

Intégration et probabilités
L3 MASS, Université de Nice-Sophia Antipolis

Sylvain Rubenthaler

8 septembre 2007

Table des matières

Introduction	iii
1 Dénombrement (rappels)	1
1.1 Ensembles dénombrables	1
1.2 Exercices	2
1.2.1 Énoncés	2
1.2.2 Corrigés	3
2 Théorie de la mesure	5
2.1 Tribus et mesures	5
2.1.1 Tribus	5
2.2 Mesures	6
2.3 Intégrales des fonctions étagées mesurables positives.	9
2.4 Fonctions mesurables et intégrales	10
2.4.1 Intégrales des fonctions mesurables positives	10
2.4.2 Intégrales des fonctions mesurables de signe quelconque.	11
2.5 Fonction de répartition	13
2.6 Exercices	13
2.6.1 Énoncés	13
2.6.2 Corrigés	15
3 Ensembles négligeables	17
4 Théorèmes limites	19
4.1 Stabilité de la mesurabilité par passage à la limite.	19
4.2 Théorèmes de convergence pour les intégrales.	19
4.3 Intégrales dépendant d'un paramètre	23
4.4 Exercices	25
4.4.1 Énoncés	25
4.4.2 Corrigés	26
5 Mesure produit et théorèmes de Fubini	29
5.1 Théorèmes de Fubini et Fubini-Tonelli	29
5.2 Changement de variable	31
5.3 Exercices	34
5.3.1 Énoncés	34
5.3.2 Corrigés	35
6 Fondements de la théorie des probabilités	37
6.1 Définitions générales	37
6.2 Espérance d'une v.a.	41
6.3 Inégalités	43
6.4 Lois classiques	44
6.4.1 Lois discrètes	44
6.4.2 Lois continues	45
6.5 Fonctions caractéristiques	45
6.6 Fonctions génératrices	46

6.7	Exercices	46
6.7.1	Énoncés	46
6.7.2	Corrigés	47
7	Variables indépendantes	53
7.1	Définitions générales	53
7.1.1	Événements et variables indépendantes	53
7.1.2	Densités de variables indépendantes	54
7.2	Lemme de Borel-Cantelli	55
7.3	Somme de deux variables indépendantes	56
7.4	Exercices	58
7.4.1	Énoncés	58
7.4.2	Corrigés	58
8	Convergence de variables aléatoires	61
8.1	Les différentes notions de convergence	61
8.2	Loi des grands nombres	62
8.3	Théorème central-limite	63
8.4	Exercices	66
8.4.1	Énoncés	66
8.4.2	Corrigés	67
9	Conditionnement	69
9.1	Conditionnement discret	69
9.2	Espérance conditionnelle	70
9.3	Exercice	72
9.3.1	Énoncé	72
9.3.2	Corrigé	72
10	Variables gaussiennes	75
10.1	Définitions et propriétés	75
10.2	Gaussiennes et espérance conditionnelle	76
A	Table de la loi normale	79

Introduction

Le but de ce cours est d'introduire les notions de théorie de la mesure qui seront utiles en calcul des probabilités et en analyse. Il est destiné aux étudiants qui veulent poursuivre leurs études dans un master à composante mathématique. Pour un cours plus complet, se reporter à la bibliographie.

Informations utiles (partiels, barèmes, annales, corrigés, ...) :
[http ://math.unice.fr/~rubentha/cours.html](http://math.unice.fr/~rubentha/cours.html).

PRÉREQUIS : Pour pouvoir suivre ce cours, l'étudiant doit connaître les développements limités, les équivalents, les études de fonction, le dénombrement, les nombre complexes, la théorie des ensembles.

Chapitre 1

Dénombrément (rappels)

1.1 Ensembles dénombrables

Définition 1.1.1. *Injection.*

Soit E, F des ensembles, $f : E \rightarrow F$ est une injection si $\forall x, y \in E, f(x) = f(y) \Rightarrow x = y$.

Définition 1.1.2. *Surjection.*

Soit E, F des ensembles, $f : E \rightarrow F$ est une surjection si $\forall z \in F, \exists x \in E$ tel que $f(x) = z$.

Définition 1.1.3. *Bijection.*

Soit E, F des ensembles, $f : E \rightarrow F$ est une bijection si f est une injection et une surjection.

Proposition 1.1.4. Soient E, F, G des ensembles. Soient $f : E \rightarrow F, g : F \rightarrow G$. Alors $[f \text{ et } g \text{ injectives}] \Rightarrow [g \circ f \text{ injective}]$.

Démonstration. Soient x, y tels que $g \circ f(x) = g \circ f(y)$. L'application g est injective donc $f(x) = f(y)$. L'application f est injective donc $x = y$. \square

Définition 1.1.5. On dit qu'un ensemble E est dénombrable s'il existe une injection de E dans \mathbb{N} . On démontre qu'il existe alors une bijection de E dans \mathbb{N} .

(Cela revient à dire que l'on peut compter un à un les éléments de E .)

Exemple 1.1.6. Tout ensemble fini est dénombrable.

Exemple 1.1.7. \mathbb{Z} est dénombrable car l'application

$$f : \mathbb{Z} \rightarrow \mathbb{N} \\ k \mapsto \begin{cases} 2n & \text{si } n \geq 0 \\ -2n + 1 & \text{si } n < 0 \end{cases}$$

est bijective (donc injective).

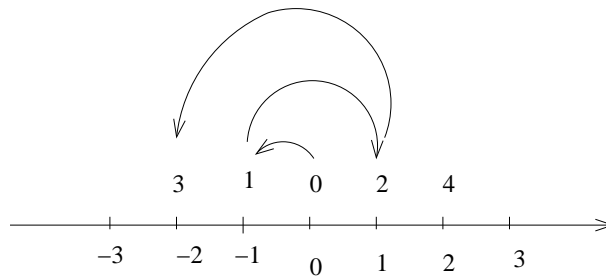


FIG. 1.1 – Énumération des éléments de \mathbb{Z} .

Exemple 1.1.8. $\mathbb{N} \times \mathbb{N}$ est dénombrable car l'application

$$f : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \\ (p, q) \mapsto \frac{(p+q)(p+q+1)}{2} + q$$

est bijective (donc injective).

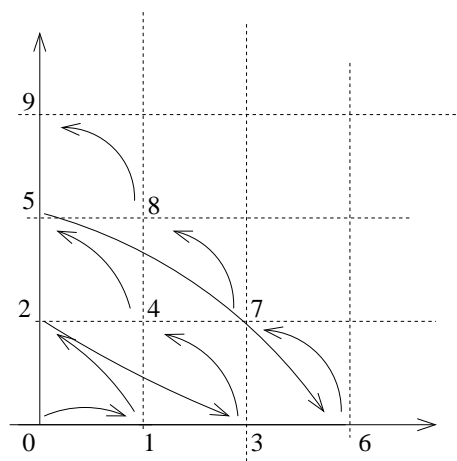


FIG. 1.2 – Énumération des éléments de $\mathbb{N} \times \mathbb{N}$.

Exemple 1.1.9. L'ensemble \mathbb{Q} est dénombrable. L'ensemble \mathbb{R} n'est pas dénombrable.

Proposition 1.1.10. Si on a $E_0, E_1, \dots, E_n, \dots$ des ensembles dénombrables alors $E = E_0 \cup E_1 \cup E_2 \cup \dots = \bigcup_{n \geq 0} E_n$ est un ensemble dénombrable.

(En d'autres termes, une réunion dénombrable d'ensembles dénombrables est dénombrable.)

Démonstration. S Pour tout $i \geq 0$, E_i est dénombrable donc $\exists f_i : E_i \rightarrow \mathbb{N}$ injective. Soit

$$\begin{aligned} F : \bigcup_{n \geq 0} E_n &\rightarrow \mathbb{N} \times \mathbb{N} \\ x &\mapsto (i, f_i(x)) \text{ si } x \in E_i \end{aligned}$$

Cette application F est injective. L'ensemble $\mathbb{N} \times \mathbb{N}$ est dénombrable donc il existe $g : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ injective. Par la proposition 1.1.4, $g \circ F$ est injective. Donc $\bigcup_{n \geq 0} E_n$ est dénombrable. \square

1.2 Exercices

Tous les exercices de ce chapitre n'ont pas un lien direct avec le cours. Par contre, ils constituent des révisions nécessaires à la suite du cours.

1.2.1 Énoncés

- 1) **Rappel :** Si $f : E \rightarrow F$ et $A \subset F$, $f^{-1}(A) = \{x \in E : f(x) \in A\}$. Si $C \subset E$, $f(C) = \{f(x), x \in C\}$.

On considère l'application $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$.

- Déterminer $f([-3, -1])$, $f([-3, 1])$, $f([-3, 1])$.
- Déterminer $f^{-1}(]-\infty, 2])$, $f^{-1}(]1, +\infty[)$, $f^{-1}(]-1, 0] \cup [1, 2])$.

- 2) Calculer les limites suivantes :

- $\lim_{x \rightarrow 0+} \frac{\sin(x)}{\log(1+x)}$
- $\lim_{x \rightarrow +\infty} \left(1 + \frac{2}{x}\right)^x$
- $\lim_{x \rightarrow 0} \frac{1 - \cos(x)}{x \sin(x)}$
- $\lim_{x \rightarrow 0} \frac{1 - (1+x)^\alpha}{1 - (1+x)^\beta}$ pour $\alpha, \beta > 0$.

- 3) Calculer les intégrales suivantes :

- $\int_0^{+\infty} x^2 e^{-x} dx$
- $\int_{e^1}^{+\infty} \frac{1}{(\log(z))^2 z} dz$

- (c) $\int_0^1 \frac{1}{(2-x)(1+x)} dx$
 (d) $\int_0^{\pi/4} \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} dx$.

4) Intégrales de Wallis

Pour tout $n \in \mathbb{N}$, on pose :

$$I_n = \int_0^{\pi/2} \sin^n(x) dx .$$

- (a) Calculer I_0 et I_1 .
 (b) Donner une relation de récurrence entre I_n et I_{n+2} .
 (c) En déduire que :

$$\forall p \in \mathbb{N}, I_{2p} = \frac{(2p-1)(2p-3)\dots 1}{2p(2p-2)\dots 2} \frac{\pi}{2} \text{ et } I_{2p+1} = \frac{2p(2p-2)\dots 2}{(2p+1)(2p-1)\dots 1} .$$

- (d) Montrer que $\forall p \in \mathbb{N}, I_{2p+1} \leq I_{2p} \leq I_{2p-1}$. En déduire que $\lim_{p \rightarrow +\infty} \frac{I_{2p}}{I_{2p+1}} = 1$.
 (e) En déduire la formule de Wallis :

$$\lim_{p \rightarrow +\infty} \frac{1}{p} \left[\frac{2p(2p-2)\dots 2}{(2p-1)(2p-3)\dots 1} \right]^2 = \pi .$$

- (f) Montrer que $\forall n \in \mathbb{N}, I_n \underset{n \rightarrow +\infty}{\sim} \sqrt{\frac{\pi}{2n}}$.

1.2.2 Corrigés

- (1) (a) $f([-3, -1]) = [1, 9], f([-3, 1]) = [0, 9], f([-3, 1]) = [0, 9]$.
 (b) $f^{-1}(]-\infty, 2]) = [-\sqrt{2}, \sqrt{2}], f^{-1}(]1, +\infty[) =]-\infty, -1] \cup]1, +\infty[, f^{-1}(]-1, 0]) \cup [1, 2]) = \{0\} \cup]-\sqrt{2}, -1] \cup [1, \sqrt{2}]$.
 (2) (a) $\frac{\sin(x)}{\log(1+x)} \underset{x \rightarrow 0+}{\sim} \frac{x}{x} = 1 \xrightarrow{x \rightarrow 0+} 1$
 (b) $(1 + \frac{x}{x})^x = e^{x \log(1 + \frac{x}{x})}$ et $x \log(1 + \frac{x}{x}) \underset{x \rightarrow 0}{\sim} \frac{2x}{x} \xrightarrow{x \rightarrow 0} 2$ donc par continuité de la fonction exp : $(1 + \frac{x}{x})^x \xrightarrow{x \rightarrow 0} e^2$
 (c) $\frac{1 - \cos(x)}{x \sin(x)} = \frac{(x^2/2) + o(x^2)}{x^2 + o(x^2)} \underset{x \rightarrow 0}{\sim} \frac{x^2}{2x^2} = 1/2$
 (d) $\frac{1 - (1+x)^\alpha}{1 - (1+x)^\beta} = \frac{\alpha x + o(x)}{\beta x + o(x)} \underset{x \rightarrow 0}{\sim} \frac{\alpha x}{\beta x} = \frac{\alpha}{\beta}$
 (a) on intègre par parties :

$$\begin{aligned} \int_0^{+\infty} x^2 e^{-x} dx &= [-x^2 e^{-x}]_0^{+\infty} + \int_0^{+\infty} 2x e^{-x} dx \\ &= 0 + [-2x e^{-x}]_0^{+\infty} + \int_0^{+\infty} 2e^{-x} dx \\ &= [-2e^{-x}]_0^{+\infty} = 2 \end{aligned}$$

- (b) changement de variable : $t = \log(z), z = e^t, dz = e^t dt$

$$\begin{aligned} \int_{e^1}^{+\infty} \frac{1}{(\log(z))^2 z} dz &= \int_1^{+\infty} \frac{1}{t^2} dt \\ &= [-1/t]_1^{+\infty} = 1 \end{aligned}$$

- (c) on décompose $\frac{1}{(2-x)(1+x)} = \frac{1/3}{2-x} + \frac{1/3}{1+x}$ (toujours possible pour une fraction rationnelle à pôles simples) et donc :

$$\int_0^1 \frac{1}{(2-x)(1+x)} dx = \left[-\frac{1}{3} \log(2-x) + \frac{1}{3} \log(1+x) \right]_0^1 = \frac{1}{3} \log(4)$$

(d) changement de variable : $t = \tan(x)$, $x = \arctan(t)$, $dx = \frac{1}{1+t^2} dt$

$$\begin{aligned} \int_0^{\pi/4} \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} dx &= \int_0^{\pi/4} 1 + \tan^2(x) dx \\ &= [\tan(x)]_0^{\pi/4} = 1 \end{aligned}$$

(3) (a) $I_0 = \int_0^{\pi/2} 1 dx = \frac{\pi}{2}$, $I_1 = \int_0^{\pi/2} \sin(x) dx = [-\cos(x)]_0^{\pi/2} = 1$.

(b) On intègre par parties pour tout $n \geq 2$:

$$\begin{aligned} I_{n+2} &= \int_0^{\pi/2} \sin^{n+1}(x) \sin(x) dx \\ &= [-\sin^{n+1}(x) \cos(x)]_0^{\pi/2} + (n+1) \int_0^{\pi/2} \sin^n(x) \cos^2(x) dx \\ &= (n+1)(I_n - I_{n+2}) \end{aligned}$$

d'où $I_{n+2} = \frac{n+1}{n+2} I_n$.

(c) Démonstration par récurrence de la formule pour I_{2p} (démonstration similaire pour I_{2p+1}) :

– c'est vrai en $p = 0$

– si c'est vrai jusqu'au rang p alors $I_{2p+2} = \frac{2p+1}{2p+2} I_{2p} = \frac{(2p+1)(2p-1)\dots 1}{(2p+2)(2p)\dots 2} \frac{\pi}{2}$

(d) $\forall p \in \mathbb{N}$, $\forall x \in [0, \pi/2]$, $0 \leq \sin^{2p+1}(x) \leq \sin^{2p}(x) \leq \sin^{2p-1}(x)$ donc par intégration $\forall p \in \mathbb{N}$, $I_{2p+1} \leq I_{2p} \leq I_{2p-1}$, donc $1 \leq \frac{I_{2p}}{I_{2p+1}} \leq \frac{I_{2p-1}}{I_{2p+1}} = \frac{2p+1}{2p}$, donc

$$\lim_{p \rightarrow +\infty} \frac{I_{2p}}{I_{2p+1}} = 1$$

(e) on déduit de la question précédente : $\lim_{p \rightarrow +\infty} \frac{\pi}{2} \left[\frac{(2p-1)(2p-3)\dots 1}{2p(2p-2)\dots 2} \right]^2 (2p+1) = 1$, d'où la formule de Wallis

(f) On fait la démonstration pour n impair . Soit $n = 2p + 1$:

$$\begin{aligned} I_{2p+1} &= \frac{2p(2p-2)\dots 2}{(2p+1)\dots 1} \\ &= \frac{\sqrt{p}}{2p+1} \sqrt{\frac{1}{p} \left(\frac{2p(2p+2)\dots 2}{(2p-1)\dots 1} \right)^2} \\ &\underset{p \rightarrow +\infty}{\sim} \frac{1}{\sqrt{2(2p+1)}} \sqrt{\pi} . \end{aligned}$$

Chapitre 2

Théorie de la mesure

La théorie de la mesure est l'outil utilisé pour modéliser le hasard.

2.1 Tribus et mesures

2.1.1 Tribus

Dans la suite, on utilisera un ensemble Ω que l'on appellera « univers ». Il contient tous les aléas possibles.

Définition 2.1.1. Une famille \mathcal{A} de parties de Ω est une tribu (sur Ω) si elle vérifie

1. $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ (stabilité par passage au complémentaire)
3. $A_0, A_1, A_2, \dots \in \mathcal{A} \Rightarrow \bigcup_{n \geq 0} A_n \in \mathcal{A}$ (une réunion dénombrable d'éléments de \mathcal{A} est dans \mathcal{A})

Remarque 2.1.2. On rappelle que :

- $A^c := \{x \in \Omega : x \notin A\}$
- Une tribu est un ensemble de parties. Ces parties sont appelées « événements ».

Proposition 2.1.3. Stabilité par intersection dénombrable.

Soient \mathcal{A} une tribu et $A_0, A_1, A_2, \dots \in \mathcal{A}$, alors $\bigcap_{n \geq 0} A_n \in \mathcal{A}$.

Démonstration. On note pour tout n , $B_n = A_n^c$. Donc, par définition d'une tribu, $B_n \in \mathcal{A}, \forall n$ et $\bigcup_{n \geq 0} B_n \in \mathcal{A}$.

$$\begin{aligned} \bigcap_{n \geq 0} A_n &= \bigcap_{n \geq 0} B_n^c \\ &= \left(\bigcup_{n \geq 0} B_n \right)^c \\ (\text{par définition}) &\in \mathcal{A}. \end{aligned}$$

□

Exemple 2.1.4. Pour n'importe quel ensemble Ω , $\mathcal{A} = \{\emptyset, \Omega\}$ est une tribu.

Exemple 2.1.5. Pour n'importe quel ensemble Ω , $\mathcal{A} = \mathcal{P}(\Omega)$ (les parties de Ω) est une tribu.

Proposition 2.1.6. Soit $\mathcal{A} \subset \mathcal{P}(\Omega)$, il existe une tribu notée $\sigma(\mathcal{A})$ telle que si \mathcal{B} est une tribu telle que $\mathcal{A} \subset \mathcal{B}$ alors $\sigma(\mathcal{A}) \subset \mathcal{B}$.

On dira que $\sigma(\mathcal{A})$ est la plus petite tribu contenant \mathcal{A} , ou encore que $\sigma(\mathcal{A})$ est la tribu engendrée par \mathcal{A} .

Définition 2.1.7. Soit l'ensemble de parties de \mathbb{R} suivant :

$$\mathcal{A} = \{]a, b[: a, b \in \mathbb{R} \cup \{+\infty, -\infty\}\}$$

(c'est l'ensemble des intervalles ouverts). La tribu $\sigma(\mathcal{A})$ s'appelle la tribu des boréliens et se note $\mathcal{B}(\mathbb{R})$.

Exemple 2.1.8. Soit $[a, b]$ intervalle fermé de \mathbb{R} . Les intervalles $] -\infty, a[$, $]b, +\infty[$ sont dans $\mathcal{B}(\mathbb{R})$. La famille $\mathcal{B}(\mathbb{R})$ est une tribu donc $] -\infty, a[\cup]b, +\infty[\in \mathcal{B}(\mathbb{R})$ (stabilité par réunion dénombrable), et donc aussi $(] -\infty, a[\cup]b, +\infty[)^c = [a, b] \in \mathcal{B}(\mathbb{R})$ (stabilité par passage au complémentaire).

De même, on peut montrer que tous les intervalles de \mathbb{R} sont dans $\mathcal{B}(\mathbb{R})$, ainsi que tous les singletons (les ensembles de la forme $\{x\}$, $x \in \mathbb{R}$).

2.2 Mesures

Notation 2.2.1. Dans le calcul des mesures, on adopte les conventions de calcul suivantes (qui ne sont pas valables ailleurs) : $\forall x \in \mathbb{R}$, $x + \infty = +\infty$, $0 \times \infty = 0$.

Définition 2.2.2. Soit Ω un ensemble muni d'une tribu \mathcal{A} . On dit que μ est une mesure (positive) sur (Ω, \mathcal{A}) si :

1. $\mu : \mathcal{A} \rightarrow [0, +\infty]$ (elle peut prendre la valeur ∞)
2. $\mu(\emptyset) = 0$
3. si $A_0, A_1, A_2, \dots \in \mathcal{A}$ et sont deux à deux disjoints alors $\mu(\bigcup_{n \geq 0} A_n) = \sum_{n \geq 0} \mu(A_n)$.

Quand μ est une mesure sur (Ω, \mathcal{A}) est telle que $\mu(\Omega) = 1$, on dit que μ est une mesure de probabilité (cette définition sera rappelée plus tard dans le cours). La tribu \mathcal{A} contient tous les événements possibles et, pour $A \in \mathcal{A}$, $\mu(A)$ est la probabilité que A se produise.

Définition 2.2.3. Quand μ est telle que $\mu(\Omega) < \infty$, on dit que μ est une mesure finie.

Définition 2.2.4. Quand on a un ensemble Ω avec une tribu \mathcal{A} sur Ω , on dit que (Ω, \mathcal{A}) est un espace mesurable. Si on a de plus, une mesure μ sur (Ω, \mathcal{A}) , on dit que $(\Omega, \mathcal{A}, \mu)$ est un espace mesuré.

Exemple 2.2.5. Le triplet $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{card})$ est un espace mesuré. Nous avons vu (exemple 2.1.5) que $\mathcal{P}(\mathbb{N})$ est une tribu sur \mathbb{N} . De plus :

1. Pour $A \in \mathcal{P}(\mathbb{N})$, $\text{card}(A)$ (= le nombre d'éléments de A) est bien dans $[0, +\infty]$.
2. La partie \emptyset est de cardinal 0.
3. Si $A_0, A_1, \dots \in \mathcal{P}(\mathbb{N})$ sont deux à deux disjoints, $\text{card}(\bigcup_{n \geq 0} A_n) = \sum_{n \geq 0} \text{card}(A_n)$.

Proposition 2.2.6. Croissance et mesure d'une différence

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Soit $A, B \in \mathcal{A}$ tels que $\mu(A) < +\infty$ et $B \subset A$.

- Alors $\mu(B) \leq \mu(A)$.
- Si, de plus $\mu(A) < +\infty$, alors $\mu(A \setminus B) = \mu(A) - \mu(B)$.

(Rappel : $A \setminus B = \{x : x \in A, x \notin B\}$.)

Démonstration. On a $\mu(A) = \mu(A \setminus B) + \mu(B)$ (car $A \setminus B$ et B sont disjoints). Donc $\mu(B) \leq \mu(A)$. Si $\mu(A) < +\infty$, nous avons alors $\mu(A \setminus B) = \mu(A) - \mu(B)$. \square

Proposition 2.2.7. Sous-additivité.

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Si $A_0, A_1, A_2, \dots \in \mathcal{A}$ (pas forcément deux à deux disjoints). Alors $\mu(\bigcup_{n \geq 0} A_n) \leq \sum_{n \geq 0} \mu(A_n)$.

Démonstration. On pose pour tout entier $k \geq 1$, $B_k = A_k \setminus \cup_{0 \leq i \leq k-1} A_i$ (et nous avons alors, par convention, $B_0 = A_0$). Les ensembles B_0, B_1, B_2, \dots sont deux à deux disjoints. Nous avons

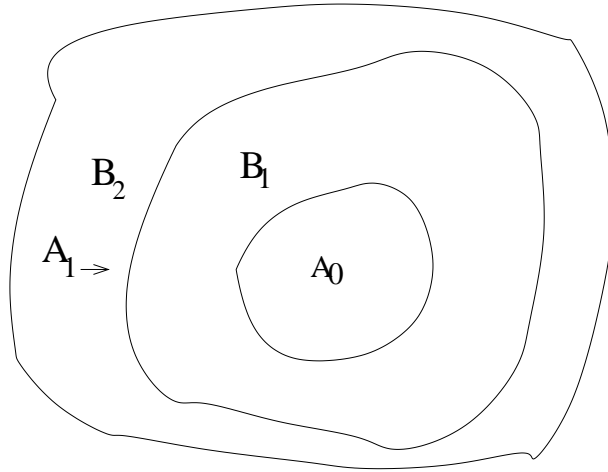
$$\begin{aligned} \mu\left(\bigcup_{n \geq 0} A_n\right) &= \mu\left(\bigcup_{n \geq 0} B_n\right) \\ (\text{car } B_0, B_1, B_2, \dots \text{ deux à deux disjoints}) &= \sum_{n \geq 0} \mu(B_n) \\ (\text{car } \forall n, B_n \subset A_n) &\leq \sum_{n \geq 0} \mu(A_n) \end{aligned}$$

□

Proposition 2.2.8. *Mesure d'une réunion croissante.*

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Soient $A_0, A_1, \dots \in \mathcal{A}$ tels que $A_0 \subset A_1 \subset \dots \subset A_n \subset A_{n+1} \subset \dots$. Alors $\mu\left(\bigcup_{k \geq 0} A_k\right) = \lim_{n \rightarrow \infty} \mu(A_n)$

Démonstration. Posons pour tout $k \geq 1$, $B_k = A_k \setminus A_{k-1} (= \{x : x \in A_k, x \notin A_{k-1}\})$ et $B_0 = A_0$.



Les ensembles B_0, B_1, B_2, \dots sont deux à deux disjoints. Donc

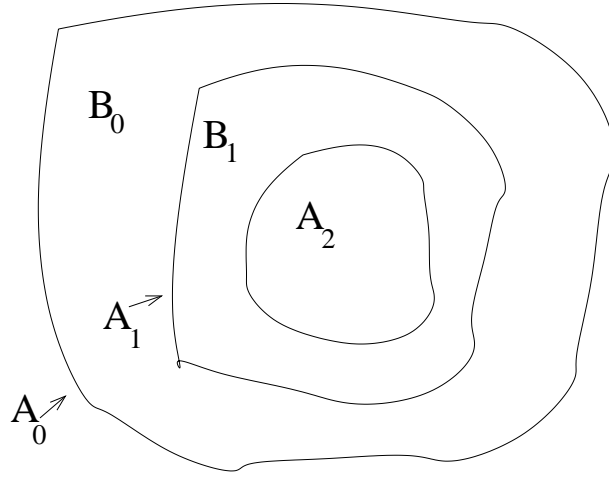
$$\begin{aligned} \mu\left(\bigcup_{k \geq 0} A_k\right) &= \mu\left(\bigcup_{k \geq 0} B_k\right) \\ &= \sum_{k \geq 0} \mu(B_k) \\ &= \lim_{n \rightarrow +\infty} \sum_{k=0}^n \mu(B_k) \end{aligned}$$

On a $\forall n, \sum_{k=0}^n \mu(B_k) = \mu(A_n)$. Donc $\mu\left(\bigcup_{k \geq 0} A_k\right) = \lim_{n \rightarrow +\infty} \mu(A_n)$. □

Proposition 2.2.9. *Mesure d'une intersection décroissante.*

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Soient $A_0, A_1, \dots \in \mathcal{A}$ tels que $A_0 \supset A_1 \supset \dots \supset A_n \supset A_{n+1} \supset \dots$ et tels que $\mu(A_0) < +\infty$. Alors $\mu\left(\bigcap_{k \geq 0} A_k\right) = \lim_{n \rightarrow +\infty} \mu(A_n)$.

Démonstration. Posons pour tout k , $B_k = A_k \setminus A_{k+1}$. Les ensembles B_0, B_1, B_2, \dots sont deux à deux disjoints.



Nous avons $\bigcap_{k \geq 0} A_k = A_0 \setminus \bigcup_{k \geq 0} B_k$, donc (par la proposition 2.2.6)

$$\begin{aligned}
 \mu\left(\bigcap_{k \geq 0} A_k\right) &= \mu(A_0) - \mu\left(\bigcup_{k \geq 0} B_k\right) \\
 (\text{mesure d'une réunion disjointe}) &= \mu(A_0) - \sum_{k \geq 0} \mu(B_k) \\
 &= \mu(A_0) - \lim_{n \rightarrow +\infty} \sum_{k=0}^n \mu(B_k) \\
 &= \lim_{n \rightarrow +\infty} (\mu(A_0) - \mu(B_0) - \cdots - \mu(B_n)) \\
 (\text{mesure d'une réunion disjointe}) &= \lim_{n \rightarrow +\infty} (\mu(A_0) - \mu(\bigcup_{0 \leq k \leq n} B_k)) \\
 (\text{cf. Prop. 2.2.6}) &= \lim_{n \rightarrow +\infty} \mu(A_{n+1}) .
 \end{aligned}$$

□

Théorème 2.2.10. *Mesure de Lebesgue.*

Il existe une mesure λ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ vérifiant

1. pour tout intervalle $]a, b[$, $\lambda(]a, b[) = b - a$
2. $\forall A \in \mathcal{B}(\mathbb{R}), \forall x \in \mathbb{R}, \lambda(\{y : y - x \in A\}) = \lambda(A)$.

Cette mesure λ s'appelle la mesure de Lebesgue.

Exemple 2.2.11. *Mesure de Lebesgue d'un intervalle quelconque.*

Soient $a \leq b$ des éléments de \mathbb{R} . Nous avons

$$\begin{aligned}
 \lambda([a, b]) &= \lambda(]a-1, b+1[\setminus]a-1, a[\cup]b, b+1[) \\
 (\text{par Prop. 2.2.6}) &= \lambda(]a-1, b+1[) - \lambda(]a-1, a[\cup]b, b+1[) \\
 (\text{réunion disjointe}) &= \lambda(]a-1, b+1[) - \lambda(]a-1, a[) - \lambda(]b, b+1[) \\
 &= (b+1 - (a-1)) - (a - (a-1)) - (b+1 - b) \\
 &= b - a .
 \end{aligned}$$

De même, $\lambda([a, b]) = \lambda(]a, b]) = b - a$.

Exemple 2.2.12. *Mesure de Lebesgue d'un singleton.*

Soit $x \in \mathbb{R}, \forall n \geq 1, \{x\} \subset [x-1/n, x+1/n]$. Donc $\forall n \geq 1, \lambda(\{x\}) \leq \lambda([x-1/n, x+1/n]) = 2/n$. Donc $\lambda(\{x\}) = 0$.

Exemple 2.2.13. *Mesure de Lebesgue de \mathbb{Q} .*

On sait que \mathbb{Q} est dénombrable. Donc on peut numéroter ses éléments : $\mathbb{Q} = \{u_0, u_1, u_2, \dots\}$. Pour tout entier $n \geq 1$, on définit $A_n = \bigcup_{i \geq 0} [u_i - \frac{1}{n2^i}, u_i + \frac{1}{n2^i}]$. On a pour tout $n, \mathbb{Q} \subset A_n$ et $\lambda(A_n) \leq \sum_{i \geq 0} \lambda([u_i - \frac{1}{n2^i}, u_i + \frac{1}{n2^i}]) = \frac{2}{n}$. Et donc $\lambda(\mathbb{Q}) = 0$.

2.3 Intégrales des fonctions étagées mesurables positives.

On se donne un espace mesuré $(\Omega, \mathcal{A}, \mu)$.

Définition 2.3.1. Soit $f : \Omega \rightarrow \mathbb{R}^+$. On dit que f est étagée (positive) s'il existe une famille finie A_1, \dots, A_n de \mathcal{A} telle que

- les A_i forment une partition de Ω (ce qui veut dire que A_1, \dots, A_n sont deux à deux disjoints et que $\Omega = \bigcup_{1 \leq i \leq n} A_i$)
- $\forall i \in \{1, \dots, n\}, \exists a_i$ tel que $f(x) = a_i, \forall x \in A_i$.

Remarque 2.3.2. Si f est une fonction étagée définie avec une partition A_1, \dots, A_n , il peut exister une autre partition B_1, \dots, B_m (différente de A_1, \dots, A_n) telle que f est constante sur chacun des B_i .

Définition 2.3.3. Soit $A \subset \Omega$. La fonction indicatrice de A est la fonction

$$\begin{aligned} \mathbf{1}_A : \Omega &\rightarrow \{0, 1\} \\ x &\mapsto \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A. \end{cases} \end{aligned}$$

Il existe d'autres notations. Par exemple si $A = [0, 1] \subset \mathbb{R}$, on peut écrire $\mathbf{1}_A(x) = \mathbf{1}_{x \in [0, 1]} = \mathbf{1}_{0 \leq x \leq 1}$.

Exemple 2.3.4. La fonction

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \begin{cases} 0 & \text{si } x < 0 \\ \lfloor x \rfloor & \text{si } x \in [0, 2] \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

est une fonction positive étagée ($\lfloor x \rfloor$ signifie « partie entière »). En effet, elle est constante sur $]-\infty, 0[$, $[0, 1[$, $[1, 2[$, $\{2\}$, $]2, +\infty[$.

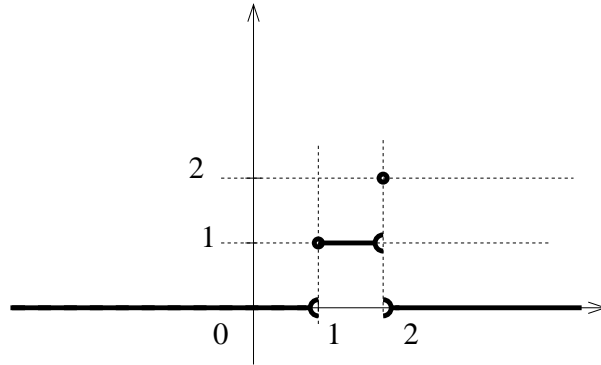


FIG. 2.1 – Dessin de f .

Avec des fonctions indicatrice, nous pouvons écrire f de manière plus compacte :

$$f(x) = \lfloor x \rfloor \mathbf{1}_{[0, 2]}(x) = \mathbf{1}_{[0, 2]}(x) + 2 \times \mathbf{1}_{\{2\}}(x) = \dots$$

Définition 2.3.5. Soit f une fonction positive étagée associée à une partition A_1, \dots, A_n (avec $f(x) = a_i$ si $x \in A_i$). On appelle intégrale de f par rapport à μ le nombre suivant

$$\int_{\Omega} f(x) \mu(dx) := \sum_{i=1}^n a_i \mu(A_i).$$

Ce nombre peut être $+\infty$. Une fonction positive étagée f est dite intégrable si $\int_{\Omega} f(x) \mu(dx) < +\infty$.

