Espérance conditionnelle

Dans le cas discret, si X variable aléatoire $\in L^1$, Y variable aléatoire, $\mathbb{E}(X|Y)$ ~~est~~ une variable aléatoire mesurable par rapport à $\sigma(Y)$.

Si de plus $X \in L^2$, $\mathbb{E}(X/Y)$ est la projection orthogonale de X sur l'espace vectoriel des variables aléatoires $\in L^2$ qui sont mesurables par rapport à $\sigma(Y)$

exemple n entier $n \geq 2$, U, V variables aléatoires uniformes sur $\{1, 2 \dots n\}$. $S = U+V$, $P = UV$; $\mathbb{E}(P|S)$?

$$\{S = k\} = \bigcup_{i=1}^{n} \{U = i, V = k-i\}$$

$$\mathbb{P}(S = k) = \sum_{i=1}^{n} \mathbb{P}(U = i, V = k-i) = \sum_{i=1}^{n} \underbrace{\mathbb{P}(U = i)}_{\frac{1}{n}} \underbrace{\mathbb{P}(V = k-i)}_{\begin{cases} \frac{1}{n} & \text{si } k-i \in [1,n] \\ 0 & \text{sinon} \end{cases}}$$

$$\mathbb{E}(P | S = k) = \mathbb{E}\left(\sum_{i=1}^{n} \mathbb{1}_{\{U=i, V=k-i\}} \, i(k-i)\right)$$

$$= \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{1}_{\{U=i, V=k-i\}} \, i(k-i)\right]$$

$$= \sum_{i=1}^{n} i(k-i) \, \mathbb{P}(U=i, V=k-i) = \underbrace{\sum_{i=1}^{n} \frac{i(k-i)}{n^2} \mathbb{1}_{\{k-i \in [1,n]\}}}_{\varphi(k)}$$

$$\mathbb{E}(P | S = k) = \varphi(k)$$

$\mathbb{E}(P|S) = \varphi(S) \rightarrow$ fonction de S, donc c'est mesurable par rapport à la tribu $\sigma(S)$

Cas général

On a montré dans le cas discret que pour toute v.a. Z qui est mesurable $/\sigma(Y)$, $\mathbb{E}(ZX) = \mathbb{E}(Z\,\mathbb{E}(X|Y))$ $(*)$

si $X, Z \in L^2$, $\mathbb{E}(ZX) = \phi(Z,X)$ ϕ : forme bilinéaire associée au produit scalaire qui définit l'espace de Hilbert L^2

$(*)$ se réécrit

$\phi(Z,X) = \phi(Z, \pi_E(X))$ où π_E est la projection orthogonale sur l'espace vectoriel des variables aléatoires mesurables/ $\sigma(Y)$ $(Z \in E)$

__Th et déf__ : $(\Omega, \mathcal{A}, P)$ espace de proba, $\mathcal{B}$ sous-tribu de $\mathcal{A}$.

Soit $X \in L^1(\Omega,\mathcal{A},P)$, 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, $\forall Z$ v.a. bornée mesurable $/\mathcal{B}$, $\quad (*) \Rightarrow (1)$ évident

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

__dém__. unicité : si X et X' vérifient (1), $B = \{X > X'\}$

$$\mathbb{E}(X\,\mathbb{1}_B - X'\mathbb{1}_B) = \mathbb{E}((X-X')\mathbb{1}_B) = 0$$

comme $X - X' > 0$ sur B, alors $\mathbb{P}(B) = 0$

donc $\mathbb{P}(X \leq X') = 1$; de même $\mathbb{P}(X \geq X') = 1$

$\mathbb{P}(X = X') = 1$

· Existence

Rappel : version faible du th. de Radon-Nikodym

si ν et μ sont des mesures ≥ 0 sur $(\Omega, \mathcal{F})$

si : $\nu \ll \mu$ [absolument continue : $\forall B \in \mathcal{F}$, $\mu(B) = 0 \Rightarrow \nu(B) = 0$]

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

$$\nu(A) = \int_A f \, d\mu$$

f : (densité
dérivée de Radon-Nikodym

p.52 de Le Gall

· $B \in \mathcal{B}$ on pose $Q(B) = \mathbb{E}(X \mathbb{1}_B)$

Q est une mesure sur $(\Omega, \mathcal{B})$ (facile à démontrer)

Q est σ-additive si $(B_i)_{i \in \mathbb{N}}$ sont deux à deux disjoints

$$\sum Q(B_i) = \mathbb{E}(X \sum \mathbb{1}_{B_i}) = \mathbb{E}(X \mathbb{1}_{\cup B_i}) = Q(\cup B_i)$$

$Q \ll P$: si $P(B) = 0$, $Q(B) = \mathbb{E}(\underbrace{X \mathbb{1}_B}) = 0$

v.a qui est nulle p.s.

