

MATHÉMATIQUES EN DEVENIR

Mathématiques en devenir

101. — Jacques Faraut. *Analyse sur les groupes de Lie. Une introduction.*
Deuxième édition revue et augmentée
102. — Patrice Tauvel. *Corps commutatifs et théorie de Galois*
103. — Jean Saint Raymond. *Topologie, calcul diff. et variable complexe*
104. — Clément de Seguins Pazzis. *Invitation aux formes quadratiques*
105. — Bruno Ingrao. *Coniques projectives, affines et métriques*
106. — Wolfgang Bertram. *Calcul différentiel topologique élémentaire*
107. — Henri Lombardi & Claude Quitté. *Algèbre commutative. Méthodes constructives. Modules projectifs de type fini.*
108. — Frédéric Testard. *Analyse mathématique. La maîtrise de l'implicite*
109. — Grégory Berhuy. *Modules : théorie, pratique... et un peu d'arithmétique.* Deuxième édition
110. — Bernard Candelpergher. *Théorie des probabilités. Une introduction élémentaire (Nouveau tirage bonifié)*
111. — Philippe Caldero et Jérôme Germoni. *Histoires hédonistes de groupes et de géométries. Tome premier*
112. — Gema-Maria Díaz-Toca, Henri Lombardi & Claude Quitté. *Modules sur les anneaux commutatifs*
113. — Philippe Caldero et Jérôme Germoni. *Histoires hédonistes de groupes et de géométries. Tome second – encores*
114. — Alain Debreil. *Groupes finis et treillis de leurs sous-groupes*
115. — François Rouvière. *Initiation à la géométrie de Riemann*
116. — Nikolaï Nikolski. *Matrices et opérateurs de Toeplitz*
117. — Philippe Caldero et Jérôme Germoni. *Nouvelles histoires hédonistes de groupes et de géométries. Tome premier (Nouveau tirage)*
118. — Martine et Hervé Queffélec. *Analyse complexe et applications.*
Nouvelle édition revue et bonifiée
119. — Alain Debreil, Jean-Denis Eiden, Rached Mneimné et Tuong-Huy NGuyen. *Formes quadratiques et géométrie*
120. — Christian Leruste. *Topologie algébrique – Une introduction, et au-delà*
121. — Grégory Berhuy. *Algèbre : le grand combat.* Deuxième édition
122. — Philippe Caldero et Jérôme Germoni. *Nouvelles histoires hédonistes de groupes et de géométries. Tome second*
123. — Charles-Michel Marle. *Géométrie symplectique et géométrie de Poisson*
124. — Pascal Boyer. *Petit compagnon des nombres et de leurs applications*
125. — Laurent Le Floch, Frédéric Testard. *Probabilités 1 – Le hasard est la nécessité*
126. — Laurent Le Floch, Frédéric Testard. *Probabilités 2 – Le hasard est la nécessité*

Laurent Le Floch & Frédéric Testard

Probabilités 2

Le hasard *est* la nécessité



Calvage & Mounet

Géomètre de formation, Laurent LE FLOCH est maître de conférences à l'université de La Rochelle. Il a partagé une grande partie des enseignements de probabilités de la licence de mathématiques de La Rochelle avec Frédéric TESTARD de 2010 à 2018. laurent.lefloch@univ-lr.fr

Frédéric TESTARD (1961-2018) a été coauteur avec R. Mneimné de la célèbre « Introduction à la théorie des groupes de Lie classiques », avec L. Salem et C. Salem du grand succès de librairie « Les plus belles formules mathématiques » (traduit en finnois, anglais, japonais, chinois et arabe) et auteur de « Analyse mathématique – La maîtrise de l'implicite ».

Mathematics Subject Classification (1991) – Primary :

05-01 Combinatorics (Instructional exposition)
28-01 Measure and Integration (Instructional exposition)
28-A05 Measurable sets
28-A25 Integration with respect to measures
28-A35 Measures and integrals in product spaces
60-01 Probability Theory and Stochastic Processes (Instructional exposition)
60-A05 Foundations of probability theory (General questions)
60-A10 Foundations of probability theory (Probabilistic measure theory)
60-B10 Foundations of probability theory (Convergence of probability measures)
60-E05 Distribution theory (General theory)
60-E05 Distribution theory (Characteristic functions)
60-F05 Central limit and other weak theorems
60-F15 Strong theorems
60-G05 Foundations of stochastic processes
60-G42 Martingales with discrete parameter
60-G42 Sums of independent random variables ; random walks

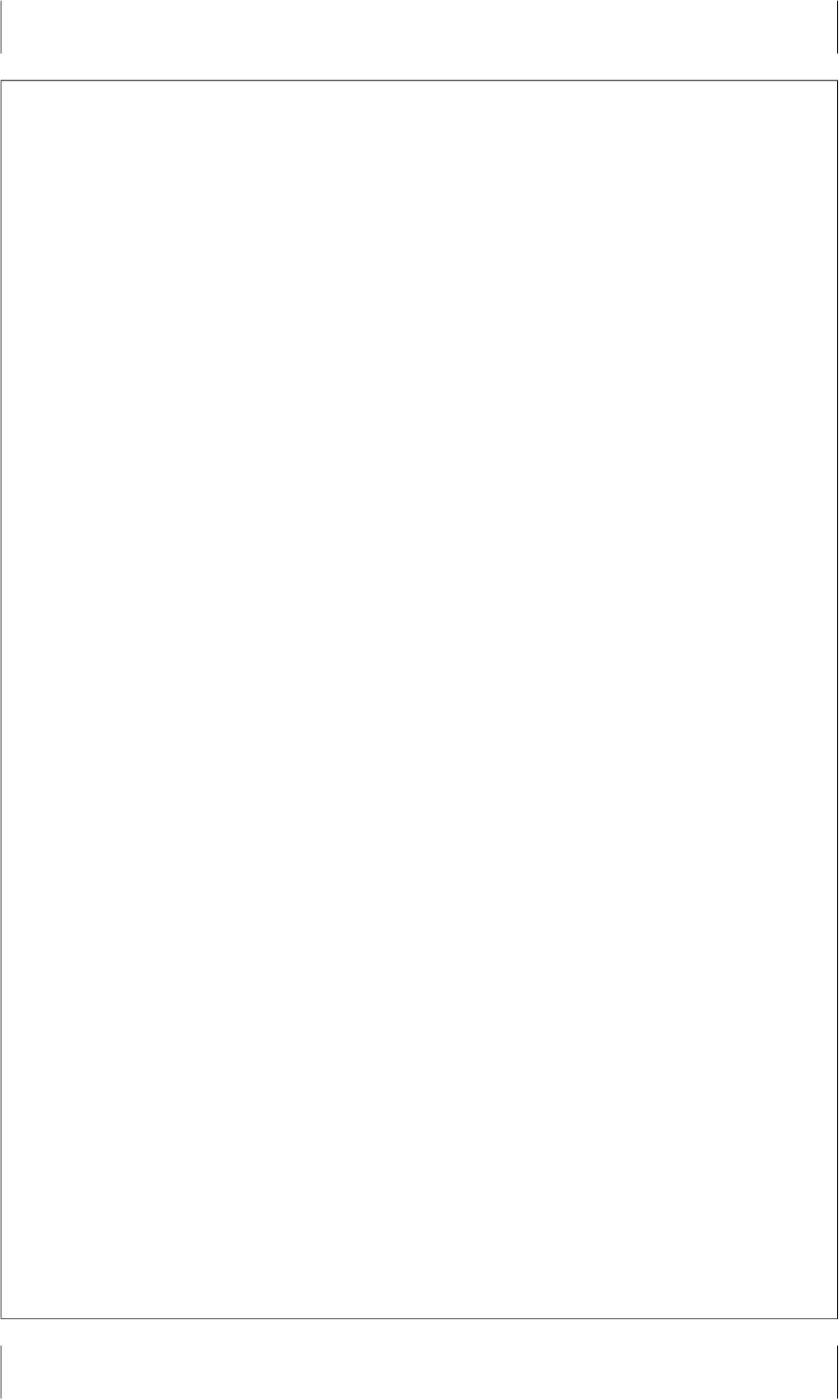
ISBN 978-2-9163-5274-9



∞ Imprimé sur papier permanent

© Calvage & Mounet, Paris, 2021

*À Sylvie, Marie et Paul
À Dominique*



Avant-propos

*Le hasard est le plus grand romancier du monde ;
pour être fécond, il n'y a qu'à l'étudier.*

Honoré de Balzac

Ce livre est destiné aux élèves de lycée (pour une partie du premier tome), aux étudiants de licence ou de master de mathématiques (en particulier les candidats aux concours du CAPES ou de l'agrégation), aux élèves de CPGE, aux enseignants en mathématiques et plus généralement à toute personne intéressée par la théorie des probabilités. Le premier tome de ce livre correspond à l'approche « élémentaire » des probabilités telle qu'elle est classiquement enseignée jusqu'en deuxième année de licence de mathématiques. Le deuxième tome correspond à l'approche moderne des probabilités, initiée par Andreï Kolmogorov, qui s'appuie sur la théorie de la mesure.

