Cours de CALCUL STOCHASTIQUE

Ciprian TUDOR
Université de Panthéon-Sorbonne Paris 1

October 6, 2007

MASTER M2: Mathématiques Appliquées à l'Economie et à la Finance

CHAPITRE 1

ATTENTION: Ces notes de cours représentent une version rémaniée du cours enseigné par Bernard De Meyer dans le cadre du DEA MMME 2004-2005, à l'Université de Paris 1.

1 CHAPITRE I: Introduction et rappels

1.1 Variables Aléatoires: définition, loi, fonction caractéristique, indépendance, espaces L^p

Le cadre général est le suivant: on considère un espace probabilisé $(\Omega, \mathcal{F}, P)$ où:

Ω est un ensemble non vide, appelé "univers", qui représente l'ensemble des résultats possibles d'une expérience

$\mathcal{F}$ est une σ-algèbre (tribu), c.à. d. $\mathcal{F} \subset \mathcal{P}(\Omega)$ et

a) $\Phi \in \mathcal{F}$

b) si $A \in \mathcal{F}$, alors $A^c \in \mathcal{F}$.

c) si $A_1, A_2, \ldots \in \mathcal{F}$, alors $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

P est une probabilité, c.à.d P est une application $P : \mathcal{F} \to [0,1]$ qui satisfait:

a) $P(\Omega) = 1$

b) si $A_1, A_2, \ldots \in \mathcal{F}$ sont disjoints, alors $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$.

Definition 1 *Une variable aléatoire est une application X mesurable de $(\Omega, \mathcal{F})$ vers $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$.*

On rappelle que X est mesurable si et seulement si, pour tout $B \in \mathcal{B}_{\mathbb{R}^d}$, $X^{-1}(B) := \{\omega | X(\omega) \in B\}$ est un élément de $\mathcal{F}$.

La variable X "transporte" la probabilité P sur $(\Omega, \mathcal{F})$ en une probabilité P_X sur $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$:

$$\forall B \in \mathcal{B}_{\mathbb{R}^d} : P_X(B) := P(X^{-1}(B))$$

P_X est appelée loi de la v.a X.

Exercice 1.1 *Montrez que P_X est une probabilité sur $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$.*

Exercice 1.2 *Montrez que $\sigma(X) := \{X^{-1}(B) | B \in \mathcal{B}_{\mathbb{R}^d}\}$ est une σ-algèbre d'ensembles. C'est la plus petite tribu sur Ω qui rend X mesurable.*

Exercice 1.3 *Montrez que si $\{\mathcal{F}_\lambda\}_{\lambda \in \Lambda}$ est une famille de σ-algèbres sur Ω, alors $\cap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ est également une σ-algèbre.*

En particulier, si $\mathcal{G} \subset \mathcal{P}(\Omega)$, on définit $\sigma(\mathcal{G})$

$$\sigma(\mathcal{G}) = \bigcap_{\mathcal{F} \textit{tribu} \supset \mathcal{G}} \mathcal{F}.$$

$\sigma(\mathcal{G})$ est donc la plus petite σ-algèbre qui contient $\mathcal{G}$.

Définition 1.4 *On définit l'espérance d'une v.a.$X \geq 0$ par:*

$$E(X) = \int_{\Omega} X(\omega)\, dP(\omega)$$

Si $p \in [1, \infty[$, on définit $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ comme l'espace des v.a. X telles que $\|X\|_{L^p}^p := E(|X|^p) < \infty$.

Sur $\mathcal{L}^1(\Omega, \mathcal{F}, P)$, on définit par

$$E(X) = \int_{\Omega} X(\omega)\, dP(\omega)$$

Exercice 1.5 *Montrez que $\|X\|_{L^p} = 0$ est équivalent à $X = 0$ P-pp.*

$\|.\|_{L^p}$ ne peut donc être une norme sur $\mathcal{L}^p(\Omega, \mathcal{F}, P)$. On définit:

Définition 1.6 *$L^p(\Omega, \mathcal{F}, P)$ est l'espace des classes d'équivalence de v.a. de $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ pour la relation d'équivalence = P-pp (presque partout).*

Théorème 1.7 *Pour tout $p \in [1, \infty[$, l'espace $L^p(\Omega, \mathcal{F}, P)$ muni de la norme $\|.\|_{L^p}$ est un espace de Banach.*

$L^2(\Omega, \mathcal{F}, P)$ muni du produit scalaire $\langle X, Y \rangle := E[X \cdot Y]$ est un espace de Hilbert

Exercice 1.8 *Montrez que si $X \in L^1(\Omega, \mathcal{F}, P)$, alors*

$$E(X) = \int_{\mathbb{R}^d} x\, dP_X(x) :$$

L'espérance d'une v.a. ne dépend que de sa loi.