On peut appliquer le th de Radon-Nikodym à Q, P : $\exists f : \Omega \to \mathbb{R}_+$ mesurable / $\mathcal{B}$, $\forall B \in \mathcal{B}$, $Q(B) = \int_B f \, dP$

donc $(\Omega, \mathcal{B}) \xrightarrow{f} \mathbb{R}$ est une fonction mesurable : c'est une variable aléatoire mesurable / $\mathcal{B}$, on appelle cette v.a. $\hat{X}$

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

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

Pour passer de (1) à (*), on approxime les variables aléatoires mesurables $/\mathcal{B}$ par des fonctions étagées. (fonctions de la forme $\sum_{i=1}^{n} c_i \mathbb{1}_{B_i}$ où les $B_i \in \mathcal{B}$)

<u>exemple</u> : $\Omega =]0,1[$, $\mathcal{A}$ = boréliens, $n \geq 2$, $\mathcal{B}$ tribu engendrée par les intervalles $\left[\frac{i}{n}, \frac{i+1}{n}\right]$, $0 \leq i \leq n-1$ [$\mathcal{B}$ est l'ensemble des réunions d'intervalles dont les bornes sont de la forme $\frac{k}{n}$] ; P mesure uniforme

$f \in L^1(\Omega, \mathcal{A}, P)$ ssi $\int_0^1 f(x)\,dx < \infty$

variable aléatoire

$$\underline{\mathbb{E}(f \mid \mathcal{B})} = \underbrace{\sum_{i=0}^{n-1} c_i \underbrace{\mathbb{1}_{\left[\frac{i}{n}, \frac{i+1}{n}\right]}}_{\text{variable aléatoire}}}_{\text{mesurable } /\mathcal{B}} \quad \text{où } c_i = \int_{i/n}^{\frac{i+1}{n}} f(x)\,dx$$

les v.a mesurables $/\mathcal{B}$ sont les fonctions constantes sur tout intervalle $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ $\rightarrow$ donc ce sont les fonctions de la forme $\sum_{i=0}^{n-1} d_i \mathbb{1}_{\left[\frac{i}{n}, \frac{i+1}{n}\right]}$

<u>Propriétés</u> :

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

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

<u>dém</u> : (a) X vérifie (*) et comme on montré l'unicité, $X = \mathbb{E}(X \mid \mathcal{B})$

(b) si X, X' sont des v.a, $Z = \lambda X + \lambda X'$, on pose

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

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

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

(c) $B = \Omega$, $\mathbb{1}_B = 1$, (1) $\Rightarrow$ (c)

(d) si $X \geq 0$, $\mathbb{E}(X \mid \mathcal{B}) \geq 0$; dans le cas général

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

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

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

Variables aléatoires ≥ 0

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

Th X v.a à valeurs ≥ 0 alors

$$\mathbb{E}(X \mid \mathcal{B}) = \lim_{n \to \infty} \nearrow \mathbb{E}(X \wedge n \mid \mathcal{B}) \qquad (2)$$

est une variable aléatoire ≥ 0 caractérisée par : $\forall Z$ $\mathcal{B}$-mesurable positive

$$\mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(X \mid \mathcal{B}) Z) \qquad (3)$$

remarque : les variables aléatoires ≥ 0 peuvent être égales à $+\infty$

dém d'après (e) $\mathbb{E}(X \wedge n \mid \mathcal{B}) \nearrow$ en n $\rightarrow$ la limite dans (2) est bien définie dans $\mathbb{R}_+ \cup \{+\infty\}$

