
INTÉGRATION ET MESURES

Licence de Mathématiques 3ème année

Jean-Christophe BRETON

Université de LA ROCHELLE

Septembre–Décembre 2009

version du 21 décembre 2009

Table des matières

I	Intégrale de Riemann	1
1	Intégrales des fonctions en escalier	3
1.1	Fonctions en escalier	3
1.2	Intégrale des fonctions en escalier	4
1.3	Sommes de Darboux et de Riemann	6
2	Intégrale de Riemann	9
2.1	Fonctions Riemann-intégrables	9
2.2	Propriétés de l'intégrale de Riemann	13
2.3	Intégrale et primitive	18
3	Fonctions réglées	23
3.1	Définition	23
3.2	Propriétés des fonctions réglées	24
3.3	Intégrale des fonctions réglées	24
4	Intégrales impropres	27
4.1	Définition et propriétés	27
4.2	Intégrales impropres des fonctions positives	30
5	Suites et séries de fonctions Riemann-intégrables	35
5.1	Différentes convergences de fonctions	35
5.2	Intégrabilité des suites et séries de fonctions	37
5.3	Quelques inconvénients de l'intégrale de Riemann	41
II	Intégrale de Lebesgue	43
6	Tribus (σ-algèbres) et mesures	45
6.1	Introduction à la théorie de la mesure	45
6.2	Algèbres et σ -algèbres	46
6.3	Mesure	49
6.4	Propriétés des mesures	51

7	Fonctions mesurables	53
7.1	Définition	53
7.2	Propriétés des applications mesurables	54
7.3	Limite de fonctions mesurables	57
7.4	Variables aléatoires	58
8	Intégrales des fonctions mesurables positives	61
8.1	Fonctions étagées (simples)	61
8.2	Intégrale des fonctions positives	63
8.3	Propriétés de l'intégrale	65
9	Théorèmes limites pour l'intégrale des fonctions mesurables positives	69
9.1	Convergence monotone	69
9.2	Lemme de Fatou	73
10	Intégrales des fonctions mesurables de signe quelconque	77
10.1	Définition et propriétés	77
10.2	Transfert	79
10.3	Loi des variables aléatoires	80
10.4	Ensembles négligeables	83
10.5	Convergence dominée et applications	85
10.6	Espérance probabiliste	87
11	Lien avec l'intégrale de Riemann	93
11.1	Fonctions en escalier	93
11.2	Fonctions continues de $[a, b] \rightarrow \mathbb{R}$	94
11.3	Fonctions Riemann-intégrables	94
11.4	Cas des intégrales de Riemann impropres	96
12	Espace produit	99
12.1	σ -algèbre produit	99
12.2	Mesure produit	101
12.3	Théorèmes de Fubini	102
12.4	Vecteurs aléatoires	105
12.5	Indépendance de variables aléatoires	108
12.6	Changement de variable	111
12.6.1	Rappel : intégrale de Riemann	111
12.6.2	Changement de variable	111
12.6.3	Coordonnées polaires et sphériques	112
13	Espaces L^p	115
13.1	Convexité	115
13.2	Espace \mathcal{L}^p	116
13.3	Espaces L^p	117

13.4	Inégalités (Hölder, Minkowski)	118
13.5	Parties denses	122
13.6	Moments des variables aléatoires	124
14	Convolution	129
14.1	Définition et propriétés	129
14.1.1	Normes des convolutions	130
14.2	Dérivation des convolutions	134
14.3	Approximation	136
14.4	Sommes de variables aléatoires indépendantes	139
15	Transformée de Fourier	143
15.1	Transformée de Fourier $L^1(\mathbb{R}^n)$	143
15.2	Inversion	146
15.3	Transformée de Fourier $L^2(\mathbb{R}^n)$	147
15.4	Fonction caractéristique	148

Introduction et rappels

Le calcul d'intégrales a déjà été rencontré les années précédentes dans des cas bien concrets, pour des intégrales de fonctions usuelles. Depuis le L1-L2 (deug), les techniques de calcul de primitives, d'aires, d'intégration par parties ou de changements de variables permettent de mener à bien les calculs effectifs d'intégrales de fonctions usuelles.

On se propose dans ce cours de donner une construction théorique de l'intégration qui recouvre les méthodes de calculs déjà connues.

Il y a plusieurs théories de l'intégration. Le premier, Cauchy en donne une pour les fonctions continues de $[a, b] \rightarrow \mathbb{R}$ (1823). Son approche est géométrique, il considère $\int_a^b f(x)dx$ comme une aire. Un peu plus tard, Riemann constate que la condition de continuité de l'intégrand f pour le calcul de $\int_a^b f(x)dx$ est inutile : il suffit que les fonctions soient limites de fonctions en escalier. Il donne donc une théorie plus générale pour les fonctions limites de fonctions en escalier (1854).

C'est dans le cadre de cette théorie que se font tous les calculs d'intégrale rencontrés jusqu'à maintenant depuis le Deug.

L'intégrale de Riemann sera l'objet du début de ce cours. On la présentera comme Darboux l'a fait (1875).

Ce type d'intégrales se calcule sur des domaines bornés $\int_a^b f(x)dx$. Quand ce n'est pas le cas, on peut avoir recours à des intégrales généralisées $\int_{-\infty}^{+\infty} f(x)dx$ (dites intégrales impropres) mais elles ne vérifient pas toutes les propriétés des intégrales classiques (sur les intervalles bornés).

Dans la théorie de Riemann, certains calculs posent des problèmes. En particulier, pour les problèmes d'interversion de somme et d'intégration (soulevés par Fourier) :

$$\sum_n \int f_n(x)dx = \int \sum_n f_n(x)dx,$$

ou de limite et d'intégrale :

$$\lim_{n \rightarrow +\infty} \int f_n(x)dx = \int \lim_{n \rightarrow +\infty} f_n(x)dx$$

on ne dispose d'aucun résultat satisfaisant alors qu'en pratique ce type de problème se pose souvent.

Par ailleurs, l'intégrale de Riemann concerne l'intégrale des fonctions de $[a, b]$ (ou \mathbb{R}) dans \mathbb{R} et se généralise mal au cas de fonctions de \mathbb{R}^n dans \mathbb{R} ou plus généralement de fonctions f définies sur des espaces plus abstraits.

Ces problèmes trouvent une réponse élégante et efficace avec une nouvelle théorie de l'intégration, introduite par Henri Lebesgue en 1902. Elle jouira d'outils très puissants : le théorème de convergence dominée répondra au problème d'interversion $\lim \int \cdots = \int \lim \cdots$, le théorème de Fubini au calcul des intégrales de fonctions de \mathbb{R}^n dans \mathbb{R} . Nous en présenterons une version simplifiée dans la seconde partie de ce cours.

L'intégrale de Lebesgue prend place dans une théorie de l'intégration beaucoup plus générale appelée la théorie de la mesure. Dans cette théorie, on définit des intégrales de fonctions définies sur des espaces abstraits par rapport à des mesures abstraites.

Par exemple ces mesures peuvent être discrètes et on retrouve la notion de série, ces mesures peuvent être des probabilités et on retrouve la notion d'espérance probabiliste.

Première partie
Intégrale de Riemann

Chapitre 1

Intégrales des fonctions en escalier

En Deug, plusieurs techniques de calculs sont étudiées et utilisées pour calculer des intégrales (intégration par parties, changements de variable). Ces calculs relèvent de l'intégrale de Riemann. On va en donner dans cette première partie, une construction plus théorique et rappeler (ou démontrer) les principales propriétés de ce type de calcul.

Pour cela, on commence par s'intéresser à des fonctions étagées (fonctions en escalier) pour lesquelles on va définir l'intégrale et vérifier ses principales propriétés. Puis on généralisera cette construction à une classe de fonctions plus large (dite intégrables) qui contient toutes les fonctions usuelles (en particulier les fonctions continues).

Dans ce chapitre, $[a, b]$ désigne un intervalle fini.

1.1 Fonctions en escalier

Définition 1.1.1 (Subdivision) Soit $[a, b]$ un intervalle fini, on appelle subdivision S de $[a, b]$ toute suite finie ordonnée $(t_i)_{0 \leq i \leq n}$ de $[a, b]$: $a = t_0 \leq t_1 \leq \dots \leq t_n = b$.

On appelle pas de la subdivision

$$\rho(S) = \max_{0 \leq i \leq n-1} |t_{i+1} - t_i|.$$

Exemples : • $S_1 = \{0, \frac{1}{2}, 1\}$, $S_2 = \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\}$, $S_3 = \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, $S_4 = \{0, \frac{2}{5}, \frac{1}{2}, \frac{5}{6}, 1\}$ sont des subdivisions de $[0, 1]$.

• La subdivision uniforme sur $[a, b]$ est celle de points $t_i = a + i \frac{b-a}{n}$, $0 \leq i \leq n$, elle est de pas $\frac{b-a}{n}$.

Précédemment, S_1, S_2, S_3 sont uniformes de pas respectivement $\rho(S_1) = 1/2$, $\rho(S_2) = 1/4$ et $\rho(S_3) = 1/3$. Par contre, S_4 n'est pas uniforme et de pas $\rho(S_4) = 2/5$.

Définition 1.1.2 Soient S et S' deux subdivisions de $[a, b]$, S' est dite plus fine que S si $S \subset S'$, c'est à dire si la suite $(t_i)_i$ qui définit S est incluse dans celle $(t'_j)_j$ de S' .

Précédemment, S_2 est plus fine que S_1 ou encore $S' = \{0, 1, 2, 3, 4\}$ est plus fine que $S = \{0, 2, 4\}$, subdivisions de $[0, 4]$.

En particulier, si on a deux subdivisions S_1 et S_2 alors la subdivision $S_3 = S_1 \cup S_2$ raffine à la fois S_1 et S_2 .

On rappelle que pour un ensemble A , la fonction indicatrice $\mathbf{1}_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$ est celle qui indique si x est dans A ou non.

Définition 1.1.3 On appelle fonction en escalier sur $[a, b]$ toute fonction $f : [a, b] \rightarrow \mathbb{R}$ telle qu'il existe une subdivision $S = (t_i)_{0 \leq i \leq n}$ de $[a, b]$ avec f constante sur chaque intervalle $[t_i, t_{i+1}[$. La fonction f s'écrit alors :

$$f(x) = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}(x). \quad (1.1)$$

Par exemple $f(x) = 2\mathbf{1}_{[0,2[}(x) - 5\mathbf{1}_{[2,4[}(x) + 3\mathbf{1}_{[4,5]}(x)$, $g(x) = \mathbf{1}_{[0,1/2[}(x) + 2\mathbf{1}_{[1/2,3[}(x) - 3\mathbf{1}_{[3,5]}(x)$.

Le nom de ces fonctions vient de l'allure de leur représentation graphique. Dessiner par exemple les graphes de f et de g .

Proposition 1.1.1 L'ensemble $\mathcal{E}([a, b])$ des fonctions en escalier sur $[a, b]$ est un sous-espace vectoriel de l'ensemble des fonctions de $[a, b]$ dans \mathbb{R} .

En effet c'est stable par multiplication par un scalaire par exemple

$$5f(x) = 10\mathbf{1}_{[0,2[}(x) - 25\mathbf{1}_{[2,4[}(x) + 15\mathbf{1}_{[4,5]}(x), \quad -2g(x) = -2\mathbf{1}_{[0,1/2[}(x) - 4\mathbf{1}_{[1/2,3[}(x) + 6\mathbf{1}_{[3,5]}(x)$$

sont encore en escalier. Et c'est stable par addition :

$$f(x) + g(x) = 3\mathbf{1}_{[0,1/2[}(x) + 4\mathbf{1}_{[1/2,2[}(x) - 3\mathbf{1}_{[2,3[}(x) - 8\mathbf{1}_{[3,4[}(x) + 0\mathbf{1}_{[4,5]}(x)$$

où il faut bien comprendre d'abord que si f se décompose sur $[0, 2[\cup [2, 4[\cup [4, 5]$ et g sur $[0, 1/2[\cup [1/2, 3[\cup [3, 5]$ alors $f + g$ se décompose sur $[0, 1/2[\cup [1/2, 2[\cup [2, 3[\cup [3, 4[\cup [4, 5]$, partition qui raffine à la fois celles de f et de g . En tout cas $f + g$ est bien en escalier.

Plus généralement, il faudrait voir que pour toutes fonctions en escalier f et g alors $\alpha f + \beta g$ l'est encore pour tous réels α, β (exercice).

1.2 Intégrale des fonctions en escalier

Définition 1.2.1 On appelle intégrale de f fonction en escalier donnée par (1.1) le nombre réel

$$\int_a^b f(x) dx = \sum_{i=0}^{n-1} \alpha_i (t_{i+1} - t_i)(x).$$

Remarque 1.2.1 • Une fonction en escalier peut avoir plusieurs écritures du type (1.1) en raffinant la subdivision, par exemple

$$f(x) = 2\mathbf{1}_{[0,2[}(x) - 5\mathbf{1}_{[2,4[}(x) + 3\mathbf{1}_{[4,5]}(x) = \underbrace{2\mathbf{1}_{[0,3/4[}(x) + 2\mathbf{1}_{[3/4,2[}(x)} - \underbrace{5\mathbf{1}_{[2,3[}(x) - 5\mathbf{1}_{[3,4[}(x)} + 3\mathbf{1}_{[4,5]}(x).$$

On montre cependant que l'intégrale $\int_a^b f(x)dx$ définie précédemment ne dépend pas de l'écriture particulière de f : deux écritures différentes de la même fonction en escalier donnent la même valeur de l'intégrale. La définition a donc bien un sens.

• Une fonction constante sur $[a, b]$ est une fonction en escalier particulière $f(x) = \alpha\mathbf{1}_{[a,b]}$, avec la subdivision triviale $\{a, b\}$ de $[a, b]$. Son intégrale est alors

$$\int_a^b f(x)dx = \alpha(b - a)$$

si α est la valeur constante de cette fonction.

- Par exemple

$$\begin{aligned} \int_0^5 f(x)dx &= 2 \times 2 - 5 \times 2 + 3 \times 1 = 4 - 10 + 3 = -3, \\ \int_0^5 g(x)dx &= 1 \times \frac{1}{2} + 2 \times \left(3 - \frac{1}{2}\right) - 3 \times 2 = \frac{1}{2} + 5 - 6 = -\frac{1}{2}. \end{aligned}$$

Propriétés des intégrales de fonctions en escalier

• **(linéarité)** L'application $f \mapsto \int_a^b f(x)dx$ est une application linéaire de $\mathcal{E}([a, b])$ l'ensemble des fonctions en escalier dans \mathbb{R} . L'intégrale est donc linéaire :

$$\int_a^b (\alpha f + \beta g)(x)dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx.$$

Pour cela, il faut raffiner les subdivisions S de f et S' de g en $S'' = S \cup S'$, adaptée à la fois à f et à g et donc à $f + g$, puis utiliser la linéarité de la somme.

• **(positivité)** Si f est une fonction en escalier positive alors son intégrale $\int_a^b f(x)dx$ est positive.

En effet, dans ce cas, son intégrale est une somme de termes positifs donc elle est positive.

• **(ordre)** Si f et g sont des fonctions en escalier avec $f \leq g$ alors $\int_a^b f(x)dx \leq \int_a^b g(x)dx$. Il s'agit de la propriété de positivité appliquée à $g - f \geq 0$.

• **(valeurs absolues)** Pour toute fonction en escalier f , on a

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

En effet, d'abord, si f est en escalier, $|f|$ l'est aussi et

$$f(x) = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}, \quad |f(x)| = \sum_{i=0}^{n-1} |\alpha_i| \mathbf{1}_{[t_i, t_{i+1}[}.$$

Comme la valeur absolue d'une somme est majorée par la somme des valeurs absolues, le résultat suit facilement.

- **(Relation de Chasles)** Si f est une fonction en escalier sur $[a, b]$ et $c \in [a, b]$ alors

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

En effet, en rajoutant c à la subdivision S , la somme définissant $\int_a^b f(x)dx$ se scinde en deux, l'une correspondant à la partie de subdivision avant c (égale à $\int_a^c f(x)dx$), l'autre à celle après c (égale à $\int_c^b f(x)dx$).

- Soit f une fonction en escalier sur $[a, b]$ et $u : \mathbb{R} \rightarrow \mathbb{R}$ une application linéaire. Alors $u \circ f$ est en escalier et

$$\int_a^b (u \circ f)(x)dx = u \left(\int_a^b f(x)dx \right).$$

En effet, $u \circ f$ est constante sur $[t_i, t_{i+1}]$ et y vaut $u(\alpha_i)$ si f y vaut α_i . Mais alors

$$\int_a^b (u \circ f)(x)dx = \sum_{i=1}^{n-1} u(\alpha_i)(t_{i+1} - t_i) = u \left(\sum_{i=1}^{n-1} \alpha_i(t_{i+1} - t_i) \right) = u \left(\int_a^b f(x)dx \right).$$

- Soit f en escalier sur $[a, b]$ alors

$$\left| \int_a^b f(x)dx \right| \leq (b - a) \sup_{x \in [a, b]} |f(x)|.$$

Notons $f = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}$. On a $\sup_{x \in [a, b]} |f(x)| = \max_{1 \leq i \leq n} |\alpha_i| =: A$. On a

$$\left| \int_a^b f(x)dx \right| = \left| \sum_{i=0}^{n-1} \alpha_i(t_{i+1} - t_i) \right| \leq \sum_{i=0}^{n-1} |\alpha_i|(t_{i+1} - t_i) = A \sum_{i=0}^{n-1} (t_{i+1} - t_i) = A(b - a).$$

1.3 Sommes de Darboux et de Riemann

On se donne f une fonction de $[a, b]$ dans \mathbb{R} bornée. Pour définir son intégrale, on va approcher f par des fonctions en escalier. Etant donnée une subdivision S , on définit des fonctions en escalier qui minorent f et qui majorent f : soient

$$E_{(f, S)}^-(x) = \sum_{i=0}^{n-1} m_i \mathbf{1}_{[t_i, t_{i+1}[}(x). \quad (1.2)$$

où $m_i = \inf_{x \in [t_i, t_{i+1}[} f(x)$ et avec $M_i = \sup_{x \in [t_i, t_{i+1}[} f(x)$, on considère

$$E_{(f,S)}^+(x) = \sum_{i=0}^{n-1} M_i \mathbf{1}_{[t_i, t_{i+1}[}(x). \quad (1.3)$$

Plus généralement, on peut approcher f par

$$\tilde{E}_{(\alpha,f,S)}(x) = \sum_{i=0}^{n-1} f(\alpha_i) \mathbf{1}_{[t_i, t_{i+1}[}(x) \quad (1.4)$$

où $\alpha_i \in [t_i, t_{i+1}[$ est quelconque. S'ils existent, avec α_i qui réalise le minimum de f sur $[t_i, t_{i+1}[$, on obtient la première fonction en escalier (1.2), avec α_i qui réalise le maximum, on obtient la seconde (1.3).

Souvent, on prend $\alpha_i = t_i$ pour tous les i ou $\alpha_i = t_{i+1}$ (c'est à dire la gauche ou la droite des intervalles). Et quand la subdivision est uniforme, on considère fréquemment les fonctions en escalier suivantes :

$$\tilde{E}_{(f,S_{\text{unif}})}(x) = \sum_{i=0}^{n-1} f\left(a + i \frac{b-a}{n}\right) \mathbf{1}_{\left[a + i \frac{b-a}{n}, a + (i+1) \frac{b-a}{n}\right]}(x).$$

Exemple : Donner ces fonctions en escalier pour la fonction $f(x) = \sin x$ et la subdivision $\{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{3\pi}{4}, \pi\}$ de $[0, \pi]$.

Ces fonctions en escalier vérifient les propriétés élémentaires suivantes :

Proposition 1.3.1

- 1) $E_{(f,S)}^-(x) \leq f(x) \leq E_{(f,S)}^+(x)$.
- 2) $E_{(f,S)}^-(x) \leq \tilde{E}_{(\alpha,f,S)}(x) \leq E_{(f,S)}^+(x)$.
- 3) Si $S \subset S'$, on a $E_{(f,S)}^-(x) \leq E_{(f,S')}^-(x)$ et $E_{(f,S')}^+(x) \leq E_{(f,S)}^+(x)$.

Le 3) se comprend de la façon suivante : quand on affine la subdivision, les fonctions en escalier deviennent plus précises et se rapprochent de f .

Définition 1.3.1 Etant donnée une subdivision S , on appelle somme de Darboux inférieure l'intégrale de la fonction en escalier (1.2)

$$A^-(f, S) = \sum_{i=0}^{n-1} m_i (t_{i+1} - t_i).$$

On appelle somme de Darboux supérieure l'intégrale de la fonction en escalier (1.3)

$$A^+(f, S) = \sum_{i=0}^{n-1} M_i (t_{i+1} - t_i)$$

toujours avec

$$m_i = \inf_{x \in [t_i, t_{i+1}[} f(x), \quad M_i = \sup_{x \in [t_i, t_{i+1}[} f(x).$$

Ce sont en fait les intégrales des fonctions en escalier $E_{(f,S)}^-$ et $E_{(f,S)}^+$.

Définition 1.3.2 *Etant donnée une subdivision S , on appelle somme de Riemann l'intégrale d'une fonction en escalier du type (1.4)*

$$R(f, S, \alpha) = \sum_{i=0}^{n-1} f(\alpha_i) (t_{i+1} - t_i)$$

où pour chaque i , $\alpha_i \in [t_i, t_{i+1}[$.

Il s'agit de l'intégrale de $\tilde{E}_{(\alpha, f, S)}$. Pour définir une telle somme, f n'a pas besoin d'être bornée.

S'il existe, avec α_i qui réalise le minimum de f sur $[t_i, t_{i+1}[$, on obtient la somme de Darboux inférieure $A^-(f, S)$.

S'il existe, avec α_i qui réalise le maximum de f sur $[t_i, t_{i+1}[$, on obtient la somme de Darboux supérieure $A^+(f, S)$.

Avec la subdivision uniforme, la somme de Riemann prend la forme classique :

$$R(f, S_{\text{unif}}) = \frac{b-a}{n} \sum_{i=0}^{n-1} f\left(a + i \frac{b-a}{n}\right).$$

De la Proposition 1.3.1 pour les fonctions en escalier, on déduit que ces sommes vérifient les propriétés élémentaires suivantes :

Proposition 1.3.2

1) $A^-(f, S) \leq R(f, S, \alpha) \leq A^+(f, S)$.

2) Si $S \subset S'$, on a $A^-(f, S) \leq A^-(f, S')$ et $A^+(f, S') \leq A^+(f, S)$.

Plus la subdivision est fine, plus les sommes de Darboux inférieures sont grandes, plus les sommes de Darboux supérieures sont petites (et se rapprochent en fait chacune d'une valeur commune ...).

Chapitre 2

Intégrale de Riemann

2.1 Fonctions Riemann-intégrables

Définition 2.1.1 Une fonction $f : [a, b] \rightarrow \mathbb{R}$ est Riemann-intégrable (sur $[a, b]$) si pour tout $\varepsilon > 0$, il existe une subdivision S telle que ses sommes de Darboux vérifient :

$$A^+(f, S) - A^-(f, S) \leq \varepsilon. \quad (2.1)$$

A fortiori si (2.1) est vraie pour S , c'est vrai pour toute subdivision S' plus fine que S .

En effet d'après la Prop. 1.3.2, pour $S \subset S'$, on a $A^+(f, S') \leq A^+(f, S)$ et $A^-(f, S') \geq A^-(f, S)$ et donc

$$A^+(f, S') - A^-(f, S') \leq A^+(f, S) - A^-(f, S).$$

Les fonctions Riemann-intégrables sont celles pour lesquelles les sommes de Darboux inférieure et supérieure peuvent être rendues arbitrairement proches à condition de choisir des subdivisions suffisamment fines (les surfaces contenantes et contenues peuvent serrer d'aussi près que voulu la vraie surface).

Proposition 2.1.1 Quand $f : [a, b] \rightarrow \mathbb{R}$ est intégrable, en prenant, les sup et inf sur l'ensemble des subdivisions de $[a, b]$, on a alors

$$\sup_S A^-(f, S) = \inf_S A^+(f, S).$$

Démonstration : En effet comme $A^-(f, S)$ est croissant quand S se raffine et est borné par les sommes de Darboux supérieures, $\sup_S A^-(f, S)$ est bien défini. De même, $\inf_S A^+(f, S)$ est bien défini. Puis d'après (2.1), il est facile de voir que pour tout $\varepsilon > 0$, on a $\inf_S A^+(f, S) - \sup_S A^-(f, S) \leq \varepsilon$. D'où l'égalité. \square

Définition 2.1.2 Par définition, l'intégrale (de Riemann) de f , Riemann-intégrable, est la valeur commune précédente

$$\int_a^b f(x)dx = \sup_S A^-(f, S) = \inf_S A^+(f, S).$$

De façon équivalente, si f est Riemann-intégrable, on a

$$\begin{aligned} \int_a^b f(x)dx &= \lim_{\rho(S) \rightarrow 0} A^-(f, S) = \lim_{\rho(S) \rightarrow 0} A^+(f, S) \\ &= \lim_{\rho(S) \rightarrow 0} R(f, S, \alpha) \end{aligned}$$

pour toute somme de Riemann $R(f, S, \alpha)$ où on rappelle que $\rho(S)$ désigne le pas de la subdivision S .

Théorème 2.1.1 *La convergence des sommes de Darboux est équivalente à celle des sommes de Riemann.*

Démonstration : • D'abord, soit f dont les sommes de Riemann convergent vers L . La fonction f est bornée et pour tout $\varepsilon > 0$, il existe une partition $S = \{a_0, \dots, a_n\}$ de $[a, b]$ telle que

$$-\varepsilon/2 \leq R(f, S, \alpha) - L \leq \varepsilon/2$$

pour toute somme de Riemann $R(f, S, \alpha)$ de pas $\rho(S)$ assez petit. Soit alors, pour chaque j , $x_j \in]a_j, a_{j+1}[$ tel que $M_j - f(x_j) \leq \varepsilon/(2(b-a))$, on a

$$\begin{aligned} A^+(f, S) - \sum_{j=0}^{n-1} f(x_j)(a_{j+1} - a_j) &= \sum_{j=0}^{n-1} (M_j - f(x_j))(a_{j+1} - a_j) \\ &= \frac{\varepsilon}{2(b-a)} \sum_{j=0}^{n-1} (a_{j+1} - a_j) \\ &\leq \varepsilon/2. \end{aligned}$$

De même, soit ensuite, pour chaque j , $y_j \in]a_j, a_{j+1}[$ tel que $m_j - f(y_j) \leq \varepsilon/(2(b-a))$, on a

$$\begin{aligned} \sum_{j=0}^{n-1} f(y_j)(a_{j+1} - a_j) - A^-(f, S) &= \sum_{j=0}^{n-1} (f(y_j) - m_j)(a_{j+1} - a_j) \\ &= \frac{\varepsilon}{2(b-a)} \sum_{j=0}^{n-1} (a_{j+1} - a_j) \\ &\leq \varepsilon/2. \end{aligned}$$

On en déduit aisément que

$$A^+(f, S) - L \leq \varepsilon, \quad L - A^-(f, S) \leq \varepsilon$$

et donc

$$A^+(f, S) - A^-(f, S) \leq 2\varepsilon.$$

Autrement dit les sommes de Darboux (inférieures et supérieures) sont arbitrairement proches et convergent vers L . La fonction f est donc Riemann intégrable (au sens de la définition de Darboux) et $L = \int_a^b f(x)dx = \sup_S A^-(f, S) = \inf_S A^+(f, S)$.

• Réciproquement, si les sommes de Darboux convergent vers $L = \inf_S A^+(f, S) = \sup_S A^-(f, S)$. Alors pour tout $\varepsilon > 0$, il existe une subdivision S telle que $A^+(f, S) - A^-(f, S) \leq \varepsilon$ et a fortiori, on a

$$L - A^-(f, S) \leq \varepsilon, \quad A^+(f, S) - L \leq \varepsilon.$$

Alors pour toute subdivision $S' \supset S$ et toute somme de Riemann $R(f, S', \alpha)$, on a

$$-\varepsilon \leq A^-(f, S) - L \leq A^-(f, S') - L \leq R(f, S', \alpha) - L \leq A^+(f, S') - L \leq A^+(f, S) - L \leq \varepsilon.$$

On a donc pour toute subdivision $S' \supset S$, $|R(f, S', \alpha) - L| \leq \varepsilon$, autrement dit les sommes de Riemann de f converge vers $L = \int_a^b f(x)dx$ (quand on raffine la subdivision). \square

Remarque 2.1.1 La fonction f est donc aussi intégrable si ses sommes de Riemann convergent. L'intégrale de f est alors aussi la limite des sommes de Riemann.

Souvent, on prend la subdivision uniforme et $\int_a^b f(x)dx$ est vue comme limite des sommes de Riemann classiques :

$$\lim_{n \rightarrow +\infty} \frac{b-a}{n} \sum_{i=0}^{n-1} f\left(a + i \frac{b-a}{n}\right) = \int_a^b f(x)dx$$

car le pas de la subdivision uniforme est $(b-a)/n$ qui tend vers 0 quand $n \rightarrow +\infty$.

Exemples : Calculer les limites des sommes de Riemann suivantes :

$$\sum_{k=1}^n \frac{k}{n^2}, \quad \sum_{k=1}^n \frac{\ln(k/n)}{n}, \quad \frac{2}{\pi} \sum_{k=1}^n \frac{\cos(\frac{k\pi}{2}n)}{n}.$$

Indication : prendre $a = 0$ et $b = 1$.

Contre-exemple : une fonction qui n'est pas Riemann-intégrable

Soit sur $[0, 1]$, la fonction

$$f(x) = \mathbf{1}_{\mathbb{Q}}(x) = \begin{cases} 0 & \text{si } x \notin \mathbb{Q} \\ 1 & \text{si } x \in \mathbb{Q}. \end{cases}$$

Soit $S = (t_i)_{1 \leq i \leq n}$ une subdivision de $[0, 1]$. Pour tout $1 \leq i \leq n$, l'intervalle (t_i, t_{i+1}) contient un rationnel α_i et un irrationnel β_i (par densité de \mathbb{Q} et de \mathbb{Q}^c dans \mathbb{R}).

On a alors $f(\alpha_i) = 1$ d'où $\sup_{x \in [t_i, t_{i+1}[} f(x) = 1$ et $f(\beta_i) = 0$ d'où $\inf_{x \in [t_i, t_{i+1}[} f(x) = 0$.

On en déduit facilement que

$$A^-(f, S) = 0 \text{ et } A^+(f, S) = 1$$

pour toute subdivision S . On ne peut donc pas vérifier la définition de la Riemann-intégrabilité : avec $\varepsilon = 1/2$, comment trouver une subdivision S telle que $1 = 1 - 0 = A^+(f, S) - A^-(f, S) \leq \varepsilon$.

La fonction f n'est pas Riemann-intégrable. On verra qu'elle est Lebesgue-intégrable, on pourra alors donner un sens à une expression du type :

$$\ll \int_{[0,1]} \mathbf{1}_{\mathbb{Q}}(x) dx \gg.$$

Définition 2.1.3 (Autre définition de la Riemann-intégrabilité) Soit $f : [a, b] \rightarrow \mathbb{R}$ alors les deux assertions suivantes sont équivalentes :

i) Pour tout $\varepsilon > 0$, il existe deux fonctions en escalier φ_ε et ϕ_ε telles que

$$|f - \varphi_\varepsilon| \leq \phi_\varepsilon, \quad \int_a^b \phi_\varepsilon(x) dx \leq \varepsilon.$$

ii) Il existe une suite de fonctions en escalier $(\varphi_n)_n$ et une suite $(\phi_n)_n$ telles que pour tout $n \in \mathbb{N}$,

$$|f - \varphi_n| \leq \phi_n, \quad \lim_{n \rightarrow +\infty} \int_a^b \phi_n(x) dx = 0.$$

Toute application $f : [a, b] \rightarrow \mathbb{R}$ vérifiant ces deux assertions est Riemann-intégrable et l'intégrale de Riemann de f est donnée par

$$\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} \int_a^b \varphi_n(x) dx.$$

Remarque : • La définition précédente de $\int_a^b f(x) dx$ a bien un sens : on vérifie qu'elle ne dépend pas du choix des suites $(\varphi_n)_n$ et $(\phi_n)_n$.

• Bien sûr, les deux formulations de cette nouvelle définition sont équivalentes : prendre $\varepsilon = 1/n$ pour déduire la 2ème formulation de la 1ère, ou prendre n assez grand, pour déduire la 1ère formulation de la 2ème.

• Dans cette définition, f n'a pas besoin (a priori) d'être bornée contrairement à la définition via les sommes de Darboux.

Théorème 2.1.2 Lorsque f est bornée, les deux définitions de Riemann-intégrabilité coïncident.

Démonstration : • Si f , bornée, est Riemann-intégrable alors pour S de pas assez petit, on a

$$E^-(f, S) \leq f \leq E^+(f, S), \quad \text{et} \quad A^+(f, S) - A^-(f, S) \leq \varepsilon. \quad (2.2)$$

Notons alors

$$\varphi_\varepsilon = \frac{E^+(f, S) + E^-(f, S)}{2}, \quad \phi_\varepsilon = \frac{E^+(f, S) - E^-(f, S)}{2}.$$

Ce sont bien des fonctions en escalier car combinaisons linéaires de telles fonctions. De plus, on réécrit alors (2.2) sous la forme

$$\varphi_\varepsilon - \phi_\varepsilon \leq f \leq \varphi_\varepsilon + \phi_\varepsilon, \quad \text{et} \quad \int_a^b \phi_\varepsilon(x) dx \leq \varepsilon/2.$$

La définition 2.1.3 précédente est donc vérifiée avec $\varepsilon/2$. Cela justifie le sens direct.

• Soit f vérifiant la définition précédente. Notons $(\varphi_n)_n$ et $(\phi_n)_n$ les suites en escalier données par la définition. On a

$$\varphi_n - \phi_n \leq f \leq \varphi_n + \phi_n$$

Soit S_n une subdivision adaptée à la fois à φ_n et à ϕ_n . En étudiant $\varphi_n - \phi_n, f, \varphi_n + \phi_n$, sur un intervalle $[a_i^n, a_{i+1}^n]$ de la subdivision S_n , on constate que

$$\varphi_n - \phi_n \leq E^-(f, S_n) \leq f \leq E^+(f, S_n) \leq \varphi_n + \phi_n.$$

Puis

$$A^+(f, S_n) - A^-(f, S_n) = \int_a^b E^+(f, S_n)(x) - E^-(f, S_n)(x) dx \leq 2 \int_a^b \phi_n(x) dx.$$

On déduit $\lim_{\rho(S) \rightarrow 0} A^+(f, S) - A^-(f, S) = 0$, et donc f est Riemann-intégrable.

Montrons enfin que les deux définitions de $\int_a^b f(x) dx$ coïncident. D'après la première partie de la preuve, on peut choisir dans la deuxième définition les suites de fonctions en escalier

$$\varphi_n = \frac{E^+(f, S_n) + E^-(f, S_n)}{2}, \quad \phi_n = \frac{E^+(f, S_n) - E^-(f, S_n)}{2}$$

et donc $\int_a^b \varphi_n(x) dx = \frac{1}{2}(A^+(f, S_n) + A^-(f, S_n))$. Le membre de gauche converge vers $\int_a^b f(x) dx$ au sens de la deuxième définition tandis que le membre de droite converge vers $\frac{1}{2}(\int_a^b f(x) dx + \int_a^b f(x) dx)$ au sens de la première définition. D'où l'égalité. \square

2.2 Propriétés de l'intégrale de Riemann

On donne ici les propriétés principales de l'intégrale de Riemann. On se contente d'esquisser les preuves. On renvoie si nécessaire à tout manuel de Deug ou de L2.

Proposition 2.2.1 *Toute fonction Riemann-intégrable sur un intervalle $[a, b]$ est bornée.*

Démonstration : Soit, on passe par les sommes de Darboux, et alors implicitement la fonction est supposée bornée.

Soit on passe par la définition 2.1.3 alternative et alors avec $\varepsilon = 1$, il existe φ_1 et ϕ_1 telle que

$$|f - \varphi_1| \leq \phi_1, \quad \int_a^b \phi_1(x) dx \leq 1.$$

On a donc $|f| \leq |\varphi_1| + \phi_1$. Mais comme les fonctions en escalier φ_1 et ϕ_1 sont (par nature) bornées, f l'est aussi. \square

La proposition suivante garantit que la notion d'intégrale de Riemann a bien un intérêt : les fonctions usuelles le sont !

Proposition 2.2.2 (Exemples de fonctions Riemann-intégrables)

- 1) Les fonctions continues sont Riemann-intégrables.
- 2) Les fonctions monotones (croissantes ou décroissantes) sont Riemann-intégrables.

Démonstration :

1) Si f est continue sur $[a, b]$ alors elle est uniformément continue (théorème de Heine, cf. Topologie). Soit pour $\varepsilon > 0$ donné $\alpha > 0$ donné par l'uniforme continuité de f . On prend une subdivision S de pas $\rho(S) \leq \alpha$. Pour tout $x, y \in [a_i, a_{i+1}]$, on a $|x - y| \leq \alpha$ et donc $|f(x) - f(y)| \leq \varepsilon$.

On déduit alors facilement que

$$\begin{aligned} A^+(f, S) - A^-(f, S) &= \sum_{i=0}^{n-1} M_i(t_{i+1} - t_i) - \sum_{i=0}^{n-1} m_i(t_{i+1} - t_i) = \sum_{i=0}^{n-1} (M_i - m_i)(t_{i+1} - t_i) \\ &= \sum_{i=0}^{n-1} (f(b_i) - f(a_i))(t_{i+1} - t_i) \leq \sum_{i=0}^{n-1} \varepsilon(t_{i+1} - t_i) = (b - a)\varepsilon \end{aligned}$$

où $a_i, b_i \in [t_i, t_{i+1}]$ réalisent le sup et l'inf de f sur $[t_i, t_{i+1}]$, puis on a utilisé $|b_i - a_i| \leq t_{i+1} - t_i \leq \rho(S) \leq \alpha$.

Les sommes de Darboux peuvent donc être rendues arbitrairement proches.

2) Soit f croissante alors avec les notations des sommes de Darboux, on a $m_i = f(a_i)$ et $M_i = f(a_{i+1})$. En utilisant les subdivisions uniformes de pas $1/n$, on a $a_i = a + i(b - a)/n$, si bien que

$$A^+(f, S) - A^-(f, S) = \frac{b - a}{n} (f(b) - f(a)) \leq \varepsilon$$

dés que n est assez grand. \square

Systématiquement, pour prouver les propriétés suivantes des fonctions Riemann-intégrables, on se ramène par approximation aux sommes de Darboux, intégrales de fonction en escalier pour lesquelles ces propriétés ont déjà été vues.

Proposition 2.2.3 (Linéarité) *Si f et g sont Riemann-intégrables alors les combinaisons linéaires le sont aussi et pour tout α, β réels*

$$\int_a^b \alpha f(x) + \beta g(x) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx.$$

Autrement dit, l'ensemble des fonctions Riemann-intégrables est un espace vectoriel sur lequel l'intégrale définit une application linéaire.

Démonstration : Soient $(\varphi_n)_n$ et $(\phi_n)_n$ les suites de fonctions en escalier données par la définition 2.1.3 de la Riemann-intégrabilité de f . Puis soient $(\tilde{\varphi}_n)_n$ et $(\tilde{\phi}_n)_n$ celles associées à g . Alors

$$\begin{aligned} |(\alpha f + \beta g) - (\alpha \varphi + \beta \tilde{\varphi})| &\leq |\alpha| |f - \varphi| + |\beta| |g - \tilde{\varphi}| \\ &\leq |\alpha| \phi_n + |\beta| \tilde{\phi}_n \end{aligned}$$

puis

$$\int_a^b (|\alpha| \phi_n(x) + |\beta| \tilde{\phi}_n(x)) dx = |\alpha| \int_a^b \phi_n(x) dx + |\beta| \int_a^b \tilde{\phi}_n(x) dx \rightarrow 0, \quad n \rightarrow +\infty$$

La fonction $\alpha f + \beta g$ vérifie donc la définition 2.1.3 de la Riemann-intégrabilité avec les suites de fonctions en escalier $(\alpha \varphi_n + \beta \tilde{\varphi}_n)_n$ et $(|\alpha| \phi_n + |\beta| \tilde{\phi}_n)_n$, cette dernière étant positive.

Puis

$$\begin{aligned} \int_a^b (\alpha f(x) + \beta g(x)) dx &= \lim_{n \rightarrow +\infty} \int_a^b (\alpha \varphi_n(x) + \beta \tilde{\varphi}_n(x)) dx \\ &= \alpha \lim_{n \rightarrow +\infty} \int_a^b \varphi_n(x) dx + \beta \lim_{n \rightarrow +\infty} \int_a^b \tilde{\varphi}_n(x) dx \\ &= \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx. \end{aligned}$$

□

Proposition 2.2.4 (Relation de Chasles) *Si f est Riemann-intégrable sur $[a, b]$ et $c \in]a, b[$ alors f l'est encore sur $[a, c]$ et sur $[c, b]$ et*

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Démonstration : D'abord si f est Riemann-intégrable sur $[a, b]$, elle l'est sur $[a, c]$ et $[c, b]$. En effet soit pour $\varepsilon > 0$, S la subdivision donnée par la définition de l'intégrabilité sur $[a, b]$. Quitte à raffiner encore S , on peut supposer que $c \in S$. Mais alors en notant S_1

la partie de la subdivision correspondant à $[a, c]$ et S_2 celle à $[c, b]$, on déduit que S_1 et S_2 vérifient la définition de l'intégrabilité de f sur $[a, c]$ et $[c, b]$ car il est facile de voir que

$$A^\pm(f, S) = A^\pm(f, S_1) + A^\pm(f, S_2).$$

On a donc

$$A^+(f, S_1) - A^-(f, S_1) + A^+(f, S_2) - A^-(f, S_2) \leq A^+(f, S) - A^-(f, S)$$

et donc $A^+(f, S_1) - A^-(f, S_1)$ et $A^+(f, S_2) - A^-(f, S_2)$ peuvent être rendus arbitrairement proche. Puis, on a

$$\begin{aligned} \int_a^b f(x)dx &= \lim_{\rho(S) \rightarrow 0} A^+(f, S) = \lim_{\rho(S) \rightarrow 0} A^+(f, S_1) + A^+(f, S_2) \\ &= \int_a^c f(x)dx + \int_c^b f(x)dx. \end{aligned}$$

□

Proposition 2.2.5 (Positivité) *Si f est positive et Riemann-intégrable sur $[a, b]$ alors*

$$\int_a^b f(x)dx \geq 0.$$

Démonstration : En effet, si f est positive alors ses sommes de Darboux le sont aussi. Mais alors comme $\int_a^b f(x)dx$ est limite de ses sommes (quand le pas des subdivisions tend vers 0), l'intégrale est positive. □

Corollaire 2.2.1 *Si f et g sont Riemann-intégrables et $f \leq g$ alors*

$$\int_a^b f(x)dx \leq \int_a^b g(x)dx$$

C'est la propriété précédente appliquée à $g - f$, fonction positive.

Proposition 2.2.6 (Intégrale et valeurs absolues) *Si f est Riemann-intégrable*

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|dx.$$

En effet la Riemann-intégrabilité de $|f|$ est facile à justifier (exercice), puis :

$$\begin{aligned} \left| \int_a^b f(x)dx \right| &= \left| \lim_{\rho(S) \rightarrow 0} A^+(f, S) \right| = \lim_{\rho(S) \rightarrow 0} |A^+(f, S)| \\ &\leq \lim_{\rho(S) \rightarrow 0} A^+(|f|, S) = \int_a^b |f(x)|dx. \end{aligned}$$

Proposition 2.2.7 Soit f une fonction Riemann-intégrable de $[a, b]$ dans \mathbb{R} et $u : \mathbb{R} \rightarrow \mathbb{R}$ une application linéaire. Alors $u \circ f$ est Riemann-intégrable et

$$\int_a^b u \circ f(x) dx = u \left(\int_a^b f(x) dx \right).$$

Démonstration : Notons M la borne de l'application linéaire u . Soient $(\varphi_n)_n$ et $(\phi_n)_n$ les suites d'applications en escalier données par la définition 2.1.3 de la Riemann-intégrabilité de f . Alors $u \circ \varphi_n$ et $M\phi_n$ sont aussi en escalier. Puis $\lim_{n \rightarrow +\infty} \int_a^b M\phi_n(x) dx = 0$ et

$$|u \circ f(t) - u \circ \varphi_n(t)| = |u \circ (f - \varphi_n)(t)| \leq M|(f - \varphi_n)(t)| \leq M\phi_n(t)$$

et donc $|u \circ f - u \circ \varphi_n| \leq M\phi_n$, si bien que la définition (alternative) de la Riemann-intégrabilité est vérifiée pour $u \circ f$ avec les fonctions en escalier $(u \circ \varphi_n)_n$ et $(M\phi_n)_n$. \square

Proposition 2.2.8 Soit f Riemann-intégrable sur $[a, b]$ positive et telle que $\int_a^b f(x) dx = 0$. Alors f prend la valeur 0 en tout point où elle est continue.

Démonstration : Soit t_0 un point de continuité de f . Si $f(t_0) \neq 0$ alors par continuité de f en t_0 , on peut trouver un intervalle I de longueur $l(I) > 0$ où $f(t) \geq f(t_0)/2 > 0$. Soit alors g la fonction en escalier égale à $f(t_0)/2$ sur I et 0 ailleurs. Alors $f \geq g$ et donc

$$\int_a^b f(x) dx \geq \int_a^b g(x) dx = l(I)f(t_0)/2 > 0$$

ce qui est absurde. On a donc bien $f(t_0) = 0$. \square

Corollaire 2.2.2 Si f et g sont Riemann-intégrables et continues avec $f \leq g$ alors $\int_a^b f(x) dx = \int_a^b g(x) dx$ entraîne $f = g$.

Proposition 2.2.9 (Première formule de la moyenne) Soient f et g deux fonctions Riemann-intégrables sur $[a, b]$, g étant une fonction positive. On désigne par m (resp. M) la borne inférieure (resp. supérieure) de f sur $[a, b]$. Alors il existe $m \leq k \leq M$ tel que

$$\int_a^b f(x)g(x) dx = k \int_a^b g(x) dx.$$

De plus, si f est continue alors il existe $c \in [a, b]$ tel que

$$\int_a^b f(x)g(x) dx = f(c) \int_a^b g(x) dx.$$

Démonstration : Comme $g \geq 0$, on a $mg \leq fg \leq Mg$ et donc

$$m \int_a^b g(x)dx \leq \int_a^b f(x)g(x)dx \leq M \int_a^b g(x)dx.$$

Si $\int_a^b g(x)dx = 0$, alors l'encadrement précédent assure que $\int_a^b f(x)g(x)dx = 0$ et $k = \frac{m+M}{2}$ convient par exemple.

Sinon, $\int_a^b f(x)g(x)dx / \int_a^b g(x)dx$ est un élément de $[m, M]$. Si en plus f est continue, on conclut par le théorème des valeurs intermédiaires pour trouver c . \square

Définition 2.2.1 Le réel $\frac{1}{b-a} \int_a^b f(x)dx$ est dit valeur moyenne de la fonction Riemann-intégrable f sur $[a, b]$.

(C'est l'espérance (probabiliste) de f pour la loi uniforme sur $[a, b]$.)

Proposition 2.2.10 (Deuxième formule de la moyenne) Soit f une fonction positive décroissante de $[a, b]$ dans \mathbb{R} et g une application Riemann-intégrable sur $[a, b]$. Alors, il existe $c \in [a, b]$ tel que

$$\int_a^b f(x)g(x)dx = f(c) \int_a^b g(x)dx.$$

(exercice difficile)

2.3 Intégrale et primitive

Proposition 2.3.1 Soit $f : [a, b] \rightarrow \mathbb{R}$ Riemann-intégrable, alors l'application intégrale $t \mapsto \int_a^t f(x)dx$ est continue.

Démonstration : Soit $t_0 \in [a, b]$ et $[\alpha, \beta] \subset [a, b]$ un voisinage de t_0 . La fonction f est intégrable et bornée par M sur $[\alpha, \beta]$. Par la relation de Chasles, on a pour tout $t \in [\alpha, \beta]$

$$\left| \int_a^t f(x)dx - \int_a^{t_0} f(x)dx \right| = \left| \int_t^{t_0} f(x)dx \right| \leq M|t - t_0|.$$

La restriction de $t \mapsto \int_a^t f(x)dx$ à $[\alpha, \beta]$ est donc lipschitzienne et en particulier elle est continue en t_0 . \square

Proposition 2.3.2 Si f admet une limite (resp. à droite, à gauche) en $t_0 \in [a, b]$ alors l'application intégrale $t \mapsto \int_a^t f(x)dx$ est dérivable (resp. à droite, à gauche) en t_0 de dérivée $l = \lim_{t \rightarrow t_0} f(t)$ (resp. $f(t_0^+)$, $f(t_0^-)$).

