Magistere d'Economiste Statisticien

Mesure et Intégration.
Applications aux probabilités et statistiques.

Jean-Marc Bardet (LSP)

Plan du cours

Laboratoire de Statistique et Probabilités — UMR C5583
Université Paul Sabatier — 31062 – Toulouse cedex 4.

1 Introduction

Il demeure des choses inconnues à partir des connaissances antérieures en probabilités :

- Qu'est-ce qu'un événement et l'ensemble de tous les événements ?
- Que se passe-t-il pour des probabilités d'événements moins classiques (par exemple l'ensemble des décimaux) ?
- Comment traiter une variable aléatoire qui est continue et discrète à la fois (par exemple le nombre de minutes passées devant la TV) ?

2 Espaces mesurables

2.1 Tribus

Notations. – Ω est un ensemble (fini ou infini).
- $\mathcal{P}(\Omega)$ est l'ensemble de tous les sous-ensembles (parties) de Ω.

Rappel. Soit E un ensemble. E est dit dénombrable s'il existe une bijection entre E et $\mathbb{N}$ ou un sous-ensemble de $\mathbb{N}$. Par exemple, un ensemble fini, $\mathbb{Z}$, $\mathbb{D}$, $\mathbb{Z} \times \mathbb{Z}$, $\mathbb{Q}$ sont dénombrables. En revanche, $\mathbb{R}$ n'est pas dénombrable.

Définition. Soit une famille $\mathcal{F}$ de parties de Ω (donc $\mathcal{F} \subset \mathcal{P}(\Omega)$). On dit que $\mathcal{F}$ est une algèbre si :

- $\Omega \subset \mathcal{F}$;
- lorsque $A \in \mathcal{F}$ alors $(\Omega \setminus A) \in \mathcal{F}$;
- pour tout $n \in \mathbb{N}^*$, lorsque $(A_1, \cdots, A_n) \in \mathcal{F}^n$ alors $A_1 \cup \cdots \cup A_n \in \mathcal{F}$.

Définition. Soit une famille $\mathcal{A}$ de parties de Ω (donc $\mathcal{A} \subset \mathcal{P}(\Omega)$). On dit que $\mathcal{A}$ est une tribu (ou σ-algèbre) sur Ω si :

- $\Omega \subset \mathcal{F}$;
- lorsque $A \in \mathcal{F}$ alors $(\Omega \setminus A) \in \mathcal{F}$;
- pour $I \subset \mathbb{N}$, lorsque $(A_i)_{i \in I} \in \mathcal{F}^I$ alors $\bigcup_{i \in I} A_i \in \mathcal{A}$.

Exemple. – Cas du Pile ou Face.
- Cas où Ω est infini : $\Omega = \mathbb{N}$ par exemple.

Propriété. Avec les notations précédentes :

1. $\emptyset \in \mathcal{A}$;
2. si A et B sont dans la tribu $\mathcal{A}$, alors $A \cap B$ est dans $\mathcal{A}$;
3. si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux tribus sur Ω, alors $\mathcal{A}_1 \cap \mathcal{A}_2$ est une tribu sur Ω . Plus généralement, pour $I \subset \mathbb{N}$, si $(\mathcal{A}_i)_{i \in I}$ ensemble de tribus sur Ω, alors $\bigcap_{i \in I} \mathcal{A}_i$ est une tribu sur Ω ;

4. si $\mathcal{A}_1$ et $\mathcal{A}_2$ sont deux tribus sur Ω, alors $\mathcal{A}_1 \cup \mathcal{A}_2$ n'est pas forcément une tribu sur Ω.

Définition. Si $\mathcal{E}$ est une famille de parties de Ω (donc $\mathcal{E} \subset \mathcal{P}(\Omega)$), alors on appelle tribu engendrée par $\mathcal{E}$, notée $\sigma(\mathcal{E})$, la tribu engendrée par l'intersection de toutes les tribus contenant $\mathcal{E}$ (on peut faire la même chose avec des algèbres).

Remarque. La tribu engendrée est la "plus petite" tribu (au sens de l'inclusion) contenant la famille $\mathcal{E}$.

Rappel.
- Un ensemble ouvert U dans un espace métrique X est telle que pour tout $x \in U$, il existe $r > 0$ tel que $B(x, r) \subset U$.
- On dit qu'un ensemble dans un espace métrique X est fermé si son complémentaire dans X est ouvert.

Définition. Soit Ω un espace métrique. On appelle tribu borélienne sur Ω, notée, $\mathcal{B}(\Omega)$, la tribu engendrée par les ouverts de Ω. Un ensemble de $\mathcal{B}(\Omega)$ est appelé borélien.

Exemple.
- Boréliens sur $\mathbb{R}$, sur $]0, 1[$.
- Boréliens sur $\mathbb{R}^2$.

2.2 Espace mesurable

Définition. Soit Ω un ensemble et soit $\mathcal{A}$ une tribu sur Ω. On dit que $(\Omega, \mathcal{A})$ est un espace mesurable.

Corollaire. Quand on s'intéressera aux probabilités, on dira que $(\Omega, \mathcal{A})$ est un espace probabilisable.

Propriété. Si $(\Omega_i, \mathcal{A}_i)_i$ sont n espaces mesurables, alors un ensemble élémentaire de $\Omega = \Omega_1 \times \cdots \times \Omega_n$ est une réunion finie d'ensembles $A_1 \times \cdots \times A_n$ où chaque $A_i \in \mathcal{A}_i$. L'ensemble des ensembles élémentaires est une algèbre et on note $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ (on dit $\mathcal{A}_1$ tensoriel $\mathcal{A}_2$ ... tensoriel $\mathcal{A}_n$) la tribu sur Ω engendrée par ces ensembles élémentaires.

Exemple. Pavés de $\mathbb{R}^d$.

Définition. On appelle espace mesurable produit des $(\Omega_i, \mathcal{A}_i)_i$ l'espace mesurable $\left(\prod_{i=1}^{n} \Omega_i, \bigotimes_{i=1}^{n} \mathcal{A}_i\right)$.

