Introduction aux probabilités

Modèles et applications

Quentin Berger
Maître de conférences à Sorbonne Université.

Francesco Caravenna
Professeur à l'université Milan-Bicocca.

Paolo Dai Pra
Professeur à l'université de Vérone.

DUNOD

Illustration de couverture : ©Constantinos-Adobe Stock.com

First published in Italian under the title
Probabilità; Un primo corso attraverso esempi, modelli e applicazioni
by Quentin Berger, Francesco Caravenna and Paolo Dai Pra, edition: 2

Traduction : Quentin Berger

11 rue Paul Bert, 92240 Malakoff
www.dunod.com
ISBN 978-2-10-081486-2

Préface

L'objectif de ce livre est de fournir une introduction à la théorie des probabilités et à ses applications, sans avoir recours à la théorie de la mesure. Il s'adresse à des étudiants de licence en filière scientifique, par exemple en mathématique, en physique ou en école d'ingénieur. Le choix des thèmes et l'approche que nous avons adoptés sont le fruit de plusieurs années d'expérience d'enseignement de cours que nous avons donnés en licence de mathématique, dans les universités de Padoue, Milan-Bicocca, Vérone et Sorbonne Université.

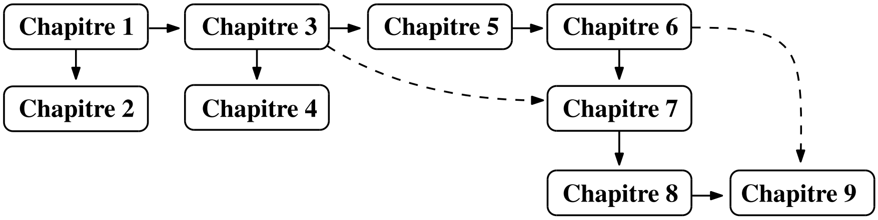
Ce qui nous a guidé pour la rédaction de cet ouvrage est d'une part notre intention de présenter les notions fondamentales de façon rigoureuse, d'autre part notre volonté d'introduire le plus tôt possible des exemples et applications intéressantes, qui motivent les développements de la théorie. Pour ces raisons, nous avons décidé d'insister particulièrement sur les probabilités *discrètes*, c'est-à-dire sur des espaces finis ou dénombrables : les quatre premiers chapitres y sont consacrés. Dans ce contexte, en effet, il suffit de très peu d'outils d'analyse pour présenter la théorie de manière complète et rigoureuse (seules les suites et les séries sont nécessaires). Cela permet en particulier d'introduire le langage et les notions de base des probabilités sans trop de complications techniques, et de se concentrer sur les principales difficultés conceptuelles rencontrées par les étudiants quand ils commencent à étudier ce sujet.

Le Chapitre 1 est dédié aux espaces probabilisés discrets et le Chapitre 3 présente les variables aléatoires discrètes. Notons que quelques notions de probabilités discrètes suffisent pour aborder des problèmes et des modèles extrêmement intéressants, certains proches de sujets de recherche. Une première sélection d'exemples est présentée dans le Chapitre 2, et quelques problèmes plus avancés, utilisant des variables aléatoires, sont décrits dans le Chapitre 4. Soulignons que le traitement d'un ou de plusieurs de ces exemples, déjà dans la première partie d'un cours, est très formateur.

Le Chapitre 5 décrit de manière synthétique les espaces probabilisés généraux, en particulier ceux où la probabilité *n'est pas* définie sur tous les sous-ensembles de l'espace d'états. L'impossibilité de traiter certains modèles très intéressants sur des espaces probabilisés discrets, en particulier les variables aléatoires à densité, nous a conduit à introduire ce chapitre, mais en se limitant aux définitions essentielles.

Le Chapitre 6 est dédié aux variables aléatoires à densité. Dans cette partie du livre, nous avons choisi d'omettre certaines démonstrations, mais nous avons cherché à toujours donner les définitions mathématiques précises, en mentionnant les questions techniques qui ne peuvent pas être résolues avec les outils à disposition. Les prérequis sont au niveau d'un cours d'analyse mathématique de première année de licence (limites, dérivées, intégrale de Riemann), à l'exception des paragraphes finaux sur les vecteurs aléatoires, signalés par un astérisque *, pour lesquels une connaissance d'un peu d'analyse multivariée est nécessaire (dérivés partielles, intégrale multidimensionnelle).

Le Chapitre 7 présente les théorèmes limites classiques du calcul des probabilités, c'est-à-dire la loi (faible) des grands nombres et le théorème central limite. Pour ce dernier, nous fournissons une démonstration complète (avec l'hypothèse d'un troisième moment fini), et nous développons en détail la technique de l'approximation normale.



Ordres possibles de lecture.

La voie d'accès la plus naturelle au Chapitre 7 est celle qui provient du Chapitre 6 ; ceci dit, la première moitié du Chapitre 7, sur la loi des grands nombres, est déjà accessible après le Chapitre 3. Une bonne partie du Chapitre 9 de simulation est accessible après le Chapitre 6 : certains résultats des Chapitres 7 et 8 sont utilisés dans les Sections 9.4-9.5.

Avec le Chapitre 7 se conclut le programme standard d'un cours introductif aux probabilités. Les derniers Chapitres 8 et 9 peuvent soit être intégrés, complètement ou en partie, dans un cours d'introduction, dans le but d'illustrer plus avant les applications des probabilités, soit former le cœur d'un cours plus avancé, avec quelques-uns des exemples plus avancés présentés dans le Chapitre 4.

Le Chapitre 8 est dédié à quelques applications à la statistique. Nous y exposons les sujets classiques de statistique inférentielle univariée, notamment les estimateurs et intervalles de confiance.

Enfin, dans le Chapitre 9, nous décrivons les techniques principales pour la *simulation* de variables aléatoires par ordinateur, en les appliquant à l'étude numérique d'une sélection de modèles probabilistes. Tous les exemples sont accompagnés du code correspondant dans le langage de programmation Python.

Le diagramme ci-dessus présente les ordres de lecture conseillés.

