A) la tribu engendrée par le système complet (B_n)

L'ensemble $\mathcal{B}$ de tous les événements B qui s'écrivent

$$B = \bigcup_{i \in I} B_i \quad \circledast$$

(réunion d'un certain nombre d'événements B_n) où $I \subset \{1, 2, \dots, N\}$ éventuellement $I = \emptyset$ et dans ce cas par convention $\bigcup_{i \in \emptyset} B_i = \emptyset$,

Cet ensemble $\mathcal{B}$ est une tribu on l'appelle tribu engendrée par le système complet B_n

la vérification que $\mathcal{B}$ est une tribu est laissée en exercice.

De telles tribus sont importantes pour les applications. On les appelle tribus atomiques car les B_i $(i = 1, \dots, N)$ sont parfois appelés les atomes de $\mathcal{B}$.

Remarque : le système complet (B_n) peut être dénombrable. On définit la tribu $\mathcal{B}$ engendrée par les $(B_n)_{n \in \mathbb{N}^*}$ de la même façon que ci-dessus (avec $I \subset \mathbb{N}^*$ et la

réunion $\bigcup_{i \in I} B_i$ étant dénombrable si I est un sous-ensemble dénombrable). Le résultat suivant est très utile :

Théorème : Toute variable aléatoire $X : \Omega \to \mathbb{R}$ $\mathcal{B}$-mesurable (i.e. $\forall x \in \mathbb{R}$, $[X \leq x] \in \mathcal{B}$) est de la forme

$$\circledast \quad X = \sum_{i=1}^{N} x_n \mathbb{1}_{B_n} \; ,$$

où $x_1, \dots, x_n \in \mathbb{R}$ et réciproquement toute v.a. de cette forme est $\mathcal{B}$-mesurable.

démⁿ : la réciproque est facile à démontrer : supposons X de la forme $\circledast$ et soit $x \in \mathbb{R}$ alors $[X \leq x] = \{\omega \in \Omega \, ; \, X(\omega) \leq x\}$

$$= \bigcup_{k \in I_x} B_k$$

où $I_x = \{n \in \{1, \dots, N\} \, ; \, x_n \leq x\}$

Donc $[X \leq x] \in \mathcal{B}$ (comme réunion de certains atomes B_n).

La démonstration de la partie directe n'est pas plus difficile : Supposons X $\mathcal{B}$-mesurable. Alors sur chaque B_k, X est constante. En effet par l'absurde supposons X non constante sur B_k. Il existerait alors $x \in \mathbb{R}$ tel que

$$\mathbb{P}(B_k \cap [X > x]) > 0 \quad (1)$$

et $\mathbb{P}(B_k \cap [X < x]) > 0$ (2)

Mais $[X > x] \in \mathcal{B}$ et $[X < x] \in \mathcal{B}$ donc

$$[X > x] = \bigcup_{i \in I} B_i \quad \text{et} \quad [X < x] = \bigcup_{i \in J} B_i$$

avec $I \cap J = \emptyset$. Or d'après (1) $k \in I$ et d'après (2) $k \in J$ ce qui est absurde.

Donc X est constante sur B_k : il existe $x_k \in \mathbb{R}$ tel que $\forall \omega \in B_k$, $X(\omega) = x_k$.

Comme les B_k forment une partition de Ω, on en déduit que

$$X = \sum_{k=1}^{N} x_k \mathbb{1}_{B_k} \quad \underline{\text{CQFD}}.$$

B) <u>La meilleure approximation en moyenne quadratique par une variable aléatoire $\mathcal{B}$-mesurable</u> :

Supposons que la v.a. X a un <u>moment d'ordre 2</u> i.e. $X \in L^2$. On rappelle que L^2 est un espace de Hilbert réel pour

le produit scalaire

$$\langle X, Y \rangle = \mathbb{E}(XY)$$

et la norme associée :

$$\|X\|_2 = \sqrt{\langle X, X \rangle} = \left(\mathbb{E}(X^2)\right)^{1/2}$$

la distance en moyenne quadratique des v.a. X et $Y \in L^2$ est ainsi définie par :

$$\|X - Y\|_2 = \sqrt{E((X-Y)^2)}$$

<u>Théorème</u> : la v.a. $Y = \sum_{n=1}^{N} \mathbb{E}(X|B_n) \mathbb{1}_{B_n}$ est la <u>meilleure approximation</u> en moyenne quadratique de la v.a. X par une v.a. $\mathcal{B}$-mesurable