Exemple. Pile / Face 2 fois.

3 Mesures

3.1 Définitions - Propriétés

Définition. Soit $(\Omega, \mathcal{A})$ un espace mesurable. L'application $\mu \ : \mathcal{A} \to [0, +\infty]$ est une mesure si :
- $\mu(\emptyset) = 0$.
- Pour tout $I \subset \mathbb{N}$ et pour $(A_i)_{i \in I}$ famille disjointe de $\mathcal{A}$ (telle que $A_i \cup A_j = \emptyset$ pour $i \neq j$, alors $\mu\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mu(A_i)$ (propriété dite de σ-additivité).

Définition. Avec les notations précédentes :
- Si $\mu(\Omega) < +\infty$, on dit que μ est finie.
- Si $\mu(\Omega) < M$ avec $M < +\infty$, on dit que μ est bornée.
- Si $\mu(\Omega) = 1$, on dit que μ est une mesure de probabilité.

Exemple. Cas de $\Omega = \mathbb{R}$, de $\mathbb{N}$, ou $\mathbb{R}^2$.

Définition. Si $(\Omega, \mathcal{A})$ est un espace mesurable (resp. probabilisable) alors $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré (resp. probabilisé quand μ est une probabilité).

Remarque. Sur $(\Omega, \mathcal{A})$, on peut définir une infinité de mesures.

Propriété. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(A_i)_{i \in I}$ pour $I \subset \mathbb{N}$, une famille de $\mathcal{A}$.
1. Si $A_1 \subset A_2$, alors $\mu(A_1) \leq \mu(A_2)$.
2. Si $\mu(A_1) < +\infty$ et $\mu(A_2) < +\infty$, alors $\mu(A_1 \cup A_2) + \mu(A_1 \cap A_2) = \mu(A_1) + \mu(A_2)$.
3. On a $\mu\left(\bigcup_{i \in I} A_i\right) \leq \sum_{i \in I} \mu(A_i)$.
4. Si $I = \mathbb{N}$ et si $A_i \subset A_{i+1}$ pour tout $i \in \mathbb{N}$ (suite croissante en sens de l'inclusion), alors $(\mu(A_n))_{n \in \mathbb{N}}$ est une suite croissante convergente telle que $\mu\left(\bigcup_{i \in \mathbb{N}} A_i\right) = \lim_{i \to +\infty} \mu(A_i)$ (même si cette limite est $+\infty$).
5. Si $I = \mathbb{N}$ et si $A_{i+1} \subset A_i$ pour tout $i \in \mathbb{N}$ (suite décroissante en sens de l'inclusion) et $\mu(A_0) < +\infty$, alors $(\mu(A_n))_{n \in \mathbb{N}}$ est une suite décroissante convergente telle que $\mu\left(\bigcap_{i \in \mathbb{N}} A_i\right) = \lim_{i \to +\infty} \mu(A_i)$.

Exemple.
1. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. On définit $\nu(A) = \mu(A \cap B)$ où $B \in \mathcal{A}$. ν mesure ?
2. Si μ_1 et μ_2 mesures sur $(\Omega, \mathcal{A})$, $\mu_1 + \mu_2$ et $\alpha\mu$ sont-elles des mesures ?

Définition. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et $(A_i)_{i \in I}$ pour $I \subset \mathbb{N}$, une famille de $\mathcal{A}$.

1. On définit $\limsup (A_n)_n = \bigcap_{n \in \mathbb{N}} \bigcup_{m \geq n} A_m$ (intuitivement, $\limsup (A_n)_n$ est l'ensemble des $\omega \in \Omega$ tels que ω appartienne à une infinité de A_n).
2. On définit $\liminf (A_n)_n = \bigcup_{n \in \mathbb{N}} \bigcap_{m \geq n} A_m$ (intuitivement, $\liminf (A_n)_n$ est l'ensemble des $\omega \in \Omega$ tels que ω appartienne à tous les A_n sauf à un nombre fini d'entre eux).

Exemple. Cas des suites croissantes et décroissantes d'ensembles.

Théorème (Théorème d'extension de Hahn - Caratheodory). Si Ω est un ensemble, $\mathcal{F}$ une algèbre sur Ω, et ν une application de $\mathcal{F}$ dans $[0, +\infty]$ additive (telle que $\nu(A \cup B) = \nu(A) + \nu(B)$ pour $A \cup B = \emptyset$), alors si $\mathcal{A}$ est la tribu engendrée par $\mathcal{F}$, il existe une mesure $\widehat{\nu}$ sur la tribu $\mathcal{A}$ qui coïncide avec ν sur $\mathcal{F}$ (c'est-à-dire que pour tout $F \in \mathcal{F}$, $\widehat{\nu(F)} = \nu(F)$). On dit que $\widehat{\nu}$ prolonge ν sur la tribu $\mathcal{A}$.

Exemple. Définition de la mesure de Lebesgue sur $\mathbb{R}$, $\mathbb{R}^n$,...

Définition. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré.

1. Pour $A \in \mathcal{A}$, on dit que A est μ-négligeable si $\mu(A) = 0$.
2. Soit une propriété $\mathcal{P}$ dépendant des éléments ω de Ω. On dit que $\mathcal{P}$ est vraie μ-presque partout (μ-presque sûrement sur un espace probabilisé) si l'ensemble des ω pour laquelle elle n'est pas vérifiée est μ-négligeable.

Exemple.
- Mesure de Lebesgue sur $\mathbb{N}$ ou $\mathbb{Q}$
- La propriété " la suite de fonction $f_n(x) = x^n$ converge vers la fonction $f(x) = 0$" est vraie λ-presque partout sur $[0, 1]$.
- Soit $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ et soit F la fonction définie par $F(x) = \mu(] - \infty, x])$ pour $x \in \mathbb{R}$.

