MESURE, INTÉGRATION ET PROBABILITÉS

MAT311–MAT321–MAP311–MAP312

Promotion X2017

Résumé

Ce texte est un résumé de cours sur les notions de mesure, intégrale de Lebesgue et de probabilité. Son but premier est de faire le parallèle entre le point de vue de l'analyse et celui des probabilités sur ces questions. Ainsi, la première partie « Mesures et intégrales » sera suivie d'une partie « Probabilité et espérance » où des notions et des résultats analogues sont introduits avec le vocabulaire de la théorie des probabilités, plutôt qu'avec celui de la théorie des ensembles et de l'analyse.

A. Mesures et intégrales

On distingue essentiellement deux points de vue pour l'intégration de Lebesgue des fonctions.
- celui des mesures (adapté aux probabilités, systèmes dynamiques etc.) ;
- celui des formes linéaires sur les espaces de fonctions (adapté par exemple aux équations aux dérivées partielles etc.).

Dans le second point de vue [1], on voit l'opération $f \mapsto \int_U f(x)\,\mathrm{d}x$ d'intégration sur un ouvert U de $\mathbb{R}^d$ comme une *forme linéaire* sur un espace de fonctions (convenables) sur U.

Dans cette approche, on travaille sur des espaces de fonctions de plus en plus gros et on généralise l'intégrale par passages à la limite successifs. Les espaces fonctionnels ont en commun de contenir l'ensemble $\mathscr{C}_c(U)$ des fonctions continues, à valeurs réelles, et nulles en-dehors d'un compact (non fixé) de U. En particulier, on impose que le calcul intégral sur $\mathscr{C}_c(U)$ soit celui que l'on connaît déjà.

Une fois l'intégrale de Lebesgue construite par cette approche, on définit la mesure d'une partie (convenable) A de U comme étant $\int_U \mathbf{1}_A(x)\,\mathrm{d}x$, où $\mathbf{1}_A$ est la fonction caractéristique [2] de A.

Dans ce texte, nous présentons le point de vue des mesures, en définissant d'abord une notion de *volume* pour des parties (convenables) de l'ensemble sur lequel on veut intégrer des fonctions (convenables). Nous allons motiver ce point de vue en revenant à la définition de deux intégrales classiques.

1. Intégrale de Riemann

1.1 Notion d'intégrale au sens de Riemann

Rappelons qu'une fonction sur un intervalle $[a,b]$ est dite *réglée* si elle admet une limite à droite en tout point de $[a,b[$ et une limite à gauche en tout point de $]a,b]$. Un des résultats de base sur cette classe de fonctions est qu'elle coïncide avec l'ensemble des limites uniformes de fonctions en escalier [3] sur $[a,b]$. La définition initiale permet de vérifier concrètement qu'une fonction

1. C'est le point de vue adopté dans le polycopié *Introduction à l'analyse réelle*, distribué pour MAT311 et MAT321

2. La *fonction caractéristique* $\mathbf{1}_A$ d'une partie $A \subset \Omega$ d'un ensemble Ω est l'application de Ω dans $\mathbb{R}$ qui associe 1 aux élements de A et 0 aux éléments de $\Omega \setminus A$.

3. Les fonctions en escalier sont les fonctions constantes par morceaux.

donnée est réglée. Le critère de limite uniforme permet de définir l'intégrale des fonctions réglées à partir de celle (évidente) des fonctions en escalier : si f est une telle fonction, on l'écrit comme limite uniforme $f = \lim_{n\to\infty} f_n$ et on pose :

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x.$$

Enfin, on vérifie que cette limite d'intégrales ne dépend pas du choix de la suite $(f_n)_{n\geq 0}$.

Par ailleurs, on rappelle que si f est une fonction $[a,b] \to \mathbb{R}$ bornée, l'*intégrale supérieure* (resp. *inférieure*) de f est la borne inférieure (resp. supérieure) des intégrales des fonctions en escalier supérieures à f (resp. inférieures à f). L'intégrale supérieure majore toujours l'intégrale inférieure, et f est dite *intégrable au sens de Riemann* (ou *Riemann-intégrable*) quand il y a égalité.

L'intégrale de Riemann est une généralisation de l'intégrale précédente dans le sens où toute fonction réglée est Riemann-intégrable, et dans ce cas les deux d'intégrales coïncident.

Notons qu'il existe des fonctions Riemann-intégrables qui ne sont pas réglées, comme par exemple la fonction $f : [0,1] \to \mathbb{R}$ qui vaut $\sin(\frac{1}{x})$ pour $x \in]0,1]$.

Les fonctions Riemann-intégrables sur un segment $[a,b]$ donné forment un espace vectoriel et l'application $f \mapsto \int_a^b f(x)\,\mathrm{d}x$ est une forme linéaire positive. L'application $f \mapsto \int_a^b |f(x)|\,\mathrm{d}x$ est une semi-norme : elle vérifie les axiomes d'une norme, sauf celui de *séparation*.

1.2 Inconvénients de l'intégrale de Riemann

L'intégrale de Riemann présente un certain nombre d'inconvénients. Une observation classique est que la fonction caractéristique de $[0,1] \cap \mathbb{Q}$ n'est pas Riemann-intégrable alors qu'elle est nulle sur une partie de $[0,1]$ de complémentaire dénombrable. Pire : en numérotant les nombres rationnels de $[0,1] \cap \mathbb{Q}$ on produit une suite $(f_n)_{n\geq 0}$ de fonctions Riemann-intégrables : f_n est la fonction caractéristique des n « premiers » nombres rationnels entre 0 et 1. Ces fonctions sont uniformément bornées, chacune est d'intégrale nulle et $(f_n)_{n\geq 0}$ converge simplement ; cependant, la limite n'est pas Riemann-intégrable.

Le point de départ de la théorie de l'intégration de Lebesgue est d'observer que toute fonction bornée est limite uniforme d'une suite de fonctions *ne prenant qu'un nombre fini de valeurs*. Ces dernières fonctions sont de la forme

$$\sum_{i=1}^{k} \lambda_i \mathbf{1}_{\{x\in[a,b]\,:\, f(x)=\lambda_i\}}.$$

Sous ce point de vue, par linéarité, il s'agit de savoir calculer en premier lieu l'intégrale d'une fonction caractéristique d'une partie A de $[a,b]$; cette intégrale peut être vue comme une « mesure » de A. Ce point de vue se généralise facilement à des ensembles quelconques, bien plus généraux que $\mathbb{R}$ ou $\mathbb{R}^d$ et sera particulièrement bien adapté à la théorie des probabilités.

2. Tribus de parties d'un ensemble

2.1 Rappels : dénombrabilité et limite inférieure

La notion de dénombrabilité est essentielle pour la théorie de la mesure : rappelons qu'un ensemble Ω est dit *dénombrable* s'il est fini ou en bijection avec $\mathbb{N}$. Pour cela, il suffit qu'il existe une application surjective $\mathbb{N} \to X$ ou injective $X \to \mathbb{N}$.

Rappelons les propriétés suivantes :

— *Tout produit fini d'ensembles dénombrables est dénombrable* ;

— *Toute réunion dénombrable d'ensembles dénombrables est dénombrable.*

Nous aurons aussi besoin de la notion de *limite inférieure* (éventuellement infinie) d'une suite numérique. La droite numérique achevée [4] $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ est munie d'une distance qui induit la

4. On utilisera aussi par la suite la notation $\overline{\mathbb{R}}_+ = [0,+\infty[\cup \{+\infty\}$

topologie usuelle sur $\mathbb{R}$ et assure qu'une suite réelle $(x_n)_{n\geq 0}$ tend vers $\pm\infty$ au sens de la distance de $\overline{\mathbb{R}}$ si, et seulement si, $\lim_{n\to+\infty} x_n = \pm\infty$ au sens usuel. Une telle distance est par exemple donnée par $d(x,y) = |f(x) - f(y)|$ pour $f(x) = \frac{x}{1+|x|}$ et $f(\pm\infty) = \pm 1$. Elle fait de $\overline{\mathbb{R}}$ un espace compact, si bien que toute suite numérique $(x_n)_{n\geq 0}$ a des valeurs d'adhérence dans $\overline{\mathbb{R}}$, notamment une *plus petite valeur d'adhérance* : celle-ci est notée $\liminf_{n\to\infty} x_n$ et est aussi égale à la limite de la suite croissante $(z_n)_{n\geq 0}$ définie par $z_n = \inf\{x_k : k \geq n\}$.

2.2 Les axiomes d'une tribu

Pour définir une notion cohérente de mesure de parties, on a recours à la notion de tribu.

Définition 1. Une tribu (ou σ-algèbre) sur un ensemble Ω est un ensemble $\mathscr{A}$ de parties de Ω satisfaisant les axiomes suivants :

- La partie vide $\varnothing$ est dans $\mathscr{A}$;
- Une réunion dénombrable de parties de Ω appartenant à $\mathscr{A}$ est encore une partie de Ω appartenant à $\mathscr{A}$;
- Le complémentaire d'une partie de Ω appartenant à $\mathscr{A}$ est encore une partie de Ω appartenant à $\mathscr{A}$.