Si Z est ≥ 0 et $\mathcal{B}$-mesurable,

$$\mathbb{E}(\mathbb{E}(X|\mathcal{B})Z) = \mathbb{E}\Big(\lim_{n\to\infty} \underbrace{\mathbb{E}(X\wedge n|\mathcal{B})\cdot Z}_{\text{suite monotone (croissante) de variables aléatoires } \geq 0}\Big)$$

par th de CV monotone

$$\mathbb{E}\Big(\lim_{n\to\infty} \mathbb{E}(X\wedge n|\mathcal{B})\cdot Z\Big) = \lim_{n\to\infty} \mathbb{E}\Big(\underbrace{\mathbb{E}(X\wedge n|\mathcal{B})}_{\text{bornée} \geq 0} \underbrace{Z}_{\geq 0}\Big)$$

propriété de l'espérance conditionnelle

$$= \lim_{n\to\infty} \mathbb{E}((X\wedge n)Z)$$

cv monotone

$$= \mathbb{E}(XZ)$$

. Unicité de la variable aléatoire $\mathcal{B}$-mesurable vérifiant (3)

si X, X' vérifient (3) et sont $\mathcal{B}$-mesurables, alors

$\forall\, q < q' \in \mathbb{Q}$, si $Z = \mathbb{1}_{\{X' \leq q < q' \leq X\}}$,

$$q\,\mathbb{P}(X' \leq q < q' \leq X) \geq q'\,\mathbb{P}(X' \leq q < q' \leq X)$$

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

$$\mathbb{P}\Big(\bigcup_{\substack{q,q'\in\mathbb{Q}\\ q<q'}} \underbrace{\{X' \leq q < q' \leq X\}}_{}\Big) = 0$$

réunion dénombrable d'ensembles négligeables

. Si on reprend l'exemple $\Omega =]0,1[$, $\mathcal{A}$ tribu borélienne, $\mathcal{B}$ sous-tribu engendrée par les intervalles $\left[\frac{i}{n}, \frac{i+1}{n}\right[$

Si on prend la fonction positive $f(x) = \frac{1}{x}$; f a ses valeurs dans $\mathbb{R}_+$

$$\mathbb{E}(f \mid \mathcal{B}) = \sum_{i=0}^{n-1} c_i \, \mathbb{1}_{[\frac{i}{n}, \frac{i+1}{n}]} \qquad c_i = \int_{\frac{i}{n}}^{\frac{i+1}{n}} f(t)\,dx$$

$c_0 = +\infty$ et $c_1 \dots c_{n-1} \in \mathbb{R}$

$+\infty + c = +\infty$

$c \cdot +\infty = +\infty$

<u>Prop</u> : (a) si X, X' v.a positives, $a, b \geq 0$,

$$\mathbb{E}(aX + bX' \mid \mathcal{B}) = a\,\mathbb{E}(X \mid \mathcal{B}) + b'\,\mathbb{E}(X' \mid \mathcal{B})$$

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

(c) X_n suite croissante de variables aléatoires positives, $X = \lim_{n\to\infty} X$

alors $\mathbb{E}(X \mid \mathcal{B}) = \lim_{n\to\infty} \nearrow \mathbb{E}(X_n \mid \mathcal{B})$ (version du Th de CV monotone pour l'espérance conditionnelle)

(d) Si X_n suite de v.a. positives

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

(e) X_n suite de v.a $\in L^1$ qui CV p.s vers X

S'il existe une v.a positive Z tq $\forall n, |X_n| \leq Z$ p.s et $\mathbb{E}(Z) < \infty$

alors $\mathbb{E}(X \mid \mathcal{B}) = \lim_{n\to\infty} \mathbb{E}(X_n \mid \mathcal{B})$ p.s et dans L^1

(f) si la fonction f est positive et convexe, si $X \in L^1$

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

[version de l'inégalité de Jensen pour l'espérance conditionnelle]

<u>dém</u> (c) si $0 \leq X_1 \leq X_2$, alors $\mathbb{E}(X_1 \mid \mathcal{B}) \leq \mathbb{E}(X_2 \mid \mathcal{B})$

donc on peut poser $X' = \lim_{n\to\infty} \nearrow \mathbb{E}(X_n \mid \mathcal{B})$ qui est $\mathcal{B}$-mesurable