3.2 Fonctions mesurables

Rappel. Soit $f : E \mapsto F$, où E et F sont 2 espaces métriques.
- Pour $I \subset F$, on appelle ensemble réciproque de I par f, l'ensemble $f^{-1}(I) = \{x \in E, f(x) \in I\}$.
- (f continue) $\Longleftrightarrow$ (pour tout ouvert U de F alors $f^{-1}(U)$ est un ouvert de E).

Définition. Soit $f : E \mapsto F$ et soit $\mathcal{I}$ une tribu sur F. On note $f^{-1}(\mathcal{I})$ l'ensemble de sous-ensembles de Ω tel que $f^{-1}(\mathcal{I}) = \{f^{-1}(I), I \in \mathcal{I}\}$.

Propriété. Soit $(\Omega', \mathcal{A}')$ un espace mesurable et soit $f : \Omega \mapsto \Omega'$. Alors $f^{-1}(\mathcal{A})$ est une tribu sur Ω appelée tribu engendrée par f.

Définition. Soit $(\Omega, \mathcal{A})$ et $(\Omega', \mathcal{A}')$ deux espaces mesurables. Une fonction $f : \Omega \mapsto \Omega'$ est dite mesurable pour les tribus $\mathcal{A}$ et $\mathcal{A}'$ si et seulement si $f^{-1}(\mathcal{A}') \subset \mathcal{A}$ (donc si et seulement si $\forall A' \in \mathcal{A}'$, alors $f^{-1}(A') \in \mathcal{A}$).

Exemple.
- Fonction indicatrice.
- Combinaison linéaire de fonctions indicatrices.

Remarque. Dans le cas où $(\Omega, \mathcal{A})$ est un espace probabilisable, et si $f : \Omega \mapsto \mathbb{R}$, alors si f est une fonction mesurable sur $\mathcal{A}$ et $\mathcal{B}(\mathbb{R})$, alors f est une variable aléatoire.

Exemple. Nombre de Piles dans un jeu de Pile/Face.

Remarque. Dans le cas où $(\Omega, \mathcal{A})$ est un espace mesurable, et si $f : \Omega \mapsto (\Omega', \mathcal{B}(\Omega'))$, où Ω' est un espace métrique et $\mathcal{B}(\Omega')$ l'ensemble des boréliens de Ω', si f est une fonction mesurable sur $\mathcal{A}$ et $\mathcal{B}(\Omega')$, alors f est dite fonction borélienne.

Proposition. Soit $(\Omega, \mathcal{A})$ et $(\Omega', \mathcal{A}')$ deux espaces mesurables et $f : \Omega \mapsto \Omega'$. Soit $\mathcal{F}$ une famille de sous-ensembles de Ω' telle que $\sigma(\mathcal{F}) = \mathcal{A}'$. Alors

1. $f^{-1}(\mathcal{F})$ engendre la tribu $f^{-1}(\mathcal{A})$.
2. (f mesurable) $\Longleftrightarrow$ ($f^{-1}(\mathcal{F}) \subset \mathcal{A}$)

Conséquence.
- Si $(\Omega, \mathcal{A})$ et $(\Omega', \mathcal{A}')$ sont deux espaces mesurables boréliens, alors toute application continue de $\Omega \mapsto \Omega'$ est mesurable.
- Pour montrer qu'une fonction $f : \Omega \mapsto \mathbb{R}$ est mesurable, il suffit de montrer que la famille d'ensemble $(\{\omega \in \Omega, f(\omega) \leq a\})_{a \in \mathbb{R}} \in \mathcal{A}$.

Propriété.
- Soit f mesurable de $(\Omega, \mathcal{A})$ dans $(\Omega', \mathcal{A}')$ et g mesurable de $(\Omega', \mathcal{A}')$ dans $(\Omega'', \mathcal{A}'')$. Alors $g_0 f$ est mesurable dans $\mathcal{A}$ et $\mathcal{A}'$.
- Soit f_1 mesurable de $(\Omega, \mathcal{A})$ dans $(\Omega_1, \mathcal{A}_1)$ et f_2 mesurable de $(\Omega, \mathcal{A})$ dans $(\Omega_2, \mathcal{A}_2)$. Alors $h : \Omega \mapsto \Omega_1 \times \Omega_2$ telle que $h(\omega) = (f_1(\omega), f_2(\omega))$ est mesurable dans $\mathcal{A}$ et $\mathcal{A}_1 \otimes \mathcal{A}_2$.
- Soit $(f_n)_{n \in \mathbb{N}}$ une suite de fonctions mesurables de $(\Omega, \mathcal{A})$ dans $(\Omega', \mathcal{B}(\Omega'))$, où Ω' est un espace métrique, telle qu'il existe une fonction f limite simple de (f_n) (donc $\forall \omega \in \Omega$, $\lim_{n \to \infty} f_n(\omega) = f(\omega)$). Alors f est mesurable dans $\mathcal{A}$ et $\mathcal{B}(\Omega')$.

Définition. Soit f mesurable de $(\Omega, \mathcal{A}, \mu)$ dans $(\Omega', \mathcal{A}')$ et soit $\mu_f : \mathcal{A}' \mapsto [0, +\infty]$ telle que pour tout $A' \in \mathcal{A}'$, on ait $\mu_f(A') = \mu(f^{-1}(A'))$. Alors μ_f est une mesure sur $(\Omega', \mathcal{A}')$ appelée mesure image de μ par f.

Cas particulier. Si μ est une mesure de probabilité et si X est une variable aléatoire alors μ_X est la mesure (loi) de probabilité de la variable aléatoire X.

3.3 Cas des fonctions réelles mesurables

Propriété. Soit f et g deux fonctions réelles mesurables (de $(\Omega, \mathcal{A}, \mu)$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$). Alors $\alpha.f$, $f+g$, $\min(f,g)$ et $\max(f,g)$ sont des fonctions réelles mesurables.

Propriété. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions réelles mesurables. Alors $\inf(f_n)$ et $\sup(f_n)$ sont des fonctions réelles mesurables.