Démonstration : D'après l'existence de la limite l en t_0 , pour $\varepsilon > 0$ assez petit, il existe α tel que si $t_0 - \alpha \leq t \leq t_0 + \alpha$, alors $|f(t) - l| \leq \varepsilon$. Mais alors pour tout $t \in [t_0 - \alpha, t_0 + \alpha]$, on a (pour $t \geq t_0$) :

$$\left| \int_a^t f(x)dx - \int_a^{t_0} f(x)dx - (t - t_0)l \right| \left| \int_{t_0}^t f(x)dx - \int_{t_0}^t ldx \right| \leq \left| \int_{t_0}^t |f(x) - l|dx \right| \leq \varepsilon|t - t_0|.$$

Il suit

$$\left| \frac{\int_a^t f(x)dx - \int_a^{t_0} f(x)dx}{t - t_0} - l \right| \leq \varepsilon,$$

ce qui conclut d'après la définition de la dérivée comme limite des taux d'accroissement. \square

Corollaire 2.3.1 *L'application intégrale $t \mapsto \int_a^t f(x)dx$ est dérivable en tout point $t_0 \in [a, b]$ en lequel f est continue, de dérivée $F'(t_0) = f(t_0)$.*

Définition 2.3.1 *Soit $f : [a, b] \rightarrow \mathbb{R}$. On appelle primitive de f toute application $F : [a, b] \rightarrow \mathbb{R}$ qui est dérivable sur $[a, b]$ et de dérivée $F' = f$.*

Proposition 2.3.3 *Soit $f : [a, b] \rightarrow \mathbb{R}$ admettant F pour primitive sur $[a, b]$. Alors f admet une infinité de primitives sur $[a, b]$ qui sont les applications de la forme $G(x) = F(x) + k$ où $k \in \mathbb{R}$ est une constante. Par contre, étant donné $c \in [a, b]$ et $\alpha \in \mathbb{R}$, f admet une unique primitive $F_{c,\alpha}$ qui vaut α en c , il s'agit de $F_c(x) = F(x) - F(c) + \alpha$.*

En effet, $G(x) = F(x) + k$ est de dérivée $G' = F' = f$ donc il s'agit bien d'une primitive de f .

Puis si G est une primitive de f alors $(F - G)' = F' - G' = f - f = 0$. La fonction $F - G$ est donc constante. On a donc $G(x) = F(x) + k$ pour une certaine constante k .

Proposition 2.3.4 *Soit f une application continue sur $[a, b]$. Alors pour tout point $c \in [a, b]$, l'application intégrale $t \mapsto \int_c^t f(x)dx$ est une primitive de f sur $[a, b]$. Il s'agit de l'unique primitive de f qui s'annule en c .*

L'ensemble des primitives d'une fonction f continue sur $[a, b]$ est donc de la forme $t \mapsto \int_c^t f(x)dx + k$ pour tout $c \in [a, b]$ et toute constante $k \in \mathbb{R}$.

Remarque 2.3.1 Si f est intégrable, l'application intégrale est seulement continue. A priori, elle n'est pas dérivable et il ne s'agit pas alors d'une primitive de f .

On a vu que les applications intégrales sont exactement les primitives des fonctions continues. Si f n'est pas continue, a priori, les applications intégrales ne sont pas des primitives. Cependant si f admet des primitives elles sont forcément données par les applications intégrales. En effet, on a :

Théorème 2.3.1 (Théorème fondamental du calcul différentiel) Soit f une fonction Riemann-intégrable sur $[a, b]$. Si F est une primitive de f alors

$$\int_a^b f(x)dx = F(b) - F(a).$$

Autrement dit, $x \mapsto \int_a^x f(t)dt$ est la primitive de f nulle en a . On a donc

$$\left(\int_a^x f(t)dt \right)' = f(x).$$

Ce résultat relie deux types d'opération a priori sans rapport : l'intégrale qui via le calcul d'aire est une notion d'origine géométrique et la dérivation qui est une notion analytique.

Le théorème indique donc que ces deux opérations sont inverses.

Démonstration : Si f est continue, le résultat est vraie d'après la Prop. 2.3.4. De façon générale, si f est seulement Riemann-intégrable, montrons pour $\varepsilon > 0$ fixé que

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| \leq \varepsilon.$$

D'après la définition de la Riemann-intégrabilité de f , il existe une subdivision S telle que pour les sommes de Darboux correspondantes, on a

$$A^+(f, S) - A^-(f, S) \leq \varepsilon$$

Soit

$$g_1 = \frac{E^+(f, S) + E^-(f, S)}{2}, \quad g_2 = \frac{E^+(f, S) - E^-(f, S)}{2}.$$

On a

$$\int_a^b g_2(x)dx \leq \varepsilon/2 \quad \text{puis} \quad |f - g_1| \leq g_2$$

car $E^-(f, S) \leq f \leq E^+(f, S)$. Par ailleurs, g_1, g_2 sont en escalier et valent respectivement $\frac{m_i + M_i}{2}$ et $\frac{M_i - m_i}{2}$ sur chaque $[a_i, a_{i+1}]$. Considérons alors la fonction G qui vaut $F(t) - \frac{m_i + M_i}{2}t$ sur chaque intervalle $[a_i, a_{i+1}]$. La dérivée $G' = f - \frac{m_i + M_i}{2}$ est majorée en valeurs absolues sur $[a_i, a_{i+1}]$ par $\frac{M_i - m_i}{2}$. Par le théorème des accroissements finis, on a

$$\left| F(a_{i+1}) - F(a_i) - \frac{M_i + m_i}{2}(a_{i+1} - a_i) \right| = |G(a_{i+1}) - G(a_i)| \leq \frac{M_i - m_i}{2}(a_{i+1} - a_i)$$

et donc en sommant sur chaque intervalle de la subdivision S

$$|F(b) - F(a) - \int_a^b g_1(x)dx| \leq \sum_{i=0}^{n-1} \frac{M_i - m_i}{2}(a_{i+1} - a_i) = \int_a^b g_2(x)dx \leq \varepsilon/2.$$

Par ailleurs, on a

$$\left| \int_a^b f(x)dx - \int_a^b g_1(x)dx \right| \leq \int_a^b |f(x) - g_1(x)|dx \leq \int_a^b g_2(x)dx \leq \varepsilon/2.$$

Si bien qu'il vient

$$\left| F(b) - F(a) - \int_a^b f(x)dx \right| \leq \varepsilon,$$

ce qui conclut en faisant tendre ε vers 0. □

Remarque 2.3.2 • Quand f est positive, l'intégrale $\int_a^b f(x)dx$ s'interprète comme l'aire de la portion de plan $\{(x, y) \mid 0 \leq y \leq f(x), a \leq x \leq b\}$.

- Si f est de signe quelconque, $\int_a^b f(x)dx$ est l'aire algébrique où les parties sous l'axe des abscisses sont d'aires négatives et celles au dessus de l'axe des abscisses sont positives.
- Parfois un calcul d'aire est une méthode efficace de calcul d'intégrale.
- Méthodes de calcul : IPP, changements de variable.

Chapitre 3

Fonctions réglées

3.1 Définition

Soit I un intervalle, notons $B(I, \mathbb{R})$ l'ensemble des fonctions bornées de I dans \mathbb{R} (on pourrait remplacer \mathbb{R} par E un espace vectoriel normé quelconque). L'ensemble $B(I, \mathbb{R})$ est un espace vectoriel, on le norme en définissant $\|\cdot\|_\infty$ sur $B(I, \mathbb{R})$:

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in I\}.$$

On montre que :

Proposition 3.1.1 *L'espace $B(I, \mathbb{R})$ est complet pour la norme $\|\cdot\|_\infty$.*

Notons que l'ensemble $\mathcal{E}(I, \mathbb{R})$ des fonctions en escalier sont dans $B(I, \mathbb{R})$. On peut alors définir :

Définition 3.1.1 *On appelle fonction réglée tout élément g dans l'adhérence de $\mathcal{E}(I, \mathbb{R})$, les fonctions en escalier sur I dans $B(I, \mathbb{R})$ pour $\|\cdot\|_\infty$. On note $R(I, \mathbb{R})$ cet ensemble :*

$$R(I, \mathbb{R}) = \overline{\mathcal{E}(I, \mathbb{R})}.$$

Par continuité pour $\|\cdot\|_\infty$ de $+$ et de la multiplication par un scalaire il est facile de voir que l'ensemble des fonctions réglées est un espace vectoriel.

En pratique : une fonction f est réglée sur I , ssi il existe une suite $(g_n)_n$ de fonctions en escalier qui convergent uniformément sur I vers f , ou encore

$$\forall \varepsilon > 0, \exists g \text{ en escalier, telle que } \|f - g\|_\infty = \sup_{x \in I} |f(x) - g(x)| \leq \varepsilon.$$

Proposition 3.1.2 *Une limite uniforme de fonctions réglées est réglée.*

Démonstration : Soit f limite uniforme de $(f_n)_n$, suite de fonctions réglées. Comme $f_n \in R(I, \mathbb{R})$, f est dans l'adhérence de $R(I, \mathbb{R})$ pour la norme $\|\cdot\|_\infty$. Mais par définition, cet espace est fermé, il est donc égale à son adhérence :

$$f \in \overline{R(I, \mathbb{R})} = R(I, \mathbb{R}).$$

La fonction f est donc réglée. □

3.2 Propriétés des fonctions réglées

Proposition 3.2.1 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction réglée. Alors l'ensemble de ses points de discontinuité est au plus dénombrable.

Pour être Riemann-intégrable, il faut (et il suffit) que cet ensemble soit négligeable, c'est à dire pour n'importe quel $\varepsilon > 0$, qu'il soit inclus dans un ensemble de longueur totale inférieure à ε , ce qui est forcément le cas pour les ensembles dénombrables (exercice).

Proposition 3.2.2 Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction réglée. Alors en tout point ses limites à gauche et à droite existent :

$$\lim_{x \rightarrow x_0^+} f(x) \text{ existe pour } x_0 \in [a, b[, \quad \lim_{x \rightarrow x_0^-} f(x) \text{ existe pour } x_0 \in]a, b].$$

La réciproque est vraie (car \mathbb{R} est complet).

Conséquences :

- Soit

$$f(x) = \cos(1/x) \text{ si } x \in]0, 1] \text{ et } f(0) = 0. \quad (3.1)$$

D'après la Prop. 3.2.2, cette fonction n'est pas réglée (elle n'a pas de limite à droite en 0) mais elle est Riemann-intégrable (car l'ensemble de ses points de discontinuité est négligeable).

- Les fonctions continues sont réglées.
- Les fonction monotones sont réglées.

3.3 Intégrale des fonctions réglées

Pour une fonction en escalier $f(x) = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}$, on définit son intégrale en tant que fonction réglée par

$$\theta_{a,b}(f) = \sum_{i=0}^{n-1} \alpha_i (t_{i+1} - t_i). \quad (3.2)$$

On a

$$|\theta_{a,b}(f)| \leq \sum_{i=0}^{n-1} (t_{i+1} - t_i) \|f\|_\infty \leq \|f\|_\infty \sum_{i=0}^{n-1} (t_{i+1} - t_i) = (b - a) \|f\|_\infty.$$

L'application $\theta_{a,b}$ est donc linéaire et lipschitzienne sur $\mathcal{E}(I, \mathbb{R})$:

$$|\theta_{a,b}(f)| \leq (b - a) \|f\|_\infty. \quad (3.3)$$

On va pouvoir étendre cette intégrale $\theta_{a,b}$ à toutes les fonctions réglées grâce au résultat suivant d'extension (cf cours de Topologie) :

Théorème 3.3.1 (Extension des applications uniformément continues) Soient E et F deux espaces métriques, F complet et X une partie dense de E ($\bar{X} = E$).

On considère $f : X \rightarrow F$ une application uniformément continue sur X . Alors il existe une unique extension g uniformément continue définie sur E qui prolonge f (i.e. $g(x) = f(x)$ si $x \in X$).

On applique ce résultat à

$$\theta_{a,b} : \begin{cases} \mathcal{E}(I, \mathbb{R}) & \rightarrow \mathbb{R} \\ f & \mapsto \int_a^b f(x)dx. \end{cases}$$

C'est une application uniformément continue (car lipschitzienne d'après (3.3)). Elle est définie sur $\mathcal{E}(I, \mathbb{R})$ qui est dense dans $R(I, \mathbb{R})$ pour $\|\cdot\|_\infty$ d'après la définition de l'ensemble des fonctions réglées. On étend alors $\theta_{a,b}$ en une application (toujours notée) $\theta_{a,b}$ sur $R(I, \mathbb{R})$.

C'est l'intégrale d'une fonction réglée.

Pour une fonction f réglée, il existe g_n une suite de fonctions en escalier qui converge uniformément vers f et par définition l'intégrale de f (en tant que fonction réglée) est

$$\theta_{a,b}(f) = \lim_{n \rightarrow +\infty} \int_a^b g_n(x)dx.$$

où le membre de droite est bien définie par (3.2) car g_n est en escalier. De plus cette définition de $\theta_{a,b}(f)$ ne dépend pas de la suite g_n en escalier qui converge vers f , si bien que la définition a bien un sens.

Cette nouvelle notion d'intégrabilité coïncide en fait avec l'intégrabilité de Riemann, en effet

Théorème 3.3.2 Si $f : [a, b] \rightarrow \mathbb{R}$ est une fonction réglée alors f est Riemann-intégrable. De plus, son intégrale de Riemann et son intégrale en tant que fonction réglée coïncident :

$$\theta_{a,b}(f) = \int_a^b f(x)dx.$$

Démonstration : Si f est en escalier, il est clair que $\theta_{a,b}(f) = \int_a^b f(t)dt$.

Si f est réglée, il existe f_n en escalier telle que $f_n \rightarrow f$ uniformément. Mais comme $\theta_{a,b}(f_n) = \int_a^b f_n(t)dt$, l'égalité est conservée en passant à la limite. \square

Attention : la réciproque est fautive : il existe des fonctions Riemann-intégrables qui ne sont pas réglées. Par exemple, la fonction donnée en (3.1) n'est pas réglée mais est Riemann-intégrable.

Chapitre 4

Intégrales impropres

Dans ce chapitre, $[a, b[$ désigne un intervalle ouvert en b , éventuellement infini ($[a, +\infty[$, $] -\infty, b]$, \mathbb{R}). On s'intéresse aux intégrales de fonctions définies sur $[a, b[$ mais pas en b . Il se pose alors un problème de convergence pour l'intégrale $\int_a^b f(t)dt$ en b .

4.1 Définition et propriétés

Définition 4.1.1 Une fonction de $[a, b[\rightarrow \mathbb{R}$ est dite localement intégrable si elle est Riemann-intégrable sur tout sous-intervalle compact $[c, d]$ de $[a, b[$.

En pratique, la locale intégrabilité sera donnée par

Proposition 4.1.1 Une fonction $f : [a, b[\rightarrow \mathbb{R}$ continue est localement intégrable.

Démonstration : Soit $[c, d] \subset [a, b[$. La fonction f est continue sur le compact $[c, d]$ donc Riemann-intégrable sur $[c, d]$. Comme c'est vrai pour tout $[c, d] \subset [a, b[$, elle est localement intégrable. \square

Nous notons alors $F(x) = \int_a^x f(t)dt$ l'application intégrale de f sur $[a, b[$. Elle est bien définie car f est Riemann-intégrable sur $[a, x]$.

Souvent dans ce genre de situation, la fonction f n'est pas définie en b . C'est pourquoi, on ne peut pas étudier directement l'intégrale $\int_a^b f(t)dt$. Cependant, on peut donner un sens généralisé à l'intégrale sur $[a, b[$:

Définition 4.1.2 • Si la fonction intégrale F admet une limite I en b , on dit que l'intégrale impropre $\int_a^b f(t)dt$ est convergente. On attribue alors à cette intégrale dite impropre la valeur I .

• Si F n'a pas de limite en b , on dit que l'intégrale impropre $\int_a^b f(t)dt$ est divergente.

On parle parfois d'intégrales généralisées plutôt que d'intégrales impropres.

Exemples.

- La fonction $f(x) = e^{-x}$ est localement intégrable sur $[0, +\infty[$. L'intégrale impropre $\int_0^{+\infty} f(t)dt$ converge et vaut 1.
- La fonction $f(x) = 1/\sqrt{1-x^2}$ est localement intégrable sur $[0, 1[$. L'intégrale impropre $\int_0^{+\infty} f(t)dt$ converge et vaut $\pi/2$.
- La fonction $f(x) = \cos x$ est localement intégrable sur $[0, +\infty[$. Mais l'intégrale impropre $\int_0^{+\infty} f(t)dt$ diverge. De même pour $\int_0^{+\infty} \sin t dt$.
- Toute fonction P polynôme est localement intégrable sur $[0, +\infty[$, mais d'intégrale impropre divergente sur $[0, +\infty[$.
- La fonction $f(x) = 1/x^2$ est localement intégrable sur $]0, 1]$ et sur $[1, +\infty[$. Son intégrale impropre est convergente sur $[1, +\infty[$, elle est divergente sur $]0, 1]$.
- La fonction $f(x) = 1/\sqrt{x}$ est localement intégrable sur $]0, 1]$ et sur $[1, +\infty[$. Son intégrale impropre est convergente sur $]0, 1]$, elle est divergente sur $[1, +\infty[$.

Proposition 4.1.2 *L'ensemble des fonctions de $[a, b[$ dans \mathbb{R} dont l'intégrale impropre sur $[a, b[$ converge est un sous-espace vectoriel de $\mathcal{F}([a, b[, \mathbb{R})$. L'intégration $f \mapsto \int_a^b f(t)dt$ est une application linéaire sur ce sous-espace vectoriel.*

Proposition 4.1.3 *Soit $f : [a, b[\rightarrow \mathbb{R}$ d'intégrale impropre convergente et $u : \mathbb{R} \rightarrow \mathbb{R}$ une application linéaire continue. Alors $u \circ f : [a, b[\rightarrow \mathbb{R}$ a une intégrale impropre convergente avec*

$$\int_a^b (u \circ f)(t)dt = u \left(\int_a^b f(t)dt \right).$$

Dans les deux cas, il s'agit juste d'un simple passage à la limite des propriétés analogues déjà vues pour les intégrales classiques. C'est donc la linéarité du passage à la limite qui prouve ces résultats.

Proposition 4.1.4 (Critère de Cauchy) *Soit $f : [a, b[\rightarrow \mathbb{R}$ une application localement intégrable. L'intégrale impropre de f sur $[a, b[$ converge ssi f vérifie la condition suivante (dite critère de Cauchy) : Pour tout $\varepsilon > 0$, il existe $c \in [a, b[$ tel que*

$$\forall x', x'' \in [c, b[, \quad \left| \int_{x'}^{x''} f(t)dt \right| \leq \varepsilon.$$

Il s'agit du critère général d'existence d'une limite pour les fonctions à valeurs dans un espace complet appliquée à l'application intégrale F à valeurs dans \mathbb{R} , espace complet (cf cours de Topologie ou de Compléments d'Analyse).

Définition 4.1.3 *Soit $f : [a, b[\rightarrow \mathbb{R}$ une application localement intégrable. L'intégrale impropre $\int_a^b f(t)dt$ de f est dite absolument convergente si l'intégrale impropre de $|f|$ converge.*

Par exemple, l'intégrale $\int_0^{+\infty} \cos t e^{-t} dt$ converge absolument car $|\cos t e^{-t}| \leq e^{-t}$ est d'intégrale convergente sur $[0, +\infty[$.

Proposition 4.1.5 *Soit $f : [a, b[\rightarrow \mathbb{R}$ une application localement intégrable. Une condition suffisante pour que l'intégrale impropre de f sur $[a, b[$ converge est qu'elle converge absolument sur $[a, b[$.*

C'est immédiat par le critère de Cauchy : celui de $|f|$ donne celui de f .

Par un passage à la limite, on a aussi pour les intégrales impropres absolument convergentes l'inégalité :

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt.$$

Définition 4.1.4 *Une intégrale impropre convergente mais non absolument convergente est dite semi-convergente.*

Exemple : L'intégrale $\int_0^{+\infty} \frac{\sin t}{t} dt$ est semi-convergente. En effet, d'abord il n'y a pas convergence absolue car

$$\int_0^{+\infty} \left| \frac{\sin t}{t} \right| dt \geq \sum_{k=0}^{+\infty} \int_{\frac{\pi}{6} + 2k\pi}^{\frac{5\pi}{6} + 2k\pi} \left| \frac{\sin t}{t} \right| dt \geq \frac{1}{2} \sum_{k=0}^{+\infty} \frac{4\pi}{6} \frac{1}{\frac{\pi}{6} + 2k\pi} = +\infty$$

car $\cup_{k=0}^{+\infty} [\frac{\pi}{6} + 2k\pi, \frac{5\pi}{6} + 2k\pi] \subset \mathbb{R}$ et sur $[\frac{\pi}{6} + 2k\pi, \frac{5\pi}{6} + 2k\pi]$, on a $|\sin t| \geq \frac{1}{2}$ et $\frac{1}{t} \geq 1/(\frac{\pi}{6} + 2k\pi)$. On montre qu'il y a quand même convergence simple et que la valeur de cette intégrale est $\frac{\pi}{2}$.

Il se peut qu'une intégrale soit plusieurs fois impropre. Donnons nous par exemple un intervalle $]a, b[$ avec $-\infty \leq a \leq b \leq +\infty$ et une application localement intégrable f sur $]a, b[$. Le problème de la convergence de l'intégrale $\int_a^b f(t) dt$ se pose à la fois en a et en b .

On vérifie que s'il existe $c \in]a, b[$ tel que les intégrales impropres de f sur $]a, c[$ et $]c, b[$ convergent alors pour tout $d \in]a, b[$, les intégrales impropres de f sur $]a, d[$ et $]d, b[$ convergent et

$$\int_a^c f(t) dt + \int_c^b f(t) dt = \int_a^d f(t) dt + \int_d^b f(t) dt.$$

Pour $c \in]a, b[$ fixé, notons

$$F(x) = \int_c^x f(t) dt, \quad G(y) = \int_y^c f(t) dt, \quad \Phi(y, x) = \int_y^x f(t) dt.$$

Théorème 4.1.1 *Les deux assertions suivantes sont équivalentes :*

- i) *Les intégrales impropres de f sur $]a, c[$ et $]c, b[$ convergent.*
- ii) *L'application Φ admet une limite au point (a, b) de $\overline{\mathbb{R}}^2$.*

Lorsque ces assertions sont vraies, la limite de Φ est égale à la somme des deux intégrales impropres. On dit alors que l'intégrale doublement impropre $\int_a^b f(t)dt$ converge et on lui attribue la valeur (indépendante de c) :

$$\int_a^b f(t)dt = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Démonstration : Si i) est vraie Alors $\lim_{x \rightarrow b} F(x) = \int_c^b f(t)dt$ et $\lim_{y \rightarrow a} G(y) = \int_a^c f(t)dt$. Mais comme $\Phi(y, x) = G(y) + F(x)$, il existe donc

$$\lim_{(y,x) \rightarrow (a,b)} \Phi(y, x) = \int_a^c f(t)dt + \int_c^b f(t)dt.$$

Inversement si ii) est vraie, alors pour tout $\varepsilon > 0$, il existe (c, c') avec $a < c < c' < b$ tels que pour tout (y, x) avec $a < y \leq c \leq c' \leq x < b$, on a $|\Phi(y, x) - l| \leq \varepsilon/2$. On en déduit que pour tout (u, v) avec $c' \leq u < v < b$, on a

$$|\Phi(c, u) - \Phi(c, v)| \leq |\Phi(c, u) - l| + |\Phi(c, v) - l| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Mais comme $\Phi(c, v) - \Phi(c, u) = \int_u^v f(t)dt$, le critère de Cauchy est vérifié pour l'intégrale $\int_a^b f(t)dt$. Elle converge donc en b .

De même pour l'intégrale en a : $\int_a f(t)dt$. □

Exemple : • L'intégrale $\int_{-\infty}^{+\infty} e^{-|t|}dt$ est impropre en $-\infty$ et en $+\infty$. Mais elle converge à la fois en $-\infty$ et en $+\infty$.

• L'intégrale $\int_0^{+\infty} \frac{e^{-t}}{\sqrt{t}}dt$ est impropre en 0 et en $+\infty$.

Elle converge en 0 car $\frac{e^{-t}}{\sqrt{t}} \simeq_0 1/\sqrt{t}$ intégrable en 0.

Elle converge en $+\infty$ car (en anticipant sur le critère de Riemann) $\lim_{t \rightarrow +\infty} t^2 \frac{e^{-t}}{\sqrt{t}} = 0$. L'intégrale est donc convergente.

4.2 Intégrales impropres des fonctions positives

Proposition 4.2.1 Soit $f : [a, b[\rightarrow \mathbb{R}^+$ une application localement intégrable à valeurs positives. L'intégrale impropre de f converge ssi l'application intégrale $F(x)$ est majorée sur $[a, b[$. L'intégrale $\int_a^b f(t)dt$ est alors la borne sup de F sur $[a, b[$ (atteinte en b)

Démonstration : Ici $F(x) = \int_a^x f(t)dt$ est croissante en effet pour $a \leq x' \leq x'' < b$, on a

$$F(x'') - F(x') = \int_{x'}^{x''} f(t)dt \geq 0.$$

Puis F , croissante, bornée est convergente. □

Proposition 4.2.2 (Critère de comparaison) Soient $f, g : [a, b[\rightarrow \mathbb{R}^+$ des applications localement intégrables à valeurs positives vérifiant $f \leq g$. Alors

i) Si l'intégrale impropre de g converge sur $[a, b[$, celle de f aussi et

$$\int_a^b f(t)dt \leq \int_a^b g(t)dt.$$

ii) Si l'intégrale impropre de f diverge, celle de g aussi.

Plus généralement, on a aussi

Proposition 4.2.3 Soient $f, g : [a, b[\rightarrow \mathbb{R}^+$ des applications localement intégrables à valeurs positives vérifiant $f = O(g)$ au voisinage de b . Alors

i) Si l'intégrale impropre de g converge sur $[a, b[$, celle de f aussi et

ii) Si l'intégrale impropre de f diverge, celle de g aussi.

Il y a de multiples versions de ces critères de comparaison. En particulier, on peut intégrer les relations de comparaison. La version la plus importante de ces critères est celle liée à l'équivalence :

Proposition 4.2.4 Soient $f, g : [a, b[\rightarrow \mathbb{R}^+$ deux applications localement intégrables positives. Si f et g sont équivalentes en b alors leurs intégrales impropres sont de même nature.

Remarque : Par contre, en cas de convergence des intégrales, on ne peut pas comparer les valeurs de ces intégrales impropres.

Démonstration : Si $f(x) \simeq_b g(x)$ alors par exemple avec $\varepsilon = 1/2$, il existe un voisinage $[b', b'']$ de b tel que pour $x \in [b', b'']$, on a $(1/2)g(x) \leq f(x) \leq (3/2)g(x)$. D'où pour $b' \leq x' \leq x'' < b$, on a

$$1/2 \int_{x'}^{x''} g(t)dt \leq \int_{x'}^{x''} f(t)dt \leq 3/2 \int_{x'}^{x''} g(t)dt.$$

Le critère de Cauchy de f donne celui de g et vice versa. Les intégrales impropres de f et de g en b sont donc de même nature. \square

En pratique pour déterminer la nature d'une intégrale impropre de fonctions positives, on commence par simplifier au maximum l'intégrand en cherchant l'équivalent le plus simple. Attention, l'équivalent est à prendre en le point b qui est critique pour la convergence de l'intégrale.

Exemples : l'échelle de Riemann

• En $+\infty$, les fonctions $1/x^\alpha$ sont intégrables en $+\infty$ ssi $\alpha > 1$.

Par exemple, $1/\sqrt{x}$ ne l'est pas, pas plus que $1/x$ tandis que $1/x^2$ l'est.

• En 0, les fonctions $1/x^\alpha$ sont intégrables en 0 ssi $\alpha < 1$.

Par exemple, $1/\sqrt{x}$ l'est mais pas $1/x$ ni $1/x^2$.

• En un point fini b , c'est l'analogie de 0 : $1/(b-x)^\alpha$ est intégrable en b si $\alpha < 1$. Sinon, l'intégrale diverge en b .

En particulier, noter qu'aucune fonction $1/x^\alpha$ n'est intégrable à la fois en 0 et en $+\infty$.

Une conséquence du critère de comparaison et de l'échelle de Riemann sont les critères suivants :

Proposition 4.2.5 (Critère de Riemann en 0) Soit $f :]0, b] \rightarrow \mathbb{R}$ localement intégrable.

- S'il existe $\alpha < 1$ tel que $\lim_{t \rightarrow 0} t^\alpha f(t) = 0$ alors f est intégrable en 0.
- S'il existe $\alpha > 1$ tel que $\lim_{t \rightarrow 0} t^\alpha f(t) = +\infty$ alors f n'est pas intégrable en 0.

Démonstration : En effet dans le premier cas, il existe $x_1 > 0$ tel que pour $t \leq x_1$, on a $t^\alpha f(t) \leq 1$ donc $f(t) \leq 1/t^\alpha$ qui est intégrable en 0 car $\alpha < 1$.

Puis dans le deuxième cas, il existe $x_2 > 0$ tel que pour $t \leq x_2$, on a $t^\alpha f(t) \geq 1$ donc $f(t) \geq 1/t^\alpha$ qui n'est pas intégrable en 0 car $\alpha > 1$. \square

Puis on a un critère semblable en $+\infty$, mais attention les conditions sont les exactes opposées.

Proposition 4.2.6 (Critère de Riemann en $+\infty$) Soit $f : [a, +\infty[\rightarrow \mathbb{R}$ localement intégrable.

- S'il existe $\alpha > 1$ tel que $\lim_{t \rightarrow +\infty} t^\alpha f(t) = 0$ alors f est intégrable en $+\infty$.
- S'il existe $\alpha < 1$ tel que $\lim_{t \rightarrow +\infty} t^\alpha f(t) = +\infty$ alors f n'est pas intégrable en $+\infty$.

Démonstration : En effet dans le premier cas, il existe $x_3 > 0$ tel que pour $t \geq x_3$, on a $t^\alpha f(t) \leq 1$ donc $f(t) \leq 1/t^\alpha$ qui est intégrable en $+\infty$ car $\alpha > 1$.

Puis dans le deuxième cas, il existe $x_4 > 0$ tel que pour $t \geq x_4$, on a $t^\alpha f(t) \geq 1$ donc $f(t) \geq 1/t^\alpha$ qui n'est pas intégrable en $+\infty$ car $\alpha > 1$. \square

En général, on ne se complique pas la vie : on essaye d'appliquer ces critères avec $\alpha = 2$ lorsqu'on cherche un $\alpha > 1$ et avec $\alpha = 1/2$ lorsqu'on cherche un $\alpha < 1$.

Par exemple, e^{-x} est intégrable en $+\infty$ car $\lim_{t \rightarrow +\infty} t^2 e^{-t} = 0$. Puis $\ln t$ l'est en 0 car $\lim_{t \rightarrow 0} \sqrt{t} \ln t = 0$.

Autres fonctions de référence : les intégrales de Bertrand

- L'intégrale $\int_e^{+\infty} \frac{1}{t^\alpha (\ln t)^\beta} dt$ converge en $+\infty$ ssi $\alpha > 1$ ou $\alpha = 1$ et $\beta > 1$.
- L'intégrale $\int_0^{1/e} \frac{1}{t^\alpha |\ln^\beta t|} dt$ converge en 0 ssi $\alpha < 1$ ou $\alpha = 1$ et $\beta > 1$.

En effet pour le problème de convergence en $+\infty$, si $\alpha > 1$, alors il existe $\frac{\alpha+1}{2} > 1$ avec

$$t^{\frac{\alpha+1}{2}} \times \frac{1}{t^\alpha (\ln t)^\beta} = \frac{1}{t^{\frac{\alpha-1}{2}} (\ln t)^\beta} \rightarrow 0, \quad t \rightarrow +\infty.$$

car $\frac{\alpha-1}{2} > 0$. Alors que si $\alpha < 1$, alors il existe $\frac{\alpha+1}{2} < 1$ avec

$$t^{\frac{\alpha+1}{2}} \times \frac{1}{t^\alpha (\ln t)^\beta} = \frac{t^{\frac{1-\alpha}{2}}}{(\ln t)^\beta} \rightarrow +\infty, \quad t \rightarrow +\infty.$$

car $\frac{1-\alpha}{2} > 0$. Puis si $\alpha = 1$, alors le changement de variable $u = \ln t$ donne

$$\int_e^{+\infty} \frac{1}{t(\ln t)^\beta} dt = \int_1^{+\infty} \frac{du}{u^\beta}$$

qui converge si $\beta > 1$.

On fait de même pour le problème de convergence en 0 : si $\alpha > 1$, alors il existe $\frac{\alpha+1}{2} > 1$ avec

$$t^{\frac{\alpha+1}{2}} \times \frac{1}{t^\alpha |\ln t|^\beta} = \frac{1}{t^{\frac{\alpha-1}{2}} |\ln t|^\beta} \rightarrow +\infty, \quad t \rightarrow 0.$$

car $\frac{\alpha-1}{2} > 0$. Alors que si $\alpha < 1$, alors il existe $\frac{\alpha+1}{2} < 1$ avec

$$t^{\frac{\alpha+1}{2}} \times \frac{1}{t^\alpha |\ln t|^\beta} = \frac{t^{\frac{1-\alpha}{2}}}{|\ln t|^\beta} \rightarrow 0, \quad t \rightarrow 0.$$

car $\frac{1-\alpha}{2} > 0$. Puis si $\alpha = 1$, alors le changement de variable $u = \ln t$ donne

$$\int_0^{1/e} \frac{1}{t|\ln t|^\beta} dt = \int_{-\infty}^{-1} \frac{du}{|u|^\beta}$$

qui converge si $\beta > 1$.

Chapitre 5

Suites et séries de fonctions Riemann-intégrables

On considère dans ce chapitre une suite de fonctions $(f_n)_n$ de $[a, b]$ dans \mathbb{R} . Elle pourrait être à valeurs dans \mathbb{C} ou dans un espace de Banach E .

5.1 Différentes convergences de fonctions

Définition 5.1.1 (Convergence simple) Une suite de fonctions $(f_n)_n$ converge simplement vers f si pour chaque $x \in [a, b]$ fixé,

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

Exemples :

- Soit la suite de fonctions f_n données sur $[0, +\infty[$ par $f_n(0) = 0$, $f_n(t) = 1$ pour $t \geq 1/n$ et linéaire entre 0 et $1/n$. La suite f_n converge simplement vers f nulle en 0 et égale à 1 ailleurs.

- Soit la suite de fonctions f_n données sur $]0, 1]$ par $f_n(t) = 1/t$ sur $[1/(n+1), 1]$ et $f_n(t) = n+1$ sur $]0, 1/(n+1)[$. La suite f_n converge simplement sur $]0, 1]$ vers $f(t) = 1/t$.

La convergence s'écrit avec le formalisme des quantificateurs « $\forall \exists$ »,

$$\forall x \in [a, b], \forall \varepsilon > 0, \exists n_0 \text{ tel que pour } n \geq n_0, \text{ on a } |f_n(x) - f(x)| \leq \varepsilon.$$

Notons que dans cette assertion, $n_0 = n_0(\varepsilon, x)$ dépend de ε mais aussi du x fixé.

Définition 5.1.2 (Convergence uniforme) Une suite de fonctions $(f_n)_n$ converge uniformément sur $[a, b]$ vers la fonction f si

$$\|f_n - f\|_\infty := \sup_{x \in [a, b]} |f_n(x) - f(x)| \rightarrow 0, \quad n \rightarrow +\infty.$$

Exemples :

- Dans l'exemple 1 précédent, $\sup_{x \in [0,1]} |f_n(x) - f(x)| = 1$. Il n'y a donc pas convergence uniforme.
- Dans l'exemple 2, il n'y a pas convergence uniforme non plus car $\sup_{x \in]0,1[} |f_n(x) - f(x)| = +\infty$.

La convergence uniforme s'écrit avec le formalisme « $\forall \exists$ »,

$$\forall \varepsilon > 0, \exists n_0, \forall n \geq n_0, \forall x \in [a, b] \text{ on a } |f_n(x) - f(x)| \leq \varepsilon.$$

Dans cette assertion, $n_0 = n_0(\varepsilon)$ dépend de ε mais plus de $x \in [a, b]$: la convergence est uniforme en x .

Pour avoir la convergence uniforme, on a juste déplacé la quantification « $\forall x \in [a, b]$ » d'avant à après « $\exists n_0$ ».

Remarque :

- La convergence uniforme entraîne la convergence simple.
- La convergence uniforme d'une suite de fonctions peut se montrer par le **critère de Cauchy uniforme** : pour tout $\varepsilon > 0$, il existe n_0 tel que pour tout $n, m \geq n_0$, on a

$$\sup_{x \in [a, b]} |f_n(x) - f_m(x)| \leq \varepsilon.$$

- Pour montrer la convergence uniforme de $(f_n)_n$, on commence par chercher la limite simple (ponctuelle) de $f_n : f(x) = \lim_{n \rightarrow +\infty} f_n(x)$. Puis on étudie $\sup_{x \in [a, b]} |f_n(x) - f(x)|$ pour déterminer s'il y a ou non convergence uniforme.

De nombreuses propriétés peuvent être transférées par convergence uniforme d'une suite de fonctions $(f_n)_n$ à une fonction limite f : la continuité, la continuité uniforme, le caractère lipschitzien, être bornée, la dérivabilité (sous certaines conditions : convergence uniforme des dérivées et convergence en un point). On verra dans la section suivante ce qu'il en est de l'intégrabilité.

La convergence des séries de fonctions est un cas particulier de celle des suites de fonctions car une série se voit comme la limite de la suite des sommes partielles. Mais comme pour les séries numériques, pour les séries de fonctions, il y a d'autres notions de convergence : la convergence absolue et la convergence normale.

Définition 5.1.3 (Convergence absolue) *La série de fonction $\sum_n f_n$ est dite absolument convergente si la série $\sum_n |f_n|$ converge simplement.*

La convergence absolue de la série $\sum_n f_n$ entraîne la convergence simple de la série $\sum_n f_n$. Mais, elle n'a aucun lien avec la convergence uniforme : par exemple, $f_n(x) = \frac{(-1)^n}{n+x}$ converge uniformément car

$$\left| \sum_{k=1}^{+\infty} f_k(x) - \sum_{k=1}^n f_k(x) \right| \leq \frac{1}{n+1+x} \leq \frac{1}{n+1}$$

mais pas absolument car $|f_n(x)| \simeq 1/n$ et $\sum_n 1/n$ diverge.

Définition 5.1.4 (Convergence normale) Soit $\sum_n f_n$ une série de fonctions. S'il existe une série numérique positive $\sum_n a_n$ (avec $a_n \geq 0$) telle que

i) la série $\sum_n a_n$ est convergente

ii) pour tout n , $\sup_{x \in [a,b]} |f_n(x)| \leq a_n$

alors la série $\sum_n f_n$ est dite normalement convergente.

Proposition 5.1.1 Il y a convergence normale de la série de fonctions $\sum_n f_n$ ssi la série numériques des $\|f_n\|_\infty$ est convergente.

La convergence normale de la série $\sum_n f_n$ entraîne la convergence uniforme de la série $\sum_n f_n$: en effet il suffit d'appliquer le critère de Cauchy uniforme à partir du critère de Cauchy pour la convergence de $\sum_n a_n$.

Souvent pour montrer la convergence uniforme d'une série de fonction $\sum_n f_n$, on cherche à voir sa convergence normale.

La convergence normale de la série $\sum_n f_n$ entraîne aussi la convergence absolue de la série $\sum_n f_n$.

Attention, les implications réciproques entre les différents modes de convergence sont fausses. En particulier, la convergence absolue et uniforme d'une série de fonctions n'entraîne pas sa convergence normale. En effet, considérons par exemple la fonction f_n affine par morceaux pour $n \geq 2$, nulle sur

$$\left[0, \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right)\right] \cup \left[\frac{1}{2}\left(\frac{1}{n} + \frac{1}{n-1}\right), 1\right]$$

et valant $1/n$ en $1/n$. On complète la suite $(f_n)_n$ avec $f_0(x) = f_1(x) = 0$.

On définit F sur $[0, 1]$ par $F(0) = 0$ et par $F(t) = f_n(t)$ pour

$$t \in \left[\frac{1}{2}\left(\frac{1}{n} + \frac{1}{n+1}\right), \frac{1}{2}\left(\frac{1}{n} + \frac{1}{n-1}\right)\right].$$

Notons $F_N = \sum_{n=0}^N f_n$. La fonction $F - F_N$ est nulle sur $[\frac{1}{2}(\frac{1}{N} + \frac{1}{N+1}) + 1]$ et égale à F sur $[0, \frac{1}{2}(\frac{1}{N} + \frac{1}{N+1})]$. On a donc $\sup_{t \in [0,1]} |F(t) - F_N(t)| \leq 1/(N+1)$.

Il y a donc convergence uniforme de F_N vers F . La convergence est absolue car tout est positif. Elle n'est pas normale car $\sup_{t \in [0,1]} |f_n(t)| = 1/n$ et $\sum_n 1/n$ diverge.

5.2 Intégrabilité des suites et séries de fonctions

D'abord, pour les intégrales sur les intervalles finis, on a :

Théorème 5.2.1 Si $(f_n)_n$ est une suite de fonctions intégrables sur un intervalle fini $[a, b]$ compact qui converge uniformément vers f sur $[a, b]$ alors f est intégrable sur $[a, b]$ et la suite des intégrales converge :

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx.$$

Démonstration :

• Soit $\varepsilon > 0$. Comme la suite $(f_n)_n$ converge uniformément vers f sur $[a, b]$, il existe n_0 tel que

$$\forall n \geq n_0, \sup_{x \in [a, b]} |f(x) - f_n(x)| \leq \frac{\varepsilon}{2(b-a)}.$$

Puisque f_{n_0} est intégrable, il existe deux applications en escalier φ et $\phi \geq 0$ telles que

$$|f_{n_0} - \varphi| \leq \phi, \quad \int_a^b \phi(x) dx \leq \varepsilon/2.$$

Il existe donc deux applications en escalier $\varphi_1 = \varphi$ et $\phi_1 = \phi + \varepsilon/(2(b-a))$ telles que

$$|f - \varphi_1| \leq |f - f_{n_0}| + |f_{n_0} - \varphi| \leq \phi + \varepsilon/(2(b-a)) = \phi_1, \quad \int_a^b \phi_1(x) dx \leq \varepsilon,$$

ce qui prouve que f est Riemann-intégrable.

• On peut aussi montrer la Riemann-intégrabilité de f par la première définition de la Riemann-intégrabilité :

D'abord, les f_k sont intégrables donc bornées. Comme les f_k convergent uniformément vers f , la limite f l'est aussi. Puis pour tout $\varepsilon > 0$, il existe k_0 tel que pour tout $k \geq k_0$, on a pour tout x ,

$$|f_k(x) - f(x)| \leq \frac{\varepsilon}{b-a}.$$

On en déduit que

$$\left| \sup_{x \in [a, b]} f_k(x) - \sup_{x \in [a, b]} f(x) \right| \leq \frac{\varepsilon}{b-a}, \quad \left| \inf_{x \in [a, b]} f_k(x) - \inf_{x \in [a, b]} f(x) \right| \leq \frac{\varepsilon}{b-a}.$$

Mais alors pour une subdivision $S = \{a_0, \dots, a_n\}$ de $[a, b]$ et les sommes de Darboux associées, on a

$$\begin{aligned} |A^+(f, S) - A^+(f_k, S)| &\leq \sum_{i=0}^{n-1} |M_i^k - M_i|(a_{i+1} - a_i) \\ &\leq \frac{\varepsilon}{b-a} \sum_{i=0}^{n-1} (a_{i+1} - a_i) \\ &\leq \varepsilon \end{aligned}$$

en notant M_i et M_i^k les sup de f et de f_k sur $[a_i, a_{i+1}]$. On ferait de même pour les sommes de Darboux inférieures. On en déduit pour $k \geq k_0$,

$$A^+(f_k, S) - A^-(f_k, S) - 2\varepsilon \leq A^+(f, S) - A^-(f, S) \leq A^+(f_k, S) - A^-(f_k, S) + 2\varepsilon.$$

Pour k_0 , comme f_{k_0} est Riemann-intégrable, il existe une subdivision S_{k_0} telle que

$$A^+(f, S_{k_0}) - A^-(f, S_{k_0}) \leq \varepsilon.$$

On a alors $0 \leq A^+(f, S) - A^-(f, S) \leq 3\varepsilon$ pour toute subdivision $S \supset S_{k_0}$.

La fonction f est donc Riemann-intégrable.

• Pour la convergence des intégrales : En effet, si f_k converge uniformément vers f , alors

$$\sup_{x \in [a, b]} |f_k(x) - f(x)| \rightarrow 0, \quad k \rightarrow +\infty.$$

Pour $\varepsilon > 0$ fixé, il existe k_0 tel que pour tout $k \geq k_0$ et tout $x \in [a, b]$, on a $|f_k(x) - f(x)| \leq \varepsilon/(b-a)$. D'où

$$\left| \int_a^b f_k(x) dx - \int_a^b f(x) dx \right| \leq \int_a^b |f_k(x) - f(x)| dx \leq \int_a^b \frac{\varepsilon}{b-a} dx = \varepsilon.$$

On a donc $\forall \varepsilon > 0, \exists k_0$, tel que pour $k \geq k_0$, $\left| \int_a^b f_k(x) dx - \int_a^b f(x) dx \right| \leq \varepsilon$, ce qui prouve la convergence cherchée. \square

• Sans convergence uniforme, la convergence de la suite des intégrales est fautive :

Considérons sur l'intervalle $[0, 1]$, $f(x) = 0$ et f_n la fonction nulle sur $[2/n, 1]$ et linéaire entre 0 où $f(0) = 0$ et $1/n$ où $f(1/n) = n$ et linéaire encore entre $1/n$ et $2/n$.

$$f_n(x) = \begin{cases} n^2 x & \text{si } 0 \leq x \leq 1/n \\ -n^2(x - 2/n) & \text{si } 1/n \leq x \leq 2/n \\ 0 & \text{si } 2/n \leq x \leq 1. \end{cases}$$

Les fonctions f_n et f sont continues et intégrables. La convergence n'est pas uniforme car

$$\sup_{x \in [0, 1]} |f_n(x) - f(x)| = n \not\rightarrow 0$$

et puis

$$\int_0^1 f(x) dx = 0, \quad \int_0^1 f_n(x) dx = \frac{2}{n} \times n \times 1/2 = 1.$$

• Si l'intervalle est infini, la convergence des intégrales est encore en défaut, même avec la convergence uniforme :

Considérons sur l'intervalle $[0, +\infty[$, la fonction $f(x) = 0$ et

$$f_n(x) = \begin{cases} 1/n & \text{si } 0 \leq x \leq n \\ 1/n + 1/n - x/n & \text{si } n \leq x \leq n+1 \\ 0 & \text{si } x \geq n+1. \end{cases}$$

On a la convergence uniforme car $\sup_{x \in [0, +\infty[} |f_n(x) - f(x)| = 1/n \rightarrow 0, n \rightarrow +\infty$ mais pas la convergence de la suite des intégrales puisque

$$\int_0^{+\infty} f(x)dx = 0, \quad \int_0^{+\infty} f_n(x)dx = n \times 1/n + 1/2 + 0 = 3/2 \not\rightarrow 0.$$

Le seul résultat positif pour des intégrales impropres est lorsque l'intervalle est fini :

Théorème 5.2.2 *Soit $(f_n)_n$ une suite de fonctions définies et localement intégrables sur $[a, b[$ (intervalle borné) qui converge uniformément sur $[a, b[$ vers f . Alors f est localement intégrable. Puis si les f_n ont des intégrales impropres convergentes sur $[a, b[$ alors l'intégrale impropre $\int_a^b f(x)dx$ converge et*

$$\int_a^b f(x)dx = \lim_{n \rightarrow +\infty} \int_a^b f_n(x)dx.$$

Démonstration : Sur un intervalle fini $[a, c] \subset [a, b[$, f est Riemann-intégrable sur $[a, c]$ d'après le théorème 5.2.1 car limite uniforme des f_n qui sont Riemann intégrables sur $[a, c]$.

