Mathématiques en devenir

101. — Jacques Faraut. *Analyse sur les groupes de Lie. Une introduction*
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*
110. — Bernard Candelpergher. *Théorie des probabilités. Une introduction élémentaire*
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*
118. — Martine et Hervé Queffélec. *Analyse complexe et applications*
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*
122. — Philippe Caldero et Jérôme Germoni. *Nouvelles histoires hédonistes de groupes et de géométries. Tome second*
123. — Jacques Faraut. *Analyse sur les groupes de Lie, une introduction.* Nouvelle édition revue et augmentée
124. — Charles-Michel Marle. *Géométrie symplectique et géométrie de Poisson*
125. — Pascal Boyer. *Petit compagnon des nombres et de leurs applications*

Bernard Candelpergher

Théorie des probabilités

Une introduction élémentaire

Nouveau tirage, corrigé et bonifié

Calvage & Mounet

Bernard Candelpergher est maître de conférences honoraire à l'université de Nice. Ses travaux de recherche se situent dans le domaine de l'Analyse et concernent les procédés de sommation de séries divergentes et l'étude de certains aspects mathématiques de la physique quantique. Il est l'auteur de *Fonctions d'une variable complexe*, chez Armand Colin (1995), du *Calcul intégral*, chez Cassini (2009), et plus récemment (en septembre 2017) d'un « lecture notes » chez Springer, intitulé *Ramanujan Summation of Divergent Series.*

`candel@unice.fr`

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 978-2-9163-5213-8

9 782916 352138

⊗ Imprimé sur papier permanent

Nouveau tirage corrigé, Paris, 2019

À mon cher Arthur

Avant-propos

Le Calcul des Probabilités intervient dans de nombreux domaines où l'on doit étudier des phénomènes soumis à des variations aléatoires. Ce calcul est né du besoin de quantifier le probable. La première question qui se pose est celle de la définition de la probabilité d'un événement A. Dans les situations où le nombre d'éventualités possibles est fini, on pourrait espérer se ramener à un problème de dénombrement, en utilisant la formule :

$$\text{probabilité de } A = \frac{\text{nombre d'éventualités favorables à } A}{\text{nombre d'éventualités possibles}}.$$

Mais l'on se rend immédiatement compte qu'elle ne peut constituer une définition générale de la probabilité. En effet, elle n'est admissible que si toutes les éventualités possibles ont la même chance de se réaliser, ce qui suppose déjà une définition de la probabilité.

Le formalisme mathématique qui permet de contourner cette difficulté a été introduit en 1933 par Kolmogorov et est aujourd'hui communément utilisé.

Le but de ce livre est de donner une présentation rigoureuse des bases mathématiques de cette théorie tout en essayant de garder une interprétation intuitive des probabilités.

Donnons maintenant un brève présentation des idées essentielles qui sont présentées dans les chapitres de ce livre.

Le formalisme de Kolmogorov consiste à adopter une définition axiomatique de la probabilité en utilisant la théorie des ensembles. Dans ce formalisme, les éventualités possibles, lors d'une expérience, forment un ensemble Ω et les événements sont représentés par des parties de Ω. Chaque occurrence de l'expérience donne lieu au choix aléatoire d'une éventualité ω dans Ω et un événement A est réalisé si cette éventualité appartient à la partie A. La distribution de la probabilité sur l'ensemble des éventualités se traduit, dans ce formalisme, par la donnée d'une mesure sur Ω, c'est-à-dire d'une application

$$P : A \mapsto P(A)$$

qui vérifie le même axiome que la mesure « géométrique » de longueur ou d'aire pour les parties d'une droite ou d'un plan, l'axiome d'additivité :

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \ldots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \ldots$$

si les A_n sont des parties deux à deux disjointes.

La justification de cet axiome réside dans l'interprétation de la probabilité $P(A)$ en termes de fréquence d'apparition de l'événement A sur un grand nombre d'expériences. Cette fréquence est définie comme le rapport N_A/N, où N_A est nombre de fois où A s'est produit lors de N d'expériences. L'hypothèse physique fondamentale est que les conditions de l'expérience sont telles que ce rapport tend vers une limite quand N devient de plus en plus grand, cette limite est la probabilité de A :

$$P(A) = \lim_{N \to +\infty} \frac{N_A}{N} \cdot$$

Comme Ω est l'ensemble de toutes les éventualités qui peuvent se produire, il est clair que $N_\Omega = N$ et donc l'application P doit vérifier la condition supplémentaire de normalisation : $P(\Omega) = 1$.