Définition. Soit $f : \Omega \to \mathbb{R}$. Alors f est dite étagée s'il existe une famille d'ensembles disjoints $(A_i)_{1\leq i\leq n}$ de Ω et une famille de réels $(\alpha_i)_{1\leq i\leq n}$ telles que pour tout $\omega \in \Omega$, on ait $f(\omega) = \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i}(\omega)$.

Remarque. Si les A_i sont tous dans $\mathcal{A}$ tribu sur Ω, alors f est $\mathcal{A}$-mesurable.

Théorème. Toute fonction réelle mesurable à valeurs dans $[0, +\infty]$ est limite simple d'une suite croissante de fonctions étagées.

Conséquence. Soit f une fonction réelle mesurable. Alors f est limite simple de fonctions étagées.

4 Intégration de Lebesgue

Dans toute la suite, on considère $(\Omega, \mathcal{A}, \mu)$ un espace mesuré.

4.1 Intégrale de Lebesgue d'une fonction positive

Définition.

1. Soit $f = \mathbb{I}_a$, où $A \in \mathcal{A}$. Alors :
$$\int f\, d\mu = \int_\omega f(\omega) d\mu(\omega) = \mu(A).$$
2. Soit $f = \mathbb{I}_a$, où $A \in \mathcal{A}$ et soit $B \in \mathcal{A}$. Alors :
$$\int_B f\, d\mu = \int_B f(\omega) d\mu(\omega) = \int \mathbb{I}_B \mu(A)(\omega) f(\omega) d\mu(\omega) = \mu(A \cap B).$$
3. Soit f une fonction étagée positive telle que $f = \sum_{i=1}^{n} \alpha_i \mathbb{I}_{A_i}$, où les $A_i \in \mathcal{A}$ et $\alpha_i > 0$ et soit $B \in \mathcal{A}$. Alors :
$$\int_B f\, d\mu = \int_B f(\omega) d\mu(\omega) = \int \mathbb{I}_B(\omega) f(\omega) d\mu(\omega) = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap B).$$

Exemple. Fonction $\mathbb{I}_{\mathbb{Q}}$ fonctions en escalier,...

Définition. Soit f une fonction $\mathcal{A}$-mesurable positive et soit $B \in \mathcal{A}$. Alors l'intégrale de Lebesgue de f par rapport à μ sur B est :

$$\int_B f\, d\mu = \int \mathbb{I}_B(\omega) f(\omega) d\mu(\omega) = \sup \left\{ \int_B g\, d\mu, \text{ pour } g \text{ étagée positive telle que } g \leq f \right\}.$$

Propriété. Soit f une fonction $\mathcal{A}$-mesurable positive et soit A et $B \in \mathcal{A}$. Alors :

1. Pour $c \geq 0$, $\int_B cf\, d\mu = c \int_B f\, d\mu$.
2. Si $A \subset B$, alors $\int_A f\, d\mu \leq \int_B f\, d\mu$.
3. Si g est une fonction $\mathcal{A}$-mesurable positive telle que $0 \leq f \leq g$ alors $0 \leq \int_B f\, d\mu \leq \int_B g\, d\mu$.
4. Si $\mu(B) = 0$ alors $\int_B f\, d\mu = 0$.

Théorème (Théorème de convergence monotone (Beppo-Lévi)). Si $(f_n)_n$ est une suite croissante de fonctions mesurables positives convergeant simplement vers f sur Ω, alors :

$$\lim_{n\to\infty} \left(\int f_n d\mu \right) = \int f\, d\mu = \int \lim_{n\to\infty} f_n d\mu.$$

Conséquence. Pour les séries de fonctions mesurables positives, on peut toujours appliquer le Théorème de convergence monotone et donc inverser la somme et l'intégrale.

Lemme (Lemme de Fatou). Soit $(f_n)_n$ est une suite de fonctions mesurables positives alors :

$$\int \left(\liminf_{n\to\infty} f_n \right) d\mu \leq \liminf_{n\to\infty} \int f_n d\mu.$$

Exemple. Appliquer Fatou à (f_n) telle que $f_{2n} = \mathbb{I}_A$ et $f_{2n+1} = \mathbb{I}_B$.

4.2 Intégrale de Lebesgue d'une fonction réelle et propriétés

Définition. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré, $B \in \mathcal{A}$ et soit f une fonction $\mathcal{A}$-mesurable à valeurs réelles telle que $f = f^+ - f^-$ avec $f^+ = \max(f, 0)$ et $f^- = \max(-f, 0)$. On dit que f est μ-intégrable sur B si $\int_B |f|\, d\mu < +\infty$. On a alors

$$\int_B f\, d\mu = \int_B f^+\, d\mu - \int_B f^-\, d\mu.$$

Notations. Lorsque f est μ-intégrable sur B, soit $\int |f|\, d\mu < +\infty$, on note $f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ (on dit que f est $\mathcal{L}^1$).

Exemple. Intégrale de Riemann et intégrale de Lebesgue.
Cas de la masse de Dirac.

Propriété. On suppose que f et $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. Alors :

1. $\displaystyle\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$ pour $(\alpha, \beta) \in \mathbb{R}^2$.
2. Si $f \leq g$ alors $\displaystyle\int f\, d\mu \leq \int g\, d\mu$.

Théorème (Théorème de convergence dominée de Lebesgue). Soit $(f_n)_n$ est une suite de fonctions de $\mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ telles que pour tout $n \in \mathbb{N}$, $|f_n| \leq g$ avec $g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$. Si on suppose que (f_n) converge simplement vers f sur Ω alors :

$$\lim_{n \to \infty} \int f_n d\mu = \int f\, d\mu.$$

Extension. Le Théorème de Lebesgue s'applique également dans le cas où $(f_n)_n$ converge presque partout vers f.

Exemple. Convergence d'intégrale dépendant d'un paramètre : par exemple $\displaystyle\int_0^\infty \frac{f(x)}{1 + x^n} dx$.