Dans les discussions qui ont précédé l'écriture du livre, avant même de définir précisément son contenu, nous avons retenu trois grands principes :

- une approche spiralée dans la présentation des différents concepts probabilistes (variables aléatoires, indépendance, conditionnement) : ces concepts reviennent régulièrement dans des cadres différents (discret fini, discret infini, continu, abstrait) ;
- des énoncés d'exercices volontairement très détaillés. Ils permettent au lecteur d'approfondir sa compréhension des notions rencontrées. Les très nombreux exercices (plus de 700 au total) font partie intégrante du livre ;
- un recours important à l'outil informatique. Nous avons fait le choix d'utiliser principalement le langage Python et plus occasionnellement le tableur. L'outil informatique est un outil incontournable pour la simulation d'expériences aléatoires. Il nous faut cependant signaler qu'il ne s'agit pas d'un livre d'algorithmique. Nous avons en particulier systématiquement

privilegié l'aspect « naturel » des programmes au détriment parfois de leur « efficacité ».

Les cinq premières parties du livre correspondent à l'approche spiralée évoquée précédemment. Ces parties correspondent à la nécessité d'introduire des modèles mathématiques de plus en plus sophistiqués pour résoudre des problèmes eux-mêmes de plus en plus complexes (souvent motivés par la solution de problèmes plus simples traités auparavant). Les derniers chapitres (en général les deux derniers) de chacune de ces parties constituent un petit *vade-mecum* d'outils nécessaires à la description ou l'étude de ces modèles.

La notion de modèle mathématique associé à une expérience aléatoire joue un rôle important dans les premières années d'apprentissage des probabilités. Dans la théorie mathématique des probabilités, la modélisation consiste initialement à construire un ensemble Ω , appelé l'*univers probabiliste* et contenant tous les résultats possibles d'une expérience aléatoire, et à le munir d'une probabilité P , qui associe à certaines parties de Ω appelées *événements* leur probabilité, en respectant un certain nombre de règles d'additivité. Quand l'ensemble Ω est fini (comme c'était, voilà quelques années, systématiquement le cas au lycée), il est naturel de munir toutes les parties de Ω d'une probabilité. Dans ce contexte, les parties à un élément sont appelées *événements élémentaires*, et l'*équiprobabilité* correspond au cas où tous les événements élémentaires ont la même probabilité. On calcule alors les probabilités en dénombrant cas favorables et cas possibles. Dans ce cadre (développé dans les parties 1 et 2 du livre), la théorie des probabilités s'identifie très fortement à l'analyse combinatoire.

Lorsque l'ensemble Ω est infini mais dénombrable (par exemple \mathbb{N}), on peut encore en général calculer la probabilité de toutes les parties de Ω , mais il est impossible cette fois qu'il y ait équiprobabilité. Ce cadre est développé dans la troisième partie du livre.

Dans le cas (très fréquent en pratique) où $\Omega = \mathbb{R}$ (ou \mathbb{R}^n), des « obstructions techniques » liées à la *théorie de la mesure* font que l'on ne peut pas en général calculer la probabilité de toutes les parties de Ω . On se limite à calculer les probabilités de certaines parties (notamment les intervalles dans le cas réel, ou les « pavés » — produits cartésiens d'intervalles — dans le cas multidimensionnel). Ce sont ces parties que l'on appelle alors des événements, et l'ensemble de tous les événements, que nous noterons \mathcal{T} est appelé une *tribu*, ou une σ -*algèbre*. La probabilité est alors une *mesure* P , définie sur \mathcal{T} , et de « masse totale » égale à 1. Le vocabulaire employé ici est celui de la théorie de la mesure. Le cas des variables à densité est développé dans la quatrième partie du livre (la dernière du premier tome). Le cadre le plus général constitue la cinquième partie du livre (la première

du deuxième tome). Nous indiquons ci-dessous les problèmes que peut poser le choix de Ω , dans les divers cadres mentionnés ci-dessus.

Dans l'approche classique, on choisit un univers Ω lié aux observations d'une expérience aléatoire. Cette approche soulève quelques difficultés, dont nous énumérons quelques-unes ci-après.

- Elle pose un problème lié à la subjectivité de l'observateur. Le choix de Ω peut dépendre du choix de l'observateur, de sa plus ou moins grande naïveté, mais aussi du choix des observations qu'il souhaite réaliser. Dans ce cadre, on peut concevoir que, à partir d'un même modèle expérimental, un changement de questionnement amène à modifier le choix de Ω .
- Elle pose un problème lié à « l'instabilité » éventuelle de Ω . Si l'expérience consiste en une suite de jeux non bornée d'avance (par exemple, on joue à pile ou face jusqu'au moment où l'on obtient dix piles consécutifs), la description de certains événements peut nécessiter que l'on « change de Ω à chaque étape », sauf si l'on se résout à introduire un univers Ω infini tenant compte de l'ensemble de l'expérience (et de sa durée potentiellement infinie), mais ceci pose — au minimum — d'autres problèmes, techniques en particulier.

Dans l'approche moderne (qui s'appuie sur la théorie de la mesure), l'univers Ω est un ensemble abstrait, en général inconnu (sinon par son nom). Une variable aléatoire X est une application de Ω dans \mathbb{R} (par exemple), mais la variable ω n'est en général pas nommée, et l'écriture $P(X \in [0, 1])$ représente la mesure de l'ensemble des éléments ω de Ω tels que $X(\omega)$ appartienne à $[0, 1]$. Dans cette approche, on ne cesse de raisonner sur des parties de Ω , de leur appliquer le langage ensembliste, sans jamais référer à ces parties comme à des ensembles mais en les décrivant souvent par un discours non formalisé et événementiel.

Cette situation est unique dans le cadre du travail mathématique : c'est à peu près le seul cas où la variable dont dépendent les fonctions n'a aucune importance (d'où le fait qu'elle ne soit pas nommée en général), ceci alors même que les fonctions subissent la plupart des traitements de l'analyse classique, comme notamment les estimations d'intégrales qui permettent d'obtenir espérance et variance.

Ce changement de statut de l'ensemble Ω est à l'origine de nombreuses difficultés d'apprentissage. Certaines situations permettent une approche mixte : du type théorie de la mesure avec un univers Ω explicite permettant de tout calculer. Cette approche, d'un certain point de vue fondamentale si l'on veut se convaincre que les objets dont parle l'approche moderne ont effectivement une existence mathématique (comme l'existence d'un modèle pour le jeu de pile ou face infini par exemple), est néanmoins peu utilisée en pratique. Elle conduit en général à des calculs en fait inutilement

complicés, précisément parce qu'elle donne une identité aux éléments abstraits de la théorie, détournant ainsi l'attention des concepts essentiels. On peut penser qu'il s'agit de la pire manière de faire des probabilités, mais que le passage par cette étape peut constituer une phase intéressante dans l'apprentissage, en motivant les approches plus abstraites indiquées plus haut.

Donnons un exemple. On choisit $\Omega = [0, 1[$, \mathcal{T} est la tribu borélienne (plus petite tribu contenant tous les intervalles de $[0, 1]$, complétée pour contenir aussi toutes les parties contenues dans un événement de mesure de Lebesgue nulle), et P la mesure de Lebesgue, qui à un intervalle associe sa longueur. On définit les variables aléatoires X_1, X_2, \dots de la manière suivante :

- ▷ $X_1(t) = 1$ si $t \in [0, 1/2[$, $X_1(t) = 0$ sinon ;
- ▷ $X_2(t) = 1$ si $t \in [0, 1/4[$ ou si $t \in [1/2, 3/4[$, $X_2(t) = 0$ sinon ;

et ainsi de suite. Chaque X_{i+1} vaut 1 sur la "première moitié" des 2^i intervalles associés à X_i et 0 sur la deuxième moitié.

La famille (X_1, X_2, \dots) constitue un modèle du jeu de pile ou face infini (où l'on suppose la pièce équilibrée et l'indépendance des résultats). Dans ce modèle, le choix initial de t détermine tout le déroulement du jeu (jusqu'à l'infini). On peut voir là au moins deux difficultés.

- La signification de ce choix de t peut poser des difficultés variées, liées au manque de compatibilité entre l'idée du hasard qui préside au résultat de chaque partie et celle au contraire d'une forme de prédestination...
- D'un point de vue didactique et épistémologique, le dénuement du modèle rend difficile à accepter l'idée de hasard associée au jeu.

Toutes ces remarques indiquent les difficultés que pose le choix d'un modèle quantitatif quand il s'agit de théorie des probabilités. Pourtant ces modèles ont du succès, notamment dans leurs applications. Ce succès tient au fait que, une fois établi et assimilé un dictionnaire entre « *vision naïve des probabilités* » et « *propriétés mathématiques du modèle* », la capacité prédictive des modèles probabilistes devient aussi bonne que celle des modèles mathématiques de la physique, par exemple.

Cette capacité prédictive des modèles mathématiques des probabilités à travers des lois du hasard, quantitatives et asymptotiques, c'est l'objet de la sixième partie du livre. Nous parlons ici de prédiction quantitative. Examinons un premier énoncé quantitatif :

« *Si on lance un dé équilibré, chaque face a 1 chance sur 6, ou encore une probabilité 1/6, d'apparaître* ».

Cet énoncé quantitatif a un sens précis dans le cadre d'un modèle. Il n'en a aucun, sinon l'expression de notre croyance en une symétrie géométrique a priori que l'expérience aléatoire va se faire un plaisir de briser, dans le cadre du jeu de dé lui-même. Dans l'expérience consistant à lancer un dé et à observer le numéro écrit sur la face qui se trouve sur le dessus, il n'y a rien qui soit égal à $1/6$.