<u>dém^{n}</u> : D'après le théorème précédent, les variables aléatoires $\mathcal{B}$-mesurables forment un sous espace vectoriel V de L^2. Cet espace V a une base évidente donnée par les v.a. $\mathbb{1}_{B_n}$ $(n = 1, 2, \dots, N)$. De plus ces v.a. forment une <u>base orthogonale</u> de V car

$$\forall n \neq m, \ \langle \mathbb{1}_{B_n}, \mathbb{1}_{B_m} \rangle = \mathbb{E}(\mathbb{1}_{B_n} \mathbb{1}_{B_m}) = \mathbb{E}(\mathbb{1}_{B_n \cap B_m}) = \mathbb{E}(\mathbb{1}_{\emptyset}) = 0.$$

Les v.a. $E_n = \frac{1}{\sqrt{\mathbb{P}(B_n)}} \mathbb{1}_{B_n}$ constituent une base orthonormale de V. D'après le Théorème de projection (voir le cours de L3 sur les espaces de Hilbert) la meilleure approximation de X au sens de la norme $\|\cdot\|_2$ est la projection orthogonale $P_V(X)$ de X sur V. Elle est donnée par :

$$P_V(X) = \sum_{n=1}^{N} \langle X, E_n \rangle E_n$$

$$= \sum_{n=1}^{N} \frac{1}{\mathbb{P}(B_n)} \langle X, \mathbb{1}_{B_n} \rangle \mathbb{1}_{B_n}$$

Mais $\frac{1}{\mathbb{P}(B_n)} \langle X, \mathbb{1}_{B_n} \rangle = \frac{1}{\mathbb{P}(B_n)} \mathbb{E}(X \mathbb{1}_{B_n})$

$$= \mathbb{E}(X / B_n)$$

(voir feuille de TD n° 6). CQFD.

Notation : on note $Y = \mathbb{E}^{\mathcal{B}}(X)$ et on l'appelle l'espérance conditionnelle de X sachant la sous-tribu $\mathcal{B}$.

Signification pratique du résultat : Considérons le problème concret suivant : Soit

X une variable aléatoire relative à une certaine expérience aléatoire modélisée par un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. Pour connaître la valeur $X(\omega)$ prise par X il faut attendre que l'expérience ait pris fin et qu'on connaisse son résultat c'est à dire l'éventualité $\omega \in \Omega$ qui s'est effectivement réalisée.

Supposons qu'on ne connaisse pas le résultat définitif mais seulement un résultat partiel de la forme :

« l'un des événements B_n $(n = 1, \dots, N)$ s'est réalisé ».

La connaissance de X s'est un peu précisée mais elle reste aléatoire : on est amené, d'après ce qui précède, à considérer que la meilleure approximation de X "sachant l'information $\mathcal{B}$" est $\mathbb{E}^{\mathcal{B}}(X)$.

Remarque : Il est important de noter que $\mathbb{E}^{\mathcal{B}}(X)$ est une v.a. et non pas un nombre !

(4) Etude générale de l'espérance conditionnelle pour les variables aléatoires de L^2.

On considère l'espace L^2 des v.a. définies sur l'espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ et ayant un moment d'ordre 2, muni du produit scalaire

$$\langle X, Y\rangle = \mathbb{E}(XY) = \int_\Omega X \cdot Y \, d\mathbb{P}$$

C'est un espace de Hilbert si on considère comme égales deux v.a. qui sont égales $\mathbb{P}$-presque sûrement.

Soit $\mathcal{B} \subset \mathcal{F}$ une sous-tribu de $\mathcal{F}$. On considère l'espace $L^2(\mathcal{B}) := L^2(\Omega, \mathcal{B}, \mathbb{P})$ des v.a. $\mathcal{B}$-mesurables et de carré intégrable. $L^2(\mathcal{B})$ est un sous-espace vectoriel de L^2.

<u>Proposition</u>: $L^2(\mathcal{B})$ est un s.e.v. <u>fermé</u> de L^2

<u>démⁿ</u>: $L^2(\mathcal{B})$ est complet donc fermé <u>CQFD</u>

<u>Notation</u> on note $\mathbb{E}^{\mathcal{B}}$ l'opérateur $P_{L^2(\mathcal{B})}$ de projection orthogonale de L^2 sur $L^2(\mathcal{B})$.