Définition 1.9 *Si X et Y sont des v.a. de $L^2(\Omega, \mathcal{F}, P)$, on définit*

$$var[X] := E[(X - E[X])^2] \text{ et } cov[X, Y] := E[(X - E[X]) \cdot (Y - E[Y])].$$

Si Z est un vecteur aléatoire, considéré comme une matrice colonne, dont toutes les composantes sont dans $L^2(\Omega, \mathcal{F}, P)$, la matrice de covariance $var[Z]$ est définie par

$$var[Z] := E[(Z - E[Z])(Z - E[Z])^\top].$$

Exercice 1.10 *Montrez que si $v \in \mathbb{R}^d$, alors $var[\langle v, Z \rangle] = v^\top var[Z] v$. En déduire que $var[Z]$ est une matrice définie positive.*

Exercice 1.11 Inégalité de Tchebychev: *si $\lambda > 0$, alors $\forall p \geq 1$*

$$P(|X| > \lambda) \leq \frac{1}{\lambda^p} E(|X|^p).$$

Inégalité de Schwartz:
$$E(XY) \leq \sqrt{E(X^2)E(Y^2)}$$

Inégalité de Hölder:
$$E(XY) \leq [E|X|^p]^{\frac{1}{p}}[E|Y|^q]^{\frac{1}{q}} \qquad si\ \frac{1}{p} + \frac{1}{q} = 1.$$

Inégalité de Jensen: *si $f : \mathbb{R}^d \to \mathbb{R}$ est une fonction convexe et Y est un vecteur aléatoire d dimensionnel, alors $E[f(Y)] \geq f(E[Y])$.*

En déduire que si $\alpha > \beta$ alors $\|X\|_{L^\alpha} \geq \|X\|_{L^\beta}$. En particulier $L^\alpha(\Omega, \mathcal{F}, P) \subset L^\beta(\Omega, \mathcal{F}, P)$

Définition 1.12 *Une variable aléatoire Z à valeurs réelles suit une loi normale centrée réduite, ce que l'on note $Z \sim \mathcal{N}(0,1)$ si sa densité par rapport à la mesure de Lebesgue s'écrit sous la forme:*
$$f_Z(z) = \frac{1}{\sqrt{2\pi}} exp(\frac{-z^2}{2})$$

Définition 1.13 *Une variable aléatoire Y suit une loi normale de moyenne μ et de variance σ^2 si $Y = \mu + \sigma Z$ où Z suit une loi normale centrée et réduite. On note alors $Y \sim \mathcal{N}(\mu, \sigma^2)$. μ et σ^2 sont respectivement l'espérance et la variance de la v. a. Y.*

Définition 1.14 *La fontion caractéristique φ_Z d'une variable aléatoire Z à valeurs dans $\mathbb{R}^d$ est définie par: $\varphi_Z : \mathbb{R}^d \to \mathcal{C} : \alpha \to \varphi_Z(\alpha)$:*
$$\varphi_Z(\alpha) = E(exp(i \cdot \langle \alpha, Z \rangle)).$$

Théorème 1.15 *La fonction caractéristique caractérise la loi: Si les variables aléatoires X et Y ont mêmes fonctions caractéristiques (i.e. $\varphi_X = \varphi_Y$), alors elles ont mêmes lois: $P_X = P_Y$.*

Exercice 1.16 *Montrez que si Z est une variable normale centrée réduite alors $\varphi_Z(\alpha) = \exp(-\alpha^2/2)$. Calculez $\varphi_Y(\alpha)$ lorsque $Y \sim \mathcal{N}(\mu, \sigma^2)$.*

Définition 1.17 *Deux événements A et B de $\mathcal{F}$ sont indépendants si $P(A \cap B) = P(A) \cdot P(B)$. On écrit alors $A \perp\!\!\!\perp B$. Deux sous tribus $\mathcal{B}$ et $\mathcal{C}$ de $\mathcal{F}$ sont indépendantes, ce que l'on note $\mathcal{B} \perp\!\!\!\perp \mathcal{C}$, si $\forall B \in \mathcal{B}, \forall C \in \mathcal{C} : B \perp\!\!\!\perp C$. Deux variables aléatoires X et Y sur Ω sont indépendantes si $\sigma(X) \perp\!\!\!\perp \sigma(Y)$.*