Le couple $(\Omega, \mathscr{A})$ est appelé *espace mesurable*.

En bref : une tribu de parties de Ω est un ensemble de parties de Ω contenant l'ensemble vide et Ω lui-même, et stable par réunion *dénombrable* et par passage au complémentaire.

2.3 Propriétés élémentaires et exemples

Propriétés élémentaires : il découle des axiomes qu'une tribu est stable par deux opérations ensemblistes importantes, la différence et la différence symétrique. Rappelons que si A et B sont des parties de Ω, la *différence* de A et de B est $A \setminus B = \{x \in X : x \in A \text{ mais } x \notin B\}$, et que leur *différence symétrique* est $A \Delta B = (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A)$.

Une tribu est aussi stable par des opérations ensemblistes plus subtiles : plus grande et plus petite limite d'une suite de parties. Si $\{A_n\}_{n\geq 0}$ est une suite de parties de Ω, la plus grande limite $\overline{\lim}_{n\to\infty} A_n$ de $\{A_n\}_{n\geq 0}$ est l'ensemble des points qui appartiennent à A_n pour une infinité d'indices n, et la plus petite limite $\underline{\lim}_{n\to\infty} A_n$ de $\{A_n\}_{n\geq 0}$ est l'ensemble de ceux qui appartiennent à tous les A_n à partir d'un certain rang.

Exemple : la paire $\{\varnothing; X\}$, ainsi que l'ensemble $\mathscr{P}(\Omega)$ de *toutes* les parties de Ω, sont des tribus triviales sur Ω.

Remarque 1. Soit (Ω, d) un espace métrique (par exemple, la droite réelle munie de la valeur absolue). Dans ce cas, la distance d, via les boules ouvertes associées, permet de définir la collection des ouverts de Ω : un ouvert est une réunion de boules ouvertes. La topologie de Ω est l'ensemble de ses ouverts ; ce n'est en général pas une tribu sur Ω.

2.4 Tribus engendrées

Il y a une relation d'ordre pour les tribus sur un ensemble Ω fixé : on dit qu'une tribu $\mathscr{A}$ *contient* une tribu $\mathscr{B}$ si toute partie de Ω qui est dans $\mathscr{B}$ est aussi dans $\mathscr{A}$. Pour un ensemble $\mathscr{E}$ *quelconque* de parties de Ω, on peut alors définir la plus petite tribu sur Ω contenant $\mathscr{E}$, ou tribu *engendrée* par $\mathscr{E}$.

En effet si l'on considère $(\mathscr{A}_i)_{i\in I}$ une famille de tribus sur Ω, on peut montrer que $\bigcap_{i\in I} \mathscr{A}_i$ est une tribu de Ω. Le vide étant dans toutes les tribus $\mathscr{A}_i$, il est dans $\bigcap_{i\in I} \mathscr{A}_i$; si $A \in \bigcap_{i\in I} \mathscr{A}_i$ alors par stabilité par passage au complémentaire des tribus, on a $\Omega \setminus A \in \mathscr{A}_i$ pour tout $i \in I$ soit $\Omega \setminus A \in \bigcap_{i\in I} \mathscr{A}_i$, et la stabilité de $\bigcap_{i\in I} \mathscr{A}_i$ par réunion dénombrable se voit de façon similaire.

Définition 2. La *tribu engendrée* par $\mathscr{E}$ est l'intersection de toutes les tribus contenant $\mathscr{E}$.

On note que cette définition est très simple, mais elle n'est pas constructive. En général une tribu engendrée contient beaucoup de parties et il ne faut pas chercher à se la représenter en entier.

Remarquons qu'il existe toujours au moins une tribu contenant $\mathscr{E}$, à savoir $\mathscr{P}(\Omega)$.

2.5 Ensembles boréliens d'un espace métrique

Considérons maintenant le cas particulier où l'ensemble Ω est un espace métrique muni d'une distance d.

Définition 3. Soit (Ω, d) un espace métrique. La *tribu borélienne* de Ω est la tribu engendrée par la topologie de Ω, c'est-à-dire par la collection des ouverts de Ω. La tribu borélienne est notée $\mathrm{Bor}(\Omega, d)$, ou $\mathrm{Bor}(\Omega)$ s'il n'a pas de risque de confusion.

Un ensemble *borélien* de Ω est un élément de la tribu borélienne de Ω.

Par définition, les ouverts de Ω sont des boréliens, et par stabilité d'une tribu par passage au complémentaire, les fermés de Ω sont aussi des boréliens.

On rappelle que dans $\mathbb{R}^d$, tout ouvert est *réunion dénombrable de cubes ouverts*. En particulier, il est équivalent de définir la *tribu borélienne sur* $\mathbb{R}$, notée $\mathrm{Bor}(\mathbb{R})$ ou $\mathscr{B}(\mathbb{R})$ (respectivement, sur $\mathbb{R}^d$, notée $\mathrm{Bor}(\mathbb{R}^d)$ ou $\mathscr{B}(\mathbb{R}^d)$) comme la tribu engendrée par les intervalles ouverts de $\mathbb{R}$ (respectivement, les cubes ouverts de $\mathbb{R}^d$).

Remarque 2. On peut prouver qu'il y a beaucoup plus de parties de $\mathbb{R}$ que de boréliens, mais paradoxalement la question d'exhiber une partie non borélienne de $\mathbb{R}$ est délicate.

2.6 Tribu image réciproque

Nous souhaitons savoir comment la structure de tribu se comporte vis-à-vis des applications. Nous introduisons la notion de tribu image réciproque.

Proposition 1. *Soit $f : \Omega_1 \to \Omega_2$ une application d'un ensemble Ω_1 vers un ensemble Ω_2.*

- *L'image réciproque par f de toute tribu de Ω_2 est une tribu sur Ω_1.*
- *Si $\mathscr{A}_1$ est une tribu sur Ω_1, alors l'ensemble $\mathscr{A}_2$ des parties de Ω_2 dont l'image réciproque est dans $\mathscr{A}_1$ est une tribu de Ω_2.*

Preuve. On prouve le second point, la démonstration étant similaire pour le premier. La partie vide de Ω_2 a pour image réciproque par f la partie vide de Ω_1, qui est dans $\mathscr{A}_1$; ainsi $\varnothing \in \mathscr{A}_2$. Soit $A \in \mathscr{A}_2$; alors le complémentaire $\Omega_2 \setminus A$ vérifie $f^{-1}(\Omega_2 \setminus A) = \Omega_1 \setminus f^{-1}(A)$. Or $f^{-1}(A) \in \mathscr{A}_1$ par définition de $\mathscr{A}_2$, et donc $\Omega_1 \setminus f^{-1}(A) \in \mathscr{A}_1$ par stabilité d'une tribu par passage au complémentaire ; ainsi $\Omega_2 \setminus A \in \mathscr{A}_2$. La stabilité de $\mathscr{A}_2$ par réunion dénombrable vient de ce que l'image réciproque d'une réunion est la réunion des images réciproques. □

3. Mesures positives sur les tribus

3.1 Notion de mesure et propriétés élémentaires

Définition 4. Soit $(\Omega, \mathscr{A})$ un espace mesurable. On appelle mesure (positive) sur $(\Omega, \mathscr{A})$ une application $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ satisfaisant les axiomes suivants.

- Mesure de l'ensemble vide : $\mu(\varnothing) = 0$.
- Additivité dénombrable : si $\{A_n\}_{n \geq 0}$ est une suite (finie ou dénombrable) d'éléments de $\mathscr{A}$ deux à deux disjoints, alors :

$$\mu\Big(\bigcup_{n \geq 0} A_n\Big) = \sum_{n \geq 0} \mu(A_n).$$

La quantitée $\mu(\Omega)$ est appelée la *masse totale*. On dit que la mesure μ est finie (ou bornée) si $\mu(\Omega) < +\infty$; La mesure μ est appelée *probabilité* si $\mu(\Omega) = 1$ (voir Partie B).

S'il existe une suite $(A_n)_n$ d'éléments de $\mathscr{A}$ telle que $\cup_n A_n = \Omega$, et si pour tout $n \in \mathbb{N}$, $\mu(A_n)$ est fini, on dit que μ est une mesure σ-finie.

Définition 5. On appelle *espace mesuré* tout triplet $(\Omega, \mathscr{A}, \mu)$ où $(\Omega, \mathscr{A})$ est un espace mesurable et μ est une mesure positive sur $(\Omega, \mathscr{A})$.

Définition 6. Pour un espace mesuré $(\Omega, \mathscr{A}, \mu)$, une partie de Ω est dite *négligeable* (ou μ-négligeable) si elle est contenue dans une partie mesurable de mesure nulle.

Définition 7. Une propriété est vraie *μ-presque partout* (p.p.) si elle est vraie en dehors d'un ensemble μ-négligeable.

3.2 Exemples de mesures

Nous donnons quelques exemples de mesures, qui illustrent la grande variété de possibilités.

