Jeudi 17 mars : cours de probabilités au lieu du cours de L. Mérel

21 et 23 mars : cours avec L. Mérel

---

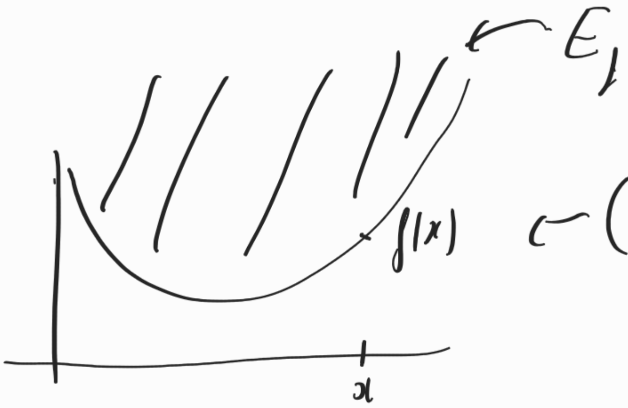
(8) Inégalité de Jensen pour l'espérance conditionnelle :

X variable aléatoire positive, $\mathcal{B}$ sous-tribu, $f : \mathbb{R}_+ \to \mathbb{R}_+$ fonction convexe. Alors

$$\mathbb{E}\big(f(X)\,|\,\mathcal{B}\big) \geq f\big(\mathbb{E}(X\,|\,\mathcal{B})\big) \quad \text{p.s.} \qquad (1)$$

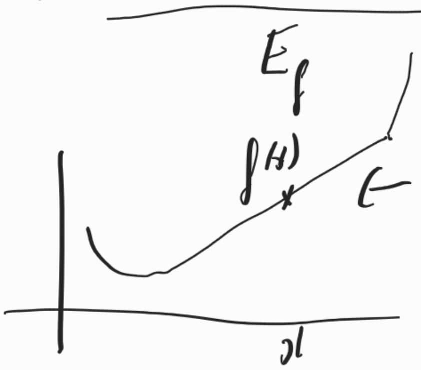
. si f est une fonction affine : $f(x) = ax + b$, on a égalité dans (1)

sinon



$\leftarrow (x, f(x))$ est extrémal

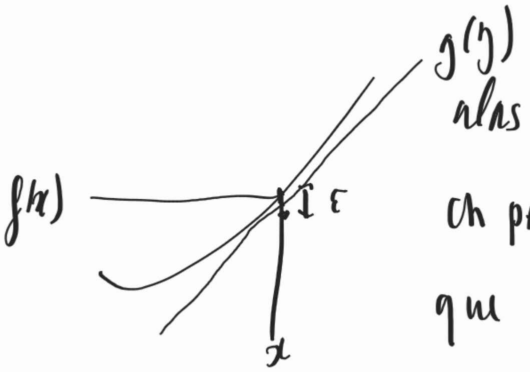
On dit que le point $(x, f(x)) \in \mathbb{R}^2$ est extrémal si l'ensemble $\underline{\{(x,y),\ x \in \mathbb{R}^+,\ y \geq f(x)\} \setminus (x, f(x))}$ est convexe



$\leftarrow (x, f(x))$ non extrémal

$(x, f(x))$ non extrémal si f est une fonction affine au voisinage de x

. si $(x, f(x))$ extrémal

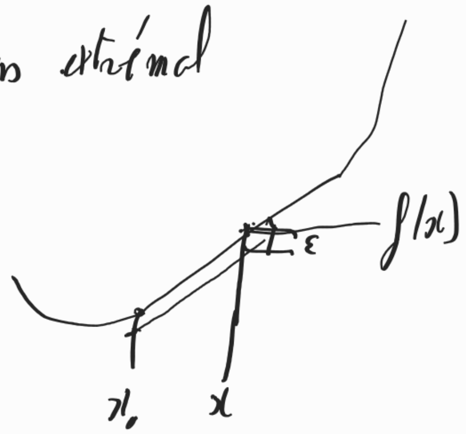


alors pour tout $\varepsilon > 0$, on peut trouver $q', q \in \mathbb{Q}$ tel que la fonction $g(y) = qy + q'$