Exercice 1.18 *Montrez que $X \perp\!\!\!\perp Y$ ssi $\forall \alpha, \beta$:*
$$\varphi_{(X,Y)}(\alpha, \beta) = \varphi_X(\alpha) \cdot \varphi_Y(\beta)$$

Exercice 1.19 *Montrez que si $X \perp\!\!\!\perp Y$, alors $cov(X,Y) = 0$. Donnez un exemple de vecteur (X,Y) tel que $cov(X,Y) = 0$, mais tel que X et Y ne soient pas indépendants.*

Exercice 1.20 *Si $\mathcal{B}$ est une σ-algèbre d'ensembles et $\mathcal{C}$ une algèbre d'ensembles qui vérifient $\forall B \in \mathcal{B}, \forall C \in \mathcal{C} : B \perp\!\!\!\perp C$, alors $\mathcal{B} \perp\!\!\!\perp \sigma(\mathcal{C})$.*

Exercice 1.21 *si $Y_1 \sim N(\mu_1, {\sigma^2}_1)$ et $Y_2 \sim N(\mu_2, {\sigma^2}_2)$ sont deux v.a. indépendantes, calculer φ_S où $S := Y_1 + Y_2$. En déduire que $S \sim N(\mu_1 + \mu_2, {\sigma^2}_1 + {\sigma^2}_2)$.*

Types de convergence pour les variables aléatoires:

Convergence en loi: Une suite de variables aléatoires $\{X_n\}_{n \in I\!N}$ à valeurs dans $I\!R^d$ converge en loi vers X ssi pour toute fonction continue bornée $f : I\!R^d \to I\!R$, on a la convergence de $E[f(X_n)]$ vers $E[f(X)]$.

Convergence en probabilité: Une suite de variables aléatoires $\{X_n\}_{n \in I\!N}$ à valeurs dans $I\!R^d$ converge en probabilité vers X si pour tout $\varepsilon > 0$,

$$P\left(|X_n - X| > \varepsilon\right) \to_{n\to\infty} 0$$

Convergence presque sûre: Une suite de variables aléatoires $\{X_n\}_{n \in I\!N}$ à valeurs dans $I\!R^d$ converge presque surement vers X si

$$X_n(\omega) \to_{n\to\infty} X(w).$$

pour tout $\omega \notin N$ où $P(N) = 0$.

Convergence dans L^p, $p \geq 1$: Une suite de variables aléatoires $\{X_n\}_{n \in I\!N}$ à valeurs dans $I\!R^d$ converge dans L^p vers X si

$$E\left(|X_n - X|^p\right) \to_{n\to\infty} 0.$$

Exercice 1.22 *Rappeller les relations entre les différentes types de convergence pour les variables aléatoires.*

Exercice 1.23 *Montrez que si $X_n \stackrel{Loi}{\to} X$, et $P[X = a] = 0$, alors $P[X_n \leq a] \to P[X \leq a]$, lorsque $n \to \infty$.*

Théorème 1.24 *Si X_n est une suite de variables aléatoire à valeurs dans $I\!R^d$, alors $X_n \stackrel{Loi}{\to} X$ ssi*

$$\forall \alpha \in I\!R^d : \varphi_{X_n}(\alpha) \to \varphi_X(\alpha).$$

Le théorème précédent permet de montrer aisément le théorème central limite:

Théorème 1.25 *(théorème central limite TCL) Si* $\{X_n\}_{n\in\mathbb{N}}$ *est une suite de variables de* $L^2(\Omega,\mathcal{F},P)$*, indépendantes, identiquement distribuées d'espérance nulle et de variance 1, alors*

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i$$

converge en loi vers $\mathcal{N}(0,1)$.

Ce théorème explique l'apparition fréquente de la loi normale dans la nature.

Définition 1.26 *Un vecteur aléatoire* Y *dans* $\mathbb{R}^d$ *est gaussien si* $\forall v \in \mathbb{R}^d$, $< v, Y >$ *est une v.a gaussienne.*

Exercice 1.27 *Montrez que si* Y *est un vecteur gaussien tel que* $E(Y) = \mu$ *et la matrice de variance covariance* $var(Y) = \Sigma$*, alors*

$$\varphi_Y(v) = exp[i < v, \mu > -\frac{v^t \Sigma\, v}{2}]$$

(les vecteurs v *de* $\mathbb{R}^d$ *sont considérés dans cette expression comme des matrices colonnes). Ceci montre que la loi d'un vecteur gaussien est entièrement déterminée par* μ *et* Σ.