Soit $\varepsilon > 0$, il existe n_0 tel que pour $n \geq n_0$, on a pour tout $x \in [a, b[$, $|f(x) - f_n(x)| \leq \varepsilon/(2(b-a))$. On a alors pour $a \leq c \leq d < b$,

$$\begin{aligned} \left| \int_c^d f(t)dt \right| &\leq \left| \int_c^d f(t) - f_{n_0}(t)dt \right| + \left| \int_c^d f_{n_0}(t)dt \right| \\ &\leq (d-c) \frac{\varepsilon}{2(b-a)} + \left| \int_c^d f_{n_0}(t)dt \right| \\ &\leq \frac{\varepsilon}{2} + \left| \int_c^d f_{n_0}(t)dt \right|. \end{aligned} \quad (5.1)$$

Comme l'intégrale impropre $\int_a^b f_{n_0}(t)dt$ est convergente, il existe x_0 tel que pour $x, x' \in [x_0, b[$, on ait $\left| \int_x^{x'} f_{n_0}(t)dt \right| \leq \varepsilon/2$. On a donc avec (5.1) pour $c, d \in [x_0, b[$, $\left| \int_c^d f(t)dt \right| \leq \varepsilon$, ce qui d'après le critère de Cauchy justifie que l'intégrale impropre $\int_a^b f(t)dt$ converge.

Puis pour $n \geq n_0$, et $x \in [a, b[$, on a

$$\left| \int_a^x f(t)dt - \int_a^x f_n(t)dt \right| \leq \int_a^x |f(t) - f_n(t)|dt \leq (x-a) \frac{\varepsilon}{b-a} \leq \varepsilon.$$

Pour $n \geq n_0$ fixé, le passage à la limite $x \rightarrow b^-$ donne

$$\left| \int_a^b f(t)dt - \int_a^b f_n(t)dt \right| \leq \varepsilon.$$

Comme c'est vrai pour tout $\varepsilon > 0$, on a $\lim_{n \rightarrow +\infty} \int_a^b f_n(t) dt = \int_a^b f(t) dt$. \square

Les résultats précédents (Théorèmes 5.2.1 et 5.2.2) admettent des versions pour les séries de fonctions $\sum_n f_n$. Leur preuve est immédiate en interprétant une série $\sum_{n=0}^{+\infty} f_n$ comme la limite de la suite des sommes partielles $\sum_{n=0}^k f_n$.

Corollaire 5.2.1 *Soit $\sum_n f_n$ une série de fonctions toutes intégrables sur un intervalle fini $[a, b]$ compact. Si la série $\sum_n f_n$ converge uniformément sur $[a, b]$ alors $F = \sum_n f_n$ est intégrable sur $[a, b]$ et on a*

$$\int_a^b \sum_{n=0}^{+\infty} f_n(x) dx = \sum_{n=0}^{+\infty} \int_a^b f_n(x) dx.$$

Corollaire 5.2.2 *Soit $\sum_n f_n$ une série de fonctions toutes localement intégrables sur un intervalle fini $[a, b[$ compact. Si les intégrales impropres $\int_a^b f_n(t) dt$ convergent toutes et si la série $\sum_n f_n$ converge uniformément sur $[a, b[$ alors $F = \sum_n f_n$ est localement intégrable sur $[a, b[$ et d'intégrale impropre convergente, avec*

$$\int_a^b \sum_{n=0}^{+\infty} f_n(x) dx = \sum_{n=0}^{+\infty} \int_a^b f_n(x) dx.$$

Les résultats de cette section seront à comparer avec les résultats d'interversion limite / intégrale pour l'intégrale de Lebesgue où les hypothèses de convergence seront beaucoup plus faibles.

5.3 Quelques inconvénients de l'intégrale de Riemann

En général l'intégrale de Riemann passe mal à la limite sans convergence uniforme ou si l'intervalle n'est pas borné : si on considère une suite $(f_n)_n$ de fonctions Riemann-intégrables sur $[a, b]$ qui converge vers f alors il n'est pas vrai en général que

$$\lim_{n \rightarrow +\infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx. \quad (5.2)$$

Autrement dit, l'intégrale (de Riemann) et la limite s'intervertissent mal.

Trois problèmes peuvent survenir en fait :

- La limite à gauche de (5.2) peut ne pas exister.
- Même si la limite existe, la fonction f peut ne pas être Riemann-intégrable.
- Même si les deux membres existent, ils peuvent ne pas être égaux.

Illustrons ces problèmes par des exemples.

Exemple. Soit $(r_n)_n$ la suite des rationnels de $[0, 1]$. Notons alors

$$f_n(x) = \begin{cases} 1 & \text{si } x = r_k, k \leq n \\ 0 & \text{sinon.} \end{cases}$$

Il est facile de voir que f_n est Riemann-intégrable (elle est même réglée car elle a un nombre fini de point de discontinuité) et que $\int_0^1 f_n(x) dx = 0$.

Puis pour tout x , $\lim_{n \rightarrow +\infty} f_n(x) = f(x)$, fonction indicatrice de $\mathbb{Q} \cap [0, 1]$. Or f n'est pas Riemann intégrable (ses sommes de Darboux inférieures valent 0, ses supérieures valent 1).

Exemple. Soit f_n donnée par

$$f_n(t) = \begin{cases} 1/t & \text{sur } [1, n] \\ 0 & \text{sur } [1 + 1/n, +\infty[\\ \text{affine} & \text{entre.} \end{cases}$$

La limite est $f(t) = 1/t$ sur $[1, +\infty[$ et $\sup_{t \in [1, +\infty[} |f(t) - f_n(t)| \leq 1/n$. Il y a donc convergence uniforme.

Cependant $\int_1^{+\infty} f_n(t) dt$ existent mais pas $\int_1^{+\infty} f(t) dt$

Exemple. Etudier la convergence de $\lim_{n \rightarrow +\infty} \int_1^{+\infty} \frac{n}{t+n} e^{-t} dt$.

Exemple. Soit

$$f_n(x) = -2n^2 x e^{-n^2 x^2} + 2(n+1)^2 x e^{-(n+1)^2 x^2}.$$

La suite de fonction $(f_n)_n$ est télescopique et comme $\lim_{n \rightarrow +\infty} 2n^2 x e^{-n^2 x^2} = 0$, on a

$$g(x) = \sum_{n=1}^{+\infty} f_n(x) = -2x e^{-x^2}.$$

Puis

$$\int_0^1 g(x) dx = \int_0^1 (-2x e^{-x^2}) dx = e^{-1} - 1,$$

tandis que

$$\sum_{n=1}^{+\infty} \int_0^1 f_n(x) dx = \sum_{n=1}^{+\infty} (e^{-n^2} - e^{-(n+1)^2}) = e^{-1}.$$

L'égalité (5.2) n'est pas valide pour cet exemple.

Deuxième partie
Intégrale de Lebesgue

Chapitre 6

Tribus (σ -algèbres) et mesures

6.1 Introduction à la théorie de la mesure

Historiquement, comme l'indique le nom, le but de cette théorie est de mesurer des ensembles. Sans s'en rendre compte, plusieurs types de « mesures » ont déjà été rencontrées :

- Le cardinal d'un ensemble discret, par exemple le cardinal de $\{1, 2, 3, 4\}$ est 4, celui de $\{1, 7, 19, 74, 106\}$ est 5, celui de \mathbb{N} est $+\infty$.
- La longueur, l'aire, le volume d'une courbe, d'une figure plane, d'un solide en dimension 3.

Par exemple, la longueur de l'intervalle $[-3, 5]$ est $5 - (-3) = 8$, l'aire du disque $D(0, R)$ est πR^2 , le volume du cylindre de base $D(0, 1)$ et de hauteur 4 est 4π .

- La probabilité d'un évènement : par exemple si on lance une dé équilibré la probabilité d'avoir un quatre est $1/6$, celle d'avoir une face impaire est $1/2$, celle de gagner au loto (au premier rang) est $1/C_{49}^5$.

Ces mesures sont des cas particuliers d'une notion plus générale de mesure, outil de base pour une nouvelle théorie de l'intégration, dite intégrale de Lebesgue (1902). Elle généralise la notion déjà vue de l'intégrale de Riemann, donc ce qui est déjà connu avec Riemann n'est pas perdu mais généralisé. Cependant, cette nouvelle théorie

- s'applique à une classe de fonctions beaucoup plus grande (les fonctions *mesurables*)
- a des théorèmes de convergence beaucoup plus forts : théorème de convergence monotone, théorème de convergence dominée
- unifie les différentes façon de mesurer, par exemple le calcul d'une espérance de v.a., d'une série, d'une intégrale classique sont des cas particuliers d'intégrales au sens de Lebesgue.

Cette théorie unifiante éclaire les analogies souvent constatées en Deug entre les résultats liés aux séries et aux intégrales de Riemann.

Remarque 6.1.1 Notons qu'une mesure est toujours associée à une famille d'ensembles à mesurer. On appellera bientôt ces familles des *tribus* ou des σ -*algèbres*.

6.2 Algèbres et σ -algèbres

On considère dans la suite X un ensemble fixé. On pourra penser à $X = \mathbb{N}$ ou à $X = \mathbb{R}$.

On rappelle que $\mathcal{P}(X)$ désigne l'ensemble de toutes les parties de l'ensemble X . Par exemple si $X = \{a, b, c\}$ alors

$$\mathcal{P}(X) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$

Noter que de façon générale, si $\text{card}(X) = n$ alors $\text{card}\mathcal{P}(X) = 2^n$. Rappelons encore les comportements de \cup et \cap par rapport au complémentaire c :

$$\begin{aligned}(A \cup B)^c &= A^c \cap B^c, \\ (A \cap B)^c &= A^c \cup B^c, \\ (A^c)^c &= A.\end{aligned}$$

En termes logiques, une réunion \cup se comprend comme un « ou », une intersection \cap comme un « et », un complémentaire comme un « contraire de ».

Définition 6.2.1 (Algèbre) $\mathcal{A} \subset \mathcal{P}(X)$ est une algèbre si

- $X \in \mathcal{A}$,
- \mathcal{A} est stable par réunion finie,
- \mathcal{A} est stable par complémentaire.

Remarque 6.2.1

- L'ensemble vide $\emptyset \in \mathcal{A}$.
- Si $A, B \in \mathcal{A}$ alors $A \cap B \in \mathcal{A}$.
- Si $A, B \in \mathcal{A}$ alors $A \setminus B \in \mathcal{A}$.

Par exemple $\mathcal{A} = \{\emptyset, X\}$, $\mathcal{P}(X)$, $\{A \subset X, A \text{ fini ou } X \setminus A \text{ fini}\}$ sont des algèbres.

Définition 6.2.2 (σ -algèbre, tribu) $\mathcal{A} \subset \mathcal{P}(X)$ est une tribu ou une σ -algèbre si

- $X \in \mathcal{A}$
- Si pour tout $i \in \mathbb{N}$ $A_i \in \mathcal{A}$ alors $\cup_{i \in \mathbb{N}} A_i \in \mathcal{A}$: \mathcal{A} est stable par réunion dénombrable.
- Si $A \in \mathcal{A}$ alors $A^c \in \mathcal{A}$: \mathcal{A} est stable par complémentaire.

Conséquences :

- L'ensemble vide $\emptyset \in \mathcal{A}$.
- Si pour tout $i \in \mathbb{N}$, $A_i \in \mathcal{A}$, alors $\cap_{i \in \mathbb{N}} A_i \in \mathcal{A}$: stabilité par intersection dénombrable.
- En particulier, $A \cap B \in \mathcal{A}$ et $A \cup B \in \mathcal{A}$ quand A et B sont dans \mathcal{A} .
- Une σ -algèbre est une algèbre.
- Si $A, B \in \mathcal{A}$ alors $A \setminus B \in \mathcal{A}$.

Exemple : $\{\emptyset, X\}$ (la tribu triviale), $\mathcal{P}(X)$, famille d'observables \mathcal{F} en probabilité sont des σ -algèbres.

Remarque 6.2.2 (explication des axiomes d'une tribu) Une tribu est une famille d'ensembles sur laquelle une « mesure » va être définie. C'est donc une famille d'ensembles à mesurer, ce qu'on a appelé en probabilité une famille d'observables.

On comprend bien les axiomes en les interprétant en termes d'évènements probabilistes :

- Si un évènement A est observable, alors l'évènement contraire A^c doit l'être aussi : c'est ce que dit la stabilité par complémentaire.
- Si deux évènements A et B sont observables alors l'évènement A ou B (c'est à dire $A \cup B$) doit l'être aussi : c'est ce que dit la stabilité par réunion.
- La stabilité par réunion finie ne peut suffir. L'exemple suivant illustre ce fait. On lance un dé jusqu'à l'obtention du premier as. Un tel évènement ne peut être décrit par un nombre fini d'évènements élémentaires (a priori le numéro du lancer du premier as peut être arbitrairement grand). Si on note A l'évènement « on obtient un as » et A_i « on obtient un as au i -ème lancer », alors

$$A = A_1 \cup A_2 \cdots \cup A_n \cdots \cup \cdots = \bigcup_{n=1}^{+\infty} A_n$$

et pour que A soit observable quand les A_n le sont, il faut la stabilité par réunion dénombrable de la tribu (des observables) \mathcal{A} .

L'axiome de stabilité par réunion dit donc pour une collection dénombrable d'évènements $(A_n)_n$: si chaque A_n est observable alors l'évènement A_1 ou A_2 ou \cdots ou A_n ou \cdots l'est encore.

- L'évènement certain X est observable, c'est ce qui se cache sous l'axiome X est dans la tribu.
- Les autres opérations sur les évènements observables comme « si A et B sont observables alors A et B aussi » se déduisent des axiomes de base de la définition.

Les axiomes de la définition d'une tribu ne sont donc rien d'autre que la formalisation mathématique des opérations (logiques) qu'on peut faire sur des ensembles « à mesurer » (comme les évènements).

Définition 6.2.3 *Un ensemble muni d'une tribu (X, \mathcal{A}) est appelé un espace mesurable.*

Comme l'ensemble des parties $\mathcal{P}(X)$ de X est une tribu, cela semble la tribu la plus naturelle à considérer sur un ensemble X . Malheureusement, pour beaucoup d'ensembles, cette tribu est trop grande pour définir de bons outils dessus (les mesures) sans incohérence interne. Par exemple, pour généraliser la notion de longueur sur \mathbb{R} , on ne peut pas considérer $\mathcal{P}(\mathbb{R})$, il va falloir introduire une nouvelle tribu : la tribu borélienne. De façon générale, pour définir, les bonnes tribus à utiliser (ni trop pauvre ni trop riche, comme la tribu borélienne), on introduit la notion de tribu engendrée et ce sont les tribus engendrées par de bonnes familles qui nous intéresseront.

Lemme 6.2.1 *Une intersection quelconque de tribus est une tribu.*

Démonstration : Soit $\mathcal{A}_i, i \in I$, des tribus. On montre que $\mathcal{A} = \bigcap_{i \in I} \mathcal{A}_i$ en est une aussi.

- D'abord $X \in \mathcal{A}_i$ pour tout $i \in I$ donc $X \in \bigcap_{i \in I} \mathcal{A}_i = \mathcal{A}$.
- Soit $A \in \mathcal{A}$ alors pour tout $i \in I$, $A \in \mathcal{A}_i$, stable par complémentaire donc $A^c \in \mathcal{A}_i$ pour chaque $i \in I$. Mais alors, $A^c \in \mathcal{A} = \bigcap_{i \in I} \mathcal{A}_i$. La famille \mathcal{A} est donc stable par complémentaire.
- Soit pour $j \in \mathbb{N}$, $A_j \in \mathcal{A} = \bigcap_{i \in I} \mathcal{A}_i$. Pour chaque $i \in I$, $A_j \in \mathcal{A}_i$ donc $\bigcup_{j \in \mathbb{N}} A_j \in \mathcal{A}_i$ car \mathcal{A}_i est une tribu. Finalement, $\bigcup_{j \in \mathbb{N}} A_j \in \bigcap_{i \in I} \mathcal{A}_i = \mathcal{A}$, qui est stable par réunion quelconque. La famille \mathcal{A} satisfait toutes les conditions des tribus, c'en est donc une. \square

Définition 6.2.4 $\mathcal{M} \subset \mathcal{P}(X)$. On note $\sigma(\mathcal{M})$ la plus petite tribu de X contenant \mathcal{M} . On l'appelle la tribu engendrée par \mathcal{M} .

Proposition 6.2.1 $\sigma(\mathcal{M}) = \bigcap_{\mathcal{A} \text{ tribu contenant } \mathcal{M}} \mathcal{A}$.

Démonstration : Notons $\mathcal{B} = \bigcap_{\mathcal{A} \text{ tribu contenant } \mathcal{M}} \mathcal{A}$. D'après la proposition précédente \mathcal{B} est une tribu et elle contient, clairement, \mathcal{M} , si bien que d'après la définition de $\sigma(\mathcal{M})$, on a $\sigma(\mathcal{M}) \subset \mathcal{B}$.

Puis $\sigma(\mathcal{M})$ est une tribu contenant \mathcal{M} donc par définition de \mathcal{B} , on a $\mathcal{B} \subset \sigma(\mathcal{M})$.

Finalement, $\mathcal{B} = \sigma(\mathcal{M})$. \square

Exemple fondamental : la tribu borélienne

On considère un ensemble X muni d'une topologie $\mathcal{T} \subset \mathcal{P}(X)$. On rappelle qu'une topologie est une famille d'ensembles

- stable par réunion quelconque
- stable par intersection finie
- contient X .

Cette définition est à comparer avec celle d'une tribu. Les ensembles de \mathcal{T} sont appelés les ouverts de la topologie. D'habitude, les ouverts de \mathbb{R} sont les réunions (dénombrables) d'intervalles ouverts $]a_i, b_i[$.

Définition 6.2.5 (Tribu borélienne) La σ -algèbre engendrée par une topologie (c'est à dire engendrée par la famille $\mathcal{M} = \mathcal{T}$ des ouverts d'une topologie) est la tribu (ou σ -algèbre) borélienne, notée $\mathcal{B}(X)$. Les éléments de la tribu borélienne $\mathcal{B}(X)$ s'appellent les boréliens de X .

Il s'agit donc de la plus petite tribu contenant tous les ouverts de X .

La tribu borélienne est donc engendrée par les ouverts de la topologie. La plupart du temps, l'ensemble \mathbb{R} est muni de sa tribu borélienne $\mathcal{B}(\mathbb{R})$ engendrée par les ouverts de sa topologie usuelle (topologie de l'ordre ou, c'est la même, la topologie engendrée par la distance $|\cdot|$). On rappelle que les ouverts de \mathbb{R} sont des réunions dénombrables d'intervalles

ouverts $\bigcup_{i=1}^{+\infty}]a_i, b_i[$ (réunion finie ou dénombrable).

Remarque 6.2.3 La tribu borélienne $\mathcal{B}(X)$ de X contient

- les ouverts U_i ,
- les intersections $\cap_i U_i$ d'ouverts
- les réunions d'intersection $\cup_j \cap_i U_i$ d'ouverts
- en généralisant le procédé, les réunions d'intersections de réunions de ... d'ouverts

$$\dots \bigcap_i \bigcup_j \dots \bigcap_k \bigcup_l \dots \bigcap_m \bigcup_n \dots U_n$$

Comme le processus ne s'arrête pas, on ne peut pas décrire tous les boréliens par des réunions d'intersections etc d'ouverts. Par contre, dans le cas de $X = \mathbb{R}$, on peut optimiser la famille qui engendre $\mathcal{B}(\mathbb{R})$, c'est à dire choisir une famille encore plus petite que celle des ouverts qui suffit pour retrouver toute la tribu borélienne $\mathcal{B}(\mathbb{R})$ avec les opérations $\cup, \cap, ^c$.

Proposition 6.2.2 *Les boréliens de \mathbb{R} sont engendrés par*

- les ouverts
- les fermés
- les intervalles ouverts $]a, b[$
- les intervalles fermés $[a, b]$
- les intervalles semi ouverts $[a, b[,]a, b]$.

Remarque 6.2.4 Ces familles suffisent en appliquant les opérations licites dans les tribus (complémentaires, réunion, intersection) pour retrouver tous les ensembles de la tribu borélienne. Attention toutefois, cela n'empêche pas que certains ensembles de cette tribu peuvent être très bizarres (ensemble de Cantor). Cependant, pour nous, les boréliens seront la plupart du temps des intervalles finis ou non, fermés ou non :

$$[a, b], [a, b[,]a, b],]a, b[,] - \infty, b],] - \infty, b[, [a, +\infty[,]a, +\infty[.$$

De là vient qu'en L2, on a considéré pour observables de \mathbb{R} que les intervalles.

Exemples : La tribu borélienne $\mathcal{B}(\mathbb{R})$ contient aussi les singletons $\{a\}$, les ensembles dénombrables, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$.

6.3 Mesure

On a déjà évoqué que le cardinal (d'un ensemble), la longueur (d'une courbe), l'aire (d'une surface plane), le volume (d'un solide) ou encore la probabilité (d'un évènement) sont différentes façon de mesurer. Toutes ces notions sont des cas particuliers de la notion générale de mesure.

Ces mesures particulières sont associées aux types d'ensemble qu'elles mesurent. Pour une mesure abstraite, la famille d'éléments « mesurables » sera une tribu. On définit donc une mesure sur une tribu.

Considérons un espace mesurable (X, \mathcal{A}) .

Définition 6.3.1 Une mesure μ sur (X, \mathcal{A}) est une application de $\mathcal{A} \rightarrow [0, +\infty]$ telle que

- $\mu(\emptyset) = 0$
- si $(A_n)_{n \in \mathbb{N}}$ est une suite dénombrable d'ensembles de \mathcal{A} deux à deux disjoints alors

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mu(A_n) \quad \sigma\text{-additivité.}$$

En particulier, la mesure peut prendre la valeur $+\infty$ (ce n'est pas choquant, par exemple \mathbb{N} a un cardinal infini, et \mathbb{R} une longueur infinie), mais elle doit absolument être positive. Les ensembles $A \in \mathcal{A}$ s'appellent aussi les mesurables.

Définition 6.3.2 Le triplet (X, \mathcal{A}, μ) est appelé un espace mesuré (espace mesurable + mesure).

Remarque :

- Si l'espace X est discret, par exemple \mathbb{N} , $\mathcal{P}(X)$ est une bonne tribu à considérer.
- Si l'espace X est topologique, la tribu borélienne est une bonne tribu à considérer : par exemple pour \mathbb{R} , $\mathcal{B}(\mathbb{R})$.

Exemples :

- mesure de comptage (ou de dénombrement) sur $(X, \mathcal{P}(X))$:

$$\eta(A) = \begin{cases} \text{card } A & \text{si } A \text{ est fini} \\ +\infty & \text{sinon.} \end{cases}$$

- mesure de Dirac sur $(X, \mathcal{P}(X))$: soit $a \in X$,

$$\delta_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{si } a \notin A. \end{cases}$$

La mesure de Dirac δ_a indique si un ensemble contient ou non le point a .

Par exemple $\delta_0([-1, 1]) = 1$, $\delta_0(]0, 1]) = 0$, $\delta_0(\mathbb{R}^*) = 0$, $\delta_2(\mathbb{N}) = 1$.

- Exemple fondamental : la mesure de Lebesgue

Théorème 6.3.1 (Mesure de Lebesgue) Il existe une unique mesure λ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ telle que

- pour tout intervalle $[a, b]$ borné, on a

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

- stabilité par translation : pour tout $A \in \mathcal{B}(\mathbb{R})$, $\lambda(A+x) = \lambda(A)$ où $A+x$ est l'ensemble des $a+x$ pour $a \in A$.

La mesure de Lebesgue λ généralise la notion de longueur d'un intervalle à tous les boréliens $A \in \mathcal{B}(\mathbb{R})$. Il aurait été vain de chercher à généraliser la longueur à une mesure qui mesure toutes les parties de \mathbb{R} (c'est à dire une mesure sur tout $\mathcal{P}(\mathbb{R})$) car on montre qu'une telle généralisation sur $\mathcal{P}(\mathbb{R})$ contient une incohérence. Il faut se contenter d'une généralisation sur la tribu borélienne $\mathcal{B}(\mathbb{R})$.

• **Probabilité** Traditionnellement, dans le cadre probabiliste, on note $(\Omega, \mathcal{F}, \mathbb{P})$ plutôt que (X, \mathcal{A}, μ) . Dans $(\Omega, \mathcal{F}, \mathbb{P})$, Ω est un ensemble (dit espace de probabilité), \mathcal{F} une tribu appelée la famille des observables, et \mathbb{P} une mesure appelée probabilité.

Définition 6.3.3 Soit μ une mesure sur (X, \mathcal{A}) .

- Si $\mu(X) < +\infty$, alors la mesure μ est dite finie.
- Si X se décompose en $X = \bigcup_{n=1}^{+\infty} X_n$ avec $\mu(X_n) < +\infty$ alors la mesure est dite σ -finie.

Exemple sur \mathbb{R} , la mesure de Lebesgue est σ -finie car

$$\mathbb{R} = \bigcup_{n \in \mathbb{N}} [-n, n] \text{ et } \lambda([-n, n]) = 2n.$$

- Si $\mu(X) = 1$, alors la mesure μ est dite de probabilité.

Par exemple, une mesure de Dirac δ_a est une mesure de probabilité (dégénérée) car $\delta_a(X) = 1$.

Autre exemple : $([0, 1], \mathcal{B}([0, 1]), \lambda)$ est un espace de probabilité car $\lambda([0, 1]) = 1$.

Un exemple de mesure finie est la mesure de dénombrement sur un ensemble fini X puisque la mesure de X est $\text{card}(X)$ qui est fini.

6.4 Propriétés des mesures

- Une mesure μ est une application croissante : si $A \subset B$, alors $\mu(A) \leq \mu(B)$.
- Si $A, B \in \mathcal{A}$ avec $A \cap B = \emptyset$ alors

$$\mu(A \cup B) = \mu(A) + \mu(B) \quad (\text{additivité}).$$

- μ est sous-additive : si $A_n \in \mathcal{A}$ pour $n \in \mathbb{N}$, alors

$$\mu\left(\bigcup_{n=1}^{+\infty} A_n\right) \leq \sum_{n=1}^{+\infty} \mu(A_n) \quad (\text{sous-additivité}).$$

- $B \subset A$ alors $\mu(A \setminus B) = \mu(A) - \mu(B)$.
- Croissance séquentielle : si $A_n \in \mathcal{A}$ et $A_n \subset A_{n+1}$ alors

$$\mu\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_n \mu(A_n).$$

- Décroissance séquentielle : si $A_n \in \mathcal{A}$ et $A_{n+1} \subset A_n$ et si $\mu(A_0) < +\infty$ alors

$$\mu\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_n \mu(A_n).$$

Contre-exemple pour la décroissance sans hypothèse de finitude : sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, soit $A_n =]n, +\infty[$, on a $\bigcap_{n=1}^{+\infty} A_n = \emptyset$ de mesure nulle tandis que $\lim_{n \rightarrow +\infty} \lambda(A_n) = +\infty$.

Chapitre 7

Fonctions mesurables

7.1 Définition

Rappel (Image réciproque) Si $f : X \rightarrow Y$ est une application quelconque pour toute partie B de Y , l'image réciproque de B par f est

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}.$$

Si $(A_i)_{i \in I}$ et $(B_j)_{j \in J}$ sont des parties de X et Y respectivement, on rappelle que le comportement de f vis à vis de \cup, \cap et de c

$$f(\cup_{i \in I} A_i) = \cup_{i \in I} f(A_i) \quad f(\cap_{i \in I} A_i) \subset \cap_{i \in I} f(A_i)$$

tandis que le comportement de f^{-1} est

$$f^{-1}(\cup_{j \in J} B_j) = \cup_{j \in J} f^{-1}(B_j) \quad f^{-1}(\cap_{j \in J} B_j) = \cap_{j \in J} f^{-1}(B_j) \quad f^{-1}(B_j^c) = (f^{-1}(B_j))^c.$$

Définition 7.1.1 Soient (X, \mathcal{A}) , (Y, \mathcal{B}) deux espaces munis de tribus. Une fonction $f : X \rightarrow Y$ est dite mesurable si et seulement si $\forall B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$.

Remarque 7.1.1 • Cette définition est à comparer avec la définition de la continuité : l'image réciproque d'un ouvert doit être ouverte.

• Quand Y est un espace topologique et que rien n'est précisé, on prendra la tribu borélienne de Y .

Exemple : Une fonction indicatrice $\mathbf{1}_A$ d'un ensemble A est la fonction définie par

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases}.$$

Cette fonction ne prend donc que deux valeurs 1 ou 0 selon qu'elle est évaluée sur A ou non.

Vérifier que $\mathbf{1}_A \mathbf{1}_B = \mathbf{1}_{A \cap B}$, $1 - \mathbf{1}_A = \mathbf{1}_{A^c}$, si A et B sont disjoints que $\mathbf{1}_A + \mathbf{1}_B = \mathbf{1}_{A \cup B}$, si $A \subset B$, que $\mathbf{1}_A \leq \mathbf{1}_B$, et que $A = B$ ssi $\mathbf{1}_A = \mathbf{1}_B$.

Proposition 7.1.1 *La fonction indicatrice $\mathbf{1}_A$ de A est mesurable de (X, \mathcal{A}) dans \mathbb{R} quand A est mesurable (en tant qu'ensemble).*

Démonstration : En effet soit $B \in \mathcal{B}(\mathbb{R})$ alors $(\mathbf{1}_A)^{-1}(B) = \{x \in X \mid \mathbf{1}_A(x) \in B\}$ et

- si 0 et $1 \in B$ alors $(\mathbf{1}_A)^{-1}(B) = X \in \mathcal{A}$,
- si $1 \in B$ mais $0 \notin B$ alors $(\mathbf{1}_A)^{-1}(B) = A \in \mathcal{A}$,
- si $0 \in B$ mais $1 \notin B$ alors $(\mathbf{1}_A)^{-1}(B) = A^c \in \mathcal{A}$,
- si 0 et $1 \notin B$ alors $(\mathbf{1}_A)^{-1}(B) = \emptyset \in \mathcal{A}$,

Finalement on a pour tout $B \in \mathcal{A}$, $(\mathbf{1}_A)^{-1}(B) \in \mathcal{A}$: la fonction $\mathbf{1}_A$ est mesurable. \square

Proposition 7.1.2 *Si \mathcal{M} engendre la tribu \mathcal{B} de Y , f est mesurable ssi $f^{-1}(B) \in \mathcal{A}$ pour tout $B \in \mathcal{M}$.*

Démonstration : En effet notons \mathcal{C} l'ensemble des $B \in \mathcal{B}$ tels que $f^{-1}(B) \in \mathcal{A}$. Alors \mathcal{C} est une tribu de Y car

- $f^{-1}(Y) = X \in \mathcal{A}$ donc $Y \in \mathcal{C}$,
- si $B \in \mathcal{C}$, $f^{-1}(B^c) = \{x \in X \mid f(x) \in B^c\} = \{x \in X \mid f(x) \notin B\} = X \setminus \{x \in X \mid f(x) \in B\} = X \setminus f^{-1}(B) = (f^{-1}(B))^c \in \mathcal{A}$ car $f^{-1}(B) \in \mathcal{A}$ et \mathcal{A} est stable par complémentaire.
- Si $A_n, n \in \mathbb{N}$, sont dans \mathcal{C} alors $f^{-1}(A_n) \in \mathcal{A}$ d'où $f^{-1}(\cup_{n \in \mathbb{N}} A_n) = \cup_{n \in \mathbb{N}} f^{-1}(A_n) \in \mathcal{A}$. Et donc $\cup_{n \in \mathbb{N}} A_n \in \mathcal{A}$, qui est stable par réunion.

Finalement, \mathcal{C} est une tribu puis par hypothèse \mathcal{C} contient \mathcal{M} qui engendre \mathcal{B} . Donc $\mathcal{B} \subset \mathcal{C}$ et on a en particulier pour tout $B \in \mathcal{B}$, $f^{-1}(B) \in \mathcal{A}$, c'est à dire f est mesurable. \square

Quand $Y = \mathbb{R}$ ou \mathbb{C} (ou un espace topologique) muni de la tribu borélienne, f est mesurable

- ssi $\forall B \in \mathcal{B}(Y)$ alors $f^{-1}(B) \in \mathcal{A}$.
- ssi $\forall O$ ouvert, $f^{-1}(O) \in \mathcal{A}$
- ssi $\forall]a, +\infty[$ alors $f^{-1}(]a, +\infty[) \in \mathcal{A}$ dans le cas $Y = \mathbb{R}$.

Corollaire 7.1.1 *Une fonction continue de (X, \mathcal{T}) dans (Y, \mathcal{T}') est mesurable pour les tribus boréliennes $\mathcal{B}(X)$ et $\mathcal{B}(Y)$ associées à X et à Y .*

On connaît donc maintenant beaucoup de fonctions mesurables (pour les tribus boréliennes) : toutes les fonctions continues.

Démonstration : En effet, comme les ouverts de Y engendrent $\mathcal{B}(Y)$, il suffit de voir que l'image réciproque $f^{-1}(O)$ d'un ouvert O est dans $\mathcal{B}(X)$. Or par continuité de f , $f^{-1}(O)$ est ouvert dans X donc borélien. \square

7.2 Propriétés des applications mesurables

La mesurabilité des fonctions est une propriété stable par toutes les opérations classiques sur les fonctions :

Proposition 7.2.1 *Si $f : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ et $g : (Y, \mathcal{B}) \rightarrow (Z, \mathcal{C})$ sont mesurables alors $g \circ f : (X, \mathcal{A}) \rightarrow (Z, \mathcal{C})$ est mesurable.*

Démonstration : Soit $C \in \mathcal{C}$, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. Or par mesurabilité de g , $g^{-1}(C) \in \mathcal{B}$, puis par celle de f , $f^{-1}(g^{-1}(C)) \in \mathcal{A}$. \square

En particulier :

Proposition 7.2.2 *Si $f : (X, \mathcal{A}) \rightarrow Y$ espace topologique est mesurable et $g : Y \rightarrow Z$, espaces topologiques est continue alors $g \circ f : (X, \mathcal{A}) \rightarrow Z$ est mesurable.*

Démonstration : Soit $C \in \mathcal{T}_Z$, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$. Or par continuité de g , $g^{-1}(C) \in \mathcal{T}_Y$, puis par mesurabilité de f , $f^{-1}(g^{-1}(C)) \in \mathcal{A}$. \square

Par exemple si $f : (X, \mathcal{A}) \rightarrow \mathbb{C}$ est mesurable alors $|f|$, $\text{Im}(f)$, $\text{Re}(f)$ le sont car

$$\left\{ \begin{array}{l} \mathbb{C} \rightarrow \mathbb{R} \\ z \mapsto |z| \end{array} \right\}, \quad \left\{ \begin{array}{l} \mathbb{C} \rightarrow \mathbb{R} \\ z \mapsto \text{Re}(z) \end{array} \right\}, \quad \left\{ \begin{array}{l} \mathbb{C} \rightarrow \mathbb{R} \\ z \mapsto \text{Im}(z) \end{array} \right\} \quad \text{sont continues.}$$

Proposition 7.2.3 *Soit (X, \mathcal{A}) un espace mesurable et $f : (X, \mathcal{A}) \rightarrow \mathbb{R}$, $g : (X, \mathcal{A}) \rightarrow \mathbb{R}$ sont mesurables, alors $h = (f, g) : (X, \mathcal{A}) \rightarrow \mathbb{R}^2$ est mesurable.*

Démonstration : On munit \mathbb{R}^2 de la topologie produit pour laquelle les ouverts sont des produits d'ouverts $U \times V$. Soit donc $U \times V$ un ouvert produit de \mathbb{R}^2 , (f, g) est mesurable si

$$(f, g)^{-1}(U \times V) \in \mathcal{A}.$$

Or

$$\begin{aligned} (f, g)^{-1}(U \times V) &= \{x \in X \mid (f, g)(x) \in U \times V\} \\ &= \{x \in X \mid (f(x), g(x)) \in U \times V\} \\ &= \{x \in X \mid f(x) \in U, g(x) \in V\} \\ &= \{x \in X \mid f(x) \in U\} \cap \{x \mid g(x) \in V\} \\ &= f^{-1}(U) \cap g^{-1}(V) \in \mathcal{A} \end{aligned}$$

car $f^{-1}(U) \in \mathcal{A}$ et $g^{-1}(V) \in \mathcal{A}$ par mesurabilité de f, g . \square

On en déduit si f et g sont mesurables, a est scalaire

- af , $f + g$, $f - g$, $f \times g$, f/g (si $g(x) \neq 0, \forall x$), $\max(f, g)$, $\min(f, g)$ sont mesurables.
- Une combinaison linéaire de fonctions mesurables est mesurable.

- $f : (X, \mathcal{A}) \rightarrow \mathbb{C}$ est mesurable ssi $\operatorname{Re}(f)$ et $\operatorname{Im}(f)$ le sont.
- $f : (X, \mathcal{A}) \rightarrow \mathbb{R}$, est mesurable ssi $f^+ = \max(f, 0)$ et $f^- = \min(f, 0)$ le sont.

Exemple : (Fonction étagée ou étagée) Une fonction $f : (X, \mathcal{A}) \rightarrow \mathbb{R}$ est dite étagée si c'est une combinaison linéaire finie de fonctions indicatrices $\mathbf{1}_{A_i}$ pour des ensembles mesurables A_i deux à deux disjoints (pour simplifier) :

$$f(x) = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}, \quad \alpha_i \in \mathbb{R}, A_i \in \mathcal{A}.$$

Elle est étagée positive si en plus $\alpha_i \geq 0$ pour $i = 1, \dots, n$.

Comme on sait que $\mathbf{1}_{A_i}$ est mesurable, les fonctions étagées sont aussi mesurables (car combinaison linéaire de telles fonctions).

• Quand $X = \mathbb{R}$, les fonctions en escalier sont des cas particuliers de fonctions étagées avec des A_i égaux à des intervalles disjoints, elles sont donc en particulier mesurables.

• On considère \mathbb{N} muni de la tribu $\mathcal{P}(\mathbb{N})$ et les suites suivantes

- $u = (u_n)_n$ avec $u_n = 0$ pour $n \geq 6$,
- $v = (v_n)_n$ avec $v_n = 3$ pour $n > 4$,
- $w = (w_n)_n$

Montrer que u, v sont étagées mais pas w .

On énonce des résultats analogues sur \mathbb{C} et en fait sur tout espace topologique Y qui a une base dénombrable d'ouverts.

Définition 7.2.1 (Mesure image) Soit $f : (X, \mathcal{A}, \mu) \rightarrow (Y, \mathcal{B})$ une fonction mesurable. On définit sur (Y, \mathcal{B}) la mesure image de f notée μf^{-1} ou μ_f :

$$\mu_f(B) = \mu(f^{-1}(B)).$$

Topologie métrique sur $\overline{\mathbb{R}}$ On considère les ensembles $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$ et $[0, +\infty] = [0, +\infty[\cup \{+\infty\}$. On définit une distance (et donc une topologie métrique associée à cette distance) sur ces ensembles : Soit ϕ une bijection de $\overline{\mathbb{R}}$ sur un compact (par exemple $\phi = \arctan$ bijection de $\overline{\mathbb{R}}$ sur $[-\pi/2, \pi/2]$).

$$d(x, y) = |\phi(x) - \phi(y)|$$

avec $\arctan(\pm\infty) = \pm l$. Alors $(\overline{\mathbb{R}}, d)$ et $([0, +\infty], d)$ sont métriques.

Les boréliens associés à ces ensembles (avec la topologie définie par la métrique indiquée) sont engendrés par $\{]a, +\infty], a \in \mathbb{R}\}$.

Règles de calcul dans $[0, +\infty]$

$$\begin{aligned} a \times b &= ab & \text{si } a, b \neq +\infty, \\ a \times (+\infty) &= +\infty & \text{si } a \neq 0, \\ 0 \times (+\infty) &= 0 & \text{si } a = 0. \end{aligned}$$

Remarque 7.2.1 • Le produit n'est pas continu dans $[0, +\infty]$ en effet avec $a_n = n$ et $b_n = 1/n$ on a $a_n \rightarrow +\infty$, $b_n \rightarrow 0$ puis $a_n b_n = 1 \not\rightarrow ab = +\infty \times 0 = 0$.

• La convention $0 \times +(\infty) = 0$ est naturelle malgré tout quand on pense à $0 \times (+\infty)$ comme le calcul de la surface de \mathbb{R} de longueur $+\infty$ et de largeur 0.

7.3 Limite de fonctions mesurables

Proposition 7.3.1 Soient $f_n : (X, \mathcal{A}) \rightarrow \overline{\mathbb{R}}$ alors $\sup_n f_n$ et $\inf_n f_n$ sont mesurables.

Démonstration : On le montre pour le sup, le raisonnement s'adapterait facilement à l'inf.

$$\begin{aligned} (\sup_n f_n)^{-1}(]a, +\infty]) &= \{x \mid (\sup_n f_n)(x) \in]a, +\infty]\} = \{x \in X \mid \sup_n f_n(x) \in]a, +\infty]\} \\ &= \{x \in X \mid \sup_n f_n(x) > a\} = \{x \in X \mid \exists n \in \mathbb{N}, f_n(x) > a\} \\ &= \bigcup_{n \in \mathbb{N}} \{x \in X \mid f_n(x) > a\} = \bigcup_{n \in \mathbb{N}} f_n^{-1}(]a, +\infty]) \in \mathcal{A}. \end{aligned}$$

car pour chaque n , $f_n^{-1}(]a, +\infty]) \in \mathcal{A}$ qui est stable par réunion. \square

Rappel : limites inférieure et supérieure

On rappelle que pour une suite réelle $(u_n)_n$, la limite supérieure $\overline{\lim}_n u_n$ et la limite inférieure $\underline{\lim}_n u_n$ désignent respectivement la plus grande valeur d'adhérence et la plus petite valeur d'adhérence de la suite (u_n) . On les obtient par les formules suivantes :

$$\overline{\lim}_n u_n = \inf_k \sup_{n \geq k} u_n, \quad \underline{\lim}_n u_n = \sup_k \inf_{n \geq k} u_n.$$

Parmi les limites de toutes les sous-suites de $(u_n)_n$, $\underline{\lim}_n u_n$ est exactement la plus petite et $\overline{\lim}_n u_n$ est exactement la plus grande.

• Les limites inférieures et supérieures existent toujours (contrairement à la limite) mais elles peuvent valoir $\pm\infty$.

• On a toujours $\underline{\lim}_n u_n \leq \overline{\lim}_n u_n$.

• Les deux types de limites coïncident ssi la limite existe. Dans ce cas

$$\lim_{n \rightarrow +\infty} u_n = \underline{\lim}_n u_n = \overline{\lim}_n u_n.$$

Considérons par exemple

• la suite $(u_n)_n$ donnée par $u_{2n} = 1$ et $u_{2n+1} = 0$ alors sa limite supérieure est 1, sa limite inférieure est 0.

• la suite $(v_n)_n$ donnée par $v_n = \cos(n)$, on montre que sa limite supérieure est 1, sa limite inférieure est -1 .

• la suite $(w_n)_n$ donnée par $w_{2n} = n$ et $w_{2n+1} = \frac{2n}{n+1}$, on montre que sa limite supérieure est $+\infty$, sa limite inférieure est 2.

On en déduit maintenant :

Proposition 7.3.2 Soit (X, \mathcal{A}) un espace mesurable, $f_n : (X, \mathcal{A}) \rightarrow \overline{\mathbb{R}}$ ou $[0, +\infty]$ des fonctions mesurables. Alors $\sup_n f_n$, $\inf_n f_n$, $\overline{\lim}_n f_n$ et $\underline{\lim}_n f_n : (X, \mathcal{A}) \rightarrow \overline{\mathbb{R}}$ ou $[0, +\infty]$ sont mesurables.

Démonstration : Pour $\sup_n f_n$, on montre que $\{x \in X \mid \sup_n f_n(x) \in]a, +\infty[\} \in \mathcal{A}$. Or $\sup_n f_n(x) > a$ ssi $\exists n_0$ tel que $f_{n_0}(x) > a$. D'où

$$\{x \in X \mid \sup_n f_n(x) \in]a, +\infty[\} = \bigcup_{n \in \mathbb{N}} \underbrace{\{x \in X \mid f_n(x) > a\}}_{\text{mesurable car } f_n \text{ l'est}} \in \mathcal{A}.$$

On fait de même pour $\inf_n f_n$.

Pour $\overline{\lim}_n f_n$, on note $g_k = \sup_{n \geq k} f_n(x)$. D'après le résultat pour le sup, les fonctions g_k sont mesurables. Puis $\overline{\lim}_n f_n = \inf_k g_k$ est mesurable car inf de fonctions mesurables.

Pour $\underline{\lim}_n f_n$ enfin, l'argument est analogue. \square

On en déduit que la mesurabilité se garde même en passant à la limite :

Théorème 7.3.1 Soit (X, \mathcal{A}) un espace mesurable, $f_n : (X, \mathcal{A}) \rightarrow Y$, métrique, mesurable. On suppose que $\forall x \in X, f_n(x) \rightarrow f(x)$ (limite simple). Alors f est mesurable.

C'est un résultat très agréable si on le compare avec l'analogue pour la continuité où on a besoin de la convergence uniforme pour que la continuité se garde à la limite. Une fonction obtenue comme limite (simple) de mesurables est donc mesurable.

Démonstration : Si la limite simple $f(x)$ existe alors elle coïncide avec les limites inférieure et supérieure :

$$f(x) = \overline{\lim}_n f_n(x) = \underline{\lim}_n f_n(x).$$

La fonction f est égale à des fonctions mesurables, elle est donc mesurable.

On peut aussi le voir directement en utilisant le lemme suivant sur les suites réelles :

Lemme 7.3.1 Si une suite u_n converge vers l alors avec $v_m = \sup\{u_m, u_{m+1}, \dots\}$ on a

$$l = \inf\{v_0, v_1, \dots, v_n, \dots\} = \inf_m \sup_{k \geq m} u_k.$$

Ici, on a $f(x) = \inf_{m \in \mathbb{N}} \sup_{k \geq m} f_k(x)$. On a alors « f_n mesurable » implique « $\sup_{m \geq k} f_k$ mesurable » et donc « $f = \inf_m \sup_{m \geq k} f_k$ mesurable ». \square

7.4 Variables aléatoires

Grâce à la théorie de la mesure, on unifie la présentation du cadre discret et du cadre continu des variables aléatoires qui sont maintenant tout simplement des fonctions mesurables sur un espace de probabilité.

Définition 7.4.1 *Un espace de probabilité est un espace mesurable (Ω, \mathcal{F}) (notation typique des probabilistes) muni d'une mesure de probabilité \mathbb{P} , c'est à dire une mesure de masse totale 1 : $\mathbb{P}(\Omega) = 1$.*

Les ensembles mesurables $A \in \mathcal{F}$ sont appelés les évènements (ou les observables).

Définition 7.4.2 *On appelle variable aléatoire toute application mesurable X d'un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ dans \mathbb{R} .*

Il s'agit en fait d'une application mesurable sur un espace de probabilité.

Définition 7.4.3 *Soit $X : \Omega \rightarrow \mathbb{R}$, on appelle loi de X la mesure \mathbb{P}_X , mesure image sur \mathbb{R} de X par rapport à \mathbb{P} :*

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \mathbb{P}(\omega \in \Omega \mid X(\omega) \in A).$$

Définition 7.4.4 *Une variable aléatoire X est discrète si elle est à valeur dans un ensemble au plus dénombrable (en bijection avec une partie de \mathbb{N}) : $X(\Omega)$ est fini ou dénombrable (on peut compter ses éléments).*

Exemples :

- $\Omega = \{a_1, \dots, a_n\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ et \mathbb{P} définie par $\mathbb{P}(a_i) = p_i$ avec $p_1 + \dots + p_n = 1$.
- La variable X qui indique la face 1, ..., 6 obtenue par le lancer d'un dé est discrète.

$$X = \sum_{i=1}^6 i \mathbf{1}_{A_i}$$

où $A_i = \{\text{avoir un } i\}$.

- La variable Y qui indique le numéro du premier lancer où on obtient un 6 lors d'une suite infinie de lancer de dé est discrète (les valeurs possibles sont \mathbb{N}).

$$Y = \sum_{i=1}^{+\infty} i \mathbf{1}_{B_i}$$

où $B_i = \{\text{le premier 6 est au } i\text{ème lancer}\}$.

- La variable Z qui indique le nombre de succès dans une suite de n épreuves est discrète (et prend ses valeurs dans $\{1, \dots, n\}$).

$$Z = \sum_{i=0}^n i \mathbf{1}_{C_i}$$

où $C_i = \{\text{avoir } i \text{ succès en } n \text{ épreuves}\}$.

- Si $A \in \mathcal{F}$ est un évènement, alors $X = \mathbf{1}_A$ est une variable aléatoire. Elle vaut 1 si l'évènement A est réalisé 0 sinon.

Définition 7.4.5 (variable aléatoire à densité) X est une variable aléatoire de densité f si

$$\mathbb{P}(X \in A) = \int_A f(x)dx, \quad \mathbb{P}(X \in [a, b]) = \int_a^b f(x)dx.$$

Remarque 7.4.1 Pour l'instant, la fonction f est une fonction Riemann-intégrable et l'intégrale précédente est une intégrale au sens de Riemann. Bientôt, on pourra considérer des densités qui sont intégrables au sens de Lebesgue.