— La *mesure grossière* sur une tribu $\mathscr{A}$ est la mesure qui vaut $+\infty$ sur toutes les parties de $\mathscr{A}$ différentes de $\varnothing$.
— La *mesure nulle* est celle qui vaut 0 tout le temps.
— La *mesure de Dirac* en un point $a \in X$ est la mesure qui vaut 1 sur tout $A \in \mathscr{A}$ tel que $a \in A$, et 0 autrement.
— La *mesure de comptage* sur Ω est la mesure – définie sur $\mathscr{P}(\Omega)$ tout entier – qui vaut le cardinal de $A \subset X$ si A est fini, et $+\infty$ sinon.

Un exemple fondamental qui conduit à la bonne généralisation de l'intégrale de Riemann est le cas de la *mesure de Lebesgue*.

3.3 Définition de la mesure de Lebesgue

Définition 8. Il existe une unique mesure λ_d sur la tribu borélienne de $\mathbb{R}^d$ telle que la mesure de tout pavé $A = [a_1, b_1] \times \ldots \times [a_d, b_d]$ est $\lambda_d(A) = \prod_{i=1}^{d}(b_i - a_i)$. Cette mesure est appelée *mesure de Lebesgue* de $\mathbb{R}^d$.

L'existence et l'unicité de cette mesure ne sont pas du tout évidentes.

De manière équivalente, la mesure de Lebesgue est l'unique mesure définie sur la tribu borélienne de $\mathbb{R}^d$ et possédant les propriétés suivantes.

— Elle est invariante par toute translation de $\mathbb{R}^d$.
— Elle attribue au cube unité la mesure 1.

Quelques propriétés pratiques.

— Pour tout borélien A de $\mathbb{R}^d$, on a $\lambda_d(A) = \inf_U \lambda_d(U)$ où U parcourt l'ensemble des ouverts contenant A.
— Pour tout borélien A de $\mathbb{R}^d$, on a $\lambda_d(A) = \sup_K \lambda_d(K)$ où K parcourt l'ensemble des compacts contenus dans A.
— Pour tout point $x \in \mathbb{R}^d$, $\lambda_d(\{x\}) = 0$.
— Pour des réels $a < b$, on a $\lambda_1([a,b]) = \lambda_1(]a,b]) = \lambda_1(]a,b[) = \lambda_1([a,b[) = b - a$.

3.4 Propriétés des mesures

Énonçons les propriétés des mesures qui se déduisent directement des axiomes.

Dans les énoncés qui suivent, $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ désigne une mesure sur un espace mesurable $(\Omega, \mathscr{A})$.

Lemme 1 (croissance des mesures)**.** *Une mesure $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ est une fonction croissante pour la relation d'ordre donnée par l'inclusion sur les parties de la tribu $\mathscr{A}$; autrement dit, si A et B dans $\mathscr{A}$ sont tels que $A \subset B$, alors $\mu(A) \leq \mu(B)$.*

Preuve. On écrit A comme réunion disjointe de A et $B \setminus A$ (en observant que $B \setminus A$ est encore dans $\mathscr{A}$) et on applique la propriété d'additivité à la suite finie $\{A; B \setminus A\}$ d'éléments de $\mathscr{A}$. □

Remarquons que les conventions de calculs dans $\overline{\mathbb{R}}_+$ permettent d'écrire que

$$\mu(B \setminus A) = \mu(B) - \mu(A) \quad \text{si} \quad \mu(A) < +\infty,$$

mais qu'on ne peut rien conclure de tel si A est de μ-mesure infinie.

Dans un contexte où la tribu $\mathscr{A}$ sur un ensemble Ω est fixée, on dit qu'un ensemble est *mesurable* s'il appartient à $\mathscr{A}$.

Lemme 2 (forte additivité des mesures)**.** *Si A et B sont des ensembles mesurables quelconques, alors on a :*

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B).$$

Preuve. On écrit $A \cup B$ comme réunion disjointe des trois ensembles (mesurables) $A \setminus B$, $B \setminus A$ et $A \cap B$. En appliquant l'axiome d'additivité des mesures, on obtient :

$$\mu(A \cup B) = \mu(A \setminus B) + \mu(B \setminus A) + \mu(A \cap B).$$

Il reste à ajouter aux deux membres de l'égalité le terme $\mu(A \cap B)$ et à utiliser, comme dans la preuve précédente, que $\mu(A \setminus B) + \mu(A \cap B) = \mu(A)$ et $\mu(B \setminus A) + \mu(A \cap B) = \mu(B)$. □

En théorie de la mesure la lettre σ est souvent utilisée comme référence à la dénombrabilité, comme dans le résultat qui suit. La différence avec le second axiome des mesures est qu'on ne suppose pas que les A_n sont deux à deux disjoints, mais le résultat est une inégalité.

Lemme 3 (σ-sous-additivité des mesures)**.** *Pour toute suite (au plus) dénombrable d'ensembles mesurables $(A_n)_{n \geq 0}$ on a :*

$$\mu\Big(\bigcup_{n \geq 0} A_n\Big) \leq \sum_{n \geq 0} \mu(A_n).$$

Preuve. C'est évident dès que l'un des termes A_n est de μ-mesure infinie, ce qu'on exclut désormais. On fait alors la manipulation naturelle pour produire, à partir de $(A_n)_{n \geq 0}$, une suite $(B_n)_{n \geq 0}$ satisfaisant la condition de disjonction du second axiome. Précisément, on pose $B_1 = A_1$ et $B_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k$ pour $n \geq 2$. Cela fournit :

$$\mu\Big(\bigcup_{n \geq 0} A_n\Big) = \mu\Big(\bigcup_{n \geq 0} B_n\Big) = \sum_{n \geq 0} \mu(B_n) \leq \sum_{n \geq 0} \mu(A_n).$$

Le lemme est démontré. □

Lemme 4 (suites croissantes d'ensembles mesurables)**.** *Pour toute suite croissante d'ensembles mesurables $(A_n)_{n \geq 0}$ et de limite (i.e. réunion) A, on a :*

$$\mu(A) = \lim_{n \to \infty} \mu(A_n) \leq +\infty.$$

Preuve. C'est la même manipulation. On pose $B_1 = A_1$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 2$. La suite $(B_n)_{n \geq 0}$ est formée d'ensembles mesurables et deux à deux disjoints, donc on a :

$$\lim_{n \to \infty} \mu(A_n) = \sum_{m=1}^{\infty} \mu(B_m) = \mu\Big(\bigcup_{m=1}^{\infty} B_m\Big) = \mu(A).$$

□

Nous pouvons en déduire le résultat suivant par passage au complémentaire.

Lemme 5 (suites décroissantes d'ensembles mesurables)**.** *Si $(A_n)_{n\geq 0}$ est une suite décroissante d'ensembles mesurables de limite (i.e. intersection) A, et si l'un de ses termes est de μ-mesure finie, on a :*

$$\mu(A) = \lim_{n\to\infty} \mu(A_n) < +\infty.$$

Preuve. Par hypothèse, pour un indice p, on a $\mu(A_p) < +\infty$. Le lemme précédent pour la suite croissante $(A_p \setminus A_n)_{n\geq p}$ fournit alors

$$\mu(A_p) - \mu(A) = \mu(A_p \setminus A) = \lim_{n\to\infty} \mu(A_p \setminus A_n) = \mu(A_p) - \lim_{n\to\infty} \mu(A_n),$$

et l'on peut conclure précisément car on peut simplifier par $\mu(A_p) < +\infty$. □

Remarque 3. La condition de finitude est indispensable. Pour le voir, prendre pour μ la mesure de comptage sur $\mathbb{N}$ et la suite donnée par $A_n = \{\text{entiers} \geq n\}$; dans ce cas, $A = \bigcap_n A_n = \varnothing$, et donc $\mu(A) = 0$, alors que $\mu(A_n) = +\infty$ pour tout $n \geq 0$.

4. Applications mesurables

La définition d'application mesurable est fondamentale pour la construction de l'intégrale de Lebesgue.

Définition 9. Si $(\Omega_1, \mathscr{A}_1)$ et $(\Omega_2, \mathscr{A}_2)$ sont des espaces mesurables, on dit qu'une application $f : \Omega_1 \to \Omega_2$ est mesurable si pour tout $A \in \mathscr{A}_2$ on a $f^{-1}(A) \in \mathscr{A}_1$.

4.1 Propriétés et exemples de fonctions mesurables

Les ensembles $\mathbb{R}$ et $\mathbb{R}^d$ seront toujours munis de la tribu borélienne et la mesure de référence est la mesure de Lebesgue. Voici quelques exemples ou énoncés de stabilité pour la mesurabilité des fonctions réelles de variable réelle, mais qui peuvent facilement se généraliser.

— Si $f, g : (\Omega_1, \mathscr{A}_1) \to \mathbb{R}$ sont mesurables et si $a, b \in \mathbb{R}$, alors

$$af + bg, \quad fg, \quad \max(f,g) \quad \text{et} \quad \min(\text{f},\text{g})$$

sont des fonctions mesurables. En particulier

$$f^+ := \max(f,0), \quad f^- := -\min(f,0) \quad \text{et} \quad |\text{f}|$$

sont mesurables.