1.2 Processus aléatoires:

Définition 1.28 *Un processus aléatoire sur* $(\Omega,\mathcal{F},P)$*, indicé par un ensemble de temps* $T \subset \mathbb{R}$ *est une famille* $\{X_t\}_{t\in T}$ *de v.a aléatoires de* $(\Omega,\mathcal{F},P)$.

X_t représente l'état du processus au temps T.

A ω fixé, l'application $t \to X_t(\omega)$ s'appelle trajectoire du processus.

On peut définir différentes relations d'équivalence entre deux processus ainsi:

Définition 1.29 *Les processus* X *et* Y *sur* $(\Omega,\mathcal{F},P)$ *sont des modifications l'un de l'autre (notation* $Y \overset{modif}{\equiv} Y$*) si:*

$$\forall t \geq 0, X_t(\omega) = Y_t(\omega)\ P - pp.$$

Définition 1.30 *Les processus* X *et* Y *sur* $(\Omega,\mathcal{F},P)$ *sont indistinguables si et seulement si:*

$$\{\omega |\, \forall\, 0 \leq t < \infty,\ X_t(\omega) = Y_t(\omega)\}$$

est mesurable et a pour probabilité 1.

Définition 1.31 *Les processus* X *et* Y *sur* $(\Omega,\mathcal{F},P)$ *ont mêmes lois fini-dimensionnelles si:*

$$\forall\, 0 \leq t_1 < t_2 < \ldots < t_n < \infty, (X_{t_1}, \ldots, X_{t_n})\ et\ (Y_{t_1}, \ldots, Y_{t_n})$$

ont la même loi.

Exercice 1.32 *Monter que, si X et Y sont indistinguables, alors ils sont des modifications l'un de l'autre.*

Exercice 1.33 *Montrer que la réciproque est fausse.*

Preuve: En effet soit $\Omega = [0, 1]$ et P la mesure de Lebesgue, posons:

$$X_t(\omega) = \begin{cases} 1 & \text{si } t = \omega \\ 0 & \text{sinon.} \end{cases}$$

et $Y_t(\omega) = 0$.
Alors on a :$P(\{\omega | X_t(\omega) \neq Y_t(\omega)\})$=$P(\{t = \omega\}) = 0$ donc $X_t(\omega) = Y_t(\omega)$ P-pp d'où $X_t(\omega)$ et $Y_t(\omega)$ sont des modifications l'un de l'autre.

Par contre

$$\{\omega | \, \forall\, 0 \leq t < \infty, X_t(\omega) = Y_t(\omega)\} = \{\omega | \, \forall\, 0 \leq t < \infty, t \neq \omega\} = \emptyset$$

donc $P(\{\omega | X_t(\omega) = Y_t(\omega)\}) = 0 \neq 1$ d'où X_t et Y_t ne sont pas indistinguables . ■

Exercice 1.34 *Montrez que si deux processus ont toutes leurs trajectoires continues (même continues à droite), alors $X \overset{modif}{\equiv} Y$ implique que X et Y sont indistinguables.*

Preuve: Soit

$$N_t = \{\omega \in \Omega / X_t(\omega) \neq Y_t(\omega)\}$$

et

$$N = \bigcup_{t \in Q} N_t$$

Q étant les rationels de $[0, \infty)$. Alors

$$P(N) = 0$$

donc $X_t(\omega) = Y_t(\omega), \forall t \in Q$, $\forall \omega \notin N$. Si $t \notin Q$, on considère t_n une suite de Q telle que $t_n \downarrow t$. Alors, pour tout $\omega \notin N$, on a $X_{t_n} = Y_{t_n}$ pour tout n et donc

$$X_t(\omega) = \lim_{n \to \infty} X_{t_n}(\omega) = \lim_{n \to \infty} Y_{t_n}(\omega) = Y_t(\omega)$$

et par conséquent X et Y sont indistinguable. ■

Exercice 1.35 *(***A faire en TD***) Montrez que si deux processus sont des modifications l'un de l'autre, alors ils ont mêmes lois fini-dimensionnelles.*

1.3 Filtration

Définition 1.36 *Une filtration $\{\mathcal{F}_t\}_{t\in T}$ est une famille croissante de sous-σ-algèbres de $\mathcal{F}$. C'est à dire $\forall 0 \leq s < t < \infty$ $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$.*

$\mathcal{F}_t$ représente le niveau de connaissance du processus X au temps t, c.à. d $\mathcal{F}_t$ est la classe des événements que l'on peut "identifier" au temps t.