Intuitivement, le calcul des probabilités est un calcul de fréquences sur des grands nombres ; dans le formalisme de Kolmogorov, il peut être vu comme un cas particulier d'une théorie générale de la mesure.

Ce qui fait sa différence essentielle avec le calcul des mesures géométriques est la notion d'indépendance.

La définition intuitive de l'indépendance de deux événements A et B est que la réalisation de l'un d'entre eux ne doit apporter aucune information sur la réalisation de l'autre. La probabilité de B, calculée en ayant l'information que A est réalisé, n'est autre que la mesure relative de B par rapport à A ; elle est définie par la probabilité conditionnelle

$$P(B \text{ si } A) = \frac{P(B \cap A)}{P(A)} \cdot$$

La condition d'indépendance des événements A et B se traduit alors par

$$P(A \cap B) = P(A) \times P(B).$$

Autrement dit, l'indépendance de deux événements se traduit par le fait que la mesure de l'intersection des deux parties qui les représentent doit être égale au produit des mesures de chaque partie. On voit que l'on a là une condition qui est totalement étrangère aux propriétés habituelles des mesures géométriques. Ces considérations font l'objet du premier chapitre de ce livre.

Le chapitre II donne une présentation succincte de la théorie de la mesure ; il peut être passé ou parcouru rapidement dans une première lecture.

Dans le chapitre III, on aborde la notion centrale du calcul des probabilités, celle de variable aléatoire et de loi de probabilité. Intuitivement, une variable aléatoire réelle est un nombre réel X dont les variations sont soumises au hasard. Dans le formalisme de Kolmogorov l'intervention du hasard se traduit par le choix d'un élément ω de Ω, le nombre X va dépendre de ce choix. On va donc représenter une variable aléatoire X par la donnée pour toute éventualité ω d'un nombre réel $X(\omega)$; autrement dit, X devient une simple fonction

$$X : \Omega \to \mathbb{R}.$$

Le fait de représenter l'objet mouvant qu'est le nombre X dont les valeurs sont aléatoires par une fonction sur Ω semble figer le hasard. Mais celui-ci se retrouve dans la distribution des probabilités des valeurs de X qui est alors définie par la donnée pour tout intervalle I de $\mathbb{R}$ de la probabilité

$$P_X(I) = P(\{\omega \in \Omega \text{ tel que } X(\omega) \in I\}).$$

Cela définit une mesure de probabilité sur $\mathbb{R}$,

$$P_X : I \mapsto P_X(I),$$

que l'on appelle la loi de probabilité de X.

Les lois de probabilités sur $\mathbb{R}$ sont donc des mesures. Intuitivement, une mesure de probabilité sur $\mathbb{R}$ est l'analogue d'une répartition de masse sur une droite. Si celle-ci est concentrée sur un ensemble dénombrable de points, on dit que la mesure est discrète ; elle peut aussi être diluée avec une certaine densité. Ces mesures P_X sur $\mathbb{R}$ peuvent être représentées plus concrètement par leurs fonctions de répartition

$$F_X(x) = P(X \leqslant x).$$

La fonction F_X est une fonction croissante dont l'allure traduit la répartition des probabilités des valeurs de X. Par exemple, si F_X est une fonction en escalier, les points de discontinuité de F_X correspondent aux valeurs possibles x_i de X et la valeur du saut de F_X en un point x_i est la probabilité que $X = x_i$. Mais la fonction F_X peut aussi être continue sur $\mathbb{R}$, par exemple si F_X est la primitive d'une fonction continue f ; dans ce cas, la loi de X est diluée sur $\mathbb{R}$ avec la densité f, ce qui se traduit par le fait que la loi de probabilité de X est donnée par

$$P_X([a, b]) = \int_a^b f(x)dx.$$

La représentation des variables aléatoires comme fonctions sur Ω permet d'utiliser un calcul intégral que l'on peut définir plus généralement sur tout ensemble muni d'une mesure.

Nous développerons l'essentiel de cette théorie de l'intégration au chapitre IV. On pourra ainsi définir la moyenne ou espérance, notée $E(X)$, d'une variable aléatoire comme l'intégrale de la fonction X sur Ω :

$$E(X) = \int_\Omega X(\omega) dP(\omega).$$

Cette intégrale est un outil théorique utile dans les démonstrations de théorèmes généraux. On démontrera ainsi très simplement les inégalités classiques concernant la probabilité qu'a une variable de s'écarter suffisamment de sa moyenne. Mais du point de vue pratique, le résultat essentiel est que l'on peut transférer l'intégration sur $\mathbb{R}$ pour obtenir

$$E(X) = \int_\mathbb{R} x \, dP_X(x),$$

où l'on intègre par rapport à la mesure P_X ; autrement dit, la connaissance de la loi de X suffit à elle seule au calcul de l'intégrale.