De toute façon, la densité f doit vérifier $f(x) \geq 0$ et $\int_{\mathbb{R}} f(x)dx = 1$.

Par exemple,

$$\frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad \frac{1}{b-a} \mathbf{1}_{[a,b]}, \quad \alpha e^{-\alpha x} \mathbf{1}_{\mathbb{R}^+}(x), \quad \frac{1}{\pi(1+x^2)}$$

sont des densités des lois normale standard $\mathcal{N}(0, 1)$, uniforme $\mathcal{U}([a, b])$, exponentielle $\mathcal{E}(\alpha)$ et de Cauchy $\mathcal{C}(1)$.

Chapitre 8

Intégrales des fonctions mesurables positives

Dans la première partie de ce cours, l'intégrale de Riemann a commencé par être définie pour les fonctions en escalier puis généralisée ensuite à une classe de fonctions plus grande.

Pour l'intégration par rapport à une mesure abstraite μ (on parle d'intégrale de Lebesgue), on commence aussi par définir l'intégrale pour des fonctions assez simples : les fonctions étagées.

On considère dans toute la suite un espace mesuré (X, \mathcal{A}, μ) .

8.1 Fonctions étagées (simples)

Définition 8.1.1 Une fonction indicatrice $\mathbf{1}_A$ d'un ensemble mesurable $A \in \mathcal{A}$ est la fonction définie par

$$\mathbf{1}_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases} .$$

Cette fonction ne prend donc que deux valeurs 1 ou 0 selon qu'elle est évaluée sur A ou non.

Une fonction indicatrice $\mathbf{1}_A$ d'un ensemble quelconque est mesurable ssi l'ensemble A l'est.

Remarque 8.1.1 Vérifier que $\mathbf{1}_A \mathbf{1}_B = \mathbf{1}_{A \cap B}$, $1 - \mathbf{1}_A = \mathbf{1}_{A^c}$, si A et B sont disjoints que $\mathbf{1}_A + \mathbf{1}_B = \mathbf{1}_{A \cup B}$, si $A \subset B$, que $\mathbf{1}_A \leq \mathbf{1}_B$, et que $A = B$ ssi $\mathbf{1}_A = \mathbf{1}_B$.

Définition 8.1.2 (Fonction simple ou étagée) Une fonction $f : (X, \mathcal{A}) \rightarrow \mathbb{R}^+$ est dite étagée positive si c'est une combinaison linéaire finie à coefficients positifs de fonctions indicatrices $\mathbf{1}_{A_i}$ pour des ensembles mesurables A_i deux à deux disjoints (pour simplifier) :

$$f(x) = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}, \quad \alpha_i \in \mathbb{R}^+, \quad A_i \in \mathcal{A}, \quad n \in \mathbb{N}.$$

• Cela ressemble à une fonction en escalier mais c'est plus général car pour une fonction en escalier les ensembles A_i devaient être des intervalles de \mathbb{R} , alors qu'ici, il s'agit d'ensembles mesurables quelconques sur X quelconque. En fait quand $X = \mathbb{R}$, les fonctions en escalier sont des cas particuliers de fonctions étagées (avec des A_i égaux à des intervalles disjoints, plutôt qu'à des ensembles mesurables généraux). On pourra penser aux fonctions en escalier quand on parlera de fonctions étagées.

• Une fonction étagée est mesurable car combinaison linéaire de fonctions indicatrices qui le sont clairement.

• Une fonction étagée prend un nombre fini de valeur : elle vaut α_i sur l'ensemble A_i .

• Si les α_i , $1 \leq i \leq n$, sont les valeurs possibles pour une fonction étagée f , alors avec $A_i = f^{-1}(\alpha_i)$, qui est mesurable par mesurabilité de f , on peut écrire

$$f = \sum_{i=1}^n \alpha_i \mathbf{1}_{f^{-1}(\alpha_i)} = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}.$$

• Une combinaison linéaire de fonctions étagées est encore une fonction étagée.

Exemples :

• Avec $E(x)$ désignant la partie entière de x , soit

$$f(x) = \begin{cases} 0 & x < 0 \\ E(x), & x \in [0, 5] \\ 3 & \text{sinon} \end{cases}$$

Montrer que f est étagée en l'écrivant comme combinaison linéaire de fonctions indicatrices.

• La fonction $x \mapsto E(x)$ est-elle étagée ?

• On lance une 5 fois un dé et on note X la variable aléatoire égale au nombre d'as obtenue. Montrer que X est étagée sur $(\Omega, \mathcal{F}, \mathbb{P})$.

• Dans l'exemple précédent, X est elle encore étagée si on lance le dé une infinité de fois ?

• On définit une variable aléatoire X de la manière suivante : on joue à pile ou face. Si pile est sorti au moins une fois avant le 100ème coup, X est le numéro du coup où pile est sorti la première fois. Sinon, $X = 100$. Montrer que X est une variable aléatoire étagée positive.

• Même exemple où X est le numéro du premier pile que ce soit avant ou après le 100ème coup. La variable X est-elle encore étagée ?

• On considère \mathbb{N} muni de la tribu $\mathcal{P}(\mathbb{N})$ et les suites suivantes

– $u = (u_n)_n$ avec $u_n = 0$ pour $n \geq 6$,

– $v = (v_n)_n$ avec $v_n = 3$ pour $n > 4$,

– $w = (w_n)_n$

Montrer que u, v sont étagées mais pas w .

Pour l'intégrale de Riemann, on a défini une fonction Riemann intégrable quand elle était (uniformément) approchable par une fonction en escalier et l'intégrale se définit comme la limite de celles en escalier.

Les fonctions étagées vont jouer un rôle analogue en théorie de la mesure de Lebesgue, mais les choses se passent mieux : toute fonction mesurable est limite de fonctions étagées.

Théorème 8.1.1 *Soit $f : (X, \mathcal{A}) \rightarrow [0, +\infty]$ mesurable. Alors, il existe des fonctions étagées s_n telles que la suite $(s_n)_n$ est croissante ($s_n(x) \leq s_{n+1}(x)$) majorée par f et convergeant simplement vers f*

$$\forall x, \lim_{n \rightarrow +\infty} s_n(x) = f(x).$$

Démonstration : Soit $n \geq 1$, et pour $i \in \{1, \dots, n2^n\}$, $A_{i,n} = f^{-1}([\frac{i-1}{2^n}, \frac{i}{2^n}[)$, $B_n = f^{-1}([n, +\infty[)$. $A_{i,n}$ et $B_n \in \mathcal{A}$ car f est mesurable. On prend alors

$$s_n = \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \mathbf{1}_{A_{i,n}} + n \mathbf{1}_{B_n}.$$

C'est une suite de fonctions étagées dont on peut vérifier qu'elle convient : elle vérifie en effet

- 1) $s_n(x) \leq s_{n+1}(x)$
- 2) $\lim_{n \rightarrow +\infty} s_n(x) = f(x)$
- 3) $s_n(x) \leq f(x)$.

Pour le point 1) : si $x \in B_n$ alors $f(x) \geq n$. Si $f(x) \geq n+1$ alors $s_{n+1}(x) = n+1 \geq s_n(x)$ et si $n \leq f(x) < n+1$ alors n s'écrit $(i-1)/2^{n+1}$ et $s_{n+1}(x) \geq n = s_n(x)$. Puis pour $x \in A_{i,n}$, on a

$$f(x) \in \left[\frac{2(i-1)}{2^{n+1}}, \frac{2i-1}{2^{n+1}} \right[\cup \left[\frac{2i-1}{2^{n+1}}, \frac{2i}{2^{n+1}} \right[$$

mais alors $s_{n+1} = \frac{2(i-1)}{2^{n+1}} = s_n(x)$ ou $s_{n+1} = \frac{2i-1}{2^{n+1}} \geq s_n(x)$. En tout cas, $s_n \leq s_{n+1}$.

Pour le 2) et 3) : si $f(x) = +\infty$ alors $s_n(x) = n$ pour tout n et $s_n(x) \rightarrow +\infty = f(x)$. Si $f(x)$ est bornée alors pour n assez grand $f(x) \leq n$ et pour ces n , on a $x \in A_{i,n}$, c'est à dire $s_n(x) = (i-1)/2^n$ avec $f(x) \in [\frac{i-1}{2^n}, \frac{i}{2^n}[$. On a donc $f(x) - s_n(x) < 1/2^n$ et $f(x) = \lim_{n \rightarrow +\infty} s_n(x)$. \square

Ce résultat va permettre de définir l'intégrale d'une fonction mesurable à partir de celle des fonctions étagées car toute fonction mesurable est limite de fonctions étagées croissantes.

8.2 Intégrale des fonctions positives

Définition 8.2.1 *Soit $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ une fonction étagée positive ($\alpha_i \geq 0$), on définit l'intégrale de f par rapport à la mesure μ par*

$$\int f d\mu = \int_X f d\mu = \sum_{i=1}^n \alpha_i \mu(A_i).$$

L'intégrale de f sur une partie mesurable $E \in \mathcal{A}$ de X est

$$\int_E f d\mu = \int (f \times \mathbf{1}_E) d\mu = \sum_{i=1}^n \alpha_i \mu(A_i \cap E).$$

Remarque 8.2.1 • $(\sum_{i=1}^n \mathbf{1}_{A_i}) \mathbf{1}_E = \sum_{i=1}^n (\mathbf{1}_{A_i} \mathbf{1}_E) = \sum_{i=1}^n \mathbf{1}_{A_i \cap E}$.

• L'intégrale de f étagée ne dépend pas de son écriture sous forme de combinaison linéaire de fonctions étagées, une autre écriture de f donne la même valeur à l'intégrale : cette définition a donc bien un sens.

• Par exemple

$$\int (2\mathbf{1}_A + 3\mathbf{1}_B + \frac{1}{2}\mathbf{1}_C) d\mu = 2\mu(A) + 3\mu(B) + \frac{1}{2}\mu(C).$$

• Noter en particulier que pour une fonction indicatrice :

$$\int \mathbf{1}_A d\mu = \mu(A), \quad \int_E \mathbf{1}_A d\mu = \mu(A \cap E).$$

Ce résultat est à la fois élémentaire et important : il fait le lien entre une mesure $\mu(A)$ et une intégrale $\int \mathbf{1}_A d\mu$, toute mesure d'un ensemble peut se voir comme une intégrale.

Définition 8.2.2 Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ mesurable. On définit son intégrale comme le sup des intégrales de fonctions étagées majorées par f :

$$\int f d\mu = \int_X f d\mu = \sup \left\{ \int s d\mu \mid s \text{ étagée} \leq f \right\}.$$

Pour une partie mesurable $E \in \mathcal{A}$ de X , on définit son intégrale sur E par

$$\int_E f d\mu = \int f \mathbf{1}_E d\mu = \sup \left\{ \int_E s d\mu \mid s \text{ étagée} \leq f \right\}.$$

Définition 8.2.3 Une fonction mesurable positive est dite μ -intégrable si $\int f d\mu < +\infty$.

Exemple : • On considère (X, \mathcal{A}) muni de la mesure δ_a . Dans ce cas,

$$\int f d\delta_a = f(a).$$

En effet, si $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ (avec $(A_i)_i$ une partition de X est étagée alors il existe un unique indice i_0 tel que $a \in A_{i_0}$ et

$$\int f d\delta_a = \sum_{i=1}^n \alpha_i \delta_a(A_i) = \alpha_{i_0} \delta_a(A_{i_0}) = \alpha_{i_0} = f(a).$$

car $f(a) = \alpha_{i_0} \mathbf{1}_{A_{i_0}}(a) = \alpha_{i_0}$. De façon générale,

$$\begin{aligned} \int f d\delta_a &= \sup \left(\int s d\delta_a \mid s \text{ étagée} \leq f \right) \\ &= \sup (s(a) \mid s \text{ étagée} \leq f) \\ &= f(a). \end{aligned}$$

• On considère $(X, \mathcal{A}, \mu) = (\mathbb{N}, \mathcal{P}(\mathbb{N}), \eta)$ où η est la mesure de dénombrement, alors si on se donne une suite $u = (u_n)$, on peut l'écrire $u = \sum_{n=0}^{+\infty} u_n \mathbf{1}_{\{n\}}$. Ainsi, $u(k) = \sum_{n=0}^{+\infty} u_n \mathbf{1}_{\{n\}}(k) = u_k$ car le seul terme non nul dans la somme est $\mathbf{1}_{\{k\}}(k) = 1$. On a alors

$$\int_{\mathbb{N}} u d\eta = \sum_{n=0}^{+\infty} \int_{\{n\}} u d\eta = \sum_{n=0}^{+\infty} u(n) \eta(\{n\}) = \sum_{n=0}^{+\infty} u_n.$$

Autrement dit une série se voit comme une intégrale par rapport à une mesure discrète : la mesure de dénombrement.

Une série positive converge ssi la suite associée est intégrable pour la mesure de dénombrement.

8.3 Propriétés de l'intégrale

On considère des fonctions mesurables positives f et g et des ensembles E, F mesurables :

• **(Croissance 1)** Si $f \leq g$ sur X , alors $\int f d\mu \leq \int g d\mu$.

En effet si $f \leq g$ toute fonction étagée s majorée par f l'est aussi par g si bien que le sup qui définit $\int g d\mu$ est pris sur un ensemble plus grand que celui définissant $\int f d\mu$, il est donc plus grand :

$$\int f d\mu = \sup \left\{ \int s d\mu \mid s \text{ étagée} \leq f \right\} \leq \sup \left\{ \int s d\mu \mid s \text{ étagée} \leq g \right\} = \int g d\mu$$

car $\{s \text{ étagée} \leq f\} \subset \{s \text{ étagée} \leq g\}$.

Proposition 8.3.1 (Inégalité de Markov) Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ mesurable et $\alpha > 0$, alors

$$\mu(x \in X \mid f(x) \geq \alpha) \leq \frac{\int_X f d\mu}{\alpha}.$$

Démonstration : On a

$$\alpha \mathbf{1}_{\{x, f(x) \geq \alpha\}}(x) \leq f(x) \tag{8.1}$$

car soit x vérifie $f(x) \geq \alpha$ et (8.1) devient $\alpha \leq f(x)$, donc est vraie ; soit x ne vérifie pas $f(x) \geq \alpha$ et (8.1) devient $0 \leq f(x)$, ce qui est encore vraie car $f \geq 0$.

En intégrant l'inégalité (8.1), il vient

$$\alpha \int_X \mathbf{1}_{\{x \mid f(x) \geq \alpha\}} d\mu \leq \int_X f d\mu.$$

Or $\int_X \mathbf{1}_{\{x \mid f(x) \geq \alpha\}} d\mu = \mu(x \in X \mid f(x) \geq \alpha)$. □

• **(Croissance 2)** Si $E \subset F$ alors $\int_E f d\mu \leq \int_F f d\mu$.

Comme $E \subset F$, ensembles mesurables, on a $\mathbf{1}_E \leq \mathbf{1}_F$. Puis comme f est une fonction positive, on a aussi $f\mathbf{1}_E \leq f\mathbf{1}_F$. Et donc par croissance de l'intégrale (premier point) $\int f\mathbf{1}_E d\mu \leq \int f\mathbf{1}_F d\mu$, c'est à dire pour f mesurable positive :

$$E \subset F \implies \int_E f d\mu \leq \int_F f d\mu.$$

• **(Nullité 1)** Si $f(x) = 0$ pour $x \in E$ alors $\int_E f d\mu = 0$.

En effet $\int_E f d\mu = \int f\mathbf{1}_E d\mu = \sup \int s\mathbf{1}_E d\mu$ où le sup est pris sur les fonctions étagées positives s majorées par $f\mathbf{1}_E : s(x) \leq f(x)\mathbf{1}_E(x)$.

Or sur E , f est nulle donc nécessairement, $s(x) = 0$ sur E et pour les fonctions s à considérer l'intégrale est $\int s\mathbf{1}_E d\mu = 0$, dont le sup ne peut manquer d'être aussi 0.

• **(Nullité 2)** Si $\mu(E) = 0$ alors $\int_E f d\mu = 0$.

En effet

$$\int_E f d\mu = \int f\mathbf{1}_E d\mu = \sup \int s\mathbf{1}_E d\mu$$

où le sup est pris sur les fonctions étagées positives s majorées par $f\mathbf{1}_E$. Or pour une telle fonction $s = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$, on a

$$\int s\mathbf{1}_E d\mu = \sum_{i=1}^n a_i \mu(A_i \cap E) = 0$$

car $\mu(A_i \cap E) \leq \mu(E) = 0$. Finalement, on prend le sup sur 0, ce qui donne 0.

• $\int_E f d\mu = \int f\mathbf{1}_E d\mu$: c'est la définition de l'intégrale sur E .

• **(Linéarité 1)** Si $c \geq 0$ alors $\int c f d\mu = c \int f d\mu$.

C'est clair d'abord si $f = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ est une fonction étagée car alors cf l'est aussi et

$$\int c f d\mu = \sum_{i=1}^n c \alpha_i \mu(A_i) = c \sum_{i=1}^n \alpha_i \mu(A_i) = c \int f d\mu.$$

Puis si f est quelconque, s est une fonction étagée majorée par f ssi cs en est une majorée par cf . Donc l'ensemble des fonctions étagées majorées par cf est exactement $\{cs, s \text{ étagée } \leq f\}$,

$$\int c f d\mu = \sup \left\{ \int s' d\mu \mid s' \text{ étagée } \leq c f \right\}$$

$$\begin{aligned}
&= \sup\left\{\int s'd\mu \mid s' = cs, s \text{ étagée} \leq f\right\} \\
&= \sup\left\{\int csd\mu \mid s \text{ étagée} \leq f\right\} \\
&= \sup\left\{c \int sd\mu \mid s \text{ étagée} \leq f\right\} \\
&= c \sup\left\{\int sd\mu \mid s \text{ étagée} \leq f\right\} \\
&= c \int fd\mu
\end{aligned}$$

où on a utilisé le fait (déjà vérifié) pour une fonction étagée que $\int csd\mu = c \int sd\mu$.

• **(Linéarité 2)** $\int(f + g)d\mu = \int fd\mu + \int gd\mu$ (On le prouve d'abord pour des fonctions étagées puis on utilisera le théorème de convergence monotone (Beppo Levi) pour généraliser).

D'abord donc si f et g sont étagées :

$$f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}, \quad g = \sum_{j=1}^p b_j \mathbf{1}_{B_j}$$

avec les A_i deux à deux disjoints et les B_j aussi. On peut réécrire f et g comme combinaisons des mêmes indicatrices $\mathbf{1}_{A_i \cap B_j}$ $1 \leq i \leq n, 1 \leq j \leq p$:

$$f = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} a_i \mathbf{1}_{A_i \cap B_j}, \quad g = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} b_j \mathbf{1}_{A_i \cap B_j}.$$

D'où $f + g = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} a_i \mathbf{1}_{A_i \cap B_j} + \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} b_j \mathbf{1}_{A_i \cap B_j} = \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} (a_i + b_j) \mathbf{1}_{A_i \cap B_j}$ et

$$\begin{aligned}
\int (f + g)d\mu &= \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} (a_i + b_j) \mu(A_i \cap B_j) \\
&= \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} a_i \mu(A_i \cap B_j) + \sum_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} b_j \mu(A_i \cap B_j) \\
&= \sum_{i=1}^n a_i \sum_{j=1}^p \mu(A_i \cap B_j) + \sum_{j=1}^p b_j \sum_{i=1}^n \mu(A_i \cap B_j) \\
&= \sum_{i=1}^n a_i \mu(A_i \cap \cup_{j=1}^p B_j) + \sum_{j=1}^p b_j \mu(\cup_{i=1}^n A_i \cap B_j) \\
&= \sum_{i=1}^n a_i \mu(A_i) + \sum_{j=1}^p b_j \mu(B_j)
\end{aligned}$$

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

Si f et g sont des fonctions mesurables positives quelconques, soient $(s_n)_n$ et $(t_n)_n$ des suites croissantes de fonctions étagées positives qui convergent vers f et g . Alors $s_n + t_n$ sont aussi des fonctions étagées positives et elles convergent vers $f + g$. D'après le théorème de convergence monotone (cf Théorème 9.1.1), $\int s_n d\mu$, $\int t_n d\mu$, $\int (s_n + t_n) d\mu$ convergent respectivement vers $\int f d\mu$, $\int g d\mu$, $\int (f + g) d\mu$. Donc la linéarité vient en passant à la limite dans $\int (s_n + t_n) d\mu = \int s_n d\mu + \int t_n d\mu$.

• **(Relation de Chasles)** Si E et F sont mesurables disjoints, alors pour une fonction mesurable

$$\int_{E \cup F} f d\mu = \int_E f d\mu + \int_F f d\mu.$$

On le prouve pour l'instant pour les fonctions étagées positives : soit $f = \sum_{i=1}^n a_i \mathbf{1}_{A_i}$ étagée alors

$$\begin{aligned} \int_{E \cup F} f d\mu &= \int f \mathbf{1}_{E \cup F} d\mu = \sum_{i=1}^n a_i \mu(A_i \cap (E \cup F)) \\ &= \sum_{i=1}^n a_i \mu((A_i \cap E) \cup (A_i \cap F)) \\ &= \sum_{i=1}^n a_i \mu(A_i \cap E) + \sum_{i=1}^n a_i \mu(A_i \cap F) \\ &= \int_E f d\mu + \int_F f d\mu. \end{aligned}$$

Si f est mesurable positive, alors, on considère $(s_n)_n$ une suite croissante de fonctions étagées positives qui converge vers f et l'égalité vient en passant à la limite dans celle de s_n par convergence monotone.

On généralise ensuite aisément au cas de fonctions mesurables à valeurs dans \mathbb{R} ou \mathbb{C} .

Chapitre 9

Théorèmes limites pour l'intégrale des fonctions mesurables positives

Avec cette nouvelle notion d'intégrale, les théorèmes de convergence pour les intégrales que nous allons voir sont beaucoup plus forts.

9.1 Convergence monotone

Théorème 9.1.1 (Convergence monotone, Beppo Levi) Soit (X, \mathcal{A}, μ) un espace mesuré et $f_n : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ une suite de fonctions mesurables, croissante ($f_n \leq f_{n+1}$).

On pose $f(x) = \lim_{n \rightarrow +\infty} f_n(x)$ la limite étagée dans $[0, +\infty]$. Alors

$$\int f d\mu = \lim_{n \rightarrow +\infty} \int f_n d\mu. \quad (9.1)$$

Remarque 9.1.1 • Dans l'énoncé, c'est la suite $(f_n)_n$ qui est croissante (i.e. $f_n(x) \leq f_{n+1}(x)$) et non pas les fonctions f_n (ce n'est pas $f_n(x) \leq f_n(y)$ pour $x \leq y$).

- La limite simple (ponctuelle) f des f_n peut aussi prendre la valeur $+\infty$!

On a d'abord besoin d'un lemme :

Lemme 9.1.1 Soit s étagée alors $\varphi : E \mapsto \int_E s d\mu$ est une mesure.

Démonstration : On commence par vérifier les deux axiomes d'une mesure.

- $\varphi(\emptyset) = \int_{\emptyset} s d\mu = 0$ car on obtient 0 en intégrant sur un ensemble de mesure 0.
- Soient $E_k, k \in \mathbb{N}$, des ensembles mesurables deux à deux disjoints de réunion E . On suppose que s s'écrit

$$s = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}.$$

Alors

$$\varphi(E) = \varphi(\cup_k E_k) = \int_{\cup_k E_k} s d\mu = \sum_{i=1}^n \alpha_i \mu(A_i \cap (\cup_k E_k)) = \sum_{i=1}^n \alpha_i \mu(\cup_k (A_i \cap E_k))$$

$$\begin{aligned}
&= \sum_{i=1}^n \alpha_i \sum_k \mu(A_i \cap E_k) = \sum_k \sum_{i=1}^n \alpha_i \mu(A_i \cap E_k) = \sum_k \int_{E_k} s d\mu \\
&= \sum_k \varphi(E_k).
\end{aligned}$$

L'application φ est donc une mesure. □

Passons à la preuve du théorème de convergence monotone 9.1.1.

Démonstration : On a pour tout x et tout n , $f_n(x) \leq f_{n+1}(x)$ et en passant à la limite $f_n(x) \leq f(x)$, ce qui donne en intégrant

$$\int_X f_n d\mu \leq \int_X f d\mu. \quad (9.2)$$

D'autre part, on a aussi $\int_X f_n d\mu \leq \int_X f_{n+1} d\mu$ donc $\int_X f_n d\mu$ est une suite (numérique) croissante de $[0, +\infty]$. Elle a donc une limite, notée $\alpha \in [0, +\infty]$:

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \alpha.$$

En passant à la limite dans (9.2), on a déjà $\alpha \leq \int_X f d\mu$.

Soit maintenant, s une fonction étagée majorée par f et $c \in]0, 1[$, on introduit les ensembles mesurables

$$E_n = \{x \in X \mid f_n(x) \geq cs(x)\}.$$

Comme $(f_n)_n$ est une suite croissante, on constate que $E_n \subset E_{n+1}$ car si $f_n(x) \geq cs(x)$ a fortiori $f_{n+1}(x) \geq cs(x)$. De plus $f_n(x) \rightarrow f(x)$ ce qui donne l'existence d'un entier N tel que pour tout $n \geq N$, $f_n(x) \geq cf(x) \geq cs(x)$. Autrement dit

$$X = \bigcup_{n=1}^{+\infty} E_n.$$

Avec $\varphi(E) = \int_E s d\mu$ qui est une mesure d'après le lemme précédent, on a

$$\int_X f_n d\mu \geq \int_{E_n} f_n d\mu \geq \int_{E_n} cs d\mu = c \int_{E_n} s d\mu = c\varphi(E_n). \quad (9.3)$$

Comme $(E_n)_n$ est croissante pour l'inclusion, on a par croissance monotone de la mesure φ ,

$$\lim_{n \rightarrow +\infty} \varphi(E_n) = \varphi(\cup_n E_n) = \varphi(X) = \int_X s d\mu,$$

on a donc en passant à la limite dans (9.3) :

$$\alpha \geq c \int_X s d\mu$$

Comme c'est vrai pour tout $c \in]0, 1[$, en faisant tendre c vers 1, on obtient $\alpha \geq \int_X s d\mu$ puis en prenant le sup sur les fonctions s étagées majorées par f , on déduit $\alpha \geq \int_X f d\mu$.

Finalement, on a obtenu

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \alpha = \int_X f d\mu.$$

□

Quelques conséquences de la convergence monotone :

Corollaire 9.1.1 (Linéarité) Soient f et g des fonctions mesurables positives, on a

$$\int (f + g) d\mu = \int f d\mu + \int g d\mu.$$

Démonstration : Si f et g sont des fonctions mesurables positives quelconques, soient $(s_n)_n$ et $(t_n)_n$ des suites croissantes de fonctions étagées positives qui convergent vers f et g . Alors $s_n + t_n$ sont aussi des fonctions étagées positives et elles convergent vers $f + g$. D'après le théorème de convergence monotone (cf Théorème 9.1.1), $\int s_n d\mu$, $\int t_n d\mu$ et $\int (s_n + t_n) d\mu$ convergent respectivement vers $\int f d\mu$, $\int g d\mu$, $\int (f + g) d\mu$. Donc la linéarité vient en passant à la limite dans $\int (s_n + t_n) d\mu = \int s_n d\mu + \int t_n d\mu$, due à la linéarité déjà vue pour les fonctions étagées. □

Corollaire 9.1.2 (Relation de Chasles) Soient E et F des ensembles mesurables dis-joints et f une fonction mesurable positive

$$\int_{E \cup F} f d\mu = \int_E f d\mu + \int_F f d\mu.$$

Démonstration : Si f est mesurable positive, alors, on considère $(s_n)_n$ une suite croissante de fonctions étagées positives qui converge vers f et l'égalité vient en passant à la limite dans celles déjà vues pour les fonctions étagées s_n . On généralise ensuite aisément au cas de fonctions mesurables à valeurs dans \mathbb{R} ou dans \mathbb{C} . □

Corollaire 9.1.3 Soient f_n une suite de fonctions mesurables de X dans $[0, +\infty]$ alors

$$\int \sum_{n=1}^{+\infty} f_n d\mu = \sum_{n=1}^{+\infty} \int f_n d\mu.$$

Démonstration : Appliquer le théorème de convergence monotone à $g_p = \sum_{n=0}^p f_n$ qui est mesurable, et forme une suite croissante car $g_{p+1} - g_p = f_p \geq 0$. Comme la limite g de g_p est $g = \sum_{n=0}^{+\infty} f_n$, on a

$$\int \sum_{n=1}^{+\infty} f_n d\mu = \int g d\mu = \lim_{p \rightarrow +\infty} \int g_p d\mu = \lim_{p \rightarrow +\infty} \int \sum_{n=0}^p f_n d\mu = \lim_{p \rightarrow +\infty} \sum_{n=0}^p \int f_n d\mu = \sum_{n=0}^{+\infty} \int f_n d\mu$$

où on a juste utiliser la linéarité de l'intégrale pour échanger la somme finie $\sum_{n=0}^p$ et \int . \square

Corollaire 9.1.4 Soit $f : (X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ mesurable alors la fonction sur \mathcal{A}

$$\varphi(E) = \int_E f d\mu$$

est une mesure (elle est finie ssi f est intégrable). Puis pour une fonction mesurable g

$$\int_X g d\varphi = \int_X g f d\mu. \quad (9.4)$$

Remarque 9.1.2 • En quelque sorte, on a $d\varphi = f d\mu$.

• Le théorème de convergence monotone est la clef de nombreux raisonnements typiques. Pour justifier une propriété, souvent, on la montre d'abord pour les fonctions indicatrices $\mathbf{1}_A$, on la généralise par linéarité aux fonctions étagées puis enfin aux fonctions mesurables quelconques par convergence monotone.

Illustrons cette façon de faire par la deuxième partie de cette preuve.

Démonstration : φ est une mesure car

- $\varphi(\emptyset) = \int_{\emptyset} f d\mu = 0$ car on obtient 0 en intégrant sur un ensemble de mesure 0.
- Soient $E_k, k \in \mathbb{N}$, des ensembles mesurables deux à deux disjoints de réunion E . On a $\mathbf{1}_E = \sum_k \mathbf{1}_{E_k}$ car les E_k sont deux à deux disjoints

$$\varphi(E) = \int_X \mathbf{1}_E f d\mu = \int_X \sum_k \mathbf{1}_{E_k} f d\mu = \sum_k \int_X \mathbf{1}_{E_k} f d\mu = \sum_k \varphi(E_k).$$

L'application φ est donc une mesure.

Pour la seconde partie, si $g = \mathbf{1}_A$ est une indicatrice, on a

$$\int_X \mathbf{1}_A d\varphi = \varphi(A) = \int_X \mathbf{1}_A f d\mu = \int_X g f d\mu$$

et (9.4) est satisfaite dans ce cas. Pour une fonction étagée, ça reste vraie par linéarité en utilisant le premier cas : si $g = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}$ alors

$$\int_X g d\varphi = \int_X \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i} d\varphi = \sum_{i=1}^n \alpha_i \int_X \mathbf{1}_{A_i} d\varphi = \sum_{i=1}^n \alpha_i \int_X \mathbf{1}_{A_i} f d\mu = \int_X \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i} f d\mu = \int_X g f d\mu.$$

Pour g mesurable positive, on prend $(s_n)_n$ suite de fonctions étagées croissantes qui converge vers g et on applique le théorème de convergence monotone

$$\lim_{n \rightarrow +\infty} \int_X s_n d\varphi = \int_X g d\varphi, \quad \lim_{n \rightarrow +\infty} \int_X s_n f d\mu = \int_X g f d\mu$$

car $s_n \nearrow g$ et $s_n f \nearrow g f$, ($f \geq 0$). Pour s_n , d'après le cas étagé, on a

$$\int_X s_n d\varphi = \int_X s_n f d\mu$$

ce qui donne en passant à la limite par convergence monotone :

$$\int_X g d\varphi = \int_X g f d\mu.$$

Si g est de signe quelconque on écrit $g = g^+ - g^-$ et on utilise le cas positif et la linéarité pour obtenir (9.4) dans ce cas le plus général.

Pour g complexe, on écrit $g = \operatorname{Re}(g) + i\operatorname{Im}(g)$ et on applique le cas réel à chaque terme réel. \square

9.2 Lemme de Fatou

Rappel (liminf et limsup) : Pour une suite réelle $u = (u_n)_n$, on définit ses limites supérieures et inférieures

$$\begin{aligned} \underline{\lim}_{n \rightarrow +\infty} u_n &= \sup_n \inf_{k \geq n} u_k, \\ \overline{\lim}_{n \rightarrow +\infty} u_n &= \inf_n \sup_{k \geq n} u_k. \end{aligned}$$

Ce sont les plus petites et plus grandes valeur d'adhérence de la suite u . On a toujours

$$\underline{\lim}_{n \rightarrow +\infty} u_n \leq \overline{\lim}_{n \rightarrow +\infty} u_n$$

et il y a égalité ssi la suite u converge ; de plus si tel est le cas

$$\lim_{n \rightarrow +\infty} u_n = \underline{\lim}_{n \rightarrow +\infty} u_n = \overline{\lim}_{n \rightarrow +\infty} u_n.$$

En plus, en changeant le signe, les limites inférieure et supérieure s'échangent :

$$\begin{aligned} \underline{\lim}_{n \rightarrow +\infty} (-u_n) &= -\overline{\lim}_{n \rightarrow +\infty} u_n, \\ \overline{\lim}_{n \rightarrow +\infty} (-u_n) &= -\underline{\lim}_{n \rightarrow +\infty} u_n. \end{aligned}$$

Pour une suite de fonctions $(f_n)_n$, on définit des fonctions limites inférieure et supérieure de la façon suivante :

$$(\underline{\lim}_{n \rightarrow +\infty} f_n)(x) = \underline{\lim}_{n \rightarrow +\infty} f_n(x), \quad (\overline{\lim}_{n \rightarrow +\infty} f_n)(x) = \overline{\lim}_{n \rightarrow +\infty} f_n(x).$$

Théorème 9.2.1 (Lemme de Fatou) Soit $(f_n)_n : (X, \mathcal{A}, \mu) \rightarrow [0, \infty]$ une suite de fonctions mesurables positives, alors

$$\int \underline{\lim}_{n \rightarrow +\infty} f_n d\mu \leq \underline{\lim}_{n \rightarrow +\infty} \int f_n d\mu.$$

Corollaire 9.2.1 Soit $(f_n)_n$ une suite de fonctions mesurables positives qui converge simplement vers f et telle que la suite des intégrales $\int f_n d\mu$ est majorée par M alors

$$\int f d\mu \leq M.$$

Démonstration : $\underline{\lim}_n f_n = \sup_n \inf_{k \geq n} f_k = \sup_n g_n$ avec $g_n = \inf_{k \geq n} f_k$. Les fonctions g_n sont mesurables, positives et $g_n \leq g_{n+1}$. Par le théorème de convergence monotone 9.1.1,

$$\int \lim_{n \rightarrow +\infty} g_n d\mu = \lim_{n \rightarrow +\infty} \int g_n d\mu.$$

Or comme la suite $(g_n)_n$ est croissante sa limite est égale à son sup et

$$\lim_{n \rightarrow +\infty} g_n = \sup_{n \in \mathbb{N}} g_n = \sup_{n \in \mathbb{N}} \inf_{k \geq n} f_k = \underline{\lim}_n f_n.$$

Puis comme $g_n \leq f_n$, on a $\int g_n d\mu \leq \int f_n d\mu$. D'où

$$\int \underline{\lim}_n f_n d\mu = \lim_{n \rightarrow +\infty} \int g_n d\mu = \underline{\lim}_n \int g_n d\mu \leq \underline{\lim}_{n \rightarrow +\infty} \int f_n d\mu$$

où on a utilisé que la limite de $\int g_n d\mu$ coïncide avec sa liminf car la limite existe. \square

Ce résultat explique que par passage à la limite, les intégrales de fonctions positives ne peuvent que diminuer.

Application :

Lemme 9.2.1 (Scheffé) Soit $f_n : (X, \mathcal{A}, \mu) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ une suite de fonctions mesurables sur un espace mesuré (X, \mathcal{A}, μ) . On suppose que f_n converge presque partout vers une fonction f . Alors

$$\lim_{n \rightarrow +\infty} \int_X |f_n| d\mu = \int_X |f| d\mu \quad \text{ssi} \quad \lim_{n \rightarrow +\infty} \int_X |f_n - f| d\mu = 0. \quad (9.5)$$

Démonstration : L'inégalité triangulaire donne $|f_n| = |f + (f_n - f)| \leq |f| + |f - f_n|$, puis $|f| = |f_n + (f - f_n)| \leq |f_n| + |f - f_n|$. On en déduit

$$|f_n| - |f| \leq |f - f_n| \quad \text{et} \quad |f_n| - |f| \leq |f - f_n|$$

et donc la même majoration est vraie pour $||f_n| - |f||$.

En intégrant, on a

$$\left| \int_X |f_n| d\mu - \int |f| d\mu \right| \leq \left| \int_X |f_n| - |f| d\mu \right| \leq \int_X ||f_n| - |f|| d\mu \leq \int_X |f - f_n| d\mu.$$

Le sens direct s'en déduit.

Pour la réciproque, par l'inégalité triangulaire, $|f - f_n| \leq |f| + |-(f_n)| = |f| + |f_n|$. On a donc $g_n = |f| + |f_n| - |f - f_n| \geq 0$. Puis comme $f_n \rightarrow f$ pp, on a $g_n \rightarrow 2|f|$ pp.

D'après le lemme de Fatou

$$\begin{aligned} 2 \int_X |f| d\mu &= \int_X \liminf_n g_n \leq \liminf_n \int_X g_n d\mu = \liminf_n \int_X |f| + |f_n| - |f - f_n| d\mu \\ &= 2 \int_X |f| d\mu + \liminf_n \left(- \int_X |f - f_n| d\mu \right). \end{aligned}$$

On en déduit $\liminf_n \left(- \int_X |f - f_n| d\mu \right) \geq 0$.

Et donc $\limsup_n \int_X |f - f_n| d\mu \leq 0$. Mais comme on a des intégrales positives, leur liminf est positive. On a

$$0 \leq \liminf_n \int_X |f - f_n| d\mu \leq \limsup_n \int_X |f - f_n| d\mu \leq 0$$

et donc $\lim_{n \rightarrow +\infty} \int_X |f - f_n| d\mu = 0$. □

Chapitre 10

Intégrales des fonctions mesurables de signe quelconque

10.1 Définition et propriétés

Définition 10.1.1 Une fonction f est dite μ -intégrable si la fonction positive $|f|$ l'est :

$$\int |f| d\mu < +\infty.$$

Une fonction $f : (X, \mathcal{A}, \mu) \rightarrow \mathbb{C}$ s'écrit

$$\begin{aligned} f &= u + iv \\ &= u^+ - u^- + i(v^+ - v^-) \end{aligned} \tag{10.1}$$

avec $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$. Comme les fonctions positives u^+, u^-, v^+, v^- sont toutes majorées par $|f|$ alors leurs intégrales sont toutes finies et on pose par définition

$$\begin{aligned} \int f d\mu &= \int u^+ d\mu - \int u^- d\mu + i \left(\int v^+ d\mu - \int v^- d\mu \right) \\ &= \int u d\mu + i \int v d\mu \end{aligned}$$

ou si $f : X \rightarrow \mathbb{R}$, $\int f d\mu = \int f^+ d\mu - \int f^- d\mu$.

Puis l'intégrale de f , fonction intégrable, sur $E \in \mathcal{A}$ est définie par

$$\int_E f d\mu = \int f \mathbf{1}_E d\mu$$

ce qui a bien un sens car comme $|f \mathbf{1}_E| \leq |f|$, la fonction $f \mathbf{1}_E$ est intégrable.

Remarque 10.1.1 • Notons que si une fonction mesurable est positive, on peut toujours définir son intégrale mais elle peut valoir $+\infty$. Elle est dite alors μ -intégrable si son intégrale par rapport à μ est finie.

• Si f est de signe quelconque, on ne définit son intégrale que si elle est intégrable, c'est à dire $|f|$ d'intégrale (forcément définie) finie.

• En particulier dans la théorie de l'intégrale de Lebesgue, il n'y a pas de notion de simple intégrabilité, c'est à dire de fonction non intégrable dont l'intégrale converge simplement comme par exemple l'intégrale de Riemann $\int_0^{+\infty} \frac{\sin x}{x} dx$ (à voir).

Proposition 10.1.1 (Quelques propriétés)

- Si f, g sont mesurables réelles et $f \leq g$ alors $\int f d\mu \leq \int g d\mu$.
- Si $f = 0$ alors f est intégrable et $\int f d\mu = 0$.
- Si $\mu(E) = 0$, alors $\int_E f d\mu = 0$ pour toute fonction intégrable.
- **(Chasles)** Si E et F sont disjoints, alors $\int_{E \cup F} f d\mu = \int_E f d\mu + \int_F f d\mu$.

Démonstration : Tout ceci provient facilement des propriétés déjà vues pour les intégrales de fonctions positives et passe au cas des fonctions de signes quelconques en les écrivant sous la forme (10.1) et en appliquant à chaque terme les propriétés correspondantes des intégrales de fonctions positives. \square

Proposition 10.1.2 (Linéarité) Soient $\alpha, \beta \in \mathbb{C}$ et f, g intégrables de X dans \mathbb{C} alors

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu.$$

Démonstration : Fastidieux : écrire $f = Re(f)^+ - Re(f)^- + i(Im(f)^+ - Im(f)^-)$, $\alpha = Re(\alpha)^+ - Re(\alpha)^- + i(Im(\alpha)^+ - Im(\alpha)^-)$ et de même pour g et β . Développer $(\alpha f + \beta g)$ et utiliser la linéarité pour les fonctions positives avec des coefficients positifs et reformer α, β, f, g à la fin. \square

Proposition 10.1.3 (Intégrale et valeurs absolues) Soit $f : (X, \mathcal{A}, \mu) \rightarrow \mathbb{C}$ intégrable alors

$$\left| \int f d\mu \right| \leq \int |f| d\mu.$$

Démonstration : Si $\int f d\mu = 0$, c'est clair. Sinon, soit $\alpha = \frac{\int f d\mu}{|\int f d\mu|}$. Alors $|\alpha| = 1$. Puis

$$\begin{aligned} \left| \int f d\mu \right| &= \alpha \int f d\mu = \int \alpha f d\mu = \int Re(\alpha f) d\mu + i \int Im(\alpha f) d\mu \\ &= \int Re(\alpha f) d\mu = \int Re(\alpha f)^+ d\mu - \int Re(\alpha f)^- d\mu \\ &\leq \int Re(\alpha f)^+ d\mu + \int Re(\alpha f)^- d\mu = \int |Re(\alpha f)| d\mu \\ &\leq \int |\alpha f| d\mu = \int |f| d\mu \end{aligned}$$

car $|\alpha| = 1$. \square

10.2 Transfert

Soient (X, \mathcal{A}) et (Y, \mathcal{B}) deux espaces mesurables et $\varphi : (X, \mathcal{A}) \rightarrow (Y, \mathcal{B})$ une fonction mesurable.

Si on considère une mesure μ sur (X, \mathcal{A}) alors la mesure image $\nu = \mu\varphi^{-1} = \mu_\varphi$ est une mesure sur (Y, \mathcal{B}) . On a un lien entre les intégrales par rapport à μ sur X et celle par rapport à $\nu = \mu\varphi^{-1} = \mu_\varphi$ sur Y :

Théorème 10.2.1 (Transfert) *Soit $h : (Y, \mathcal{B}) \rightarrow K = \mathbb{R}, \bar{\mathbb{R}}, \mathbb{C}$ mesurable alors h est ν -intégrable ssi $h \circ \varphi$ est μ -intégrable et*

$$\int_X h \circ \varphi d\mu = \int_Y h d(\mu_\varphi) = \int_Y h d\nu. \quad (10.2)$$

Il s'agit d'une formule de changement de variable abstraite, très générale puisque la seule condition pour le changement de variable $y = \varphi(x)$ est que φ soit mesurable !

Démonstration : La stratégie est à nouveau de commencer par $h = \mathbf{1}_B$ une indicatrice, puis par linéarité de généraliser aux fonctions étagées et par convergence monotone aux fonctions mesurables positives. Enfin par linéarité aux fonctions mesurables quelconques.

- Si $h = \mathbf{1}_B$, $B \in \mathcal{B}$ alors $A = \varphi^{-1}(B) \in \mathcal{A}$ car φ est mesurable et

$$\int_Y \mathbf{1}_B d\nu = \nu(B) = \mu(\varphi^{-1}(B)) = \mu(A) = \int_X \mathbf{1}_A d\mu$$

et (10.2) est vérifiée car $\mathbf{1}_A = \mathbf{1}_B \circ \varphi$.

- Si $h = \sum_{i=1}^n \alpha_i \mathbf{1}_{B_i}$ est étagée, alors

$$h \circ \varphi = \left(\sum_{i=1}^n \alpha_i \mathbf{1}_{B_i} \right) \circ \varphi = \sum_{i=1}^n \alpha_i (\mathbf{1}_{B_i} \circ \varphi) = \sum_{i=1}^n \alpha_i \mathbf{1}_{\varphi^{-1}(B_i)}$$

et

$$\begin{aligned} \int_X (h \circ \varphi) d\mu &= \sum_{i=1}^n \alpha_i \int_X \mathbf{1}_{\varphi^{-1}(B_i)} d\mu = \sum_{i=1}^n \alpha_i \mu(\varphi^{-1}(B_i)) \\ &= \sum_{i=1}^n \alpha_i \nu(B_i) = \sum_{i=1}^n \alpha_i \int_Y \mathbf{1}_{B_i} d\nu \\ &= \int_Y h d\nu. \end{aligned}$$

et (10.2) est vérifiée à nouveau.

• Si h est mesurable positive alors il existe s_n des fonctions étagées positives qui forment une suite croissante qui converge vers h . Mais alors $s_n \circ \varphi$ sont aussi des fonctions étagées qui forment une suite croissante et qui converge vers $h \circ \varphi$. On a donc $\lim_n s_n = h$ et $\lim_n s_n \circ \varphi = h \circ \varphi$. D'après le point 2, on a

$$\int_X s_n \circ \varphi d\mu = \int_Y s_n d\nu.$$

Par convergence monotone, le membre de gauche converge vers $\int_Y h \circ \varphi d\mu$, celui de droite vers $\int_Y h d\nu$, ce qui justifie (10.2) encore dans ce cas.

• Si h est quelconque alors $|h \circ \varphi| = |h| \circ \varphi$ et le point 3, appliquée à $|h|$ positive, donne $\int_X |h \circ \varphi| d\mu = \int_Y |h| d\nu$ et donc l'équivalence de la ν -intégrabilité de h et de la μ -intégrabilité de $h \circ \varphi$.

On écrit alors $(h \circ \varphi)^+ = h^+ \circ \varphi$ et $(h \circ \varphi)^- = h^- \circ \varphi$ puis

$$\begin{aligned} \int_X (h \circ \varphi) d\mu &= \int_X (h \circ \varphi)^+ d\mu - \int_X (h \circ \varphi)^- d\mu \\ &= \int_Y h^+ d\nu - \int_Y h^- d\nu \\ &= \int_Y h d\nu \end{aligned}$$

en utilisant le point 3 pour les fonctions positives h^+ et h^- , ce qui achève la preuve de (10.2) dans le cas général. \square

10.3 Loi des variables aléatoires

Application : Les lois de probabilités \mathbb{P}_X sont les mesures images de la mesure de probabilité \mathbb{P} par les variables aléatoires X .

Définition 10.3.1 Soit $X : \Omega \rightarrow \mathbb{R}$, on appelle loi de X la mesure \mathbb{P}_X , mesure image sur \mathbb{R} de X par rapport à \mathbb{P} :

$$\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(\omega \in \Omega \mid X(\omega) \in B).$$

Exemples :

- $\Omega = \{a_1, \dots, a_n\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ et \mathbb{P} définie par $\mathbb{P}(a_i) = p_i$ avec $p_1 + \dots + p_n = 1$:

$$\mathbb{P}_X = p_1 \delta_{a_1} + \dots + p_n \delta_{a_n}.$$

- La variable X qui indique la face $1, \dots, 6$ obtenue par le lancer d'un dé est discrète :

$$\mathbb{P}_X = \frac{1}{6} \delta_1 + \dots + \frac{1}{6} \delta_6.$$

- La variable Y qui indique le numéro du premier lancer où on obtient un 6 lors d'une suite infinie de lancer de dé est discrète (les valeurs possibles sont \mathbb{N}) :

$$\mathbb{P}_Y = \sum_{k=1}^{+\infty} p(1-p)^k \delta_k.$$

• La variable Z qui indique le nombre de succès dans une suite de n épreuves est discrète (et prend ses valeurs dans $\{1, \dots, n\}$) :

$$\mathbb{P}_Y = \sum_{k=0}^n C_n^k p^k (1-p)^{n-k} \delta_k.$$

• Si $A \in \mathcal{F}$ est un évènement, alors $X = \mathbf{1}_A$ est une variable aléatoire. Elle vaut 1 si l'évènement A est réalisé 0 sinon :

$$\mathbb{P}_X = \delta_A.$$

Par exemple, précédemment, X est de loi équirépartie sur $\{1, \dots, 6\}$, Y est de loi géométrique $\mathcal{G}(1/6)$ et Z est de loi binomiale $\mathcal{B}(n, p)$.