Définition 1.37 *Si $\mathcal{F}_t$ est une filtration, alors on note:*

- $\mathcal{F}_\infty = \sigma(\bigcup_{t\geq 0} \mathcal{F}_t)$
- $\mathcal{F}_{t^+} = \bigcap_{\epsilon>0} \mathcal{F}_{t+\epsilon}$
- $\mathcal{F}_{t^-} = \sigma(\bigcup_{s<t} \mathcal{F}_s)$.

Définition 1.38 *Une filtration $\{\mathcal{F}_t\}$ est continue à droite (respectivement à gauche) si $\mathcal{F}_t = \mathcal{F}_{t^+}$ (respectivement $\mathcal{F}_t = \mathcal{F}_{t^-}$) $\forall t \geq 0$.*

Définition 1.39 *Une filtration $\{\mathcal{F}_t\}$ est complète si $\forall A \in \mathcal{F}_\infty$ tel que $P(A) = 0$, $A \in \mathcal{F}_0$.*

Définition 1.40 *Le processus X est adapté à la filtration $\{\mathcal{F}_t\}_{t\geq 0}$ si et seulement si $\forall$ $0 \leq t < \infty$, X_t est $\mathcal{F}_t$-mesurable.*

Définition 1.41 *Soit $\mathcal{F}$ une sous σ-algèbre de $\mathcal{G}$ et soit $\mathcal{N}$ la classe des ensembles de $\mathcal{G}$ de probabilité nulle. Alors la tribu $\bar{\mathcal{F}} := \sigma(\mathcal{F} \cup \mathcal{N})$ est appelée complètion de la tribu $\mathcal{F}$.*

Exercice 1.42 *Montrez que toute variable $\bar{X}$ $\bar{\mathcal{F}}$-mesurable est égale P-pp à une variable X $\mathcal{F}$-mesurable.*

Définition 1.43 *La filtration $\{\bar{\mathcal{F}}_t\}_{t\in T}$ est appelée la complétion de $\{\mathcal{F}_t\}_{t\in T}$.*

Exercice 1.44 *Montrez que toute modification X' d'un processus X adapté à $\mathcal{F}_t$ est adapté à $\{\bar{\mathcal{F}}_t\}_{t\in T}$*

1.4 Espérance conditionnelle:

La probabilité conditionnelle d'un événement $A \in \mathcal{F}$ par un événement $B \in \mathcal{F}$ avec $P(B) > 0$ est définie par

$$P(A|B) = \frac{P(A \bigcap B)}{P(B)}.$$

On voit que A et B sont indépendants ssi $P(A|B) = P(A)$. La probabilité conditionnelle $P(A|B)$ représente la probabilité que A se réalise sachant que B s'est réalisé. L'application

$$A \to P(A|B)$$

définit alors une nouvelle probabilité sur $\mathcal{F}$. On peut alors définir l'espérance conditionnelle d'une v.a. X par rapport à un événement B par

$$E(X|B) = \frac{1}{P(B)} E(X1_B).$$

Un concept plus compliqué est le conditionnement par une σ algèbre $\mathcal{B} \subset \mathcal{F}$ d'une v.a. X.

Exercice 1.45 *Soit X une variable aléatoire de $L^2(\Omega, \mathcal{F}, P)$. Montrez que $C := E[X]$ est la constante qui approxime le mieux X au sens $L^2(\Omega, \mathcal{F}, P)$.*

Interprétons une sous σ-algèbre $\mathcal{B}$ de $\mathcal{F}$ comme un "niveau d'information" concernant l'expérience aléatoire: A ce niveau d'information, on peut décider de tout événement $B \in \mathcal{B}$ s'il a ou non été réalisé. On peut donc également calculer la valeur de toute fonction $\mathcal{B}$ mesurable. Au vu de l'exercice précédent, il est assez naturel de définir l'espérance conditionnelle comme suit:

Définition 1.46 *Soit $\mathcal{B} \subset \mathcal{F}$ et $X \in L^2(\mathcal{F})$. La variable dans $L^2(\mathcal{B})$ qui approxime le mieux X au sens de L^2 est notée $E[X|\mathcal{B}]$ et se lit espérance conditionnelle de X sachant $\mathcal{B}$.*

$E[X|\mathcal{B}]$ est ainsi la projection orthogonale de X sur l'espace $L^2(\mathcal{B})$ donc $E[X|\mathcal{B}]$ existe toujours, puisque $L^2(\mathcal{B})$ est un sous espace vectoriel fermé de $L^2(\mathcal{F})$, et est unique au sens L^2.