Remarque 2.3.6. La valeur de $\int_{\Omega} f(x) \mu(dx)$ est indépendante de la partition associée à f .

2.4 Fonctions mesurables et intégrales

2.4.1 Intégrales des fonctions mesurables positives

Définition 2.4.1. *Application mesurable.*

Soient (Ω, \mathcal{A}) , (Ω', \mathcal{A}') deux espaces mesurés. On dit qu'une application $f : \Omega \rightarrow \Omega'$ est mesurable (par rapport aux tribus \mathcal{A} , \mathcal{A}') si $\forall B \in \mathcal{A}'$, $f^{-1}(B) := \{x \in \Omega : f(x) \in B\} \in \mathcal{A}$.

Proposition 2.4.2.

- Toute fonction continue $f : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est mesurable.
- Si f et g sont des fonction mesurables $(\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ alors $f + g$, $f \times g$, $\frac{f}{g}$ sont mesurables.
- Si $f : (\Omega, \mathcal{A}) \rightarrow (\Omega', \mathcal{A}')$ est mesurable et $g : (\Omega', \mathcal{A}') \rightarrow (\Omega'', \mathcal{A}'')$ est mesurable alors $g \circ f : (\Omega, \mathcal{A}) \rightarrow (\Omega'', \mathcal{A}'')$ est mesurable.

De manière générale, toute fonction $(\mathbb{R}, \mathcal{B}(\mathbb{R})) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ définie par une formule est mesurable.

Proposition 2.4.3. *Mesure image.*

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Soit (Ω', \mathcal{B}) un espace mesurable. Soit $f : \Omega \rightarrow \Omega'$ mesurable. L'application $\nu : \mathcal{B} \rightarrow [0, +\infty]$ définie par $\nu(B) = \mu(f^{-1}(B))$ est une mesure appelée mesure image de μ par f .

(Rappel : $f^{-1}(B) := \{x \in \Omega : f(x) \in B\}$.)

Démonstration. Vérifions d'abord que ν est bien définie : $\forall B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$ car f est mesurable, donc $\nu(B)$ est bien défini. On a donc $\nu : \mathcal{B} \rightarrow [0, +\infty]$.

Puis $\nu(\emptyset) = \mu(f^{-1}(\emptyset)) = \mu(\emptyset) = 0$ car μ est une mesure.

Enfin, si $B_0, B_1, B_2, \dots \in \mathcal{B}$ sont deux à deux disjoints, $\nu(\bigcup_{n \geq 0} B_n) = \mu(f^{-1}(\bigcup_{n \geq 0} B_n)) = \mu(\bigcup_{n \geq 0} f^{-1}(B_n))$. En effet $f^{-1}(\bigcup_{n \geq 0} B_n) = \{x \in \Omega : f(x) \in \bigcup_{n \geq 0} B_n\} = \bigcup_{n \geq 0} \{x \in \Omega : f(x) \in B_n\}$. Soient $m \neq n$, si $x \in f^{-1}(B_n)$, $f(x) \in B_n$, donc $f(x) \notin f^{-1}(B_m)$ (car B_0, B_1, B_2, \dots sont deux à deux disjoints), donc $x \notin f^{-1}(B_m)$, donc $f^{-1}(B_n) \cap f^{-1}(B_m) = \emptyset$. Donc, puisque μ est une mesure,

$$\begin{aligned} \nu(\bigcup_{n \geq 0} B_n) &= \mu(\bigcup_{n \geq 0} f^{-1}(B_n)) \\ &= \sum_{n \geq 0} \mu(f^{-1}(B_n)) \\ &= \sum_{n \geq 0} \nu(B_n) . \end{aligned}$$

Donc ν est une mesure. □

Définition 2.4.4. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Si $f : \Omega \rightarrow [0, +\infty]$ est mesurable (par rapport aux tribus \mathcal{A} et $\mathcal{B}(\mathbb{R})$) positive, l'intégrale de f sur Ω par rapport à la mesure μ est définie par

$$\int_{\Omega} f(x) \mu(dx) := \sup_{\phi \in \mathcal{E}(f)} \int_{\Omega} \phi(x) \mu(dx)$$

où $\mathcal{E}(f) := \{\phi \text{ étagée positive} : \phi(x) \leq f(x), \forall x \in \Omega\}$. Cette intégrale peut prendre sa valeur dans $[0, +\infty]$.

Pour $B \in \mathcal{A}$, on note

$$\int_B f(x) \mu(dx) = \int_{\Omega} f(x) \mathbf{1}_B(x) \mu(dx) .$$

Définition 2.4.5. Une fonction mesurable positive f est dite intégrable si $\int_{\Omega} f(x) \mu(dx) < \infty$.

Proposition 2.4.6. *Croissance de l'intégrale.*

Soient f, g deux fonctions positives mesurables sur $(\Omega, \mathcal{A}, \mu)$. Si $f \leq g$ (ce qui veut dire $f(x) \leq g(x), \forall x$) alors $\int_{\Omega} f(x) \mu(dx) \leq \int_{\Omega} g(x) \mu(dx)$.

Démonstration. Nous avons $\mathcal{E}(f) \subset \mathcal{E}(g)$ car $f \leq g$. Donc

$$\sup_{\phi \in \mathcal{E}(f)} \int_{\Omega} \phi(x) \mu(dx) \leq \sup_{\phi \in \mathcal{E}(g)} \int_{\Omega} \phi(x) \mu(dx) .$$

□

Cette proposition admet comme corollaire le théorème suivant.

Théorème 2.4.7. *Théorème de comparaison.*

Soient f, g deux fonctions positives mesurables sur $(\Omega, \mathcal{A}, \mu)$. Si $f \leq g$ et g est intégrable alors f est intégrable.

Définition 2.4.8. Soit μ mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. La mesure μ est dite avoir pour densité la fonction $f \geq 0$ sur \mathbb{R} (par rapport à λ) si $\forall \phi$ mesurable positive $\mathbb{R} \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}} \phi(x) \mu(dx) = \int_{\mathbb{R}} \phi(x) f(x) \lambda(dx) .$$

Ceci implique, en particulier, que $\forall B \in \mathcal{B}(\mathbb{R})$,

$$\mu(B) = \int_B \phi(x) f(x) \lambda(dx) .$$

Théorème 2.4.9. *Linéarité de l'intégrale.*

Soient f, g deux fonctions positives mesurables sur $(\Omega, \mathcal{A}, \mu)$ et $a \geq 0$, alors :

$$\int_{\Omega} f(x) + g(x) \mu(dx) = \int_{\Omega} f(x) \mu(dx) + \int_{\Omega} g(x) \mu(dx)$$

et

$$\int_{\Omega} a f(x) \mu(dx) = a \int_{\Omega} f(x) \mu(dx) .$$

En particulier, si f et g sont intégrables alors $f + g$ aussi.

Théorème 2.4.10. *Inégalité de Markov.*

Soient f, g deux fonctions positives mesurables sur $(\Omega, \mathcal{A}, \mu)$. Soit $a > 0$. Alors :

$$\mu(\{x \in \Omega : f(x) \geq a\}) \leq \frac{1}{a} \int_{\Omega} f(x) \mu(dx) .$$

Démonstration. On a $a \mathbf{1}_{\{y: f(y) \geq a\}} \leq f$ donc par théorème de comparaison (théorème 2.4.7) :

$$\int_{\Omega} a \mathbf{1}_{\{y: f(y) \geq a\}}(x) \mu(dx) \leq \int_{\Omega} f(x) \mu(dx) .$$

La fonction $a \mathbf{1}_{\{y: f(y) \geq a\}}$ est une fonction étagée et on calcule son intégrale :

$$\int_{\Omega} a \mathbf{1}_{\{y: f(y) \geq a\}}(x) \mu(dx) = a \times \mu(\{y : f(y) \geq a\}) + 0 \times \mu(\{y : f(y) < a\}) .$$

D'où le résultat. □

2.4.2 Intégrales des fonctions mesurables de signe quelconque.

Soit une espace mesuré $(\Omega, \mathcal{A}, \mu)$. Soit $f : \Omega \rightarrow \mathbb{R}$ mesurable. Elle peut toujours s'écrire $f = f^+ - f^-$ avec f^+ et f^- mesurables positives :

$$f^+(x) = \begin{cases} f(x) & \text{si } f(x) \geq 0 \\ 0 & \text{sinon} \end{cases}$$

$$f^-(x) = \begin{cases} 0 & \text{si } f(x) \geq 0 \\ -f(x) & \text{sinon.} \end{cases}$$

Définition 2.4.11. Une fonction f mesurable sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$ est dite intégrable si f^+ et f^- le sont (voir définition 2.4.5 de l'intégrabilité des fonctions mesurables positives) et dans ce cas, on définit l'intégrale de f (sur Ω par rapport à μ) par

$$\int_{\Omega} f(x) \mu(dx) := \int_{\Omega} f^+(x) \mu(dx) - \int_{\Omega} f^-(x) \mu(dx)$$

et, $\forall A \in \mathcal{A}$, l'intégrale de f sur A par

$$\int_A f(x) \mu(dx) := \int_{\Omega} f(x) \mathbf{1}_A(x) \mu(dx) .$$

Lemme 2.4.12. Soit f une fonction mesurable sur un espace mesuré $(\Omega, \mathcal{A}, \mu)$ et intégrable. Alors

$$\left| \int_{\Omega} f(x) \mu(dx) \right| \leq \int_{\Omega} |f(x)| \mu(dx)$$

Démonstration.

$$\begin{aligned} \left| \int_{\Omega} f(x) \mu(dx) \right| &= \left| \int_{\Omega} f^+(x) \mu(dx) - \int_{\Omega} f^-(x) \mu(dx) \right| \\ &\leq \left| \int_{\Omega} f^+(x) \mu(dx) \right| + \left| \int_{\Omega} f^-(x) \mu(dx) \right| \\ &= \int_{\Omega} f^+(x) \mu(dx) + \int_{\Omega} f^-(x) \mu(dx) \\ &= \int_{\Omega} |f(x)| \mu(dx) . \end{aligned}$$

□

Ce lemme peut aussi être vu comme une conséquence de l'inégalité de Jensen (cf. exercice 4 du chapitre 4 et théorème 6.3.1).

Théorème 2.4.13. Linéarité et croissance.

Pour l'intégrale d'une fonction de signe quelconque, on a encore la linéarité et la croissance comme dans la proposition 2.4.6 et le théorème 2.4.9.

Remarque 2.4.14. Lien intégrale de Lebesgue/intégrale de Riemann.

Quand $(\Omega, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, l'intégrale $\int_{\Omega} f(x) \mu(dx) = \int_{\mathbb{R}} f(x) \lambda(dx)$ que nous venons de définir s'appelle l'intégrale de Lebesgue sur \mathbb{R} . Vu la définition 2.4.11, l'intégrale de Lebesgue sur un intervalle $[a, b]$ est donnée par

$$\int_{[a,b]} f(x) \lambda(dx) := \int_{\mathbb{R}} f(x) \mathbf{1}_{[a,b]}(x) \lambda(dx) .$$

L'intégrale de Riemann est celle qui se calcule avec la primitive. Si f admet une primitive F alors son intégrale de Riemann est

$$\int_a^b f(x) dx = [F(x)]_a^b = F(b) - F(a)$$

avec la convention que si F n'est pas définie en a (et pareil en b), par exemple parce que $a = -\infty$, alors $F(a) = \lim_{x \rightarrow a, x \in [a,b]} F(x)$. On parle alors d'intégrale généralisée (ou d'intégrale de Riemann généralisée). L'intégrale de Riemann n'est définie que si $F(a)$ et $F(b)$ sont finis.

On a les règles de signe suivantes :

$$\begin{aligned} \int_a^b f(x) dx &= - \int_b^a f(x) dx \\ \int_{[a,b]} f(x) \lambda(dx) &= \int_{[b,a]} f(x) \lambda(dx) . \end{aligned}$$

Dans le cas où f a une intégrale de Riemann, nous avons l'égalité suivante entre les deux types d'intégrales si $a \leq b$

$$\int_{[a,b]} f(x) \lambda(dx) = \int_a^b f(x) dx .$$

C'est en général avec cette formule que l'on calculera les intégrales. On écrira parfois :

$$\int_{[a,b]} f(x) \lambda(dx) = \int_{[a,b]} f(x) dx .$$

2.5 Fonction de répartition

L'étude de la fonction de répartition d'une mesure va nous permettre de mettre en œuvre les théorèmes de ce chapitre.

Définition 2.5.1. Soit μ mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\mu(\mathbb{R}) < +\infty$. On définit la fonction de répartition de μ par :

$$\begin{aligned} F_\mu : \mathbb{R} &\rightarrow [0, +\infty[\\ x &\mapsto F_\mu(x) = \mu(]-\infty, x]) . \end{aligned}$$

Proposition 2.5.2. Soit μ mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que $\mu(\mathbb{R}) < +\infty$. La fonction F_μ est croissante, càdlàg (continue à droite avec une limite à gauche), $\lim_{x \rightarrow +\infty} F_\mu(x) = \mu(\mathbb{R})$, $\lim_{x \rightarrow -\infty} F_\mu(x) = 0$.

Démonstration. Soient $x \leq y$. Nous avons $]-\infty, x] \subset]-\infty, y]$ donc, par la proposition 2.2.6, $F_\mu(x) = \mu(]-\infty, x]) \leq \mu(]-\infty, y]) = F_\mu(y)$.

Soit $x \in \mathbb{R}$ et $(u_n)_{n \geq 0}$ suite de \mathbb{R} telle que $u_n \geq x$ et $u_n \geq u_{n+1}$, $\forall n$ et $\lim_{n \rightarrow +\infty} u_n = x$. Pour tout n , $]-\infty, u_{n+1}] \subset]-\infty, u_n]$, $\bigcap_{n \geq 0}]-\infty, u_n] =]-\infty, x]$ et $\mu(]-\infty, u_0]) \leq \mu(\mathbb{R}) < \infty$, donc, par la proposition sur l'intersection décroissante (prop. 2.2.9) $\lim_{n \rightarrow +\infty} \mu(]-\infty, u_n]) = \mu(\bigcap_{n \geq 0}]-\infty, u_n]) = \mu(]-\infty, x])$. En d'autres termes : $\lim_{n \rightarrow +\infty} F_\mu(u_n) = F_\mu(x)$. Ceci prouve que F est continue à droite.

Soit $x \in \mathbb{R}$ et $(u_n)_{n \geq 0}$ suite de \mathbb{R} telle que $u_n < x$ et $u_n \leq u_{n+1}$, $\forall n$ et $\lim_{n \rightarrow +\infty} u_n = x$. Pour tout n , $]-\infty, u_{n+1}] \supset]-\infty, u_n]$, $\bigcup_{n \geq 0}]-\infty, u_n] =]-\infty, x[$, donc par la propriété de réunion croissante (prop. 2.2.8), $\lim_{n \rightarrow +\infty} F(u_n) = \mu(]-\infty, x[)$. Ceci prouve que F_μ a une limite à gauche (égale à $\mu(]-\infty, x[)$).

On trouve également la limite de F_μ en $+\infty$ en utilisant la propriété de réunion croissante et la limite de F_μ en $-\infty$ en utilisant la propriété de réunion décroissante. \square

Remarque 2.5.3. Dans la proposition précédente, la limite à gauche en x de F_μ est $\mu(]-\infty, x[)$ et $F_\mu(x) = \mu(]-\infty, x])$. Par la proposition 2.2.6, $\mu(]-\infty, x]) - \mu(]-\infty, x[) = \mu(\{x\})$. Donc $F_\mu(x) = \mu(]-\infty, x])$ si et seulement si $\mu(\{x\}) = 0$.

2.6 Exercices

2.6.1 Énoncés

- 1) Rappel : Pour une famille d'ensemble $(A_n)_{n \in \mathbb{N}}$, on note $\bigcap_{n \geq 0} A_n = \{x : \forall n, x \in A_n\}$ et $\bigcup_{n \geq 0} A_n = \{x : \exists n \text{ tel que } x \in A_n\}$
 - (a) Déterminer $\bigcap_{n \geq 0}]1, 1 + 1/(n+1)]$.
 - (b) Déterminer $\bigcap_{n \geq 0}]1, 2 + 1/(n+1)]$.
 - (c) Déterminer $\bigcap_{n \geq 0}]1 - 1/(n+1), 2]$.
 - (d) Soit $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^2$. Déterminer $f^{-1}(\bigcup_{n \geq 0} [1/(n+1), +\infty[)$.
- 2) Soit Ω un ensemble et soient A_0, A_1, \dots des parties de Ω .
 - (a) On suppose dans cette question que $A_0 \subset A_1 \subset \dots \subset A_n \subset A_{n+1} \subset \dots$. Posons pour tout $n \geq 1$, $B_n = A_n \setminus A_{n-1}$ (rappel : $A \setminus C = \{x \in A : x \notin C\}$). Montrer que les ensembles B_n sont deux à deux disjoints.

- (b) On note : $\forall A \subset \Omega, A^c = \{x \in \Omega : x \notin A\}$. Montrer que $\bigcup_{n \geq 0} A_n^c = (\bigcap_{n \geq 0} A_n)^c$.
- (c) Montrer que $(\bigcup_{n \geq 0} A_n^c)^c = \bigcap_{n \geq 0} A_n$.
- 3) Soit A_1, \dots, A_n une partition de \mathbb{R} . Montrer que $\mathcal{A} = \{\bigcup_{i \in I} A_i : I \subset \{1, \dots, n\}\}$ est une tribu. (\mathcal{A} est constitué de toutes les réunions possibles d'ensembles A_i .)
- 4) Soit

$$\begin{aligned} \text{Card} : \mathcal{P}(\mathbb{N}) &\rightarrow [0, +\infty] \\ A &\mapsto \text{Card}(A) = \text{le nombre d'éléments de } A. \end{aligned}$$

Montrer que Card est une mesure sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

- 5) On se donne un espace mesurable (E, \mathcal{A}) .

- (a) Soit $x \in E$, on note

$$\begin{aligned} \delta_x : \mathcal{A} &\rightarrow [0, +\infty] \\ B &\mapsto \delta_x(B) \begin{cases} = 1 & \text{si } x \in B \\ = 0 & \text{sinon.} \end{cases} \end{aligned}$$

Montrer que δ_x est une mesure sur (E, \mathcal{A}) .

- (b) Soient x_1, \dots, x_k des éléments distincts de E et $p_1, \dots, p_k \in \mathbb{R}_+^*$. On note

$$\begin{aligned} \mu : \mathcal{A} &\rightarrow [0, +\infty] \\ B &\mapsto \mu(B) = \sum_{1 \leq i \leq k} p_i \delta_{x_i}(B) \end{aligned}$$

Montrer que μ est une mesure sur (E, \mathcal{A}) .

- 6) Soit $A = \bigcup_{n \geq 0} [n, n + \frac{1}{2^n}[$. Calculer $\lambda(A)$. (On se servira du fait que A est réunion d'ensembles disjoints et on utilisera la propriété d'additivité.)
- 7) (a) Soit $x \in \mathbb{R}$, calculer $\lambda(\{x\})$ (utiliser la propriété de croissance).
- (b) Soit $x_0, x_1, x_2, \dots \in \mathbb{R}$, calculer

$$\lambda(\bigcup_{n \geq 0} \{x_n\})$$

(utiliser la propriété de sous-additivité).

- (c) En déduire que $\lambda(\mathbb{Q}) = 0$. Calculer $\lambda([0, 1] \setminus \mathbb{Q})$.

- 8) Un ensemble de Cantor.

Pour $n \geq 1$, on note :

$$A_n = \{x \in [0, 1[, x \text{ n'a que des } 1 \text{ ou des } 5 \text{ dans son développement décimal jusqu'à l'ordre } n\}$$

A_n est donc l'ensemble des $x \in [0, 1[$ qui s'écrivent $x = 0, u_1 u_2 \dots u_n u_{n+1} \dots$ avec $u_1, \dots, u_n \in \{1, 5\}$.

- (a) Calculer $\lambda(A_n)$ pour tout n .
- (b) Soit $B = \bigcap_{n \geq 1} A_n$, calculer $\lambda(B)$ (utiliser la propriété d'intersection décroissante).
- 9) Mesures à densité.
- (a) Soit μ mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ de densité $\mathbf{1}_{[0,1]}(x)$ par rapport à la mesure de Lebesgue. Calculer $\mu([0, 1]), \mu([0, 2]), \mu([0, 1/2]), \mu(\{1/2\})$.
- (b) Soit μ mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ de densité $\mathbf{1}_{x>0}e^{-x}$ par rapport à la mesure de Lebesgue. Calculer $\mu(\mathbb{R}), \mu(\{1\}), \mu([0, 1]), \mu([1, +\infty[)$.
- (c) Soit μ mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ de densité $\mathbf{1}_{x>0}xe^{-x^2/2}$ par rapport à la mesure de Lebesgue. Calculer $\mu([0, 1])$.

2.6.2 Corrigés

- (1) (a) $\bigcap_{n \geq 0}]1, 1 + 1/(n+1)] = \emptyset$ car $1 \notin \bigcap_{n \geq 0}]1, 1 + 1/(n+1)]$ et $\forall x \neq 1, \exists n$ tel que $x \notin]1, 1 + 1/(n+1)]$ et donc $x \notin \bigcap_{n \geq 0}]1, 1 + 1/(n+1)]$
- (b) $\bigcap_{n \geq 0}]1, 2 + 1/(n+1)] =]1, 2]$
- (c) $\bigcap_{n \geq 0}]1 - 1/(n+1), 2] = [1, 2]$
- (d) $\bigcup_{n \geq 0} [1/(n+1), +\infty[=]0, +\infty[$ donc $f^{-1}(\bigcup_{n \geq 0} [1/(n+1), +\infty]) = f^{-1}(]0, +\infty]) = \mathbb{R} \setminus \{0\} = \mathbb{R}^*$
- (2) (a) Soient $k \neq n, k < n$. $A_k \subset A_{n-1}$ donc $\forall x \in A_k, x \notin B_n$. Comme $B_k \subset A_k$, alors $B_k \cap B_n = \emptyset$
- (b) – Si $x \in (\bigcap_{n \geq 0} A_n)^c$ alors $x \notin \bigcap_{n \geq 0} A_n$ donc $\exists n$ tel que $x \notin A_n$. Donc $\exists n$ tel que $x \in A_n^c$. Donc $x \in \bigcup_{n \geq 0} A_n^c$.
- Si $x \in \bigcup_{n \geq 0} A_n^c$ alors $\exists n$ tel que $x \notin A_n$. Donc $x \notin \bigcap_{n \geq 0} A_n$. Donc $x \in (\bigcap_{n \geq 0} A_n)^c$.
- Conclusion : $\bigcup_{n \geq 0} A_n^c = (\bigcap_{n \geq 0} A_n)^c$.
- (c) Par passage au complémentaire dans le résultat précédent : $(\bigcup_{n \geq 0} A_n^c)^c = \bigcap_{n \geq 0} A_n$.
- (3) On rappelle que " A_1, \dots, A_n partition de \mathbb{R} " signifie que les ensembles A_i sont 2 à 2 disjoints et que $A_1 \cup \dots \cup A_n = \mathbb{R}$.
- (i) $\mathbb{R} = A_1 \cup \dots \cup A_n \in \mathcal{A}$
- (ii) Soit $\bigcup_{i \in I} A_i \in \mathcal{A}$, $(\bigcup_{i \in I} A_i)^c = \bigcap_{i \in I} A_i^c \in \mathcal{A}$.
- (iii) Si on fait une réunion dénombrable d'éléments de \mathcal{A} :

$$\bigcup_{n \geq 0} (\bigcup_{i \in I_n} A_i) = \bigcup_{i \in \bigcup_{n \geq 0} I_n} A_i \in \mathcal{A}.$$

(4) Fait en cours

- (5) (a) Remarque : δ_x s'appelle la mesure de Dirac en x .
- (i) δ_x est bien une fonction de \mathcal{A} dans $[0, +\infty]$
- (ii) $\delta_x(\emptyset) = 0$ car $x \notin \emptyset$
- (iii) Si on a des éléments 2 à 2 disjoints de $\mathcal{A} : A_0, A_1, \dots$

$$\begin{aligned} \delta_x\left(\bigcup_{n \geq 0} A_n\right) &= \begin{cases} = 1 & \text{si } x \in \bigcup_{n \geq 0} A_n \\ = 0 & \text{sinon} \end{cases} \\ &= \begin{cases} = 1 & \text{si } \exists n \text{ tel que } x \in A_n \\ = 0 & \text{sinon} \end{cases} \\ &= \sum_{n \geq 0} \delta_x(A_n) \end{aligned}$$

car les A_n sont 2 à 2 disjoints (et donc au plus un seul d'entre eux contient x , c'est à dire au plus un seul d'entre eux est tel que $\delta_x(A_n) = 1$).

- (b) On remarque que $\forall i, \delta_{x_i}$ est une mesure par la question précédente.
- (i) μ est bien une fonction de \mathcal{A} dans $[0, +\infty]$
- (ii) $\mu(\emptyset) = \sum_{1 \leq i \leq k} p_i \delta_{x_i}(\emptyset) = 0$
- (iii) Si on a des éléments 2 à 2 disjoints de $\mathcal{A} : A_0, A_1, \dots$:

$$\begin{aligned} \mu\left(\bigcup_{n \geq 0} A_n\right) &= \sum_{1 \leq i \leq k} p_i \delta_{x_i}\left(\bigcup_{n \geq 0} A_n\right) \\ &= \sum_{1 \leq i \leq k} p_i \sum_{n \geq 0} \delta_{x_i}(A_n) \\ &= \sum_{n \geq 0} \sum_{1 \leq i \leq k} p_i \delta_{x_i}(A_n) \\ &= \sum_{n \geq 0} \mu(A_n). \end{aligned}$$

- (6) Les ensembles $[n, n + \frac{1}{2^n}[$ sont 2 à 2 disjoints donc $\lambda(A) = \sum_{n \geq 0} \lambda([n, n + \frac{1}{2^n}[) = \sum_{n \geq 0} \frac{1}{2^n} = 2$.
- (7) (a) $\forall \varepsilon > 0, \{x\} \subset [x, x + \varepsilon]$ donc $\lambda(\{x\}) \leq \lambda([x, x + \varepsilon]) = \varepsilon$. Donc $\lambda(\{x\}) = 0$.
 (b) $\lambda(\cup_{n \geq 0} \{x_n\}) \leq \sum_{n \geq 0} \lambda(\{x_n\}) = 0$ par la question précédente.
 (c) \mathbb{Q} est dénombrable donc on peut écrire $\mathbb{Q} = \{x_0, x_1, \dots, x_n, \dots\}$ donc $\lambda(\mathbb{Q}) = 0$ par la question précédente.
- (8) (a) On remarque que

$$\begin{aligned} A_n &= \{[x, x + 10^{-n}[: x = 0, u_1 \dots u_n \text{ avec } u_1, \dots, u_n \in \{1, 5\}\} \\ &= \bigcup_{x \in B_n} [x, x + 10^{-(n+1)}[\end{aligned}$$

où $B_n = \{x = 0, u_1 \dots u_n \text{ avec } u_1, \dots, u_n \in \{1, 5\}\}$. On remarque que B_n est fini et que les intervalles $([x, x + 10^{-n}[)_{x \in B_n}$ sont 2 à 2 disjoints. Donc :

$$\begin{aligned} \lambda(A_n) &= \sum_{x \in B_n} \lambda([x, x + 10^{-n}[) \\ &= \text{Card}(B_n) \times 10^{-n} = 2^n \times 10^{-n} . \end{aligned}$$

- (b) $\forall n, A_n \subset A_{n+1}$ donc par intersection décroissante : $\lambda(B) = \lim_{n \rightarrow +\infty} \lambda(A_n) = 0$.
- (9) (a) $\mu([0, 1]) = \int_{\mathbb{R}} \mathbf{1}_{[0,1]}(x) \mathbf{1}_{[0,1]}(x) dx = \int_0^1 1 dx = 1$
 $\mu([0, 2]) = \int_{\mathbb{R}} \mathbf{1}_{[0,2]}(x) \mathbf{1}_{[0,1]}(x) dx = \int_0^1 1 dx = 1$
 $\mu([0, 1/2]) = \int_{\mathbb{R}} \mathbf{1}_{[0,1/2]}(x) \mathbf{1}_{[0,1]}(x) dx = \int_0^{1/2} 1 dx = 1/2$
 $\mu(\{1/2\}) = \int_{\mathbb{R}} \mathbf{1}_{\{1/2\}}(x) \mathbf{1}_{[0,1]}(x) dx = \int_{\mathbb{R}} \mathbf{1}_{\{1/2\}}(x) dx = 0$ car $\lambda(\{1/2\}) = 0$
- (b) $\mu(\mathbb{R}) = \int_{\mathbb{R}} \mathbf{1}_{x>0}(x) e^{-x} dx = 1$
 $\mu(\{1\}) = \int_{\mathbb{R}} \mathbf{1}_{\{1\}}(x) \mathbf{1}_{x>0}(x) e^{-x} dx = \int_{\mathbb{R}} \mathbf{1}_{\{1\}}(x) e^{-1} dx = 0$ car $\lambda(\{1\}) = 0$
 $\mu([0, 1]) = \int_{\mathbb{R}} \mathbf{1}_{[0,1]}(x) \mathbf{1}_{x>0}(x) e^{-x} dx = \int_0^1 e^{-x} dx = 1 - e^{-1}$
 $\mu([1, +\infty[) = \int_{\mathbb{R}} \mathbf{1}_{[1,+\infty]}(x) \mathbf{1}_{x>0}(x) e^{-x} dx = \int_1^{+\infty} e^{-x} dx = e^{-1}$
- (c) $\mu([0, 1]) = \int_{\mathbb{R}} \mathbf{1}_{[0,1]}(x) \mathbf{1}_{x>0}(x) x e^{-x^2/2} dx = \int_0^1 x e^{-x^2/2} dx = \left[-e^{-x^2/2} \right]_0^1 = (1 - e^{-1/2})$

Chapitre 3

Ensembles négligeables

Définition 3.0.1. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Un élément A de \mathcal{A} est dit négligeable (pour la mesure μ) si $\mu(A) = 0$.

Soit $f : \Omega \rightarrow \mathbb{R}$ une fonction mesurable. Elle est dite μ -presque partout nulle si $\exists A \in \mathcal{A}$ négligeable tel que $x \in A^c \Rightarrow f(x) = 0$. On dira aussi que f est : presque partout nulle, μ -presque sûrement nulle, presque sûrement nulle, p.p. nulle, p.s. nulle.

Remarque 3.0.2. Une fonction positive d'intégrale finie est finie p.p.

Si f est une fonction mesurable positive $\Omega \rightarrow \mathbb{R}^+$ telle que $\exists A \in \mathcal{A}$, $\mu(A) > 0$ et $f(x) = +\infty$ si $x \in A$, alors $\int_{\Omega} f(x) \mu(dx) = +\infty$. En effet, la fonction $\phi(x) = +\infty \times \mathbf{1}_A(x)$ est une fonction étagée vérifiant $\phi \leq f$, $\int_{\Omega} \phi(x) \mu(dx) = +\infty$. D'où $\int_{\Omega} f(x) \mu(dx) = +\infty$ par la définition ci-dessus.

Nous avons donc que si $\int_{\Omega} f(x) \mu(dx) < +\infty$ alors il n'existe pas d'ensemble A ayant les propriétés ci-dessus, ce qui veut donc dire que f est finie presque partout.

Théorème 3.0.3. Espace complet.

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Il existe une tribu \mathcal{B} sur Ω et une mesure ν sur \mathcal{B} telles que

- $\mathcal{A} \subset \mathcal{B}$
- $\forall N \subset \Omega$ tel que $N \subset A$ avec $A \in \mathcal{A}$, $\mu(A) = 0$, on a $N \in \mathcal{B}$ et $\nu(N) = 0$.

La tribu \mathcal{B} est alors appelée tribu complétée de \mathcal{A} et ν est appelée mesure complétée de μ . Un espace mesuré $(\Omega, \mathcal{A}, \mu)$ pour lequel

$$[N \subset A \text{ avec } A \in \mathcal{A}, \mu(A) = 0] \Rightarrow [N \in \mathcal{A}]$$

est appelé un espace mesuré complet.

Théorème 3.0.4. Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré et f fonction mesurable sur cet espace. Alors f est p.p. nulle si et seulement si $\int_{\Omega} f(x) \mu(dx) = 0$.

Démonstration. Nous ne faisons la démonstration que dans le cas $f \geq 0$.

- Si f est p.p. nulle alors $\exists A \in \mathcal{A}$ tel que $\mu(A) = 0$ et f est nulle sur A^c . Soit $\phi \in \mathcal{E}(f)$ et B_1, \dots, B_p partition associée à ϕ . On note $B'_i = B_i \cap A$ et $B''_i = B_i \cap A^c$, $\forall i \in \{1, \dots, p\}$. Les ensembles $B'_1, \dots, B'_p, B''_1, \dots, B''_p$ sont deux à deux disjoints et ϕ est constante sur chacun d'entre eux. Pour $x \in B'_i$, on note $\phi(x) = c_i$. Pour tout x dans B''_1, \dots, B''_p , $f(x) = 0$. Pour tout $i \in \{1, \dots, p\}$, $\mu(B'_i) \leq \mu(A)$ (par proposition 2.2.6) donc $\mu(B'_i) = 0$. Donc

$$\int_{\Omega} \phi(x) \mu(dx) = 0 \times \mu(B''_1) + \dots + 0 \times \mu(B''_p) + c_1 \times \mu(B'_1) + \dots + c_p \times \mu(B'_p) = 0.$$

Cela est vrai pour toute $\phi \in \mathcal{E}(f)$ donc $\int_{\Omega} f(x) \mu(dx) = 0$.