Comme tout opérateur de projection, <u>la propriété caractéristique de $\mathbb{E}^{\mathcal{B}}$</u> est la suivante:

$\mathbb{E}^{\mathcal{B}}(X)$ est <u>l'unique élément</u> de $L^2(\mathcal{B})$ tel que:

$$\forall X \in L^2, \quad X - \mathbb{E}^{\mathcal{B}}(X) \perp L^2(\mathcal{B})$$

Autrement dit:

$$\forall X \in L^2, \forall Z \in L^2(\mathcal{B}), \langle X - \mathbb{E}^{\mathcal{B}}(X), Z\rangle = 0$$

compte tenu de la forme du produit scalaire ceci peut encore s'écrire

$$\forall X \in L^2, \forall Z \in L^2(\mathcal{B}), \mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}^{\mathcal{B}}(X) Z)$$

On peut montrer facilement (voir en exercice) que dans la condition précédente on peut se restreindre aux v.a. Z de la forme $\mathbb{1}_B$ avec $B \in \mathcal{B}$. Ce qui donne finalement la condition simplifiée:

<u>Propriété caractéristique de $\mathbb{E}^{\mathcal{B}}$</u>: Pour toute $X \in L^2$, $\mathbb{E}^{\mathcal{B}}(X)$ est l'unique v.a. de $L^2(\mathcal{B})$ telle que

$$\int_B X \, d\mathbb{P} = \int_B \mathbb{E}^{\mathcal{B}}(X) \, d\mathbb{P}.$$

(A) Propriétés générales de l'espérance conditionnelle des v.a. de L^2:

On va rassembler dans un théorème les principales propriétés de l'opérateur de projection $\mathbb{E}^{\mathcal{B}}$ sous une forme adaptée aux probabilités.

Théorème : Soit $X \in L^2$. Alors :

1) $\mathbb{E}^{\mathcal{B}}(X) \in L^2(\mathcal{B})$ est la projection orthogonale de X sur $L^2(\mathcal{B})$ (en fait c'est la définition)

2) $\|\mathbb{E}^{\mathcal{B}}(X)\|_2 \leq \|X\|_2$

3) Si $\mathcal{C} \subset \mathcal{B} \subset \mathcal{F}$ sont des sous-tribus alors
$$\mathbb{E}^{\mathcal{C}}(\mathbb{E}^{\mathcal{B}}(X)) = \mathbb{E}^{\mathcal{C}}(X) \text{ (\underline{thm des 3} } \perp\text{)}.$$

4) $\mathbb{E}(\mathbb{E}^{\mathcal{B}}(X)) = \mathbb{E}(X)$ (formule de l'espérance totale)

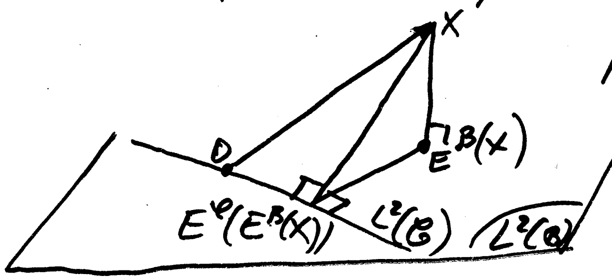
5) $\forall Z \in L^\infty(\mathcal{B})$, $\mathbb{E}^{\mathcal{B}}(Z.X) = Z\,\mathbb{E}^{\mathcal{B}}(X)$

__Dém__ : 1) rien à démontrer

2) la projection diminue la norme (c'est une propriété générale)

3) C'est la propriété dite des __trois perpendiculaires__

$L^2(\mathcal{C}) \subset L^2(\mathcal{B}) \subset L^2$ Voir la figure :
s.e.v fermé   s.e.v fermé



c'est encore une propriété générale des opérateurs de projection

4) Considérons $\mathcal{C} = \{\phi, \Omega\}$ la tribu grossière

alors Y est $\mathcal{C}$-mesurable $\Leftrightarrow$ $Y = c^{te}$.

Donc $\mathbb{E}^{\mathcal{C}}(X) = C$ ($= c^{te}$) et d'après la propriété caractéristique p. 18 en prenant $B = \Omega$, on obtient :
$$\int_\Omega X\,d\mathbb{P} = \int_\Omega \mathbb{E}^{\mathcal{C}}(X)\,d\mathbb{P} = \int_\Omega C\,d\mathbb{P} = C.$$

Donc $\mathbb{E}^{\mathcal{C}}(X) = \mathbb{E}(X)$ et 4) est prouvé.