Les relations d'orthogonalité s'écrivent:

$$\forall Z \in \mathbf{L^2}(\mathcal{B}) : E[ZX] = E[ZE[X|\mathcal{B}]].$$

Exercice 1.47 *Montrez que, si $X \geq Y$ $P-pp$, alors $E[X|\mathcal{B}] \geq E[Y|\mathcal{B}]$ $P-pp$.*

Idée de la preuve: On pose

$$A_\varepsilon = \{E[Y|\mathcal{B}] - E[X|\mathcal{B}] \geq \varepsilon > 0\}$$

et on montre aisément que

$$P(A_\varepsilon) = 0.$$

■

Exercice 1.48 *(**A faire en TD**)Si $U \in L^2(\mathcal{B})$, montrez que $U = E[X|\mathcal{B}]$ si et seulement si $\forall B \in \mathcal{B} : E[U\mathbf{1}_B] = E[X\mathbf{1}_B]$.*

Exercice 1.49 *On considère $(\Omega_j)_{j\geq 1}$ une partition de Ω telle que $P(\Omega_j) > 0$ pour tout $j \geq 1$. Soit $\mathcal{F} = \sigma(\Omega_j, j \geq 1)$. Alors*

$$E(X/\mathcal{F}) = \sum_{j\geq 1} \frac{E(X1_{\Omega_j})}{P(\Omega_j)} 1_{\Omega_j} := Y$$

Preuve: Comme $1_{\Omega_j} \in \mathcal{F}$, on a $\sum_{j\geq 1} \frac{E(X1_{\Omega_j})}{P(\Omega_j)} 1_{\Omega_j} \in \mathcal{F}$. Il suffit de vérifier la propriété d'orthogonalité pour $A = \Omega_n$, n fixé. Mais

$$E(Y1_{\Omega_n}) = E\left(\frac{X1_{\Omega_n}}{P(\Omega_n)} 1_{\Omega_n}\right) = \frac{X1_{\Omega_n}}{P(\Omega_n)} E(1_{\Omega_n}) = E(X1_{\Omega_n}).$$

■

Théorème 1.50 *1. $\forall \alpha, \beta \in I\!R$ $E[\alpha X + \beta Y|\mathcal{B}] = \alpha E[X|\mathcal{B}] + \beta E[X|\mathcal{B}]$.*

2. Si $\mathcal{B} \subset \mathcal{C} \subset \mathcal{F}$, alors $E[X|\mathcal{B}] = E[E[X|\mathcal{C}]|\mathcal{B}]$.

3. Si X est indépendant de $\mathcal{B}$, alors $E[X|\mathcal{B}] = E[X]$.

4. $\forall Z \in L^2(\mathcal{B})$:

$$E[(X - Z)^2] = E[(X - E[X|\mathcal{B}])^2] + E[(E[X|\mathcal{B}] - Z)^2].$$

5. $\forall\, Y \in \mathbf{L}^\infty(\mathcal{B})$: $E[YX|\mathcal{B}] = YE[X|\mathcal{B}]$.

6. Soit $\phi : I\!R \to I\!R$ une fonction convexe telle que $\phi(X) \in \mathbf{L^2}(\mathcal{F})$, alors :

$$E[\phi(X)|\mathcal{B}] \geq \phi(E[X]\mathcal{B}])\ ps.$$

Preuve:

1. Résulte du fait que la projection orthogonale sur $L^2(\mathcal{B})$ est une application linéaire.

2. Soit $Z \in L^2(\mathcal{B})$, alors $E[ZE[X|\mathcal{B}]] = E[ZX]$. Puisque $\mathcal{B} \subset \mathcal{C}$, on a aussi $E[ZX] = E[ZE[X|\mathcal{C}]]$, d'où $E[ZE[X|B]] = E[ZE[X|\mathcal{C}]]$ $\forall Z \in L^2(\mathcal{B})$. Puisque par définition $E[X|\mathcal{B}] \in L^2(\mathcal{B})$, la dernière relation n'est autre que la relation d'orthogonalité qui définit $E[E[X|\mathcal{C}]|\mathcal{B}]$. Ainsi $E[X|\mathcal{B}] = E[E[X|\mathcal{C}]|\mathcal{B}]$.

3. $\forall$ Z $\in L^2(\mathcal{B})$ et X indépendant de $\mathcal{B}$, $E[ZX] = E[X]E[Z] = E[E[X]Z]$, donc $E[X|\mathcal{B}] = E[X]$. Car la seule v.a Y de $\mathbf{L^2}(\mathcal{B})$ vérifiant pour tout $Z \in \mathbf{L^2}(\mathcal{B})$ $E[ZX] = E[ZY]$ est $Y = E[X|\mathcal{B}]$.