soit inférieure à f $(\forall y,\ g(y) \leq f(y))$ et telle que

$f(x) - \varepsilon \leq g(x) \leq f(x)$

. si $(x, f(x))$ n'est pas extrémal



$x_0 = \inf \{ t,$ la fonction f est affine sur $[t, x]$

pour tout $\varepsilon > 0$, on peut trouver $q, q' \in \mathbb{Q}$, tq la fonction $g(y) = qy + q'$

soit inférieure à f et $f(x) - \varepsilon \leq g(x) \leq f(x)$

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

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

$$\underbrace{\mathbb{E}\left(\sup_{(q, q') \in A} (qX + q') / \mathcal{B}\right) \geq \mathbb{E}(q_1 X + q_2 / \mathcal{B})}$$ pour tout $(q_1, q_2) \in A$

et l'inégalité est vraie sur $\Omega - N_{q_1, q_2}$ où $P(N_{q_1, q_2}) = 0$

$$\mathbb{E}\Big(\sup_{(q,q')\in A} qX+q' \,/\, \mathcal{B}\Big) \geq \underbrace{\sup_{(q_1,q_2)\in A} \mathbb{E}(qX_1+q_2 \,/\, \mathcal{B})}$$

et l'inégalité est vraie sur $\Omega - \underbrace{\bigcup_{(q_1,q_2)\in A} N_{q_1,q_2}}$

réunion dénombrable d'ensembles négligeables $\longrightarrow$ négligeable

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

<u>Th</u> si $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ et $\mathcal{B}$ sous-tribu de $\mathcal{F}$, alors $\mathbb{E}(X/\mathcal{B})$ est la projection orthogonale de X sur $\underbrace{L^2(\Omega,\mathcal{B},P)}$

sous espace vectoriel des variables aléatoires de carré intégrable mesurables / $\mathcal{B}$

<u>dém</u> . $\mathbb{E}(X/\mathcal{B})^2 \leq \mathbb{E}(X^2/\mathcal{B})$

d'après (J)

$$\mathbb{E}(\mathbb{E}(X/\mathcal{B})^2) \leq \mathbb{E}(\mathbb{E}(X^2/\mathcal{B})) = \mathbb{E}(X^2) < \infty$$

$$\Rightarrow \mathbb{E}(X/\mathcal{B}) \in L^2$$

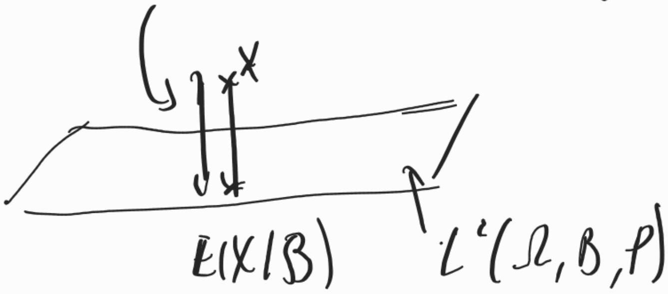
. si Z v.a $\mathcal{B}$-mesurable bornée,

$$\mathbb{E}(Z(X-\mathbb{E}(X/\mathcal{B}))) = \mathbb{E}(ZX) - \underbrace{\mathbb{E}(Z\mathbb{E}(X/\mathcal{B}))}_{\mathbb{E}(ZX)} = 0$$

Z orthogonal à $X - \mathbb{E}(X/\mathcal{B})$ pour tout Z $\mathcal{B}$-mesurable bornée par densité, c'est vrai pour tout Z $\mathcal{B}$-mesurable, L^2

(approximer Z par $\sup(\inf(Z,n),-n)$)

donc $X - \mathbb{E}(X|\mathcal{B})$ est orthogonal à $L^2(\Omega,\mathcal{B},P)$