Définition 10.3.2 On appelle fonction de répartition d'une v.a. $X : \Omega \rightarrow \mathbb{R}$ la fonction F_X définie sur \mathbb{R} par

$$F_X(t) = \mathbb{P}(X \leq t) = \mathbb{P}_X(]-\infty, t]).$$

Proposition 10.3.1 (Propriétés de la fonction de répartition) a) F_X est croissante.

b) $\lim_{t \rightarrow -\infty} F_X(t) = 0$ et $\lim_{t \rightarrow +\infty} F_X(t) = 1$.

c) F_X est continue à droite et $\lim_{t \rightarrow t_0, t \leq t_0} F_X(t) = \mathbb{P}(X < t_0)$.

Démonstration : • a) est dû à la croissance de la mesure.

b) et c) s'obtiennent en appliquant les propriétés de monotonie séquentielle des mesures.

• Pour b), prendre d'abord $A_n =]-\infty, -n]$, il est clair que $\cap_n A_n = \emptyset$, de mesure nulle, si bien que

$$\lim_{t \rightarrow -\infty} F_X(t) = \lim_{n \rightarrow +\infty} F_X(-n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq -n) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(A_n) = \mathbb{P}_X(\cap_n A_n) = 0.$$

Puis prendre $B_n =]-\infty, n]$ de réunion $\cup_n B_n =]-\infty, +\infty[= \mathbb{R}$ de mesure $\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X \in \mathbb{R}) = 1$ si bien que

$$\lim_{t \rightarrow +\infty} F_X(t) = \lim_{n \rightarrow +\infty} F_X(n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq n) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(B_n) = \mathbb{P}_X(\cup_n B_n) = 1.$$

• Pour le c), prendre $A_n =]-\infty, x + 1/n]$ d'intersection $\cap_n A_n =]-\infty, x]$ de mesure $\mathbb{P}_X(\cap_n A_n) = F_X(x)$ si bien que

$$\lim_{t \rightarrow x^+} F_X(t) = \lim_{n \rightarrow +\infty} F_X(x + 1/n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq x + 1/n) = \lim_{n \rightarrow +\infty} \mathbb{P}_X(A_n) = \mathbb{P}_X(\cap_n A_n) = F_X(x).$$

Ensuite, prendre $B_n =]-\infty, x - 1/n]$ de réunion $\cap_n B_n =]-\infty, x[$ de mesure $\mathbb{P}_X(\cup_n B_n) = \mathbb{P}(X < x)$ qui peut être distinct de $\mathbb{P}(X \leq x)$ car

$$\mathbb{P}(X \leq x) - \mathbb{P}(X < x) = \mathbb{P}(\{X \leq x\} \setminus \{X < x\}) = \mathbb{P}(X = x)$$

qui peut être non nul. On a alors

$$\lim_{t \rightarrow x^-} F_X(t) = \lim_{n \rightarrow +\infty} F_X(x - 1/n) = \lim_{n \rightarrow +\infty} \mathbb{P}(X \leq x - 1/n) = \lim_{n \rightarrow +\infty} \mathbb{P}(B_n) = \mathbb{P}(\cup_n B_n) = \mathbb{P}(X < x).$$

□

Cas des variables aléatoires à densité

Définition 10.3.3 X est une variable aléatoire de densité f si sa loi est donnée par

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \int_A f(x) d\lambda(x), \quad \mathbb{P}(X \in [a, b]) = \int_a^b f(x) dx.$$

Remarque 10.3.1 La fonction de densité f est une fonction mesurable intégrable au sens de Lebesgue. Cependant la plupart du temps, il s'agira d'une fonction Riemann-intégrable et même d'une fonction usuelle.

De toute façon, la densité f doit vérifier $f(x) \geq 0$ et $\int_{\mathbb{R}} f(x) d\lambda(x) = 1$.

Par exemple,

$$\frac{e^{-x^2/2}}{\sqrt{2\pi}}, \quad \frac{1}{b-a} \mathbf{1}_{[a,b]}, \quad \alpha e^{-\alpha x} \mathbf{1}_{\mathbb{R}^+}(x), \quad \frac{1}{\pi(1+x^2)}$$

sont des densités des lois normale standard $\mathcal{N}(0, 1)$, uniforme $\mathcal{U}([a, b])$, exponentielle $\mathcal{E}(\alpha)$ et de Cauchy $\mathcal{C}(1)$.

Dans le cas à densité, la densité f est reliée à la fonction de répartition F_X de la façon suivante.

Proposition 10.3.2 Si X est une v.a. de densité f , sa fonction de répartition F_X vérifie :

- (i) $\forall x \in \mathbb{R}, F_X(x) = \int_{-\infty}^x f(t) dt.$
- (ii) F est continue sur \mathbb{R} .
- (iii) Si f est continue au point x_0 , alors F est dérivable en x_0 de dérivée $F'(x_0) = f(x_0)$.

D'après (ii), la fonction de répartition est continue. De là, vient le nom qu'on donne parfois aux variables aléatoires à densité : variable aléatoire continue.

Démonstration : Puisque X a pour densité f , et comme

$$F(b) = \mathbb{P}(X \in]-\infty, b]) = \mathbb{P}(X \in]-\infty, a] \cup]a, b]) = F(a) + \mathbb{P}(X \in]a, b]),$$

on a pour tous réels $a < b$:

$$\mathbb{P}(\omega, X(\omega) \in]a, b]) = \mathbb{P}(X \in]a, b]) = F(b) - F(a) = \int_a^b f(t) dt. \quad (10.3)$$

(i) : Il suffit d'appliquer (10.3) avec $b = x$ fixé et $a = -n$ pour chaque $n \in \mathbb{N}$ tel que $x > -n$. La suite d'évènements

$$A_n = \{\omega, X(\omega) \in]-n, x]\}, \quad n > -x,$$

est croissante pour l'inclusion et de réunion $A = \{\omega, X(\omega) \in]-\infty, x]\} = \{X \leq x\}$. Par la propriété de continuité monotone séquentielle, on a $\mathbb{P}(A_n) \uparrow \mathbb{P}(A)$, d'où

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(A) = \lim_{n \rightarrow +\infty} \mathbb{P}(A_n) = \lim_{n \rightarrow +\infty} \int_{-n}^x f(t) dt = \int_{-\infty}^x f(t) dt$$

en notant que l'intégrale généralisée de la densité f converge en $-\infty$.

(ii) : On fixe $x_0 \in \mathbb{R}$ quelconque. D'abord F est continue à droite en tout point car c'est une fonction de répartition

Il reste à voir la continuité à gauche. On se contente de le faire avec l'hypothèse supplémentaire suivante : « il existe $a < x_0$ tel que f soit définie et Riemann-intégrable sur tout intervalle $[a, a'] \subset [a, x_0]$ ». On a alors :

$$\lim_{x \uparrow x_0} \int_a^x f(t) dt = \int_a^{x_0} f(t) dt,$$

ou la deuxième intégrale est soit une intégrale de Riemann ordinaire soit une intégrale de Riemann généralisée convergente. On peut réécrire

$$\lim_{x \uparrow x_0} (F(x) - F(a)) = F(x_0) - F(a).$$

On conclut en rajoutant des deux côtés $F(a)$.

(iii) : Comme par hypothèse f est continue en x_0 , elle est définie sur tout un voisinage de x_0 et donc sur un intervalle $[a, b]$ qui contient x_0 . La continuité de f en x_0 s'écrit : $\forall \varepsilon > 0, \exists \delta > 0$ tel que $]x_0 - \delta, x_0 + \delta[\subset]a, b[$ et

$$\forall t \in]x_0 - \delta, x_0 + \delta[, \quad |f(t) - f(x_0)| < \varepsilon.$$

Pour tout h tel que $0 < |h| < \delta$, on a alors $F(x_0 + h) - F(x_0) = \int_{x_0}^{x_0+h} f(t) dt$. D'où

$$|F(x_0 + h) - F(x_0) - hf(x_0)| = \left| \int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt \right| \leq \int_{x_0}^{x_0+h} |(f(t) - f(x_0))| dt \leq h\varepsilon.$$

En divisant par h puis en faisant $h \rightarrow 0$, on constate que

$$\lim_{h \rightarrow 0} \frac{F(x_0 + h) - F(x_0)}{x_0 + h - x_0} = f(x_0)$$

c'est à dire : F est dérivable en x_0 , de dérivée $f'(x_0)$. □

10.4 Ensembles négligeables

Définition 10.4.1 *Un ensemble $A \in \mathcal{A}$ est dit μ -négligeable si $\mu(A) = 0$.*

Exemples : Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $\{x_0\}, \mathbb{N}, \mathbb{Z}, \mathbb{Q}$ sont négligeables. Sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \eta)$, où η est la mesure de comptage, seul \emptyset est négligeable.

Définition 10.4.2 *Une fonction f mesurable est dite μ -négligeable s'il existe A négligeable tel que $\forall x \notin A$ alors $f(x) = 0$ (l'ensemble des points où f est non nulle est dans un négligeable).*

Exemples : Sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $f(x) = \mathbf{1}_{\mathbb{Q}}(x)$, $g(x) = x\mathbf{1}_{\mathbb{N}}(x)$ sont négligeables.

Définition 10.4.3 Une propriété est dite vraie (μ) -presque partout, si l'ensemble des points où elle n'est pas vraie est μ -négligeable.

En particulier, $f = g$ presque partout si $\{x \in X \mid f(x) \neq g(x)\}$ est dans un négligeable.

On écrit p.p. pour presque partout.

Proposition 10.4.1 • Si $f = g$ p.p. alors f est intégrable ssi g l'est et

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

- Si $f \leq g$ p.p. alors $\int f d\mu \leq \int g d\mu$.
- Si $f \geq 0$ et $\int f d\mu < +\infty$ alors f est finie p.p.
- Si $|\int f d\mu| = \int |f| d\mu$ alors il existe $\alpha \in \mathbb{C}$ de module 1 tel que $\alpha f = |f|$.

Démonstration :

• Soit $A = \{x \in X \mid f(x) \neq g(x)\}$ alors $\mu(A) = 0$. Si $x \notin A$, on a $f(x) = g(x)$ donc $|f(x)| = |g(x)|$. Donc si f est intégrable :

$$\begin{aligned} \int |g| d\mu &= \int_A |g| d\mu + \int_{A^c} |g| d\mu = \int_{A^c} |g| d\mu \\ &= \int_{A^c} |f| d\mu = \int_{A^c} |f| d\mu + \int_A |f| d\mu = \int |f| d\mu < +\infty \end{aligned}$$

alors g l'est. Puis le même calcul sans les $|\cdot|$ montre que $\int f d\mu = \int g d\mu$.

- Si $A = \{x \in X \mid f(x) = +\infty\}$ n'est pas négligeable alors

$$\int f d\mu \geq \int_A f d\mu = +\infty$$

car $f = +\infty$ sur A qui est de mesure $\mu(A) > 0$. On contredit la finitude de $\int f d\mu$.

• Reprendre la preuve de $|\int f d\mu| \leq \int |f| d\mu$. Pour avoir égalité, il faut qu'il y ait égalité dans toutes les étapes de cette preuve. Ce qui nécessite $Re(\alpha f)^- = 0$ et $|Re(\alpha f)| = |\alpha f|$. On a donc $Re(\alpha f) = |\alpha f|$, ce qui implique $Im(\alpha f) = 0$ et donc $|f| = |\alpha f| = Re(\alpha f) = \alpha f$, ce qui prouve le résultat car dans cette preuve $|\alpha| = 1$. \square

Remarque 10.4.1 (Tribu complète) Il se peut que A mesurable de mesure $\mu(A)$ nulle contienne des sous ensembles B non mesurables, ce qui cause bien des soucis. On dira que la tribu est complète si pour tout A mesurable de mesure nulle, dès que $N \subset A$ alors N est mesurable aussi (et alors forcément de mesure nulle).

On peut toujours compléter la tribu \mathcal{A} en considérant la tribu $\tilde{\mathcal{A}}$ définie par

$$\tilde{\mathcal{A}} = \{A \in \mathcal{A}, A \cup N \text{ avec } N \subset B \in \mathcal{A} \text{ et } \mu(B) = 0\}.$$

On prolonge alors μ sur $\tilde{\mathcal{A}}$ par $\mu(A \cup N) = \mu(A)$.

Dans la suite, on supposera toujours que les tribus sont complètes.

10.5 Convergence dominée et applications

Théorème 10.5.1 (Convergence dominée) Soit $f_n : (X, \mathcal{A}, \mu) \rightarrow \mathbb{C}$ mesurable telle que $f_n \rightarrow f$ p.p. quand $n \rightarrow +\infty$. S'il existe g une fonction mesurable $(X, \mathcal{A}, \mu) \rightarrow [0, +\infty]$ telle que

- 1) $|f_n| \leq g$ p.p.
- 2) g est μ -intégrable.

Alors

$$\lim_{n \rightarrow +\infty} \int f_n d\mu = \int f d\mu. \quad (10.4)$$

Conséquence : Si la convergence est dominée, on peut intervertir limite et intégrale.

En fait, on a même mieux que (10.4) : on a $\lim_{n \rightarrow +\infty} \int |f_n - f| d\mu = 0$.

Démonstration : On considère la fonction $2g - |f - f_n|$. Comme $|f_n| \leq g$ et $|f| \leq g$, il vient facilement $|f - f_n| \leq 2g$, c'est à dire $2g - |f - f_n| \geq 0$. De plus $\lim_n 2g - |f - f_n| = 2g$ car $f_n \rightarrow f$ simplement. On applique alors le lemme de Fatou à cette suite de fonctions :

$$\begin{aligned} \int \underline{\lim}_{n \rightarrow +\infty} (2g - |f - f_n|) d\mu &\leq \underline{\lim}_{n \rightarrow +\infty} \int 2g - |f - f_n| d\mu \\ \int 2g d\mu &\leq \underline{\lim}_{n \rightarrow +\infty} \int 2g d\mu - \int |f - f_n| d\mu \\ \int 2g d\mu &\leq \int 2g d\mu + \underline{\lim}_{n \rightarrow +\infty} \left(- \int |f - f_n| d\mu \right) \\ 0 &\leq -\overline{\lim}_{n \rightarrow +\infty} \int |f - f_n| d\mu \\ \overline{\lim}_{n \rightarrow +\infty} \int |f - f_n| d\mu &\leq 0. \end{aligned}$$

Comme en plus $\underline{\lim}_{n \rightarrow +\infty} \int |f - f_n| d\mu \geq 0$, on a $\lim_{n \rightarrow +\infty} \int |f - f_n| d\mu = 0$ et *a fortiori*

$$\lim_{n \rightarrow +\infty} \int f_n d\mu = \int f d\mu. \quad \square$$

Exemple : Montrer que si $f_n : [0, 1] \rightarrow [0, 1]$ est continue et pour tout $x \in [0, 1]$; $f_n(x) \rightarrow 0$ alors $\int_{[0,1]} f_n d\lambda \rightarrow 0$. Essayer sans convergence dominée.

Exemple : Considérer la suite de fonctions $(f_n)_{n \in \mathbb{N}}$ mesurables définies sur \mathbb{R} et telle que pour tout $n \in \mathbb{N}$, et $x \in \mathbb{R}$: $|f_n(x)| \leq 1/(1+x^2)$. Supposons que pour tout $x \in \mathbb{R}$, $f_n(x)$ converge vers $f(x)$. Montrer alors que $\lim_{n \rightarrow +\infty} \int_{-\infty}^{+\infty} f_n(x) dx = \int_{-\infty}^{+\infty} f(x) dx$.

Application aux intégrales à paramètres

Les résultats suivants donnent des conditions assez faibles sur f pour que la fonction $F(t) = \int_X f(t, x) d\mu(x)$ soit continue ou dérivable.

Théorème 10.5.2 (Continuité sous l'intégrale) *Soit (X, \mathcal{A}, μ) un espace mesuré et T un espace métrique (penser à \mathbb{R} ou à \mathbb{C}).*

On considère $f : T \times X \rightarrow \mathbb{C}$ tel que pour tout $t, x \mapsto f(t, x)$ est mesurable et

$$\forall(t, x), \quad |f(t, x)| \leq g(x)$$

avec g intégrable ($\int_X g(x) d\mu(x) < +\infty$). On pose $F(t) = \int_X f(t, x) d\mu(x)$.

Soit $t_0 \in T$, on suppose que pour presque tout $x \in X$, $t \mapsto f(t, x)$ est continue en t_0 . Alors F est continue en t_0 .

Démonstration : Comme \mathbb{R} et \mathbb{C} sont métriques, la continuité est donnée par la continuité séquentielle (c'est à dire avec les suites : F est continue en t ssi pour toute suite $(t_n)_n$ qui converge vers t , on a $\lim_{n \rightarrow +\infty} F(t_n) = F(t)$).

Il faut alors voir que pour toute suite $(t_n)_n$ qui converge vers t_0 , on a $F(t_n) \rightarrow F(t_0)$. Mais $F(t_n) = \int_X f(t_n, x) d\mu(x)$. Les fonctions $x \mapsto f(t_n, x)$ sont mesurables et dominées par g , intégrable. De plus quand $n \rightarrow +\infty$, $f(t_n, x) \rightarrow f(t_0, x)$ pour presque chaque $x \in X$ par la continuité presque partout de $f(t_0, \cdot)$. La conclusion découle du théorème de convergence dominée :

$$\lim_{n \rightarrow +\infty} F(t_n) = \lim_{n \rightarrow +\infty} \int_X f(t_n, x) d\mu(x) = \int_X f(t_0, x) d\mu(x) = F(t_0).$$

□

Exemples. La fonction $F(t) = \int_0^{+\infty} e^{-tx} dx$ est continue sur \mathbb{R}_+^* .

La fonction Gamma $\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx$ est continue sur \mathbb{R}_+^* .

Théorème 10.5.3 (Dérivabilité sous l'intégrale) *Soit (X, \mathcal{A}, μ) , I un intervalle ouvert de \mathbb{R} et $f : I \times X \rightarrow \mathbb{C}$ telle que*

- $\forall t \in I, x \mapsto f(t, x)$ est mesurable
- $\exists t_0 \in I, x \mapsto f(t_0, x)$ est intégrable.

On suppose qu'il existe $A \in \mathcal{A}$ tel que $\mu(A^c) = 0$ et

- $\forall x \in A, t \mapsto f(t, x)$ est dérivable en tout $t \in I$
- $\forall x \in A, \forall t \in I, \left| \frac{\partial f}{\partial t}(t, x) \right| \leq g(x)$, avec $\int_X g(x) d\mu < +\infty$.

Alors $F(t) = \int_X f(t, x) d\mu(x)$ est finie et F est dérivable de dérivée

$$F'(t) = \int_X \frac{\partial f}{\partial t}(t, x) d\mu(x).$$

Démonstration : Pour $x \in A$, $\frac{\partial f}{\partial t}(t, x) = \lim_{t_n \rightarrow t} \frac{f(t_n, x) - f(t, x)}{t_n - t}$, est mesurable car limite de fonctions mesurables.

Pour $t_1 \in I$ et t_0 donné dans l'énoncé, par le théorème des accroissements finis, on a $\theta \in (t_0, t_1)$:

$$|f(t_1, x) - f(t_0, x)| = \left| \frac{\partial f}{\partial t}(\theta, x) \right| |t_1 - t_0| \leq g(x) |t_1 - t_0|.$$

Il suit $|f(t_1, x)| \leq g(x) |t_1 - t_0| + |f(t_0, x)|$ qui est intégrable car g et $f(t_0, \cdot)$ le sont. L'intégrale $F(t_1)$ est donc finie pour tout $t_1 \in I$.

Soit maintenant $t \in I$ fixé et $t_n \rightarrow t$, on a

$$\frac{F(t_n) - F(t)}{t_n - t} = \int_A \frac{f(t_n, x) - f(t, x)}{t_n - t} d\mu(x).$$

Or $\frac{f(t_n, x) - f(t, x)}{t_n - t}$ est une suite de fonctions qui tend vers $\frac{\partial f}{\partial t}(t, x)$ et qui vaut par le théorème des accroissements finis $\frac{\partial f}{\partial t}(\theta_n, x)$ pour un certain $\theta_n \in (t, t_n)$. Or $\frac{\partial f}{\partial t}(\theta_n, x)$ est dominée par $g(x)$, intégrable. Le théorème de convergence dominée s'applique et donne

$$F'(t) = \lim_{n \rightarrow +\infty} \frac{F(t_n) - F(t)}{t_n - t} = \int_A \frac{\partial f}{\partial t}(t, x) d\mu(x) = \int_X \frac{\partial f}{\partial t}(t, x) d\mu(x).$$

□

Exemples. La fonction $F(t) = \int_0^{+\infty} e^{-tx} dx$ est dérivable sur \mathbb{R}_+^* .
La fonction Gamma $\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx$ est dérivable sur \mathbb{R}_+^* .

10.6 Espérance probabiliste

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}^+$ une variable aléatoire positive.

L'intégrale de X par rapport à la mesure \mathbb{P} est son espérance :

Définition 10.6.1 $E[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega) = \int X d\mathbb{P}.$

Une variable $X \geq 0$ est dite intégrable si son espérance est finie.

Un exemple de variable aléatoire positive est $X = \mathbf{1}_A$ où $A \in \mathcal{F}$ est un évènement on a alors

$$E[X] = E[\mathbf{1}_A] = \int_{\Omega} \mathbf{1}_A d\mathbb{P} = \mathbb{P}(A).$$

La variable $\mathbf{1}_A$ qui indique si l'évènement A se réalise ou non a pour espérance $\mathbb{P}(A)$.

Définition 10.6.2 Soit X une variable de signe quelconque. Elle est dite intégrable si $|X|$ est d'espérance (forcément définie) finie. On note alors

$$E[X] = \int X d\mathbb{P}.$$

La quantité $E[|X|]$ s'appelle aussi le moment d'ordre 1. On peut formuler de la façon suivante : X est intégrable si son moment d'ordre 1 est fini.

Conséquences : des propriétés de l'intégration, on déduit pour des variables aléatoires intégrables X, Y et des réels a, b :

- $E[aX + bY] = aE[X] + bE[Y]$ (linéarité de E).
- Inégalité de Markov : Si X est une v.a. positive $\mathbb{P}(X \geq t) \leq \frac{E[X]}{t}$.
- L'espérance $E[X]$ n'est rien d'autre que l'intégrale (au sens de Lebesgue) de la fonction mesurable X par rapport à la mesure de probabilité \mathbb{P} .

D'après le théorème de transfert, $E[X]$ peut s'écrire comme l'intégrale par rapport à la loi :

Corollaire 10.6.1 Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ une variable aléatoire de loi \mathbb{P}_X . Alors X est \mathbb{P} -intégrable ssi

$$\int_{\mathbb{R}} |x| d\mathbb{P}_X < +\infty.$$

Et l'espérance est alors

$$E[X] = \int X d\mathbb{P} = \int_{\mathbb{R}} x d\mathbb{P}_X(x) < +\infty.$$

Démonstration : Appliquer le théorème de transfert avec la fonction mesurable $\varphi = X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$, $\mathbb{P}_X = \mathbb{P}X^{-1}$, $h(x) = x$:

$$E[X] = \int_{\Omega} h \circ X d\mathbb{P} = \int_{\mathbb{R}} h d\mathbb{P}_X = \int_{\mathbb{R}} x d\mathbb{P}_X(x).$$

□

Plus généralement, si $h : \mathbb{R} \rightarrow \mathbb{R}$ est borélienne (i.e. mesurable) et $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ est une variable aléatoire. Alors $Y = h(X)$ est une variable aléatoire car c'est une application $(\Omega, \mathcal{F}, \mathbb{P}) \xrightarrow{X} \mathbb{R} \xrightarrow{h} \mathbb{R}$, c'est à dire $h \circ X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$. Puis la variable Y est \mathbb{P} -intégrable ssi h est \mathbb{P}_X -intégrable ($\int_{\mathbb{R}} |h(x)| d\mathbb{P}_X < +\infty$) et dans ce cas

$$E[Y] = E[h(X)] = \int_{\Omega} h(X) d\mathbb{P} = \int_{\mathbb{R}} h(x) d\mathbb{P}_X(x) = \int_{\mathbb{R}} y d\mathbb{P}_{h(X)}(y).$$

Pour cela, appliquer le théorème de transfert avec $(X, \mathcal{A}, \mu) = (\Omega, \mathcal{F}, \mathbb{P})$, $(Y, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et $\varphi = X$.

Variable discrète

Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ avec $X(\Omega)$ discret. La loi de X est donnée par la mesure discrète

$$\mathbb{P}_X = \sum_{x_k \in X(\Omega)} \mathbb{P}(X = x_k) \delta_{x_k}$$

Il s'agit d'une somme de mesures de Dirac : en chaque atome $x_k \in X(\Omega)$, il y a la masse $\mathbb{P}(X = x_k)$.

Alors X est intégrable ssi

$$E[|X|] = \sum_{x_k \in X(\Omega)} |x_k| \mathbb{P}(X = x_k) < +\infty$$

et dans ce cas $E[X] = \sum_{x_k \in X(\Omega)} x_k \mathbb{P}(X = x_k)$ où la somme est au plus dénombrable car $X(\Omega)$ est discret (la v.a. X est discrète).

Si $h : \mathbb{R} \rightarrow \mathbb{R}$ est mesurable, alors $h(X)$ est une variable discrète, elle est intégrable ssi $\sum_{x_k \in X(\Omega)} |h(x_k)| \mathbb{P}(X = x_k) < +\infty$. Son espérance est alors

$$E[h(X)] = \sum_{x_k \in X(\Omega)} h(x_k) \mathbb{P}(X = x_k).$$

Exemples : (lois de variables discrètes classiques)

- Loi de Bernoulli de paramètre p notée $b(p)$

Une v.a. X suit une loi de Bernoulli de paramètre $p \in [0, 1]$ si elle ne prend que deux valeurs, la plupart du temps 0 et 1 avec :

$$\mathbb{P}(X = 1) = p, \quad \mathbb{P}(X = 0) = 1 - p := q.$$

Exemple : pile ou face avec $p = 1/2$ si la pièce est équilibrée, $p \neq 1/2$ si elle est truquée. Son espérance est $E[X] = p$.

- Loi equirépartie sur un ensemble fini $\{x_1, \dots, x_n\}$ notée $\mathcal{U}\{x_1, \dots, x_n\}$.

Une v.a. X prenant un nombre fini de valeurs x_1, \dots, x_n suit une loi equirépartie quand

$$\mathbb{P}(X = x_i) = \mathbb{P}_X(\{x_i\}) = \frac{1}{n}, \quad 1 \leq i \leq n.$$

Exemple : jet d'un dé

Son espérance est $E[X] = \frac{x_1 + \dots + x_n}{n}$.

- Loi binomiale de paramètres n, p notée $\mathcal{B}(n, p)$

Une v.a. suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0, 1]$ si l'ensemble de ses valeurs possibles est :

$$X(\Omega) = \{0, 1, 2, \dots, n\}$$

et pour tout $k = 0, 1, \dots, n$, on a

$$\mathbb{P}(X = k) = C_n^k p^k (1 - p)^{n-k} \quad (10.5)$$

où $C_n^k = \frac{n!}{k!(n-k)!}$ est le coefficient binomial. Il s'agit bien d'une loi de probabilité car la formule du binôme de Newton (d'où le nom de la loi) donne :

$$\sum_{k=0}^n C_n^k p^k (1-p)^{n-k} = (p + (1-p))^n = 1^n = 1.$$

Il est souvent pratique de voir cette loi comme celle du nombre de succès obtenus dans une suite de n épreuves répétées indépendantes avec pour chaque épreuve une probabilité p de succès.

Son espérance est $E[X] = np$.

- Loi géométrique de paramètre $p \in]0, 1[$

Une v.a. X suit la loi géométrique de paramètre $p \in]0, 1[$ notée $\mathcal{G}(p)$ si $X(\Omega) = \mathbb{N}^*$ et

$$\mathbb{P}(X = k) = (1-p)^{k-1} p, \quad k \in \mathbb{N}^*.$$

Exemple : dans une suite infinie d'épreuves indépendantes avec probabilité p de succès à chacune, elle modélise le rang du premier succès.

Son espérance est $E[X] = 1/p$.

- Loi de Poisson de paramètre λ

Une v.a. discrète X suit une loi de Poisson de paramètre $\lambda > 0$ notée $\mathcal{P}(\lambda)$ si l'ensemble de ses valeurs possibles est $X(\Omega) = \mathbb{N}$ et

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}.$$

Son espérance est $E[X] = \lambda$.

Variable à densité

Soit $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ une variable de densité f . La loi de X est donnée par la mesure

$$\mathbb{P}_X(A) = \int_A f d\lambda = \int_A f(x) dx, \quad \forall A \in \mathcal{B}(\mathbb{R})$$

où f est une fonction mesurable positive d'intégrale égale à 1. Alors X est intégrable ssi

$$E[|X|] = \int_{\mathbb{R}} |x| f(x) dx < +\infty$$

et dans ce cas $E[X] = \int_{\mathbb{R}} x f(x) dx$.

Si $h : \mathbb{R} \rightarrow \mathbb{R}$ est mesurable, alors $h(X)$ est une variable aléatoire, elle est intégrable ssi $\int_{\mathbb{R}} |h(x)|f(x)dx < +\infty$. Son espérance est alors

$$E[h(X)] = \int_{\mathbb{R}} h(x)f(x)dx.$$

Noter que l'intégrale de Lebesgue unifie les deux cas : variables aléatoires discrètes et à densité. La différence n'était donc que formelle.

Exemples : (Lois à densité classiques)

- Loi uniforme

La v.a. X suit une loi uniforme sur l'intervalle $[a, b]$ ($-\infty < a < b < +\infty$) si elle a une densité f constante sur cet intervalle et nulle en dehors. Sa densité est alors

$$f(t) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(t) = \begin{cases} 1/(b-a) & \text{si } t \in [a, b], \\ 0 & \text{si } t \notin [a, b]. \end{cases}$$

Son espérance est $E[X] = \frac{b+a}{2}$.

En fait on peut définir une loi uniforme sur un ensemble borélien A quelconque (pas forcément un intervalle), c'est la loi de densité $\frac{1}{\lambda(A)} \mathbf{1}_A$.

- Loi exponentielle

La v.a. X suit une loi exponentielle de paramètre $a > 0$, notée $\mathcal{E}(a)$ si elle admet pour densité :

$$f(t) = ae^{-at} \mathbf{1}_{[0,+\infty[}(t).$$

Elle est utilisée pour modéliser un temps d'attente d'un phénomène aléatoire. Le temps d'attente moyen est alors $E[X] = 1/a$.

- Loi de Cauchy

Une variable aléatoire réelle suit une loi de Cauchy de paramètre $a \in \mathbb{R}_+^*$ si elle admet pour densité :

$$f(t) = \frac{a}{\pi} \frac{1}{a^2 + t^2}.$$

Son espérance n'est pas définie car $E[|X|] = \int_{\mathbb{R}} \frac{a|x|}{\pi(a^2 + x^2)} dx = +\infty$, la fonction $x/(1+x^2) \simeq 1/x$ n'est pas intégrable en $\pm\infty$. Alors que par parité, on aurait facilement (mais erronément) $E[X] = 0$!

- Loi normale (standard)

Une variable aléatoire X_0 suit la loi normale standard $\mathcal{N}(0, 1)$ si elle admet pour densité

$$t \mapsto \frac{1}{\sqrt{2\pi}} e^{-t^2/2}.$$

Elle admet un moment d'ordre 1 et son espérance est $E[X_0] = 0$

Une variable aléatoire X de loi normale $\mathcal{N}(m, \sigma^2)$ peut alors se définir comme une translatée et dilatée de X_0 par

$$X = m + \sigma X_0.$$

Son espérance est $E[X] = m + \sigma E[X_0] = m$.

Chapitre 11

Lien avec l'intégrale de Riemann

L'intégrale de Riemann est une intégrale pour les fonctions définies sur \mathbb{R} . La mesure classique à considérer sur \mathbb{R} est la mesure de Lebesgue λ . On va donc faire le lien entre l'intégrale de Riemann et l'intégrale de Lebesgue pour la mesure de Lebesgue λ .

11.1 Fonctions en escalier

On a déjà remarqué que les fonctions en escalier

$$f(x) = \sum_{i=0}^{n-1} \alpha_i \mathbf{1}_{[t_i, t_{i+1}[}(x).$$

sont des cas spéciaux de fonctions étagées. Leur intégrale de Riemann sur $[a, b]$ est donnée par

$$\int_a^b f(x) dx = \sum_{i=0}^{n-1} \alpha_i (t_{i+1} - t_i).$$

Tandis (qu'en tant que fonctions étagées), leur intégrale de Lebesgue sur $[a, b]$ ($[a, b]$ est un intervalle donc un ensemble borélien de \mathbb{R}) est donnée par

$$\int_{[a,b]} f d\lambda = \sum_{i=0}^{n-1} \alpha_i \lambda([t_i, t_{i+1}[) = \sum_{i=0}^{n-1} \alpha_i (t_{i+1} - t_i).$$

Les deux types d'intégrale de Riemann et de Lebesgue coïncident donc pour les fonctions en escalier :

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

11.2 Fonctions continues de $[a, b] \rightarrow \mathbb{R}$

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction continue. Elle est mesurable et bornée par $M : |f(x)| \leq M$. Donc f est Lebesgue-intégrable sur $[a, b]$ fini :

$$\int_{[a,b]} |f(x)| d\lambda \leq \int_{[a,b]} M d\lambda = M\lambda([a, b]) = (b - a)M < +\infty.$$

De plus, la continuité sur $[a, b]$ donne en fait l'uniforme continuité de f (théorème de Heine). La fonction f est donc limite uniforme de fonctions en escalier E_n (en fait, on a dit dans ce cas que f est une fonction réglée).

On a vu que pour les fonctions en escalier E_n , les intégrales de Riemann $\int_a^b E_n(x) dx$ et de Lebesgue $\int_{[a,b]} E_n d\lambda$ coïncident. Par définition de $\int_a^b f(x) dx$ limite de $\int_a^b E_n(x) dx$ et de $\int_{[a,b]} f d\lambda$ sup, donc limite, des intégrales $\int_{[a,b]} E_n d\lambda$, on a

$$\int_a^b f(x) dx = \lim_{n \rightarrow +\infty} \int_a^b E_n(x) dx = \lim_{n \rightarrow +\infty} \int_{[a,b]} E_n(x) d\lambda = \int_{[a,b]} f d\lambda$$

Les intégrales de Riemann et de Lebesgue coïncident donc pour des fonctions continues sur des intervalles bornés. Elles sont même Lebesgue-intégrables.

Remarque 11.2.1 • La même chose reste vraie (plus généralement) pour les fonctions réglées, limites uniformes de fonctions en escalier.

• On a deux notations différentes pour les intégrales de Riemann et de Lebesgue d'une fonction (continue) f mais finalement les deux intégrales coïncident :

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

• Tous les calculs que l'on sait faire avec l'intégrale de Riemann restent donc valables avec l'intégrale de Lebesgue. Les calculs de primitives, intégration par parties, changements de variable etc ne disparaissent pas.

11.3 Fonctions Riemann-intégrables

Mesurabilité

D'abord, on montre que les fonctions Riemann-intégrables ont un sens pour l'intégrale de Lebesgue : elles sont mesurables (pour la tribu borélienne de \mathbb{R}).

On commence par le critère suivant de Riemann-intégrabilité.

Théorème 11.3.1 (Lebesgue) *Une fonction bornée f sur $[a, b]$ est Riemann-intégrable ssi l'ensemble de ses points de discontinuité est de mesure nulle. On dira que f est continue presque partout.*

Démonstration admise.

Proposition 11.3.1 *Les fonctions Riemann-intégrables $f : [a, b] \rightarrow \mathbb{R}$ sont mesurables pour les tribus boréliennes de $[a, b]$ et de \mathbb{R} .*

Démonstration : On rappelle qu'étant donnée une subdivision $S = \{t_0, \dots, t_n\}$ de $[a, b]$, on encadre f Riemann-intégrable par les fonctions en escalier de type Darboux :

$$E_{(f,S)}^-(x) = \sum_{i=0}^{n-1} m_i \mathbf{1}_{[t_i, t_{i+1}[}(x), \quad E_{(f,S)}^+(x) = \sum_{i=0}^{n-1} M_i \mathbf{1}_{[t_i, t_{i+1}[}(x)$$

où $m_i = \inf_{x \in [t_i, t_{i+1}[} f(x)$ et avec $M_i = \sup_{x \in [t_i, t_{i+1}[} f(x)$. On a alors

$$E_{(f,S)}^- \leq f \leq E_{(f,S)}^+$$

et aussi

$$\sup_S E_{(f,S)}^- \leq f \leq \inf_S E_{(f,S)}^+$$

où le sup est pris sur les subdivisions S de $[a, b]$. Puis comme f est Riemann-intégrable, $\lim_{\rho(S) \rightarrow 0} A^+(f, S) = \lim_{\rho(S) \rightarrow 0} A^-(f, S)$.

$$\lim_{\rho(S) \rightarrow 0} \int_a^b E_{(f,S)}^-(x) dx = \lim_{\rho(S) \rightarrow 0} A^-(f, S) = \lim_{\rho(S) \rightarrow 0} A^+(f, S) = \lim_{\rho(S) \rightarrow 0} \int_a^b E_{(f,S)}^+(x) dx.$$

Comme pour une subdivision S fixée, on a

$$E_{(f,S)}^- \leq \sup_S E_{(f,S)}^- \leq \inf_S E_{(f,S)}^+ \leq E_{(f,S)}^+$$

on a en intégrant

$$\int_a^b E_{(f,S)}^-(x) dx \leq \int_a^b \sup_S E_{(f,S)}^-(x) dx \leq \int_a^b \inf_S E_{(f,S)}^+(x) dx \leq \int_a^b E_{(f,S)}^+(x) dx$$

et en passant à la limite

$$\lim_{\rho(S) \rightarrow 0} \int_a^b E_{(f,S)}^-(x) dx \leq \int_a^b \sup_S E_{(f,S)}^-(x) dx \leq \int_a^b \inf_S E_{(f,S)}^+(x) dx \leq \lim_{\rho(S) \rightarrow 0} \int_a^b E_{(f,S)}^+(x) dx.$$

Comme les deux côtés de l'inégalité ont des limites égales, il y a égalité partout. On a donc $\int_a^b \sup_S E_{(f,S)}^-(x) dx = \int_a^b \inf_S E_{(f,S)}^+(x) dx$. Comme en plus $\sup_S E_{(f,S)}^- \leq f \leq \inf_S E_{(f,S)}^+$, on a égalité

$$\sup_S E_{(f,S)}^- = f = \inf_S E_{(f,S)}^+$$

en tous les points de continuité, c'est à dire d'après le théorème précédent de Lebesgue, égalité presque partout.

On en déduit que pour presque chaque $x \in [a, b]$, on a

$$f(x) = \sup_S E_{(f,S)}^-(x) = \inf_S E_{(f,S)}^+(x).$$

Comme $E_{(f,S)}^-$ et $E_{(f,S)}^+$ sont mesurables pour les tribus boréliennes (car en escalier), la fonction f est donc limite p.p. de fonctions mesurables. À ce titre, elle est mesurable aussi. \square

Comparaison des intégrales

Proposition 11.3.2 *Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction Riemann-intégrable. Alors elle est λ -intégrable au sens de Lebesgue et les intégrales de Riemann et de Lebesgue sont égales :*

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

Démonstration : Supposons pour simplifier f positive. Quand $\rho(S) \rightarrow 0$, on a p.p.

$$E_{(f,S)}^- \nearrow f, \quad E_{(f,S)}^+ \searrow f.$$

Mais par définition de l'intégrale de Riemann de f , on a $\int_a^b f(x) dx = \lim_{\rho(S) \rightarrow 0} A^-(f, S) =$

$$\lim_{\rho(S) \rightarrow 0} \int_a^b E_{(f,S)}^-(x) dx.$$

Puis d'après le théorème de convergence monotone, on a $\int_{[a,b]} f d\lambda = \lim_{\rho(S) \rightarrow 0} \int_{[a,b]} E_{(f,S)}^- d\lambda.$

Comme pour les fonctions en escalier, les deux types d'intégrales coïncident $\int_{[a,b]} E_{(f,S)}^- d\lambda = \int_a^b E_{(f,S)}^-(x) dx$, la même chose reste vraie en passant à la limite :

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

Si f n'est pas positive, on écrit $f = f^+ - f^-$ avec f^+ et f^- ses parties positive et négative, puis on applique ce qui précède à chacune. \square

Dans la suite, pour simplifier, on utilisera souvent la notation $\int_a^b f(x) dx$ pour l'intégrale $\int_{[a,b]} f d\lambda$ par rapport à la mesure λ .

11.4 Cas des intégrales de Riemann impropres

Soit $-\infty \leq a < b \leq +\infty$ et $f :]a, b[\rightarrow \mathbb{C}$ une fonction localement Riemann-intégrable, c'est à dire qu'on suppose que pour tout $a < c < d < b$, on a $\int_c^d |f(t)| dt < +\infty$.

Par définition, la fonction f a une intégrale (de Riemann) impropre convergente sur $]a, b[$ ssi

$$\lim_{x \searrow a, y \nearrow b} \int_x^y f(t) dt$$

existe, c'est à dire ssi pour toutes suites $x_n \searrow a, y_n \nearrow b$,

$$\int_{x_n}^{y_n} f(t) dt = \int_{[x_n, y_n]} f d\lambda$$

a une limite, notée $\int_a^b f(t) dt$.

Or si f est intégrable sur $]a, b[$ (au sens de Lebesgue), par le théorème de convergence dominée, on a

$$\int_{[x_n, y_n]} |f| d\lambda = \int \mathbf{1}_{[x_n, y_n]} |f| d\lambda \rightarrow \int \mathbf{1}_{]a, b[} |f| d\lambda = \int_{]a, b[} |f| d\lambda$$

car $f \mathbf{1}_{]a, b[}$ est intégrable. Dans ce cas, un passage à la limite dans (??) montre que f a une intégrale impropre de Riemann absolument convergente.

Par contre si f n'est pas intégrable (au sens de Lebesgue) sur $]a, b[$, rien ne permet de passer à la limite dans $\int_{[x_n, y_n]} f d\lambda$ et on ne peut pas considérer l'intégrale de Lebesgue de f sur $]a, b[$.

Il y a alors deux cas de figure :

- Les intégrales impropres convergent absolument, dans ce cas, c'est que $f \in L^1(]a, b[)$ (= elles sont Lebesgue-intégrables).
- Les intégrales sont seulement semi-convergentes et $\int_{]a, b[} f d\lambda$ n'est pas défini au sens de Lebesgue.

Par exemple $\int_0^\infty \frac{\sin x}{x} dx$ est seulement semi-convergente et n'a pas de sens en tant qu'intégrale de Lebesgue alors qu'en tant qu'intégrale impropre de Riemann, elle vaut $\pi/2$.

Il faudrait définir une notion d'intégrale impropre de Lebesgue pour généraliser les intégrales semi-convergente. On ne le fera pas.

Chapitre 12

Espace produit

12.1 σ -algèbre produit

Définition 12.1.1 *Le produit cartésien de deux ensembles A et B noté $A \times B$ est l'ensemble des couples (a, b) avec $a \in A$ et $b \in B$:*

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Soient (X, \mathcal{A}) et (Y, \mathcal{B}) deux espaces mesurables. On considère l'espace produit

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\}.$$

Pour faire de l'intégration sur cet espace, il faut considérer une tribu sur $X \times Y$, la plus naturelle est la tribu produit.

En pratique, on travaille avec $X = Y = \mathbb{R}$ et les tribus boréliennes $\mathcal{A} = \mathcal{B} = \mathcal{B}(\mathbb{R})$.

Définition 12.1.2 *La σ -algèbre produit $\mathcal{A} \otimes \mathcal{B}$ est la tribu de $X \times Y$ engendrée par les pavés, c'est à dire les produits de mesurables $A \times B$, $A \in \mathcal{A}$, $B \in \mathcal{B}$:*

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\mathcal{P}), \quad \mathcal{P} = \{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}\}.$$

- Le produit de la tribu borélienne sur \mathbb{R} par elle-même donne la tribu borélienne sur \mathbb{R}^2 en effet $\mathcal{B}(\mathbb{R}^2)$ est engendrée par les produits d'intervalles ouverts $]a, b[\times]c, d[$ qui engendrent aussi la tribu produit $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. On a donc $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$.

- Plus généralement, la tribu borélienne sur \mathbb{R}^n (engendrée donc par les ouverts de \mathbb{R}^n) s'obtient comme le produit de celles de \mathbb{R} :

$$\mathcal{B}(\mathbb{R}^n) = \underbrace{\mathcal{B}(\mathbb{R}) \otimes \cdots \otimes \mathcal{B}(\mathbb{R})}_{n \text{ fois}}.$$

- Les ensembles mesurables de base de $X \times Y$ sont donc les produits $A \times B$ de mesurables de X et de Y . Lorsque $X = Y = \mathbb{R}$, les plus simples sont mêmes les produits d'intervalles

$[a, b] \times [c, d]$. Cependant, il y a des mesurables de $\mathcal{A} \otimes \mathcal{B}$ plus généraux qui ne peuvent se voir comme des produits de mesurables, par exemple le disque unité $D(0, 1) = \{(x, y) | x^2 + y^2\} \in \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ puisque on l'écrit comme l'image réciproque d'un intervalle (donc mesurable) par une fonction mesurable :

$$D(0, 1) = F^{-1}([0, 1])$$

avec $F(x, y) = x^2 + y^2$ qui est mesurable car continue. De même $K(0, 1) = \{(x, y) | |x| + |y| \leq 1\} \in \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ puisque

$$K(0, 1) = G^{-1}([0, 1])$$

avec $G(x, y) = |x| + |y|$ qui est mesurable car continue.

Noter que $D(0, 1)$ est la boule unité pour la norme euclidienne (ou ℓ_2), $K(0, 1)$ est celle de la norme ℓ_1 et celle de la norme uniforme ℓ_∞ est $[-1, 1] \times [-1, 1]$ mesurable car produit de mesurables.

Définition 12.1.3 Soit $E \subset X \times Y$ et $f : X \times Y \rightarrow Z$, $x \in X$, $y \in Y$. On définit les sections de E :

$$E_x = \{y \in Y | (x, y) \in E\} \subset Y, \quad E^y = \{x \in X | (x, y) \in E\} \subset X$$

et celles de f

$$f_x : \begin{cases} (Y, \mathcal{B}) & \rightarrow Z \\ y & \mapsto f(x, y) \end{cases} \quad f^y : \begin{cases} (X, \mathcal{B}) & \rightarrow Z \\ x & \mapsto f(x, y) \end{cases}$$

Exemples : Avec $X = Y = \mathbb{R}$

$$\begin{aligned} ([1, 3] \times [-4, -2])_{x=2} &= [-4, -2], & ([1, 3] \times [-4, -2])_{x=5} &= \emptyset \\ ([1, 3] \times [-4, -2])^{y=-3} &= [1, 3], & ([1, 3] \times [-4, -2])^{y=0} &= \emptyset. \end{aligned}$$

Plus généralement, soit $A \in \mathcal{A}$, $B \in \mathcal{B}$ des mesurables de X et de Y , alors

$$(A \times B)_x = \begin{cases} B & \text{si } x \in A \\ \emptyset & \text{si } x \notin A \end{cases} \quad (A \times B)^y = \begin{cases} A & \text{si } y \in B \\ \emptyset & \text{si } y \notin B \end{cases}.$$

Proposition 12.1.1 Soit $E \in \mathcal{A} \otimes \mathcal{B}$ et $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow (Z, \mathcal{C})$ mesurable alors

- 1) $E_x \in \mathcal{B}$, $E^y \in \mathcal{A}$ pour tout $x \in X$ et $y \in Y$.
- 2) $f_x : (Y, \mathcal{B}) \rightarrow (Z, \mathcal{C})$ est mesurable et $f^y : (X, \mathcal{A}) \rightarrow (Z, \mathcal{C})$ est mesurable.