- Si $\int_{\Omega} f(x) \mu(dx) = 0$. Soit $\varepsilon > 0$, soit $A_{\varepsilon} = \{x \in \Omega : f(x) \geq \varepsilon\} = f^{-1}([\varepsilon, +\infty[)$. L'ensemble $[\varepsilon, +\infty[$ appartient à $\mathcal{B}(\mathbb{R})$ car c'est un intervalle. La fonction f est mesurable donc $A_{\varepsilon} \in \mathcal{A}$. Soit ϕ étagée telle que

$$\phi(x) \begin{cases} = 0 & \text{si } x \in A_{\varepsilon}^c \\ = \varepsilon & \text{si } x \in A_{\varepsilon} \end{cases}.$$

L'ensemble A_ε^c appartient à \mathcal{A} . Pour tout x , $\phi(x) \leq f(x)$ donc

$$0 \leq \int_{\Omega} \phi(x) \mu(dx) \leq \int_{\Omega} f(x) \mu(dx)$$

donc $\int_{\Omega} \phi(x) \mu(dx) = 0$. Par ailleurs,

$$\int_{\Omega} \phi(x) \mu(dx) = 0 \times \mu(A_\varepsilon^c) + \varepsilon \times \mu(A_\varepsilon)$$

donc $\mu(A_\varepsilon) = 0$. Les ensembles $A_{1/n}$ pour $n \in \mathbb{N}^*$ vérifient $A_{1/n} \subset A_{1/n+1}$. Donc par la proposition sur la réunion croissante (proposition 2.2.8), $\mu(\{x \in \Omega : f(x) > 0\}) = \mu(\cup_{n \geq 1} A_{1/n}) = \lim_{n \rightarrow +\infty} \mu(A_{1/n}) = 0$. Donc f est nulle p.p. □

Proposition 3.0.5. *Intégrale sur un ensemble négligeable.*

Soit $(\Omega, \mathcal{A}, \mu)$ un espace mesuré. Soit $A \in \mathcal{A}$ négligeable. Soit $f, g : \Omega \rightarrow \mathbb{R}$ mesurables. On suppose que $\int_{\Omega} f(x) \mu(dx)$ est définie (ce qui a lieu, par définition, quand f^+ et f^- sont d'intégrales finies) ainsi que $\int_{\Omega} g(x) \mu(dx)$. On suppose que $f(x) = g(x)$ si $x \notin A$ (donc f et g sont presque partout égales). Alors

$$\int_A f(x) \mu(dx) = 0 ,$$

$$\int_{\Omega} f(x) \mu(dx) = \int_{\Omega} g(x) \mu(dx) .$$

Démonstration. – Par définition,

$$\int_A f(x) \mu(dx) = \int_{\Omega} f(x) \mathbf{1}_A(x) \mu(dx) .$$

Donc par le théorème précédent, $\int_A f(x) \mu(dx) = 0$.

– Par linéarité, $\int_{\Omega} f(x) \mu(dx) - \int_{\Omega} g(x) \mu(dx) = \int_{\Omega} (f(x) - g(x)) \mu(dx)$. La fonction $f - g$ est nulle presque partout donc, par le théorème précédent $\int_{\Omega} (f(x) - g(x)) \mu(dx) = 0$. □

On retient de la proposition précédente que deux fonctions égales presque partout ont la même intégrale.

Chapitre 4

Théorèmes limites

On se donne $(\Omega, \mathcal{A}, \mu)$ un espace mesuré complet. On supposera à partir de maintenant, pour des raisons techniques, que Ω est réunion dénombrable d'éléments de \mathcal{A} de mesure finie. On dit alors que Ω est σ -fini.

4.1 Stabilité de la mesurabilité par passage à la limite.

Théorème 4.1.1. Soit $(f_n)_{n \geq 0}$ une suite de fonctions $\Omega \rightarrow \mathbb{R}$ une suite de fonctions mesurables positives. Alors $\sup_n f_n$ et $\inf_n f_n$ sont des fonctions mesurables.

Démonstration partielle. On pose $f(x) = \sup_n f_n(x)$. Nous allons montrer que $\forall a \in \mathbb{R}$, $f^{-1}(] - \infty, a]) \in \mathcal{A}$. Cela est en fait suffisant pour montrer que f est mesurable mais nous ne démontrerons pas ce point.

Fixons donc $a \in \mathbb{R}$ et prenons $A = f^{-1}(] - \infty, a])$. On remarque que

$$\begin{aligned} A &= \{x \in \Omega : f(x) \leq a\} \\ &= \{x \in \Omega : f_n(x) \leq a, \forall n\} \\ &= \bigcap_{n \geq 0} \{x \in \Omega : f_n(x) \leq a\}. \end{aligned}$$

Pour tout n , $\{x \in \Omega : f_n(x) \leq a\} = f_n^{-1}(A) \in \mathcal{A}$ car f_n est mesurable. La famille \mathcal{A} est une tribu, elle est donc stable par intersection dénombrable donc $A \in \mathcal{A}$. \square

Définition 4.1.2. Soit $(f_n)_{n \geq 0}$ une suite de fonctions $\Omega \rightarrow \mathbb{R}$. On dit que (f_n) convergence presque sûrement vers f (et on note $f_n \xrightarrow[n \rightarrow +\infty]{p.s.} f$) s'il existe A négligeable tel que $[x \notin A] \Rightarrow [f_n(x) \xrightarrow[n \rightarrow +\infty]{} f(x)]$.

Définition 4.1.3. Soit $(f_n)_{n \geq 0}$ une suite de fonctions $\Omega \rightarrow \mathbb{R}$. On dit que (f_n) convergence simplement vers f si $\forall x$, $f_n(x) \xrightarrow[n \rightarrow +\infty]{} f(x)$.

Remarque 4.1.4. La convergence simple implique la convergence presque sûre.

Corollaire 4.1.5. Si on a une suite (f_n) de fonctions $\Omega \rightarrow [0, +\infty[$ mesurables (positives) telle que $f_n \xrightarrow[n \rightarrow +\infty]{p.s.} f$ alors f est mesurable.

Démonstration. On ne va faire la démonstration que dans le cas où (f_n) converge simplement vers f . Pour tout x et pour tout n , on pose $v_n(x) = \sup\{f_n(x), f_{n+1}(x), f_{n+2}(x), \dots\}$. Par le théorème précédent, les fonctions v_n sont mesurables. Pour tout x , $f(x) = \inf\{v_0(x), v_1(x), v_2(x), \dots\}$. Donc par le théorème précédent, f est mesurable. \square

4.2 Théorèmes de convergence pour les intégrales.

Théorème 4.2.1. *Théorème de convergence monotone*

Soit (f_n) une suite croissante (c'est à dire que $\forall x, \forall n, f_n(x) \leq f_{n+1}(x)$) de fonctions

mesurables positives $\Omega \rightarrow [0, +\infty[$ convergeant presque sûrement vers une fonction f . Alors

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n(x) \mu(dx) = \int_{\Omega} f(x) \mu(dx) .$$

Démonstration. Soit $\alpha \in]0, 1[$. La suite $(\int_{\Omega} f_n(x) \mu(dx))$ est croissante (par croissance de l'intégrale) donc elle a une limite $l \in [0, +\infty]$. Soit pour tout n , $A_n = \{x \in \Omega : f_n(x) \geq \alpha f(x)\}$. Pour tout n et pour tout x , $f_n(x) \geq f_n(x) \mathbf{1}_{A_n}(x)$ donc

$$\int_{\Omega} f_n(x) \mu(dx) \geq \int_{\Omega} f_n(x) \mathbf{1}_{A_n}(x) \mu(dx) = \int_{A_n} f_n(x) \mu(dx) \geq \alpha \int_{A_n} f(x) \mu(dx) \quad (4.2.1)$$

Montrons que

$$\int_{A_n} f(x) \mu(dx) \xrightarrow{n \rightarrow +\infty} \int_{\Omega} f(x) \mu(dx) . \quad (4.2.2)$$

Soit $\varepsilon > 0$. Soit ϕ une fonction étagée telle que $\phi \leq f$, $\int_{\Omega} \phi(x) \mu(dx) \geq \int_{\Omega} f(x) \mu(dx) - \varepsilon$ (il en existe par définition de l'intégrale). Nous avons

$$\int_{\Omega} \mathbf{1}_{A_n}(x) \phi(x) \mu(dx) \leq \int_{\Omega} f(x) \mathbf{1}_{A_n}(x) \mu(dx) \leq \int_{\Omega} f(x) \mu(dx) . \quad (4.2.3)$$

On suppose que ϕ se décompose sur une certaine partition B_1, \dots, B_p :

$$\phi(x) = \sum_{1 \leq i \leq p} b_i \mathbf{1}_{B_i}(x) .$$

Alors $\forall n$, $\phi \mathbf{1}_{A_n}$ est une fonction étagée qui se décompose en

$$\phi(x) \mathbf{1}_{A_n}(x) = 0 \times \mathbf{1}_{A_n^c}(x) + \sum_{1 \leq i \leq p} b_i \mathbf{1}_{B_i \cap A_n}(x) .$$

Et donc

$$\int_{\Omega} \phi(x) \mathbf{1}_{A_n}(x) \mu(dx) = 0 \times \mu(A_n^c) + \sum_{1 \leq i \leq p} b_i \times \mu(B_i \cap A_n) \quad (4.2.4)$$

Pour tout n , nous avons $A_n \subset A_{n+1}$ et donc $\forall i$, $B_i \cap A_n \subset B_i \cap A_{n+1}$. Par la propriété de convergence croissante de la mesure,

$$\mu(B_i \cap A_n) \xrightarrow{n \rightarrow +\infty} \mu(\cup_{n \geq 0} (B_i \cap A_n)) = \mu(B_i \cap \cup_{n \geq 0} A_n) . \quad (4.2.5)$$

On remarque que $\cup_{n \geq 0} A_n = \{x \in \Omega : \exists n, f_n(x) \geq \alpha f(x)\} \supset \{x \in \Omega : f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)\}$. Donc $\{x \in \Omega : f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)\}^c \subset (\cup_{n \geq 0} A_n)^c$. Donc $0 = \mu(\{x \in \Omega : f_n(x) \xrightarrow{n \rightarrow +\infty} f(x)\}^c) \geq \mu((\cup_{n \geq 0} A_n)^c)$. Donc $\mu((\cup_{n \geq 0} A_n)^c) = 0$, $\mu(B_i \cap (\cup_{n \geq 0} A_n)^c) \leq \mu((\cup_{n \geq 0} A_n)^c) = 0$. Puis $\mu(B_i) = \mu(B_i \cap (\cup_{n \geq 0} A_n)^c) + \mu(B_i \cap (\cup_{n \geq 0} A_n))$ donc $\mu(B_i) = \mu(B_i \cap \cup_{n \geq 0} A_n)$. On déduit donc de (4.2.4) et (4.2.5)

$$\int_{\Omega} \phi(x) \mathbf{1}_{A_n}(x) \mu(dx) \xrightarrow{n \rightarrow +\infty} \sum_{1 \leq i \leq p} b_i \times \mu(B_i) = \int_{\Omega} \phi(x) \mu(dx) .$$

Donc par (4.2.3)

$$\begin{aligned} \int_{\Omega} f(x) \mu(dx) - \varepsilon &\leq \int_{\Omega} \phi(x) \mu(dx) \\ &\leq \liminf_{n \rightarrow +\infty} \int_{A_n} f(x) \mu(dx) \\ &\leq \limsup_{n \rightarrow +\infty} \int_{A_n} f(x) \mu(dx) \\ &\leq \int_{\Omega} f(x) \mu(dx) \end{aligned}$$

Cela est vrai pour tout $\varepsilon > 0$ donc nous avons donc montré (4.2.2). Alors, par (4.2.1),

$$l \geq \alpha \int_{\Omega} f(x) \mu(dx) . \quad (4.2.6)$$

Pour presque tout x , $f_n(x) \nearrow_{n \rightarrow +\infty} f(x)$ donc $f_n(x) \leq f(x)$. Soit $\forall n$, \bar{f}_n définie par

$$\bar{f}_n(x) = \begin{cases} f_n(x) & \text{si } f_n(x) \leq f(x) \\ 0 & \text{sinon} \end{cases}$$

Les fonctions f_n et \bar{f}_n sont égales presque partout donc leurs intégrales sont égales. La fonction \bar{f}_n vérifie $\bar{f}_n(x) \leq f(x)$ ($\forall x$) donc en particulier

$$\int_{\Omega} f_n(x) \mu(dx) = \int_{\Omega} \bar{f}_n(x) \mu(dx) \leq \int_{\Omega} f(x) \mu(dx) .$$

Donc

$$\int_{\Omega} f(x) \mu(dx) \geq l .$$

Et comme l'équation (4.2.6) est vraie pour tout $\alpha \in]0, 1[$, ceci finit la démonstration. \square

Théorème 4.2.2. Lemme de Fatou

Soit $(f_n)_{n \geq 0}$ une suite de fonctions mesurables positives. On note $f = \liminf_{n \rightarrow +\infty} f_n$. Alors f est mesurable positive et

$$\int_{\Omega} f d\mu \leq \liminf_{n \rightarrow +\infty} \int_{\Omega} f_n d\mu$$

Démonstration. Par définition de la \liminf , nous avons pour tout x ,

$$f(x) = \lim_{n \rightarrow +\infty} \left(\inf_{k \geq n} f_k(x) \right)$$

(cette limite existe dans $[-\infty, +\infty[$ car c'est la limite d'une suite décroissante). Par le théorème 4.1.1, les fonctions $x \mapsto \inf_{k \geq n} f_k(x)$ sont mesurables pour tout n . Par le corollaire 4.1.5, la fonction f est mesurable.

Soit $m \geq 1$. Soit pour tout n , $A_n = \{x : \forall p \geq n, f_p(x) \geq (f(x) - \frac{1}{m})_+\}$. Pour tout x , $\exists N \in \mathbb{N}$ tel que $n \geq N \Rightarrow f_n(x) \geq f(x) - \frac{1}{m}$. Nous avons donc $\bigcup_{n \geq 1} A_n = \Omega$. On remarque que pour tout n , $A_n \subset A_{n+1}$. Et donc pour tout x ,

$$\left(f(x) - \frac{1}{m} \right)_+ \mathbf{1}_{A_n}(x) \nearrow_{n \rightarrow +\infty} \left(f(x) - \frac{1}{m} \right)_+ .$$

Donc, par théorème de convergence monotone,

$$\int_{\Omega} \left(f(x) - \frac{1}{m} \right)_+ \mathbf{1}_{A_n}(x) \mu(dx) \xrightarrow{n \rightarrow +\infty} \int_{\Omega} \left(f(x) - \frac{1}{m} \right)_+ \mu(dx) .$$

Pour tout n , nous avons

$$\int_{\Omega} f_n(x) \mu(dx) \geq \int_{\Omega} f_n(x) \mathbf{1}_{A_n}(x) \mu(dx) \geq \int_{\Omega} \left(f(x) - \frac{1}{m} \right)_+ \mathbf{1}_{A_n}(x) \mu(dx)$$

et donc

$$\liminf_{n \rightarrow +\infty} \int_{\Omega} f_n(x) \mu(dx) \geq \int_{\Omega} \left(f(x) - \frac{1}{m} \right)_+ \mu(dx) .$$

Nous avons pour tout x , $\left(f(x) - \frac{1}{m} \right)_+ \nearrow_{m \rightarrow \infty} f(x)$. Donc, par théorème de convergence monotone, $\int_{\Omega} \left(f(x) - \frac{1}{m} \right)_+ \mu(dx) \xrightarrow{m \rightarrow \infty} \int_{\Omega} f(x) \mu(dx)$. Et donc

$$\liminf_{n \rightarrow +\infty} \int_{\Omega} f_n(x) \mu(dx) \geq \int_{\Omega} f(x) \mu(dx) .$$

\square

Théorème 4.2.3. *Théorème de convergence dominée (appelé aussi théorème de Lebesgue)*

Soit $(f_n)_{n \geq 0}$ une suite de fonctions mesurables sur Ω . Si :

- il existe g positive mesurable et intégrable telle que $\forall n \in \mathbb{N}, \forall x \in \Omega, |f_n(x)| \leq g(x)$
- et $f_n \xrightarrow[n \rightarrow +\infty]{p.s.} f$

alors

- $\int_{\Omega} |f(x)| \mu(dx) < \infty$
- $\lim_{n \rightarrow +\infty} \int_{\Omega} |f_n(x) - f(x)| \mu(dx) = 0$.

Ce qui implique en particulier

$$\lim_{n \rightarrow +\infty} \int_{\Omega} f_n(x) \mu(dx) = \int_{\Omega} f(x) \mu(dx) .$$

Démonstration. Pour simplifier la démonstration, nous allons supposer que (f_n) converge simplement vers f . Nous avons alors pour tout x , $|f(x)| \leq g(x)$, donc $\int_{\Omega} |f(x)| \mu(dx) < \infty$. Pour tout x , $2g(x) - |f(x) - f_n(x)| \geq 0$ et $\liminf_{n \rightarrow +\infty} (2g(x) - |f(x) - f_n(x)|) = 2g(x)$ donc par le lemme de Fatou

$$\liminf_{n \rightarrow +\infty} \int_{\Omega} (2g(x) - |f(x) - f_n(x)|) \mu(dx) \geq \int_{\Omega} 2g(x) \mu(dx) .$$

Mais par linéarité de l'intégrale,

$$\liminf_{n \rightarrow +\infty} \int_{\Omega} (2g(x) - |f(x) - f_n(x)|) \mu(dx) = \int_{\Omega} 2g(x) \mu(dx) - \limsup_{n \rightarrow +\infty} \int_{\Omega} |f(x) - f_n(x)| \mu(dx) .$$

Donc

$$\begin{aligned} \limsup_{n \rightarrow +\infty} \int_{\Omega} |f(x) - f_n(x)| \mu(dx) &= 0 \\ \lim_{n \rightarrow +\infty} \int_{\Omega} |f(x) - f_n(x)| \mu(dx) &= 0 . \end{aligned}$$

Puis

$$\begin{aligned} \left| \int_{\Omega} f_n(x) \mu(dx) - \int_{\Omega} f(x) \mu(dx) \right| &= \left| \int_{\Omega} f(x) - f_n(x) \mu(dx) \right| \\ (\text{par lemme 2.4.12}) &\leq \int_{\Omega} |f(x) - f_n(x)| \mu(dx) \xrightarrow[n \rightarrow +\infty]{} 0 . \end{aligned}$$

□

Exemple 4.2.4. Soit l'espace mesuré $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{card})$. Soit $f(k) = \frac{1}{(k+1)^2}$ et pour tout $n \geq 0$, $f_n(k) = \frac{1}{(k+1)^2} \mathbf{1}_{k \leq n}$. Pour tout k , $f_n(k) \nearrow_{n \rightarrow +\infty} f(k)$. Fixons $n \geq 0$, la fonction f_n est étagée et son intégrale vaut

$$\begin{aligned} \int_{\mathbb{N}} f_n(x) \text{card}(dx) &= \frac{1}{1} \times \text{card}(\{0\}) + \frac{1}{2^2} \times \text{card}(\{1\}) + \dots \\ &\quad \dots + \frac{1}{(n+1)^2} \times \text{card}(\{n\}) + 0 \times \text{card}(\{n+1, n+2, \dots\}) \\ &= \sum_{k=1}^n \frac{1}{(k+1)^2} . \end{aligned}$$

Par théorème de convergence monotone,

$$\int_{\mathbb{N}} f_n(x) \text{card}(dx) \xrightarrow[n \rightarrow +\infty]{} \int_{\mathbb{N}} f(x) \text{card}(dx)$$

et donc

$$\int_{\mathbb{N}} f(x) \text{card}(dx) = \sum_{k=0}^{+\infty} \frac{1}{(k+1)^2} .$$

On peut ainsi montrer que pour n'importe quelle fonction $g : \mathbb{N} \rightarrow \mathbb{R}^+$,

$$\int_{\mathbb{N}} g(x) \text{card}(dx) = \sum_{k=0}^{+\infty} g(k)$$

et donc, pour l'espace mesuré $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{card})$, calculer une intégrale d'une fonction positive revient à faire la somme d'une série.

Exemple 4.2.5. Soit l'espace mesuré $([0, 1], \mathcal{B}([0, 1]), \lambda)$. Soient les fonctions (pour $n \geq 1$)

$$\begin{aligned} f_n & : [0, 1] \rightarrow \mathbb{R}^+ \\ x & \mapsto 1 - x^{1/n} \end{aligned}$$

Pour tout $x \in]0, 1]$, $\lim_{n \rightarrow +\infty} f_n(x) = 0$ et $f_n(0) = 1$ pour tout $n \geq 1$. Donc $f_n \xrightarrow[n \rightarrow +\infty]{p.s.} f$ (sur $[0, 1]$) avec f la fonction nulle. Pour tout $n \geq 1$, $|f_n(x)| \leq 1$ qui est une fonction intégrable sur $[0, 1]$. En effet

$$\int_{[0,1]} 1 dx = 1 < \infty .$$

Donc, par théorème de convergence dominée,

$$\int_{[0,1]} f_n(x) \mu(dx) \xrightarrow[n \rightarrow +\infty]{} 0 .$$

4.3 Intégrales dépendant d'un paramètre

Soit $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, on définit une fonction $F(u) = \int_{\mathbb{R}} f(u, x) \lambda(dx)$. Cette fonction F s'appelle, suivant les auteurs, une « intégrale à paramètre », « intégrale dépendant d'un paramètre », ... Dans cette partie, nous allons démontrer diverses propriétés des intégrales à paramètre à l'aide du théorème de convergence dominée.

Théorème 4.3.1. *Continuité sous l'intégrale*

Soit $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ telle que

- (i) $\forall u \in \mathbb{R}, x \mapsto f(u, x)$ est mesurable
- (ii) $\exists u_{\infty}$ tel que pour presque tout x , $u \mapsto f(u, x)$ est continue en u_{∞}
- (iii) $\exists g$ positive intégrable telle que $\forall u \in \mathbb{R}, |f(u, x)| \leq g(x)$.

Alors la fonction F définie par $F(u) = \int_{\mathbb{R}} f(u, x) \lambda(dx)$ est définie en tout point $u \in \mathbb{R}$ et est continue en u_{∞} .

Démonstration. Il suffit de montrer que $F(u_n) \xrightarrow[n \rightarrow +\infty]{} F(u_{\infty})$ pour toute suite $(u_n)_{n \geq 0}$ convergeant vers u_{∞} . Prenons donc une telle suite $(u_n)_{n \geq 0}$. Posons $\forall n, f_n(x) = f(u_n, x)$. Nous avons $f_n \xrightarrow[n \rightarrow +\infty]{p.s.} h$ avec $h(x) := f(u_{\infty}, x)$ par (ii). Les fonctions f_n sont mesurables par (i). Par (iii), nous avons $\forall n, \forall x, |f_n(x)| \leq g(x)$ avec g intégrable. Donc par théorème de convergence dominée,

$$F(u_n) = \int_{\mathbb{R}} f_n(x) \lambda(dx) \xrightarrow[n \rightarrow +\infty]{} \int_{\mathbb{R}} h(x) \lambda(dx) = F(u_{\infty}) .$$

□

Corollaire 4.3.2. *Théorème de continuité « globale » sous l'intégrale*

Soit $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ telle que

- (i) $\forall u \in \mathbb{R}, x \mapsto f(u, x)$ est mesurable
- (ii) pour presque tout x , $u \mapsto f(u, x)$ est continue
- (iii) $\exists g$ positive intégrable telle que $\forall u \in \mathbb{R}, |f(u, x)| \leq g(x)$.

Alors la fonction F définie par $F(u) = \int_{\mathbb{R}} f(u, x) \lambda(dx)$ est définie et continue en tout point $u \in \mathbb{R}$.

Remarque 4.3.3. Ces théorèmes restent vrais si on remplace $u \in \mathbb{R}$ par $u \in I$ avec I intervalle ouvert de \mathbb{R} .

Exemple 4.3.4. *Convolution*

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ intégrable et $\phi : \mathbb{R} \rightarrow \mathbb{R}$ bornée et continue. La convolée de f et ϕ est définie par

$$u \mapsto (f \star \phi)(u) := \int_{\mathbb{R}} \phi(u-x)f(x)\lambda(dx)$$

Notons $h(u, x) = \phi(u-x)f(x)$. Pour tout x , $u \mapsto \phi(u-x)f(x)$ est continue. Pour tout u , $|\phi(u-x)f(x)| \leq \|\phi\|_{\infty}|f(x)|$ et $\int_{\Omega} \|\phi\|_{\infty}|f(x)|\lambda(dx) < \infty$ par hypothèse. On rappelle que $\|\phi\|_{\infty} := \sup_{v \in \mathbb{R}} \phi(v)$. Pour tout $u \in \mathbb{R}$, $x \mapsto \phi(u-x)f(x)$ est mesurable comme produit de fonctions mesurables. Donc par le théorème de continuité globale, $f \star \phi$ est continue sur \mathbb{R} .

Théorème 4.3.5. *Dérivation sous l'intégrale*

Soit I un intervalle ouvert non vide de \mathbb{R} , $u_{\infty} \in I$. Soit $f : I \times \mathbb{R} \rightarrow \mathbb{R}$ telle que

- (i) $\forall u \in I$, $x \mapsto f(u, x)$ est intégrable
- (ii) pour presque tout x , $\frac{\partial f}{\partial u}(u_{\infty}, x)$ existe
- (iii) $\exists g$ positive intégrable telle que $\forall u \in I$, $\forall x \in \mathbb{R}$, $|f(u, x) - f(u_{\infty}, x)| \leq g(x)|u - u_{\infty}|$.

Alors $F(u) := \int_{\mathbb{R}} f(u, x)\lambda(dx)$ existe pour tout $u \in I$ et est dérivable en u_{∞} . De plus

$$F'(u_{\infty}) = \int_{\mathbb{R}} \frac{\partial f}{\partial u}(u_{\infty}, x)\lambda(dx) .$$

Démonstration. L'existence de F est assurée par (i).

En ce qui concerne la dérivation, il suffit de montrer que pour toute suite $(u_n)_{n \geq 0}$ convergeant vers u_{∞} avec $\forall n$, $u_n \neq u_{\infty}$, $\frac{F(u_n) - F(u_{\infty})}{u_n - u_{\infty}} \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}} \frac{\partial f}{\partial u}(u_{\infty}, x)\lambda(dx)$. Prenons donc une telle suite $(u_n)_{n \geq 0}$. Posons $\forall n$

$$\phi_n(x) = \frac{f(u_n, x) - f(u_{\infty}, x)}{u_n - u_{\infty}} .$$

Par (ii), $\phi_n \xrightarrow[n \rightarrow +\infty]{p.s.} \frac{\partial f}{\partial u}(u_{\infty}, \cdot)$. Par (iii), nous avons pour p.t. x , $|\phi_n(x)| \leq g(x)$. Donc par théorème de convergence dominée,

$$\frac{F(u_n) - F(u_{\infty})}{u_n - u_{\infty}} = \int_{\mathbb{R}} \frac{f(u_n, x) - f(u_{\infty}, x)}{u_n - u_{\infty}} \lambda(dx) \xrightarrow{n \rightarrow +\infty} \int_{\mathbb{R}} \frac{\partial f}{\partial u}(u_{\infty}, x)\lambda(dx) .$$

□

Corollaire 4.3.6. *Dérivation « globale » sous l'intégrale*

Soit I un intervalle ouvert non vide de \mathbb{R} , $u_{\infty} \in I$. Soit $f : I \times \mathbb{R} \rightarrow \mathbb{R}$ telle que

- (i) $\exists u_0 \in I$, $x \mapsto f(u_0, x)$ est intégrable
- (ii) pour p.t. x , $u \mapsto f(u, x)$ est dérivable sur I
- (iii) $\forall x, \forall u$, $\left| \frac{\partial f}{\partial u}(u, x) \right| \leq g(x)$ avec g intégrable.

Alors $F(u) := \int_{\mathbb{R}} f(u, x)\lambda(dx)$ existe et est dérivable sur I . De plus

$$F'(u) = \int_{\mathbb{R}} \frac{\partial f}{\partial u}(u, x)\lambda(dx) .$$

Démonstration. Pour tout $u \in I$,

$$\begin{aligned} |f(u, x)| &\leq |f(u_0, x)| + |f(u, x) - f(u_0, x)| \\ &\leq |f(u_0, x)| + |u - u_0| \sup_{v \in [u, u_0]} \left| \frac{\partial f}{\partial u}(v, x) \right| \\ &\leq |f(u_0, x)| + |u - u_0| g(x) . \end{aligned}$$

Donc, par (i) et (iii), F est bien définie. Pour tous $u, u_\infty \in I$, pour tout x ,

$$\begin{aligned} |f(u, x) - f(u_\infty, x)| &\leq |u - u_\infty| \sup_{v \in [u, u_0]} \left| \frac{\partial f}{\partial u}(v, x) \right| \\ &\leq g(x) |u - u_\infty| \end{aligned}$$

par (iii). Et le théorème précédent finit la démonstration. \square

Exemple 4.3.7. Soit, pour $u > 0$, $F(u) = \int_0^{+\infty} e^{-ut} \times \frac{\sin(t)}{t} dt$. La fonction $x \mapsto e^{-1 \times t} \times \frac{\sin(t)}{t}$ est intégrable sur $]0, +\infty[$ car $\left| e^{-1 \times t} \times \frac{\sin(t)}{t} \right| \leq e^{-t}$ (car $\left| \frac{\sin(t)}{t} \right| \leq 1$). Pour tout $t > 0$, $u \mapsto e^{-ut} \times \frac{\sin(t)}{t}$ est dérivable sur $]0, +\infty[$ et $\frac{\partial}{\partial u} \left(e^{-ut} \times \frac{\sin(t)}{t} \right) = -e^{-ut} \sin(t)$.

Soit $\varepsilon > 0$. Pour tout $u > \varepsilon$, $|-e^{-ut} \sin(t)| \leq e^{-\varepsilon t}$ (car $|\sin(t)| \leq 1$) qui est intégrable sur $]0, +\infty[$. Donc par théorème de dérivation globale, nous avons pour $u > \varepsilon$

$$F'(u) = \int_0^{+\infty} -e^{-ut} \sin(t) dt .$$

Cela est vrai $\forall \varepsilon > 0$ donc $\forall u > 0$, $F'(u) = \int_0^{+\infty} -e^{-ut} \sin(t) dt$. Calculons

$$\begin{aligned} F'(u) &= [e^{-ut} \cos(t)]_0^{+\infty} + \int_0^{+\infty} u e^{-ut} \cos(t) dt \\ &= -1 + [u e^{-ut} \sin(t)]_0^{+\infty} + \int_0^{+\infty} u^2 e^{-ut} \sin(t) dt \\ &= -1 - u^2 F'(u) . \end{aligned}$$

Donc $F'(u) = \frac{-1}{1+u^2}$. Donc il existe une constante C telle que $F(u) = C - \arctan(u)$. Posons pour $n \in \mathbb{N}^*$, $f_n(t) = \exp(-nt) \frac{\sin(t)}{t}$. Les fonctions f_n sont mesurables. Pour tout $t > 0$, $f_n(t) \xrightarrow{n \rightarrow +\infty} 0$ et $|f_n(t)| \leq e^{-t} \times 1$ qui est intégrable sur $[0, +\infty[$. Donc, par théorème de convergence dominée, $F(n) = \int f_n(t) \mu(dt) \xrightarrow{n \rightarrow +\infty} 0$. Nous avons $\lim_{n \rightarrow +\infty} \arctan(n) = \frac{\pi}{2}$ donc $C = \frac{\pi}{2}$. Donc

$$F(u) = \frac{\pi}{2} - \arctan(u) .$$

4.4 Exercices

4.4.1 Énoncés

1) Calculer les limites suivantes :

- (a) $\lim_{n \rightarrow +\infty} \int_0^1 \frac{1}{\sqrt{x}} \sin\left(\frac{1}{nx}\right) dx$
- (b) $\lim_{n \rightarrow +\infty} \int_0^1 \left(1 - \frac{x}{n}\right)^n dx$
- (c) $\lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} \sin\left(\frac{x}{n}\right) \frac{n}{x(1+x^2)} dx$
- (d) $\lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} e^{1+\cos^2(x)} e^{-|x|} dx$.
- (e) $\lim_{n \rightarrow +\infty} \int_0^{+\infty} \arctan(x/n) e^{-x} dx$

2) On pose : $I(\alpha) = \lim_{n \rightarrow +\infty} \int_0^n \left(1 - \frac{x}{n}\right)^n e^{\alpha x} dx$ pour $n \in \mathbb{N}$ et $\alpha \in \mathbb{R}$.

- (a) On pose pour $n \in \mathbb{N}$, $f_n : \mathbb{R}^+ \rightarrow \mathbb{R}$ telle que $f_n(x) = \left(1 - \frac{x}{n}\right)^n e^{\alpha x} \mathbf{1}_{x \leq n}$. Montrer que $(f_n)_{n \geq 0}$ est une suite croissante de fonctions. (On pourra notamment étudier : $g_n(x) = (n+1) \ln\left(1 - \frac{x}{n+1}\right) - n \ln\left(1 - \frac{x}{n}\right)$.)
- (b) En déduire la valeur de $I(\alpha)$ en fonction de α .

3) Soit μ la mesure de comptage ("Card") sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Pour toute suite $(u_n)_{n \geq 0}$, on a : $\sum_{n \geq 0} u_n = \int_{\mathbb{N}} u_n d\mu$.

- (a) Calculer $\lim_{k \rightarrow +\infty} \left[\sum_{n \geq 0} \frac{1}{3^n} \left(1 - \frac{1}{k(n+1)} \right) \right]$.
- (b) Calculer $\lim_{k \rightarrow +\infty} \left[\sum_{n \geq 0} \frac{\sin(n/k)}{2^n} \right]$.

4) Inégalité de Jensen.

Soit (E, \mathcal{A}, μ) un espace mesuré avec $\mu(E) = 1$. Soit $\phi : \mathbb{R} \rightarrow \mathbb{R}^+$ convexe et dérivable deux fois (et donc $\phi'' \geq 0$). Soit (E, \mathcal{A}, μ) un espace mesuré avec $\mu(E) = 1$. Soit $f : (E, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mesurable et telle que $\int_E f(x) d\mu(x) < +\infty$.

- (a) Montrer que $\forall z, y \in I$ avec $z \leq y$, $\phi(y) \geq \phi(z) + \phi'(z)(y - z)$
- (b) En prenant $z = \int_E f(t) d\mu(t)$ et $y = f(x)$ dans l'inégalité précédente, montrer que :

$$\phi \left(\int_E f(x) d\mu(x) \right) \leq \int_E \phi \circ f(x) d\mu(x).$$

- (c) En déduire que pour toute fonction $f : [0, 1] \rightarrow \mathbb{R}$ telle que $\int_0^1 |f(x)| dx < +\infty$:

$$\left(\int_0^1 |f(x)| dx \right)^2 \leq \int_0^1 f(x)^2 dx.$$

- 5) (a) Montrer que $\forall z \geq 0$, $0 \leq 1 - e^{-z} \leq z$.
- (b) En déduire que $\forall y > 0$, $x \mapsto \frac{1 - e^{-x^2 y}}{x^2}$ est intégrable sur $[0, +\infty[$.
- (c) Pour tout $y > 0$, on pose

$$F(y) = \int_0^{+\infty} \frac{1 - e^{-x^2 y}}{x^2} dx.$$

Montrer que F est dérivable sur $]0, +\infty[$. Calculer $F'(y)$. On rappelle que

$$\int_0^{+\infty} e^{-x^2} dx = \sqrt{\pi}/2.$$

- (d) En déduire $F(y)$ à une constante près.
- (e) Calculer cette constante en regardant $\lim_{n \rightarrow +\infty} F(1/n)$.

