Inégalité de Jensen, version espérance conditionnelle

$X \in L^1(\Omega, \mathcal{A}, \mathbb{P})$, $\mathcal{B}$ sous-tribu de $\mathcal{A}$

$f : \mathbb{R} \to \mathbb{R}$ convexe, positive

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

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

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

$x \longmapsto x^2$ convexe, positive

$$\mathbb{E}(\mathbb{E}(X / \mathcal{B})^2) \leq \mathbb{E}(\mathbb{E}(X^2 / \mathcal{B})) = \mathbb{E}(X^2) < \infty \to \mathbb{E}(X / \mathcal{B}) \in L^2(\Omega, \mathcal{B}, P)$$

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

$$\mathbb{E}[Z\,(X - \mathbb{E}(X / \mathcal{B}))] = \mathbb{E}(ZX) - \mathbb{E}[Z\,\mathbb{E}(X / \mathcal{B})] = 0$$

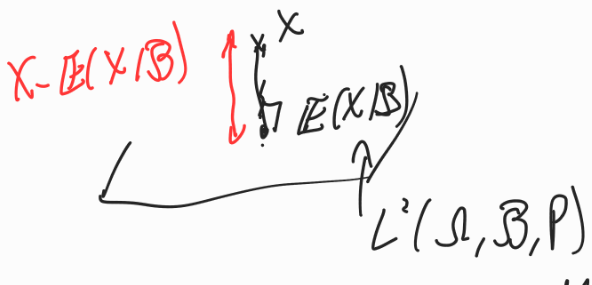
$Q(Z, X - \mathbb{E}(X / \mathcal{B})) = 0$ $\qquad Q(Z, Z') = \mathbb{E}(Z Z')$

vrai $\forall Z$ $\mathcal{B}$ mesurable bornée $\to$ vrai $\forall Z$ $\mathcal{B}$-mesurable, L^2

[densité des v.a. $\mathcal{B}$-mesurables bornées dans $L^2(\Omega, \mathcal{B}, P)$]

$$Z^{(n)} = \sup(\inf(Z, n), -n) \to Z$$

$X - \mathbb{E}(X / \mathcal{B})$ est orthogonal à $L^2(\Omega, \mathcal{B}, P) \to \mathbb{E}(X / \mathcal{B}) = \Pi_{L^2(\Omega, \mathcal{B}, P)}(X)$



. remarque : souvent, on a des variables aléatoires X, Y, et on écrit

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

de même $\mathbb{E}(X | Y_1, Y_2 \ldots Y_n) = \mathbb{E}(X / \sigma(Y_1, Y_2 \ldots Y_n))$

Prop. X, Y v.a. dans L^1 ou ≥ 0, $\mathcal{B}$ sous-tribu, si Y est $\mathcal{B}$-mesurable, alors $\mathbb{E}(YX|\mathcal{B}) = Y\,\mathbb{E}(X|\mathcal{B})$

dem cas où X et $Y \geq 0$ soit Z $\mathcal{B}$-mesurable, positive

$$\mathbb{E}\left(Z\left(Y\mathbb{E}(X|\mathcal{B})\right)\right) = \mathbb{E}\big(\underbrace{(ZY)}_{\mathcal{B}\text{ mesurable}}\mathbb{E}(X|\mathcal{B})\big) = \mathbb{E}(ZYX) = \mathbb{E}(Z\,\mathbb{E}(YX|\mathcal{B}))$$

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

Prop! $\mathcal{B}_1 \subset \mathcal{B}_2$ sous-tribus alors, $\forall X \begin{cases} \in L^1(\Omega, \mathcal{A}, P) \\ \text{ou } X \geq 0 \end{cases}$, $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1) = \mathbb{E}(X|\mathcal{B}_1)$

remarque : si $X \in L^2(\Omega, \mathcal{A}, P)$, $\mathbb{E}(X|\mathcal{B}_1) = \Pi_{L^2(\Omega,\mathcal{B}_1,P)}(X)$

$$= \Pi_{L^2(\Omega,\mathcal{B}_1,P)}\Pi_{L^2(\Omega,\mathcal{B}_2,P)}(X)$$

$$L^2(\Omega,\mathcal{B}_1,P) \subset L^2(\Omega,\mathcal{B}_2,P)$$

dem cas ≥ 0 soit $Z \geq 0$, $\mathcal{B}_1$ mesurable, alors Z est $\mathcal{B}_2$-mesurable