On donnera plus de sens à ce nombre $1/6$ en répétant 1000 fois l'expérience et en déterminant la fréquence d'apparition de chaque face. Cette fréquence représente maintenant une donnée quantitative que l'on peut comparer à $1/6$, et l'expérience prouve alors que les six fréquences obtenues au bout des 1000 lancers sont en général assez voisines de $1/6$. Et si on lance 10 000 fois au lieu de 1000, ce sera encore plus vrai...

C'est ce type de constatation qui conduit à dire que les lois mathématiques du hasard sont forcément des lois asymptotiques, et que ce sont ces lois qui ont une force prédictive.

La plus connue de ces lois asymptotiques est la loi des grands nombres. Cette loi possède deux variantes, connues sous le nom de loi faible et loi forte, que nous énonçons ci-dessous sans entrer dans les détails techniques.

Dans la suite, X_1, X_2, \dots désignent les résultats (quantitatifs) d'expériences aléatoires reproduites à l'identique et indépendamment. On note M_n la moyenne empirique

$$M_n = \frac{X_1 + \dots + X_n}{n}.$$

Loi forte des grands nombres.— Avec probabilité 1, la moyenne empirique tend vers l'espérance $E(X)$. On dit aussi que la suite de terme général M_n converge presque sûrement vers $E(X)$.

Loi faible des grands nombres.— Pour tout $\varepsilon > 0$, la probabilité pour que $|M_n - E(X)| > \varepsilon$ tend vers 0 quand n tend vers l'infini. On dit aussi que la suite de terme général M_n converge en probabilité vers $E(X)$.

Pour des variables aléatoires X_n de carré intégrable, la loi faible se démontre très facilement en utilisant l'additivité de la variance pour des variables indépendantes et l'inégalité de Bienaymé-Tchebycheff. La loi forte est plus difficile, même pour des variables de carré intégrable. Elle est vraie — mais la démonstration est plus difficile — pour des variables seulement supposées intégrables (en théorie des probabilités, c'est une hypothèse plus faible). Par ailleurs, il est facile de prouver que la convergence presque sûre implique la convergence en probabilités, et par conséquent que la loi forte implique la loi faible. Cette dernière est donc, elle aussi, vraie pour des variables seulement intégrables.

Ces deux lois ont des significations très différentes.

- La première n'a de sens que si l'on suppose que chaque « occurrence » ω du hasard détermine l'ensemble infini de tous les résultats des expériences réalisées. La loi forte dit alors que cette suite converge en moyenne de Césaro vers $E(X)$ sauf pour un ensemble de valeurs de ω dont la probabilité vaut 0. On retrouve les difficultés épistémologiques liées à l'infini mentionnées plus haut.

- La deuxième pose beaucoup moins de problèmes d'interprétation. Elle dit que, lorsque l'on réalise n fois l'expérience et que l'on considère tous les résultats possibles pour la moyenne empirique M_n , la plupart des valeurs prises par la variable M_n sont voisines de $E(X)$ quand n est grand.

Au lycée, les élèves se sont habitués à l'idée de stabilisation des fréquences mentionnée plus haut. Ils s'y sont habitués :

- ▷ dans un cadre d'expérimentation (jeux ou expériences de hasard réalisés dans la classe, le nombre d'élèves permettant d'obtenir un grand nombre de tirages) ;
- ▷ dans un cadre de simulation : calculatrices, tableurs et autres simulateurs ad hoc permettent de « *faire comme si* » en économisant du temps et sans doute un peu de désordre...

L'adéquation des résultats obtenus dans les deux cadres montre a priori uniquement que les simulateurs sont bien programmés (ceci étant d'ailleurs plus ou moins vrai d'une technologie à l'autre). En revanche, la stabilité des fréquences est un fait scientifique, de nature expérimentale, qui ne relève pas du modèle mathématique mais de ce que nous appellerons des « lois du hasard », auxquelles il ne semble pas plus inadéquat de croire qu'à la loi de la gravitation universelle, dans la mesure où dans les deux cas n'existe guère qu'une validation expérimentale, mais où celle-ci n'a jamais été mise en défaut. Les élèves débutants manquent sans doute d'outils et de connaissances pour se poser de telles questions, mais on peut considérer que l'adéquation entre les résultats expérimentaux (les lois du hasard) et les prévisions du modèle mathématique (la loi faible des grands nombres) constitue une validation raisonnable du modèle. Cette adéquation permet en particulier de parler de *probabilité a priori* (celle du modèle : *subjective et souvent géométrique*) et de *probabilité a posteriori* (celle de l'expérience : *objective et asymptotique*) et de les comparer.

Ce type de questionnement a été retenu par les auteurs du programme de première et apparaît dans les commentaires de ce programme : « la simulation permet d'une part d'avoir des estimations de résultats impossibles à calculer explicitement et d'autre part, par la comparaison de résultats simulés à des résultats expérimentaux, de valider des modèles ».

Enfin, la septième partie du livre est consacrée à des chapitres plus « avancés » de la théorie des probabilités, comme les chaînes de Markov, les martingales ou les marches aléatoires, qui intéresseront principalement les étudiants de master. On y trouve également des chapitres concernant quelques thèmes chers aux auteurs comme la loi de Benford, des problèmes de dés ou encore quelques développements autour du problème des anniversaires.

Nos collègues de l'université de La Rochelle nous ont soutenus tout au long de notre projet. Nous pensons notamment à Laurence Cherfils, Fabienne Marotte et Jean-Marc Garnier. Qu'ils en soient remerciés. Nous remercions particulièrement nos collègues Gilles Bailly-Maitre, Noel Fraisseix (tous deux également soutien de la première heure) ainsi que Gregory Liorit qui ont relu quelques chapitres du livre et suggéré quelques modifications.

Un grand merci enfin aux éditions Calvage & Mounet, et plus particulièrement à Rached Mneimné, pour le soin apporté à l'aboutissement de ce livre.

À la mémoire de Frédéric Testard (1961-2018)

La décision d'écrire ce livre date de la fin de l'année 2015. Frédéric et moi partageons l'essentiel de nos enseignements depuis une dizaine d'années et l'idée d'écrire un livre ensemble date des débuts de cette période. Cependant, l'ampleur de la tâche m'a longtemps fait peur. C'était sans compter sur l'opiniâtreté et la force de persuasion de Frédéric. Il est le principal artisan de ce livre. Il en est à l'origine et à la conclusion. C'est lui encore qui rompt les périodes d'inactivité. Je me souviens du moment où il m'a décrit les grandes lignes du projet. Et, devant ma mine déconfite, pour me rassurer, il m'a dit qu'il y avait un millier de pages rédigées et disséminées dans son ordinateur. C'était devenu un sujet de plaisanterie entre nous : Frédéric mène mille à zéro. Durant ces presque trois années de rédaction, son envie d'expliquer des mathématiques n'a jamais faibli.

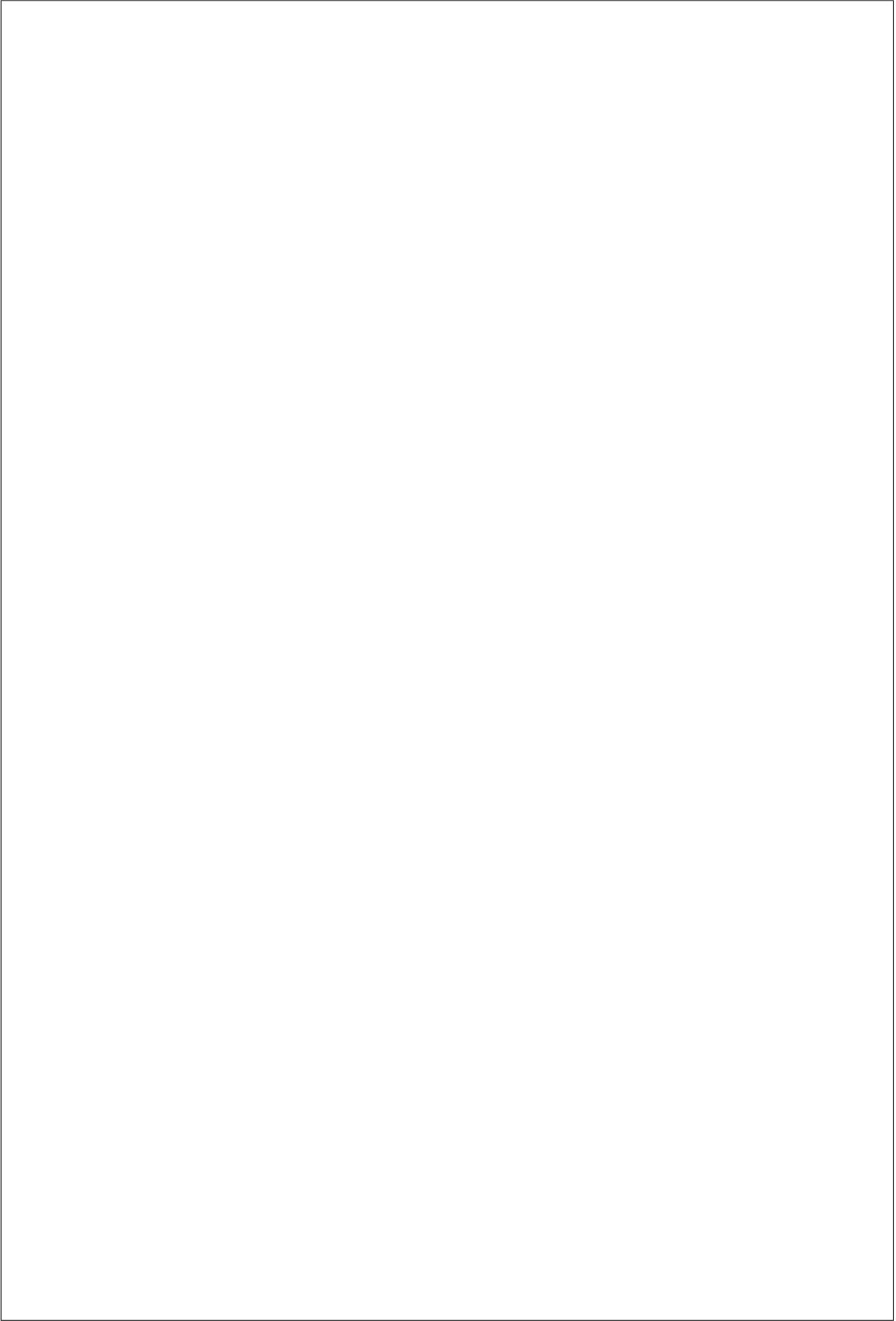


Table des matières

Probabilités 1 - P'têt ben qu'non...