4.4.2 Corrigés

- (1) (a) $\forall x \in]0, 1]$, $\left| \frac{1}{\sqrt{x}} \sin(1/nx) \right| \leq \frac{1}{\sqrt{x}}$ et $\frac{1}{\sqrt{x}}$ intégrable sur $[0, 1]$. $\forall x \in]0, 1]$,

$$\frac{1}{\sqrt{x}} \sin(1/nx) \xrightarrow{n \rightarrow +\infty} 0$$

donc par convergence dominée $\lim_{n \rightarrow +\infty} \int_0^1 \frac{1}{\sqrt{x}} \sin\left(\frac{1}{nx}\right) dx = 0$

- (b) $\forall x \in [0, 1]$, $\left| \left(1 - \frac{x}{n}\right)^n \right| \leq 1$ et la fonction constante égale à 1 est intégrable sur $[0, 1]$. On a $\forall x \in [0, 1]$, $\left(1 - \frac{x}{n}\right)^n = \exp(n \log(1 - \frac{x}{n})) = \exp(n(-x/n + o(1/n))) = \exp(-x + o(1)) \xrightarrow{n \rightarrow +\infty} e^{-x}$ par continuité de la fonction exponentielle. Donc par convergence dominée,

$$\int_0^1 \left(1 - \frac{x}{n}\right)^n dx \xrightarrow{n \rightarrow +\infty} \int_0^{+\infty} e^{-x} dx = 1.$$

- (c) $\forall x \in \mathbb{R}$, $\left| \sin\left(\frac{x}{n}\right) \frac{n}{x(1+x^2)} \right| \leq \frac{1}{(1+x^2)}$ qui est une fonction intégrable sur $] -\infty, +\infty[$.
 $\forall x \in \mathbb{R}$, $\sin\left(\frac{x}{n}\right) \frac{n}{x(1+x^2)} \xrightarrow{n \rightarrow +\infty} \frac{1}{(1+x^2)}$ car $\sin(u) \underset{u \rightarrow 0}{\sim} u$ donc par convergence dominée,

$$\lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} \sin\left(\frac{x}{n}\right) \frac{n}{x(1+x^2)} dx = \int_{-\infty}^{+\infty} \frac{1}{(1+x^2)} dx = [\arctan(x)]_{-\infty}^{+\infty} = \pi.$$

(d) $\forall x \in \mathbb{R}$,

$$e^{1+\cos^{2n}(x)}e^{-|x|} \leq e^{2-|x|}$$

qui est une fonction intégrable sur \mathbb{R} . Pour tout $x \in \mathbb{R}$, $e^{1+\cos^{2n}(x)}e^{-|x|} \xrightarrow{n \rightarrow +\infty} e^{1-|x|}$ donc par convergence dominée,

$$\lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} e^{1+\cos^{2n}(x)}e^{-|x|} dx = \int_{-\infty}^{+\infty} e^{1-|x|} dx = 2e^1.$$

(e) $\forall x \geq 0$, $\arctan(x/n)e^{-x} \leq (\pi/2)e^{-x}$ qui est une fonction intégrable sur $[0, +\infty[$.
Pour tout $x \geq 0$, $\arctan(x/n)e^{-x} \xrightarrow{n \rightarrow +\infty} 0$ donc par convergence dominée,

$$\lim_{n \rightarrow +\infty} \int_0^{+\infty} \arctan(nx)e^{-x} dx = 0.$$

(2) (a) On a pour $0 \leq x \leq n$, $f_{n+1}(x)/f_n(x) = \exp(g_n(x))$.

$$g'_n(x) = \left(\frac{1}{n} - \frac{1}{n+1} \right) \frac{x}{(1-x/n)(1-x/(n+1))} \geq 0$$

pour $0 \leq x \leq n$ donc g_n croissante sur $[0, n]$. $g_n(0) = 0$ donc $g_n(x) \geq 0 \forall x \in [0, n]$.
Donc $f_{n+1}(x) \geq f_n(x) \forall x \in [0, n]$. C'est également vrai sur $[n, +\infty[$ donc f_n suite de fonctions croissantes.

(b) On a $\int_0^n (1 - \frac{x}{n})^n e^{\alpha x} dx = \int_0^{+\infty} f_n(x) dx$. $\forall x \geq 0$, $f_n(x) \xrightarrow{n \rightarrow +\infty} e^{-x+\alpha x}$ donc par convergence monotone, $\lim_{n \rightarrow +\infty} \int_0^{+\infty} f_n(x) dx = \int_0^{+\infty} e^{-x+\alpha x} dx$, donc :

$$I(\alpha) \begin{cases} +\infty & \text{si } \alpha \geq 1 \\ \frac{1}{-1+\alpha} & \text{sinon.} \end{cases}$$

(3) (a) Pour tout n, k , $0 \leq \frac{1}{3^n} \left(1 - \frac{1}{k(n+1)} \right) \leq \frac{1}{3^n}$ qui est le terme général d'une série convergente. Pour tout n , $\frac{1}{3^n} \left(1 - \frac{1}{k(n+1)} \right) \xrightarrow{k \rightarrow +\infty} \frac{1}{3^n}$ donc par convergence dominée :

$$\lim_{k \rightarrow +\infty} \left[\sum_{n \geq 0} \frac{1}{3^n} \left(1 - \frac{1}{k(n+1)} \right) \right] = \sum_{n \geq 0} \frac{1}{3^n} = \frac{3}{2}.$$

(b) Pour tout n, k , $\left| \frac{\sin(n/k)}{2^n} \right| \leq \frac{1}{2^n}$ qui est le terme général d'une série convergente.
Pour tout n , $\frac{\sin(n/k)}{2^n} \xrightarrow{k \rightarrow +\infty} 0$ donc par convergence dominée :

$$\lim_{k \rightarrow +\infty} \left[\sum_{n \geq 0} \frac{\sin(n/k)}{2^n} \right] = 0.$$

(4) Inégalité de Jensen.

(a) $\forall z, y \in I$ avec $z \leq y$, $\phi(y) - \phi(z) = \int_z^y \phi'(t) dt \geq \int_z^y \phi'(z) dt$ (car ϕ convexe), donc $\phi(y) - \phi(z) \geq \phi'(z)(y - z)$

(b) On prend $z = \int_E f(t) d\mu(t)$ et $y = f(x)$ dans l'inégalité précédente et on a :

$$\phi(f(x)) \geq \phi \left(\int_E f(t) d\mu(t) \right) + \phi' \left(\int_E f(t) d\mu(t) \right) (y - z).$$

On intègre ensuite par rapport à $d\mu(x)$:

$$\begin{aligned}
 \int \phi(f(x))d\mu(x) &\geq \int \phi\left(\int_E f(t)d\mu(t)\right)d\mu(x) \\
 &\quad + \int \phi'\left(\int_E f(t)d\mu(t)\right)(y-z)d\mu(x) \\
 &= \phi\left(\int_E f(t)d\mu(t)\right) \\
 &\quad + \phi'\left(\int_E f(t)d\mu(t)\right)\left(\int f(x)d\mu(x) - \int f(x)d\mu(x)\right) \\
 &= \phi\left(\int_E f(t)d\mu(t)\right).
 \end{aligned}$$

- (c) La fonction $\phi : x \in [0, 1] \mapsto x^2$ est convexe. Donc par le résultat précédent, pour toute fonction $f : [0, 1] \rightarrow \mathbb{R}$ intégrable,

$$\left(\int_0^1 |f(x)|dx\right)^2 \leq \int_0^1 f(x)^2 dx.$$

(5) (a) $0 \leq 1 - e^{-z} = \int_0^z e^{-t} dt \leq \int_0^z 1 dt = z$

- (b) Par la question précédente, $\forall y > 0$, $0 \leq \frac{1-e^{-x^2 y}}{x^2} \leq y$ et $\leq \frac{1}{x^2}$ donc $0 \leq \frac{1-e^{-x^2 y}}{x^2} \leq \inf(1, 1/x^2)$ donc $x \mapsto \frac{1-e^{-x^2 y}}{x^2}$ est intégrable

- (c) Soit $\varepsilon > 0$,

– $\forall y > \varepsilon$, $x \mapsto \frac{1-e^{-x^2 y}}{x^2}$ est intégrable

– $\forall x > 0$ (et donc pour presque tout $x \geq 0$), $y \mapsto \frac{1-e^{-x^2 y}}{x^2}$ est dérivable

– $\forall x > 0$, $\forall y > \varepsilon$, $\frac{\partial}{\partial y} \left(\frac{1-e^{-x^2 y}}{x^2} \right) = e^{-x^2 y}$ et $|e^{-x^2 y}| \leq e^{-\varepsilon x^2}$ qui est intégrable sur $[0, +\infty[$

Donc (théorème de dérivation globale) F est dérivable sur $]\varepsilon, +\infty[$ et F' vaut :

$$F'(y) = \int_0^{+\infty} e^{-x^2 y} dx$$

Cela est vrai $\forall \varepsilon > 0$ donc cette dérivée est valable pour tout $y \in]0, +\infty[$. Par changement de variable ($u = \sqrt{y}x$), $F'(y) = \frac{1}{\sqrt{y}} \int_0^{+\infty} e^{-u^2} du = \frac{\sqrt{\pi}}{2\sqrt{y}}$.

- (d) On en déduit $F(y) = \sqrt{\pi y} + C$ pour une certaine constante C .

- (e) $F(1/n) = \int_0^{+\infty} f_n(x) dx$ avec $f_n(x) = \frac{1-e^{-x^2/n}}{x^2}$. Pour tout $x > 0$, $f_n(x) \xrightarrow{n \rightarrow +\infty} 0$.

Pour tout $x > 0$, $|f_n(x)| \leq \inf(1, 1/x^2)$ (voir question 1). Donc, par théorème de convergence dominée :

$$F(1/n) \xrightarrow{n \rightarrow +\infty} 0$$

donc $C = 0$.

Chapitre 5

Mesure produit et théorèmes de Fubini

On se donne deux espaces mesurés $(\Omega, \mathcal{A}, \mu)$ et $(\Omega', \mathcal{A}', \mu')$.

5.1 Théorèmes de Fubini et Fubini-Tonelli

Théorème 5.1.1. *Sur l'ensemble $\Omega \times \Omega'$, il existe une « plus petite tribu » \mathcal{C} contenant tous les ensembles de la forme $A \times B$ avec $A \in \mathcal{A}$, $B \in \mathcal{A}'$. On note $\mathcal{C} = \mathcal{A} \otimes \mathcal{A}'$.*

Il existe une unique mesure, notée $\mu \otimes \mu'$ sur \mathcal{C} telle que, si $(A, A') \in \mathcal{A} \times \mathcal{A}'$, $\mu \otimes \mu'(A \times A') = \mu(A)\mu'(A')$.

Définition 5.1.2. *La mesure $\mu \otimes \mu'$ définie par le théorème ci-dessus s'appelle la mesure produit de μ et μ' . La tribu \mathcal{C} définie par le théorème ci-dessus s'appelle la tribu produit.*

Définition 5.1.3. *On notera $\mathcal{B}(\mathbb{R}^d)$ la tribu $\mathcal{B}(\mathbb{R}) \otimes \dots \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R})^{\otimes d}$ (produit d fois).*

La mesure $\lambda \otimes \lambda$ sur $\mathcal{B}(\mathbb{R}^2)$ mesure les aires, la mesure $\lambda \otimes \lambda \otimes \lambda = \lambda^{\otimes 3}$ sur $\mathcal{B}(\mathbb{R}^3)$ mesure les volumes, ...

Théorème 5.1.4. *Théorème de Fubini-Tonelli*

Soit $f : \Omega \times \Omega' \rightarrow [0, +\infty]$ mesurable positive. On définit les fonctions ϕ et ψ sur Ω et Ω' respectivement par

$$\phi(x) = \int_{\Omega'} f(x, y) \mu'(dy), \quad \psi(y) = \int_{\Omega} f(x, y) \mu(dx) .$$

Ces fonctions sont mesurables positives et vérifient

$$\int_{\Omega} \phi(x) \mu(dx) = \int_{\Omega \times \Omega'} f(x, y) \mu \otimes \mu'(dx, dy) = \int_{\Omega'} \psi(y) \mu'(dy)$$

(et cette quantité $\in [0, +\infty]$).

On retient que pour des fonctions positives, on peut intervertir l'ordre des intégrations.

Théorème 5.1.5. *Théorème de Fubini (ou Fubini-Lebesgue)*

Soit $f : \Omega \times \Omega' \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$ une fonction mesurable. On définit les fonctions f_1 et f_2 sur Ω et Ω' respectivement par

$$f_1(x) = \int_{\Omega'} |f(x, y)| \mu'(dy), \quad f_2(y) = \int_{\Omega} |f(x, y)| \mu(dx) .$$

(i) Si l'un des fonctions f_1 ou f_2 est intégrable alors l'autre l'est aussi et dans ce cas, f , ϕ et ψ sont intégrables. De plus, nous avons alors

$$\int_{\Omega} \phi(x) \mu(dx) = \int_{\Omega \times \Omega'} f(x, y) \mu \otimes \mu'(dx, dy) = \int_{\Omega'} \psi(y) \mu'(dy) .$$

(ii) Si f est intégrable (contre la mesure $\mu \otimes \mu'$) alors f_1 et f_2 sont intégrables et nous avons encore l'égalité ci-dessus.

Exemple 5.1.6. Soit

$$\begin{aligned} f &: [0, 1] \times [0, 1] \rightarrow \mathbb{R}^+ \\ (x, y) &\mapsto e^{-(x+y)} \mathbf{1}_{x+y \leq 1} . \end{aligned}$$

Cette fonction est mesurable positive. Par Fubini-Tonelli

$$\begin{aligned} \int_{[0,1] \times [0,1]} f(x, y) \lambda \otimes \lambda(dx, dy) &= \int_0^1 \left(\int_0^1 e^{-(x+y)} \mathbf{1}_{x+y \leq 1} dx \right) dy \\ &= \int_0^1 e^{-y} \left(\int_0^1 e^{-x} \mathbf{1}_{x+y \leq 1} dx \right) dy \\ &= \int_0^1 e^{-y} \left(\int_0^{1-y} e^{-x} dx \right) dy \\ &= \int_0^1 e^{-y} (1 - e^{-(1-y)}) dy \\ &= \int_0^1 e^{-y} - e^{-1} dy \\ &= 1 - e^{-1} - e^{-1} = 1 - \frac{2}{e} . \end{aligned}$$

Notation 5.1.7. Intégrale multiple

Pour toute fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}$ intégrable, on notera indifféremment

$$\begin{aligned} \int_{\mathbb{R}^d} f(x_1, \dots, x_d) \lambda^{\otimes d}(dx_1, \dots, dx_d) &= \int_{\mathbb{R}^d} f(x_1, \dots, x_d) dx_1 \dots dx_d \\ &= \int_0^1 \dots \int_0^1 f(x_1, \dots, x_d) dx_1 \dots dx_d \\ &= \int_{\mathbb{R}^d} f(u) du \end{aligned}$$

(on a remplacé, dans cette écriture, (x_1, \dots, x_d) par u).

Définition 5.1.8. Soit μ mesure sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. La mesure μ est dite avoir pour densité la fonction $f : \mathbb{R}^d \rightarrow \mathbb{R}^+$ (par rapport à $\lambda^{\otimes d}$) si $\forall \phi$ mesurable positive $\mathbb{R}^d \rightarrow \mathbb{R}$,

$$\int_{\mathbb{R}} \phi(x) \mu(dx) = \int_{\mathbb{R}} \phi(x) f(x) \lambda(dx) .$$

Ceci implique, en particulier, que $\forall B \in \mathcal{B}(\mathbb{R}^d)$,

$$\mu(B) = \int_B \phi(x) f(x) \lambda(dx) .$$

Exemple 5.1.9. Soit

$$\begin{aligned} f &: \mathbb{R}_+ \times [0, 1] \rightarrow \mathbb{R} \\ (x, y) &\mapsto 2e^{-2xy} - e^{-xy} . \end{aligned}$$

Cette fonction est mesurable et n'est pas de signe constant. Calculons pour tout $y > 0$

$$\begin{aligned} \int_0^{+\infty} f(x, y) dx &= \int_0^{+\infty} 2e^{-2xy} - e^{-xy} dx \\ &= \left[\frac{-e^{-2xy} + e^{-xy}}{y} \right]_0^{+\infty} \\ &= 0 \end{aligned}$$

Nous avons donc pour p.t. $y \in [0, 1]$, $\int_0^{+\infty} f(x, y) dx = 0$ donc, par le théorème 3.0.4,

$$\int_0^1 \left(\int_0^{+\infty} f(x, y) dx \right) dy = 0 .$$

Calculons pour tout $x > 0$

$$\begin{aligned} \int_0^1 f(x, y) dy &= \left[\frac{-e^{-2xy} + e^{-xy}}{x} \right]_0^1 \\ &= \frac{e^{-x} - e^{-2x}}{x} . \end{aligned}$$

Pour $x > 0$, $\frac{e^{-x} - e^{-2x}}{x} > 0$. Nous avons pour p.t. $x \in [0, 1]$, $\int_0^1 f(x, y) dy = \frac{e^{-x} - e^{-2x}}{x} > 0$ donc, par le théorème 3.0.4,

$$\int_0^{+\infty} \left(\int_0^1 f(x, y) dy \right) dx > 0 .$$

Donc

$$\int_0^{+\infty} \left(\int_0^1 f(x, y) dy \right) dx \neq \int_0^1 \left(\int_0^{+\infty} f(x, y) dx \right) dy .$$

Exemple 5.1.10. *Interversion de somme et d'intégrale*

Soit $f : \Omega \times \Omega' \rightarrow \mathbb{R}^+$ mesurable positive. Nous supposons dans cet exemple que $(\Omega', \mathcal{A}', \mu') = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \text{card})$. Comme nous l'avons vu dans l'exemple 4.2.4, pour toute fonction g positive sur Ω' ,

$$\int_{\Omega'} g(x) \mu'(dx) = \sum_{k \geq 0} g(k) .$$

Par Fubini-Tonelli, nous avons alors

$$\int_{\Omega} \left(\sum_{k \geq 0} f(x, k) \right) \mu(dx) = \sum_{k \geq 0} \left(\int_{\Omega} f(x, k) \mu(dx) \right) .$$

5.2 Changement de variable

Définition 5.2.1. Soient U et V deux ouverts de \mathbb{R}^d . Un difféomorphisme ϕ de U dans V est une bijection $\phi : U \rightarrow V$ qui est \mathcal{C}^1 telle que ϕ^{-1} est \mathcal{C}^1 aussi.

Rappel : \mathcal{C}^1 veut dire que la fonction est continue et ses dérivées partielles du premier ordre existent et sont continues. De manière plus explicite, la fonction

$$\begin{aligned} \phi : U &\rightarrow V \\ (u_1, \dots, u_d) &\mapsto (\phi_1(u_1, \dots, u_d), \dots, \phi_d(u_1, \dots, u_d)) \end{aligned}$$

est \mathcal{C}^1 si ϕ_1, \dots, ϕ_d sont continues et $\forall i, j$, $\frac{\partial \phi_i}{\partial u_j}$ existe et est continue.

Définition 5.2.2. Si ϕ est un difféomorphisme de U dans V (deux ouverts de \mathbb{R}^d), on appelle matrice jacobienne la matrice suivante (fonction de (u_1, \dots, u_d))

$$J_{\phi} = \begin{bmatrix} \frac{\partial \phi_1}{\partial u_1}(u_1, \dots, u_d) & \dots & \frac{\partial \phi_d}{\partial u_1}(u_1, \dots, u_d) \\ \frac{\partial \phi_1}{\partial u_2}(u_1, \dots, u_d) & \dots & \frac{\partial \phi_d}{\partial u_2}(u_1, \dots, u_d) \\ \dots & \dots & \dots \\ \frac{\partial \phi_1}{\partial u_d}(u_1, \dots, u_d) & \dots & \frac{\partial \phi_d}{\partial u_d}(u_1, \dots, u_d) \end{bmatrix}$$

Théorème 5.2.3. *Théorème de changement de variable.*

Soient U, V deux ouverts de \mathbb{R}^d . Soit $\phi : U \rightarrow V$ un difféomorphisme. Soit f une fonction intégrable $V \rightarrow \mathbb{R}$. Alors la fonction $f \circ \phi : U \rightarrow \mathbb{R}$ est intégrable et

$$\int_V f(y) dy = \int_U (f \circ \phi)(x) \times |\det(J_{\phi}(x))| dx$$

(Attention à ne pas oublier la valeur absolue dans les calculs.)

Remarque 5.2.4. *Lien avec le changement de variable en dimension 1.*

Soient $]a, b[,]c, d[$ deux intervalles ouverts de \mathbb{R} . Soit $\phi :]a, b[\rightarrow]c, d[$ un difféomorphisme tel que $\lim_{x \rightarrow a} \phi(x) = c$, $\lim_{x \rightarrow b} \phi(x) = d$. Nous connaissons le changement de variable pour les intégrales de Riemann, pour $f :]c, d[\rightarrow \mathbb{R}$

$$\int_c^d f(x) dx = \int_a^b f \circ \phi(y) \phi'(y) dy .$$

Et d'après le théorème précédent,

$$\int_{[c, d]} f(x) dx = \int_{[a, b]} f \circ \phi(y) |\phi'(y)| dy$$

car la matrice jacobienne est ici une matrice 1×1 . Supposons $a \leq b$, $c \geq d$. La fonction ϕ est donc monotone décroissante donc $\forall y, \phi'(y) \leq 0$. D'après la remarque 2.4.14,

$$\int_c^d f(x) dx = - \int_{[c, d]} f(x) dx$$

ce qui est cohérent avec le fait que

$$\int_a^b f \circ \phi(y) \phi'(y) dy = - \int_{[a, b]} f \circ \phi(y) |\phi'(y)| dy .$$

Donc, en dimension 1, on peut faire le changement de variable avec le théorème ci-dessus ou directement sur l'intégrale de Riemann.

Exemple 5.2.5. *Changement de variables en coordonnées polaires.*

Soit

$$\begin{aligned} \phi :]0, +\infty[\times]0, \frac{\pi}{2}[&\rightarrow \mathbb{R}_+^* \times \mathbb{R}_+^* \\ (\rho, \theta) &\rightarrow (\rho \cos(\theta), \rho \sin(\theta)) . \end{aligned}$$

L'application ϕ est un difféomorphisme (on l'admet sans démonstration). Calculons sa matrice jacobienne

$$J_\phi(\rho, \theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\rho \sin(\theta) & \rho \cos(\theta) \end{bmatrix} .$$

Nous avons donc $|\det J_\phi|(\rho, \theta) = |\rho \cos^2(\theta) + \rho \sin^2(\theta)| = |\rho|$. Donc, par le théorème 5.2.3

$$\begin{aligned} \int_0^{+\infty} \int_0^{+\infty} e^{-(x^2+y^2)} dx dy &= \int_0^{+\infty} \int_0^{\frac{\pi}{2}} e^{-\rho^2} |\rho| d\theta d\rho \\ &= \frac{\pi}{2} \int_0^{+\infty} \rho e^{-\rho^2} d\rho \\ &= \frac{\pi}{2} \left[\frac{1}{2} e^{-\rho^2} \right]_0^{+\infty} = \frac{\pi}{4} . \end{aligned}$$

Or

$$\begin{aligned} \int_0^{+\infty} \int_0^{+\infty} e^{-(x^2+y^2)} dx dy &= \int_0^{+\infty} e^{-x^2} \left(\int_0^{+\infty} e^{-y^2} dy \right) dx \\ &= \left(\int_0^{+\infty} e^{-y^2} dy \right) \times \int_0^{+\infty} e^{-x^2} dx \\ &= \left(\int_0^{+\infty} e^{-y^2} dy \right)^2 . \end{aligned}$$

Donc

$$\int_0^{+\infty} e^{-y^2} dy = \frac{\sqrt{\pi}}{2} . \quad (5.2.1)$$

Exemple 5.2.6. *Convolution*

Soient $f, g : \mathbb{R} \rightarrow \mathbb{R}$ deux fonctions intégrables. Rappelons que la convolée de f et g est $(f \star g)(x) = \int_{\mathbb{R}} f(y)g(x-y)dy$. Montrons que cette fonction est bien définie (c'est à dire que $f \star g < \infty$ p.p.). Nous avons

$$\begin{aligned}
 \int_{\mathbb{R}} |(f \star g)(x)| dx &= \int_{\mathbb{R}} \left| \int_{\mathbb{R}} f(y)g(x-y)dy \right| dx \\
 &\leq \int_{\mathbb{R}} \int_{\mathbb{R}} |f(y)| \times |g(x-y)| dy dx \\
 (\text{Fubini-Tonelli}) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(y)| \times |g(x-y)| dx \right) dy \\
 &= \int_{\mathbb{R}} |f(y)| \int_{\mathbb{R}} |g(x-y)| dx dy \\
 &= \int_{\mathbb{R}} |f(y)| \left(\int_{\mathbb{R}} |g(x-y)| dx \right) dy .
 \end{aligned}$$

Pour y fixé, nous avons par changement de variable en dimension 1 ($u = x - y$, $x = u + y$, $dx = du$)

$$\begin{aligned}
 \int_{\mathbb{R}} |g(x-y)| dx &= \int_{-\infty}^{+\infty} |g(x-y)| dx \\
 &= \int_{-\infty}^{+\infty} |g(u)| du \\
 &= \int_{\mathbb{R}} |g(u)| du .
 \end{aligned}$$

Donc

$$\begin{aligned}
 \int_{\mathbb{R}} |(f \star g)(x)| dx &= \int_{\mathbb{R}} f(y) \left(\int_{\mathbb{R}} |g(u)| du \right) dy \\
 &= \left(\int_{\mathbb{R}} |g(u)| du \right) \times \left(\int_{\mathbb{R}} f(y) dy \right) < \infty
 \end{aligned}$$

car f et g sont intégrables. Par la remarque 3.0.2, $|f \star g|$ est donc finie presque partout, donc $f \star g$ est p.p. finie.

Fixons x et opérons un changement de variable $y = x - u$ dans l'intégrale :

$$\begin{aligned}
 \int_{\mathbb{R}} f(y)g(x-y)dy &= \int_{-\infty}^{+\infty} f(y)g(x-y)dy \\
 &= \int_{+\infty}^{-\infty} f(x-u)g(u)(-du) \\
 &= \int_{-\infty}^{+\infty} f(x-u)g(u)du \\
 &= \int_{\mathbb{R}} f(x-u)g(u)du .
 \end{aligned}$$

Donc

$$f \star g = g \star f .$$

Exemple 5.2.7. *Volume de la boule unité*

On note $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ la boule unité de \mathbb{R}^2 . Calculons

$$\begin{aligned}
 \lambda \otimes \lambda(B) &= \int_{\mathbb{R}^2} \mathbf{1}_B(x, y) \lambda \otimes \lambda(dx, dy) \\
 (\text{Fubini-Tonelli}) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbf{1}_{x^2+y^2 \leq 1} dy \right) dx \\
 &= \int_{\mathbb{R}} \mathbf{1}_{|x| \leq 1} \left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1 dy \right) dx \\
 &= \int_{-1}^{+1} 2\sqrt{1-x^2} dx \\
 (\text{changement de variable } x = \sin u) &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} 2 \cos(u) \cos(u) du \\
 &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(2u) + 1 du \\
 &= \left[\frac{\sin(2u)}{2} \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \pi \\
 &= \pi .
 \end{aligned}$$

Exemple 5.2.8. Calculons $I = \int_{[0, +\infty[\times [0, +\infty[} e^{-(x+y)^2 - (x-y)^2} dx dy$. Changement de variables

$$\begin{cases} u = x + y \\ v = x - y \end{cases}, \quad \begin{cases} x = \frac{u+v}{2} \\ y = \frac{u-v}{2} \end{cases}.$$

Le difféomorphisme est $\phi : (u, v) \in \{(u, v) \in \mathbb{R}^2 : u \geq 0, |v| \leq u\} \mapsto \left(\frac{u+v}{2}, \frac{u-v}{2}\right) \in [0, +\infty[\times [0, +\infty[$. Sa matrice jacobienne est

$$J_\phi = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}.$$

Donc

$$\begin{aligned}
 I &= \int_{(u,v) \in \mathbb{R}^2 : u \geq 0, |v| \leq u} e^{-u^2} e^{-v^2} \frac{1}{2} du dv \\
 (\text{Fubini-Tonelli}) &= \int_{u \in [0, +\infty[} \frac{e^{-u^2}}{2} \left(\int_{-u}^u e^{-v^2} dv \right) du.
 \end{aligned}$$

Posons $F(u) = \int_{-u}^u e^{-v^2} dv = 2 \int_0^u e^{-v^2} dv$ (par symétrie). Nous avons $F'(u) = 2e^{-u^2}$. Donc

$$\begin{aligned}
 I &= \int_0^{+\infty} \frac{1}{4} F'(u) F(u) du \\
 &= \left[\frac{1}{8} F(u)^2 \right]_0^{+\infty} \\
 &= \frac{1}{8} \left(\int_{-\infty}^{+\infty} e^{-v^2} dv \right)^2 \\
 (\text{par l'égalité 5.2.1}) &= \frac{\pi}{8}.
 \end{aligned}$$

5.3 Exercices

5.3.1 Énoncés

1) (a) Montrer que pour tout $y > 0$:

$$\int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dx = \frac{\pi}{2} \frac{1}{\sqrt{y}(1+y)}.$$

(b) Montrer que :

$$\int_0^{+\infty} \left(\int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dx \right) dy = \frac{\pi^2}{2} .$$

(c) Montrer que pour tout $x > 0, x \neq 1$:

$$\int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dy = \frac{2 \log(x)}{x^2 - 1} .$$

(d) En déduire que :

$$\int_0^{+\infty} \frac{\log(x)}{x^2 - 1} dx = \frac{\pi^2}{4} .$$

2) On rappelle que : $\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$. En utilisant un changement de variable, calculer :

$$\int_{\mathbb{R} \times \mathbb{R}} e^{-(x+y)^2} e^{-(x-y)^2} dx dy .$$

5.3.2 Corrigés

(1) (a)

$$\begin{aligned} \int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dx &= \frac{1}{(1+y)} \left[\frac{1}{\sqrt{y}} \arctan(x\sqrt{y}) \right]_0^{+\infty} \\ &= \frac{\pi}{2} \frac{1}{\sqrt{y}(1+y)} . \end{aligned}$$

(b)

$$\begin{aligned} \int_0^{+\infty} \int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dx dy &= \int_0^{+\infty} \frac{\pi}{2} \frac{1}{\sqrt{y}(1+y)} dy \\ &= \int_0^{+\infty} \frac{\pi}{2} \frac{1}{1+u^2} 2du \\ &= \pi [\arctan(u)]_0^{+\infty} \\ &= \frac{\pi^2}{2} \end{aligned}$$

où l'on a fait un changement de variable en $u = \sqrt{y}$, $du = \frac{1}{2\sqrt{y}} dy$.

(c) Pour tout $x > 0, x \neq 1$, on a par décomposition en éléments simples :

$$\begin{aligned} \int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dy &= \frac{1}{1-x^2} \int_0^{+\infty} \frac{1}{1+y} - \frac{x^2}{1+x^2y} dy \\ &= \frac{1}{1-x^2} [\log(1+y) - \log(1+x^2y)]_0^{+\infty} \\ &= \frac{1}{1-x^2} \left[\log \left(\frac{1+y}{1+x^2y} \right) \right]_0^{+\infty} \\ &= \frac{1}{1-x^2} \log \left(\frac{1}{x^2} \right) \\ &= \frac{2 \log(x)}{x^2 - 1} . \end{aligned}$$

(d) Par Fubini-Tonelli et puisque $\int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dy = \frac{2 \log(x)}{x^2-1}$ pour p.t. $x \in [0, +\infty[$:

$$\begin{aligned} \int_0^{+\infty} \int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dx dy &= \int_0^{+\infty} \int_0^{+\infty} \frac{1}{(1+y)(1+x^2y)} dy dx \\ \frac{\pi^2}{2} &= \int_0^{+\infty} \frac{2 \log(x)}{x^2-1} dx \\ \frac{\pi^2}{4} &= \int_0^{+\infty} \frac{\log(x)}{x^2-1} dx . \end{aligned}$$

(2) Changement de variable :

$$\begin{cases} u = x + y \\ v = x - y \end{cases} \quad \begin{cases} x = \frac{u+v}{2} \\ y = \frac{u-v}{2} . \end{cases}$$

L'application :

$$\begin{aligned} \phi : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ (u, v) &\mapsto \left(\frac{u+v}{2}, \frac{u-v}{2} \right) \end{aligned}$$

est bijective. On calcule le jacobien (c'est à dire que l'on écrit dans une matrice les dérivées partielles de ϕ en u et v) :

$$J(u, v) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

On fait le changement de variable dans l'intégrale et on utilise Fubini-Tonelli :

$$\begin{aligned} \int_{\mathbb{R} \times \mathbb{R}} e^{-(x+y)^2} e^{-(x-y)^2} dx dy &= \int_{\mathbb{R} \times \mathbb{R}} e^{-u^2} e^{-v^2} |\det(J(u, v))| du dv \\ &= \int_{\mathbb{R} \times \mathbb{R}} e^{-u^2} e^{-v^2} \frac{1}{2} du dv \\ &= \int_{\mathbb{R}} e^{-u^2} \frac{1}{2} \sqrt{\pi} du \\ &= \frac{\pi}{2} . \end{aligned}$$

Chapitre 6

Fondements de la théorie des probabilités

6.1 Définitions générales

Définition 6.1.1. On appelle espace probabilisé un espace mesuré $(\Omega, \mathcal{A}, \mathbb{P})$ où la mesure \mathbb{P} est telle que $\mathbb{P}(\Omega) = 1$. On dit alors que \mathbb{P} est une mesure de probabilité (et c'est pour cela qu'on la note \mathbb{P}). Les éléments de \mathcal{A} sont appelés événements.

Définition 6.1.2. On appelle variable aléatoire toute application mesurable X d'un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ dans un espace mesurable (E, \mathcal{E}) . On dit alors que X est à valeurs dans E .

On notera v.a. pour « variable aléatoire » et v.a.r. pour « variable aléatoire réelle » (variable aléatoire à valeurs dans \mathbb{R}).

Dans toute la suite du chapitre, si on ne précise rien, $(\Omega, \mathcal{A}, \mathbb{P})$ sera un espace probabilisé.