Prop : X, Y v.a réelles, $\mathcal{B}$ sous-tribu. Si Y est $\mathcal{B}$-mesurable, alors

$$\mathbb{E}(YX|\mathcal{B}) = Y\,\mathbb{E}(X|\mathcal{B})$$ [on suppose que Y et $X \geq 0$ ou que Y et $X \in L^1$]

remarque : si $\lambda \in \mathbb{R}$ $\mathbb{E}(\lambda X) = \lambda\,\mathbb{E}(X)$

dém : on se place dans le cas où $X, Y \geq 0$. Soit Z v.a $\mathcal{B}$-mesurable positive

$$\mathbb{E}(Z(Y\,\mathbb{E}(X|\mathcal{B}))) = \mathbb{E}[\underbrace{(ZY)}_{\mathcal{B}\text{-mesurable } \geq 0}\,\mathbb{E}(X|\mathcal{B})] = \mathbb{E}(ZYX) = \mathbb{E}(Z\,\mathbb{E}(XY|\mathcal{B}))$$

par définition de l'espérance conditionnelle, on en déduit que

$$Y\,\mathbb{E}(X|\mathcal{B}) = \mathbb{E}(XY|\mathcal{B})$$

Prop si $\mathcal{B}_1, \mathcal{B}_2$ sont deux sous-tribus $\mathcal{B}_1 \subset \mathcal{B}_2$, alors pour toute v.a X positive (ou L^1), $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1) = \mathbb{E}(X|\mathcal{B}_1)$

remarque : si on est dans L^2, $L^2(\Omega,\mathcal{B}_1,P) \subset L^2(\Omega,\mathcal{B}_2,P)$, la prop. dit que si on projette sur $L^2(\Omega,\mathcal{B}_2,P)$ puis sur $L^2(\Omega,\mathcal{B}_1,P)$, c'est la même chose que de projeter sur $L^2(\Omega, \mathcal{B}_1, P)$

dém : cas ≥ 0, si $Z \geq 0$ est $\mathcal{B}_1$-mesurable, alors Z est aussi $\mathcal{B}_2$-mesurable

$$\mathbb{E}(Z\, \mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1)) = \mathbb{E}(Z\, \mathbb{E}(X|\mathcal{B}_2)) = \mathbb{E}(ZX)$$

Th ① si $\mathcal{B}_1, \mathcal{B}_2$ sont indépendantes, alors $\forall X$ v.a positive, $\mathcal{B}_2$-mesurable, $\mathbb{E}(X|\mathcal{B}_1) = \mathbb{E}(X)$

② réciproquement, si $\mathcal{B}_1, \mathcal{B}_2$ sont des sous-tribus et si $\forall X$ v.a ≥ 0, $\mathcal{B}_2$-mesurable, $\mathbb{E}(X|\mathcal{B}_1) = \mathbb{E}(X)$, alors $\mathcal{B}_1, \mathcal{B}_2$ sont indépendantes

dém ① $\forall Z \geq 0$, $\mathcal{B}_1$-mesurable,

$$\mathbb{E}(ZX) = \mathbb{E}(Z)\mathbb{E}(X) = \mathbb{E}(Z\, \mathbb{E}(X|\mathcal{B}_1))$$

donc $\mathbb{E}(X) = \mathbb{E}(X|\mathcal{B}_1)$

② en particulier $\forall A \in \mathcal{B}_2$, $\mathbb{1}_A$ est ≥ 0, $\mathcal{B}_2$-mesurable

$$\mathbb{E}(\mathbb{1}_A|\mathcal{B}_1) = \mathbb{E}(\mathbb{1}_A) = P(A)$$

Si $B \in \mathcal{B}_1$, $\underline{P(A\cap B)} = \mathbb{E}(\mathbb{1}_{A\cap B}) = \underline{\mathbb{E}(\mathbb{1}_A \mathbb{1}_B)}$

$$= \mathbb{E}(\underbrace{\mathbb{E}(\mathbb{1}_A|\mathcal{B}_1)}_{P(A)}\mathbb{1}_B) = \mathbb{E}(P(A)\mathbb{1}_B) = P(A)\,\mathbb{E}(\mathbb{1}_B) = \underline{P(A)P(B)}$$