Théorème (Inégalité de Jensen). Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé, soit $\phi : \mathbb{R} \mapsto \mathbb{R}$ une fonction convexe et soit $f : \Omega \mapsto \mathbb{R}$ mesurable telle que $\phi(f)$ soit une fonction intégrable par rapport à P. Alors :

$$\phi\left(\int f\, dP\right) \leq \int \phi(f)\, dP.$$

Exemple. Soit X une v.a. sur $(\Omega, \mathcal{A}, P)$. Alors $\phi\left(\mathbb{E}X\right) \leq \mathbb{E}\left(\phi(X)\right)$.

4.3 Mesures induites et densités

Théorème (Théorème du Transport). Soit f une fonction mesurable de $(\Omega, \mathcal{A}, \mu)$ dans $(\Omega', \mathcal{A}')$ telle que μ_f soit la mesure induite par f (donc $\mu_f(A') = \mu(f^{-1}(A'))$ pour $A' \in \mathcal{A}'$) et soit ϕ une fonction mesurable de $(\Omega', \mathcal{A}')$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Alors, si $\phi_0 f \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$,

$$\int_{\Omega'} \phi\, d\mu_f = \int_\Omega \phi_0 f\, d\mu.$$

Définition. Soit μ et ν deux mesures sur $(\Omega, \mathcal{A})$. On dit que μ domine ν (ou ν est dominée par μ) et que ν est absolument continue par rapport à μ lorsque pour tout $A \in \mathcal{A}$, $\mu(A) = 0 \Longrightarrow \nu(A) = 0$.

Propriété. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et f une fonction définie sur $(\Omega, \mathcal{A})$ mesurable et positive. On suppose que pour $A \in \mathcal{A}$, $\nu(A) = \displaystyle\int_A f\, d\mu$. Alors, ν est une mesure sur $(\Omega, \mathcal{A})$, dominée par μ. De plus, pour toute fonction g définie sur $(\Omega, \mathcal{A})$ mesurable et positive,

$$\int g d\nu = \int g.f d\mu.$$

Enfin, g est ν intégrable si et seulement si $g.f$ est μ intégrable.

Définition. On dit que μ mesure sur $(\Omega, \mathcal{A})$ est σ-finie lorsqu'il existe une famille $(A_i)_{i\in I}$, avec I dénombrable, d'ensembles de $\mathcal{A}$ telle que $\bigcup A_i = \Omega$ et $\mu(A_i) < +\infty$ pour tout $i \in I$.

Théorème (Théorème de Radon-Nikodym). On suppose que μ et ν sont deux mesures σ-finies sur $(\Omega, \mathcal{A})$ telles que μ domine ν. Alors il existe une fonction f définie sur $(\Omega, \mathcal{A})$ mesurable et positive, appelée densité de ν par rapport à μ, telle que pour tout $A \in \mathcal{A}$, $\nu(A) = \displaystyle\int_A f\, d\mu$.

Théorème (Théorème de Fubini). Soit $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{A} = \mathcal{A}_1 \otimes \mathcal{A}_2$ et $\mu = \mu_1 \otimes \mu_2$ (mesures σ finies), où $(\Omega_1, \mathcal{A}_1, \mu_1)$ et $(\Omega_2, \mathcal{A}_2, \mu_2)$ sont des espaces mesurés. Soit une fonction $f : \Omega \mapsto \mathbb{R}$, $\mathcal{A}$-mesurable et μ-intégrable. alors :

$$\int_\Omega f d\mu = \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) d\mu_2(\omega_2) \right) d\mu_1(\omega_1) = \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) d\mu_1(\omega_1) \right) d\mu_2(\omega_2).$$

4.4 Espaces $\mathcal{L}^p$

Définition. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. On appelle espace $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, où $p > 0$, l'ensemble des fonctions $f : \Omega \mapsto \mathbb{R}$, mesurables et telles que $\int |f|^p d\mu < +\infty$.

Définition. Pour $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$, où $p > 0$, on note $\| f \|_p = \left(\displaystyle\int |f|^p d\mu \right)^{1/p}$.

Propriété (Inégalité de Hölder). Soit $p > 1$ et $q > 1$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$, et soit $f \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ et $g \in \mathcal{L}^q(\Omega, \mathcal{A}, \mu)$. Alors, $f\, g \in \mathcal{L}^1(\Omega, \mathcal{A}, \mu)$ et

$$\| f\, g \|_1 \leq \| f \|_p \, . \, \| g \|_q \, .$$

Propriété (Inégalité de Minkowski). Soit $p > 1$ et soit f et $g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. Alors, $f + g \in \mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ et

$$\| f + g \|_p \leq \| f \|_p + \| g \|_p \, .$$

Remarque. Pour $p > 1$, $\| \,.\, \|_p$ définie ainsi sur une semi-norme sur $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$. Pour obtenir une norme, il faut se place dans l'espace $\mathbb{L}^p(\Omega, \mathcal{A}, \mu)$ obtenu en "quotientant" $\mathcal{L}^p(\Omega, \mathcal{A}, \mu)$ par la relation d'équivalence $f = g$ μ-presque partout (c'est-à-dire que dans $\mathbb{L}^p(\Omega, \mathcal{A}, \mu)$ on dira que $f = g$ lorsque $f = g$ μ-presque partout).

Définition. Pour f et $g \in \mathbb{L}^2(\Omega, \mathcal{A}, \mu)$, on définit le produit scalaire $< f, g >= \displaystyle\int f.g\, d\mu$. On muni ainsi $\mathbb{L}^2(\Omega, \mathcal{A}, \mu)$ d'une structure d'espace de Hilbert. On dira que f est orthogonale à g lorsque $< f, g >= 0$.