Si Z positive, $\mathcal{B}$ - mesurable

$$\mathbb{E}(ZX') = \mathbb{E}(Z \lim_{n\to\infty}\nearrow \mathbb{E}(X_n | \mathcal{B})) \overset{\text{CV monotone}}{=} \lim_{n\to\infty} \mathbb{E}(Z\,\mathbb{E}(X_n|\mathcal{B}))$$
$$= \lim_{n\to\infty} \mathbb{E}(Z X_n)$$
$$= \mathbb{E}(ZX)$$

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

(d) on utilise le fait que pour toute suite de réels (u_n),

$$\liminf (u_n) = \lim_{k\to\infty}\nearrow \left(\inf_{n\geq k} u_n\right)$$

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

(e) on applique (d)

$$\mathbb{E}(\liminf (Z - X_n) | \mathcal{B}) \leq \mathbb{E}(Z|\mathcal{B}) - \limsup \mathbb{E}(X_n|\mathcal{B})$$
$$\mathbb{E}(\liminf (Z + X_n) | \mathcal{B}) \leq \mathbb{E}(Z|\mathcal{B}) + \liminf \mathbb{E}(X_n|\mathcal{B})$$

$$\mathbb{E}(X|\mathcal{B}) \leq \liminf \mathbb{E}(X_n|\mathcal{B}) \leq \limsup \mathbb{E}(X_n|\mathcal{B}) \leq \mathbb{E}(X|\mathcal{B})$$

→ on a bien CV ps

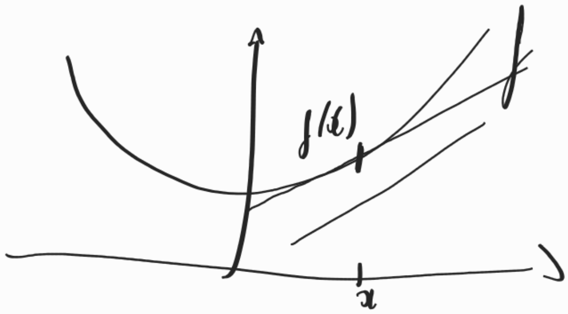
montrons la CV dans L^1 :

$$\mathbb{E}(X_n \mid \mathcal{B}) \leq \mathbb{E}(|X_n| \mid \mathcal{B}) \leq \mathbb{E}(Z \mid \mathcal{B}) < \infty$$

$$\mathbb{E}(\mathbb{E}(Z \mid \mathcal{B})) = \mathbb{E}(Z) < \infty$$

on applique le th de CV dominée.

(f) $E_f = \{(a,b) \in \mathbb{R}^2, \forall x \in \mathbb{R}, f(x) \geq ax+b$



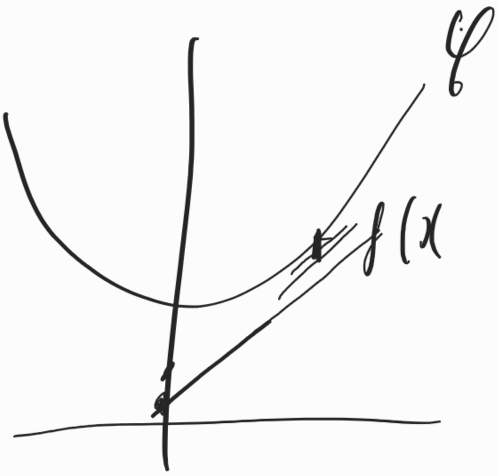
[on considère l'ensemble des droites qui sont en dessous de la courbe de f]

on vérifie que $\forall x$, $f(x) = \sup\limits_{(a,b) \in E_f} (ax+b) = \sup\limits_{(a,b) \in E_f \cap \mathbb{Q}^2} (ax+b)$ dénombrable

$$\mathbb{E}(f(X) \mid \mathcal{B}) = \mathbb{E}\left(\sup_{(a,b) \in E_f \cap \mathbb{Q}^2} aX+b \mid \mathcal{B}\right) \geq \sup_{(a,b) \in E_f \cap \mathbb{Q}^2} \mathbb{E}(aX+b \mid \mathcal{B})$$

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

Δ



1 nb réel peut être approché par des rationnels