# Espérance conditionnelle, cas général

rappel : si $X, Y, Z$ sont des variables aléatoires discrètes et $L^2$, de moyenne nulle

si $Z$ est mesurable $/\sigma(Y)$, alors [en notant $V_Y$ l'espace vectoriel des v.a $L^2$, de moyenne nulle, mesurables $/\sigma(Y)$]

$$\mathbb{E}(ZX) = \mathbb{E}(Z\,\underbrace{\mathbb{E}(X|Y)})$$

$$\phi(Z,X) = \phi(Z, \pi_{V_Y}(X)) \quad \text{vrai si } Z \in V_Y$$

⇝ on utilise cette propriété pour définir l'espérance conditionnelle dans le cas général

Th et déf : $(\Omega, \mathcal{A}, P)$ espace de probabilité. $\mathcal{B}$ sous-tribu de $\mathcal{A}$

$X \in L^1(\Omega, \mathcal{A}, P)$, alors il existe une unique variable aléatoire dans $L^1(\Omega, \mathcal{B}, P)$, qu'on note $\mathbb{E}(X|\mathcal{B})$, telle que $\forall B \in \mathcal{B}$,

$$\mathbb{E}(X \mathbb{1}_B) = \mathbb{E}(\mathbb{E}(X|\mathcal{B})\mathbb{1}_B) \qquad (1)$$

plus généralement, pour toute v.a. $Z$ bornée, mesurable $/\mathcal{B}$,

$$\mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(X|\mathcal{B})Z) \qquad (2)$$

dém unicité : supposons que $U, U'$ variables aléatoires mesurables $/\mathcal{B}$, et $\forall B \subset \mathcal{B}$, $\mathbb{E}(X\mathbb{1}_B) = \mathbb{E}(U\mathbb{1}_B) = \mathbb{E}(U'\mathbb{1}_B)$

prenons $B = \{U > U'\} \in \mathcal{B}$ car $U$ et $U'$ sont $\mathcal{B}$-mesurables

$$\mathbb{E}(U\mathbb{1}_{U>U'}) = \mathbb{E}(U'\mathbb{1}_{U>U'}) \Leftrightarrow \mathbb{E}((U-U')\mathbb{1}_{(U>U')}) = 0$$

$$\Rightarrow \mathbb{P}(U>U') = 0 \quad \text{et de même } \mathbb{P}(U<U') = 0$$

existence On utilise le th de Radon-Nikodym (p 52 de Le Gall)

si $\nu$ et $\mu$ sont des mesures $\geq 0$ sur $(\Omega, \mathcal{F})$ avec $\nu \ll \mu$

$[\forall A \in \mathcal{F},\ \mu(A) = 0 \Rightarrow \nu(A) = 0]$

alors il existe $f : \Omega \to \mathbb{R}_+$, $\mathcal{F}$-mesurable, telle que $\forall A \in \mathcal{F}$,

$\nu(A) = \int_A f \, d\mu$ $f$ est appelée $\left|\begin{array}{l}\text{dérivée de Radon-Nikodym} \\ \text{densité}\end{array}\right.$

$\boxed{\text{Application de ce th : on suppose } X \geq 0}$ : pour tout $B \in \mathcal{B}$, on pose $Q(B) = \mathbb{E}(X \mathbb{1}_B)$

$Q : \mathcal{B} \to \mathbb{R}_+$ est une mesure :

$$\left[\begin{array}{l} Q(B^c) = Q(\Omega) - Q(B) \\ Q(\emptyset) = 0 \\ Q\left(\bigsqcup_{n\in\mathbb{N}} B_n\right) = \sum_{n\in\mathbb{N}} Q(B_n) \end{array}\right.$$

$Q \ll P$ : si $P(B) = 0$, $Q(B) = \mathbb{E}(\underbrace{X \mathbb{1}_B}_{\text{v.a. nulle p.s.}}) = 0$

Radon-Nikodym : $\exists f : \Omega \to \mathbb{R}_+$, mesurable $/\mathcal{B}$, $\forall B \in \mathcal{B}$,

$Q(B) = \int_B f \, P$

fonction mesurable $/\mathcal{B}$ = variable aléatoire mesurable $/\mathcal{B}$ → on l'appelle $\tilde{X}$

$$\forall B \in \mathcal{B} \qquad \mathbb{E}(\tilde{X} \mathbb{1}_B) = \int_B f \, dP = Q(B) = \mathbb{E}(X \mathbb{1}_B)$$

→ on a démontré le th pour $X$ v.a $\geq 0$

cas général : $X = X_+ - X_-$ où $X_+ \geq 0$ p.s. $\quad X_+ = X \mathbb{1}_{\{X>0\}}$