De façon cohérente avec l'esprit de ce livre, l'exposition de la théorie est enrichie de nombreux exemples, qui constituent une partie fondamentale de la présentation. Dans chaque chapitre, nous présentons de plus une vaste sélection d'exercices, pour lesquels une solution détaillée est accessible via la page de présentation de l'ouvrage sur le site dunod.com. Les derniers exercices, signalés comme étant « plus difficiles », en plus de constituer des défis pour les étudiants plus motivés, peuvent servir de base pour un enseignant qui souhaiterait proposer des problèmes plus articulés, sous la forme d'approfondissements ou de projets. Certaines parties, plus techniques (ou que nous avons considérées comme non essentielles), apparaissent en caractère plus petit, ou sont reportées dans l'Appendice.

Cette deuxième édition, même si elle conserve la même structure générale que la première édition italienne, présente des changements significatifs. Outre la présence de nouveaux Chapitres (un chapitre introductif aux chaînes de markov non présent dans l'édition française, et le Chapitre 9) signalons certaines variations notables.

- Les Chapitres 2 et 4 ont été remaniés, en remplaçant certains modèles développés dans la première édition par d'autres modèles pertinents et actuels, parmi lesquels la *percolation* (Section 2.3), le *processus de branchement de Galton–Watson* (Section 4.5) et le *graphe aléatoire d'Erdős–Rényi* (Section 4.6).

- Dans le Chapitre 3, nous illustrons la *méthode des moments*, qui joue un rôle clé dans le traitement de nombreux modèles probabilistes, comme nous le verrons dans le Chapitre 4 ; nous avons aussi ajouté la définition de *médiane* et de *quantiles* (Section 3.4.3).
- Les Chapitres 6, 7 et 8 ont aussi été enrichis par de nouveaux exemples et de nouvelles sections.

Une autre nouveauté importante est le choix de formuler de nombreuses définitions et résultats des Chapitres 1 et 3 sans faire explicitement référence à la nature discrète des espaces probabilisés et des variables aléatoires, afin de rendre évident (rétrospectivement) quand ceux-ci restent valables pour des espaces probabilisés généraux.

Ce livre est évidemment le miroir de notre formation, de nos intérêts de recherche, et de nos goûts personnels. Nous avons été inspirés et aidés par les commentaires et observations de nombreuses et nombreux collègues, que nous remercions sincèrement pour leurs suggestions. Un merci particulier à Wolfgang J. Runggaldier et Tiziano Vargiolu, pour les nombreuses et utiles discussions qui ont accompagné la rédaction de la première édition, et à Sébastien Martineau et Lorenzo Zambotti pour l'enthousiasme avec lequel ils ont contribué à améliorer la présentation de plusieurs points dans la deuxième édition. Merci aussi à Omer Adelman, Nicolas Pétrélis, Damien Simon, Camille Tardif et Pierre Tarrago, pour leurs relectures et leurs commentaires concernant la version française.

En outre, nous devons beaucoup aux autrices et auteurs d'articles et de livres desquels nous avons appris une grande partie des mathématiques que nous présentons ici. En particulier, nous croyons, nous espérons, avoir été influencés par les excellents livres de William Feller [22, 23] et Patrick Billingsley [6].

Pour finir, nous sommes reconnaissants aux étudiantes et étudiants de licence en mathématiques des universités de Padoue, Milan-Bicocca et Sorbonne Université et de la licence en mathématiques appliquées de l'université de Vérone : vos commentaires, vos critiques et vos signalements d'erreurs ont contribué au projet, à la construction et à la révision de ce livre.

Paris, Milan, Vérone
Juin 2021

Quentin Berger
Francesco Caravenna
Paolo Dai Pra

Table des matières

Notions préliminaires

Sommaire Dans ce chapitre, nous fixons les notations importantes que nous utiliserons tout au long de ce livre et nous rappelons quelques résultats d'analyse qui reviendrons fréquemment. Nous définissons notamment la notion de somme pour une famille infinie de réels indexée par un ensemble quelconque, et nous en commentons les propriétés.

Notations

Pour un ensemble Ω et deux sous-ensembles $A, B \subseteq \Omega$ donnés, on utilisera les notations standard

$$\begin{aligned}
A \cup B &:= \{\omega \in \Omega : \ \omega \in A \text{ ou } \omega \in B\}, \\
A \cap B &:= \{\omega \in \Omega : \ \omega \in A \text{ et } \omega \in B\}, \\
A^c &:= \{\omega \in \Omega : \ \omega \notin A\}, \\
A \setminus B &:= A \cap B^c, \\
A \triangle B &:= (A \setminus B) \cup (B \setminus A) \ = \ (A \cup B) \setminus (A \cap B),
\end{aligned}$$

où le symbole « := » indique une définition. Les notions d'union et d'intersection s'étendent à une famille quelconque $(A_i)_{i \in I}$ de sous-ensembles de Ω :

$$\begin{aligned}
\bigcup_{i \in I} A_i &:= \{\omega \in \Omega : \exists i \in I \text{ tel que } \omega \in A_i\}, \\
\bigcap_{i \in I} A_i &:= \{\omega \in \Omega : \forall i \in I \text{ on a } \omega \in A_i\}.
\end{aligned}$$

Avec un léger abus de terminologie, on dira qu'une *union* $\bigcup_{i \in I} A_i$ *est disjointe* si les ensembles $(A_i)_{i \in I}$ sont *deux à deux disjoints*, c'est-à-dire si $A_i \cap A_j = \emptyset$ quels que soient $i \neq j$; on dira simplement disjoints dans la suite, par souci de brévité.

Nous utiliserons souvent les propriétés de distributivité des intersections et unions :

$$\big(A \cup B\big) \cap C = (A \cap C) \cup (B \cap C), \qquad \big(A \cap B\big) \cup C = (A \cup C) \cap (B \cup C),$$

et les lois de De Morgan :

$$(A \cup B)^c \ = \ A^c \cap B^c, \qquad (A \cap B)^c \ = \ A^c \cup B^c,$$

qui restent valables plus généralement pour des familles quelconques d'ensembles :

$$\Big(\bigcup_{i\in I} A_i\Big)\cap C = \bigcup_{i\in I}(A_i\cap C)\,, \qquad \Big(\bigcap_{i\in I} A_i\Big)\cup C = \bigcap_{i\in I}(A_i\cup C)\,,$$
$$\Big(\bigcup_{i\in I} A_i\Big)^c = \bigcap_{i\in I} A_i^c\,, \qquad \Big(\bigcap_{i\in I} A_i\Big)^c = \bigcup_{i\in I} A_i^c\,.$$

On notera $\mathbb{N} := \{0,1,2,\dots\}$ l'ensemble des entiers naturels et $\mathbb{N}^* := \{1,2,\dots\}$ l'ensemble des entiers naturels *zéro exclu*. On adoptera les notations standard pour les ensembles des nombres entiers relatifs, rationnels et réels, respectivement notés $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{R}$, et on écrira $\mathbb{R}_+ := [0,\infty[= \{x\in\mathbb{R} : x \geq 0\}$; on notera aussi $\mathbb{Z}^* = \mathbb{Z}\setminus\{0\}$, $\mathbb{Q}^* = \mathbb{Q}\setminus\{0\}$ et $\mathbb{R}^* = \mathbb{R}\setminus\{0\}$.