Dans le chapitre V, on examinera les situations où plusieurs variables aléatoires sont liées à une même expérience. La notion fondamentale qui décrira la variation de ces variables est celle de loi conjointe. Dans le cas de deux variables réelles X et Y, la loi conjointe du couple (X, Y) est la mesure de probabilité sur $\mathbb{R}^2$, définie par

$$P_{(X,Y)}(I \times J) = P(\{\omega \in \Omega \text{ tel que } X(\omega) \in I \text{ et } Y(\omega) \in J\}).$$

On peut déduire de cette mesure les lois de probabilité de X et de Y. L'indépendance des variables X et Y est définie par le fait que la loi conjointe de (X, Y) est la mesure produit des lois de X et de Y, définie par

$$(P_X \otimes P_Y)(I \times J) = P_X(I) \times P_Y(J).$$

Dans ce cas, la loi de la somme $X + Y$ s'obtient simplement à l'aide des lois de X et de Y par une opération que l'on appelle le produit de convolution, et la moyenne du produit XY est donnée par

$$E(XY) = E(X)E(Y),$$

ou encore

$$\int_\Omega XY dP = \Big(\int_\Omega X dP\Big) \times \Big(\int_\Omega Y dP\Big),$$

ce qui montre combien l'indépendance des variables X et Y peut modifier les règles du calcul intégral. On généralisera ces notions à des variables aléatoires à valeurs dans $\mathbb{R}^n$.

La transformation de Fourier est une notion centrale en mathématiques ; dans les Probabilités, elle donne lieu à la notion de fonction caractéristique d'une loi de probabilité P_X. C'est la fonction définie sur $\mathbb{R}$ par

$$\varphi_X(t) = E(e^{itX}).$$

Si X et Y sont des variables aléatoires indépendantes, la fonction caractéristique de la somme $X+Y$ s'obtient simplement comme le produit des fonctions caractéristiques de X et Y :

$$\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t).$$

Le point essentiel est que la fonction caractéristique d'une loi suffit pour déterminer cette loi. On peut ainsi déterminer directement la loi de la somme de variables indépendantes. Cette notion, très utile, fera l'objet du chapitre VI.

Les variables aléatoires gaussiennes jouent un grand rôle dans la théorie des probabilités. La plus simple des lois gaussiennes est la loi normale, définie par la densité bien connue

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Les généralisations de cette loi et leurs propriétés seront étudiées au chapitre VII.

Les résultats d'une suite infinie de tirages successifs et indépendants les uns des autres, d'une variable aléatoire X, donnent naissance à une suite infinie de variables aléatoires $X_1, X_2, \ldots, X_n, \ldots$, qui sont indépendantes et de même loi que X. On peut comparer les « moyennes empiriques »

$$\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

à la « moyenne théorique » $E(X)$.

Les lois des grands nombres, dont il est question au chapitre IX, sont des résultats de convergence du type

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to E(X).$$

Les divers modes de convergence communément utilisés sont introduits dans le chapitre VIII. En particulier, la convergence en loi permet d'énoncer le théorème central limite, qui affirme que l'écart entre $\bar{X}_n$ et $E(X)$, convenablement normalisé, a pour loi limite une loi gaussienne.

Dans le cas particulier où la variable aléatoire X est la variable indicatrice d'un événement A, c'est-à-dire prend la valeur 1 lorsque A est réalisé et 0 sinon, la moyenne empirique $\bar{X}_n$ est la traduction mathématique de la fréquence d'apparition de A lors de n épreuves.

Les lois des grands nombres affirment, dans ce cas, que l'on a

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \to P(A).$$

Le théorème central limite donne alors une information sur l'écart entre ces fréquences et la probabilité. Cela montre la puissance de la théorie et sa cohérence avec les hypothèses sur lesquelles elle a été fondée.