Démonstration :

1) Soit $x \in X$, alors $\mathcal{F} = \{E \subset X \times Y | E_x \in \mathcal{B}\}$ est une σ -algèbre car \mathcal{B} en est une (à vérifier) et pour $A \in \mathcal{A}$, $B \in \mathcal{B}$, $x \in X$, on a

$$(A \times B)_x = \begin{cases} B & \text{si } x \in A, \\ \emptyset & \text{sinon.} \end{cases} \quad (12.1)$$

\mathcal{F} contient donc tous les ensembles du type $A \times B$ pour $A \in \mathcal{A}$ et $B \in \mathcal{B}$. Par définition de la tribu produit, on a $\mathcal{F} \supset \mathcal{A} \otimes \mathcal{B}$. Donc, pour tout $x \in X$ et $E \in \mathcal{A} \otimes \mathcal{B}$, on a $E_x \in \mathcal{B}$. On fait de la même façon pour E^y .

2) $f_x^{-1}(C) = \{y \in Y \mid f(x, y) \in C\} = \{y \in Y \mid (x, y) \in f^{-1}(C)\} = f^{-1}(C)_x$. Or $C \in \mathcal{C}$ implique $f^{-1}(C) \in \mathcal{A} \otimes \mathcal{B}$ par mesurabilité de f et $f^{-1}(C)_x \in \mathcal{B}$. La section f_x est donc mesurable. De même pour f^y . \square

12.2 Mesure produit

Rappel (mesure σ -finie) : Une mesure μ sur un espace mesurable (X, \mathcal{A}) est σ -finie s'il existe une décomposition dénombrable de X en $X = \bigcup_{n \in \mathbb{N}} X_n$ avec les X_n mesurables et de mesures $\mu(X_n) < +\infty$.

Un exemple fondamentale est la mesure de Lebesgue λ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ puisque $X = \bigcup_{n \in \mathbb{N}} [-n, n]$ avec $\lambda([-n, n]) = 2n < +\infty$.

Un autre exemple de mesure σ -finie est une mesure de probabilité \mathbb{P} , la mesure est même finie puisque $\mathbb{P}(\Omega) = 1$.

Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés avec μ et ν des mesures σ -finies.

Proposition 12.2.1 Soit $E \in \mathcal{A} \otimes \mathcal{B}$ alors les applications

$$\left\{ \begin{array}{l} (X, \mathcal{A}) \rightarrow [0, +\infty] \\ x \mapsto \nu(E_x) \end{array} \right\}, \quad \left\{ \begin{array}{l} (Y, \mathcal{B}) \rightarrow [0, +\infty] \\ y \mapsto \mu(E_y) \end{array} \right\}$$

sont mesurables.

Admis

Théorème 12.2.1 Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) deux espaces mesurés avec des mesures μ et ν σ -finies. Alors il existe une seule mesure (dite mesure produit) notée $\mu \otimes \nu$ sur $\mathcal{A} \otimes \mathcal{B}$ telle que

$$\mu \otimes \nu(A \times B) = \mu(A)\nu(B), \quad A \in \mathcal{A}, B \in \mathcal{B}.$$

De plus pour tout $E \in \mathcal{A} \otimes \mathcal{B}$,

$$(\mu \otimes \nu)(E) = \int_X \nu(E_x) d\mu = \int_Y \mu(E^y) d\nu.$$

Démonstration : • L'unicité est admise : il faut le théorème dit des π -systèmes.

• D'après le résultat précédent pour $E \in \mathcal{A} \otimes \mathcal{B}$,

$$(\mu \otimes \nu)_1(E) = \int_X \nu(E_x) d\mu, \quad (\mu \otimes \nu)_2(E) = \int_Y \mu(E^y) d\nu$$

ont un sens car $x \mapsto \nu(E_x)$ est \mathcal{A} -mesurable donc sa μ -intégrale existe et $y \mapsto \mu(E^y)$ est \mathcal{B} -mesurable donc sa ν -intégrale existe. On montre qu'il s'agit de mesures :

- $(\mu \otimes \nu)_1(\emptyset) = (\mu \otimes \nu)_2(\emptyset) = 0$ car

$$(\mu \otimes \nu)_1(\emptyset) = \int_X \nu(\emptyset_x) d\mu = \int_X \nu(\emptyset) d\mu = \int_X 0 d\mu = 0.$$

- Soit $E = \bigcup_{n=1}^{+\infty} E_n$ avec E_n deux à deux disjoints on a

$$E_x = \left(\bigcup_{n=1}^{+\infty} E_n \right)_x = \bigcup_{n=1}^{+\infty} (E_n)_x$$

d'où

$$\begin{aligned} (\mu \otimes \nu)_1(E) &= \int_X \nu(E_x) d\mu = \int_X \nu\left(\bigcup_{n=1}^{+\infty} (E_n)_x\right) d\mu \\ &= \int_X \sum_{n=1}^{+\infty} \nu((E_n)_x) d\mu = \sum_{n=1}^{+\infty} \int_X \nu((E_n)_x) d\mu = \sum_{n=1}^{+\infty} (\mu \otimes \nu)_1(E_n) \end{aligned}$$

où on a utilisé le théorème de convergence monotone pour échanger $\int \sum_{n=1}^{+\infty} = \sum_{n=1}^{+\infty} \int$. De même

$$(\mu \otimes \nu)_2\left(\bigcup_{n=1}^{+\infty} E_n\right) = \sum_{n=1}^{+\infty} (\mu \otimes \nu)_2(E_n).$$

Les applications $(\mu \otimes \nu)_1$ et $(\mu \otimes \nu)_2$ sont donc deux mesures. Puis comme avec (12.1) :

$$(\mu \otimes \nu)_1(A \times B) = \int_X \nu((A \times B)_x) d\mu = \int_A \nu(B) d\mu + \int_{A^c} \nu(\emptyset) d\mu = \int_A \nu(B) d\mu = \mu(B)\mu(A)$$

et de même $(\mu \otimes \nu)_2(A \times B) = \mu(A)\nu(B)$, ces mesures conviennent pour le théorème.

Par unicité (admise), la mesure cherchée est $\mu \otimes \nu = (\mu \otimes \nu)_1 = (\mu \otimes \nu)_2$. \square

12.3 Théorèmes de Fubini

Sous de bonnes conditions, le théorème de Fubini permet de permuter les intégrations dans des intégrales multiples. Ainsi les intégrales multiples, ou par rapport à des mesures produits, se ramènent à des intégrales simples emboîtées.

Théorème 12.3.1 (Fubini-Tonelli) Soit $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}) \rightarrow [0, +\infty]$ une fonction mesurable, avec μ et ν des mesures σ -finies sur (X, \mathcal{A}) et (Y, \mathcal{B}) . Alors

- 1) $x \mapsto \int_Y f_x d\nu$ est \mathcal{A} -mesurable et $y \mapsto \int_X f^y d\mu$ est \mathcal{B} -mesurable.

2) Puis, on a

$$\int_{X \times Y} f d(\mu \otimes \nu) = \int_X \left(\int_Y f_x d\nu \right) d\mu = \int_Y \left(\int_X f^y d\mu \right) d\nu.$$

Démonstration : • Pour $f = \mathbf{1}_E$, il s'agit du théorème précédent.

• Pour f étagée positive, on obtient 1) et 2) par linéarité grâce au cas précédent des fonctions indicatrices.

• Pour f mesurable positive quelconque, on trouve s_n une suite croissante de fonctions étagées positives qui converge vers f . Mais alors la suite des sections $(s_n)_x$ est étagée, croissante, positive et converge vers f_x . D'après le théorème de convergence monotone

$$\int_Y f_x d\nu = \lim_n \int_Y (s_n)_x d\nu$$

est donc mesurable car limite de mesurables. Puis en appliquant trois fois la convergence monotone (pour $\mu \otimes \nu$, pour μ et pour ν) et le résultat déjà vérifié pour les fonctions étagées s_n , on a

$$\begin{aligned} \int f d(\mu \otimes \nu) &= \lim_n \int s_n d(\mu \otimes \nu) = \lim_n \int_X \left(\int_Y (s_n)_x d\nu \right) d\mu \\ &= \int_X \lim_n \left(\int_Y (s_n)_x d\nu \right) d\mu = \int_X \left(\int_Y \lim_n (s_n)_x d\nu \right) d\mu \\ &= \int_X \left(\int_Y f_x d\nu \right) d\mu. \end{aligned}$$

Le résultat avec les intégrations dans l'autre sens se justifie rigoureusement de la même façon. \square

Le théorème de Fubini-Tonelli ne s'applique qu'à des fonctions mesurables positives (sans conditions supplémentaires). Pour des fonctions quelconques, on a le résultat suivant

Théorème 12.3.2 (Fubini) Soient (X, \mathcal{A}, μ) et (Y, \mathcal{B}, ν) des espaces σ -finis et $f : (X \times Y, \mathcal{A} \otimes \mathcal{B}, \mu \otimes \nu) \rightarrow \mathbb{R}$ ou \mathbb{C} . Alors

1) Pour μ -presque chaque x , f_x est ν -intégrable et pour ν -presque chaque y , f^y est μ -intégrable.

Posons $I(x) = \int_Y f_x d\nu$ et $J(y) = \int_X f^y d\mu$, alors

2) I et J sont intégrables.

3)

$$\int_{X \times Y} |f| d(\mu \otimes \nu) = \int_X I d\mu = \int_Y J d\nu.$$

On a donc en écrivant les variables d'intégration :

$$\int_{X \times Y} f(x, y) d(\mu \otimes \nu) = \int_X \int_Y f(x, y) d\nu(y) d\mu(x) = \int_Y \int_X f(x, y) d\mu(x) d\nu(y).$$

Démonstration : En appliquant Fubini-Tonelli à $|f|$ qui est mesurable positive, on a

$$\int_X \left(\int_Y |f_x| d\nu \right) d\mu = \int_Y \left(\int_X |f^y| d\mu \right) d\nu = \int_{X \times Y} |f| d(\mu \otimes \nu) < +\infty$$

car f est intégrable pour $\mu \otimes \nu$. On en déduit que $x \mapsto \int_Y |f_x| d\nu$ est finie μ -p.p. et $y \mapsto \int_X |f^y| d\mu$ est finie ν -p.p., car ces fonctions sont positives et d'intégrales finies. Cela justifie le point 1).

Pour le 2), on écrit $f = u^+ - u^- + i(v^+ - v^-)$ où $u = \operatorname{Re}(f)$ et $v = \operatorname{Im}(f)$. On a alors

$$I(x) = \int f_x d\nu = \int u_x^+ d\nu - \int u_x^- d\nu + i \int v_x^+ d\nu - i \int v_x^- d\nu.$$

Par Fubini-Tonelli, les 4 intégrales sont mesurables donc I aussi puis

$$|I(x)| \leq \int_Y |f_x| d\nu$$

On a donc I intégrable. On fait de même pour J .

Pour 3), on écrit à nouveau $f = u^+ - u^- + i(v^+ - v^-)$, on utilise la linéarité et Fubini-Tonelli pour les intégrales de u^+, u^-, v^+, v^- puis on reforme f à la fin. \square

Remarque 12.3.1 • En pratique, on raisonne de la façon suivante :

- 1) On montre que f est mesurable (arguments généraux).
- 2) Pour montrer que f est intégrable, on calcule $\int |f| d(\mu \otimes \nu)$ en appliquant Fubini-Tonelli à la fonction positive $|f|$:

$$\int |f| d(\mu \otimes \nu) = \int_Y \left(\int_X |f(x, y)| d\mu \right) d\nu = \int_X \left(\int_Y |f(x, y)| d\nu \right) d\mu$$

en choisissant la forme la plus convenable (intégrer d'abord en x ou en y) pour faire le calcul.

- 3) On applique Fubini.

• Si F est positive, on peut intervertir directement les intégrations (par la version Fubini-Tonelli du résultat) . Si f ne l'est pas, il faut vérifier l'intégrabilité en calculant l'intégrale de $|f|$ en appliquant par exemple la version Fubini-Tonelli à $|f| > 0$ pour se ramener à des intégrales simples.

- L'utilisation du théorème de Fubini permet de ramener de nombreux calculs d'intégrales doubles (ou triples ou plus généralement multiples) à des calculs successifs d'intégrales simples (aussi bien pour des calculs effectifs que pour montrer des convergences d'intégrales).

- Un autre outil essentiel est le changement de variables que nous verrons en section 12.6.

- Application : grâce au théorème de Fubini, on peut justifier des interversions $\sum \int = \int \sum$, en effet une somme peut se voir comme une intégrale par rapport à une mesure discrète de comptage et il s'agit bien dès lors d'invertir deux intégrales.

12.4 Vecteurs aléatoires

Définition 12.4.1 On appelle vecteur aléatoire toute application de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ mesurable.

Définition 12.4.2 Si on note $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ la i -ème projection qui à $x = (x_1, \dots, x_n)$ associe $p_i(x) = x_i$, on appelle i -ème marginale du vecteur X la variable aléatoire $X_i = p_i(X)$

$$X = (X_1, \dots, X_n), \quad X_i \text{ est la } i\text{-ème marginale.}$$

Définition 12.4.3 La loi du vecteur aléatoire X est la mesure image sur \mathbb{R}^n de la probabilité :

$$\mathbb{P}_X(A_1 \times \dots \times A_n) = \mathbb{P}(X \in A_1 \times \dots \times A_n) = \mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n), \quad A_i \in \mathcal{B}(\mathbb{R}).$$

C'est une mesure dans $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$.

Définition 12.4.4 Un vecteur aléatoire X est discret si l'ensemble de ses valeurs $X(\Omega)$ est discret dans \mathbb{R}^n .

Un vecteur aléatoire X est de loi à densité si

$$\mathbb{P}_X(A) = \int_A f(x_1, \dots, x_n) dx_1 \dots dx_n, \quad A \in \mathcal{B}(\mathbb{R}^n).$$

On vérifie sans peine que comme $\mathbb{P}_X(\mathbb{R}) = \mathbb{P}(X \in \mathbb{R}^n) = 1$, une densité en dimension n satisfait $f(x_1, \dots, x_n) \geq 0$ et

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) dx_1 \dots dx_n = 1.$$

Proposition 12.4.1 Si (X, Y) est un couple aléatoire de loi de densité f , ses lois marginales $\mathbb{P}_X, \mathbb{P}_Y$ sont données par :

$$\forall A \in \mathcal{B}(\mathbb{R}), \mathbb{P}_X(A) = \mathbb{P}(X \in A) = \int_A \int_{\mathbb{R}} f(x, y) dx dy,$$

$$\forall B \in \mathcal{B}(\mathbb{R}), \mathbb{P}_Y(B) = \mathbb{P}(Y \in B) = \int_B \int_{\mathbb{R}} f(x, y) dy dx.$$

Autrement dit, la loi de X est à densité de densité $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$, celle de Y est de densité $f_Y(y) = \int_{-\infty}^{+\infty} f(x, y) dx$.

Démonstration : La preuve est une application directe du théorème de Fubini-Tonelli sur les intégrales doubles une fois qu'on a remarqué que

$$\begin{aligned} \mathbb{P}_X(A) &= \mathbb{P}(X \in A) = \mathbb{P}(X \in A, Y \in \mathbb{R}) = \mathbb{P}((X, Y) \in A \times \mathbb{R}) = \mathbb{P}_{(X, Y)}(A \times \mathbb{R}) \\ &= \int_{A \times \mathbb{R}} f(x, y) dx dy = \int_A \left(\int_{\mathbb{R}} f(x, y) dy \right) dx = \int_A f_X(x) dx \end{aligned}$$

avec la densité annoncée $f_X(x) = \int_{-\infty}^{+\infty} f(x, y) dy$. Ce théorème s'applique sans problème car par définition d'une densité, f est positive (et même intégrable sur \mathbb{R}^2). Idem pour Y . \square

De même si $X = (X_1, \dots, X_n)$ est de densité f , la i -ème marginale X_i est de densité

$$f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_n) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_n.$$

Remarque 12.4.1 Attention, si la connaissance de la loi du couple ou d'un vecteur permet d'en déduire celle des lois marginales, la réciproque est en général fautive. Voir l'exemple suivant.

Exemples : Pour des variables discrètes.

On donne le tableau de $\mathbb{P}(X = x_i, Y = y_j)$:

$X \setminus Y$	$y_1 = -1$	$y_2 = 2$	$y_3 = 3$	$y_4 = 5$	
$x_1 = 0$	0, 1	0, 05	0, 15	0	0, 3
$x_2 = 2$	0, 05	0, 2	0, 05	0, 1	0, 4
$x_3 = 3$	0, 1	0	0, 1	0, 1	0, 3
	0, 25	0, 25	0, 3	0, 2	1

On en déduit la loi de X : $X(\Omega) = \{0, 2, 3\}$ et

$$\mathbb{P}(X = 0) = 0, 3, \quad \mathbb{P}(X = 2) = 0, 4, \quad \mathbb{P}(X = 3) = 0, 3$$

et celle de Y : $Y(\Omega) = \{-1, 2, 3, 5\}$ et

$$\mathbb{P}(Y = -1) = 0, 25, \quad \mathbb{P}(Y = 2) = 0, 25, \quad \mathbb{P}(Y = 3) = 0, 3, \quad \mathbb{P}(Y = 5) = 0, 2.$$

Notons qu'il n'y a pas unicité des couples (X, Y) donnant les mêmes marginales. Ainsi, le couple suivant différent du précédent partage les mêmes marginales.

$X \setminus Y$	$y_1 = -1$	$y_2 = 2$	$y_3 = 3$	$y_4 = 5$	
$x_1 = 0$	0, 1	0, 1	0	0, 1	0, 3
$x_2 = 2$	0, 1	0, 1	0, 1	0, 1	0, 4
$x_3 = 3$	0, 05	0, 05	0, 2	0	0, 3
	0, 25	0, 25	0, 3	0, 2	1

Exemples : Pour des variables à densité.

• Considérons $f(x, y) = \frac{1}{3}\mathbf{1}_{[0,1] \times [-1,2]}(x, y)$. Il s'agit bien d'une densité car f est positive et

$$\begin{aligned}
 \int \int_{\mathbb{R}^2} f(x, y) \, dx dy &= \frac{1}{3} \int \int_{\mathbb{R}^2} \mathbf{1}_{[0,1] \times [-1,2]}(x, y) \, dx dy \\
 &= \frac{1}{3} \int \int_{\mathbb{R}^2} \mathbf{1}_{[0,1]}(x) \times \mathbf{1}_{[-1,2]}(y) \, dx dy \\
 &= \frac{1}{3} \underbrace{\int_{-\infty}^{+\infty} \mathbf{1}_{[0,1]}(x) \, dx}_{=1} \underbrace{\int_{-\infty}^{+\infty} \mathbf{1}_{[-1,2]}(y) \, dy}_{=2-(-1)=3} \\
 &= 1.
 \end{aligned}$$

Considérons un couple (X, Y) de loi de densité f . La loi de X est alors de densité donnée par :

$$\begin{aligned}
 f_X(x) &= \int_{-\infty}^{+\infty} f(x, y) \, dy = \frac{1}{3} \int_{-\infty}^{+\infty} \mathbf{1}_{[0,1] \times [-1,2]}(x, y) \, dy = \frac{1}{3} \int_{-\infty}^{+\infty} \mathbf{1}_{[0,1]}(x) \times \mathbf{1}_{[-1,2]}(y) \, dy \\
 &= \mathbf{1}_{[0,1]}(x) \times \underbrace{\frac{1}{3} \int_{-\infty}^{+\infty} \mathbf{1}_{[-1,2]}(y) \, dy}_{=1} \\
 &= \mathbf{1}_{[0,1]}(x).
 \end{aligned}$$

De la même façon, $f_Y(y) = \frac{1}{3}\mathbf{1}_{[-1,2]}(y)$.

• Soit $f(x, y) = \frac{1}{3\pi} e^{-\frac{x^2+2xy+5y^2}{6}}$. Il s'agit d'une densité car

$$\begin{aligned}
 \int \int_{\mathbb{R}^2} f(x, y) \, dx dy &= \int \int_{\mathbb{R}^2} e^{-\frac{x^2+2xy+5y^2}{6}} \frac{dx dy}{3\pi} \\
 &= \int \int_{\mathbb{R}^2} e^{-\frac{(x+y)^2+4y^2}{6}} \frac{dx dy}{3\pi} = \int \int_{\mathbb{R}^2} e^{-\frac{(x+y)^2}{2 \times 3}} e^{-\frac{4y^2}{2 \times 3}} \frac{dx dy}{3\pi} \\
 &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-\frac{(x+y)^2}{2 \times 3}} \, dx \right) e^{-\frac{4y^2}{2 \times 3}} \frac{dy}{3\pi} = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{-\frac{z^2}{2 \times 3}} \, dz \right) e^{-\frac{4y^2}{2 \times 3}} \frac{dy}{3\pi}
 \end{aligned}$$

$$= \int_{\mathbb{R}} \sqrt{2\pi \times 3} e^{-\frac{4y^2}{2 \times 3}} \frac{dy}{3\pi} = \int_{\mathbb{R}} e^{-\frac{y^2}{2 \times (3/4)}} \frac{dy}{\sqrt{2\pi \times 3/4}} = 1$$

en utilisant que $\int_{\mathbb{R}} e^{-\frac{t^2}{2\sigma^2}} dt = \sqrt{2\pi\sigma^2}$ d'après la normalisation de la loi normale $\mathcal{N}(0, \sigma^2)$.

Considérons un couple (X, Y) de densité f , alors X est de densité

$$\begin{aligned} f_X(x) &= \int_{\mathbb{R}} f(x, y) dy = \int_{\mathbb{R}} e^{-\frac{x^2 + 2xy + 5y^2}{6}} \frac{dy}{3\pi} = \int_{\mathbb{R}} e^{-\frac{(\frac{1}{\sqrt{5}}x + \sqrt{5}y)^2 + 4x^2/5}{6}} \frac{dy}{3\pi} \\ &= \int_{\mathbb{R}} e^{-\frac{(\frac{1}{\sqrt{5}}x + \sqrt{5}y)^2}{6}} e^{-\frac{4x^2}{30}} \frac{dy}{3\pi} = e^{-\frac{4x^2}{30}} \int_{\mathbb{R}} e^{-\frac{z^2}{2 \times 3}} \frac{dz}{3\pi\sqrt{5}} = e^{-\frac{4x^2}{30}} \frac{\sqrt{2\pi \times 3}}{3\pi\sqrt{5}} \\ &= \frac{1}{\sqrt{15\pi/2}} e^{-\frac{4x^2}{30}}. \end{aligned}$$

et Y de densité

$$\begin{aligned} f_Y(y) &= \int_{\mathbb{R}} f(x, y) dx = \int_{\mathbb{R}} e^{-\frac{x^2 + 2xy + 5y^2}{6}} \frac{dx}{3\pi} = \int_{\mathbb{R}} e^{-\frac{(x+y)^2 + 4y^2}{6}} \frac{dx}{3\pi} = \int_{\mathbb{R}} e^{-\frac{(x+y)^2}{2 \times 3}} e^{-\frac{4y^2}{6}} \frac{dx}{3\pi} \\ &= e^{-\frac{4y^2}{6}} \int_{\mathbb{R}} e^{-\frac{(x+y)^2}{2 \times 3}} \frac{dx}{3\pi} = e^{-\frac{4y^2}{6}} \frac{\sqrt{2\pi \times 3}}{3\pi} = \frac{1}{\sqrt{3\pi/2}} e^{-\frac{4y^2}{6}}. \end{aligned}$$

Les marginales X et Y sont donc de lois $\mathcal{N}(0; 15/4)$ et $\mathcal{N}(0; 3/4)$.

12.5 Indépendance de variables aléatoires

Cette notion a déjà été vue en L1-L2. Rappelons d'abord que

Définition 12.5.1 Deux évènements $A, B \in \mathcal{F}$ d'un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ sont indépendants si

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

En particulier, A et B incompatibles ne peuvent pas être indépendants à moins que l'un de deux ne soit de probabilité nulle. Sinon $\mathbb{P}(A \cap B) = \mathbb{P}(\emptyset) = 0$, tandis que $\mathbb{P}(A)\mathbb{P}(B) > 0$.

Il ne faut donc pas confondre les deux notions.

Pour des variables aléatoires, on a :

Définition 12.5.2 (Indépendance de deux v.a.) Deux v.a. X, Y sont dites indépendantes si pour $A, B \in \mathcal{B}(\mathbb{R})$, mesurables de \mathbb{R} , les évènements $\{X \in A\}, \{Y \in B\}$ sont indépendants :

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B).$$

Définition 12.5.3 (Indépendance d'une famille finie de v.a.) Les m variables aléatoires X_1, \dots, X_m sont dites (mutuellement) indépendantes si pour tout boréliens A_1, \dots, A_m , les évènements $\{X_1 \in A_1\}, \dots, \{X_m \in A_m\}$ sont mutuellement indépendants :

$$\mathbb{P}(X_1 \in A_1, \dots, X_m \in A_m) = \mathbb{P}(X_1 \in A_1) \dots \mathbb{P}(X_m \in A_m).$$

Définition 12.5.4 (Indépendance d'une suite de v.a.) Une suite $(X_i)_{i \in \mathbb{N}}$ de v.a. est dite indépendante si toute sous-suite finie de $(X_i)_{i \in \mathbb{N}}$ est indépendante au sens de la définition 12.5.3.

Proposition 12.5.1 Un vecteur aléatoire $X = (X_1, \dots, X_n)$ est à composantes indépendantes ssi sa loi \mathbb{P}_X est une loi produit (de ses lois marginales) :

$$\mathbb{P}_X = \mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n}. \quad (12.2)$$

Démonstration : Soit $B = B_1 \times \cdots \times B_n$ pavé de $\mathcal{B}(\mathbb{R}^n)$, alors

$$\begin{aligned} P_{(X_1, \dots, X_n)}(B) &= \mathbb{P}((X_1, \dots, X_n) \in B) = \mathbb{P}((X_1, \dots, X_n) \in B_1 \times \cdots \times B_n) \\ &= \mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_n \in B_n) \\ &= \mathbb{P}_{X_1}(B_1) \times \cdots \times \mathbb{P}_{X_n}(B_n) = (\mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n})(B_1 \times \cdots \times B_n) \\ &= (\mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n})(B). \end{aligned}$$

Comme les pavés $B = B_1 \times \cdots \times B_n$ engendrent $\mathcal{B}(\mathbb{R}^n)$ alors $\mathbb{P}_{(X_1, \dots, X_n)}$ et $\mathbb{P}_{X_1} \otimes \cdots \otimes \mathbb{P}_{X_n}$ coïncident.

Pour la réciproque, prendre B un pavé et remonter les étapes. \square

Corollaire 12.5.1 Les v.a. X_1, \dots, X_n sont indépendantes ssi pour tout x_1, \dots, x_n , on a

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \mathbb{P}(X_1 \leq x_1) \times \cdots \times \mathbb{P}(X_n \leq x_n).$$

Démonstration : Appliquer le résultat précédent avec $B =]-\infty, x_1] \times \cdots \times]-\infty, x_n]$. C'est suffisant car on montre que cette famille d'ensembles engendre $\mathcal{B}(\mathbb{R}^n)$. \square

Remarque 12.5.1 Dans les cas discret et à densité, on peut préciser les critères d'indépendances :

– Les v.a. discrètes X et Y sont indépendantes si et seulement si

$$\forall x_i \in X(\Omega), \forall y_j \in Y(\Omega), \quad \mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j). \quad (12.3)$$

– Les v.a. X, Y de densités respectives f et g sont indépendantes si et seulement si le couple (X, Y) est de densité $f \otimes g : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto f(x)g(y)$.

Remarque 12.5.2 Une conséquence importante : si on connaît les lois de X et Y , des variables supposées **indépendantes**, on peut reconstruire la loi du couple (X, Y) à partir des marginales par (12.2) (dans le cas discret par (12.3) ou par le produit des densités dans le cas à densité). Insistons sur le fait que ce n'est pas vrai en général quand X et Y ne sont pas indépendantes.

Dans les deux exemples de la page 106, X et Y ne sont pas indépendantes car par exemple pour le premier :

$$\mathbb{P}(X = 2, Y = 2) = 0,2, \quad \text{tandis que} \quad \mathbb{P}(X = 2) \times \mathbb{P}(Y = 2) = 0,4 \times 0,25 = 0,1.$$

et pour le second :

$$\mathbb{P}(X = 3, Y = 5) = 0, \quad \text{tandis que} \quad \mathbb{P}(X = 3) \times \mathbb{P}(Y = 5) = 0,3 \times 0,2 = 0,06.$$

Exemples : • On donne le tableau de la loi d'un couple (X, Y) en donnant les probabilités ponctuelles $\mathbb{P}(X = x_i, Y = y_j)$:

$X \setminus Y$	y_1	y_2	y_3	
x_1	0,12	0,08	0,20	0,4
x_2	0,18	0,12	0,30	0,6
	0,3	0,2	0,5	= 1

On vérifie ici que X et Y sont indépendantes car pour tout $i = 1, 2$ et $j = 1, 2, 3$, on a

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i) \mathbb{P}(Y = y_j).$$

• Considérons le couple (X, Y) de loi donnée par la densité $f_{(X,Y)}(x, y) = \frac{1}{3} \mathbf{1}_{[0,1] \times [-1,2]}(x, y)$. On a vu que X et Y avaient pour densité $f_X(x) = \mathbf{1}_{[0,1]}(x)$ et $f_Y(y) = \frac{1}{3} \mathbf{1}_{[-1,2]}(y)$. On a alors

$$f_{(X,Y)}(x, y) = \frac{1}{3} \mathbf{1}_{[0,1] \times [-1,2]}(x, y) = \mathbf{1}_{[0,1]}(x) \times \frac{1}{3} \mathbf{1}_{[-1,2]}(y) = f_X(x) f_Y(y).$$

Les variables X et Y sont donc indépendantes.

• Soit (X, Y) le couple aléatoire de loi donnée par la densité $f_{(X,Y)}(x, y) = \frac{1}{3\pi} e^{-\frac{x^2 + 2xy + 5y^2}{6}}$. On a vu que les densités marginales sont

$$f_X(x) = \frac{1}{\sqrt{15\pi/2}} e^{-\frac{4x^2}{30}}, \quad f_Y(y) = \frac{1}{\sqrt{3\pi/2}} e^{-\frac{4y^2}{6}}.$$

On a alors

$$f_X(x) f_Y(y) = \frac{1}{\sqrt{15\pi/2}} e^{-\frac{4x^2}{30}} \times \frac{1}{\sqrt{3\pi/2}} e^{-\frac{4y^2}{6}} \neq \frac{1}{3\pi} e^{-\frac{x^2 + 2xy + 5y^2}{6}} = f_{(X,Y)}(x, y).$$

Dans ce cas, X et Y ne sont pas indépendantes.

Proposition 12.5.2 Soient X_i $1 \leq i \leq n$, des v.a. indépendantes et $h_i : \mathbb{R} \rightarrow \mathbb{C}$ des fonctions boréliennes telles que $h_i(X_i)$ soient \mathbb{P} -intégrables. Alors $\prod_{i=1}^n h_i(X_i)$ est \mathbb{P} -intégrable et

$$E\left[\prod_{i=1}^n h_i(X_i)\right] = \prod_{i=1}^n E[h_i(X_i)].$$

Démonstration : On utilise le théorème de transfert, l'indépendance et Fubini :

$$\begin{aligned} E\left[\prod_{i=1}^n h_i(X_i)\right] &= \int \prod_{i=1}^n h_i(x_i) d\mathbb{P}_{(X_1, \dots, X_n)} = \int \prod_{i=1}^n h_i(x_i) d\mathbb{P}_{X_1}(x_1) \dots d\mathbb{P}_{X_n}(x_n) \\ &= \int h(x_1) d\mathbb{P}_{X_1}(x_1) \times \dots \times \int h(x_n) d\mathbb{P}_{X_n}(x_n) \\ &= \prod_{i=1}^n E[h_i(X_i)]. \end{aligned}$$

□

Corollaire 12.5.2 Soient X_1, \dots, X_n des v.a. réelles ou complexes indépendantes

$$E[X_1 \dots X_n] = E[X_1] \dots E[X_n].$$

Démonstration : Appliquer le résultat précédent avec la fonction $h_i(x) = x$. □

La réciproque est fautive : soit X_1 de loi uniforme sur $[-1, 1]$ et $X_2 = X_1^2$, on a $E(X_1 X_2) = E(X_1^3) = \int_{-1}^{+1} x_1^3 dx_1 = 0$ et $E(X_1) = 0$ si bien qu'on a $E(X_1 X_2) = 0 = E(X_1)E(X_2)$ mais X_1 et X_2 ne sont pas indépendantes car par exemple

$$\mathbb{P}(X_1 \in [0, 1/2], X_2 \in [0, 1/4]) = 1/4, \quad \mathbb{P}(X_1 \in [0, 1/2]) \times \mathbb{P}(X_2 \in [0, 1/4]) = 1/8.$$

12.6 Changement de variable

On a vu un résultat de changement de variable formel (le théorème de transfert) avec pour seule condition d'avoir un changement de variable $y = \varphi(x)$ mesurable. Malheureusement, la nouvelle mesure $\nu = \mu \circ \varphi^{-1} = \mu \circ \varphi^{-1}$ n'est pas explicite du tout. Ce résultat est donc essentiellement abstrait et difficile à utiliser pour des calculs explicites.

On propose ici des résultats plus explicites, avec des conditions plus restrictives sur le changement de variables.

12.6.1 Rappel : intégrale de Riemann

Soit I un intervalle de \mathbb{R} et $\varphi : I \rightarrow \mathbb{R}$ strictement monotone et C^1 tel que φ' ne s'annule pas sur I . Alors on a

$$\int_I f(x) dx = \int_{\varphi(I)} f(\varphi^{-1}(y)) |(\varphi^{-1})'(y)| dy.$$

Pour cela, on pose $y = \varphi(x)$ ou $x = \varphi^{-1}(y)$ et en dérivant on a la relation entre dx et dy dérivant

$$\frac{dx}{dy} = (\varphi^{-1})'(y) \quad \text{c'est à dire} \quad dx = (\varphi^{-1})'(y)dy.$$

12.6.2 Changement de variable

Définition 12.6.1 Soit $F : D \subset \mathbb{R}^n \rightarrow D' \subset \mathbb{R}^n$ où D et D' sont des ouverts. F est appelé un *difféomorphisme* si c'est une bijection de classe C^1 dont la bijection réciproque est aussi de classe C^1 .

Définition 12.6.2 La matrice jacobienne d'un changement de variable

$$y = F(x) \iff (y_1, \dots, y_n) = (F_1(x_1, \dots, x_n), \dots, (F_n(x_1, \dots, x_n)))$$

est

$$J_F(x) = J_F(x_1, \dots, x_n) = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1} & \dots & \frac{\partial F_n}{\partial x_n} \end{bmatrix}.$$

Le jacobien est le déterminant de la matrice jacobienne

La matrice jacobienne est la matrice des dérivées partielles.

Rappel : Calculs des déterminants d'ordre 2 et 3 :

- $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$

- Règle de Sarrus : $\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3 - a_3b_2c_1 - b_3c_2a_1 - c_3a_2b_1.$

- Développements selon une ligne ou une colonne pour se ramener à des déterminants d'ordre inférieur.

Théorème 12.6.1 (Changement de variable) Soit V un ouvert de \mathbb{R}^d et φ un C^1 -difféomorphisme de V dans $\varphi(V) \subset \mathbb{R}^d$. Alors, on a les formules de changements de variables

$$\begin{aligned} \int_V f(x)d\lambda(x) &= \int_{\varphi(V)} f(\varphi^{-1}(y)) |J_{\varphi^{-1}}(y)|d\lambda(y) \\ \int_V h(\varphi(x))d\lambda(x) &= \int_{\varphi(V)} h(y) |J_{\varphi^{-1}}(y)|d\lambda(y) \end{aligned}$$

pour toutes fonctions f et h mesurables telle que f est λ -intégrable et $h \circ \varphi$ est λ -intégrable.

Démonstration : Admis.

12.6.3 Coordonnées polaires et sphériques

• Un changement de variables utile dans le plan \mathbb{R}^2 est le changement de variables en polaire qui consiste à passer de (x, y) représentant des coordonnées cartésiennes dans un repère orthonormé à (r, θ) les coordonnées polaires correspondantes données par

$$(r, \theta) = \varphi(x, y) \iff \varphi^{-1} : \begin{cases} x = r \cos \theta \\ y = r \sin \theta \end{cases}, \quad r \in [0, +\infty[, \theta \in [0, 2\pi[.$$

On remplace alors $dxdy$ par $rdrd\theta$ car le jacobien du changement de variable est r :

$$J_{\varphi^{-1}}(r, \theta) = \begin{vmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{vmatrix} = r \cos^2 \theta + r \sin^2 \theta = r.$$

Ainsi :

$$\begin{aligned} \int \int_{\mathbb{R}^2} f(x, y) dx dy &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy \\ &= \int \int_{[0, +\infty[\times [0, 2\pi[} f(r \cos \theta, r \sin \theta) r dr d\theta. \end{aligned}$$

Exemples :

- Normalisation de la loi normale $\int_{-\infty}^{+\infty} e^{-x^2/2} dx = \sqrt{2\pi}$.

Notons $I = \int_{-\infty}^{+\infty} e^{-x^2/2} dx$ et montrons que $I^2 = 2\pi$. On a

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} e^{-x^2/2} dx \times \int_{-\infty}^{+\infty} e^{-y^2/2} dy \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2/2} e^{-y^2/2} dx dy = \int \int_{\mathbb{R} \times \mathbb{R}} e^{-(x^2+y^2)/2} dx dy \\ &= \int_0^{2\pi} \int_0^{+\infty} e^{-r^2/2} r dr d\theta \\ &= \int_0^{2\pi} d\theta \int_0^{+\infty} r e^{-r^2/2} dr = 2\pi \left[-e^{-r^2/2} \right]_0^{+\infty} = 2\pi \end{aligned}$$

où on a utilisé le théorème de Fubini à la 2ème ligne puis on a fait un changement de variables en polaires à la 3ème ligne.

- Aire d'un disque : $\Delta = \{(x, y) \mid x^2 + y^2 \leq R^2\}$:

$$\lambda_2(B(0, R^2)) = \iint_{B(0, R)} dx dy = \iint_{[0, R] \times [0, 2\pi[} r dr d\theta = \int_0^{2\pi} d\theta \int_0^R r dr = 2\pi \left[\frac{r^2}{2} \right]_0^R = \pi R^2.$$

• En dimension 3, le changement de variables utile est le changement en coordonnées sphériques donné par

$$(r, \theta, \varphi) = \phi(x, y, z) \iff \phi^{-1} : \begin{cases} x = r \cos \theta \cos \varphi \\ y = r \cos \theta \sin \varphi \\ z = r \sin \theta \end{cases}$$

où $\theta \in]-\pi/2, \pi/2[$ est la latitude, $\varphi \in [0, 2\pi[$ est la longitude et $r \in [0, +\infty[$ la distance à l'origine.

Le jacobien du changement de variable est

$$J_{\phi^{-1}}(r, \theta, \varphi) = \begin{bmatrix} \cos \theta \cos \varphi & -r \sin \theta \cos \varphi & -r \cos \theta \sin \varphi \\ \cos \theta \sin \varphi & -r \sin \theta \sin \varphi & r \sin \theta \cos \varphi \\ \sin \theta & r \cos \theta & 0 \end{bmatrix} = r^2 \cos \theta.$$

Ainsi :

$$\begin{aligned} & \int \int \int_{\mathbb{R}^3} f(x, y, z) dx dy dz \\ &= \int \int \int_{[0, +\infty[\times [0, 2\pi[\times]-\frac{\pi}{2}, \frac{\pi}{2}[} f(r \cos \theta \cos \varphi, r \cos \theta \sin \varphi, r \sin \theta) r^2 \cos \theta dr d\theta d\varphi. \end{aligned}$$

Ce type de changement de variable (polaire en dimension 2, sphérique en dimension 3) se généralise en dimension n avec

$$\begin{cases} x_1 &= r \cos \theta_1 \cos \theta_2 \dots \cos \theta_{n-2} \cos \theta_{n-1}, \\ x_2 &= r \cos \theta_1 \cos \theta_2 \dots \cos \theta_{n-2} \sin \theta_{n-1}, \\ x_3 &= r \cos \theta_1 \cos \theta_2 \dots \cos \theta_{n-3} \sin \theta_{n-2}, \\ x_4 &= r \cos \theta_1 \cos \theta_2 \dots \cos \theta_{n-4} \sin \theta_{n-3}, \\ \dots &= \dots \\ x_{n-1} &= r \cos \theta_1 \sin \theta_2, \\ x_n &= r \sin \theta_1. \end{cases}$$

Exemple : Calcul du volume d'une boule euclidienne de rayon R est

$$\begin{aligned} \lambda_3(B(0, R)) &= \iiint_{B(0, R)} dx dy dz = \iiint_{[0, R] \times]-\pi/2, \pi/2[\times [0, 2\pi[} r^2 \cos \theta dr d\theta d\varphi \\ &= \int_0^{2\pi} d\varphi \int_{-\pi/2}^{\pi/2} \cos \theta d\theta \int_0^R r^2 dr = 2\pi [\sin \theta]_{-\pi/2}^{\pi/2} \left[\frac{r^3}{3} \right]_0^R \\ &= \frac{4}{3} \pi R^3 \end{aligned}$$

où λ_3 désigne la mesure de Lebesgue en dimension 3.

Autre application du changement de variable : calcul de loi par changement de variables.

Chapitre 13

Espaces L^p

13.1 Convexité

Définition 13.1.1 Soit I un intervalle. Une fonction $\varphi : I \rightarrow \mathbb{R}$ est convexe ssi $\forall x, y \in I$ et $\forall t \in [0, 1]$ alors

$$\varphi(tx + (1-t)y) \leq t\varphi(x) + (1-t)\varphi(y). \quad (13.1)$$

φ est dite concave si $-\varphi$ est convexe.

Une fonction φ est convexe si entre deux points x et y , la corde entre $(x, f(x))$ et $(y, f(y))$ est sous le segment entre $(x, f(x))$ et $(y, f(y))$.

Proposition 13.1.1 Une fonction φ est convexe sur I ssi $\forall x < y < z$ dans I , on a

$$\frac{\varphi(y) - \varphi(x)}{y - x} \leq \frac{\varphi(z) - \varphi(x)}{z - x}.$$

Exemples :

- Une fonction $\varphi \in C^1$ est convexe ssi sa dérivée φ' est croissante.
- Une fonction $\varphi \in C^2$ est convexe ssi $\varphi'' \geq 0$.
- $f(x) = x^{2n}$ est convexe sur \mathbb{R} , $g(x) = x^{2n+1}$ n'est convexe que sur \mathbb{R}^+ , $h(x) = ax + b$ affine est convexe, \exp est convexe sur \mathbb{R} , \log est concave (c'est à dire $-\log$ est convexe).

Théorème 13.1.1 (Jensen) Soit (X, \mathcal{A}, μ) un espace de probabilité et $f : X \rightarrow I \subset \mathbb{R}$ intégrable et $\varphi : I \rightarrow \mathbb{R}$ convexe. Alors si $\varphi \circ f$ est intégrable ou positive, on a

$$\varphi\left(\int_X f d\mu\right) \leq \int_X \varphi \circ f d\mu.$$

Démonstration : Posons $y = \int f d\mu$. Par convexité de φ , en interprétant $\int f d\mu$ comme un barycentre, on montre que $y \in I$ car f est à valeurs dans I . Pour y fixé, par convexité de φ , il existe $\alpha \in \mathbb{R}$ (cf. (13.1)) tel que

$$\sup_{x < y} \frac{\varphi(y) - \varphi(x)}{y - x} \leq \alpha \leq \inf_{z > y} \frac{\varphi(z) - \varphi(y)}{z - y}.$$

En particulier, on a pour tout $x \in I$:

$$\varphi(x) \geq \varphi(y) + \alpha(x - y).$$

Avec $x = f(u) \in I$, on a : $\varphi \circ f(u) \geq \varphi(y) + \alpha(f(u) - y)$.

Comme φ est convexe, elle est mesurable et $\varphi \circ f$ l'est aussi. On a ainsi

$$\begin{aligned} \int \varphi \circ f(u) d\mu(u) &\geq \int \varphi(y) d\mu(u) + \alpha \left(\int f(u) d\mu(u) - y\mu(X) \right) \\ &\geq \varphi(y)\mu(X) + \alpha \times 0 \\ &= \varphi \left(\int f d\mu \right) \end{aligned}$$

car $\mu(X) = 1$ (μ est une mesure de probabilité) et par définition $y = \int f d\mu$. □

13.2 Espace \mathcal{L}^p

Définition 13.2.1 Soit (X, \mathcal{A}, μ) un espace mesuré et $f : X \rightarrow \mathbb{C}$ mesurable.

- On définit pour $p \in [1, +\infty[$

$$\|f\|_p = \left(\int_X |f|^p d\mu \right)^{1/p}.$$

- Puis M est un majorant essentiel de f si $\mu\{x \in X \mid |f(x)| > M\} = 0$.
- On définit alors

$$\|f\|_\infty = \begin{cases} +\infty & \text{si } f \text{ n'a pas de majorant essentiel,} \\ \text{l'inf des majorants essentiels} & \text{sinon.} \end{cases}$$

Par exemple, sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $\|\mathbf{1}_\mathbb{Q}\|_\infty = 0$.

Définition 13.2.2 • On note $\mathcal{L}^p(X, \mathcal{A}, \mu)$ l'ensemble des fonctions mesurables de puissance p -ième μ -intégrable.

- On note aussi $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ l'ensemble des fonctions mesurables bornées.

Proposition 13.2.1 Les ensembles $\mathcal{L}^p(X, \mathcal{A}, \mu)$ et $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ sont des espaces vectoriels.

Démonstration : On montre qu'il s'agit d'un sous-espace vectoriel de l'espace des fonctions de X dans \mathbb{C} . Pour cela, on montre la stabilité par addition et par multiplication par un scalaire.

C'est clair pour $p = \infty$. Pour $p \in [1, \infty[$: Par convexité de la fonction $x \mapsto x^p$ sur \mathbb{R}^+ , on a pour $a, b > 0$: $(\frac{a+b}{2})^p \leq \frac{1}{2}a^p + \frac{1}{2}b^p$. Avec $a = |f(x)|$ et $b = |g(x)|$, il vient

$$\left(\frac{|f(x) + g(x)|}{2}\right)^p \leq \left(\frac{|f(x)| + |g(x)|}{2}\right)^p \leq \frac{1}{2}|f(x)|^p + \frac{1}{2}|g(x)|^p.$$

En intégrant par rapport à μ , et en multipliant des deux côtés par 2^p , on déduit :

$$\int_X |f + g|^p d\mu \leq 2^{p-1} \int_X |f|^p d\mu + 2^{p-1} \int_X |g|^p d\mu$$

qui est fini lorsque f et g sont dans $\mathcal{L}^p(X, \mathcal{A}, \mu)$. C'est donc que $f + g \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ comme par ailleurs $af \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ quand $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$, on a bien un espace vectoriel. \square

Rappel (Norme et semi-norme) : $\|\cdot\| : E \rightarrow \mathbb{R}_+$ est une norme sur un espace vectoriel E si

- $x = 0 \iff \|x\| = 0$
- Pour tout $x \in E$ et $\lambda \in \mathbb{C}$, $\|\lambda x\| = |\lambda| \|x\|$
- Inégalité triangulaire : pour tout $x, y \in E$ $\|x + y\| \leq \|x\| + \|y\|$.

$\|\cdot\|$ est seulement une semi-norme si le premier point est remplacé par $x = 0 \Rightarrow \|x\| = 0$, c'est à dire un vecteur x peut être de norme $\|x\| = 0$ sans être nul.

En fait, on définit des (semi) normes à partir de ces quantités $\|\cdot\|_p$.

On vérifie facilement que $\|\alpha f\|_p = |\alpha| \|f\|_p$.

Par contre notons que $\|f\|_p = 0$ n'entraîne pas $f = 0$ mais seulement que $f = 0$ p.p. En effet, par l'inégalité de Markov, on a

$$\mu\{x \mid |f(x)| \geq 1/n\} \leq n^p \int |f|^p d\mu = 0$$

et alors comme $\{x \in X \mid |f(x)| > 0\} = \bigcup_{n=1}^{+\infty} \{x \in X \mid |f(x)| \geq 1/n\}$, par la croissance séquentielle de la mesure μ :

$$\mu\{x \mid f(x) \neq 0\} = \mu\{x \mid |f(x)| > 0\} = \lim_{n \rightarrow +\infty} \mu\{x \mid |f(x)| \geq 1/n\} = \lim_{n \rightarrow +\infty} 0 = 0.$$

Pour avoir vraiment une norme il faudrait que $\|f\|_p = 0$ entraîne $f = 0$, c'est à dire $f(x) = 0$ pour tout x et pas seulement pour μ -presque tous les x . Pour remédier à ce problème, on va identifier les fonctions qui ont la même valeur presque partout. Ainsi, $f = 0$ p.p. est identifiée à la fonction nulle (qui vaut zéro toujours).