On dira qu'un réel $x\in\mathbb{R}$ est *positif* si $x \geq 0$ et *strictement positif* si $x > 0$; de manière analogue, on dira qu'un réel x est *négatif* si $x \leq 0$ et *strictement négatif* si $x < 0$. Notons qu'avec ces conventions, 0 est aussi bien positif que négatif. La partie positive et négative d'un réel $x\in\mathbb{R}$ sont définies respectivement par

$$x^+ := \max\{x,0\}\,, \qquad x^- := -\min\{x,0\} = \max\{-x,0\}\,,$$

de sorte que $x^+, x^- \geq 0$, et $x = x^+ - x^-$. La valeur absolue de x est donnée par $|x| = x^+ + x^-$.

Pour $x\in\mathbb{R}$, on notera $\lfloor x\rfloor$ la *partie entière* de x, c'est-à-dire le plus grand entier inférieur à x :

$$\lfloor x\rfloor := \max\{m\in\mathbb{Z} : m \leq x\}\,. \tag{0.1}$$

On notera aussi $\lceil x\rceil := \min\{m\in\mathbb{Z} : m \geq x\}$ le plus petit entier supérieur à x, appelé *partie entière supérieure* de x.

On utilisera les adjectifs *croissant* et *décroissant* dans le sens faible : une fonction $f : \mathbb{R}\to\mathbb{R}$ est croissante (resp. décroissante) si pour tous $x \geq y$ on a $f(x) \geq f(y)$ (resp. $f(x) \leq f(y)$). Une fonction constante est donc à la fois croissante et décroissante.

Le produit cartésien $A\times B$ de deux ensembles A et B est l'ensemble des couples ordonnés (a,b) avec $a\in A$ et $b\in B$. On définit de manière analogue le produit de plus de deux ensembles. On notera A^n le produit $A\times\ldots\times A$ (n fois), pour $n\in\mathbb{N}^*$.

Pour un ensemble quelconque Ω, pour tout sous-ensemble $A\subseteq\Omega$ on définit la fonction $\mathbb{1}_A : \Omega\to\mathbb{R}$, appelée *indicatrice de A*, qui vaut 1 sur A et 0 sur A^c :

$$\mathbb{1}_A(x) := \begin{cases} 1 & \text{si } x\in A\,, \\ 0 & \text{si } x\notin A\,. \end{cases} \tag{0.2}$$

Dans d'autres contextes, cette fonction est appelée fonction caractéristique et est notée χ_A. En probabilités, on utilise en général « fonction caractéristique » pour désigner un autre objet (que nous ne décrirons pas dans ce livre). Nous nous tiendrons donc à la terminologie *fonction indicatrice* et à la notation $\mathbb{1}_A$.

Cardinalité et dénombrabilité

Pour un ensemble *fini* A, on notera $|A|$ son cardinal, c'est-à-dire son nombre d'éléments. On écrira $|A| < +\infty$ pour indiquer que l'ensemble est fini, et $|A| = +\infty$ quand l'ensemble est infini.

Un ensemble A est dit *dénombrable* s'il est fini ou bien en bijection avec $\mathbb{N}$, autrement dit s'il existe une application injective $f : A\to\mathbb{N}$. De manière concrète, un ensemble est dénombrable si on peut lister ses éléments, c'est-à-dire si on peut les écrire comme une suite finie $x_0, x_1, \dots, x_N$ ou infinie $(x_n)_{n\in\mathbb{N}}$. Rappelons ici quelques résultats importants, permettant de voir si des ensembles sont dénombrables.

Tout d'abord, si $A_1, \dots, A_n$ sont des ensembles dénombrables ($n \in \mathbb{N}^*$), alors le produit cartésien $A_1 \times \dots \times A_n$ est également dénombrable. Un corollaire assez simple est que l'ensemble $\mathbb{Z}$ des entiers relatifs et l'ensemble $\mathbb{Q}$ des nombres rationnels sont tous les deux dénombrables.

Si $(A_i)_{i\in I}$ est une famille dénombrable d'ensembles dénombrables, c'est-à-dire si I est dénombrable et si A_i est dénombrable pour tout $i \in I$, alors l'union $\bigcup_{i\in I} A_i$ est également dénombrable.

Rappelons enfin que l'ensemble des réels $\mathbb{R}$ (ou n'importe quel intervalle de $\mathbb{R}$ non réduit à un point) n'est pas dénombrable.

Quelques rappels d'analyse

Étant donnée une suite de réels $(a_n)_{n\in\mathbb{N}}$, la *série* $\sum_{n=0}^{\infty} a_n$ est définie comme la limite de la suite des sommes partielles $(s_N := \sum_{n=0}^{N} a_n)_{N\in\mathbb{N}}$, quand une telle limite existe dans $\mathbb{R} \cup \{\pm\infty\}$. Rappelons le critère, dit *de Riemann* : pour tout $a \in \mathbb{R}$, on a

$$\sum_{n=1}^{\infty} \frac{1}{n^a} \begin{cases} < +\infty & \text{si } a > 1\,, \\ = +\infty & \text{si } a \leq 1\,, \end{cases} \tag{0.3}$$

comme on peut le montrer, par exemple, en comparant la série à une intégrale appropriée.

Les sommes partielles de la *série géométrique* (à partir de 0 ou de 1) sont connues :

$$\sum_{n=0}^{N} x^n = \frac{1 - x^{N+1}}{1 - x}\,, \qquad \sum_{n=1}^{N} x^n = \frac{x(1 - x^N)}{1 - x}\,, \qquad \forall x \in \mathbb{R} \setminus \{1\}\,, \tag{0.4}$$

ce que l'on peut montrer facilement par récurrence. En particulier, il s'ensuit que

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1 - x}\,, \qquad \sum_{n=1}^{\infty} x^n = \frac{x}{1 - x}\,, \qquad \forall x \in \mathbb{R} \text{ tel que } |x| < 1\,. \tag{0.5}$$

Ces deux séries valent $+\infty$ si $x \geq 1$, alors qu'elles ne sont pas définies si $x \leq -1$ (la suite des sommes partielles n'a pas de limite).

