

Lecture Notes on STOCHASTIC CALCULUS

Ciprian TUDOR
Université de Panthéon-Sorbonne Paris 1

October 6, 2007

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

CHAPTER 1

Attention: the proofs and the exercises are not yet translated

1 CHAPTER 1: Introduction and preliminaries

1.1 Random variables : definition, law, characteristic function, independence, L^p spaces

The general context is as follows: let (Ω, \mathcal{F}, P) be a probability space, that is, :

Ω is a non-empty set

\mathcal{F} is a σ - algebra, that is $\mathcal{F} \subset \mathcal{P}(\Omega)$ and

- a) $\Phi \in \mathcal{F}$
- b) if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$.
- c) if $A_1, A_2, \dots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

P is a probability, that means that P is an application $P : \mathcal{F} \rightarrow [0, 1]$ that satisfies:

- a) $P(\Omega) = 1$
- b) if $A_1, A_2, \dots \in \mathcal{F}$ are disjoint set , then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Definition 1 A random variable is an application X which is measurable from (Ω, \mathcal{F}) onto $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$.

We recall that X is measurable if and only if for every event $B \in \mathcal{B}_{\mathbb{R}^d}$, the set $X^{-1}(B) := \{\omega | X(\omega) \in B\}$ is an element in \mathcal{F} .

The random variable X "transfers" the probability P on (Ω, \mathcal{F}) into a probability P_X on $(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$:

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

P_X is called the law of the r.v. 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 $\bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ est également une σ -algèbre.

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

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

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

Definition 1.4 We define the expectation of a random variables $X \geq 0$ by:

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

If $p \in [1, \infty[$, we define $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ as the set of r.v. X such that $\|X\|_{L^p}^p := E(|X|^p) < \infty$.
On $\mathcal{L}^1(\Omega, \mathcal{F}, P)$, we introduce

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

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

Remark 1 $\|\cdot\|_{L^p}$ cannot be a norm $\mathcal{L}^p(\Omega, \mathcal{F}, P)$.

We define

Definition 1.6 $L^p(\Omega, \mathcal{F}, P)$ is the set of equivalence classes of $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ with respect to the equivalence relation $= P$ -a.s. (almost surely).

Theorem 1.7 For every $p \in [1, \infty[$, the space $L^p(\Omega, \mathcal{F}, P)$ endowed with the norm $\|\cdot\|_{L^p}$ is a Banach space .

$L^2(\Omega, \mathcal{F}, P)$ endowed by the scalar product $\langle X, Y \rangle := E[X \cdot Y]$ is a Hilbert space.

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.

Definition 1.9 If X and Y are random variables in $L^2(\Omega, \mathcal{F}, P)$, we define

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

If Z is a random vector, (viewed as a column matrix) , whose components are in $L^2(\Omega, \mathcal{F}, P)$, the covariance matrix $\text{var}[Z]$ is given by

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

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

Exercice 1.11 Tchebychev inequality : if $\lambda > 0$, then $\forall p \geq 1$

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

Schwartz inequality :

$$E(XY) \leq \sqrt{E(X^2)E(Y^2)}$$

Hölder inequality :

$$E(XY) \leq [E|X|^p]^{\frac{1}{p}} [E|Y|^q]^{\frac{1}{q}} \quad \text{si } \frac{1}{p} + \frac{1}{q} = 1.$$

Jensen inequality : if $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is a convex function and Y is a random vector (d dimensional), then $E[f(Y)] \geq f(E[Y])$.

Deduce that if $\alpha > \beta$ then $\|X\|_{L^\alpha} \geq \|X\|_{L^\beta}$. In particular $L^\alpha(\Omega, \mathcal{F}, P) \subset L^\beta(\Omega, \mathcal{F}, P)$

Definition 1.12 A random variable Z follows a standard normal law, denoted $Z \sim \mathcal{N}(0, 1)$ if its density with respect to the Lebesgue measure can be written as:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

Definition 1.13 A r.v. Y follows a normal distribution with mean zero μ and variance σ^2 si $Y = \mu + \sigma Z$ where Z is a standard normal r.v.. Then we write $Y \sim \mathcal{N}(\mu, \sigma^2)$. μ and σ^2 are respectively the expectation and the variance of the random variable Y .

Definition 1.14 The characteristic fonction φ_Z of a r.v. Z with values in \mathbb{R}^d is defined by : $\varphi_Z : \mathbb{R}^d \rightarrow \mathbb{C} : \alpha \rightarrow \varphi_Z(\alpha)$:

$$\varphi_Z(\alpha) = E(\exp(i \cdot \langle \alpha, Z \rangle)).$$

Theorem 1.15 The characteristic fonction characterizes the law of the random variables: if the random variables X and Y have the same characteristic fonction (i.e. $\varphi_X = \varphi_Y$), then they have the same law: $P_X = P_Y$.

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 $Y \sim \mathcal{N}(\mu, \sigma^2)$.