Conséquence. Si A est un sous-espace vectoriel fermé de $\mathbb{L}^2(\Omega, \mathcal{A}, \mu)$ (par exemple un sous-espace de dimension finie), alors pour tout $f \in \mathbb{L}^2(\Omega, \mathcal{A}, \mu)$, il existe un unique projeté orthogonal de f sur A, noté f_A, qui vérifie $f_A = \operatorname*{Arginf}_{g\in A} \| g - f \|_2$.

5 Applications en Probabilités

5.1 Espérance de variables aléatoires

Définition. Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Alors si $X \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$, on définit l'espérance de X par le nombre $\mathbb{E}X = \int X dP$. Plus généralement, si $\phi : \mathbb{R} \mapsto \mathbb{R}$ est borélienne et si $\phi(X) \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$, on définit l'espérance de $\phi(X)$ par $\mathbb{E}\,\phi(X) = \int \phi(X) dP$.

Propriété. Si X est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, si $\phi : \mathbb{R} \mapsto \mathbb{R}$ est borélienne telle que $\phi(X) \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$, et si P_X est la mesure de probabilité de X alors :

$$\mathbb{E}\,\phi(X) = \int_{\mathbb{R}} \phi(x)\, dP_X(x).$$

Conséquence.
- Si P_X est absolument continue par rapport à la mesure de Lebesgue (donc X est une v.a. dite absolument continue), de densité f_X, alors $\mathbb{E}\,\phi(X) = \int_{\mathbb{R}} \phi(x) f_X(x) dx$.
- Si P_X est absolument continue par rapport à la mesure de comptage sur $\mathbb{N}$ (donc X est une v.a. dite discrète), de densité p_X, alors $\mathbb{E}\,\phi(X) = \sum_{k=0}^{\infty} p_X(k)\,\phi(k)$.

Propriété.
1. Soit X et Y des variables aléatoires telles que X et $Y \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$. Alors pour tout $(a, b) \in \mathbb{R}^2$, $aX + bY \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$ et
$$\mathbb{E}(aX + bY) = a\mathbb{E}X + b\mathbb{E}Y.$$
2. Soit X une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, et soit $A \in \mathcal{A}$. Alors $\mathbb{E}(\mathbb{I}_A(X)) = P(X \in A)$.
3. Soit X et Y des variables aléatoires telles que $X \in \mathbb{L}^p(\Omega, \mathcal{A}, P)$ et $Y \in \mathbb{L}^q(\Omega, \mathcal{A}, P)$ avec $\frac{1}{p} + \frac{1}{q} = 1$ et $p > 1$, $q > 1$. Alors $X.Y \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$ et
$$\mathbb{E}|X.Y| \leq (\mathbb{E}|X|^p)^{1/p}(\mathbb{E}|Y|^q)^{1/q}.$$
4. Soit X et Y des variables aléatoires telles que X et $Y \in \mathbb{L}^p(\Omega, \mathcal{A}, P)$, avec $p \geq 1$. Alors $X + Y \in \mathbb{L}^p(\Omega, \mathcal{A}, P)$ et
$$(\mathbb{E}|X + Y|^p)^{1/p} \leq (\mathbb{E}|X|^p)^{1/p} + (\mathbb{E}|Y|^p)^{1/p}.$$
5. Soit X une variable aléatoire telle que $X \in \mathbb{L}^p(\Omega, \mathcal{A}, P)$ pour $p > 0$. Alors pour tout $0 < r \leq p$, $X \in \mathbb{L}^r(\Omega, \mathcal{A}, P)$ et
$$(\mathbb{E}|X|^r)^{1/r} \leq (\mathbb{E}|X|^p)^{1/p}.$$

6. Si X est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, si $\phi : \mathbb{R} \mapsto \mathbb{R}$ est une fonction borélienne convexe telle que X et $\phi(X) \in \mathbb{L}^1(\Omega, \mathcal{A}, P)$, alors
$$\mathbb{E}(\phi(X)) \geq \phi(\mathbb{E}X).$$

Définition. Pour X et Y des variables aléatoires telles que X et $Y \in \mathbb{L}^2(\Omega, \mathcal{A}, P)$, on définit la covariance de X et Y par
$$\text{cov}(X, Y) = \mathbb{E}\left[(X - \mathbb{E}X)(Y - \mathbb{E}Y)\right];$$
On appelle variance de X, $\text{var}(X) = \text{cov}(X, X) = \mathbb{E}\left[(X - \mathbb{E}X)^2\right] = \mathbb{E}(X^2) - (\mathbb{E}X)^2$.

Propriété. Sur $\mathbb{L}^2(\Omega, \mathcal{A}, P)$, $\text{cov}(.,.)$ définit un produit scalaire. De plus
$$|\text{cov}(X, Y)|^2 \leq \text{var}(X).\text{var}(Y).$$

5.2 Vecteurs aléatoires

Définition. On dit que X est un vecteur aléatoire sur $(\Omega, \mathcal{A}, P)$, un espace probabilisé, si X est une fonction mesurable de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Définition. Soit X un vecteur aléatoire sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}^d$. Alors la loi (ou mesure) de probabilité de X, P_X, est définie de façon univoque à partir de la fonction de répartition de X, telle que pour $x = (x_1, \cdots, x_d)$,
$$F_X(x) = P_X(\prod_{i=1}^{d}]-\infty, x_i]) = P(X \in \prod_{i=1}^{d}]-\infty, x_i]).$$

Propriété. Soit X un vecteur aléatoire sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}^d$. On suppose que $X = (X_1, \cdots, X_d)$. Alors les X_i sont des variables aléatoires sur $(\Omega, \mathcal{A}, P)$, de fonction de répartition
$$F_{X_i}(x_i) = \lim_{\substack{x_j \to +\infty \\ j \neq i}} F_X(x_1, \cdots, x_i, \cdots, x_d).$$
Les mesures de probabilités P_{X_i} déterminées de façon univoque à partir des F_{X_i} sont appelées lois marginales de X.