Rappelons les séries de Taylor des fonctions exponentielle et logarithme (on utilisera la notation anglo-saxonne log pour le logarithme népérien ln) :

$$\mathrm{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad \forall x \in \mathbb{R}\,, \tag{0.6}$$

$$\log(1 + x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \qquad \forall x \in \,]-1, 1]\,, \tag{0.7}$$

où le symbole $n!$ (« factorielle n ») est rappelé dans la Section 1.2.4.

Rappelons aussi les formules explicites pour les sommes des entiers de 1 à n et de leurs carrés, qui se démontrent facilement par récurrence :

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}\,, \qquad \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}\,, \qquad \forall n \in \mathbb{N}^*\,. \tag{0.8}$$

Énonçons pour finir le comportement asymptotique précis de la série harmonique :

$$H_n := \sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + o(1) \qquad \text{quand } n \to \infty, \tag{0.9}$$

où $\gamma \simeq 0,577$ est appelée *constante d'Euler–Mascheroni* †, et $o(1)$ désigne une quantité qui tend vers 0 quand $n \to \infty$. Il est souvent suffisant d'utiliser l'estimée élémentaire suivante :

$$\log n \le H_n \le \log n + 1, \qquad \forall n \ge 1, \tag{0.10}$$

que l'on obtient en sommant les inégalités $\frac{1}{k+1} \le \int_k^{k+1} \frac{1}{x} dx \le \frac{1}{k}$ sur k.

Sommes infinies

Il nous arrivera régulièrement de considérer *la somme d'une famille de réels* $(a_i)_{i\in I}$ indexée par un ensemble I général, qui peut être non dénombrable et sans ordre, on procède de la manière suivante. Nous en donnons une définition précise, en considérant dans un premier temps le cas fondamental où tous les termes sont positifs, c'est-à-dire $a_i \ge 0$ pour tout $i \in I$.

Familles à termes positifs

Pour une famille $(a_i)_{i\in I}$ à termes positifs donnée, c'est-à-dire telle que $a_i \ge 0$ pour tout $i \in I$, *la somme* $\sum_{i\in I} a_i$ *est toujours bien définie*, comme le supremum sur toutes les sommes finies possibles d'éléments a_i :

$$\sum_{i\in I} a_i \; := \sup_{A\subseteq I,\, |A|<\infty} \; \sum_{j\in A} a_j \;\in\; [0,+\infty].$$

Soulignons que $\sum_{i\in I} a_i$ peut prendre la valeur $+\infty$.

Cette définition s'étend au cas où $a_i \in [0,+\infty]$, avec la convention que la somme vaut $+\infty$ dès que l'un (au moins) des éléments vaut $+\infty$.

Une première remarque, élémentaire mais fondamentale, est que *si la somme est finie, il n'y a nécessairement qu'un nombre dénombrable de termes non nuls.*

Observation 0.1. Si une famille $(a_i)_{i\in I}$ à termes positifs admet une somme finie, c'est-à-dire si $\sum_{i\in I} a_i < \infty$, alors l'ensemble $\tilde{I} := \{i \in I : a_i \neq 0\}$ est dénombrable. En effet, pour tout $\varepsilon > 0$ fixé, les termes tels que $a_i > \varepsilon$ sont nécessairement en nombre fini ; sinon on montrerait facilement que $\sum_{i\in I} a_i = +\infty$. Ainsi, en écrivant

$$\tilde{I} := \{i \in I : a_i \neq 0\} \;=\; \bigcup_{n\in\mathbb{N}^*} \{i \in I : a_i > \tfrac{1}{n}\}, \tag{0.11}$$

on obtient que l'ensemble $\tilde{I}$ est une union dénombrable d'ensembles finis, et est donc dénombrable. □

Observation 0.2. Si l'ensemble $\tilde{I} = \{i \in I : a_i \neq 0\}$ est dénombrable, on peut énumérer ses éléments en une suite $(i_n)_{n\ge 1}$. On peut alors montrer que la somme à termes positifs $\sum_{i\in I} a_i$

†. On peut montrer que $\gamma = -\int_0^\infty \log x \, e^{-x} \, dx$.

coïncide avec la série habituelle $\sum_{n=1}^{\infty} a_{i_n} = \lim_{N\to\infty} \sum_{n=1}^{N} a_{i_n}$, qui donc *ne dépend pas* du choix de l'énumération $(i_n)_{n\geq 1}$. □

Les sommes de familles infinies à termes positifs possèdent de nombreuses propriétés des sommes habituelles, que nous utiliserons fréquemment dans ce livre, et que nous énonçons maintenant. Les démonstrations ne sont pas difficiles, mais les détails sont fastidieux : le lecteur intéressé peut les trouver dans l'Appendice A.1.

Soient $(a_i)_{i\in I}$ et $(b_i)_{i\in I}$ deux familles à termes positifs.

- *Linearité.* Pour tous $\alpha, \beta \geq 0$: $\displaystyle\sum_{i\in I}(\alpha a_i + \beta b_i) = \alpha\Big(\sum_{i\in I} a_i\Big) + \beta\Big(\sum_{i\in I} b_i\Big)$.
- *Monotonie.* Si $a_i \leq b_i$ pour tout $i \in I$, alors $\displaystyle\sum_{i\in I} a_i \leq \sum_{i\in I} b_i$.
- *Associativité (sommation par paquets).* Soit $(I_j)_{j\in J}$ une *partition* de I, c'est-à-dire

$$I = \bigcup_{j\in J} I_j\,, \qquad I_j \cap I_k = \emptyset \quad \text{pour } j \neq k\,.$$

Alors, en posant $s_j := \sum_{i\in I_j} a_i$ pour $j \in J$, on a

$$\sum_{i\in I} a_i = \sum_{j\in J} s_j = \sum_{j\in J}\left(\sum_{i\in I_j} a_i\right). \tag{0.12}$$

Comme cas particulier de l'associativité, notons que tout ensemble produit $I = E \times F$ peut être partitionné en $I = \bigcup_{x\in E}\{x\} \times F$, ou bien en $I = \bigcup_{y\in F} E \times \{y\}$. Par conséquent, pour toute famille $(a_{x,y})_{(x,y)\in E\times F}$ à termes positifs, on a

$$\sum_{(x,y)\in E\times F} a_{x,y} = \sum_{x\in E}\left(\sum_{y\in F} a_{x,y}\right) = \sum_{y\in F}\left(\sum_{x\in E} a_{x,y}\right), \tag{0.13}$$

qui est une version discrète du théorème de Fubini–Tonelli. Cette relation s'étend au cas où $I = E_1 \times \cdots \times E_n$ est le produit cartésien d'un nombre fini d'ensembles.

