal{B}$ on a

$$\int_B \mathbb{I}_A \mathbb{E}(Y/\mathcal{B}) d\mathbb{P} = \int_{A\cap B} \mathbb{E}(Y/\mathcal{B}) d\mathbb{P} = \int_{A\cap B} Y d\mathbb{P} \quad \text{car } A \cap B \in \mathcal{B}$$

$$= \int_B \mathbb{I}_A Y d\mathbb{P}$$

Étape (2) : $X = \sum_{i=1}^{p} \lambda_i \mathbb{I}_{A_i}$. Le résultat découle directement de la linéarité de l'intégrale.

Étape (3) : X est mesurable positive. Il existe une suite croissante de variables aléatoires étagées positives $(X_n)_{n\in\mathbb{N}}$ telle que $lim \uparrow X_n = X$. L'étape précédante assure que pour tout n, on a

$$\int_B X_n \mathbb{E}(Y/\mathcal{B})d\mathbb{P} = \int_B X_n Y d\mathbb{P} \tag{1.6}$$

$\diamond$ La variable aléatoire Y est positive. La suite $(YX_n)_n$ est donc croissante positive et converge vers XY. On peut donc appliquer le théoréme de Beppo Levi aux deux membres de (1.6) et on obtient

$$\int_B X\mathbb{E}(Y/\mathcal{B})d\mathbb{P} = \int_B XY d\mathbb{P}$$

$\diamond$ Y quelconque. On applique le resultat précédente à Y^+ et Y^- i.e.

$$\int_B X\mathbb{E}(Y^{\pm}/\mathcal{B})d\mathbb{P} = \int_B XY^{\pm} d\mathbb{P}$$

En suite, le resultat découle de la linéarité de l'intégrale et de l'espérance conditionnelle.
*É*tape (4) : X est intégrable. On applique l'étape précédente à X^+ et X^- et on utilise à nouveau les propriétés de linéarité.

Proposition 1.3.2 *Si X est indépendante de $\mathcal{B}$ alors $\mathbb{E}(X/\mathcal{B}) = \mathbb{E}(X)$ presque sûrement.*

D*É*MONSTRATION. La variable aléatoire constante et égale à $\mathbb{E}(X)$ est $\mathcal{B}$-mesurable. De plus, pour tout $B \in \mathcal{B}$, on a

$$\begin{aligned}\int_B \mathbb{E}(X)d\mathbb{P} = \mathbb{E}(X)\int_B d\mathbb{P} &= \int Xd\mathbb{P}\int \mathbb{I}_B d\mathbb{P} \\ &= \textstyle\int X\mathbb{I}_B d\mathbb{P} \qquad \text{car } X \amalg \mathbb{I}_B \\ &= \int_B Xd\mathbb{P}\end{aligned}$$

Proposition 1.3.3 *Soit $\mathcal{B}$ et $\grave{\mathcal{B}}$ deux tribus telles que $\mathcal{B} \subset \grave{\mathcal{B}}$, on a*

$$\mathbb{E}(X/\mathcal{B}) = \mathbb{E}(\mathbb{E}(X/\grave{\mathcal{B}})/\mathcal{B}) = \mathbb{E}(\mathbb{E}(X/\mathcal{B})/\grave{\mathcal{B}}) \tag{1.7}$$

D*É*MONSTRATION. Rappelons que la condition $\mathcal{B} \subset \grave{\mathcal{B}}$ implique que si une variable aléatoire est $\mathcal{B}$-mesurable alors elle est aussi $\grave{\mathcal{B}}$-mesurable. Ce résultat implique que $\mathbb{E}(X/\mathcal{B})$

est $\grave{\mathcal{B}}$-mesurable et donc d'aprés la proposition (1.3.1) on a $\mathbb{E}(X/\mathcal{B})\mathbb{E}(\mathbb{E}(X/\mathcal{B})/\grave{\mathcal{B}})$.

Montrons maintenant que $\mathbb{E}(X/\mathcal{B}) = \mathbb{E}(\mathbb{E}(X/\grave{\mathcal{B}})/\mathcal{B})$. La $\mathcal{B}$-mesurabilité de $\mathbb{E}(X/\mathcal{B})$ est acquise, de plus pour tout $B \in \mathcal{B}$

$$\begin{aligned}\int_B \mathbb{E}(X/\grave{\mathcal{B}})d\mathbb{P} &= \int_B Xd\mathbb{P} \qquad \text{car } B \in \grave{\mathcal{B}} \supset \mathcal{B} \\ &= \int_B \mathbb{E}(X/\mathcal{B})d\mathbb{P} \quad \text{car } B \in \mathcal{B}\end{aligned}$$

1.3.2 Existence de $\mathbb{E}(X/\mathcal{B})$ et intrprétation géométrique :

Dans cette partie, nous montrons que si X est de carré intégrable $i.e.X \in \mathbf{L}^2(\Omega, \mathcal{A}, \mathbb{P})$, la meilleure approximation de X par un élément de $\in \mathbf{L}^2(\Omega, \mathcal{B}, \mathbb{P})$ est une version de l'éspérance conditionnele.