5.3 Fonctions génératrices

Définition. Soit X un vecteur aléatoire sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}^d$. La fonction caractéristique de X est la fonction $\phi_X : \mathbb{R}^d \mapsto \mathbb{C}$ telle que
$$\phi_X(t) = \mathbb{E}\left[\exp(i < t, X >)\right] = \int_{\mathbb{R}^d} e^{i<t,x>} dP_X(x),$$
où $< . >$ désigne le produit scalaire euclidien sur $\mathbb{R}^d$ tel que $< t, x >= \sum_{i=1}^{d} t_i x_i$ pour $t = (t_1, \cdots, t_d)$ et $x = (x_1, \cdots, x_d)$.

Remarque. La fonction génératrice existe sur $\mathbb{R}$ et $\phi_X(0) = 1$. ϕ_X est aussi la transformée de Fourier de la mesure P_X.

Théorème. Soit X et Y des vecteurs aléatoires sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}^d$, de lois P_X et P_Y. Alors $P_X = P_Y$ si et seulement si $\phi_X = \phi_Y$.

Théorème (Théorème d'inversion). Si X est un vecteur aléatoire sur $(\Omega, \mathcal{A}, P)$ à valeurs dans $\mathbb{R}^d$ et si ϕ_X est une fonction intégrable par rapport à la mesure de Lebesgue λ_d sur $\mathbb{R}^d$, alors X admet une densité f_X par rapport à λ_d telle que pour $x \in \mathbb{R}^d$,

$$f_X(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i<t,x>} \phi_X(t) dt.$$

Proposition. Si X est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$ de fonction génératrice ϕ_X. Alors si $\mathbb{E}(|X|^n) < +\infty$ (ou $X \in \mathbb{L}^n(\Omega, \mathcal{A}, P)$), ϕ_X est n fois dérivable et $\phi_X^{(n)}(t) = i^n \mathbb{E}(X^n e^{itX})$.

Remarque. Lorsque ces moments existent, on a $i^n \mathbb{E}(X^n) = \phi_X^{(n)}(0)$.

5.4 Indépendance

Définition. Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

- Soit $(A_i)_{i \in I}$ une famille dénombrable d'événements de $\mathcal{A}$. On dit que les événements $(A_i)_{i \in I}$ sont indépendants si et seulement si pour tous les sous-ensembles finis $K \subset I$,
$$P\left(\bigcap_{i \in K} A_i\right) = \prod_{i \in K} P(A_i).$$
- Soit $(\mathcal{A}_i)_{i \in I}$ une famille de sous-tribus de $\mathcal{A}$ (donc pour tout $i \in I$, $\mathcal{A}_i \subset \mathcal{A}$). On dit que les tribus $(\mathcal{A}_i)_{i \in I}$ sont indépendantes si et seulement si pour tous les sous-ensembles finis $K \subset I$, et pour tous les événements $A_k \in \mathcal{A}_k$ avec $k \in K$, les A_k sont indépendants.
- Soit $(X_i)_{i \in I}$ des variables aléatoires sur $(\Omega, \mathcal{A})$ à valeurs dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On dit que les v.a. $(X_i)_{i \in I}$ sont indépendantes si et seulement si les tribus engendrées $(X_i^{-1}(\mathcal{B}(\mathbb{R})))_{i \in I}$ sont indépendantes.

Proposition. Si $(X_1, \cdots, X_n)$ sont des variables aléatoires sur $(\Omega, \mathcal{A}, P)$. Alors les (X_i) sont indépendantes si et seulement si $P_{(X_1, \cdots, X_n)} = \bigotimes_{i=1}^{n} P_{X_i}$.

Proposition. Si $(X_i)_{i \in I}$ sont des variables aléatoires indépendantes sur $(\Omega, \mathcal{A}, P)$. Alors les (X_i) sont indépendantes si et seulement si pour tout $J \subset I$, J fini, pour toutes fonctions boréliennes $(g_j)_{j \in J}$ telles que $g_j(X_j)$ soit intégrable, alors

$$\mathbb{E}\left(\prod_{j \in J} g_j(X_j)\right) = \prod_{j \in J} \mathbb{E}(g_j(X_j)).$$

Corollaire. $(X_1, \cdots, X_n)$ sont des variables aléatoires indépendantes si et seulement si pour tout $(t_1, \cdots, t_n) \in \mathbb{R}^n$,

$$\phi_{(X_1,\cdots,X_n)}(t_1, \cdots, t_n) = \prod_{j=1}^{n} \phi_{X_j}(t_j).$$

Lemme (Lemme de Borel-Cantelli). Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements sur $(\Omega, \mathcal{A}, P)$.

1. Si $\sum P(A_n) < +\infty$ alors $P(\limsup A_n) = 0$.
2. Si les (A_n) sont indépendants, $\sum P(A_n) = +\infty$ implique que $P(\limsup A_n) = 1$.

5.5 Convergence de suites de variables aléatoires

Définition. Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires sur $(\Omega, \mathcal{A}, P)$. On dit que

– (X_n) converge en probabilité vers X, noté $X_n \xrightarrow[n\to+\infty]{\mathcal{P}} X$, lorsque pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0.$$

– (X_n) converge dans $\mathbb{L}^p(\Omega, \mathcal{A}, P)$ vers X, noté $X_n \xrightarrow[n\to+\infty]{\mathbb{L}^p} X$, avec $p > 0$, lorsque

$$\lim_{n\to\infty} \mathbb{E}|X_n - X|^p = 0.$$

– (X_n) converge en loi vers X, noté $X_n \xrightarrow[n\to+\infty]{\mathcal{L}} X$, lorsque,

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \quad \text{où } x \in \mathbb{R} \text{ et } F_X \text{ continue en } x.$$

– (X_n) converge presque sûrement vers X, noté $X_n \xrightarrow[n\to+\infty]{p.s.} X$, lorsque pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} P(\sup_{m\geq n} |X_m - X| > \varepsilon) = 0.$$

Propriété.