Première partie - Des réponses aux problématiques issues des jeux et de la vie courante

I. Le hasard fait bien les choses

1. Panorama	7
2. Expérience aléatoire	7
3. La loi naïve des grands nombres	9
4. Deux problèmes fondateurs	11
4.1. Le problème de Galilée – ou du duc de Toscane	11
4.2. Le problème du chevalier de Méré	12
5. Expérience et simulation : l'importance de la modélisation	12
6. Exercices	14

II. Probabilités élémentaires

1. Panorama	15
2. Univers associé à une expérience aléatoire finie	16
3. Définition constructive de la probabilité	17
4. Le cas de l'équiprobabilité	18
5. Un modèle quasi-universel : le modèle d'urne	19
6. Formulaire	20
7. Petit dictionnaire	24
8. Définition axiomatique de la probabilité	25
9. Et si l'on modifie l'expérience ?	26

10. La solution des deux problèmes fondateurs	28
10.1. Le problème de Galilée, ou du duc de Toscane	28
10.2. Problème du chevalier de Méré	29
11. Exercices	30

III. Outil 1 - Utilisation du tableur

1. Panorama	37
2. Pratique du tableur	37
2.1. Affichage des entêtes de lignes et colonnes	38
2.3. Saisie des formules au tableur	39
2.5. Nommer une zone	40
2.6. Deux fonctions utiles pour les probabilités	41
3. Simulation et fluctuation d'échantillonnage	42
4. Ne peut-on simuler que ce que l'on connaît ?	46
5. Exercices	47

IV. Outil 2 – Éléments d'analyse combinatoire

1. Panorama	51
2. Principe additif	51
3. Principe multiplicatif	53
4. Parties, arrangements, combinaisons	55
5. Quelques beaux problèmes	64
5.1. Comment partager l'addition ?	64
5.5. Le jeu de dobble	69
5.9. Il y a crible et crible : Poincaré et Ératosthène	74
5.12. Dénumérer les dérangements	76
6. Exercices	78

Deuxième partie - Aux urnes, citoyens !

V. Variables aléatoires finies

1. Panorama	103
2. Variable aléatoire et loi de probabilité finies	103
2.1. Variable aléatoire finie	103
2.3. Loi de probabilité finie	105
3. Fonctions de répartition et de survie	107
4. Espérance	110
5. Variance et écart-type	114
6. Fonction génératrice des moments	116
7. Des inégalités	118
8. Automatisation des calculs	120
8.1. Représentation d'une loi finie	121
8.2. Calcul des caractéristiques numériques d'une loi finie	122
8.3. Simulation d'une loi finie	123
9. Exercices	125

VI. Conditionnement et indépendance

1. Panorama	133
2. Probabilité conditionnelle. Événements indépendants	134
2.1. Une approche statistique du conditionnement	134
2.6. Extension à des univers aléatoires finis généraux	138
3. Des conditionnements paradoxaux	143
3.1. Le rouge et le noir...	144
3.2. <i>Let's make a deal</i> : le paradoxe de Monty Hall	144
3.3. P'tetbenq oui...	147
4. Lois conditionnelles. Variables aléatoires indépendantes	148
4.1. Conditionnement d'une variable par un événement	148
4.2. Conditionnement d'une variable par une autre variable	148
5. Caractérisation « intégrale » de l'indépendance	156
6. Espérance conditionnelle	161
6.1. Espérance conditionnelle sachant un événement	161
6.3. Espérance conditionnelle sachant une variable aléatoire	162
7. Propriétés de l'espérance conditionnelle	163

8. Caractérisation par les espérances de l'espérance conditionnelle .	167
9. Exercices	170

VII. Lois finies classiques – Petit catalogue

1. Panorama	183
2. Loi uniforme	184
3. Loi de Bernoulli	184
4. Loi binomiale	185
5. Loi hypergéométrique	190
6. Sommes de variables finies indépendantes	193
6.1. Loi de la somme : convolution discrète	194
6.7. Somme « binomiale » de variables de Bernoulli	196
7. Automatisation des calculs	198
7.1. Loi binomiale : calculs de probabilités avec le tableur	198
7.2. Loi binomiale : calculs de probabilités avec Geogebra	199
7.3. Loi hypergéométrique	200
7.4. Sommes de variables indépendantes : simulation avec le tableur	200
7.5. Sommes de variables indépendantes : calcul avec Python	205
8. Exercices	207

VIII. Outil 3 – Simuler et visualiser avec Python

1. Panorama	231
2. Jupyter Notebook	231
2.1. Installer le Jupyter Notebook	232
2.2. Lancer le Jupyter Notebook	232
2.3. Créer un notebook	233
2.4. Utiliser le Jupyter Notebook	234
3. Quelques éléments de syntaxe du langage Python	236
3.1. Structures de données	236
3.2. Bases de programmation	243
4. Modules complémentaires	250
4.1. Comment importer un module	251

4.2. Le module <code>math</code>	251
4.3. Le module <code>numpy</code> et le sous-module <code>numpy.random</code>	252
4.4. Le sous-module <code>stats</code> du module <code>scipy</code>	256
4.5. Le module <code>matplotlib</code>	258
5. Exercices	261

IX. Outil 4 - Fonctions génératrices

1. Panorama	267
2. Fonction génératrice d'une variable finie à valeurs dans \mathbb{N}	268
3. Fonction génératrice d'une somme de variables aléatoires	270
3.1. Somme d'un nombre fixe de variables	270
3.3. Somme d'un nombre aléatoire de variables	271
4. Fonction génératrice, espérance et variance	273
5. Fonction génératrice d'un couple et indépendance	274
6. Exercices	275

Troisième partie - À petits pas vers l'infini

X. On ne peut pas se contenter du fini

1. Panorama	283
2. Espaces probabilisés discrets	284
2.1. Ensembles dénombrables	284
2.4. Univers aléatoires discrets	285
2.6. Loi de probabilité discrète	286
3. Formulaire	288
4. Un modèle pour le jeu de pile ou face infini ?	295
5. Exercices	296

XI. Variables aléatoires discrètes

1. Panorama	305
2. Variables et lois de probabilité discrètes	306
2.1. Variables aléatoires discrètes.	306
2.5. Loi de probabilité	308
2.8. Fonction de répartition et de survie	309

3. Variables indépendantes	311
4. Espérance d'une loi discrète	312
4.1. Définition de l'espérance	312
4.3. Propriétés de l'espérance	314
5. Espérance et indépendance	323
6. Espérance et convergence	325
7. Variance d'une loi discrète	332
8. Fonction génératrice des moments	335
9. Inégalités	338
10. Conditionnement	338
11. Exercices	340

XII. Lois discrètes classiques – Petit catalogue

1. Panorama	351
2. Loi de Poisson	352
2.1. L'opiniâtreté du malchanceux	352
2.3. Loi de Poisson	354
3. Loi géométrique	358
4. Loi de Pascal, loi binomiale négative	366
5. Sommes de variables discrètes indépendantes	372
5.1. Somme d'un nombre fixe de variables	372
5.5. Somme d'un nombre aléatoire de variables	374
6. Automatisation des calculs	378
6.1. Loi de Poisson : probabilités et fonction de répartition . . .	378
6.2. Loi binomiale négative et loi de Pascal	380
6.3. Simulation avec Python	381
7. Exercices	384

XIII. Temps d'attente

1. Panorama	393
2. Temps d'attente du $n^{\text{ème}}$ succès	394
2.1. Temps d'attente du premier succès	395
2.5. Temps d'attente du $n^{\text{ème}}$ succès	396
3. Un processus à accroissements homogènes et indépendants	398
3.1. Loi des accroissements	398
3.6. Loi conjointe et indépendance des accroissements	403
4. Deux applications	406
4.1. Les partis de Pascal – Énoncé	406
4.2. Les partis de Pascal – Solution récursive	407
4.4. Les partis de Pascal – Solution de Pascal et Fermat	408
4.6. Les partis de Pascal – Solution par les temps d'attente	409
4.8. Les allumettes de Banach	411
5. Suites de piles consécutifs	413
5.1. Simulation	413
5.2. Calculs	414
5.5. Temps d'attente d'une suite de piles consécutifs : loi et espérance	416
6. Paradoxe de Penney et théorème de Conway	422
6.1. Tous les motifs ne sont pas égaux	422
6.2. Théorème de Conway	422
6.5. Détermination du temps d'attente moyen	424
6.10. Affrontement de deux joueurs	427
6.14. Simulation en Python du théorème de Conway	430
7. Exercices	433