Le chapitre X donne une petite introduction au très vaste domaine des processus stochastiques. L'exemple le plus élémentaire d'un tel processus est la marche aléatoire consistant à faire un pas vers la droite ou vers la gauche, suivant les résultats d'une suite de tirages successifs de pile ou face. Nous donnerons quelques résultats classiques sur les trajectoires de cette marche. On présentera brièvement une généralisation de ce processus que l'on appelle les martingales à temps discret. On prendra comme exemple de processus à temps continu le processus de Poisson, dont on donnera les diverses représentations classiques. Un processus à temps continu célèbre est le mouvement brownien, dont on donnera une présentation heuristique, due à Paul Levy, sous la forme d'une limite de marche aléatoire de pas infinitésimal dt et d'amplitude $\sqrt{dt}$.

Ce qui précède montre que le Calcul des Probabilités ne se réduit pas à des problèmes de dénombrement ou de calcul intégral, bien qu'une bonne maîtrise des techniques de l'Analyse soit en général nécessaire. L'expérience de l'enseignement montre que la difficulté principale que rencontre l'étudiant lors de son premier contact avec cette théorie n'est pas seulement technique, c'est celle de se forger une vision claire des notions de probabilité, indépendance, variable aléatoire, lois conjointes, convergences, etc.

Nous avons essayé dans ce livre de donner les résultats de base du Calcul des Probabilités, en insistant sur la compréhension du statut des objets de notre étude. Le Calcul des Probabilités fourmille de paradoxes qui montrent combien il est difficile d'avoir une intuition sûre dans ce domaine. Pour favoriser le développement d'une telle intuition, nous avons illustré les notions introduites par de nombreux exemples et exercices. Ceux-ci n'ont aucune prétention à l'originalité et la plupart se retrouvent dans les très bons ouvrages cités en bibliographie.

Quelques développements plus techniques sont souvent placés en fin de chaque chapitre et peuvent être passés en première lecture, certains points plus difficiles sont simplement évoqués ; le lecteur en trouvera le développement dans la vaste littérature des Probabilités, par exemple dans le livre-référence de William Feller.

Le présent ouvrage est issu des cours de Probabilités donnés en Licence de Mathématiques entre les années 1995 et 2010.

Je voudrais exprimer toute ma gratitude à Michel Miniconi, pour l'aide qu'il m'a apportée en me faisant bénéficier, lors de nos enseignements, de ses vastes connaissances dans le domaine des Probabilités. Combien de problèmes en apparence compliqués se sont éclaircis lors de nos agréables discussions. Sa vision des Mathématiques, très humaine et loin de tout dogmatisme, a été pour moi une source de réconfort. Lorsqu'il était directeur du département de Mathématiques, son soutien constant aux « proba » a permis l'éclosion à l'Université de Nice de nouveaux enseignements des probabilités.

Mes vifs remerciements vont aussi à Claude Lobry, avec qui j'ai eu le plaisir de travailler. Dans nos enseignements des probabilités, sa recherche de la présentation simple des idées de base et de démonstrations « en deux lignes », évitant tout formalisme inutile, a été pour moi très instructive.

Merci à tous les autres collègues avec qui j'ai enseigné les Probabilités, et en particulier à Frédéric Testard [1], Jean-François Burnol, Jean-Philippe Labrousse et Yannick Baraud.

Merci à mes collègues physiciens Thierry Grandou et Jacques Rubin pour les passionnantes discussions sur la physique quantique et les Inégalités de Bell.

Merci à Jean-Paul Pradère pour sa disponibilité et sa gentillesse.

Je voudrais aussi exprimer toute ma reconnaissance aux Éditions Calvage et Mounet pour le soin apporté à l'achèvement de ce livre, et en particulier à Fulbert Mignot et Bernard Randé pour leurs remarques pertinentes. Un mot enfin à l'adresse d'Alberto Arabia et de Rached Mneimné, qui m'ont prodigué avec beaucoup de gentillesse leurs conseils et encouragements.

Ce second tirage a été l'occasion de corriger un certain nombre de coquilles et d'inclure un chapitre sur le formalisme probabiliste de la Mécanique quantique et des compléments sur les intégrales stochastiques. Une amélioration sensible a été apportée par ailleurs à la mise en page.

Septembre 2019

1. Une pensée émue pour mon ami Frédéric Testard, disparu bien trop tôt en l'automne 2018. J'ai eu la chance d'enseigner avec Frédéric durant plusieurs années. Avec lui, un problème à l'abord difficile devenait une expérience joyeuse pour l'esprit. Sa puissance de travail et son engagement passionné pour l'enseignement n'ont cessé de susciter mon admiration.

Table des matières

III. Variables aléatoires réelles

IV. Intégrale et moments

V. Variables aléatoires à valeurs dans $\mathbb{R}^n$

Petit appendice mathématique