Definition 1.17 Two events (sets) A and B in \mathcal{F} are independent if $P(A \cap B) = P(A) \cdot P(B)$. We then write $A \perp\!\!\!\perp B$. Two σ -sub algebras \mathcal{B} and \mathcal{C} in \mathcal{F} are independent (we denote $\mathcal{B} \perp\!\!\!\perp \mathcal{C}$, if $\forall B \in \mathcal{B}, \forall C \in \mathcal{C} : B \perp\!\!\!\perp C$). Two r.v. X and Y on Ω are independent if $\sigma(X) \perp\!\!\!\perp \sigma(Y)$.

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 $\text{cov}(X, Y) = 0$. Donnez un exemple de vecteur (X, Y) tel que $\text{cov}(X, Y) = 0$, 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 $\{X_n\}_{n \in \mathbb{N}}$ à valeurs dans \mathbb{R}^d converge en loi vers X ssi pour toute fonction continue bornée $f : \mathbb{R}^d \rightarrow \mathbb{R}$, on a la convergence de $E[f(X_n)]$ vers $E[f(X)]$.

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

$$P(|X_n - X| > \varepsilon) \rightarrow_{n \rightarrow \infty} 0$$

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

$$X_n(\omega) \rightarrow_{n \rightarrow \infty} X(\omega).$$

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

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

$$E(|X_n - X|^p) \rightarrow_{n \rightarrow \infty} 0.$$

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

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

Theorem 1.24 Si X_n est une suite de variables aléatoire à valeurs dans \mathbb{R}^d , alors $X_n \xrightarrow{Loi} X$ ssi

$$\forall \alpha \in \mathbb{R}^d : \varphi_{X_n}(\alpha) \rightarrow \varphi_X(\alpha).$$

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

Theorem 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.

Definition 1.26 Un vecteur aléatoire Y dans \mathbb{R}^d est gaussien si $\forall v \in \mathbb{R}^d$, $\langle v, Y \rangle$ 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 $\text{var}(Y) = \Sigma$, alors

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

(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 Stochastic processes :

Definition 1.28 A stochastic process (Ω, \mathcal{F}, P) , indexed by a set $T \subset \mathbb{R}$ is a family $\{X_t\}_{t \in T}$ of random variables on (Ω, \mathcal{F}, P) .

X_t represents the situation at time t .

For fixed ω , the application $t \rightarrow X_t(\omega)$ is called the trajectory of the stochastic process X .

We can introduce different types of equivalence relations for stochastic processes:

Definition 1.29 The processes X and Y on (Ω, \mathcal{F}, P) are modifications (notation $Y \stackrel{\text{modif}}{\equiv} X$) if:

$$\forall t \geq 0, X_t(\omega) = Y_t(\omega) \text{ } P - \text{a.s.}$$

Definition 1.30 The processes X and Y on (Ω, \mathcal{F}, P) are indistinguishable if and only if:

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

is measurable and its probability measure is equal to 1.

Definition 1.31 The processes X and Y on (Ω, \mathcal{F}, P) have the same finite-dimensional distributions :

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

ont the same law.

Exercice 1.32 *Montrer 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 fautive.*

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 $X \stackrel{\text{modif}}{\equiv} Y$ 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 rationnels 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

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

et par conséquent X et Y sont indistinguables. ■

Exercice 1.35 (TO BE DONE IN CLASS) *Montrez que si deux processus sont des modifications l'un de l'autre, alors ils ont mêmes lois fini-dimensionnelles.*

1.3 Filtration

Definition 1.36 A filtration $\{\mathcal{F}_t\}_{t \in T}$ is an increasing family of sub- σ -algebras of \mathcal{F} . That is $\forall 0 \leq s < t < \infty \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$.

\mathcal{F}_t represents "how much the observer X knows at time t , i.e. \mathcal{F}_t is a class of events that we can identify at time t .

Definition 1.37 If \mathcal{F}_t is a filtration, then we denote by:

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

Definition 1.38 A filtration $\{\mathcal{F}_t\}$ is right continuous (respectively left continuous) if $\mathcal{F}_t = \mathcal{F}_{t+}$ (respectively $\mathcal{F}_t = \mathcal{F}_{t-}$) $\forall t \geq 0$.

Definition 1.39 A filtration $\{\mathcal{F}_t\}$ is complete if $\forall A \in \mathcal{F}_\infty$ such that $P(A) = 0, A \in \mathcal{F}_0$.

Definition 1.40 The stochastic process X is adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ if and only if $\forall 0 \leq t < \infty, X_t$ is \mathcal{F}_t -measurable.

Definition 1.41 Let \mathcal{F} a sub- σ -algebra of \mathcal{G} and let \mathcal{N} be the class of subsets of \mathcal{G} having its probability measure equal to 0. Then the σ -algebra $\bar{\mathcal{F}} := \sigma(\mathcal{F} \cup \mathcal{N})$ is called the completion of the σ algebra \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.

Definition 1.43 The filtration $\{\bar{\mathcal{F}}_t\}_{t \in T}$ is called the completion of $\{\mathcal{F}_t\}_{t \in T}$.