XIV. Outil 5 – Séries numériques

1. Panorama	457
2. Séries numériques – Convergence	458
3. Convergence des séries à termes positifs	460
3.4. Théorèmes de comparaison et d'équivalence	461
3.7. Séries de référence	462

3.12. Quelques critères supplémentaires	463
4. Convergences absolue et commutative	465
4.1. Convergence absolue	465
4.4. Convergence commutative	466
5. Séries semi-convergentes	470
5.1. Définition et exemples	470
5.5. Les mystères de « la » somme d'une série alternée	472
6. Séries doubles – Théorèmes de Fubini	474
7. Série produit	484
8. Exercices	485

XV. Outil 6 – Séries entières et fonctions génératrices

1. Panorama	497
2. Séries entières – Convergence	497
3. Propriétés de la somme d'une série entière	500
4. Les séries classiques à connaître	501
5. Fonction génératrice des variables entières	504
5.1. Définition et propriétés élémentaires	504
5.4. Moments d'ordre 1 et 2 et fonction génératrice	505
5.6. Fonction génératrice et sommes de variables entières	507
6. Fonction génératrice d'une suite	509
7. Fonctions génératrices en analyse combinatoire	514
8. Exercices	517

Quatrième partie - Le continu chasse le discret

XVI. L'indispensable continu

1. Panorama	527
2. Des problèmes historiques	528
2.1. Le problème de l'aiguille de Buffon	529
2.2. Le paradoxe de Bertrand	531
3. Faire du discret avec du continu	533
4. Un modèle pour le jeu de pile ou face	536
5. Pour finir : la généricité des variables continues	538
6. Exercices	539

XVII. Variables aléatoires continues à densité

1. Panorama	547
2. Lois continues à densité	548
3. L'existence de la loi uniforme : un premier regard	551
4. Fonctions de répartition et de survie d'une loi à densité	552
5. Espérance d'une loi à densité	554
6. Variance d'une loi à densité	564
7. Fonction génératrice des moments	565
8. Inégalités	567
9. Exercices	568

XVIII. Lois à densité classiques – Petit catalogue

1. Panorama	589
2. Lois uniformes	589
2.1. Lois uniformes	589
2.5. Loi uniforme et simulation	592
3. Lois exponentielles	593
4. Lois Gamma	596
5. Lois de Cauchy	598
6. Lois gaussiennes	600
7. Automatisation des calculs	604
7.1. Simulation en utilisant le tableur	604
7.2. Simulation avec Python	605
8. Exercices	612

XIX. Vecteurs aléatoires, indépendance

1. Panorama	629
2. Vecteurs à densité, lois marginales	630
3. Variables indépendantes	633
4. Caractérisation intégrale de l'indépendance	634
5. Quelques problèmes classiques	635
5.1. L'aiguille de Buffon	635
5.2. Le paradoxe de Bertrand	637

5.3. Le jeu de « franc-carreau »	638
6. Simulation par la méthode de rejet	638
7. Exercices	643
XX. Sommes de variables indépendantes	
1. Panorama	663
2. Loi de la somme de deux variables indépendantes	664
3. Une autre méthode : les fonctions caractéristiques	666
3.1. Introduction	666
3.2. Fonction caractéristique d'une variable aléatoire	666
3.9. Application aux sommes de variables aléatoires	669
3.12. Un petit catalogue de fonctions caractéristiques	670
4. Sommes de variables usuelles	671
5. Exercices	675
XXI. Conditionnement	
1. Panorama	693
2. Conditionnement entre variables discrète et à densité	693
2.1. Conditionnement d'une variable à densité par une variable discrète	694
2.8. Conditionnement d'une variable discrète par une variable à densité	697
3. Conditionnement entre deux variables à densité	700
4. Caractérisation intégrale de l'espérance conditionnelle	703
5. Appendice : conditionnement pour un couple sans densité	707
6. Exercices	708
XXII. Outil 7 – Intégrales généralisées	
1. Panorama	715
2. Intégrales généralisées – Convergence	715
3. Intégrales généralisées et fonctions positives	718
4. Intégrales généralisées de fonctions de signe quelconque	720
4.1. Convergence absolue et intégrabilité	720
4.4. Semi-convergence	720

5. Suites d'intégrales généralisées	721
6. Application : fonctions définies par une intégrale	722
7. Application : fonction caractéristique des lois classiques	723
8. Exercices	728

XXIII. Outil 8 – Intégrales multiples

1. Panorama	731
2. Intégrales doubles	732
2.1. Définition géométrique des intégrales doubles	732
2.2. Quelques propriétés générales	733
3. Calcul des intégrales doubles : théorèmes de Fubini	733
4. Calcul en coordonnées polaires	736
5. Calcul des intégrales par changement de variables	738
6. Exercices	738

Probabilités 2 - ... P'têt ben qu'oui

Cinquième partie - La grande unification

XXIV. Espaces probabilisés abstraits

1. Panorama	5
2. Espaces probabilisés abstraits	6
2.1. Les événements	6
2.2. Définition axiomatique d'une probabilité	8
2.3. Propriétés des probabilités	9
3. Exemple fondamental : l'intervalle semi-ouvert $[0, 1[$ et la mesure de Lebesgue	12
4. D'autres exemples classiques	13
5. Variables aléatoires	15
6. Loi d'une variable aléatoire	17
7. Fonction de répartition d'une variable réelle	18
7.1. Propriétés générales des fonctions de répartition	18
7.5. Inversion des fonctions de répartition et simulation	19
8. Densité des variables finies	23
9. Exercices	27

XXV. Espérance et variance

1. Panorama	31
2. Définition abstraite de l'espérance	32
3. Variance	36
4. Inégalités	37
5. Convergences	39
6. Fonction génératrice des moments	40
7. Exercices	43

XXVI. Indépendance

1. Panorama	47
2. Variables aléatoires indépendantes	47
3. Caractérisation intégrale de l'indépendance	55
4. Exercices	57

XXVII. Conditionnement

1. Panorama	63
2. Le théorème de Radon-Nykodym	63
3. Probabilité conditionnelle	68
4. Espérance conditionnelle	75
5. Théorèmes de convergence conditionnels	76
6. Calcul des espérances conditionnelles	78
7. Caractérisation intégrale de l'espérance conditionnelle	82
8. L'espérance conditionnelle comme projection orthogonale	84
9. Exercices	87

XXVIII. Existence de modèles probabilistes

1. Panorama	91
2. Construction de mesures par prolongement	92
3. Unicité des prolongements	99
4. La mesure de Lebesgue	102
5. Un premier modèle pour le jeu de pile ou face	110
6. Construction d'une suite de variables uniformes indépendantes	112
7. Un autre modèle pour le jeu de pile ou face	115
8. Le théorème de Kolmogorov	122
9. Exercices	128

XXIX. Outil 9 - Propriétés de l'intégrale de Lebesgue

1. Panorama	137
2. Rappel — Construction de l'intégrale de Riemann	138
3. Le cadre : espaces mesurés	140
4. L'intégrale de Lebesgue	145
4.1. Intégrale des fonctions étagées positives	145

4.7. Intégrale des fonctions mesurables positives	148
4.14. Intégration des fonctions de signe quelconque	153
4.23. Propriétés algébriques de l'intégrale de Lebesgue	157
5. Ensembles de mesure nulle, égalité presque partout	160
6. Des passages à la limite simplifiés	163
7. Les espaces L^p	174
7.1. Préliminaire : des espaces dont les éléments sont des classes de fonctions	174
7.2. Les espaces L^p : définitions et propriétés topologiques . . .	175
7.9. Relations d'inclusion entre les $L^p(\Omega)$	181
8. Théorèmes de densité et applications	182
9. Mesures produit, intégrales multiples et théorèmes de Fubini . .	184
10. Applications à la théorie des probabilités	188
10.1. Mesures à densité	188
10.4. Mesures images et théorème de transfert	191
11. Exercices	194

Sixième partie - Le hasard dompté

XXX. Divers modes de convergence

1. Panorama	213
2. Convergence en loi	214
3. Convergence de lois binomiales vers une loi de Poisson	226
4. Une propriété de compacité	227
5. Convergence en probabilité	230
6. Loi faible des grands nombres	232
7. Convergence presque sûre	233
8. Convergence dans $L^p(\Omega)$	235
9. Relations entre les divers modes de convergence	235
10. Le rôle de l'équi-intégrabilité	238
11. Exercices	244

XXXI. Le théorème central-limite

1. Panorama	259
2. Fonction caractéristique et convergence	260
2.1. Rappels	260
2.2. Inversion de la fonction caractéristique	260
2.3. Convergence et fonction caractéristique : le théorème de Paul Lévy	261
3. Théorème central-limite	264
3.2. Le cas particulier du jeu de pile ou face équilibré	265
3.3. Le cas général	270
4. Estimation de l'erreur : le théorème de Berry-Esseen	271
4.1. Inversion de la transformation de Fourier et conséquences .	271
4.7. Le théorème de Berry-Esseen	279
5. Approximation normale de lois binomiales	285
6. Exercices	285