Théorème 1.3.1 *Soit H un espace de Hilbert (i.e. un espace normé complet dont la norme provient d'un produit scalaire). Soit E un sous espace complet de H pour tout $x \in H$, il esxiste un unique élément $y \in E$ tel que :*

$$\|y - x\| = \inf_{\omega \in E} \|\omega - x\| \tag{1.8}$$

$$y - x \perp E \tag{1.9}$$

On note $y = P_E(x)$. C'est la projection orthogonal sur E.

Rappelons que $\mathbf{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de Hilbert. Soit $\mathcal{B}$ une sous tribu de $\mathcal{A}$ et $\mathbf{L}^2(\mathcal{B}) = \mathbf{L}^2(\Omega, \mathcal{B}, \mathbb{P})$ le sous espace $\mathbf{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ constitué des classe d'équivalence d'applications $\mathcal{B}$-mesurables. C'est un sous espace complet de $\mathbf{L}^2(\Omega, \mathcal{A}, \mathbb{P})$. Les hypothéses du théorème (1.3.1) *sont donc satisfaites. Posons $P_{\mathbf{L}^2(\mathcal{B})}$ la projection orthogonale sur $\mathbf{L}^2(\mathcal{B})$.*

Théorème 1.3.2 *Dans le contexte décrit ci-dessus, si X une variable aléatoire dans $\mathbf{L}^2(\Omega, \mathcal{A}, \mathbb{P})$ alors la variable aléatoire $Y = P_{\mathbf{L}^2(\mathcal{B})}(X)$ vérifie la condition $[\pi_2]$ (c'est une version de l'éspérance conditionnelle de X sachant $\mathcal{B}$).*

DÉMONSTRATION. Montrons que la variable aléatoire Y vérifie la condition $[\pi_2]$. Par construction,la variable aléatoire $Y = P_{\mathbf{L}^2(\mathcal{B})}(X)$ est $\mathcal{B}$-mesurable. Comme la variable $Y - X$ est orthogonale à $\mathcal{B}$, on a pour tout $B \in \mathcal{B}$, les variables aléatoires $\mathbb{I}_B$ et $Y - X$ sont orthogonales i.e.

$$\mathbb{E}((X - Y)\mathbb{I}_B) = 0$$

Ceci implique que

$$\mathbb{E}(Y\mathbb{I}_B) = \mathbb{E}(X\mathbb{I}_B)$$

ou encore

$$\int_B Y d\mathbb{P} = \int_B X d\mathbb{P}$$

Par conséquent, la variable aléatoire Y vérifie $[\pi_2]$.

Remarque 1.3.1 *L'existence reste vrai dans* $\mathbf{L}^1$. *On admet ici ce resultat. Il est important de remarquer que l'interprétation en terme de projection n'est plus valide.*

1.3.3 Conditionnement par une variable aléatoire

. Soit Y une variable aléatoire à valeur dans l'espace E (muni de la tribu ξ). On note $\sigma(Y)$ la tribu engendrée par Y. Pour simplifier les notations, on écrit $\mathbb{E}(X/Y)$ au lieu de $\mathbb{E}(X/\sigma(Y))$.

Lemme 1.3.1 *(Lemme de Doob). Soit U une variable aléatoire sur Ω. U est $\sigma(Y)$-mesurable si et seulement si'il existe une application mesurable g de E dans $\mathbb{R}$ telle que $U = g(Y)$*

L'éspérance conditionnelle $\mathbb{E}(X/Y)$ est donc de la forme $g(Y)$ avec g mesurable et

$$\int_B g(Y) d\mathbb{P} = \int_B X d\mathbb{P} \qquad \forall B \in \sigma(Y)$$

L'évenement B peut aussi s'exprimer de la forme $Y^{-1}(C)$ avec $C \in \Im$. En utilisant le théorème du transfére, on obtient

$$[\pi_3] : \int_C g(y) d\mathbb{P}_Y(y) = \int_{Y^{-1}(C)} X d\mathbb{P}$$

Remarque 1.3.2 *Par abus de langage, on utilise aussi la notation* $g(y) = \mathbb{E}(X/Y = y)$ *qui est différente de* $g(y) = \mathbb{E}(\{X/Y = y\})$

Remarque 1.3.3 *Rappel : une* π*-classe est une collection de parties de* E *contenant* E *et stable par intersection fini. La propriété* $[\pi_1]$ $(resp.[\pi_2])$ *est vraie si et seulement si elle est vérifiée pour tout* $C \in \mathcal{C}$ *où* $\mathcal{C}$ *est une* π*-classe qui engendre* $\mathcal{B}$*. On admet ce resultat.*
Rappel : L'image réciproque d'une π*-classe est aussi une* π*-classe.*
Par conséquent $[\pi_3]$ *est vraie si et seulement si elle est vérifiée sur une* π*-classe qui engendre* $\Im$ *.*
Dans ce contexte, les propriétés de l'éspérance conditionnelle s'écrivent
1) *Si* X *et* Y *sont indépendantes alors* $\mathbb{E}(X/Y) = \mathbb{E}(X)$
2) *Pour toute aplication* h *de* E *dans* $\mathbb{R}$ *mesurable,*$\mathbb{E}(h(Y)Z/Y) = h(Y)\mathbb{E}(Z/Y)$
3) *Si* $Z = h(Y)$ *alors* $\mathbb{E}(X/Z) = \mathbb{E}(\mathbb{E}(X/Y)/Z) = \mathbb{E}(\mathbb{E}(X/Z)/Y)$