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

Th ① si $\mathcal{B}_1, \mathcal{B}_2$ tribus indépendantes, $\forall X \geq 0$, $\mathcal{B}_2$-mesurable,

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

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

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

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

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

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

$$\mathbb{E}(\mathbb{1}_A / \mathcal{B}_2) = \mathbb{E}(\mathbb{1}_A) = \mathbb{P}(A)$$

si $B \in \mathcal{B}_1$ $\quad \mathbb{P}(A \cap B) = \mathbb{E}(\mathbb{1}_{A\cap B}) = \mathbb{E}(\mathbb{1}_A \cdot \mathbb{1}_B)$

$$= \mathbb{E}\left(\mathbb{E}(\mathbb{1}_A / \mathcal{B}_1)\, \mathbb{1}_B\right)$$

$$= \mathbb{E}(\mathbb{P}(A)\, \mathbb{1}_B) = \mathbb{P}(A)\, \mathbb{E}(\mathbb{1}_B) = \mathbb{P}(A)\mathbb{P}(B)$$

$\to \mathcal{B}_1, \mathcal{B}_2$ indépendantes

<u>Th</u> X, Y v.a $\quad Y$ $\mathcal{B}$ mesurable, X indép de $\mathcal{B}$

X à valeurs dans E
Y $\qquad\qquad\quad F$ $\quad$, $g : E \times F \longrightarrow \mathbb{R}_+$

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

où P_X : loi de X

remarque : si Y déterministe, alors $\int g(x,Y)\, P_X(dx)$ est déterministe
et (1) s'écrit $\mathbb{E}(g(X,Y)/\mathcal{B}) = \underbrace{\mathbb{E}(h(X)/\mathcal{B})}_{\to \text{ déterministe}}$ avec $h(x) = g(x, Y)$

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

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

$$\mathbb{E}\left[Z\, g(X,Y)\right] = \int z\, g(x,y)\, P(dx, dy, dz)$$

$$= \int z\, g(x,y)\, P(dx)\, P(dy, dz)$$

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

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

$\phi(Y)$ $\mathcal{B}$-mesurable $\to \phi(Y) = \mathbb{E}(g(X,Y)/\mathcal{B})$

Application : calcul de l'espérance conditionnelle pour des variables à densité

X à valeurs dans $\mathbb{R}^m$
Y $\qquad\qquad\qquad\quad \mathbb{R}^n$

on suppose que (X,Y) a une densité dans $\mathbb{R}^{m+n}$

on appelle p cette densité : $p : \mathbb{R}^{m+n} \to \mathbb{R}_+$ mesurable $\int_{\mathbb{R}^{m+n}} p(z)\, dz = 1$

$\to$ X et Y ont une densité : densité q de Y est la fonction

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{R}_+ \\ y &\longmapsto \int_{\mathbb{R}^m} p(x,y)\, dx \end{aligned}$$

Prop 1. $h : \mathbb{R}^m \to \mathbb{R}_+$ mesurable, alors

$$\mathbb{E}[h(X) \mid Y] = \int h(x) \frac{p(x,Y)}{q(Y)}\, dx = F(Y)$$

variable aléatoire $\sigma(Y)$-mesurable $\Leftrightarrow$ fonction mesurable de Y

$$F(y) = \int h(x) \frac{p(x,y)}{q(y)}\, dy$$

dém : $Z \geq 0$, $\sigma(Y)$-mesurable $\to$ Z s'écrit $g(Y)$

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

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

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

on note $\varphi(y) = 0$ si $q(y) = 0$

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

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

$\rightarrow \varphi(Y) = \mathbb{E}(h(X) | Y)$

Espérance conditionnelle pour les vecteurs gaussiens

rappels ! . $(X_1, \dots X_d) \in \mathbb{R}^d$ est un vecteur gaussien si toute combinaison linéaire des X_i est une gaussienne