$X_- \geq$ p.s

remarque : dans le th, on a démontré (1). Pour démontrer (2) avec $Z$ v.a. $\mathcal{B}$-mesurable bornée, on approxime $Z$ par des combinaisons linéaires d'indicatrices : $Z^{(n)} = \sum_k \frac{k}{n} \mathbb{1}\left(Z \in \left[\frac{k-1}{n}, \frac{k}{n}\right]\right)$

exemple : $\Omega = ]0,1[$, $\mathcal{A}$ tribu borélienne, $\mathcal{B}$ sous-tribu engendrée par les intervalles $\left[\frac{i}{n}, \frac{i+1}{n}\right)$ $i$ entier, $0 \leq i \leq n-1$, $P$ mesure uniforme

$X \in L^1(\Omega, \mathcal{A}, P)$

$\mathbb{E}(X \mid \mathcal{B})$ est une v.a. mesurable $/ \mathcal{B}$ qui est engendrée par les fonctions $\underbrace{\mathbb{1}_{[\frac{i}{n}, \frac{i+1}{n}]}}_{Y_i}$ $\rightarrow \mathbb{E}(X \mid \mathcal{B}) = f(\mathbb{1}_{[\frac{i}{n}, \frac{i+1}{n})})$

$$\mathbb{E}(X \mid \mathcal{B}) = f(X_0, X_1 \dots X_{n-1})$$

On vérifie, en prenant $\mathbb{E}(X X_i) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{B}) X_i)$, que

$$\mathbb{E}(X \mid \mathcal{B}) = \sum_i \mathbb{E}(X X_i) X_i$$

Propriétés :

(a) si $X$ est $\mathcal{B}$-mesurable, alors $\mathbb{E}(X \mid \mathcal{B}) = X$

(b) $X \longmapsto \mathbb{E}(X \mid \mathcal{B})$ linéaire

(c) $\mathbb{E}(\mathbb{E}(X \mid \mathcal{B})) = \mathbb{E}(X)$

(d) $\mathbb{E}(X \mid \mathcal{B}) \leq \mathbb{E}(|X| \mid \mathcal{B})$

(e) si $X \leq X'$, alors $\mathbb{E}(X \mid \mathcal{B}) \leq \mathbb{E}(X' \mid \mathcal{B})$

dém (a) $X$ vérifie toutes les propriétés de la v.a. qui définit l'espérance conditionnelle

(b) $X, X' \in L^1$ $\quad Z = \lambda X + \lambda' X'$

$Y = \lambda \mathbb{E}(X \mid \mathcal{B}) + \lambda' \mathbb{E}(X' \mid \mathcal{B})$ est $\mathcal{B}$-mesurable

$$\forall B \in \mathcal{B}, \quad \mathbb{E}(Y \mathbb{1}_B) = \lambda \mathbb{E}(\mathbb{E}(X \mid \mathcal{B}) \mathbb{1}_B) + \lambda' \mathbb{E}(\mathbb{E}(X' \mid \mathcal{B}) \mathbb{1}_B)$$
$$= \lambda \mathbb{E}(X \mathbb{1}_B) + \lambda' \mathbb{E}(X' \mathbb{1}_B)$$
$$= \mathbb{E}[(\lambda X + \lambda' X') \mathbb{1}_B] = \mathbb{E}(Z \mathbb{1}_B)$$

$\rightarrow Y = \mathbb{E}(Z \mid \mathcal{B})$

(c) on prend $B = \Omega$

(d) si $X \geq 0$ alors $\mathbb{E}(X \mid \mathcal{B}) \geq 0$

sinon on écrit $X = X_+ - X_-$ $\qquad X_+ = X \mathbb{1}_{(X \geq 0)}$

$$|X| = X_+ + X_-$$

$$\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(X_+|\mathcal{B}) - \mathbb{E}(X_-|\mathcal{B}) \leq \mathbb{E}(X_+|\mathcal{B}) + \mathbb{E}(X_-|\mathcal{B})$$
$$= \mathbb{E}(|X|\,|\mathcal{B})$$

(e) si $X \leq X'$ alors $\mathbb{E}(\underbrace{X' - X}_{\geq 0}|\mathcal{B}) \geq 0 \Rightarrow \mathbb{E}(X'|\mathcal{B}) \geq \mathbb{E}(X|\mathcal{B})$

---

Extension aux variables aléatoires $\geq 0$ [qui peuvent valoir $+\infty$]