Exemple 6.1.3. Soit $\Omega = \{1, 2, \dots, 6\} \times \{1, 2, \dots, 6\}$ muni de la tribu $\mathcal{P}(\Omega)$ et de la mesure \mathbb{P} telle que $\mathbb{P}((i, j)) = \frac{1}{36}$, $\forall (i, j) \in \Omega$. La mesure \mathbb{P} est une mesure de probabilité car $\text{card}(\Omega) = 36$. L'ensemble Ω est l'ensemble des combinaisons que l'on peut obtenir en jetant un dé deux fois (« ensemble de tous les possibles »). La quantité $\mathbb{P}(3, 2) = 1/36$ est la probabilité d'obtenir 3 puis 2. C'est du moins une modélisation raisonnable de ce qui se passe quand on jette un dé deux fois. Nous pouvons calculer diverses quantités en utilisant la propriété 3 de la définition d'une mesure (déf. 2.2.2) :

- $\mathbb{P}(\{(1, 1), (2, 2)\}) = 2/36 = 1/18$ est la probabilité d'avoir (1 puis 1) ou (2 puis 2)
- $\mathbb{P}(\{(1, 1), (1, 2), \dots, (1, 6)\}) = 6 \times 1/36 = 1/6$ est la probabilité d'avoir 1 au premier tirage .

Introduisons deux variables aléatoires

$$X : (i, j) \in \Omega \mapsto i \in \mathbb{R}, \quad Y : (i, j) \in \Omega \mapsto i + j \in \mathbb{R}.$$

La variable X est le résultat du premier tirage et Y est la somme des deux tirages. Remarquons aussi une variable aléatoire triviale

$$Z : (i, j) \in \Omega \mapsto (i, j) \in \Omega.$$

Définition 6.1.4. Soit $X : \Omega \rightarrow (E, \mathcal{E})$ une variable aléatoire. On appelle loi de X la mesure \mathbb{P}_X sur (E, \mathcal{E}) définie par

$$\begin{aligned} \mathbb{P}_X(A) &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}) \\ &= \mathbb{P}(X^{-1}(A)). \end{aligned}$$

(On rappelle que, par définition, $\{\omega \in \Omega : X(\omega) \in A\} = X^{-1}(A)$.) On notera $\mathbb{P}_X(A) = \mathbb{P}(X \in A)$. C'est un abus de notation très courant.

Remarque 6.1.5. La mesure \mathbb{P}_X est la mesure image de \mathbb{P} par X (cf. prop. 2.4.3). La mesure \mathbb{P}_X est une mesure de probabilité.

Exemple 6.1.6. Reprenons l'exemple précédent. Nous pouvons décrire complètement la loi de Y (toujours à l'aide de la propriété 3 de la définition 2.2.2) :

$$\begin{aligned}\mathbb{P}_Y(\{1\}) &= \mathbb{P}(Y = 1) = 0 \\ \mathbb{P}_Y(\{2\}) &= \mathbb{P}(Y = 2) = \mathbb{P}((i, j) = (1, 1)) = 1/36 \\ \mathbb{P}_Y(\{3\}) &= \mathbb{P}(Y = 3) = \mathbb{P}(\{(1, 2), (2, 1)\}) = 2/36 \\ \mathbb{P}_Y(\{4\}) &= \mathbb{P}(Y = 4) = \mathbb{P}(\{(1, 3), (2, 2), (3, 1)\}) = 3/36 \\ &\dots\end{aligned}$$

Définition 6.1.7. Soit X une v.a. à valeurs dans \mathbb{R} . On appelle fonction de répartition de X la fonction de répartition associée à la mesure \mathbb{P}_X (cf. déf. 2.5) , c'est à dire la fonction

$$\begin{aligned}F &: \mathbb{R} \rightarrow \mathbb{R}_+ \\ t &\mapsto F_X(t) = \mathbb{P}_X(]-\infty, t]) = \mathbb{P}(X \leq t)\end{aligned}$$

Le théorème suivant est une conséquence de la proposition 2.5.2.

Théorème 6.1.8. Soit X une v.a.r.. Alors

- (i) F_X est croissante
- (ii) $\lim_{t \rightarrow -\infty} F_X(t) = 0$, $\lim_{t \rightarrow +\infty} F_X(t) = 1$
- (iii) F_X est càdlàg et $\lim_{t \rightarrow t_0, t < t_0} F_X(t) = \mathbb{P}(X < t_0)$
 F_X est continue en t_0 si, et seulement si, $\mathbb{P}(X = t_0) = 0$.

Définition 6.1.9. Soit X une v.a. à valeurs dans \mathbb{R}^d . On dit que X a une densité $f_X : \mathbb{R}^d \rightarrow \mathbb{R}^+$ si $\forall \phi$ mesurable $\mathbb{R}^d \rightarrow \mathbb{R}$,

$$\mathbb{E}(\phi(X)) = \int_{\mathbb{R}^d} \phi(x) f_X(x) dx .$$

Ceci implique en particulier

$$\mathbb{P}(X \in B) = \int_{\mathbb{R}^d} f_X(x) \mathbf{1}_B(x) dx .$$

La densité de X est la densité de \mathbb{P}_X (cf. les déf. 2.4.8, 5.1.8 de la densité d'une mesure).

Remarque 6.1.10. Si X est une v.a.r. avec une densité f_X alors

$$F_X(t) = \int_{-\infty}^t f_X(u) du .$$

La densité d'un variable aléatoire détermine complètement sa loi.

Par définition, une densité f_X (d'une v.a. X à valeurs dans \mathbb{R}^d) est toujours positive et vérifie $\int_{\mathbb{R}^d} f_X(x) dx = 1$.

Exemple 6.1.11. Soit X une v.a.r. avec la densité $f_X : x \in \mathbb{R} \mapsto e^{-x} \mathbf{1}_{x>0}$.

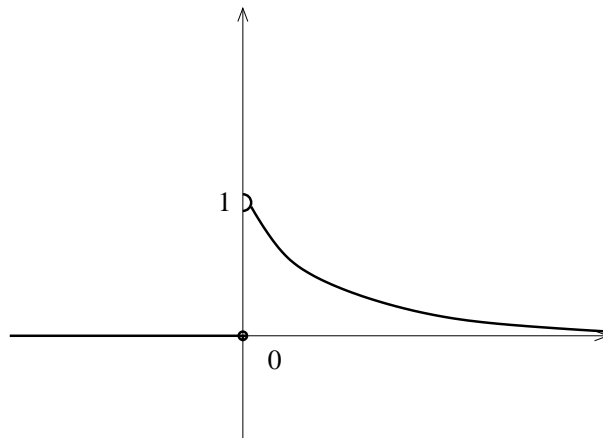


FIG. 6.1 – Dessin de f_X .

Calculons

$$\begin{aligned}
 \mathbb{P}(X \geq 1) &= \int_{\mathbb{R}} e^{-x} \mathbf{1}_{x \geq 0} \mathbf{1}_{x \geq 1} dx \\
 &= \int_1^{+\infty} e^{-x} dx \\
 &= [-e^{-x}]_1^{+\infty} \\
 &= 1 - e^{-1}.
 \end{aligned}$$

Calculons la fonction de répartition de X . Si $t < 0$,

$$\mathbb{P}(X \leq t) = \int_{\mathbb{R}} e^{-x} \mathbf{1}_{x \geq 0} \mathbf{1}_{x \leq t} dx = 0.$$

Si $t \geq 0$,

$$\begin{aligned}
 \mathbb{P}(X \leq t) &= \int_{\mathbb{R}} e^{-x} \mathbf{1}_{x \geq 0} \mathbf{1}_{x \leq t} dx \\
 &= \int_0^t e^{-x} dx \\
 &= 1 - e^{-t}.
 \end{aligned}$$

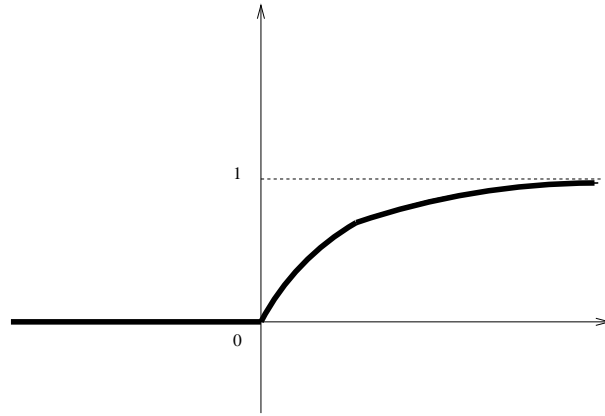


FIG. 6.2 – Dessin de F_X .

Exemple 6.1.12. Soit X v.a.r. de densité $x \mapsto \mathbf{1}_{[0,1]}(x)$.

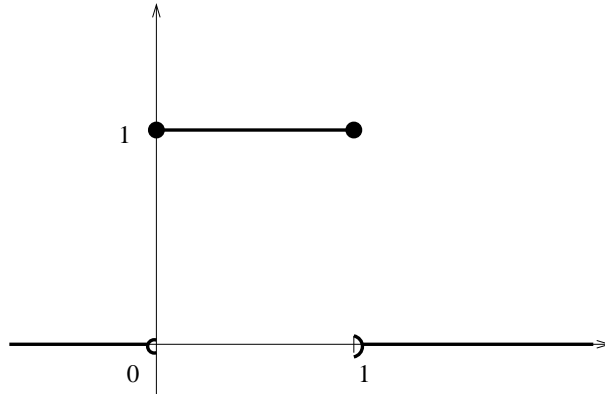


FIG. 6.3 – Dessin de la densité de X .

$$\begin{aligned}
 \text{Si } t < 0, \mathbb{P}(X \leq t) &= \int_{\mathbb{R}} \mathbf{1}_{]-\infty, t]}(x) \mathbf{1}_{[0,1]}(x) dx = 0. \\
 \text{Si } t \geq 1, \mathbb{P}(X \leq t) &= \int_{\mathbb{R}} \mathbf{1}_{]-\infty, t]}(x) \mathbf{1}_{[0,1]}(x) dx = \int_{\mathbb{R}} \mathbf{1}_{[0,1]}(x) dx = 1.
 \end{aligned}$$

Si $t \in [0, 1]$, $\mathbb{P}(X \leq t) = \int_{\mathbb{R}} \mathbf{1}_{]-\infty, t]}(x) \mathbf{1}_{[0, 1]}(x) dx = \int_0^t 1 dx = t$.

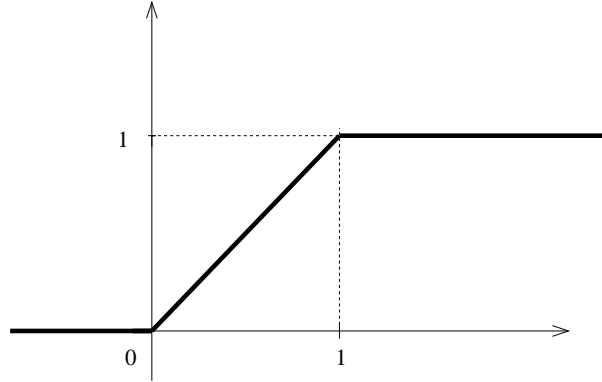


FIG. 6.4 – Dessin de la fonction de répartition de X .

Exemple 6.1.13. Soit X v.a.r. de densité $x \mapsto \mathbf{1}_{[-1, 1]}(x) \sqrt{1 - x^2} \frac{2}{\pi}$.

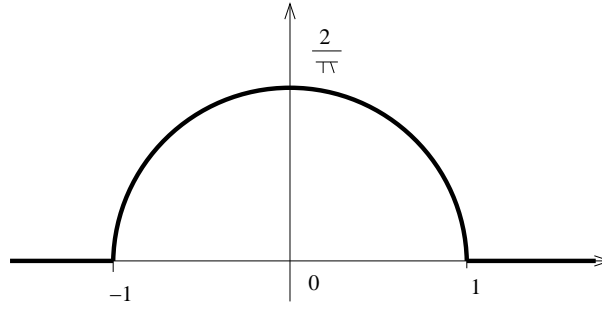
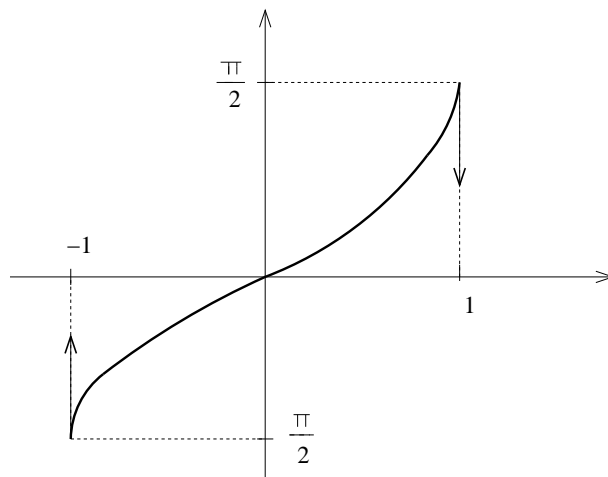
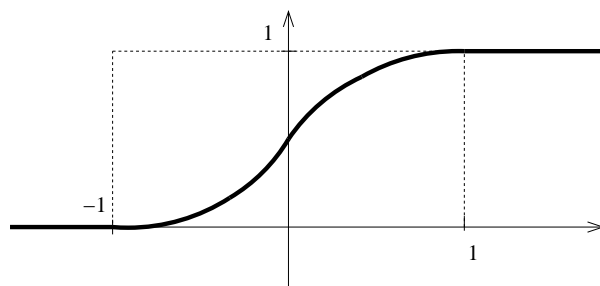


FIG. 6.5 – Dessin de la densité de X .

Si $t \leq -1$, $\mathbb{P}(X \leq t) = 0$. Si $t \in [-1, 1]$,

$$\begin{aligned}
 \mathbb{P}(x \leq t) &= \int_{-\infty}^{+\infty} \mathbf{1}_{]-\infty, t]}(x) \mathbf{1}_{[-1, 1]}(x) \sqrt{1 - x^2} \frac{2}{\pi} dx \\
 &= \int_{-1}^t \sqrt{1 - x^2} \frac{2}{\pi} dx \\
 &= \left[\frac{1}{\pi} \left(x \sqrt{1 - x^2} + \arcsin(x) \right) \right]_{-1}^t \\
 &= \frac{1}{\pi} \left(t \sqrt{1 - t^2} + \arcsin(t) \right) + \frac{1}{2}
 \end{aligned}$$

car sur $[-1, 1]$, $\arcsin'(x) = \frac{1}{\sqrt{1 - x^2}}$ et $\arcsin(-1) = -\frac{\pi}{2}$.

FIG. 6.6 – Dessin de \arcsin . C'est une fonction impaire.FIG. 6.7 – Dessin de la fonction de répartition de X .

6.2 Espérance d'une v.a.

Définition 6.2.1. Soit X v.a.r. On note

$$\mathbb{E}(X) = \int_{\Omega} X(\omega) \mathbb{P}(d\omega)$$

qui est bien définie dans les cas suivants (cf. déf. 2.4.4, 2.4.11)

- $X \geq 0$ (et dans ce cas $\mathbb{E}(X) \in [0, +\infty]$)
- X de signe quelconque et $\int_{\Omega} |X(\omega)| \mathbb{P}(d\omega) < \infty$.

On dit que X est intégrable si $\mathbb{E}(|X|) < \infty$.

Remarque 6.2.2. L'espérance est une intégrale. Réécrivons les propriétés de l'intégrale avec le symbole \mathbb{E} .

- (i) *Linéarité* : si X et Y sont deux v.a.r. et $a, b \in \mathbb{R}$, $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ (cf. th. 2.4.13).
- (ii) *Croissance* : si X et Y sont deux v.a.r. telles que $X \leq Y$ (c'est à dire $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$) alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$ (cf. th. 2.4.13).
- (iii) *Variable aléatoire constante* : si X v.a.r. et $a \in \mathbb{R}$ tels que $X(\omega) = a, \forall \omega$, alors $\mathbb{E}(X) = a$ (cf. déf. 2.3.5).
- (iv) Si X et Y v.a.r. telle que $X = Y$ p.p. alors $\mathbb{E}(X) = \mathbb{E}(Y)$ (cf. prop. 3.0.5).
- (v) Si X variable aléatoire à valeurs dans $[0, +\infty]$ telle que $\mathbb{E}(X) < \infty$ alors X est finie p.s. (cf. rem. 3.0.2).

Proposition 6.2.3. Soit X une variable aléatoire à valeurs dans (E, \mathcal{E}) . Soit f mesurable $E \rightarrow [0, +\infty]$. La fonction $f(X) : \omega \in \Omega \mapsto f(X(\omega)) \in [0, +\infty]$ est une variable aléatoire. Nous avons

$$\mathbb{E}(f(X)) = \int_E f(x) \mathbb{P}_X(dx) .$$

Si $E = \mathbb{R}^d$ et X a une densité g alors

$$\mathbb{E}(f(X)) = \int_{\mathbb{R}^d} f(x)g(x)dx .$$

Définition 6.2.4. Si X est une v.a.r. telle que X^2 est intégrable alors la variance de X est la quantité

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 .$$

Lemme 6.2.5. $\text{Var}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$

Démonstration. Nous allons utiliser les propriétés (i) et (iii) de la remarque 6.2.2.

$$\begin{aligned} \mathbb{E}((X - \mathbb{E}(X))^2) &= \mathbb{E}(X^2 + \mathbb{E}(X)^2 - 2X\mathbb{E}(X)) \\ &= \mathbb{E}(X^2) + \mathbb{E}(\mathbb{E}(X)^2) - 2\mathbb{E}(X\mathbb{E}(X)) \\ &= \mathbb{E}(X^2) + \mathbb{E}(X)^2 - 2\mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{aligned}$$

□

Exemple 6.2.6. Soit X v.a.r. de densité $x \mapsto e^{-x}\mathbf{1}_{x \geq 0}$,

$$\begin{aligned} \mathbb{E}(X) &= \int_{\mathbb{R}} xe^{-x}\mathbf{1}_{x \geq 0}dx \\ &= \int_0^{+\infty} xe^{-x}dx \\ (\text{intégration par parties}) &= [-xe^{-x}]_0^{+\infty} + \int_0^{+\infty} e^{-x}dx \\ &= 0 + [-e^{-x}]_0^{+\infty} = 1 . \end{aligned}$$

Exemple 6.2.7. Soit X v.a. à valeurs dans $\{0, \dots, n\}$ (n un entier fixé) avec $\forall 0 \leq k \leq n$, $\mathbb{P}(X = k) = C_n^k p^k (1-p)^{n-k}$ ($p \in [0, 1]$ fixé). Alors

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^n k \mathbb{P}(X = k) \\ &= \sum_{k=0}^n k C_n^k p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} p^k (1-p)^{n-k} \\ (\text{changement d'indice en } q = k-1) &= \sum_{q=0}^{n-1} C_{n-1}^q p^{q+1} (1-p)^{n-1-q} \\ &= p(p+1-p)^{n-1-q} = p . \end{aligned}$$

Rappel sur le binôme de Newton : $(a+b)^n = \sum_{i=0}^n C_n^i a^i b^{n-i}$.

Exemple 6.2.8. Soit X v.a. à valeurs dans \mathbb{N} avec $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ ($\lambda > 0$ fixé). Alors

$$\begin{aligned} \mathbb{E}(X) &= \sum_{k=0}^{+\infty} \frac{k \lambda^k e^{-\lambda}}{k!} \\ &= \sum_{k=1}^n \frac{1}{(k-1)!} \lambda^k e^{-\lambda} \\ (\text{changement d'indice en } q = k-1) &= \sum_{q=0}^{+\infty} \frac{\lambda^{q+1} e^{-\lambda}}{q!} = \lambda \end{aligned}$$

Proposition 6.2.9. *La loi d'une variable aléatoire X à valeurs dans \mathbb{R}^d est uniquement déterminée par le calcul de $\mathbb{E}(\phi(X))$ pour toute fonction $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continue positive bornée. Autrement dit :*

Soit X variable aléatoire à valeurs dans \mathbb{R}^d . S'il existe $g : \mathbb{R}^d \rightarrow \mathbb{R}$ telle que $\forall \phi : \mathbb{R}^d \rightarrow \mathbb{R}$ continue positive bornée,

$$\mathbb{E}(\phi(X)) = \int_{\mathbb{R}^d} \phi(x) f(x) dx ,$$

alors g est la densité de X .

Notation 6.2.10. On note $\mathcal{C}_b^+(\mathbb{R}^d)$ l'ensemble des fonctions continues positives bornées $\mathbb{R}^d \rightarrow \mathbb{R}^+$.

Exemple 6.2.11. Soit X v.a.r. de densité $x \mapsto \frac{e^{-x^2/2}}{\sqrt{2\pi}}$. Soient $(a, b) \in \mathbb{R}^* \times \mathbb{R}$. Calculons la loi de $aX + b$. Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue et bornée (on dit que f est une « fonction test »). Par la proposition 6.2.3, nous avons

$$\begin{aligned} \mathbb{E}(f(aX + b)) &= \int_{-\infty}^{+\infty} f(ax + b) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ (\text{changement de variable } y = ax + b) &= \int_{-\infty}^{+\infty} f(y) \frac{e^{-(\frac{y-b}{a})^2 \frac{1}{2}}}{\sqrt{2\pi} \times a} dy \end{aligned}$$

Donc, par la proposition 6.2.9, la variable $aX + b$ a une loi de densité $y \mapsto \frac{\exp\left(-\frac{1}{2}\left(\frac{y-b}{a}\right)^2\right)}{\sqrt{2\pi} \times a}$.

Exemple 6.2.12. Soit (X, Y) v.a. à valeurs dans \mathbb{R}^2 de densité $(x, y) \mapsto \frac{1}{\pi} \mathbf{1}_{x^2+y^2 \leq 1}$. Calculons la loi de $X + Y$. Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue. Soit $F : (x, y) \in \mathbb{R}^2 \mapsto f(x + y) \in \mathbb{R}^+$. Alors, par la proposition 6.2.3,

$$\begin{aligned} \mathbb{E}(f(X + Y)) &= \mathbb{E}(F(X, Y)) \\ &= \int_{\mathbb{R}^2} F(x, y) \frac{1}{\pi} \mathbf{1}_{x^2+y^2 \leq 1} dx dy . \end{aligned}$$

Opérons un changement de variable

$$\begin{cases} u &= x + y \\ v &= x - y \end{cases} , \quad \begin{cases} x &= \frac{u+v}{2} \\ y &= \frac{u-v}{2} \end{cases}$$

Difféomorphisme $\phi : (x, y) \in \mathbb{R}^2 \mapsto (x + y, x - y) \in \mathbb{R}^2$. Matrice jacobienne :

$$\begin{bmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{bmatrix} .$$

Donc

$$\begin{aligned} \mathbb{E}(f(X + Y)) &= \int_{\mathbb{R}} f(u) \frac{1}{\pi} \mathbf{1}_{\frac{u^2+v^2}{2} \leq 1} du dv \\ (\text{Fubini-Tonelli}) &= \int_{\mathbb{R}} \frac{f(u)}{\pi} \left(\int_{\mathbb{R}} \mathbf{1}_{u^2+v^2 \leq 2} dv \right) du \\ &= \int_{\mathbb{R}} \frac{f(u)}{\pi} 2\sqrt{2-u^2} \mathbf{1}_{|u| \leq 2} du \end{aligned}$$

Donc $X + Y$ a pour densité $u \mapsto \frac{2}{\pi} \sqrt{2-u^2} \mathbf{1}_{|u| \leq 2}$.

6.3 Inégalités

Théorème 6.3.1. *Inégalité de Jensen*

Soit $\phi : \mathbb{R} \rightarrow \mathbb{R}$ mesurable convexe. Soit X v.a.r. intégrable telle que $\phi(X)$ est intégrable. Alors

$$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X)) .$$

Pour la démonstration de ce théorème, voir l'exercice 4 du chapitre 4.

Théorème 6.3.2. *Inégalité de Bienaymé-Tchebichev*
Soit X v.a.r. positive, intégrable. Soit $\lambda > 0$. Alors

$$\mathbb{P}(X \geq \lambda) \leq \frac{1}{\lambda} \mathbb{E}(X) .$$

Corollaire 6.3.3. *Soit X v.a.r. telle que X^2 est intégrable. Alors*

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) \leq \frac{\text{Var}(X)}{\lambda^2} .$$

Démonstration du théorème 6.3.2. Pour tout ω , $X(\omega) \geq \lambda \mathbf{1}_{X(\omega) \geq \lambda}$ donc, par la propriété de croissance (cf. rem. 6.2.2, (iii)),

$$\begin{aligned} \mathbb{E}(X) &\geq \mathbb{E}(\lambda \mathbf{1}_{X \geq \lambda}) \\ &= \lambda \mathbb{P}(X \geq \lambda) . \end{aligned}$$

□

Démonstration du corollaire 6.3.3.

$$\begin{aligned} \mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda) &= \mathbb{P}((X - \mathbb{E}(X))^2 \geq \lambda^2) \\ (\text{par inégalité de Bienaymé-Tchebichev}) &\leq \frac{1}{\lambda^2} \mathbb{E}((X - \mathbb{E}(X))^2) . \end{aligned}$$

□

Théorème 6.3.4. *Inégalité de Markov*
Si X v.a.r. avec X^2 intégrable et si $\lambda > 0$ alors

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{\mathbb{E}(X^2)}{\lambda^2} .$$

Démonstration.

$$\begin{aligned} \mathbb{P}(|X| \geq \lambda) &= \mathbb{P}(X^2 \geq \lambda^2) \\ (\text{par inégalité de Bienaymé-Tchebichev}) &\leq \frac{\mathbb{E}(X^2)}{\lambda^2} . \end{aligned}$$

□

6.4 Lois classiques

6.4.1 Lois discrètes

- Loi uniforme. Soit E ensemble fini de cardinal n , X est une variable uniforme sur E si $\forall x \in E, \mathbb{P}(X = x) = \frac{1}{n}$.
- Loi de Bernoulli de paramètre p ($\in [0, 1]$), notée $\mathcal{B}(p)$: X à valeurs dans $\{0, 1\}$ telle que $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$.
- Loi binômiale de paramètres n, p ($n \in \mathbb{N}^*, p \in [0, 1]$), notée $\mathcal{B}(n, p)$: X à valeurs dans $\{0, \dots, n\}$ telle que $\forall k \in \{0, \dots, n\}, \mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k}$.
- Loi géométrique de paramètre p ($\in [0, 1]$), notée $\mathcal{G}(p)$: X à valeurs dans \mathbb{N}^* telle que $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = (1 - p)^{k-1} p$.
- Loi de Poisson de paramètre λ (> 0), notée $\mathcal{P}(\lambda)$: X à valeurs dans \mathbb{N} telle que $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

6.4.2 Lois continues

- a) Loi uniforme sur $[a, b]$ ($a < b$), notée $\mathcal{U}([a, b])$: de densité $x \mapsto \frac{1}{b-a} \mathbf{1}_{[a, b]}(x)$.
b) Loi exponentielle de paramètre λ ($\lambda > 0$), notée $\mathcal{E}(\lambda)$: de densité $x \mapsto \lambda e^{-\lambda x} \mathbf{1}_{\mathbb{R}^+}(x)$.
c) Loi gaussienne (ou normale) de moyenne m ($\in \mathbb{R}$) et de variance σ^2 ($\in \mathbb{R}^+$), notée $\mathcal{N}(m, \sigma^2)$: de densité $x \mapsto \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$

6.5 Fonctions caractéristiques

Définition 6.5.1. Soit X v.a.r, la fonction caractéristique de X est

$$\begin{aligned} \Phi_X &: \mathbb{C} \rightarrow \mathbb{C} \\ \xi &\mapsto \int_{\mathbb{R}} e^{i\xi x} \mathbb{P}_X(dx) = \mathbb{E}(e^{i\xi X}) . \end{aligned}$$

Remarque 6.5.2. Pour une fonction $f : \Omega \rightarrow \mathbb{C}$ avec $(\Omega, \mathcal{A}, \mu)$ un espace mesuré quelconque, on note

$$\int_{\Omega} f(x) \mu(dx) = \int_{\Omega} \operatorname{Re}(f)(x) \mu(dx) + i \int_{\Omega} \operatorname{Im}(f)(x) \mu(dx) .$$

et donc dans la définition précédente

$$\Phi_X(\xi) = \int_{\mathbb{R}} \operatorname{Re}(e^{i\xi x}) \mathbb{P}_X(dx) + i \int_{\mathbb{R}} \operatorname{Im}(e^{i\xi x}) \mathbb{P}_X(dx) .$$

Lemme 6.5.3. Soit X de loi $\mathcal{N}(m, \sigma^2)$. Alors $\Phi_X(\xi) = \exp\left(i\xi m - \frac{\sigma^2 \xi^2}{2}\right)$.

Démonstration. Nous ne ferons la démonstration que dans le cas $m = 0, \sigma = 1, \xi \in \mathbb{R}$. Nous avons

$$\begin{aligned} \Phi_X(\xi) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} e^{i\xi x} dx \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \operatorname{Re}(e^{-x^2/2} e^{i\xi x}) dx + i \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} \operatorname{Im}(e^{-x^2/2} e^{i\xi x}) dx \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(x\xi) dx + i \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \sin(x\xi) dx \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(x\xi) dx + 0 \end{aligned}$$

car l'intégrale d'une fonction impaire sur \mathbb{R} est nulle.

Pour tout ξ , $\left| \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(x\xi) \right| \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ qui est intégrable sur \mathbb{R} . Pour tout $x \in \mathbb{R}$, $\xi \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \cos(x\xi)$ est dérivable, de dérivée $\xi \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2} (-x \sin(x\xi))$. Pour tous ξ, x , $\left| \frac{1}{\sqrt{2\pi}} e^{-x^2/2} (-x \sin(x\xi)) \right| \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2} |x|$ qui est intégrable sur \mathbb{R} . En effet, par symétrie,

$$\begin{aligned} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} |x| dx &= 2 \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} x dx \\ &= \left[\frac{1}{\sqrt{2\pi}} e^{-x^2/2} \right]_0^{+\infty} \\ &= \frac{1}{\sqrt{2\pi}} . \end{aligned}$$

Donc par théorème de dérivation globale (cf. cor. 4.3.6)

$$\begin{aligned} \Phi'_X(\xi) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} (-x \sin(x\xi)) dx \\ &= \left[e^{-x^2/2} \sqrt{2\pi} \sin(x\xi) \right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} e^{-x^2/2} \sqrt{2\pi} \xi \cos(x\xi) dx \\ &= 0 - \xi \Phi_X(\xi) . \end{aligned}$$

Nous avons donc l'équation

D'où

$$\begin{aligned}\log(\Phi_X(\xi)) - \log(\Phi_X(0)) &= -\frac{\xi^2}{2} \\ \Phi_X(\xi) &= \Phi_X(0)e^{-\xi^2/2}.\end{aligned}$$

Remarquons que $\Phi_X(0) = \mathbb{E}(1) = 1$. Nous avons donc $\Phi_X(\xi) = e^{-\xi^2/2}$. \square

Théorème 6.5.4. *La fonction caractéristique d'une v.a.r. caractérise entièrement la loi de cette variable. C'est à dire que si X et Y des v.a.r. ont même fonction caractéristique alors X et Y ont même loi.*

6.6 Fonctions génératrices

Définition 6.6.1. *Soit X une v.a. à valeurs dans \mathbb{N} . On appelle fonction génératrice de X la fonction*

$$\begin{aligned}g_X &: [0, 1] \rightarrow \mathbb{R} \\ r &\mapsto \mathbb{E}(r^X) = \sum_{n=0}^{+\infty} \mathbb{P}(X = n)r^n\end{aligned}$$

Proposition 6.6.2. *Si X est une v.a. à valeurs dans \mathbb{N} , la fonction génératrice caractérise la loi de X .*

Exemple 6.6.3. *Soit $X \sim \mathcal{P}(\lambda)$ (ce qui veut dire que X est de loi $\mathcal{P}(\lambda)$). Calculons*

$$\begin{aligned}g_X(u) &= \sum_{n=0}^{+\infty} u^n \frac{\lambda^n e^{-\lambda}}{n!} \\ &= e^{\lambda u} e^{-\lambda} = e^{-\lambda(1-u)}.\end{aligned}$$

6.7 Exercices

6.7.1 Énoncés

- 1) Soit X variable aléatoire réelle de loi de densité $\mathbf{1}_{x \geq 0} \lambda e^{-\lambda x}$, $\lambda > 0$ fixé (loi exponentielle de paramètre λ). Calculer $\mathbb{E}(X)$ et $\text{Var}(X)$. Calculer la densité de la loi de $2X$. Calculer $\mathbb{E}(2X)$, $\text{Var}(2X)$.
- 2) Soit X variable aléatoire réelle de loi de densité $\frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-m)^2}{2\sigma^2}}$, $\sigma, m \in \mathbb{R}$ fixés (loi $\mathcal{N}(m, \sigma^2)$). Soit U variable aléatoire réelle de loi de densité $\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.
 - (a) Montrer que $\sigma U + m$ a même loi que X .
 - (b) Calculer $\mathbb{E}(X)$ et $\text{Var}(X)$.
 - (c) Calculer la densité de la loi de $Y = aX + b$ pour a et b réels.
 - (d) Calculer $\mathbb{E}(Y)$ et $\text{Var}(Y)$.
- 3) Soit X variable aléatoire à valeurs dans \mathbb{N} telle que $\forall k \geq 0, \mathbb{P}(X = k) = \frac{\theta^k e^{-\theta}}{k!}$ ($\theta > 0$ fixé). Calculer $\mathbb{E}(X)$. Pour $u \geq 0$, calculer $\mathbb{E}(e^{-uX})$.
Rappel : $\forall t \in \mathbb{R}, \sum_{n=0}^{+\infty} \frac{t^n}{n!} = e^t$.
- 4) Soit (X, Y) variable aléatoire à valeurs dans \mathbb{R}^2 de loi de densité

$$(x, y) \mapsto \frac{3}{4} \exp(-|x + 2y| - |x - y|).$$

Calculer la densité de la loi de $(X + 2Y, X - Y)$ puis les densités des lois de X et Y . (On pourra utiliser un changement de variable approprié.)

- 5) Soit Y variable aléatoire réelle de densité $\frac{1}{\pi(1+x^2)}$. Montrer que $1/Y$ a même loi que Y .

- 6) M. Dupond attend son bus en moyenne 10 min. tous les matins. Donner une majoration de la probabilité que M. Dupond attende son bus plus de 20 min.
- 7) Soit (X, Y) variable aléatoire à valeurs dans \mathbb{R}^2 densité $\frac{1}{\pi^2} \frac{1}{1+(1+x^2)^2 y^2}$. Calculer la loi de X .
- 8) Soit (X, Y) variable aléatoire à valeurs dans \mathbb{R}^2 de densité $\frac{1}{2\pi} e^{-x^2/2} e^{-y^2/2}$. Calculer la loi de X/Y . Cette variable est-elle intégrable ?
- 9) Soit $\delta > 0$ et Y de loi $\mathcal{N}(0, 1)$.