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

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

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

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

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:

Definition 1.46 Let $\mathcal{B} \subset \mathcal{F}$ and $X \in L^2(\mathcal{F})$. The random variable in $L^2(\mathcal{B})$ that is the closest to X in the L^2 sense is denoted by $E[X|\mathcal{B}]$ and it is called as the conditional expectation of X given \mathcal{B} .

In this way $E[X|\mathcal{B}]$ is the orthogonal projection of X on the space $L^2(\mathcal{B})$ then $E[X|\mathcal{B}]$ always exists because $L^2(\mathcal{B})$ is a closed vector subspace of $L^2(\mathcal{F})$, and it unique in the L^2 sense L^2 .

The orthogonality relationship can be written:

$$\forall Z \in L^2(\mathcal{B}) : E[Z(X - E[X|\mathcal{B}])] = 0.$$

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 (TO BE DONE IN CLASS) Si $U \in L^2(\mathcal{B})$, montrez que $U = E[X|\mathcal{B}]$ si et seulement si $\forall B \in \mathcal{B} : E[U1_B] = E[X1_B]$.

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

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

■

Theorem 1.50 1. $\forall \alpha, \beta \in \mathbb{R} E[\alpha X + \beta Y | \mathcal{B}] = \alpha E[X | \mathcal{B}] + \beta E[Y | \mathcal{B}]$.

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

3. If X is independent by \mathcal{B} , then $E[X | \mathcal{B}] = E[X]$.

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. Let $\phi : \mathbb{R} \rightarrow \mathbb{R}$ be a convex function such that $\phi(X) \in \mathbf{L}^2(\mathcal{F})$, then :

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

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

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

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

Exercice 1.51 (TO BE DONE IN CLASS) Sur l'espace probabilisé (Ω, \mathcal{F}, P) , on considère la classe

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

1. Montrez que \mathcal{B} est une σ -algèbre.
2. Caractériser les fonctions \mathcal{B} -mesurables.

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

Exercice 1.52 (TO BE DONE IN CLASS) Si (X, Y) est un vecteur aléatoire dont la densité $f_{(X,Y)}$ vérifie $\forall x, y \in \mathbb{R} : f_{(X,Y)}(x, y) > 0$, si $\mathcal{B} := \sigma(Y)$;

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

$$g : \mathbb{R} \rightarrow \mathbb{R} : y \rightarrow g(y) := \frac{\int_{\mathbb{R}} x f_{(X,Y)}(x, y) dx}{\int_{\mathbb{R}} f_{(X,Y)}(x, y) dx}$$

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

$$\begin{aligned} E(g(Y)) &= E(g(Y)1_B(Y)) \\ &= \int_B \int_{\mathbb{R}} g(y) f(x, y) dx dy \\ &= \int_B dy \int_{\mathbb{R}} \left(\frac{z f(z, y) dz}{\int f(z, y) dz} \right) f(x, y) dx \\ &= \int_B dy \int_{\mathbb{R}} z f(z, y) dz = E(Y1_B(Y)) \end{aligned}$$

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

Exercice 1.53 (TO BE DONE IN CLASS) On prend $\Omega = (0, 1)$, $\mathcal{F}_0 = \mathcal{B}((0, 1))$ et $P = \lambda$ la mesure de Lebesgue. On pose $X(\omega) = \cos(\omega\pi)$ et

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

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

Exercice 1.54 (TO BE DONE IN CLASS) 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 (TO BE DONE IN CLASS) La loi d'un vecteur aléatoire $X = (X_1, \dots, X_n)$ est dite échangeable si pour toute permutation π de $\{1, \dots, n\}$ les vecteurs X et X_π ont même loi, où $X_\pi := (X_{\pi(1)}, \dots, X_{\pi(n)})$.

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

Exercice 1.56 (TO BE DONE IN CLASS) Si Z_1, Z_2, \dots est une suite i.i.d. de variables $\mathcal{N}(0, 1)$, si $\mathcal{F}_n := \sigma(Z_1, \dots, Z_n)$ et si $X_n := Z_1 + \dots + Z_n$,

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.)

The conditional expectation on 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 (Ω, \mathcal{B}) , alors les deux énoncés suivants sont équivalents:

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

Exercice 1.58 Soient X_1, \dots, X_n , n variables indépendantes uniformes sur $[0, 1]$. Soit $T := \max_{i=1, \dots, n} X_i$. Calculez $E[X_1 | T]$. (Piste: $T = T' \cdot \mathbf{1}_{\{X_1 < T'\}} + X_1 \cdot \mathbf{1}_{\{X_1 > T'\}}$, où $T' := \max_{i=2, \dots, n} X_i$. En calculant la loi de (X_1, T') , on montre que pour toute fonction mesurable bornée $g : \mathbb{R} \rightarrow \mathbb{R} : E[X_1 g(T)] = \int_0^1 g(s) \cdot \frac{s^n (n+1)}{2} ds$.)