XXXII. Loi forte des grands nombres

1. Panorama	299
2. Lemme de Borel-Cantelli	300
3. Loi du tout ou rien	303
4. La longueur des longues suites de succès à pile ou face	304
5. Limites de sommes de variables indépendantes	306
6. Loi faible des grands nombres	311
7. Loi forte des grands nombres	317
7.2. Calculs préliminaires	317
7.7. Démonstration de la loi forte des grands nombres	321
8. Exercices	323

Septième partie - Pour aller plus loin

XXXIII. Les dés sont jetés

- 1. Panorama 341
- 2. Paradoxe de Condorcet et dés non transitifs 342
- 3. Des trucages possibles et d'autres non 344
 - 3.1. Les dés de Sicherman 345
 - 3.4. D'autres décompositions de sommes de dés équilibrés 350
 - 3.18. La somme n'est jamais uniforme 370
- 4. Décompositions de lois uniformes en sommes de lois finies 373
- 5. Exercices 387

XXXIV. Des coïncidences pas si troublantes

- 1. Panorama 411
- 2. Des classes de 23 pour économiser les bougies 412
 - 2.1. L'énoncé du problème 412
 - 2.2. Un univers aléatoire 412
- 3. Le problème des anniversaires – Solution générale 415
 - 3.1. Modélisation et expression de la probabilité 415
 - 3.2. Un outil d'analyse : la formule de Stirling 416
 - 3.3. Calcul approché de $P(N, n)$ et du seuil $n(N)$ 417
 - 3.4. Une estimation plus précise de $P(N, n)$ et $n(N)$ 422
 - 3.5. Pour aller plus loin : des doubles coïncidences? 425
- 4. Quelques applications 427
 - 4.1. Comptage de population 427
 - 4.2. Sondage avec ou sans répétition? 429
 - 4.3. Fonctions de hachage et sécurité informatique 429
 - 4.4. Fiabilité des tests ADN 431
- 5. Exercices 432

XXXV. Marches aléatoires

1. Panorama	443
2. Théorème du scrutin et principe de réflexion	444
2.1. Le théorème du scrutin : problématique et modèles	444
2.2. Le théorème du scrutin : principe de réflexion	450
3. Passages en 0 des marches dans \mathbb{Z}^2	452
3.1. Marche associée à une partie finie de pile ou face	452
3.2. Le problème des retours en 0	455
4. Exercices	467

XXXVI. Chaînes de Markov

1. Panorama	493
2. Définitions et propriétés élémentaires	494
2.1. Chaînes de Markov	494
2.2. Distributions conjointes finies	496
3. Irréductibilité	499
4. Période d'une chaîne de Markov	500
5. Distributions stationnaires	502
6. États transitoires et récurrents	503
7. Distributions stationnaires, ergodicité et retours	508
8. Un regard algébrique – Le théorème de Perron-Frobenius	518
9. Retour sur le temps moyen d'obtention d'un motif à pile ou face	523
10. Automatisation des calculs	528
10.1. Calcul de la distribution stationnaire	529
10.2. Calcul des lois itérées et convergence vers π	531
10.3. Simulation de l'évolution de la chaîne	532
11. Exercices	535

XXXVII. Martingales

1. Panorama	575
2. Martingales, sous-martingales, surmartingales	576
3. Inégalités – Cas des sous-martingales	581
4. Convergence des martingales et sous-martingales	584
5. Martingales renversées	589
6. Arrêt optionnel, temps d'arrêt	590
7. Inégalités – Cas des surmartingales	596
8. Convergence des surmartingales	598
9. Tout le monde ne peut pas perdre indéfiniment	599
10. Exercices	603

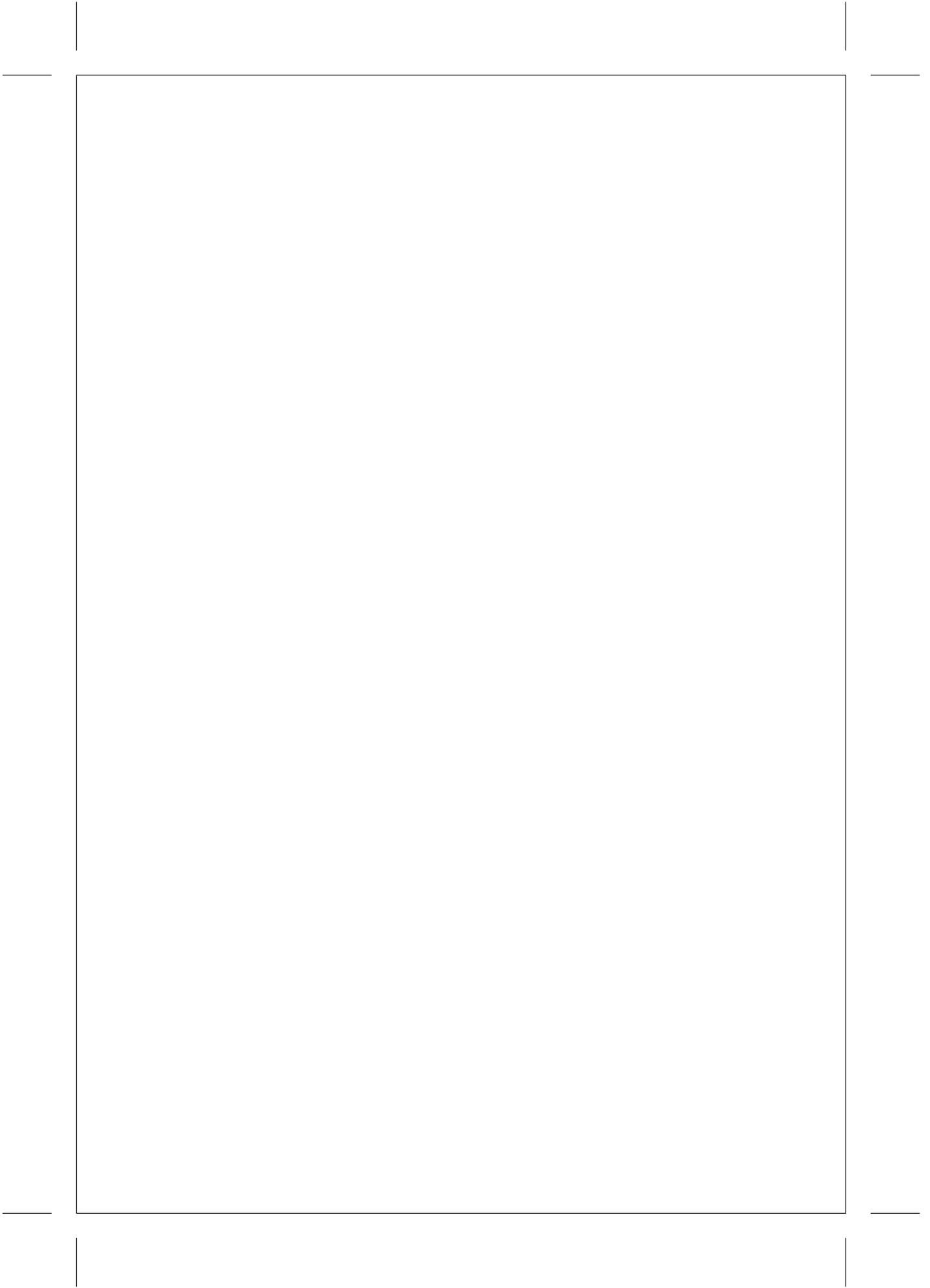
XXXVIII. Splendeurs et misères de la loi de Benford