Familles de signe quelconque

Considérons maintenant une famille $(a_i)_{i\in I}$ de réels, non nécessairement positifs. On dira que la famille $(a_i)_{i\in I}$ *admet une somme* si au moins une des deux sommes $\sum_{i\in I} a_i^+$, $\sum_{i\in I} a_i^-$ est finie (ces deux sommes sont biens définies, car elles sont à termes positifs). Si tel est le cas, la somme $\sum_{i\in I} a_i$ est définie par

$$\sum_{i\in I} a_i := \sum_{i\in I} a_i^+ - \sum_{i\in I} a_i^- \in [-\infty, +\infty]\,. \tag{0.14}$$

Notons qu'une famille à termes tous positifs ou tous négatifs admet toujours une somme.

Si une famille $(a_i)_{i\in I}$ admet une somme et si $\sum_{i\in I} a_i \in\,]-\infty, +\infty[$, on dira que la famille *admet une somme finie*, ou *est sommable*. Clairement, c'est le cas si et seulement si *les deux sommes* $\sum_{i\in I} a_i^+$ et $\sum_{i\in I} a_i^-$ sont finies, ce qui par linéarité est équivalent à $\sum_{i\in I} |a_i| < \infty$.

Observation 0.3. Pour une famille $(a_n)_{n\in\mathbb{N}}$ indexée par les entiers naturels, c'est-à-dire une suite, la notion d'*admettre une somme finie* correspond à la *convergence absolue* de la série correspondante, $\sum_{n=0}^{\infty} |a_n| < \infty$. □

Observation 0.4. La définition de somme infinie $\sum_{i\in I} a_i$ donnée plus haut s'étend sans aucune modification au cas où les éléments de la famille $(a_i)_{i\in I}$ appartiennent à la *ligne réelle étendue* $\overline{\mathbb{R}} := [-\infty, +\infty]$, avec la convention donnée plus haut qu'une somme d'éléments de $[0, +\infty]$ vaut $+\infty$ dès que l'un (au moins) des éléments vaut $+\infty$. $\square$

Une conséquence de l'Observation 0.1 est que *si une famille $(a_i)_{i\in I}$ admet une somme non nulle, alors les termes non nuls $a_i \neq 0$ sont nécessairement en nombre dénombrable* : on peut alors les énumérer en une suite $(a_{i_n})_{n\geq 1}$ et on peut montrer que la somme $\sum_{i\in I} a_i$ coïncide avec la série $\sum_{n=1}^{\infty} a_{i_n} = \lim_{N\to\infty} \sum_{n=1}^{N} a_{i_n}$. Cela n'est plus vrai si la famille n'admet pas de somme (voir l'Observation 0.5).

Les propriétés de linéarité, monotonie et associativité de la somme restent valables pour les familles à signe quelconque, avec les hypothèses suivantes :

- *Linéarité* : pour des familles $(a_i)_{i\in I}$, $(b_i)_{i\in I}$ qui *admettent une somme finie* et $\alpha, \beta \in \mathbb{R}$;
- *Monotonie* : pour des familles $(a_i)_{i\in I}$, $(b_i)_{i\in I}$ qui *admettent une somme (finie ou infinie)* ;
- *Associativité* : pour des familles $(a_i)_{i\in I}$ qui *admettent une somme (finie ou infinie)* ;

Les énoncés précis (et les démonstrations) se trouvent dans l'Appendice A.1.

Observation 0.5. Supposons qu'une famille $(a_i)_{i\in I}$ possède *un nombre dénombrable de termes non nuls*, on peut alors énumérer ces termes non nuls en une suite $(a_{i_n})_{n\in\mathbb{N}}$, mais le choix d'une telle suite, c'est-à-dire de l'ordre dans lequel énumérer ces termes, est arbitraire.

Si la famille $(a_i)_{i\in I}$ admet une somme (finie ou infinie), on peut montrer que la somme $\sum_{i\in I} a_i$ coïncide avec la série $\sum_{n=0}^{\infty} a_{i_n} = \lim_{N\to\infty} \sum_{n=0}^{N} a_{i_n}$, qui donc *ne dépend pas* de la façon de choisir la suite $(a_{i_n})_{n\in\mathbb{N}}$.

Ceci n'est plus vrai si la famille $(a_i)_{i\in I}$ n'admet pas de somme : des ordres différents de la suite $(a_{i_n})_{n\in\mathbb{N}}$ peuvent produire des valeurs différentes de la série $\sum_{n=0}^{\infty} a_{i_n}$, ou même des séries qui ne convergent pas. C'est la raison pour laquelle il faut être attentif dans la définition d'une somme d'une famille infinie $(a_i)_{i\in I}$ quand il n'y a pas d'ordre canonique sur l'ensemble des indices I (ou, dans le cas où I possède un ordre, quand on souhaite que la somme ne dépende pas de l'ordre dans lequel sont énumérés les termes).

Un exemple classique est donné par la famille $(a_i := (-1)^i \frac{1}{i})_{i\in\mathbb{N}^*}$. La série correspondante est convergente : en effet, la relation (0.7) donne

$$\sum_{i=1}^{\infty} \frac{(-1)^i}{i} := \lim_{N\to\infty} \sum_{i=1}^{N} \frac{(-1)^i}{i} = -\log 2 .$$

Cependant, selon la définition, la famille *n'admet pas de somme* : en effet,

$$\sum_{i\in\mathbb{N}^*} a_i^+ = \sum_{i\in 2\mathbb{N}^*} \frac{1}{i} = +\infty, \qquad \sum_{i\in\mathbb{N}^*} a_i^- = \sum_{i\in 2\mathbb{N}+1} \frac{1}{i} = +\infty,$$

où $2\mathbb{N}^*$ et $2\mathbb{N}+1$ désignent respectivement les entiers naturels pairs et impairs.