$\Rightarrow \mathcal{B}_1, \mathcal{B}_2$ sont indépendantes

Th X v.a à valeurs dans E | on suppose que Y est $\mathcal{B}$-mesurable
Y —————— F | X indépendante de $\mathcal{B}$

Alors pour toute fonction mesurable $g : E \times F \to \mathbb{R}_+$

$$\mathbb{E}(g(X,Y) \mid \mathcal{B}) = \int g(x,Y)\, P_X(dx) \qquad (1)$$

où P_X est la loi de X

<u>remarque</u> : si Y est déterministe alors la variable aléatoire donnée par (1) est aussi déterministe → on se ramène au résultat précédent : on peut écrire $\mathbb{E}(g(X,Y)/\mathcal{B}) = \mathbb{E}(h(X) \mid \mathcal{B})$

Une variable aléatoire X est déterministe si $\exists C \in \mathbb{R}$ $X = C$ p.s

<u>dém</u> : On note $\phi : y \longmapsto \int g(x,y)\, P_X(dx)$

soit $Z \geq 0$, $\mathcal{B}$-mesurable.

Y, Z $\mathcal{B}$-mesurables et X indépendante de $\mathcal{B}$

$$\mathbb{E}\big(g(X,Y)Z\big) = \int g(x,y)\, z\; P(dx, dy, dz)$$

$$= \int g(x,y) z\; P_X(dx)\, P_{(Y,Z)}(dy,dz)$$

$$= \int_{F \times \mathbb{R}_+} z \underbrace{\left(\int_E g(x,y)\, P_X(dx)\right)} P_{Y,Z}(dy\, dz)$$

$$= \int_{F \times \mathbb{R}_+} z\; \phi(y)\; P_{Y,Z}(dy\, dz) = \mathbb{E}(\phi(Y) Z)$$

et $\phi(Y)$ $\mathcal{B}$-mesurable donc $\phi(Y)$ est l'espérance conditionnelle de $g(X,Y) \mid \mathcal{B}$

<u>Comment calculer ces espérances conditionnelles ?</u>

- On a vu le cas discret
- cas des variables aléatoires à densité

on suppose que $\begin{cases} X & \text{à valeurs dans } \mathbb{R}^m \\ Y & \text{———— } \mathbb{R}^n \end{cases}$ et que (X,Y) a une

densité $p : \mathbb{R}^{m+n} \longrightarrow \mathbb{R}$ mesurable, positive, $\int_{\mathbb{R}^{m+n}} p(x)\,dx = 1$

Alors X et Y ont une densité. La densité de Y est la fonction

$$q : \mathbb{R}^n \to \mathbb{R}$$
$$y \longmapsto \int_{\mathbb{R}^m} p(x,y)\,dx$$

<u>Prop</u> : pour toute fonction $h : \mathbb{R}^m \to \mathbb{R}^+$, mesurable

$$\mathbb{E}(h(X)\,|\,Y) = \int h(x) \frac{p(x,Y)}{q(Y)}\,dx$$ (variable aléatoire mesurable $/ \sigma(Y) \to$ fonction de Y)

<u>dém</u> : on prend Z v.a ≥ 0, $\sigma(Y)$-mesurable, alors Z s'écrit comme une fonction de Y : $Z = g(Y)$ $\quad g : \mathbb{R}^n \to \mathbb{R}_+$

$$\mathbb{E}(Z\,h(X)) = \mathbb{E}(h(X)\,g(Y)) = \int_{\mathbb{R}^m \times \mathbb{R}^n} h(x)\,g(y)\,p(x,y)\,dx\,dy$$

$$= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} h(x)\,p(x,y)\,dx \right) g(y)\,dy$$