5) Cette propriété fait appel aux raisonnements introduits dans le cours d'espace de Hilbert (voir L3)

Soit M_Z l'opérateur de multiplication par Z sur l'espace L^2 :
$$\forall X \in L^2,\ M_Z(X) = Z.X$$

Comme Z est borné, il est clair que
$$M_Z(L^2) \subset L^2$$
et que M_Z est un opérateur linéaire de L^2 dans lui même. De plus
$$\forall Y \in L^2,\ \mathbb{E}((ZX).Y) = \mathbb{E}(X.(ZY)) \text{ (évident)}$$
i.e. $\forall Y \in L^2$, $\langle M_Z(X), Y\rangle = \langle X, M_Z(Y)\rangle$.

donc M_Z est un __opérateur autoadjoint__.

Or $M_Z(L^2(\mathcal{B})) \subset L^2(\mathcal{B})$

i.e. M_Z stabilise le s.e.v $L^2(\mathcal{B})$ donc (puisqu'il est autoadjoint) il stabilise aussi son orthogonal $L^2(\mathcal{B})^\perp$. Pour $X \in L^2$, on a sa décomposition

$$X = \mathbb{E}^{\mathcal{B}}(X) + X_2$$

avec $X_2 = X - \mathbb{E}^{\mathcal{B}}(X) \in L^2(\mathcal{B})^\perp$. Alors:

$$M_Z(X) = \underbrace{Z\,\mathbb{E}^{\mathcal{B}}(X)}_{\in L^2(\mathcal{B})} + \underbrace{Z X_2}_{\in L^2(\mathcal{B})^\perp}$$

$$\Rightarrow Z \cdot \mathbb{E}^{\mathcal{B}}(X) = \mathbb{E}^{\mathcal{B}}(ZX) \qquad \text{CQFD}.$$

B) Positivité et espérance conditionnelle

Soit $X \in L^2$ une variable aléatoire avec moment d'ordre 2.

$$\boxed{\text{Théorème : si } X \geq 0 \text{ alors } \mathbb{E}^{\mathcal{B}}(X) \geq 0 \text{ p.s.}}$$

dém^n : $\forall B \in \mathcal{B}$; par la propriété caractéristique, on a

$$\mathbb{E}(\mathbb{E}^{\mathcal{B}}(X)\mathbb{1}_B) = \mathbb{E}(X\mathbb{1}_B) \geq 0$$

Posons $\mathbb{E}^{\mathcal{B}}(X) = Y$. Ainsi :

$$\forall B \in \mathcal{B}, \quad \mathbb{E}(Y\mathbb{1}_B) \geq 0 \,. \quad (*)$$

Soit $B_n = \{\omega \in \Omega \,;\, Y(\omega) < -\frac{1}{n}\}$. On a $B_n \in \mathcal{B}$ car Y est $\mathcal{B}$-mesurable et

$$\mathbb{E}(Y\mathbb{1}_{B_n}) \geq 0 \text{ d'après } (*)$$

Mais on a aussi

$$\mathbb{E}(Y\mathbb{1}_{B_n}) = \int_{B_n} Y\,dP \leq -\frac{1}{n}\int_{B_n} 1\,d\mathbb{P}$$

donc $\mathbb{E}(Y\mathbb{1}_{B_n}) \leq -\frac{1}{n}\mathbb{P}(B_n)$, ce qui exige $\mathbb{P}(B_n) = 0$ (sinon contradiction avec $(*)$)

Ainsi $\mathbb{P}([Y<0]) = \mathbb{P}(\bigcup_n B_n) \leq \sum \mathbb{P}(B_n) = 0$

ce qui prouve que $Y \geq 0$ p.s. CQFD.

Corollaire : Soient X, X_1 et $X_2 \in L^2$. On a

1) $X_1 \leq X_2 \Rightarrow \mathbb{E}^{\mathcal{B}}(X_1) \leq \mathbb{E}^{\mathcal{B}}(X_2)$ p.s.

2) $|\mathbb{E}^{\mathcal{B}}(X)| \leq \mathbb{E}^{\mathcal{B}}(|X|)$

dém^n : 1) $0 \leq X_2 - X_1 \Rightarrow \mathbb{E}^{\mathcal{B}}(X_2 - X_1) \geq 0$ p.s.
d'où $\mathbb{E}^{\mathcal{B}}(X_1) \leq \mathbb{E}^{\mathcal{B}}(X_2)$ (linéarité de $\mathbb{E}^{\mathcal{B}}$).