— Si $f_1, \ldots, f_d$ sont des fonctions définies sur un ouvert U de $\mathbb{R}$ et mesurables, et si $\Phi : \mathbb{R}^d \to \mathbb{R}$ est continue, alors

$$\Phi\big(f_1, \ldots, f_d\big)(x) := \Phi\big(f_1(x), \ldots, f_N(x)\big)$$

est mesurable.

— Si $f(x) \neq 0$ p.p. sur U ouvert de $\mathbb{R}$, alors $x \mapsto 1/f(x)$ est mesurable.

— Toute fonction continue sur U ouvert de $\mathbb{R}$ est mesurable.

— Toute fonction continue par morceaux sur un intervalle $I \subset \mathbb{R}$ est mesurable.

— Soit $(f_n)_{n\geq 0}$ une suite de fonctions mesurables, définies sur un ouvert U de $\mathbb{R}$ et f telles que $f_n \to f$ p.p. sur U. Alors, f est mesurable.

Remarque 4. Il existe des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ non mesurables (leur construction repose toujours sur l'axiome du choix).

5. Intégration contre une mesure

5.1 Fonctions étagées

Dans cette section, on considère un espace mesuré $(\Omega, \mathscr{A}, \mu)$; autrement dit, Ω est un ensemble, $\mathscr{A}$ est une tribu de parties de Ω et $\mu : \mathscr{A} \to \overline{\mathbb{R}}_+$ est une mesure sur les ensembles mesurables (parties de Ω appartenant à $\mathscr{A}$).

Définition 10. Etant donné l'espace mesurable $(\Omega, \mathscr{A})$, on dit qu'une fonction $\Omega \to \mathbb{R}$ est *étagée* si elle est mesurable et si elle ne prend qu'un nombre fini de valeurs. On note $\mathscr{E}^+$ l'ensemble des fonctions étagées à valeurs positives.

Une fonction étagée est une fonction qui s'écrit $\sum_{i \in I} \alpha_i \mathbf{1}_{A_i}$ avec I fini et $A_i \in \mathscr{A}$ pour tout $i \in I$; cette écriture n'est pas unique.

Nous allons utiliser ces fonctions pour définir l'intégrale des fonctions μ-mesurables. Les fonctions étagées sont en effet des fonctions auxquelles on associe facilement une intégrale, et le cas général s'obtiendra en prolongeant la construction par passage à la limite.

5.2 Intégrale des fonctions étagées positives

Définition 11. Soit $f \in \mathscr{E}^+$. L'*intégrale* de f par rapport à la mesure μ est l'élément $\int f \, \mathrm{d}\mu \in \overline{\mathbb{R}}_+$ défini comme somme finie par la formule :

$$\int f \, \mathrm{d}\mu = \sum_{\alpha \text{ valeur de } f} \alpha \, \mu(\{f = \alpha\}) \leq +\infty,$$

où $\{f = \alpha\}$ désigne la partie mesurable $f^{-1}(\{\alpha\})$.

On dit que f est *intégrable* si $\int f \, \mathrm{d}\mu < +\infty$.

Pour $f \in \mathscr{E}^+$ admettant l'écriture $f = \sum_{i \in I} \alpha_i \mathbf{1}_{A_i}$ avec I fini et avec les $A_i \in \mathscr{A}$ mesurables et deux à deux disjoints, on a : $\int f \, \mathrm{d}\mu = \sum_{i \in I} \alpha_i \, \mu(A_i)$.

5.3 Propriétés de l'intégrale des fonctions étagées

L'intégrale des fonctions étagées positives possède les propriétés suivantes.

1. **Croissance :** si $f, g \in \mathscr{E}^+$ sont telles que $f \leq g$, alors
$$\int f \, \mathrm{d}\mu \leq \int g \, \mathrm{d}\mu.$$
2. **Additivité :** pour $f, g \in \mathscr{E}^+$, on a
$$\int (f+g) \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu + \int g \, \mathrm{d}\mu.$$
3. **Homogénéité :** pour $f \in \mathscr{E}^+$ et $a \geq 0$, on a
$$\int (af) \, \mathrm{d}\mu = a \int f \, \mathrm{d}\mu.$$

Justification. Les points 1 et 3 sont faciles. Pour 2, on écrit $f = \sum_{i \in I} \alpha_i \mathbf{1}_{A_i}$ avec I fini et avec les A_i mesurables partitionnant Ω ; idem pour $g = \sum_{j \in J} \beta_j \mathbf{1}_{B_j}$. Ceci permet d'obtenir $f + g = \sum_{i,j} (\alpha_i + \beta_j) \mathbf{1}_{A_i \cap B_j}$, écriture similaire aux précédentes. Alors :

$$\begin{aligned} \int (f+g) \, \mathrm{d}\mu &= \sum_{i,j} (\alpha_i + \beta_j) \, \mu(A_i \cap B_j) \\ &= \sum_i \alpha_i \sum_j \mu(A_i \cap B_j) + \sum_j \beta_j \sum_i \mu(A_i \cap B_j). \end{aligned}$$

Le premier terme vaut $\sum_i \alpha_i \, \mu(A_i) = \int f \mathrm{d}\mu$ puisque les B_j partitionnent Ω, et de même le second terme vaut $\int g \mathrm{d}\mu = \sum_j \beta_j \, \mu(B_j)$ puisque les A_i partitionnent Ω.

5.4 Approximation par des fonctions étagées

Énonçons maintenant le résultat d'approximation qui permet de prolonger la construction de l'intégrale par passage à la limite.

Théorème 1. *Les fonctions mesurables à valeurs dans $\mathbb{R}$, $\mathbb{C}$ ou $\overline{\mathbb{R}}_+$ sont les fonctions partout limites de suites de fonctions étagées. En outre :*

- *toute fonction mesurable bornée est limite uniforme d'une suite de fonctions étagées ;*
- *toute fonction mesurable positive est limite d'une suite croissante de fonctions étagées positives.*

Preuve (abrégée). Les fonctions étagées sont mesurables, donc toutes les fonctions limites de l'énoncé sont bien mesurables, comme limites simples de suites de fonctions mesurables. Étudions maintenant les réciproques.

Soit f mesurable. S'il existe M tel que $0 \leq f(x) \leq M$ pour tout $x \in X$, alors pour chaque entier $n \geq 0$ et chaque $k = 0, 1, \ldots, 2^n - 1$, on note $E_k^n = \{2^{-n}kM \leq f < 2^{-n}(k+1)M\}$. La suite des fonctions étagées $f_n = \sum_k \frac{kM}{2^n} \mathbf{1}_{E_k^n}$ est croissante et converge uniformément vers f.

Si f est positive à valeurs éventuellement infinies, pour chaque entier $n \geq 0$ on introduit $g_n = \inf\{n; f_n\}$ pour obtenir une fonction mesurable *bornée*. Par ce qui précède, on sait trouver une fonction étagée h_n satisfaisant $g_n - \frac{1}{n} \leq h_n \leq g_n$, qui fournit une suite qui converge vers f. Pour obtenir une suite croissante on pose enfin $f_n = \max\{h_1; h_2; \ldots h_n\}$.

Pour les fonctions à valeurs réelles non nécessairement positives, puis complexes, on travaille sur les parties positives et négatives, puis sur les parties réelles et imaginaires. □

5.5 Intégration des fonctions mesurables positives

On note $\mathscr{M}^+$ l'ensemble des fonctions mesurables à valeurs dans $\overline{\mathbb{R}}_+$. On vient juste de voir que les fonctions dans $\mathscr{M}^+$ sont les limites de suites croissantes de fonctions dans $\mathscr{E}^+$.

Définition 12. Soit $f \in \mathscr{M}^+$. On appelle *intégrale* de f par rapport à la mesure μ, qu'on note $\int f \, \mathrm{d}\mu$ ou $\int f(x) \, \mathrm{d}\mu(x)$, la borne supérieure, finie ou $+\infty$, de l'ensemble des intégrales des fonctions étagées positives majorées par f.

On dit que f est *intégrable* si $\int f \, \mathrm{d}\mu < +\infty$.

Dans le cas où la fonction est étagée, on retrouve la définition précédente. La définition en termes de borne supérieure permet de déduire la croissance de cette intégrale de celle des fonctions étagées.

5.6 Premiers théorèmes limites

Si $f \in \mathscr{E}^+$ et si A est mesurable, on note $\int_A f \, \mathrm{d}\mu$ l'intégrale de la fonction $f \mathbf{1}_A$ (qui est encore une fonction étagée).

Si A et B sont des parties mesurables disjointes, l'additivité de l'intégrale assure que

$$\int_{A \cup B} f \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu + \int_B f \, \mathrm{d}\mu.$$

Voici un premier résultat de passage à la limite monotone.