Plus précisément, dans ce cas, il est possible de montrer que, si l'on choisit un réarrangement bien choisi $(a_{i_n})_{n\in\mathbb{N}}$ de la famille, on peut faire converger la série correspondante $\sum_{n=1}^{\infty} a_{i_n}$ vers n'importe quel nombre de la droite réelle étendue $x \in \mathbb{R} \cup \{\pm\infty\}$ fixé à l'avance ! Il s'agit d'un cas particulier du *théorème de réarrangement de Riemann*. $\square$

Chapitre 1
Espaces probabilisés discrets : la théorie

Sommaire Dans ce chapitre, nous introduisons la notion d'espace probabilisé discret, utilisé pour décrire une expérience aléatoire dont l'ensemble des résultats possibles est dénombrable, et nous en étudions les propriétés principales. Nous détaillons ici les techniques classiques de dénombrement pour les ensembles finis, appelé combinatoire. À ce stade, il est déjà possible, et peut-être même conseillé, de s'attaquer à l'analyse d'un ou plusieurs des modèles probabilistes présentés dans le Chapitre 2. Enfin, nous développons les notions cruciales de conditionnement et d'indépendance d'événements, et nous les illustrons avec une sélection d'exemples et de paradoxes.

1.1 Modèles probabilistes

1.1.1 Considérations introductives

Dans ce livre, on parlera d'*expérience aléatoire* en référence à un phénomène quelconque (physique, économique, social, etc.) dont le résultat ne peut pas être déterminé au préalable avec certitude. Notre objectif est de fournir une description mathématique d'une expérience aléatoire, en définissant un *modèle probabiliste*.

La première étape consiste à identifier un ensemble Ω, appelé *espace d'états* ou *univers*, qui contient tous les résultats possibles de l'expérience.

Exemple 1.1. (i) Pour le lancer d'un dé ordinaire à six faces, l'espace d'états naturel est $\Omega = \{1,2,3,4,5,6\}$.

(ii) Pour le nombre de voix obtenues par Selina Meyer aux prochaines élections, on peut choisir comme espace d'états $\Omega = \{0,1,\dots,M\}$, où $M \in \mathbb{N}^*$ est le nombre d'électeurs, mais l'on peut aussi prendre $\Omega = \mathbb{N}$.

(iii) Pour la mesure du temps auquel une particule radioactive est émise par un atome donné (disons mesuré en secondes et fractions de seconde), un choix naturel d'espace d'états est donné par $\Omega = [0,+\infty[$. $\square$

Le deuxième ingrédient d'un modèle probabiliste est constitué des *événements*, qui sont de manière informelle des *affirmations sur le résultat d'une expérience aléatoire*. D'un point de vue mathématique, les événements sont décrits par des *sous-ensembles de l'espace d'états* Ω : chaque affirmation est identifiée au sous-ensemble de Ω constitué des résultats de l'expérience pour lesquels l'affirmation est vérifiée. Quand aucun des résultats possibles de l'expérience ne vérifie l'affirmation, l'événement est dit impossible, et identifié à l'ensemble vide $\emptyset$. À l'extrême opposé, une affirmation qui est vérifiée quel que soit le résultat de l'expérience correspond à l'événement certain, donné par l'ensemble Ω tout entier.

Exemple 1.2. Faisons référence à l'exemple précédent.

(i) Lors du lancer d'un dé à six faces, voici quelques exemples d'événements :
$A :=$ « Le résultat du dé est un nombre pair » $= \{2,4,6\}$,
$B :=$ « Le résultat du dé est un multiple de 3 » $= \{3,6\}$,
$C :=$ « Le résultat du dé est un nombre pair et multiple de 3 » $= \{6\} = A \cap B$.

(ii) L'événement $D =$ « Selina Meyer n'obtient aucun vote aux prochaines élections » correspond au sous-ensemble $D = \{0\} \subseteq \Omega$. Un exemple d'événement impossible est donné par $E :=$ « Selina Meyer obtient $M+1$ votes » $= \emptyset$, M étant le nombre total d'électeurs.

(iii) L'événement $F =$ « La particule est émise après plus d'un an » correspond au sous-ensemble $F = [60 \cdot 60 \cdot 24 \cdot 365, +\infty[= [31\,536\,000, +\infty[\subseteq \Omega$. □

Notons que les opérations logiques de *conjonction*, de *disjonction* et de *négation* des affirmations correspondent à l'*intersection*, à l'*union* et au *complémentaire* des ensembles :

$$\begin{aligned} \text{« } A \text{ et } B \text{ sont tous les deux vérifiés »} &\longrightarrow A \cap B\,, \\ \text{« au moins un parmi } A \text{ ou } B \text{ est vérifié »} &\longrightarrow A \cup B\,, \\ \text{« } A \text{ n'est pas vérifié »} &\longrightarrow A^c\,. \end{aligned} \tag{1.1}$$

Observation 1.3. Une question délicate est de savoir s'il est opportun de considérer que *tous* les sous-ensembles d'un espace d'états Ω sont des événements. C'est bien sûr possible quand Ω est dénombrable, mais si Ω est infini non dénombrable (comme dans l'exemple (iii) ci-dessus), il peut se révéler nécessaire de ne considérer comme événements qu'une classe restreinte de sous-ensembles de Ω, pour des raisons qui seront claires par la suite. Dans ce cas, l'analyse devient plus technique et compliquée, et c'est pour cette raison que, dans la première partie de ce livre, nous nous concentrerons principalement sur les espaces d'états dénombrables. □

Le troisième et dernier ingrédient d'un modèle probabiliste, le plus important, est la donnée d'un *degré de confiance*, ou *(mesure de) probabilité*, aux événements de l'expérience aléatoire. Cela permet de formaliser les expressions du type « la probabilité qu'en lançant un dé à six faces on obtienne un nombre pair vaut $1/2$ ». D'un point de vue mathématique, une probabilité est décrite par une application P qui à tout événement $A \subseteq \Omega$ associe un nombre $\mathrm{P}(A) \in [0,1]$. Parfois, la probabilité d'un événement est exprimée par un pourcentage, en écrivant 50% au lieu de 1/2, 10% au lieu de 0,1, et ainsi de suite.