$$= \int_{\mathbb{R}^n} \left(\frac{\int_{\mathbb{R}^m} h(x)\,p(x,y)\,dx}{q(y)} \right) g(y)\,q(y)\,\mathbb{1}_{\{q(y)>0\}}\,dy$$

$$= \int_{\mathbb{R}^n} g(y)\,q(y)\,\mathbb{1}_{\{q(y)>0\}}\,\varphi(y)$$

$$\varphi(y) = \frac{1}{q(y)} \int_{\mathbb{R}^m} h(x)\,p(x,y)\,dx \quad \text{si } q(y) > 0$$
$$= 0 \quad \text{si } q(y) = 0$$

$= \mathbb{E}(g(Y)\varphi(Y)) = \mathbb{E}(Z\,\varphi(Y))$ donc $\varphi(Y) = \mathbb{E}(h(X)\,|\,Y)$

Rappel : dans le cas discret : (X,Y) variables aléatoires discrètes $\forall y$, on peut définir la loi conditionnelle de $X\,|\,Y=y$ et l'espérance conditionnelle est l'espérance de la loi conditionnelle : $\mathbb{E}(X\,|\,Y=y)$

→ on va essayer de généraliser la notion de loi conditionnelle dans le cas non discret.

déf : $(E,\mathcal{E})$, $(F,\mathcal{F})$ espaces mesurables. on dit que

$\nu : E\times\mathcal{F} \longrightarrow [0,1]$ est une probabilité de transition si elle vérifie :

- $\forall x\in E$, $\begin{array}{l}\mathcal{F}\longrightarrow[0,1]\\ B\longmapsto \nu(x,B)\end{array}$ est une mesure de probabilité [probabilité conditionnelle à x]

- $\forall B\in\mathcal{F}$, $\begin{array}{l}E\longrightarrow[0,1]\\ x\longmapsto \nu(x,B)\end{array}$ est $\mathcal{E}$-mesurable

remarque : cas précédent $p(x,y)$ $p : \mathbb{R}^m\times\mathbb{R}^n\longrightarrow\mathbb{R}^+$; $x\in\mathbb{R}^m$, $V\in\mathcal{B}(\mathbb{R}^n)$

$\nu(x,B) = \int_B p(x,y)\,dy$ vérifie ces conditions

Prop : (1) si h mesurable positive sur $(F,\mathcal{F})$, alors

$\varphi : x\longmapsto \int \nu(x,dy)\,h(y)$ est une fonction mesurable positive sur E

(2) si λ est une mesure de proba sur $(E,\mathcal{E})$ alors

$\mu : A\in\mathcal{F}\longmapsto \int\lambda(dx)\,\nu(x,A)$ est une mesure de proba sur $(F,\mathcal{F})$

déf : $\left| \begin{array}{l} X \\ Y \end{array} \right.$ v.a sur $(E, \mathcal{E})$ / $(F, \mathcal{F})$ on appelle loi conditionnelle de Y sachant X une probabilité de transition ν telle que $\forall h$ $F \to \mathbb{R}_+$ $\mathcal{F}$-mesurable, $\mathbb{E}\left(h(Y)|X\right) = \int \nu(X, dy)\, h(y)$

espérance conditionnelle = espérance de la loi conditionnelle

<u>unicité</u> : si ν et ν' sont des lois conditionnelles, $\forall A \in \mathcal{F}$

$$\nu(X, A) = P(Y \in A / X) = \nu'(X, A) \quad \text{p.s}$$

$$\nu(X, A) = \nu'(X, A) \quad \text{p.s}$$

de même, $\underbrace{\nu(x, \cdot)}_{\text{mesure de proba sur } (F, \mathcal{F})} = \nu'(x, \cdot)$ p.s

<u>Th</u> si $(E, \mathcal{E})$, $(F, \mathcal{F})$ sont métriques complets séparables, $\mathcal{E}, \mathcal{F}$ tribus boréliennes, alors il existe une loi conditionnelle de Y sachant X