. si $(X_1 \dots X_d)$ vecteur gaussien, il est caractérisé par sa moyenne $M = \mathbb{E}(X_1, \dots, X_d) \in \mathbb{R}^d$ et par sa matrice de covariance Q

$$Q_{ij} = \mathrm{cov}(X_i, X_j)$$

. si $(X_1 \dots X_d)$ vecteur gaussien, il est de carré intégrable

$\rightarrow$ si $Y \in L^2$, $\mathbb{E}(Y | X_1, \dots, X_d) = \mathbb{E}(Y | \sigma(X_1 \dots X_d))$ est la projection orthogonale sur $L^2(\Omega, \sigma(X_1 \dots X_d), P)$

Prop ! si $(X_1 \dots X_m, Y_1 \dots Y_n)$ vecteur gaussien $\in \mathbb{R}^{m+n}$, alors $X = (X_1 \dots X_m)$ et $Y = (Y_1 \dots Y_n)$ sont des vecteurs gaussiens et X et Y sont indépendants si et seulement si $\{ \forall i \in [1, m], \forall j \in [1, n], \mathrm{cov}(X_i Y_j) = 0 \quad (*) \}$

dém : $\mathbb{E} \exp i \langle \xi, X_1 \dots X_m, Y_1 \dots Y_n \rangle = \exp\left(\frac{1}{2} {}^t\xi Q \xi\right)$

si $(*)$ alors $Q = \begin{pmatrix} \overbrace{\quad}^{m} & 0 \\ 0 & \underbrace{\quad}_{n} \end{pmatrix}$ $\hookrightarrow \mathbb{E} \exp i \langle \xi^{(1)}, X_1 \dots X_m \rangle$

$\mathbb{E} \exp i \langle \xi^{(2)}, Y_1 \dots Y_n \rangle$

$\xi^{(1)} = (\xi_1 \dots \xi_m)$ $\quad \xi^{(2)} = (\xi_{m+1} \dots \xi_{m+n})$

$\rightarrow (X_1 \dots X_m)$ indépendant de $(Y_1 \dots Y_n)$

Th (1) $(X, Y_1, Y_2 \dots Y_n)$ vecteur gaussien centré (espérance nulle)

alors $\mathbb{E}(X | Y_1 \dots Y_n) = \Pi_V(X)$ où V est l'espace vectoriel engendré par $Y_1 \dots Y_n$. on notera $\Pi_V(X) = \lambda_1 Y_1 + \dots + \lambda_n Y_n$

(2) si $h : \mathbb{R} \to \mathbb{R}_+$ mesurable, alors

$$\mathbb{E}(h(X) \mid Y_1 \ldots Y_n) = \int_{\mathbb{R}} h(x) \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{(x-m)^2}{2\sigma^2}\right) dx$$

$\sigma^2 = \mathbb{E}(X - \pi_V(X))^2$ [carré de la distance de X à V]

$m = \sum \lambda_i Y_i = \pi_V(X)$

<u>dém</u> (1) on peut supposer X non mesurable $/ \sigma(Y_1 \ldots Y_n)$

$\hat{X} = \pi_V(X) = \lambda_1 Y_1 \ldots + \lambda_n Y_n$

$\forall i, \quad (X - \hat{X}) \perp Y_i$

$$\operatorname{cov}(X - \hat{X}, Y_i) = \mathbb{E}((X - \hat{X}) Y_i) = 0$$

donc $(X - \hat{X}, Y_1 \ldots Y_n)$ vecteur gaussien et $X - \hat{X}$ est indépendant de $(Y_1, \ldots Y_n)$

$$\mathbb{E}(X \mid Y_1 \ldots Y_n) = \mathbb{E}((X - \hat{X}) + \hat{X} \mid Y_1 \ldots Y_n) = \mathbb{E}(X - \hat{X} \mid Y_1 \ldots Y_n) + \mathbb{E}(\hat{X} \mid Y_1 \ldots Y_n)$$