- (a) Montrer que $\mathbb{P}(Y > \delta) = \frac{1}{\sqrt{2\pi}} \int_{\delta}^{+\infty} \frac{1}{x} \frac{\partial}{\partial x} (-e^{-\frac{x^2}{2}}) dx$. En déduire une intégration par parties de cette intégrale qui donne que $\mathbb{P}(Y > \delta) = \frac{1}{\delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2}} - (\text{intégrale positive})$. En déduire que

$$\mathbb{P}(Y > \delta) \leq \frac{1}{\delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2}}.$$

- (b) On remarque que

$$\mathbb{P}(Y > \delta) = \int_{\delta}^{+\infty} y \left(\frac{1}{y} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \right) dy.$$

Déduire de la question précédente que

$$\mathbb{P}(Y > \delta) \geq \delta \int_{\delta}^{+\infty} F(y) dy$$

avec

$$F(y) = \int_y^{+\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx.$$

- (c) Intégrer par parties $\int_{\delta}^{+\infty} 1 \times F(y) dy$ (en intégrant le 1 et dérivant le F) pour trouver

$$\int_{\delta}^{+\infty} 1 \times F(y) dy = -\delta \int_{\delta}^{+\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx + \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}}.$$

- (d) En déduire que

$$\mathbb{P}(Y > \delta) \geq \frac{1}{(\delta + \frac{1}{\delta})} \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}}.$$

6.7.2 Corrigés

- (1) On fait des intégrations par parties :

$$\begin{aligned} \mathbb{E}(X) &= \int_{\mathbb{R}} \lambda x e^{-\lambda x} \mathbf{1}_{x \geq 0} dx \\ &= \int_0^{+\infty} \lambda x e^{-\lambda x} dx \\ &= [-x e^{-\lambda x}]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx \\ &= 0 + \left[-\frac{1}{\lambda} e^{-\lambda x} \right]_0^{+\infty} = \frac{1}{\lambda} \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx - \frac{1}{\lambda^2} \\ &= [-x^2 e^{-\lambda x}]_0^{+\infty} + \int_0^{+\infty} 2x e^{-\lambda x} dx - \frac{1}{\lambda^2} \\ &= 0 + \left[-2x \frac{e^{-\lambda x}}{\lambda} \right]_0^{+\infty} + \int_0^{+\infty} 2 \frac{e^{-\lambda x}}{\lambda} dx - \frac{1}{\lambda^2} \\ &= 0 + \left[-2 \frac{e^{-\lambda x}}{\lambda^2} \right]_0^{+\infty} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \end{aligned}$$

Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue bornée.

$$\begin{aligned}\mathbb{E}(f(2X)) &= \int_{-\infty}^{+\infty} f(2x) \mathbf{1}_{x \geq 0} \lambda e^{-\lambda x} dx \\ (\text{changement de variable } t = 2x) &= \int_{-\infty}^{+\infty} f(t) \mathbf{1}_{t/2 \geq 0} \lambda e^{-\lambda t/2} \frac{1}{2} dt .\end{aligned}$$

Cela est vrai $\forall f$ donc la densité de la loi de $2X$ est $t \mapsto \mathbf{1}_{t \geq 0} \frac{\lambda}{2} e^{-\lambda t/2}$.

Calculons : $E(2X) = 2E(X) = \frac{2}{\lambda}$ (par linéarité de l'espérance) et $\text{Var}(2X) = \mathbb{E}((2X)^2) - \mathbb{E}(2X)^2 = 4\mathbb{E}(X^2) - 4(\mathbb{E}(X))^2 = 4\text{Var}(X) = \frac{4}{\lambda^2}$ (par linéarité de l'espérance).

(2) (a) Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue bornée.

$$\begin{aligned}\mathbb{E}(f(\sigma U + m)) &= \int_{-\infty}^{+\infty} f(\sigma x + m) \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ (\text{changement de variable } t = \sigma x + m) &= \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-m)^2}{2\sigma^2}} dt\end{aligned}$$

Cela est vrai $\forall f$ donc la densité de la loi de $\sigma U + m$ est la même que celle de la loi de X donc $\sigma U + m$ et X ont même loi

(b) $E(X) = \mathbb{E}(\sigma U + m) = \sigma \mathbb{E}(U) + m = m$ (car $\mathbb{E}(U) = 0$ car intégrale de fonction impaire) et

$$\begin{aligned}\text{Var}(X) &= \text{Var}(\sigma U + m) = \mathbb{E}((\sigma U + m)^2) - \mathbb{E}(\sigma U + m)^2 \\ &= \sigma^2 \mathbb{E}(U^2) + m^2 + 2m\sigma \mathbb{E}(U) - m^2 = \sigma^2 \mathbb{E}(U^2) \\ &= \sigma^2 \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \sigma^2 \left[-x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right]_0^{+\infty} + \sigma^2 \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= 0 + \sigma^2 .\end{aligned}$$

(c) Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue bornée.

$$\begin{aligned}\mathbb{E}(f(Y)) &= \mathbb{E}(f(aX + b)) \\ &= \int_{-\infty}^{+\infty} f(at + b) \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(t-m)^2}{2\sigma^2}} dt \\ (\text{changement de variable } x = at + b) &= \int_{-\infty}^{+\infty} f(x) \frac{1}{\sqrt{2\pi}a^2\sigma^2} e^{-\frac{(x-b-m)^2}{2a^2\sigma^2}} dx\end{aligned}$$

Cela est vrai $\forall f$ donc la densité de la loi de Y est $x \mapsto \frac{1}{\sqrt{2\pi}a^2\sigma^2} e^{-\frac{(x-b-m)^2}{2a^2\sigma^2}}$.

(d) $\mathbb{E}(Y) = \mathbb{E}(aX + b) = a\mathbb{E}(X) + b = am + b$ et

$$\begin{aligned}\text{Var}(Y) &= \text{Var}(aX + b) = \mathbb{E}((aX + b)^2) - \mathbb{E}(aX + b)^2 \\ &= a^2 \mathbb{E}(X^2) + b^2 + 2ab\mathbb{E}(X) - a^2 \mathbb{E}(X)^2 - b^2 - 2ab\mathbb{E}(X) \\ &= a^2 \mathbb{E}(X^2) - a^2 \mathbb{E}(X)^2 = a^2 \text{Var}(X) = a^2 \sigma^2 .\end{aligned}$$

(3)

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k \geq 0} k \frac{\theta^k e^{-\theta}}{k!} \\ &= \sum_{k \geq 1} \theta \frac{\theta^{k-1} e^{-\theta}}{(k-1)!} \\ &= \theta \sum_{q \geq 0} \frac{\theta^q e^{-\theta}}{q!}\end{aligned}$$

(somme de série exponentielle) $= \theta$

$$\begin{aligned}
\mathbb{E}(e^{-uX}) &= \sum_{k \geq 0} e^{-uk} \frac{\theta^k e^{-\theta}}{k!} \\
&= \sum_{k \geq 0} (e^{-u}\theta)^k \frac{e^{-\theta}}{k!} \\
(\text{somme de série exponentielle}) &= \exp(e^{-u}\theta) e^{-\theta} \\
&= \exp(\theta(e^{-u} - 1))
\end{aligned}$$

(4) Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ continue bornée.

$$\mathbb{E}(f(X + 2Y, X - Y)) = \int_{\mathbb{R}^2} f(x + 2y, x - y) \frac{3}{4} \exp(-|x + 2y| - |x - y|) dx dy$$

On fait un changement de variable en :

$$\begin{cases} u = x + 2y \\ v = x - y \end{cases} \quad \begin{cases} x = \frac{u+2v}{3} \\ y = \frac{u-v}{3} \end{cases} .$$

L'application :

$$\begin{aligned}
\phi : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\
(u, v) &\mapsto \left(\frac{u+2v}{3}, \frac{u-v}{3} \right)
\end{aligned}$$

est bijective. On calcule le jacobien (c'est à dire que l'on écrit dans une matrice les dérivées partielles de ϕ en u et v) :

$$J(u, v) = \begin{bmatrix} 1/3 & 1/3 \\ 2/3 & -1/3 \end{bmatrix}$$

On fait le changement de variable dans l'intégrale :

$$\begin{aligned}
\mathbb{E}(f(X + 2Y, X - Y)) &= \int_{\mathbb{R}^2} f(u, v) \frac{3e^{-|u|}e^{-|v|}}{4} |\det(J(u, v))| du dv \\
&= \int_{\mathbb{R}^2} f(u, v) \frac{e^{-|u|}e^{-|v|}}{4} du dv .
\end{aligned}$$

Cela est vrai $\forall f$ donc la densité de la loi de $(X + 2Y, X - Y)$ est $(u, v) \mapsto \frac{e^{-|u|}e^{-|v|}}{4}$. Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue bornée.

$$\begin{aligned}
\mathbb{E}(f(X)) &= \\
&\int_{\mathbb{R}^2} f(x) \frac{3}{4} \exp(-|x + 2y| - |x - y|) dx dy = \\
&(\text{Fubini-Tonelli}) \\
&\frac{3}{4} \int_{\mathbb{R}} f(x) \left(\int_{\mathbb{R}} \exp(-|x + 2y| - |x - y|) dy \right) dx
\end{aligned}$$

On veut calculer $\psi(x) := \int_{\mathbb{R}} \exp(-|x + 2y| - |x - y|) dy$. Commençons par montrer que c'est une fonction paire. On fait un changement de variable en $t = -y$ dans l'intégrale suivante :

$$\begin{aligned}
\psi(-x) &= \int_{-\infty}^{+\infty} \exp(-|-x + 2y| - |-x - y|) dy \\
&= \int_{+\infty}^{-\infty} \exp(-|-x - 2t| - |-x + t|) (-1) dt \\
&= \int_{-\infty}^{+\infty} \exp(-|x + 2t| - |x - t|) dt = \psi(x) .
\end{aligned}$$

On se contente donc de calculer $\psi(x)$ pour $x \geq 0$:

$$\begin{aligned}\psi(x) &= \int_{-\infty}^{-x/2} e^{(x+2y)} e^{-(x-y)} dy \\ &\quad + \int_{-x/2}^x e^{-(x+2y)} e^{-(x-y)} dy + \int_x^{+\infty} e^{-(x+2y)} e^{(x-y)} dy \\ &= \frac{e^{-3x/2}}{3} + e^{-3x/2} - e^{-3x} + \frac{e^{-3x}}{3} \\ &= \frac{4e^{-3x/2}}{3} - \frac{2}{3}e^{-3x}.\end{aligned}$$

Donc par parité, $\psi(x) = \frac{4e^{-3|x|/2}}{3} - \frac{2}{3}e^{-3|x|} \quad \forall x \in \mathbb{R}$. Donc :

$$\mathbb{E}(f(X)) = \frac{3}{4} \int_{\mathbb{R}} f(x) \left(\frac{4e^{-3|x|/2}}{3} + e^{-5|x|/2} - \frac{2}{3}e^{-3|x|} \right) dx.$$

Cela est vrai $\forall f$ donc la densité de la loi de X est $x \mapsto e^{-3|x|/2} - \frac{e^{-3|x|}}{2}$.

Des calculs analogues conduisent à la densité suivante pour $Y : y \mapsto \frac{3}{4}e^{-3|y|}(1+3|y|)$.

(5) Soit $f : \mathbb{R} \rightarrow \mathbb{R}^+$ continue bornée.

$$\begin{aligned}\mathbb{E}(f(1/Y)) &= \int_{-\infty}^{+\infty} f(1/x) \frac{1}{\pi(1+x^2)} dx \\ &= \int_{-\infty}^0 f(1/x) \frac{1}{\pi(1+x^2)} dx + \int_0^{+\infty} f(1/x) \frac{1}{\pi(1+x^2)} dx \\ &\quad \text{(changement de variable en } u = 1/x) \\ &= \int_0^{-\infty} f(u) \frac{1}{\pi(1+1/u^2)} \left(-\frac{1}{u^2} \right) du \\ &\quad \text{(changement de variable en } v = 1/x) \\ &\quad + \int_{+\infty}^0 f(v) \frac{1}{\pi(1+1/v^2)} \left(-\frac{1}{v^2} \right) dv \\ &= \int_{-\infty}^0 f(u) \frac{1}{\pi(1+u^2)} du + \int_0^{+\infty} f(v) \frac{1}{\pi(1+v^2)} dv \\ &= \int_{-\infty}^{+\infty} f(u) \frac{1}{\pi(1+u^2)} du\end{aligned}$$

(Remarque : on est obligé de découper l'intégrale en deux morceaux pour faire des changements de variables bien définis.) On a donc que $u \mapsto \frac{1}{\pi(1+u^2)}$ est la densité de $1/Y$.

(6) On utilise l'inégalité de Bienaymé-Tchebichev : $\mathbb{P}(X \geq 20) \leq \frac{1}{20} \mathbb{E}(X) = \frac{1}{2}$.

(7) Soit $f \in C_b^+(\mathbb{R})$, on calcule :

$$\begin{aligned}\mathbb{E}(f(X)) &= \int_{\mathbb{R}^2} f(x) \frac{1}{\pi^2} \frac{1}{1+(1+x^2)^2 y^2} dx dy \\ \text{par Fubini-Tonelli} &= \int_{-\infty}^{+\infty} f(x) \left(\frac{1}{\pi^2} \int_{-\infty}^{+\infty} \frac{1}{1+(1+x^2)^2 y^2} dy \right) dx.\end{aligned}$$

Donc la densité de X est la fonction de x suivante :

$$\begin{aligned}\frac{1}{\pi^2} \int_{-\infty}^{+\infty} \frac{1}{1+(1+x^2)^2 y^2} dy &= \frac{1}{\pi^2} \left[\frac{1}{1+x^2} \arctan((1+x^2)y) \right]_{-\infty}^{+\infty} \\ &= \frac{1}{\pi} \frac{1}{1+x^2}.\end{aligned}$$

(8) Soit $f \in C_b^+(\mathbb{R})$, on calcule :

$$\mathbb{E}(f(X/Y)) = \int_{\mathbb{R}^2} f(u/v) \frac{1}{2\pi} e^{-u^2/2} e^{-v^2/2} du dv$$

On fait un changement de variable en $s = u/v, t = v, u = st, v = t$. La matrice jacobienne est :

$$J(s, t) = \begin{bmatrix} t & 0 \\ s & 1 \end{bmatrix}$$

de déterminant t . On a donc :

$$\begin{aligned} \mathbb{E}(f(X/Y)) &= \int_{\mathbb{R}^2} \frac{1}{2\pi} f(s) |t| e^{-(st)^2/2} e^{-t^2/2} ds dt \\ \text{par Fubini-Tonelli} &= \int_{-\infty}^{+\infty} f(s) \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-(st)^2/2} e^{-t^2/2} |t| dt \right) ds. \end{aligned}$$

Donc la densité de X/Y est la fonction de s suivante (par parité) :

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-(st)^2/2} e^{-t^2/2} |t| dt &= \frac{1}{\pi} \int_0^{+\infty} e^{-(st)^2/2} e^{-t^2/2} t dt \\ (\text{changement de variable } z = \sqrt{1+s^2} \times t) &= \frac{1}{\pi} \int_0^{+\infty} e^{-z^2/2} z \frac{1}{1+s^2} dz \\ &= \frac{1}{\pi} \left[-e^{-z^2/2} \frac{1}{1+s^2} \right]_0^{+\infty} \\ &= \frac{1}{\pi} \frac{1}{1+s^2}. \end{aligned}$$

On calcule :

$$\begin{aligned} \mathbb{E}(|X/Y|) &= \int_{-\infty}^{+\infty} |s| \frac{1}{\pi} \frac{1}{1+s^2} ds \\ (\text{parité}) &= \int_0^{+\infty} s \frac{2}{\pi} \frac{1}{1+s^2} ds \\ &= +\infty \end{aligned}$$

car $\frac{s}{1+s^2} \underset{s \rightarrow +\infty}{\sim} \frac{1}{s}$ qui n'est pas intégrable en $+\infty$. Donc X/Y n'est pas intégrable.

(9) (a) On a $\frac{\partial}{\partial x}(-e^{-\frac{x^2}{2}}) = xe^{-\frac{x^2}{2}}$. On fait une intégration par parties :

$$\begin{aligned} \mathbb{P}(Y > \delta) &= \frac{1}{\sqrt{2\pi}} \int_{\delta}^{+\infty} \frac{1}{x} \frac{\partial}{\partial x}(-e^{-\frac{x^2}{2}}) dx \\ &= \frac{1}{\sqrt{2\pi}} \left[\frac{1}{x} e^{-\frac{x^2}{2}} \right]_{\delta}^{\infty} - \frac{1}{\sqrt{2\pi}} \int_{\delta}^{+\infty} \frac{1}{x^2} e^{-\frac{x^2}{2}} dx \\ &\leq \frac{1}{\delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{\delta^2}{2}}. \end{aligned}$$

(b) Par la question précédente :

$$\begin{aligned} \mathbb{P}(Y > \delta) &= \int_{\delta}^{\infty} y \left(\frac{1}{y} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \right) dy \\ &\leq \int_{\delta}^{\infty} y \times \mathbb{P}(Y > y) dy \\ &= \delta \int_{\delta}^{+\infty} F(y) dy \end{aligned}$$

(c)

$$\begin{aligned}
\int_{\delta}^{+\infty} 1 \times F(y) dy &= [yF(y)]_{\delta}^{\infty} + \int_{\delta}^{\infty} y \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \\
&= -\delta F(\delta) + \left[-\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} \right]_{\delta}^{\infty} \\
&= -\delta \int_{\delta}^{\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx + \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}}
\end{aligned}$$

(d) D'où

$$\begin{aligned}
\mathbb{P}(Y > \delta) &\geq -\delta^2 \mathbb{P}(Y > \delta) + \delta \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}} \\
\mathbb{P}(Y > \delta) &\geq \frac{\delta}{1 + \delta^2} \frac{e^{-\frac{\delta^2}{2}}}{\sqrt{2\pi}} .
\end{aligned}$$

Chapitre 7

Variables indépendantes

On se donne dans tout le chapitre un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

7.1 Définitions générales

7.1.1 Événements et variables indépendantes

Définition 7.1.1. On dit que $A_1, A_2, \dots \in \mathcal{A}$ sont indépendants si $\forall j_1, \dots, j_p$ (indices distincts) :

$$\mathbb{P}(A_{j_1} \cap \dots \cap A_{j_p}) = \mathbb{P}(A_{j_1}) \times \dots \times \mathbb{P}(A_{j_p}) .$$

On notera $A_1 \perp\!\!\!\perp A_2 \perp\!\!\!\perp \dots$.

On dit que deux événements A_1, A_2 sont indépendants si $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \times \mathbb{P}(A_2)$.

On notera $A_1 \perp\!\!\!\perp A_2$.

Remarque 7.1.2. Pour que les événements ci-dessus soient indépendants, il ne suffit pas qu'ils soient deux à deux indépendants (c'est à dire $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j), \forall i \neq j$).

Lemme 7.1.3. $A_1, A_2, \dots \in \mathcal{A}$ sont indépendants alors A_1^c, A_2^c, \dots sont indépendants.

Démonstration. Nous ne faisons la démonstration que pour deux événements A_1, A_2 . Nous avons (en utilisant les propriétés des mesures)

$$\begin{aligned} \mathbb{P}(A_1^c \cap A_2^c) &= \mathbb{P}((A_1 \cup A_2)^c) \\ &= 1 - \mathbb{P}(A_1 \cup A_2) \\ &= 1 - \mathbb{P}((A_1 \setminus A_2) \cup (A_2 \setminus A_1) \cup (A_1 \cap A_2)) \\ &= 1 - \mathbb{P}(A_1 \setminus A_2) - \mathbb{P}(A_2 \setminus A_1) - \mathbb{P}(A_1 \cap A_2) \\ &= 1 - (\mathbb{P}(A_1) - \mathbb{P}(A_1 \cap A_2)) \\ &\quad - (\mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)) - \mathbb{P}(A_1 \cap A_2) \\ (\text{car } A_1 \text{ indépendant de } A_2) &= 1 - \mathbb{P}(A_1) - \mathbb{P}(A_2) + \mathbb{P}(A_1)\mathbb{P}(A_2) \\ &= (1 - \mathbb{P}(A_1))(1 - \mathbb{P}(A_2)) = \mathbb{P}(A_1^c)\mathbb{P}(A_2^c) \end{aligned}$$

□

Définition 7.1.4. Soient X_1, \dots, X_n des variables aléatoires à valeurs (respectivement) dans des espaces mesurables $(E_1, \mathcal{E}_1), \dots, (E_n, \mathcal{E}_n)$. On dit que X_1, \dots, X_n sont indépendantes si $\forall (F_1, \dots, F_n) \in \mathcal{E}_1 \times \dots \times \mathcal{E}_n$,

$$\mathbb{P}(\{X_1 \in F_1\} \cap \dots \cap \{X_n \in F_n\}) = \mathbb{P}(\{X_1 \in F_1\}) \times \dots \times \mathbb{P}(\{X_n \in F_n\}) .$$

On notera $X_1 \perp\!\!\!\perp \dots \perp\!\!\!\perp X_n$.

Remarque 7.1.5. Pour que X_1, \dots, X_n soient indépendants, il ne suffit pas qu'ils soient deux à deux indépendants.

Théorème 7.1.6. Soient X_1, \dots, X_n des variables indépendantes comme dans la définition ci-dessus. Alors $\forall f_1 : E_1 \rightarrow \mathbb{R}$ mesurable, ..., $\forall f_n : E_n \rightarrow \mathbb{R}$ mesurable :

$$\mathbb{E}(f_1(X_1) \dots f_n(X_n)) = \mathbb{E}(f_1(X_1)) \times \dots \times \mathbb{E}(f_n(X_n)) .$$

Corollaire 7.1.7. Si X, Y sont des v.a.r. indépendantes alors

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) .$$

Démonstration.

$$\begin{aligned} \text{Var}(X + Y) &= \mathbb{E}((X + Y)^2) - (\mathbb{E}(X + Y))^2 \\ &= \mathbb{E}(X^2 + Y^2 + 2XY) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(X)\mathbb{E}(Y) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) \\ &= \mathbb{E}(X^2) + \mathbb{E}(Y^2) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 \\ &= \text{Var}(X) + \text{Var}(Y) . \end{aligned}$$

□

Définition 7.1.8. Soit Y v.a. à valeurs dans un espace mesurable quelconque (E, \mathcal{E}) . La tribu engendrée par Y est $\sigma(Y) = \{Y^{-1}(A), A \in \mathcal{E}\}$. La famille $\sigma(Y)$ est une tribu et $\sigma(Y) \subset \mathcal{A}$.

Proposition 7.1.9. Soient X_1, \dots, X_m des variables indépendantes comme dans la définition 7.1.4. Alors $\forall A_1 \in \sigma(X_1), \dots, A_n \in \sigma(X_n)$, A_1, \dots, A_n sont indépendants. (En d'autres termes, des événements relatifs à des variables indépendantes sont indépendants.)

Et, de plus, $\forall f_1 : E_1 \rightarrow \mathbb{R}$ mesurable, ..., $\forall f_n : E_n \rightarrow \mathbb{R}$ mesurable, les variables $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes.

7.1.2 Densités de variables indépendantes

Théorème 7.1.10. Soient X_1, \dots, X_n des v.a.r.

(i) Si $\forall i$, X_i a la densité p_i et X_1, \dots, X_n indépendantes alors (X_1, \dots, X_n) a la densité

$$(x_1, \dots, x_n) \mapsto p(x_1, \dots, x_n) = p_1(x_1) \times \dots \times p_n(x_n) .$$

(ii) Si X_1, \dots, X_n sont telles que (X_1, \dots, X_n) a une densité de la forme

$$(x_1, \dots, x_n) \mapsto p(x_1, \dots, x_n) = q_1(x_1) \times \dots \times q_n(x_n) ,$$

alors X_1, \dots, X_n sont indépendantes et $\forall i$, X_i a une densité $p_i = C_i q_i$ pour une certaine constante C_i .

Remarque 7.1.11. Quand on se trouve dans le cas (ii) du th. ci-dessus, on détermine les constantes C_i à l'aide de la propriété : $\forall i$, $\int_{\mathbb{R}^d} p_i(x) dx = 1$ (cf. rem 6.1.10). Ce qui donne

$$C_i = \frac{1}{\int_{\mathbb{R}^d} q_i(x) dx} .$$

Exemple 7.1.12. Soit $U \sim \mathcal{E}(1)$ et $V \sim \mathcal{U}([0, 1])$. Les variables U, V sont supposées indépendantes. Soient $X = \sqrt{U} \cos(2\pi V)$, $Y = \sqrt{U} \sin(2\pi V)$. Soit $\phi \in C^+(\mathbb{R}^2)$. Calculons

$$\begin{aligned} \mathbb{E}(\phi(X, Y)) &= \mathbb{E}(\phi(\sqrt{U} \cos(2\pi V), \sqrt{U} \sin(2\pi V))) \\ &= \int_0^{+\infty} \int_0^1 \phi(\sqrt{u} \cos(2\pi v), \sqrt{u} \sin(2\pi v)) e^{-u} du dv . \end{aligned}$$

Changement de variable :

$$\begin{cases} u &= r^2 \\ v &= \frac{2\theta}{\pi} \end{cases} , \quad \begin{cases} r &= \sqrt{u} \\ \theta &= \frac{\pi}{2} v \end{cases} .$$

Difféomorphisme :

$$F : \begin{array}{ccc} [0, +\infty[\times [0, \frac{\pi}{2}[& \rightarrow & [0, +\infty[\times [0, 1] \\ (r, \theta) & \mapsto & (r^2, \frac{2\theta}{\pi}) \end{array} .$$

Matrice jacobienne :

$$\begin{bmatrix} 2r & 0 \\ 0 & \frac{2}{\pi} \end{bmatrix} .$$

Donc

$$\mathbb{E}(\phi(X, Y)) = \int_0^{+\infty} \int_0^{\pi/2} \phi(r \cos(\theta), r \sin(\theta)) |r| e^{-r^2} \frac{4}{\pi} d\theta dr .$$

Puis par changement de variables en coordonnées polaires (comme dans l'exemple 5.2.5) :

$$\mathbb{E}(\phi(X, Y)) = \int_0^{+\infty} \int_0^{+\infty} \phi(x, y) e^{-x^2-y^2} \frac{4}{\pi} dx dy .$$

Donc la densité de (X, Y) est $(x, y) \mapsto \frac{4}{\pi} e^{-x^2} e^{-y^2}$ (par (5.2.1), on peut vérifier que c'est bien une fonction d'intégrale sur \mathbb{R}^2 égale à 1). C'est un produit d'une fonction de x et d'une fonction de y donc X et Y sont indépendantes.

7.2 Lemme de Borel-Cantelli

Théorème 7.2.1. *Lemme de Borel-Cantelli*

(i) Soient A_1, A_2, \dots une famille dénombrable d'événements telle que $\sum_{n \geq 1} \mathbb{P}(A_n) < \infty$. Alors

$$\mathbb{P}(\{\omega : \omega \in \text{une infinité de } A_n\}) = 0 .$$

Ce qui s'énonce aussi : p.s., seul un nombre fini d'événements A_n est réalisé.

(ii) Si on a A_1, A_2, \dots une famille dénombrable d'événements indépendants tels que

$$\sum_{n \geq 1} \mathbb{P}(A_n) = \infty$$

alors

$$\mathbb{P}(\{\omega : \omega \in \text{une infinité de } A_n\}) = 1 .$$

Ce qui s'énonce aussi : p.s., une infinité d'événements A_n est réalisée.

Démonstration. (i) Le symbole \mathbb{E} est une intégrale, nous avons donc, d'après l'exemple 5.1.10 :

$$\begin{aligned} \mathbb{E} \left(\sum_{n \geq 1} \mathbf{1}_{A_n} \right) &= \sum_{n \geq 1} \mathbb{E}(\mathbf{1}_{A_n}) \\ &= \sum_{n \geq 1} \mathbb{P}(A_n) < \infty \end{aligned}$$

donc, par la propriété (v) de la remarque 6.2.2, la variable $Y = \sum_{n \geq 1} \mathbf{1}_{A_n}$ est finie p.s.

(ii) Calculons

$$\begin{aligned} \{\omega : \omega \in \text{infinité de } A_n\} &= \{\omega : \forall n_0, \exists k \geq n_0, \omega \in A_k\} \\ &= \{\omega : \forall n_0, \omega \in \bigcup_{k \geq n_0} A_k\} \\ &= \bigcap_{n_0 \geq 1} \left(\bigcup_{k \geq n_0} A_k \right) . \end{aligned}$$

Soit n_0 fixé, nous avons par indépendance $\forall n \geq n_0$

$$\begin{aligned} \mathbb{P} \left(\bigcap_{n_0 \leq k \leq n} A_k^c \right) &= \prod_{n_0 \leq k \leq n} \mathbb{P}(A_k^c) \\ &= \prod_{n_0 \leq k \leq n} (1 - \mathbb{P}(A_k)) \end{aligned}$$

donc $\log \left(\mathbb{P} \left(\bigcap_{n_0 \leq k \leq n} A_k^c \right) \right) = \sum_{n_0 \leq k \leq n} \log(1 - \mathbb{P}(A_k))$. Nous avons

$$\log(1 - \mathbb{P}(A_k)) \leq -\mathbb{P}(A_k)$$

donc la série précédente diverge. Donc

$$\lim_{n \rightarrow +\infty} \sum_{n_0 \leq k \leq n} \log(1 - \mathbb{P}(A_k)) = -\infty$$

donc

$$\lim_{n \rightarrow +\infty} \prod_{n_0 \leq k \leq n} (1 - \mathbb{P}(A_k)) = \lim_{n \rightarrow +\infty} \mathbb{P} \left(\bigcap_{n_0 \leq k \leq n} A_k^c \right) = -\infty .$$

Pour tout $n \geq n_0$, $\bigcap_{n_0 \leq k \leq n+1} A_k^c \subset \bigcap_{n_0 \leq k \leq n} A_k^c$. Donc par intersection décroissante (cf. prop. 2.2.9)

$$\mathbb{P} \left(\bigcap_{n_0 \leq k} A_k^c \right) = \lim_{n \rightarrow +\infty} \mathbb{P} \left(\bigcap_{n_0 \leq k \leq n} A_k^c \right) = 0 .$$

Et donc par réunion,

$$\mathbb{P} \left(\bigcup_{n_0 \geq 1} \bigcap_{n_0 \leq k} A_k^c \right) \leq \sum_{n_0 \geq 1} \mathbb{P} \left(\bigcap_{n_0 \leq k} A_k^c \right) = 0 .$$

Donc par passage au complémentaire

$$\mathbb{P} \left(\bigcap_{n_0 \geq 1} \bigcup_{n_0 \leq k} A_k \right) = 0$$

□

7.3 Somme de deux variables indépendantes

Définition 7.3.1. *Convolution de deux mesures*

Si μ et ν sont deux mesures sur \mathbb{R}^d , on définit $\mu \star \nu$ (la convolée de μ et ν) par la relation suivante : $\forall \phi \in \mathcal{C}_b^+(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \phi(z) \mu \star \nu(dz) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x+y) \mu(dx) \nu(dy) .$$

(Cette relation détermine complètement $\mu \star \nu$.)

Remarque 7.3.2. Par un changement de variable, on montre que $\mu \star \nu = \nu \star \mu$.

Lemme 7.3.3. Si μ et ν sont deux mesures de probabilité sur \mathbb{R}^d de densités respectivement f et g alors $\mu \star \nu$ est une mesure de probabilité de densité $f \star g$ (cf. ex. 4.3.4 et 5.2.6 pour la définition de la convolée de deux fonctions).

Démonstration. Soit $\phi \in \mathcal{C}_b^+(\mathbb{R}^d)$,

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(z) \mu \star \nu(dz) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x+y) \mu(dx) \nu(dy) \\ &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(x+y) f(x) g(y) \mu(dx) \nu(dy) . \end{aligned}$$

Changement de variable $\mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $u = x + y, v = y, x = u - v, y = v$. Matrice jacobienne :

$$\begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix} .$$

Donc

$$\begin{aligned} \int_{\mathbb{R}^d} \phi(z) \mu \star \nu(dz) &= \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \phi(u) f(u-v) g(v) du dv \\ (\text{Borel-Cantelli}) &= \int_{\mathbb{R}^d} \phi(u) \left(\int_{\mathbb{R}^d} f(u-v) g(v) dv \right) du \\ &= \int_{\mathbb{R}^d} \phi(u) f \star g(u) du . \end{aligned}$$

Donc $f \star g$ est la densité de $\mu \star \nu$. □

Proposition 7.3.4. Soient X et Y deux variables indépendantes à valeurs dans \mathbb{R}^d .

- i) La loi de $X + Y$ est $\mathbb{P}_X \star \mathbb{P}_Y$. Si, de plus, X, Y ont des densités respectivement p_X, p_Y , alors $X + Y$ a pour densité $p_X \star p_Y$.
- ii) La fonction caractéristique de $X + Y$ est $\Phi_{X+Y} = \Phi_X \times \Phi_Y$.
- iii) Si X, Y à valeurs dans \mathbb{N} , la fonction génératrice de $X + Y$ est $g_{X+Y} = g_X \times g_Y$.

Démonstration. (i) vient du lemme précédent.