13.3 Espaces L^p

Formellement, on définit la relation d'équivalence \sim par $f \sim g$ ssi il existe $A \in \mathcal{A}$ tel que $\mu(X \setminus A) = 0$ (c'est à dire A^c est négligeable) et pour tout $x \in A$, $f(x) = g(x)$ (c'est à dire f égale g p.p.) et on considère désormais l'espace quotient

Définition 13.3.1

$$\begin{aligned} L^p(X, \mathcal{A}, \mu) &= \mathcal{L}^p(X, \mathcal{A}, \mu) / \sim \\ L^\infty(X, \mathcal{A}, \mu) &= \mathcal{L}^\infty(X, \mathcal{A}, \mu) / \sim \end{aligned}$$

• $L^p(X, \mathcal{A}, \mu)$ est l'espace vectoriel des classes d'équivalence des fonctions modulo $f = g$ p.p. telles que $\|f\|_p < +\infty$ (i.e. telles que $|f|^p$ est μ -intégrable).

• $L^\infty(X, \mathcal{A}, \mu)$ est l'espace vectoriel des classes d'équivalence des fonctions modulo $f = g$ p.p. telles que $\|f\|_\infty < +\infty$ (i.e. bornées μ -p.p.).

Par exemple $\mathbf{1}_\mathbb{Q}$ est identifiée à 0 dans les espaces $L^p(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Plus simplement :

Définition 13.3.2 L'ensemble $L^p(X, \mathcal{A}, \mu)$ est l'ensemble des fonctions $\mathcal{L}^p(X, \mathcal{A}, \mu)$ où on identifie toutes les fonctions qui se valent μ -presque partout. C'est l'espace vectoriel des fonctions telles que $\|f\|_p < +\infty$ (i.e. telles que $|f|^p$ est μ -intégrable).

L'ensemble $L^\infty(X, \mathcal{A}, \mu)$ est l'ensemble des fonctions $\mathcal{L}^\infty(X, \mathcal{A}, \mu)$ où on identifie toutes les fonctions qui se valent μ -presque partout. C'est l'espace vectoriel des fonctions telles que $\|f\|_\infty < +\infty$ (c'est à dire bornées presque partout).

Si $\mathcal{A} = \mathcal{P}(X)$ et μ est la mesure de comptage (ou de dénombrement), on note plutôt $\ell^p(X)$, ainsi

$$\begin{aligned} \ell^p(\mathbb{N}) &= \{(u_n)_n \text{ avec } \sum_{n=0}^{+\infty} |u_n|^p < +\infty\} \\ \ell^\infty(\mathbb{N}) &= \{\text{suites bornées}\}. \end{aligned}$$

Sur les espaces $L^p(X, \mathcal{A}, \mu)$, les semi normes $\|\cdot\|_p$ deviennent de vraies normes car on a vu que $\|f\|_p = 0$ entraîne $f = 0$ p.p., c'est à dire que f est la classe nulle (f est identifiée à la fonction nulle).

Il reste quand même à voir l'inégalité triangulaire pour $\|\cdot\|_p$. On l'obtiendra par l'inégalité de Minkowski à venir. Avant, on définit :

Définition 13.3.3 Soient $p, q \in [1, +\infty]$. Ce sont des exposants conjugués ssi

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Par exemple, 1 et $+\infty$ sont conjugués, 2 est son propre conjugué, 3 a pour conjugué $\frac{3}{2}$.

13.4 Inégalités (Hölder, Minkowski)

Théorème 13.4.1 Soit (X, \mathcal{A}, μ) un espace mesuré et $f, g : X \rightarrow \mathbb{C}$ des fonctions mesurables, p, q des exposants conjugués dans $[1, +\infty]$. Alors si $f \in L^p(X, \mathcal{A}, \mu)$ et $g \in L^q(X, \mathcal{A}, \mu)$, $fg \in L^1(X, \mathcal{A}, \mu)$ avec

$$\|fg\|_1 = \int_X |fg| d\mu \leq \|f\|_p \|g\|_q \quad (\text{inégalité de Hölder}).$$

Pour $p = q = 2$, il s'agit de l'inégalité de **Cauchy-Schwarz** :

$$\left(\int fg d\mu \right)^2 \leq \int |f|^2 d\mu \int |g|^2 d\mu.$$

Puis, on a aussi pour $f, g \in L^p(X, \mathcal{A}, \mu)$:

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \quad (\text{inégalité de Minkowski}).$$

Démonstration :(Holdër) Il suffit de le faire pour $f, g \geq 0$.

• Si p (ou q) = $+\infty$, on a $f(x) \leq \|f\|_\infty$ p.p. d'où $f(x)g(x) \leq \|f\|_\infty g(x)$, ce qui donne en intégrant

$$\int_X fg d\mu \leq \|f\|_\infty \int_X g d\mu = \|f\|_\infty \|g\|_1.$$

Puis si $f, g \in L^1(X, \mathcal{A}, \mu)$, en intégrant, on obtient directement

$$\int_X (f + g) d\mu = \int_X f d\mu + \int_X g d\mu \iff \|f + g\|_1 = \|f\|_1 + \|g\|_1.$$

pour f et g de signe quelconque l'égalité devient une inégalité.

• Soient $p, q \neq +\infty, 1$ avec pour commencer $\|f\|_p = \|g\|_q = 1$. Si $f(x), g(x) \neq 0, +\infty$ on peut écrire $f(x)^p = e^u$ et $g(x)^q = e^v$ (en prenant $u = \ln(f(x)^p)$ et $v = \ln(g(x)^q)$). Par convexité de \exp , on a alors :

$$e^{\frac{1}{p}u + \frac{1}{q}v} \leq \frac{1}{p}e^u + \frac{1}{q}e^v,$$

ce qui se réécrit $f(x)g(x) \leq \frac{1}{p}f(x)^p + \frac{1}{q}g(x)^q$. L'inégalité est en fait encore vraie si $f(x)$ ou $g(x)$ vaut 0 ou $+\infty$. En l'intégrant alors sur tout X , on a

$$\|fg\|_1 = \int_X fg d\mu \leq \frac{1}{p} \int_X f^p d\mu + \frac{1}{q} \int_X g^q d\mu = \frac{1}{p} \|f\|_p^p + \frac{1}{q} \|g\|_q^q = \frac{1}{p} + \frac{1}{q} = 1$$

car d'abord $\|f\|_p = \|g\|_q = 1$ puis ensuite parce que p, q sont conjugués. Ce qui prouve l'inégalité de Hölder dans ce cas.

• Si $p, q \neq 1, +\infty$ et si $\|f\|_p$ et $\|g\|_q \neq 0, +\infty$ alors on considère

$$\tilde{f} = f/\|f\|_p, \quad \tilde{g} = g/\|g\|_q.$$

On a facilement $\|\tilde{f}\|_p = \|\tilde{g}\|_q = 1$ si bien que le cas précédent donne :

$$\|\tilde{f}\tilde{g}\|_1 \leq 1 \iff \|fg\|_1 \leq \|f\|_p \|g\|_q.$$

• Si $\|f\|_p = 0$ alors $f = 0$ p.p. et donc $fg = 0$ p.p. si bien que $\int_X fg d\mu = \|fg\|_1 = 0$ et l'inégalité cherchée se réécrit $0 \leq 0$, ce qui est vraie.

• Si $\|f\|_p = +\infty$ alors il suffit de voir le cas $\|g\|_q \neq 0$ sinon on se ramène au cas précédent (avec g à la place de f). Mais dans ce cas, l'inégalité devient une majoration par $+\infty$

ce qui est encore vrai. \square

Démonstration : (Minkowski) Remarquons que

$$(f + g)^p = (f + g)(f + g)^{p-1} = f(f + g)^{p-1} + g(f + g)^{p-1}.$$

Soit q conjugué de p alors d'après l'inégalité de Hölder

$$\begin{aligned} \int (f + g)^p d\mu &= \int f(f + g)^{p-1} d\mu + \int g(f + g)^{p-1} d\mu \\ &\leq \left(\int f^p d\mu \right)^{1/p} \left(\int (f + g)^{(p-1)q} d\mu \right)^{1/q} + \left(\int g^p d\mu \right)^{1/p} \left(\int (f + g)^{(p-1)q} d\mu \right)^{1/q} \\ &\leq \left[\left(\int f^p d\mu \right)^{1/p} + \left(\int g^p d\mu \right)^{1/p} \right] \left(\int (f + g)^{(p-1)q} d\mu \right)^{1/q}. \end{aligned}$$

Comme $(p - 1)q = p$ et $1/q = (p - 1)/p$, on a

$$\begin{aligned} \int (f + g)^p d\mu &\leq \left[\left(\int f^p d\mu \right)^{1/p} + \left(\int g^p d\mu \right)^{1/p} \right] \left(\int (f + g)^p d\mu \right)^{\frac{p-1}{p}} \\ \|f + g\|_p^p &\leq [\|f\|_p + \|g\|_p] \|f + g\|_p^{p-1}. \end{aligned} \quad (13.2)$$

• Si $\|f + g\|_p \neq 0, +\infty$, l'inégalité de Minkowski suit en simplifiant par $\|f + g\|_p^{p-1}$ dans (13.2).

• Si $\|f + g\|_p = 0$, l'inégalité de Minkowski est immédiate.

• Si $\|f + g\|_p = +\infty$, par convexité de la fonction $x \mapsto x^p$ pour $p > 1$, on a

$$\left(\frac{f + g}{2} \right)^p \leq \frac{1}{2} f^p + \frac{1}{2} g^p \iff (f + g)^p \leq 2^{p-1} f^p + 2^{p-1} g^p.$$

On a donc $\int (f + g)^p d\mu \leq 2^{p-1} \int f^p d\mu + 2^{p-1} \int g^p d\mu$. Mais alors si $\int (f + g)^p d\mu$ est infini, $2^{p-1} \int f^p d\mu + 2^{p-1} \int g^p d\mu$ l'est aussi donc $\int f^p d\mu$ ou $\int g^p d\mu$ l'est si bien que l'inégalité de Minkowski devient $+\infty \leq +\infty$, ce qui est encore vraie. \square

Corollaire 13.4.1 $f \mapsto \|f\|_p$ est une norme sur $L^p(X, \mathcal{A}, \mu)$.

Démonstration : Il ne manquait plus que l'inégalité triangulaire, mais c'est exactement l'inégalité de Minkowski. \square

L'espace $L^p(X, \mathcal{A}, \mu)$ est donc un espace vectoriel normé. On a mieux avec le résultat suivant (culturel, admis) :

Théorème 13.4.2 (Riesz-Fisher) $L^p(X, \mathcal{A}, \mu)$ est un espace complet (pour la norme $\|\cdot\|_p$). Il s'agit donc d'un espace de Banach.

Le résultat suivant donne un (petit) lien entre la convergence d'une suite de fonctions $(f_n)_n$ vers f dans $L^p(X, \mathcal{A}, \mu)$ et la convergence μ -presque partout.

Théorème 13.4.3 Si $f_n \rightarrow f$ dans $L^p(X, \mathcal{A}, \mu)$, c'est à dire $\|f_n - f\|_p \rightarrow 0$ quand $n \rightarrow +\infty$ alors il existe une sous-suite $(f_{n_p})_p$ de $(f_n)_n$ qui converge p.p. vers f .

Comparaison des espaces L^p

En général, il n'y a pas de relations d'inclusion entre les espaces $L^p(X, \mathcal{A}, \mu)$ pour différents exposants p . Par exemple, considérons $(X, \mathcal{A}, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et les espaces $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Alors avec

$$f(x) = \frac{1}{\sqrt{x}} \mathbf{1}_{]0,1[}(x), \quad g(x) = \frac{1}{x} \mathbf{1}_{[1,+\infty[}(x)$$

on a $f \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ mais $f \notin L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ en effet

$$\begin{aligned} \|f\|_1 &= \int |f(x)| d\lambda = \int_0^1 \frac{dx}{\sqrt{x}} = \frac{1}{2} \\ \|f\|_2^2 &= \int |f(x)|^2 d\lambda = \int_0^1 \frac{dx}{x} = +\infty \end{aligned}$$

alors que $g \in L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ mais $g \notin L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ en effet

$$\begin{aligned} \|g\|_1 &= \int |g(x)| d\lambda = \int_1^{+\infty} \frac{dx}{x} = +\infty \\ \|g\|_2^2 &= \int |g(x)|^2 d\lambda = \int_1^{+\infty} \frac{dx}{x^2} = 1. \end{aligned}$$

On n'a donc aucune inclusion entre les espaces $L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ et $L^2(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$.

Par contre si μ est une mesure **finie** (par exemple une probabilité), les espaces $L^p(X, \mathcal{A}, \mu)$ sont ordonnés pour l'inclusion :

Théorème 13.4.4 Soit (X, \mathcal{A}, μ) un espace fini (c'est à dire $\mu(X) < +\infty$) alors pour $p \geq q$:

$$L^p(X, \mathcal{A}, \mu) \subset L^q(X, \mathcal{A}, \mu).$$

Démonstration : • D'abord si $p = +\infty$: si $f \in L^\infty(X, \mathcal{A}, \mu)$ alors $f \in L^q(X, \mathcal{A}, \mu)$, en effet $|f(x)| \leq \|f\|_\infty$ pour presque chaque x et donc,

$$\|f\|_q^q = \int |f(x)|^q d\mu(x) \leq \left(\int \|f\|_\infty^q d\mu(x) \right)^{1/q} \leq \left(\|f\|_\infty^q \int d\mu \right)^{1/q} = \|f\|_\infty \mu(X)^{1/q} < +\infty.$$

Donc $f \in L^q(X, \mathcal{A}, \mu)$.

• Puis si $f \in L^p(X, \mathcal{A}, \mu)$ alors $f \in L^q(X, \mathcal{A}, \mu)$, en effet $\frac{q}{p} + \frac{p-q}{p} = 1$

$$\|f\|_q = \left(\int |f|^q \times 1 d\mu \right)^{1/q} \leq \left(\int |f|^{q \times \frac{p}{q}} d\mu \right)^{q/p} \times \left(\int 1^{p/(p-q)} d\mu \right)^{1-q/p}$$

en appliquant l'inégalité d'Hölder avec p/q et son conjugué $\frac{p}{p-q}$. D'où

$$\begin{aligned} \|f\|_q^q &\leq \mu(X)^{1-q/p} \left(\int |f|^p d\mu \right)^{q/p} \\ \|f\|_q &\leq \mu(X)^{1/q-1/p} \left(\int |f|^p d\mu \right)^{1/p} < +\infty. \end{aligned}$$

□

Plus précisément, le théorème montre une inclusion topologique car l'inclusion

$$f \in L^p(X, \mathcal{A}, \mu) \mapsto f \in L^q(X, \mathcal{A}, \mu)$$

est continue.

13.5 Parties denses

On a les résultats suivants – admis eux aussi – de densité dans les espaces $L^p(X, \mathcal{A}, \mu)$:

Théorème 13.5.1 Soit $S_p = \{\text{combinaisons linéaires finies de fonctions indicatrices d'ensembles } A_i\}$ avec en plus $\mu(A_i) < +\infty$ si $p < +\infty$. Alors S_p est dense dans $L^p(X, \mathcal{A}, \mu)$.

Démonstration : 1) Si $f \geq 0$, il existe une suite croissante de fonctions étagées s_n qui converge vers f . Comme $f \in L^p(X, \mathcal{A}, \mu)$ on a aussi $s_n \in L^p(X, \mathcal{A}, \mu)$ car

$$0 \leq \int s_n^p d\mu \leq \int f^p d\mu < +\infty.$$

• si $p = +\infty$, on a $\|f\|_\infty \leq M$, par construction de s_n , on a $|s_n(x) - f(x)| \leq 1/2^n$ donc $\|s_n - f\|_\infty \leq 1/2^n$, si bien que $s_n \rightarrow f$ dans L^∞ .

• Si $p < +\infty$, on a $|s_n - f|^p \leq 2^p |f|^p$ donc par convergence dominée

$$\|s_n - f\|_p^p = \int |f - s_n|^p d\mu \rightarrow 0.$$

2) Pour f quelconque dans L^p , on écrit

$$f = \operatorname{Re}(f)^+ - \operatorname{Re}(f)^- + i\operatorname{Im}(f)^+ - i\operatorname{Im}(f)^-$$

et on applique le 1) à chaque morceau positif par linéarité de l'intégrale. □

Théorème 13.5.2 Soit $1 \leq p < +\infty$, alors $C_c^0(\mathbb{R}^n) = \{\text{fonctions de } \mathbb{R}^n \rightarrow \mathbb{C} \text{ continues à support compact}\}$ est dense dans $L^p(\mathbb{R}^n, \lambda^n)$.

L'espace $C_c^0(\mathbb{R}^n)$ n'est pas complet pour la norme $\|\cdot\|_p$. Cependant on peut voir $L^p(\mathbb{R}^n, \lambda^n)$ comme son complété pour la norme $\|\cdot\|_p$.

Idée de la preuve : On approche les fonctions indicatrices par des fonctions de $C_c^0(\mathbb{R}^n)$ et on applique le résultat précédent. Pour cela, on dispose de

Lemme 13.5.1 (Uryshon) Soit (X, d) un espace métrique, V ouvert et $K \subset V$ compact. Alors il existe $f \in C_c^0(X)$ telle que

$$K \subset \text{supp}(f) \subset V \quad \text{et} \quad 0 \leq f \leq 1.$$

Démonstration : Soit

$$f(x) = \frac{d(x, V^c)}{d(x, V^c) + d(x, K)}.$$

La fonction f est continue car $x \mapsto d(x, F)$ est continue quand F est fermé. De plus $0 \leq f \leq 1$.

$$\text{Si } x \in K, f(x) = \frac{d(x, V^c)}{d(x, V^c) + 0} = 1 \text{ alors que si } x \in V^c, f(x) = \frac{0}{0 + d(x, K)} = 0.$$

On a donc $f|_K = 1$ et $f|_{V^c} = 0$ et $0 \leq f \leq 1$. \square

Ce résultat reste vrai si l'espace est seulement séparé localement compact, mais la justification est bien plus difficile que celle donnée ci-dessus pour un espace métrique.

On a en plus des propriétés (admisses) dites de régularité (intérieure et extérieure) de la mesure de Lebesgue :

Lemme 13.5.2 La mesure de Lebesgue vérifie

$$\begin{aligned} \lambda(A) &= \sup\{\lambda(K) \mid K \text{ compact } \subset A\} && (\text{régularité intérieure}) \\ &= \inf\{\lambda(U) \mid U \text{ ouvert } \supset A\} && (\text{régularité extérieure}) \end{aligned}$$

Démonstration du théorème 13.5.2 Comme on sait déjà que l'espace des fonctions indicatrices est dense pour $\|\cdot\|_p$ dans $L^p(X, \mathcal{A}, \mu)$ (cf. théorème 13.5.1), il suffit de montrer que toute fonction indicatrice peut-être approchée en norme $\|\cdot\|_p$ par une fonction de $C_c^0(\mathbb{R}^n)$.

On considère donc la fonction indicatrice $\mathbf{1}_A$ pour $A \in \mathcal{A}$ et on se fixe $\varepsilon > 0$ arbitraire. D'après le lemme de régularité de la mesure de Lebesgue, il existe K compact inclu dans A et U ouvert contenant A tels que $\lambda_n(U \setminus K) \leq \varepsilon^p$ où λ_n désigne la mesure de Lebesgue (en dimension n) sur \mathbb{R}^n .

Appliquons le lemme d'Uryshon dans $X = \mathbb{R}^n$ à $K \subset U$: il existe $f \in C_c^0(\mathbb{R}^n)$ telle que $0 \leq f \leq 1$ et

$$K \subset \text{supp}(f) \subset U.$$

Mais alors

$$\mathbf{1}_A - f = \begin{cases} 1 - 1 = 0 & \text{si } x \in K \\ 0 - 0 = 0 & \text{si } x \notin U \\ \in [0, 1] & \text{si } x \in U \setminus K. \end{cases}$$

On a donc

$$\|\mathbf{1}_A - f\|_p^p = \int_{U \setminus K} |\mathbf{1}_A - f|^p d\lambda_n \leq \int_{U \setminus K} d\lambda_n = \lambda_n(U \setminus K) \leq \varepsilon^p.$$

Comme $\varepsilon > 0$ est arbitraire, on a réussi à approcher en norme $\|\cdot\|_p$ toute indicatrice par une fonction de $C_c^0(\mathbb{R}^n)$. La fin s'obtient comme expliqué précédemment. \square

Remarque 13.5.1 • Le résultat est faux pour $L^\infty(\mathbb{R}^n)$: prendre $f = \mathbf{1}_{]-\infty, 0]}$ la fonction qui vaut 1 si $x \leq 0$ et 0 sinon. Il y a manifestement une discontinuité en 1 pour f et si on avait $\|g - f\|_\infty < 1/2$ pour une fonction g , on devrait avoir $g(0^+) > 1/2$ et $g(0^-) < 1/2$ ce qui empêchera g d'être continue en 0. Et donc $f \in L^\infty(\mathbb{R}^n)$ ne peut pas être approchée par des fonctions g continues.

• En fait, on a même (admis pour l'instant, cf. après) : l'ensemble des fonctions C^∞ à support compact est dense dans L^p pour $p < +\infty$.

13.6 Moments des variables aléatoires

Définition 13.6.1 Une v.a. $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ a un moment d'ordre $p \geq 1$ ssi

$$E[|X|^p] = \int_{\Omega} |X|^p d\mathbb{P} < +\infty.$$

Définition 13.6.2 $L^p(\Omega, \mathcal{F}, \mathbb{P}) = \{X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R} \mid E[|X|^p] < +\infty\}$.

Les résultats généraux sur les espaces $L^p(X, \mathcal{A}, \mu)$ se reformulent dans le cadre probabiliste de la façon suivante.

Proposition 13.6.1 On dispose de

- Inégalité de Hölder : $\|XY\|_1 \leq \|X\|_p \|Y\|_q$ pour p, q exposant conjugués.
- Inégalité de Cauchy-Schwarz : $\|XY\|_1 \leq \|X\|_2 \|Y\|_2$ ($p = q = 2$).
- Inégalité de Minkowski : $\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$ ($1 \leq p < +\infty$).
- Si une v.a. est bornée, elle admet des moments de tous les ordres.
- Si X possède un moment d'ordre r , pour tout $n \leq r$, X en possède un d'ordre n .
- $(L^p(\Omega, \mathcal{F}, \mathbb{P}), \|\cdot\|_p)$ est un e.v.n. complet, c'est à dire un espace de Banach.

Définition 13.6.3 (Variance) Si $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, on définit la variance de X par

$$\text{Var}(X) = E[(X - E[X])^2]. \quad (13.3)$$

On définit aussi l'écart-type $\sigma_X = \sqrt{\text{Var}(X)}$.

Remarque 13.6.1 La définition de la variance est unifiée entre les deux principaux cas (discret et à densité) grâce à cette nouvelle théorie de la mesure.

• Si X est discrète de domaine $X(\Omega) = \{x_1, \dots, x_n, \dots\}$, la loi de X est $\mathbb{P}_X = \sum_{i=1}^{+\infty} \mathbb{P}(X = x_i) \delta_{x_i}$ et la variance en (13.3) devient

$$\text{Var}(X) = \sum_{i=1}^{+\infty} (x_i - E[X])^2 \mathbb{P}_X\{x_i\} = \sum_{i=1}^{+\infty} (x_i - E[X])^2 \mathbb{P}(X = x_i).$$

• Si X est une v.a. de densité f alors la loi de X est la mesure de densité f , $d\mathbb{P}_X = f(x)dx$ et la variance en (13.3) devient

$$\text{Var}(X) = \int_{\mathbb{R}} (x - E[X])^2 f(x) dx.$$

Proposition 13.6.2 (Propriétés de la variance)

- $\text{Var}(X) \geq 0$.
- $\text{Var}(X) = E[X^2] - (E[X])^2$ (Formule de Koenig).
- $\text{Var}(\lambda X) = \lambda^2 \text{Var}(X)$.
- $\text{Var}(X + a) = \text{Var}(X)$ pour toute constante $a \in \mathbb{R}$.
- $\text{Var}(X) = 0$ ssi X est constante p.p. (et vaut alors $E[X]$).

Démonstration : • Le premier point est clair (par définition, c'est une somme de carrés).

- On développe $\text{Var}(X)$, en notant $\mu = E[X]$:

$$\begin{aligned} \text{Var}(X) = E[(X - \mu)^2] &= E[X^2 - 2X\mu + \mu^2] \\ &= E[X^2] - 2E[X\mu] + \mu^2 \\ &= E[X^2] - 2E[X]\mu + \mu^2 \\ &= E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - E[X]^2. \end{aligned}$$

- Pour les troisième et quatrième points :

$$\text{Var}(aX) = E[aX - E[aX]]^2 = E[a(X - E[X])]^2 = a^2 E[(X - E[X])^2] = a^2 \text{Var}(X)$$

$$\text{Var}(X + b) = E[(X + b - E(X + b))]^2 = E[(X + b - E[X] - b)^2] = E[(X - E[X])^2] = \text{Var}(X).$$

• Si $X = c$ une constante p.p. alors $E[X] = E[c] = c$ et $E[X^2] = E[c^2] = c^2$ si bien que $\text{Var}(X) = c^2 - c^2 = 0$. Réciproquement, si $\text{Var}(X) = E[(X - E[X])^2] = 0$ alors la variable $(X - E[X])^2$, positive d'espérance nulle, est elle même nulle p.p., c'est à dire $X = E[X]$ p.p. □

Attention, en général, on n'a pas $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$. Prendre par exemple $X = Y$. Par contre

Théorème 13.6.1 Si X et Y sont deux v.a. indépendantes avec des moments d'ordre deux alors

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

Démonstration :

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

car par indépendance de X et Y , on a : $E[XY] = E[X] \times E[Y]$. □

Théorème 13.6.2 (Inégalité de Tchebychev) Si $\text{Var}(X)$ existe, on a

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

Démonstration : Par l'inégalité de Markov, on a

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

□

Application : On jette 3600 fois un dé. Minorer la probabilité que le nombre d'apparitions du 1 soit compris strictement entre 480 et 720.

Notons S le nombre d'apparitions du 1. On peut voir S comme la somme de 3600 v.a. de Bernoulli indépendantes de paramètre $p = 1/6$ (probabilité d'apparition du 1 au cours d'un lancer). Par un raisonnement maintenant classique, S suit une loi $\mathcal{B}(3600, p)$. On cherche ici

$$\mathbb{P}(480 < S < 720) = \sum_{k=481}^{719} C_n^k p^k (1-p)^{n-k}.$$

Ce résultat exact ne peut être calculé en pratique, même un ordinateur très puissant ne pouvant calculer tous ces coefficients binomiaux pour des chiffres aussi grands.

On peut penser à approximer la loi $\mathcal{B}(3600, 1/6)$ par $\mathcal{P}(600)$ mais il resterait à calculer

$$\sum_{k=481}^{719} e^{-600} \frac{600^k}{k!},$$

ce qui n'est pas évident.

On a alors recours à l'inégalité de Tchebychev : notons que $E(S) = np = 3600/6 = 600$ et $\text{Var}(X) = npq = 3600 \times 5/6 \times 1/6 = 500$. Remarquons de plus que

$$480 < S < 720 \iff -120 < S - 600 < 120.$$

D'où

$$\begin{aligned} \mathbb{P}(480 < S < 720) &= \mathbb{P}(-120 < S - 600 < 120) = \mathbb{P}(|S - 600| < 120) \\ &= 1 - \mathbb{P}(|S - 600| \geq 120) \\ &\geq 1 - \frac{500}{120^2} \\ &\geq 0,95833\dots \end{aligned}$$

Remarque 13.6.2 Les valeurs 480 et 720 sont symétriques par rapport à la moyenne 600 de la v.a. considérée, ce sont 600 ± 120 . Ce n'est pas nécessaire : on peut aussi appliquer l'inégalité de Tchebychev sur un intervalle non centré autour de l'espérance. Il suffit pour cela d'utiliser le plus grand intervalle centré sur l'espérance qu'il contient. Ainsi pour minorer $\mathbb{P}(550 < S < 700)$, il suffit de remarquer que

$$550 < S < 700 \iff \underbrace{550 < S < 650}_{\text{intervalle centré autour de 600}} \iff -50 < S - 600 < 50.$$

et

$$\begin{aligned} \mathbb{P}(550 < S < 700) &\geq \mathbb{P}(550 < S < 650) = \mathbb{P}(-50 < S - 600 < 50) \\ &= \mathbb{P}(|S - 600| < 50) \\ &= 1 - \mathbb{P}(|S - 600| \geq 50) \\ &\geq 1 - \frac{500}{50^2} = 0,2. \end{aligned}$$

Tableau comparatif des formules pour des v.a. discrètes et continues à densité

Lorsque les intégrales et les séries concernées sont absolument convergentes, on a le tableau comparatif suivant entre le cas général et les déclinaisons discrètes et à densité (cas continu) :

X	Cas général	Variable discrète	Variabl
$X(\Omega)$	quelconque	$\{x_1, x_2, \dots, x_k, \dots\}$	\mathbb{R} ou
$\mathbb{P}(a \leq X \leq b)$	$\int_{\Omega} \mathbf{1}_{[a,b]}(X(\omega)) d\mathbb{P}(\omega)$	$\sum_{a \leq x_k \leq b} \mathbb{P}(X = x_k)$	\int_a^b
$F(x) = \mathbb{P}(X \leq x)$	$\int_{\Omega} X(\omega) d\mathbb{P}(\omega)$	$\sum_{x_k \leq x} \mathbb{P}(X = x_k)$	$\int_{-\infty}^x$
$E[X]$	$\int \mathbf{1}_{]-\infty, x]}(X(\omega)) d\mathbb{P}(\omega)$	$\sum_{k=1}^{+\infty} x_k \mathbb{P}(X = x_k)$	$\int_{-\infty}^{+\infty}$
$Eg(X)$	$\int_{\Omega} g(X(\omega)) d\mathbb{P}(\omega)$	$\sum_{k=1}^{+\infty} g(x_k) \mathbb{P}(X = x_k)$	$\int_{-\infty}^{+\infty}$
$E(X^2)$	$\int_{\Omega} X(\omega)^2 d\mathbb{P}(\omega)$	$\sum_{k=1}^{+\infty} x_k^2 \mathbb{P}(X = x_k)$	$\int_{-\infty}^{+\infty}$
$\text{Var}(X)$	$\int_{\Omega} (X(\omega) - E[X])^2 d\mathbb{P}(\omega)$	$\sum_{k=1}^{+\infty} (x_k - E[X])^2 \mathbb{P}(X = x_k)$	$\int_{-\infty}^{+\infty} (t - E[X])^2$

Chapitre 14

Convolution

14.1 Définition et propriétés

On considère dans ce chapitre $(X, \mathcal{A}, \mathbb{P}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), \lambda^n)$.

Définition 14.1.1 *La convolution de deux fonctions f et g réelles est la fonction $f * g$ définie sur \mathbb{R}^n et donnée par*

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y) dy.$$

*On parle encore de la convolée $f * g$ de f et de g .*

Il n'est pas clair pour quels x la fonction $f * g$ est bien définie. Certains résultats suivent pour donner des conditions d'existence de $(f * g)(x)$.

Proposition 14.1.1 *Soient f, g des fonctions mesurables, alors*

- $(f * g)(x) = (g * f)(x)$
- $(f, g) \mapsto f * g$ est bilinéaire.

Démonstration : • Pour le premier point, le changement de variable $z = x - y$ donne

$$(f * g)(x) = \int_{\mathbb{R}^n} f(x - y)g(y)dy = \int_{\mathbb{R}^n} f(z)g(x - z)|(-1)^n|dz = (g * f)(x).$$

- Pour le second point, par linéarité de l'intégrale, il est facile de voir que

$$(a_1f_1 + a_2f_2) * g(x) = a_1(f_1 * g)(x) + a_2(f_2 * g)(x).$$

La linéarité par rapport à g se montre de la même façon ou en utilisant celle par rapport à f avec la symétrie de $*$ vue au premier point. \square

Proposition 14.1.2 *Si f est nulle hors de A et g nulle hors de B , alors $(f * g)(x)$ existe et vaut 0 en dehors de $A + B$.*

Démonstration : $(f * g)(x)$ existe si l'intégrale

$$\int_{\mathbb{R}^n} |f(y)| |g(x - y)| dy$$

converge. Or si $x = y + (x - y) \notin A + B$, alors soit $y \notin A$ et $f(y) = 0$, soit $y \in A$ mais alors $x - y \notin B$ et $g(x - y) = 0$. Donc dans tous les cas pour $x \notin A + B$, on a $f(y)g(x - y) = 0$ pour tout y et en intégrant :

$$\int_{\mathbb{R}^n} |f(y)| |g(x - y)| dy = 0$$

si bien que $(f * g)(x)$ est bien définie et vaut forcément 0 quand $x \notin A + B$. \square

Proposition 14.1.3 *Soient f et g bornées et nulles respectivement hors de A et de B compacts alors $(f * g)(x)$ existe pour tout x et $(f * g)(x) = 0$ pour $x \notin A + B$.*

Démonstration : Pour $x \notin A + B$, c'est le résultat précédent qui s'applique. Puis pour $x \in A + B$, on a

$$\int_{\mathbb{R}^n} |f(y)| |g(x - y)| dy = \int_A |f(y)| |g(x - y)| dy < \lambda(A) \|f\|_{\infty} \|g\|_{\infty} < +\infty.$$

On en déduit que $y \mapsto f(y)g(x - y)$ est intégrable pour tout x . La fonction $f * g$ est donc bien définie sur tout \mathbb{R}^n . \square

14.1.1 Normes des convolutions

Les résultats suivants donnent des inégalités de type Hölder pour les convolutions.

Proposition 14.1.4 *Soient $f, g \in L^1(\mathbb{R}^n)$, alors $f * g$ est définie p.p. et*

$$\|f * g\|_1 \leq \|f\|_1 \|g\|_1.$$

Démonstration : En appliquant le théorème de Fubini-Tonelli, on a

$$\begin{aligned} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y)| |g(x - y)| dy dx &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y)| |g(x - y)| dx dy \\ &= \int_{\mathbb{R}^n} |f(y)| \left(\int_{\mathbb{R}^n} |g(x - y)| dx \right) dy \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}^n} |f(y)| \left(\int_{\mathbb{R}^n} |g(z)| dz \right) dy \\
&= \int_{\mathbb{R}^n} |f(y)| dy \int_{\mathbb{R}^n} |g(z)| dz \\
&= \|f\|_1 \|g\|_1.
\end{aligned}$$

Par conséquent, $I(x) = \int_{\mathbb{R}^n} |f(y)| |g(x-y)| dy$ est d'intégrale finie bornée par $\|f\|_1 \|g\|_1$, c'est donc que $I(x) = +\infty$ sur un ensemble de mesure nulle et $f * g$ est bien définie p.p. Puis

$$\begin{aligned}
\|f * g\|_1 &= \int_{\mathbb{R}^n} |(f * g)(x)| dx = \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(y) g(x-y) dy \right| dx \\
&\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(y) g(x-y)| dy dx \\
&\leq \|f\|_1 \|g\|_1.
\end{aligned}$$

□

Ainsi $*$ est une loi de composition interne de $L^1(\mathbb{R}^n)$. On peut vérifier qu'elle est associative. $(L^1(\mathbb{R}^n), +, \cdot, *)$ est alors une algèbre. Comme elle est complète pour $\|\cdot\|_1$, il s'agit même d'une algèbre de Banach.

On a aussi, le résultat suivant :

Corollaire 14.1.1 Soit $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$ avec $r \neq +\infty$ alors pour $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$, $f * g$ existe p.p. et $f * g \in L^r(\mathbb{R}^n)$, avec

$$\|f * g\|_r \leq \|f\|_p \|g\|_q.$$

Si $q = 1$, on constate que la convolution par une fonction de $g \in L^1(\mathbb{R})$ est un opérateur de $L^p(\mathbb{R})$ dans $L^p(\mathbb{R})$:

$$f \in L^p(\mathbb{R}) \mapsto f * g \in L^p(\mathbb{R}).$$

Démonstration : On a $\frac{1}{r} + \left(\frac{1}{p} - \frac{1}{r}\right) + \left(\frac{1}{q} - \frac{1}{r}\right) = 1$. Posons

$$\alpha = r, \quad \beta = \frac{rp}{r-p}, \quad \gamma = \frac{rq}{r-q}.$$

On a alors $\frac{1}{\alpha} + \frac{1}{\beta} + \frac{1}{\gamma} = 1$. Puis une inégalité de Hölder généralisée avec trois termes donnent :

$$\begin{aligned}
&\int_{\mathbb{R}^n} |f(y)| |g(x-y)| dy \\
&= \int_{\mathbb{R}^n} |f(y)|^{p/r} |g(x-y)|^{q/r} |f(y)|^{1-p/r} |g(y)|^{1-q/r} dy
\end{aligned}$$

$$\leq \left(\int_{\mathbb{R}^n} |f(y)|^p |g(x-y)|^q dy \right)^{1/r} \left(\int_{\mathbb{R}^n} |f(y)|^p dy \right)^{(r-p)/(rp)} \left(\int_{\mathbb{R}^n} |g(x-y)|^q dy \right)^{(r-q)/(rq)}. \quad (14.1)$$

Comme $f \in L^p$ et $g \in L^q$, le premier facteur est fini par le théorème précédent, le deuxième facteur est $\|f\|_p^{(r-p)/r} < +\infty$ et le dernier (après un changement de variable $z = x - y$) est $\|g\|_q^{(r-q)/r} < +\infty$. Finalement, on a

$$\int_{\mathbb{R}^n} |f(y)| |g(x-y)| dy < +\infty$$

et la convolée $f * g$ existe p.p. Puis

$$\begin{aligned} \int_{\mathbb{R}^n} |(f * g)(x)|^r dx &= \int_{\mathbb{R}^n} \left| \int_{\mathbb{R}^n} f(y)g(x-y) dy \right|^r dx \\ &\leq \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(y)g(x-y)| dy \right)^r dx \\ &\leq \left(\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |f(y)|^p |g(x-y)|^q dy \right) dx \right) \|f\|_p^{r-p} \|g\|_q^{r-q} \\ &\leq \| |f|^p * |g|^q \|_1 \|f\|_p^{r-p} \|g\|_q^{r-q} \\ &\leq \|f\|_p^p \|g\|_q^q \|f\|_p^{r-p} \|g\|_q^{r-q} \\ &\leq \|f\|_p^r \|g\|_q^r \end{aligned}$$

en utilisant (14.1) à la troisième ligne. D'où $\|f * g\|_r \leq \|f\|_p^p \|g\|_q$. \square

Théorème 14.1.1 Soit $f \in L^p(\mathbb{R})$ alors pour $h \in \mathbb{R}$, $\tau_h f(x) = f(x+h)$ converge vers f dans $L^p(\mathbb{R})$ quand $h \rightarrow 0$.

Démonstration : • Pour $p < \infty$, les fonctions continues à support compact sont denses dans L^p (cf. théorème 13.5.2). On commence donc par considérer φ une telle fonction, de support dans $B(0, R)$. Pour $\|h\| \leq 1/2$, on a

$$\begin{aligned} \int_{\mathbb{R}^n} |\tau_h \varphi(x) - \varphi(x)| dx &= \int_{\mathbb{R}^n} |\varphi(x+h) - \varphi(x)| dx \\ &\leq \int_{\|x\| \leq R+1/2} |\varphi(x+h) - \varphi(x)| dx \end{aligned}$$

car φ est nulle hors de $B(0, R)$ et $\varphi(\cdot + h)$ l'est hors de $B(0, R + \frac{1}{2})$ quand $|h| < \frac{1}{2}$. Comme φ est uniformément continue sur le compact $\bar{B}(0, R + 1/2)$ (th. de Heine), pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que si $\|h\| \leq \alpha$ alors pour tout $x \in B(0, R + 1/2)$ on a

$$|\varphi(x+h) - \varphi(x)| \leq \varepsilon.$$

D'où

$$\begin{aligned} \int |\tau_h \varphi(x) - \varphi(x)|^p dx &\leq \varepsilon^p \text{vol}(B(0, R + 1/2)) \\ \|\tau_h \varphi - \varphi\|_p &\leq \varepsilon (\text{vol}(B(0, R + 1/2)))^{1/p}. \end{aligned}$$

C'est à dire $\lim_{h \rightarrow 0} \|\tau_h \varphi - \varphi\|_p = 0$.

• Pour $f \in L^p$ et $\varepsilon > 0$, par densité des fonctions continues à support compact, il existe une telle fonction φ telle que $\|f - \varphi\|_p \leq \varepsilon/3$. Puis

$$\begin{aligned} \|\tau_h f - f\|_p &\leq \|\tau_h f - \tau_h \varphi\|_p + \|\tau_h \varphi - \varphi\|_p + \|\varphi - f\|_p \\ &= 2\|f - \varphi\|_p + \|\tau_h \varphi - \varphi\|_p \\ &\leq 2\varepsilon/3 + \|\tau_h \varphi - \varphi\|_p. \end{aligned}$$

D'après le premier cas, $\|\tau_h \varphi - \varphi\|_p$ tend vers 0 quand $h \rightarrow 0$, donc pour h assez petit, $\|\tau_h \varphi - \varphi\|_p \leq \varepsilon/3$ si bien que

$$\|\tau_h f - f\|_p \leq 2\varepsilon/3 + \varepsilon/3 = \varepsilon.$$

Le résultat suit car $\varepsilon > 0$ est arbitraire. □

Si p et q sont conjugués, on a le résultat de convolution suivant :

Théorème 14.1.2 Soit $\frac{1}{p} + \frac{1}{q} = 1$ et $f \in L^p(\mathbb{R}^n)$, $g \in L^q(\mathbb{R}^n)$ alors

- $(f * g)(x)$ existe pour presque tout x et $|(f * g)(x)| \leq \|f\|_p \|g\|_q$;
- $f * g$ est continue et même uniformément continue ;
- Pour $p, q \neq +\infty$, on a de plus $(f * g)(x) \rightarrow 0$ quand $\|x\| \rightarrow +\infty$.

Démonstration :

- Par l'inégalité de Hölder, on a :

$$\int_{\mathbb{R}^n} |f(y)| |g(x-y)| dy \leq \left(\int_{\mathbb{R}^n} |f(y)|^p dy \right)^{1/p} \left(\int_{\mathbb{R}^n} |g(x-y)|^q dy \right)^{1/q} \leq \|f\|_p \|g\|_q.$$

D'où $|(f * g)(x)| \leq \|f\|_p \|g\|_q < +\infty$ existe pour tout x .

•

$$\begin{aligned} |(f * g)(x+h) - (f * g)(x)| &\leq \int_{\mathbb{R}^n} |f(y)| |g(x+h-y) - g(x-y)| dy \\ &\leq \|f\|_p \left(\int_{\mathbb{R}^n} |g(x+h-y) - g(x-y)|^q dy \right)^{1/q} \\ &\leq \|f\|_p \|\tau_h g - g\|_q. \end{aligned}$$

Quitte à échanger le rôle de f et g , on peut supposer $q \neq +\infty$. Or la proposition précédente donne : pour tout $\varepsilon > 0$, il existe $\alpha > 0$ tel que pour $|h| \leq \alpha$, on a $\|\tau_h g - g\|_q \leq \varepsilon$, c'est-à-dire

$$|(f * g)(x + h) - (f * g)(x)| \leq \varepsilon \|f\|_p.$$

On a donc pour tout x et $|h|$ assez petit $|(f * g)(x + h) - (f * g)(x)|$ petit, ce qui donne la continuité uniforme de $f * g$.

• Soient φ et ϕ des fonctions continues à support compact qui approchent $f \in L^p$ et $g \in L^q$ en norme $\|\cdot\|_p$, $\|\cdot\|_q$ respectivement. On approche alors $f * g$ par $\varphi * \phi$ de la façon suivante :

$$\begin{aligned} |(f * g)(x) - (\varphi * \phi)(x)| &\leq |(f * g)(x) - (\varphi * g)(x)| + |(\varphi * g)(x) - (\varphi * \phi)(x)| \\ &\leq \|g\|_q \|f - \varphi\|_p + \|\varphi\|_p \|g - \phi\|_q \end{aligned}$$

en majorant grâce au premier point. On a alors

$$\begin{aligned} |(f * g)(x)| &\leq |(f * g)(x) - (\varphi * \phi)(x)| + |(\varphi * \phi)(x)| \\ &\leq \|g\|_q \|f - \varphi\|_p + \|\varphi\|_p \|g - \phi\|_q + |(\varphi * \phi)(x)| \end{aligned}$$

Choisissons φ telle que

$$\|f - \varphi\|_p \leq \frac{\varepsilon}{2\|g\|_q},$$

puis ϕ telle que

$$\|g - \phi\|_q \leq \frac{\varepsilon}{2\|\varphi\|_p}$$

ce qui est possible par densité des fonctions continues à support compact dans L^p et L^q ($p, q < +\infty$). Mais alors comme $\varphi * \phi$ est nulle hors d'un compact (car φ et ϕ sont elles-mêmes nulles hors d'un compact, cf. Prop. 14.1.3), on en déduit que $|(f * g)(x)| < \varepsilon$ pour $\|x\|$ assez grand. Autrement dit

$$\lim_{\|x\| \rightarrow +\infty} (f * g)(x) = 0.$$

□

14.2 Dérivation des convolutions

Théorème 14.2.1 • Si $f \in L^1(\mathbb{R}^n)$ et g est C^1 bornée de dérivées partielles $\partial g / \partial x_i$ bornées, alors $f * g$ est C^1 de dérivées

$$\frac{\partial}{\partial x_i} (f * g) = f * \frac{\partial g}{\partial x_i}.$$

• Si $f \in L^p(\mathbb{R}^n)$ et g est C^1 à support compact, la même conclusion est vraie.

Démonstration :

1) La fonction $f * g$ existe et est continue d'après le théorème précédent car on convole $L^1 * L^\infty$. Puis, pour tout y , $x \mapsto f(y)g(x - y)$ est dérivable par rapport à x_i de dérivée

$$f(y) \frac{\partial g}{\partial x_i}(x - y).$$

Comme $|f(y) \frac{\partial g}{\partial x_i}(x - y)| \leq M|f(y)|$, le théorème de dérivation sous l'intégrale s'applique et donne la dérivabilité de $f * g$ par rapport à x_i avec la dérivée

$$f * \frac{\partial g}{\partial x_i}(x) = \int f(y) \frac{\partial g}{\partial x_i}(x - y) dy,$$

convolée à nouveau de $L^1 * L^\infty$ donc continue. Finalement, $f * g$ est de classe C^1 .

2) Comme g est à support compact dans $B(0, R)$, $g \in L^q$ (q conjugué de p). Puis pour $\|x\| \leq K$:

$$|(f * g)(x)| \leq \int_{\mathbb{R}^n} |f(y)| |g(x - y)| dy = \int_{\|y\| \leq R-K} |f(y)| |g(x - y)| dy$$

existe car on convole L^p et L^q , de plus $f * g$ est continue. Puis pour tout y , $x \mapsto f(y)g(x - y)$ est dérivable par rapport à x_i de dérivée $f(y) \frac{\partial g}{\partial x_i}(x - y)$ dominée par $\|f(y)\| \left\| \frac{\partial g}{\partial x_i} \right\|_\infty \mathbf{1}_{\{\|y\| \leq R-K\}}$ avec par Hölder

$$\begin{aligned} \int_{\mathbb{R}^n} |f(y)| \left\| \frac{\partial g}{\partial x_i} \right\|_\infty \mathbf{1}_{\{\|y\| \leq R+K\}} dy &\leq \|f\|_p \left(\int \left(\left\| \frac{\partial g}{\partial x_i} \right\|_\infty \mathbf{1}_{\{\|y\| \leq R-K\}} \right)^p dy \right)^{1/q} \\ &\leq \|f\|_p \left\| \frac{\partial g}{\partial x_i} \right\|_\infty^{p/q} \lambda(\|y\| \leq R - K)^{1/q} < +\infty. \end{aligned}$$

On peut appliquer le théorème de dérivation sous \int : on en déduit que $f * g$ est dérivable par rapport à x_i de dérivée $f * \frac{\partial g}{\partial x_i}$, puis on en déduit aussi la continuité. Finalement, $f * g$ est de classe C^1 . \square

Corollaire 14.2.1 Soit $f \in L^p$ et g de classe C^∞ à support compact alors $f * g$ est C^∞ avec $\partial^\alpha(f * g) = f * \partial^\alpha g$ pour un multi-indice $\alpha \in \{1, \dots, n\}^p$.

Démonstration : Immédiat par utilisation successive du théorème 14.2.1. \square

Définition 14.2.1 Une fonction $g : \mathbb{R}$ ou $\mathbb{C} \rightarrow \mathbb{C}$ est entière si elle s'écrit $g(x) = \sum_{n \geq 0} c_n x^n$ avec un rayon de convergence infini.

Pour les convolées de fonctions entières, on a la version suivante du résultat de dérivation :

Corollaire 14.2.2 Si $f \in L^1(\mathbb{C})$ (ou $L^1(\mathbb{R})$) est à support compact et g est entière. Alors $f * g$ est entière.