Lemme 6 (limite monotone des fonctions étagées)**.** *Soit $f = \sum_{i \in I} \alpha_i \mathbf{1}_{A_i}$ une fonction étagée positive et soit $\{E_n\}_{n \geq 0}$ une suite croissante de parties mesurables dont la limite (i.e. la réunion) vaut Ω tout entier. Alors on a*

$$\int f \, \mathrm{d}\mu = \lim_{n \to \infty} \int_{E_n} f \, \mathrm{d}\mu.$$

Preuve. On se ramène au cas où les A_i dans l'écriture de f sont deux à deux disjoints ; alors :

$$\int_{E_n} f \,\mathrm{d}\mu = \int f \mathbf{1}_{E_n} \,\mathrm{d}\mu = \sum_i \alpha_i \,\mu(A_i \cap E_n).$$

Pour chaque $i \in I$, la suite croissante de parties mesurables $\{E_n \cap A_i\}_{n\geq 0}$ tend vers A_i (i.e. est de réunion égale à A_i), donc par propriété des mesures, pour chaque $i \in I$, on a :

$$\mu(A_i) = \lim_{n\to\infty} \mu(E_n \cap A_i).$$

Ceci permet finalement d'écrire :

$$\int f \,\mathrm{d}\mu \;=\; \sum_i \alpha_i \,\mu(A_i) \;=\; \lim_{n\to\infty} \sum_i \alpha_i \,\mu(E_n \cap A_i) \;=\; \lim_{n\to\infty} \int_{E_n} f \,\mathrm{d}\mu,$$

l'interversion centrale entre limite et somme ne posant pas de problème puisque I est fini. □

Nous en déduisons le théorème de convergence croissante de Beppo Levi.

Théorème 2 (de convergence croissante)**.** *Soit $(f_n)_{n\geq 0}$ une suite croissante de fonctions dans $\mathscr{M}^+$ convergeant ponctuellement vers f. Alors, on a :*

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

Preuve. La croissance de l'intégrale et les inégalités $f_n \leq f$ pour tout $n \geq 0$ fournissent $\int f_n \,\mathrm{d}\mu \leq \int f \,\mathrm{d}\mu$, et donc $\lim_{n\to\infty} \int f_n \,\mathrm{d}\mu \leq \int f \,\mathrm{d}\mu$ par passage à la limite.

Pour l'inégalité inverse, par définition de l'intégrale sur $\mathscr{M}^+$, il suffit de voir que pour toute $\varphi \in \mathscr{E}^+$ telle que $\varphi \leq f$ on a :

$$\int \varphi \,\mathrm{d}\mu \;\leq\; \lim_{n\to\infty} \int f_n \,\mathrm{d}\mu.$$

On conclut alors en passant à la borne supérieure des $\int \varphi \,\mathrm{d}\mu$ pour $\varphi \in \mathscr{E}^+$ telle que $\varphi \leq f$.

Pour prouver cette dernière inégalité, on se donne $a \in]0,1[$ et pour chaque $n \geq 0$ on pose $E_n = \{f_n \geq a\varphi\}$. Par définition des E_n on a ainsi $f_n \geq (a\varphi)\mathbf{1}_{E_n}$ pour tout $n \geq 0$. Ceci implique

$$\int f_n \,\mathrm{d}\mu \;\geq\; \int_{E_n} (a\,\varphi) \,\mathrm{d}\mu \;=\; a \int_{E_n} \varphi \,\mathrm{d}\mu.$$

En outre la suite de parties mesurables $\{E_n\}_{n\geq 0}$ est croissante pour l'inclusion puisque la suite de fonctions $(f_n)_{n\geq 0}$ est croissante, et par hypothèse de convergence ponctuelle (et puisque $a < 1$), les E_n tendent vers Ω tout entier. Ainsi le lemme précédent implique que

$$\lim_{n\to\infty} \int_{E_n} \varphi \,\mathrm{d}\mu = \int \varphi \,\mathrm{d}\mu,$$

ce qui permet de passer à la limite dans l'inégalité précédente pour obtenir

$$\lim_{n\to\infty} \int f_n \,\mathrm{d}\mu \;\geq\; a \int \varphi \,\mathrm{d}\mu,$$

et finalement l'inégalité cherchée puisque a est arbitraire dans $]0,1[$. □

5.7 Propriétés de l'intégrale

1. **Croissance :** si $f, g \in \mathscr{M}^+$ sont telles que $f \leq g$, alors

$$\int f \,\mathrm{d}\mu \leq \int g \,\mathrm{d}\mu.$$

2. **Additivité :** pour $f,g \in \mathscr{M}^+$, on a
$$\int (f+g)\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu.$$
3. **Homogénéité :** pour $f \in \mathscr{M}^+$ et $a \geq 0$, on a
$$\int (af)\,\mathrm{d}\mu = a \int f\,\mathrm{d}\mu.$$

Justification. Le premier point a été vu juste après la définition de l'intégrale sur $\mathscr{M}^+$ et le troisième est immédiat. Pour le deuxième, on sait qu'il existe des suites croissantes de fonctions dans $\mathscr{E}^+$, disons $(f_n)_{n\geq 0}$ et $(g_n)_{n\geq 0}$, qui convergent vers f et g respectivement (théorème d'approximation). Alors $(f_n+g_n)_{n\geq 0}$ est une suite croissante dans $\mathscr{E}^+$ qui converge vers $f+g$, et on a
$$\int (f+g)\,\mathrm{d}\mu = \lim_{n\to\infty} \int (f_n+g_n)\,\mathrm{d}\mu = \lim_{n\to\infty} \int f_n\,\mathrm{d}\mu + \lim_{n\to\infty} \int g_n\,\mathrm{d}\mu = \int f\,\mathrm{d}\mu + \int g\,\mathrm{d}\mu,$$
le théorème de Beppo Levi assurant les deux égalités aux extrémités de la chaîne ; l'égalité du milieu provient de l'additivité de l'intégrale sur $\mathscr{E}^+$. □

Proposition 2. *Soit f une fonction dans $\mathscr{M}^+$. Alors,*

1. *L'intégrale $\int f\,\mathrm{d}\mu$ est nulle si, et seulement si, f est nulle μ-presque partout.*
2. *Si f est intégrable, alors l'ensemble $\{f = +\infty\}$ est négligeable.*

Preuve. Le premier point est vrai pour les fonctions étagées positives. Dans le cas général, la condition est suffisante par définition de l'intégrale comme borne supérieure. Réciproquement, soit $f \in \mathscr{M}^+$ d'intégrale nulle. On se donne une suite croissante dans $\mathscr{E}^+$, disons $(f_n)_{n\geq 0}$, qui converge vers f. Alors pour tout $n \geq 0$, on a $\int f_n\,\mathrm{d}\mu = 0$ et donc $f_n = 0$ μ-p.p. puisque f_n est étagée. Puisque $f(x) = \lim_{n\to\infty} f_n(x)$, on en conclut que f est nulle μ-presque partout.

Pour le deuxième point, il suffit de considérer une fonction $f \in \mathscr{M}^+$ telle que l'ensemble $E = \{f = +\infty\}$ satisfasse $\mu(E) > 0$. On pose alors $f_n = n \cdot \mathbf{1}_E \leq f$, pour $n \geq 0$. On a $\int f\,\mathrm{d}\mu \geq \int f_n\,\mathrm{d}\mu = n \cdot \mu(E)$ et on conclut en passant à la limite $n \to +\infty$.

5.8 Intégration des fonctions à valeurs réelles et complexes

Les parties positive $f^+ = \max\{f;0\}$ et négative $f^- = \max\{-f;0\}$ sont définies pour avoir $f = f^+ - f^-$ et $|f| = f^+ + f^-$. Pour $f : X \to \mathbb{R}$ mesurable f^+ et f^- sont dans $\mathscr{M}^+$.

Définition 13. Soit $f : X \to \mathbb{R}$ mesurable. On dit que f est *intégrable* par rapport à μ si f^+ et f^- le sont au sens qui précède. On pose alors
$$\int f\,\mathrm{d}\mu = \int f^+\,\mathrm{d}\mu - \int f^-\,\mathrm{d}\mu.$$
On note $\mathscr{L}^1(\mu,\mathbb{R})$ ou $\mathscr{L}^1(\mu)$ *l'ensemble des fonctions intégrables à valeurs dans* $\mathbb{R}$.

Pour $f : X \to \mathbb{R}$ mesurable, on observe : $f \in \mathscr{L}^1(\mu)$ si et seulement si $|f| \in \mathscr{L}^1(\mu)$.

Définition 14. Soit $f : X \to \mathbb{C}$ mesurable. On dit que f est *intégrable* par rapport à μ si $\mathrm{Re}f$ et $\mathrm{Im}f$ sont intégrables au sens ci-dessus. Si tel est le cas, on définit :
$$\int f\,\mathrm{d}\mu = \int \mathrm{Re}f\,\mathrm{d}\mu + i \int \mathrm{Im}f\,\mathrm{d}\mu.$$
On note $\mathscr{L}^1(\mu,\mathbb{C})$ *l'ensemble des fonctions intégrables à valeurs dans* $\mathbb{C}$.