4. On a:

$$\begin{aligned} E[(X-Z)^2] &= E[(X-E[X|\mathcal{B}]+E[X|\mathcal{B}]-Z)^2] \\ &= E[(X-E[X|\mathcal{B}])^2]+E[(E[X|\mathcal{B}]-Z)^2] \\ &\quad +2E[(X-E[X|\mathcal{B}])(E[X|\mathcal{B}]-Z)] \end{aligned}$$

Or: $(E[X|\mathcal{B}]-Z) \in L^2(\mathcal{B})$. On en déduit que

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

et la relation est vérifiée .

5. Posons $H = YE[X|\mathcal{B}]$. Comme $Y \in \mathbf{L}^\infty(\mathcal{B})$ et $E[X|\mathcal{B}] \in L^2(\mathcal{B})$, nous avons: $H \in \mathbf{L^2}(\mathcal{B})$. Ainsi, si $Z \in L^2(\mathcal{B})$, on a:

$$E[ZH] = E[ZYE[X|\mathcal{B}]] = E[Z(YX)]$$

(car $ZY \in \mathbf{L^2}(\mathcal{B})$). Donc $H = E[YX|\mathcal{B}] = YE[X|\mathcal{B}]$.

6. Nous nous limitons ici à des fonction ϕ différentiables. On sait que, si ϕ est convexe, alors $\forall y, z$:

$$\phi(z) \geq \phi(y) + \phi'(y)(z-y)$$

donc $\forall \omega$ on a :

$$\phi(X(\omega)) \geq \phi(E[X|\mathcal{B}](\omega)) + \phi'(E[X|\mathcal{B}](\omega))(X(\omega) - E[X|\mathcal{B}](\omega))$$

donc, en prenant l'espérance conditionnelle, nous obtenons:

$$\begin{aligned} E[\phi(X)|\mathcal{B}] &\geq E[(\phi(E[X|\mathcal{B}]) + \phi'(E[X|\mathcal{B}])(X-E[X|\mathcal{B}])|\mathcal{B}] \\ &= \phi(E[X|\mathcal{B}]) + \phi'(E[X|\mathcal{B}])E[(X-E[X|\mathcal{B}])|\mathcal{B}] \end{aligned}$$

Mais

$$E[(X-E[X|\mathcal{B}])|B] = E[X|\mathcal{B}] - E[E[X|\mathcal{B}]|\mathcal{B}] = 0,$$

et partant $E[\phi(X)|B] \geq \phi[E[X|\mathcal{B}]], P-ps.$

Exercice 1.51 *(***A faire en TD***) Sur l'espace probabilisé $(\Omega, \mathcal{F}, P)$, on considère la classe*

$$\mathcal{B} := \{B \in \mathcal{F} | P(B) = 0 \ ou \ P(B) = 1\}.$$

1. *Montrez que $\mathcal{B}$ est une σ-algèbre.*

2. *Caractérisez les fonctions $\mathcal{B}$-mesurables.*

3. *Pour* $X \in L^2(\Omega, \mathcal{F}, P)$, *calculez* $E[X|\mathcal{B}]$.

Exercice 1.52 *(***A faire en TD***) Si* (X,Y) *est un vecteur aléatoire dont la densité* $f_{(X,Y)}$ *vérifie* $\forall x, y \in I\!R : f_{(X,Y)}(x,y) > 0$, *si* $\mathcal{B} := \sigma(Y)$;

1. *Montrez que* $E[X|\mathcal{B}] = g(Y)$, *où*

$$g : I\!R \to I\!R : y \to g(y) := \frac{\int_{I\!R} x f_{(X,Y)}(x,y)dx}{\int_{I\!R} f_{(X,Y)}(x,y)dx}$$

   Idée de la preuve: *On a* $g \in \mathcal{B}(\mathbb{R})$ *et donc* $g(Y) \in \sigma(Y)$.

   Soit maintenant $A \in \sigma(Y)$. *Alors* A *est de la forme* $A = \{\omega, Y(\omega) \in B\}$ *et* $1_A = 1_B(Y)$. *Ensuite*

$$\begin{aligned} E(g(Y)) &= E\left(g(Y)1_B(Y)\right) \\ &= \int_B \int_{\mathbb{R}} g(y) f(x,y) dx dy \\ &= \int_B dy \int_{\mathbb{R}} \left(\frac{z f(z,y)dz}{\int f(z,y)dz} \right) f(x,y)dx \\ &= \int_B dy \int_{\mathbb{R}} z f(z,y)dz = E(Y 1_B(Y)) \end{aligned}$$