Démonstration : Admis.

14.3 Approximation

Un des inconvénients majeurs de l'opération convolution est de ne pas avoir d'unité, c'est à dire de fonction E telle que toute fonction $f : f * E = E * f = f$. Cependant, pour compenser ce défaut, on introduit la notion d'approximation de l'unité :

Définition 14.3.1 Soit $e_k : \mathbb{R}^n \rightarrow \mathbb{R}^+$ mesurable (continue en pratique) vérifiant

$$\int_{\mathbb{R}^n} e_k(x) dx = 1 \quad \text{et pour tout } \eta > 0, \lim_{k \rightarrow +\infty} \int_{|x| \geq \eta} e_k(x) dx = 0.$$

La suite $(e_k)_k$ est appelée approximation de l'unité.

Exemples : • Soit

$$e(x) = \frac{e^{-x^2}}{\sqrt{\pi}}. \quad (14.2)$$

Il s'agit d'une fonction C^∞ entière d'intégrale 1. On vérifie facilement qu'on définit alors une approximation de l'unité en prenant $e_k(x) = ke(kx)$. En effet

- d'abord e_k est mesurable positive,
- puis $\int_{\mathbb{R}^n} e_k(x) dx = \int_{\mathbb{R}^n} e(y) dy = 1$ avec le changement de variable $y = kx$.
- $\int_{\{\|x\| > \eta\}} e_k(x) dx = \int_{\{\|y\| > k\eta\}} e(y) dy \rightarrow 0$ quand $k \rightarrow +\infty$ par convergence dominée.

• Soit

$$\varphi(x) = \begin{cases} \mathbb{R} \rightarrow \mathbb{R} \\ x \mapsto e^{-1/(1-x^2)} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

On montre que φ est C^∞ et donc à support compact, ce qui permet de définir sur \mathbb{R}^n

$$e(x) = \frac{\varphi(\|x\|^2)}{\int_{\mathbb{R}^n} \varphi(\|y\|^2) dy} \quad (14.3)$$

où $\|x\|^2 = x_1^2 + \dots + x_n^2$. Alors e est C^∞ , à support dans $B(0, 1)$ et $\int_{\mathbb{R}^n} e(x) dx = 1$. On définit alors une approximation de l'unité en prenant $e_k(x) = ke(kx)$.

La terminologie « approximation de l'unité » est justifiée par les deux résultats suivants :

Théorème 14.3.1 Soit $f \in L^\infty(\mathbb{R}^n)$ et $(e_k)_k$ une approximation de l'unité, on suppose que

- soit f est continue sur $\{\|x\| \leq R\}$,
- soit f est uniformément continue sur \mathbb{R}^n .

Alors $f * e_k$ converge uniformément vers f .

Démonstration : Notons que comme $\int_{\mathbb{R}^n} e_k(y)dy = 1$, on peut écrire $f(x) = \int_{\mathbb{R}^n} f(x)e_k(y)dy$. Puis par continuité (uniforme) de f : pour $\varepsilon > 0$, il existe $\eta > 0$ tel que pour $\|y\| \leq \eta$, on a $|f(x-y) - f(x)| \leq \varepsilon$, si bien que

$$\begin{aligned} |f * e_k(x) - f(x)| &\leq \left| \int_{\mathbb{R}^n} f(x-y)e_k(y)dy - \int_{\mathbb{R}^n} f(x)e_k(y)dy \right| \\ &= \int_{\mathbb{R}^n} |f(x-y) - f(x)|e_k(y)dy \\ &\leq \int_{\|y\| < \eta} |f(x-y) - f(x)|e_k(y)dy + \int_{\|y\| \geq \eta} |f(x-y) - f(x)|e_k(y)dy \\ &\leq \varepsilon \int_{\|y\| < \eta} e_k(y)dy + 2\|f\|_\infty \int_{\|y\| \geq \eta} e_k(y)dy \\ &\leq \varepsilon + 2\|f\|_\infty \int_{\|y\| \geq \eta} e_k(y)dy \\ &\leq 2\varepsilon \end{aligned}$$

pour k assez grand car $\lim_{k \rightarrow +\infty} \int_{\|y\| \geq \eta} e_k(y)dy \rightarrow 0$ donc $\int_{\|y\| \geq \eta} e_k(y)dy \leq \varepsilon/(2\|f\|_\infty)$ pour k assez grand. \square

On a aussi

Théorème 14.3.2 Soit $f \in L^p(\mathbb{R}^n)$ pour $p < +\infty$ et $(e_k)_k$ une approximation de l'unité alors $f * e_k \rightarrow f$ dans $L^p(\mathbb{R}^n)$.

Démonstration :

Comme $e_k \in L^1(\mathbb{R}^n)$, $f \in L^p(\mathbb{R}^n)$, on a $f * e_k \in L^p(\mathbb{R}^n)$. Puis

$$\begin{aligned} |f * e_k(x) - f(x)| &\leq \int_{\mathbb{R}^n} |f(x-y) - f(x)|e_k(y)^{1/p}e_k(y)^{1-1/p}dy \\ &\leq \left(\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p e_k(y)dy \right)^{1/p} \left(\int_{\mathbb{R}^n} e_k(y)dy \right)^{1/q} \end{aligned}$$

par l'inégalité de Hölder, puis comme $\int e_k(y)dy = 1$, on a

$$\begin{aligned} \int_{\mathbb{R}^n} |f * e_k(x) - f(x)|^p dx &\leq \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p e_k(y)dydx \\ &\leq \int_{\mathbb{R}^n} e_k(y) \left(\int_{\mathbb{R}^n} |f(x-y) - f(x)|^p dx \right) dy \\ \|f * e_k - f\|_p^p &\leq \int_{\|y\| < \eta} \dots dy + \int_{\|y\| \geq \eta} \dots dy. \end{aligned}$$

Par le théorème 14.1.1, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour $\|y\| < \eta$, $\|\tau_y f - f\|_p^p = \int_{\mathbb{R}^n} |f(x-y) - f(x)|^p dx \leq \varepsilon^p$. D'où,

$$\begin{aligned} \|f * e_k - f\|_p^p &\leq \varepsilon^p \int_{\|y\| < \eta} e_k(y) dy + (2\|f\|_p)^p \int_{\|y\| \geq \eta} e_k(y) dy \\ &\leq \varepsilon^p + \varepsilon^p = 2\varepsilon^p \end{aligned}$$

dès que k est assez grand car $\int_{|y| \geq \eta} e_k(y) dy \rightarrow 0$, ce qui conclut la preuve. \square

Applications

- Avec l'approximation de l'unité $e_k(x) = ke(kx)$ où e est donnée par (14.2), on a :

Théorème 14.3.3 (Stone-Weierstrass) *Soit $f : K \rightarrow \mathbb{R}$ continue sur K compact. Alors f peut être uniformément approchée par des polynômes.*

Démonstration : On commence par prolonger f en une fonction \tilde{f} définie sur \mathbb{R} et à support compact. On obtient ainsi une fonction \tilde{f} uniformément continue. Les fonctions $\tilde{f} * e_k$ sont entières car les e_k le sont puis elles convergent uniformément vers \tilde{f} sur \mathbb{R} donc convergent sur K vers f . Or toute fonction entière est limite uniforme sur un compact de ses sommes partielles qui sont des polynômes. Finalement, f est limite uniforme de polynômes sur le compact K . \square

- Avec l'approximation de l'unité $e_k(x) = ke(kx)$ où e est donnée par (14.3), on a :

Théorème 14.3.4 *Soit $p < +\infty$, $\mathcal{D}(\mathbb{R}^n) = \{\text{fonctions } C^\infty \text{ à support compact}\}$ est dense dans $L^p(\mathbb{R}^n)$.*

Démonstration : Troncature et convolution (les deux mamelles de la régularisation)

- Troncature : soit $f \in L^p$ alors par convergence dominée $f\mathbf{1}_{B(0,m)} \rightarrow f$ dans L^p quand $m \rightarrow +\infty$:

$$\int |f\mathbf{1}_{B(0,m)} - f|^p dx \rightarrow 0.$$

Pour $\varepsilon > 0$ fixé, soit m tel que $\|f\mathbf{1}_{B(0,m)} - f\|_p \leq \varepsilon/2$.

- La fonction $(f\mathbf{1}_{B(0,m)}) * \varphi_k$ est C^∞ et à support compact donc $(f\mathbf{1}_{B(0,m)}) * \varphi_k \rightarrow f\mathbf{1}_{B(0,m)}$ dans L^p , on trouve donc k tel que $\|(f\mathbf{1}_{B(0,m)}) * \varphi_k - f\mathbf{1}_{B(0,m)}\|_p \leq \varepsilon/2$. Finalement,

$$\begin{aligned} \|f - (f\mathbf{1}_{B(0,m)}) * \varphi_k\|_p &\leq \|f\mathbf{1}_{B(0,m)} - f\|_p + \|(f\mathbf{1}_{B(0,m)}) * \varphi_k - f\mathbf{1}_{B(0,m)}\|_p \\ &\leq \varepsilon/2 + \varepsilon/2 = \varepsilon. \end{aligned}$$

Dans tout voisinage de $f \in L^p$ pour la norme $\|\cdot\|_p$, il y a une fonction $(f\mathbf{1}_{B(0,m)}) * \varphi_k$ de classe C^∞ et à support compact, ce qui prouve le résultat cherché de densité. \square

•

Théorème 14.3.5 (Fejer) Soit f une fonction 2π -périodique, continue de \mathbb{R} dans \mathbb{C} et

$$\tilde{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt, \quad s_n(x) = \sum_{k=-n}^n \tilde{f}(k) e^{ikx}, \quad \sigma_N(x) = \frac{1}{N+1} \sum_{n=0}^N s_n(x).$$

Alors σ_N converge uniformément vers f quand $N \rightarrow \infty$.

On écrit donc f comme limite uniforme de polynômes trigonométriques.

Démonstration : Admise.

14.4 Sommes de variables aléatoires indépendantes

La convolution de deux densités permet de calculer la densité de la somme de deux variables aléatoires indépendantes qui ont des densités :

Proposition 14.4.1 Soient X, Y deux v.a. indépendantes et de densités f et g . La loi de $X + Y$ est donnée par :

$$\forall A \in \mathcal{B}(\mathbb{R}), \quad \mathbb{P}_{X+Y}(A) = \mathbb{P}(X + Y \in A) = \int_A (f * g)(x) dx.$$

Autrement dit, $X + Y$ a pour densité la fonction $f * g$.

Démonstration : On fait la preuve pour $A = [a, b]$. Comme (X, Y) est de densité $(x, y) \mapsto f(x)g(y)$, on a

$$\mathbb{P}(X + Y \in [a, b]) = \mathbb{P}((X, Y) \in D) = \iint_D f(x)g(y) dx dy = \iint_{(x,y); x+y \in [a,b]} f(x)g(y) dx dy$$

où $D = \{(x, y); x+y \in [a, b]\}$. On fait le changement de variable $(x, y) \rightarrow (t, s) = (x, x+y)$. Comme (x, y) varie dans \mathbb{R}^2 de façon que $x + y \in [a, b]$, t décrit tout \mathbb{R} et s décrit $[a, b]$. On a alors :

$$\mathbb{P}(X + Y \in [a, b]) = \int_a^b \int_{-\infty}^{+\infty} f(t)g(s-t) ds dt = \int_a^b (f * g)(s) ds,$$

car le jacobien du changement de variable est

$$Jac = \begin{vmatrix} \frac{\partial t}{\partial x} & \frac{\partial s}{\partial x} \\ \frac{\partial t}{\partial y} & \frac{\partial s}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1,$$

ce qui prouve la proposition. □

Remarque 14.4.1 On connaît bien la loi de la somme $X+Y$ si X et Y sont indépendantes. Sinon, il faut connaître la loi du couple (X, Y) et par exemple sa densité $h(x, y)$ si elle existe pour avoir la loi de $X + Y$ par

$$\mathbb{P}(X + Y \in [a, b]) = \int_{(x,y); x+y \in [a,b]} h(x, y) dx dy = \int_a^b \int_{-\infty}^{+\infty} h(x, y - x) dx dy.$$

Exemples : • Soient X, Y des v.a. indépendantes suivant des lois exponentielles de paramètres respectifs a et b . Quelle est la loi de $S = X + Y$?

Les lois de X et Y sont de densités

$$f(x) = ae^{-ax} \mathbf{1}_{[0, +\infty[}(x), \quad g(y) = be^{-by} \mathbf{1}_{[0, +\infty[}(y).$$

Comme X et Y sont indépendantes, la densité de $X + Y$ est, si $a \neq b$:

$$\begin{aligned} (f * g)(x) &= \int_{-\infty}^{+\infty} g(y) f(x - y) dy \\ &= \int_0^{+\infty} be^{-by} ae^{-a(x-y)} \mathbf{1}_{[0, +\infty[}(x - y) dy \\ &= abe^{-ax} \mathbf{1}_{\{x \geq 0\}} \int_0^x e^{(a-b)y} dy \\ &= \frac{abe^{-ax}}{a - b} (e^{(a-b)x} - 1) \mathbf{1}_{\{x \geq 0\}} \\ &= \frac{ab}{a - b} (e^{-bx} - e^{-ax}) \mathbf{1}_{\{x \geq 0\}}. \end{aligned}$$

Si $a = b$, la densité est

$$\begin{aligned} (f * g)(x) &= \int_{-\infty}^{+\infty} g(y) f(x - y) dy = a^2 \int_{-\infty}^{+\infty} e^{-ay} \mathbf{1}_{\{y \geq 0\}} e^{-a(x-y)} \mathbf{1}_{\{x-y \geq 0\}} dy \\ &= a^2 \mathbf{1}_{\{x \geq 0\}} \int_0^x e^{-ay} e^{-a(x-y)} dy = a^2 \mathbf{1}_{\{x \geq 0\}} e^{-ax} \int_0^x dy = a^2 x \mathbf{1}_{\{x \geq 0\}} e^{-ax}. \end{aligned}$$

•

Proposition 14.4.2 Soient X_1 de loi $\mathcal{N}(m_1, \sigma_1^2)$ et X_2 de loi $\mathcal{N}(m_2, \sigma_2^2)$ alors $X_1 + X_2$ est de loi normale $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$.

Comme $X_1 - m_1 \simeq \mathcal{N}(0, \sigma_1^2)$ et $X_2 - m_2 \simeq \mathcal{N}(0, \sigma_2^2)$, il suffit de voir le cas où $m_1 = m_2 = 0$. Notons alors f_1 et f_2 les densités de X_1 et de X_2 . Celle de $X_1 + X_2$ est donnée par

$$f_1 * f_2(x) = \int_{-\infty}^{+\infty} f_1(t) f_2(x - t) dt = \int_{-\infty}^{+\infty} e^{-t^2/(2\sigma_1^2)} e^{-(x-t)^2/(2\sigma_2^2)} \frac{dt}{\sqrt{2\pi\sigma_1^2} \sqrt{2\pi\sigma_2^2}}$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} \exp - \left(\frac{(\sigma_1^2 + \sigma_2^2)t^2 - 2\sigma_1^2 xt + \sigma_1^2 x^2}{2\sigma_1^2 \sigma_2^2} \right) \frac{dt}{2\pi\sigma_1\sigma_2} \\
&= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp - \left(\frac{((\sigma_1^2 + \sigma_2^2)^{1/2}t - \frac{\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)^{1/2}}x)^2 - \frac{\sigma_1^4}{(\sigma_1^2 + \sigma_2^2)}x^2 + \sigma_1^2 x^2}{2\sigma_1^2 \sigma_2^2} \right) dt \\
&= \frac{1}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp - \left(\frac{((\sigma_1^2 + \sigma_2^2)^{1/2}t - \frac{\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)^{1/2}}x)^2 + \frac{\sigma_1^2 \sigma_2^2}{(\sigma_1^2 + \sigma_2^2)}x^2}{2\sigma_1^2 \sigma_2^2} \right) dt \\
&= \frac{\exp - \frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp - \left(\frac{((\sigma_1^2 + \sigma_2^2)^{1/2}t - \frac{\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)^{1/2}}x)^2}{2\sigma_1^2 \sigma_2^2} \right) dt \\
&= \frac{\exp - \frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}}{2\pi\sigma_1\sigma_2} \int_{-\infty}^{+\infty} \exp - \left(\frac{u^2}{2\sigma_1^2 \sigma_2^2} \right) \frac{du}{(\sigma_1^2 + \sigma_2^2)^{1/2}}
\end{aligned}$$

avec le changement de variable $u = (\sigma_1^2 + \sigma_2^2)^{1/2}t - \frac{\sigma_1^2}{(\sigma_1^2 + \sigma_2^2)^{1/2}}x$. Puis d'après la normalisation de la loi normale $\mathcal{N}(0, \sigma_1^2 \sigma_2^2)$, la dernière intégrale vaut $\frac{\sqrt{2\pi\sigma_1^2 \sigma_2^2}}{(\sigma_1^2 + \sigma_2^2)^{1/2}}$, on a finalement :

$$f_1 * f_2(x) = \frac{\exp - \frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}}{2\pi\sigma_1\sigma_2} \frac{\sqrt{2\pi\sigma_1^2 \sigma_2^2}}{(\sigma_1^2 + \sigma_2^2)^{1/2}} = \frac{\exp - \frac{x^2}{2(\sigma_1^2 + \sigma_2^2)}}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}}.$$

On a obtenu la densité de la loi $\mathcal{N}(0, \sigma_1^2 + \sigma_2^2)$.

Cas de variables aléatoires discrètes

Si X et Y sont des variables aléatoires discrètes indépendantes, alors $X + Y$ est encore une variable discrète de domaine

$$(X + Y)(\Omega) = X(\Omega) + Y(\Omega) = \{x_i + y_j \mid x_i \in X(\Omega), y_j \in Y(\Omega)\}$$

et on peut calculer ses probabilités ponctuelles par une « convolution discrète ».

Supposons pour simplifier que $X(\Omega) = Y(\Omega) = \mathbb{N}$, et notons $(p_k)_{k \in \mathbb{N}}$, $(q_k)_{k \in \mathbb{N}}$ les probabilités ponctuelles de X et Y alors $(X + Y)(\Omega) = \mathbb{N}$ et comme on a pour chaque $k \in \mathbb{N}$ la partition suivante $\{X + Y = n\} = \bigcup_{k=0}^n \{X = k, Y = n - k\}$, il vient

$$\begin{aligned}
\mathbb{P}(X + Y = n) &= \mathbb{P} \left(\bigcup_{k=0}^n \{X = k, Y = n - k\} \right) \\
&= \sum_{k=0}^n \mathbb{P}(X = k, Y = n - k)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{k=0}^n \mathbb{P}(X = k) \mathbb{P}(Y = n - k) \\
&= \sum_{k=0}^n p_k q_{n-k}.
\end{aligned}$$

Les probabilités ponctuelles de $X + Y$ sont donc donnée par $\left(\sum_{k=0}^n p_k q_{n-k} \right)_{k \in \mathbb{N}}$.

Exemple : Soient X, Y des v.a. indépendantes suivant des lois de Poisson de paramètres respectifs α et β . Quelle est la loi de $S = X + Y$?

Les lois de X et Y sont données par

$$\mathbb{P}(X = i) = \frac{e^{-\alpha} \alpha^i}{i!}, \quad \mathbb{P}(Y = j) = \frac{e^{-\beta} \beta^j}{j!}, \quad i, j \in \mathbb{N}.$$

Comme X et Y sont indépendantes, on a en utilisant la formule du binome de Newton :

$$\begin{aligned}
\mathbb{P}(S = n) &= \sum_{i=0}^n \mathbb{P}(X = i) \mathbb{P}(Y = n - i) = \sum_{i=0}^n \frac{e^{-\alpha} \alpha^i}{i!} \frac{e^{-\beta} \beta^{n-i}}{(n-i)!} \\
&= \frac{e^{-(\alpha+\beta)}}{n!} \sum_{i=0}^n C_n^i \alpha^i \beta^{n-i} \\
&= \frac{e^{-(\alpha+\beta)} (\alpha + \beta)^n}{n!}.
\end{aligned}$$

Ainsi $S = X + Y$ suit la loi de Poisson de paramètre $\alpha + \beta$.

Chapitre 15

Transformée de Fourier

15.1 Transformée de Fourier $L^1(\mathbb{R}^n)$

Définition 15.1.1 (Transformée de Fourier) Si $f \in L^1(\mathbb{R}^n)$, la transformée de Fourier de f est la fonction $\mathcal{F}(f)$ définie par

$$\mathcal{F}(f)(u) = \int_{\mathbb{R}^n} e^{-2i\pi\langle x, u \rangle} f(x) dx \quad (15.1)$$

où $\langle x, u \rangle = \sum_{i=1}^n x_i u_i$ est le produit scalaire usuel sur \mathbb{R}^n .

Remarque 15.1.1

- Attention aux notations changeantes dans la littérature pour la transformée de Fourier.

- Dans la suite, pour simplifier on considèrera $\mathbb{R}^n = \mathbb{R}$.

Proposition 15.1.1 (Propriétés de la transformée de Fourier)

1. $\mathcal{F}(f)(u)$ est définie pour tout u et borné par $\|f\|_1$.

2. \mathcal{F} est linéaire.

3. $\mathcal{F}(\tau_h f)(u) = e^{2i\pi hu} \mathcal{F}(f)(u)$ pour $h \in \mathbb{R}$.

4. Si $g(x) = f(\lambda x)$ alors $\mathcal{F}(g)(u) = \frac{1}{|\lambda|^n} \mathcal{F}(f)\left(\frac{u}{\lambda}\right)$.

5. $\mathcal{F}(f * g) = \mathcal{F}(f)\mathcal{F}(g)$.

6. Si $xf(x) \in L^1(\mathbb{R})$ alors $\mathcal{F}(f)$ est dérivable de dérivée

$$\mathcal{F}(f)'(x) = -2i\pi \mathcal{F}(xf).$$

7. Soit f une fonction C^1 , de limites nulle en ∞ et de dérivée $f' \in L^1(\mathbb{R})$ alors

$$\mathcal{F}\left(\frac{\partial f}{\partial x_j}\right)(u) = 2i\pi u_j \mathcal{F}(f)(u).$$

Démonstration :

1. 1) Comme $|e^{2i\pi ux} f(x)| = |f(x)|$, l'intégrale (15.1) est bien définie et

$$|\mathcal{F}(f)(u)| \leq \int_{\mathbb{R}} |e^{2i\pi ux}| |f(x)| dx \leq \int_{\mathbb{R}} |f(x)| dx = \|f\|_1.$$

2. 2) C'est la linéarité de l'intégrale.

3. 3) Avec le changement de variable $y = x + h$, on a :

$$\begin{aligned} \mathcal{F}(\tau_h(f))(u) &= \int_{\mathbb{R}} e^{-2i\pi xu} f(x+h) dx = \int_{\mathbb{R}} e^{-2i\pi(y-h)u} f(y) dy \\ &= e^{2i\pi hu} \int_{\mathbb{R}} e^{-2i\pi yu} f(y) dy \\ &= e^{2i\pi hu} \mathcal{F}(f)(u). \end{aligned}$$

4. 4) Faire le changement de variable $y = \lambda x$:

$$\mathcal{F}(g)(u) = \int e^{-2i\pi ux} f(\lambda x) dx = \int e^{-2i\pi(y/\lambda)u} f(y) \frac{dy}{\lambda} = \frac{1}{\lambda} \mathcal{F}(f)(u/\lambda).$$

5. 5) Grâce au théorème de Fubini (pourquoi s'applique-t-il ?)

$$\begin{aligned} \mathcal{F}(f * g)(u) &= \int_{\mathbb{R}} e^{-2i\pi xu} \int_{\mathbb{R}} f(y) g(x-y) dy dx \\ &= \int_{\mathbb{R}} e^{-2i\pi yu} f(y) \left(\int_{\mathbb{R}} e^{-2i\pi(x-y)u} g(x-y) dx \right) dy \\ &= \int_{\mathbb{R}} e^{-2i\pi yu} f(y) dy \left(\int_{\mathbb{R}} e^{-2i\pi z u} g(z) dz \right) \quad \text{avec } z = x - y \\ &= \mathcal{F}(g)(u) \int_{\mathbb{R}} e^{-2i\pi yu} f(y) dy \\ &= \mathcal{F}(g)(u) \mathcal{F}(f)(u). \end{aligned}$$

6. 6) On dérive sous \int grâce au théorème de convergence dominée. C'est légitime car $|e^{2i\pi ux} f(u)| \leq |f(u)|$, intégrable.

7. 7) On intègre par parties ?? :

$$\begin{aligned} \mathcal{F}(f')(u) &= \int_{\mathbb{R}} e^{-2i\pi xu} f'(x) dx \\ &= [e^{-2i\pi xu} f(x)]_{\mathbb{R}} + 2i\pi u \int_{\mathbb{R}} e^{-2i\pi xu} f(x) dx \\ &= 0?? + 2i\pi u \mathcal{F}(f)(u). \end{aligned}$$

□

Remarque 15.1.2 D'après ses propriétés, la transformée de Fourier vérifie

$$\mathcal{F} : \begin{cases} L^1(\mathbb{R}^n) & \rightarrow \mathcal{B}(\mathbb{R}^n) \\ f & \mapsto \mathcal{F}(f) \end{cases} \quad (\text{applications linéaires continues bornées sur } \mathbb{R}^n)$$

et $\|\mathcal{F}(f)\|_\infty \leq \|f\|_\infty$. La transformée de Fourier \mathcal{F} est linéaire continue.

En fait, on a mieux :

Théorème 15.1.1 (Riemann-Lebesgue) Soit $f \in L^1(\mathbb{R})$, alors $\mathcal{F}(f)(u) \rightarrow 0$ quand $|u| \rightarrow +\infty$.

Démonstration :

- Soit d'abord $f \in \mathcal{D}(\mathbb{R})$, alors

$$\mathcal{F}(f - f'') = \mathcal{F}(f) - \mathcal{F}(f'') = \mathcal{F}(f) - (2i\pi u)^2 \mathcal{F}(f) = (1 + 4\pi^2 u^2) \mathcal{F}(f).$$

Or en tant que transformée de Fourier, $\mathcal{F}(f - f'')$ est bornée par (disons) K . Il suit

$$|\mathcal{F}(f)(u)| \leq \frac{K}{1 + 4\pi^2 u^2} \rightarrow 0, \quad u \rightarrow \pm\infty.$$

- Si $f \in L^1(\mathbb{R})$, par densité de $\mathcal{D}(\mathbb{R})$ dans $L^1(\mathbb{R})$, pour $\varepsilon > 0$, il existe $g \in \mathcal{D}(\mathbb{R})$ tel que $\|f - g\|_1 \leq \varepsilon/2$ et

$$|\mathcal{F}(f)(u) - \mathcal{F}(g)(u)| \leq \|f - g\|_1 \leq \varepsilon/2.$$

Or d'après le premier cas, pour $g \in \mathcal{D}(\mathbb{R})$, il existe A tel que $|u| \geq A$ implique $|\mathcal{F}(g)(u)| < \varepsilon/2$. Mais alors,

$$|\mathcal{F}(f)(u)| \leq |\mathcal{F}(f)(u) - \mathcal{F}(g)(u)| + |\mathcal{F}(g)(u)| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

ce qui conclut car $\varepsilon > 0$ est arbitraire. □

Remarque 15.1.3 On a donc

$$\mathcal{F} : L^1(\mathbb{R}^n) \rightarrow C_0(\mathbb{R}^n)$$

où $C_0(\mathbb{R}^n)$ est l'ensemble des fonctions continues qui tendent vers 0 en $\pm\infty$.

Exemple : Soit $f(x) = e^{-ax^2}$ alors $\mathcal{F}(f)(u) = \sqrt{\pi/a} e^{-(\pi^2/a^2)u^2}$. En particulier, $x \mapsto e^{-\pi x^2}$ est un point fixe de la transformée de Fourier \mathcal{F} .

En effet, la fonction f vérifie l'équation différentielle $f'(x) = -2axf(x)$. En appliquant \mathcal{F} , on a

$$2i\pi u \mathcal{F}(f)(u) = \mathcal{F}(f')(u) = \mathcal{F}(-2axf)(u) = -2a \mathcal{F}(xf)(u) = \frac{-2a}{-2i\pi} \mathcal{F}(f)'(u).$$

La fonction $\mathcal{F}(f)$ est donc solution de $g'(u) = -\frac{2\pi^2}{a}ug(u)$. D'où

$$\mathcal{F}(f)(u) = g(u) = g(0)e^{-\frac{\pi^2}{a}u^2}.$$

Puis, comme

$$g(0) = \mathcal{F}(f)(0) = \int_{\mathbb{R}} e^{2i\pi x \times 0} f(x) dx = \int_{\mathbb{R}} f(x) dx = \int_{\mathbb{R}} e^{-ax^2} dx = \sqrt{\pi/a},$$

on en déduit l'expression de $\mathcal{F}(f)(u) = \sqrt{\pi/a} e^{-\frac{\pi^2}{a}u^2}$.

15.2 Inversion

Comment trouver une fonction dont la transformée de Fourier est une fonction donnée ? C'est l'objet du résultat suivant d'inversion qui, sous certaines conditions, indique qu'il suffit de réappliquer la transformée de Fourier pour trouver (à une symétrie $x \mapsto -x$ près) la fonction cherchée. Précisément, on a le résultat suivant, admis

Théorème 15.2.1 (Inversion) Soit $f \in L^1(\mathbb{R})$ telle que $\mathcal{F}(f) \in L^1(\mathbb{R})$. Alors

$$g(t) = \int_{\mathbb{R}} e^{2i\pi tu} \mathcal{F}(u) du$$

est égale p.p. à f .

En notant $\bar{\mathcal{F}} : h \in L^1(\mathbb{R}) \mapsto \int_{\mathbb{R}} e^{2i\pi tu} h(u) du$ alors si $f \in L^1(\mathbb{R})$ et $\mathcal{F}(f) \in L^1(\mathbb{R})$, on a

$$\bar{\mathcal{F}} \circ \mathcal{F}(f) = f \text{ p.p.}$$

ou encore

$$\mathcal{F} \circ \mathcal{F}(f)(t) = f(-t) = \check{f}(t) \text{ p.p.}$$

Corollaire 15.2.1 La transformation de Fourier \mathcal{F} est injective.

Démonstration : Si $\mathcal{F}(f) = 0$ alors le théorème d'inversion s'applique car $\mathcal{F}(f) \in L^1(\mathbb{R})$ et donc $f(-t) = \mathcal{F}(\mathcal{F}(f))(t) = \mathcal{F}(0)(t) = 0$. D'où $f = 0$. \square

15.3 Transformée de Fourier $L^2(\mathbb{R}^n)$

On a défini la transformée de Fourier pour les fonctions de $L^1(\mathbb{R})$. On va maintenant étendre la transformée sur $L^2(\mathbb{R})$. Commençons par :

Théorème 15.3.1 (Plancherel) *Soit $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ alors on a $\mathcal{F}(f) \in L^2(\mathbb{R})$ et*

$$\|\mathcal{F}(f)\|_2 = \|f\|_2.$$

Démonstration : À voir.

Corollaire 15.3.1 *Il existe une unique isométrie de $L^2(\mathbb{R})$ dans $L^2(\mathbb{R})$ appelée transformation de Fourier L^2 et notée $f \mapsto \hat{f}$ telle que*

- si $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ alors $\mathcal{F}(f) = \hat{f}$ (La transformation $\widehat{}$ prolonge \mathcal{F}).
- $\|\hat{f}\|_2 = \|f\|_2$.

Démonstration : On a vu précédemment que $\mathcal{F} : L^1(\mathbb{R}) \cap L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ est une isométrie pour la norme $\|\cdot\|_2$.

Or $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ est dense dans $L^2(\mathbb{R})$ pour $\|\cdot\|_2$ car contient, S_2 les combinaisons linéaires de fonctions indicatrices (cf. les résultats de densité) qui l'est déjà.

L'espace $(L^2(\mathbb{R}), \|\cdot\|_2)$ étant complet (théorème de Riesz-Fisher), on utilise maintenant le théorème d'extension 3.3.1 (prolongement des applications uniformément continues) pour prolonger $\mathcal{F}|_{L^1(\mathbb{R}) \cap L^2(\mathbb{R})}$ en une isométrie de $L^2(\mathbb{R})$, ce qui prouve le résultat. \square

Proposition 15.3.1 *La transformation de Fourier L^2 $f \mapsto \hat{f}$ est une bijection.*

Démonstration : L'injectivité est due au corollaire 15.2.1. Pour la surjectivité, considérons h_k une approximation de l'unité. D'après les résultats de convolution, on a pour $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, $f * h_k \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. D'où

$$\mathcal{F}(f * h_k) = \widehat{f * h_k} = \mathcal{F}(f)\mathcal{F}(h_k) = \hat{f}\hat{h}_k \in L^1(\mathbb{R}),$$

car \hat{f} et \hat{h}_k sont dans $L^2(\mathbb{R})$ et donc $\hat{f}\hat{h}_k \in L^1(\mathbb{R})$.

On a donc $f * h_k \in L^1(\mathbb{R})$ et $\mathcal{F}(f * h_k) \in L^1(\mathbb{R})$. Par le théorème d'inversion, il suit $\mathcal{F}(\mathcal{F}(f * h_k)) = (f * h_k)^\sim$, c'est à dire

$$\widehat{\widehat{f * h_k}} = (f * h_k)^\sim$$

Or $f * h_k \rightarrow f$ dans $L^2(\mathbb{R})$, par continuité de $\widehat{}$ et de $^\sim$, à la limite on a pour toute $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$

$$\hat{\hat{f}} = f.$$

Par densité de $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ et par continuité de $\widehat{}$ et de $^\sim$, pour $\|\cdot\|_2$, le résultat subsiste pour toute $f \in L^2(\mathbb{R})$. \square

À voir : Espace $S(\mathbb{R}^n)$. ????

15.4 Fonction caractéristique

Définition et propriétés

Définition 15.4.1 Soit X une variable aléatoire définie sur l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle fonction caractéristique de X la fonction de \mathbb{R} dans \mathbb{C}

$$\phi_X(t) = E[e^{itX}] = \int_{\Omega} e^{itX} d\mathbb{P}. \quad (15.2)$$

Remarque 15.4.1 • ϕ_X est toujours bien définie car $|e^{itX}| = 1$ est \mathbb{P} -intégrable.

- Si X est de densité f , la fonction caractéristique s'écrit

$$\phi_X(t) = \int_{\mathbb{R}} e^{itx} f(x) dx = \mathcal{F}(f)\left(\frac{t}{2\pi}\right).$$

À quelques notations près, il s'agit de la transformée de Fourier de la densité f de la variable X .

- Si X est une variable discrète, la fonction caractéristique peut se voir encore comme une transformée de Fourier mais il faut alors introduire une transformée de Fourier par rapport à une mesure discrète.

- De façon générale

$$\phi_X(t) = E[e^{itX}] = \int_{\Omega} e^{itx} d\mathbb{P}_X(x).$$

La fonction caractéristique ϕ_X est donc la transformée de Fourier de la mesure \mathbb{P}_X , loi de la v.a. X .

Proposition 15.4.1 La fonction caractéristique caractérise la loi : si $\phi_X = \phi_Y$ alors les variables X et Y ont même loi et réciproquement.

Démonstration : Voyons cela dans le cas à densité. Comme $|\phi_X(t)| \leq E[|e^{itX}|] = 1$, ϕ_X est bornée donc intégrable par rapport à \mathbb{P} . Or $\phi_X(t) = \mathcal{F}(f)(t/(2\pi))$ si X est de densité f . Le théorème d'inversion s'applique et donne

$$f(-t/(2\pi)) = \mathcal{F}(\phi_X)(t).$$

La densité s'obtient donc à partir de la fonction caractéristique. Comme la densité caractérise la loi, la fonction caractéristique aussi : si X et Y ont même fonction caractéristique ϕ alors leur densité f et g se valent :

$$f(t) = \mathcal{F}(\phi)(-2\pi t) = g(t).$$

De façon générale, en appliquant le théorème d'inversion aux transformées de Fourier des lois \mathbb{P}_X et \mathbb{P}_Y , on montre qu'elle s'exprime identiquement à partir de leur fonction caractéristique commune. On a donc $\mathbb{P}_X = \mathbb{P}_Y$. \square

Exemples :

- Pour des variables à densité :

Variabes	Densité f	Intervalle	Fonction caractéristique ϕ_X
Loi normale standard	$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$	$-\infty < x < +\infty$	$e^{-t^2/2}$
Loi normale	$\frac{1}{\sqrt{2\pi\sigma^2}}e^{-(x-m)^2/(2\sigma^2)}$	$-\infty < x < +\infty$	$e^{imt-\sigma^2t^2/2}$
Loi uniforme	1	$0 \leq x \leq 1$	$\frac{e^{it}-1}{it}$
Loi exponentielle	e^{-x}	$0 \leq x < +\infty$	$\frac{1}{1-it}$
Loi de Cauchy	$\frac{1}{\pi(1+x^2)}$	$-\infty < x < +\infty$	$e^{- t }$

- Pour des variables discrètes X de domaine $X(\Omega) = \{x_1, \dots, x_n, \dots\}$ de masse ponctuelle $p_k = \mathbb{P}(X = x_k)$ en x_k , on a

$$\phi_X(t) = E[e^{itX}] = \sum_{k=1}^{+\infty} e^{itx_k} p_k$$

Variabes	Domaine $X(\omega)$	Probabilités ponctuelles p_k	Fonction caractéristique ϕ_X
Loi de Bernoulli	$\{0, 1\}$	$1 - p, p$	$1 - p + pe^{it}$
Loi binomiale	$\{0, \dots, n\}$	$C_n^k p^k (1-p)^{n-k}$	$(1 - p + pe^{it})^n$
Loi de Poisson	\mathbb{N}	$e^{-\alpha} \frac{\alpha^k}{k!}$	$e^{\alpha(e^{it}-1)}$

La régularité de la fonction caractéristique ϕ_X est liée à l'intégrabilité de la v.a. X :

Proposition 15.4.2 • $\phi_X(t)$ est définie pour tout $t \in \mathbb{R}$.

- Si X a un moment d'ordre 1, alors ϕ_X est dérivable (et même C^1) avec

$$\phi_X'(t) = iE[Xe^{itX}].$$

En particulier : $iE[X] = \phi_X'(0)$.

- Plus généralement, si f a un moment d'ordre p alors ϕ_X est p fois dérivable (et même de classe C^p), avec

$$\phi_X^{(p)}(t) = i^p E[X^p e^{itX}].$$

En particulier $i^p E[X^p] = \phi_X^{(p)}(0)$.

Démonstration : Pour les résultats de dérivation, il suffit d'appliquer le théorème de convergence dominée pour dériver sous le signe E . La domination est à chaque fois garantie par l'existence des moments nécessaires :

$$|X^p e^{itX}| \leq |X|^p \quad \text{est } \mathbb{P}\text{-intégrable}$$

par existence du moment d'ordre p . □

Propriétés

Théorème 15.4.1 *Si deux variables aléatoires X et Y sont indépendantes alors*

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t).$$

Plus généralement, pour n variables aléatoires X_1, \dots, X_n indépendantes on a

$$\phi_{X_1+\dots+X_n}(t) = \phi_{X_1}(t) \dots \phi_{X_n}(t).$$

Démonstration : On a

$$\phi_{X+Y}(t) = E[e^{it(X+Y)}] = E[e^{itX}e^{itY}] = E[e^{itX}]E[e^{itY}]$$

d'après la proposition 12.5.2. La généralisation au cas de n variables est immédiate. □

On définit aussi la fonction caractéristique d'un vecteur de la façon suivante :

Définition 15.4.2 *Soit (X, Y) un vecteur aléatoire, on définit sa fonction caractéristique comme la fonction de 2 variables de \mathbb{R}^2 dans \mathbb{C} :*

$$\phi_{(X,Y)}(t, s) = E[e^{i\langle (t,s), (X,Y) \rangle}] = E[e^{i(tX+sY)}]$$

où $\langle (t, s), (x, y) \rangle = tx + sy$ est le produit scalaire euclidien de \mathbb{R}^2 .

Il s'agit de la transformée de Fourier de la mesure de \mathbb{R}^2 $\mathbb{P}_{(X,Y)}$ loi du vecteur (X, Y) . On a alors un critère d'indépendance de deux variables aléatoires :

Proposition 15.4.3 *Deux variables aléatoires X et Y sont indépendantes ssi*

$$\phi_{(X,Y)}(t, s) = \phi_X(t)\phi_Y(s).$$

Démonstration : On raisonne par équivalence :

$$\begin{aligned} X \perp Y &\Leftrightarrow d\mathbb{P}_{X,Y}(dx, dy) = d\mathbb{P}_X(dx)d\mathbb{P}_Y(dy) \\ &\Leftrightarrow \int_{\mathbb{R}^2} e^{i(tx+sy)} d\mathbb{P}_{X,Y}(dx, dy) = \int_{\mathbb{R}^2} e^{i(tx+sy)} d\mathbb{P}_X(dx)d\mathbb{P}_Y(dy) \\ &\Leftrightarrow \int_{\mathbb{R}^2} e^{i(tx+sy)} d\mathbb{P}_{X,Y}(dx, dy) = \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X(dx) \int_{\mathbb{R}} e^{isy} d\mathbb{P}_Y(dy) \\ &\Leftrightarrow \int_{\Omega} e^{i(tX+sY)} d\mathbb{P} = \int_{\Omega} e^{itX} d\mathbb{P} \int_{\Omega} e^{isY} d\mathbb{P} \\ &\Leftrightarrow \phi_{(X,Y)}(t, s) = \phi_X(t)\phi_Y(s) \end{aligned}$$

où on admet l'équivalence de la deuxième ligne (c'est la même que celle sur la caractérisation de la loi de X par ϕ_X) et où on a utilisé ensuite les théorèmes de Fubini, puis de transfert. \square

On généralise immédiatement au cas de n variables la définition de la fonction caractéristique du vecteur (X_1, \dots, X_n)

$$\phi_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = E[e^{i(t_1 X_1 + \dots + t_n X_n)}]$$

et le critère d'indépendance de n variables aléatoires.

On a maintenant le critère suivant pour la convergence en loi :

Théorème 15.4.2 (Paul Lévy) *La convergence en loi des variables aléatoires est équivalente à la convergence simple de leur fonction caractéristique :*

$$X_n \xrightarrow{\mathcal{L}} X \quad \text{ssi} \quad \forall t \in \mathbb{R}, \phi_{X_n}(t) \rightarrow \phi_X(t).$$

Démonstration : • Dans le sens direct : on sait que la convergence en loi est équivalente à $E[f(X_n)] \rightarrow E[f(X)]$ pour toute fonction f continue bornée. Avec $f(x) = e^{itx}$, on a $\phi_{X_n}(t) \rightarrow \phi_X(t)$.

• Le sens indirect est plus difficile et admis. \square

Théorème central limite (TCL)

Dans la suite i.i.d. signifiera indépendant(e)s et identiquement distribué(e)s, c'est à dire de même loi.

On déduit du théorème de Paul Lévy un résultat fondamental en probabilité et en statistiques : le théorème central limite (TCL) dû à Paul Lévy.

Théorème 15.4.3 (TCL) *Soit $(X_n)_n$ une suite de variables aléatoires réelles indépendantes et de même loi, d'espérance m_1 et de variance finie σ^2 . Soit $S_n = X_1 + \dots + X_n$ la somme partielle. Alors quand $n \rightarrow +\infty$*

$$\frac{S_n - nm_1}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Remarque 15.4.2 • D'un certain point de vue ce résultat complète la loi des grands nombres : la LGN dit que la moyenne arithmétique

$$\frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n}$$

converge quand $n \rightarrow +\infty$ vers l'espérance $m_1 = E[X_1]$ (la moyenne probabiliste) en probabilité ou presque sûrement selon la version de la LGN. On sait donc que S_n est équivalent à nm_1 . Le TCL indique (en un sens faible) que la vitesse de cette convergence est en $1/\sqrt{n}$

(sens faible car la convergence du TCL est une convergence assez faible : la convergence en loi).

• En plus dans le TCL, apparaît à la limite la loi $\mathcal{N}(0, 1)$ alors que les v.a. X_i sont de lois arbitraires : ce résultat justifie le rôle universel de la loi normale. Elle modélise les petites variations de n'importe quelle loi (avec un moment d'ordre 2) par rapport à sa moyenne.

Démonstration : Posons $Y_i = X_i - m_1$, si bien que les v.a. Y_i sont indépendantes de même loi avec $E[Y_i] = 0$, $\text{Var}(Y_i) = \text{Var}(X_i)$. Notons $S'_n = Y_1 + \dots + Y_n$ et $Z_n = \frac{S_n - nm_1}{\sigma\sqrt{n}} = \frac{S'_n}{\sigma\sqrt{n}}$. On a

$$\begin{aligned}\phi_{Z_n}(t) &= E[\exp\{it\frac{S'_n}{\sigma\sqrt{n}}\}] = E[\exp\{i\frac{t}{\sigma\sqrt{n}}S'_n\}] = \phi_{S'_n}(\frac{t}{\sigma\sqrt{n}}) \\ &= \phi_{Y_1}(\frac{t}{\sigma\sqrt{n}}) \dots \phi_{Y_n}(\frac{t}{\sigma\sqrt{n}}) \\ &= \left(\phi_{Y_1}(\frac{t}{\sigma\sqrt{n}})\right)^n\end{aligned}$$

en utilisant $\phi_{Y_1+\dots+Y_n} = \phi_{Y_1} \dots \phi_{Y_n} = \phi_{Y_1}^n$ par indépendance et identique distribution des variables Y_i .

Comme Y_1 a un moment d'ordre 2, ϕ_{Y_1} est dérivable 2 fois avec $\phi_{Y_1}(0) = 1$, $\phi'_{Y_1}(0) = iE[Y_1] = 0$ et $\phi''_{Y_1}(0) = i^2E[Y_1^2] = -\sigma^2$. La formule de Taylor à l'ordre 2 en 0 donne alors

$$\begin{aligned}\phi_{Y_1}(x) &= \phi_{Y_1}(0) + x\phi'_{Y_1}(0) + \frac{x^2}{2}\phi''_{Y_1}(0) + x^2\epsilon(x) \\ &= 1 - \frac{\sigma^2x^2}{2} + x^2\epsilon(x)\end{aligned}$$

où la fonction ϵ vérifie $\lim_{x \rightarrow 0} \epsilon(x) = 0$. On a donc

$$\begin{aligned}\phi_{Z_n}(t) &= \left(\phi_{Y_1}(\frac{t}{\sigma\sqrt{n}})\right)^n = \left(1 - \frac{\sigma^2t^2}{2\sigma^2\sqrt{n}^2} + \frac{t}{\sigma^2\sqrt{n}^2}\epsilon(t/(\sigma\sqrt{n}))\right)^n = \left(1 - \frac{t^2}{2n} + \frac{1}{n}\epsilon(1/\sqrt{n})\right)^n \\ &= \exp\left(n \ln\left(1 - \frac{t^2}{2n} + \frac{1}{n}\epsilon(1/\sqrt{n})\right)\right) = \exp\left(n\left(-\frac{t^2}{2n} + \frac{1}{n}\epsilon(1/\sqrt{n})\right)\right) \\ &= \exp\left(-\frac{t^2}{2} + \epsilon(1/\sqrt{n})\right).\end{aligned}$$

D'où pour chaque $t \in \mathbb{R}$,

$$\lim_{n \rightarrow +\infty} \phi_{Z_n}(t) = \exp\left\{-\frac{t^2}{2}\right\} = \phi_{\mathcal{N}(0,1)}(t).$$

Le théorème de Paul Lévy affirme alors que Z_n converge en loi vers $\mathcal{N}(0, 1)$, c'est la conclusion du TCL. \square

Remarque 15.4.3 En général, lorsque n est grand, on approxime la loi d'une somme de v.a. i.i.d. de $L^2(\Omega)$ par une loi normale grâce au TCL de la façon suivante. Soit $S_n = X_1 + \dots + X_n$ la somme de v.a. i.i.d. X_i avec $\sigma_X^2 < +\infty$, on a d'après le TCL

$$\frac{X_1 + \dots + X_n - nE[X_1]}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Quand n est grand on approxime alors la loi de $\frac{X_1 + \dots + X_n - nE[X_1]}{\sigma\sqrt{n}}$ par celle de $Y = \mathcal{N}(0, 1)$. Si bien que la loi de la somme $S_n = X_1 + \dots + X_n$ est approximée par celle de

$$nE[X_1] + \sigma\sqrt{n}Y \simeq \mathcal{N}(nE[X_1], \sigma^2n).$$

Règle statistique : La somme S_n d'une suite de v.a.i.i.d. L^2 de moyenne m_1 et de variance σ^2 s'approxime par

$$S_n \approx \mathcal{N}(nm_1, \sigma^2n).$$