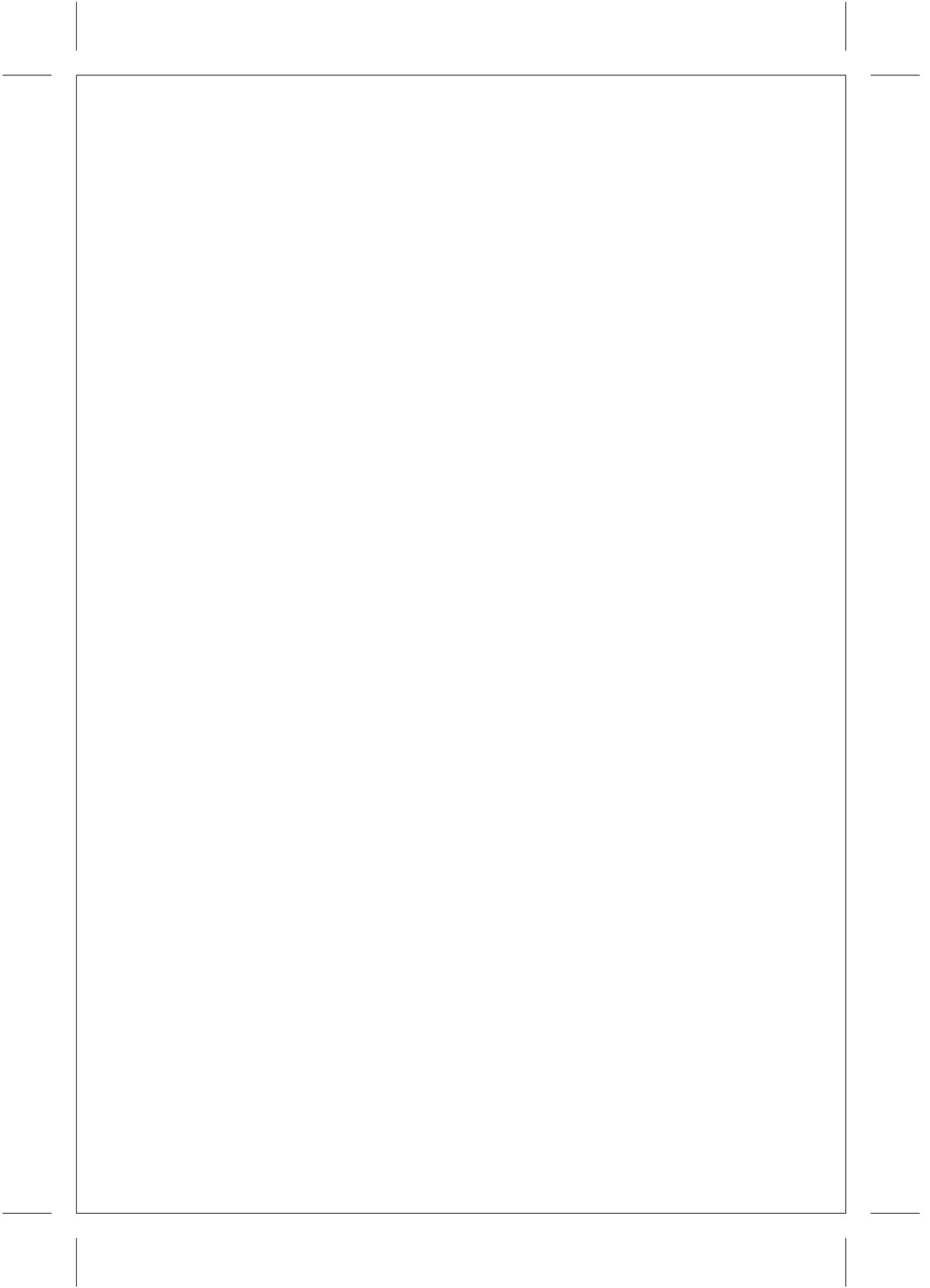
1. Panorama	615
2. La loi de Benford : un peu d'histoire	616
3. La question des fondements : intra ou extra-mathématique?	616
4. Une autre approche : les moyennes de Cesaro itérées	622
5. Une loi de la nature?	637
6. Des applications et des questions	652
6.1. Fraudes fiscales, comportements des marchés	652
6.2. La loi de Benford doit-elle faire la loi?	652
7. Une généralisation : suites d'entiers benfordiennes	656
8. Exercices	658

Bibliographie	671
----------------------	------------

Notations (celles du tome 2 sont en italique)	679
--	------------

Index	681
--------------	------------

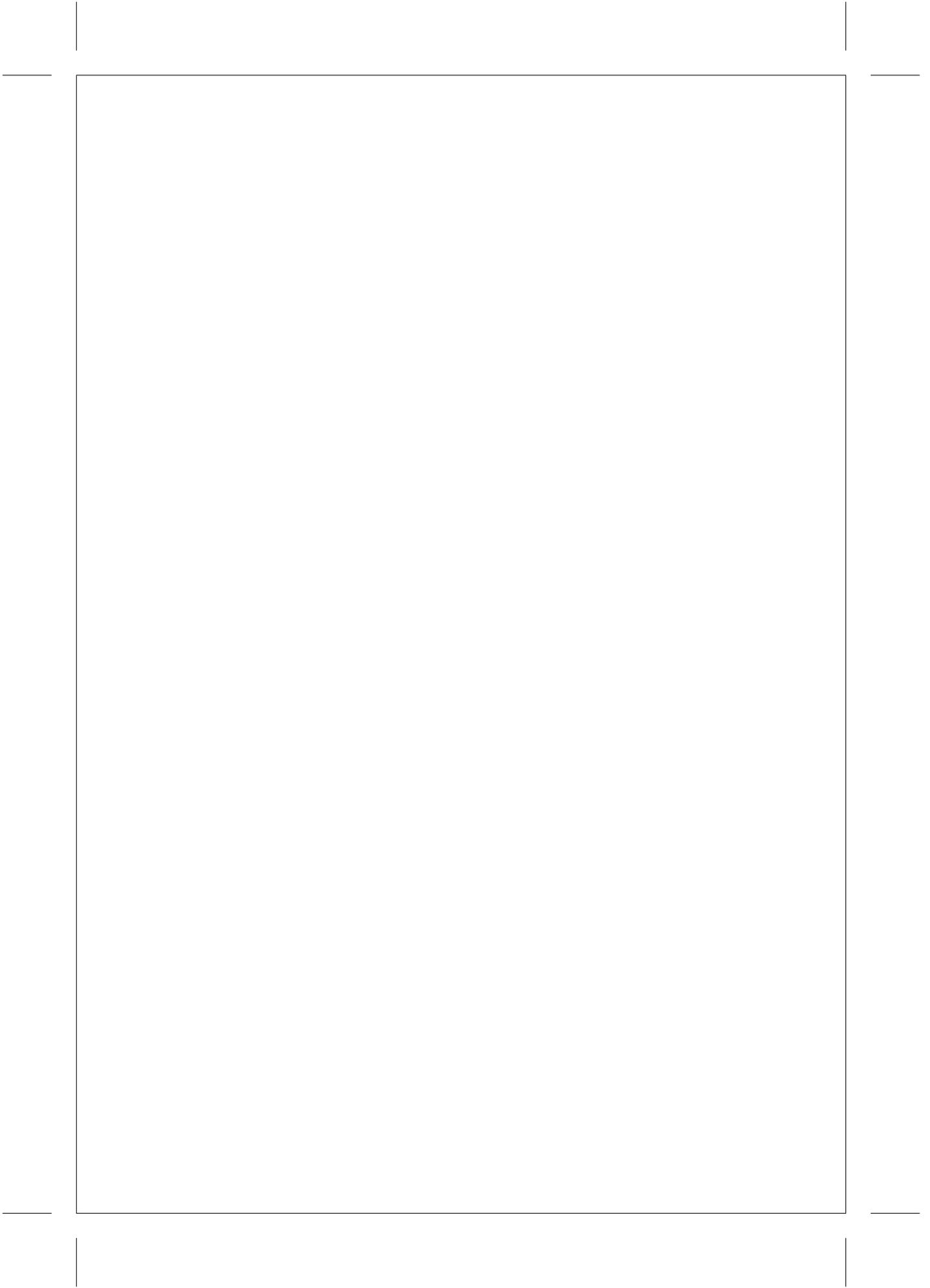




Le hasard *est* la nécessité

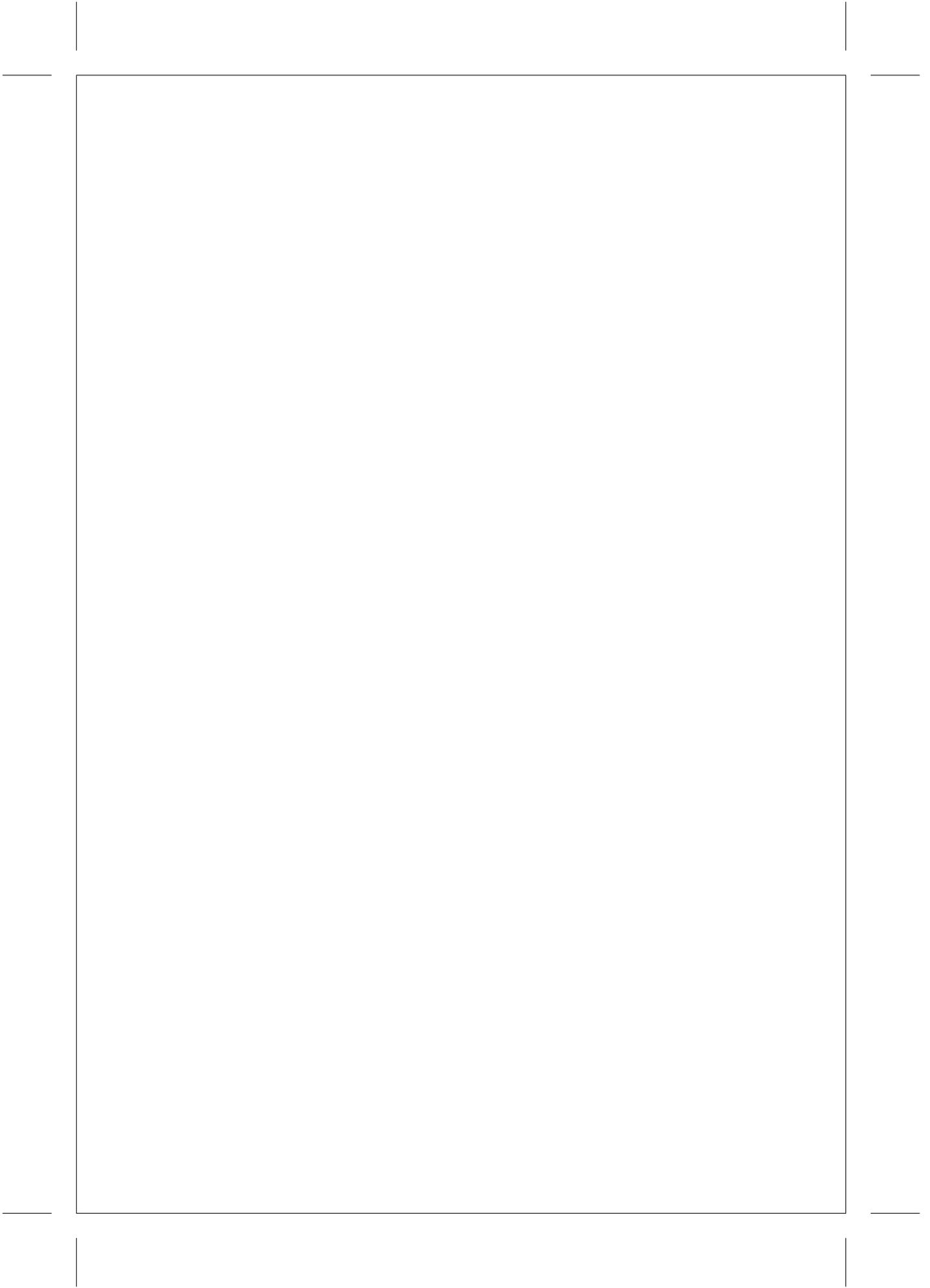
PROBABILITÉS 2

... P'TÊT BEN QU'OUI



Cinquième partie

LA GRANDE UNIFICATION



Chapitre XXIV

Espaces probabilisés abstraits

Comprendre c'est avant tout unifier.
Albert Camus

1. Panorama

À relire rapidement les points de vue développés dans les trois parties précédentes, on serait tenté de penser qu'il n'y a pas *une* mais *plusieurs* théories des probabilités, à utiliser chacune en fonction d'un contexte particulier. Des théories certes pas étanches comme nous l'avons vu à plusieurs reprises, mais possédant chacune des outils *ad hoc* et parfois guère transférables.