5.9 Propriétés de l'intégration des fonctions à valeurs réelles

- L'ensemble $\mathscr{L}^1(\mu)$ est un $\mathbb{R}$-espace vectoriel.
- L'application $f \mapsto \int f\,\mathrm{d}\mu$ est une forme $\mathbb{R}$-linéaire sur $\mathscr{L}^1(\mu)$.
- Pour toute $f \in \mathscr{L}^1(\mu)$, on a $|f| \in \mathscr{L}^1(\mu)$ et

$$\left|\int f\,\mathrm{d}\mu\right| \leq \int |f|\,\mathrm{d}\mu.$$

5.10 Intégrale sur un intervalle par rapport à la mesure de Lebesgue

Soit f une fonction continue sur un intervalle $[a,b]$ et soit F une primitive de f. On a alors

$$\int_{[a,b]} f\,\mathrm{d}\lambda_1 = F(b) - F(a).$$

De plus, toute fonction continue par morceaux (ou plus généralement intégrable au sens de Riemann) sur un intervalle borné $[a,b]$, est intégrable au sens de Lebesgue et son intégrale au sens de Lebesgue coïncide avec son intégrale au sens de Riemmann. On utilisera indifféremment la notation $\int f(x)\,\mathrm{d}x$ ou $\int f\,\mathrm{d}\lambda_1$.

6. Théorèmes de convergence

6.1 Théorème de convergence monotone de Beppo Levi, version série

On considère un espace mesuré $(\Omega, \mathscr{A}, \mu)$. Toutes les fonctions considérées ici sont définies sur cet espace.

Théorème 3 (convergence monotone, version série)**.** *Soit $(u_n)_{n\geq 0}$ une suite de fonctions μ-mesurables à valeurs dans $\overline{\mathbb{R}}_+$. Alors on a*

$$\int \Big(\sum_{n\geq 0} u_n\Big)\,\mathrm{d}\mu \;=\; \sum_{n\geq 0} \Big(\int u_n\,\mathrm{d}\mu\Big) \;\leq\; +\infty.$$

Ce théorème se déduit immédiatement du Théorème 2, via les sommes partielles.

Proposition 3. *Soit $(f_n)_{n\geq 0}$ une suite croissante de fonctions μ-intégrables réelles. Alors on a*

$$\lim_{n\to\infty} \int f_n\,\mathrm{d}\mu \;=\; \int \big(\lim_{n\to\infty} f_n\big)\,\mathrm{d}\mu \;\leq\; +\infty.$$

Preuve. Par croissance, la suite $(f_n)_{n\geq 0}$ converge ponctuellement vers une fonction $f : X \to \overline{\mathbb{R}}$ dont la partie négative f^- est intégrable (car $(f_n^-)_n$ est décroissante). Donc $\int f\mathrm{d}\mu = \int f^+\mathrm{d}\mu - \int f^-\mathrm{d}\mu$ est bien définie μ-presque partout. On se ramène au théorème de Beppo Levi en considérant la suite des fonctions $f_n - f_1$, qui sont elles aussi bien définies μ-presque partout car f_1 est intégrable et donc ne vaut $\pm\infty$ que sur un ensemble négligeable. □

L'hypothèse de monotonie est cruciale, comme l'indique l'exemple qui suit.

Exemple : On considère la suite de fonctions $f_n(x) = 2nx(1-x^2)^{n-1}$ définies sur $]0,1[$. On vérifie que $\lim_{n\to+\infty} f_n(x) = 0$ pour tout $x \in]0,1[$ et que $\int_0^1 f_n(x)\mathrm{d}x = 1$ pour tout $n \geq 1$. Donc :

$$0 = \int_0^1 \lim_{n\to+\infty} f_n(x)\,\mathrm{d}x < \lim_{n\to+\infty} \int_0^1 f_n(x)\,\mathrm{d}x = 1.$$

L'inégalité qui subsiste est expliquée par le lemme suivant, très utile en pratique.

6.2 Lemme de Fatou

Le résultat suivant fournit une inégalité, et non pas exactement une interversion limite-somme. Il peut être utile notamment pour prouver des divergences de suites numériques définies par des intégrales. Dans beaucoup de cas pratiques, la suite de fonctions converge et la première limite inférieure (i.e. la fonction intégrée) est alors une vraie fonction limite (ponctuelle).

Lemme 7 (lemme de Fatou)**.** *Soit $(f_n)_{n\geq 0}$ une suite de fonctions μ-mesurables positives. Alors,*

$$\int \liminf_{n\to+\infty} f_n \,\mathrm{d}\mu \;\leq\; \liminf_{n\to+\infty} \int f_n \,\mathrm{d}\mu.$$

Preuve. La suite de fonctions mesurables $(g_n = \inf_{k\geq n} f_k)_{n\geq 0}$ est croissante et converge vers $\liminf_{n\to+\infty} f_n$. Par le théorème de Beppo Levi, on a donc : $\int \liminf_{n\to+\infty} f_n \,\mathrm{d}\mu = \int \lim_{n\to\infty} g_n \,\mathrm{d}\mu = \lim_{n\to\infty} \int g_n \,\mathrm{d}\mu$. Enfin, le fait que $g_n \leq f_n$ implique que $\lim_{n\to\infty} \int g_n \,\mathrm{d}\mu \leq \liminf_{n\to+\infty} \int f_n \,\mathrm{d}\mu$; d'où le résultat. □

6.3 Théorème de la convergence dominée de Lebesgue

Théorème 4 (théorème de convergence dominée)**.** *Soit $(f_n)_{n\geq 0}$ une suite de fonctions de $\mathscr{L}^1(\mu,\mathbb{C})$. On suppose que :*

- *la suite $\left(f_n\right)_{n\geq 0}$ converge vers une fonction f μ-presque partout sur Ω ;*
- *il existe $g \in \mathscr{L}^1(\mu,\mathbb{R})$ telle que $|f_n| \leq g$ μ-presque partout sur Ω.*

Alors, $f \in \mathscr{L}^1(\mu,\mathbb{C})$ et

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

Il s'agit d'un énoncé très important, s'appliquant à une *vaste classe de fonctions*, et utile dans de nombreux calculs de limite impliquant une opération d'intégration.

Parmi les applications théoriques importantes de ce théorème, on compte des résultats généraux de continuité et de dérivation sous le signe somme.

B. Probabilité et espérance

Nous introduisons dans cette partie des notions analogues en utilisant le vocabulaire de la théorie de la mesure en probabilité.

L'ensemble Ω est appelé *espace fondamental* ou *espace d'états*. On note $\omega \in \Omega$ un élément de Ω, aussi appelé *résultat possible* ou *expérience*. L'ensemble Ω est muni d'une tribu $\mathscr{A}$. Ainsi $(\Omega,\mathscr{A})$ est un espace mesurable (ou *espace probabilisable*). Les éléments de la tribu $\mathscr{A}$ sont appelés *événements*.

7. Notion de probabilité

On munit l'ensemble mesurable $(\Omega,\mathscr{A})$ d'une mesure de masse totale 1. Une telle mesure est appelée mesure de probabilité ou plus simplement *probabilité*.

Définition 15. Soit $(\Omega,\mathscr{A})$ un espace mesurable. On appelle *probabilité* sur $(\Omega,\mathscr{A})$ une application $\mathbb{P} : \mathscr{A} \to [0,1]$ satisfaisant les axiomes suivants.

- $\mathbb{P}(\Omega) = 1$;
- Si $(A_n)_{n\geq 0}$ est une suite (finie ou dénombrable) d'éléments de $\mathscr{A}$ deux à deux disjoints :

$$\mathbb{P}\big(\bigcup_{n\geq 0} A_n\big) = \sum_{n\geq 0} \mathbb{P}(A_n).$$

On appelle *espace de probabilité* le triplet $(\Omega,\mathscr{A},\mathbb{P})$.

8. Notion de variable aléatoire

Définition 16. On appelle *variable aléatoire réelle* une application mesurable d'un espace mesurable $(\Omega, \mathscr{A})$ dans $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$.

On note en général les variables aléatoires par une lettre majuscule (par exemple, X ou Y). Les valeurs prises par la variable aléatoire, appelées *réalisations* de la variable aléatoire, sont notées en lettres minuscules. On écrira par exemple $X(\omega) = x$ pour $\omega \in \Omega$ et x une réalisation de X.

9. Espérance des variables aléatoires

9.1 Variables aléatoires étagées

Dans cette section, on travaille avec un espace de probabilité $(\Omega, \mathscr{A}, \mathbb{P})$; autrement dit, Ω est un ensemble, $\mathscr{A}$ est une tribu de parties de Ω et $\mathbb{P} : \mathscr{A} \to [0,1]$ est une mesure de masse 1 sur les parties de Ω appartenant à la tribu $\mathscr{A}$.