Que signifie concrètement un *degré de confiance* $\mathrm{P}(A)$? Si l'expérience aléatoire pouvait être répétée un très grand nombre $N \gg 1$ de fois, de façon « identique et indépendante », on pourrait (en principe) compter le nombre de fois où l'événement A est vérifié, que l'on notera

$$\mathcal{S}_N(A) \in \{0, \ldots, N\}\,.$$

On peut alors interpréter $\mathrm{P}(A)$ comme *la proportion de fois où l'événement A est vérifié*,

$$\mathrm{P}(A) \simeq \frac{\mathcal{S}_N(A)}{N}\,,$$

et idéalement, $\mathrm{P}(A) = \lim_{N\to+\infty} \frac{\mathcal{S}_N(A)}{N}$. Il s'agit de l'*interprétation fréquentiste* des probabilités (qui est d'une certaine manière justifiée a posteriori par la *loi des grands nombres*, un théorème que nous verrons dans le Chapitre 7).

Soulignons cependant qu'il ne s'agit pas d'une définition opérationnelle : tout d'abord, toutes les expériences aléatoires ne peuvent pas forcément être répétées dans des conditions *identiques et indépendantes* (on peut penser au cas d'une élection, comme dans l'Exemple 1.1 (ii)). De plus, même s'il était possible de répéter l'expérience, rien ne garantit que le rapport $\mathcal{S}_N(A)/N$ converge vers une limite, et même si c'était le cas, il n'est pas clair de savoir à quel point N doit être grand pour que l'approximation $\mathrm{P}(A) \simeq \mathcal{S}_N(A)/N$ soit bonne. Cependant, toutes précautions dûes, il est bon de garder à l'esprit cette interprétation fréquentiste, qui permettra de donner un contenu intuitif aux résultats que nous rencontrerons.

Dans les faits, le choix de la probabilité est un point délicat. Dans certains cas, il existe un choix « naturel », sur la base de considérations sur la nature de l'expérience aléatoire considérée, notamment de symétries. Très souvent, pourtant, ce n'est pas le cas et *le choix peut dépendre de facteurs subjectifs*. L'observation cruciale est que, quelle que soit la manière dont elle est choisie, une probabilité P devra satisfaire certaines propriétés, par exemple :

$$\mathrm{P}(\Omega) = 1\,, \qquad \mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) \quad \text{si} \quad A \cap B = \emptyset\,. \tag{1.2}$$

Ces relations sont tout à fait naturelles si l'on pense à l'interprétation fréquentiste $\mathrm{P}(A) \simeq \mathcal{S}_N(A)/N$: d'une part on a $\mathcal{S}_N(\Omega) = N$ (car Ω contient *tous* les résultats possibles de l'expérience) ; d'autre part $\mathcal{S}_N(A \cup B) = \mathcal{S}_N(A) + \mathcal{S}_N(B)$ si $A \cap B = \emptyset$.

Le point de vue moderne consiste à *appeler probabilité toute fonction* P *qui satisfait une version renforcée des propriétés* (1.2). Cette approche a des conséquences riches et profondes, comme on aura l'occasion de l'apprécier.

Observation 1.4. Une autre interprétation des probabilités est celle *subjectiviste*, basée sur un *schéma de paris*. Dans cette approche, la probabilité $\mathrm{P}(A)$ d'un événement A représente le *prix équitable* qu'un individu attribue à un pari qui rapporte 1 si l'événement A est vérifié, et qui rapporte 0 sinon †. Par *prix équitable*, on veut dire que que l'individu est disposé non seulement à acheter, mais aussi à vendre des paris (c'est-à-dire à « faire la banque ») aux prix établis $\mathrm{P}(A)$. Avec cette interprétation aussi, les propriétés (1.2) sont naturelles : en effet, si ces propriétés ne sont pas vérifiées, on peut construire une combinaison opportune de paris qui produit un gain sûr où une perte sûre, quel que soit le résultat de l'expérience (on omet ici les détails par souci de brévité). □

Observation 1.5. Le *choix* du modèle probabiliste pour une expérience aléatoire peut dépendre de considérations extra-mathématiques. Cependant, une fois le modèle choisi, son *étude* est un problème purement mathématique. De plus, quand c'est possible, il convient de soumettre le modèle choisi à des vérifications, sur la base de données expérimentales. Cela constitue l'un des objectifs principaux de la *statistique*, dont nous développerons certains aspects dans le Chapitre 8. □

1.1.2 Axiomes des probabilités

Donnons quelques définitions fondamentales, motivées par les considérations précédentes.

Définition 1.6 (Axiomes des probabilités, I). Soit Ω un ensemble non vide arbitraire et soit $\mathcal{P}(\Omega)$ la famille de tous les sous-ensembles de Ω. Une fonction

$$\mathrm{P} : \mathcal{P}(\Omega) \to [0,1]$$

†. On utilise une formulation équivalente dans les paris sportifs : pour un pari sur le fait que l'événement A soit vérifié, on fixe par convention à 1 le prix du pari, tandis que le bookmaker établit le montant associé au *gain* ; on parle aussi de *cote*, qui correspond alors à $1/\mathrm{P}(A)$.

est appelée *probabilité* sur Ω si elle vérifie les propriétés suivantes :

(P1) $\mathrm{P}(\Omega) = 1$.

(P2) (σ-*additivité*) Pour toute famille (infinie) *dénombrable* $(A_n)_{n\geq 1}$ de sous-ensembles de Ω disjoints, c'est-à-dire tels que $A_n \cap A_m = \emptyset$ pour $n \neq m$, on a

$$\mathrm{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathrm{P}(A_n)\,.$$

On utilisera la terminologie suivante :

- l'ensemble Ω est appelé *espace d'états* (ou *univers*) ;
- les sous-ensembles $A \subseteq \Omega$ sont appelés *événements* ;
- le couple (Ω, P) est appelé *espace probabilisé*.

L'interprétation d'un espace probabilisé est celle donnée plus haut : l'ensemble Ω contient tous les résultats possibles ω de l'expérience aléatoire, et pour tout événement $A \subseteq \Omega$, le réel $\mathrm{P}(A) \in [0,1]$ exprime le *degré de confiance* que l'on attribue au fait que l'événement A se réalise (c'est-à-dire à la possibilité que le résultat ω de l'expérience soit un élément de A).