Nous allons ici nous attacher à montrer ce qui unifie ces points de vue. On ne peut qu'être frappé par la similarité des résultats : pour citer un simple exemple, des variables indépendantes ont une covariance nulle, que l'on soit dans le cadre fini, discret ou continu. Ceci s'applique même aux outils : la ressemblance entre les théorèmes de comparaison pour la convergence des séries ou des intégrales généralisées suggère l'existence d'un énoncé général les englobant.

C'est au début du XX^{ème} siècle qu'a été élaborée cette unification, avec la création par Lebesgue¹ de la théorie de la mesure et de l'intégration qui porte son nom. Au prix d'un sérieux effort d'abstraction, ces théories fournissent un cadre pour l'intégrale ordinaire dans lequel beaucoup de théorèmes, notamment sur la convergence, se simplifient, mais aussi une écriture unifiée pour tous les contextes probabilistes.

Dans ce nouveau cadre, la variété des univers aléatoires décrits précédemment n'a plus lieu d'être : un univers « *générique* », encore noté Ω , les

1. Henri Lebesgue (1875-1941) est le créateur de la théorie de l'intégration étudiée dans cette cinquième partie du livre. Il présenta ses résultats essentiels lors de la soutenance de sa thèse, en 1902.

remplace tous. Cet univers, muni d'une mesure de probabilité P , est alors la source de variables aléatoires qui se chargeront, par leurs propriétés particulières, de prêter vie aux divers contextes — fini, discret, continu — que nous avons décrits jusqu'ici. L'outil fondamental que constitue le calcul des espérances est incarné dans ce cadre par la théorie de l'intégration, et des théorèmes aux preuves multiples tels que la linéarité de l'espérance, deviennent de simples propriétés de calcul intégral.

Enfin, il n'est pas désagréable d'observer que théorie des séries et des intégrales généralisées relevant elles aussi (mais cette fois dans un contexte non probabiliste) de cette même conception unifiée, les théorèmes de comparaison évoqués plus haut apparaissent effectivement comme des variations d'un même énoncé général sur l'intégrabilité au sens de Lebesgue.

2. Espaces probabilisés abstraits

Un espace probabilisé est un ensemble Ω , *a priori* quelconque — mais de préférence assez gros si on veut pouvoir décrire des phénomènes complexes, comme nous le verrons au chapitre suivant — dans lequel on a identifié certaines parties appelées des **événements**, et défini une manière de calculer la probabilité de ces événements, non pas constructive comme la donnée des p_i sur un univers fini par exemple, mais axiomatique : une probabilité doit satisfaire certaines règles de calcul (d'ailleurs sans grande surprise).

Nous décrivons ci-dessous les propriétés exigées tant des événements que de leurs probabilités : ces propriétés constituent l'axiomatique de la théorie abstraite des probabilités.

2.1. Les événements

Dans les contextes finis ou dénombrables, toute partie de l'univers aléatoire était un événement. Dans le contexte continu, nous nous sommes surtout intéressés à des événements tels que « $X \in [a, b]$ », ou des intersections ou réunions d'événements de ce type. Dans toutes ces situations, les événements étaient définis *après* l'expérience aléatoire, la conception — plus ou moins implicite — de l'univers aléatoire faisant en sorte que l'on puisse toujours calculer leur probabilité.

Le point de vue ici est à peu près opposé. On se donne d'abord Ω , puis les événements, et les variables aléatoires n'arrivent qu'après. C'est, comme indiqué plus haut, le choix d'un Ω suffisamment riche qui permettra d'étudier les variables qui nous intéressent (ce sera décrit au chapitre suivant, où nous prouverons que la théorie n'est pas construite sur du sable et que les objets dont nous décrivons les propriétés existent effectivement).

2.1.1. Définition. Soit Ω un ensemble et soit \mathcal{A} une famille de parties de Ω .

On dit que \mathcal{A} est une **tribu**, ou une **σ -algèbre** si elle vérifie les propriétés suivantes.

- (a) La partie Ω appartient à \mathcal{A} .
- (b) Si $A \in \mathcal{A}$, la partie A^C (complémentaire de A) appartient aussi à \mathcal{A} .
- (c) Si les parties A_n ($n \in \mathbb{N}$) appartiennent à \mathcal{A} , leur réunion aussi.

On reformule souvent la propriété (c) en disant que \mathcal{A} est « *stable par réunion dénombrable* », ou « *σ -stable* ».

Les éléments de \mathcal{A} sont appelés des **événements**. L'élément Ω est appelé **événement certain**.

Exemples

L'ensemble $\mathcal{A} = \mathcal{P}(\Omega)$ toutes les parties de Ω est évidemment une tribu. C'est la plus grosse possible. On l'appelle la **tribu discrète**.

L'ensemble $\mathcal{A} = \{\Omega, \emptyset\}$ est aussi une tribu (nous laissons le lecteur le vérifier). C'est la plus petite. On l'appelle la **tribu grossière**.

Les propriétés (a) à (c) ont quelques conséquences simples et utiles.

2.1.2. Théorème. Soit Ω un ensemble et soit $\mathcal{A} \subset \mathcal{P}(\Omega)$ une tribu.

(i) La partie \emptyset (ensemble vide) appartient à \mathcal{A} . On l'appelle **événement impossible**.

(ii) Si les parties A_n ($n \in \mathbb{N}$) appartiennent à \mathcal{A} , leur intersection aussi.

En d'autres termes, \mathcal{A} est aussi « *stable par intersection dénombrable* ».

Démonstration. (i) C'est une conséquence immédiate de (a) et (b).

(ii) On utilise l'égalité ensembliste :

$$\left(\bigcap_{n=0}^{+\infty} A_n \right) = \left(\bigcup_{n=0}^{+\infty} A_n^C \right)^C$$

et les points (a) et (c). □

En pratique, on étudie souvent des réunions (ou intersections) finies d'événements : le théorème suivant indique qu'il s'agit là encore d'événements.

2.1.3. Théorème. *Toute tribu est stable par réunion ou intersection finies.*

Démonstration. Pour montrer que si $A_1, \dots, A_n \in \mathcal{A}$, leur réunion

$$F = A_1 \cup \dots \cup A_n$$

appartient à \mathcal{A} , on pose $A_{n+1} = A_{n+2} = \dots = A_n$ et on observe que F est égal à la réunion dénombrable des A_i .

Pour une intersection, soit on reprend la démonstration du point (ii) du théorème précédent, soit on « complète » encore l'intersection avec des A_i égaux à A_n ou Ω . \square

Nous laissons le lecteur vérifier la propriété suivante.

2.1.4. Proposition. *Si A et B sont des événements, $A \setminus B$ (éléments de A qui ne sont pas dans B) est un événement.*

Ces propriétés permettent de faire sur les événements les opérations habituelles : le fait qu'un événement ne se produise pas est encore un événement (propriété (ii) de la définition), le fait que l'un des événements d'une famille dénombrable se produise est encore un événement (propriété (iii) : par exemple, le fait qu'une pièce « finisse par tomber sur pile » est la réunion dénombrable des événements « la pièce tombe sur pile au $n^{\text{ème}}$ coup »), le fait que tous se produisent aussi, etc.

La nécessité que Ω soit un événement peut ne pas apparaître aussi claire : mais elle est nécessaire si on veut qu'existe au moins un événement A (car alors A^C est un événement et $A \cup A^C = \Omega$ en est aussi un), mais aussi pour pouvoir écrire des phrases telles que « la probabilité totale est égale à 1 : il s'agit bien de la probabilité de l'événement certain Ω ».

2.1.5. Définition. Espace probabilisable

On appelle **espace probabilisable** tout couple (Ω, \mathcal{A}) , où Ω est un ensemble et \mathcal{A} une tribu de parties de Ω .

2.2. Définition axiomatique d'une probabilité

Une fois que l'on dispose d'un espace probabilisable, on peut associer aux événements leur probabilité.

2.2.1. Définition. Mesure de probabilité

Soit (Ω, \mathcal{A}) un espace probabilisable. Une application $P : \mathcal{A} \rightarrow \mathbb{R}$ est une **mesure de probabilité** si elle vérifie les conditions suivantes.