(ii)

$$\begin{aligned} \Phi_{X+Y}(\xi) &= \mathbb{E}(e^{i\xi(X+Y)}) \\ &= \mathbb{E}(e^{i\xi X} e^{i\xi Y}) \\ (X \perp Y, \text{ cf. cor. 7.1.6}) &= \mathbb{E}(e^{i\xi X}) \mathbb{E}(e^{i\xi Y}) \\ &= \Phi_X(\xi) \Phi_Y(\xi) . \end{aligned}$$

(iii) De même

$$\begin{aligned} g_{X+Y}(t) &= \mathbb{E}(t^{X+Y}) \\ &= \mathbb{E}(t^X t^Y) \\ &= \mathbb{E}(t^X) \mathbb{E}(t^Y) = g_X(t) g_Y(t) . \end{aligned}$$

□

Exemple 7.3.5. Somme de gaussiennes

Soient $X \sim \mathcal{N}(m_1, \sigma_1^2)$, $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ indépendantes. Nous avons (cf. lem. 6.5.3)

$$\Phi_X(\xi) = \exp\left(i\xi m_1 - \frac{\xi^2 \sigma_1^2}{2}\right), \Phi_Y(\xi) = \exp\left(i\xi m_2 - \frac{\xi^2 \sigma_2^2}{2}\right) .$$

Donc, par la proposition précédente,

$$\Phi_{X+Y}(\xi) = \exp\left(i\xi(m_1 + m_2) - \frac{\xi^2(\sigma_1^2 + \sigma_2^2)}{2}\right)$$

Et donc (cf. th. 6.5.4)

$$X + Y \sim \mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2) .$$

Exemple 7.3.6. Si X et Y de loi $\mathcal{G}(p)$ indépendantes alors

$$\begin{aligned}
 \mathbb{P}(X + Y = n) &= \mathbb{P}(\cup_{0 \leq k \leq n} \{X = k, Y = n - k\}) \\
 (\text{car év. disjoints}) &= \sum_{k=0}^n \mathbb{P}(X = k, Y = n - k) \\
 (\text{car } X \perp Y) &= \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = n - k) \\
 &= \sum_{k=0}^n p^k (1-p) p^{n-k} (1-p) \\
 &= (n+1) p^n (1-p)^2.
 \end{aligned}$$

7.4 Exercices

7.4.1 Énoncés

- 1) Soient X et Y deux variables aléatoires indépendantes de loi $\mathcal{N}(0, 1)$. Montrer que $X + Y$ et $X - Y$ sont indépendantes.
- 2) Soient X et Y deux variables aléatoires réelles indépendantes. On suppose que X suit une loi de Poisson de paramètre λ et que Y suit une loi de Poisson de paramètre μ . Calculer la loi de $X + Y$.
- 3) X une variable aléatoire dans \mathbb{R} est dite symétrique si $-X$ a même loi que X .
 - (a) Si X a une densité f , montrer que : X est symétrique si et seulement si $f(x) = f(-x)$ pour presque tout x .
 - (b) Donner un exemple de loi symétrique.
 - (c) Montrer que X est symétrique si et seulement si le nombre $\mathbb{E}(e^{iuX})$ est réel $\forall u \in \mathbb{R}$.
 - (d) Soit X variable aléatoire dans \mathbb{R} symétrique. On suppose $\mathbb{P}(X = 0) = 0$. On note :

$$\varepsilon = \begin{cases} 1 & \text{si } X > 0 \\ 0 & \text{si } X = 0 \\ -1 & \text{si } X < 0 \end{cases}.$$

Montrer que ε et $|X|$ sont indépendantes.

- (e) Si Y et Y' sont deux variables aléatoires réelles de même loi et indépendantes, montrer que $Y - Y'$ est symétrique.
- 4) Soient U de loi uniforme sur $[0, 1]$ et X de loi exponentielle de paramètre 1 deux variables aléatoires réelles indépendantes.
 - (a) Calculer $\mathbb{P}(\sup(U, X) \leq t)$ dans les 3 cas suivants : $t < 0$, $t \in [0, 1]$, $t > 1$.
 - (b) Dessiner la fonction de répartition de $\sup(U, X)$.

7.4.2 Corrigés

- (1) Soit $f \in C_b^+(\mathbb{R}^2)$, on calcule :

$$\begin{aligned}
 \mathbb{E}(f(X + Y, X - Y)) &= \int_{\mathbb{R}^2} f(x + y, x - y) \frac{1}{2\pi} e^{-x^2/2} e^{-y^2/2} dx dy \\
 &\quad (\text{changement de variable déjà vu : } u = x + y, v = x - y) \\
 &= \int_{\mathbb{R}^2} \frac{1}{2\pi} f(u, v) e^{-(u+v)^2/8} e^{-(u-v)^2/8} \frac{1}{2} du dv \\
 &= \int_{\mathbb{R}^2} f(u, v) e^{-u^2/4} e^{-v^2/4} \frac{1}{4\pi} du dv
 \end{aligned}$$

Donc la densité de $(X + Y, X - Y)$ est la fonction $(u, v) \mapsto e^{-u^2/4} e^{-v^2/4} \frac{1}{4\pi}$. C'est un produit d'une fonction de u et d'une fonction de v donc $X + Y$ et $X - Y$ sont indépendantes.

(2) Les variables X et Y sont à valeurs dans \mathbb{N} donc $X + Y$ aussi. Soit $n \in \mathbb{N}$, calculons :

$$\begin{aligned}\mathbb{P}(X + Y = n) &= \mathbb{P}(\{X = 0 \text{ et } Y = n\} \cup \{X = 1 \text{ et } Y = n - 1\} \cup \dots \\ &\quad \dots \cup \{X = n \text{ et } Y = 0\}) \\ \text{événements disjoints} &= \sum_{k=0}^n \mathbb{P}(X = k \text{ et } Y = n - k) \\ \text{indépendance} &= \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = n - k) \\ &= \sum_{k=0}^n \frac{\lambda^k e^{-\lambda}}{k!} \frac{\mu^{n-k} e^{-\mu}}{(n-k)!} \\ &= \frac{e^{-\lambda-\mu}}{n!} \sum_{k=0}^n C_n^k \lambda^k \mu^{n-k} \\ &= \frac{e^{-\lambda-\mu}}{n!} (\lambda + \mu)^n.\end{aligned}$$

Donc $X + Y \sim \mathcal{P}(\lambda + \mu)$.

(3) (a) – Si X est symétrique :

$$\forall \phi \in C_b^+(\mathbb{R}),$$

$$\begin{aligned}\mathbb{E}(\phi(X)) &= \mathbb{E}(\phi(-X)) \\ \int_{-\infty}^{+\infty} \phi(t) f(t) dt &= \int_{-\infty}^{+\infty} \phi(-t) f(t) dt \\ \int_{-\infty}^{+\infty} \phi(t) f(t) dt &= \int_{-\infty}^{+\infty} \phi(u) f(-u) du \text{ (changement de variable } u = -t) \text{ .}\end{aligned}$$

Donc $\int_{-\infty}^{+\infty} \phi(t)(f(t) - f(-t)) dt = 0$. Cela est vrai $\forall \phi \in C_b^+(\mathbb{R})$ donc $f(t) - f(-t)$ est nulle presque partout donc $f(t) = f(-t)$ pour presque tout t .

– Si $f(t) = f(-t)$ pour presque tout t :

$$\forall \phi \in C_b^+(\mathbb{R}),$$

$$\begin{aligned}\mathbb{E}(\phi(-X)) &= \int_{-\infty}^{+\infty} \phi(-t) f(t) dt \\ &= \int_{-\infty}^{+\infty} \phi(t) f(-t) dt \text{ (par changement de variable)} \\ &= \int_{-\infty}^{+\infty} \phi(t) f(t) dt \\ &\quad \text{(car } f(t) \text{ et } f(-t) \text{ coïncident presque partout)}\end{aligned}$$

donc $-X$ est de densité $t \mapsto f(t)$ comme X donc X est symétrique.

(b) Exemple de loi symétrique : $X = 1$ avec probabilité $1/2$ et $X = -1$ avec probabilité $1/2$.

(c) – Si X est symétrique :

Pour tout u :

$$\begin{aligned}\mathbb{E}(e^{iuX}) &= \mathbb{E}(e^{-iuX}) \\ &= \overline{\mathbb{E}(e^{iuX})}.\end{aligned}$$

Donc $\mathbb{E}(e^{iuX}) \in \mathbb{R}$.

– Si $\mathbb{E}(e^{iuX}) \in \mathbb{R}, \forall u$:

$$\begin{aligned}\mathbb{E}(e^{iu(-X)}) &= \overline{\mathbb{E}(e^{iuX})} \\ &= \mathbb{E}(e^{iuX}).\end{aligned}$$

Donc X et $-X$ ont même fonction caractéristique donc X et $-X$ ont même loi donc X est symétrique.

(d) Soient $A_1, A_2 \in \mathcal{B}(\mathbb{R})$.

- Si $1 \in A_1$ et $-1 \in A_1$, $\mathbb{P}(\varepsilon \in A_1, |X| \in A_2) = \mathbb{P}(|X| \in A_2) = \mathbb{P}(\varepsilon \in A_1)\mathbb{P}(|X| \in A_2)$.
- Si $1 \in A_1$ et $-1 \notin A_1$, $\mathbb{P}(\varepsilon \in A_1, |X| \in A_2) = \mathbb{P}(\varepsilon = 1, |X| \in A_2) = \mathbb{P}(X > 0, X \in A_2) = \mathbb{P}(X < 0, -X \in A_2)$ car X symétrique donc $\mathbb{P}(\varepsilon \in A_1, |X| \in A_2) = \frac{1}{2}(\mathbb{P}(X > 0, X \in A_2) + \mathbb{P}(X < 0, -X \in A_2)) = \mathbb{P}(\varepsilon \in A_1)\mathbb{P}(|X| \in A_2)$.
- Si $1 \notin A_1$ et $-1 \in A_1$, on montre de même que $\mathbb{P}(\varepsilon \in A_1, |X| \in A_2) = \mathbb{P}(\varepsilon \in A_1)\mathbb{P}(|X| \in A_2)$.
- Si $1 \notin A_1$ et $-1 \notin A_1$, $\mathbb{P}(\varepsilon \in A_1, |X| \in A_2) = 0 = \mathbb{P}(\varepsilon \in A_1)\mathbb{P}(|X| \in A_2)$.

On a donc toujours $\mathbb{P}(\varepsilon \in A_1, |X| \in A_2) = \mathbb{P}(\varepsilon \in A_1)\mathbb{P}(|X| \in A_2)$, donc ε et $|X|$ sont indépendants.

(e) On calcule la fonction caractéristique ;

$$\begin{aligned}
 \mathbb{E}(e^{iu(Y-Y')}) &= \mathbb{E}(e^{iuY} e^{-iuY'}) \\
 &= \mathbb{E}(e^{iuY})\mathbb{E}(e^{-iuY'}) \text{ (par indépendance)} \\
 &= \mathbb{E}(e^{iuY'})\mathbb{E}(e^{-iuY}) \text{ (car } Y \text{ et } Y' \text{ ont même loi)} \\
 &= \mathbb{E}(e^{iu(Y'-Y)}) \\
 &= \overline{\mathbb{E}(e^{iu(Y-Y')})} .
 \end{aligned}$$

Donc par la question 3c, $Y - Y'$ est symétrique.

(4) (a)

$$\begin{aligned}
 F(t) = \mathbb{P}(\sup(U, X) \leq t) &= \mathbb{P}(U \leq t, X \leq t) \\
 \text{(indépendance)} &= \mathbb{P}(U \leq t)\mathbb{P}(X \leq t) \\
 &= \begin{cases} 0 & \text{si } t \leq 0 \\ t(1 - e^{-t}) & \text{si } t \in]0, 1[\\ 1 - e^{-t} & \text{si } t > 1 \end{cases}
 \end{aligned}$$

(b) On remarque que F est continue et qu'elle a un point anguleux en 1.

Chapitre 8

Convergence de variables aléatoires

On se donne dans tout le chapitre un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

8.1 Les différentes notions de convergence

On se donne $X, (X_n)_{n \geq 0}$ v.a. à valeurs dans \mathbb{R}^d .

Définition 8.1.1. (C'est une réécriture de la définition 4.1.2.)

On dit que X_n converge presque sûrement vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X$ si

$$\mathbb{P}(\{\omega \in \Omega : X(\omega) = \lim_{n \rightarrow +\infty} X_n(\omega)\}) = 1 .$$

Définition 8.1.2. On dit que X_n converge dans L^p vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{L^p} X$ si $\mathbb{E}(\|X - X_n\|^p) \xrightarrow[n \rightarrow +\infty]{} 0$ (ici, $\|\cdot\|$ est la norme usuelle sur \mathbb{R}^d).

Définition 8.1.3. On dit que X_n converge en probabilité vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{proba.} X$ si $\forall \varepsilon > 0, \mathbb{P}(\|X - X_n\| \geq \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.

Définition 8.1.4. On dit que X_n converge en loi vers X et on note $X_n \xrightarrow[n \rightarrow +\infty]{loi} X$ si $F_{X_n}(t) \xrightarrow[n \rightarrow +\infty]{loi} F_X(t)$ en tout point t où F_X est continue (c'est à dire en tout point t tel que $\mathbb{P}(X = t) = 0$).

Définition 8.1.5. Soit (μ_n) une suite de mesures de probabilité sur \mathbb{R}^d . On dit que (μ_n) converge étroitement vers μ et on note $\mu \xrightarrow[n \rightarrow +\infty]{étr.} \mu$ si $\forall \phi \in \mathcal{C}_b^+(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \phi(x) \mu_n(dx) \xrightarrow[n \rightarrow +\infty]{} \int_{\mathbb{R}^d} \phi(x) \mu(dx) .$$

Théorème 8.1.6. Nous avons l'équivalence

$$\left[X_n \xrightarrow[n \rightarrow +\infty]{loi} X \right] \Leftrightarrow \left[\forall f \in \mathcal{C}_b^+(\mathbb{R}^d), \mathbb{E}(f(X_n)) \xrightarrow[n \rightarrow +\infty]{} \mathbb{E}(f(X)) \right] .$$

Corollaire 8.1.7. Pour une suite de v.a. à valeurs dans \mathbb{R}^d ,

$$\left[X_n \xrightarrow[n \rightarrow +\infty]{loi} X \right] \Leftrightarrow \left[\mathbb{P}_{X_n} \xrightarrow[n \rightarrow +\infty]{étr.} \mathbb{P}_X \right]$$

Théorème 8.1.8. i) $\left[X_n \xrightarrow[n \rightarrow +\infty]{p.s.} X \right] \Rightarrow \left[X_n \xrightarrow[n \rightarrow +\infty]{proba.} X \right]$

$$\begin{aligned}
ii) \quad & \forall p \geq 1, \left[X_n \xrightarrow[n \rightarrow +\infty]{L^p} X \right] \Rightarrow \left[X_n \xrightarrow[n \rightarrow +\infty]{\text{proba.}} X \right] \\
iii) \quad & \left[X_n \xrightarrow[n \rightarrow +\infty]{\text{proba.}} X \right] \Rightarrow \left[\exists \text{ sous-suite } (X_{g(n)}) : X_{g(n)} \xrightarrow[n \rightarrow +\infty]{p.s.} X \right] \\
iv) \quad & \left[X_n \xrightarrow[n \rightarrow +\infty]{\text{proba.}} X \right] \Rightarrow \left[X_n \xrightarrow[n \rightarrow +\infty]{\text{loi}} X \right]
\end{aligned}$$

Rappel : une sous-suite d'une suite $(u_n)_{n \geq 0}$ est donnée par une application strictement croissante $g : \mathbb{N} \rightarrow \mathbb{N}$, la sous-suite s'écrit alors $(u_{g(n)})_{n \geq 0}$.

Diagramme :

$$\begin{array}{ccccc}
\text{convergence } L^p & \Rightarrow & \text{convergence en probabilité} & \Leftarrow & \text{convergence p.s.} \\
& & \Downarrow & & \\
& & \text{convergence en loi.} & &
\end{array}$$

Toutes les autres implications sont fausses.

8.2 Loi des grands nombres

Notation 8.2.1. Soient X_1, X_2, \dots des variables indépendantes et de même loi. On dira que ces variables sont indépendantes et identiquement distribuées et on utilisera la notation « i.i.d. ».

Théorème 8.2.2. Loi faible des grands nombres

Soient X_1, X_2, \dots des v.a.r. i.i.d. Si $\mathbb{E}(X_n^2) < \infty$, on a

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{L^2} \mathbb{E}(X_1) .$$

Démonstration.

$$\begin{aligned}
\mathbb{E} \left(\left(\frac{X_1 + \dots + X_n}{n} - \mathbb{E}(X_1) \right)^2 \right) &= \mathbb{E} \left(\left(\frac{(X_1 - \mathbb{E}(X_1)) + \dots + (X_n - \mathbb{E}(X_n))}{n} \right)^2 \right) \\
&= \sum_{k=1}^n \frac{1}{n^2} \mathbb{E}((X_k - \mathbb{E}(X_k))^2) \\
&= \frac{\text{Var}(X_1)}{n} \xrightarrow[n \rightarrow +\infty]{} 0
\end{aligned}$$

□

Théorème 8.2.3. Loi forte des grands nombres

Soient X_1, X_2, \dots des v.a.r. i.i.d. Si $\mathbb{E}(|X_1|) < \infty$ (en d'autres termes. Si X_1 est intégrable) alors

$$\frac{X_1 + \dots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1) .$$

Démonstration. Nous ne ferons la démonstration que dans le cas $\mathbb{E}(X_1^4) < \infty$. Nous voulons montrer que

$$\frac{(X_1 - \mathbb{E}(X_1)) + \dots + (X_n - \mathbb{E}(X_n))}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} 0 .$$

Posons pour tout i , $X'_i = X_i - \mathbb{E}(X_i)$. Calculons

$$\mathbb{E} \left(\left(\frac{X'_1 + \dots + X'_n}{n} \right)^4 \right) = \frac{1}{n^4} \sum_{i_1, i_2, i_3, i_4 \in \{1, \dots, n\}} \mathbb{E}(X'_{i_1} X'_{i_2} X'_{i_3} X'_{i_4}) .$$

Remarquons que dans cette dernière somme, certains termes sont nuls. Par exemple, en utilisant les propriétés des variables indépendantes (cf. cor. 7.1.6)

$$\begin{aligned}
\mathbb{E}(X'_1 X'_2 X'_2 X'_2) &= \mathbb{E}(X'_1) \mathbb{E}((X'_2)^3) = 0 \\
\mathbb{E}(X'_1 X'_2 X'_3 X'_3) &= \mathbb{E}(X'_1) \mathbb{E}(X'_2) \mathbb{E}((X'_3)^2) = 0 .
\end{aligned}$$

Après regroupement des termes identiques, nous obtenons

$$\begin{aligned}\mathbb{E}\left(\left(\frac{X'_1 + \cdots + X'_n}{n}\right)^4\right) &= \frac{1}{n^4}(n\mathbb{E}((X'_1)^4) + 6n(n-1)\mathbb{E}((X_1)^2(X_2)^2)) \\ &\leq \frac{7}{n^2}.\end{aligned}$$

Et donc $\sum_{n \geq 1} \mathbb{E}\left(\left(\frac{X'_1 + \cdots + X'_n}{n}\right)^4\right) < \infty$. Par Fubini-Tonelli (cf. ex. 5.1.10)

$$\mathbb{E}\left(\sum_{n \geq 1} \left(\frac{X'_1 + \cdots + X'_n}{n}\right)^4\right) = \sum_{n \geq 1} \mathbb{E}\left(\left(\frac{X'_1 + \cdots + X'_n}{n}\right)^4\right) < \infty.$$

Donc la variable $\sum_{n \geq 1} \left(\frac{X'_1 + \cdots + X'_n}{n}\right)^4$ est finie p.s. (cf. rem. 6.2.2, (v)). Donc le terme général de la série converge vers 0, p.s. \square

Exemple 8.2.4. Soient U_1, U_2, \dots i.i.d. de loi $\mathcal{U}([0, 1])$. Soient $0 \leq a < b \leq 1$. Soit pour tout i , $X_i = \mathbf{1}_{[a, b]}(U_i)$. Les variables X_1, X_2, \dots sont i.i.d. de loi $\mathcal{G}(b-a)$ et vérifient $\mathbb{E}(|X_i|) < \infty$ puisqu'elles sont bornées. Par la loi des grands nombres

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1) = \mathbb{P}(a \leq U_1 \leq b) = b - a.$$

Ce qui veut dire que la proportion de points tombant dans $[a, b]$ converge vers $b - a$. Illustration, la densité empirique de $\mathcal{U}([0, 1])$:

<http://www-sop.inria.fr/mefisto/java/tutorial1/node7.html>

#SECTION00031010000000000000

De même,

$$\frac{\mathbf{1}_{X_1 \leq 1/2} + \cdots + \mathbf{1}_{X_n \leq 1/2}}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \frac{1}{2}.$$

Illustration, le jeu de pile ou face : <http://www-sop.inria.fr/mefisto/java/tutorial1/node8.html>

#SECTION00031020000000000000

Une autre illustration : l'aiguille de Buffon <http://www-sop.inria.fr/mefisto/java/tutorial1/node14.html>

#SECTION00033110000000000000

8.3 Théorème central-limite

Définition 8.3.1. Soit μ mesure de probabilité sur \mathbb{R}^d , on appelle fonction caractéristique de μ la fonction suivante

$$x \in \mathbb{R}^d \mapsto \hat{\mu}(x) = \int_{\mathbb{R}^d} e^{itx} \mu(dt) \in \mathbb{C}.$$

Si X est une v.a. de loi μ alors $\Phi_X = \hat{\mu}$.

Théorème 8.3.2. (dû à Paul Lévy) Soit (μ_n) une suite de mesures de probabilité sur \mathbb{R}^d ,

$$\left[\mu_n \xrightarrow[n \rightarrow +\infty]{\text{étr.}} \mu \right] \Leftrightarrow \left[\forall \xi \in \mathbb{R}^d, \hat{\mu}_n(\xi) \xrightarrow[n \rightarrow +\infty]{} \hat{\mu}(\xi) \right].$$

Ce qui s'énonce aussi

$$\left[X_m \xrightarrow[m \rightarrow +\infty]{\text{loi}} X \right] \Leftrightarrow \left[\forall \xi \in \mathbb{R}^d, \Phi_{X_m}(\xi) \xrightarrow[m \rightarrow +\infty]{} \Phi_X(\xi) \right]$$

Théorème 8.3.3. Théorème central-limite (aussé noté TCL)

Soit (X_n) une suite de v.a.r. i.i.d. avec $\mathbb{E}(X_1) = m$ et $\text{Var}(X_1) = \sigma^2$ ($m, \sigma^2 < \infty$). Alors

$$\frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow +\infty]{\text{loi}} Z \text{ de loi } \mathcal{N}(0, 1),$$

(où $\sigma > 0$ est la racine carrée de la variance).

Il existe des résultats raffinés sur la « vitesse » de cette convergence en loi. Voir, par exemple, le théorème de Berry-Esseen dans [Dur96].

Remarque 8.3.4. *Sous les hypothèses du théorème précédent, prenons $a < b$, $f(x) = \mathbf{1}_{[a,b]}(x)$. Par le théorème 8.1.6,*

$$\mathbb{E} \left(f \left(\frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \right) \right) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(f(Z)) ,$$

c'est à dire

$$\mathbb{P} \left(a \leq \frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} \leq b \right) \xrightarrow{n \rightarrow +\infty} \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx .$$

C'est cette propriété qui sera le plus souvent utilisée dans les exercices.

Démonstration du théorème 8.3.3. Posons $\forall n$, $Y_n = X_n - m$. Soient

$$S'_n = Y_1 + \cdots + Y_n, \quad Z_n = \frac{X_1 + \cdots + X_n - nm}{\sigma\sqrt{n}} = \frac{S'_n}{\sigma\sqrt{n}} .$$

Nous avons

$$\begin{aligned} \Phi_{Z_n}(t) &= \mathbb{E} \left(\exp \left(\frac{itS'_n}{\sigma\sqrt{n}} \right) \right) \\ &= \mathbb{E} \left(\exp \left(\frac{it}{\sigma\sqrt{n}} (Y_1 + \cdots + Y_n) \right) \right) \\ (\text{par indépendance des } Y_j) &= \prod_{1 \leq j \leq n} \mathbb{E} \left(\exp \left(\frac{it}{\sigma\sqrt{n}} Y_j \right) \right) \\ (\text{car les } Y_j \text{ sont identiquement distribués}) &= \Phi_{Y_1} \left(\frac{t}{\sigma\sqrt{n}} \right)^n . \end{aligned}$$

Regardons la fonction $\Phi_{Y_1}(u) = \mathbb{E}(e^{iuY_1})$ pour $u \in \mathbb{R}$. Pour tout u , $\mathbb{E}(|e^{iuY_1}|) = 1 < \infty$. Pour tout ω , $u \mapsto e^{iuY_1(\omega)}$ est dérivable et de dérivée $u \mapsto iY_1 e^{iuY_1(\omega)}$. Pour tous u , ω , $|Y_1 e^{iuY_1(\omega)}| \leq |Y_1(\omega)|$ qui est intégrable (et qui ne dépend pas de u). Donc, par théorème de dérivation (cf. cor. 4.3.6)

$$\Phi'_{Y_1}(u) = \mathbb{E}(iY_1 e^{iY_1 u}) .$$

De même, $\Phi''_{Y_1}(u) = \mathbb{E}(-Y_1^2 e^{iY_1 u})$. Donc $\Phi'_{Y_1}(0) = \mathbb{E}(iY_1) = i\mathbb{E}(Y_1) = 0$, $\Phi''_{Y_1}(0) = -\mathbb{E}(Y_1^2) = -\sigma^2$. Supposons que Φ_{Y_1} admette un développement limité en 0 (ce n'est pas toujours le cas). Ce développement est alors :

$$\begin{aligned} \Phi_{Y_1}(u) &= \Phi_{Y_1}(0) + u\Phi'_{Y_1}(0) + \frac{u^2}{2}\Phi''_{Y_1}(0) + o(u^2) \\ &= 1 - \frac{u^2\sigma^2}{2} + o(u^2) . \end{aligned}$$

Donc

$$\begin{aligned} \Phi_{Z_n}(t) &= \left(1 - \frac{t^2}{\sigma^2 n} + o\left(\frac{1}{n}\right) \right)^n \\ &= \exp \left(n \log \left(1 - \frac{t^2}{\sigma^2 n} + o\left(\frac{1}{n}\right) \right) \right) \\ &= \exp \left(-\frac{t^2}{\sigma^2} + o(1) \right) \\ &\xrightarrow{n \rightarrow +\infty} e^{-t^2/\sigma^2} \end{aligned}$$

par continuité de l'exponentielle. □

Exemple 8.3.5. On s'intéresse au nombre de gens qui achètent de la lessive Ariel en France. On ne peut pas interroger toute la population et on se contente donc d'un échantillon de personnes. Introduisons la variable

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ème personne interrogée achète Ariel} \\ 0 & \text{si la } i\text{-ème personne interrogée n'achète pas Ariel.} \end{cases}$$

Les variables X_i sont supposées i.i.d. avec $\mathbb{P}(X_i = 1) = p$ (ce sont nos hypothèses de modélisation). La quantité p est celle que nous cherchons à déterminer. Remarquons que $\mathbb{E}(X_1) = p \times 1 + (1 - p) \times 0 = p$. Par la loi (forte) des grands nombres

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n \rightarrow +\infty]{p.s.} \mathbb{E}(X_1) = p.$$

Quelle taille n d'échantillon sélectionner pour que $\frac{X_1 + \cdots + X_n}{n}$ soit proche de p ? Supposons que l'on veuille n tel que la probabilité de se tromper de plus de 0,01 dans notre estimée de p soit plus petite que 0,05, c'est à dire

$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - p\right| \geq 0,01\right) \leq 0,05. \quad (8.3.1)$$

Notons $\sigma^2 = \text{Var}(X_1)$. Nous avons

$$\begin{aligned} \mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{n} - p\right| \geq 0,01\right) &= \mathbb{P}\left(\left|\frac{(X_1 - p) + \cdots + (X_n - p)}{\sigma\sqrt{n}}\right| \geq \frac{\sqrt{n} \times 0,01}{\sigma}\right) \\ (\text{par TCL}) &\approx \mathbb{P}\left(Z \geq \frac{\sqrt{n} \times 0,01}{\sigma}\right) \text{ avec } Z \sim \mathcal{N}(0, 1) \\ &= \int_{\frac{\sqrt{n} \times 0,01}{\sigma}}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ &= 1 - \int_{-\infty}^{\frac{\sqrt{n} \times 0,01}{\sigma}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx. \end{aligned} \quad (8.3.2)$$

Nous voyons sur une table (cf. annexe A) qu'il suffit de prendre n tel que $\sqrt{n} \times 0,01 / \sigma = 1.65$. Calculons

$$\begin{aligned} \text{Var}(X_1) &= \mathbb{E}(X_1^2) - \mathbb{E}(X_1)^2 \\ &= p \times 1^2 + (1 - p) \times 0^2 - p^2 \\ &= p - p^2 = p(1 - p). \end{aligned}$$

Nous avons alors que

$$n = \left(\frac{1,65 \times \sqrt{p(1-p)}}{0,01}\right)^2$$

réalise (8.3.1). Mais justement, nous ne connaissons pas p . Nous étudions la fonction

$$p \in [0, 1] \mapsto p(1 - p).$$

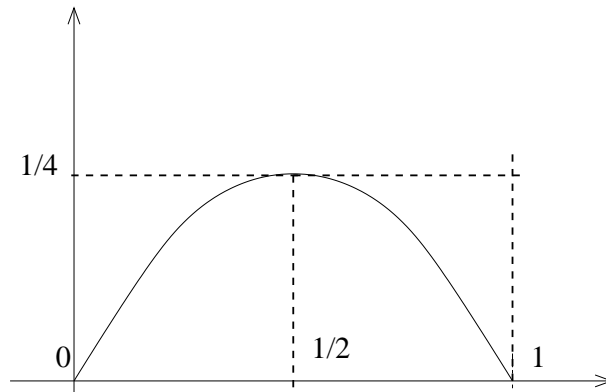


FIG. 8.1 –

C'est une parabole qui atteint son max. en $1/2$. Donc, $\forall p \in [0, 1]$,

$$\left(\frac{1,65 \times \sqrt{p(1-p)}}{0,01} \right)^2 \leq \left(\frac{1,65 \times \sqrt{0,5 \times 0,5}}{0,01} \right)^2.$$

Remarquons, au vu de (8.3.2), que si (8.3.1) est réalisée pour un certain n_1 alors elle est réalisée pour tout $n_2 \geq n_1$; donc il suffit de prendre $n = \left(\frac{1,65 \times \sqrt{0,5 \times 0,5}}{0,01} \right)^2$.

Exemple 8.3.6. Théorème de Moivre

Soient X_1, X_2, \dots i.i.d. $\sim \mathcal{B}(1/2)$. Soit $S_n = X_1 + \dots + X_n$. Calculons

$$\begin{aligned} \Phi_{S_n}(u) &= \mathbb{E}(e^{iu(X_1 + \dots + X_n)}) \\ (\text{par indépendance des } X_j) &= \mathbb{E}(e^{iuX_1})^n \\ &= \left(\frac{1}{2}(1 + e^{iu}) \right)^n \\ &= \sum_{k=0}^n C_n^k \left(\frac{1}{2} \right)^{n-k} \left(\frac{1}{2} \right)^k e^{iku} \end{aligned}$$

qui est la fonction caractéristique de $\mathcal{B}(n, \frac{1}{2})$. Donc $S_n \sim \mathcal{B}(n, 1/2)$.

Nous avons $\mathbb{E}(X_1) = 1/2$, $\text{Var}(X_1) = 1/4$ (cf. ex. précédent). Donc le TCL nous dit que pour $a \leq b$

$$\mathbb{P} \left(a \leq \frac{S_n - n/2}{(1/2)\sqrt{n}} \leq b \right) \xrightarrow{n \rightarrow +\infty} \int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx.$$

Illustration : la planche de Galton,

<http://www-sop.inria.fr/mefisto/java/tutorial1/node11.html>

#SECTION00032010000000000000 Si on règle le paramètre n à 8, chaque bille arrive en bas en une abscisse aléatoire de même loi que $S_8 - 8 \times (1/2)$. Donc l'histogramme représente la densité empirique de cette loi, qui se rapproche du dessin d'une gaussienne.

8.4 Exercices

8.4.1 Énoncés

- 1) Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ telle que $\forall x, y, |f(x) - f(y)| \leq C \inf(1, |x - y|)$ pour une certaine constante C .
 - (a) Si $X_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} X$ (rappel : pour p.t. ω , $X_n(\omega) \xrightarrow[n \rightarrow +\infty]{} X(\omega)$), montrer que $\mathbb{E}(f(X)) - \mathbb{E}(f(X_n)) \xrightarrow[n \rightarrow +\infty]{} 0$.
 - (b) Soit $\varepsilon > 0$, toujours sous l'hypothèse $X_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} X$, déduire de la question précédente que $\mathbb{P}(|f(X_n) - f(X)| \geq \varepsilon) \xrightarrow[n \rightarrow +\infty]{} 0$.
- 2) On achète un stock d'ampoules pour un lampadaire. Les ampoules ont une durée de vie de loi $\mathcal{E}(\lambda)$. La première ampoule dure un temps X_1 , on la remplace immédiatement et la deuxième qui dure un temps $X_2 \dots$. Soit $T > 0$. On admet que le nombre d'ampoules N grillées pendant le temps T est tel que N est de loi $\mathcal{P}(\lambda T)$. On suppose que $\lambda T \in \mathbb{N}$.
 - (a) Calculer $m = \mathbb{E}(N)$.
 - (b) Soit $p \in \mathbb{N}^*$. Montrer que $\mathbb{P}(N \geq m + p) = \mathbb{P}(X_1 + \dots + X_{m+p} \leq T)$.
 - (c) On suppose maintenant que $\lambda = 1$, $T = 20$, $p = 5$. Donner une valeur numérique approchée de $\mathbb{P}(N \geq m + p)$ à l'aide de la table jointe.
 - (d) Avec les mêmes valeurs numériques que ci-dessus, combien d'ampoules faut-il acheter au minimum pour que $\mathbb{P}(\text{se retrouver à court d'ampoules avant le temps } T) < 0.05$?

- 3) On rappelle que la somme de deux variables gaussiennes indépendantes, respectivement de lois $\mathcal{N}(m_1, \sigma_1^2)$ et $\mathcal{N}(m_2, \sigma_2^2)$ est une variable gaussienne de loi $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$. Soient X_1, X_2, X_3, \dots des variables indépendantes et identiquement distribuées (i.i.d.) de loi $\mathcal{N}(m, \sigma^2)$. On suppose que l'on connaît σ mais pas m , que l'on veut estimer par $S_n = \frac{1}{n}(X_1 + \dots + X_n)$.

- (a) Montrer que $\sqrt{n} \left(\frac{S_n - m}{\sigma} \right)$ est (exactement) de loi $\mathcal{N}(0, 1)$.
 (b) Dédire de l'exercice précédent que

$$\forall \delta > 0, \mathbb{P} \left(m - \delta \frac{\sigma}{\sqrt{n}} \leq S_n \leq m + \delta \frac{\sigma}{\sqrt{n}} \right) \geq 1 - \sqrt{\frac{2}{\pi}} \frac{1}{\delta} \exp \left(-\frac{\delta^2}{2} \right)$$

- (c) En déduire que

$$\forall \varepsilon > 0, \mathbb{P}(m - \varepsilon \leq S_n \leq m + \varepsilon) \geq 1 - \sqrt{\frac{2}{n\pi}} \frac{\sigma}{\varepsilon} \exp \left(-\frac{n\varepsilon^2}{2\sigma^2} \right).$$

- (d) On suppose que $\varepsilon = 0.01$, $\sigma = 1$, $n = 10000$, minorer $\mathbb{P}(|S_n - m| \leq \varepsilon)$ par une valeur numérique.