Définition 17. On dit qu'une variable aléatoire réelle définie sur l'espace mesurable $(\Omega, \mathscr{A})$ est *étagée* si elle ne prend qu'un nombre fini de valeurs. On note $\mathscr{E}^+$ l'ensemble des variables aléatoires étagées positives.

Une variable aléatoire étagée s'écrit $\sum_{i\in I} \alpha_i \mathbf{1}_{A_i}$ avec I fini et $A_i \in \mathscr{A}$ pour tout $i \in I$; cette écriture n'est pas unique.

Nous allons utiliser ces variables aléatoires pour définir l'espérance des variables aléatoires. Les variables étagées sont en effet des variables aléatoires auxquelles on associe facilement une espérance, et le cas général s'obtiendra en prolongeant la construction par passage à la limite.

9.2 Espérance des variables aléatoires étagées positives

Définition 18. Soit $X \in \mathscr{E}^+$. L'*espérance* de X par rapport à la probabilité $\mathbb{P}$ est l'élément $\mathbb{E}(X) \in \overline{\mathbb{R}}_+$ défini comme somme finie par la formule :

$$\mathbb{E}(X) = \sum_{\alpha \text{ valeur de } X} \alpha \, \mathbb{P}(\{X = \alpha\}) \leq +\infty,$$

où $\{X = \alpha\}$ désigne la partie mesurable $X^{-1}(\{\alpha\})$. On note aussi $\int X \, \mathrm{d}\mathbb{P}$ cette quantité.

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

Pour une écriture de $X \in \mathscr{E}^+$ sous la forme $X = \sum_{i\in I} \alpha_i \mathbf{1}_{A_i}$ avec I fini et les $A_i \in \mathscr{A}$ deux à deux disjoints, on a : $\mathbb{E}(X) = \sum_{i\in I} \alpha_i \mathbb{P}(A_i)$.

9.3 Propriétés de l'espérance des variables aléatoires étagées

L'espérance des variables aléatoires étagées positives possède les propriétés suivantes.

1. **Croissance :** si $X, Y \in \mathscr{E}^+$ sont telles que $X \leq Y$, alors
$$\mathbb{E}(X) \leq \mathbb{E}(Y).$$
2. **Additivité :** pour $X, Y \in \mathscr{E}^+$, on a
$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$
3. **Homogénéité :** pour $X \in \mathscr{E}^+$ et $a \geq 0$, on a
$$\mathbb{E}(aX) = a\,\mathbb{E}(X).$$

Justification. Les points 1 et 3 sont faciles. Pour 2, on écrit $X = \sum_{i\in I} \alpha_i \mathbf{1}_{A_i}$ avec I fini et les $A_i \in \mathscr{A}$ partitionnant Ω ; idem, pour $Y = \sum_{j\in J} \beta_j \mathbf{1}_{B_j}$. Ceci permet d'obtenir $X+Y = \sum_{i,j}(\alpha_i+\beta_j)\,\mathbf{1}_{A_i\cap B_j}$, écriture similaire aux précédentes. Alors :

$$\mathbb{E}(X+Y) = \sum_{i,j}(\alpha_i+\beta_j)\,\mathbb{P}(A_i\cap B_j) = \sum_i \alpha_i \sum_j \mathbb{P}(A_i\cap B_j) + \sum_j \beta_j \sum_i \mathbb{P}(A_i\cap B_j).$$

Le premier terme vaut $\sum_i \alpha_i \mathbb{P}(A_i) = \mathbb{E}(X)$ puisque les B_j partitionnent Ω et de même le second terme vaut $\mathbb{E}(Y) = \sum_j \beta_j\,\mu(B_j)$ puisque les A_i partitionnent Ω.

9.4 Approximation par des variables aléatoires étagées

Énonçons maintenant le résultat d'approximation qui permet de prolonger la construction de l'espérance par passage à la limite.

Théorème 5. *Les variables aléatoires à valeurs dans $\mathbb{R}$ (ou $\overline{\mathbb{R}}_+$) sont les variables aléatoires partout limites de suites de valeurs aléatoires étagées. En outre :*

— *toute variable aléatoire bornée est limite uniforme d'une suite de variables aléatoires étagées ;*

— *toute variable aléatoire positive est limite d'une suite croissante de variables aléatoires étagées positives.*

Preuve (abrégée). Toutes les limites de l'énoncé sont bien des variables aléatoires, comme limites simples de suites de variables aléatoires. Etudions maintenant les réciproques.

Soit X une variable aléatoire. S'il existe M tel que $0 \le X(\omega) \le M$ pour tout $\omega\in\Omega$, alors pour chaque entier $n\ge 0$ et chaque $k = 0,1,\ldots,2^n-1$, on note $E_k^n = \{2^{-n}kM \le X < 2^{-n}(k+1)M\}$. La suite des variables aléatoires étagées $X_n = \sum_k \frac{kM}{2^n}\mathbf{1}_{E_k^n}$ est croissante et converge uniformément vers la variable aléatoire X.

Si X est positive à valeurs éventuellement infinies, pour chaque entier $n\ge 0$ on introduit $Y_n = \inf\{n;X\}$ pour obtenir une variable aléatoire *bornée*. Par ce qui précède, on sait trouver une variable aléatoire étagée Z_n satisfaisant $Y_n - \frac{1}{n} \le Z_n \le Y_n$, qui fournit une suite qui converge vers X. Pour obtenir une suite croissante on pose enfin $X_n = \max\{Z_1, Z_2,,\ldots Z_n\}$. $\square$

9.5 Espérance des variables aléatoires positives

On note $\mathscr{M}^+$ l'ensemble des variables aléatoires à valeurs dans $\overline{\mathbb{R}}_+$. Les variables aléatoires dans $\mathscr{M}^+$ sont les limites de suites croissantes de variables aléatoires étagées positives.

Définition 19. Soit $X\in\mathscr{M}^+$. On appelle *espérance* de X par rapport à la probabilité $\mathbb{P}$, notée $\mathbb{E}(X)$ ou $\int X\,\mathrm{d}\mathbb{P}$, la borne supérieure, finie ou $+\infty$, de l'ensemble des intégrales des variables aléatoires étagées positives majorées par X. On dit que X est *intégrable* si $\mathbb{E}(X) < +\infty$.

Dans le cas où la variable aléatoire est étagée, on retrouve la définition précédente. La définition en termes de borne supérieure permet de déduire la croissance de cette espérance de celle des variables aléatoires étagées.

9.6 Premiers théorèmes limites

Si $X\in\mathscr{E}^+$ et si $A\in\mathscr{A}$, alors le produit $X\,\mathbf{1}_A$ est encore une variable aléatoire étagée et on a $\mathbb{E}(X\,\mathbf{1}_A) = \int_A X\,\mathrm{d}\mathbb{P}$.

Si A et B sont des parties mesurables *disjointes*, l'additivité de l'espérance assure que

$$\mathbb{E}(X\,\mathbf{1}_{A\cup B}) = \mathbb{E}(X\,\mathbf{1}_A) + \mathbb{E}(X\,\mathbf{1}_B).$$

Lemme 8 (limite monotone des variables aléatoires étagées)**.** *Soit* $X = \sum_{i\in I}\alpha_i \mathbf{1}_{A_i}$ *une variable aléatoire étagée positive et soit* $\{E_n\}_{n\geq 0}$ *une suite croissante de parties mesurables dont la limite (i.e. la réunion) vaut* Ω *tout entier. Alors on a*

$$\mathbb{E}(X) = \lim_{n\to\infty} \mathbb{E}(X\,\mathbf{1}_{E_n}).$$

Preuve. On se ramène au cas où les A_i sont deux à deux disjoints ; alors :

$$\mathbb{E}(X\,\mathbf{1}_{E_n}) = \sum_i \alpha_i\,\mathbb{P}(A_i \cap E_n).$$

Pour chaque $i \in I$, la suite croissante de parties mesurables $\{E_n \cap A_i\}_{n\geq 0}$ tend vers A_i (i.e. est de réunion égale à A_i), donc

$$\mathbb{P}(A_i) = \lim_{n\to\infty} \mathbb{P}(E_n \cap A_i).$$

Ceci permet finalement d'écrire :

$$\mathbb{E}(X) = \sum_i \alpha_i\,\mathbb{P}(A_i) = \lim_{n\to\infty}\sum_i \alpha_i\,\mathbb{P}(E_n\cap A_i) = \lim_{n\to\infty}\mathbb{E}(X\,\mathbf{1}_{E_n}),$$

l'interversion entre limite et somme ne posant pas de problème puisque I est fini. □

Nous en déduisons le théorème de convergence croissante de Beppo Levi.

Théorème 6 (de convergence croissante)**.** *Soit* $(X_n)_{n\geq 0}$ *une suite croissante de variables aléatoires dans* $\mathscr{M}^+$ *convergeant presque-sûrement vers* X*. Alors, on a :*

$$\mathbb{E}(X) = \lim_{n\to\infty}\mathbb{E}(X_n) \leq +\infty.$$

Preuve. La croissance de l'espérance et les inégalités $X_n \leq X$ pour tout $n \geq 0$ fournissent $\mathbb{E}(X_n) \leq \mathbb{E}(X)$, et donc $\lim_{n\to\infty}\mathbb{E}(X_n) \leq \mathbb{E}(X)$ par passage à la limite.