$a \wedge b = \inf(a,b)$

Th $X$ v.a $\geq 0$, alors on peut définir

$$\mathbb{E}(X|\mathcal{B}) = \lim_{n\to\infty}\nearrow \mathbb{E}((X\wedge n)|\mathcal{B})$$

$\mathbb{E}(X|\mathcal{B})$ est une v.a. $\mathcal{B}$-mesurable et $\forall Z$ v.a $\mathcal{B}$-mesurable positive,

$$\mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(X|\mathcal{B})Z) \leq \infty \quad (3)$$

dém : comme $X \geq 0$, $X \wedge n$ est bornée donc $\in L^1$ et $\mathbb{E}(X\wedge n|\mathcal{B})$ est bien définie. $n < n'$, $(X \wedge n) \leq (X \wedge n') \Rightarrow \mathbb{E}(X\wedge n|\mathcal{B}) \leq \mathbb{E}(X\wedge n'|\mathcal{B})$

donc la suite $(\mathbb{E}(X\wedge n|\mathcal{B}))_{n\geq 1}$ est une suite croissante de variables aléatoires $\mathcal{B}$-mesurables. Elle admet une limite p.s qui est $\geq 0$, $\mathcal{B}$ mesurable

$$\mathbb{E}(XZ) = \lim\nearrow \mathbb{E}((X\wedge n)Z) \quad \text{(Th de CV monotone)}$$
$$= \lim\nearrow \mathbb{E}(\mathbb{E}(X\wedge n|\mathcal{B})\, Z)$$
$$= \mathbb{E}(\lim\nearrow \mathbb{E}(X\wedge n|\mathcal{B})\, Z)$$

Th : à un ensemble de mesure nulle près, il existe une unique v.a. $\geq 0$, $\mathcal{B}$ mesurable qui vérifie (3)

dém : si $Y, Y'$ sont $\geq 0$, $\mathcal{B}$ mesurables et si $\forall Z \geq 0$, $\mathcal{B}$-mesurable,

$$\mathbb{E}(YZ) = \mathbb{E}(Y'Z) \qquad (4)$$

alors $\forall q, q'$ rationnels, $q < q'$, en prenant $Z = \mathbb{1}_{\{Y' \leq q < q' \leq Y\}}$

$$(4) \Rightarrow \quad q\, \mathbb{P}(Y' \leq q < q' < Y) \geq q'\, \mathbb{P}(Y' \leq q < q' < Y)$$

$$\mathbb{P}(Y' \leq q < q' < Y) = 0$$

L'ensemble $\{Y' < Y\} = \bigcup_{\substack{q<q' \\ q,q' \in \mathbb{Q}}} \{Y' \leq q < q' \leq Y\}$ est de probabilité nulle

même chose $\{Y < Y'\}$ donc $Y = Y'$ p.s.

exemple : $\Omega = ]0,1[$, $\mathcal{A}$ tribu borélienne, $\mathcal{B}$ tribu engendrée par les

$\left[\frac{i}{n}, \frac{i+1}{n}\right]$, $0 \leq i \leq n-1$ $\qquad$ $X$ uniforme sur $\Omega$, $\quad Y = \frac{1}{X} \geq 0$

$X_i = \mathbb{1}_{\left[\frac{i}{n}, \frac{i+1}{n}\right]}$ $\qquad$ $Y \notin L^1$

$$\mathbb{E}(Y \mid \mathcal{B}) = f(X_i) \qquad \mathbb{E}(Y X_i) = \mathbb{E}(\mathbb{E}(Y/\mathcal{B})\, X_i)$$

$$\mathbb{E}(Y) = \sum_{i=0}^{n-1} c_i X_i \qquad c_i = \int_{\frac{i}{n}}^{\frac{i+1}{n}} \frac{1}{x}\, dx$$

$$c_0 = \infty$$

Prop : pour des v.a. positives

(a) $\lambda, \lambda' \geq 0$, $\quad \mathbb{E}(\lambda X + \lambda' X' \mid \mathcal{B}) = \lambda\, \mathbb{E}(X \mid \mathcal{B}) + \lambda'\, \mathbb{E}(X' \mid \mathcal{B})$

(b) si $X$ est $\mathcal{B}$-mesurable, $\quad \mathbb{E}(X/\mathcal{B}) = X$

(c) $X_n$ suite $\nearrow$ de variables aléatoires $\geq 0$, $X = \lim X_n$

alors $\mathbb{E}(X \mid \mathcal{B}) = \lim_{n \to \infty} \nearrow \mathbb{E}(X_n \mid \mathcal{B})$