2. *Si* (X,Y) *est un vecteur gaussien tel que* $E[X] = E[Y] = 0$, $var[X] = 1 = var[Y]$ *et* $cov[X,Y] = \rho \in [0,1]$, *calculez* $E[X|\mathcal{B}]$. *(voir également l'exercice* **??***).*

Exercice 1.53 *(***A faire en TD***) On prend* $\Omega = (0,1)$, $\mathcal{F}_0 = B((0,1))$ *et* $P = \lambda$ *la mesure de Lebesque. On pose* $X(\omega) = cos(\omega\pi)$ *et*

$$\mathcal{F} = \{A \subset (0,1), A \text{ ou } A^c \text{ denombrable } \}.$$

Alors $E(X/\mathcal{F}) = 0$.

Exercice 1.54 *(***A faire en TD***) Si* $\mathcal{B} \subset \mathcal{C} \subset \mathcal{F}$ *et* $X \in L^2(\Omega, \mathcal{F}, P)$

1. *Montrez que* $E[X|\mathcal{B}] = E[(E[X|\mathcal{C}])|\mathcal{B}]$.
2. *Montrez que la relation précédente n'est en général pas vérifiée si* $\mathcal{B}$ *n'est pas inclus dans* $\mathcal{C}$, *en considérant le point 2 de l'exercice précédent 1.52 avec* $\mathcal{B} := \sigma(X)$ *et* $\mathcal{C} := \sigma(Y)$.

Exercice 1.55 *(***A faire en TD***) La loi d'un vecteur aléatoire* $X = (X_1, \ldots, X_n)$ *est dite échangeable si pour toute permutation* π *de* $\{1, \ldots, n\}$ *les vecteurs* X *et* X_π *ont même loi, où* $X_\pi := (X_{\pi(1)}, \ldots, X_{\pi(n)})$.

Si X *a une loi échangeable, et* $S := X_1 + \ldots + X_n$, *calculez* $E[X_1|S]$ *(*$E[X_1|S]$ *est une notation abrégée pour* $E[X_1|\sigma(S)]$*).*

Exercice 1.56 *(***A faire en TD***)Si $Z_1, Z_2, \ldots$ est une suite i.i.d. de variables $\mathcal{N}(0,1)$, si $\mathcal{F}_n := \sigma(Z_1, \ldots, Z_n)$ et si $X_n := Z_1 + \ldots + Z_n$,*

1. *Montrez que $\forall n \geq m : E[X_n|\mathcal{F}_m] = X_m$.*

2. *Montrez que $\forall n \geq m : E[Y_n|\mathcal{F}_m] = Y_m$, où $Y_n := X_n^2 - n$.*

3. *Montrez que $\forall n \geq m : E[M_n|\mathcal{F}_m] = M_m$, où $M_n := \exp(X_n - n/2)$.*

(Les relations précédentes s'expriment en disant que les processus X, Y et M sont des martingales.)

L'espérance conditionnelle sur L^1

L'espérance conditionnelle $E[X|\mathcal{B}]$ a été définie plus haut pour les variables X de $L^2(\Omega, \mathcal{F}, P)$. Le théorème de Radon Nikodym est utilisé dans l'exercice suivant pour définir l'espérance conditionnelle des variables X de $L^1(\Omega, \mathcal{F}, P)$. Ce théorème affirme que si μ est une mesure signée bornée et P une probabilité sur $(\Omega, \mathcal{B})$, alors les deux énoncés suivants sont équivalents:

1)μ admet une densité Y par rapport à P (i.e. $\exists Y \in L^1(\Omega, \mathcal{B}, P) : \forall U \in L^1(\Omega, \mathcal{B}, \mu) : \int_\Omega U d\mu = E_P[U.Y]$)

2) μ est absolument continue par rapport à P (i.e. $\forall B \in \mathcal{B} : P(B) = 0 \Rightarrow \mu(B) = 0.$)

Exercice 1.57 *Soit $X \in L^1(\Omega, \mathcal{F}, P)$ et $\mathcal{B}$ une sous σ-algèbre de $\mathcal{F}$. Montrez que la mesure μ sur $(\Omega, \mathcal{B})$ définie par $\mu(B) := E_P[X \cdot \mathbf{1}_B]$ est absolument continue par rapport à P. Montrez que la densité Y de μ par rapport à P, qui éxiste par le théorème de Radon Nikodym, vérifie $\forall Z \in L^\infty(\Omega, \mathcal{B}, P)$: $E[Z \cdot X] = E[Z \cdot Y]$ et Y est donc un candidat naturel pour définir $E[X|\mathcal{B}]$.*