8.4.2 Corrigés

- (1) (a)

$$\begin{aligned} |\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| &\leq \mathbb{E}(|f(X_n) - f(X)|) \\ &\leq C \mathbb{E}(\inf(1, |X_n - X|)) \end{aligned}$$

Pour p.t. ω , $\inf(1, |X_n(\omega) - X(\omega)|) \xrightarrow{n \rightarrow +\infty} 0$ et $\forall \omega$, $\inf(1, |X_n(\omega) - X(\omega)|) \leq 1$.

Donc par théorème de convergence dominée, $\mathbb{E}(\inf(1, |X_n - X|)) \xrightarrow{n \rightarrow +\infty} 0$. Donc $|\mathbb{E}(f(X_n)) - \mathbb{E}(f(X))| \xrightarrow{n \rightarrow +\infty} 0$.

- (b) $\mathbb{P}(|f(X_n) - f(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon} \mathbb{E}(|f(X_n) - f(X)|)$ (inégalité de Bienaymé-Tchebycheff)

- (2) (a)

$$\begin{aligned} \mathbb{E}(N) &= \sum_{n \geq 0} n \frac{(\lambda T)^n e^{-\lambda T}}{n!} \\ &= \sum_{n \geq 1} n \frac{(\lambda T)^n e^{-\lambda T}}{n!} \\ &= (\lambda T) e^{-\lambda T} \sum_{k \geq 0} \frac{(\lambda T)^k}{k!} \\ &= \lambda T \end{aligned}$$

- (b)

$$\begin{aligned} \mathbb{P}(N \geq m + p) &= \mathbb{P}(\text{on a grillé plus de } m + p \text{ ampoules dans } [0, T]) \\ &= \mathbb{P}(\text{les } m + p \text{ premières ampoules ont déjà grillé} \\ &\quad \text{quand on arrive en } T) \\ &= \mathbb{P}(X_1 + \dots + X_{m+p} < T) \end{aligned}$$

- (c) On remarque que $\text{Var}(X_1) = 1/\lambda^2$, $\mathbb{E}(X_1) = 1/\lambda$.

$$\begin{aligned} \mathbb{P}(N \geq m + p) &= \mathbb{P}(X_1 + \dots + X_{m+p} \leq T) \\ &= \mathbb{P} \left(\frac{X_1 - \mathbb{E}(X_1) + \dots + X_{m+p} - \mathbb{E}(X_{m+p})}{(1/\lambda^2)\sqrt{m+p}} \right) \\ &\quad < \frac{T - (m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}} \end{aligned}$$

$$(\text{TCL}) \approx \int_{-\infty}^{\frac{T - (m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt.$$

On calcule $\frac{T-(m-1+p)/\lambda}{(1/\lambda)\sqrt{m-1+p}} = -1$. On a par parité :

$$\begin{aligned} \int_{-\infty}^{-1} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt &= \int_1^{+\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= 1 - \int_{-\infty}^1 \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ \text{(d'après la table)} &= 1 - 0,8413 = 0,1587 . \end{aligned}$$

(d) Ici, on cherche p pour que $\mathbb{P}(N \geq m+p) \leq 0,05$. Comme avant :

$$\begin{aligned} \mathbb{P}(N \geq m+p) &\underset{\text{TCL}}{\approx} \int_{-\infty}^{\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt \\ &= 1 - \int_{-\infty}^{-\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}}} \frac{e^{-t^2/2}}{\sqrt{2\pi}} dt . \end{aligned}$$

On regarde la table et on voit qu'il faut prendre $-\frac{T-(m+p)/\lambda}{(1/\lambda^2)\sqrt{m+p}} \geq 1,65$. Une rapide étude de fonction montre qu'il faut prendre $m+p \geq 29$.

(3) (a) $S_n \sim \mathcal{N}(nm, n\sigma^2)$ donc $\sqrt{n} \left(\frac{S_n - m}{\sigma} \right) \sim \mathcal{N}(0, 1)$.

(b) Par symétrie et par les résultats précédents :

$$\begin{aligned} \mathbb{P} \left(m - \delta \frac{\sigma}{\sqrt{n}} \leq S_n \leq m + \delta \frac{\sigma}{\sqrt{n}} \right) &= \mathbb{P} \left(\left| \sqrt{n} \left(\frac{S_n - m}{\sigma} \right) \right| \leq \delta \right) \\ &= 1 - 2\mathbb{P} \left(\sqrt{n} \left(\frac{S_n - m}{\sigma} \right) > \delta \right) \\ &\geq 1 - \sqrt{\frac{2}{\pi}} \frac{1}{\delta} \exp \left(-\frac{\delta^2}{2} \right) . \end{aligned}$$

(c) Avec $\delta = \sqrt{n}\varepsilon/\sigma$, on a (par la question précédente) :

$$\mathbb{P}(m - \varepsilon \leq S_n \leq m + \varepsilon) \geq 1 - \sqrt{\frac{2}{n\pi}} \frac{\sigma}{\varepsilon} \exp \left(-\frac{n\varepsilon^2}{2\sigma^2} \right) .$$

(d) À l'aide d'une calculatrice, on trouve :

$$\mathbb{P}(|S_n - m| \leq \varepsilon) \geq 0,8920 .$$

Chapitre 9

Conditionnement

On se donne toujours un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

9.1 Conditionnement discret

Définition 9.1.1. Soient $A, B \in \mathcal{A}$, $\mathbb{P}(B) > 0$, la probabilité de B sachant A est

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

Définition 9.1.2. Si X est une v.a. et $B \in \mathcal{A}$, $\mathbb{P}(B) > 0$, l'espérance de X sachant B est la nombre suivant

$$\mathbb{E}(X|B) = \frac{\mathbb{E}(X \mathbf{1}_B)}{\mathbb{P}(B)} .$$

Définition 9.1.3. Soit X v.a.r. et Y v.a. prenant un nombre dénombrable de valeurs. On définit l'espérance conditionnelle de X sachant Y de la manière suivante : $\mathbb{E}(X|Y)$ est une v.a. qui peut s'écrire $\mathbb{E}(X|Y) = \phi(Y)$ avec

$$\begin{aligned} \phi &: \mathbb{R} \rightarrow \mathbb{R} \\ y &\mapsto \begin{cases} \mathbb{E}(X|Y=y) = \frac{\mathbb{E}(X \mathbf{1}_{Y=y})}{\mathbb{P}(Y=y)} & \text{si } \mathbb{P}(Y=y) > 0 \\ 0 & \text{sinon} \end{cases} \end{aligned}$$

Exemple 9.1.4. Soit, $\Omega = \{1, 2, \dots, 6\}$ et $\forall \omega \in \Omega$, $\mathbb{P}(\{\omega\}) = 1/6$. Soient les v.a.

$$X(\omega) = \omega, \quad Y(\omega) = \begin{cases} 1 & \text{si } \omega \text{ impair} \\ 0 & \text{si } \omega \text{ pair} \end{cases} .$$

Si $\omega \in \{1, 3, 5\}$, alors $Y = 1$ et

$$\begin{aligned} \mathbb{E}(X|Y)(\omega) &= \frac{\mathbb{E}(X \mathbf{1}_{Y=1})}{\mathbb{P}(Y=1)} \\ &= \frac{\frac{1}{6}(1+3+5)}{\frac{3}{6}} = 3 . \end{aligned}$$

Si $\omega \in \{2, 4, 6\}$, alors $Y = 0$ et

$$\begin{aligned} \mathbb{E}(X|Y)(\omega) &= \frac{\mathbb{E}(X \mathbf{1}_{Y=0})}{\mathbb{P}(Y=0)} \\ &= \frac{\frac{1}{6}(2+4+6)}{\frac{3}{6}} = 4 . \end{aligned}$$

9.2 Espérance conditionnelle

Définition 9.2.1. Soit Y v.a. à valeurs dans un espace mesurable quelconque (E, \mathcal{E}) . La tribu engendrée par Y est $\sigma(Y) = \{Y^{-1}(A), A \in \mathcal{E}\}$. La famille $\sigma(Y)$ est une tribu et $\sigma(Y) \subset \mathcal{A}$.

On dit d'une v.a. Z à valeurs dans un espace mesurable quelconque (E', \mathcal{E}') qu'elle est $\sigma(Y)$ -mesurable si $\forall A \in \mathcal{E}', Z^{-1}(A) \in \sigma(Y)$.

Soit \mathcal{B} une tribu $\subset \mathcal{A}$, on dit que Z est \mathcal{B} -mesurable si $\forall A \in \mathcal{E}', Z^{-1}(A) \in \mathcal{B}$.

Remarque 9.2.2. Prenons une variable Z $\sigma(Y)$ -mesurable comme dans la définition ci-dessus. La tribu $\sigma(Y)$ représente les événements relatifs à Y (tous ceux qui peuvent se décrire en terme de « il est arrivé telle chose à Y »). Dire que Z est Y -mesurable revient à dire que tous les événements relatifs à Z peuvent se décrire comme des événements relatifs à Y et donc que Z est une fonction de Y .

Théorème 9.2.3. Soit \mathcal{B} une tribu $\subset \mathcal{A}$. Soit X une v.a.r. intégrable. Il existe une et une seule v.a.r. intégrable, appelée espérance conditionnelle de X sachant \mathcal{B} et notée $\mathbb{E}(X|\mathcal{B})$, qui vérifie

$$\forall B \in \mathcal{B}, \mathbb{E}(X\mathbf{1}_B) = \mathbb{E}(\mathbb{E}(X|\mathcal{B})\mathbf{1}_B).$$

La variable $\mathbb{E}(X|\mathcal{B})$ vérifie en outre que $\forall Z$ v.a. à valeurs dans \mathbb{R}^d , \mathcal{B} -mesurable et bornée,

$$\mathbb{E}(XZ) = \mathbb{E}(\mathbb{E}(X|\mathcal{B})Z).$$

Définition 9.2.4. Soit X une v.a.r. et Y une v.a. quelconque, l'espérance conditionnelle de X sachant Y est la variable suivante

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

Remarque 9.2.5. La définition ci-dessus inclut la définition 9.1.3 (les deux définitions coïncident dans le cas où Y ne prend qu'un nombre dénombrable de valeurs).

Proposition 9.2.6. Soit X, Y des v.a.r. et \mathcal{B} tribu $\subset \mathcal{A}$,

- (i) si X est \mathcal{B} -mesurable alors $\mathbb{E}(X|\mathcal{B}) = X$
- (ii) linéarité : $\forall a, b \in \mathbb{R}, \mathbb{E}(aX + bY|\mathcal{B}) = a\mathbb{E}(X|\mathcal{B}) + b\mathbb{E}(Y|\mathcal{B})$
- (iii) $\mathbb{E}(\mathbb{E}(X|\mathcal{B})) = \mathbb{E}(X)$
- (iv) $|\mathbb{E}(X|\mathcal{B})| \leq \mathbb{E}(|X||\mathcal{B})$
- (v) croissance : $X \geq Y \Rightarrow \mathbb{E}(X|\mathcal{B}) \geq \mathbb{E}(Y|\mathcal{B})$, p.s.
- (vi) si $X \perp Y$, $\mathbb{E}(XY|\mathcal{B}) = \mathbb{E}(X|\mathcal{B})\mathbb{E}(Y|\mathcal{B})$.

Démonstration. (partielle)

- (i) X est \mathcal{B} -mesurable et $\forall B \in \mathcal{B}, \mathbb{E}(X\mathbf{1}_B) = \mathbb{E}(X\mathbf{1}_B)$ donc $\mathbb{E}(X|\mathcal{B}) = X$
- (ii) soit $B \in \mathcal{B}$,

$$\begin{aligned} \mathbb{E}((a\mathbb{E}(X|\mathcal{B}) + b\mathbb{E}(Y|\mathcal{B}))\mathbf{1}_B) &= a\mathbb{E}(\mathbb{E}(X|\mathcal{B})\mathbf{1}_B) + b\mathbb{E}(\mathbb{E}(Y|\mathcal{B})\mathbf{1}_B) \\ &= a\mathbb{E}(X\mathbf{1}_B) + b\mathbb{E}(Y\mathbf{1}_B) \\ &= \mathbb{E}((aX + bY)\mathbf{1}_B) \end{aligned}$$

et $a\mathbb{E}(X|\mathcal{B}) + b\mathbb{E}(Y|\mathcal{B})$ est \mathcal{B} -mesurable (car la somme de deux variables \mathcal{B} -mesurable est \mathcal{B} -mesurable, cf. prop. 2.4.2), donc $\mathbb{E}(aX + bY|\mathcal{B}) = a\mathbb{E}(X|\mathcal{B}) + b\mathbb{E}(Y|\mathcal{B})$.

- (iii) $\Omega \in \mathcal{B}$ (car \mathcal{B} tribu) donc

$$\begin{aligned} \mathbb{E}(\mathbb{E}(X|\mathcal{B})) &= \mathbb{E}(\mathbf{1}_\Omega \mathbb{E}(X|\mathcal{B})) \\ &= \mathbb{E}(\mathbf{1}_\Omega X) = \mathbb{E}(X). \end{aligned}$$

□

Proposition 9.2.7. i) Si X, Y v.a.r. avec Y \mathcal{B} -mesurable (\mathcal{B} tribu $\subset \mathcal{A}$), alors

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

ii) Si $\mathcal{B}_1, \mathcal{B}_2$ tribus $\subset \mathcal{A}$ avec $\mathcal{B}_1 \subset \mathcal{B}_2$, alors pour toute v.a.r. X

$$\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1) = \mathbb{E}(X|\mathcal{B}_1) .$$

Démonstration. (i) Soit $B \in \mathcal{B}$, la variable $Y\mathbf{1}_B$ est \mathcal{B} -mesurable comme produit de variables \mathcal{B} -mesurables (cf. prop. 2.4.2), donc

$$\mathbb{E}(Y\mathbb{E}(X|\mathcal{B})\mathbf{1}_B) = \mathbb{E}(YX\mathbf{1}_B) .$$

La variable $Y\mathbb{E}(X|\mathcal{B})$ est \mathcal{B} -mesurable comme produit de variables \mathcal{B} -mesurables (cf. prop. 2.4.2. D'où le résultat.

(ii) Soit $B \in \mathcal{B}_1$

$$\begin{aligned} \mathbb{E}(\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1)\mathbf{1}_B) &= \mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)\mathbf{1}_B) \\ (\text{car } B \in \mathcal{B}_2) &= \mathbb{E}(X\mathbf{1}_B) . \end{aligned}$$

La variable $\mathbb{E}(\mathbb{E}(X|\mathcal{B}_2)|\mathcal{B}_1)$ est \mathcal{B}_1 -mesurable, d'où le résultat. □

Exemple 9.2.8. Reprenons l'exemple 9.1.4. Soit

$$Z = \begin{cases} 1 & \text{si } X \in \{1, 3\} \\ 2 & \text{si } X = 5 \\ 3 & \text{si } X \in \{2, 4\} \\ 4 & \text{si } X = 6 . \end{cases}$$

Remarquons que la connaissance de Z implique la connaissance de Y et que donc $\sigma(Y) \subset \sigma(Z)$. Si $\omega \in \{1, 3\}$, alors $Z = 1$ et

$$\begin{aligned} \mathbb{E}(X|Z)(\omega) &= \frac{\mathbb{E}(X\mathbf{1}_{Z=1})}{\mathbb{P}(Z=1)} \\ &= \frac{\frac{1}{6}(1+3)}{\frac{2}{6}} = 2 . \end{aligned}$$

Si $\omega = 5$, $Z = 2$ et

$$\begin{aligned} \mathbb{E}(X|Z)(\omega) &= \frac{\mathbb{E}(X\mathbf{1}_{Z=2})}{\mathbb{P}(Z=2)} \\ &= \frac{\frac{1}{6}5}{\frac{1}{6}} = 5 . \end{aligned}$$

De même, $\mathbb{E}(X|Z)(\omega) = 3$ si $\omega \in \{2, 4\}$ et $\mathbb{E}(X|Z)(\omega) = 6$ si $\omega = 6$. Calculons pour ω tel que $Y = 1$ (c'est à dire $\omega \in \{1, 3, 5\}$)

$$\begin{aligned} \mathbb{E}(\mathbb{E}(X|Z)|Y)(\omega) &= \frac{\frac{2}{6} \times 2 + \frac{1}{6} \times 5}{\frac{3}{6}} \\ &= 3 . \end{aligned}$$

De même, pour ω tel que $Y = 0$ (c'est à dire $\omega \in \{2, 4, 6\}$) : $\mathbb{E}(\mathbb{E}(X|Z)|Y)(\omega) = 4$. Par ailleurs, nous avons vu dans l'exemple 9.1.4,

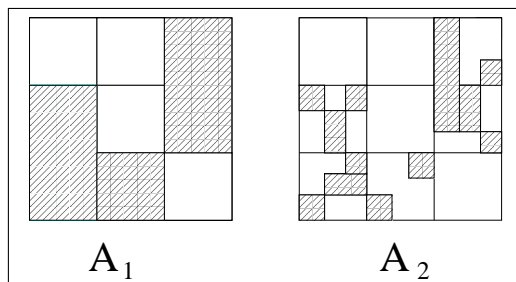
$$\mathbb{E}(X|Y) = \begin{cases} 3 & \text{si } Y = 1 \\ 4 & \text{si } Y = 0 . \end{cases}$$

Donc on a $\mathbb{E}(\mathbb{E}(X|Z)|Y) = \mathbb{E}(X|Y)$ comme annoncé dans prop. 9.2.7, (ii).

9.3 Exercice

9.3.1 Énoncé

- 1) Soit $p \in [0, 1]$. Soit A_0 le carré $[0, 1]^2 \subset \mathbb{R}^2$. L'ensemble A_1 est un ensemble aléatoire construit de la manière suivante : on découpe A_0 en 9 carrés, chaque petit carré appartient à A_1 avec probabilité p (indépendamment des autres). On recommence l'opération sur les carrés de A_1 pour former A_2 (de manière indépendante de ce qui s'est passé avant) et ainsi de suite, on obtient des ensembles A_1, A_2, A_3, \dots . Si $A_n = \emptyset$ alors $\forall k \geq n, A_k = \emptyset$. La figure ci-dessous représente une réalisation de A_1 et A_2 (hachurés) pour une certaine valeur de p .



- Pour tout n , on note Z_n le nombre de carrés de côté $1/3^n$ formant A_n . Soit $n \geq 1$, montrer que $\forall r \in [0, 1], g_{Z_n}(r) = g_{Z_{n-1}}(f(r))$ où $\forall r \in [0, 1], f(r) = (pr + 1 - p)^9$.
- En déduire que $g_{Z_n}(r) = f^{\circ n}(r)$ ("o" veut dire que l'on compose n fois).
- Montrer que f est convexe (c'est à dire que sa dérivée est une fonction croissante).
- Calculer $f(0), f(1), f'(1)$. Faire un dessin de f .
- On suppose que $p \leq 1/9$.
 - Montrer que $\forall r \in [0, 1], g_{Z_n}(r) \xrightarrow{n \rightarrow +\infty} 1$.
 - En déduire que $\mathbb{P}(Z_n = 0) \xrightarrow{n \rightarrow +\infty} 1$.
 - En déduire que $Z_n \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} 0$. (On pourra considérer l'événement $\{\omega : Z_n(\omega) \xrightarrow[n \rightarrow +\infty]{} 0\}$ comme une réunion croissante d'événements.)

On pourra se reporter à [Wil91] pour une étude plus complète de ce problème, appelé « arbre de Galton-Watson ».

9.3.2 Corrigé

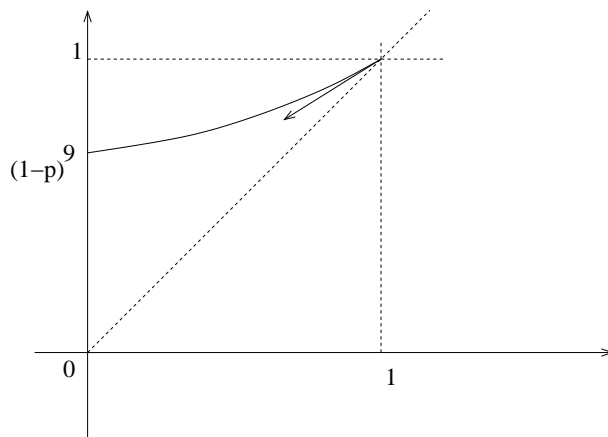
- (1) (a) Calculons

$$\begin{aligned} g_{Z_n}(r) &= \mathbb{E}(r^{Z_n}) \\ &= \mathbb{E}(\mathbb{E}(r^{Z_n} | Z_{n-1})) . \end{aligned}$$

Dans l'ensemble A_{n-1} , on numérote les carrés (de 1 à Z_{n-1}). On note pour tout $i \in \{1, \dots, Z_{n-1}\}$, X_i le nombre de carrés de A_n qui sont dans le carré numéro i de A_{n-1} . À Z_{n-1} fixé, les variables X_i sont i.i.d. de loi $\mathcal{B}(9, p)$. Nous avons donc :

$$\begin{aligned} g_{Z_n}(r) &= \mathbb{E}(\mathbb{E}(r^{X_1 + \dots + X_{Z_{n-1}}} | Z_{n-1})) \\ &= \mathbb{E}(\mathbb{E}(r^{X_1} | Z_{n-1}) \dots \mathbb{E}(r^{X_{Z_{n-1}}} | Z_{n-1})) \\ &= \mathbb{E}(\mathbb{E}(r^{X_1} | Z_{n-1})^{Z_{n-1}}) \\ &= \mathbb{E}(f(r)^{Z_{n-1}}) \\ &= g_{Z_{n-1}}(f(r)) . \end{aligned}$$

- (b) Par récurrence : $g_{Z_{n-1}}(r) = g_{Z_0}(f^{\circ n}(r))$. Or Z_0 est constante égale à 1, donc $g_{Z_0}(r) = r$, donc $g_{Z_{n-1}}(r) = f^{\circ n}(r)$.
- (c) Calculons $f'(r) = 9p(pr + 1 - p)^8$. La fonction f' est positive (pour $r \in [0, 1]$) donc f est convexe (sur $[0, 1]$).
- (d) Calculons $f(0) = (1 - p)^9$, $f(1) = 1$, $f'(1) = 9p$.

FIG. 9.1 – Dessin de f pour un $p < 1.9$.

- (e) i. (Pas de démonstration, on le voit sur le dessin.)
- ii. Nous avons $\mathbb{P}(Z_n = 0) = g_{Z_n}(0) \xrightarrow{n \rightarrow +\infty} 1$ par la question précédente.
- iii. Soit $B_n = \{\omega : Z_n(\omega) = 0\}$. Si $Z_n(\omega) = 0$ alors $Z_{n+1}(\omega) = 0$ donc $B_n \subset B_{n+1}$. Par réunion croissante $\mathbb{P}(\cup_{n \geq 0} B_n) = \lim_{n \rightarrow +\infty} \mathbb{P}(B_n) = 1$ par la question précédente. Si $\omega \in \cup_{n \geq 0} B_n$ alors $Z_n(\omega) \xrightarrow{n \rightarrow +\infty} 0$, d'où le résultat.

Chapitre 10

Variables gaussiennes

On se donne toujours un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

Les variables gaussiennes sont très utilisées en modélisation à cause de leurs propriétés, que nous allons détailler dans ce chapitre.

10.1 Définitions et propriétés

Définition 10.1.1. Une v.a. X à valeurs dans \mathbb{R}^d est dite gaussienne si $\forall u \in \mathbb{R}^d$, $\langle u, X \rangle$ est une v.a.r. gaussienne. (On dit aussi que X est un vecteur gaussien.)

Théorème 10.1.2. La loi d'une v.a. gaussienne $X = (X_1, \dots, X_d)$ dans \mathbb{R}^d est entièrement déterminée par le vecteur $m = \mathbb{E}(X) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_d))$ et la matrice carrée $\Sigma_X = ((\mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)))_{1 \leq i, j \leq d}$ (dite matrice de covariance). On note $\text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)$. Sa fonction caractéristique est alors

$$\forall u \in \mathbb{R}^d, \quad \Phi(u) = \mathbb{E}(e^{i\langle u, X \rangle}) = \exp \left(i\langle u, m \rangle - \frac{1}{2} \langle \Sigma_X u, u \rangle \right).$$

Remarque 10.1.3. Le symbole $\langle \cdot, \cdot \rangle$ est le produit scalaire usuel dans \mathbb{R}^d . Pour $u = (u_1, \dots, u_d)$ et $m = (m_1, \dots, m_d)$:

$$\langle u, m \rangle = u_1 m_1 + \dots + u_d m_d.$$

Proposition 10.1.4.

$$\begin{aligned} \text{les v.a. } X_1, \dots, X_d \text{ sont indépendantes} &\Leftrightarrow \Sigma_X \text{ est diagonale} \\ &\Leftrightarrow \forall i \neq j, \mathbb{E}(X_i X_j) = \mathbb{E}(X_i)\mathbb{E}(X_j) \end{aligned}$$

Démonstration partielle. Supposons que Σ_X est diagonale. Écrivons

$$\Sigma_X = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & \dots \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \sigma_d^2 \end{bmatrix}.$$

Soient Y_1, \dots, Y_d des v.a.r. telles que X_j et Y_j ont même loi pour tout j et Y_1, \dots, Y_d sont

indépendantes. Calculons

$$\begin{aligned}
 \Phi_X(u) &= \exp\left(i(u_1 m_1 + \cdots + u_d m_d) - \frac{1}{2}(\sigma_1^2 u_1^2 + \cdots + \sigma_d^2 u_d^2)\right) \\
 &= \prod_{j=1}^d \exp\left(iu_j m_j - \frac{1}{2}\sigma_j^2 u_j^2\right) \\
 &= \prod_{j=1}^d \Phi_{X_j}(u_j) \\
 &= \prod_{j=1}^d \Phi_{Y_j}(u_j) \\
 (\text{car les } Y_j \text{ ind.}) &= \Phi_{(Y_1, \dots, Y_d)}(u) .
 \end{aligned}$$

De manière analogue au théorème 6.5.4, ceci prouve que $X = (X_1, \dots, X_d)$ et (Y_1, \dots, Y_d) ont même loi et donc X_1, \dots, X_d sont indépendants. \square

Proposition 10.1.5. *Soit X vecteur gaussien sur \mathbb{R}^d .*

- *La loi de X a une densité (par rapport à la mesure de Lebesgue) si, et seulement si, $\forall u \in \mathbb{R}^d \setminus \{0\}, \langle u, \Sigma_X u \rangle > 0$.*
- *Dans le cas où X a une densité, celle-ci est*

$$x \in \mathbb{R}^d \mapsto \frac{1}{\sqrt{\det(2\pi\Sigma_X)}} \exp\left(-\frac{1}{2}\langle \Sigma_X^{-1}(x - m), x - m \rangle\right) .$$

10.2 Gaussiennes et espérance conditionnelle

Théorème 10.2.1. *Soit (Y_1, \dots, Y_n, X) un vecteur gaussien centré (c'est à dire que $\mathbb{E}(Y_1) = \cdots = \mathbb{E}(Y_n) = \mathbb{E}(X) = 0$). Alors, $\exists \lambda_1, \dots, \lambda_n \in \mathbb{R}$ tels que*

$$\mathbb{E}(X|\sigma(Y_1, \dots, Y_n)) = \sum_{j=1}^n \lambda_j Y_j .$$

De plus, pour toute fonction mesurable $h : \mathbb{R} \rightarrow \mathbb{R}^+$,

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

avec

$$\sigma^2 = \mathbb{E}\left(\left(X - \sum_{j=1}^n \lambda_j Y_j\right)^2\right), \quad m = \sum_{j=1}^n \lambda_j Y_j .$$

Remarque 10.2.2. *Comme exposé dans la remarque 9.2.2, $\mathbb{E}(h(X)|\sigma(Y_1, \dots, Y_n))$ est une v.a. qui s'écrit comme une fonction de Y_1, \dots, Y_n .*

Exemple 10.2.3. *Calcul des λ_i apparaissant dans le théorème ci-dessus.*

Notons $Z = \mathbb{E}(X|\sigma(Y_1, \dots, Y_n))$. Nous avons $\forall i \in \{1, \dots, n\}$,

$$\mathbb{E}(ZY_i) = \mathbb{E}(XY_i) = \text{Cov}(X, Y_i) .$$

Et par ailleurs

$$\begin{aligned}
 \mathbb{E}(ZY_i) &= \mathbb{E}\left(\sum_{k=1}^n \lambda_k Y_i Y_k\right) \\
 &= \sum_{k=1}^n \lambda_k \text{Cov}(Y_i, Y_k) .
 \end{aligned}$$

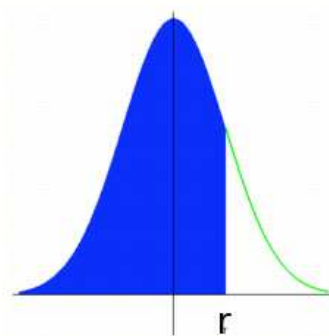
Donc

$$\begin{aligned}\Sigma_Y \times \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{bmatrix} &= \begin{bmatrix} \text{Cov}(X, Y_1) \\ \dots \\ \text{Cov}(X, Y_n) \end{bmatrix} \\ \begin{bmatrix} \lambda_1 \\ \dots \\ \lambda_n \end{bmatrix} &= \Sigma_Y^{-1} \times \begin{bmatrix} \text{Cov}(X, Y_1) \\ \dots \\ \text{Cov}(X, Y_n) \end{bmatrix} .\end{aligned}$$

Annexe A

Table de la loi normale

$$P(X \leq r) \text{ avec } X \sim N(0,1)$$



	0	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

Bibliographie

- [DMRR06] Pierre Del Moral, Bruno Rémillard, and Sylvain Rubenthaler. *Une introduction aux probabilités*. Ellipses, Paris, 2006.
- [Dur96] Richard Durrett. *Probability : theory and examples*. Duxbury Press, Belmont, CA, second edition, 1996.
- [JP03] Jean Jacod and Philip Protter. *L'essentiel en théorie des probabilités*. Cassini, Paris, 2003.
- [Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991.

Index

- A^c , 5
- L^p , 61
- $\mathbf{1}$, 9
- \mathbb{E} , 41
- Ω , 5
- $\|\cdot\|$, 61
- $\|\cdot\|_\infty$, 24
- \circ , 1, 2, 10, 31
- \perp , 53
- $\langle \cdot, \cdot \rangle$, 75
- $\mathcal{B}(\cdot)$, 44
- $\mathcal{B}(\cdot, \cdot)$, 44
- $\mathcal{C}_b^+(\mathbb{R}^d)$, 43
- \mathcal{C}^1 , 31
- $\mathcal{E}(\cdot)$, 45
- $\mathcal{G}(\cdot)$, 44
- $\mathcal{N}(\cdot, \cdot)$, 45
- $\mathcal{P}(\cdot)$, 44
- $\mathcal{U}(\cdot)$, 44, 45
- \otimes , 29
- \mathbb{P} , 37
- σ -fini, 19
- $\sigma(\cdot)$, 5, 54, 70
- \sim , 46
- \star , 24, 33, 56
- $f^{-1}(A)$, 2, 10
- aléas, 5
- Application mesurable, 10
- Bijection, 1
- Binôme de Newton, 42
- Boréliens, 6
- càdlàg, 13, 38
- Calcul de loi, 43
- Calcul de volume, 33
- Changement de variable, 31
- Convergence
 - L^p , 61
 - en loi, 61
 - en probabilité, 61
 - étroite, 61
 - presque sûre, 19, 61
 - simple, 19
- Convolution, 24, 33, 56
- Coordonnées polaires, 32
- Covariance, 75
- Densité, 11, 30, 38
- Difféomorphisme, 31
- Ensemble dénombrable, 1
- Ensemble négligeable, 17
- Espérance conditionnelle, 69, 70
- Espace complet, 17
- Espace mesuré, 6
- Espace mesurable, 6
- Espace probabilisé, 37
- Espérance, 41
- Événement, 5
- Événements indépendants, 53
- Fonction étagée, 9
- Fonction caractéristique, 45, 57, 63
- Fonction de répartition, 13, 38
- Fonction génératrice, 46, 57
- Fonction indicatrice, 9
- Fonction intégrable, 9, 10, 12
- Fonction test, 43
- Fubini, 29
- Gaussienne, 75
- i.i.d., 62
- Indépendance, 75
- Inégalité
 - de Bienaymé-Tchebichev, 44
 - de Jensen, 26, 43
 - de Markov, 11, 44
- Injection, 1
- Intégrale multiple, 30
- Intégrabilité, 10, 12, 41
- Intégrale
 - d'une fonction étagée positive, 9
 - d'une fonction mesurable positive, 10
 - de Lebesgue, 12
 - de Riemann, 12, 32
- Intégrales dépendant d'un paramètre, 23
- Intégration sur \mathbb{N} , 22
- Intersection décroissante, 7
- Lancer de dé, 37, 69, 71
- Lemme de Borel-Cantelli, 55
- Lemme de Fatou, 21
- Loi
 - binômiale, 44
 - de Bernoulli, 44
 - de Poisson, 44
 - exponentielle, 45

- faible des grands nombres, 62
- forte des grands nombres, 62
- géométrique, 44
- gaussienne, 45
- normale, 45
- uniforme, 44, 45
- Loi d'une variable aléatoire, 37
- Lois classiques, 44
- Lois discrètes, 44

- Matrice de covariance, 75
- Matrice jacobienne, 31
- Mesurabilité, 10, 19, 54, 70
- Mesure, 6
- Mesure d'une intersection décroissante, 7
- Mesure d'une réunion croissante, 7
- Mesure de Lebesgue, 8
- Mesure de probabilité, 6, 37
- Mesure image, 10, 37
- Mesure produit, 29
- Modélisation, 5, 37

- p.p, 17
- p.s., 17
- presque partout, 17
- presque sûrement, 17
- Probabilité, 37
- Probabilité conditionnelle, 69

- Réunion croissante, 7

- Singleton, 6
- Sondages, 65
- Surjection, 1

- TCL, 63
- Théorème
 - central-limite, 63
 - de comparaison, 11
 - de continuité globale sous l'intégrale, 23
 - de continuité sous l'intégrale, 23
 - de convergence domin'ee, 22
 - de convergence monotone, 19
 - de dérivation globale sous l'intégrale, 24
 - de dérivation sous l'intégrale, 24
 - de Fubini, 29
 - de Fubini Tonelli, 29
 - de Moivre, 66
- Tribu, 5
- Tribu complétée, 17
- Tribu des Boréliens, 6, 29
- Tribu engendrée, 5, 54, 70
- Tribu produit, 29
- Tribu, plus petite, 5, 29

- Univers, 5

- v.a., 37
- v.a.r., 37
- Variable aléatoire, 37
- Variable aléatoire intégrable, 41
- Variable finie p.s., 41
- Variables indépendantes, 53
- Variables indépendantes identiquement distribuées, 62
- Variance, 42
- Vecteur gaussien, 75