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

(e) $X_n$ suite v.a. dans $L^1$, $X_n \xrightarrow{p.s.} X$

si $\exists Y$ v.a $\geq 0$, tq $\forall n$, $|X_n| < Y$ p.s. et $\mathbb{E}(Y) < \infty$

alors $\mathbb{E}(X \mid \mathcal{B}) = \mathbb{E}(\lim X_n \mid \mathcal{B}) = \lim \mathbb{E}(X_n \mid \mathcal{B})$

où la limite est une limite p.s et dans $L^1$

<u>dém</u> (a) et (b) : mêmes démonstrations que la proposition précédente

(c) si $0 \leq Y \leq Y'$, alors $\mathbb{E}(Y' - Y \mid \mathcal{B}) \geq 0 \Rightarrow \mathbb{E}(Y \mid \mathcal{B}) \leq \mathbb{E}(Y' \mid \mathcal{B})$

donc si $X_n \nearrow$, alors $\mathbb{E}(X_n \mid \mathcal{B}) \nearrow$, il existe une limite p.s qu'on appelle $X'$. Si $Z \geq 0$, $\mathcal{B}$ mesurable,

$$\mathbb{E}(Z X') = \mathbb{E}(Z \lim^{\nearrow} \mathbb{E}(X_n \mid \mathcal{B})) \overset{\text{th de CV monotone}}{=} \lim^{\nearrow} \mathbb{E}(Z\, \mathbb{E}(X_n \mid \mathcal{B}))$$
$$= \lim^{\nearrow} \mathbb{E}(Z X_n)$$
$$= \mathbb{E}(Z \lim^{\nearrow} X_n) = \mathbb{E}(Z X)$$

donc $X' = \mathbb{E}(X \mid \mathcal{B})$

(d) si $(u_n)$ suite de réels, $\liminf u_n = \lim\limits_{k \to \infty}^{\nearrow} \inf\limits_{n \geq k} u_n$

$$\mathbb{E}(\liminf X_n \mid \mathcal{B}) = \mathbb{E}\left(\lim_k^{\nearrow} \left(\inf_{n \geq k} X_n\right) \mid \mathcal{B}\right) = \lim_k^{\nearrow} \mathbb{E}\left(\inf_{n \geq k} X_n \mid \mathcal{B}\right)$$
$$\leq \lim_k^{\nearrow} \inf_{n \geq k} \mathbb{E}(X_n \mid \mathcal{B})$$
$$= \liminf_{n \to \infty} \mathbb{E}(X_n \mid \mathcal{B})$$

(e) CV ps

$$\mathbb{E}(\liminf (Y - X_n) \mid \mathcal{B}) \leq \mathbb{E}(Y \mid \mathcal{B}) - \limsup \mathbb{E}(X_n \mid \mathcal{B})$$
$$\mathbb{E}(\liminf Y + X_n \mid \mathcal{B}) \leq \mathbb{E}(Y \mid \mathcal{B}) + \liminf \mathbb{E}(X_n \mid \mathcal{B})$$
$$\mathbb{E}(X \mid \mathcal{B}) = \mathbb{E}(\lim X_n \mid \mathcal{B}) = \mathbb{E}(\liminf X_n \mid \mathcal{B})$$
$$\leq \liminf \mathbb{E}(X_n \mid \mathcal{B}) \leq \limsup \mathbb{E}(X_n \mid \mathcal{B}) \leq \mathbb{E}(X \mid \mathcal{B})$$

$\to$ on a CV p.s.

c.v dans $L^1$

$$\mathbb{E}(X_n|\mathcal{B}) \leq \mathbb{E}(|X_n||\mathcal{B}) \leq \mathbb{E}(Y|\mathcal{B})$$

$\mathbb{E}(Y|\mathcal{B})$ est une variable aléatoire $\in L^1$ car

$$\mathbb{E}(\mathbb{E}(Y|\mathcal{B})) = \mathbb{E}(Y) < \infty$$

La suite de v.a $\mathbb{E}(X_n|\mathcal{B})$ cv p.s. vers $\mathbb{E}(X|\mathcal{B})$ est elle est dominée par une v.a. dans $L^1$

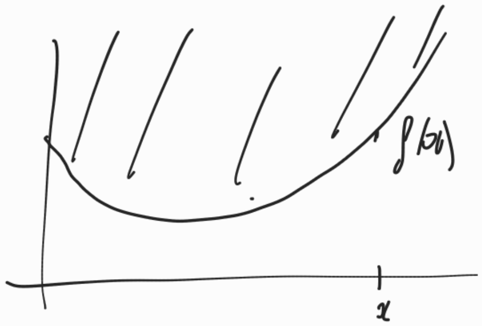
donc $\mathbb{E}(X_n|\mathcal{B}) - \mathbb{E}(X|\mathcal{B}) \xrightarrow{p.s} 0$ et par th de CV dominée, $\mathbb{E}(X_n|\mathcal{B}) - \mathbb{E}(X|\mathcal{B}) \xrightarrow{L^1} 0$

<u>Prop</u> : Inégalité de Jensen pour l'espérance conditionnelle.