2) $-|X| \leq X \leq |X| \Rightarrow -\mathbb{E}^{\mathcal{B}}(|X|) \leq \mathbb{E}^{\mathcal{B}}(X) \leq \mathbb{E}^{\mathcal{B}}(|X|)$
d'où $|\mathbb{E}^{\mathcal{B}}(X)| \leq \mathbb{E}^{\mathcal{B}}(|X|)$.

(5) Extension de la notion d'espérance conditionnelle aux variables aléatoires de L^1

On considère ici l'espace $L^1 = L^1(\Omega, \mathcal{F}, \mathbb{P})$ des variables aléatoires X ayant un moment d'ordre 1. C'est un espace de Banach quand on le munit de la norme

$$\|X\|_1 = \int_\Omega |X| \, d\mathbb{P}$$

(<u>attention</u> il n'y a pas de produit scalaire sur L^1 !).

Notons que L^2 est un sous-espace vectoriel de L^1. En effet si $X \in L^2$, on a :

$$\int_\Omega |X| \, d\mathbb{P} = \int_\Omega |X| \cdot 1 \, d\mathbb{P} = \langle |X|, 1 \rangle_{L^2} \leq \sqrt{\mathbb{E}(|X|^2)} \cdot 1 < +\infty$$

(inég. de Cauchy-Schwarz)

donc $X \in L^1$.

Le résultat suivant est bien connu : c'est le point clé de l'extension à L^1 de la notion d'espérance conditionnelle :

<u>Proposition</u> : L^2 est dense dans L^1 (pour la norme $\|\cdot\|_1$)

<u>dém^n</u> : Soit $Y \in L^1$. Pour tout n, soit

$$A_n = [|Y| \leq n] = \{\omega \in \Omega ; |Y(\omega)| \leq n\}$$

Posons $Y_n = Y \cdot \mathbb{1}_{A_n}$. On a $Y_n \in L^2$ (c'est clair car Y_n est bornée par n). De plus

1) $\lim\limits_{n\to\infty} Y_n(\omega) = Y(\omega)$ (car $A_n \nearrow \Omega$)

2) $\forall n, \ |Y_n| \leq |Y| \in L^1$

Le théorème de convergence dominée implique donc $\int_\Omega |Y_n - Y| \, d\mathbb{P} \to 0$ i.e. $\|Y_n - Y\|_1 \to 0$

CQFD.

<u>Théorème</u> : Soit $\mathcal{B}$ une sous-tribu de $\mathcal{F}$. L'opérateur $\mathbb{E}^{\mathcal{B}}$ défini sur L^2 possède un prolongement continu $\widetilde{\mathbb{E}}^{\mathcal{B}}$ défini sur L^1 et à valeurs dans L^1 et vérifiant les propriétés :

1) $\forall X \in L^1(\mathcal{B}), \ X = \widetilde{\mathbb{E}}^{\mathcal{B}}(X)$

2) $\|\widetilde{\mathbb{E}}^{\mathcal{B}}(X)\|_1 \leq \|X\|_1$

3) si $\mathcal{C} \subset \mathcal{B} \subset \mathcal{F}$ sont des sous-tribus,

$\forall X \in L^1$, $\tilde{E}^{\mathcal{C}}(\tilde{E}^{\mathcal{B}}(X)) = \tilde{E}^{\mathcal{C}}(X)$

4) $E(\tilde{E}^{\mathcal{B}}(X)) = E(X)$

5) $\forall Z \in L^\infty(\mathcal{B})$, $\tilde{E}^{\mathcal{B}}(ZX) = Z\tilde{E}^{\mathcal{B}}(X)$

6) (pour mémoire car $\tilde{E}^{\mathcal{B}}$ prolonge $E^{\mathcal{B}}$)

$\forall X \in L^2(\mathcal{B})$, $\tilde{E}^{\mathcal{B}}(X) = E^{\mathcal{B}}(X)$

<u>dém^n</u> : $\forall X \in L^2$, $|E^{\mathcal{B}}(X)| \leq E^{\mathcal{B}}(|X|)$ p.s.