1. $p.s.$ et $\mathbb{L}^p \longrightarrow \mathcal{P} \longrightarrow \mathcal{L}$.
2. pour $q \geq p$, $\mathbb{L}^q \longrightarrow \mathbb{L}^p$.
3. La convergence en loi n'entraîne pas la convergence en probabilité. Mais $(X_n \xrightarrow[n\to+\infty]{\mathcal{P}} C) \iff (X_n \xrightarrow[n\to+\infty]{\mathcal{L}} C)$ pour C une constante.

Propriété.

1. Si g est une fonction borélienne continue alors $(X_n \xrightarrow[n\to+\infty]{\mathcal{P}} X) \Longrightarrow (g(X_n) \xrightarrow[n\to+\infty]{\mathcal{P}} g(X))$.

2. Si pour tout $\varepsilon > 0$, $\sum_{n=0}^{\infty} P(|X_n - X| > \varepsilon) < +\infty$ alors $X_n \xrightarrow[n\to+\infty]{p.s.} X$

3. Si il existe $r > 0$ tel que $\mathbb{E}(|X_n|^r) < +\infty$ et $\sum_{n=0}^{\infty} \mathbb{E}(|X_n - X|^r) < +\infty$ alors $X_n \xrightarrow[n\to+\infty]{p.s.} X$

Théorème (Loi faible des Grands Nombres). Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées. Alors si $\mathbb{E}(|X_i|) < +\infty$,

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{\mathcal{P}} m = \mathbb{E}X_i.$$

Théorème (Loi forte des Grands Nombres). Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées. Alors si $\mathbb{E}(|X_i|) < +\infty$,

$$\bar{X}_n = \frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to+\infty]{p.s.} m = \mathbb{E}X_i.$$

Théorème (Théorème de la limite centrale). Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires indépendantes et identiquement distribuées. Alors si $\sigma^2 = \mathbb{E}X_i^2 < +\infty$, et $m = \mathbb{E}X_i$,

$$\sqrt{n}\frac{\bar{X}_n - m}{\sigma} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

5.6 Espérance conditionnelle

Définition. Soit Y une variable aléatoire sur $(\Omega, \mathcal{A}, P)$. Si $\mathcal{B}$ est une sous-tribu de $\mathcal{A}$ et si $Y \in \mathbb{L}^2(\Omega, \mathcal{A}, P)$. Alors on note $\mathbb{E}(Y \,|\, \mathcal{B})$ la projection orthogonale de Y sur $\mathbb{L}^2(\Omega, \mathcal{B}, P)$, appelée espérance conditionnelle de Y sachant $\mathcal{B}$. Ainsi :

$$\mathbb{E}\,|Y - \mathbb{E}(Y \,|\, \mathcal{B})|^2 = \inf_{Z\in\mathbb{L}^2(\Omega,\mathcal{B},P)} \left\{\mathbb{E}\,|Y - Z|^2\right\}.$$

Définition. Par convention, si X un vecteur aléatoire à valeurs dans $\mathbb{R}^n$ sur $(\Omega, \mathcal{A}, P)$ et si Y une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, on note $\mathbb{E}(Y \,|\, X) = \mathbb{E}(Y \,|\, X^{-1}(\mathcal{B}(\mathbb{R})))$.

Propriété.
1. $\mathbb{E}(Y \,|\, \mathcal{B})$ est une variable aléatoire de $\mathbb{L}^2(\Omega, \mathcal{B}, P)$; de plus, $\mathbb{E}(Y \,|\, X) = h(X)$, avec h une fonction borélienne.
2. Pour tout $B \in \mathcal{B}$, alors $\int_B \mathbb{E}(Y \,|\, \mathcal{B})\, dP = \int_B Y\, dP$.
3. Pour Y_1 et Y_2 deux variables aléatoires sur $(\Omega, \mathcal{A}, P)$, et $(a, b, c) \in \mathbb{R}^3$, alors

$$\mathbb{E}(aY_1 + bY_2 + c \,|\, \mathcal{B}) = a\mathbb{E}(Y_1 \,|\, \mathcal{B}) + b\mathbb{E}(Y_2 \,|\, \mathcal{B}) + c.$$

4. Si $Y_1 \leq Y_2$, alors $\mathbb{E}(Y_1 \,|\, \mathcal{B}) \leq \mathbb{E}(Y_2 \,|\, \mathcal{B})$.
5. Le Lemme de Fatou, les théorèmes de Beppo-Levi, Lebesgue et Jensen s'appliquent avec l'espérance conditionnelle.
6. Si $Y \in \mathbb{L}^2(\Omega, \mathcal{B}, P)$, alors $\mathbb{E}(Y \,|\, \mathcal{B}) = Y$; ainsi $\mathbb{E}(g(X) \,|\, X) = g(X)$ pour g une fonction mesurable réelle.
7. On a $\mathbb{E}(\mathbb{E}(Y \,|\, \mathcal{B})) = \mathbb{E}Y$.
8. Si $Y^{-1}(\mathcal{B}(\mathbb{R}))$ et $\mathcal{B}$ sont indépendantes alors $\mathbb{E}(Y \,|\, \mathcal{B}) = \mathbb{E}Y$; ainsi, si X et Y sont indépendantes, $\mathbb{E}(Y \,|\, X) = \mathbb{E}Y$.

Proposition. Si $(Y, X_1, \cdots, X_n)$ est un vecteur gaussien, alors $\mathbb{E}(Y \,|\, (X_1, \cdots, X_n)) = a_0 + a_1 X_1 + \cdots + a_n X_n$, où les a_i sont des réels.

Bibliographie

– Ansel et Ducel, *Exercices corrigés en théorie de la mesure et de l'intégration*, Ellipses.
– Barbe et Ledoux, *Probabilités*, Belin.
– Dacunha-Castelle et Duflo, *Probabilités et Statistiques (I)*, Masson
– Jacod, *Cours d'intégration*, http ://www.proba.jussieu.fr/pageperso/jacod.html.
– Jacod, *Cours de Probabilités*, http ://www.proba.jussieu.fr/pageperso/jacod.html.