. $\hat{X}$ est fonction de $Y_1 \ldots Y_n$, donc mesurable $/ \sigma(Y_1 \ldots Y_n)$

donc $\mathbb{E}(\hat{X} \mid Y_1 \ldots Y_n) = \hat{X}$

. $X - \hat{X}$ est indépendant de $(Y_1 \ldots Y_n)$ donc

$$\mathbb{E}(X - \hat{X} \mid Y_1 - Y_n) = \mathbb{E}(X - \hat{X}) = 0$$

$\to \mathbb{E}(X / Y_1 \ldots Y_n) = \hat{X} = \pi_V(X)$

(2) $\mathbb{E}(h(X) \mid Y_1 \ldots Y_n) = \mathbb{E}\left[h((X - \hat{X}) + \hat{X}) \mid Y_1 \ldots Y_n)\right]$

on note $Z = X - \hat{X}$

$$= \mathbb{E}(h(Z + \hat{X}) \mid Y_1 \ldots Y_n]$$

$$= \int h(\hat{X} + z) \underbrace{P_Z(dz)}_{\text{loi de } Z}$$

$= \pi_V(X) = \lambda_1 Y_1 \ldots + \lambda_n Y_n$

existence de lois conditionnelles

- rappel si (X,Y) couple de v.a. discrètes, si $P(Y=y)>0$
  on peut définir la variable aléatoire $X_y = X \mid Y=y$ et
  $$\mathbb{E}(X \mid Y=y) = \mathbb{E}(X_y)$$

on essaye de généraliser

- <u>def</u> $(E,\mathcal{E})$, $(F,\mathcal{F})$ espaces mesurables. On dit que
  $\nu : E \times \mathcal{F} \longrightarrow [0,1]$ est une probabilité de transition si

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

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

cas important : pour une variable aléatoire ayant une densité $p : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}_+$

si on pose $\nu(x,B) = \dfrac{\int_B p(x,y)\,dy}{\int_{\mathbb{R}^n} p(x,y)\,dy}$ $\quad x \in \mathbb{R}^m$, B borélien de $\mathbb{R}^n$

ν est une probabilité de transition [si $\int_{\mathbb{R}^n} p(x,y)\,dy = 0$, $\nu(x,\cdot) = \delta_y$, $y \in F$]

<u>Prop</u> : (1) si h fonction mesurable positive $(F,\mathcal{F}) \longrightarrow \mathbb{R}_+$, alors
$$\varphi : x \longmapsto \int \nu(x,dy)\,h(y) \text{ est mesurable } \geq 0$$
$$(E,\mathcal{E}) \longrightarrow \mathbb{R}_+$$

(2) si λ mesure de probabilités sur $(E,\mathcal{E})$, alors
$$\mu : A \in \mathcal{F} \longmapsto \int \lambda(dx)\,\nu(x,A) \text{ est une mesure de proba}$$
$$(F,\mathcal{F}) \longrightarrow [0,1]$$

déf : X v.a sur $(E,\mathcal{E})$ une loi conditionnelle de Y sachant X est
Y $(F,\mathcal{F})$

une probabilité de transition ν telle que pour toute fonction $h : F \to \mathbb{R}_+$ mesurable $/ \mathcal{F}$, $\mathbb{E}(h(Y)|X) = \int h(y)\, \nu(X, dy)$

remarque : unicité de la loi conditionnelle . ν et ν' lois conditionnelles, $\forall A \in \mathcal{F}$

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

$$\rightsquigarrow \nu(x,A) = \nu'(x,A) \quad p.s.$$

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

Th si $(E,\mathcal{E})$ et $(F,\mathcal{F})$ sont métriques complets séparables , $\mathcal{E}, \mathcal{F}$ tribus boréliennes alors il existe une loi conditionnelle.