(voir corollaire p.22). Ceci implique

$$E(|E^{\mathcal{B}}(X)|) \leq E\,E^{\mathcal{B}}(|X|) = E(|X|)$$

$$\Rightarrow \|E^{\mathcal{B}}(X)\|_1 \leq \|X\|_1 \quad (\text{ceci } \underline{\forall X \in L^2}) \; ⊛$$

Mais L^2 est dense dans L^1 et $E^{\mathcal{B}}$ est un opérateur continu sur L^2 d'après ⊛ donc $E^{\mathcal{B}}$ se prolonge <u>de manière unique</u> en un opérateur $\tilde{E}^{\mathcal{B}} : L^1 \to L^1$ et l'égalité $E^{\mathcal{B}}(X) = X$ $(\forall X \in L^2(\mathcal{B}))$ prolonge donc en

$$\forall X \in L^1(\mathcal{B}), \quad \tilde{E}^{\mathcal{B}}(X) = X$$

Toutes les autres propriétés se prolongent aussi par continuité <u>CQFD</u>.

<u>Notation</u> : pour ne pas surcharger les notations nous noterons $\tilde{E}^{\mathcal{B}} = E^{\mathcal{B}}$ l'espérance conditionnelle sur L^1.

<u>Remarque</u> : Pour $X \in L^2$, $E^{\mathcal{B}}(X)$ est la v.a. de $L^2(\mathcal{B})$ qui est la meilleure approximation de X au sens de la norme $\|\cdot\|_2$. Dans le cas où $X \in L^1$, $E^{\mathcal{B}}(X)$ n'a pas la même interprétation néanmoins les propriétés formelles de l'opération $E^{\mathcal{B}}$ sur L^1 sont analogues à celle de $E^{\mathcal{B}}$ sur L^2 comme on va le voir dans le paragraphe qui suit :

(6) <u>Propriétés de l'espérance conditionnelle des variables aléatoires de L^1</u> :

<u>Théorème</u> (Approximation croissante de $E^{\mathcal{B}}(X)$) :

Soit (X_n) et $X \in L^1$ des v.a. telles que

$X_n \nearrow X$ p.s. $(n \to +\infty)$

(i.e. $X_n \leq X_{n+1}$ et $X_n \to X$ p.s). <u>Alors</u>

$$E^{\mathcal{B}}(X_n) \nearrow E^{\mathcal{B}}(X) \text{ p.s. } (n \to +\infty)$$

<u>dém^n</u> : Posons $Y_n = X - X_n$ alors $Y_n \searrow 0$ p.s.

Mais $Y_n \geq Y_{n+1} \Rightarrow \mathbb{E}^{\mathcal{B}}(Y_n) \geq \mathbb{E}^{\mathcal{B}}(Y_{n+1})$

Donc $\mathbb{E}^{\mathcal{B}}(Y_n)$ converge en décroissant vers une v.a. $Z \geqslant 0$. Mais le théorème de convergence monotone implique alors

$$\lim_{n\to\infty} \mathbb{E}(\mathbb{E}^{\mathcal{B}}(Y_n)) = \mathbb{E}(Z)$$

Or $\mathbb{E}(\mathbb{E}^{\mathcal{B}}(Y_n)) = \mathbb{E}(Y_n) \searrow 0$ donc $\mathbb{E}(Z)=0$

Comme $Z \geq 0$ ceci implique $Z = 0$ p.s.

On a donc montré que

$$\mathbb{E}^{\mathcal{B}}(X - X_n) \searrow 0 \quad (n \to +\infty) \quad \underline{\text{CQFD}}.$$

<u>Corollaire 1</u> (propriété caractéristique de $\mathbb{E}^{\mathcal{B}}(X)$)

Pour tout $X \in L^1$, $\mathbb{E}^{\mathcal{B}}(X)$ est l'unique élément de $L^1(\mathcal{B})$ tel que :

$$\forall B \in \mathcal{B}, \quad \int_B \mathbb{E}^{\mathcal{B}}(X)\, d\mathbb{P} = \int_B X\, d\mathbb{P} \quad \circledast$$

<u>démo</u>: on utilise le cas où $X \in L^2$ puis on démontre $\circledast$ pour les v.a. $X \in L^1$ avec $X \geq 0$ par approximation croissante (exercice).

<u>Corollaire 2</u>: $\forall X \in L^2$ et $\forall Z \in L^2$,

$$\mathbb{E}^{\mathcal{B}}(Z \cdot X) = Z \cdot \mathbb{E}^{\mathcal{B}}(X)$$