Pour l'inégalité inverse, par définition de l'intégrale sur $\mathscr{M}^+$, il suffit de voir que pour toute $Y \in \mathscr{E}^+$ telle que $Y \leq X$ on a :

$$\mathbb{E}(Y) \leq \lim_{n\to\infty}\mathbb{E}(X_n).$$

On conclut alors en passant à la borne supérieure des $\mathbb{E}(Y)$ pour $Y \in \mathscr{E}^+$ telle que $Y \leq X$.

Pour prouver cette dernière inégalité, on se donne $a \in]0,1[$ et pour chaque $n \geq 0$ on pose $E_n = \{X_n \geq aY\}$. Par définition des E_n on a ainsi $X_n \geq (aY)\,\mathbf{1}_{E_n}$ pour tout $n \geq 0$. Ceci implique

$$\mathbb{E}(X_n) \geq \mathbb{E}(\mathbf{1}_{E_n} aY), = a\mathbb{E}(\mathbf{1}_{E_n} Y).$$

En outre la suite des évenements $(E_n)_{n\geq 0}$ est croissante pour l'inclusion puisque la suite de variables aléatoires $(X_n)_{n\geq 0}$ est croissante, et par hypothèse de convergence ponctuelle (et puisque $a < 1$), les E_n tendent vers Ω tout entier. Ainsi le lemme précédent implique que

$$\lim_{n\to\infty}\mathbb{E}(\mathbf{1}_{E_n} Y) = \mathbb{E}(Y),$$

ce qui permet de passer à la limite dans l'inégalité précédente pour obtenir

$$\lim_{n\to\infty}\mathbb{E}(X_n) \geq a\mathbb{E}(Y),$$

et finalement l'inégalité cherchée puisque a est arbitraire dans $]0,1[$. □

9.7 Propriétés de l'espérance

1. **Croissance :** si $X,Y \in \mathscr{M}^+$ sont telles que $X \le Y$, alors
$$\mathbb{E}(X) \le \mathbb{E}(Y).$$
2. **Additivité :** pour $X,Y \in \mathscr{M}^+$, on a
$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$
3. **Homogénéité :** pour $X \in \mathscr{M}^+$ et $a \ge 0$, on a
$$\mathbb{E}(aX) = a\mathbb{E}(X).$$

Justification. Le premier point a été vu juste après la définition de l'espérance sur $\mathscr{M}^+$ et le troisième est immédiat. Pour le deuxième, on sait qu'il existe des suites croissantes de variables aléatoires dans $\mathscr{E}^+$, disons $(X_n)_{n\ge 0}$ et $(Y_n)_{n\ge 0}$, qui convergent vers X et Y respectivement (théorème d'approximation). Alors $(X_n + Y_n)_{n\ge 0}$ est une suite croissante dans $\mathscr{E}^+$ qui converge vers $X+Y$, et on a

$$\mathbb{E}(X+Y) = \lim_{n\to\infty} \mathbb{E}(X_n+Y_n) = \lim_{n\to\infty} \mathbb{E}(X_n) + \lim_{n\to\infty} \mathbb{E}(Y_n) = \mathbb{E}(X) + \mathbb{E}(Y),$$

le théorème de Beppo Levi assurant les deux égalités aux extrémités de la chaîne ; l'égalité du milieu provient de l'additivité de l'intégrale sur $\mathscr{E}^+$. □

Proposition 4. *Soit X une variable aléatoire dans $\mathscr{M}^+$. L'espérance $\mathbb{E}(X)$ est nulle si, et seulement si, X est nulle $\mathbb{P}$-presque sûrement, c'est-à-dire si $\mathbb{P}(X=0)=1$.*

Preuve. C'est vrai pour les variables aléatoires étagées positives. Dans le cas général, la condition est suffisante par définition de l'espérance comme borne supérieure. Réciproquement, soit $X \in \mathscr{M}^+$ d'espérance nulle. On se donne une suite croissante $(X_n)_{n\ge 0}$ dans $\mathscr{E}^+$ qui converge vers X. Alors pour tout $n \ge 0$, on a $\mathbb{E}(X_n) = 0$ et donc $X_n = 0$ $\mathbb{P}$-p.s. puisque X_n est étagée. Puisque $X(\omega) = \lim_{n\to\infty} X_n(\omega)$, on en conclut que X est nulle $\mathbb{P}$-presque sûrement.

9.8 Espérance des variables aléatoires à valeurs réelles

Pour $X : \Omega \to \mathbb{R}$ une variable aléatoire, X^+ et X^- sont dans $\mathscr{M}^+$.

Définition 20. Soit $X : \Omega \to \mathbb{R}$ variable aléatoire. On dit que X est *intégrable* par rapport à $\mathbb{P}$ si X^+ et X^- le sont au sens qui précède. On pose alors

$$\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-).$$

On note $\mathscr{L}^1$ l'ensemble des variables aléatoires intégrables à valeurs réelles.

9.9 Propriétés de l'intégration des variables aléatoires à valeurs réelles

— L'ensemble $\mathscr{L}^1$ est un $\mathbb{R}$-espace vectoriel.
— L'application $X \mapsto \mathbb{E}(X)$ est une forme $\mathbb{R}$-linéaire sur $\mathscr{L}^1$.
— Pour toute $X \in \mathscr{L}^1$, on a $|X| \in \mathscr{L}^1$ et
$$|\mathbb{E}(X)| \le \mathbb{E}(|X|).$$

9.10 Inégalité de Markov

Proposition 5 (inégalité de Markov)**.** *Soit X une variable aléatoire dans $\mathscr{M}^+$ et soit α un nombre réel strictement positif. Alors*

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

Preuve. On note A l'événement $\{X \geq \alpha\} = \{\omega \in \Omega : X(\omega) \geq \alpha\}$. On a $X \geq \alpha \cdot \mathbf{1}_A$ et on conclut en prenant l'espérance. □

10. Théorèmes de convergence

10.1 Théorème de la convergence monotone de Beppo Levi, version série

Théorème 7 (convergence monotone, version série)**.** *Soit $(U_n)_{n\geq 0}$ une suite de variables aléatoires à valeurs dans $\overline{\mathbb{R}}_+$. Alors on a*

$$\mathbb{E}\big(\sum_{n\geq 0} U_n\big) \;=\; \sum_{n\geq 0} \big(\mathbb{E}(U_n)\big) \;\leq\; +\infty.$$

Ce théorème se déduit immédiatement du Théorème 2, via les sommes partielles.

Proposition 6. *Soit $(X_n)_{n\geq 0}$ une suite croissante de variables aléatoires réelles intégrables. Alors on a*

$$\lim_{n\to\infty} \mathbb{E}(X_n) = \mathbb{E}(\lim_{n\to\infty} X_n) \;\leq\; +\infty.$$

Preuve. Par croissance, la suite $(X_n)_{n\geq 0}$ converge ponctuellement vers une variable aléatoire $X : \Omega \to \overline{\mathbb{R}}$ dont la partie négative X^- est intégrable (car $(X_n^-)_n$ est décroissante). Donc $\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-)$ est bien définie. On se ramène au théorème de Beppo Levi en considérant la suite des fonctions $X_n - X_1$, qui sont elles aussi bien définies car X_1 est intégrable. □

10.2 Lemme de Fatou

Lemme 9 (lemme de Fatou)**.** *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires positives. Alors on a :*

$$\mathbb{E}\big(\liminf_{n\to+\infty} X_n\big) \leq \liminf_{n\to+\infty} \mathbb{E}\big(X_n\big).$$

Preuve. La suite de fonctions mesurables $(Y_n = \inf_{k\geq n} X_k)_{n\geq 0}$ est croissante et converge vers $\liminf_{n\to+\infty} X_n$.

Par le théorème de Beppo Levi, on a donc : $\mathbb{E}(\liminf_{n\to+\infty} X_n) = \mathbb{E}(\lim_{n\to\infty} Y_n) = \lim_{n\to\infty} \mathbb{E}(Y_n)$. Enfin, le fait que $Y_n \leq X_n$ implique que $\lim_{n\to\infty} \mathbb{E}(Y_n) \leq \liminf_{n\to+\infty} \mathbb{E}(X_n)$; d'où le résultat. □

10.3 Théorème de la convergence dominée de Lebesgue

Théorème 8 (théorème de convergence dominée)**.** *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires intégrables. On suppose que :*

- *La suite $\big(X_n\big)_{n\geq 0}$ converge presque-sûrement vers X ;*
- *Il existe Y intégrable telle que $|X_n| \leq Y$ presque-sûrement.*

Alors, X est intégrable et

$$\lim_{n\to+\infty} \mathbb{E}(X_n) = \mathbb{E}(X).$$

Il s'agit d'un énoncé très important et utile dans de nombreux calculs de limite impliquant une espérance.