$X$ v.a positive, $\in L^1$, $\mathcal{B}$ sous-tribu, $f : \mathbb{R}_+ \to \mathbb{R}_+$ convexe, alors $\mathbb{E}(f(X)|\mathcal{B}) \geq f(\mathbb{E}(X|\mathcal{B}))$ p.s. (5)

<u>dém</u> : si $f = ax+b$, on a égalité dans (5)

rappel :




$(x, f(x)) \in \mathbb{R}_+^2$ est extrémal si $\underbrace{\{(a,b) \in \mathbb{R}_+^2, b \geq f(a)\}}_{E_f} \setminus \{(x, f(x))\}$

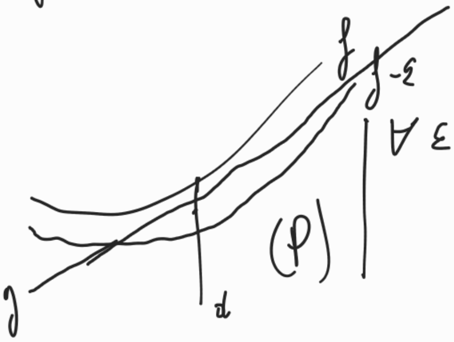
est convexe


$\leftarrow (x, f(x))$ non extrémal

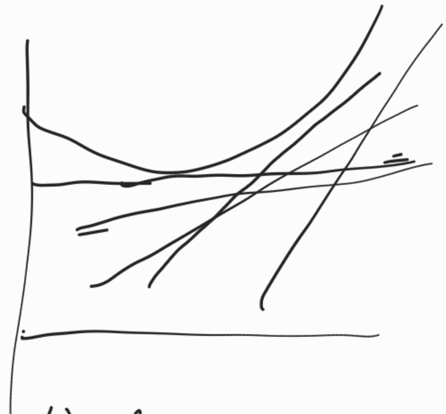
Si $f$ est strictement convexe, tous les points $(x, f(x))$ sont extrémaux.

$$x < y \quad , \quad f\left(\frac{x+y}{2}\right) < \frac{f(x)+f(y)}{2}$$

On fait la démonstration pour $f$ strictement convexe.



(P) $\forall \varepsilon > 0$, $\exists\, q, q' \in \mathbb{Q}$ tels que la fonction $g(y) = qy + q'$ soit $\leq f$ et $f(x) - \varepsilon \leq g(x) \leq f(x)$



$$\mathbb{E}(f(X)/\mathcal{B}) = \mathbb{E}\Big[\sup_{(q,q')\in A} (qX+q')/\mathcal{B})\Big)$$

$$A = \{(q,q') \in \mathbb{Q}^2,\ \forall x,\ qx+q' \leq f(x)\}$$

par définition de $A$, $f(X) \geq qX+q'$ si $(q,q') \in A$

donc $f(X) \geq \sup\limits_{(q,q')\in A} qX+q'$

mais d'après (P), $\forall \varepsilon > 0$, $f(x) \leq \sup\limits_{(q,q')\in A} (qX+q') + \varepsilon$

$$\mathbb{E}\Big[\sup_{(q,q')\in A} (qX+q')/\mathcal{B}\Big] \geq \mathbb{E}(qX+q'/\mathcal{B}) \quad \forall (q,q') \in A$$

sauf sur un ensemble négligeable $N_{(q,q')}$

$$\mathbb{E}\Big(\sup_{(q,q')\in A} (qX+q')/\mathcal{B}\Big) \geq \sup_{q,q'\in A} \mathbb{E}((qX+q')/\mathcal{B}) = \sup_{(q,q')\in A}(q' + q\,\mathbb{E}(X/\mathcal{B}))$$

sauf sur $\bigcup\limits_{(q,q')\in A} N_{(q,q')}$ qui est négligeable

$$\mathbb{E}(f(X)/\mathcal{B}) \geq f(\mathbb{E}(X/\mathcal{B}))$$

[remplacer $qx+q'$ par $(qx+q') \wedge 0$]

$$\mathbb{E}\left((gX+g')\wedge 0 \mid \mathcal{B}\right) \overset{?}{=} \left(g\mathbb{E}(X\mid\mathcal{B})+g'\right)\wedge 0$$