Observation 1.7. Dans la première partie de ce livre, nous nous concentrerons principalement sur les espaces d'états Ω *dénombrables*. Dans ce contexte, la Définition 1.6 est parfaitement adéquate. Plus loin, nous verrons que si Ω n'est pas dénombrable, la Définition 1.6 devient trop restrictive, et ne permet pas de décrire certains modèles probabilistes importants, comme l'Exemple 1.1 (iii). Pour cette raison, à partir du Chapitre 5, nous donnerons une *définition plus générale d'une probabilité* P : la seule différence par rapport à la Définition 1.6 résidera dans le fait que P sera définie définie *non pas sur tous les sous-ensembles de* Ω, mais seulement sur une classe d'ensembles $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ opportune. □

Revenons à la Définition 1.6. La propriété (P1) exprime le fait que l'espace d'états entier est un événement *certain*. La propriété (P2) est une extension de l'identité (1.2) et requiert une discussion plus avancée. Commençons par déduire quelques conséquences de la Définition 1.6.

Proposition 1.8. *Soit* P *une probabilité. Alors on a les propriétés suivantes :*

(i) $\mathrm{P}(\emptyset) = 0$.

(ii) (Additivité finie) *Si* $A_1, A_2, \ldots, A_k$ *est une famille finie d'événements disjoints (c'est-à-dire* $A_i \cap A_j = \emptyset$ *pour* $i \neq j$*), alors*

$$\mathrm{P}\left(A_1 \cup A_2 \cup \cdots \cup A_k\right) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots + \mathrm{P}(A_k)\,. \tag{1.3}$$

Démonstration. Commençons par le point (i). Posons $x := \mathrm{P}(\emptyset) \in [0,1]$ et prenons $A_n = \emptyset$ pour tout $n \in \mathbb{N}^*$. Clairement, $(A_n)_{n\geq 1}$ est une suite de sous-ensembles disjoints de Ω. Ainsi, grâce à l'axiome (P2) et au fait que $\bigcup_{n=1}^{+\infty} A_n = \emptyset$, on obtient

$$x = \mathrm{P}(\emptyset) = \mathrm{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mathrm{P}(A_n) = \sum_{n=1}^{+\infty} x\,.$$

Une telle égalité n'est possible que si $x = 0$; on a en effet $\sum_{n=1}^{+\infty} x = \infty$ si $x > 0$.

Passons maintenant au point (ii). On peut prolonger la famille d'événements disjoints $A_1, A_2, \ldots, A_k$ à une suite infinie d'événements disjoints, en posant $A_n := \emptyset$ pour $n > k$. Alors, on a $\bigcup_{i=1}^{k} A_i = \bigcup_{i=1}^{+\infty} A_i$, et grâce à l'axiome (P2)

$$\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_k) \;=\; \mathrm{P}\left(\bigcup_{i=1}^{+\infty} A_i\right) \;=\; \sum_{i=1}^{+\infty} \mathrm{P}(A_i) \;=\; \sum_{i=1}^{k} \mathrm{P}(A_j)\,,$$

qui est ce que l'on souhaitait démontrer. □

Observation 1.9. Si on réécrit (1.3) dans le cas spécial $k = 2$, on obtient

$$\mathrm{P}(A \cup B) \;=\; \mathrm{P}(A) + \mathrm{P}(B)\,, \qquad \forall A, B \subseteq \Omega \;\text{ tels que }\; A \cap B = \emptyset\,. \tag{1.4}$$

Remarquons d'ailleurs que la relation (1.3) dans le cas général ($k \geq 2$) peut être obtenue à partir de (1.4), en utilisant une démonstration par récurrence (laissée en exercice). □

La propriété (1.3) ou de manière équivalente (1.4), appelée *additivité finie* ou simplement *additivité*, est une condition « naturelle » qui correspond à une idée intuitive de probabilité, comme on l'a déjà observé dans (1.2). Il est donc légitime de se demander si les deux hypothèses {(P1), (P2)} et {(P1), (1.3)} sont équivalentes, c'est-à-dire si chacune peut se déduire de l'autre. La réponse est affirmative dans le cas où Ω est un ensemble fini, car il n'existe pas de suite infinie d'événements disjoints qui ne soient pas vides à partir d'un certain rang — en effet, $\mathcal{P}(\Omega)$ possède un nombre fini d'éléments, égal à $2^{|\Omega|}$ (voir l'Observation 1.31).

Si en revanche Ω est infini (dénombrable ou non), alors les axiomes {(P1), (P2)} sont strictement plus forts que {(P1), (1.3)}, dans le sens où il existe des fonctions $\mathrm{P} : \Omega \to [0,1]$ qui satisfont (1.3) mais pas (P2). Un exemple avec $\Omega = \mathbb{N}^*$ est décrit dans l'Appendice A.2 (dont la lecture peut être omise, les arguments utilisés étant assez sophistiqués).

Ainsi, la σ-additivité *n'est pas* une conséquence de l'additivité finie. Bien que la théorie des probabilités finiment additives soit développée dans une partie de la littérature mathématique, nous demanderons toujours dans cet ouvrage que la σ-additivité soit vérifiée, car elle est plus adaptée pour une grande partie des applications — il s'agit de l'approche adoptée par une large majorité des auteurs. Les raisons pour lesquelles l'axiome (P2) est plus pertinent par rapport au plus faible (1.3) sont variées et difficilement compréhensibles à ce point de la présentation de la théorie. Toutefois, une conséquence importante de la σ-additivité est déjà donnée dans la Proposition 1.24 plus bas.

Observation 1.10. La propriété d'additivité (1.4) est satisfaite par certaines notions à première vue distantes des probabilités. On peut penser par exemple à l'*aire* d'une figure plane : si on divise la figure en deux parties disjointes, l'aire totale coïncide avec la somme des aires des deux parties. Un raisonnement analogue vaut pour le *volume* d'un solide, ou pour la *masse* de matière contenue dans une région de l'espace. Ces notions (et d'autres) ont en commun une structure mathématique, la *théorie de la mesure*, à laquelle nous ferons référence dans le Chapitre 5. Pour le moment, observons que l'analogie avec la notion d'aire (ou de volume), permet d'interpréter de manière « géométrique » de nombreuses propriétés des probabilités, comme celles décrites plus bas dans la Section 1.1.5. □